

Diplomado Big Data y Ciencias de datos Visualización de información

Profesor: Cristián Hernández

Ayudante: Yeison Salinas

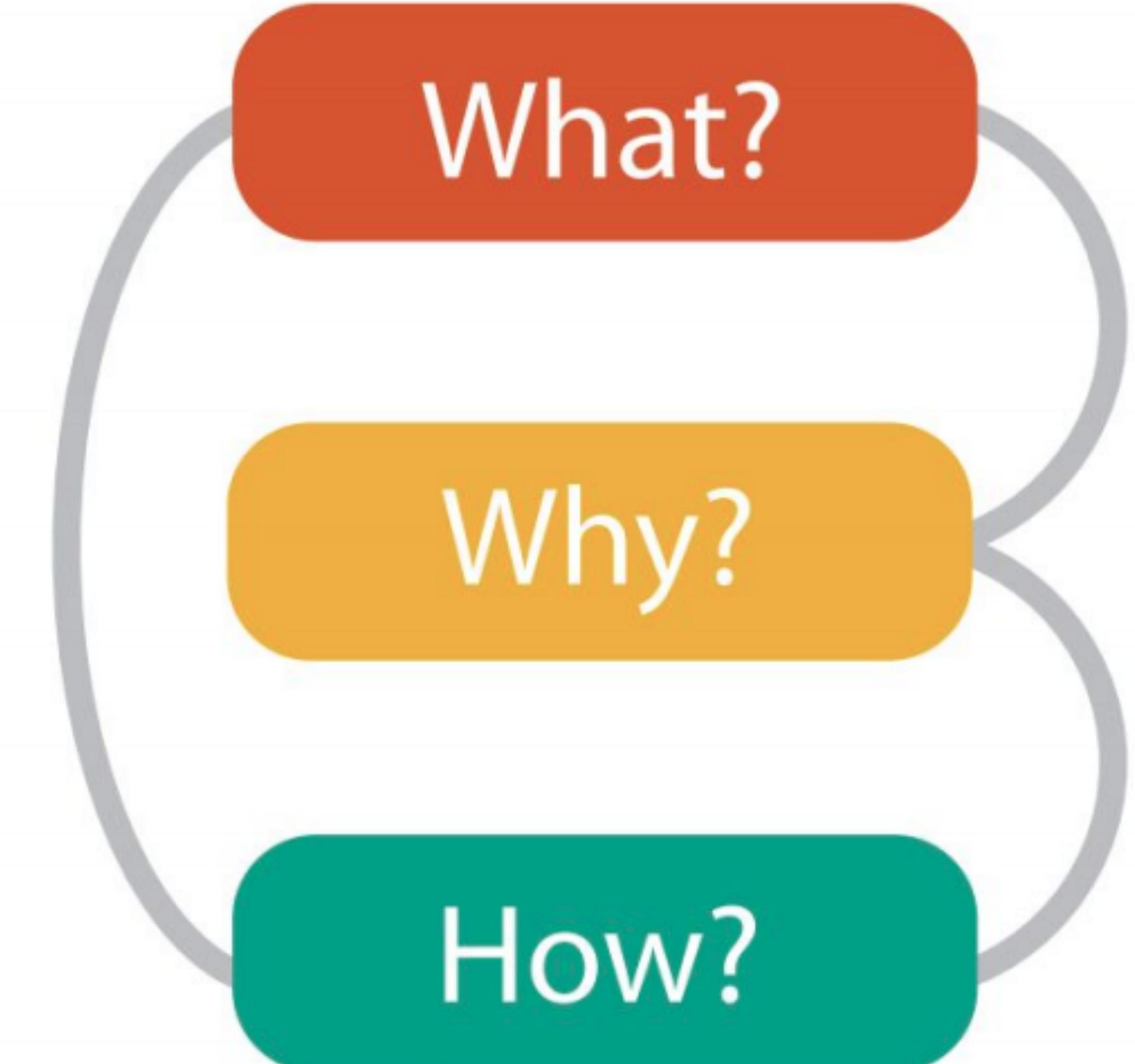
Resumen clase 3

Munzner's Framework

What qué información se está mostrando (*data*).

Why por qué queremos visualizar esto (*task*).

How cómo se construye la visualización (*idiom*).



What

What?

Datasets

→ Data Types

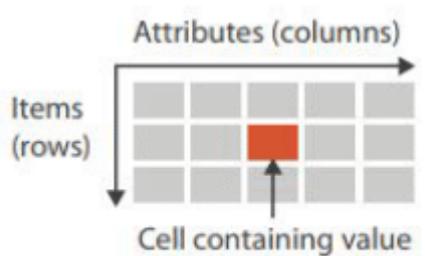
→ Items → Attributes → Links → Positions → Grids

→ Data and Dataset Types

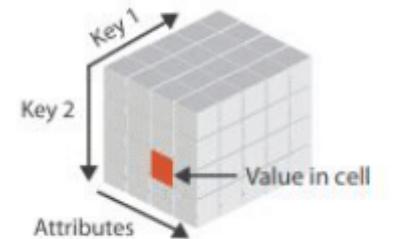


→ Dataset Types

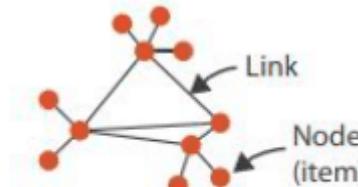
→ Tables



→ Multidimensional Table



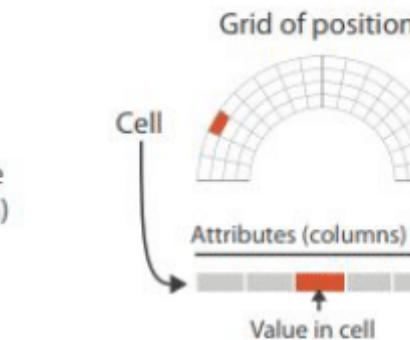
→ Networks



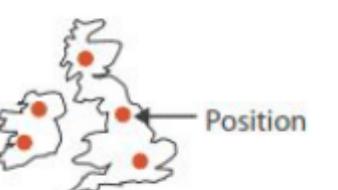
→ Trees



→ Fields (Continuous)



→ Geometry (Spatial)



→ Dataset Availability

→ Static



→ Dynamic



Attributes

→ Attribute Types

→ Categorical



→ Ordered

→ Ordinal



→ Quantitative



→ Ordering Direction

→ Sequential



→ Diverging



→ Cyclic



What?

Why?

How?

Clase 4: Formalizando el proceso de diseño (continuación)

Why

Why?

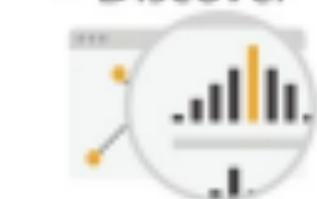
Actions

Targets

→ Analyze

→ Consume

→ Discover



→ Present

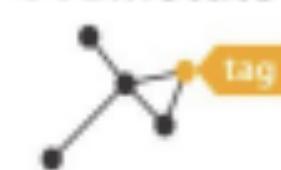


→ Enjoy



→ Produce

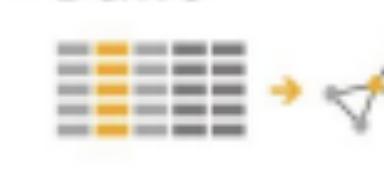
→ Annotate



→ Record



→ Derive

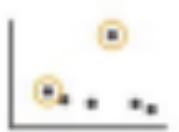


→ Search

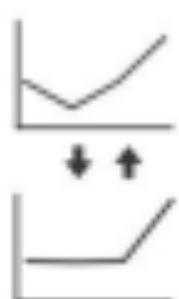
	Target known	Target unknown
Location known	••• <i>Lookup</i>	•• <i>Browse</i>
Location unknown	⟨•••⟩ <i>Locate</i>	⟨•••⟩ <i>Explore</i>

→ Query

→ Identify



→ Compare



→ Summarize



→ All Data

→ Trends



→ Outliers



→ Features



→ Attributes

→ One



→ Many



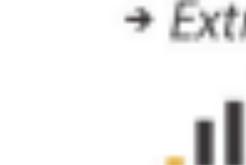
→ Dependency



→ Correlation

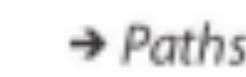
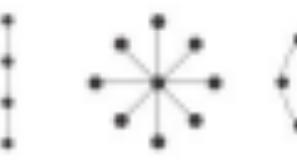
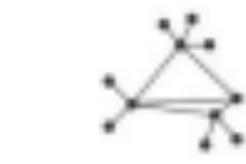


→ Similarity



→ Network Data

→ Topology



→ Paths



→ Spatial Data

→ Shape



Why

Define tres niveles de acciones (del más alto al más bajo):

Analizar

Buscar

Consultar

Estos tres niveles son independientes entre ellos. Incluso, es recomendable describir acciones en todos ellos.

Se describen como pares {Acción, Objetivo}.

Los objetivos deben estar en el contexto del dataset.

Descubrir la distribución de la temperatura.

Localizar los días outliers del año.

Why - ¿por qué un “por qué”?

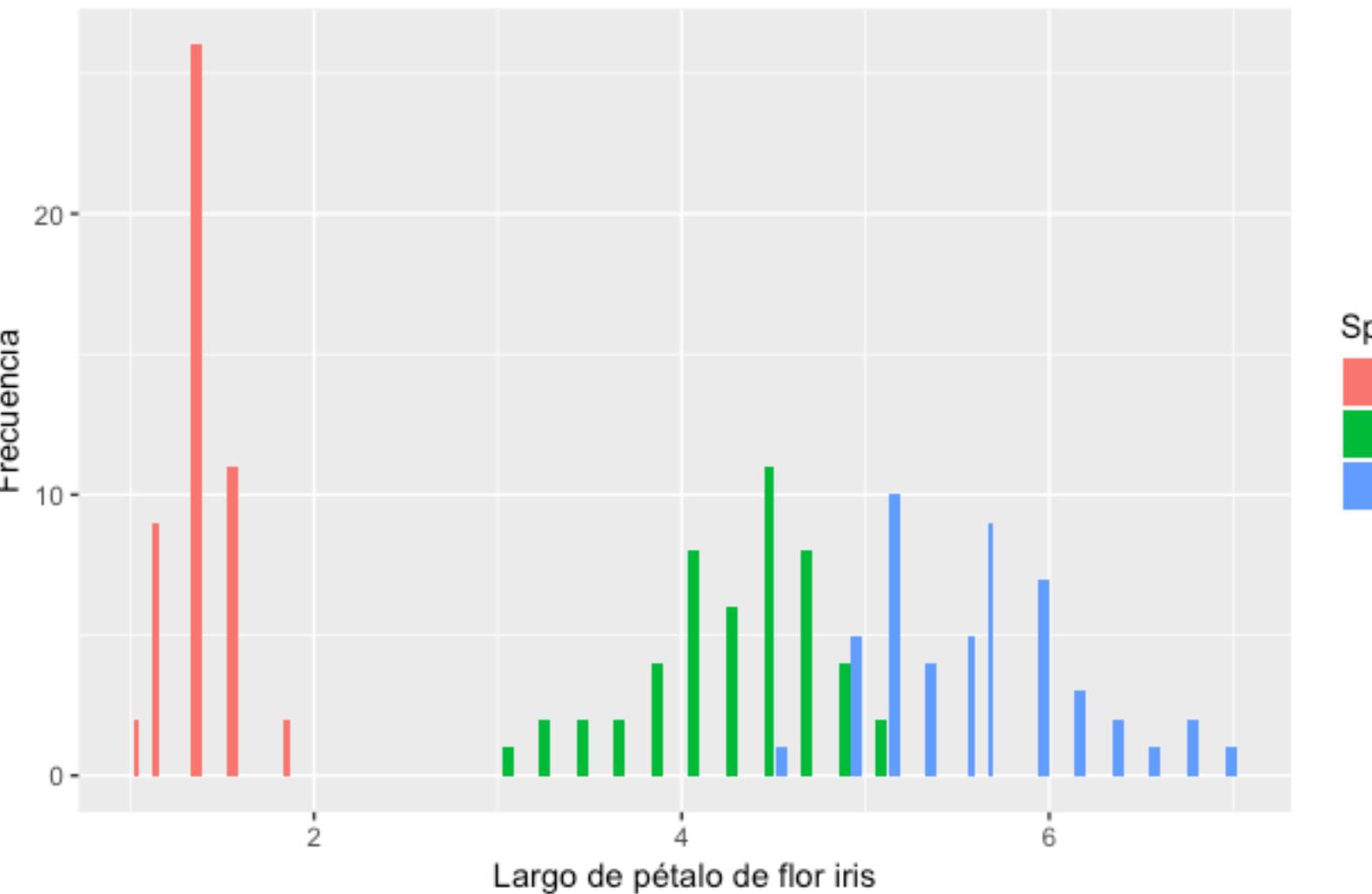
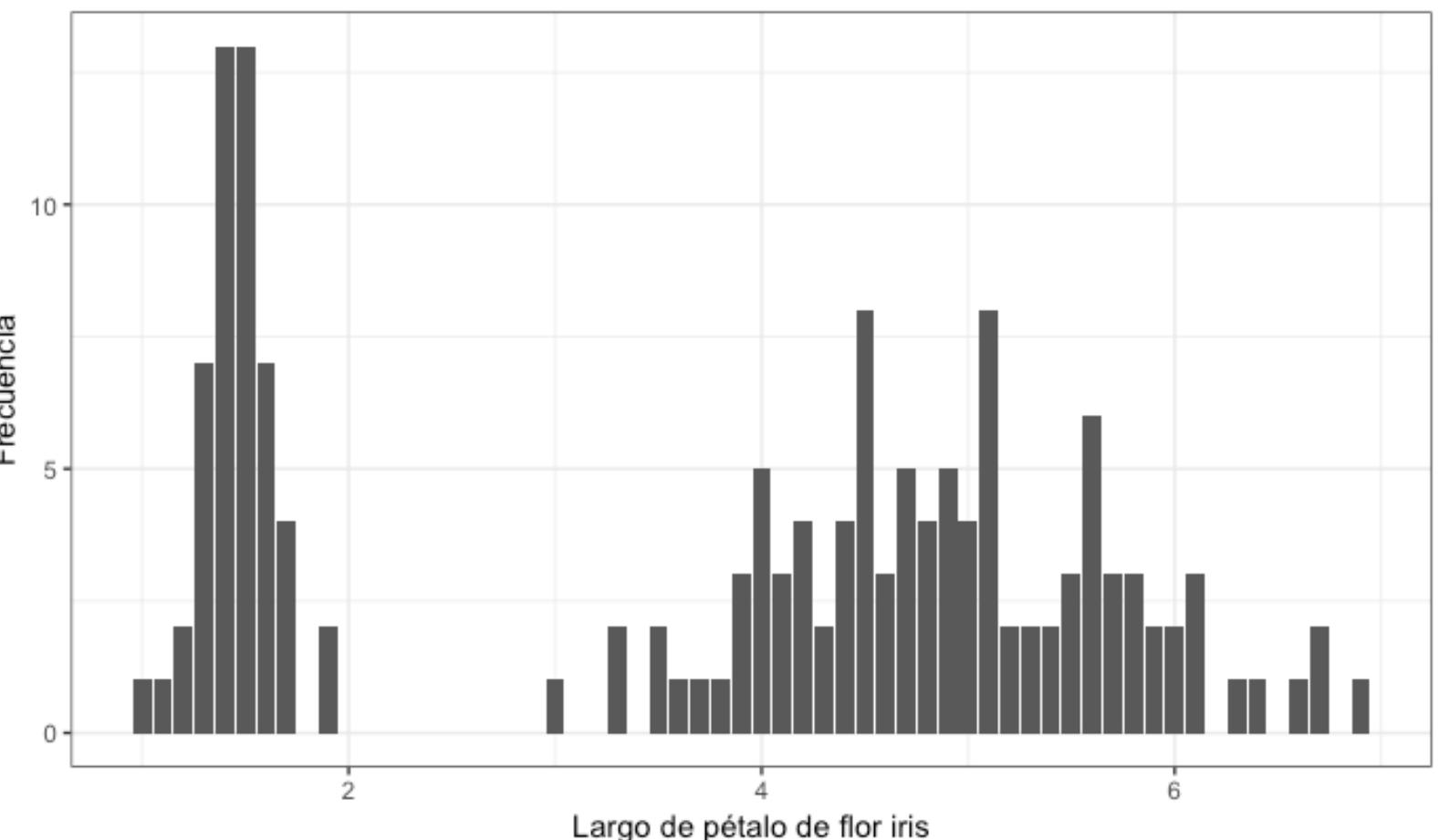
Caso 1: Si nos interesa visualizar la distribución del tamaño de pétalos de una flor,

Un histograma podría ser una buena opción.

Caso 2: Si queremos comparar la distribución de tamaños de pétalos de 3 tipos de flores,

Una visualización que agregue el canal de color y sobreponga 3 histogramas es una mejor opción.

El por qué del gráfico, modifica la decisión del tipo de gráfico para los mismos datos.



¿Por qué tareas genéricas?

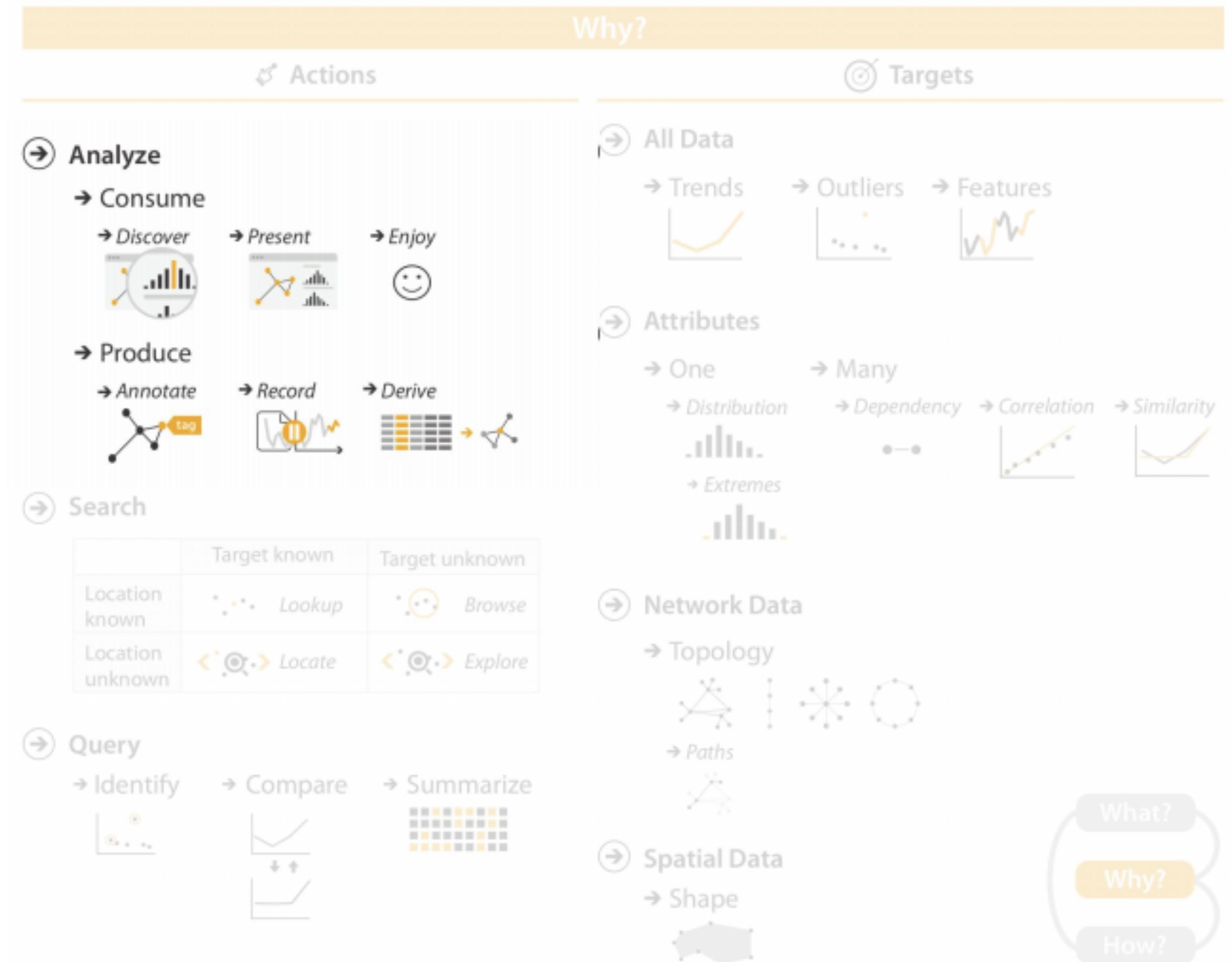
Ganancias de la empresa durante el año en relación la anterior

> Identificar la tendencia de un atributo en función del tiempo

Nivel alto: Analyze

Consumir información ya
existente.

Producir nueva información.



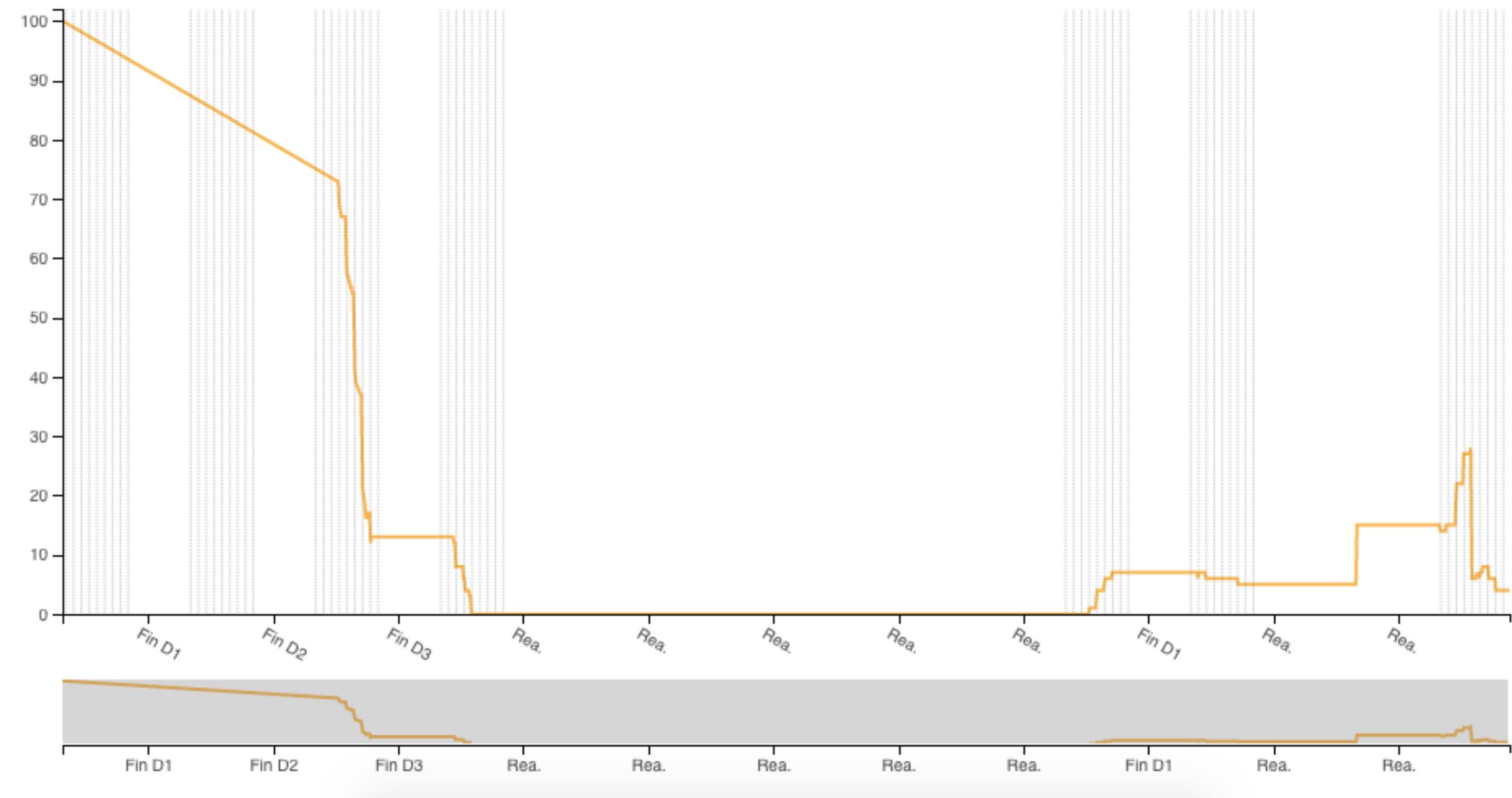
Consumir

Descubrir

La visualización como herramienta para encontrar conocimiento escondido

Puede ayudar a generar, confirmar o desmentir una hipótesis

Consumir - Descubrir



<https://observablehq.com/@hernan4444/visualizacion-de-toma-de-cursos-uc-banner?collection=@hernan4444/tiempo-libre>

Consumir

Descubrir

La visualización como herramienta para encontrar conocimiento escondido

Puede ayudar a generar, confirmar o desmentir una hipótesis

Presentar

La visualización como herramienta para comunicar información

Puede servir como storytelling o para guiar a través de una serie de operaciones cognitivas

Lo importante de esta acción es que se utiliza para comunicar algo específico y que la audiencia ya conoce

Consumir - Presentar

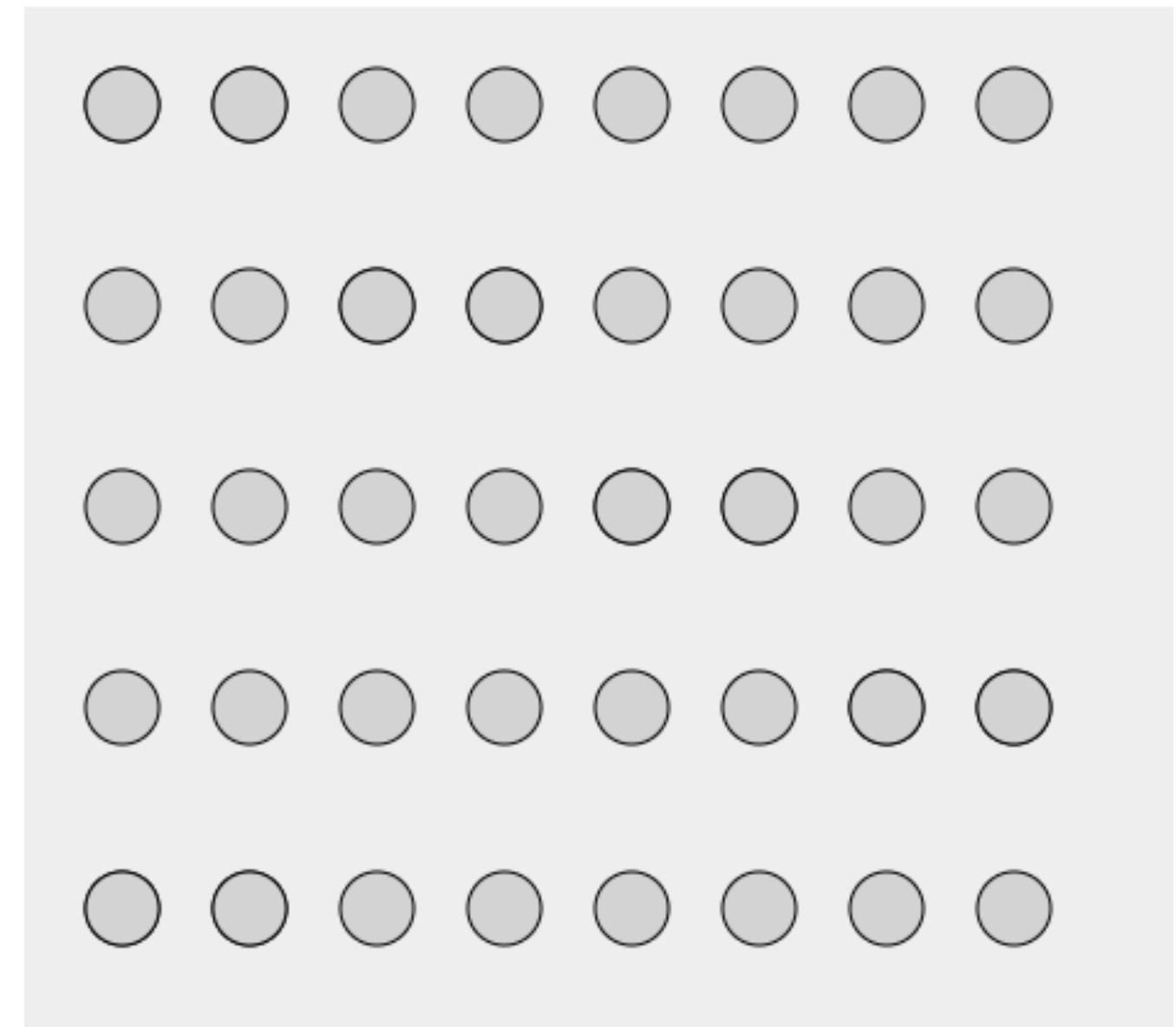
Gestalt Principles for Data Visualization

Similarity, Proximity & Enclosure

Introduction

At a recent talk I challenged the audience to define several gestalt principles based solely on representative figures. This "academic" approach to data visualization seems in opposition to a "pragmatic" approach that focuses on best practices and prior art demonstrated in the growing library of data visualization books and 2-day seminars.

But let me suggest that gestalt is very much a pragmatic aspect of creating data visualization, in fact a necessary aspect if you plan to do more than simple bar and line charts (and perhaps even for those simple charts). This exploration of three of the most simple gestalt principles focuses on how they operate and how they might act in tandem with and in opposition to each other. I also include some gestures toward how the gestalt may already be influencing what we think of not as cognitive qualities but as design and style in data visualization.



<http://emeeks.github.io/gestaltdataviz/section1.html>

Consumir

Descubrir

La visualización como herramienta para encontrar conocimiento escondido

Puede ayudar a generar, confirmar o desmentir una hipótesis

Presentar

La visualización como herramienta para comunicar información

Puede servir como storytelling o para guiar a través de una serie de operaciones cognitivas

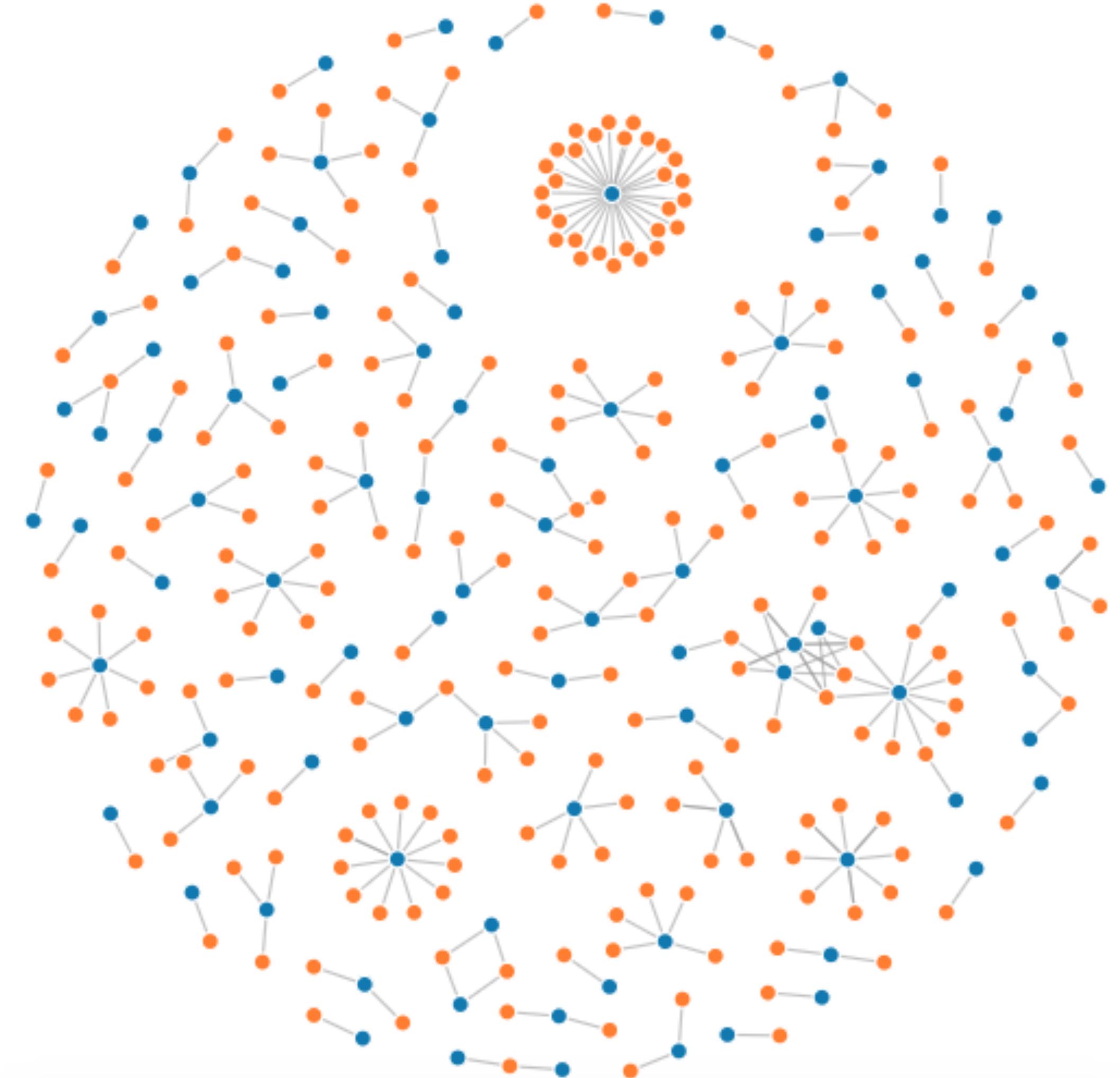
Lo importante de esta acción es que se utiliza para comunicar algo específico y que la audiencia ya conoce

Disfrutar

El usuario es atraído por curiosidad (e.g. satisfacción de ver una infografía bien diseñada)

Puede ocurrir que la visualización haya sido creada para cierta audiencia, pero otro grupo de gente la utilice por placer

Consumir - Disfrutar

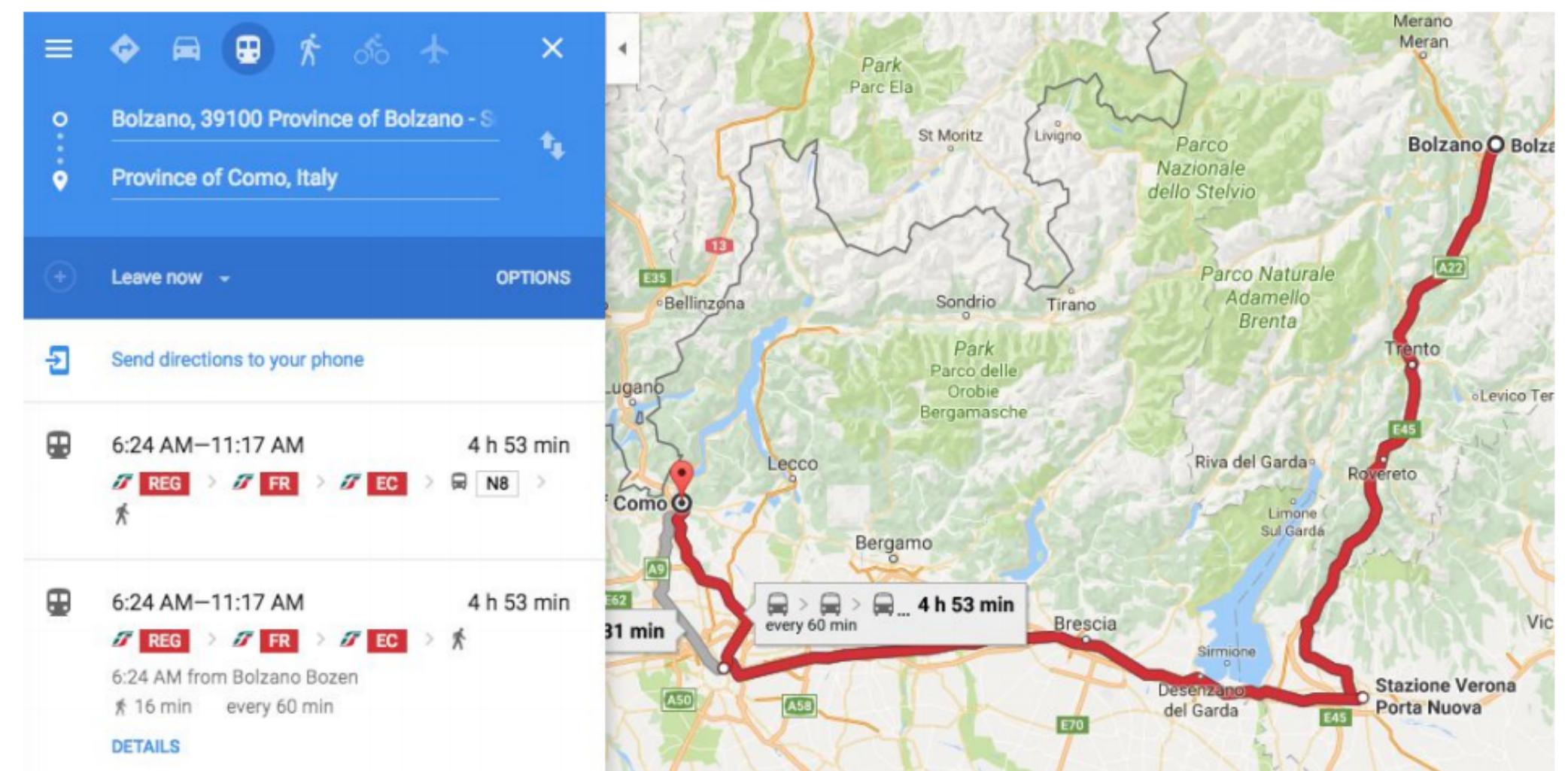


<https://observablehq.com/@d3/disjoint-force-directed-graph>

Producir Comentar/etiqueta

Se agregan comentarios gráficos o textuales

Son temporales



Producir

Comentar/etiqueta

Se agregan comentarios gráficos o textuales

Son temporales

Grabar

La idea es generar artefactos persistentes a partir de elementos de la visualización

Por ejemplo, screenshots, lista de bookmarks, ajuste de parámetros, entre otros

La diferencia con comentar es que estas anotaciones quedan almacenadas

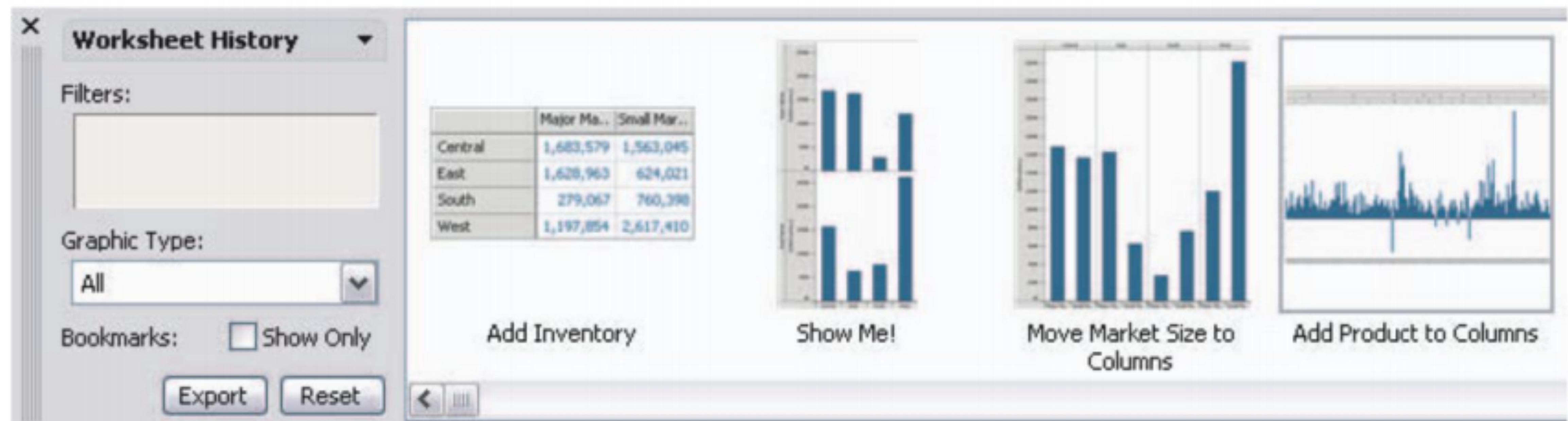


Fig. 1. A Graphical History Interface. Thumbnails show previous visualization states and labels describe the actions performed.

Producir

Comentar/etiqueta

Se agregan comentarios gráficos o textuales

Son temporales

Grabar

La idea es generar artefactos persistentes a partir de elementos de la visualización

Por ejemplo, screenshots, lista de bookmarks, ajuste de parámetros, entre otros

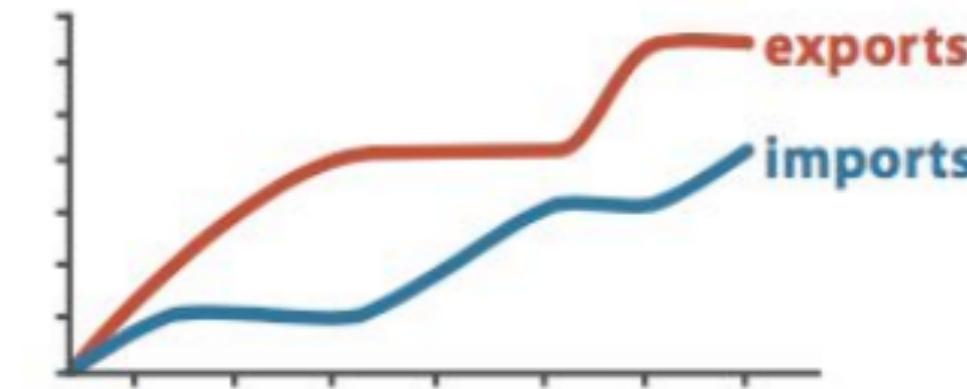
La diferencia con comentar es que estas anotaciones quedan almacenadas

Derivar (Transformar)

El objetivo es producir nuevos datos basados en los existentes.

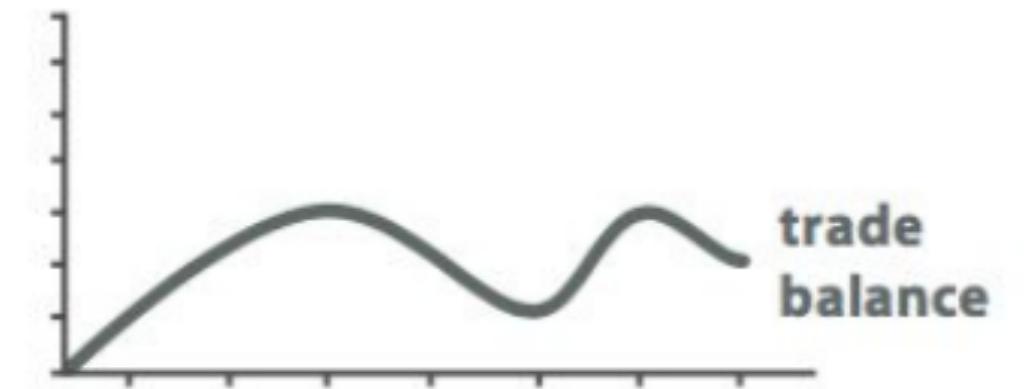
Por ejemplo, diferencia de exportaciones con importaciones para tener el valor neto.

Producir - Calcular ganancias



Original Data

(a)



$trade\ balance = exports - imports$

Derived Data

(b)

Figure 3.5. Derived attributes can be directly visually encoded. (a) Two original data attributes are plotted, imports and exports. (b) The quantitative derived attribute of trade balance, the difference between the two originals, can be plotted directly.

Nivel medio: Search

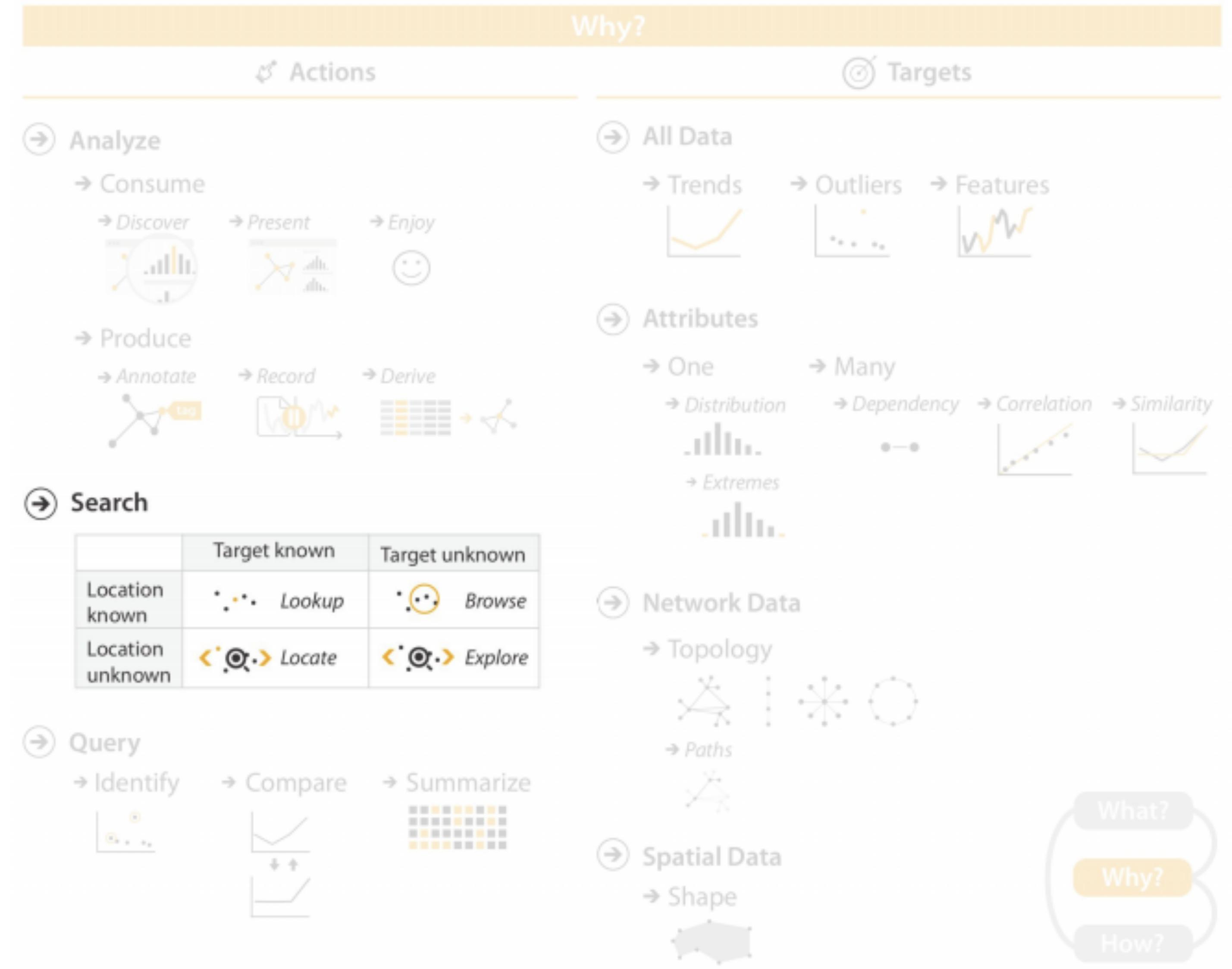
Hay dos variables:

Qué buscamos (target).

Dónde estamos buscando (location).

En ambas variables, ellas pueden ser conocidas o desconocidas.

Se generan cuatro combinaciones en total



Nivel medio: Search

Lookup: sé que estoy buscando y también dónde está

Por ejemplo, buscar humanos en una jerarquía taxonómica.

Locate: sé que estoy buscando, pero no dónde está

Ahora, buscar conejos.

- **Browse:** no tengo claro que estoy buscando, pero sí dónde debería estar
 - Por ejemplo, buscar por ítems que hagan match con un rango particular de atributos
- **Explore:** sólo sé que nada sé (?)
 - Por ejemplo, buscar outliers en un scatterplot

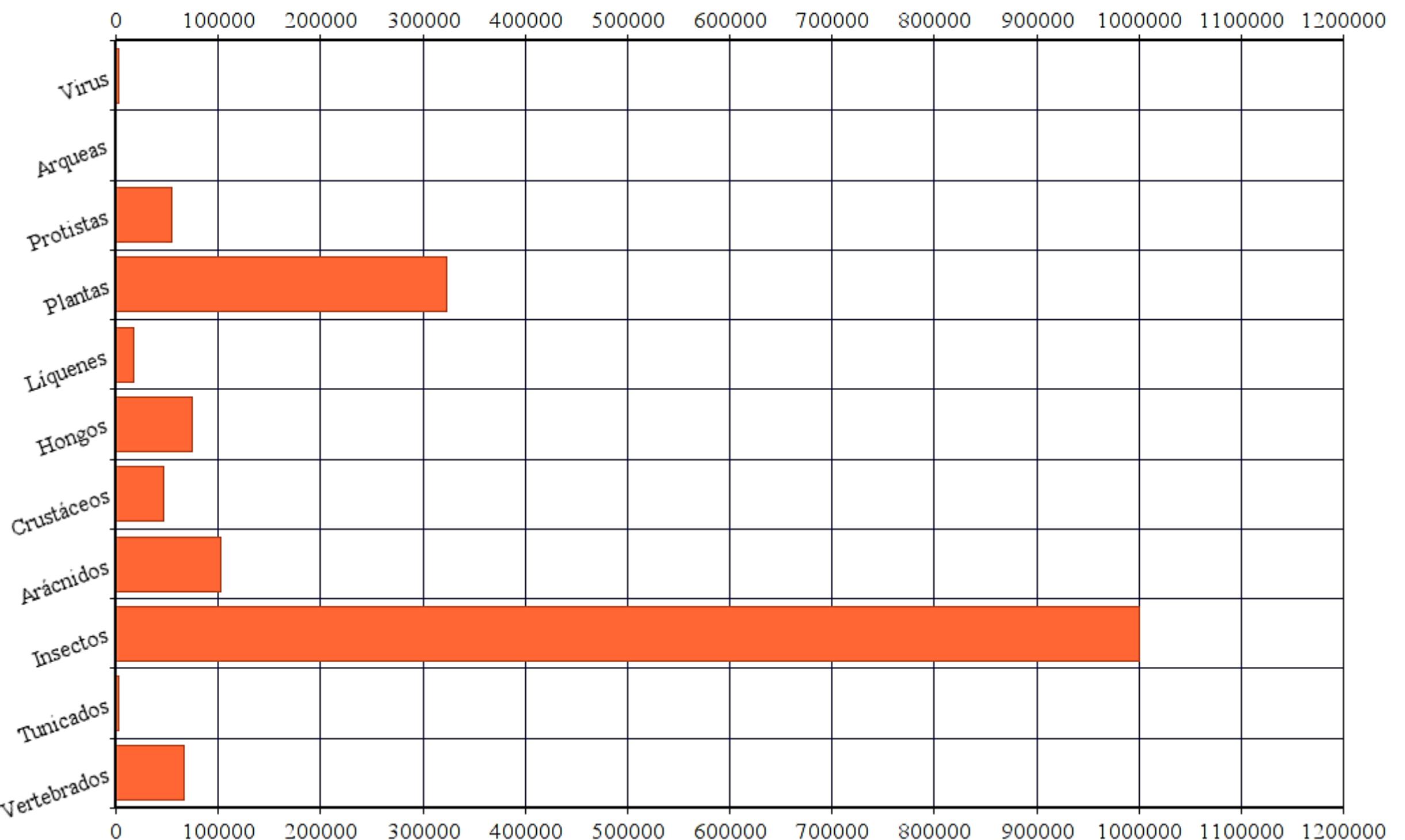
	Target known	Target unknown
Location known	••• <i>Lookup</i>	•• <i>Browse</i>
Location unknown	◁•○▷ <i>Locate</i>	◁•○▷ <i>Explore</i>

Search - Locate y Lookup

Contexto: el gráfico muestra la cantidad de especies diferentes en cada categoría.

Locate: Si queremos buscar la cantidad de especies de las plantas, dado que el **gráfico no está ordenado**, será necesario buscar en todo el gráfico, pero sabemos el dato en que buscamos.

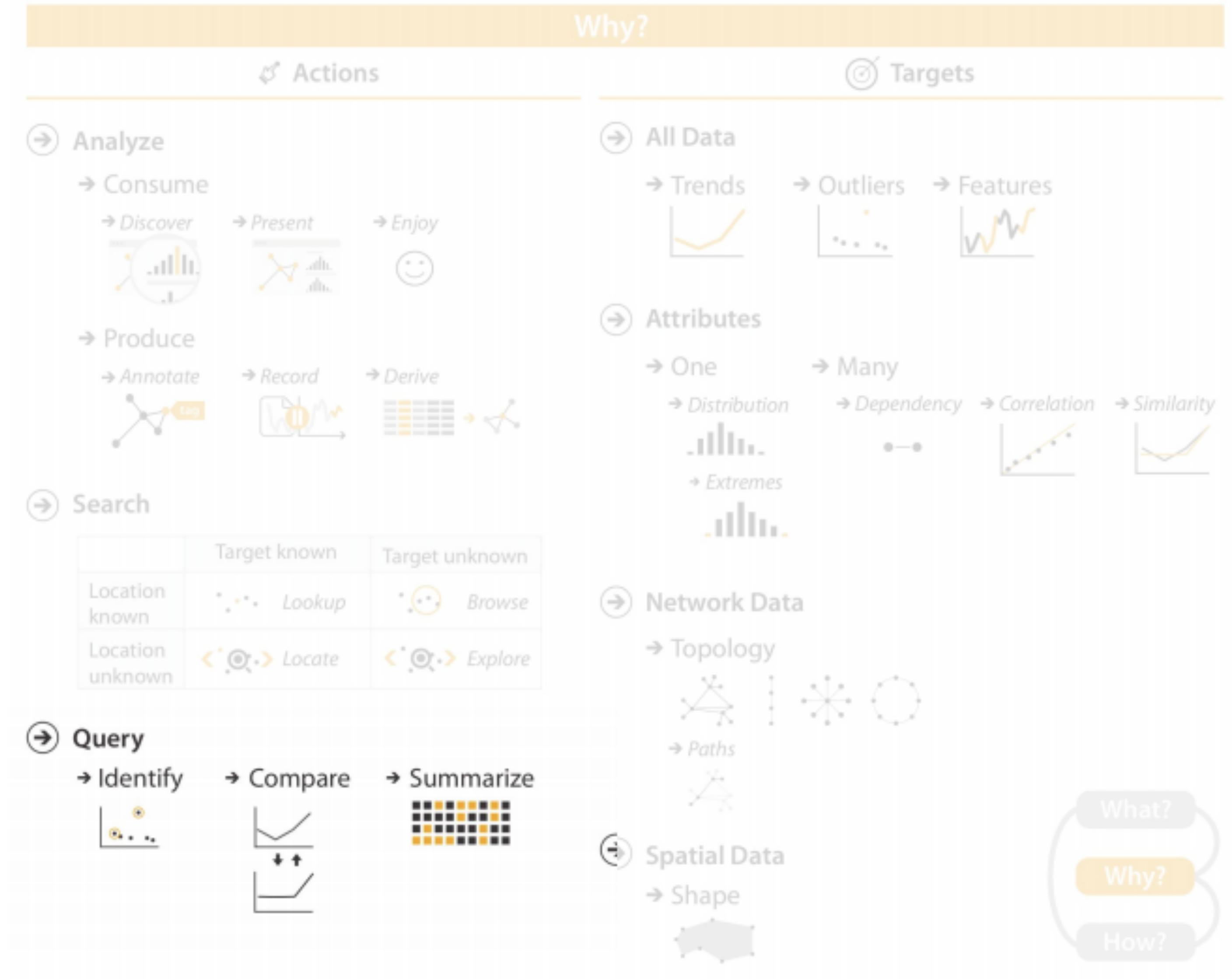
Lookup: Ahora, si ordenamos las categorías alfabéticamente y queremos buscar la cantidad de especies de las plantas, sabemos qué dato queremos y dónde encontrarlo.



Nivel bajo: Query

Cuando el target de la búsqueda ya fue identificado, ahora se puede consultar en tres niveles:

- Identificar**
- Comparar**
- Resumir.**



Nivel bajo: Query

Ellos se diferencian en la cantidad de información considerada.

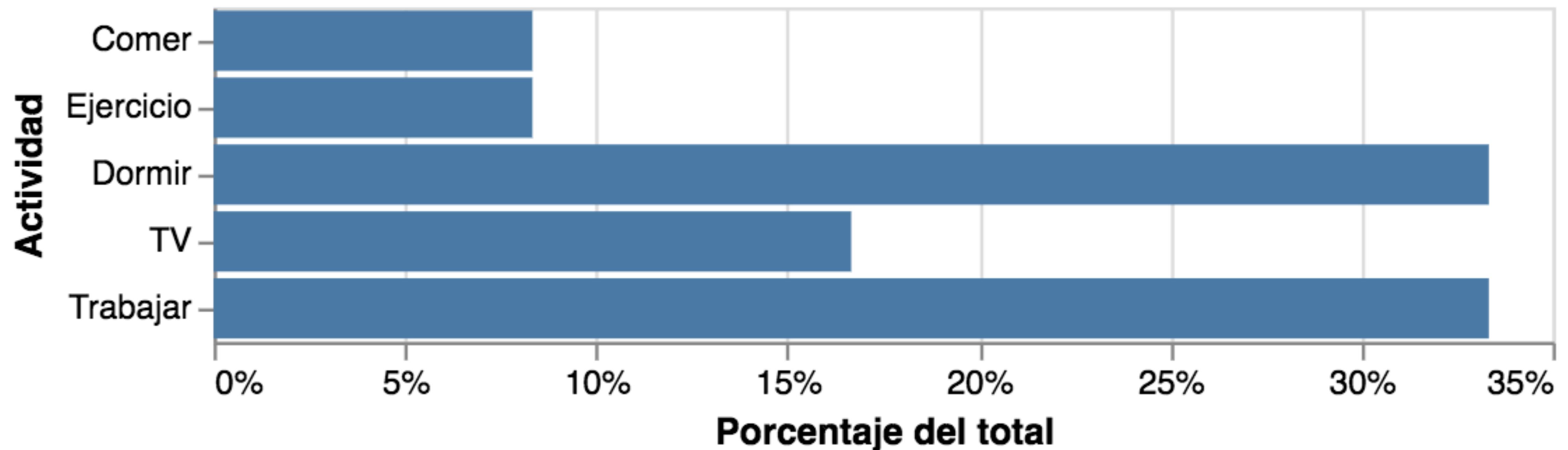
Para exemplificar, tenemos el caso del resultado de las elecciones:

Identificar: se escoge una única comuna para mostrar información.

Comparar: se eligen dos o más comunas para realizar comparaciones.

Resumir: se recogen todas las comunas para tener un resultado global.

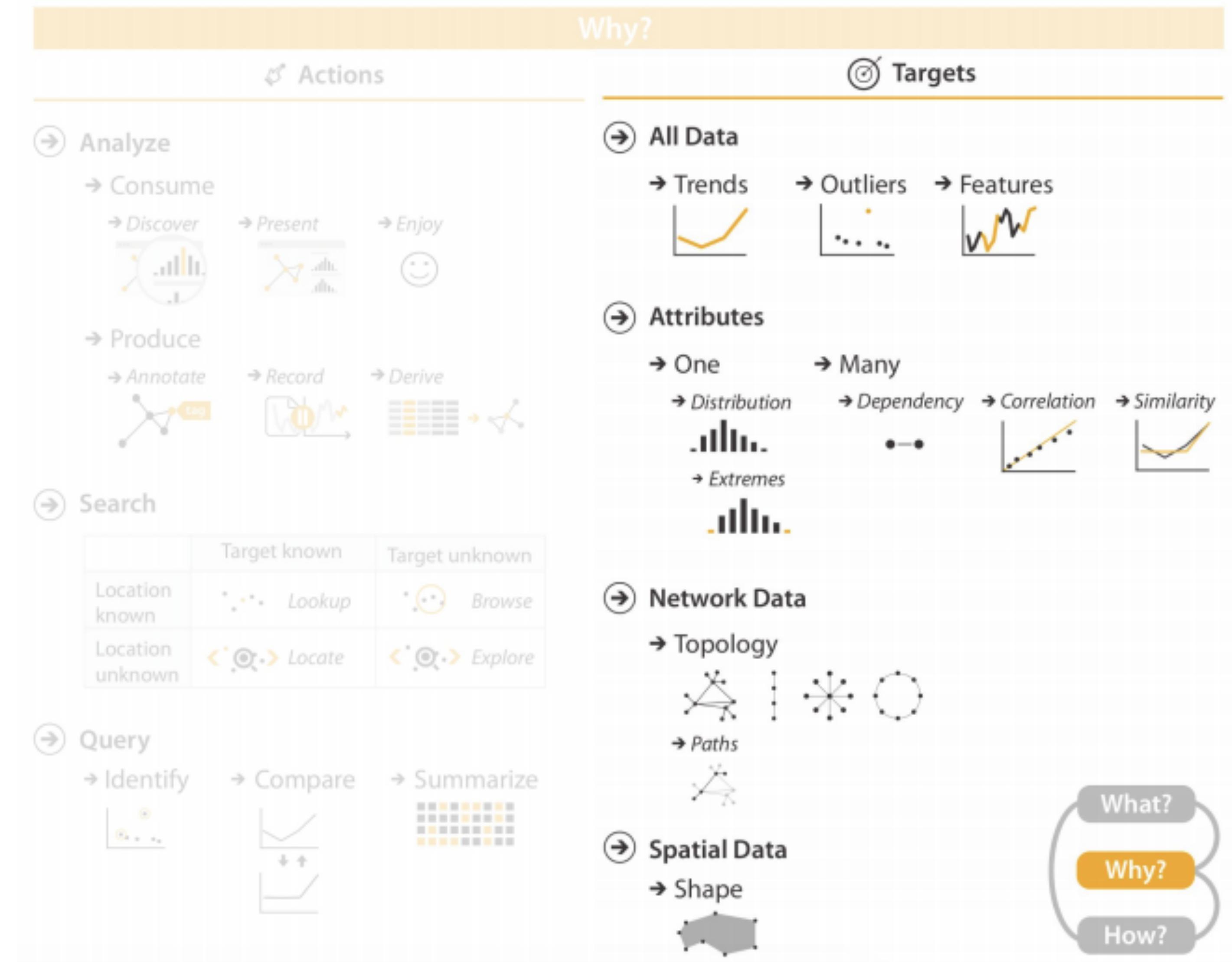
Nivel bajo: Query



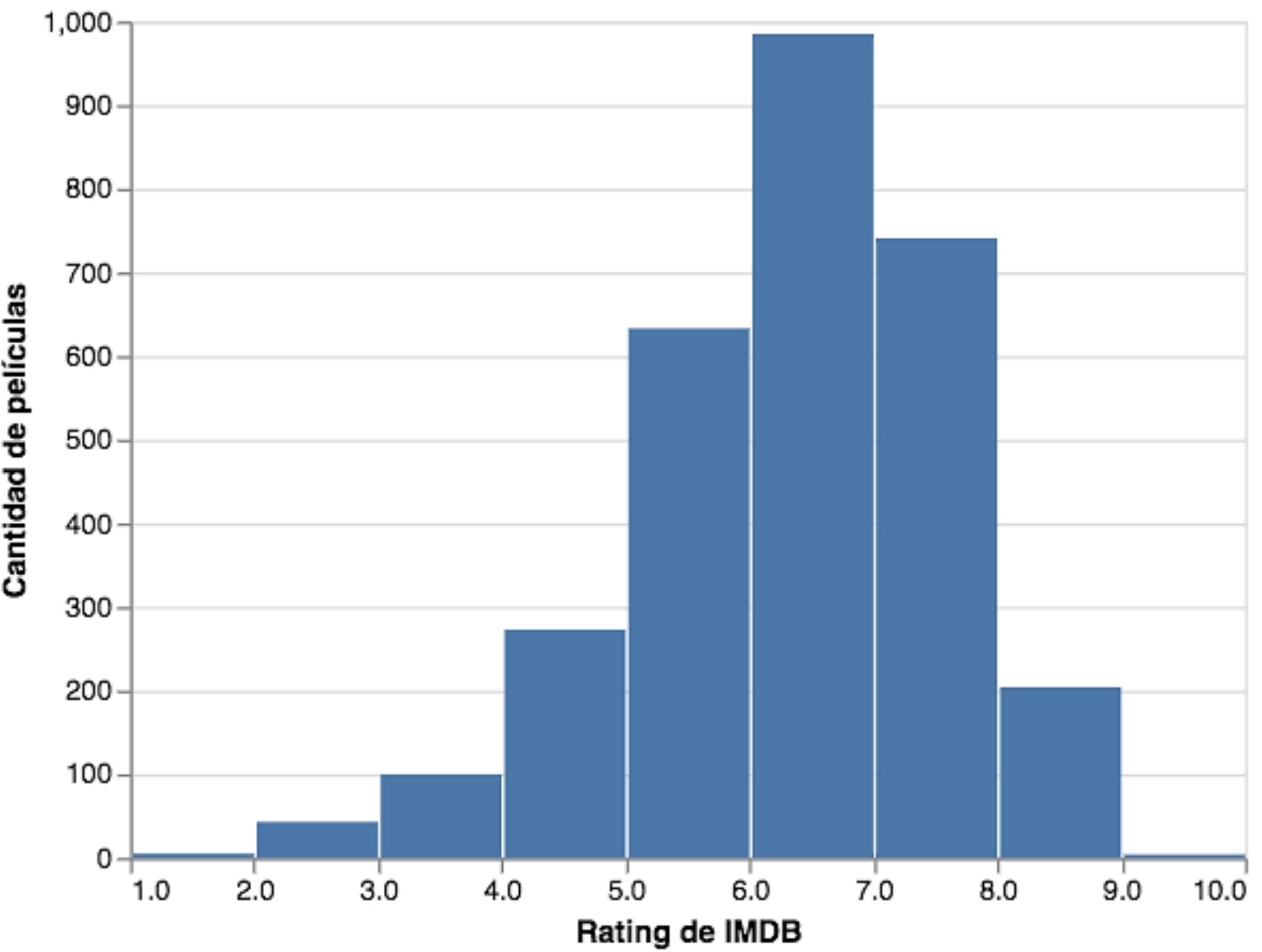
https://altair-viz.github.io/gallery/percentage_of_total.html

Targets (objetivos)

Los *targets* se refieren a algún **aspecto del dato** que queremos visualizar. Como puede ser la tendencia de los datos, la distribución de un atributo, etc.



Targets (objetivos) I



Targets (objetivos) III

si fuera posible. Judíos y gentiles, blancos o negros. Tenemos que ayudarnos unos a otros. Los seres humanos somos así.
todos los seres. El camino de la vida puede ser libre y hermoso, pero lo hemos perdido. La codicia ha envenenado las almas.
nosotros. Ahora mismo mi voz llega a millones de seres en todo el mundo, a millones de hombres desesperados, mujeres y niños.
a la felicidad. ¡Soldados, en nombre de la democracia, debemos unirnos todos!

y encarcelar a gentes inocentes. A los que puedan oírme, les digo: no desesperéis. La desdicha que padecemos no es más que la pasajera codicia y la amargura de hombres que temen seguir el camino del progreso humano.
hombres trabajo. Y dé a la juventud un futuro. Y a la vejez, seguridad. Con la promesa de esas cosas, las fieras alcanzaron el poder.
demás, no hacerlos desgraciados. No queremos odiar ni despreciar a nadie. En este mundo hay sitio para todos.
que puedan oírme, les digo: no desesperéis. La desdicha que padecemos no es más que la pasajera codicia y la amargura de hombres que temen seguir el camino del progreso humano.
los juventud un futuro. Y a la vejez, seguridad. Con la promesa de esas cosas, las fieras alcanzaron el poder.
vejez, seguridad. Con la promesa de esas cosas, las fieras alcanzaron el poder. Pero mintieron. No han cumplido sus promesas ni nunca las cumplirán.
felicidad. ¡Soldados, en nombre de la democracia, debemos unirnos todos!

, sino ayudar a todos si fuera posible. Judíos y gentiles, blancos o negros. Tenemos que ayudarnos unos a otros.
nadie . En este mundo hay sitio para todos. La Tierra es rica y puede alimentar a todos los seres. El camino de la vida puede ser libre y hermoso, pero lo hemos perdido.

seres en todo el mundo, a millones de hombres desesperados, mujeres y niños. Víctimas de un sistema que hace torturar a los hombres y encarcelar a gentes inocentes.
millones de hombres desesperados, mujeres y niños. Víctimas de un sistema que hace torturar a los hombres y encarcelar a gentes inocentes.

hombres! que en realidad os desprecian, os esclavizan, reglamentan vuestras vidas y os dicen lo que tenéis que hacer, que pensar y que sentir.
esos individuos inhumanos, hombres máquinas, con cerebros y corazones de máquinas. Vosotros no sois máquinas; no sois ganado.
otros. Los seres humanos somos así. Queremos hacer felices a los demás, no hacerlos desgraciados. No queremos odiar ni despreciar a nadie.
gentes inocentes. A los que puedan oírme, les digo: no desesperéis. La desdicha que padecemos no es más que la pasajera codicia y la amargura de hombres que temen seguir el camino del progreso humano.
ganado. Y como a carne de cañón. No os entreguéis a esos individuos inhumanos, hombres máquinas, con cerebros y corazones de máquinas.
carne de cañón. No os entreguéis a esos individuos inhumanos, hombres máquinas, con cerebros y corazones de máquinas. Vosotros no sois máquinas; no sois ganado.
luchar para libertar al mundo. Para derribar barreras nacionales. Para eliminar la ambición, el odio y la intolerancia. Luchemos por el mundo de la razón.

Contenidos

Formalizando el proceso de diseño - *Framework* de Tamara Munzner

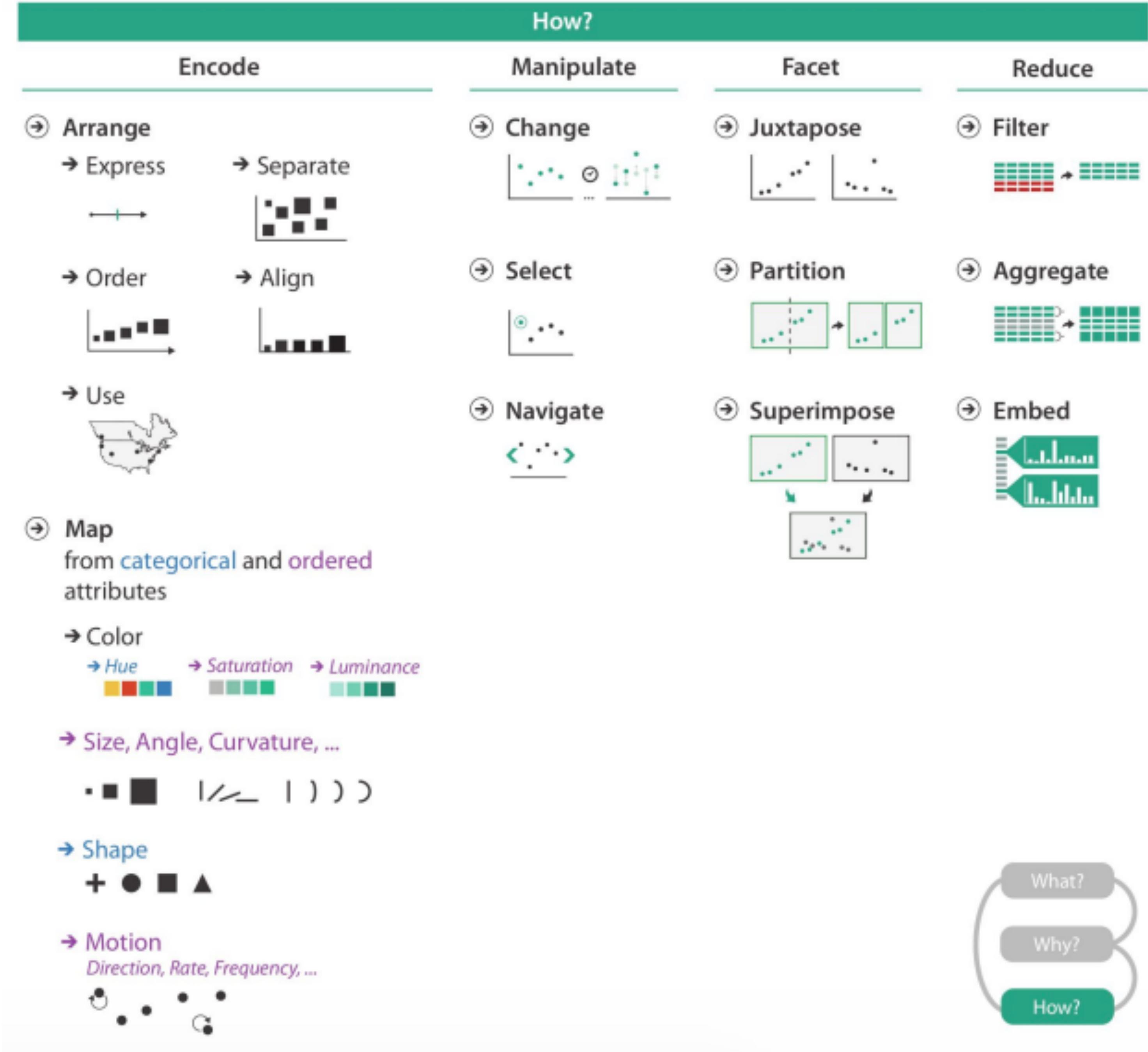
Parte 2 de 3: Why?

Parte 3 de 3: How?

Control 1 - No se preguntará lo visto en esta clase.

How

Cómo un *visualization idiom* puede ser construido
a partir de un conjunto de
elecciones de diseño.

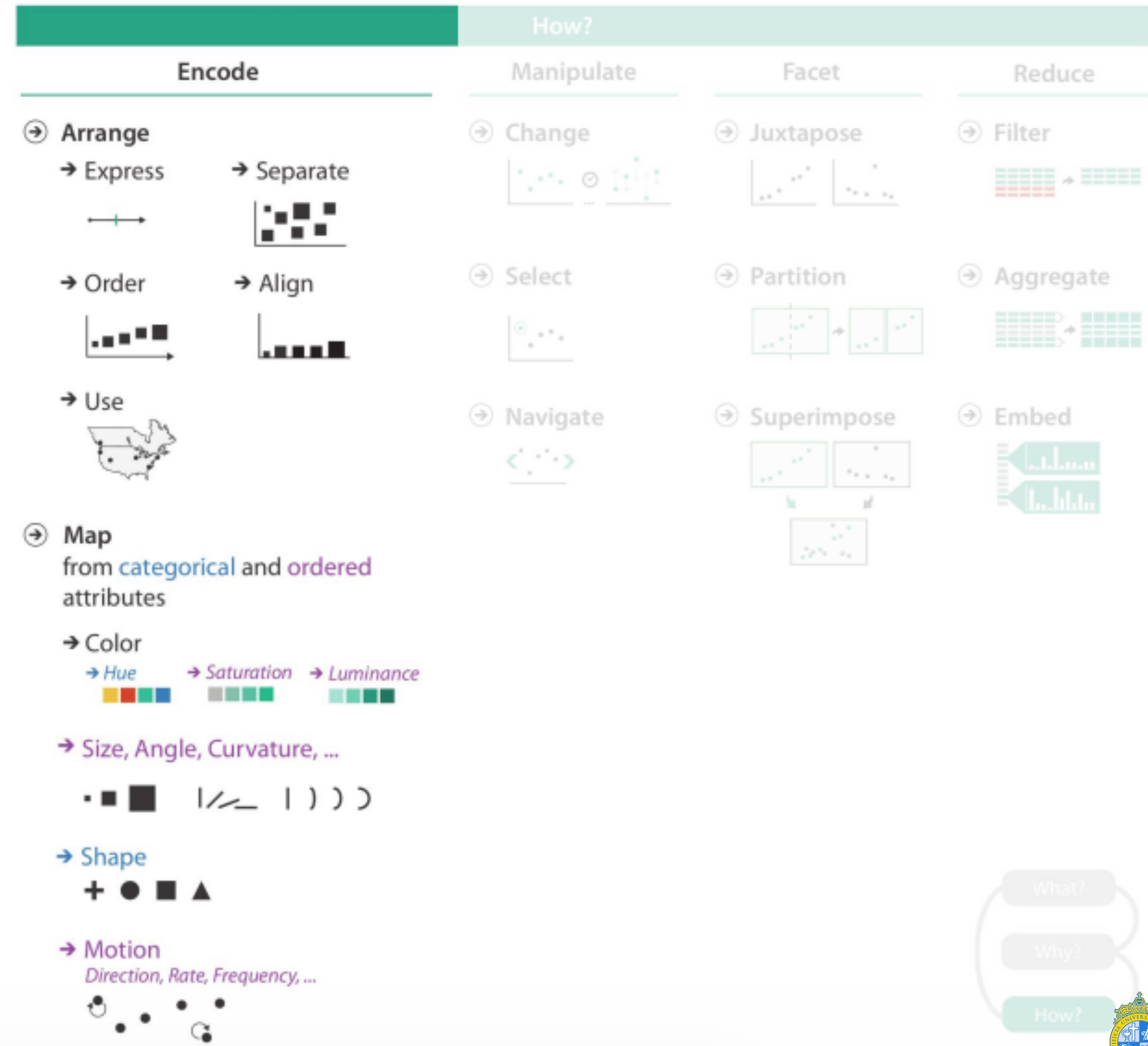


How

Las decisiones se pueden separar en dos áreas.

Visual encoding

Establece una traducción entre datos y elementos visuales

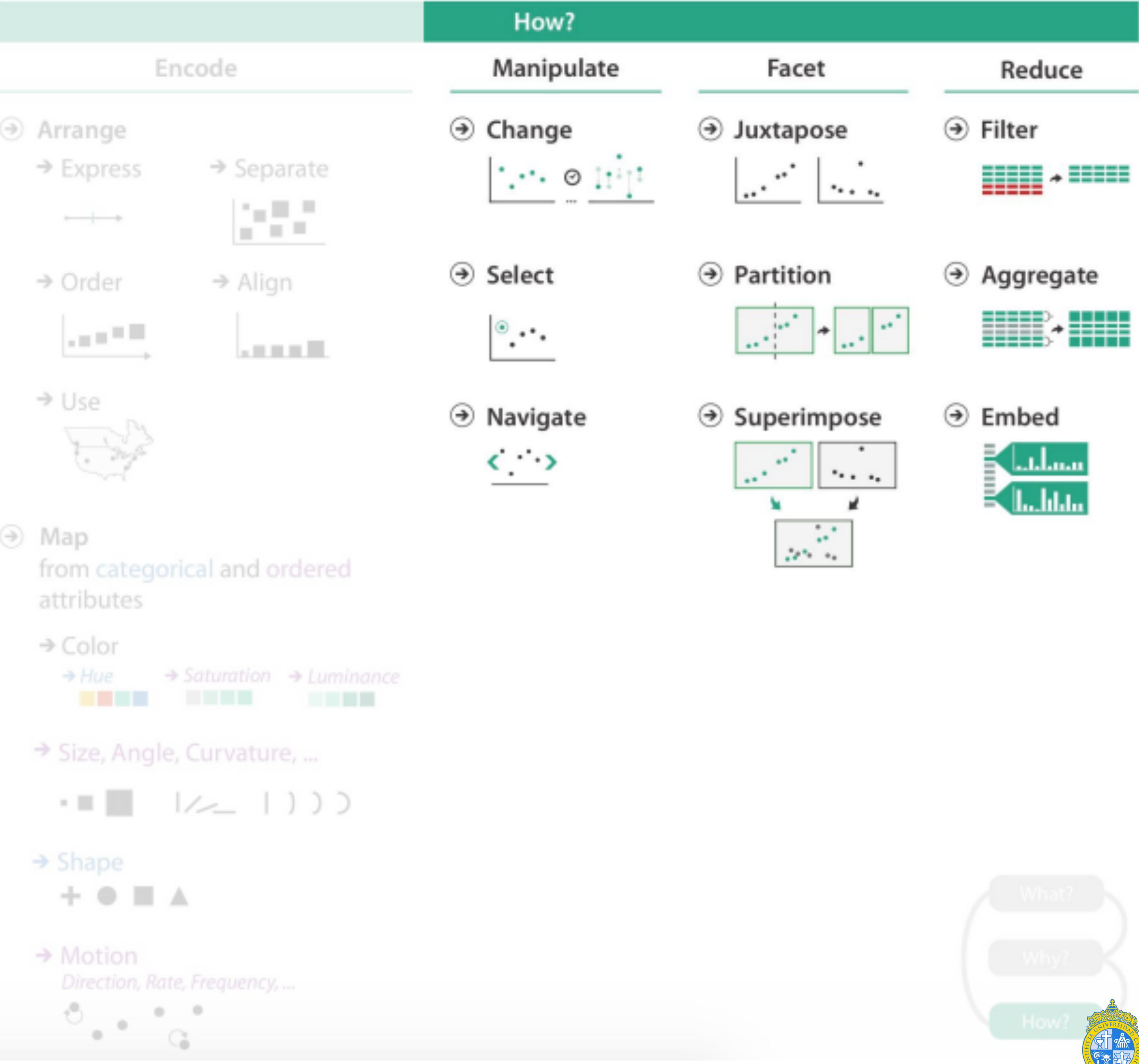


How

Las decisiones se pueden separar en dos áreas.

Interactions

Establece cómo el usuario va a interactuar con los elementos visuales

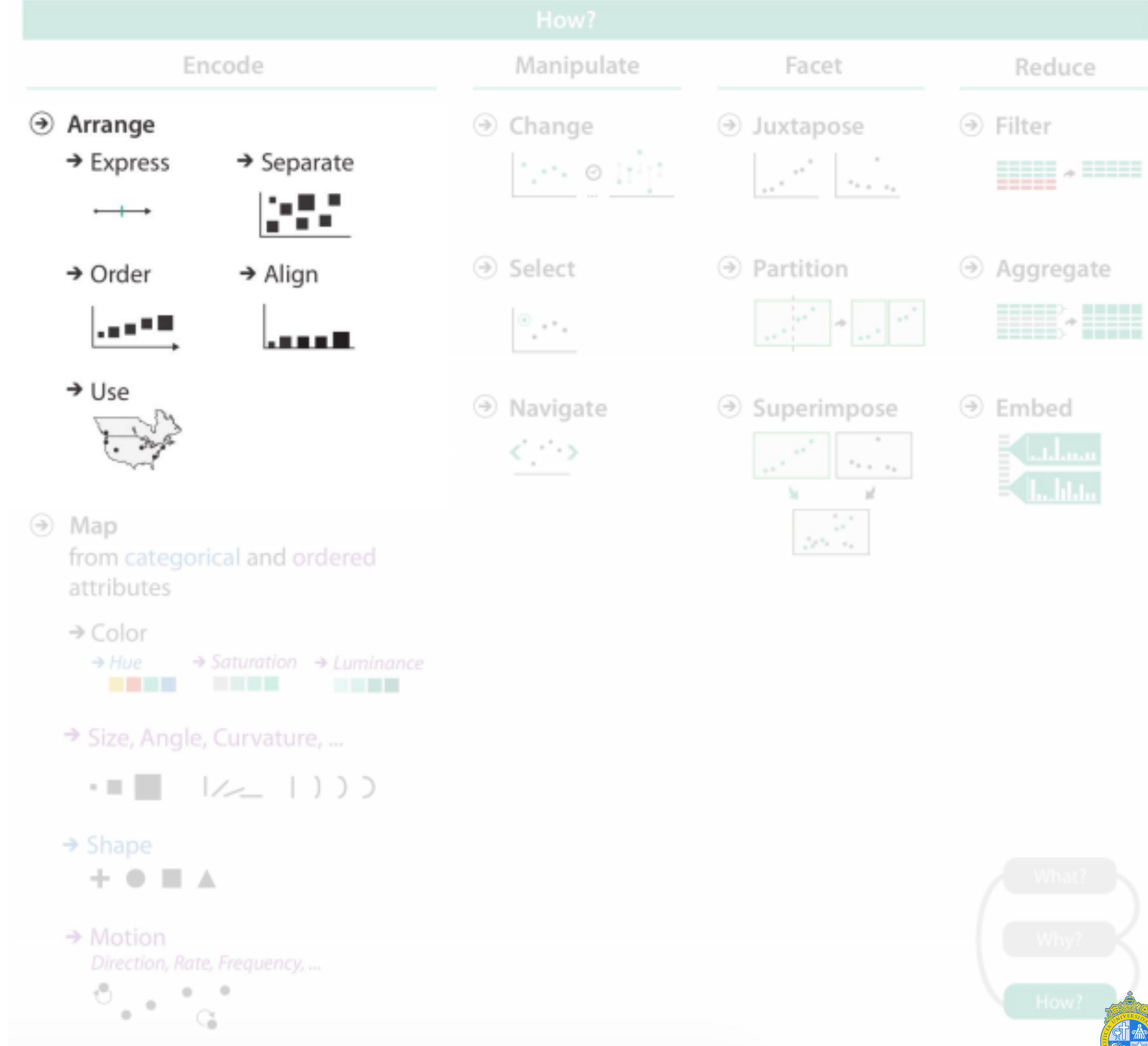


Arrange

Con **arrange**, buscamos saber cómo **organizar los datos** en el espacio.

De todos los encodings, es el más crucial porque el uso del espacio domina el modelo mental que tiene el usuario de los datos.

Queremos saber cómo expresar los valores, cómo separar, ordenar y alinear las regiones, y cómo usar un espacio dado (e.g. dataset geográfico)



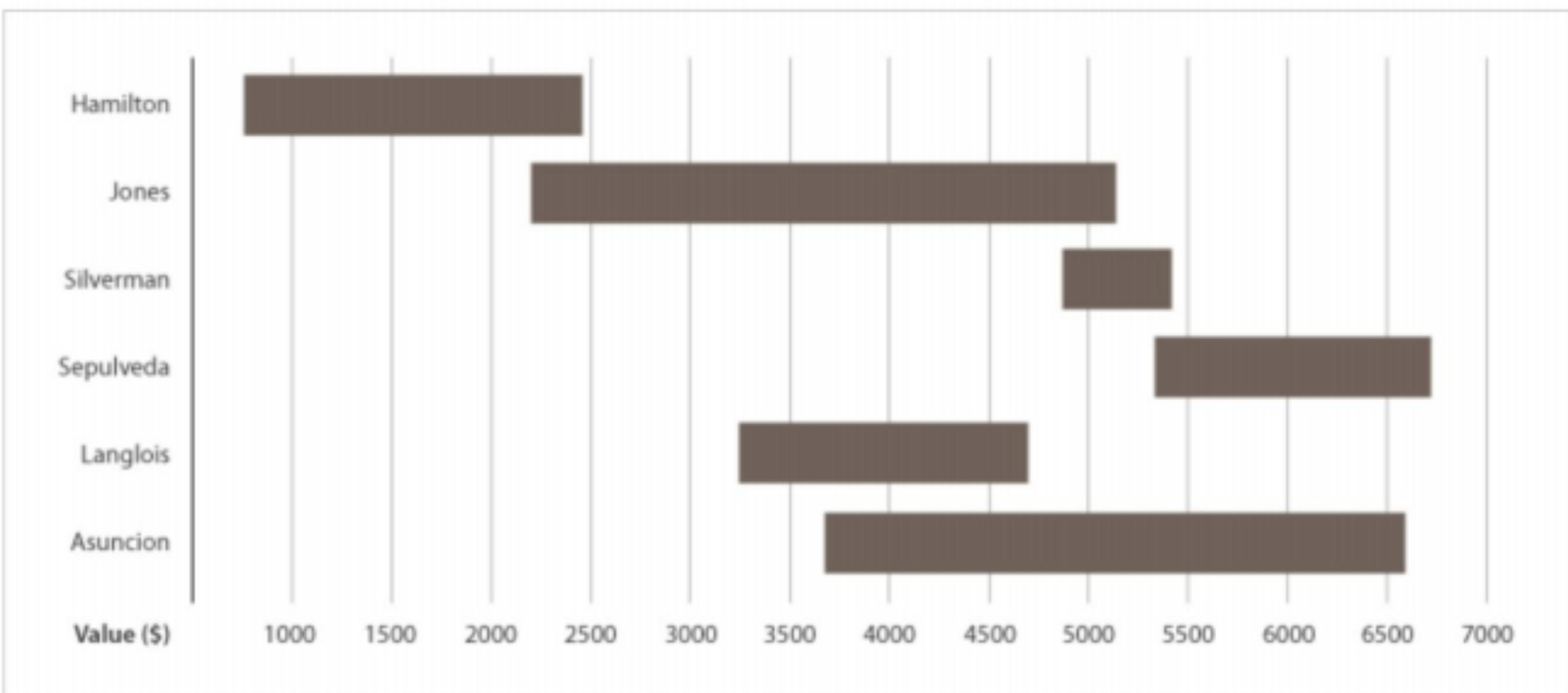
Arrange

Las posiciones de las barras

expresan el valor mínimo y

máximo

[https://datavizcatalogue.com/methods/
span_chart.html](https://datavizcatalogue.com/methods/span_chart.html)

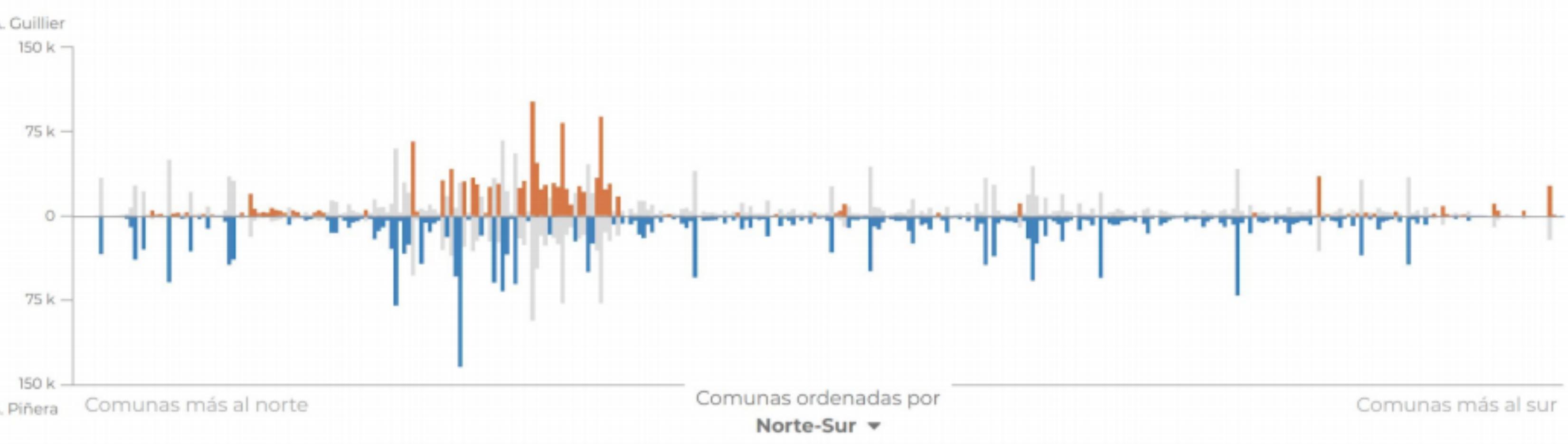


Arrange

- Se **separan** los votos de los candidatos orientando las barras hacia arriba y hacia abajo

Se **alinean** las barras desde el centro de la visualización

Se **ordenan** las comunas de acuerdo a la latitud

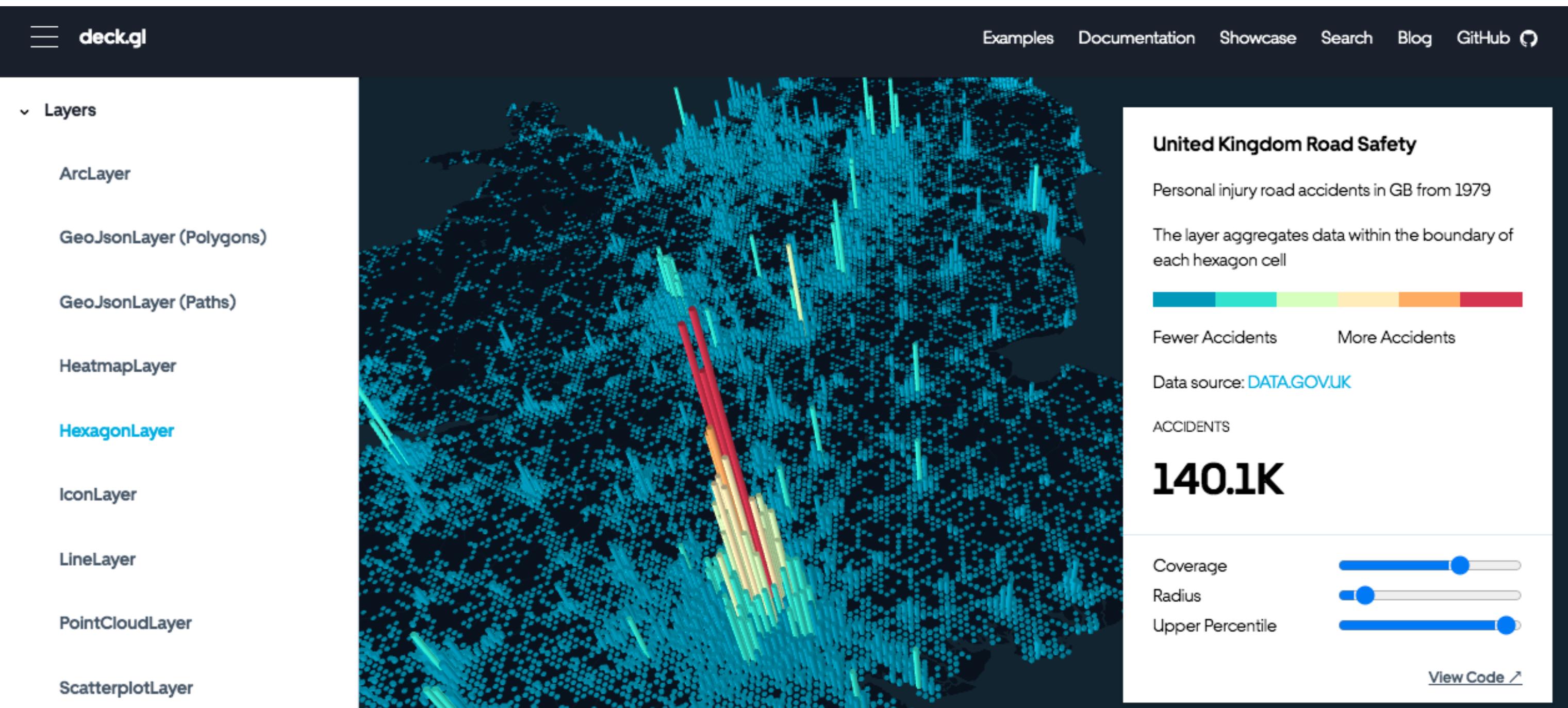


https://www.decidechile.cl/index.html#/ev/2017/territory_panorama

Arrange

Usamos la latitud y longitud
para ubicar la barra en el
mapa de Reino Unido

<https://deck.gl/examples/hexagon-layer/>

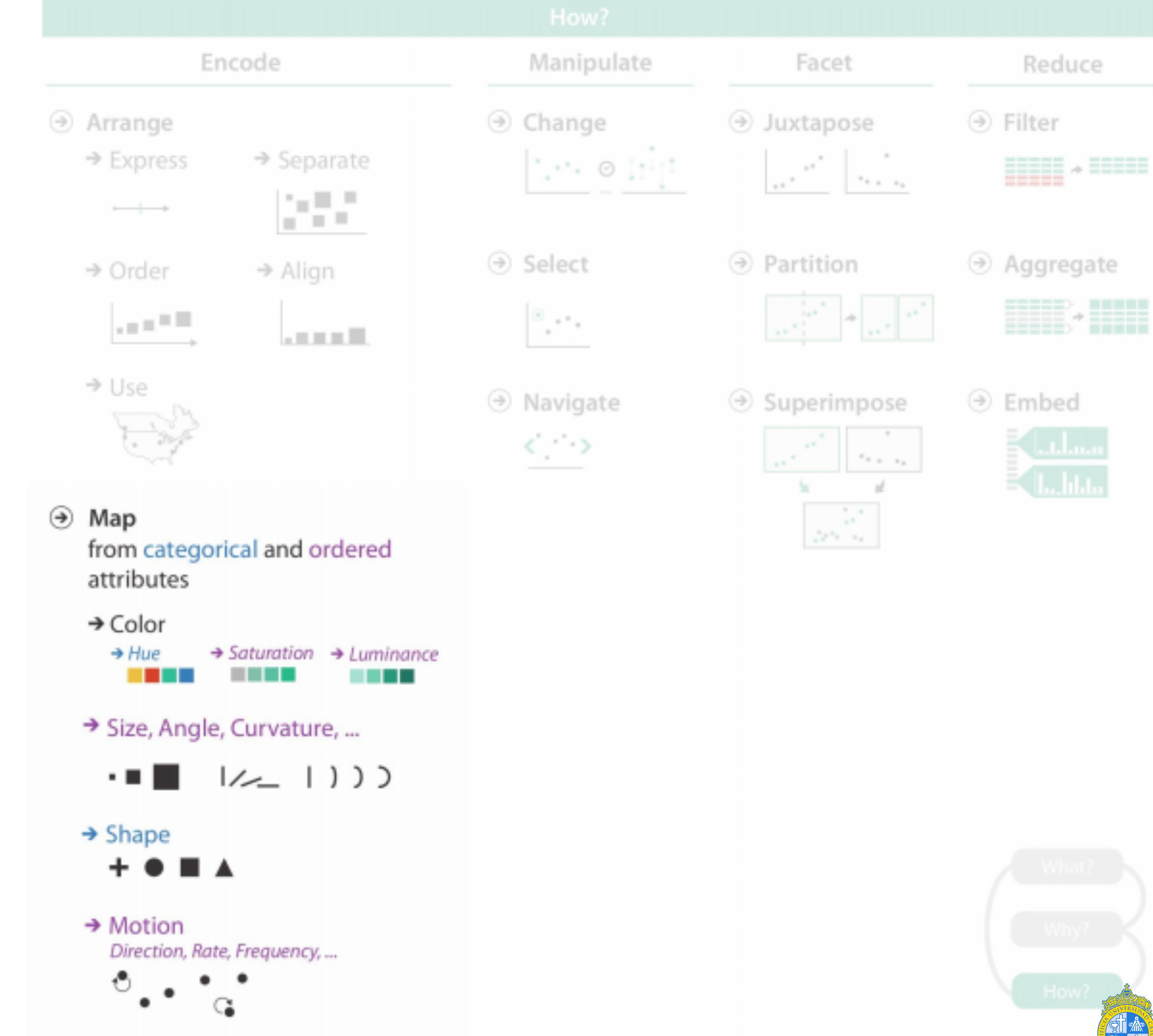


Map

Con map, buscamos aprovechar los canales visuales no-espaciales.

Podemos trabajar con color (hue, saturation, luminance), tamaño, ángulo, curvatura, formas.

Pero también con atributos dinámicos: dirección, frecuencia, tasa de aparición.



Manipulate

Describen acciones que puede realizar el usuario con los elementos de la visualización (marcas).

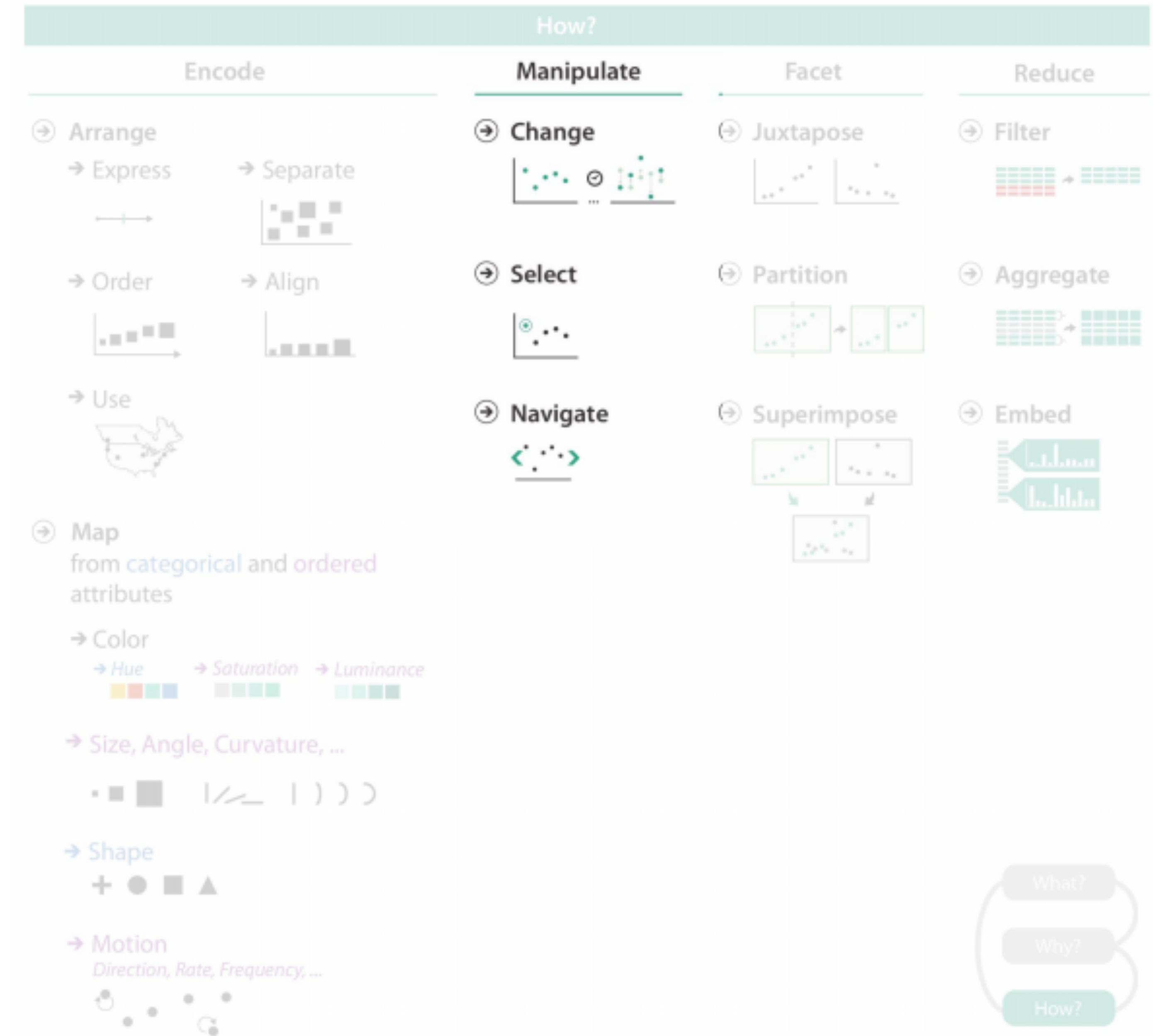
Change

Cambiar la codificación de los datos.

Cambiar la disposición de los datos (por ejemplo, ordenar).

Cambiar el/los canales utilizados (Pasar de color a tamaño).

Modificar el canal utilizado (pasar de una escala secuencial de 3 colores a 5 colores).



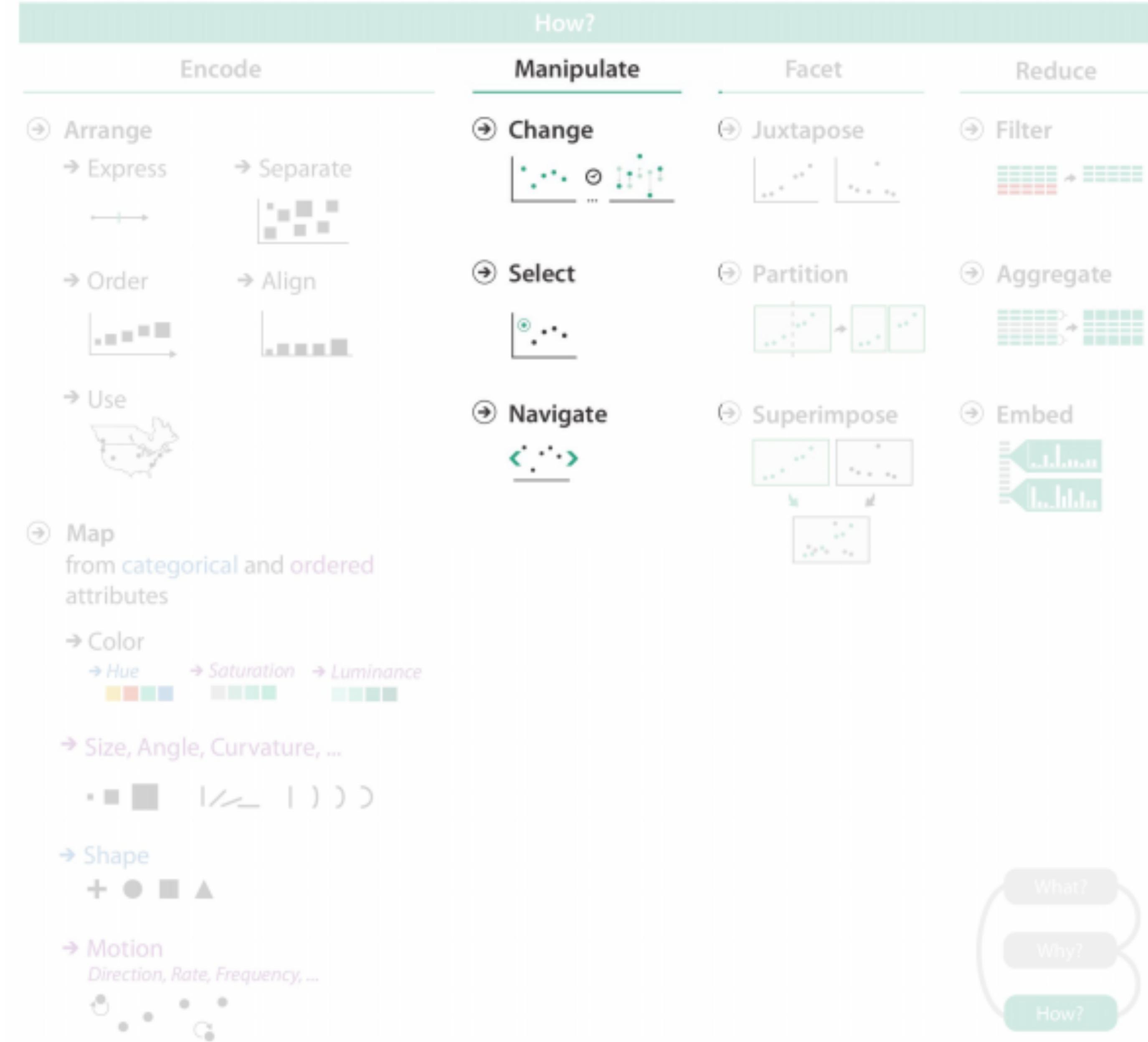
Manipulate

Describen acciones que puede realizar el usuario con los elementos de la visualización (marcas).

Select

Seleccionar/enfatizar un elemento.

Seleccionar el atributos a observar.



Manipulate

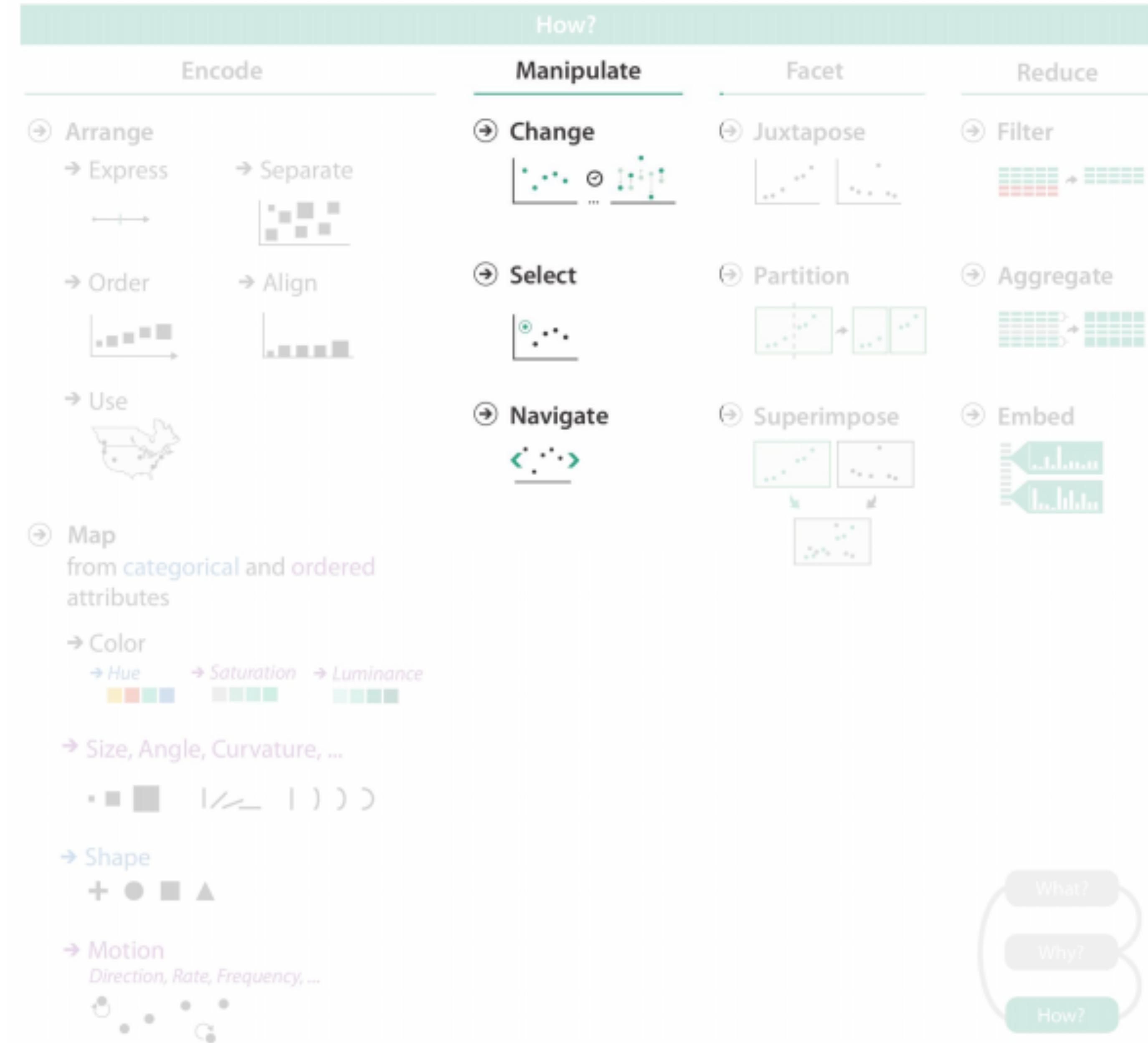
Describen acciones que puede realizar el usuario con los elementos de la visualización (marcas).

Navigate

Cambiar el *viewpoint* (punto de vista) con *panning*, *translating* y/o *zooming*.

Panning (cambiar la sección de la visualización que se ve) . Cuando estamos en un contexto de tres dimensiones, se hablará de *translating*.

Zooming (acercarse o alejarse de una zona)

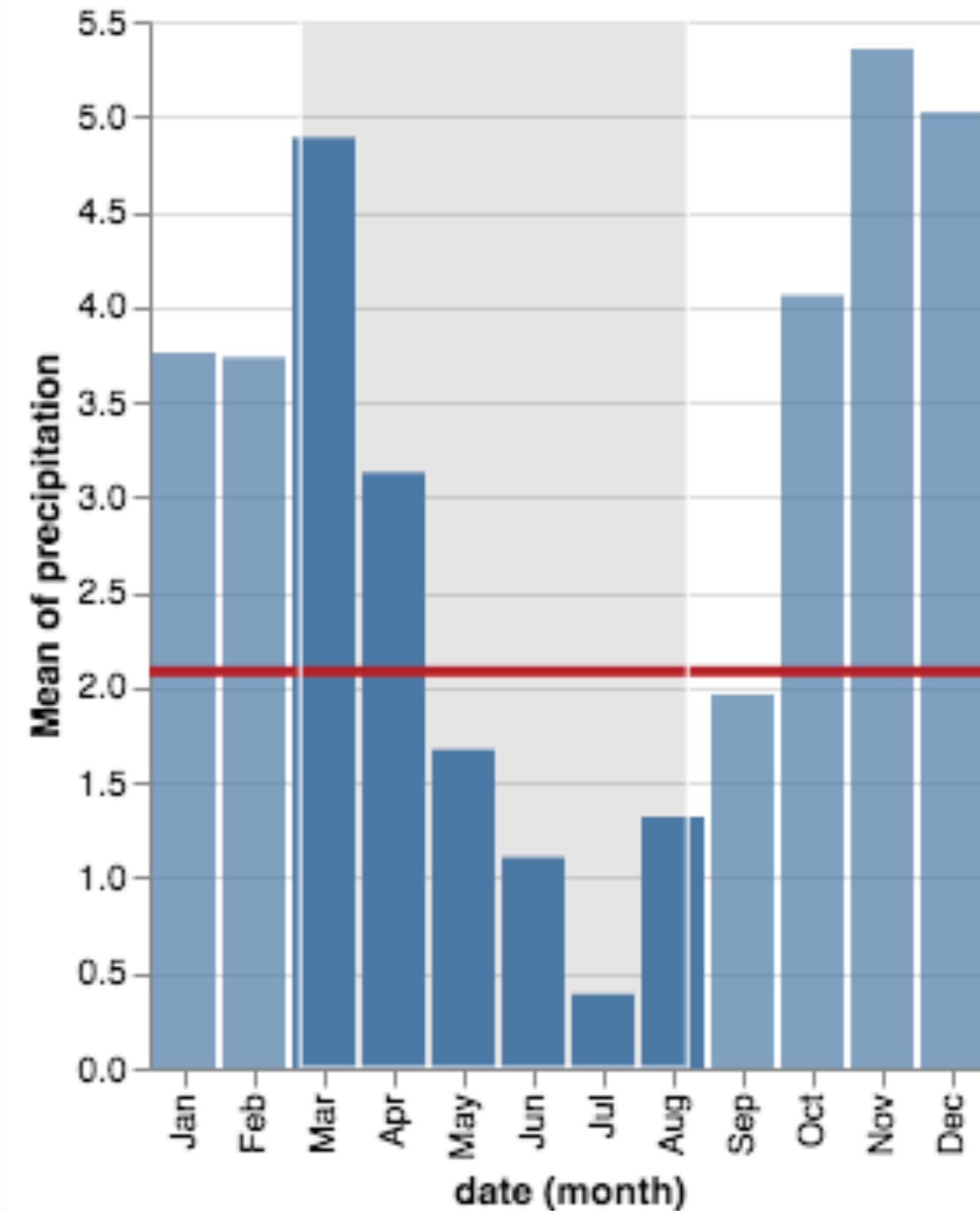


What?

Why?

How?

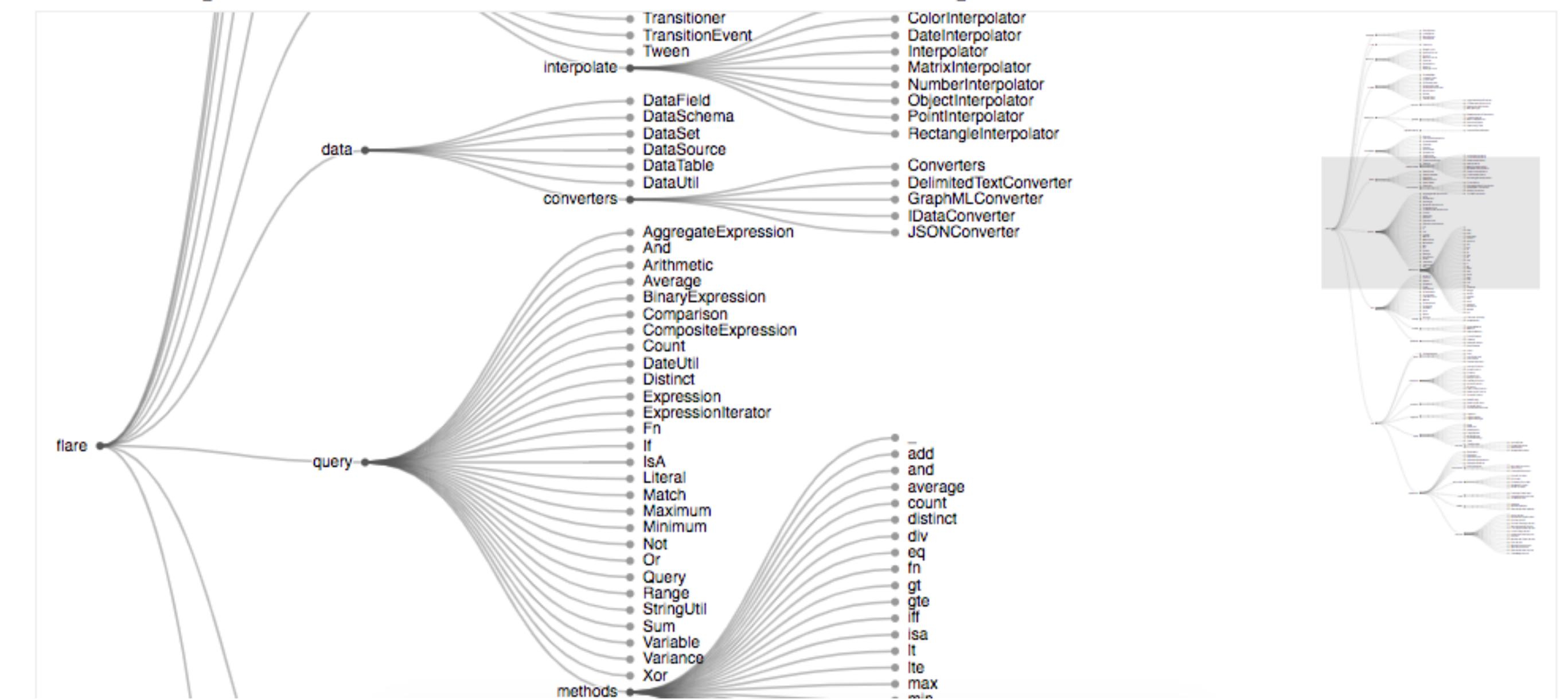
Manipulate - Ejemplos



https://altair-viz.github.io/gallery/selection_layer_bar_month.html

Manipulate - Ejemplos

Collapsible Tree with Minimap



<https://bl.ocks.org/bwswedberg/464a7dbc471ee2a94dd6278bc7d94710>

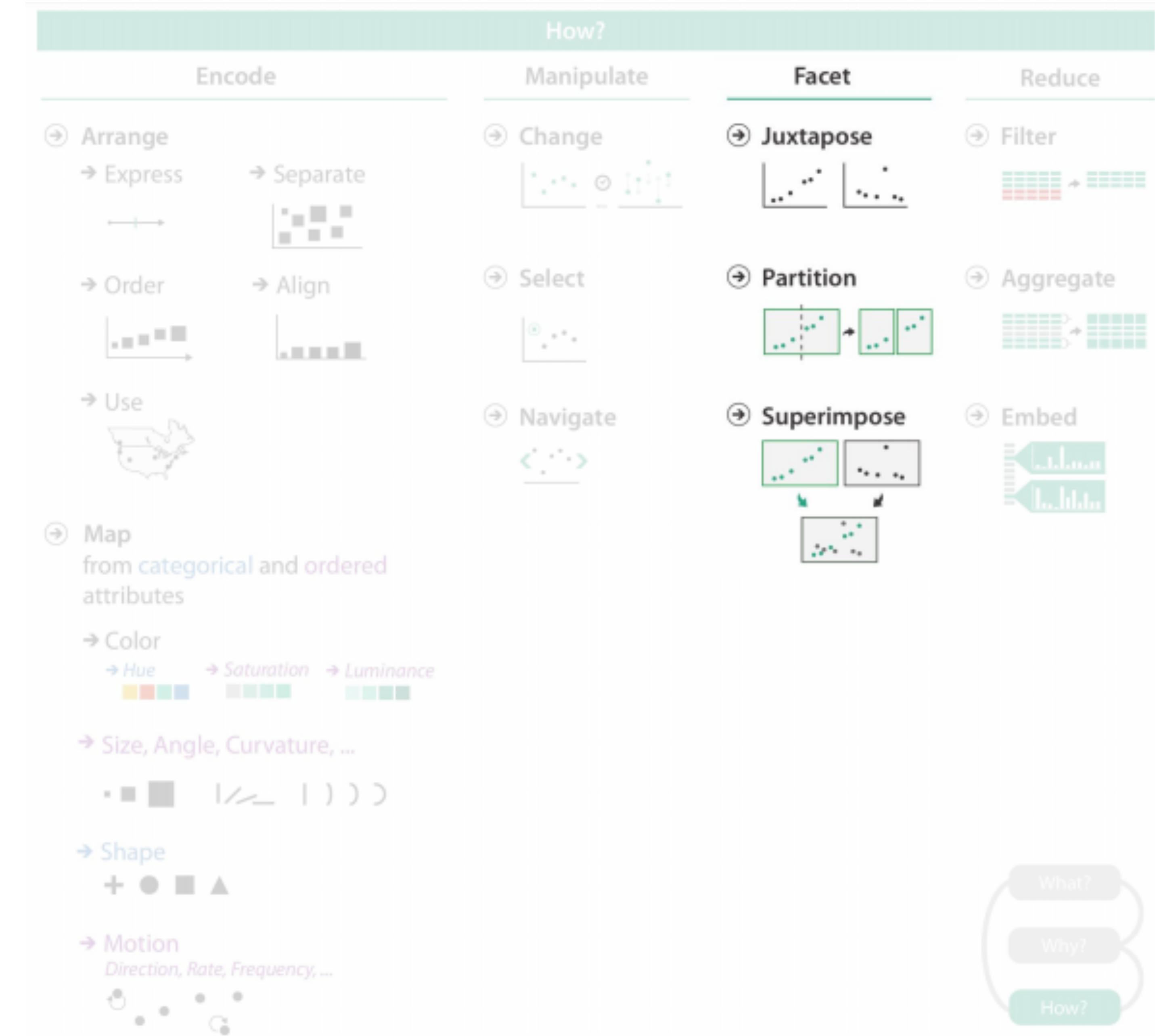
Facet

La idea de facet es mostrar diferentes ángulos de un dataset, dividiendo la visualización en diferentes vistas.

juxtapose

Una o más vistas simultáneas. Generalmente coordinadas entre ellas.

Se debe elegir cómo coordinar las vistas entre ellas, cuántos datos/atributos compartir, qué canales utilizar, etc.



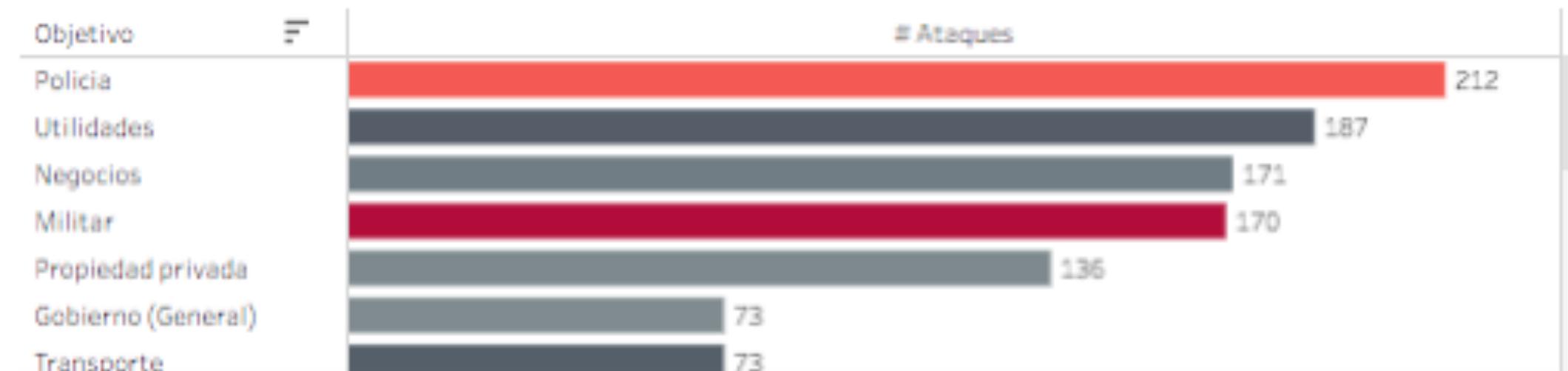
¿Cómo afecta el terrorismo en el Mundo?

Región Década

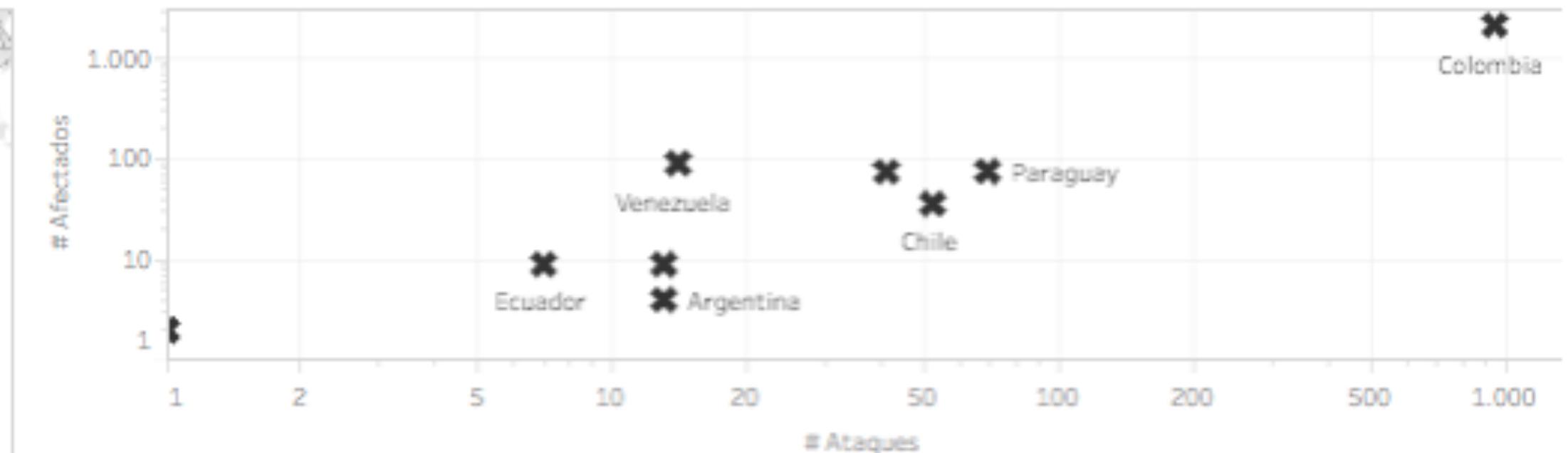
Cantidad de ataques en América del Sur (2010-2016)



Cantidad de ataques por tipo de objetivo en América del Sur (2010-2016)

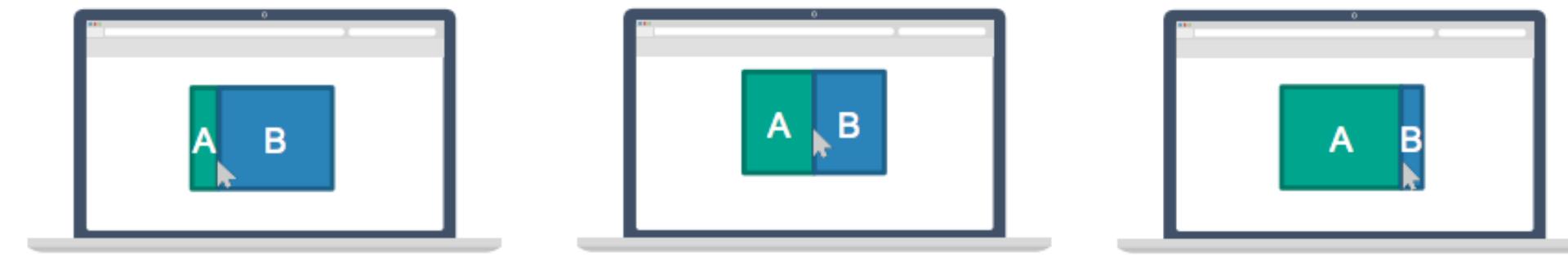


Relación de afectados por ataque en América del Sur (2010-2016)



Evolución en el tiempo de la cantidad de ataques, por categoría de grupo terrorista en América del Sur (2010-2016)



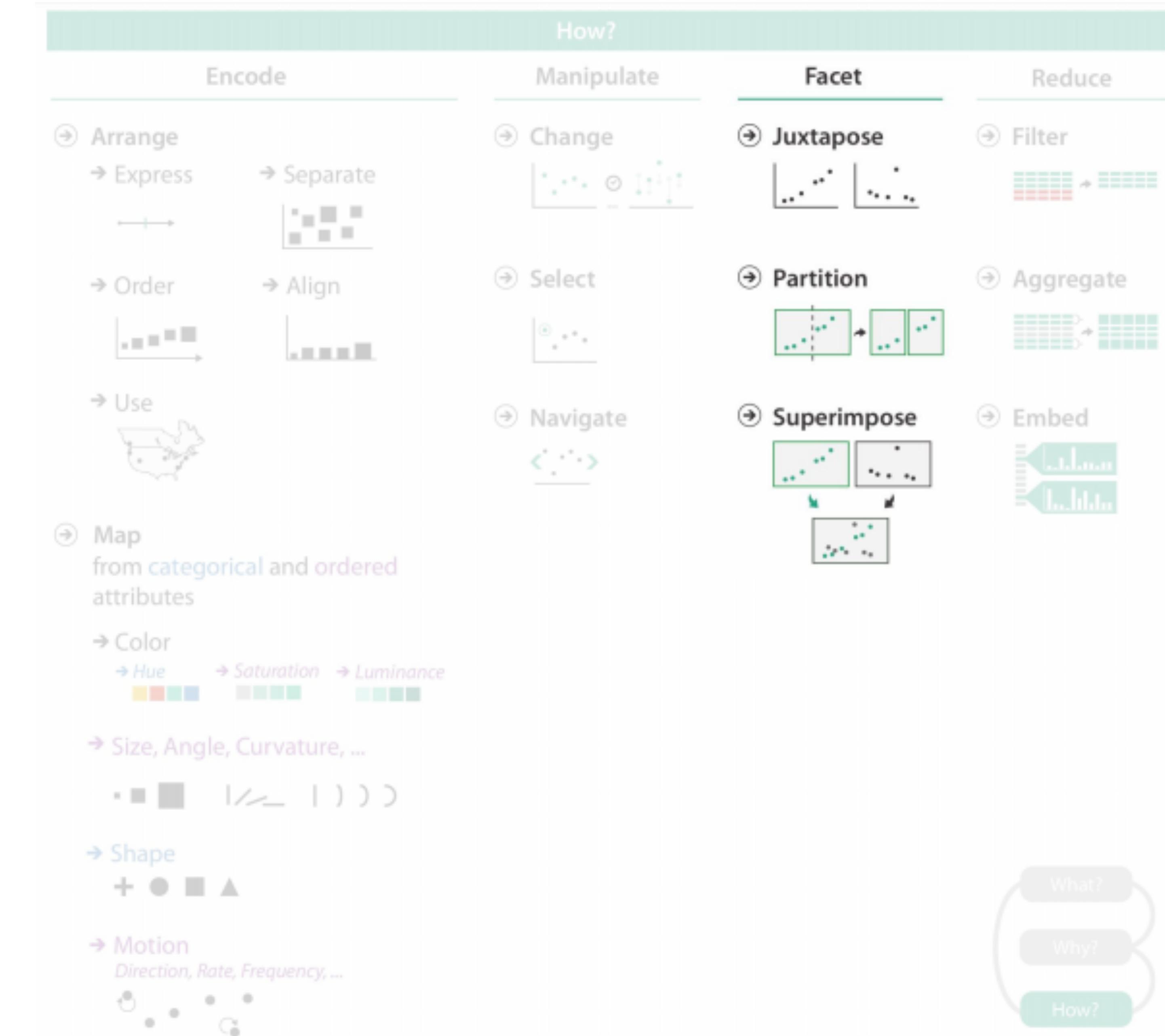


Facet

La idea de facet es mostrar diferentes ángulos de un dataset, dividiendo la visualización en diferentes vistas.

partition

Se debe elegir cuántas regiones utilizar, cómo dividir los datos entre ellas, o el orden de los atributos al utilizar.



What?
Why?
How?

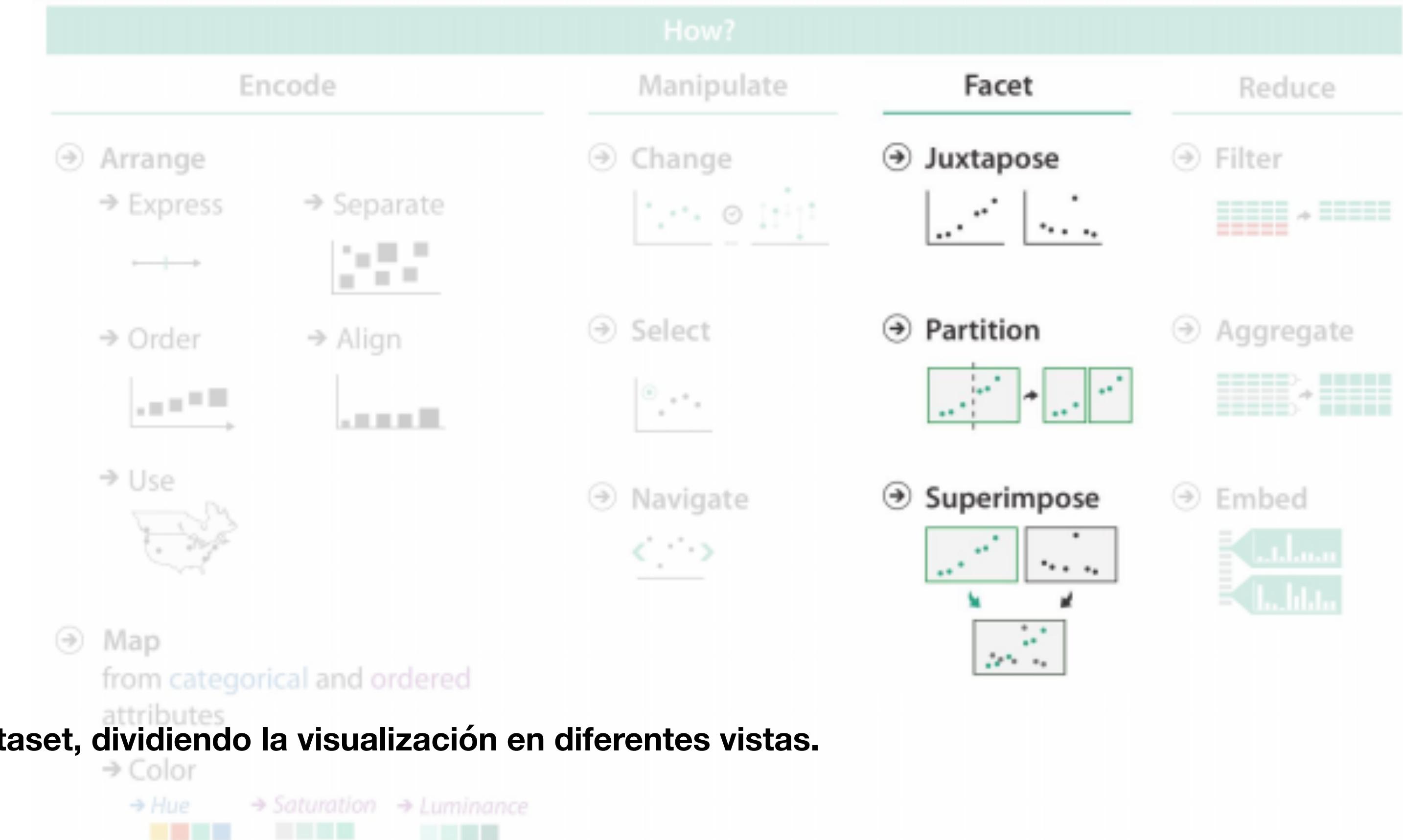
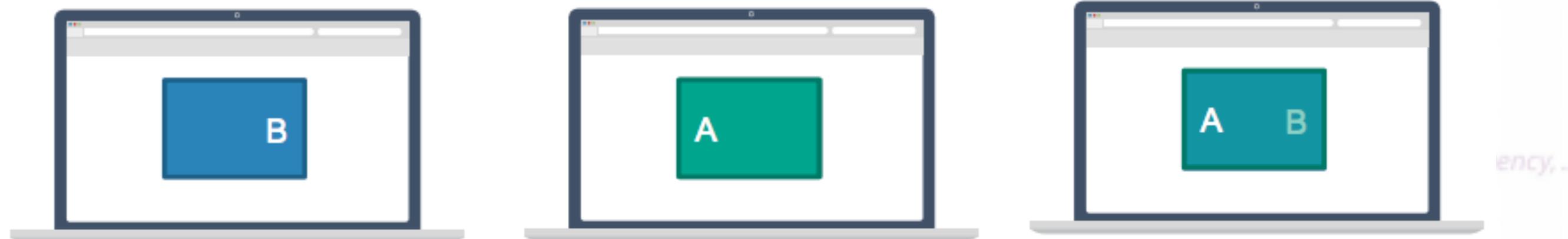
Facet

La idea de facet es mostrar diferentes ángulos de un dataset, dividiendo la visualización en diferentes vistas.

Superimpose

Se sitúan N visualizaciones una encima de otra.

Se debe elegir cómo los elementos serán partitionados en las distintas capas, cuántas capas usar, etc.



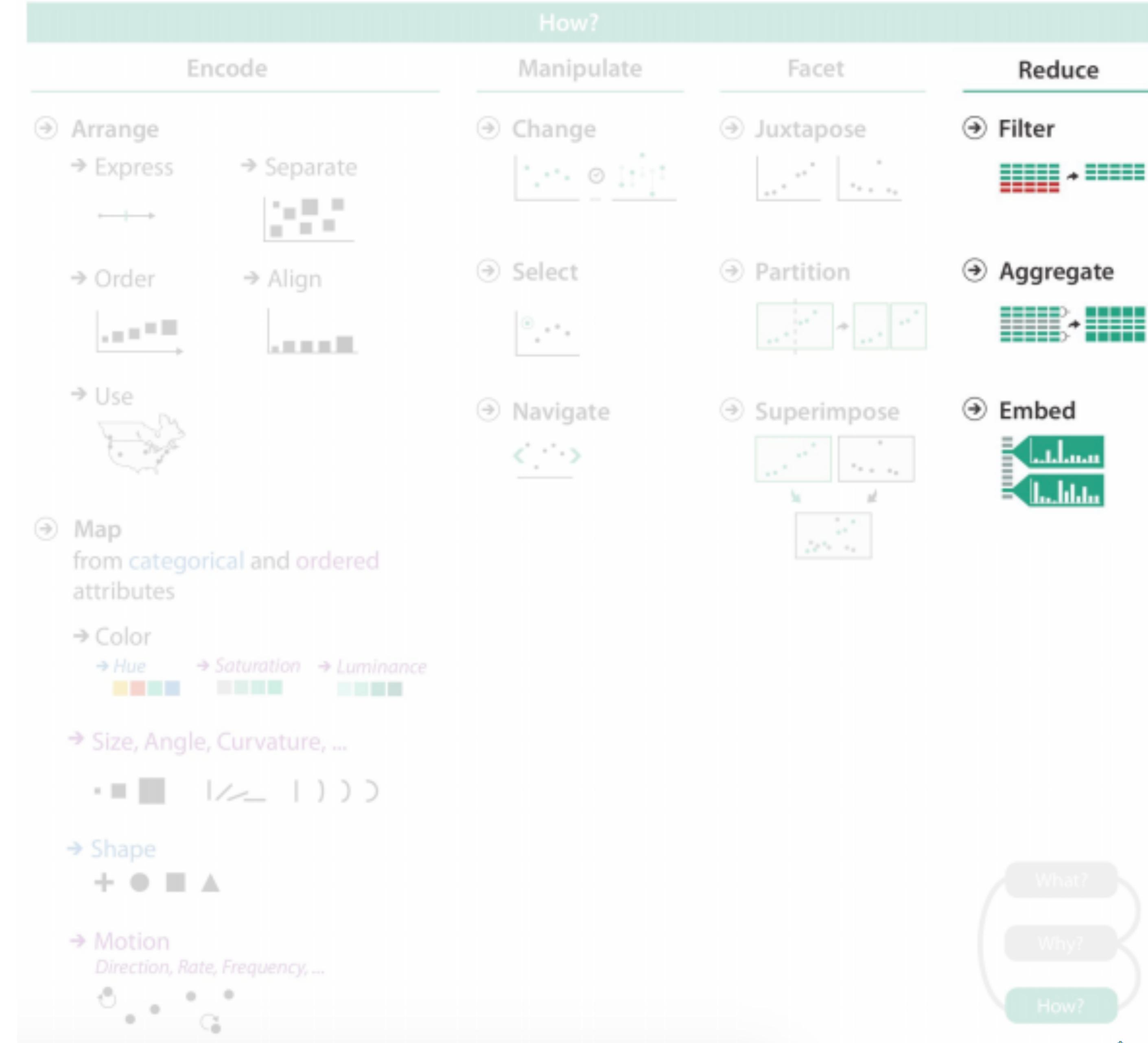
Reduce

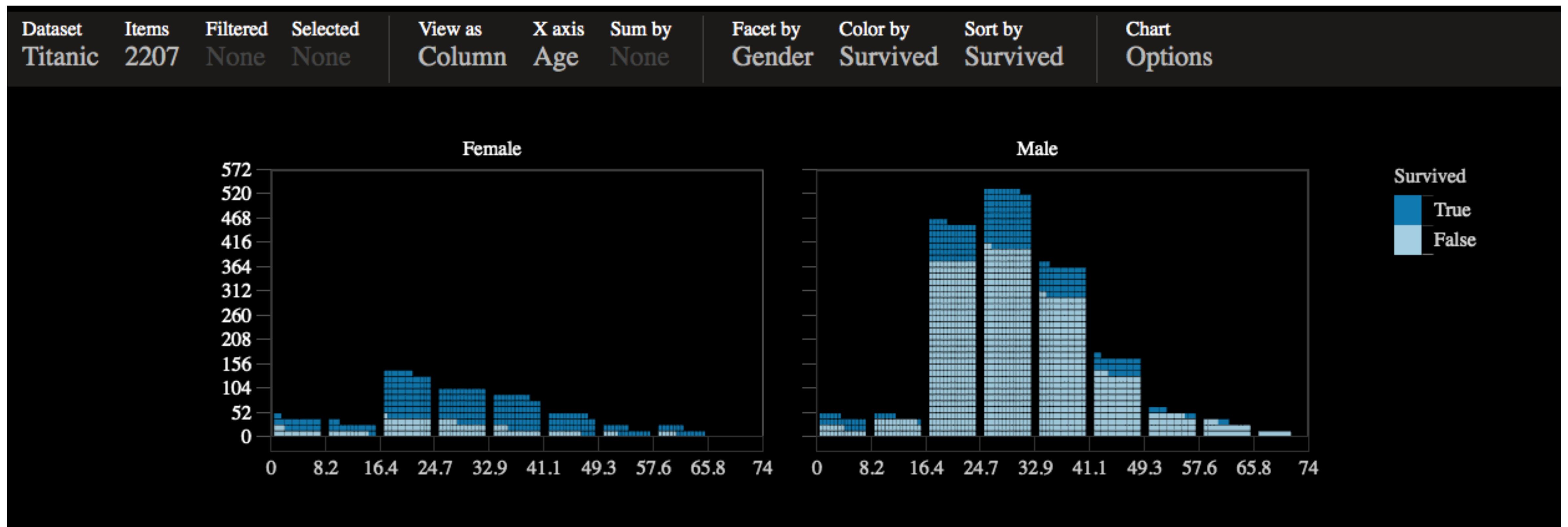
Esta familia tiene por objetivo manejar la complejidad del dataset.

Filter permite eliminar la cantidad de elementos mostrados (e.g. por uno o más rangos de interés)

Aggregate busca que un grupo de elementos sea representado por un nuevo elemento que los represente; de esta forma, se hace un *merge* (e.g. obtener el promedio)

Embeber permite reducir la cantidad de elementos mediante una sofisticada combinación de filtrado y agregación.





<https://sanddance.azurewebsites.net/BeachPartyApp/BeachPartyApp.html>

How

¿Cómo elegimos entre las opciones que tenemos disponibles?

Debemos considerar al menos: -

Percepción

Memoria

Tarea a resolver

Canales disponibles

Marcas disponibles

Interacciones entre marcas y canales

Eficiencia de canales

Algunas reglas basadas en la experiencia (Veremos la próxima clase)

Diplomado Big Data y Ciencias de datos Visualización de información

Profesor: Cristián Hernández

Ayudante: Yeison Salinas