

**Comparing and evaluating data-driven journalism:
data visualization performance from the perspective of web analytics**

Ilo Aguiar Reginaldo Alexandre

**Doctoral Thesis in Digital Media
Tese de Doutoramento em Media Digitais**

Julho, 2019

Thesis submitted in accordance with the requirements for the degree of
Doctor of Philosophy in Digital Media, under scientific supervision of Prof.
António Granado.

Tese apresentada para cumprimento dos requisitos necessários à obtenção
do grau de Doutor em Media Digitais, realizada sob a orientação científica
de Prof. António Granado.

Financial support from:

FCT

Fundação para a Ciência e a Tecnologia

UT Austin | Portugal

INTERNATIONAL COLLABORATORY FOR EMERGING TECHNOLOGIES, CoLAB

Declarações

Declaro que esta tese é o resultado da minha investigação pessoal e independente. O seu conteúdo é original e todas as fontes consultadas estão devidamente mencionadas no texto, nas notas e na bibliografia.

O candidato,

A handwritten signature in blue ink, appearing to read 'D. Aguiar Reginaldo Alexandre'.

Lisboa, 19 de Julho de 2019

Declaro que esta tese se encontra em condições de ser apreciado pelo júri a designar.

O orientador,

A handwritten signature in black ink, appearing to read 'A. Branco'.

Lisboa, 19 de Julho de 2019

Acknowledgments

First, I would like to thank Fundação para Ciência e a Tecnologia and the UT Austin-Portugal program for funding this research. Without this support, this work might never have seen the light of day. Thanks to Professor António Granado for being my thesis supervisor and for supporting me since the very beginning of this journey, I appreciate your patience and good advice.

I would like to thank Professor Rosental Alves for the insights and guidance during my semester at UT Austin. Thanks to *The Texas Tribune*, especially Rodney Gibbs, for allowing me the access the data that I needed for this research. Thanks to *O Globo*, especially Vivianne Cohen and Mariana Musa, for helping me in collecting the data from the newspaper. Thanks to *Público*, especially Elisabeth Fernandes, for making available the data I required and for the availability and kindness in explaining concepts and techniques of web analytics. Still on the topic of web analysis, I also have to thank the contributions of colleagues Tiago Dias Pinheiro, Mariana Ruggeri, and Emanuel Lima.

Finally, I would like to express my gratitude to my family, especially Caroline Ribeiro, for their continuous encouragement and support, and for the informal reviews and contributions and advice throughout all the PhD process. Thank you very much.

Abstract

Data journalism is one of the hottest trends in journalism. Although the practice of using numbers to tell stories is not particularly new to newsrooms, the interaction present in data visualizations allows the active participation of the user, rarely possible before. The production of data-driven and interactive pieces increases every year, as well as the number of studies in the field. However, there is still a lack of research on user behavior in interactive news pieces. This gap can be explained in part by researchers' difficulty in accessing real quantitative data from media outlets since such data are considered sensitive and strategic. In this thesis, we compare the performance of interactive data journalism pieces with the most read news articles, and with news stories of the same topic published in the same period as the interactive ones. This analysis is based on 90 pieces of data journalism pieces and on the access to Google Analytics, an online web analytics tool, from three media outlets from three different countries: *The Texas Tribune* (United States), *O Globo* (Brazil) and *Público* (Portugal). The investigation begins by recalling important moments in the use of data in journalism in the three analyzed countries. Then, we present a brief history of media audience measurement and the relationship of journalists with the public. We present the individual results of each media outlet analyzed and all together to try to identify patterns and trends in the consumption of interactive news pieces. Among the key findings of this research are: data journalism pieces continue to be visited years after publication in a much higher proportion than other news on the same topic; users spend more time on interactive news than on the most read articles and related news stories; the recirculation rate in data journalism pieces is higher than the rate in the related news; the bounce rate is higher in the data visualization than in the most read news articles and the related news stories.

Keywords: data journalism, web analytics, The Texas Tribune, O Globo, Público.

Resumo

O jornalismo de dados é definitivamente uma das principais tendências do jornalismo atualmente. Apesar de a prática de usar dados para contar histórias não ser algo particularmente novo nas redações, o recurso da interação nas visualizações de dados permite uma participação ativa do usuário, como poucas vezes foi possível anteriormente. A produção de peças orientadas por dados e interativas aumenta a cada ano, assim como o número de estudos na área. Todavia, ainda há uma ausência de investigações sobre o comportamento do usuário em notícias interativas de jornalismo de dados. Essa lacuna pode ser explicada em parte pela dificuldade dos investigadores em aceder a dados quantitativos reais dos meios de comunicação social, uma vez que esses dados são considerados sensíveis e estratégicos. Nesta tese, comparamos o desempenho de notícias interativas de jornalismo de dados com as notícias mais lidas e com notícias do mesmo tópico publicadas no mesmo período em que as matérias interativas. Essa análise é realizada com base em 90 peças de jornalismo de dados e no acesso ao Google Analytics, ferramenta *online* de análise de métricas, de três veículos de comunicação de três países diferentes: *The Texas Tribune* (Estados Unidos), *O Globo* (Brasil) e *Público* (Portugal). A investigação começa por recontar momentos importantes do uso de dados no jornalismo nos três países analisados. Em seguida é realizado um breve histórico da mensuração da audiência nos meios de comunicação social e da relação dos jornalistas com o público. Apresentamos os resultados individuais de cada veículo de comunicação e em conjunto para tentar identificar padrões e tendências no consumo de notícias interativas de jornalismo de dados. Entre as principais descobertas dessa investigação estão: visualizações de dados seguem sendo visitadas anos após a publicação em uma proporção muito superior do que outras notícias do mesmo tópico; os usuários passam mais tempo nas visualizações interativas do que nas notícias mais lidas e nas notícias relacionadas; a taxa de “recirculation” nas peças de jornalismo de dados é superior à taxa das notícias relacionadas; a taxa de rejeição é maior nas matérias com visualização de dados do que nas notícias mais lidas e nas notícias relacionadas.

Palavras-chave: jornalismo de dados, web analytics, The Texas Tribune, O Globo, Público.

Contents

List of Figures	x
List of Tables	xviii
Introduction	1
I – Data Journalism: past and present	6
1.1. The roots of Data Journalism	8
1.2. Databases, statistical graphics and the “Golden Age” of data visualization	19
1.3. The legacy of early social surveys to journalism and data visualization	23
1.4. Computer-assisted reporting and the Precision Journalism.....	38
1.5. Data journalism and data visualization.....	49
II – Web Analytics and the quantitative audience	58
2.1. Brief history of traditional media measurement.....	61
2.2. Online audience measurements	63
2.3. From values-oriented to analytics-driven news	66
2.4. From data-driven to data-informed newsrooms	74
III – Methodology	79
3.1. Quantitative content analysis	82
3.2. Quantitative analysis	84
3.3. Selection of the case studies and limitations	89
IV – The Texas Tribune	92
4.1. History	92
4.2. Data journalism pieces from 2014 and the average of the 500 most read articles.....	93
4.3. Data journalism pieces from 2014 and their related stories	96
4.4. Data journalism pieces from 2015 and the average of the 500 most read articles.....	100
4.5. Data journalism pieces from 2015 and their related stories	103
4.6. Data journalism pieces from 2016 and the average of the 500 most read articles.....	106
4.7. Data journalism pieces from 2016 and their related stories	110
V – O Globo	114
5.1. History	114
5.2. Data journalism pieces from 2014 and the average of the 500 most read articles.....	115
5.3. Data journalism pieces from 2014 and their related stories	118
5.4. Data journalism pieces from 2015 and the average of the 500 most read articles.....	121
5.5. Data journalism pieces from 2015 and their related stories	124

5.6.	Data journalism pieces from 2016 and the average of the 500 most read articles.....	128
5.7.	Data journalism pieces from 2016 and their related stories	131
VI – Público		136
6.1.	History	136
6.2.	Data journalism pieces from 2014 and the average of the 500 most read articles.....	137
6.3.	Data journalism pieces from 2014 and their related stories	140
6.4.	Data journalism pieces from 2015 and the average of the 500 most read articles.....	144
6.5.	Data journalism pieces from 2015 and their related stories	147
6.6.	Data journalism pieces from 2016 and the average of the 500 most read articles.....	150
6.7.	Data journalism pieces from 2016 and their related stories	153
VII – Discussion and Conclusions		157
7.1.	<i>The Texas Tribune's</i> data journalism pieces performed impressively; those from <i>O Globo</i> and <i>Público</i> , not so much	157
7.2.	Data journalism pieces are not a landing page	158
7.3.	Related news dies as (some) data journalism pieces flourish	159
7.4.	Users spend more time on data journalism pieces than on other news.....	160
7.5.	Data journalism pieces visits are single-page visits.....	162
7.6.	Data journalism pieces drive more recirculation	164
7.7.	A quarter of the data journalism pieces are accessed again and again.....	166
7.8.	Select is the most commonly used feature in data journalism pieces.....	167
7.9.	Politics/governance is the most covered subject among data journalism pieces.....	170
7.10.	Data journalism pieces are appealing to international audiences.....	171
7.11.	Data journalism pieces shift from the desktop dominance to mobile devices.....	173
7.12.	Once the main source for data journalism pieces, Google is dethroned by Facebook	174
7.13.	Over the years Chrome reinforces its sovereignty among data journalism pieces.....	175
7.14.	Final considerations and further research.....	176
Bibliography		180
Appendix A: <i>The Texas Tribune</i> data journalism pieces		221
A.1.	Interactive pieces from 2014	221
A.2.	Interactive pieces from 2015	231
A.3.	Interactive pieces from 2016	241

Appendix B: <i>O Globo</i> data journalism pieces.....	251
B.1. Interactive pieces from 2014	251
B.2. Interactive pieces from 2015	261
B.3. Interactive pieces from 2016	271
Appendix C: <i>Público</i> data journalism pieces	281
C.1. Interactive pieces from 2014	281
C.2. Interactive pieces from 2015	291
C.3. Interactive pieces from 2016	301

List of Figures

Figure 1 – Search interest related to the topic of Data Journalism between 2004 and 2018 worldwide. A value of 100 is the peak popularity for the term. A score of 0 means there was not enough data for the term.	1
Figure 2 – “Afghanistan war logs: IED attacks on civilians, coalition and Afghan troops”, a data journalism story published on July 26, 2010, by <i>The Guardian</i>	7
Figure 3 – “Table of Casualties” from <i>Natural and Political Observations on the Bills of Mortality</i> printed in <i>The Economic Writings of Sir William Petty</i> (1986).....	9
Figure 4 – The highlight shows a data journalism piece published on page 2 of <i>The Guardian</i> (then <i>The Manchester Guardian</i>) on May 5, 1821.....	11
Figure 5 – The highlight shows a data journalism piece published on page 2 of the <i>New York Herald</i> on February 5, 1842.....	13
Figure 6 – The highlight shows a data visualization published on the front page of the <i>New York Tribune</i> on September 29, 1849.....	14
Figure 7 – The highlight shows a data journalism piece published on pages 1 to 3 of <i>O Observador Constitucional</i> on January 17, 1831.....	15
Figure 8 – The highlight shows a data journalism piece published on front page of <i>Diário Popular</i> on March 9, 1868.....	17
Figure 9 – The highlight shows a data journalism piece published on the front page of <i>Repórter</i> on April 21, 1888.....	18
Figure 10 – Reproduction of the original map made by the physician John Snow in 1854. Cholera cases are highlighted in black.	21
Figure 11 – Reproduction of “Diagram of the causes of mortality in the army in the East” made by the nurse Florence Nightingale in 1858.	22
Figure 12 – Reproduction of the engineer Charles Joseph Minard’s map from 1869 depicting Napoleon’s disastrous Russian campaign of 1812.	23
Figure 13 – Reproduction of the industrialist Charles Booth’s <i>Descriptive Map of London Poverty</i> published in 1889.....	27
Figure 14 – Reproduction of Hull-House’s Wage Map 1-4 by Residents of Hull-House (1895)..	29
Figure 15 – Compilation of charts and tables present in the sociologist W. E. B. Du Bois’ study <i>The Philadelphia Negro</i> (2010).....	30
Figure 16 – Compilation of charts and tables present in <i>A Modern Church to Meet a Modern Situation</i> published in 1912 and organized by Reverend Charles Stelzle.	32

Figure 17 – The highlight shows a data visualization published on page 10, 4 th section of <i>The Sun</i> , on March 24, 1912.	33
Figure 18 – The highlight shows a data visualization published on the front page of <i>The New York Times</i> on November 3, 1920.	34
Figure 19 – Reproduction from <i>Fortune's</i> November 1940 edition, pages 74 and 75.	35
Figure 20 – The highlights of charts in <i>Seara Nova's</i> March 22, 1941 edition (above) and in the edition of May 24, 1947.	36
Figure 21 – The highlights of charts and maps in <i>Realidade's</i> November 1967 edition (above) and the edition of January 1968.	37
Figure 22 – Reproduction of page 4b of the <i>Detroit Free Press</i> of August 20, 1967.	39
Figure 23 – The highlight shows data used in Mário Rosa's scoop published on page 2 of <i>Jornal do Brasil</i> on August 27, 1991.....	43
Figure 24 – Reproduction of the article on the computerization of <i>O Comércio do Porto</i> , published on November 29, 1985.....	44
Figure 25 – The highlight of the <i>24 Horas'</i> front page, published on January 13, 2006.	45
Figure 26 – Reproduction of the <i>Miami Herald's</i> page 16SR, published on December 20, 1992.	46
Figure 27 – Reproduction of the <i>Chicago Tribune's</i> webpage "1995 Chicago Homicides".	49
Figure 28 – Reproduction of Adrian Holovaty's website "Chicago Crime Map".....	50
Figure 29 – Reproduction of the <i>Estadão's</i> webpage "Basômetro".	55
Figure 30 – Reproduction of the <i>Público's</i> webpage "Doze Anos de Incêndios".....	56
Figure 31 – Reproduction of W3Techs survey showing the top 10 traffic analysis tools for websites.	59
Figure 32 – Reproduction of Pew Research Center's graph showing the estimated advertising and circulation revenue of the newspaper industry in the United States.....	70
Figure 33 – Reproduction of Pew Research Center's graph showing the employment in newspaper newsrooms in the United States.	70
Figure 34 – Reproduction of "A conta dos passarálhos" graph showing the number of journalists fired from newsrooms in Brazil.	71
Figure 35 – Chart with the estimated circulation between 2007 and 2018 of the 3 largest daily newspapers in Portugal.....	71
Figure 36 – Reproduction of <i>The Guardian's</i> Ophan system.	75

Figure 37 – Reproduction of the <i>Financial Time’s</i> Lantern system.....	75
Figure 38 – Performance in pageviews of <i>The Texas Tribune</i> interactive pieces compared to the average of the most read articles in 2014 (represented as 100%).....	94
Figure 39 – Performance in pageviews of <i>The Texas Tribune</i> interactive pieces (represented as 100%) compared to their related news stories in 2014.....	96
Figure 40 – Evolution in pageviews of <i>The Texas Tribune</i> 2014 interactive pieces and their related news stories over time.	97
Figure 41 – Performance in pageviews of <i>The Texas Tribune</i> interactive pieces compared to the average of the most read articles in 2015 (represented as 100%).....	101
Figure 42 – Performance in pageviews of <i>The Texas Tribune</i> interactive pieces (represented as 100%) compared to their related news stories in 2015.....	103
Figure 43 – Evolution in pageviews of <i>The Texas Tribune</i> 2015 interactive pieces and their related news stories over time.	104
Figure 44 – Performance in pageviews of <i>The Texas Tribune</i> interactive pieces compared to the average of the most read articles in 2016 (represented as 100%).....	108
Figure 45 – Performance in pageviews of <i>The Texas Tribune</i> interactive news pieces (represented as 100%) compared to their related news stories in 2016.	110
Figure 46 – Evolution in pageviews of <i>The Texas Tribune</i> 2016 interactive pieces and their related news stories over time.	111
Figure 47 – Performance in pageviews of <i>O Globo</i> interactive news pieces compared to the average of the most read articles in 2014 (represented as 100%).....	116
Figure 48 – Performance in pageviews of <i>O Globo</i> interactive pieces (represented as 100%) compared to their related news stories in 2014.....	118
Figure 49 – Evolution in pageviews of <i>O Globo</i> 2014 interactive pieces and their related news stories over time.	119
Figure 50 – Performance in pageviews of <i>O Globo</i> interactive pieces compared to the average of the most read articles in 2015 (represented as 100%).....	123
Figure 51 – Performance in pageviews of <i>O Globo</i> interactive pieces (represented as 100%) compared to their related news stories in 2015.....	125
Figure 52 – Evolution in pageviews of <i>O Globo</i> 2015 interactive pieces and their related news stories over time.	126
Figure 53 – Performance in pageviews of <i>O Globo</i> interactive pieces compared to the average of the most read articles in 2016 (represented as 100%).....	130

Figure 54 – Performance in pageviews of <i>O Globo</i> interactive pieces (represented as 100%) compared to their related news stories in 2016.....	132
Figure 55 – Evolution in pageviews of <i>O Globo</i> 2016 interactive pieces and their related news stories over time.	133
Figure 56 – Performance in pageviews of <i>Público</i> interactive pieces compared to the average of the most read articles in 2014 (represented as 100%).	138
Figure 57 – Performance in pageviews of <i>Público</i> interactive pieces (represented as 100%) compared to their related news stories in 2014.....	140
Figure 58 – Evolution in pageviews of <i>Público</i> 2014 interactive pieces and their related news stories over time.	141
Figure 59 – Performance in pageviews of <i>Público</i> interactive pieces compared to the average of the most read articles in 2015 (represented as 100%).	145
Figure 60 – Performance in pageviews of <i>Público</i> interactive pieces (represented as 100%) compared to their related news stories in 2015.....	147
Figure 61 – Evolution in pageviews of <i>Público</i> 2015 interactive pieces and their related news stories over time.	148
Figure 62 – Performance in pageviews of <i>Público</i> interactive pieces compared to the average of the most read articles in 2016 (represented as 100%).	152
Figure 63 – Performance in pageviews of <i>Público</i> interactive news pieces (represented as 100%) compared to their related news stories in 2016.....	154
Figure 64 – Evolution in pageviews of <i>Público</i> 2016 interactive pieces and their related news stories over time.	154
Figure 65 – Average time on page. Most read news are represented as 100%. Interactive pieces’ results are relative to that value.	161
Figure 66 – Average bounce rate. Most read news are represented as 100%. Interactive pieces’ results are relative to that value.	163
Figure 67 – Average exit rate. Most read news are represented as 100%. Interactive pieces’ results are relative to that value.	165
Figure 68 – Interaction techniques most frequent in the data journalism pieces analyzed. ...	168
Figure 69 – Topics most frequent in the data journalism pieces analyzed.....	170
Figure 70 – “Government Salaries Explorer” screenshot.....	221
Figure 71 – “Election Night” screenshot.	222
Figure 72 – “Elected Officials Directory” screenshot.	223

Figure 73 – “Texas Prison Inmates” screenshot.....	224
Figure 74 – “Public Schools Explorer” screenshot.	225
Figure 75 – “Interactive: Search CSCOPE Lesson Plans” screenshot.....	226
Figure 76 – “Map: Comparing the 2010 and 2014 Governor's Races” screenshot.....	227
Figure 77 – “Texas Hospitals Face Penalties Over Infections” screenshot.	228
Figure 78 – “Higher Education Explorer” screenshot.....	229
Figure 79 – “Interactive Map: Find Texas' Remaining Abortion Clinics” screenshot.	230
Figure 80 – “Interactive: Undocumented Students on In-State Tuition” screenshot.	231
Figure 81 – “Ethics Explorer” screenshot.....	232
Figure 82 – “Faces of Death Row” screenshot.....	233
Figure 83 – “Texas Hospitals Explorer” screenshot.....	234
Figure 84 – “Texas Reservoir Levels” screenshot.....	235
Figure 85 – “Interactive: Demographics of Texas' Undocumented Population” screenshot. ..	236
Figure 86 – “The 84 th Texas Legislature, by the Numbers” screenshot.	237
Figure 87 – “Interactive: How Many Texans Resemble You?” screenshot.	238
Figure 88 – “Texas vs. the Feds” screenshot.....	239
Figure 89 – “See Vaccine Exemptions in Texas by School District” screenshot.....	240
Figure 90 – “Ballpark Figures” screenshot.	241
Figure 91 – “Nov. 8 general election results” screenshot.....	242
Figure 92 – “Track how many Texans are voting early” screenshot.	243
Figure 93 – “2016 Primary Election Results” screenshot.....	244
Figure 94 – “See which counties in Texas Trump and Clinton won” screenshot.....	245
Figure 95 – “The Price of Admission” screenshot.....	246
Figure 96 – “Why Isn’t Texas Ready for the Next Big Hurricane?” screenshot.....	247
Figure 97 – “Can you tell real news from fake news? Take our quiz to find out” screenshot..	248
Figure 98 – “Bordering on Insecurity” screenshot.....	249
Figure 99 – “Who Said It: Donald Trump or Mark Cuban?” screenshot.	250

Figure 100 – “Apuração das eleições 2014 – 1º Turno” screenshot.....	251
Figure 101 – “O peso eleitoral dos estados” screenshot.	252
Figure 102 – “As divisões socioeconômicas da votação para presidente” screenshot.	253
Figure 103 – “Evolução das bancadas” screenshot.....	254
Figure 104 – “Os números da Pnad 2013” screenshot.	255
Figure 105 – “A seleção dos sonhos da Copa do Mundo 2014” screenshot.....	256
Figure 106 – “O Brasil nos Mundiais” screenshot.....	257
Figure 107 – “Você sabe qual a sua expectativa de vida?” screenshot.	258
Figure 108 – “Mapa da taxa de suicídio no mundo” screenshot.	259
Figure 109 – “Campanha versus realidade” screenshot.	260
Figure 110 – “DNA do Congresso Nacional” screenshot.....	261
Figure 111 – “A distribuição dos clientes e valores do ‘Swissleaks’” screenshot.	262
Figure 112 – “Ataques terroristas no mundo desde 1970” screenshot.....	263
Figura 113 – “Raio-x dos atrasos dos voos no Brasil” screenshot.....	264
Figura 114 – “Qual o futuro do PT?” screenshot.	265
Figura 115 – “Consulte a media da sua escola no Enem 2014” screenshot.	266
Figura 116 – “Os medalhistas brasileiros” screenshot.....	267
Figura 117 – “Duelo entre os lados sombrio e luminoso da Força” screenshot.....	268
Figura 118 – “Ranking da liberdade de imprensa no mundo em 2015” screenshot.	269
Figura 119 – “Personagens da Lava-Jato” screenshot.	270
Figure 120 – “Placar do Impeachment” screenshot.	271
Figure 121 – “O mapa das coligações” screenshot.	272
Figure 122 – “A votação no Rio de Janeiro, por zona eleitoral” screenshot.....	273
Figure 123 – “Você consegue escalar um time melhor que o Dream Team?” screenshot.....	274
Figure 124 – “Partido do ‘você não me representa’” screenshot.....	275
Figure 125 – “Calculadora de aposentadoria” screenshot.....	276
Figure 126 – “A votação para presidente dos EUA” screenshot.....	277

Figure 127 – “Com qual candidato a prefeito do Rio você mais se identifica?” screenshot. ...	278
Figure 128 – “O trajeto da toca olímpica no Brasil” screenshot.	279
Figure 129 – “Mapa das manifestações do dia 13 de março” screenshot.....	280
Figure 130 – “O ranking das escolas 2014” screenshot.	281
Figure 131 – “O ranking dos preços da água em 2013” screenshot.	282
Figure 132 – “A vida desde 1820” screenshot.	283
Figure 133 – “Taxa de sobrevivência de cânceros em Portugal” screenshot.	284
Figure 134 – “80 anos de Verão” screenshot.....	285
Figure 135 – “As linhas da liberdade” screenshot.	286
Figure 136 – “Retrato dos aeroportos nacionais” screenshot.	287
Figure 137 – “VIH: O vírus que apareceu em Kinshasa em 1920 e alastrou para o mundo inteiro” screenshot.....	288
Figure 138 – “A I Liga vista de dez maneiras diferentes” screenshot.	289
Figure 139 – “Os salários dos gestores do PSI 20 à lupa” screenshot.....	290
Figure 140 – “A água que gastamos mas não vemos” screenshot.	291
Figure 141 – “Radiografia da TAP” screenshot.	292
Figure 142 – “As melhores praias para este Verão” screenshot.....	293
Figure 143 – “O que pomos no prato... não é o que devíamos pôr” screenshot.....	294
Figure 144 – “Relações que matam” screenshot.	295
Figure 145 – “Uma década de pobreza e algumas medidas” screenshot.....	296
Figure 146 – “Esta Primavera vá ver os passarinhos” screenshot.	297
Figure 147 – “Cristiano Ronaldo em golos” screenshot.....	298
Figure 148 – “Menos alunos para as provas do secundário” screenshot.	299
Figure 149 – “Todos os resultados das legislativas” screenshot.....	300
Figure 150 – “Os exames chegaram” screenshot.	301
Figure 151 – “Portugal no pódio olímpico” screenshot.	302
Figure 152 – “Um século de energia em Portugal” screenshot.	303
Figure 153 – “As preocupações e motivações dos professores” screenshot.....	304

Figure 154 – “Como se elege um Presidente nos EUA” screenshot.	305
Figure 155 – “De que é feita a cerveja” screenshot.....	306
Figure 156 – “ZZZZZZ... Zika” screenshot.	307
Figure 157 – “A crise dos refugiados em números” screenshot.	308
Figure 158 – “Os 16 desportos de Portugal no Rio 2016” screenshot.....	309
Figure 159 – “Presidenciais 2016” screenshot.....	310

List of Tables

Table 1 – Definition of Google Analytics metrics used in this study.....	81
Table 2 – Definition of the interaction techniques used in this study.....	83
Table 3 – Hypothetical performances in 2014.....	86
Table 4 – The 10 interactive pieces chosen in 2014 from <i>The Texas Tribune</i> and their specific aspects.....	93
Table 5 – The 10 interactive pieces chosen in 2015 from <i>The Texas Tribune</i> and their specific aspects.....	100
Table 6 – The 10 interactive pieces chosen in 2016 from <i>The Texas Tribune</i> and their specific aspects.....	107
Table 7 – The 10 data journalism pieces chosen in 2014 from <i>O Globo</i> and their specific aspects.....	115
Table 8 – The 10 data journalism pieces chosen in 2015 from <i>O Globo</i> and their specific aspects.....	122
Table 9 – The 10 data journalism pieces chosen in 2016 from <i>O Globo</i> and their specific aspects.....	129
Table 10 – The 10 data journalism pieces chosen in 2014 from <i>Público</i> and their specific aspects.....	137
Table 11 – The 10 data journalism pieces chosen in 2015 from <i>Público</i> and their specific aspects.....	144
Table 12 – The 10 data journalism pieces chosen in 2016 from <i>Público</i> and their specific aspects.....	151

Introduction

The Google Trends website¹ enables the analysis and comparison of search queries popularity in Google engine. In other words, “this allows us to measure the interest in a particular topic across search, from around the globe” (Rogers, 2016, p. n.p.). Figure 1 shows the interest in the topic of “Data Journalism” over time. Insignificant until 2010, curiosity about the subject begins to grow in the period in which major news organizations produce a series of interactive visualizations to make sense of the large volume of data leaked by WikiLeaks.²

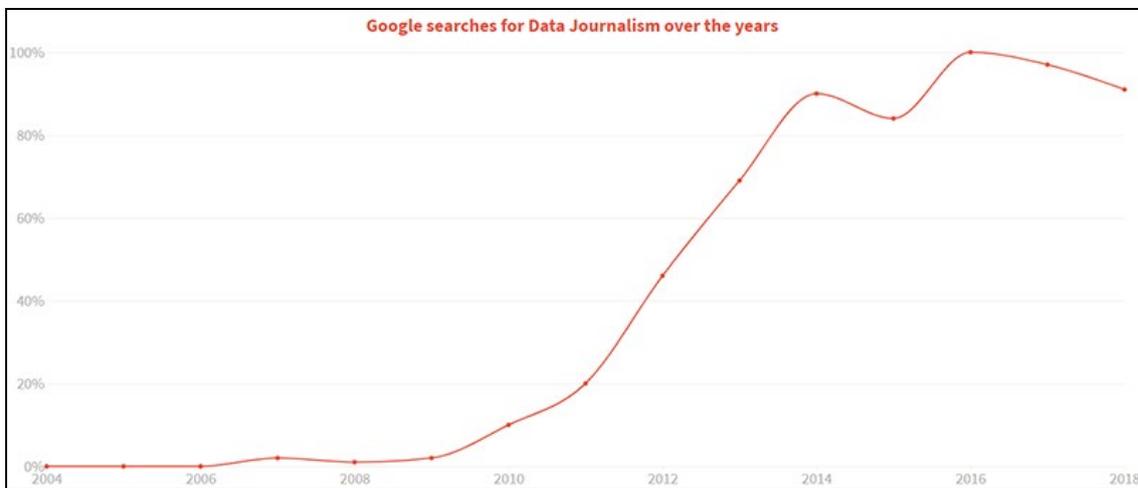


Figure 1 – Search interest related to the topic of Data Journalism between 2004 and 2018 worldwide. A value of 100 is the peak popularity for the term. A score of 0 means there was not enough data for the term.

The line chart makes it clear that people’s interest in data journalism has never ceased since. Neither has the creation of interactive pieces in newspapers. On the contrary, data journalism has gained increasing popularity in large and small newsrooms around the world (Gray, Bounegru, & Chambers, 2012; Mair, Keeble, Lucero, & Moore, 2017; Radcliffe, 2017). The 8th edition of the Data Journalism Awards received the submission of 608 projects. Organized by the Global Editors Network,³ the

¹ Available at <https://trends.google.com/trends/>. Accessed on February 25, 2019.

² WikiLeaks is self-described as “multi-jurisdictional public service designed to protect whistleblowers, journalists and activists who have sensitive materials to communicate to the public.” Available at <https://wikileaks.org/wiki/Wikileaks:About>. Accessed on February 25, 2019.

³ Global Editors Network is an “international association of over 6,000 editors-in-chief and media executives with the mission of fostering digital innovation in newsrooms all over the world.” Available at <https://www.globaleditorsnetwork.org/about-us/>. Accessed on February 25, 2019.

Data Journalism Awards are the first international awards in the field of data journalism. In 2012, in the first edition, 200 pieces were submitted. In 2019 the United States was the country with the highest number of submitted projects: 139. 47 were submitted from Brazil, the third country with the highest number of projects. Portugal had 3.

The number of scholarly publications on data journalism has also increased in recent years (Ojo & Heravi, 2018; Souza et al., 2018) showing that researchers are aware of the growing interest in the field in newsrooms. There are studies investigating the production of data journalism pieces in the United States (Fink & Anderson, 2015; Parasie & Dagiral, 2013; Royal, 2010; W. Weber & Rall, 2013; Young & Hermida, 2015), Brazil (Abrás, 2018; Moura, 2018; A. P. B. de Oliveira, 2018; Träsel, 2014a), Portugal (Abrás, 2018; Alexandre, 2014; Martinho, 2013; Moura, 2018), Canada (Hermida & Young, 2017; Roy, 2016), Argentina (Vázquez-Herrero, Negreira-Rey, & López-García, 2019), sub-Saharan Africa (Cheruiyot & Ferrer-Conill, 2018), Spain (Rodríguez & Castilla, 2016), the United Kingdom (Borges-Rey, 2016, 2017; Dick, 2014; Hannaford, 2015), Norway (Karlsen & Stavelin, 2014), Sweden (Appelgren & Nygren, 2014), Netherlands (Smit, de Haan, & Buijs, 2014), Germany (Ausserhofer, 2015; Weinacht & Spiller, 2014), Belgium (De Maeyer, Libert, Domingo, Heinderyckx, & Le Cam, 2015), Russia (Valeeva, 2017), China (Zhang & Feng, 2019), and Australia (Wright & Doyle, 2018).

Often working with public data obtained through the right to access public information and/or with leaked databases, scholars recognize that data journalism can reinforce transparency, accountability, and journalism's watchdog role through scrutiny of those in power (Anderson, Bell, & Shirky, 2014; S. Cohen, Hamilton, & Turner, 2011; Loosen, Reimer, & De Silva-Schmidt, 2017). Works like "The Panama Papers"⁴ and "The Swiss Leaks",⁵ for instance, show that "watchdog journalism can

⁴ The Panama Papers is "an unprecedented investigation that reveals the offshore links of some of the globe's most prominent figures." Available at <https://www.icij.org/investigations/panama-papers/pages/panama-papers-about-the-investigation/>. Accessed on February 25, 2019.

⁵ The Swiss Leaks is an investigation based on "leaked files that provide details on over 100,000 HSBC clients and their bank accounts." Available at <https://www.icij.org/investigations/swiss-leaks/explore-swiss-leaks-data/>. Accessed on February 25, 2019.

rightly be considered a flag-bearer of the growing interest in data journalism” (Waisbord, 2016, p. 4).

Another critical characteristic of data journalism is “its participatory openness and cross-field hybridity” (Coddington, 2015a, p. 337). Interactive news stories encourage active user participation in the exploration of data (Toledo, 2014), and many projects are tools to assist the user in making decisions (Parasie & Dagiral, 2013).

There is research that examines the main topics addressed by data journalism pieces, the interaction techniques implemented, the most used visualization types, the sort of the data sources used in the article, and the number of people involved in each data project (M. Knight, 2015; Loosen et al., 2017; Ojo & Heravi, 2018; Stalph, 2018; Tabary, Provost, & Trottier, 2016; Tandoc & Oh, 2017; Young, Hermida, & Fulda, 2018).

However, there are research divisions within the field of data journalism that have not yet been properly explored. There are not many studies that investigate how the audience behaves and interacts with data visualization. Does the user explore the data when given the opportunity? What are the audience’s favorite interaction techniques? Does the user understand and feel comfortable when faced with more sophisticated data visualization?

Likewise, there is an insufficient number of works that address the performance of data journalism pieces and whether there is a distinct behavior between who accesses interactive stories and other news content. Who performs best in pageviews, the “metric of success” (Groves & Brown, 2011, p. 16)? Who has the highest bounce rate? Does the audience spend more time on interactive news than on static pieces?

All these questions are related and are very important in order for journalists to improve their work by creating data journalism pieces in a clearer, more comprehensive, engaging and memorable way. Thus, it may be that journalists’ fear of dismissing or alienating the data-averse audience decreases.

This work attempts to shed light on the points related to the second group: how is the consumption and performance of data journalism pieces in comparison to other news articles? To answer this main question, we used web analytics data to compare performance, engagement, and audience behavior. We selected interactive

pieces from 3 newspapers from 3 different countries: *The Texas Tribune* (United States), *O Globo* (Brazil), and *Público* (Portugal). 90 interactive pieces were chosen: 10 from each news organization in 2014, another 10 from 2015 and 10 more from 2016. Each interactive piece was compared: (1) with the average of 500 most read news in the year it was published; (2) with two news articles on the same topic and published in the same period as the data journalism piece.

This approach allows, on the one hand, to evaluate the performance of data journalism pieces compared to the most read content of the website. On the other hand, general news organizations such as *O Globo* and *Público* have a wide range of content and the comparison with news pieces on the same topic allows for a more targeted comparison, theoretically with an audience interested in that subject.

We had access to the Google Analytics⁶ data from *The Texas Tribune*, *O Globo*, and *Público*, which allowed us to use real data for analysis, which is rare in audience analytics studies (Zamith, 2015). However, in the results we tend to concentrate on proportions rather than absolute numbers because of potential commercial implications and because it is a way to give more context and meaning to the data.

In any case, we present an overview of the consumption of data journalism pieces in 3 newspapers from 3 different countries. We explain what brings together and what separates the interactive pieces from the other news in each newspaper and what patterns and divergences can be observed in the audience behavior in data journalism pieces from different news organizations.

To accomplish this, we begin by demonstrating that the news pieces with data are quite old. Rich and complex maps and graphics were often used as early as the 19th century. We show that long before Philip Meyer borrowed methodological procedures from social sciences, this field played an important role in the development of data visualization. Finally, we argue that interactive stories of data, rather than being something completely new, are a new phase of a long path, featuring the key points of this journey in the countries of the case study media outlets. This is Chapter I.

⁶ Google Analytics is a web analytics service by Google that “provides extensive data with regard to the characteristics of individuals who visit your website or blog” (Mokalis & Davis, 2018, p. iv).

In Chapter II, a brief history of the audience measurement in traditional media is performed. From the virtual monopoly players in the press, radio, and television to the cacophony of services to measure the online audience. The 3 main methods to measure audience ratings online are described, highlighting their strengths and weaknesses. We review the literature on the increasing influence of web analytics in the newsrooms, showing that, once ignored, the audience is now tracked to the minute, click to click. We conclude by emphasizing that impact and engagement are gaining space among the most important metrics.

We present the methodology used in this work in Chapter III. We describe the 11 metrics used in this work and introduce the questions they raise. We explain how the data journalism pieces will be compared with the most read news stories and the related news pieces. In addition to this quantitative analysis, content analysis is also performed in this work. In this way, 10 interactivity techniques are presented, which will be used to identify the most used functions in the interactive pieces. We also feature 10 topics that will serve to determine the most recurrent main subject of data projects. Finally, the contributions of the adopted approach and its limitations are stated.

Chapters IV, V, and VI present the results of the analyses carried out in *The Texas Tribune*, *O Globo* and *Público*, respectively. In Chapter VII the results presented separately in the previous chapters are discussed together. In this way, we try to understand the points of convergence and divergence of the data journalism pieces from the 3 news organizations in relation to: (1) the performance of this content; (2) level of audience engagement; (3) and the behavior of users. The key findings are stated, the contributions to the field of data journalism, audience data and metrics in online journalism are reiterated and future research on the topic is recommended.

I – Data Journalism: past and present

WikiLeaks played a key role in the popularization of the term “Data Journalism”. In July 2010, the organization (founded by the computer programmer Julian Assange) shared with the German magazine *Der Spiegel*, the British newspaper *The Guardian* and the American newspaper *The New York Times* some 92,000 leaked documents related to the War in Afghanistan. Three months later, in October 2010, Wikileaks released more than 391,000 reports on the Iraq War (Wikileaks & Assange, 2016). The Qatar-based broadcaster *Al Jazeera*, the British broadcaster *Channel 4* and the French newspaper *Le Monde* were added to the list of news organizations with earlier access (Meikle, 2012). The solution found by some of these media outlets to make sense of this large volume of data and tell complex stories in a compelling way was to create interactive data visualization that would help the reader navigate through the information.

The interactive piece “Afghanistan war logs: IED attacks on civilians, coalition and Afghan troops”,⁷ shown in Figure 2, is a good example of how to present a lot of information in a didactic, compelling and interactive way for the reader. The piece displays a map with the 16,000 incidents caused by improvised explosive devices (IED) between January 4, 2004, and December 31, 2009, with data on the location of IEDs and the number of casualties. Along with the interactive map, the visualization features a timeline of all attacks and highlights that contextualize important moments. The reader can zoom in and out to see each incident better, and clicking on a particular attack shows details of the number of casualties and victims. By pressing start, all incidents appear on the map chronologically and automatically.

⁷ Available at <http://www.theguardian.com/world/datablog/interactive/2010/jul/26/ied-afghanistan-war-logs>. Accessed on February 25, 2019.

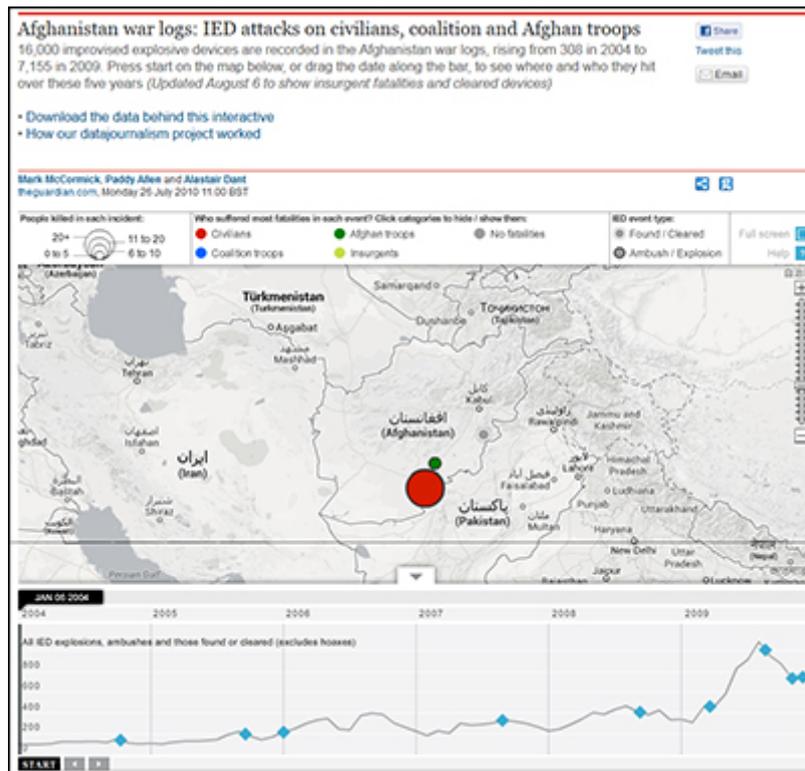


Figure 2 – “Afghanistan war logs: IED attacks on civilians, coalition and Afghan troops”, a data journalism story published on July 26, 2010, by *The Guardian*.

Wikileaks’ exposé – and collaboration with established media organizations – was important because it compelled journalists to move forward in a new approach. Whether for the bold way in which the data were presented or because of the importance of the information and its repercussion, the fact is the episode of the War Logs was a milestone in data journalism, raising this form of reporting to a new level, with much more prominence. Simon Rogers, then *The Guardian*’s data editor and creator of the *Datablog*,⁸ one of the first blogs dedicated to data journalism, acknowledges this.

But the game-changer for data journalism happened in spring 2010, with the WikiLeaks war logs. That is just one set of stories, a single high-point in data journalism’s road to acceptance. There are still reporters out there who don’t know what all the fuss is about, who really don’t want to know about maths or spreadsheets. But for others, this new wave represents a way to save journalism. It represents a new role for journalists as a bridge and a guide

⁸ Available at <https://www.theguardian.com/data>. Accessed on February 25, 2019.

between those in power who have the data (and are rubbish at explaining it) and the public who desperately want to understand the data and access it but need help. We can be that bridge. (Rogers, 2013, p. 43)

Nevertheless, the *Chicago Tribune*, *The Texas Tribune*, *The Washington Post*, *The New York Times* and *The Guardian* are some media outlets that were already producing data journalism before the War Logs (Aitamurto, Sirkkunen, & Lehtonen, 2011; Parasie, 2011). In fact, the collection and display of numerical data and the lineage of quantitative research can be traced back some centuries.

1.1. The roots of Data Journalism

In his report about data journalism, Alexander Howard points out that at least since the Han dynasty – with the *Tipao* – and the Roman Empire – with the *Acta Diurna* –, “governments produced and circulated news of military campaigns, politics, trials, and executions”⁹ (Howard, 2014, p. 9). Scott Klein, deputy managing editor at *ProPublica*, highlights *Bills of Mortality* as another example of data journalism before newspapers (Klein 2016a). The bills were documents with the number of baptisms and mortalities and the cause of death in London. It is not easy to say when these documents began to circulate since no bills survived the Great Fire of London (1666), but since 1603 these bills began to be written by the printer E. Cotes and published weekly by the Company of Parish Clerks (Poovey, 2004), two years before the publication of the first newspaper.¹⁰

Nevertheless, it was neither the first nor the most complete report of the time. Since the plague became endemic in Europe – in the 14th century –, some cities organized disease surveillance and started to identify each victim individually, typically by name, social status, and location of death (Heitman, 2017). Nonetheless, it was

⁹ *Tipao* and *Acta Diurna* were primitive newspapers with a daily record of political and social events. The ancient Chinese gazette (also called *Dibao*) began to circulate during the 3rd century B.C., while the minutes about ancient Rome were first published in 59 B.C. (Demir, 2012; Stephens, 2007).

¹⁰ Historians consider Johann Carolus' *Relation aller Fürnemmen und gedenckwürdigen Historien*, established in Strasbourg, Germany, in 1605, the world's first newspaper (Borchard, 2019; J. Weber, 2006).

would help answer questions of clear civic value.
(Heitman, 2017, p. 13)

If we have Meredith Broussard's (2018) definition of data journalism in mind, "the practice of finding stories in numbers and using numbers to tell stories", Graunt's *Natural and Political Observations on the Bills of Mortality* fits these demands better than *Acta Diurna*, *Tipao* or *Bills of Mortality* itself. Graunt not only finds stories in his investigations but also, by publishing the "Table of Casualties", allows the readers to scrutinize the data, verify his claims and make their own observations.

Newspapers began to print the price of commodities, information on ship arrivals and departures, currencies quotes, etc., since their very beginning. In such cases, as in the examples of Howard and Klein, numbers are records, reproduced without analysis or context. There is no investigative work, nor a "story" as enunciated by Broussard's definition. Only in the 19th century would newspapers start looking for stories in numbers and report them as news, not simply reproducing data.

One of the first examples is from *The Manchester Guardian* on May 5, 1821, shown in Figure 4. The piece contains a table with data about schools in the cities of Manchester and Salford, shows the number of pupils in each one and the annual spending of every institution and analyzes the state of public education in these cities, looking at how many pupils received free education – and how many poor children there were in the city (Rogers, 2013).

In the United States, data journalism began to emerge after the popularization of the penny press¹² in the 1830s. Before that few newspapers reported facts.

It is now widely agreed that the 1830s, a remarkable decade in so many ways, marked a revolution in American journalism. That revolution led to the triumph of "news" over the editorial and "facts" over opinion, a change which was shaped by the expansion of democracy and the market, and which would lead, in time, to the journalist's uneasy allegiance to objectivity. (Schudson 1981)

¹² The newspapers used to be sold for six cents and by annual subscription, while the penny papers began to be sold for a penny on the streets every day by noisy newsboys. Quickly the circulation of the penny papers became far greater than the six-penny journals (Schudson, 1981).

commerce is what we aim at, combined with accuracy, brevity, and spirit.”¹³ With this approach, not only printing records but contextualizing it with analysis, the *Herald* added human interest to the abstractions of the market and its daily tables while developing the best financial section in the penny press (P. Knight, 2016; Roush, 2012).

In the edition of February 5, 1842, in the section “National Statistics – Wealth of the Country”, the *New York Herald* showed what we consider to be one of the earliest examples of data use beyond reproduction in the United States. The newspaper presented in an exclusive and detailed way the value of the annual Gross Domestic Product (GDP) of the United States in 1839, as shown in Figure 5. It then projected the GDP of 1841 to a value 20 percent higher, adding up the annual productions of the United States to \$1.6 trillion and questioning: “Can such a country be long distressed by a debt of only \$200,000,000? Ought such a country to repudiate?”¹⁴

¹³ *New York Herald* (then *Morning Herald*), March 5, 1838. Available at <https://chroniclingamerica.loc.gov/lccn/sn83030312/1838-03-05/ed-1/seq-1/>. Accessed on February 25, 2019.

¹⁴ *New York Herald*, February 5, 1842. Available at <https://chroniclingamerica.loc.gov/lccn/sn83030313/1842-02-05/ed-1/seq-2/>. Accessed on February 25, 2019.

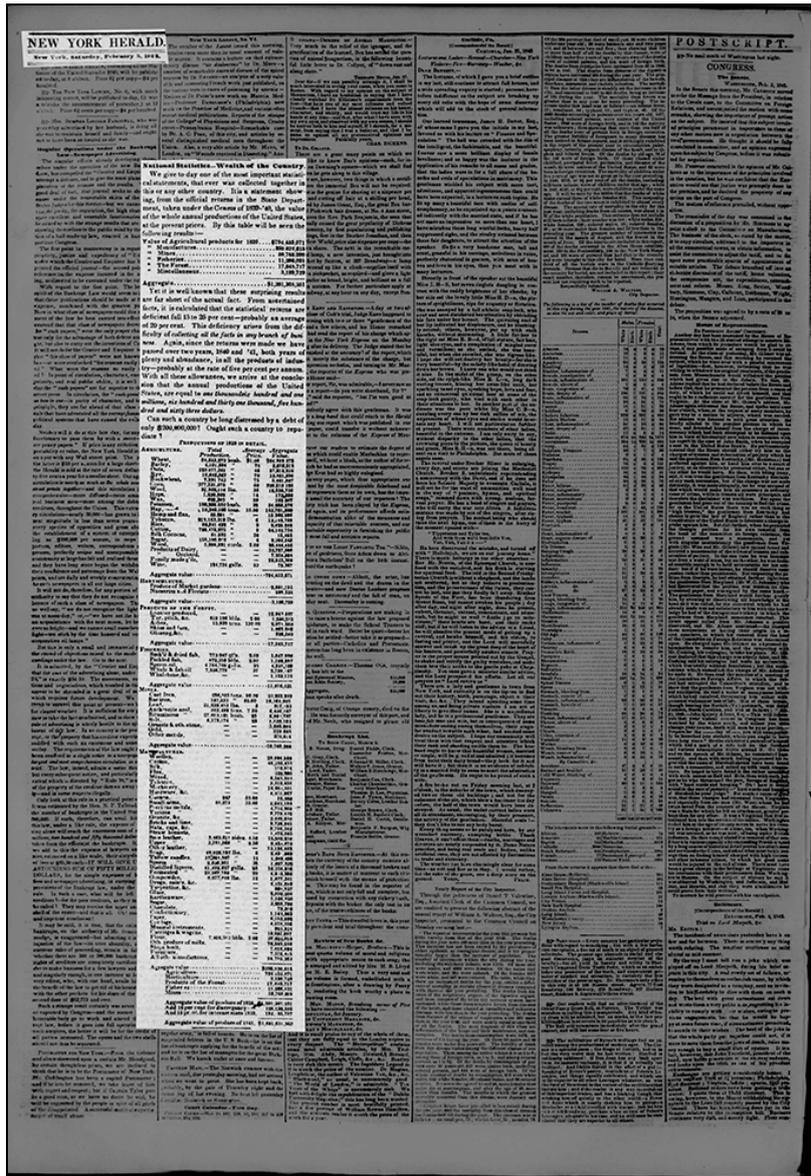


Figure 5 – The highlight shows a data journalism piece published on page 2 of the *New York Herald* on February 5, 1842.

Klein (2016b) credits to one of *Herald's* competitors one of the first data visualizations printed in American newspapers history. The *New York Tribune* published on September 29, 1849, a line chart tracking the deaths in New York City from the cholera epidemic that summer (Figure 6). The graph begins on May 19 and runs through September 22 – a week before its publication – and is accompanied by an explanatory text of how the chart should be read. Klein points out how innovative this chart is:

Although graphical displays of data were not unheard of in scholarly and engineering books, they were

exceedingly rare in U.S. newspapers of that time. Maps became common during the U.S. Civil War in the 1860s, but in the 1840s, illustrations of any kind (let alone data visualization) were rare. I've spent the past few years studying the history of data visualization in news. I have seen hundreds of examples of data displays in antique newspapers. Few have struck me as much as the cholera chart has. It is rare almost to the point of anachronism. (...) The Tribune graphic is a snapshot in miniature of a standard yet to come, revealing a growing sense of awareness of the potential for well-designed information displays to help people understand and solve problems. (Klein 2016b)

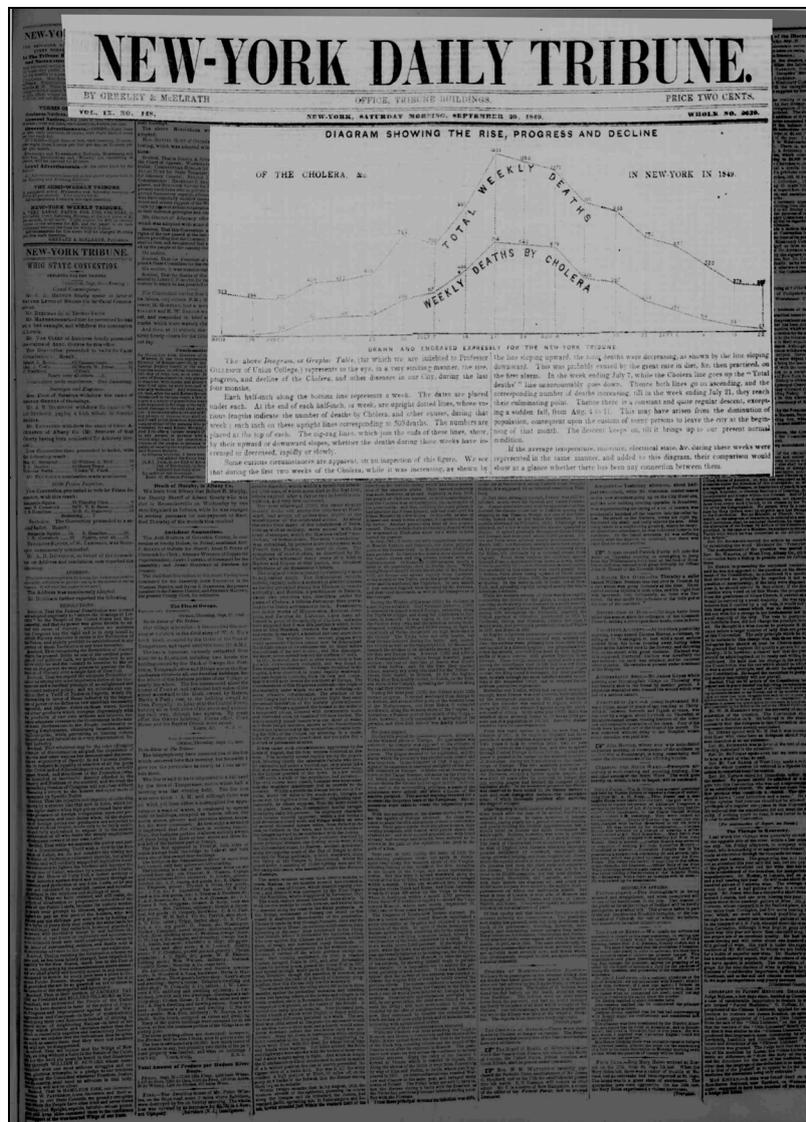


Figure 6 – The highlight shows a data visualization published on the front page of the *New York Tribune* on September 29, 1849.

While in the United States news pieces from data begin to appear with the rise of a “democratic market society” (Schudson 1981, 30), in Brazil the first examples occur still during the reign of Dom Pedro I.¹⁵

Founded in October 1829 by the Italian doctor Giovanni Battista Libero Badaró, the Brazilian newspaper *Observador Constitucional* had a short life, closing three years later in 1832 (C. E. F. de Oliveira, 2009; Sodr , 1999). But in that brief period, the conservative and economically liberal newspaper fought for the freedom of the press and antagonized the president of the province of S o Paulo (Cruz, 2000; Filho, 2013; Sodr , 1999). One of these confrontations took place in the edition of January 17, 1831 (Figure 7).

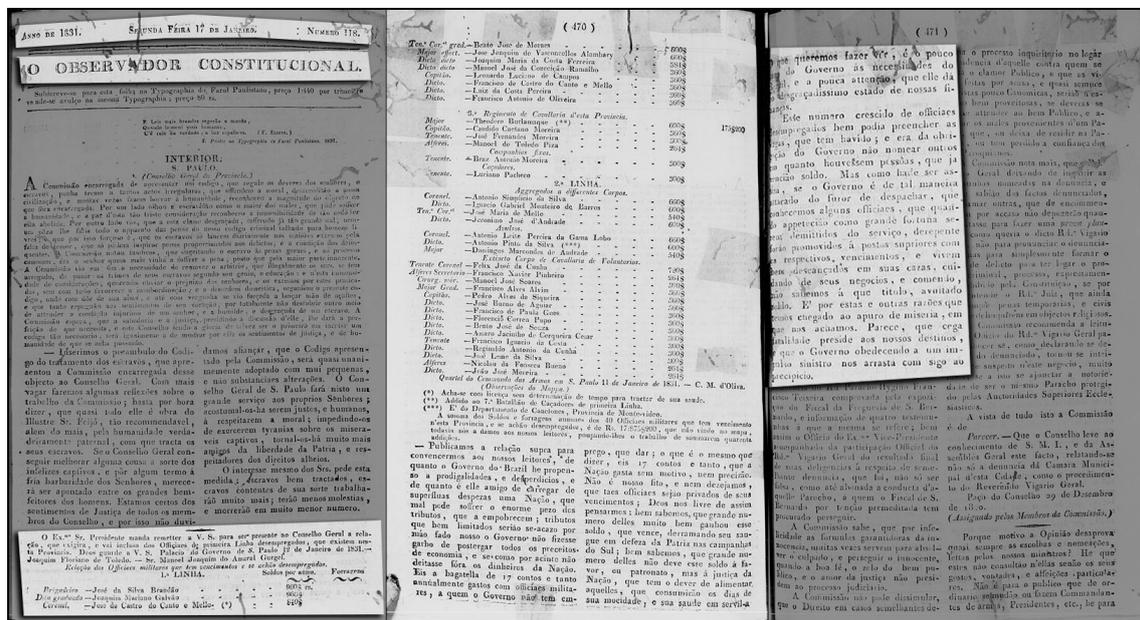


Figure 7 – The highlight shows a data journalism piece published on pages 1 to 3 of *O Observador Constitucional* on January 17, 1831.

The paper printed a news piece with the name, position, and salary of 40 military officers who were unemployed but continued to be paid by the province of S o Paulo, totaling more than 17 *contos de r is*. At one point the editor complained about these expenses: “We publish the above account to convince our readers of how much the Government of Brazil is prone to prodigalities and waste (...) Here is the bagatelle of 17 *contos* annually spent with military officers to whom the Government

¹⁵ Dom Pedro I reigned the Empire of Brazil from 1822 to 1831.

has no job to give; which is the same as saying, here are 17 *contos* that the Nation spends without reason or need.”^{16,17}

In Portugal, the penny press and the early newspapers that were predominantly interested in news and facts began to circulate in 1865, following the pioneer *Diário de Notícias* (Sousa, 2009; Tengarrinha, 2013; Traquina, 2004). One of the earliest examples of data journalism can be found in the *Diário Popular*. Founded by the politician and professor Mariano de Carvalho in 1866, *Diário Popular* was one of the most prestigious newspapers of the monarchy (P. J. Fernandes, 2011a, 2011b; Tengarrinha, 2013).

In the edition of March 9, 1868, shown in Figure 8, the newspaper dismantled the proposal of a commission that said it was possible to build popular houses whose rents would cost between \$15 and \$30 thousand *réis*. With the reproduction of a possible model of a house for workers and the list of costs of works (land, material, and professionals), it shows that the cost would not be less than \$45 thousand *réis*. The piece ends in a tone of challenge: “Here, then, is the utopia undone. They want to build houses for rents of 15 and 30 thousand *réis* and a stingy budget does not allow them to get it for less than 10 pounds a year. If you can respond.”¹⁸

¹⁶ In the original: “Publicamos a relação supra para convenceremos aos nossos leitores, de quanto o Governo do Brazil he propenso a prodigalidades, e desperdícios (...) Eis a bagatella de 17 contos e tanto annualmente gastos com officiaes militares, a quem o Governo não tem emprego, que dar; o que é o mesmo que dizer, eis 17 contos e tanto, que a Nação gasta sem motivo, nem precizão.”

¹⁷ *Observador Constitucional*, January 17, 1831. Available at <http://memoria.bn.br/docreader/814326/261>. Accessed on February 25, 2019.

¹⁸ In the original: “Aqui está pois a utopia desfeita. Querem construir casas para rendas de 15 e 30 mil réis, e um orçamento mesquinho não lhes permite obtel as para menos de 10 libras por anno. Se poderem, respondam.”

to be easily accessible in other countries: “one deficiency in our customs statistics is that we do not classify imports into the three major categories that are of most interest to a country's economy – food, industrial raw materials and factory products – as we see in English or French trade.”¹⁹



Figure 9 – The highlight shows a data journalism piece published on the front page of *O Repórter* on April 21, 1888.

¹⁹ In the original: “Um defeito, que tem as nossas estatísticas aduaneiras, é não classificarem as importações nas tres grandes categorias que mais interessam a economia de um paiz – os alimentos, as materias primas industriaes e os productos fabris – como se ve nos quadros commerciaes ingleses ou francezes.”

1.2. Databases, statistical graphics and the “Golden Age” of data visualization

As Klein (2016b) points out, statistical graphs were very rare in newspapers in the 1840-1850s, in part for technical reasons. The technological advances brought by the industrialization of the newspapers in the mid to late nineteenth century made the impression of maps, images and charts less laborious. This period also witnessed an outbreak of data provided by scientists, companies, state and federal bureaus, etc. (Reilly 2017; Örnebring 2010; Douglas 1999; Schudson 1981). Although databases are present in journalism from the beginning, it is in this period, with the use of documents (Anderson, 2015), the invention of the interview (Schudson, 1996), and the commitment to objectivity and facts (Schudson, 1981, 2001), that databases begin to gain some prominence in newsrooms. At the end of the 1870s, early box scores for sports arose, and more complex tables appeared with the rise of specialized business journals, such as the *Wall Street Journal* in 1889, and by 1896 maps with electoral information emerged on front pages (Usher, 2016). Nevertheless, it is outside the newsrooms that databases, charts, and maps develop more during this period.

In the United States, for example, Ellen Gruber Garvey (2013) imagines a prehistory of the database claiming that by isolating and recontextualizing data collected in a giant mass of clippings, the book *American Slavery As It Is: Testimony of a Thousand Witnesses*, published anonymously in 1839, helped to create the modern concept of information. The authors of the book, the abolitionists Theodore Dwight Weld and the Grimké sisters, gathered thousands of runaway slave advertisements on papers and firsthand testimonies of southern slaveholders to expose the tremendous brutality perpetrated against slaves in the United States (Trudeau & Morrissey, 2017). “It was the work of trimming, sifting, and aggregating the material that recreated it as a database and not just a collection of anecdotes. This work allowed for its recontextualization and analysis” (Garvey 2013: p. 96-97).

In Europe, since Graunt, statistics and databases began to be used in several countries in different areas, to calculate life insurance and war finance, imports and exports, in public finance and public health, in astronomy and physics, etc. (Hudson & Ishizu, 2017; Poovey, 2004). Only in the 19th century, however, great advancements

start to happen in the collection and display of statistics. First with the “statistical movement” in the United Kingdom, a mobilization that appears in the 1830s and the 1840s as an attempt to better understand the consequences of industrialization and to promote social reform. The Board of Trade's Statistical Department, the General Register Office, the London (later Royal) Statistical Society, the Manchester Statistical Society and 20 other provincial societies in Great Britain were created in those years and, together with private bodies, conducted surveys of health and sanitary conditions of the working classes (Cullen, 2016; O'Brien, 2011; Porter, 2011).

Some of the leaders of this movement, such as Richard Jones and Charles Babbage, explained that their methods of investigation differed from others by embracing the facts and rejecting speculation and opinions (Poovey, 2004). Although, as some scholars point out (Cullen, 2016; Hudson & Ishizu, 2017; Porter, 2011), these early statisticians often confused propaganda with facts, leaning their research to reinforce their own prejudices and opinions.

The statistical movement as a whole and the nature of data collected and classified in the nineteenth century whether by public or voluntary agencies shared a common purpose: to assist economic and social engineering and social reform. (...) In addition, the social, political and economic preoccupations of nineteenth-century bureaucrats, civil servants and reformers conditioned and controlled the nature of the information gathered and of that left out. Even the very act of quantification itself, across the many areas of social and economic data, necessarily ignored the array of meanings and connotations that attached to things measured. (...) One of the major biases of statistics gathered at this time came from the tendency amongst reformers to vindicate industrial progress by blaming social problems on other causes such as the growth of cities, alcohol consumption, the moral weakness of the poor or the evils of Anglicanism. (Hudson & Ishizu, 2017, pp. 32–33)

In 1854, the physician John Snow disclosed a detailed street map of cholera deaths in his Soho neighborhood of London (Figure 10). He argued that the disease was exclusively waterborne and not airborne, and his map illustrates a geographical relationship between cholera cases and the public water pump on Broad Street, the source of contamination (Koch, 2017; Sederstrom, 2018). Snow obtained much of his data from the Registrar-General’s *Weekly Return* that was compiled by William Farr, member of the statistical movement (Bingham, Verlander, & Cheal, 2004; Cicak & Tynan, 2015).

Another member of the statistical movement was Edwin Chadwick, who had as his disciple the nurse Florence Nightingale (Gill & Gill, 2005). In 1858, Nightingale published her rose diagram, which showed that more British soldiers died from illnesses than from battle wounds (Figure 11): “with her diagrams, Nightingale was able to capture the whole picture of the disaster, from the high mortality rate to the cause of death to the reasons for the disaster and its solution” (Brasseur, 2005, p. 180).

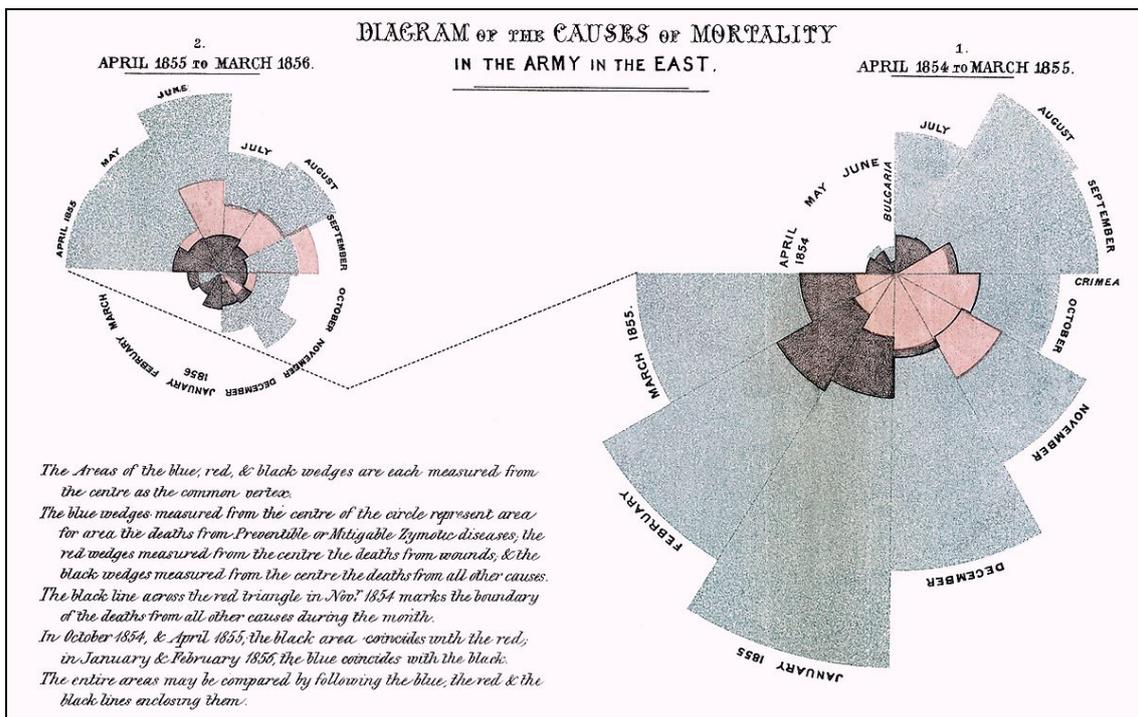


Figure 11 – Reproduction of “Diagram of the causes of mortality in the army in the East” made by the nurse Florence Nightingale in 1858.

Another well-known historical representation of wartime mortality was designed not by a British, but a French engineer. Charles Joseph Minard in 1869 presented a flow map that illustrated Napoleon’s deadly Russian campaign of 1812

journalists' belief in facts was so strong by the end of the nineteenth century, but why it was no stronger" (1981, 77).

Part of the answer lies in the fact that American newspapers at the time advocate that the news had to have a human interest, must be both informative and entertaining, with facts, but also lively and colorful, even if it required some elements of fiction. Therefore, some editors and reporters were less interested in facts than in increasing the paper circulation and creating popular styles of writing (Douglas 1999; Schudson 1981). By the mid-nineteenth century, Bennett's *Herald* and Charles Dana's *New York Sun* routinely used picturesque interviews with fictional elements, celebrity gossip, sensational headlines, and other tricks to retain readers. These techniques will be appropriated and developed some years later by Joseph Pulitzer's *New York World* and William Hearst's *New York Journal*. These techniques will also cross the Atlantic and succeed in Great Britain much because of W. T. Stead, editor of the London-based *Pall Mall Gazette* (Wiener, 2011).

In his famed series on child prostitution and trafficking in London, "The Maiden Tribute of Modern Babylon" (1885), Stead used some of the techniques he learned from American newspapers such as fictionalization and narrativization, sensational and provoking subheadings – "The Violation of Virgins", "Strapping Girls Down", and "Virgins Willing and Unwilling". The articles were the high point of Stead's journalistic career and promoted a debate that led to Parliament passing legislation raising the age for sexual consent for girls from 13 to 16 years old. At the same time, Stead was charged for "peddling pornography" and imprisoned for using illegal investigative methods (Baylen, 1972; Regard, 2014; Westwood, 2010).²⁰

A couple of years later, in 1887, the English poet Matthew Arnold published a critique of this style of reporting that he credited to W. T. Stead and called it New Journalism.²¹

We have had the opportunities of observing a new journalism²² which a clever and energetic man has

²⁰ Stead purchased a child virgin from her impoverished mother to show how easily it could be done (Baylen, 1972; Regard, 2014).

²¹ Although Stead cannot be referred to as inventor of the New Journalism, he was definitely one of the pioneers and one of the great enthusiasts in Britain of the sensational style (Griffiths, 2015).

lately invented. It has much to recommend it; it is full of ability, novelty, variety, sensation, sympathy, generous instincts; its one great fault is that it is feather-brained. It throws out assertions at a venture because it wishes them to be true; does not correct either them or itself, if they are false; and to get at the state of things as they truly are seems to feel no concern whatever. (Arnold, 1887, p. 638)

Arnold was concerned to establish a clear line between journalistic style and literary writing, since now the frontier between news and romance, fact and fiction was deliberately blurred. But the novelization of the news was a success, with reporters portraying London as a dense and dangerous urban forest, describing the slums and the degrading living conditions of the poor working class (Griffiths, 2015; Soares, 2017; Topalov, 1993).

Ever since the 1830s and the 1840s Britain and other urbanizing countries had seen the growth of fictional accounts, journalists' reports, and inquiries which had purported to take the middle-class reader into the exotic world of the urban worker. By the late 1880s the textual features of this established tradition were firmly fixed. The author takes the reader by the hand for a guided tour through the poor districts, passing through streets, courts, and dead-ends, and sometimes through dwellings and workshops. A variety of sad or shocking episodes are observed, and the reader is progressively introduced to another human race, with its habits, warts, hopes, and all. Dickens and Mayhew were experts in this genre, but it also attracted authors who were neither novelists nor journalists. (Topalov, 1993, p. 411)

The scholar Isabel Soares (2017) argues that the New Journalists were giving visibility and exposing to the broader public the living conditions of determined population segments while analyzing the causes and consequences. Notwithstanding, the New Journalism failed to understand and explain the Trafalgar Square riots in February 1886. During that very harsh winter in London, the industrial depression was

²² Currently this genre from the 1800s is better known as Literary Journalism, while the movement that emerged in the 1960s with authors such as Tom Wolfe and Truman Capote is most commonly acknowledged as "new journalism" (Soares, 2017).

at its worst and the lack of work and the hunger pushed the inhabitants of the East End to unprecedented public actions (Bales, 1999; Topalov, 1993). In the following days, Stead's *Pall Mall Gazette* conducted inquiries and during February and March published a series of reports that precipitated frenzy, drawing attention to the chronic poverty of London and asserting that at least 25% of the population was continually on the limit of distress (Hennock, 1976). As Bales put it: "the working class of the East End was perceived as a threat by much of the rest of London, by opinion shapers and policy makers. It was generally believed that a serious threat to public order existed, and reputable journals discussed the possibility of social revolution" (1999, p. 154).

This widespread panic was one of the reasons why the industrialist and shipowner Charles Booth embarked on his own survey to answer a simple question: how many people lived in poverty in London? Booth was extremely critical of this literature and news stories mainly because they blocked a detailed appreciation of the phenomenon of poverty and Booth "wanted first of all to establish the facts objectively, and to determine the character and the exact dimensions of the evil" (Topalov, 1993, p. 400). The answer was published between 1889 and 1903 in 17 volumes in his massive survey *Life and Labour of the People of London* that provided a portrait of the lives and occupations of Londoners.

Booth and his 20 researchers collected data from over 120,000 households and found that as much as one-third of London was living in poverty, while among those living in the city's East End the proportion was 35% of the population. These results were seen as a breakthrough, nevertheless, Booth's survey also showed that the assumptions disseminated by politicians and newspapers were wrong: poverty was not concentrated in one place, there were pockets of poverty throughout the city; the poorer working classes were unlikely to organize any form of action, let alone a revolution; poverty was caused by social problems (unemployment or underemployment) rather than by moral problems, such as alcohol abuse (Bales, 1999; Dorling, Mitchell, Shaw, Orford, & Davey Smith, 2000; Vaughan, 2018).

While the Statistical Movement confused propaganda with facts and the New Journalism "[threw] out assertions at a venture because it wishe[d] them to be true" (Arnold, 1887, p. 638), Booth created a database of statistics on wages and households

and expressed surprise with his findings on poverty, since “he had not expected the figure to be so high, because he had expected the difference between East London and the rest of London to be greater than it turned out to be” (Hennock, 1976, p. 72). The methods of social statistics developed by Booth were innovative. On the one hand, he created a taxonomy of classes, defining and identifying seven social classes, based on income combined with employment patterns and status. On the other hand, his methods of visual representation also went beyond anything already attempted by social scientists (Topalov, 1993; Vaughan, 2018).

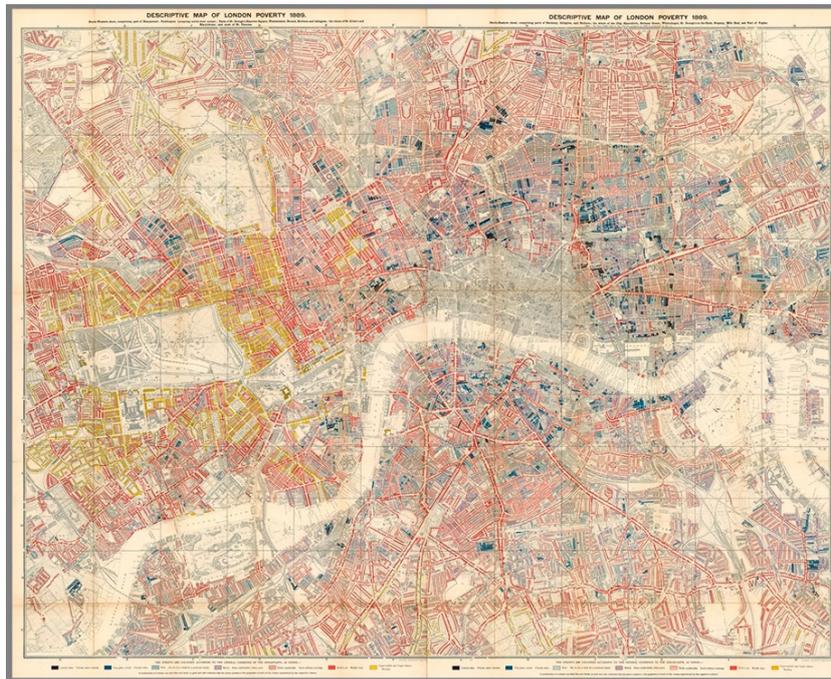


Figure 13 – Reproduction of the industrialist Charles Booth’s *Descriptive Map of London Poverty* published in 1889.

Booth plotted those classifications by households in a series of maps, the most important being the aforementioned *Descriptive Map of London Poverty*, shown in Figure 13, and which exposed, building by building, the streets of London colored to correspond to one of seven categories reflecting the condition of poverty of the resident of the time (Orford, Dorling, Mitchell, Shaw, & Smith, 2002). Booth’s poverty map played a fundamental role drawing attention to the social problems of metropolises.

This was probably the first thematic map to cover an entire city together with the spatial distribution of the ‘social conditions’ of its population. It was a new

language, the properties of which were visible with a unique clarity at this very beginning. This way of representing the city was indeed very different from the prevailing methods of the day and, in some respects, went against them. It implied formal constraints which delineated the content of its silent discourse, and produced entirely distinct cognitive and social effects. (...) The 'poverty map', however, innovated in terms of its global character and its objective. It did not portray a particular problem, but urban society itself, in its full spatial extension. It was thus a comprehensive discourse on the relationship between the social classes and urban space, probably the first synthetic representation of a social ecology. Heir to a sanitary tradition, it launched the methods of an urban sociology. (Topalov, 1993, pp. 408–412)

Booth's poverty map not only sheds light on a much-debated but little-researched subject and tells a story of the phenomenon of poverty in the *fin-de-siècle* London but also allows readers to come to their own conclusions, something Booth did consciously.

I do not wish to add any colour that can be avoided to the white light in which I have persistently tried to work in making researches into existing facts. What I have endeavoured to present to my readers is a picture or a way of looking at things, rather than a doctrine or an argument. I have been glad to see my book furnish weapons and ammunition for absolutely opposed schools, and can even make shift to stifle my annoyance when it is occasionally quoted in support of doctrines which I abhor. (Booth, 1897, p. 440)

Heir to Nightingale and Snow, Charles Booth is acknowledged as the father of social investigation and one of those responsible for popularizing the idea of a poverty line (Desmond & Western, 2018; Gillie, 1996; Vaughan, 2018). Booth's groundbreaking work had a great influence on the "Social Survey movement" in the United States. According to the communication researcher C. W. Anderson, "Booth's *Life and Labour* would inaugurate a mania for gathering data about urban living conditions, an obsession that intersected with deep belief that quantifying the existence of poverty would provide reformers with tools to end it" (Anderson, 2018, p. 24). Indeed,

between the turn of the century and the early 1930s, nearly 3,000 studies were undertaken by large and small cities across the United States (Friedland & Campbell, 2011).

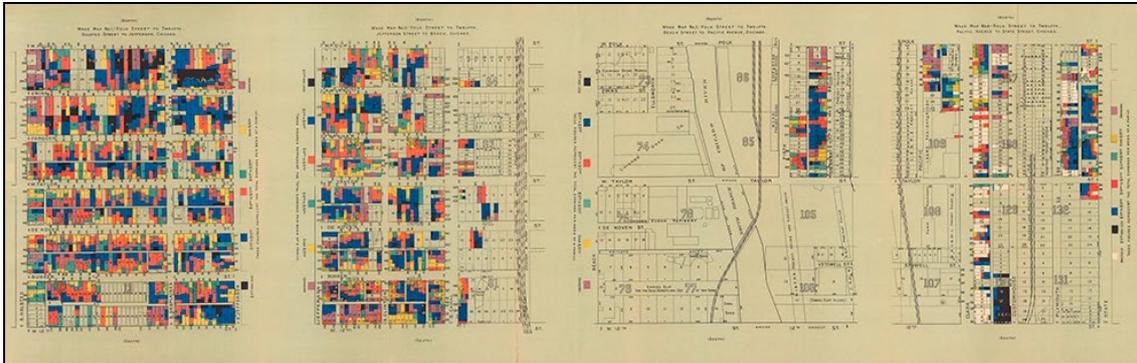


Figure 14 – Reproduction of Hull-House’s Wage Map 1-4 by Residents of Hull-House (1895).

Hull-House Maps and Paper (1895), for example, published by the residents of the settlement house Hull House in Chicago and led by the activists Jane Addams and Florence Kelley, was modeled on Booth’s work to create the survey and maps with statistical information about the residents such as nationality, household size, income, ethnicity, etc. (M. Deegan, 2004; Friedland & Campbell, 2011). Figure 14 gathers the four maps with the weekly income of each family in the neighborhood.

Aware of the work of Jane Addams and Florence Kelley, W. E. B. Du Bois published *The Philadelphia Negro* (2010), a study that intended to understand the causes of the social problems of African Americans living in the Philadelphia’s Seventh Ward (M. J. Deegan, 1988; O’Connor, 2009). Like Booth and Addams and Kelley, Du Bois developed detailed, scaled maps of demographic characteristics and throughout the book repeatedly used charts, maps, and graphs to present the data collected in the most effective way. Figure 15 shows a compilation with some graphs present in the study.

All these visual representations – the majority quite sophisticated – allow the reader to gain insights from the visualization (Card, Mackinlay, & Shneiderman, 2007). Unfortunately, it is very difficult to find in the newspapers of the 19th and early 20th century news graphs or maps with consistency of data and clarity of information that we find in these data visualization created by sanitarians, engineers and social

reformers. Only in the mid-twentieth century will data visualization as informative and appealing as these begin to appear with some frequency in the newspapers.

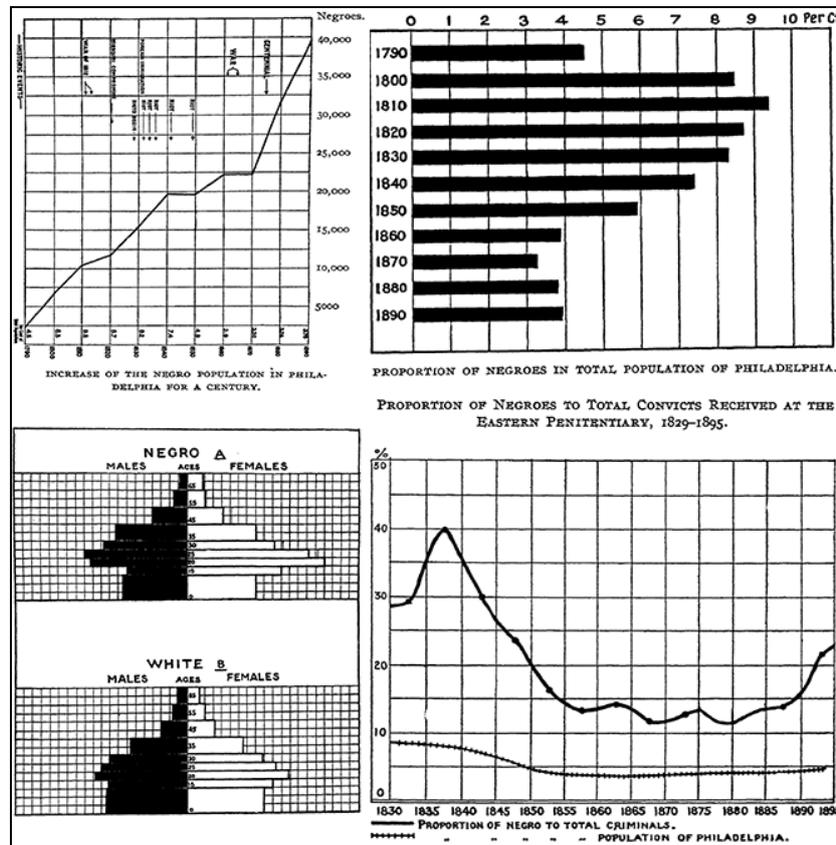


Figure 15 – Compilation of charts and tables present in the sociologist W. E. B. Du Bois’ study *The Philadelphia Negro* (2010).

Nevertheless, Anderson (2018) sees in these earliest social scientists and social reformers crucial players in the development of quantitative information in journalism,²³ since the lines separating journalism, social reform, data collection (and visualization), and the social sciences were largely blurred in the early 20th century. Some social scientists used journalistic techniques in their research (e.g. do interviews, pursue facts) while some newspapers were beginning to share some ideas from the Social Survey movement, such as the belief in the natural and unambiguous truth of quantitative data and the visual power of graphical representation of statistics. Not only that, but also some of the most prominent earliest social scientists were journalists. Du Bois worked as a journalist, the major sociological study *Pittsburgh Survey* (1907-1908) – full of graphs, charts and maps – was conducted by journalist

²³ Some scholars consider data journalism as one of the quantitative forms of journalism, with computer-assisted reporting and computational journalism being the others (Coddington, 2015a).

Paul U. Kellogg, and the academic discipline of sociology in the University of Chicago was created with the help of the former journalist Robert Park, one of the most influential figures in sociology (Friedland & Campbell, 2011).

Anderson (2018) mentions the social surveys of Reverend Charles Stelzle, a key figure in the Men and Religion Forward Movement (MRFM), as an *avant-garde* example in the use and publicity of quantitative data and data visualization. MRFM began as an effort to figure out how to increase male participation in the church, since the leaders of the movement feared the “feminization” of American society. However, the Movement’s attention soon shifted to a broader concern for social justice, documenting not only church attendance, but also poverty and alcohol consumption, for example (Bederman, 1989; Putney, 2003).

Although there are no charts, tables, or maps on the seven volumes with the MRFM findings, Anderson (2018) argues that Reverend Stelzle acknowledged the power of visual representations and that charts and graphs were widely used in pamphlets and posters on the survey exhibits the Movement held at more than 7,000 meetings in about 70 American cities. Figure 16 shows some of these graphs and charts present in *A Modern Church to Meet a Modern Situation* (1912), organized by Stelzle. These data visualizations were used to (1) attract the attention of people who did not read newspapers; (2) transmit information in few words; (3) stimulate debate among visitors; (4) draw the awareness of the press to the survey findings.

First the data had to be gathered, then displayed, and then transmitted to news organizations so it could be publicized more broadly. For these quantitative reformers, “social survey reportage”²⁴ was a device that both represented reality and mobilized the American public to action in support of progressive goals. Because the MRFM straddled the line between journalism, data gathering, and activism, the role quantitative information played in assembling a particular public was effaced – good science made for

²⁴ Anderson defines social survey reportage as a “phenomenon that stands at the crossroads of a variety of important epistemological paths: a progressive and ameliorative thrust in segments of the larger US political culture, a belief in the importance of collecting on-the-ground empirical evidence, and an obsession with counting and with the power of numbers to illuminate important truths.” (Anderson 2018, 21)

good visualization, and good visualization would activate the public for political ends, ends which were not in question. Key to all this, however, was the publicity-generating potential of newspapers. (Anderson, 2018, p. 41)

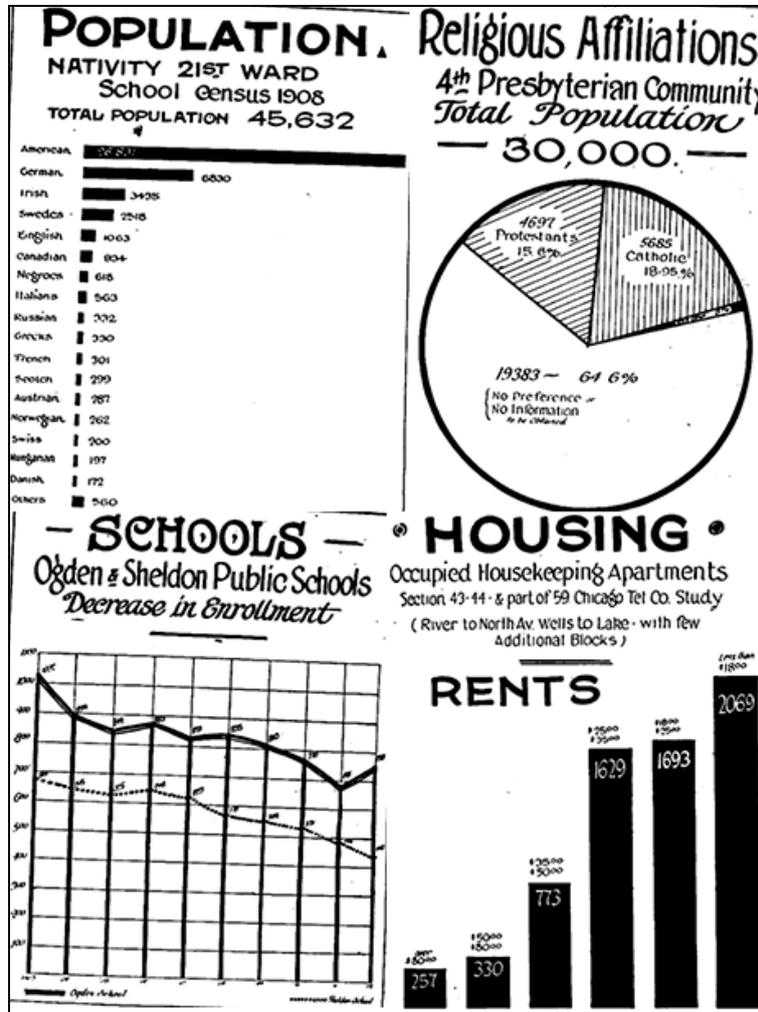


Figure 16 – Compilation of charts and tables present in *A Modern Church to Meet a Modern Situation* published in 1912 and organized by Reverend Charles Stelzle.

The strategy of using data visualization to publicize the findings in the newspapers worked very well. Anderson (2018) found 409 news stories about MRFM between 1911 and 1912, with vastly positive coverage. However, according to Anderson, few newspapers analyzed the data visualization techniques applied by the Movement or critically scrutinized the survey statistics and almost no newspaper reproduced graphs and charts produced by MRFM. The map shown in Figure 17 with data about religious beliefs in the United States, published on March 24, 1912, in the *Sun*, is a rare exception.

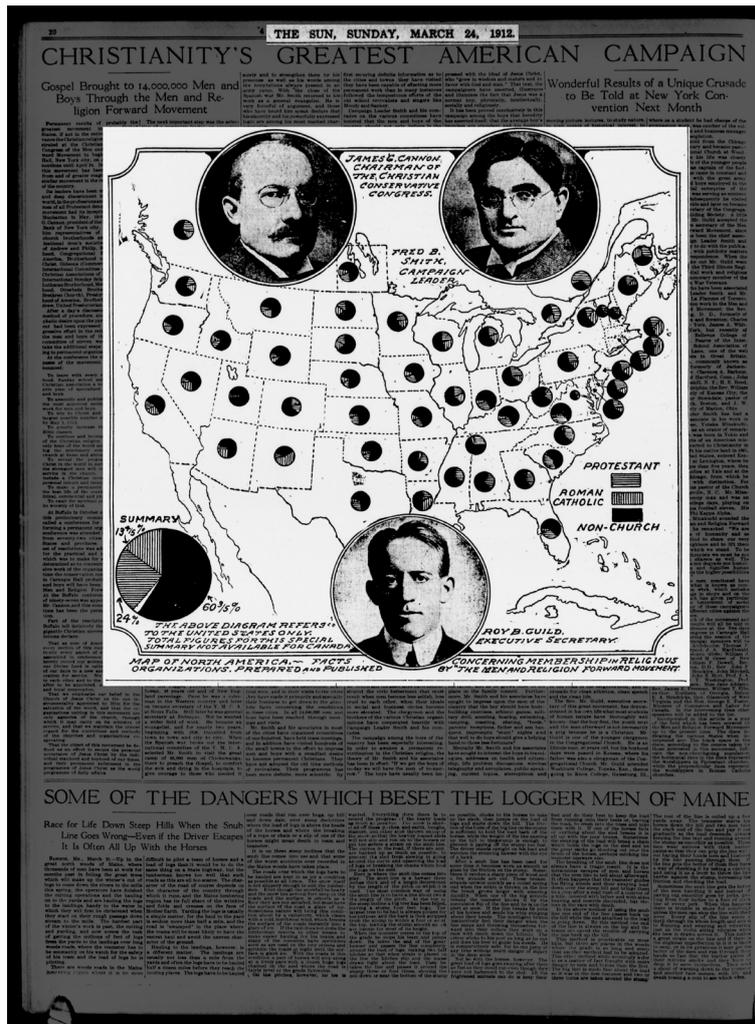


Figure 17 – The highlight shows a data visualization published on page 10, 4th section of *The Sun*, on March 24, 1912.

After World War I, social science began to become professionalized and to move away from journalism, an illegitimate science. If the use of information visualization was a constant in the work of earliest social scientists, from the 1920s onwards, visual representations began to be used less and less by sociologists because it was not scientifically objective enough.

There were few graphical innovations, and, by the mid-1930s, the enthusiasm for visualization which characterized the late 1800s had been supplanted by the rise of quantification and formal, often statistical, models in the social sciences. Numbers, parameter estimates, and, especially, those with standard errors were precise. Pictures were — well, just pictures: pretty or evocative, perhaps, but incapable of stating

a “fact” to three or more decimals. Or so it seemed to statisticians. (Friendly, 2008, p. 37)



Figure 18 – The highlight shows a data visualization published on the front page of *The New York Times* on November 3, 1920.

In journalism, on the one hand, visual representations were a huge success in the Yellow Press²⁵ since the end of the 19th century. As stated before, newspapers such as Pulitzer’s *New York World* and Hearst’s *New York Journal* preferred the sensational over the factual. Hence, illustrators had creative freedom and valued illustrations and pictorial representations more than data charts and news graphics to attract readers (Spencer 2007; Schudson 1981). On the other hand, news graphics, especially maps, became less unusual in newspapers after World War II and very common only with the popularization of computers and computer graphics, by the mid-1980s (Monmonier,

²⁵ This genre emerges a few years after New Journalism and Ted C. Smythe describes it as follows: “It was New Journalism carried to an extreme. Headlines were larger and bolder and scare heads attracted readers. Illustrations no longer reflected reality” (Smythe, 2003, p. 210).

1999). Of course, there are exceptions like the front page of *The New York Times* edition of November 3, 1920, that, as shown in Figure 18, featured the piece “Map Showing How the States Voted”, with the results of the US presidential election on a map.

If the newspapers used few statistical graphs in this period, some magazines used them quite frequently and with high quality. A good example is the *Fortune* surveys section, a regular and famous feature in the US business magazine displaying statistical visualization of public opinion polls on several topics. The November 1940 issue, for example, features a series of data visualization with public opinion on the candidates of the 1940 US presidential election, as shown in Figure 19.

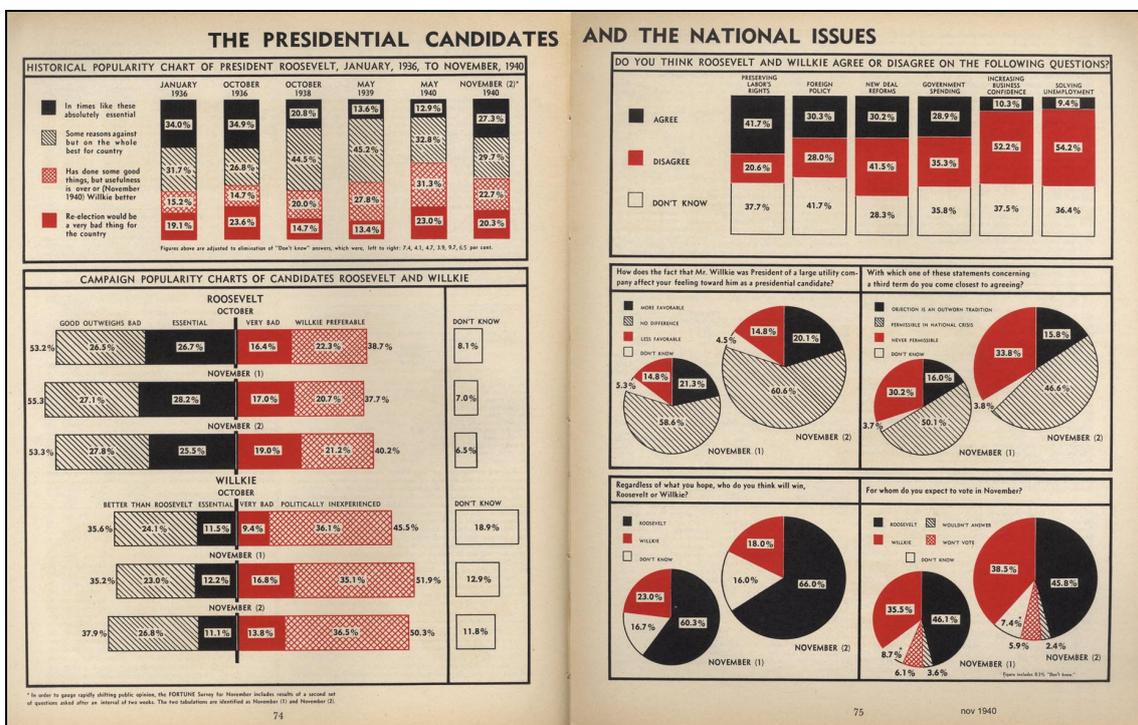


Figure 19 – Reproduction from *Fortune's* November 1940 edition, pages 74 and 75.

News charts were also used in Portuguese and Brazilian magazines, albeit more sporadically. Figure 20 shows two examples of timeline graph used by the Portuguese magazine *Seara Nova* during the 1940s. The top one is from the edition of March 22, 1941 and brings the fluctuations of American public opinion to the situation of the United States in the face of World War II. The bottom one is from the edition of May 24, 1947 and shows the variation of the popularity of the 33rd US president, Harry Truman.

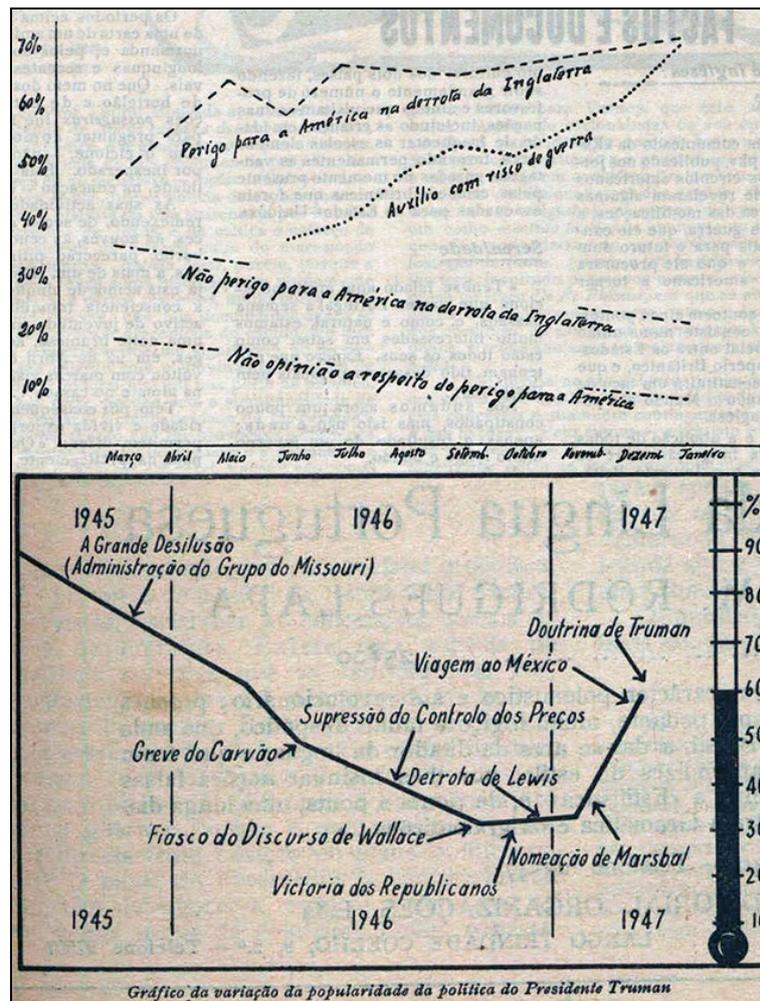


Figure 20 – The highlights of charts in *Seara Nova*'s March 22, 1941 edition (above) and in the edition of May 24, 1947.

Figure 21 shows two examples from the Brazilian magazine *Realidade*.²⁶ The top one is from the November 1967 edition and shows where the revenue comes from and how the budget of the Brazilian federal government will be spent for the year 1968. The bottom one is from the January 1968 edition and displays four maps with the incidence of different endemic diseases in Brazilian states.

²⁶ The original print version is colored, but unfortunately it was not possible to reproduce here the color version of the charts and maps.

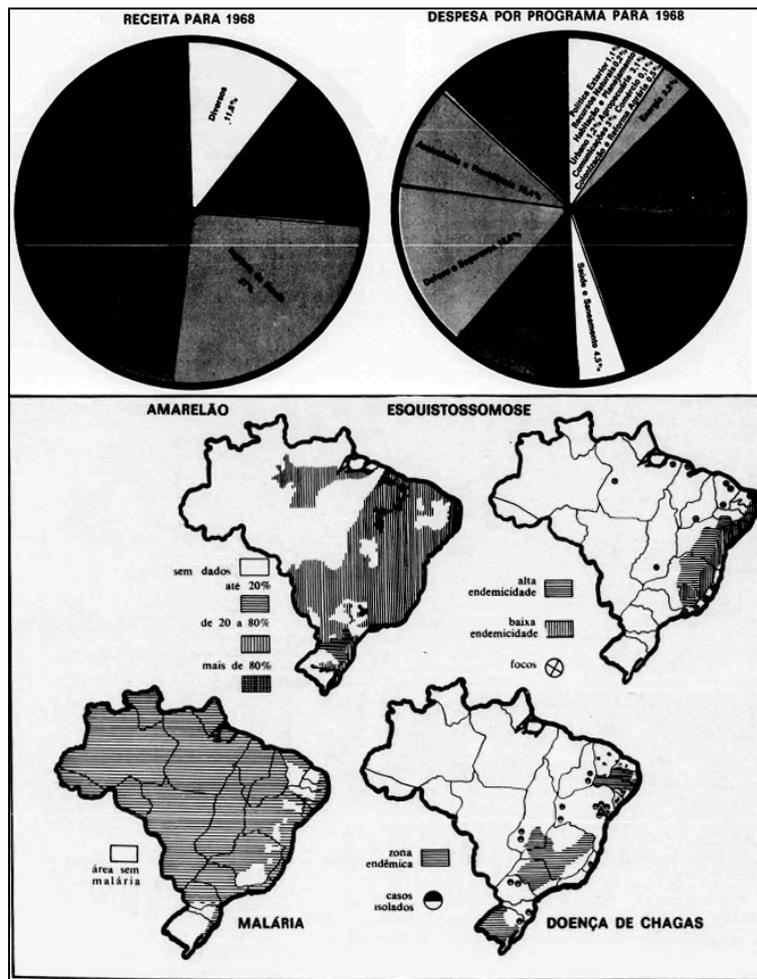


Figure 21 – The highlights of charts and maps in *Realidade*'s November 1967 edition (above) and the edition of January 1968.

Nevertheless, illustration and pictorial charts gain even more prominence with the popularization of a new form of picture language called Isotype.

Starting on 1927, Neurath and Arntz, and later Reidemeister, began developing a “universal visual language” to depict statistics and other information in a pictorial fashion. They called it ISOTYPE, which stands for “International System of Picture Education.” (...) ISOTYPE data charts are usually unidimensional and quite simplistic in comparison to the visualizations designed by statisticians and social scientists of the 19th century, but they are also very clear and elegant. Their impact on news graphics was significant. (Cairo, 2017, p. 49)

From Austria to the Netherlands and then Britain, where the Isotype Institute was established, Otto and Marie Neurath promoted the Isotype in many exhibitions,

commissions and publications where they exerted editorial control, as in the *Future* books and magazine published in London between 1946 and 1952 (Kindel, 2017). By that time Isotype had already crossed the Atlantic and influenced the style of infographics regularly displayed in major publications in the United States, such as *Time*, *National Geographic*, *Newsweek* and *Fortune* (E. K. Meyer, 1997). Isotype also shaped the infographics in Brazil and Portugal, although in these countries it took a few more years to become a trend (S. A. Ribeiro, 2008; Teixeira, 2010).

Isotype has an educational and explanatory purpose, thus it seeks to simplify complex information rather than being a means of instilling facts or showing exact data (Kinross, 2017). As Cairo stated above, the Isotype had a tremendous effect on the visual representations of the newspapers that – just as Sociology – started to use fewer and fewer data charts and statistical visualizations to produce more and more explanatory illustrations. Still, data visualization would have a new chapter with the emergence of computers.

1.4. Computer-assisted reporting and the Precision Journalism

Launched in 1951, Universal Automatic Computer (UNIVAC) was the first commercial computer. The following year, one of the UNIVAC models was first used by a communication vehicle. The American television channel CBS used the machine to predict the outcome of the 1952 presidential election. Election polls pointed to a tight victory of Democrat candidate Adlai Stevenson over Republican Dwight D. Eisenhower, but UNIVAC calculated that the Republican candidate would be the winner for a good advantage. The directors of CBS were afraid to divulge the result obtained by the computer and, after the result of the election, had to confess that UNIVAC correctly predicted the winner (Alfred, 2010; Chinoy, 2010). This was the first Computer-Assisted Reporting (CAR) experience, and since then all US news channels have started to use computers during Election Night reporting (Cox 2000).

Fifteen years later, in 1967, one of the most violent and destructive riots in American history took place in Detroit. Officially, 43 people died, around 1,200 were injured and over 7,200 arrested, and more than 2,000 buildings were destroyed over the course of five days (Hammer, Coleman, & Albom, 2014). The dominant theory, present both in the newspapers and in common sense, was that the riots were carried

out by the less educated people who had no other way of expressing themselves. Another theory in vogue was that the riots were a consequence of the difficulty that southern African Americans had in being assimilated into northern culture. The journalist Philip Meyer, from the *Detroit Free Press*, proved the fallacy of both theories (P. Meyer, 2002).



Figure 22 – Reproduction of page 4b of the *Detroit Free Press* of August 20, 1967.

With the help of a psychology professor at the University of Michigan's Institute for Social Research (ISR), Nathan Caplan, Meyer deployed a quantitative survey research, recruited about 30 interviewers for the field work and used an IBM 360 mainframe computer to compile and analyze the data (Nguyen & Lugo-Ocampo, 2016; Rosegrant, 2011). A month after the riots, the *Detroit Free Press* published the results

in a series of reportages. Figure 22 shows one of them, displaying Meyer's findings in text, charts and maps: there were people from different social classes in the disturbance and they were mostly from the North. The series of reportages won a Pulitzer Prize for its coverage of the riot and years later Meyer presented the concept of Precision Journalism, in which he theorized Computer-Assisted Reporting.

The new precision journalism is scientific journalism. (In France, the term "precision journalism" has been translated as "le journalisme scientifique.") It means treating journalism as if it were a science, adopting scientific method, scientific objectivity, and scientific ideals to the entire process of mass communication. If that sounds absurdly pretentious, remember that science itself is restrained about its achievements and its possibilities and has its own sanctions against pretension. (P. Meyer, 2002, p. 5)

Almost half a century after social science moved away from journalism, Meyer attempted to rebuild this bridge. As Anderson puts it: "Meyer sought to increase journalistic claims to factual certainty, and reduce the uncertainty of social life, by mobilizing the techniques of social science on journalism's behalf" (2018, p. 92). By adopting procedures of scientific rigor from the social sciences, such as statistical methods, sample surveys, etc., Precision Journalism ends up opposing New Journalism.²⁷ As the scholars Dennis and Rivers explain, "the new journalists are subjective to a degree that disturbs conventional journalists and horrifies precision journalists. In essence, all the other new journalists push reporting toward art. Precision journalists push it toward science" (2017, p. 11).

There are points of contact between the groundbreaking work of Charles Booth and Philip Meyer. Booth conducted surveys, created a database, and categorized and mapped poverty in late Victorian London because at the time the newspapers did not address the subject objectively and Booth wanted to know the facts. Meyer deployed a quantitative survey research, filtered and organized the data (on a computer) to find out what had really happened during the Detroit riots and, on the way, dismantled

²⁷ Arisen in the early 1960s, New Journalism, also called "the nonfiction novel" and "the literature of fact", had promoted a reevaluation of journalist practice. As stated before, Truman Capote and Tom Wolfe are some of the well know figures of the movement (Hollowell, 2017).

fallacious theories with facts, charts and maps. Both, social scientist and journalist, gathered information, analyzed what they collected and disseminated information using graphs and maps to facilitate the transmission of the findings.

If, in the 1880s, Stead used fictional elements to make “The Maiden Tribute of Modern Babylon” series more appealing, the 1960s and 1970s New Journalism doubled down with more subjectivity, first-person point of view and proximity between subject and object (Hartsock, 2000). In this way, Meyer saw New Journalism as an antagonist of Precision Journalism.

For decades, as a precision journalist I considered narrative journalists my natural enemies. It didn’t help that the early practitioners sometimes got caught making things up. For example, Gail Sheehy wrote an article for New York magazine in 1973 that described in great detail the sexual and financial escapades of a prostitute in New York who was called “Redpants.” Then the Wall Street Journal revealed that there was no “Redpants.” Sheehy had used a composite of several different prostitutes to provide the dramatic compression needed to give her story the pace and depth of fiction.²⁸ (P. Meyer, 2012)

During the 1960s and 1970s computers were mostly used for organization purposes such as payroll, inventory and management of circulation data and subscriber lists (Garrison, 2001). Thus, the use of computers to search and gather information was scarce in the 1970s – although there are examples of CAR in this decade (M. Cox, 2000) – and only in the late 1980s, with the popularization of personal computers in newsrooms, CAR stories become more common and relevant. According to the journalist Rose Ciotta (1996), between 1989 and 1995, at least one journalist that used computer techniques was among Pulitzer winners with stories about racism in mortgage loans, medical malpractice, government waste, arson fraud and lax building codes.

²⁸ In 2016, Gay Talese disowned his own new book – *The Voyeur’s Motel* – saying that its credibility is “down the toilet” after an investigation by *The Washington Post* (Flood, 2016).

In 1989, profession recognized the value of computer-assisted reporting when it gave a Pulitzer Prize to *The Atlanta Journal-Constitution* in Georgia for its stories on racial disparities in home loan practices. During the same year, [Elliot] Jaspin established the institute at the Missouri School of Journalism now known as the National Institute for Computer-Assisted Reporting (NICAR), and in 1990 Indiana University professor James Brown held the first computer-assisted reporting conference in Indianapolis. Since that time, use of computer-assisted reporting blossomed, primarily due to the seminars conducted throughout the world by NICAR and Investigative Reporters and Editors (IRE) as part of a joint program. The Global Investigative Journalism Network also has worked with IRE on training worldwide. (Houston, 2014, p. 9)

During the 1990s NICAR trained more than 12,000 journalists at some 300 conferences and seminars (Houston, 1999). Some of these conferences and seminars took place in Brazilian newsrooms and a few years later the *Associação Brasileira de Jornalismo Investigativo* (Abraji, the Brazilian Association of Investigative Journalism) began to conduct its own CAR courses, having already trained more than four thousand journalists (Träsel, 2014b). According to scholar Marcelo Träsel (2014b), one of the first examples of the use of computers on newsgathering in Brazil occurs in *Jornal do Brasil* in 1991, with the journalist Mário Rosa using database analysis to find out a corruption case involving the First Lady Rosane Collor, wife of the 32nd President of Brazil, Fernando Collor de Melo.

Rosa had access to a database that stored financial data of the Federal Government of Brazil and, analyzing the accounts of a public assistance body that was chaired by Rosane Collor, he found out that public funds were being diverted. On August 27, 1991, *Jornal do Brasil* began to publish a series of stories proving that the money destined for the poor of the hinterland of the state of Alagoas was being diverted to the first lady's family. Rosane Collor was convicted of the crimes of corruption and embezzlement and the story earned Rosa the Esso Award. Figure 23 shows one of Rosa's pieces on the corruption case with a table showing atypical

Terceira Vaga

129 (11.85)
129 (11.85)

O Comércio do Porto

A PRIMEIRA REDACÇÃO COMPUTORIZADA EM PORTUGAL

A DIGICOMP COM O STRIDE NA INFORMATIZAÇÃO DOS JORNAIS

«O Comércio do Porto» acaba de dar mais um importante passo em frente em matéria de renovação tecnológica, investindo — sem quaisquer apoios externos à Empresa — na informatização da sua Redacção. Este jornal centenário colocou-se assim, e mais uma vez, na vanguarda, como a primeira Redacção computizada em território nacional.

STRIDE A TECNOLOGIA ADAPTADA AO ACESSO DO PRESIDENTE

A Digicomp é uma empresa desportiva que nasceu em 1984, com o propósito de desenvolver soluções tecnológicas para o mundo empresarial. A sua especialidade é a implementação de sistemas de informação que permitem a gestão eficiente dos recursos humanos e materiais. A Digicomp tem vindo a crescer rapidamente, graças ao sucesso das suas soluções e ao apoio dos seus clientes. A empresa está atualmente a trabalhar na implementação de um novo sistema de informação que permitirá a gestão integrada de todos os processos da empresa.

STRIDE 400

O STRIDE 400 é um sistema de informação desenvolvido pela Digicomp para a gestão de recursos humanos. Este sistema permite a gestão de todos os processos relacionados com o pessoal, desde a contratação até à despedida. O sistema é muito fácil de utilizar e permite a gestão de grandes volumes de dados. O STRIDE 400 é um sistema muito versátil e pode ser adaptado às necessidades de qualquer empresa.

SOFTWARE DE GESTÃO DE RECURSOS

O software de gestão de recursos humanos desenvolvido pela Digicomp é um sistema muito completo e eficiente. Este sistema permite a gestão de todos os processos relacionados com o pessoal, desde a contratação até à despedida. O sistema é muito fácil de utilizar e permite a gestão de grandes volumes de dados. O software de gestão de recursos humanos é um sistema muito versátil e pode ser adaptado às necessidades de qualquer empresa.

CONTABILIDADE

O sistema de contabilidade desenvolvido pela Digicomp é um sistema muito completo e eficiente. Este sistema permite a gestão de todos os processos relacionados com a contabilidade, desde a abertura de contas até à elaboração de balanços. O sistema é muito fácil de utilizar e permite a gestão de grandes volumes de dados. O sistema de contabilidade é um sistema muito versátil e pode ser adaptado às necessidades de qualquer empresa.

GESTÃO PESSOAL

O sistema de gestão pessoal desenvolvido pela Digicomp é um sistema muito completo e eficiente. Este sistema permite a gestão de todos os processos relacionados com o pessoal, desde a contratação até à despedida. O sistema é muito fácil de utilizar e permite a gestão de grandes volumes de dados. O sistema de gestão pessoal é um sistema muito versátil e pode ser adaptado às necessidades de qualquer empresa.





Figure 24 – Reproduction of the article on the computerization of *O Comércio do Porto*, published on 8 November 29, 1985.

In Portugal, the introduction of computers in newsrooms also began in the 1980s (Bastos, 2010b; J. L. Fernandes & Cascais, 2006), and *O Comércio do Porto* was the first newspaper in Portugal to computerize the entire editorial staff. In the following decade, all the newsrooms of the major Portuguese newspapers were computerized (Bastos, 2010b).

However, as in other European countries, such as England, examples of the use of computers on newsgathering in Portugal appear only in the 2000s. In Europe, the Center for Investigative Journalism (CIJ), linked with NICAR and IRE, conducted CAR training in several countries during the 2000s (Hewett, 2017). In 2010, Steve Doig, specializing in computer-assisted reporting and Cronkite School’s founding Knight Chair in Journalism, taught courses and lectures on CAR in several Portuguese cities (Doig & Cascais, 2011). However, there is an absence in Portugal of networks such as NICAR in the United States and Abraji in Brazil. During the literature review or the reading of

conference papers and proceedings focused on the topic, very little production on CAR in Portugal was found. There are exceptions (Fidalgo, 2004, 2007), but theoretical and without citing examples from the Portuguese press. Perhaps this helps to explain the fact that one of the first stories that we found that used computers to analyze a database and had great repercussion occurred only in January 2006.



Figure 25 – The highlight of the *24 Horas*’ front page, published on January 13, 2006.

The *24 Horas* had access to files with thousands of phone records related to the Casa Pia affair.²⁹ After cleaning and analyzing the data, a step that lasted months, the journalists were able to identify several prominent figures that had their conversations illegally intercepted by the Public Ministry, including the 18th President of Portugal, Jorge Sampaio (Alexandre, 2014). The series of stories that exposed the illegal wiretapping began to be published on January 13, 2006 and became known as the

²⁹ Scandal involving cases of sexual abuse of children who had been under the care of a public institution for minors in risk of social exclusion, the “Casa Pia de Lisboa” (Barroso, Leite, Manita, & Nobre, 2011).

“Envelope 9” scandal (Figure 25).³⁰ The following year, there was a reform in the Portuguese Code of Criminal Procedure, which altered, among other things, the system of wiretapping and the secrecy of justice. Some authors question whether the new Code was created because of the Casa Pia affair and the “Envelope 9” scandal (Andrade, 2009).

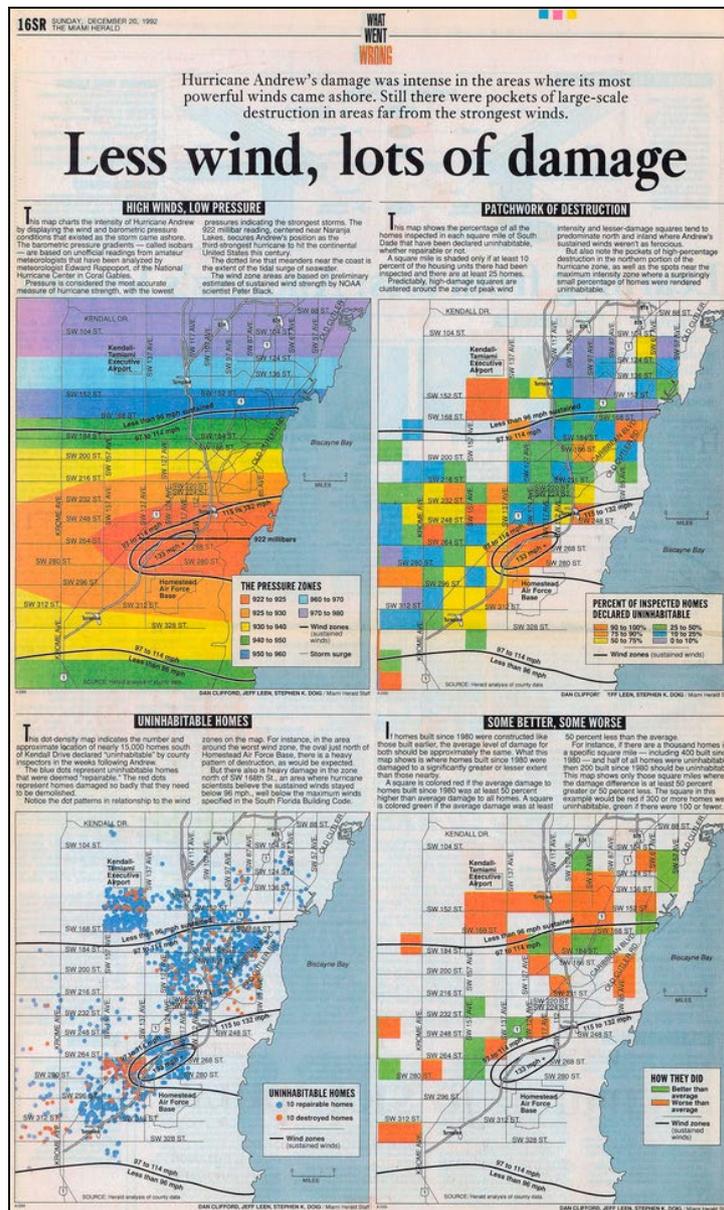


Figure 26 – Reproduction of the *Miami Herald's* page 16SR, published on December 20, 1992.

³⁰ “Envelope 9” was an annex to the Casa Pia affair which contained a list of private telephone numbers of people unrelated to the investigation (Alexandre, 2014).

CAR techniques and the *Precision Journalism* have revolutionized the way news is produced, but the communication of information has not undergone major changes, maintaining a traditional structure, although there are exceptions with CAR stories using data charts and maps, such as Meyer’s previously mentioned pioneering series of reportages on the 1967 riots in Detroit (shown in Figure 22); the Pulitzer-winning articles from Bill Dedman “The Color of Money”, an investigation from *The Atlanta Journal-Constitution* in 1988 that revealed the racial discrimination by credit institutions in Atlanta;³¹ and “What Went Wrong”, in which *The Miami Herald* proved in 1992 that there was a lack of inspection in Florida’s housing, where the newer the house, the easier they broke down (Figure 26).

As we can see, *Precision Journalism* and the CAR techniques also had a great impact on investigative journalism, reinforcing the foundations to fact check data and boost journalism’s commitment to fact and increasing the accountability function of journalism (watchdog journalism).

CAR gives journalists the opportunity to dig for truth in data, and the comparative analysis that a computer can do often reveals pertinent questions. What reporters are able to learn from using CAR provides readers with knowledge and insights that can cut through the clutter of opinionated noise and celebrity obsession. It also can allow even relatively small news operations to delve into problems affecting the global community, yet speak to readers and viewers right around the block. (Method, 2008, p. 14)

At the beginning of the 21st century, computer-assisted reporting skills, such as how to access government records, navigate database managers and find, retrieve, organize and communicate information were no longer a novelty in the United States, being taught at many journalism programs at universities and used on a daily basis in newsrooms (Davenport, Fico, & DeFleur, 2002; Houston, 2014). The trivialization of the use of computers and social sciences methods to tell stories from data or with data made Meyer declare the term CAR obsolete.

³¹ The author has put together all the content here: <http://powerreporting.com/color/>. Accessed on February 25, 2019.

After three decades, the time has come to move away from the CAR frame and toward a social science frame. I don't know what label to substitute for CAR. "Precision journalism" sounds dated. We need something that captures the notion in a fresh way. (...) Whatever the new term, let's give up on CAR. Computers are used in so many different ways, even in the newsroom, that it no longer defines us, if it ever did. The time has come to declare CAR victorious and move on to a fresher, more ambitious concept. We need it, and so does the world. (P. Meyer, 1999, p. 5)

It was not long before new terms to define quantitative journalism began to emerge: database journalism (Suzana Barbosa, 2007; Fidalgo, 2004, 2007; Garrison, 1995; Loosen, 2002), computational journalism (Anderson, 2011b; S. Cohen et al., 2011; Diakopoulos, 2015; Flew, Spurgeon, Daniel, & Swift, 2012; J. T. Hamilton & Turner, 2009), data-driven journalism or just data journalism (Aitamurto et al., 2011; Gray et al., 2012; Lorenz, 2010; Parasie & Dagiral, 2013).

There are a lot of overlaps in practice among these news terms. Leaving database journalism aside, Mark Coddington (2015a) offers a typology for evaluating CAR, computational journalism, and data journalism and highlights the key differences between these forms of quantitative journalism: "CAR is rooted in social science methods and the deliberate style and public-affairs orientation of investigative journalism, data journalism is characterized by its participatory openness and cross-field hybridity, and computational journalism is focused on the application of the processes of abstraction and automation to information" (2015a, p. 337).

In order to create these interactive and participatory data-driven pieces, special attention was required for the presentation of the stories. As José Roberto de Toledo (2014) puts it: "the data journalism product is not a headline or a story. It is a tool. (...) Maps, infographics, charts and timelines help users to create a narrative for the facts – not the narrative of the journalist, but his/her own. (...) The user's point of view becomes necessarily part of the narrative"³² (Toledo, 2014, p. 8). Thus, if part of the

³² In the original: "O produto do jornalismo de dados não é uma manchete ou uma reportagem. É uma ferramenta. (...) Mapas, infográficos, tabelas e linhas do tempo ajudam o internauta a criar uma

20th century is called “the modern dark ages” of data visualization (Friendly, 2008; Friendly & Denis, 2000), with the rise of data journalism (and computational journalism) statistical graphs and thematic maps began to have an unparalleled protagonism in journalism.

1.5. Data journalism and data visualization

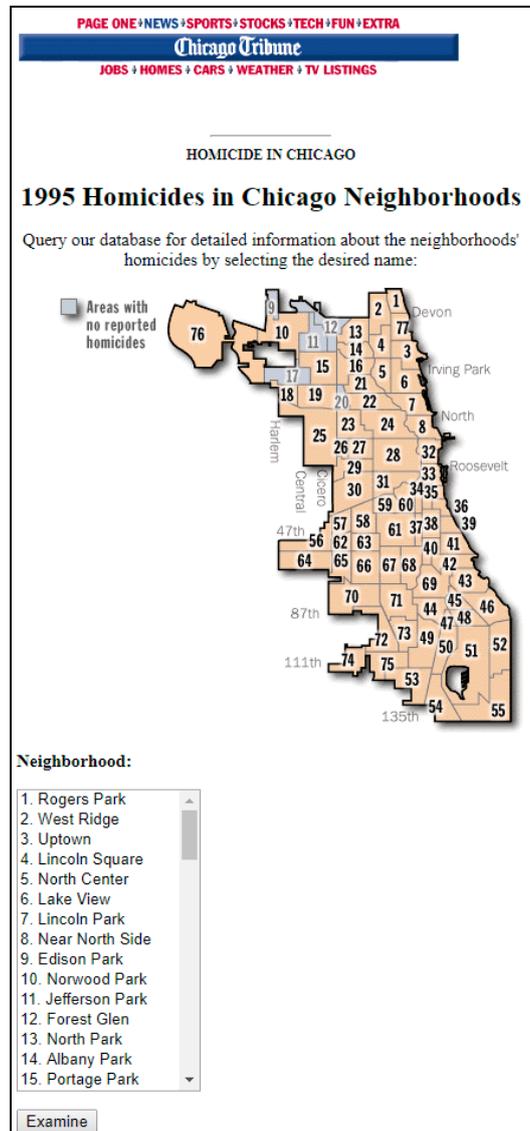


Figure 27 – Reproduction of the *Chicago Tribune's* webpage “1995 Chicago Homicides”.

The earliest examples of interactive news using database and graphics arise even in the 1990s. Scholar Nikki Usher (2016) points to the *Chicago Tribune's* “Homicide in Chicago” as one of the pioneers. An interactive map shows all homicides

narrativa para os fatos — não a narrativa do jornalista, mas a sua própria. (...) O ponto de vista do usuário passa a ser necessariamente parte da narrativa.”

in the city from 1965 to 1995; another section allows the user to search an interactive database for detailed information about the neighborhoods' homicides, as shown in Figure 27. The project continues in a different form through today.

Almost a decade later, the web developer Adrian Holovaty launched a project called "Chicago Crime Map" (chicagocrime.org),³³ shown in Figure 28, a mashup that combined data available from the Chicago Police Department with Google Maps: "it offered a page and RSS feed for every city block in Chicago and a multitude of ways to browse crime data — by type, by location type (e.g., sidewalk or apartment), by ZIP code, by street/address, by date, and even by an arbitrary route" (Holovaty, 2008). The project won a prize for innovation in journalism (Holovaty, 2008) and was widely debated in the media.

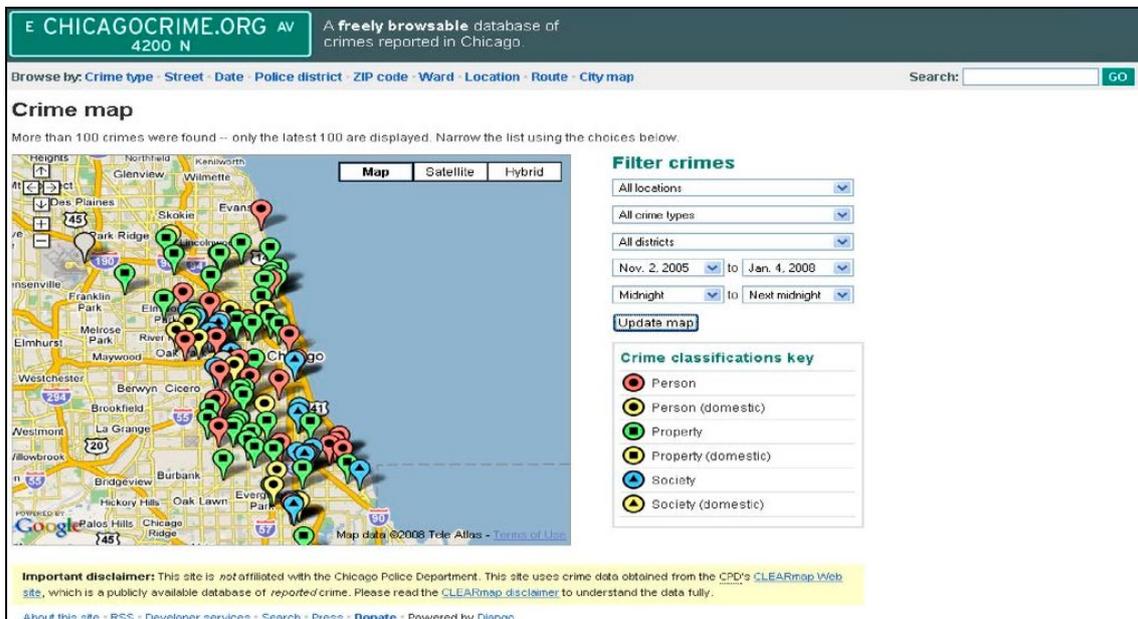


Figure 28 – Reproduction of Adrian Holovaty’s website “Chicago Crime Map”.

If the map of the *Chicago Tribune* brought the number of homicides, an interactivity restricted to accessing more information about a particular neighborhood and had to be manually updated by the newspaper’s staff, “Chicago Crime Map” provided details on all crimes occurring on every street in the city, allowing the user to receive alerts of new crimes per street, and was updated automatically.

³³ The project has been terminated and the URL redirects to another page.

“Chicago Crime Map” is a good example of computational journalism since it considers its audience as a collection of rational, participatory users who are autonomous and creative people capable to do searches and analysis on their own (Coddington, 2015a; Gynnild, 2014). And Holovaty began to call attention to journalism as programming.

He was the first high-profile programmer journalist, the first person who really got programming and journalism and treated them both as equally valuable. (...) The coverage Holovaty received and his constant presence on the news-innovation conference circuit helped spread the message that programming and journalism were complimentary, necessary, and critical to the editorial project. (Usher, 2016, p. 57)

In 2006, Holovaty published “The fundamental way newspapers need to change”, which can be considered a manifesto about the use of database in journalism. If Meyer's Precision Journalism marked a turning point in how journalism uses data to tell stories, Holovaty's post marks a new phase in quantitative journalism.

So much of what local journalists collect day-to-day is *structured information*: the type of information that can be sliced-and-diced, in an automated fashion, by computers. Yet the information gets distilled into a big blob of text — a newspaper story — that has no chance of being repurposed. (...) For example, say a newspaper has written a story about a local fire. Being able to read that story on a cell phone is fine and dandy. Hooray, technology! But what I really want to be able to do is explore the raw facts of that story, one by one, with layers of attribution, and an infrastructure for comparing the details of the fire — date, time, place, victims, fire station number, distance from fire department, names and years experience of firemen on the scene, time it took for firemen to arrive — with the details of previous fires. And subsequent fires, whenever they happen. (Holovaty, 2006)

Database was the buzzword at that moment. Building on work by theorist Lev Manovich³⁴ (2002), Brazilian researchers also discussed the potential of databases for journalism (Suzana Barbosa, 2007; Larrondo, Mielnickzuk, & Barbosa, 2008; Machado, 2005, 2007; Palacios, 2003). In 2007, Elias Machado argued that with the growing use of databases, interactive content would become more common in journalism.

The interactive character of the narrative in cyberspace transforms the displacements through the navigable space as a central instrument of observation, exploration, narration and, ultimately, the composition of the narrative itself. As it remains tied to the narrative forms of conventional media, the journalistic narrative in cyberspace hardly employs navigable space formatted on databases as the standard interface. The progressive use of databases as a format by journalistic organizations and navigable space as a support for interactive narratives goes through the recognition that, although they were existing forms before computers, both databases and navigable space, by assuming functions distinct from those performed so far, have enabled the narrative in cyberspace to be understood as an environment for creative interactive actions. (Machado, 2007, p. 114)³⁵

To produce content such as “Chicago Crime Map”, it is essential to bring together a wide range of computer techniques that the vast majority of journalists do not have. To fill this gap, news organizations began hiring hacker/programmers such as Adrian Holovaty (Daniel & Flew, 2010; Royal, 2010). The software developer and author Eric Raymond (2004) defines hackers as non-authoritarian programmers who believe in sharing information as a virtue and therefore contribute to or work with free

³⁴ On *The Language of new media* (Manovich, 2002), Manovich argues that database is to the digital age what the narrative, in literature and cinema, was to the modern era.

³⁵ In the original: “O caráter interativo da narrativa no ciberespaço transforma os deslocamentos pelo espaço navegável como um instrumento central da observação, exploração, narração e, em última instância, da composição da narrativa propriamente dita. Como permanece atrelada às formas narrativas dos meios convencionais, a narrativa jornalística no ciberespaço pouco emprega o espaço navegável formatado sobre bancos de dados como interface padrão. A progressiva utilização dos bancos de dados como formato pelas organizações jornalísticas e do espaço navegável como suporte para narrativas interativas passa pelo reconhecimento de que, embora se tratando de formas existentes antes dos computadores, tanto bancos de dados quanto o espaço navegável, ao assumirem funções distintas das desempenhadas até aqui, têm possibilitado que a narrativa no ciberespaço seja compreendida como um ambiente para criativas ações interativas.”

and open-source software. Himanen (2001) identifies seven values that characterize a hacker: passion, freedom, social worth, openness, activity, caring, and creativity. For Himanen, the latter value is the most important and fundamental: “the imaginative use of one's own abilities, the surprising continuous surpassing of oneself, and the giving to the world of a genuinely valuable new contribution” (2001, p. 141).

These “programmer-journalists” or “hacker journalists” (Boyer, 2008; Royal, 2010) bring their values and world views to the newsrooms (S. C. Lewis & Usher, 2013; Parasie & Dagiral, 2013; Porlezza, 2019; Russell, 2016; Träsel, 2014a). More than telling a static traditional story showing few data points (like the CAR stories), these professionals are interested in making all data points available in an interactive way, increasing the transparency and participation of the readers/users.

Because of their involvement in open software communities and open government advocacy, programmer-journalists view public access to information – whether it is code or public information – as a major issue. This is why they all put the emphasis on giving the audience the largest and easiest access to the data. They claim that readers should be given the opportunity not only to check the data but also to combine them and to use them for other goals. (...) Such news projects offer readers a convenient “decision-making tool” or “research tool” in their daily life through simple and standardized access to data. (Parasie & Dagiral, 2013, pp. 863–864)

If with CAR techniques journalism began moving toward a more scientific stance, with data journalism the craft moved toward a more hacker stance – open and collaborative. And if the 20th century was marked by explanatory illustrations, the 21st century begins by prioritizing data visualization.

As previously stated, in the 1980s computer graphics boosted both quantity and quality of the visual representations in newspapers, although production remained focused on pictorial representations and explanatory illustrations, rather than on statistical graphs and data maps (Cairo, 2017; Henriquez, 2014). This means that often the form was more important than the content and the information could be impaired. Emily Arnette Vines analyzed the integrity, accuracy and credibility of the information

present in some editions of *USA Today Snapshots* in 2001 and found that many of these graphs “deviated from journalism values, misled readers with inaccurate data, sources that were not credible, and images that did not wholly reflect the data” (2002, p. 45).

The commitment of programmer-journalists to the transparency, consistency and reliability of the published data seems to make them more interested in creating interactive tools and data-driven pieces than illustration-driven visual explanations. More than a hundred years after the pioneers of data visualization in the 18th century and the survey movement of the 1900s offered new techniques for gathering data and visualizing information, journalism turns its attention to this field.

By the 2010s, the major newsrooms in the United States, United Kingdom and Germany had teams dedicated to the production of data journalism pieces, as we saw in WikiLeaks coverage. The information designer Juan Velasco argues that *The New York Times* played a prominent role in popularizing data visualization.

Juan Velasco, the Art Director at *National Geographic* who worked at *The New York Times* from 1996 to 2001, believes that newspapers moved to data visualizations and away from explanatory graphics because of the influence of *The New York Times*, which slowly reduced the presence of large explanatory graphics, especially illustrated ones, during the 2000s. Data-driven and unadorned charts replaced them, and over time, this trend was widely imitated as illustrated graphics came to be seen as less “serious”. Due to that trend, many newspapers stopped hiring people who could draw well, or who can design complex infographics with multiple elements. Developers are preferred to personnel with artistic and design skills. Velasco believes that graphics are resolved as “data visualizations” even at times when an illustrated graphic would be better, due to lack of vision and skill. (Henriquez, 2014, pp. 10–11)

Usually data journalism teams have journalists, developers, and designers (Gray et al., 2012). In Brazil, the first team of data journalism in a newsroom had four people: three journalists and one programmer. The *Estadão Dados* was created in May 2012

and is affiliated to the newspaper *O Estado de S. Paulo* (Träsel, 2014a). One of the team’s first work was the “Basômetro”,³⁶ shown in Figure 29, an interactive tool that allows the reader to monitor how a senator or a federal deputy voted and the level of support a party gives to the federal government in each voting bill since 2003.



Figure 29 – Reproduction of the *Estadão’s* webpage “Basômetro”.

In Portugal there are still no dedicated data journalism teams in the newsrooms of the major newspapers (Abras, 2018; Moura, 2018). In any case, *Público* was one of the first newspapers to have a data journalist on its team in 2012.³⁷ Usher defines data journalist as follows:

they are primarily working with data in the service of stories, actively trying to tell stories with data, and spend most of their time working specifically with data. Certainly, some of these data journalists may be programmer journalists or hacker journalists, who may also at times work with data, but data journalists employ data as their primary focus in their work – and many of these journalists do not see themselves primarily as working with code. In fact, some data

³⁶ Available at <http://estadaodados.com/basometro/>. Accessed on February 25, 2019.

³⁷ They were research fellows of the project *REACTION* (Retrieval, Extraction and Aggregation Computing Technology for Integrating and Organizing News), under the UT Austin-Portugal program, supported by Fundação de Ciência e Tecnologia (Portuguese Foundation for Science and Technology), and which aimed to develop computer journalism in Portugal.

journalists may not do any coding at all. (Usher 2016: p. 90-91)

“Doze Anos de Incêndios”,³⁸ shown in Figure 30, is from that period. Part of a special on forest fires in Portugal, this interactive piece allows the reader to explore the most affected regions by year and to examine the time of alert from more than 61,000 fires.



Figure 30 – Reproduction of the *Público*'s webpage “Doze Anos de Incêndios”.

There are examples of data journalism in Brazil before the work of *Estadão Dados* (Susana Barbosa, 2012; Canavilhas, Satuf, Luna, & Torres, 2014), but this team from *O Estado de S. Paulo* was the first dedicated to producing data visualization in a systematic way. Similarly, there are examples of interactive pieces before “Doze Anos de Incêndios” (Alexandre, 2016), but, although without a team dedicated to producing data journalism, *Público* was for some years the only Portuguese newspaper to produce this content with some regularity (Alexandre, 2014). In Brazil, several news organizations currently have data journalism teams and in Portugal there are now

³⁸ Available at <https://acervo.publico.pt/floresta-em-perigo/doze-anos-de-incendios>. Accessed on February 25, 2019.

some media outlets with data journalists in their newsrooms (Abrás, 2018; Moura, 2018).

This delay may be compromising Portuguese newsrooms. Data journalism may not only increase transparency and engagement of the readers, as stated before, but also boost audience and revenue generated. Out of the 20 most visited content pieces of *The New York Times* in 2014, 8 were interactive (Usher, 2016). In 2018, out of the 10 most read stories, 8 were interactive.³⁹ The dialect quiz “How Y'all, Youse and You Guys Talk”⁴⁰ was published by *The New York Times* at the end of December 2013 and in just 11 days got more views than any other article from the newspaper that year, quickly becoming the most viewed page in the paper’s history (Katz, 2016). *Time* magazine launched in 2015 a website called “TIME Labs”,⁴¹ a page dedicated to interactive data journalism. Here is the explanation for the creation of the site in financial terms:

Advertisers are gaining interest in how much time readers are spending on a site versus just clicks, and Time Labs also is a way for the publisher to capitalize on that interest. Buyers said more time spent can lead to a greater advertiser benefit, which in turn could help a site command higher ad rates, or at least more advertising. (Moses, 2015, p. n.p.)

In the next chapter, we discuss how the audience, once neglected or even completely ignored by journalists, has become a central figure in editorial decisions. Likewise, and complementarily, we examine how hits, unique pageviews, average time on page and other metrics are tracked by the minute and have become an obsession in some newsrooms around the world.

³⁹ Available at: <https://www.nytimes.com/interactive/2018/12/19/reader-center/top-stories.html>. Accessed on February 25, 2019.

⁴⁰ Available at: <https://www.nytimes.com/interactive/2014/upshot/dialect-quiz-map.html>. Accessed on February 25, 2019.

⁴¹ Available at: <http://labs.time.com/>. Accessed on February 25, 2019.

II – Web Analytics and the quantitative audience

The earliest computing technology was conceptualized by the British mathematicians Charles Babbage and Ada Lovelace during the mid-nineteenth century (Gleick, 2011; Hammerman & Russell, 2015; Swade, 2002). The first programmable computers began to emerge a century later in the 1940s in Germany (Z3), the United Kingdom (Colossus) and the United States (ABC and ENIAC) (O'Regan, 2016a; Rojas, 2002; Smiley, 2010). During the 1960s, two graduate students from the Massachusetts Institute of Technology (MIT) developed and presented to the U.S. Defense Advanced Research Projects Agency (then ARPA, now DARPA) the idea that it was possible to transfer tiny information packets from one computer to another. Initially, this interconnectivity – the Internet – was called ARPANET (Abbate, 2000; Leiner et al., 2009). In 1989, Tim Berners-Lee, a researcher working at the European Organization for Nuclear Research (CERN) in Switzerland, created the World Wide Web. Among the features created by Berners-Lee and present in the World Wide Web are the Universal Resource Locator (URL), the Hypertext Transfer Protocol (HTTP), the Hypertext Markup Language (HTML), and the Internet Browser (Abbate, 2000; O'Regan, 2016b). The World Wide Web is the modern Internet.

With the overwhelming popularity of the World Wide Web, home users and especially small to mid-sized businesses began using *counters* on their webpages. Page counters allowed any visitor to see how many people had accessed that page, since every time the page was loaded the number on the webpage counter increased by one. Around 1995–96, page counters were very popular (Kaushik, 2007; Ledford, Teixeira, & Tyler, 2010), although the tool tended to inflate the number of visitors because not every new page load meant it was a new visitor. By the year 2000, web analytics was a real business, with companies charging from \$200 to a couple thousand a month. Accrue, Coremetrics, HitBox Professional, WebTrends, and WebSideStory were some key vendors at the time, but at some point, during the dot-com boom, there were around 200 vendors providing solutions that reported massive amounts of data (Kaushik, 2007; Ledford et al., 2010).

On 2008, the Digital Analytics Association (then Web Analytics Association) Standards Committee proposed a definition to web analytics: “Web Analytics is the measurement, collection, analysis and reporting of Internet data for the purposes of understanding and optimizing Web usage”.⁴² At that time Google Analytics had already been released three years earlier, making web analytics trendier than ever, and the web analytics experts were very excited with the tool and the consolidation of the field. Justin Cutroni (2010) said “Let’s face it, Google Analytics is sexy as hell right now! Everyone wants a piece of it”, while Avinash Kaushik (2010) compared web analytics to Angelina Jolie: “that comparison should suggest how sexy it is, how powerful it is, and what a force for good I think it is.”

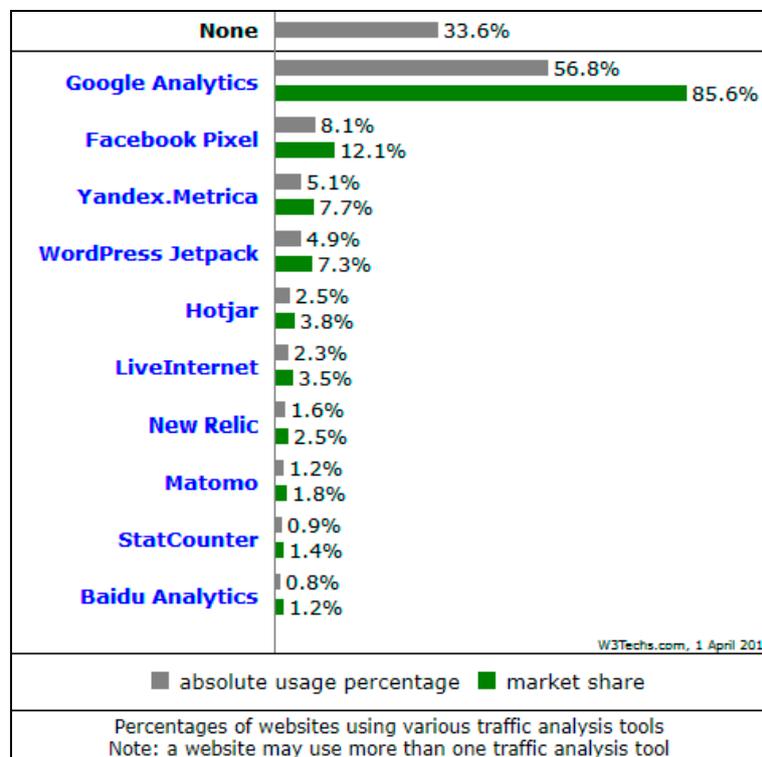


Figure 31 – Reproduction of W3Techs survey showing the top 10 traffic analysis tools for websites.

In April 2005, Google acquired a leading web analytics firm called Urchin Software Corporation (Weller & Calcott, 2012). Seven months later, in November, Google rebranded Urchin’s service and launched Google Analytics, but while Urchin Analytics had a monthly cost of about \$200, Google Analytics (GA) was released as a free application. GA counted the most popular pages, how long a regular user stayed

⁴² Available at https://www.digitalanalyticsassociation.org/Files/PDF_standards/WebAnalyticsDefinitions.pdf. Accessed on February 25, 2019.

on the page, the percentage of users who bounced the web page from a specific page, and more. The response was huge and a quarter of a million new accounts were created in two days (Ledford et al., 2010). GA is updated regularly with new features, like real-time reports, and new versions, like Universal Analytics and a paid version, Google Analytics 360 (Alhlou, Asif, & Fettman, 2016; Clifton, 2015; Mokalis & Davis, 2018). According to a report from World Wide Web Technology Surveys (W3Techs), a company that tracks the usage of different web technologies, by April 2019, GA was “used by 85.6% of all the websites whose traffic analysis tool we know. This is 56.8% of all websites” (Figure 31).⁴³

However, although it is the most widely used software that generates metrics, GA is rarely the only tool used by those who really want to know their audience. Companies and tools such as Comscore, Nielsen/NetRatings, Alexa, Hitwise, Omniture (now part of Adobe Marketing Cloud and called Adobe Analytics), Visual Revenue (now integrated with Outbrain Engage), Webtrends, AT Internet, Webtrekk, Chartbeat, Parse.ly, Linkpulse, EzyInsights, and many others have been adopted by newsrooms around the world (see, for instance, Moyo, Mare, and Matsilele 2019; Blanchett Neheli 2018; Christin 2018; N. S. Cohen 2018; Nguyen and Vu 2018; Vieira 2018; Cherubini and Nielsen 2016; Stockleben and Lugmayr 2016; Petre 2015; Tandoc 2014; Usher 2013; Anderson 2011b; Graves, Kelly, and Gluck 2010; Schaudt and Carpenter 2009; MacGregor 2007).

There are two main reasons for this plethora of web analysis services: (1) each tool has its own methodology and specific features; (2) the results are inconsistent and vary widely from service to service. For example, in August 2014, the *Huffington Post* reached 368 million global unique visitors by the internal numbers, but the webpage made a post celebrating the 115 million global unique visitors according to Comscore (Cherubini and Nielsen 2016).⁴⁴ Rick Hirsch, then multimedia editor of the *Miami Herald*, recalls to Graves, Kelly, and Gluck (2010) another anecdotal case.

⁴³ Available at: https://w3techs.com/technologies/overview/traffic_analysis/all. Accessed on April 1, 2019.

⁴⁴ Available at: https://www.huffpost.com/entry/100-million-thank-you-to-huffposters-around-the-world_b_5822998. Accessed on April 1, 2019.

As an example, Hirsch points to January of 2010, the month of the Haitian earthquake that claimed an estimated 230,000 lives. “We know our traffic went through the roof, because of our history of coverage in the region,” Hirsch says. The paper’s internal figures matched expectations: as recorded by Omniture, traffic spiked 36 percent over December, to 35 million pageviews, while unique visitors jumped 11 percent, to almost 6 million people. Meanwhile, though, comScore recorded less than half as much traffic for January, and fewer than a third as many unique visitors. Unaccountably, comScore had the Herald’s pageviews dropping by 40 percent the month of the earthquake — and falling again in February, despite the fact that Miami hosted the Super Bowl that month. (Graves et al., 2010, p. 19)

This diversity of companies and great discrepancy in results between them is not typical in the history of media measurement in the United States, notable for relying heavily on third-party measurement firms, which led the audience measurement industry in the 20th century to virtual monopolies in each media technology (Beville, 1988; K. S. F. Buzzard, 2002; Gluck & Roca Sales, 2008; Rossman, 2015).

2.1. Brief history of traditional media measurement

In order to establish media transparency, accountability, and confidence between the newspaper and advertising companies, the nonprofit organization Audit Bureau of Circulations was founded in 1914 as a collaboration of publishers, advertisers, and advertising agencies. Since then, the now called Alliance for Audited Media (AAM) is the *standard* that provides audited circulation figures for newspapers and magazines in the United States (Beville, 1988; Napoli & Bjur, 2017). Census-based approach is the main methodology applied by the AAM.

Census-based measures, so-called because they purport to reflect the entire universe of media users rather than just a sample, are possible only where distribution offers some clue about that the size of universe — for instance, in records of the number of copies of a newspaper printed and sold each day. (Graves et al., 2010, p. 8)

Many countries followed the United States with one virtual monopolist player providing print circulation figures. In 1961 the *Instituto Verificador de Circulação* (now *Instituto Verificador de Comunicação*)⁴⁵ was founded in Brazil and in 1986 the *Associação Portuguesa para Controlo de Tiragem* was created in Portugal.⁴⁶ Both nonprofit organizations are members of the International Federation of Audit Bureaux of Certification (IFABC), “a voluntary cooperative federation of industry-sponsored organizations established in nations throughout the world to verify and report facts about the circulations of publications and related data” founded in Sweden in 1963.⁴⁷

In 1949, Jim Seiler founded the American Research Bureau (later Arbitron and now Nielsen Audio), the dominant provider of radio measurement data in the United States for more than 60 years (K. S. F. Buzzard, 2003; Napoli, 2011; Webster, Phalen, & Lichty, 2014). In Brazil, the undisputed provider of radio ratings is the *Instituto Brasileiro de Opinião Pública e Estatística* (IBOPE, now Kantar Ibope Media), founded in 1942 (Carniello, Mota, & Santos, 2017; Eduardo, 1990; Gardim, 2012). In Portugal, the standard that has been providing statistics about radio broadcasting audience for more than 30 years is the *Bareme Rádio*,⁴⁸ from Marktest, launched in 1983 (Santos, 2006).

The singular measurement “currency” for television in the United States has long been Nielsen Media Research: “from the 1950s Nielsen has been the dominant monopoly player in the U.S. market for audience ratings” (Napoli and Bjur 2017, 6; see also Greene and Yao 2016; K. S. F. Buzzard 2015; Wilbur 2015; K. Buzzard 2012). In Brazil, from 1954 to 2015 Kantar Ibope Media was the only company in the market for television audience measurement (Carniello & Santos, 2016; Leite, Moraes, & Machado, 2017). Audience measurement in the television sector in Portugal had successive monopolies. During the 1980s, the public service broadcaster *Rádio e Televisão de Portugal* (RTP) was responsible for the measurement. From 1998 to 2012 Marktest did the job, and since then Growth from Knowledge (GfK) has been in charge of the audience ratings (Abreu, 2017; Romeu, 2015).

⁴⁵ Available at: <https://ivcbrasil.org.br/#/institucional/fundacao>. Accessed on April 1, 2019.

⁴⁶ Available at: http://www.apct.pt/APCT_artigo.php?idCategoria=667. Accessed on April 1, 2019.

⁴⁷ Available at: <http://www.ifabc.org/about-us/mission-and-values>. Accessed on April 1, 2019.

⁴⁸ Available at: <https://www.marktest.com/wap/a/grp/p~4.aspx>. Accessed on April 1, 2019.

The audience measurement technology and methodology used in radio during the 1930s and 1940s migrated to television in the 1950s (Beville, 1988; K. S. Buzzard, 1990; Greene & Yao, 2016). Napoli and Bjur (2017) summarize the approach developed for radio and applied and refined by the television.

Continuity between radio and television measurement is strong; again illustrating the tendency for audience ratings systems to be transferred from established to new media. (...) the first measurement solution was to ask a random sample of individuals about their media consumption. The second major approach that was developed was to construct meters that could monitor when the radio (and later TV) sets were turned on and what they were tuned to, and to set them up in a sample panel of households selected to represent the broader audience. (Napoli & Bjur, 2017, pp. 5–6)

This second major approach, known as the sample-based methodology (also called panel-based), “is both predictive and much less expensive to develop for a mass medium like television” than the census-based measures (Gluck & Roca Sales, 2008, p. 65). However, on the Internet, neither the sample-based nor census-based approach fully meets the needs of the online media outlets.

2.2. Online audience measurements

Unlike traditional media, there is no monopolies to measure the online audience. There are 3 main methods to measure audience ratings online: panel-based (also called user-based), census-based (also called server-based), and through Internet Service Providers (ISP).

Starting with the latter, ISP-based measurement uses anonymous data that is collected by internet service providers. The best example of company that has deals with a lot of the world’s largest ISPs and use ISP-based measurement is Hitwise. The big benefit of using this approach is that there is no sampling, it has data of all users who use an ISP (and it is Hitwise’s partner). Another advantage is that Hitwise combines the data collected with panel data, delivering also demographic information. One of the major shortcomings of this approach is that ISP-based measurement does

not provide data about behavior deep into the website (Hindman, 2018; Jackson, 2016; Kaushik, 2007).

Panel-based audience measurement is very similar to the Nielsen ratings systems: a longitudinal survey in which the same sample of individuals is tracked over some period of time, then their habits are extrapolated to the broader population. The panel members agree to install a software metering on their computers that tracks their online activities by the click. This methodology offers two key advantages. First, it provides accurate information about the socio-demographic characteristics of the users. Second, user-based approach allows better comparisons across competing sites and over time, showing the potential audience of the website. There are three main disadvantages. First, websites' numbers of visitors tend to be underestimated since the software metering is likely installed only on home computers (leaving work environment, schools, universities, etc., out of the sample). Second, the panel tends to be biased, not a real random statistical sample of the population. Third, this methodology tends to favor websites with huge amount of traffic, while websites with small audiences are likely not represented with accurate numbers. Comscore and Nielsen/NetRatings are the two major audience measurement companies that use this offsite approach. (Christin, 2014; Clifton, 2012; Gluck & Roca Sales, 2008; Graves et al., 2010; Kaushik, 2007; Webster et al., 2014).

Finally, census-based methodology uses *cookies*⁴⁹ to track every activity of the user in the website, as Christin explains:

When a reader visits a website for the first time, her computer is generally tracked by “cookies” linked to each website that recognize her each time she comes to the website. Cookies store data that can later be exploited to provide fine-grained information about the website’s visitors (a cookie equals a unique visitor) for any given period. (Christin, 2014, p. 126)

Some of the most popular server-based tools used in newsrooms are Google Analytics, Adobe Analytics and Chartbeat (Cherubini & Nielsen, 2016; Vieira, 2018;

⁴⁹ Brian Clifton (2015) defines cookie as: “a small text file that a web server transmits to a web browser so that it can keep track of the user’s activity on a specific website. The visitor’s browser stores the cookie information on the local hard drive as name – value pairs.”

Wang, 2018). This onsite web analytics approach has one major inconvenience: it offers very limited (and reliable) information about who is visiting the website. On the other hand, this site-centric method has two advantages. First, it provides a level of users' behavioral detail that panels cannot match (e.g. from where the user came to the website, time on page, depth on site, etc.). Second, it allows a more precise measurement of the audience, although the number of visitors may be inflated (Christin, 2014; Clifton, 2012; Graves et al., 2010; Kaushik, 2010; Napoli & Bjur, 2017; Webster et al., 2014).

Discrepancies in levels are, first and foremost, due to the fact that site-centric methods work by putting cookies on people's browsers. If an individual uses several browsers and several devices each device-browser combination will get a cookie and be counted as a unique visitor. The same goes for users that apply cookie deletion; once a cookie is removed the system will read the visitor as a new one. (Napoli & Bjur, 2017, p. 14)

No single method can measure everything on the Internet with accuracy, which is why media outlets tend to combine different methods and tools to complement each other. Audience measurement was an essential tool for the marketing departments of media outlets during the 20th century and remains indispensable. Christin (2014) points out that marketing departments of media companies tend to subscribe to panel-based companies in order to know how many monthly unique visitors they have and their socio-demographic characteristics. The marketing departments also rely on census-based measurement to gather accurate data about the behavior of the users. That way, they can get a glimpse of the website's audience profile, which is very important for advertisers.

Marketing departments use online audience measurements in two main ways. First, they rely on web metrics in order to construct a data-informed portrait of the website's audience that they can then market to the companies who are interested in buying ad space on the website. (...) Second, marketing departments rely on web metrics to set the rate of their different online inventories. They can

charge higher advertising rates for larger numbers of unique visitors; likewise, they can charge more for matching the target audience of the advertisers. At present, news websites' marketing departments typically rely on a mix of panel-based and server-based data. (Christin, 2014, pp. 127–128)

The absence of monopolies in the ratings is not the only novelty regarding the audience that the Internet has generated in the media landscape. Another critical transformation is related to the relationship between journalists and the audience. If during the 20th century journalists mostly ignored the audience wants and desires, now the newsrooms watch – and worry in real time with – the moods of the audience.

2.3. From values-oriented to analytics-driven news

In the early newspapers large spaces were dedicated for opinions and objections from the public and, according to some scholars, this active participation played an important role in the history of freedom of the press.

Because if it hadn't been for the publication of opinions from the public, it is doubtful that freedom of public expression would have achieved its status as an inherent human right. Historically, the public arguments in the late 17th and early 18th centuries calling for an end to government control of the media were essentially letters to the editor. (Reader, 2015, pp. 1–2)

Even during the competitive mid-nineteenth century, American newspapers continued to devote extensive room to readers' writings.

The eighteenth-century newspaper reader was, often enough, also a newspaper writer, and this was not uncommon for the associational press of the nineteenth century. Even in the metropolitan press, letters to the editor continued to occupy large chunks of newsprint. The country edition of the *New York Tribune* in the 1870s regularly devoted a page or more of its sixteen pages to agricultural correspondence, a kind of mass-mediated communal self-help column (...) these were correspondents writing directly for publication for the benefit of other

readers and correspondents. (Schudson, 1996, pp. 50–51)

However, as the press became more professionalized, self-conscious, and profitable business, the distance between the journalists and the readers increased. And by the late 20th century, the segments previously dedicated to the community members shifted into exclusive spaces only available to an elite. “More and more, the voice of the professional newspaper was separated out from the voice of the readers; where once the two were undifferentiated, they became sharply divided” (Schudson, 1996, p. 51).

The press has a critical role in liberal democratic societies (Bennett, Lawrence, & Livingston, 2008; Habermas, 2012; Kovach & Rosenstiel, 2007; Schudson, 2008; Traquina, 2004), since “just as democracy without a free press is unthinkable, journalism without freedom is either farce or tragedy” (Traquina, 2004, p. 23).⁵⁰ Besides being a public good essential for democracies, journalism also plays a crucial role as a system of checks and balances on government and as watchdog and guardian.

Watchdog journalism puts the spotlight on wrongdoing and injustices with the expectation that they will cause public outrage and prompt legislative and judicial action. By doing so, it serves democratic goals of accountability, honesty, truth-telling, and transparency. Watchdog journalism exemplifies one of the best contributions of journalism to democracy for it contributes to monitoring power and revealing information that helps citizens understand how government, corporations, and other powerful actors function. International surveys of professional values show that reporters hold it as one of the highest ideals. (Waisbord, 2016, p. 1)

These public functions combined with the development of core professional norms (e.g. objectivity, ethics, autonomy, etc.) (Deuze, 2005) and the system of “news values” (e.g. impact, surprise, timeliness, tragedy, proximity, etc.) (Galtung & Ruge, 1965; Harcup & O’Neill, 2001, 2017; Traquina, 2005) that determined whether a story was worthy of publication contributed to journalists ignoring – and even disregarding –

⁵⁰ In the original: “Tal como a democracia sem uma imprensa livre é impensável, o jornalismo sem liberdade ou é farsa ou é tragédia.”

the public's opinion. In the mid-twentieth century, the elitist attitude that they know what is best for the audience is reinforced among journalists (Schudson, 2012). In his article on the experience of writing news in the 1960s, Robert Darnton recalls: "We really wrote for one another"(Darnton, 1975, p. 176). "I know we have twenty million viewers, but I don't know who they are. I don't know what the audience wants, and I don't care. I can't know, so I can't care", said a top television producer (Gans, 2004, p. 234). In his classic newsroom ethnography, the sociologist Herbert Gans sums up the attention given to the audience feedback by journalists as follows:

I began this study with the assumption that journalists, as commercial employees, take the audience directly into account when selecting and producing stories (...) I was surprised to find, however, that they had little knowledge about the actual audience and rejected feedback from it. Although they had a vague image of the audience, they paid little attention to it; instead, they filmed and wrote for their superiors and for themselves, assuming (...) that what interested them would interest the audience. (...) Journalists have access to formal feedback from the audience, but they use it only rarely. (...) Television journalists are even more dubious about their correspondents, believing them to be largely "nuts" or "cranks." (Gans, 2004, pp. 229–231)

While these examples are centered in the United States, the tendency to snub audience opinion can also be found elsewhere. In Brazilian newsrooms the roles were well established: "journalist and public in their original functions – the first with absolute control over what will or will not be published, the latter in a posture merely reactive to journalist work" (Madureira, 2010, p. 42).⁵¹ In Portugal, a former ombudsman from *Público* states that "the journalist has a certain tendency to despise [the letters to the editor], that is not for them, it was not produced by them, so it is

⁵¹ In the original: "Jornalista e público em suas funções originais – o primeiro com absoluto controle sobre o que será ou não publicado, o segundo em uma postura meramente reativa ao trabalho jornalista."

something that does not matter because they do not believe that readers are more interested in the letters section than in what journalists do” (Silva, 2014, p. 314).⁵²

In the last decade, the attitude towards audience feedback has changed. However, the possibility of monitoring readers’ behavior in real time is not the only reason for this shift of mindset on journalists’ approach. The economic and business priorities have become more palpable and present in the newsrooms.

The economic factor is an important force in journalistic activity. While the ideological pole defines journalism as a public service, journalism is done at companies that, in their overwhelming majority, aim to end the year with profits. While the ideological pole defines journalism as a public service, the economic pole defines journalism as a business, which tends to define the news as a commodity that sells newspapers or gets a good share of the audience. (Traquina, 2004, p. 207)⁵³

According to the Pew Research Center, the total estimated circulation of U.S. daily newspapers in 2017 is the lowest since 1940.⁵⁴ Even combining print and digital circulation still results in an overall drop in circulation by 4% in 2017 compared to 2016. The total estimated newspaper industry advertising revenue for 2017 was \$16.5 billion, a decrease of 10% compared to the previous year, while the total estimated circulation revenue was \$11 billion, which shows a slight increase of 3% compared to 2016 (Figure 32).⁵⁵

⁵² In the original: “o jornalista tem uma certa tendência para desprezar [as cartas], aquilo não é para eles, não foi produzido por eles, portanto é uma coisa que não interessa, porque eles não acreditam que os leitores tenham mais interesse pela secção de cartas do que por aquilo que os jornalistas fazem.”

⁵³ In the original: “O fator econômico é uma força importante na atividade jornalística. Enquanto o pólo ideológico define o jornalismo como um serviço público, o jornalismo é feito em empresas que, na sua esmagadora maioria, têm como objetivo acabar o ano com lucros. Enquanto o pólo ideológico define o jornalismo como um serviço público, o pólo econômico define o jornalismo como um negócio, que tem tendência para definir as notícias como uma mercadoria que vende jornais ou consegue um bom share da audiência.”

⁵⁴ Available at <https://www.journalism.org/chart/sotnm-newspapers-total-estimated-circulation-for-u-s-daily-newspapers/>. Accessed on March 25, 2019.

⁵⁵ Available at <https://www.journalism.org/chart/sotnm-newspapers-newspaper-industry-estimated-advertising-and-circulation-revenue/>. Accessed on March 25, 2019.

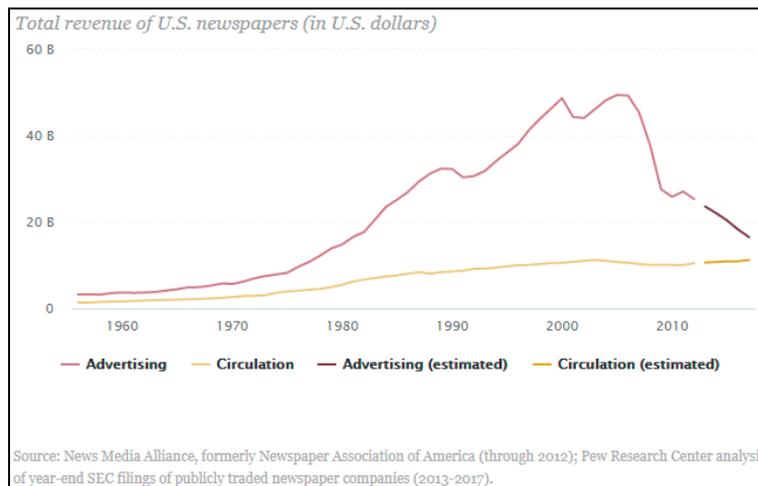


Figure 32 – Reproduction of Pew Research Center’s graph showing the estimated advertising and circulation revenue of the newspaper industry in the United States.

The digital advertising accounted for 31% of newspaper advertising revenue in 2017, while in 2011 that portion stood at 17%.⁵⁶ Finally, still according to the Pew Research Center, 39,210 people worked as reporters, editors, photographers, etc. in the newspaper industry in 2017.⁵⁷ That represents a decrease of 45% from 2004 (Figure 33).



Figure 33 – Reproduction of Pew Research Center’s graph showing the employment in newspaper newsrooms in the United States.

In Brazil, according to *Instituto Verificador de Comunicação*, the total estimated circulation of the country’s 11 main newspapers fell 41.4% between December 2014

⁵⁶ Available at <https://www.journalism.org/chart/sotnm-newspapers-percentage-of-newspaper-advertising-revenue-coming-from-digital/>. Accessed on March 25, 2019.

⁵⁷ Available at <https://www.journalism.org/chart/sotnm-newspapers-newsroom-employment/>. Accessed on March 25, 2019.

and December 2017 (from 1,256,322 to 736,346 copies per day).⁵⁸ Meanwhile, the digital circulation of these newspapers grew by only 31,768 in the same period. The project “A conta dos passaralhos” estimates 2,327 journalists laid off between 2012 and 2018 in Brazil (Figure 34).⁵⁹

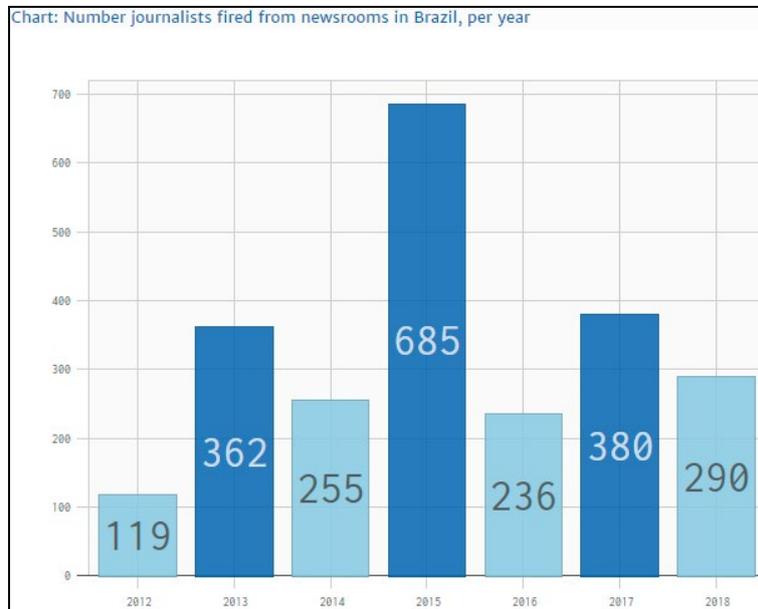


Figure 34 – Reproduction of “A conta dos passaralhos” graph showing the number of journalists fired from newsrooms in Brazil.

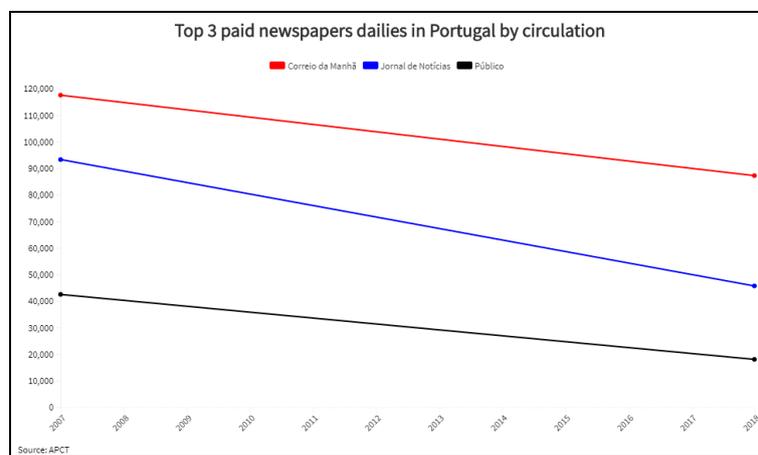


Figure 35 – Chart with the estimated circulation between 2007 and 2018 of the 3 largest daily newspapers in Portugal.

In Portugal, the total estimated circulation of the top 3 paid daily newspapers in Portugal dropped 40.5% between 2007 and 2018 (from 253,281 to 150,867 copies per

⁵⁸ Available at <https://www.poder360.com.br/midia/tiragem-imprensa-dos-maiores-jornais-perde-520-mil-exemplares-em-3-anos/>. Accessed on March 25, 2019.

⁵⁹ Available at <http://passaralhos.voltdata.info/>. Accessed on March 25, 2019.

day) (Figure 35).⁶⁰ According to Reuters Digital News Report, in 2018 only 8.6% of respondents paid for online news, 0.9% less than in 2017, one of the lowest figures in the survey (Cardoso, Paisana, & Martinho, 2018).

As is perceptible, the news industry is in crisis. It is not recent, but has been intensified in the last decades with the continuous decline in print circulation of the press, the migration of advertising revenue to online, the striking increase in competition with other players (video streaming, games, social media, etc.) (see Wu 2017), and the deterioration of the journalistic authority (Gitlin, 2011). This perfect storm accelerated the process of erosion of the traditional “wall” between the editorial and business sides of news organizations (Coddington, 2015b; Gade, 2004), since, as Tandoc (2014, p. 569) put it: “The bottom line is that the online audience translates into advertising revenues that translate into paychecks”. He elaborates:

What is different now is how journalists in these online newsrooms can no longer ignore the audience. It used to be that news editors did not have to think about how many people are attending to their news content, consistent with the wall of separation between the editorial department that protected its journalistic autonomy and the business department that took care of audience size and revenue. But things have changed. (Tandoc, 2014, p. 570)

The first studies showed that metrics had little impact on editorial decisions (Boczkowski & Peer, 2011; Graves et al., 2010; Lee-Wright, 2010; Singer, 2011). However, with the emergence of analytics services such as Chartbeat and Visual Revenue that provide real-time statistics, things began to change. Groves and Brown (2011) and Usher (2012) conducted ethnographic studies on the *Christian Science Monitor*, the United States’ largest nonprofit daily newspaper, when it went web-first in 2009. Pageviews became a key metric of success and the journalists “felt a loss of autonomy and authority over news decision making (...) The economic and business priorities had, at least for these journalists, become too central to the priorities of the newsroom” (Usher, 2012, p. 1916). One person on the business side of the organization said: “It’s a hard mindshift to take, that readers are not as interested in

⁶⁰ Data available at http://www.apct.pt/Analise_simples.php. Accessed on March 25, 2019.

what you want them to be interested in. Maybe they really are more interested in Paris Hilton than the deficit, even if you don't like it. You are giving up that control" (Groves & Brown, 2011, p. 111).

"We need to build traffic or lose our jobs", summarizes a journalist at the *Monitor* (Usher, 2012, p. 1909). A similar dissatisfaction is expressed by a reporter at the *Philly.com*, where C.W. Anderson conducted an ethnographic study: "We're probably headed toward a new model where reporters get paid by clicks" (Anderson, 2011a, p. 559). Website traffic numbers became an obsession at *Philly.com* and often the key ingredient in the editorial decision-making (Anderson, 2011a). On another ethnographic observation, a *Gawker Media* writer said he was "actually concerned by the extent to which my emotional well-being is dictated by the number of hits on my posts. I talk to my therapist about it!" (Petre, 2015, p. 25). And a managing editor compared, to Edson C. Tandoc, using web analytics with taking drugs: "It's like crack (...) you can sit here and watch it, popping all night" (Tandoc, 2014, p. 567).

Gawker and other digital-born companies such as *Buzzfeed* and *Upworthy* have based "their business models on crafting content that increases measures of user attention and interaction" (Martin & Dwyer, 2019, p. 9). *Gawker*, a gossip website founded in 2002, was shut down in 2016 after the media company declared bankruptcy. *Gawker* posted a sex tape of the former professional wrestler Hulk Hogan. Hogan sued, arguing that this was a violation of his privacy, and a Florida jury agreed, awarding Hogan \$140 million in damages. Ultimately, *Gawker* and Hogan reached a \$31 million settlement (Ember, 2016). In early 2019 *Buzzfeed* cut 15% of its workforce, more than 200 people were laid off (Alpert & Mullin, 2019).

Scholars have shown that editors are using metrics to decide the news placement, headline adjustment, lifespan of news stories, and subsequent reporting (see Welbers et al. 2016; Lee, Lewis, and Powers 2014; Tandoc 2014; Vu 2014; Anderson 2011b; MacGregor 2007). Surveying newsroom editors across the United States, Vu (2014) found out that 84% monitor web traffic on regular basis and that to some extent "editors are willing to adjust their editorial decision-making based on web metrics. This willingness is influenced by their perceived economic benefits of getting readership" (Vu, 2014, p. 1104).

In this manner, scholars are concerned that, with this movement toward “market-driven”, prioritizing quantity over quality, journalism is disregarding its role of not only giving what audiences want, but also giving what they need.

Under a market framework, the survival of excellent journalism depends largely on the desire for excellent journalism. (...) The evolution of the audience has understandably jolted traditional and familiar journalistic roles. The role, however, should not swing mindlessly into providing what the audience wants. The role should be about understanding what the audience wants and how journalists can take that information and balance this against what the audience needs. It is an unusual responsibility, but journalism is an unusual public good. Journalists need to be responsive to the public they serve, but they also should serve the higher goal of public interest. (Tandoc & Thomas, 2015, p. 253)

2.4. From data-driven to data-informed newsrooms

Aware of this situation, some newsrooms seek to be data-informed, and not data-driven, since “the most sophisticated audience teams are keenly aware that *analytics are not perfect*” (Cherubini & Nielsen, 2016, p. 7). Thus, some media organizations have started to develop their own analytics system to have not only a quantitative analysis, but also a more understandable overview of the audience, and so that the analytics tool can help in short-term decision optimization and longer-term planning (Cherubini & Nielsen, 2016).

Journalists today not only *need analytics* to navigate an ever-more competitive battle for attention. Many journalists also *want analytics*, as an earlier period of scepticism seems to have given way to interest in how data and metrics can help newsrooms reach their target audiences and do better journalism. That is encouraging, because analytics and data metrics will continue to evolve, and if journalists are not part of that process, the tools and techniques developed will continue to reflect and empower commercial and technological priorities more than editorial priorities. (Cherubini & Nielsen, 2016, p. 7)

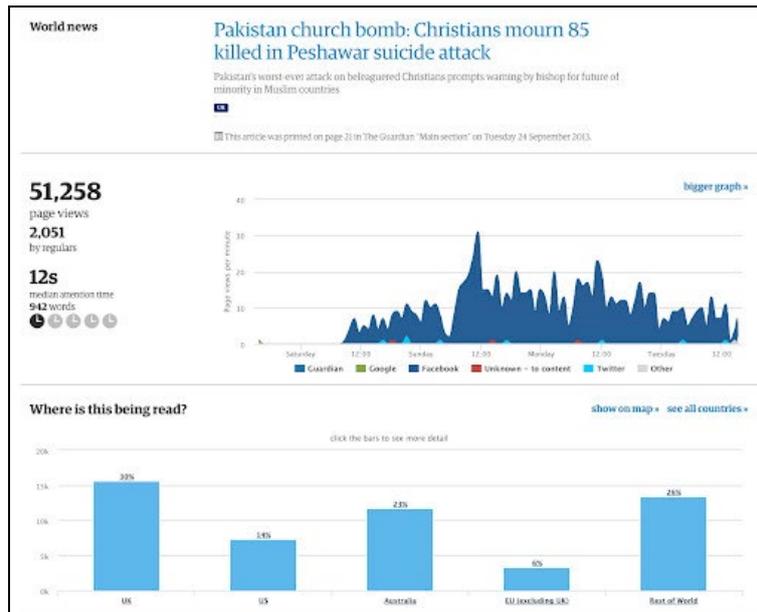


Figure 36 – Reproduction of *The Guardian's* Ophan system.

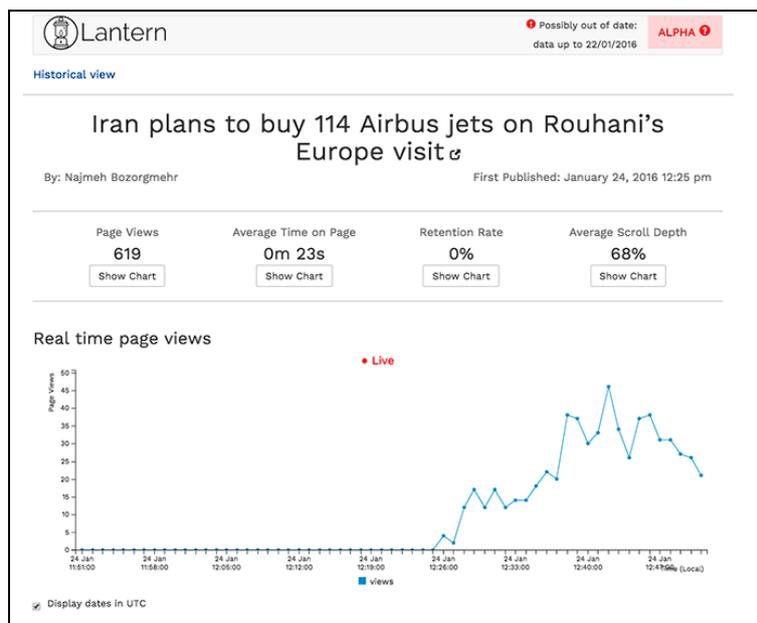


Figure 37 – Reproduction of the *Financial Times's* Lantern system.

The Guardian's in-house real-time analytics tool Ophan was launched in 2012 (Figure 36). In 2016 the *Financial Times* launched Lantern (Figure 37). Both platforms are an attempt to communicate data in a very clear way to journalists. They show not only traditional metrics (pageviews, unique visitors, etc.) but also metrics of engagement (time spent, number of comments, etc.) to have a more holistic picture of the success of a story. According to Renée Kaplan, the head of audience and new content strategies at the *FT*, the goal is not only growing the reach of *FT* journalism,

but also drive quality traffic to the website, expand the engagement and the impact of the newspaper (Cherubini & Nielsen, 2016; Lichterman, 2016).

As mentioned by Kapan, among the other ways of assessing the success of a story is the measurement of the impact on individuals, groups, or institutions. The notion of journalistic impact is not new, but in recent years it has been emphasized and gained more relevance – especially in nonprofit newsrooms (see Powers 2018; Konieczna and Powers 2017; Pitt and Green-Barber 2017; Schiffrin and Zuckerman 2015; C. Lewis and Niles 2013).

While news organizations increasingly track audience metrics and user engagement, and may define what impact means in their own newsroom, they rarely share this definition publicly or engage in a broader conversation about how journalists should define (or redefine) impact in the digital age. (Konieczna & Powers, 2017, p. 3)

Indeed, defining and measuring journalistic impact is hard enough (C. Lewis & Niles, 2013; Simons et al., 2017). Still, the impact that some stories have on the community is undeniable, and there are some forms of journalism more “impact-oriented” than others, such as investigative journalism (J. Hamilton, 2016; D. Merritt & McCombs, 2014). There are some tools and projects that try to measure the impact of news, such as *Chalkbeat’s* MORI (Measures of Our Reporting’s Influence),⁶¹ NewsLynx,⁶² Media Impact Project Measurement System,⁶³ and Impacto.jor⁶⁴ (Green, Cramer, & Anand, 2014; Keller & Abelson, 2015; Torres, 2017; Vieira, 2018). But these projects are still very incipient and there are no studies on their performance.

Lindsay Green-Barber (2019) argues that data journalism can also contribute greatly to various types of social change (i.e. on individuals; on networks; on institutions; and on public discourse). She mentions *ProPublica’s* “Dollars for Docs”⁶⁵ as a classic example of data journalism: an interactive story with the basic aim to encourage the audience to conduct their own exploration and possibly take action.

⁶¹ Available at <https://chalkbeat.org/mori/>. Accessed on March 25, 2019.

⁶² Available at <https://newslynx.readthedocs.io/en/latest/>. Accessed on March 25, 2019.

⁶³ Available at <http://mediaimpactproject.org/measurement-system.html>. Accessed on March 25, 2019.

⁶⁴ Available at <https://www.impacto.jor.br/>. Accessed on March 25, 2019.

⁶⁵ Available at <https://projects.propublica.org/docdollars/> Accessed on March 25, 2019.

“We want to know how you've used or might use this information in your day to day lives. Have you talked to your doctor? Do you plan to? Tell us”, she says on the project page.⁶⁶

Another case that combines investigative journalism and data journalism which had extensive repercussion and impact was the *International Consortium of Investigative Journalists' "The Panama Papers"* (Pitt & Green-Barber, 2017). Investigations were sparked in more than 80 countries, the prime ministers of Iceland and Pakistan resigned, some countries have adopted new rules forcing corporate transparency, and more than \$1.2 billion has been recouped in 22 countries since the publication of ICIJ's investigation in 2016 (Wilson-Chapman, Cucho, & Fitzgibbon, 2019).

To stay relevant, journalism must not only accept that it has an impact on society, but embrace that fact. By working to understand the ecosystem of change in which journalism functions, and its specific role within this system, the industry can work to maximize its positive impact and demonstrate its value to audiences. Data journalists, with their understanding for the value and importance of both quantitative and qualitative data, are well positioned for this endeavor. By articulating the goals of data journalism projects, developing creative audience engagement and distribution strategies, and building sophisticated methods for measuring success into these projects, reporters can lead this movement from within. (Green-Barber, 2019, p. n.p.)

Nevertheless, most of the newsrooms that are investing and embracing impact measurement are nonprofit, particularly because of the pressure from funders and donors to show a return on their investment (Konieczna & Powers, 2017). The *State of Technology in Global Newsrooms*, a 2017 study conducted by the International Center for Journalists (ICFJ) that received more than 2,700 responses from journalists in 130 countries shows that pageviews is the metric that gets the most attention from newsrooms (73%) and that only 1% of the organizations employ analytics editors (ICFJ,

⁶⁶ Available at <https://projects.propublica.org/docdollars/>. Accessed on March 25, 2019.

2017). As Chris Moran, *The Guardian's* editor for strategic projects, former audience editor, and co-creator of Ophan, put it:

Pageviews are often dismissed as vanity metrics — a measure whose sole purpose is to boost the ego of writers. But we don't write journalism for ourselves. Reaching a wide audience is a crucial part of the journalistic act. Of course, metrics at their core, are simply measurements, and have to be contextualized by expert editorial judgment. But they are a crucial part of making sure that news is disseminated both widely and responsibly. (...) The fact that audience metrics don't reflect the quality of a piece or necessarily capture its full impact isn't a weakness. Those things are subjective and a newsroom should care about them in any circumstances. Instead, metrics tell us things we absolutely don't, or can't, know from gut instinct. (Moran, 2019, p. n.p.)

In the next chapter, we present the metrics that we will analyze in this work; the journalistic contents that will be scrutinized; the comparisons that will be made; the limitations of this approach and the contributions of this work to the field of data journalism and audience data and metrics in online journalism.

III – Methodology

Stories with data have been in the newspapers for at least 150 years, as we saw in Chapter I. First with tables and numbers, then with charts and maps, and more recently with interactive features, characterized by its participatory openness. This form of journalism has gained more and more prominence. Since 2012, *The Washington Post* publishes every year a list with the most important data visualizations created by the newspaper that year. In 2018 the list had more than 40 graphs.⁶⁷ And the audience seems to like it. In 2018, out of the 10 most read stories in *The New York Times*, 8 were interactive.⁶⁸ But does data journalism also spark interest in audiences from other newspapers elsewhere? There are still few studies on the reach of data journalism pieces.

In the same way, as we saw in Chapter 2, since the beginning of the 20th century media outlets have been tracking their audience for commercial reasons. In recent years, however, this interest in the audience feedback has reached the newsrooms and there are studies on the use of web analytics in newsrooms in North America, Europe, South America, Africa, Asia, and Oceania (see, for instance, Moyo, Mare, and Matsilele 2019; Vieira 2018; Blanchett Neheli 2018; Antunes 2017; Hanusch 2017; Torres 2017; Cherubini and Nielsen 2016; Bunce 2015; Usher 2013; Groves and Brown 2011; Lee-Wright 2010; Schaudt and Carpenter 2009). However, most of these studies focus on the impact that quantitative audience is having in journalists' editorial decisions.

There is also some research on journalistic content that use lists of popular news to identify audience's interests (Bright & Nicholls, 2014; Lee et al., 2014; Schaudt & Carpenter, 2009; Welbers et al., 2016; Zamith, 2018). However, this approach has obvious limitations in using ordinal data. For example, relying on secondary data, it is not possible to know the distance between the most viewed items (e.g. how many

⁶⁷ Available at: <https://www.washingtonpost.com/graphics/2018/ns/best-graphics/> Accessed on February 25, 2019.

⁶⁸ Available at: <https://www.nytimes.com/interactive/2018/12/19/reader-center/top-stories.html> Accessed on February 25, 2019.

visits has the most popular item had in relation to the second?) or even compare the performance of those items with popular items from other days or weeks. These lists are often used to circumvent the fact that real numbers are difficult to access, since they are considered sensitive and strategic data by the news companies and have potential commercial implications (Couldry & Turow, 2014; Napoli, 2011).

Although page-view data are generally readily available to editors and managers at those organizations, those data are often out of the reach of scholars. (...) In lieu of that prized data, scholars often turn instead to the lists of popular items — typically titled “most viewed,” “most clicked,” or “most popular” — that appear on the homepages of many news organizations. (Zamith, 2015, p. 83)

In this work we analyze the performance of interactive pieces using continuous data in three newspapers: *The Texas Tribune* (the United States), *O Globo* (Brazil) and *Público* (Portugal). Among the audience measurement tools used in those newsrooms, Google Analytics is used by all three. We had Google Analytics access to all three newspapers, which means that continuous data was used to compare the performance of the news pieces. However, in the results we tend to concentrate on proportions rather than absolute numbers because of potential commercial implications and because it is a way to give more context and meaning to the data (Beasley, 2013).

Google Analytics has over 100 default reports at hand and it is possible to create practically infinite combinations of custom segments (Clifton, 2015). For example, it is possible to create a report that identifies how many users interacted with the data visualization; or create an event to monitor when the user scrolled down the page (Mokalis & Davis, 2018). Unfortunately, segmentations and events like those were not implemented in all three case studies in 2014 – year of the oldest data collected. Therefore, we have chosen to use only some of the default metrics.

Table 1 – Definition of Google Analytics metrics used in this study.⁶⁹

Variable	Description
Pageviews	Total number of times a page has been loaded in a browser.
Unique pageviews	Number of visits in which the page was viewed at least once.
Average time on page	Average amount of time a user spends on a specified page.
Entrances	Number of times visitors entered the site through a specific page.
Bounce rate	Percentage of visits in which the user leaves the site from the same page she/he entered it, i.e. single-page visit.
Exit rate	Percentage of how often users exit the site from a specific page.
Traffic sources	Source of referrals to a specified page.
Devices	Type of device used to visit the page.
Location	Country from which the visit originated.
New vs. Returning	Number of new visitors and returning visitors to a specific page.
Browsers	Web browser used to visit the page.

We analyzed 6 metrics: pageviews, unique pageviews, average time on page, entrances, bounce rate, and exit rates; and 5 dimensions: traffic sources, devices, location, new vs. returning visitors, and browsers (Table 1). According to Google,⁷⁰ metrics are quantitative measurements, while dimensions describe the data collected.⁷¹ These 11 variables can be divided into three groups: (1) success metrics; (2) engagement variables; and (3) behavioral reports.

Scholars have systematically shown that pageviews is the most important metric in the newsrooms, the “metric of success” (Groves and Brown 2011, 16; see also Blanchett Neheli 2018; Tandoc 2014; Usher 2012; Anderson 2011b; MacGregor 2007). The popularity of news items are commonly determined by the number of pageviews and unique pageviews, and those two metrics remain as the most tracked

⁶⁹ The book *Google Analytics Demystified* (2018) and the Google Analytics website (available at <https://support.google.com/analytics/>; accessed on February 25, 2019) were used as reference for the production of this table.

⁷⁰ Available at: <https://support.google.com/analytics/answer/1033861>. Accessed on February 25, 2019.

⁷¹ Metrics and variables will be used hereafter as equivalents and to the detriment of dimensions.

by newsrooms (ICFJ, 2017). Nevertheless, an increasing number of newsrooms are paying more and more attention to engagement metrics.

A news industry that used to obsess over print circulation and broadcast ratings has circled through a quick succession of preferred digital metrics, from clicks, pageviews, and unique users to engaged time. (...) Currently, pageviews and unique browsers are falling out of favour, and attention or engagement is seen as the future. (Cherubini & Nielsen, 2016, pp. 34–35)

Average time on page, new vs. returning visitors, bounce and exit rates are some of the variables used to try to understand user's engagement with the content. Finally, behavioral metrics such as devices, location, and browsers help to draw a profile of the users who access that content.

To analyze the dimensions, we use unique pageviews, which also returns a value closer to the whole website experience than pageviews (Mokalis and Davis 2018). In the following sections we will explain how the comparisons will be made.

3.1. Quantitative content analysis

In data journalism pieces, interaction is a key component (Appelgren, 2018; Bradshaw, 2014; Broussard & Boss, 2018; Usher, 2016), hence the benchmark used for the selection of data visualization analyzed was interactivity, “a measure of a media’s potential ability to let the user exert an influence on the content and/or form of the mediated communication” (Jensen, 1998, p. 201). We identify the interactive features deployed on the articles using the interaction techniques described by Boy et al. (2015) and Figueiras (2015), both works built on the taxonomy of Yi et al. (2007). Table 2 defines the 10 categories we used. Considering some of the news pieces often have multiple interaction techniques, in the results we focus on the 3 most prominent elements. In the Appendix we list all interaction techniques for each data journalism piece.

Ten interactive pieces were selected from each news media in 2014 in an arbitrary way; another 10 from 2015 and another 10 from 2016, totaling 90 data journalism pieces. Analyzing data visualization of different years allows us to observe

the evolution of interactive elements in the pieces. Have interaction techniques remained the same or have they become more sophisticated? Previous work has shown that the elements abstract/elaborate, inspect and filter are among the most common in data visualizations (Loosen et al., 2017; Stalph, 2018; Young et al., 2018).

Table 2 – Definition of the interaction techniques used in this study.⁷²

Category	Description
Filter	Show to the user some data conditionally (e.g. selecting a city from a drop-down menu).
Select	Mark something as interesting to keep track of it (e.g. clicking on a line graph).
Inspect	Show details on demand (e.g. hovering over a bar chart to get more information about the data).
Abstract/Elaborate	Adjust the level of abstraction (e.g. zoom in on a map to display more details).
Connect	Show related data (e.g. clicking on an element and highlighting all similar items).
Reconfigure/Encode	Display a different arrangement and/or representation (e.g. convert a pie chart to a bar graph and/or change the color).
Narrate/History	Show the steps of the story (e.g. clicking on a stepper-button to go back or forward in the visualization).
Explore	Display a different or specific subset (e.g. clicking on a query-button to scrutinize the data).
Collaboration	Allow the user to share their insights (e.g. post in social media a customized version of the data).
Gamification	Playable or playful way to present data (e.g. visualizations with game mechanics).

We also used content analysis to identify the topics of data visualizations. 10 categories were adopted, with 7 hard news: politics/governance, social issues, economy/business, environment/science/technology, foreign affairs, education, and health; and 3 soft news: entertainment, lifestyle/culture, and sports (J. B. Cox, 2014;

⁷² Boy et al. (2015) and Figueiras (2015) classifications and definitions were used as reference for the production of this table.

Newman, Fletcher, Levy, & Nielsen, 2016; North, 2016). There are already some studies on the main topics of data journalism pieces.

Analyzing 106 data-driven stories published in the UK by 15 mainstream news media, Knight (2015) found that stories covering social issues were dominant (27%). Tandoc and Oh (2017) examined 260 interactive articles published in *The Guardian's Datablog* and found that the most frequent topic was politics (19%). Loosen, Reimer, and De Silva-Schmidt (2017) analyzed 179 data projects nominated for the Data Journalism Award⁷³ between 2013 and 2015 and almost half of the articles cover a political topic (48.2%). Focusing on 26 Canadian stories that were finalists in 3 data journalism awards, Young, Hermida and Fulda (2018) found that the majority of the pieces addressed issues ranging from local news (23%) and social issues (23%). Finally, Stalph (2018) analyzed data-driven stories published by European quality news media and found that most of the articles are about politics (39.3%).

3.2. Quantitative analysis

In order to analyze the performance of data visualization, we collected the data described in Table 1 from the date of publication until the last day of the respective year.⁷⁴ Each data journalism piece was contrasted with two groups: (1) the 500 most read stories of that year in the same newspaper; (2) two news pieces from the same newspaper, on the same topic, and published in the same period as the interactive story. Thus, we performed a total of 180 analyses: 90 data visualizations vs. group 1; 90 data visualizations vs. group 2. We adopted these approaches because only by comparing numbers is it possible to understand the data, since isolated numbers are meaningless (Beasley, 2013).

In the end, all you can do is compare data to get context. (...) context meaning understanding data in relation to other, equivalent data, whether across time, between two entities such as pages, or both. (...) Therefore, you can't make sense of the numbers until you compare them — page to page, or the same pages over time. By seeing that, for example, users

⁷³ The competition is organized by the Global Editors Network and the website is available at <https://datajournalismawards.org/>. Accessed on March 25, 2019.

⁷⁴ Unfortunately, the data from *The Texas Tribune* in 2014 only go back as far as June 25.

spent an average of 50 seconds on all of the pages on your website, it becomes clear that a page with an average time of 1 minute and 20 seconds is well above average. (Beasley, 2013, p. 19)

For comparison between the selected data visualizations and the first group, we have chosen to apply the following variables: pageviews, unique pageviews, average time on page, entrances, bounce and exit rates.⁷⁵ The comparison with the first group allows us to observe:

- the performance (hits and visits) of the data journalism pieces in relation to the performance of the 500 most read stories;
- if users spend more or less time in interactive piece than they spend (on average) on the 500 most read stories;
- if the rate of entrances on the site through data journalism pieces is higher or lower than the rate of entrances on the site through the 500 most read stories;
- if the users that visit the interactive pieces are bouncing at a higher or lower rate than the visitors of the 500 most read stories;
- if the users are leaving the site from a data journalism pieces at a higher or lower rate than the visitors of the 500 most read stories.

As a result, a set of performance snapshots of different types of data visualizations was composed using success variables and engagement metrics and put into perspective when compared to the 500 most read news articles from the same newspaper in the same year.

To illustrate how the comparisons will work in the subsequent chapters, Table 3 shows the hypothetical performances of a data journalism piece in 2014 and the average of the 500 most read news in the same year. The difference to the comparisons that will be presented in the following chapters is that in these there will be no table with data.

⁷⁵ Considering the time available to access Google Analytics in some newsrooms was limited, we decided it would be best to leave out the metrics traffic sources, devices, location, new vs. returning visitors, browsers, since collecting these on the most read stories would be too laborious.

Table 3 – Hypothetical performances in 2014.

Variable	Data Journalism piece	Average of the 500 most read stories
Pageviews	18,500	10,000
Unique pageviews	10,600	9,500
Average time on page	00:00:50	00:02:00
Entrances	780	5620
Bounce rate	60%	80%
Exit	20%	50%

The hypothetical data journalism piece (H-DJ) registered more clicks and visits than the hypothetical average of the 500 most read (H-500). The interactive piece reported 85% more pageviews than the H-500, and 11% more unique pageviews.

Comparing pageviews and unique pageviews, it is possible to check if the article was accessed multiple times by the same user.

Unique pageviews represent the number of visits during which a page was viewed at least once, for example: If a visitor views the same web page seven times during the same visit, then it will count as seven pageviews but only one unique pageview. If the same visitor exits your site, but come back later after the session expires and views the same web page two more times, the metric will increased to nine pageviews and two unique pageviews. (Mokalis & Davis, 2018, p. 222)

That is, a discrepancy between pageviews and unique pageviews means that the article was clicked/refreshed multiple times by the user during the same session. A session ends in Google Analytics after 30 minutes of inactivity or at midnight (Mokalis & Davis, 2018). Finally, dividing pageviews by unique pageviews will show on average how many times a page was viewed during users' sessions.

According to the GOV.UK website, a "high ratio (above 1.4) indicates that users have to come back to that page within their session".⁷⁶ The ratio of the H-DJ is 1.7,

⁷⁶ Available at: <https://gov.uk/guidance/content-design/data-and-analytics>. Accessed on May 25, 2019.

meaning that the interactive piece registered a high return rate, while the H-500 reported a ratio of 1.05.

The ratio entrances/pageviews shows entrances to the webpage as a percentage of the pageviews. That is, more than half (56%) of the users who accessed the H-500 entered the website from the H-500 pieces. On the other hand, H-DJ had a very low ratio and nearly 96% of H-DJ visitors had accessed another page of the website before accessing the interactive piece.

On average, users who visited H-500 spent on these pieces more than twice the time of those who accessed H-DJ. H-500 registered an average of 2 minutes, while H-DJ reported an average time on page of 50 seconds.

The H-DJ bounce rate is lower than in H-500. While in H-500 8 out of 10 visitors that entered the website from these news pieces left the website without visiting another page, in the H-DJ were 6 out of 10. Finally, H-DJ registered a lower exit rate than the H-500, 20%. Meaning that only 2 in 10 users who visited the interactive piece left the website from it. The average exit rate in H-500 is 50%.

For comparison between the chosen data journalism piece and the second group we used all 11 variables. In this way it was possible to examine:

- the performance (hits and visits) of the data visualizations in relation to the performance of similar news pieces;
- if users spend more or less time in data journalism pieces than they spend (on average) on similar news articles;
- if the rate of entrances on the site through interactive pieces is higher or lower than the rate of entrances on similar news stories;
- if the users that visit data visualizations are bouncing at a higher or lower rate than the visitors of similar news articles;
- if the users are leaving the site from an interactive piece at a higher or lower rate than the visitors of similar news pieces;
- if there is overlap between the most frequent traffic sources of visitors of the interactive pieces and of users that access the similar news stories;

- if there is overlap between the devices most used by users that visit data visualizations and visitors of similar news articles;
- if data journalism pieces receive more or less hits from outside the country than the similar news articles;
- if there is overlap between the web browsers most used by users that visit data visualizations and visitors of similar news articles;
- if the users who access the interactive pieces return to visit it again more or less times than the visitors of similar news.

While as a consequence of the comparison with the first group we have an overview of the performance of the data visualizations, in the second approach we have a more detailed investigation, having not only access to success metrics and engagement metrics, but also a glimpse of the users' behavior and profile. Dimensions such as traffic sources, devices, location and browsers have some constraints that will be covered in the next section.

As long as in the first study data journalism pieces are compared to all types and genres of news, in the second study each interactive piece is contrasted to similar news stories on the same subject and published in the same period. That is, while in the former approach we compare users that access data journalism pieces with the regular audience of the media outlet website, in the second analysis we compare data visualization visitors with users that accessed similar news and are potentially interested in the same topics.

Finally, as stated before, all these comparisons are done only for the year of the publication of the news pieces. However, in the second analysis it is possible to examine the lifetime of the data journalism pieces and related news stories. Most of the online news are not visited after two days of having been published. And even popular stories normally having a lifespan of one to two weeks (Ho & Liu, 2015). Hence, pageviews were used to calculate the lifespan of each data journalism piece and the respective two similar news pieces (from the date it was published until December 31, 2017).

3.3. Selection of the case studies and limitations

The choice of the case study news sites is best understood through the contextualization of this PhD. The PhD in Digital Media is a joint program between Universidade NOVA de Lisboa and Universidade do Porto. During 2007-2018 the University of Texas was also in the partnership, and the program included visiting researchers in Austin, Texas, in the United States. Being a doctoral candidate from Brazil, it seemed natural to choose Brazilian, Portuguese and American news sites for case studies.

This decentralized approach with case studies from different parts of the world enriches the research and allows news questions such as: (1) is there a pattern in performance of data journalism pieces across the different news sites? (2) Is it possible to trace similar behaviors in the visitors of interactive pieces in the case studies?

The Texas Tribune (United States), *O Globo* (Brazil) and *Público* (Portugal) will be presented appropriately in subsequent chapters, but it is worth noting that they all are references in the production of data journalism pieces in their respective countries (Aitamurto et al., 2011; Alexandre, 2014; Gray et al., 2012; Moura, 2018). *O Globo* and *Público* are national legacy news media, while *The Texas Tribune* is a nonprofit media organization that mostly covers the state of Texas. With two traditional newspapers and one digital-born media, it is possible to propose other questions, such as: (1) does the performance of *The Texas Tribune* data journalism pieces stand out from the performance of interactive pieces from the other media outlets? (2) The lifespan of the data journalism pieces of *The Texas Tribune* are dissonant from the lifetime of the interactive pieces of the legacy media?

In this manner, this work presents a series of contributions to the field of data journalism and audience data and metrics in online journalism: (1) performance analysis of data journalism pieces compared with most read stories; (2) performance of data journalism pieces compared with news on the same topic; (3) comparison of the performance of interactive pieces in different newspapers; (4) investigation of the lifespan of news stories; (5) study of news media visitor behavior; (6) use of Google Analytics real data; (7) the most used interaction techniques in different newspapers;

(8) the main topics of data journalism pieces in different media outlets; (9) web archiving of the analyzed interactive news pieces.

In the Appendix, all 90 data journalism pieces are displayed with their respective interaction techniques, subjects, and screenshots. We also provide the original URL and a link on the Internet Archive's Wayback Machine.⁷⁷ In this way, we try to contribute for this information to be stored, accessible, and discoverable for future works in the field. Still, we know these efforts are limited and that a solution to the preservation of data journalism content is much more complex, as Broussard and Boss point out:

Now as ever, newspapers are undergoing a rapid and continuous evolution of journalism production, publishing, and distribution. This digital shift is a rich site of study for current and future communication scholars. However, scholarly work proceeds from the assumption that the scholar is able to access a robust corpus of artifacts. It is not currently safe to make this assumption, given that newsrooms and libraries cannot currently archive dynamic digital news content at scale in any way. (...) It is imperative that newsrooms, libraries, and cultural memory institutions work together to find a solution for capturing and archiving these works as soon as possible. (Broussard & Boss, 2018, pp. 1216–1217)

This methodological approach, however, has several limitations. Chapter II has shown that server-based tools are heavily used in newsrooms, but web analytics like Google Analytics have some crucial restraints. Pageviews tends to inflate the number of hits as it counts each time the page has been refreshed. In the average time on page metric, it is not possible to conclude “whether users are actively interacting with a site or simply on it (while potentially doing something else)” (Cherubini & Nielsen, 2016, p. 35). Google Analytics does not generate reports of some metrics from the total number of visits, but from small samples of the total (e.g. location, web browser) (Mokalis & Davis, 2018). Also geographic data may not be reliable, since users may be accessing the page using a virtual private network, which hides the user's real location.

⁷⁷ The Wayback Machine is a digital archive of the internet. Anyone can add a URL and save a webpage. Available at <https://archive.org/web/>. Accessed on February 25, 2019.

These limitations reach all case studies in a cross-sectional way, since we used the same tool and the same parameters to collect the data in the three news organizations.

The study analyzes only quantitative data. We investigated *what* happened and *when*, but we did not study *why*. In other words, we analyze whether users clicked more or less on interactive pieces compared to other contents, but it is not wise to conclude *why* people clicked (or not) on the data journalism pieces. In the same way, the term *performance* is used in this work in a very reductive way. It is defined quantitatively: a high number of pageviews and time on page is good; a high bounce rate is bad. Finally, as Loosen, Reimer, and De Silva-Schmidt (2017, p. 14) explain: “quantitative content analyses are designed to reduce the complexity of their objects of investigation and are unable to detect developments that occur ‘below the radar’ of the variables and categories used”. In this manner, the quality of the design or of the interactive feature, the impacts on the community or even the efficiency of the data visualization (e.g. understand, engagement, memorability), none of this is taken into account in this work.

IV – *The Texas Tribune*

4.1. History

The Texas Tribune “is the only member-supported, digital-first, nonpartisan media organization that informs Texans — and engages with them — about public policy, politics, government and statewide issues”.⁷⁸ The nonprofit media outlet was founded by venture capitalist John Thornton and veteran journalists Evan Smith (now chief executive officer) and Ross Ramsey (now executive editor) after they observed that the number of full-time, professional journalists covering government and public affairs in the State Capitol was diminishing (Batsell, 2017). Launched on November 3, 2009 and based in Austin, according to the 2018 Annual Report,⁷⁹ in almost 10 years the *Tribune’s* team has grown from 17 employees to almost 70; raised more than \$60 million; won more than 80 awards for outstanding enterprise, investigative and data journalism; and has a monthly audience of 5.8 million pageviews and 2 million unique visits.

From the early stage interactive pieces have been a key feature on the *Tribune*, making “an immediate splash through its use of data — namely, creating searchable (and SEO-friendly) databases of public records, the sort usually only accessible through FOIA⁸⁰ requests. (...) That use of data was a way to have an impact quickly even as the traditional newsroom work was growing” (Ellis, 2014, p. n.p.). It worked so well that by 2014 *The Texas Tribune* had more than 200 different data applications and visualizations, driving as much as two-thirds of the website’s traffic (Howard, 2014; Stencel, Aidar, & Kamalakanthan, 2014).

The Texas Tribune acknowledges the participatory openness of data journalism and, in many “news apps”, invites the user to explore the databases. Those databases also have a crucial role in holding those in power accountable: “These data visualizations provide context for our reporters’ watchdog reporting, fact-based

⁷⁸ Available at: <https://www.texastribune.org/about/>. Accessed on February 25, 2019.

⁷⁹ Available at:

<https://s3.amazonaws.com/static.texastribune.org/media/documents/2018%20Texas%20Tribune%20-%20Annual%20Report.pdf>. Accessed on February 25, 2019.

⁸⁰ Freedom of Information Act (FOIA) is a U.S. law that provides citizens the right to access information.

journalism and deep-dive investigations — and serve as useful standalone tools for our readers.”⁸¹

4.2. Data journalism pieces from 2014 and the average of the 500 most read articles

As explained in Chapter III, unfortunately, the data from 2014 only go back as far as June 25. Also, some of the interactive pieces chosen were originally published in previous years but were updated and relaunched in 2014. Table 4 shows the title of the piece, up to 3 interaction techniques, the main topic and the link to the article (when available). In the Appendix there are all the interaction techniques and all topics of each data visualization, as well as a screenshot of the webpage.

Table 4 – The 10 interactive pieces chosen in 2014 from *The Texas Tribune* and their specific aspects.

Title	Interaction techniques	Topic	URL
Government Salaries Explorer	Explore.	Politics/governance.	http://bit.ly/2VBtCsz
Election Night	Filter, Select.	Politics/governance.	http://bit.ly/2JeDqlr
Elected Officials Directory	Filter, Explore, Reconfigure/Encode	Politics/governance.	http://bit.ly/2VIINRz
Texas Prison Inmates	Explore, Inspect.	Social issues.	http://bit.ly/2Y9sauZ
Texas Public Schools Explorer	Explore.	Education.	http://bit.ly/2J1oE24
Interactive: Search CSCOPE Lesson Plans	Filter, Explore.	Education.	http://bit.ly/2POsD2S
Map: Comparing the 2010 and 2014 Governor's Races	Inspect, Select, Abstract/Elaborate.	Politics/governance.	http://bit.ly/2ZXdZuA
Texas Hospitals Face Penalties Over Infections	Inspect, Select, Abstract/Elaborate.	Health.	http://bit.ly/2H2vD7Q
Higher Education Explorer	Explore, Inspect, Abstract/Elaborate.	Education.	Not available.
Interactive Map: Find Texas' Remaining Abortion Clinics	Inspect, Filter, Abstract/Elaborate.	Health.	http://bit.ly/2LlvPUX

⁸¹ Available at:

<https://s3.amazonaws.com/static.texastribune.org/media/documents/2018%20Texas%20Tribune%20-%20Annual%20Report.pdf>. Accessed on February 25, 2019.

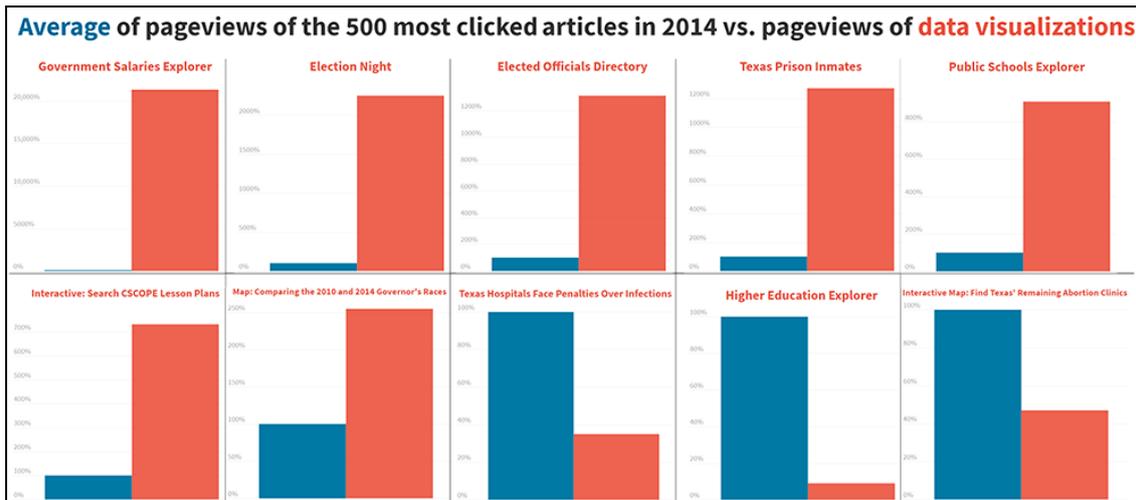


Figure 38 – Performance in pageviews of *The Texas Tribune* interactive pieces compared to the average of the most read articles in 2014 (represented as 100%).⁸²

Politics/governance was the most recurrent topic, present in 40% of the interactive pieces, followed by education, the subject matter of 30% of the data journalism pieces. Explore was the most used interaction technique, introduced in 60% of the data visualizations, followed by inspect, present in half of the interactive pieces.

Figure 38 shows pageviews performance of the data journalism pieces compared to the average of the 500 most read stories in 2014. Seven out of 10 interactive pieces registered more pageviews than the average of the 500 articles and 6 data journalism pieces are among the 10 most read news pieces of 2014: “Government Salaries Explorer”, “Election Night”, “Elected Officials Directory”, “Texas Prison Inmates”, “Public Schools Explorer”, and “Interactive: Search CSCOPE Lesson Plans”.

“Government Salaries Explorer” reported the best performance: 213 times more pageviews than the average of the 500 most clicked stories in 2014. “Higher Education Explorer” registered the worst performance: about one-tenth pageviews of the average of the 500 articles. Finally, 4 of those interactive pieces are also among the top 10 most read stories of 2015 and 3 are among the 10 most clicked news articles of 2016.

⁸² Available at: <https://iloaguardotcom1.files.wordpress.com/2019/12/djxavg-2014.png>. Accessed on December 19, 2019.

The ratio of pageviews/unique pageviews shows how many times a page was accessed during users' sessions, i.e. it is possible to check if a news article was viewed multiple times by the same user. GOV.UK website considers a high ratio above 1.4.⁸³

The ratio of the average of the 500 most read stories in 2014 is 1.2. "Elected Officials Directory" has the highest ratio between the data journalism pieces, 1.8. Other 4 interactive pieces registered a ratio higher than 1.4: "Election Night", "Texas Prison Inmates", "Public Schools Explorer", and "Higher Education Explorer". "Interactive Map: Find Texas' Remaining Abortion Clinics" reported the lowest ratio, 1.07.

Looking at the ratio entrances/pageviews, "Interactive: Search CSCOPE Lesson Plans" is proportionately the interactive piece through which more visitors entered *The Texas Tribune's* website, 81%. The proportion of the average of the 500 most clicked news stories in 2014 is 56%. There are 4 data journalism pieces that reported a percentage lower than 56% and "Elected Officials Directory" had the lowest: 10%. Which indicates that 9 out of 10 visitors of this data visualization had visited another page of *The Texas Tribune* before accessing it.

On average, users who visited "Election Night" and "Interactive: Search CSCOPE Lesson Plans" spent on these data journalism pieces more than twice the time of those who accessed the 500 most read news articles in 2014. On the other hand, 5 interactive pieces registered visits that lasted less than the average of the 500 articles. Most notably is "Government Salaries Explorer", where users spent on average 83% less time on page than the visitors of the most clicked news stories.

The average bounce rate of the 500 most read news articles in 2014 is 72%. Which means that about 7 out of 10 visitors that entered the website from these news stories left *The Texas Tribune's* website without visiting any other page. Among the data journalism pieces, "Interactive Map: Find Texas' Remaining Abortion Clinics" reported the highest value: 91%. Half of the interactive pieces registered a lower bounce rate than the 500 news stories, and "Government Salaries Explorer" had the lowest, 26%.

⁸³ Available at: <https://gov.uk/guidance/content-design/data-and-analytics>. Accessed on May 25, 2019.

“Government Salaries Explorer” also registered the lowest exit rate, 10%. That is, only 1 in 10 users who visited this data journalism piece left *The Texas Tribune’s* website from it. The average exit rate from the 500 most read news articles in 2014 is 58%. Half of the interactive pieces reported a percentage higher than 58% and “Interactive: Search CSCOPE Lesson Plans” had the highest exit rate: 82%.

4.3. Data journalism pieces from 2014 and their related stories

Figure 39 shows the performance (in pageviews) of the interactive pieces compared to their respective related news stories. The data journalism pieces are displayed as 100%. That way, related articles that registered a higher number of pageviews than the interactive piece appear in the figure with more than 100%; and those related news stories that registered fewer clicks than the data visualization appear in the figure with less than 100%.

Nine out of 10 data journalism pieces reported more clicks than their respective related news stories. “Government Salaries Explorer” is the interactive piece that performed best: almost 400 times more pageviews than the related story with better performance. “Higher Education Explorer” is the only data journalism piece that registered fewer hits than the related stories. Information on related news stories such as title, publication date, and link to the article are available in the Appendix.

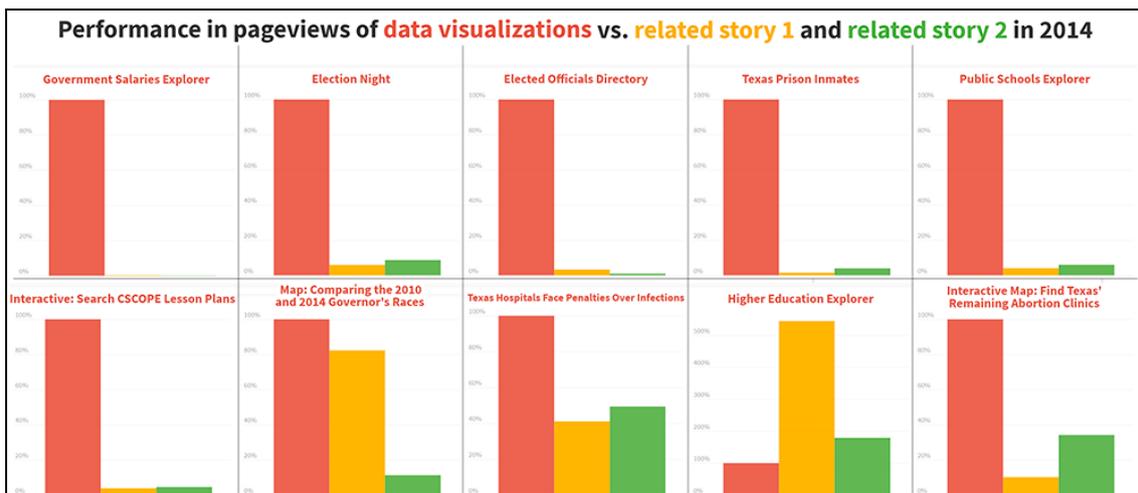


Figure 39 – Performance in pageviews of *The Texas Tribune* interactive pieces (represented as 100%) compared to their related news stories in 2014.⁸⁴

⁸⁴ Available at: <https://iloaguardotcom1.files.wordpress.com/2019/12/djxrel-2014.png>. Accessed on December 19, 2019.

With the related news stories, it is also possible to look at the lifetime of the articles. Figure 40 shows the percentage of pageviews in 2015, 2016 and 2017 compared to 2014, represented as 100%. That is, we can observe the evolution of clicks that the data journalism pieces and their related news stories have registered over the years.

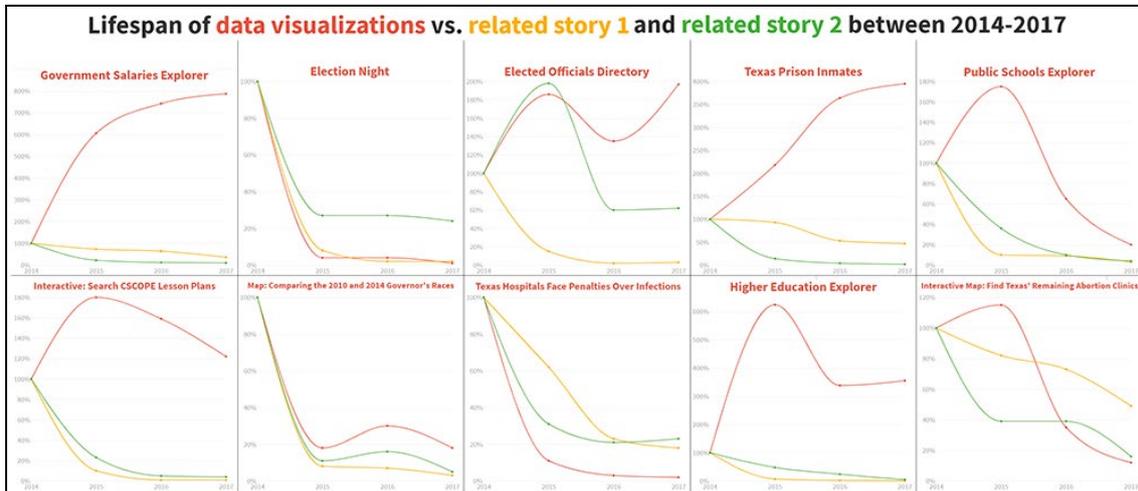


Figure 40 – Evolution in pageviews of *The Texas Tribune* 2014 interactive pieces and their related news stories over time.⁸⁵

Nineteen of the 20 related news stories registered fewer pageviews in 2015 than in the previous year, which is the normal behavior of news articles. However, 7 of the 10 interactive pieces reported more clicks in 2015 than in 2014. And half of the data journalism pieces reported more pageviews in 2017 than in 2014: “Government Salaries Explorer”, “Elected Officials Directory”, “Texas Prison Inmates”, “Interactive: Search CSCOPE Lesson Plans”, and “Higher Education Explorer”. Once again “Government Salaries Explorer” is the interactive piece that performed best: almost 8 times more pageviews in 2017. Among the 10 data journalism pieces, “Election Night” had the worst performance in 2017: fewer than 1% of the number of hits registered in the year of its publication.

All the related news stories in 2014 reported a ratio of pageviews/unique pageviews below 1.2, the average ratio of the 500 most clicked news articles that year. “Map: Comparing the 2010 and 2014 Governor’s Races” and “Interactive Map: Find Texas’ Remaining Abortion Clinics” showed a proportion lower than at least 1 of the

⁸⁵ Available at: <https://iloaguardotcom1.files.wordpress.com/2019/12/djxrel-2014-17.png>. Accessed on December 19, 2019.

related articles. That is, in the other 8 data visualizations, the visitors returned to the page during the same session more times than those who accessed their related news stories.

However, over the years, the number of returning visitors is lower on most data journalism pieces compared to related news. According to Google Analytics, the user that visits the website within the past two years and returns from the same device is marked as a returning visitor. Between 2014 and 2017, “Interactive: Search CSCOPE Lesson Plans”, “Higher Education Explorer”, and “Interactive Map: Find Texas’ Remaining Abortion Clinics” are the interactive pieces with returning visitor rates higher than their related news stories. That is, the other 7 interactive pieces had, in that period, a higher proportion of new visitors than their related news stories.

The ratio of entrances/pageviews is higher in 4 data journalism pieces than in their related news articles: “Public Schools Explorer”, “Interactive: Search CSCOPE Lesson Plans”, “Map: Comparing the 2010 and 2014 Governor’s Races”, and “Texas Hospitals Face Penalties Over Infections”. That is, the proportion of users who entered the website through these interactive pieces was higher than those who came to *The Texas Tribune* through their related news. In the other 6 data visualizations the rate is lower than in their related stories, which indicates that the visitors to these interactive pieces visited another page of the website before accessing it more times than the visitors of the related articles.

Users who visited “Election Night” and “Interactive: Search CSCOPE Lesson Plans” spent on these data journalism pieces about twice the time of those who accessed their related news stories. The other 8 interactive pieces reported less time on the page than at least 1 of the related articles. The users spent about 7 times more time on related news of “Government Salaries Explorer” than on this data journalism piece.

There are 4 interactive pieces where the bounce rate is higher than in at least 1 related news article. Which means that on these 4 data journalism pieces the visitors entered the website through these data visualization and left the website without visiting any other *The Texas Tribune’s* page in a proportion higher than in the respective related news. “Government Salaries Explorer”, “Election Night”, “Elected

Officials Directory”, “Texas Prison Inmates”, “Public Schools Explorer”, and “Higher Education Explorer” registered a lower bounce rate than their related stories.

The exit rates are similar. Here 6 interactive pieces also had a lower proportion than their related articles: “Government Salaries Explorer”, “Election Night”, “Elected Officials Directory”, “Texas Prison Inmates”, “Higher Education Explorer”, and “Interactive Map: Find Texas’ Remaining Abortion Clinics”. That is, users who visited these data journalism pieces then visited other pages of *The Texas Tribune*’s website in a greater proportion than those who visited the related news.

In all 10 interactive pieces, more than 50% of the visitors used the desktop to access the story. In “Higher Education Explorer”, about 93% of visitors used the desktop. Mobile and tablet were most frequently used by users who accessed “Election Night”: 37% and 12%, respectively. 12 of the 20 news articles registered more than 50% of their traffic coming from desktops.

Google is the place where most users were before visiting 8 of the interactive pieces in 2014. Only “Texas Hospitals Face Penalties Over Infections” and “Interactive Map: Find Texas’ Remaining Abortion Clinics” reported more visitors arriving from another location. 17 of the 20 related news stories also registered most of the users coming from Google.

About 30% of “Government Salaries Explorer” visitors in 2014 were from outside the United States. Among the other interactive pieces, only “Texas Hospitals Face Penalties Over Infections” reported a 100% U.S. audience. All the 20 related articles recorded more than 90% of users accessing from the United States.

Google Chrome was the browser most used by visitors of 7 of the data journalism pieces in 2014. “Government Salaries Explorer”, “Texas Hospitals Face Penalties Over Infections”, and “Higher Education Explorer” registered other browsers as the most used. 13 of the 20 related news were also accessed mostly by visitors using Chrome.

4.4. Data journalism pieces from 2015 and the average of the 500 most read articles

Some of the interactive pieces chosen were originally published in previous years but were updated and relaunched in 2015. Table 5 shows some information about the data journalism pieces from 2015: the title of the piece, up to 3 interaction techniques, the main topic and the link to the article (when available). In the Appendix there are all the interaction techniques and all topics of each data visualization, as well as a screenshot of the webpage.

Table 5 – The 10 interactive pieces chosen in 2015 from *The Texas Tribune* and their specific aspects.

Title	Interaction techniques	Topic	URL
Interactive: Undocumented Students on In-State Tuition	Explore.	Education.	http://bit.ly/2Y5PWb6
Ethics Explorer	Filter, Explore, Reconfigure/Encode.	Politics/governance.	Not available.
Faces of Death Row	Filter, Connect, Inspect.	Social issues.	http://bit.ly/2Vdyyo7
Texas Hospitals Explorer	Explore.	Health.	http://bit.ly/2GVeAo6
Texas Reservoir Levels	Inspect, Select, Abstract/Elaborate.	Environment / science/technology.	http://bit.ly/2vC1QgH
Interactive: Demographics of Texas' Undocumented Population	Inspect, Filter, Abstract/Elaborate.	Social issues.	http://bit.ly/2UZQ6yT
The 84 th Texas Legislature, by the Numbers	Inspect, Filter.	Politics/governance.	http://bit.ly/2H1SJww
Interactive: How Many Texans Resemble You?	Gamification, Filter.	Social issues.	http://bit.ly/2H1AY0g
Texas vs. the Feds	Filter, Select, Connect.	Politics/governance.	http://bit.ly/2ZVgDRw
See Vaccine Exemptions in Texas by School District	Explore.	Health.	http://bit.ly/2VlpQxs

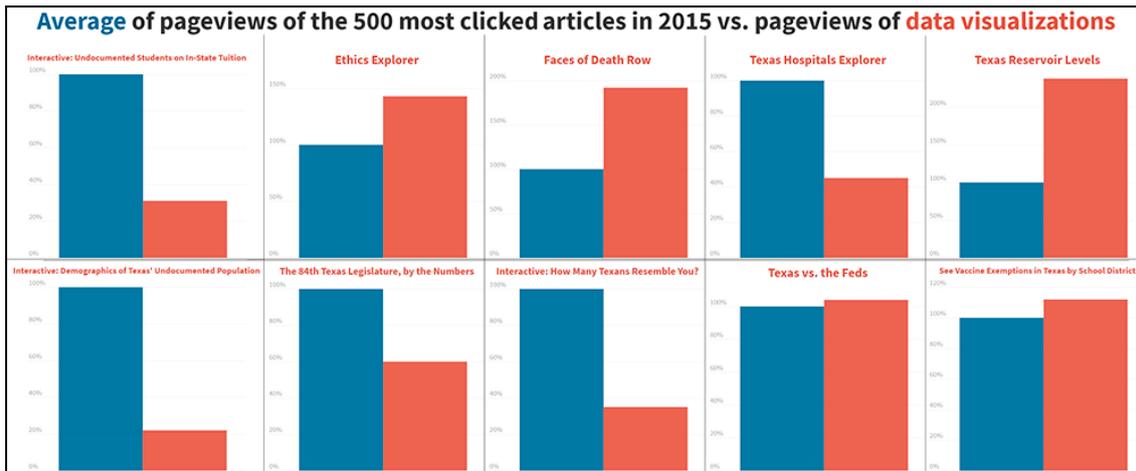


Figure 41 – Performance in pageviews of *The Texas Tribune* interactive pieces compared to the average of the most read articles in 2015 (represented as 100%).⁸⁶

Politics/governance and social issues were the most frequent topics, each one present in 3 interactive pieces, followed by health and education, both topics the subject of 2 data visualizations each. Filter was the most used interaction technique, introduced in 60% of the data journalism pieces, followed by explore, present in half of the interactive pieces.

Figure 41 shows pageviews performance of the data journalism pieces compared to the average of the 500 most read news articles in 2015. Half of the interactive pieces had more hits than the average of the 500 most read news stories and 2 are among the 30 most clicked articles of 2015: “Texas Reservoir Levels” and “Faces of Death Row”.

“Texas Reservoir Levels” registered the best performance: more than twice as many pageviews as the average of the 500 most read stories; while “Interactive: Demographics of Texas’ Undocumented Population” presented the worst, about a quarter of the clicks of the average of the 500 news articles.

The average ratio of pageviews/unique pageviews of the 500 most clicked news stories in 2015 is 1.2. “Ethics Explorer” reported the highest ratio between the data journalism pieces, 3.2. That is, for each unique pageview more than 3 pageviews have been recorded. “Texas Hospitals Explorer” also registered a ratio higher than the average, 1.7. The other 8 interactive pieces registered a proportion lower than the

⁸⁶ Available at: <https://iloaguardotcom1.files.wordpress.com/2019/12/djxavg-2015.png>. Accessed on December 19, 2019.

average of the 500 most read articles, and “Interactive: How Many Texans Resemble You?” reported the lowest ratio, 1.07.

Looking at the ratio of entrances/pageviews, “See Vaccine Exemptions in Texas by School District” is proportionately the data journalism piece through which more visitors entered *The Texas Tribune’s* website, 78%. The average percentage of the 500 most clicked news stories in 2015 is 63%. Half of the interactive pieces registered a ratio lower than 63% and “Ethics Explorer” reported the lowest: 13%. Which indicates that about 9 out of 10 visitors of this interactive piece had visited another page of *The Texas Tribune* before accessing it.

On average, users who visited “Faces of Death Row” spent on this data journalism piece almost twice the time of those who accessed the 500 most read news articles in 2015. There are 3 interactive pieces that registered visits that lasted less than the average of the 500 news stories. Most notable is “Ethics Explorer”, where users spent on average 75% less time on page than the visitors of the most clicked news stories.

The average bounce rate of the 500 most read news articles in 2015 is 74%. Which means that – as in 2014 – about 7 out of 10 visitors that entered the website through these news stories left *The Texas Tribune’s* website without visiting any other page. Among the data journalism pieces, “Texas Reservoir Levels” reported the highest value: 90%. Only 2 interactive pieces registered a lower bounce rate than the 500 most clicked news stories, and “Ethics Explorer” recorded the lowest proportion, 28%.

“Ethics Explorer” also reported the lowest exit rate, 9%. That is, about 1 in 10 users who visited this data visualization left *The Texas Tribune’s* website from it. The average exit rate from the 500 most read news articles in 2015 is 63%. There are 7 interactive pieces that registered a percentage higher than 63%, and “Texas Reservoir Levels” had the highest exit rate: 81%.

4.5. Data journalism pieces from 2015 and their related stories

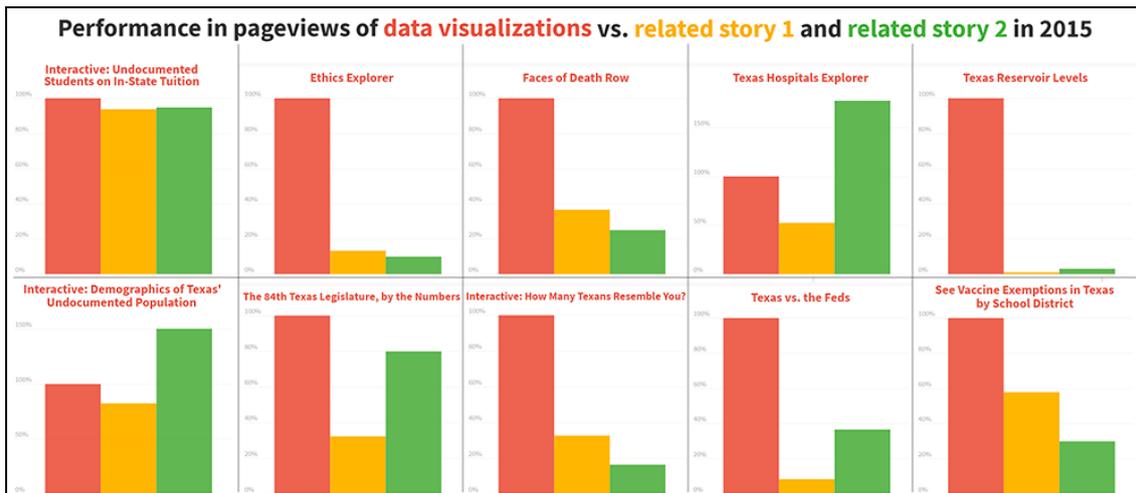


Figure 42 – Performance in pageviews of *The Texas Tribune* interactive pieces (represented as 100%) compared to their related news stories in 2015.⁸⁷

Figure 42 shows the performance (in pageviews) of the interactive pieces compared to their respective related news stories. The data journalism pieces are displayed as 100%. That way, related articles that had a higher number of pageviews than the data visualization appear in the figure with more than 100%; and those related news stories that registered fewer hits than the interactive pieces appear in the figure with less than 100%.

Eight out of 10 data journalism pieces registered more clicks than their respective related news stories. “Texas Reservoir Levels” is the interactive piece that performed best: almost 33 times more pageviews than the related story with better performance. “Texas Hospitals Explorer” and “Interactive: Demographics of Texas’ Undocumented Population” are the only data visualizations that reported fewer hits than at least one of the related stories. Information on related news stories such as title, publication date, and link to the article are available in the Appendix.

⁸⁷ Available at: <https://iloaguardotcom1.files.wordpress.com/2019/12/djxrel-2015.png>. Accessed on December 19, 2019.

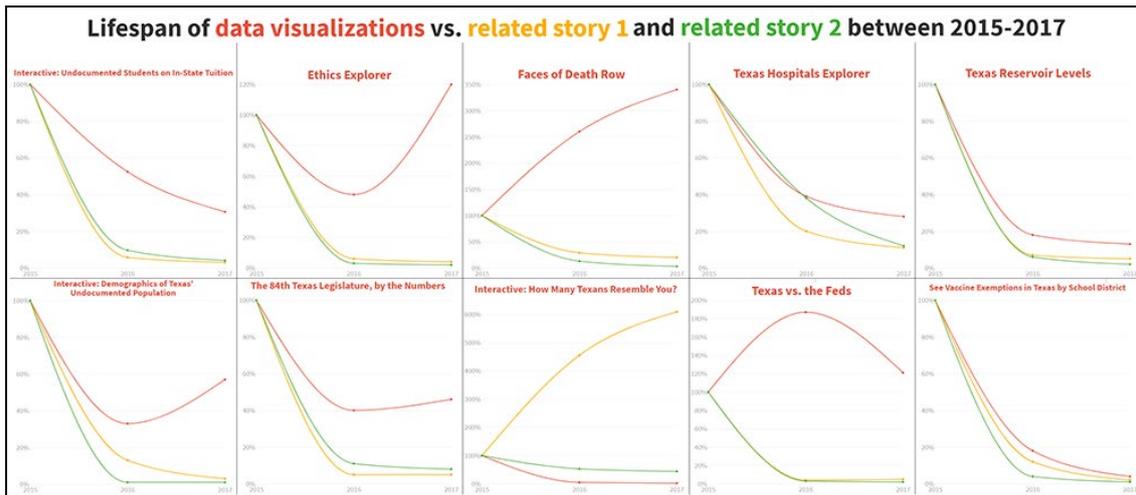


Figure 43 – Evolution in pageviews of *The Texas Tribune* 2015 interactive pieces and their related news stories over time.⁸⁸

Figure 43 shows the percentage of pageviews in 2016 and 2017 compared to 2015, represented as 100%. That is, we can observe the evolution of hits that the interactive pieces and their related news stories have registered over the years.

Nineteen of the 20 related news stories had fewer pageviews in 2016 than in the previous year. There are 2 data journalism pieces that reported more hits in 2016 than in 2015. And 3 interactive pieces registered more pageviews in 2017 than in 2015: “Ethics Explorer”, “Faces of Death Row”, and “Texas vs. the Feds”. “Faces of Death Row” is the data visualization that performed best: about 3.5 times more pageviews in 2017. Among the 10 interactive pieces, “Interactive: How Many Texans Resemble You?” presented the worst performance in 2017: less than 2% of the number of hits it registered in the year of its publication.

Looking at the ratio of pageviews/unique pageviews in 2015, 19 of the 20 related news stories reported a number below 1.2, the average ratio of the 500 most clicked news articles that year. Half of the data journalism pieces reported a proportion lower than at least 1 of the related articles: “Texas Reservoir Levels”, “Interactive: Demographics of Texas’ Undocumented Population”, “The 84th Texas Legislature, by the Numbers”, “Interactive: How Many Texans Resemble You?”, and “Texas vs. the Feds”. That is, in the other 5 interactive pieces visitors returned to the

⁸⁸ Available at: <https://iloaguardotcom1.files.wordpress.com/2019/12/djxrel-2015-17.png>. Accessed on December 19, 2019.

page during the same session more times than those who accessed their related news stories.

Between 2015 and 2017 the rate of returning visitors is higher on half of the data journalism pieces compared to their related news: “Interactive: Undocumented Students on In-State Tuition”, “Ethics Explorer”, “Texas Hospitals Explorer”, “Interactive: Demographics of Texas’ Undocumented Population”, and “Interactive: How Many Texans Resemble You?”. “Texas Reservoir Levels”, “The 84th Texas Legislature, by the Numbers”, and “Texas vs. the Feds” are the interactive pieces with a proportion of new visitors higher than their related articles in that period.

The ratio of entrances/pageviews is higher in 3 data visualizations than in their related news articles: “Texas Reservoir Levels”, “Interactive: How Many Texans Resemble You?”, and “Texas vs. the Feds”. The other 7 interactive pieces recorded a proportion lower than at least 1 of the related news, which implies that the visitors to these data journalism pieces visited another page of *The Texas Tribune’s* website before accessing it more times than the visitors of the related articles.

“Interactive: Undocumented Students on In-State Tuition” was the only data journalism piece in which users spent more time on page than those who accessed the related news stories. The other 9 interactive pieces reported less time on the page than at least 1 of the related articles. The users spent about 5 times more time on related articles of “Ethics Explorer” than on this data visualization.

In half of data journalism pieces, the bounce rate is higher than in at least 1 related news story. Which means that on these interactive pieces the visitors entered the website through these data visualizations and left the website without visiting any other *The Texas Tribune’s* page in a proportion higher than in the respective related news. “Interactive: Undocumented Students on In-State Tuition”, “Ethics Explorer”, “Faces of Death Row”, “Texas Hospitals Explorer”, and “Interactive: Demographics of Texas’ Undocumented Population” had a bounce rate lower than their related stories.

The exit rates are quite similar. The same 5 interactive pieces showed a lower proportion than their related news stories: “Interactive: Undocumented Students on In-State Tuition”, “Ethics Explorer”, “Faces of Death Row”, “Texas Hospitals Explorer”,

and “Interactive: Demographics of Texas’ Undocumented Population”. That is, users who visited these data journalism pieces then visited other pages of *The Texas Tribune’s* website in a greater proportion than those who visited the related articles.

The desktop was the most used device in 9 of the interactive pieces. Mobile was the device more frequently used in “See Vaccine Exemptions in Texas by School District”, 46%. “Texas Hospitals Explorer” recorded the highest proportion of visitors that used tablet: 15%. Fourteen of the 20 related news stories recorded more than 50% of their traffic coming from the desktops.

Google is the place where most users were before visiting 7 of the interactive pieces in 2015. Only “Texas Hospitals Explorer”, “Interactive: How Many Texans Resemble You?”, and “See Vaccine Exemptions in Texas by School District” recorded more visitors arriving from another location. Fifteen of the 20 related news stories also registered most of the users coming from Google.

About 22% of “Faces of Death Row” visitors in 2015 were from outside the United States. None of the interactive pieces reported a 100% U.S. audience. 17 of the 20 related news articles recorded more than 90% of users accessing from the United States.

Google Chrome was the browser most used by visitors of all 10 data journalism pieces in 2015. Eighteen of the 20 related news stories were also accessed mostly by visitors using Chrome.

4.6. Data journalism pieces from 2016 and the average of the 500 most read articles

Table 6 shows some information about the chosen data journalism pieces from 2016: the title of the piece, up to 3 interaction techniques, the main topic and the link to the article. In the Appendix there are all the interaction techniques and all topics of each interactive piece, as well as a screenshot of the webpage.

As in 2014 and 2015, politics/governance was once again the most common topic among the data journalism pieces, present in half of the articles, followed by economy/business, theme of 3 interactive pieces. Filter was the most used interaction

technique, introduced in half of the data visualizations, followed by inspect, present in 4 interactive pieces.

Table 6 – The 10 interactive pieces chosen in 2016 from *The Texas Tribune* and their specific aspects.

Title	Interaction techniques	Topic	URL
Ballpark Figures	Explore.	Sports.	http://bit.ly/2LnmuMk
Nov. 8 general election results	Filter.	Politics/governance.	http://bit.ly/2YeRHDb
Track how many Texans are voting early	Inspect.	Politics/governance.	http://bit.ly/2J4fcLI
2016 Primary Election Results	Select, Filter.	Politics/governance.	http://bit.ly/2JriQyD
See which counties in Texas Trump and Clinton won	Inspect.	Politics/governance.	http://bit.ly/2J1We8e
The Price of Admission	Inspect, Filter, Explore.	Education.	http://bit.ly/2J31fgY
Why Isn't Texas Ready for the Next Big Hurricane?	Narrate/History, Explore, Select.	Environment / science/technology.	http://bit.ly/2IYKrYg
Can you tell real news from fake news? Take our quiz to find out	Gamification, Collaboration, Narrate/History.	Entertainment.	http://bit.ly/2VQzE8s
Bordering on Insecurity	Select, Filter, Inspect.	Social issues.	http://bit.ly/2JjOBcL
Who Said It: Donald Trump or Mark Cuban?	Gamification, Collaboration, Narrate/History.	Entertainment.	http://bit.ly/2vEU45R

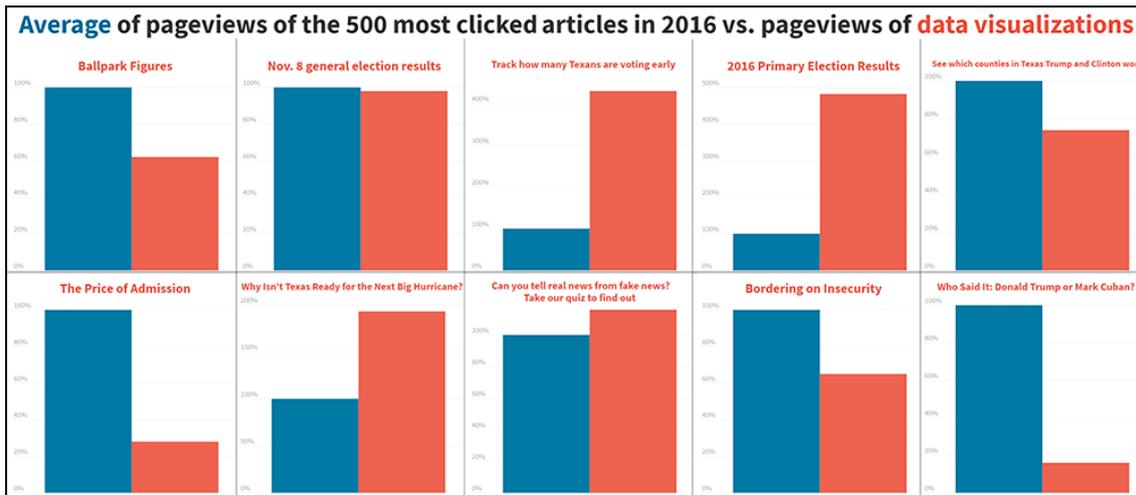


Figure 44 – Performance in pageviews of *The Texas Tribune* interactive pieces compared to the average of the most read articles in 2016 (represented as 100%).⁸⁹

Figure 44 shows pageviews performance of the data journalism pieces compared to the average of the 500 most read news stories in 2016. Four out of 10 data visualizations had more hits than the average of the 500 articles and 2 are among the 15 most read news stories of 2016: “2016 Primary Election Results” and “Track how many Texans are voting early”.

These two interactive pieces reported the best performance: more than 4 times more pageviews than the average of the 500 most clicked stories in 2016. “Who Said It: Donald Trump or Mark Cuban?” registered the worst: about a sixth of the hits of the average of the most read news articles.

The average ratio of pageviews/unique pageviews of the 500 most read news stories in 2016 is 1.3. “Ballpark Figures” and “Track how many Texans are voting early” reported the highest ratio between the data journalism pieces, 2.2. That is, for each unique pageview more than 2 pageviews have been recorded. “Ballpark Figures”, “Nov. 8 general election results” and “2016 Primary Election Results” also registered a proportion higher than the average. The other 6 interactive pieces registered a ratio lower than the average of the 500 most read news articles, and “Can you tell real news from fake news? Take our quiz to find out” and “Who Said It: Donald Trump or Mark Cuban?” registered the lowest ratio, 1.05.

⁸⁹ Available at: <https://iloquiardotcom1.files.wordpress.com/2019/12/djxavg-2016.png>. Accessed on December 19, 2019.

Looking at the ratio of entrances/pageviews, “Can you tell real news from fake news? Take our quiz to find out” is proportionately the interactive piece through which more visitors entered *The Texas Tribune’s* website, 91%. Which indicates that, for 9 out of 10 users who accessed this interactive piece, that was the first page on *The Texas Tribune’s* website they visited during their session. The average proportion of the 500 most read news stories in 2016 is 64%. There are 6 data journalism pieces that reported a percentage lower than 64%, and “Nov. 8 general election results” had the lowest: 33%. That is, about 7 out of 10 visitors of this data visualization had visited another page of *The Texas Tribune* before accessing it.

Half of the interactive pieces registered more average time on page than the 500 most read news articles in 2016. “Nov. 8 general election results” reported the best performance: users who visited this data visualization spent 1.6 more time on page than those who accessed the 500 news stories. “See which counties in Texas Trump and Clinton won” had the worst performance: visitors spent on average about 99% less time on page on this interactive piece than the visitors of the most clicked articles.

The average bounce rate of the 500 most read news articles in 2016 is 68%. Which means that about 7 out of 10 visitors that entered the website through these news stories left *The Texas Tribune’s* website without visiting any other page. Among the data journalism pieces, “Who Said It: Donald Trump or Mark Cuban?” reported the highest value: 93%. There are 3 interactive pieces that registered a lower bounce rate than the 500 news stories, and “Track how many Texans are voting early” showed the lowest, 0.1%.

“See which counties in Texas Trump and Clinton won” registered the lowest exit rate, 2.5%. That is, almost 98% of the users who visited this interactive piece did not leave *The Texas Tribune’s* website from it. The average exit rate from the 500 most read news articles in 2016 is 63%. Half of the interactive pieces reported a percentage higher than 63% and “Can you tell real news from fake news? Take our quiz to find out” recorded the highest exit rate: 92%.

4.7. Data journalism pieces from 2016 and their related stories

Figure 45 shows the performance (in pageviews) of the interactive pieces compared to their respective related news stories. The data journalism pieces are displayed as 100%. That way, related articles that registered a higher number of pageviews than the interactive pieces appear in the figure with more than 100%; and those related news stories that registered fewer hits than the data visualization appear in the figure with less than 100%.

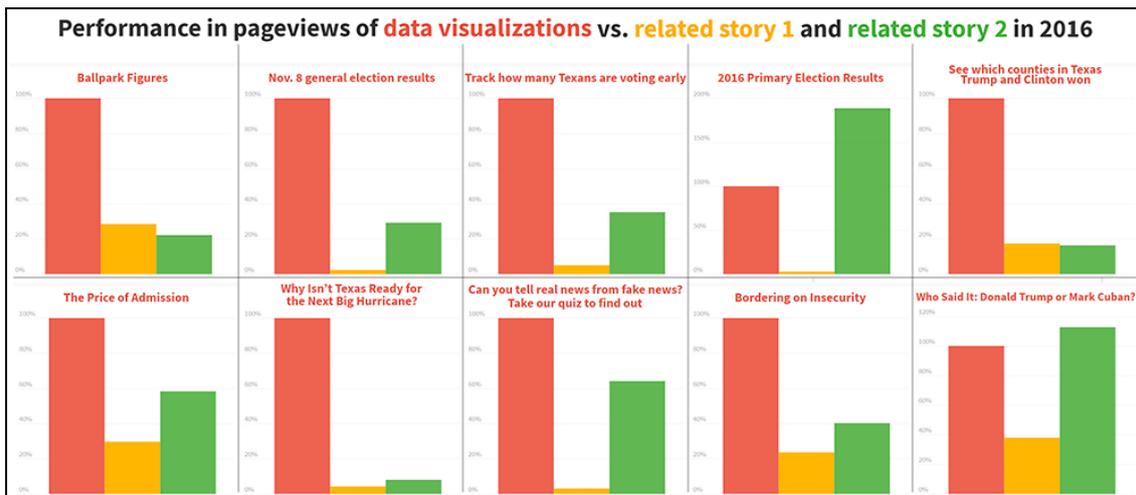


Figure 45 – Performance in pageviews of *The Texas Tribune* interactive news pieces (represented as 100%) compared to their related news stories in 2016.⁹⁰

Eight out of 10 data journalism pieces had more clicks than their respective related news stories. “Why Isn’t Texas Ready for the Next Big Hurricane?” is the interactive piece that performed best: 12 times more pageviews than the related story with better performance. “2016 Primary Election Results” and “Who Said It: Donald Trump or Mark Cuban?” are the only interactive pieces that reported fewer hits than at least one of the related stories. Information on related news stories such as title, publication date, and link to the article are available in the Appendix.

⁹⁰ Available at: <https://iloquiardotcom1.files.wordpress.com/2019/12/djxrel-2016.png>. Accessed on December 19, 2019.



Figure 46 – Evolution in pageviews of *The Texas Tribune* 2016 interactive pieces and their related news stories over time.⁹¹

Figure 46 shows the percentage of pageviews in 2017 compared to 2016, represented as 100%. That means we can observe the evolution of hits that the data journalism pieces and their related news stories have registered over the years.

All the 20 related news stories recorded fewer pageviews in 2017 than in the previous year. There are 2 interactive pieces that reported more hits in 2017 than in 2016: “Ballpark Figures” and “Why Isn’t Texas Ready for the Next Big Hurricane?”. “Ballpark Figures” is the data journalism piece that performed best: about twice as many pageviews in 2017 than in 2016. Among the 10 interactive pieces, “Track how many Texans are voting early” had the worst performance in 2017: less than 1% of the number of hits it registered in the year of its publication.

All the related news stories in 2016 reported a ratio of pageviews/unique pageviews below 1.3, the average ratio of the 500 most read news articles that year. “The Price of Admission”, “Can you tell real news from fake news? Take our quiz to find out”, and “Who Said It: Donald Trump or Mark Cuban?” registered a proportion lower than at least 1 of the related articles. That is, in the other 7 data journalism pieces the visitors returned to the page during the same session more times than those who accessed their related news stories.

⁹¹ Available at: <https://iloaguardotcom1.files.wordpress.com/2019/12/djxrel-2016-17.png>. Accessed on December 19, 2019.

Between 2016 and 2017, “Nov. 8 general election results”, “Track how many Texans are voting early”, and “Bordering on Insecurity” are the interactive pieces with returning visitors rate higher than their related news stories. Half of the data journalism pieces, in that period, the proportion of new visitors was higher than their related news stories: “See which counties in Texas Trump and Clinton won”, “The Price of Admission”, “Why Isn’t Texas Ready for the Next Big Hurricane?”, “Can you tell real news from fake news? Take our quiz to find out”, and “Who Said It: Donald Trump or Mark Cuban?”.

The ratio of entrances/pageviews is higher in 3 data journalism pieces than in their related news articles: “Why Isn’t Texas Ready for the Next Big Hurricane?”, “Can you tell real news from fake news? Take our quiz to find out”, and “Who Said It: Donald Trump or Mark Cuban?”. The other 7 interactive pieces presented a proportion lower than at least 1 of the related stories, which implies that the visitors to these data visualizations visited another page of *The Texas Tribune’s* website before accessing it more times than the visitors of the related articles.

“Nov. 8 general election results” and “2016 Primary Election Results” were the data visualizations in which users spent more time on page than those who accessed the related news stories. The other 8 interactive pieces reported less time on page than at least 1 of the related news articles. The users spent about 70 times more time on related articles of “See which counties in Texas Trump and Clinton won” than on this interactive piece.

There are 4 data journalism pieces where the bounce rate is higher than in at least 1 related news story, which means that on these interactive pieces the visitors entered the website through these data visualizations and left the website without visiting any other *The Texas Tribune’s* page in a proportion higher than in the respective related news. “Ballpark Figures”, “Track how many Texans are voting early”, “2016 Primary Election Results”, “See which counties in Texas Trump and Clinton won”, “The Price of Admission”, and “Bordering on Insecurity” registered a bounce rate lower than their related articles.

The exit rates are similar. Here 5 interactive pieces had a lower proportion than their related news stories: “Ballpark Figures”, “Track how many Texans are voting

early”, “2016 Primary Election Results”, “See which counties in Texas Trump and Clinton won”, and “Bordering on Insecurity”. That is, users who visited these data journalism pieces then visited other pages of *The Texas Tribune’s* website in a greater proportion than those who visited the related articles. The other 5 data visualizations had higher exit rate than their related news.

The mobile was the most used device in 6 of the interactive pieces: “2016 Primary Election Results”, “See which counties in Texas Trump and Clinton won”, “The Price of Admission”, “Why Isn’t Texas Ready for the Next Big Hurricane?”, “Can you tell real news from fake news? Take our quiz to find out”, and “Who Said It: Donald Trump or Mark Cuban?”. “Ballpark Figures” registered the highest proportion of visitors that used tablet: 13%. Half of the related news stories recorded more than 50% of their traffic coming from the desktops.

In 2016, most users were at Google and Facebook before visiting interactive pieces, 3 each. “Ballpark Figures”, “Can you tell real news from fake news? Take our quiz to find out”, “Bordering on Insecurity”, and “Who Said It: Donald Trump or Mark Cuban?” registered more visitors arriving from another place. Thirteen of the 20 related news stories reported most of the users coming from Google.

About 13% of “Who Said It: Donald Trump or Mark Cuban?” visitors in 2016 were from outside the United States. None of the interactive pieces registered a 100% U.S. audience. Four of the 20 related news articles registered 100% of the visitors from the United States and other 14, more than 90% of users accessing from the U.S.

Finally, Google Chrome was the browser most used by visitors of 6 of the data visualizations in 2016. “See which counties in Texas Trump and Clinton won”, “The Price of Admission”, “Why Isn’t Texas Ready for the Next Big Hurricane?”, and “Can you tell real news from fake news? Take our quiz to find out” registered Safari as the most used browser. Eleven of the 20 related news stories were also accessed mostly by visitors using Google Chrome.

V – *O Globo*

5.1. History

The Brazilian newspaper *O Globo* was first published on July 29, 1925. Founder Irineu Marinho died a few days later and his eldest son, Roberto Marinho, took control in 1931, after the death of the editor-in-chief Eurycles de Mattos (Carvalho, 2012). In the decade of 1950, the newspaper positioned itself against President Getúlio Vargas (1951-1954) and in 1964 supported the military coup against President João Goulart (Magnolo & Pereira, 2016). In 2013 the newspaper “acknowledged that, in the light of history, this support was a mistake” (O Globo, 2013).⁹²

On July 29, 1996, when the newspaper completed 71 years, *O Globo’s* website was launched. In mid-2017, occurred the integration of the editorial rooms of the newspapers *O Globo*, *Extra*, *Expresso* and the weekly magazine *Época*. As part of the process, the outlets fired more than 30 journalists (I. Ribeiro, 2018). During 2018 *O Globo’s* website underwent a wide change, presenting a new layout, and among the main innovations was the production of exclusive content for the subscribers (O Globo, 2018). The newspaper ended 2018 as the largest newspaper in circulation in Brazil (Rosa, 2019).

In 2014 it launched the blog “Na Base dos Dados”, focused on data journalism.⁹³ In 2016 the interactive article “DNA do Congresso Nacional”⁹⁴ won the *Grupo de Diários de América* (GDA) award in the Data Journalism category (Briso, 2016). Currently there is a team dedicated to fact-checking and data journalism in the newspaper (Moura, 2018).

⁹² In the original: “reconhece que, à luz da História, esse apoio foi um erro.”

⁹³ Available at: <https://blogs.oglobo.globo.com/na-base-dos-dados/>. Accessed on May 25, 2019.

⁹⁴ Available at: <https://infograficos.oglobo.globo.com/brasil/dna-do-congresso.html>. Accessed on May 25, 2019.

5.2. Data journalism pieces from 2014 and the average of the 500 most read articles

Unfortunately, the Google Analytics data we had access to from 2014 was divided into 2 groups: (1) hits made from a desktop; (2) accesses made from a mobile device. Thus, the analysis of 2014 was carried out only with data from accesses made from a desktop.

Table 7 shows the title of the interactive pieces, up to 3 interaction techniques, the main topic and the link to the article. In the Appendix there are all the interaction techniques and topics of each piece, as well as a screenshot of the webpage.

Table 7 – The 10 data journalism pieces chosen in 2014 from *O Globo* and their specific aspects.

Title	Interaction techniques	Topic	URL
Apuração das eleições 2014 – 1º Turno	Inspect, Filter, Explore.	Politics/governance.	https://glo.bo/2VPhUKN
O peso eleitoral dos estados	Select, Inspect.	Politics/governance.	http://bit.ly/2Y92vT2
As divisões socioeconômicas da votação para presidente	Inspect, Filter, Abstract/Elaborate.	Politics/governance.	http://bit.ly/2vCxn2e
Evolução das bancadas	Narrate/History, Inspect, Select.	Politics/governance.	http://bit.ly/2LnuLzX
Os números da Pnad 2013	Inspect, Select, Abstract/Elaborate.	Social issues.	http://bit.ly/2H3Vjlp
A seleção dos sonhos da Copa do Mundo 2014	Gamification, Collaboration, Narrate/History.	Sports.	http://bit.ly/2V4O1lp
O Brasil nos Mundiais	Inspect, Select, Filter.	Sports.	http://bit.ly/2Vktmi4
Você sabe qual a sua expectativa de vida?	Gamification.	Social issues.	http://bit.ly/2VktJsY
Mapa da taxa de suicídio no mundo	Inspect, Select, Filter.	Social issues.	http://bit.ly/2H0bLSD
Campanha versus realidade	Inspect, Select, Filter.	Politics/governance.	http://bit.ly/2YeTZ5f

Politics/governance was the most recurrent topic, present in half of the interactive pieces, followed by social issues, the subject matter of 40% of the data journalism pieces. Inspect and select were the most used interaction techniques, each one introduced in 80% of the data visualizations, followed by filter, present in 7 of the interactive articles.

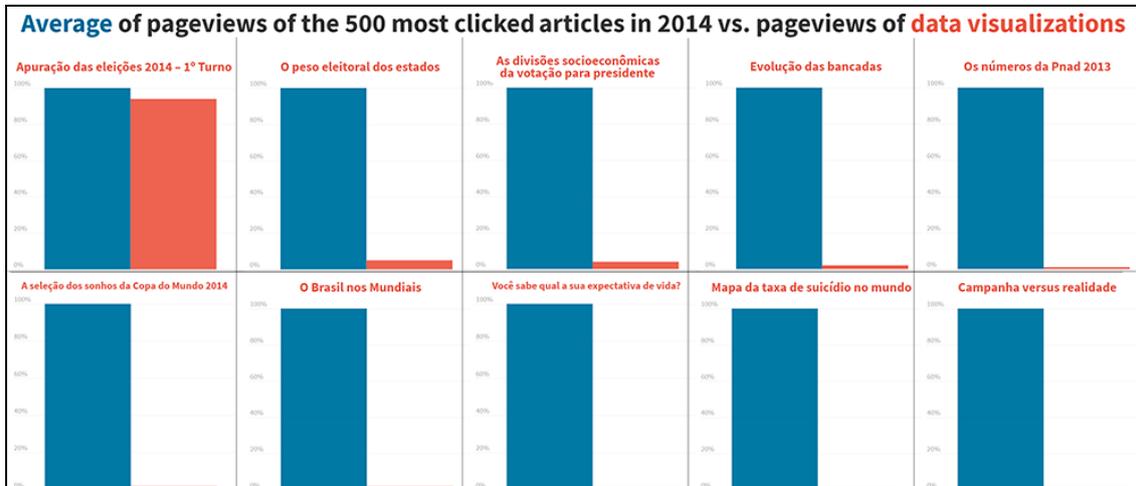


Figure 47 – Performance in pageviews of *O Globo* interactive news pieces compared to the average of the most read articles in 2014 (represented as 100%).⁹⁵

Figure 47 shows pageviews performance of the data journalism pieces compared to the average of the 500 most read stories in 2014. All 10 interactive pieces had fewer hits than the average of the 500 articles. “Apuração das eleições 2014 – 1º Turno” registered the best performance: 6% fewer pageviews than the average of the 500 most read news stories. “Mapa da taxa de suicídio no mundo” reported the worst: fewer than 1% of the average hits of the 500 articles.

The ratio of pageviews/unique pageviews shows how many times a page was accessed during users’ sessions, i.e. it is possible to check if a news article was viewed multiple times by the same user. GOV.UK website considers a high ratio above 1.4.⁹⁶

The average ratio pageviews/unique pageviews of the 500 most clicked stories in 2014 is 1.1. “Apuração das eleições 2014 – 1º Turno” registered the highest ratio between the data journalism pieces, 1.43. All the other 9 interactive pieces reported a

⁹⁵ Available at: <https://iloquiardotcom1.files.wordpress.com/2019/12/djxavg-2014-1.png>. Accessed on December 19, 2019.

⁹⁶ Available at: <https://gov.uk/guidance/content-design/data-and-analytics>. Accessed on May 25, 2019.

proportion below 1.4. Nevertheless, all data visualization had a ratio higher than the average of the 500 most read articles in 2014.

Looking at the ratio of entrances/pageviews, “As divisões socioeconômicas da votação para presidente” is proportionately the data journalism piece through which more visitors entered *O Globo*’s website, 55%. The average proportion of the 500 most read news stories in 2014 is 73%. All 10 interactive pieces recorded a percentage lower than 73% and “O peso eleitoral dos estados” registered the lowest: 7%. This indicates that more than 9 out of 10 visitors of this interactive piece had visited another page of *O Globo* before accessing it.

On average, users who visited “Apuração das eleições 2014 – 1º Turno” spent on this data journalism piece 50% more time than those who accessed the 500 most clicked news articles in 2014. On the other hand, all the other 9 interactive pieces registered visits that lasted less than the average of the most read stories. Most notable is “O peso eleitoral dos estados”, where users spent on average 56% less time than the visitors of the most clicked news.

The average bounce rate of the 500 most read news articles in 2014 is 72%. This means that about 7 out of 10 visitors that entered the website through these news stories left *O Globo*’s website without visiting any other page. Among the data journalism pieces, “As divisões socioeconômicas da votação para presidente” reported the highest value: 85%. Only 2 interactive pieces registered a lower bounce rate than the 500 news stories, and “O peso eleitoral dos estados” had the lowest, 61%.

“O peso eleitoral dos estados” registered the lowest exit rate, 30%. That is, 3 in 10 users who visited this data journalism piece left *O Globo*’s website from it. The average exit rate from the 500 most read news articles in 2014 is 70%. There are 9 interactive pieces that reported a percentage lower than 70%, only “As divisões socioeconômicas da votação para presidente” also registered an exit rate of 70%.

5.3. Data journalism pieces from 2014 and their related stories

Figure 48 shows the performance (in pageviews) of the interactive pieces compared to their respective related news stories. The data journalism pieces are displayed as 100%. That way, related articles that achieved a higher number of pageviews than the data visualization appear in the figure with more than 100%; and those related news stories that registered fewer hits than the interactive pieces appear in the figure with less than 100%.

Half of the data journalism pieces had more clicks than their respective news stories. “Apuração das eleições 2014 – 1º Turno” is the interactive piece that performed best: 27 times more pageviews than the related story with better performance. “Evolução das bancadas” is the only interactive piece that reported fewer hits than both related stories. Information on related news stories such as title, publication date, and link to the article are available in the Appendix.

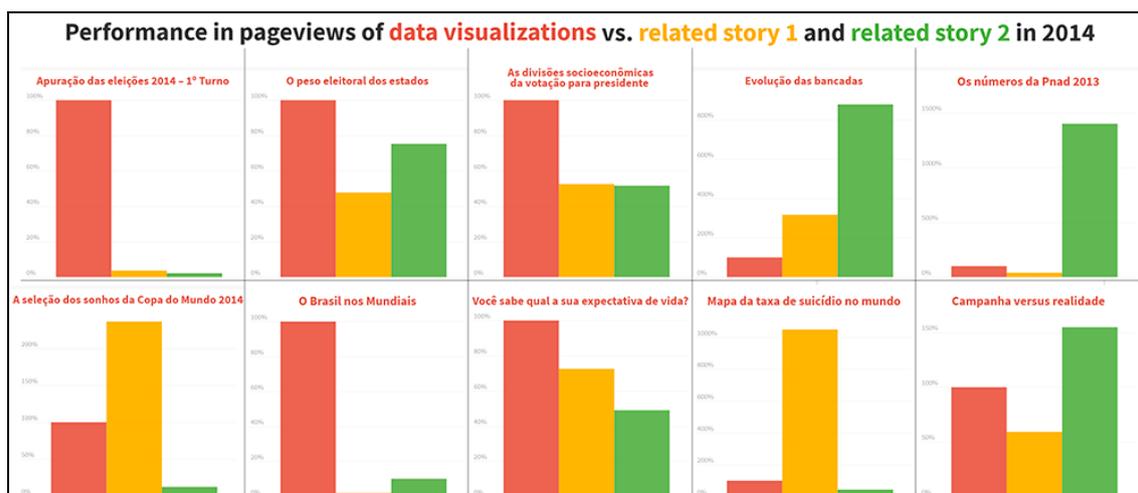


Figure 48 – Performance in pageviews of *O Globo* interactive pieces (represented as 100%) compared to their related news stories in 2014.⁹⁷

⁹⁷ Available at: <https://iloaguiardotcom1.files.wordpress.com/2019/12/djxrel-2014-1.png>. Accessed on December 19, 2019.

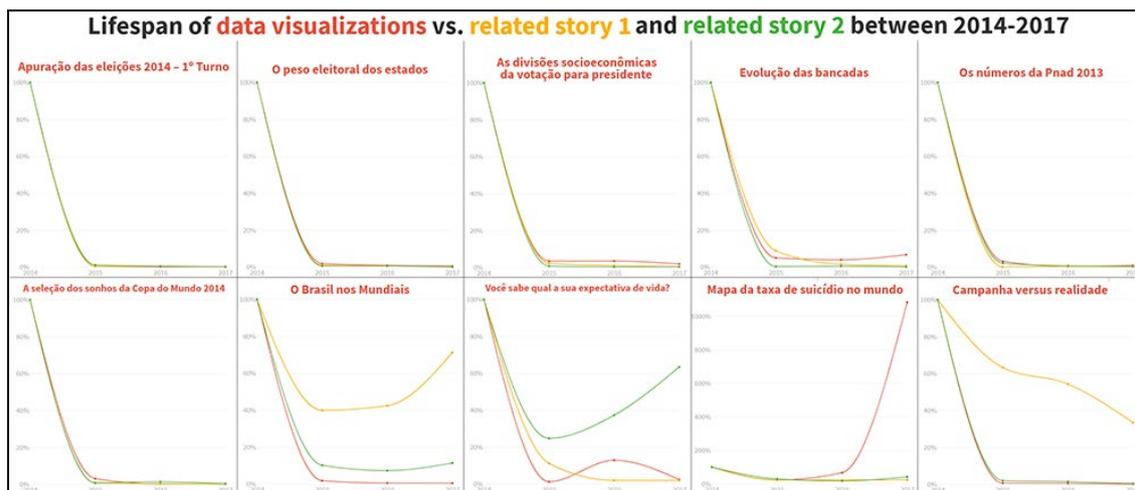


Figure 49 – Evolution in pageviews of *O Globo* 2014 interactive pieces and their related news stories over time.⁹⁸

With the related news stories, it is also possible to look at the lifetime of the articles. Figure 49 shows the percentage of pageviews in 2015, 2016 and 2017 compared to 2014, represented as 100%. That is, we can observe the evolution of hits that the data journalism pieces and their related news stories have registered over the years.

All the related news stories reported fewer pageviews in 2015 than in the previous year, which is the normal behavior of news articles. Nevertheless, 1 interactive piece registered more hits in 2017 than in 2014: “Mapa da taxa de suicídio no mundo”. This data journalism piece registered about 11 times more pageviews in 2017 than in the year of its publication. Among the 10 interactive pieces, “Campanha versus realidade” had the worst performance in 2017: less than 0.1% of the number of hits it registered in 2014.

Looking at the ratio of pageviews/unique pageviews in 2014, only 3 of the 20 related news stories reported a number below 1.1, the average ratio of the 500 most clicked news articles that year. “As divisões socioeconômicas da votação para presidente”, “Evolução das bancadas”, “Os números da Pnad 2013”, and “O Brasil nos Mundiais” had a proportion lower than at least 1 of the related articles. That is, in the other 7 data journalism pieces, the visitors returned to the page during the same session more times than those who accessed their related news stories.

⁹⁸ Available at: <https://iloaguardotcom1.files.wordpress.com/2019/12/djxrel-2014-17-1.png>. Accessed on December 19, 2019.

However, over the years the number of returning visitors is lower on most interactive pieces compared to related news. According to Google Analytics, the user that visited the website within the past two years and returns from the same device is marked as a returning visitor. Between 2014 and 2017, “Apuração das eleições 2014 – 1º Turno”, “Evolução das bancadas”, and “Mapa da taxa de suicídio no mundo” are the data journalism pieces with returning visitors rate higher than their related news stories. That is, the other 7 interactive pieces received a higher proportion of new visitors than their related news stories.

The ratio of entrances/pageviews is higher in only 1 data journalism piece than in their related news articles: “Você sabe qual a sua expectativa de vida?”. That is, the proportion of users who entered the website through this interactive piece was higher than those who came to *O Globo* through their related news. In the other 9 interactive pieces the rate is lower than at least 1 of the related stories, which indicates that the visitors to these data visualizations visited another page of *Globo’s* website before accessing it more times than the visitors of the related articles.

Users who visited “Apuração das eleições 2014 – 1º Turno” spent on this data journalism piece about twice the time of those who accessed the related news stories. “As divisões socioeconômicas da votação para presidente”, “Os números da Pnad 2013”, “A seleção dos sonhos da Copa do Mundo 2014” and “Campanha versus realidade” also registered more average time on the page than at least 1 of the related articles. The users spent about 70% more time on related news of “O peso eleitoral dos estados” than on this interactive piece.

In half of data journalism pieces, the bounce rate is higher than in at least 1 related news story, which means that on these interactive pieces visitors entered the website through these data visualizations and left the website without visiting any other *O Globo’s* page in a proportion higher than in the respective related news. “Apuração das eleições 2014 – 1º Turno”, “O peso eleitoral dos estados”, “O Brasil nos Mundiais”, “Mapa da taxa de suicídio no mundo”, and “Campanha versus realidade” registered a bounce rate lower than their related stories.

“As divisões socioeconômicas da votação para presidente” and “Você sabe qual a sua expectativa de vida?” are the only data journalism pieces that reported an exit

rate higher than at least 1 related story. The other 8 interactive pieces had a lower exit rate than their related articles. That is, users who visited these 8 data visualizations then visited other pages of *O Globo's* website in a greater proportion than those who visited the related news.

In 2014 we analyzed *O Globo's* data only from the desktops, so it is not possible to measure access by devices.

Google is the place where most users were before visiting half of the interactive pieces in 2014: “O peso eleitoral dos estados”, “Evolução das bancadas”, “A seleção dos sonhos da Copa do Mundo 2014”, “O Brasil nos Mundiais”, and “Mapa da taxa de suicídio no mundo”. Globo and Facebook appear on second place as path for 3 data journalism pieces each. Fourteen of the 20 related news stories also registered most of the users coming from Google.

About half of “A seleção dos sonhos da Copa do Mundo 2014” visitors in 2014 were from outside Brazil. Among the other interactive pieces, “Os números da Pnad 2013”, “Você sabe qual a sua expectativa de vida?”, and “Mapa da taxa de suicídio no mundo” recorded a 100% Brazilian audience. Seven of the 20 related news stories reported 100% of the visitors from Brazil and other 9, more than 90% of users accessing from there.

Google Chrome was the browser most used by visitors of 9 of the data journalism pieces in 2014. Only “Mapa da taxa de suicídio no mundo” registered another browser as the most used. Nineteen of the 20 related news were also accessed mostly by visitors using Chrome.

5.4. Data journalism pieces from 2015 and the average of the 500 most read articles

Table 8 shows some information about the chosen data journalism pieces from 2015: the title of the piece, up to 3 interaction techniques, the main topic and the link to the article. In the Appendix there are all the interaction techniques and all topics of each interactive piece, as well as a screenshot of the webpage.

Politics/governance and foreign affairs were the most frequent topics, each one present in 3 interactive pieces, followed by economy/business and social issues, both

topics the subject of 2 data visualizations each one. Like in 2014, inspect was once again the most used interaction technique, introduced in 80% of the data journalism pieces, followed by select, present in 7 of the interactive pieces.

Figure 50 shows pageviews performance of the data journalism pieces compared to the average of the 500 most read news articles in 2015. All 10 interactive pieces had fewer hits than the average of the 500 most clicked news stories. “Consulte a média da sua escola no Enem 2014” registered the best performance: 83% fewer pageviews than the average of the 500 most clicked stories. “Raio-x dos atrasos dos voos no Brasil” and “Os medalhistas brasileiros” reported the worst, fewer than 1% of the hits of the average of the 500 news articles.

Table 8 – The 10 data journalism pieces chosen in 2015 from *O Globo* and their specific aspects.

Title	Interaction techniques	Topic	URL
DNA do Congresso Nacional	Inspect, Select, Reconfigure/Encode.	Politics/governance.	http://bit.ly/2vDXs0V
A distribuição dos clientes e valores do “Swissleaks”	Inspect, Select, Abstract/Elaborate.	Economy/business.	http://bit.ly/2DSUXMz
Ataques terroristas no mundo desde 1970	Inspect, Select, Connect.	Social issues.	http://bit.ly/2V8PyHe
Raio-x dos atrasos dos voos no Brasil	Inspect, Select, Filter.	Economy/business.	http://bit.ly/2VLs2nH
Qual o futuro do PT?	Select, Inspect.	Politics/governance.	http://bit.ly/2H0HxyZ
Consulte a média da sua escola no Enem 2014	Filter, Explore, Reconfigure/Encode.	Education.	http://bit.ly/2VENriH
Os medalhistas brasileiros	Gamification, Inspect, Select.	Sports.	http://bit.ly/2HdtyGf
Duelo entre os lados sombrio e luminoso da Força	Gamification, Collaboration.	Entertainment.	http://bit.ly/2H6q21y
Ranking da liberdade de imprensa no mundo em 2015	Inspect, Select, Abstract/Elaborate.	Social issues.	http://bit.ly/2vGuBco
Personagens da Lava-Jato	Inspect, Filter.	Politics/governance.	http://bit.ly/2VKOPIE

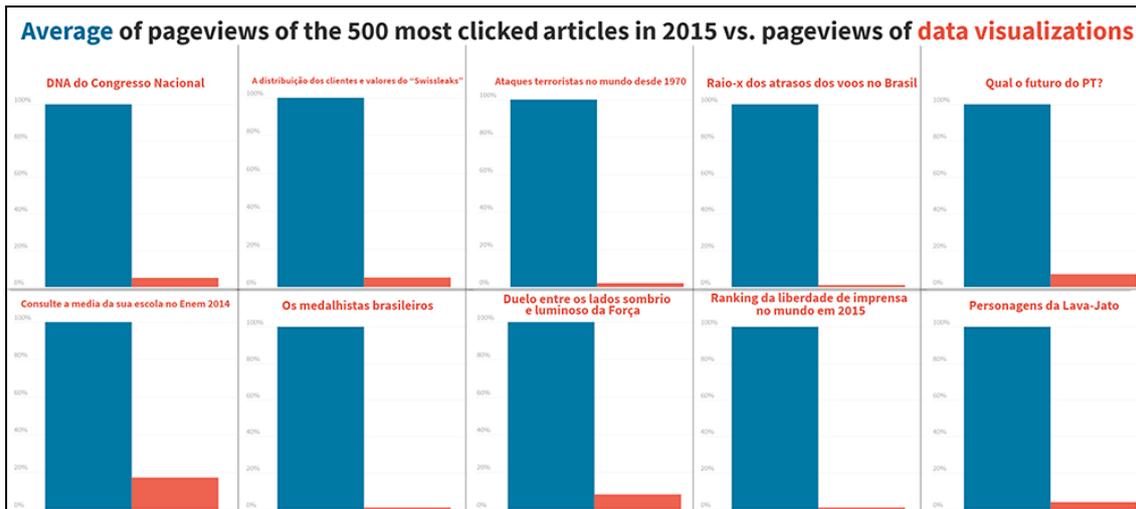


Figure 50 – Performance in pageviews of *O Globo* interactive pieces compared to the average of the most read articles in 2015 (represented as 100%).⁹⁹

As in 2014, the average ratio of pageviews/unique pageviews of the 500 most clicked stories in 2015 is 1.1. “Duelo entre os lados sombrio e luminoso da Força” reported the highest ratio between the data journalism pieces, 1.86. Other 7 interactive pieces had a proportion higher than the average of the 500 most read news articles. “Qual o futuro do PT?” registered the lowest ratio, 1.07.

Looking at the ratio entrances/pageviews, “Ranking da liberdade de imprensa no mundo em 2015” is proportionately the data journalism piece through which more visitors entered *O Globo*’s website, 61%. The average percentage of the 500 most read news stories in 2015 is 80%. As in the previous year, in 2015 none of the 10 interactive pieces had a percentage higher than the average of the 500 most read articles and “Os medalhistas brasileiros” reported the lowest: 10%. This indicates that 9 out of 10 visitors of this interactive piece had visited another page of *O Globo* before accessing it.

On average, users who visited “Ranking da liberdade de imprensa no mundo em 2015” spent on this data journalism piece 42% more time than those who accessed the 500 most clicked news articles in 2015. There are 6 interactive pieces with visits that lasted less than the average of the most read stories. Most notable is “A

⁹⁹ Available at: <https://iloquiardotcom1.files.wordpress.com/2019/12/djxavg-2015-1.png>. Accessed on December 19, 2019.

distribuição dos clientes e valores do ‘Swissleaks’”, where users spent on average 43% less time on page than the visitors of the most clicked news.

The average bounce rate of the 500 most read news articles in 2015 is 80%. This means that 8 out of 10 visitors that landed the website from these news stories left *O Globo’s* website without visiting any other page. Among the data journalism pieces, “Ataques terroristas no mundo desde 1970” presented the highest value: 88%. Half of the interactive pieces reported a lower bounce rate than the 500 news stories, and “Duelo entre os lados sombrio e luminoso da Força” registered the lowest, 61%.

“Ranking da liberdade de imprensa no mundo em 2015” reported the highest exit rate, 73%. That is, about 7 in 10 users who visited this data journalism piece left *O Globo’s* website from it. The average exit rate from the 500 most read news articles in 2015 is 76%. All the 10 interactive pieces registered a percentage lower than 76%, and “Os medalhistas brasileiros” registered the lowest exit rate, 31%.

5.5. Data journalism pieces from 2015 and their related stories

Unfortunately, there is no data available in *O Globo’s* Google Analytics for 6 related news and 1 data journalism piece – “Os medalhistas brasileiros” – on the following metrics: traffic sources, devices, location, new vs. returning visitors, and browsers. There is still information about pageviews, unique pageviews, average time on page, entrances, bounce rate, and exit rate concerning all 20 related news articles and all 10 interactive pieces.

Figure 51 shows the performance (in pageviews) of the interactive pieces compared to their respective related news stories. The data visualizations are displayed as 100%. That way, related articles that showed a higher number of pageviews than the data journalism pieces appear in the figure with more than 100%; and those related news stories that registered fewer hits than the interactive pieces appear in the figure with less than 100%.

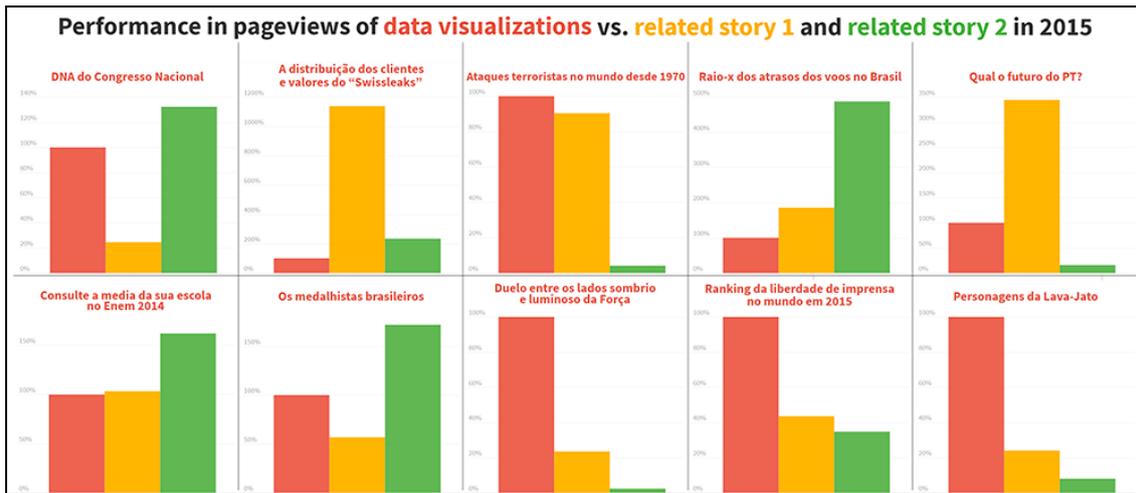


Figure 51 – Performance in pageviews of *O Globo* interactive pieces (represented as 100%) compared to their related news stories in 2015.¹⁰⁰

Four out of 10 data journalism pieces registered more clicks than their respective news stories. “Duelo entre os lados sombrio e luminoso da Força” and “Personagens da Lava-Jato” are the interactive pieces that performed best: both registered 4 times more pageviews than their related story with better performance. “A distribuição dos clientes e valores do ‘Swissleaks’”, “Raio-x dos atrasos dos voos no Brasil”, and “Consulte a média da sua escola no Enem 2014” are the data journalism pieces that reported fewer hits than their related stories. Information on related news stories such as title, publication date, and link to the article are available in the Appendix.

¹⁰⁰ Available at: <https://iloaguiardotcom1.files.wordpress.com/2019/12/djxrel-2015-1.png>. Accessed on December 19, 2019.

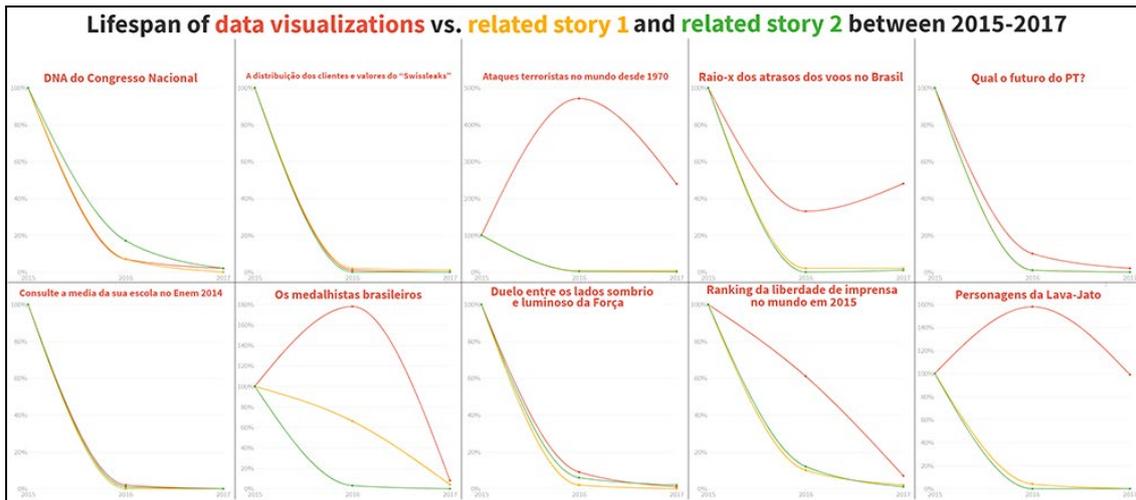


Figure 52 – Evolution in pageviews of *O Globo* 2015 interactive pieces and their related news stories over time.¹⁰¹

Figure 52 shows the percentage of pageviews in 2016 and 2017 compared to 2015, represented as 100%. That means we can observe the evolution of hits that the data visualizations and their related news stories have registered over the years.

All the 20 related news stories recorded fewer pageviews in 2016 than in the previous year. There are 3 interactive pieces that reported more hits in 2016 than in 2015: “Ataques terroristas no mundo desde 1970”, “Os medalhistas brasileiros”, and “Personagens da Lava-Jato”. “Ataques terroristas no mundo desde 1970” is the data journalism piece that performed best in 2017: about 2.4 times more pageviews than in 2015. Among the 10 interactive pieces, “A distribuição dos clientes e valores do ‘Swissleaks’” and “Consulte a média da sua escola no Enem 2014” had the worst performance in 2017: less than 0.5% of the number of hits it registered in the year of its publication.

Looking at the ratio of pageviews/unique pageviews in 2015, 7 of the 20 related news stories reported a number below 1.1, the average ratio of the 500 most clicked news articles that year. “Ataques terroristas no mundo desde 1970”, “Qual o futuro do PT?”, “Os medalhistas brasileiros”, and “Personagens da Lava-Jato” had a proportion lower than at least 1 of the related articles. That is, in the other 6 interactive pieces, visitors returned to the page during the same session more times than those who accessed their related news stories.

¹⁰¹ Available at: <https://iloaguardotcom1.files.wordpress.com/2019/12/djxrel-2015-17-1.png>. Accessed on December 19, 2019.

Between 2015 and 2017 the rate of returning visitors is higher in 2 of the data journalism pieces compared to their related news: “A distribuição dos clientes e valores do ‘Swissleaks’” and “Duelo entre os lados sombrio e luminoso da Força”. “DNA do Congresso Nacional” and “Raio-x dos atrasos dos voos no Brasil” are the interactive pieces with a proportion of new visitors higher than their related articles in that period.

The ratio of entrances/pageviews is higher in 3 data journalism pieces than in their related news articles: “DNA do Congresso Nacional”, “Raio-x dos atrasos dos voos no Brasil”, and “Ranking da liberdade de imprensa no mundo em 2015”. The other 7 interactive pieces showed a proportion lower than at least 1 of the related news, which implies that the visitors to these data visualizations visited another page of the *O Globo’s* website before accessing it more times than the visitors of the related articles.

Users who visited “Consulte a média da sua escola no Enem 2014” and “Ranking da liberdade de imprensa no mundo em 2015” spent on these data journalism pieces about twice the time of those who accessed their related news stories. Only 1 interactive piece registered less time on the page than its related articles. The users spent about 80% more time on related news of “A distribuição dos clientes e valores do ‘Swissleaks’” than on this data visualization.

There are 7 interactive pieces where the bounce rate is higher than in at least 1 related news article. Which means that on these data journalism pieces the visitors entered the website through these data visualization and left them without visiting any other *O Globo’s* page in a proportion higher than in the respective related news. “A distribuição dos clientes e valores do ‘Swissleaks’”, “Duelo entre os lados sombrio e luminoso da Força” and “Personagens da Lava-Jato” had a bounce rate lower than their related stories.

There are 3 interactive pieces that registered an exit rate lower than their related articles: “A distribuição dos clientes e valores do ‘Swissleaks’”, “Os medalhistas brasileiros” and “Duelo entre os lados sombrio e luminoso da Força”. That is, users who visited these data journalism pieces then visited other pages of *O Globo’s* website in a greater proportion than those who visited the related news stories.

The desktop was the most used device in 7 of the 9 interactive pieces. Mobile was the device more frequently used in “Ataques terroristas no mundo desde 1970” and “Consulte a média da sua escola no Enem 2014”: 60% and 51%, respectively. “DNA do Congresso Nacional” reported the highest proportion of visitors that used tablet: 42%. Ten of the 14 related news stories recorded more than 50% of their traffic coming from the desktops. There are no data in this metric for “Os medalhistas brasileiros” and other 6 related articles.

Globo is the place where most users were before visiting 4 of the interactive pieces in 2015: “A distribuição dos clientes e valores do ‘Swissleaks’”, “Ataques terroristas no mundo desde 1970”, “Raio-x dos atrasos dos voos no Brasil”, and “Personagens da Lava-Jato”. Google and Facebook appear on second place as path for 3 data journalism pieces each. Five of the 14 related news stories also registered most of the users coming from Globo. There are no data in this metric for “Os medalhistas brasileiros” and other 6 related articles.

About 40% of “Ataques terroristas no mundo desde 1970” visitors in 2015 were from outside Brazil. Among the other interactive pieces, “Raio-x dos atrasos dos voos no Brasil”, “Qual o futuro do PT?”, “Consulte a média da sua escola no Enem 2014”, and “Duelo entre os lados sombrio e luminoso da Força” reported a 100% Brazilian audience. Nine of the 14 related news stories registered 100% of the visitors from Brazil and other 3, more than 90% of users accessing from there. There are no data in this metric for “Os medalhistas brasileiros” and other 6 related articles.

Google Chrome was the browser most used by visitors of all 9 data journalism pieces in 2015. Eleven of the 14 related news stories were also accessed mostly by visitors using Google Chrome. There are no data in this metric for “Os medalhistas brasileiros” and other 6 related articles.

5.6. Data journalism pieces from 2016 and the average of the 500 most read articles

Table 9 shows some information about the chosen data journalism pieces from 2016: the title of the piece, up to 3 interaction techniques, the main topic and the link to the article. In the Appendix there are all the interaction techniques and all topics of each interactive piece, as well as a screenshot of the webpage.

Table 9 – The 10 data journalism pieces chosen in 2016 from *O Globo* and their specific aspects.

Title	Interaction techniques	Topic	URL
Placar do Impeachment	Filter, Select, Explore.	Politics/governance.	http://bit.ly/2vEe2xz
O mapa das coligações	Inspect, Select.	Politics/governance.	http://bit.ly/2H6kcM6
A votação no Rio de Janeiro, por zona eleitoral	Inspect, Select, Filter.	Politics/governance.	http://bit.ly/2VMp2rm
Você consegue escalar um time melhor que o Dream Team?	Gamification, Collaboration.	Sports.	http://bit.ly/2H4deIU
Partido do “você não me representa”	Gamification, Collaboration, Narrate/History.	Politics/governance.	http://bit.ly/2WISm53
Calculadora de aposentadoria	Gamification, Collaboration.	Social issues.	http://bit.ly/2H3f24W
A votação para presidente dos EUA	Inspect, Explore, Select.	Foreign affairs.	http://bit.ly/2JjUyqc
Com qual candidato a prefeito do Rio você mais se identifica?	Gamification, Filter.	Politics/governance.	http://bit.ly/2WqjFet
O trajeto da tocha olímpica no Brasil	Inspect, Select, Abstract/Elaborate.	Sports.	http://bit.ly/2GXIIQM
Mapa das manifestações do dia 13 de março	Inspect, Explore, Connect.	Social issues.	http://bit.ly/2LoN6wv

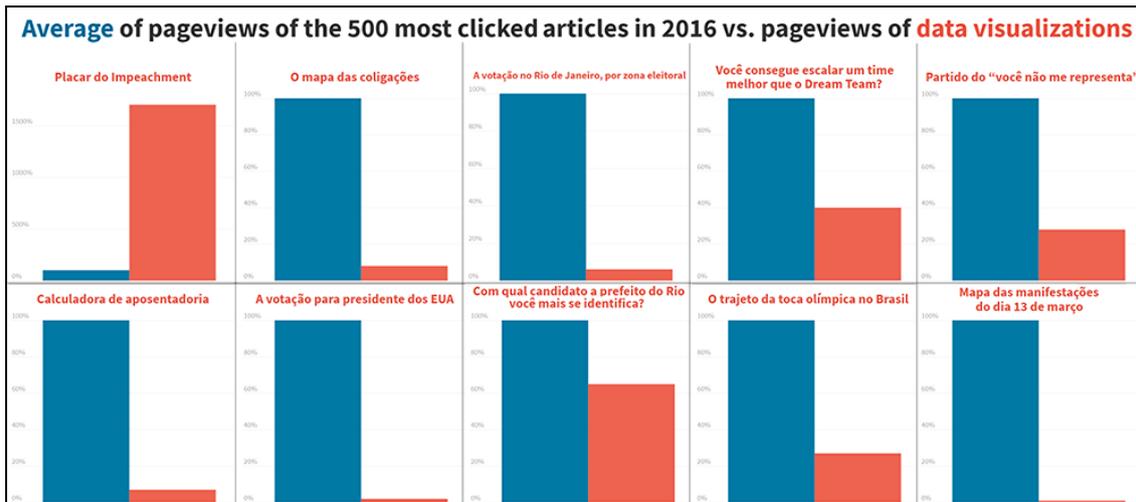


Figure 53 – Performance in pageviews of *O Globo* interactive pieces compared to the average of the most read articles in 2016 (represented as 100%).¹⁰²

As in 2014 and 2015, politics/governance was once again the most common topic among the data journalism pieces, present in 70% of the articles, followed by social issues, theme of 4 interactive pieces. Collaboration was the most used interaction technique, introduced in 60% of the data visualizations, followed by inspect, present in half of the data journalism pieces.

Figure 53 shows pageviews performance of the interactive pieces compared to the average of the 500 most read news stories in 2016. Only 1 data journalism piece had more hits than the average of the 500 articles and is among the 10 most read news stories of 2016: “Placar do Impeachment”.

“Placar do Impeachment” reported the best performance: more than 17 times more pageviews than the average of the 500 most clicked stories in 2016. “Mapa das manifestações do dia 13 de março” registered the worst: 1% of the average hits of the most read news articles.

The average ratio of pageviews/unique pageviews of the 500 most clicked news stories in 2016 is 1.2. “Você consegue escalar um time melhor que o Dream Team?” reported the highest ratio among the data visualizations, 2.8. That is, for each unique pageview more than 2 pageviews have been recorded. Other 6 interactive pieces had a

¹⁰² Available at: <https://iloaguardotcom1.files.wordpress.com/2019/12/djxavg-2016-1.png>. Accessed on December 19, 2019.

proportion higher than the average of the 500 most read news articles. “O mapa das coligações” registered the lowest ratio, 1.08.

Looking at the ratio entrances/pageviews, “O mapa das coligações” is proportionately the data journalism piece through which more visitors entered *O Globo*’s website, 88%. The average proportion of the 500 most read news stories in 2016 is 60%. Half of the interactive pieces reported a ratio lower than 60% and “A votação para presidente dos EUA” had the lowest: 7%. This indicates that more than 9 out of 10 visitors of this interactive piece had visited another page of *O Globo* before accessing it.

On average, users who visited “Placar do Impeachment” spent on this data journalism piece more than twice the time of those who accessed the 500 most read news articles in 2016. There are 3 interactive pieces that registered visits that lasted less than the average of the 500 articles. Most notable is “Você consegue escalar um time melhor que o Dream Team?”, where users spent on average 36% less time than the visitors of the most clicked news stories.

The average bounce rate of the 500 most read news articles in 2016 is 62%. This means that about 6 out of 10 visitors that entered the website through these news stories left *O Globo*’s website without visiting any other page. Among the data journalism pieces, “O mapa das coligações” reported the highest value: 86%. Only 2 interactive pieces registered a lower bounce rate than the 500 news stories, and “A votação para presidente dos EUA” presented the lowest, 51%.

“A votação para presidente dos EUA” also registered the lowest exit rate, 27%. That is, fewer than 3 in 10 users who visited this data journalism piece left *O Globo*’s website from it. The average exit rate from the 500 most read news articles in 2016 is 62%. Half of the interactive pieces reported a percentage higher than 62% and “O mapa das coligações” had the highest exit rate: 83%.

5.7. Data journalism pieces from 2016 and their related stories

Unfortunately, there is no data available in *O Globo*’s Google Analytics for 3 related news on the following metrics: traffic sources, devices, location, new vs. returning visitors, and browsers.

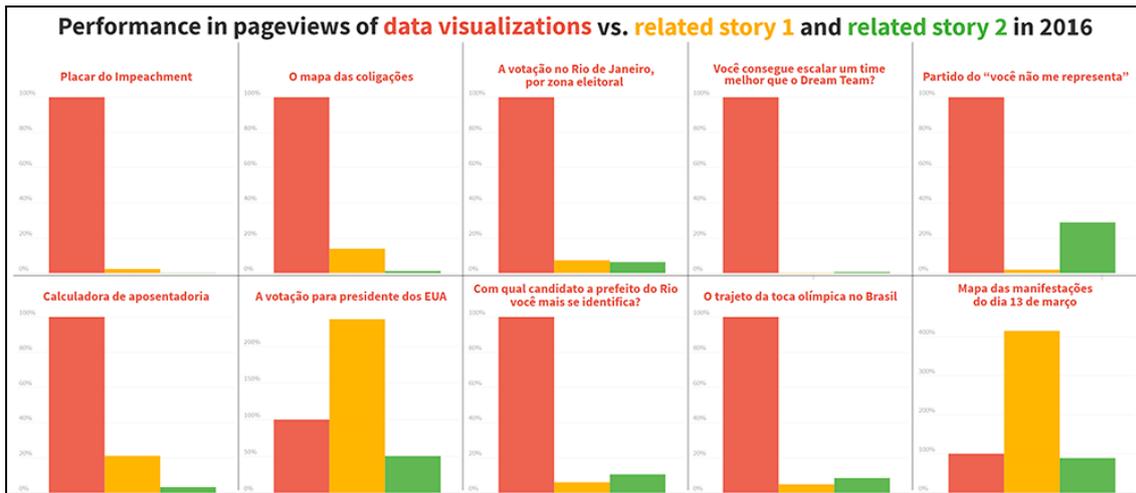


Figure 54 – Performance in pageviews of *O Globo* interactive pieces (represented as 100%) compared to their related news stories in 2016.¹⁰³

Figure 54 shows the performance (in pageviews) of the interactive pieces compared to their respective related news stories. The data journalism pieces are displayed as 100%. That way, related articles that recorded a higher number of pageviews than the data visualizations appear in the figure with more than 100%; and those related news stories that registered fewer hits than the interactive pieces appear in the figure with less than 100%.

Eight out of 10 data journalism pieces had more clicks than their respective news stories. “Você consegue escalar um time melhor que o Dream Team?” is the interactive piece that performed best: 177 times more pageviews than the related story with better performance. “A votação para presidente dos EUA” and “Mapa das manifestações do dia 13 de março” are the data visualization that reported fewer hits than at least 1 of the related stories. Information on related news stories such as title, publication date, and link to the article are available in the Appendix.

Figure 55 shows the percentage of pageviews in 2017 compared to 2016, represented as 100%. That means we can observe the evolution of hits that the data visualizations and their related news stories have registered over the years.

¹⁰³ Available at: <https://iloaguardotcom1.files.wordpress.com/2019/12/djxrel-2016-1.png>. Accessed on December 19, 2019.

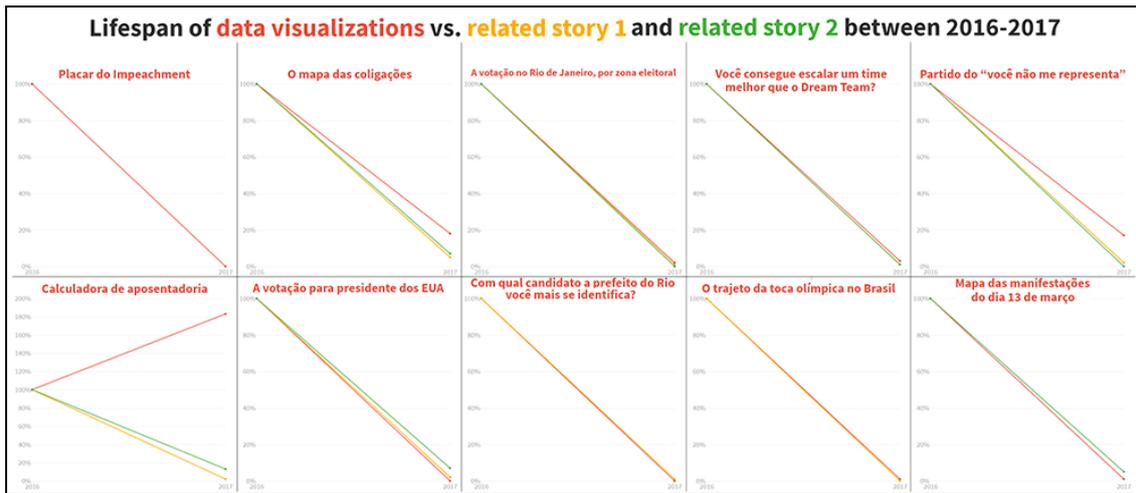


Figure 55 – Evolution in pageviews of *O Globo* 2016 interactive pieces and their related news stories over time.¹⁰⁴

All the 20 related news stories had fewer pageviews in 2017 than in the previous year. There is only 1 data visualization piece that reported more hits in 2017 than in 2016: “Calculadora de aposentadoria”. This interactive piece registered 83% more pageviews in 2017 than in 2016. Among the 10 data journalism pieces, “Placar do Impeachment” had the worst performance in 2017: less than 0.5% of the number of hits it registered in the year of its publication.

Looking at the ratio of pageviews/unique pageviews in 2016, 19 of the 20 related news stories reported a number below the average ratio of the 500 most clicked news articles that year, 1.2. “O mapa das coligações” and “Mapa das manifestações do dia 13 de março” presented a proportion lower than at least 1 of the related articles. That is, in the other 8 data journalism pieces the visitors returned to the page during the same session more times than those who accessed their related news stories.

Between 2016 and 2017, “Placar do Impeachment” and “A votação para presidente dos EUA” are the interactive pieces with returning visitors rate higher than their related news stories. Half of the data journalism pieces reported in that period a proportion of new visitors higher than their related news stories: “O mapa das coligações”, “Partido do ‘você não me representa’”, “Com qual candidato a prefeito do

¹⁰⁴ Available at: <https://iloaguardotcom1.files.wordpress.com/2019/12/djxrel-2016-17-1.png>. Accessed on December 19, 2019.

Rio você mais se identifica?”, “O trajeto da tocha olímpica no Brasil”, and “Mapa das manifestações do dia 13 de março”.

The ratio of entrances/pageviews is higher in half of the data journalism pieces than in their related news articles: “Placar do Impeachment”, “O mapa das coligações”, “A votação no Rio de Janeiro, por zona eleitoral”, “Partido do ‘você não me representa’”, “Com qual candidato a prefeito do Rio você mais se identifica?”, and “O trajeto da tocha olímpica no Brasil”. The other 5 interactive pieces showed a proportion lower than at least 1 of the related news, which implies that the visitors to these data visualization visited another page of *O Globo*’s website before accessing it more times than the visitors of the related articles.

Users who visited “Placar do Impeachment” and “O mapa das coligações” spent on these interactive pieces about twice the time of those who accessed their related news stories. Only 1 data journalism piece registered less time on the page than its related articles. The users spent about 37% more time on related news of “Mapa das manifestações do dia 13 de março” than on this interactive piece.

There are 8 data journalism pieces where the bounce rate is higher than in at least 1 related news article. This means that on these interactive pieces the visitors entered the website through these interactive pieces and left the website without visiting any other *O Globo*’s page in a proportion higher than in the respective related news. “Você consegue escalar um time melhor que o Dream Team?” and “A votação para presidente dos EUA” had lower bounce rate than their related stories.

The exit rates are similar. There are 3 interactive pieces that registered an exit rate lower than their related articles: “Você consegue escalar um time melhor que o Dream Team?”, “A votação para presidente dos EUA”, and “Mapa das manifestações do dia 13 de março”. That is, users who visited these data journalism pieces then visited other pages of *O Globo*’s website in a greater proportion than those who visited the related news stories.

The desktop was the most used device in 6 of the interactive pieces. Mobile was the device more frequently used in the other 4 data journalism pieces: “O mapa das coligações”, “Partido do ‘você não me representa’”, “Com qual candidato a

prefeito do Rio você mais se identifica?”, and “O trajeto da tocha olímpica no Brasil”. “A votação no Rio de Janeiro, por zona eleitoral” had the highest proportion of visitors that used tablet: 9%. Nine of the 17 related news stories had more than 50% of their traffic coming from the desktops. There are no data in this metric for 3 related articles.

Facebook is the place where most users were before visiting 4 of the interactive pieces in 2016: “O mapa das coligações”, “A votação no Rio de Janeiro, por zona eleitoral”, “Partido do ‘você não me representa’”, and “Com qual candidato a prefeito do Rio você mais se identifica?”. Globo and Google appear in second place as path for 3 data visualization each. 8 of the 17 related news stories registered most of the users coming from Globo. There are no data in this metric for 3 related articles.

About 25% of “Mapa das manifestações do dia 13 de março” visitors in 2016 were from outside Brazil. Among the other interactive pieces, “Calculadora de aposentadoria” and “A votação para presidente dos EUA” reported a 100% Brazilian audience. Six of the 17 related news stories registered 100% of the visitors from Brazil and other 5, more than 90% of users accessing from there. There are no data in this metric for 3 related articles.

Google Chrome was the browser most used by visitors of 9 of the data journalism pieces in 2016. Only “O mapa das coligações” registered another browser as the most used, Safari. All 17 related news stories were also accessed mostly by visitors using Google Chrome. There are no data in this metric for 3 related articles.

VI – *Público*

6.1. History

The Portuguese newspaper *Público* was first published on March 5, 1990. Founded and owned by the multinational business group Sonae, then led by Belmiro de Azevedo, the journalist Vicente Jorge Silva was the first editor-in-chief (Nobre, 2015). In its editorial statute, published in the first edition, the newspaper claims to be independent and aware of technological changes.

Público is an information project in tune with the process of technological change and civilization in the contemporary public space. (...) A daily newspaper of great information, guided by criteria of rigor and editorial creativity, without any dependence of ideological, political and economic order. (...) It is inscribed in a European tradition of demanding and quality journalism, refusing the sensationalism and the commercial exploitation of information material. (Público, 1990)¹⁰⁵

On September 22, 1995, *Público*'s website was officially launched, the second general-interest daily newspaper to do so in Portugal. About four years later, the newspaper began to produce special content for the online version and launched the service “Última Hora” (Bastos, 2010a). In 2012, during the restructuring process, 36 journalists were dismissed, about 28% of the newsroom staff (Nobre, 2015). In the first 2 months of 2019, *Público* recorded a rise in the sale of print editions (8%) and online subscriptions (13%) compared to the same period of the previous year (Público, 2019).

Público began investing in data journalism in 2012, when research fellows of the *REACTION* project, under the UT Austin-Portugal program, funded by Fundação de Ciência e Tecnologia, started to collaborate in the production of data-driven stories in

¹⁰⁵ In the original: “Público é um projecto de informação em sintonia com o processo de mudanças tecnológicas e de civilização no espaço público contemporâneo. (...) É um jornal diário de grande informação, orientado por critérios de rigor e criatividade editorial, sem qualquer dependência de ordem ideológica, política e económica. (...) Inscreve-se numa tradição europeia de jornalismo exigente e de qualidade, recusando o sensacionalismo e a exploração mercantil da matéria informativa.”

the newspaper (Alexandre, 2014). In 2014 the hotspot “Floresta em Perigo”,¹⁰⁶ with interactive pieces such as “Doze anos de incêndio”,¹⁰⁷ won several awards, including innovation category (Público, 2014). Currently there is no dedicated team to produce data journalism in *Público*; however, there is a journalist that occasionally produces data journalism pieces. (Moura, 2018).

6.2. Data journalism pieces from 2014 and the average of the 500 most read articles

Table 10 – The 10 data journalism pieces chosen in 2014 from *Público* and their specific aspects.

Title	Interaction techniques	Topic	URL
O ranking das escolas 2014	Inspect, Filter, Explore.	Education.	http://bit.ly/2H74VdU
O ranking dos preços da água em 2013	Explore, Inspect, Select.	Economy/business.	http://bit.ly/2WsgoeR
A vida desde 1820	Select, Explore, Filter.	Social issues.	http://bit.ly/2VjeAIA
Taxa de sobrevivência de cancro em Portugal	Narrate/History, Select, Filter.	Health.	http://bit.ly/2DSMqZW
80 anos de Verão	Inspect, Select, Filter.	Environment / science/technology.	http://bit.ly/2YbMNGM
As linhas da liberdade	Inspect, Select.	Social issues.	http://bit.ly/2V6hJX7
Retrato dos aeroportos nacionais	Select, Inspect.	Economy/business.	http://bit.ly/2Lrnv67
VIH: O vírus que apareceu em Kinshasa em 1920 e alastrou para o mundo inteiro	Select, Narrate/History.	Health.	http://bit.ly/2Lu6CaV
A I Liga vista de dez maneiras diferentes	Filter, Select.	Sports.	http://bit.ly/2VNeLv3
Os salários dos gestores do PSI 20 à lupa	Filter, Select, Narrate/History.	Economy/business.	http://bit.ly/2YeU350

¹⁰⁶ Available at: <http://publico.pt/floresta-em-perigo>. Accessed on May 25, 2019.

¹⁰⁷ Available at: <https://acervo.publico.pt/floresta-em-perigo/doze-anos-de-incendios>. Accessed on May 25, 2019.

Table 10 shows some information about the chosen interactive pieces from 2014: the title of the piece, up to 3 interaction techniques, the main topic and the link to the article. In the Appendix there are all the interaction techniques and all topics of each interactive piece, as well as a screenshot of the webpage.

Economy/business and social issues were the most recurrent topic, present in 30% of the data journalism pieces, followed by health, the subject matter of 20% of the interactive pieces. Select was the most used interaction techniques, introduced in all 10 data journalism pieces, followed by filter, present in 70% of the interactive pieces.

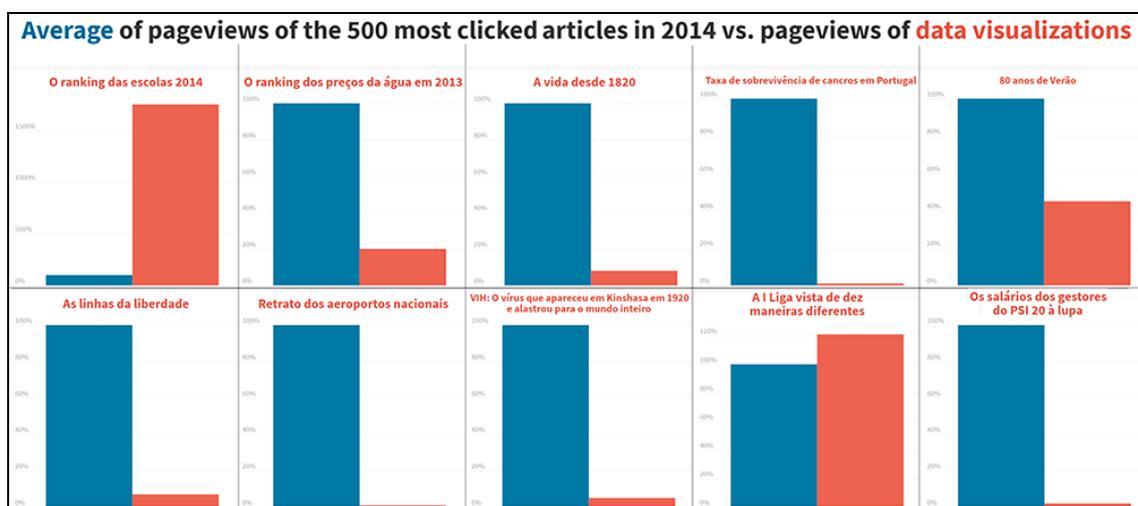


Figure 56 – Performance in pageviews of *Público* interactive pieces compared to the average of the most read articles in 2014 (represented as 100%).¹⁰⁸

Figure 56 shows pageviews performance of the data journalism pieces compared to the average of the 500 most read stories in 2014. Two out of 10 interactive pieces had more pageviews than the average of the 500 articles and 1 data visualization is among the 10 most read news pieces of 2014: “O ranking das escolas 2014”. This piece is also among the top 10 most read news articles of 2015.

“O ranking das escolas 2014” achieved the best performance among the interactive pieces: 17 times more pageviews than the average of the 500 most clicked stories in 2014. “Retrato dos aeroportos nacionais” reported the worst: fewer than 1% of the average hits of the 500 articles.

¹⁰⁸ Available at: <https://iloaguardotcom1.files.wordpress.com/2019/12/djxavg-2014-2.png>. Accessed on December 19, 2019.

The ratio of pageviews/unique pageviews shows how many times a page was accessed during users' sessions, i.e. it is possible to check if a news article was viewed multiple times by the same user. As stated before, GOV.UK website considers a high ratio above 1.4.¹⁰⁹

The average ratio of the 500 most read stories in 2014 is 2.9. "O ranking das escolas 2014" registered the highest ratio between the data journalism pieces, 17.5. That is, for each unique pageview more than 17 pageviews have been reported. "A I Liga vista de dez maneiras diferentes" also presented a higher proportion than the average, 7.8. The other 8 interactive pieces reported a lower ratio than the average of the 500 most read news articles, and "A vida desde 1820" registered the lowest proportion, 1.07.

Looking at the ratio of entrances/pageviews, "A vida desde 1820" is proportionately the interactive piece through which more visitors entered *Público's* website, 31%. This indicates that for 3 out of 10 users who accessed this data visualization, that was the first page on *Público's* website they visited during their session. The average proportion of the 500 most clicked news stories in 2014 is 26%. Only "A vida desde 1820" had a higher percentage than the average of the 500 most read articles and "A I Liga vista de dez maneiras diferentes" reported the lowest ratio: 2.5%.

On average, users who visited "As linhas da liberdade" spent on this interactive piece almost 4 times more time than those who accessed the 500 most clicked news articles in 2014. And users who visited "A vida desde 1820" and "Os salários dos gestores do PSI 20 à lupa" spent on these data journalism pieces more than twice the time of those who accessed the most read news stories. Half of the interactive pieces recorded visits that lasted less than the average of the 500 articles. Most notable is "A I Liga vista de dez maneiras diferentes", where users spent on average 86% less time on page than the visitors of the most clicked news stories.

The average bounce rate of the 500 most read news articles in 2014 is 28%, which means that almost 3 out of 10 visitors that entered the website through these

¹⁰⁹ Available at: <https://www.gov.uk/guidance/content-design/data-and-analytics>. Accessed on May 25, 2019.

news stories left *Público's* website without visiting any other page. Among the data journalism pieces, “A vida desde 1820” reported the highest value: 69%. There are 7 interactive pieces that registered a lower bounce rate than the 500 news stories, and “A I Liga vista de dez maneiras diferentes” had the lowest, 14%.

“O ranking das escolas 2014” registered the lowest exit rate, 5%. That is, 95% of the users who visited this interactive piece did not leave *Público's* website from it. The average exit rate from the 500 most read news articles in 2014 is 32%. There are 3 interactive pieces that reported a percentage higher than 32% and “A vida desde 1820” had the highest exit rate: 55%.

6.3. Data journalism pieces from 2014 and their related stories

Unfortunately, there is no data available in *Público's* Google Analytics for 2 related news on the following metrics: traffic sources, devices, location, new vs. returning visitors, and browsers.

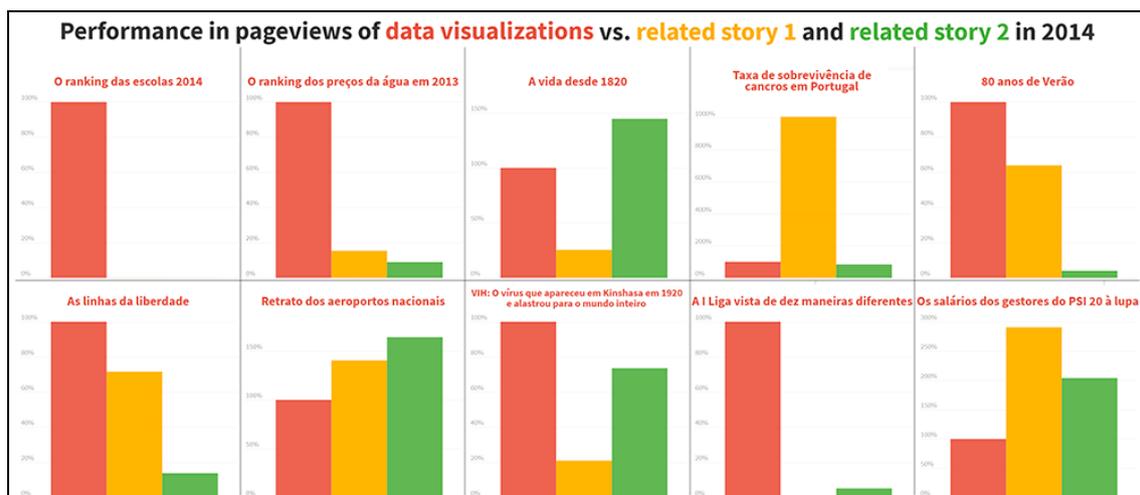


Figure 57 – Performance in pageviews of *Público* interactive pieces (represented as 100%) compared to their related news stories in 2014.¹¹⁰

Figure 57 shows the performance (in pageviews) of the interactive pieces compared to their respective related news stories. The data journalism pieces are displayed as 100%. That way, related articles that achieved a higher number of pageviews than the interactive pieces appear in the figure with more than 100%; and

¹¹⁰ Available at: <https://iloaguardotcom1.files.wordpress.com/2019/12/djxrel-2014-2.png>. Accessed on December 19, 2019.

those related news stories that registered fewer hits than the data journalism piece appear in the figure with less than 100%.

Six out of 10 interactive pieces had more clicks than their respective news stories. “O ranking das escolas 2014” is the data journalism piece that performed best: 1,722 times more pageviews than the related story with better performance. “Retrato dos aeroportos nacionais” and “Os salários dos gestores do PSI 20 à lupa” are the data visualization that reported fewer hits than their related stories. Information on related news stories such as title, publication date, and link to the article are available in the Appendix.

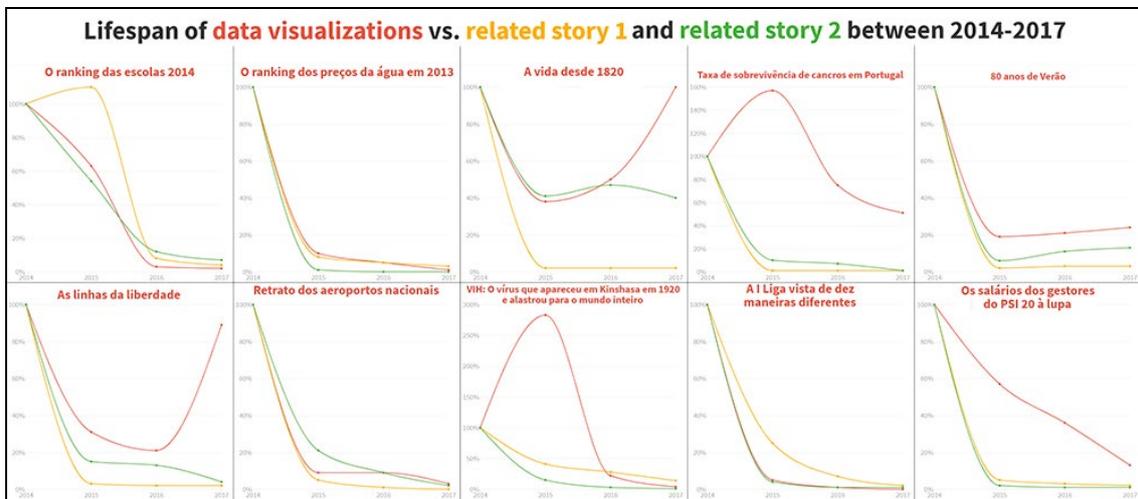


Figure 58 – Evolution in pageviews of *Público* 2014 interactive pieces and their related news stories over time.¹¹¹

With the related news stories, it is also possible to look at the lifetime of the articles. Figure 58 shows the percentage of pageviews in 2015, 2016 and 2017 compared to 2014, represented as 100%. That way, we can observe the evolution of hits that the data journalism pieces and their related news stories have registered over the years.

Nineteen of the 20 related news stories recorded fewer pageviews in 2015 than in the previous year, which is the normal behavior of news articles. However, 2 of the 10 interactive pieces reported more hits in 2015 than in 2014: “Taxa de sobrevivência de cancro em Portugal” and “VIH: O vírus que apareceu em Kinshasa em 1920 e

¹¹¹ Available at: <https://iloaguardotcom1.files.wordpress.com/2019/12/djxrel-2014-17-2.png>. Accessed on December 19, 2019.

alastrou para o mundo inteiro”. And “A vida desde 1820” registered more pageviews in 2017 than in 2014: 0.5% more pageviews in 2017. Among the 10 data journalism pieces, “A I Liga vista de dez maneiras diferentes” presented the worst performance in 2017: fewer than 0.5% of the number of hits it registered in the year of its publication.

All the related news stories in 2014 reported a ratio of pageviews/unique pageviews below 2.9, the average ratio of the 500 most clicked news articles that year. “A vida desde 1820”, “As linhas da liberdade”, and “Os salários dos gestores do PSI 20 à lupa” showed a proportion lower than at least 1 of the related articles. That is, in the other 7 data journalism pieces the visitors returned to the page during the same session more times than those who accessed their related news stories.

However, over the years, the number of returning visitors is lower on most interactive pieces compared to related news. According to Google Analytics, the user that visited the website within the past two years and returns from the same device is marked as a returning visitor. Between 2014 and 2017, “VIH: O vírus que apareceu em Kinshasa em 1920 e alastrou para o mundo inteiro” is the interactive piece with higher returning visitors rate than its related news stories. “O ranking das escolas 2014”, “80 anos de Verão”, and “As linhas da liberdade” are the data journalism pieces that achieved in that period a higher proportion of new visitors than their related news stories.

The ratio of entrances/pageviews is lower in all 10 data visualizations than at least 1 of their related news articles, which indicates that the visitors of interactive pieces visited another page of *Público*'s website before accessing it more times than the visitors of the related articles.

Users who visited “As linhas da liberdade” spent on this interactive piece 3 times more time than those who accessed the related news stories. “A vida desde 1820” and “Os salários dos gestores do PSI 20 à lupa” also registered more time on the page than at least 1 of the related articles. The users spent about 8 times more time on related news of “A I Liga vista de dez maneiras diferentes” than on this interactive piece.

There are 8 data journalism pieces where the bounce rate is higher than in at least 1 related news article. This means that on these interactive pieces the visitors entered the website from these data visualization and left the website without visiting any other *Público's* page in a proportion higher than in the respective related news. “Taxa de sobrevivência de cancro em Portugal” and “As linhas da liberdade” registered a bounce rate lower than their related stories.

“A vida desde 1820”, “As linhas da liberdade”, and “Os salários dos gestores do PSI 20 à lupa” are the data journalism pieces that reported an exit rate higher than their related articles. The other 7 interactive pieces registered a lower exit rate than their related stories. That is, users who visited these 7 data visualizations then visited other pages of *Público's* website in a greater proportion than those who visited the related news.

The desktop was the most used device in all 10 data journalism pieces. Mobile was the device more frequently used in “O ranking das escolas 2014”, 26%. “A vida desde 1820” had the highest proportion of visitors that used tablet: 36%. Fourteen of the 18 related news stories recorded more than 50% of their traffic coming from desktops. There are no data in this metric for 2 related articles.

Google is the place where most users were before visiting 4 of the interactive pieces in 2014: “O ranking dos preços da água em 2013”, “VIH: O vírus que apareceu em Kinshasa em 1920 e alastrou para o mundo inteiro”, “A Liga vista de dez maneiras diferentes”, and “Os salários dos gestores do PSI 20 à lupa”. The other 6 data journalism pieces registered more visitors arriving from another location. Thirteen of the 18 related news stories also registered most of the users coming from Google. There are no data in this metric for 2 related articles.

About half of “Taxa de sobrevivência de cancro em Portugal” visitors in 2014 were from outside Portugal. Among the other interactive pieces, only “Retrato dos aeroportos nacionais” reported a 100% Portuguese audience. Seven of the 18 related news stories registered 100% of the visitors from Portugal and other 2, more than 90% of users accessing from there. There are no data in this metric for 2 related articles.

Google Chrome was the browser most used by visitors of 6 of the data journalism pieces in 2014. “O ranking dos preços da água em 2013”, “A vida desde 1820”, “Taxa de sobrevivência de cancro em Portugal”, and “Retrato dos aeroportos nacionais” registered other browsers as the most used. Eleven of the 18 related news stories were also accessed mostly by visitors using Google Chrome. There are no data in this metric for 2 related articles.

6.4. Data journalism pieces from 2015 and the average of the 500 most read articles

Table 11 – The 10 data journalism pieces chosen in 2015 from *Público* and their specific aspects.

Title	Interaction techniques	Topic	URL
A água que gastamos mas não vemos	Inspect, Select.	Environment / science/technology.	http://bit.ly/2Jt7IAW
Radiografia da TAP	Narrate/History, Select.	Economy/business.	http://bit.ly/2H6Lrri
As melhores praias para este Verão	Filter, Select.	Lifestyle/culture.	http://bit.ly/2vLeXMC
O que pomos no prato... não é o que devíamos pôr	Select, Narrate/History.	Health.	http://bit.ly/2VNqA4w
Relações que matam	Narrate/History, Inspect.	Social issues.	http://bit.ly/2J4HSnN
Uma década de pobreza e algumas medidas	Select, Inspect, Narrate/History.	Economy/business.	http://bit.ly/301VhC2
Esta Primavera vá ver os passarinhos	Narrate/History, Inspect, Select.	Environment / science/technology.	http://bit.ly/2YcNRtY
Cristiano Ronaldo em golos	Inspect, Select, Narrate/History,	Sports.	http://bit.ly/2HaqDiN
Menos alunos para as provas do secundário	Select, Narrate/History.	Education.	http://bit.ly/2J8jUIf
Todos os resultados das legislativas	Inspect, Select.	Politics/governance.	http://bit.ly/2VlwR83

Table 11 shows some information about the chosen data journalism pieces from 2015: the title of the piece, up to 3 interaction techniques, the main topic and the link to the article. In the Appendix there are all the interaction techniques and all topics of each interactive piece, as well as a screenshot of the webpage.

Environment/science/technology was the most frequent topic, present in 3 interactive pieces, followed by economy/business, lifestyle/culture and social issues, each topic the subject of 2 data visualization. Select was the most used interaction technique, introduced in 90% of the data journalism pieces, followed by narrate/history and inspect, each one present in half of the interactive pieces.

Figure 59 shows pageviews performance of the data journalism pieces compared to the average of the 500 most read news articles in 2015. All 10 interactive pieces had fewer hits than the average of the 500 most clicked news stories. “As melhores praias para este Verão” registered the best performance: 83% less pageviews than the average of the 500 most read stories. “Todos os resultados das legislativas” reported the worst, fewer than 1% of the hits of the average of the 500 news articles.

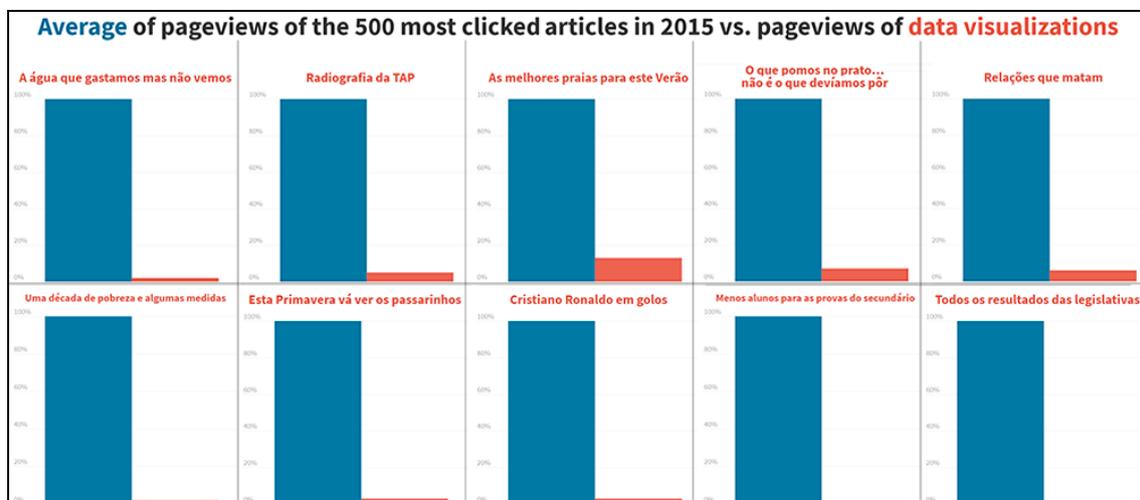


Figure 59 – Performance in pageviews of *Público* interactive pieces compared to the average of the most read articles in 2015 (represented as 100%).¹¹²

The average ratio of pageviews/unique pageviews of the 500 most read news stories in 2015 is 4.4. That is, for each unique pageview more than 4 pageviews have

¹¹² Available at: <https://iloaguardotcom1.files.wordpress.com/2019/12/djxavg-2015-2.png>. Accessed on December 19, 2019.

been registered. All 10 data journalism pieces reported a proportion lower than 4.4. “Todos os resultados das legislativas” had the highest ratio among the interactive pieces, 2.1. “Radiografia da TAP”, “Cristiano Ronaldo em golos”, and “Menos alunos para as provas do secundário” had the lowest ratio, 1.11.

Looking at the ratio of entrances/pageviews, “O que pomos no prato... não é o que devíamos pôr” is proportionately the data journalism piece through which more visitors entered *Público’s* website, 50%. This indicates that for half of the users who accessed this interactive piece, that was the first page on *Público’s* website they visited during their session. The average percentage of the 500 most read news stories in 2015 is 18%. There are 3 data visualization that reported a ratio lower than 18% and “Radiografia da TAP” had the lowest proportion: 6%. That is, more than 9 out of 10 visitors of this data journalism piece had visited another page of *Público* before accessing it.

All 10 interactive pieces registered more average time on page than the 500 most read news articles in 2015. “O que pomos no prato... não é o que devíamos pôr” reported the best performance: users who visited this data visualization spent on average 3.5 more time on page than those who accessed the 500 news stories. “Todos os resultados das legislativas” had the worst performance: users spent on average about 38% more time on this interactive piece than the visitors of the most read articles.

The average bounce rate of the 500 most read news articles in 2015 is 17%, which means almost 2 out of 10 visitors that entered the website through these news stories left *Público’s* website without visiting any other page. All 10 interactive pieces reported a higher bounce rate than the 500 news stories. “Esta Primavera vá ver os passarinhos” had the highest value: 70%. “Relações que matam” registered the lowest rate, 30%.

“Todos os resultados das legislativas” reported the lowest exit rate, 18%. That is, fewer than 2 in 10 users who visited this interactive piece left *Público’s* website from it. The average exit rate from the 500 most read news articles in 2015 is 24%. There are 9 data journalism pieces that registered a percentage higher than 24% and “As melhores praias para este Verão” registered the highest exit rate: 49%.

6.5. Data journalism pieces from 2015 and their related stories

Unfortunately, there is no data available in *Público's* Google Analytics for 1 related news on the following metrics: traffic sources, devices, location, new vs. returning visitors, and browsers.

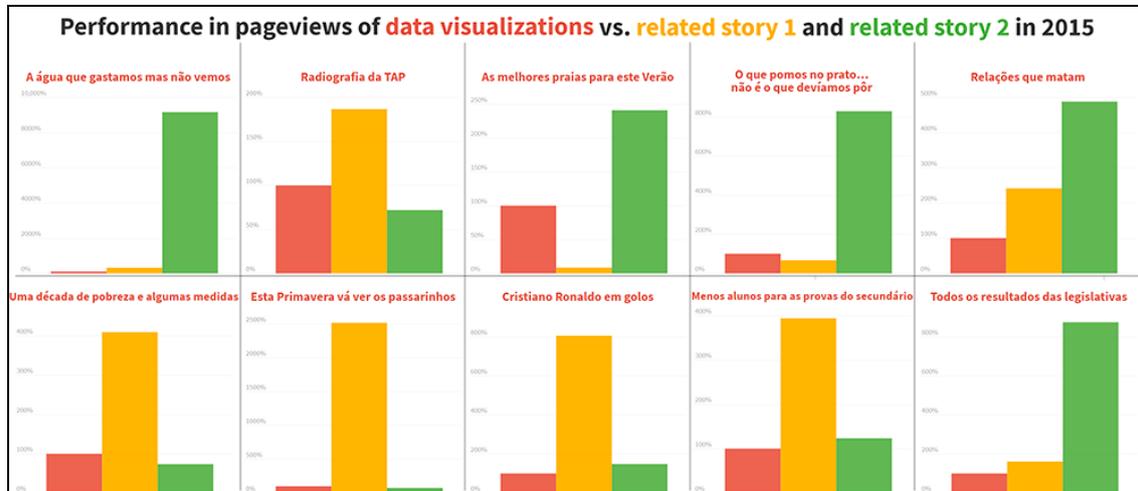


Figure 60 – Performance in pageviews of *Público* interactive pieces (represented as 100%) compared to their related news stories in 2015.¹¹³

Figure 60 shows the performance (in pageviews) of the interactive pieces compared to their respective related news stories. The data journalism pieces are displayed as 100%. That way, related articles that achieved a higher number of pageviews than the data visualization appear in the figure with more than 100%; and those related news stories that registered fewer hits than the interactive pieces appear in the figure with less than 100%.

Half of the data journalism pieces registered more clicks than at least 1 of their related stories: “Radiografia da TAP”, “As melhores praias para este Verão”, “O que pomos no prato... não é o que devíamos pôr”, “Uma década de pobreza e algumas medidas”, and “Esta Primavera vá ver os passarinhos”. The other interactive pieces reported fewer hits than their related stories. Information on related news stories such as title, publication date, and link to the article are available in the Appendix.

¹¹³ Available at: <https://iloaguardotcom1.files.wordpress.com/2019/12/djxrel-2015-2.png>. Accessed on December 19, 2019.

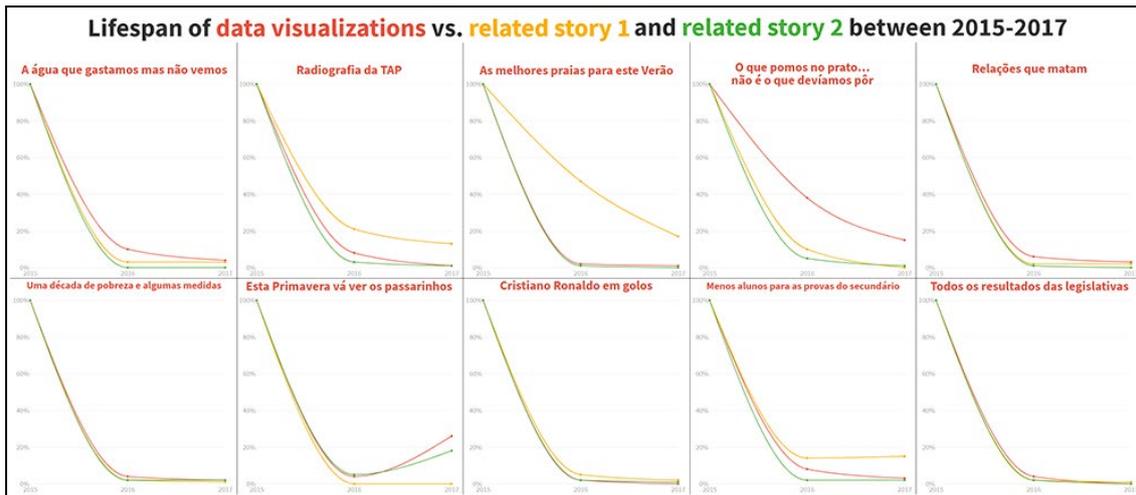


Figure 61 – Evolution in pageviews of *Público* 2015 interactive pieces and their related news stories over time.¹¹⁴

Figure 61 shows the percentage of pageviews in 2016 and 2017 compared to 2015, represented as 100%. That way, we can observe the evolution of hits that data journalism pieces and their related news stories have registered over the years.

All the 20 related news stories reported fewer pageviews in 2016 than in the previous year. None of the interactive pieces reported more hits in 2016 than in 2015. In 2017 once again none of the data journalism pieces registered more pageviews than in 2015. “Esta Primavera vá ver os passarinhos” is the data visualization that performed best in 2017: 26% of the number of hits it reported in the year of its publication. “Todos os resultados das legislativas” reported the worst performance in 2017: less than 0.1% of pageviews it registered in 2015.

Looking at the ratio of pageviews/unique pageviews in 2015, 19 of the 20 related news stories reported a number below 4.4, the average ratio of the 500 most clicked news articles that year. Half of the data journalism pieces registered a proportion higher than at least 1 of the related articles: “A água que gastamos mas não vemos”, “As melhores praias para este Verão”, “O que pomos no prato... não é o que devíamos pôr”, “Relações que matam”, and “Todos os resultados das legislativas”. That is, in the other 5 interactive pieces, visitors returned to the page during the same session fewer times than those who accessed their related news stories.

¹¹⁴ Available at: <https://iloaguardotcom1.files.wordpress.com/2019/12/djxrel-2015-17-2.png>. Accessed on December 19, 2019.

Between 2015 and 2017 the rate of returning visitors is higher in 3 of the data journalism pieces compared to their related news: “O que pomos no prato... não é o que devíamos pôr”, “Relações que matam”, and “Menos alunos para as provas do secundário”. “A água que gastamos mas não vemos”, “Cristiano Ronaldo em golos”, and “Todos os resultados das legislativas” are the interactive pieces with a proportion of new visitors higher than their related articles in that period.

The ratio of entrances/pageviews is higher in only 1 data journalism piece than in its related news articles: “O que pomos no prato... não é o que devíamos pôr”. The other 9 interactive pieces registered a proportion lower than at least 1 of the related news, which implies that visitors of these data visualizations visited another page of *Público’s* website before accessing it more times than the visitors of the related news articles.

Users who visited “O que pomos no prato... não é o que devíamos pôr” spent on this data journalism piece almost 4 times more time than those who accessed the related news stories. “A água que gastamos mas não vemos”, “Esta Primavera vá ver os passarinhos”, and “Cristiano Ronaldo em golos” also registered more time on the page than at least 1 of their related articles. The users spent almost 3 times more time on related news of “Radiografia da TAP” than on this interactive piece.

All 10 data journalism pieces had a bounce rate higher than at least 1 related news article. This means that in all interactive pieces the visitors entered the website through these data visualizations and left the website without visiting any other *Público’s* page in a proportion higher than in the respective related news.

Half of the data journalism pieces registered an exit rate lower than their related articles: “Radiografia da TAP”, “As melhores praias para este Verão”, “Relações que matam”, “Menos alunos para as provas do secundário”, and “Todos os resultados das legislativas”. That is, users who visited these interactive pieces then visited other pages of *Público’s* website in a greater proportion than those who visited the related news stories.

The desktop was the most used device in all 10 data journalism pieces. Mobile was more frequently used in “Menos alunos para as provas do secundário”, 44%.

“Todos os resultados das legislativas” had the highest proportion of visitors that used tablet: 50%. Eleven of the 19 related news stories registered more than 50% of their traffic coming from desktops. There are no data in this metric for 1 related article.

Google is the place where most users were before visiting 7 of the interactive pieces in 2015. Only “As melhores praias para este Verão”, “O que pomos no prato... não é o que devíamos pôr”, and “Uma década de pobreza e algumas medidas” registered more visitors arriving from another location. Fourteen of the 19 related news stories registered most of the users coming from Facebook. There are no data in this metric for 1 related article.

About 33% of “Esta Primavera vá ver os passarinhos” visitors in 2015 were from outside Portugal. Among the other interactive pieces, “Menos alunos para as provas do secundário” and “Todos os resultados das legislativas” reported a 100% Portuguese audience. Three of the 19 related news stories recorded more than 90% of the visitors from Portugal. There are no data in this metric for 1 related article.

Google Chrome was the browser most used by visitors of 9 of the data journalism pieces in 2015. Only “Todos os resultados das legislativas” registered other browsers as the most used, Safari and Internet Explorer. Fifteen of the 19 related news articles were also accessed mostly by visitors using Google Chrome.

6.6. Data journalism pieces from 2016 and the average of the 500 most read articles

Table 12 shows some information about the chosen data journalism pieces from 2016: the title of the piece, up to 3 interaction techniques, the main topic and the link to the article. In the Appendix there are all the interaction techniques and all topics of each interactive piece, as well as a screenshot of the webpage.

Politics/governance, education, sports and foreign affairs were the most common topics among the interactive pieces, each one present in 2 articles. Select was the most used interaction technique, introduced in 9 of the data journalism pieces, followed by narrate/history, present in half of data visualizations.

Table 12 – The 10 data journalism pieces chosen in 2016 from *Público* and their specific aspects.

Title	Interaction techniques	Topic	URL
Os exames chegaram	Select.	Education.	http://bit.ly/3058QB0
Portugal no pódio olímpico	Select.	Sports.	http://bit.ly/2WtfpLm
Um século de energia em Portugal	Inspect, Select, Filter.	Environment / science/technology.	http://bit.ly/2H4S8de
As preocupações e motivações dos professores	Select.	Education.	http://bit.ly/2WyPmm8
Como se elege um Presidente nos EUA	Narrate/History.	Foreign affairs.	http://bit.ly/2DVjiS1
De que é feita a cerveja	Select, Narrate/History.	Lifestyle/culture.	http://bit.ly/2LrO2A9
ZZZZZ... Zika	Select, Narrate/History.	Health.	http://bit.ly/2POarGz
A crise dos refugiados em números	Select, Narrate/History.	Foreign affairs.	http://bit.ly/2LrO4bf
Os 16 desportos de Portugal no Rio 2016	Select.	Sports.	http://bit.ly/2LHrrQv
Presidenciais 2016 – Resultados completos	Filter, Inspect, Explore.	Politics/governance.	http://bit.ly/2Luq4En

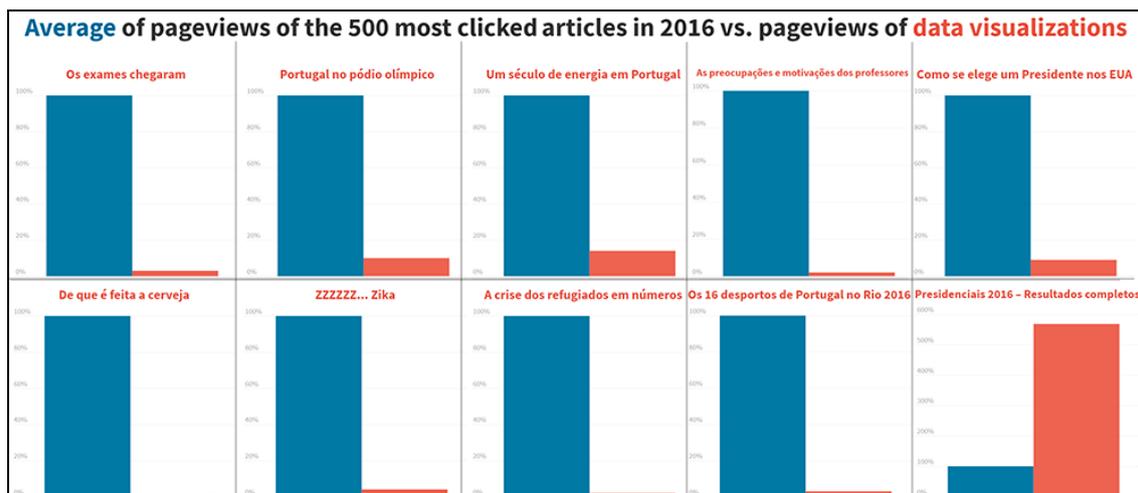


Figure 62 – Performance in pageviews of *Público* interactive pieces compared to the average of the most read articles in 2016 (represented as 100%).¹¹⁵

Figure 62 shows pageviews performance of the interactive pieces compared to the average of the 500 most read news stories in 2016. Only 1 data journalism piece had more hits than the average of the 500 articles and is among the 10 most read news stories of 2016: “Presidenciais 2016 – Resultados completos”.

“Presidenciais 2016 – Resultados completos” reported the best performance: more than 5 times more pageviews than the average of the 500 most clicked stories in 2016. “De que é feita a cerveja” registered the worst: fewer than 1% of the average hits of the most read news articles.

The average ratio of pageviews/unique pageviews of the 500 most clicked news stories in 2016 is 3.6. That is, for each unique pageview more than 3 pageviews have been registered. All 10 data journalism pieces reported a proportion lower than 3.6. “Um século de energia em Portugal” reported the highest ratio among the interactive pieces, 2.4. “Como se elege um Presidente nos EUA” had the lowest ratio, 1.07.

Looking at the ratio of entrances/pageviews, “Como se elege um Presidente nos EUA” is proportionately the interactive piece through which more visitors entered *Público*’s website, 61%. The average proportion of the 500 most read news stories in 2016 is 22%. There are 4 data journalism pieces that registered a percentage lower than 22% and “Os 16 desportos de Portugal no Rio 2016” reported the lowest ratio:

¹¹⁵ Available at: <https://iloaguardotcom1.files.wordpress.com/2019/12/djxavg-2016-2.png>. Accessed on December 19, 2019.

12%, which indicates that about 9 out of 10 visitors of this interactive piece had visited another page of *Público* before accessing it.

Nine out of 10 data journalism pieces reported more average time on page than the 500 most read news articles in 2016. “Como se elege um Presidente nos EUA” registered the best performance: users who visited this data visualization spent on average 3.8 more time on page than those who accessed the 500 news stories. Only “Presidenciais 2016 – Resultados completos” reported visits that lasted less than the average of the most read articles: users spent on average about 18% less time on this interactive piece than the visitors of the 500 articles.

The average bounce rate of the 500 most read news articles in 2016 is 33%. Which means that about 3 out of 10 visitors that entered the website through these news stories left *Público*'s website without visiting any other page. All 10 interactive pieces reported a higher bounce rate than the 500 news stories. “Como se elege um Presidente nos EUA” achieved the highest value: 82%. “De que é feita a cerveja” registered the lowest rate, 38%.

“Presidenciais 2016 – Resultados completos” reported the lowest exit rate, 18%. That is, about 2 in 10 users who visited this data visualization left *Público*'s website from it. The average exit rate from the 500 most read news articles in 2016 is 28%. There are 7 interactive pieces that had a percentage higher than 28% and “Como se elege um Presidente nos EUA” registered the highest exit rate: 70%.

6.7. Data journalism pieces from 2016 and their related stories

Figure 63 shows the performance (in pageviews) of the interactive pieces compared to their respective related news stories. The data journalism pieces are displayed as 100%. That way, related articles that achieved a higher number of pageviews than the data visualization appear in the figure with more than 100%; and those related news stories that registered fewer hits than the interactive pieces appear in the figure with less than 100%.

Half of the data journalism pieces had more clicks than their respective news stories. “Presidenciais 2016 – Resultados completos” is the interactive piece that performed best: almost 28 times more pageviews than the related story with better

performance. “As preocupações e motivações dos professores” is the only data visualization that reported fewer hits than the both related stories. Information on related news stories such as title, publication date, and link to the article are available in the Appendix.

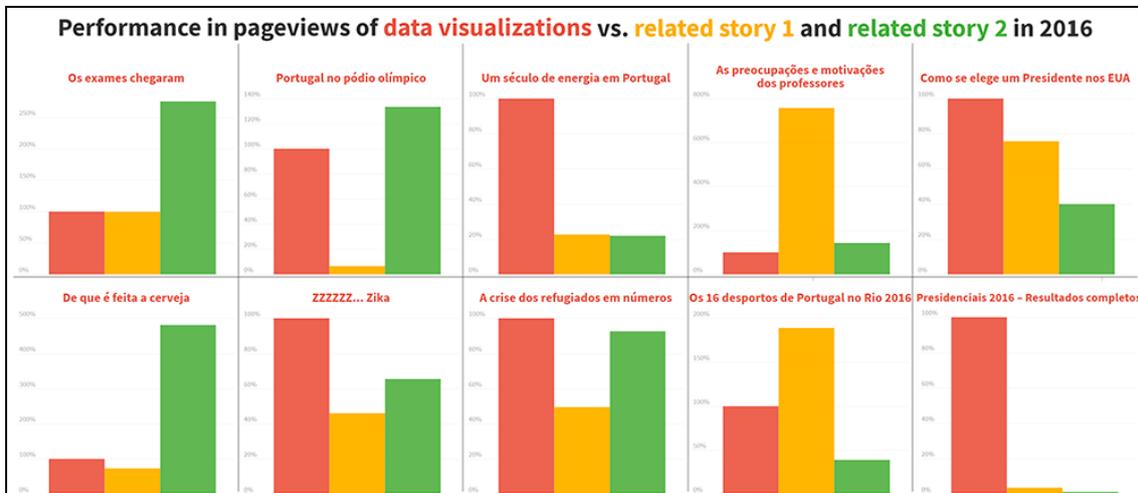


Figure 63 – Performance in pageviews of *Público* interactive news pieces (represented as 100%) compared to their related news stories in 2016.¹¹⁶

Figure 64 shows the percentage of pageviews in 2017 compared to 2016, represented as 100%. That way, we can observe the evolution of hits that the data journalism pieces and their related news stories have registered over the years.

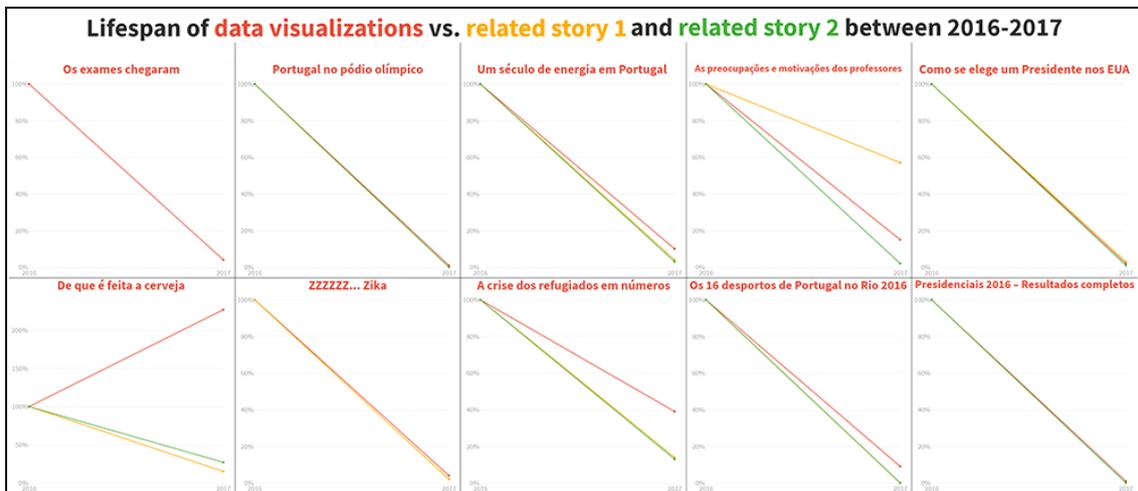


Figure 64 – Evolution in pageviews of *Público* 2016 interactive pieces and their related news stories over time.¹¹⁷

¹¹⁶ Available at: <https://iloaguardotcom1.files.wordpress.com/2019/12/djxrel-2016-2.png>. Accessed on December 19, 2019.

¹¹⁷ Available at: <https://iloaguardotcom1.files.wordpress.com/2019/12/djxrel-2016-17-2.png>. Accessed on December 19, 2019.

All 20 related news stories had fewer pageviews in 2017 than in the previous year. There is 1 interactive piece that reported more hits in 2017 than in 2016: “De que é feita a cerveja”. This interactive piece registered about 11 times more pageviews in 2017 than in the year of its publication. Among the 10 interactive pieces, “Portugal no pódio olímpico” had the worst performance in 2017: fewer than 1% of the number of hits it registered in the year of its publication.

Looking at the ratio of pageviews/unique pageviews in 2016, all 20 related news stories reported a number below 3.6, the average ratio of the 500 most clicked news articles that year. Only “ZZZZZZ... Zika” and “Os 16 desportos de Portugal no Rio 2016” registered a proportion higher than their related articles. That is, in these 2 data journalism pieces the visitors returned to the page during the same session more times than those who accessed their related news stories. The other 8 interactive pieces registered a ratio lower than at least 1 of the related articles.

Between 2016 and 2017, none of the interactive pieces had a returning visitors rate higher than their related news stories. There are 4 data journalism pieces that reported, in that period, a proportion of new visitors higher than their related news stories: “Portugal no pódio olímpico”, “A crise dos refugiados em números”, “Os 16 desportos de Portugal no Rio 2016”, and “Presidenciais 2016 – Resultados completos”.

The ratio of entrances/pageviews is higher in only 1 data journalism piece than in its related news articles: “Portugal no pódio olímpico”. The other 9 interactive pieces registered a lower proportion than at least 1 of the related stories, which implies that the visitors to these data visualization visited another page of *Público*'s website before accessing it more times than the visitors of the related articles.

Users who visited “Como se elege um Presidente nos EUA” spent on this interactive piece about twice the time of those who accessed the related news. Other 7 data journalism pieces also registered more average time on the page than at least 1 of their related articles: “Os exames chegaram”, “Portugal no pódio olímpico”, “Um século de energia em Portugal”, “De que é feita a cerveja”, “A crise dos refugiados em números”, “Os 16 desportos de Portugal no Rio 2016”, and “Presidenciais 2016 – Resultados completos”. The users spent about 50% more time on related news of “As

preocupações e motivações dos professores” and “ZZZZZZ... Zika” than on these interactive pieces.

There are 8 data journalism pieces where the bounce rate is higher than in at least 1 related news article. This means that on these interactive pieces the visitors entered the website through these data visualization and left the website without visiting any other *Publico's* page in a proportion higher than in the respective related news. “As preocupações e motivações dos professores” and “A crise dos refugiados em números” reported a bounce rate lower than their related stories.

There are 3 interactive pieces that registered an exit rate lower than their related articles: “As preocupações e motivações dos professores”, “ZZZZZZ... Zika”, and “Os 16 desportos de Portugal no Rio 2016”. That is, users who visited these data journalism pieces then visited other pages of *Público's* website in a greater proportion than those who visited the related news stories.

The desktop was the most used device in 6 of the interactive pieces. Mobile was the device more frequently used in “Portugal no pódio olímpico”, “As preocupações e motivações dos professores”, and “Os 16 desportos de Portugal no Rio 2016”. “De que é feita a cerveja” registered the highest proportion of visitors that used tablet: 33%. Fourteen of the 20 related news stories had more than 50% of their traffic coming from the desktops.

Google is the place where most users were before visiting 6 of the interactive pieces in 2016. “Portugal no pódio olímpico”, “Como se elege um Presidente nos EUA”, “A crise dos refugiados em números”, and “Presidenciais 2016 – Resultados completos” registered more visitors arriving from another location. Twelve of the 20 related news stories also reported most of the users coming from Google.

About 45% of “ZZZZZZ... Zika” visitors in 2016 were from outside Portugal. “Os exames chegaram” and “De que é feita a cerveja” had a 100% Portuguese audience. Seven of the 20 related news recorded more than 90% of the visitors from Portugal.

Chrome was the browser most used by visitors of 9 of the interactive pieces in 2016. Only “ZZZZZZ... Zika” registered another browser as the most used, Safari. Nineteen of the 20 related news were also accessed mostly by visitors using Chrome.

VII – Discussion and Conclusions

In the last three chapters, we have presented the performance results of data journalism pieces grouped by year. In this chapter, we will examine the overall results in each newspaper. We also compare the general results of *The Texas Tribune*, *O Globo*, and *Público* to observe if there are similarities in the behavior of readers of different newspapers who access interactive pieces. In this manner, the main conclusions of this investigation are presented here, answering the research questions introduced in Chapter III.

7.1. *The Texas Tribune's* data journalism pieces performed impressively; those from *O Globo* and *Público*, not so much

The Texas Tribune's interactive pieces are among the most clicked content on the website: 16 of the 30 data journalism pieces analyzed between 2014 and 2016 registered more pageviews than the average of the 500 most read news stories in the year of its publication. The 5 most clicked news articles of 2014 were data journalism pieces, and 12 interactive pieces appear among the top 10 news stories in the analyzed period – 6 in 2014, 4 in 2015 and 2 in 2016. This performance of *The Texas Tribune's* data journalism pieces is quite divergent from that registered by *O Globo* and *Público*.

Between 2014 and 2016 only 1 interactive piece from *O Globo* and 3 from *Público* reported more pageviews than the average of 500 most read articles. Still, the performance of these data journalism pieces was astonishing: “Placar do Impeachment” from *O Globo* and “O ranking das escolas 2014” from *Público* were the most read news stories in their websites in their respective years.

Eleven out of 30 *Público* data journalism pieces registered more hits than their related news articles in the year they were published. *O Globo's* interactive pieces performed a little better: 17 out of 30. In *The Texas Tribune* 25 of the data visualizations reported more pageviews than the related news.

Such contrasting results between *The Texas Tribune* on the one hand and *O Globo* and *Público* on the other may be partially explained by the treatment that these media outlets give to their data journalism pieces. The U.S. platform seems to consider

the interactive pieces as crown jewels: the data visualizations were usually the main story and they regularly get featured on *The Texas Tribune's* homepage. On the other hand, in *O Globo* and *Público*, interactive pieces were produced as a sidebar, as complements to the main story and with limited – and sometimes none – featured on the homepage.¹¹⁸ And in the case of the legacy media, 1 of the 2 selected related news articles was often the main story and the data journalism piece was its accompaniment. Taking this into account, the performance of the interactive pieces in these newspapers is notable.

7.2. Data journalism pieces are not a landing page

56 of the 90 interactive pieces registered an entrance/pageviews ratio lower than the average of the 500 most clicked news stories. That is, 62% of the data journalism pieces were not the first page the user landed on when entering the website. This proportion is even higher when the data journalism pieces from *The Texas Tribune*, *O Globo* and *Público* are compared to their related news: in 75% of the 90 interactive pieces, the rate of users who visited before another page from the website is higher than in their related news.

The different approach to data journalism pieces in the news organizations can be spotted here. A third of *The Texas Tribune's* data journalism pieces had a higher entrance/pageviews ratio than their respective related news, while 93% of *Público* interactive pieces registered a proportion lower than their related stories. This means that in only 2 of 30 analyzed data journalism pieces the proportion of users that entered *Público's* website through the data visualization is higher than in their related news.

These results suggest that *Público's* interactive pieces do not appeal to a broader audience as their related news since data journalism pieces visitors overwhelmingly browse the site before entering the interactive piece. On the other hand, *The Texas Tribune's* data visualizations appear to have more autonomy, being the gateway of the website to a wider audience more often than some of its related stories.

¹¹⁸ In 2016 *O Globo's* interactive pieces received more importance and prominence in the homepage.

It is worth noting the evolution of the entrance/pageviews ratio in the interactive pieces of *O Globo*. While in 2014 and 2015 none of the data journalism pieces registered a ratio higher than the average of the 500 most read news stories, in 2016, 5 of the 10 interactive pieces did it. A similar performance occurred when comparing with the related news. While in 2014 only 1 data journalism piece achieved a proportion higher than their related articles, in 2016, 6 out of 10 interactive pieces did it.

“Interactive: Search CSCOPE Lesson Plans” and “Can you tell real news from fake news? Take our quiz to find out” from *The Texas Tribune* and *O Globo*’s “O mapa das coligações” registered more than 80% of their visitors entering the website through them. This implies that such data journalism pieces have gone beyond the homepage and links in related news and were widely shared by visitors.

7.3. Related news dies as (some) data journalism pieces flourish

177 of the 180 related news articles registered fewer pageviews in the year following their publication. This accounts for 98% of related articles. This performance is in line with previous studies (Ho & Liu, 2015). Nevertheless, 19 of the 90 interactive pieces analyzed had more hits in the year following their publication. This represents 21% of the data journalism pieces. And there are cases like *Globo*’s “Mapa da taxa de suicídio no mundo”, published in 2014, that in 2017 reported about 11 times more pageviews than in the year of its publication.

The Texas Tribune has the largest number of interactive pieces that registered a growth in the pageviews over the years: 11 out of 30. *O Globo* had 5 and *Público*, 3. Many of *The Texas Tribune*’s data journalism pieces may be classified as “decision-making tool” or “research tool” (Parasie & Dagiral, 2013, p. 864). These tools are updated regularly and continue to get featured on *The Texas Tribune*’s homepage.

It is interesting to note that some interactive pieces done by the legacy media that can also be classified as decision tools like *O Globo*’s “Calculadora de aposentadoria” and *Público*’s “O ranking das escolas 2014” were not updated, but new data visualizations virtually with the same features were created sometime later. For instance, “O Ranking das Escolas 2015” and “O Ranking das Escolas 2016” appear

among the most clicked articles of 2016 along with “O ranking das escolas 2014”. An interactive piece that presented all the ranking would be a better solution for users. The development of a new data journalism piece every year with virtually the same features may be justified by a technical issue or it can also be a way to get more pageviews, since to have a panorama of the previous years the user will have to access the other news. This last possibility is a practice well-documented by scholars (Anderson, 2011a; Tandoc, 2014; Usher, 2012).

Although few interactive pieces from *O Globo* and *Público* registered audience growth over the years, it should be noted that, unlike *The Texas Tribune’s* data journalism pieces, *O Globo’s* and *Público’s* data visualizations have not been updated or re-featured on the website. That is, even without any highlight, some interactive pieces reported more pageviews in the years following their publication than when they were published and promoted.

7.4. Users spend more time on data journalism pieces than on other news

61% of the data journalism pieces analyzed between 2014 and 2016 registered more average time on page than the average of the 500 most read news stories in the year they were published. The audience of *Público* spent more time in 24 interactive pieces than in the most read articles. In *O Globo* 14 of the 30 data journalism pieces had a higher average time than the most clicked news, and in *The Texas Tribune*, 17.

In Figure 65 the average time on page of the 500 most read news pieces is represented as 100% and the average results of the data journalism pieces in that year range accordingly. The interactive pieces from *Público* registered more average time on the page every year and in 2015 and 2016 had more than twice the time on page than the average of the most read news stories.

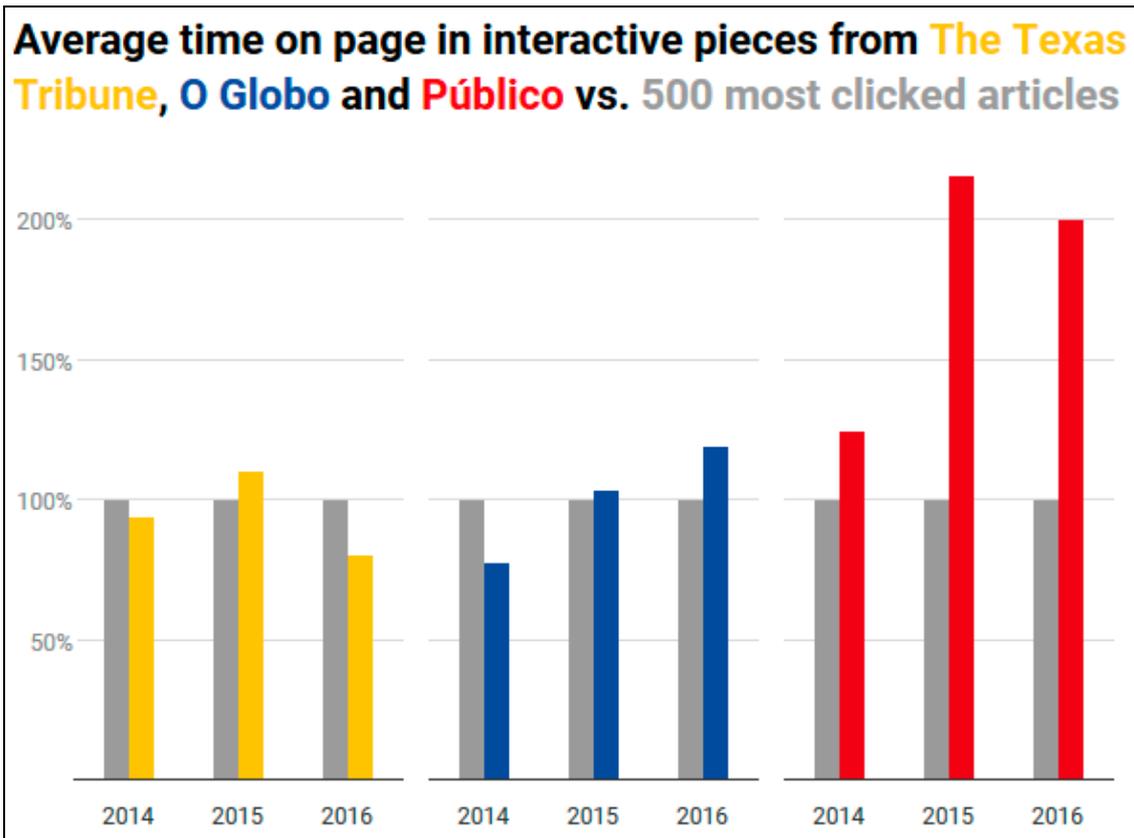


Figure 65 – Average time on page. Most read news are represented as 100%. Interactive pieces’ results are relative to that value.

In comparison to related news stories, 55 of the 90 interactive pieces reported more average time on the page than at least 1 of the related articles. These 2 findings meet the results of 3 web-based filed experiments made by Boy, Detienne, and Fekete (2015), where the average time spent on the page in the data visualizations was higher than in the standard version of the news story.

60% of the people who visited *O Globo’s* interactive pieces spent more time in these stories than in their related articles. The proportion of users who spend more time on data journalism pieces than in their related news stories on *P ublico* is 36% and 16% on *The Texas Tribune*.

Along with pageviews, time spent is among the top 5 metrics used regularly by newsrooms (ICFJ, 2017). And advertisers that have always been more interested in clicks are beginning to show enthusiasm in also counting time (Chartbeat, 2015; Ingram, 2015; B. Merritt, 2017).

Unfortunately, it is not possible to state categorically whether people spent more time in data visualizations because they were actively exploring the tool, or trying to understand the visualization proposal or something else, since, using the average time on page metric, it is not possible to conclude “whether users are actively interacting with a site or simply on it (while potentially doing something else)” (Cherubini & Nielsen, 2016, p. 35).

Although metrics such as engaged time, which reports the average time visitors spend actively interacting with a story, are being increasingly used in newsrooms (Cherubini & Nielsen, 2016; ICFJ, 2017), it was not possible to collect data from this kind of metric for all analyzed articles. Even so, metrics such as bounce rate show whether there was some kind of interaction on the page.

7.5. Data journalism pieces visits are single-page visits

A quarter of the interactive pieces analyzed between 2014 and 2016 registered a lower bounce rate than the average of the 500 most read news stories in the year they were published. This means that in 75% of the data journalism pieces visitors that entered the website through these data visualization left the website without visiting any other page in a higher proportion than in the average of the 500 most read articles. *The Texas Tribune* achieved the best performance: 10 of the 30 interactive pieces reported a lower bounce rate than the average of most clicked news. *O Globo* registered 9, and *Público* had 3 data journalism pieces with a proportion inferior than the average of the 500 most read articles in the year they were published.

In Figure 66 the average bounce rate of the 500 most read news pieces is represented as 100% and the average bounce rate of the data journalism pieces in that year range accordingly. The interactive pieces from *The Texas Tribune* are the only ones where in all the years analyzed the bounce rate is equal to or below the average of the most read news stories.

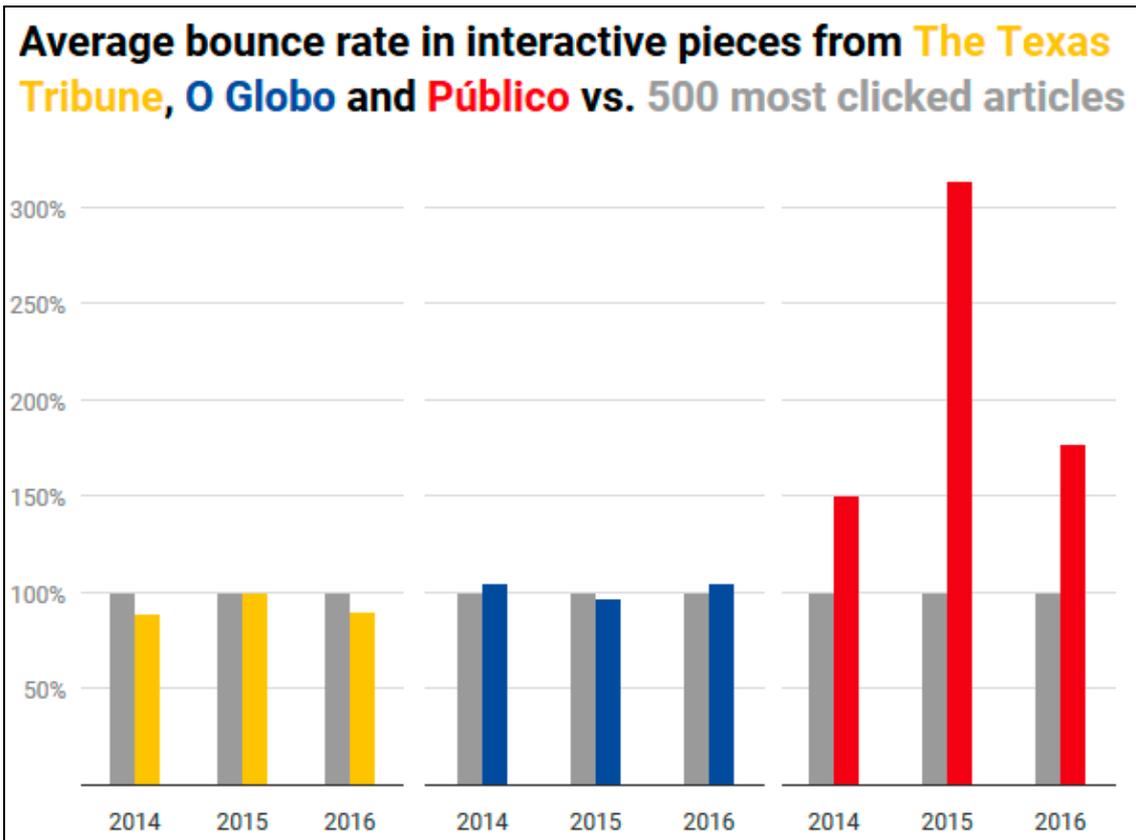


Figure 66 – Average bounce rate. Most read news are represented as 100%. Interactive pieces’ results are relative to that value.

In comparison to related news stories, 31 of the 90 interactive pieces reported a lower bounce rate than their related news articles. Once again, *The Texas Tribune* registered the best performance: 56% of the data journalism pieces had a lower proportion than their related news. *O Globo* reported 33% and *Público*, 13% of data visualizations in which the proportion of visitors that left the website without visiting any other page is lower than in those who accessed their related articles.

Unfortunately, from the data we collected it is not possible to identify the reasons for the bounce. But our findings showing that users spend more time on interactive pieces than on the most read news or the related articles suggest that readers interact with the data journalism pieces before leaving the site.

We do not know if *The Texas Tribune*, *O Globo*, or *Público* apply any technique (tracking metrics, surveys, etc.) to know how (and even if there is) interaction with the data visualization, but studies show that there is no such evaluation by the newsrooms. For instance, in their pursuit to compare the practice of data journalism at

multiple large, medium, and small-sized U.S. news organizations, Fink and Anderson (2015, p. 8) claim that, in fact, “data journalists knew little about how their audiences interacted with their stories. (...) Another journalist said she had no way to know how users interacted with data, but she did know that having data tended to increase the time they spent on a page”.

Finally, it is important to emphasize that – just like most metrics – bounce rate needs to be put into context. If the entrance/pageviews ratio is low, the bounce rate can be an outlier, not describing the behavior of most users. For instance, the entrance/pageviews ratio of *Público* interactive pieces is very low, and the bounce rate is very high. That is, the number of people who entered *Público*’s website through the interactive pieces is tiny, but this modest group left the website without visiting any other page in a large proportion.

Likewise, the data journalism pieces of *The Texas Tribune* registered a high proportion of entrance/pageviews and the lowest bounce rate of the 3 news organizations. In this case, the opposite occurred: the number of people who entered *The Texas Tribune*’s website through the interactive pieces is large and this big group actually visited another page before leaving the website.

7.6. Data journalism pieces drive more recirculation

53% of the interactive pieces analyzed registered a lower exit rate than the average of the 500 most clicked news stories in the year they were published. That is, in more than half of the data journalism pieces the user visited another page of the website in greater proportion than who accessed the most read stories. 24 of the 30 interactive pieces from *O Globo* had a lower exit rate than the average of the most clicked articles. In *The Texas Tribune* was 13 and in *Público*, 11.

In Figure 67 the average exit rate of the 500 most read news pieces is represented as 100% and the average exit rate of the data journalism pieces in that year range accordingly. The interactive pieces from *Público* in 2015 and 2016 are the only ones that registered a higher exit rate than the average of the most read news stories.

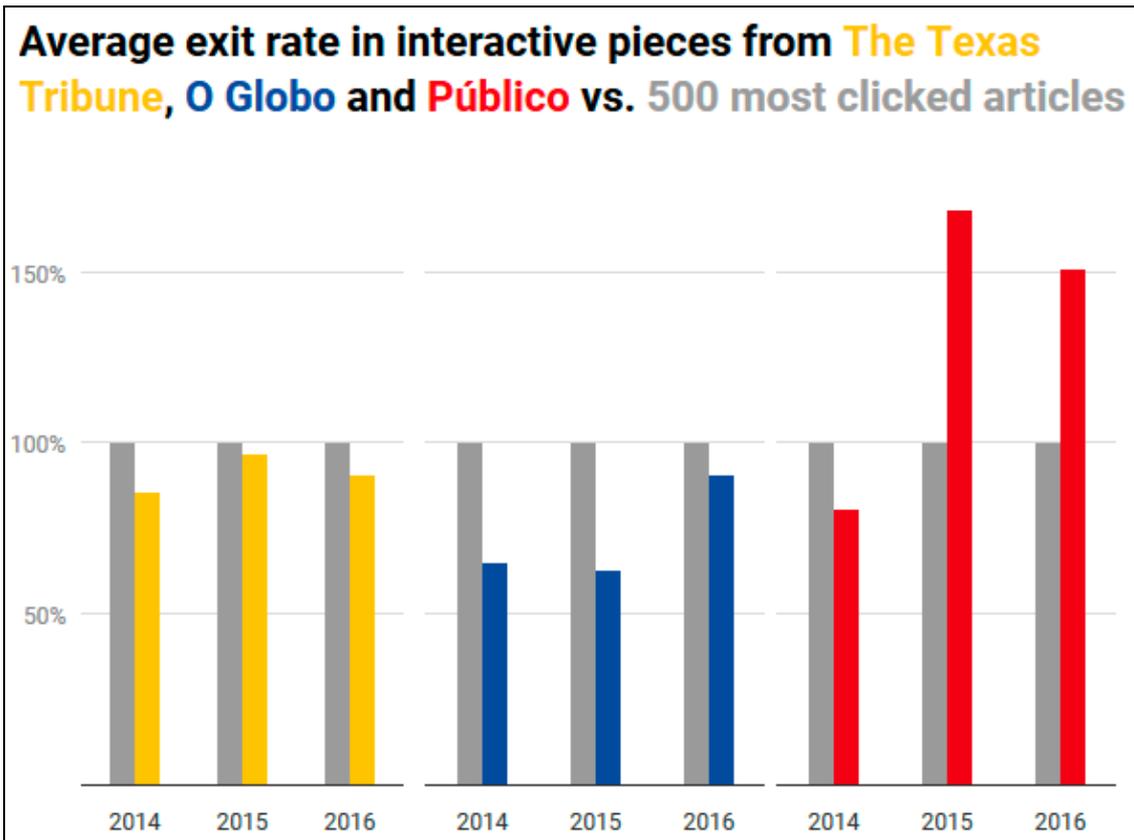


Figure 67 – Average exit rate. Most read news are represented as 100%. Interactive pieces’ results are relative to that value.

In comparison to related news stories, half of the data journalism pieces reported a lower exit ratio than both of their related articles. 21 of the 30 interactive pieces from *Público* registered a lower exit rate than at least 1 of the related stories. *O Globo* had 20 and *The Texas Tribune*, 19 of the data visualizations that the proportion of exit rate is lower than at least 1 of the related stories.

It may be important to clarify the difference between bounce rate and exit rate:

A bounce occurs when a user comes to your website and only looks at a single page before leaving the site. When a visitor bounces, he does not visit any other pages on your site, nor does he interact with anything on the single page viewed. The exit rate does not take into consideration how many pages the visitor looked at. It only looks at exits compared to total visits. % Exit reports how many people leave your site from a particular page or from the site overall. (Mokalis & Davis, 2018, pp. 222–223)

In other words, bounce rate refers to the first page a user lands on when entering a website, while exit rate concerns the last page a visitor is on before leaving the website. Our findings show that, in general, those who visit data journalism pieces are more interested in continuing to navigate the website than those who access the most read news or the related articles. Keeping the user moving through the website is one of the main goals of news organizations.

This behavior of proceeding to another content piece from the website is called recirculation: “percentage of the audience that has engaged with a particular piece of content (article, video, etc.) who proceeds to engage with another piece of content” (Cherubini & Nielsen, 2016, p. 35). Here we are using the exit rate to measure recirculation, but web analytics tools such as Chartbeat report recirculation rate by default (Petre, 2015). Although recirculation is not among the metrics used regularly by newsrooms (ICFJ, 2017), it has gained increasing importance in news organizations like *Vox Media*, *The Guardian* and the *Financial Times* (Cherubini & Nielsen, 2016; B. Merritt, 2017; Schmidt, 2018).

Finally, it is also important to put some data into context. *The Texas Tribune* interactive pieces “Government Salaries Explorer”, “Ethics Explorer”, “See which counties in Texas Trump and Clinton won”, and “Ballpark Figures”, for example, registered a low average time on page, but also had very low exit rates. In these specific data visualizations, the user requests information on a page and gets the answer in a second page. The data collected suggest that these interactive pieces were very successful in its proposal.

7.7. A quarter of the data journalism pieces are accessed again and again

38 of the 90 interactive pieces registered higher pageviews/unique pageviews ratio than the average of the 500 most read news stories. That is, in 42% of the data journalism pieces the user came back during the same session in a higher proportion than those who accessed the most read news.¹¹⁹ The ratio goes to 57% when facing related news articles.

¹¹⁹ A session ends in Google Analytics after 30 minutes of inactivity or at midnight (Mokalis and Davis 2018).

25 of the 30 interactive pieces from *O Globo* showed higher pageviews/unique pageviews ratio than the average of the 500 most clicked news, while only 2 of the data visualizations from *Público* reported a proportion higher than the most read articles; *The Texas Tribune* had 11. In comparison to related news stories, 67% of *The Texas Tribune* and *O Globo* data journalism pieces registered a higher ratio. *Público* had 36% of the interactive pieces with a higher ratio than the related news.

Nevertheless, according to the GOV.UK website, a high rate is above 1.4.¹²⁰ 24 of the 90 interactive pieces presented pageviews/unique pageviews proportion higher than 1.4. That is, in 27% of the data journalism pieces, the rate of users coming back to the data visualization is meaningful. There are 3 interactive pieces from *The Texas Tribune*, another 3 from *Público* and 1 from *O Globo* that recorded a rate higher than 2. *Público*'s "O ranking das escolas 2014" registered the incredible rate of 17.5. Unfortunately, it is not possible to determine the reason for the high return on these data visualization. Among the 180 related news, only 3 from *Público* recorded a pageviews/unique pageviews proportion higher than 1.4.

Still comparing to the related news, over the years the proportion of returning visitors is lower on most data journalism pieces. Between the publication date and December 31, 2017, 55% of the interactive pieces reported a higher ratio of new visitors than their related news stories.¹²¹ As discussed previously, the lifetime of related news is quite short. The majority gain almost no click after a few months of their publication, unlike some interactive pieces, which come to collect more pageviews years after they have been published. This helps to understand why more than half of the interactive pieces registered more visitors than their related news, and almost a third had a higher ratio of return visitors than their related articles.

7.8. Select is the most commonly used feature in data journalism pieces

Select was the most used interaction technique, introduced in 65% of the interactive pieces, followed by inspect, present in 49 of the 90 data journalism pieces analyzed, and filter, featured in 39 of the data visualizations. Select highlights

¹²⁰ Available at: <https://www.gov.uk/guidance/content-design/data-and-analytics>. Accessed on May 25, 2019.

¹²¹ According to Google Analytics, the user that visited the website within the past two years and returns from the same device is marked as a returning visitor.

something as interesting, to keep track of it, while inspect allows the user to get more information by clicking on an element. Filter shows to the user some data conditionally (Boy et al., 2015; Figueiras, 2015).

Select is the most frequent interaction technique in *Público*, appearing in 28 of the 30 data journalism pieces, although it is featured in 11 of *The Texas Tribune*'s data visualizations. Inspect was the most used function in *O Globo*, 21 of the interactive pieces, while 13 of *The Texas Tribune*'s data visualization introduced this interaction technique. Filter was the most common function used by *The Texas Tribune*, featuring half of the data journalism pieces, but appearing in a third of *Público*'s interactive pieces. Figure 68 shows the most used interaction techniques in each media outlet.

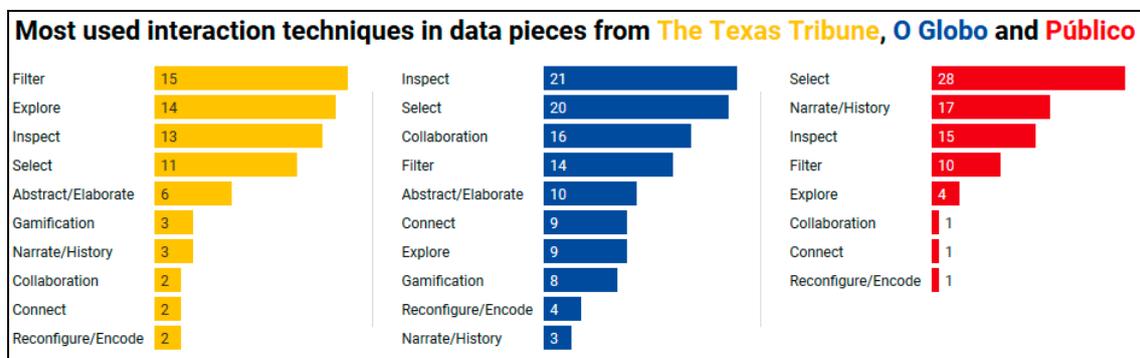


Figure 68 – Interaction techniques most frequent in the data journalism pieces analyzed.

These results are in line with previous observations. Select, inspect, and filter, along with abstract/elaborate, are considered some of the simplest interaction techniques and appear also as the most used in other studies. Loosen, Reimer, and De Silva-Schmidt (2017) analyzed 179 data projects nominated for the Data Journalism Award between 2013 and 2015 and zoom/details on demand and filter were the most used interaction techniques.¹²² Focusing on 26 Canadian stories that were finalists in 3 data journalism awards, Young, Hermida and Fulda (2018) found that inspect and filter were the most used functions. Ojo and Heravi (2018) analyzed 44 cases of award-winning data journalism work, comprising winning entries of the Data Journalism Award from 2013 to 2016, and found out that filter and select were among the most used functions. Finally, Stalph (2018) analyzed data-driven stories published by

¹²² Zoom/details on demand in our classification corresponds to inspect and abstract/elaborate.

European quality news media and found that aside from maps, most of the articles offered basic interactive features.

Scholars studying data visualizations from Canada, the UK, and Germany found that a great majority of data projects contain maps (M. Knight, 2015; Loosen et al., 2017; Stalph, 2018; Tabary et al., 2016; Young et al., 2018). 46% of *O Globo's* interactive pieces and 40% of *Público's* data journalism pieces have maps. On the other hand, 23% of *The Texas Tribune's* data visualizations include maps.

Some scholars point out the relationship between the predominance of interaction techniques such as inspect, filter, and abstract/elaborate and the use of maps, since most of the maps are created with free tools such as Google Maps and these features are part of the default functionality (Young et al., 2018). Indeed, most maps from *The Texas Tribune* and *O Globo* was made from Mapbox and OpenStreetMap.¹²³ The maps from *Público* interactive pieces were developed internally.

By analyzing the top 3 most clicked interactive piece from each news organization in 2014, 2015 and 2016, there is a map in 55% of the data journalism pieces. A third of the most read *The Texas Tribune* data visualizations between 2014 and 2016 contain a map. Same in *O Globo*, while in *Público* all 3 data visualizations with more pageviews have maps.

Still scrutinizing only for the data visualization with more pageviews each year, filter is the most commonly used interaction technique, present in 6 out of 9 data journalism pieces, followed by select, introduced in 5 interactive pieces. Select is present in 2 of the 3 most read *The Texas Tribune* data visualizations, explore and filter are featured in all 3 most clicked data journalism pieces from *O Globo*, and select and filter are displayed in all 3 *Público* interactive pieces with more pageviews.

Explore – the interaction technique featured in the most clicked *O Globo* interactive piece in 2014, 2015 and 2016 – displays different or specific subsets from the data and is considered a sophisticated interactive feature. This feature is present in 27 of the 90 data visualizations, almost a third. Functions like explore have a more

¹²³ In fact, most *O Globo* maps were made in Tableau, powered by Mapbox.

“reader-driven approach” (Segel & Heer, 2010, p. 1146), inviting “readers to discover the underlying data-set themselves, find data points of interest and draw their own conclusions” (Young et al., 2018, p. 6), in this manner assisting them to gain insight in the process (Dove & Jones, 2012).

All 10 interaction techniques categorized in Chapter III were found in the analyzed data journalism pieces. Reconfigure/encode was the less featured, introduced in 8 of the 90 data visualizations, followed by gamification, present in 11 interactive pieces.

7.9. Politics/governance is the most covered subject among data journalism pieces

News stories covering politics/governance were more likely to contain data elements. Counting only the topics that were the main subject, politics/governance led with 26 of the 90 interactive pieces, followed by social issues, the theme of 15 of the data visualizations, and education and sports, each one being the focus of 10 data journalism pieces.

Politics/governance is the most frequent main subject in *O Globo’s* interactive pieces, the focus of 13 out of 30 data visualizations. The majority of *The Texas Tribune’s* data journalism pieces also cover politics/governance, 11 out of 30 interactive articles. Economy/business is the most recurrent leading topic of *Público’s* data visualizations, the focus of 5 data journalism pieces. Figure 69 shows the most frequent topics in each media outlet.

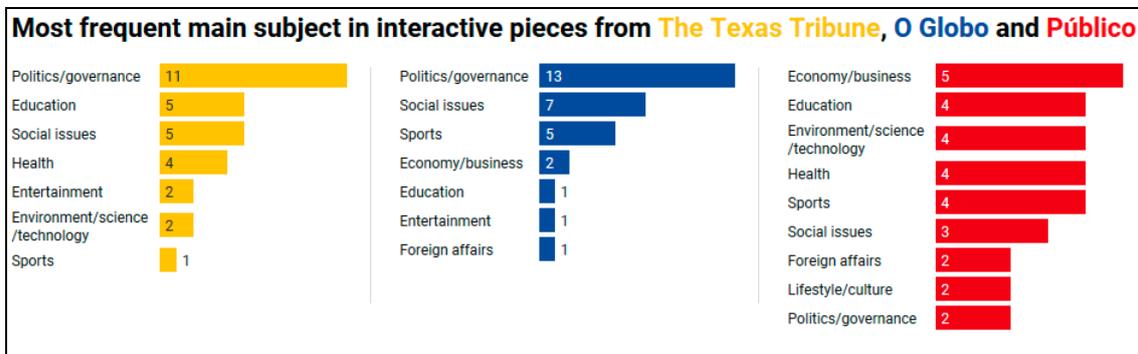


Figure 69 – Topics most frequent in the data journalism pieces analyzed.

The findings showing that politics/governance and social issues are the most frequent main subject of data journalism pieces are in line with previous studies.

Tandoc and Oh (2017) examined 260 interactive articles published in *The Guardian's Datablog* and found that the most frequent topic was politics. Loosen, Reimer, and De Silva-Schmidt (2017) analyzed 179 data projects nominated for the Data Journalism Award between 2013 and 2015 and almost half of the articles cover a political topic and more than a third deal with social issues. Analyzing 106 data-driven stories published in the UK by 15 mainstream news media, Knight (2015) found that stories covering social issues were dominant. Focusing on 26 Canadian stories that were finalists in 3 data journalism awards, Young, Hermida and Fulda (2018) found that social issues were among the dominant topics. Finally, Stalph (2018) analyzed data-driven stories published by European quality news media and found that most of the articles are about politics.

Although politics/governance, social issues, education, and sports corresponded to two-thirds of the data journalism pieces leading topics, all 10 topics categorized in Chapter III were the main subjects of at least 2 interactive pieces. This is the case of lifestyle/culture, the focus of 2 *Público* data visualization. *Público* was the most diversified news organization among the analyzed, with data journalism pieces with 9 different main themes. It was also the newspaper that recorded the shortest distance between the most recurrent and less frequent main subject ($SD=1.12$).¹²⁴

By analyzing the top 3 most clicked interactive piece from each news organization in 2014, 2015 and 2016, politics/governance is the leading topic in 55% of the data journalism pieces. Two-thirds of the most read *The Texas Tribune* data visualizations between 2014 and 2016 cover politics/governance. Same in *O Globo*, while in *Público* education, lifestyle/culture, and politics/governance were the main subjects in respective years among the data visualizations with more pageviews.

7.10. Data journalism pieces are appealing to international audiences

Two-thirds of the data journalism pieces analyzed registered more unique pageviews from outside the country from which they were published than at least 1 of the related articles. Fourteen of 30 interactive pieces from *The Texas Tribune* had more hits from outside the United States than their related news stories. Same in *Público*: almost half of the data journalism pieces reported more audience from outside

¹²⁴ *The Texas Tribune* registered $SD=3.35$ and *O Globo* had $SD=4.50$.

Portugal than their respective related news articles. In *O Globo* they were 38% of the data visualizations. Recalling that there are no data in this metric for “Os medalhistas brasileiros” and 9 of the 60 related news from *O Globo* and other 3 related articles from *Público*.

In *The Texas Tribune*, among the 30 data journalism pieces only “Texas Hospitals Face Penalties Over Infections” registered 100% of hits from the United States, while 10 out of 60 related stories reported 100% U.S. audience. *Público* reported 5 data visualizations with the audience totally originating from the Portuguese territory and a quarter of the related news articles with all unique pageviews from Portugal. 31% of *O Globo*’s interactive pieces had 100% Brazilian audience, while 43% of the related stories registered all accesses from Brazil.

Público’s “Taxa de sobrevivência de câncros em Portugal” and *O Globo*’s “A seleção dos sonhos da Copa do Mundo 2014” registered half of the unique pageviews from outside the country from which they were published. *The Texas Tribune* interactive piece with more hits from outside the United States was “Government Salaries Explorer” with 30% of the accesses from elsewhere.

How to reach a broader audience has always been a concern in news organizations, but it became a daily obsession in newsrooms with the rise of web analytics (see, for instance, Moyo, Mare, and Matsilele 2019; Blanchett Neheli 2018; Christin 2018; N. S. Cohen 2018; Nguyen and Vu 2018; Vieira 2018; Cherubini and Nielsen 2016; Stockleben and Lugmayr 2016; Petre 2015; Tandoc 2014; Usher 2013; Anderson 2011b; Graves, Kelly, and Gluck 2010; Schaudt and Carpenter 2009; MacGregor 2007). The results presented in this section imply that data visualizations have a more worldwide appeal than traditional news. Even *The Texas Tribune*, with a more local perspective, focused on public policy, politics and the government in the state of Texas, has a good audience share of its interactive pieces outside the country. Hence, these findings suggest that investing in interactive pieces may be a valuable asset to get a more diverse audience.

Nevertheless, it is important to stress that geographic data may not be 100% reliable, since users may be accessing the page using a virtual private network, which hides the user’s real location. These limitations reach all case studies in a cross-

sectional way, since we used the same tool and the same parameters to collect the data in the three news organizations.

7.11. Data journalism pieces shift from the desktop dominance to mobile devices

Between 2014 and 2016 interactive pieces and related news stories were mostly accessed by desktop computers. 70% of the data journalism pieces registered the desktop as the most used device to access the articles. The proportion is even higher in the related news stories: 82%. However, the behavior in the news organizations is not homogeneous. Recalling that in 2014 we scrutinized *O Globo's* data only from desktops, and so they are not part of these results. There are also no data for all the metrics in this section for *O Globo's* "Os medalhistas brasileiros" and 9 of the 60 related news from *O Globo* and other 3 related articles from *Público*.

Público interactive pieces registered the highest proportion of unique pageviews that mostly used desktops to access the page: 87%. The ratio of the related articles was a little lower: 84%. *O Globo's* data journalism pieces between 2015 and 2016 reported the lowest percentage of accesses from desktops: 68% of the data visualizations were mainly visited in a desktop. The proportion of the related news stories was a little lower: 65%. 77% of the interactive pieces from *The Texas Tribune* were predominantly accessed from desktops and the ratio of the related article was higher: 88%.

The results show that data journalism pieces from the legacy news media were slightly more accessed from computers than their respective related news stories, while in *The Texas Tribune* the access to the interactive pieces was achieved in a more diverse way than in the related news.

It is interesting to note that while in 2014 none of *The Texas Tribune* or *Público* data journalism pieces were mostly accessed by mobile devices, in 2016, 6 out of 10 *The Texas Tribune* and 4 *Público* data visualizations registered the largest flow of traffic from that device. Among the 29 interactive pieces of *The Texas Tribune*, *O Globo* and *Público* analyzed in 2015, only 3 recorded accesses mostly from mobile devices. In 2016 they were 14 of the 30 data journalism pieces, almost 5 times more than in the previous year. This is in line with recent research that shows that mobile devices have

become the most common way to get news in the United States,¹²⁵ Brazil (Newman, Fletcher, Kalogeropoulos, Levy, & Nielsen, 2018) and Portugal (Cardoso et al., 2018).

Nevertheless, this behavior was not observed with the same intensity in the related news stories. While 2 of the 20 *The Texas Tribune* related news articles were mainly visited from mobile devices in 2014, only 3 of the 20 related news in 2016 registered the largest flow of traffic from that device. In *Público* 11% of the related news stories were predominantly accessed from mobile in 2014 and in 2016 this proportion rose to 25%.

7.12. Once the main source for data journalism pieces, Google is dethroned by Facebook

Interactive pieces and related news stories were mostly reached by Google. 52% of the data journalism pieces registered the search engine as the place where most users were before visiting the page. The proportion is even higher in the related news stories: 56%. Recalling that there is also no data for this metric for *O Globo's* "Os medalhistas brasileiros" and 9 of the 60 related news from *O Globo* and other 3 related articles from *Público*.

The Texas Tribune's data journalism pieces registered the highest proportion of unique pageviews that mostly used Google to visit the page: 60%. The ratio of the related articles was higher: 75%. *O Globo* interactive pieces reported the lowest percentage of accesses coming from Google: 38% of the interactive pieces were mainly visited arriving from Google. The proportion of the related news stories was higher: 41%. 53% of the data visualizations from *Público* were predominantly accessed after the users visited Google and the ratio of the related article was lower: 49%.

Looking at the results of each year, it is possible to see that the percentage of *The Texas Tribune's* interactive pieces that registered Google as the location where most users were before visiting the page dropped from 80% in 2014 to 30% in 2016. In a similar move, if in 2014 half of *O Globo's* data visualizations reported the search engine as the place where most visitors were before entering the page, in 2016 it was

¹²⁵ Available at <https://www.pewresearch.org/fact-tank/2018/07/17/use-of-mobile-devices-for-news-continues-to-grow-outpacing-desktops-and-laptops/>. Accessed on March 25, 2019.

a third. Related news stories from these news organizations also reported a drop in the volume of traffic coming from Google.

Público reported a different motion, increasing from 40% in 2014 to 50% in 2016 the proportion of interactive pieces that registered Google as the location where most users were before visiting the page. The related news articles dropped from 72% to 60%.

Our findings show that while Google's dominance fell in 2 of the 3 news media analyzed, Facebook's influence as the location where most users were before visiting the page more than doubled in interactive pieces and related news stories from all 3 news organizations.

Among the 30 interactive pieces of *The Texas Tribune*, *O Globo* and *Público* analyzed in 2014, only 5 recorded Facebook as the location where most users were before visiting the page. In 2016 they were 12, making the social network the most recurrent place where most data journalism pieces visitors were before entering the page. According to a recent report, this behavior goes against the expectation of the majority of editors, since publishers are “looking to diversify away from Facebook and towards other platforms” (Newman, 2019, p. 17).

7.13. Over the years Chrome reinforces its sovereignty among data journalism pieces

Interactive pieces and related news stories were mostly accessed by visitors using Chrome. 83% of the data visualizations registered the Google browser as the most used to access the articles. The proportion is lower in the related news stories: 78%. Recalling that there are also no data for this metric for *O Globo's* “Os medalhistas brasileiros” and 9 of the 60 related news from *O Globo* and other 3 related articles from *Público*.

O Globo's interactive pieces registered the highest proportion of unique pageviews that mostly used Chrome to access the page: 93%. The ratio of the related articles was slightly higher: 94%. *The Texas Tribune's* data journalism pieces reported the lowest percentage of accesses from the Google browser: 77% of the data visualizations were mainly visited using Chrome. The proportion of the related news

stories was lower: 67%. 80% of the data visualizations from *Público* were predominantly accessed using Google Chrome and the ratio of the related news articles was lower: 77%.

Looking at the results of each year, it is possible to see that Chrome's domain has been strengthened throughout the years. Among the 30 interactive pieces of *The Texas Tribune*, *O Globo* and *Público* analyzed in 2014, 73% recorded the Google browser as the principal one to access the page. In 2016, it had 80% of the data visualizations.

Público is largely responsible for this rise. In 2014, 6 of the 10 interactive pieces were mostly accessed from Chrome, while in 2016 it was the most used for 9 data journalism pieces from the Portuguese newspaper. *O Globo* kept the ratio unchanged: 9 in 2014 and in 2016. On the other hand, interactive pieces from *The Texas Tribune* registered a decline in Chrome as the primary browser: from 7 in 2014 to 6 in 2016.

Globally, the related news stories also recorded a boost of Chrome's domain as the main navigator: from 71% in 2014 to 81% in 2016. Related articles from *Público* and *O Globo* reported a rise in the Google browser's usage: from 55% to 95% and from 95% to 100%, respectively. On the other hand, *The Texas Tribune* registered a drop in the proportion of visitors who used predominantly Chrome to access related news: from 60% in 2014 to 50% in 2016.

7.14. Final considerations and further research

By using Google Analytics data to measure the performance of data journalism pieces and the behavior of its visitors in news organizations from 3 different countries – United States, Brazil, and Portugal – we performed unprecedented work. We compared the performance of interactive pieces with the most clicked news stories from each newspaper so that we could have an overview of the interest in data visualizations in comparison to the most accessed content in their respective websites. Each interactive piece was also contrasted to 2 similar news articles on the same subject and published in the same period. In this manner, it was possible to measure the performance of data journalism pieces with contents that could also interest the visitor of the interactive pieces and vice versa.

Taking as case studies a nonprofit digital-born media and a couple of legacy media, it was possible to observe the performance of data journalism pieces in news organizations with different goals and from different countries. And we tried to identify if there are performance patterns among the interactive pieces from these different news media.

Our findings show that *The Texas Tribune* interactive pieces have more pageviews than their related news stories and reported also a better performance than the average of the 500 most clicked news in the year they were published. Data journalism pieces from *O Globo* and *Público* did not achieve the same performance facing the average of its 500 most read articles, but it should be recognized that while *The Texas Tribune's* data visualizations were widely publicized, the interactive pieces of the legacy media were commonly complementary to other news stories and did not receive the same kind of prominence on their websites.

Also, *The Texas Tribune* is focused on public policy, politics and the government in the state of Texas, therefore the variation of topics is not so great, unlike the legacy media, which are general-interest newspapers. When comparing with related news stories, a third of *Público's* interactive pieces and 56% of *O Globo's* data journalism pieces registered more pageviews.

27% of all 90 interactive pieces reported a high proportion of pageviews divided by the unique pageviews, i.e. in these data visualizations the users returned to the page several times in the same session. And 21% of the pieces registered growth in hits over the years, while fewer than 2% of related news articles recorded more accesses in the years following publication.

These groups of results considering pageviews – “metric of success” (Groves & Brown, 2011, p. 16) – show that data journalism pieces have performed well when they are published and continue to be visited over the years even without any updates or highlights in the homepage, as in *O Globo* and *Público*.

Examining entrances, time on page, and exit rate metrics, our findings show that users typically (1) visited a page from the website before entering an interactive piece; (2) spent time above average in the interactive piece; (3) and then moved on to

another page from the website. This behavior of visiting a page of content and then proceeding to another piece of content is known as recirculation and has gained increasing importance in news organizations (Schmidt 2018; B. Merritt 2017; Cherubini and Nielsen 2016). The recirculation and higher average time on page occur in the data journalism pieces in a higher proportion than in the average of the most clicked news and in the related articles.

Although the interactive pieces registered a low rate of entrances (especially in *O Globo* and *Público*), those who landed on the website in a data visualization mostly left the website without visiting any other page. Unfortunately, from the data we collected it is not possible to identify the reasons for the bounce. But our findings showing that users spend more time on interactive pieces than on the most read news or the related articles suggest that readers interact with the data journalism pieces before leaving the site.

Select was the most used interaction technique, introduced in almost two-thirds of the interactive pieces, followed by inspect, present in more than half of the 90 data journalism pieces analyzed. Select and inspect are among the simplest interaction techniques and appear also as the most used in other studies evaluating data visualizations (Loosen et al., 2017; Stalph, 2018; Young et al., 2018).

Politics/governance was the most frequent main subject of the interactive pieces, 28%, followed by social issues, the theme of 17% of the data journalism pieces. These findings are in line with previous studies classifying data visualization (M. Knight, 2015; Loosen et al., 2017; Stalph, 2018; Tandoc & Oh, 2017; Young et al., 2018).

Finally, interactive pieces attracted more visitors from outside the country of origin of the data piece than the average of the most clicked news and the related articles. The device, traffic source, and browser used to access the data visualization was not much different from those used by users who accessed the related news articles.

We believe that this body of findings brings contributions to the study in the fields of data journalism and audience data and metrics in online journalism. There is very little research that uses actual Google Analytics data (or any other web analytics

for this matter), and even less that uses it to investigate data journalism pieces performance. Likewise, the way the methodology and the results of this work are structured facilitates the reproduction and replicability of this research.

The approach adopted in this work has some restraints. Google Analytics reports – and web analytics data in general – are richer and more enlightening when combined with qualitative data. By adopting a more quantitative bias, it was not possible to know how the user interacts with the data visualizations, or if the interactive pieces were understood, or whether they are efficient. And consequently, it was not possible to realize, for instance, if the interactive pieces that best communicate are those with better performance.

Investigations that focus and combine web analytics data with usability testing to understand how the user interacts with the data visualization are critical. The production of data journalism pieces in newsrooms has grown considerably in the last years and tends to continue expanding. Journalists will only be able to create more comprehensive and memorable interactive pieces if they know how these data visualizations are received and evaluated by the audience.

Bibliography

- Abbate, J. (2000). *Inventing the Internet*. Cambridge, Mass.: MIT Press.
- Abras, F. B. (2018). *How Brazilian and Portuguese newsrooms use information visualization in journalism: A qualitative comparative analysis* (Master Thesis, Universidade Nova de Lisboa). Retrieved from <https://run.unl.pt/handle/10362/49133>
- Abreu, A. P. C. de. (2017). *A televisão para lá do televisor: Novas práticas, novos consumos?* (Master Thesis, Universidade Católica Portuguesa). Retrieved from <https://repositorio.ucp.pt/handle/10400.14/21868>
- Aitamurto, T., Sirkkunen, E., & Lehtonen, P. (2011). Trends In Data Journalism. *Next Media*. Espoo, VTT. Retrieved from http://virtual.vtt.fi/virtual/nextmedia/Deliverables-2011/D3.2.1.2.B_Hyperlocal_Trends_In%20Data_Journalism.pdf
- Alexandre, I. (2014). *Jornalismo de Dados: O estado da arte nos jornais generalistas diários em Portugal* (Master Thesis, Universidade Nova de Lisboa). Retrieved from <https://run.unl.pt/handle/10362/13615>
- Alexandre, I. (2016). Promoting Insight: A Case Study of How to Incorporate Interaction in Existing Data Visualizations. *2016 20th International Conference Information Visualisation (IV)*, 203–208. <https://doi.org/10.1109/IV.2016.15>
- Alfred, R. (2010, November 4). Nov. 4, 1952: Univac Gets Election Right, But CBS Balks. *Wired*. Retrieved from <https://www.wired.com/2010/11/1104cbs-tv-univac-election/>
- Alhlou, F., Asif, S., & Fettman, E. (2016). *Google Analytics breakthrough: From zero to business impact*. Hoboken, New Jersey: Wiley.

- Alpert, L. I., & Mullin, B. (2019, January 24). BuzzFeed to Cut 15% of Its Workforce. *Wall Street Journal*. Retrieved from <https://www.wsj.com/articles/buzzfeed-to-cut-15-of-its-workforce-11548286211>
- Anderson, C. W. (2011a). Between creative and quantified audiences: Web metrics and changing patterns of newswork in local US newsrooms. *Journalism: Theory, Practice & Criticism*, 12(5), 550–566.
<https://doi.org/10.1177/1464884911402451>
- Anderson, C. W. (2011b). Notes Towards an Analysis of Computational Journalism. *SSRN Electronic Journal*. <https://doi.org/10.2139/ssrn.2009292>
- Anderson, C. W. (2015). Between the Unique and the Pattern: Historical tensions in our understanding of quantitative journalism. *Digital Journalism*, 3(3), 349–363.
<https://doi.org/10.1080/21670811.2014.976407>
- Anderson, C. W. (2018). *Apostles of certainty: Data journalism and the politics of doubt*. New York, NY: Oxford University Press.
- Anderson, C. W., Bell, E. J., & Shirky, C. (2014). Post Industrial Journalism: Adapting to the Present. *Columbia University*. <https://doi.org/10.7916/d8n01js7>
- Andrade, M. da C. (2009). *Bruscamente no verão passado, a reforma do código de processo penal: Observações críticas sobre uma lei que podia e devia ter sido diferente*. Coimbra: Coimbra Editora.
- Antunes, A. P. S. (2017). *Jornalismo mensurado: Uma investigação sobre os impactos dos sistemas de medição de audiência em sites de notícias*. Retrieved from <http://tede2.pucrs.br:80/tede2/handle/tede/7393>
- Appelgren, E. (2018). An Illusion of Interactivity: The paternalistic side of data journalism. *Journalism Practice*, 12(3), 308–325.
<https://doi.org/10.1080/17512786.2017.1299032>

- Appelgren, E., & Nygren, G. (2014). Data Journalism in Sweden: Introducing new methods and genres of journalism into “old” organizations. *Digital Journalism*, 2(3), 394–405. <https://doi.org/10.1080/21670811.2014.884344>
- Arnold, M. (1887). Up to Easter. In J. Knowles (Ed.), *Nineteenth Century 21* (pp. 629–643). London.
- Ausserhofer, J. (2015). “Die Methode liegt im Code”: Routinen und digitale Methoden im Datenjournalismus. <https://doi.org/10.17174/dcr.v2.5>
- Baldessar, M. J. (2003). *A mudança anunciada: O cotidiano dos jornalistas com o computador na redação*. Florianópolis, SC: Insular.
- Bales, K. (1999). Popular Reactions to Sociological Research: The Case of Charles Booth. *Sociology*, 33(1), 153–168. <https://doi.org/10.1177/S0038038599000085>
- Barbosa, Susana. (2012). Agentes de inovação, renovação e de reconfiguração para o jornalismo em tempos de convergência. *Cibercomunicación*, 1(1), 1–8.
- Barbosa, Suzana. (2007). Jornalismo digital em base de dados (JDBD): Um paradigma para produtos jornalísticos digitais dinâmicos. *Doctoral thesis*. Retrieved from <http://repositorio.ufba.br/ri/handle/ri/11299>
- Barroso, R. G., Leite, A. L., Manita, C., & Nobre, P. (2011). Between public agenda and the emergence of intervention programmes: Sexual offenders within the Portuguese context. *Sexual Offender Treatment*, 6(2), 1–10.
- Bastos, H. (2010a). *Ciberjornalistas em Portugal: Práticas, papéis e ética*. Lisboa: Livros Horizonte.
- Bastos, H. (2010b). *Origens e evolução do ciberjornalismo em Portugal: Os primeiros quinze anos (1995-2010)*. Porto: Edições Afrontamento.
- Batsell, J. (2017). Nonprofit model development. In M. Ferrier & E. Mays (Eds.), *Media innovation and entrepreneurship*. Retrieved from

<https://open.bccampus.ca/find-open-textbooks/?uuid=a2f103a2-acb0-4764-bf5c-3be1cb66e663&contributor=&keyword=&subject=>

Bayatrizi, Z. (2008). *Life sentences: The modern ordering of mortality*. Toronto: Univ. of Toronto Press.

Baylen, J. O. (1972). The 'New Journalism' in Late Victorian Britain. *Australian Journal of Politics & History*, 18(3), 367–385. <https://doi.org/10.1111/j.1467-8497.1972.tb00602.x>

Beasley, M. (2013). *Practical web analytics for user experience: How analytics can help you understand your users*. Amsterdam: Elsevier/Morgan Kaufmann.

Bederman, G. (1989). "The Women Have Had Charge of the Church Work Long Enough": The Men and Religion Forward Movement of 1911-1912 and the Masculinization of Middle-Class Protestantism. *American Quarterly*, 41(3), 432. <https://doi.org/10.2307/2713149>

Beniger, J. R., & Robyn, D. L. (1978). Quantitative Graphics in Statistics: A Brief History. *The American Statistician*, 32(1), 1–11. <https://doi.org/10.1080/00031305.1978.10479235>

Bennett, W. L., Lawrence, R. G., & Livingston, S. (2008). *When the press fails: Political power and the news media from Iraq to Katrina*. Chicago, Ill.: University of Chicago Press.

Beville, H. M. (1988). *Audience ratings: Radio, television, and cable* (rev. ed). Hillsdale, NJ [u.a]: Erlbaum.

Bingham, P., Verlander, N. Q., & Cheal, M. J. (2004). John Snow, William Farr and the 1849 outbreak of cholera that affected London: A reworking of the data highlights the importance of the water supply. *Public Health*, 118(6), 387–394. <https://doi.org/10.1016/j.puhe.2004.05.007>

- Blanchett Neheli, N. (2018). News by Numbers: The evolution of analytics in journalism. *Digital Journalism*, 6(8), 1041–1051.
<https://doi.org/10.1080/21670811.2018.1504626>
- Boczkowski, P. J., & Peer, L. (2011). The Choice Gap: The Divergent Online News Preferences of Journalists and Consumers. *Journal of Communication*, 61(5), 857–876. <https://doi.org/10.1111/j.1460-2466.2011.01582.x>
- Booth, C. (1897). *Life and labour of the people in London*. London and New York: Macmillan and Co.
- Borchard, G. A. (2019). *A narrative history of the American press*.
- Borges-Rey, E. (2016). Unravelling Data Journalism: A study of data journalism practice in British newsrooms. *Journalism Practice*, 10(7), 833–843.
<https://doi.org/10.1080/17512786.2016.1159921>
- Borges-Rey, E. (2017). Towards an epistemology of data journalism in the devolved nations of the United Kingdom: Changes and continuities in materiality, performativity and reflexivity. *Journalism: Theory, Practice & Criticism*, 146488491769386. <https://doi.org/10.1177/1464884917693864>
- Boy, J., Detienne, F., & Fekete, J.-D. (2015). Storytelling in Information Visualizations: Does it Engage Users to Explore Data? *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems - CHI '15*, 1449–1458.
<https://doi.org/10.1145/2702123.2702452>
- Boyer, B. (2008, July 2). The hacker journalist: In whom programming and prose intersect. Retrieved March 11, 2019, from Hacker journalist website:
<https://hackerjournalist.net/2008/07/02/the-hacker-journalist-in-whom-programming-and-prose-intersect/>

- Bradshaw, P. (2014). Data Journalism. In L. Zion & D. Craig (Eds.), *Ethics for Digital Journalists: Emerging Best Practices* (1st ed., pp. 202–219).
<https://doi.org/10.4324/9780203702567>
- Brasseur, L. (2005). Florence Nightingale's Visual Rhetoric in the Rose Diagrams. *Technical Communication Quarterly*, 14(2), 161–182.
https://doi.org/10.1207/s15427625tcq1402_3
- Bright, J., & Nicholls, T. (2014). The Life and Death of Political News: Measuring the Impact of the Audience Agenda Using Online Data. *Social Science Computer Review*, 32(2), 170–181. <https://doi.org/10.1177/0894439313506845>
- Briso, C. B. (2016, August 27). O GLOBO vence duas categorias no Prêmio GDA de Jornalismo. Retrieved February 27, 2019, from Jornal O Globo website:
<https://oglobo.globo.com/rio/o-globo-vence-duas-categorias-no-premio-gda-de-jornalismo-20003562>
- Broussard, M. (2018). *Artificial unintelligence: How computers misunderstand the world*. Cambridge, Massachusetts London: The MIT Press.
- Broussard, M., & Boss, K. (2018). Saving Data Journalism: New strategies for archiving interactive, born-digital news. *Digital Journalism*, 6(9), 1206–1221.
<https://doi.org/10.1080/21670811.2018.1505437>
- Bunce, M. (2015). Africa in the click stream: Audience metrics and foreign correspondents in Africa. *African Journalism Studies*, 36(4), 12–29.
<https://doi.org/10.1080/23743670.2015.1119487>
- Buzzard, K. S. (1990). *Chains of gold: Marketing the ratings and rating the markets*. Metuchen, NJ: Scarecrow Press.

- Buzzard, K. S. F. (2002). The Peoplemeter Wars: A Case Study of Technological Innovation and Diffusion in the Ratings Industry. *Journal of Media Economics*, 15(4), 273–291. https://doi.org/10.1207/S15327736ME1504_4
- Buzzard, K. S. F. (2003). James W. Seiler of the American Research Bureau. *Journal of Radio Studies*, 10(2), 186–201. https://doi.org/10.1207/s15506843jrs1002_4
- Buzzard, K. S. F. (2012). *Tracking the Audience: The Ratings Industry from Analog to Digital*. Taylor & Francis Inc.
- Buzzard, K. S. F. (2015). The Rise of Market Information Regimes and the Historical Development of Audience Ratings. *Historical Journal of Film, Radio and Television*, 35(3), 511–517. <https://doi.org/10.1080/01439685.2015.1052219>
- Cairo, A. (2017). *Nerd journalism: How data and digital technology transformed news graphics* (Ph.D. Thesis, Universitat Oberta de Catalunya). Retrieved from <http://www.tdx.cat/handle/10803/404809>
- Canavilhas, J., Satuf, I., Luna, D., & Torres, V. (2014). Jornalistas e tecnoatores: Dois mundos, duas culturas, um objetivo. *Esferas*, 0(5). <https://doi.org/10.19174/esf.v0i5.5690>
- Card, S. K., Mackinlay, J. D., & Shneiderman, B. (2007). *Readings in information visualization: Using vision to think* (Nachdr.). San Francisco, Calif.: Morgan Kaufmann.
- Cardoso, G., Paisana, M., & Martinho, A. P. (2018). *Digital News Report 2018: Portugal*. Retrieved from Reuters Institute for the Study of Journalism website: https://obercom.pt/wp-content/uploads/2018/09/DNR_PT_2018.pdf
- Carniello, M. F., Mota, H. A. da, & Santos, M. J. dos. (2017). PERFIL DE AUDIÊNCIA DE RÁDIO EDUCATIVA NO MUNICÍPIO DE TAUBATÉ: A MÍDIA NÃO COMERCIAL COMO DIFUSORA DE CONTEÚDO

- REGIONAL. *Revista Latinoamericana de Ciencias de la Comunicación*, 14(26). Retrieved from <http://www.alaic.org/revista/index.php/alaic/article/view/920>
- Carniello, M. F., & Santos, M. J. dos. (2016). Proposta metodológica para avaliação de perfil de audiência para veículos de comunicação públicos, comunitários e governamentais. *Intercom – Sociedade Brasileira de Estudos Interdisciplinares Da Comunicação*, 39. Retrieved from <http://portalintercom.org.br/anais/nacional2016/resumos/R11-0882-1.pdf>
- Carvalho, M. A. R. de. (2012). *Irineu Marinho: Imprensa e cidade*.
- Chartbeat. (2015). *Why Publishers Are Killing Pageviews to Capitalize on Reader Attention* [White Paper]. Retrieved from Chartbeat website: https://www.intime-media-services.net/app/download/20752434/attention_white_paper.pdf
- Cherubini, F., & Nielsen, R. K. (2016). Editorial Analytics: How News Media are Developing and Using Audience Data and Metrics. *SSRN Electronic Journal*. <https://doi.org/10.2139/ssrn.2739328>
- Cheruiyot, D., & Ferrer-Conill, R. (2018). “Fact-Checking Africa”: Epistemologies, data and the expansion of journalistic discourse. *Digital Journalism*, 6(8), 964–975. <https://doi.org/10.1080/21670811.2018.1493940>
- Chinoy, I. (2010). *Battle of the Brains: Election-Night Forecasting at the Dawn of the Computer Age*. Retrieved from <http://drum.lib.umd.edu/handle/1903/10504>
- Christin, A. (2014). *Clicks or Pulitzers? Web Journalists and Their Work in the United States and France* (Ph.D. Thesis, Princeton University). Retrieved from <http://dataspace.princeton.edu/jspui/handle/88435/dsp01wd375z524>
- Christin, A. (2018). Counting Clicks: Quantification and Variation in Web Journalism in the United States and France. *American Journal of Sociology*, 123(5), 1382–1415. <https://doi.org/10.1086/696137>

- Cicak, T., & Tynan, N. (2015). Mapping London's Water Companies and Cholera Deaths. *The London Journal*, 40(1), 21–32.
<https://doi.org/10.1179/0305803414Z.00000000057>
- Ciotta, R. (1996). Baby you should drive this CAR. *American Journalism Review*, 18(2), 34–40.
- Clifton, B. (2012). *Advanced Web metrics with Google Analytics* (3. ed). Hoboken, N.J: Wiley.
- Clifton, B. (2015). *Successful analytics: Gain business insights by managing Google Analytics*.
- Coddington, M. (2015a). Clarifying Journalism's Quantitative Turn: A typology for evaluating data journalism, computational journalism, and computer-assisted reporting. *Digital Journalism*, 3(3), 331–348.
<https://doi.org/10.1080/21670811.2014.976400>
- Coddington, M. (2015b). The wall becomes a curtain. In M. Carlson (Ed.), *Boundaries of Journalism* (1st ed., pp. 67–82; By S. C. Lewis).
<https://doi.org/10.4324/9781315727684-5>
- Cohen, N. S. (2018). At Work in the Digital Newsroom. *Digital Journalism*, 1–21.
<https://doi.org/10.1080/21670811.2017.1419821>
- Cohen, S., Hamilton, J. T., & Turner, F. (2011). Computational journalism. *Communications of the ACM*, 54(10), 66.
<https://doi.org/10.1145/2001269.2001288>
- Couldry, N., & Turow, J. (2014). Advertising, big data and the clearance of the public realm: Marketers' new approaches to the content subsidy. *International Journal of Communication*, 8, 1710–1726.

- Cox, J. B. (2014). News You Need to Know: Examining the Prioritization of News Content in Print and Online Publications. *Journal of Mass Communication & Journalism*, 04(08). <https://doi.org/10.4172/2165-7912.1000210>
- Cox, M. (2000). The development of computer-assisted reporting. *Newspaper Division of the Association for Education in Journalism and Mass Communication's Southeast Colloquium*. Presented at the University of North Carolina, Chapel Hill. University of North Carolina, Chapel Hill.
- Crouthamel, J. L., & Jackson, A. (1973). James Gordon Bennett, the “New York Herald”, and the Development of Newspaper Sensationalism. *New York History*, 54(3), 294–316.
- Cruz, H. de F. (2000). *São Paulo em papel e tinta: Periodismo e vida urbana, 1890 - 1915*. São Paulo: EDUC Imprensa Oficial SP.
- Cullen, M. J. (2016). *The statistical movement in early Victorian Britain: The foundations of empirical social research*.
- Cutroni, J. (2010). *Google analytics: Understanding visitor behavior* (1. ed). Beijing: O'Reilly.
- Daniel, A., & Flew, T. (2010). The Guardian Reportage of the UK MP Expenses Scandal: A Case Study of Computational Journalism. In F. Papandrea & M. Armstrong (Eds.), *Record of the communications policy and research forum 2010* (pp. 186–194). Sydney: Network Insight Pty. Ltd.
- Darnton, R. (1975). Writing News and Telling Stories. *Daedalus*, 104(2), 175–194. Retrieved from JSTOR.
- Davenport, L. D., Fico, F., & DeFleur, M. H. (2002). Computer-Assisted Reporting in Classrooms: A Decade of Diffusion and a Comparison to Newsrooms. *Journalism and Mass Communication Educator*, 57(1), 6–22.

- De Maeyer, J., Libert, M., Domingo, D., Heinderyckx, F., & Le Cam, F. (2015).
Waiting for Data Journalism: A qualitative assessment of the anecdotal take-up
of data journalism in French-speaking Belgium. *Digital Journalism*, 3(3), 432–
446. <https://doi.org/10.1080/21670811.2014.976415>
- Deegan, M. (2004). Hull-House Maps and Papers. *Sociology Department, Faculty
Publications*. Retrieved from
<http://digitalcommons.unl.edu/sociologyfacpub/321>
- Deegan, M. J. (1988). W.E.B. Du Bois and the women of hull-house, 1895–1899. *The
American Sociologist*, 19(4), 301–311. <https://doi.org/10.1007/BF02691827>
- Demir, M. (2012). The Role Played by Italy in Media History: Its Historical Problems
and Its Recents Innovations in Journalism and Television World. *Cross-Cultural
Communication*, 8(3), 8–16.
<http://dx.doi.org/10.3968/j.ccc.1923670020120803.2830>
- Dennis, E. E., & Rivers, W. L. (2017). *Other Voices: The New Journalism in America*.
London: Routledge.
- Desmond, M., & Western, B. (2018). Poverty in America: New Directions and Debates.
Annual Review of Sociology, 44(1), 305–318. <https://doi.org/10.1146/annurev-soc-060116-053411>
- Deuze, M. (2005). What is journalism?: Professional identity and ideology of journalists
reconsidered. *Journalism: Theory, Practice & Criticism*, 6(4), 442–464.
<https://doi.org/10.1177/1464884905056815>
- Diakopoulos, N. (2015). Algorithmic Accountability: Journalistic investigation of
computational power structures. *Digital Journalism*, 3(3), 398–415.
<https://doi.org/10.1080/21670811.2014.976411>

- Dick, M. (2014). Interactive Infographics and News Values. *Digital Journalism*, 2(4), 490–506. <https://doi.org/10.1080/21670811.2013.841368>
- Doig, S., & Cascais, F. (2011). Entrevista a Steve Doig. *Comunicação & Cultura*, 11: *Sagrado e modernidade*, 149–158.
- Dorling, D., Mitchell, R., Shaw, M., Orford, S., & Davey Smith, G. (2000). The Ghost of Christmas Past: Health effects of poverty in London in 1896 and 1991. *BMJ*, 321(7276), 1547–1551. <https://doi.org/10.1136/bmj.321.7276.1547>
- Douglas, G. H. (1999). *The Golden Age of the newspaper*. Westport, Conn.: Greenwood Press.
- Dove, G., & Jones, S. (2012). *Narrative Visualization: Sharing Insights into Complex Data*. Presented at the Interfaces and Human Computer Interaction (IHCI 2012), Lisbon, Portugal. Retrieved from <http://openaccess.city.ac.uk/1134/>
- DuBois, W. E. B. (2010). *The Philadelphia Negro: A Social Study*.
- Eduardo, O. da C. (1990). O desenvolvimento da pesquisa de propaganda no Brasil. In R. Castelo Branco, R. L. Martensen, & F. Reis (Eds.), *História da propaganda no Brasil* (pp. 98–112). São Paulo: T.A. Queiroz, Editor.
- Ellis, J. (2014, November 3). The Texas Tribune is 5 years old and sustainable. Now what? Retrieved May 2, 2019, from Nieman Lab website: <https://www.niemanlab.org/2014/11/the-texas-tribune-is-5-years-old-and-sustainable-now-what/>
- Ember, S. (2016, November 2). Gawker and Hulk Hogan Reach \$31 Million Settlement. *The New York Times*. Retrieved from <https://www.nytimes.com/2016/11/03/business/media/gawker-hulk-hogan-settlement.html>

- Eric Raymond. (2004). The jargon file, version 4.4.8, Oct. 2004. Retrieved March 11, 2019, from <http://www.catb.org/jargon/>
- Fernandes, J. L., & Cascais, F. (2006). *A digitalização no sector da comunicação: Um desafio europeu*. Centro Protocolar de Formação Profissional para Jornalistas (CENJOR) and Sindicato dos Jornalistas.
- Fernandes, P. J. (2011a). A súbita vocação 'africanista' de um ex-ministro. A viagem de Mariano de Carvalho a Moçambique em 1890. *Africana Studia: Revista Internacional de Estudos Africanos*, 17(First semester), 17–40.
- Fernandes, P. J. (2011b). How to Rule a Small Town in Early Twentieth Century Portugal. *The International Journal of Regional and Local Studies*, 7(1–2), 115–139. <https://doi.org/10.1179/jrl.2011.7.1-2.115>
- Fidalgo, A. (2004). Sintaxe e Semântica das Notícias Online: Para um Jornalismo Assente em Base de Dados. In A. Lemos, Â. Pryston, J. M. da Silva, & S. F. de Sá (Eds.), *Mídia.br: Livro do XII Compós, 2003* (pp. 180–192). Porto Alegre: Editora Sulina.
- Fidalgo, A. (2007). Data Mining e um novo jornalismo de investigação. In Suzana Barbosa (Ed.), *Jornalismo Digital de Terceira Geração* (pp. 143–156). Covilhã: LabcomBooks.
- Figueiras, A. (2015). Towards the Understanding of Interaction in Information Visualization. *2015 19th International Conference on Information Visualisation*, 140–147. <https://doi.org/10.1109/iV.2015.34>
- Filho, J. C. (2013). A Imprensa Liberal [Memória da Imprensa]. Retrieved February 25, 2019, from Arquivo Público do Estado de São Paulo website: http://www.arquivoestado.sp.gov.br/memoria_imprensa/pdf/A%20Imprensa%20Liberal.pdf

- Fink, K., & Anderson, C. W. (2015). Data Journalism in the United States: Beyond the “usual suspects.” *Journalism Studies*, 16(4), 467–481.
<https://doi.org/10.1080/1461670X.2014.939852>
- Flew, T., Spurgeon, C., Daniel, A., & Swift, A. (2012). THE PROMISE OF COMPUTATIONAL JOURNALISM. *Journalism Practice*, 6(2), 157–171.
<https://doi.org/10.1080/17512786.2011.616655>
- Flood, A. (2016, July 1). Gay Talese says credibility of new book is “down the toilet.” *The Guardian*. Retrieved from
<https://www.theguardian.com/books/2016/jul/01/gay-talese-says-credibility-of-new-book-is-down-the-toilet-voyeurs-motel>
- Friedland, L. A., & Campbell, K. B. (2011). Connected Research: The Chicago School Precedent. In S. H. Iorio (Ed.), *Qualitative research in journalism: Taking it to the streets*. Retrieved from
<https://www.taylorfrancis.com/books/9781410609557>
- Friendly, M. (2008). A Brief History of Data Visualization. In *Handbook of Data Visualization* (pp. 15–56). https://doi.org/10.1007/978-3-540-33037-0_2
- Friendly, M., & Denis, D. (2000). Discussion and comments. Approche graphique en analyse des données. The roots and branches of modern statistical graphics. *Journal de La Société Française de Statistique*, 141(4), 51–60.
- Gade, P. J. (2004). Newspapers and Organizational Development: Management and Journalist Perceptions of Newsroom Cultural Change. *Journalism & Communication Monographs*, 6(1), 3–55.
<https://doi.org/10.1177/152263790400600101>

- Galtung, J., & Ruge, M. H. (1965). The Structure of Foreign News: The Presentation of the Congo, Cuba and Cyprus Crises in Four Norwegian Newspapers. *Journal of Peace Research*, 2(1), 64–90. <https://doi.org/10.1177/002234336500200104>
- Gans, H. J. (2004). *Deciding what's news: A study of CBS evening news, NBC nightly news, Newsweek, and Time* / Herbert J. Gans. Evanston, Ill: Northwestern University Press.
- Gardim, N. (2012). *O setor publicitário no Brasil: Construção e reconfiguração dos principais grupos e instituições de influência (1920-1980)*. Retrieved from <http://repositorio.unicamp.br/jspui/handle/REPOSIP/287032>
- Garrison, B. (1995). Online Services as Reporting Tools: Daily Newspaper Use of Commercial Databases in 1994. *Newspaper Research Journal*, 16(4), 74–86. <https://doi.org/10.1177/073953299501600406>
- Garrison, B. (2001). Computer-Assisted Reporting near Complete Adoption. *Newspaper Research Journal*, 22(1), 65–79. <https://doi.org/10.1177/073953290102200106>
- Garvey, E. G. (2013). “facts and FACTS”: Abolitionists’ Database Innovations. In L. Gitelman (Ed.), *Raw Data Is an Oxymoron* (pp. 89–102). MIT Press.
- Gill, C. J., & Gill, G. C. (2005). Nightingale in Scutari: Her Legacy Reexamined. *Clinical Infectious Diseases*, 40(12), 1799–1805. <https://doi.org/10.1086/430380>
- Gillie, A. (1996). The origin of the poverty line. *The Economic History Review*, 49(4), 715–730. <https://doi.org/10.1111/j.1468-0289.1996.tb00589.x>
- Gitlin, T. (2011). A surfeit of crises: Circulation, revenue, attention, authority, and deference. In R. W. McChesney & V. W. Pickard (Eds.), *Will the last reporter please turn out the lights: The collapse of journalism and what can be done to fix it* (pp. 92–102). New York: New Press : Distributed by Perseus Distribution.

- Gleick, J. (2011). *The information: A history, a theory, a flood* (1st Vintage Books ed., 2012). New York: Vintage Books.
- Gluck, M., & Roca Sales, M. (2008). *The future of television: Advertising, technology and the pursuit of audiences*. Retrieved from <http://openaccess.uoc.edu/webapps/o2/handle/10609/6801>
- Graves, L., Kelly, J., & Gluck, M. (2010). Confusion online: Faulty metrics and the future of digital journalism. *Columbia University*.
- Gray, J., Bounegru, L., & Chambers, L. (Eds.). (2012). *The Data Journalism Handbook* (First Edition). Sebastopol, CA: O'Reilly & Associates.
- Green, E., Cramer, P., & Anand, A. (2014). *What We Talk About When We Talk About Impact: One News Organization's Approach to Practicing Journalism with a Purpose*. Retrieved from Chalkbeat website: <https://chalkbeat.org/wp-content/uploads/2016/04/Chalkbeat-White-Paper-on-Impact-042914.pdf>
- Green-Barber, L. (2019). Beyond Clicks and Shares: How and why to measure the impact of data journalism projects. In J. Gray & L. Bounegru (Eds.), *The Data Journalism Handbook 2*. European Journalism Centre.
- Greene, H., & Yao, D. A. (2016). Navigating Natural Monopolies: Market Strategy and Nonmarket Challenges in Radio and Television Audience Measurement Markets. In J. M. De Figueiredo, M. Lenox, F. Oberholzer-Gee, & R. G. Vanden Bergh (Eds.), *Strategy Beyond Markets* (Vol. 34, pp. 367–411). <https://doi.org/10.1108/S0742-332220160000034011>
- Griffiths, A. (2015). *The New Journalism, the New Imperialism and the Fiction of Empire, 1870-1900*.

- Groves, J., & Brown, C. (2011). Stopping the Presses: A Longitudinal Case Study of the Christian Science Monitor Transition From Print Daily to Web Always. *International Symposium on Online Journalism, 1*(2), 86–128.
- Gynnild, A. (2014). Journalism innovation leads to innovation journalism: The impact of computational exploration on changing mindsets. *Journalism: Theory, Practice & Criticism, 15*(6), 713–730.
<https://doi.org/10.1177/1464884913486393>
- Habermas, J. (2012). *A transformação estrutural da esfera pública: Investigações sobre uma categoria da sociedade burguesa* (L. Nahodil, Trans.). Lisboa: Fundação Calouste Gulbenkian.
- Hamilton, J. (2016). *Democracy's detectives: The economics of investigative journalism*. Cambridge, Massachusetts: Harvard University Press.
- Hamilton, J. T., & Turner, F. (2009). *Developing the Field of Computational Journalism*. Retrieved from Center For Advanced Study in the Behavioral Sciences Summer Workshop, July 27-31, 2009 website:
<https://web.stanford.edu/~fturner/Hamilton%20Turner%20Acc%20by%20Alg%20Final.pdf>
- Hammer, P. J., Coleman, T. W., & Albom, M. (2014). *Crusader for justice: Federal judge Damon J. Keith*. Detroit, Michigan: Wayne State University Press.
- Hammerman, R., & Russell, A. L. (Eds.). (2015). *Ada's Legacy*.
<https://doi.org/10.1145/2809523>
- Hannaford, L. (2015). Computational Journalism in the UK Newsroom: Hybrids or Specialists? *Journalism Education, 4*(1), 6–21.
- Hanusch, F. (2017). Web analytics and the functional differentiation of journalism cultures: Individual, organizational and platform-specific influences on

- newswork. *Information, Communication & Society*, 20(10), 1571–1586.
<https://doi.org/10.1080/1369118X.2016.1241294>
- Harcup, T., & O’Neill, D. (2001). What Is News? Galtung and Ruge revisited. *Journalism Studies*, 2(2), 261–280. <https://doi.org/10.1080/14616700118449>
- Harcup, T., & O’Neill, D. (2017). What is News?: News values revisited (again). *Journalism Studies*, 18(12), 1470–1488.
<https://doi.org/10.1080/1461670X.2016.1150193>
- Hartsock, J. C. (2000). *A history of American literary journalism: The emergence of a modern narrative form*. Amherst: University of Massachusetts.
- Heitman, K. (2017). Reductionism at the Dawn of Population Health. In A. M. El-Sayed & S. Galea (Eds.), *Systems Science and Population Health* (pp. 9–24).
<https://doi.org/10.1093/acprof:oso/9780190492397.003.0002>
- Hennock, E. P. (1976). Poverty and social theory in England: The experience of the eighteen-eighties*. *Social History*, 1(1), 67–91.
<https://doi.org/10.1080/03071027608567369>
- Henriquez, H. (2014). *The Importance of Explanatory Infographics in Journalism* (Master Thesis, Universitat Oberta de Catalunya). Retrieved from
<http://ecollections.scad.edu/iii/cpro/DigitalItemViewPage.external;jsessionid=54CA095AAB35B5C177183D1628472266?lang=eng&sp=1002406&sp=T&sp=1&suite=def>
- Hermida, A., & Young, M. L. (2017). Finding the Data Unicorn: A hierarchy of hybridity in data and computational journalism. *Digital Journalism*, 5(2), 159–176. <https://doi.org/10.1080/21670811.2016.1162663>

- Hewett, J. (2017). Collaborative learning: From CAR to data journalism and Hacks/Hackers. In J. Mair, R. L. Keeble, & M. Lucero (Eds.), *Data journalism: Past, present and future*. Bury St Edmunds, Suffolk: Abramis.
- Himanen, P. (2001). *The hacker ethic, and the spirit of the information age* (1st ed). New York: Random House.
- Hindman, M. S. (2018). *The internet trap: How the digital economy builds monopolies and undermines democracy*. Princeton: Princeton University Press.
- Ho, B., & Liu, P. (2015). Herd journalism: Investment in novelty and popularity in markets for news. *Information Economics and Policy*, 31, 33–46.
<https://doi.org/10.1016/j.infoecopol.2015.04.004>
- Hollowell, J. (2017). *Fact and Fiction: The New Journalism and the Nonfiction Novel* (New edition edition). The University of North Carolina Press.
- Holovaty, A. (2006, September 6). A fundamental way newspaper sites need to change. Retrieved March 11, 2019, from Holovaty.com website:
<http://www.holovaty.com/writing/fundamental-change/>
- Holovaty, A. (2008, January 31). In memory of chicagocrime.org. Retrieved March 11, 2019, from Holovaty.com website:
<http://www.holovaty.com/writing/chicagocrime.org-tribute/>
- Houston, B. (1999). Changes in Attitudes, Changes in Latitudes. In *When Nerds and Words Collide: Reflections on the Development of Computer Assisted Reporting*. Poynter Institute for Media Studies.
- Houston, B. (2014). *Computer-Assisted Reporting: A Practical Guide*. Taylor and Francis.
- Howard, A. B. (2014). *The Art and Science of Data-Driven Journalism*. Retrieved from <https://academiccommons.columbia.edu/catalog/ac:zcrjdfn317>

- Hudson, P., & Ishizu, M. (2017). *History by numbers: An introduction to quantitative approaches* (Second edition). London Oxford New York New Delhi Sydney: Bloomsbury Academic.
- ICFJ. (2017). *2017 State of Technology in Global Newsrooms Survey*. Retrieved from International Center for Journalists website:
<https://www.icfj.org/sites/default/files/2018-04/ICFJTechSurveyFINAL.pdf>
- Ingram, M. (2015, May 19). The Financial Times bets on new ad format: Brands pay for time, not pageviews or clicks. Retrieved June 5, 2019, from Fortune website:
<http://fortune.com/2015/05/19/financial-times-advertising/>
- Jackson, S. (2016). *Cult of analytics data analytics for marketing*. Retrieved from <https://ezproxy.aub.edu.lb/login?url=https://www.taylorfrancis.com/books/9781317561897>
- Jensen, J. F. (1998). "Interactivity": Tracking a New Concept in Media and Communication Studies. *Nordicom Review*, 19(1), 185–204.
- Karlsen, J., & Stavelin, E. (2014). Computational Journalism in Norwegian Newsrooms. *Journalism Practice*, 8(1), 34–48.
<https://doi.org/10.1080/17512786.2013.813190>
- Katz, J. (2016). *Speaking American: How y'all, youse, and you guys talk: a visual guide*. Boston ; New York: Houghton Mifflin Harcourt.
- Kaushik, A. (2007). *Web analytics: An hour a day*. Indianapolis, Ind: Wiley.
- Kaushik, A. (2010). *Web analytics 2.0 the art of online accountability & science of customer centrality*. Retrieved from <http://nbn-resolving.de/urn:nbn:de:101:1-201412278534>
- Keller, M., & Abelson, B. (2015). *NewsLynx: A Tool for Newsroom Impact Measurement*. Retrieved from Tow Center for Digital Journalism website:

https://www.cjr.org/tow_center_reports/newslynx_a_tool_for_newsroom_impact_measurement.php/

- Kindel, E. (2017). Future, Fortune, and the graphic design of information. In A. Black, P. Luna, O. Lund, & S. Walker (Eds.), *Information Design: Research and Practice* (1st ed.). <https://doi.org/10.4324/9781315585680>
- Kinross, R. (2017). The lessons of Isotype for information design. In A. Black, P. Luna, O. Lund, & S. Walker (Eds.), *Information Design: Research and Practice* (1st ed.). <https://doi.org/10.4324/9781315585680>
- Klein, Scott. 2016a. *The Forgotten History of Visualization in News*. Tapestry: The Data Storytelling Conference.
https://www.youtube.com/watch?v=V_dgGbiSe8I.
- . 2016b. “Infographics in the Time of Cholera.” ProPublica. March 16, 2016.
<https://www.propublica.org/nerds/infographics-in-the-time-of-cholera>.
- Knight, M. (2015). Data journalism in the UK: A preliminary analysis of form and content. *Journal of Media Practice*, 16(1), 55–72.
<https://doi.org/10.1080/14682753.2015.1015801>
- Knight, P. (2016). *Reading the market: Genres of financial capitalism in Gilded Age America*. Baltimore: Johns Hopkins University Press.
- Koch, T. (2017). A thank-you, long overdue. *Cartographic Perspectives*, 0(87), 5–8.
<https://doi.org/10.14714/CP87.1407>
- Konieczna, M., & Powers, E. (2017). What can Nonprofit Journalists Actually do for Democracy? *Journalism Studies*, 18(12), 1542–1558.
<https://doi.org/10.1080/1461670X.2015.1134273>

- Kovach, B., & Rosenstiel, T. (2007). *The elements of journalism: What newspeople should know and the public should expect* (1st rev. ed., Completely updated and rev). New York: Three Rivers Press.
- Larrondo, A., Mielnickzuk, L., & Barbosa, S. (2008). Narrativa jornalística e base de dados: Discussão preliminar sobre gêneros textuais no ciberjornalismo de quarta geração. *Encontro Nacional de Pesquisadores Em Jornalismo, VI*. São Bernardo do Campo.
- Ledford, J., Teixeira, J., & Tyler, M. E. (2010). *Google analytics* (3. ed). Indianapolis, Ind: Wiley.
- Lee, A. M., Lewis, S. C., & Powers, M. (2014). Audience Clicks and News Placement: A Study of Time-Lagged Influence in Online Journalism. *Communication Research, 41*(4), 505–530. <https://doi.org/10.1177/0093650212467031>
- Lee-Wright, P. (2010). Culture Shock: New media and organizational change in the BBC. In N. Fenton (Ed.), *New media, old news: Journalism and democracy in the digital age* (pp. 73–86). London: SAGE.
- Leiner, B. M., Cerf, V. G., Clark, D. D., Kahn, R. E., Kleinrock, L., Lynch, D. C., ... Wolff, S. (2009). A brief history of the internet. *ACM SIGCOMM Computer Communication Review, 39*(5), 22. <https://doi.org/10.1145/1629607.1629613>
- Leite, Y. V. P., Moraes, W. F. A. de, & Machado, A. G. C. (2017). Ibope e os desafios competitivos. *Revista Pensamento Contemporâneo Em Administração, 11*(3), 101. <https://doi.org/10.12712/rpca.v11i3.670>
- Lewis, C., & Niles, H. (2013). *Measuring Impact: The art, science and mystery of nonprofit news assessment*. Retrieved from Investigative Reporting Workshop at American University website:
<https://irw.s3.amazonaws.com/uploads%2Fmeasuring-impact-final-pdf.pdf>

- Lewis, S. C., & Usher, N. (2013). Open source and journalism: Toward new frameworks for imagining news innovation. *Media, Culture & Society*, 35(5), 602–619. <https://doi.org/10.1177/0163443713485494>
- Lichterman, J. (2016, March 17). The FT is launching a new analytics tool to make metrics more understandable for its newsroom. Retrieved April 19, 2019, from Nieman Lab website: <https://www.niemanlab.org/2016/03/the-ft-is-launching-a-new-analytics-tool-to-make-metrics-more-understandable-for-its-newsroom/>
- Loosen, W. (2002). The Second-Level Digital Divide of the Web and Its Impact on Journalism. *First Monday*, 7(8). <https://doi.org/10.5210/fm.v7i8.977>
- Loosen, W., Reimer, J., & De Silva-Schmidt, F. (2017). Data-driven reporting: An ongoing (r)evolution? An analysis of projects nominated for the *Data Journalism Awards 2013–2016*. *Journalism: Theory, Practice & Criticism*, 146488491773569. <https://doi.org/10.1177/1464884917735691>
- Lorenz, M. (2010). Status and Outlook for data-driven journalism. *Data-Driven Journalism: What Is There to Learn?* Presented at the Data-driven journalism roundtable, Amsterdam. Retrieved from http://mediapusher.eu/datadrivenjournalism/pdf/ddj_paper_final.pdf
- MacGregor, P. (2007). TRACKING THE ONLINE AUDIENCE: Metric data start a subtle revolution. *Journalism Studies*, 8(2), 280–298. <https://doi.org/10.1080/14616700601148879>
- Machado, E. (2005). A Base de Dados como Formato no Jornalismo Digital. *Actas Do VII Lusocom*, 301–310. Covilhã.
- Machado, E. (2007). A Base de Dados como espaço de composição multimídia. In Suzana Barbosa (Ed.), *Jornalismo Digital de Terceira Geração* (pp. 111–126). Covilhã: LabcomBooks.

- Madureira, F. B. (2010). *Cidadão-fonte ou cidadão-repórter? O engajamento do público no jornalismo colaborativo dos grandes portais brasileiros*. (Master Thesis, Universidade de São Paulo). <https://doi.org/10.11606/D.27.2010.tde-08112010-115607>
- Magnolo, T. S., & Pereira, A. A. (2016). O papel desempenhado pelo jornal O Globo ao golpe de 64. *Intercom – Sociedade Brasileira de Estudos Interdisciplinares Da Comunicação, 21*. Retrieved from <http://www.portalintercom.org.br/anais/sudeste2016/resumos/R53-0160-1.pdf>
- Mair, J., Keeble, R. L., Lucero, M., & Moore, M. (Eds.). (2017). *Data journalism: Past, present and future*. Bury St Edmunds, Suffolk: Abramis.
- Manovich, L. (2002). *The language of new media* (1st MIT Press pbk. ed). Cambridge, Mass: MIT Press.
- Martin, F., & Dwyer, T. (2019). The Numbers Game: Social News Analytics. In F. Martin & T. Dwyer, *Sharing News Online* (pp. 61–90). https://doi.org/10.1007/978-3-030-17906-9_3
- Martinho, A. I. P. (2013). *Jornalismo de dados: Contributo para uma caracterização do estado da arte em Portugal*. Retrieved from <https://repositorio.iscte-iul.pt/handle/10071/8329>
- Matos, S. C. (1993). Leite de Vasconcelos no debate acerca da formação de Portugal: Um confronto com Oliveira Martins. *O Arqueólogo Português, 11–12*, 11–33.
- Meikle. (2012). Continuity and Transformation in Convergent News: The Case of Wikileaks. *Media International Australia, 144*(1), 52–59. <https://doi.org/10.1177/1329878X1214400109>
- Merritt, B. (2017). *The Rise of Attention Metrics: Can a New Digital Currency Help Sustain Journalism?* [White Paper]. Retrieved from The George Washington

University website:

https://smpa.gwu.edu/sites/g/files/zaxdzs2046/f/downloads/Merritt_whitepaper_for_publication.pdf

Merritt, D., & McCombs, M. (2014). *The Two W's of Journalism: The Why and What of Public Affairs Reporting* (1st ed.). <https://doi.org/10.4324/9781410609298>

Method, J. (2008). *The Benefits of Computer-Assisted Reporting* (Nieman Reports No. Vol. 62 No. 3 Fall 2008; pp. 12–14). Retrieved from The Nieman Foundation for Journalism at Harvard University website: <https://niemanreports.org/articles/the-benefits-of-computer-assisted-reporting/>

Meyer, E. K. (1997). *Designing infographics*. Indianapolis, Ind.: Hayden Books.

Meyer, P. (1999). The Future of CAR: Declare Victory and Get out! In *When Nerds and Words Collide: Reflections on the Development of Computer Assisted Reporting* (pp. 4–5). Poynter Institute for Media Studies.

Meyer, P. (2002). *Precision journalism: A reporter's introduction to social science methods* (4th ed). Lanham, Md: Rowman & Littlefield Publishers.

Meyer, P. (2012, October 18). Precision Journalism and Narrative Journalism: Toward a Unified Field Theory. Retrieved March 8, 2019, from Nieman Reports website: <https://niemanreports.org/articles/precision-journalism-and-narrative-journalism-toward-a-unified-field-theory/>

Mokalis, A. L., & Davis, J. J. (2018). *Google Analytics demystified* (4th Edition). CreateSpace Independent Publishing Platform.

Monmonier, M. S. (1999). *Maps with the news: The development of American journalistic cartography* (Paperback ed). Chicago: University of Chicago Press.

Moran, C. (2019, April 2). You may hate metrics. But they're making journalism better. Retrieved April 22, 2019, from Columbia Journalism Review website:

<https://www.cjr.org/innovations/you-may-hate-metrics-guardian-audience-twitter-images.php>

Moses, L. (2015, June 22). "Narcissism works really well": Why Time magazine created a site for its interactive stories. Retrieved March 11, 2019, from Digiday website: <https://digiday.com/media/narcissism-works-really-well-time-magazine-created-site-interactive-stories/>

Moura, I. C. (2018). *O estado do jornalismo de dados no cenário luso-brasileiro* (Master Thesis, Universidade Nova de Lisboa). Retrieved from <https://run.unl.pt/handle/10362/47623>

Moyo, D., Mare, A., & Matsilele, T. (2019). Analytics-Driven Journalism? Editorial Metrics and the Reconfiguration of Online News Production Practices in African Newsrooms". *Digital Journalism*, 1–17.
<https://doi.org/10.1080/21670811.2018.1533788>

Napoli, P. M. (2011). *Audience evolution: New technologies and the transformation of media audiences*. New York: Columbia University Press.

Napoli, P. M., & Bjur, J. (2017). Audience Ratings, Media Industries. In J. Matthes, C. S. Davis, & R. F. Potter (Eds.), *The International Encyclopedia of Communication Research Methods* (pp. 1–19).
<https://doi.org/10.1002/9781118901731.iecrm0010>

Newman, N. (2019). *Journalism, Media and Technology Trends and Predictions 2019*. Retrieved from Reuters Institute for the Study of Journalism website: <https://www.reuterscommunity.com/wp-content/uploads/2019/03/journalism-media-and-technology-trends-and-predictions-2019.pdf>

Newman, N., Fletcher, R., Kalogeropoulos, A., Levy, D. A. L., & Nielsen, R. K. (2018). *Digital News Report 2018*. Retrieved from Reuters Institute for the Study of

Journalism website:

<https://agency.reuters.com/content/dam/openweb/documents/pdf/news-agency/report/dnr-18.pdf>

- Newman, N., Fletcher, R., Levy, D. A. L., & Nielsen, R. K. (2016). *Digital News Report 2016*. Retrieved from Reuters Institute for the Study of Journalism website: <http://www.digitalnewsreport.org/survey/2016/hard-soft-news-2016/>
- Nguyen, A., & Lugo-Ocando, J. (2016). The state of data and statistics in journalism and journalism education: Issues and debates. *Journalism: Theory, Practice & Criticism*, 17(1), 3–17. <https://doi.org/10.1177/1464884915593234>
- Nguyen, A., & Vu, H. T. (2018). Numbers behind the news: Audience metrics and the changing nature of gatekeeping. In A. Nguyen (Ed.), *News, numbers and public opinion in a data-driven world*. New York: Bloomsbury Academic.
- Nobre, A. (2015, March 5). “Público”: 25 anos em papel, num jornal cada vez mais digital. Retrieved June 4, 2019, from Expresso website: <http://leitor.expresso.pt/diario/05-03-2015/html/caderno-1/temas-principais/publico-25-anos-em-papel-num-jornal-cada-vez-mais-digital>
- North, L. (2016). The Gender of “soft” and “hard” news: Female journalists’ views on gendered story allocations. *Journalism Studies*, 17(3), 356–373. <https://doi.org/10.1080/1461670X.2014.987551>
- O Globo. (2013, August 31). Apoio editorial ao golpe de 64 foi um erro. Retrieved February 27, 2019, from Jornal O Globo website: <https://oglobo.globo.com/brasil/apoio-editorial-ao-golpe-de-64-foi-um-erro-9771604>
- O Globo. (2018, July 29). Site do GLOBO ganha nova estrutura e layout. Retrieved February 27, 2019, from Jornal O Globo website:

<https://oglobo.globo.com/brasil/site-do-globo-ganha-nova-estrutura-layout-22927697>

- O'Brien, C. (2011). The origins and originators of early statistical societies: A comparison of Liverpool and Manchester: Origins of Early Statistical Societies. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 174(1), 51–62. <https://doi.org/10.1111/j.1467-985X.2010.00649.x>
- O'Connor, S. (2009). Methodological Triangulation and the Social Studies of Charles Booth, Jane Addams, and W.E.B. Du Bois. *Sociation Today*, 7(1). Retrieved from <http://www.ncsociology.org/sociationtoday/dubois/three.htm>
- Ojo, A., & Heravi, B. (2018). Patterns in Award Winning Data Storytelling: Story Types, Enabling Tools and Competences. *Digital Journalism*, 6(6), 693–718. <https://doi.org/10.1080/21670811.2017.1403291>
- Oliveira, A. P. B. de. (2018). *Uma cartografia interativa do jornalismo de dados no Brasil: Percepções sobre competências e habilidades no mercado de trabalho e na academia* (Master Thesis, Universidade Municipal de São Caetano do Sul). Retrieved from <http://repositorio.uscs.edu.br/handle/123456789/1202>
- Oliveira, C. E. F. de. (2009). *Poder local e palavra impressa: A dinâmica política em torno dos Conselhos Provinciais e da imprensa periódica em São Paulo, 1824-1834* (Master Thesis, Universidade de São Paulo). <https://doi.org/10.11606/D.8.2009.tde-08022010-123937>
- O'Regan, G. (2016a). The First Digital Computers. In G. O'Regan, *Introduction to the History of Computing* (pp. 55–72). https://doi.org/10.1007/978-3-319-33138-6_4
- O'Regan, G. (2016b). The Internet Revolution. In G. O'Regan, *Introduction to the History of Computing* (pp. 163–178). https://doi.org/10.1007/978-3-319-33138-6_14

- Orford, S., Dorling, D., Mitchell, R., Shaw, M., & Smith, G. D. (2002). Life and death of the people of London: A historical GIS of Charles Booth's inquiry. *Health & Place, 8*(1), 25–35. [https://doi.org/10.1016/S1353-8292\(01\)00033-8](https://doi.org/10.1016/S1353-8292(01)00033-8)
- Örnebring, H. (2010). Technology and journalism-as-labour: Historical perspectives. *Journalism: Theory, Practice & Criticism, 11*(1), 57–74. <https://doi.org/10.1177/1464884909350644>
- Palacios, M. (2003). Ruptura, Continuidade e Potencialização no Jornalismo Online: O Lugar da Memória. In E. Machado & M. Palacios (Eds.), *Modelos de Jornalismo Digital*. Salvador: Editora Calandra.
- Parasie, S. (2011, October 14). 'Hacker' Journalism—A New Utopia for the Press? *Books & Ideias*. Retrieved from <https://booksandideas.net/Hacker-Journalism-A-New-Utopia-for.html>
- Parasie, S., & Dagiral, E. (2013). Data-driven journalism and the public good: “Computer-assisted-reporters” and “programmer-journalists” in Chicago. *New Media & Society, 15*(6), 853–871. <https://doi.org/10.1177/1461444812463345>
- Pelling, M. (2016). John Graunt, the Hartlib circle and child mortality in mid-seventeenth-century London. *Continuity and Change, 31*(03), 335–359. <https://doi.org/10.1017/S0268416016000278>
- Petre, C. (2015). The Traffic Factories: Metrics at Chartbeat, Gawker Media, and The New York Times. *Columbia University*. <https://doi.org/10.7916/d80293w1>
- Petty, W., Hull, C. H., & Graunt, J. (1986). *The economic writings of Sir William Petty: Together with the Observations upon the bills of mortality more probably by John Graunt* (Reprinted in 1 vol. Orig. publ. Cambridge, 1899). Fairfield, NJ: Kelley.

- Pitt, F., & Green-Barber, L. (2017). *The Case for Media Impact: A Case Study of ICIJ's Radical Collaboration Strategy*. <https://doi.org/10.7916/d85q532v>
- Poovey, M. (2004). *A history of the modern fact: Problems of knowledge in the sciences of wealth and society* (3th pr). Chicago: Univ. of Chicago Press.
- Porlezza, C. (2019). Data journalism and the ethics of open source. In A. Daly, S. K. Devitt, & M. Mann (Eds.), *Good Data. Theory on Demand* (pp. 189–201). Amsterdam: Institute of Network Cultures.
- Porter, T. M. (2011). *The rise of statistical thinking: 1820 - 1900* (Nachdr.). Princeton, NJ: Princeton Univ. Press.
- Powers, E. (2018). Selecting Metrics, Reflecting Norms: How journalists in local newsrooms define, measure, and discuss impact. *Digital Journalism*, 6(4), 454–471. <https://doi.org/10.1080/21670811.2018.1445002>
- Público. (1990, March 5). Estatuto Editorial. Retrieved June 4, 2019, from PÚBLICO website: <https://www.publico.pt/nos/estatuto-editorial>
- Público. (2014, March 25). PÚBLICO vence prémio internacional de design digital. Retrieved June 4, 2019, from Público website: <https://www.publico.pt/2014/03/25/sociedade/noticia/publico-vence-premio-internacional-de-design-digital-1629638>
- Público. (2019, April 30). Vendas em banca do PÚBLICO crescem num mercado em queda. Retrieved June 4, 2019, from Público website: <https://www.publico.pt/2019/04/30/sociedade/noticia/vendas-banca-publico-crescem-mercado-queda-1870995>
- Putney, C. (2003). *Muscular Christianity: Manhood and sports in Protestant America, 1880 - 1920* (2. print). Cambridge, Mass.: Harvard Univ. Press.

- Radcliffe, D. (2017). A importância dos dados para o jornalismo local. *Comunicação & Educação*, 22(1), 85. <https://doi.org/10.11606/issn.2316-9125.v22i1p85-97>
- Reader, B. (2015). *Audience Feedback in the News Media* (1st ed.). <https://doi.org/10.4324/9781315773858>
- Regard, F. (2014). The Sexual Exploitation of the Poor in WT Stead's "New Journalism": Humanity, Democracy and the Tabloid Press. In B. Korte & F. Regard (Eds.), *Narrating poverty and precarity in Britain*. Berlin: De Gruyter.
- Reilly, S. (2017). The Need to Help Journalists with Data and Information Visualization. *IEEE Computer Graphics and Applications*, 37(2), 8–10. <https://doi.org/10.1109/MCG.2017.32>
- Residents of Hull-House. (1895). *Hull-House maps and papers: A presentation of nationalities and wages in a congested district of Chicago, together with comments and essays on problems growing out of the social conditions*. New York: T.Y. Crowell.
- Ribeiro, I. (2018, July 24). O Globo apresenta projeto voltado à redação integrada – Meio & Mensagem. Retrieved May 29, 2019, from Meio & Mensagem website: <https://www.meioemensagem.com.br/home/midia/2018/07/24/o-globo-apresenta-projeto-voltado-a-redacao-integrada.html>
- Ribeiro, S. A. (2008). *Infografia de imprensa: História e análise ibérica comparada*. Coimbra: MinervaCoimbra.
- Rodríguez, L. T., & Castilla, E. B. (2016). El periodismo de datos en los medios de referencia españoles. *Trípodos*, 0(38), 13-30–30.
- Rogers, S. (2013). *Facts are Sacred* (Main edition). London: Guardian Faber Publishing.

- Rogers, S. (2016, July 1). What is Google Trends data—And what does it mean?
Retrieved June 12, 2019, from Medium website: <https://medium.com/google-news-lab/what-is-google-trends-data-and-what-does-it-mean-b48f07342ee8>
- Rojas, R. (Ed.). (2002). *The first computers: History and architectures*. Cambridge, Mass: MIT Press.
- Romeu, A. M. D. (2015). *A medição das audiências televisivas em Portugal: Novas práticas, novos consumos, novos desafios* (Master Thesis, Universidade Católica Portuguesa). Retrieved from <https://repositorio.ucp.pt/handle/10400.14/17241>
- Rosa, B. (2019, January 25). O GLOBO é o jornal que mais cresceu em 2018. Retrieved February 27, 2019, from Jornal O Globo website:
<https://oglobo.globo.com/economia/o-globo-o-jornal-que-mais-cresceu-em-2018-23400125>
- Rosegrant, S. (2011, April 24). Revealing the roots of a riot. Retrieved March 5, 2019, from Institute for Social Research website: <https://isr.umich.edu/news-events/insights-newsletter/article/revealing-the-roots-of-a-riot/>
- Rosenberg, D. (2007). Joseph Priestley and the Graphic Invention of Modern Time. *Studies in Eighteenth Century Culture*, 36(1), 55–103.
<https://doi.org/10.1353/sec.2007.0013>
- Rossmann, G. (2015). *Climbing the charts: What radio airplay tells us about the diffusion of innovation*.
- Roush, C. (2012). *Profits and losses: Business journalism & its role in society*. Retrieved from <http://public.eblib.com/choice/publicfullrecord.aspx?p=784666>
- Roy, J.-H. (2016). *Le journalisme informatique au Québec: Expansion du journalisme ou nouveau territoire professionnel?* (Master Thesis, Université Laval). Retrieved from <https://corpus.ulaval.ca/jspui/handle/20.500.11794/27205>

- Royal, C. (2010). A Case Study of The New York Times Interactive News Technology Department. *International Symposium in Online Journalism*. Presented at the International Symposium in Online Journalism.
- Russell, A. (2016). *Journalism as activism: Recoding media power*. Retrieved from <http://proxy.uqtr.ca/login.cgi?action=login&u=uqtr&db=ebsco&ezurl=http://search.ebscohost.com/login.aspx?direct=true&scope=site&db=nlebk&db=nlabk&AN=1358580>
- Santos, R. (2006). *A Rádio em Portugal—Estado da Arte em 2006* (pp. 220–223). OberCom.
- Schardt, S., & Carpenter, S. (2009). The News That's Fit to Click: An Analysis of Online News Values and Preferences Present in the Most-viewed Stories on azcentral.com. *Southwestern Mass Communication Journal*, 24(2), 17–26.
- Schiffrin, A., & Zuckerman, E. (2015). Can We Measure Media Impact? Surveying the Field. *Stanford Social Innovation Review*, 13(4).
- Schmidt, C. (2018, March 26). Recirculate! Vox Media's new structure for story packages gives readers context (and helps them stick around). Retrieved May 2, 2019, from Nieman Lab website: <https://www.niemanlab.org/2018/03/recirculate-vox-medias-new-structure-for-story-packages-gives-readers-context-and-helps-them-stick-around/>
- Schudson, M. (1981). *Discovering The News: A Social History of American Newspapers*. Retrieved from <http://public.eblib.com/choice/publicfullrecord.aspx?p=904009>
- Schudson, M. (1996). *The power of news*. Cambridge, Mass: Harvard University Press.

- Schudson, M. (2001). The objectivity norm in American journalism*. *Journalism: Theory, Practice & Criticism*, 2(2), 149–170.
<https://doi.org/10.1177/146488490100200201>
- Schudson, M. (2008). *Why democracies need an unlovable press*. Cambridge, UK ; Malden, MA: Polity.
- Schudson, M. (2012). *The sociology of news* (2nd ed). New York: W.W. Norton & Company.
- Sederstrom, E. (2018). Data Visualization. In M. L. Sylvia & M. F. Terhaar (Eds.), *Clinical analytics and data management for the DNP* (2nd ed., p. 396). Springer Publishing Company.
- Segel, E., & Heer, J. (2010). Narrative Visualization: Telling Stories with Data. *IEEE Transactions on Visualization and Computer Graphics*, 16(6), 1139–1148.
<https://doi.org/10.1109/TVCG.2010.179>
- Silva, M. T. da. (2014). *As cartas dos leitores na imprensa portuguesa: Uma forma de comunicação e debate do público*. Retrieved from http://www.labcom-ifp.ubi.pt/ficheiros/20150224-201407_mtsilva_cartasleitores.pdf
- Simons, M., Tiffen, R., Hendrie, D., Carson, A., Sullivan, H., Muller, D., & McNair, B. (2017). Understanding the civic impact of journalism: A realistic evaluation perspective. *Journalism Studies*, 18(11), 1400–1414.
<https://doi.org/10.1080/1461670X.2015.1129284>
- Singer, J. B. (2011). COMMUNITY SERVICE: Editor pride and user preference on local newspaper websites. *Journalism Practice*, 5(6), 623–642.
<https://doi.org/10.1080/17512786.2011.601938>
- Smiley, J. (2010). *The man who invented the computer: The biography of John Atanasoff, digital pioneer* (1. ed). New York: Doubleday.

- Smit, G., de Haan, Y., & Buijs, L. (2014). Visualizing News: Make it work. *Digital Journalism*, 2(3), 344–354. <https://doi.org/10.1080/21670811.2014.897847>
- Smythe, T. C. (2003). *The Gilded Age press, 1865-1900*. Westport, Conn.: Praeger.
- Soares, I. (2017). At the intersection of risk: When literary journalism and sociology study urban problems by means of akin methodologies. *Sociologia, Problemas e Práticas*, 2017(84). <https://doi.org/10.7458/SPP2017843466>
- Sodré, N. W. (1999). *História da imprensa no Brasil*. Mauad Editora Ltda.
- Sousa, J. P. (2009). Diário de Notícias: A revolução de Eduardo Coelho no jornalismo português oitocentista. *Revista PJ:Br – Jornalismo Brasileiro*, VI(12).
- Souza, D. R. de, Leuck, L. P., Santos, C. Q., Silveira, M. S., Manssour, I. H., & Tietzmann, R. (2018). Interacting with Data to Create Journalistic Stories: A Systematic Review. In S. Yamamoto & H. Mori (Eds.), *Human Interface and the Management of Information. Interaction, Visualization, and Analytics* (Vol. 10904, pp. 685–704). https://doi.org/10.1007/978-3-319-92043-6_54
- Spence, R. (2014). *Information visualization: An introduction* (Third edition). Cham Heidelberg New York Dordrecht London: Springer.
- Spencer, D. R. (2007). *The yellow journalism: The press and America's emergence as a world power*. Evanston, Ill: Northwestern University Press.
- Stalph, F. (2018). Classifying Data Journalism: A content analysis of daily data-driven stories. *Journalism Practice*, 12(10), 1332–1350. <https://doi.org/10.1080/17512786.2017.1386583>
- Stencel, M., Aidar, B., & Kamalakanthan, P. (2014). *The Goat Must Be Fed*. Retrieved from the DeWitt Wallace Center for Media & Democracy website: <http://www.goatmustbefed.com/resources/pdf/goat-must-be-fed.pdf>
- Stephens, M. (2007). *A history of news* (3. ed). New York: Oxford Univ. Press.

- Stockleben, B., & Lugmayr, A. (2016). The Impact of Fluid Publishing on Media Information Management—A Survey of Latest Journalistic Trends as Data-Driven Journalism, Journalism as Process and Metrics-Driven Journalism. In A. Lugmayr, E. Stojmenova, K. Stanoevska, & R. Wellington (Eds.), *Information Systems and Management in Media and Entertainment Industries* (pp. 299–318). https://doi.org/10.1007/978-3-319-49407-4_15
- Swade, D. (2002). *The difference engine: Charles Babbage and the quest to build the first computer*. New York: Penguin Books.
- Tabary, C., Provost, A.-M., & Trottier, A. (2016). Data journalism's actors, practices and skills: A case study from Quebec. *Journalism: Theory, Practice & Criticism*, 17(1), 66–84. <https://doi.org/10.1177/1464884915593245>
- Tandoc, E. C. (2014). Journalism is twerking? How web analytics is changing the process of gatekeeping. *New Media & Society*, 16(4), 559–575. <https://doi.org/10.1177/1461444814530541>
- Tandoc, E. C., & Oh, S.-K. (2017). Small Departures, Big Continuities?: Norms, values, and routines in *The Guardian*'s big data journalism. *Journalism Studies*, 18(8), 997–1015. <https://doi.org/10.1080/1461670X.2015.1104260>
- Tandoc, E. C., & Thomas, R. J. (2015). The Ethics of Web Analytics: Implications of using audience metrics in news construction. *Digital Journalism*, 3(2), 243–258. <https://doi.org/10.1080/21670811.2014.909122>
- Teixeira, T. (2010). *Infografia e jornalismo: Conceitos, análises e perspectivas*. Retrieved from <http://repositorio.ufba.br/ri/handle/ri/20642>
- Tengarrinha, J. (2013). *Nova história da imprensa portuguesa: Das origens a 1865* (1. ed). Lisboa: Temas e Debates [u.a.].

- Toledo, J. R. de. (2014). “Pirando” nos dados. In H. Dantas, J. R. de Toledo, & M. A. C. Teixeira (Eds.), *Análise política & jornalismo de dados: Ensaio a partir do Basômetro* (1ª edição). Rio de Janeiro, RJ, Brasil: FGV Editora.
- Tominski, C., Aigner, W., Miksch, S., & Schumann, H. (2017). Images of time. In A. Black, P. Luna, O. Lund, & S. Walker (Eds.), *Information Design: Research and Practice* (1st ed., p. 766). <https://doi.org/10.4324/9781315585680>
- Topalov, C. (1993). The city as *terra incognita*: Charles Booth’s poverty survey and the people of London, 1886–1891*. *Planning Perspectives*, 8(4), 395–425.
<https://doi.org/10.1080/02665439308725782>
- Torres, V. (2017). *Mensuração Editorial: O uso de métricas e dados sobre o público na produção de informação jornalística* (Ph.D. Thesis, Universidade Federal da Bahia). Retrieved from http://poscom.tempsite.ws/wp-content/uploads/2011/05/Mensura%C3%A7%C3%A3o-Editorial-Vitor-Torres_Vers%C3%A3oFinal.pdf
- Traquina, N. (2004). *Teorias do jornalismo—Volume I*. Florianópolis, SC: Insular.
- Traquina, N. (2005). *Teorias do jornalismo—Volume II*. Florianópolis, SC: Insular.
- Träsel, M. R. (2014a). *Entrevistando planilhas: Estudo das crenças e do ethos de um grupo de profissionais de jornalismo guiado por dados no Brasil* (Ph.D. Thesis, Pontifícia Universidade Católica do Rio Grande do Sul). Retrieved from <http://repositorio.pucrs.br:80/dspace/handle/10923/6841>
- Träsel, M. R. (2014b). O jornalismo guiado por dados numa perspectiva brasileira. In J. Gray, L. Bounegru, L. Chambers, European Journalism Centre, & Open Knowledge Foundation (Eds.), *Manual de Jornalismo de Dados* (Edição em português coordenada pela Abraji). O’Reilly Media.

- Trudeau, J. T., & Morrissey, M. E. (2017). “Bring in an Honest Verdict”: Prosecuting Southern Whiteness in *American Slavery As It Is*. *Southern Communication Journal*, 82(5), 312–323. <https://doi.org/10.1080/1041794X.2017.1332093>
- Tufte, E. R. (2001). *The visual display of quantitative information* (2. ed). Cheshire, Conn: Graphics Press.
- Usher, N. (2012). Going Web-First at The Christian Science Monitor: A Three-Part Study of Change. *International Journal of Communication*, 6(0), 20.
- Usher, N. (2013). AL JAZEERA ENGLISH ONLINE: Understanding Web metrics and news production when a quantified audience is not a commodified audience. *Digital Journalism*, 1(3), 335–351. <https://doi.org/10.1080/21670811.2013.801690>
- Usher, N. (2016). *Interactive journalism: Hackers, data, and code*. Urbana: University of Illinois Press.
- Valeeva, A. (2017). Investigative open data journalism in Russia. In J. Mair, R. L. Keeble, & M. Lucero (Eds.), *Data journalism: Past, present and future*. Bury St Edmunds, Suffolk: Abramis.
- Vaughan, L. (2018). *Mapping society: The spatial dimensions of social cartography*. Retrieved from <http://www.jstor.org/stable/10.2307/j.ctv550dcj>
- Vázquez-Herrero, J., Negreira-Rey, M.-C., & López-García, X. (2019). La innovación multimedia e interactiva en el ciberperiodismo argentino. *Revista de Comunicación*, 18(1), 191–214. <https://doi.org/10.26441/RC18.1-2019-A10>
- Vieira, L. de S. (2018). *Métricas editoriais no jornalismo online: Ética e cultura profissional na relação com audiências ativas* (Ph.D. Thesis, Universidade Federal de Santa Catarina). Retrieved from <https://repositorio.ufsc.br/handle/123456789/189167>

- Vines, E. (2002). Is this journalism?: A study of whether the snapshots on the front page of USA Today adhere to journalistic standards. *LSU Master's Theses*. Retrieved from https://digitalcommons.lsu.edu/gradschool_theses/1244
- Vu, H. T. (2014). The online audience as gatekeeper: The influence of reader metrics on news editorial selection. *Journalism: Theory, Practice & Criticism*, 15(8), 1094–1110. <https://doi.org/10.1177/1464884913504259>
- Waisbord, S. (2016). Watchdog Journalism. In G. Mazzoleni (Ed.), *The International Encyclopedia of Political Communication* (pp. 1–5). <https://doi.org/10.1002/9781118541555.wbiepc046>
- Wang, Q. (2018). Dimensional Field Theory: The adoption of audience metrics in the journalistic field and cross-field influences. *Digital Journalism*, 6(4), 472–491. <https://doi.org/10.1080/21670811.2017.1397526>
- Weber, J. (2006). Strassburg, 1605: The Origins of the Newspaper in Europe. *German History*, 24(3), 387–412. <https://doi.org/10.1191/0266355406gh380oa>
- Weber, W., & Rall, H. (2013). "We are journalists." Production Practices, Attitudes and a Case Study of the New York Times Newsroom. In W. Weber, M. Burmester, & R. Tille (Eds.), *Interaktive Infografiken* (pp. 161–172). https://doi.org/10.1007/978-3-642-15453-9_9
- Webster, J. G., Phalen, P. F., & Lichty, L. W. (2014). *Ratings analysis: Audience measurement and analytics* (Fourth edition). New York, NY: Routledge.
- Weinacht, S., & Spiller, R. (2014). Datenjournalismus in Deutschland: Eine explorative Untersuchung zu Rollenbildern von Datenjournalisten. *Publizistik*, 59(4), 411–433. <https://doi.org/10.1007/s11616-014-0213-5>
- Welbers, K., van Atteveldt, W., Kleinnijenhuis, J., Ruigrok, N., & Schaper, J. (2016). News selection criteria in the digital age: Professional norms versus online

- audience metrics. *Journalism: Theory, Practice & Criticism*, 17(8), 1037–1053.
<https://doi.org/10.1177/1464884915595474>
- Weller, B., & Calcott, L. (2012). Google Analytics and Actionable Data. In B. Weller & L. Calcott, *The Definitive Guide to Google AdWords* (pp. 291–333).
https://doi.org/10.1007/978-1-4302-4015-0_11
- Westwood, J. L. (2010). *The social construction of risk in child trafficking discourses: A study of melodramatic tactics in child trafficking narratives* (Ph.D. Thesis, University of Central Lancashire). Retrieved from
<http://www.uclan.ac.uk/research/index.php>
- Wiener, J. H. (2011). *The Americanization of the British press, 1830s-1914: Speed in the age of transatlantic journalism*. Retrieved from
<http://www.dawsonera.com/depp/reader/protected/external/AbstractView/S9780230347953>
- Wikileaks, & Assange, J. (2016). *The WikiLeaks files: The world according to US Empire*. London: Verso.
- Wilbur, K. C. (2015). Recent Developments in Mass Media. In *Handbook of Media Economics* (Vol. 1, pp. 205–224). <https://doi.org/10.1016/B978-0-444-62721-6.00005-6>
- Wilson-Chapman, A., Cucho, A., & Fitzgibbon, W. (2019, April 3). What happened after the Panama Papers? Retrieved April 23, 2019, from ICIJ website:
<https://www.icij.org/investigations/panama-papers/what-happened-after-the-panama-papers/>
- Wright, S., & Doyle, K. (2018). The Evolution of Data Journalism: A Case Study of Australia. *Journalism Studies*, 1–17.
<https://doi.org/10.1080/1461670X.2018.1539343>

- Wu, T. (2017). *The attention merchants: The epic scramble to get inside our heads* (First Vintage Books Edition). New York: Vintage Books, a division of Penguin Random House LLC.
- Yi, J. S., Kang, Y. ah, & Stasko, J. (2007). Toward a Deeper Understanding of the Role of Interaction in Information Visualization. *IEEE Transactions on Visualization and Computer Graphics*, 13(6), 1224–1231.
<https://doi.org/10.1109/TVCG.2007.70515>
- Young, M. L., & Hermida, A. (2015). From Mr. and Mrs. Outlier To Central Tendencies: Computational journalism and crime reporting at the *Los Angeles Times*. *Digital Journalism*, 3(3), 381–397.
<https://doi.org/10.1080/21670811.2014.976409>
- Young, M. L., Hermida, A., & Fulda, J. (2018). What Makes for Great Data Journalism?: A content analysis of data journalism awards finalists 2012–2015. *Journalism Practice*, 12(1), 115–135.
<https://doi.org/10.1080/17512786.2016.1270171>
- Zamith, R. (2015). *Editorial Judgment in an Age of Data: How Audience Analytics and Metrics are Influencing the Placement of News Products* (Ph.D. Thesis, University of Minnesota). Retrieved from
<http://conservancy.umn.edu/handle/11299/175385>
- Zamith, R. (2018). On Metrics-Driven Homepages: Assessing the relationship between popularity and prominence. *Journalism Studies*, 19(8), 1116–1137.
<https://doi.org/10.1080/1461670X.2016.1262215>
- Zhang, S., & Feng, J. (2019). A Step Forward?: Exploring the diffusion of data journalism as journalistic innovations in China. *Journalism Studies*, 20(9), 1281–1300. <https://doi.org/10.1080/1461670X.2018.1513814>

Appendix A: *The Texas Tribune* data journalism pieces

A.1. Interactive pieces from 2014

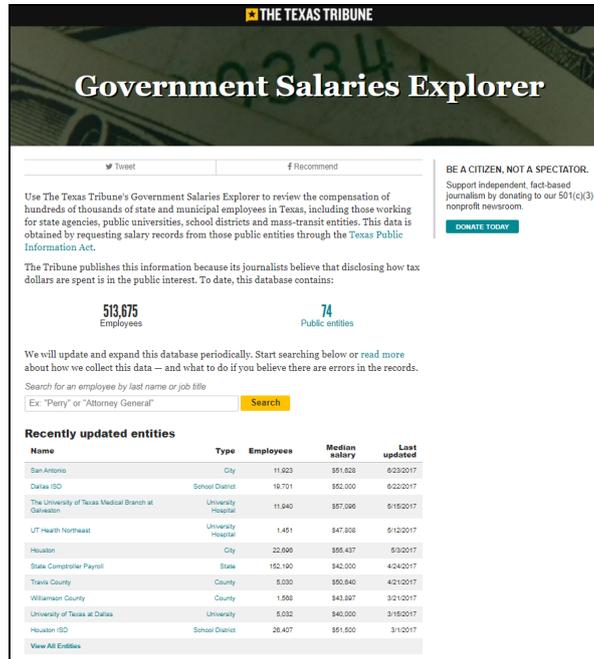


Figure 70 – “Government Salaries Explorer” screenshot.

Title: “Government Salaries Explorer”.

Original url: <https://salaries.texastribune.org/>

Updated in 10/21/2014.

Archived url: <http://bit.ly/2Sj51V2>

Interaction technique: Explore.

Topics: Politics/governance, economy/business.

Related news story 1: “First Wind, Now Gas: Tax Breaks Face Scrutiny”.

Url: <https://www.texastribune.org/2014/10/20/big-tax-break-natural-gas-get-scrutiny/>

Published in 10/20/2014.

Related news story 2: “Little Help for Those Who Can't Make the Rent”.

Url: <https://www.texastribune.org/2014/10/22/high-demand-expected-austin-section-8-waiting-list/>

Published in 10/22/2014.

THE TEXAS TRIBUNE PRESENTED BY at&t

2014 ELECTIONS

Welcome to The Texas Tribune's elections scoreboard. Track Texas general election race results in near real time and find out when races have been called for particular candidates. Sort races by party; search for races by candidate, district or address; and save specific races for easy check-ins. Create a customized list of saved races and click "embed" to add the Tribune's election results widget to your website. Top races ★ are those deemed by our editors to be particularly competitive or newsworthy. To get updates from our reporters, be sure to follow our election liveblog.

Twitter Recommend

View all
View my saved races
View undecided races
View called races
View top races ★
View located races 📍

Filter by candidate or district name
Enter your address to find your districts
Share saved races

All races Statewide Congress SBOE Texas Senate Texas House

Time to next update: 0:26

Race	Candidate	Votes	%
U.S. Sen. ★	R - J. Cornyn (R)	1,155,050	61.6%
	D - D. Alameel	1,584,252	34.4%
	L - R. Padgett	125,462	2.3%
	G - E. Sanchez	54,387	0.9%
U.S. House District 1	R - L. Gohmert (R)	114,923	77.5%
	D - S. McKellar	32,469	22.5%
U.S. House District 2	R - T. Poe (R)	101,195	68.0%
	D - N. Letcos	44,549	29.6%
	L - J. B. Wassaw	2,115	1.3%
	G - M. Roberts	1,107	0.3%
U.S. House District 3	R - E. Johnson (R)	412,726	82.0%
	G - P. Blair	24,759	16%
U.S. House District 4	R - J. Ratcliffe (R)	9	100%
U.S. House District 5	R - J. Hensarling (R)	55,633	85.3%
	L - K. Ashby	15,252	14.5%
U.S. House District 6	R - J. Barton (R)	82,205	61.4%
	D - D. Cozad	54,886	36.4%
	L - R. Chausen	3,534	2.4%
U.S. House District 7	R - J. Culberson (R)	80,422	63.3%
	D - J. Cargen	49,241	34.5%
U.S. House District 8	R - K. Brady (R)	144,831	89.3%
	L - K. Petty	14,930	10.7%
U.S. House District 9	R - A. Green (R)	71,979	90.8%
	L - J. Johnson	7,886	9.2%
U.S. House District 10	R - M. McZul (R)	109,311	82.4%
	D - T. Cadan	60,020	34.1%
	L - B. Kelley	6,475	3.1%
U.S. House District 11	R - M. Conaway (R)	197,752	90.5%
	L - R. Lange	11,607	9.7%
U.S. House District 12	R - K. Granger (R)	111,105	71.3%
	D - M. Greene	41,720	26.3%
	L - E. Collier	3,793	2.4%
U.S. House District 13	R - M. Thornberry (R)	110,636	84.9%
	D - M. Minlar	16,803	12.8%
	L - E. Pevada	2,969	2.2%
	G - D. Cook	926	0.7%
U.S. House District 14	R - R. Weber (R)	83,376	61.8%
	D - D. Brown	52,420	36.1%
	L - J. Wieder	3,025	2.1%
U.S. House District 15	R - R. Himes (R)	45,561	64.2%
	R - E. Zamora	38,551	43.1%
U.S. House District 16	R - K. O'Rourke (R)	49,251	61.5%
	R - C. Roan	21,293	28.2%
U.S. House District 17	R - B. Flores (R)	55,055	64.9%
	D - N. Haynes	42,896	32.4%
U.S. House District 18	D - S. Jackson-Lee (D)	75,963	71.8%
	R - S. Seibert	29,194	24.3%
	L - W. Duncan	2,159	2.1%
U.S. House District 19	R - R. Neugebauer (R)	89,201	77.2%
	D - N. Marchbanks	21,250	18.4%

Figure 71 – “Election Night” screenshot.

Title: “Election Night”.

Original url: <https://www.texastribune.org/2014/elections/scoreboard/>

Published in 11/04/2014.

Archived url: <http://bit.ly/2YWCVS4>

Interaction techniques: Filter, select.

Topic: Politics/governance.

Related news story 1: “Patrick Wins Big in Bid for Lieutenant Governor”.

Url: <https://www.texastribune.org/2014/11/04/patrick-wins-big-bid-lieutenant-governor/>

Published in 11/04/2014.

Related news story 2: “Greg Abbott Crushes Wendy Davis in GOP Sweep”.

Url: <https://www.texastribune.org/2014/11/04/abbott-crushes-wendy-davis-gop-sweep/>

Published in 11/04/2014.

Register | Login

FRONT PAGE DATA DIRECTORY EVENTS VIDEO NEWSLETTERS TRIBUTALK PROJECTS #HARVEY FESTIVAL Search

THE TEXAS TRIBUNE SUPPORT US

Elected Officials Directory SHARE THIS PAGE: | | |

As part of The Texas Tribune's commitment to help Texans know more about their representatives, we've gathered the bios, contact information and more for the occupants of the state's highest offices in our Elected Officials Directory. You can learn about the districts they represent, who's serving on their staff, where they sit in the House and Senate, and their financial disclosures.

Who represents you?

Enter your ZIP code

State Officials Texas House Texas Senate U.S. Congress

Governor			Supreme Court of Texas		
Abbott, Greg	Governor	R	Boyd, Jeffrey S.	Justice	R
Lieutenant Governor			Brown, Jeff	Justice	R
Patrick, Dan	Lieutenant Governor	R	Devine, John	Justice	R
Texas Comptroller of Public Accounts			Green, Paul	Justice	R
Hegar, Glenn	Comptroller	R	Guzman, Eva	Justice	R
Attorney General			Hecht, Nathan L.	Chief Justice	R
Paxton, Ken	Attorney General	R	Johnson, Phil	Justice	R
Texas General Land Office			Lehmann, Debra	Justice	R
Bush, George P.	Commissioner	R	Willett, Don	Justice	R
Texas Department of Agriculture			Texas Court of Criminal Appeals		
Miller, Sid	Commissioner	R	Alcala, Elsa	Judge	R
Texas Railroad Commission			Keasler, Michael	Judge	R
Christlan, Wayne	Commissioner	R	Keel, Mary Lou	Judge	R
Craddock, Christi	Commissioner	R	Keller, Sharon	Presiding Judge	R
Silton, Ryan	Commissioner	R	Newell, David	Judge	R
			Parker Harvey, Barbara	Judge	R
			Patrick Yeary, Kevin	Judge	R
			Richardson, Bert	Judge	R
			Walker, Scott	Judge	R

Figure 72 – “Elected Officials Directory” screenshot.

Title: “Elected Officials Directory”.

Original url: <https://www.texastribune.org/directory/>

Updated in 12/21/2014.

Archived url: <http://bit.ly/2XRnrh9>

Interaction techniques: Filter, explore, reconfigure/encode.

Topic: Politics/governance.

Related news story 1: “Farenthold's Challenges May Go Beyond Lawsuit”.

Url: <https://www.texastribune.org/2014/12/17/farentholds-challenges-may-go-beyond-lawsuit/>

Published in 12/17/2014.

Related news story 2: “Dan Patrick Fills Out Staff Roster”.

Url: <https://www.texastribune.org/2014/12/19/dan-patrick-fills-out-staff-roster/>

Published in 12/19/2014.

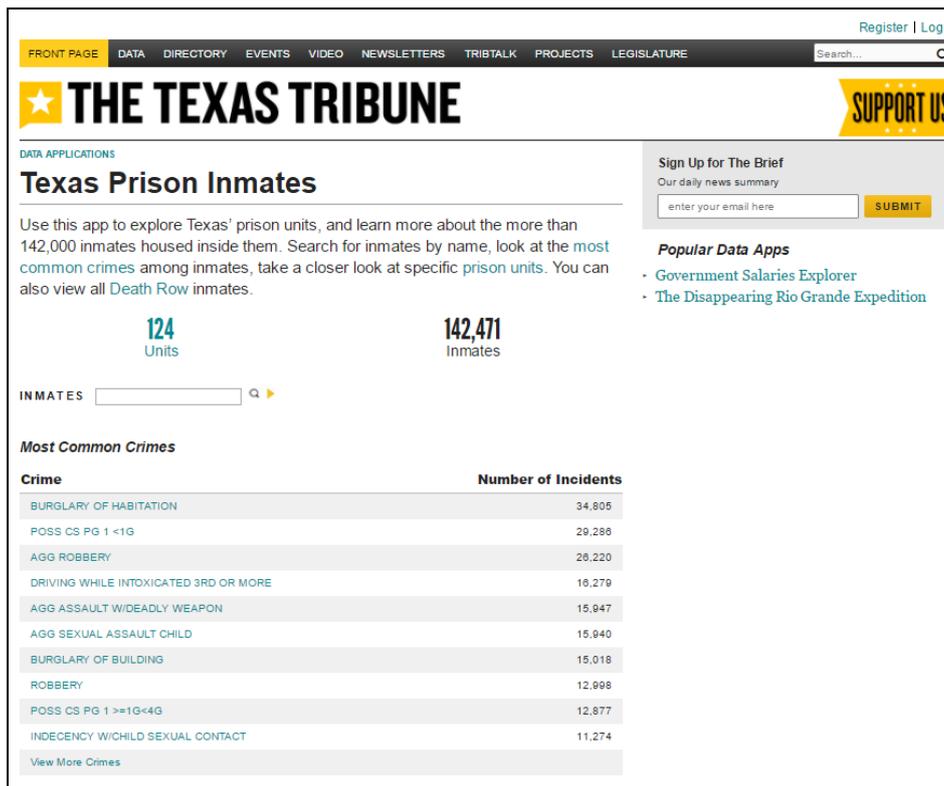


Figure 73 – “Texas Prison Inmates” screenshot.

Title: “Texas Prison Inmates”.

Original url: <https://www.texastribune.org/library/data/texas-prisons/>

Updated in 11/04/2014.

Archived url: <http://bit.ly/2Ls9unP>

Interaction techniques: Explore, inspect.

Topic: Social issues.

Related news story 1: “Prison Officials Seek Ways to Recruit, Retain Officers”.

Url: <https://www.texastribune.org/2014/11/14/prison-officials-seek-ways-recruit-retain-guard/>

Published in 11/14/2014.

Related news story 2: “Criminal Appeals Judge Price: I Oppose Death Penalty”.

Url: <https://www.texastribune.org/2014/11/26/criminal-appeals-judge-price-i-oppose-death-penalt/>

Published in 11/26/2014.

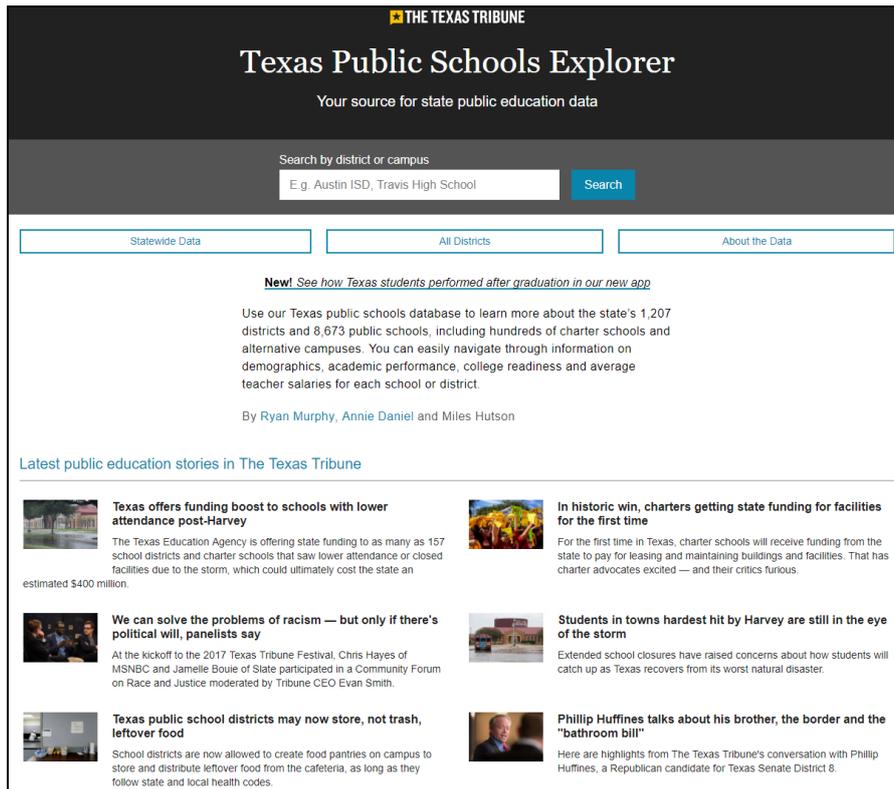


Figure 74 – “Public Schools Explorer” screenshot.

Title: “Public Schools Explorer”.

Original url: <https://schools.texastribune.org/>

Updated in 07/06/2014.

Archived url: <http://bit.ly/2Y9VHZ0>

Interaction technique: Explore.

Topic: Education.

Related news story 1: “Despite Success, Some Shortcomings in Texas Higher Ed”.

Url: <https://www.texastribune.org/2014/07/08/deadline-nears-closing-gaps-higher-education/>

Published in 07/08/2014.

Related news story 2: “As Teacher Pay Lags, Attrition and Class Size Grow”.

Url: <https://www.texastribune.org/2014/07/11/teacher-pay-lags-attrition-and-class-sizes-grow/>

Published in 07/11/2014.

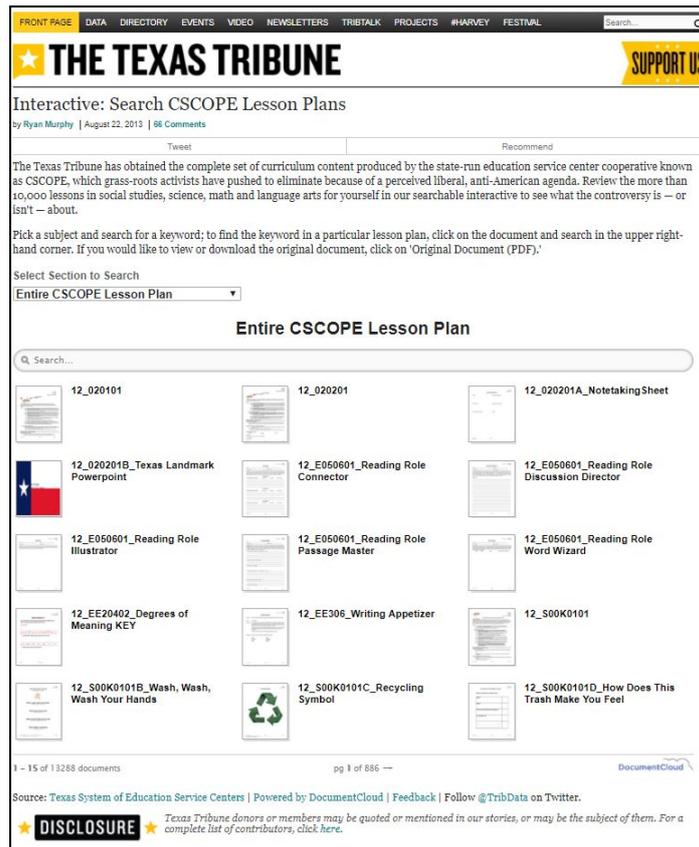


Figure 75 – “Interactive: Search CSCOPE Lesson Plans” screenshot.

Title: “Interactive: Search CSCOPE Lesson Plans”.

Original url: <https://www.texastribune.org/interactive/search-cscope-lesson-plans/>

Updated in 08/26/2014.

Archived url: <http://bit.ly/2JBRYLG>

Interaction techniques: Filter, explore.

Topic: Education.

Related news story 1: “Davis Draws Praise, Criticism for Education Plans”.

Url: <https://www.texastribune.org/2014/08/27/wendy-davis-gets-praise-education-plans/>

Published in 08/27/2014.

Related news story 2: “A Pass on State Math Exams for Fifth- and Eighth-Graders”.

Url: <https://www.texastribune.org/2014/08/22/pass-state-math-exams-fifth-and-eighth-graders/>

Published in 08/28/2014.

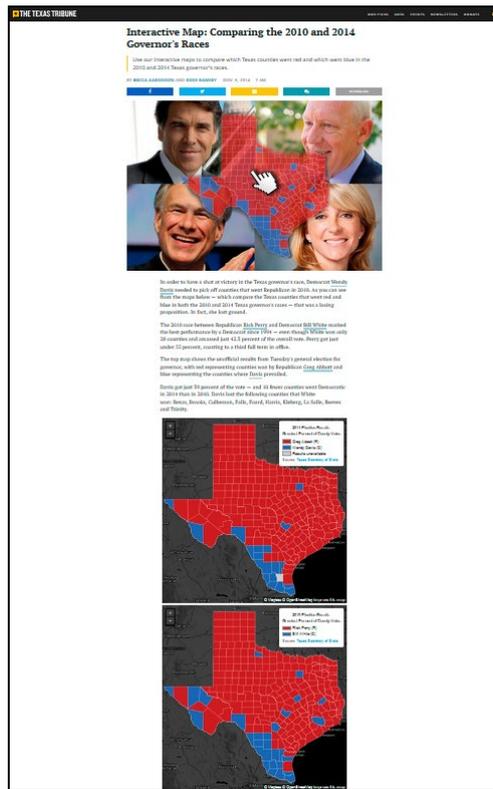


Figure 76 – “Map: Comparing the 2010 and 2014 Governor's Races” screenshot.

Title: “Map: Comparing the 2010 and 2014 Governor's Races”.

Original url: <https://www.texastribune.org/2014/11/05/interactive-compare-governors-race-2010-2014/>

Published in 11/05/2014.

Archived url: <http://bit.ly/32zbiAQ>

Interaction techniques: Inspect, abstract/elaborate, select.

Topic: Politics/governance.

Related news story 1: “What Went Wrong With Battleground Texas?”.

Url: <https://www.texastribune.org/2014/11/05/little-show-battleground-texas-fights/>

Published in 11/05/2014.

Related news story 2: “Republicans Extend Statewide Streak to 16 Years”.

Url: <https://www.texastribune.org/2014/11/04/republicans-extend-statewide-streak-16-years/>

Published in 11/04/2014.

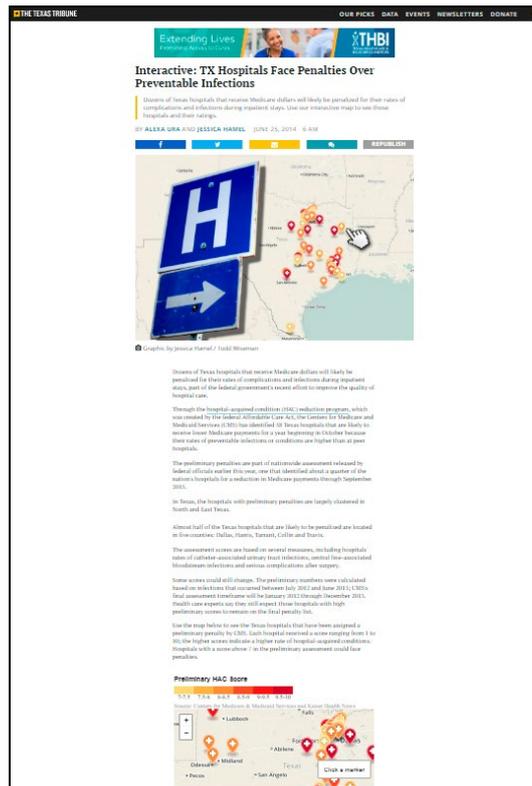


Figure 77 – “Texas Hospitals Face Penalties Over Infections” screenshot.

Title: “Texas Hospitals Face Penalties Over Infections”.

Original url: <https://www.texastribune.org/2014/06/25/interactive-tx-hospitals-likely-lose-some-medicare/>

Published in 06/25/2014.

Archived url: <http://bit.ly/2Y2hBJS>

Interaction techniques: Inspect, abstract/elaborate, select.

Topic: Health.

Related news story 1: “Texas Conservatives Laud Court Ruling on Birth Control”.

Url: <https://www.texastribune.org/2014/06/30/cases-highlight-continued-uphill-battle-womens-hea/>

Published in 06/30/2014.

Related news story 2: “Timeline: A History of Lethal Drug Use in Texas”.

Url: <https://www.texastribune.org/2014/07/08/history-lethal-drug-use-texas/>

Published in 07/08/2014.

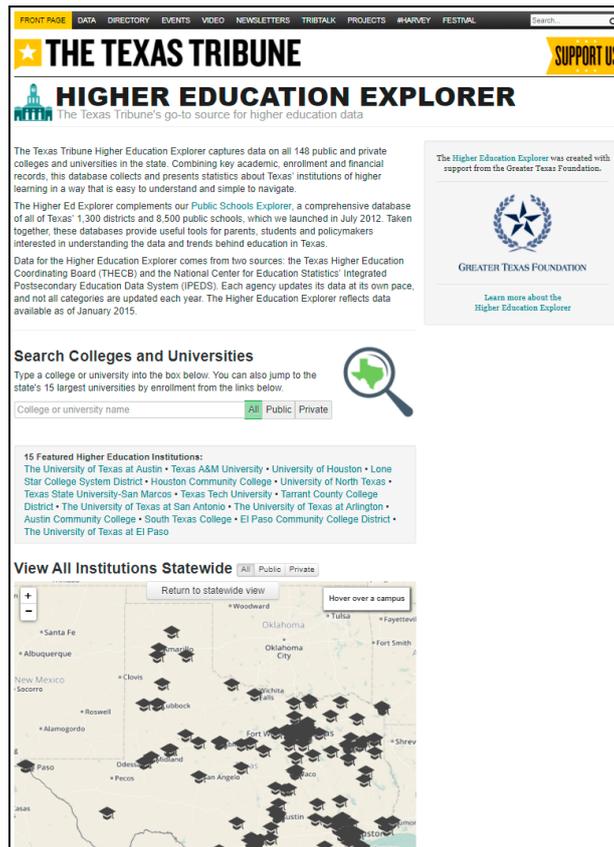


Figure 78 – “Higher Education Explorer” screenshot.

Title: “Higher Education Explorer”.

Original url: Not available.

Updated in 07/03/2014.

Archived url: <http://bit.ly/2YTbAAh>

Interaction techniques: Explore, inspect, abstract/elaborate, select.

Topic: Education.

Related news story 1: “TEA Chief Circumvents State Board Charter School Veto”.

Url: <https://www.texastribune.org/2014/07/02/tea-chief-goes-around-state-board-charter-school-v/>

Published in 07/02/2014.

Related news story 2: “Bruce Zimmerman: The TT Interview”.

Url: <https://www.texastribune.org/2014/07/03/bruce-zimmerman-tt-interview/>

Published in 07/03/2014.

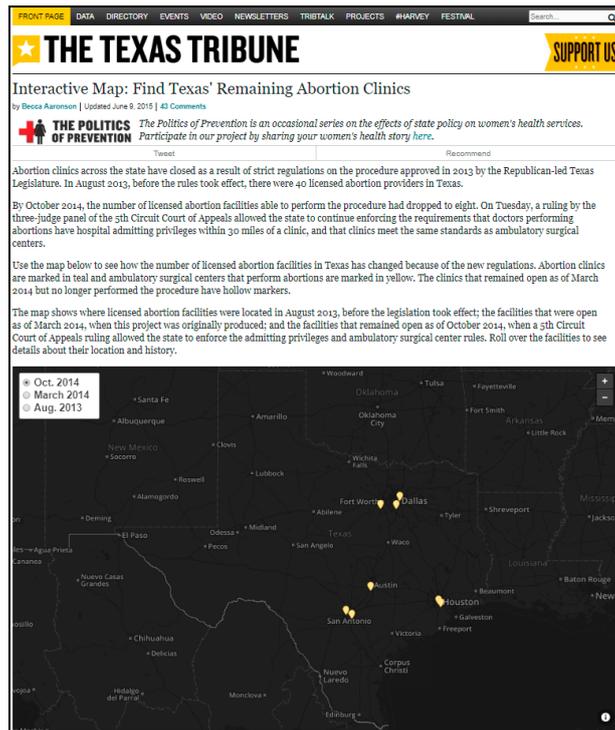


Figure 79 – “Interactive Map: Find Texas' Remaining Abortion Clinics” screenshot.

Title: “Interactive Map: Find Texas' Remaining Abortion Clinics”.

Original url: <https://www.texastribune.org/2014/03/19/impact-hb2-regulations-abortion-facilities-over-time/>

Archived url: <http://bit.ly/2SmWTmq>

Updated in 03/19/2014.

Interaction techniques: Inspect, filter, abstract/elaborate.

Topic: Health.

Related news story 1: “Abortion Provider Shuttters McAllen and Beaumont Clinics”.

Url: <https://www.texastribune.org/2014/03/06/whole-womans-shuttters-mcallen-and-beaumont-clinics/>

Published in 03/06/2014.

Related news story 2: “As Abortion Clinics Close, Student Creates Travel Fund”.

Url: <https://www.texastribune.org/2014/03/12/abortion-clinics-close-advocates-fund-farther-trav/>

Published in 03/12/2014.

A.2. Interactive pieces from 2015



Figure 80 – “Interactive: Undocumented Students on In-State Tuition” screenshot.

Title: “Interactive: Undocumented Students on In-State Tuition”.

Original url: <https://www.texastribune.org/2015/04/16/colleges-undocumented-students-with-state-tuition/>

Published in 04/16/2015.

Archived url: <http://bit.ly/2Ydx9hP>

Interaction technique: Explore.

Topic: Education.

Related news story 1: “UT Must Tussle With Outsiders for Strong and Smart Trademarks”.

Url: <https://www.texastribune.org/2015/04/15/ut-must-tussle-outsiders-strong-and-smart-trademar/>

Published in 04/15/2015.

Related news story 2: “Fenves Named Next President of UT-Austin”.

Url: <https://www.texastribune.org/2015/04/20/fenves-named-next-president-ut-austin/>

Published in 04/20/2015.

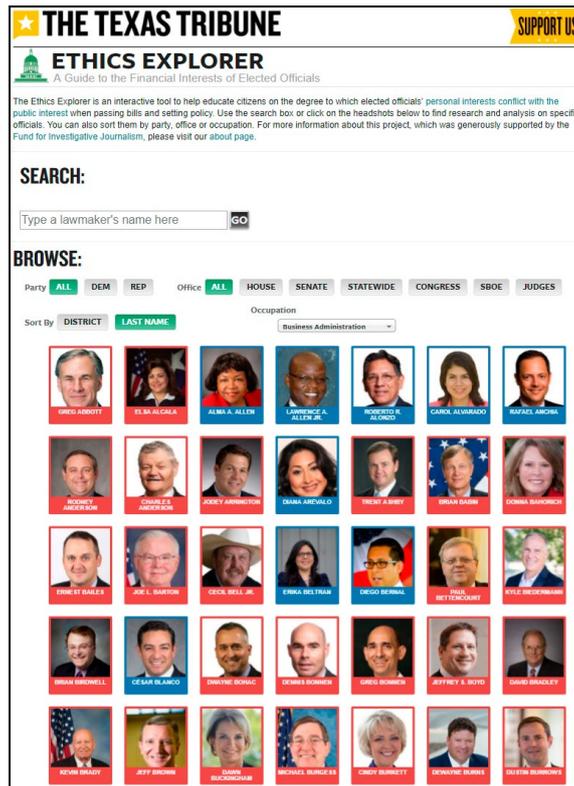


Figure 81 – “Ethics Explorer” screenshot.

Title: “Ethics Explorer”.

Original url: Not available.

Updated in 01/13/2015.

Archived url: <http://bit.ly/2Y58UCB>

Interaction techniques: Filter, explore, reconfigure/encode.

Topics: Politics/governance, economy/business.

Related news story 1: “UT Dream Act Supporters Rally Against Repeal Efforts”.

Url: <https://www.texastribune.org/2015/01/14/dream-act-supporters-rally-fight-repeal-efforts/>

Published in 01/14/2015.

Related news story 2: “Rule on Natural Gas Rate Hikes is Fuel for Debate”.

Url: <https://www.texastribune.org/2015/01/11/cities-say-rule-could-raise-natural-gas-bills/>

Published in 01/11/2015.

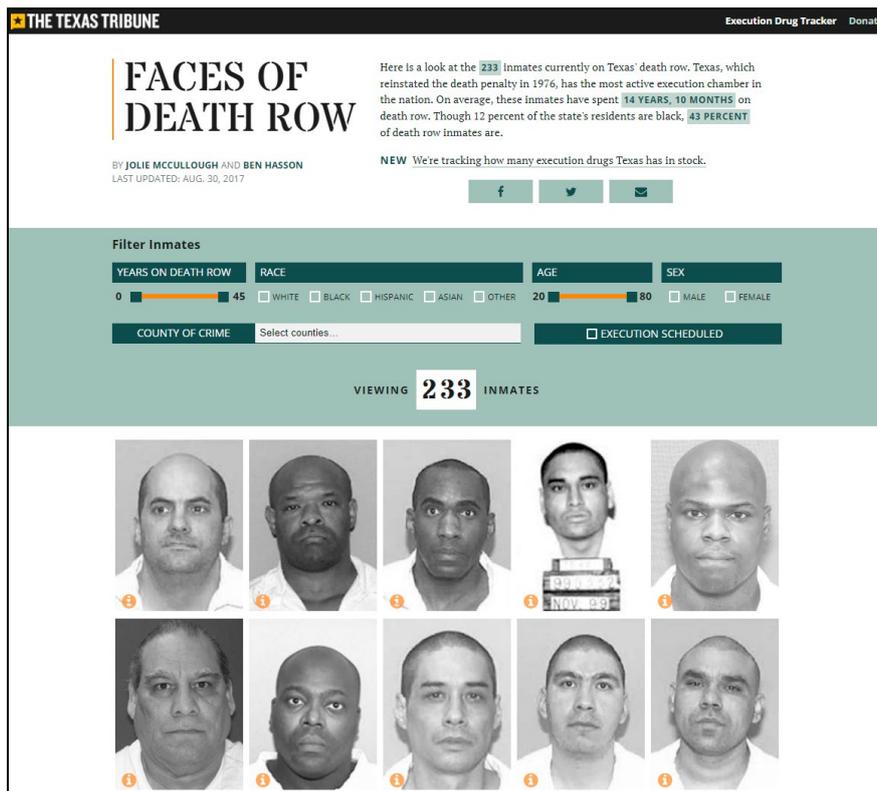


Figure 82 – “Faces of Death Row” screenshot.

Title: “Faces of Death Row”.

Original url: <https://apps.texastribune.org/death-row/>

Published in 06/15/2015.

Archived url: <http://bit.ly/30EKpK8>

Interaction techniques: Filter, connect, inspect, explore, select.

Topic: Social issues.

Related news story 1: “Preying on Texas Prisoners: When Guards Demand Sex”.

Url: <https://www.texastribune.org/2015/06/17/preying-texas-prisoners-when-guards-demand-sex/>

Published in 06/17/2015.

Related news story 2: “Prosecutor Disbarred for Wrongful Death Row Conviction”.

Url: <https://www.texastribune.org/2015/06/12/prosecutor-anthony-graves-case-disbarred/>

Published in 06/12/2015.

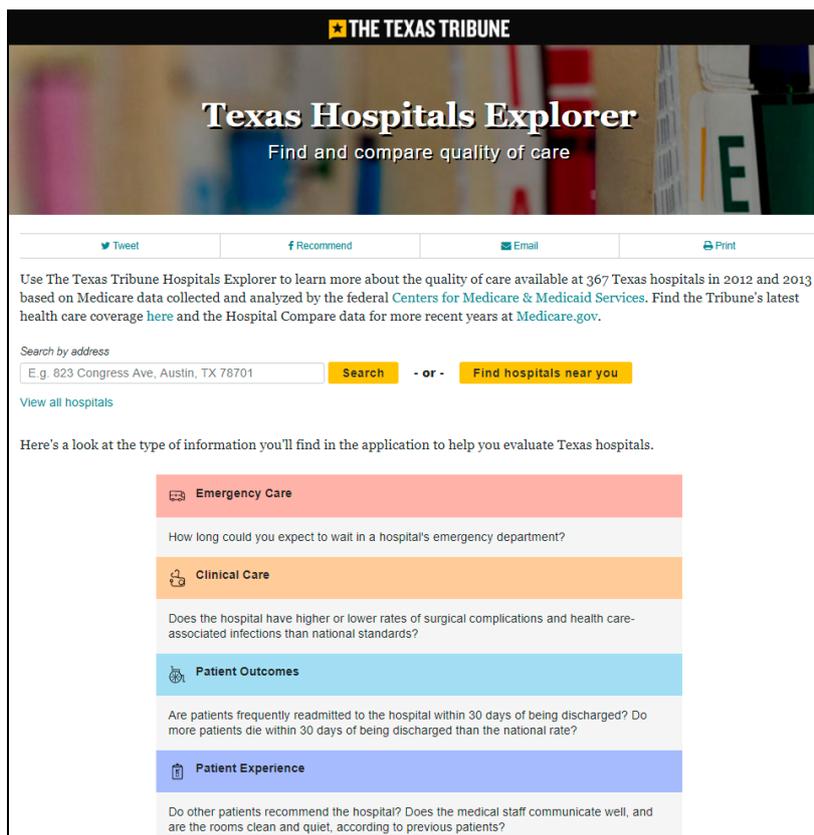


Figure 83 – “Texas Hospitals Explorer” screenshot.

Title: “Texas Hospitals Explorer”.

Original url: <https://hospitals.texastribune.org>

Updated in 02/03/2015.

Archived url: <http://bit.ly/2Ybe6oj>

Interaction techniques: Explore.

Topic: Health.

Related news story 1: “The Health Care Budget: Four Things to Know”.

Url: <https://www.texastribune.org/2015/01/27/four-things-know-about-new-health-budget/>

Published in 01/27/2015.

Related news story 2: “Drug Testing for Welfare Benefits Back on the Table”.

Url: <https://www.texastribune.org/2015/02/05/drug-testing-welfare-benefits-back-table/>

Published in 02/05/2015.

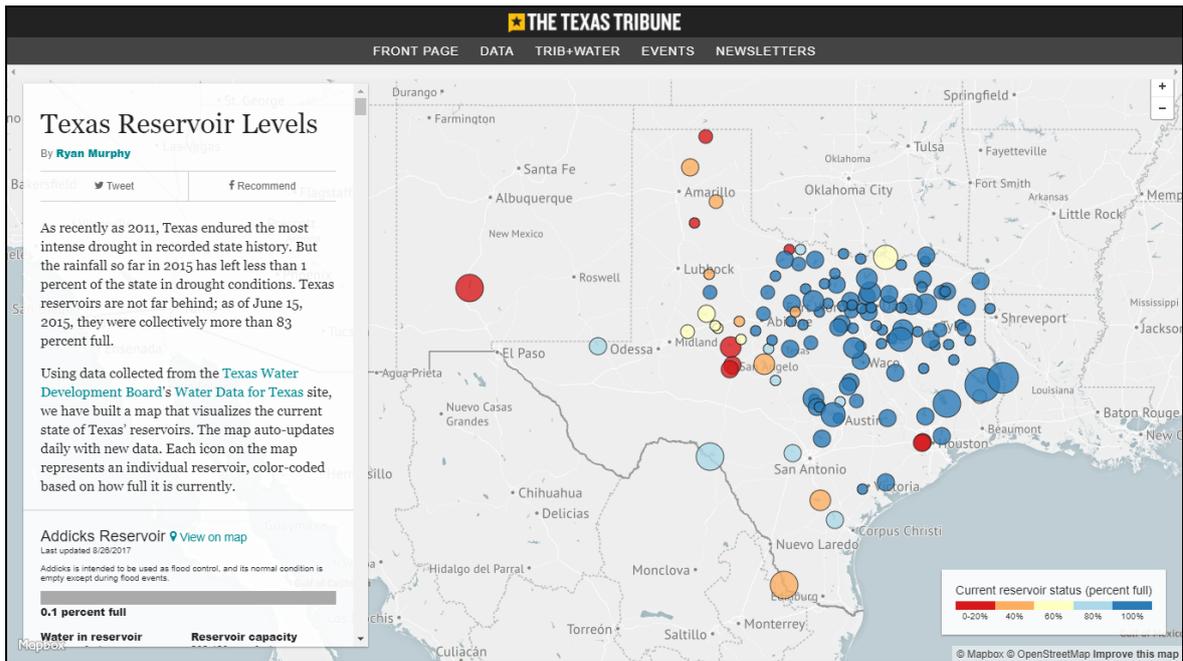


Figure 84 – “Texas Reservoir Levels” screenshot.

Title: “Texas Reservoir Levels”.

Original url: <https://apps.texastribune.org/reservoirs/>

Updated in 05/15/2015.

Archived url: <http://bit.ly/2XT7BCw>

Interaction techniques: Inspect, select, abstract/elaborate.

Topic: Environment/science/technology.

Related news story 1: “The Q&A: Mike Gershon”.

Url: <https://www.texastribune.org/2015/05/19/q-mike-gershon/>

Published in 05/19/2015.

Related news story 2: “Harris County in Crosshairs of Pollution Lawsuit Limits”.

Url: <https://www.texastribune.org/2015/05/20/senate-backs-bill-cap-pollution-payouts/>

Published in 05/20/2015.

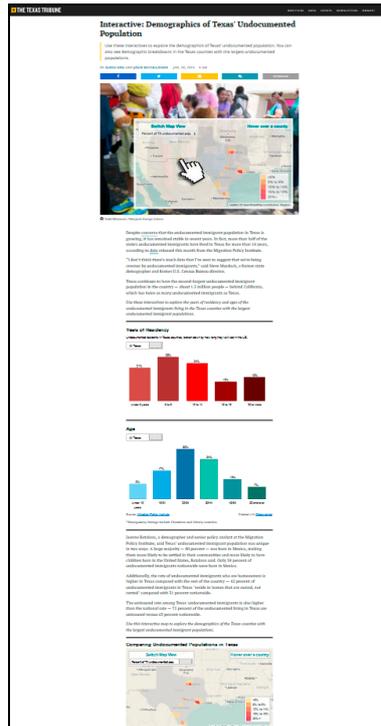


Figure 85 – “Interactive: Demographics of Texas' Undocumented Population” screenshot.

Title: “Interactive: Demographics of Texas' Undocumented Population”.

Original url: <https://www.texastribune.org/2015/01/28/undocumented-population-demographics/>

Published in 01/28/2015.

Archived url: <http://bit.ly/2LqpE1d>

Interaction techniques: Inspect, filter, abstract/elaborate.

Topic: Social issues.

Related news story 1: “GOP Hopefuls Eyeing the Texas Hispanic Vote”.

Url: <https://www.texastribune.org/2015/01/25/gop-presidential-contenders-and-hispanic-vote/>

Published in 01/25/2015.

Related news story 2: “Mexican Governors Won't Be at Abbott's Inauguration”.

Url: <https://www.texastribune.org/2015/01/19/mexican-governors-wont-be-abbotts-inauguration/>

Published in 01/19/2015.

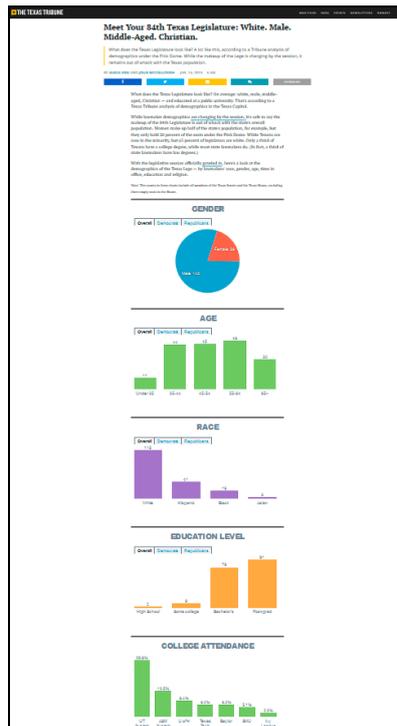


Figure 86 – “The 84th Texas Legislature, by the Numbers” screenshot.

Title: Interactive: “The 84th Texas Legislature, by the Numbers”.

Original url: <https://www.texastribune.org/2015/01/14/demographics-2015-texas-legislature/>

Published in 01/14/2015.

Archived url: <http://bit.ly/32CcrYy>

Interaction techniques: Inspect, filter.

Topic: Politics/governance.

Related news story 1: “Lt. Gov.-Elect Taps Business Leaders for New Advisory Board”.

Url: <https://www.texastribune.org/2015/01/15/lt-gov-elect-creates-special-policy-advisory-board/>

Published in 01/15/2015.

Related news story 2: “Straus Re-elected House Speaker Over Tea Party Challenger”.

Url: <https://www.texastribune.org/2015/01/13/straus-re-elected-house-speaker-over-tea-party-cha/>

Published in 01/13/2015.

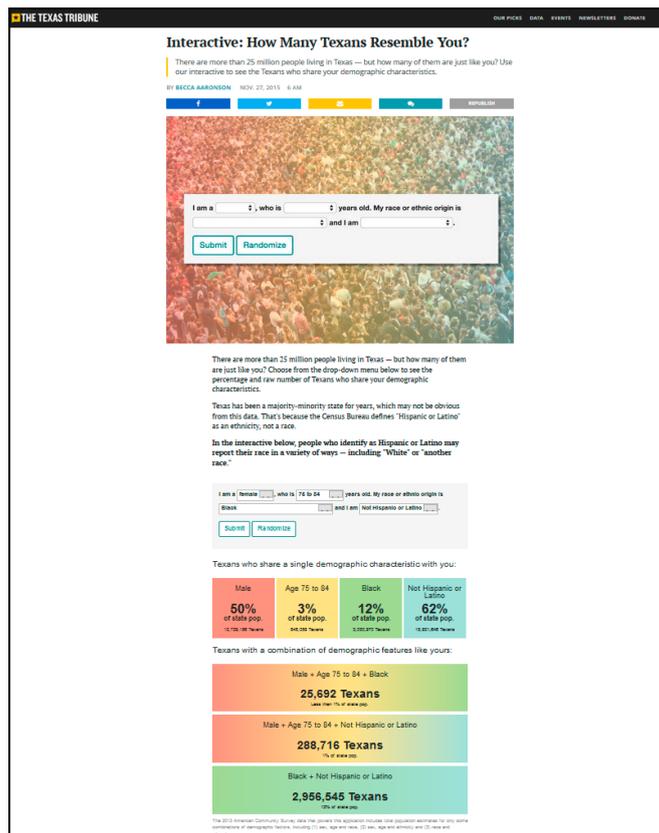


Figure 87 – “Interactive: How Many Texans Resemble You?” screenshot.

Title: “Interactive: How Many Texans Resemble You?”.

Original url: <https://www.texastribune.org/2015/11/27/how-many-texans-are-me/>

Published in 11/27/2015.

Archived url: <http://bit.ly/30B0L6r>

Interaction techniques: Gamification, filter.

Topic: Social issues.

Related news story 1: “As Texas Grows, More Languages are Spoken at Home”.

Url: <https://www.texastribune.org/2015/11/26/languages-spoken-texas-homes/>

Published in 11/26/2015.

Related news story 2: “Video: Same-Sex Couples Navigate Immigration Benefits”.

Url: <https://www.texastribune.org/2015/11/21/video-gay-couples-navigate-immigration-benefits/>

Published in 11/21/2015.



Figure 88 – “Texas vs. the Feds” screenshot.

Title: “Texas vs. the Feds”.

New Url: <https://www.texastribune.org/2017/01/17/texas-federal-government-lawsuits/>

Published in 02/14/2015.

Archived url: <http://bit.ly/2YWQfpr>

Interaction techniques: Filter, select, connect.

Topic: Politics/governance.

Related news story 1: “Ethics Commission Approves Pay Increase for Lawmakers”.

Url: <https://www.texastribune.org/2015/02/13/ethics-commission-approves-higher-diems-lawmakers/>

Published in 02/13/2015.

Related news story 2: “Report: Powers Overruled UT Admissions Office”.

Url: <https://www.texastribune.org/2015/02/12/ut-system-releases-admissions-report/>

Published in 02/12/2015.



Figure 89 – “See Vaccine Exemptions in Texas by School District” screenshot.

Title: “See Vaccine Exemptions in Texas by School District”.

Original url: <https://www.texastribune.org/2015/02/05/school-vaccine-exemptions-high-pockets-texas/>

Published in 02/05/2015.

Archived url: <http://bit.ly/30Eniz4>

Interaction techniques: Filter, select, connect.

Topic: Politics/governance.

Related news story 1: “Texas Republican Calls for Limiting Vaccine Exemptions”.

Url: <https://www.texastribune.org/2015/02/06/texas-republican-calls-close-vaccine-exemptions/>

Published in 02/06/2015.

Related news story 2: “Texas Has High Stakes in Lawsuit Over Health Law”.

Url: <https://www.texastribune.org/2015/02/10/texas-has-high-stakes-battle-over-health-law/>

Published in 02/10/2015.

A.3. Interactive pieces from 2016



Figure 90 – “Ballpark Figures” screenshot.

Title: “Ballpark Figures”.

Original url: <https://college-sports.texastribune.org>

Published in 03/17/2016.

Archived url: <http://bit.ly/2LrZrPB>

Interaction technique: Explore.

Topics: Sports, economy/business.

Related news story 1: “At Texas Colleges, Football Revenue Props Up Other Sports”.

Url: <https://www.texastribune.org/2016/03/17/how-college-football-props-entire-athletic-departm/>

Published in 03/17/2016.

Related news story 2: “Students Pitch In More as Texas Colleges’ Athletics Costs Climb”.

Url: <https://www.texastribune.org/2016/03/18/texas-universities-are-hiking-student-fees-support/>

Published in 03/18/2016.

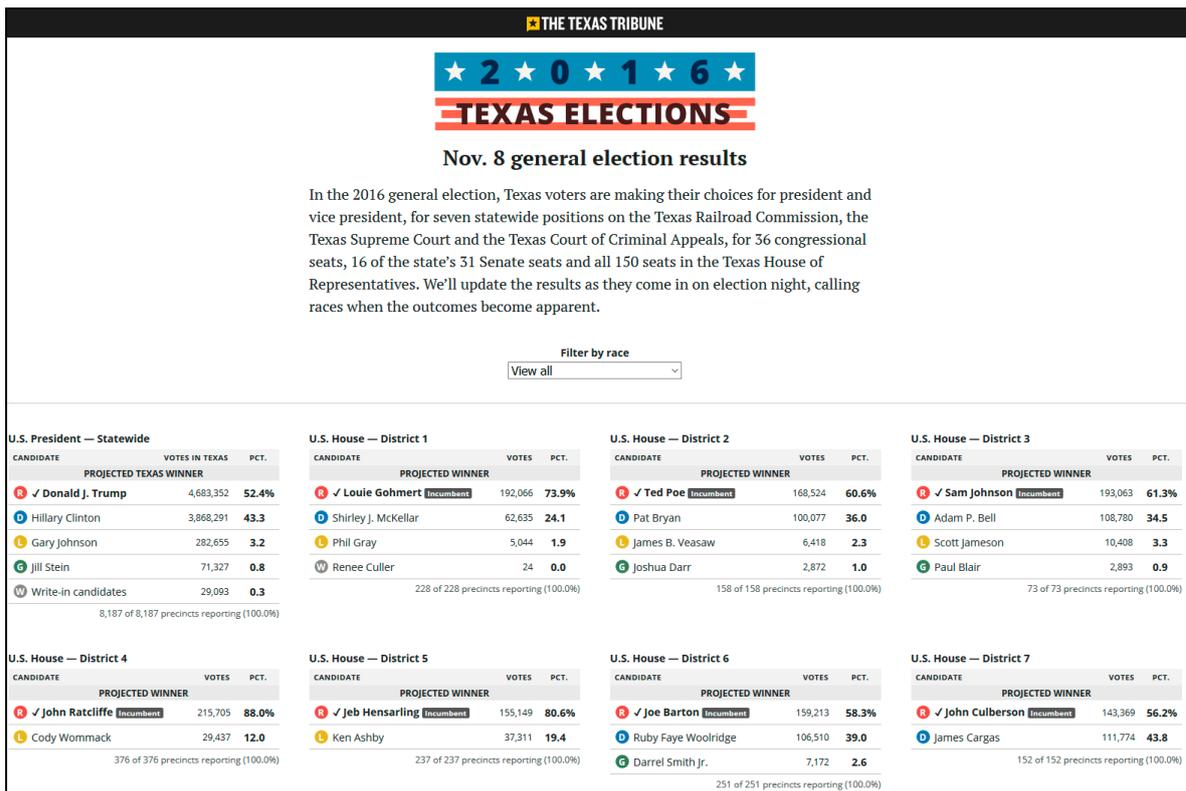


Figure 91 – “Nov. 8 general election results” screenshot.

Title: “Nov. 8 general election results”.

Original url: <https://elections.texastribune.org/2016/general-election-results/>

Published in 11/08/2016.

Archived url: <http://bit.ly/2XZNEK6>

Interaction technique: Filter.

Topic: Politics/governance.

Related news story 1: “Despite resignation, Dukes handily wins re-election to Texas House”.

Url: <https://www.texastribune.org/2016/11/08/dawna-dukes-gabriel-nila/>

Published in 11/08/2016.

Related news story 2: “Democrats pick up four Texas House seats”.

Url: <https://www.texastribune.org/2016/11/08/texas-house-race-results-election-day-2016/>

Published in 11/08/2016.

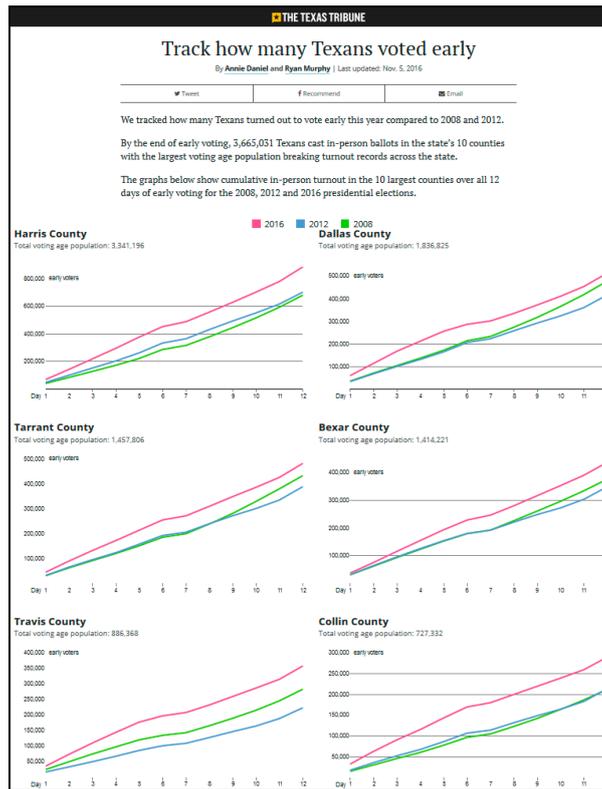


Figure 92 – “Track how many Texans are voting early” screenshot.

Title: “Track how many Texans are voting early”.

Original url: <https://elections.texastribune.org/2016/general-election-results/>

Published in 10/26/2016.

Archived url: <http://bit.ly/2xTWFkS>

Interaction technique: Inspect.

Topic: Politics/governance.

Related news story 1: “Amid early voting rush, Texas sees voter ID hiccups”.

Url: <https://www.texastribune.org/2016/10/25/amid-early-voting-rush-texas-sees-voter-id-hiccups/>

Published in 10/25/2016.

Related news story 2: “Early voting is breaking records in Texas’ 10 biggest counties”.

Url: <https://www.texastribune.org/2016/10/25/early-voting-breaking-records-texas-10-biggest-cou/>

Published in 10/25/2016.

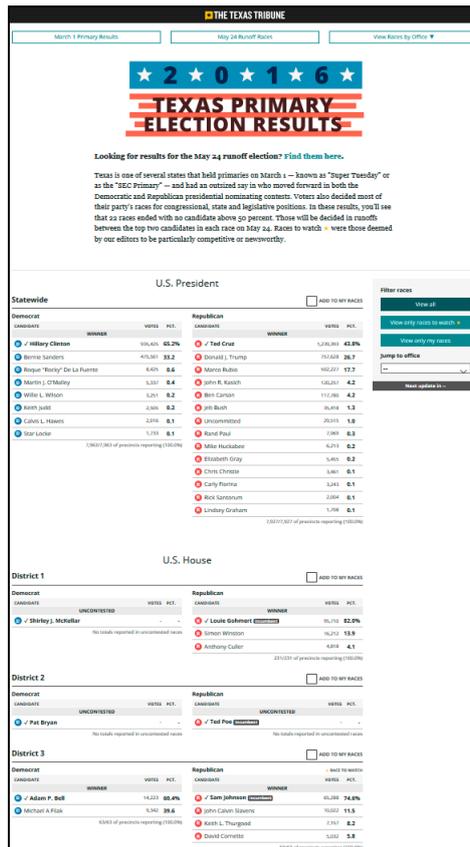


Figure 93 – “2016 Primary Election Results” screenshot.

Title: “2016 Primary Election Results”.

Original url: <https://elections.texastribune.org/2016/primary-election-results/>

Published in 03/01/2016.

Archived url: <http://bit.ly/2LtfexS>

Interaction techniques: Select, filter.

Topic: Politics/governance.

Related news story 1: “Cruz, Clinton Grab Most Votes in Almost Every Texas County”.

Url: <https://www.texastribune.org/2016/03/02/cruz-clinton-grab-most-votes-almost-every-county/>

Published in 03/02/2016.

Related news story 2: “Travis County GOP Apoplectic Over New Chairman”.

Url: <https://www.texastribune.org/2016/03/02/newly-elected-gop-chair-texas-capitol/>

Published in 03/02/2016.

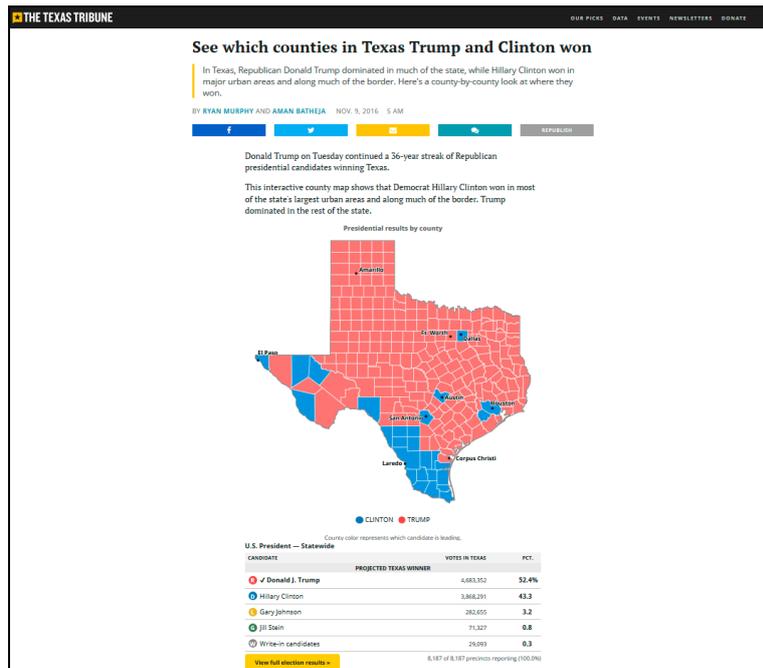


Figure 94 – “See which counties in Texas Trump and Clinton won” screenshot.

Title: “See which counties in Texas Trump and Clinton won”.

Original url: <https://www.texastribune.org/2016/11/09/see-which-counties-texas-trump-and-clinton-won/>

Published in 11/09/2016.

Archived url: <http://bit.ly/2Lpulbm>

Interaction technique: Inspect.

Topic: Politics/governance.

Related news story 1: “Trump's shocking upset emboldens Texas GOP”.

Url: <https://www.texastribune.org/2016/11/09/trumps-impact-texas-gop-unlikely-soon-fade/>

Published in 11/09/2016.

Related news story 2: “George W. Bush congratulates Trump for win after not voting for him”.

Url: <https://www.texastribune.org/2016/11/09/george-w-bush-congratulates-trump-win-after-not-vo/>

Published in 11/09/2016.

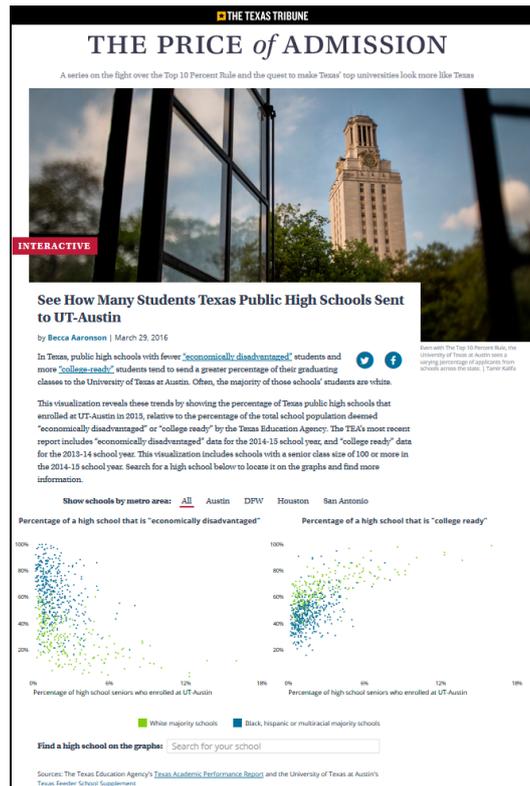


Figure 95 – “The Price of Admission” screenshot.

Title: “The Price of Admission”.

Original url: <https://apps.texastribune.org/price-of-admission/interactive/>

Published in 03/29/2016.

Archived url: <http://bit.ly/2O3Mfmd>

Interaction techniques: Inspect, filter, explore, select.

Topics: Education, economy/business.

Related news story 1: “Guns Up: Texas Tech Will Allow Firearms in Classrooms, Some Dorms”.

Url: <https://www.texastribune.org/2016/03/29/guns-texas-tech-wont-ban-guns-classrooms-some-dorm/>

Published in 03/29/2016.

Related news story 2: “Texas Behind in Preparing Kids for College, Panel Told”.

Url: <https://www.texastribune.org/2016/03/29/senators-tackle-college-readiness/>

Published in 03/29/2016.

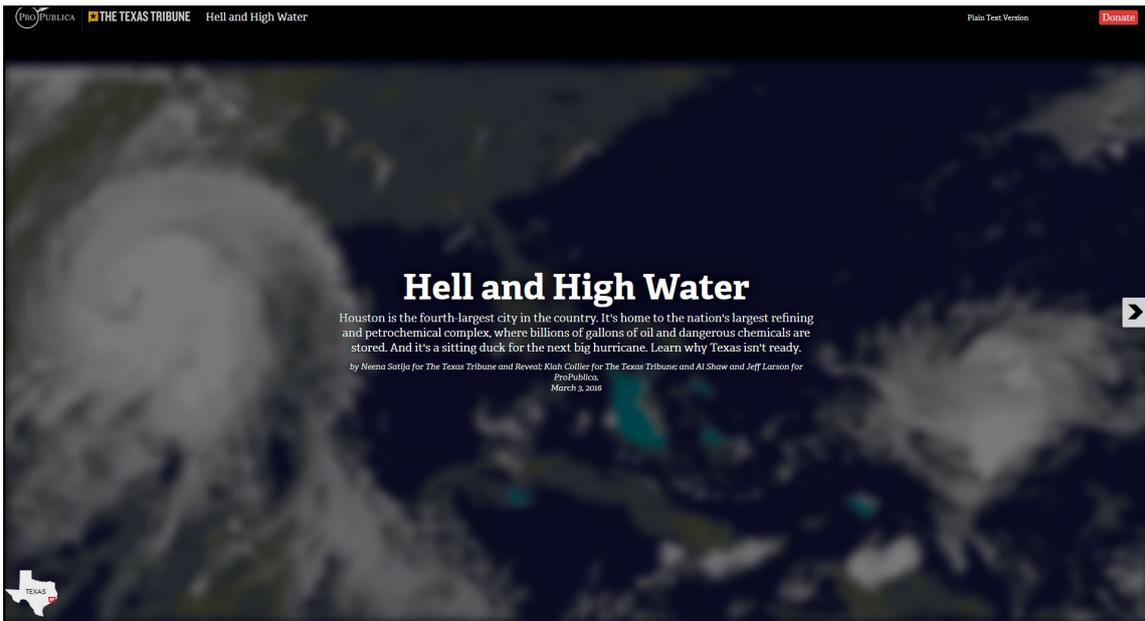


Figure 96 – “Why Isn’t Texas Ready for the Next Big Hurricane?” screenshot.

Title: “Why Isn’t Texas Ready for the Next Big Hurricane?”.

Original url: <https://houston.texastribune.org/hell-and-high-water/>

Published in 03/03/2016.

Archived url: <http://bit.ly/32wc8yp>

Interaction techniques: Narrate/history, explore, select, filter.

Topics: Environment/science/technology, politics, governance, economy/business.

Related news story 1: “Reveal Radio: A Monster Storm in the Making”.

Url: <https://www.texastribune.org/2016/03/05/reveal-radio-mighty-ike-monster-hurricane/>

Published in 03/05/2016.

Related news story 2: “Billions At Stake for State in Tax Challenge”.

Url: <https://www.texastribune.org/2016/03/07/billions-dollars-stake-court-hear-oil-drillers-tax/>

Published in 03/07/2016.

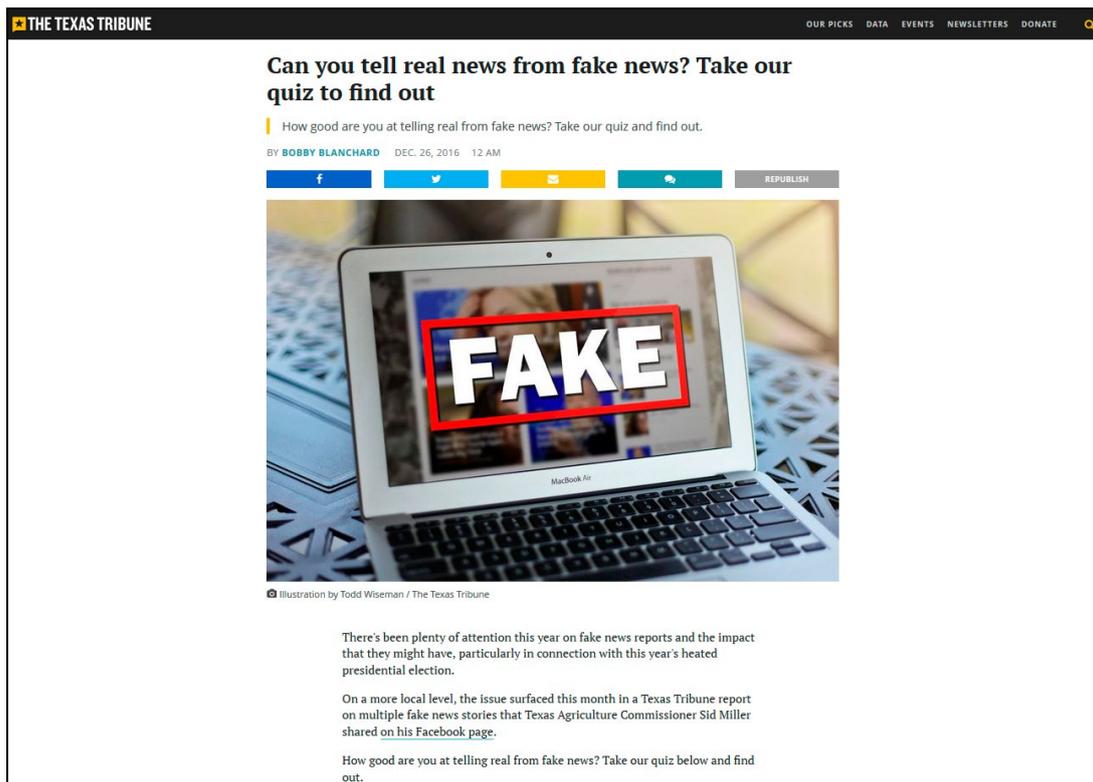


Figure 97 – “Can you tell real news from fake news? Take our quiz to find out” screenshot.

Title: “Can you tell real news from fake news? Take our quiz to find out”.

Original url: <https://www.texastribune.org/2016/12/26/can-you-tell-real-news-fake-news/>

Published in 12/26/2016.

Archived url: <http://bit.ly/2O2Ki9x>

Interaction techniques: Gamification, collaboration, narrate/history.

Topic: Entertainment.

Related news story 1: “Five stories that marked 2016 in Texas politics”.

Url: <https://www.texastribune.org/2016/12/25/Five-stories-marked-2016-Texas-politics/>

Published in 12/25/2016.

Related news story 2: “On Texas ag chief Sid Miller's Facebook, fake news flows freely”.

Url: <https://www.texastribune.org/2016/12/03/texas-ag-chiefs-facebook-account-fake-news-flows/>

Published in 12/03/2016.

BORDERING ON  INSECURITY A PROJECT BY  THE TEXAS TRIBUNE



Cracks in the Wall: When Border Watchdogs Turn Criminal

The Texas Tribune and Reveal | July 7, 2016 Photo credit: Martin de Noverano

Rudy Soliz did it for sex. **Sergio Lopez Hernandez** blamed depression and financial trouble. **Daphney Caganap** allegedly received cash and a “deluxe” hot tub worth \$10,000.

THE CIRCUMSTANCES OF THESE CRIMINAL ACTS, FROM EMPLOYERS THEY ALLEGEDLY WORKED FOR, ARE COMMON: The corrupt actions they’re accused of weakened the same U.S. borders and ports of entry they were tasked with protecting.

They are just three of the dozens of U.S. Customs and Border Protection (CBP) officials and Border Patrol Agents who have been arrested, charged or convicted of corruption in the past 12 years. A joint investigation by The Texas Tribune and Reveal from the Center for Investigative Reporting has identified at least 140 officials who were arrested or convicted for acts of corruption that allegedly compromised their mission to stop crime and keep the nation secure.

Click on the highlighted names throughout this story to learn more about the individual cases. You can also hover on a name to find that individual on the chart.

Cases in which officials were found not guilty or had their charges dismissed were not included in the Tribune/Reveal review. Neither were examples of routine theft or graft.

In 134 of the 140 corruption cases, officials either pleaded guilty or were convicted at trial. Four have pleaded not guilty and are awaiting trial. The outcome in the two remaining cases is uncertain: In one case, federal authorities would not identify the defendant; in another, they would not provide information on his status.

The cases represent a tiny percentage of the total force of more than 40,000 CBP officers and Border Patrol agents. Agency officials, who declined an interview request, said in an emailed statement that the “overwhelming majority of CBP officers and agents perform their duties with honor and distinction, working tirelessly every day to keep our country safe.”

The Tribune/Reveal joint investigation identified at least 140 officials who were arrested or convicted for acts of corruption that allegedly compromised their mission to stop crime and keep the nation secure.

LENGTH OF PRISON SENTENCE

THE TRIBUNE/REVEAL JOINT INVESTIGATION IDENTIFIED AT LEAST 140 OFFICIALS WHO WERE ARRESTED OR CONVICTED FOR ACTS OF CORRUPTION THAT ALLEGEDLY COMPROMISED THEIR MISSION TO STOP CRIME AND KEEP THE NATION SECURE.

LENGTH OF PRISON SENTENCE

LOCATION OF DUTY STATION

ALLEGED SCHEME*

YEARS OF SERVICE

NONE (32)

LESS THAN 2 YEARS (17)

2 TO 10 YEARS (61)

MORE THAN 10 YEARS (22)

SENTENCE PENDING OR NOT APPLICABLE (8)

*See Methodology

Figure 98 – “Bordering on Insecurity” screenshot.

Title: “Bordering on Insecurity”.

Original url: <https://apps.texastribune.org/bordering-on-insecurity/when-border-watchdogs-turn-criminal/>

Published in 07/07/2016.

Archived url: <http://bit.ly/2LriI RI>

Interaction techniques: Select, filter, inspect.

Topics: Social issues.

Related news story 1: “Border Corruption Often Has Family Roots”.

Url: <https://www.texastribune.org/2016/07/08/family-ties-can-lead-border-corruption/>

Published in 07/08/2016.

Related news story 2: “Agent, Smuggler Joined Together for Love and Money”.

Url: <https://www.texastribune.org/2016/07/06/former-border-patrol-agent-recalls-job-landed-her-/>

Published in 07/06/2016.

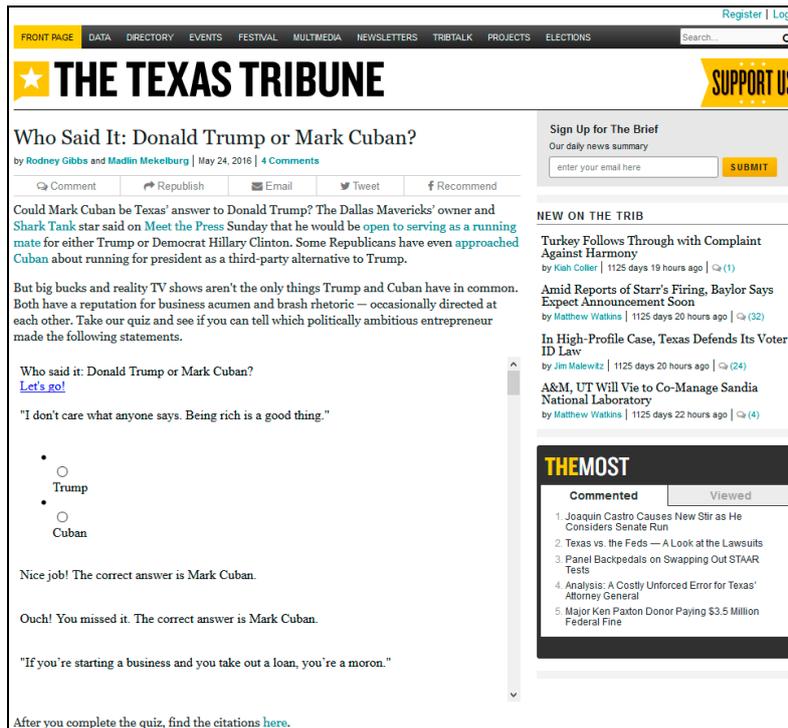


Figure 99 – “Who Said It: Donald Trump or Mark Cuban?” screenshot.

Title: “Who Said It: Donald Trump or Mark Cuban?”.

Original url: <https://www.texastribune.org/2016/05/24/who-said-donald-trump-or-mark-cuban/>

Published in 05/24/2016.

Archived url: <http://bit.ly/2LtPlxR>

Interaction techniques: Gamification, collaboration, narrate/history.

Topics: Entertainment.

Related news story 1: “Analysis: Texas Judge Tweets While Trump Auditions for GOP”.

Url: <https://www.texastribune.org/2016/05/20/analysis-texas-judge-tweets-while-trump-auditions-/>

Published in 05/20/2016.

Related news story 2: “In Texas Senate Races, Hughes and Buckingham Win Republican Runoffs”.

Url: <https://www.texastribune.org/2016/05/24/texas-senate-runoff-results/>

Published in 05/24/2016.

Appendix B: *O Globo* data journalism pieces

B.1. Interactive pieces from 2014

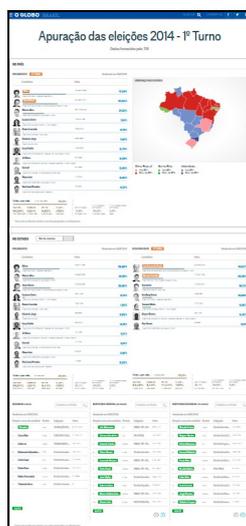


Figure 100 – “Apuração das eleições 2014 – 1º Turno” screenshot.

Title: “Apuração das eleições 2014 – 1º Turno”.

Original url: <https://oglobo.globo.com/brasil/eleicoes-2014/apuracao-votos-1-turno/>

Published in 10/06/2014.

Archived url: <http://bit.ly/2xRv10q>

Interaction techniques: Inspect, filter, explore, select.

Topic: Politics/governance.

Related news story 1: “Dilma ganha em cinco estados a mais que Aécio, mas tucano vence em SP, maior colégio eleitoral”.

Url: <https://oglobo.globo.com/brasil/dilma-ganha-em-cinco-estados-mais-que-aecio-mas-tucano-vence-em-sp-maior-colegio-eleitoral-14157797>

Published in 10/06/2014.

Related news story 2: “Aécio venceu no berço do sindicalismo paulista, enquanto Dilma dominou nas áreas pobres de MG”.

Url: <https://oglobo.globo.com/brasil/aecio-venceu-no-berco-do-sindicalismo-paulista-enquanto-dilma-dominou-nas-areas-pobres-de-mg-14165314>

Published in 10/06/2014.



Figure 101 – “O peso eleitoral dos estados” screenshot.

Title: “O peso eleitoral dos estados”.

Original url: <http://infograficos.oglobo.globo.com/brasil/eleicoes-2014/o-peso-eleitoral-dos-estados.html>

Published in 10/09/2014.

Archived url: <http://bit.ly/2XOwyUD>

Interaction techniques: Select, inspect.

Topic: Politics/governance.

Related news story 1: “Votação em São Paulo, Minas e Pernambuco deve ser decisiva para Dilma e Aécio”.

Url: <https://oglobo.globo.com/brasil/votacao-em-sao-paulo-minas-pernambuco-deve-ser-decisiva-para-dilma-aecio-14165054>

Published in 10/07/2014.

Related news story 2: “Dilma tem mais apoio que Aécio nos estados”.

Url: <https://oglobo.globo.com/brasil/dilma-tem-mais-apoio-que-aecio-nos-estados-14177942>

Published in 10/08/2014.

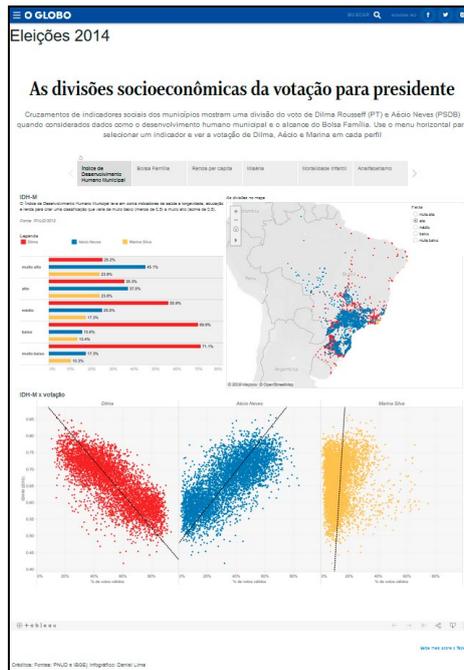


Figure 102 – “As divisões socioeconômicas da votação para presidente” screenshot.

Title: “As divisões socioeconômicas da votação para presidente”.

Original url: <http://infograficos.oglobo.globo.com/brasil/eleicoes-2014/as-divisoes-socioeconomicas-da-votacao-para-presidente.html>

Published in 10/07/2014.

Archived url: <http://bit.ly/2O5HAQE>

Interaction techniques: Inspect, filter, abstract/elaborate, explore, select, connect, collaboration.

Topics: Politics/governance.

Related news story 1: “Dilma venceu em áreas com mais Bolsa Família”.

Url: <https://oglobo.globo.com/brasil/dilma-venceu-em-areas-com-mais-bolsa-familia-14165436>

Published in 10/07/2014.

Related news story 2: “Aécio sai na frente na disputa pelos evangélicos”.

Url: <https://oglobo.globo.com/brasil/aecio-sai-na-frente-na-disputa-pelos-evangelicos-14177477>

Published in 10/08/2014.

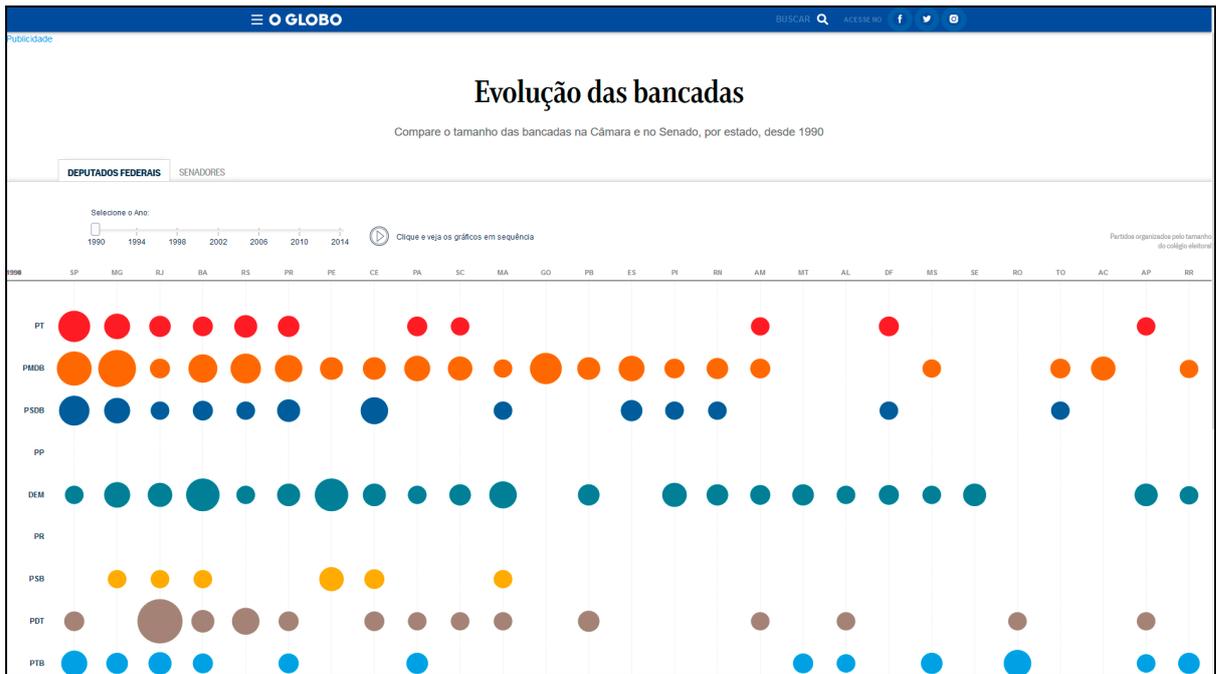


Figure 103 – “Evolução das bancadas” screenshot.

Title: “Evolução das bancadas”.

Original url: <https://infograficos.oglobo.globo.com/brasil/evolucao-das-bancadas-parlamentares.html>

Published in 10/06/2014.

Archived url: <http://bit.ly/2Ls6zvq>

Interaction techniques: Narrate/history, inspect, select.

Topic: Politics/governance.

Related news story 1: “Bancada evangélica cresce 14% e deve prejudicar causas LGBT”.

Url: <https://oglobo.globo.com/brasil/bancada-evangelica-cresce-14-deve-prejudicar-causas-lgbt-14178049>

Published in 10/08/2014.

Related news story 2: “PT e PMDB ainda dominam as cadeiras do Senado”.

Url: <https://oglobo.globo.com/brasil/pt-pmdb-ainda-dominam-as-cadeiras-do-senado-14155231>

Published in 10/06/2014.



Figure 104 – “Os números da Pnad 2013” screenshot.

Title: “Os números da Pnad 2013”.

Original url: <http://infograficos.oglobo.globo.com/economia/os-numeros-da-pnad-2013.html>

Published in 09/18/2014.

Archived url: <http://bit.ly/2LY1Ot2>

Interaction techniques: Inspect, filter, abstract/elaborate, select, connect, collaboration.

Topic: Social issues.

Related news story 1: “Pnad 2013: Desemprego sobe pela primeira vez desde a crise”.

Url: <https://oglobo.globo.com/economia/pnad-2013desemprego-sobe-pela-primeira-vez-desde-crise-13971679>

Published in 09/18/2014.

Related news story 2: “Pnad 2013: Renda cresce, mas desigualdade aumenta; analfabetismo cai”.

Url: <https://oglobo.globo.com/economia/pnad-2013-renda-cresce-mas-desigualdade-aumenta-analfabetismo-cai-13971768>

Published in 09/18/2014.



Figure 105 – “A seleção dos sonhos da Copa do Mundo 2014” screenshot.

Title: “A seleção dos sonhos da Copa do Mundo 2014”.

Original url: <https://infograficos.oglobo.globo.com/esportes/copa2014/copa-do-mundo-montar-selecao.html>

Published in 07/01/2014.

Archived url: <http://bit.ly/30Eukns>

Interaction techniques: Gamification, collaboration, narrate/history.

Topics: Sports, entertainment.

Related news story 1: “Choro divide especialistas sobre controle emocional da seleção”.

Url: <https://oglobo.globo.com/esportes/copa-2014/choro-divide-especialistas-sobre-controle-emocional-da-selecao-13084361>

Published in 07/01/2014.

Related news story 2: “Destaque do jogo, goleiro americano lamenta eliminação”.

Url: <https://oglobo.globo.com/esportes/copa-2014/destaque-do-jogo-goleiro-americano-lamenta-eliminacao-13098577>

Published in 07/01/2014.

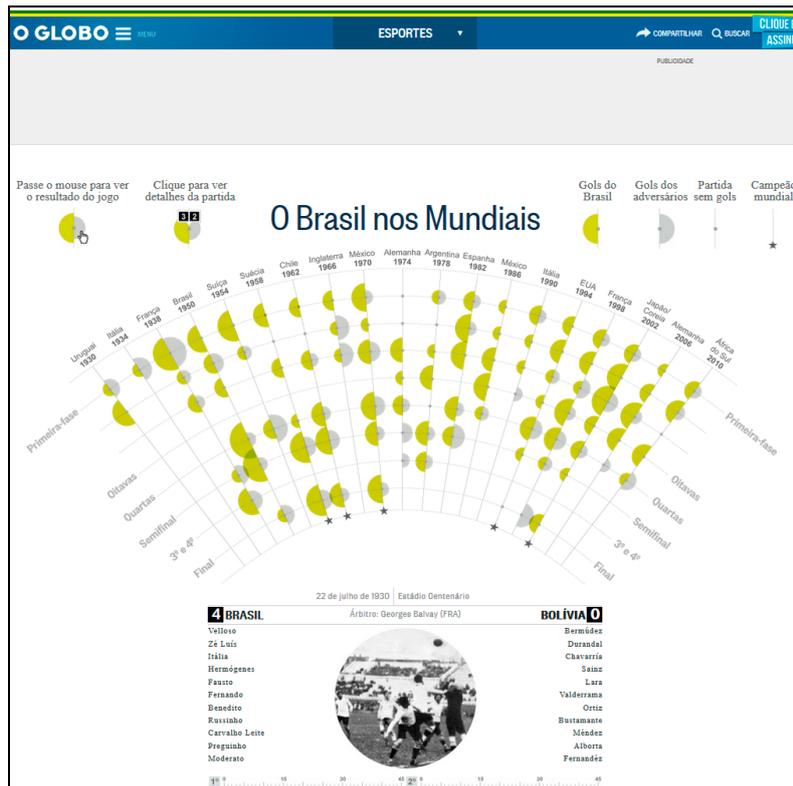


Figure 106 – “O Brasil nos Mundiais” screenshot.

Title: “Brasil nos Mundiais”.

Original url: <https://infograficos.oglobo.globo.com/esportes/brasil-jogo-a-jogo.html>

Published in 05/30/2014.

Archived url: <http://bit.ly/2SnfGhG>

Interaction techniques: Inspect, select, filter.

Topic: Sports.

Related news story 1: “Copa de 1958: Início da era de ouro do futebol brasileiro”.

Url: <https://oglobo.globo.com/esportes/copa-de-1958-inicio-da-era-de-ouro-do-futebol-brasileiro-12655412>

Published in 05/29/2014.

Related news story 2: “Copa de 1970: A Jules Rimet para sempre”.

Url: <https://oglobo.globo.com/esportes/copa-2014/copa-de-1970-jules-rimet-para-sempre-12680301>

Published in 06/01/2014.



Figure 107 – “Você sabe qual a sua expectativa de vida?” screenshot.

Title: “Você sabe qual a sua expectativa de vida?”.

Original url: <http://infograficos.oglobo.globo.com/economia/expectativa-vida.html>

Published in 11/16/2014.

Archived url: <http://bit.ly/2SwXemZ>

Interaction techniques: Gamification, filter.

Topic: Social issues.

Related news story 1: “Implosão demográfica: Brasil mais idoso vai exigir o triplo dos investimentos”.

Url: <https://oglobo.globo.com/economia/implosao-demografica-brasil-mais-idoso-vai-exigir-triplo-dos-investimentos-14574302>

Published in 11/16/2014.

Related news story 2: “‘Nasce cada vez menos e morre cada vez menos. E aí, vamos ser imortais?’, indaga pesquisadora”.

Url: <https://oglobo.globo.com/economia/nasce-cada-vez-menos-morre-cada-vez-menos-ai-vamos-ser-imortais-indaga-pesquisadora-14569682>

Published in 11/16/2014.



Figure 108 – “Mapa da taxa de suicídio no mundo” screenshot.

Title: “Mapa da taxa de suicídio no mundo”.

Original url: <http://infograficos.oglobo.globo.com/sociedade/mapa-da-taxa-de-suicidio-no-mundo.html>

Published in 09/04/2014.

Archived url: <http://bit.ly/2Y9zE4W>

Interaction techniques: Inspect, select, filter, connect, abstract/elaborate, collaboration.

Topics: Social issues, foreign affairs.

Related news story 1: “Uma pessoa comete suicídio a cada 40 segundos no mundo, diz OMS”.

Url: <https://oglobo.globo.com/sociedade/saude/uma-pessoa-comete-suicidio-cada-40-segundos-no-mundo-diz-oms-13826787>

Published in 09/04/2014.

Related news story 2: “Tabu é entrave a políticas públicas contra suicídio”.

Url: <https://oglobo.globo.com/sociedade/saude/tabu-entrave-politicas-publicas-contra-suicidio-13836147>

Published in 09/05/2014.

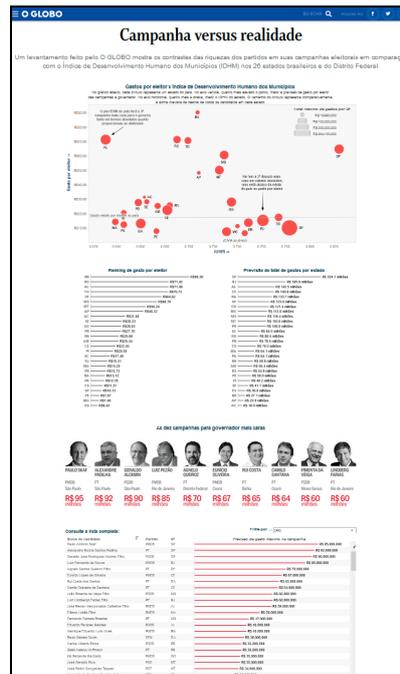


Figure 109 – “Campanha versus realidade” screenshot.

Title: “Campanha versus realidade”.

Original url: <http://infograficos.oglobo.globo.com/brasil/o-custo-das-campanhas-nos-estados.html>

Published in 07/16/2014.

Archived url: <http://bit.ly/30ILMqT>

Interaction techniques: Inspect, select, filter, reconfigure/encode..

Topics: Politics/governance, economy/business.

Related news story 1: “Com pior IDH do país, população de Alagoas pede mais investimento em saúde e educação”.

Url: <https://oglobo.globo.com/brasil/com-pior-idh-do-pais-populacao-de-alagoas-pede-mais-investimento-em-saude-educacao-13271607>

Published in 07/16/2014.

Related news story 2: “Presidenciais divulgam jingles para as eleições de outubro”.

Url: <https://oglobo.globo.com/brasil/presidenciais-divulgam-jingles-para-as-eleicoes-de-outubro-13279309>

Published in 07/16/2014.

B.2. Interactive pieces from 2015

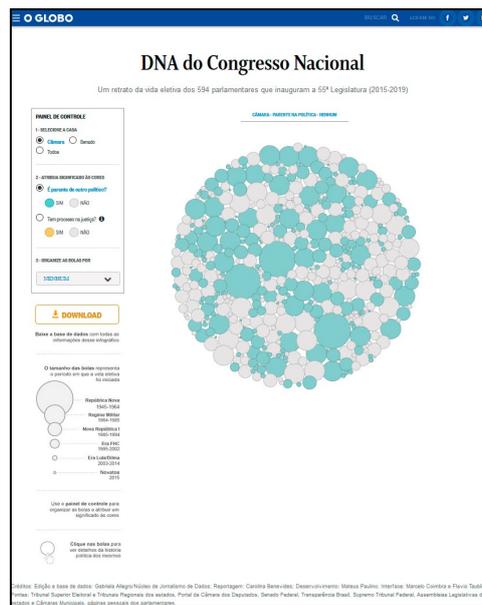


Figure 110 – “DNA do Congresso Nacional” screenshot.

Title: “DNA do Congresso Nacional”.

Original url: <http://infograficos.oglobo.globo.com/infograficos/dna-do-congresso.html>

Published in 02/01/2015.

Archived url: <http://bit.ly/2YXlj7C>

Interaction techniques: Inspect, select, reconfigure/encode, filter, collaboration.

Topic: Politics/governance.

Related news story 1: “Eleito para o quarto mandato, Renan Calheiros é visto como uma fênix”.

Url: <https://oglobo.globo.com/brasil/dilma-ganha-em-cinco-estados-mais-que-aecio-mas-tucano-vence-em-sp-maior-colegio-eleitoral-14157797>

Published in 02/01/2015.

Related news story 2: “Eduardo Cunha derrota o governo e é eleito presidente da Câmara”.

Url: <https://oglobo.globo.com/brasil/aecio-venceu-no-berco-do-sindicalismo-paulista-enquanto-dilma-dominou-nas-areas-pobres-de-mg-14165314>

Published in 02/01/2015.

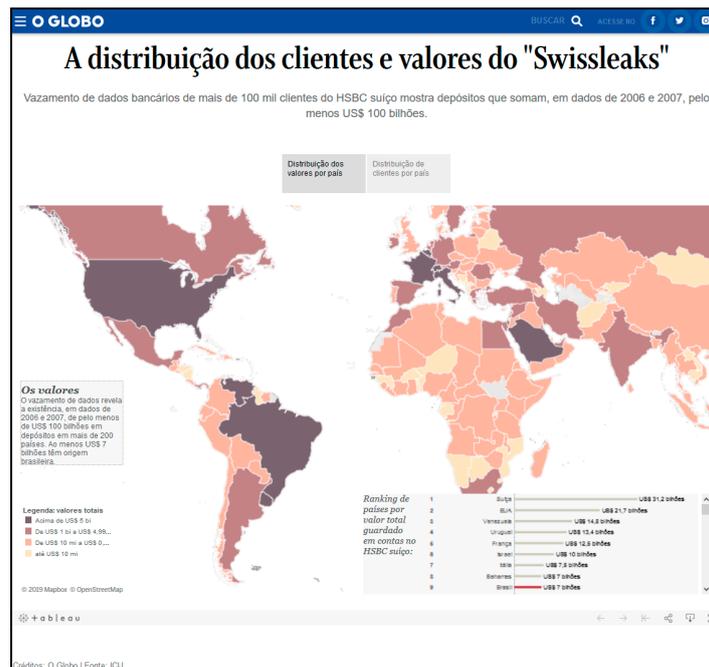


Figure 111 – “A distribuição dos clientes e valores do ‘Swissleaks’” screenshot.

Title: “A distribuição dos clientes e valores do ‘Swissleaks’”.

Original url: <http://infograficos.oglobo.globo.com/brasil/a-distribuicao-dos-clientes-e-valores-do-swissleaks.html>

Published in 03/12/2015.

Archived url: <http://bit.ly/2LUWPJH>

Interaction techniques: Inspect, select, abstract/elaborate, connect, collaboration.

Topics: Economy/business, foreign affairs.

Related news story 1: “SwissLeaks: lista de brasileiros com contas secretas tem vários acusados de fraudes”.

Url: <https://oglobo.globo.com/brasil/swissleaks-lista-de-brasileiros-com-contas-secretas-tem-varios-acusados-de-fraudes-15565841>

Published in 03/12/2015.

Related news story 2: “Brasileiros tinham ao menos US\$ 7 bilhões na Suíça”.

Url: <https://oglobo.globo.com/brasil/brasileiros-tinham-ao-menos-us-7-bilhoes-na-suica-15571612>

Published in 03/12/2015.



Figure 112 – “Ataques terroristas no mundo desde 1970” screenshot.

Title: “Ataques terroristas no mundo desde 1970”.

Original url: <http://infograficos.oglobo.globo.com/mundo/ataques-terroristas-no-mundo-desde-1970.html>

Published in 11/19/2015.

Archived url: <http://bit.ly/2LXSepV>

Interaction techniques: Inspect, select, abstract/elaborate, connect, explore, collaboration.

Topics: Social issues, foreign affairs.

Related news story 1: “A participação das mulheres em grupos terroristas”.

Url: <https://oglobo.globo.com/mundo/a-participacao-das-mulheres-em-grupos-terroristas-18093310>

Published in 11/19/2015.

Related news story 2: “Itália eleva nível de segurança após alerta do FBI sobre ataque terrorista”.

Url: <https://oglobo.globo.com/mundo/italia-eleva-nivel-de-seguranca-apos-alerta-do-fbi-sobre-ataque-terrorista-18095266>

Published in 11/19/2015.

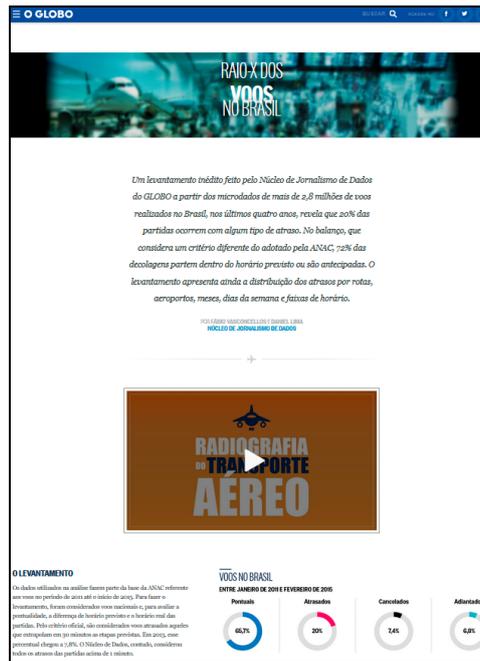


Figura 113 – “Raio-x dos atrasos dos voos no Brasil” screenshot.

Title: “Raio-x dos atrasos dos voos no Brasil”.

Original url: <http://infograficos.oglobo.globo.com/economia/raio-x-dos-atrasos-dos-voos.html>

Published in 07/19/2015.

Archived url: <http://bit.ly/2YgdkGy>

Interaction techniques: Inspect, select, abstract/elaborate, connect, explore, collaboration.

Topic: Economy/business.

Related news story 1: “Mapa interativo mostra os voos que mais atrasam”.

Url: <https://oglobo.globo.com/economia/mapa-interativo-mostra-os-voos-que-mais-atrasam-16835185>

Published in 07/19/2015.

Related news story 2: “Passageiro de negócios some, e aéreas recorrem a promoções”.

Url: <https://oglobo.globo.com/economia/passageiro-de-negocios-some-aereas-recorrem-promocoes-16835005>

Published in 07/19/2015.



Figura 114 – “Qual o futuro do PT?” screenshot.

Title: “Qual o futuro do PT?”.

Original url: <http://infograficos.oglobo.globo.com/brasil/futuro-pt.html>

Published in 09/13/2015.

Archived url: <http://bit.ly/30IMWCL>

Interaction techniques: Select, inspect.

Topic: Politics/governance.

Related news story 1: “Parlamentares do PT já estão de malas prontas para deixar o partido”.

Url: <https://oglobo.globo.com/brasil/parlamentares-do-pt-ja-estao-de-malas-prontas-para-deixar-partido-17477487>

Published in 09/13/2015.

Related news story 2: “Para ex-ministro Moreira Franco, está cada vez mais difícil Dilma terminar o mandato”.

Url: <https://oglobo.globo.com/brasil/para-ex-ministro-moreira-franco-esta-cada-vez-mais-dificil-dilma-terminar-mandato-17477655>

Published in 09/13/2015.

RANKING	ESCOLA	MUNICÍPIO	MÉDIA Prova objetiva	MÉDIA Redação
1	OBJETIVO COLEGIO INTEGRADO	SAO PAULO	743	774,6
2	FARIAS BRITO COLEGIO DE APLICACAO	FORTALEZA	737,9	786,4
3	COLEGIO OLIMPO INTESIRAL	GOIANA	735	837,6
4	CRISTUS COLEGIO PRE-UNIVERSITARIO	FORTALEZA	731,4	836,6
5	COLEGIO BERNHULLI - UNIDADE LOURDES	BELO HORIZONTE	730,3	799,4
6	ARZ DE SA CAVALCANTE COLEGIO - SAUJUR FACUNDO	FORTALEZA	725,1	888,3
7	COLEGIO E CURSO PONTO DE ENENHO	RIO DE JANEIRO	720,7	716,8
8	COLEGIO ELITE VALE DO ADO	BRITANIA	718,8	778,3
9	COLEGIUM	BELO HORIZONTE	718,7	807,8
10	OBJETIVO INTEGRADO DE NOSSAS CRUIZES COLEGIO	MOGI DAS CRUIZES	718,7	797,7
11	COLEGIO BONATUS I	CAMPUS GRAJAE	712,9	870,4
12	COLEGIO SANTO ANTONIO	BELO HORIZONTE	712	780,9
13	COLEGIO E CURSO PONTO DE ENENHO	MITEROI	711,4	788
14	VERTICE COLEGIO UNIDADE I	SAO PAULO	711	747,1
15	COLEGIO E CURSO PONTO DE ENENHO	RIO DE JANEIRO	709,4	788,1
16	INSTIT DOM BARRETO	TERESINA	705,3	775,2
17	COL DE SAO BENTO	RIO DE JANEIRO	704,9	886,6
18	SEB COC UNIDADE ALVARES CABRAL	IBERAPO NETO	701,7	776,5
19	ANGELO LEONARDO DA VINHO COLEGIO	CABARCUBA	701,1	734,4
20	COLEGIO LEROTTE LTDA	TERESINA	700,9	844,6

Figura 115 – “Consulte a media da sua escola no Enem 2014” screenshot.

Title: “Consulte a media da sua escola no Enem 2014”.

Original url: <https://infograficos.oglobo.globo.com/sociedade/educacao/consulte-a-media-da-sua-escola-no-enem-2014.html>

Published in 08/05/2015.

Archived url: <http://bit.ly/2Smrkt2>

Interaction techniques: Filter, explore, reconfigure/encode.

Topic: Education.

Related news story 1: “Colégio paulista lidera ranking do Enem 2014, mas Nordeste e Centro-Oeste ganham posições”.

Url: <https://oglobo.globo.com/sociedade/educacao/enem-e-vestibular/colégio-paulista-lidera-ranking-do-enem-2014-mas-nordeste-centro-oeste-ganham-posicoes-17087541>

Published in 08/05/2015.

Related news story 2: “Em cinco das dez melhores escolas no Enem 2014, 80% dos alunos ou mais saíram de outros colégios”.

Url: <https://oglobo.globo.com/sociedade/educacao/enem-e-vestibular/em-cinco-das-dez-melhores-escolas-no-enem-2014-80-dos-alunos-ou-mais-sairam-de-outros-colegios-17088237>

Published in 08/05/2015.

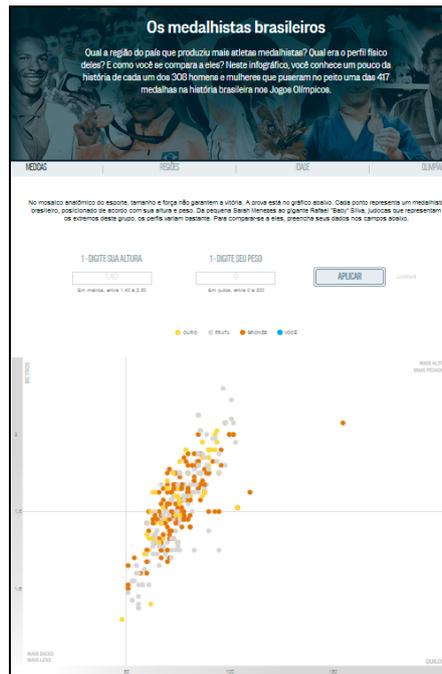


Figura 116 – “Os medalhistas brasileiros” screenshot.

Title: “Os medalhistas brasileiros”.

Original url: <http://infograficos.oglobo.globo.com/esportes/os-medalhistas-brasileiros.html>

Published in 08/05/2015.

Archived url: <http://bit.ly/2JNRMbd>

Interaction techniques: Gamification, inspect, select, filter.

Topic: Sports.

Related news story 1: “Medalhistas olímpicos do passado escolhem favoritos do Brasil ao pódio na Rio-2016”.

Url: <https://oglobo.globo.com/esportes/medalhistas-olimpicos-do-passado-escolhem-favoritos-do-brasil-ao-podio-na-rio-2016-17083454>

Published in 08/05/2015.

Related news story 2: “Cinco grandes obstáculos a um ano dos Jogos Olímpicos do Rio, em 2016”.

Url: <https://oglobo.globo.com/esportes/cinco-grandes-obstaculos-um-ano-dos-jogos-olimpicos-do-rio-em-2016-17082852>

Published in 08/05/2015.

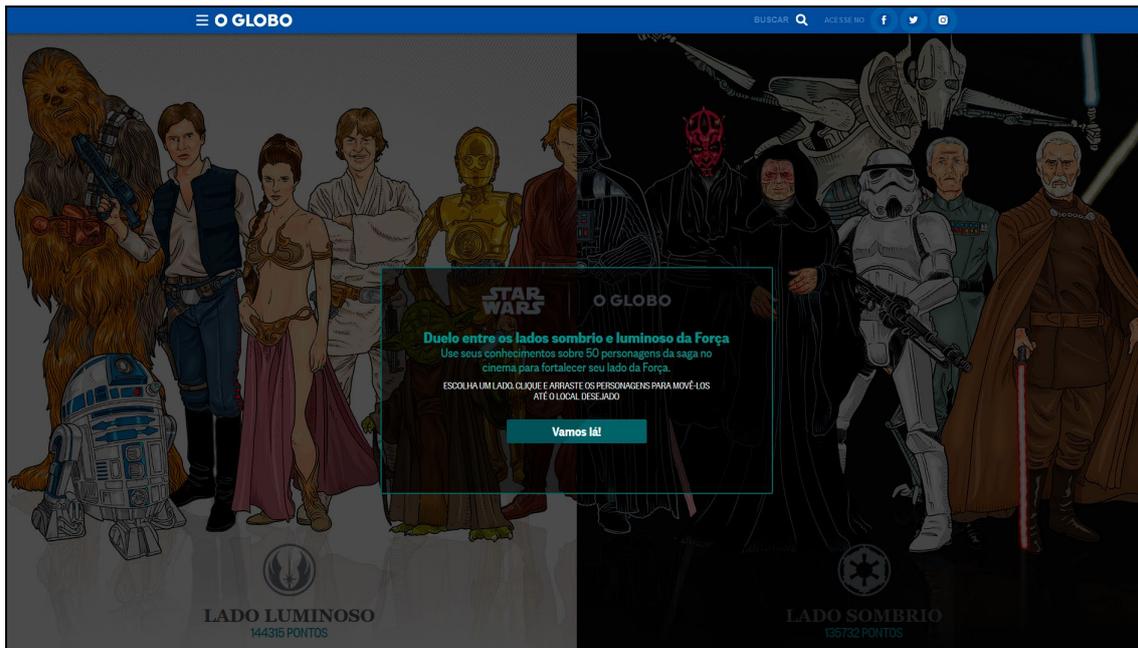


Figura 117 – “Duelo entre os lados sombrio e luminoso da Força” screenshot.

Title: “Duelo entre os lados sombrio e luminoso da Força”.

Original url: <http://infograficos.oglobo.globo.com/cultura/jogo-star-wars.html>

Published in 12/15/2015.

Archived url: <http://bit.ly/30FV2we>

Interaction techniques: Gamification, collaboration.

Topics: Entertainment, lifestyle/culture.

Related news story 1: “Uma nova esperança: 'Episódio VII' de 'Star Wars' agrada fãs em première”.

Url: <https://oglobo.globo.com/cultura/filmes/uma-nova-esperanca-episodio-vii-de-star-wars-agrada-fas-em-premiere-18298808>

Published in 12/15/2015.

Related news story 2: “Disney busca público feminino para 'O despertar da Força’”.

Url: <https://oglobo.globo.com/cultura/filmes/disney-busca-publico-feminino-para-despertar-da-forca-18299407>

Published in 12/15/2015.

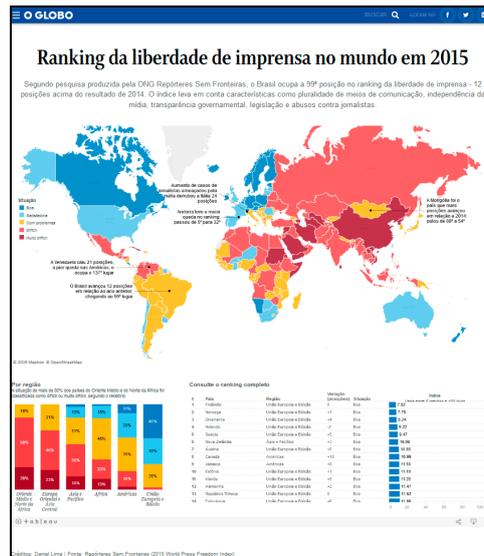


Figura 118 – “Ranking da liberdade de imprensa no mundo em 2015” screenshot.

Title: “Ranking da liberdade de imprensa no mundo em 2015”.

Original url: <http://infograficos.oglobo.globo.com/sociedade/ranking-da-liberdade-de-imprensa-no-mundo-em-2015.html>

Published in 02/12/2015.

Archived url: <http://bit.ly/30COJcH>

Interaction techniques: Inspect, select, abstract/elaborate, connect, collaboration, reconfigure/encode.

Topics: Social issues, foreign affairs.

Related news story 1: “Maioria dos países perde posições em ranking de liberdade de imprensa”.

Url: <https://oglobo.globo.com/sociedade/midia/maioria-dos-paises-perde-posicoes-em-ranking-de-liberdade-de-imprensa-15317201>

Published in 02/12/2015.

Related news story 2: “Brasil sobe 12 posições em ranking de liberdade de imprensa, mas enfrenta desafios”.

Url: <https://oglobo.globo.com/sociedade/midia/brasil-sobe-12-posicoes-em-ranking-de-liberdade-de-imprensa-mas-enfrenta-desafios-15324556>

Published in 02/13/2015.

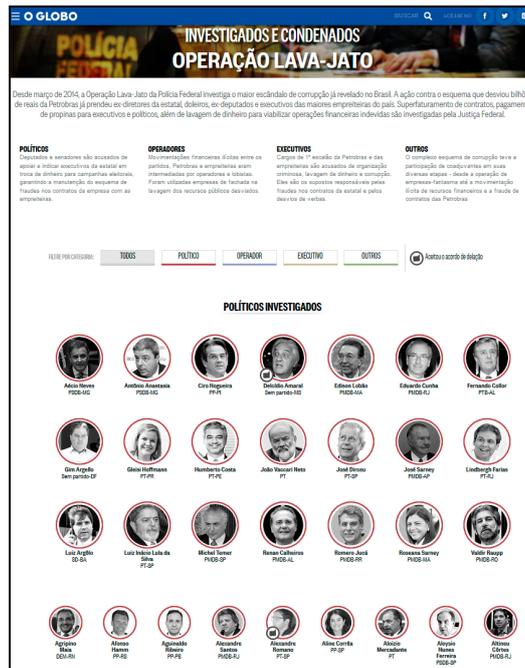


Figura 119 – “Personagens da Lava-Jato” screenshot.

Title: “Personagens da Lava-Jato”.

Original url: <http://infograficos.oglobo.globo.com/brasil/lava-jato-personagens.html>

Published in 08/22/2015.

Archived url: <http://bit.ly/2Yjbsx6>

Interaction techniques: Inspect, filter.

Topic: Politics/governance.

Related news story 1: “Vice-presidente diz que apoia investigações ‘sérias’ da Lava-Jato”.

Url: <https://oglobo.globo.com/brasil/vice-presidente-diz-que-apoia-investigacoes-serias-da-lava-jato-17274259>

Published in 08/22/2015.

Related news story 2: “Senadores apostam na recondução de Janot, mas sabatina ainda é uma incógnita”.

Url: <https://oglobo.globo.com/brasil/senadores-apostam-na-reconducao-de-janot-mas-sabatina-ainda-uma-incognita-17271700>

Published in 08/22/2015.

B.3. Interactive pieces from 2016

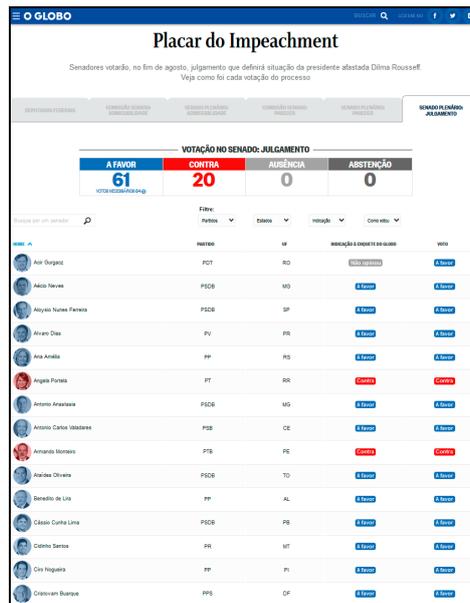


Figure 120 – “Placar do Impeachment” screenshot.

Title: “Placar do Impeachment”.

Original url: <http://infograficos.oglobo.globo.com/brasil/o-placar-do-impeachment.html>

Published in 04/13/2016.

Archived url: <http://bit.ly/2GiNzvn>

Interaction techniques: Filter, select, explore.

Topic: Politics/governance.

Related news story 1: “Oposição diz ter 349 votos a favor do impeachment, sete além do necessário”.

Url: <https://oglobo.globo.com/brasil/oposicao-diz-ter-349-votos-favor-do-impeachment-sete-alem-do-necessario-19079261>

Published in 04/13/2016.

Related news story 2: “Manifestantes já acampam no DF à espera da votação do impeachment”.

Url: <https://oglobo.globo.com/brasil/manifestantes-ja-acampam-no-df-espera-da-votacao-do-impeachment-19072625>

Published in 04/13/2016.



Figure 121 – “O mapa das coligações” screenshot.

Title: “O mapa das coligações”.

Original url: <http://infograficos.oglobo.globo.com/brasil/grafico-coligacoes.html>

Published in 09/18/2016.

Archived url: <http://bit.ly/2JDO2Kr>

Interaction techniques: Inspect, select.

Topic: Politics/governance.

Related news story 1: “Nas alianças para prefeito, a lógica do vale tudo”.

Url: <https://oglobo.globo.com/brasil/nas-aliancas-para-prefeito-logica-do-vale-tudo-20133037>

Published in 09/18/2016.

Related news story 2: “Sem empresas, Fundo Partidário é fonte de verba que mais cresceu”.

Url: <https://oglobo.globo.com/brasil/sem-empresas-fundo-partidario-fonte-de-verba-que-mais-cresceu-20132514>

Published in 09/18/2016.

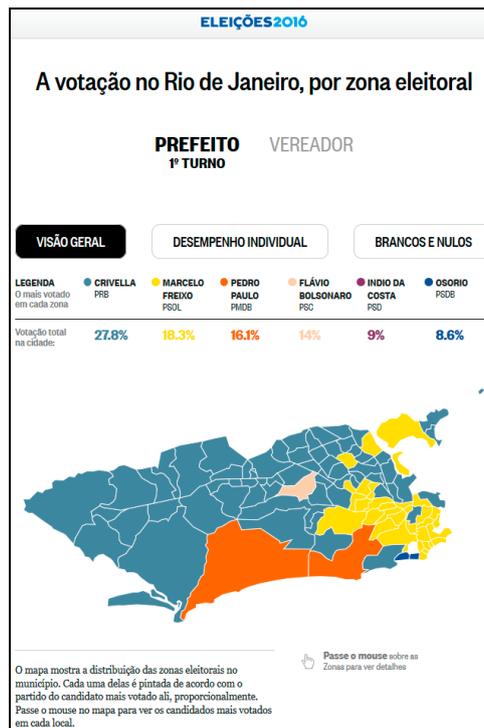


Figure 122 – “A votação no Rio de Janeiro, por zona eleitoral” screenshot.

Title: “A votação no Rio de Janeiro, por zona eleitoral”.

Original url: <http://infograficos.oglobo.globo.com/brasil/a-votacao-no-rio-de-janeiro-por-zona-eleitoral.html>

Published in 10/03/2016.

Archived url: <http://bit.ly/2O2OGpi>

Interaction techniques: Inspect, select, filter.

Topic: Politics/governance.

Related news story 1: “Renovação de vereadores na Câmara do Rio cai para 35%”.

Url: <https://oglobo.globo.com/brasil/renovacao-de-vereadores-na-camara-do-rio-cai-para-35-20222458>

Published in 10/03/2016.

Related news story 2: “No Rio, esquerda foi mais eficiente no voto útil”.

Url: <https://oglobo.globo.com/brasil/no-rio-esquerda-foi-mais-eficiente-no-voto-util-20222232>

Published in 10/03/2016.

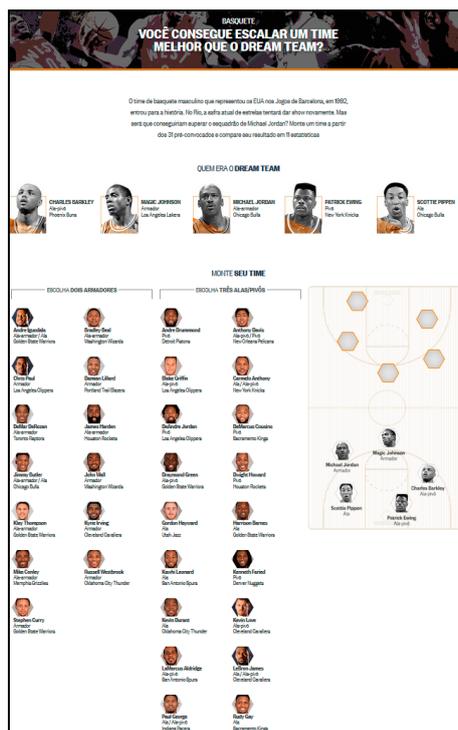


Figure 123 – “Você consegue escalar um time melhor que o Dream Team?” screenshot.

Title: “Você consegue escalar um time melhor que o Dream Team?”.

Original url: <http://infograficos.oglobo.globo.com/esportes/rio-2016/voce-consegue-escalar-um-time-melhor-que-o-dream-team.html>

Published in 02/26/2016.

Archived url: <http://bit.ly/2XQxvf7>

Interaction techniques: Gamification, collaboration.

Topics: Sports, entertainment.

Related news story 1: “Basquete: Flamengo vence Brasília na Liga das Américas”.

Url: <https://oglobo.globo.com/esportes/basquete-flamengo-vence-brasilia-na-liga-das-americas-18763175>

Published in 02/26/2016.

Related news story 2: “Anderson Varejão vai jogar no Golden State Warriors”.

Url: <https://oglobo.globo.com/esportes/anderson-varejao-vai-jogar-no-golden-state-warriors-18721128>

Published in 02/22/2016.



Figure 124 – “Partido do ‘você não me representa’” screenshot.

Title: “Partido do ‘você não me representa’”.

Original url: <http://infograficos.oglobo.globo.com/brasil/partido-do-voce-nao-me-representa.html>

Published in 03/22/2016.

Archived url: <http://bit.ly/2O0z8SM>

Interaction techniques: Gamification, collaboration, narrate/history.

Topics: Politics/governance, social issues.

Related news story 1: “Prefeito de Niterói deixa o PT e se filia ao PV”.

Url: <https://oglobo.globo.com/rio/bairros/prefeito-de-niteroi-deixa-pt-se-filia-ao-pv-18934068>

Published in 03/22/2016.

Related news story 2: “Renan diz que impeachment sem provas tem ‘outro nome’”.

Url: <https://oglobo.globo.com/brasil/renan-diz-que-impeachment-sem-provas-tem-outro-nome-18935008>

Published in 03/22/2016.



Figure 125 – “Calculadora de aposentadoria” screenshot.

Title: “Calculadora de aposentadoria”.

Original url: <http://infograficos.oglobo.globo.com/economia/calculadora-aposentadoria.html>

Published in 08/27/2016.

Archived url: <http://bit.ly/2JStJaX>

Interaction techniques: Gamification, collaboration.

Topic: Social issues.

Related news story 1: “Governo estuda regime especial para aposentado que trabalha”.

Url: <https://oglobo.globo.com/economia/governo-estuda-regime-especial-para-aposentado-que-trabalha-20006435>

Published in 08/28/2016.

Related news story 2: “Regras mais iguais para homens e mulheres, rotinas ainda diferentes”.

Url: <https://oglobo.globo.com/economia/regras-mais-iguais-para-homens-mulheres-rotinas-ainda-diferentes-20006533>

Published in 08/28/2016.

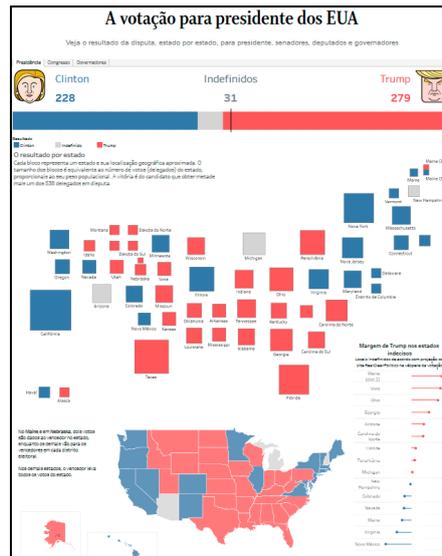


Figure 126 – “A votação para presidente dos EUA” screenshot.

Title: “A votação para presidente dos EUA”.

Original url: <http://infograficos.oglobo.globo.com/mundo/eleicoes-eua-2016-presidente-hillary-clinton-donald-trump.html>

Published in 11/09/2016.

Archived url: <http://bit.ly/2JFaVgJ>

Interaction techniques: Inspect, explore, select, connect, abstract/elaborate, collaboration.

Topics: Foreign affairs, politics/governance.

Related news story 1: “Trump é eleito presidente dos EUA: ‘Vamos renovar o sonho americano’”.

Url: <https://oglobo.globo.com/mundo/trump-eleito-presidente-dos-eua-vamos-renovar-sonho-americano-20435804>

Published in 11/09/2016.

Related news story 2: “Hillary faz discurso de derrota: ‘EUA mais divididos do que pensávamos’”.

Url: <https://oglobo.globo.com/mundo/hillary-faz-discurso-de-derrota-eua-mais-divididos-do-que-pensavamos-20438149>

Published in 11/09/2016.



Figure 128 – “O trajeto da tocha olímpica no Brasil” screenshot.

Title: “O trajeto da tocha olímpica no Brasil”.

Original url: <https://infograficos.oglobo.globo.com/esportes/rio-2016/o-trajeto-da-tocha-olimpica-no-brasil.html>

Published in 02/24/2016.

Archived url: <http://bit.ly/2JFZwgJ>

Interaction techniques: Inspect, select, abstract/elaborate, explore, collaboration.

Topic: Sports.

Related news story 1: “Clubes vão poder entrar em licitação para administrar o Maracanã”.

Url: <https://oglobo.globo.com/esportes/clubes-va-o-poder-entrar-em-licitacao-para-administrar-maracana-18739715>

Published in 02/24/2016.

Related news story 2: “Rio-2016 revela cidades no revezamento da tocha olímpica”.

Url: <https://oglobo.globo.com/esportes/rio-2016-revela-cidades-no-revezamento-da-tocha-olimpica-18737625>

Published in 02/24/2016.

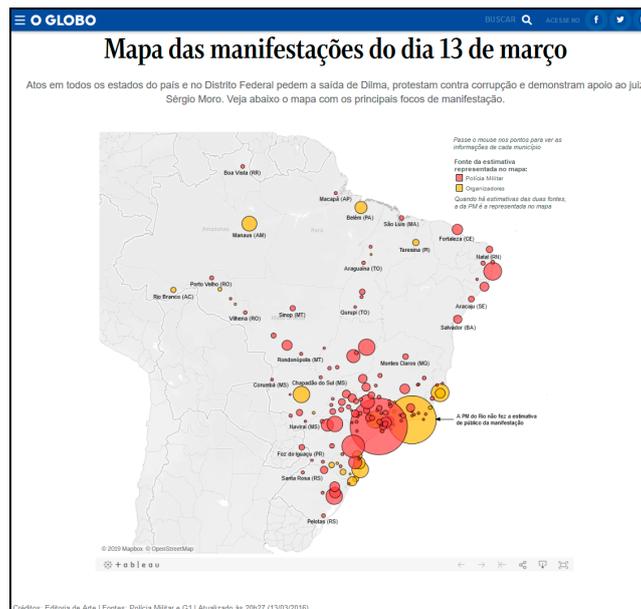


Figure 129 – “Mapa das manifestações do dia 13 de março” screenshot.

Title: “Mapa das manifestações do dia 13 de março”.

Original url: <http://infograficos.oglobo.globo.com/brasil/mapa-das-manifestacoes-do-dia-13-de-marco.html>

Published in 03/13/2016.

Archived url: <http://bit.ly/30Galol>

Interaction techniques: Inspect, explore, connect, abstract/elaborate, collaboration.

Topics: Social issues, politics/governance.

Related news story 1: “Protestos contra Dilma reúnem 3,6 milhões em todos os estados”.

Url: <https://oglobo.globo.com/brasil/protestos-contradilma-reunem-36-milhoes-em-todos-os-estados-18865889>

Published in 03/13/2016.

Related news story 2: “Presentes em manifestações de 2013, jovens contam por que hoje estão em lados diferentes”.

Url: <https://oglobo.globo.com/brasil/presentes-em-manifestacoes-de-2013-jovens-contam-por-que-hoje-estao-em-lados-diferentes-18862146>

Published in 03/13/2016.

Appendix C: *Público* data journalism pieces

C.1. Interactive pieces from 2014

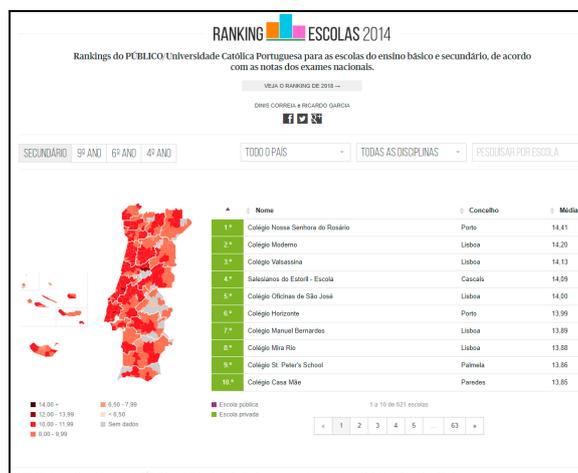


Figure 130 – “O ranking das escolas 2014” screenshot.

Title: “O ranking das escolas 2014”.

Original url: <https://acervo.publico.pt/ranking-das-escolas-2014/listas>

Published in 11/29/2014.

Archived url: <http://bit.ly/2SkZw8s>

Interaction techniques: Inspect, filter, explore, select, reconfigure/encode, collaboration, narrate/history.

Topic: Education.

Related news story 1: “Uma em cada dez escolas do privado teve 4 nos exames nacionais”.

Url: <https://www.publico.pt/2014/11/29/sociedade/noticia/4-ano-uma-em-cada-dez-escolas-do-privado-teve-4-nos-exames-nacionais-1677751>

Published in 11/29/2014.

Related news story 2: “6.º ano: Só 1% das públicas consegue melhores resultados nos exames do que na escola”.

Url: <https://www.publico.pt/2014/11/29/sociedade/noticia/6-ano-so-1-das-publicas-consegue-melhores-resultados-nos-exames-do-que-na-escola-1677755>

Published in 11/29/2014.

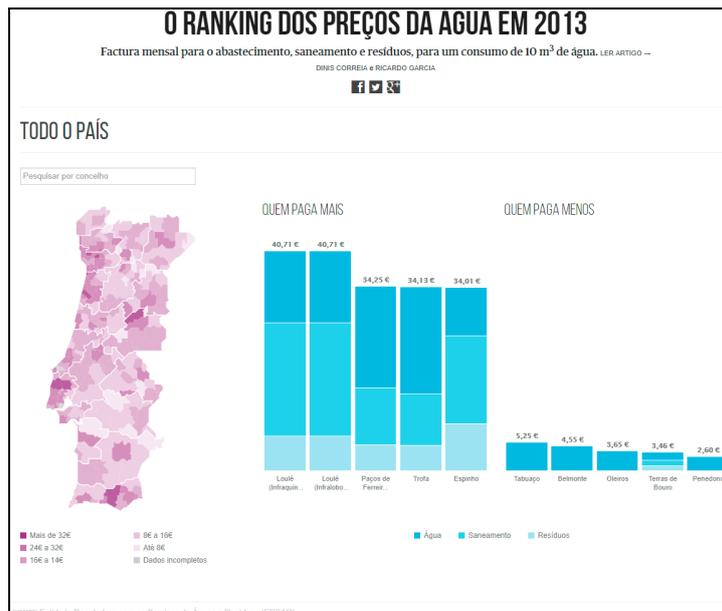


Figure 131 – “O ranking dos preços da água em 2013” screenshot.

Title: “O ranking dos preços da água em 2013”.

Original url: <https://acervo.publico.pt/multimedia/infografia/ranking-dos-precos-da-agua-em-2013>

Published in 11/22/2014.

Archived url: <http://bit.ly/30IJ70K>

Interaction techniques: Inspect, filter, explore, select, narrate/history.

Topics: Economy/business, social issues.

Related news story 1: “Metade dos municípios aumentou a factura da água para as famílias em 2013”.

Url: <https://www.publico.pt/2014/11/22/sociedade/noticia/metade-dos-municipios-aumentou-a-factura-da-agua-para-as-familias-em-2013-1677031>

Published in 11/22/2014.

Related news story 2: “Ministro do Ambiente acredita que “culpa não morrerá solteira” no surto de Legionella”.

Url: <https://www.publico.pt/2014/11/18/sociedade/noticia/ministro-do-ambiente-acredita-que-culpa-nao-morrera-solteira-no-surto-de-legionella-1676661>

Published in 11/18/2014.

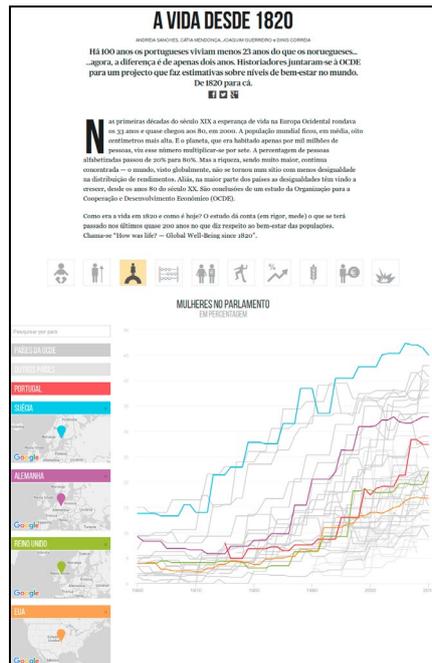


Figure 132 – “A vida desde 1820” screenshot.

Title: “A vida desde 1820”.

Original url: <https://acervo.publico.pt/multimedia/infografia/a-vida-desde-1820>

Published in 11/03/2014.

Archived url: <http://bit.ly/2GllkYa>

Interaction techniques: Select, explore, filter.

Topic: Social issues.

Related news story 1: “Há um século os portugueses viviam menos 23 anos do que os noruegueses”.

Url: <https://www.publico.pt/2014/11/03/sociedade/noticia/ha-um-seculo-os-portugueses-viviam-menos-23-anos-do-que-os-noruegueses-1674666>

Published in 11/03/2014.

Related news story 2: “Mais de 70% dos jovens portugueses com sinais de dependência da Internet”.

Url: <https://www.publico.pt/2014/11/03/sociedade/noticia/quase-tres-quartos-dos-jovens-portugueses-apresentam-sinais-de-dependencia-da-internet-1674907>

Published in 11/03/2014.

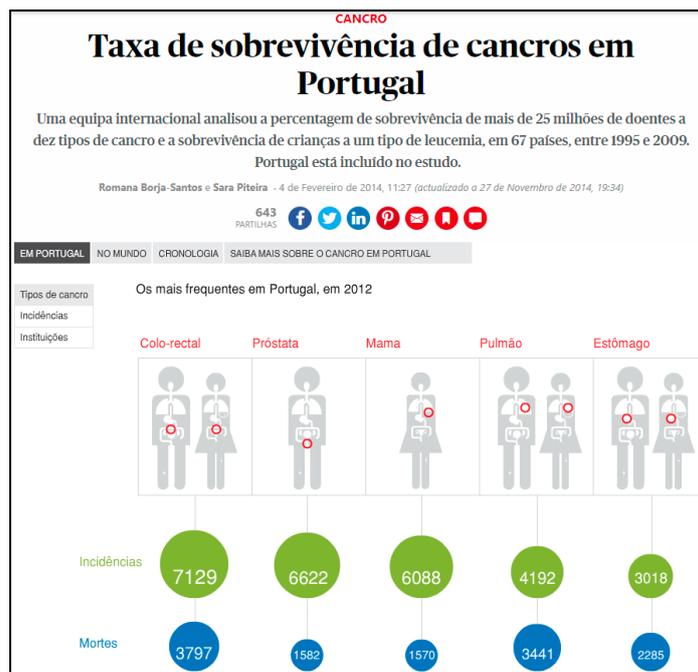


Figure 133 – “Taxa de sobrevivência de cancros em Portugal” screenshot.

Title: “Taxa de sobrevivência de cancros em Portugal”.

Original url: <https://www.publico.pt/2014/02/04/infografia/cancro-em-portugal-114>

Published in 02/04/2014.

Archived url: <http://bit.ly/2O2RyCw>

Interaction techniques: Narrate/history, select, filter, inspect.

Topic: Health.

Related news story 1: “Vacinação contra vírus do cancro do colo do útero fez desaparecer lesões pré-cancerosas”.

Url: <https://www.publico.pt/2014/02/04/ciencia/noticia/vacinacao-contra-virus-do-cancro-do-colo-do-utero-fez-desaparecer-lesoes-precancerosas-1622182>

Published in 02/04/2014.

Related news story 2: “Percentagem de portuguesas fumadoras quase duplicou em cinco anos”.

Url: <https://www.publico.pt/2014/02/06/sociedade/noticia/percentagem-de-portuguesas-fumadoras-quase-duplicou-em-cinco-anos-1622589>

Published in 02/06/2014.

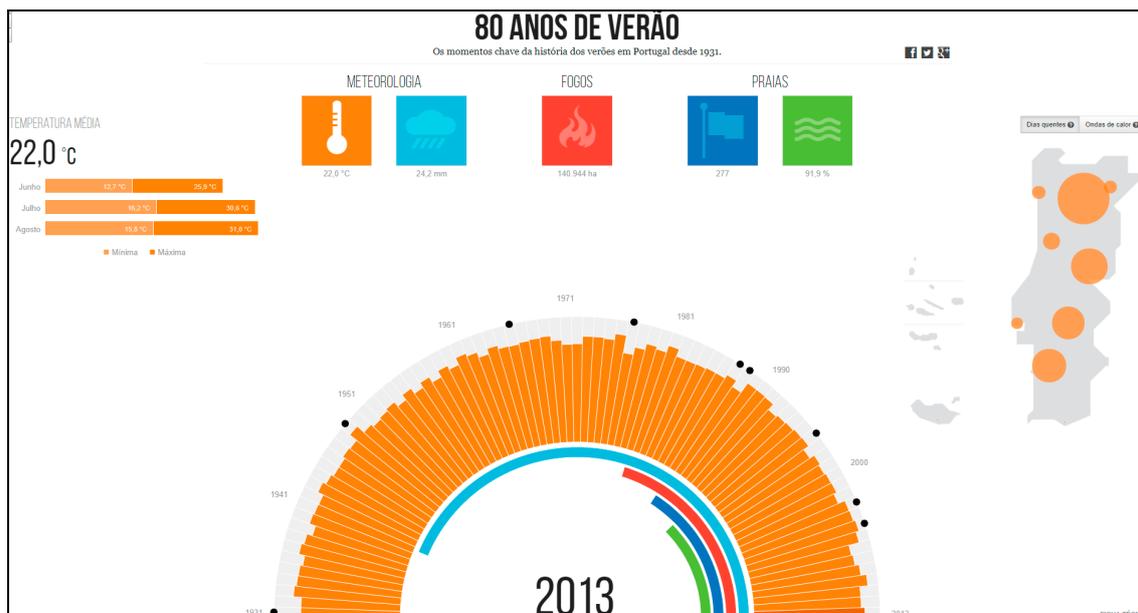


Figure 134 – “80 anos de Verão” screenshot.

Title: “80 anos de Verão”.

Original url: <https://acervo.publico.pt/multimedia/infografia/80-anos-de-verao>

Published in 07/27/2014.

Archived url: <http://bit.ly/2JPupxV>

Interaction techniques: Inspect, select, filter.

Topic: Environment/science/technology.

Related news story 1: “Nunca se viu um Verão assim? Não é verdade”.

Url: <https://www.publico.pt/2014/07/27/ecosfera/noticia/nunca-se-viu-um-verao-assim-nao-e-verdade-1664372>

Published in 07/27/2014.

Related news story 2: “1949, o ano em que as senhoras deixaram as meias em casa”.

Url: <https://www.publico.pt/2014/07/27/ecosfera/noticia/1949-o-ano-em-que-as-senhoras-deixaram-as-meias-em-casa-1664242>

Published in 07/27/2014.



Figure 135 – “As linhas da liberdade” screenshot.

Title: “As linhas da liberdade”.

Original url: <https://acervo.publico.pt/25abril/as-linhas-da-liberdade>

Published in 04/25/2014.

Archived url: <http://bit.ly/2xW193a>

Interaction techniques: Inspect, select.

Topics: Social issues, politics/governance.

Related news story 1: “O descontentamento, a revolta e a alegria saíram à rua em Lisboa”.

Url: <https://www.publico.pt/2014/04/25/politica/noticia/o-descontentamento-a-revolta-e-a-alegria-sairam-a-rua-em-lisboa-1633640>

Published in 04/25/2014.

Related news story 2: “Primeiras leis, como se construiu o Portugal democrático”.

Url: <https://www.publico.pt/2014/04/25/politica/noticia/como-se-construiu-o-portugal-democratico-1633173>

Published in 04/25/2014.



Figure 136 – “Retrato dos aeroportos nacionais” screenshot.

Title: “Retrato dos aeroportos nacionais”.

Original url: <https://www.publico.pt/2014/02/09/infografia/retrato-dos-aeroportos-nacionais-115>

Published in 02/09/2014.

Archived url: <http://bit.ly/2YVhukw>

Interaction techniques: Select, inspect.

Topic: Economy/business.

Related news story 1: “Vinci investe 271 milhões nos aeroportos para aumentar tráfego em 2,6%”.

Url: <https://www.publico.pt/2014/04/08/economia/noticia/vinci-investe-271-milhoes-nos-aeroportos-para-aumentar-trafego-em-26-1631410>

Published in 04/08/2014.

Related news story 2: “Eixo Sines-Badajoz concentra investimento na ferrovia”.

Url: <https://www.publico.pt/2014/04/08/economia/noticia/ligacao-que-permite-aproximar-sines-setubal-e-lisboa-a-badajoz-devera-estar-concluida-em-2019-1631405>

Published in 04/08/2014.



Figure 137 – “VIH: O vírus que apareceu em Kinshasa em 1920 e alastrou para o mundo inteiro” screenshot.

Title: “VIH: O vírus que apareceu em Kinshasa em 1920 e alastrou para o mundo inteiro”.

Original url: <https://www.publico.pt/2014/11/30/infografia/vih-o-virus-que-apareceu-em-kinshasa-em-1920-e-alastrou-para-o-mundo-inteiro-153>

Published in 11/30/2014.

Archived url: <http://bit.ly/2Sp5khg>

Interaction techniques: Select, narrate/history.

Topic: Health.

Related news story 1: “Médicos vão generalizar análises ao VIH e testes rápidos chegarão a mais centros de saúde”.

Url: <https://www.publico.pt/2014/12/01/sociedade/noticia/medicos-vaio-generalizar-analises-ao-vih-e-testes-rapidos-chegarao-a-mais-centros-de-saude-1678126>

Published in 12/01/2014.

Related news story 2: “Um quarto dos inquiridos com VIH não tem parceiro amoroso ou sexual”.

Url: <https://www.publico.pt/2014/12/01/sociedade/noticia/cerca-de-um-quarto-dos-inquiridos-infectados-com-hiv-nao-tem-parceiro-amoroso-ou-sexual-1677803>

Published in 12/01/2014.



Figure 138 – “A I Liga vista de dez maneiras diferentes” screenshot.

Title: “A I Liga vista de dez maneiras diferentes”.

Original url: <https://acervo.publico.pt/desporto/a-i-liga-vista-de-dez-maneiras-diferentes>

Published in 12/02/2014.

Archived url: <http://bit.ly/32CMNTd>

Interaction techniques: Filter, select.

Topic: Sports.

Related news story 1: “Abel Xavier é o novo treinador do Farense”.

Url: <https://www.publico.pt/2014/12/01/desporto/noticia/abel-xavier-e-o-novo-treinador-do-farense-1678035>

Published in 12/01/2014.

Related news story 2: “O Super Mario, os judeus e a nova polémica de Balotelli”.

Url: <https://www.publico.pt/2014/12/02/desporto/noticia/o-super-mario-os-judeus-e-a-nova-polemica-de-balotelli-1678144>

Published in 12/02/2014.



Figure 139 – “Os salários dos gestores do PSI 20 à lupa” screenshot.

Title: “Os salários dos gestores do PSI 20 à lupa”.

Original url: <https://www.publico.pt/2014/06/08/infografia/psi-20-134>

Published in 06/08/2014.

Archived url: <http://bit.ly/2O1wwnN>

Interaction techniques: Filter, select, narrate/history.

Topic: Economy/bisness.

Related news story 1: “Presidentes executivos do PSI20 ganharam 33,5 vezes mais do que os trabalhadores”.

Url: <https://www.publico.pt/2014/06/09/economia/noticia/presidentes-executivos-do-psi20-ganharam-335-vezes-mais-do-que-os-trabalhadores-1639190>

Published in 06/09/2014.

Related news story 2: “A revolução dos serviços públicos está em marcha”.

Url: <https://www.publico.pt/2014/06/08/economia/noticia/a-revolucao-dos-servicos-publicos-esta-em-marcha-e-agora-quem-paga-a-factura-1639069>

Published in 06/08/2014.

C.2. Interactive pieces from 2015



Figure 140 – “A água que gastamos mas não vemos” screenshot.

Title: “A água que gastamos mas não vemos”.

Original url: <https://www.publico.pt/2015/03/21/infografia/pegada-hidrica-nos-produtos-de-consumo-160>

Published in 03/21/2015.

Archived url: <http://bit.ly/2xXvDI2>

Interaction techniques: Inspect, select.

Topic: Environment/science/technology.

Related news story 1: “426 milhões de litros de água deitados fora num dia”.

Url: <https://www.publico.pt/2015/03/22/sociedade/noticia/426-milhoes-de-litros-de-agua-deitados-fora-num-dia-1689925>

Published in 03/22/2015.

Related news story 2: “O espectáculo do mar na ‘maré do século””.

Url: <https://www.publico.pt/2015/03/21/fotogaleria/o-espectaculo-do-mar-na-mare-do-seculo-346248>

Published in 03/21/2015.

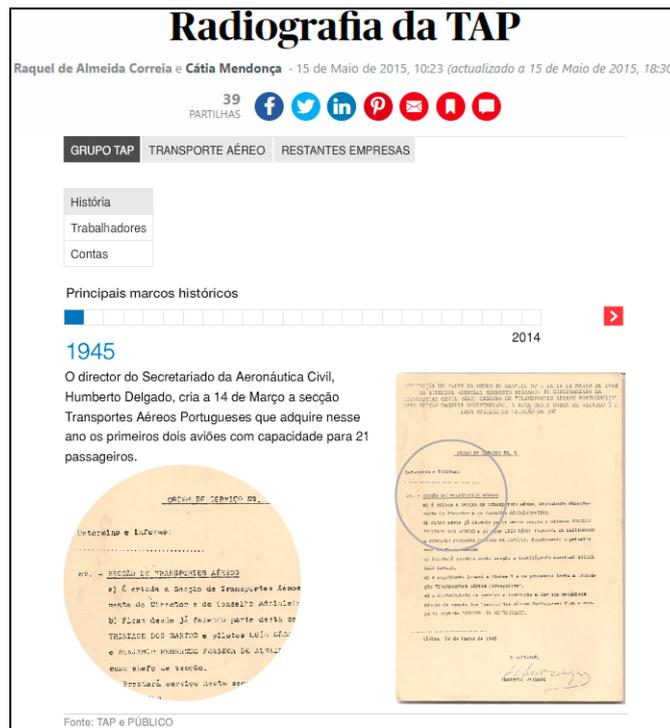


Figure 141 – “Radiografia da TAP” screenshot.

Title: “Radiografia da TAP”.

Original url: <https://www.publico.pt/2015/05/15/infografia/-166>

Published in 05/15/2015.

Archived url: <http://bit.ly/2xSNmdx>

Interaction techniques: Narrate/history, select, inspect.

Topic: Economy/business.

Related news story 1: “Vamos tomar café a Évora no meu avião”.

Url: <https://www.publico.pt/2015/05/10/sociedade/noticia/anda-com-eles-ver-os-avioes-1693977>

Published in 05/10/2015.

Related news story 2: “A Humberto Pedrosa só falta a TAP para ter um império nos transportes”.

Url: <https://www.publico.pt/2015/05/18/economia/noticia/humberto-pedrosa-o-empresario-a-quem-so-falta-comprar-a-tap-1695920>

Published in 05/18/2015.

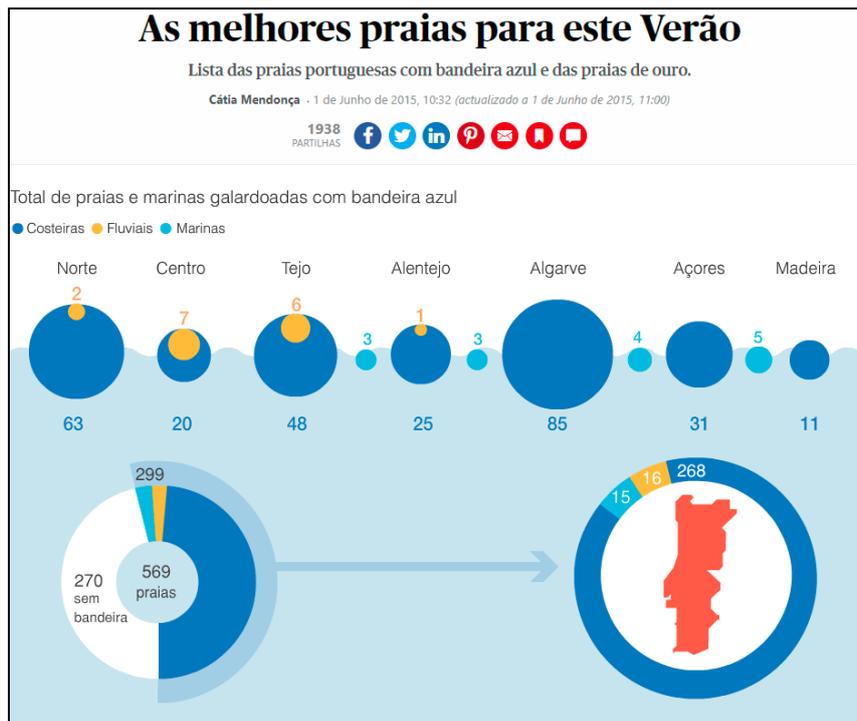


Figure 142 – “As melhores praias para este Verão” screenshot.

Title: “As melhores praias para este Verão”.

Original url: <https://www.publico.pt/2015/06/01/infografia/bandeiras-azuis-em-2014-133>

Published in 06/01/2015.

Archived url: <http://bit.ly/2NYTADw>

Interaction techniques: Filter, select.

Topics: Lifestyle/culture, environment/science/technology.

Related news story 1: “Tanto tempo depois, a barrinha de Esmoriz voltará a ter uma ponte e um cais”.

Url: <https://www.publico.pt/2015/06/01/local/noticia/tanto-tempo-depois-a-barrinha-de-esmoriz-voltara-a-ter-uma-ponte-e-um-cais-1697302>

Published in 06/01/2015.

Related news story 2: “Lisboetas sentem-se cada vez mais acoçados pelos turistas”.

Url: <https://www.publico.pt/2015/06/01/local/noticia/lisboetas-sentemse-cada-vez-mais-acossados-pelos-turistas-1697332>

Published in 06/01/2015.

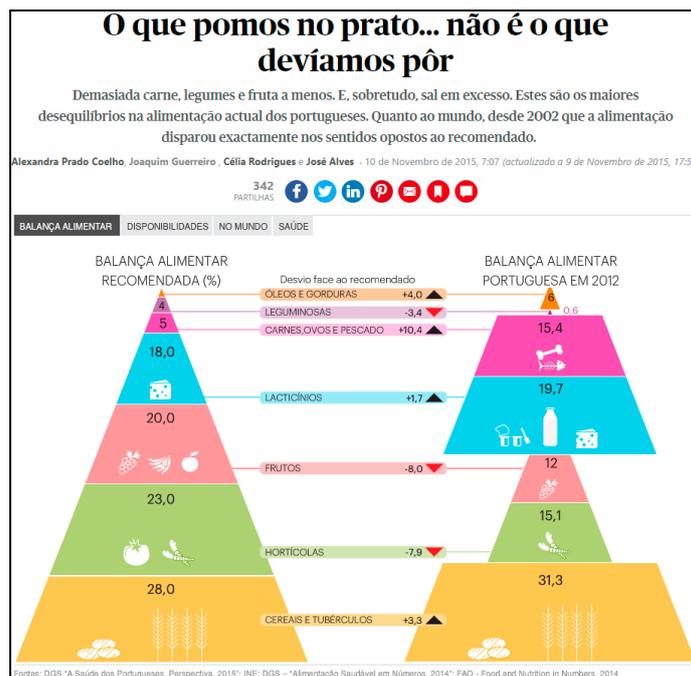


Figure 143 – “O que pomos no prato... não é o que devíamos pôr” screenshot.

Title: “O que pomos no prato... não é o que devíamos pôr”.

Original url: <https://acervo.publico.pt/multimedia/infografia/o-que-pomos-no-prato-nao-e-o-que-deviamos-por-177>

Published in 11/10/2015.

Archived url: <http://bit.ly/2O0EHk8>

Interaction techniques: Narrate/history, select.

Topic: Health.

Related news story 1: “O que pomos no prato... não é o que devíamos pôr”.

Url: <https://www.publico.pt/2015/11/10/sociedade/noticia/o-que-pomos-no-prato-nao-e-o-que-deviamos-por-1713865>

Published in 11/10/2015.

Related news story 2: “Nesta escola, a comida é vegetariana há 42 anos e ninguém estranha ou reclama”.

Url: <https://www.publico.pt/2015/11/02/sociedade/noticia/nesta-escola-a-comida-e-vegetariana-ha-42-anos-e-ninguem-estranha-ou-reclama-1713024>

Published in 11/02/2015.



Figure 144 – “Relações que matam” screenshot.

Title: “Relações que matam”.

Original url: <https://www.publico.pt/2015/04/09/infografia/relacoes-que-matam-156>

Published in 04/09/2015.

Archived url: <http://bit.ly/2XRJFEj>

Interaction techniques: Narrate/history, inspect.

Topic: Social issues.

Related news story 1: “Nós, homens, pensamos que as mulheres são nossas”.

Url: <https://www.publico.pt/2015/04/12/sociedade/noticia/nos-homens-pensamos-que-as-mulheres-sao-nossas-1691505>

Published in 04/12/2015.

Related news story 2: “Bebé terá sido morto à facada pelo pai em Oeiras”.

Url: <https://www.publico.pt/2015/04/08/sociedade/noticia/bebe-tera-sido-morto-a-facada-pelo-pai-em-lisboa-1691770>

Published in 04/08/2015.



Figure 145 – “Uma década de pobreza e algumas medidas” screenshot.

Title: “Uma década de pobreza e algumas medidas”.

Original url: <https://www.publico.pt/2015/02/15/infografia/uma-decada-de-pobreza-e-algumas-medidas-158>

Published in 02/15/2015.

Archived url: <http://bit.ly/2SkjaRS>

Interaction techniques: Select, inspect, narrate/history.

Topics: Economy/business, social issues.

Related news story 1: “Estado paga quase 50 mil refeições por dia a famílias carenciadas”.

Url: <https://www.publico.pt/2015/02/15/sociedade/noticia/estado-paga-quase-50-mil-refeicoes-a-familias-carenciadas-por-dia-1686076>

Published in 02/15/2015.

Related news story 2: “Todos os dias Maria Madalena caminha até casa com as marmitas”.

Url: <https://www.publico.pt/2015/02/15/sociedade/noticia/todos-os-dias-maria-madalena-caminha-ate-casa-com-as-marmitas-1686109>

Published in 02/15/2015.

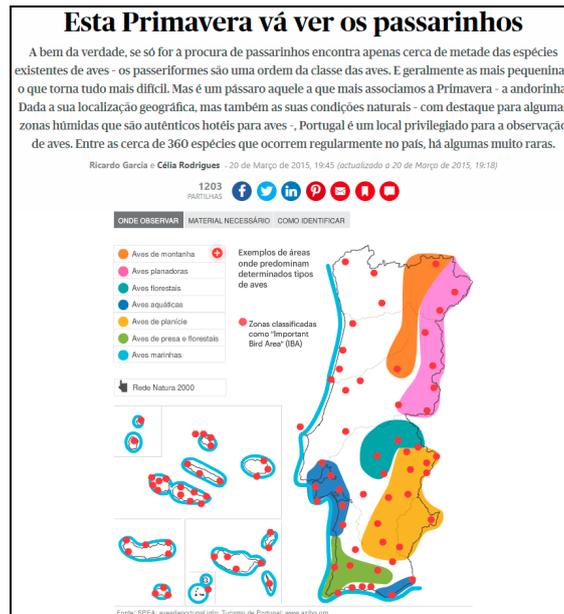


Figure 146 – “Esta Primavera vá ver os passarinhos” screenshot.

Title: “Esta Primavera vá ver os passarinhos”.

Original url: <https://www.publico.pt/2015/03/20/infografia/esta-primavera-va-ver-os-passarinhos-161>

Published in 03/20/2015.

Archived url: <http://bit.ly/30Dh42z>

Interaction techniques: Narrate/history, inspect, select.

Topics: Environment/science/technology, lifestyle/culture.

Related news story 1: “No equinócio da Primavera, a Lua vai tapar o Sol e as marés vão ser maiores”.

Url: <https://www.publico.pt/2015/03/19/ciencia/noticia/no-equinocio-da-primavera-a-lua-vai-tapar-o-sol-e-as-mares-va-ser-maiores-1689600>

Published in 03/19/2015.

Related news story 2: “Desvendado o mistério dos estranhos mamíferos sul-americanos de Darwin”.

Url: <https://www.publico.pt/2015/03/20/ciencia/noticia/desvendado-o-misterio-dos-estranhos-mamiferos-sulamericanos-de-darwin-1689786>

Published in 03/20/2015.



Figure 147 – “Cristiano Ronaldo em golos” screenshot.

Title: “Cristiano Ronaldo em golos”.

Original url: <https://www.publico.pt/2015/01/12/infografia/cristiano-ronaldo-em-golos-157>

Published in 01/12/2015.

Archived url: <http://bit.ly/2XPxzvE>

Interaction techniques: Select, inspect, narrate/history.

Topic: Sports.

Related news story 1: “O que mais distingue Cristiano Ronaldo dos outros”.

Url: <https://www.publico.pt/2015/01/12/desporto/noticia/o-que-mais-distingue-cristiano-ronaldo-dos-outros-a-potencia-1681928>

Published in 01/12/2015.

Related news story 2: “Desvendado Ronaldo é a marca portuguesa mais valiosa de sempre”.

Url: <https://www.publico.pt/2015/01/12/desporto/noticia/ronaldo-e-a-marca-portuguesa-mais-valiosa-de-sempre-1681979>

Published in 01/12/2015.

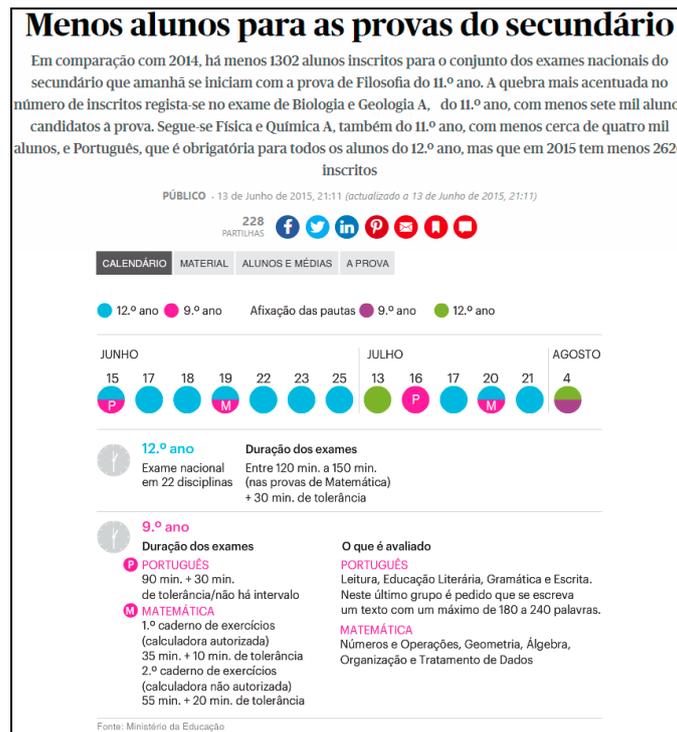


Figure 148 – “Menos alunos para as provas do secundário” screenshot.

Title: “Menos alunos para as provas do secundário”.

Original url: <https://www.publico.pt/2015/06/13/infografia/menos-alunos-para-as-provas-do-secundario-168>

Published in 06/13/2015.

Archived url: <http://bit.ly/2Yb8eLS>

Interaction techniques: Select, narrate/history.

Topic: Education.

Related news story 1: “Classificar alunos de 1 a 5 é penalizado”.

Url: <https://www.publico.pt/2015/06/14/sociedade/noticia/classificar-alunos-de-1-a-5-e-penalizador-1698814>

Published in 06/14/2015.

Related news story 2: “Demasiada pressão para um teste igual aos outros”.

Url: <https://www.publico.pt/2015/06/14/sociedade/noticia/demasiada-pressao-para-um-teste-igual-aos-outros-1698884>

Published in 06/14/2015.

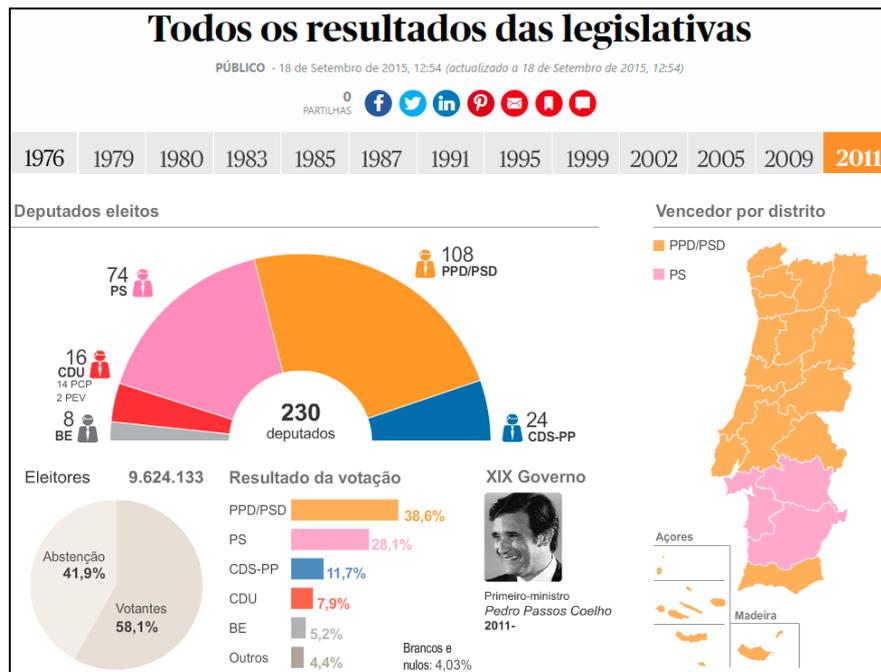


Figure 149 – “Todos os resultados das legislativas” screenshot.

Title: “Todos os resultados das legislativas”.

Original url: <http://www.publico.pt/multimedia/infografia/todos-os-resultados-das-legislativas-171>

Published in 09/18/2015.

Archived url: <http://bit.ly/2JPKsf8>

Interaction techniques: Inspect, select.

Topic: Politics/governance.

Related news story 1: “Catarina, a grande nas ruas”.

Url: <https://www.publico.pt/2015/09/19/politica/noticia/catarina-a-grande-nas-ruas-1708374>

Published in 09/19/2015.

Related news story 2: “Costa assume baixa de impostos para a classe média, Passos não se compromete”.

Url: <https://www.publico.pt/2015/09/17/politica/noticia/ultimo-frenteafrente-entre-passos-e-costa-nas-radios-1708068>

Published in 09/17/2015.

C.3. Interactive pieces from 2016

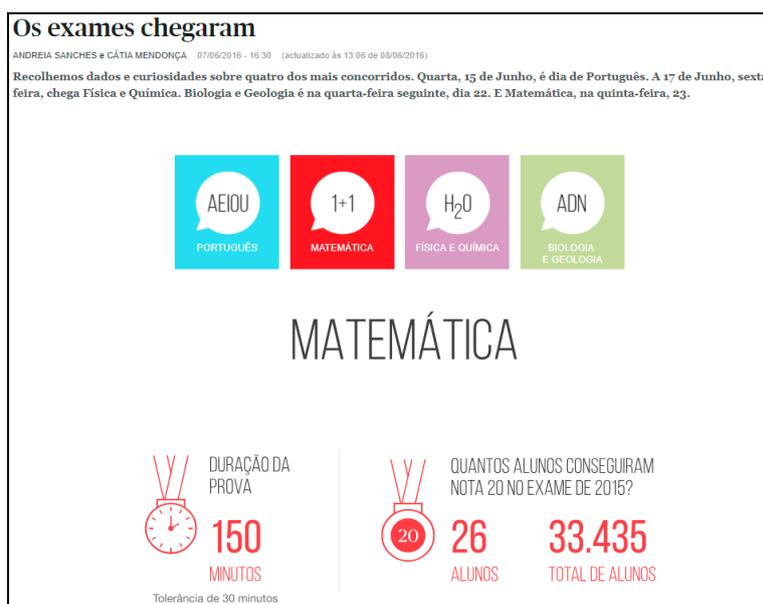


Figure 150 – “Os exames chegaram” screenshot.

Title: “Os exames chegaram”.

Original url: <https://acervo.publico.pt/multimedia/infografia/os-exames-chegaram-192>

Published in 06/07/2016.

Archived url: <http://bit.ly/2SmcBOV>

Interaction technique: Select.

Topic: Education.

Related news story 1: “Alunos do secundário continuam a fugir de Física e Matemática”.

Url: <https://www.publico.pt/2016/06/08/sociedade/noticia/alunos-do-secundario-continuam-a-fugir-de-fisica-e-matematica-1734481>

Published in 06/08/2016.

Related news story 2: “Fácil, básica, acessível. Assim foi a prova de aferição de Português”.

Url: <https://www.publico.pt/2016/06/06/sociedade/reportagem/afericao-2-e-5-ano-provas-acessiveis-para-muita-concentracao-1734230>

Published in 06/06/2016.

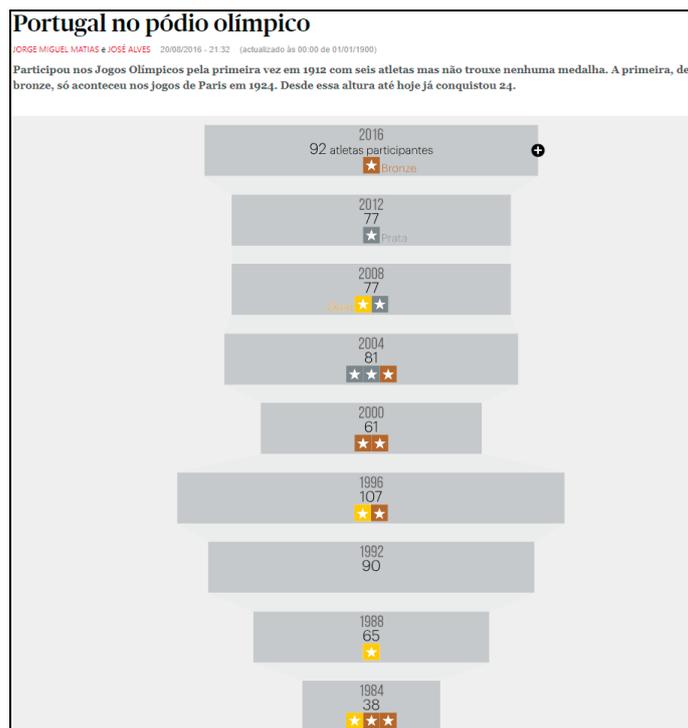


Figure 151 – “Portugal no pódio olímpico” screenshot.

Title: “Portugal no pódio olímpico”.

Original url: <https://acervo.publico.pt/multimedia/infografia/portugal-no-podio-olimpico-197>

Published in 08/20/2016.

Archived url: <http://bit.ly/32xbQXS>

Interaction technique: Select.

Topic: Sports.

Related news story 1: “COI: Brasil realizou ‘Jogos Olímpicos inesquecíveis’”.

Url: <https://www.publico.pt/2016/08/20/desporto/noticia/coi-brasil-realizou-jogos-olimpicos-inesqueciveis-1741938>

Published in 08/20/2016.

Related news story 2: “Usain Bolt: três vezes três antes dos 30”.

Url: <https://www.publico.pt/2016/08/20/desporto/noticia/bolt-completa-triplotriplo-com-vitoria-a-estafeta-1741892>

Published in 08/20/2016.

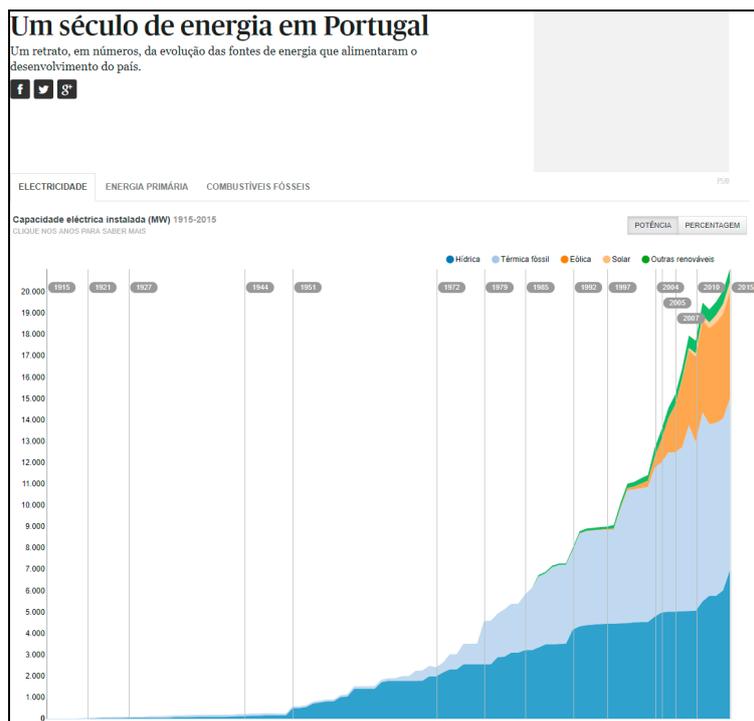


Figure 152 – “Um século de energia em Portugal” screenshot.

Title: “Um século de energia em Portugal”.

Original url: <https://acervo.publico.pt/ecosfera/interactivo/um-seculo-de-energia-em-portugal>

Published in 06/26/2016.

Archived url: <http://bit.ly/2XPyrAq>

Interaction techniques: Inspect, select, filter, connect.

Topic: Environment/science/technology.

Related news story 1: “Um vazio legal que pode valer 400 milhões”.

Url: <https://www.publico.pt/2016/06/23/politica/noticia/um-vazio-legal-que-pode-valer-400-milhoes-1736150>

Published in 06/23/2016.

Related news story 2: “Joana Vasconcelos e Vhils pintam torres em ‘terras do demo’”.

Url: <https://www.publico.pt/2016/07/05/local/noticia/joana-vasconcelos-e-vhils-pintam-torres-em-terras-do-demo-1737362>

Published in 07/05/2016.

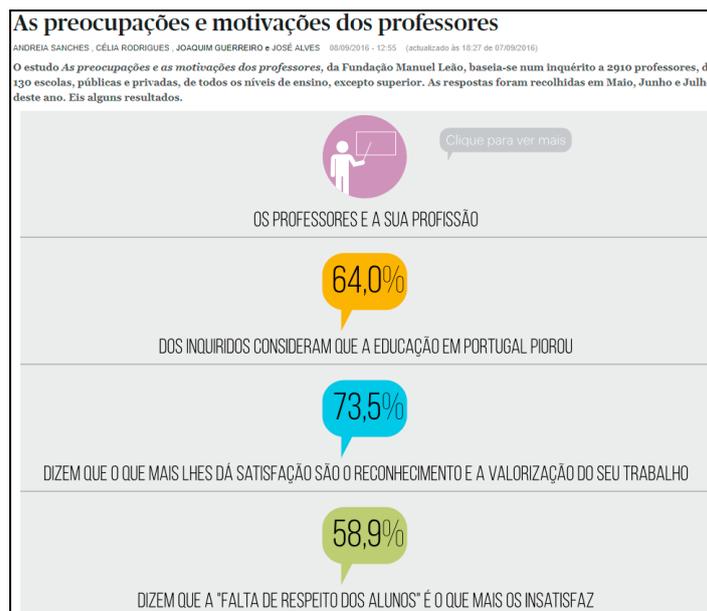


Figure 153 – “As preocupações e motivações dos professores” screenshot.

Title: “As preocupações e motivações dos professores”.

Original url: <https://acervo.publico.pt/multimedia/infografia/as-preocupacoes-e-motivacoes-dos-professores-199>

Published in 09/08/2016.

Archived url: <http://bit.ly/2JFBthV>

Interaction technique: Select.

Topic: Education.

Related news story 1: “Exaustos, desiludidos ou baralhados. Um terço dos professores sente-se assim”.

Url: <https://www.publico.pt/2016/09/08/sociedade/noticia/exaustos-desiludidos-baralhados-e-assim-com-um-terco-dos-professores-1743362>

Published in 09/08/2016.

Related news story 2: ““Os miúdos continuam com muita vontade de aprender, são é menos focados””.

Url: <https://www.publico.pt/2016/09/08/sociedade/noticia/os-miudos-continuam-com-muita-vontade-de-aprender-sao-e-menos-focados-1743432>

Published in 09/08/2016.



Figure 154 – “Como se elege um Presidente nos EUA” screenshot.

Title: “Como se elege um Presidente nos EUA”.

Original url: <https://acervo.publico.pt/multimedia/infografia/como-se-elege-um-presidente-nos-eua-202>

Published in 11/04/2016.

Archived url: <http://bit.ly/2GiUIB3>

Interaction technique: Narrate/history.

Topics: Foreign affairs, politics/governance.

Related news story 1: “Um Trump a menos em Atlantic City, a cidade onde os casinos vão para morrer”.

Url: <https://www.publico.pt/2016/11/05/mundo/reportagem/um-trump-a-menos-em-atlantic-city-a-cidade-onde-os-casinos-va-para-morrer-1750008>

Published in 11/05/2016.

Related news story 2: “Hillary tem Beyonce, mas as ‘Latinas for Trump’ dão espectáculo na Quinta Avenida”.

Url: <https://www.publico.pt/2016/11/05/mundo/reportagem/hillary-tem-beyonce-mas-as-latinas-for-trump-dao-espectaculo-na-quinta-avenida-1750134>

Published in 11/05/2016.



Figure 155 – “De que é feita a cerveja” screenshot.

Title: “De que é feita a cerveja”.

Original url: <https://acervo.publico.pt/multimedia/infografia/o-segredo-da-cerveja-201>

Published in 10/28/2016.

Archived url: <http://bit.ly/2LXeZKI>

Interaction techniques: Select, narrate/history.

Topic: Lifestyle/culture.

Related news story 1: “Cerveja artesanal: pequenos, mas resistentes”.

Url: <https://www.publico.pt/2016/10/31/economia/noticia/cerveja-artesanal-pequenos-mas-resistentes-1749401>

Published in 10/31/2016.

Related news story 2: “Consumo de cerveja em Portugal caiu 25% em dez anos”.

Url: <https://www.publico.pt/2016/10/31/economia/noticia/consumo-de-cerveja-em-portugal-caiu-25-em-dez-anos-1749399>

Published in 10/31/2016.

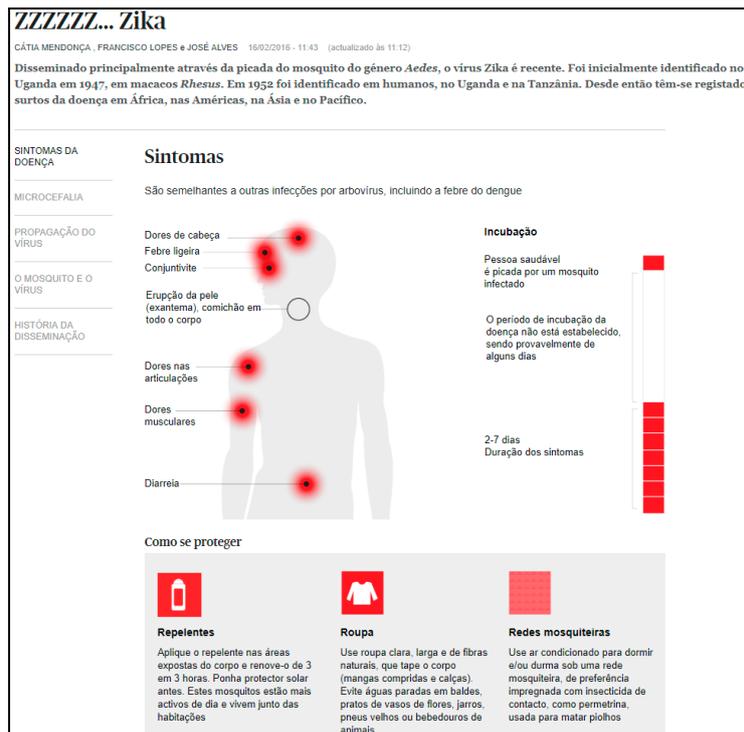


Figure 156 – “ZZZZZZ... Zika” screenshot.

Title: “ZZZZZZ... Zika”.

Original url: <https://acervo.publico.pt/multimedia/infografia/zzzzzzika-185>

Published in 02/16/2016.

Archived url: <http://bit.ly/2JSBzRV>

Interaction techniques: Select, narrate/history.

Topic: Health.

Related news story 1: “65% dos médicos trabalham no Serviço Nacional de Saúde”.

Url: <https://www.publico.pt/2016/02/10/sociedade/noticia/so-65-dos-medicos-trabalham-no-servico-nacional-de-saude-1722697>

Published in 02/10/2016.

Related news story 2: “Foco na produtividade ‘prejudica doentes’ e ‘torna cuidados mais impessoais’”.

Url: <https://www.publico.pt/2016/02/18/sociedade/noticia/livro-alerta-que-foco-em-indicadores-e-tempos-de-consulta-prejudica-os-doentes-1723604>

Published in 02/18/2016.

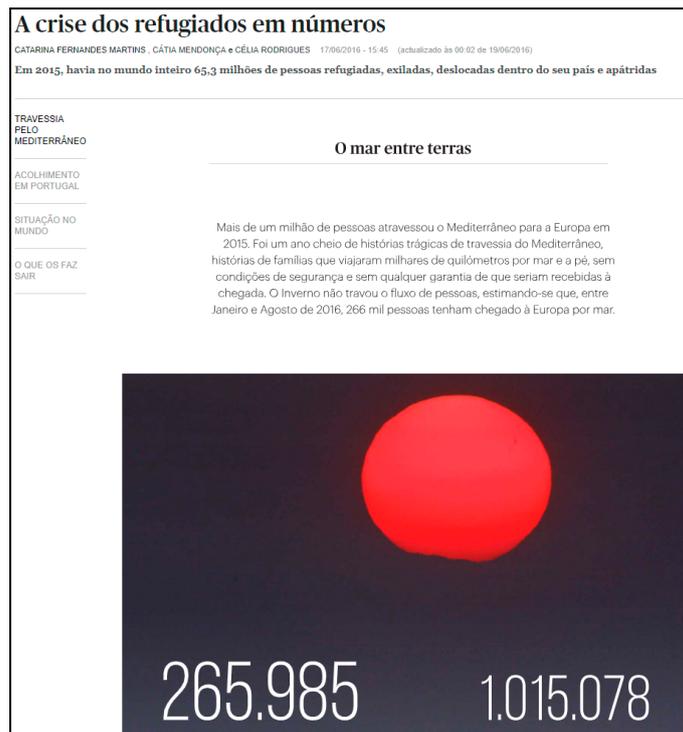


Figure 157 – “A crise dos refugiados em números” screenshot.

Title: “A crise dos refugiados em números”.

Original url: <https://acervo.publico.pt/multimedia/infografia/dia-mundial-do-refugiado-193>

Published in 06/17/2016.

Archived url: <http://bit.ly/2XWbmw1>

Interaction techniques: Select, narrate/history.

Topics: Foreign affairs, social issues.

Related news story 1: “Médicos Sem Fronteiras viram costas a dezenas de milhões em fundos europeus”.

Url: <https://www.publico.pt/2016/06/17/mundo/noticia/medicos-sem-fronteiras-recusam-fundos-da-ue-em-protesto-contr-a-acordo-com-a-turquia-1735432>

Published in 06/17/2016.

Related news story 2: “Portugal está a acolher. Como vai integrar?”.

Url: <https://acervo.publico.pt/sociedade/noticia/portugal-esta-a-acolher-como-vai-integrar-1735470>

Published in 06/19/2016.



Figure 158 – “Os 16 desportos de Portugal no Rio 2016” screenshot.

Title: “Os 16 desportos de Portugal no Rio 2016”.

Original url: <https://acervo.publico.pt/multimedia/infografia/olimpicos-teste-196>

Published in 08/05/2016.

Archived url: <http://bit.ly/2JC23Z4>

Interaction technique: Select.

Topic: Sports.

Related news story 1: “Acabou cedo a estreia olímpica de Shao Jieni”.

Url: <https://www.publico.pt/2016/08/07/desporto/noticia/acabou-cedo-a-estreia-olimpica-de-shao-jieni-1740541>

Published in 08/07/2016.

Related news story 2: “Por dois pontos, João Costa falhou final do tiro”.

Url: <https://www.publico.pt/2016/08/06/desporto/noticia/joao-costa-fora-da-final-do-tiro-1740523>

Published in 08/06/2016.

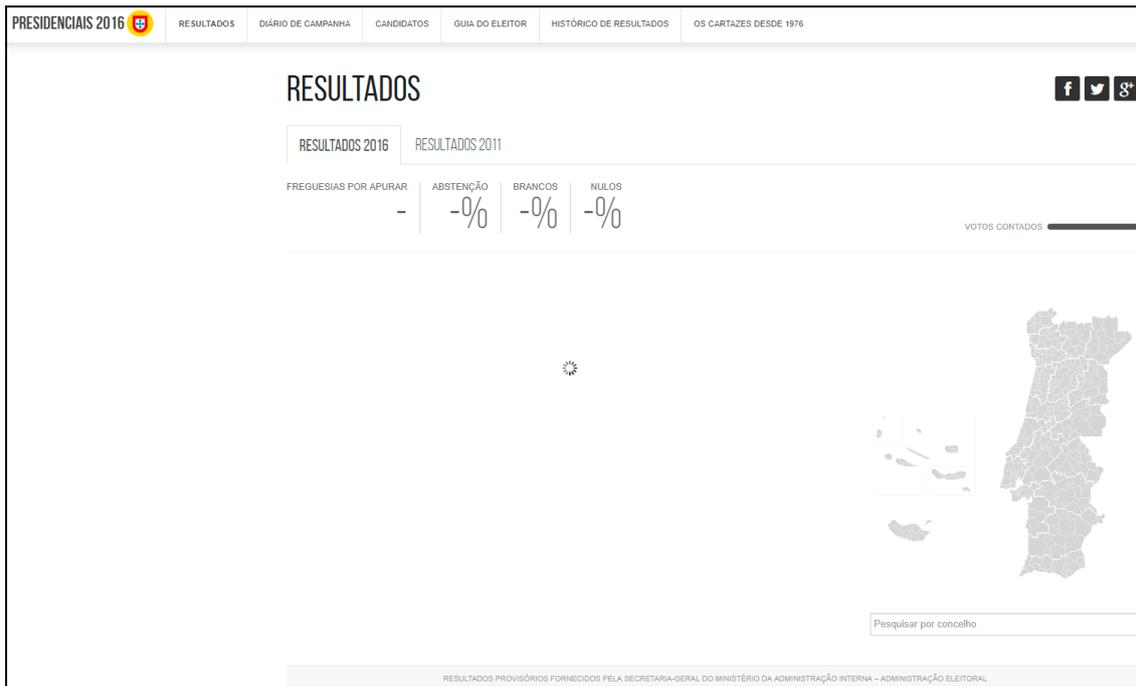


Figure 159 – “Presidenciais 2016” screenshot.

Title: “Presidenciais 2016”.

Original url: <https://acervo.publico.pt/presidenciais-2016/resultados>

Published in 01/24/2016.

Archived url: <http://bit.ly/2Z28c6m>

Interaction techniques: Filter, inspect, explore, select, narrate/history.

Topic: Politics/governance.

Related news story 1: “Marcelo ganha à primeira com dobro dos votos de Nóvoa”.

Url: <https://www.publico.pt/2016/01/24/politica/noticia/marcelo-rebelo-de-sousa-eleito-presidente-1721277>

Published in 01/24/2016.

Related news story 2: “Marisa e o BE mudaram ‘o mapa político de Portugal’”.

Url: <https://www.publico.pt/2016/01/24/politica/noticia/marisa-o-be-mudaram-o-mapa-politico-de-portugal-1721299>

Published in 01/24/2016.