DATA JOURNALISM AND DEMOCRACY: HOW AMERICAN MASS MEDIA FRAMED THE 2016 PRESIDENTIAL CAMPAIGN IN THE UNITED STATES USING DATA VISUALIZATION

A Thesis presented to

the Faculty of the Graduate School

at the University of Missouri-Columbia

In Partial Fulfillment

of the Requirements for the Degree

Master of Arts

by

IULIIA ALIEVA

Dr. Randall Smith, Thesis Supervisor

MAY 2017

The undersigned, appointed by the Dean of the Graduate School, have examined the thesis entitled:

DATA JOURNALISM AND DEMOCRACY: HOW AMERICAN MASS MEDIA FRAMED THE 2016 PRESIDENTIAL CAMPAIGN IN THE UNITED STATES USING DATA VISUALIZATION

presented by Iuliia Alieva,

a candidate for the degree of Master of Arts

and hereby certify that in their opinion it is worthy of acceptance

Professor Randall Smith

Professor Yong Volz

Professor David Herzog

Professor Antonie Stam

ACKNOWLEDGEMENTS

Undoubtedly, the University of Missouri School of Journalism gave me an unparalleled advantage by immersing me into a new educational experience in the United States, which helped me to gain new qualifications, abilities to work in different cultural spaces and interact with professionals who can teach and inspire me into the future. With the help of the MU Journalism School Faculty, I improved my skills in data journalism and international journalism, as well as economic journalism and documentary filmmaking. I also gained experience in new technologies in the media, and I'm excited to apply my knowledge in the process of practical work in my home country.

Working on my thesis, I learned a lot and I would like to thank all the members of my thesis committee for their help and patience. My advisor and committee chair Professor Randall Smith carefully guided me through the process of working on my thesis, Professor Yong Volz shared valuable knowledge about methodology for the research, Professor David Herzog provided valuable feedback and expertise in data journalism, while Professor Antonie Stam added a unique perspective that helped me analyze my work from a different angle.

TABLE OF CONTENTS

Acknowled	gements	ii
List of Figu	res	vi
Abstract		vii
Chapter		
1.	Introduction	1
2.	Literature Review: Theory/Concept Explication	3
	Theoretical Approaches to Framing: Historical Perspective	4
	Traditional Interpretation of the Framing Effect	5
	Framing in the System of Media Effects	6
	The Implication of Framing in Political Communications	7
	Understanding the Main Principles of Data Visualization	9
	Studying the Implementation of Infographic Techniques	13
	Framing Effect in Data Visualization	15
3.	Methodology	17
	Introduction	17
	Research Design	18
	The Role of Textual and Visual Analysis	19
	The Implementation of the Interview as a Method	20
	Defining the Sample: Participant Selection and Information	21
	Data Collection Process	24
	Data Analysis	25
	Validity and Challenges	26

Limitations of the Study	27
Chapter Summary	.27
Findings	.28
Introduction	.28
Textual/Visual Analysis: Data coding	.28
Textual/Visual Analysis: Findings	31
Analysis of Graphics at The New York Times Before the 2016	
Election	.31
Analysis of Graphics at The New York Times After the 2016	
Election	.38
Analysis of Graphics at The Washington Post Before the 2016	
Election	.40
Analysis of Graphics at The Washington Post After the 2016	
Election	.44
Interview Information	.46
Participant Information	.47
Data Analysis of the Interviews	.48
Structure of the Data Department	.49
Purpose of Data Visualization	.52
Work Process	.55
Challenges and Limitations	57
Frames and Trends	.60
Outreach	.62
	Limitations of the Study Chapter Summary Findings Introduction Textual/Visual Analysis: Data coding Textual/Visual Analysis: Findings Analysis of Graphics at The New York Times Before the 2016 Election Analysis of Graphics at The New York Times After the 2016 Election Analysis of Graphics at The Washington Post Before the 2016 Election Analysis of Graphics at The Washington Post Before the 2016 Election Analysis of Graphics at The Washington Post After the 2016 Election Interview Information Participant Information Data Analysis of the Interviews Structure of the Data Department Purpose of Data Visualization Work Process Challenges and Limitations Frames and Trends Outreach

]	Recommendations for Future Coverage	64
(Chapter Summary	66
5. Disc	cussion	67
(Summary	67
]	Limitations	73
r	Theoretical Implications	74
]	Practical Implications	74
]	Recommendations	75
References		76
Appendices		83
A. IRB Ap	proval	83
B. Participa	ant Information	85
C. Email P	rotocol	
D. Intervie	w script	

LIST OF FIGURES

TABLES				
1.	Descriptive codes and sub-codes for textual and visual analysis			
2.	Descriptive codes for the interview			
ILLUSTRATIONS				
1.	The types of infographics used by The New York Times before election32			
2.	Packed bubbles demonstrate colors that were used in The New York Times			
	visuals before election			
3.	Treemap of visualization types used by The New York Times in the month			
	after the election			
4.	The most used types of infographics by The Washington Post before			
	election			
5.	Packed bubbles demonstrate colors used in The Washington Post visuals			
	before election			
6.	Treemap of visualization types used by The Washington Post during the			
	month after the election			

DATA JOURNALISM AND DEMOCRACY: HOW AMERICAN MASS MEDIA FRAMED THE 2016 PRESIDENTIAL CAMPAIGN IN THE UNITED STATES USING DATA VISUALIZATION

Iuliia Alieva

Professor Randall Smith, Thesis Supervisor

ABSTRACT

The purpose of this study is to explore the role of data visualization in the media coverage of the 2016 presidential campaign in the United States by focusing on datavisualization projects from The New York Times and The Washington Post. The research is focused on how journalists implemented data-visualization techniques and how the theory of framing is connected with that process. A secondary purpose of the research is to collect opinions from journalists working in the field about how data visualization influenced coverage of the campaign and how future reporting can be improved.

This study consists of two parts: textual/visual analysis of data visualization examples from coverage of 2016 U.S. presidential campaign and interviews with the journalists involved in infographics production. Visual analysis was used for analyzing various design elements such as type of graphics, colors, fonts and how they helped to frame issues during the campaign. Textual analysis was used to identify the main issues and frames that were covered and considered important for the audience. The interviews provided information about the professional experience of data journalists and editors and their opinions about the role of data visualization and the problems and limitations that they experienced while working with it.

Chapter 1. Introduction

As technology becomes an increasingly integral part of society, graphics and data have come to play a significant role in journalism. In order to facilitate readers' perception of information and increase accuracy of their own reporting, journalists are looking for more effective methods of data presentation. Data visualization has become one of the strongest tools for the media to cope with the rising flood of information. Leading journalism universities, such as University of Missouri, Stanford University and Columbia University, as well as journalism organizations such as Investigative Reporters and Editors, invented new programs and trainings for data journalists to improve their knowledge of new tools that are constantly developing in the age of Internet and online media.

The SAS Institute created the Statistical Analysis System (SAS) for advanced analytics, multivariate analysis, business intelligence, data management and predictive analytics. It defines data visualization this way:

Data visualization is the presentation of data in a pictorial or graphical format. It enables decision makers to see analytics presented visually, so they can grasp difficult concepts or identify new patterns. With interactive visualization, you can take the concept a step further by using technology to drill down into charts and graphs for more detail, interactively changing what data you see and how it's processed (SAS, n. d.).

Data-visualization techniques help design and structure information, help readers perceive information, and focus attention on a particular element of a story. The concept

1

of data visualization is not new: Data was used for maps and graphs in the 17th century, while the pie chart was invented in the early 1800s. One of the first examples of statisticalbased graphics occurred when Charles Minard in 1869 mapped Napoleon's 1812 invasion of Russia. The map depicted the size of the army as well as the path of Napoleon's retreat from Moscow, and it tied that information to temperature and time scales for a more indepth understanding of the event (Tufte, 1983).

It is hard to ignore that the amount of available data is huge and growing every day. Computers have made it possible to process large amounts of data with lightning speed: "Today, data visualization has become a rapidly evolving blend of science and art that is certain to change the corporate landscape over the next few years" (SAS, n. d.).

For journalists, data visualizations can be a useful and effective way to visualize large amounts of complex data. Data visualizations makes it possible to illustrate multiple processes and patterns, as well as to identify the factors that influenced them.

The 2016 presidential election in the United States sparked an important debate concerning the use of data reporting by the American mass media. Despite all the sophisticated methods of election and poll data presentation during the presidential campaign, the media came to the conclusion that it is impossible to cover all of the aspects of the issue relying only on quantitative data (Rutenberg, 2016). Data visualization was actively used by journalists in their reporting. The purpose of this study is to explore the role of data visualization in the media coverage of the 2016 presidential campaign in the United States by focusing on data-visualization projects from The New York Times and The Washington Post. The research is focused on how journalists implemented data-visualization techniques and how the theory of framing is connected with that process.

Chapter 2. Literature Review: Theorizing/Concept Explication

Although data visualization is a relatively new topic in academia, various approaches to researching this recent phenomenon have already appeared. Data visualization was a valuable tool for presenting graphical information during the 2016 United States presidential campaign. Many newsrooms, such as The New York Times and The Washington Post, produced data-visualization content for illustrating the candidates' positions and voters' preferences. Visual information in journalism such as graphics and interactives can frame issues and promote particular perspectives (Diakopoulos & Hullman, 2011). One of the main aspects of this research is a focus on framing effect in data visualization during the American presidential campaign. The framing effect is not a unique phenomenon in mass media research: Journalists have individual biases that sometimes affect their reporting and lead to other related mass-communication effects such as priming and agenda-setting (Entman, 1993). It is important that journalists understand this natural tendency and improve the objectivity of their work.

The news is not valuable for the audience without the right context, which makes particular issues more or less important for readers. Stories include facts and narratives, but they also contain specific messages to the audience. Through data visualizations, mass media formulate the ways that the audience understands political processes. Mass media create an agenda and play a role in involving people in the democratic process (McCombs, 2005; Nelson, Lecheler, Schuck, & De Vreese, 2012). Data visualization is a way to present information, in many cases in concert with text, to form new types of content and use data with elements of infographic design.

Scientific research of data journalism is not yet a popular field of study, but there are researchers who have concentrated on different aspects of data visualization (Tufte, 1983; Segel & Heer, 2010; Burmester, Mast, Tille, & Weber, 2010) and framing in presentation of visual information (Hullman & Diakopoulos, 2011). To understand the phenomenon of framing in data visualization, it is necessary to explore theoretical approaches to framing as a theory in mass-media communication in order to identify particular issues and mechanisms of how it can be applied in data visualizations.

Theoretical Approaches to Framing: Historical Perspective

Framing theory was first put forth by Erving Goffman, who divided frames into two groups: social and natural. Goffman argued that frames help people interpret information, that this process goes on unintentionally, and the audience and creators of information sometimes are not aware of this standard of perceiving information (Goffman, 1974, p. 23).

In many cases, framing is illustrated in the system of other media effects, such as priming and agenda-setting (Bennett & Iyengar, 2008; Brewer, Graf & Willnat, 2003; Carragee & Roefs, 2004; McCombs, 2005; Scheufele & Tewksbury, 2007). The role of framing in political communications was studied by Bennett & Iyengar (2008), Chong & Druckman (2007), Entman (2007); the ability of the media to shape people's opinions and political preferences was examined by Iyengar & Kinder (1987), McCombs & Shaw (1972).

Bennett and Iyengar (2008) studied the dynamics of media effects in a retrospective manner. The research is focused on the media-effects influence on political communication; Bennett and Iyengar illustrated how present media effects were developed on the foundation of the minimal-effects era of 1920-1950 when scientists wrote about the minimal impact of the media on the audience. The strong-effects era began with the research of McCombs and Shaw, who investigated the agenda-setting effect (McCombs & Shaw, 1972). This new direction was taken by Goffman (1974) and Edelman (1988), who studied the framing and language aspect of this phenomenon. They then passed this sociological tradition to Entman (1993), Iyengar (1991) and others.

Traditional Interpretation of the Framing Effect

Scheufele (1999) explored basic definitions and implications of the framing effect in the media. The study made an attempt to form a classification of previous research about the same issues and divided them into two categories: the type of frame (media/audience frames) and the way that these frames worked (p. 104). The author identified four stages of framing: frame building, frame setting, individual-level processes of framing, and a feedback loop from audiences to journalists (p. 114-118). Scheufele accumulated scientific groundwork from past research, differentiated the definition of the framing from other close concepts such as priming and agenda-setting, and developed the principal core of this media effect.

The main point of the framing effect as an integrated process that consists of production, content and media-use perspectives was stated by De Vreese (2005). He used typology of generic and issue-specific frames and wrote about how generic news frames could be used for understanding cross-national differences in news coverage.

According to Entman (1993), framing can be seen from different dimensions, but generally, it includes selection and salience. Entman stated that journalists selected certain details of informational reality and modified them into text, often over-stressing particular definitions and interpretations, moral evaluations and recommendations. The author provided an exact order of the process which led to framing and consisted of four stages. The first stage was defining problems (advantages and disadvantages of the situation), second was diagnosing them (forces that led to the problem), the third was moral judgments (actors and effects), and the final one was remedies (treatment for the problem) (p. 52).

Framing in the System of Media Effects

Some researchers have different interpretations of the framing effect, particularly in cases with intersections of priming and agenda-setting (Carragee & Roefs, 2004; Scheufele & Iyengar 2014; Scheufele & Tewksbury, 2007).

Scheufele & Tewksbury (2007) studied theoretical explanations of news framing, agenda setting and priming effects in an attempt to determine whether connections between these concepts exist. Carragee & Roefs (2004) emphasized the significance of the definitions of frames and conceptual problems in mass-media information. In the research, agenda setting is considered an extension of the framing effect, and the role of framing in political and social processes is emphasized. Chong and Druckman (2007) analyzed how people's opinions are formed in different environments. Connections and disconnections between media effects were found in the research of Scheufele & Iyengar (2012), who analyzed effects such as agenda setting, priming, salience-based effects and how they are correlated with framing. They described the evolution of these theories from previous concepts (the theory of Spiral of Silence, Cultivation) and concluded that all of them were

based on cognitive processes with the scheme of stimuli responses (p. 8-9). The investigation of the role of visualizations and their influence on decisions about political preferences was a significant part of the research.

McCombs (2005) provided a more detailed analysis of the agenda-setting effect and examined an experiment from the previous paper, conducted by McCombs and Shaw. This case-study demonstrated changes in people's attitudes during the 1968 U.S. Presidential campaign. It documented the agenda-setting effect by analyzing mass-media materials and interviews with voters. The author supported hypotheses with empirical findings, which demonstrated framing by mass-media publications during presidential campaigns, especially in regards to issues such as arms control, civil rights and unemployment.

The above-mentioned studies will help to identify particular issues and mechanisms of how framing was applied in data visualizations during the 2016 presidential campaign by focusing on particular political issues.

The Implication of Framing in Political Communications

It was not only Bennett and Iyengar (2008) who marked out the possibility of using media framing in the political process. Entman (2007) concentrated on framing, priming, agenda-setting and the implication of these effects in political issues and democracy. The author divided bias into three categories: distortion or falsification of reality (distortion bias), favoring one side in a political conflict (content bias), and motivations of journalists (decision-making bias) (p. 163).

The work of Valkenburg, Semetko & De Vreese (1999) demonstrated that journalists unavoidably structure their representations of political events to make them understandable to a broad audience. Journalists use four methods: pointing out the conflict between parties or individuals (conflict frame); focusing on emotions (human-interest frame); writing about responsibility, sometimes blaming certain political figures (responsibility frame); and emphasizing economic issues (economic-consequences frame) (p. 551).

Iyengar (1987) presented an experiment in which U.S. residents participated in a series of studies that were described as investigations into selective perception of television news. The research demonstrated people's concerns about national affairs, especially in issues such as terrorism, poverty and unemployment. It may be of interest to observe whether similar patterns took place in data visualization during the last U.S. presidential campaign.

In another study, Iyengar (1989) expanded his analysis to focus on other concepts of media effect, such as causal responsibility and treatment responsibility: "Casual responsibility focuses on the origin of the problem, while treatment responsibility focuses on who or what has the power either to alleviate or to forestall alleviation of the problem" (p. 879). Iyengar analyzed additional factors that could influence people's perceptions about the issues. His model contained four modules: poverty, racial inequality, crime and terrorism (p. 880). When all four modules were taken into account, opinions were affected by various factors such as partisanship, liberal/conservative orientation or factual knowledge, as well as socioeconomic status, level of education and race (p. 888-889). Nelson, Lecheler, Schuck & De Vreese (2012) considered the impact of framing as a communicator's attempt to influence public opinion by guiding impressions. There is an extensive field of research in the sphere of mass-media framing, but the subject of many studies is textual or television representation. Data visuals, which have become popular in mass-media reporting, are not investigated enough in connection with framing. That is why the main purpose of this research is to use the previous theoretical foundation to inspect the role of data visualization in media framing. To make it possible to effectively investigate the patterns of data-visualization framing, it is necessary to cover the main concepts of data visualization.

Understanding the Main Principles of Data Visualization

One of the patriarchs of the studies in the sphere of data visualization is Edward Tufte, with his book "Visual Display of Quantitative Information". Tufte wrote about the philosophical meaning of design but at the same time provided many valuable points for the practical work of graphic creators. According to the author's opinion, it is necessary to study graphical and statistical components, but also how to present information through words, numbers and pictures. He described how graphics can be demonstrated in different forms using mapping, time and space by presenting various examples from the history of information graphics. Among the principles of graphical excellence, Tufte mentioned a matter of substance, statistics and design. He stated that it "consists of complex ideas communicated with clarity, precision and efficiency" (Tufte, 1983, p. 51). Throughout the book, the author tells us that decorative elements of the graphics are absolutely useless, or, as he called them, "chartjunk". These give a designer an opportunity to exercise artistic skill but don't contain any valuable meaning. Tufte supported minimalism in graphic design: representation of maximum information with the minimum details. He also mentioned that very often graphic designers underestimate their audience: It is better not

to oversimplify and distort data to make it easier to digest, but to show exact numbers and choose the right way to do it – this is a principal skill and the meaning of graphical representation.

He provided some rules for graphic designers to improve their everyday routine graphics. Among them there are right format and design; the correlation of words, numbers and drawing; precise reflection of balance and proportions; sense of relevant scale; demonstration of complexity of detail; narrative quality and story to tell about data; professional manner with carefully created technical details in production; and avoidance of meaningless decoration and chartjunk.

Data journalism plays a significant role in serving the democratic goals of society. The Associated Press published the insight and thoughts of Julia Angwin of ProPublica, an independent, nonprofit newsroom that produces investigative and data stories in the public interest. Stories with data resonate more strongly with audiences because people see the depth of the story and expertise with which it is written (Associated Press, 2016). It can explain why journalists and audiences relied the most on data analysis and visualization during the presidential campaign.

In the work "Letting the data speak", Boyles and Meyer demonstrated how data journalists perceive their social responsibility within the democratic conversation and building dialogue with the audience (2016). Today, data journalists have a great responsibility as translators of big data and the technical language of datasets. Reporters are formulating clear and concise narratives with interactive and visual elements, and organizing debates in the social sphere. As it happens with traditional reporting, using new forms of data language leads to the issue of salience of information which is presented with it. Data graphics and interactives allow readers not only to get involved in the social processes and discussions, but also to build interpretations through different manipulations with data: "Activating audiences in conversation around data journalism has been most successful when focusing on highly personalized, local-level, policy-driven news products" (Boyles & Meyer, 2016, p. 949).

The role of data visualization in journalism is discussed in the article "Narrative Visualization: Telling Stories with Data" by Segel and Heer, who mention the meaning of emerging data visualization in storytelling, which is particularly important for online media. They are developing the basics of understanding data visualization and pointing out different ways of integrating it into journalistic narratives, using design space for the new element, as well as the ways for balancing interactive and traditional written content (Segel & Heer, 2010). They mentioned that journalists today have to acquire new skills in computer science, statistics and artistic design, as well as traditional storytelling. They identified and investigated various techniques for telling stories with infographics; they also analyzed five case studies and 58 examples that provide valuable context for understanding interactivity and the use of graphical techniques.

It is mentioned that where infographics are used, framing can appear:

Of course, not all elements in a scene are of equal importance throughout a story, and so authors often manipulate a scene to direct attention to a point of interest. Psychologists have extensively studied phenomena of visual salience, showing that outliers among visual features such as color, size, and orientation preferentially attract one's attention (Segel & Heer, 2010, p. 1140).

11

For this research, data-visualization examples were taken from leading media organizations such as The New York Times, The Guardian, The Financial Times, The Washington Post and Slate. The main components of data visualizations are identified as genre (visual narrative type), visual narrative tactics (visual structuring, highlighting and transition guidance) and narrative structure tactics (ordering, interactivity and messaging). Research explored modern platforms such as GeoTime Stories and Tableau, which help to enrich storytelling with data visualizations. The authors focused on the components of infographic representation but didn't cover the cognitive and emotional experience of the audience. Considering this study, the similar strategy for analysis with the abovementioned components was used for the present analysis of data visualizations during the 2016 presidential campaign.

The complete definition of informational graphics was stated in "How Users Perceive and Use Interactive Information Graphics: An Exploratory Study" by Burmester, Mast, Tille and Weber (2010), who wrote that interactive information graphics are "a hybrid presentation of different modalities: text, photos, audios, videos, charts, maps, graphs, illustrations are combined; the interplay of these is more than the sum of the parts. In addition to multimodality, interactivity and hypertextuality are main characteristics of interactive information graphics" (p. 361). In this research, it is possible to find not only the main information about the principles of informational graphics, but also an attempt to understand the process of the audience's perception and experience. Researchers used an eye-tracking system to monitor the audience's attention during the experiment with users. The research found that the main problems identified in the perception process were connected with website usability, design of the text (size, font), length of the sentences and the amount of text. The main challenge for journalists and designers nowadays is "to combine the 'old' media (text, photos, videos, audios, maps, diagrams) with the concept of interactivity in an effective way so users will understand the message easily" (p. 367).

Studying the Implementation of Infographic Techniques

Data journalism is often referred to by various terms, such as data-driven journalism, computational journalism or computer-assisted reporting. In the research by Weber and Rall "Data Visualization in Online Journalism and Its Implications for the Production Process", data-driven journalism is defined as a part of computational journalism and seen "as a workflow that consists of: digging deep into data, analyzing and filtering the found data, visualizing it, and forming a story" (2012, p. 349). A visual component of data-driven journalism is called an "interactive information graphic", "interactive feature" or "interactive graphic" (p. 349). According to their definition, an interactive information graphic is "a visual representation of information or knowledge with the following elements: verbal elements (text, audio, typography) and visual elements (photo, illustration, diagram, map, symbol, icon, pictogram, video, moving image) are combined in such a way that they create a new hybrid form" (p. 349).

Data visualization was studied as a phenomenon within journalism by Khan and Khan (2011), who described and classified standard visualization techniques. Their research deals with many definitions and aspects of visualization, why visualization is used, different steps of the visualization process, and problems that are confronted in visualization. They also explored the categorization of visualization techniques on the basis of distinct perspective, common data and information, basic interactive methods for

visualization, interactivity process, and the scope of visualization in different fields of research.

Interactive Graphics Development (IGD) framework, a "theoretical model for researchers studying the effectiveness of interactive graphics" (George-Palilonis & Spilman, 2011, p. 167) was created by George-Palilonis and Spilman in their work "Interactive Graphics Development: A framework for studying innovative visual story forms". It was formed after exploration of the data-visualization work of American newsrooms such as The New York Times, The South Florida Sun-Sentinel, The Detroit News, The Los Angeles Times, USA Today and MSNBC.com. The process of analysis contains two phases: story evaluation and story-form selection, which cover both a theoretical and practical analysis of infographic creation. The first phase is focused on news values, objectives and audience; the second consists of types of interactive presentation: "In summary, the IGD framework suggests that editors judging the interactive potential of a story should first consider and rate the traditional news characteristics of significance, prominence, proximity, conflict, novelty/human interest, and timeliness, and select stories that rate highly in these areas" (p. 170). As researchers explained, data visualization has to be used to help journalists exercise their societal roles, such as improving community engagement and citizen participation in the work of government (p. 170). They used McNair's classification of journalists' roles in society that can use implementation of interactive visual content, such as the journalist as a "supplier of information," a "resource for, support to and often participant in public life and political debate," and a "medium of education, enlightenment and entertainment" (McNair, 2005, p. 28 as cited in George-Palilonis & Spilman, 2011).

To analyze how data visualization can frame a presidential campaign, it is necessary to connect the studies of the framing effect with research of data visualization to find out the forms in which data visuals can play a framing role.

Framing Effect in Data Visualization

There is a limited amount of research about the framing effect resulting from the use of infographic visual content. "Visualization Rhetoric: Framing Effects in Narrative Visualization" by Hullman and Diakopoulos (2011) is devoted to this particular phenomenon and demonstrates how narrative visualization can affect readers' interpretations of information by using the data, visual representation, textual annotations and interactivity. This research provides a picture of forms and types of techniques in data visualization and how they are connected to readers' perception and interaction in a sociocultural context. Bias, rhetoric and framing are discussed in the context of studying audience interpretation and considering interconnections with representational, individual and social forces (p. 2232). Bias is defined as "a systematic error introduced into sampling or testing by selecting or encouraging one outcome or answer over others". Framing is interpreted as a "construct". Rhetoric is explained as "persuasion as a result of the implicit motivation of the speaker to gain other adherents to a preconceived view or conclusion" (p. 2232): "Information representations can influence interpretation in diverse ways, such as by presenting a preliminary statistic before a decision, or by manipulating the anchor points on a survey scale" (p.2232). They also described how semiotics theory applies the analysis of visualization to help understand the way information visualizations incorporate meaning by using signs (word, color choice, visual icon, etc.). It helps in the interpretation of the visual, textual, cultural and perceptual codes in data visualization. A sample of 51

data visuals from mass-media organizations such as The New York Times, BBC, news magazines, local news and political outlets were collected and analyzed for framing and bias techniques. Omission techniques (oversimplifying information, filtering), metonymy techniques and averaging techniques (categorizing, binning, aggregating, textual and visual summaries), obscuring ("noise", third dimension, sizing transformation), visual metaphor and metonymy (color mapping such as red and blue for political parties' representation), use of contrast, classification (by size, color), and redundancy were identified as the main techniques for framing media issues. There are also linguistic-based tools for framing, such as typographic emphases (bolding, italicizing) and irony (rhetorical questions and headlines).

Considering the previous research basis, multiple research questions appear appropriate for analysis:

• Research Question 1: What is the role of data visualization in framing of coverage of the 2016 U.S. presidential campaign?

 Research Question 2: How did the structure of data departments influence their work during the coverage of the 2016 U.S. presidential campaign?

 Research Question 3: What challenges and limitations do journalists struggle with in creating the data visualizations?

 Research Question 4: What measures can be implemented in newsrooms' activities to improve data-visualization reporting?

16

Chapter 3. Methodology

Introduction

This study consists of two parts of analysis: textual/visual analysis of data visualization examples during the coverage of the 2016 U.S. presidential campaign and interviews with the journalists who were involved in infographics production. It led to a triangulation of the research, a combination of multiple methodological practices, empirical evidence, and perspectives that will be used to increase the quality and validity of the research (Denzin & Lincoln, 2007). Triangulation helped to collect more data from different sources and "strengthen study" (Yinn, 2011, p. 57). According to Maxwell's seven-point checklist for expanding the validity of research, it was necessary to have an intensive long-term involvement, analysis of rich data, and search for evidence and cases (2009, pp. 244–245). These aspects were considered and implemented during the in-depth interviews and textual and visual analysis.

Opinions were collected from graphics editors and data journalists at two major newsrooms of public interest: The New York Times and The Washington Post, where they were asked about their experiences while working with data visualization during the presidential campaign.

A primary goal of the research is to identify the purposes and problems that newsrooms are dealing with in visualizing the data, as well as to analyze how framing can be displayed in the graphic content. A secondary purpose of the research is to collect opinions of the participants about how the journalistic work can be improved and what skills can be implemented by journalists to improve their data-visualization reporting.

Research Design

This study was conducted following the qualitative research design because qualitative methods allow for the collection of information through interviews and descriptive textual and visual analysis about professional experience and practices from the people who are working in the field. Clifford, Christians & Carey (1981) pointed out that the main aim of qualitative research is to study the creative process of humans and interpretations to understand the meaning that people use to guide their activities.

The research questions for this thesis were evolved from a larger idea: How does this form of election coverage help journalists convey their narratives, and are there ways for development and improvement of data-visualization reporting?

The questions for this research were developed to focus on the method of data visualization reporting and collect opinions from the professionals in graphics and data departments to investigate their experiences and the structure of their teams. The primary research question is broader than the others:

• What is the role of data visualization in framing of coverage of the 2016

U.S. presidential campaign?

The secondary research questions are more detailed and are focused on various aspects of reporting using data visualization:

- How did the structure of data departments influence their work during the coverage of the 2016 U.S. presidential campaign?
- What challenges and limitations do journalists struggle with in creating the

data visualizations?

 What measures can be implemented in newsrooms' activities to improve data-visualization reporting?

The Role of Textual and Visual Analysis

For this research, visual and textual analysis was conducted of graphics from The New York Times and The Washington Post in order to identify the main patterns during coverage of the presidential campaign. Visual analysis was used for analyzing various design elements such as type of graphics, colors, fonts and how they helped to frame issues during the campaign. With textual analysis, the researcher was able to identify the main issues and frames that were covered and considered important for the audience. Using textual analysis helped to identify "various linguistic categories" which "can be seen to come together as a coherent structure, a text with a message to be interpreted" (Jankowski, 1991, p. 35):

Through language, reality becomes social; through genres, social reality becomes the object of specific forms of story-telling, argument, and action. Genre may be the analytical level where social-scientific and humanistic modes of inquiry can be said to converge, with implications for both theory and methodology. (p. 37)

In regards to visual analysis, it is important to note that visual communication contains a complex process of encoding and decoding in different forms of representation and perception. One of the main tasks is to explore this representation in order to see the process of constructing various modes of socialization through visual communication (Jankowski, 1991). For this research, data-visualization examples from The New York Times and The Washington Post were collected and analyzed during the period of six months before and one month after the presidential election. These particular newsrooms were chosen because they both provide national coverage and produce content of general interest. The decision about the time period for analysis was made in order to see the patterns in the data visualizations of the presidential campaign during the period of six months before the election as the most active phase of the campaign. The period of one month after the election was chosen because it helped to explore whether and how data visualizations were affected by the outcome of the election.

The textual and visual analysis helped to identify certain contexts for using data visualization in the coverage of the presidential campaign and discover particular patterns and cases. Data visualizations can change people's perceptions of topics and move readers' focus on toward particular aspects of topics. To understand this phenomenon, it is important to analyze media content that includes data-visualization examples. For textual and visual analysis, there was a need to explore graphics and data-visualization examples of coverage from the 2016 presidential campaign. In the process of the textual and visual analysis, various patterns and specific topics and cases were analyzed and various methods of how journalists implemented data-visualization techniques were identified.

The Implementation of the Interview as a Method

In-depth interviews with semi-structured questions were chosen as a method to improve the quality of the data and to provide an "extent of spontaneous, rich, specific, and relevant answers from the interviewee" (Roulston, 2010, p. 202). A significant positive aspect of the interview is that it facilitates the conveyance of rich content: "The interview is 'self-communicating' – it is a story contained in itself that hardly requires much extra descriptions and explanations" (Kvale, 1996, p. 145). The interviews provided information

about the professional experience of data journalists and editors and their opinions on the role of data visualization, problems and limitations which they were experiencing while working with it.

Defining the Sample: Participant Selection and Information

Interviews were conducted with seven journalists from the two newsrooms that produced infographics and data visualizations for the 2016 presidential campaign. The number of interviews was defined in the process according to whether sufficient data to identify patterns has been reached.

The number of journalists from the two newsrooms interviewed for this research is small, but it covered multiple perspectives and allowed for identification of patterns in the participants' responses. There are differing opinions about the appropriate number of interviews for academic research. Tracy (2010) pointed out that with too many interviews, the data can be difficult to analyze and to interpret. Tracy made an assumption that five to eight interviews is a minimum number to be "pedagogically valuable" (p. 138). In this research, data collected during the seven interviews is adequate for observing patterns and obtaining valuable knowledge. The researcher also considered the saturation point for the research and found the collected data to be satisfactory for completing of the analysis.

Journalists were contacted by e-mail to arrange times and dates for the interviews. Their e-mail addresses were collected through Internet searches and were found on their personal websites, social networks and the websites of their news organizations.

The snowball sampling method was also used to recruit participants. The researcher asked interview subjects for recommendations and contacts of colleagues who could be appropriate for the research (Tracy, 2010). This method was found as an effective way to

21

reach the subjects; however, it became apparent during this process that it is difficult to reach people remotely. People who are working in the news industry are constantly busy, which poses an obstacle to getting a response from them. Also, the small number of people working with data visualization in the selected newsrooms made finding more interviewees a challenge. As a result, a total of seven people were interviewed, and description of the seven interview subjects is listed in Appendix B.

In the first e-mail, potential respondents were provided with information about the study, the Consent Form as required by the IRB, and a list of interview questions to help interviewees to prepare for the interview.

Three of the participants (Respondent No. 2, Respondent No. 3 and Respondent No. 4) were interviewed personally during the National Institute for Computer-Assisted Reporting Conference (NICAR), which gathers data journalists every year for a series of panels and master classes. Three interviews (with Respondent No. 1, Respondent No. 5 and Respondent No. 6) were conducted by Skype, allowing the interviewee and interviewer to hear and see each other as in an in-person conversation. According to Paul Hanna, "using Skype as a research medium can allow the researcher to reap the well-documented benefits of traditional face-to-face interviews in qualitative research, while also benefiting from the aspects Holt suggests telephone interviews bring to such research" (Hanna, 2012, p. 240). As Hanna pointed out, Skype provides not only synchronous interaction between the researcher and participant but also the ability to save visual and interpersonal aspects of the interaction (Evans et al., 2008 as cited in Hanna, 2012): "Further, using technologies such as Skype, the researcher can easily record both the visual and audio interaction of the interview through simple software downloaded onto their workstation" (Hanna, 2012, p.

242). One interview (Respondent No. 7) was conducted by e-mail at the subject's request because of time constraints. Using Skype allowed interviews to be conducted without additional costs. For the interviews, the researcher used semi-structured questions because they offered flexibility in allowing questions to be adapted and changed depending on the respondents' answers. This type of interview can also generate qualitative data through the use of open-ended questions, which allow respondents to talk without restriction on various issues. This method helps to develop a real sense of the person's understanding of a situation, ask for clarifications, and guide the interview in directions that will add to the validity of the research.

Interview subjects needed to be employed either at The New York Times or The Washington Post and work with data visualization reporting in their professional work. Among the interviewees, there are graphics editors and journalists from both newsrooms.

In forming the question list for the interview, since focus on data journalism has only recently developed, the researcher started with questions about the structure of the data desks and departments in each newsroom and the skills that journalists need to have to be able to do their job. Other questions were related to the use of data visualization, specifically for the coverage of 2016 presidential campaign; possible challenges, trends and frames in data visualization which they have noticed; the process of working on data visualizations; and their impressions about the effects that data visualization has on the audience. The researcher also found it valuable to explore possible recommendations from professionals about potential improvements to the process of reporting with data visualization and advice for coverage of future elections. The approval form for conducting this study was received from the Institutional Review Board (IRB) (see Appendix A).

Data Collection Process

For this research, data was collected using methods such as visual and textual analysis and semi-structured interviews. For the first method, data was collected by gathering the examples from The New York Times and The Washington Post during the period of six months before election and one month after. Data was collected through the websites of the news organizations as well as the Twitter accounts of the data desks. Examples that were related to coverage of the presidential campaign were selected. As a result, 136 examples were chosen for analysis: 97 published before the election and 39 after the election.

The primary manner recruiting interview subjects was by e-mail. The researcher sent an introductory e-mail directly to the prospective subject. The first e-mail also contained a request for recommendations of colleagues who could also participate in the research. After receiving consent from the participant, the researcher sent the second email in which the most convenient date and time were negotiated. The first e-mail also stated a guarantee of anonymity for all the participants. Appendix C contains the e-mail protocol.

For the three interviews that were conducted at the NICAR Conference, participants were contacted by e-mail, and the date and time were negotiated. After that, meetings were set and interviews took place at the Conference on a one-on-one basis.

One interview was conducted by e-mail. In this case questions, were sent and answers were received using this mode of communication. In each case, permission for recording was stated in the e-mail and was obtained before the interview. The length of the interviews conducted by Skype and at the NICAR Conference varied, but averaged about 37 minutes. The shortest interview was with Subject No. 2, lasting 23 minutes. The longest interview was with the Subject No. 4, lasting 64 minutes.

The list of questions for the interviews was predetermined, but interviewees were given the opportunity to speak about additional issues if they considered them important to note for the current study. The list of questions can be found on Appendix D.

The concept of "informant interviews" was used for these interviews, as Tracy (2010) stated that participants are the professionals in their sphere "who are experienced and savvy in the scene, can articulate stories and explanations that others would not, and are especially friendly and open to providing information" (pp. 140-141).

Data Analysis

For the first part of the research, graphic and visual analysis, each example was analyzed by specific categories, such as type of graphic, colors used, font, use of text, video, audio, tone that was used in text, intention of the graphic, topic, interactivity, key words and purpose. After analyzing the criteria, the results were organized and codified. The most used and uncommon examples were identified and stated in this research, along with the main topics and trends that were found in visual examples.

In the second part of the research, seven interviews were transcribed by the researcher. The online platform oTranscribe was used to facilitate the process. Using this platform, it is possible to slow the speed of recording, which makes transcribing it more

efficient. It was especially useful for the researcher, who is not a native English speaker. During the process, anonymity was ensured by the researcher.

The data was collected and analyzed with a focus on particular aspects for efficient coding of information. After the textual/visual analysis, particular patterns and frames were identified and coded. After the in-depth interviews, it was possible to enrich this data with expanded answers from various practitioners within the journalism industry. Using both methods made it possible to formulate answers for the stated research questions.

For data analysis, an open coding method was used. Using this method, the researcher can shape the basic concepts from the data into a simple form and develop categories (Strauss & Corbin, 1990). The method of open coding is based on patterns found during the analysis. The coding formulation consisted of examining each example of data visualization in terms of type of infographic (map, chart, line graph, interactive or static), visual components (numbers, illustrations, colors, forms), textual explanations, purpose of the infographic (ability to state the possible purpose after analysis of data visualizations that can be used during the in-depth interviews), and type of frame that the infographic contains (particular social phenomena).

Coding of the interviews was based on the main topics of the interview, as well as on the issues that were stated in the open questions.

Validity and Challenges

For this research, an inductive approach was used to create concepts during the extensive analysis of the collected data. To ensure that the findings possess high validity, it was necessary to conduct visual and textual analysis to identify patterns and elements of data framing. In the next stage, interviews with journalists were conducted to increase the

accuracy and validity of the previous method, as well as to obtain new data for possible revelations about additional concepts. An elicitation technique was employed while conducting in-depth interviews, in which journalists were asked to describe the process of producing data visualizations using examples of their own work. It helped to increase the quality of generalization, which is a desirable goal for qualitative social researchers, according to Clive Seale (1999). Generalization is important because it "involves the extent to which you can make some form of wider claim on the basis of your research and analysis" (Mason, 2002, p.41).

Limitations of the Study

Among limitations, there is an assumption that generalization is difficult to make in a broader sense. The work of two of the best newsrooms in the industry was analyzed, and the results cannot easily be translated to the entire mass media industry; however, they can help other organizations in the news industry improve their practices.

The biases of the interviewees as well as the researcher's subjectivity have to be considered as well. As a limitation of the qualitative studies, objectivity could be decreased since the research collects personal experiences and their implementations.

Chapter Summary

This chapter described the research design of the study, including using the framing theory and triangulation. The research questions as well as the main procedures such as methods of textual/visual analysis and the interview were presented. The process of data collection and analysis is explained. Limitations of the research were identified.
Chapter 4. Findings

Introduction

This study consists of two parts of analysis: textual/visual analysis of data visualization examples during the coverage of the 2016 U.S. presidential campaign and interviews with the journalists who were involved in infographics production. The textual/visual analysis was conducted before the interviews. It allowed examples of graphics to be collected and analyzed before talking to the people who participated in their production. The textual and visual analysis helped to identify certain contexts for using data visualization in the coverage of the presidential campaign and to discover particular patterns and cases.

For this research, it was necessary to conduct visual and textual analysis of graphics from The New York Times and The Washington Post to identify the main patterns during the coverage of the presidential campaign. Visual analysis was used for analyzing various design elements such as type of graphics, colors, fonts and how they helped to frame issues during the campaign. With textual analysis the researcher was able to identify the main topics that were covered and considered newsworthy for the audience.

Textual/Visual Analysis: Data coding

Data was collected by gathering the examples from The New York Times and The Washington Post during the period of six months before the election and one month after it through the websites of the news organizations as well as through the Twitter accounts of the data desks. Examples that were related to the coverage of the presidential campaign were selected. As a result, 136 examples were selected for analysis: 97 from before the election and 39 after it.

The data was collected and analyzed with a focus on particular aspects for the efficient coding of information. After the textual/visual analysis, particular patterns and frames were identified and coded. For data analysis, the open coding method was used. Open coding is based on patterns found during the analysis. The coding formulation consisted of examining each example of data visualization in terms of type of infographic (map, chart, line graph, interactive or static), visual components (numbers, illustrations, colors and forms), textual explanations, purpose of the infographic (ability to state the possible purpose after analysis of data visualizations that could be used during the in-depth interviews), and type of frame that the infographic contains (particular social phenomena).

Data was broken into categories, codes were developed, and themes were generated and used in the final findings. Microanalysis was used to create primary concepts and categories (Strauss and Corbin, 1990). Using microanalysis, the researcher analyzed the data collected from examples, interpreted the data and formed connections between observation and interpretation. Microanalysis was conducted during open coding to analyze data and produce new data for the findings.

Here is the table with descriptive codes and sub-codes for the analysis of visual examples.

Code	Sub-codes	Explanation
Visual narrative type	Illustration,	The visual narrative type is
	horizontal/vertical bar, line	a description of the type of
	graph, map, cartogram,	visual that was used for the
	bubbles, picture, scatter	piece.
	plot, histogram, interactive,	

Table 1. Descriptive codes and sub-codes for textual and visual analysis

	animation, treemap, circle graph, table, bullet graph, 3D model, area chart, quiz	
Colors	Red, blue, purple, green, yellow, orange, grey	Color is analyzed to see the possible connotations and meanings with political parties
Font	Classic, highlighted, bold	The style of the font was analyzed to illustrate how text can be used as an element for data visualization
Text	Yes, No	The presence of text was taken into account to draw conclusions about using visual elements with traditional textual reporting
Video	Yes, No	This aspect was checked in order to find out if video is used in data visualization
Audio	Yes, No	This component was analyzed to figure out if audio is used in data visualization
Tone	Positive, negative, neutral	The tone of the piece was used to analyze the character of the presentation of information
Intention	To present polls, facts, analysis, forecast, opinion, data, quotes, fact-checking, entertainment	The intention showed the main focus of the piece
General issues	Election issues, legislative issues, government, political issues, economic issues, behavior, social issues, security, immigration, foreign issues	The general issues code was used to categorize the type of specific issues that were covered
Interactivity	Yes, no	This code illustrated the presence of interactivity
Key words	Multiple sub-codes in different pieces	The key words were analyzed not by their usage count but by presenting the main meaning of the example. They were

				analy frame	zed to fin the ex	d possible amples
Purpose of the visual	Multiple different p	sub-codes ieces	in	The more narra	purpose details abo tive of the o	contained ut the main example

Textual/Visual Analysis: Findings

The codes were analyzed and the most used and unique examples were identified and stated in the findings. The researcher found many traditional and simple methods for infographics as well as some experimental examples in the content produced by both newsrooms. Some were used for elicitation to ask the journalists and editors who were working on them about the production process and how existing frames were chosen. All the examples were gathered with regards to the period in which they were chosen and the newsroom in which they were produced. After that, the most recurring patterns in design elements as well as in narrative type were identified and stated in the findings.

Analysis of Graphics at The New York Times Before the 2016 Election

For analysis of data visualizations in the pre-election period, the researcher looked through the content during the six months before election day, as this was the most active phase of the campaign, and one month after the election to see how the topics and graphics had changed. As a result, 131 data visualization examples from The New York Times and The Washington Post were analyzed for this research. 46 examples were included in the analysis of graphics produced by The New York Times before the 2016 election.

The system of coding for analysis consisted of visual narrative type (map, bars, illustration, etc.), colors that were used (red, blue, purple and others), specific features of font (bold, highlighted), size of text (large or small), presence of video or audio content on the page, tone of the content (positive, negative or neutral), the intention (present facts,

opinion, entertainment), the issue presented in the content (election, political or economic issues), interactivity, key words and the main purpose of the piece.

Considering the results of visual analysis of The New York Times graphics before the election, the most popular types of graphics during the coverage of presidential campaign were tables, vertical/horizontal bars, line graphs, geographic maps, pictures and illustrations.



Figure 1. The types of infographics used by The New York Times before election. This figure illustrates the most frequently used types of visuals during the coverage of presidential campaign.

The most popular type of visualization was horizontal and vertical bars, that were used in various forms and colors such as stacked bar and side-by-side bar. There are several examples, such as a vertical side-by-side bar that illustrated that Donald Trump was supported most strongly by white people without college degree who live mostly in rural areas, according to the polls (Fessenden, 2016). The vertical stacked bar was used to illustrate changes in voter ID laws in different states and how they had changed since 2000 (Lee, 2016). The horizontal side-by-side bar was published to demonstrate the results of

early voting (Patel, 2016). Vertical and horizontal bars remain a meaningful tool for clear presentation of complex information, especially when there is a purpose to demonstrate a relationship between sets of parameters, such as actors and related numbers. In the examples, it is possible to see that bars helped to illustrate the various groups of voters and their proportional size.

Bar charts are easy for our eyes to read. Our eyes compare the end points of the bars, so it is easy to see quickly which category is the biggest, which is the smallest, and also the incremental difference between categories. (Nussbaumer, 2015, p. 50)

Stacked bars allow readers to compare numbers in various categories and also see several components within a category. Side-by-side bars present an ability to show several objects within one category and compare them with each other.

The New York Times actively used tables, line graphs and maps – these visuals were used numerously on the website.

Tables are used for presenting a range of data, usually with colors for a more effective visual presentation, such as an example that demonstrates the primary results and which candidate won the nomination from the major parties (Andrews, Bennett & Parlapiano, 2016).

Tables present an opportunity to illustrate several categories and items of data in a particular order and help the readers identify and compare information:

Tables interact with our verbal system, which means that we *read* them. When I have a table in front of me, I typically have my index finger out: I'm reading across rows and down columns or I'm comparing values. Tables are great for just that—communicating to a mixed audience whose members will each look for their

particular row of interest. If you need to communicate multiple different units of measure, this is also typically easier with a table than a graph (Nussbaumer, 2015, p. 40).

The New York Times demonstrated how it is possible to present a wide range of information, such as a Senate Election Forecast, in one table using color and illustration elements to show the state-by-state predicted outcomes of the election by multiple polls (Aisch, Cox, Katz, Pearce, Quealy, Bostock & Carter, 2016). In this article, heatmap used, in addition to the numbers, colored cells to "convey the relative magnitude of the numbers" (Nussbaumer, 2015, p. 42). According to Nussbaumer, to make it easier for a human brain to process information, creators of visuals can use color saturation to provide visual cues and help readers to identify the points of interest faster (Nussbaumer, 2015, p. 42).

Line graphs are the most commonly presented graphs in the mass media and help to visualize continuous data. A line graph is often applied in cases where there is a comparison over time and is used to show many periods with non-cyclical data. The New York Times example of a line graph presents how the number of early voters has changed over two decades (Patel, 2016).

The line graph can also show several series of data, which helps to compare numbers from two parties on the same graph, as it is made in the visual example showing the exit poll results and the number of white voters with no college degree who voted for either the Democratic or Republican candidate (Fessenden, 2016).

Line graphs in most of the cases were used to show changes across time, often called time lines or time series. According to Chen, it is useful "to differentiate between value measurements at particular time points and summary measurements over a period" (Chen, Hrdle & Unwin, 2008, p. 74).

Maps are frequently used to describe election and poll results. This type of visual is used to demonstrate a geographic category of data. The New York Times presented a chronological map to illustrate how the divide between Republicans and Democrats in America has grown in recent years (Aisch, Pearce & Yourish, 2016). In this example, information was broken down by counties; however, in the case of the U.S. election, the size of a geographical territory does not correlate with the territories' electoral votes, which are determined by population:

The main problems to be solved lie in fact that areas do not reflect the relative importance of regions (e.g. Montana has fewer people than New York City but is much bigger) and spatial distance is not directly associated with similarity or nearness (e.g. where countries are divided by natural borders, like mountain ranges). (Chen, Hrdle & Unwin, 2008, p. 75).

In certain examples information is presented with geographical category by states and does not accurately represent the situation, as in an example that covers the percentage of white people with no college degree who voted for the Republican candidate in the 2012 election. These voters were considered Donald Trump's main voter base (Cohn, 2016).

The high frequency of usage of visuals such as lines, bars and maps can be explained by the fact that they are the most visually effective and are not overwhelmed by other elements. They help readers dive into the data and quickly understand it. Despite this fact, The New York Times has produced more sophisticated graphics that are interactive. One example is an interactive dial that showed the chances of victory during the presidential race in real time (Aisch, Cohn, Cox, Katz, Pearce & Quealy, 2016). There is also an example of an interactive diagram that was published before election day and gave users an opportunity to make projections about candidates' potential path to victory (Aisch, Cox, Katz, Pearce, Quealy, Bostock & Carter, 2016). Another example of data visualization was created during the 2012 presidential race between Barack Obama and Mitt Romney and was used again for 2016 campaign. It combines methods of animation and interaction, which can entertain and help readers be involved in the forecasting process. One of the authors of this infographic, Mike Bostock, said that this visual gives people "a good overview of the full tree of paths" (Heller & Landers, 2014, p. 46).

The New York Times also used elements of animation, which work well for engagement, as motion catches the eyes of the readers, but usually it is more time consuming and difficult to produce (Simmon, 2013). In one case, animation was used to analyze Trump's maneuvers to reduce his tax liabilities. In this case, animation did not present quantitative information, but followed Trump's cash flows between various actors in his business, which makes the information easy to absorb (Dance & Byrd, 2016).

The other component that was analyzed in the examples was color. Color is "one of the most effective ways of displaying data" (Chen, Hrdle & Unwin, 2008, p. 68). Colors often have specific associations in U.S. elections: Red is mostly used for the Republicans and blue for Democrats, something that was illustrated in numerous data visualization examples across many years. However, the same colors can have different connotations, such as when red is for danger or losses and blue is for more neutral issues, along with other colors (Chen, Hrdle & Unwin, 2008, p. 69).



Figure 2. Packed bubbles demonstrate colors that were used in The New York Times visuals before election.

Text can be a visualization tool. For instance, it can be highlighted, or the size can be bigger than average or bolded, as in fact checks, where different statements from the candidates were analyzed and labeled with highlighted text depending on how true or false the statement was (The New York Times, 2016).

Another observation is that there is a tendency to combine traditional reporting with the graphics. In most cases, graphics are an additional element to support facts from form the main text. Text can also be a short explanation of a data visualization. Videos are rarely published with graphics: In The New York Times, graphics videos were used in coverage of debates and showed the candidates' answers on the most complex issues (Gold, Huang, Parshina-Kottas & Yourish, 2016). Audio was not included in data visualization examples.

An analysis of the purposes of the data visualization examples showed that several types of purposes existed, such as presenting the results of polls and forecasts, facts, opinions, analysis, quotes and live results. Most of the examples were considered neutral. Coverage with negative connotations was found mostly about the growing divide between Republicans and Democrats, candidates' statements that were at odds with the U.S. Constitution, Trump's tax maneuvers, and how Trump was unpopular among some members of Republican Party. Most of the topics were connected with the election, legal issues, economic issues such as unemployment, candidates' behavior, immigration and healthcare.

Among the most popular topics in The New York Times graphics during the six months before the election were poll results and forecasts (in the majority of them, Clinton had higher chances to win), Donald Trump's business ventures, the lack of transparency in his tax activities, the lack of Republican Party support for Donald Trump, the political divide in America, white voters with no college degree as the main Republican electorate, the FBI investigation of Hillary Clinton e-mails, donations and campaign finance, as well as voting laws (such as voter ID laws and laws that prevented felons from voting).

Analysis of Graphics at The New York Times After the 2016 Election

21 examples were included in the analysis of graphics produced by The New York Times after the 2016 election. Among the data visualization types after election, there are maps, tables, lines, bullet graph, vertical/horizontal graphs, illustration, charts and 3D models.



Figure 3. Treemap of visualization types used by The New York Times in the month after the election.

The most popular type of data visualization for The New York Times during the month after the election was mapping. The purpose of this type of graphic was to illustrate the results of election. The New York Times maps presented different breakdowns of election data, such as by state, county, size of lead and change from 2012 (The New York Times, 2016).

Mapping was used for coverage of possible outcomes of certain issues because of Trump's election, such as impact of new climate change policies (Strauss, 2016). The map shows what could happen if new climate change policies were implemented and the United States withdrew from the Paris Agreement. There is also a successful attempt to create an election map without using red and blue colors but only with the colors of natural surface (Wallace, 2016).

Covering the political issues after election, The New York Times used animation to show the number of immigrants who could be deported if Donald Trump fulfills his campaign promises. Animation played a significant role in illustrating the numbers and scale of the issue (Park & Griggs, 2016). Among colors, blue and red were the most frequently used after election as well. In most of the cases, graphics were an additional element in a long-form article. No video or audio was used with graphics during the month after the election. More than half of publications with graphics were considered negative. These examples covered the divide between Republicans and Democrats, the protests against Trump's election, the potential consequences of the decisions and actions of President-elect Trump according to his promises during the campaign, such as his plan to reverse climate policies and leave the Paris Agreement, his decision to change immigration policies, and his call to terminate NAFTA. Such issues as Trump's statements that contradicted traditional interpretations the U.S. Constitution, analogies to Trump with European politicians and political trends of populism and movements against immigration and Islam, and possible conflicts of interests of a Trump presidency and his global business empire.

The key topics of publications with graphics were environmental issues, immigration, health care, Trump's behavior, legal issues and economic issues. Most of the examples demonstrate facts, forecast and analyze possible outcomes of the new president's decisions and promises. Graphics also serve the purpose of demonstrating an analysis of the election results, the shift to the Republican side, and the reasons and conditions that helped Donald Trump to win.

Analysis of Graphics at The Washington Post Before the 2016 Election

Analysis of The Washington Post graphics examples showed that the most used types of visuals before the election were illustrations, horizontal/vertical bars, images, line graphs, maps, cartograms and treemaps. Among the least popular used were scatter plots,

40

tables and interactive visuals. 51 examples were included in the analysis of graphics produced by The Washington Post before the 2016 election.



Figure 4. The most used types of infographics by The Washington Post before election.

Illustration was generally used more by The Washington Post than pictures of the candidates (The Washington Post, 2016). Illustration was also implemented for other issues, such as informational purposes and entertainment. There is an example from The Washington Post which was used in a humorous story of a possible plan of how to rig the election, a then-popular topic in media (Berkowitz, Rivero & Meko, 2016).

One of the unique examples of vertical bars used by The Washington Post showed how Trump's campaign bars on Twitter presented distorted information (Muyskens, 2016). This example demonstrated a detailed analysis of the incorrectly built bar charts that were used on the Trump's Twitter account and showed his leading in different states. This example raised issues of infographic literacy, which will also be covered later in the interview section of the thesis.

The Washington Post extensively experimented with various types of maps. For coverage of this election, a new way of presenting votes was applied: the cartogram, which was used numerous times by The Post. The Washington Post created a piece that compared the two methods of presentation, such as traditional geographical map and cartogram, providing a historical perspective of election map development (Gamio, 2016).

The election polling results were presented in two ways by The Washington Post. The cartogram presents the results in a more precise way but at the same time it does not have the recognizable shapes of the states that possibly can create a confusion for the readers. In the interview part of the thesis, data journalists from both newsrooms covered the issue of using both ways with more details (Uhrmacher & Gamio, 2016).

The Washington Post used treemaps to compare numbers, such as an example with the results of early voting by different states (Uhrmacher & Gamio, 2016). The Washington Post has published multiple examples of interactive work which has different forms of quizzes and interactive text. In one of the interactives, users can become moderators and simulate debates. They are asked to pick a topic and candidate to see what is proposed by a particular candidate on the chosen issue (The Washington Post, 2016). After the GOP Convention, The Washington Post created an interactive quiz with facts about each state using quotes from the speeches form the convention (Uhrmacher & Schaul, 2016).

There is also an example of interactive quiz that allows users to compare their predictions with an average opinion of American voter which was created considering the national polling result. The quiz consisted of questions about the election and candidates (Guskin & Tan, 2016).

There was also an article with elements of infographics that consisted of interactive rolling text with false statements that Trump made. Users can put the cursor on the statement and see accurate information (Fahrenthold & O'Harrow Jr., 2016).

42

Text was an important tool in data visualization, such as in an example with information about election procedures and restrictions in each state. Users could choose any state from the list and the text would be change to the related information (Cameron, Vitkovskaya & Yang, 2016).

The Washington Post transformed textual information to quantitative data and presented it in a visual way, such as in a piece about the main frames in speeches at the Republican National Convention (Muyskens & Soffen, 2016).

The colors used most were the traditional red and blue for the parties, but purple, orange, yellow and grey were also found in the visuals.



Figure 5. Packed bubbles demonstrate colors used in The Washington Post visuals before election.
In most cases, the text was an element of the piece with infographics, sometimes highlighted or colored. Video was used twice in infographics before election. The majority of the pieces had neutral tone, and the main purpose was to provide analysis, facts, quotes, forecasts and entertainment. They covered political, economic and election issues, as well as the candidates' behavior. The most popular topics were possible scenarios and polling, Democratic and Republican Conventions, candidates' debates, candidates' statements and

presentations, analysis and extensive coverage of Donald Trump's false statements and exaggerations, the results of early voting, and campaign finance.

Analysis of Graphics at The Washington Post After the 2016 Election

In covering the outcome of the election, The Washington Post produced data visualization using mostly illustration, mapping, horizontal/vertical bars and line graphs. Other types of visuals used include cartograms, pictures, scatter plots, interactives, histograms, alternative reversed map, side-by-side circles, tables and animation. 18 examples were included in the analysis of graphics produced by The Washington Post after the 2016 election.



Figure 6. Treemap of visualization types used by The Washington Post during the month after the election.

Illustration was mostly used for the candidates' images and for drawing small objects. In one example, an illustration was used to draw a story about how marketing skills helped Trump win the election. Illustration was used to draw various examples of Trump using marketing approaches to promote his candidacy (Rivero, 2016).

After the election, The Washington Post produced a unique type of alternative map that was reversed and focused on the two coasts. In that article, journalists covered how large number of counties from the East Coast to the West Coast voted more Republican than in 2012 and where the highest peaks of support for both candidates were (Gamio & Keating, 2016). The journalist focused on this example in the interview part of the thesis.

As a result of multiple experiments with maps, The Washington Post published four versions of electoral maps for the final results of the election. The first one illustrates geographical shape, the second showed the margin with bubbles for each state, the third had arrows that showed where counties had flipped, and the fourth was a cartogram that demonstrated the weight of votes for each state (The Washington Post, 2016).

There were not as many examples of scatter-plots and interactive features, which are more time-consuming to build. There was only one example of a scatter-plot in postelection coverage, which showed how much better or worse candidates fared with voters compared to the previous election (Soffen, Mellnik, Granados, Muyskens, 2016). According to the data editor who was interviewed, scatter-plots are not used much because of the overwhelming information which is sometimes hard to read.

Interactive features served mostly for entertainment purposes in post-election coverage by The Washington Post. The newsroom created a music app where users could become DJs and produce their own "remix" of election sounds (Lu & Williams, 2016). Another example of an entertainment interactive was created for users who wanted to be distracted from the election and extensive news coverage of it. The app contains the series of various political jokes, funny videos, quotes and commercials (The Washington Post, 2016). The main colors in post-election coverage were red and blue. The text was mostly classic, sometimes highlighted or bold. Video and audio were not used in the overwhelming majority of cases. Most of the examples have neutral character and cover election and legislative issues, present facts, analysis, investigation and occasionally provide entertainment. The main topics and frames in the post-election coverage were contributions and campaign finance, Barack Obama's executive-orders strategy and how Donald Trump could use it, electoral college, new cabinet nominations, the divide between Republican and Democratic America, polling and underestimated chances for Donald Trump to win the election.

The results of textual and visual analysis of the data visualization examples from The New York Times and The Washington Post were presented. Among the pieces, there were many simple charts, as well as some experimental features and interactives about presidential election. Each piece contained visual and textual framing using design elements, particular topics and angles. In the next section, the results of the interviews with data journalists and data editors from both newsrooms will be presented and analyzed.

Interview Information

The duration of each interview was expected to be about 30-40 minutes. The length of the interviews varied, but averaged about 37 minutes. The shortest interview lasted 23 minutes. The longest interview was 64 minutes.

The researcher began the interviews with a question about the structure of the data department within their organization. Interviewees were also asked to provide their opinion about the purpose of data visualization during the 2016 presidential campaign and any problems or limitations while working with it. Other questions were focused on practical examples from the journalists' work, the interaction between data visuals and textual reporting, the process of making decisions about how data visuals should look, as well as what angle it should have. At the end, participants were asked about the outcomes of using data visualization for covering the presidential campaign and how data reporting can be improved in the future. As a final question, the researcher asked whether there were any issues that had not been addressed but that the journalists would like to cover to give the interviewees the opportunity add more of their thoughts about the process and add more information and personal experience. A complete script of the interviews with questions can be found in Appendix D.

Interviews were conducted individually, by Skype and by e-mail. Before the interviews were conducted, subjects were assured that their personal information and privacy would be secure. All the interviews were recorded using a digital audio recorder and then transcribed manually in a private environment using the oTranscribe platform, which helps to regulate the speed of the interviews and transcription process. After that, transcripts were kept in a secure environment.

Participant Information

Seven professionals were interviewed for this research. All of them were involved in the production of data visualization during the presidential campaign in the United States. The main criterion for selection was the affiliation with The New York Times or The Washington Post data departments. Such details as names, titles and other personal information are kept anonymous.

Data Analysis of the Interviews

As in the previous method, data from the interviews was analyzed and structured with the various categories, codes and broad topics. The system of codes was created. As in the previous method, microanalysis was used during the coding process for the interviews. With this approach, textual data was analyzed and broke into particular concepts and categories. Three aspects of microanalysis were included, such as data analysis, data interpretation, and finding the connection between the perceived and interpreted information. As in the textual and visual analysis, open coding was used in the analysis of the interviews.

The codes, which were found during the process of analyzing the interviews, were formulated according to the research questions. The primary research question was: What is the role of data visualization in framing of coverage of the 2016 U.S. presidential campaign? Then, all of the concepts and codes were stated in relation to the research questions.

Code	Definition	Examples
Structure of the Data Department	How data departments are	"It is a mix of journalists who are
	structured, what skills journalists	working on reporting, designing,
	need, how work on data	writing code, and some people
	visualization is divided	are specialized in cartography,
		statistics, VR nowadays and 3D
		modeling
Purpose of Data Visualization	What is the purpose of using data	<i>"it says a lot more than anything</i>
	visualization for coverage of	else about who or what the
	presidential campaign and	United States is as a country,
	election	what the country looks like"
Work Process	The description of the work	"We are collaborating with our
	process on data visualization,	editors, we have like a daily
	what stories journalists	meeting where it will be like a
	discovered and how data	panel of editors and if you have a
	visualization helped them to	graphic you sign up for a slot and
	improve their reporting	show them your sketch or pitch
		an idea"

 Table 2. Descriptive codes for the Interview

Challenges	Various problems and limitations which journalists dealt with covering the campaign with data visualization	"Then technically it is a challenge to present the data in a way that make sense across a tiny mobile screen and a big desktop screen"
Frames and trends	Frames, perspectives and trends in data visualization in coverage of the presidential campaign	"Framing is really powerful because it cuts out all the unnecessary details you don't need to tell your story, they're just going to distract"
Outreach	Considering the results and reactions from readers on data visualization after election	"The most traffic I've ever seen, almost definitely near the top of the most traffic I've ever seen on anything"
Recommendations for Future Coverage	Possible recommendations and advice about how to improve data reporting in the future coverage	"There is a room for visual fact- checking too and it is unexplored area within visualization"

Structure of the Data Department

Graphics desk exists in both newsrooms, although there are different forms. The structure can vary depending on how the work with data is planned. According to the interviewees, graphics departments consist of several teams, such as a data team ("work with databases, data cleaning and analysis"), infographics team ("data visualization"), story team ("handle creative storytelling within the narrative"). Each of them has professionals with various specializations, as Respondent No. 1 noted: "My team is composed with data scientists, developers, designers, reporters; most people have skills that kind of overlap those disciplines". Respondent No. 4 said that two people were hired in the graphics department specifically for the election the year before and were focused only on politics, but at the same time everybody in a team was working on election pieces: "I also find that it can be tough and stressful but doing graphic on deadline is also really rewarding because you have a ton of people interested in that and you get a big audience for it". Respondent No. 2 pointed out that journalists collaborate with each other to mix various skills:

"It is a mix of journalists who are working on reporting, designing, writing code, and some people are specialized in cartography, statistics, VR nowadays and 3D modeling. We usually work in combination of 2 or 3 people on a project to combine these skills together and make graphics".

Data teams consist of people with different skills such as journalists, scientists, illustrators, designers and developers. Departments work independently on their own projects but often collaborate with other units on big projects where data is a component and must be visualized. Everybody in the team has general knowledge about data, but also advanced skills in one or two narrow disciplines:

"We have data scientists who will actually do the analysis on the data, we have people who do databases, who do back end development, getting the data formatted into databases that can be used in applications, front end developers, design and then storytelling and reporting".

Another model, described by the Respondent No. 5, consists of five teams that work with data in different variations, such as graphics desk ("data visualization work"), Upshot ("data reporters and regular reporters that wrote heavily on the presidential election"), computer-assisted reporting team ("data analysis or investigations"), digital design desk ("front end programming"), and interactive news desk ("combination of all of the things").

Respondent No. 5 stated that having storytelling skills is important for data journalist: "*The ability to find a story just like you would when the narrative is important, that's probably the most important thing*". Respondent No. 2, as well as others in their interviews, pointed out that the skills of collecting and analyzing data are valuable. Expertise in working with spreadsheets and databases, programming languages such as

HTML, CSS, JavaScript and D3 library for producing interactive graphics, Adobe Illustrator, or other software like QGIS are helpful for producing effective data visualization. As Respondent No. 4 said, the infographics tools save time in producing data visualization, especially when someone is using coding in D3, which allows for updating data very quickly: "It's impractical to rebuild the graphic every single time we have new data which is in some cases every day, there's new polling data, you want your graphic automatically update, then we're building a graphic with D3". Respondent No. 3 also noted that the front end development background is the core for future employee, as well as the ability to adapt and learn new skills. One of the most important things that was noted by Respondent No. 4 is that the journalist must know how to tell a story visually and how to produce an effective visual component: "You may not even need to know how to produce whatever graphic you're making, but you have a sense for how you use your visuals to effectively do your storytelling or explanatory journalism". Respondent No. 5 also commented that the standard skillset of a reporter applies in working with data as well: "Being able to work with government agencies to request data, knowing how to do it, knowing what data is good and useful and what is not, being able to find truth of that data, find narratives in it".

Participants of the research noted that there is no formal trainings conducted in newsrooms, but self-learning is very useful. As Respondent No.1 said: "*My background is in design but I picked up the development in informal way and now I develop almost all of my own stories*". Several interviewees stated that there are skill-share meetings in which everybody on the team has 30 minutes to share their knowledge with other members of the

group. Respondent No. 4 pointed out that everybody on a team should have general knowledge of various skills:

"On our team most people eventually learn everything and everyone is expected to have a basic level and be able to do most things but we are definitely focused. People on the data team are going to focus more on doing regional data work, data analysis and bigger data projects while the infographics team is working on a lot of quick turnaround daily projects but we are often also collaborating".

Majority of interviewees mentioned that there is strong team unity. People on teams rely on each other and help each other learn new skills. For example, Respondent No. 3 shared: "*I didn't study any development in college, I've learned that in my career, doing projects and experimenting with patient people who helped me to learn*".

Purpose of Data Visualization

Data visualization is generally defined as a tool that helps present complex information in visually understandable ways for the audience. With the development of new technologies, the definition of data visualization is growing as well. Data visualization is also defined as a tool that helps present a different angle and, as Respondent No. 3 noted, *"get above 'he said, she said'*". It is always a complicated process of looking for perspective in a big amount of data, but according to Respondent No.1, it is important to present the most important aspect in it: *"I feel that the data can be so big and broad that what works well in a lot of stories is making it a little more specific with finding an angle on that data"*. The data can be framed and presented with the different connotations. For data visualization of the presidential campaign, different types of graphics were used, such

as maps, graphs and charts, as participants noted and the researcher found during the visual and textual analysis.

In coverage of the presidential campaign, data visualization was a highly important way of presenting information, especially when the topic was related to financial issues and quantitative information. Respondent No. 2 stated that it was much easier to convey this kind of information, such as polling numbers, with data visualization: *"It is easier to see it in a chart than in a sentence"*. Data visualization also helps to illustrate trends over time, as well as providing geographical references such as using maps for illustration of election results and comparing them with the results from previous elections. According to the Respondent No. 6, Dave Leip's US Election Atlas can be useful for doing retrospective analyses of the election results.

Elections present a massive amount of data (such as polling, political contributions, campaign finance and election results) and data visualization allows people to understand it easily. Detailed analysis of this data tells a lot about the country, according to the Respondent No. 4: *"It says a lot more than anything else about who or what the United States is as a country, what the country looks like"*.

As Respondent No. 3 pointed out, coverage of the presidential campaign and election provided many opportunities for experimenting with new tools and data visualization: *"There was room to try different stuff because there's so much going on"*. Many types of data visualization were used, such as line charts, bar charts and polling charts, but maps played a significant role, according to many respondents. As Respondent No. 3 pointed out, mapping is the area that gives the most opportunities for creative growth in election coverage:

"I think we were really good at being creative this year with map shapes and we made some cartograms that were by electoral college for each state and how come that they have huge state sizes but they don't have many electoral votes".

Cartograms became a subject of much experimentation for election coverage, but at the same time geographic maps were still used to facilitate the process of understanding the data visuals: "*I think that kind of thing is experimental and it's useful, it's a good use of interactivity to help people make connections between what they have seen, especially when it is something which is really conceptual and not obvious*".

Respondent No. 3 emphasized that coverage of the campaign with data visualization raised an important issue of data literacy for the audience. It led to extensive coverage of this issue, such as how to understand the maps and how to read the graphics. There was an example in which a journalist covered the deception in charts presented on Donald Trump's Twitter account, which exaggerated his lead in the polls. Respondent No. 4 stated an example of the piece about mapping election results:

"If you make some of these counties that voted for Trump red and the counties who voted for Clinton blue, you're losing a lot of nuance because what you will see is a map that's just full of red with a couple of little blue areas. Now, actually, more people voted for Hillary Clinton than for Donald Trump even though on the map it looks like there's a ton more red but the Clinton counties are a lot more highly populated. There are cities, they're much denser, and Trump counties are maybe rural and they may have fewer people".

Work Process

Work process is described in different ways. As Respondent No. 3 stated, it starts with investigating what data means and what aspects of it can be newsworthy, talking to people who keep the data and know what it means. After that, there is a process of visualizing it and trying different forms of it. Participants stated that usually it consists of interaction between journalists and editors who are working on a particular piece providing discussions and feedback:

"We are collaborating with our editors, we have a daily meeting where it will be a panel of editors and if you have a graphic you sign up for a slot and show them your sketch or pitch an idea and they are giving you feedback about the right way or angle, recommend the shape and try to make it better so we are working together and getting the editing".

Working on stories, journalists structure their work in various ways. As participants noted, sometimes they have data first and try to find ideas for stories inside the data but in other cases, journalists have a topic in mind and try to find data to illustrate it. Respondent No. 1 mentioned that usually at the beginning of the working on a story, there is a need to examine the data you have and try to look for a story in it or newsworthy pattern that may be interesting to the audience. The interviewee stated a problem that can arise in the working process when journalists try to find data that supports their existing opinion:

"People are trying to make the data say what they wanted to say, so I think it's better and more honest to just start from looking at what the data itself is saying, and then narrowing that down to something specific and newsworthy and interesting, that you can tell a story around. And then thinking of the design perspective of what's the best visual way to communicate that story and communicate what's happening in the data which will be different for every type of the data and different type of the story that you are telling".

The participant also noted that data visualization sometimes helps to see the degree that a problem or trend has and that it can be a surprise even if the problem was already known:

"There was another piece that one of my colleagues did, that was about the different weight of electoral votes so basically somebody in California's vote counts for a lot less than someone in Wyoming's. In the way she visualized it, was using state map that had bars coming out of it that showed you the difference in like what kind of more, what kind of less than what an actual vote you use in popular vote count. So you can really see that degree to which a person in California's vote counts for a lot less than a person in Wyoming's. Even if you knew that already just from looking at the spreadsheet of data, that really emphasized the degree to which it's true".

Several journalists stated that data visualization in most cases is published with traditional reporting in order to explain the meaning of data or trend. Respondent No. 1 also mentioned the importance of having traditional reporting along with data visuals:

"Usually the data doesn't just stay until you act to it. It tells you what's happening but not why it is happening. And from the other side, if you are reporting in a more traditional way, you are talking to people and they have anecdotes, you can use the data to say how this person fits into this larger picture whether they are part of the trend or against it; or you know an example of certain larger scene that they made by data that include a lot more people".

Respondent No. 3 confirmed that everything that the graphics department does usually involves traditional reporting, and it is always necessary to clarify data with researchers' opinions and people's experience.

Respondent No. 4 stated that quantitative data works the best for infographics but that at the same time there is the possibility of converting qualitative data into quantitative. During election coverage, it happened with speeches by political figures:

"We did this for the Republican Convention. So here, basically, we had text, we had the words people spoke but not necessarily numbers. There's a bunch of different ways you can turn speech into data but we decided to pick out some really common phrases and highlight how people use them".

Challenges and Limitations

The biggest challenge for data reporters, which was repeated by almost every participant, is cleaning the data, which in many cases can be messy and raw, especially when the journalists sometimes are gathering data from different sources. It is a very time-consuming process to prepare data before working with it. Often the data, which is necessary, isn't housed anywhere and in some cases, journalists must create the database by themselves. For example, Respondent No. 1 said: "*We have a police shootings database that tracks when a person is killed by a police officer and that data didn't exist before we started tracking it and the effort putting the tracking it is pretty significant*". The challenge can be also providing responsive design: "*Then technically it is a challenge to present the data in a way that make sense across a tiny mobile screen and a big desktop screen*".

Another problem that was considered by the journalists was the huge amount of data that has to be analyzed and presented for the audience in a digestible format, which is where data visualization plays important role. In coverage of the election, getting results from precincts was a challenge for data journalists and data scientists in the newsrooms. Each state has different procedures of reporting, and there is no standard way for counties to do it. There is also the challenge of covering the data in real time, which is not always possible, as Respondent No. 1 stated:

"But a lot of data is not in real time, it's from the last census or American Community Survey, which comes out once a year, government data usually comes out once a year, sometimes only even four years, every seven years or every ten years so it is hard to tell stories in real time just through data depending on the type of data it is. That is where it is really important to do the reporting side of it where you can actually talk to people and see how those experiences are represented and applied to the data itself".

Several respondents stated the difficulty of presenting the polling results because there are various methodologies from multiple companies and different ways of conducting it, such as by phone, mail or Internet, that made it difficult to convey the usefulness of the results.

Mapping became a challenge in coverage of the election results, which created debates even inside newsrooms, as Respondent No. 2 stated. This year, two types of presentations were used, such as traditional geographical map and cartogram:

"Having the state shaped as it is geographically, is much more recognizable by people who recognize the shapes and when you create a cartogram you have to move things around and it is harder for people to understand the geographical references. But then the cartogram represents the weight better".

Among other challenges, journalists stated that it is hard to find new ideas for data visualization to present information correctly and at the same time in unique way. As Respondent No. 3 pointed out, data literacy of the audience is one of the limitations that data visualization workers face, and sometimes there is a need to sacrifice the form in order to improve the meaning and accuracy:

"Making things that are not so conceptual that somebody will be able to look at it and get the story out of it. So we try really hard to annotate the graphics and to not make things so interactive that we're hiding information, to make it more like you can see something instead of having to go over to get the data".

The issue of data literacy of the audience was mentioned by the interviewees as a challenge. According to the Respondent No. 4, journalists cannot expect the same data expertise from their readers, which is why there is a responsibility of journalists to bring the audience there and explain how it is working.

One of the challenges that was mentioned by Respondent No. 4 was that there is no perfect visualization tool that allows for a perfect explanation of the data, that there will be always limitations: "So that might mean visualizing your data in more than one way without overwhelming the reader that might mean doing careful annotations and storytelling in text".

Another limitation, that was mentioned by Respondent No. 6, is that there is always a time limit, which restricts the ability to produce sophisticated visualizations that take a lot of time to do. At the same time, as it was stated previously, it needs to be responsive for different devices: "It's only 20% of people are looking at pages on Desktop. That has required rethinking a lot of ideas and things that I'd like to do".

Frames and Trends

Respondent No. 5 stated in the interview that a trend for simplifying graphics is appearing:

"If we're going to ask a user to click on something or to interact with it in some way that goes beyond scrolling, the payoff for doing that has to be worthwhile and all we want people to do is to see some numbers in a chart, crafted over time for trend purposes, there's no reason for us to ask people to click on that to be able to make that happen or to interact with it in some way rather we are going to show them the information we want them to see and show them the information that best tells the story they can understand".

Respondent No. 3 also observed a trend in the decreasing priority of interactivity in data visualization. Newsrooms realized that scrolling is more native behavior for users and additionally, it is always more difficult and time consuming to make interactives responsive on different kinds of screens such as mobile phones or tablets.

As Respondent No. 1 stated, during the coverage of the election, current trends in data visualization appeared, as well as many new approaches.

"That was a map, but that was vertically presented so the east coast was on the top and the west coast was on the bottom. That was showing with slashes, these diagonal lines, areas of the country that had swung either Democrat or Republican after the last election and they had a text embedded throughout the map so, I think, that is just an example of taking something that is really traditional way of covering the election like showing precinct results and doing it in the way that really tells a story by focusing on more narrow aspect of the data and giving it an angle, taking the time to carefully explain it".

Traditional reporting can help to explain data as well as to provide a newsworthy frame related to the topic:

"I think that's what you see pretty frequently with these projects that they are so going to have the explanation, that you would have if you were reporting the story without visuals but you use the visuals to tell a larger part of that story and then the text is more supporting for framing and explaining different aspects of it that need more nuance".

In many, cases several people work on one piece, which allows a mix of different skills and makes the final content rich in terms of presenting different kinds of reporting and angles. Respondent No. 1 described the process of making decisions about the possible angle of the story: *"[We decide] is it newsworthy or do people care about it or [is it] relevant. And then you construct the story around it to give it the most impact and make it more clear".*

In considering framing, Respondent No. 3 stated that colors can be used in this process, especially with blue and red colors for election coverage: "*We are trying to not use blue and red in our nonelection graphics*".

Respondent No. 4 covered the style that the newsroom follows to make its graphics recognizable and easy to digest, such as sticking to simple forms such as lines, bars and charts without overwhelming the viewer. The style guide also addresses using colors:

"The most important things to remember in when using color, it is really tricky in elections because we've made this association between Democrats are blue and Republicans are red, that's not been the case throughout history, that's more recent thing but it's such a strong association now that you should be really careful when you're using those colors for anything else".

At the same time, it is better not to rely on color because of issues like color blindness making graphics understandable and visible for all readers.

As Respondent No. 4 also mentioned that framing can be useful for making particular elements of a story more salient:

"Framing is really powerful because it cuts out all the unnecessary details you don't need to tell your story, they're just going to distract. But at the same time the power of data visualizations is putting your data in context so often times you can make the context small and then you make your story big, I guess, that the basic of what we do".

Among the main frames of the election campaign coverage, Respondent No. 1 identified the focus on urban vs. rural, areas that are more white and more middle class, coverage of working-class manufacturing jobs, which were major issues in the campaign. Respondent No. 4 mentioned that election results in general in many different forms provided many frames.

Outreach

Professionally well-done data visualization resonates with people, but otherwise people can be confused if a visual is presented in an unclear way or with an overwhelming amount of information. Considering the election coverage, as a result, pieces that focused on possible mistakes in predictions and reporting appeared, as Respondent No. 1 said: "There was more analysis of using the data to explain that larger trend. There was a lot of focus on exactly what the results meant and looking at them from different lens".

Speaking about elections, respondents pointed out that there was a negative reaction from the audience after unexpected results took place. This phenomenon attracted attention to another problem in data visualization, which is how to present uncertainty in polls. As Respondent No. 2 pointed out, people blamed the media for showing an incorrect forecast:

"At the point of the election Trump was at 16 percent chance of winning and that number looked very low to people and people told our newsroom had helped create this situation getting Trump elected. People don't really understand what 16 percent chance means, even 10 percent chance is pretty high".

In terms of attracting the audience, the election coverage is very popular, especially election maps, as Respondent No. 4 said: "*The most traffic I've ever seen, almost definitely near the top of the most traffic I've ever seen on anything.*" Another interesting trend that was mentioned by the same respondent was that sometimes the simplest content has very high traffic: "*One of our best pages that got the most traffic, that was a graphics page, was just a schedule of when the debates were, very simple, important service journalism*". The complexity or uniqueness of a graphic piece does not mean that it will have high outreach, according to the Respondent No. 4: "*Most of our audience is not going to the page because like*" *oh it's a cool interactive*", *they come to the page because it is compelling story or because it's something they want to learn about*".

Several participants provided information that the number of subscribers increased in both newsrooms despite the incorrect forecast before the election.
Recommendations for Future Coverage

Respondent No. 1 raised the issue of lack of trust in data from the audience. As the interviewee stated, there was a readers' response which stated that data journalism does not seem objective to the audience. It led to a recommendation for future improvement, such as being more *"transparent with how you analyze the data and how you are presenting it and making sure that you are really clear and precise"*, providing sources of data and details for the methodology of data analysis.

Respondent No. 3 underlined that it is always necessary to cover all sides of a problem and to be as nonbiased as possible in coverage. A good practice that can be implemented is fact-checking within a newsroom and, according to Respondent No. 3, *"there is a room for visual fact-checking too and it is unexplored area within visualization"*.

Respondent No. 4 concluded that in the future, data journalism and data visualization should be focused more on reporting candidates' policies and explaining what they mean:

"Explaining the impact of each candidate's policies to people, for instance, clearly explaining what Republicans want to do for health care versus Democrats, I think, there were a lot of voters who may have been confused by talk of Obamacare versus the Affordable Care Act. These are things that in terms of policy may be complicated for most people to understand but I think through graphics we have potential to explain complex topics in more easily understandable ways".

The Respondent emphasized the strong need for explanatory visual journalism and providing more information about the consequences of electing a candidate, because choosing a particular party can change the direction of the country. Respondent No. 4 also pointed out that it is important in the future to cover uncertainty when reporting on polling:

"There was a lot of interest in predicting the outcome of the election and I'm not going to pick on any particular outlet in this right now but the overall picture that was painted by these forecasts was one of high confidence in Hillary Clinton winning the election. That was at least, I think, what the audience took away. However, the audience may not be familiar or fully understand probability. That is just fundamentally difficult for humans to understand".

As Respondent No. 5 stated, the advice would be to approach the data with a more open mind:

"Being open to the idea that when you're working with data sometimes narratives are the right way to tell a story and sometimes it's not, and how do you build out in your newsroom the right forms to be able to tell the stories using visual means in the way that you are better to tell them. On the reporting side, being open minded about trying to find stories and data that you might not have thought of before rather than just coming to the question and having the data to answer that question on visualization side".

Respondent No. 6 raised the question of math and statistical literacy among the audience and that was something that journalists knew but learned more about during the election coverage: *"I think a lot of people who are making these things don't have an intuitive sense that people don't understand what they're communicating"*. As a solution, journalists have to present the data in the audience's terms and make it easier for readers to understand.

Chapter Summary

In this chapter, the researcher stated the results of visual and textual analysis, as well as the outcomes of interviews conducted for this research. The chapter contains descriptive codes for both methods. The research data formed major components, such as analysis of visual examples from The New York Times and The Washington Post, and various themes for the interview data such as structure of the data department, purpose of data visualization, work process, challenges and limitations, frames and trends, outreach, and recommendations for future coverage.

The results showed that the data that was collected by different methods and then analyzed provided coverage of the research topic from different sides, and both methods enriched each other.

Chapter 5. Discussion

Summary

Among the main findings that were found during the research were multiple ideas and observations:

- Technology is developing every day, and many sophisticated methods of data visualization are constantly appearing, which means that journalists have to be flexible and adapt and implement new skills. At the same time, traditional storytelling skills are still important;
- Even with the existence of highly developed tools, data journalists and editors tend to use simple forms more often to make visuals easier for readers to understand and avoid overwhelming them with very complicated graphics and interactives;
- Another advantage of simple forms is that it takes less time to create a high-quality chart. At the same time, professionally produced interactive pieces have a stronger influence on the audience;
- Data visualization is mostly presented as part of textual reporting, with both components containing graphic and textual framing;
- Visual framing in election coverage is mostly found through the use of colors (blue and red), as well as the type of infographic (maps, cartograms);
- Textual framing contains frames about the electorate (race, gender, education), parties and candidates' positions and projected policies, and personal background and issues (such as Trump's taxes and Clinton's FBI investigation);

- Election coverage showed a trend toward multiple experiments with data visualization, such as interactive trees and cartograms. Using geographic maps and cartograms became a controversial issue of coverage with graphics for two newsrooms. The Washington Post implemented more experimental forms while The New York Times followed more traditional format of data visualization;
- Infographics can be used not only for presenting complex data but also for entertainment;
- Text became a valuable tool for data visualization and can be converted to quantitative information;
- Challenges of data and visual literacy appeared during the coverage of the last presidential campaign;
- Wide diversity of participants' expertise helped to analyze various perspectives on data visualization from journalistic, design and computer/mathematical science background;
- A strong need for better interaction with audience and users appeared;
- The lack of audience trust in data reporting raised an issue of transparency in methodology of data analysis and citing of data sources;
- The question of how to present uncertainty and probability emerged after the election;
- Data reporting presented new challenges, such as working with data cleaning and dealing with massive amounts of information and responsive design;
- Journalists realized that they cannot expect the same data expertise from their readers that they have;

• There is a strong need in explanatory visual journalism for providing more information about what the consequences of electing each candidate would be.

This chapter provides an analysis of the research findings in the context of framing theory and how it works in relation to data visualization techniques. Limitations and implications of the research, as well as recommendations for future research, are stated in this chapter as well.

In the process of the textual and visual analysis, various patterns, specific topics and cases were analyzed and various methods of how journalists implemented datavisualization techniques were identified. The textual and visual analysis helped in finding certain contexts for using data visualization in coverage of the campaign and discovering particular patterns and cases.

Considering the results of visual analysis of The New York Times graphics before the election, the most popular types of graphics during the coverage of presidential campaign were vertical/horizontal bars, tables, line graphs, geographic maps, pictures and illustrations. During the analysis of the purpose of data visualization examples, several types were found, such as presenting the results of polls and forecasts, facts, opinions, analysis, quotes and live results. The most popular topics in The New York Times graphics during the six months before the election were poll results and forecasts.

Examples of coverage after the election by The New York Times include the divide between Republicans and Democrats, the protests against Trump's election, the potential consequences of Trump fulfilling his promises during the campaign, such as his plan to reverse climate-change policies and leave the Paris Agreement, his decision to change immigration policies, and his call to terminate NAFTA. Such issues as Trump's statements that appear to contradict the U.S. Constitution, analogies to Trump of European politicians and populist political trends and movements against immigration and Islam, possible conflicts of interests of a Trump presidency, and his global business empire.

The key topics of publications with graphics were environmental problems, immigration, health care, Trump's behavior, legal issues and economic issues. Most of the examples demonstrate facts, forecasts and analyses of possible outcomes if the new president keeps his promises. Graphics also served to analyze election results, the national shift toward Republicans, and the reasons and conditions that helped Trump win.

Analysis of The Washington Post graphics examples demonstrated that the most used types of visuals before the election were illustrations, horizontal/vertical bars, images, line graphs, maps and cartograms, as well as treemaps. Among the less popular were scatter plots, tables and interactive visuals.

In most cases, the text was an element of the infographics, and sometimes it was highlighted or colored. Video was used twice in infographics before the election. The majority of the pieces had neutral tone, and the main purpose was to present analysis, facts, quotes, forecasts, or provide entertainment. They cover political, economic, election issues as well as candidates' behavior. The most popular topics were possible scenarios and polling, Democratic and Republican Conventions, candidates' debates, candidates' statements and presentations, analysis and extensive coverage of Trump's false statements and exaggerations, the results of early voting, and campaign finance. Most of the examples have neutral character and cover election and legislative issues, presenting facts, analysis, investigations and, in some cases, entertainment. The main topics and frames in the post-

strategy and how Donald Trump may use it, the electoral college, new cabinet selections, the divide between Republicans and Democrats, polling, and the underestimated chances of Trump winning the election.

Analysis demonstrated that The New York Times used more traditional and simple forms of data visualization while the coverage by The Washington Post contained more experimental examples. At the same time, The New York Times has successfully illustrated uncertainty of the election forecast with the interactive dial visualization. Cartograms were mostly used by The Washington Post and were presented as an alternative way to show polling and election results.

The interviews provided information about the professional experience of data journalists and editors and their opinions on the role of data visualization, problems and limitations that they experienced while working with it.

The results of the interview analysis illustrated that data teams consist of people with different skills, such as journalists, scientists, illustrators, designers and developers, who collaborate with each other in order to mix various skills. The infographics tools help save time for producing data visualizations and presenting complex information in visually understandable ways for the audience and help to present a different angle

Cartograms became a subject of much experimentation, but at the same time geographic maps are still used to facilitate the process of understanding the data visualization. This year, coverage of the election with data visualization raised an important issue of audience data literacy. It led to extensive coverage of this issue, such as how to understand maps and how to distinguish the right graphics. There is a problem that can be found in the working process when journalists try to find data that supports their existing opinion. Several journalists stated that data visualization in most of the cases is published within traditional reporting to explain the meaning of data or trends, and it is always necessary to clarify data with researchers' opinions and people's experience.

Quantitative data works the best for infographics but at the same time there is the possibility of converting qualitative data into quantitative. The biggest challenge for data reporters is cleaning the data, which in many cases can be messy and raw, especially when journalists gather data from different sources. It is a very time-consuming process to prepare data before working with it. Often the data isn't housed anywhere and in some cases, journalists create the database by themselves.

Another problem that was considered by the journalists was the huge amount of data, which has to be analyzed and presented to the audience in a digestible format, where data visualization plays an important role. In coverage of the election, getting precincts' results was a challenge for data journalists and scientists in the newsroom. Each state has different procedures in the process of reporting results, and there is no standard way for counties to do that. Several respondents stated the difficulty of presenting the polling results because there are various methodologies of polling from multiple companies and diffent ways of conducting it (by phone, mail or the Internet) that made it difficult to convey the usefulness of the results.

Mapping was a challenge in coverage of the election results, which led to two types of presentations, in the form of traditional geographic maps and cartograms. Among other challenges, journalists stated that it is hard to find new ideas for data visualization to present information correctly and at the same time in a unique way. Data literacy is one of the limitations that data visualization is faced with, and sometimes there is a need to sacrifice the form in order to improve the meaning and accuracy. Journalists cannot expect the same data expertise from their readers, that is why there is a responsibility for journalists to bring the audience there and explain how it is working.

Another limitation stated was time constraints, which add many restrictions to the ability to produce sophisticated visualizations that take a lot of time to do. At the same time, it needs to be done in responsive form for different devices, which has led to a trend in simplifying graphics and decreasing the priority of interactivity in data visualization. Newsrooms realized that scrolling is a more native behavior for users.

Colors can be used for framing, especially blue and red colors for election coverage. Among the main frames of the election campaign coverage, the focus on urban versus rural, areas that are more white and more of middle class, coverage of working-class manufacturing jobs were stated by interviewees, which were major issues during the campaign. Election results in many different forms provided many frames.

Concerning election, respondents pointed out that there was a negative reaction from the audience after unexpected results took place. This phenomenon attracted attention to another problem in data visualization, which is how to present uncertainty in polls.

In this chapter, the limitations of this research, its theoretical and practical implications and recommendations for further study are stated.

Limitations

In this research, various limitations can be found, such as problems with generalizations and that the data analysis results cannot represent the whole mass media industry. Another possible limitation can be considered the number of interviewees which is comparatively small due to the difficulties with availability of the subjects. The diversity

73

of participants' background can present the wide range of perspectives but at the same time can narrow the generalization.

The biases of interview subjects as well as researcher herself could affect the findings of the study, as a side effect of qualitative research design. In addition to the stated limitations, there is also the researcher's lack of experience in conducting academic research.

Theoretical Implications

The theory of framing served as a basis for this research. As a result, evidence that supports the theory and its implications were found and stated in research results. Particular issues and mechanisms of how it can be applied in data visualizations were stated.

As it was previously stated in the literature review, there is a lack of research in the area of visual framing in journalism. This study provided new findings and can be a source of more extensive knowledge about the topic.

Practical Implications

There is wide practical implication of this research because it contains the detailed descriptions of journalists' experiences, which can be useful for professionals working in data journalism in the future. This research revealed issues related not only to the coverage of the presidential campaign and election, but also the challenges of data visualization reporting in general. There are many recommendations stated for future coverage by the data journalists interviewed, such as being more transparent with the methods for data analysis, providing sources of data and details for the methodology. A best practice that can be implemented is fact-checking within a newsroom, and especially visual fact-

checking. Among other recommendations, there is advice to provide explanatory visual journalism about political issues and candidates' plans.

Recommendations

For future research, a wider range of newsrooms and participants can be used to enrich the knowledge and generalization of the present study. In this research, the two largest American newsrooms were chosen and their practices were analyzed but at the same time there could be more useful knowledge found about smaller organizations and their methods of dealing with the professional challenges.

Future research can cover more deeply the experience, problems and challenges for journalism in data visualization and focus more on the technical side of the issue involving software and other components in more details. In this research certain technology issues and challenges were identified such as challenges with responsive design which could be analyzed with more details and extend the number of perspectives on the data journalism work.

Additional research can be done in analyzing the education of journalists who work on data visuals. The research in this field could be helpful for the media industry because it could state more directions for journalists to improve their skills and work.

References

- Aisch, G., Cohn, N., Cox, A., Katz, J., Pearce, A., Quealy, K. (2016). Live Presidential Forecast. *The New York Times*. Retrieved from <u>http://www.nytimes.com/elections/forecast/president</u>
- Aisch, G., Cox, A., Katz, J., Pearce, A., Quealy, K., Bostock, M., Carter, S. (2016). Who Will Be President? *The New York Times*. Retrieved from <u>https://www.nytimes.com/interactive/2016/upshot/presidential-polls-forecast.html</u>
- Aisch, G., Pearce, A., Yourish, K. (2016). How Large Is the Divide Between Red and Blue America? *The New York Times*. Retrieved from <u>https://www.nytimes.com/interactive/2016/11/04/us/politics/growing-dividebetween-red-and-blue-america.html</u>
- Andrews, W, Bennett, K & Parlapiano, A. (2016). 2016 Delegate Count and Primary Results. *The New York Times*. Retrieved from <u>https://www.nytimes.com/interactive/2016/us/elections/primary-calendar-and-results.html</u>
- Associated Press. (2016). Becoming a "tech-literate" journalist. Retrieved from <u>https://insights.ap.org/industry-trends/becoming-a-tech-literate-journalist?utm_source=twitter&utm_medium=social</u>
- Bennett, L., & Iyengar, S. (2008). A New Era of Minimal Effects? The Changing Foundations of Political Communication. *Journal of Communication*, 58(2), 707-731.
- Berkowitz, B., Rivero, C., Meko, T. (2016). If you want to rig an election... *The Washington Post*. Retrieved from <u>https://www.washingtonpost.com/graphics/politics/2016-election/how-to-rig-an-election/</u>
- Boyles, J. L. & Meyer, E. (2016). Letting the Data Speak, *Digital Journalism*, 4:7, 944-954, DOI: 10.1080/21670811.2016.1166063.
- Brewer, P., Graf, J., & Willnat, L. (2003). Priming or framing: Media influences on Attitudes Toward Foreign Countries. *Gazette: the international journal for communication studies*, 65(6), 493–508.
- Burmester, M., Mast, M., Tille, R., & Weber, W. (2010). How users perceive and use interactive information graphics: an exploratory study. *Proceedings from 2010* 14th International Conference Information Visualisation, 361–368.
- Cameron, D., Vitkovskaya, J., Yang, J., L. (2016). How to vote in each state. *The Washington Post.* Retrieved from

https://www.washingtonpost.com/graphics/politics/voting-rights/

- Carragee, K., & Roefs, W. (2004). The Neglect of Power in Recent Framing Research. *Journal of Communication*, 54(2), 214-233.
- Chen, C., Hrdle, W., Unwin, A. (2008). Handbook of Data Visualization (Springer Handbooks of Computational Statistics). Springer-Verlag TELOS, Santa Clara, CA.
- Chong, D., & Druckman, J. (2007). A Theory of Framing and Opinion Formation in Competitive Elite Environments. *Journal of Communication*, 57, 99–118.
- Clifford G. Christians & James W. Carey (1981). The Logic and Aims of Qualitative Research in G.H.Stempel & B.H.Westley (eds.), Research Methods in Mass Communication (Eglewood Cliffs, NJ:Prentice Hall), p. 354-374
- Cohn, N. (2016). Donald Trump's Red-State Problem. *The New York Times*. Retrieved from <u>https://www.nytimes.com/2016/08/11/upshot/donald-trumps-red-state-problem.html</u>
- Dance, G., Byrd, A. (2016). How Donald Trump Avoided Paying Taxes Using Other People's Money. *The New York Times*. Retrieved from <u>https://www.nytimes.com/interactive/2016/us/politics/trump-taxes-</u> <u>loophole.html?smid=tw-nytimes&smtyp=cur</u>
- De Vreese, C. (2005). News framing: Theory and typology. *Information Design Journal*, 13(1), 51-62.
- Denzin, N. & Lincoln, Y. (2007). Collecting and Interpreting Qualitative Materials (3rd edition) (Eds.). Thousand Oaks: Sage.
- Diakopoulos N. & Hullman, J. (2011). Visualization Rhetoric: Framing Effects in Narrative Visualization. *IEEE Transactions on Visualization and Computer Graphics archive.* 17:12, 2231-2240.
- Entman, R. (1993). Framing: Towards Clarification of a Fractured Paradigm. *Journal of Communication*, 43(4), 51–58.
- Entman, R. (2007). Framing Bias: Media in the Distribution of Power. *Journal of Communication*, 57 (1), 163-173.
- Fahrenthold, D., O'Harrow Jr., R. (2016). Trump: A True Story. *The Washington Post*. Retrieved from <u>https://www.washingtonpost.com/graphics/politics/2016-election/trump-lies/</u>
- Fessenden, F. (2016). Donald Trump's Big Bet on Less Educated Whites. *The New York Times*. Retrieved from <u>https://www.nytimes.com/interactive/2016/11/07/us/how-trump-can-win.html</u>

Gamio, L. (2016). Election maps are telling you big lies about small things. The

Washington Post. Retrieved from

https://www.washingtonpost.com/graphics/politics/2016-election/how-electionmaps-lie/

- Gamio, L., Keating, D. (2016). How Trump redrew the electoral map, from sea to shining sea. *The Washington Post*. Retrieved from <u>https://www.washingtonpost.com/graphics/politics/2016-election/election-results-from-coast-to-coast/</u>
- George-Palilonis, J. & Spillman, M. (2011). Interactive Graphics Development: A framework for studying innovative visual story forms. *Visual Communication Quarterly*, 18:3, 167-177.
- Goffman, E. (1974). Frame Analysis. Boston: Northeastern University Press.
- Gold, G., Huang, J., Parshina-Kottas, Y., Yourish, K. (2016). Can You Tell What Questions Trump and Clinton Were Answering? *The New York Times*. Retrieved from <u>https://www.nytimes.com/interactive/2016/09/26/us/elections/trump-clintondebate-quiz.html?smid=pl-share</u>
- Guskin, E., Tan, S. (2016). How much do you know about what American voters think? *The Washington Post.* Retrieved from <u>https://www.washingtonpost.com/graphics/politics/2016-</u> <u>election/conventions/conventions-quiz/</u>
- Hanna, P. (2012). Using Internet Technologies (Such as Skype) as a Research Medium: A Research Note. *Qualitative Research*, 12, 239-242.
- Heller, S., Landers, R. (2014). Infographic designers' sketchbooks. New York: Princeton Architectural Press.
- Iyengar, S. (1987). Television News and Citizens' Explanations of National Affairs. *The American Political Science Review*, 81(3), 815-832
- Iyengar, S. (1989). How citizens think about national issues: A matter of responsibility. *American Journal of Political Science*, 33, 878–900.
- Iyengar, S., & Kinder, D. R. (1987). News that matters: Television and American opinion. Chicago: University of Chicago Press.
- Jankowski, N. (1991). Qualitative research and community media. In K. Jensen, & N. Jankowski (Eds.), Qualitative methodologies for mass communication. London: Routledge.
- Khan M. & Khan S. (2011). Data and Information Visualization Methods, and Interactive Mechanisms: A Survey. *International Journal of Computer Applications* (0975 – 8887) Volume 34– No.1
- Kvale, S. (1996). Interviews: An Introduction to Qualitative Research Interviewing. London: Sage.

- Lee, J. (2016). How States Moved Toward Stricter Voter ID Laws. *The New York Times*. Retrieved from <u>https://www.nytimes.com/interactive/2016/11/03/us/elections/how-states-moved-toward-stricter-voter-id-laws.html?smid=pl-share</u>
- Lu, D., Williams, A. (2016). The 2016 election in sound: Make your own election song from the results. *The Washington Post*. Retrieved from <u>https://www.washingtonpost.com/graphics/politics/2016-election/election-</u> <u>remix/?Dem=guitar&GOP=synth&genre=glam&pollclose=808&statecall=game</u>
- Mason, J. (2002). Qualitative Researching. London: Sage.
- Maxwell, Joseph A. (2009). Designing a qualitative study. In Leonard Bickman & Debra J. Rog (Eds.), The Sage handbook of applied social research methods (2nd ed., pp. 214–253). Thousand Oaks, CA: Sage.
- McCombs, M. (2005). The Agenda-setting Function of the Press. In The Press: Institutions of American Democracy series, Geneva Overholser and Kathleen Hall Jamieson (Eds.). New York: Oxford University Press & The Annenberg Foundation Trust at Sunnylands, Chapter 9, 156-168.
- McCombs, M., & Shaw, D. (1972). The Agenda-setting Function of Mass Media. Public Opinion Quarterly, 36, 176–187.
- Muyskens, J. (2016). Most of Trump's charts skew the data. And not always in his favor. *The Washington Post.* Retrieved from <u>https://www.washingtonpost.com/graphics/politics/2016-election/trump-charts/</u>
- Muyskens, J., Soffen, K. (2016). Inside the echo chamber. *The Washington Post*. Retrieved from <u>https://www.washingtonpost.com/graphics/politics/2016-election/gop-convention-speeches/</u>
- Nelson, T., Lecheler, S., Schuck, A., & De Vreese, C. (2012). Framing Effects on Values, Emotions, and Attitudes. Western Political Science Association.
- Nussbaumer, K. C. (2015). Storytelling with data: A data visualization guide for business professionals. Hoboken, New Jersey: Wiley.
- Pan, Z., & Kosicki, G. (1993). Framing Analysis: an Approach to News Discourse. *Political Communication*, 10, 55–75.
- Park, H., Griggs, T. (2016). Could Trump Really Deport Millions of Unauthorized Immigrants? *The New York Times*. Retrieved from <u>https://www.nytimes.com/interactive/2016/11/29/us/trump-unauthorizedimmigrants.html</u>
- Patel, J. (2016). Voting Early, and in Droves: Over 22 Million Ballots Are Already In. *The New York Times.* Retrieved from

https://www.nytimes.com/interactive/2016/10/31/us/elections/earlyvoters.html?smi d=tw-share

- Patel, J. (2016). Voting Early, and in Droves: Over 22 Million Ballots Are Already In. *The New York Times*. Retrieved from <u>https://www.nytimes.com/interactive/2016/10/31/us/elections/earlyvoters.html?smi</u> <u>d=tw-share</u>
- Rivero, C. (2016). How marketing helped Donald Trump win the 2016 election. *The Washington Post*. Retrieved from <u>https://www.washingtonpost.com/graphics/politics/2016-election/trump-campaign-marketing/</u>
- Roulston, K. (2010). Considering Quality in Qualitative Interviewing. *Qualitative Research*, 10, 199-228.
- Rutenberg, J. (2016). News Outlets Wonder Where the Predictions Went Wrong. *The New York Times*. Retrieved from <u>https://www.nytimes.com/2016/11/10/business/media/news-outlets-wonder-where-</u> <u>the-predictions-went-wrong.html</u>
- Scheufele, D. (1999). Framing as a Theory of Media Effects. *Journal of Communication*, 49(1), 103–122.
- Scheufele, D., & Iyengar, S. (2012). The State of Framing Research: A Call for New Directions. In K. Kenski & K. H. Jamieson (Eds.), The Oxford Handbook of Political Communication Theories. New York: Oxford University Press.
- Scheufele, D., & Tewksbury, D. (2007). Framing, Agenda Setting, and Priming: The Evolution of Three Media Effects Models. *Journal of Communication*, 57(1), 9– 20.
- Seale, C. (1999). The Quality of Qualitative Research. London: Sage.
- Segel, E. & Heer, J. (2010). Narrative Visualization: Telling Stories with Data. *Visualization and Computer Graphics*, 16:6, 1139-1148.
- Simmon, R. (2013). Is Animation an Effective Tool for Data Visualization? *NASA Earth Observatory*. Retrieved from <u>http://earthobservatory.nasa.gov/blogs/elegantfigures/2013/03/14/is-animation-an-effective-tool-for-data-visualization/</u>
- Soffen, K., Mellnik, T., Granados, S., Muyskens, J. (2016). In a crucial Democratic stronghold, Trump surged. Clinton didn't. *The Washington Post*. Retrieved from <u>https://www.washingtonpost.com/graphics/politics/2016-election/precinct-post-analysis/</u>
- Statistical Analysis System Institute. (n. d.). Data Visualization: What It Is and Why Matters. Retrieved from <u>http://www.sas.com/en_us/insights/big-data/data-visualization.html</u>.

- Strauss, A. & Corbin, J. (1990). "Grounded Theory Research: Procedures, Canons and Evaluative Criteria." *Qualitative sociology*. 19(6). 418-427.
- Strauss, B. (2016). What Trump's Climate Legacy Could Look Like. *The New York Times*. Retrieved from <u>https://www.nytimes.com/interactive/2016/11/20/sunday-review/donald-trump-climate-legacy.html?smid=pl-share</u>
- The New York Times. (2016). Fact Checks of the 2016 Election. Retrieved from https://www.nytimes.com/interactive/2016/us/elections/fact-check.html
- The New York Times. (2016). Presidential Election Results: Donald J. Trump Wins. Retrieved from <u>http://www.nytimes.com/elections/results/president</u>
- The Washington Post. (2016). The 2016 Election Fact Checker. Retrieved from https://www.washingtonpost.com/graphics/politics/2016-election/fact-checker/
- The Washington Post. (2016). Your 2016 Debate. Retrieved from https://www.washingtonpost.com/graphics/politics/2016-election/virtual-debate/
- The Washington Post. (2016). Live Results: Presidential Election. Retrieved from <u>https://www.washingtonpost.com/2016-election-results/us-presidential-</u> <u>race/?utm_term=.f97d242c08fa</u>
- The Washington Post. (2016). 20 Election Day distractions. Retrieved from <u>https://www.washingtonpost.com/graphics/politics/2016-election/how-to-deal-</u> <u>with-election-day/</u>
- Tracy, S. (2010). "Qualitative Quality: Eight 'Big-Tent' Criteria for Excellent Qualitative Research." *Qualitative Inquiry*. 16(10), 837-849.
- Tufte, E. (1983). The visual display of quantitative information. Cheshire, CT: Graphics Press.
- Uhrmacher, K., Gamio, L. (2016). Poll: Redrawing the electoral map. *The Washington Post.* Retrieved from <u>https://www.washingtonpost.com/graphics/politics/2016-</u> <u>election/50-state-poll/</u>
- Uhrmacher, K., Gamio, L. (2016). More than 2.5M votes were cast before the final debate. Does it matter? *The Washington Post*. Retrieved from <u>https://www.washingtonpost.com/graphics/politics/2016-election/what-earlyvoters-missed/</u>
- Uhrmacher, K., Schaul, K. (2016). 'Best state for duck hunting'? Name that GOP convention brag. *The Washington Post*. Retrieved from https://www.washingtonpost.com/graphics/politics/2016-election/conventions/name-that-gop-state-brag/
- Valkenburg, P, Semetko, H, & De Vreese, C. (1999). The Effects of News Frames on Readers' Thoughts and Recall. *Communication research*, 26(5), 550-569.

- Wallace, T. (2016). The Two Americas of 2016. *The New York Times*. Retrieved from <u>https://www.nytimes.com/interactive/2016/11/16/us/politics/the-two-americas-of-2016.html?smid=pl-share</u>
- Weber, W., Rall, H. (2012). Data Visualization in Online Journalism and Its Implications for the Production Process. 16th International Conference on Information Visualisation, 349 – 356.
- Yin, R. (2011). Qualitative Research from Start to Finish. New York: Guilford.

Appendix A: IRB Approval



190 Galena Hall; Dc074.00 Columbia, MO 65212 573-882-3181 irb@missouri.edu

January 23, 2017

Principal Investigator: Iuliia Alieva Department: Journalism-MA

Your Exempt Application to project entitled DATA JOURNALISM AND DEMOCRACY: HOW AMERICAN MASS MEDIA FRAMED THE 2016 PRESIDENTIAL CAMPAIGN IN THE UNITED STATES USING DATA VISUALIZATION was reviewed and approved by the MU Institutional Review Board according to the terms and conditions described below:

IRB Project Number	2007528
IRB Review Number	222307
Initial Application Approval Date	January 23, 2017
IRB Expiration Date	January 23, 2018
Level of Review	Exempt
Project Status	Active - Open to Enrollment
Exempt Categories	45 CFR 46.101b(2)
Risk Level	Minimal Risk
Internal Funding	Personal funds

The principal investigator (PI) is responsible for all aspects and conduct of this study. The PI must comply with the following conditions of the approval:

- No subjects may be involved in any study procedure prior to the IRB approval date or after the expiration date.
- 2. All unanticipated problems and deviations must be reported to the IRB within 5 business days.
- 3. All changes must be IRB approved prior to implementation unless they are intended to reduce immediate risk.
- 4. All recruitment materials and methods must be approved by the IRB prior to being used.
- 5. The Annual Exempt Form must be submitted to the IRB for review and approval at least 30 days prior to the project expiration date. If the study is complete, the Completion/Withdrawal Form may be submitted in lieu of the Annual Exempt Form
- 6. Maintain all research records for a period of seven years from the project completion date.
- Utilize all approved research documents located within the attached files section of eCompliance. These documents are highlighted green.

If you are offering subject payments and would like more information about research participant

payments, please click here to view the MU Business Policy and Procedure: http://bppm.missouri.edu/chapter2/2_250.html

If you have any questions, please contact the IRB at 573-882-3181 or irb@missouri.edu.

Thank you, MU Institutional Review Board



190 Galena Hall; Dc074.00 Columbia, MO 65212 573-882-3181 irb@missouri.edu

March 2, 2017

Principal Investigator: Iuliia Alieva Department: Journalism-MA

Your Exempt Amendment Form to project entitled DATA JOURNALISM AND DEMOCRACY: HOW AMERICAN MASS MEDIA FRAMED THE 2016 PRESIDENTIAL CAMPAIGN IN THE UNITED STATES USING DATA VISUALIZATION was reviewed and approved by the MU Institutional Review Board according to the terms and conditions described below:

IRB Project Number	2007528
IRB Review Number	224429
Initial Application App	roval Date January 23, 2017
Approval Date of this R	eview March 02, 2017
IRB Expiration Date	January 23, 2018
Level of Review	Exempt
Project Status	Active - Open to Enrollment
Exempt Categories	45 CFR 46.101b(2)
Risk Level	Minimal Risk
Internal Funding	Personal funds

The principal investigator (PI) is responsible for all aspects and conduct of this study. The PI must comply with the following conditions of the approval:

- No subjects may be involved in any study procedure prior to the IRB approval date or after the expiration date.
- 2. All unanticipated problems and deviations must be reported to the IRB within 5 business days.
- 3. All changes must be IRB approved prior to implementation unless they are intended to reduce immediate risk.
- 4. All recruitment materials and methods must be approved by the IRB prior to being used.
- 5. The Annual Exempt Form must be submitted to the IRB for review and approval at least 30 days prior to the project expiration date. If the study is complete, the Completion/Withdrawal Form may be submitted in lieu of the Annual Exempt Form
- 6. Maintain all research records for a period of seven years from the project completion date.
- Utilize all approved research documents located within the attached files section of eCompliance. These documents are highlighted green.

If you are offering subject payments and would like more information about research participant payments, please click here to view the MU Business Policy and Procedure: http://bppm.missouri.edu/chapter2/2_250.html

If you have any questions, please contact the IRB at 573-882-3181 or irb@missouri.edu.

Thank you, MU Institutional Review Board

Appendix B: Participant Information

In the research, all the personal information about participants such as gender, age, job title and name of the media organizations is being protected for this thesis project. The descriptions are as follows:

- No. 1: Graphics Reporter at U.S. major newsroom of general interest.
- No. 2: Graphics Editor at U.S. major newsroom of general interest.
- No. 3: Graphics Editor at U.S. major newsroom of general interest.
- No. 4: Graphics Editor at U.S. major newsroom of general interest.
- No. 5: Interactive News Editor at U.S. major newsroom of general interest.
- No. 6: Graphics Editor at U.S. major newsroom of general interest.
- No. 7: Graphics Editor at U.S. major newsroom of general interest.

Appendix C: Email Protocol

Initial Email: Dear Mr./Ms. [NAME],

My name is Iuliia, I'm a graduate student at the University of Missouri School of Journalism and now I'm doing my thesis research in data journalism field. My research will be about the meaning of data visualization for journalists in coverage of presidential campaign 2016 in the United States. I would like to explore how journalists make decisions about this issue for infographics and analyze how successfully they use new methods and techniques of data visualization in their reporting, and what limitation of such way of reporting they deal with.

Considering your expertise, I thought you could contribute to my research, participating in the Skype interview and share your professional experience. There are 10-12 journalists who I'm looking for and the interview will take about 40 minutes. Please, find attached the list of questions for the interview.

As we move forward, I would also need to schedule the interview, please, let me know when is the most convenient time for you to talk.

Please, find a consent document attached in this e-mail with the description of procedures for the research and contact information of IRB and researchers

If you have any questions, please, don't hesitate to ask.

Thanks in advance,

Iuliia Alieva

+1 (573) 554 6580

Skype: alieva_julia

Appendix D: Interview Script

Introductory section of conversation:

Hello [Name]! My name is Iuliia and I'm a graduate student at the University of Missouri School of Journalism. Now I'm doing my thesis research about data visualization of presidential campaign and elections in the United States last year. Thank you very much for your time and effort to contribute to my research. I sent you all the information about research and the list of questions in the first e-mail and I hope, you had time to check it. Do you have any questions before we start?

Interview questions

- Could you please describe the structure of your Data desk? How many people are working there? Is it separated department? How work is divided?
- What skills do journalists need to produce effective data visualization in your newsroom? Do you have special trainings for journalists or skills sharing inside the newsroom?
- What is the purpose of using data visualization techniques for the coverage of the 2016 U.S. presidential campaign?
- What kind of data was the best for infographics during the 2016 U.S. presidential campaign?
- What methods of data visualization (mapping, graphs, charts, other) did you use the most for the presentation of information during the 2016 U.S. presidential campaign?
- What stories did you discover while working with data and visualizing it?

- How did data-visualization techniques help you improve your reporting?
- What challenges did you struggle with in visualizing the data?
- Did data visualization produce the desirable outreach and effect on audience?
- What limitations did this tool have for the presentation of content during the 2016
 U.S. presidential campaign?
- In what cases did you combine data visualizations with traditional reporting?
- How did newsrooms make decisions about the components and elements of datavisualization content?
- Did you form the topic of the content using data you already have or did you look for data with a formulated topic in mind?
- What was the process for deciding which aspects and which perspective to represent and emphasize in data visualizations during presidential campaign?
- In your opinion, what were the main frames/trends in data visualization during the presidential campaign and right after the election?
- What results were reached with data visualizations before and after the 2016 U.S. presidential campaign? (audience feedback)
- What would you do or suggest to other journalists to do differently in future to improve their data reporting?
- Is there anything you would like to add which I didn't cover in my questions?