



Visualización de Información y Analítica Visual

— Hernán Valdivieso López (hfvdivieso@uc.cl) —

Clase 9: Repaso y taller final

Contenidos

- Repaso del curso
 - Marcas & Canales
 - *Framework* Tamara Munzner
 - Gráficos comunes para datos tabulares
 - Teoría del color y Principios de Diseño
- Taller final

Marcas & canales

Marcas y canales

Marcas

- Elemento **geométrico básico**, que puede ser clasificado según el número de dimensiones espaciales que requiera.
- ¿Cómo representar cada dato de forma visual?

➔ Points



➔ Lines



➔ Areas



➔ Containment



➔ Connection



Marcas y canales

Canal

- Permite controlar la **apariciencia** de las marcas, independientemente de la dimensionalidad de este elemento primitivo.
- ¿Cómo puedo caracterizar cada dato de forma visual?

➔ Position

➔ Horizontal



➔ Vertical



➔ Both



➔ Color



➔ Shape



➔ Tilt



➔ Size

➔ Length



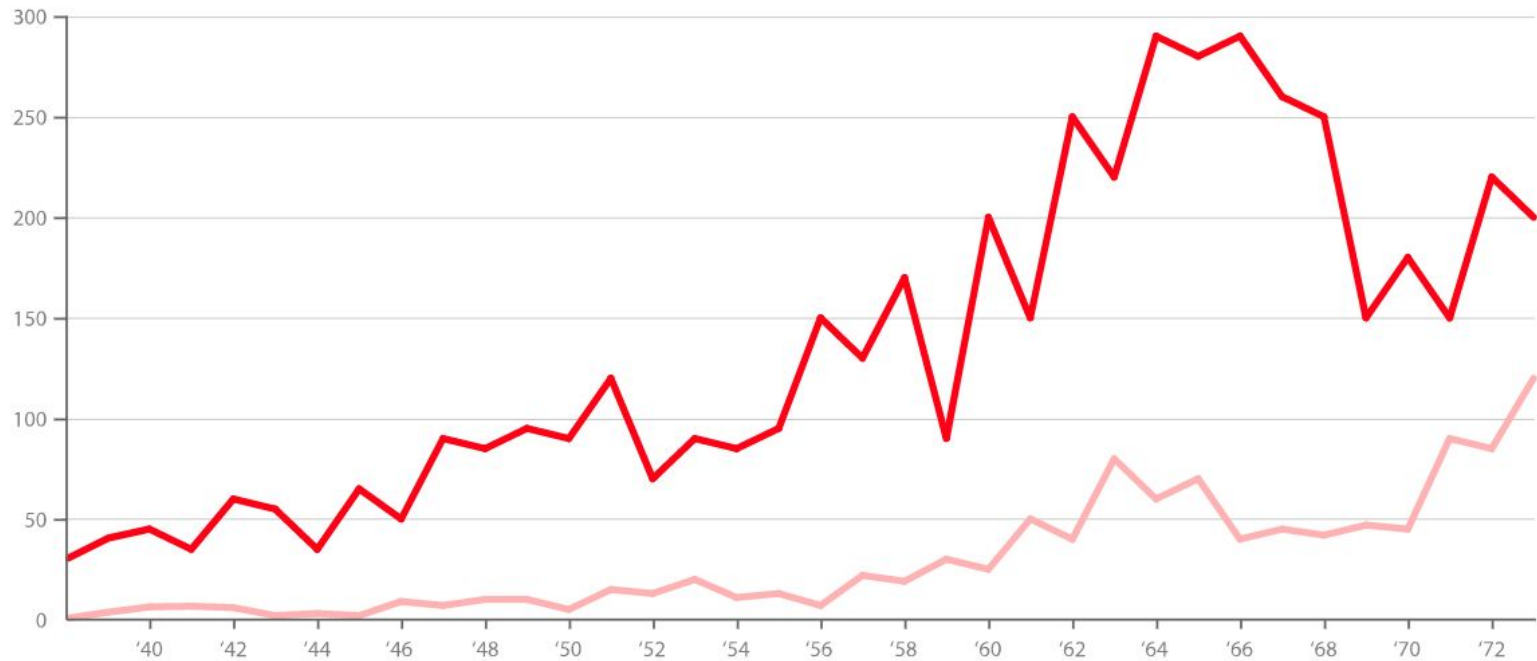
➔ Area



➔ Volume



Marcas y canales



Marcas y canales

Principio de expresividad

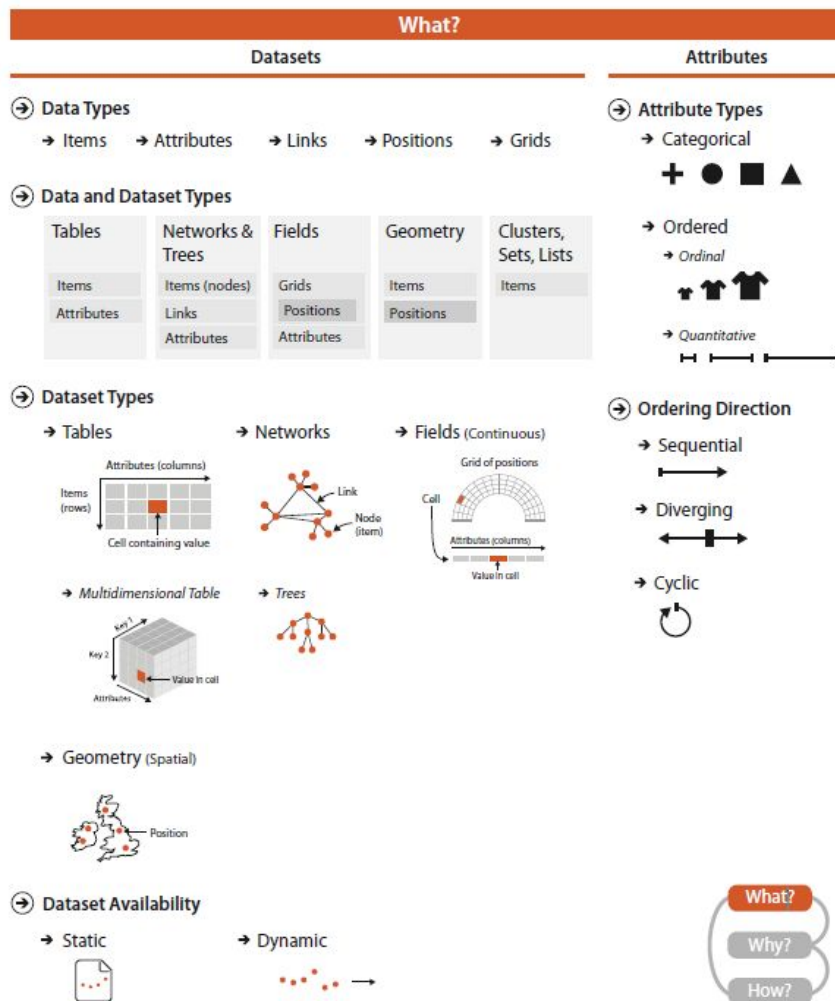
- Debe haber **coherencia** entre el tipo de canal (magnitud, identidad) con la semántica del atributo (cuantitativo, ordinal, categórico).
 - Los datos ordenados deben ser mostrados de tal forma que nuestro sistema perceptual los perciba como ordenados; inversamente, debe ocurrir lo mismo con los datos no ordenados.

Principio de efectividad

- Dicta que los **atributos más importantes** deben ser codificados con los **canales más efectivos**, para que sean más perceptibles.

Framework Tamara Munzner

What



Tipos de atributos

- **Categorico** → País, frutas, género de un libro, etc.
 - Atributos sin un orden intrínseco.
 - Puedes ordenarlos por otra información como por su nombre.
- **Ordenado:**
 - **Ordinal** → Talla ropa, ranking, etc.
 - Sin comparación aritmética exacta entre valores. Se pueden ver como datos **categoricos** que se pueden ordenar.
 - **Cuantitativo** → estatura, precio, temperatura, tiempo.
 - Puedes realizar operaciones matemáticas. (año promedio)
 - Puedes extraer datos matemáticos que tienen sentido.

Tipos de atributos

- Tipo de órdenes
 - **Secuencial:** Cantidad de un elemento, años (considerando sólo D.C), peso, altura de una persona.
 - **Divergente:** años (si se considera A.C y D.C), altura a nivel del mar.
 - **Cíclicos:** Meses, días de la semana, latitud-longitud.

Why

Why?

Actions

→ Analyze

→ Consume

→ Discover



→ Present



→ Enjoy



→ Produce

→ Annotate



→ Record



→ Derive



→ Search

	Target known	Target unknown
Location known	• • • Lookup	• • • Browse
Location unknown	◀ • • Locate	◀ • • Explore

→ Query

→ Identify



→ Compare



→ Summarize



Targets

→ All Data

→ Trends



→ Outliers



→ Features



→ Attributes

→ One

→ Distribution



→ Extremes



→ Many

→ Dependency



→ Correlation

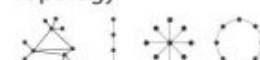


→ Similarity



→ Network Data

→ Topology



→ Paths



→ Spatial Data

→ Shape

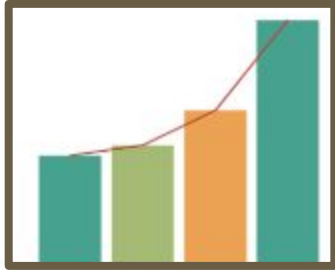


What?

Why?

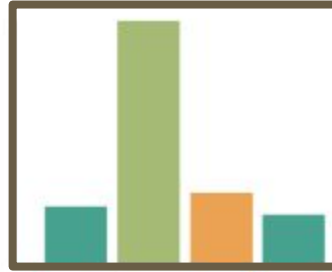
How?

Objetivos (targets)



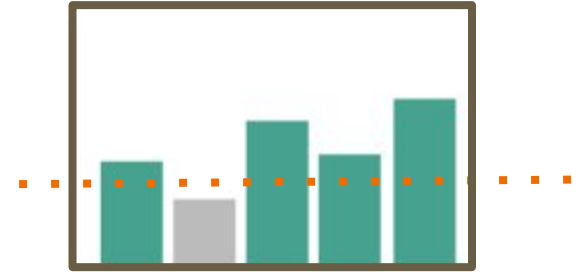
Tendencias

Descripción general
de un patrón en los
datos



Outliers

Datos atípicos, que
no siguen un patrón



Features
(Extractos)

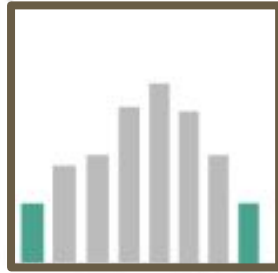
Conjunto de datos con
ciertas características
en común

Objetivos de atributos



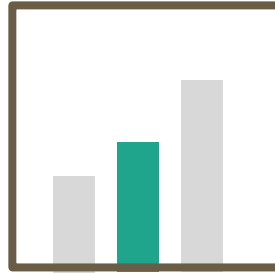
Distribución

Frecuencia de los valores presentes en un atributo



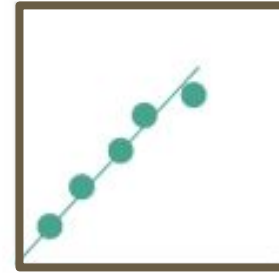
Extremos

Rango numérico en el que viven los atributos.



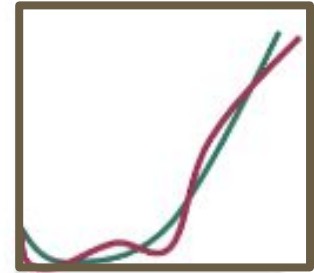
Valor singular

El valor específico de algún dato.



Correlación

Si la variación de un atributo depende de otro atributo.

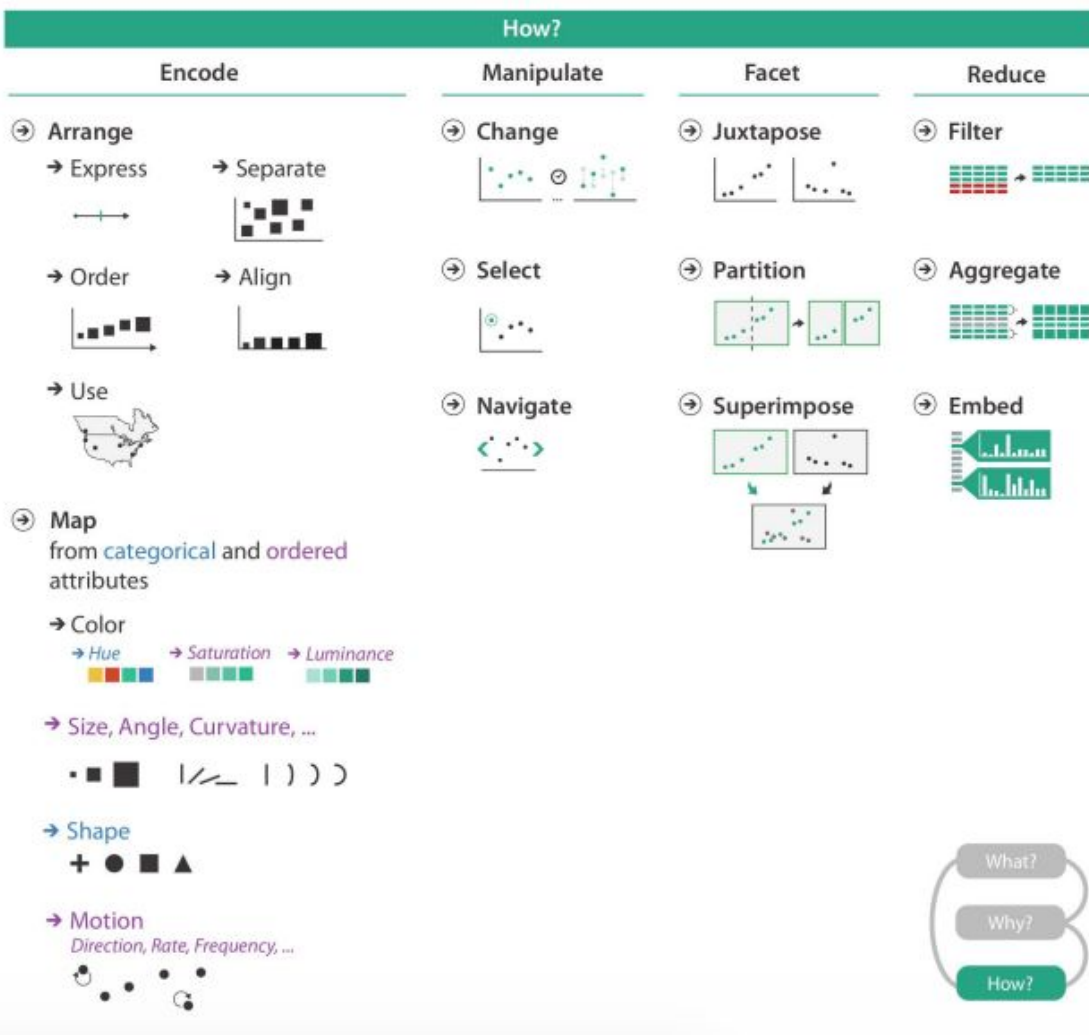


Similitud

Qué tan semejantes son los valores de 2 atributos.

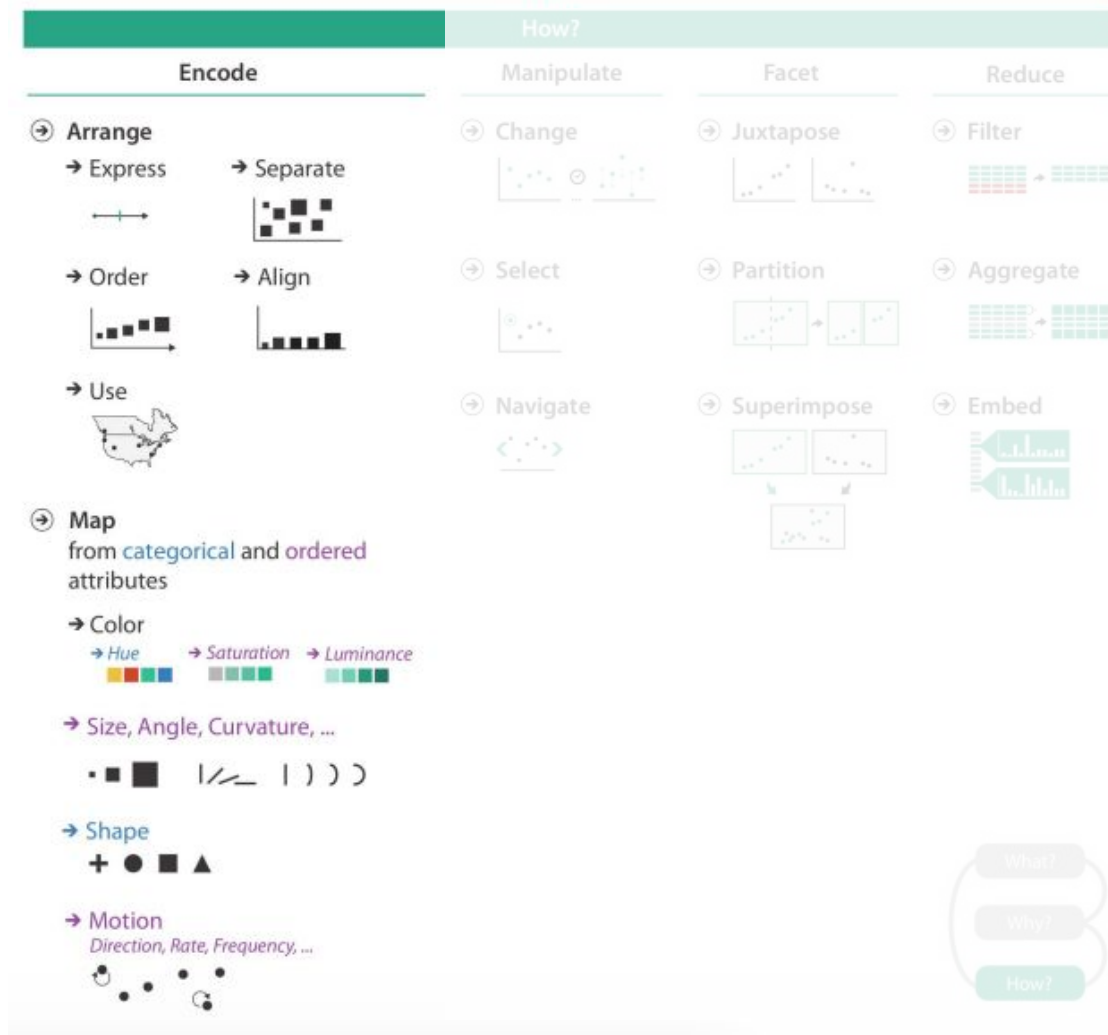
How

- **Visual encoding:** Establece una traducción entre datos y elementos visuales
- **Interactions:** Establece cómo el usuario va a interactuar con los elementos visuales



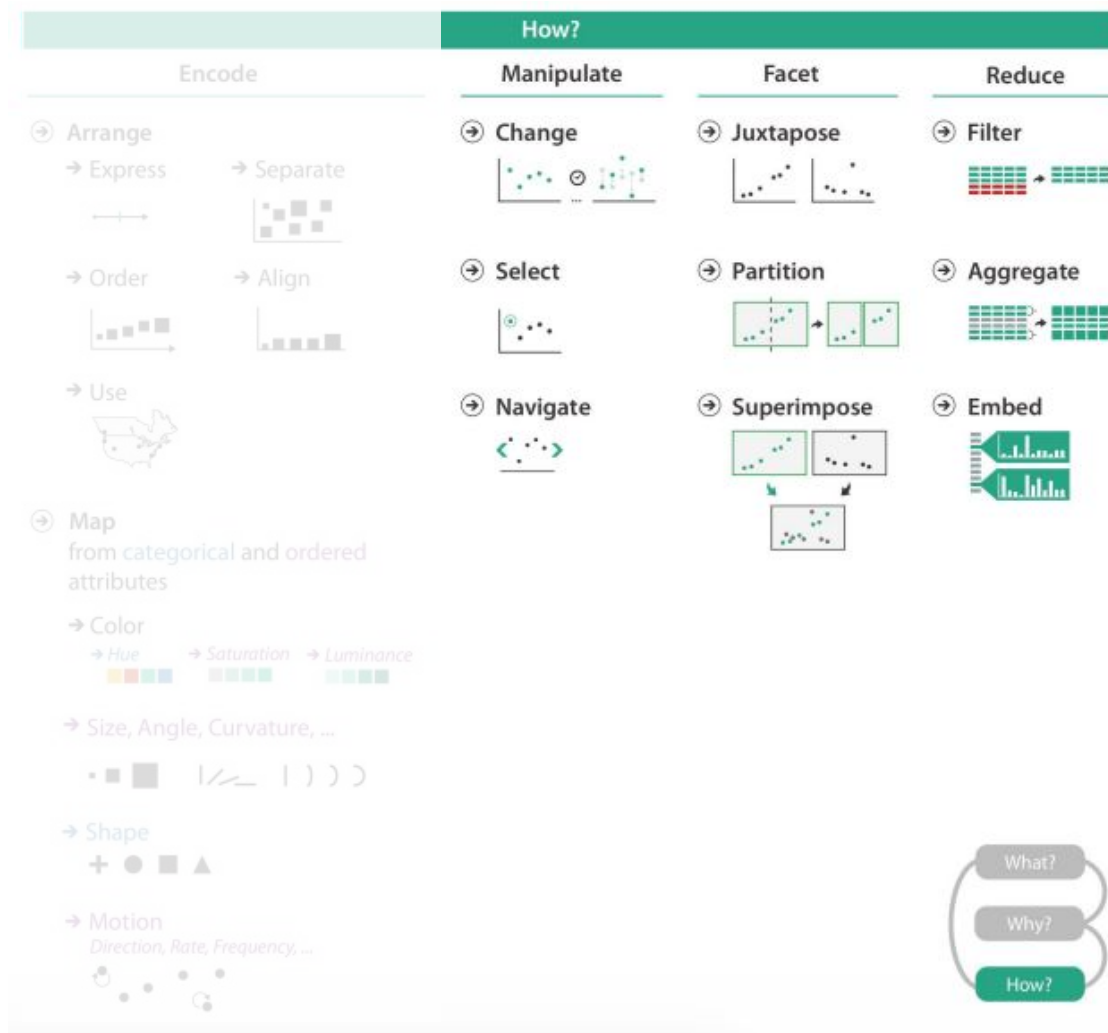
How

- **Arrange:** Saber cómo organizar los datos en el espacio, cómo expresar valores, cómo ordenarlos, cómo usar un dato (geográfico)
- **Map:** Aprovechar los canales visuales no-espaciales



How

- **Manipulate:** Acciones que puede realizar el usuario con elementos de la vis. (marcas)
- **Facet:** Mostrar diferentes ángulos de un dataset
- **Reduce:** Manejar la complejidad del dataset



Framework Tamara Munzner

¿Cómo aplicarlo?

Pasos para utilizar el *What*

1. Entender la **semántica** de los datos, ¿qué información representa cada dato?
2. Identificar el **tipo de dataset**.
3. En caso de identificar **atributos**. Para cada uno de los atributos...
 - a. Identificar el **tipo de atributo** que es: categórico, ordinal o cuantitativo.
 - b. En caso de ser ordinal o cuantitativo, identificar la **dirección de ordenamiento** (secuencial, divergente o cíclico).

Pasos para utilizar el *Why*

1. Partir desde el nivel más alto, **Analizar**, e identificar qué tarea se está presentando de dicho nivel (consumir o producir, junto con la tarea detallada).
2. Identificar el **tipo de búsqueda** que se hará: ¿el usuario sabe qué quiere buscar? ¿El usuario sabe donde encontrar lo que busca en la visualización?
3. Identificar el **tipo de consulta** que hará: ¿quiere identificar un dato en particular? ¿Quiere comparar N datos? ¿O quiere resumir toda o gran parte de la información en la visualización?
4. A las tareas identificadas, ver el **target** asociado: ¿tendencias? ¿Distribución de un atributo? ¿Caminos de un grafo?

Pasos para utilizar el *How*

1. Se recomienda iniciar con el **Mapeo** (*map*).

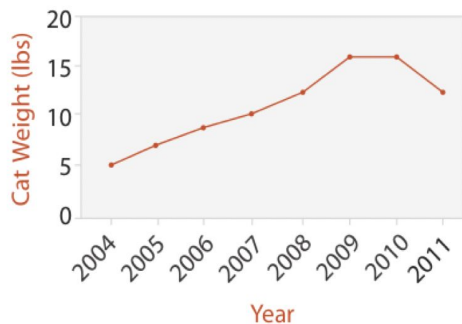
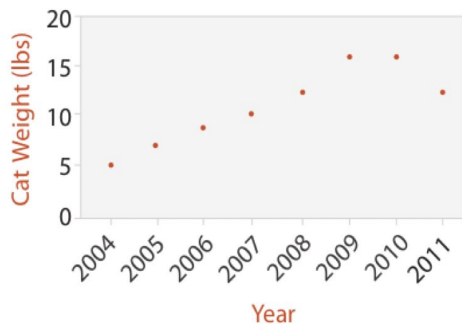
¿Qué canales se están utilizando y cuál atributo están codificando?

2. Analizar la **posición de los datos** (*arrange*): ¿la posición está expresando algún atributo? ¿están ordenados? ¿se utiliza un mapa o grilla?

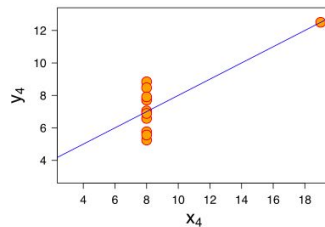
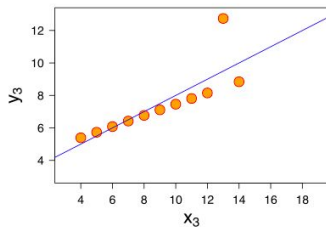
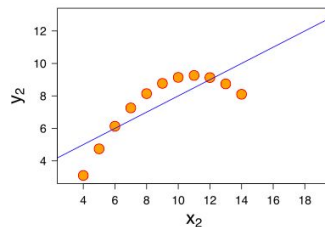
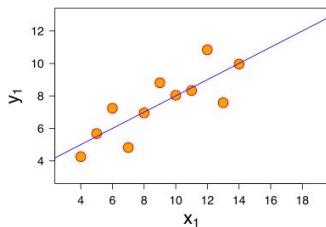
3. Finalmente, empezar con el análisis de **interactividad**: ¿hay alguna manipulación? ¿Se juega con diferentes vistas de *dataset* (múltiples gráficos por ejemplo) ? ¿Hay un sistema para agrupar o filtrar los datos dentro de la visualización?

Gráficos comunes para datos tabulares

Gráficos comunes para datos tabulares

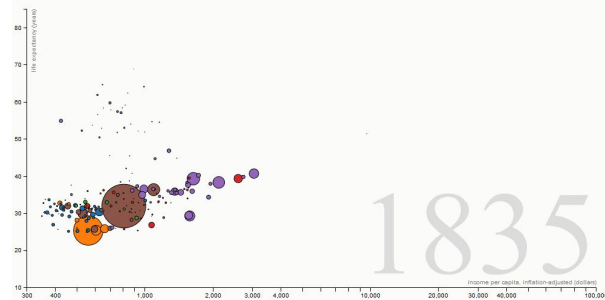


Puntos y líneas



Dispersión

The Wealth & Health of Nations



Burbuja

Gráficos comunes para datos tabulares

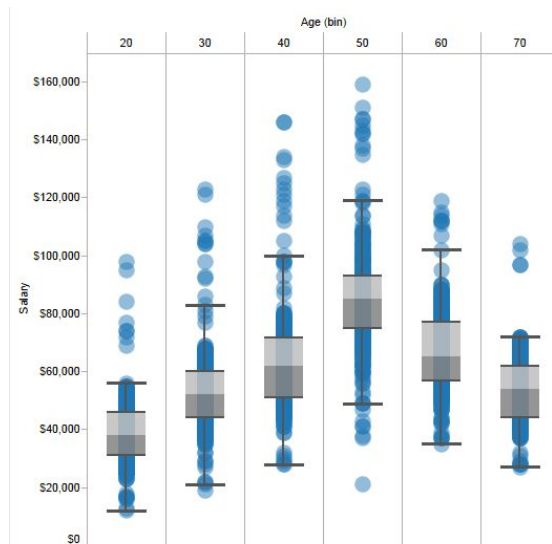


Gráfico de caja

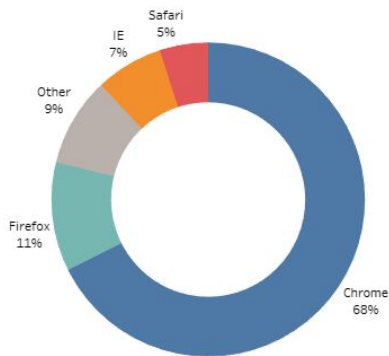
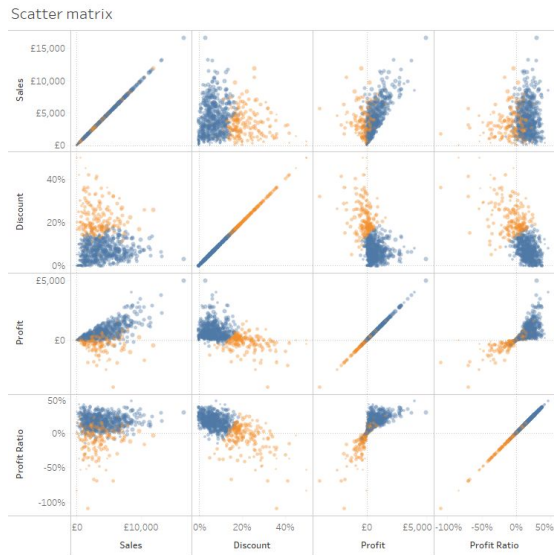


Gráfico de torta/donut

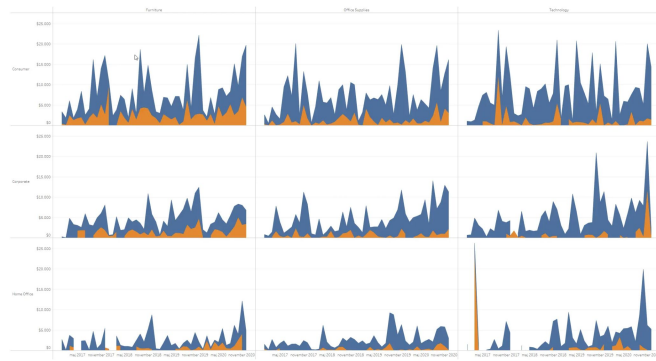


Maps de calor
Heatmap

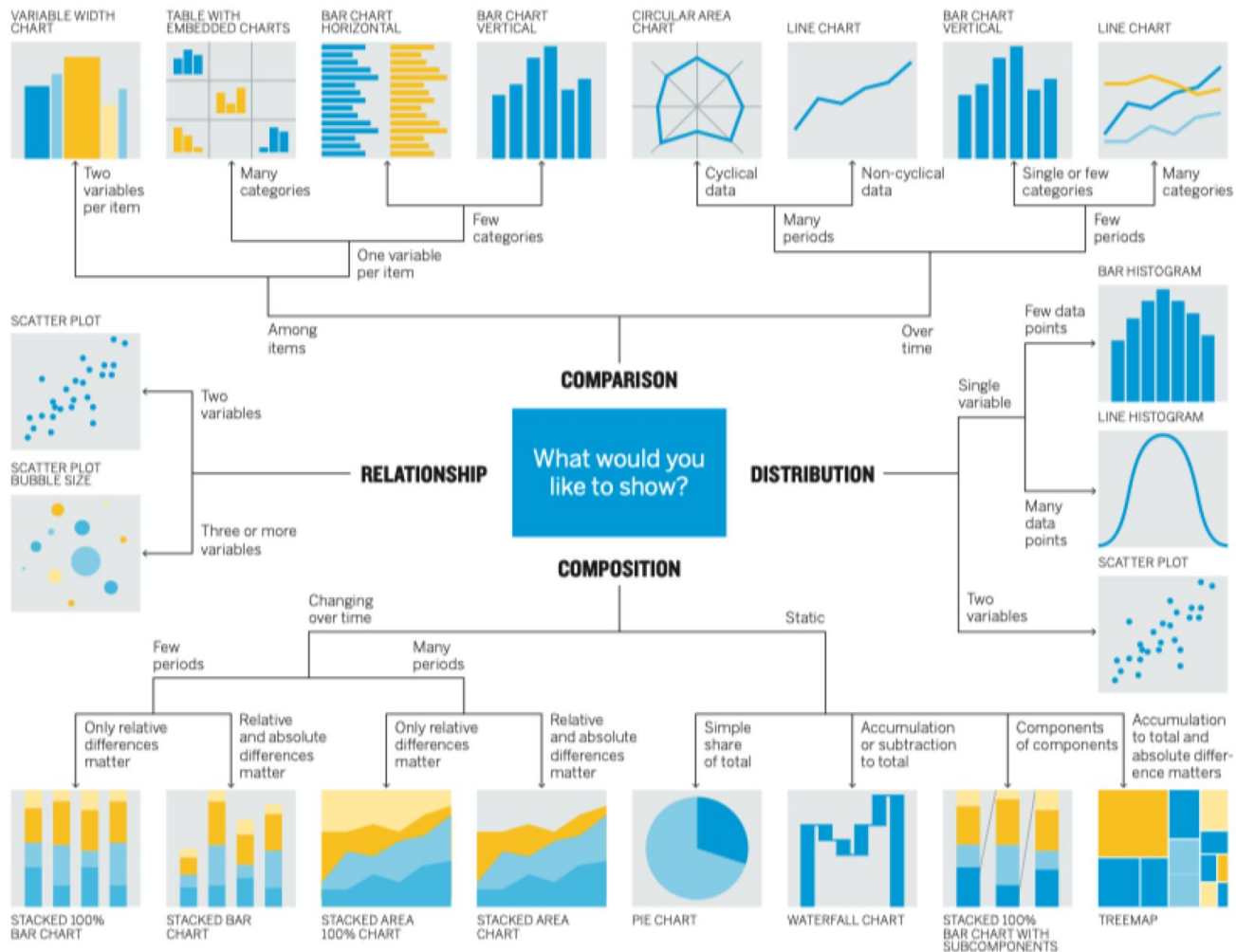
Gráficos comunes para datos tabulares



Matriz de gráficos



Small Multiples



Numeric

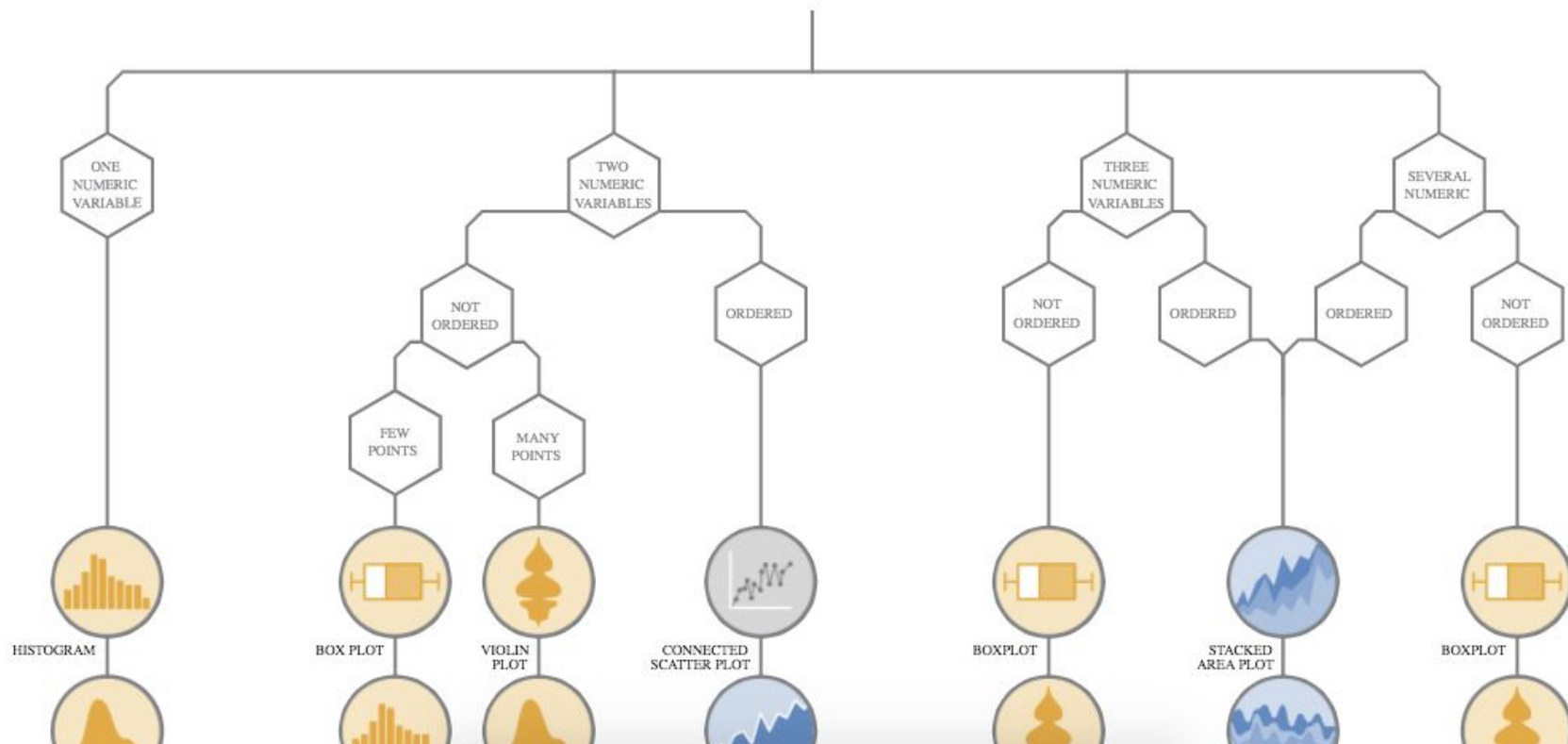
Categoric

Num & Cat

Maps

Network

Time series



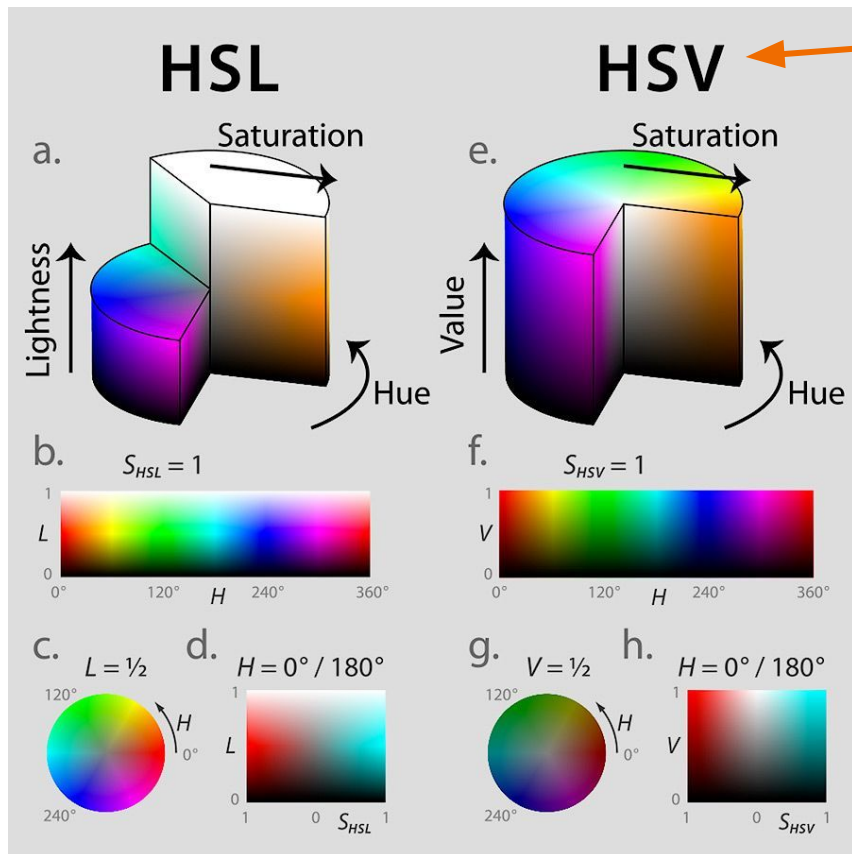
Gráficos comunes para datos tabulares

Códigos que serán de utilidad:

- [Clase 1 - Taller 1 - Pandas y Altair \(Evaluado\) \[Solución\].ipynb](#)
- [Clase 3 - Framework \(How\) aplicado en Altair](#)
- [Clase 4 - Visualizaciones Tabulares Matplotlib-Seaborn](#)
- [Clase 4 - Taller - Matplotlib, Seaborn y Crítica a visualización \(No evaluado\) \[Solución\]](#)

Teoría del color y Principios de Diseño

Modelos de color basado en propiedades



Extra - sólo con fines de que vean que hay más modelos basados en propiedades.

Principios de Diseño

... relacionados con la expresividad

- Tasa de tinta de datos (*Data ink ratio*)
- Factor de la mentira (*Lie factor*)
- Ejes engañosos
- Primero el fondo, luego la forma*

... relacionados con HCI

- Overview first, details on demand
- Los ojos le ganan a la memoria (*Eyes beat memory*)
- Tiene que ser receptivo (*Responsive is required*)

... relacionados con la efectividad

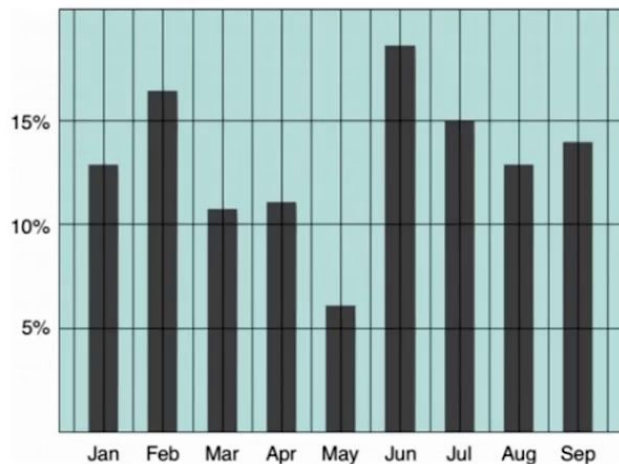
- No al 1D/2D/3D injustificado
- Lograrlo en blanco y negro (*Get it right in black and white*)
- Primero el fondo, luego la forma*

... relacionados con el diseño gráfico

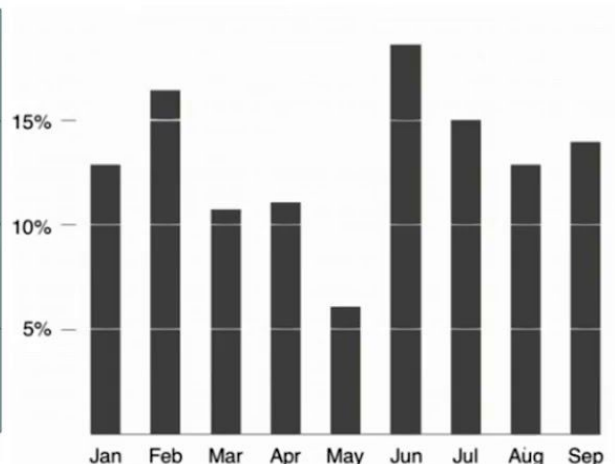
- Consistencia interna y externa
- Autocontención

Principios de Diseño

Data ink ratio (Edward Tufte)



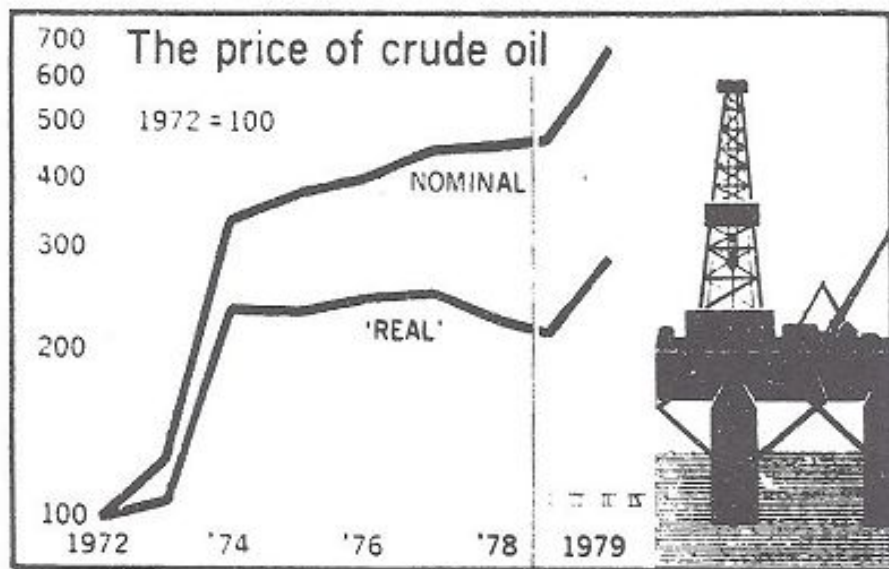
low data-ink ratio



high data-ink ratio

Principios de Diseño

$$\text{Lie Factor} = \frac{\text{Size of the effect in the visual}}{\text{Size of the effect in the data}}$$



Principios de Diseño

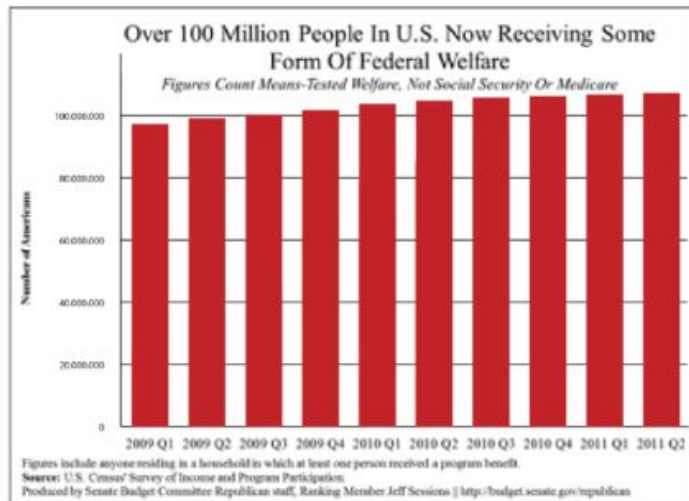
THE BLOG

Over 100 Million Now Receiving Federal Welfare

2:40 PM, AUG 8, 2012 • BY DANIEL HALPER

DETAIL PAGE PRINT LARGER TEXT SMALLER TEXT PARTS

A new chart set to be released later today by the Republican side of the Senate Budget Committee details a startling statistic: "Over 100 Million People in U.S. Now Receiving Some Form Of Federal Welfare."



lie factor: 1

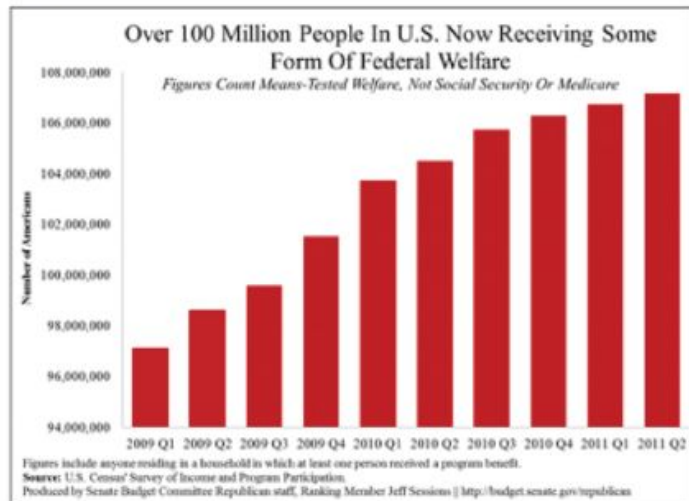
THE BLOG

Over 100 Million Now Receiving Federal Welfare

2:40 PM, AUG 8, 2012 • BY DANIEL HALPER

DETAIL PAGE PRINT LARGER TEXT SMALLER TEXT PARTS

A new chart set to be released later today by the Republican side of the Senate Budget Committee details a startling statistic: "Over 100 Million People in U.S. Now Receiving Some Form Of Federal Welfare."



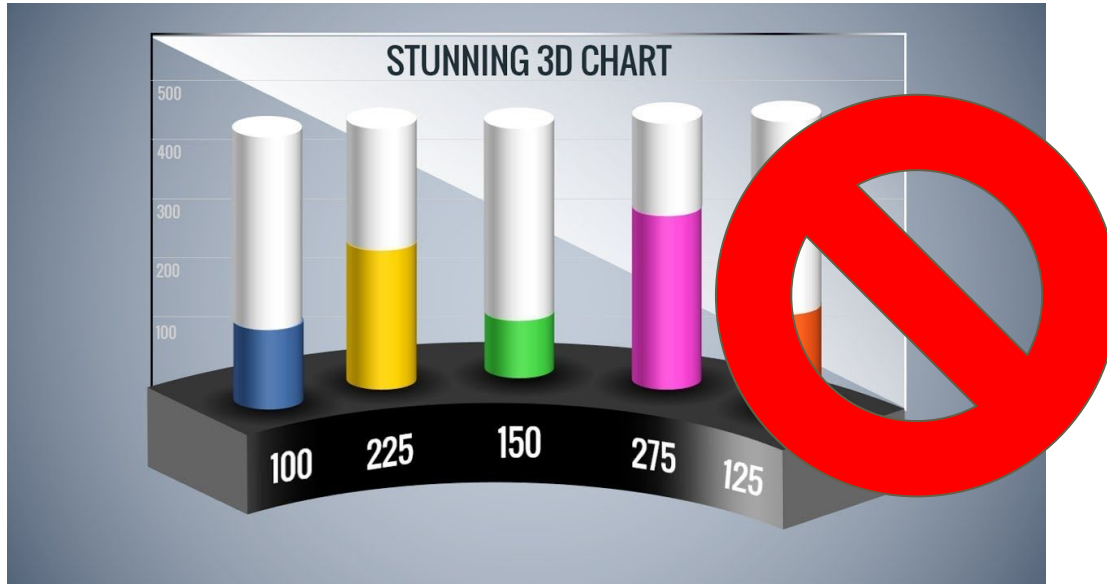
lie factor: 16,08

Principios de Diseño

Primero el fondo y luego la forma

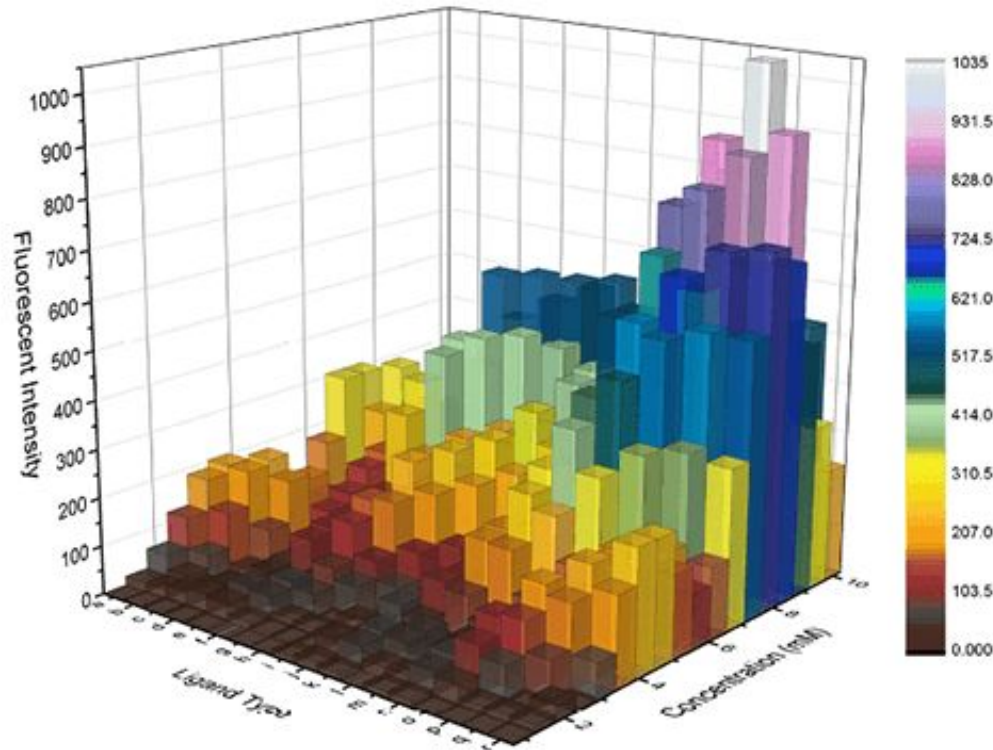
Enfocarse en la efectividad aunque sea tosco, luego se puede refinar.

Partir con los principios de efectividad y expresividad.



Principios de Diseño

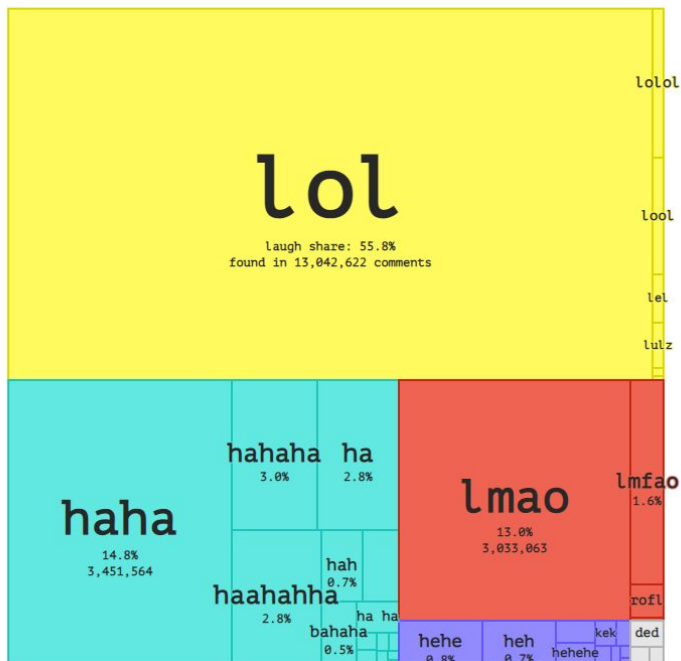
No al 3D injustificado



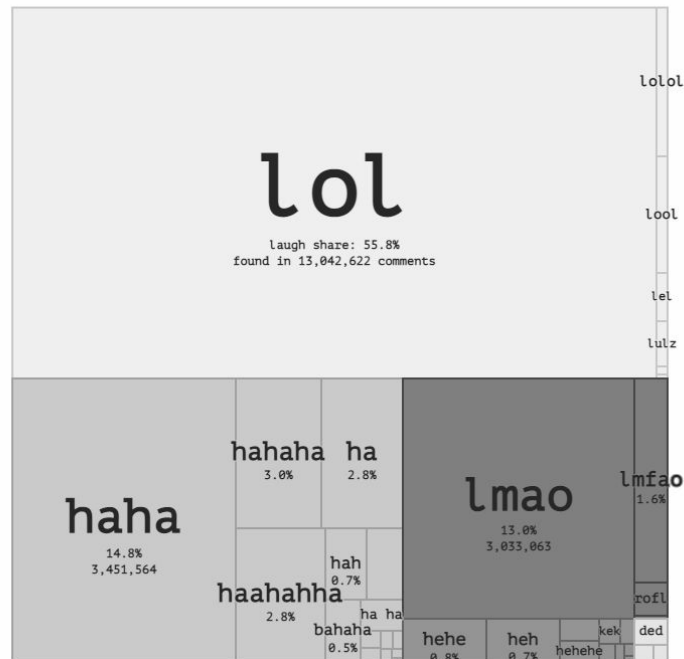
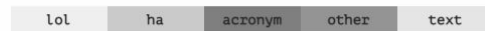
Principios de Diseño

Lograrlo en blanco y negro (*Get it right in black and white*)

2019 usage share of laughs by category on Reddit

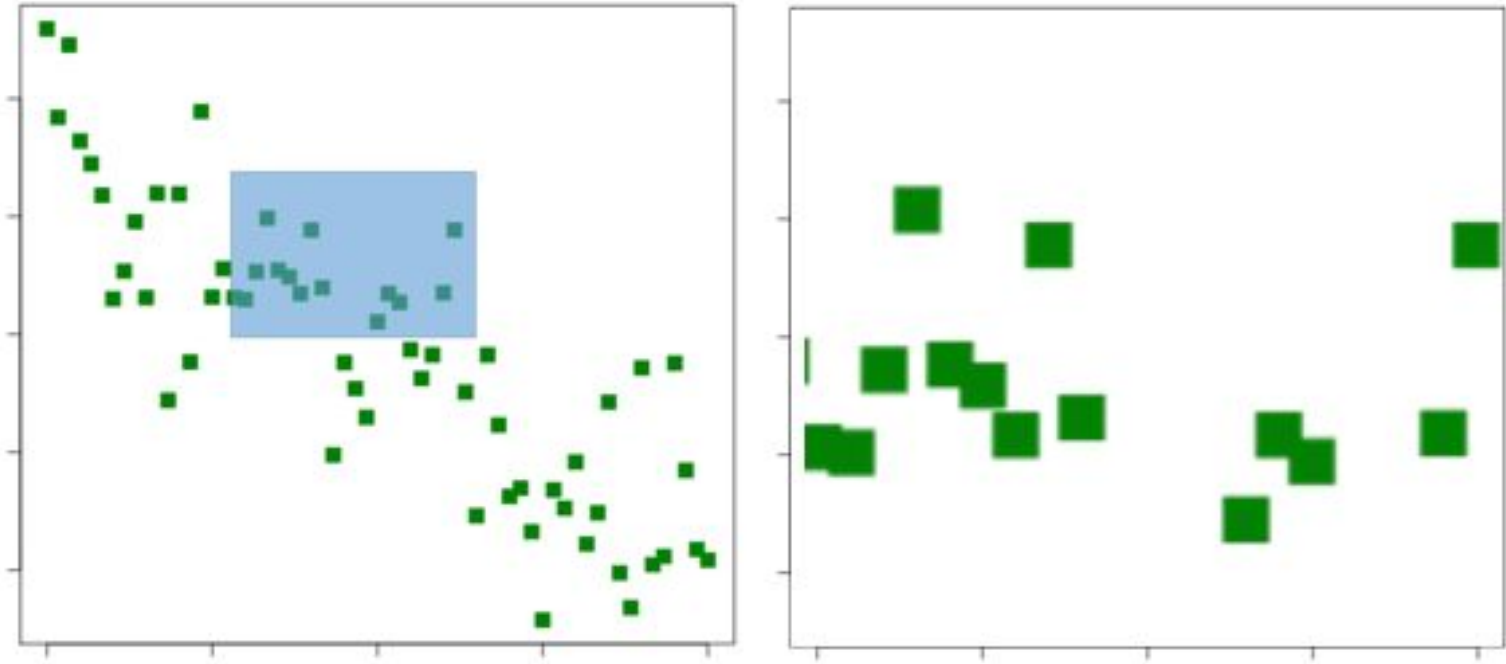


2019 usage share of laughs by category on Reddit



Principios de Diseño

“Overview first, details on demand” (shneiderman's mantra)



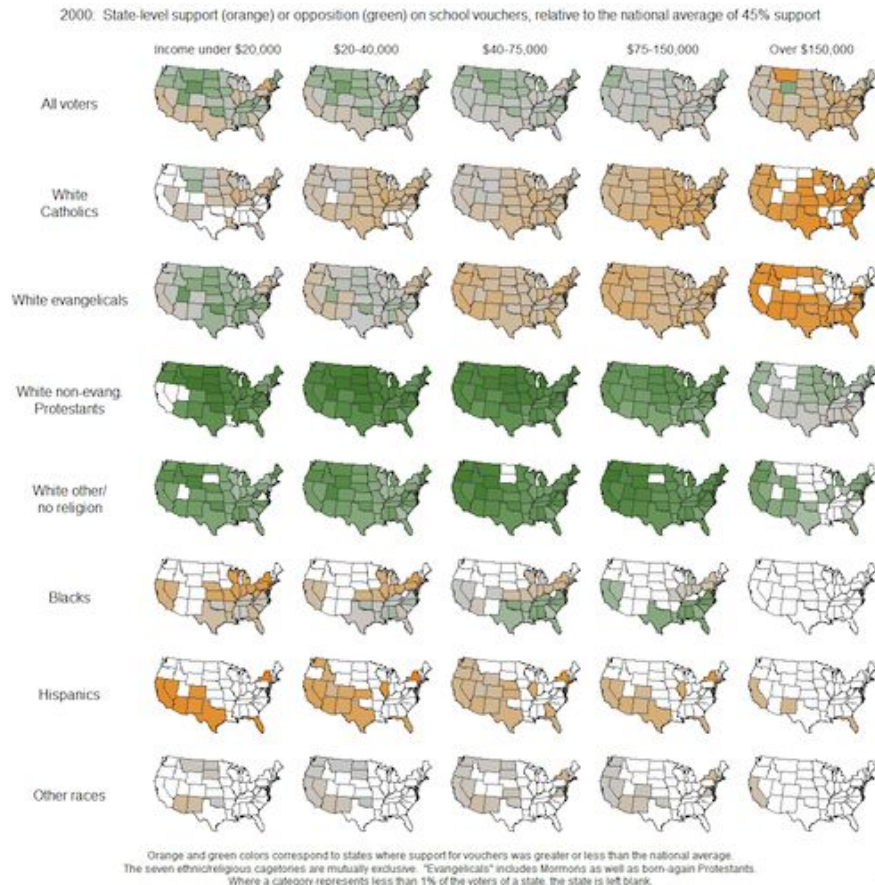
Principios de Diseño

Los ojos le ganan a la memoria (*eyes beat memory*)

Es más fácil usar la **cognición externa** que nuestra memoria interna

Por lo tanto, es más fácil comparar, moviendo nuestros ojos de lado a lado, que hacerlo tratando de recordar algo que vimos recientemente

- Ejemplo: un gráfico con diferentes instancias, variando el color según el experimento.



Principios de Diseño

Ser responsivo

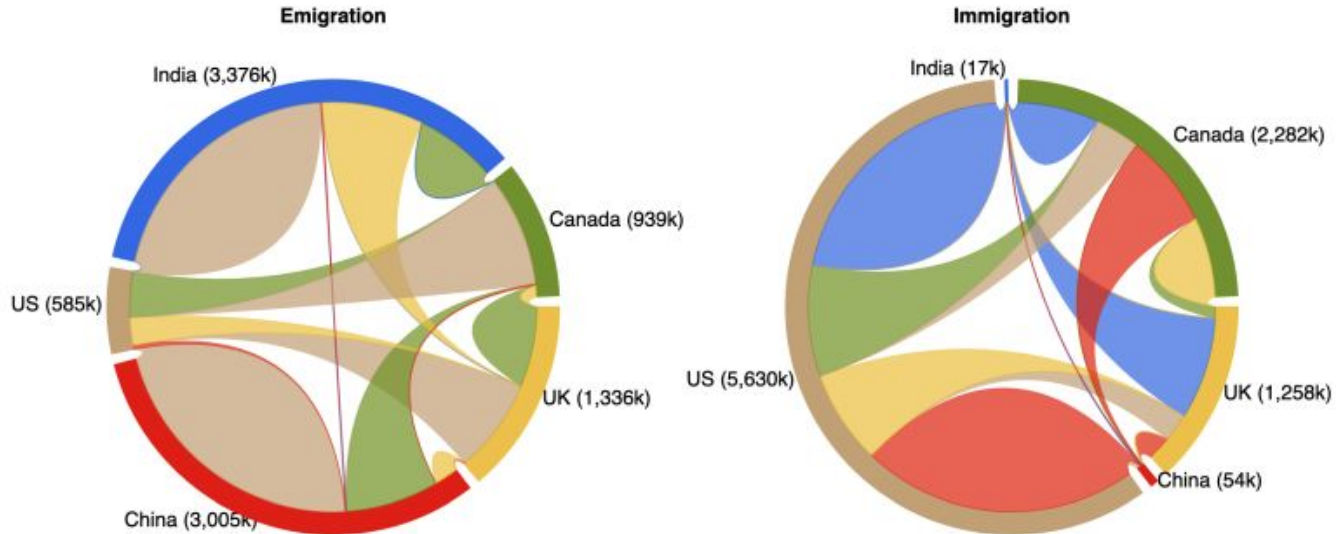
Dar señal que la visualización está procesando una acción.



Principios de Diseño

Consistencia - interna

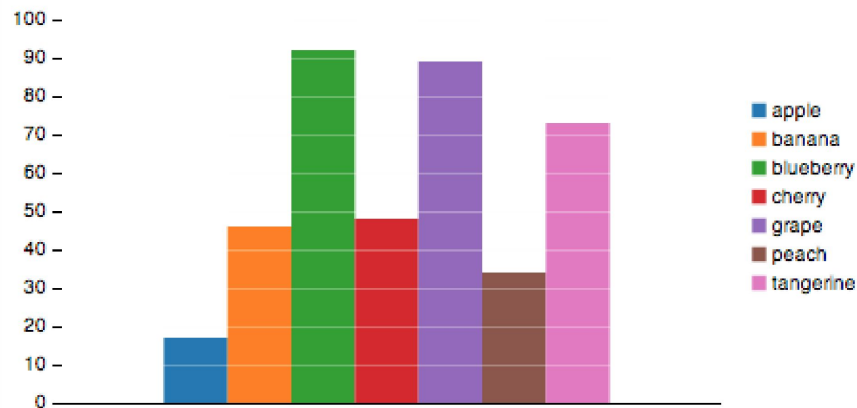
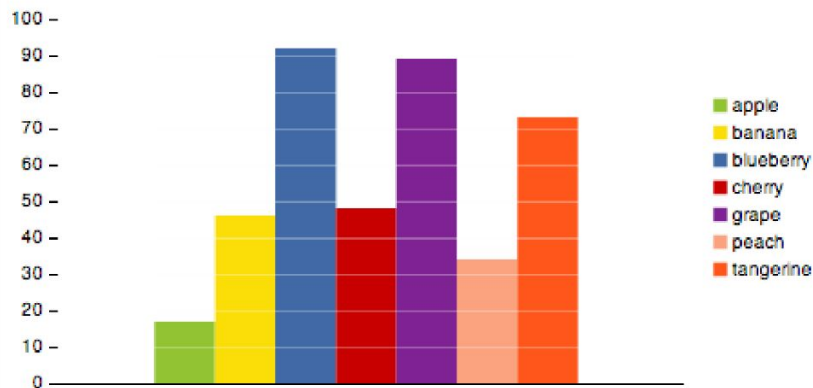
Mantén las decisiones de diseño realizadas durante todo el documento.



Principios de Diseño

Consistencia - externa

Tomar elecciones inspiradas en referentes externos al documento o visualización, y que potencialmente permiten una mejor comprensión gracias al conocimiento previo.



Principios de Diseño

Autocontención

El contenido de una visualización, en conjunto a su contenedor, debe ser **autoexplicativo**.

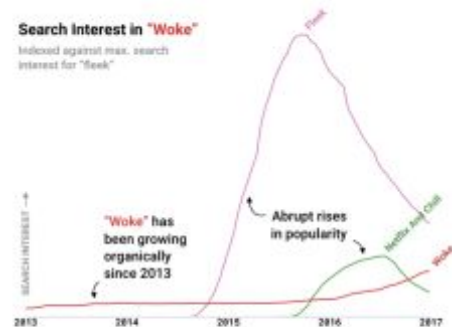
- Buen uso de leyendas.
- Buen uso de ejes.
- Buen uso de títulos en una visualización.
- Agregar contexto a la visualización (texto introductorio previo y posterior).
 - También hay una *rules of thumbs* llamada "Entender la narrativa" que pide incluir un contexto de la visualización.

Catalyst 2: Politics

In the US, last year's political circus tipped five of 2016's top slang into the mainstream.

Woke (noun) - aware of (racial) social injustice

"Woke" is the one term on 2016's list that's been growing organically for at least a decade. There was no major event that propelled it into the spotlight – it just finally crossed the threshold of mainstream use.



There were some unique events in 2016 for "woke." In terms of semantics, there's the notion that 2016's US election cycle whitewashed the term. Today, I see it used broadly as "political awareness," beyond its historic racial connotations. Childish Gambino's "Redbone" helped too (~50M streams on Spotify), which dropped in November 2016. The song's hook repeats the phrase "stay woke," helping to normalize the phrase.

Taller final

Taller final

Se compone de 2 partes:

- **Parte 1:** Criticar 2 visualizaciones en función de los principios de diseño.
- **Parte 2:** Programar visualizaciones adecuadas dada un contexto.
 - Deberán analizar los datos que tienen para cada caso.
 - Decidir la visualización a confeccionar en función de los datos, el contexto y la tarea visual presentada dentro del contexto.
 - Deberán programar la visualización usando *pandas*, *altair*, *matplotlib* y/o *seaborn*.
Disclaimer: Todos los casos tienen una solución aplicando alguna visualización vistas en el curso.

No olvidar: la nota final del curso es el promedio de los **4 mejores talleres**.



MIA

Magister en
Inteligencia Artificial

Evaluación Docente Interna 6to Bimestre 2024

“Visualización de Información y Analítica Visual”, INF3842, profesor Hernán Valdivieso.

Por favor, contesten la evaluación docente interna del curso escaneando el Qr:



Este formulario estará abierto a respuestas desde:

El jueves 25 de julio a las 20:30hrs hasta el viernes 9 de agosto a las 13hrs.



Visualización de Información y Analítica Visual

— Hernán Valdivieso López (hfvaldivieso@uc.cl) —
