

# Lecture II: Filtering & Aggregation and Focus+Context

DS 4200  
FALL 2022

Prof. Ab Mosca (*they/them*)  
NORTHEASTERN UNIVERSITY

# Last Class

**We:**

- Reviewed Interaction
- Reviewed Faceting

**Any Questions?**

# Today

- Filtering and Aggregation
- Focus+Context

# Visualizing Big Data



# Visualizing Big Data

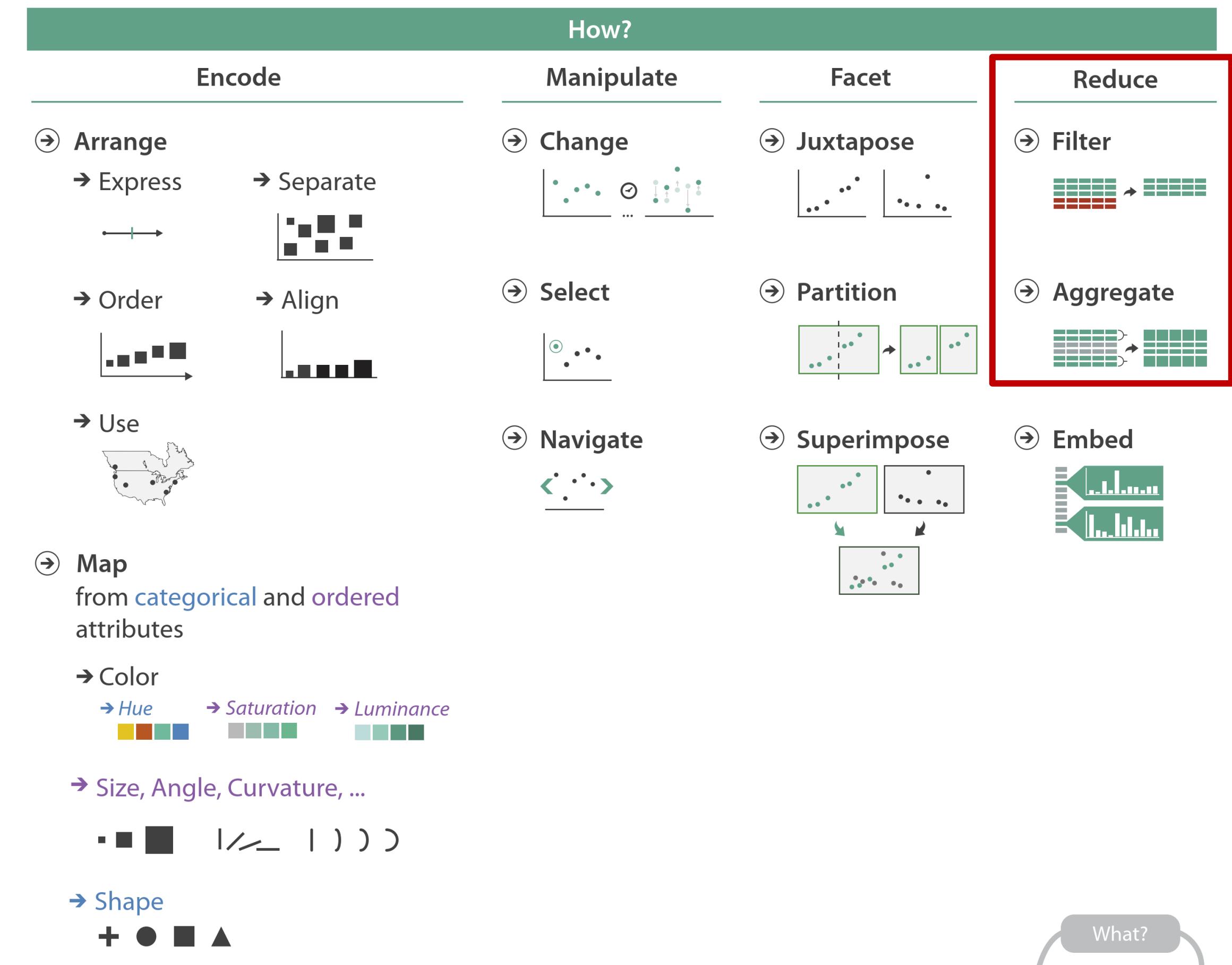
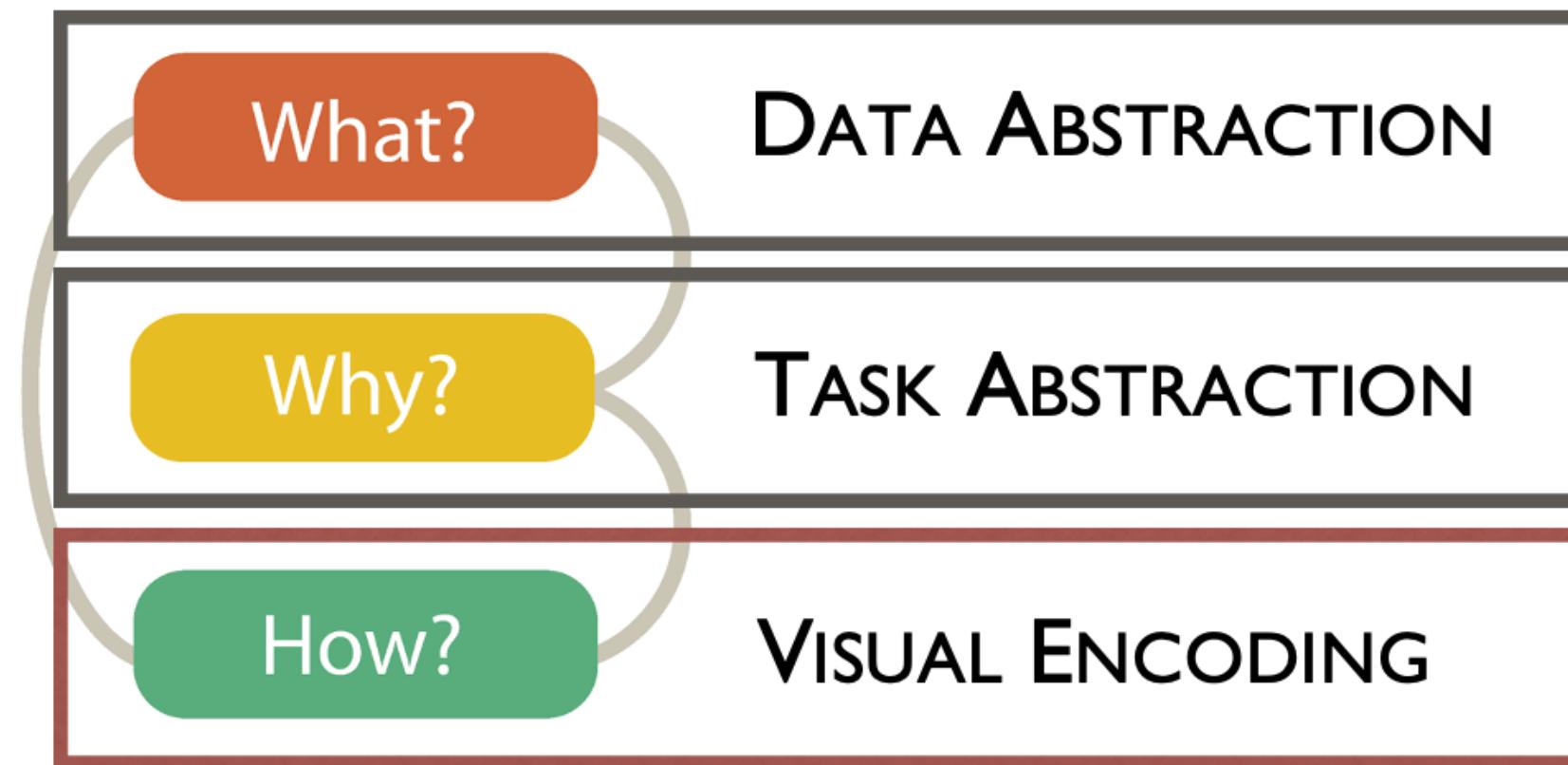
- 1.** Faceting – Split data into multiple views
- 2.** Interaction – Let user manipulate a single view
- 3.** Aggregate and Filter – Reduce amount of data visualized
- 4.** Focus+Context – Embed focused information

# Visualizing Big Data

1. Faceting – Split data into multiple views
2. Interaction – Let user manipulate a single view
3. Aggregate and Filter – Reduce amount of data visualized
4. Focus+Context – Embed focused information

# AGGREGATION AND FILTERING

# Munzner's Design Pipeline



# Caveats to Filtering & Aggregation

- Be careful not to hide important details
- Keep in mind your audience
  - Will they understand the resulting data?
  - Will they remember they're looking at reduced data?

# Filtering vs. Aggregation

**Filter** → Eliminate (hide) elements

**Aggregate** → Create new element(s)  
representing multiple

# Filtering vs. Aggregation

**Filter** → Eliminate (hide) elements

**Pros:** • Easy to understand

• Easy to compute

**Cons:** • Easy for the user to forget elements

**Aggregate** → Create new element(s) representing multiple

**Pros:** • No missing elements

• Can require sophisticated expertise from the user

**Cons:**

# Filtering

# Filtering

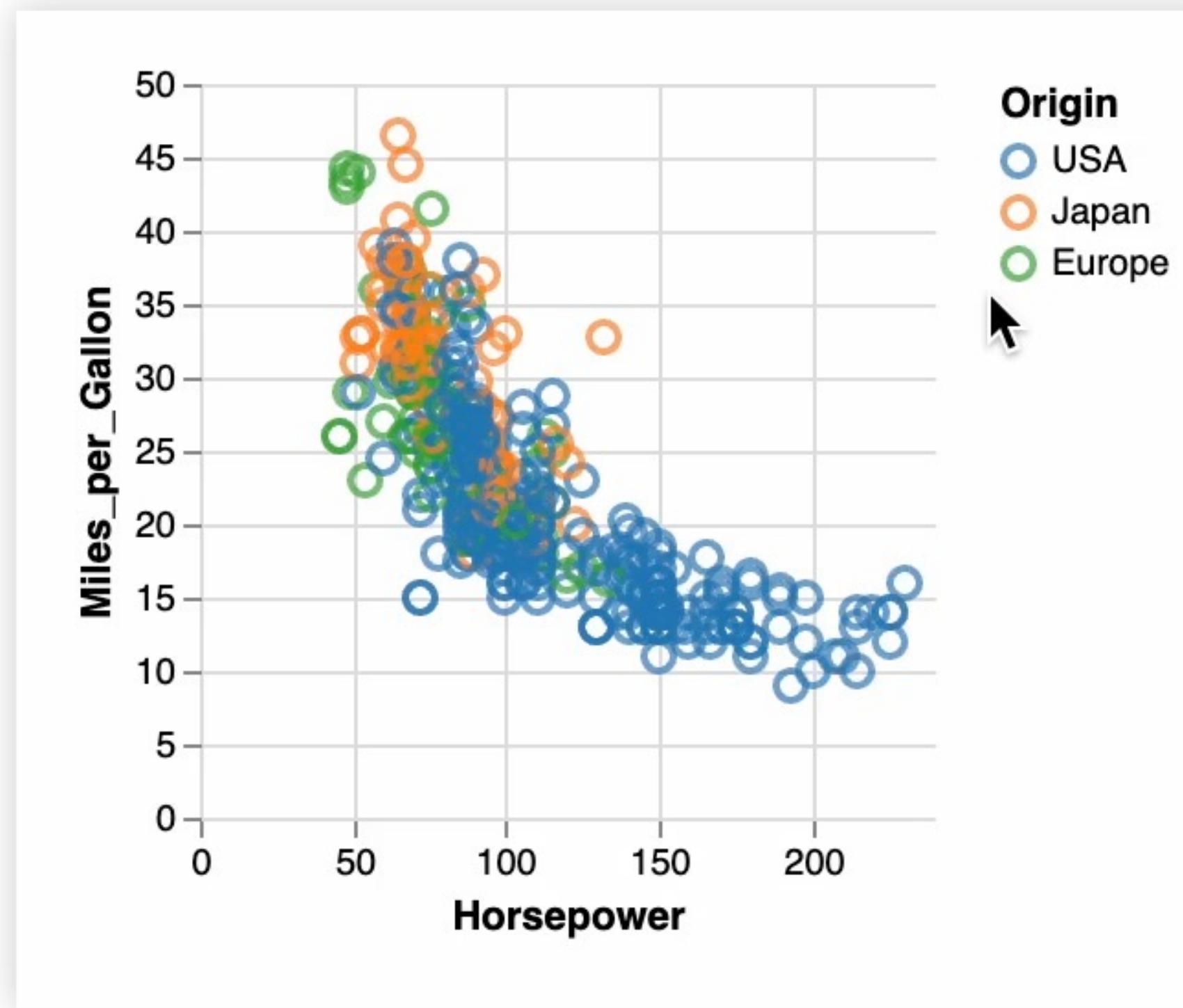
## Item filtering

→ eliminate items based on value

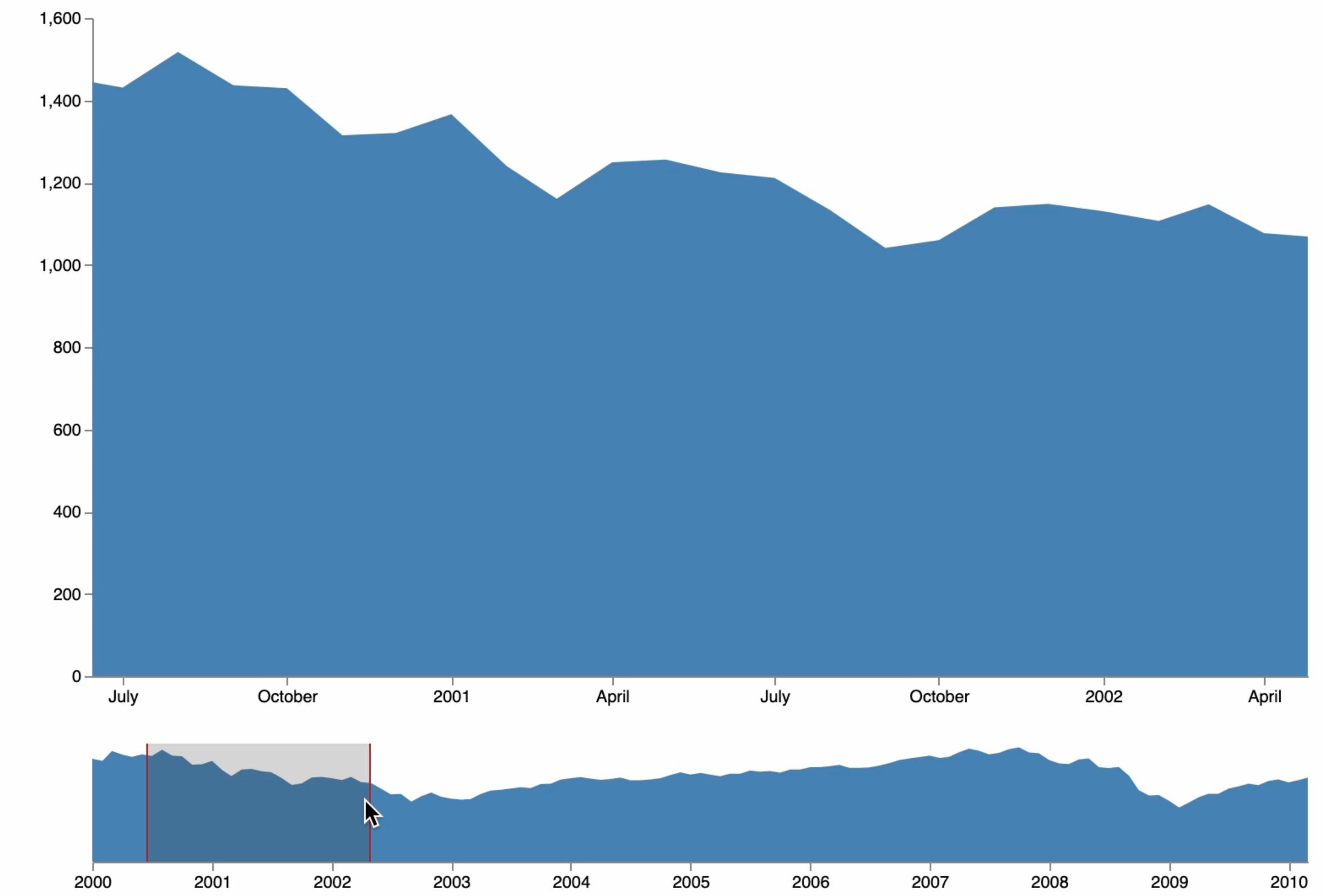
# Filtering

## Item filtering

→ eliminate items based on value



<https://vega.github.io/vega/examples/interactive-legend/>



<https://vega.github.io/vega/examples/overview-plus-detail/>

# Filtering

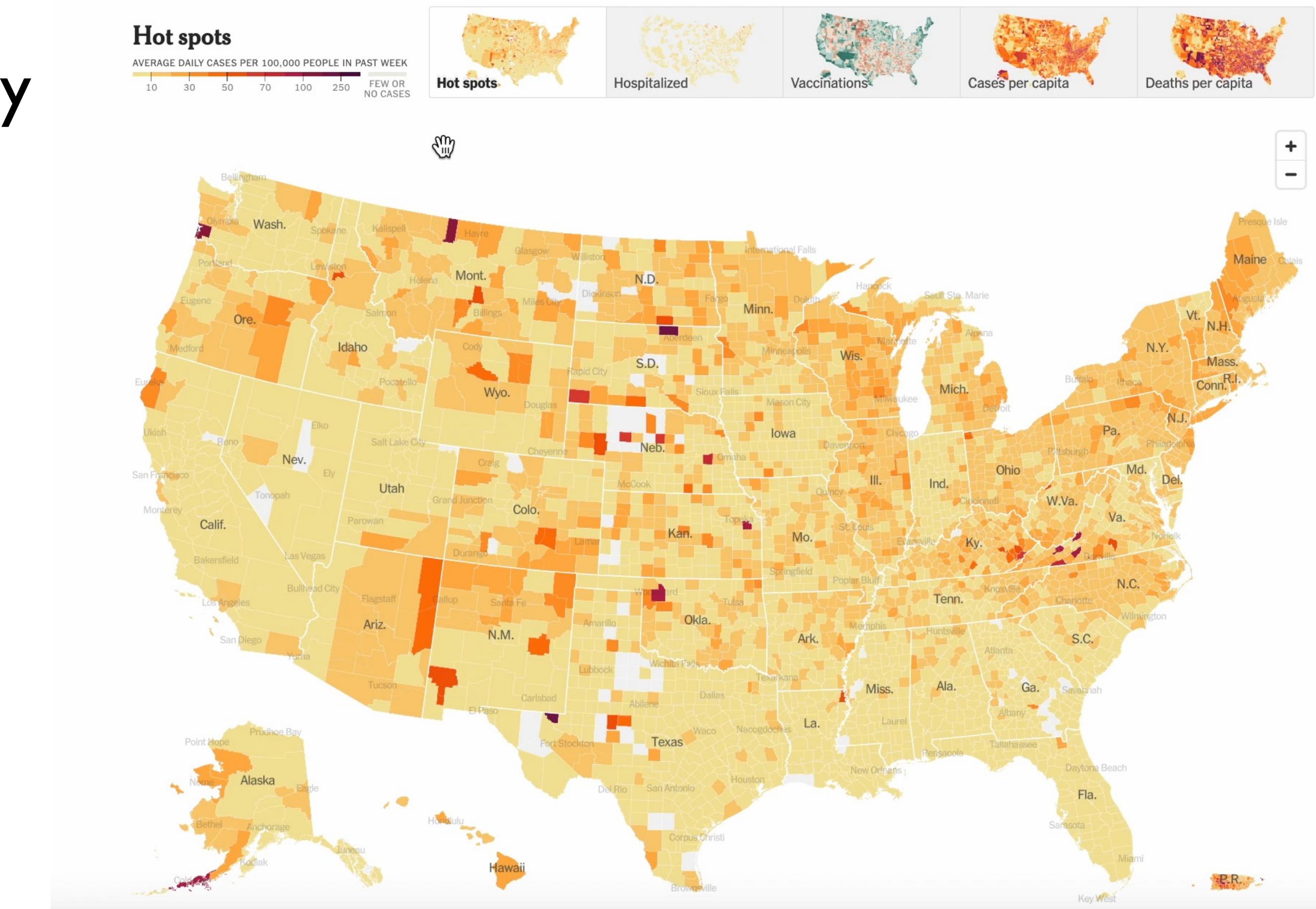
## **Attribute filtering**

→ eliminate attributes from display

# Filtering

## Attribute filtering

→ eliminate attributes from display



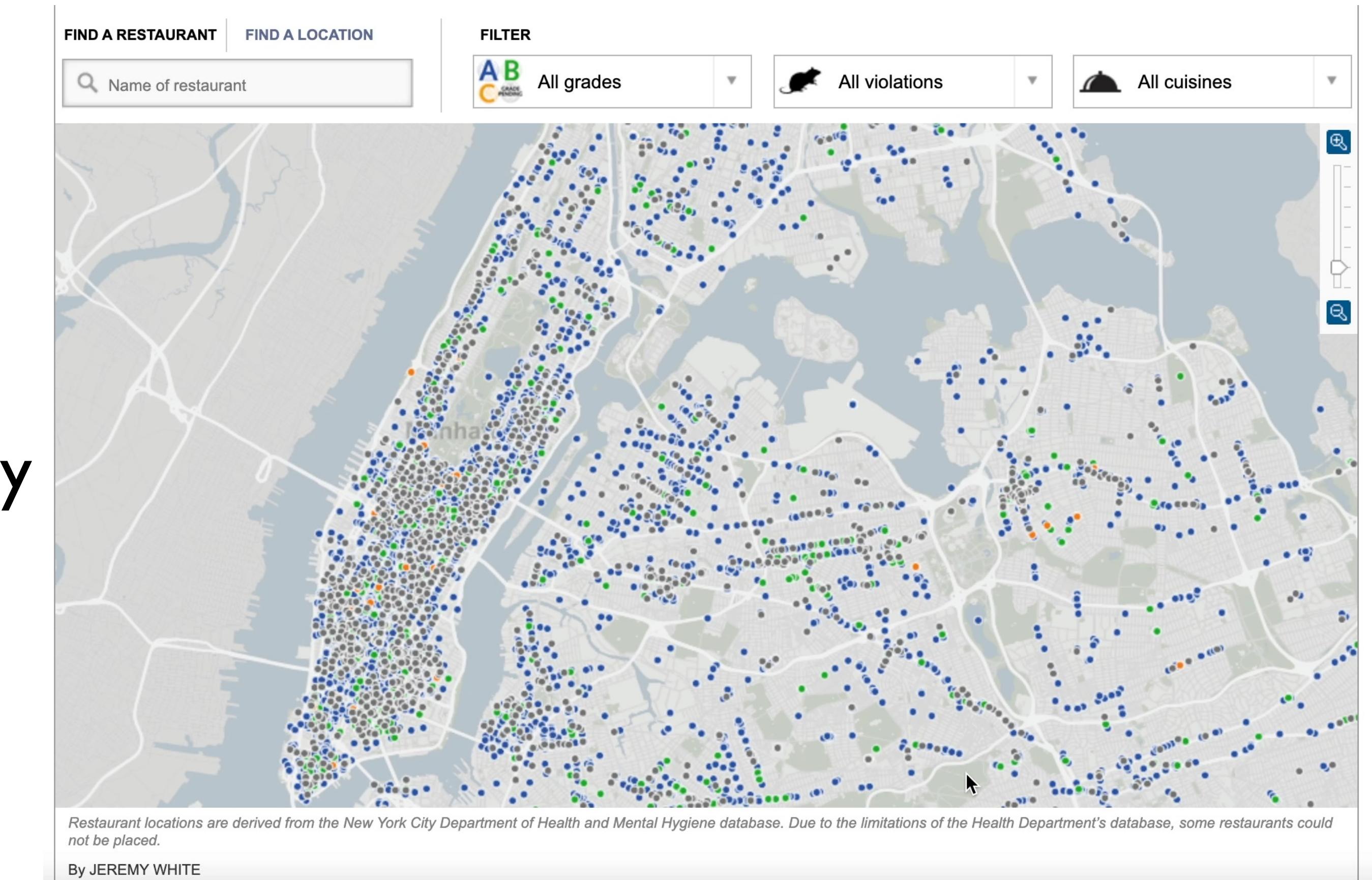
# Filtering

## Item filtering

→ eliminate items based on value

## Attribute filtering

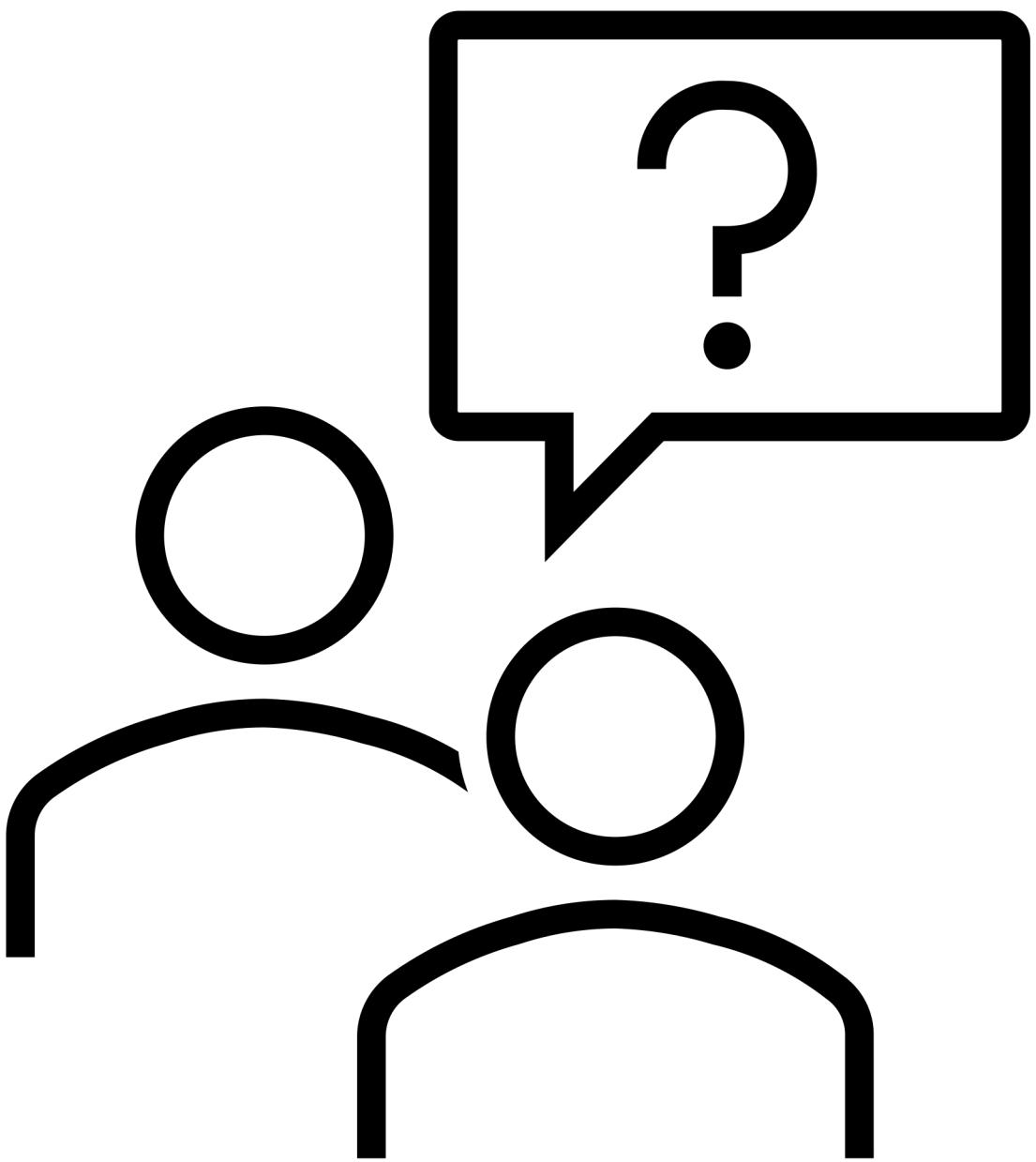
→ eliminate attributes from display



# Filtering

**How can we help the user remember  
that they are only viewing a subset of  
the data?**

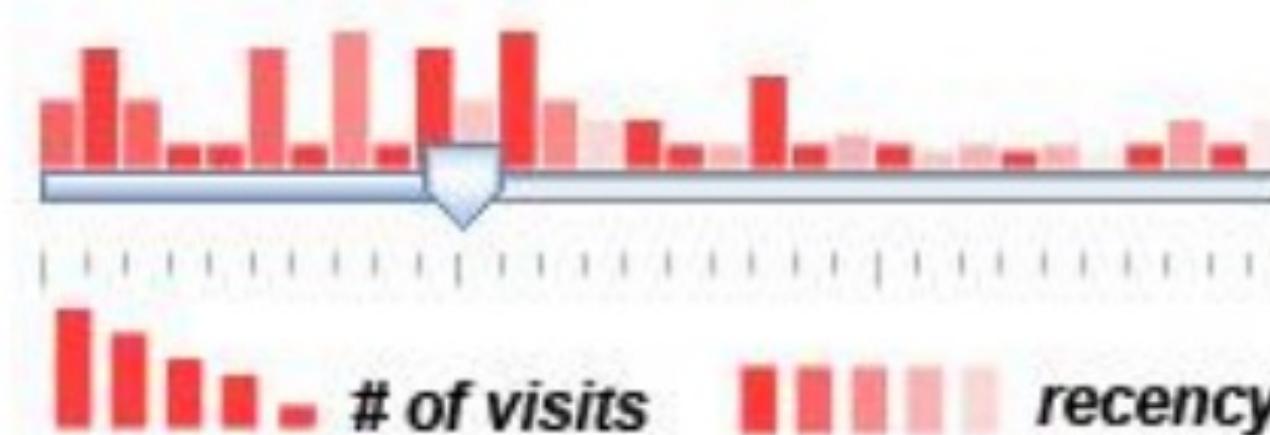
Sketch some ideas



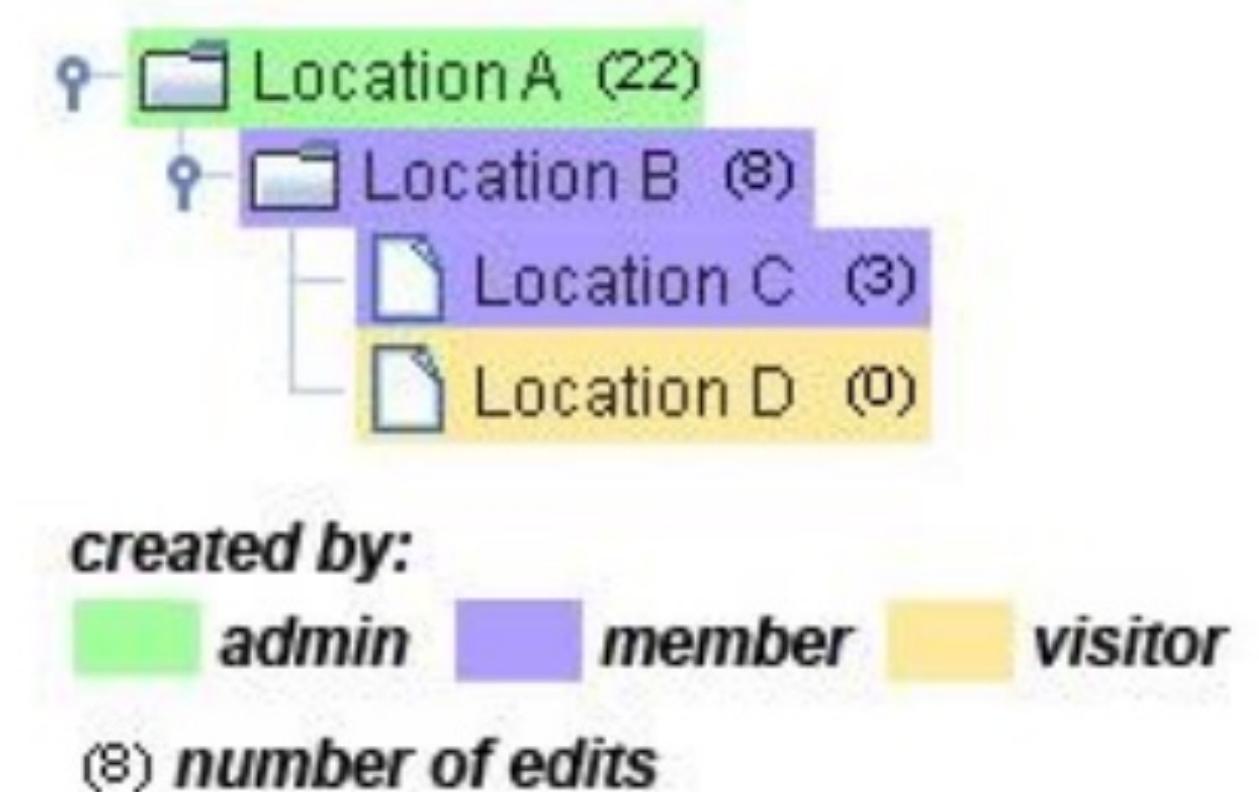
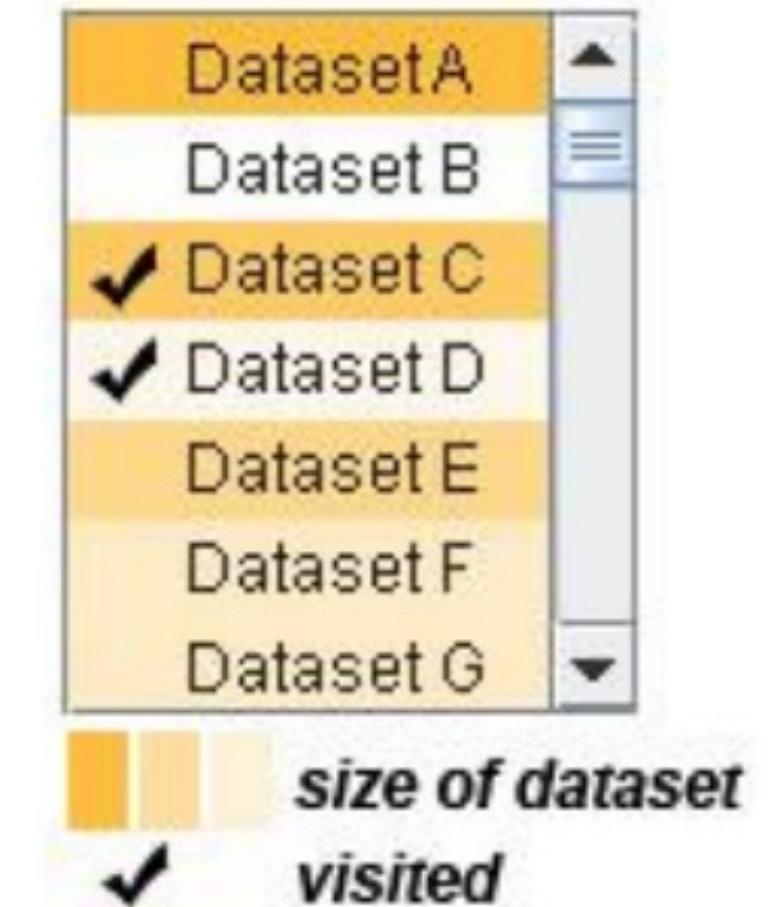
# Filtering

## Scented Widgets

→ Help address the risk of forgetting information



- 1st  Option A
- 3rd  Option B
- 2nd  Option C
- 3rd  Option D



# Aggregation

# Aggregation

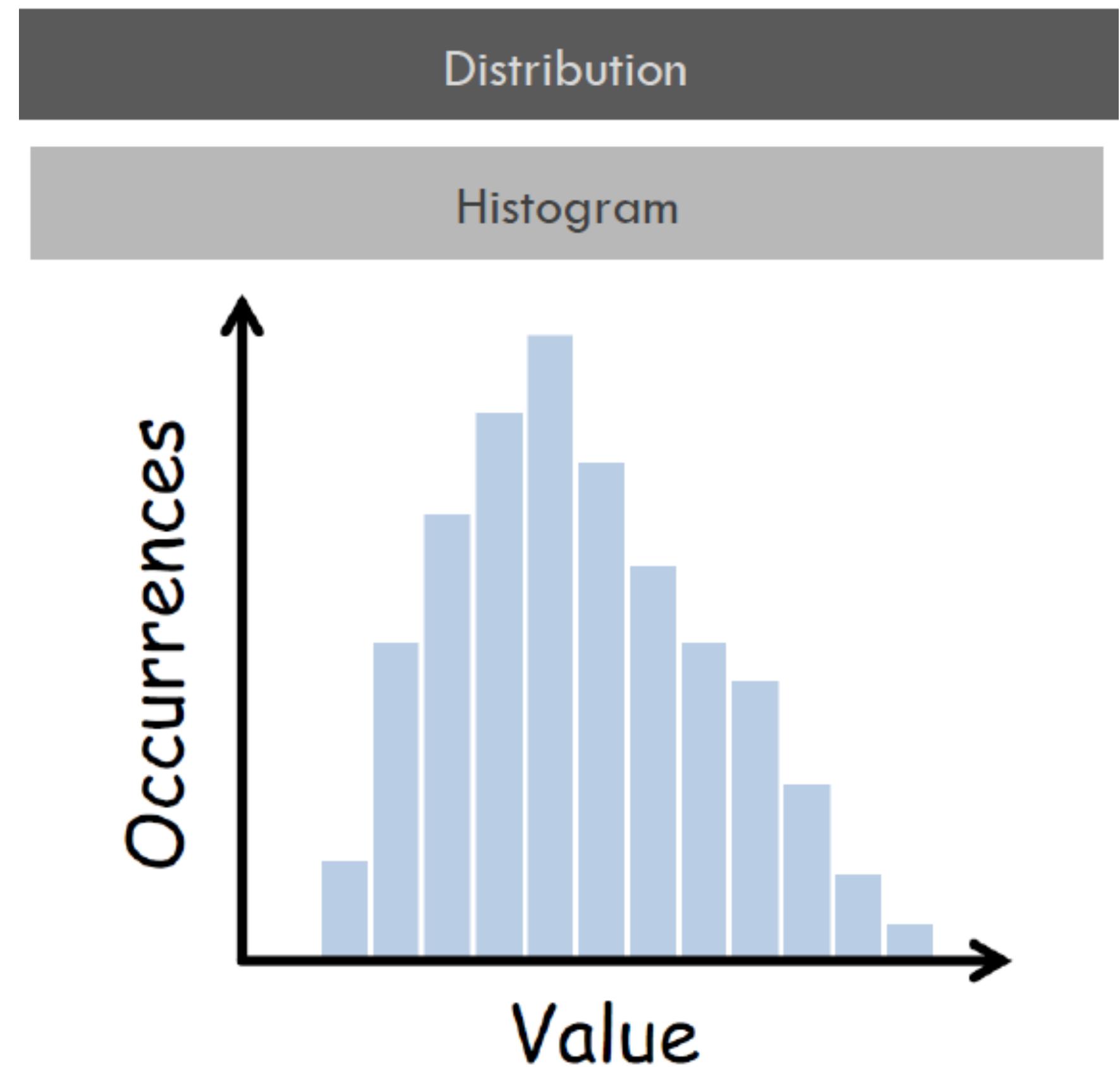
## **Item aggregation**

→ Create one item to represent many

# Aggregation

## Item aggregation

→ Create one item to represent many



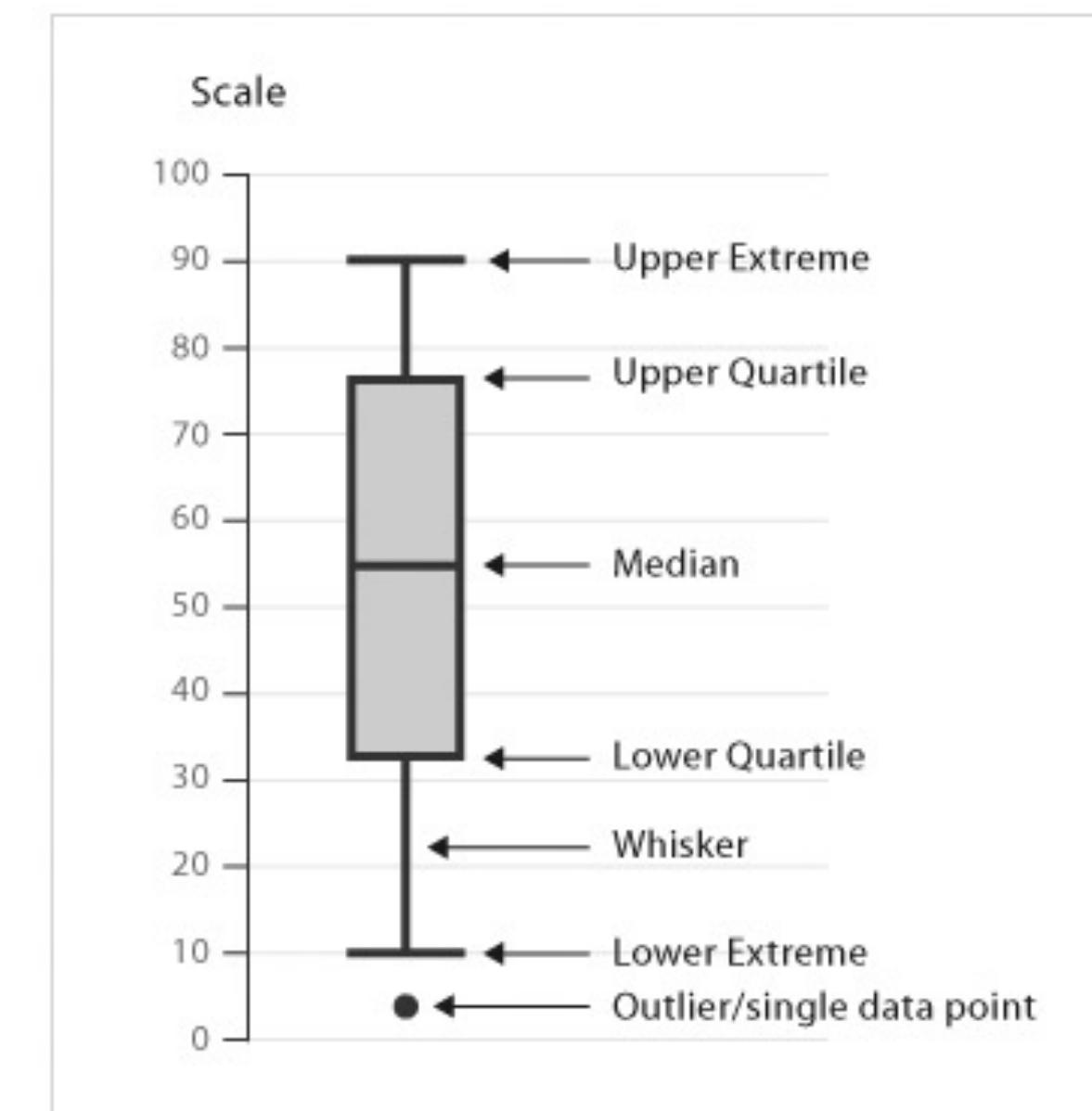
*bins (aggregates) of data*

# Aggregation

## Item aggregation

→ Create one item to represent many

Box-And-Whisker Plot



# Aggregation

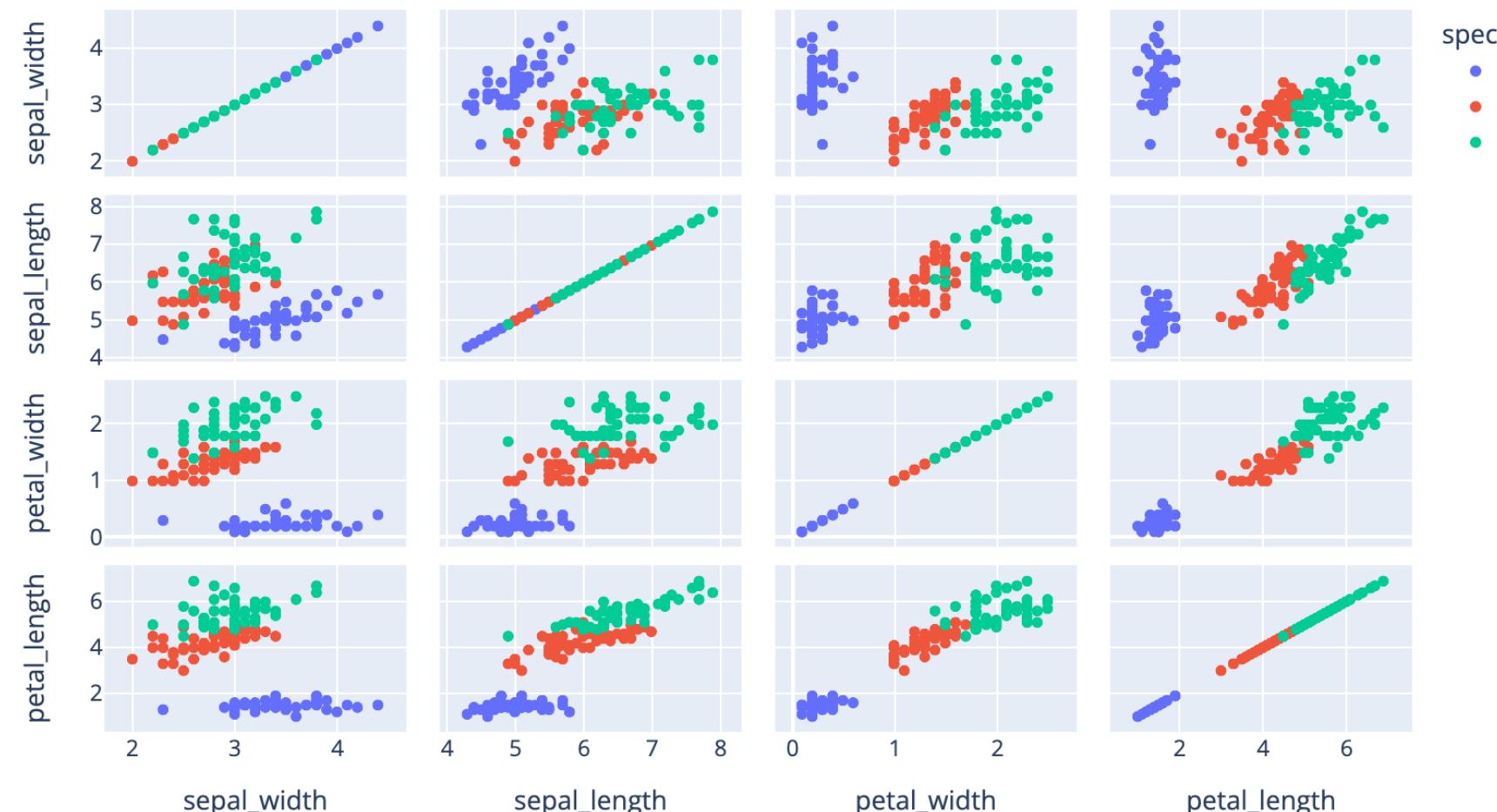
## **Attribute aggregation**

→ Collapse / reduce attributes

# Aggregation

## Attribute aggregation

→ Dimensionality reduction

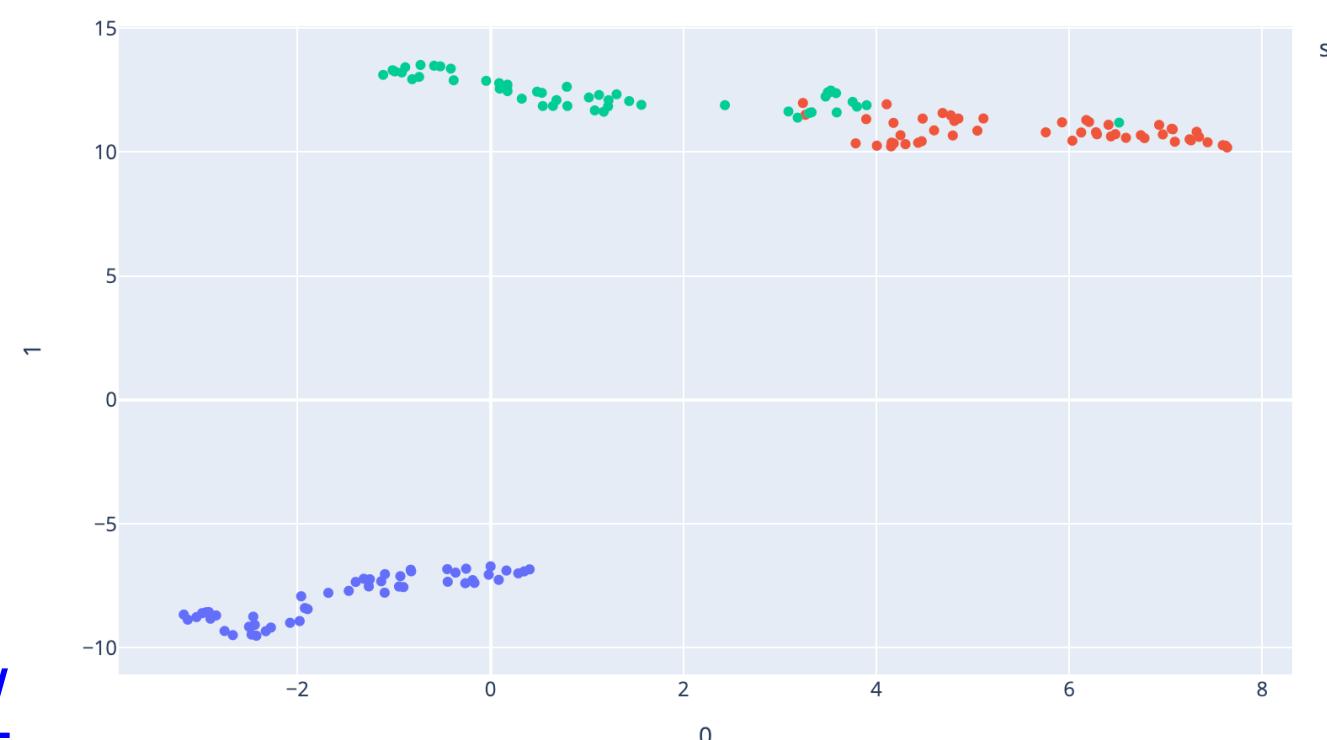


species  
• setosa  
• versicolor  
• virginica

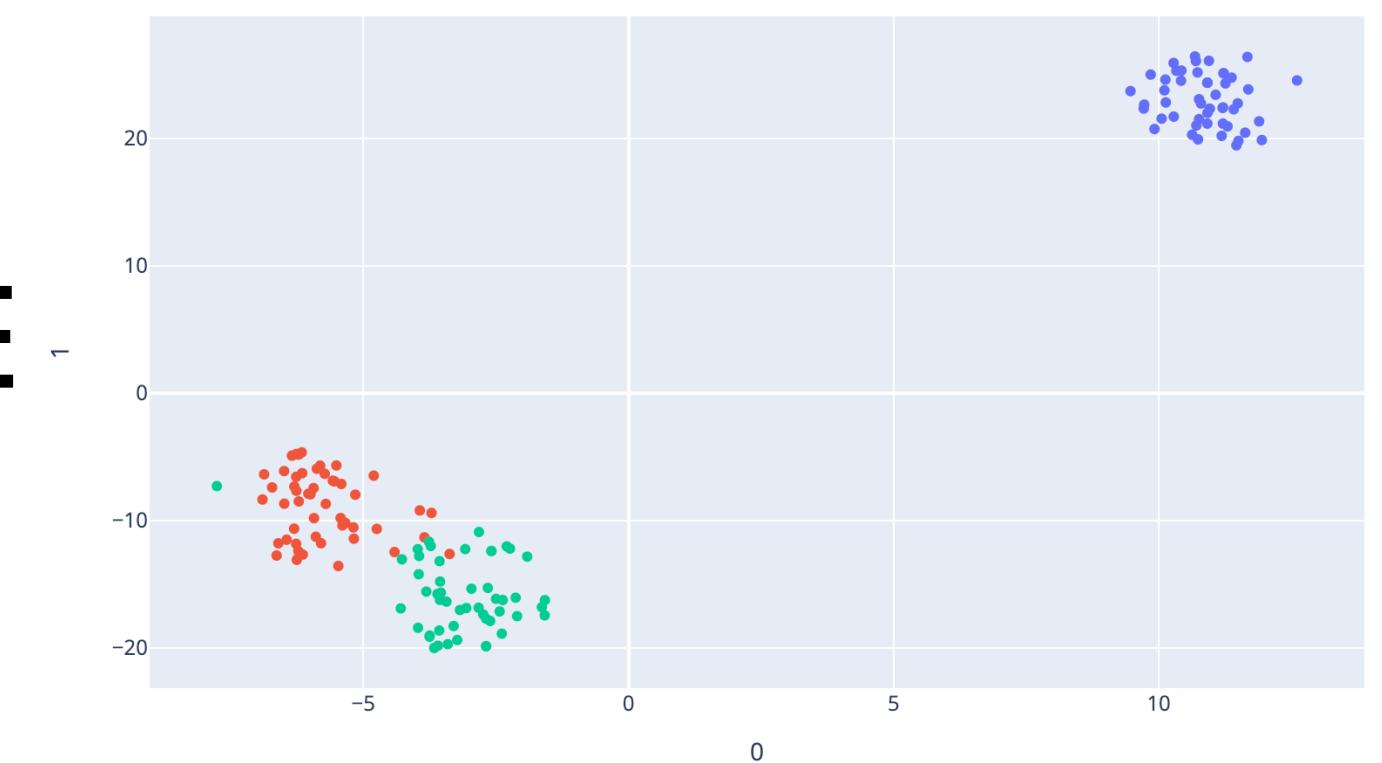


color  
• setosa  
• versicolor  
• virginica

PCA



t-SNE



UMAP

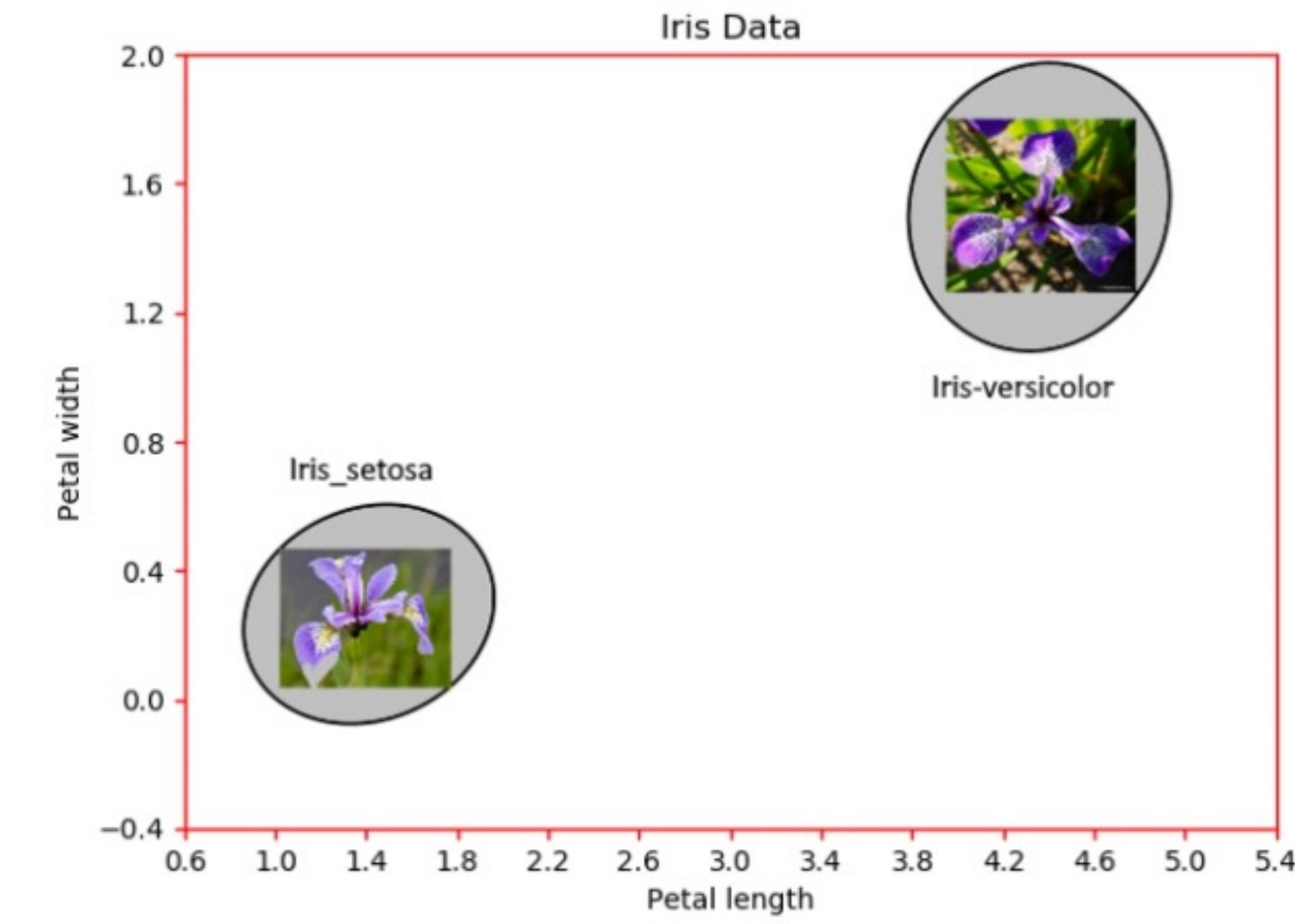
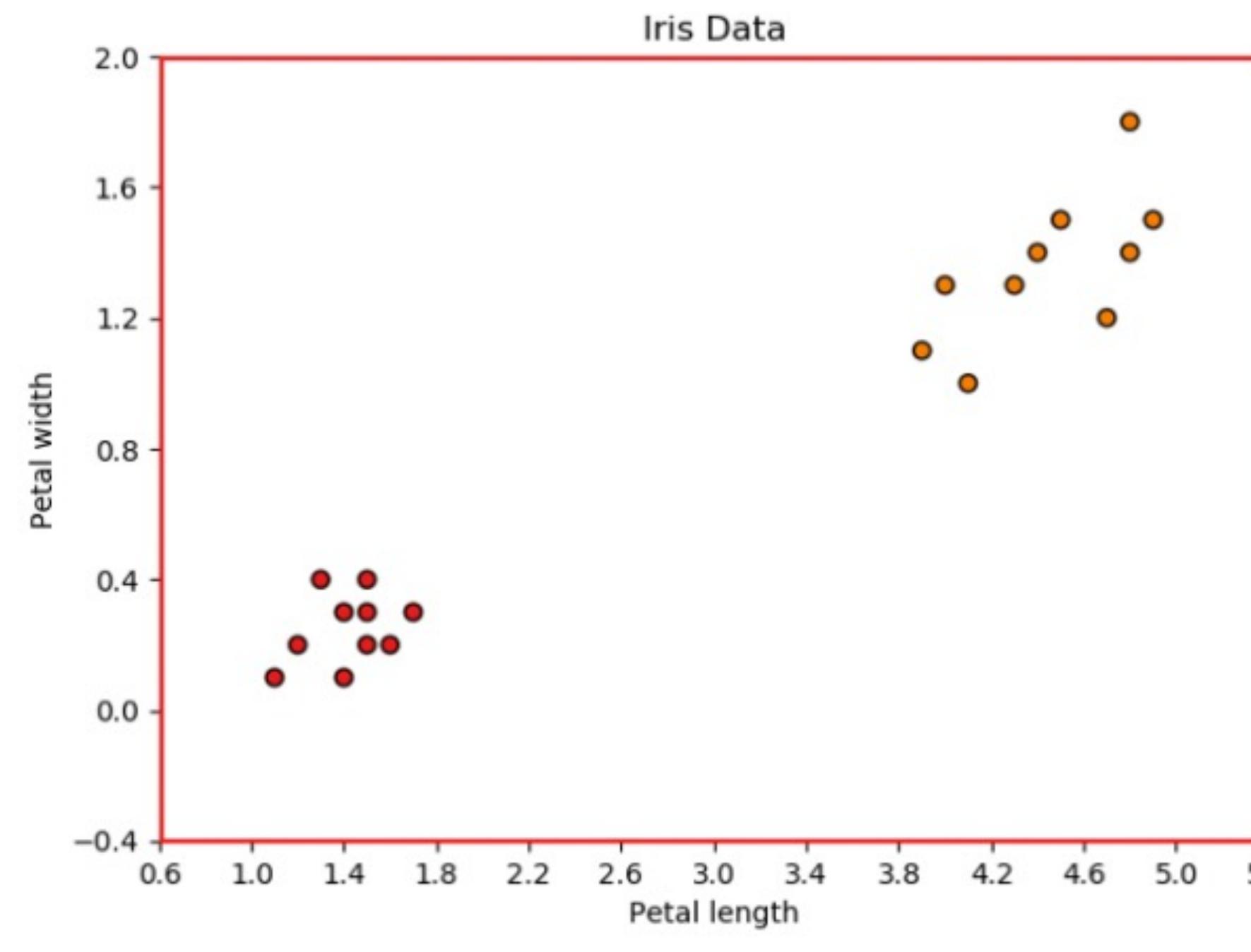
<https://plotly.com/python/pca-visualization/>

<https://plotly.com/python/t-sne-and-umap-projections/>

# Aggregation

## Attribute aggregation

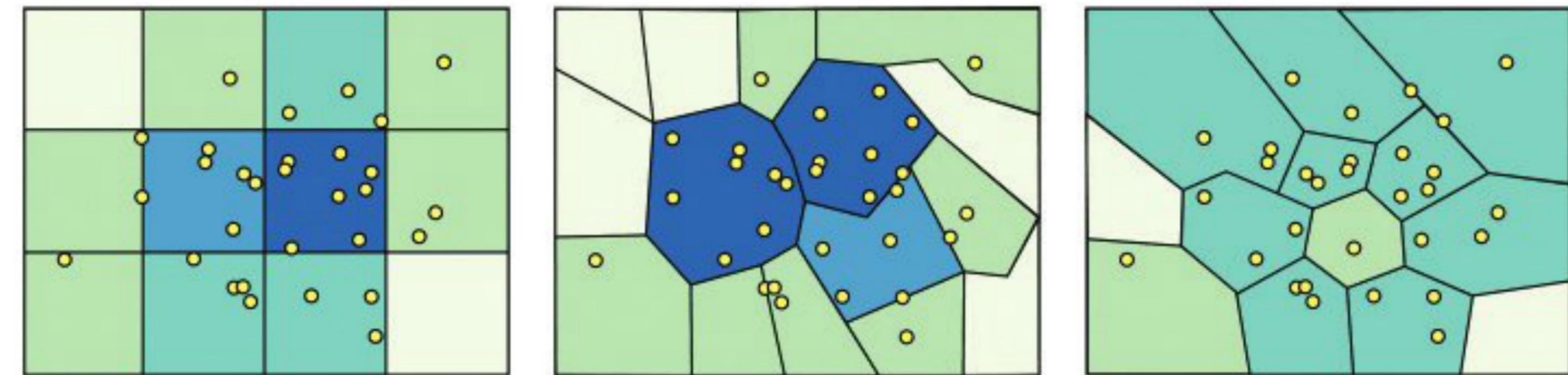
→ Group by similarity measure



# Aggregation

## Spatial aggregation

→ Collapse by spatial similarity



**Warning!** boundary definition can dramatically change data analysis

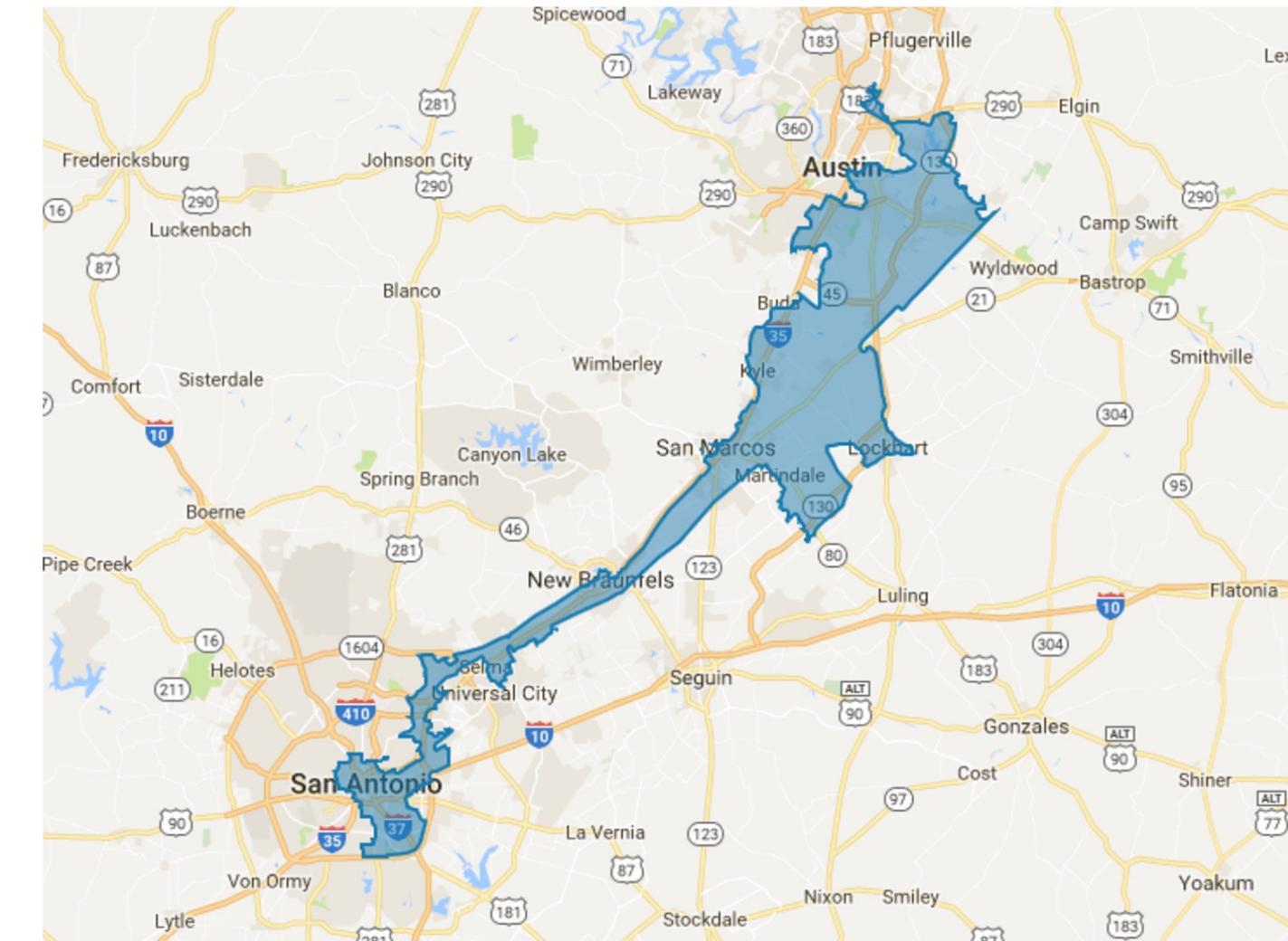
# Aggregation

**Spatial aggregation**  
→ Collapse by spatial similarity

Federal Court Rules Three Texas Congressional Districts Illegally Drawn

March 11, 2017 · 5:56 PM ET

LAUREL WAMSLEY 



A federal panel ruled Friday that three of Texas's Congressional districts, including the 35th, shown here, were illegally drawn by the state's Republicans.

Screengrab by NPR/Google Maps

**Gerrymandering**

<http://www.npr.org/sections/thetwo-way/2017/03/11/519839892/federal-court-rules-three-texas-congressional-districts-illegally-drawn>