

# Communicating with Data – Choosing Visual Channels

Dr. Ab Mosca (they/them)

Slides based off slides courtesy of Jordan Crouser (<https://jcrouser.github.io/>)

# Reminder












- Download before class on Tuesday 02/11
- Use the free 14 day trial until tableau sends an educational license
- Need help? Ping me on Slack, or come to office hours


# Plan for Today

- Recap data-visual mapping
- Expressiveness & effectiveness of visual channels
- A sampling of visualization techniques

Recall: Data → Visuals

- Remember... **Big idea behind visualization**
  - Map data dimensions to visual dimensions in a principled way**

Position on common scale	
Position on unaligned scale	
Length (1D size)	
Tilt/angle	
Area (2D size)	
Color luminance	
Color saturation	
Color hue	
Shape	

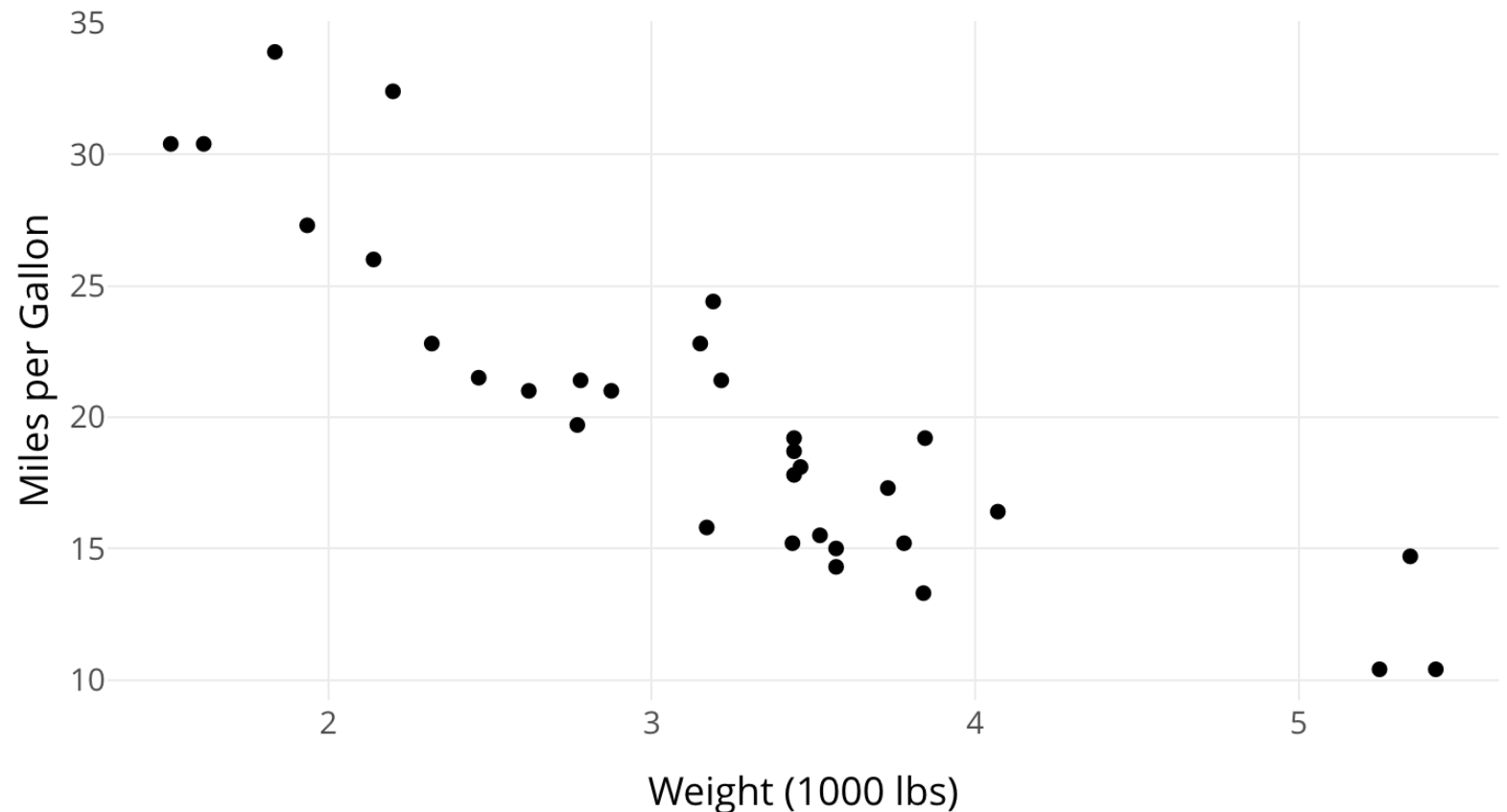
		
Categorical	Ordinal	Quantitative
✓	✓	✓
✓	✓	✓
	✓	✓
	✓	✓
	✓	✓
	✓	✓
	✓	✓
✓		
✓		

# Visual Encoding

- Remember... **Big idea behind visualization**
  - **Map data dimensions to visual dimensions in a principled way**
- The process of representing variables via visual channels is called **encoding**
- **Visual encoding** is another term for a visualization
- One way we will often analyze visualizations is by inspecting their data-visual mapping (i.e. encoding)

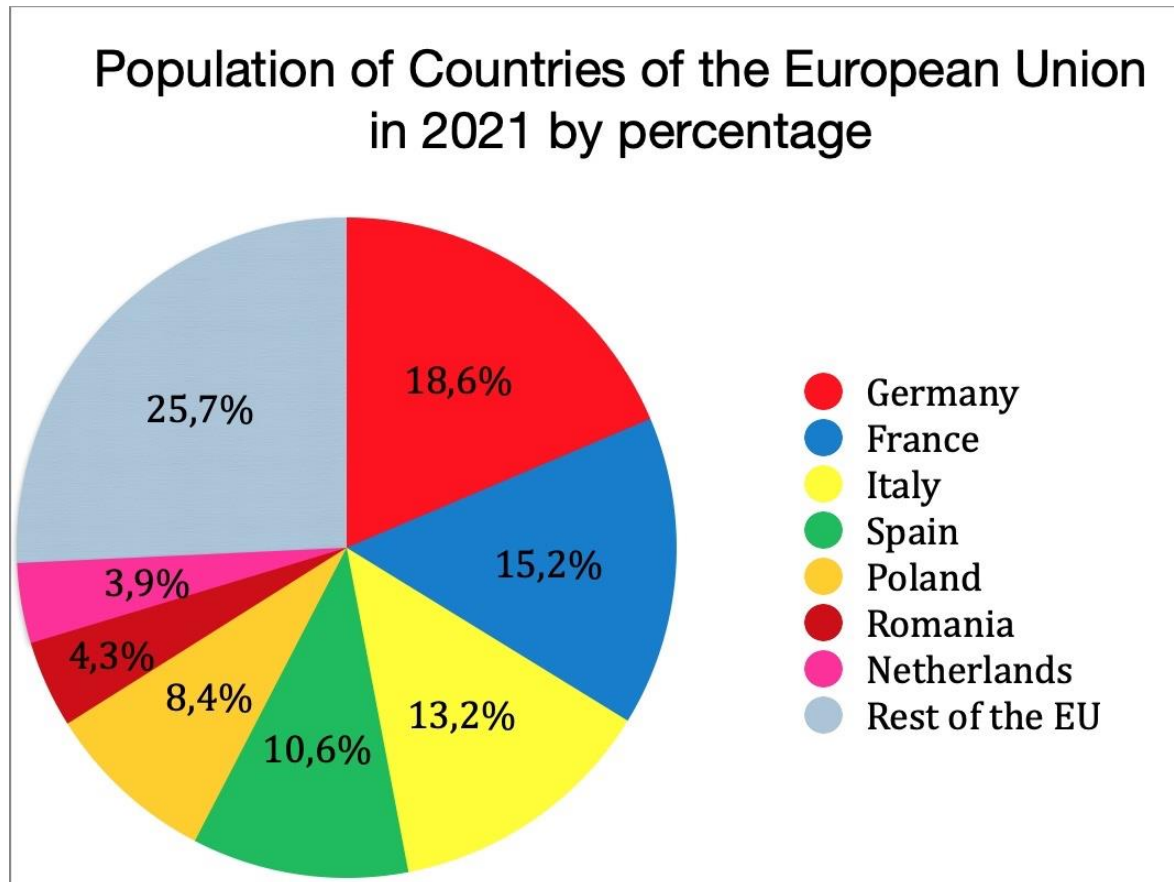
## Ex. Data-Visual Mapping

- What are the marks in this visualization?
  - What variables are shown?
  - What visual channel(s) encode(s) each variable?
- Based on data type and channel type, is the encoding appropriate?



## Ex. Data-Visual Mapping

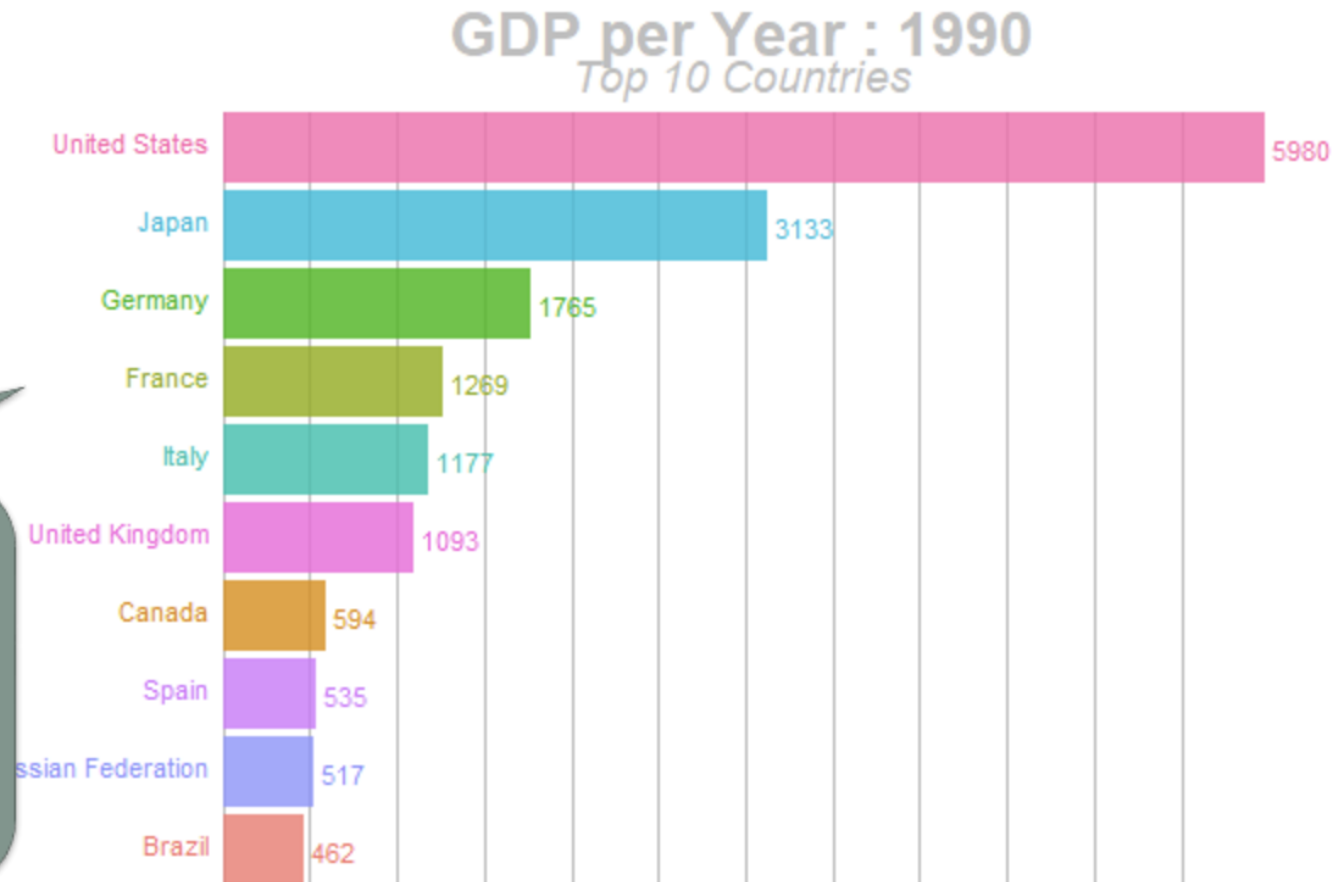
- What are the marks in this visualization?
  - What variables are shown?
- What visual channel(s) encode(s) each variable?
- Based on data type and channel type, is the encoding appropriate?



- What are the marks in this visualization?
  - What variables are shown?
  - What visual channel(s) encode(s) each variable?
- Based on data type and channel type, is the encoding appropriate?

## Ex. Data-Visual Mapping


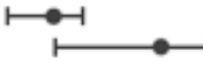



**Dual (or double) encoding** ==  
encoding one variable via  
more than one channel









# Ranking Visual Channels

## ➔ Magnitude Channels: **Ordered** Attributes

Position on common scale	
Position on unaligned scale	
Length (1D size)	
Tilt/angle	
Area (2D size)	
Depth (3D position)	
Color luminance	
Color saturation	
Curvature	
Volume (3D size)	

## ➔ Identity Channels: **Categorical** Attributes

Spatial region	
Color hue	
Motion	
Shape	

- Using the correct visual channel type for the correct variable type is called **expressiveness**
- We also consider **effectiveness** to choose visual channels

# Ranking Visual Channels

- More important attributes should be encoded with more **effective** channels.
- **Effectiveness** refers to how well a channel supports:
  - Accuracy
  - Discriminability
  - Separability
  - Visual popout
  - Grouping

# Accuracy

**Definition:** how close human perceptual judgement is to an objective measurement of the stimulus

# Accuracy

**Definition:** how close human perceptual judgement is to an objective measurement of the stimulus



How much longer is the second bar?

# Accuracy

**Definition:** how close human perceptual judgement is to an objective measurement of the stimulus



How much longer is the second bar?  
2X

# Accuracy

**Definition:** how close human perceptual judgement is to an objective measurement of the stimulus



How much bigger is the second square?

# Accuracy

**Definition:** how close human perceptual judgement is to an objective measurement of the stimulus



How much bigger is the second square?

4X

# Accuracy

**Definition:** how close human perceptual judgement is to an objective measurement of the stimulus



How much bigger is the second box?



# Accuracy

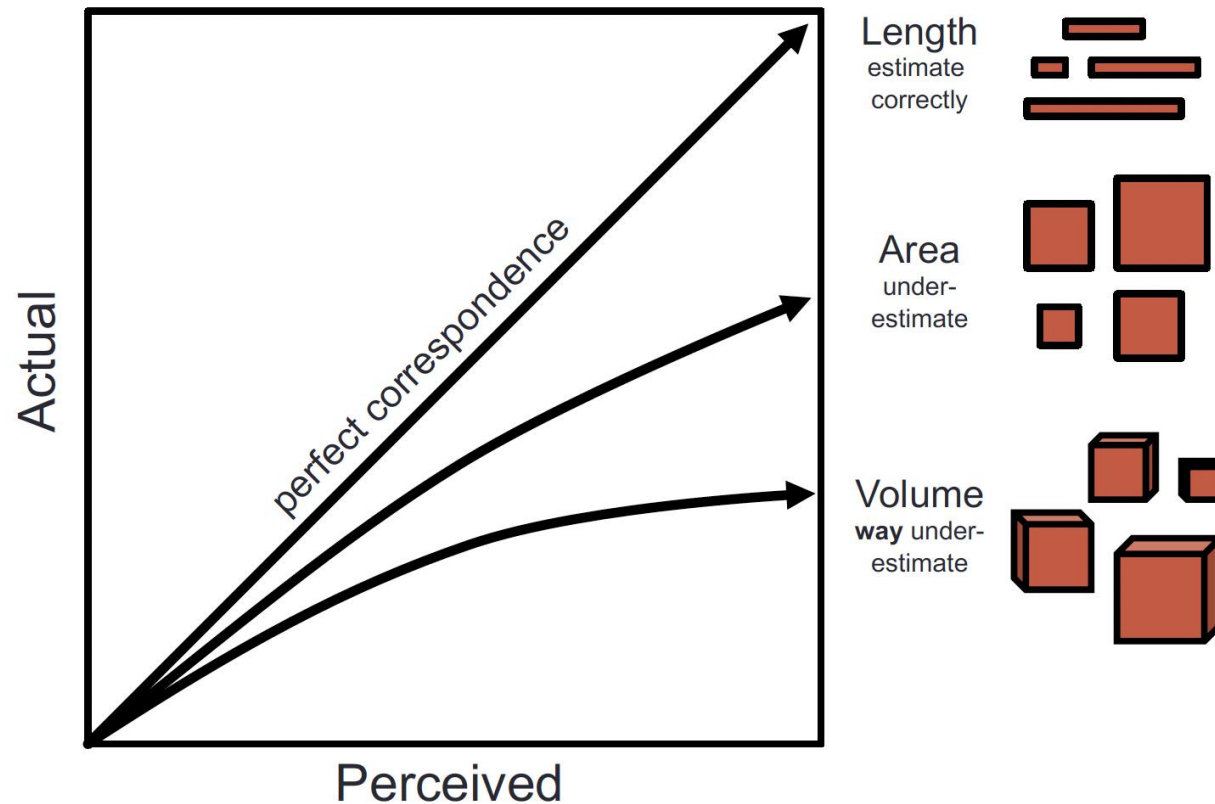
**Definition:** how close human perceptual judgement is to an objective measurement of the stimulus



How much bigger is the second box?  
27X

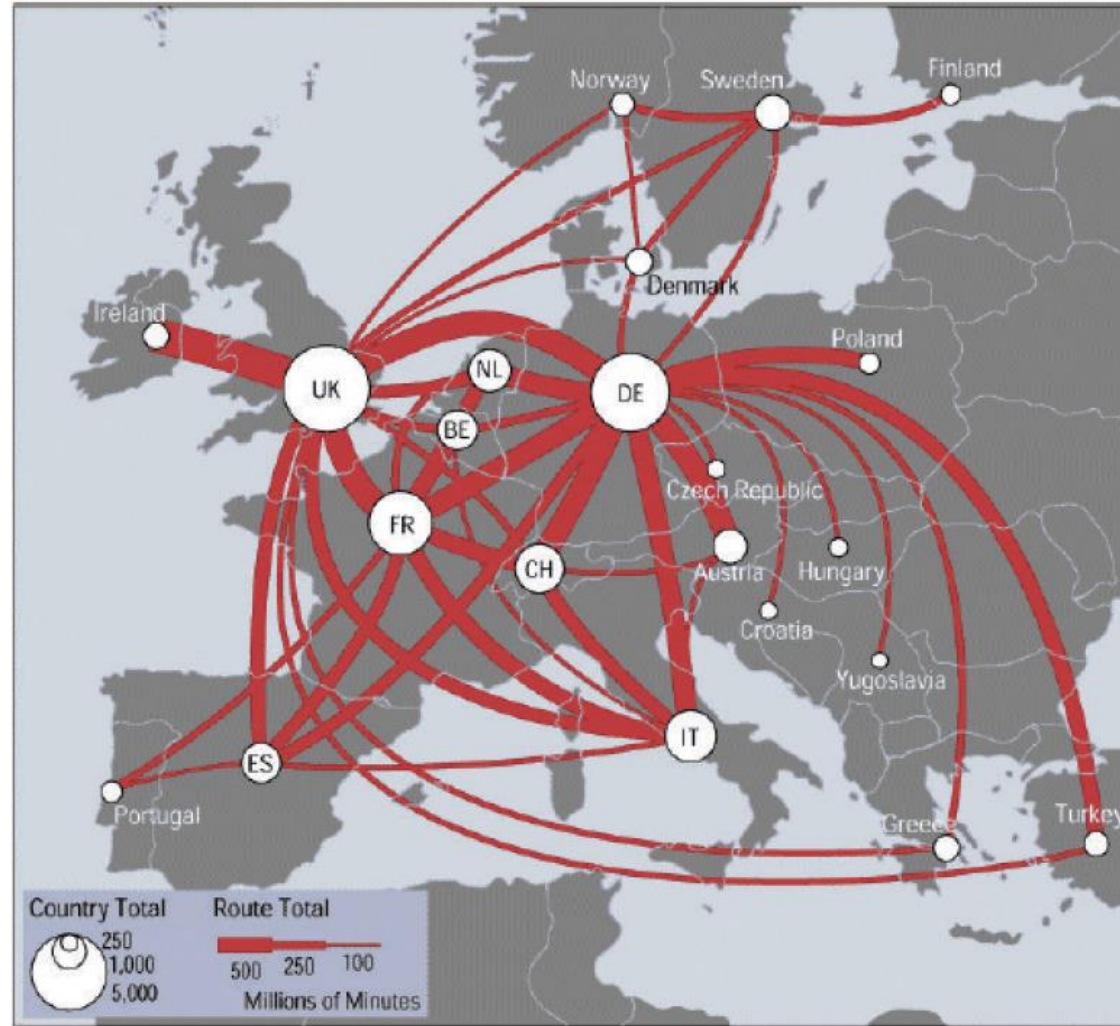
# Accuracy

**Definition:** how close human perceptual judgement is to an objective measurement of the stimulus



# Discriminability

**Definition:** how differentiable levels of the channel are

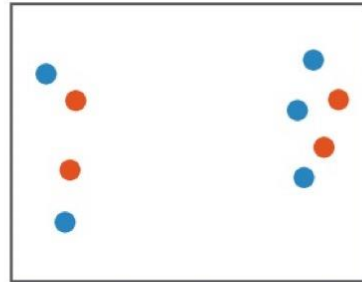


<https://web.cse.ohio-state.edu/~shen.94/Melbourne/Slides/TamaraChp5.pdf>

# Separability

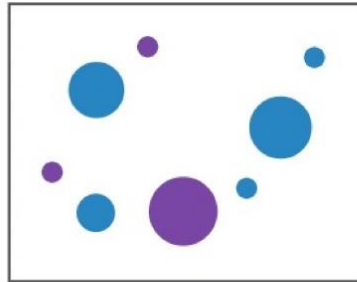
**Definition:** whether channels exist independently or integrally with others

Position  
+ Hue (Color)



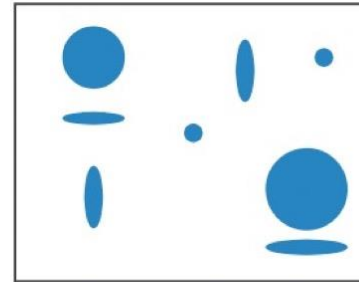
Fully separable

Size  
+ Hue (Color)



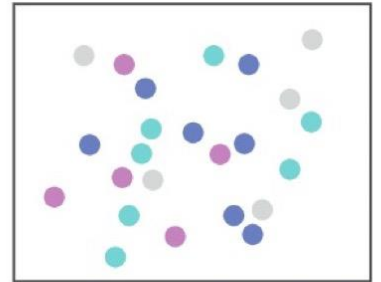
Some interference

Width  
+ Height



Some/significant  
interference

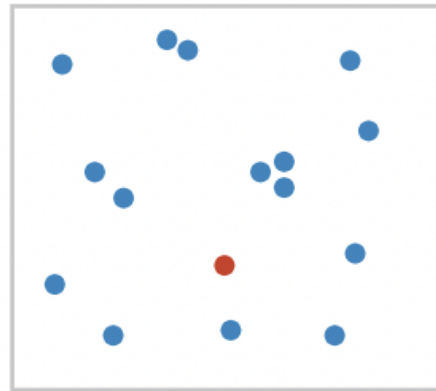
Red  
+ Green (saturation)



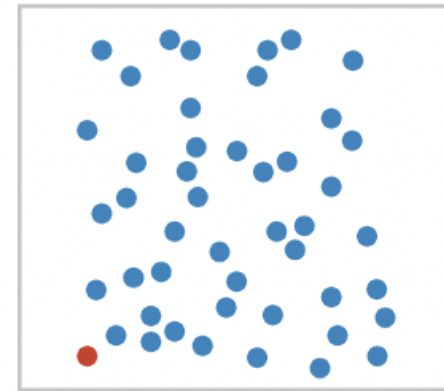
Major interference

# Visual Popout

**Definition:** how well a distinct item stands out from others



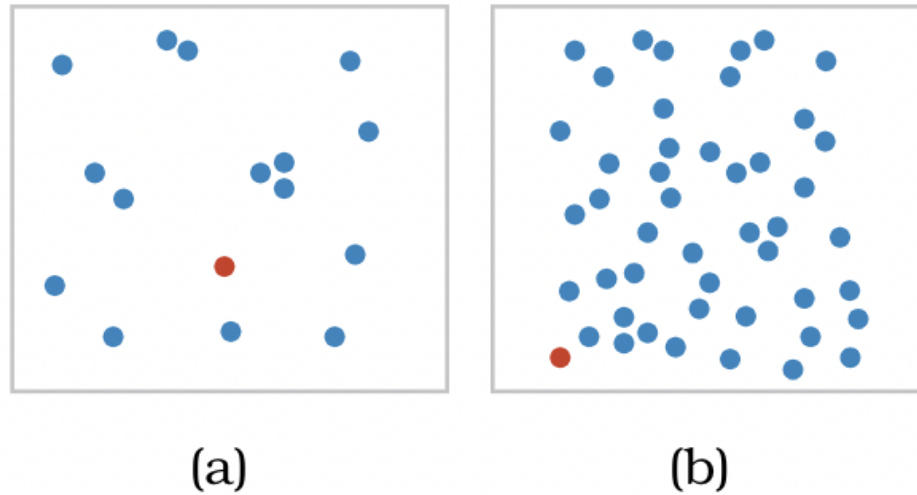
(a)



(b)

# Visual Popout

**Definition:** how well a distinct item stands out from others



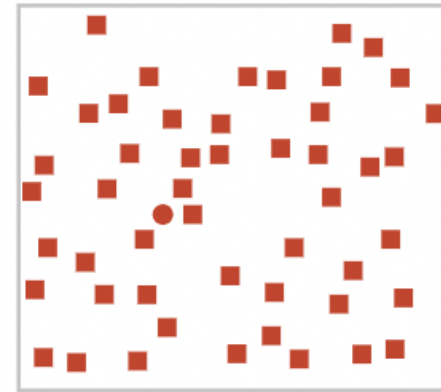
Color is a good channel for this

# Visual Popout

**Definition:** how well a distinct item stands out from others



(c)



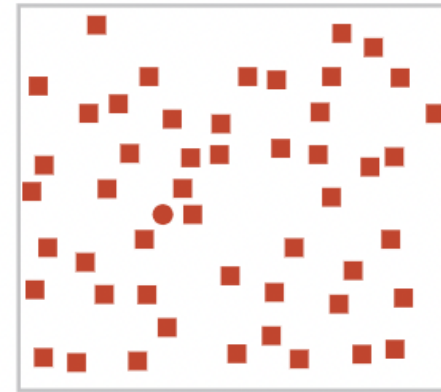
(d)

# Visual Popout

**Definition:** how well a distinct item stands out from others



(c)



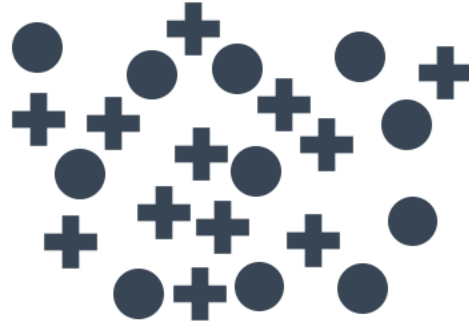
(d)

Shape is not as helpful

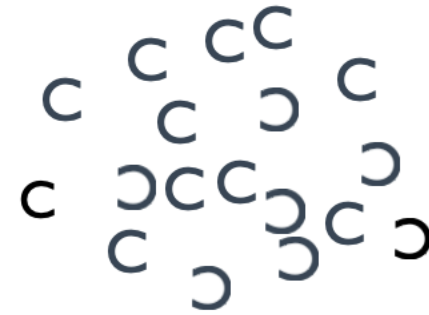


# Grouping

**Definition:** how likely people are to infer differences as representing distinct groups



Circles and +'s

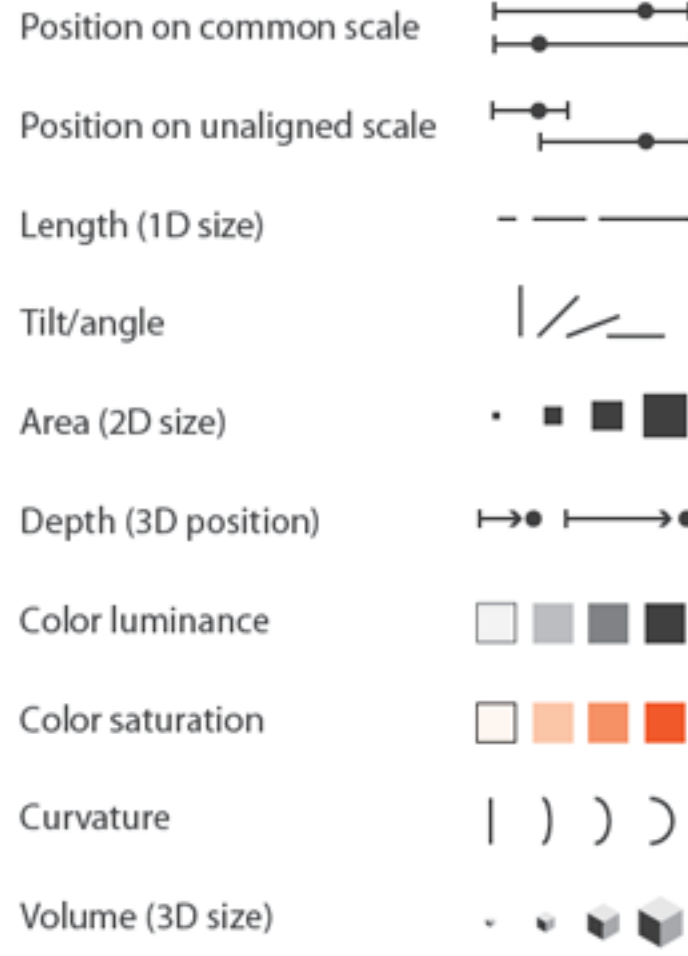


C's and D's

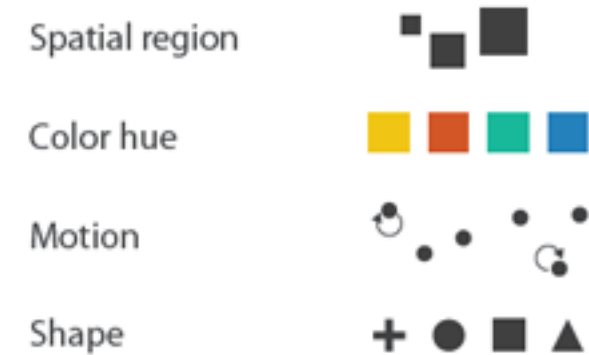
# Ranking Visual Channels

- **Design takeaway:** Prioritize most effective channels for most important data

## ➔ Magnitude Channels: Ordered Attributes



## ➔ Identity Channels: Categorical Attributes



## Your Turn!

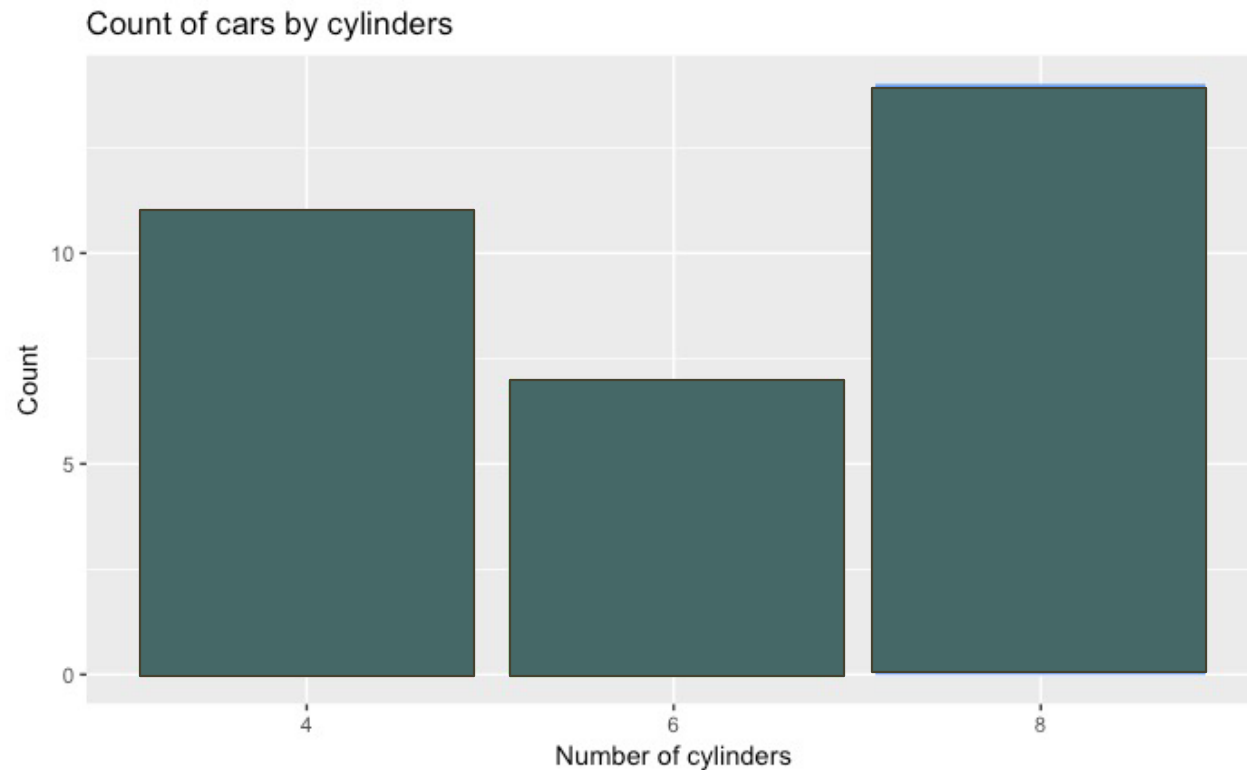
- Find someone (or two) to work with
- Choose a visualization from <https://informationisbeautiful.net/>
- What variables does the visualization show?
- What visual channel(s) encode each variable?
- Does the visualization follow the principles of effectiveness and expressiveness?

# A sampling of visualization techniques



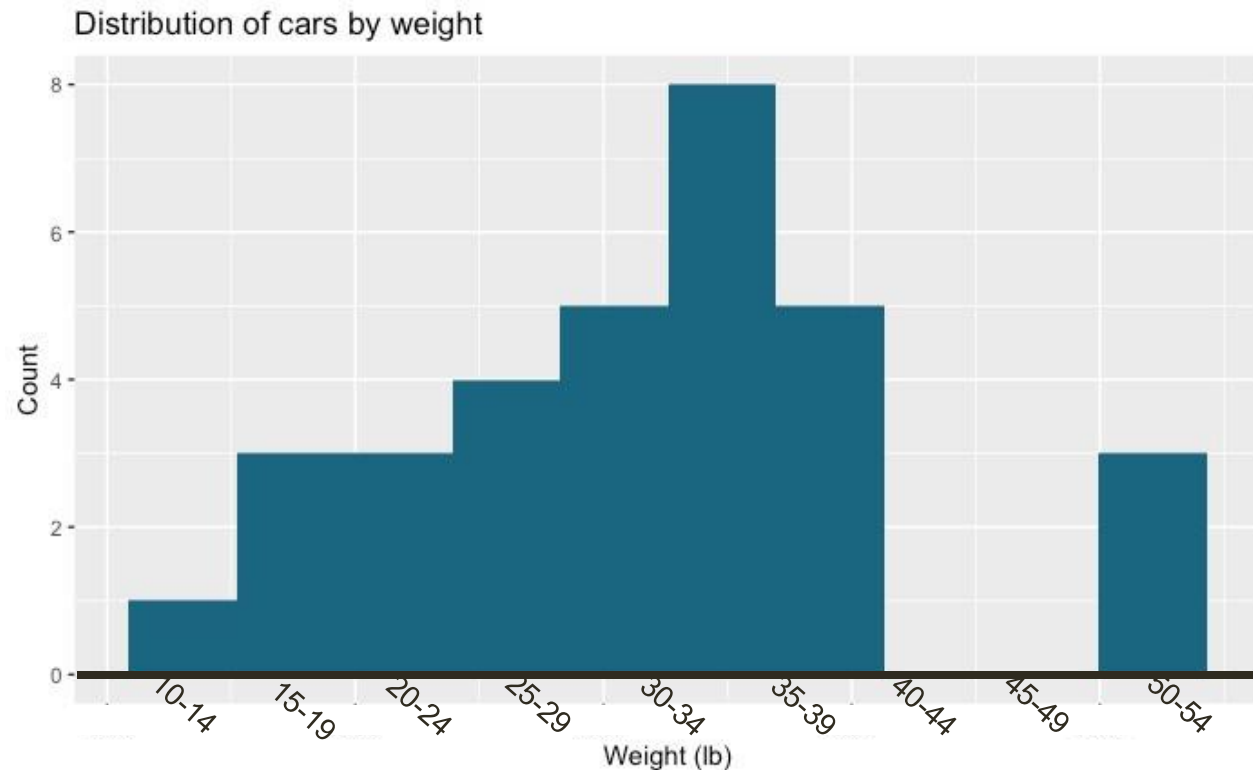
# Bar chart

- Used to show **quantitative variable X categorical or ordinal variable**
- **Compares** quantitative values for different categories
- Highlights **relative amounts**
- Grouped/stacked bars can break each **category** into sub-groups



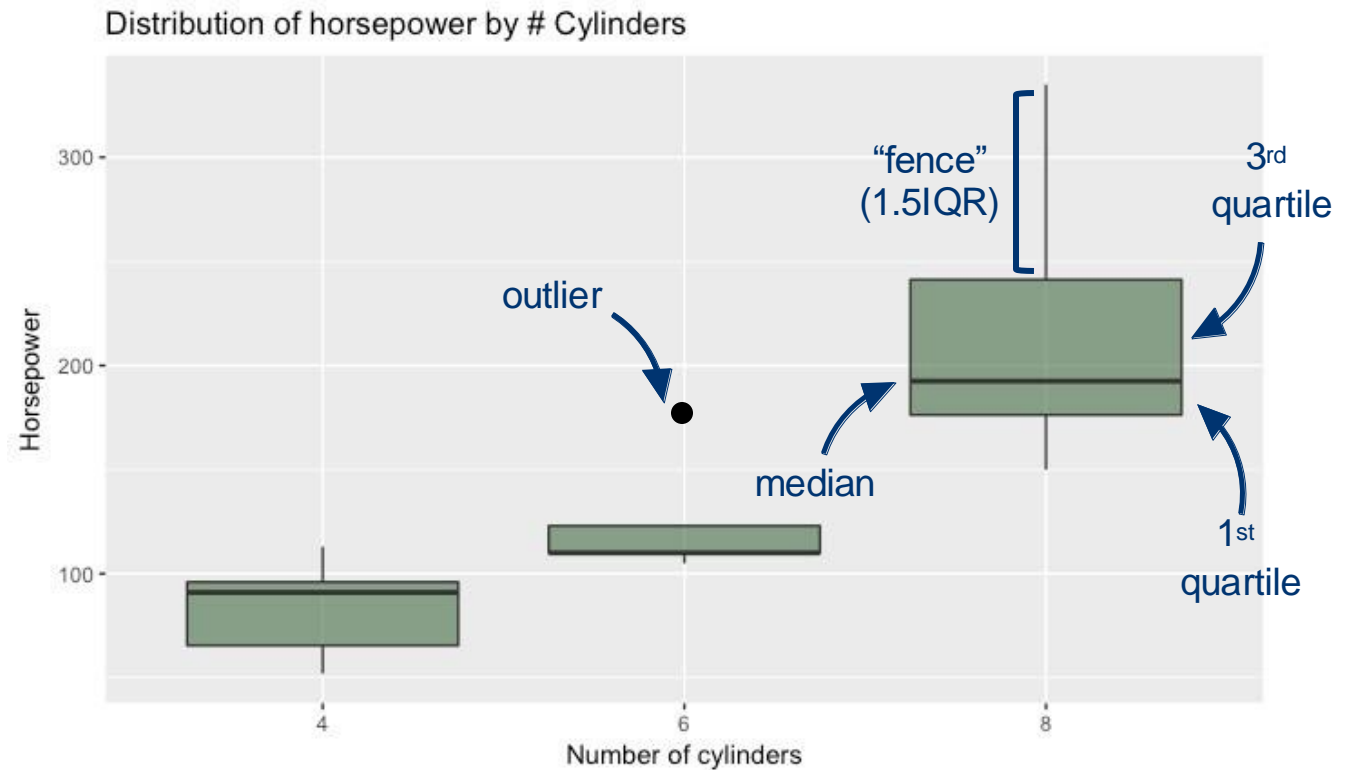
# Histogram

- Used for **one quantitative variable**
- Looks like a bar chart... but the x-axis consists of equal size ranges (bins) for the quantitative variable
- Y-axis shows count or relative frequency
- Highlights **distribution**
- Note: bin size makes a big difference!



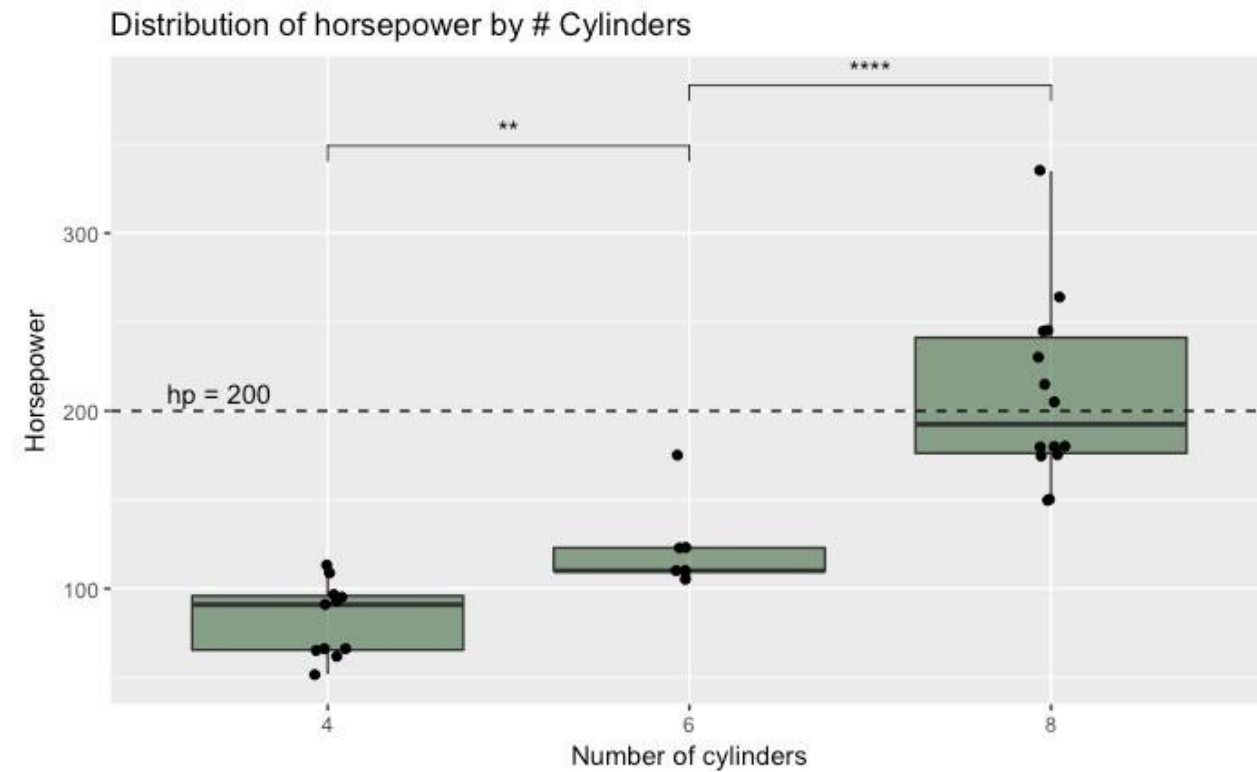
# Boxplot

- Used for **one quantitative variable**
- Also useful for highlighting **distribution**
- Calls out key values:
  - median
  - 1st & 3rd quartiles
  - “fences”
  - outliers



# Boxplot

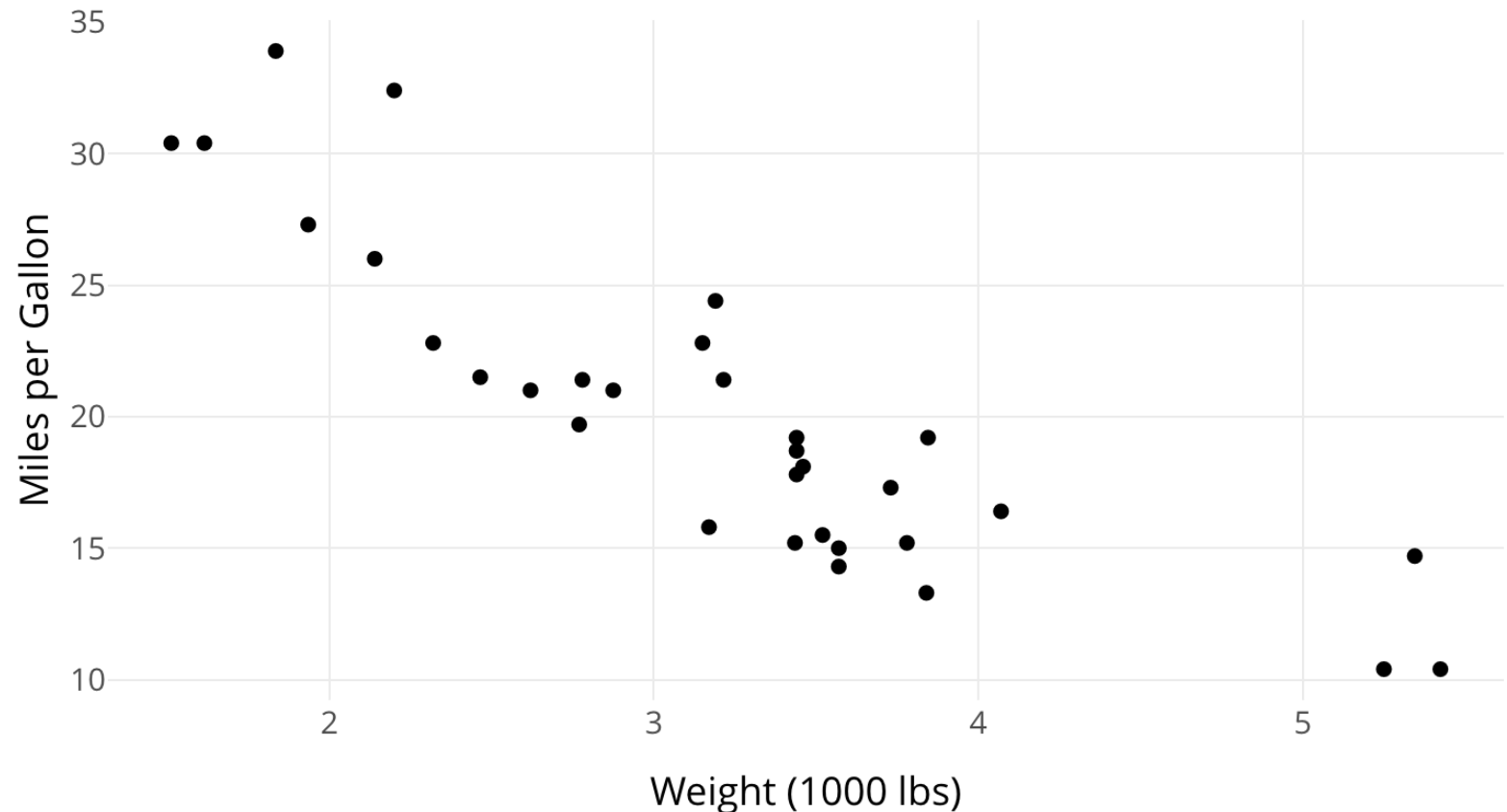
- Use “jitter” to show actual values
- Reference lines can help provide context
- Can use annotations to show statistical significance





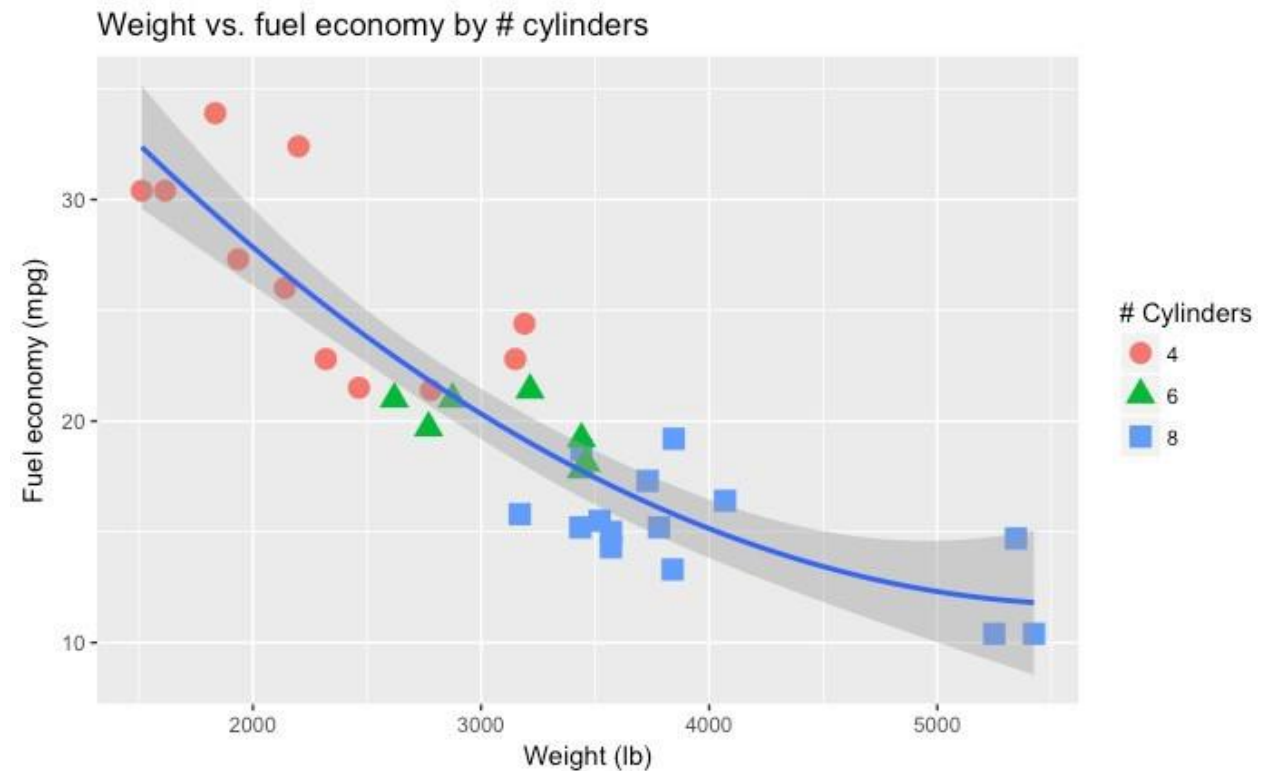
# Scatterplot

- Used to show **one quantitative variable X one quantitative variable**
- Shows the **relationship** between the two variables
- Each point in the plot represents an observation



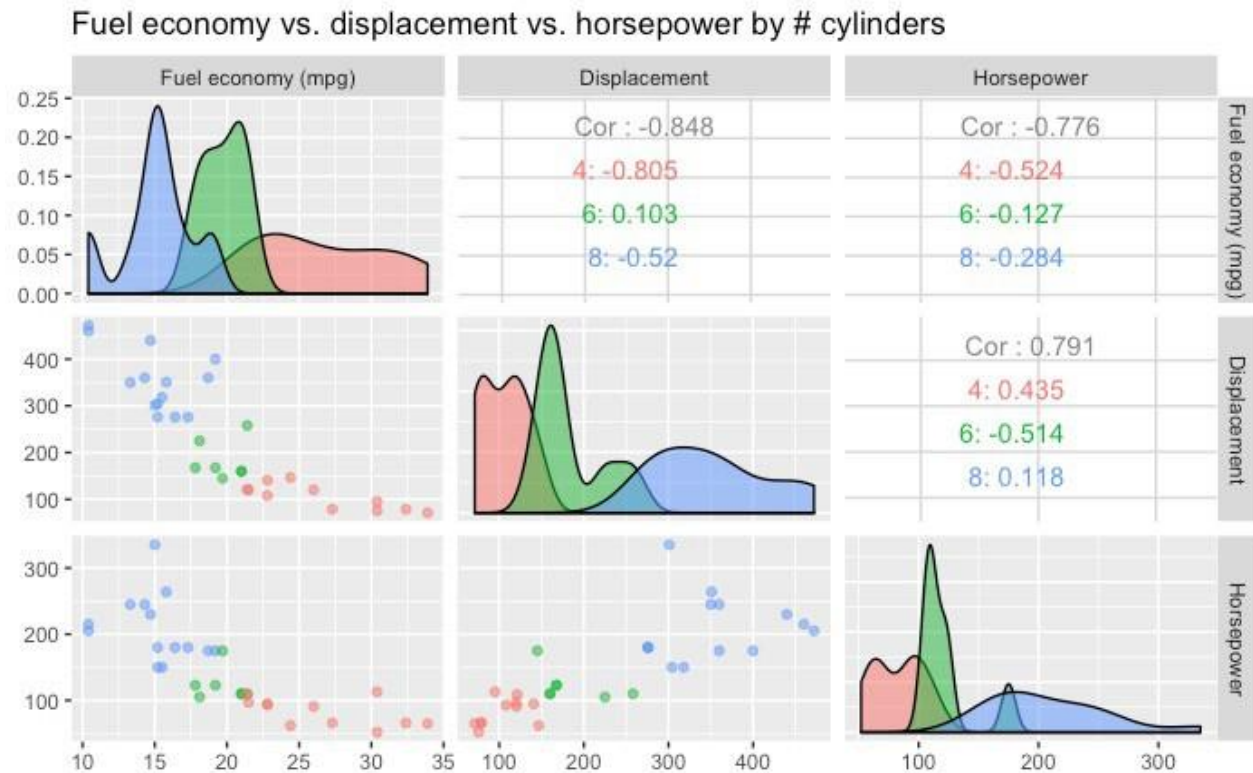
# Scatterplot

- Can use color or symbol to show **one quantitative variable X one quantitative variable X one categorical variable**
- This highlights groups
- Sometimes useful to show a trend line (regression)



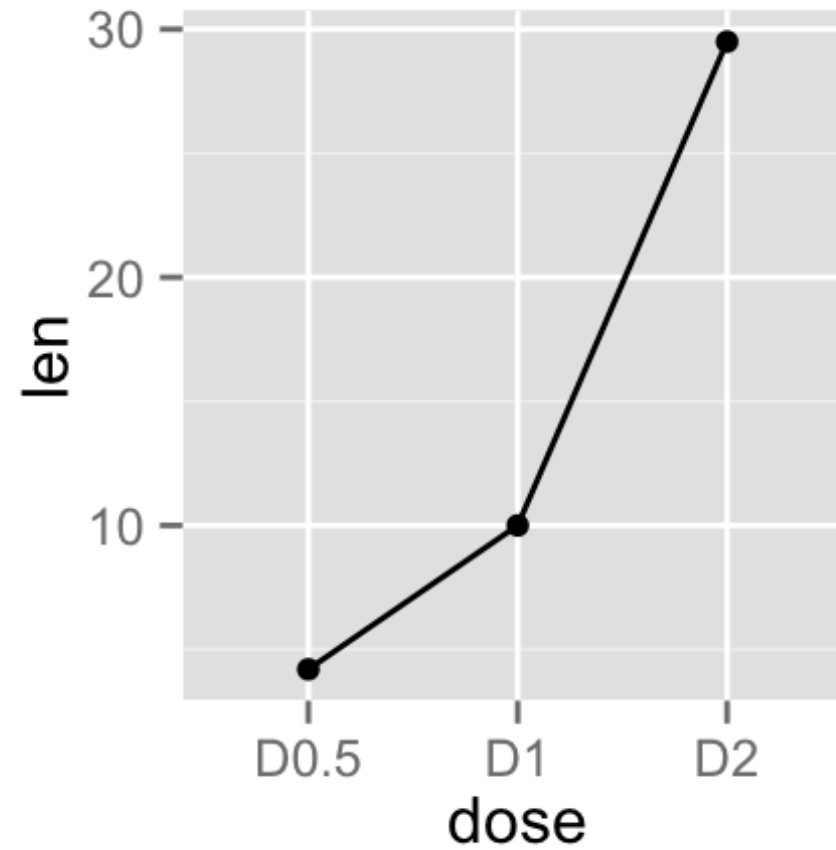
# Scatterplot matrix (SPLOM)

- Can use to show **many combinations of one quantitative variable X one quantitative variable**
- Combines multiple scatterplots into a matrix to show **additional relationships**



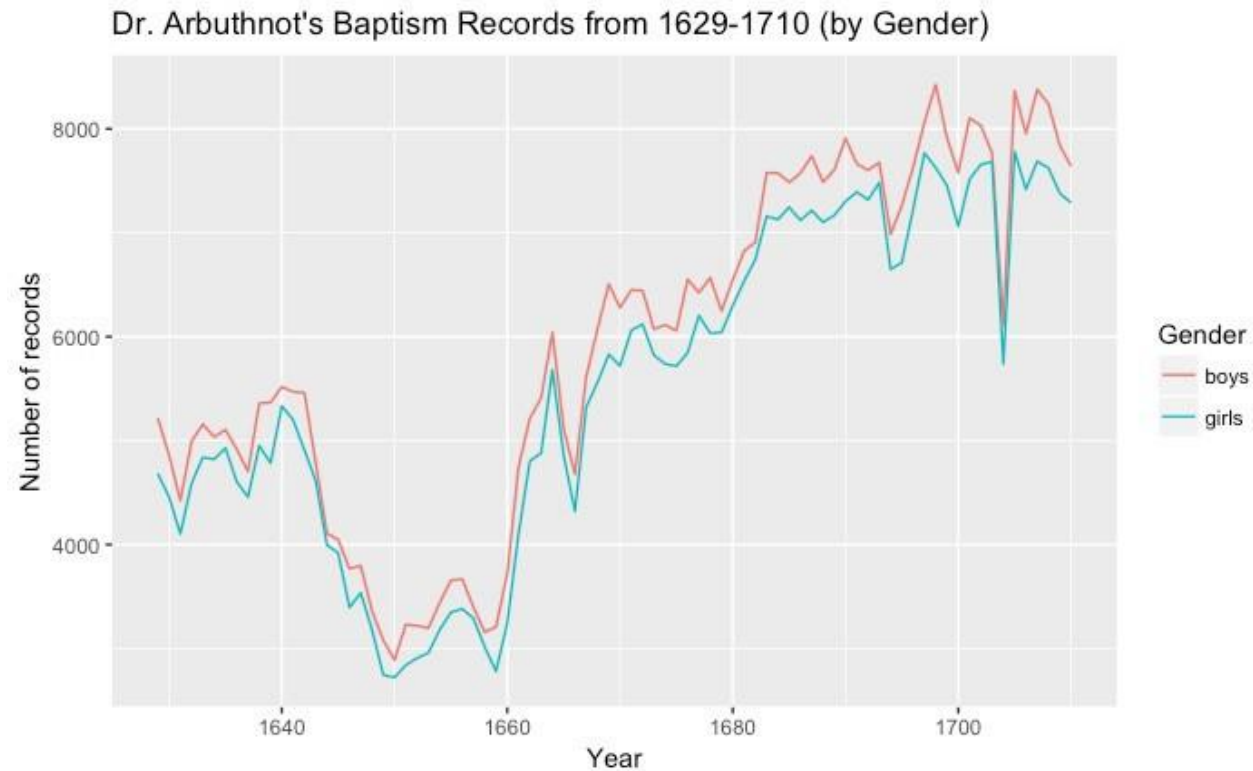
## Line chart

- Shows the **trend of one quantitative variable**, often **over time**

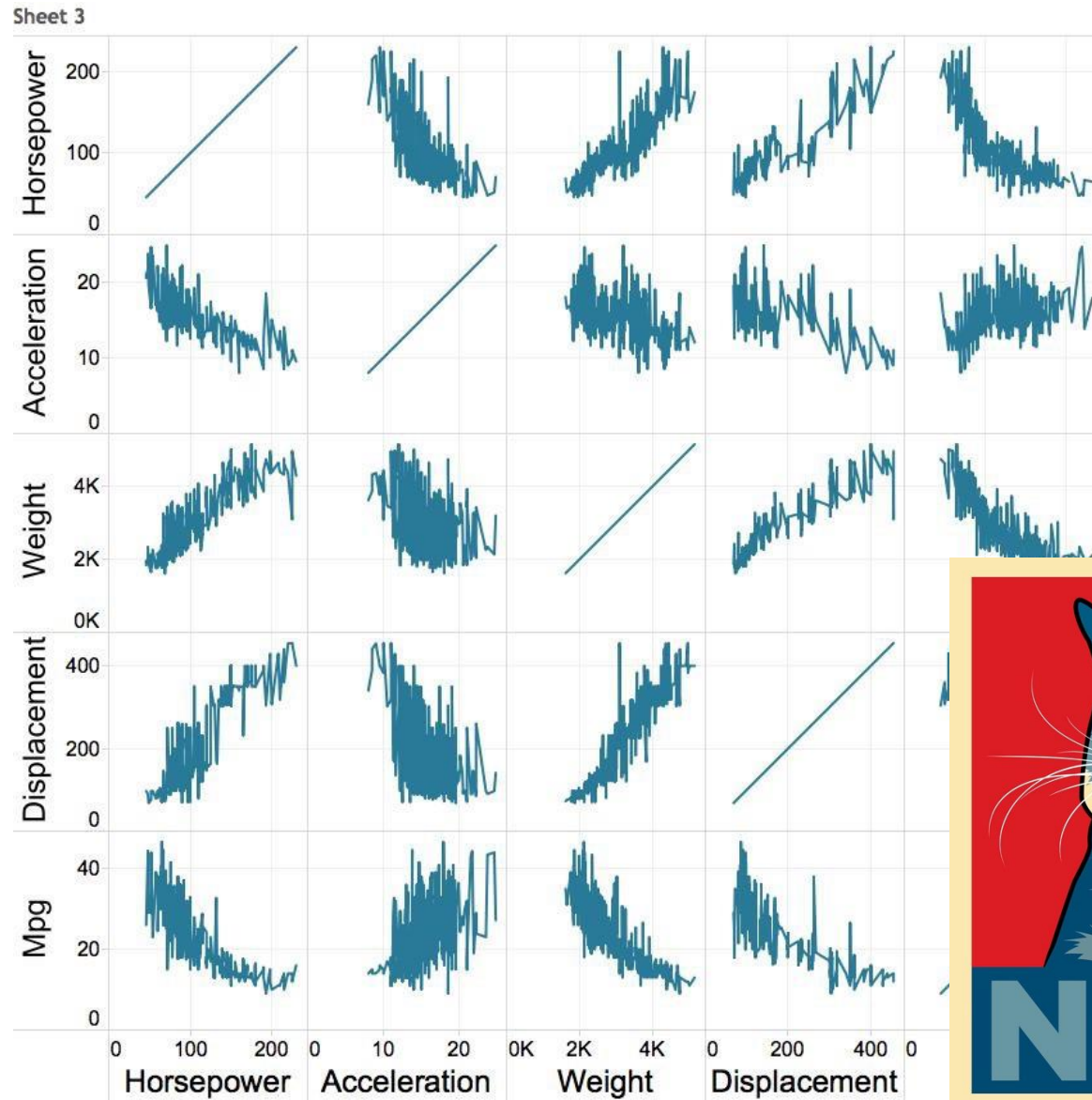


## Line chart

- Multiple lines allow **comparison of trends**
- Can show one quantitative variable across groups, or multiple quantitative variables (if they have the same scale)
- Highlights “position switches”

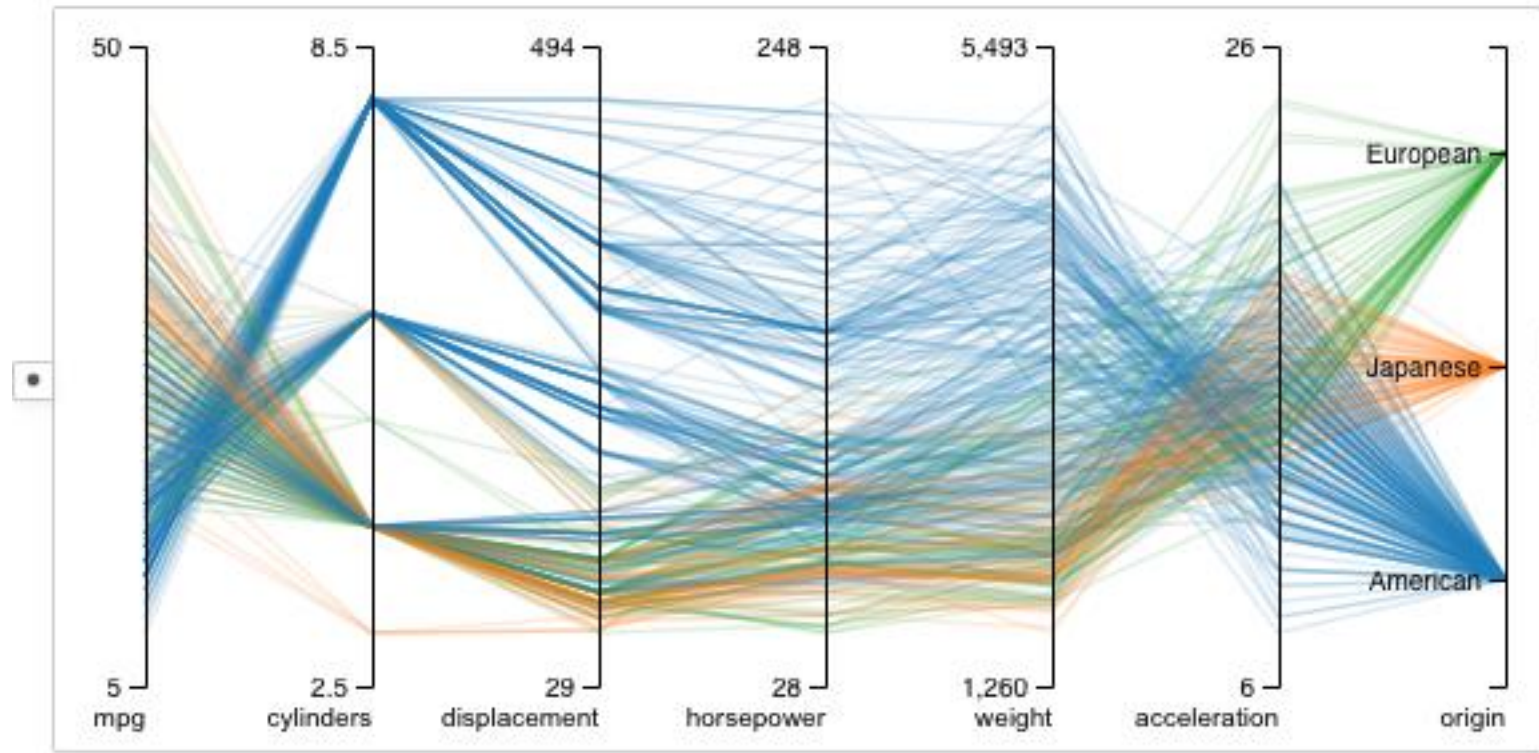


Multiple  
variables: line  
chart matrix?



# Parallel Coordinates Plot

- Supports (pairwise) **comparison of a collection of quantitative variables**
- Each axis represents one variable
  - They may have different scales, typically you normalize them
- Each line represents one observation (connecting the associated values along each axis)
- Axis order matters!



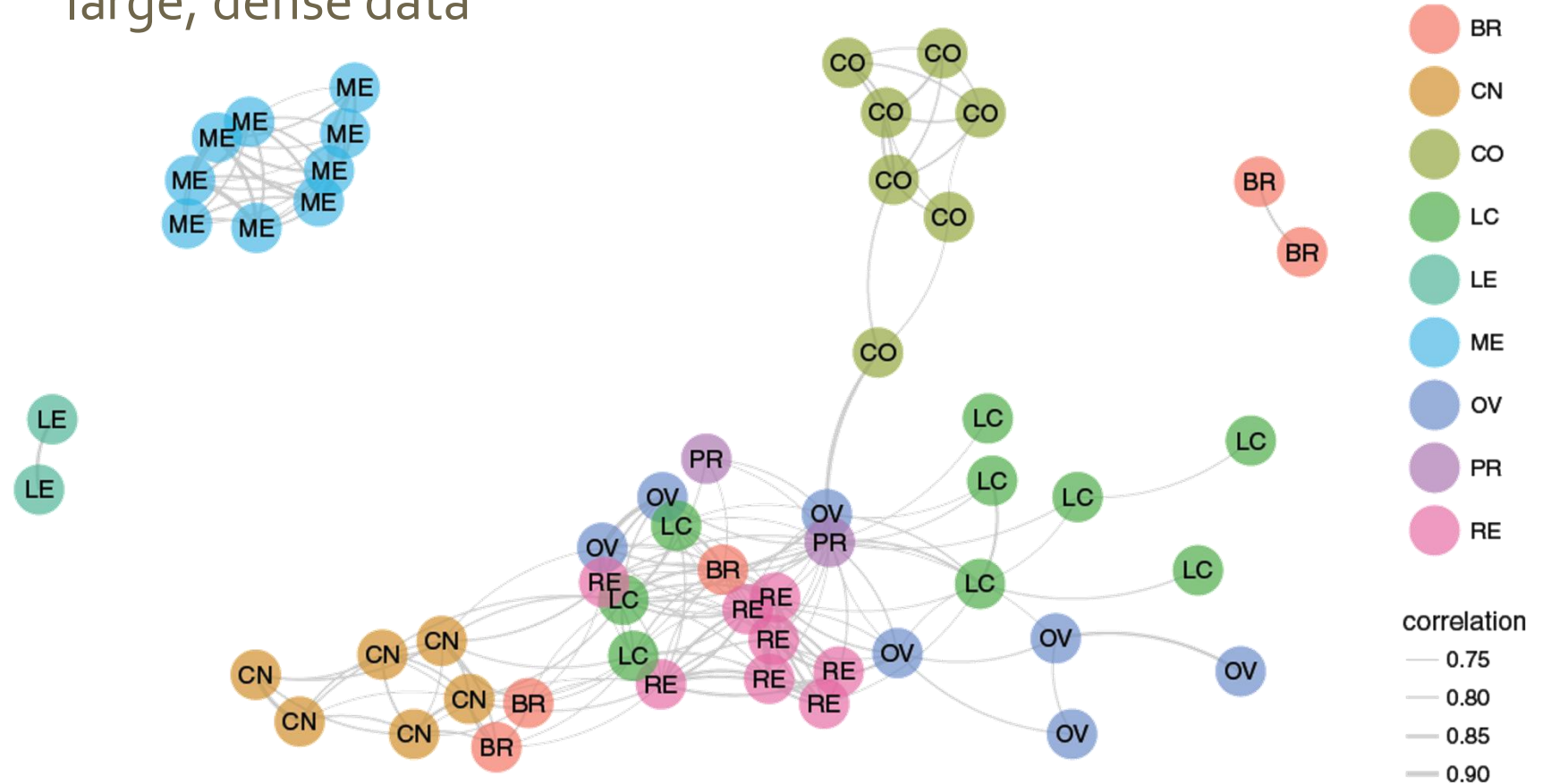
<https://visflow.org/node/visualization/parallel-coordinates.html>

ta-to-parallel.html



# Network

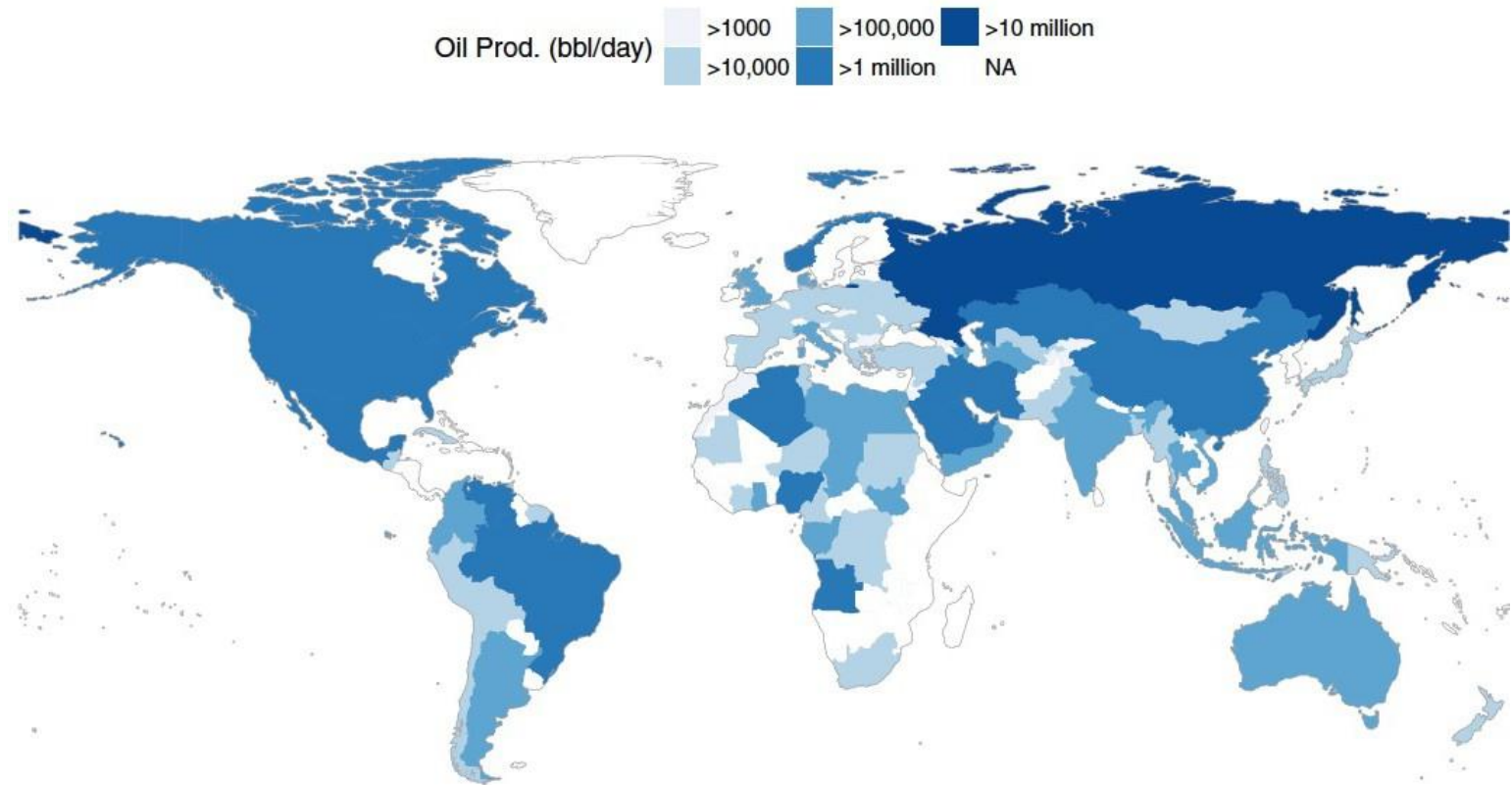
- Shows **network (graph)** data
- Useful for showing the **relationships between entities**
- Can use color, size, etc. to encode additional information
- Caveat: network diagrams quickly become hairballs for large, dense data





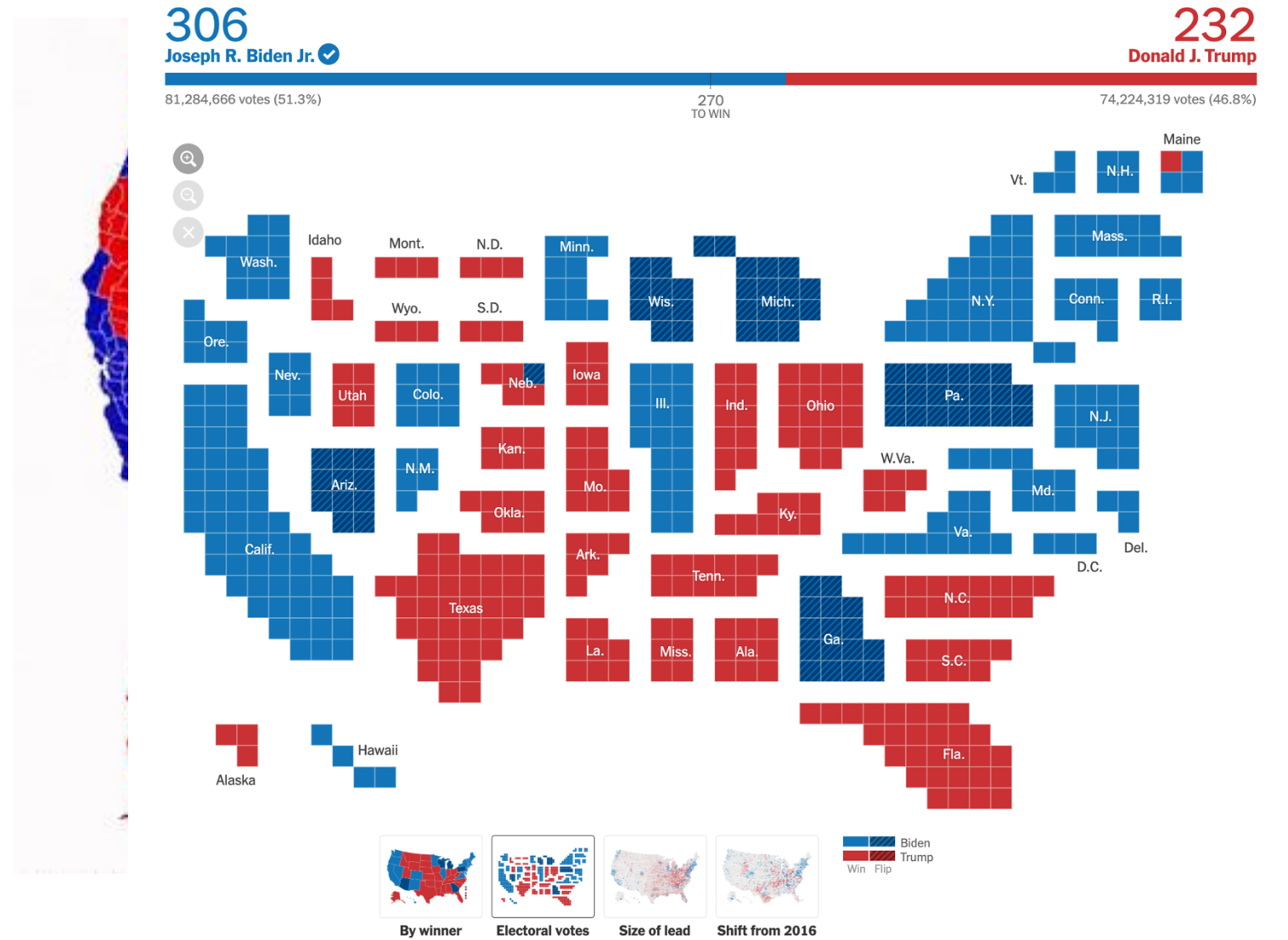
# Map

- Shows **geographic (spatial) data**
- Useful for analysis with a strong geographic component
- Filled maps (choropleth) vs. points
- Remember: color scale comparisons are harder for humans than size comparisons!



# Map

- Remember to map the correct data to your visual channels



## Your Turn!

- Find someone (or two) to work with
- Choose a visual encoding from <https://datavizproject.com/> that you *have not previously encountered*
- What variable types is your visual encoding good for?
- What types of questions do you think this visual encoding would be good for answering?
- Find an example of the visual encoding with real data. What is one interesting thing the visualization shows?