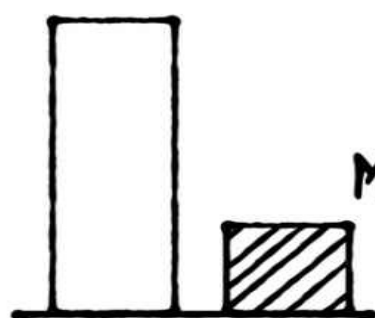
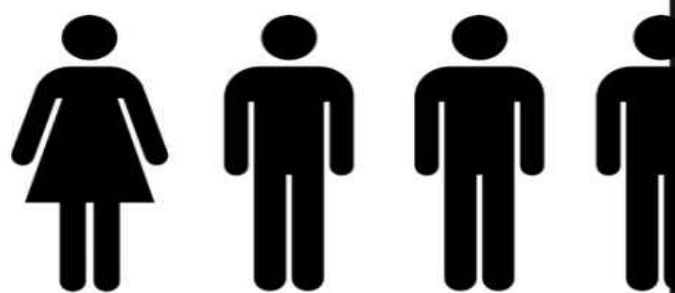


Visualización de Datos & Storytelling



Mujeres



Hombres

Ali Fenwick
Marybeth Sandell
Jose Berengueres

Editado por
Bárbara Covarrubias

Derechos de autor © 2020 Jose Berengueres

Todos los derechos reservados

Los personajes y eventos que se presentan en este libro son ficticios. Cualquier similitud con personas reales, vivas o muertas, es una coincidencia y no algo intencionado por parte del autor.

Ninguna parte de este libro puede ser reproducida ni almacenada en un sistema de recuperación, ni transmitida de cualquier forma o por cualquier medio, electrónico, o de fotocopia, grabación o de cualquier otro modo, sin el permiso expreso del editor.

ISBN-13: 9781234567890

ISBN-10: 1477123456

Diseño de la portada de: Jose Berengueres

Número de control de la Biblioteca del Congreso: 2018675309

Impreso en los Estados Unidos de América

Visualización de Datos & Storytelling

Jose Berengueres

con

Marybeth Sandell

y

Ali Fenwick

Editado

por

Bárbara Covarrubias

Traducción de la segunda edición del inglés
“Data Visualization & Storytelling”
por Angels Berengueres

Copyright texto © Jose Berengueres
Donde se indique lo contrario, obras de arte,
portada y dibujos © por Jose Berengueres

Cadena de Valor y modos de pensamiento^[1]



Índice

Índice

Prefacio

Cómo usar este libro

Capítulo 1. Las Historias, Datos y Narrativas

¿Qué son los datos?

La Sabiduría frente a los Datos

¿Cómo lograr sabiduría?

Una las palabras

Información vs. significado

Relatos e historias

Elementos de una visualización efectiva

Cuestionario: Narrativas y relatos

Cuestionario: Storytelling del cambio climático

Capítulo 2. Visualizando la Información

Visualizar género

El ajuste Gráfico-Narrativa

Gráficos Centrados en Personas

¿Sexismo en tu gráfico?

La gravedad y gráficos

Guerra de gráficos: Tortas contra Barras

Cuestionario: Hacer gráficos útiles

Capítulo 3. El Conocimiento

Marcos mentales

Visualizando la inclusión

El ganador se lleva todo

Todo o nada

El marco BRICS

Cuestionario: Resúmenes visuales

Capítulo 4. Gráficos para pensar

Utilice el ranking para crear conciencia de la situación

Proyectar en 2D

Prever con reversión a la media

[Representar la estrategia empresarial en dos dimensiones](#)

[La Matriz GAP](#)

[La Matriz de Innovación](#)

[Mapas de Wardley](#)

[Rueda de la Vida](#)

[Exploración interactiva](#)

[Cuestionario: Decisiones de apoyo](#)

[Capítulo 5. Crear gráficos que destaquen](#)

[Utilice flechas para desatar su pensamiento](#)

[Despejando gráficos](#)

[Use personas para ganarse la audiencia](#)

[Visualizando grandes diferencias](#)

[Poner el dinero global en perspectiva](#)

[¿Cuántos paneles para abastecer a todo los EE.UU.?](#)

[Storytelling del sesgo de edad con humor](#)

[Utilice la proporción áurea en todas partes](#)

[Giros](#)

[Cuestionario: Helados Suecos](#)

[Cuestionario: El calentamiento global](#)

[Cuestionario: Cuadrantes Mágicos](#)

[Prueba: Visualice la brecha](#)

[Capítulo 6 Psicología de la visualización](#)

[Contexto](#)

[Enmarcado](#)

[Gestalt](#)

[La regla del Peak-End](#)

[Capítulo 7. Sesgos y Prejuicios](#)

[El sesgo en la narrativa: ¿Una comida equilibrada?](#)

[Sesgo en la historia ¿Alemanes trabajadores?](#)

[Sesgo creencia: Las mentiras de la austeridad](#)

[Otros sesgos](#)

[El caso de incendios forestales de las Amazonas](#)

[Lista de control de diligencia debida](#)

[Cuestionario: Tweets incendiario](#)

[Reconocimiento](#)

[Referencias](#)



Científicos y diseñadores viven en mundos separados y van a escuelas diferentes, pero para hacer buenos gráficos hay que combinar las dos. Foto: Campaña I am a Mac, I am a PC.

Prefacio

Era mayo de 2018, cuando recibí un correo electrónico de [Kaggle.com](https://www.kaggle.com/), la comunidad de la Ciencia de Datos donde personas de todo el mundo compiten para ver quién es el mejor modelador de datos. Algo así como las Olimpiadas de Datos. Cada año desde el 2017, el website Kaggle ha encuestado a sus usuarios sobre todo lo relacionado con datos. Desde cuánto cobran, hasta el servicio de cloud que usan. Este año, además se decidió organizar una competencia para explicar y visualizar la encuesta misma [26]. Se ofrecían varios premios en efectivo, de los cuales a la postre se me otorgó (vía Twitter) uno de \$1500. Acto seguido a recibir el email, me puse a fisgonear las entradas que la gente había publicado ya en el sitio web de la competencia. ¡La verdad es que me impresionó la cantidad de tiempo y esfuerzo que la comunidad había invertido en el análisis, había entradas con más de 100 gráficos! Sin embargo, también me di cuenta que la mayoría de gráficos no hacían justicia a los datos de la encuesta, que (como decimos en Barcelona) tenían mucha tela. ¿Por qué gente tan lista no produce visualizaciones más interesantes?

Se trata de concientizar

El 80% de los científicos de datos que se graduará este año lo hará sin haber recibido ningún tipo de educación formal sobre visualización de datos, ni sobre cómo contar historias^[2]. Y el 60% de ellos sitúa la visualización de datos en la parte inferior de su lista de habilidades importantes para triunfar laboralmente. Además, el hecho de que la mayor parte de los cursos de Ciencia de Datos se enseñen en las escuelas STEM, no en escuelas de Arte, no ayuda. (En realidad, como dice John Maeda, las escuelas STEM deberían aspirar a ser escuelas STEAM e incluir Arte). Claramente, si queremos mejorar la calidad de la visualización de datos habrá que reformar el sistema educativo y hacerlo más tolerante a campos que como la Ciencia de Datos son tan artísticos como son científicos. Sin embargo, para visualizar efectivamente, la sensibilidad artística no es el único requisito.

En 2001, Angla Graber acuñó el término, *Muerte por PowerPoint* para referirse a presentaciones visuales que mostraban enorme potencial de mejora porque no seguían las más básicas reglas de comunicación (una idea por diapositiva, evitar poner más de cinco conceptos en una diapositiva, no leer las diapositivas y no poner demasiado texto).

Muerte por Excel



¿Muerto por PowerPoint o por Excel?

Pues bien, si en el caso de PowerPoint la responsabilidad de malas presentaciones recae primordialmente en la falta de entrenamiento de los usuarios de PowerPoint (Ver Presentation Zen). En el caso de Excel a esto hay que añadir lo complicado que es modificar el estilo de un gráfico en Excel.

Este es un aspecto al que claramente Bill Gates no prestó demasiada atención, ya que nunca asistió a una escuela de Arte, y eso queda patente en cómo Microsoft Excel trata y produce gráficos [30]. Por ejemplo, es muy difícil ajustar las opciones de los gráficos en Excel. Un solo gráfico de barras en Excel tiene la friolera +129 opciones configurables: opciones de relleno de color, ejes, grosor de línea, tipos de línea, opciones de leyenda de posicionamiento, opciones de escala, tipo barra, opciones de espaciado. Cada opción tiene entre 2 y 10 valores posibles. ¿Cuántas combinaciones posibles tenemos con esto? Son más de 100 000 000 000 000 000 000 000 000 000 000 000. (Un 1 seguido de 38 ceros).

Debido a las caprichosas Leyes de la Estética [57], muy pocas de estas combinaciones producirán un gráfico que se libre de un cero en cualquier clase de Diseño Gráfico. Para comprender lo difícil que es la tarea de ajustar las opciones gráficas solo basta decir que hay más combinaciones posibles que estrellas en la Vía Láctea y si pasamos tan solo 3 segundos en ajustar cada una de las 129 combinaciones, tardaríamos 6 minutos y 578 clics. No ayuda a que la mayoría del software de visualización en uso hoy en día (esto se aplica a la librería ggplot también) fuera desarrollado por ingenieros sin entrenamiento en los fundamentos de la Teoría del Color, Diseño de la Información o la Comunicación Visual.

La excepción fue Steve Jobs, que tomó algunas clases de caligrafía y que

atribuye a ello las fuentes hermosas de Mac. Ver biografía de Walter Isaacson [28] como un flashback de cuánto ha mejorado el campo.

Comience con un por qué

Una tercera causa es que para producir una visualización significativa necesitamos saber por qué estamos visualizando. Para llegar al por qué, es importante explorar la historia que necesita ser contada y ser capaz de hacer que sea relevante. Para tener éxito en la transformación de datos primarios en una historia convincente y fácil de recordar (sticky), se necesita conectar los datos a un contexto o marco de conocimiento. Los marcos de Economía, Arte, Sociología son particularmente útiles y veremos ejemplos más adelante. Sin embargo, hay que tener en cuenta que, con el fin de hacer estas conexiones, una educación multidisciplinaria es necesaria. Por desgracia, esto no es la norma. Para hacer frente a este desafío, en las siguientes páginas encontrará una serie de ejemplos, que me gustaría haber aprendido cuando era estudiante.



¡Felices visualizaciones!
Jose Berengueres
Estocolmo, 25 Oct 2019

Cómo usar este libro

Qué va a aprender (resultados del aprendizaje)

1. **Identificar** el papel de una narrativa en un gráfico
2. **Transformar** datos en información
3. **Sintetizar** conocimiento
4. **Aplicar** herramientas de pensamiento visual al proceso de toma de decisiones
5. **Seleccionar** técnicas de comunicación visual para persuadir
6. **Detectar** sesgos y mentiras en gráficos

Actualizaciones de esta edición

Esta edición añade dos nuevos capítulos, uno sobre el sesgo por Marybeth Sandell y otro en Psicología por Ali Fenwick. Sea como fuere, tanto si usted usa este libro como material para un curso de Visualización de Datos, como en Periodismo o en una especialización de Ciencia de Datos; este libro es una excelente introducción previo a aprender Tableau, Excel o Pandas Python.

Ejercicios

Hay varios ejercicios en este libro por lo que es útil tener papel y pluma a mano.

Diapositivas

Este libro nació de un taller de visualización y es usado en varias universidades como KAIST, UAE University y Cornell entre otras. Para conseguir las diapositivas en color envíen un correo electrónico a jse@ieee.org con el asunto “DATAVIZ101” y una captura de pantalla del texto de su *review*, opinión o valoración en website de Amazon y recibirá un e-mail con un link para descargarlas.

Capítulo 1. Las Historias, Datos y Narrativas

- Los “básicos” -



Figura 1 Estas tres revistas publican la mayoría de los gráficos más influyentes en diversas disciplinas. Ninguno de ellos está hecho con Excel^[3].

Antes de que empiece

Antes de visualizar, no hay datos. Sin embargo, antes de trabajar con datos, es importante entender qué es el conocimiento. ¿Qué es el conocimiento para usted? En este capítulo, aprenderá tres cosas; (i) La diferencia entre los Datos, la Información y el Conocimiento; (ii) Qué es la Sabiduría y (iii)Cuál que el proceso para producir Conocimiento.

¿Qué son los datos?

Ejercicio

¿Qué son los datos para usted? Siempre comenzamos los talleres de visualización de datos con este ejercicio.

Considere las siguientes ideas

	Conocimiento	
Datos		Sabiduría

Información

Ahora, ordénalas lógicamente

_____ > _____ > _____ > _____

Figura 2 *Un ejercicio para entender qué son datos en relación con el conocimiento.*

La palabra Datos tiene muchas definiciones. Sin embargo, para entenderlo, no hay nada como hacer^[4] un ejercicio. ¿Puede ordenar las cuatro palabras: ¿Datos, Conocimiento, Información y Sabiduría? La mayoría de los estudiantes se va a plantear el siguiente orden:

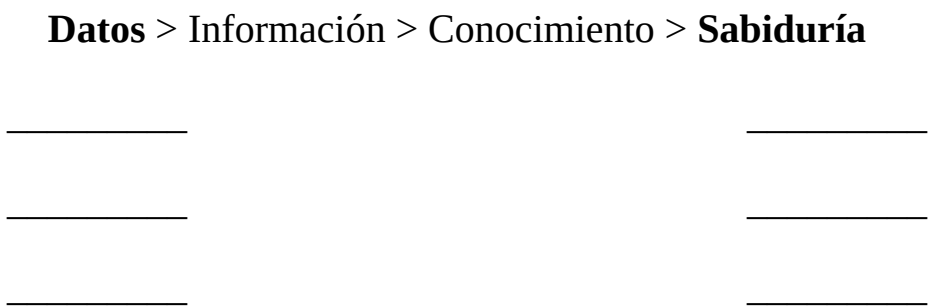
Datos > Información > Conocimiento > Sabiduría

Una vez que el orden de las palabras se ha acordado, podemos hablar de los criterios de orden. ¿Por qué los ordenamos como lo hicimos? Esta es una gran conversación de inicio. Para conectar a tierra la conversación, será aún de más ayuda identificar los atributos de las palabras de los extremos. ¿Cuáles son los atributos que distinguen a los Datos frente a la Sabiduría?

La Sabiduría frente a los Datos

Ejercicio

Dado el siguiente ordenamiento enumere atributos



***Fig. 3** Un ejercicio utilizado para entender la flecha del valor entre los Datos y la Sabiduría [56].*

Solución

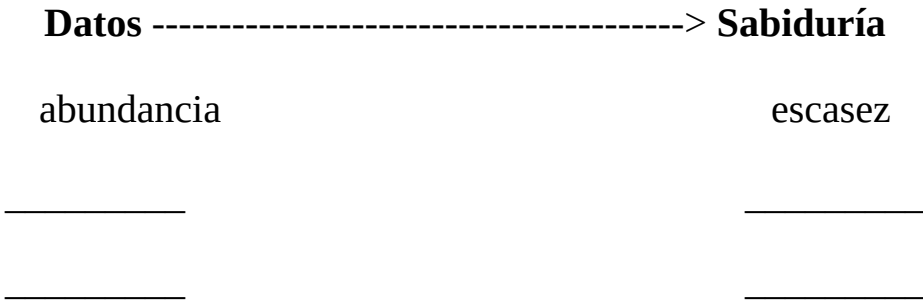


Fig 4 La flecha de valor

¿Qué es la sabiduría?

Los datos son muchos mientras que la sabiduría está concentrada. Los datos son abundantes, la sabiduría es escasa. El valor se correlaciona estrechamente con la escasez también [59].

Este ejercicio es excelente para aclarar la confusión generalizada entre los Datos, la Información, el Conocimiento y su relación con el valor, la escasez y la Sabiduría.

¿Cómo lograr sabiduría?

Datos -----> Sabiduría

Fig 5 Esta flecha representa un proceso.

Jackie Chan dice en una de sus películas, “*La información no es conocimiento, y el conocimiento no es sabiduría*”. Pero, ¿qué es la sabiduría? ¿Es la sabiduría sólo el conocimiento en su contexto? ¿Es el conocimiento meta sabiduría? - ¿conocimiento sobre el conocimiento? Y lo más importante, ¿saber en qué situación se aplica a un conocimiento determinado? Incluso si la definición no es universal, lo que interesa más aquí es cómo transformar el conocimiento en sabiduría.

¿Por qué? Debido a que es una actividad de un alto valor añadido y una de las razones (si no la única razón) por la que las empresas emplean a científicos de datos. Una forma de llegar a la sabiduría es el **proceso de síntesis**; la combinación dialéctica de la tesis y la antítesis en una verdad superior.

Una las palabras

Ejercicio

Conectar cada palabra clave con su imagen correspondiente. Tiempo de 1 minuto. Palabras: Datos, Información, Conocimiento, Sabiduría.

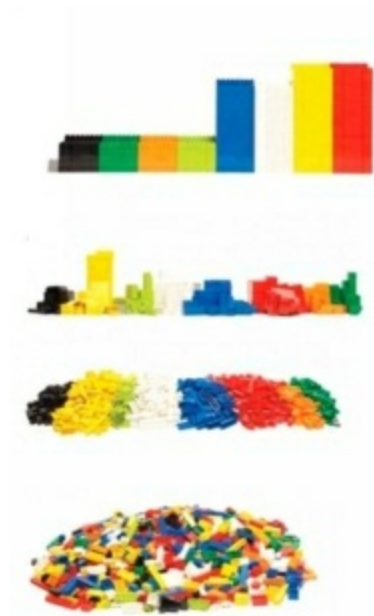


Fig 6 Un ejercicio utilizado para comprender lo que es la sabiduría^[5].

Solución

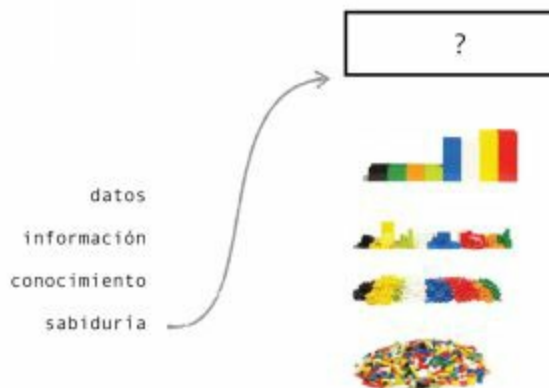


Fig 7 Para resolver este ejercicio, pensar fuera de la caja.

En este ejercicio, los estudiantes deben unir las palabras a una configuración de Lego. Cada pieza de Lego representa alguna pieza de datos. El ejercicio de Lego es una gran manera de aclarar que la Sabiduría (y en cierta medida el Conocimiento) no es sólo la representación de los datos organizados y presentados visualmente.

Vamos a hacer zoom en este caso. Imagínese que usted es el CEO de Lego y esta información le ha sido presentada. ¿Puede mapear las palabras a las imágenes? ¿Dónde mapea a la sabiduría? En este ejercicio la sabiduría no sabe regresar a casa ya que la sabiduría es algo más que la presentación y organización de datos. En otras palabras, la sabiduría es saber dónde aplicar el conocimiento. Ejemplo de la sabiduría: “No tenemos ninguna pieza de Lego de color rosa en este conjunto. ¿Por qué pasa esto? ¿Estamos ciegos a algún importante segmento de clientes? ¿Las ventas no son óptimas debido a eso?” Note cuánto conocimiento valioso surge cuando conectamos nuestra información (no hay piezas de color rosa) a otro conocimiento existente (estudios de género).

Información vs. significado

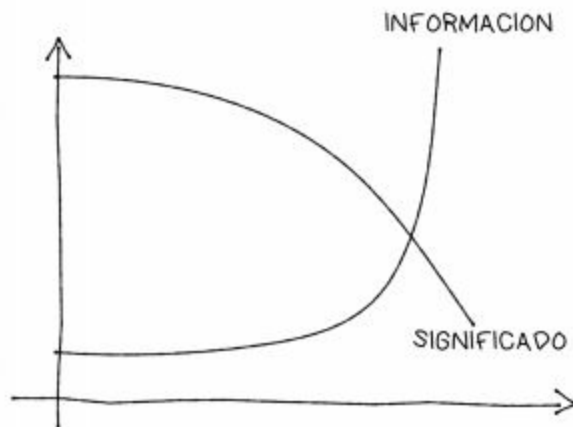


Fig 8 “No hay paz en más”.

Significado vs. Información

Una versión de esta tabla fue originalmente publicada por Hugo McLeod, el autor de “Ignorar a todos”. Cuando vi por primera vez esta tabla, me sentí fascinado por ella. Luego traté de encontrar defectos en el gráfico. Pero no hay defectos. Esta tabla es una obra maestra. Estaba desconcertado por su simplicidad^[6]. El eje ‘Y’ es el [Shannon Información Entropía](#), medida en número de bits. El eje ‘X’ es una secuencia ordenada (no continua y no tiene unidades). Es un ordenamiento de los casos (ejemplos de visualización) que el autor clasifica en sentido descendente, a lo largo del eje X y para cada elemento se representan dos puntos en el eje Y. ¡A continuación, el autor bebe de su propia medicina y simplifica el gráfico al reducir toda la información que no sea esencial para entregar el “Momento AHA!”: Menos (información) es más (significado). La línea de información se llama a veces “palo de hockey”. Es una función exponencial. La línea significativa traza un arco hacia abajo. Lo que significa esta tabla es que **no** es posible tener el significado con sobrecarga de información.

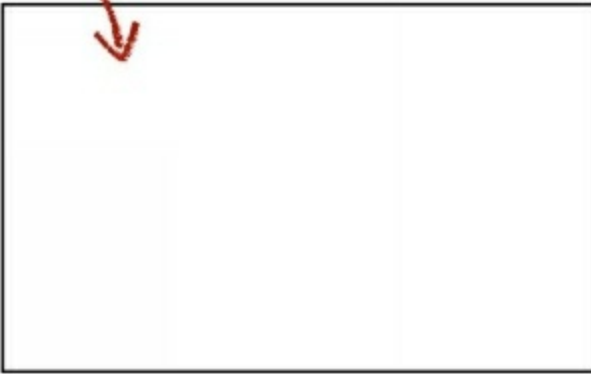


Fig 9 Una deconstrucción científica de la gráfica de McLeod.

Ejercicio

Resumir las 20 páginas anteriores en 20 palabras o menos. Tiempo de 3 minutos.

Ahora, visualice



Solución



Fig 10 Una metáfora visual basada en el modelo DIKW^[7].

Terminando

Cuando se realiza un taller, este ejercicio es una gran manera de llamar la atención sobre el punto de resumen vs. síntesis.

Un resumen es un proceso de simple reducción, mientras que la síntesis está demostrando una comprensión del tema, relacionándolo con otros sujetos, en última instancia, es la adición de valor a través del pensamiento conectivo.

En la figura 10, se resumieron la relación entre los datos y la sabiduría a través de la metáfora de la pirámide (valor = altitud). Este es un gran ejemplo de un resumen visual del capítulo, mientras que también es un buen ejemplo de síntesis.

Ahora que hemos aprendido la diferencia entre los datos y el conocimiento y cómo transformar el conocimiento en sabiduría, vamos a ver el papel de las narrativas en los gráficos.

Relatos e historias

Ejercicio

Llene el espacio con un verbo. Ejemplo: “Datos _____ historias”. Tiempo de 3 minutos.

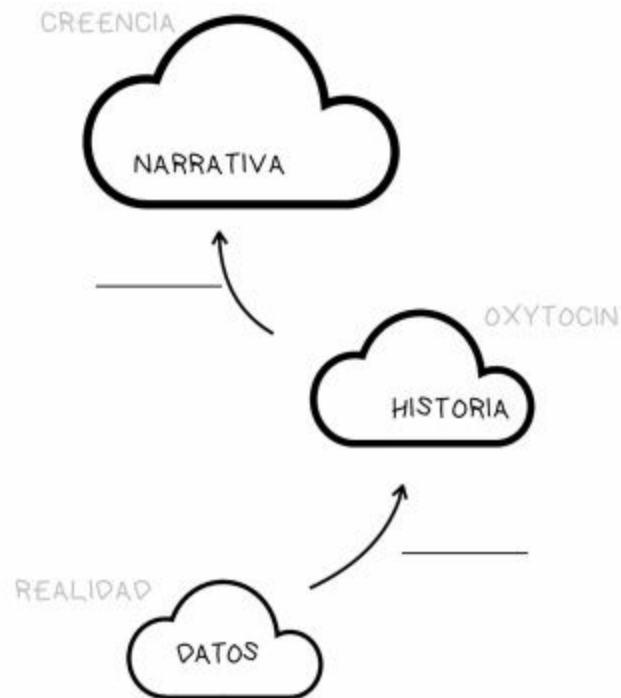


Fig 11 Un ejercicio utilizado para entender la diferencia entre los relatos y cuentos.

Solución



Fig 12 La plantilla de nube, utilizar cuando las cosas son borrosas [86].

Cuentos

Para entender lo que es una narración, en primer lugar, hay que entender lo que es una historia. Una historia es un relato de los acontecimientos. A los humanos nos encantan los cuentos. ¿Por qué? Contar y consumir historias es adictivo. Por ejemplo, escuchar una fábula de Esopo, leer un libro, y ver una película, todo ello libera oxitocina en el cerebro, hormona de sentirse bien. Es por eso que las personas se vuelven adictas a series de Netflix, telenovelas venezolanas e incluso a los libros de ficción. La historia oral se piensa que es el método más antiguo para el intercambio de narrativas. Desde una perspectiva antropológica, durante la infancia, las narrativas se utilizan para formar a los niños en comportamientos y valores. Esto se hace generalmente a través de cuentos.

Narrativas

Una narrativa es un conjunto de creencias, valores o visión del mundo. Por lo tanto, las narrativas elegidas interpretan el relato (y por lo tanto los datos subyacentes / realidad). Un ejemplo popular en la cultura europea es que los

niños, especialmente las niñas jóvenes, no deben confiar en extraños. Un cuento que promueve esta narrativa es la historia de Caperucita Roja, un cuento del siglo X. En las fábulas, la narrativa se explicita al final del cuento en forma de moraleja. Otro ejemplo de narrativa es el método Fear, Uncertainty, Doubt (FUD por sus siglas en inglés), Miedo, Incertidumbre y Duda. También se conoce como una estrategia de desinformación utilizada para frustrar el cambio en el status quo [32-34]. Se dice que IBM fue una de las primeras empresas que cotizan en el Dow Jones en usar FUD abiertamente. Una ejemplificación fue: “Nadie es despedido por comprar un IBM”.

Conexión a Aristóteles

Debido a que el objetivo de una historia es persuadir, los relatos, historias y datos están relacionados con las tres modalidades de la persuasión de Aristóteles: El Ethos, el Pathos y el Logos [49]. La narrativa se relaciona con el Ethos (apelación a los valores éticos), el relato está relacionado con el Pathos (apelación a las emociones), y los datos que apoyan la historia están relacionados con el Logos (apelación a la lógica).

Ejercicio

Identificar la narrativa, la historia, los datos y la llamada a la acción en esta foto. Tiempo de 2 minutos.



Fig 13 Un autobús del Brexit, Reino Unido. Foto original de David Beeson. En el bus se lee: Enviamos a la UE 350 millones a la semana, fundemos nuestra salud pública en vez de ello. Vote Salir de la Unión Europea. ¡Recuperemos el control!

Solución



Fig 14 *Anotaciones de un estudiante.*

A primera vista, la historia del bus Brexit es un ejemplo de una narrativa donde un grupo de fuera roba a un grupo de dentro. Pero es más que eso. De hecho, el autobús apela a mucha gente porque está conectado a cuatro relatos: (i) apelación a los instintos tribales del grupo interior para luchar contra los “extraños”; (ii) hacerlo por el bien común es ético; (iii) la lucha contra la injusticia es muy ética; y (iv) la maximización de la utilidad de un recurso es de sentido común.

Estas narrativas fueron usadas de manera brillante en el bus Brexit. Posteriormente se demostró que los números impresos en el autobús eran mentiras, pero no importaba. Esta historia apeló a los votantes a favor del Brexit y los movilizó.

¿Es ética su narrativa?

Las narrativas pueden ser evaluadas desde una perspectiva ética. Hay cuatro marcos éticos de uso común [35]. Cada uno maximiza algunos criterios éticos diferentes. Para cada marco, la opción más ética es el que maximiza una política diferente.

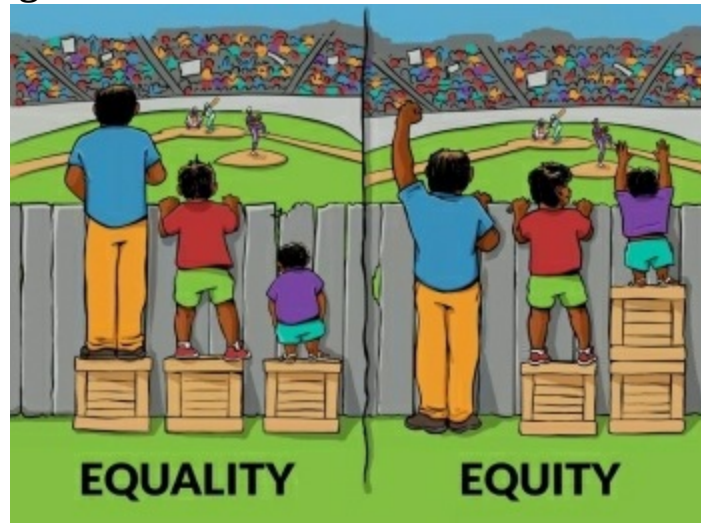
Los cuatro marcos

1. **Utilitario** → Optimiza el marco de la felicidad global
2. **Bien común** → Bienestar de la sociedad

3. **Justicia** → Equidad

4. **Virtudes** → Un conjunto de valores predefinidos

Equidad vs. igualdad



***Fig 15** La equidad (equity) no es lo mismo que igualdad (equality).
“Instituto de Interacción para el cambio social”, Angus Maguire.*

Ejemplos de marcos de virtudes famosos

- “Liberté, égalité, fraternité” [91]
- "Los diez mandamientos"
- El Bushido

¿Cuáles son los dilemas éticos?

Los llamados dilemas éticos surgen cuando el conflicto entre los enfoques aparece. Además, también pueden surgir dilemas cuando diferentes regiones en el cerebro (amígdala vs. lóbulo frontal) evalúan una situación con resultados opuestos. (Ver el Dilema del Tranvía) [36].

Opciones y ética

Para ser ético uno necesita considerar por lo menos dos opciones. Muchas veces, nos encontramos con opciones subóptimas simplemente porque las alternativas no fueron ni siquiera consideradas. Antes de la difusión de una imagen visual, consideramos siempre **dos** alternativas. A continuación, aplicar un marco apropiado para clasificar las alternativas. (Si algo va mal, al menos, se puede mostrar que siguió un proceso).

Elementos de una visualización efectiva

Ejercicio

Resumir la relación entre historia, narrativa y datos en una imagen visual titulada **Elementos de visualización efectiva** que utiliza un avión como metáfora. Llenar los espacios en blanco. Tiempo 3 minutos.

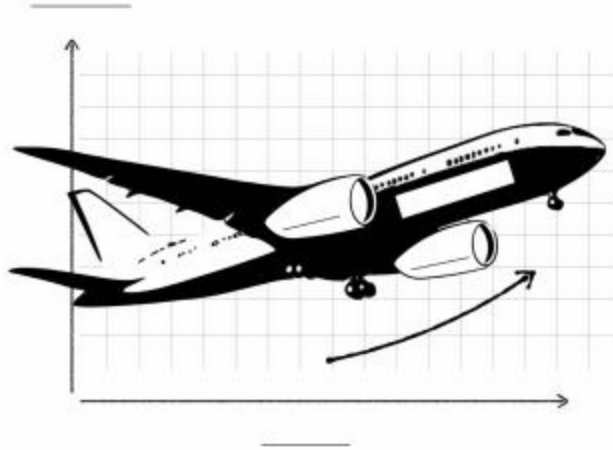


Fig 16 Metáforas visuales + anotaciones, son muy poderosas.

(Solución en la página siguiente)

Elementos de visualización efectiva

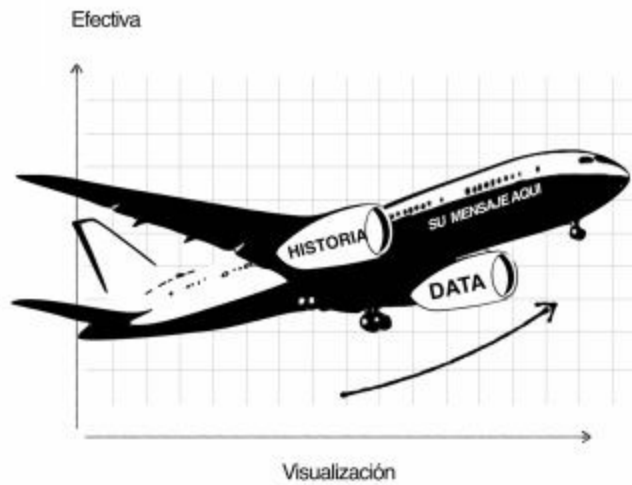


Fig 17 No permita que palabras como 'narrativa' estorben en una gran historia.

De acuerdo con esta imagen visual, para convencer a los demás se necesita:

1. Un porqué o una narrativa (mensaje)
 2. Una historia (para ayudar a asimilar el mensaje)
 3. Los datos para dar credibilidad a la historia
 4. Una manera de visualizar la historia (las metáforas ayudan)
- La carga útil es la narración (su mensaje) y usted es el piloto.

Cuestionario: Narrativas y relatos

¿Verdadero o Falso? Tiempo 5 minutos.

1. Una historia es un relato de acontecimientos.

[Verdadero/Falso]

2. El consumo de la historia libera oxitocina. [V / F]
3. El propósito de una historia es abogar (transmitir/perpetuar) una narrativa (una creencia, una cultura, una ideología). [V/ F]
4. Una narración que no se inserte en el cerebro humano, no puede ser viral, no es fácil de recordar. [V / F]
5. Si su historia (gráfico, visualización) no transmite alguna narrativa, se sentirá como si no tuviera propósito. [V / F]

(Solución en la página siguiente)

Solución

1. Una historia es un relato de acontecimientos [V / F]. Cierto. Para ganar perspectiva sobre esta cuestión véase también el trabajo de Lisa Filman “Constructed Model of the Word as a Reality” [88].
2. El consumo de la historia libera oxitocina [V / F]. Cierto. Para obtener más información sobre la bioquímica de la narración y su función en la evolución véase [89-90].
3. El propósito de una historia es transmitir/ perpetuar) una narrativa (una creencia, una cultura, una ideología) [V / F]. Cierto. Véase también [89-90].
4. Una narración que no se pegue al cerebro humano, no puede ser viral, no es fácil de recordar [V / F]. Cierto. El mensaje de una narrativa en su forma sucinta no suele ser viral. Historias en forma de anuncios, películas y libros son más adecuadas para ser viral.
5. Si su historia (gráfico, visualización) transmite/perpetua alguna narrativa, se sentirá como si no tuviera propósito [V / F]. Cierto.

Cuestionario: Storytelling del cambio climático

Tiempo 15 minutos. Búsqueda en Internet: Permitida. En agosto de 2019 Greta Thunberg zarpó de Europa a una conferencia sobre el clima de la ONU en el otro lado del Atlántico. Identificar los datos, la historia y la narrativa en esta historia.



Fig 18 A diferencia de Al Gore en 2006, Greta Thunberg no necesitaba gráficos para transmitir su mensaje^[8].

Solución

Narrativa (ejemplo de respuestas)

- “El perdedor gana^[9]”
- "Cambio de roles"
- "Familia disfuncional"
- “El cambio climático es una emergencia”

Historia

(Resumen). “Usted puede ser adulto, pero no es lo suficientemente maduro para entender esta emergencia. Si lo fuera, no habría volado a la conferencia como lo hizo. Podía haber hecho un Skype o viajar como yo para reducir su huella de carbono. Por lo tanto, usted (no yo) se está comportando como el niño inmaduro...” Ver discurso completo en [92].

Datos

- Mira las grandes olas detrás de mí, estoy seria, esto es peligroso.
- Un velero neutral en carbono = es posible reducir la huella de carbono.
- La situación es suficientemente mala para saltarme las clases de la escuela
- Viajar por aire en jet produce CO₂, pero hay alternativas. ¡Mírame!

Capítulo 2. Visualizando la Información

- Cómo transformar los Datos en Información -

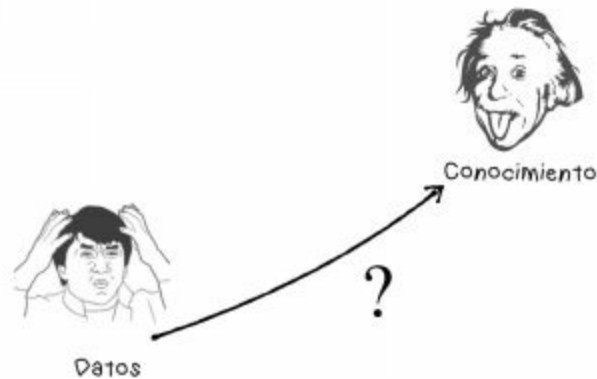


Fig 19 *La creación de conocimiento a partir de los datos. ¿El secreto para ganar el premio Nobel?*

En este capítulo, aprenderá a transformar los Datos en Información, un requisito previo para producir conocimiento. En otras palabras, cómo hacer que sus gráficos sean más fáciles de usar y útiles mostrando información **efectivamente**.

“La base de un gráfico útil es rara vez el diseño gráfico”

La mayor parte de los libros publicados sobre Visualización de Datos ponen el foco en cómo crear gráficos y cómo hacerlos legibles. Algunos de ellos dedican un considerable número de páginas a relacionar todos los tipos de gráficos disponibles por ahí y utilizar la lógica de los gráficos buenos/malos para enseñar cómo un pequeño retoque puede mejorar o romper la legibilidad de un gráfico.

Otros dan consejos sobre el color y lo importante que es no saturar su gráfico con colores, (ver también el término *Gráficos Basura* acuñado por Tufte [50]). Esto es útil para mejorar la legibilidad y la estética, pero no son suficientes para transformar la Información en Conocimiento (también conocida como Sabiduría, Análisis Prescriptivo^[10]). Y mientras que las deficientes opciones de color pueden matar la legibilidad de cualquier

gráfico, centrarse en la estética del gráfico es el equivalente a enseñar acerca de la importancia de los tipos de fuente a alguien que solo quiere ser escritor.

Después de todo, el libro de Harry Potter no se convirtió en un éxito de ventas debido al tipo de fuente que utilizaron. En la misma línea, la causa de raíz típica de un gráfico deficiente recae en la transformación de los datos en conocimiento (significativo), no en el uso de estéticas equivocadas. En este capítulo, usted aprenderá estas habilidades con ejemplos. Empecemos con los datos sobre la distribución por género de la encuesta Kaggle 2018, una de las mayores comunidades científicas de datos en todo el mundo

Visualizar género

Ejercicio

¿Cómo **visualizará** el siguiente desglose por género de los datos científicos? Tiempo de 3 minutos.

Respuestas de la encuesta	
Femenino	16,8%
Masculino	81,4%
NS	1,4%
NC	0,3%

Solución

Solución con Matplotlib

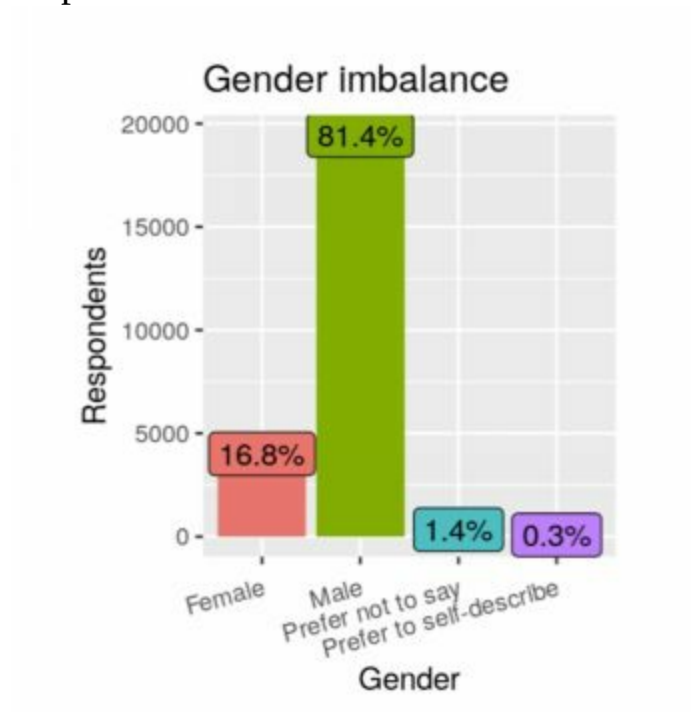


Fig 20 Esta gráfica Matplotlib predeterminada utiliza **cuatro** diferentes tamaños de fuente y **seis** colores (rojo, verde, turquesa, púrpura, negro, gris).

Ejercicio

Es la cifra sobre Datos, ¿Información o Conocimiento? ¿Por qué? Razone su respuesta. Tiemp2 minutos.

Solución

Es solo información. No es Conocimiento, ya que no es significativamente más útil que los datos originales.

Reflexión

A menos que esté en un Análisis Exploratorio de Datos preliminar (AED), no es una buena idea diseminar un gráfico si no está claro el por qué (narrativa) para el gráfico. E incluso si produce muchos gráficos como parte de un AED, resista la tentación de mostrarlos todos. En este caso, se nos pide graficar para visualizar la distribución por género de los que contestaron la encuesta 2018 Kaggle. El género fue una de las más de 30 preguntas de la encuesta que fue respondida por unos 30 mil encuestados. La figura 20 es el gráfico de configuración por defecto y que fue producido utilizando el popular *Matplotlib* de la biblioteca de Python. Este gráfico está perfectamente bien. Es informativo, pero no hay mensaje, no hay por qué. Carece de un propósito. ¿Por qué? Una de las razones es que no está conectado a ninguna narrativa. Otra razón es que no aumenta nuestro conocimiento. ¿Nos está ayudando a ser más sabios? ¿Está facilitando la función analítica prescriptiva? ¿Cómo haría este gráfico más útil?

Ejercicio

Dibuje aquí por lo menos **tres** gráficos alternativos a la figura 20. b
Tiempo 3 minutos. (Solución en la página siguiente).

El ajuste Gráfico-Narrativa

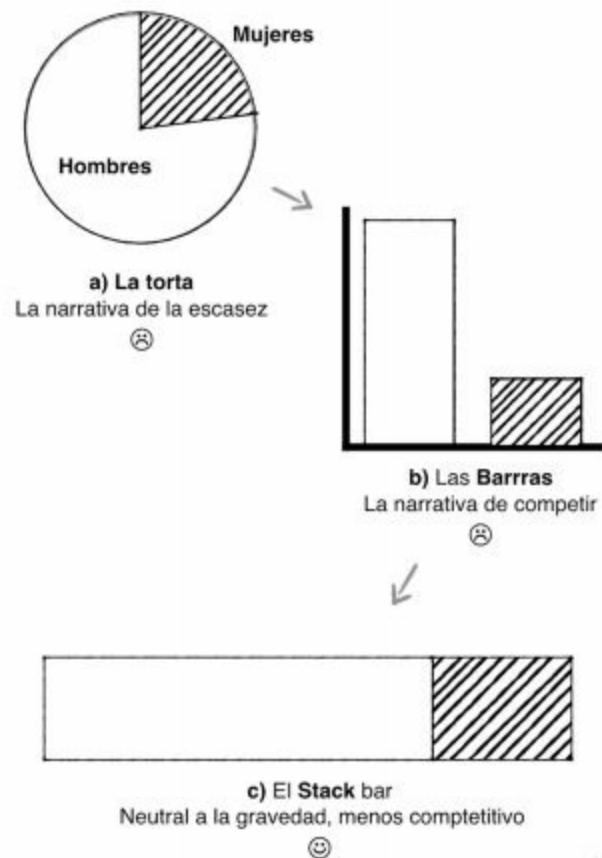


Fig 21 Tres formas de visualizar los mismos datos.

Elegir el gráfico correcto

La figura 21 muestra la misma información que la figura 20 de tres maneras diferentes a) gráfico circular b) gráfico de barras c) gráfico de barra apilada horizontal. Vamos a considerar el proceso que seguimos para producirlos.

Simplificar

Consideramos solo hombres y mujeres e ignoramos los otros dos grupos de valores atípicos.

Evitar colores

Los hombres están representados en blanco y las mujeres en la zona

rayada. Evitamos el color, porque podemos. Los colores se cargan con simbolismos [31], en escala de grises esto pasa con menos frecuencia. La conjugación de colores es un arte muy sutil, que es fácil subestimar. ¿Por qué utilizar más de seis colores sin la formación adecuada en Diseño Gráfico? También hemos eliminado las escamas, las figuras y las etiquetas de los ejes. Intercambiamos un poco de precisión por una gran ganancia en claridad.

Considerar una narrativa

Supongamos que estamos abogando por una narrativa de equidad de género. Definimos la narrativa equidad como una visión del mundo en la que el mundo será un lugar mejor si hay menos desequilibrio en la proporción de género [37].

Comprobar el ajuste de la narrativa del gráfico

Una vez que se establece una narrativa, la elección de una plantilla de gráficos **compatible** es el siguiente paso para transformar la Información en Conocimiento.

En este caso, los gráficos circulares y los de barras son una elección poco efectiva. ¿Por qué? Debido a que los gráficos circulares están conectados a la narrativa de la cantidad de pastel que cada grupo obtiene (recursos finitos), es una narración de confrontación que socava la narrativa de la equidad de género. El gráfico de diagrama de barras es también una mala elección, ya que está conectado a la narrativa de competencia y a otras narrativas ganar-perder, tales como la escasez y la visión de que los hombres y las mujeres compiten. Su utilización corre el riesgo de socavar cualquier mentalidad de crecimiento [68] o narrativa de equidad (ganar-ganar) por la que podría estar abogando. La barra apilada horizontal es una mejor opción. Es horizontal, de modo que no está sujeta a la metáfora de la gravedad. (Ver [gravedad y gráficos](#)).

Ahora que hemos encontrado un ajuste entre el gráfico y la narrativa, vamos a diseñar el gráfico de manera que sea más fácil de asimilar por personas de diferentes edades y géneros. A esto le llamamos gráficos centrados en personas. Los gráficos centrados en humanos aprovechan los mismos principios que los diseños centrados en usuarios, colocándolos en el centro del proceso de diseño.

Gráficos Centrados en Personas

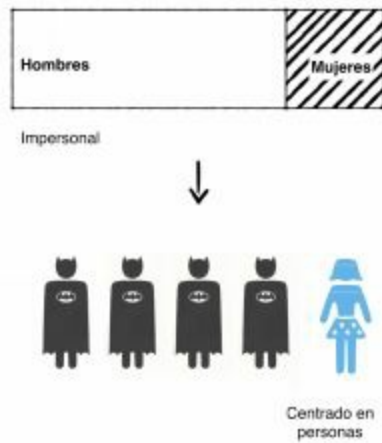


Fig 22 Cuatro Batman y una Mujer Maravilla hacen este gráfico más fácil de recordar.

El uso de superhéroes es una gran manera de visualizar el pesado tópico [1-3,52] de la equidad de género en entornos dominados por STEM. También usamos la iso-medida. En un gráfico de iso-medida, cada unidad (en este caso un superhéroe) representa la misma cantidad (en este caso, alrededor de 5.000 respuestas de la encuesta). También hemos reducido el recuento de superhéroes al mínimo posible (cinco personas) lo que significa la aproximación de los porcentajes originales a un ratio de 1:4 mujeres a hombres. Este gráfico también se puede leer de la siguiente manera: en promedio, por cada 5 personas en un equipo (habitación/reunión), uno será mujer.

Comprobación del ajuste gráfico-narrativa

Dada la misma narrativa de equidad de género utilizada anteriormente, vamos a ver cómo las opciones de diseño que hemos hecho se ajustan a tal narrativa.

- El trabajo de un superhéroe es hacer del mundo un lugar mejor, la narrativa de equidad es acerca de hacer del mundo un lugar mejor. Este tema lúdico ayudará al lector a recordar el gráfico [55].
- El iso-medida es el ser humano, al igual que los encuestados.
- El número de héroes es inferior a siete, (estamos respetando la regla de

los 7 trozos y no sobrecargamos al lector con información).

- Utilizamos el humor para mejorar la eficacia de la comunicación [55].

Aprovechando el humor

Muchos gráficos son impersonales, porque no podemos relacionarnos con ellos. Hemos resuelto esto con los superhéroes. Vemos también personas. No obstante, si además queremos que la audiencia recuerde el gráfico, podemos utilizar el humor o una broma interna como que los científicos de datos son superhéroes, porque tienen que “pelearse” con los datos, ver el término *Disputa de Datos*.

¿Sexismo en tu gráfico?

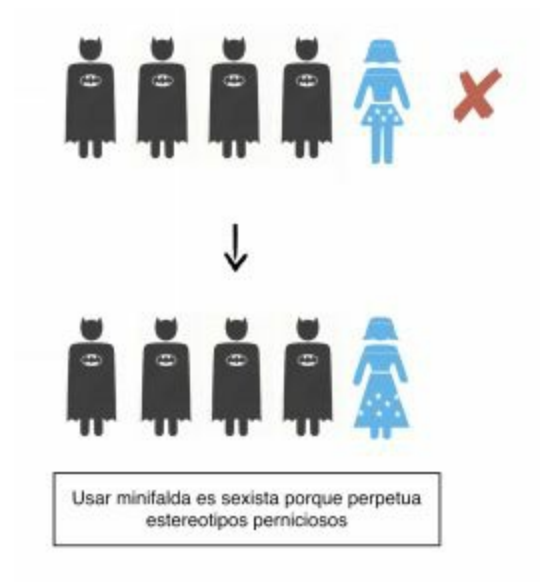


Fig 23 Verifique siempre los puntos ciegos.

Lista de verificación del gráfico

- ☒ Tipo de gráfico compatible con la narrativa (mensaje)
- ☒ No hay sobrecarga de color (3 colores máx.)
- ☒ Un gráfico, un mensaje
- ☒ Metáforas alineadas con la narrativa
- ☒ No necesito leer el subtítulo para entender el gráfico
- ☒ El subtítulo se utiliza como una oportunidad de síntesis
- ☒ El subtítulo no explica el gráfico de nuevo
- ☒ Sesgo verificado por un tercero

¿Cómo comprobar el sesgo? Es importante verificar los puntos ciegos [51-53]. Los gráficos no son indiferentes. Es prudente pedir la opinión a diversas personas, idealmente con diferentes orígenes. Véase el Capítulo 5 para más información sobre el sesgo.

La gravedad y gráficos

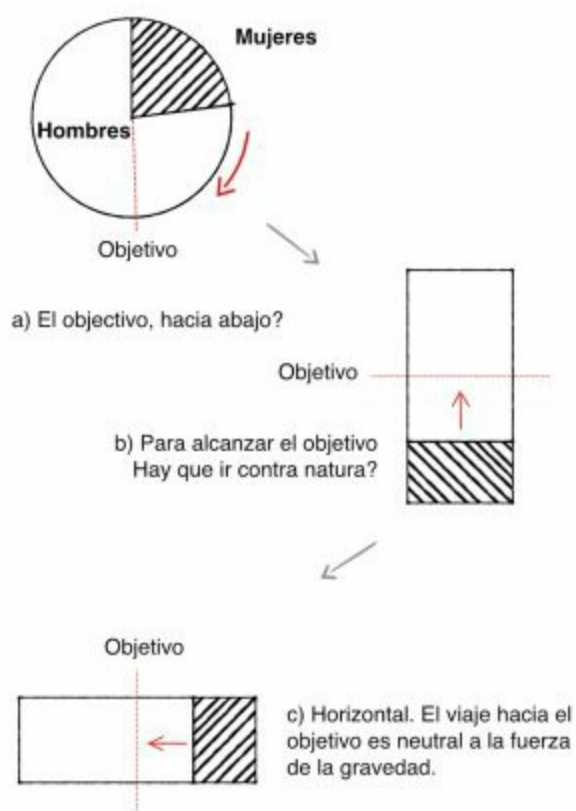


Fig 24 La gravedad da forma a todo en la Tierra, incluyendo la forma en que interpretamos los gráficos.

¿Cómo afecta la gravedad a los objetivos

Vamos a suponer que los gráficos han convencido a su organización que el logro de la paridad de género en el trabajo es una buena meta y después de una reunión de la junta directiva se ha establecido el objetivo de llegar a una ratio del 50%: paridad de género. ¿Cómo visualizarlo para **persuadir** y unir al resto de la organización? Las flechas apuntan a la meta. Note cómo el alineamiento de las flechas con la dirección de las influencias gravitatorias se percibe como la meta alcanzable. En el gráfico de sectores, la flecha hacia abajo tiene una connotación negativa. Para el gráfico vertical la flecha hacia arriba, contra la gravedad, hace que el objetivo parezca difícil de lograr. El gráfico horizontal es gravedad neutral y el más libre de connotaciones que podrían distraer de la narrativa.

Guerra de gráficos: Tortas contra Barras



Fig 25 Musk vs Bezos. Dos visiones de la exploración espacial. Dos formas de visualizar la altitud.

Cuándo utilizar las tortas

Hay una diferencia fundamental entre los gráficos circulares y gráficos de barras. El cerebro es sensible al cambio angular y (por comparación) insensible al cambio lineal [72]. Esto es particularmente cierto cuando se considera el movimiento y la sensibilidad a pequeños cambios. Si en su narrativa, destacar pequeños cambios en una variable es importante para la historia, los gráficos circulares (medidores de velocidad) son el camino a seguir. Si, por el contrario, demasiada atención al cambio es una distracción, evite gráficos circulares. Compare usted mismo. En la emisión del lanzamiento de prueba de *Blue Origin*, el cambio de actitud es apenas perceptible. Mientras que en la transmisión del lanzamiento del *Space X*, aparece en todo.

- Emisión en vivo en Blue Origin: <http://bit.ly/2NHycmf>
- Emisión en vivo en Space X: <http://bit.ly/2XwXYxY>

Cuestionario: Hacer gráficos útiles

¿Verdadero o falso? Tiempo 10 minutos.

1. Un gráfico de barras es una gran manera de visualizar las posibilidades de la Ruleta del Casino. [V / F]
2. Los seres humanos son más sensibles a los gráficos circulares que a los de barras. Sin embargo, en los gráficos animados donde las relaciones cambian, estos cambios son más obvios en un gráfico de barras. [V / F]
3. Un inconveniente de la tabla de iso-medida es que no se pueden utilizar metáforas o personas. [V / F]
4. Un camino para convertir información en conocimiento es el uso de una paleta con la mínima cantidad de colores posible. [V / F]
5. El propósito de una EDA es contar una historia acerca de los datos. [V / F]

(Solución en la página siguiente)

Solución

1. Un gráfico de barras es una gran manera de visualizar las posibilidades de la Ruleta del Casino [V / F]. **Falso.** Un gráfico de barras está vinculado a la narrativa completa. La ruleta es un juego de suma cero. Un gráfico circular, un gráfico de **iso-medida** o una foto de la propia ruleta comunican posibilidades más claramente.
2. Los seres humanos son más sensibles a los gráficos circulares que a los gráficos de barras. Sin embargo, en los gráficos animados donde las ratios cambian, dichos cambios son más evidentes en un gráfico de barras. [V / F]. **Falso.** Son más sensibles a los gráficos redondos / agujas en ambas situaciones.
3. Un inconveniente del gráfico de iso-medida es que no se pueden utilizar metáforas o personas con él. [V / F]. **Falso.** Es lo **opuesto**. El iso-medida se presta a las metáforas.
4. Una manera de convertir la información en conocimiento es el uso de una paleta con la mínima cantidad de colores posible [V / F]. **Falso.** **Evitar** la sobrecarga de información es una condición previa para el conocimiento. Sin embargo, la transformación fundamental es la utilidad. (Por lo general, conectándola a otros conocimientos).
5. El propósito de una EDA es contar una historia sobre los datos [V / F]. **Falso.** El propósito **principal** de un Análisis Exploratorio de Datos consiste en agregar y visualizar información estadística básica.

Capítulo 3. El Conocimiento

- De la Información al Conocimiento -



Fig 26 Este gráfico utiliza la metáfora de la gravedad.

En el capítulo 1 vimos que los gráficos con un por qué y un propósito claro tienen un mensaje. En el capítulo 2, vimos un ejemplo de cómo transformar los datos de género en información. En este capítulo, vamos a aprender cómo sintetizar la información en conocimiento. Si estuviéramos en una clase de idioma inglés, esta habilidad sería el equivalente a escribir el libro de síntesis. Sus datos son el libro, el gráfico es la síntesis. Una manera eficaz de hacer esto es mediante el uso de marcos de referencia, técnicas de resumen y metáforas visuales. Veamos un ejemplo que usa datos de edad a partir de una encuesta.

Marcos mentales

Vamos a echar un vistazo al gráfico de la figura 27. ¿Cuántos trozos de información se pueden contar? Tiene una sobrecarga de información. Enfoquémonos en una paleta de colores, por ejemplo, un arco iris. Sin embargo, un arcoíris no transmite ningún significado aquí. Por el contrario, mediante el uso de 12 colores, hemos aumentado la sobrecarga de información por la friolera de **12 trozos de memoria** [48] con una ganancia cero en significado, (véase información frente a significado en el cap. 1).

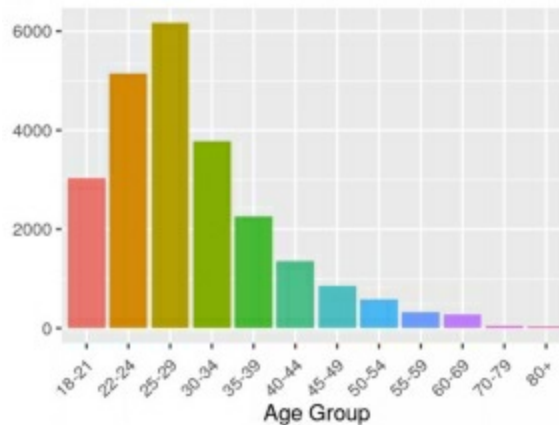


Fig 27 Una visualización por defecto ggplot de la edad de los encuestados de la encuesta Kaggle 2018. Escala de colores: Arcoíris, 12 colores.

Ejercicio

¿Cómo podríamos hacer más útil hacer la tabla anterior? Podemos empezar por la reducción de la sobrecarga de información. Dibuje soluciones. Tiempo 2 minutos.

Solución

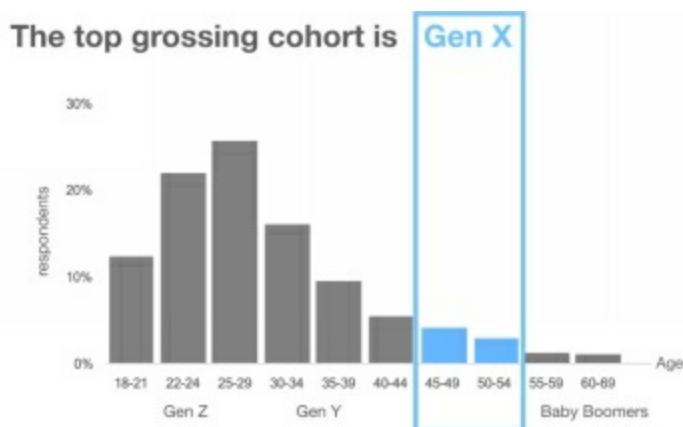


Fig 28 La misma información con una narrativa vinculada a un marco de referencia, más cerca del conocimiento. Fuente: [26]

¿Qué podemos aprender de los científicos de datos de alto nivel?

Al igual que la figura 27, la figura 28 es una distribución por edades. Sin embargo, aquí usamos un esquema de dos colores [18] para resaltar qué grupo de edad ganó la mayoría de las competencias por usuario. Sin embargo, tantas barras para representar grupos de edad pueden abrumar a cualquier lector. Una manera de suprimir elementos y estructurar los contenedores en información **utilizable** es reducir su número y agruparlos en una forma familiar conocida. Una manera de hacer esto es reducir los grupos de edad a grupos de generación. En este caso, se utilizó un marco de referencia con el que muchos están familiarizados: generaciones en edad laboral. Comprende la Generación X, Y, Z y los Boomers [5]. Por otra parte, estamos interesados en ver qué grupo es el más productivo en términos de competencias y premios en efectivo por usuario. Como todo el mundo pertenece a una generación, este gráfico puede llegar a ser muy agradable. ¿Qué podemos aprender de la sabiduría que cada generación ofrece? Fuente: 2018 Kaggle Encuesta Q2 ¿Cuál es su edad? [27].

Generación y marco atributo ético del trabajo

- Los **Baby Boomers**, nacidos 1946 - 1964, “marcados como adictos al trabajo” [6]
- **Gen X**, nacidos 1967 - 1977, “esta generación trabaja para vivir y llevan con ellos un nivel de cinismo” [7]
- **Gen Y**, “Millennials” nacidos 1980 - 2000 “considerada la generación más educada y consciente de sí mismo en el empleo” [8]

- **Gen Z**, nacidos después de 2000 - ver [52]

Narrativa

La discriminación por edades es la narrativa que dice cosas como “las personas de edad no tienen energía para ser emprendedores y no pueden innovar” [14]. Sin embargo, algunos trabajos han refutado esta afirmación [72]. El gráfico defiende la narrativa opuesta “(En la Ciencia de Datos), las generaciones mayores son tan productivas como las generaciones más jóvenes”.

Creación de conocimiento

Note cómo el paso clave para crear un significado (conocimiento) no solo es resumir y eliminar lo que no agrega valor, sino encontrar **dónde** es más útil la información vinculándola a **ese** contexto (marco de referencia). En este caso, las generaciones en el lugar de trabajo y productividad. Otra forma de crear significado es a través de metáforas visuales, tales como la pirámide. Veamos un ejemplo basado en la información sobre salarios.

Visualizando la inclusión

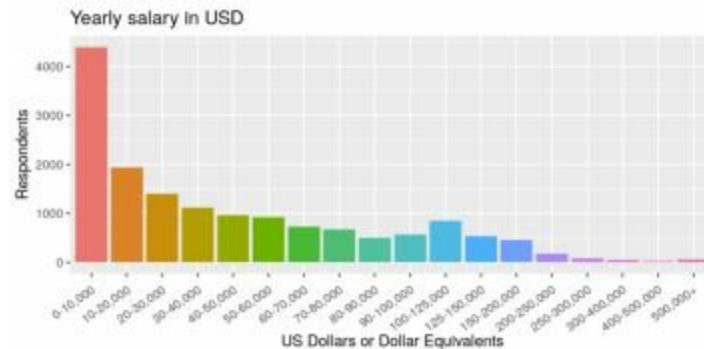


Fig 29 Una visualización ggplot2 por defecto de los sueldos de la Encuesta 2018 Kaggle. Fuente Q9 ¿Cuál es su compensación anual actual? [27].

Ejercicio

En la sección anterior, vimos un ejemplo de lo que significa la creación de significado por medio de la conexión de la información a un marco de referencia. Ahora vamos a hacer lo mismo y, además, vamos a aplicar una metáfora visual. Echemos un vistazo a los datos salariales de la misma encuesta 2018. ¿La figura 29 corresponde a Información de Datos o Conocimiento?

Dada una narrativa de inclusión ¿cómo crear un gráfico más útil? Tiempo 6 minutos.

Sugerencia: Si el gráfico fuera un edificio ¿Dónde tendrían los apartamentos los individuos altamente pagados?

Paso intermedio

(Fig 29 gira en sentido anti horario 90 grados)

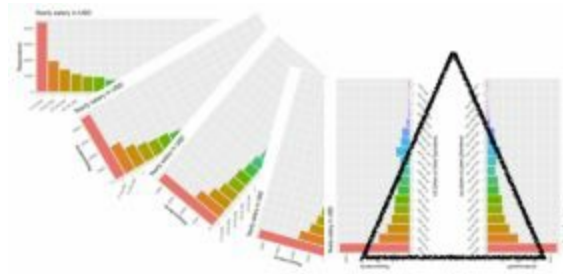


Fig 30 ¿Su gráfico se adapta a cualquier metáfora visual?

Solución

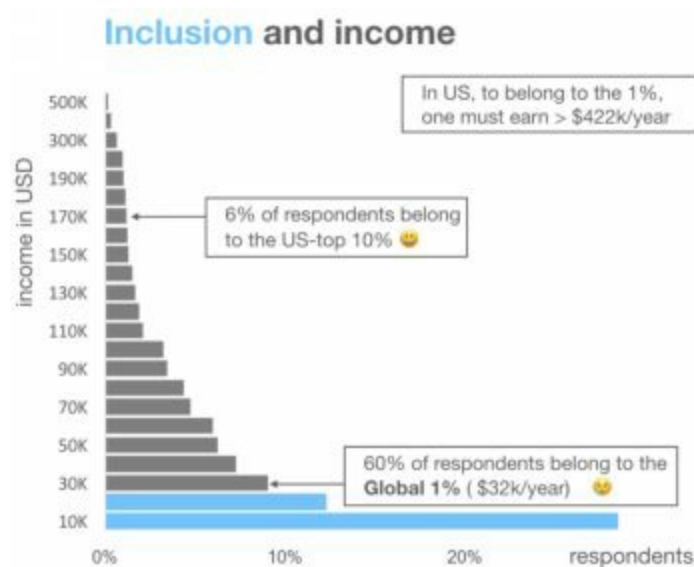


Fig 31 El uso de texto dentro de los gráficos es una gran manera de desactivar la ambigüedad. Fuente: [26]

Anteriormente mencionamos la importancia del ajuste narrativo del gráfico. La narrativa de la figura 31 trata de la brecha digital. ¿Cuán inclusiva es la Ciencia de Datos como comunidad? Como acabamos de ver en la sección anterior, una forma de crear conocimiento es relacionar nuestra información a los marcos de referencia existentes.

Las generaciones en la fuerza laboral fueron un ejemplo de un marco popular para pensar en la edad y la fuerza de trabajo. Ayudar al lector a **situ**ar la nueva información en los marcos existentes hace que la nueva información sea más fácil de asimilar, utilizar y recordar. Aquí, aplicamos el marco de referencia de los **percentiles de ingresos**, un marco de análisis común utilizado por los economistas y ajustado con la pirámide una metáfora que representa las jerarquías. Véase también el movimiento #onepercent. En EE.UU., para pertenecer a la élite del 1%, uno necesita ganar más de \$ 422K por año [10]. Cerca de 23 encuestados declararon que ellos lo hacían. Además, alrededor del 6% declaró que pertenecen al percentil 10%, un número muy inclusivo porque el 6% es similar al 10%. El 10% percentil de ingresos es de aproximadamente \$ 166k en los Estados Unidos [11], por lo que, si la muestra refleja la distribución encontrada en la sociedad, significa que es por lo menos de alguna manera inclusiva. Añadimos un emoji sonriente para tranquilizar al lector que sí, esto es bueno. Sin embargo, esos números son para ingresos de los hogares en Estados Unidos. Cuando miramos a nivel mundial, el umbral del percentil 1% es de \$ 32k por año. Esto pone al 60% de los encuestados en el 1% superior. El 60% es muy diferente del 1%, por lo que a nivel mundial este punto de datos no admite la inclusión, ya que no refleja la distribución global. Momento ¡AHA! Una forma de crear esos momentos es cambiar entre dos puntos de vista. Acabamos de ver lo poderoso que puede ser metáforas visuales. Vamos a ver tres ejemplos más.

El ganador se lleva todo

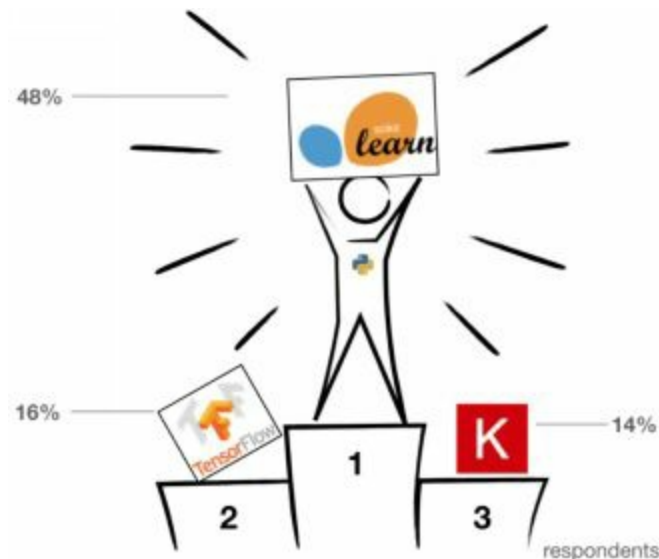


Fig 32 Las tres bibliotecas principales de Machine Learning. Fuente: [26].

En este caso, estamos visualizando las bibliotecas de Ciencia de Datos que utilizan los encuestados. Utilizando la encuesta Q20. De las opciones que seleccionó en la pregunta anterior, ¿Qué biblioteca ML utilizó más? [27]. Dada la narrativa un ganador se lleva todo (tan común en el mundo del software), ¿Qué metáfora visual podemos aplicar? Este gráfico es un ejemplo de menos, es más. En este caso, Sci-Kit (una famosa biblioteca científica de Python) tiene una participación del 48%, TensorFlow de Google tiene un 16%, seguido por Keras 14%. Veamos cómo esta visualización está conectada a la narrativa.

Ajuste gráfico-narrativa

En escenarios de negocios con fuertes externalidades en juego, tales como una red social, un sistema operativo para teléfonos o una carrera olímpica; estar en el podio (siendo primero) tiene un efecto desproporcionado sobre la recompensa. En tales casos, el ganador se lleva todo el relato. Antropomorfizar el ranking con un podio transmite una narrativa memorable y de accesibilidad^[11]: La gloria que el ganador se merece por la gran utilidad que esta biblioteca presta a la comunidad. Esta narrativa también está conectada a otros memes famosos en el mundo del software, tales como la gloria del desarrollador.

(Véase S. Balmer en “Desarrolladores, desarrolladores, desarrolladores, desarrolladores”).

Todo o nada

En la sección anterior, visualizamos datos sobre las bibliotecas de Machine Learning más populares con un ganador que se lleva toda la narrativa, aquí hacemos lo mismo con una narrativa diferente.

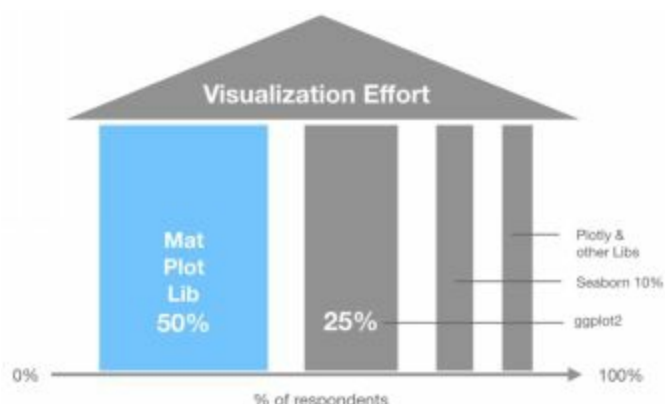


Fig 33 La casa de Shiva. Cuando el color ocupa grandes áreas, utilizar un 50% de gris y colores pastel en lugar de 100% de colores sólidos llamativos.

Casa de Shiva

La Fig 33 es una combinación de (i) una plantilla de gráfico llamado Marimekko con (ii) un gráfico simbólico llamado Casa de Shiva [13]. La Casa de Shiva se usa para enfatizar las relaciones de todo o nada. La metáfora es que el techo cae si una sola columna se derrumba.

Simbolismo

Las columnas soportan los esfuerzos de visualización de la comunidad (carga del techo, bien común). La anchura de las “columnas” expresa cuánta carga/trabajo soporta cada columna. Las columnas grises de la derecha representan las bibliotecas menos convencionales tales como: D3, Shiny, Bokeh, Leaflet, Lattice. Fuente: Encuesta Q22 ¿Qué biblioteca específica de visualización de datos o herramienta ha utilizado más? [27].

Metáfora

El objetivo es el techo. Al igual que con una casa, la integridad del mismo se deteriora claramente si una columna es débil.

Narrativa

La narrativa es que las bibliotecas de visualización no convencionales son importantes, pero en diferente grado. Nótese aquí que si se hubiera utilizado

un gráfico circular habríamos transmitido una narrativa de escasez de ganar o perder, no representando fielmente el espíritu de ganar-ganar del movimiento de código abierto [111].

El marco BRICS

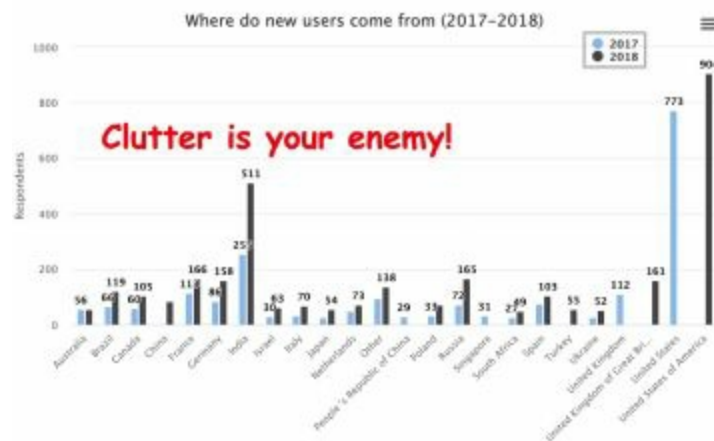


Fig 34 Un ejemplo de pérdida de significado por sobrecarga de información. Fuente: <http://bit.ly/2K7ZLBk>

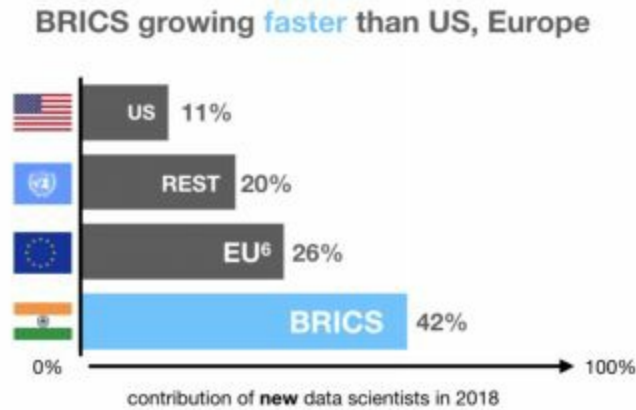
Una forma común de resumir los datos internacionales es contarlos agrupándolos por países. Sin embargo, esto a menudo conduce a la confusión porque hay más de 200 países en el mundo. Otra forma común es el uso de un mapa geográfico mundial, y modular el color del país con el recuento. Sin embargo, algunos países son muy grandes en extensión, mientras que otros países se convierten en un píxel en la pantalla. Esto no es ideal para su uso.

Ejercicio

Resumir la tabla anterior. Hay más de 200 países. Sólo 24 se muestran en esta vista [87]. ¿Cómo lo haría? Dibuje soluciones. Tiempo 5 minutos.

Solución

Si nos fijamos en la encuesta de 2018 y la comparamos con la del año 2017, hay un aumento de 1145 nuevos encuestados que se identifican como “científicos de datos”. Entonces, ¿De dónde vienen los nuevos usuarios de datos científicos? La figura 34 muestra los datos de 2018 y 2017. Sin embargo, ¿Hay demasiados países para que una persona le encuentre sentido! Recuerde que la memoria de trabajo de un cerebro humano se limita a 5-7 ítems [48]. Esto significa que no podemos hacer malabares con más de siete países al mismo tiempo y su gráfico tampoco. ¿Qué haría Marie Kondo? [38]



***Fig 35** Este cuadro fue hecho con PowerPoint porque era más rápido que ajustar los parámetros de ggplot. Observe cómo se utiliza la proporción áurea través del gráfico.*

Use Mega Regiones

Un modo de resumir de una forma más humana es agrupar los países en regiones mega económicas que tienen un cierto grado de homogeneidad (cultural o económica). En este caso, se aplica el marco BRICS. BRICS significa Brasil, Rusia, India y China. Es una agrupación acuñada por Goldman Sachs para clasificar a países con indicadores económicos similares bajo una etiqueta. (El término “BRIC” fue acuñado en 2001 por el entonces presidente de Goldman Sachs Asset Management, Jim O'Neill [16]). Este marco proporciona una manera significativa de agrupar países reuniéndolos por afinidades en el desarrollo económico y social. En este caso, hemos elegido cuatro grupos: los EE.UU., Europa, BRICS y el resto del mundo. Cuando hacemos eso, vemos que no sólo es BRICS el principal contribuyente al crecimiento con el 42% del total en 2018, porque fue también el de más rápido crecimiento entre los tres grandes. En 2018, en la categoría “usuarios que se definen a sí mismos como científicos de datos”, Europa añadió 302 usuarios, 131 Estados Unidos, el resto del mundo se sumaron 231 y BRICS añadió 481. En 2020 el crecimiento de los BRICS superará en número a Europa y los Estados Unidos combinados. Fuente de datos: hicimos una horquilla y modificamos un fragmento del código de [15] y tomamos los primeros 20 países cuyos encuestados se identificaban como “científicos de datos” [17]. UE-6 significa los 6 primeros países europeos.

Estética

Tenga en cuenta cuán diferentes tamaños de fuente son utilizados en el gráfico. Cómo el diseño horizontal neutraliza la metáfora gravedad. Y cómo

se utilizan banderas y etiquetas. El crecimiento de los países individuales sumados debe llegar al 100%, ¿Cuál sería la narrativa en la que caeríamos si el autor hubiera utilizado un gráfico circular o de barras?

Narrativa

La narrativa aquí es ser sabios, uno debe mirar donde la pelota va para no estar donde está ahora. No mire los números absolutos de 2018 solamente, mire el crecimiento. Por supuesto, una narrativa alternativa es que los países emergentes están ganando terreno en campos estratégicos tales como la Ciencia de Datos [71].

Cuestionario: Resúmenes visuales

¿Verdadero o Falso? Tiempo 5 minutos.

1. La visualización de 200 nombres de países en una pantalla 800 x 1200 es ridículo porque no hay suficientes píxeles. [Verdadero / Falso]
2. La visualización de 200 países en un gráfico viola la regla del límite de 7 trozos. [T / F]
3. El uso de PowerPoint o Adobe para construir un gráfico que uno necesita es profesional. [T / F]
4. Para visualizar la situación “un ganador se lleva todo” podemos utilizar un gráfico de Marimekko. [T / F]
5. La Casa de Shiva es la mejor manera de visualizar las relaciones de todo o nada. [T / F]

(Solución en la página siguiente)

Solución

1. Visualizar un gráfico con 200 nombres de países en una pantalla 800 x 1200 es ridículo porque no hay suficientes píxeles [Verdadero / Falso]. Falso. Es ridículo porque un gráfico con 200 etiquetas de texto es información sobrecargada.
2. Visualización de 200 países en un gráfico viola la regla del límite de 7 trozos [V / F]. Cierto.
3. El uso de PowerPoint o Adobe para construir un gráfico que uno necesita es profesional [V / F]. Cierto. Así es como se construyen los gráficos en Bloomberg Businessweek.
4. Para visualizar la situación un 'ganador se lleva todo' situación podemos utilizar un gráfico de Marimekko [V / F]. Falso. Podium 10x es más clara en este caso.
5. La Casa de Shiva es la mejor manera de visualizar 'las relaciones de todo o nada' [V / F]. Cierto. Si falta una columna el techo cae.

Capítulo 4. Gráficos para pensar

- Del Conocimiento a la Sabiduría -

Hasta ahora, hemos visto ejemplos de cómo transformar los datos en información y la información en conocimiento. Ahora consideremos análisis prescriptivos - los gráficos que los Gobiernos y los directorios utilizan para la toma de decisiones en el diseño de políticas públicas y también usados en la política interna de las compañías. Por desgracia, los desafíos políticos son difíciles de lidiar. En particular, cuando las variables involucradas tienen dependencias inesperadas que no se comprenden adecuadamente [70] o incluso no se conocen.

Una manera de hacer frente a estos problemas complejos es simplificarlos mediante la descripción en una forma más sencilla. Una manera popular es transformar el problema en un mapa: un proceso llamado cartografía. El mapeo o cartografía se puede hacer si nos limitamos a usar dos dimensiones para describir el problema. Una vez en 2D, el lienzo se convierte en un espacio de diseño en el que podemos buscar soluciones. Debido a que estamos en un entorno graficado en un mapa, podemos aprovechar las extraordinarias capacidades cognitivas espaciales con las que hemos sido dotados [82-85]. Gloriosos ejemplos de “mapeo” son las salas de guerra, la matriz de participación de crecimiento BCG, diagramas de Gantt, el Modelo de Negocio de Canvas, tableros Kanban, el Cuadrante Mágico de Gartner y los mapas Wardley, hasta la fecha, la herramienta de pensamiento más avanzado para el diseño de innovación estratégica [74 , 75].

Vamos a empezar con la cartografía 2D. Imagine que usted es un alto funcionario a cargo de la Innovación en Singapur. Solo le han dado datos de la figura 34 (número de científicos de datos por país). ¿Cómo utilizaría estos datos para informar sobre la siguiente política de innovación a implementar en el país? Pista: Un primer paso es crear conciencia situacional [75]. Una manera es clasificar los países, para que podamos ver dónde estamos. El segundo paso es utilizar un indicador significativo en el eje X. (Solución en la siguiente sección).

Utilice el ranking para crear conciencia de la situación

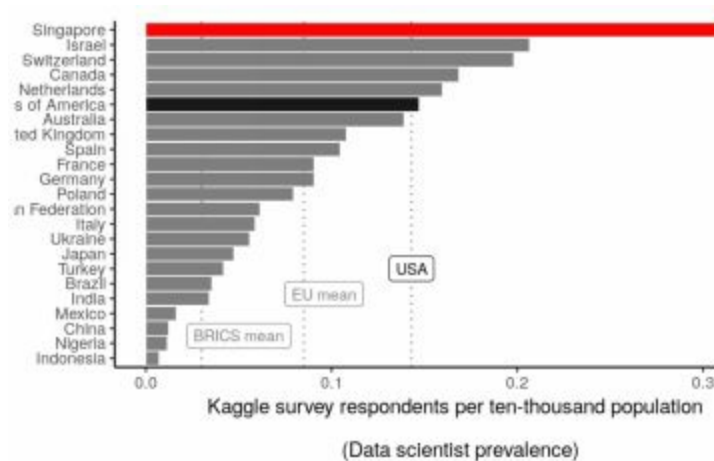


Fig 36 *Dos es la máxima cantidad de colores que puede usar en un gráfico. El gris no cuenta.*

Use medidas per cápita

Como el economista y comediante Harald Eia insinúa cortésmente en su conferencia TEDx Talk en Oslo [29], al comparar los países, casi nunca estamos interesados en números absolutos. De hecho, cualquier medida que no sea per cápita es bastante inútil. Sin embargo, ¿con qué frecuencia nos olvidamos de esto! Los gráficos no normalizados por población del país se están preparando para los estereotipos [60] y narrativas poco saludables [29]. A partir de la encuesta de la Ciencia de Datos 2018 Kaggle, es posible contar el número de científicos de datos en cada país, y la mayoría de los participantes rindió tales gráficos. Pero ¿qué tan útil es esto? Ya sabemos que China y los EE.UU. son países grandes, por lo que en números absolutos también tendrán la mayor cantidad de científicos de datos. Esto es lo que se espera. Sin embargo, si es lo esperado significa que ya lo sabíamos. Cuanto más esperado, menos información: 100% esperado significa cero información). Lo que es útil, es mirar a las medidas per cápita [60]. ¿Cómo se comparan los países en la densidad de científicos de datos? En este gráfico, destacamos Singapur en rojo y los EE.UU. (el hogar de la comunidad encuestada más grande) en negro, por lo que el lector tiene un punto de referencia. Este gráfico tiene mucho que explicar:

- EE.UU. media: 0,14 por 10.000
- UE6 media *: 0,09 por 10.000
- BRICS * media: 0,03 (5x menos que US)

Cuidado con la brecha

Cuidado con la brecha es una estrategia común para pensar acerca de las diferencias entre las categorías en los datos, en este caso, los países. Pensar acerca de por qué existe la brecha puede ayudar a explicar la realidad que el gráfico representa. Por ejemplo, un lingüista podría pensar en la brecha en términos de dominio del inglés y su correlación con la prevalencia en científicos de datos. ¿Es la barrera del idioma un factor explicativo de la brecha? ¿Cuáles son las implicaciones de la política? Nota: los BRICS y UE6 representa lo que significan los países no ponderados por los encuestados. Fuente: Banco Mundial de Datos de Población 2016, encuesta Q11. País de residencia actual. [19-21, 27].

Estética

Esta combinación de colores se llama roja sobre gris y es mi esquema preferido para los gráficos. A diferencia de otros esquemas tales como púrpura sobre gris, este es de género neutro [23]. Sin embargo, para que funcione, la superficie roja debe mantenerse al mínimo, de lo contrario, aparece como estridente.

Narrativa

Ventaja Competitiva de las Naciones de Porter [61].

Ejercicio

Vamos a dar un paso más en el gráfico de la Fig. 36. Anteriormente vimos que los marcos pueden ayudar al lector a dar sentido a nueva información. Utilizando el *Índice de Innovación Global*, ¿cómo relacionarlo con el predominio de los científicos de datos? Utilizar mapeo 2D. Tiempo 5 minutos. (Solución en la siguiente sección).

Proyectar en 2D

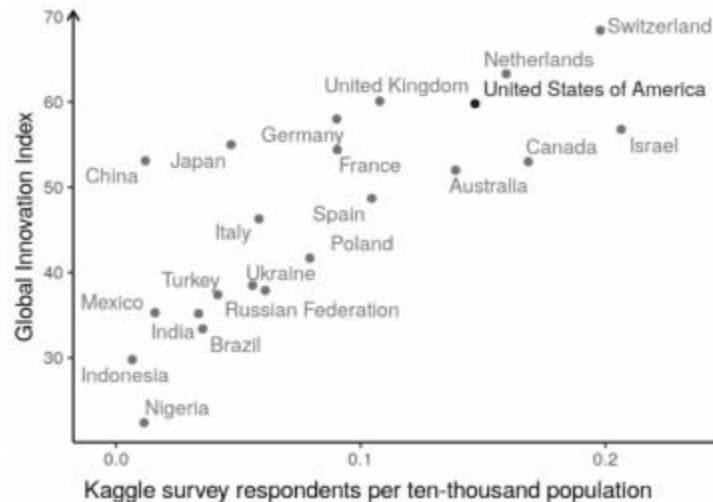


Fig 37 Un diagrama de dispersión entre dos variables correlacionadas siempre produce una nube similar.

Como el Sr. Wardley diría, “cuándo es necesario comprender el territorio ayuda tener un mapa”. Aquí se utiliza la cartografía 2D dispersando los países a lo largo de dos dimensiones^[12]. Este mapa se puede utilizar para agrupar países políticamente para ayudar a dilucidar los factores de éxito que influyen en la posición en el mapa [60].

Narrativa

Ventaja Competitiva de Porter de las Naciones [61].

Sobre el Índice de Innovación

Cada año, el INSEAD MBA, la Universidad de Cornell y la OMPI publican el Índice Global de Innovación [20]. En 2018, el país más innovador fue Suiza. Una correlación de rango de Spearman [23] del 79% entre GII y rendimientos de prevalencia del usuario.

Ejercicio

Vamos a tomar esta tabla y llevarla un paso más allá. Una de las habilidades más valiosas es la predicción. Dada la figura 37, ¿se puede predecir dónde estará Japón dentro de 10 años? Utilice una regresión lineal. Tiempo 5 minutos. (Solución en la siguiente sección).

Prever con reversión a la media

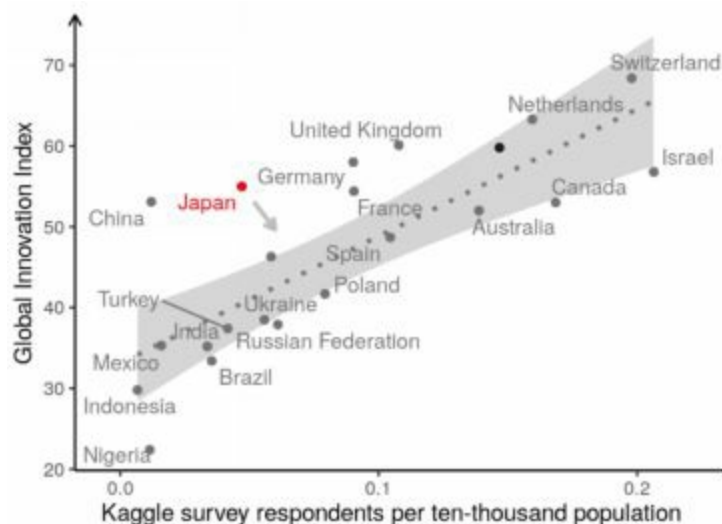


Fig 38 Reversión media, ¿siempre correcta en el largo plazo?

Aquí, acabamos de agregar una línea de regresión y hemos eliminado el valor atípico de Singapur. El 95% del margen de error estándar se muestra en gris. Algunos países están por debajo y otros por encima. Resaltado en rojo está Japón, como un valor atípico con un alto índice de innovación (eje Y), pero baja en el eje X con respecto a sus compañeros. Asumamos que el principio de reversión media se aplica aquí como un predictor de línea de base y una “mano oculta” empuja continuamente los países hacia la media [24] (línea de puntos). El principio de reversión media se basa en la idea de que no hay ventajas competitivas permanentes para las empresas [98] o naciones. Se ha demostrado su eficacia, sobre todo en las finanzas. Por ejemplo, en las apuestas en la composición del Dow Jones, muy pocas empresas tienen lo que se necesita para durar mucho tiempo en el Dow Jones. De los miembros originales del índice formado en 1896, sólo se mantiene GE.

Reflexión

¿Qué podemos pronosticar acerca de la posición de Japón en el ranking GII 2019? Aplicando el principio de reversión media, es poco probable que

Japón aumente su rango porque es ya alto. Incluso si Japón se pone al día en la prevalencia de científicos de datos, lo más probable es que todavía vaya a ir hacia abajo (hacia la media). Los índices son solo valores. Suponiendo que el valor de los Datos Científicos en la innovación de la economía sólo aumentará en las próximas décadas y que el método de cálculo del índice GII se actualizará en consecuencia, ¿qué países son más propensos a mejorar su ranking “nominal” en 2019? Cuando los valores del índice GII se reajusten, ¿es probable que países como Canadá, Australia salten algunos puestos? Fuente: Índice de Innovación Global 2018, Banco Mundial de Datos de Población de 2016, Q11 - País de residencia actual [19-21, 27] Para análisis residuales, véase [25].

Una nota sobre el origen de la regresión lineal

El nombre de regresión lineal como la línea que minimiza la suma del cuadrado de los errores, fue popularizado en un papel donde se verificó que el principio de “regresión a la media” en la altura de la descendencia está relacionada con la altura de los padres. ¡Alerta de spoiler! Sólo el 60% de la altura de la descendencia se explica por la altura de los padres. El resto se explica por la media de la carrera. Lo que significa que el principio reversión a la media aplicado a la altura, tiene una influencia del 40% aproximadamente. Sin embargo, el método matemático es completamente ajeno a cualquier concepto de regresión. El documento se hizo famoso y la palabra regresión se quedó pegada al método. Una gran “pregunta” consiste en pedir a los estudiantes que expliquen por qué se llama regresión lineal a una regresión lineal. Siempre me sorprende la creatividad de algunos estudiantes [62].

Representar la estrategia empresarial en dos dimensiones

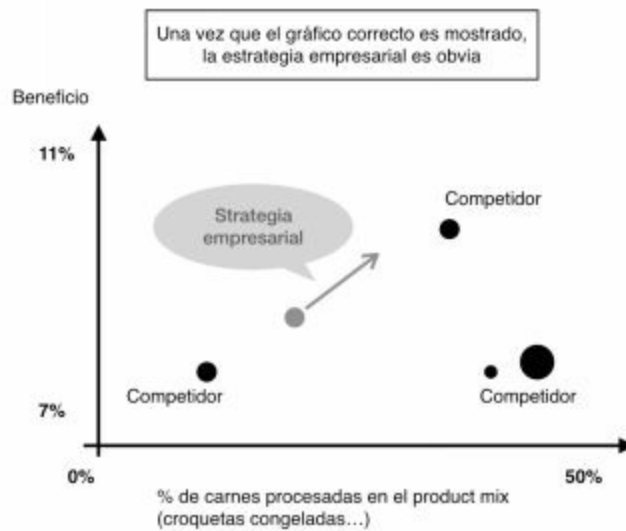


Fig 39 ¿Se puede visualizar un problema como un mapa?

Reformular el problema

Uno de los papeles más importantes para un científico de datos, es darse cuenta cuando el cliente no puede articular sus propias necesidades (véase la teoría Trabajos-por-hacer [39]). Esta habilidad es la que distingue al científico de datos A + del resto. El gráfico aquí es una adaptación del libro *The Accidental Investment Banker*. Al autor, un banquero, se le ocurrió durante una cita de negocios. Lo utilizó para trazar la estrategia M&A para un cliente [40]. Una vez que hizo esta gráfica, todo el mundo en la habitación podría visualizar dónde estaba valor. En su libro, él acredita esta tabla como un momento importante en su carrera.

La Matriz GAP

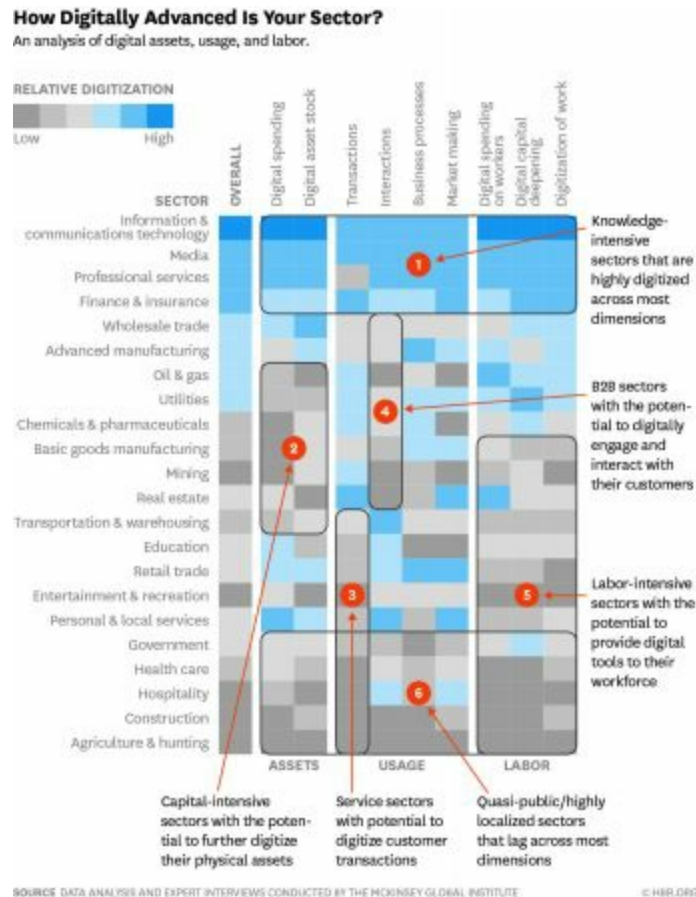


Fig 40 La innovación empresarial es a veces tan fácil como encontrar un espacio en blanco. Fuente: McKinsey Global Institute, [93]

Encontrar brechas en el mercado

Otro uso del espacio de diseño 2D es encontrar segmentos pasados por alto o desatendidos. Podemos aplicar este método para matrices también. Ejemplos de ello son una matriz de habilidades de empresa-empleado, planes de desarrollo tecnológico, hojas de ruta tecnológicas y la matriz de innovación. En la figura 40, McKinsey tomó una vista sectorial. Sin embargo, otros puntos de vista pueden producir descubrimientos también. Por ejemplo: en lugar de ver por sectores podíamos ver por el proveedor de software (Salesforce, JIRA, Autodesk ...).

La tabla periódica

Fig 40b ¿La visualización más útil en la historia de la Ciencia [\[13\]](#)?

Una famosa aplicación de la Matriz de Gap de 1869 es la Tabla Periódica de los Elementos de Mendeleev. En 1869, hace 150 años, Dimitri Mendeleev publicó una tabla periódica de los elementos químicos en base a las propiedades que aparecen con cierta regularidad, ya que él presentó a los elementos del más ligero al más pesado. Cuando Mendeleev propuso su tabla periódica, señaló lagunas. En ese momento, sólo se conocían 63 de los 118 elementos conocidos en la actualidad. Luego predijo las propiedades de los cinco elementos desconocidos - un golpe de genio para cualquier joven científico [122]. Las predicciones visuales de Mendeleev impulsaron una carrera de descubrimiento. Sorprendentemente, nunca recibió el premio Nobel - un testimonio de rivalidades científicas' sino también es un clara ejemplo de cómo el pensamiento visual ha sido históricamente despreciado por el mundo académico dominado por un pensamiento orientado a palabras.

La Matriz de Innovación

La Matriz de Innovación es una herramienta de descubrimiento del conocimiento - Leo Tschirsky

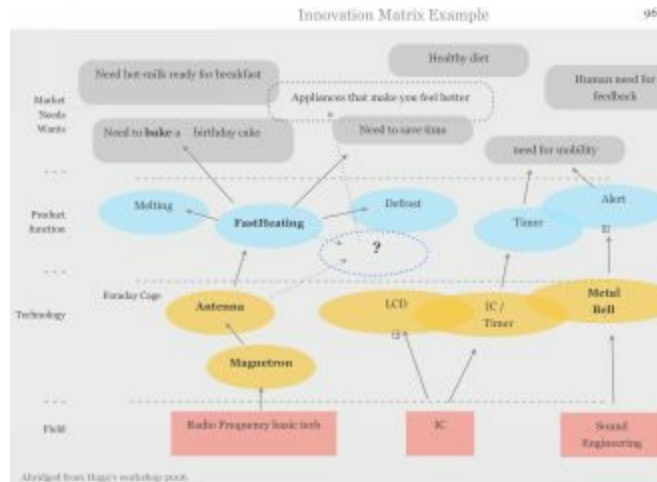


Fig 41 La Matriz de Innovación (IM). Fuente: [94]

Leo Tschirsky, Profesor Emérito de Dirección de Empresas en el Instituto Federal Suizo de Tecnología (ETH), facilitó este taller sobre el horno de microondas en el Instituto de Tecnología de Tokio en 2006. El **IM** ayuda a formalizar y organizar las relaciones funcionales entre:

- Necesidades del mercado
- Funciones del producto
- Tecnología
- Campos de la ciencia

Cómo utilizar

1. Dibuje un formato de matriz por filas

- Escribir las necesidades del mercado (Por qué la gente compra hornos)
- Funciones del producto (calentar, hervir ...)
- Tecnologías detrás de esas funciones (Magnetron, LCD ...)
- Campo de la ciencia básica que apoya esas tecnologías. (RF, IC)

2. Enlazar conceptos con flechas

- Aclarar
- Buscar verdades profundas
- Usar 5 “por qué” del análisis de raíz de la causa de la teoría empleos-por-hacer.

Una vez que su producto esté claro y trazado...

3. Innovar

Ahora que tiene una idea clara de las relaciones entre los valores, las necesidades del cliente, los costos y la tecnología, usted está en una mejor posición para innovar usando una variedad de técnicas, tales como:

- Lluvia de ideas
- La planificación de un taller al estilo “Carro de compras de Ideo”
- El uso de herramientas de creatividad de Edward de Bono
- Encontrar lagunas
- Cubrir las nuevas necesidades con funciones existentes

Ejercicio

Grupos de cuatro. Tiempo 20 minutos. Piense en este microondas y sus componentes ...



Fig 42 El ciclotrón es el elemento que genera las microondas. Fuente: [94]

Usted ha sido contratado por una marca de horno de microondas. Recientemente, debido a la competencia china, los márgenes del horno son finos como el papel, por lo que la supervivencia de la empresa podría depender de que usted presente un nuevo diseño de horno por el que los clientes quieran pagar más. **Requisito:** Utilice la Mensajería Instantánea para innovar en el microondas.

Solución

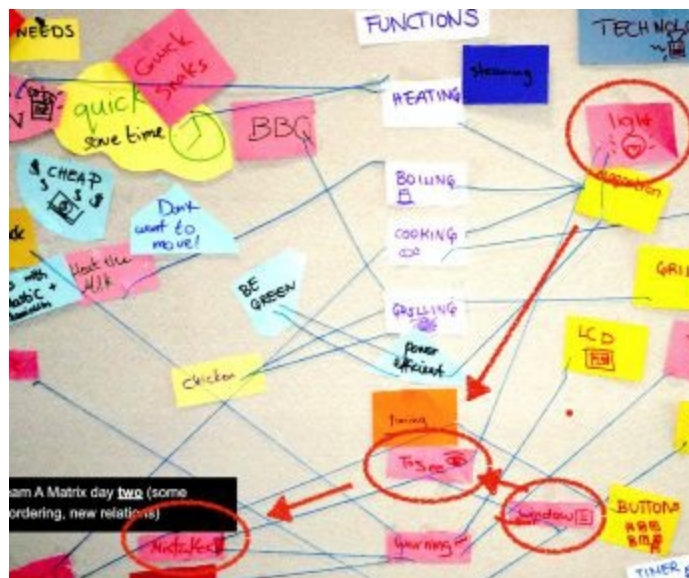


Fig 43-44 Innovación. Ejemplo funcional con post-its y cuerdas por los estudiantes de la clase de Design Thinking. Fuente: [94]

La luz y la ventana del horno parecen dos tecnologías no relacionadas en un horno. Sin embargo, responden a una necesidad común: ¡la necesidad de comprobar si hay errores! **¿El costo de (Ventana + Luz) > costo de (X)?** Considere el costo de la ventana + luz. Su único propósito está claro ahora. ¿Puede X hacer su función mejor y de forma más barata? ¡Vamos a encontrar la X!

Mapas de Wardley

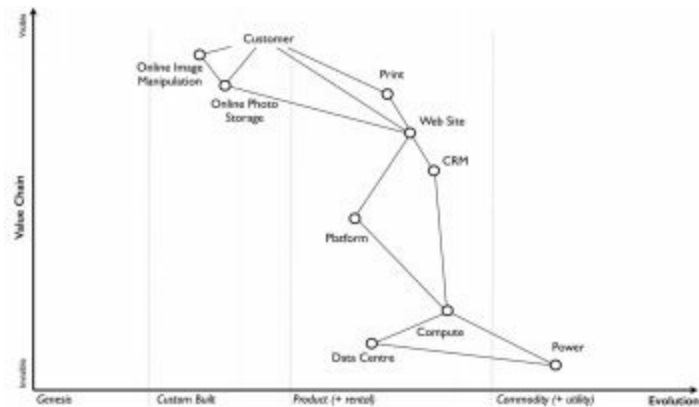


Fig 45 Un mapa Wardley, Simón Wardley CC BY-SA 4.0.

Un mapa Wardley puede ser pensado como un mapeo 2D de Comoditización vs Cadena de Valor, o como una evolución de la Matriz de Innovación de Leo Tschirsky. En cualquier caso, cuando hay dilemas sobre si externalizar o no, este es el camino para manejarlo en 2019. Un excelente tutorial se puede encontrar en [95]. Hasta ahora hemos visto un par de herramientas de visualización utilizadas en los negocios y en la formulación de políticas públicas. Sin embargo, la visualización también se utiliza en el entrenamiento. Veamos un ejemplo llamado la Rueda de la Vida.

Rueda de la Vida

“Una carta para pensar acerca de sus objetivos de vida”

Amy es una mujer que trabaja, madre soltera, que acaba de renunciar a su trabajo en una empresa Fortune 500. Una exitosa mujer de carrera por derecho propio, Amy simplemente dejó su ascendente función ejecutiva porque quería viajar menos y pasar más tiempo de calidad con su hija adolescente de 14 años. Estoy sentado con Amy. Nos hemos reunido para hablar de la vida, la carrera y los gráficos, sí ¡los gráficos! Amy está a punto de mostrarme un gráfico que nunca he visto antes. Ella lo llama la **Rueda de la Vida**. Este gráfico, me dice, fue fundamental para ayudarle a darse cuenta de que tenía que cambiar su vida.

Cómo utilizarla



Fig 46 Comience con un punto en medio de un papel.

En primer lugar, se dibuje radios. Ocho radios de una rueda. Cada radio representa una categoría diferente de su vida y le ayudará a medir su satisfacción en cada área de su vida. El primero es el Dinero (¿Cuál es su grado de satisfacción con el dinero que ha ahorrado/ganado?) En segundo lugar, la Carrera (¿Cuál es su grado de satisfacción con su trayectoria, el progreso y la carrera actual?). En tercer lugar, el Bienestar (tanto espiritual, mental y físico). Luego siga con Amigos y Familia, Amor, Diversión y Entorno ¿Le gusta el país, ciudad/casa/ vecindario que está? Y finalmente, el crecimiento espiritual y personal. Ponemos calificaciones en cada categoría marcando un punto en el radio en una escala de 1 a 10, siendo 1 el centro y 10 la parte más alejada del centro, y luego conectamos los puntos.

Conectando los puntos



Fig 47 'Citizen Kane' - tipo de Rueda de la Vida [102].

Al conectar los puntos de mi gráfico, un escalofrío recorrió mi espalda. ¡Mi rueda de la vida no es redonda! Parece un asteroide y tiene una gran hendidura donde está el radio del amor. ¿Es esto malo? “Oh, bueno” Creo que, asustado, asumo que necesito re priorizar mi vida.

Una radiografía de su vida

A continuación, Amy se inclina y dice: “Este es el resumen de tu vida. Aprendí este ejercicio en un inicio de clase en Stanford. Resultó ser una de las cosas más útiles que aprendí allí porque me ayudó a visualizar mis **puntos ciegos**. Muchos de mis compañeros hicieron esta tabla también, y estaban tan sorprendidos como tú lo estás ahora...” Luego agrega: “El comentario típico de mis compañeros era: Aquí estoy, centrado en mi carrera y desarrollo personal, como lo he hecho toda la vida, mientras que partes iguales o más importantes de mi vida están siendo descuidadas”. La Rueda de la Vida, junto con otras herramientas de visualización tales como el “*The Good Time Journal*” forman parte del **diseño de tu tendencia de vida** donde los principios de diseño de iteración, visualización y el pensamiento de diseño en general se aplican al coaching. Los resultados son espectaculares. Para aprender más, el libro del co fundador de Atari “El diseño de su vida: ¿Cómo construir una vida alegre bien vivida” es un excelente inicio? (96)

Ejercicio

Haga su Rueda de la Vida. ¿Cómo haría esta tabla más convincente que la de la figura 47? Tiempo 7 minutos. (Solución en la página siguiente).

Solución

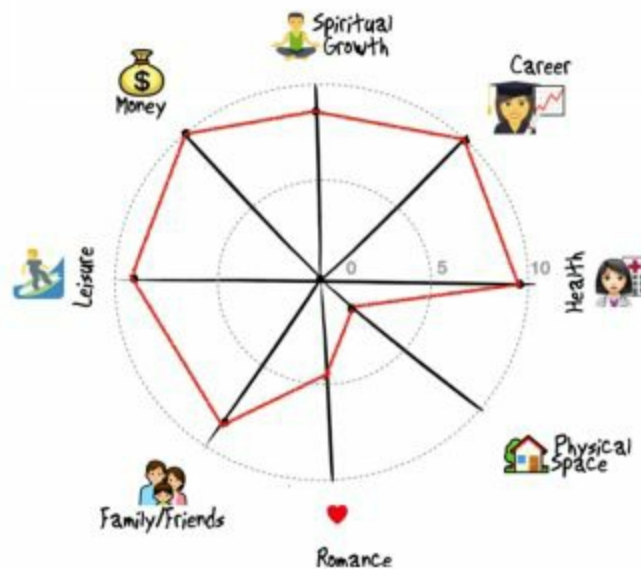


Fig 48 Los iconos se utilizan para reducir el tiempo necesario para entender un gráfico.

Los íconos y emoji son un recurso infrautilizado en la realización de gráficos. Por otro lado, el uso de emoji se correlaciona con la participación de los empleados [106].

Exploración interactiva

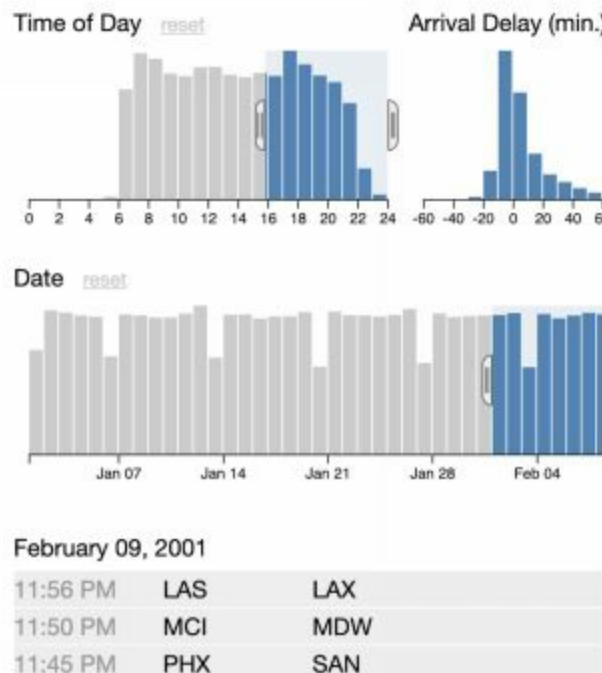


Figura 49 Crossfilter puede hacer operaciones SQL interactivas. Fuente: Square 2001. [97]

Filtro cruzado de Square

Crossfilter es una biblioteca de JavaScript para explorar grandes conjuntos de datos multivariados en el navegador. Extremadamente rápido (<30 ms), que permite la interacción “umbral Doherty” con vistas coordinadas, incluso con conjuntos de datos que contienen un millón o más de registros; Square lo construyó en 2001 para análisis de potencia para Square Register.

Ejercicio

Grupos de dos. Tiempo 12 minutos. Utilizando una demostración de Crossfilter [97], encuentre tres puntos de vista sobre los viajes aéreos. Ejemplo: Para evitar retrasos en los vuelos por la mañana. Visualizar sus resultados.

Cuestionario: Decisiones de apoyo

¿Verdadero o Falso? Tiempo 10 minutos.

1. Trabajar con medidas per cápita es la única manera racional de comparar países. [Verdadero/Falso]
2. El principio de reversión a la media establece que, en el largo plazo, una “mano oculta” empuja los valores atípicos hacia la media. [V/F]
3. La matriz de separación se utiliza principalmente para descubrir nicho o vacíos en el mercado. [V/F]
4. Los mapas Wardley han hecho obsoleta la Matriz de Innovación. [V/F]
5. La biblioteca Crossfilter es una opción apropiada para visualizar las relaciones lineales entre dos variables, tales como retrasos y la hora de salida. [V/F]

Solución

1. Trabajar con medidas per cápita es la única manera racional para comparar países. [Verdadero Falso]. Cierto.
2. El principio de reversión media establece que, en el largo plazo, una ‘mano oculta’ empuja los valores atípicos hacia la media [V / F]. Cierto. No hay ventajas competitivas de larga duración [98].
3. La matriz de separación se utiliza principalmente para descubrir nicho o vacíos en el mercado [V / F]. Cierto. Sin embargo, también se puede utilizar en otras áreas.
4. Los mapas Wardley han dejado obsoleta la Matriz de Innovación [V / F]. Falso. La Matriz de innovación sigue siendo mejor que Wardley para hacer innovación de productos y para asignar un producto a su mercado.
5. La biblioteca Crossfilter es una opción apropiada para visualizar las relaciones lineales entre dos variables, tales como retrasos y hora de salida [V / F]. Falso. Un diagrama de dispersión es más claro. Crossfilter es ideal para la EDA.

Capítulo 5. Crear gráficos que destaquen

- Consejos para que sea viral -



Fig 50 *Manos de jazz, para crear una visual atractiva.*

En este capítulo, vamos a ver estrategias de piratería para hacer su gráfico memorable. Ahora que conocemos los principios para hacer un gráfico significativo (mediante la reducción de la sobrecarga de información, vinculación a marcos, narrativas apropiadas y el apoyo de las metáforas visuales), vamos a ver sencillos consejos sobre la manera de animarlo o en la jerga del Design Thinking, hacerlo “pop” [12]. Pop significa que salta a la vista.

Mis recursos favoritos en esto son: los libros de Dan Roam, “La Información es bella” por McCandless, y la cuenta de Instagram “Chartr”. ¡Tenga cuidado! Algunos gráficos son tan hermosos por sí mismos que son como una obra de arte por derecho propio. En cualquier caso, recuerde, el gráfico nunca debería ser el personaje principal de su historia, lo que importa es la narrativa (mensaje) y lo bien que se comunica a su público. Desafortunadamente, las visualizaciones estéticamente agradables y una visualización que cumple con su trabajo no siempre coinciden. Por cierto, esta es una de las razones por las que, en una agencia de publicidad, verá al director de arte y al director de cuentas siempre en desacuerdo. El director de arte quiere ir al Festival de Cannes para ganar un León de Cannes como premio a su Creatividad, mientras que el gerente quiere un anuncio efectivo

[41]. Vamos a ver las principales técnicas a utilizar para hacer un gráfico “pop”.

Utilice flechas para desatar su pensamiento

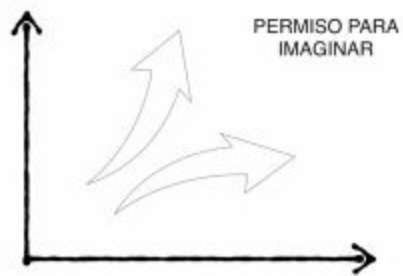
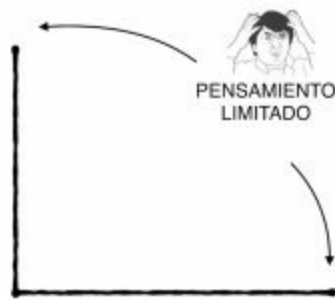


Fig 51 Las flechas le dan permiso para pensar más allá.

Despejando gráficos

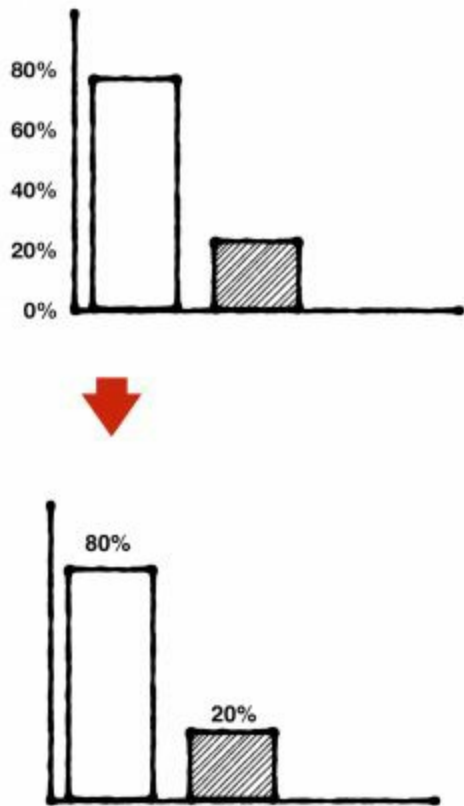


Fig 52 Marie Kondo aplica el principio de tirar cosas en la vida que no irradian alegría [45].

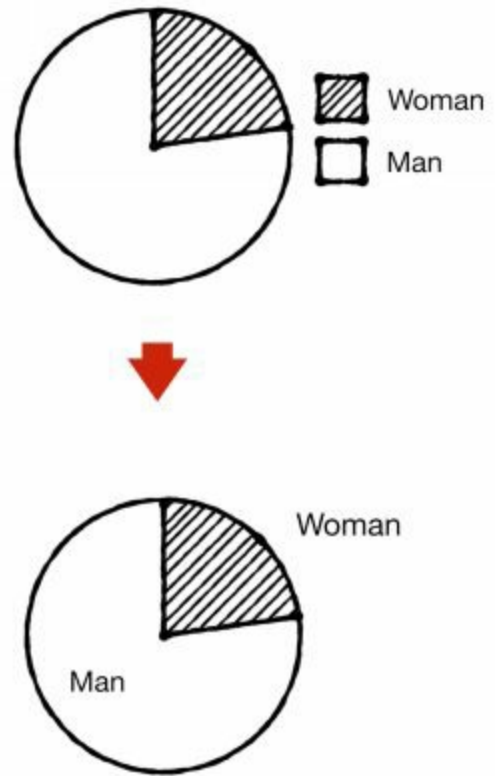


Fig 53 Animación para limpiar el desorden de un gráfico circular
bit.ly/2OgCLUO

Use personas para ganarse la audiencia

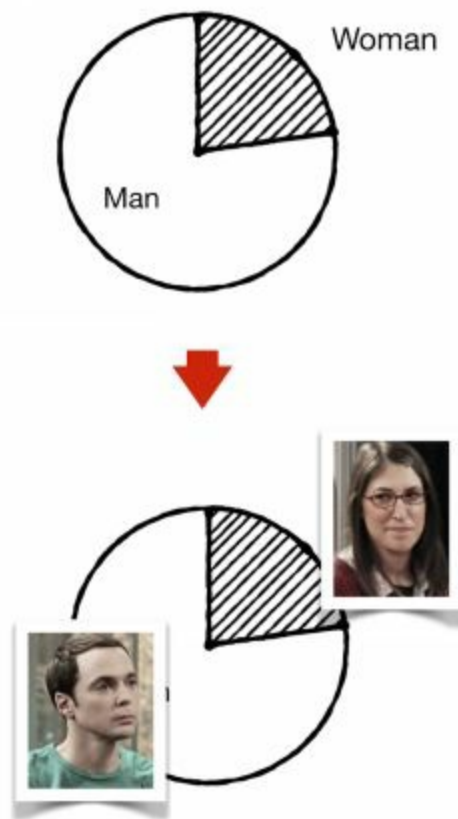


Fig 54 Las personas son adeptas al reconocimiento de rostros.

¿Sabía que podemos reconocer una cara más rápido que muchos otros objetos del mundo? [42] ¡Utilícelo! En 2007 Honda utilizó este principio cuando diseñaron una moto que, desde la parte de atrás, parecía una cara humana (antropomórfica).

Visualizando grandes diferencias

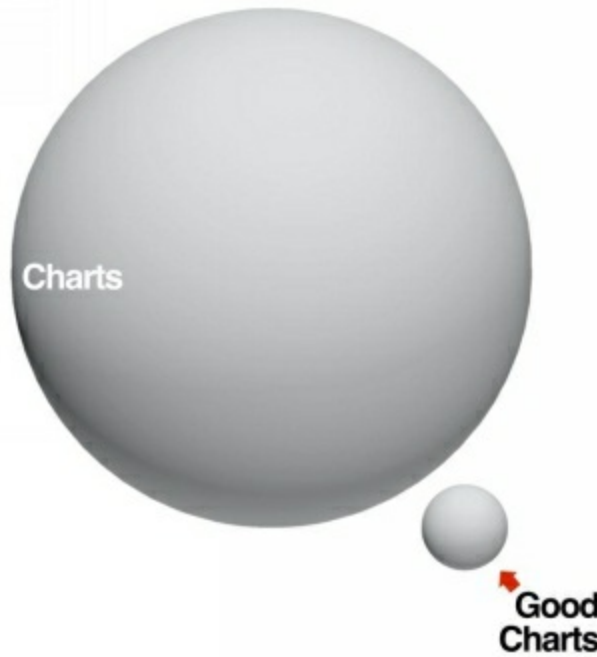


Fig 55 Este gráfico tiene una gama asombrosa de cuatro órdenes de magnitud.

¿Cómo calcular el rango dinámico en un gráfico del planeta? En un gráfico del planeta, el rango dinámico es n^3 , donde n es las veces que el radio del planeta pequeño cabe en uno grande. En la figura 56, $n = 12$. Por lo tanto, $n^3 = 12^3 = 1728$. Por lo tanto, la parte del pequeño planeta es de aproximadamente 1 en $(1728 + 1)$, 0,05%.

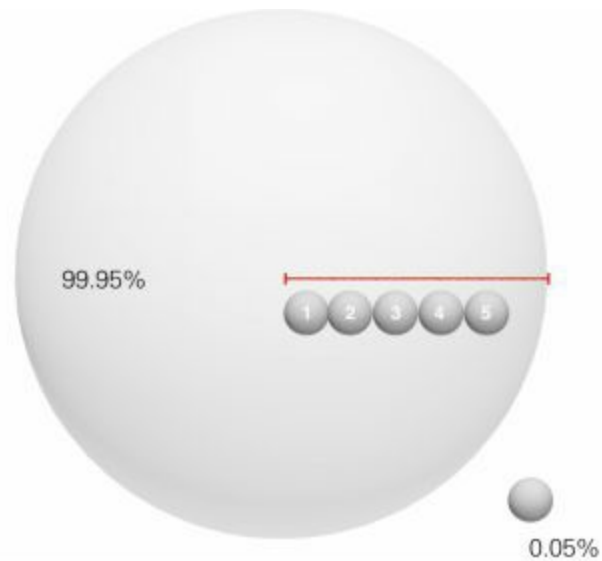


Fig 56 El pequeño planeta encaja 1720 veces en el volumen del grande.

Otros intentos exitosos para visualizar enormes diferencias son “¿Qué significado tiene para ti PhD” [76], y la película de 1977 “Powers of Ten” [77]; utilizando en ambos la técnica zoom 2D para ilustrar contrastes de tamaño gigantesco. Sin embargo, de todas las metáforas de gráficos, la metáfora del planeta es excelente para visualizar vastas diferencias. Consulte la página siguiente para averiguar por qué.

Poner el dinero global en perspectiva

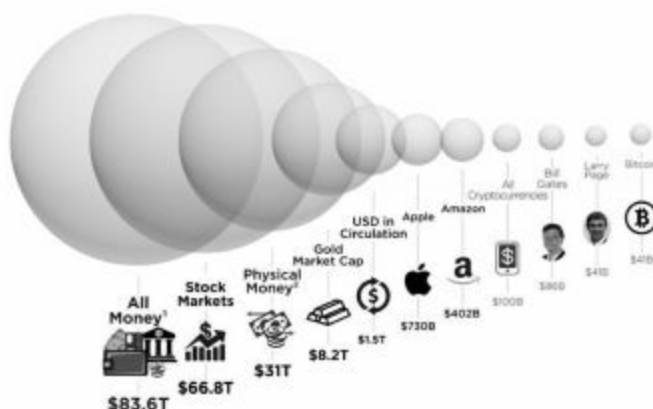


Fig 57 Los seres humanos han evolucionado adaptativamente para estimar la masa desde la altura.

¡La furia dinámica de este gráfico se encuentra en una sorprendente magnitud de cuatro órdenes! Se puede visualizar 41bn y algo 20 mil veces más grande. Este gráfico se combinó con la narrativa del Miedo de Perderse (FOMO) antes del accidente del Bitcoin de enero de 2018. Fuente original: HowMuch.net [78].

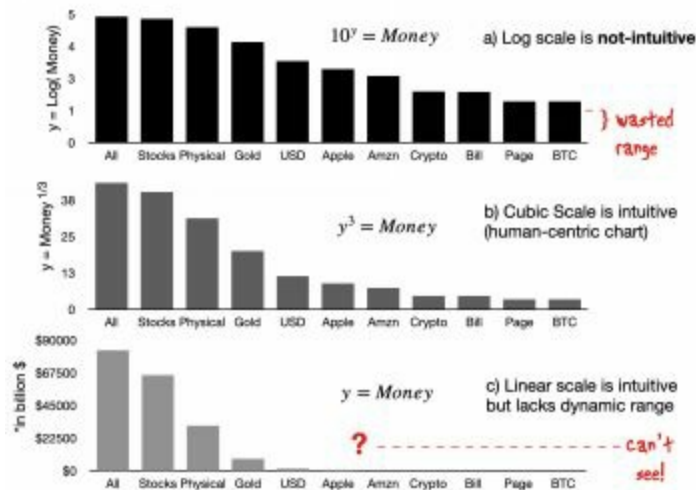


Fig 58 Comparación entre a) Registro de valor b) Raíz cúbica del valor: El radio equivale al peso de la esfera y c) Escala lineal.

En la figura 58, note cómo (a) tiene el rango más dinámico, pero no es intuitivo; (b) es el radio de la esfera equivalente. Sería intuitivo si se expresara como volumen, como en la figura 57; (C) es una escala lineal; es intuitivo, pero carece de rango dinámico.

¿Por qué las esferas 3D funcionan tan bien?

Los humanos hemos evolucionado para estimar el peso de un animal viéndolo. Por supuesto, esto era una habilidad muy útil para nuestros antepasados en la selva. Fíjese lo fácil que es entender tamaños relativos cuando usamos el volumen, en comparación con cualquier otra opción. Para los primates, la estimación visual del peso de un compañero primate era una habilidad de supervivencia fundamentalmente útil para determinar la peligrosidad de la amenaza de un oponente antes del contacto.

La mayoría de las formas de vida tienen una densidad de peso similar, una manera de hacer esto fue mediante la estimación del volumen. Al mismo tiempo, los seres humanos vamos un poco cojos para entender los gráficos de barras cuando las barras difieren en más de 2 órdenes de magnitud. Por suerte, si se muestra en 2D proyecciones de objetos 3D, la mayoría de los seres humanos pueden estimar bien el peso. Esto es útil para comparar magnitudes tan diferentes como de 3 ó 4 órdenes de magnitud sobre una superficie plana, como este libro.

Gráficos de registro

Tenga en cuenta que el gráfico de registro resuelve el problema de la gama

dinámica, pero los humanos no nacen con la intuición logarítmica incorporada (figura 58). En otras palabras, un niño va a entender las bolas, pero se necesitan muchas horas para que estudiantes de licenciatura se familiaricen con parcelas semi-log.

¿Cuántos paneles para abastecer a todo los EE.UU.?



Fig 59 En 2017, Elon Musk utilizó esta plantilla de gráfico para impulsar el uso de Energía Solar. - Fue un fracaso.

En la fig 59 hay una narrativa de escasez. En 2017, Elon Musk utilizó un gráfico como éste. Él estaba abogando por la Energía Solar. Él dijo “solo necesitamos un píxel del mapa cubierto de paneles para dar energía a todo EE.UU., recuerde solo un píxel.” Fue un fracaso. ¿Por qué? Porque conectó a una narrativa de ganar o perder. También es difícil confiar en lo que no podemos ver (un píxel no es una gran visualización). Por desgracia, los gráficos 2D no tienen suficiente rango dinámico para visualizar diferencias de más de 2 órdenes de magnitud. Él estaba tratando de visualizar 4 órdenes.

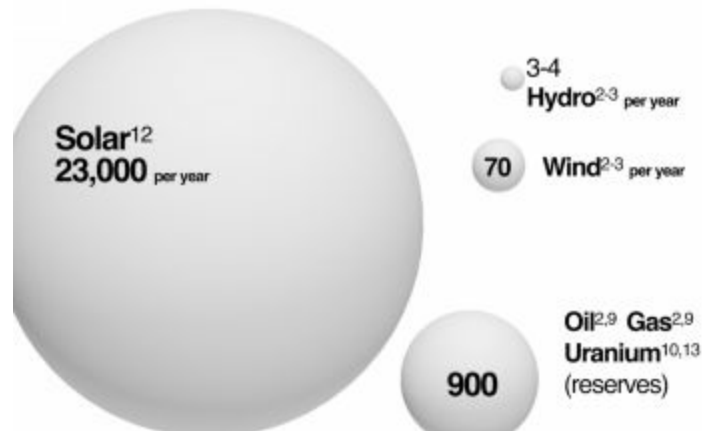


Fig 60 Un gráfico que utiliza la narrativa de mentalidad de crecimiento.
Adaptado de Q-Cells.

La narrativa de la fig 60 es la de la **mentalidad de crecimiento** [68]. Visualiza más de 5 órdenes sin esfuerzo. Alrededor del año 2005, el fabricante alemán de células solares Q utilizó una tabla similar en sus Relaciones Públicas. Esta tabla es más eficaz y confiable que la de la figura 59 porque nos conecta a la narrativa de mentalidad de crecimiento mediante la visualización de la sorprendente abundancia de energía **renovable** disponible.

Sin embargo, tanto Q-Cells como Elon Musk hicieron imágenes a escalas muy gigantescas (USA anual en el caso de Musk, y global anual en el caso de Q-Cells). Quizás se dejaron llevar por delirios de grandeza y perdieron de vista al cliente de a pie. Irónicamente Q-Cells fue a la bancarrota y Solarcity tuvo que ser rescatada a través de una fusión con Tesla. ¿Usando la misma narrativa, como se podría hacer el gráfico de Q-Cells o el de Elon Musk a una escala más humana? Una manera es aplicar el principio de gráficos centrados en el humano. Por ejemplo, usando objetos de escala más humana. Por ejemplo, mostrando una casa y mostrando cuántos paneles son necesarios para ser 100% independientes eléctricamente y en el caso de Q-Cells mostrando una casa y mostrando un tanque con todo el petróleo que una persona consume al día. En Europa se consumen un 0.8 L de gasolina por persona por día y unos 2kWh de electricidad por persona por día. En total equivale a 1L de gasolina o 9kWh por día por persona. Para generar dicha energía con paneles solares se necesita hoy en día una superficie de 8 metros por 8 metros. ¿Qué gráfico le resulta más fácil de entender?

Storytelling del sesgo de edad con humor

La fig 61 se basa en un conjunto de datos de detención proporcionada por Minneapolis Policía y otras fuentes. Fue publicado por Kaggle en 2018 [44]. Se muestra la correlación entre las detenciones sin cargos y la diferencia de edad entre la policía y el sujeto.

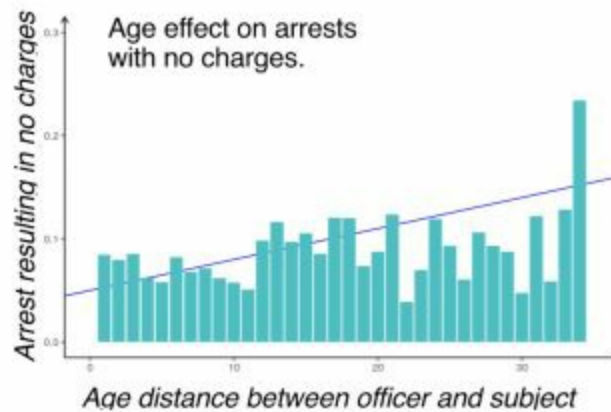


Fig 61 De todos los prejuicios, la edad es uno de los más penetrantes y del que menos se habla.

Núcleo completo: <http://bit.ly/2JXWP9d>. La regresión se muestra en azul. Si uno considera positivo ser acusado con resultado negativo, en este caso preferiría interactuar con un policía de más edad, pero si se tiene en cuenta “un arresto sin cargo” como un resultado negativo, entonces uno preferiría interactuar con un policía más cercano a la edad que tenemos. Cuanto menor sea la diferencia de edad, menor será la probabilidad de que sea inculcado.

Ejercicio

Hacer este gráfico más convincente. Tiempo 3 minutos.

Solución

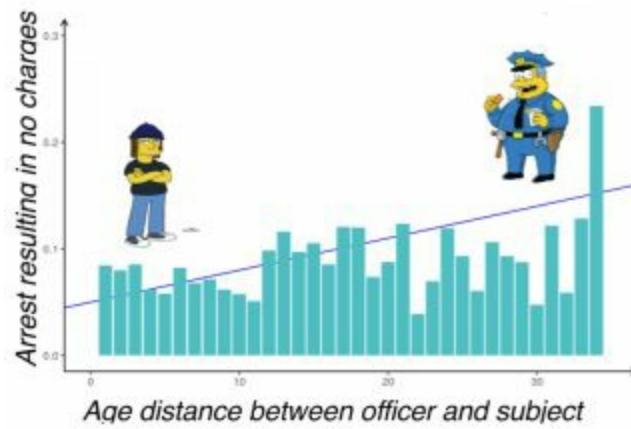


Fig 62 Use Personas para hacer un gráfico memorable.

Utilice la proporción áurea en todas partes

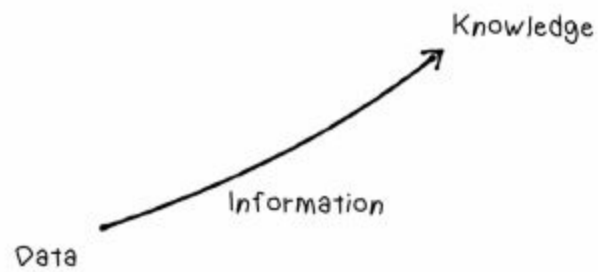
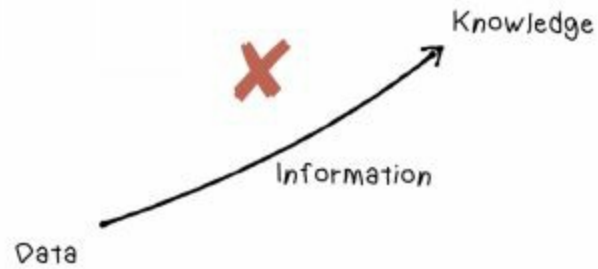


Fig 63 ¿Cuál parece más armoniosa?

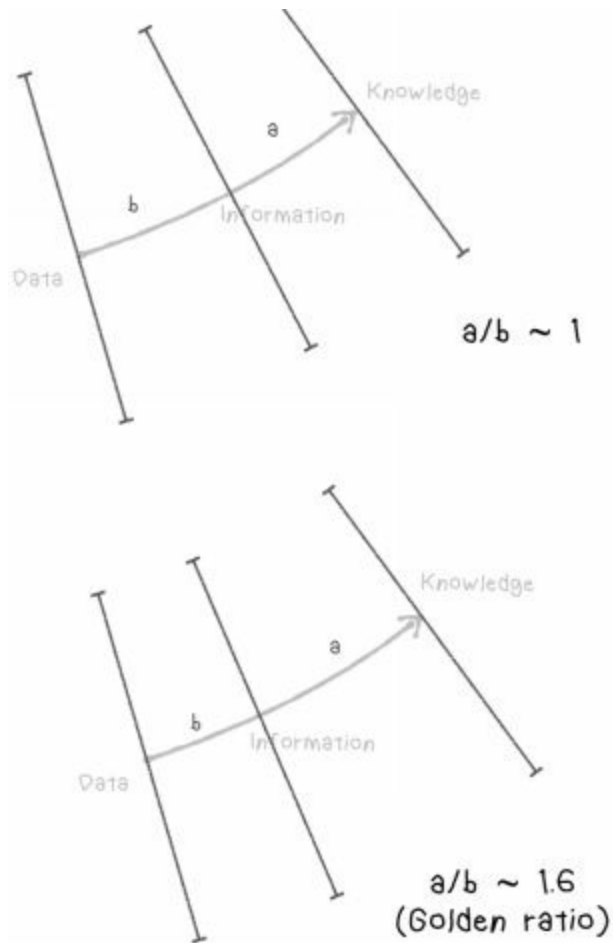


Fig 64 Nunca pierdas la oportunidad de utilizar la proporción áurea.

¿Sabía que si muestra dos tarjetas diferentes a un ser humano, elegirá aquella cuya proporción de lados esté más cercana a 1:6?

Giros

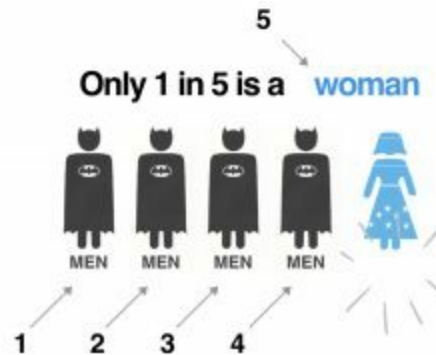


Fig 65 Utiliza “giros” para dar un toque a tu gráfico.

En el teatro de *improvisación*, siempre decimos que una buena historia debe tener un giro [46-47]. Un giro efectivo es lo que hace viral un video Tiktok o en casos de negocio, la expresión ¡Aha! en un momento determinado. Podemos utilizar un espacio negativo para crear estos giros en los gráficos. En este caso, el lector espera encontrar la etiqueta #5 en un determinado lugar, pero ¡no está aquí! Esto pone al cerebro en modo de búsqueda [55]. Cuando unos pocos milisegundos más tarde, el cerebro encuentra la etiqueta que faltaba, se liberan endorfinas. Una ventaja de este giro es que pone a prueba de fallos (Poka-Yoke [69]) el gráfico y cambiando la ubicación de la etiqueta se asegura que el lector no perderá el título.

Ahora que conocemos los principales trucos para hacer un gráfico pop, veamos algunos ejemplos.

Cuestionario: Helados Suecos

Evaluar el siguiente gráfico. Fuente: [99]. A continuación, proponga una mejor manera de visualizarlo. **Requisito:** 300 palabras de ensayo + dos bocetos de la visualización propuesta mejorada. Tiempo 25 minutos.

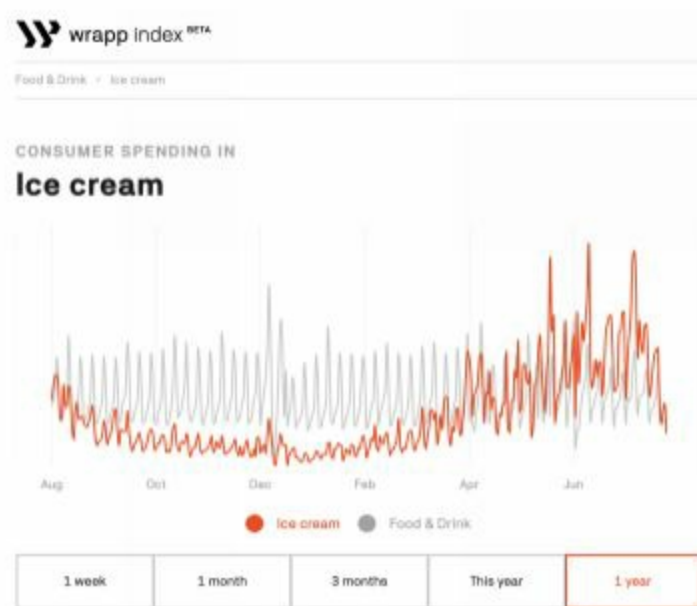


Fig 66 Ventas de helados en Suecia, Islandia y Finlandia.

(Solución en la página siguiente)

Solución

Principales observaciones

1. Este gráfico contiene información útil como: las ventas de fin de semana son más altas que las ventas de los miércoles, y las ventas en verano son más altas que en el invierno. Sin embargo, la línea gráfica del tiempo no es óptima para mostrar esto.
2. Para dar a este gráfico más utilidad, debemos conectarlo a otros conocimientos. Por ejemplo, la comparación de las ventas en España, donde la proporción de ventas en la relación Invierno: Verano es menor, 4x vs 10x en Suecia.
3. Esta gráfica carece de narrativa. Un ejemplo de narrativa podría ser cómo la cultura influye en los hábitos de consumo. Ejemplo: “Es curioso, lo que sucede en Suecia durante los meses de verano, es como si perdiéramos nuestra mente. Comemos, bebemos y nadamos como si nunca hubiéramos visto el sol - es lo que pasa cuando se vive cerca del Polo Norte y llega Julio”. Thérèse Lundquist - Jefa de Marketing & RP de Wrapp. Ver este comentario en el contexto en [103]

Opciones de color

Naranja: Lindo, a primera vista, molesto después de 20 segundos.

Línea del tiempo

El dato es periódico semanal y anual, por lo tanto, escalas periódicas de tiempo anuales o semanales reducirían el desorden. Por ejemplo, un gráfico de radar semanal.

Gráficos alternativos

4. Radar semanal
5. Diagrama de barra mensual, semanal, con iso-medidas de helado
6. La línea gris es constante, por tanto, no tiene significado (¿quitar?)

Otras sugerencias

7. Normalizar per cápita y por cada usuario de la aplicación

8. Anotar el fin de semana de Navidad y otros días peak
9. Explotar en una dispersión entre la temperatura (eje X), las ventas (eje Y), color de punto es días entre semana
10. Calcular el porcentaje de las ventas debido a la temperatura y el porcentaje de ventas debido a la estacionalidad en un análisis de regresión lineal utilizando como factores el verano y los fines de semana [101]
11. Utilizar Crossfilter para permitir a los usuarios explorar y descubrir relaciones ocultas en las variables para el día de la semana, mes, temperatura, en las ventas de helados y la variación individual en el consumo

Cuestionario: El calentamiento global

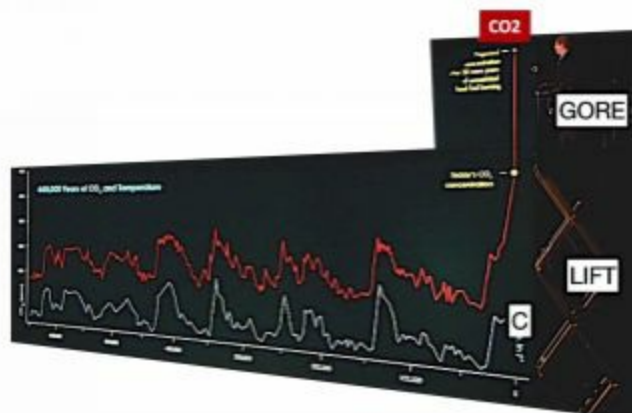


Fig 67 El gráfico de Al Gore de emisiones de CO2 no fue viral.

En 2006 Al Gore presentó este Gráfico de Emisiones de CO2. El lapso del eje X es +100 años; en el eje Y, la línea roja es la concentración de CO2; la línea azul es la temperatura. Tenga en cuenta la alta correlación entre ambas. A la derecha está Al Gore en la parte superior de un elevador de tijera. Fue un fracaso. ¿Cómo ayudaría a Al Gore a visualizar el calentamiento global? Tiempo 25 minutos.

Solución

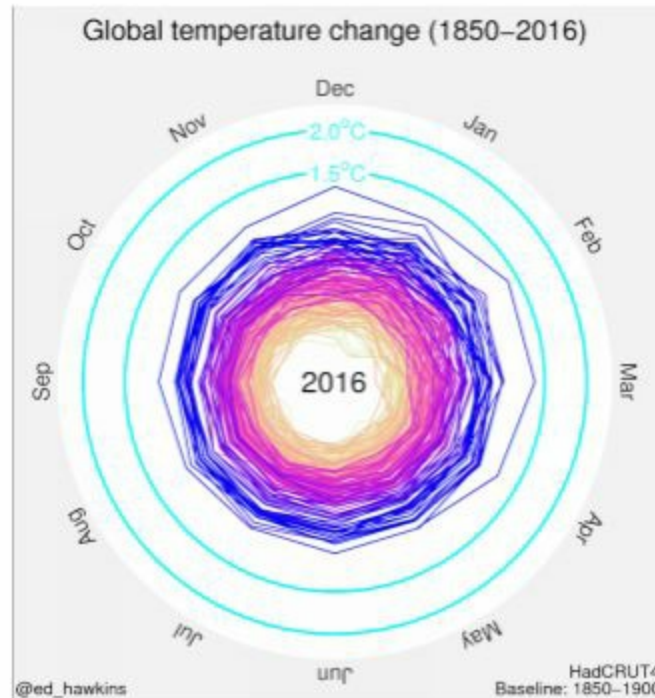


Fig 68 Ed Hawkins hizo este **gráfico en espiral** en mayo de 2016. Se hizo viral en cuestión de minutos.

De [Tortas vs Barras](#). Barras sabemos que los seres humanos somos **más** sensibles al cambio circular que al lineal. Si queremos que el gráfico se alinee con la narrativa que “el cambio climático es una emergencia” [100], ¡Aprovechemos esto! Ver también GIF animado en [[104](#), [105](#)].

Cuestionario: Cuadrantes Mágicos



Fig 69 Las Habilidades de Ciencias de Datos más buscadas en 2019

En 2019, Gregory Piatetsky, editor de Kdnuggets, publicó este gráfico [108]. Se utilizaron tres principios de visualización:

1. **Color** - para identificar las habilidades de frías/calientes,
 1. **Punto** - tamaño para indicar una magnitud
 2. **Cartografía** - en 2D (pero no utilizó su potencial de diseño)
- El gráfico se basa en una encuesta. La encuesta tiene solo dos preguntas:

1. ¿Qué habilidades/áreas de conocimiento **tiene** actualmente?
6. ¿Qué habilidades **desea** añadir o mejorar?

KDnuggets recibió 1.500 respuestas y los agregados por la mirada de habilidad de la siguiente manera.

tabla 1. Los agregados de la encuesta.

Habilidad	Tengo	Quiero	ratio
Python	71,2%	37,1%	0.52

Visualización de Datos	69,0%	25,3%	0.37
Pensamiento Crítico	66,7%	15,5%	0.23
Excel	66,5%	4,6%	0.07
...			

¿Cómo podemos visualizar estos datos en una forma más significativa?
 ¿Qué espacio de diseño es el más apropiado dados los datos? Tiempo 4 horas.
 Sugerencia: Ver Gartner [107].

Solución

Vamos a aplicar lo que hemos aprendido hasta ahora. (Encuentre un por qué para transformar los datos en información y sintetice el conocimiento mediante la vinculación a los marcos para que sea útil en la toma de decisiones). Antes de encontrar un por qué, primero vamos a explorar los datos.

La primera reacción es hacer un gráfico de dispersión para identificar grupos de interés. El eje X puede ser el porcentaje de encuestados que tienen una destreza en particular, y el eje Y, el porcentaje de encuestados que le gustaría tener esa habilidad (quiero). Sin embargo, hay demasiados puntos de datos para que un ser humano pueda darle sentido. Es el caso de la muerte del libro de texto por exceso de información y en la figura 70 se utilizó el meme de Jackie Chan para transmitirlo.

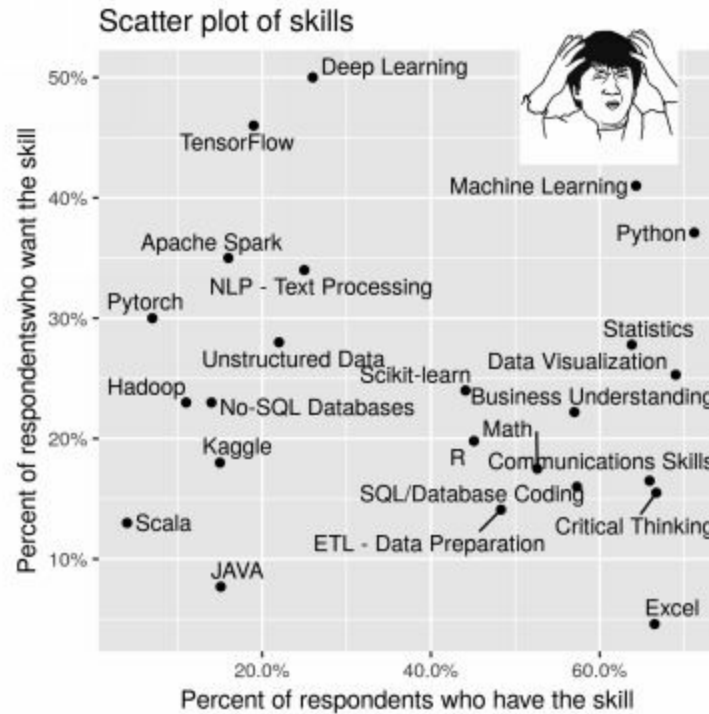


Fig 70 ¿Víctima de sobrecarga de información?

¿Cómo transformar esta información en conocimiento? Tomemos una señal del **Cuadrante Mágico de Gartner**, un marco utilizado en la inteligencia de negocios TI. Reduce la complejidad a niveles humanos mediante el uso de una jerarquía cuadrante [107] (Fig 71).

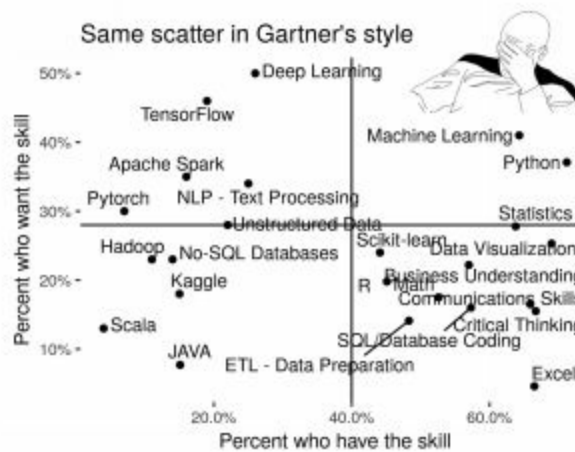


Fig 71 Gartner utiliza el esquema de cuadrantes para agrupar y reducir la complejidad y dar sentido.

Sin embargo, la figura 71 está lejos de estar lista. El eje Y está alineado con la metáfora de la gravedad (altamente deseada, alta Y). Sin embargo, el eje X no está alineado con otra regla no escrita, (ésta por Guy Kawasaki): “quiere (metas deseadas) para estar arriba y tener razón [110]”. En este caso, la habilidad más deseada (Aprendizaje Profundo) está en el lado equivocado - necesitamos voltear el eje X, (Fig 72).

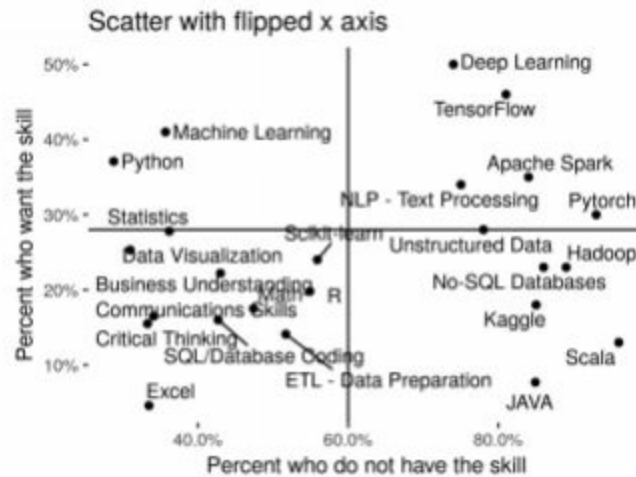


Fig 72 Los objetivos deben estar “altos y a la derecha” - Guy Kawasaki.

Si hace un gráfico y nadie se acuerda de él. ¿Sigue siendo útil? En la figura 72, se agruparon las habilidades en cuatro categorías, pero ¿para qué sirve si nadie se acuerda de ellas? Una forma de ayudar a su audiencia a recordar son las personas (memes, en el argot Gen-Z). Vamos a aplicar personas usuario. En la figura 73 cada cuadrante significa:

1. **Habilidades no deseadas** (Tengo, pero, no quiero = Excel)
2. **No, gracias** (No quieren y no tienen = JAVA)
3. **Habilidades calientes** (Quieren, pero no tienen = Flujo Tensor)
4. **Habilidades amadas** (Quieren y tienen = Python)



Fig 73 Cultura pop, utilizarla para hacer su gráfico llamativo.

Bolígrafo rojo

¡Su lienzo no debe temer al bolígrafo rojo! Los subtítulos son una oportunidad para aclarar el significado y añadir un toque a su historia (no todas son visuales). Note cómo en la figura 74 se rompió la simetría por la inclinación de la etiqueta “amado”, que es el Feng-Shui para los gráficos. También hemos añadido un toque en el cuadrante de Java al no tener una etiqueta para él. Esto asegura que el lector irá a este cuadrante después de visitar todos los demás. La etiqueta para este cuadrante se encuentra dentro del meme (No, gracias).

Estratificación

Note cómo hemos estratificado la información en jerarquías (meme, etiquetas cuadrantes, cuadrante representativo). Hemos respetado el límite de siete trozos en cada capa para evitar la sobrecarga. El significado se logró mediante la vinculación a un marco existente, organizando los datos en cuadrantes y etiquetándolos.

Narrativa

Por último, los gráficos deben tener un propósito. Es razonable avanzar en el proceso de creación de conocimiento sin saber por qué. Una vez que el conocimiento se ha encontrado y visualizado, el por qué va a ser fácil de encontrar. Mi por qué para este gráfico es: “Me gustaría ver más Python y menos Java en mi clase”.

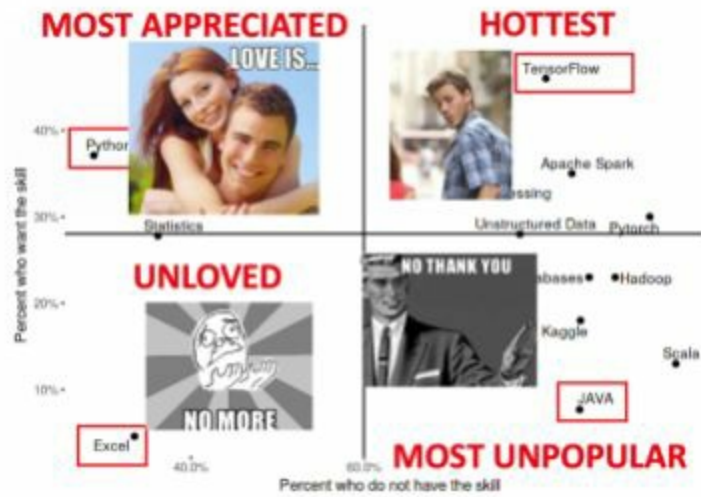


Fig 74 ¿Pueden los cuadrantes mágicos hacer memorable un diagrama de dispersión?

Pensando en la estrategia

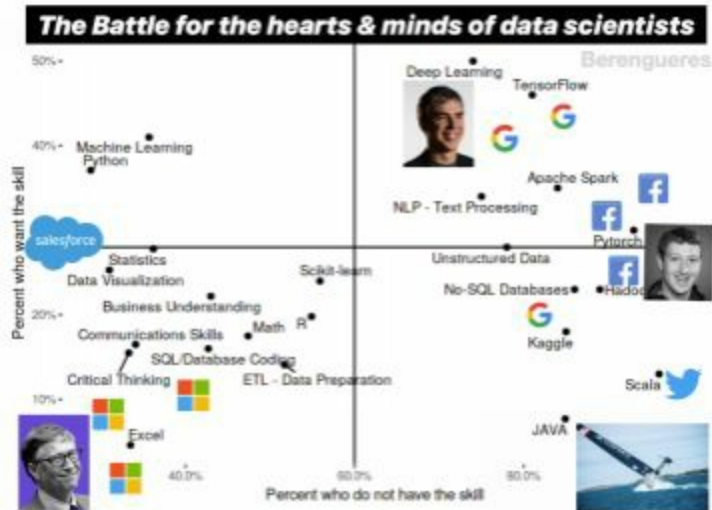


Fig 75 Salesforce adquirió Tableau. ¿Por qué?

Ahora que hemos creado el conocimiento, ¿podemos usar este gráfico como una herramienta de pensamiento? Una manera es imaginar contextos donde este gráfico podría ser útil. ¿Dónde podría ser utilizado este gráfico para crear conciencia de la situación? Las “guerras por la nube” entre Microsoft y Google se libran a través de servidores tales como PowerBI, Kaggle y otras palancas de bloqueo de la nube. En los años 90 se desarrolló un libro de jugadas similar en el mercado de bases de datos. La figura 75 visualiza cuáles son los patrocinadores y dónde se ubican los lenguajes de programación, para ver cómo están las lealtades.

Prueba: Visualice la brecha

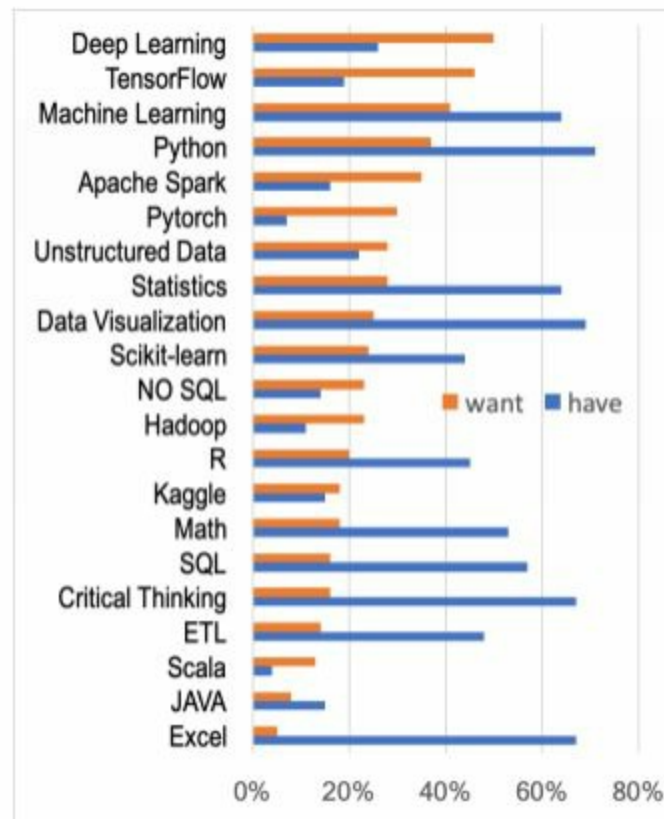


Fig 76 ¿Cómo mejorar este gráfico?

En el mismo post, KDnuggets publicó este gráfico. Su propósito era mostrar la **brecha** entre *tener* y *querer habilidades* sin perder la perspectiva de las magnitudes absolutas implicadas en cada habilidad. ¿Por qué falla este gráfico? ¿Cómo se pueden suprimir elementos de este gráfico y hacerlo más significativo y sin pérdida de información? (No hay requisitos para este gráfico). Tiempo 1 hora. Sugerencia: Véase cascada de barras de pila con componentes negativos de McKinsey.

Solución

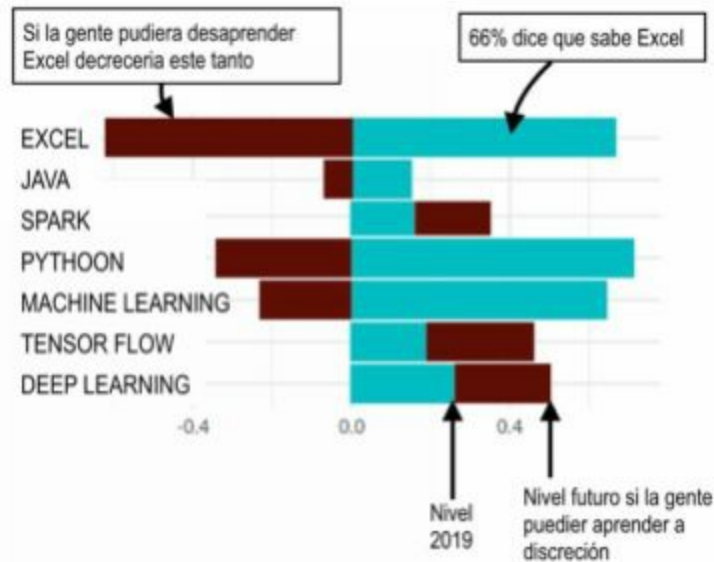


Fig 77 Si la gente pudiera olvidar cosas, ¿lo harían?

Explicación del proceso de despejar

El autor quiere mostrar para cada habilidad tres atributos:

1. **Tener** habilidades
2. **Querer** habilidades
3. Se busca **tener menos** (la brecha)

De estos tres solo dos son variables independientes. Aplicando el principio de eliminación de información superflua mostramos solo dos. Tener una habilidad y el quererla, además de la brecha entre ambos. En el caso de Excel, el 66% dice que tiene el conocimiento de Excel, pero solo el 5% lo quiere ahora. ¿Esta brecha de 61% (en granate oscuro, izquierda) indica cuántas personas querrían olvidar Excel si pudieran?

Sintetizando nuevo conocimiento

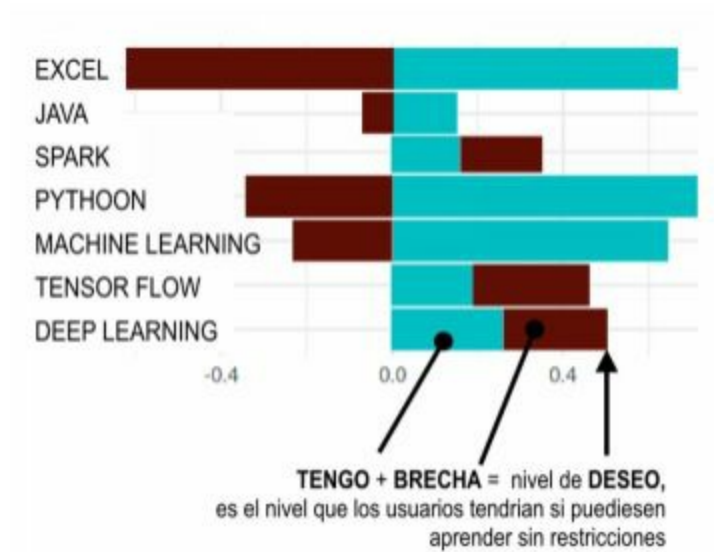


Fig 78 ¿Qué pasaría si usted pudiera aprender todo lo que quisiera?

Este gráfico se ha convertido ahora en un **vaticinador** de lo que sucedería si la gente aprendiera lo que dice que querría. *Tiene* es el nivel actual (prevalencia). *Tiene + brecha* es el nivel futuro. En el caso de las lagunas negativas, las barras de separación se representan en el otro lado del ej. Y, (un fallo de la función de apilamiento de ggplot2 o una característica, pues uno no puede desaprender). Observamos cómo, a partir de una narrativa competitiva de la Fig 71 (barras competidoras), hemos cambiado a una narrativa de **mentalidad de crecimiento** con el “y si pudiera aprender todo lo que quisiera”. ¿Cómo podemos usar este gráfico para priorizar qué habilidades enseñar? Fuente [108].

Capítulo 6 Psicología de la visualización

co-autor *Ali Fenwick*

La Psicología juega un papel central en el diseño de efectos visuales eficaces, especialmente cuando se trata de la persuasión y la influencia. En este capítulo revisaremos más a fondo cómo los principios de comportamiento se pueden utilizar para influir en el diseño visual y la interpretación de los datos. En concreto, nos centraremos en tres conceptos psicológicos que se aplican a la visualización: Contexto, Encuadre y la Teoría de la Gestalt. Conocer estos principios y técnicas y aprender a aplicarlos de manera efectiva mejorará en gran medida la forma de diseñar visualizaciones convincentes.

Ciencia del Comportamiento

El campo de la ciencia de la conducta es un amplio conjunto de especialidades que se centran en la teoría y los marcos que van desde la Antropología a la Psicología. Ideas de la Psicología Cognitiva y la Teoría Económica del comportamiento pueden ser útiles en la comprensión de cómo las técnicas de visualización pueden afectar el juicio humano, especialmente al momento de decidir cómo visualizar los datos y la forma de contar una historia. En los entornos de la vida real, las decisiones importantes no solo se basan en hechos, sino también en parcialidades subjetivas. Como vimos en el capítulo anterior, la forma en que se presenta la información, en lugar de qué es lo que se presenta, puede ser igual o incluso más influyente en el proceso a través del que las personas interpretan y actúan sobre los datos. Como Tversky y Kahneman explican en su artículo escrito hace más de 40 años, el juicio y la toma de decisiones están fuertemente afectados por factores inconscientes como los psicológicos, emocionales y sociales; y los factores contextuales [115,118]. Ser consciente de cómo estos factores afectan la cognición es un primer paso hacia el desarrollo de visualizaciones eficaces y éticas. En la siguiente sección, vamos a empezar por explorar el poder del contexto y la forma en que juega un papel crucial en la visualización de datos.

Contexto

De acuerdo con Thomas & Millar, el contexto es uno de los factores más fuertes de influencia sobre el juicio y la toma de decisiones [116]. Esto se debe a que nuestro cerebro utiliza todos los sentidos (por ejemplo, la vista, oído, olfato) para recopilar información sobre el entorno para crear una historia coherente de lo que está pasando. En las presentaciones, por ejemplo, el diseño visual, la voz del altavoz, el vestido, el lenguaje corporal y el entorno físico -todos juntos- comunican una 'narrativa única' para ayudar al cerebro dar sentido a la experiencia [114]. El cerebro también trata de dar sentido a todas estas entradas relacionándolas con nuestras narrativas existentes (por ejemplo, nuestros valores, creencias y marcos mentales). En otras palabras, la información individual no es (y no se puede) procesar de forma aislada o de forma independiente de su entorno.

Cuando se dirige a una audiencia por primera vez, el cerebro usa todas estas entradas para crear una opinión sobre alguien o algo en el salón. La ciencia muestra que esto sucede dentro de los primeros 11 segundos de su encuentro e influye en cómo las personas lo evaluarán a usted y a su presentación. Es por ello que, por ejemplo, cuando realizo una presentación ante un público que nunca he visto antes, me aseguro de ensayar el primer minuto de ella meticulosamente y recitar palabra por palabra la apertura de mi exposición, idealmente ensayo en el mismo espacio donde me presentaré, asegurándome de mirar siempre hacia el público y moverme con confianza, y si es posible tener a alguien que me presente antes del inicio. Estas técnicas que conozco bien mejoran significativamente esa primera impresión y por lo tanto mi éxito en la entrega de una presentación impactante. ¡Sir David Ogilvy, el pionero de la publicidad, era un maestro en el uso del contexto para influir en su audiencia y construyó su carrera sobre ello! En 'Confesiones de un hombre de Publicidad', su biografía de 1963 [123] relata varias técnicas que utilizó. A continuación, veremos un ejemplo de cómo el contexto, en este caso la relatividad, afecta la percepción visual.

Principio de la relatividad

Echar un vistazo al círculo azul la figura 88. ¿Lo que es más grande?



Fig 79 *La influencia del contexto sobre el tamaño percibido*^[14].

Ya se puede esperar cuál es la respuesta correcta: ambos son del mismo tamaño. Pero si tomamos en cuenta la percepción sabia, no lo son. El círculo azul se ve más grande cuando se acompaña de pequeños círculos alrededor de él y parece más pequeño cuando ponemos a su alrededor círculos mayores. Al cubrir los círculos exteriores con los dedos se verá que los dos círculos azules son del mismo tamaño.

Esta ilusión refleja cómo nuestra mente interpreta el mundo. Por tanto, es aconsejable ignorar el impacto más amplio de factores contextuales en sus presentaciones. Este efecto no solo es válido para los gráficos, sino también para las cosas más cotidianas como las emociones, actitudes y puntos de vista [124]. Esto significa que nuestro cerebro usa tanto el contexto externo (imágenes, sonidos y los olores), como el contexto interno (las emociones, las experiencias pasadas y los deseos) para interpretar la narrativa de lo que sucede.

El procesamiento mental de la información contextual, ya sea visual o no, ocurre predominantemente en el nivel inconsciente [114]. Menos del 20% de nuestro estado de alerta es la conciencia, lo que significa que factores como el contexto pueden influir fácilmente en el juicio y toma de decisiones sin que la gente sea consciente de ello, teniendo efectos intencionales y no intencionales. Vamos a ver un ejemplo de cómo el contexto puede influir en la conducta alimentaria de manera inconsciente.



Fig 80 El profesor Wansink estudió el efecto que el tamaño del plato tiene sobre su consumo de calorías^[15].

La misma cantidad de alimentos se presenta en dos platos de diferentes tamaños; uno es más grande que el otro. Visualmente, el plato más pequeño hace que la comida se vea más grande y los estudios han demostrado que las personas que comen del plato más pequeño tienen obtienen la sensación de saciedad antes y por lo tanto comen menos. Las personas que comen del plato más grande tienden a comer más que las personas que comen del plato más pequeño. Ahora piensa en esto y en cómo se aplica a los gráficos.

Sobrecarga cognitiva

Según Sweller, la sobrecarga cognitiva es un fenómeno en el cual el cerebro es incapaz de procesar la información de manera efectiva debido a la gran cantidad de datos presentados [117]. La sobrecarga cognitiva provoca fatiga mental y la dependencia de atajos mentales (que están sometidos a la heurística y sesgos). Por ello es importante que el diseño incorpore elementos visuales que ayuden a superar las limitaciones de recursos cognitivos e impidan la toma de decisiones defectuosa. Como vimos en el capítulo de [cuadrantes mágicos](#), tres son las principales trampas de sobrecarga a evitar:

1. **Limite** la cantidad de datos presentados en un gráfico o una diapositiva
 - **Estructure** sus datos en pequeñas partes de información relevante, y
 - **Realce** palabras y frases clave que ayuden al lector a recoger las partes más importantes del texto

Enmarcado

Desencadenar atajos mentales en el diseño visual también puede ser beneficioso para mejorar la eficacia visual. Una forma de activar atajos mentales en la visualización es a través de la estructura. El encuadre es un sesgo cognitivo que afecta el juicio y la toma de decisiones utilizando lo positivo (ganancia) o mensajes negativos (pérdida) (véase de nuevo el trabajo de Kahneman en [118]). La gente tiende a estar más motivada para tomar medidas cuando los mensajes se enmarcan como positivos (una ganancia) o negativos (pérdida) en función de la situación. Por ejemplo, los estudios muestran que las personas son más propensas a buscar riesgos cuando un mensaje se enmarca como negativo o una pérdida. Aquí hay un ejemplo.

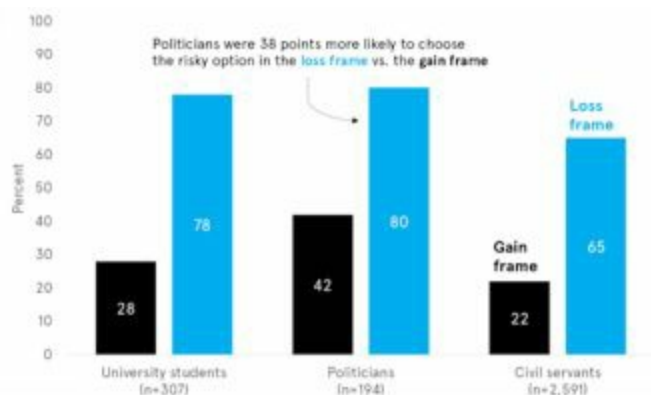


Fig 81 Evitar muertes (marco de pensamiento pérdida) conduce a una mayor persuasión que frasear una propuesta de política como “salvar vidas” (un marco de pensamiento de ganancia)

De acuerdo con el diagrama anterior [119], los funcionarios del Gobierno eran tres veces más propensos a tener opciones de mayor riesgo para combatir un brote de enfermedad, cuando el diseño de políticas se enmarcaba como la prevención de la pérdida de vidas en comparación con las vidas salvadas. En situaciones médicas, los estudios muestran que es más prudente utilizar un encuadre positivo, en lugar de un encuadre negativo, para ayudar a los pacientes a tomar mejores decisiones de salud por sí mismos, sobre todo en situaciones de alta incertidumbre. Curiosamente, los recientes hallazgos en los efectos de encuadre han encontrado que ciertos factores podrían moderar el impacto de enmarcar. Por ejemplo, un estudio realizado por Thomas y Millar en 2011 mostró que los efectos del encuadre cambian con la edad, afectando la toma de decisiones y la memoria de las personas jóvenes y mayores de forma diferente [116]. También se ha descubierto que los efectos

de encuadre desaparecen cuando la información se presenta en un segundo idioma o en un idioma que no resulta natural para el lector [120]. Esto puede explicarse por el hecho de que se necesita más pensamiento deliberado para procesar información en un idioma extranjero, lo que reduce la tendencia del cerebro para las respuestas automáticas (con prioridad)^[16].

Gestalt

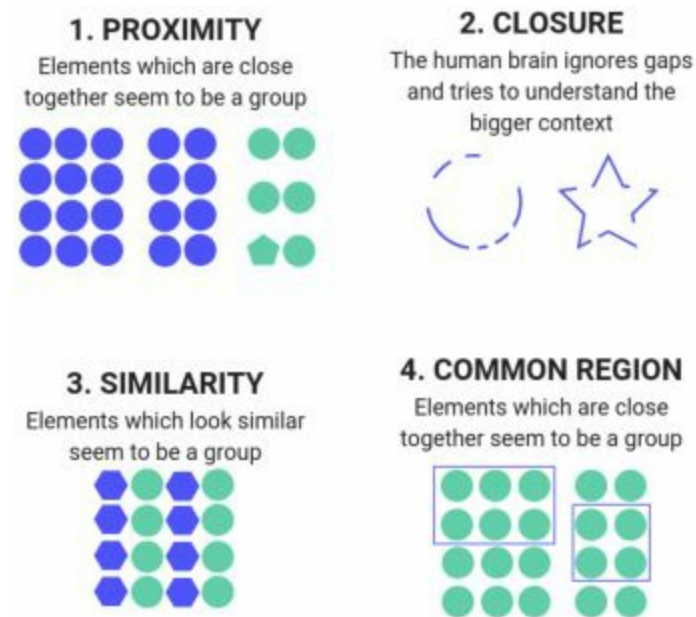


Fig 82 Cuatro de los ocho principios de la Gestalt^[17].

En el siglo XX, un grupo de psicólogos alemanes y austríacos proporcionaron una comprensión de cómo el cerebro interpreta a los humanos y organiza los elementos visuales. Su investigación condujo a la creación de los Principios de la Gestalt para un mejor diseño visual para reflejar la forma natural del cerebro de procesar la información. Gestalt es una palabra alemana que significa un patrón o conjunto unificado. Veamos cómo algunos de estos principios se pueden aplicar a la Visualización de Datos (para más principios de la Gestalt, ver [121]).

La Ley de la Proximidad

La Ley de la Proximidad refleja la necesidad de agrupar artículos similares juntos o en estrecha proximidad entre sí. La agrupación de elementos similares juntos hace que sea más fácil para que el lector entienda lo que se comunica. Cuando las partes más estrechas se colocan juntas, mayor es la

relación percibida. El gráfico de reversión a la media, por ejemplo, en el capítulo 4 (Fig [38](#)), es un gran ejemplo de la ley de la proximidad. Cuanto más cerca está la etiqueta de “Japón” de su correspondiente punto, menor es la confusión y mayor la carga cognitiva. Consejo: Tenga cuidado de no poner objetos no relacionados demasiado cerca unos de otros, ya que el ojo tendrá como objetivo crear la relación basada en la proximidad de los elementos presentados.

La ley del cercado

Agrupar elementos crea relación. Aplicar límites que funcionen como un cercado para resaltar a un grupo funciona. Por ejemplo, podemos agrupar puntos en un gráfico para indicar relación entre partes específicas de información (clúster, valores atípicos). En la Fig 83-84, hemos añadido un ejemplo. Las delimitaciones circulares de la figura 84 ayudan a reducir la carga cognitiva de la figura 83 y aumentan su significado.

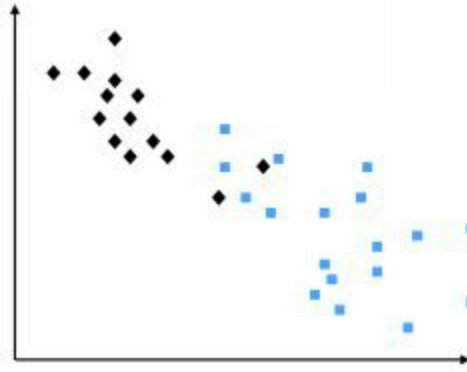


Fig 83 Un gráfico de dispersión con dos clases de puntos de datos.

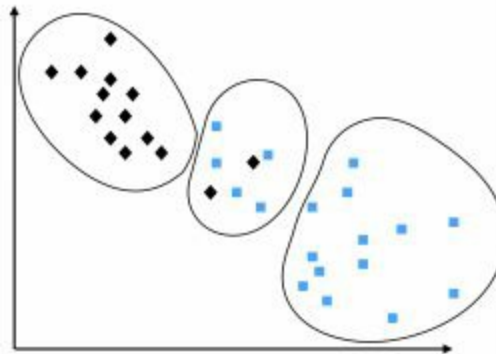


Fig 84 Mismo gráfico con recintos. De repente, parece que tiene más sentido.

La regla del Peak-End

Por último, vamos a ver un principio psicológico fascinante llamado la regla del Peak-End, una heurística psicológica en la que la gente juzga (o recuerda) una experiencia basada principalmente en dos momentos específicos: un momento peak en algún lugar en medio de la experiencia y el final de la experiencia, en lugar de evaluar la experiencia en su totalidad. Este sorprendente efecto se produce independientemente de si la experiencia es agradable o desagradable. En otras palabras, en la narración de su visualización, el público recordará la experiencia más favorable si termina con una nota alta (¡excepto el golpe más fuerte para el final!). Fig 85 Muestra

la intensidad de dos presentaciones por el Sr. A y el Sr. B.

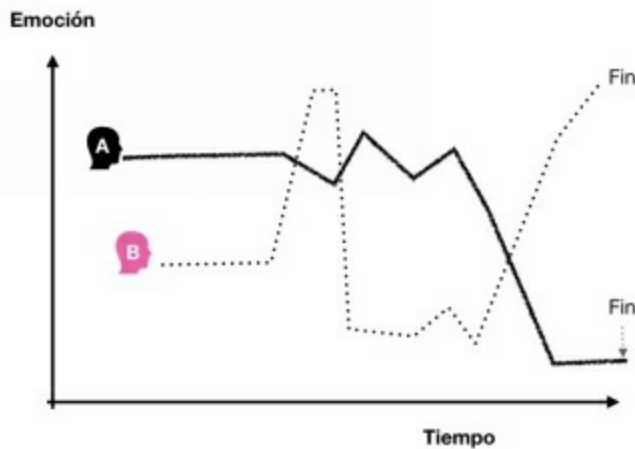


Fig 85 Regla de la Cúspide de gama, el uso a efectos de maximizar.

De acuerdo con esta regla, la presentación del Sr. B será clasificada casi tan favorablemente como la presentación del Sr. A pesar de que ésta fue más larga y tenía más momentos AHA!.

Steve Jobs utilizaba esta regla para maximizar el impacto de sus propias presentaciones, por eso siempre terminaba en una nota alta. Su “Una cosa más ...” es ahora parte de la cultura pop. Entre 1998 y 2010, Jobs la usó en no menos de 25 ocasiones^[18]. Greg Wyatt de @appleexplained realizó una compilación de una hora de duración de tales momentos. La regla del Peak-End no solo se puede aplicar a las presentaciones, también a los gráficos. Un ejemplo es la figura 65, “giros”. Allí, (a expensas de la legibilidad), el autor maximiza el puñetazo de la tabla, asegurándose de que no se alcanza el momento de AHA hasta que el público pasó por todos los otros elementos de la tabla por primera vez.

Capítulo 7. Sesgos y Prejuicios

co-autora Marybeth Sandell

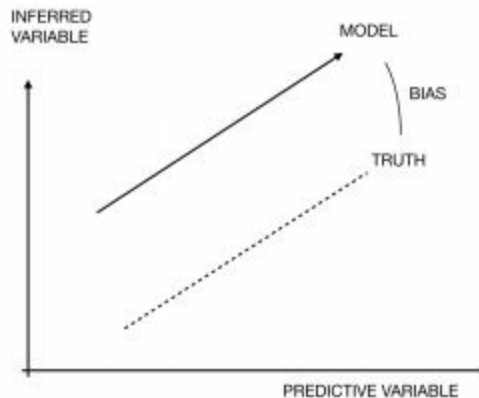


Fig 86 El sesgo (bias) es poco ético cuando es injusto con la verdad (truth), por lo general contra las minorías.

En este capítulo, aprenderá lo que es el sesgo y cómo puede afectar a una visualización basada en datos. Sesgo no solo se puede clasificar por su punto de entrada (datos, historia, narrativa), sino también por el área que explota en el sistema de la cognición (ilusiones ópticas, sesgos culturales). Es fácil asumir que el sesgo es intencional. Sin embargo, el sesgo puede surgir por muchas razones.

En primer lugar, el sesgo puede estar embebido en los datos en sí, de forma deliberada, por la forma en que se recoge la información y también accidentalmente por no darse cuenta de lo que falta.

En segundo lugar, puede aparecer un sesgo a medida que se elabora la historia. Esto también puede ser intencional seleccionando datos existentes, o accidental en los casos en los que no se dedica suficiente tiempo a explorar todos los datos disponibles (generalmente debido a la presión del tiempo).

En tercer lugar, el sesgo puede ser incrustado en la propia narración. A menudo esto es intencional, como en la propaganda. Pero también puede ser involuntario como en los casos de los prejuicios culturales.

Tipos de sesgo

En términos generales, sesgo es cualquier error sistemático. En otras palabras, una diferencia sistemática entre un modelo y la “verdad” que

supuestamente representa^[19]. En Ciencias Sociales, los sesgos se juzgan como poco éticos cuando son injustos (por lo general hacia una minoría). Véase también [marcos](#) éticos.

El sesgo puede afectar al productor del gráfico (como en el sesgo de selección en los datos) y también al consumidor del gráfico, por ejemplo en Groupthink: pensar como el rebaño; la falacia de la mano caliente; y así sucesivamente. ¡Los psicólogos y los economistas del comportamiento han identificado más de 200^[20] tipos de sesgos cognitivos! Estos se pueden clasificar en tres grupos:

1. sesgo de creencia
2. prejuicio social
3. sesgo de la memoria

Además de los sesgos cognitivos mencionados, cuando se trata de la visualización de datos, también se aplican los sesgos de percepción visual. Veamos algunos ejemplos.

El sesgo en la narrativa

Las formas más amplias de sesgo inconsciente se deben a la falta de conciencia y están tan arraigados en la sociedad que, por lo general, son normas culturales y/o morales también. Tenga en cuenta que no todas las normas culturales están sesgadas, y que la mayoría de las normas evolucionan más lento que la sociedad y por lo general van quedándose atrás con respecto a la realidad. Ejemplos de sesgo en la narrativa consciente son la propaganda y la desinformación. Las técnicas típicas utilizadas son FUD (por sus siglas en inglés) o Miedo, Duda e Incertidumbre; como se ve con la industria del tabaco y FLICC (Mentiras Falsas, Errores Lógicos, la Expectativa Imposible, Exclusión de Datos que no interesan y Teorías de Conspiración) como se ve con la negación del cambio climático ^[21]. Revisemos un ejemplo.

El sesgo en la narrativa: ¿Una comida equilibrada?

En la segunda mitad del siglo XX, se asumió que una dieta equilibrada era óptima para la salud. En la escuela, a muchos niños (a mí también) se les mostraron los gráficos con los grupos de alimentos relativamente

equilibrados. Sin embargo, otras culturas y algunos trabajos de investigación independientes muestran que la narrativa de alimentación balanceada más extendida en el mundo tal vez no es la más saludable. Por ejemplo, la Dieta de Okinawa contiene menos de un 5% de proteína animal y no incluye derivados de la leche. Esta dieta se considera peligrosamente desequilibrada en todos los sentidos occidentales. Sin embargo, los habitantes de Okinawa reportan una de las esperanzas de vida más saludables y más largas del mundo.

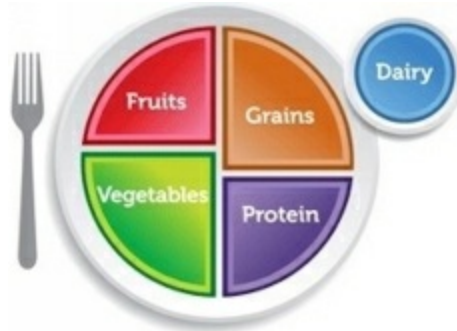


Fig 87 ¿Falacia lógica? ¿Cómo podría esta “equilibrada” dieta no ser saludable si se equilibra? [\[22\]](#)

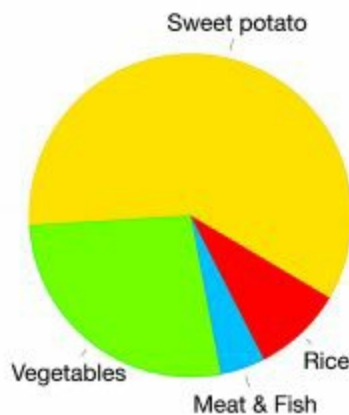


Fig 88 Dieta de Okinawa, totalmente desequilibrada, pero ¿más saludable? [\[23\]](#)

Sesgo en la historia ¿Alemanes trabajadores?

Al contar una historia visual, cómo se usan los datos y se presentan puede inducir sesgos conscientes e inconscientes (bien conocido en la publicidad y neuropsicología). Esto se puede hacer a través de la selección y la presentación de los datos, los colores utilizados y así sucesivamente. Veamos un ejemplo.

Ejercicio: Los alemanes trabajan duro

La figura 89 es un gráfico creado por la Agencia de Desarrollo Económico de Alemania. Utilizaron datos de la UE para demostrar que los alemanes trabajan más tiempo (y en labores más complejas) que la media. ¿Pero cuánto más difícil? ¿Puede identificar dónde está el sesgo? Tiempo de 4 minutos.

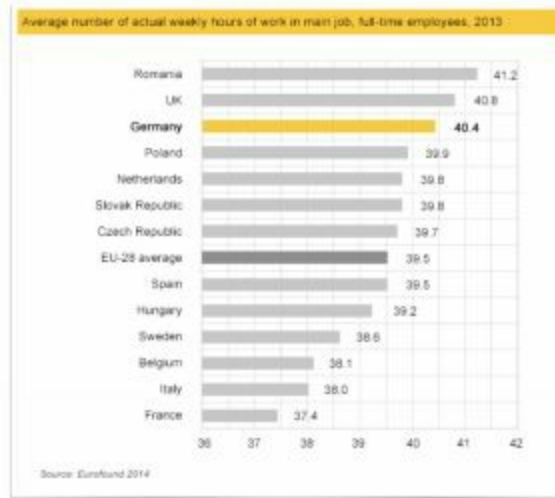


Fig 89 *¿Una vista parcial?*

Solución

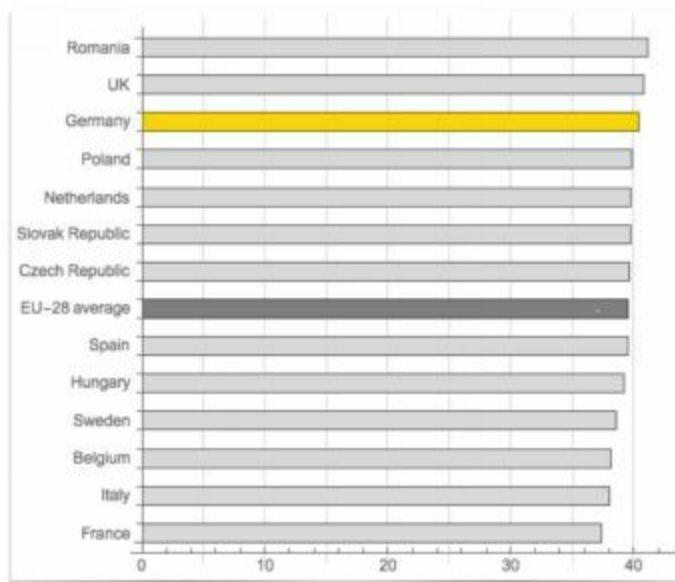


Fig 90 *Una visión más completa.*

Crecimiento en tiempos de deuda (2010) fue un paper que no fue revisado por pares publicado por Carmen Reinhart y Ken Rogoff - ex economista jefe del Fondo Monetario Internacional. Se dijo que la economía de un país disminuye significativamente cuando la deuda se eleva por encima del 90% del PIB. Esto no es verdad. Sin embargo, el paper se utiliza para justificar la austeridad en aras del “bien común”, (ver marcos éticos en el cap. 1). Un

estudiante tratando de reproducir los resultados descubrió que, de los más de 20 países considerados para hacer la estadística, los principales países se habían omitido de manera arbitraria en los cálculos. Más tarde, el profesor Michael Ash señaló otra conducta impropia en el método de promedio de ponderación utilizados. Por ejemplo: “Nueva Zelanda (un país de población inferior a 5 millones de personas) que tuvo un solo año (1951) esta condición, con un crecimiento de -8%, se mantuvo con el mismo peso que los 20 años de Gran Bretaña con un crecimiento de 2.5% (población de 60 millones)”. En base al dólar, es decir una proporción de polarización de 1 a 240. Por ahora hemos visto tres ejemplos de sesgo en la historia, los datos y la narrativa. Ahora vamos a ver un poco más a menudo sesgos al revisar diferentes gráficos.

Sesgo de los datos

En el Periodismo, el sesgo en los datos viene en tres grandes grupos: los paraguas defectos de selección, la asimetría y la omisión. Algunos ejemplos incluyen: cómo se hacen las preguntas (como en las preguntas principales); sesgo en la hipótesis o suposiciones; errores matemáticos y defectos de diseño de la encuesta. Veamos un ejemplo de un famoso sesgo de selección^[24] descubierto por el estudiante Thomas Herndon [126].

Sesgo creencia: Las mentiras de la austeridad

Crecimiento en tiempos de deuda (2010) fue un paper que no fue revisado por pares publicado por Carmen Reinhart y Ken Rogoff - ex economista jefe del Fondo Monetario Internacional. Se dijo que la economía de un país disminuye significativamente cuando la deuda se eleva por encima del 90% del PIB. Esto no es verdad. Sin embargo, el paper se utiliza para justificar la austeridad en aras del “bien común”, (ver marcos éticos en el cap. 1). Un estudiante tratando de reproducir los resultados descubrió que, de los más de 20 países considerados para hacer la estadística, los principales países se habían omitido de manera arbitraria en los cálculos. Más tarde, el profesor Michael Ash señaló otra conducta impropia en el método de promedio de ponderación utilizados. Por ejemplo: “Nueva Zelanda (un país de población inferior a 5 millones de personas) que tuvo un solo año (1951) esta condición, con un crecimiento de -8%, se mantuvo con el mismo peso que los 20 años de Gran Bretaña con un crecimiento de 2.5% (población de 60 millones)”. En base al dólar, es decir una proporción de polarización de 1 a 240. Por ahora hemos visto tres ejemplos de sesgo en la historia, los datos y la narrativa.

Ahora vamos a ver un poco más a menudo sesgos al revisar diferentes gráficos.

Otros sesgos

Sesgo de confirmación: Tendemos a creer las declaraciones y los datos que confirman nuestras creencias e ignoran la evidencia que no nos gusta. (ver apartado anterior).

Efecto de superioridad foto: Presentaciones con gráficos tienen más credibilidad.

Sesgo autogeneración: Tendemos a recordar lo que decimos mejor que lo que otros dicen. Por lo tanto, si usted hace una pregunta durante su presentación visual, y deja que el público llegue a la conclusión entonces ellos recordarán mejor la respuesta.

Sesgo de Ilusión de verdad: Es más probable que identifiquemos como verdaderas las declaraciones que vemos o escuchamos en reiteradas ocasiones. Si tomamos el caso de Twitter, que un tweet haya sido retuiteado por múltiples fuentes no significa que sea correcto. Por ejemplo, el documento de austeridad anterior fue citado más de 2000 veces antes de que se descubriera el sesgo.

Engaño óptico - Otra forma en la que los elementos visuales pueden ser ilusorios son los engaños ópticos. En 2015, en el artículo “*How Deceptive are Deceptive Visualizations?*”, el Dr. Pandey estudió cómo tales manipulaciones se utilizan en gráficos [127].

La mayoría de las manipulaciones en los gráficos consisten en jugar con los niveles de zoom de los ejes y rezar para que nadie lo note. Tenga en cuenta que no es una excusa para decir que no ocultaron la escala del eje, la forma en que Pandey mide el engaño es midiendo lo que la gente entendió después de ver una imagen, no lo que está escrito en la imagen.

El caso de incendios forestales de las Amazonas

El sesgo se puede encontrar en múltiples niveles. Por ejemplo, al recordar los incendios de la selva amazónica de Brasil en el verano de 2019, sabemos que esta historia dominó los principales medios de comunicación del mundo por varias semanas, ya que algunos lectores expresaron su indignación por el impacto que los incendios tendrían sobre el clima de la tierra. Aplicar lo que hemos aprendido hasta el momento para encontrar posibles sesgos en el

siguiente gráfico. Tiempo de 5 minutos.

Ejercicio

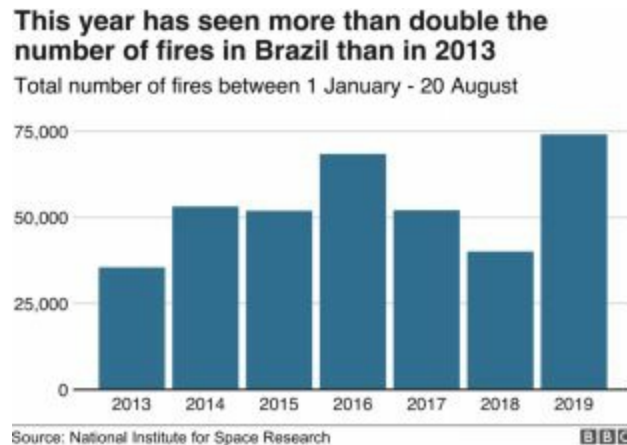


Fig 91 ¿Dónde está el sesgo?

Solución

Los datos fueron centrales en esta historia. ¿Qué tan malo fueron los incendios? Había una fuente de datos principal - el Instituto Nacional de Investigaciones Espaciales (INPE^[25]). Inicialmente, la cobertura de noticias presentó datos que alentaron al lector a indignarse y sospechar que se avecinaba una crisis.

Los datos del gráfico se muestran en el número de incendios desde 2013 hasta 2019. El gráfico lleva al lector a pensar que este era el nivel más alto de incendios de la historia. De hecho, CNN el 22 de agosto de 2019 escribió una historia en la que dice que el bosque estaba “ardiendo a un ritmo récord”. En la primera frase de esta historia CNN dijo que la velocidad de combustión fue un récord “desde que el INPE comenzó a dejar registro audiovisual de los incendios en 2013”. Entre el uso de palabras como 'doble' y 'registro' y el uso de una imagen con la barra de 2019 que se cierne sobre todos los demás años expuestos, la narrativa fue establecida. Y sin dudas, agencias como CNN agregaron al título de su nota escrita “... y los científicos advierten que podría dar un golpe devastador para la lucha contra el cambio climático.”

This year has seen the highest number of fires in Brazil since 2010

Total number of fires, 1 January - 27 August (1998-2019)

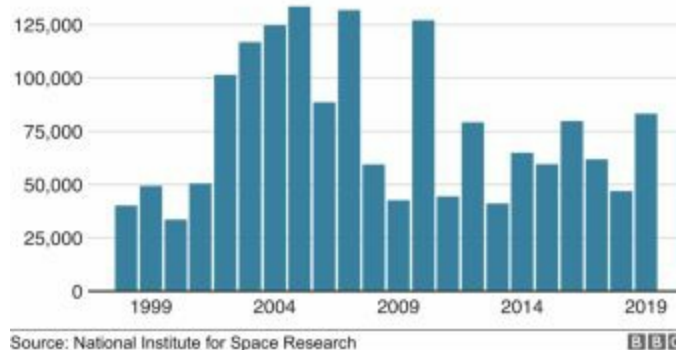


Fig 92 El buen periodismo?

Sin embargo, una señal de diligencia debida es siempre preguntar si hay más datos. A menudo están allí, pero no están disponibles en línea o no son gratis. Pueden estar en el estante de una biblioteca o detrás del muro de pago de un medio. Pueden existir, pero el reportero, siempre trabajando contra el tiempo no tuvo tiempo de mirar, o incluso posiblemente lo ignoró de manera intencional.

Annual Deforestation in the Brazilian Amazon

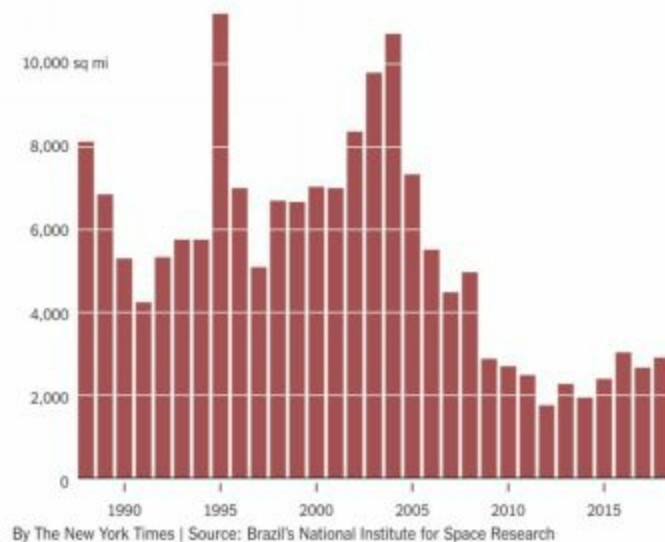


Fig 93 Un mejor periodismo.

En este caso, INSE sí tenía más datos en cuanto al número de incendios

(¡que no estaban online!). La BBC actualizó su gráfico para incluir más años. Aquí se muestra cómo se veía el gráfico al profundizaren los archivos que no estaban en línea.

Sin embargo ¿es esta la mejor información que podemos obtener? ¿Hay más? Considere lo que se está midiendo: número de incendios. ¿Quiere decir que si encendemos cinco pequeños incendios hoy y mañana uno grande, el número total de incendios está disminuyendo? A continuación, el New York Times publicó el siguiente gráfico. En este caso, tenemos la misma fuente, INSE, pero la medida usada fueron las millas cuadradas que se han incendiado. Los números utilizados inicialmente no estaban mal, sino que no estaban completos o no eran un reflejo fiel de la situación, un caso de sesgo de selección inconsciente.

Lista de control de diligencia debida

1. Prestar atención a las palabras

Cualquier conjunto de buenos datos ofrecerá transparencia respecto a la metodología con la que se han obtenido los datos. Esto significa prestar especial atención a qué preguntas hacer en las encuestas y cómo hacerlas preguntas, y también en la forma en que se presentan las declaraciones hechas. Siempre hay que levantar una bandera roja ante cualquier uso de adverbios y adjetivos. Por lo general, se cargan con un sesgo

1. Sigue el dinero. ¿Quién pagó por la investigación?

Las grandes tabacaleras nos mostraron que la organización que paga la investigación puede controlar sus resultados. Por ejemplo, los grupos de presión de la industria del huevo están pagando por una investigación en universidades acreditadas para promover evidencia que muestra que los huevos no aumentan el colesterol malo en los seres humanos. En la década de 1960, la industria azucarera pagó a los investigadores para que produjeran datos que hicieran creer a los consumidores que la grasa era un peligro mayor para la salud que el azúcar.

La lista continúa. La próxima vez que vea una investigación sobre la salud o el medioambiente, trate de descubrir la identidad del patrocinador financiero final. Siga el dinero.

2. Prestar atención a los métodos estadísticos utilizados

Como ya hemos visto, a veces los datos han sido seleccionados para apoyar intencionadamente una posición. Después de realizar un análisis estadístico, una buena regla general es corroborar siempre los resultados con un científico de datos más competente para que ayude a encontrar los defectos. Funciona de maravillas y se puede aprender mucho.

3. Tenga en cuenta la disponibilidad de los datos

El hecho de que los datos no estén disponibles en Internet de manera gratuita no significa que no existan. Los periodistas post Millennials - a los que nunca se les enseñó cómo investigar antes de que existiera Internet- son particularmente vulnerables a este sesgo.

Cuestionario: Tweets incendiario



Fig 94 *¿Dónde está el problema en este tweet?*

Responder: La foto es falsa. Al usar la búsqueda de Fotos de Google se descubrió que corresponde a la vista aérea de un incendio en la Estación Ecológica de Taim, en Río Grande do Sul, a sur de Brasil, el 27 de marzo de 2013. No se trata de los incendios forestales amazónicos de 2019. El New York Times escribió el 23 de agosto sobre cómo se difundieron imágenes engañosas imágenes durante los incendios forestales del Amazonas. Fuente: [113].



Fig 95 Dirección metáfora polo, más centrado en las personas de una lista.

Reconocimiento

Este libro recibió la ayuda de muchos amigos. Un agradecimiento especial a Cheryl Shuttle, Ricardo Bragança, Liz Ortiguera, Agatha Braja, Kumiko Sasaki, Louise Taylor, Natalia Rodríguez, Eric Nylen, Ferran Pujol, Eva Mitchell, Cliff Hazel, James Piecoway, Amal, el usuario Kaggle *Head or Tails* y Paul. Extiendo también mi gratitud a Anthony Goldblom, Carolina Ampuero, Francois Cadiou, Nicolás París, Alex Ríos, Steven Anthony, a la Biblioteca DiDi, al Comité Organizador de la Conferencia del Gobierno digital de 2019, SNASK, Susanna Vento, Kai Bruns, 710stories, al equipo de Ivory&Gold, y al equipo Kaggle. El estilo de libro electrónico se basa en la 3ª edición de “No me hagas pensar”, libro electrónico; por lo cual estoy muy agradecido del autor.

Referencias

- [1] (2020, January 1). Solving The Equation: The Variables For Women's Success In Engineering And Computing. AAUW: *Empowering Women Since 1881*. Retrieved February 22, 2020, from <https://web.archive.org/web/20200222084817/https://www.aauw.org/research/solving-the-equation/>
- [2] Statt, N. (2018, November 3). Over 20,000 Google Employees Participated In Yesterday's Mass Walkout. *The Verge*. Retrieved February 22, 2020, from <https://web.archive.org/web/20200222084920/https://www.theverge.com/2018/11/2/18057716/google-walkout-20-thousand-employees-ceo-sundar-pichai-meeting>
- [3] Zhang, V. (n.d.). Breaking Down The Gender Gap In Data Science. *Forbes*. Retrieved February 22, 2020, from <https://web.archive.org/web/20200222085154/https://www.forbes.com/sites/womensmedia/2017/08/03/down-the-gender-gap-in-data-science/>
- [4] <https://www.Kaggle.com/paultimothymooney/2018-Kaggle-machine-learning-data-science-survey>
- [5] https://en.wikipedia.org/wiki/Generations_in_the_workforce
- [6] Sinton, E (2011). 'Baby boomers are very privileged human beings' <https://www.telegraph.co.uk/finance/personalfinance/pensions/8840963/Baby-boomers-are-very-privileged-human-beings.html> retrieved October 23, 2013 from www.telegraph.co.uk
- [7] Ken Blanchard Companies. (2009). Next Generation of workers. http://www.kenblanchard.com/img/pub/Blanchard_Next_Generation_of_Workers.pdf Retrieved October 14, 2013, from kenblanchard.com
- [8] Adecco Group UK and Ireland. (nod). Managing the modern workforce. <http://www.adeccogroupuk.co.uk/SiteCollectionDocuments/Adecco-Group-Workplace-Revolution.pdf> Retrieved October 13, 2013, from www.Adeccouk.co.uk
- [9] Own sources
- [10] https://en.wikipedia.org/wiki/Affluence_in_the_United_States
- [11] <https://www.epi.org/blog/top-1-0-percent-reaches-highest-wages-ever-up-157-percent-since-1979/>
- [12] J. Berengueres, Sketch thinking. 2016
- [13] https://en.wikipedia.org/wiki/Marimekko#Marimekko_chart
- [14] Palmore, Erdman. Ageism: Negative and positive. Springer Publishing Company, 1999.
- [15] <https://www.Kaggle.com/ash316/Kaggle-journey-2017-2018>
- [16] <https://en.wikipedia.org/wiki/BRICS>
- [17] <https://www.Kaggle.com/harriken/brics-growth>
- [18] See primary vs. secondary color in <https://material.io/design/color/the-color-system.html#color-theme-creation>
- [19] Dutta, S., Reynoso, R.E., Garanasvili, A., Saxena, K., Lanvin, B., Wunsch-Vincent, S., León, L.R. and Guadagno, F., 2018. THE GLOBAL INNOVATION INDEX 2018: ENERGIZING THE WORLD WITH INNOVATION. GLOBAL INNOVATION INDEX 2018, p.1.
- [20] CSV file global innovation in <https://www.globalinnovationindex.org/analysis-indicator>
- [21] World Bank, <https://data.worldbank.org/indicator/SP.POP.TOTL>
- [22] https://en.wikipedia.org/wiki/Kendall_rank_correlation_coefficient
- [23] <https://www.pinterest.com/pin/281615782925970581/?lp=true>
- [24] https://en.wikipedia.org/wiki/Regression_toward_the_mean
- [25] <https://www.Kaggle.com/harriken/residuals-fig8b-test>
- [26] <https://www.Kaggle.com/Kaggle/Kaggle-survey-2018>
- [27] <https://www.Kaggle.com/harriken/storytelling-the-2018-Kaggle-survey>

- [28] Isaacson, Walter. Steve Jobs. 2011.
- [29] Where in the world is it easiest to get rich? | Harald Eia .
- [30] Rose, T. (2016). The end of average: How to succeed in a world that values sameness. Penguin UK.
- [31] Hyman, J., 2006. The objective eye: color, form, and reality in the theory of art. University of Chicago Press.
- [32] https://en.wikipedia.org/wiki/Fear,_uncertainty,_and_doubt
- [33] https://en.wikipedia.org/wiki/Merchants_of_Doubt
- [34] Hern, A. and Pegg, D., 2018. Facebook fined for data breaches in Cambridge Analytica scandal. The Guardian, 11.
- [35] Reynolds, George. Ethics in information technology. Nelson Education, 2011.
- [36] Greene, J. D., Sommerville, R. B., Nystrom, L. E., Darley, J. M., & Cohen, J. D. (2001). An fMRI investigation of emotional engagement in moral judgment. Science, 293(5537), 2105-2108.
- [37] Schippers, M. C., Scheepers, A. W., & Peterson, J. B. (2015). A scalable goal-setting intervention closes both the gender and ethnic minority achievement gap. Palgrave Communications, 1, 15014.
- [38] https://www.youtube.com/watch?v=w1-HMMX_NR8
- [39] Christensen, C. M., Hall, T., Dillon, K., & Duncan, D. S. (2016). Know your customers' jobs to be done. Harvard Business Review, 94(9), 54-62.
- [40] Knee, J. A. (2007). The accidental investment banker: Inside the decade that transformed Wall Street. Random House Incorporated.
- [41] Ogilvy, D. (2006). Ogilvy on Advertising.
- [42] Parvizi, J., Jacques, C., Foster, B. L., Withoft, N., Rangarajan, V., Weiner, K. S., & Grill-Spector, K. (2012). Electrical stimulation of human fusiform face-selective regions distorts face perception. Journal of Neuroscience, 32(43), 14915-14920.
- [43] <https://www.rideapart.com/articles/254402/safety-8-ways-to-make-your-motorcycle-more-visible/>
- [44] <https://www.Kaggle.com/harriken/police-dogs-and-grey-hair-will-save-you-from-jail>
- [45] Kondō, M. (2014). The life-changing magic of tidying up: The Japanese art of decluttering and organizing. Ten Speed Press.
- [46] Johnstone, Keith. Impro for storytellers. Routledge, 2014.
- [47] Johnstone, Keith. Impro: Improvisation and the theatre. Routledge, 2012.
- [48] Cowan, N., 2010. The magical mystery four: How is working memory capacity limited, and why?. Current directions in psychological science, 19(1), pp.51-57.
- [49] Gallo, C. 2019. The Art of Persuasion Hasn't Changed in 2,000 Years.
- [50] Tufte, E.R., 2001. The visual display of quantitative information (Vol. 2). Cheshire, CT: Graphics press.
- [51] <https://www.weforum.org/agenda/2017/11/why-iceland-ranks-first-gender-equality/>
- [52] Lukianoff, G. and Haidt, J., 2018. The coddling of the American mind: How good intentions and bad ideas are setting up a generation for failure. Penguin.
- [53] <https://megaphone.link/VMP1005647522>
- [54] Norman, D., 2004. Affordances and design. Unpublished article, available online at: <http://www.jnd.org/dn.mss/affordances-and-design.html>.
- [55] Norman, D., 2013. The design of everyday things: Revised and expanded edition. Basic books.
- [56] Berengueres, J., 2019, June. Visualization & Storytelling Workshop. In 20th Annual International Conference on Digital Government Research (pp. 532-533). ACM.
- [57] Rand, P., 1985. Paul Rand: A designer's art. New Haven: Yale University Press.
- [58] <https://youtu.be/Plf8slbBzzQ?t=240>
- [59] Berengueres, J. (2018). "Valuation of Crypto-Currency Mining Operations", The Journal of

Cryptocurrency and Blockchain Technology Research “Ledger”, Vol. 3, pp 60-67.

[60] Rosling, H., Rönnlund, A.R. and Rosling, O., 2018. Factfulness: wie wir lernen, die Welt so zu sehen, wie sie wirklich ist. Ullstein Buchverlage.

[61] Grant, R.M., 1991. Porter's 'competitive advantage of nations': an assessment. Strategic management journal, 12(7), pp.535-548.

[62] Galton, F., 1886. Regression towards mediocrity in hereditary stature. The Journal of the Anthropological Institute of Great Britain and Ireland, 15, pp.246-263.

[63] Knafllic, C.N., 2015. Storytelling with data: A data visualization guide for business professionals. John Wiley & Sons.

[64] Berinato, S., 2016. Good charts: The HBR guide to making smarter, more persuasive data visualizations. Harvard Business Review Press.

[65] Roam, D., 2011. Blah blah blah: What to do when words don't work. Penguin.

[66] Duarte, N., 2013. Resonate: Present visual stories that transform audiences. John Wiley & Sons.

[67] McCandless, David. Information is beautiful. London: Collins, 2012.

[68] Dweck, C., 2015. Carol Dweck revisits the growth mindset. Education Week, 35(5), pp.20-24.

[69] Berengueres, J., 2007. The Toyota production system re-contextualized. Lulu. com.

[70] Levitt, S.D. and Dubner, S.J., 2014. Freakonomics. B DE BOOKS.

[71] Harari, Y.N., 2016. Homo Deus: A brief history of tomorrow. Random House.

[72] Backes-Gellner, U. and Veen, S., 2013. Positive effects of ageing and age diversity in innovative companies—large-scale empirical evidence on company productivity. Human Resource Management Journal, 23(3), pp.279-295.

[73] Mourant, R.R. and Rockwell, T.H., 1972. Strategies of visual search by novice and experienced drivers. Human factors, 14(4), pp.325-335.

[74] Simon Wardley OSCON 2014 Keynote: "Introduction to Value Chain Mapping"

[75] EE Annual Conference 2017: Simon Wardley keynote

[76] <https://love-ely.blogspot.com/2010/08/what-phd-mean-for-you.html>

[77] Eames, Charles, and Ray Eames. "Powers of ten, 1977." (1977).

[78] <https://howmuch.net/articles/worlds-money-in-perspective-2018>

[79] Sinek, S., 2009. Start with why: How great leaders inspire everyone to take action. Penguin.

[80] Dollar Street. <https://www.gapminder.org/dollar-street/matrix?lowIncome=135&highIncome=2114>

[81] [gapminder.org](https://www.gapminder.org)

[82] Dabbs Jr, J.M., Chang, E.L., Strong, R.A. and Milun, R., 1998. Spatial ability, navigation strategy, and geographic knowledge among men and women. Evolution and human behavior, 19(2), pp.89-98.

[83] Salas, C., Broglio, C. and Rodríguez, F., 2003. Evolution of forebrain and spatial cognition in vertebrates: conservation across diversity. Brain, behavior and evolution, 62(2), pp.72-82.

[84] Wynn, T., 2000. Symmetry and the evolution of the modular linguistic mind. Evolution and the human mind: modularity, language and meta-cognition, pp.113-139.

[85] Wynn, T., 2002. Archaeology and cognitive evolution. Behavioral and brain sciences, 25(3), pp.389-402.

[86] Tanaka, K., 1997. An introduction to fuzzy logic for practical applications.

[87] <https://www.kaggle.com/harriken/kaggle-journey-2017-2018>

[88] Barrett, L.F., 2017. How emotions are made: The secret life of the brain. Houghton Mifflin Harcourt.

[89] Williams, D., 2012. The trickster brain: Neuroscience, evolution, and narrative. Lexington Books.

[90] Zak, P.J., 2014. Why your brain loves good storytelling. Harvard business review, 28.

- [91] Jefferys, S., 2003. *Liberté, Egalité and Fraternité at work: changing French employment relations and management*. Springer.
- [92] Thunberg, G., 2018. Speech at UN Climate Change COP 24 Conference, Poland 2018. Published online on YouTube by Connect4Climate.
- [93] Gandhi, P., Khanna, S. and Ramaswamy, S., 2016. Which industries are the most digital (and why). *Harvard Business Review*, 1.
- [94] Berengueres, J., 2015. *The Brown Book of Design Thinking: A workshop based approach*. UAE University College. <https://books.apple.com/mt/book/the-brown-book-of-design-thinking/id761406970>
- [95] <https://learnwardleymapping.com>
- [96] Burnett, W. and Evans, D.J., 2016. *Designing your life: How to build a well-lived, joyful life*. Knopf.
- [97] <https://square.github.io/crossfilter/>
- [98] Introduction chapter in, Kim, W.C. and Mauborgne, R.A., 2014. *Blue ocean strategy, expanded edition: How to create uncontested market space and make the competition irrelevant*. Harvard business review Press.
- [99] <http://index.wrapp.com>
- [100] Gore, A., 2006. *An inconvenient truth: The planetary emergency of global warming and what we can do about it*. Rodale.
- [101] <https://www.kaggle.com/harriken/visualizing-and-storytelling-sales>
- [102] Naremore, J. ed., 2004. *Orson Welles's citizen Kane: a casebook*. Oxford University Press on Demand.
- [103] <https://www.linkedin.com/posts/activity-6567737343908909056-nUHW/>
- [104] http://blogs.reading.ac.uk/climate-lab-book/files/2016/06/spiral_2017_large-1.gif
- [105] <http://www.met.reading.ac.uk/~ed/home/twitter.php>
- [106] Berengueres, J. and Castro, D., 2017, December. Differences in emoji sentiment perception between readers and writers. In *2017 IEEE International Conference on Big Data (Big Data)* (pp. 4321-4328). IEEE.
- [107] https://en.wikipedia.org/wiki/Magic_Quadrant
- [108] <https://www.kdnuggets.com/2019/09/core-hot-data-science-skills.html>
- [109] <https://www.kdnuggets.com/2019/10/4-quadrants-data-science-skills-data-visualization.html>
- [110] Guy Kawasaki, *The Art of Start*, TIECON 2006
- [111] David, Paul A., and Joseph S. Shapiro. "Community-based production of open-source software: What do we know about the developers who participate?." *Information Economics and Policy* 20, no. 4 (2008): 364-398.
- [112] Okinawa diet [Willcox, B. J.; Willcox, D. C.; Todoriki, H.; Fujiyoshi, A.; Yano, K.; He, Q.; Curb, J. D.; Suzuki, M. (October 2007), "Caloric Restriction, the Traditional Okinawan Diet, and Healthy Aging: The Diet of the World's Longest-Lived People and Its Potential Impact on Morbidity and Life Span" (PDF), *Annals of the New York Academy of Sciences*, 1114 (1): 434–455, Bibcode:2007NYASA1114..434W, doi:10.1196/annals.1396.037, PMID 17986602
- [113] <https://www.nytimes.com/2019/08/23/world/americas/amazon-rainforest-fire-photos.html>
- [114] Economic status cues from clothes affect perceived competence from faces, *Nature Human Behaviour* (2019). DOI: 10.1038/s41562-019-0782-4, <https://nature.com/articles/s41562-019-0782-4>
- [115] Tversky, A. and Kahneman, D., 1974. Judgment under uncertainty: Heuristics and biases. *science*, 185(4157), pp.1124-1131.
- [116] Thomas, A.K. and Millar, P.R., 2011. Reducing the framing effect in older and younger adults by encouraging analytic processing. *Journals of Gerontology Series B: Psychological Sciences and Social Sciences*, 67(2), pp.139-149.
- [117] Sweller, J., 1988. Cognitive load during problem solving: Effects on learning. *Cognitive science*, 12(2), pp.257-285.

- [118] Kahneman, D. and Tversky, A., 1981. On the study of statistical intuitions (No. TR-6). STANFORD UNIV CA, DEPT OF PSYCHOLOGY.
- [119] Source: <https://www.bi.team/blogs/in-the-frame-how-policy-choices-are-shaped-by-the-way-ideas-are-presented/> (retrieved 7 December 2019)
- [120] Matsumoto, D. and Juang, L., 2016. Culture and psychology. Nelson Education.
- [121] <https://uxcam.com/blog/gestalt-principles/> (retrieved on 7 dec 2019)
- [122] Stewart, P.J., 2019. Mendelev's predictions: success and failure. Foundations of Chemistry, 21(1), pp.3-9.
- [123] Ogilvy, D. and Horgan, P., 1963. Confessions of an advertising man (p. 45). New York: Atheneum.
- [124] Loschelder, D.D., Stuppi, J. and Trötschel, R., 2014. "€ 14,875?!": Precision boosts the anchoring potency of first offers. Social Psychological and Personality Science, 5(4), pp.491-499.
- [125] Cook, J., Supran, G., Lewandowsky, S., Oreskes, N. and Maibach, E., 2019. America Misled: How the Fossil Fuel Industry Deliberately Misled Americans about Climate Change.
- [126] Herndon, T., Ash, M. and Pollin, R., 2014. Does high public debt consistently stifle economic growth? A critique of Reinhart and Rogoff. Cambridge journal of economics, 38(2), pp.257-279.
- [127] Pandey, A.V., Rall, K., Satterthwaite, M.L., Nov, O. and Bertini, E., 2015, April. How deceptive are deceptive visualizations?: An empirical analysis of common distortion techniques. In Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems (pp. 1469-1478). ACM.
-

[1] Modos de pensar: divergentes (exploratorios), convergentes (síntesis)

[2] Estimaciones propias [27]

[3] "Toma demasiado tiempo poner el estilo que deseas en una tabla usando R o Excel. Al final es más rápido usar Adobe Illustrator"- Marybeth Sandell, ex jefe de la oficina de Londres, Bloomberg.

[4] Hay varias maneras de aprender algo. Escuchando, hablando de ello, mediante su lectura y al hacerlo. De los cuatro, "hacerlo" tiene el índice más alto en la de tasa de recuerdo.

[5] Origen de la imagen de Lego: LinkedIn, meme anónimo.

[6] Véase también "cómo detectar ideas geniales" en Ogilvy & Publicidad [41]

[7] DIKW (Datos - Información - Sabiduría) Modelo

[8] En 2006, Al Gore [100] contrató a Nancy Duarte para hacer su famosa gráfica de presentación del CO2. Un elevador de tijera se utilizó en el escenario para demostrar que el nivel CO2 está "fuera de la gráfica". Al Gore fue posteriormente criticado por la emisión de una gran cantidad de CO2, debido a su viaje en jet.

[9] Véase también David y Goliath en el libro de Samuel.

[10] La analítica se divide generalmente en tres categorías según la función: descriptiva, predictiva y prescriptiva.

[11] En el campo UX un ejemplo de una accesibilidad es algo con el que el usuario está familiarizado. La accesibilidad es utilizada para ayudar al usuario a comprender cómo funciona un nuevo producto. (Véase también Don Norman en El diseño de las cosas cotidianas [54])

[12] La técnica de proyección en dos dimensiones se ha utilizado con éxito en gráficos famosos como Wardley Maps, la Matriz BCG de cuota de crecimiento, la matriz Urgente-Importante y Cuadrante Mágico de Gartner.

[13] Fuente: Bloomberg BusinessWeek.

[14] Ilusión óptica Ebbinghaus.

[15] Fuentes: Pelle Guldberg Hansen, inudgeyou.com

[16] Khaneman lo cita como sistema de pensamiento rápido del cerebro.

[17] Fuente: Annemarie Bufer. Véase el artículo se extienden en [121]

[18] Véase también el término “Stevenote” en Wikipedia.

[19] En los campos de la ciencia y la ingeniería, el sesgo se define como opuesto a la exactitud, como una desviación sistemática de la media esperada.

[20] Fuente: https://en.wikipedia.org/wiki/List_of_cognitive_biases

[21] <https://skepticalscience.com/>, [125]

[22] Fuente: Departamento de Agricultura de EE.UU.

[23] Fuente: [112]. Dieta de Okinawa en 1950 se correlaciona con la esperanza de vida más largas y más saludables del mundo. ¿Se refuta la creencia popular de que una dieta desequilibrada no puede ser saludable?

[24] Debido a que los autores del artículo se negaron a retractarse de sus creencias después de que los datos mostraron lo contrario, esto también es evidencia de sesgo de confirmación. [Https://www.bbc.com/news/magazine-22223190]

[25] El Instituto Nacional de Investigaciones Espaciales. Se trata de una unidad de investigación del Ministerio de Ciencia, Tecnología e Innovación de Brasil

Acerca del autor

Jose Berengueres



Jose Berengueres es ingeniero de telecomunicación por la UPC de Barcelona y doctor en robótica por el TokyoTech. Kaggle competitions master y angel investor, desde el 2011 trabaja en la U.A.E University en Emiratos donde compagina su trabajo de profesor con mentoring a startups y asesoría en temas de data science, UX y design thinking.

Libros de este autor

Sketch Thinking

Are your sketching skills holding back your creativity? Do you feel constrained by your drawing skills but don't have time to enroll an illustration class? Sketching (for design) Thinking is a short, clear and brilliant collection of tips that will help you build your own graphic vocabulary. Combined with exercises and photos from the d.school workshop, this book is the perfect companion to power up your design thinking workshops.