

Visualización de Datos Efectiva con Altair

Nov 2022

@ChekosWH | @tacosdedatos

Sergio Sánchez

Hoy...



Nov 2022

Quien soy yo

Visualización de datos

Visualización de datos efectiva

Gramática de gráficos

Herramientas para esta clase

Tarea (y si hay tiempo, demo)

Quien soy yo



Mi nombre es **Sergio** Sánchez Zavala (también conocido como **Chekos**) y soy originario de **Tijuana**, Baja California - la frontera de **México** con **Estados Unidos**.

Me enfoco en crear contenido y herramientas para facilitar el trabajar datos.

La meta es que nuestros análisis sean más ***transparentes, reproducibles y accesibles***.

Trabajo como **ingeniero de datos** en una empresa sin fines de lucro (TalkingPoints) en la bahía de San Francisco, California.

Mi carrera profesional

2020

Data Visualization Analyst
Alluma

2021

Ingeniero de Datos
Alluma

2022

Ingeniero de Datos
Talking Points



2017

Python for Data Science
(UCSD)
Edx.org



2018

Research Associate
Public Policy Institute of
California

Mi carrera profesional

en el mundo de los datos

2020

Data Visualization Analyst
Alluma

2021

Ingeniero de Datos
Alluma

2022

Ingeniero de Datos
Talking Points



2017

Python for Data Science
(UCSD)
Edx.org



2018

Research Associate
Public Policy Institute of
California

Políticas Pùblicas
+
datos

Ciencias sociales y la ciencia de datos

Ingeniería de datos

Extraer, transformar / limpiar, guardar / cargar datos

Análisis de datos

Extraer información accionable de los datos
Presentar datos reportes, dashboards, gráficos.

Aprendizaje automático (*machine learning*)

Hacer predicciones y extraer patrones utilizando estadística: ANOVA, regresión linear, PCA

tacosdedatos



20k tiktok
9.4k twitter
1.1k instagram

Hay una comunidad enorme de personas de todos niveles en el mundo de datos dispuestos a apoyarte y compartir conocimiento

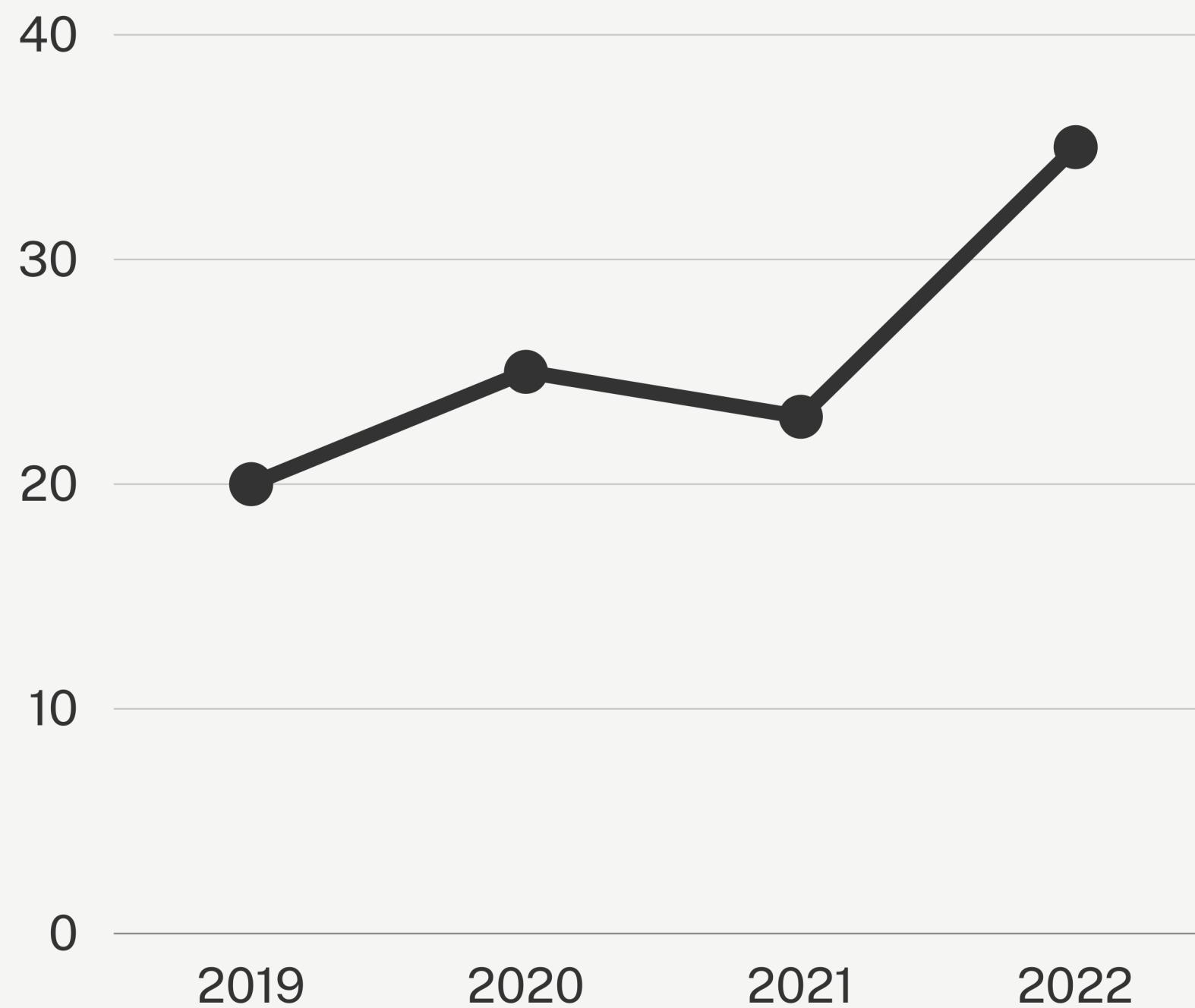
*¿Qué es la
visualización de
datos?*

¿qué es la visualización de datos?

Definición

La representación visual de información y datos

visualización de datos efectiva con altair

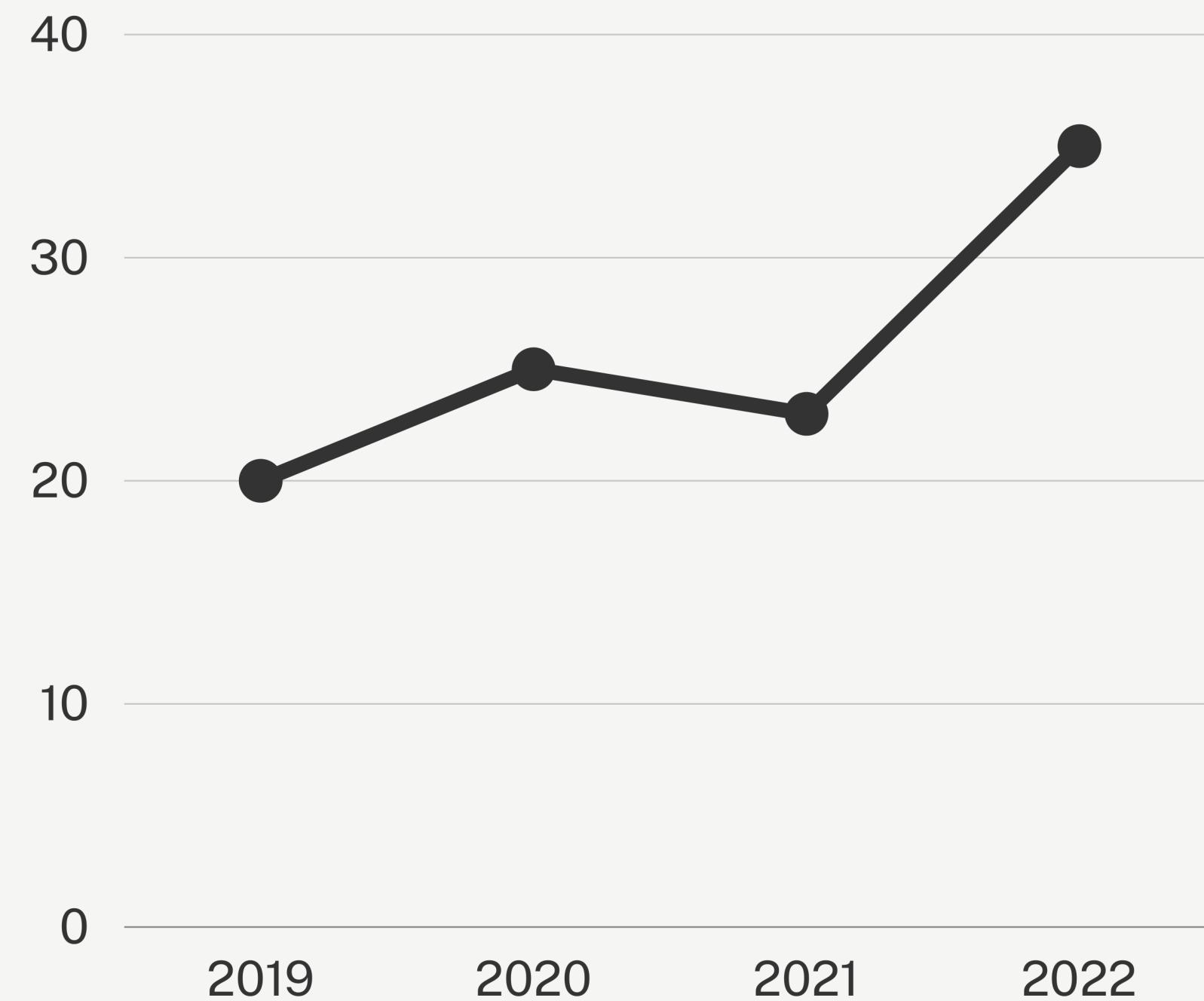


¿qué es la visualización de datos?

visualización de datos efectiva con altair

Definición

La **representación visual** de información y datos

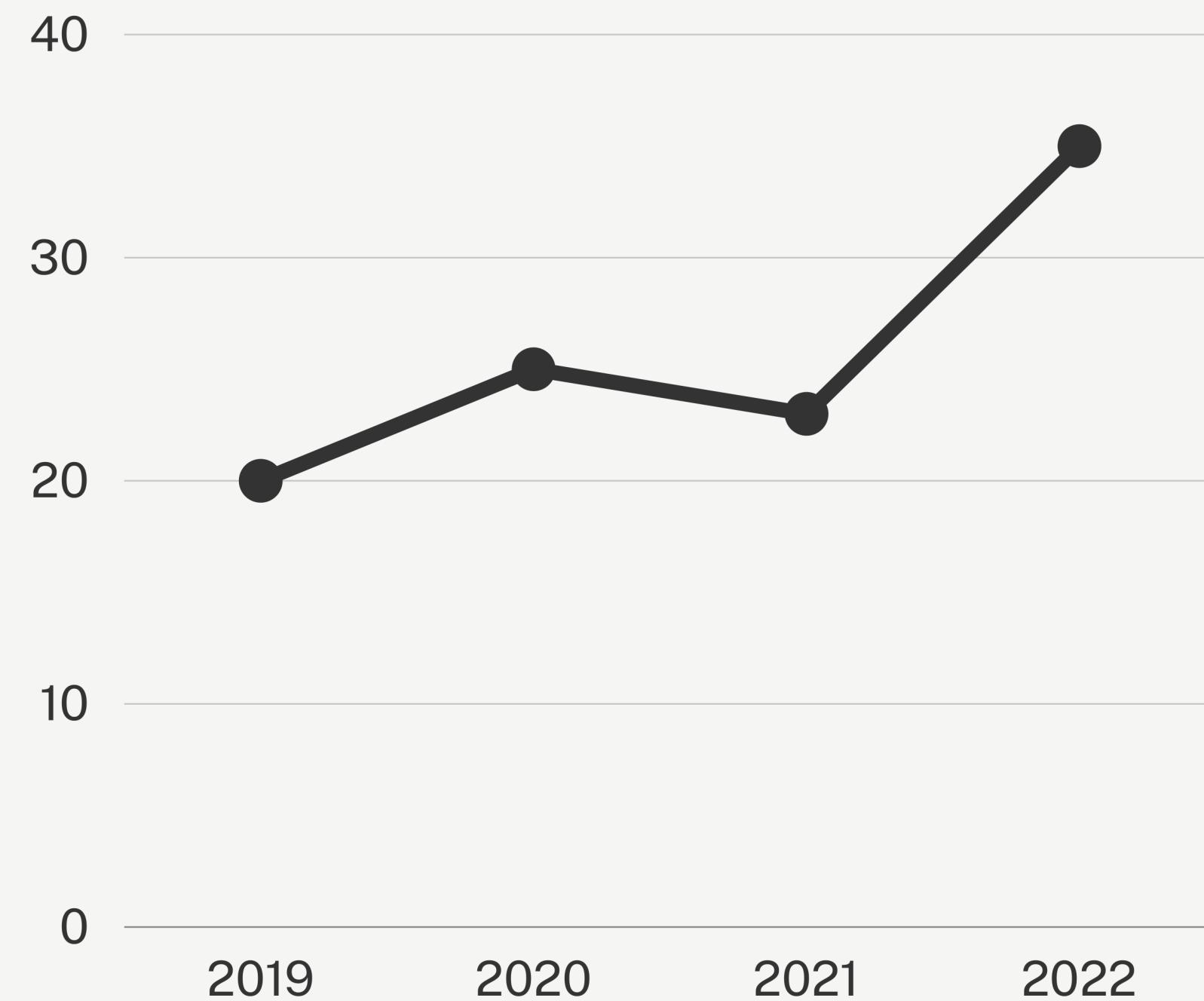


¿qué es la visualización de datos?

visualización de datos efectiva con altair

Definición

La representación visual de **información** y datos

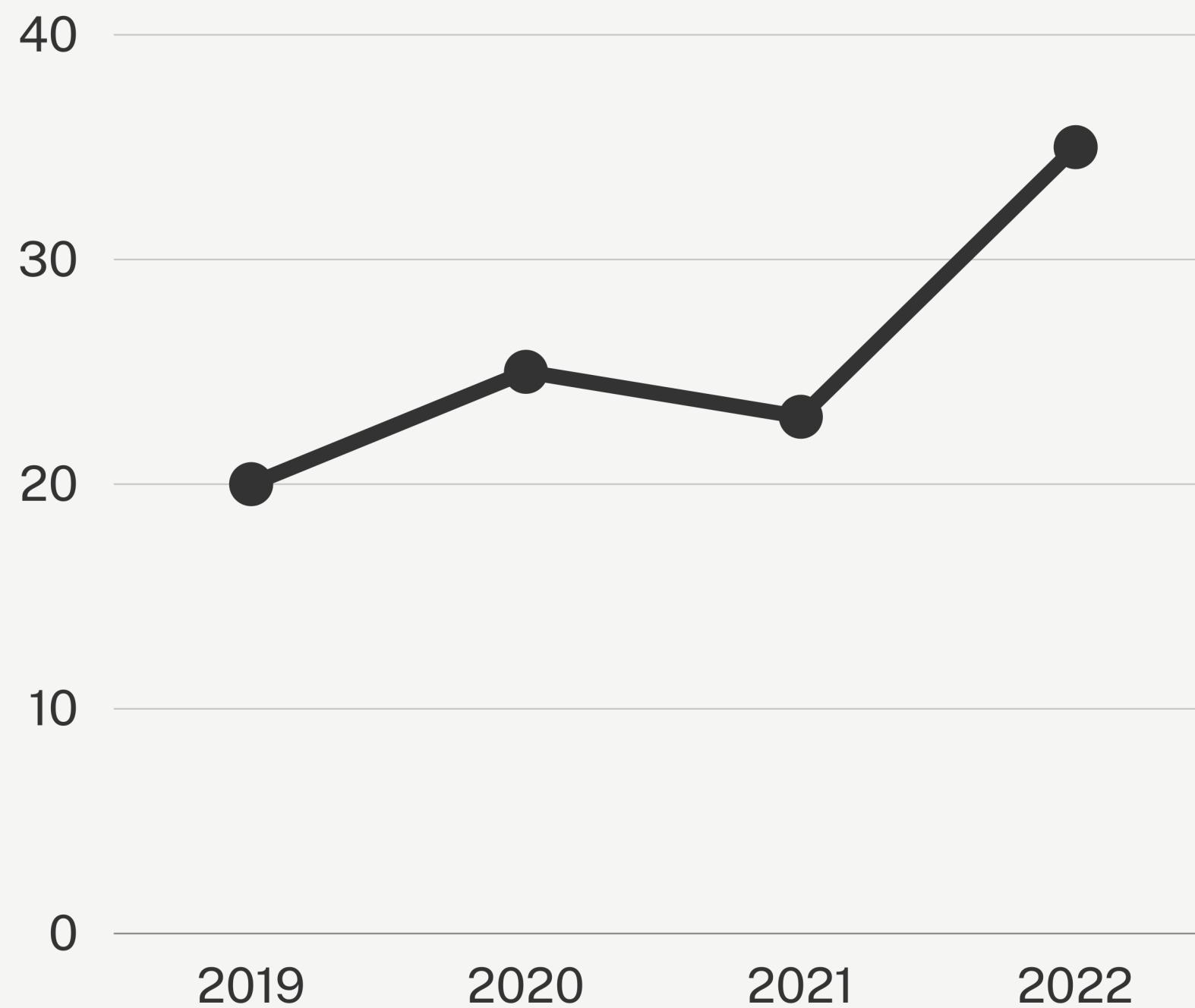


¿qué es la visualización de datos?

Definición

La representación visual de información y datos

visualización de datos efectiva con altair



Charles Minard

Carte Figurative des pertes successives en hommes de l'Armée Française dans la Campagne de Russie 1812-1813.

Dessinée par M. Minard, Inspecteur Général des Ponts et Chaussées en retraite.

Paris, le 20 Novembre 1869.

Les nombres d'hommes présents sont représentés par les largeurs des zones colorées à raison d'un millimètre pour dix mille hommes; ils sont de plus écrits en travers des zones. Le rouge désigne les hommes qui entrent en Russie, le noir ceux qui en sortent. — Les renseignements qui ont servi à dresser la carte ont été puisés dans les ouvrages de M. M. Chiers, de Ségur, de Fezensac, de Chambray et le journal inédit de Jacob, pharmacien de l'Armée depuis le 28 Octobre.

Pour mieux faire juger à l'œil la diminution de l'armée, j'ai supposé que les corps du Prince Jérôme et du Maréchal Davout, qui avaient été détachés sur Minsk et Mohilow et qui rejoignirent Orsha et Wilebsk, avaient toujours marché avec l'armée.

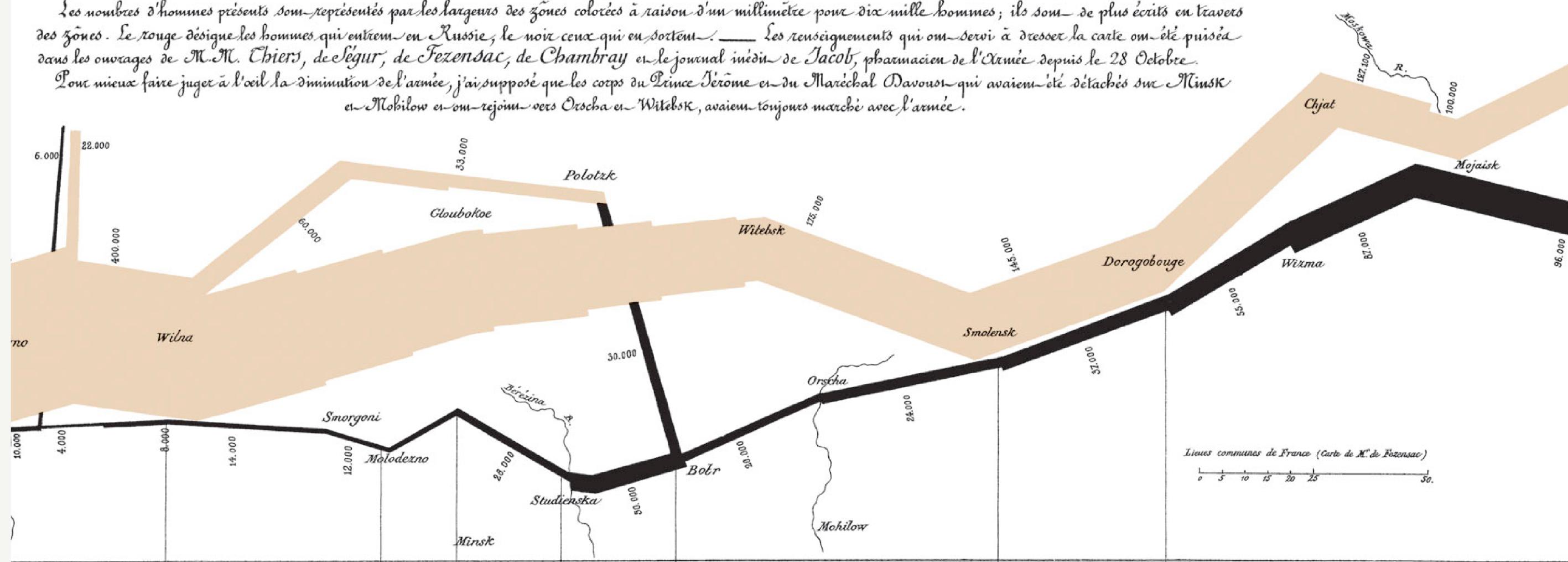
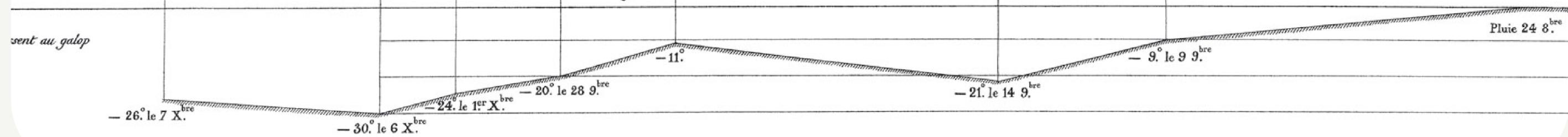
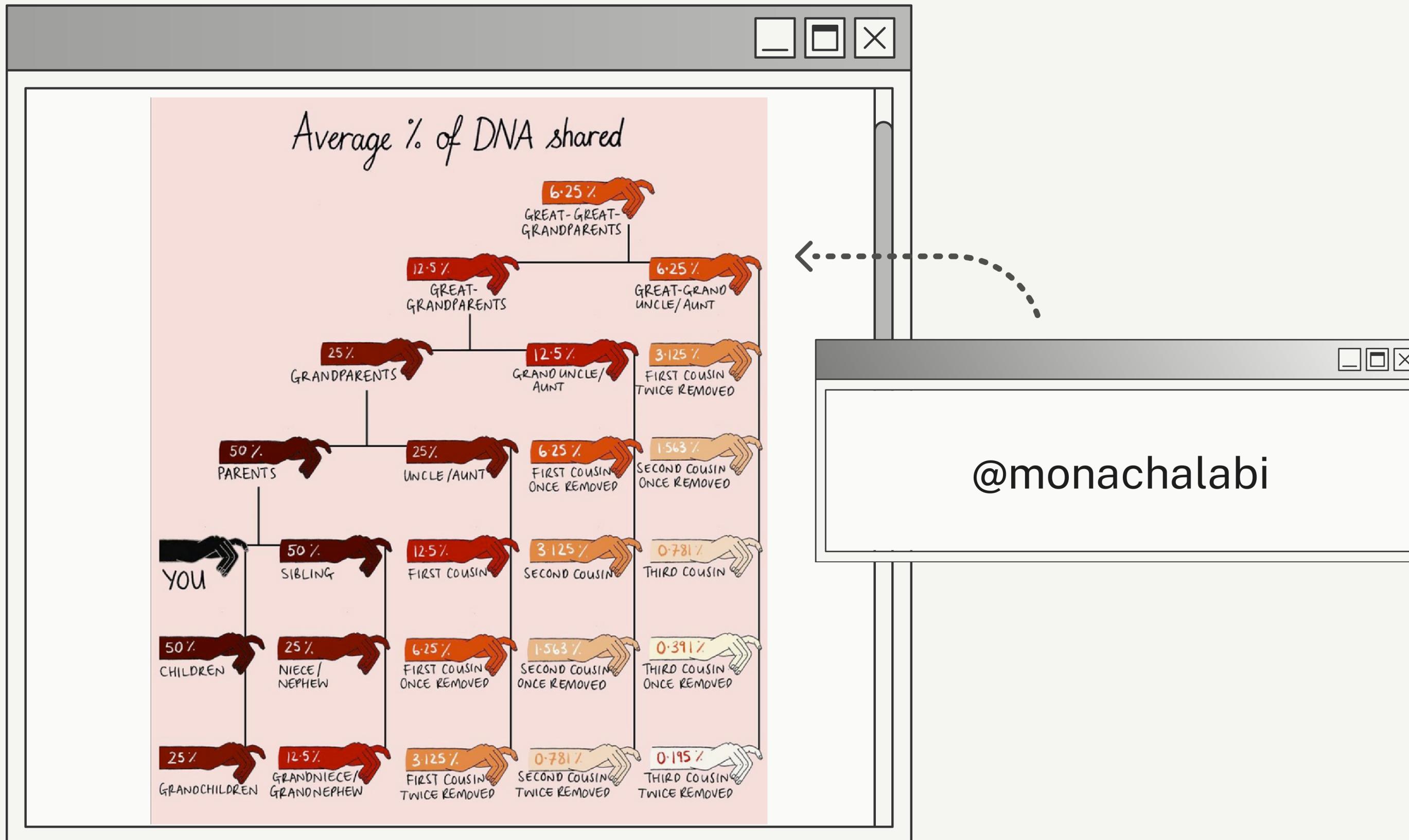


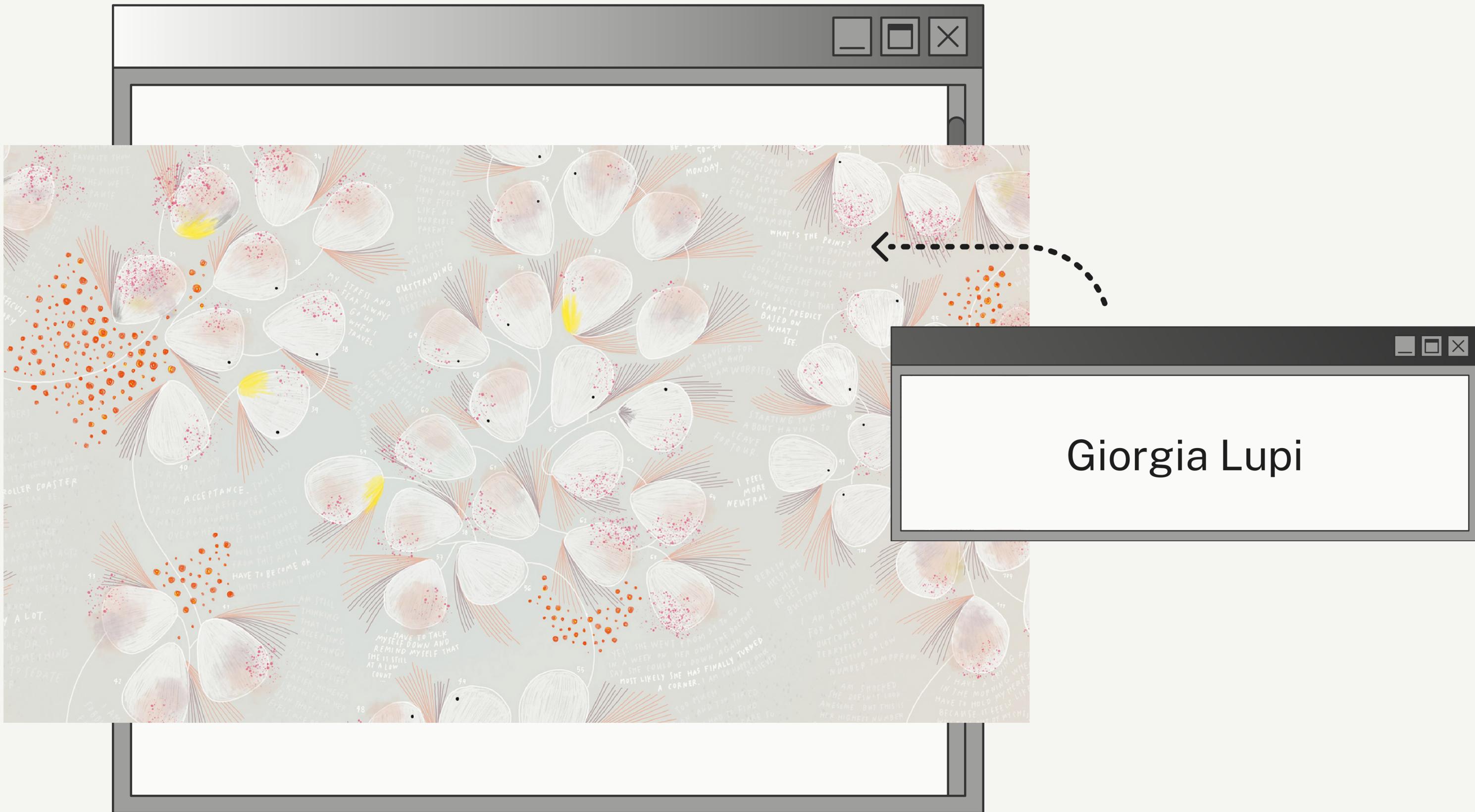
TABLEAU GRAPHIQUE de la température en degrés du thermomètre de Réaumur au dessous de zéro.



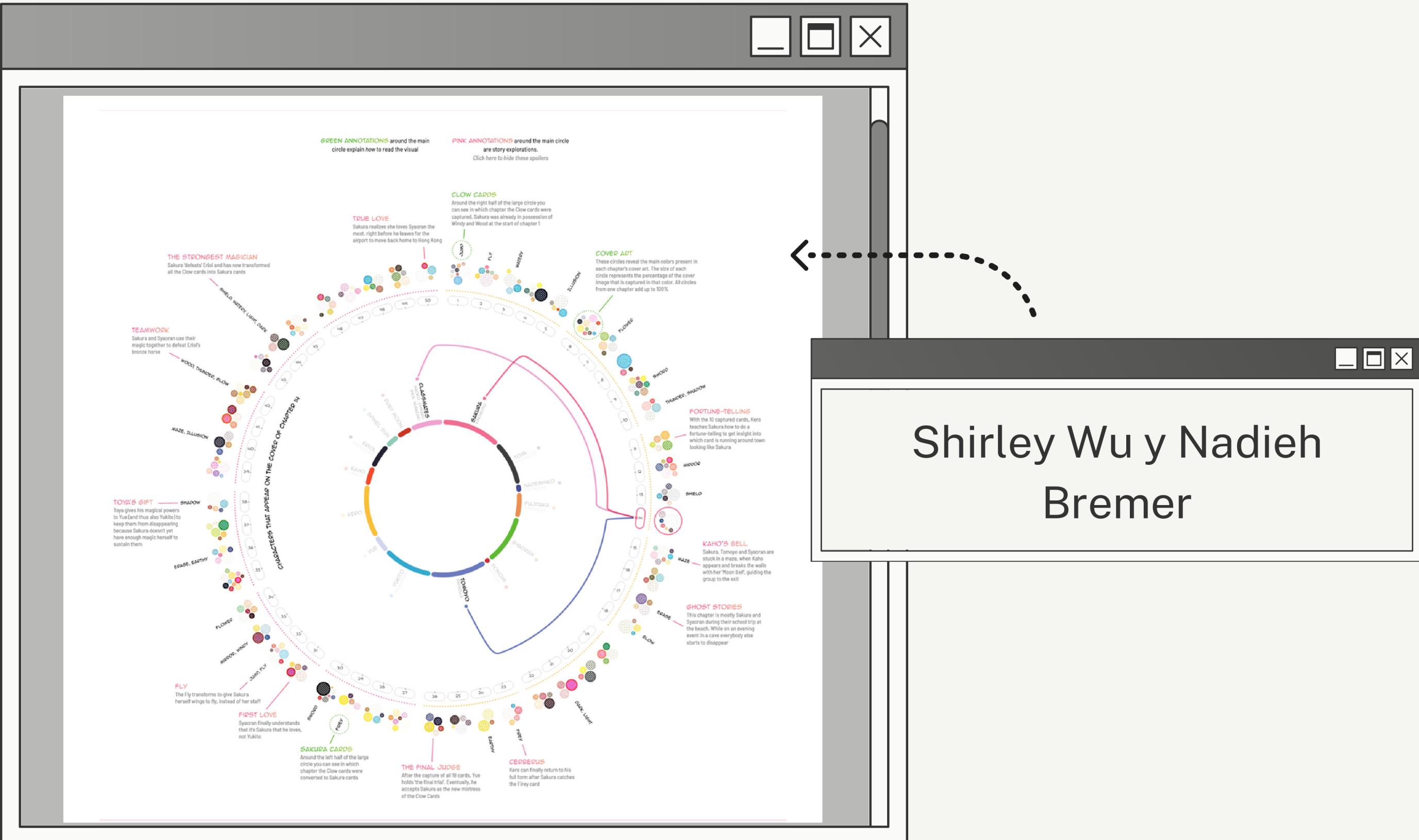
Ejemplos



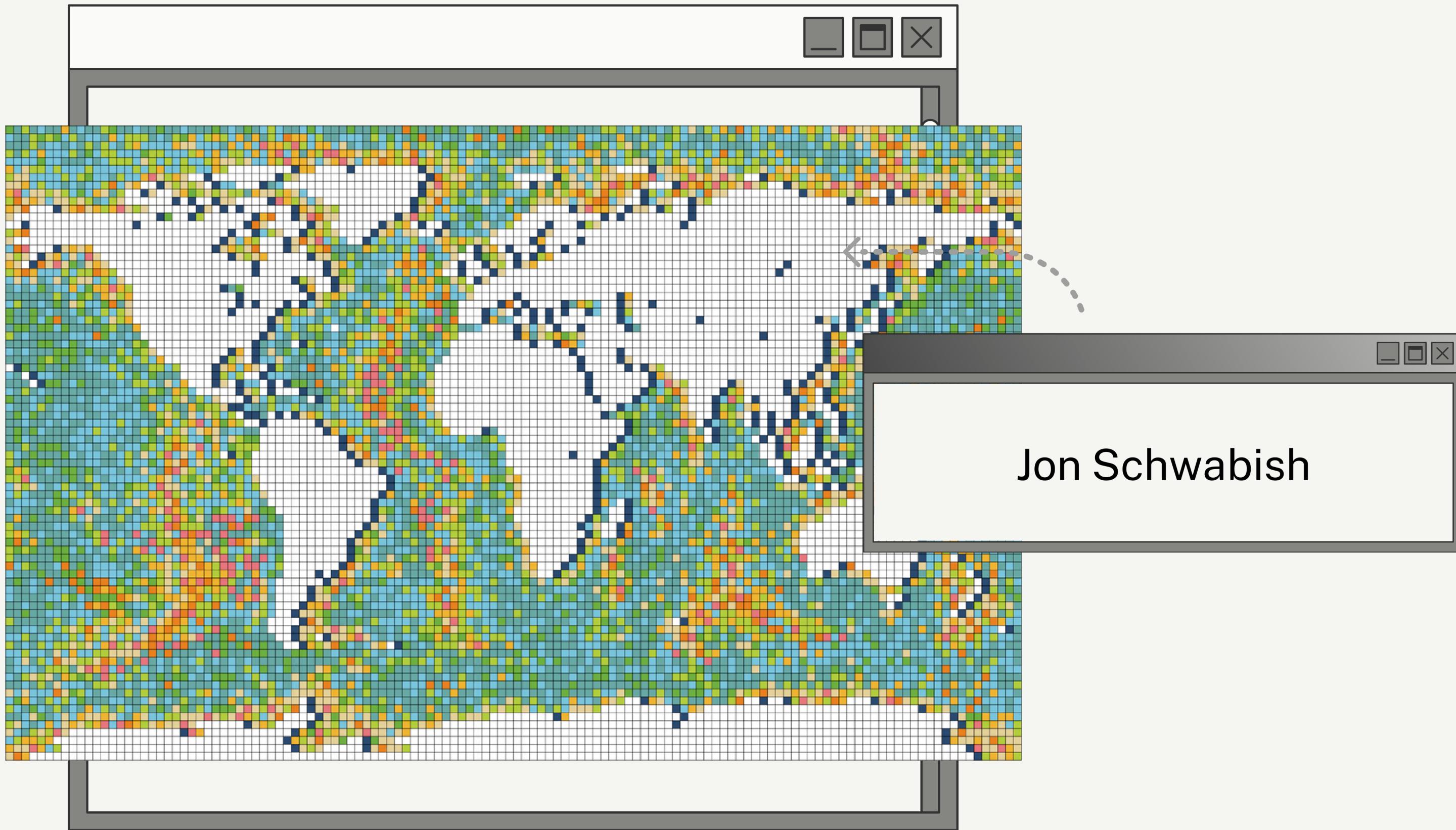
Ejemplos



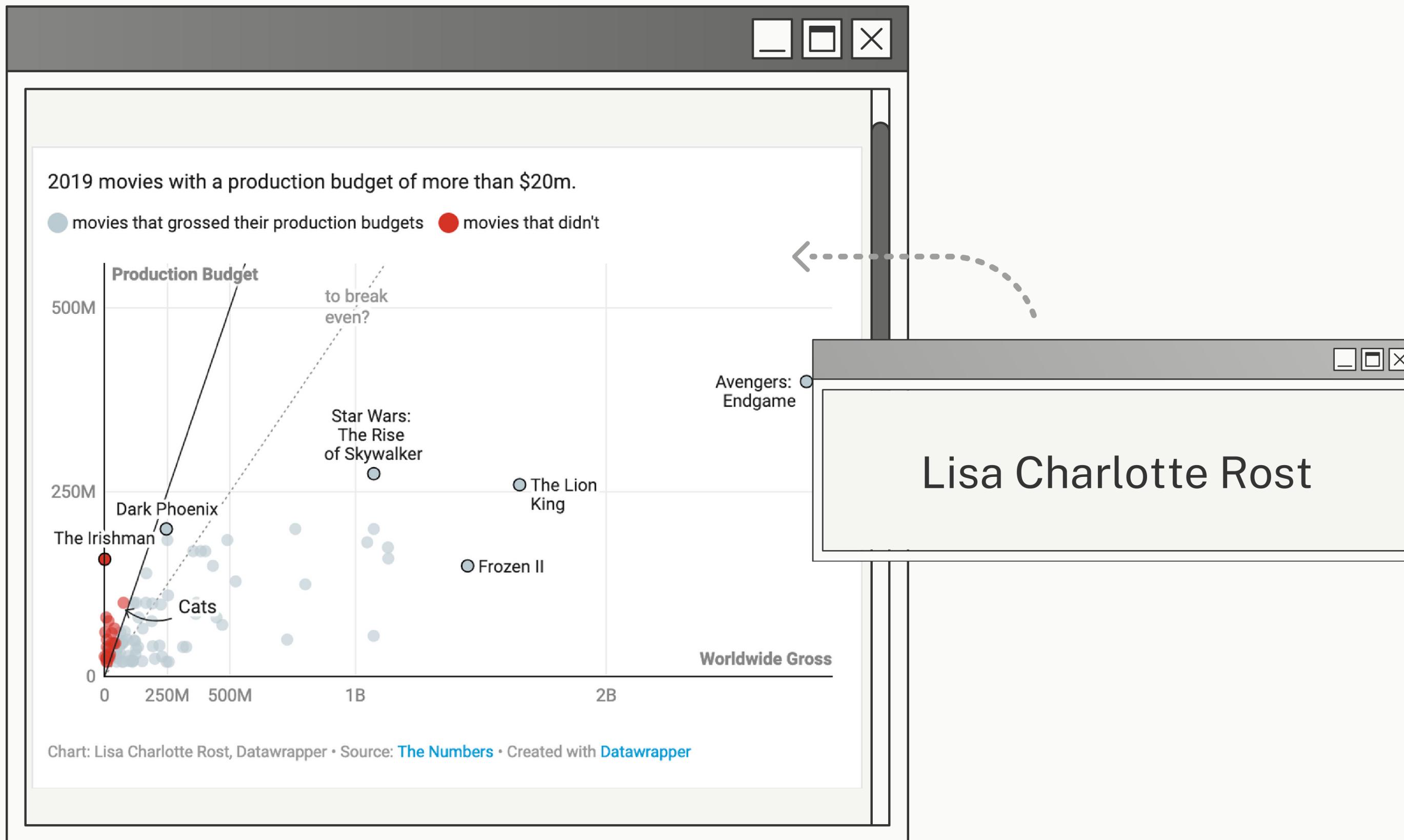
Ejemplos



Ejemplos

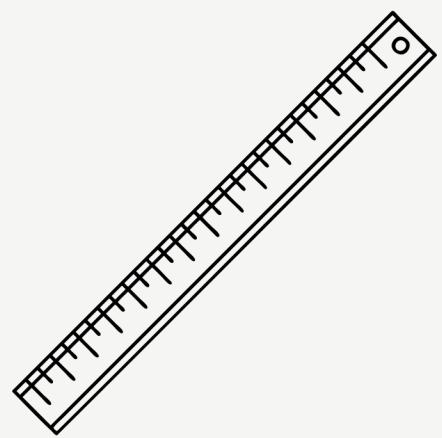


Ejemplos

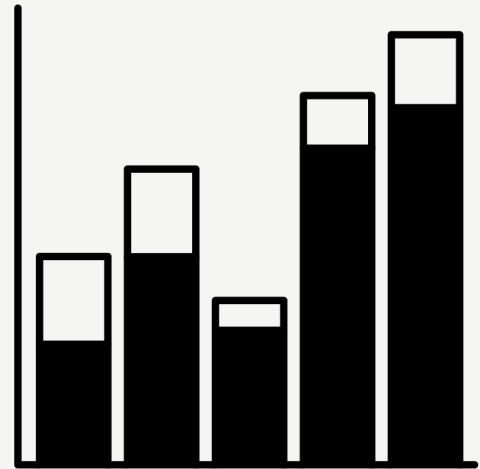
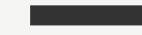


*¿Qué es la
visualización de datos
efectiva?*

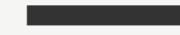
visualización de datos efectiva



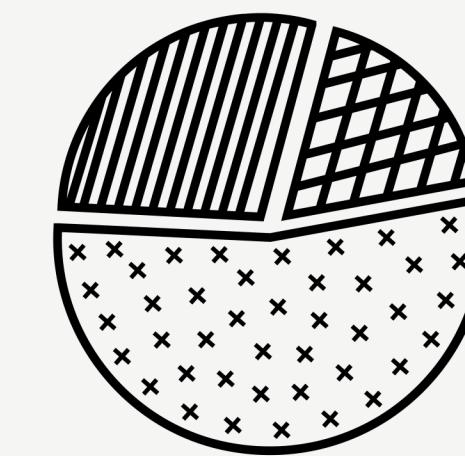
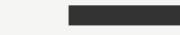
Honesta



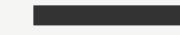
Clara



Entendible



Accesible



*¿Qué es la gramática
de gráficos?*

Una gramática de gráficos es una herramienta

que nos permite describir de forma concisa los componentes de un gráfico.

Esta gramática nos permite ir más allá de los gráficos con nombre (por ejemplo, el "diagrama de dispersión") y comprender mejor la estructura profunda que subyace a los gráficos estadísticos.

Hadley Wickham

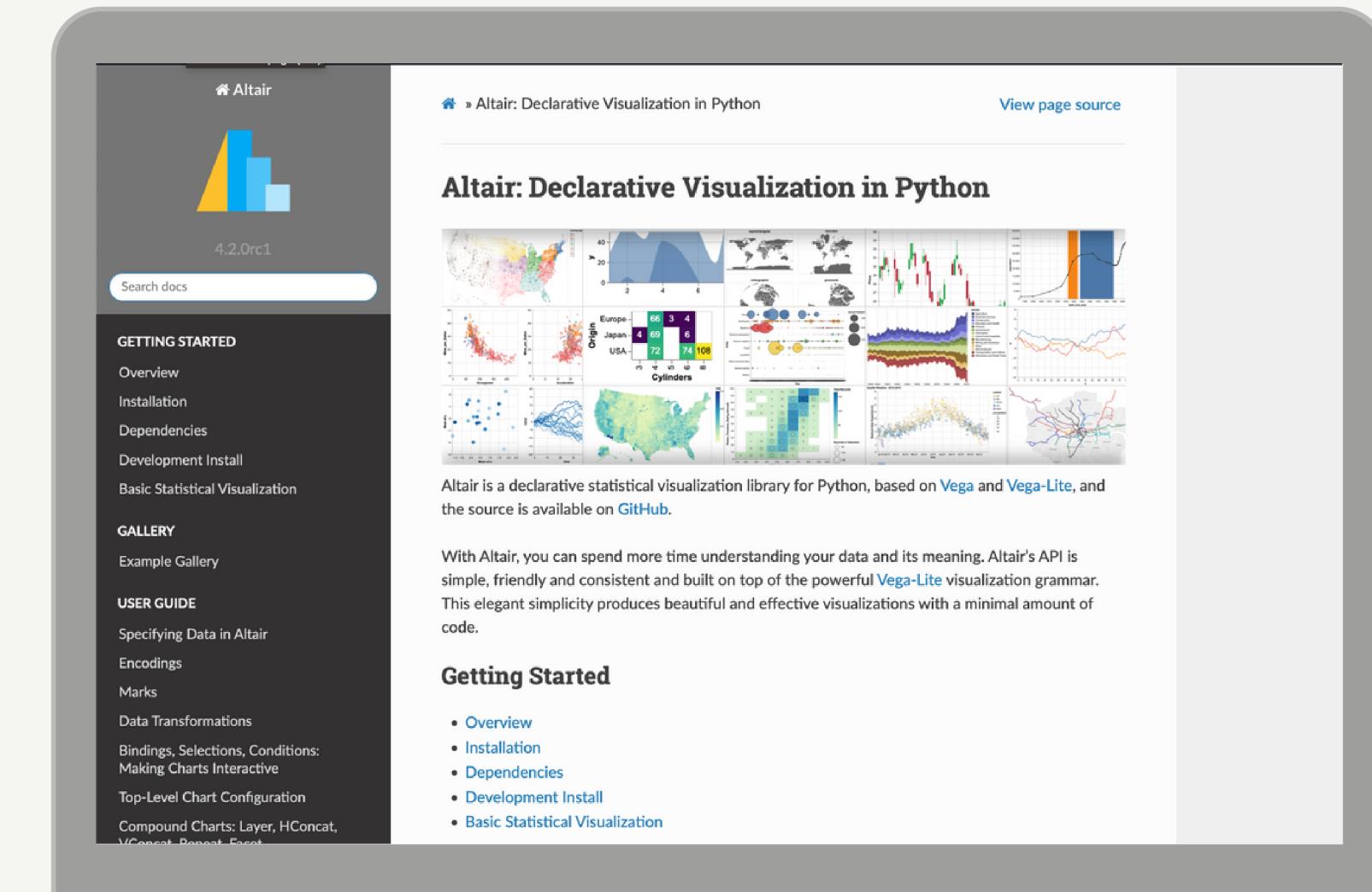
A layered grammar of graphics

Vega-Lite y Altair

Altair

Altair es una biblioteca de visualización estadística declarativa para Python, basada en Vega y Vega-Lite, y la fuente está disponible en GitHub.

Con Altair, puedes dedicar más tiempo a entender tus datos y su significado. La API de Altair es simple, amigable y consistente y está construida sobre la poderosa gramática de visualización Vega-Lite. Esta elegante simplicidad produce visualizaciones hermosas y efectivas con una cantidad mínima de código.



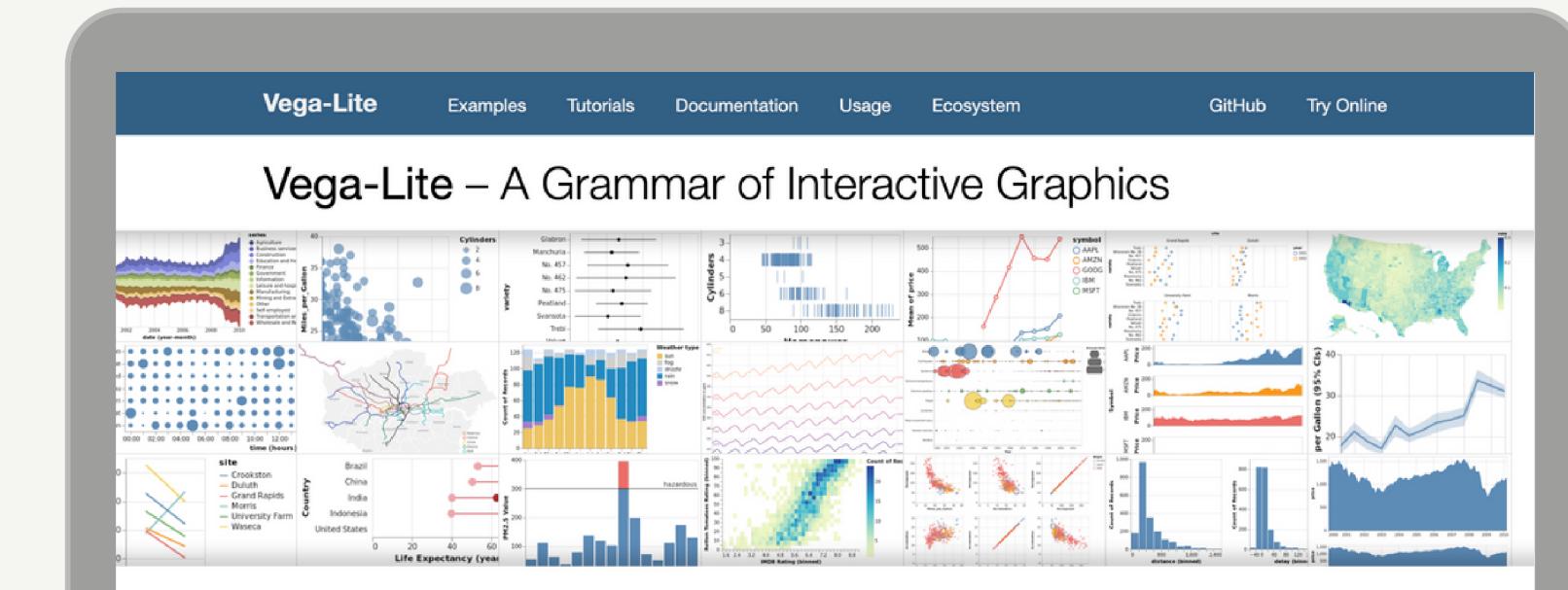
vega.github.io/vega-lite/

Vega-Lite

Vega-Lite es una gramática de alto nivel de gráficos interactivos.

Las especificaciones de Vega-Lite describen las visualizaciones como mapeos de codificación de datos a propiedades de marcas gráficas (por ejemplo, puntos o barras). El compilador de Vega-Lite produce automáticamente componentes de visualización que incluyen ejes, leyendas y escalas.

Determina las propiedades por defecto de estos componentes basándose en un conjunto de reglas cuidadosamente diseñadas.



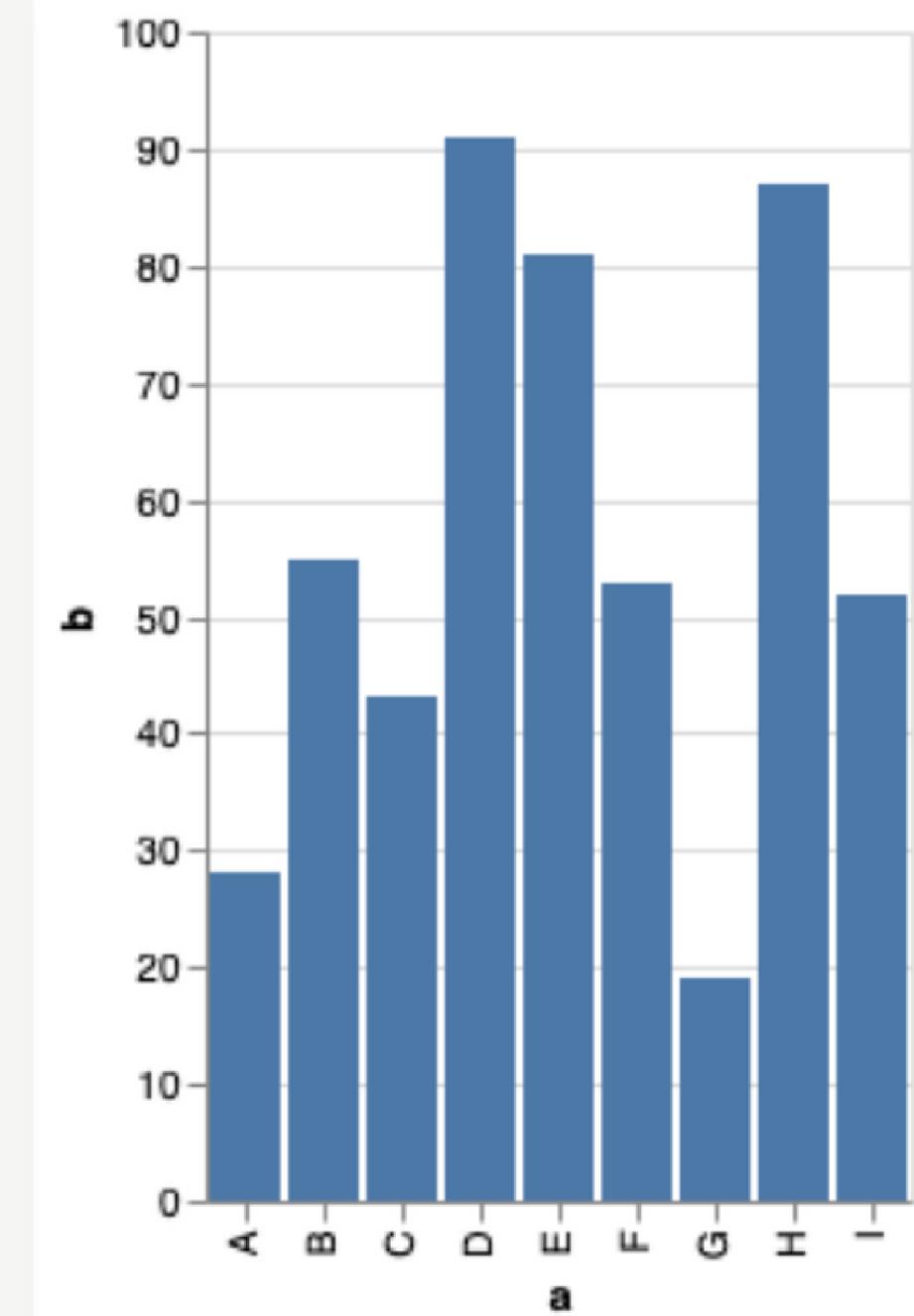
Vega-Lite is a high-level grammar of interactive graphics. It provides a concise, declarative JSON syntax to create an expressive range of visualizations for data analysis and presentation.

Vega-Lite specifications describe visualizations as encoding mappings from data to properties of graphical marks (e.g., points or bars). The Vega-Lite compiler automatically produces visualization components including axes, legends, and scales. It determines default properties of these components based on a set of carefully designed rules. This approach allows Vega-Lite specifications to be concise for quick visualization authoring, while giving user control to override defaults and customize various parts of a visualization. As we also designed Vega-Lite to support data analysis, Vega-Lite supports both data transformations (e.g., aggregation, binning, filtering, sorting) and visual transformations (e.g., stacking and faceting). Moreover, Vega-Lite specifications can be composed into layered and multi-view displays, and made interactive with selections.

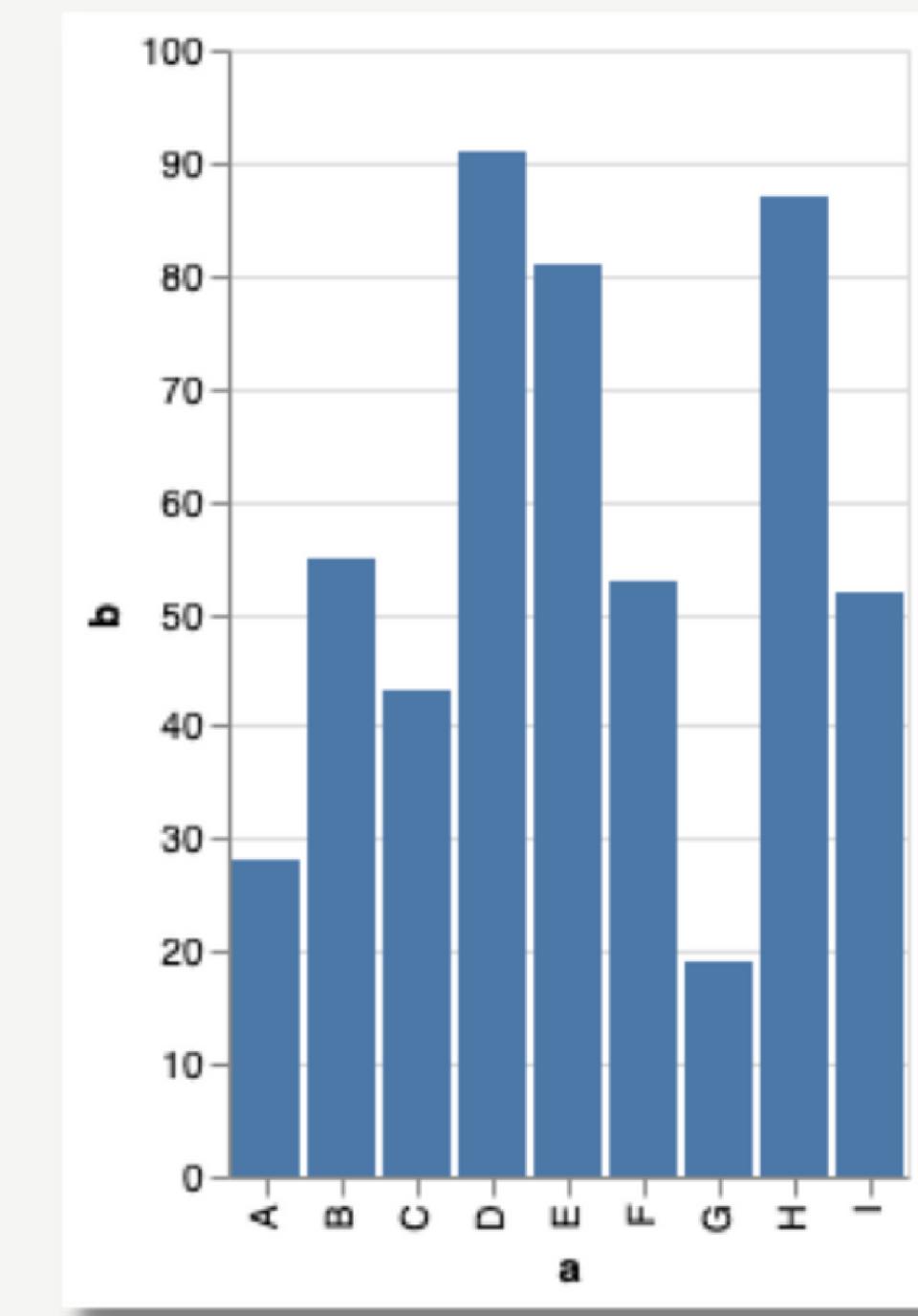
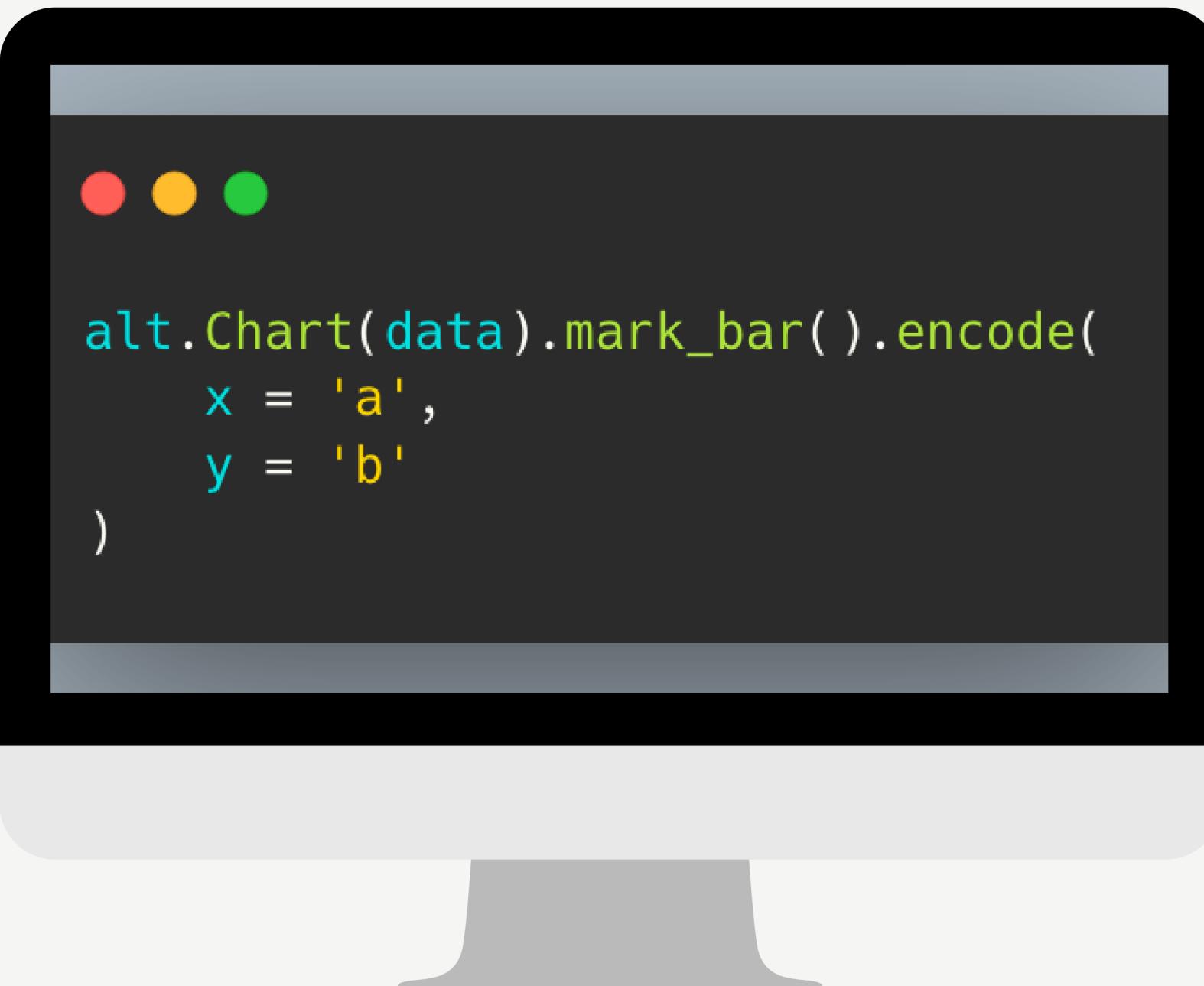
Get started
Latest Version: 5.6.0

Try online

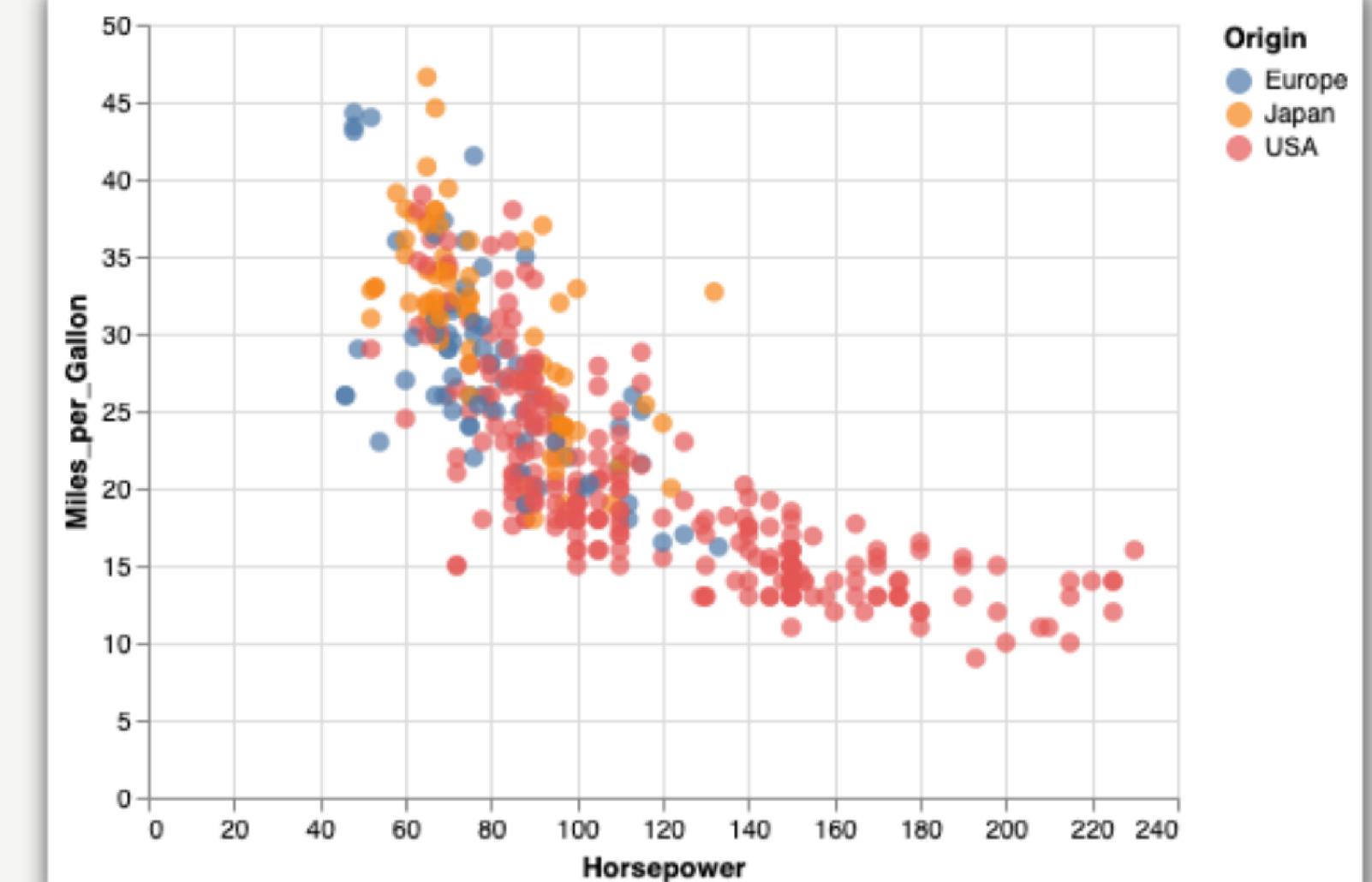
```
● ● ●  
import altair as alt  
import pandas as pd  
  
source = pd.DataFrame({  
    'a': ['A', 'B', 'C', 'D', 'E', 'F', 'G', 'H', 'I'],  
    'b': [28, 55, 43, 91, 81, 53, 19, 87, 52]  
})  
  
alt.Chart(source).mark_bar().encode(  
    x='a',  
    y='b'  
)
```



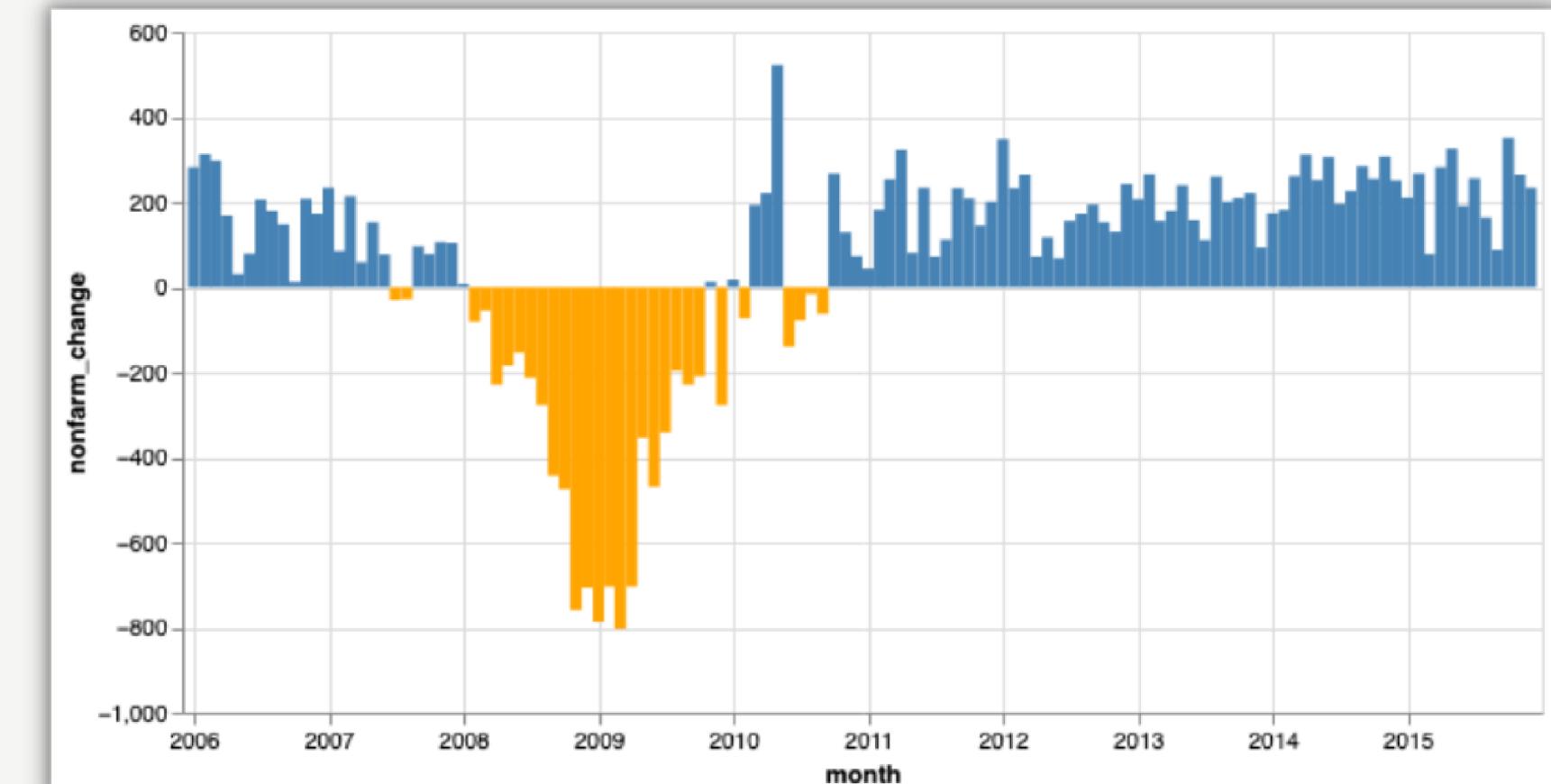
a	b
A	28
B	55
C	43
D	91
E	81
F	53
G	19
H	87
I	52

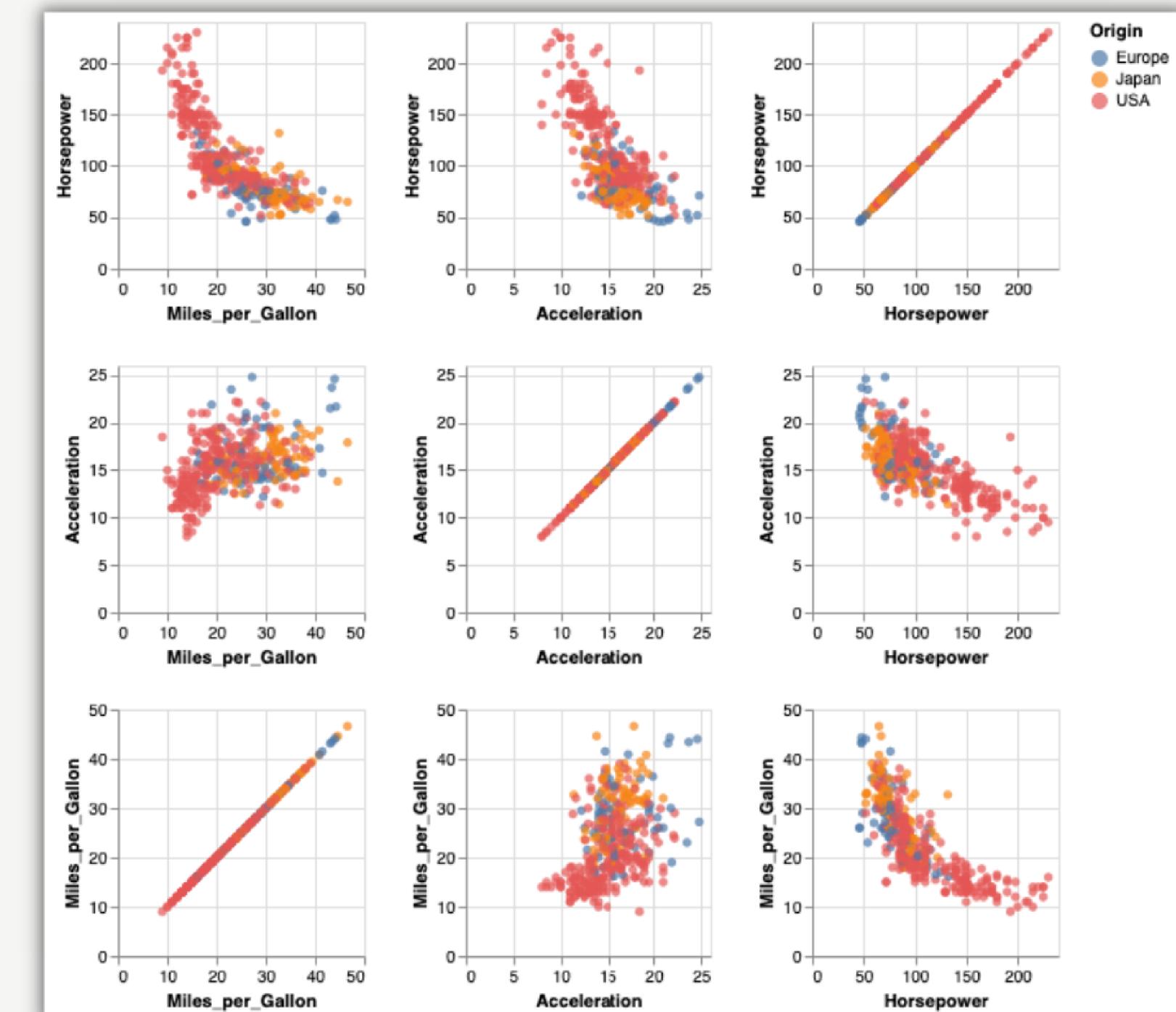
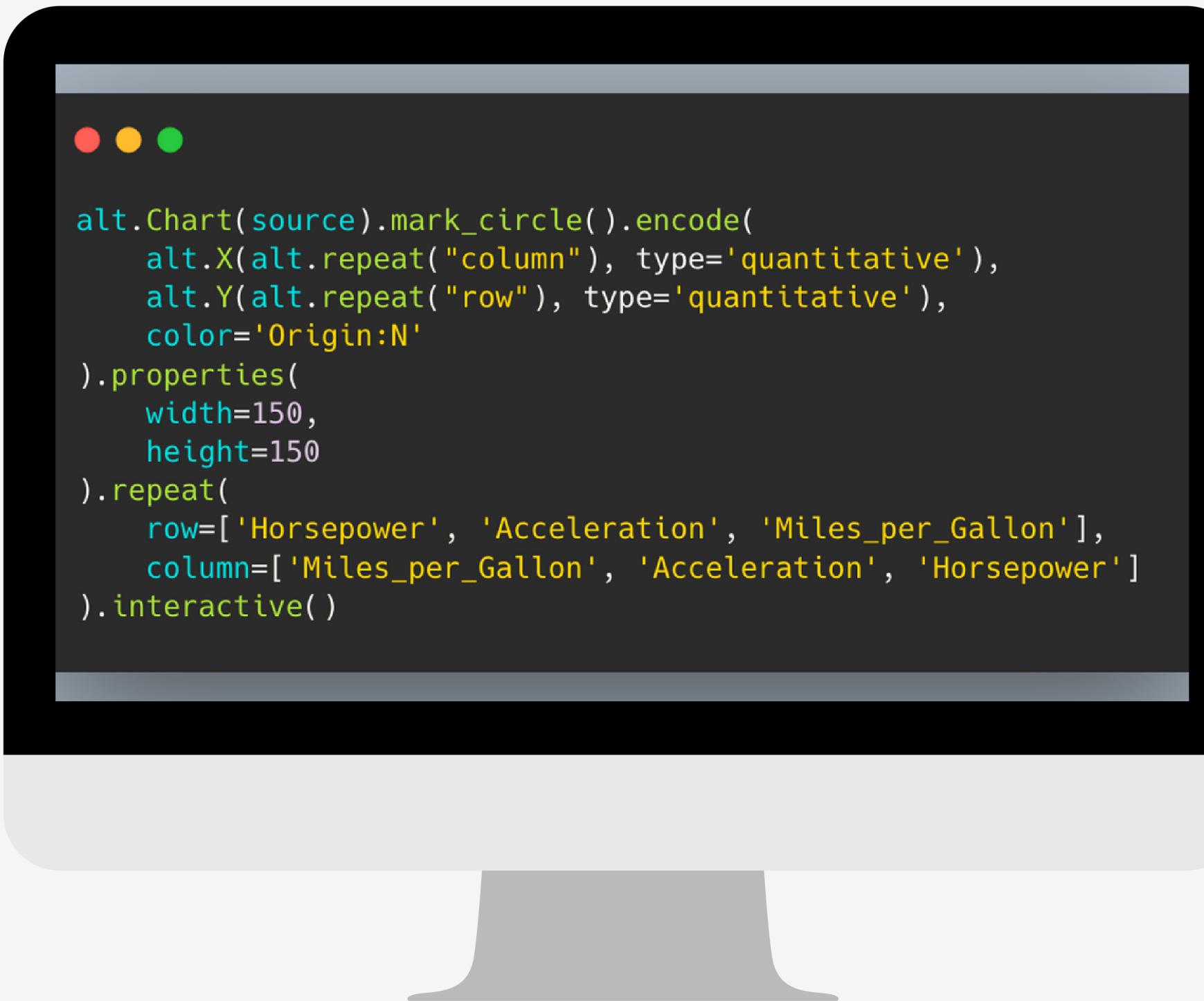


```
alt.Chart(source).mark_circle(size=60).encode(  
    x='Horsepower',  
    y='Miles_per_Gallon',  
    color='Origin',  
    tooltip=[  
        'Name',  
        'Origin',  
        'Horsepower',  
        'Miles_per_Gallon'  
    ].interactive()
```

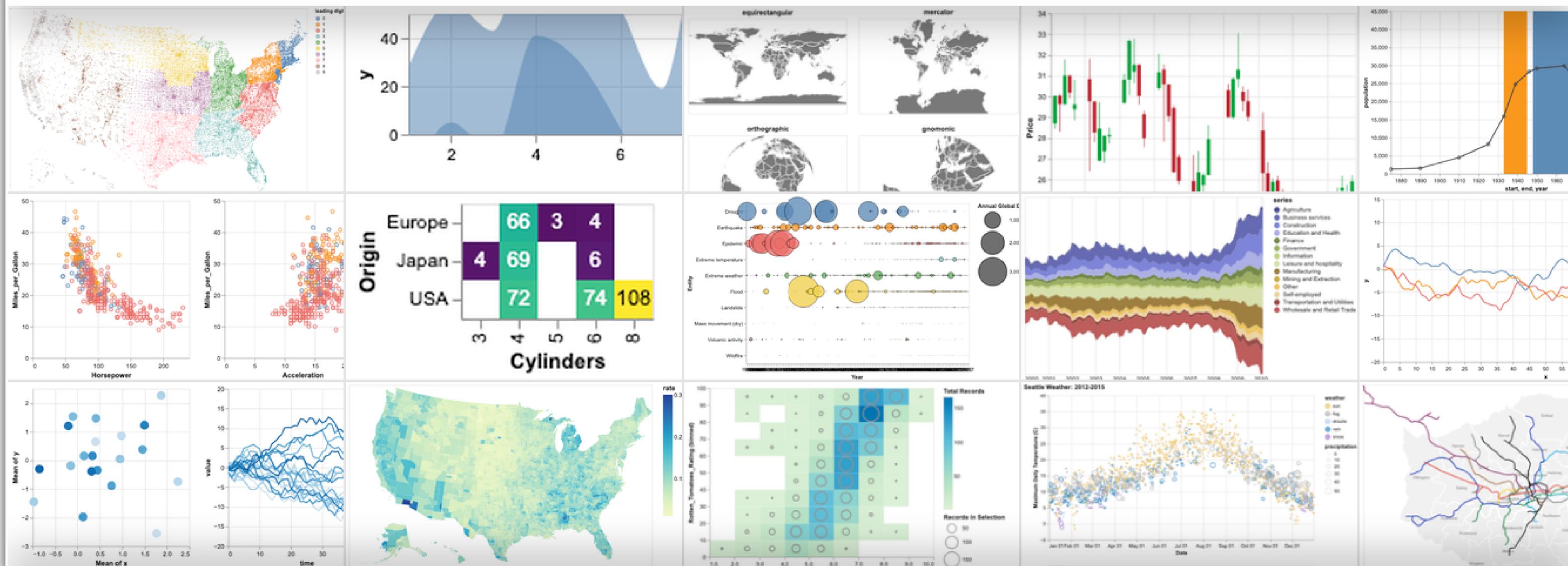


```
● ● ●  
alt.Chart(source).mark_bar().encode(  
    x="month:T",  
    y="nonfarm_change:Q",  
    color=alt.condition(  
        alt.datum.nonfarm_change > 0,  
        alt.value("steelblue"), # Positivo  
        alt.value("orange") # Negativo  
    )  
).properties(width=600)
```



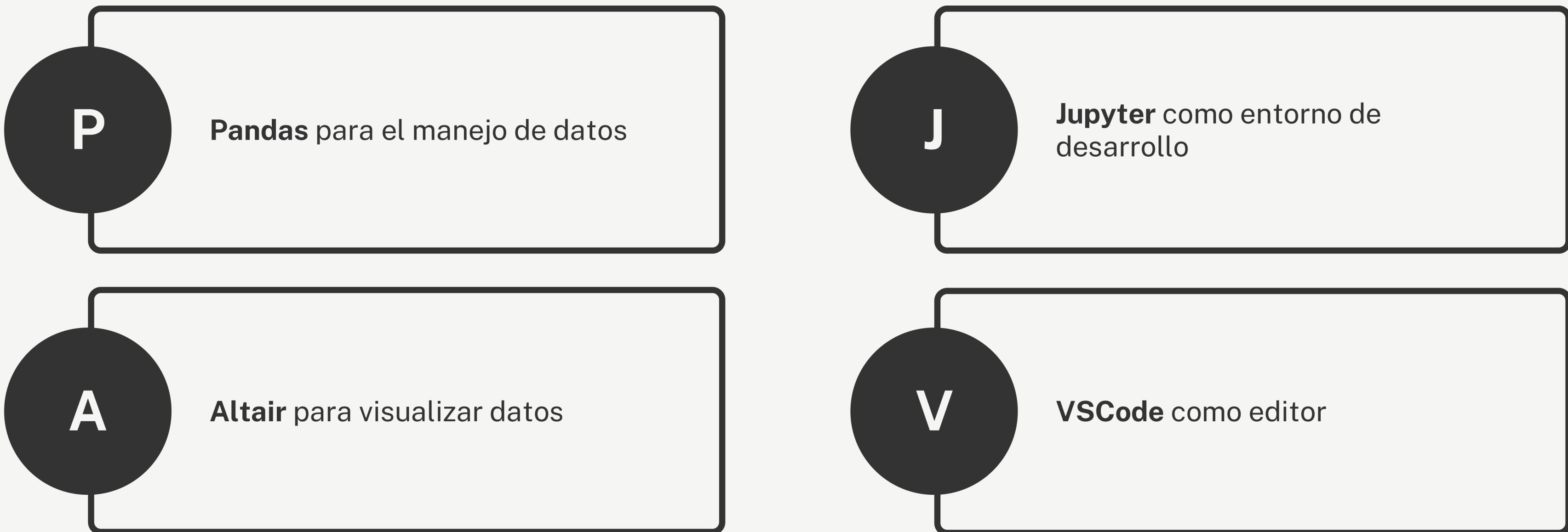


Vega-Altair: Declarative Visualization in Python



Herramientas

Herramientas



Enlaces

try.jupyter.org

mybinder.org

github.com/tacosdedatos/codigofacilito-dataviz-efectiva

altair-viz.github.io

Tarea

Tarea

Crear una visualización de datos a mano

Tarea

Crear una visualización de datos a mano

- no tardes mas de una hora
- puedes usar código/tecnología para preparar tus datos (si quieres)
- tiene que ser una gráfica "física" no digital
- cuando termines compártela en el grupo de telegram o por correo a chekos@tacosdedatos.com

Tarea

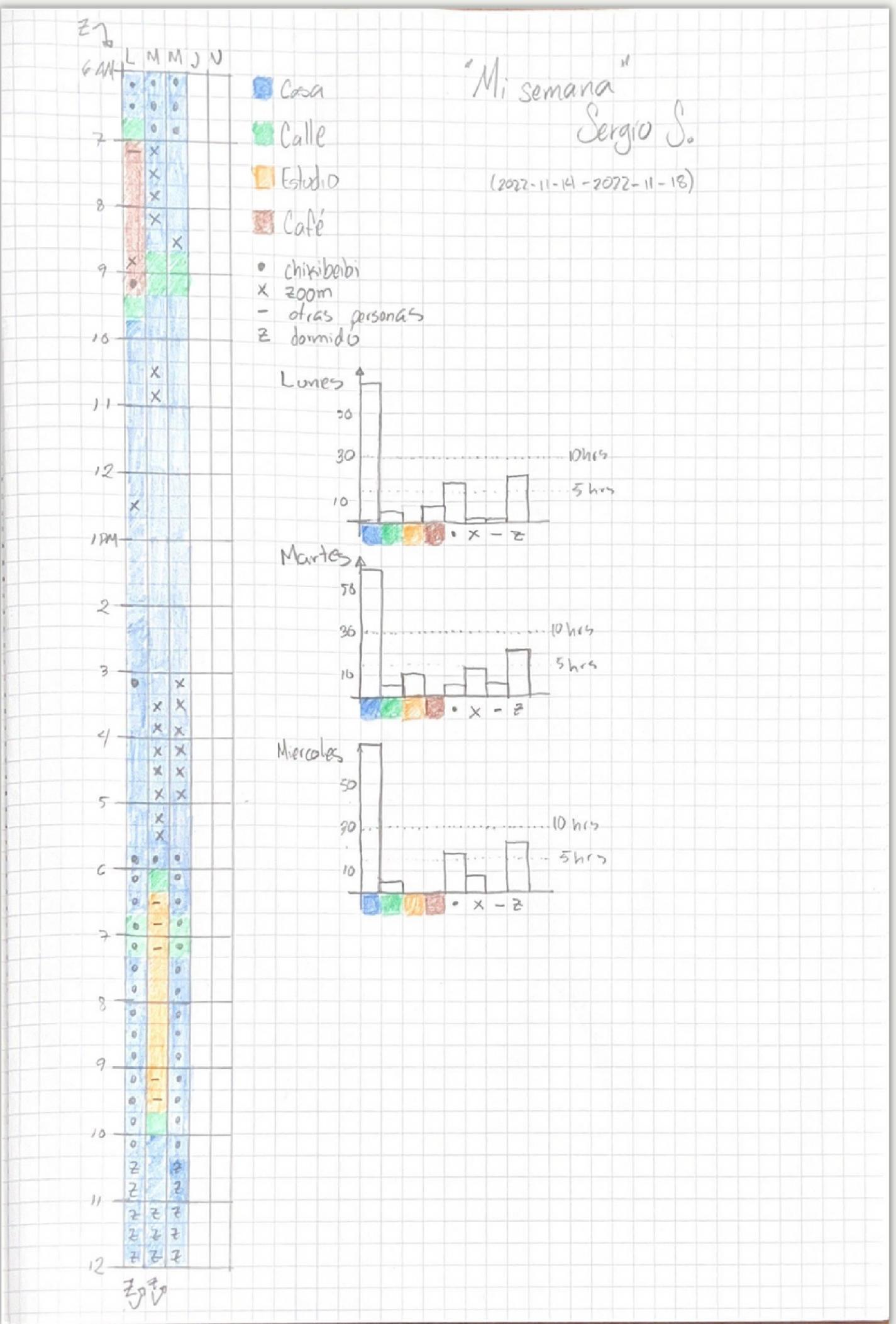
Crear una visualización de datos a mano

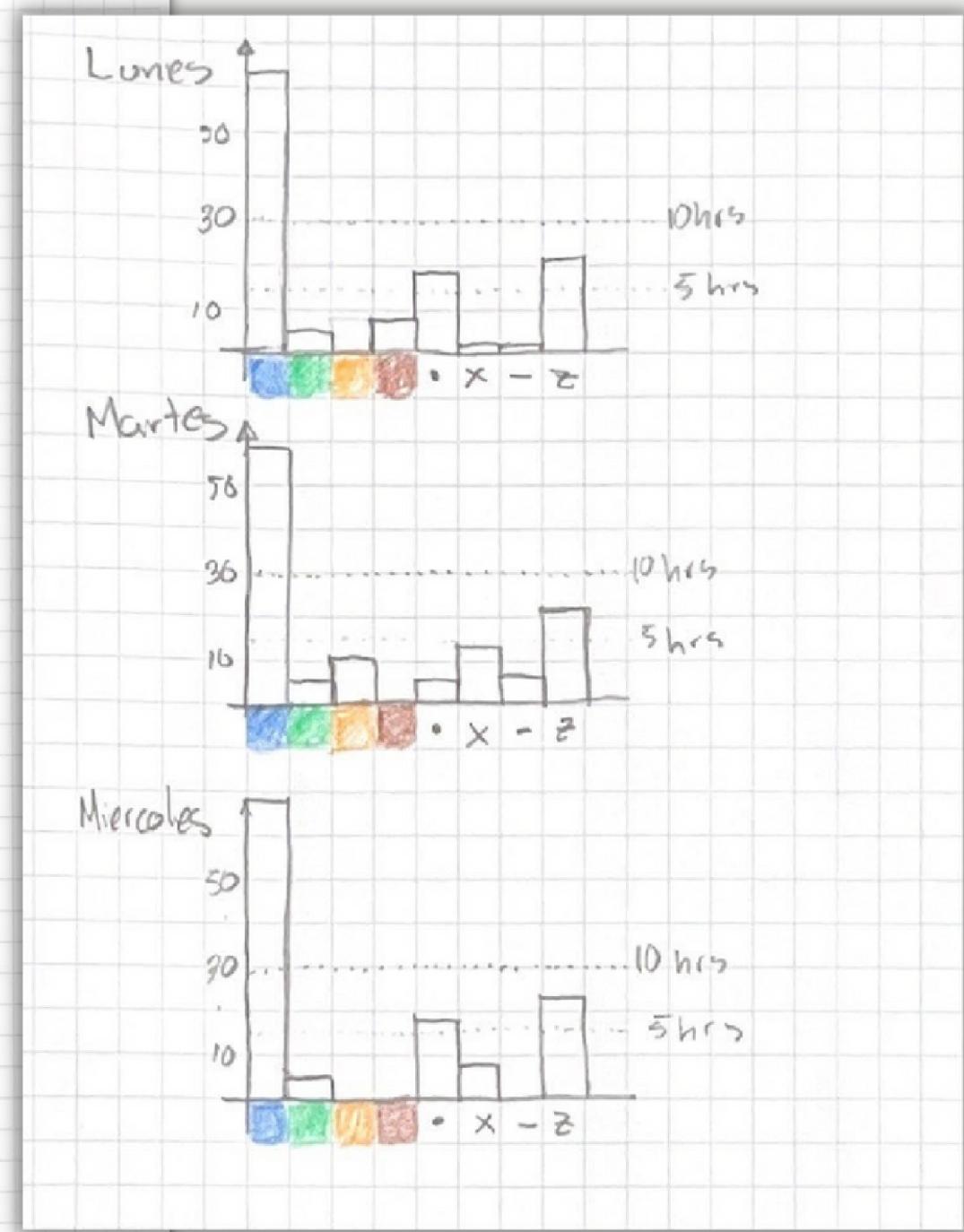
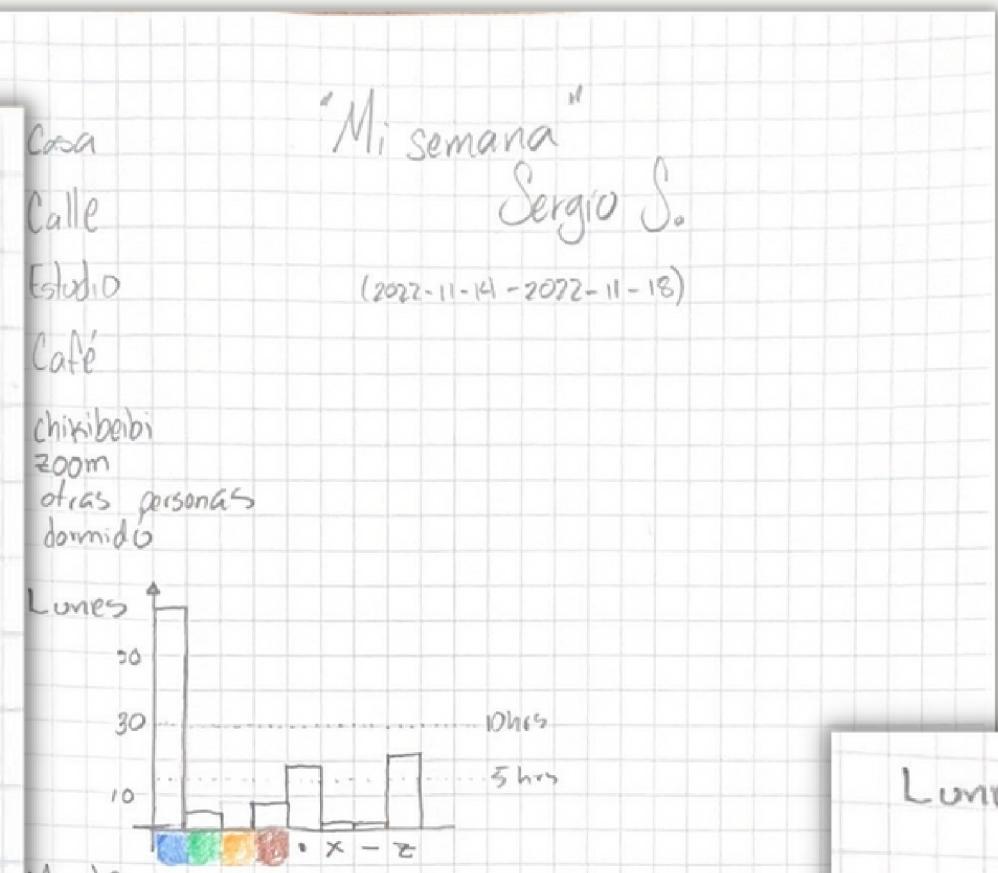
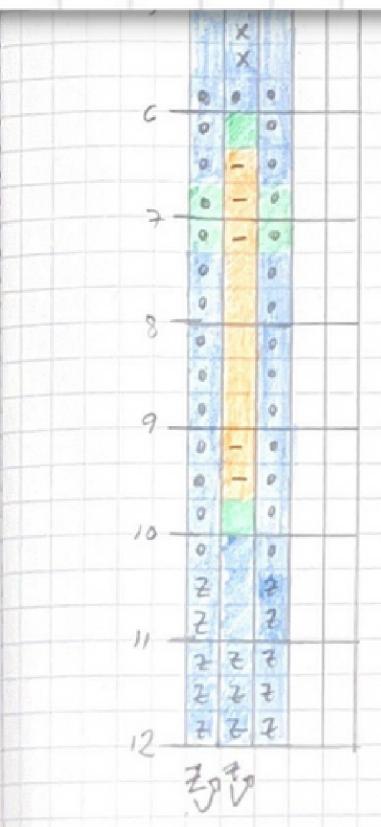
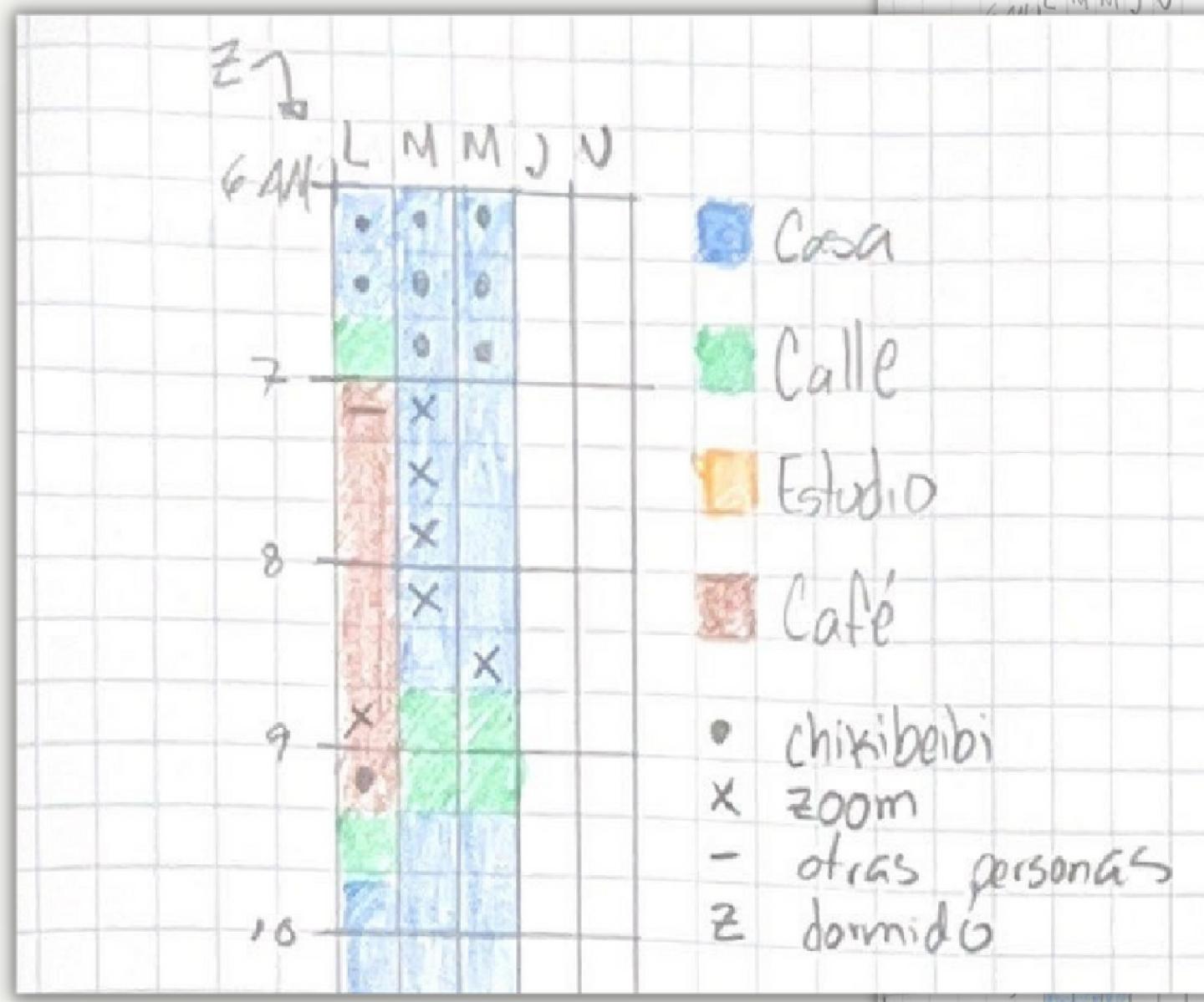
- no tardes mas de una hora
- puedes usar código/tecnología para preparar tus datos (si quieres)
- tiene que ser una gráfica "física" no digital
- cuando termines compártela en el grupo de telegram o por correo a chekos@tacosdedatos.com

Piensa en como representar tus datos de manera visual (colores, figuras y símbolos, etc.)

Como sugerencia podrías preguntarte:

- *¿cómo visualizaría los libros que tengo en mi casa?*
- *... como he vivido mis años?*
- *... mis partes favoritas de mis canciones favoritas?*
- *... donde gasto mi tiempo?*
- *... mis amistades? ¿Dónde las conocí, desde hace cuánto?*
- *... cuantas distracciones tengo al día? ¿De dónde provienen?*
- *... mi estado de ánimo? Día a día, hora a hora, etc.*





Visualización de Datos Efectiva con Altair

Nov 2022

@ChekosWH | @tacosdedatos

Sergio Sánchez