

Visualización de los resultados de tu investigación: Como construir gráficos útiles para las publicaciones científicas

Yusnelkis Milanés Guisado

Fechas: 30 May, 1 y 3 de
Junio de 9:30 a 11:30



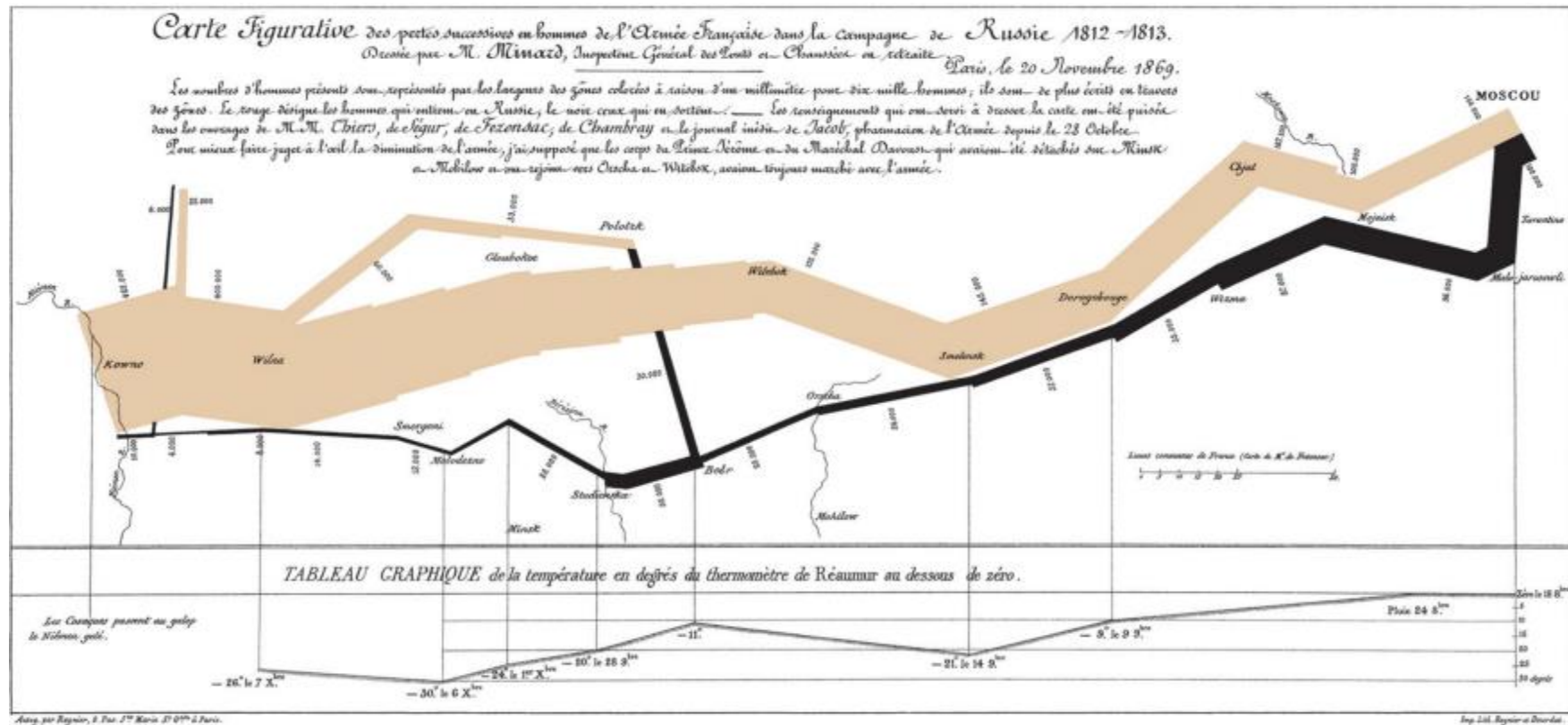
Programa

- Relevancia de la visualización de datos. Principios, procesos y buenas prácticas.
- Tablas y gráficos en tu investigación.
- Narrativa de datos
- Ejemplos a través de estudios de caso con distintas herramientas (Excel, Datawrapper, Flourish, Tableau)

Relevancia de la visualización de datos.

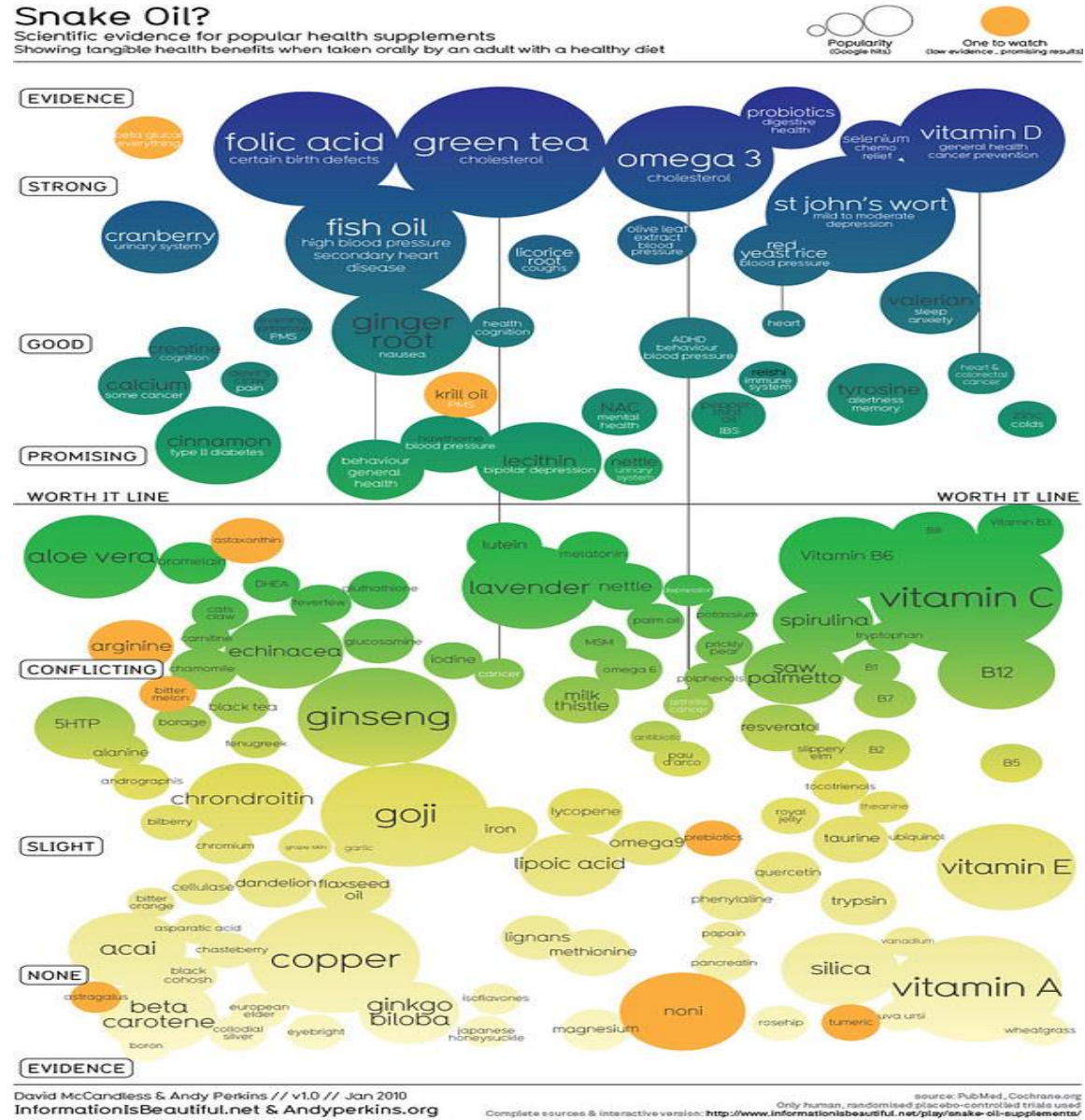
Historia de la Visualización

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.



Snake oil

The evidence for nutritional complements



<https://www.informationisbeautiful.net/visualizations/snake-oil-scientific-evidence-for-nutritional-supplements-vizsweet/>

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 and researchers, total times cited, and highly cited were counted and obtained, analyzed, and shown in Figures 2–6, and 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 an average annual growth rate of 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%.

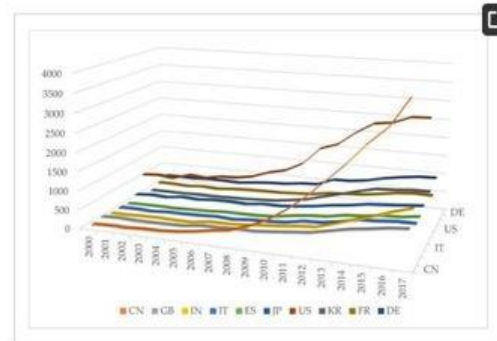


Figure 2. Temporal characteristics of China's scientific production in solar energy. (CN: China, GB: England, IN: India, IT: Italy, ES: Spain, JP: Japan, US: United States, KR: Korea, FR: France, DE: Germany)

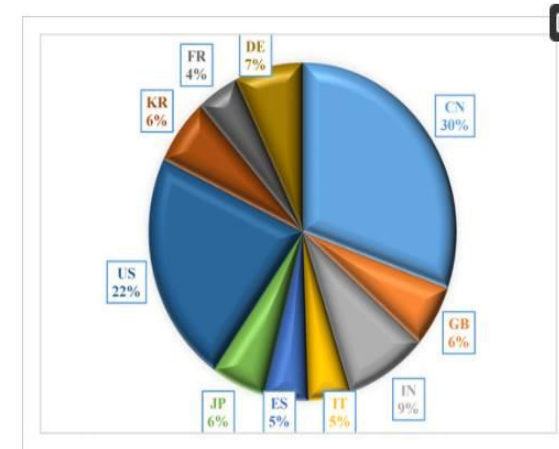
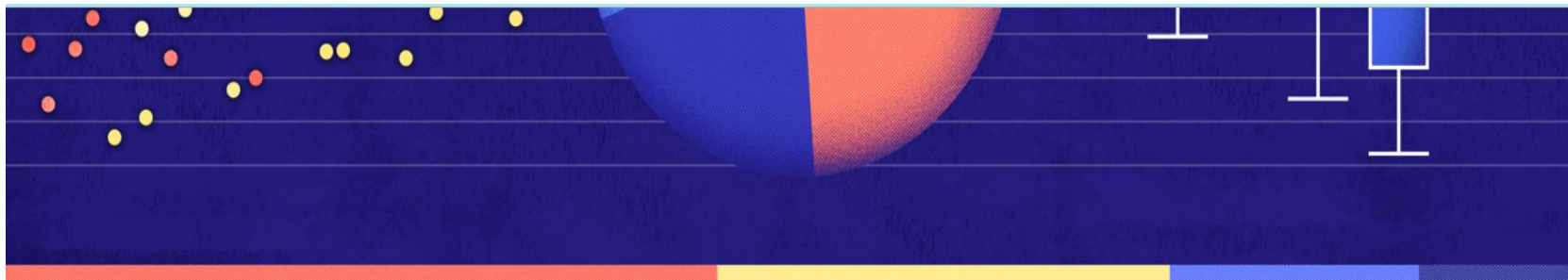


Figure 3. Scientific production of top 10 countries in 2017 by percentage.

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



— 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.

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

*...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...*

nature portfolio View all journals Search Q Login

nature > search

Search

Search Q [Advanced search](#)

Journal: **Nature Methods (7)** Article type: **All** Subject: **All** Date: **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)

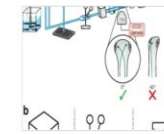
Bang Wong



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

Points of view: Points of review (part 1)

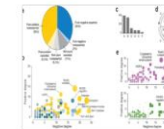
Bang Wong



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

Points of view: Points of review (part 2)

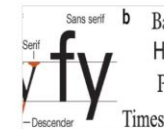
Bang Wong



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

Points of view: Typography

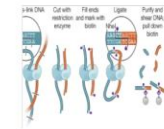
Bang Wong



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

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.



Volume 20
Issue 11
5 November 2019

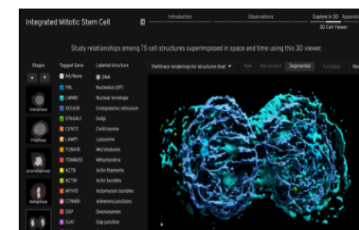
[IN THIS ISSUE](#)

[About the cover](#)

FIGURES | REFERENCES | RELATED | DETAILS

Figure 1.

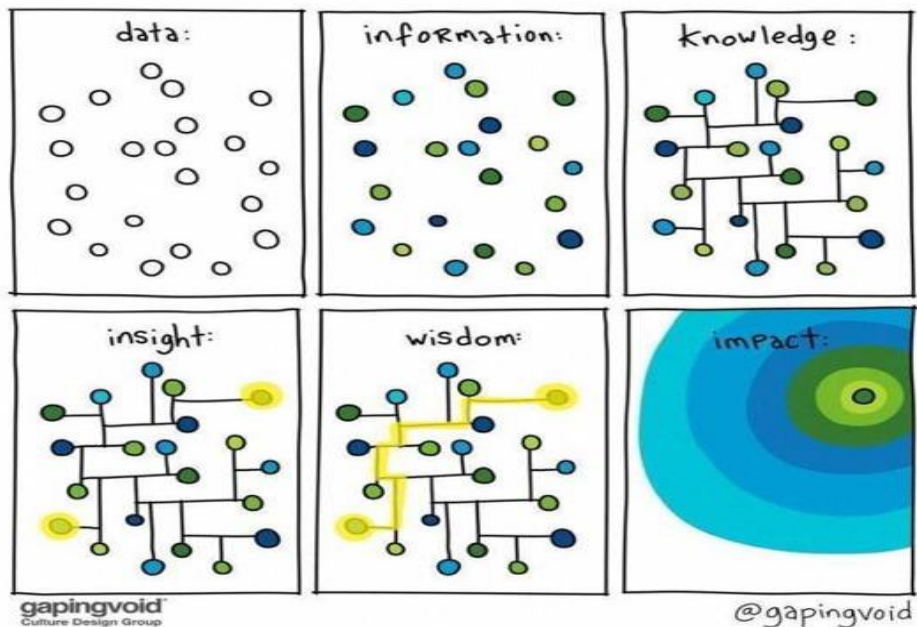
Analysis using Allen Cell Explorer



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

¿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?

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?

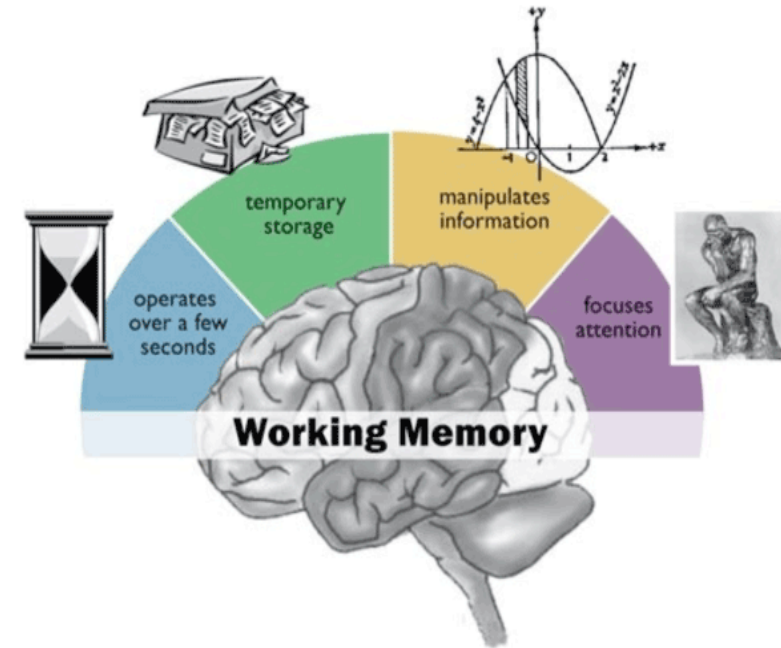
How many P's can you find in the text?

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?

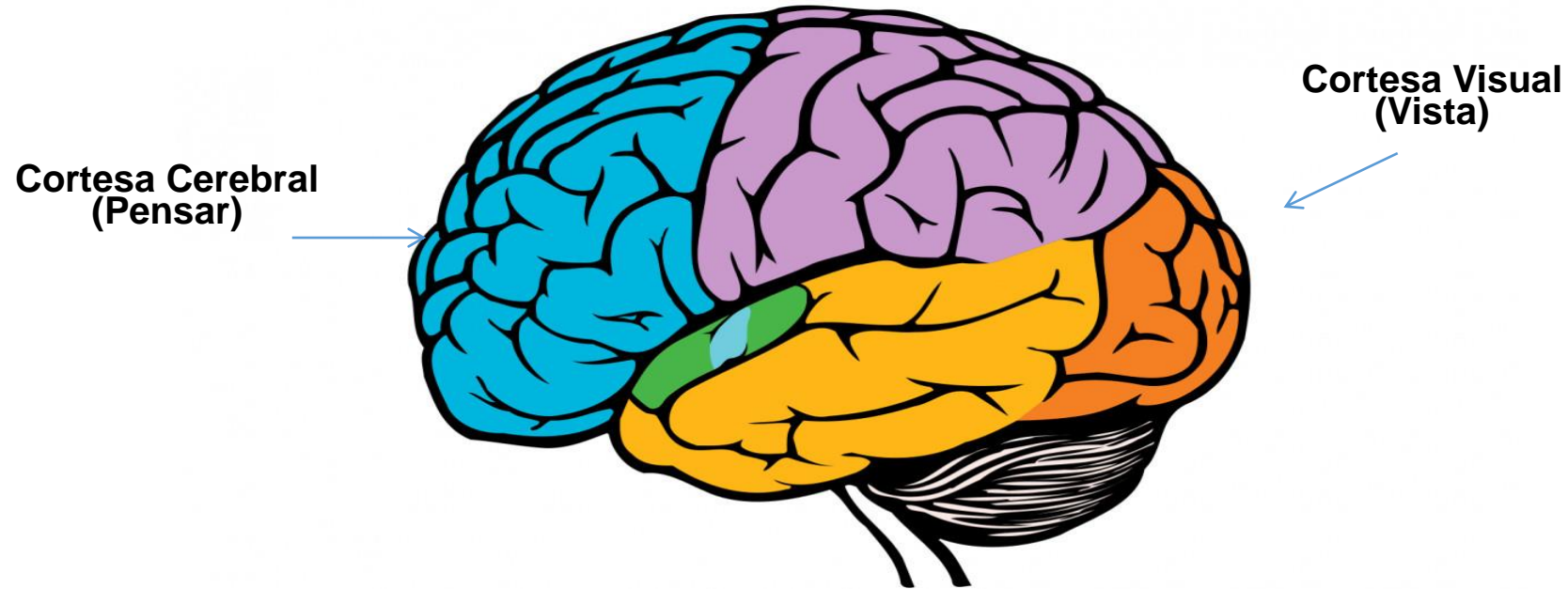
“Working Memory”

- 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....

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

¿ Cómo ocurre esto ?



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

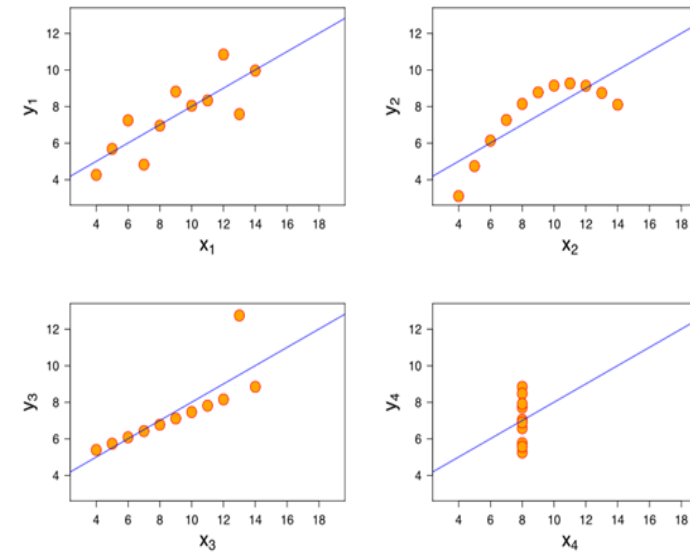
El cerebro humano es capaz de procesar una imagen completa en 13 milésimas de Segundo

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

¿Porqué la Visualización?

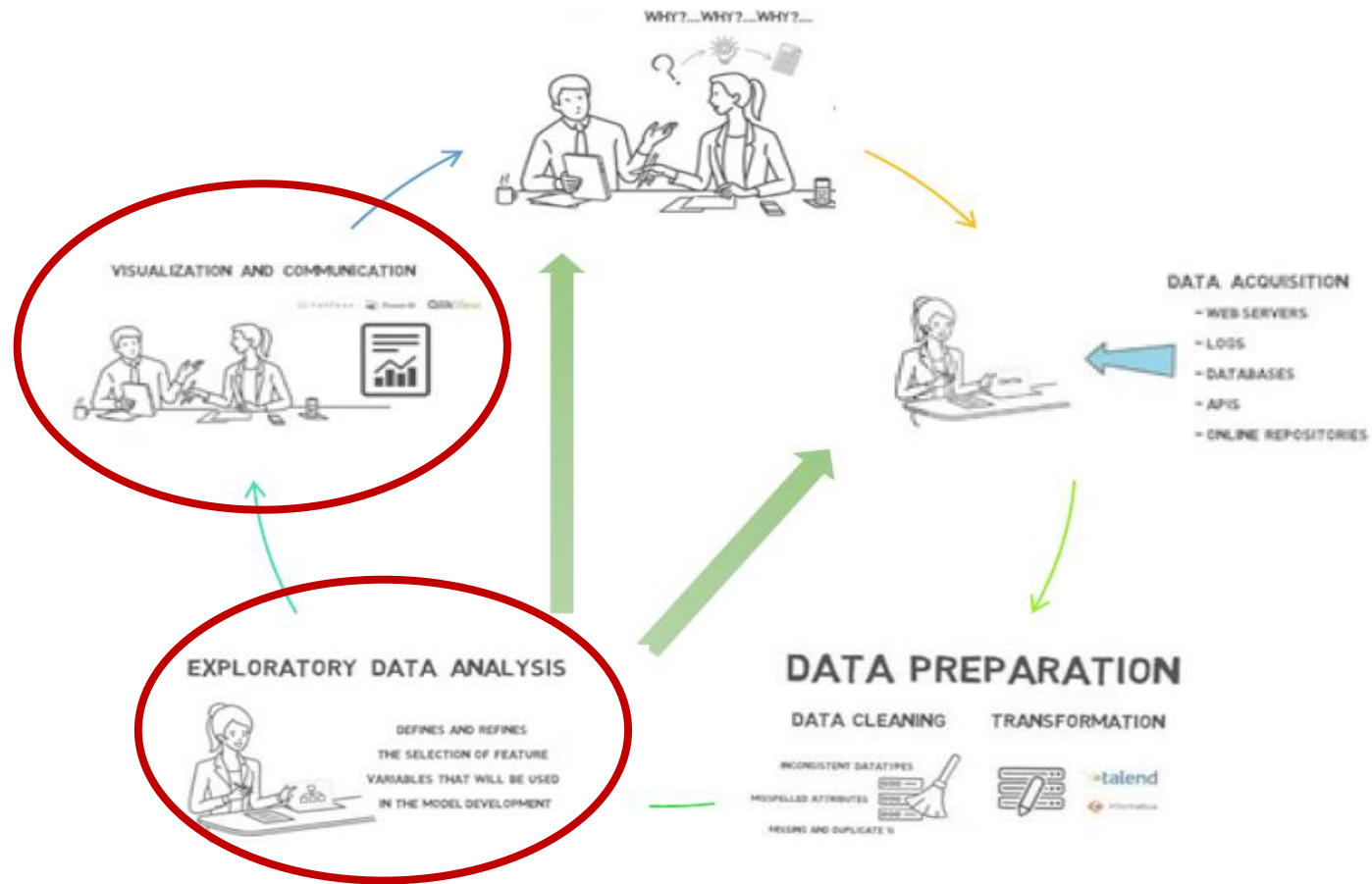
- Un resumen estadístico puede no contemplar tendencias importantes en un dataset.
- La visualización puede presentar las multiples miradas de un dataset.
- Puede funcionar como una importante primera etapa de investigación en una nueva área de estudio

Anscombe's Quartet



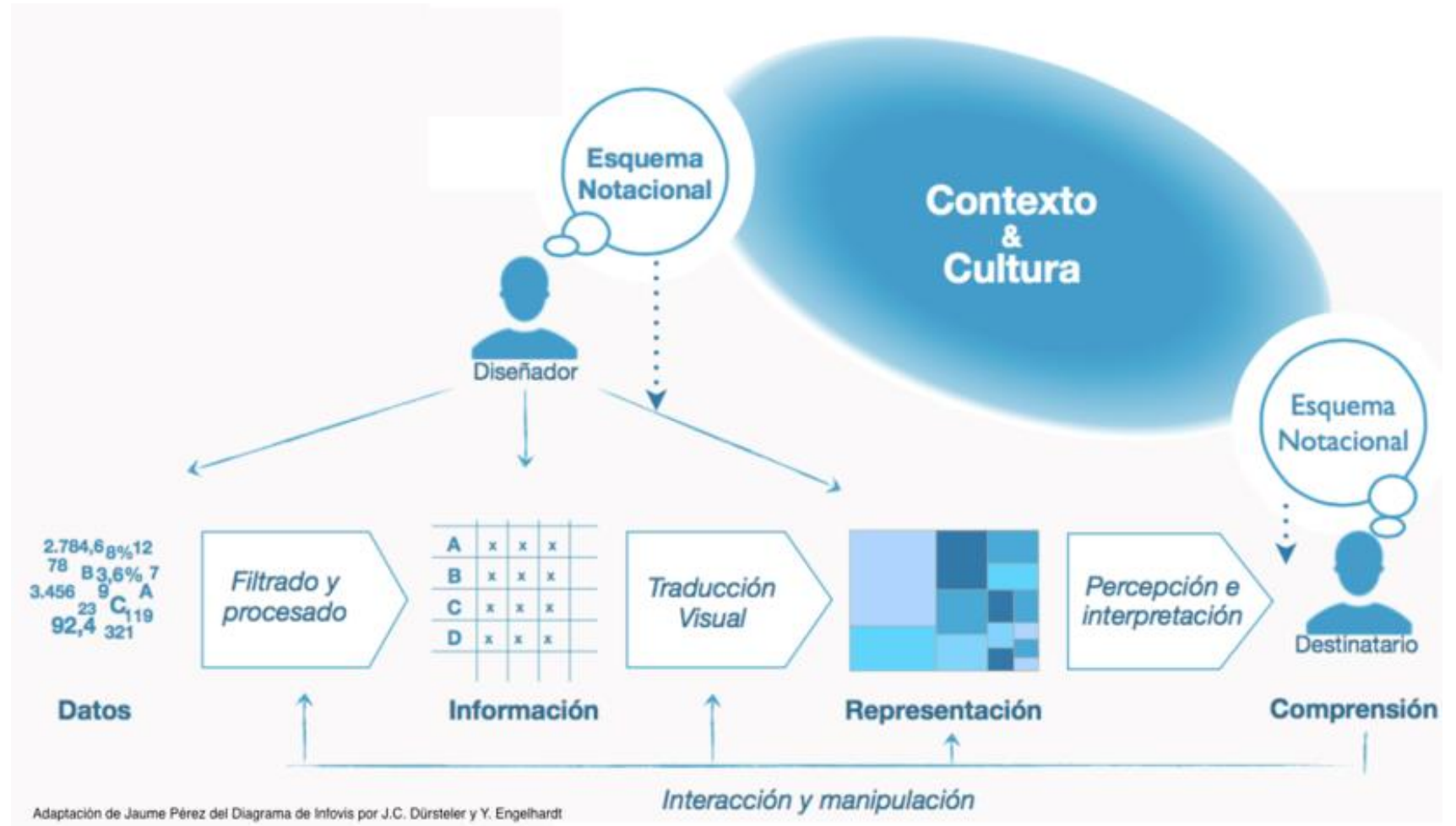
http://en.wikipedia.org/wiki/Anscombe%27s_quartet

La visualización de datos



La visualización de datos

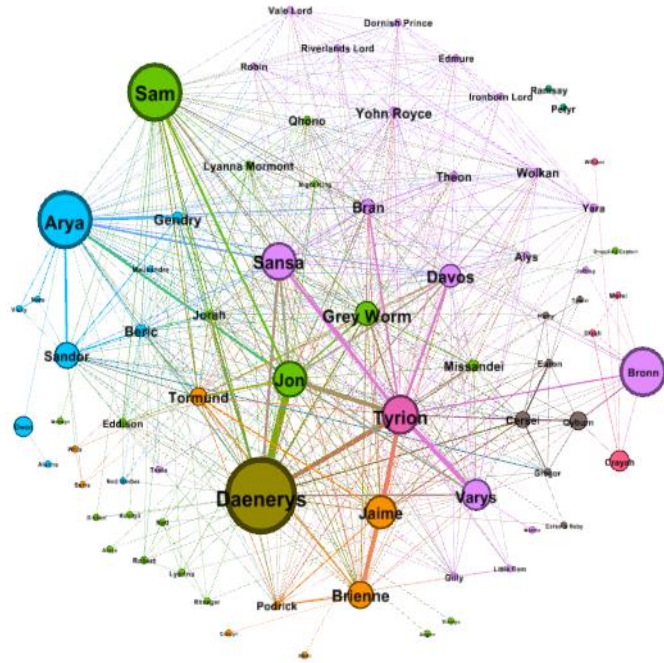
“proceso de **transformación de datos** en información, mediante elementos visuales, con dos objetivos fundamentales: **análisis y comunicación**”



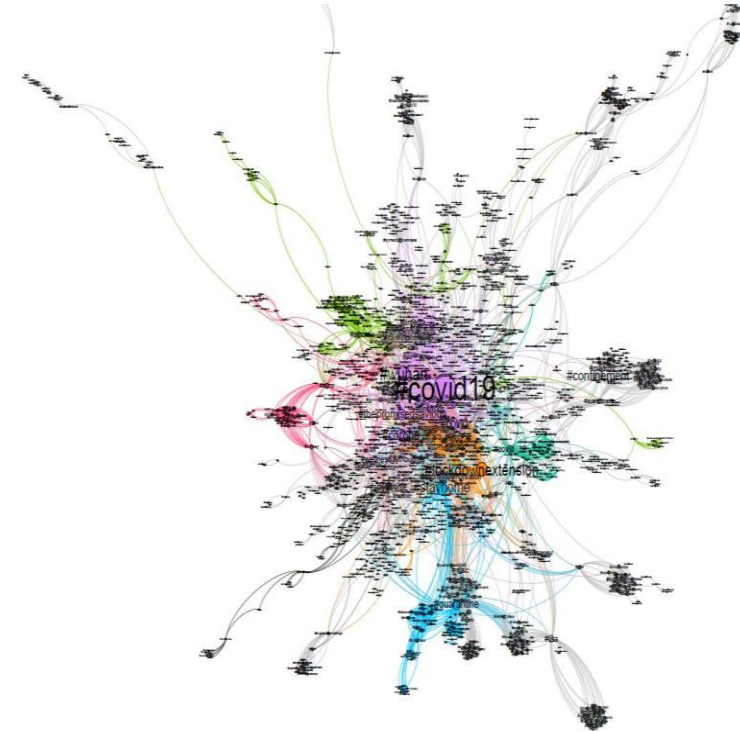
La visualización de datos

**NO visualizamos datos, sino relaciones
entre los datos**

Visualizamos relaciones entre los datos

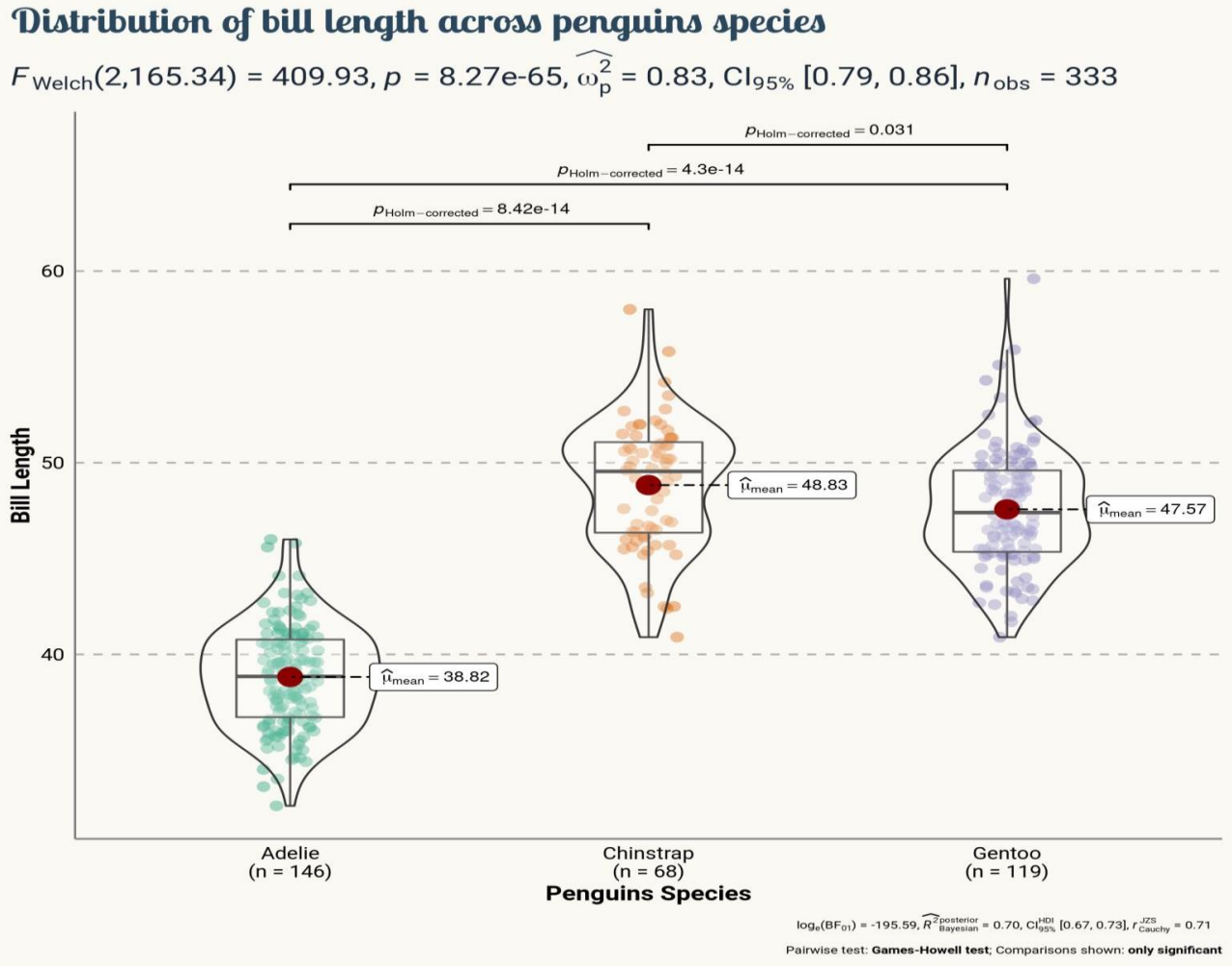


! El poder de un personaje!



! Red de comentarios en Twitter con Hashtags COVID !

Visualizamos relaciones entre los datos



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

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



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

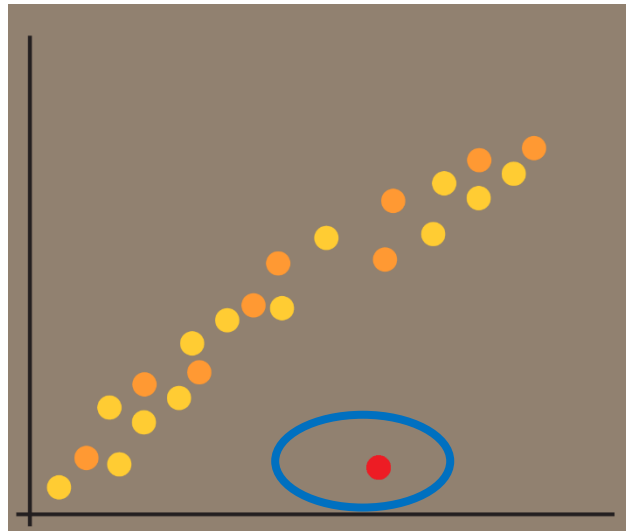
Aclaratorio o comunicar

- 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.



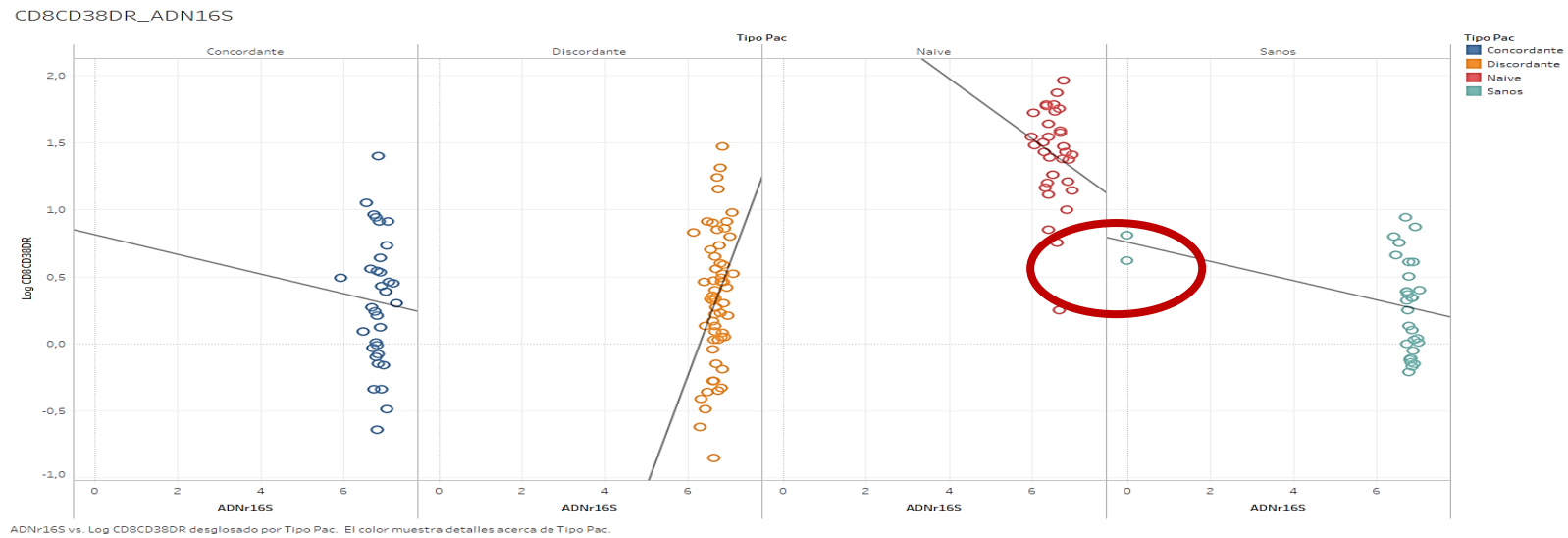
Analizar

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

Encontrar patrones

La Visualización exploratoria

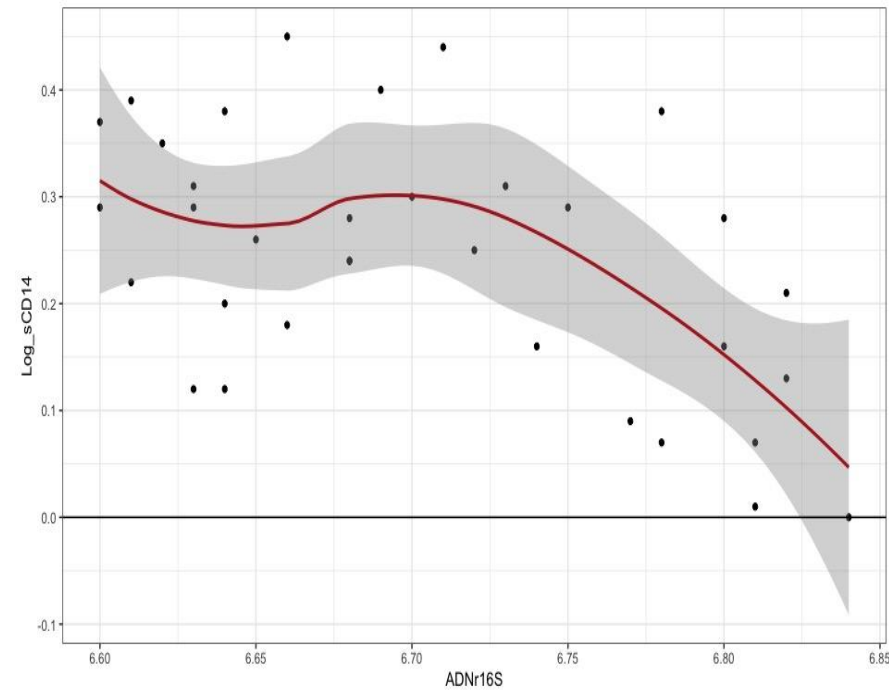
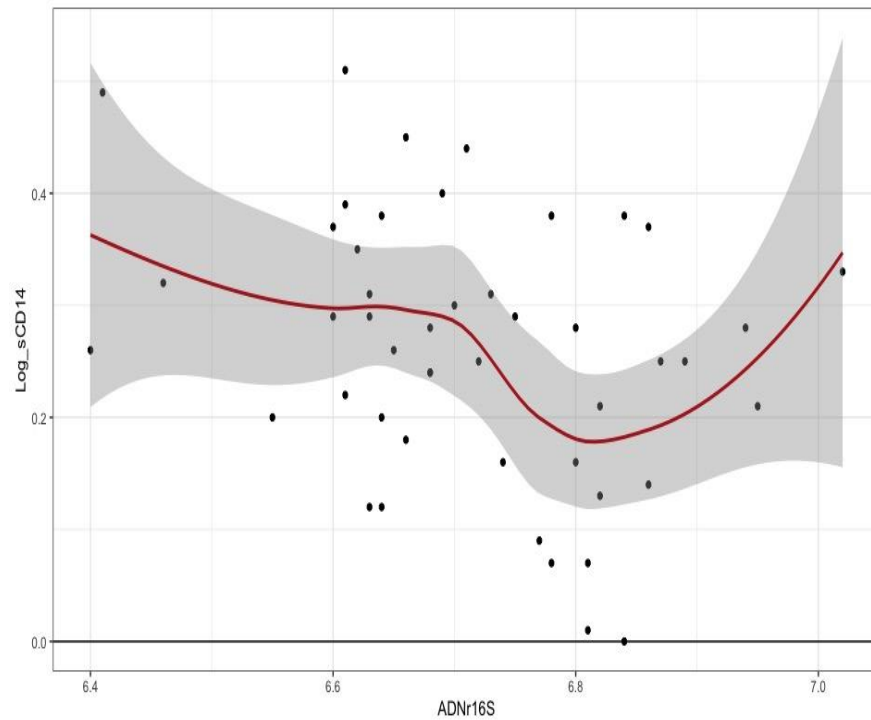
Ejemplo 1: Identificación de casos atípicos o outliers



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

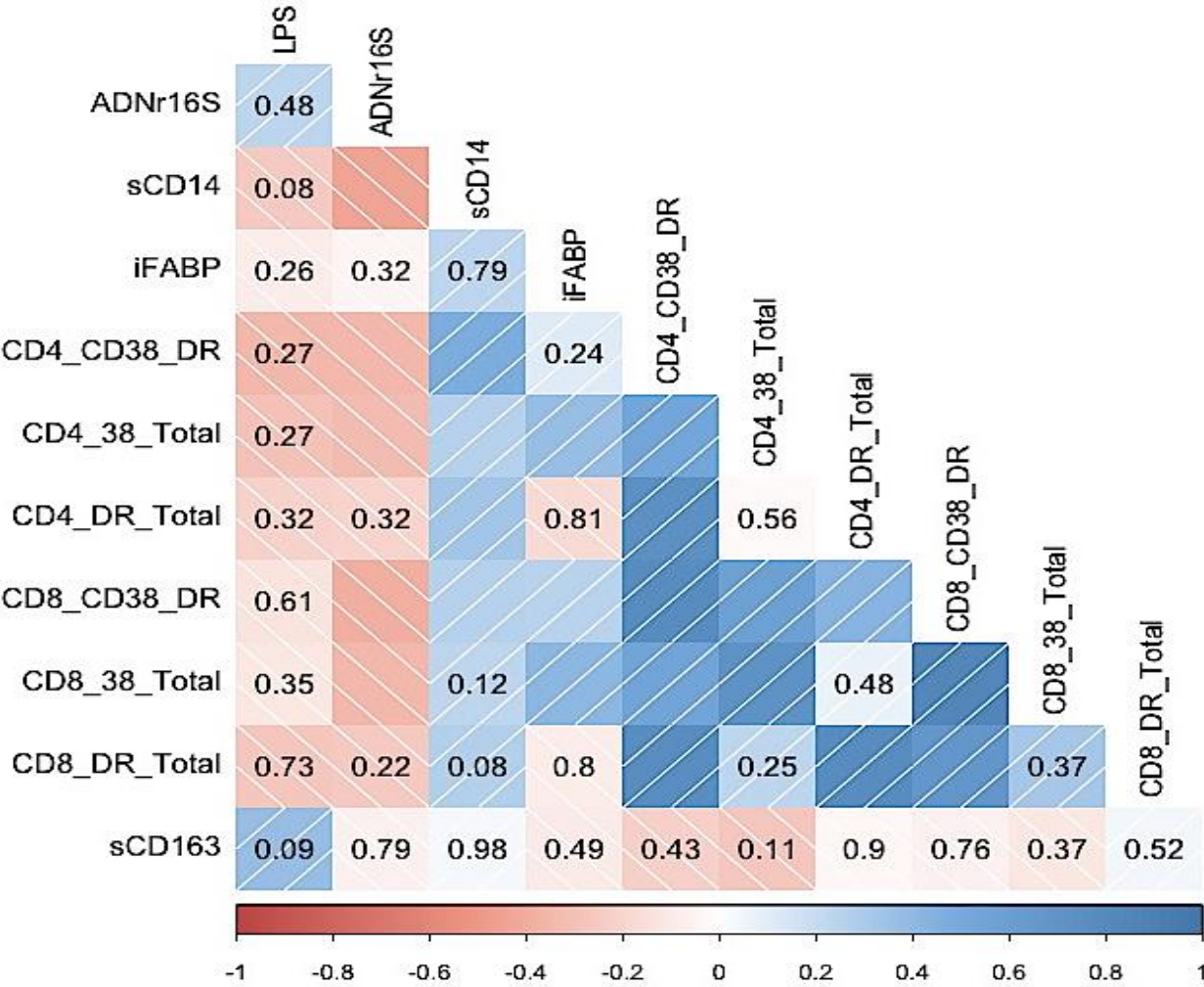
La Visualización exploratoria

Ejemplo 2: Identificación de patrones en los datos



La Visualización exploratoria

Dependencia entre variables

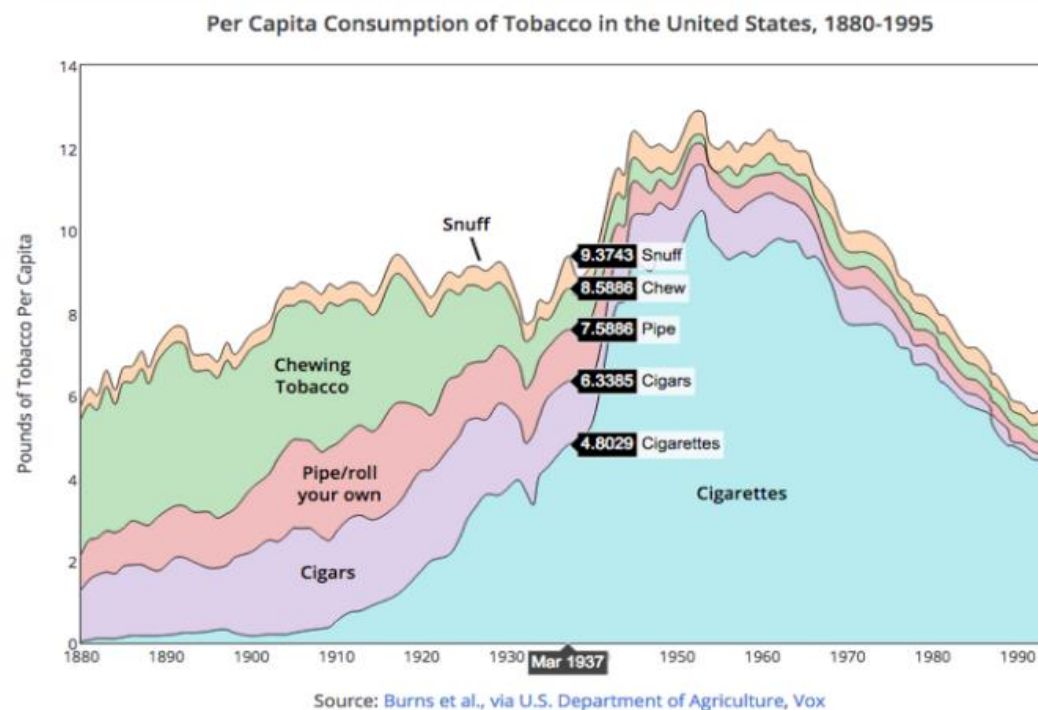


Tipos de Visualización

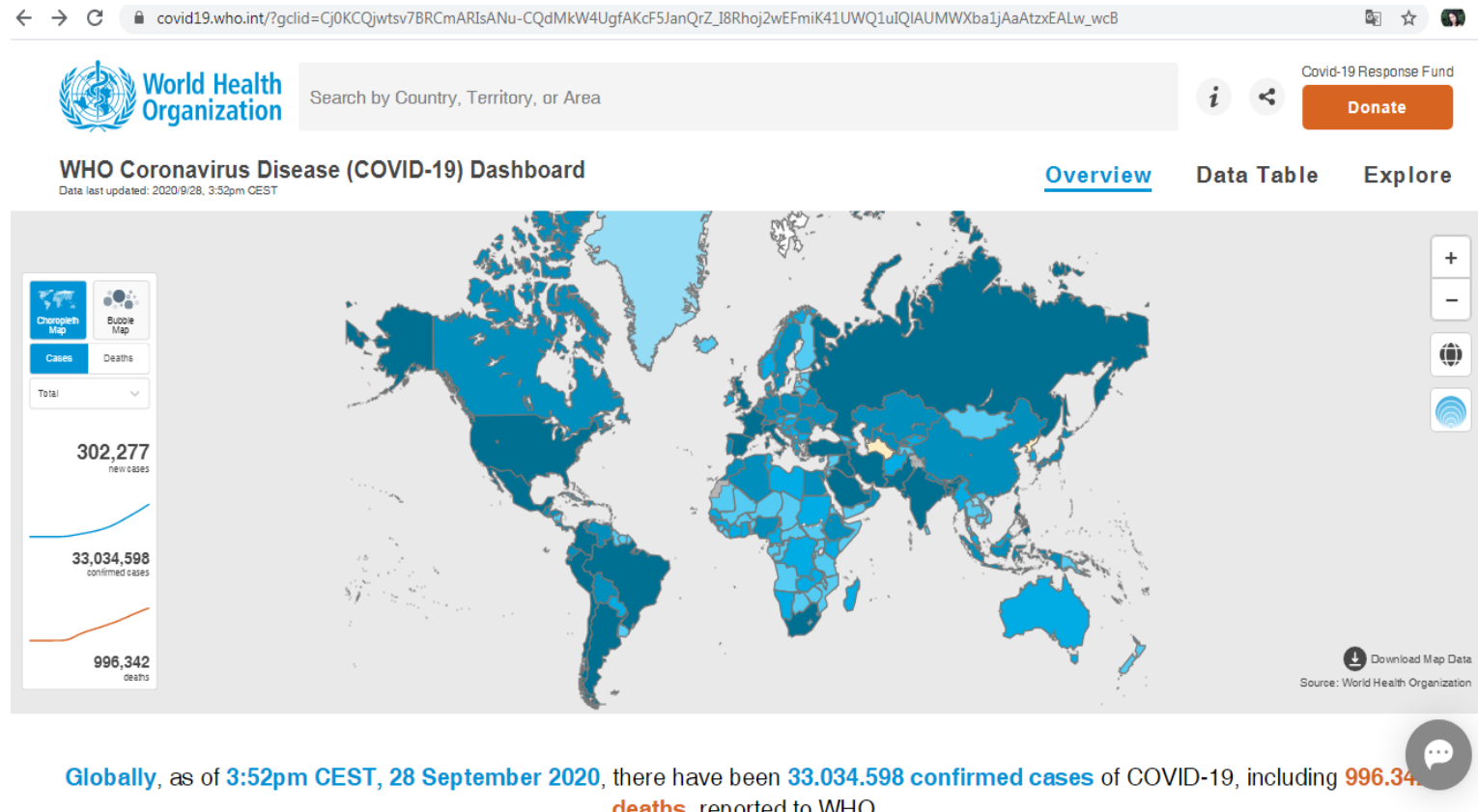
ESTÁTICAS vs DINÁMICAS

Visualizaciones estáticas

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



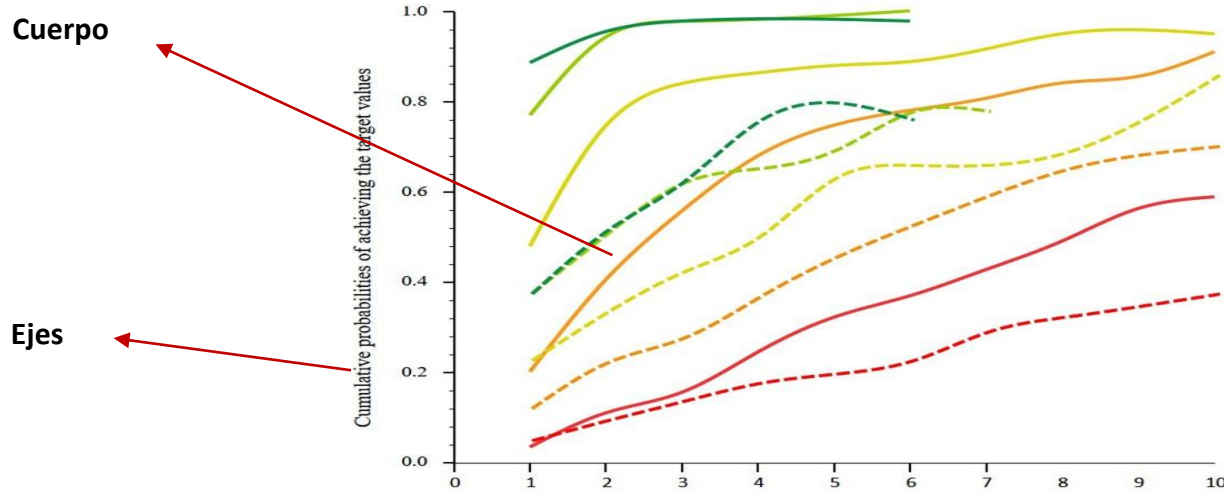
Visualizaciones dinámicas



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

Anatomía de un gráfico

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.



Título
 EL gráfico también puede incluir un subtítulo con un mensaje a destacar

Ejes

Cuerpo

Legenda

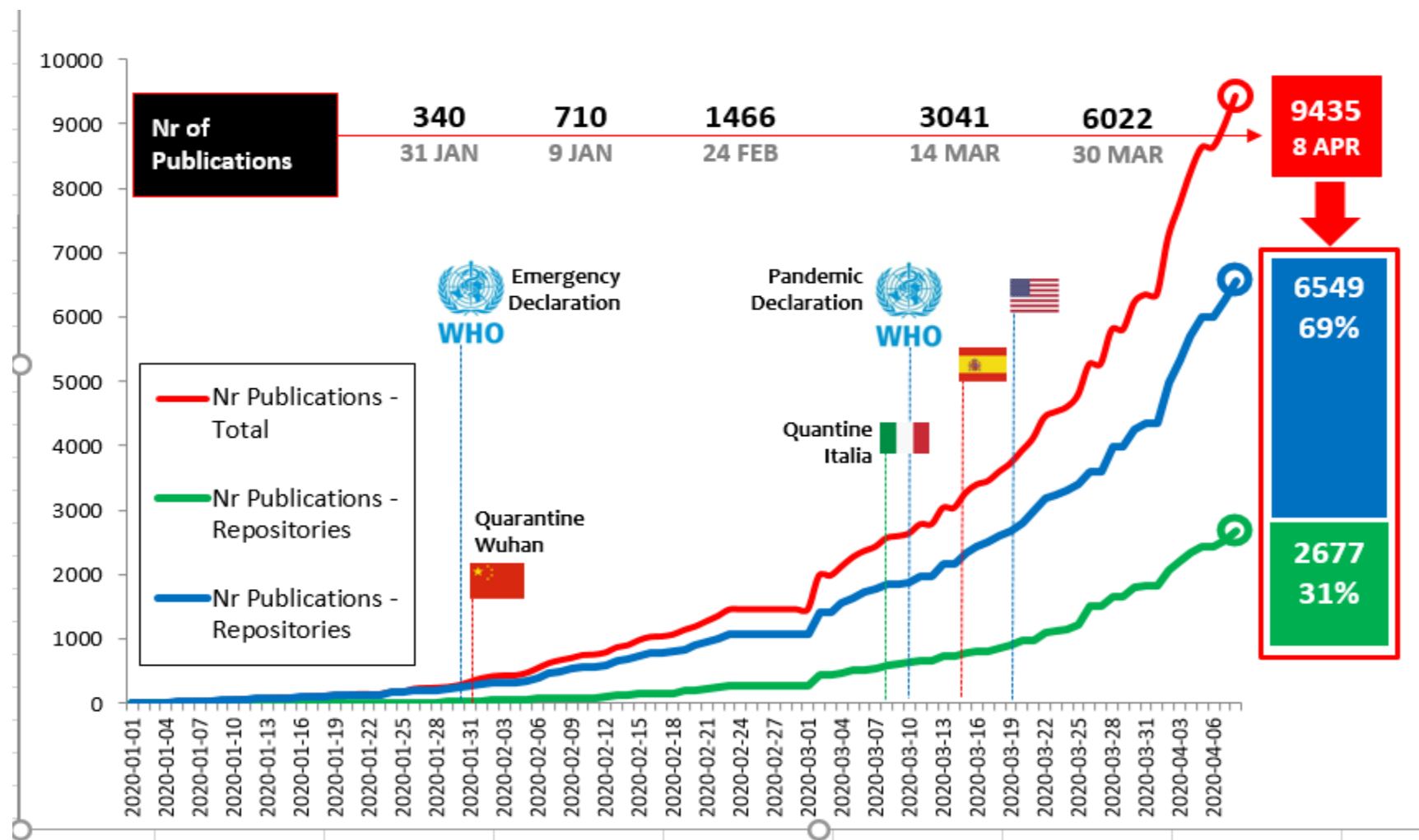
Esto pudiera ir directo asociado a las líneas. Esta forma permite agregar datos estadísticos para mejor comprensión

Basal CD4 ⁺	CD4 ⁺ ≥ 650	CD4 ⁺ /CD8 ⁺ ≥ 1	After
≤ 200/ μl	58.6%; CI ₉₅ , 51.2–66.0	36.0%; CI ₉₅ , 29.2–42.8	10 years
201–350/ μl	90.9%; CI ₉₅ , 84.9–96.8	68.0%; CI ₉₅ , 59.8–76.2	
351–500/ μl	95.6%; CI ₉₅ , 91.4–99.9	83.3%; CI ₉₅ , 66.3–100	
501–650/ μl	100%; CI ₉₅ , 100	75.5%; CI ₉₅ , 65.3–92.3	
> 650/ μl	97.7%; CI ₉₅ , 93.3–100	74.7%; CI ₉₅ , 57.5–91.9	

* the values for patients who started treatment with absolute CD4 between 500 and 650/ μl and >650/ μl 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.

Milanés, y. et al. <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6197681/>

Anatomía de un gráfico



Torres-Salinas, D. (2020). Ritmo de crecimiento diario de la producción científica sobre Covid-19. Análisis en bases de datos y repositorios en acceso abierto

Anatomía de un gráfico

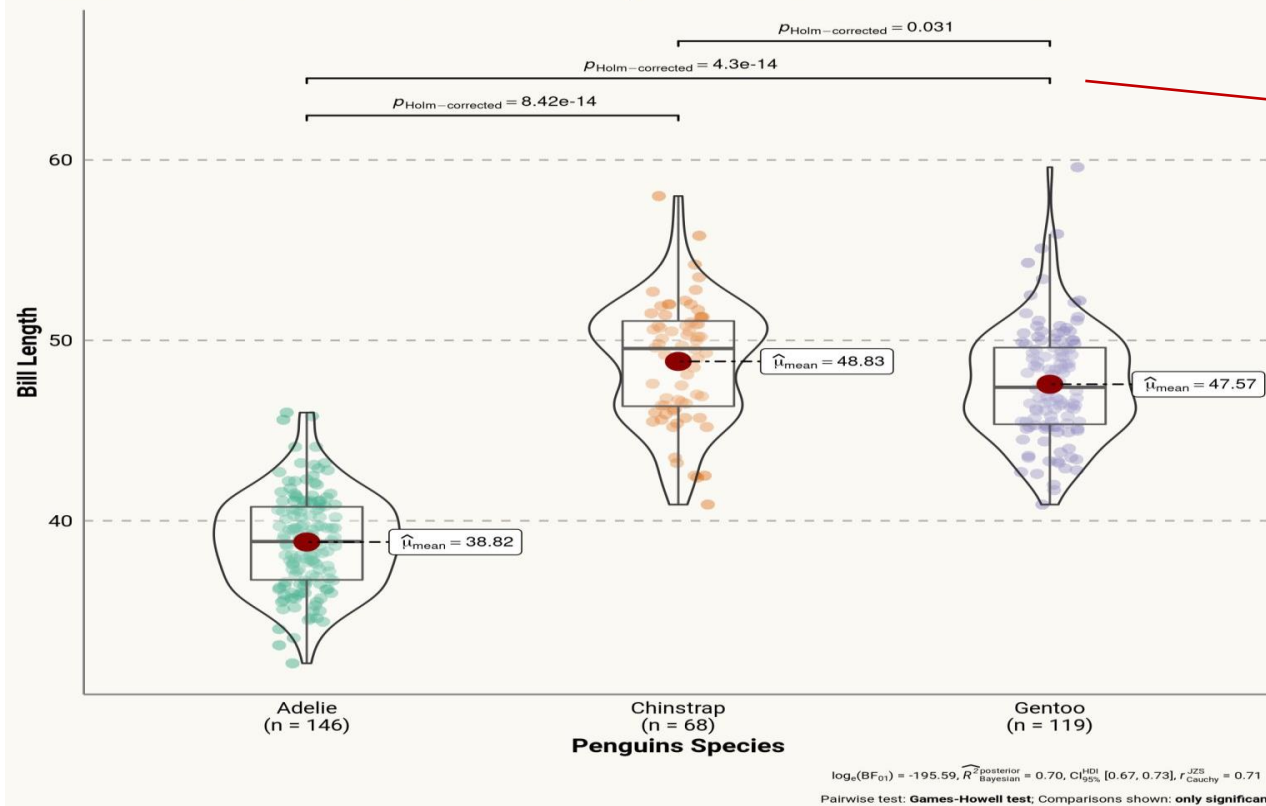
Gráfico de violines

Incluye detalles. Se transparente

- Escoge modelos visuales que te permitan mostrar la distribución real de tus datos.
- Evita caer en malas prácticas que pueden generar dudas sobre la validez de tus resultados

Distribution of bill length across penguins species

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



¿Diferencias entre los grupos?

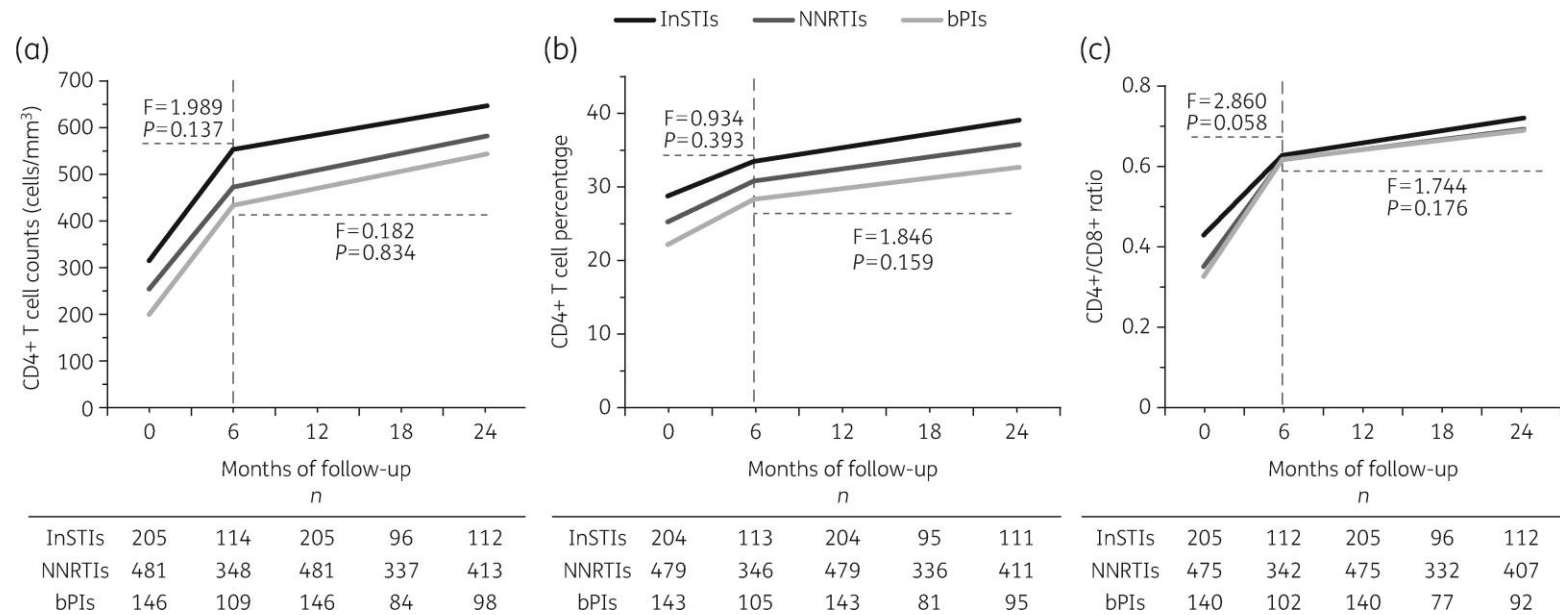
Si comparas grupos, incluye la significación estadística de haber testado si hay o no diferencias entre ellos en función de la variable que estés analizando

<https://r-graph-gallery.com/web-violinplot-with-ggstatsplot.html>

Anatomía de un gráfico

Gráficos múltiples

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



Puedes usar multigráficos para generar una única figura para tu artículo científico.

J Antimicrob Chemother, Volume 75, Issue 1, January 2020, Pages 200–207, <https://doi.org/10.1093/jac/dkz421>

The content of this slide may be subject to copyright: please see the slide notes for details.

#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.

Etapas en la Visualización de datos

Pasos...

Cuenta bien tu mensaje ...



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

Tipos de datos

Qualitative (categorical)

Arizona, New York, Texas

Sarah, John, Maria

Coors, Bud Light, Stella Artois

Qualitative (ordinal)

Gold, silver, bronze

Excellent health, good health,
poor health

Love it, like it, hate it

Quantitative

Weight (10 lbs, 20 lbs, 5000 lbs)

Cost (\$50, \$100, \$0.05)

Discount (5%, 10%, 12.8%)

Tipos de datos

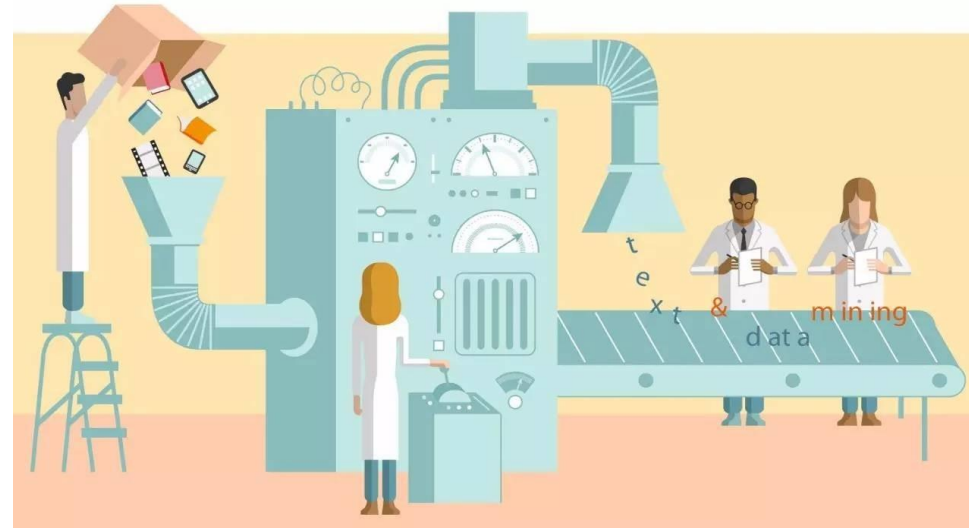
Data Types								
Numerical		Example	Categorical		Example	Time Series	Example	Text
Discrete→	Countable	Children count or count of members in family	Nominal→	Not ordered, Mutually exclusive	Sex(Male,Female)	Time series tracks the movement of data points along timestamp.	Network Logs, Tracking daily, 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
- Validéz
- 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 distinto,
Empresarios
- ¿Se enviará por email o se presencial ?
- ¿Reporte o publicación científica ?
- ¿Impresa o digital ?
- ¿Estática o Dinámica?



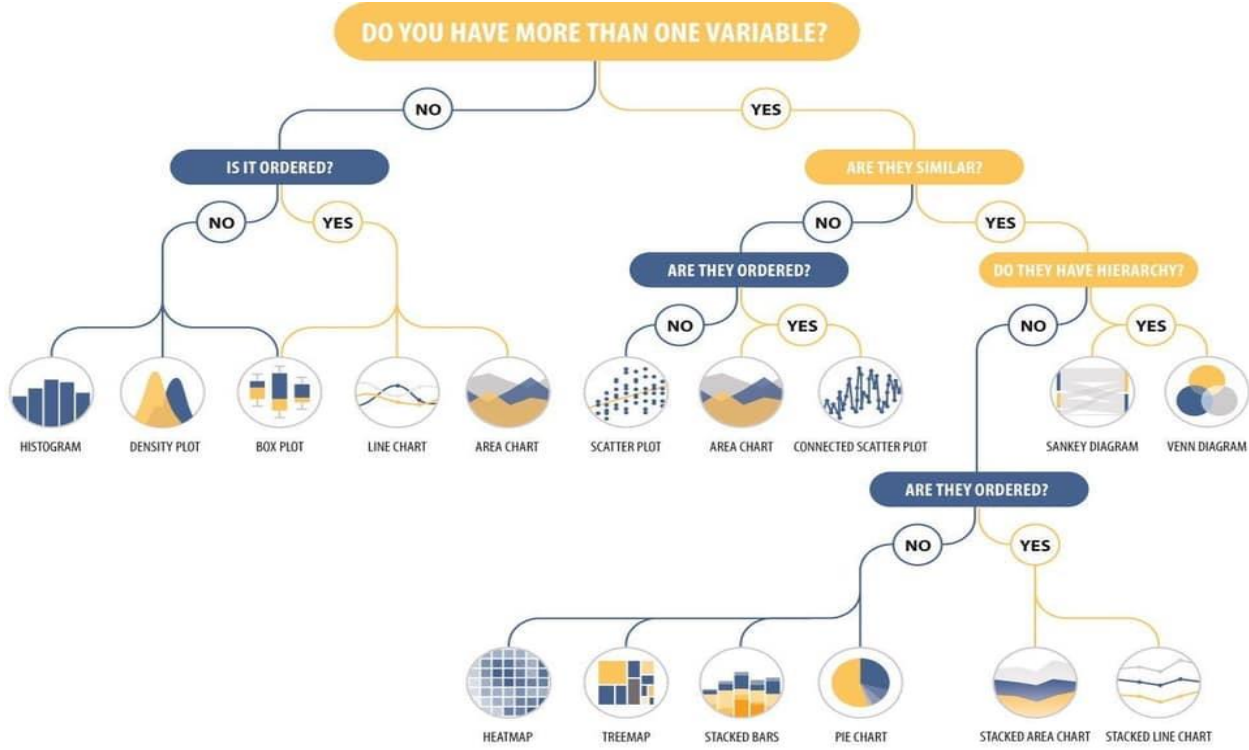
Escoge un tipo de gráfico efectivo

- ? 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 ?

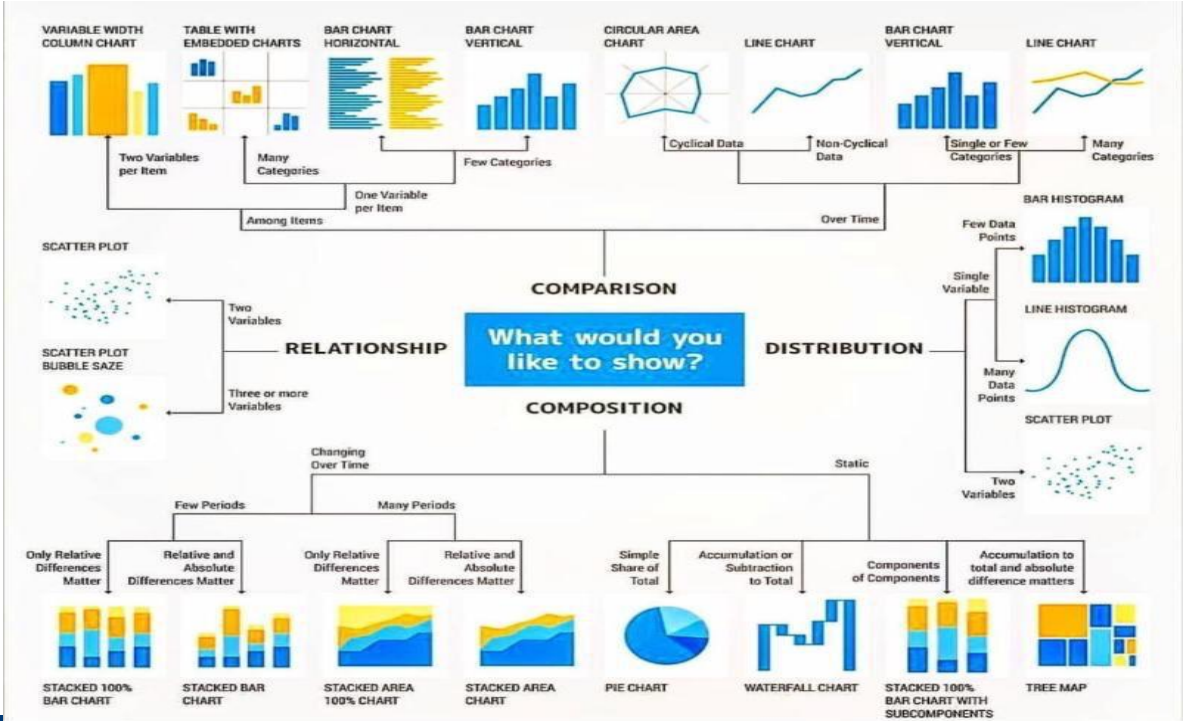
! Utiliza los diccionarios visuales online !

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

Escoge un tipo de gráfico efectivo



Created by ActiveWizards



Escoge un tipo de gráfico efectivo

Deviation
Emphasize variations (+/-) from a fixed reference point. Typically, the reference point is zero but it can also be a negative or a long-term average. Can also be used to show sentiment (growth/loss/profit/loss).

Example FT uses
Trade substitution, climate change

Diverging bar
A simple standard bar chart that can handle both negative and positive magnitude values.

Diverging stacked bar
Perfect for presenting survey results which involve sentiment (eg disagreement/agreement).

Spine
Spine is a single value into two contrasting components (eg male/female).

Supermarket filled box
The shaded area of the box represents a reference to be shown either against a bar chart or between two bar charts.

Correlation
Show the relationship between two or more variables. Be mindful that unless you tell them otherwise, many readers will assume the relationship you show them to be causal (ie one causes the other).

Example FT uses
Wages and unemployment, income and life expectancy

Scatterplot
The standard way to show the relationship between two continuous variables, one of which has its own axis.

Column + line timeline
A good way of showing the relationship between an amount (columns) and a rate (line).

Connected scatterplot
Usually used to show how the relationship between 2 variables has changed over time.

Bubble
Like a scatterplot, but by listing the values on both axes, you can compare multiple variables.

XY heatmap
Good for showing two patterns between 2 variables. The color indicates the effect of showing the relationship.

Dot plot
Like a scatter plot, but the dots are arranged in order on a single axis to represent multiple categories.

Slope
Perfect for showing how variables change over time or vary between categories.

Lollipop
Lollipop charts represent the data value that is most important and can also show rank and value effectively.

Bar chart
Effective for showing changing data trends. For larger datasets, consider showing lines using colour.

Ranking
Use where an item's position in an ordered list is more important than the absolute or relative value. Sort the absolute or relative value. Sort the distribution on a meaningful way of highlighting the lack of uniformity or equality in the data.

Example FT uses
Wages, unemployment, budget deficits, environmental election results

Ordered bar
Standard bar charts to display the rank of values. Small bars are usually ordered.

Ordered column
See above.

Ordered proportional symbol
Use when there are big variations between values. The difference between values is not so important.

Dot strip plot
Like dot strip plots, good for showing the data in a table, they are best used when comparing multiple categories.

Slope
Perfect for showing how variables change over time or vary between categories.

Lollipop
Lollipop charts represent the data value that is most important and can also show rank and value effectively.

Bar chart
Effective for showing changing data trends. For larger datasets, consider showing lines using colour.

Distribution
Show values in a dataset and how often they occur. The shape of the distribution can be a meaningful way of highlighting the lack of uniformity or equality in the data.

Example FT uses
Income distribution, population growth, revenue

Histogram
The standard way to show a distribution. The shape of the distribution can be a meaningful way of highlighting the lack of uniformity or equality in the data.

Dot plot
A simple way of showing the change or range (minimum) of data across multiple categories.

Dot strip plot
Good for showing individual values in a distribution. Can be a good way of showing when two more data have the same value.

Bar chart
Good for showing the distribution of data across multiple categories.

Population pyramid
A standard way for showing the age and sex distribution of a population. The bars extend back to back.

Cumulative curve
A good way of showing how a variable changes over time. The x-axis is time and the y-axis is the cumulative value of the variable.

Frequency polygons
For displaying multiple distributions of data. Like a regular bar chart, but instead of a histogram, it uses lines to connect the tops of the bars.

Beeswarm
Use to emphasize individual points in a distribution. Points can be colored by an additional variable. Best with medium-sized datasets.

Change over Time
Give emphasis to changing trends. These can be identified by the slope of the line. Highlighting trends or contrasts. Choosing the correct line pattern is important to provide suitable context for the reader.

Example FT uses
Share price movements, seasonal time series, sectoral changes in a market

Line
The standard way to show a changing time series. Use color and line style to distinguish between different data points.

Column
Columns work well for showing change over time, but usually best with only one series of data at a time.

Column + line timeline
Good for showing individual values in a distribution. Can be a good way of showing when two more data have the same value.

Strip
Good for showing the distribution of data across multiple categories.

Area chart
Use with care - these are good at showing change over time, but can be very difficult to read.

Candlestick
Should be reserved for showing the range and high/low points of each day.

Fan chart (spineless)
Use to show the distribution of data across multiple categories. Usually, this is used to represent the proportion of each category.

Connected scatterplot
A good way of showing changing data trends. For larger datasets, consider showing lines using colour.

Calendar heatmap
A good way of showing data over time. The x-axis is time and the y-axis is the value of the variable.

Heatmap
Good for showing the distribution of data across multiple categories. The color indicates the value of the variable.

Vertical timeline
Perfect for showing the distribution of data across multiple categories. The x-axis is time and the y-axis is the value of the variable.

Sankey diagram
Another alternative to the flow diagram for showing flows between categories in the data.

Streamgraph
A type of area chart, use when showing multiple data series over time. It is more important than individual values.

Magnitude
Show size comparisons. These can be relative (eg being able to see the difference) or absolute (eg being able to see the difference). Usually, these are used to show the relative size of the components. Consider a 'counted' number (for example, barrels, dollars or pesos) rather than a calculated rate or per cent.

Example FT uses
Community production, market capitalization, volumes in general

Column
A simple way to compare the size of components. Can be used to show that one is larger than another.

Bar
Use above. Good when you want to compare the size of components. Can be used to show that one is larger than another.

Paired column
An easy way of comparing two values. The bars are placed side-by-side.

Paired bar
See above.

Horizontal bar
A good way of showing the size and proportion of data at the same level. Can be used to show that one is larger than another.

Proportional symbol
Use when there are big variations between values. The difference between values is not so important.

Isotype (dots/pipes)
Excellent solution in some situations. Use only with small numbers. Do not use if an arrow is required.

Lollipop
Lollipop charts draw more attention to the data value than the category. They are best used to show a clear pattern of progress.

Radar
A space efficient way of showing multiple variables. But make sure they are all important. It can be hard to read.

Parallel coordinates
An alternative to radar charts. Again, the elements of the chart are usually benefits from highlighting values.

Bar
Good for showing a measurement against the context of a target or performance range.

Grouped symbol
An alternative to bar charts. Use when you want to highlight individual elements in a chart.

Part-to-whole
Show how a single entity can be broken down into its component elements. The reader's mental model of the whole is reinforced by the size of the components. Consider a 'counted' number (for example, barrels, dollars or pesos) rather than a calculated rate or per cent.

Example FT uses
Fossil budgets, company structures, national election results

Stacked column/bar
The standard way of showing part-to-whole data. The bars are stacked on top of each other. Can be used to show that one is larger than another.

Horizontal bar
A good way of showing the size and proportion of data at the same level. Can be used to show that one is larger than another.

Pie
A common way of showing part-to-whole data - but be aware that it is difficult to accurately compare the size of the segments.

Donut
Similar to a pie chart - but the center is empty. Can be used to show that one is larger than another.

Tree map
Use for hierarchical data. The size of the rectangles represents the value of the data. Can be difficult to read when there are many small segments.

Voronoi
A way of showing the relationship between two variables. The lines represent the boundaries between the two variables.

Arc
A heuristic, often used for illustrating proportions. Can be used to show that one is larger than another.

Gridmap
Good for showing % information. They are best used to show a clear pattern of progress.

Venn
Generally only used for technical representation.

Waterfall
Can be useful for showing part-to-whole data. The bars are stacked on top of each other.

Spatial
Aside from locator maps only used when precise location is important for the reader. These might be logical or statistical. These might be logical or statistical. These might be logical or statistical.

Example FT uses
Regional density, natural resource locations, natural disaster risk/impact, settlement areas, variables in election results

Basic choropleth (heatmaps)
The standard approach for putting data on a map. The color of the map represents the value of the data. Can be used to show that one is larger than another.

Proportional symbol (point/size/shape)
Use for areas where it is difficult to accurately compare the size of the segments.

Flow map
For showing movement between areas. Can be used to show that one is larger than another.

Contour map
For showing areas of equal value. Can be used to show that one is larger than another.

Equalized cartogram
Cartograms that use a map to represent data. The size of the rectangles represents the value of the data. Can be used to show that one is larger than another.

Scaled cartogram (Choropleth)
Cartograms that use a map to represent data. The size of the rectangles represents the value of the data. Can be used to show that one is larger than another.

Dot density
Used to show the location of individual elements. Can be used to show that one is larger than another.

Heat map
Grid-based data values mapped with an underlying color scale. As mentioned, they are best used to show a clear pattern of progress.

Flow
Show the reader volumes or intensity of movement between two or more points or conditions. These might be logical or statistical. These might be logical or statistical. These might be logical or statistical.

Example FT uses
Movement of funds, trade, migration, tourism, information relationship graphs

Sankey
Shows changes in flows from one condition to at least one other. Good for showing the flow of information.

Waterfall
Designed to show the sequencing of data through a flow process. Usually, lollipop charts, can include +/- components.

Chord
A complex but powerful diagram which can show the relationships between two or more categories.

Network
Used for showing the relationships between two or more categories. Can be used to show that one is larger than another.

Visual vocabulary

Designing with data

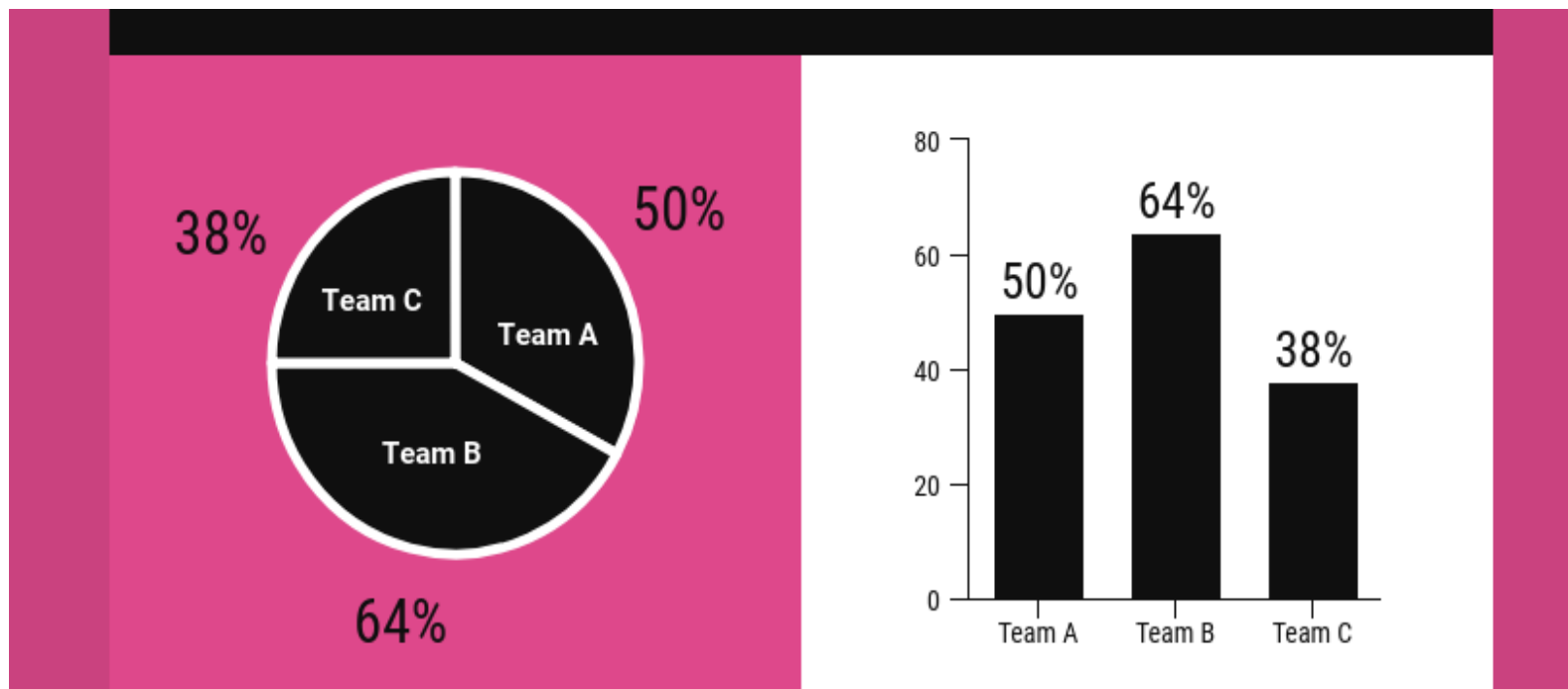
There are so many ways to visualise data - how do we know which one to pick? Use the categories across the top to decide which data relationship is most important in your story, then look at the different types of chart within the category to form some initial ideas about what might work best. This list 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

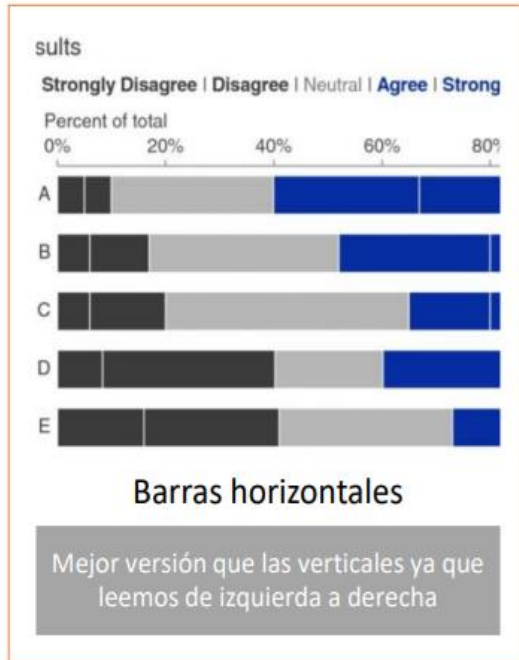
Visual Vocabulary of Financial times

Escoge un tipo de gráfico efectivo

¿Cuál es el correcto?



Tipos de gráficos para comparar datos



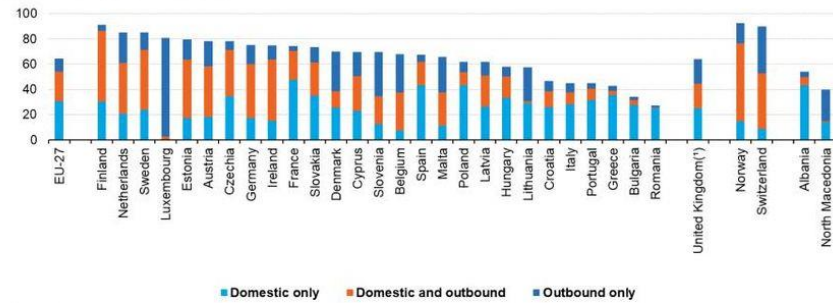
Storytelling with data: a data visualization guides por business professionals

Comparar datos

Buenas prácticas

BEFORE

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

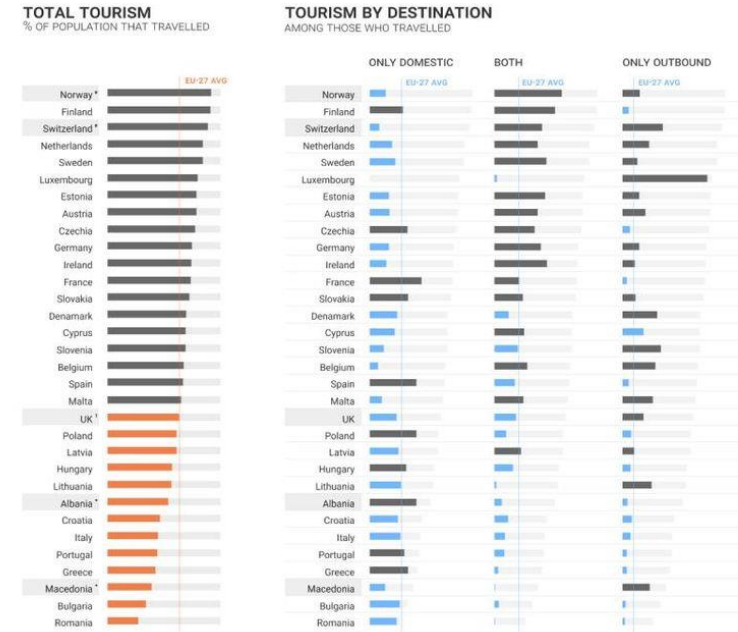


(*) 2016 data.
Source: Eurostat (online data code: tour_dem_totot)



AFTER

% EU population 15+ participating in tourism in 2018



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

Comparar datos

Buenas prácticas

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



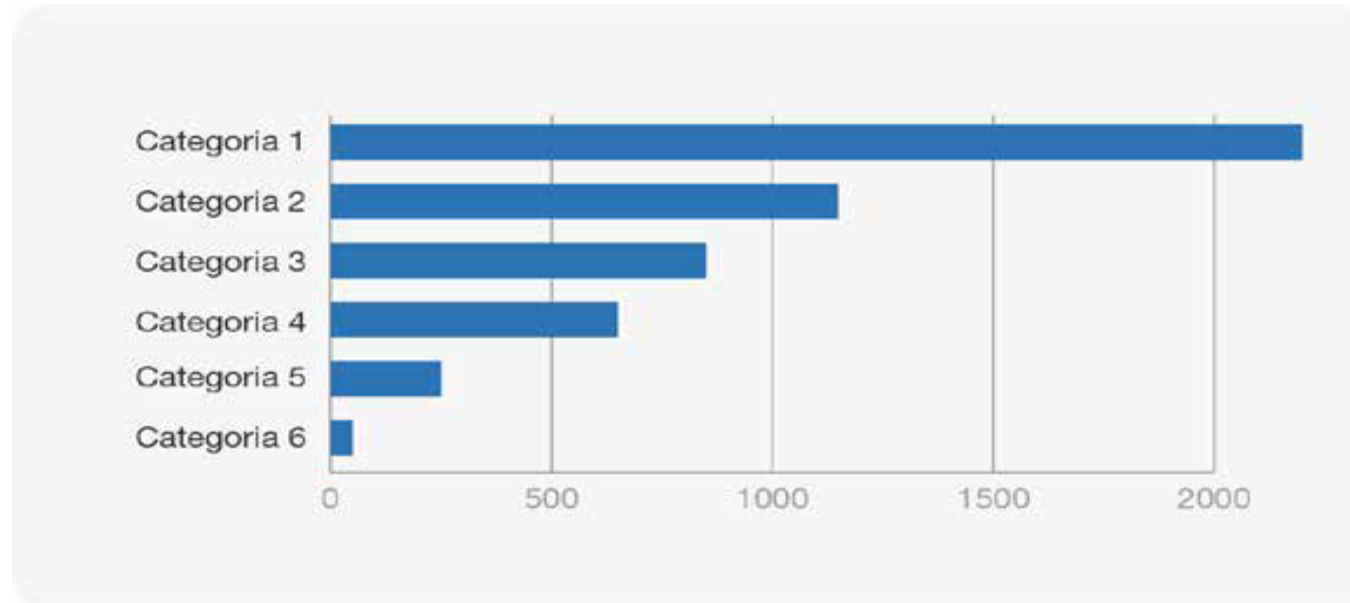
<https://venngage.com/blog/misleading-graphs/>

Comparar datos

Buenas prácticas



¡Si los nombres son largos, conviene orientarlo horizontalmente!

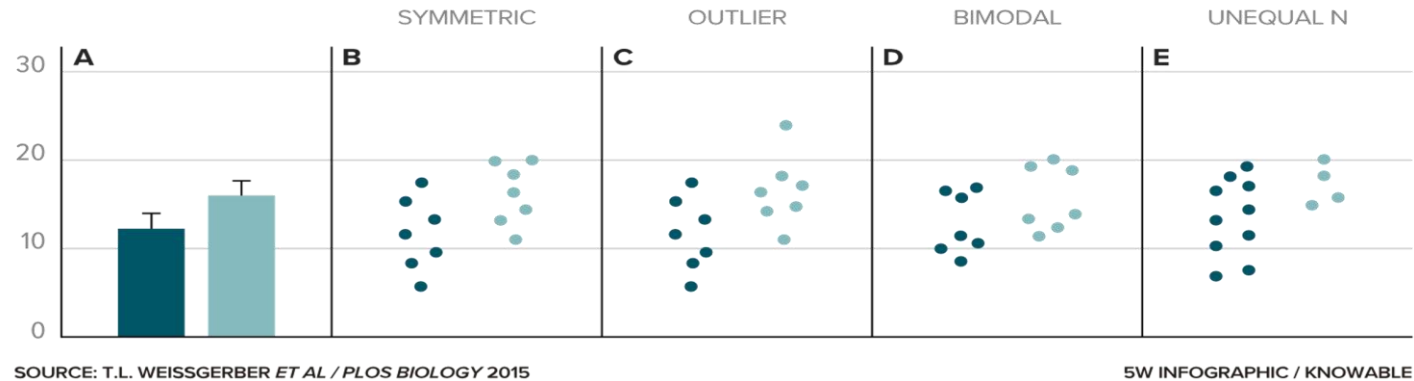


http://atenciociudadana.gencat.cat/web/.content/manuals/visualitzacio_dades/guia_visualitzacio_es.pdf

Buenas prácticas

Hidden in the bars

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



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.**

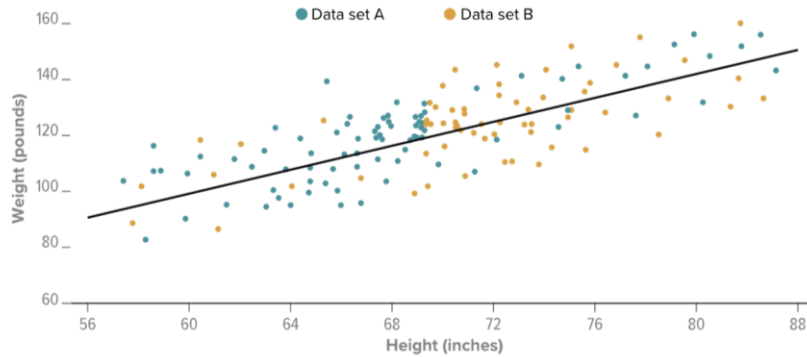
<https://knowablemagazine.org/article/mind/2019/science-data-visualization>

Comparar datos

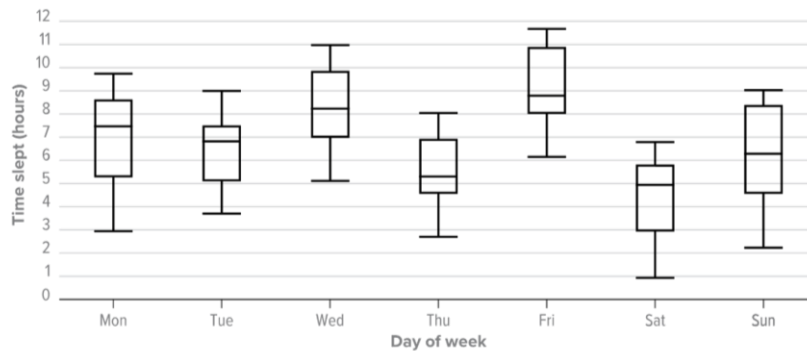
Buenas prácticas

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

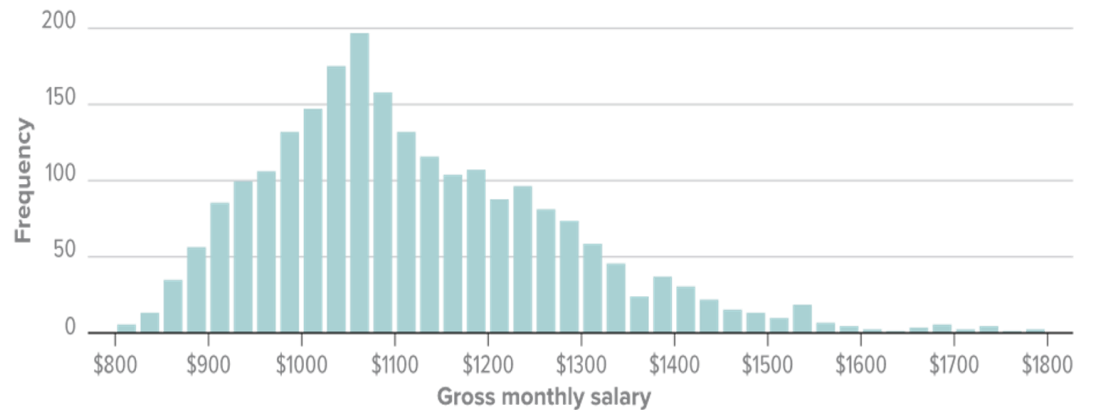
SCATTERPLOT



BOX PLOT



HISTOGRAM



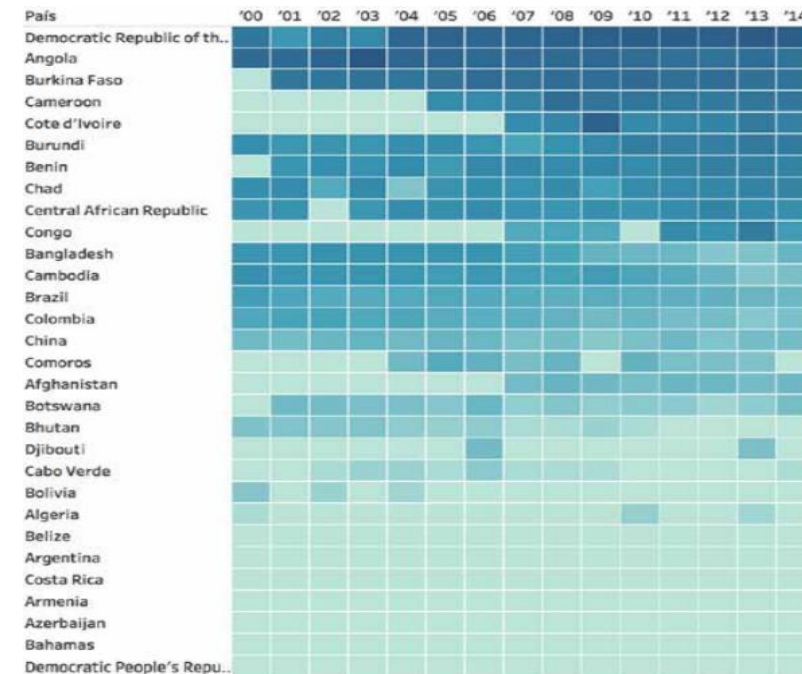
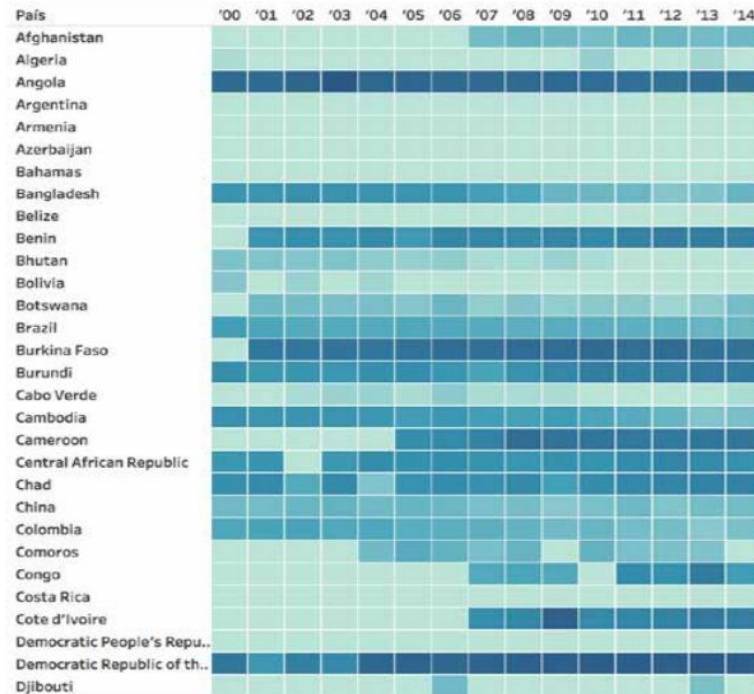
5W INFOGRAPHIC / KNOWABLE

Comparar datos

Buenas prácticas

Gráfico de intensidad de colores (Heatmap)

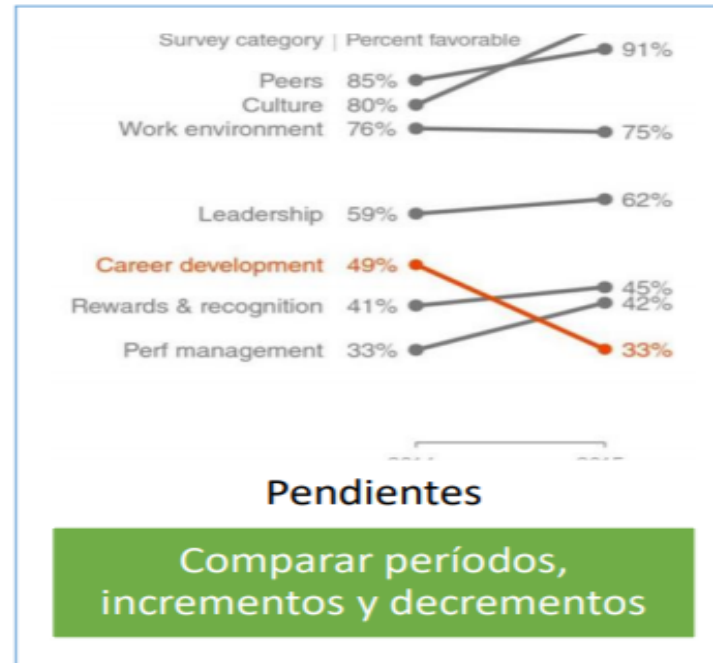
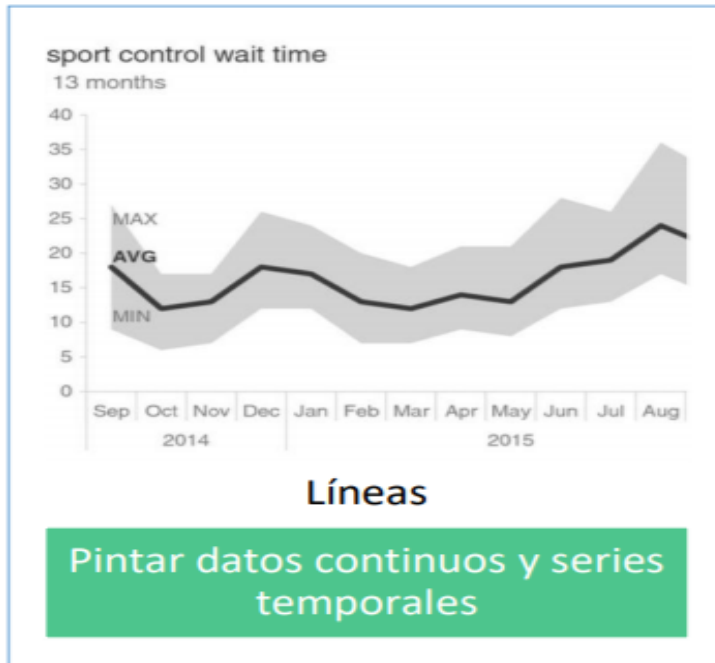
- Ordenar las filas y las columnas del gráfico de intensidades por colores según un criterio establecido
- Variables tienen escalas diferentes, **es aconsejable normalizarlas, de modo que todas pasen a tener el mismo rango de valores.**



Tipos de gráficos / Tendencias

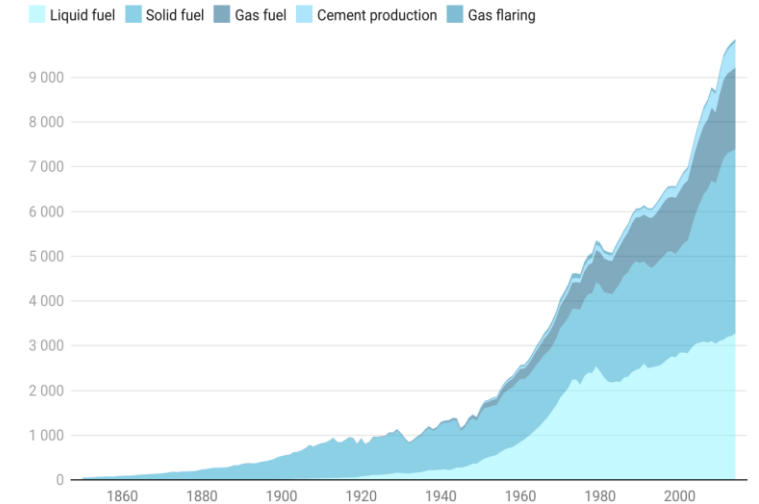
Buenas prácticas

- En lugar de barras, escoge líneas, pendientes o áreas para datos temporales.



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

Global CO2 emissions from fossil-fuel burning, cement manufacture, and gas flaring, in million metric tonnes, 1850-2014



Fuente: Carbon Dioxide Information Analysis Center • Creado con Datawrapper

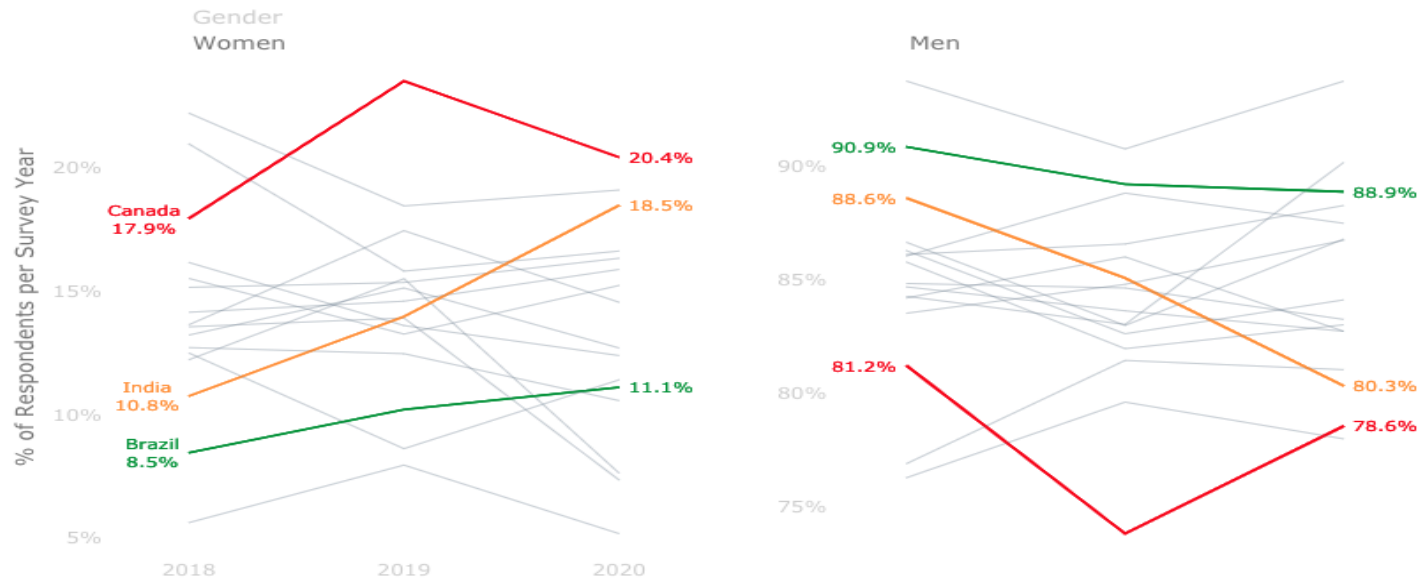
Storytelling with data: a data visualization guides por business professionals

Tipos de gráficos / Tendencias

Buenas prácticas

- 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.

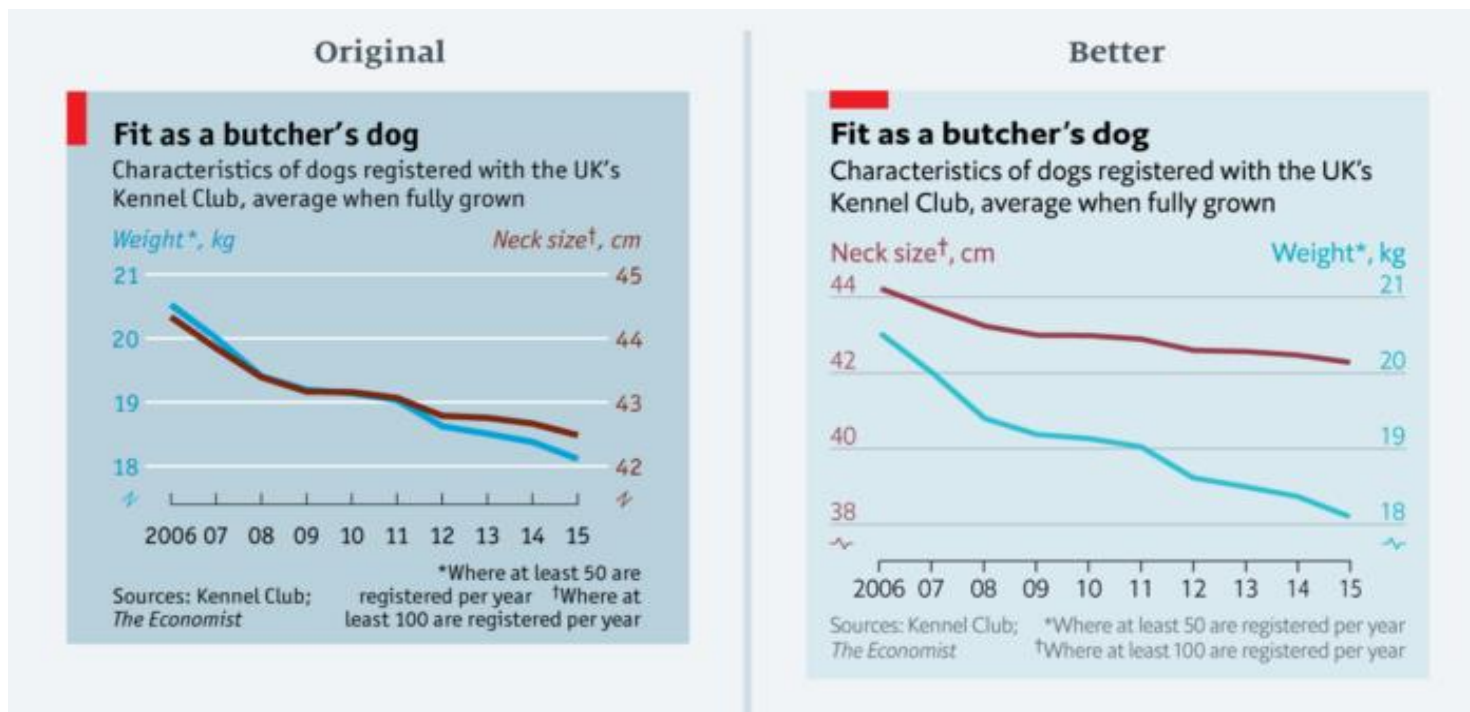
Tipos de gráficos / Tendencias

Buenas 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>

Tipos de gráficos / Mapas

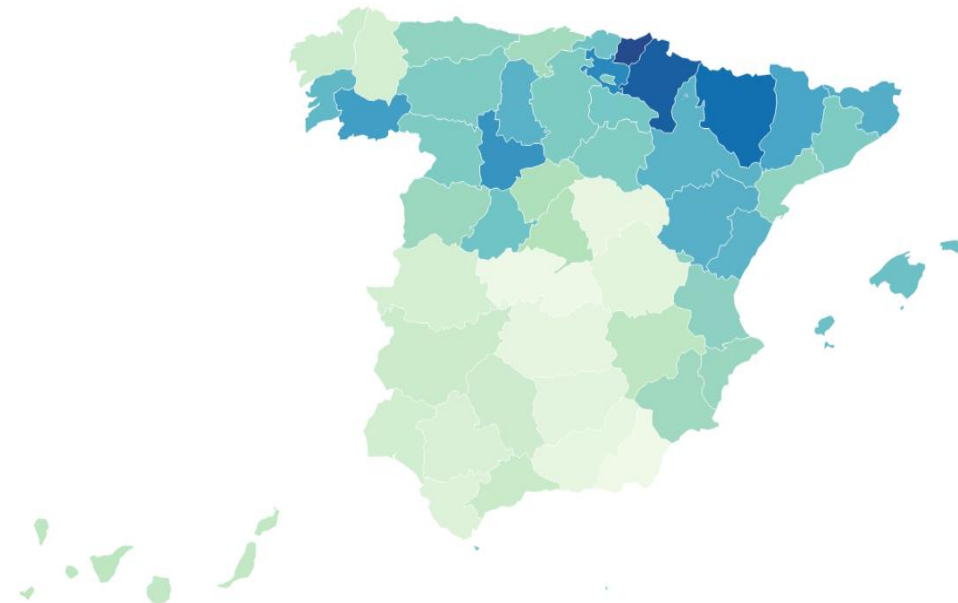
Mapa de coropletas (mapas rellenos)

- 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

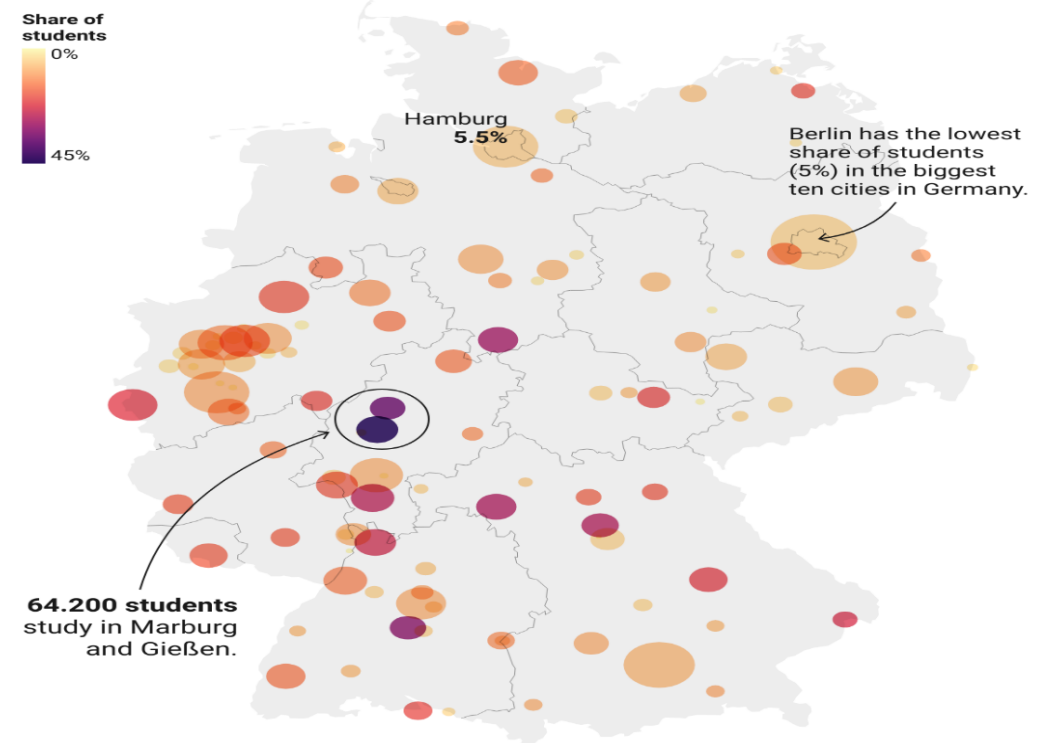
Tipos de gráficos / Mapas

Mapa, de símbolos proporcionales

- 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.**



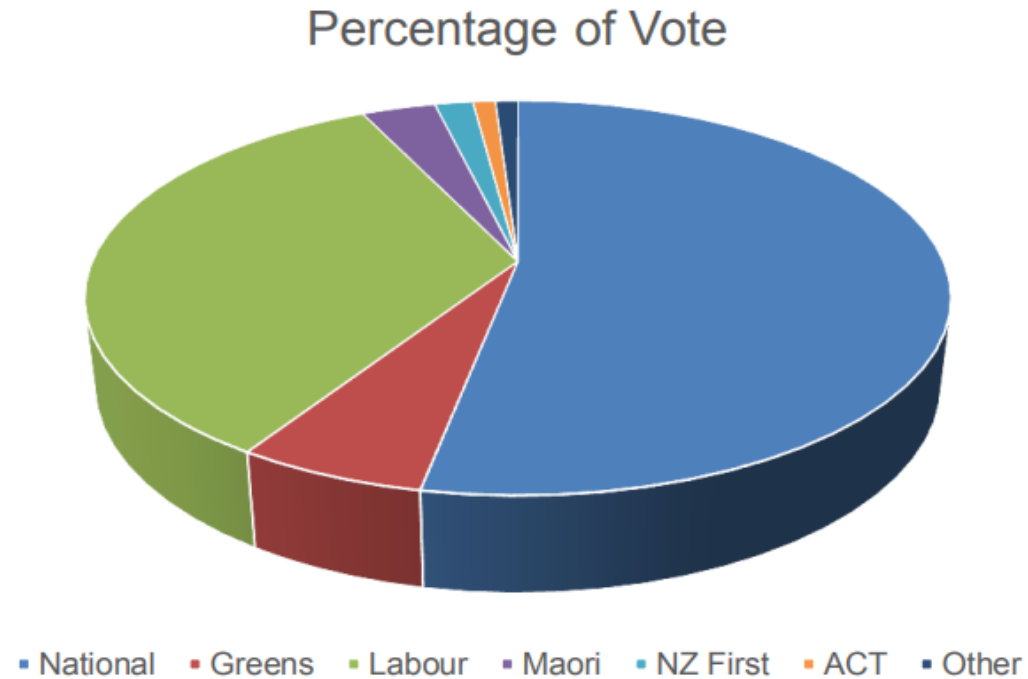
Mapa: Lisa Charlotte Rost, Datawrapper • Fuente: Eurostat 2016 • Creado con Datawrapper

- Como vas a utilizar varias símbolos y además colores, se recomienda utilizar la transparencia para facilitar la lectura del gráfico e identificar los símbolos que quedan

Tipos de gráficos / Partes de un todo

Pie Charts – NO!

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



EL ojo humano es malo leyendo ángulos.

Tipos de gráficos / Partes de un todo

Buenas prácticas

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

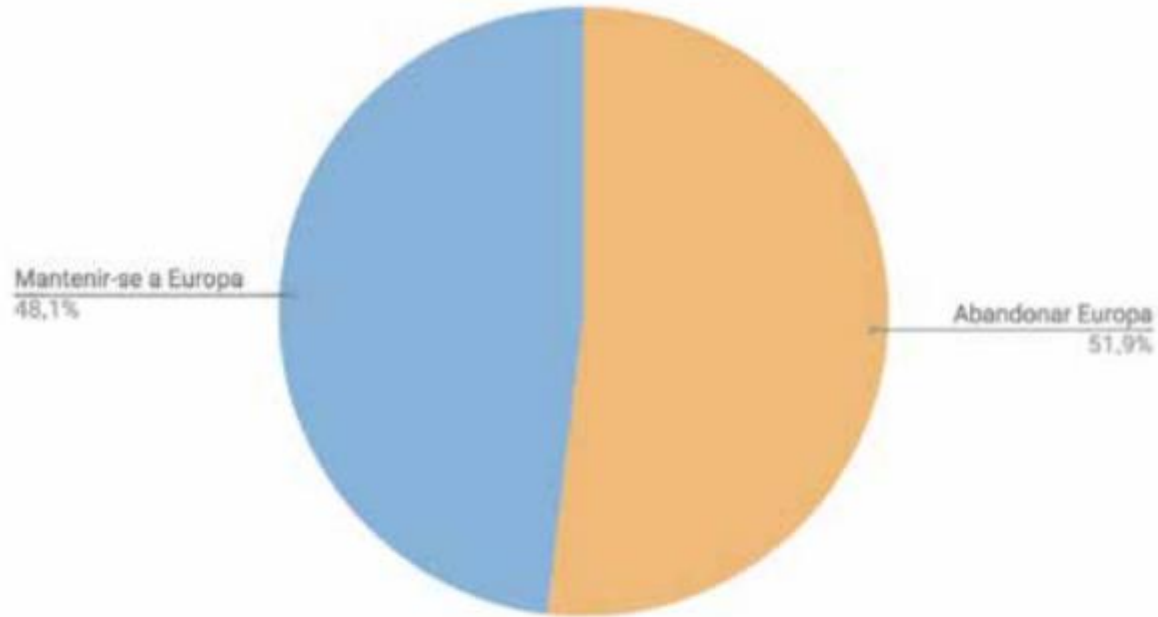


Figura 4.27. Resultado de la votación del Brexit en el Reino Unido.

Tipos de gráficos / Partes de un todo

Buenas prácticas

Cuando son datos categóricos ordinales



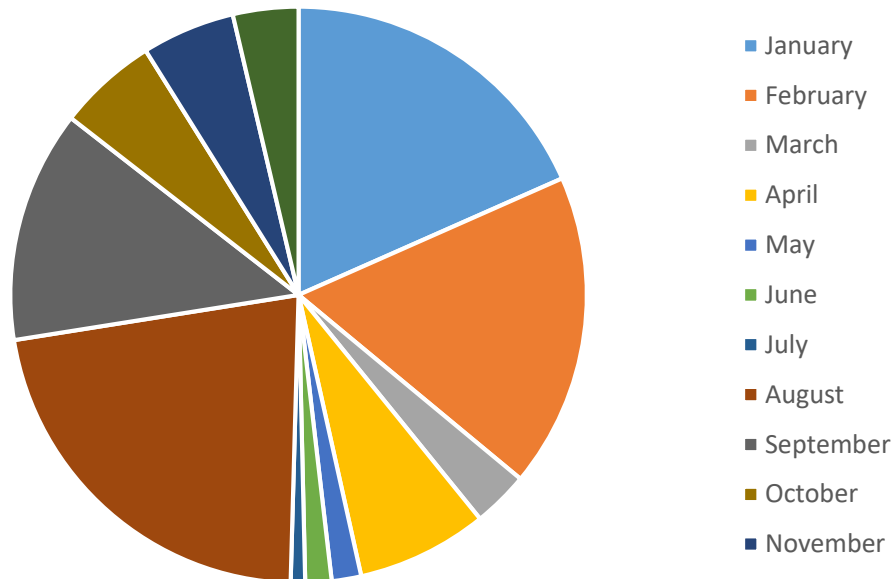
Figura 4.31. Datos ordinales representados con un gráfico de barras apiladas. Adaptación de la imagen cedida por Ann K. Emery. Fuente: <http://annkemery.com/pie-chart-guidelines/>.

Tipos de gráficos / Partes de un todo

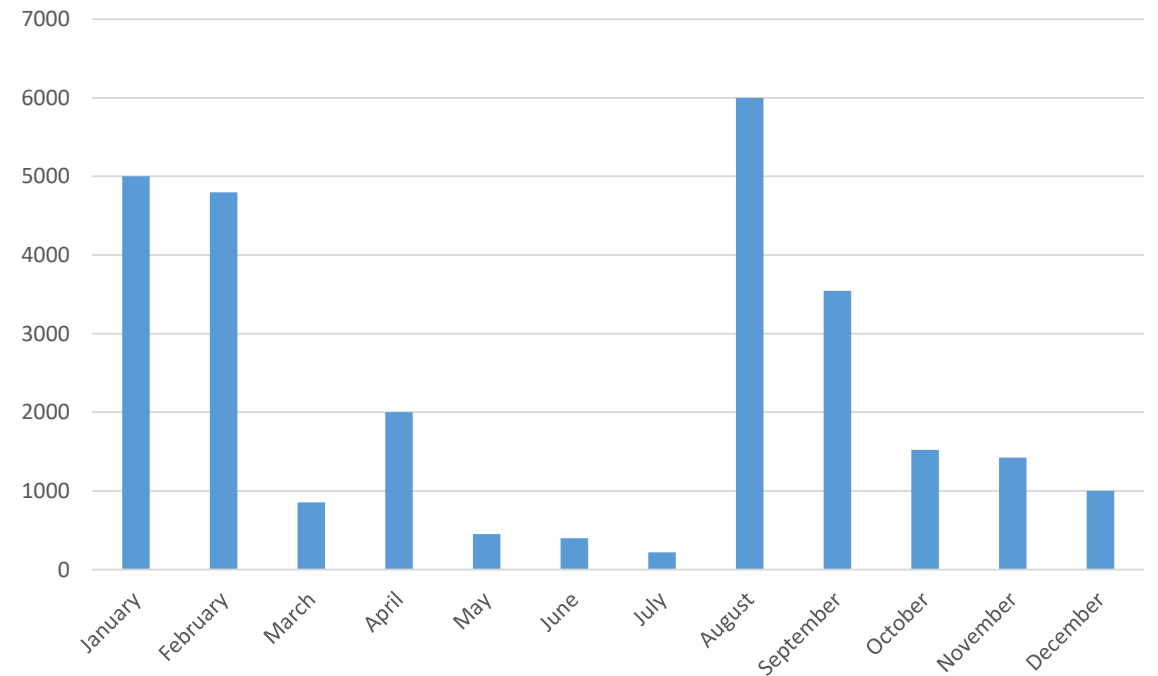
Buenas prácticas

Edward Tufte: *"the only worse design than a pie chart is several of them."*

Website Traffic by Month



Website Traffic by Month



Tipos de gráficos / Partes de un todo

Buenas prácticas

Evítalos si tienes muchas categorías

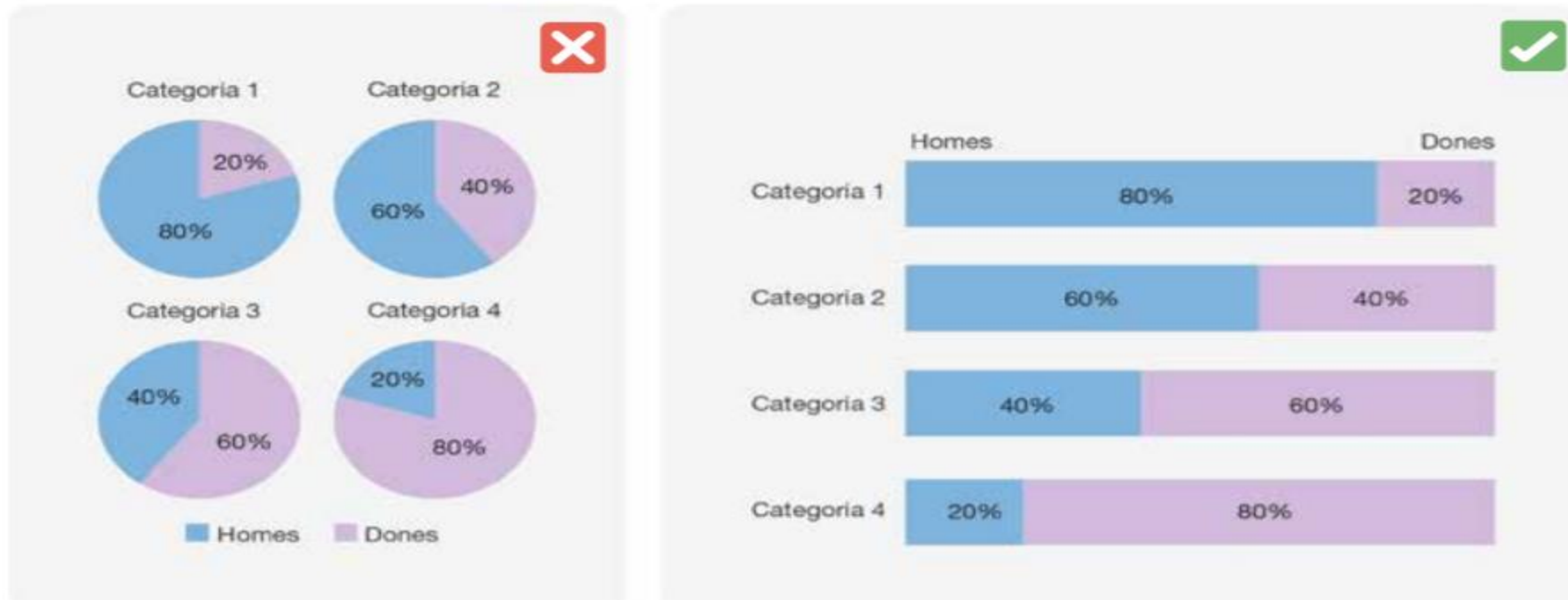


Figura 4.33. Es preferible utilizar barras para comparar valores a través de varias categorías. Adaptación de la imagen cedida por Ann K. Emery. Fuente: <http://annkemery.com/pie-chart-guidelines/>.

Tipos de gráficos / Partes de un todo

PICTOGRAMAS

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

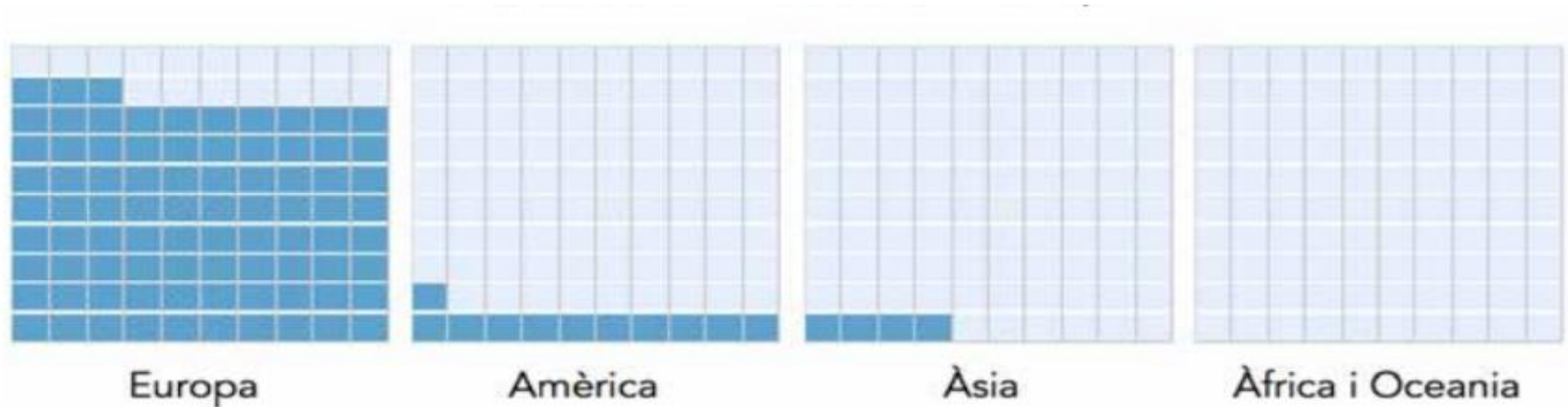


Figura 4.36. Percentatge de població que beu cava als diferents continents. Visualització feta per OneTandem.
Font: <https://public.tableau.com/profile/onetandem#!/vizhome/cava/cavaexports>.

Tipos de gráficos / Partes de un todo

TREEMAP

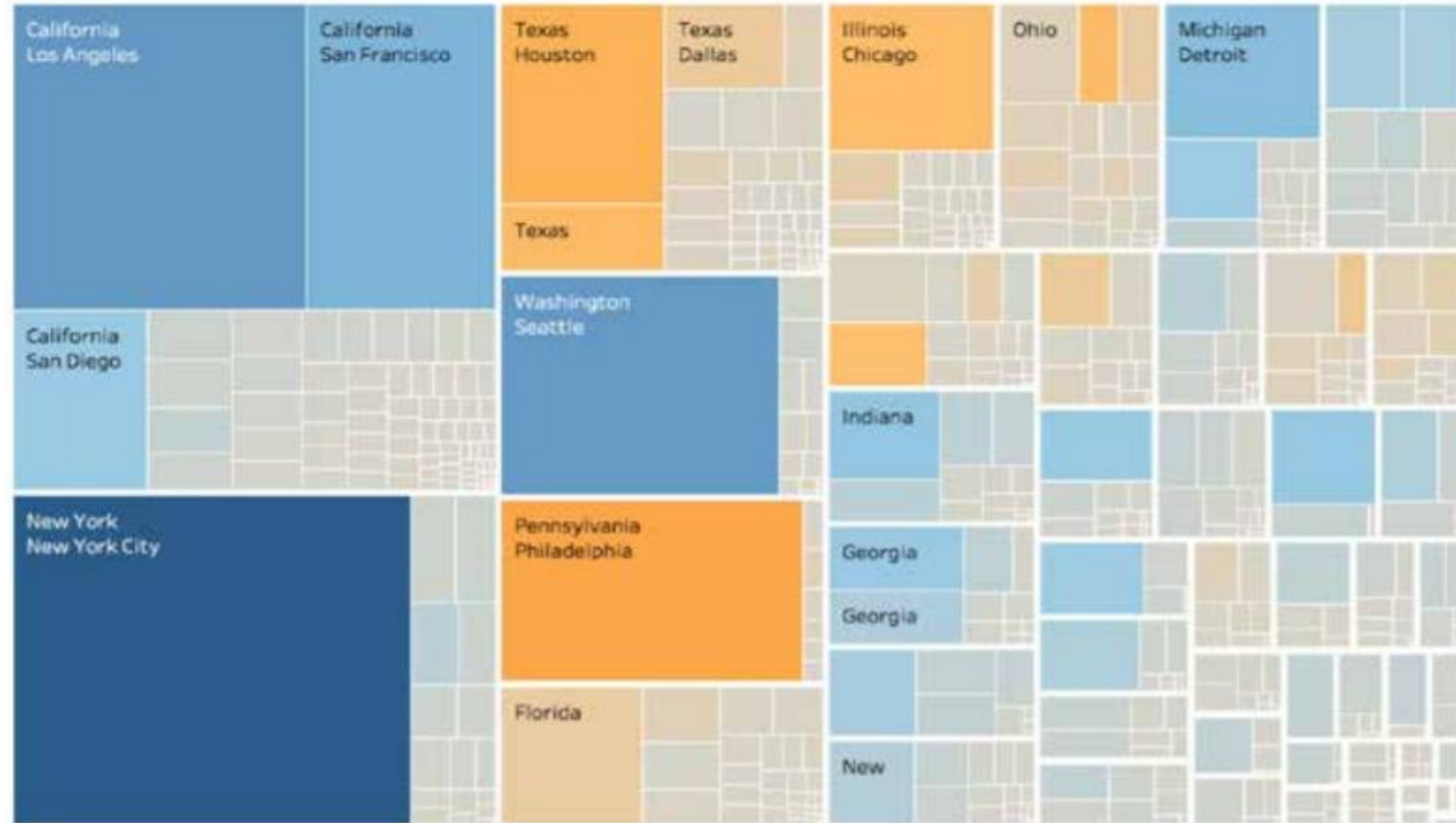
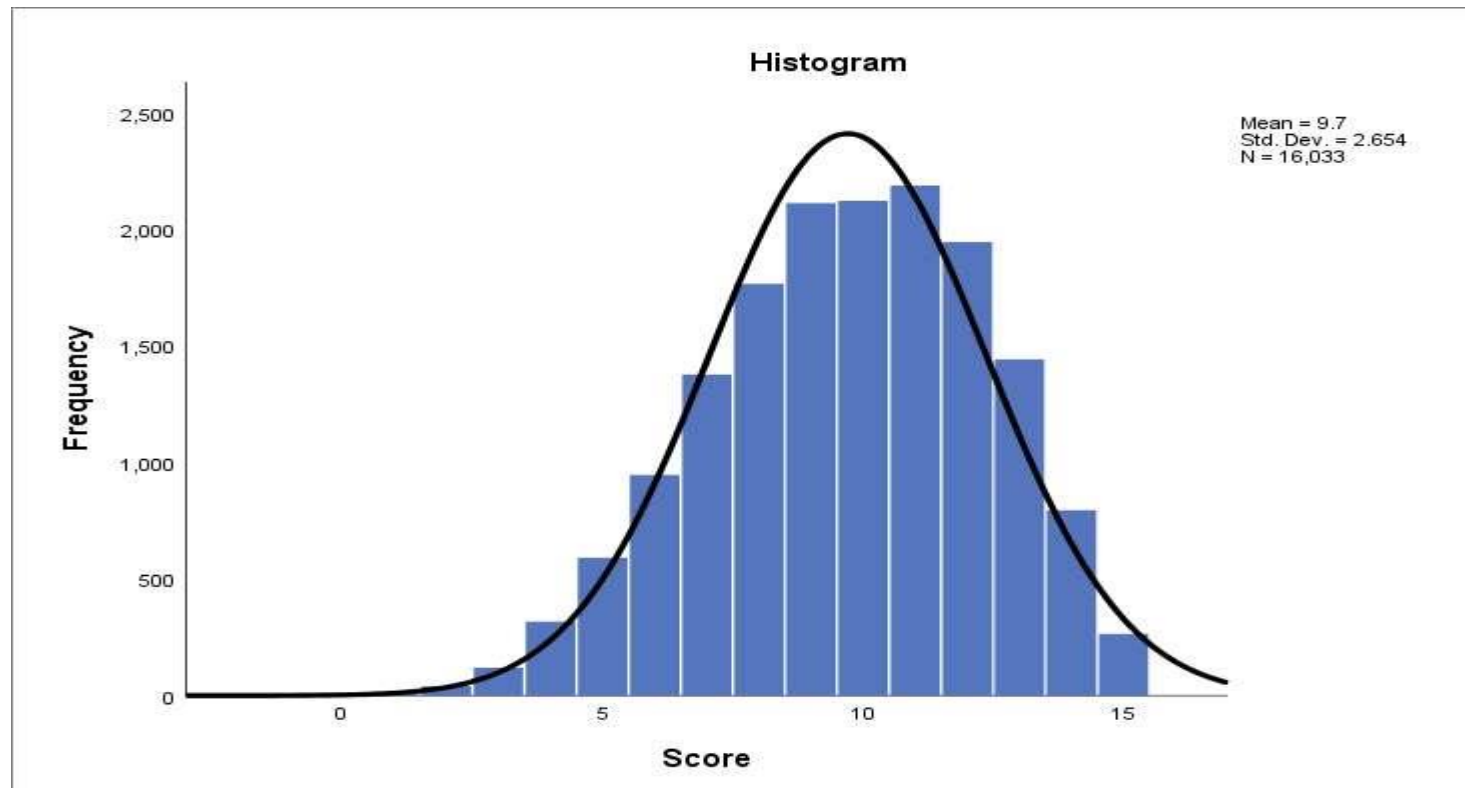


Figura 4.37. Mapa de árbol que representa el volumen de ventas (representado con la medida) y el beneficio (representado con el color) de una empresa ficticia en diferentes ciudades americanas agrupadas por estado.

Tipos de gráficos / Distribución

HISTOGRAMA

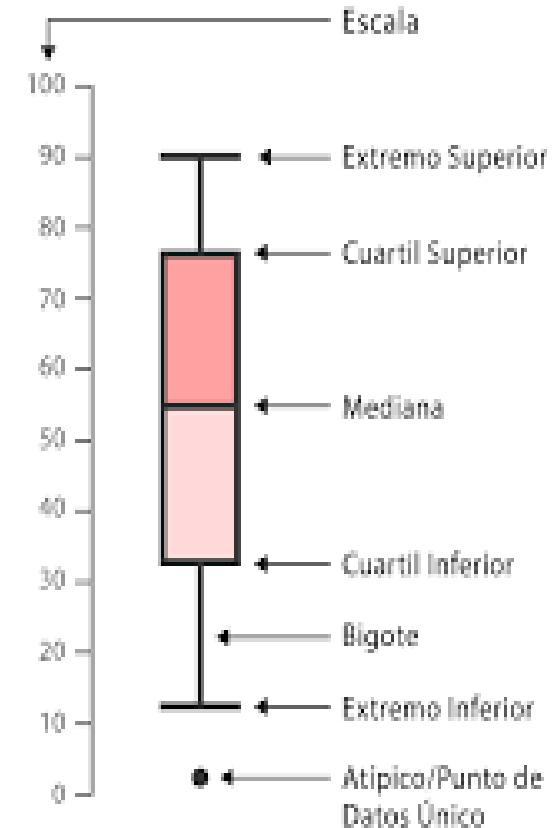
El **eje y** (a la izquierda) representa un recuento de frecuencia y el **eje x** (en la parte inferior), el valor de la variable (en este caso, el número de respuestas correctas).



Tipos de gráficos / Distribución

DIAGRAMA DE CAJAS

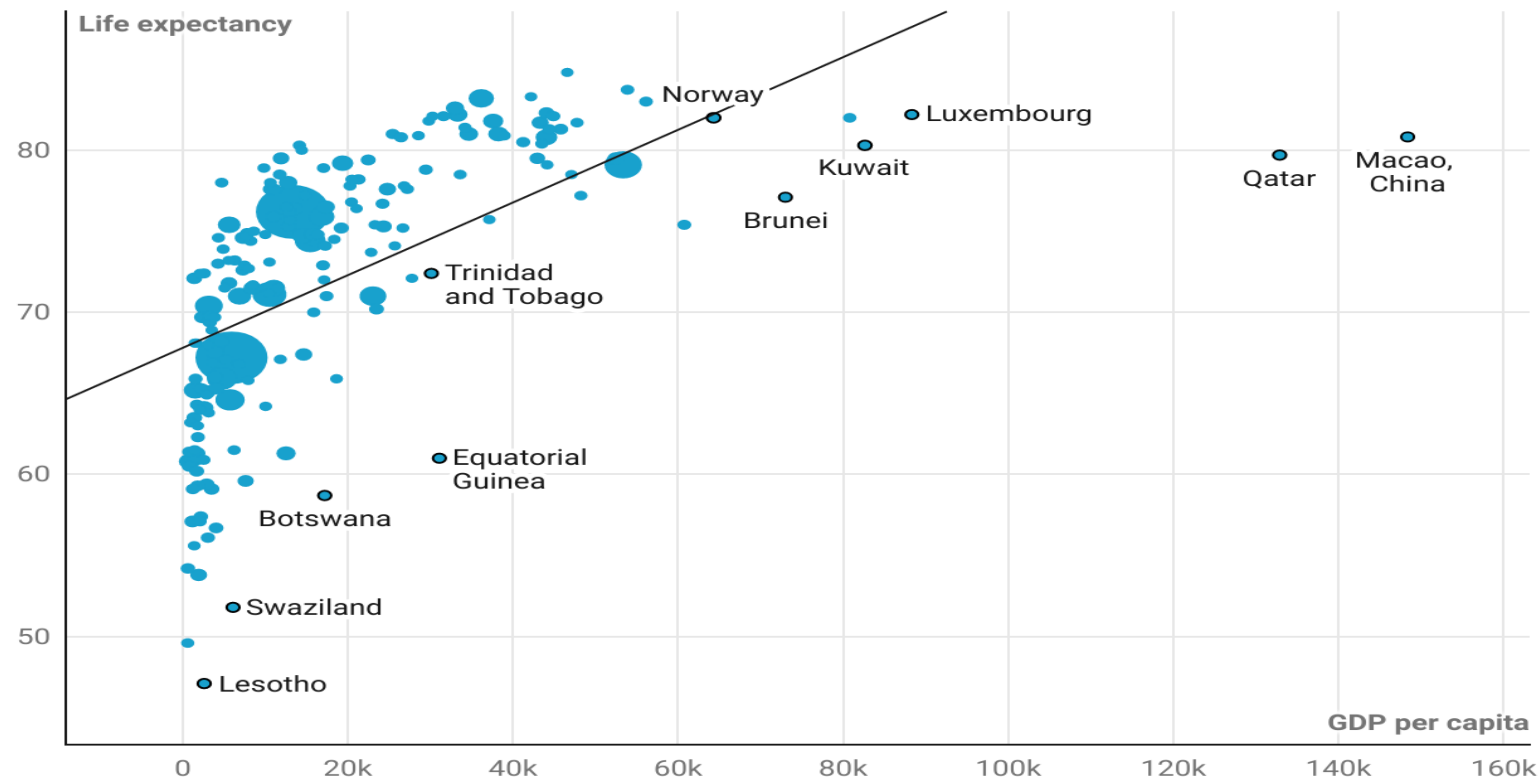
- La caja está dividida por una línea que representa la **mediana**.
- Los extremos de la caja corresponden **al primer y al tercer cuartil**.
- De la caja, emergen unas líneas denominadas "**bigotes**", que pueden calcularse de diferentes maneras y terminan en los valores **máximo y mínimo**
- Todos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



Tipos de gráficos / Distribución

GRÁFICO DE PUNTOS (SCATTER PLOT)

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



Fuente: Gapminder • Creado con Datawrapper

Tipos de gráficos / Conexiones

DIAGRAMA DE SANKEY

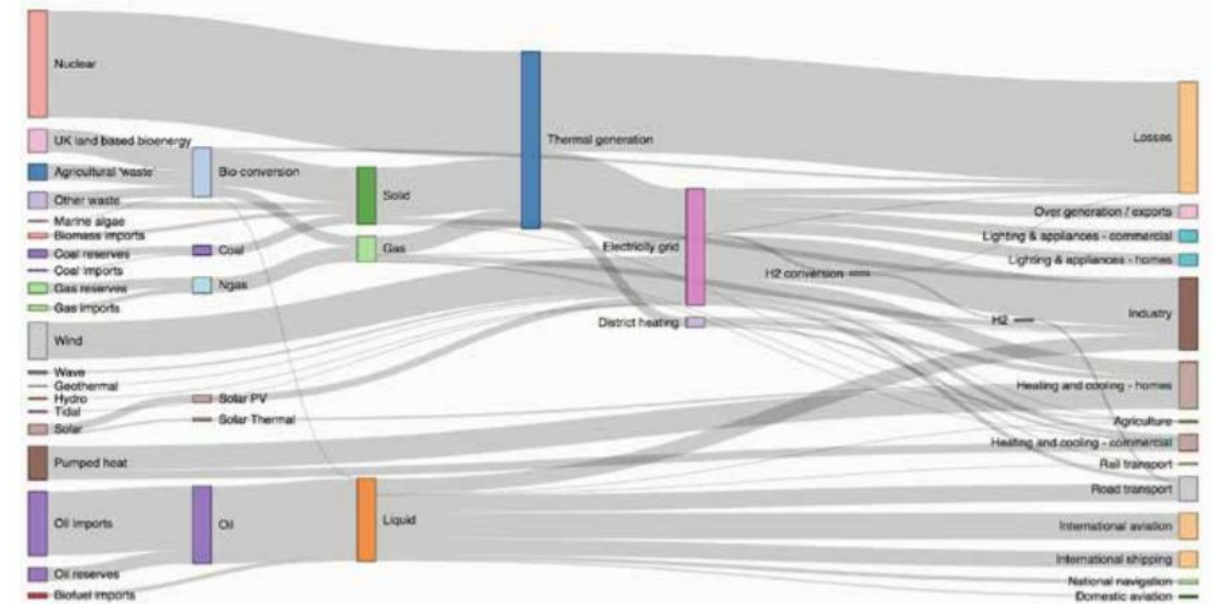


Figura 4.52. Diagrama de Sankey que muestra la cadena de producción y consumo de energía eléctrica.
Fuente: <https://bost.ocks.org/mike/sankey/>.

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.

Plantilla web: <https://cloud.squirrel365.io/.../3DD829C796E040ECBF40A976...>

COMMON MYTHS in DATA VIZ

MYTH:
LINE GRAPHS are
for CONTINUOUS
DATA ONLY

The LINES that CONNECT the POINTS HAVE to MAKE SENSE



MYTH:
BARS are
ALWAYS
BETTER

BARS are a GOOD PLACE to START... but NOT ALWAYS the BEST

Ask "what do I want my audience to see?"



Try other types of charts and decide what meets your needs

MYTH:
GRAPHS *
MUST have
a ZERO
BASELINE

* THIS is TRUE for BAR CHARTS



MYTH:
PIE CHARTS
are EVIL

When USING a PIE, ASK yourself WHY?



If you think pies make sense for your data
& audience, test them out to see!

MYTH:
UNBIASED
DATA EXISTS

We are **BIASING** our DATA at
EVERY STEP of the PROCESS

RULE! DON'T LIE
with DATA



MYTH:
MORE DATA is
ALWAYS BETTER

BEFORE CHASING after MORE DATA,
ask "WHAT will it HELP us DO or DECIDE?"

Audience & context are important when it
comes to the right amount of data.

MYTH:
AVERAGES
ALWAYS WORK to
SUMMARIZE DATA

YOU NEED to UNDERSTAND the
DISTRIBUTION, SPREAD, and VARIABILITY



MYTH:
THERE is a SINGLE
RIGHT ANSWER
when VISUALIZING
DATA

YOU SHOULD ALWAYS CONSIDER when
SHOWING DATA: WHAT is your GOAL?



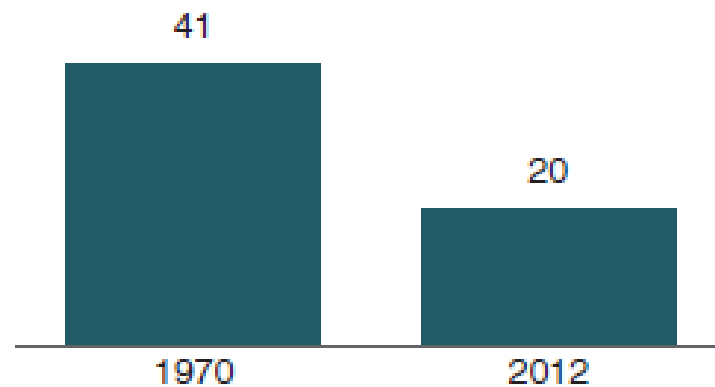
Tablas

Tablas vs gráficos

Texto simple

Children with a "Traditional" Stay-at-Home Mother

% of children with a married stay-at-home mother with a working husband



Note: Based on children younger than 18. Their mothers are categorized based on employment status in 1970 and 2012.

20%

of children had a **traditional stay-at-home mom** in 2012, compared to 41% in 1970

Storytelling With Data, pages 38-39

Tablas

- 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.


Tablas vs gráficos

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.

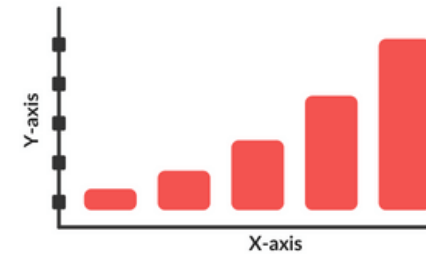
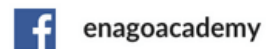
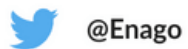


Figure 1. ABC

- 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!



<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 vs gráficos

- 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

Heavy borders

Group	Metric A	Metric B	Metric C
Group 1	\$X.X	Y%	Z,ZZZ
Group 2	\$X.X	Y%	Z,ZZZ
Group 3	\$X.X	Y%	Z,ZZZ
Group 4	\$X.X	Y%	Z,ZZZ
Group 5	\$X.X	Y%	Z,ZZZ

Light borders

Group	Metric A	Metric B	Metric C
Group 1	\$X.X	Y%	Z,ZZZ
Group 2	\$X.X	Y%	Z,ZZZ
Group 3	\$X.X	Y%	Z,ZZZ
Group 4	\$X.X	Y%	Z,ZZZ
Group 5	\$X.X	Y%	Z,ZZZ

Minimal borders

Group	Metric A	Metric B	Metric C
Group 1	\$X.X	Y%	Z,ZZZ
Group 2	\$X.X	Y%	Z,ZZZ
Group 3	\$X.X	Y%	Z,ZZZ
Group 4	\$X.X	Y%	Z,ZZZ
Group 5	\$X.X	Y%	Z,ZZZ

FIGURE 2.4 Table borders

Storytelling With Data, pages
40-41

Tablas / Buenas prácticas

Don't.

Name	ID	Status	Valuation (USD)
1. Mark Green	MG233	Active	\$ 3'234'500
2. Tom Jackson	TJ3346	Active	\$ 3'100'345
3. Tim Berg	TB3232	Active	\$ 115'600
4. Jeff Roden	JF4355	Archive	\$ 1'124'300
5. Kate Watson	KW345	Archive	\$ 123'445
6. Jessi Johnson	JJ3432	Archive	\$ 224'040
7. Mark Green	MG342	Blocked	\$ 234'500
8. Tom Jackson	TJ2343	Archive	\$ 1010'345

Do.

Name	ID	Status	Valuation (USD)
1. Mark Green	MG233	Active	\$ 3'234'500
2. Tom Jackson	TJ3346	Active	\$ 3'100'345
3. Tim Berg	TB3232	Active	\$ 115'600
4. Jeff Roden	JF4355	Archive	\$ 1'124'300
5. Kate Watson	KW345	Archive	\$ 123'445
6. Jessi Johnson	JJ3432	Archive	\$ 224'040
7. Mark Green	MG342	Blocked	\$ 234'500
8. Tom Jackson	TJ2343	Archive	\$ 1010'345

Qubstudio (2017). Data Tables Design Basics

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 / Heatmaps

Tabla

	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

LOW-HIGH

	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%

FIGURE 2.5 Two views of the same data

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

Tablas / Buenas prácticas

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	Media de tweets	Nº de lecturas en Mendeley	Media lecturas en Mendeley	Nº bookmarks en Citelik	Media bookmarks Citelik		
ABADAL-FALGUERAS, ERNEST	UB	186	✓	✓	✓	137	19509	15	18	1,3	189	9	4	0,2	
ALEIXANDRE-BENAVENT, RAFAEL	LIV	○	○	○	✓	○	○	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-ANEAGON, FELIX	CCHS-CSIC	✓	622	✓	✓	✓	○	○	52	101	2,1	750	2,9	69	0,3
DIAZ-NOCI, 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	UPVME3	✓	56	✓	✓	✓	○	○	3	0	0	8	2,7	2	0,7
GONZALEZ-PACANOWSKI, TONI	Prestigio	✓	232	✓	○	✓	40	—	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	○	✓	✓	✓	1.907	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, JOSEP-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	✓	○	✓	95	—	6	14	2,3	56	14	3	0,8
PESET-MANCEBO, FERNANDA	UPV	○	○	✓	✓	✓	27	—	11	0	0	47	4,7	5	0,5
RAMIREZ-POSADA, MARIA-VICTORIA	UPVME3	○	○	○	○	○	○	○	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, JOSEP-MANUEL	UB	○	○	✓	✓	✓	X	○	6	4	0,7	56	56	2	2
RODRIGUEZ-PARADA, CONCEPCION	UB	○	○	○	○	○	○	○	4	0	0	21	5,3	2	0,5
ROVIRA-FONTANALS, CRISTOFOL	IUPF	✓	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	CLUM	✓	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	○	✓	✓	230	22.821	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 / Buenas prácticas

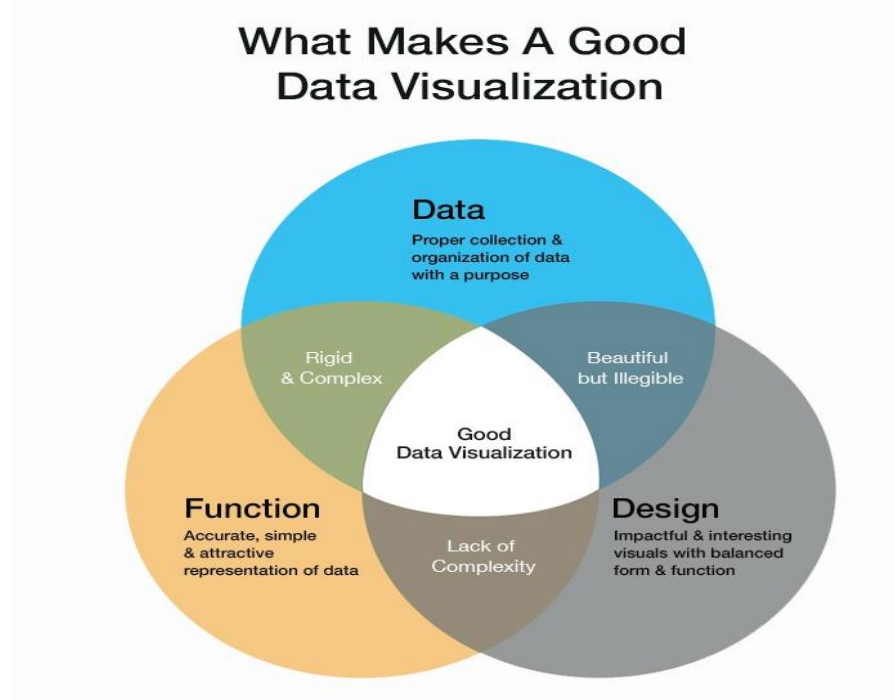
- Para condiciones en las que necesita información definida, como cifras y valores precisos.
- Cuando son mucho datos que un gráfico no mostraría con facilidad
- Muy útil cuando muestras datos por grupos y categorías
- Se relacionan con un sistema de observación, lo que nos permite procesar hechos y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

¿Cómo conseguimos visualizaciones
efectivas ?

¿ Cómo alcanzar una visualización efectiva ?

- Tips para conseguirla
- ¿ “bonita” o funcional?
- ¿ Es el color importante ?
- ¿ Influencias científicas y/o psicológicas?
- Principios

¿ Cómo alcanzar una visualización efectiva ?



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

5 cualidades de una buena visualización

- Truthful
- Functional
- Beautiful
- Insightful
- Enlightening



The Truthful Art, page 45

¿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

Data-ink ratio

$$\text{Data-ink ratio} = \frac{\text{Data-ink}}{\text{Total ink used to print the graphic}}$$

= proportion of a graphic's ink devoted to the non-redundant display of data-information

= 1.0 - proportion of a graphic that can be erased

Edward Tufte. The Visual Display of Quantitative Information
<https://www.youtube.com/watch?v=JIMUzJzqaA8>

A low Data-Ink ratio of .1 would mean that for every one pixel that communicates your message, nine pixels are wasting space and your audience's attention.

Data-ink ratio

A low Data-Ink ratio of .1 would mean that for every one pixel that communicates your message, nine pixels are wasting space and your audience's attention.

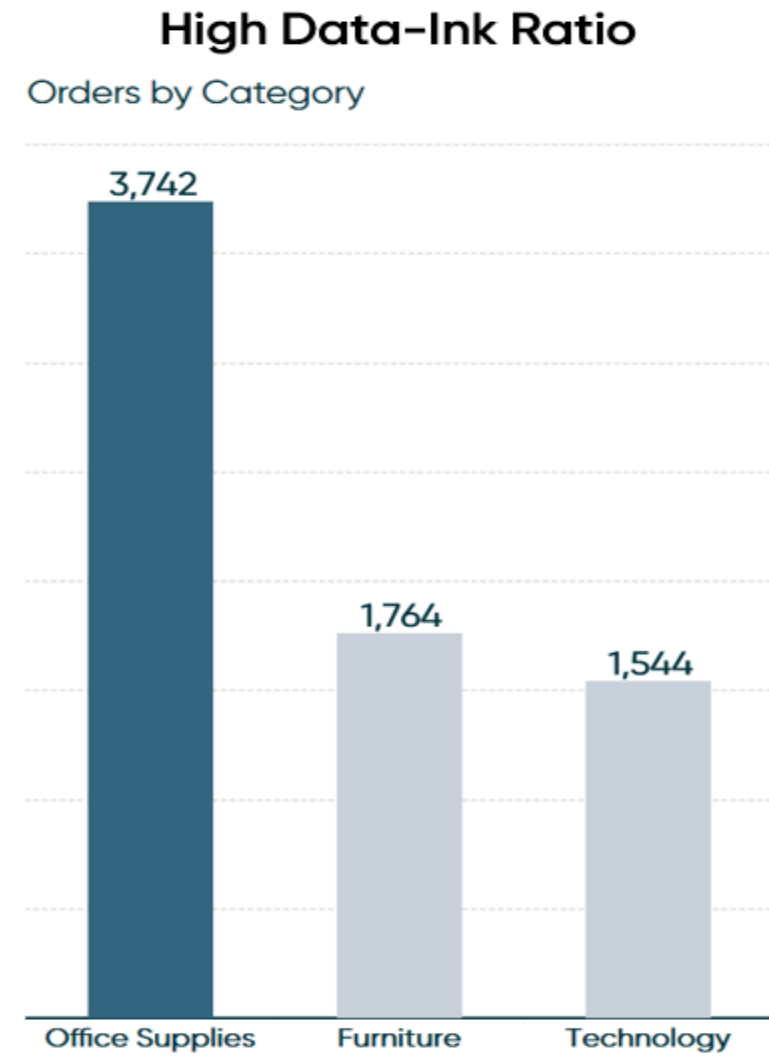
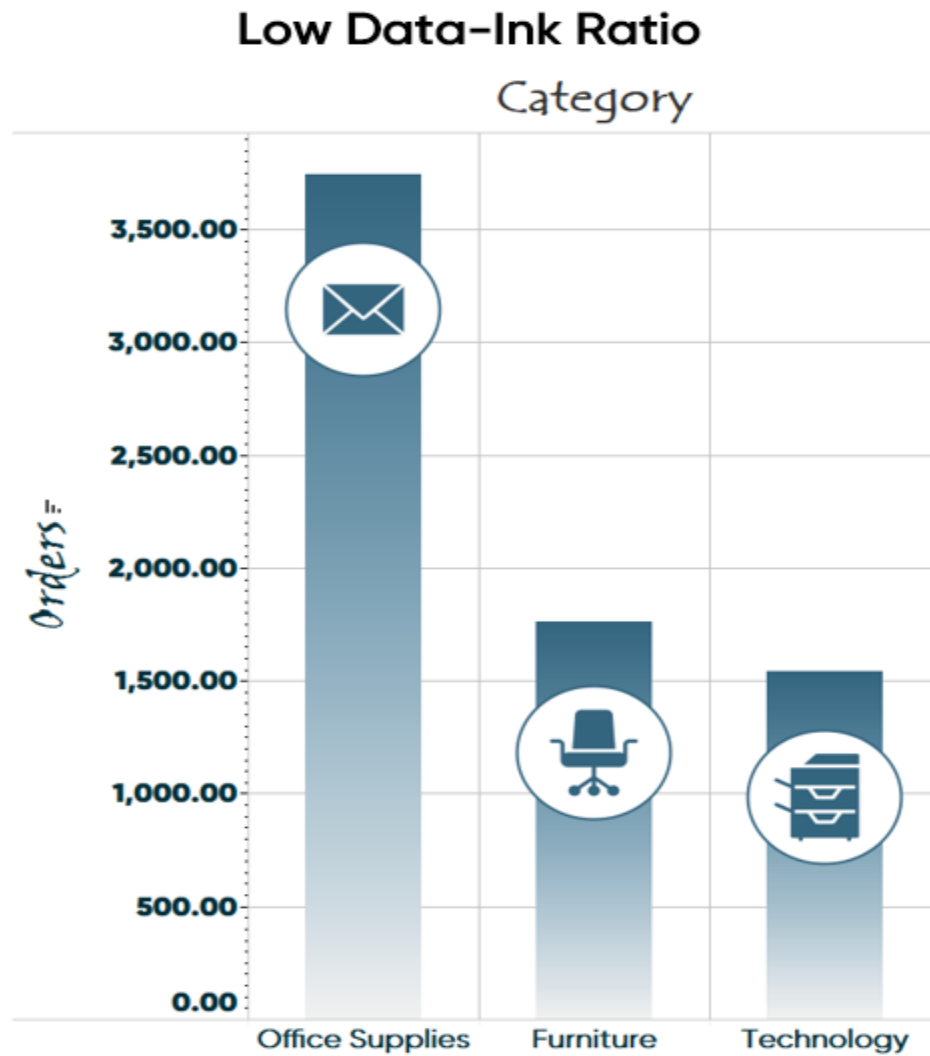
Edward Tufte. The Visual Display of Quantitative Information
<https://www.youtube.com/watch?v=JIMUzJzqaA8>

Data-ink ratio

- **Rule:** Maximize the data-ink ratio, within reason
- **Two Erasing Principles**
 - Erase non-data-ink, within reason
 - Erase redundant data-ink, within reason

Edward Tufte. The Visual Display of Quantitative Information

Data-ink ratio



<https://playfairdata.com/data-ink-ratio-animation-and-how-to-apply-it-in-tableau/>

Pensamiento visual

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

- “an effective display must be easily encoded and comprehended by the visual information-processing system” (Kosslyn [1989](#), 192)

- **Visual information processing:**

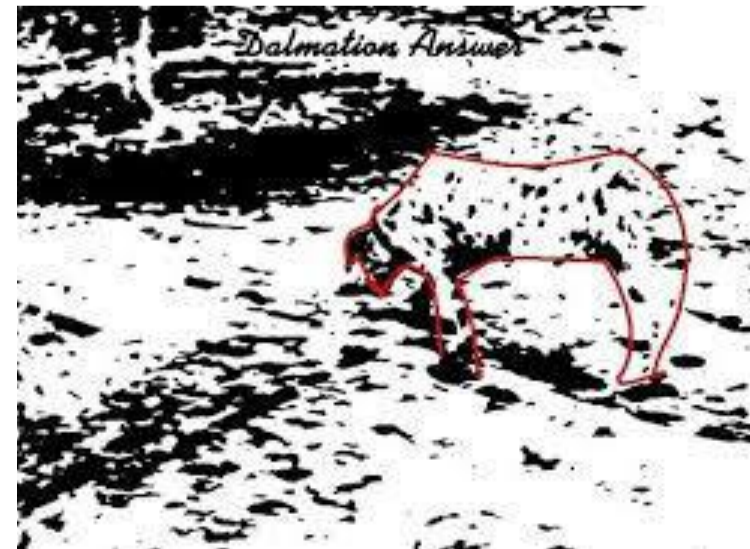
- → Perceptual image (1)
- → Short-term memory (2)
- → Long-term memory (3) (Kosslyn [1989](#), 190-91)

Kosslyn, Stephen M. 1989. “Understanding Charts and Graphs.” *Appl. Cogn. Psychol.* 3 (3): 185-225.

Percepción visual y principios de Gestalt

Formar patrones complejos a partir de reglas simples...

¿Qué pueden ver en esta figura?



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

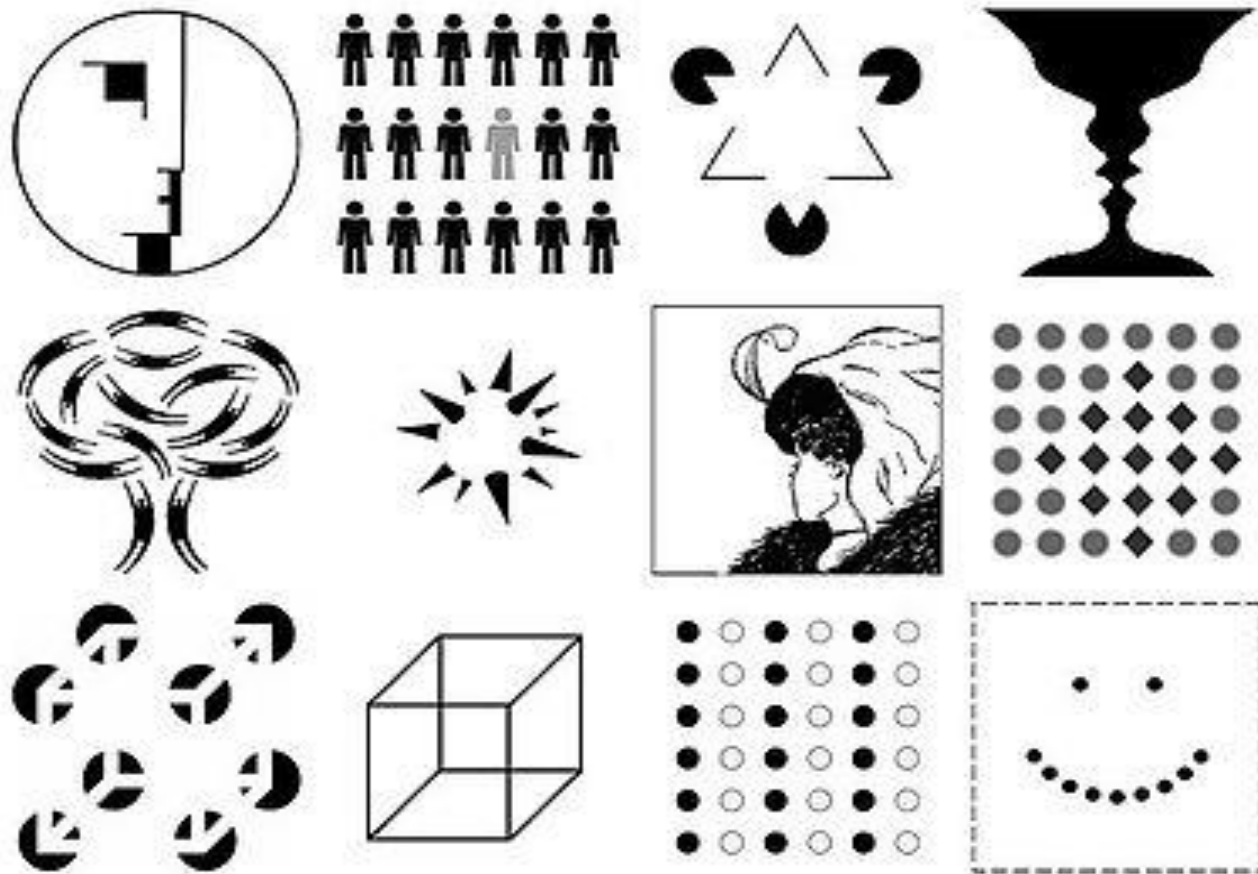
Permite:

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

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

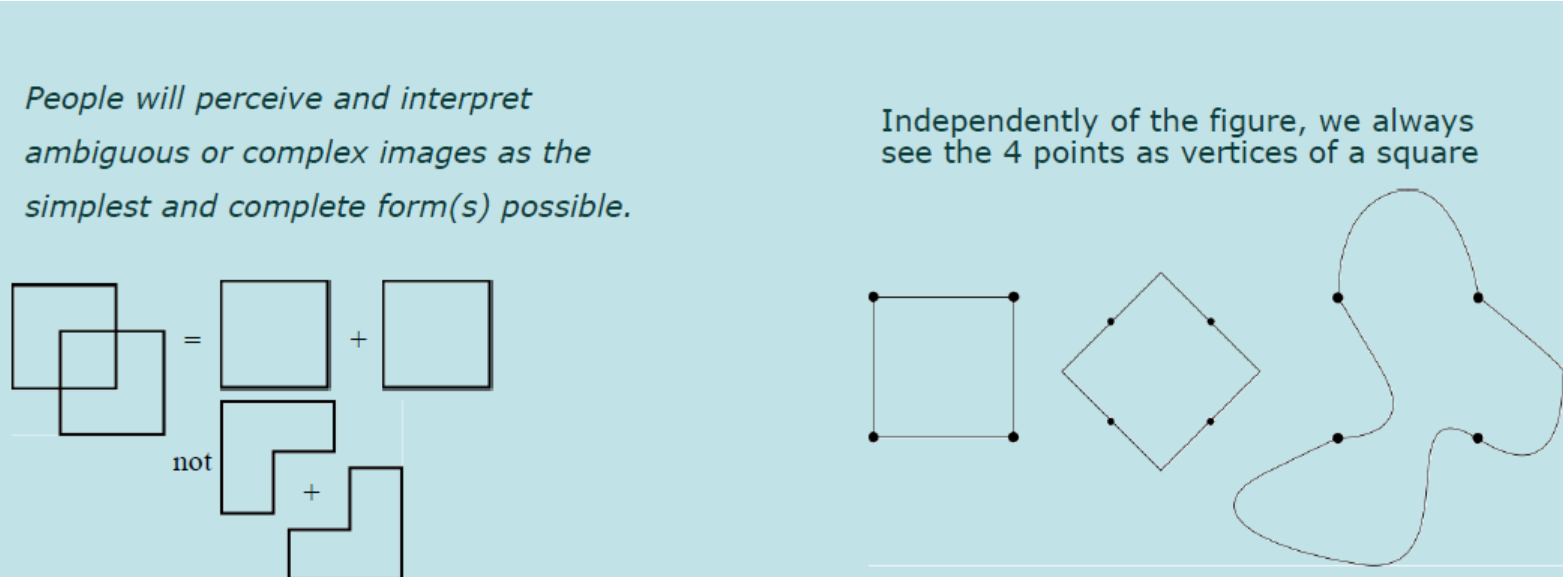
Psicología Gestalt (Berlin, años 20)



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.

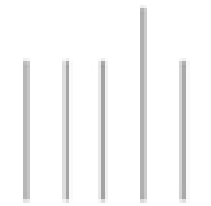


Independently of the figure, we always see the 4 points as vertices of a square

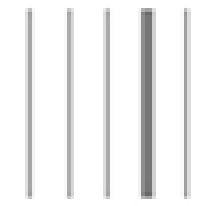
European Commission. 2017. 15TH JRC TRAINING ON COMPOSITE INDICATORS AND ANNUAL MEETING OF COP ON INDICES & SCOREBOARDS.

Atributos visuales Pre-atentivos

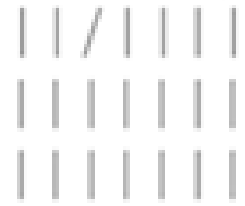
Length



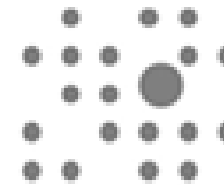
Width



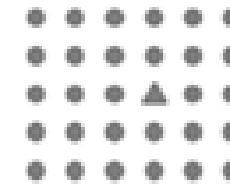
Orientation



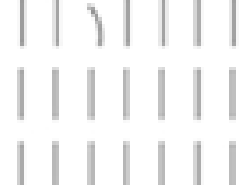
Size



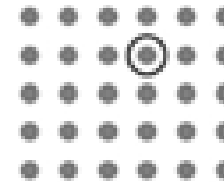
Shape



Curvature



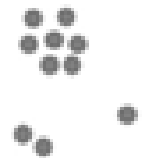
Enclosure



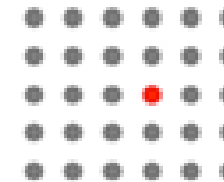
2-D Position



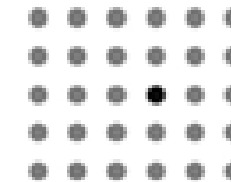
Spatial Grouping



Color (Hue)

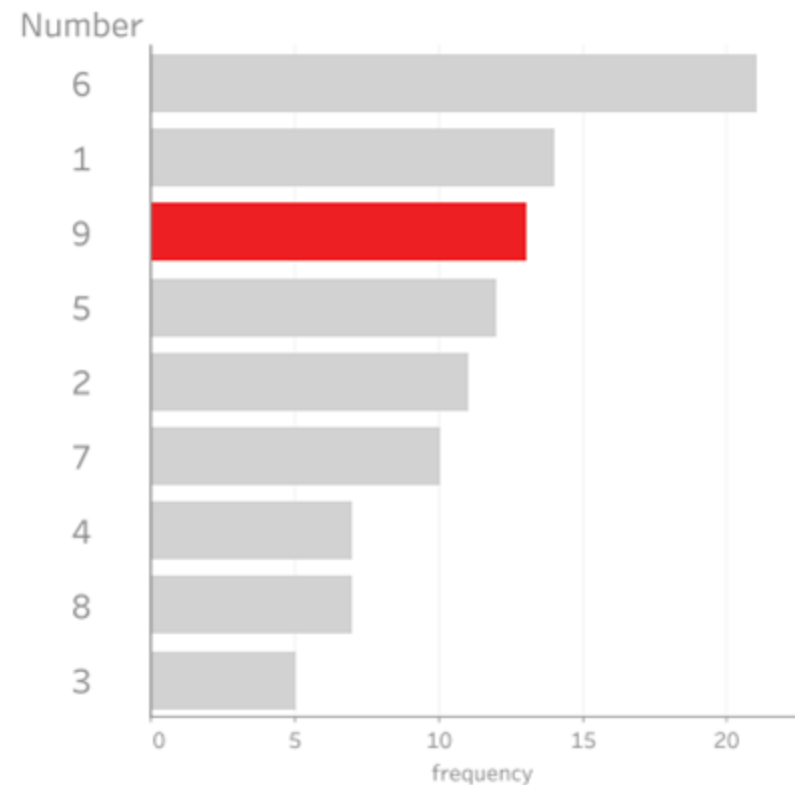


Color (Intensity)



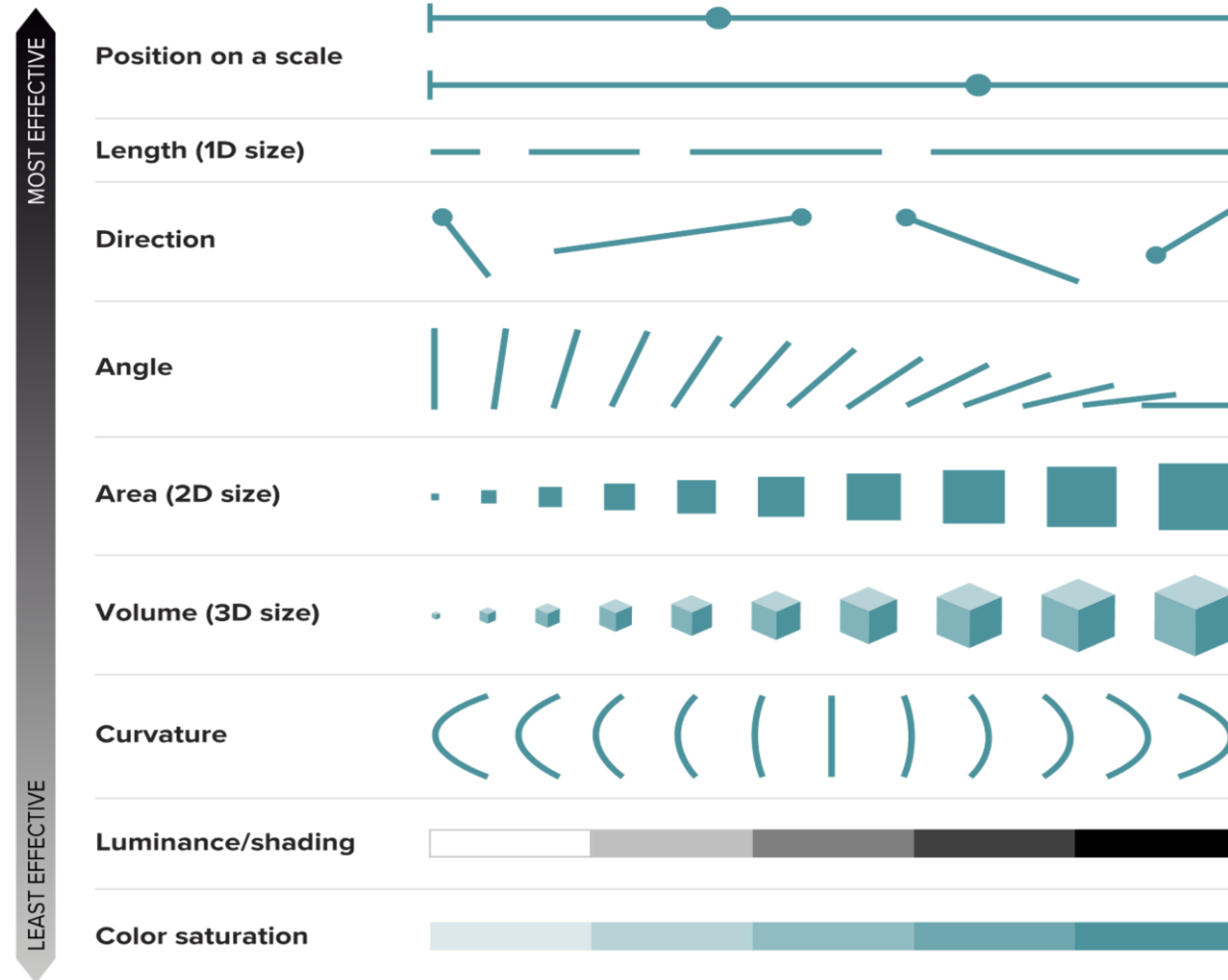
Atributos visuales Pre-atentivos

2	2	5	6	7	1	1	6	9	1
9	1	7	5	5	5	6	2	5	9
4	5	2	9	6	9	7	6	4	6
8	1	5	7	8	5	6	6	6	7
7	2	3	6	8	9	1	7	9	1
3	8	6	8	4	5	6	9	4	5
4	9	9	2	3	7	1	9	1	2
3	7	8	1	6	1	5	6	1	6
5	6	6	8	6	6	9	1	2	6
3	2	4	2	6	9	4	2	7	1



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.



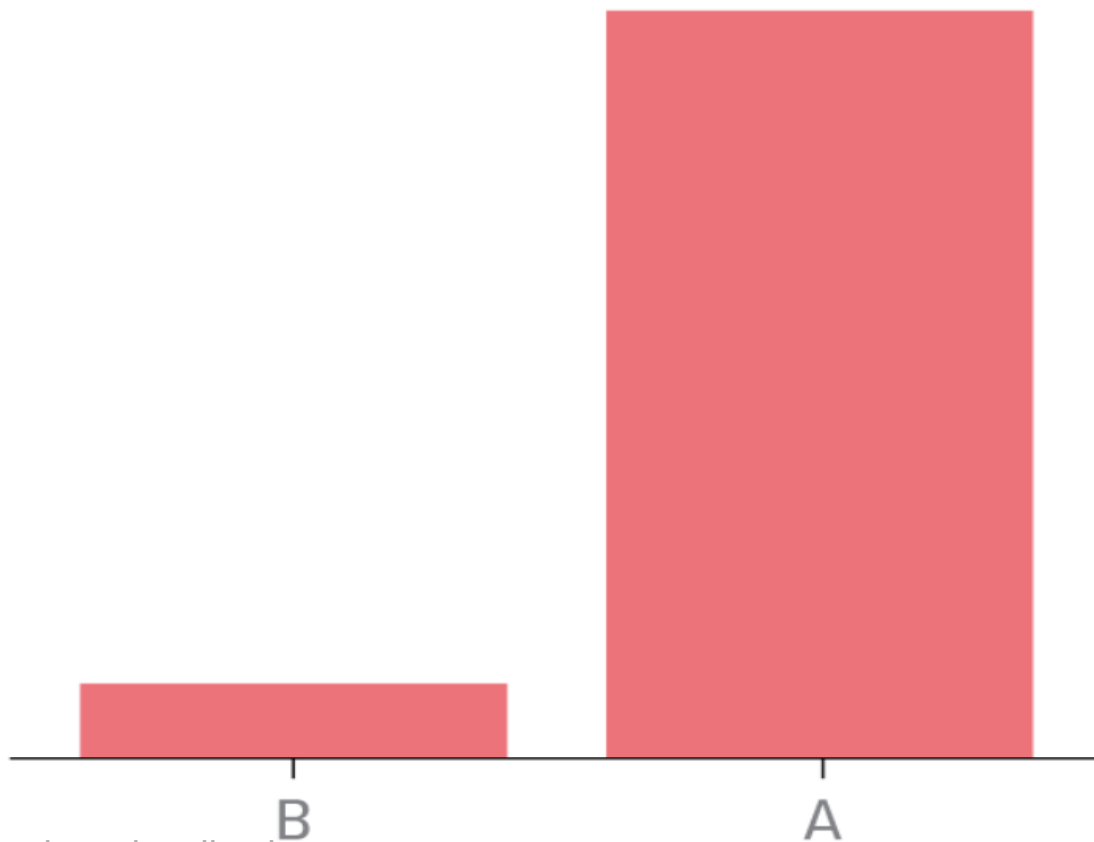
SOURCES: W.S. CLEVELAND AND R. MCGILL / JOURNAL OF THE AMERICAN STATISTICAL ASSOCIATION 1984;
S.I. O'DONOGHUE ET AL / AR BIOMEDICAL DATA SCIENCE 2018

5W INFOGRAPHIC / KNOWABLE

Expresividad atributos gráficos

¿Cuál es la relación proporcional entre A y B ?

10x



Mason, B. (2019). Why scientists need to be better at data visualization

Legibilidad



Designs **BECOME**
cluttered and *confusing*
when using too
many **FONT'S**.



Designs become
cluttered and confusing
when using too
many fonts.

<https://medium.com/global-maksimum-data-information-technologies/data-visualization-5-most-important-things-to-know-725323bd2bca>

Legibilidad

For Numbers...

Din

0123456789

Helvetica Neue

0123456789

Myriad

0123456789

Gills Sans

0123456789

Tahoma

0123456789

Azuro

1 2 3 4 5 6 7 8 9 0

Lucida Sans

0123456789

For Text...

Meta

Quip-crazy Gorbach

Tahoma

AaBbCcDdEeFfGg

Azuro

AaBbCc Handgloves

Lucida Sans

Quincy Jones vowed

[For titles] Adobe Jenson

AaBbCcDdEeFfGg

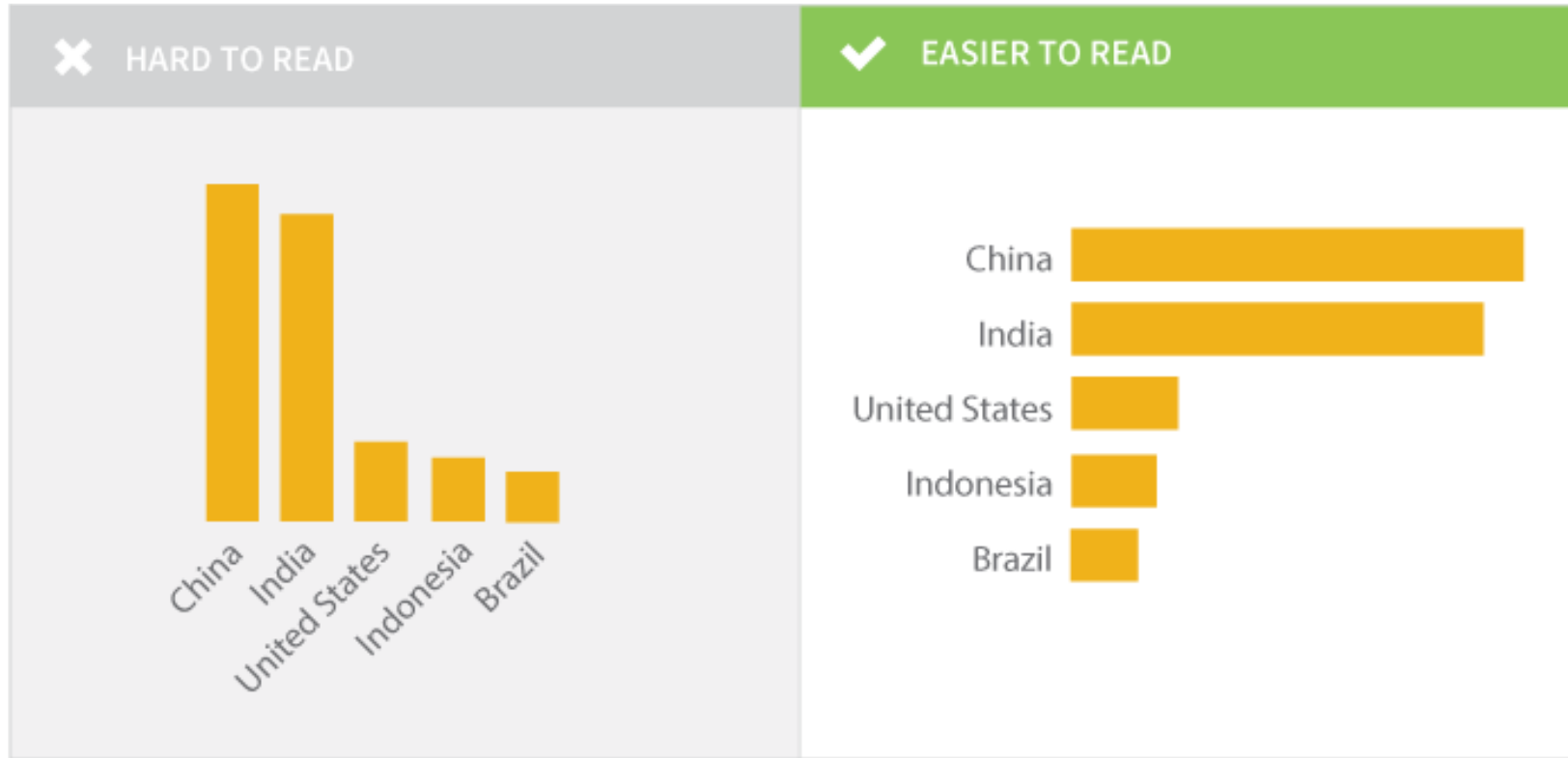
Mason, B. (2019). Why scientists need to be better at data visualization

Legibilidad

Esto resulta bastante fácil
de leer

En cambio este texto es
más complicado de leer

Legibilidad

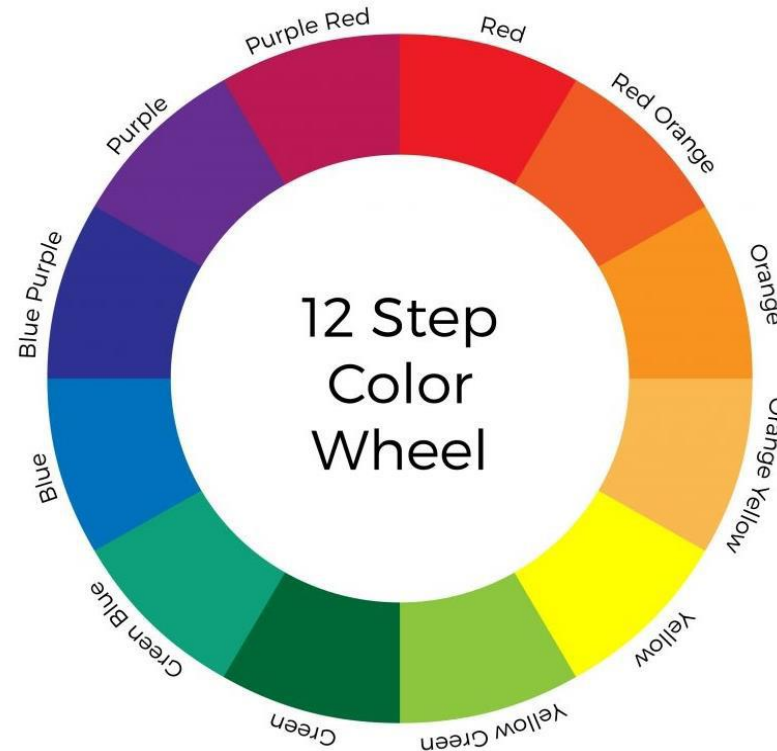


Chiasson, T., Gregory, D., et al. Data+Design.

Uso del color para ayudar a “leer” los datos

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 para ayudar a “leer” los datos

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

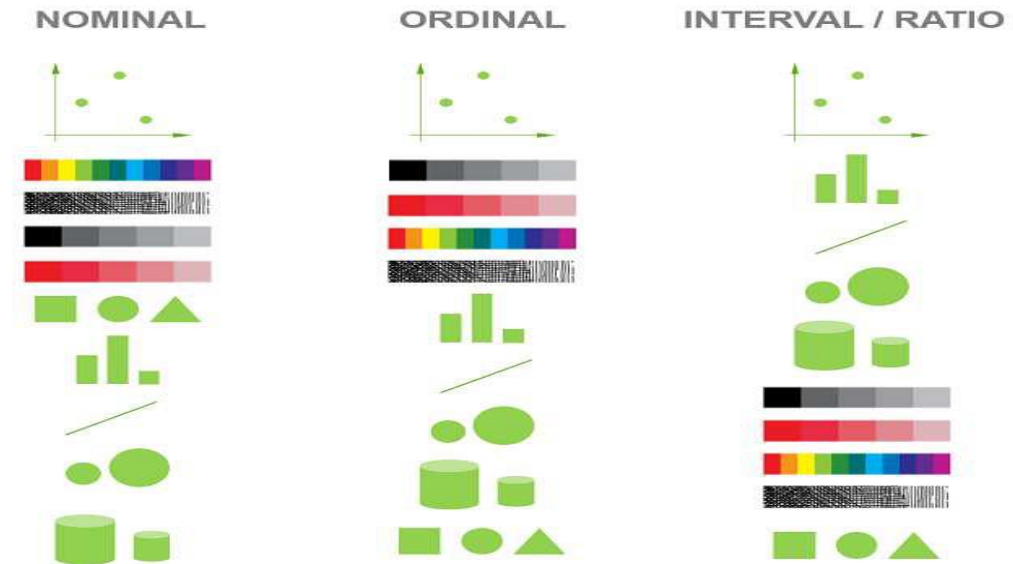


Uso del color para ayudar a "leer" los datos

Utilizamos gradientes para datos cuantitativos.

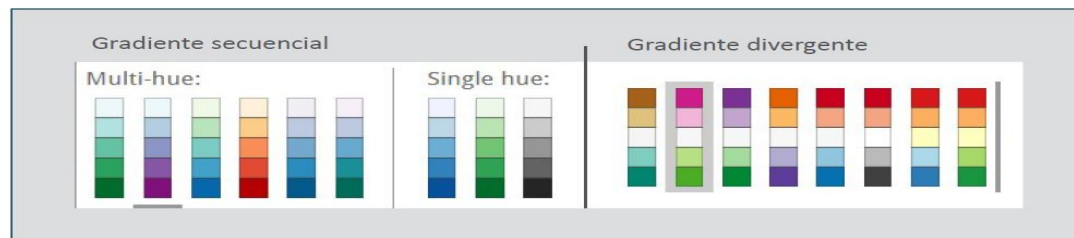
- Gradientes secuenciales.
- Gradientes divergentes.

Utilizamos una composición de colores para representar información cualitativa.

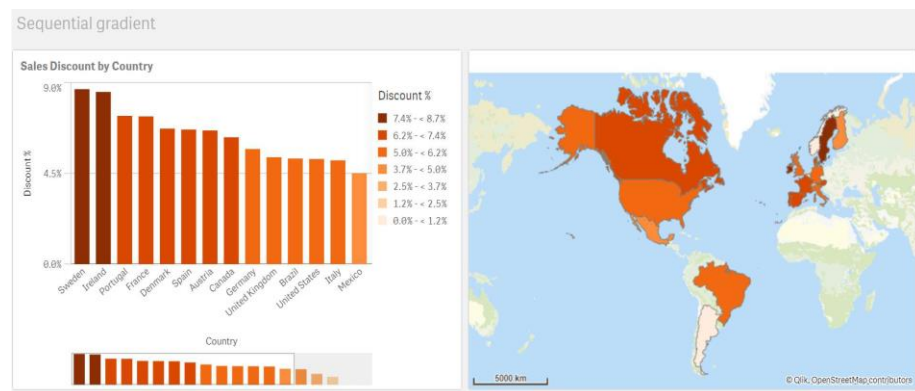


Uso del color para ayudar a "leer" los datos

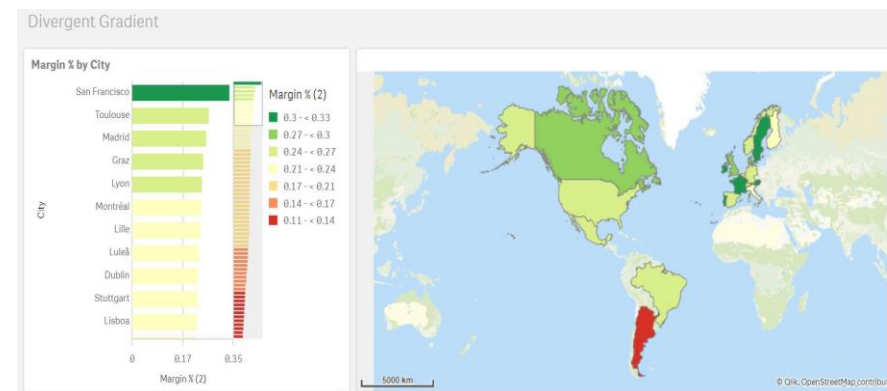
Información Cuantitativa



<http://colorbrewer2.org/>



Los datos van de pequeño a grande.



Mostrar desviaciones de un punto central.

Uso del color para ayudar a "leer" los datos

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



Herramientas para la selección del color

The screenshot displays the ColorBrewer 2.0 interface. At the top right, it says "COLORBREWER 2.0 color advice for cartography". The main interface is divided into several sections:

- Number of data classes:** A dropdown menu set to "5".
- Nature of your data:** Radio buttons for "sequential", "diverging" (selected), and "qualitative".
- Pick a color scheme:** A grid of 12 color scheme thumbnails. The "5-class RdBu" scheme is highlighted.
- 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".
- 5-class RdBu legend:** A vertical legend showing five color swatches with their corresponding hex codes: red (#ca0020), orange (#f4a582), white (#f7f7f7), light blue (#92c5de), and dark blue (#0571b0).
- EXPORT:** A vertical button on the right side of the legend.
- Color transparency:** A horizontal slider at the bottom left of the legend area.

The main map area shows a geographical map of a region, likely in Spain, colored according to the selected 5-class RdBu scheme. The colors transition from red in the northwest to blue in the southeast, with orange and white in the center.

At the bottom of the interface, there is a copyright notice: "© Cynthia Brewer, Mark Harrower and The Pennsylvania State University". Below this are links for "Source code and feedback", "Back to Flash version", and "Back to ColorBrewer 1.0". The "axismaps" logo is visible in the bottom right corner.

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

Elimina Ruido. Reduce la carga cognitiva

Destaca lo importante

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

Presta atención a los detalles. Optimiza tu gráfico

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

Performance overview

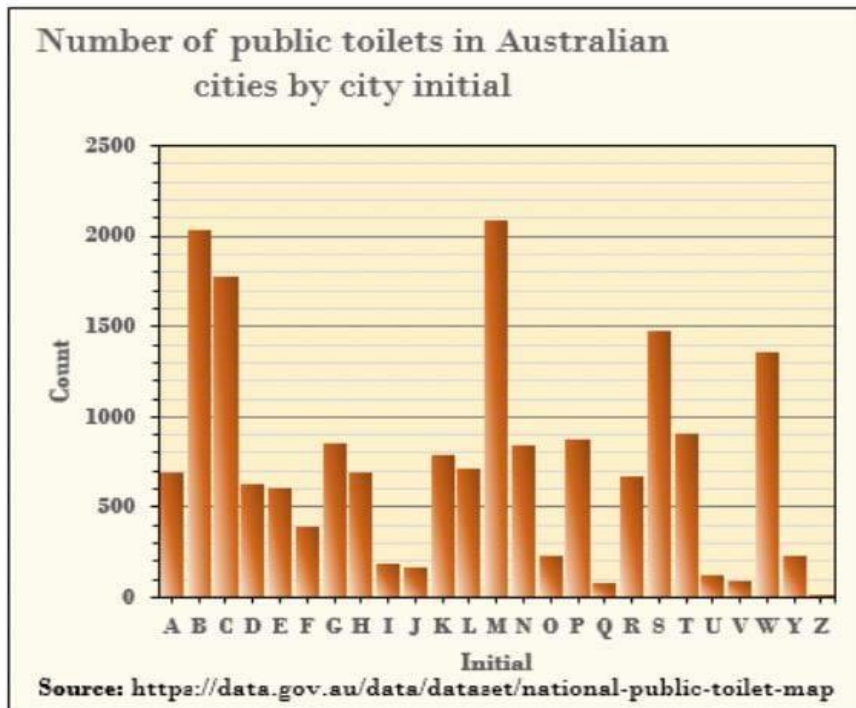
■ Our business

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



Presta atención a los detalles. Optimiza tu gráfico

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

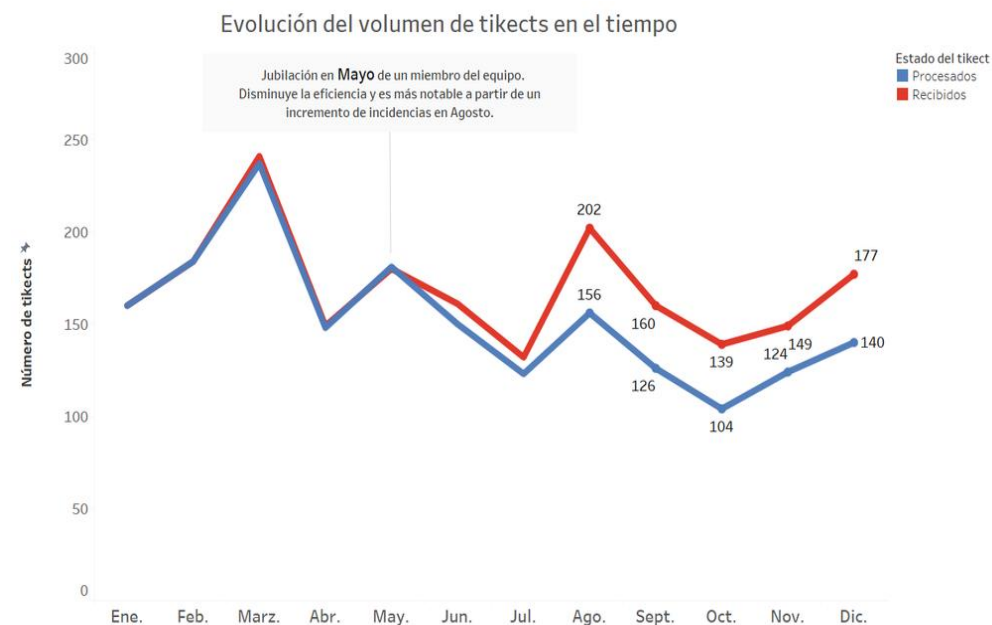
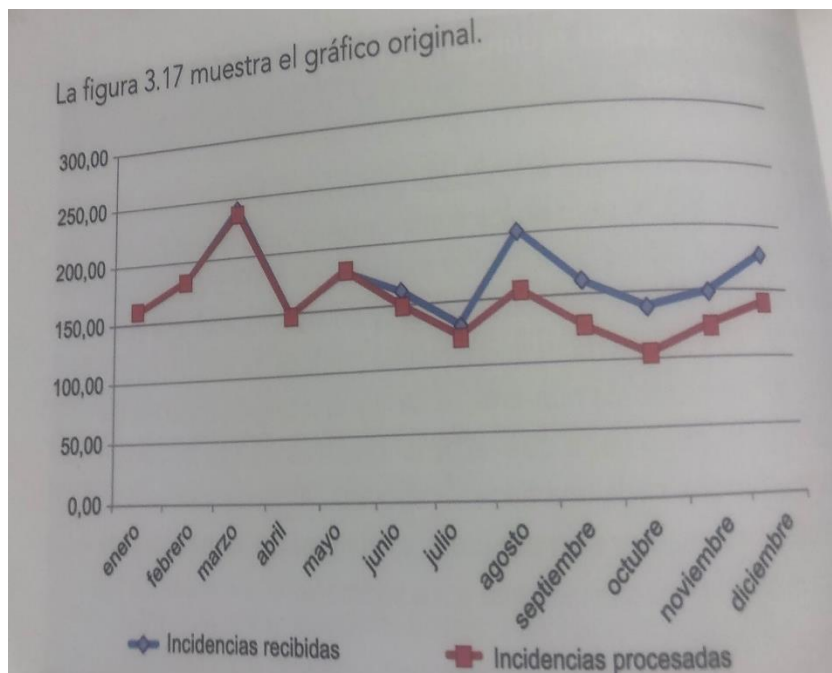


Less than 10% of public toilets are in cities starting with vowels

Number of public toilets in Australian cities, by city initial

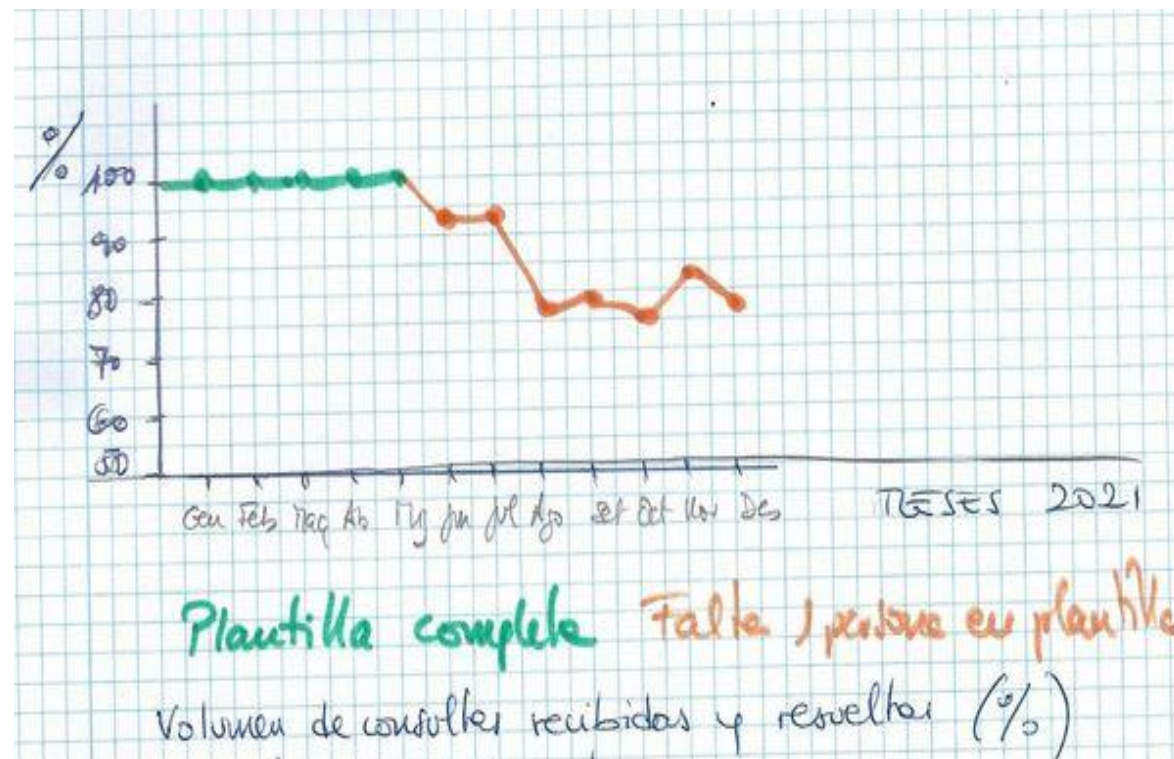
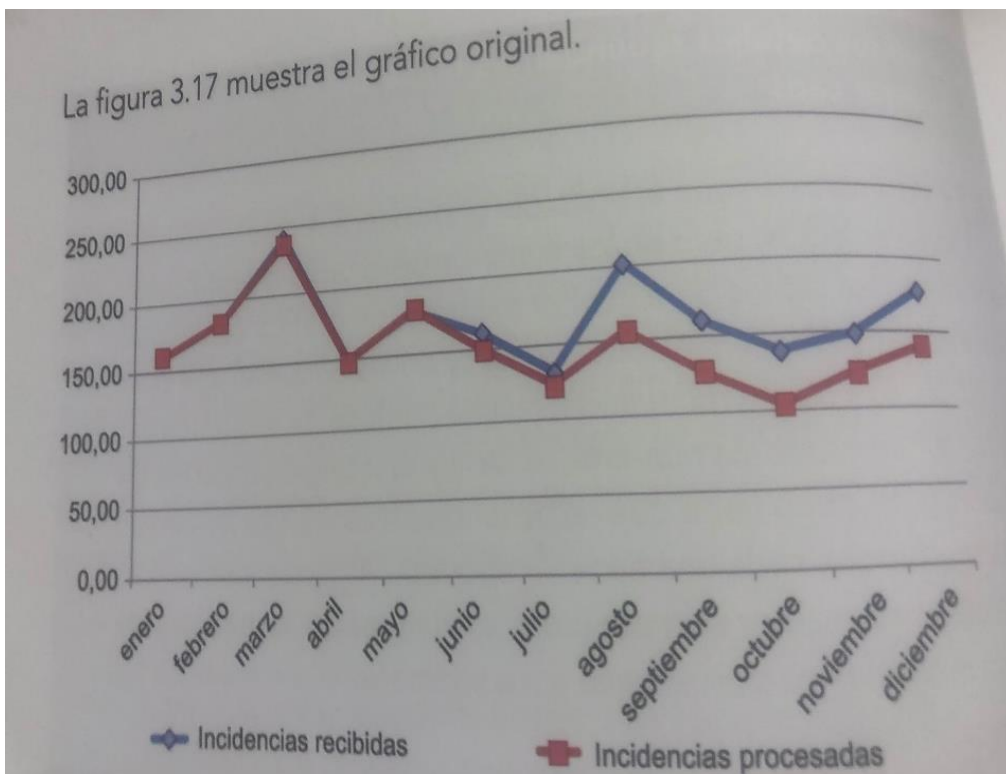


Antes y Después (Opción 1)



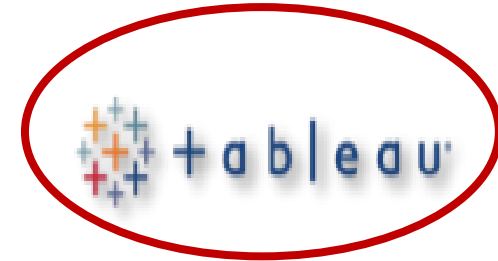
Un gráfico más orientado a una publicación científica podría incluir el análisis de las pendientes hasta que ocurre el evento y después. Por tanto, habría que incluir si la pendiente es significativa, e incluso si hay diferencias entre la primera pendiente (antes del evento) y la última

Antes y Después (Opción 2)



Un gráfico más orientado a una publicación científica podría incluir el análisis de las pendientes hasta que ocurre el evento y después. Por tanto, se podría incluir si la pendiente es significativa, e incluso si hay diferencias entre la primera pendiente (antes del evento) y la última

Herramientas ...



ADOBE ILLUSTRATOR

Lecturas sugeridas

- ***The Functional Art***, Alberto Cairo, New Riders (2012)
- ***The Truthful Art***, Alberto Cairo, New Riders (2016)
- ***How Charts Lie***, Alberto Cairo, W.W. Norton & Company (2019)
- ***The Wall Street Journal Guide to Information Graphics: The Dos and Don'ts of Presenting Data, Facts, and Figures***, Dona M. Wong, W. W. Norton & Company (2010)
- ***Information Dashboard Design: Displaying Data for At-a-Glance Monitoring***, Stephen Few, O'Reilly Media (2013)
- ***Show Me the Numbers: Designing Tables and Graphs to Enlighten***, Stephen Few, Analytics Press (2004)
- ***Now You See It***, Stephen Few, Analytics Press (2009)
- ***Storytelling with Data***, Cole Nussbaumer, Wiley (2015)
- ***Visualize This: The Flowing Data Guide to Design, Visualization, and Statistics***, Nathan Yau, Wiley (2011)
- ***The Visual Display of Quantitative Information***, Edward Tufte, Graphics Press, 2nd Edition (2001)

Lecturas sugeridas

- [Viz of the Day](#) by Tableau Software
- [The Functional Art](#) by Alberto Cairo
- [Perceptual Edge](#) by Stephen Few
- [Flowing Data](#) by Nathan Yau
- [Storytelling with Data](#) by Cole Nussbaumer
- [VizWiz](#) by Andy Kriebel
- [DataPlusScience](#) by Jeffrey Shaffer
- [Data Revelations](#) by Steve Wexler
- [Dear-Data.com](#) by Giorgia Lupi and Stefanie Posavec
- [Dear-Data-Two.com](#) by Jeffrey Shaffer and Andy Kriebel
- [Makeover Monday Project](#) by Andy Kriebel and Eva Murray

