

Visualización de datos para la investigación científica

Yusnelkis Milanés Guisado. PhD.

Servicio de Datos de Investigación

Biblioteca /CRAI



UNIVERSIDAD
**PABLO DE
OLAVIDE**
SEVILLA



Programa

- Relevancia de la visualización de datos. Principios, procesos y buenas prácticas.
- Tablas y gráficos en tu investigación.
- Narrativa de datos
- Práctica a través de estudios de caso

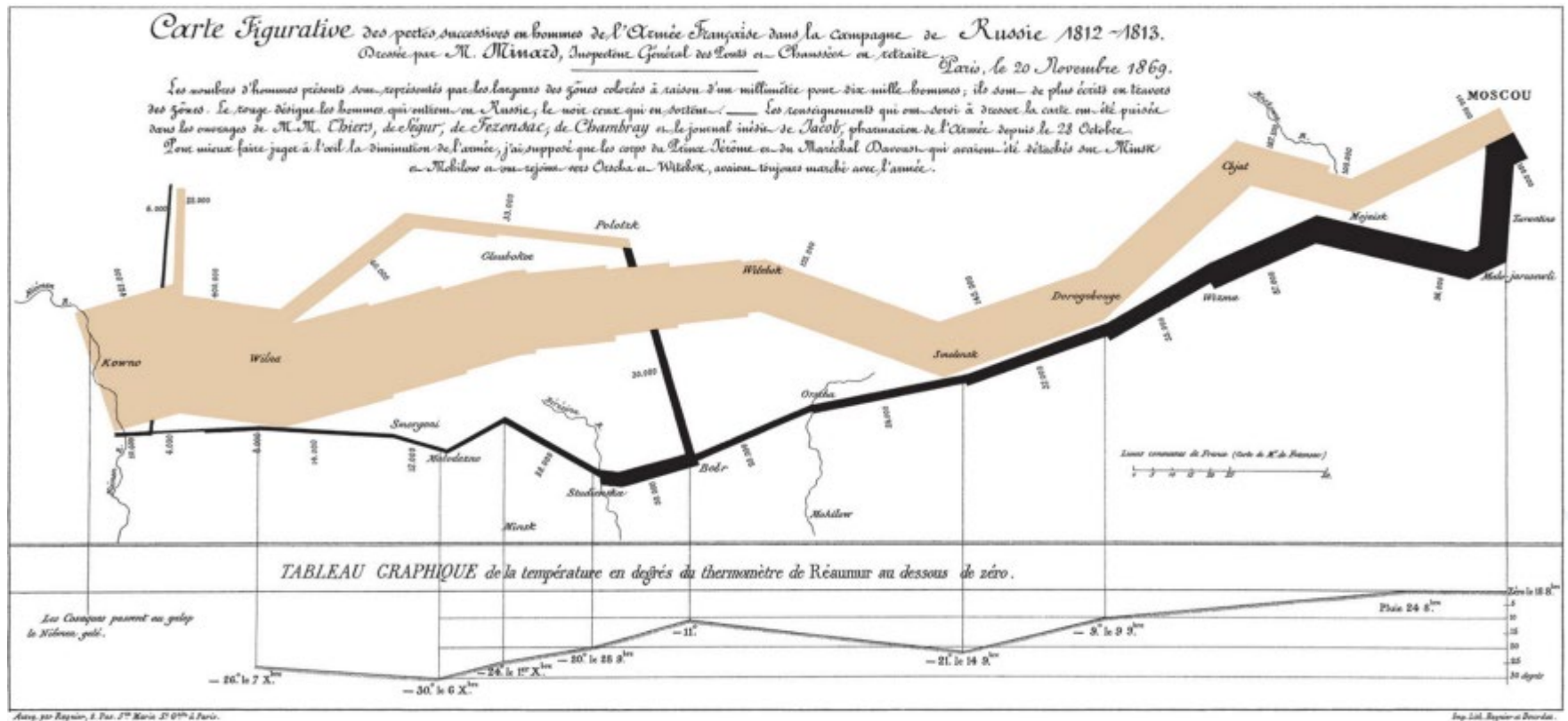


Relevancia de la visualización de datos. Principios, procesos y buenas prácticas.



Historia de la Visualización

Drawn in **1869** one of the most cited examples of statistical graphics occurred when Charles Minard mapped Napoleon's invasion of Russia. The map depicted the size of the army as well as the path of Napoleon's retreat from Moscow – and tied that information to temperature and time scales for a more in-depth understanding of the event.



Chales Joseph Minard — map of Napoleon's Russia campaign — [Source](#)



COVID-19 Dashboard by the Center for Systems Science and Engineering (CSSE) at Johns Hopkins University (JHU)



Last Updated at (M/D/YYYY)
19/12/2021 20:21

Total Cases
274.640.402

Total Deaths
5.353.784

Total Vaccine Doses Administered
8.678.554.614

Cases | Deaths by
Country/Region/Sovereignty

US
28-Day: **3.070.889** | **34.496**
Totals: **50.825.950** | **806.375**

United Kingdom
28-Day: **1.485.936** | **3.326**
Totals: **11.425.656** | **147.679**

Germany
28-Day: **1.438.841** | **9.171**
Totals: **6.812.538** | **108.285**

France
28-Day: **1.183.755** | **2.996**
Totals: **8.730.140** | **122.478**

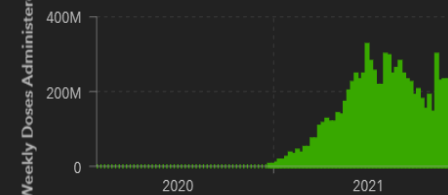
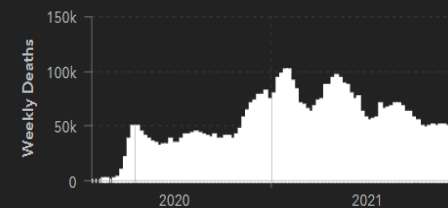
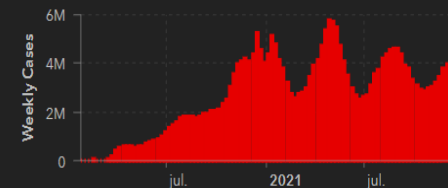
Russia
28-Day: **874.717** | **32.449**
Totals: **10.037.546** | **291.336**

Poland

28-Day Cases
17.046.865

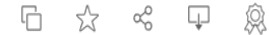
28-Day Deaths
202.624

28-Day Vaccine Doses Administered
930.975.108

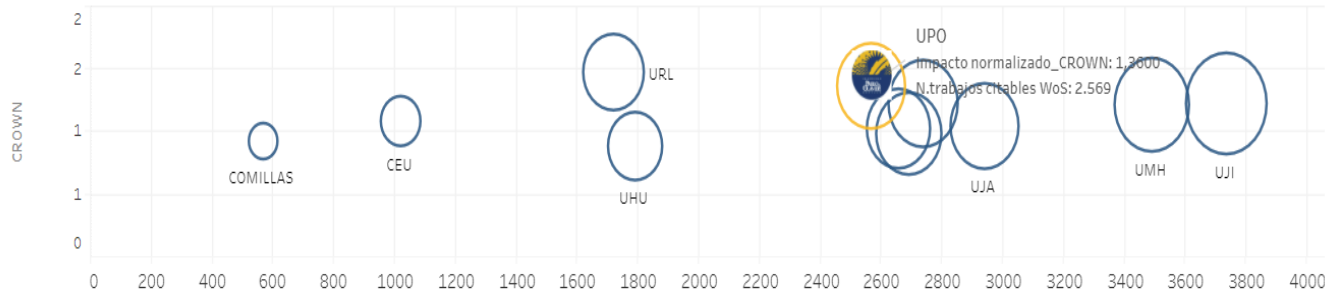


<https://coronavirus.jhu.edu/map.html>

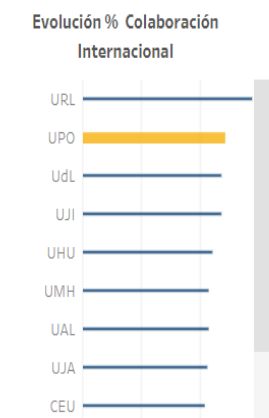
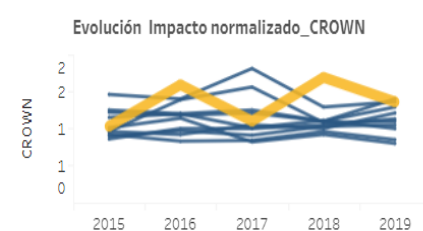
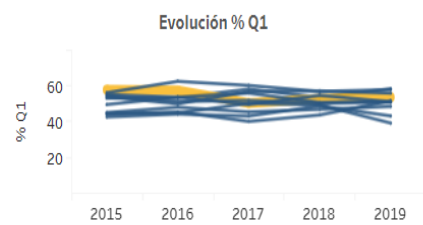
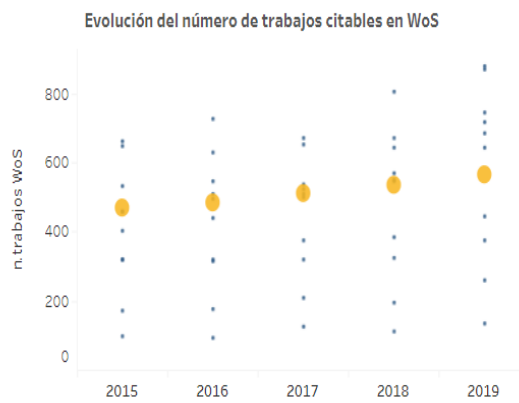
ComparativaNacional20152019_UPO_200921_V0 de [Apoyo a la Investigación. Biblioteca/CRAI UPO](#)



Posición de universidades españolas (normalizado por número de PDI*), según Impacto normalizado_CROWN** y número de trabajos publicados en la Web of Science (2015-2019).



- Publication Year
- 15_19
 - 2015
 - 2016
 - 2017
 - 2018
 -



- Universidades
- (Todo)
 - CEU
 - COMILLAS
 - UAL
 - UdL
 - UHU
 - UJA
 - UJI
 - ULE
 - UMH

*PDI: Profesor Docente Investigador

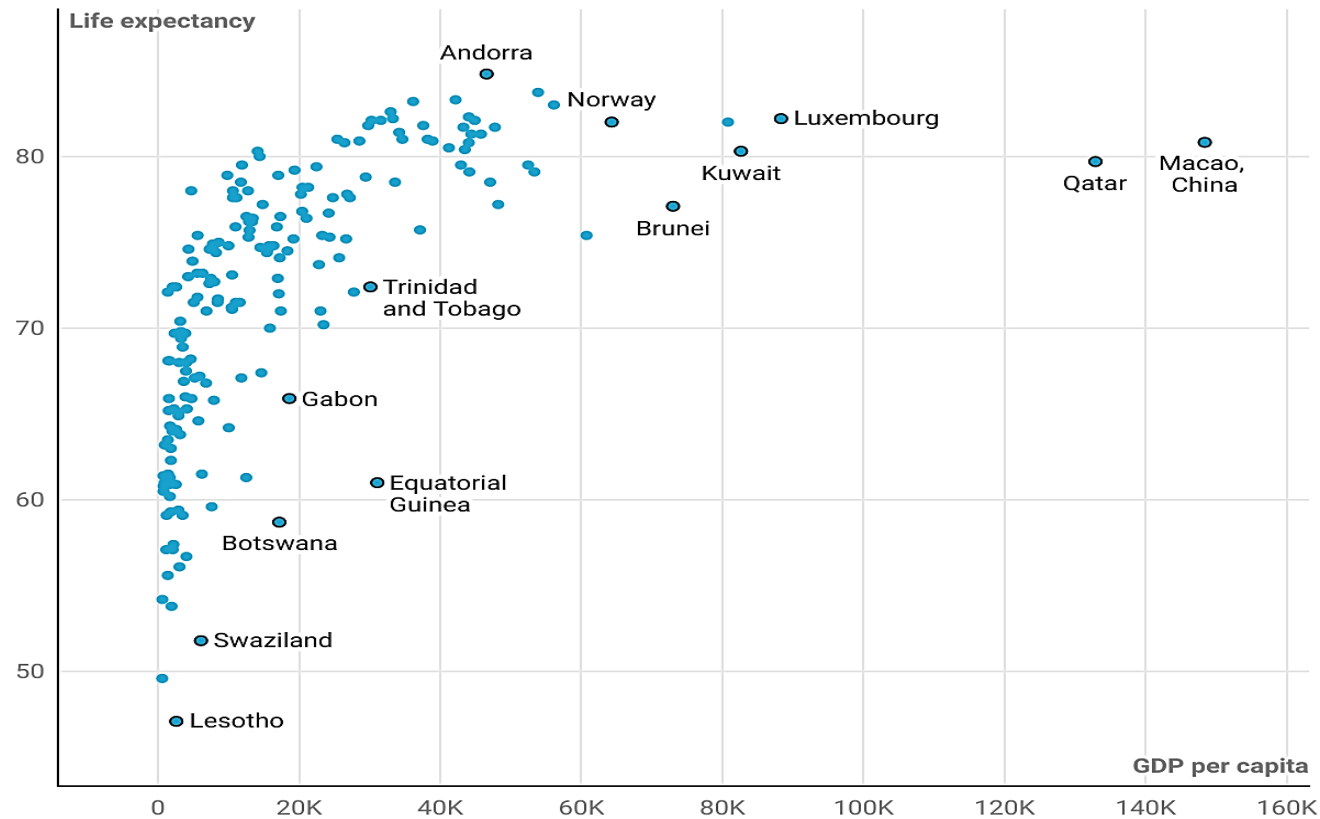
** Indicador CROWN: Mide el impacto en comparación con la media mundial. Tiene en cuenta el campo disciplinar, tipo documental y año de publicación. Por ejemplo, un CROWN de 0.9 significa que la Universidad se sitúa un 10% por debajo de la media mundial y un 1.2 indica que la universidad recibe una citación superior en un 20% de la media mundial



Datawrapper

[GDP per capita vs Life expectancy]

GDP per capita in US-Dollars and life expectancy in years for selected countries, 2015.



Source: Gapminder • Created with Datawrapper

Visualización científica / mala práctica

IK



100,928 articles are raw data collected from SCI-E. We used the visual analysis software Vantagepoint to process the raw data and conduct statistical analysis. After statistical analysis, the total scientific production, institutions, researchers, total times cited, and highly cited were counted and obtained, analyzed, and shown in Figures 2–6, 8. In addition, we reprocessed the statistical results applying formulas 1–3 to show the results in Figures 7 and 9.

3.1. The Development of Scientific Production

Figure 2 exhibits the temporal characteristics of scientific production in the area of solar energy during the period 2000–2017; it indicates that scientific publications consistently increased throughout the period. The number of scientific publications increased from 69 in 2000 to 3928 in 2017; there is a cumulative increase of 570%, and the average annual growth rate is 33% during this phase. The development of scientific production in solar energy can be divided into two stages: the incubation stage and the burgeoning stage. In the incubation stage (2000–2005), scientific production in solar energy experienced a steady increase except 2002, which saw a weak downturn. This was followed by a significant increase in the burgeoning stage (2006–2017). During the burgeoning stage, the scientific production continued to rise, with an average annual growth rate up to 46%.

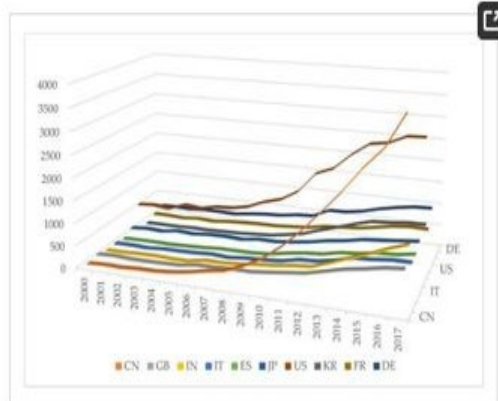


Figure 2. Temporal characteristics of China's scientific production in solar energy. (CN: China, GB: England, IN: India, IT: Italy, ES: Spain, JP: Japan, US: United States, KR: Korea, FR: France, DE: Germany)

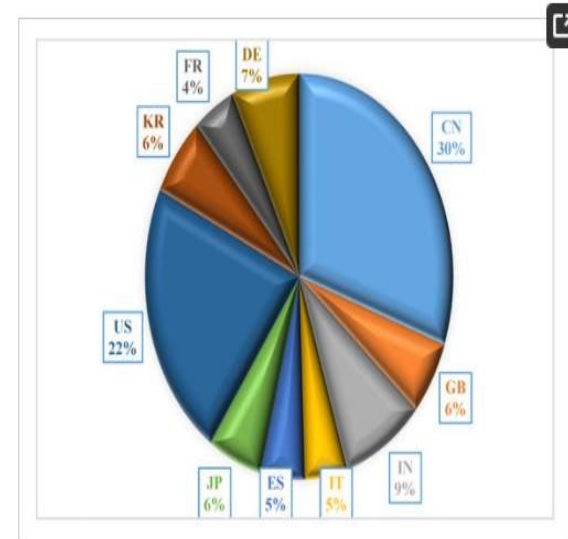
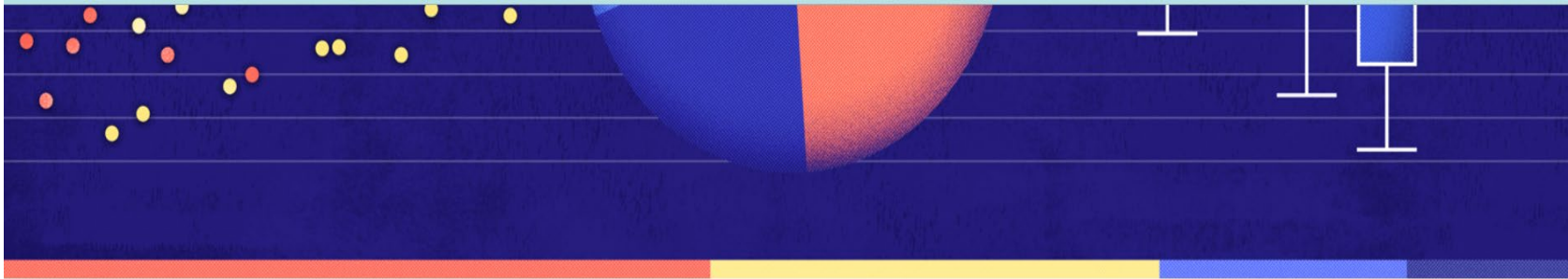


Figure 3. Scientific production of top 10 countries in 2017 by percentage.

<https://doi.org/10.3390/su11030623>

Visualización científica



Common data visualization forms include the pie chart, box plot and scatterplot. Depending on the data, some are more successful at communicating than others.

CREDIT: KNOWABLE MAGAZINE

— THE MIND —

Why scientists need to be better at data visualization

The scientific literature is riddled with bad charts and graphs, leading to misunderstanding and worse. Avoiding design missteps can improve understanding of research.

*...science is littered with **poor data visualizations that confound readers and can even mislead the scientists who make them**. Deficient data visuals can reduce the quality and impede the progress of scientific research...*

Visualización científica/ Nature resources

nature portfolio View all journals Search Login

nature > search

Search

Search [Advanced search](#)

Journal **Article type** **Subject** **Date**

Nature Methods (7) **All** **All** **All** [Clear all filters](#)

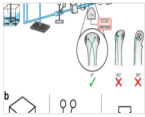
Showing 1-7 of 7 results Sort by Date — oldest first

Advanced filters: **Title: "points of view"** [Clear advanced filters](#)

News
28 Jan 2011
Nature Methods
Volume: 8, P: 101

Points of view: Points of review (part 1)

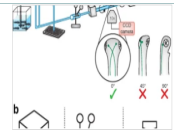
Bang Wong



News
28 Jan 2011
Nature Methods
Volume: 8, P: 101

Points of view: Points of review (part 1)

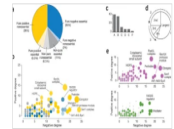
Bang Wong



News
25 Feb 2011
Nature Methods
Volume: 8, P: 189

Points of view: Points of review (part 2)

Bang Wong



News
30 Mar 2011
Nature Methods
Volume: 8, P: 277

Points of view: Typography

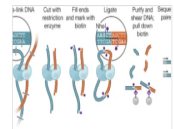
Bang Wong



News
28 Apr 2011
Nature Methods
Volume: 8, P: 365

Points of view: The overview figure

Bang Wong



https://www.nature.com/search?title=%22points%20of%20view%22&order=date_asc&journal=nmeth

Visualización científica

EMBO reports | EMBOpress | JOURNALS | This Journal | Search journal | LOGIN

CURRENT ISSUE | ABOUT | INFORMATION | ARCHIVE | ALERTS | SUBMIT

Science & Society | 14 October 2019 | [FREE ACCESS](#)

Preparing scientists for a visual future

Visualization is a powerful tool for research and communication but requires training and support

Shraddha Nayak, Janet H Iwasa

[Author Information](#)

EMBO Reports (2019) 20: e49347 | <https://doi.org/10.15252/embr.201949347>

PDF

Tools | Share

Picture a visual interface for exploring phosphatases. A search tool lets you select a phosphatase and a substrate. The display allows you to toggle between interactive and multiscale views of the phosphatase structure, its hydrophobicity, electrostatic nature, active site conformation, and other attributes, and how these change upon substrate binding. Along with structural alterations, the application shows possible downstream events, such as the activation or inactivation of signaling pathways or transcriptional regulation of specific genes. Visual interfaces like these can lead to an intuitive understanding of how different proteins work, and, in this case, provide clarity into the functional consequences of a phosphatase binding to one substrate versus another.



[About the cover](#)

Volume 20

Issue 11

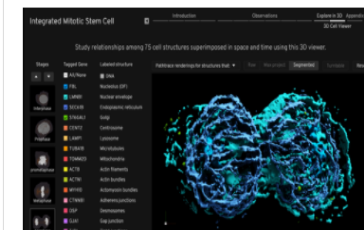
5 November 2019

[IN THIS ISSUE](#)

FIGURES | REFERENCES | RELATED | DETAILS

Figure 1.

Analysis using Allen Cell Explorer



<https://www.embopress.org/doi/full/10.15252/embr.201949347>

Visualización científica



CIBC HOME ABOUT THE CIBC ▾ RESEARCH ▾ PERSONNEL ▾ SOFTWARE ▾ PUBLICATIONS EVENTS CONTACT ▾ SCI HOME

The NIH/NIGMS Center for Integrative Biomedical Computing

Top Scientific Visualization Research Problems



Scientific visualization as currently understood and practiced is still a relatively new discipline. As a result, we visualization researchers are not necessarily accustomed to undertaking the sorts of self-examinations that other scientists routinely undergo in relation to their work. Yet if we are to create a disciplinary culture focused on matters of real scientific importance and committed to real progress, it is essential that we ask ourselves hard questions on an ongoing basis. What are the most important research issues facing us? What underlying assumptions need to be challenged and perhaps abandoned? What practices need to be reviewed? In this article, I attempt to start a discussion of these issues by proposing a list of top research problems and issues in scientific visualization. [\[PDF version\]](#)

"Study the science of art and the art of science"—Leonardo da Vinci

<https://www.sci.utah.edu/cibc-research/highlights/24-cibc-highlights/253-top-scientific-visualization-research-problems.html>

Visualización científica

Map a research topic

Get an overview - Find papers - Identify relevant concepts



- PubMed (life sciences)
- BASE (all disciplines)

Refine your search ▾

bioinformatic

GO

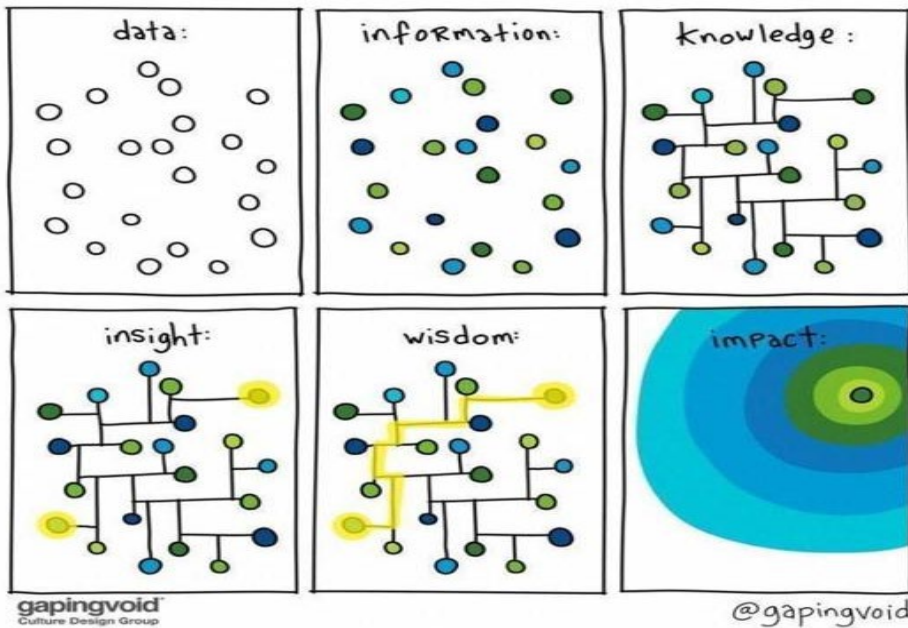
Try out: [digital education](#) [climate change](#) AND [impact](#)

<https://openknowledgemaps.org/index.php>

¿Porqué La Visualización?

Via
@gapingvoid

[Ver traducción](#)



- ✓ **Los datos son sólo un medio para plasmar la realidad . Siempre se usan como sustitutos de algo.**
- ✓ **Para comprender los datos, debemos mirar a través de ellos, darles valor.**
- ✓ **Le podemos dar forma, contexto, para innovar, transformarlos en información y conocimiento, en decisiones, que generan un impacto a diferentes niveles**

¿Porqué La Visualización?

1 DATOS



2 LIMPIOS EN UNA BASE DE DATOS



3 ANALIZADOS



4 PRESENTADOS DE FORMA VISUAL



5 EXPLICADOS CON UNA HISTORIA



¿Porqué La Visualización?

Comunicación efectiva a través de la Visualización

How many P's can you find in the text?

A M C D F G O I S H P O F Q H O P I U O I U L F K S D K
F K F J Z F C P T H B M U G I N D I C A T O R S D H D X
B E W Z C O M P O S I T E A E T F R J L M N O J K P B R
L A D X O G F J E I L K S A P R P U E D G H M X O Q B I

¿Porqué La Visualización?

Comunicación efectiva a través de la Visualización

How many P's can you find in the text?

A M C D F G O I S H P O F Q H O P I U O I U L F K S D K
F K F J Z F C P T H B M U G I N D I C A T O R S D H D X
B E W Z C O M P O S I T E A E T F R J L M N O J K P B R
L A D X O G F J E I L K S A P R P U E D G H M X O Q B I

¿Porqué La Visualización?

Comunicación efectiva a través de la Visualización

Grupo de pacientes: Discordantes

	LPS	ADNr16S	sCD14	iFABP	CD4_CD38_DR	CD4_38_T	CD4_DR_T	CD8_CD38_DR	CD8_38_T	CD8_DR_T	sCD163
LPS	1.00	0.18	0.00	-0.27	0.11	-0.20	0.05	0.26	0.14	0.25	0.10
ADNr16S	0.18	1.00	-0.22	0.12	0.19	-0.07	0.17	0.23	0.15	0.21	0.08
sCD14	0.00	-0.22	1.00	0.16	0.16	0.17	0.02	0.14	0.26	0.12	0.18
iFABP	-0.27	0.12	0.16	1.00	-0.04	-0.12	0.05	-0.15	-0.12	-0.02	-0.12
CD4_CD38_DR	0.11	0.19	0.16	-0.04	1.00	0.54	0.72	0.74	0.63	0.59	0.02
CD4_38_T	-0.20	-0.07	0.17	-0.12	0.54	1.00	0.03	0.53	0.69	0.08	-0.10
CD4_DR_T	0.05	0.17	0.02	0.05	0.72	0.03	1.00	0.46	0.22	0.70	0.15
CD8_CD38_DR	0.26	0.23	0.14	-0.15	0.74	0.53	0.46	1.00	0.91	0.68	0.08
CD8_38_T	0.14	0.15	0.26	-0.12	0.63	0.69	0.22	0.91	1.00	0.47	-0.01
CD8_DR_T	0.25	0.21	0.12	-0.02	0.59	0.08	0.70	0.68	0.47	1.00	0.21
sCD163	0.10	0.08	0.18	-0.12	0.02	-0.10	0.15	0.08	-0.01	0.21	1.00

✓ Dificultad para detectar relaciones

✓ Dificultad para comparar, extraer diferencias...

Grupo de pacientes: Sanos

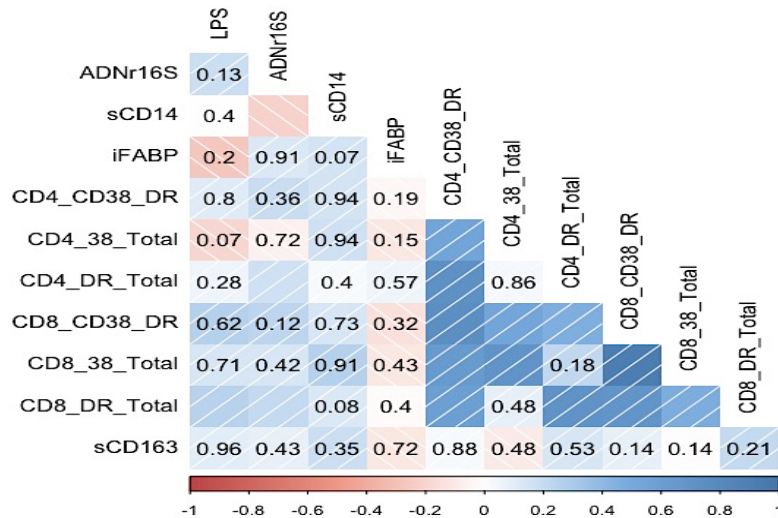
	LPS	ADNr16S	sCD14	iFABP	CD4_CD38_DR	CD4_38_T	CD4_DR_T	CD8_CD38_DR	CD8_38_T	CD8_DR_T	sCD163
LPS	1.00	0.23	-0.26	-0.09	-0.34	-0.30	-0.22	-0.14	-0.12	-0.28	0.37
ADNr16S	0.23	1.00	-0.44	-0.05	-0.35	-0.33	-0.21	-0.39	-0.34	-0.26	-0.07
sCD14	-0.26	-0.44	1.00	0.23	0.47	0.26	0.33	0.26	0.23	0.27	0.03
iFABP	-0.09	-0.05	0.23	1.00	0.11	0.37	-0.18	0.25	0.41	-0.09	-0.10
CD4_CD38_DR	-0.34	-0.35	0.47	0.11	1.00	0.53	0.73	0.77	0.55	0.76	-0.23
CD4_38_T	-0.30	-0.33	0.26	0.37	0.53	1.00	-0.03	0.58	0.71	0.23	-0.28
CD4_DR_T	-0.22	-0.21	0.33	-0.18	0.73	-0.03	1.00	0.43	0.07	0.77	-0.04
CD8_CD38_DR	-0.14	-0.39	0.26	0.25	0.77	0.58	0.43	1.00	0.83	0.67	-0.07
CD8_38_T	-0.12	-0.34	0.23	0.41	0.55	0.71	0.07	0.83	1.00	0.32	-0.12
CD8_DR_T	-0.28	-0.26	0.27	-0.09	0.76	0.23	0.77	0.67	0.32	1.00	0.04
sCD163	0.37	-0.07	0.03	-0.10	-0.23	-0.28	-0.04	-0.07	-0.12	0.04	1.00

Working paper. HIV Research group. Virgen del Rocio Hospital.

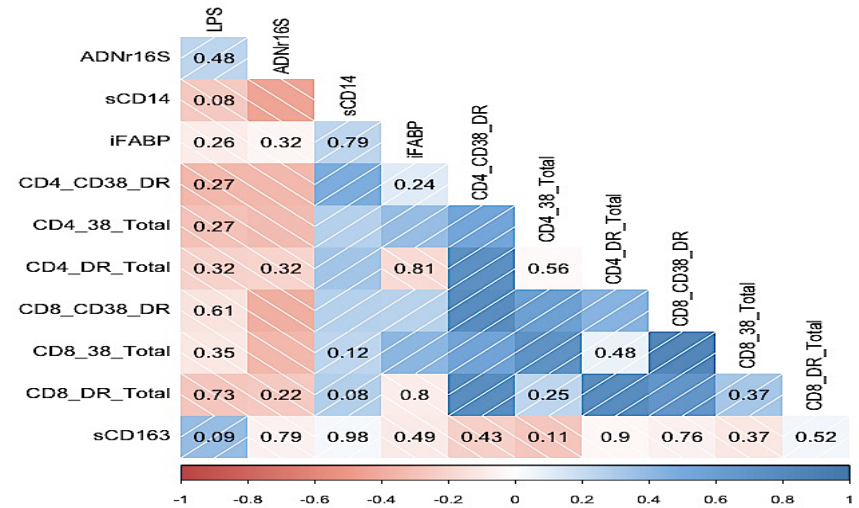
¿Porqué La Visualización?

Comunicación efectiva a través de la Visualización

Discordantes



Sanos



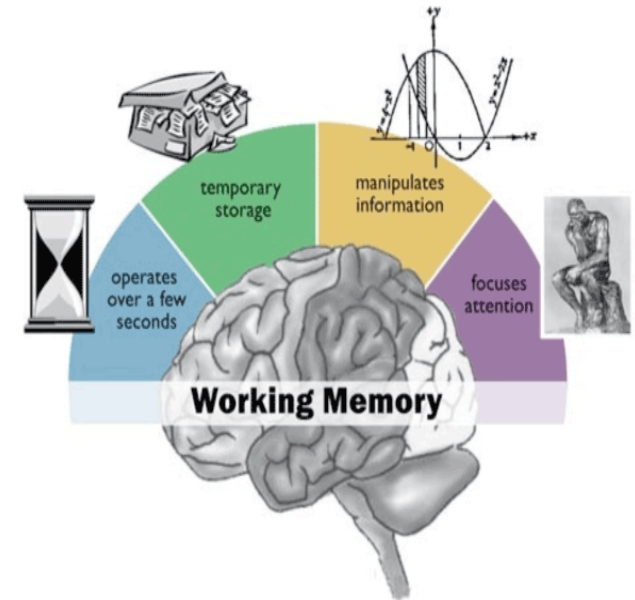
- ✓ Facilidad para detectar patrones (Relaciones positivas y negativas, color y orientación)
- ✓ Comparación entre grupos (Más relaciones negativas en Sanos, etc)

Working paper. HIV Research group. Virgen del Rocío Hospital.

¿Porqué La Visualización?

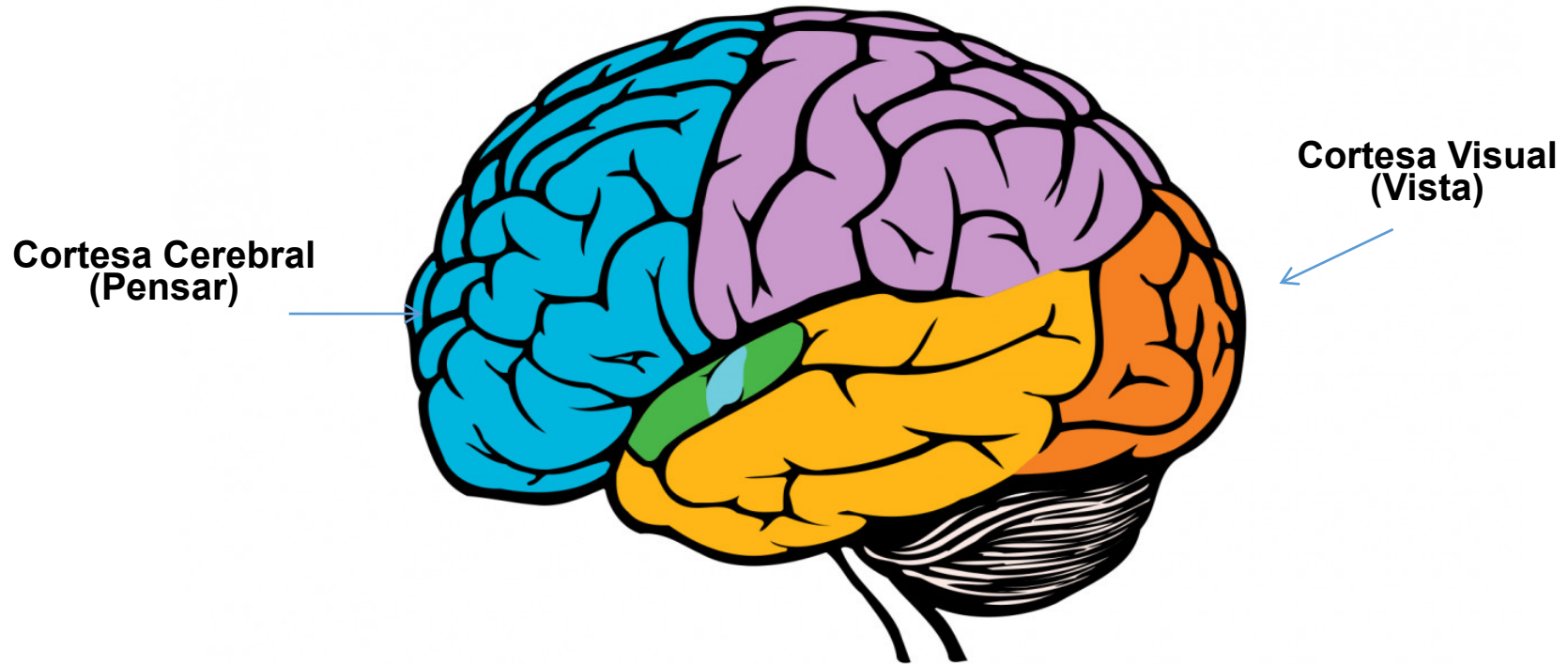
“Working Memory”

- La visualización ataca nuestros sentidos, los que utilizan la “Working memory”.
- El cerebro humano sólo puede retener y almacenar entre 5 y 8 números, elementos ...
- Somos buenos comparando alturas, distancias..



¿Porqué La Visualización?

Pensar es lento vs La vista es más rápida



El cerebro ama las diferencias....



¿Porqué La Visualización?

La visualización utiliza primero los ojos, luego surgen los conceptos.

¿ Cómo ocurre esto ?



LENGUAJE VISUAL

EL ojo es sensible a los patrones de variaciones de los colores, las formas, las distancias, etc.

*El cerebro humano es capaz de procesar una imagen completa en **13 milésimas de Segundo***

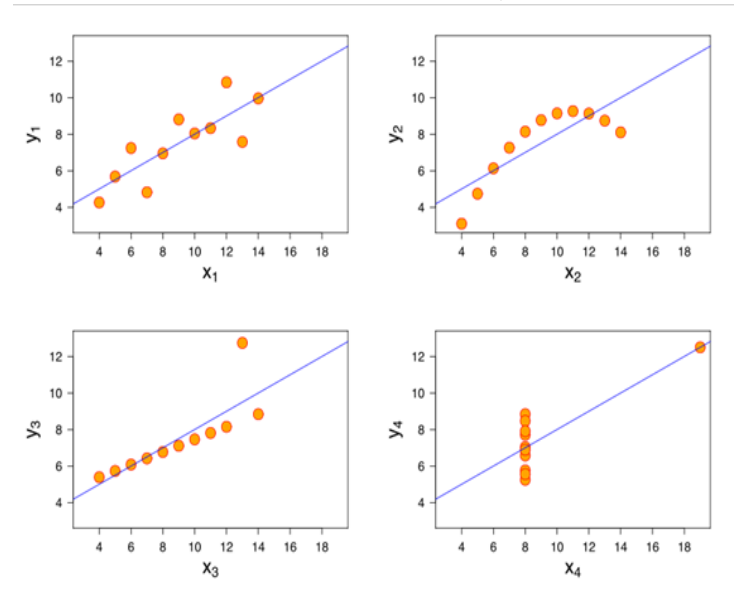
*A su vez es capaz de procesar una imagen **60000 veces más rápido** que un texto*



¿Porqué La Visualización?

- Un resumen estadístico puede no contemplar tendencias importantes en un dataset.
- Puede preservar la complejidad, o presentar las multiples miradas de un dataset.
- Puede funcionar como una importante primera etapa de investigación en una nueva área de estudio

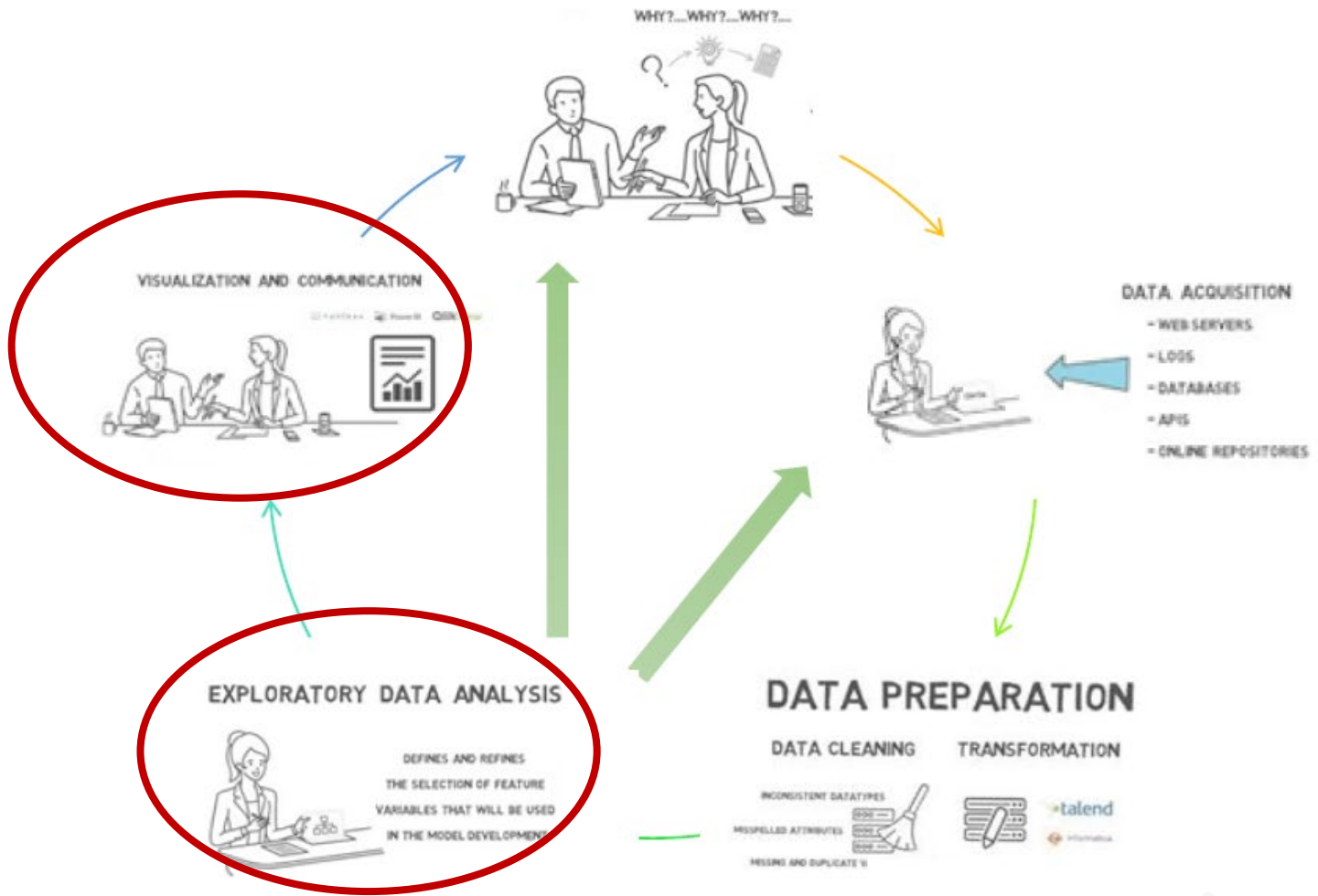
Anscombe's Quartet



http://en.wikipedia.org/wiki/Anscombe%27s_quartet



La Visualización de datos





La Visualización de datos

Conquer the **Last Mile** in the **Analytics Marathon**



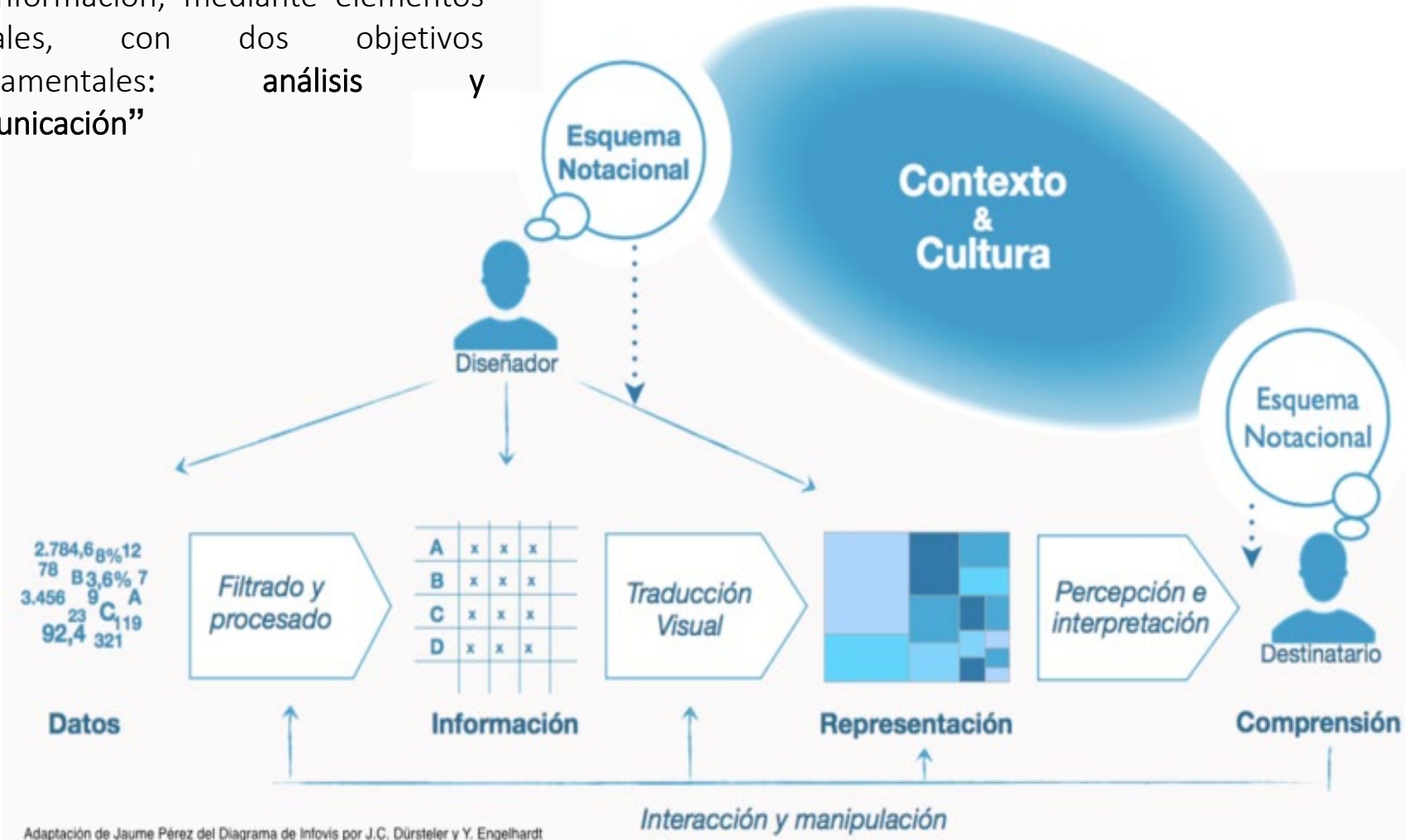


Visualización:

Asimilar la mayor cantidad de información en el menor tiempo posible

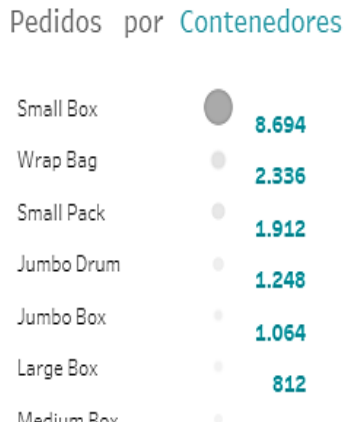
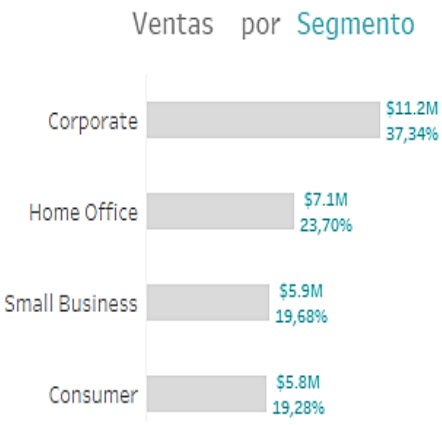
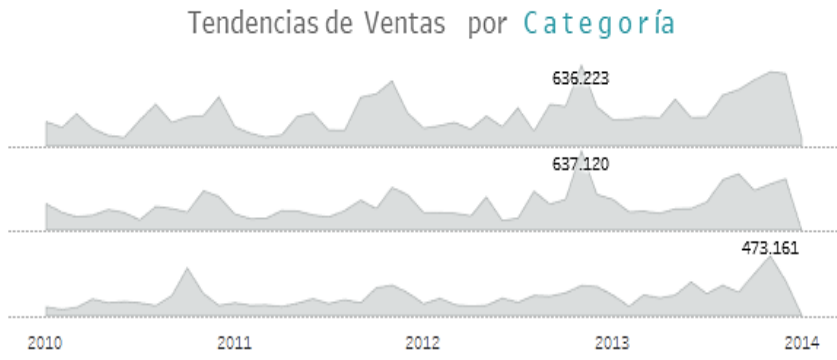
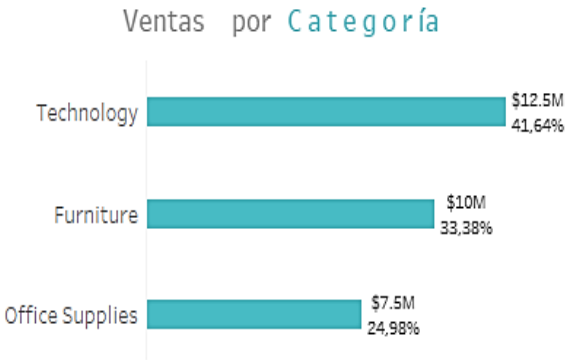
La Visualización de datos

“proceso de transformación de datos en información, mediante elementos visuales, con dos objetivos fundamentales: análisis y comunicación”



La Visualización de datos

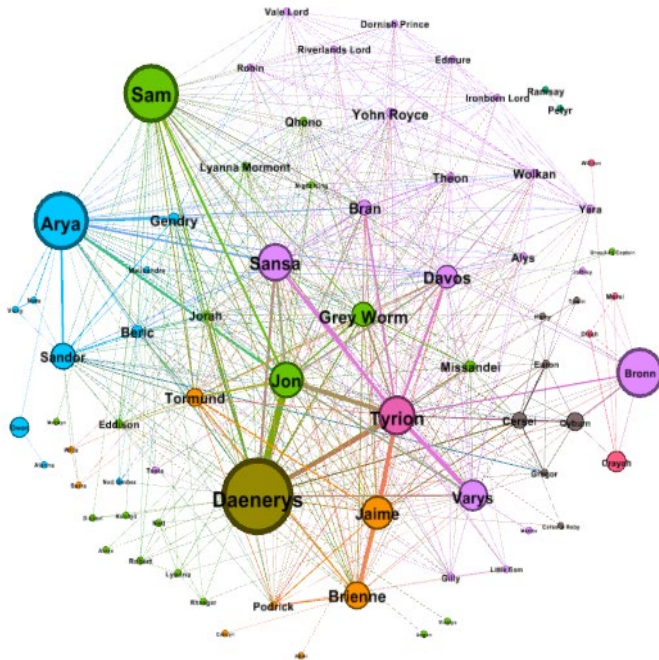
El estado de las ventas



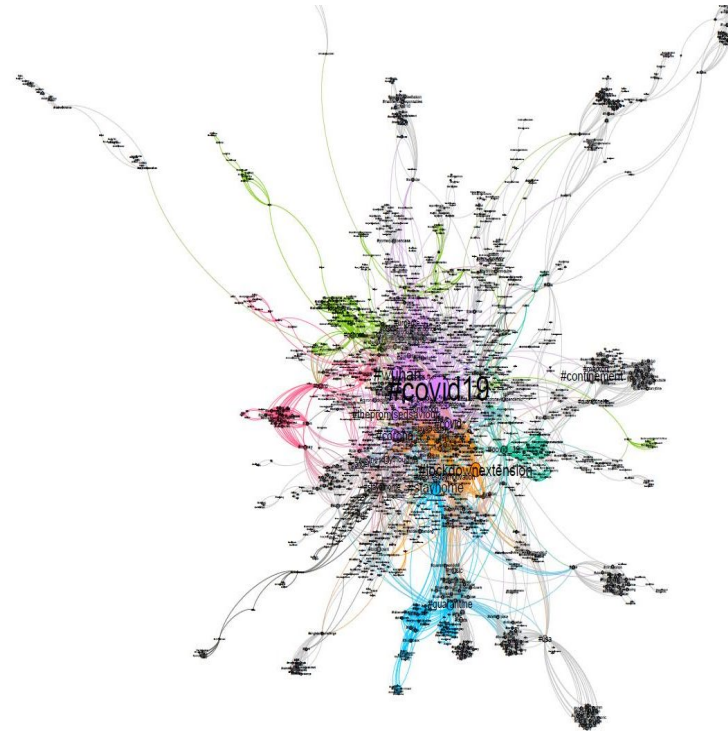
Milanés, Y. Caso práctico resuelto en certificación de Big data analytic de la EOI.
https://public.tableau.com/views/Ejercicio_Tableroventas_Dataviz_200521/Cuadro_Ventas?:language=es-ES&:display_count=n&:origin=viz_share_link

La Visualización de datos | Redes

Objetivo: Realizar una representación adecuada para obtener la mayor cantidad de información posible



! El poder de un personaje!



! Red de comentarios en Twitter con Hastags COVID !

Tipos de Visualización

Analíticas o exploratorias

- Exploración de los datos
 - Se busca entender los datos
 - Detectar patrones y relaciones
- . Uno es el productor y consumidor de las visualizaciones

Aclaratorias o explanatorias

- Comunicar los resultados
- Ayudar a otros a entender el mensaje
- Produces las visualizaciones para una audiencia



La Visualización exploratoria

- ¿Qué puede ser relevante ?
- Buscar una (s) perla (s) entre las ostras



La Visualización exploratoria

Exploración de los datos (EDA)

- **Exploración de los datos.** (E.D.A. exploratory data analysis, J.Tukey)

El objetivo de este análisis exploratorio es:

- a) Buscar posibles relaciones de la variable respuesta/dependiente con la(s) variable(s) explicativa(s);
- b) Considerar la necesidad de aplicar transformaciones de las variables;
- c) Eliminar variables explicativas que estén altamente correlacionadas.

La Visualización aclaratoria o explanatoria

Aclaratorio o comunicar

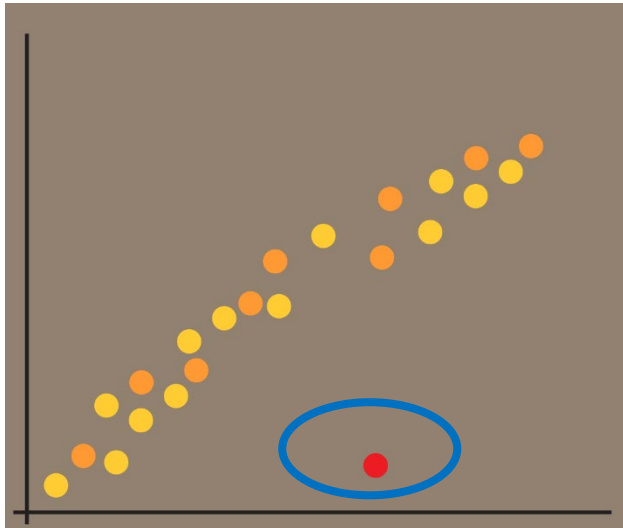
- Comunicas el resultado del análisis a través de tus gráficos
- Algo concreto que mostrar (Las dos perlas, una historia, un hecho, una causa...)
- No mostrar el exploratorio, **MOSTRAR** el aclaratorio (Dedicar tiempo a transformar los datos en info. que pueda ser comprendida, **LAS DOS PERLAS**)
- Abstenerse de la tentación de mostrar cuanto has trabajado !



La Visualización exploratoria

Visualiza (plot, map, or both)

- Relevancia de la visualización exploratoria



Explorar

Plots, trends, timelines, etc.



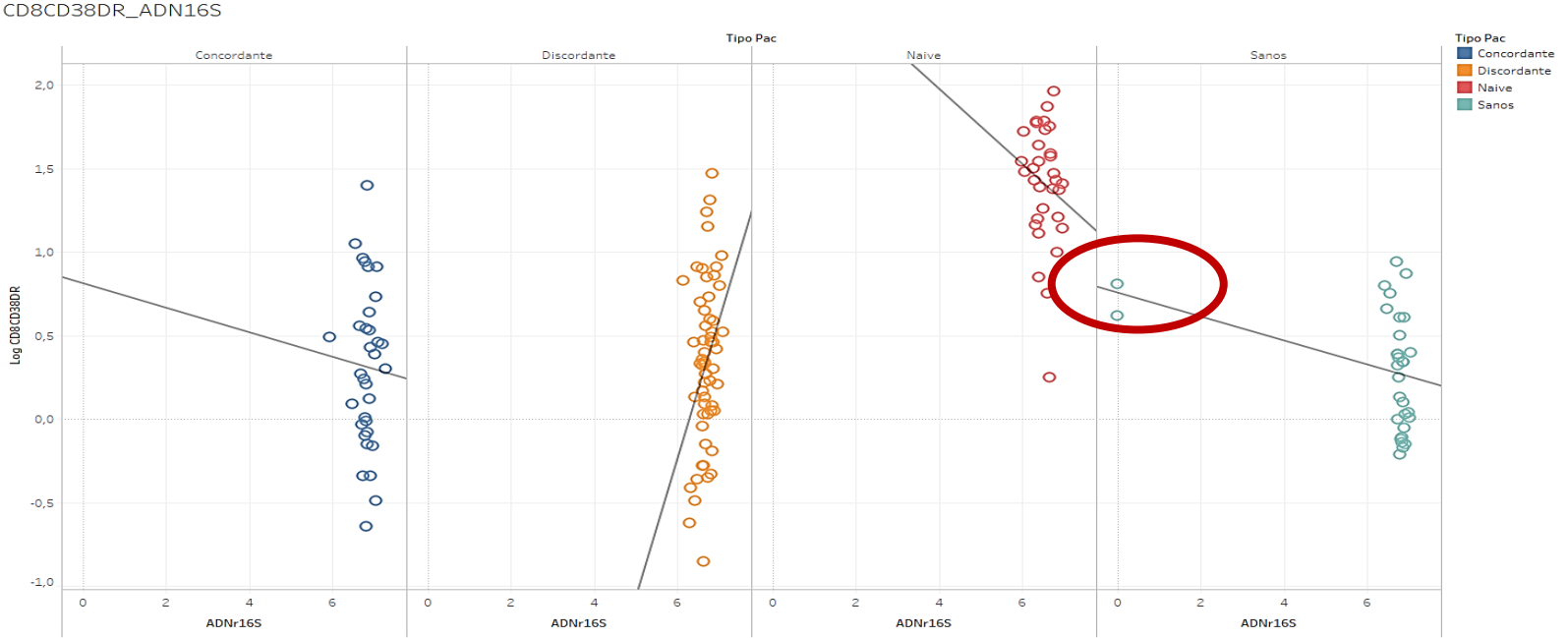
Analizar

Desarrollar y testar hipótesis
Descubrir errores en los datos,
anomalías.

Encontrar patrones

La Visualización exploratoria

Ejemplo 1: Identificación de casos atípicos o outliers

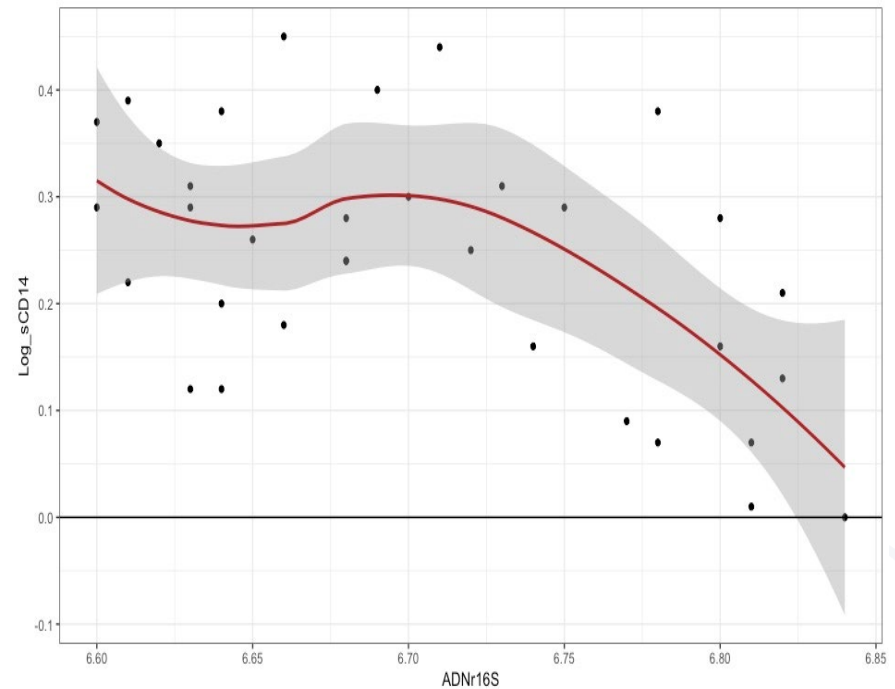
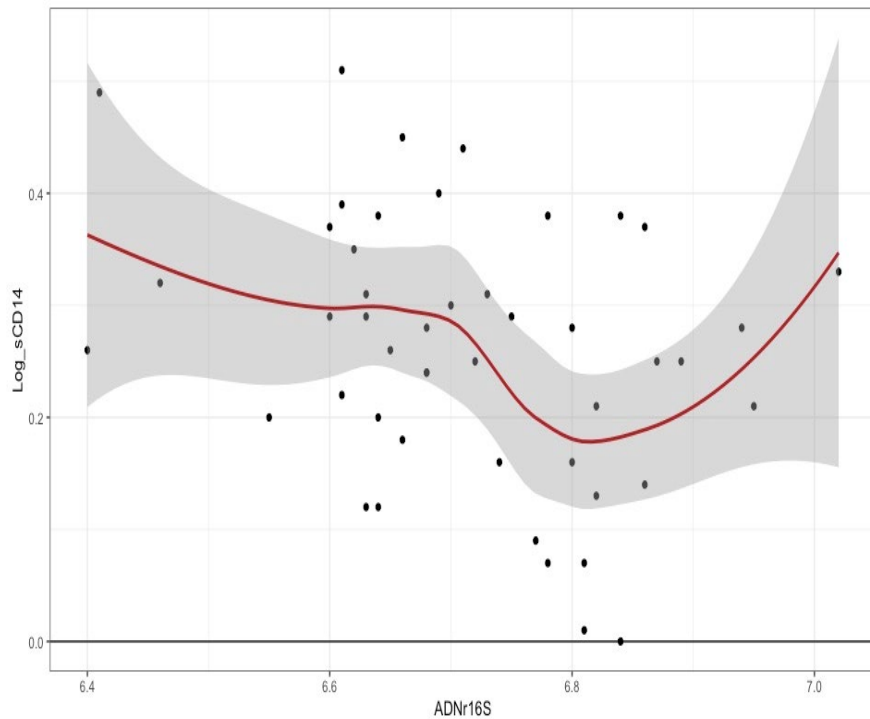


ADNr16S vs. Log CD8CD38DR desglosado por Tipo Pac. El color muestra detalles acerca de Tipo Pac.

Paneles		Línea		Coeficientes				
Fila	Columna	valor p	DF	Término	Valor	StdErr	valor t	valor p
Log CD8CD38DR	Concordante	0,845093	31	ADNr16S	-0,0734909	0,373002	-0,197026	0,845093
Log CD8CD38DR	Discordante	0,015575	57	ADNr16S	0,829366	0,332609	2,49352	0,015575
Log CD8CD38DR	Naive	0,367706	30	ADNr16S	-0,225215	0,24625	-0,914578	0,367706
Log CD8CD38DR	Sanos	0,0541853	29	ADNr16S	-0,0717642	0,0357623	-2,0067	0,0541853

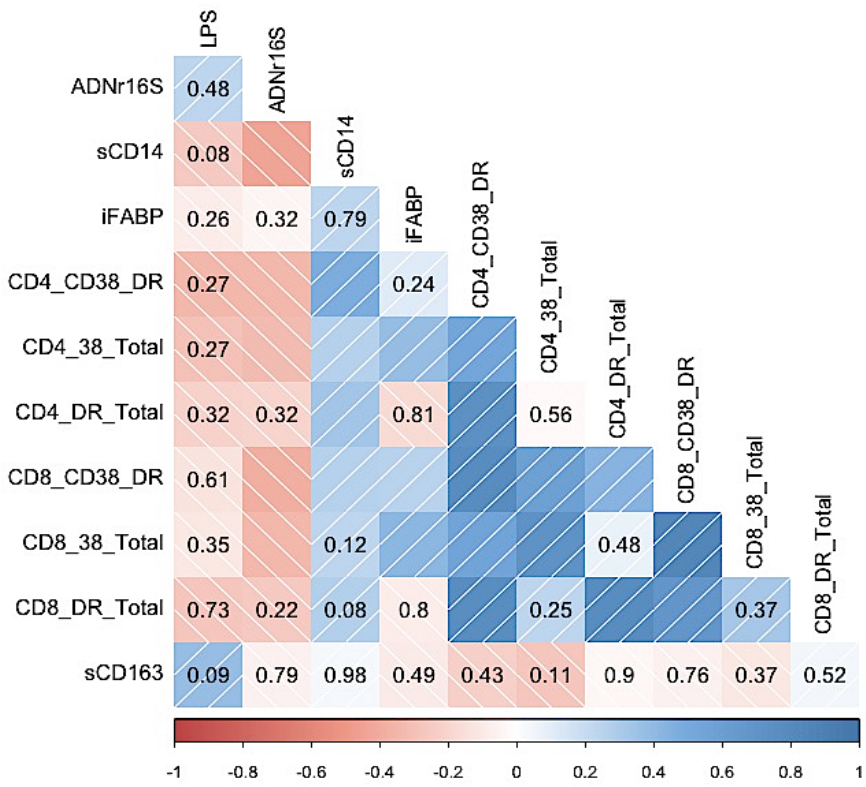
La Visualización exploratoria

Ejemplo 2: Identificación de patrones en los datos

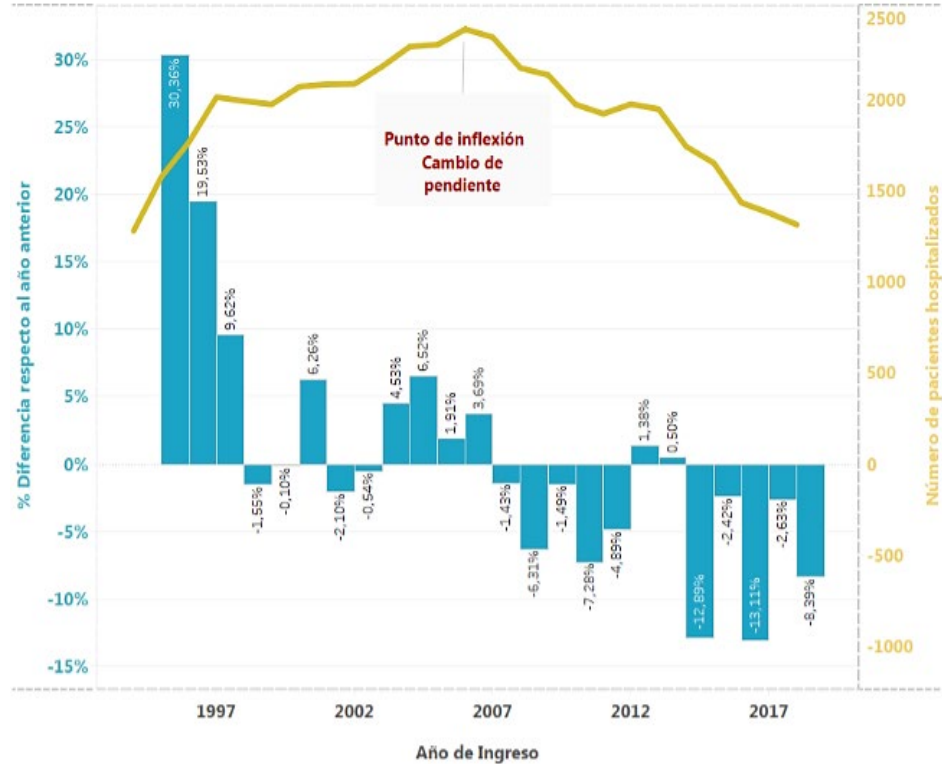


La Visualización exploratoria

Dependencia entre variables



Tendencias



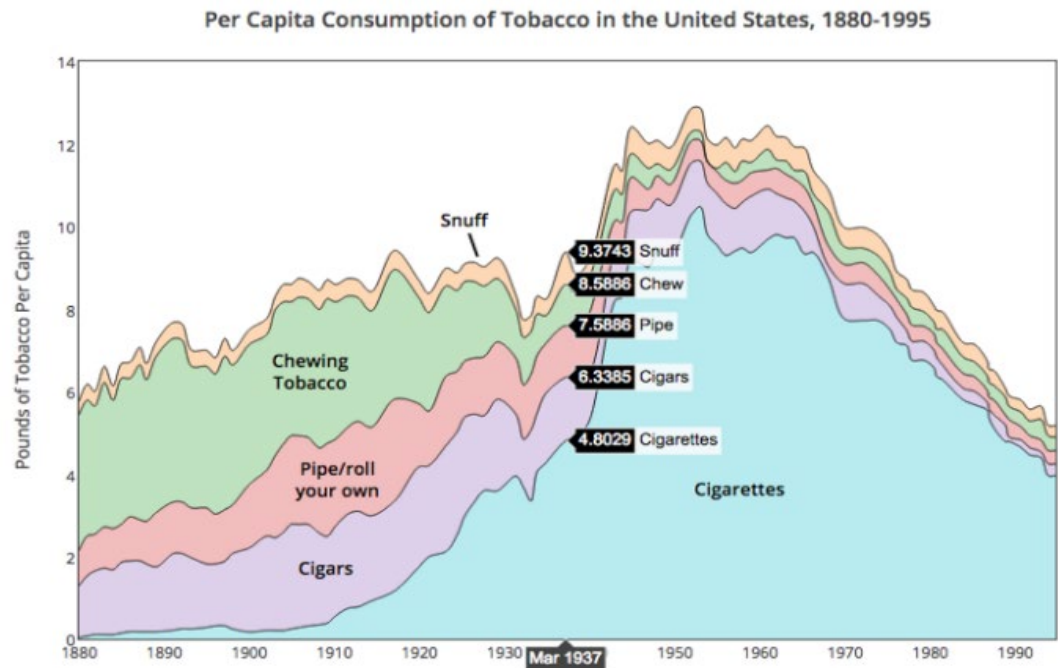
ESTÁTICAS vs DINÁMICAS



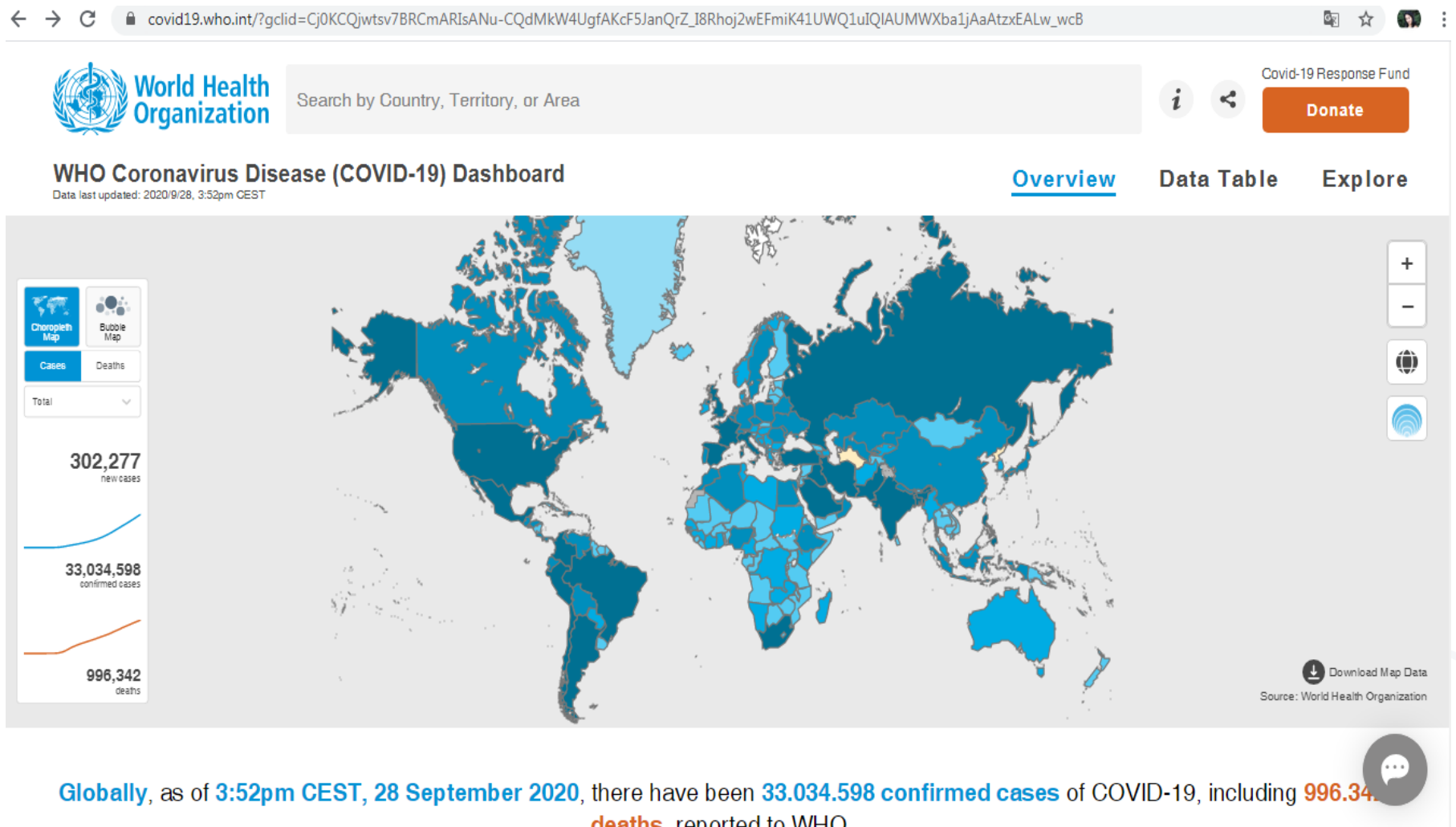
Visualizaciones estáticas

ESTÁTICAS

- Nos ayudan a detectar patrones
- Visualizar outliers
- Útil en formatos físicos (periódicos)
- Publicaciones científicas

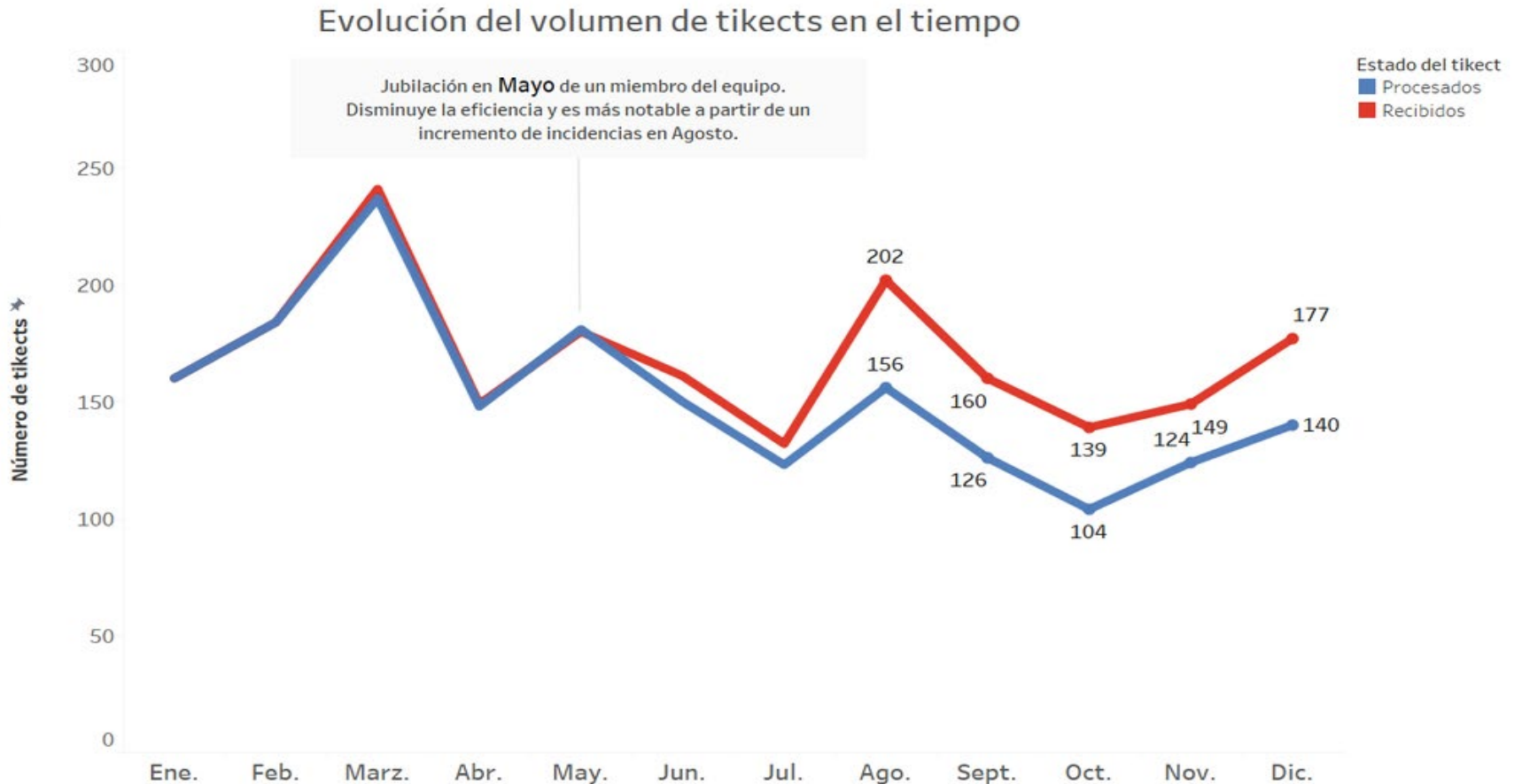


Visualizaciones dinámicas



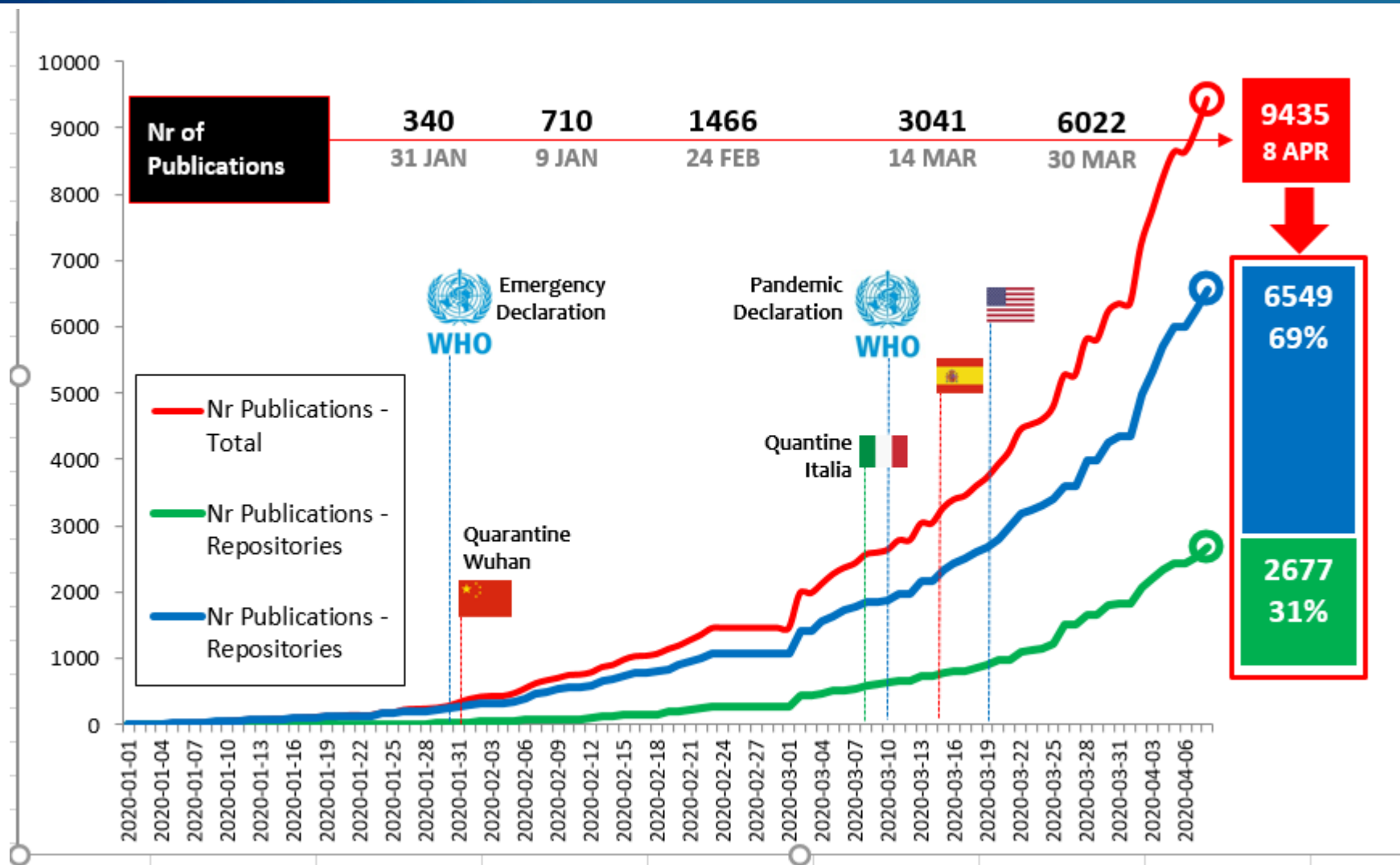
https://covid19.who.int/?gclid=Cj0KCQjwtsv7BRCmARIsANu-CQfqRn1hWBm8zYRHg5_-LTw5dLnB5yFWBwFVKVRPPVjETDQFCPpiBUYaAgImEALw_wcB

Anatomía de un gráfico



La tendencia de suma de Volumen de tikeects para Meses. El color muestra detalles acerca de Estado del tikeect.

Anatomía de un gráfico

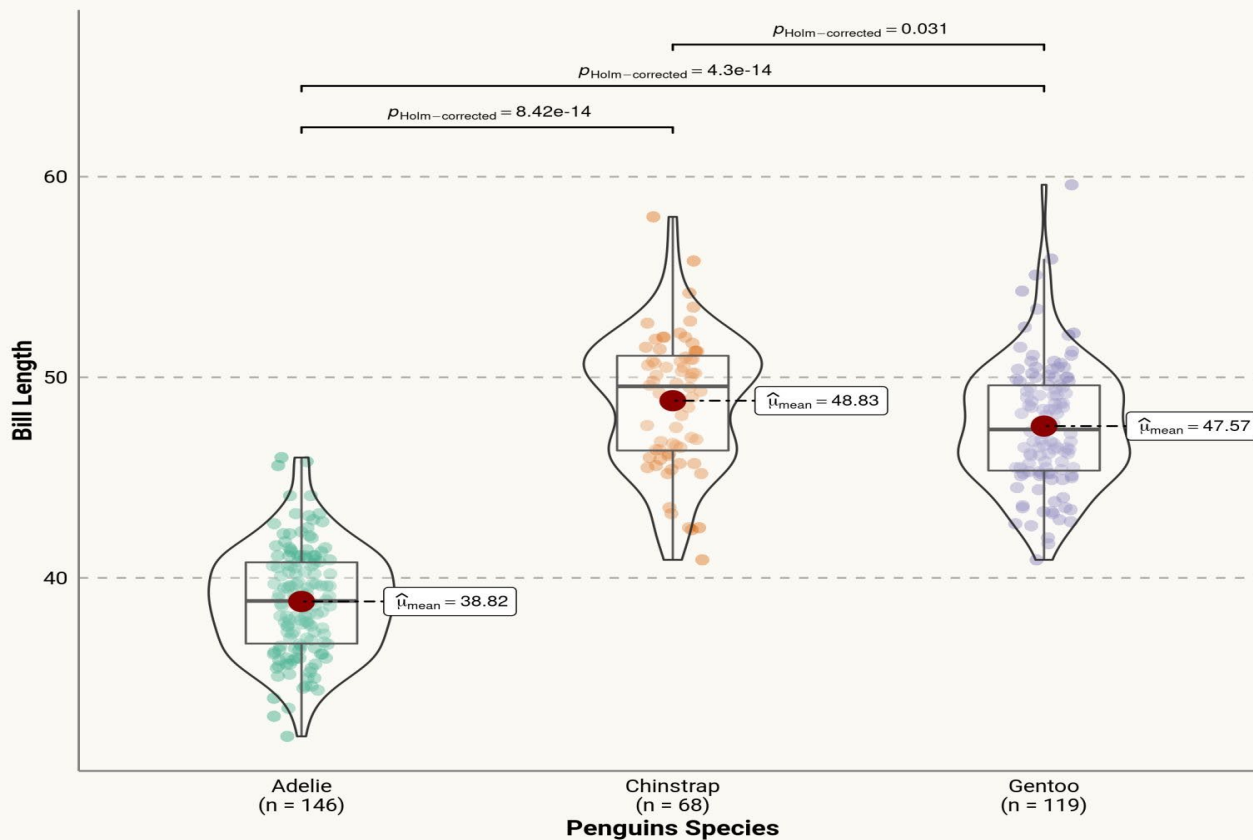


Anatomía de un gráfico

Gráfico de violines

Distribution of bill length across penguins species

$F_{\text{Welch}}(2,165.34) = 409.93, p = 8.27e-65, \hat{\omega}_p^2 = 0.83, \text{CI}_{95\%} [0.79, 0.86], n_{\text{obs}} = 333$

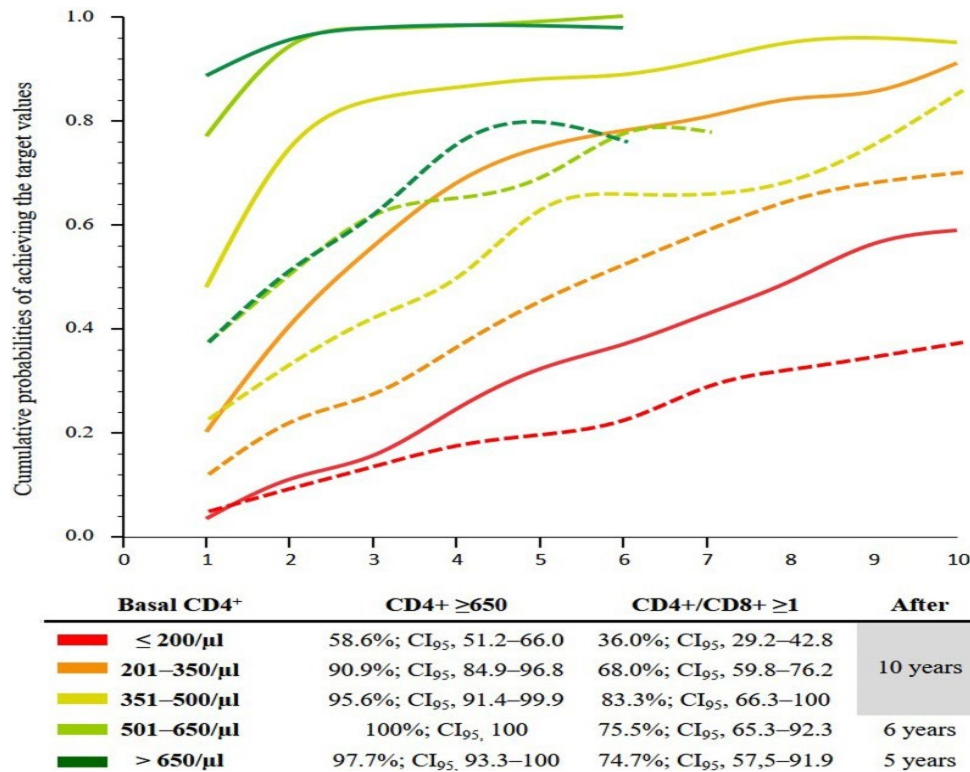


$\log_e(\text{BF}_{01}) = -195.59, R_{\text{Bayesian}}^{\text{posterior}} = 0.70, \text{CI}_{95\%}^{\text{HDI}} [0.67, 0.73], r_{\text{Cauchy}}^{\text{JZS}} = 0.71$

Pairwise test: **Games-Howell test**; Comparisons shown: **only significant**

Anatomía de un gráfico

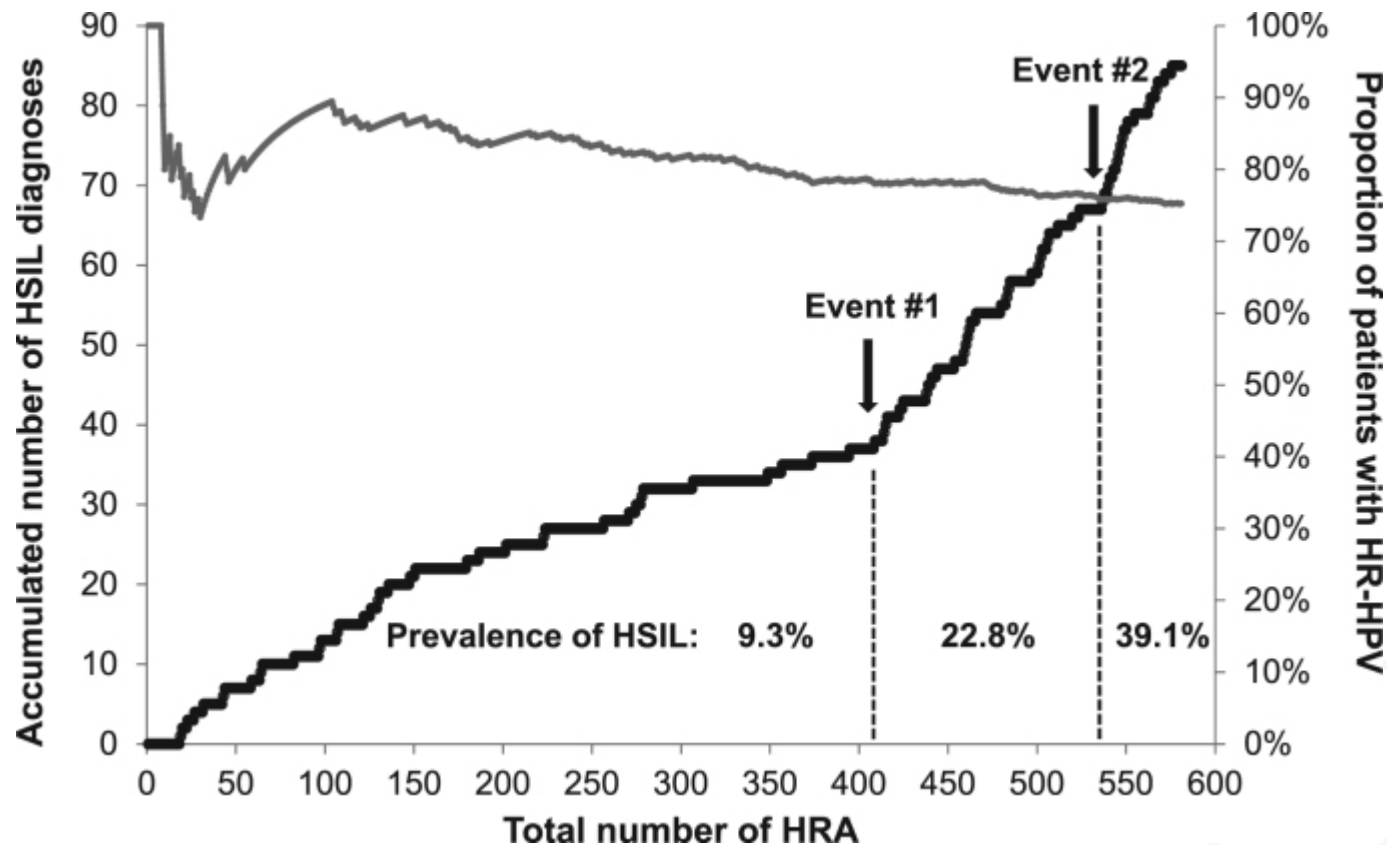
Cumulative probabilities of achieving a CD4 count $\geq 650/\mu\text{l}$ (continuous lines) and a CD4/CD8 ratio ≥ 1 (dashed lines) after starting antiretroviral therapy as function of CD4 counts at baseline.



* the values for patients who started treatment with absolute CD4 between 500 and 650/ μl and $>650/\mu\text{l}$ were displayed up to 6 and 5 years of follow-up, respectively, as the number in these groups decreased to less than 10 patients from these points on.

<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6197681/>

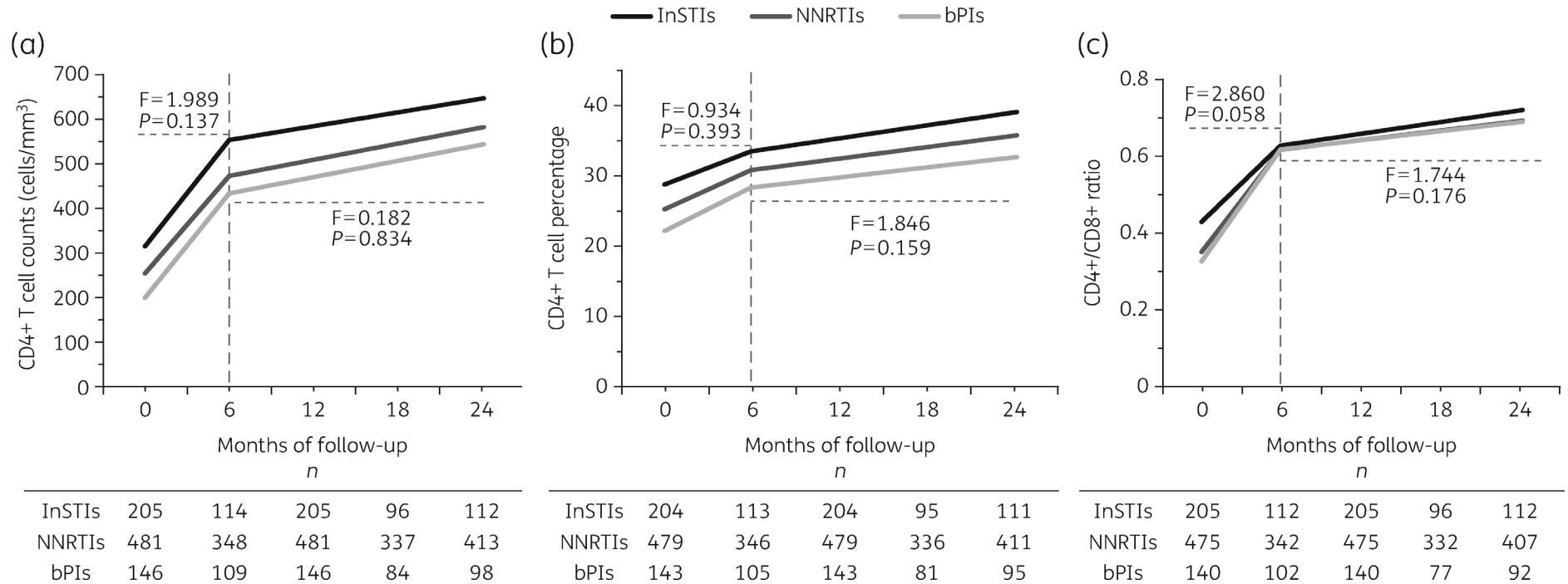
Anatomía de un gráfico



Accumulated high-grade anal squamous neoplasias (HSIL) according to the total number of high resolution anoscopies (HRA) with subsequent biopsy (black line) and the prevalence of high-risk human papillomavirus (HR-HPV) (grey line) in 581 HIV-infected men who have sex with men seen between 2010 and 2017 in an HIV outpatient clinic of a tertiary care centre in Seville, Spain. Event #1: One single expert pathologist responsible for biopsy interpretation; Event #2: Observer participates at additional one week expert training. The slope of the HR-HPV curve from the hundredth HRA onwards was -0.0003 ($r^2 = 0.97$).

<https://pubmed.ncbi.nlm.nih.gov/30716543/>

Figure 2. Slopes of increment by third antiretroviral drug class from baseline to month 6 and from month 6 to month 24. ...



#Tips gráficos publicaciones

Leyendas de las figuras:

- las figuras deben estar numeradas y tener títulos o leyendas descriptivas.
- **Las leyendas deben ser lo suficientemente concisas como para comprenderlas a primera vista.**
- Los subtítulos se colocan debajo de la figura y se justifican.

Imagen: Elija una imagen que sea simple y fácilmente comprensible. Considere el tamaño, la resolución y el atractivo visual general de la imagen.

Información adicional: Las ilustraciones en los manuscritos están numeradas por separado de las tablas. Incluya cualquier información que el lector necesite para comprender su figura, como las leyendas.

¿Gráficos simples o gráficos complejos?

Diseño acorde a propósito

Funcionalidad vs Estética.



Funcionalidad vs Estética

- Rigurosidad vs Belleza (o ambos)
- Audiencia
- Autoexplicativos
- Ideal = Algo bello + Algo que comunique
- Escribe el mensaje principal

Un subtítulo puede agregar claridad

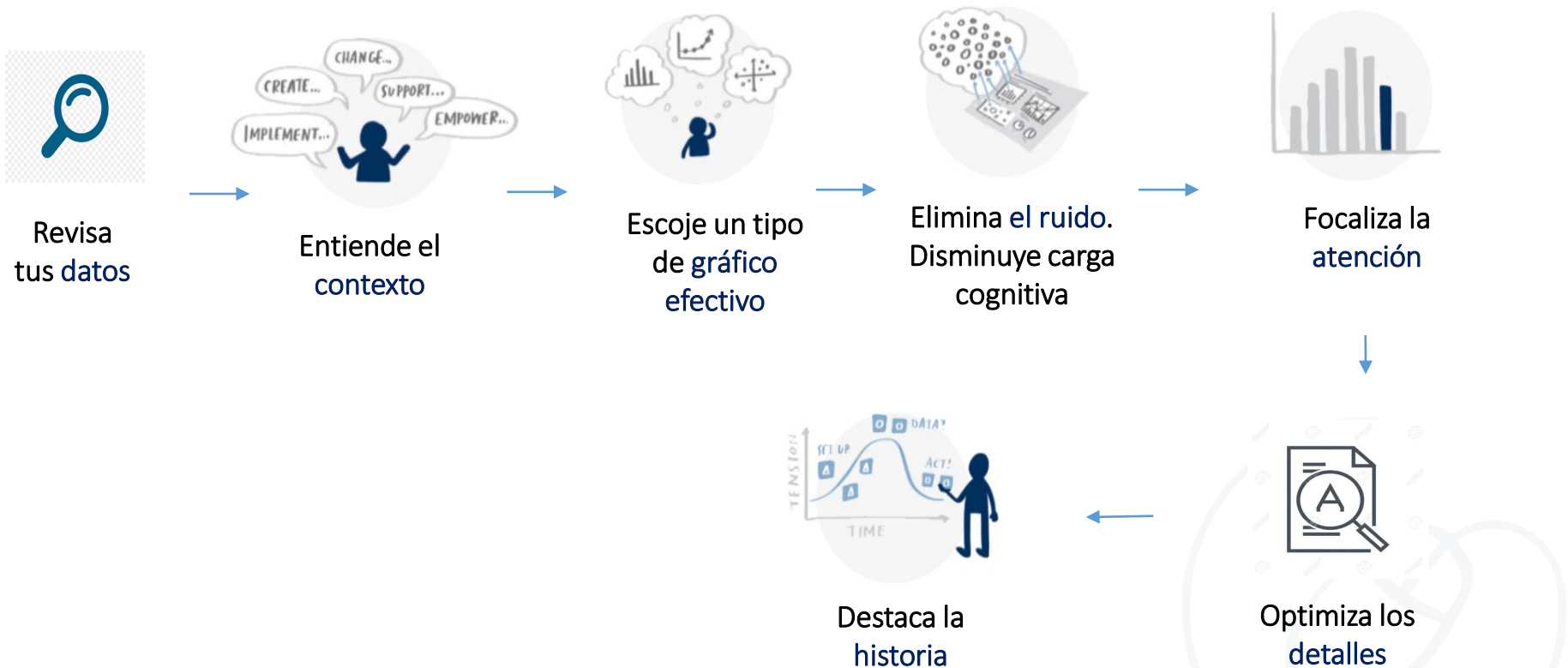


Etapas en la Visualización de datos



Pasos principales

Cuenta bien tu mensaje



<https://www.storytellingwithdata.com/>

Tipos de datos

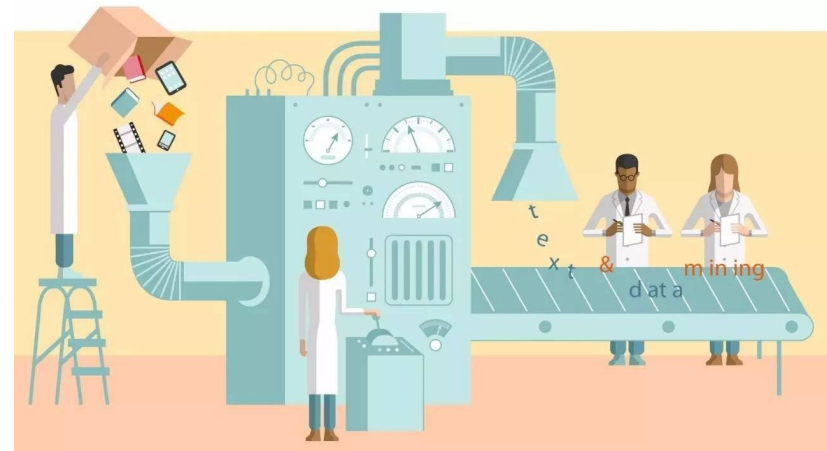
Data Types								
Numerical		Example	Categorical		Example	Time Series	Example	Text
Discrete->	Countable	Children count or count of members in family	Nominal->	Not ordered, Mutually exclusive	Sex(Male,Female)	Time series tracks the movement of data points along timestamp.	Network logs, Tracking daily, hourly, or weekly weather data	Plain text data. Example: Comments
Continuous->	Measurable	Age,Fare	Ordinal->	Ordered Values	Passenger class, Ratings			
Interval->	ordered units have same difference. Its something that's measured on a scale.	Temperature	Binary->	True/False or 0/1	Survived or not			
Ratio->	Same as interval except that it can have an absolute zero	length						

Revisa tus datos (Análisis exploratorio)

Prepara los datos, asegura su calidad.



- Accesibilidad
- Validez
- Precisión
- Relevancia
- Formato
- Consistencia
- Granularidad



Revisa tus datos (Análisis exploratorio)

Conoce tus datos, asegura su calidad.



- Es una serie temporal ? Una jerarquía?
- Cuántas dimensiones tengo , cuáles son las más importantes ?
- Tipos de datos: Categóricas, Numéricas, Ordinales, Ranking, Discretas.
- Siguen una distribución lineal o no – lineal ?
- Son correctos ? Outliers ? Valores ausentes ?

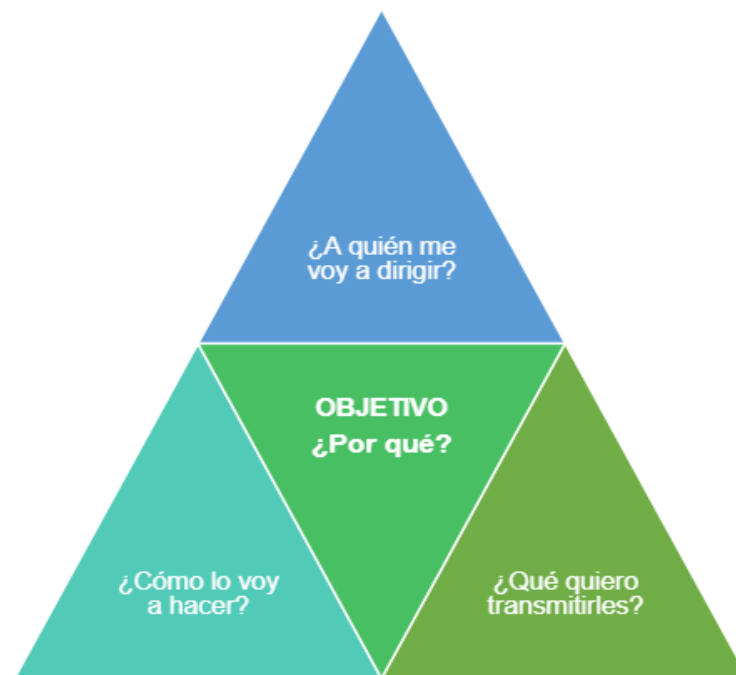


- Debe satisfacer tus objetivos
- Útil para la audiencia
- Debe estar estructurada adecuadamente
- Libre de ruido



Entiende el contexto y la audiencia

- **Audiencia:** Clínicos, Docentes, Público general; Investigadores de mismo campo o campo dis
Empresarios?
- Se enviará por email o se presencial ?
- Reporte o publicación científica ?
- Impresa o digital ?
- Estática o Dinámica?

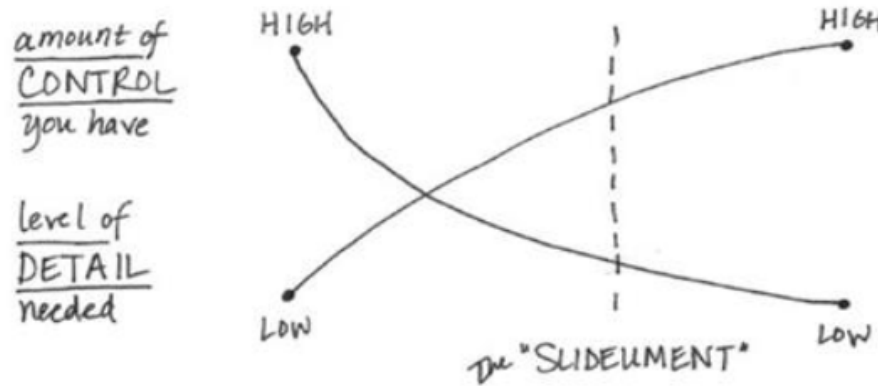


Entiende el contexto y la audiencia



La forma que toma el mensaje (medio) así como el modo en que es proporcionado a la audiencia, son elementos críticos a la hora de comunicar datos

LIVE PRESENTATION WRITTEN DOC OR EMAIL



Storytelling with data: a data visualization guides for business professionals

Escoge un tipo de gráfico efectivo

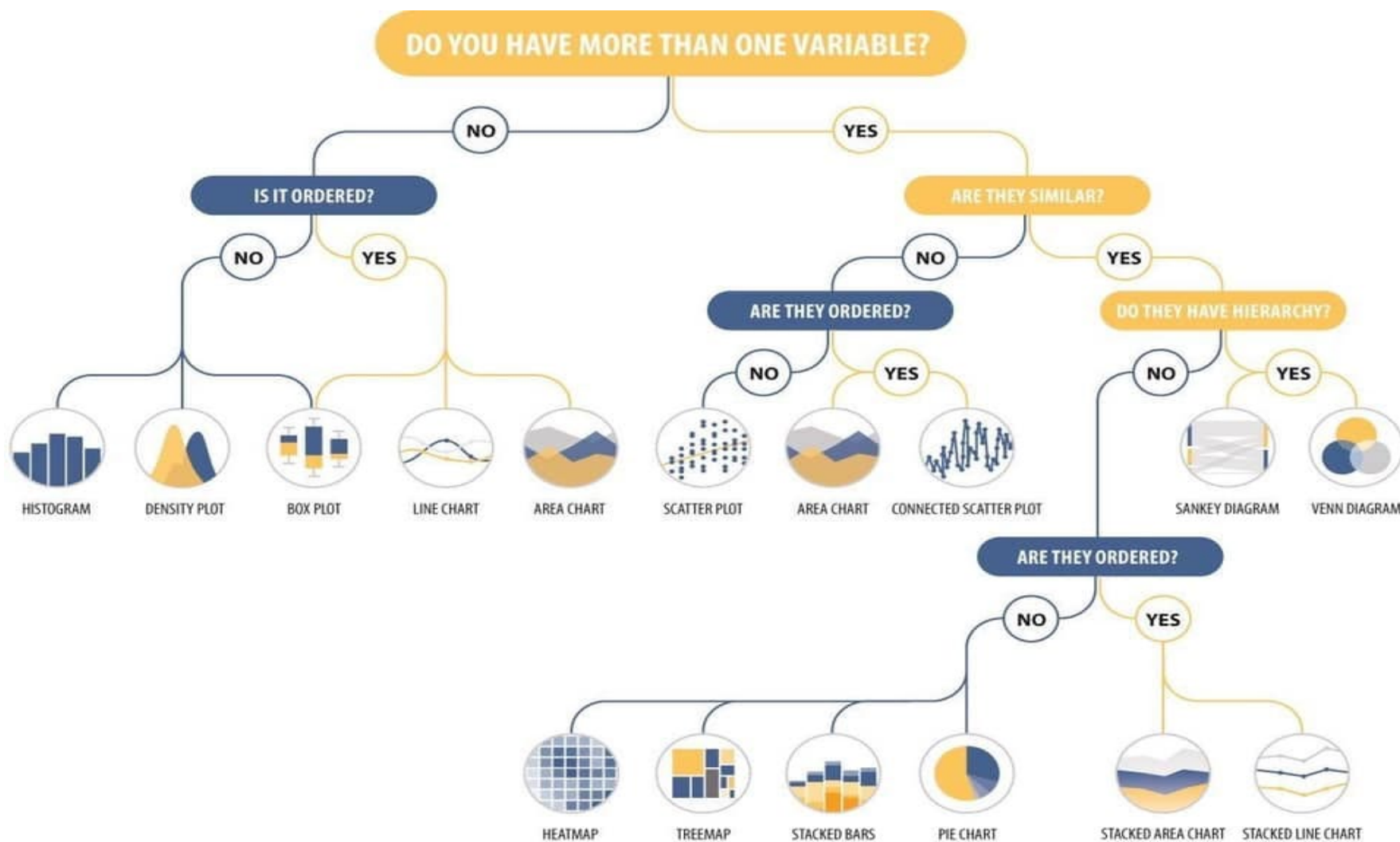


- ? Tienes una sola variable o más de una
- ? Son datos cualitativos o cuantitativos
- ? Los datos siguen un orden y/o jerarquía
- ? ¿Cómo se relacionan los datos ?

! Utiliza los diccionarios visuales online !

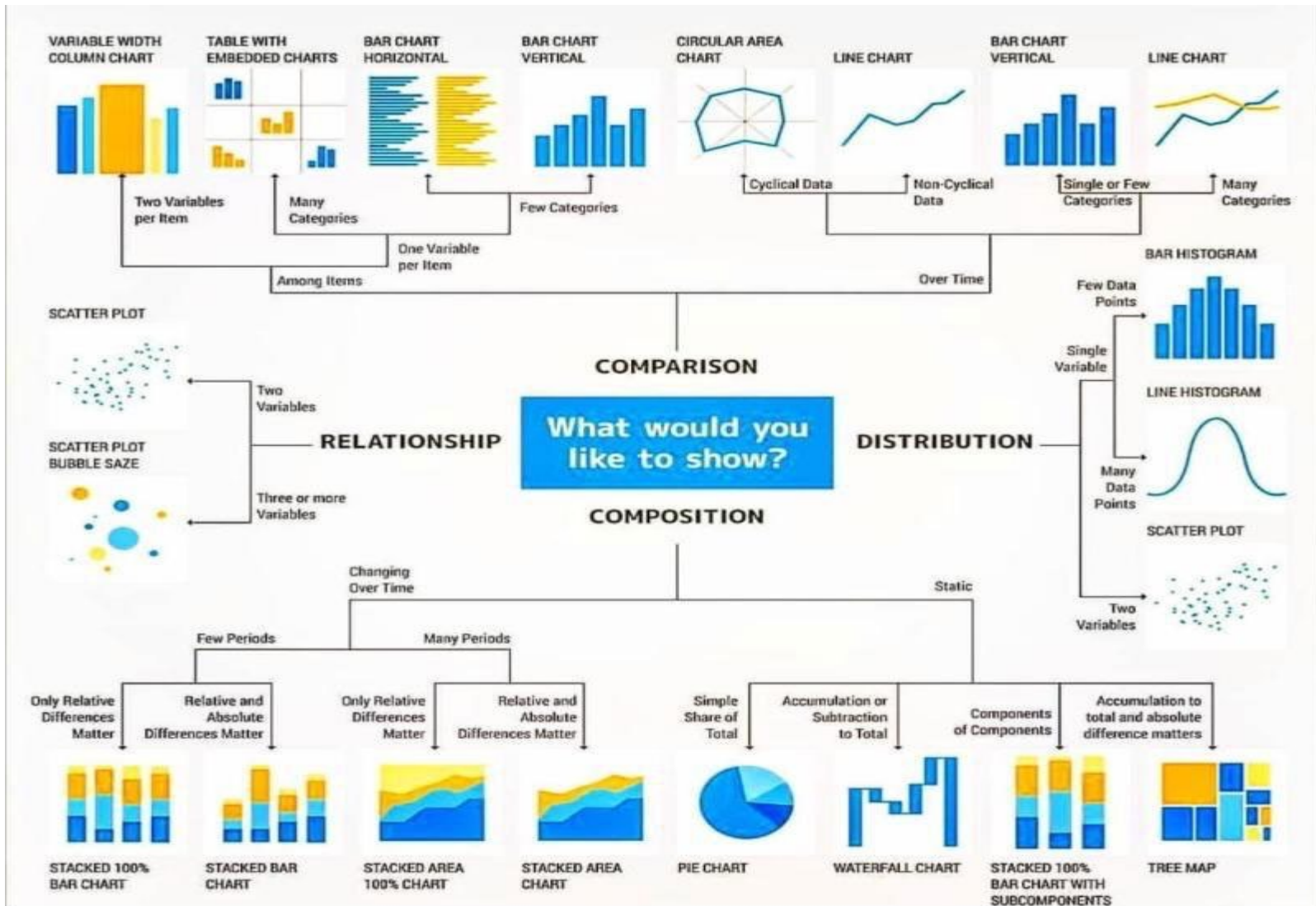
- [Visual Vocabulary of Financial times](#)
- [Graphica catalogue](#)
- [DataVizProject](#)

Escoge un tipo de gráfico efectivo



Created by ActiveWizards

Escoge un tipo de gráfico efectivo



Escoge un tipo de gráfico efectivo



<h3>Deviation</h3> <p>Emphasize variations (+/-) from a fixed reference point. Typically the reference point is zero but it can also be a target or a fixed benchmark. Can also be used to show asymmetry.</p> <p>Example FT uses Trade surplus/deficit, climate change</p> <p>Chart types: Diverging bar Diverging stacked bar Spin Surplus/deficit filled line</p>	<h3>Correlation</h3> <p>Show the relationship between two or more variables. Be mindful that, unless you list them effectively, charts readers will assume the relationship you show them will assume the relationship you show them.</p> <p>Example FT uses Inflation and unemployment, income and the expectancy</p> <p>Chart types: Scatterplot Column + line timeline Connected scatterplot Bubble XY heatmap Lollipop Rump</p>	<h3>Ranking</h3> <p>Use where an item's position in an ordered list is more important than its absolute or relative value. Don't be afraid to highlight the points of interest.</p> <p>Example FT uses Health, corruption, league tables, constituency election results</p> <p>Chart types: Ordered bar Ordered column Ordered proportional pyramid Dot strip plot Slope Lollipop Rump</p>	<h3>Distribution</h3> <p>Show values in a dataset and how often they occur. The shape (or 'skew') of a distribution can be useful for highlighting the lack of uniformity or equality in the data.</p> <p>Example FT uses Income distribution, population, logarithmic distribution, inequality</p> <p>Chart types: Histogram Dot plot Dot strip plot Barcode plot Boxplot Violin plot Population pyramid Cumulative curve Frequency polygons Baseline Circle timeline Vertical timeline Solenoigram Streamgraph</p>	<h3>Change over Time</h3> <p>Give emphasis to changing trends. These can be short (one day) movements or long-term changes. Learning from the past, forecasting the future, choosing the correct time period is important to provide context for the reader.</p> <p>Example FT uses Share price movements, economic time series, sectoral changes in a market</p> <p>Chart types: Line Column Bar Paired column Paired bar Slope Area chart Candlestick Fan chart (spirograms) Connected scatterplot Calendar heatmap Priority matrix Circle timeline Vertical timeline Solenoigram Streamgraph</p>	<h3>Magnitude</h3> <p>Show size comparisons. These can be relative (can be able to use a logarithmic scale) or absolute (read the size of the components, consider the use of a counter number (for example, barrels, dollars or people) rather than a calculated rate or per cent).</p> <p>Example FT uses Commodity production, market capitalisation, vitamins in a general</p> <p>Chart types: Column Bar Paired column Paired bar Marimekko Proportional symbol Isotype (spirograms) Lollipop Radar Puffball coordinates Bulker Grouped symbol</p>	<h3>Part-to-whole</h3> <p>Show how a single entity can be broken down into its component elements, if the reader's interest is directed toward the use of the components, consider the use of a magnifying type chart layout.</p> <p>Example FT uses Fiscal budgets, company structures, national election results</p> <p>Chart types: Stacked column Marimekko Pie Donut Treemap Voronoi Arc Gridplot Venn Waterfall</p>	<h3>Spatial</h3> <p>Addition factor maps only used when precise locations or geographical patterns in data are more important than the reader's anything else.</p> <p>Example FT uses Population density, natural resource locations, natural disaster risk, road network, catchment areas, variation in election results</p> <p>Chart types: Basic choropleth Proportional symbol (count/magnitude) Flow map Contour map Equalized cartogram Scaled cartogram (treemap) Dot density Heat map</p>	<h3>Flow</h3> <p>Show the reader volume or intensity of movement between two or more states or conditions. These might be regional sequences or geographical locations.</p> <p>Example FT uses Movement of funds, trade, migrants, accounts, information relationship graphs.</p> <p>Chart types: Sankey Waterfall Chord Network</p>
----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------	-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------	--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------	----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------	-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------	---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------	---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------	---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------	--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

Visual vocabulary

Designing with data

There are so many ways to visualise data - how do we know which one to pick? Use the categories across the top to decide which data relationship is most important in your story, then look at the different types of chart within the category to form some initial ideas about what might work best. This list is not meant to be exhaustive, nor a wizard, but is a useful starting point for making informative and meaningful data visualisations.

ft.com/vocabulary

•Visual Vocabulary of Financial times

Escoge un tipo de gráfico efectivo



91%

Simple text

	A	B	C
Category 1	15%	22%	42%
Category 2	40%	36%	20%
Category 3	35%	17%	34%
Category 4	30%	29%	26%
Category 5	55%	30%	58%
Category 6	11%	25%	49%

Table

	A	B	C
Category 1	15%	22%	42%
Category 2	40%	36%	20%
Category 3	35%	17%	34%
Category 4	30%	29%	26%
Category 5	55%	30%	58%
Category 6	11%	25%	49%

Heatmap



Scatterplot



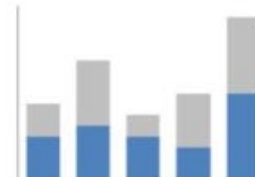
Vertical bar



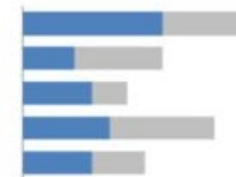
Horizontal bar



Line



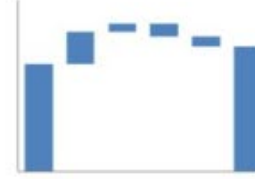
Stacked vertical bar



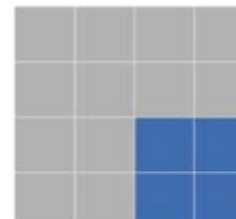
Stacked horizontal bar



Slopegraph



Waterfall



Square area

Storytelling with data: a data visualization guides por business professionals

Escoge un tipo de gráfico efectivo



Comparaciones

Para comparar diferentes variables o categorías entre si.



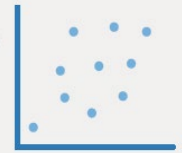
Tendencias

Para entender la evolución temporal de variables cuantitativas.



Correlaciones

Para entender la relación entre diferentes variables.



Mapas

Para encontrar patrones geográficos.



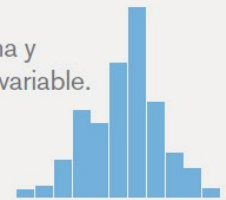
Partes de un total

Para entender la contribución de diferentes categorías en un total.



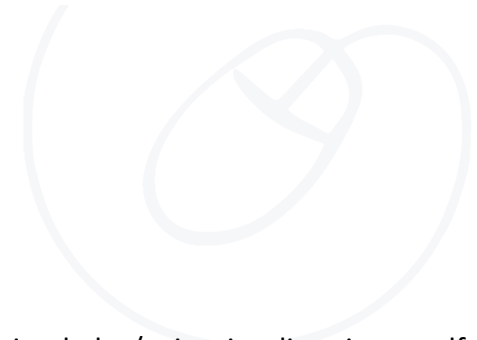
Distribuciones

Para entender la forma y propiedades de una variable.



Conexiones, relaciones y redes

Para entender la relación entre los elementos de un conjunto de datos.



http://atenciociudadana.gencat.cat/web/.content/manuals/visualitzacio_dades/guia_visualitzacio_es.pdf

Tipos de gráficos para comparar datos



Gráfico de barras

Para comparar un conjunto de valores y establecer un ranking.

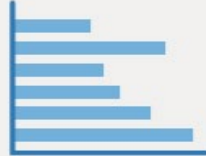


Gráfico de barras agrupadas

Para comparar el valor de los diferentes segmentos que forman parte de nuestras categorías.



Gráfico de barras apiladas

Para comparar el valor total de la suma de los segmentos que forman cada una de las barras.

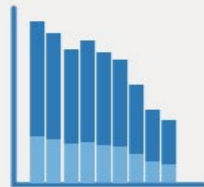


Gráfico de radar

Para entender los valores de un elemento en diferentes variables a la vez.



Gráfico de intensidades de colores

Para descubrir la variabilidad de un conjunto de variables.

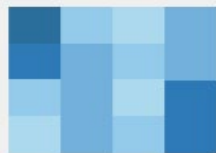


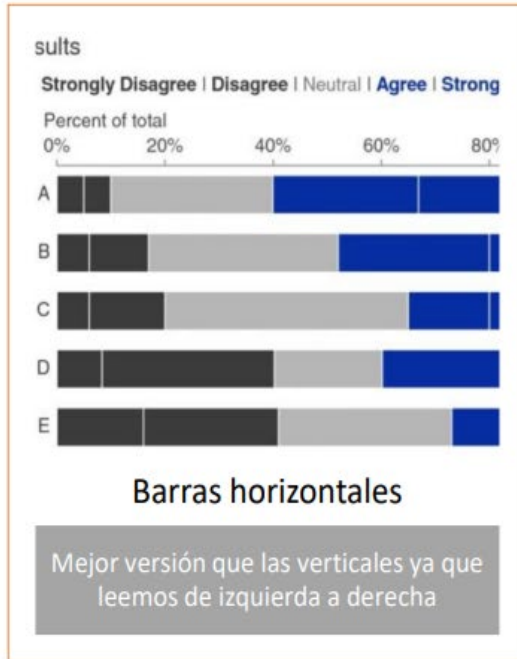
Gráfico de marcas

Para evaluar un valor concreto respecto al que se esperaba y respecto a unos rangos de calidad establecidos.



http://atenciociudadana.gencat.cat/web/.content/manuals/visualitzacio_dades/guia_visualitzacio_es.pdf

Tipos de gráficos para comparar datos

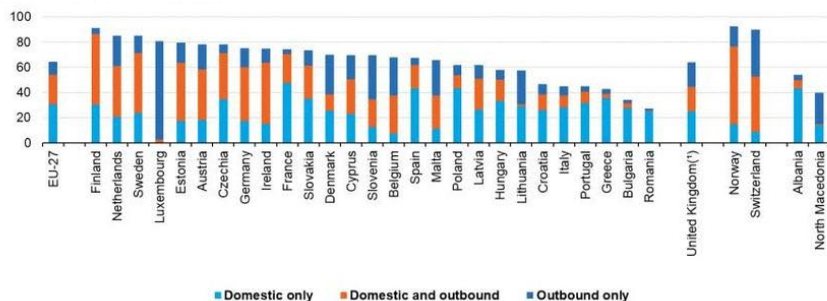


Storytelling with data: a data visualization guides for business professionals



BEFORE

Share of the EU population (aged 15 and over) participating in tourism by destination, 2018 (%)



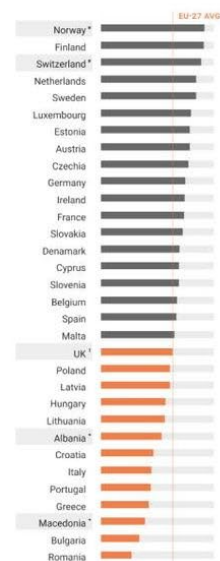
eurostat

(*) 2016 data.
Source: Eurostat (online data code: tour_dem_totot)

AFTER

% EU population 15+ participating in tourism in 2018

TOTAL TOURISM
% OF POPULATION THAT TRAVELLED



TOURISM BY DESTINATION
AMONG THOSE WHO TRAVELLED



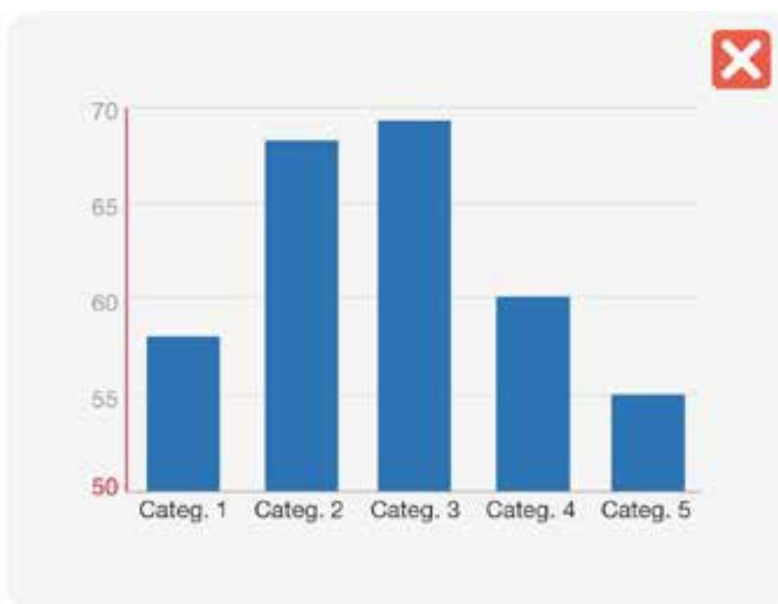
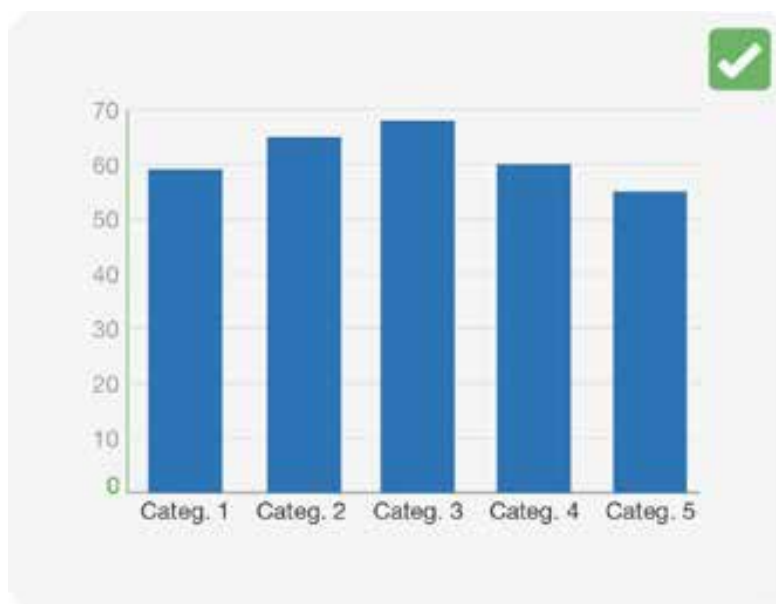
* countries outside EU / * candidate countries / * 2016 data
Source: Eurostat (online data code: tour_dem_totot)

Fuente: Eurostat.



Recomendaciones

¡Los gráficos de barras deben de ir siempre con el eje de ordenadas comenzando en el cero!

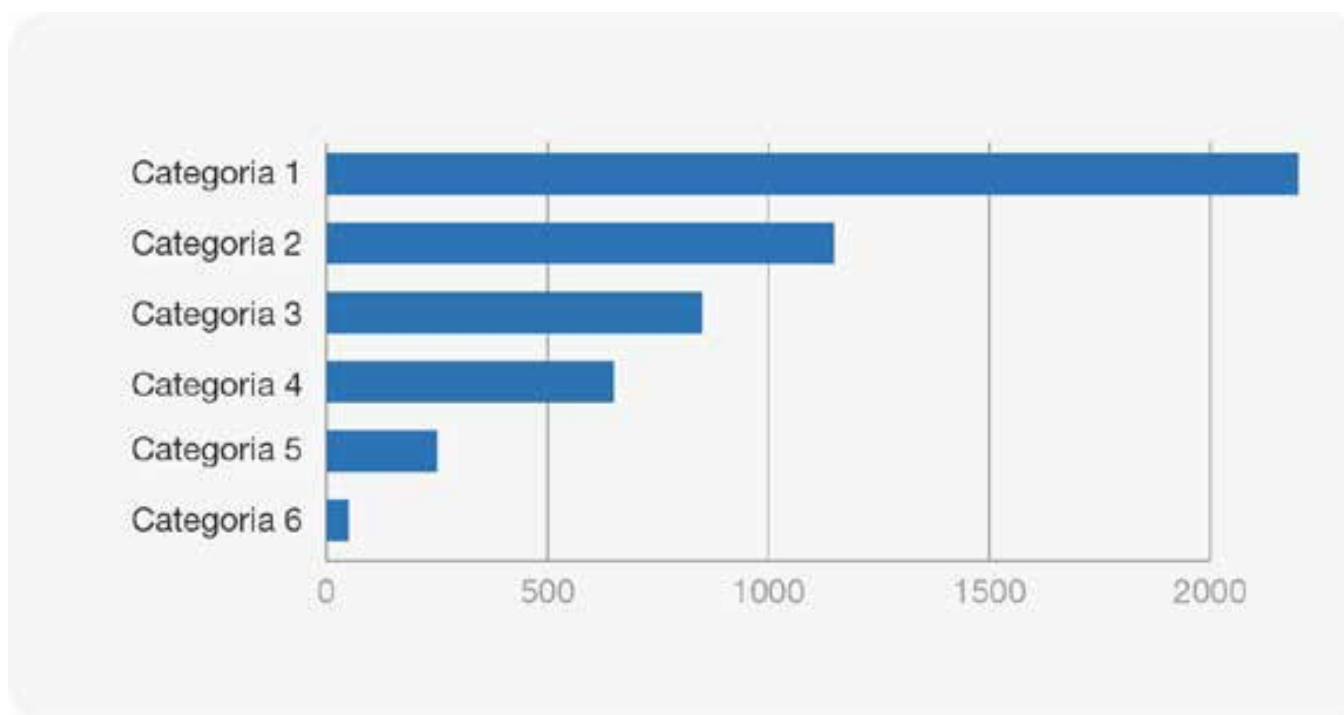


http://atenciociudadana.gencat.cat/web/.content/manuals/visualitzacio_dades/guia_visualitzacio_es.pdf



Recomendaciones

¡Si los nombres son largos, conviene orientarlo horizontalmente!

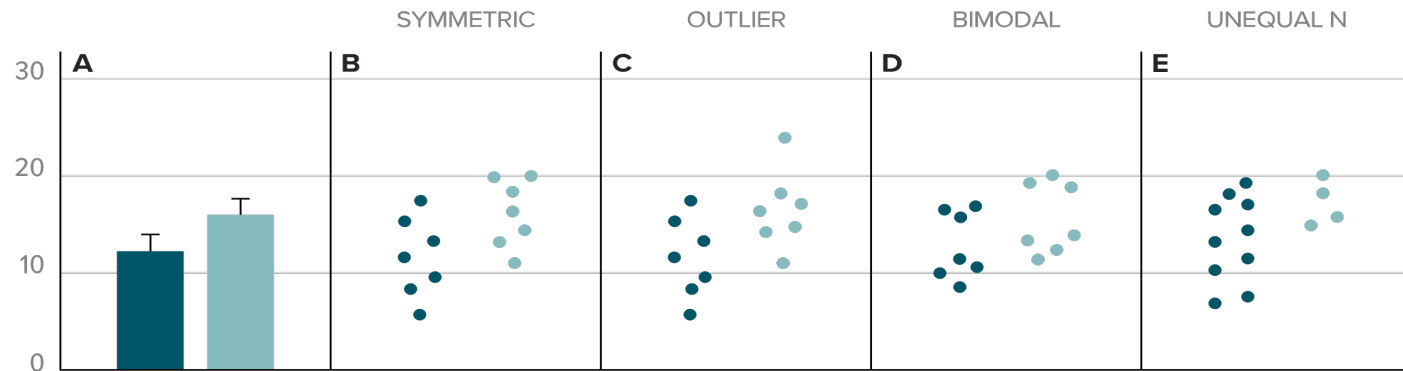


http://atenciociudadana.gencat.cat/web/.content/manuals/visualitzacio_dades/guia_visualitzacio_es.pdf

Buenas y malas prácticas

Hidden in the bars

Data revealed in scatterplots may be masked within a bar chart.



SOURCE: T.L. WEISSGERBER ET AL / PLOS BIOLOGY 2015

5W INFOGRAPHIC / KNOWABLE

Every one of the four sets of data on the right can be accurately represented by the same bar graph on the left, illustrating how bar graphs can obscure important details about the data, possibly misleading readers.

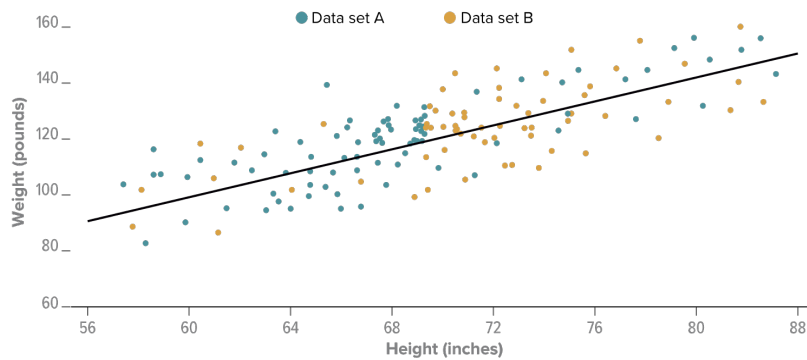
- **Bars don't show how small the sample sizes are, and outliers can have a big effect on the mean indicated by the height of a bar.**

<https://knowablemagazine.org/article/mind/2019/science-data-visualization>

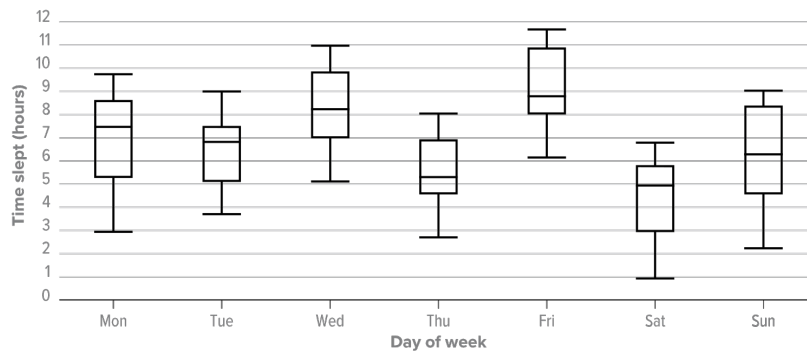
Buenas y malas prácticas

Los datos continuos es mejor presentarlos en scatterplots, diagrama de cajas y bigotes y/o histogramas.

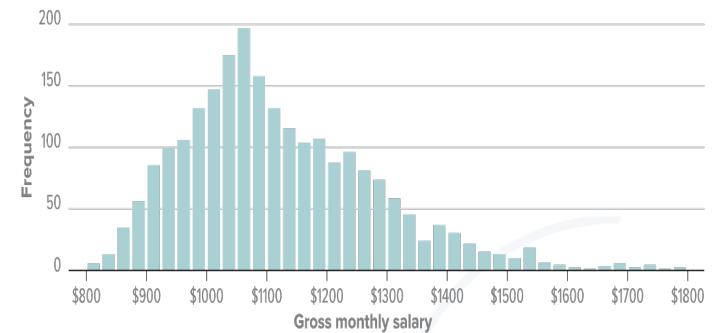
SCATTERPLOT



BOX PLOT

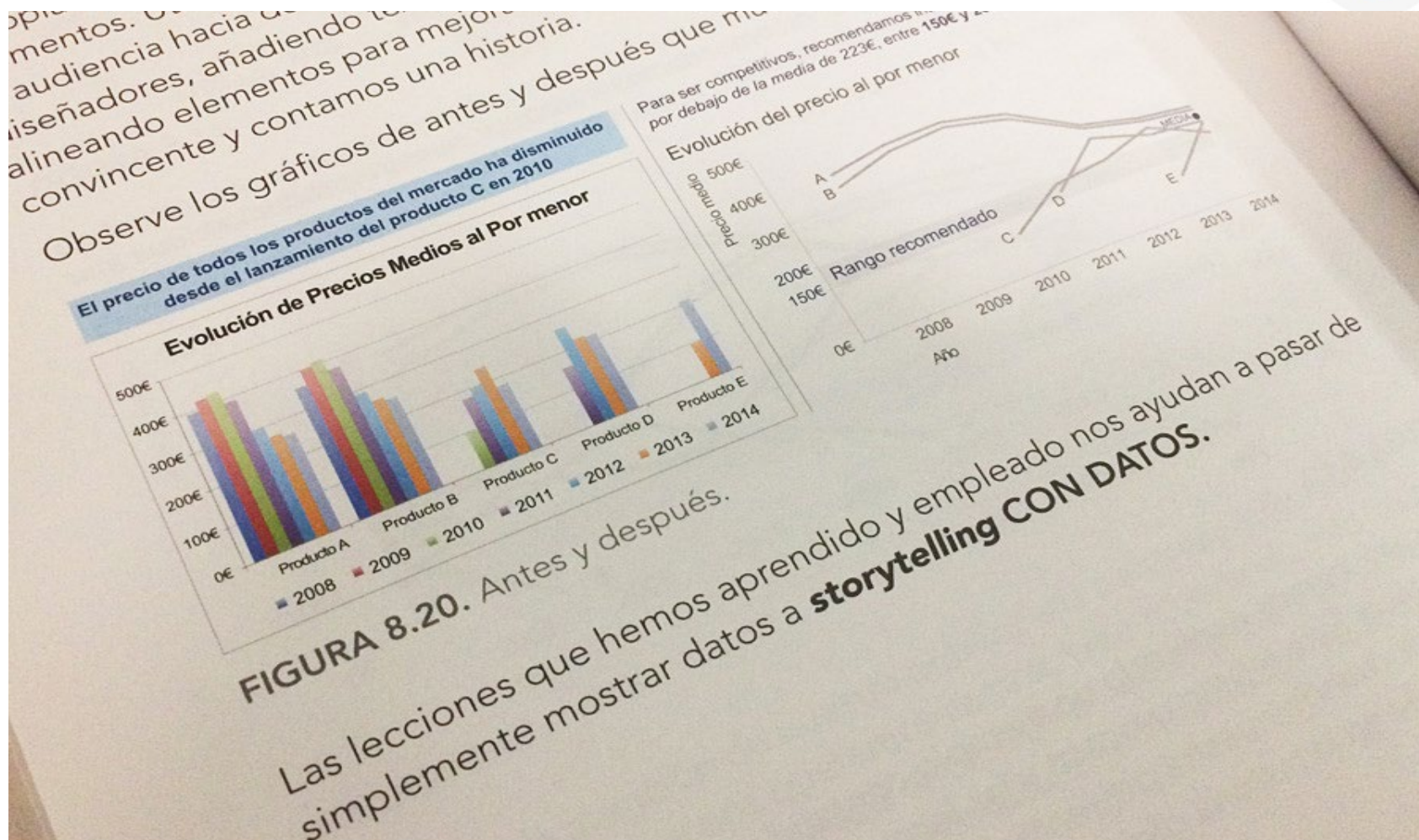


HISTOGRAM



SW INFOGRAPHIC / KNOWABLE

Buenas y malas prácticas

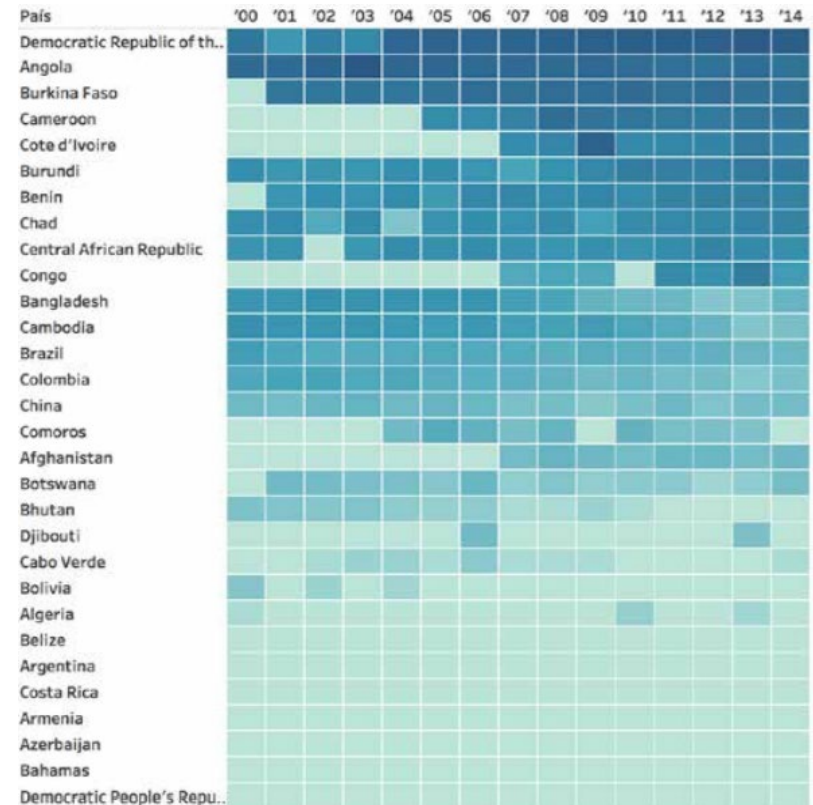
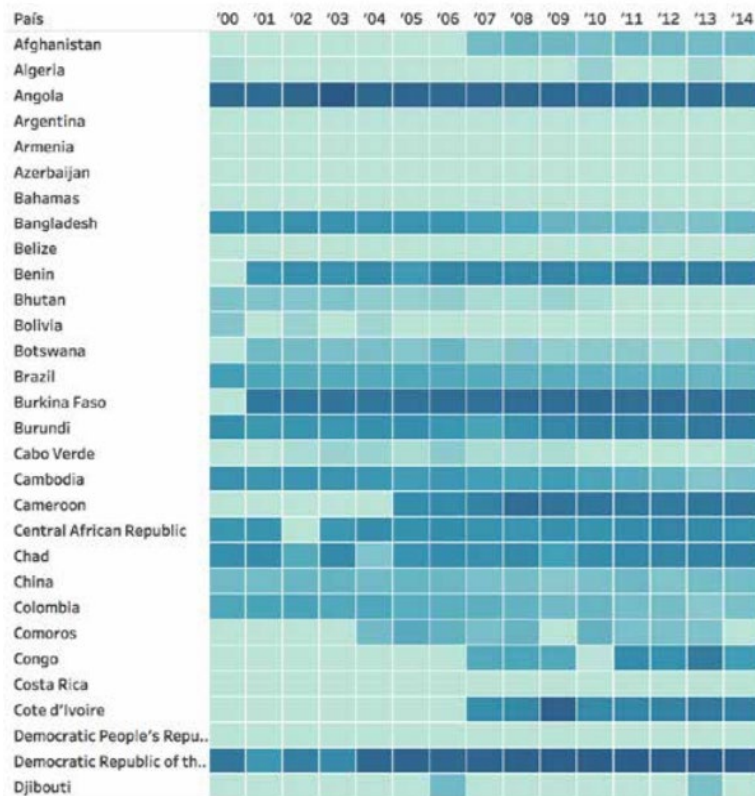


Storytelling with data: a data visualization guides por business professionals

Gráfico de intensidad de colores (Heatmap)



- Ordenar las filas y las columnas del gráfico de intensidades por colores según un criterio establecido
- Variables tienen escalas diferentes, **es aconsejable normalizarlas, de modo que todas pasen a tener el mismo rango de valores.**



Tipos de gráficos / Tendencias



Gráfico de líneas

Para ver la evolución temporal de una o más variables.



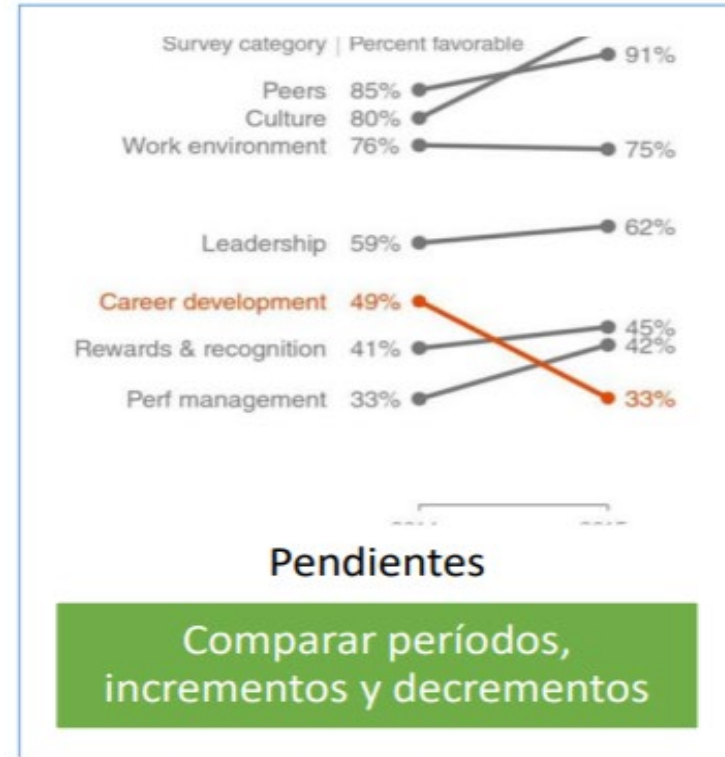
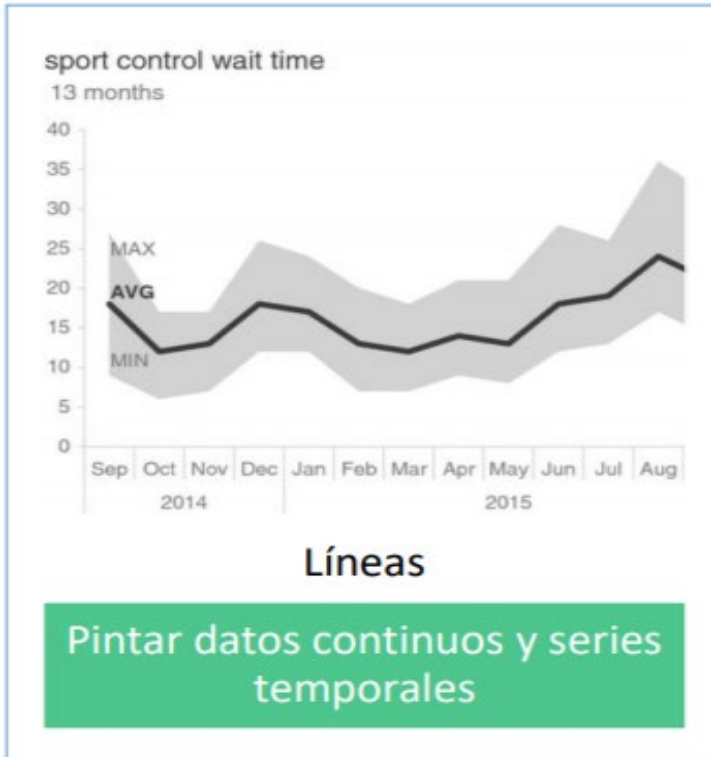
Gráfico de pendientes

Para comparar el antes y el después de una variable en varias categorías.



http://atenciociudadana.gencat.cat/web/.content/manuals/visualitzacio_dades/guia_visualitzacio_es.pdf

Gráficos de tendencia



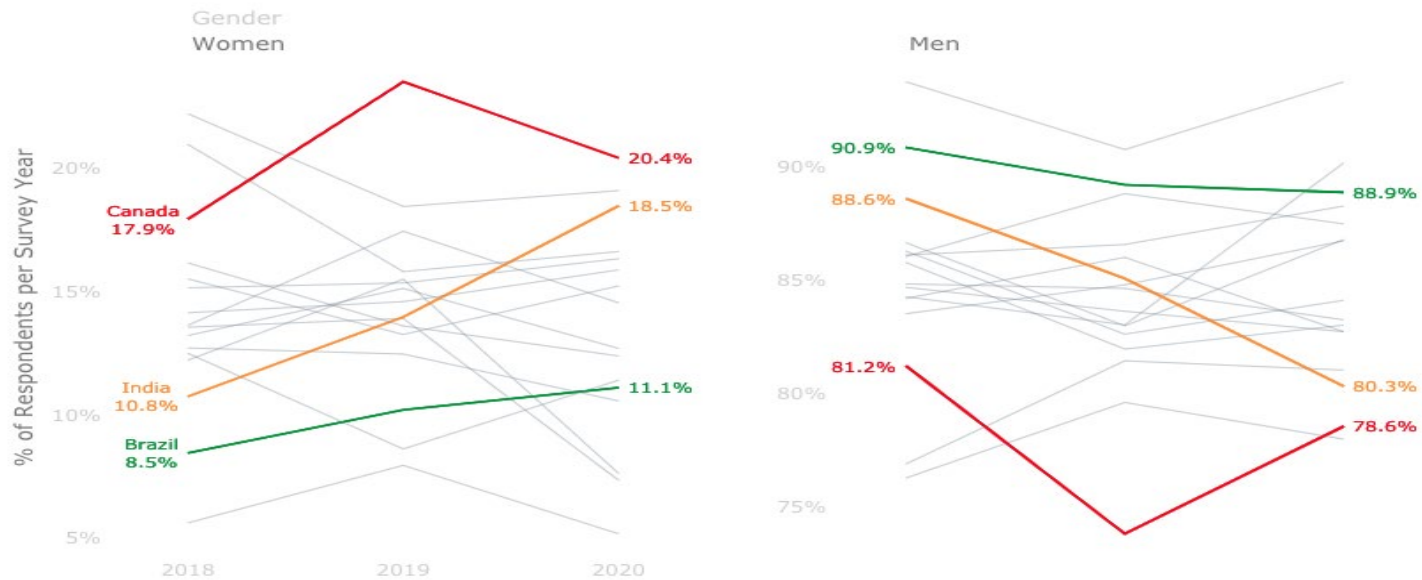
Storytelling with data: a data visualization guides por business professionals



Gráficos de tendencia / Tips

- Agrega la etiqueta usando el mismo color que las líneas correspondientes
- Si estas mostrando muchas líneas, intenta destacar sólo aquellas que enfatizan el mensaje que quieres transmitir

Gender Gap: India, Brazil and Canada are countries where the gender gap is reducing.
However changes are still very small to make any difference.
Percentage of professional respondents per country



Source: Kaggle surveys from 2018 to 2020.

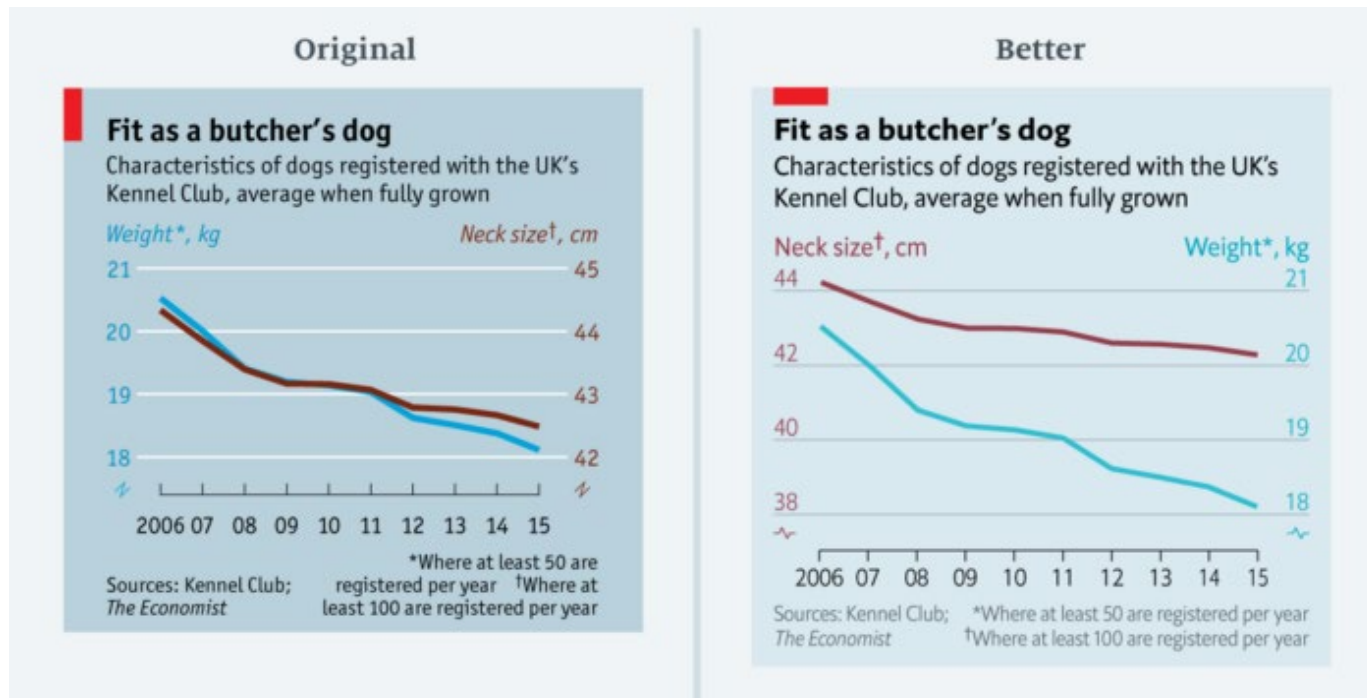
<https://medium.economist.com/mistakes-weve-drawn-a-few-8cdd8a42d368>

Buenas y malas prácticas

No forzar las escalas...

Forzar las escalas puede llevarnos a interpretaciones incorrectas.

Una correlación muy fuerte? No realmente.



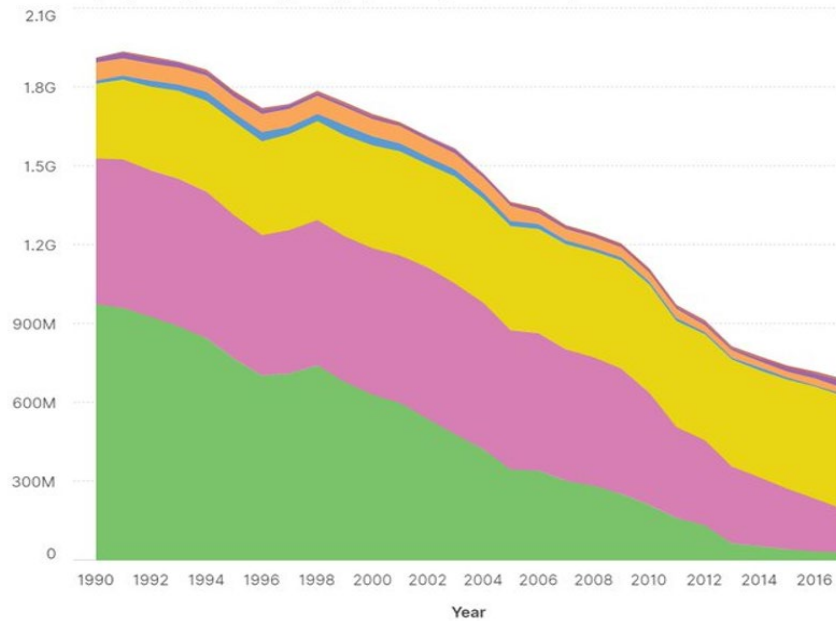
<https://medium.economist.com/mistakes-weve-drawn-a-few-8cdd8a42d368>



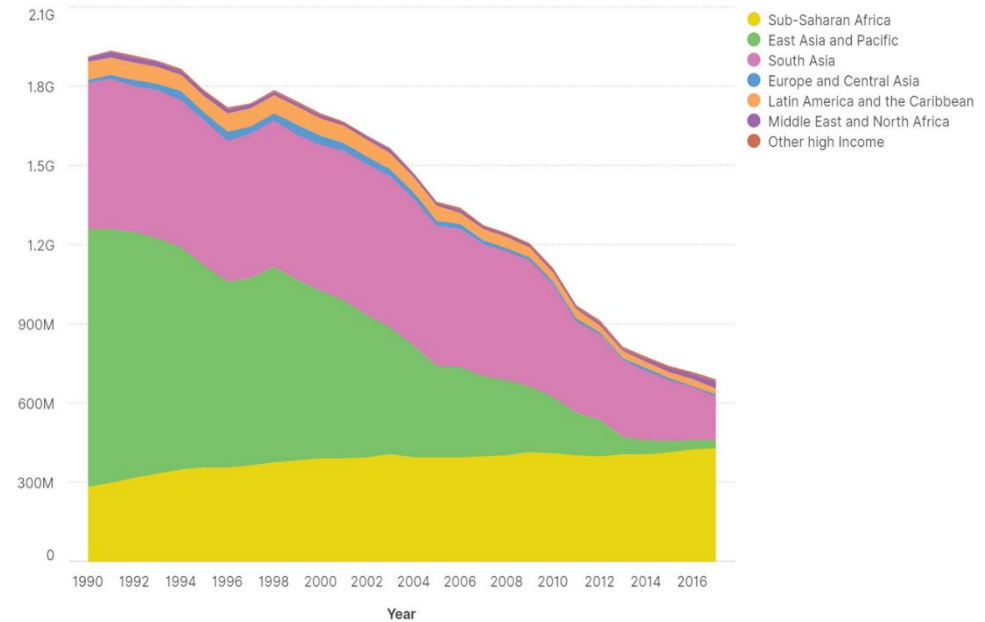
Buenas y malas prácticas

- Cuidar el orden en el que se apilan las áreas.

Number of people living in extreme poverty, by world region (World Bank, Povcalnet (2020))



Number of people living in extreme poverty, by world region (World Bank, Povcalnet (2020))



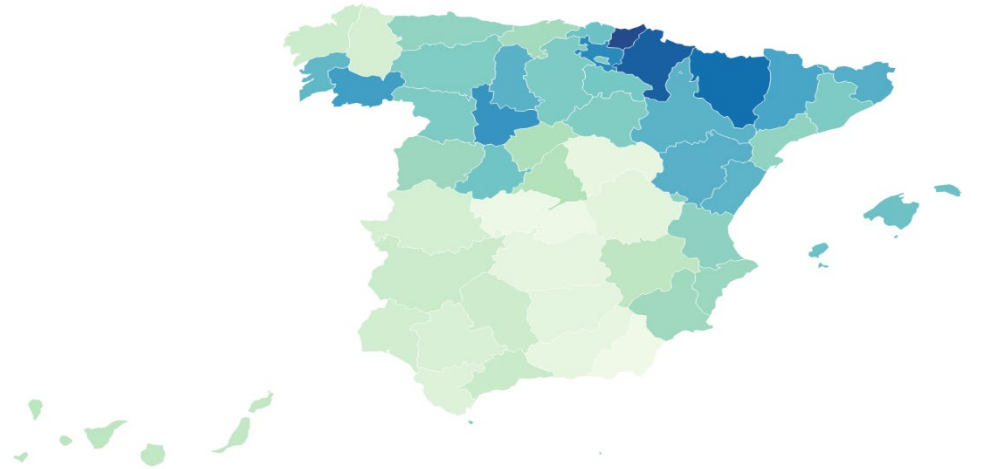


Mapa de coropletas (mapas rellenos)

- Los mapas de coropletas muestran los **valores de una variable sobre un mapa pintando las áreas de cada región** afectada de un color determinado.
- Los colores se utilizan para **representar una variable numérica o bien para representar la pertenencia de una región a una categoría concreta.**

Incidencia acumulada a 14 días en España

Mapa de incidencia acumulada por provincias en los últimos 14 días (1 al 14 de diciembre)



Mapa: A partir de La Tribuna de Ciudad Real • Fuente: Instituto de Salud Carlos III • Creado con Datawrapper

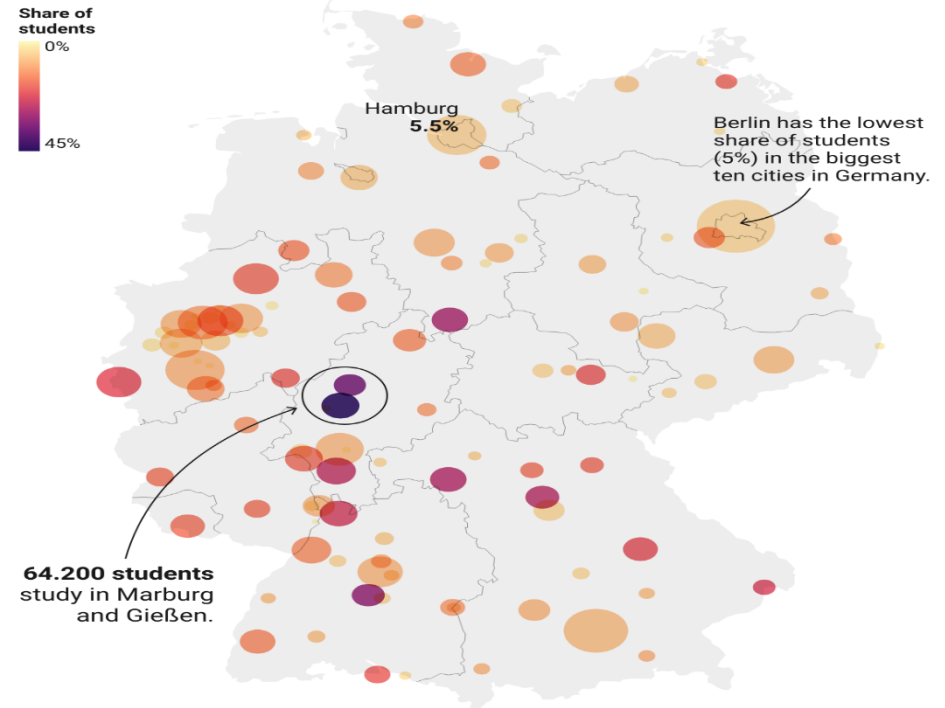


Mapa de coropletas (mapas rellenos)

- El mapa de símbolos proporcionales consiste en un mapa donde **se sitúa un icono o símbolo, generalmente un círculo (podría ser otro que se ajuste al contexto), de medida proporcional a la variable que se representa sobre el centro de la región a la que corresponde.**

Students in Germany: Where they live, how many they are.

There were 2.807.000 students in Germany in 2016. Where were they based, and which cities have the highest share of students? **The size of the circle represents the number of students in these cities. The darker the circle, the higher the share of students.**



Mapa: Lisa Charlotte Rost, Datawrapper • Fuente: Eurostat 2016 • Creado con Datawrapper

- Como vas a utilizar varios símbolos y además colores, se recomienda utilizar la transparencia para facilitar la lectura del gráfico e identificar los símbolos que quedan unos debajo de los otros.

Tipos de gráficos / Partes de un total



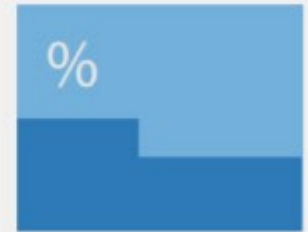
Gráfico de sectores

Para mostrar como se distribuyen proporcionalmente nuestros datos.



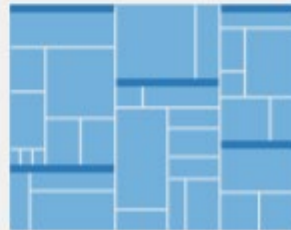
Pictograma

Para expresar un valor y a la vez ofrecer un contexto que nos permita identificar si es grande o pequeño.



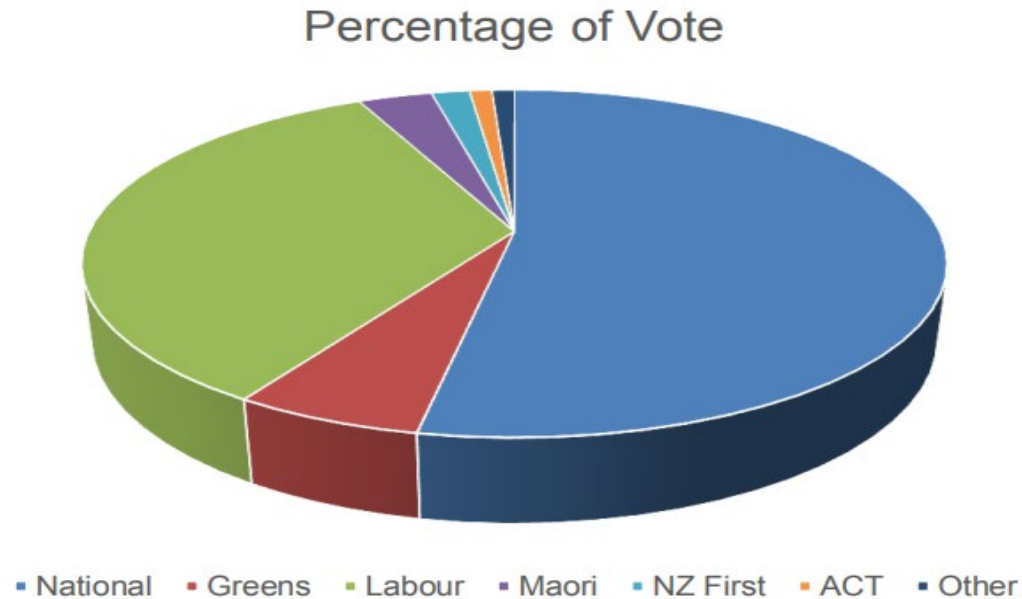
Mapa de árbol

Para representar datos jerárquicos y comparar una o dos variables entre los diferentes elementos.



Buenas y malas prácticas

- ✓ Evite gráficos de pastel y circulares (Especialmente en 3D)



EL ojo humano es malo leyendo ángulos.

Buenas y malas prácticas

Es un gráfico ideal cuando hay dos valores, como por ejemplo el número de personas que contestan sí o no a una pregunta.

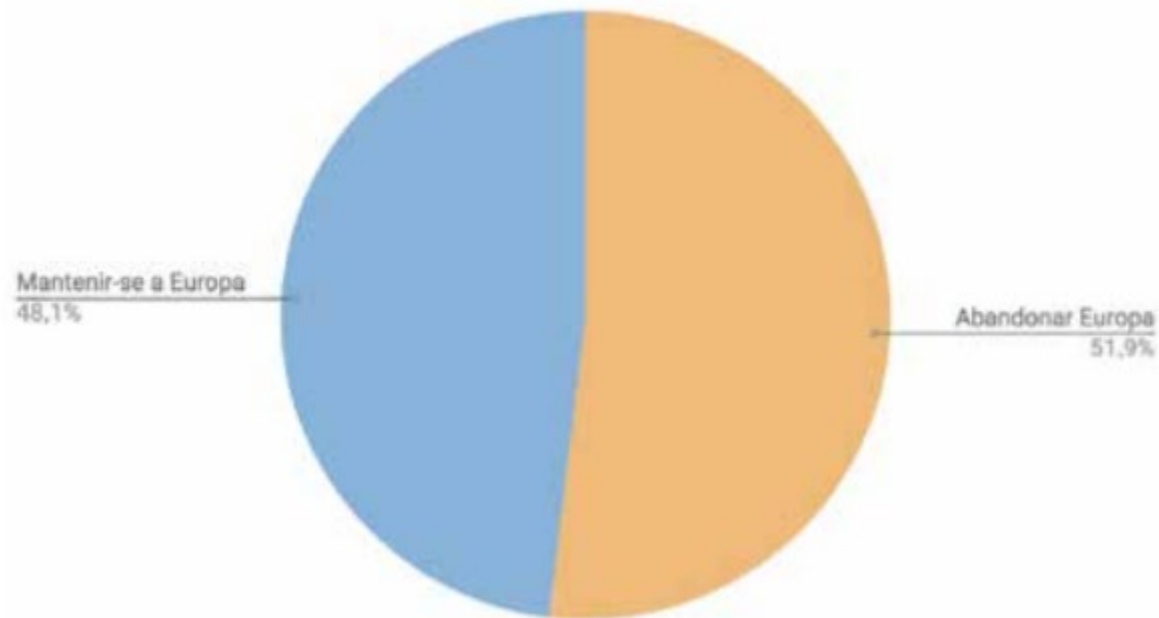


Figura 4.27. Resultado de la votación del Brexit en el Reino Unido.

Buenas y malas prácticas

Cuando son datos categóricos no ordinales

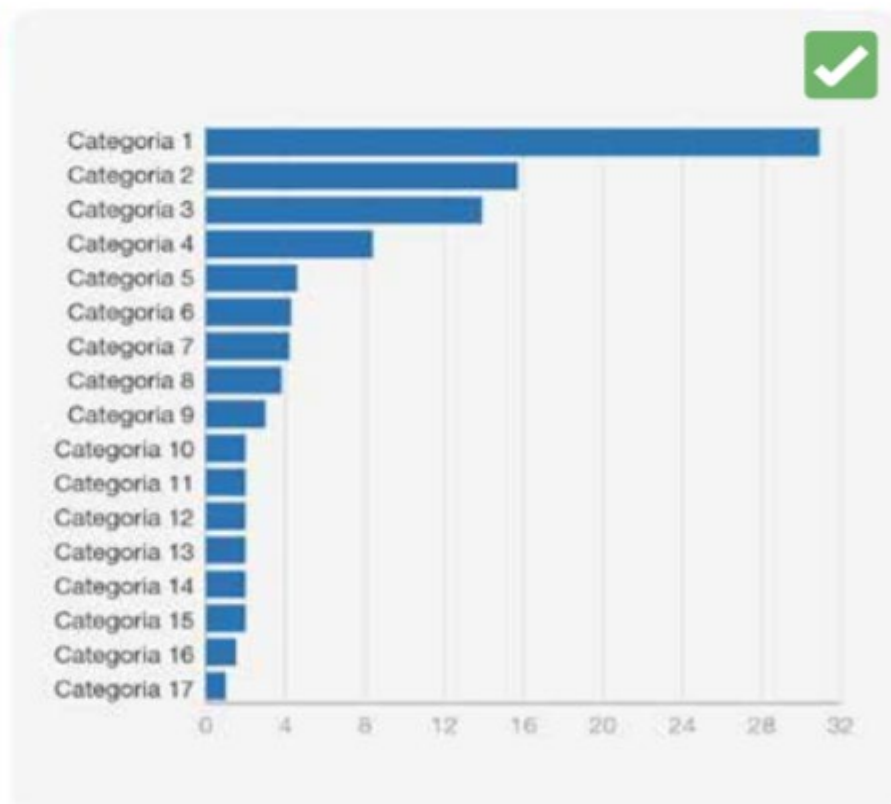
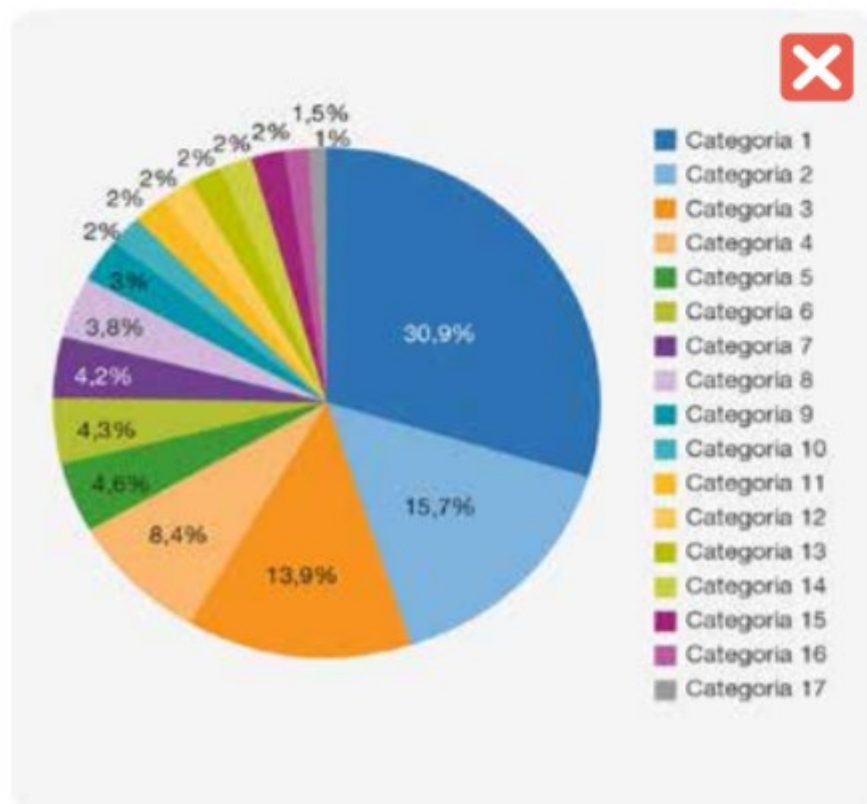


Figura 4.31. Datos ordinales representados con un gráfico de barras apiladas. Adaptación de la imagen cedida por Ann K. Emery.
Fuente: <http://annkemery.com/pie-chart-guidelines/>.



Buenas y malas prácticas

Evítalos si tienes muchas categorías



Figuras 4.29. Los gráficos de sectores con muchas divisiones no son buenas representaciones ya que confunden y no ayudan a entender la diferencia entre los valores. Fuente:

<http://www.conceptdraw.com/solution-park/resource/images/solutions/pie-charts/Graphs-and-Charts-Pie-Chart-Business-Report.png>.

Buenas y malas prácticas

Evítalos si tienes muchas categorías

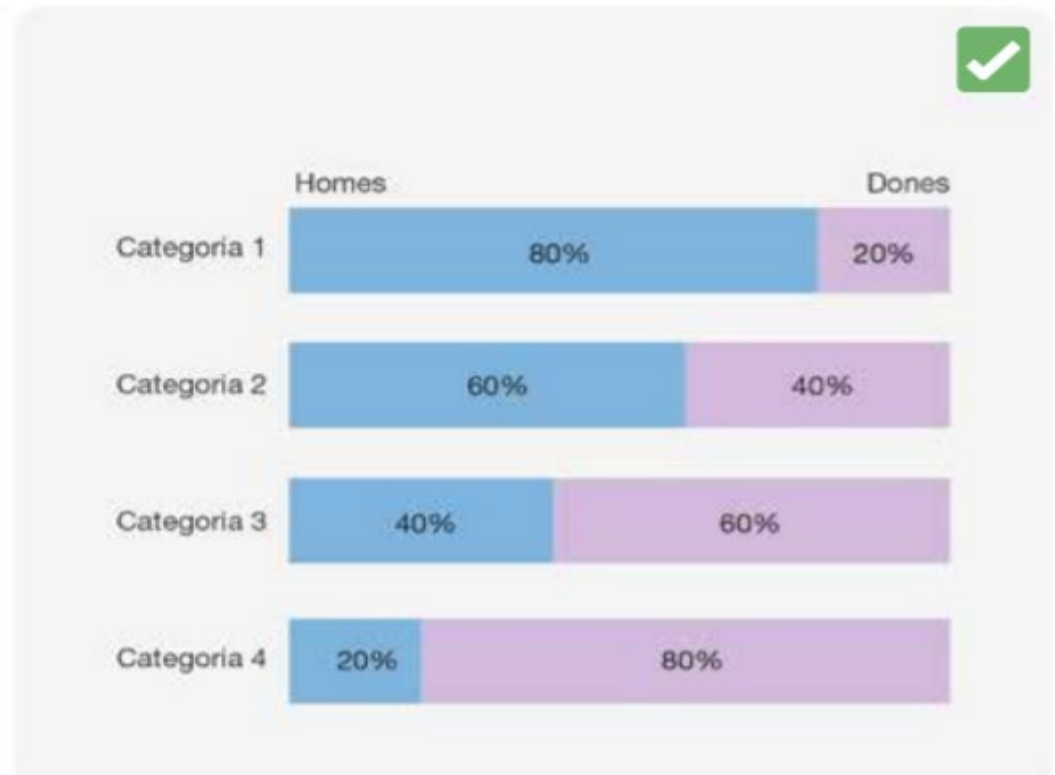
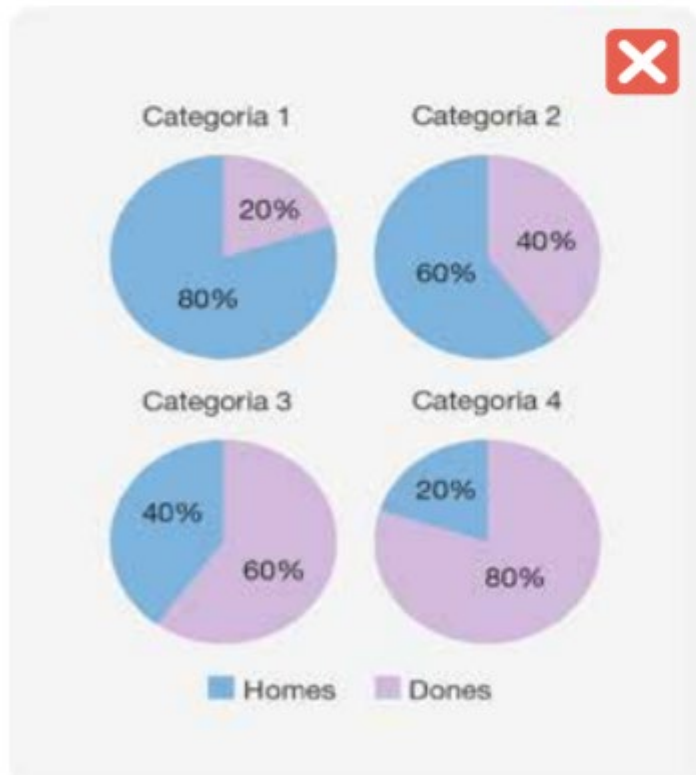


Figura 4.33. Es preferible utilizar barras para comparar valores a través de varias categorías. Adaptación de la imagen cedida por Ann K. Emery. Fuente: <http://annkemery.com/pie-chart-guidelines/>.



Tipos de gráficos / Partes de un total

PICTOGRAMAS

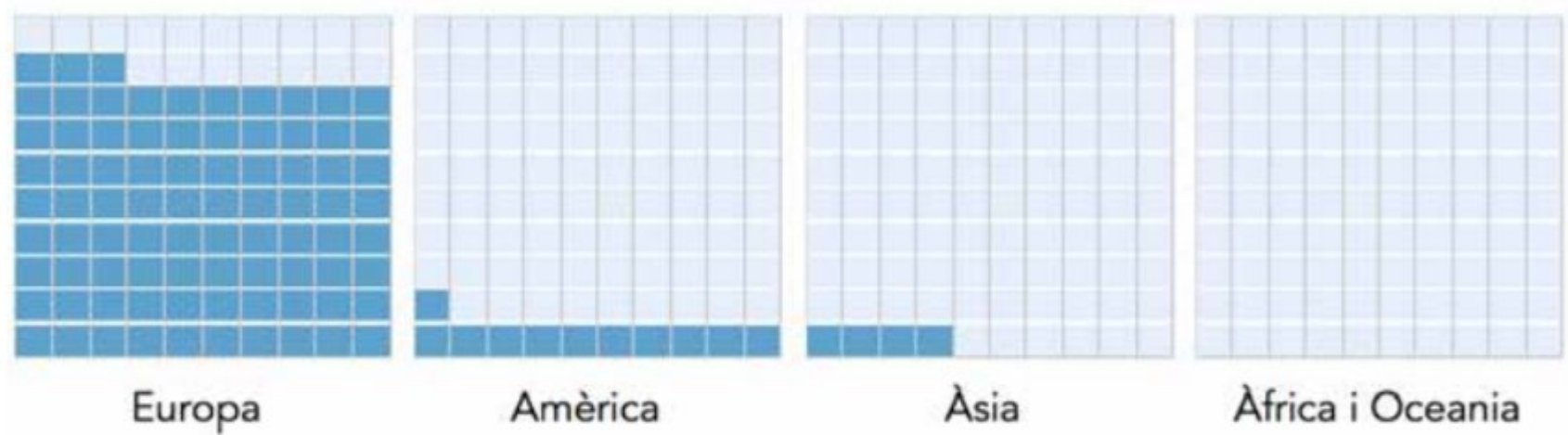


Figura 4.36. Percentatge de població que beu cava als diferents continents. Visualització feta per OneTandem.
Font: <https://public.tableau.com/profile/onetandem#!/vizhome/cava/cavaexports>.

Los pictogramas son una buena alternativa gráfica para mostrar un único número, especialmente cuando se trata de un indicador principal que va del 0% al 100%.

Tipos de gráficos / Partes de un total

Mapas de árbol (Treemap)

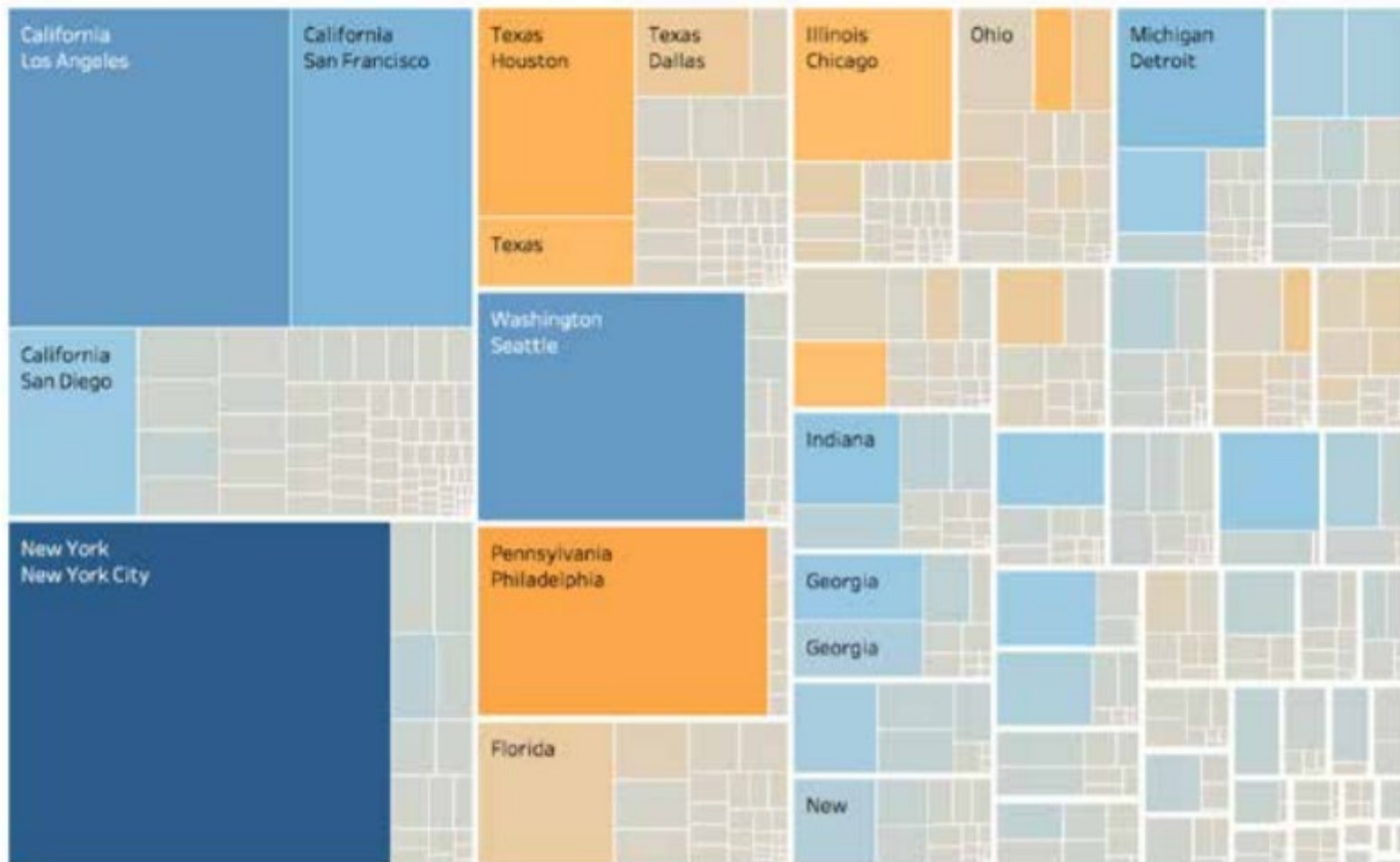


Figura 4.37. Mapa de árbol que representa el volumen de ventas (representado con la medida) y el beneficio (representado con el color) de una empresa ficticia en diferentes ciudades americanas agrupadas por estado.



Tipos de gráficos / Distribución

Histograma

Para entender como están distribuidos los valores de una variable.

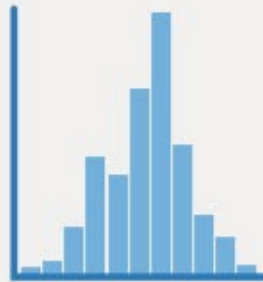
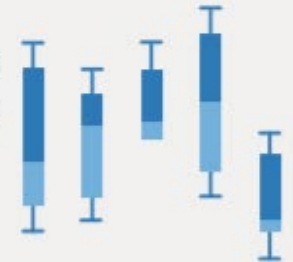


Diagrama de caja

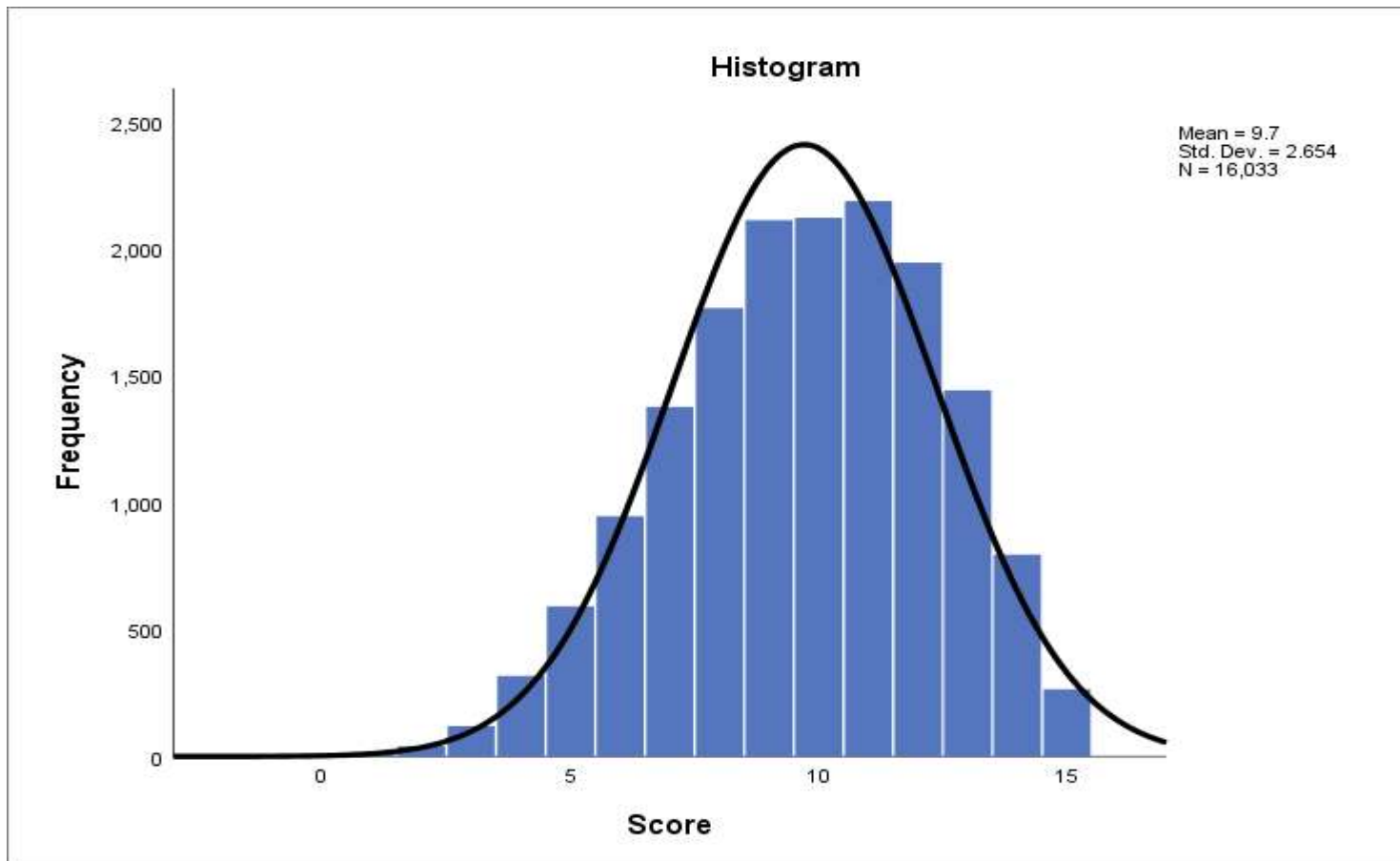
Para comparar la distribución de diferentes variables o la distribución de una de ellas en diferentes categorías.





Tipos de gráficos / Distribución

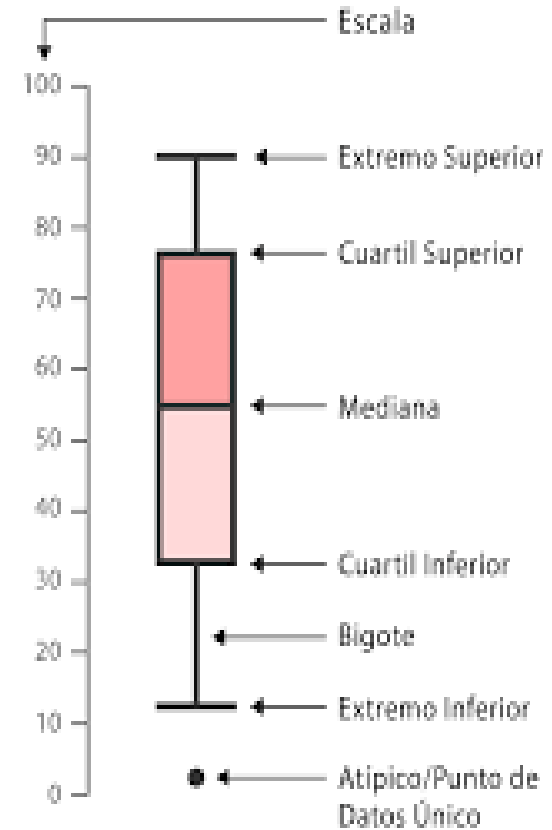
El **eje y** (a la izquierda) representa un recuento de frecuencia y el **eje x** (en la parte inferior), el valor de la variable (en este caso, el número de respuestas correctas).





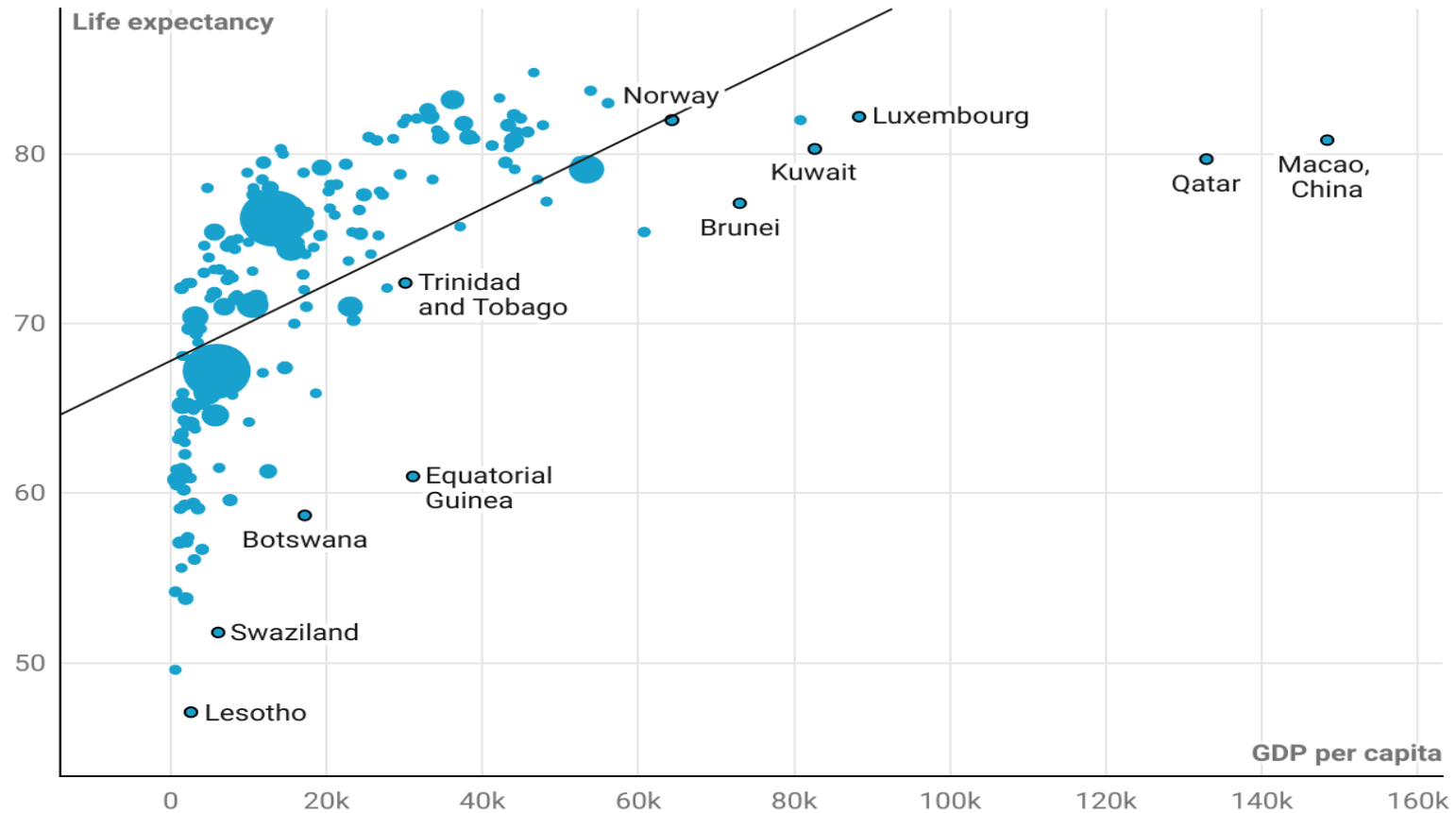
Tipos de gráficos / Distribución

- La caja está dividida por una línea que representa la **mediana**.
- Los extremos de la caja corresponden **al primer y al tercer cuartil**.
- De la caja, emergen unas líneas denominadas “**bigotes**”, que pueden calcularse de diferentes maneras y terminan en los valores **máximo y mínimo**
- Todos los puntos que se encuentren más allá de los bigotes se consideraran **valores atípicos**.



Tipos de gráficos / Dispersión

GDP per capita in US-Dollars and life expectancy in years for selected countries, 2015



Fuente: Gapminder • Creado con Datawrapper

Tipos de gráficos / Redes, conexiones

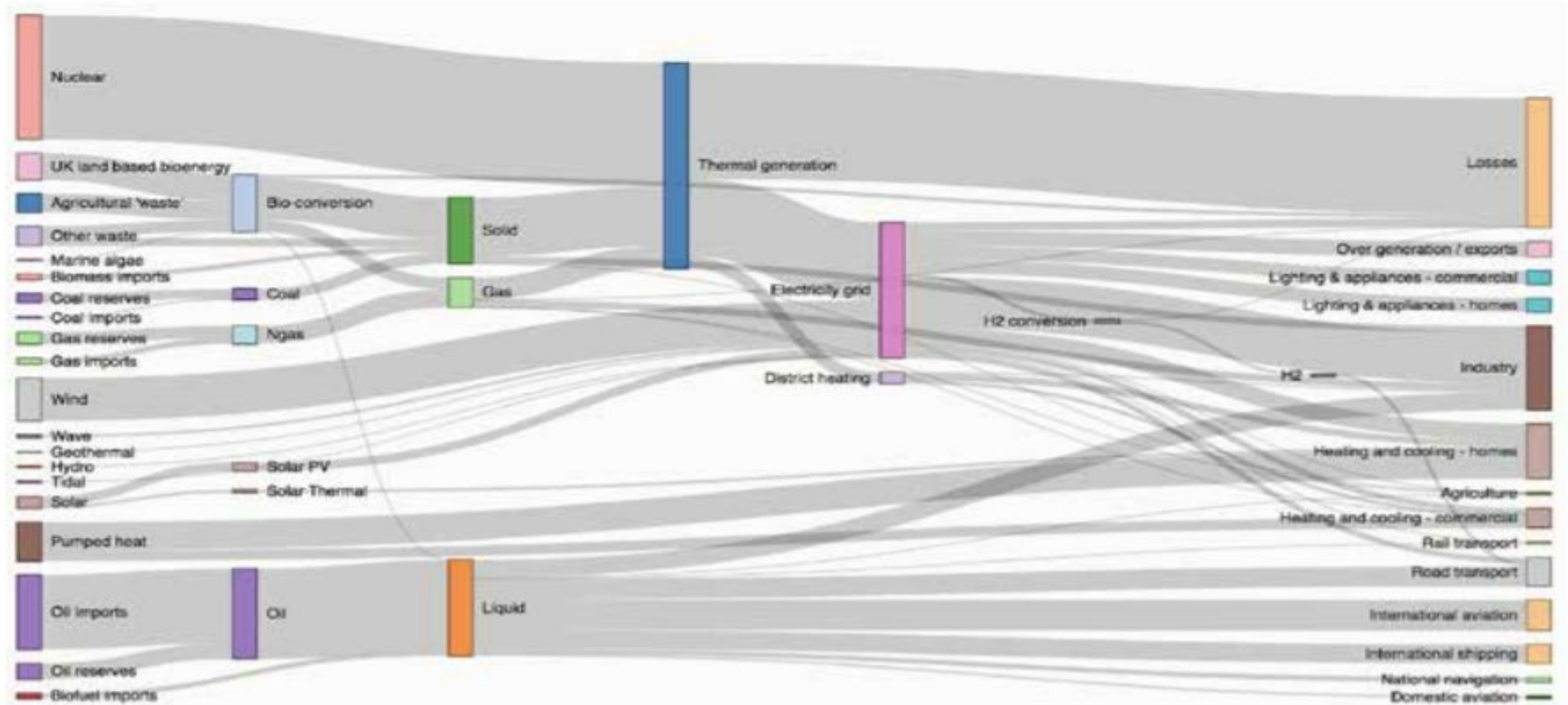


Figura 4.52. Diagrama de Sankey que muestra la cadena de producción y consumo de energía eléctrica.
Fuente: <https://bost.ocks.org/mike/sankey/>.

El diagrama de Sankey muestra diferentes categorías o estados a través de los cuales una variable va cambiando de valor. También muestra que una categoría es la suma de los valores de diferentes categorías, tal y como se aprecia en el ejemplo.

Tablas vs gráficos



- Las tablas son conjuntos ordenados y sistemáticos de números y/o palabras presentados en filas y columnas.
- Su principal finalidad es la ordenación y presentación de información de tipo repetitivo en una forma comprensible
- Las tablas constituyen un complemento muy útil del texto, pero no deben duplicarlo.



Simple Tips & Tricks

“

Using Tables and Figures in Your Manuscript

”

A (kg)	B (m)	C (s)

- A clear, descriptive, and concise title placed above the table.
- A good set of column titles (with units/unit symbols wherever applicable) that allow the reader to grasp the context of the table.
- Data should read from top to bottom.

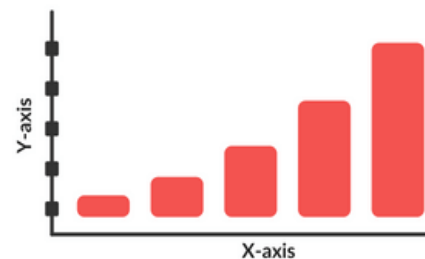


Figure 1. ABC

- A clear, descriptive, and concise caption placed below the figure.
- A high-quality image with good resolution and appropriate size.
- Can include bar graphs, histograms, maps, scatter plots, etc.

Wish to learn more about writing scientific names in your manuscript? Visit [enago.com/academy](https://www.enago.com/academy) now!



@Enago



enagoacademy



[enago.com/academy](https://www.enago.com/academy)



[enago.com/app](https://www.enago.com/app)

<https://www.enago.com/es/academy/how-to-use-tables-and-figures-to-effectively-organize-data-in-research-papers/>

Tablas vs gráficos

El propósito de las tablas y los gráficos es comunicar de manera efectiva.

!NO ENTRETENER !



Tablas / Buenas prácticas

- **Delinear columnas y filas**
- **Orden de forma lógica y secuencial**
- **Orden de izquierda a derecha**
- **Agregar datos resumen**
- **Evitar líneas**
- **Ordenar datos por categorías**
- **Tabla limpia**



Tablas / Buenas prácticas

- **Mantener consistencia entre los grupos y en los datos (Ej: n(%))**
- **Encabezamientos en negrita, cortos y precisos**
- **Autoexplicativa**
- **Sólo el título de la tabla debe estar en cursiva**
- **No olvide la fuente**



Tablas / Buenas prácticas

Tabla . Conocimiento de estudiantes sobre las herramientas de verificación de noticias o de información por universidad

	Conocimiento de las herramientas			Total, n (%)*
	Conozco nada	Conozco poco	Conozco mucho	
Universidad				
UANL, n (%)	13 (9,35)	100 (71,9)	26 (18,7)	139 (35,6)
UDEM, n (%)	56 (22,2)	150 (59,3)	47 (18,3)	252 (64,4)
Total	n = 69 (17,8)	n = 250 (63,9)	n = 72 (18,4)	n = 391

Tablas / Buenas prácticas

Tabla 2. Presencia en redes sociales y altmetrics de las publicaciones Web of Science (2009-2013) de los investigadores más productivos de EPI

INVESTIGADOR Y CENTRO	PRESENCIA EN REDES SOCIALES						ALTMETRICS PARA TRABAJOS WEB OF SCIENCE (WoS)								
	Presencia en Twitter	Seguidores Twitter	Presencia en Mendeley	Presencia en Google Scholar	Presencia en LinkedIn	Presencia en SlideShare	Número de trabajos en WoS	Número de tweets	Medio de tweets	Nº de lecturas en Mendeley	Medio lecturas en Mendeley	Nº bookmarks en CiteULike	Medio bookmarks CiteULike		
ABADAL-FALGUERAS, ERNEST UB	✓	186	✓	✓	✓	137	19509	15	18	1,3	189	9	4	0,2	
ALEIXANDRE-BENAVENT, RAFAEL UV	○	○	○	✓	✓	○	○	47	8	0,3	200	3,2	4	0,1	
ALONSO-AREVALO, JULIO USAL	✓	2.107	✓	✓	✓	✓	600	247.336	5	3	0,5	221	74	58	19
ARROYO-VAZQUEZ, NATALIA FGSR	✓	2.930	✓	✓	✓	✓	196	150.041	3	0	0	28	9,3	0	0
CABEZAS-CLAVIJO, ALVARO UGR	✓	1.154	✓	✓	✓	✓	172	88.882	10	110	11	168	7,3	14	0,6
CODINA-BONILLA, LLUIS UPF	✓	456	✓	✓	✓	✓	163	42.282	14	0	0	114	6,7	2	0,1
CORDON-GARCIA, JOSE-ANTONIO USAL	✓	780	✓	✓	○	✓	174	41.071	6	3	0,5	246	82	58	19
DELGADO-LOPEZ-COZAR, EMILIO UGR	○	○	○	○	○	✓	20	36.415	23	48	2,7	301	3,4	49	0,6
DE-MOYA-ANECON, FELIX CCHS-CSIC	✓	622	✓	✓	○	○	○	52	101	2,1	750	2,9	69	0,3	
DIAZ-NOCL, JAVIER UPF	✓	83	✓	✓	○	✓	32	302	7	4	0,6	85	4,3	0	0
FEIJOO-GONZALEZ, CLAUDIO UPM	○	○	○	○	○	○	○	5	4	0,2	144	18	10	1,3	
FERRER-SAPENA, ANTONIA UPV	○	○	○	○	○	○	○	9	24	3	124	11	2	0,2	
FUMERO-REVERON, ANTONIO UPM	✓	4.033	✓	✓	✓	✓	630	8.498	3	0	0	6	2	2	0,7
GARCIA-MARCO, FRANCISCO-JAVIER UNIZAR	○	○	○	○	○	○	○	5	3	0,6	59	30	4	2	
GIMENEZ-TOLEDO, ELEA CCHS-CSIC	○	○	○	○	○	○	○	16	8	0,7	53	1,8	3	0,1	
GOMEZ-BARROSO, JOSE-LUIS UNED	○	○	○	○	○	○	○	27	4	0,2	159	3,1	10	0,2	
GONZALES-AGUILAR, AUDILIO UPVM3	✓	56	✓	✓	○	○	○	3	0	0	8	2,7	2	0,7	
GONZALEZ-PACANOWSKI, TONI Prestigia	✓	232	✓	○	○	○	○	3	0	0	25	25	4	4	
GUALLAR-DELGADO, JAVIER UB	✓	3.784	✓	✓	✓	✓	110	52.289	11	8	0,9	56	2,2	0	0
HIDALGO-NUCHERA, ANTONIO UPM	○	○	○	○	○	○	○	9	0	0	70	35	5	2,5	
JIMENEZ-CONTRERAS, EVARISTO UGR	○	○	○	○	○	○	○	23	53	2,5	203	2,5	11	0,1	
LEYDESDORFF, LOET UvA	✓	150	✓	✓	○	○	○	114	161	1,5	2056	1,6	120	0,1	
MARCOS, MARI-CARMEN UPF	✓	1.010	○	○	○	○	○	9	0	0	38	13	0	0	
MARCOS-RECIO, JUAN-CARLOS UCM	✓	115	○	○	○	○	○	7	0	0	19	3,8	4	0,8	
MARTINEZ-MENDEZ, FRANCISCO-JAVIER UM	○	○	○	○	○	○	○	4	0	0	40	40	3	3	
MERLO-VEGA, JOSE-ANTONIO USAL	✓	3.554	✓	✓	✓	✓	58	35.276	3	0	0	33	33	5	5
MICO-SANZ, JOSEF-LLUIS URLL	○	○	○	○	○	○	○	4	0	0	31	5,2	0	0	
MOREIRO, JOSE-ANTONIO UC3M	○	○	○	○	○	○	○	6	4	0,8	63	32	8	4	
OLIVERA-ZALDUA, MARIA UCM	○	○	○	○	○	○	○	7	0	0	25	13	4	2	
ORDUNA-MALEA, ENRIQUE UPV	✓	251	✓	✓	○	○	○	11	38	3,5	56	6,2	2	0,2	
PASTOR-SANCHEZ, JUAN-ANTONIO UM	○	○	○	○	○	○	○	6	0	0	24	24	1	1	
PEDRAZA-JIMENEZ, RAFAEL UPF	✓	76	✓	✓	✓	✓	50	73.456	8	0	0	74	3,9	2	0,1
PEREZ-MONTORO, MARIO UB	✓	75	○	○	○	○	○	6	14	2,3	56	14	3	0,8	
PESET-MANCEBO, FERNANDA UPV	○	○	○	○	○	○	○	11	0	0	47	4,7	5	0,5	
RAMIREZ-POSADA, MARIA-VICTORIA UPVM3	○	○	○	○	○	○	○	4	0	0	8	2	2	0,5	
ROBINSON-GARCIA, NICOLAS UGR	✓	456	✓	✓	✓	✓	105	72.071	15	160	11	97	4,4	1	0,1
RODRIGUEZ-GAIRIN, JOSEF-MANUEL UB	○	○	○	○	○	○	○	6	4	0,7	56	56	2	2	
RODRIGUEZ-PARADA, CONCEPCION UB	○	○	○	○	○	○	○	4	0	0	21	5,3	2	0,5	
ROVIRA-FONTANALS, CRISTOFOL IUFP	✓	576	✓	✓	✓	✓	6	13.711	6	0	0	37	5,3	1	0,1
RUIZ-PEREZ, RAFAEL UGR	○	○	○	○	○	○	○	12	8	1	66	2,2	15	0,5	
SANCHEZ-CUADRADO, SONIA UC3M	○	○	○	○	○	○	○	9	11	1,6	48	16	5	1,7	
SANCHEZ-VIGIL, JUAN-MIGUEL CUM	✓	19	○	○	○	○	○	9	0	0	21	4,2	4	0,8	
SAORIN-PEREZ, TOMAS UM	✓	84	✓	✓	✓	✓	44	16.202	4	0	0	12	3	0	0
SERRANO-COBOS, JORGE UPV	✓	1.170	○	○	○	○	○	6	2	0,3	59	6,6	4	0,4	
TORRES-SALINAS, DANIEL UN	✓	1.001	✓	✓	✓	✓	270	314.488	33	137	4,4	464	4,1	60	0,5
TRAMULLAS-SAZ, JESUS UNIZAR	✓	1.656	✓	✓	✓	✓	147	223.123	8	11	1,6	46	12	2	0,5
XIFRA-TRIADU, JORDI UPF	○	○	○	○	○	○	○	18	10	0,6	79	5,3	7	0,5	

<https://revista.profesionaldelainformacion.com/index.php/EPI/article/view/epi.2014.jul.04/16961>

Tablas vs gráficos

- Para condiciones en las que necesita información definida, como cifras y valores precisos.
- Cuando son mucho datos que un gráfico no mostraría con facilidad
- Muy útil cuando muestras datos por grupos y categorías
- Se relacionan con un sistema de observación, lo que **nos permite procesar hechos y cifras de forma secuencial.**
- **Los gráficos más centrados en mostrar patrones y tendencias**
- **Un gráfico mejor opción para mostrar relaciones espaciales**





Principios de diseño



¿ Cómo alcanzar una visualización efectiva ?

- Tips para conseguirla?

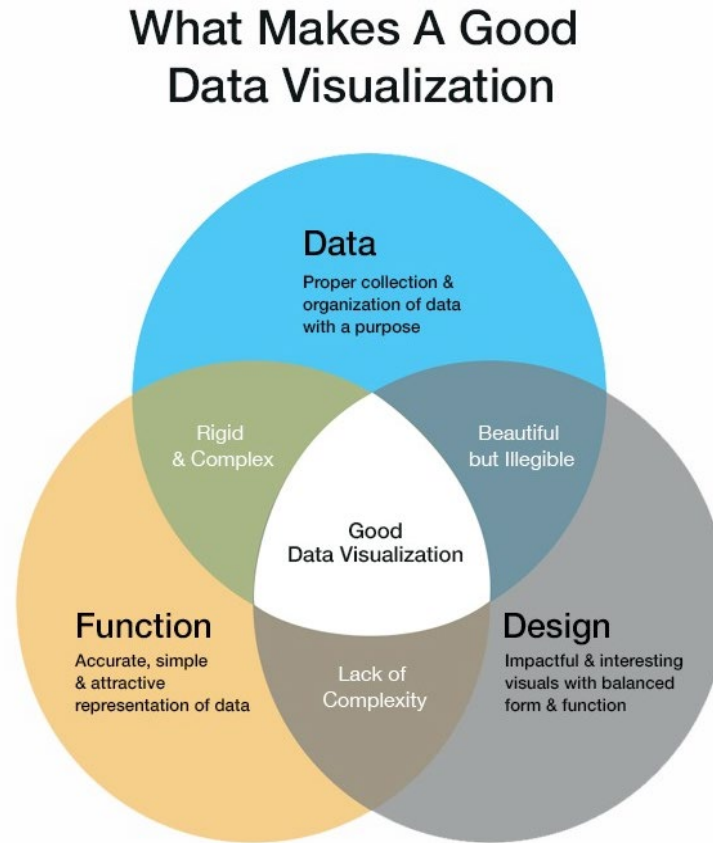
- “bonita” o funcional?

- Es el color importante ?

- Influencias científicas y/o psicológicas?



¿ Cómo alcanzar una visualización efectiva ?



<https://www.nabler.com/articles/the-data-visualization-manual-that-you-will-want-to-bookmark/>



Pensamiento visual



<https://towardsdatascience.com/how-to-create-beautiful-and-insightful-charts-e14a41c2cae0>



Al leer un gráfico sucede esto:

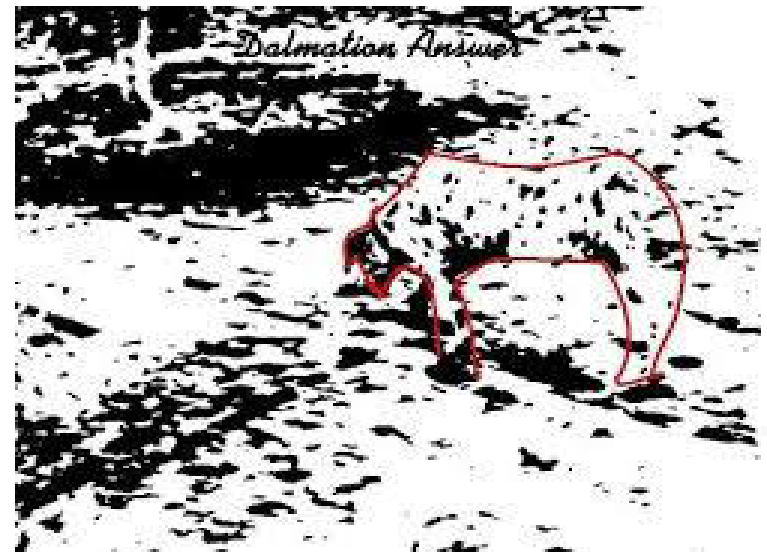
- Tus ojos buscan los números mencionados en la pregunta
- Tu cerebro los memoriza
- Tu cerebro compara los números organizados



Percepción visual y principios de Gestalt

Formar patrones complejos a partir de reglas simples...

¿Qué pueden ver en esta figura?



¿ Cómo alcanzar una visualización efectiva ?

Psicología Gestalt (Berlin, años 20)

Gestalt: Significa en alemán forma, patrón, figura o estructura unificada

Grupo de leyes que incluyen en como percibimos o intuimos patrones y conclusiones significativas de aquello que vemos

Permite:

- Mejor comprensión de la percepción humana
- Obtener mejores visualizaciones
- Acelerar el desarrollo eliminando ensayo /error





¿ Cómo alcanzar una visualización efectiva ?

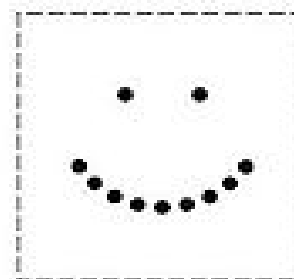
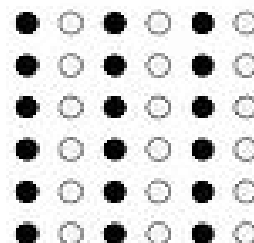
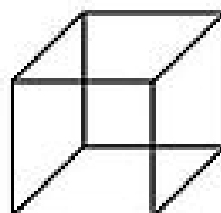
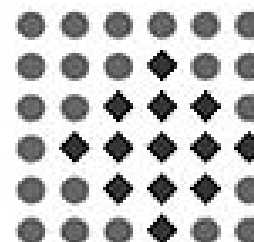
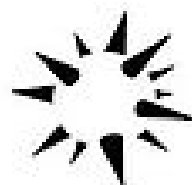
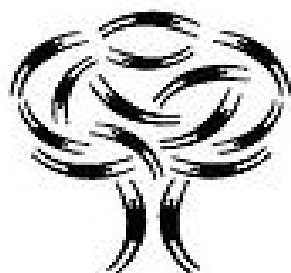
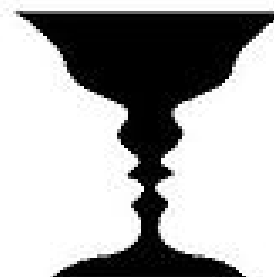
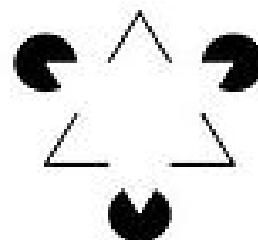
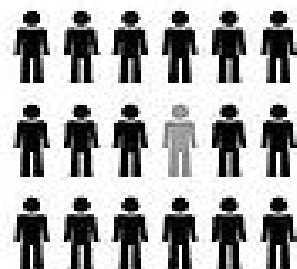
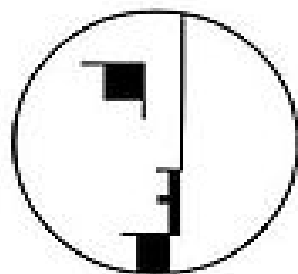
Psicología Gestalt (Berlin, años 20)

- Ley de la Proximidad*
- Ley de Similaridad*
- Ley de clausura*
- Ley de la Simetría*
- Ley de la misma dirección
- Ley de la inclusividad o figura de fondo*
- Ley de la continuidad*
- Ley de un buen Gestalt
- Experiencia pasada (Correspondencia isomórfica)
- Continuidad uniforme
- Regiones comunes





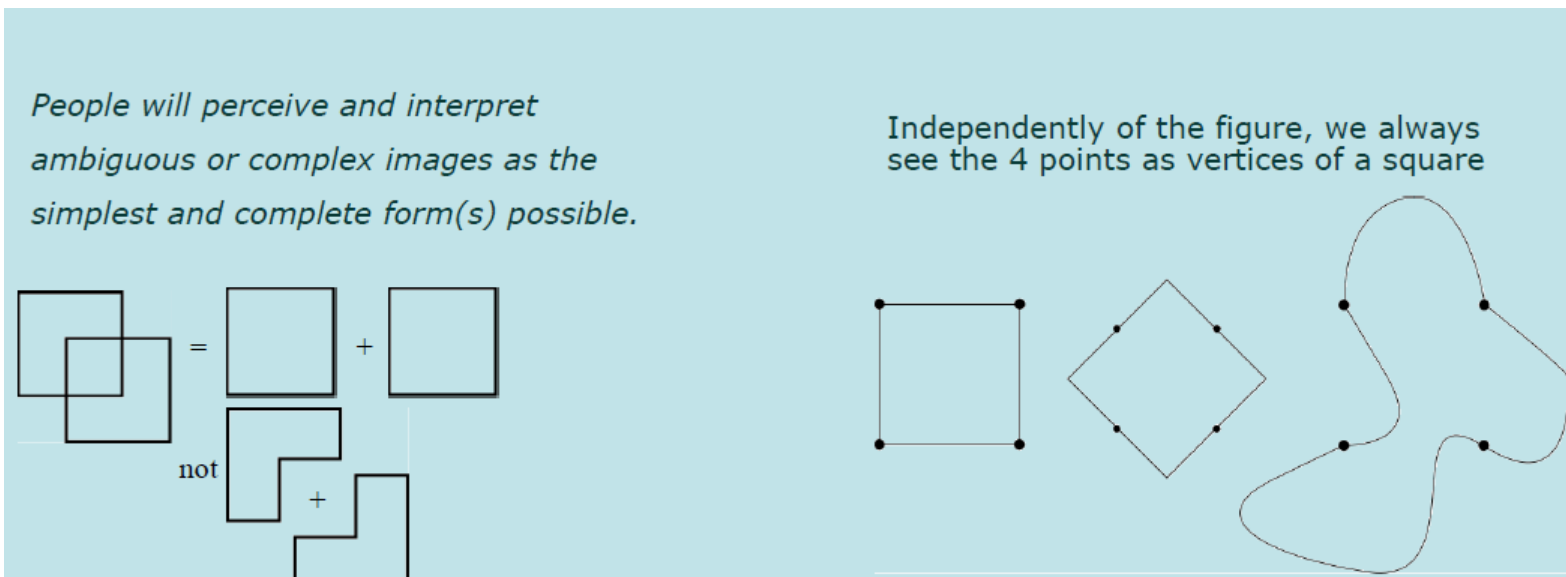
Principios de Gestalt



Pragnanz _ La simplicidad es la clave

Las personas percibirán e interpretarán imágenes ambiguas o complejas como la forma más simple posible.

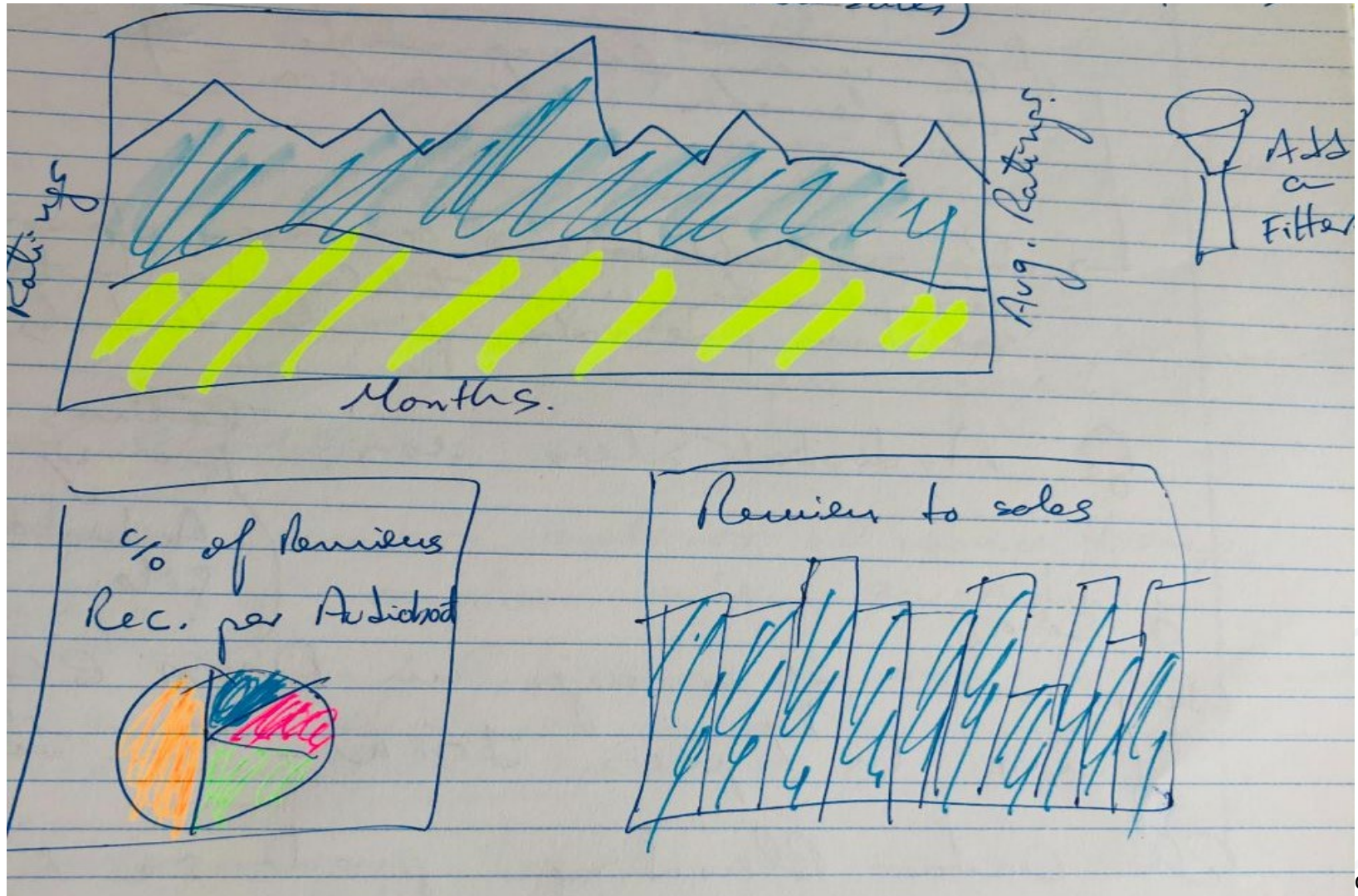
People will perceive and interpret ambiguous or complex images as the simplest and complete form(s) possible.



Independently of the figure, we always see the 4 points as vertices of a square

European Commission. 2017. 15TH JRC TRAINING ON COMPOSITE INDICATORS AND ANNUAL MEETING OF COP ON INDICES & SCOREBOARDS.

Dibuja antes. Con el Fin en mente



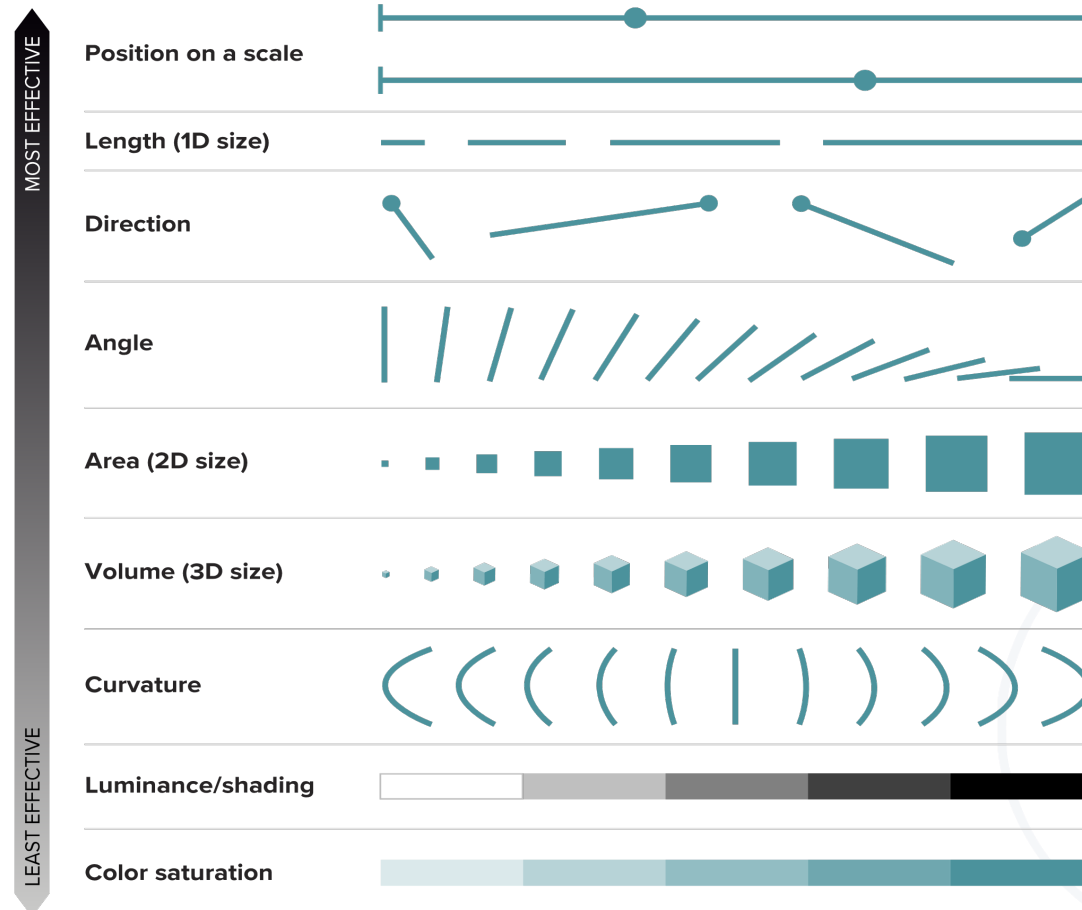
c2cae0

Elementos visuales útiles



Ranking of visual elements

Studies have identified the easiest ways for people to understand differences in quantitative data, on a scale from most effective to least.



SOURCES: W.S. CLEVELAND AND R. MCGILL / JOURNAL OF THE AMERICAN STATISTICAL ASSOCIATION 1984;
S.I. O'DONOGHUE ET AL / AR BIOMEDICAL DATA SCIENCE 2018

5W INFOGRAPHIC / KNOWABLE

Elimina Ruido. Reduce la carga cognitiva

Importancia del color.

El color siempre debe ayudar a 'leer' los datos

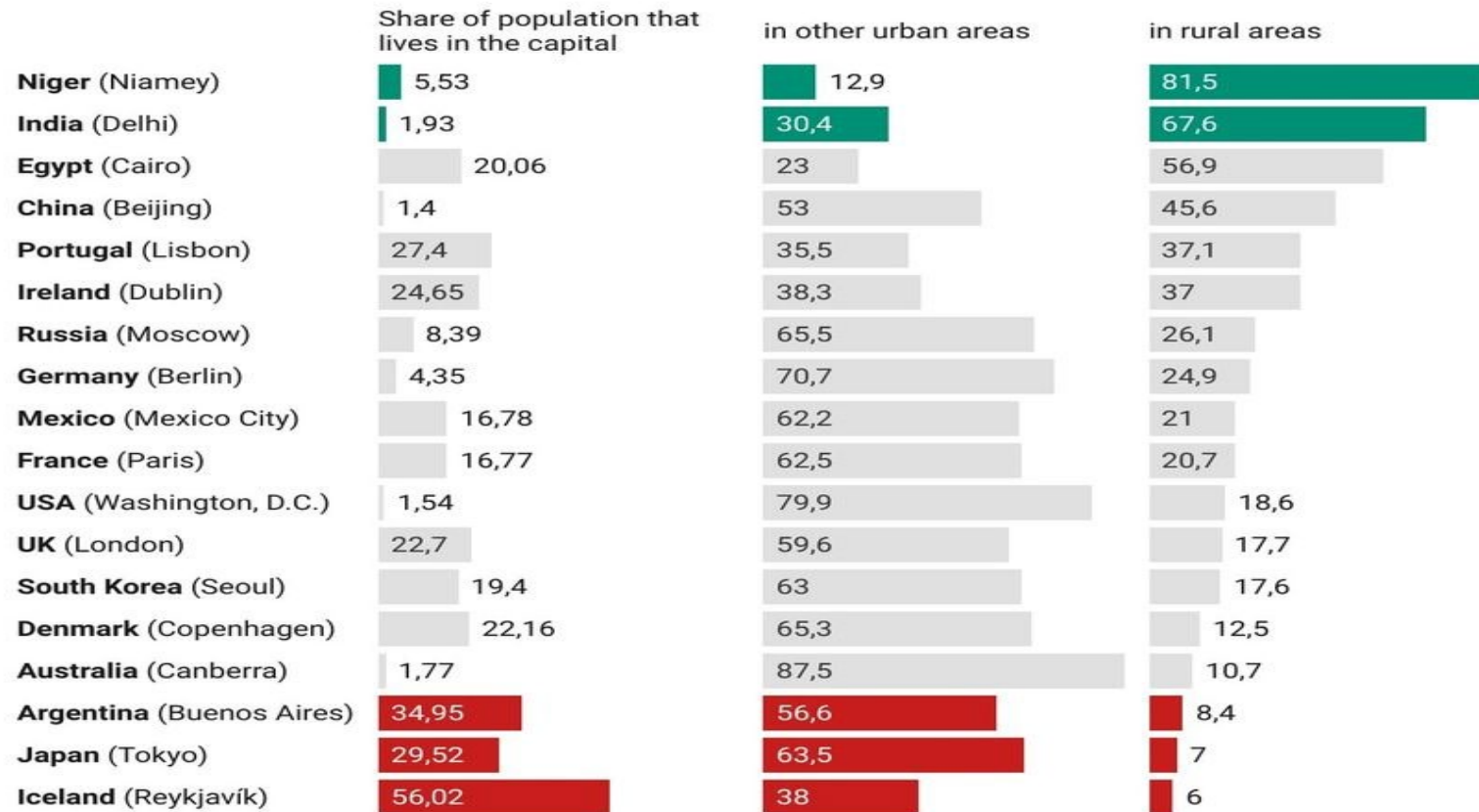
- Mostrar cantidad.
 - Información cuantitativa.
- Mostrar Pertenencia a un grupo.
 - Información cualitativa.
- Guiar al ojo (focalizar la atención).
 - Storytelling.



Uso del color

Diferencias en el porcentaje de la población que vive en la capital, otras áreas urbanas, y en áreas rurales

Una selección de países, 2014.



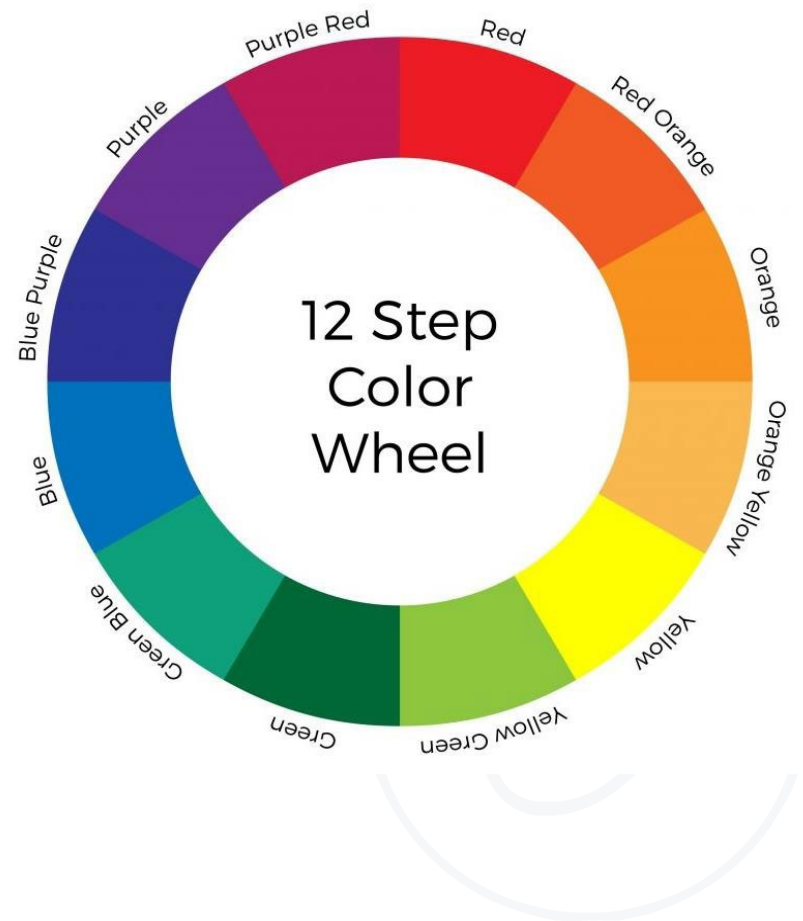
The UN defines "Urban" differently for each country. To count as "urban", Japanese settlements need to have at least 50,000 inhabitants. In Iceland, 200 inhabitants are enough.

Gráfico: Yusnelkis Milanés Guisado • Fuente: UN Population Division • Creado con Datawrapper

Uso del color

Importancia del color.

- Use el color con moderación
 - El gris es tu amigo
- Elementos relacionados tienen colores relacionados
- **Evitar:**
 - Fusionar rojo y gris (Se usa para daltónicos)
 - Estereotipos (Rosa-Mujer, Azul_Niño)



Uso del color

Importancia del color.

- Limite el uso de muchos colores, a no ser q sea necesario.
- Legibilidad: Suficiente contraste entre fondo y figura
- Uso estratégico del color para destacar mensajes
- Mantener colores culturales en mente (Ej: Semáforo)

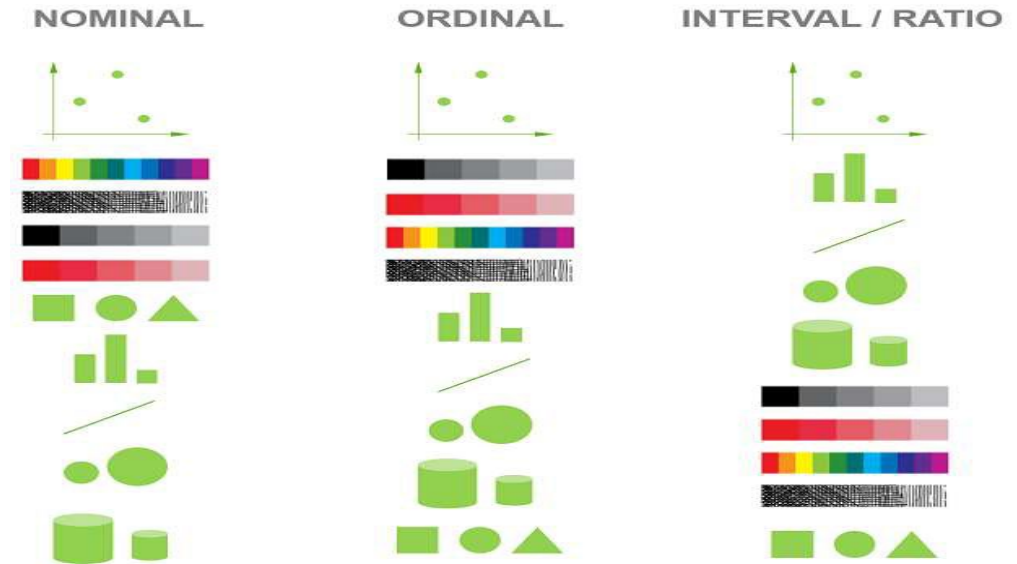


Uso del color

Utilizamos gradientes para datos cuantitativos.

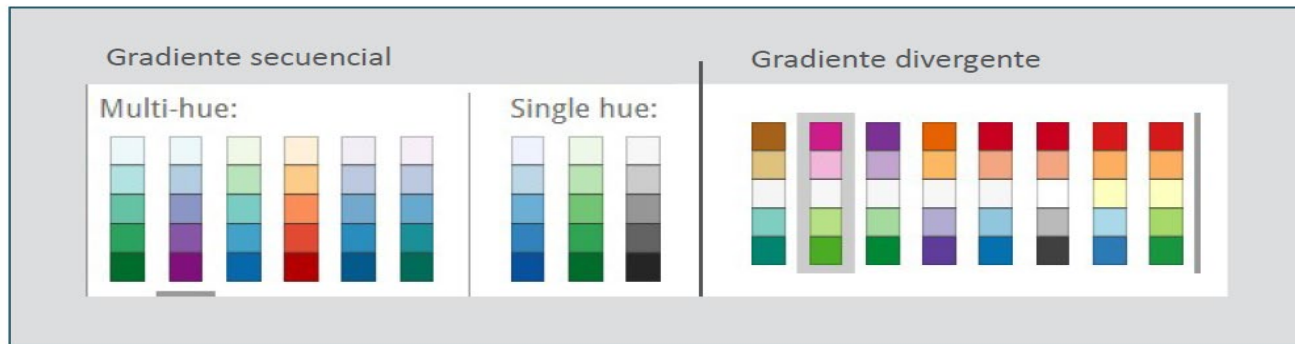
- Gradientes secuenciales.
- Gradientes divergentes.

Utilizamos una composición de colores para representar información cualitativa.

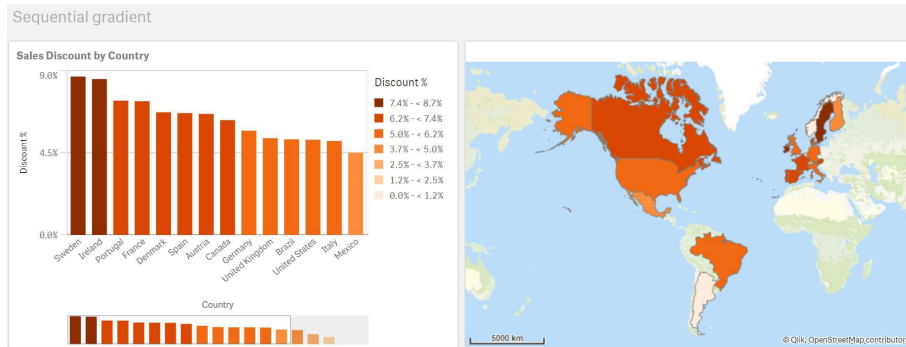


Uso del color

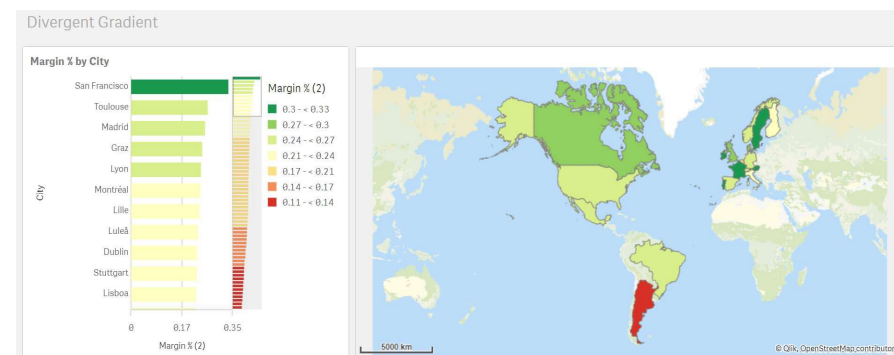
Información Cuantitativa



<http://colorbrewer2.org/>



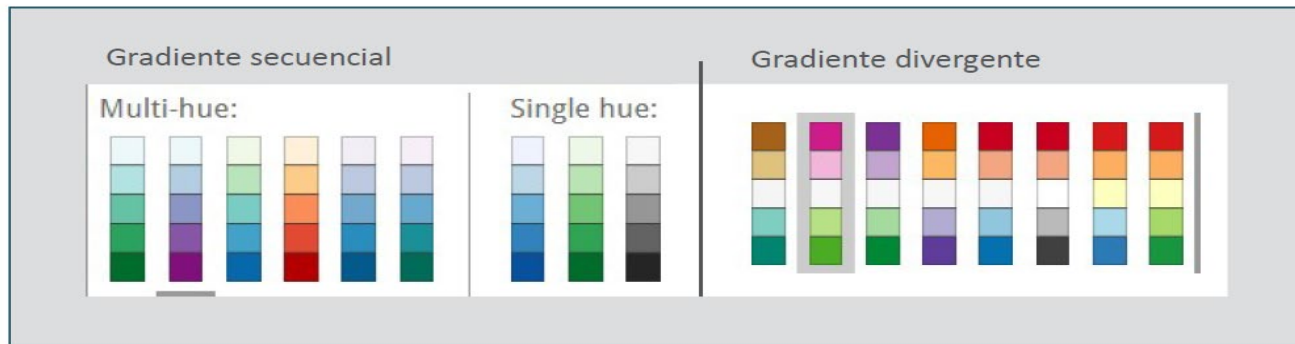
Los datos van de pequeño a grande.



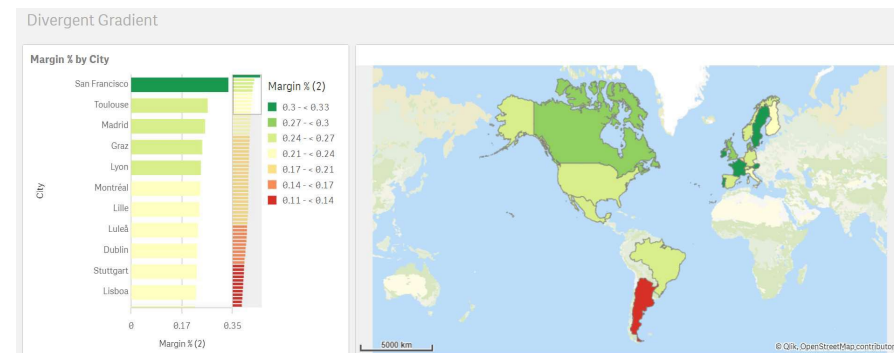
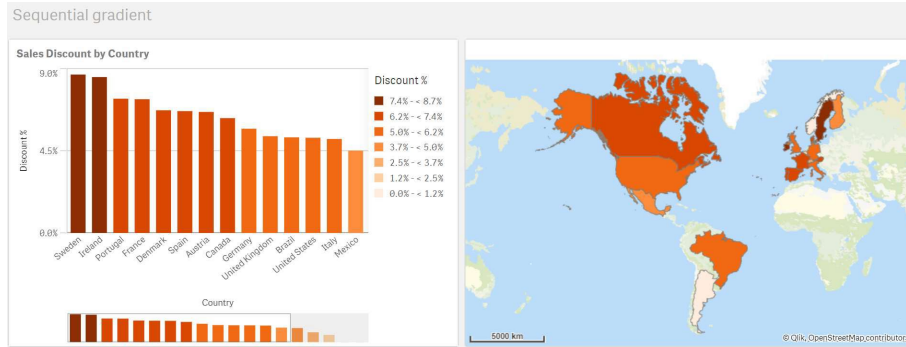
Mostrar desviaciones de un punto central.

Uso del color

Información Cuantitativa



<http://colorbrewer2.org/>

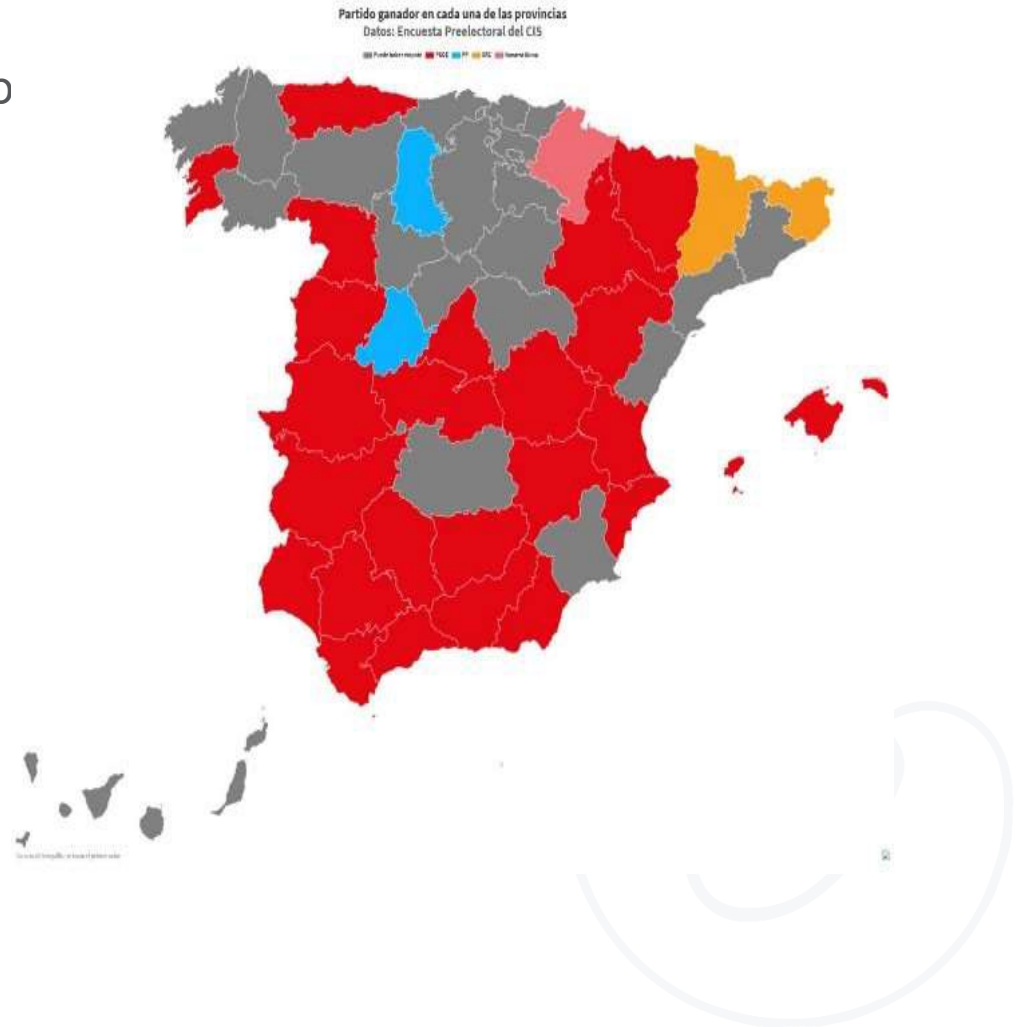


Los datos van de pequeño a grande.

Mostrar desviaciones de un punto central.

Uso del color

- Algunos Colores tienen significado propio
- Pertenencia a un grupo.
 - Métricas.
 - Dimensiones.

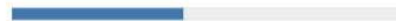


Información cualitativa / Uso del color

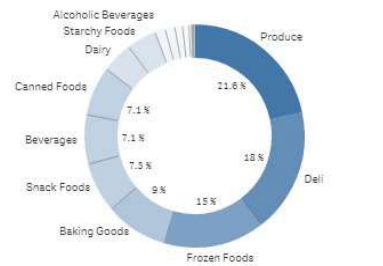
- Cada KPI tiene un color distinto.
- Los colores son persistentes.
- Evitamos leyendas.

Margin

Margin = 44%



By Product Sub Group

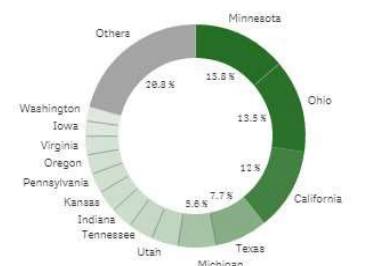


TY vs LY Sales

TY vs LY Sales = 68%



By State

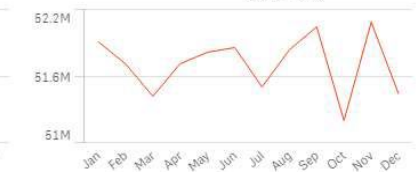
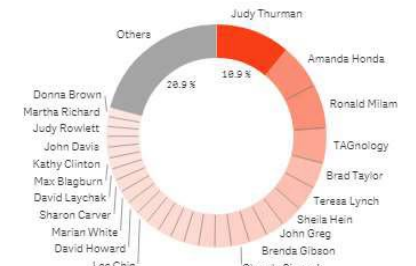


Budget

Budget vs Sales = 113%



By Sales Rep



Uso del color

Importancia del color.

Herramientas para la selección del color y sus relaciones

<http://colorbrewer2.org/#type=diverging&scheme=RdBu&n=5>

The screenshot displays the ColorBrewer 2.0 interface. At the top, it shows the URL and navigation links: 'how to use', 'updates', 'downloads', and 'credits'. The main title is 'COLORBREWER 2.0 color advice for cartography'. The interface is divided into several sections:

- Number of data classes:** Set to 5.
- Nature of your data:** Radio buttons for 'sequential', 'diverging' (selected), and 'qualitative'.
- Pick a color scheme:** A grid of 30 color scheme thumbnails.
- Only show:** Checkboxes for 'colorblind safe', 'print friendly', and 'photocopy safe'.
- Context:** Checkboxes for 'roads', 'cities', and 'borders' (checked).
- Background:** Radio buttons for 'solid color' (selected) and 'terrain', with a 'color transparency' slider below.
- 5-class RdBu:** A legend showing the color palette with corresponding HEX codes: #ca0020 (red), #f4a582 (orange), #f7f7f7 (white), #92c5de (light blue), and #0571b0 (dark blue).
- EXPORT:** A button to export the selected scheme.

The main map area shows a map of Spain with the selected 5-class RdBu diverging color scheme applied to its administrative boundaries. The map is divided into regions, with colors ranging from red in the north to blue in the south.

© Cynthia Brewer, Mark Harrower and The Pennsylvania State University
[Source code and feedback](#)
[Back to Flash version](#)
[Back to ColorBrewer 1.0](#)

axismaps

Uso del color

Data Visualization Foundations: Color

Color can be a great tool to add more information into your charts, but only if you use it wisely.



Patrik Lundblad

Fri Oct 30, 2015 | 2 min read



Hi! Welcome to Qlik. If you'd like to chat, I'm here to help you find your way.

KPI Dashboard

KPI Dashboard provides a high level look at the company with the ability to drill down into the business details.

Use this app to perform various analysis on these KPIs.

Sales analysis by segment, sales rep and product

Customer retention across products

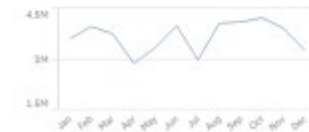
Sales rep performance by location



Margin

Margin = 44%

By Product Sub Group



TY vs LY Sales

TY vs LY Sales = 68%

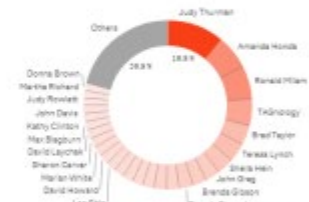
By State



Budget

Budget vs Sales = 113%

By Sales Rep



<https://www.qlik.com/blog/data-visualization-foundations-color>



Elimina Ruido. Reduce la carga cognitiva

Destaca lo importante





En una visualización gráfica, puede ser tan importante lo que muestras como lo que escondes.



Elimina Ruido. Reduce la carga cognitiva

- ✓ Escoge de 1 a 3 elementos a destacar
- ✓ Uso estratégico del contraste
- ✓ Focaliza la atención



Performance overview

■ Our business

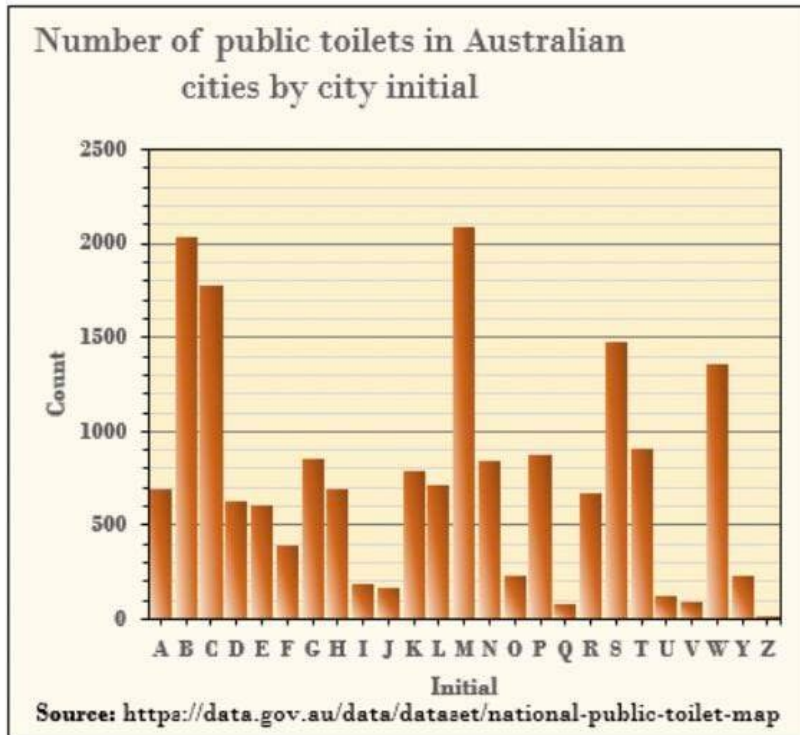
- Competitor A
- Competitor B
- Competitor C
- Competitor D
- Competitor E



Storytelling with data: a data visualization guides por business professionals

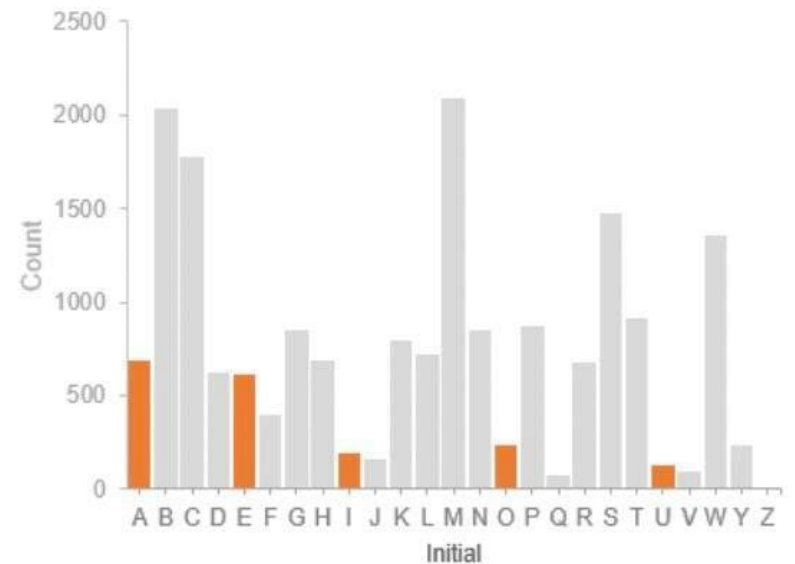
Presta atención a los detalles. Optimiza tu gráfico

Scenario: Your audience want to know how many public toilets in Australia are in cities that start with a vowel.



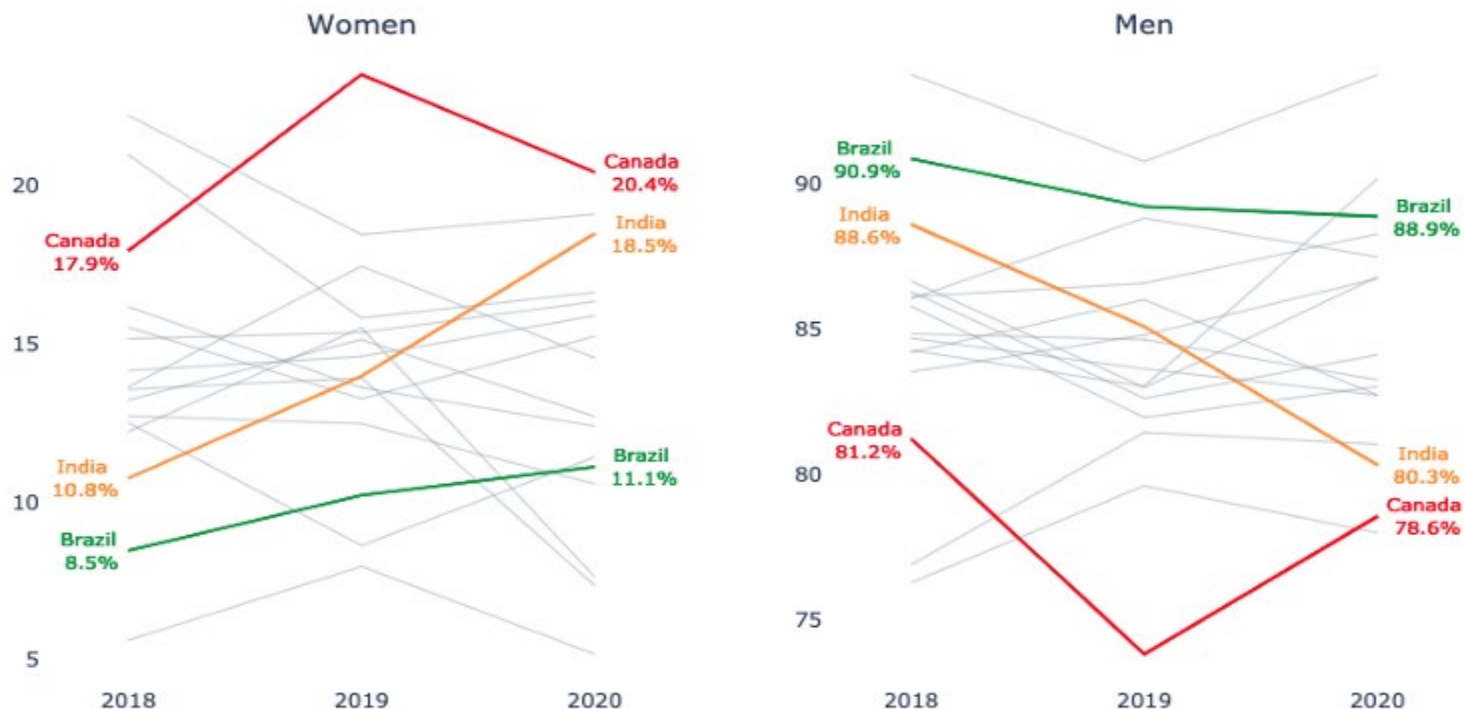
Less than 10% of public toilets are in cities starting with vowels

Number of public toilets in Australian cities, by city initial



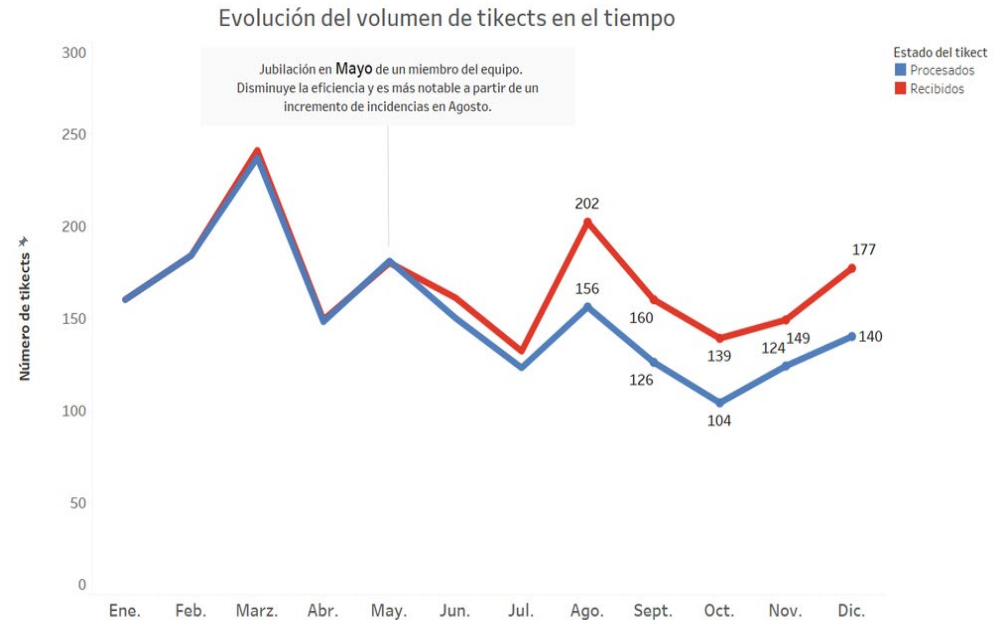
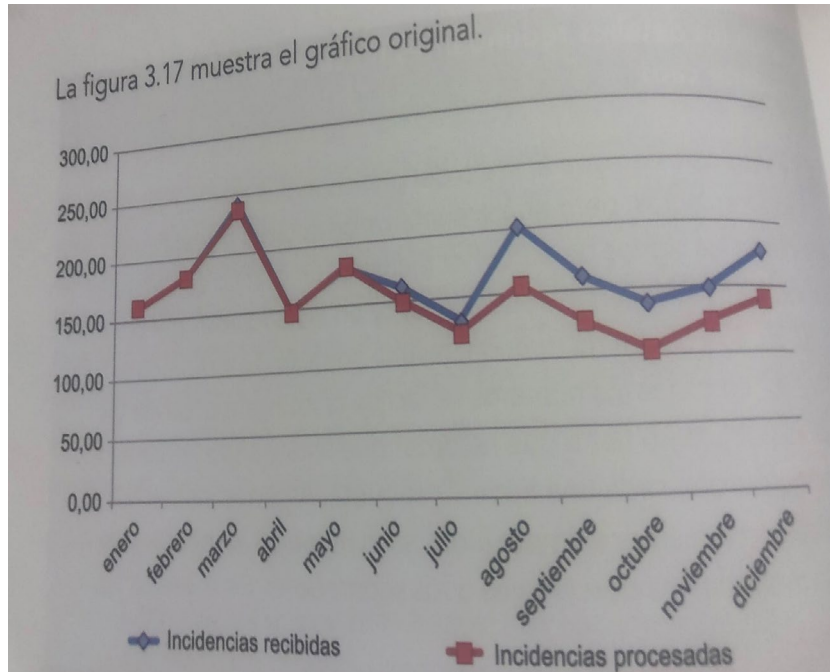
Presta atención a los detalles. Optimiza tu gráfico

Gender Gap: Kaggle members are mostly men. And there are no signs of increase in women participation



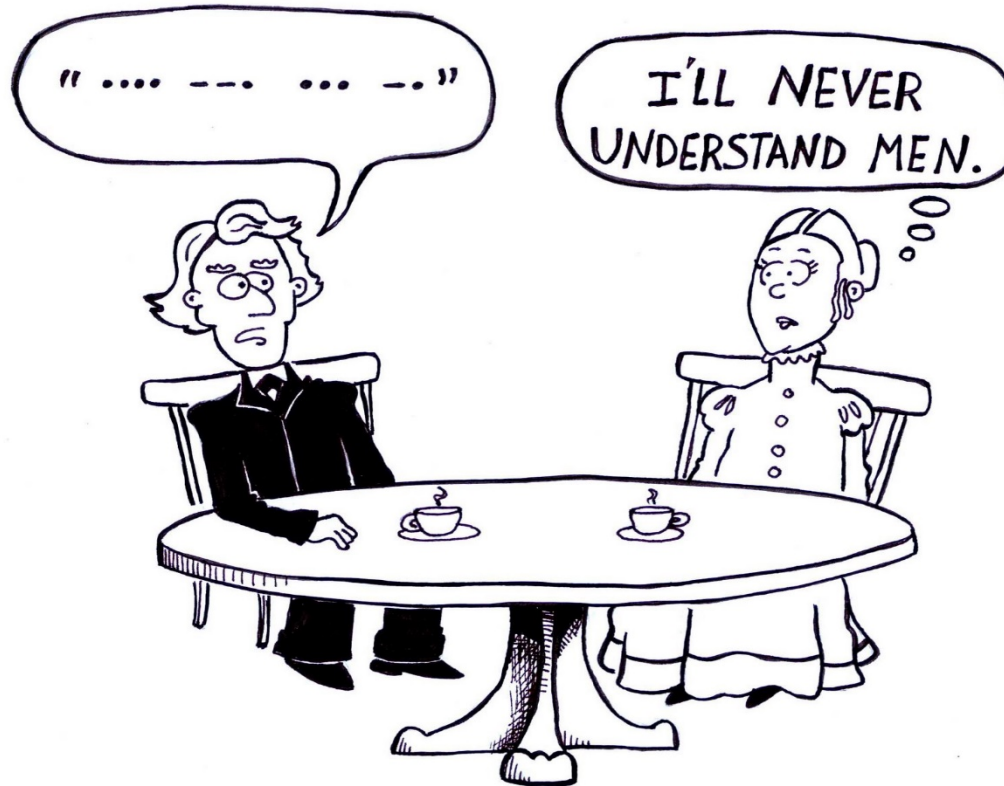
<https://medium.economist.com/mistakes-weve-drawn-a-few-8cdd8a42d368>

Ejemplo



Storytelling with data: a data visualization guides for business professionals

¿Cómo saber entonces que una visualización es buena ?



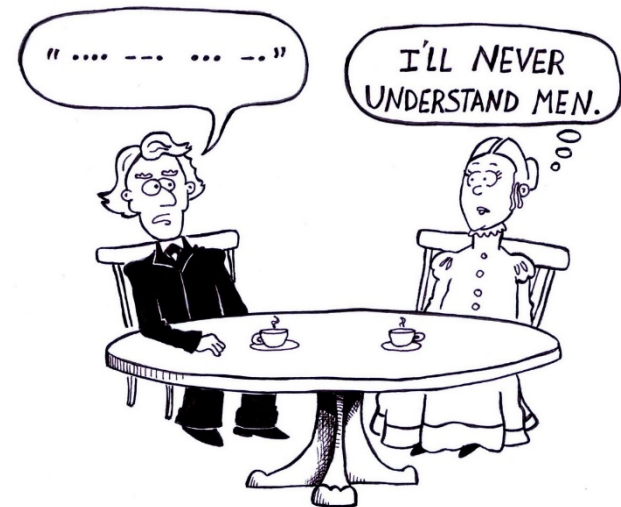
©2009 Isely/Lowrance

SAMUEL MORSE GOES ON A DATE.

<https://medium.com/multiple-views-visualization-research-explained/how-do-we-know-when-visualization-is-good-c894b5194b>

Si requiere mucha explicación, entonces **no es Bueno...**

- REFLEJA EL CONTEXTO: CONFORME A PROPÓSITO
- POCA CARGA COGNITIVA, MENOS ESFUERZO MENTAL
- Simple: MENOS ES MÁS
- SE DESTACA EL MENSAJE PRINCIPAL (La historia, las perlas)
- Tipo de gráfico adecuado en función del tipo de datos
- Detalles óptimos: No sesgos (Ejes cortados,)



©2009 ISEY/LORANGE

SAMUEL MORSE GOES ON A DATE.

<https://medium.com/multiple-views-visualization-research-explained/how-do-we-know-when-a-visualization-is-good-c894b5194b62>

Algunos tips para mejorar los gráficos

- Comienza con tu mensaje en mente. Cuál es tu mensaje?
- Intentar gráficos auto-explicativos
- Prepara bien los datos. Tipos de datos
- Facilita la lectura. Destaca lo importante
- Importancia del color
- La forma y el orden, según tipo de dato
- Elimina el ruido
- Agrega etiquetas estratégicas / contexto / Títulos y subtítulos
- Compara tus datos con (por ejemplo, con el objetivo o el promedio)
- Usa datos relativos que te permiten comparar, evitando sesgos !



Herramientas ...



ADOBE ILLUSTRATOR

Libros

1. Storytelling con datos. Visualización de datos para profesionales (TÍTULOS ESPECIALES) (Español) Tapa blanda – 14 septiembre 2017 **(ISBN: 8441539308)**
2. [Alberto Cairo. Functional Art, The: An introduction to information graphics and visualization.](#) **(ISBN-10 : 9780321834737)**
3. Stephen Few. **Show Me the Numbers: Designing Tables and Graphs to Enlighten (ISBN-10 : 0970601972)**
5. Dear Data. **(ISBN-10 : 9781846149061)**
6. Alberto Cairo. How Charts Lie: Getting Smarter about Visual Information. **(ISBN-10 : 1324001569)**
7. R for Data Science: Import, Tidy, Transform, Visualize, and Model Data.**(ISBN-10 : 1491910399)**

BLogs

- 1) Storytelling with data
- 2) Information is beautiful
- 3) FLOWingData
- 4) The Pudding
- 5) New York Times
- 6) The Data Visualization Catalogue
- 7) Dribbble
- 8) Medium.com



Datos de investigación en abierto

- [DataCite Search](#): buscador generalista de datasets abiertos con DOI
- [DataMED](#): buscador de datasets y repositorios en el ámbito biomédico
- [DataONE](#): buscador de datasets en repositorios sobre datos de la tierra y medio ambiente
- [DataSearch](#): buscador de datasets de Elsevier, en su mayoría de artículos de Elsevier
- [Dryad](#): buscador de datasets asociados a publicaciones de ciencias y medicina
- [e-cienciaDatos](#): buscador de datasets del Consorcio Madroño
- [Eudat B2FIND](#): buscador generalista de datasets mantenido por la organización
- [Figshare](#): buscador internacional en el que se pueden recuperar datos por categorías temáticas
- [Google Dataset Search](#): buscador de datasets en repositorios de datos científicos, bases de datos de gobiernos locales y nacionales, sitios web de editores y autores y otras fuentes.
- [Zenodo](#): buscador de datos de investigación de la Unión Europea
- European Open Science Cloud. [All resources - EOSC Marketplace \(eosc-portal.eu\)](#)

Datos en abierto

- [data.world](#)
- [Data is Plural](#)
- [UNdata](#)
- [Data.gov](#)
- [Kaggle](#)
- [Oficina Nacional de Administración Oceánica y Atmosférica \(NOAA\)](#)
- [Reddit](#)
- [World Fact Book](#)
- [Environmental Data Explorer de la ONU](#)
- [Organización Mundial de la Salud](#)
- [Pitney Bowes](#)
- [Work bank](#)
- <https://public.tableau.com/en-us/s/resources>



¿Preguntas?

Gracias

Yusnelkis Milanés Guisado

ymilgui@upo.es



U N I V E R S I D A D

PABLO^D
OLAVIDE

S E V I L L A