



Taller de cocina de Datos

Exploración y Visualización
con **Power BI**

Objetivos del Taller

*Vamos a enfocarnos en mejorar la manera en la que **ordenamos** y **combinamos** la **información** que debemos **analizar** para luego explorar distintas formas de **visualizarla**, según su naturaleza.*

1. Fundamentos del análisis de datos

1.1. Breve (brevisima) introducción de la Teoría de la información.

2. Limpieza, Exploración y Visualización de datos

Basándonos en la herramienta de Power BI, vamos a aprender a:

2.1. Importar datos desde distintas fuentes, limpiarlos, ordenarlos y transformarlos.

2.2. Combinar y cruzar distintas tablas para crear modelos complejos.

2.3. Crear columnas dinámicas que permitan cálculos agregados.

2.4. Reglas básicas de visualización de datos y oferta de visualizaciones en Power BI.

1. Fundamentos del análisis de datos - Flow de la Información

1. Fundamentos del análisis de datos - Flow de la Información

La utilidad de los datos reside en su potencial para transformarse en acciones concretas que mejoren la gestión.

Claud Shannon, 1948:

“La Información es lo opuesto a la incertidumbre.”



“La Comunicación es la reducción a la incertidumbre.”

“Journey” de los datos



Importar



Estructurar



Conectar



Enriquecer



Analizar



Comunicar



2. Limpieza, Exploración y Visualización de datos

2. Limpieza, Exploración y Visualización de datos

Antes de transformar los datos en información útil es necesario "tenerlos" en nuestro poder (o power bi)



Importar



Estructurar



Conectar



Enriquecer



Analizar



Comunicar

**Vamos a realizar los seis pasos
del camino de los datos
utilizando el Power Bi como
herramienta integral.**

2. Limpieza, Exploración y Visualización de datos



Importar



Estructurar



Conectar



Enriquecer



Analizar



Comunicar

Antes de transformar los datos en información útil es necesario "tenerlos" en nuestro poder (o power bi)

1. Explicación del ambiente power query [t](#)
2. Importación de un csv (PC o HTML) [t](#)
3. Importación de múltiples archivos con igual formato desde carpeta
4. 1era fila como encabezado
5. Anexar

2. Limpieza, Exploración y Visualización de datos

Preparar los datos implica entender su distribución y naturaleza



Importar



Estructurar



Conectar



Enriquecer



Analizar



Comunicar

1. Análisis estructural de las variables (perfil de columnas)
2. Agrupar / Agregar
3. Formato de campos
4. Filtrar campos innecesarios
5. Editor avanzado

2. Limpieza, Exploración y Visualización de datos



Importar



Estructurar



Conectar



Enriquecer



Analizar



Comunicar

Definir columnas compartidas entre distintas tablas permite conectar diferentes objetos de estudios

1. Combinar
2. Variables clave
3. Conexiones izquierda <<>> derecha

2. Limpieza, Exploración y Visualización de datos

El valor agregado primordial por parte del analista de datos está en la construcción de información a partir del relacionamiento entre variables y el retrabajo de los datos



Importar



Estructurar



Conectar






Enriquecer



Analizar



Comunicar

1. Columnas personalizadas
 - a. formato de las columnas
 - b. datos agregados
 - c. ratios
 - d. columnas indexadoras (orden para variables categóricas)
2. Indicadores dinámicos   

2. Limpieza, Exploración y Visualización de datos

Con buenas preguntas en mente, el siguiente paso es volcar la información en el tablero



Importar



Estructurar



Conectar




Enriquecer



Analizar



Comunicar

1. Microsoft Marketplace 
2. Selección de variables y distintos modos de insertar el dato *(recuento distintivo, no resumir, etc.)*
3. Formatos de visualizaciones

2. Limpieza, Exploración y Visualización de datos

Los datos deben traducirse en historias y para eso debemos elegir cuidadosamente los medios que usamos para visualizar la información correcta.



Importar



Estructurar



Conectar



Enriquecer



Analizar



Comunicar



Mapa de John Snow (1854)

2. Limpieza, Exploración y Visualización de datos

Los datos deben traducirse en historias y para eso debemos elegir cuidadosamente los medios que usamos para visualizar la información correcta.



Importar



Estructurar



Conectar



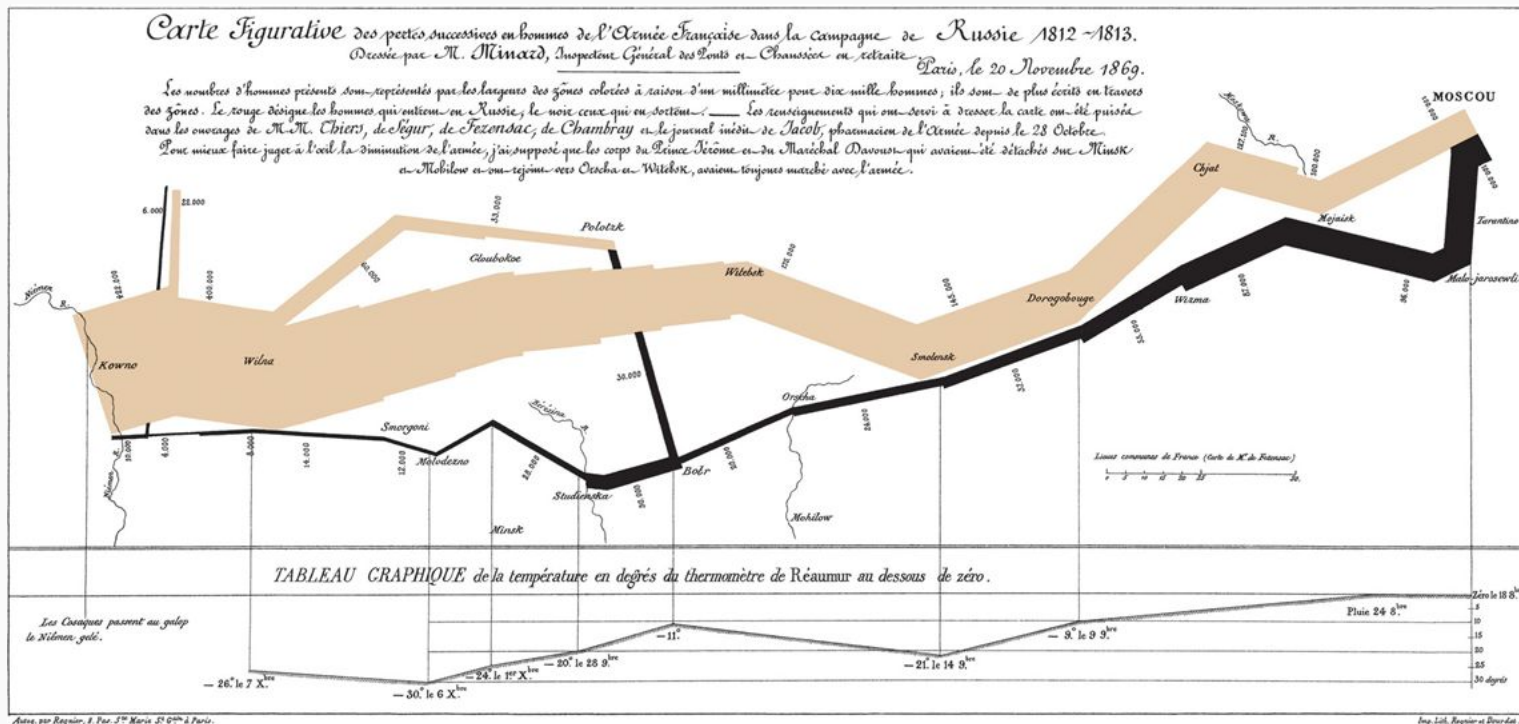
Enriquecer



Analizar



Comunicar



Mapa de Charles Minard (1869)

2. Limpieza, Exploración y Visualización de datos

Para algunos especialistas, la visualización de los datos no debe distraer al receptor del mensaje que se quiere transmitir.



Importar



Estructurar



Conectar



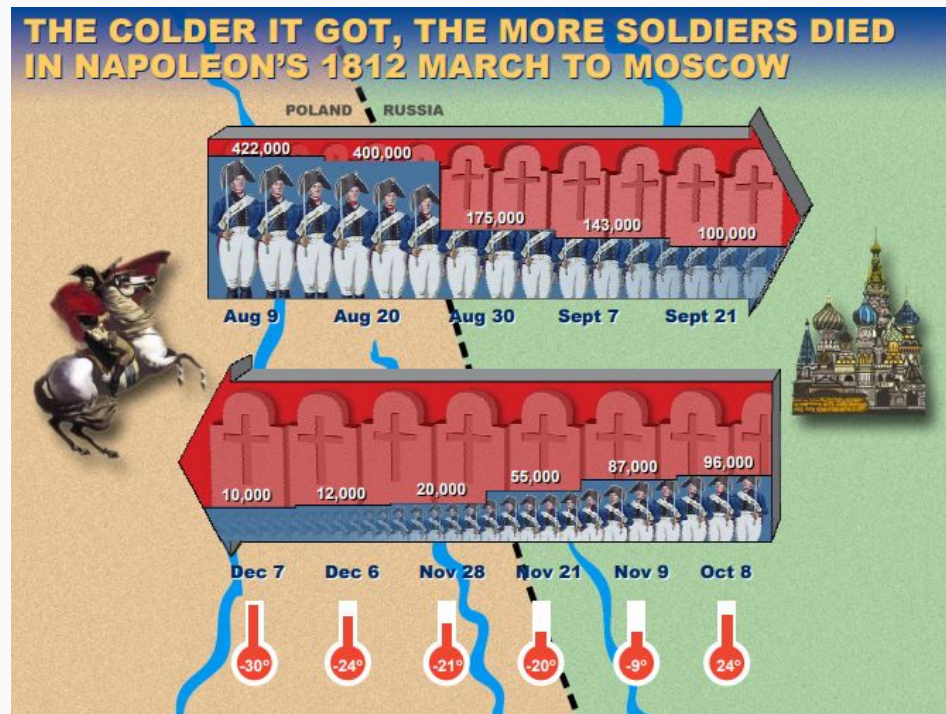
Enriquecer



Analizar



Comunicar



Gene Zelazny, "Say it with PowerPoint"

2. Limpieza, Exploración y Visualización de datos

Para algunos especialistas, la visualización de los datos no debe distraer al receptor del mensaje que se quiere transmitir.



Importar



Estructurar



Conectar



Enriquecer



Analizar



Comunicar

Remove
to improve
(the **data-ink** ratio)

Created by Darkhorse Analytics

www.darkhorseanalytics.com

2. Limpieza, Exploración y Visualización de datos

Para algunos especialistas, la visualización de los datos no debe distraer al receptor del mensaje que se quiere transmitir.



Importar



Estructurar



Conectar



Enriquecer



Analizar



Comunicar



2. Limpieza, Exploración y Visualización de datos



Importar



Estructurar



Conectar



Enriquecer



Analizar



Comunicar

Los modelos de visualización que presentan los softwares son justamente modelos: nosotros debemos adaptarlos a nuestras necesidades, agregando o quitando elementos.



2. Limpieza, Exploración y Visualización de datos

No obstante, las visualizaciones que transmiten conclusiones en forma gráfica suelen captar la atención del receptor de manera más eficiente y son más recordadas que las minimalistas.



Importar



Estructurar



Conectar



Enriquecer



Analizar



Comunicar

CHI 2010: Graphs

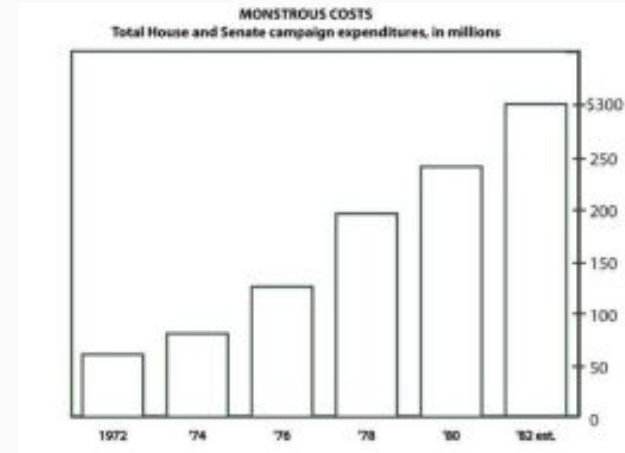
April 10–15, 2010, Atlanta, GA, USA

Useful Junk? The Effects of Visual Embellishment on Comprehension and Memorability of Charts

Scott Bateman, Regan L. Mandryk, Carl Gutwin,
Aaron Genest, David McDine, Christopher Brooks

Department of Computer Science, University of Saskatchewan, Saskatoon, Saskatchewan, Canada

scott.bateman@usask.ca, regan@cs.usask.ca, gutwin@cs.usask.ca,
aaron.genest@usask.ca, dam085@mail.usask.ca, cab938@mail.usask.ca



2. Limpieza, Exploración y Visualización de datos

Consejo #1: Para comparar datos, nunca hacer varios gráficos. Utilizar gráficos de comparación.



Importar



Estructurar



Conectar



Enriquecer

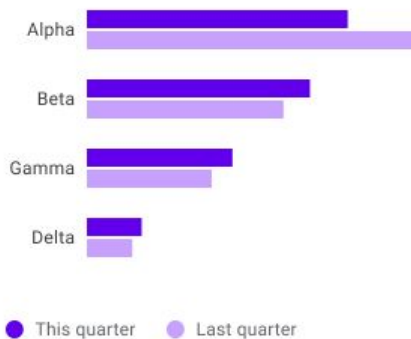


Analizar



Comunicar

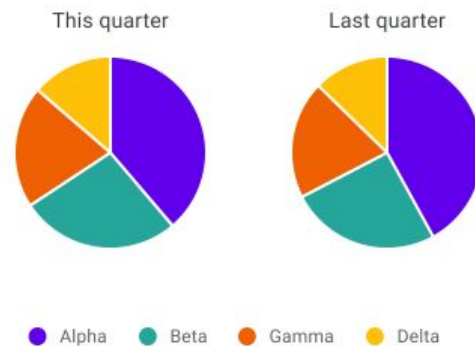
Sales by product



Do.

Use bar charts to show changes over time or differences between categories.

Sales by product



Don't.

Don't use multiple pie charts to show changes over time. It's difficult to compare the difference in size across each slice of the pie.

2. Limpieza, Exploración y Visualización de datos

Consejo #2: Cuidar que la información pueda visualizarse de forma clara e intuitiva. Evitar el overlapping de formas y colores.



Importar



Estructurar



Conectar



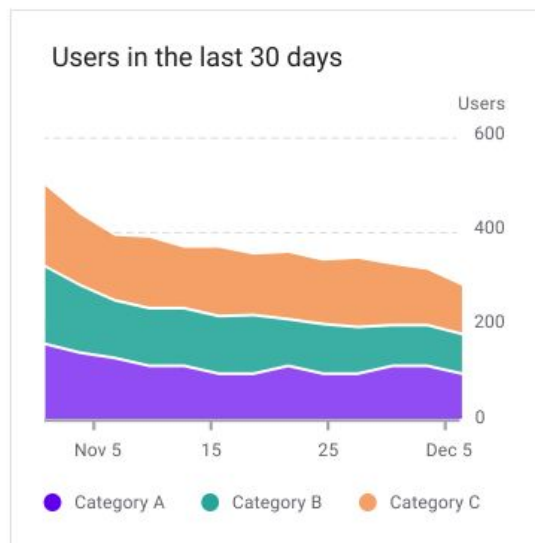
Enriquecer



Analizar

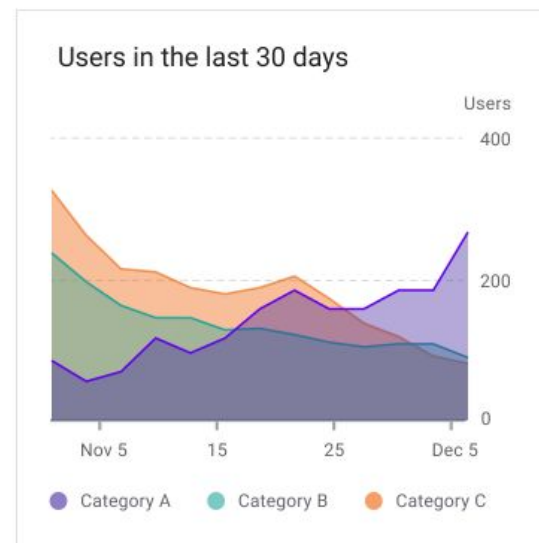


Comunicar



Do.

Use a stacked area chart to represent multiple time series and maintain a good level of legibility.



Don't.

Don't use overlapped area charts as it obscures data values and reduces readability.

2. Limpieza, Exploración y Visualización de datos

Consejo #3: Keep it simple. Las formas particulares pueden confundir el mensaje y dificultar la lectura.



Importar



Estructurar



Conectar



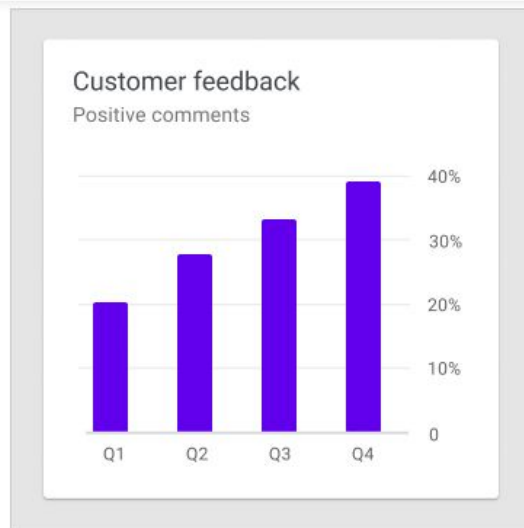
Enriquecer



Analizar

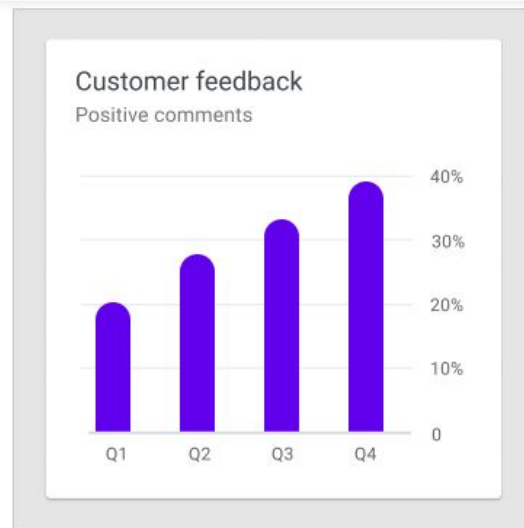


Comunicar



Do.

The bars in this chart have subtle rounded corners, ensuring that the top of the bar precisely measures the bar's length.



Don't.

Don't use shapes that make it hard to read a chart, such as bars with imprecise top edges.

2. Limpieza, Exploración y Visualización de datos

Consejo #4 Solo se usan colores distintos para significar categorías distintas. En esos casos, siempre debe haber una etiqueta o leyenda.



Importar



Estructurar



Conectar



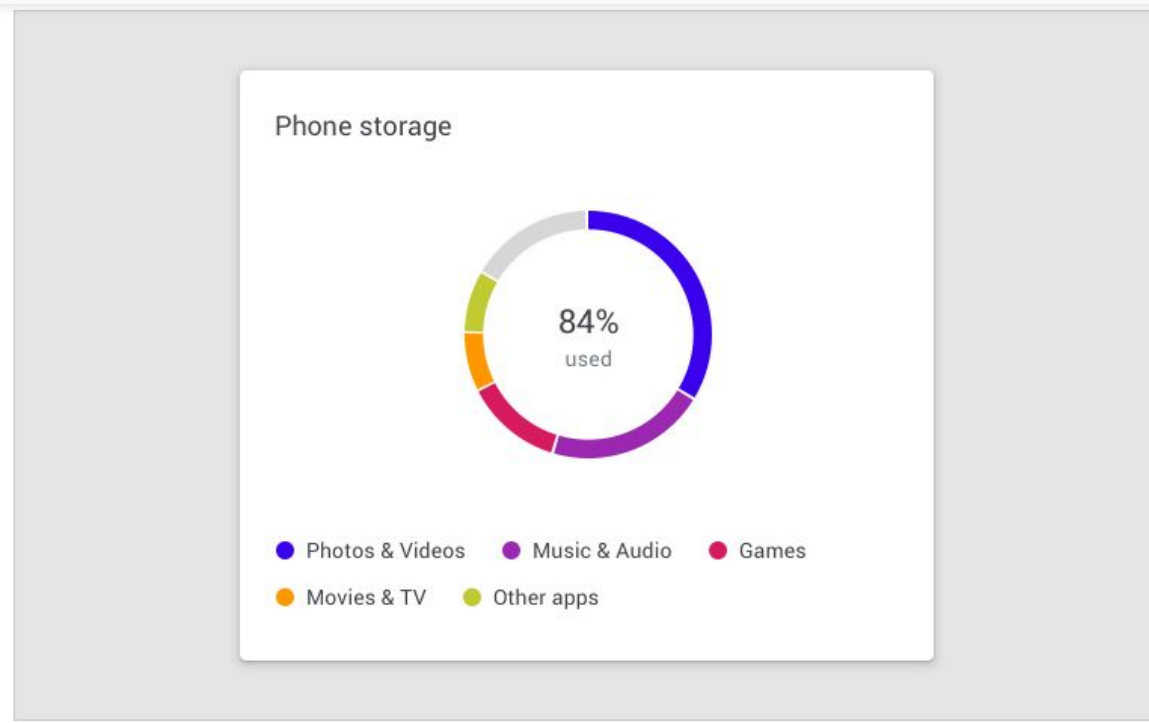
Enriquecer



Analizar



Comunicar



2. Limpieza, Exploración y Visualización de datos



Importar



Estructurar



Conectar



Enriquecer

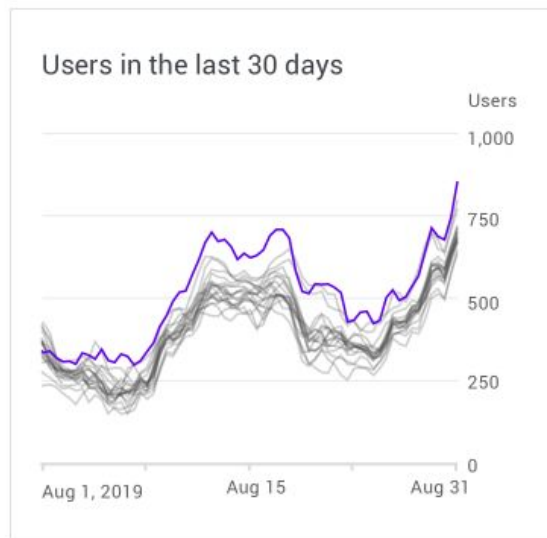


Analizar



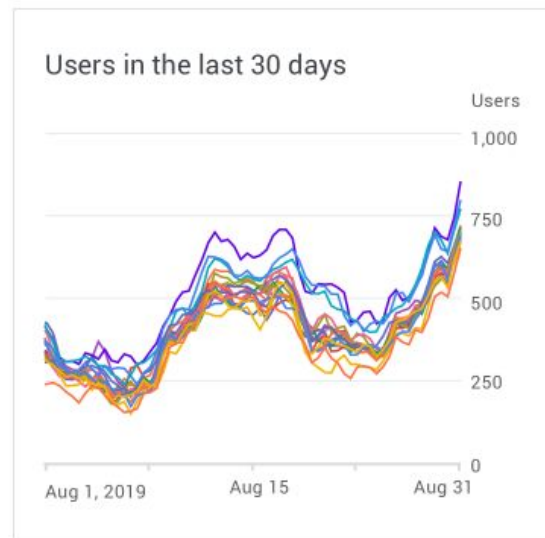
Comunicar

Consejo #5: Los colores deben ser claramente distinguibles. El uso de transparencias puede ayudar a resaltar unos por sobre otros



Do.

Use a combination of color highlights and neutral colors to provide contrast and emphasis.



Caution.

Many colors in a single chart can hinder focus.

2. Limpieza, Exploración y Visualización de datos

Consejo #6: Una variable = Un color, siempre. Para mostrar distintas variaciones de la misma variable pueden usarse, por ejemplo, diferentes texturas de línea.



Importar



Estructurar



Conectar



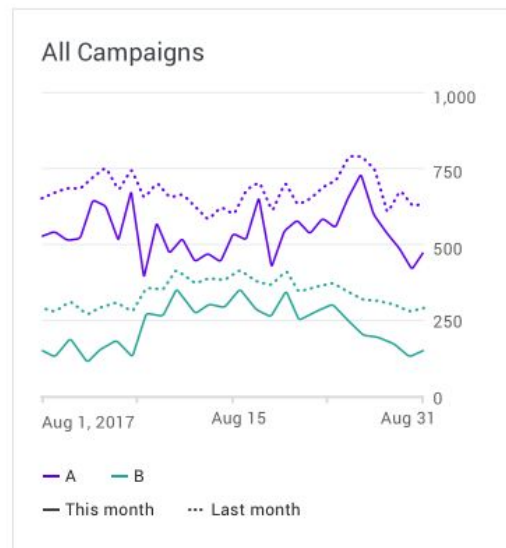
Enriquecer



Analizar

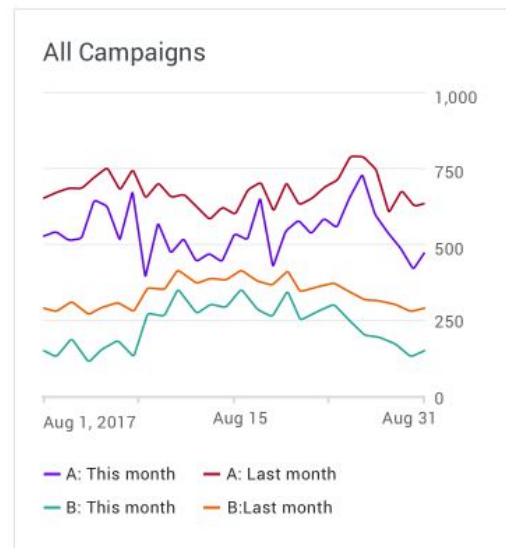


Comunicar



Do.

Vary a line's texture to represent different data types.



Don't.

Don't use different colors to show periodical variation for the same data category.

2. Limpieza, Exploración y Visualización de datos

Consejo #7: Usar íconos, pero siempre acompañados de su significado



Importar



Estructurar



Conectar



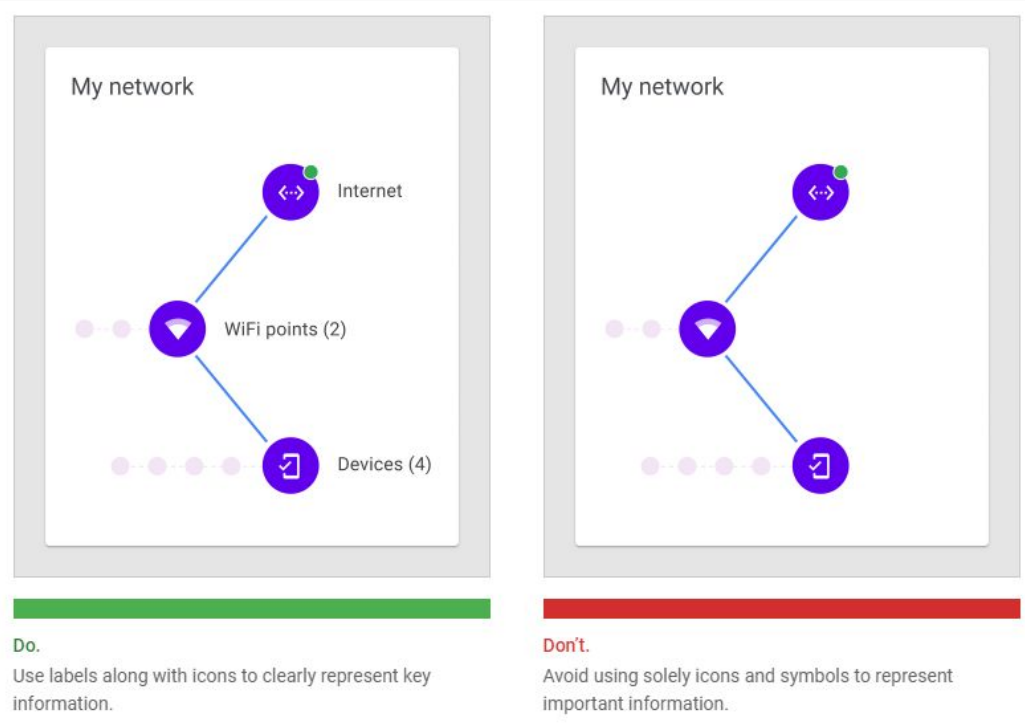
Enriquecer



Analizar



Comunicar



2. Limpieza, Exploración y Visualización de datos



Importar



Estructurar



Conectar



Enriquecer

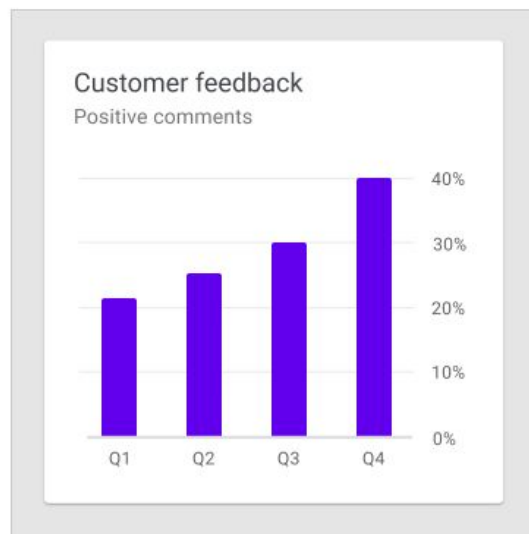


Analizar



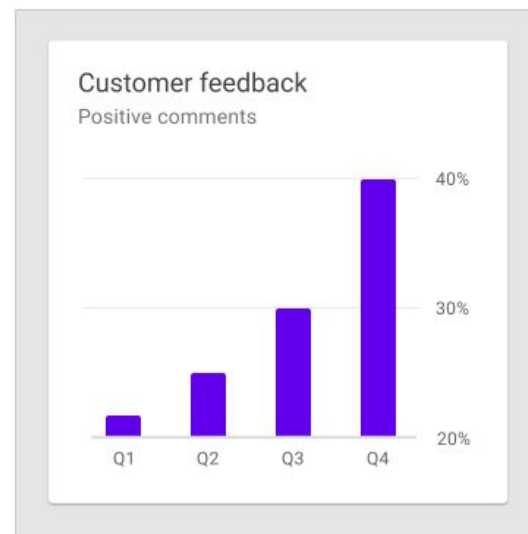
Comunicar

Consejo #8: En gráficos de barras, la línea de base siempre debe ser cero, para no exagerar las diferencias en la visualización.



Do.

A bar chart starting at the zero baseline



Don't.

Don't start the baseline at values other than zero. This baseline starts at 20%, making the bar differences look more dramatic.

2. Limpieza, Exploración y Visualización de datos

Consejo #9: No agregar demasiadas etiquetas en los ejes si no son necesarias.



Importar



Estructurar



Conectar



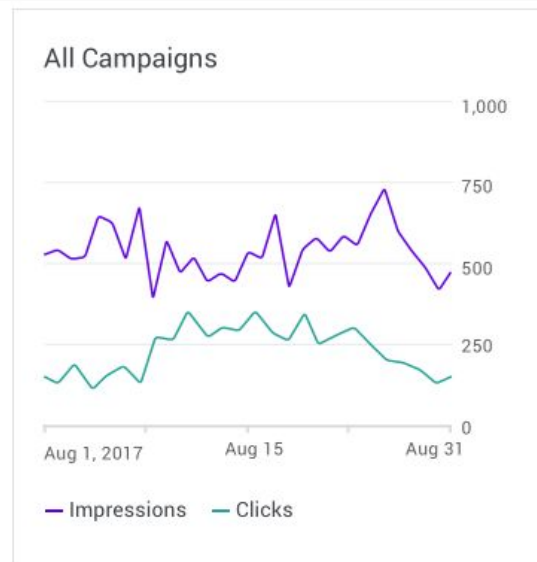
Enriquecer



Analizar

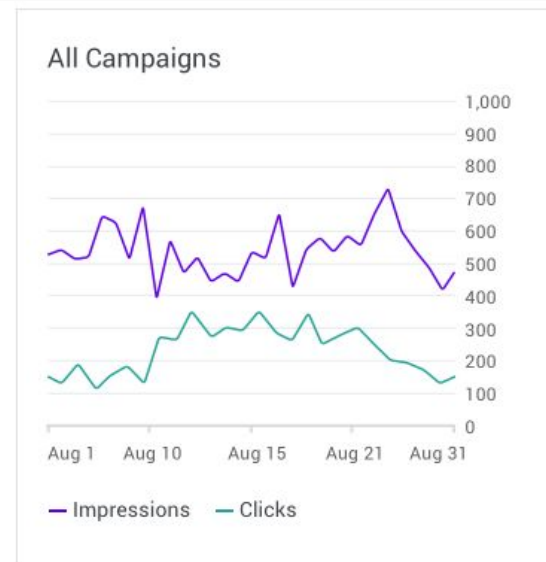


Comunicar



Do.

Support legibility by using a balanced number of axis labels.



Don't.

Don't overload the chart with numerous axis labels.

2. Limpieza, Exploración y Visualización de datos

Consejo #10: La rotación de etiquetas debe ser evitada siempre que se pueda, ya que dificulta la lectura. En esos casos es preferible rotar todo el gráfico.



Importar



Estructurar



Conectar



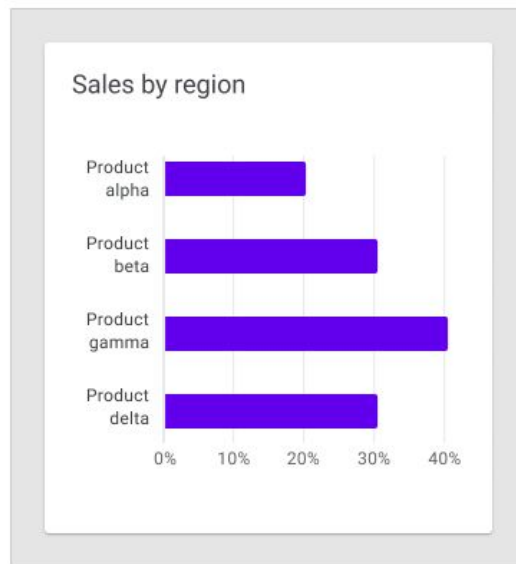
Enriquecer



Analizar



Comunicar



Do.

Orient text horizontally on bar charts, rotating the bars if needed to make space.



Caution.

Don't rotate bar labels, as it makes them difficult to read.

Taller de cocina de Datos

Exploración y Visualización



Maximiliano
Scocozza