

# Fastbook 01

## Visual Analytics

Nociones básicas de visualización



## 01. Nociones básicas de visualización



En este primer fastbook iremos viendo los términos y principios básicos sobre las técnicas utilizadas para **presentar datos y contar una historia**. El análisis descriptivo es vital en cualquier proyecto para obtener el conocimiento total sobre cualquier conjunto de datos. El objetivo de esta primera toma de contacto es entender las **ideas y métodos** detrás de una visualización efectiva.

*Autor: Daniel Pegalajar Luque*

**Introducción**

**¿Por qué es tan importante mirar los datos?**

**Gráficos engañosos o confusos**

**Storytelling**

**Conclusiones**

**Bibliografía**

**Test final**

# Introducción

X Edix Educación

---

**Siempre debes profundizar en los datos.**

Los gráficos te permiten explorar y aprender la estructura interna de la información que has reunido, mientras que las buenas visualizaciones te ayudarán a **comunicar mejor tus ideas y aprendizajes a otras personas**. No todo lo bueno termina aquí, desarrollar gráficos efectivos de tus propios datos es el mejor camino para educar ese ‘buen ojo’ que te ayudará a entender y leer gráficos —buenos o malos— realizados por otros, los presentados en artículos, papers científicos, en *slides* de negocio, o los utilizados en política o informativos de noticias. ¡En esta asignatura aprenderás cómo hacerlo!

**La importancia de conocer tus datos**

---

**Algunas visualizaciones son mejores que otras.**

---

Este mantra es algo que iremos repitiendo varias veces durante el transcurso de la asignatura. Es tentador simplificarlo todo a un ‘qué funciona y qué no’, pero por desgracia el proceso para hacer un gráfico bueno y útil no puede ser reducido a una lista de reglas simples que se puedan seguir en cualquier circunstancia. **Los gráficos que realices estarán dirigidos a alguien en particular**. No es lo mismo una imagen dirigida a un grupo de expertos, difícilmente interpretable o entendible fuera de ese círculo, que un gráfico dirigido a un grupo de estudiantes.

Trabajando con R o Python y sus librerías gráficas, tenemos buena parte de la labor hecha. Estos frameworks facilitan muchas de las virtudes gráficas y permiten centrarse únicamente en el gráfico a utilizar. De todas formas, conocer algunas de estas virtudes y cómo funcionan es el motivo de esta asignatura. Las herramientas que utilices te ayudarán a emplear unos estándares bien definidos, pero no pueden hacer la elección correcta por ti. Es por este motivo que importa (y mucho) cultivar un buen juicio sobre las visualizaciones.

---

## La primera pregunta es: ¿por qué molestarse en mirar un gráfico en lugar de confiar en mirar las tablas directamente o las estadísticas resumen?

Más adelante discutiremos algunos ejemplos sencillos para entrar en calor, como esas malas prácticas en visualizaciones que tienden a engañar e intentar comunicar mensajes de forma sesgada (seguro que te suena algún gráfico visto en los medios con polémica).

---

En resumen, los buenos **métodos de visualización son herramientas valiosas** que debemos usar en el proceso de exploración, entendimiento y explicación de cualquier conjunto de datos.

Pero estas no son herramientas mágicas que permiten ver el mundo como es en realidad. No evitarán que consigas engañar a alguien (si ese es tu propósito) y tampoco evitarán que consigas conclusiones sesgadas por ti mismo.

# ¿Por qué es tan importante mirar los datos?

X Edix Educación

“...cálculos y gráficos. Deben estudiarse ambos tipos de resultados; cada uno contribuirá a la comprensión total del problema.”

- F.J. Anscombe (1973).

Esta frase ha resonado desde entonces y sigue siendo de obligado cumplimiento en la actualidad.

Puede ser difícil demostrar la importancia de la visualización de los datos. Algunos tienen la impresión de que simplemente son ‘imágenes bonitas’, mientras que otros se centran directamente en el análisis estadístico sin importar nada más. Vamos a ver un ejemplo muy claro de lo peligroso que es ignorar la importancia de un buen análisis descriptivo.

**Alberto Cairo** es firme defensor del concepto: nunca confíes al 100% en los resultados estadísticos únicamente; visualiza siempre tus datos.

Vamos a verlo en R.

```
# Datasaurus _____  
  
# https://cran.r-project.org/web/packages/datasauRus/vignettes/Datasaurus.html  
# install.packages("datasauRus")  
  
# Caramos librerias necesarias  
library(datasauRus)  
library(dplyr)  
  
# Vamos a ver un resumen estadistico de los datos  
datasaurus_dozen %>%  
  group_by(dataset) %>%  
    summarise(media_x = mean(x),  
             media_y = mean(y),  
             desv_est_x = sd(x),  
             desv_est_y = sd(y),  
             correla_x_y = cor(x, y))
```

Que nos da el siguiente resultado:

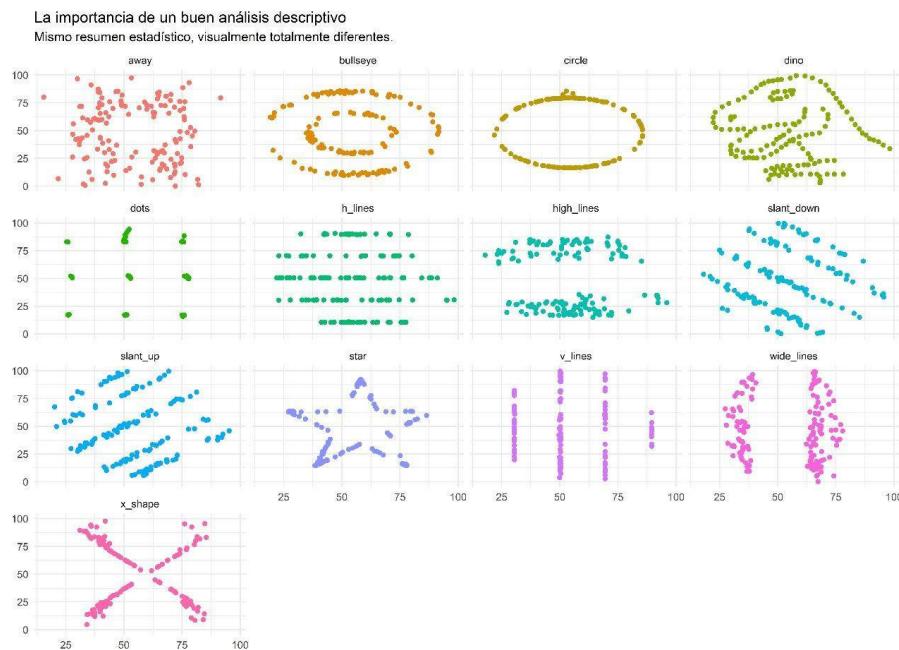
```
# A tibble: 13 x 6
  dataset    media_x media_y desv_est_x desv_est_y correla_x_y
  <chr>      <dbl>   <dbl>     <dbl>     <dbl>       <dbl>
1 away        54.3    47.8     16.8     26.9      -0.0641
2 bullseye    54.3    47.8     16.8     26.9      -0.0686
3 circle      54.3    47.8     16.8     26.9      -0.0683
4 dino         54.3    47.8     16.8     26.9      -0.0645
5 dots         54.3    47.8     16.8     26.9      -0.0603
6 h_lines      54.3    47.8     16.8     26.9      -0.0617
7 high_lines   54.3    47.8     16.8     26.9      -0.0685
8 slant_down   54.3    47.8     16.8     26.9      -0.0690
9 slant_up     54.3    47.8     16.8     26.9      -0.0686
10 star         54.3    47.8     16.8     26.9      -0.0630
11 v_lines      54.3    47.8     16.8     26.9      -0.0694
12 wide_lines   54.3    47.8     16.8     26.9      -0.0666
13 x_shape      54.3    47.8     16.8     26.9      -0.0656
```

Podemos ver que **cada uno de los datasets ofrece información muy similar**. Una conclusión peligrosa a la vista de este análisis sería afirmar que los datos son muy similares (o prácticamente iguales) y no ir más allá. Es aquí donde interviene la **visualización** de los datos.

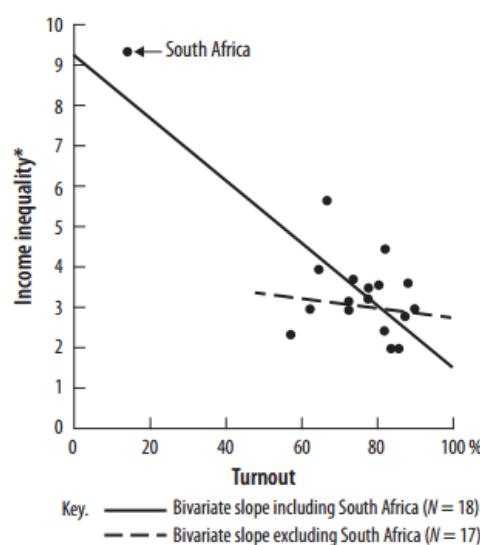
```
# Vamos a pintar los dataset
library(ggplot2)

ggplot(datasaurus_dozen, aes(x = x, y = y, colour = dataset)) +
  geom_point() +
  theme_minimal() +
  theme(legend.position = "none") +
  facet_wrap(~dataset, ncol = 4) +
  labs(x = "", y = "", title = "La importancia de un buen análisis descriptivo",
       subtitle = "Resumen estadístico igual, pero visualmente totalmente diferentes.")
```

Puedes obtener más detalle del proceso para crear este tipo de gráficos mediante perturbaciones que buscan ‘formas’, preservando ciertos estadísticos (como la media o la desviación típica), [aquí](#).



En el artículo de **Jackman, R. M. (1980)** *The impact of outliers on income inequality*. *American Sociological Review*, 45, 344–47 se recoge una crítica hacia el paper original de **Hewitt** (1977), donde se argumentaba el descubrimiento de una fuerte relación entre la participación electoral y la desigualdad de ingresos, basado en el análisis cuantitativo de 18 países. Cuando esta relación se dibujaba mediante un **scatterplot**, se obtenía lo que ves en la imagen.



---

**Scatterplot**, también conocido como diagrama de dispersión, es un tipo de visualización que se suele emplear para representar la relación entre dos variables, como altura y peso, edad y renta o tiempo y desempleo. ¡Estos diagramas son posiblemente la herramienta de visualización más utilizada en las Ciencias Sociales, así que ve habituándote a ellos!

Como decíamos, cuando se dibujaba la relación en este scatterplot, salta a la vista que todo el peso de esas conclusiones lo tiene la inclusión en el análisis de Sudáfrica, claramente un **dato atípico o outlier** en este estudio.

---

Los **diagramas de dispersión** permiten visualizar la asociación entre dos variables. Esa asociación la podemos cuantificar con la correlación, que puede ir de -1 a +1, en la que 0 indica la ausencia de relación entre las variables en cuestión. Un -1 significa asociación negativa o indirecta perfecta, y en este tipo de diagramas se podría observar si la recta que mejor representa a esa nube de puntos (la **recta de regresión**) tiene una inclinación de 45 grados hacia abajo.

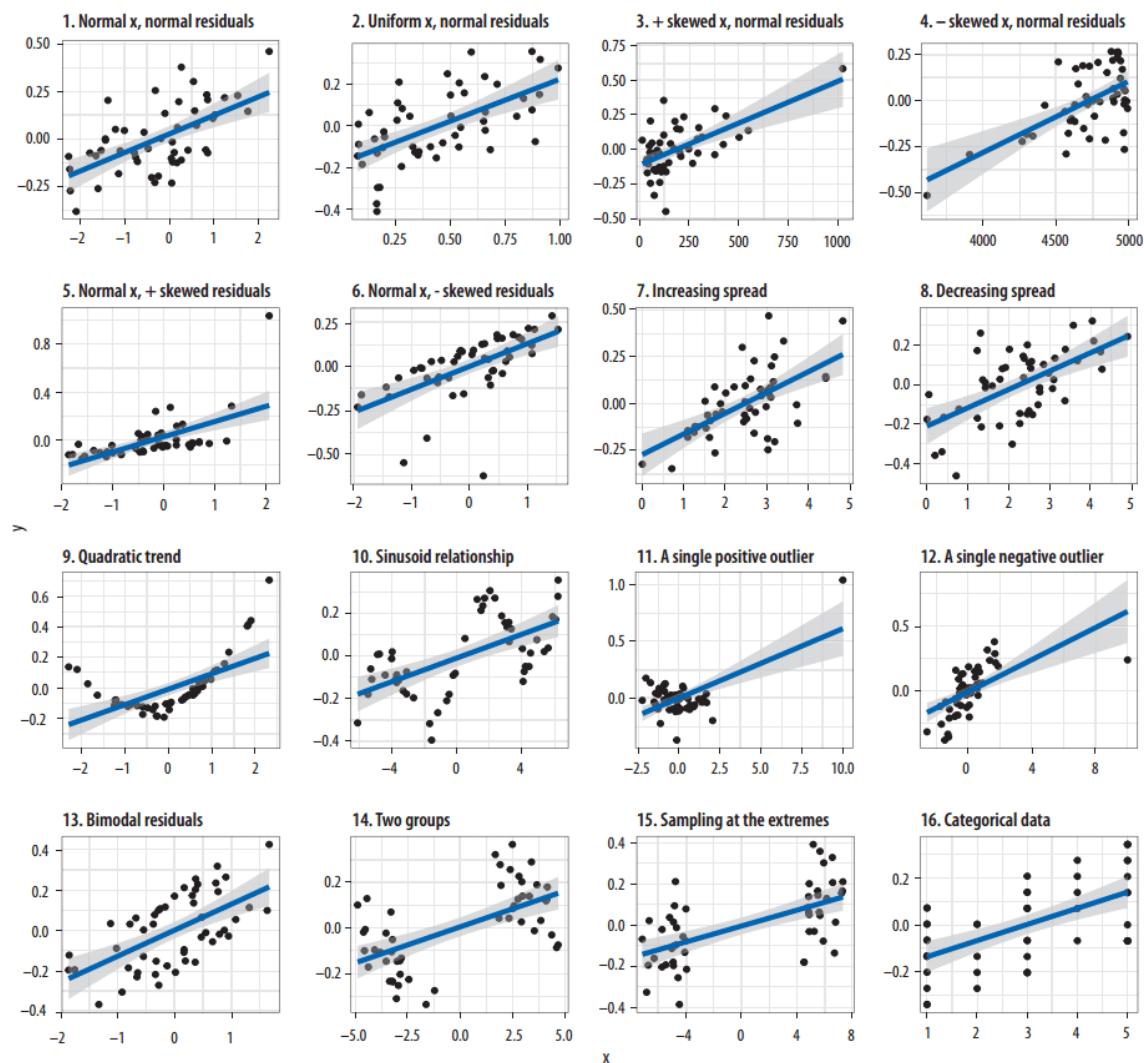
En este ejemplo, la correlación que representa la recta sólida es muy fuerte y lleva a la conclusión de que cuanta más desigualdad tenga un país, menor será la participación electoral en sus elecciones.

Si en el artículo original se hubiese realizado un análisis más exhaustivo o se hubiese utilizado un análisis de sensibilidad se habría detectado este problema. **Un simple gráfico como el de arriba realiza todo el trabajo directamente.**

La segunda línea punteada muestra la nueva relación entre los 17 países restantes, si se eliminase Sudáfrica. Se puede apreciar cómo la recta de regresión prácticamente se vuelve horizontal, reflejando que la asociación entre ambas variables es prácticamente nula (su coeficiente de correlación sería cercano a 0).

Otro ejemplo es el realizado por Jan Vanhove (2016) del que tenemos disponible su código publicado en GitHub. Podéis echarle un vistazo en [este enlace](#).

En él, ilustra el **peligro de confiar en la obtención de aprendizajes y conclusiones solo con un coeficiente de correlación**. En su artículo, podemos observar la demostración de la importancia de mirar datos y estadísticas a la vez.



En el gráfico superior podemos observar 16 casos diferentes de conjuntos de datos y todos ellos comparten la misma correlación: 0,6.

Este es un ejemplo similar al [cuarteto de Anscombe](#), donde cada panel muestra la asociación entre dos variables simuladas con un buen grado de asociación. Lo interesante es que cada grupo ha sido creado con un proceso distinto: comienzan con datos provenientes de distribuciones normales y, a partir de ahí, **tenemos outliers, dos grupos de datos o incluso variables categóricas en forma de escalón**; todas ellas muestran el mismo grado de asociación.

---

## ¿Vas entendiendo la importancia de visualizar tus datos?

El siguiente tema a debatir es la importancia y dificultad de elegir buenas visualizaciones. Este punto puede ser clave, ya que a veces la mala lección de un gráfico puede conducir a la audiencia a conclusiones sesgadas o erróneas.

# Gráficos engañosos o confusos

X Edix Educación

---

Llegamos a terreno pantanoso, es hora de hablar de **gráficos distorsionados o misleading graphs**, en inglés. Cada día somos bombardeados por noticias que, con bastante frecuencia, utilizan gráficos o visualizaciones para ilustrar unos datos determinados. En algunas ocasiones puntuales, esos gráficos son engañosos o se han construido erróneamente (no vamos a discutir su posible intencionalidad...).

---

**La pregunta es: ¿eres capaz de detectar el engaño? ¿Cuánta población tiene las nociones para identificar si lo que están viendo en un gráfico es engañoso o está exagerado para guiar las conclusiones que deben extraer? Vamos a profundizar en esto.**

Un gráfico puede ser distorsionado de muchas formas: complicándolo en exceso, por una pobre construcción, construyéndolo con muy poca o nula precisión a la hora de mostrar sus características...

Uno de los primeros autores en hablar sobre este tema fue **Darrell Huff** que en un libro maravilloso, publicado en 1954 bajo el título *How to Lie with Statistics*, profundiza en este tipo de errores. Para resumir diremos que esta obra ilustra el mal uso de la estadística y los errores más comunes de visualización o cómo dichos errores pueden conducir a conclusiones incorrectas. Un best-seller que recomiendo encarecidamente para aquel que se sienta interesado por este mundillo.

## **Lista de errores y métodos más habituales en visualizaciones engañosos**

A continuación, describimos algunos de los numerosos escenarios que tendrían que ponernos en alerta a la hora de interpretar un gráfico.

- **Uso excesivo de imágenes**

Utilizar visualizaciones donde no son necesarias puede conducir a interpretaciones o posibles confusiones. Las visualizaciones son muy importantes, pero un gráfico no siempre transmite mejor la información que una tabla sencilla. Piensa antes de dibujar.

- **Etiquetas sesgadas**

Las etiquetas en los ejes, títulos de un gráfico o pie de foto pueden estar siendo utilizadas inapropiadamente para predisponer al lector.

- **Tendencias artificiales**

En ocasiones, se recurre al uso de líneas que trazan una tendencia a través de los datos (aunque no estén correlacionados), induciendo al lector a creer que puede existir una relación cuando no la hay.

Esto solo tiene dos explicaciones posibles: o se está intentando engañar al lector o se está incurriendo en el error de una correlación espuria.

- **Gráfico de tartas o pie charts**

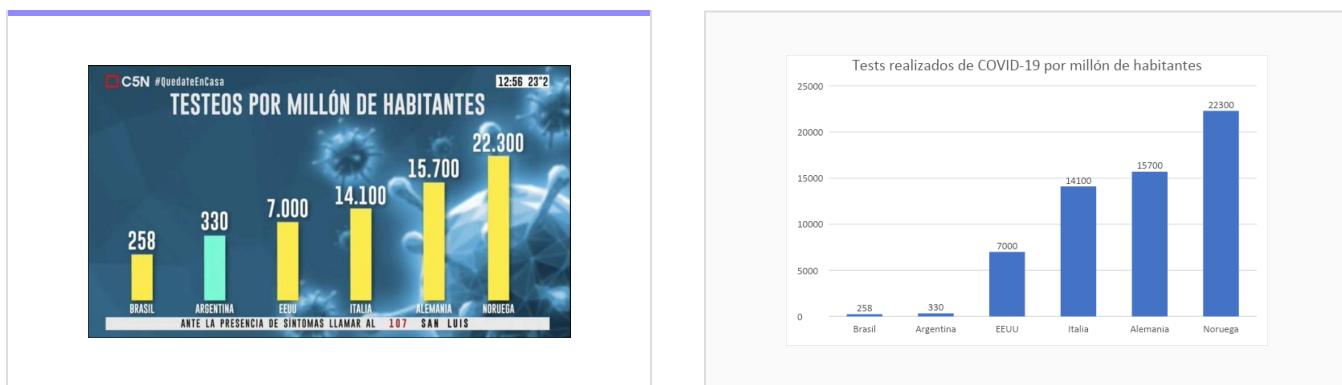
Possiblemente el gráfico más denostado de todos los que veréis en este curso. Y no será por falta de motivos:

- En ocasiones, algunos sectores de este tipo de gráficos son tan finos que es muy difícil interpretar los resultados quedando claro solo aquellos trozos con mayor peso.
- Además, si se comparan varias tartas, la diferencia de tamaños puede ser engañosa, ya que el lector no puede leer con precisión el área comparativa de ambos círculos.
- El uso de porcentajes como etiquetas puede inducir a error cuando la muestra que se está representando es pequeña.
- Si hablamos de utilizar el 3D o añadir inclinación, la cosa empeora aún más. Este efecto distorsiona tanto la perspectiva que cualquiera puede verse confundido por la información que se trata de representar.

- **Escala inapropiada**

Se suele detectar este tipo de sesgo en el uso de **gráficos de barras o pictogramas** (lo mismo, pero ilustrando la barra con algún ícono representativo).

Como puedes ver en esta tarjeta en la que mostramos el gráfico original y una versión más fiel a los datos realizada por mí, el canal C5N de Argentina manipuló el eje Y del gráfico de barras para ocultar el terrible número de tests COVID-19 realizados a la población, tal y como se explica en [Reddit](#).



Queda claro que la foto es totalmente diferente cuando la escala está bien representada y que Argentina no está ni por asomo cerca de EE. UU. en cuanto a los test realizados en su población.

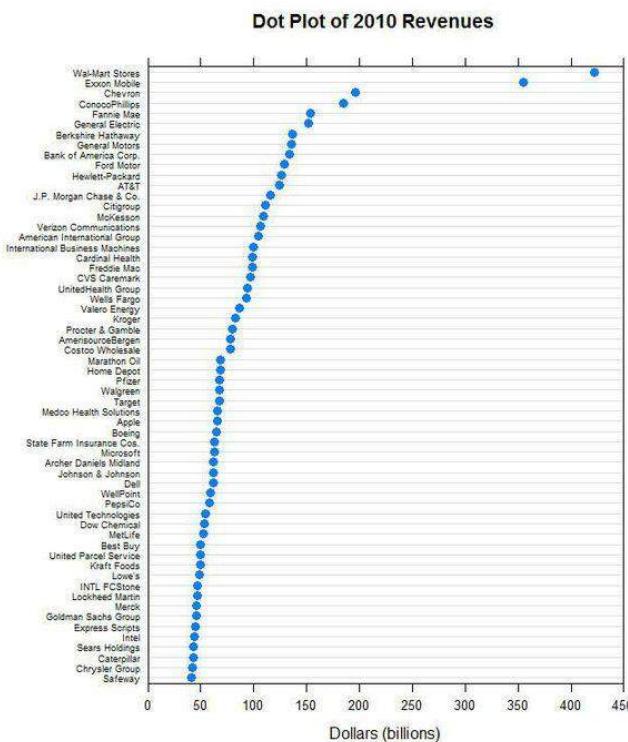
## ¿A quién le gusta escuchar malas noticias? ¿Ves alguna intención detrás de este error?

- Escala logarítmica

Suele haber dos razones para utilizar este tipo de escala: disminuir la asimetría excesiva de valores extremos, es decir, aquellos casos en los que se tiene una minoría de puntos con valores muchísimo más grandes que la mayoría de los datos. El segundo uso es mostrar un porcentaje de cambio o factores multiplicativos.

Vamos a mostrar un ejemplo para ilustrar el uso:

Esta figura utiliza un **dot plot** o **diagrama de puntos** del Top 50 de 2010 de las compañías que forman parte del grupo Fortune 500 (las empresas americanas que más *revenue/profit* obtienen en el año fiscal). La elección de este gráfico sobre un gráfico de barras es que **permite mostrar un gran número de casos sin saturar la visualización**. Paciencia, porque aprenderemos más sobre ellos en fastbook posteriores.

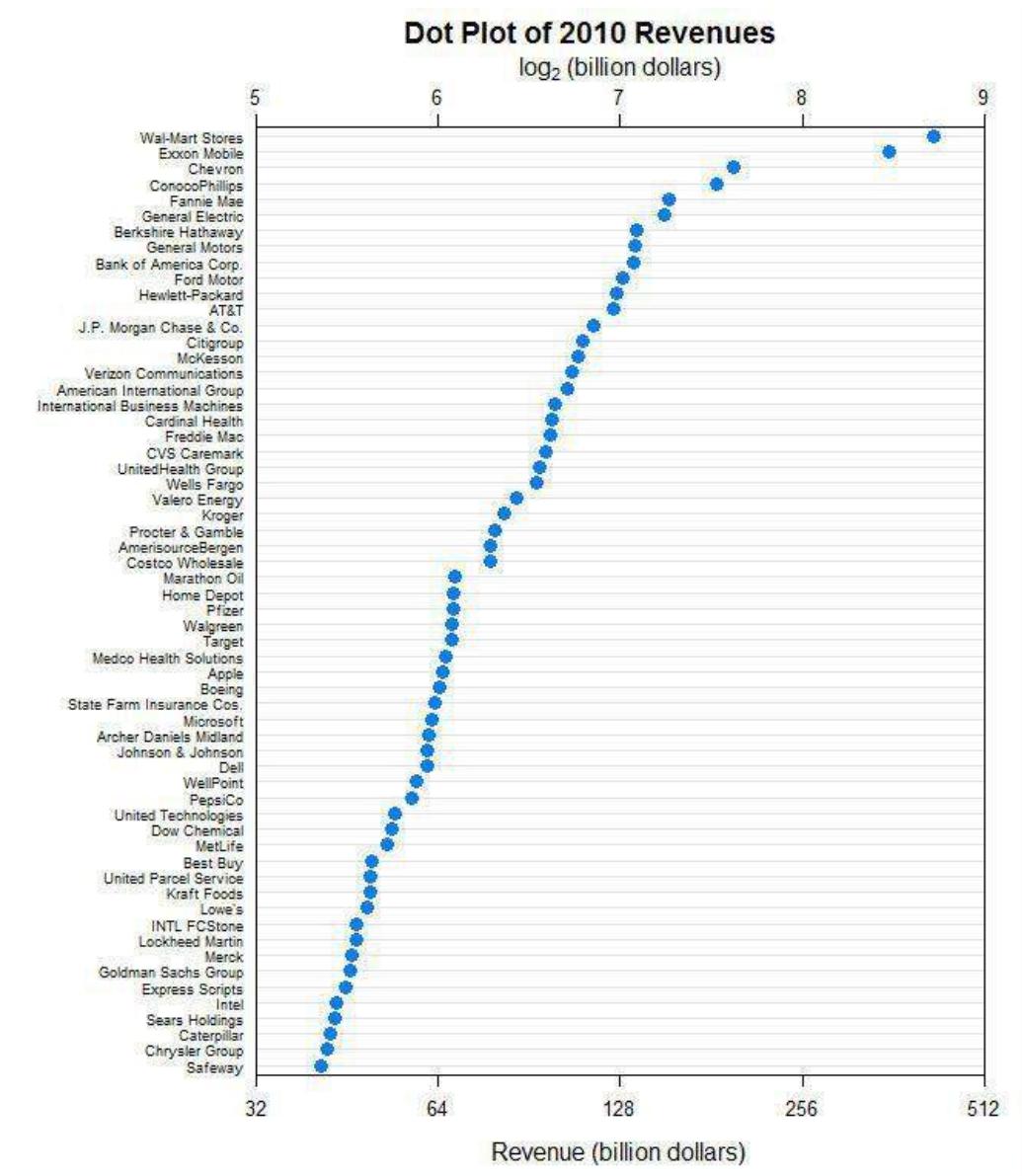


Fuente: [Forbes.com](http://Forbes.com)

Walmart y Exxon tienen un revenue mucho más elevado que el resto de compañías. Como resultado, estas diferencias hacen que las otras compañías se ‘apelotonen’ a la izquierda, algo que complica la visualización de las diferencias entre la mayoría de empresas. Sin embargo, en la

figura de abajo, tenemos los mismos datos, pero se representan utilizando la escala logarítmica.

¿Notas alguna diferencia ahora?



Fuente: [Forbes.com](#)

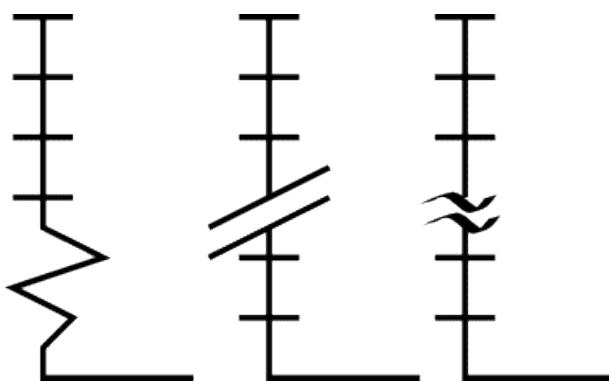
- Gráficos truncados (en inglés se conocen como *torn graph*)

Se caracterizan por tener un eje y que no comienzan en el 0. Estas visualizaciones suelen tener el efecto de **realzar una diferencia que realmente no es tan importante**. A veces, este recurso se suele utilizar para observar mejor las diferencias o ahorrar espacio en la visualización. Su uso **normalmente se desaconseja**.

Algunos software como Microsoft Excel tienden a truncar el eje Y automáticamente, generando este tipo de problema, **¿lo sabías?**



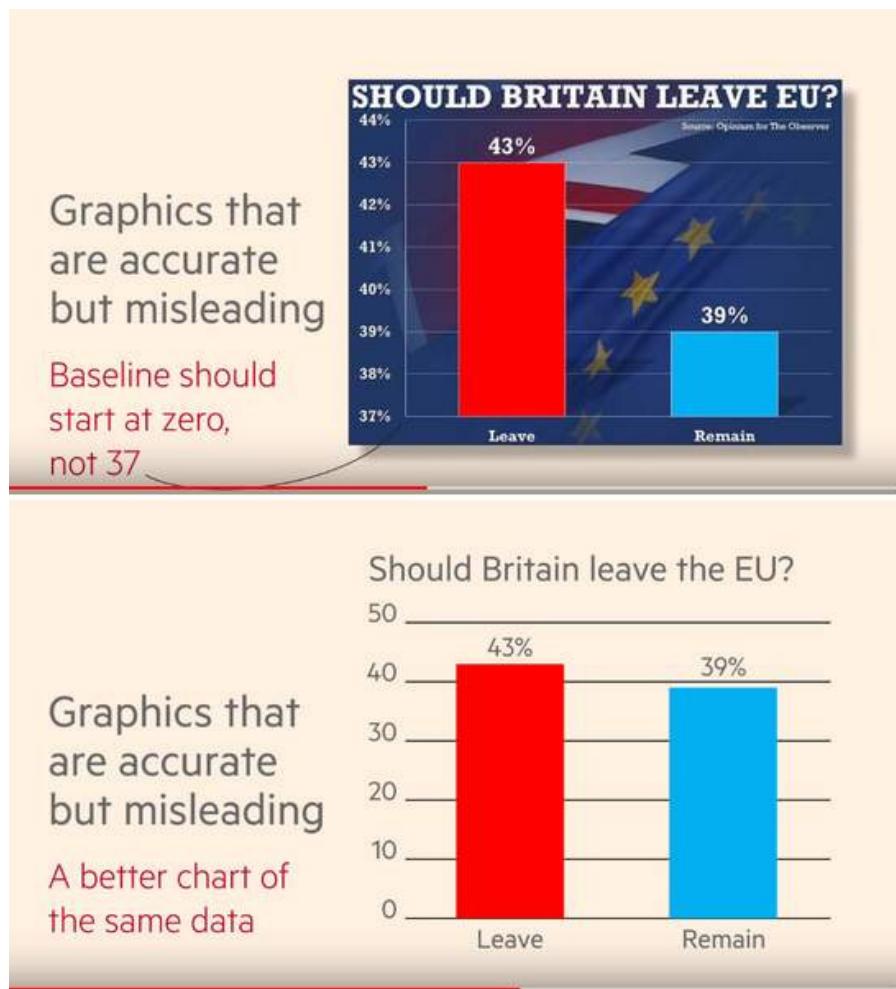
**Ambos gráficos contienen la misma información pero, ¿cuál te llama más la atención? ¿Crees que las diferencias son realmente reseñables?**



Fuente: Wikipedia.

Normalmente hay formas de señalizar que el eje Y ha sido truncado, pero no es habitual que la mayoría de la audiencia entienda y sepa interpretar las visualizaciones al 100%.

Este tipo de 'error' es muy habitual verlo en los medios de comunicación para exagerar diferencias. Por ejemplo, en este ejemplo sobre el Brexit.



Fuente: Wikipedia.

---

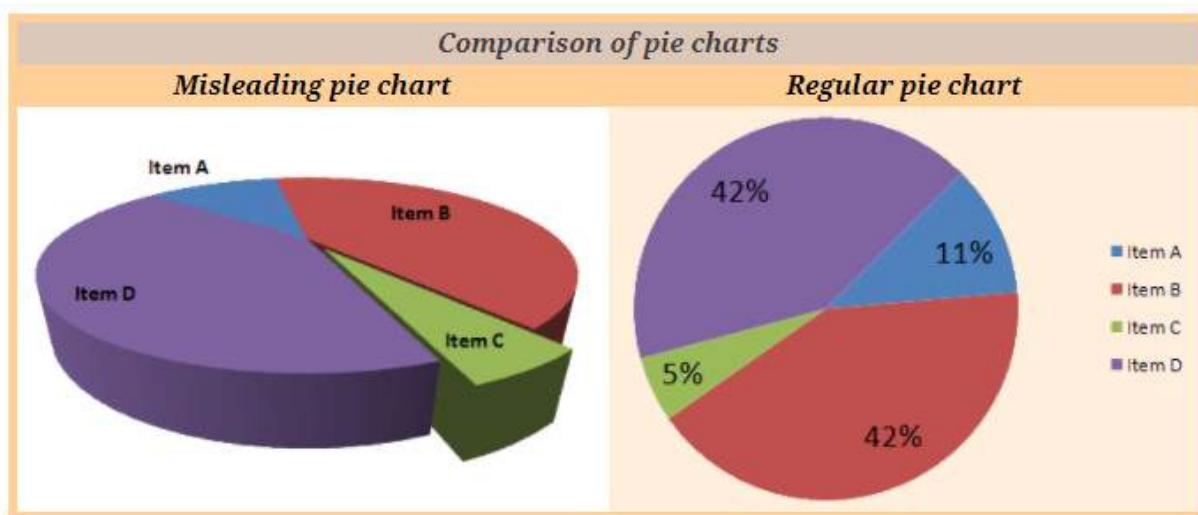
¿Qué gráfico crees que llamará más la atención de la audiencia? ¿Se le puede dar un uso indebido a alguna de las imágenes?

## Gráficos de tarta en 3D

Vamos a detenernos un poco en esta perspectiva porque ha sido introducida en multitud de programas, posiblemente la suite de Microsoft sea la más destacada. Todo el mundo que ha tenido que pintar alguna información se ha visto tentado a utilizar este gráfico, normalmente porque el efecto estético final lo merece. Sin embargo, **la tercera dimensión no hace milagros y no consigue mejorar la legibilidad de los datos**; por el contrario, este tipo de gráficos es difícil de interpretar por la distorsión introducida al usar este efecto.

**Cualquier uso de dimensiones que no sea aprovechado para mostrar información de interés está tremadamente desaconsejado a la hora de realizar visualizaciones** (este consejo no aplica únicamente a los gráficos de tartas). Veamos un ejemplo.

En la imagen podemos apreciar que el ítem C parece ser muy similar al ítem A en cuanto a tamaño; sin embargo, al eliminar el efecto 3D salta a la vista que el ítem A es más del doble que el C. Esto sería un gráfico engañoso.



---

En un gráfico 3D, los trozos que están más cerca del lector parecen ser más grandes que aquellos más alejados, debido al ángulo o perspectiva con el que se presentan. Este efecto resta capacidad de juicio al lector para evaluar la magnitud de los ítems representados.

---

Para cerrar este punto, unas palabras de **Edward Tufte**, un gran estadístico norteamericano que odiaba profundamente el uso abusivo de este tipo de gráficos. En su libro *The Visual Display of Quantitative Information*, dijo:

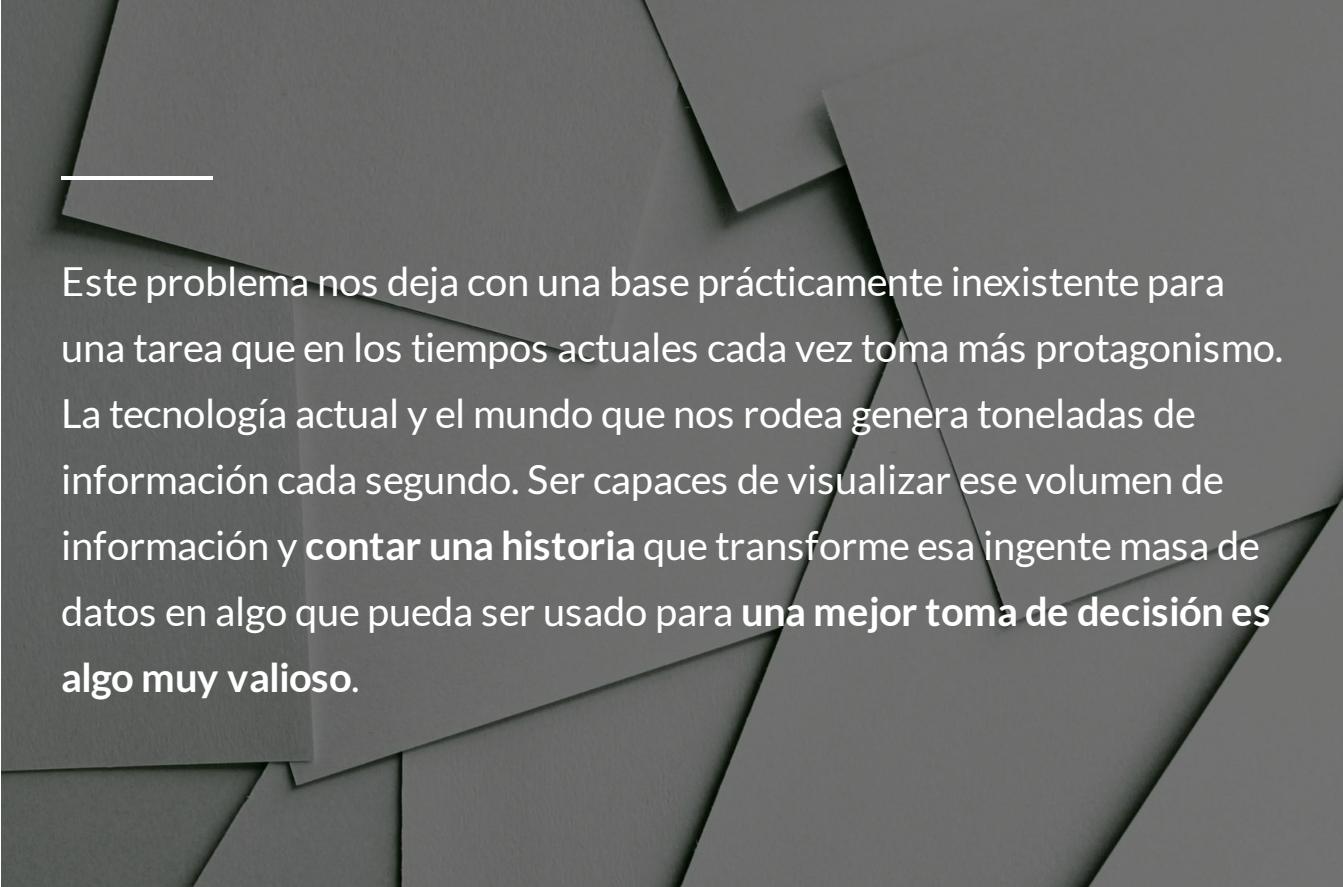
“Las tablas son siempre preferibles a los gráficos para conjuntos de datos pequeños. Una tabla es casi siempre la mejor elección frente a un estúpido gráfico de tartas; solo hay una cosa peor que este gráfico, un montón de ellos, ya que entonces el lector debe comparar las cantidades situadas en un desorden espacial entre ellos y dentro de ellos. Dada su baja densidad de datos y la dificultad para ordenar los números a lo largo de una dimensión visual, los gráficos de tarta nunca deberían utilizarse.”

# Storytelling

X Edix Educación

---

Asumamos una cosa: a lo largo de nuestra trayectoria educativa nos enseñan Lengua Castellana y Matemáticas, principalmente. La primera nos ayuda a entender y saber estructurar nuestras frases para comunicarnos bien (y a no cometer faltas de ortografía); la segunda nos (intenta) enseñar el sentido de los números y cómo definen el mundo que nos rodea. Rara vez se entrelazan: **nadie nos enseña a contar historias a través de los números.**



Este problema nos deja con una base prácticamente inexistente para una tarea que en los tiempos actuales cada vez toma más protagonismo. La tecnología actual y el mundo que nos rodea genera toneladas de información cada segundo. Ser capaces de visualizar ese volumen de información y **contar una historia** que transforme esa ingente masa de datos en algo que pueda ser usado para **una mejor toma de decisión es algo muy valioso.**

Hoy en día cualquiera puede utilizar una herramienta (la mayoría hemos usado Excel, ¿no?) para pintar algún tipo de gráfico. Pero justo ahí está el peligro... ¡cualquiera puede crear un gráfico! Sin la formación adecuada y un objetivo claro, nuestras mejores intenciones y esfuerzos pueden llevarnos a realizar decisiones realmente malas: uso del 3D, paletas de colores sin sentido o algo incluso peor, los gráficos de tartas.

---

**Quédate con lo siguiente: hay una historia en tus datos. Las herramientas que aprendas a utilizar no saben qué historia es (por este principal motivo, estamos muy lejos de reemplazar a los analistas o data scientists humanos por meras máquinas). Esa responsabilidad recae sobre ti, el analista o comunicador de la información.**

Es muy sencillo obtener algo de información, introducirla en Excel y crear un gráfico. Para muchos el proceso de visualización termina ahí. Esto puede ocasionar que una historia perfecta sea un completo bodrio o, incluso peor, que no sea posible de entender.

---

**No abandones tus datos y las historias que esconden, ¡aprende a extraerlas y contarlas debidamente!**

---

A lo largo de los siguientes temas aprenderás a desarrollar y familiarizarte con gran cantidad de tipos de gráficos. Aprenderás sus secretos, sus pros y contras, así como dónde resalta su mayor utilidad. Con estos conocimientos, adquirirás una de las herramientas más valiosas para el **storytelling**. Destacar en este campo conlleva mucha práctica, pero merece la pena. Saber contar historias te proporciona una habilidad valiosa para conseguir el éxito en cualquier posición o rol que te puedas imaginar. Así que, despierta esa curiosidad y gana práctica con cualquier tipo de visualización; al fin y al cabo, **ninguna herramienta será capaz de destapar la historia que esconden tus datos**.

## Claves para contar una historia con datos

Existen muchos gurús en este campo, pero mi preferida es **Cole Nussbaumer** (os recomiendo sus obras, que podéis encontrar referenciadas en la bibliografía de este fastbook).

Cole resume las claves de **una historia efectiva en seis lecciones** que debes interiorizar cada vez que quieras extraer la historia que esconde un conjunto de datos. Vamos a verlas.

1

### Entiende el contexto

Antes de sumergirte en la visualización, debes realizarte una serie de preguntas a las que deberías ser capaz de dar una respuesta concisa:



¿A quién va a ir dirigido tu mensaje?



¿Qué necesitas saber o hacer?

Es importante entender el contexto situacional al que te enfrentas, esto incluye la audiencia, cómo comunicar el mensaje o incluso el tono a utilizar. Este primer paso reduce los posibles bloqueos que pueda sufrir tu historia y es el primer paso al éxito cuando se trata de crear contenido visual.

2

### Elección de la visualización

O cuál es la mejor forma de mostrar la información que pretendes comunicar. Con este fin, analizaremos en futuros fastbooks toda la gama de gráficos disponibles para comunicar una historia. Es importante que sepas diferenciar cuándo ciertos gráficos deben ser evitados. Para ello, usaremos ejemplos reales que ayuden a reafirmar estas palabras.

3

### El desorden es tu peor enemigo

Cuando generas un gráfico, ten siempre en la mente el impacto que tendrá en la audiencia. Esta debe ser capaz de entender la visualización de un vistazo y, para lograr esto, tú **debes ser capaz de eliminar todos aquellos elementos que son innecesarios**, los que solo ‘ensucian’ la imagen. Este es uno de los puntos clave; por ello utilizaremos ejemplos reales que ilustren cuándo una visualización está demasiado sobrecargada y distrae la atención de los aprendizajes clave.

4

### Capta la atención de tu audiencia

A la hora de crear visualizaciones que centren la atención, tú eres el arquitecto. **Atributos como el tamaño, color y posición** son dimensiones útiles y estratégicas para dirigir la mirada de cualquier persona que vea tu imagen. Debes ser capaz de centrar la atención en la información que quieras comunicar y en la forma de hacerlo.

5

### Piensa como un diseñador

**Forma y función**, nociones básicas en el mundo del diseño que también son aplicables en este mundo. En primer lugar, debemos pensar **qué es lo que queremos que nuestra audiencia pueda hacer con los datos** (función) y **crear la visualización** (forma) que permitirá esto con facilidad.

6

### Cuenta tu historia

**Una historia debe ser clara desde el principio.** Si has seguido los cinco principios anteriores y lo has combinado, estás listo para pasar de mostrar simplemente datos a contar una historia extraída de tus datos. Evidentemente, la práctica hace al maestro. Además necesitarás otras nociones más destinadas a construir presentaciones de trabajo efectivas, a mantener un tono atractivo y captar a tu audiencia desde el principio. La práctica vale más que la gramática: lánzate a la piscina y comienza (oblígate) a practicar contando una historia.

# Conclusiones

X Edix Educación

---

En este primer fastbook, hemos introducido la importancia que tiene la visualización en el mundo de los datos. Hemos comenzado ilustrando **cómo graficar tus datos puede hacer aflorar conclusiones que no obtendrías mirando solo el resumen estadístico de estos**. Además, hemos aprendido con varios ejemplos que no es algo que ocurra en pocas ocasiones, ¡ni algunos papers científicos se salvan de caer en este error!

Además, hemos atravesado la galería del horror y el engaño, enumerando un listado de los recursos que suelen generar visualizaciones engañosas y que, en muchas ocasiones, son utilizados para generar conclusiones erróneas y guiar la opinión de la audiencia. ¡Asegúrate de entenderlas perfectamente y aprende a identificarlas!

Finalmente, hemos dado las primeras nociones de storytelling y lo importante que es la práctica para dominar este campo tan útil en la actualidad.

# Bibliografía

X Edix Educación

---

- [\*Same Stats, Different Graphs: Generating Datasets with Varied Appearance and Identical Statistics through Simulated Annealing\*](#), por **Justin Matejka y George Fitzmaurice**.
- [\*How to lie with statistics\*](#), por **Darrell Huff**.
- [\*The Visual Display of Quantitative Information\*](#), por **Edward R. Tufte**.
- [\*When Should I Use Logarithmic Scales in My Charts and Graphs?\*](#), por **Naomi Robbins**.
- [\*Las mejores visualizaciones cada día\*](#), de Reddit.
- [\*Storytelling with Data\*](#), por **Cole Nussbaumer**.

Lección 7 de 7

## Test final

**X** Edix Educación

---

Pon a prueba lo aprendido en este fastbook.

---

*Pregunta*

01/05

¿Qué fase es indispensable en cualquier análisis?

---

- Análisis estadístico de los datos.
- Análisis descriptivo de los datos.
- Las dos anteriores son correctas.

Pregunta

02/05

¿Qué tipo de gráfico evalúa la relación entre dos variables cuantitativas?

---

- Un gráfico de tartas o *pie chart*.
- Un *scatterplot* o diagrama de dispersión.
- Un diagrama de barras.
- Un *dot plot*.

Pregunta

03/05

¿Qué puede ocurrir si solo confío en el análisis puro de mis datos?

---

- Puedo obtener conclusiones engañosas. Debo complementarlo con visualizaciones.
- Nada, si tengo argumentos numéricos que sostengan mis resultados.
- Nada, si justifico mis resultados con una tabla.

Pregunta

04/05

¿Cuál de las siguientes opciones puede ocasionar gráficos o visualizaciones engañosas?

---

- Escala truncada.
- Escala lineal/logarítmica.
- Etiquetas sesgadas.
- Todas son correctas.

Pregunta

05/05

¿Cuál de los siguientes puntos no es importante para contar una historia con tus datos?

---

- Elegir la visualización adecuada.
- Entender el contexto.
- Elección del software (R/Python/Excel).
- Captar la atención.



Respuestas: 1-C, 2-B, 3-A, 4-D, 5-C

edix

Creamos Digital Workers