IIC2026 Visualización de Información

Hernán F. Valdivieso López (2023 - 2 / Clase 12)

Temas de la clase - Framework, marcas y canales

- 1. Introducción al modelo Anidado de Tamara Munzner
- 2. ¿Qué es un *Idiom*?
- 3. Marcas y canales
 - a. Tipos de marcas
 - b. Tipos de canales
 - c. Principio de efectividad y expresividad

Clase Grabada

Esta clase fue realizada el 21 de septiembre. Cómo médica experimental para sobrellevar la semana previa a la semana de receso, se tomó la decisión de hacerla <u>híbrida</u> (presencial y por *zoom* simultáneamente).

En la siguiente playlist estará la grabación de dicha clase

Link a Playlist

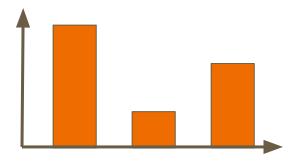
Se definen 3 preguntas fundamentales:

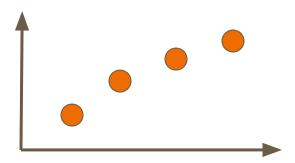
- 1. What? (¿Qué?)
 - a. ¿Qué datos vamos a visualizar?
 - b. ¿Cómo se caracteriza el *dataset* y los datos?
- 2. Why? (¿Por qué?)
 - a. ¿Por qué se hará la visualización?
 - b. ¿Cuáles acciones o tareas visuales necesita realizar el usuario?
- 3. How? (¿Cómo?)
 - a. ¿Cómo será nuestra visualización?
 - b. ¿Qué codificaciones o idioms utilizaremos?

Estas preguntas se resumen en:

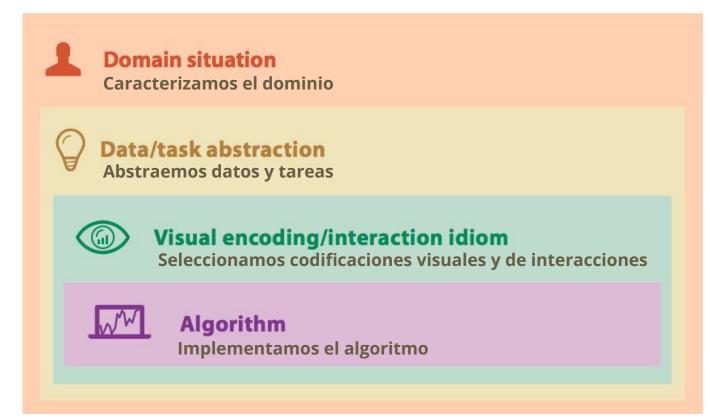
(dato, tarea visual, codificación)

También permiten analizar una visualización





- (llave + valor cuantitativo, comparar valor, alto de barras)
 - + (2 atributos numéricos, buscar correlación, posición de puntos)



Modelo anidado o en cascada

Proceso para diseñar y validar visualizaciones

Es anidado/cascada porque los resultados de una etapa son utilizados en la etapa siguiente.



Domain situationCaracterizamos el dominio



Data/task abstraction Abstraemos datos y tareas



Visual encoding/interaction idiomSeleccionamos codificaciones visuales y de interacciones



AlgorithmImplementamos el algoritmo

Entender el campo de conocimiento y contexto de aplicación

Entender datos, usuarios y tareas.

Entrevistas, observaciones o investigación mediante literatura



Domain situationCaracterizamos el dominio



Data/task abstraction
Abstraemos datos y tareas



Visual encoding/interaction idiomSeleccionamos codificaciones visuales y de interacciones

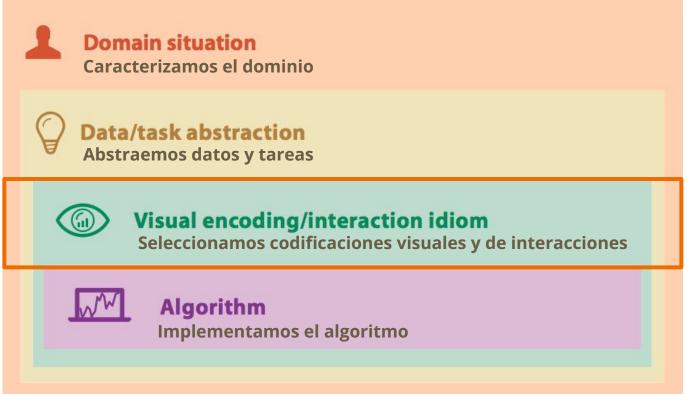


AlgorithmImplementamos el algoritmo

Traer las particularidades del contexto al mismo lenguaje.

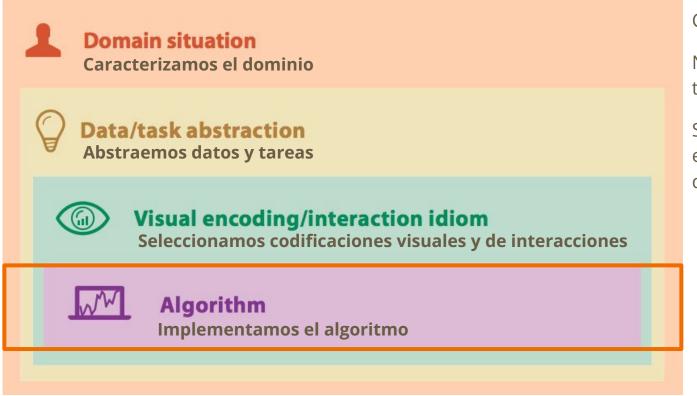
Definir el qué y por qué de la herramienta.

Seleccionar los aspectos a consideración



Definir el cómo.

Hay decisiones visuales y decisiones de interacción que debemos realizar



Código.

Nivel técnico tecnológico.

Se revisa eficiencia, escalabilidad y dificultad

¿Qué es un Idiom?

¿Qué es un idiom?

- Enfoque distintivo de codificar información o manipularla.
- *Tldr*: Son las decisiones tomadas para transformar los datos en una visualización.

Idioms visuales

- Decisiones elementales, cómo escoger si utilizar un círculo o una línea. O escoger qué colores utilizar en la visualización
- Decisiones complejas, como utilizar un gráfico de torta para representar las proporciones de datos agregados.
- También hay decisiones en esta familia de cómo posicionar y organizar los distintos elementos en una visualización.

Idioms de interacción

- Animar un gráfico cuando se carga.
- Agregar la capacidad de seleccionar elementos en la visualización.
- Reducir la cantidad de información mostrada en pantalla.
- Poder navegar por la visualización (por ejemplo, hacer zoom).

¿Qué es un idiom?

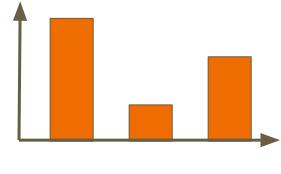


Datos

 Representar cada dato por una barra.

 Usar largo de barra para mostrar el valor

Idiom



= Visualización



burbujas.
No incluiré el color en la visualización.

Hacer un gráfico de



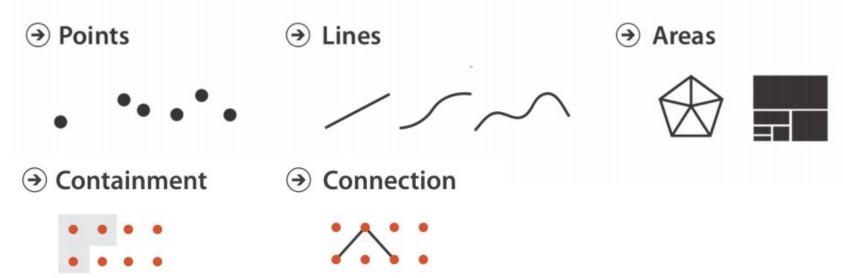
Datos

+

Idiom

Marcas

 Elemento geométrico básico, que puede ser clasificado según el número de dimensiones espaciales (largo/ancho/profundidad) necesarias para poder dibujar la marca.



Marcas cero-dimensional

- Punto
- No presenta variación de ancho, largo o profundidad (son constantes)

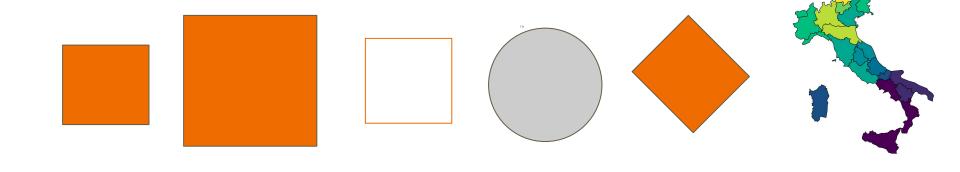
Marcas uni-dimensional

- Línea
- Tenemos una dimensión: largo
- No presenta variación de ancho ni profundidad (es constante)



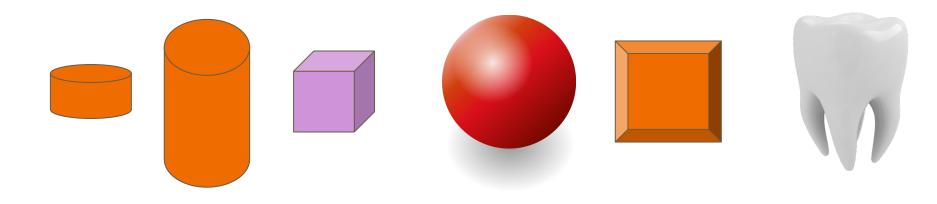
Marcas bi-dimensional

- Áreas
- Tenemos dos dimensiones: largo y ancho.
- No presenta profundidad (es constante).



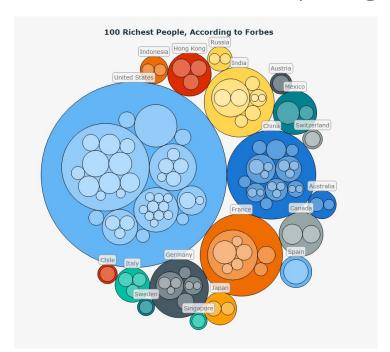
Marcas tri-dimensional

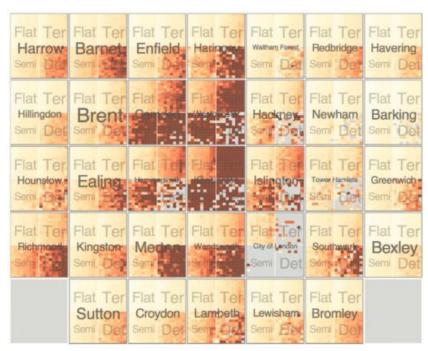
- Volumen.
- Tenemos tres dimensiones: largo, ancho y profundidad.



Marcas de contención

Marca de áreas utilizada para agrupar elementos dentro de ella.

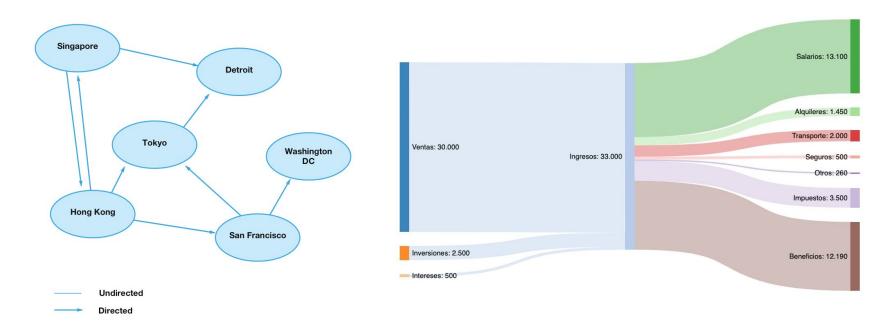




Fuente: Creating Circle Packing Chart with JavaScript y Visualization Analysis & Design, Tamara Munzner

Marcas de conexión

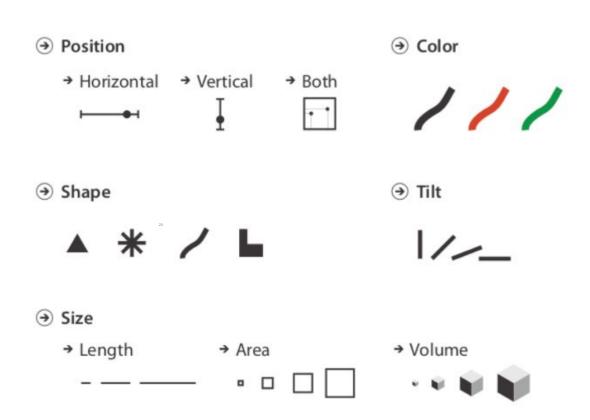
Marca de línea o áreas utilizada para relacionar 2 elementos.



Fuente: Swift Algorithm Club: Graphs with Adjacency List y Un generador de diagramas de Sankey

Canal

 Permite controlar la apariencia de las marcas, independientemente de la dimensionalidad de este elemento primitivo



Marcas y canales - Ejemplo

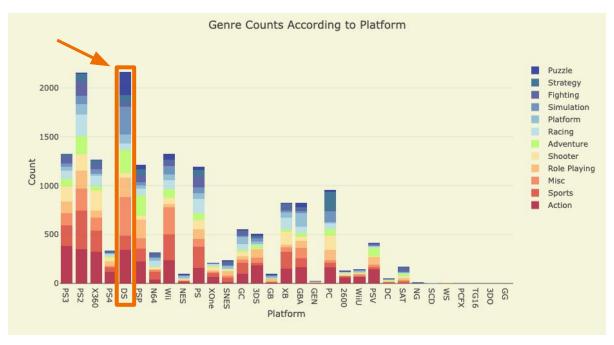
Nuestra información la vamos a representar por un área circular (marca).



Podemos personalizar (canales) estos puntos con:

	Tamaño Borde	

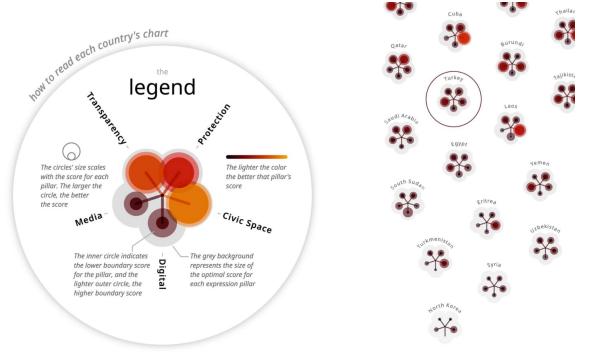
Glifos → Objeto gráfico **complejo** que representa algún dato, y con una estructura interna **formada por múltiples marcas** primitivas que también representan otra información.



Fuente: Video Game Sales Exploratory Data Analysis | Kaggle

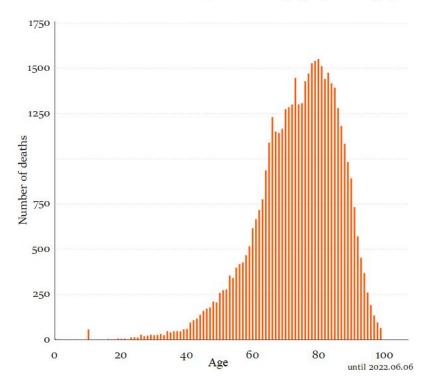
Glifos → Objeto gráfico **complejo** que representa algún dato, y con una estructura interna **formada por múltiples marcas** primitivas que también representan otra

información.



Fuente: The Freedom of Expression

Number of deaths from Covid-19 in different age groups in Hungary



Marca

 Líneas/barras para representar el dato (cantidad de fallecidos según edad)

Canal

- Largo para codificar el valor.
- Posición horizontal para codificar la edad asociada a dicho valor.

Fuente: Publicación en reddit (Data is Beautiful)



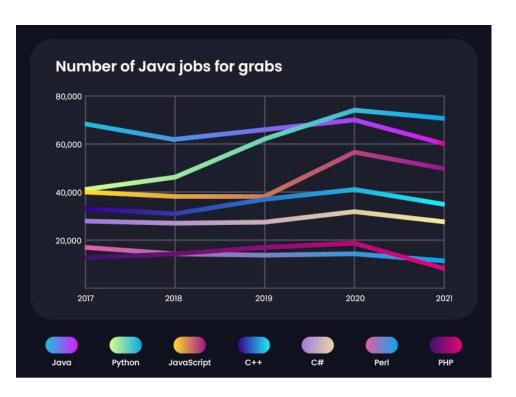
Marca

• **Imágen** (área con forma y color específico) para representar cada conjunto de emoji.

Canal

 Tamaño para codificar la frecuencia de cada emoji

Fuente: Wordcloud with emojis



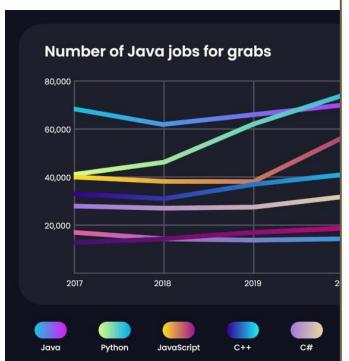
Marca

 Puntos para mostrar el valor y líneas para enfatizar la relación entre cada valor.

Canal

- Color en degradé para separar cada lenguaje.
- Posición horizontal para codificar cada año.
- Posición vertical para codificar el valor en dicho año.

Fuente: Publicación en reddit (Data is Ugly)





Short answer: Point marks

Longer answer: point marks to show values, with connecting line marks emphasizing the relationship between successive items

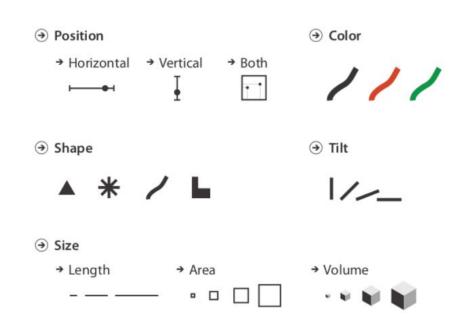
Fuente: Publicación en reddit (Data is Ugly)

Tipos de canales

- El sistema de percepción humano tiene dos tipos de modalidades:
 - El identity channel permite discernir información sobre qué es algo o dónde se encuentra,
 - El magnitude channel, por otra parte, nos permite saber cuánto de ese algo existe.
- Con estas dos modalidades, podemos saber, por ejemplo:
 - ¿Qué figura es? ¿Un círculo, un triángulo, una cruz o un heptágono? [identity]
 - ¿De qué hue/matiz es? ¿Rojo, verde, caqui o gris? [identity]
 - ¿Cuánta saturación tiene ese azul? ¿Celeste, azul marino o turquesa? [magnitude]
 - ¿Dónde está? ¿En qué región se encuentra la marca? [identity]
 - ¿Qué tan larga es aquella línea con respecto a esta? [magnitude]
 - ¿Cuánto espacio hay entre ambos rectángulos? [magnitude]

¿Cómo usarlos?

- No todos los canales son iguales: los mismos datos codificados con dos canales visuales distintos resultará en información diferente.
- Dos principios guían el uso de canales visuales: expresividad y efectividad



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 visualmente se entiendan 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

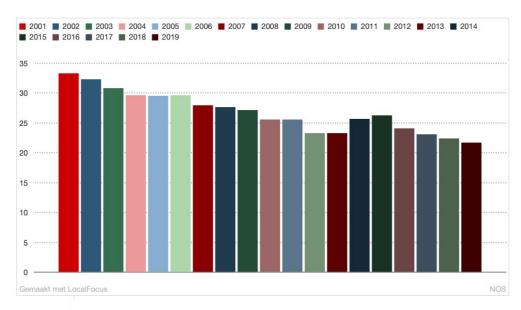
^{*}La semántica del atributos los veremos con más profundidad la próxima clase

Ejemplos reales



El uso de área no el más efectivo en esta situación

No cumple el **principio de efectividad**



Porcentaje fumadores en Países Bajos

Los años presentan un orden, pero la matiz de color no. ¿Por qué el verde va después que el azul y el rojo?

No cumple el **principio de expresividad**

Fuente: https://viz.wtf/

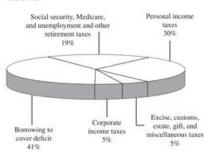
Ejemplos reales

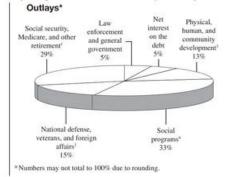
Major Categories of Federal Income and Outlays for Fiscal Year 2021

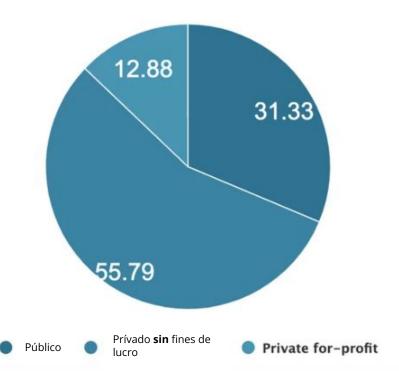
Income and Outlays. These pie charts show the relative sizes of the major categories of federal income and outlays for fiscal year 2021.

Income

Outlays*







El uso de volumen + perceptiva no el más efectivo para comparar los valores

No cumple el **principio de efectividad**

Se ocupa una gama de azules para 3 datos categóricos que no tienen una relación de orden.

No cumple el **principio de expresividad**

Fuente: r/dataisugly (izquierda) y r/dataisugly (derecha)

Antes de salir... Revisión de contenidos (RC)

Se publicó un mini control de alternativas en Canvas sobre lo que **vimos en la clase de hoy**.

- **Duración**: 2 semanas para realizarlo a partir de hoy.
- Intentos para responder: ilimitados.
- Extensión: 9 preguntas de 1 punto c/u.
- Condición para obtener el punto RC: Al menos 8 puntos de 9.
- Cada vez que respondan, verán el puntaje total logrado, pero no cuáles preguntas están correctas e incorrectas.

Próximos eventos

Próxima clase

Framework Tamara Munzner: Abstracción de datos.

Ayudantía de mañana

Sala de ayuda para la T2

Tarea 2

- Se entrega este domingo a las 20:00. Luego comienzan los 3 días de atraso, hasta el miércoles a las 20:00, para entregar.
- Se entrega este lunes a las 20:00. Luego comienzan los 2 días de atraso, hasta el miércoles a las 20:00, para entregar.

IIC2026 Visualización de Información

Hernán F. Valdivieso López (2023 - 2 / Clase 12)