Communicating with Data – Mental Models & Data-Visual Mapping

Dr. Ab Mosca (they/them)

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

Plan for Today

- ~~A quick history lesson~~
- Mental models: how we process information
- Visualization building blocks
- Takeaways

Checking in

- Are you on Slack?
  - All communication will be through Slack (see syllabus for details)
- Did you join Gradescope?

Looking forward

- Hw01 is released today! (Due next week)
- Find instructions on the course website under the "Homework" tab
- Submit on Gradescope – work in pairs or get approval to work individually

Looking forward

- Tableau is a drag-and-drop visualization tool suite
- Tableau for Teaching has donated license keys (good for one year) for everyone enrolled in this course
- Instructions for downloading are on the next slide. Download before class on Tuesday 09/16
- Need help? Ping me on Slack, or come to office hours

Preparing Tableau

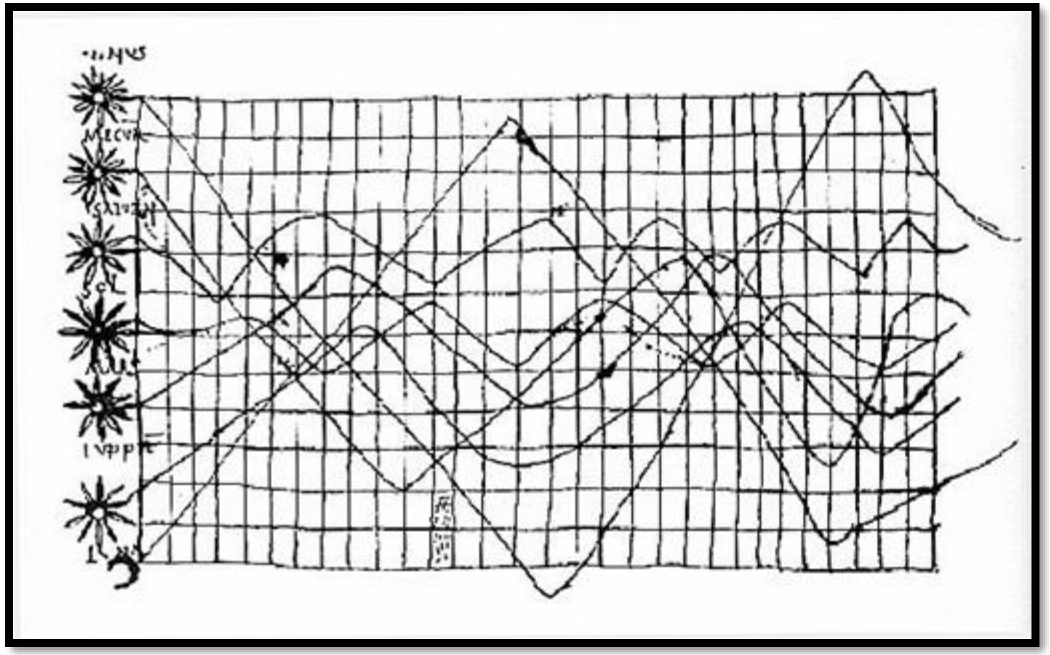
- Sign into an existing Tableau.com account, or create a new account using your school-issued email
- Once signed in, visit the Academic Quick Start page to download the latest versions of Tableau Desktop and Tableau Prep Builder
- Activate with product key: **TC10-80CE-B700-DF1B-AF80**
- Already have a copy of Tableau Desktop installed? Update the license key in the application: Help menu → Manage Product Keys

Having trouble? The Academic Quick Start page includes FAQs and help articles.

(Incomplete) History of Visualization: 15,000BC

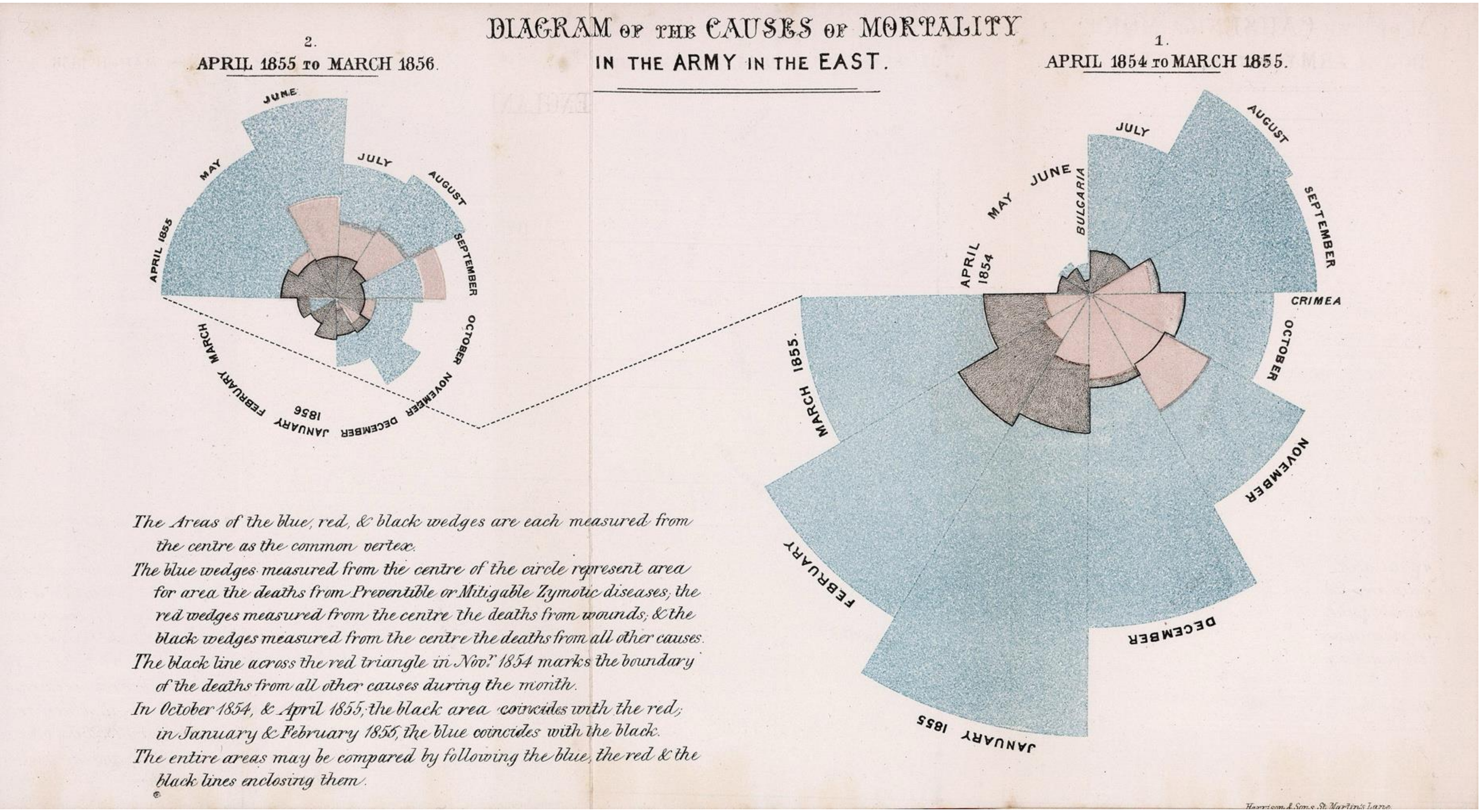
15,000 BC. Laxcaux, France

(Incomplete) History of Visualization: 900s



- Oldest known attempt to show changing values graphically
- Inclinations of the planetary orbits over time

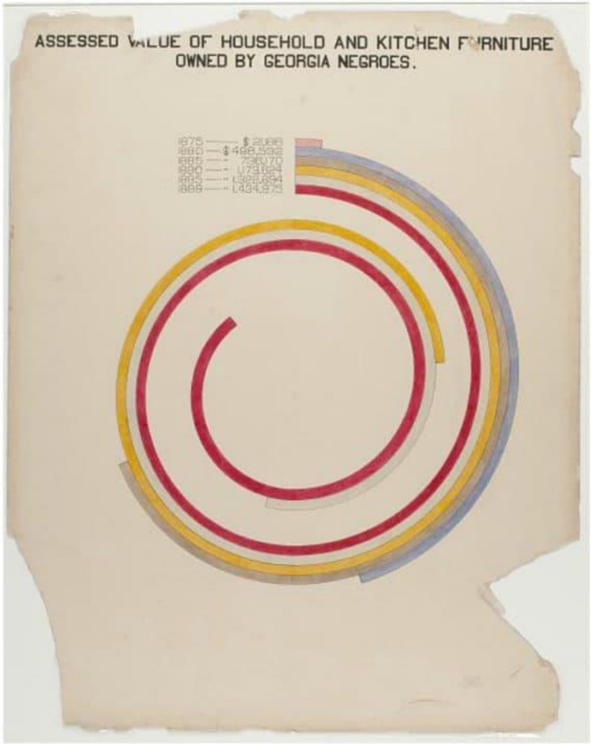
(Incomplete) History of Visualization: mid-1800s

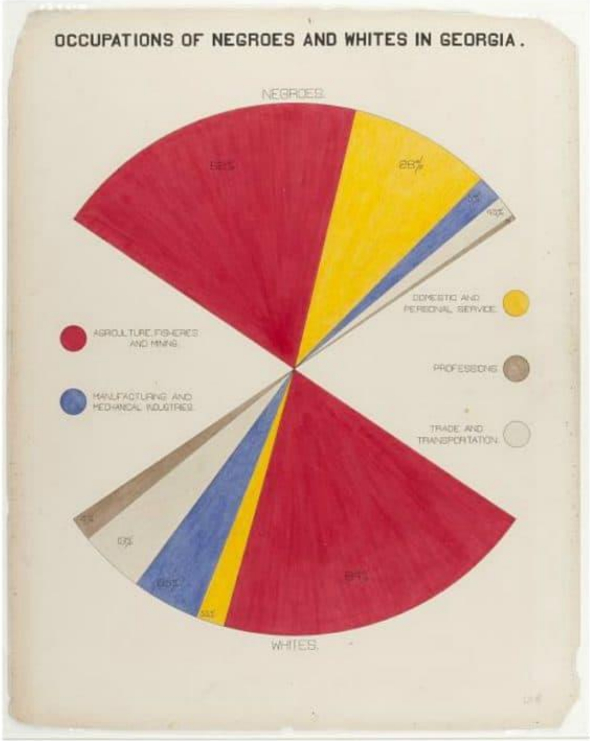


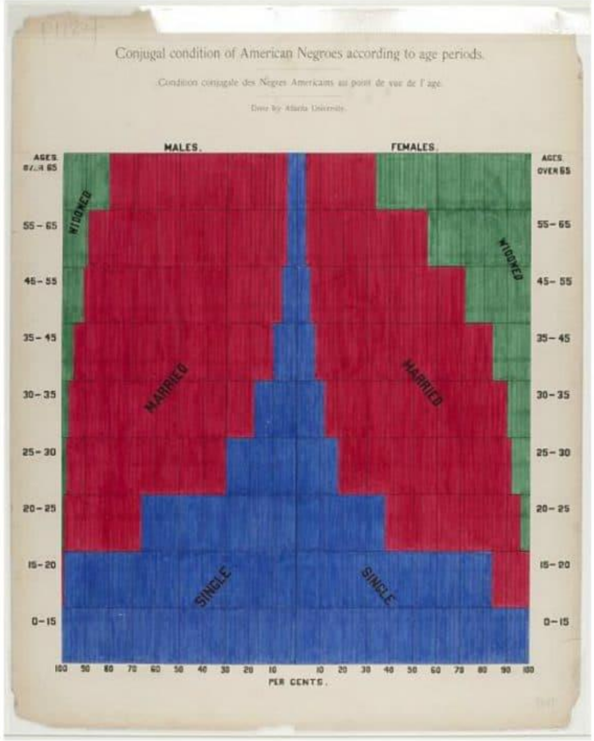
- Florence Nightingale
- Statistician, early adopter of visualization, founder of modern nursing
- Used this visualization to argue for better sanitary practices

https://en.wikipedia.org/wiki/Florence_Nightingale

(Incomplete) History of Visualization: mid-1800s



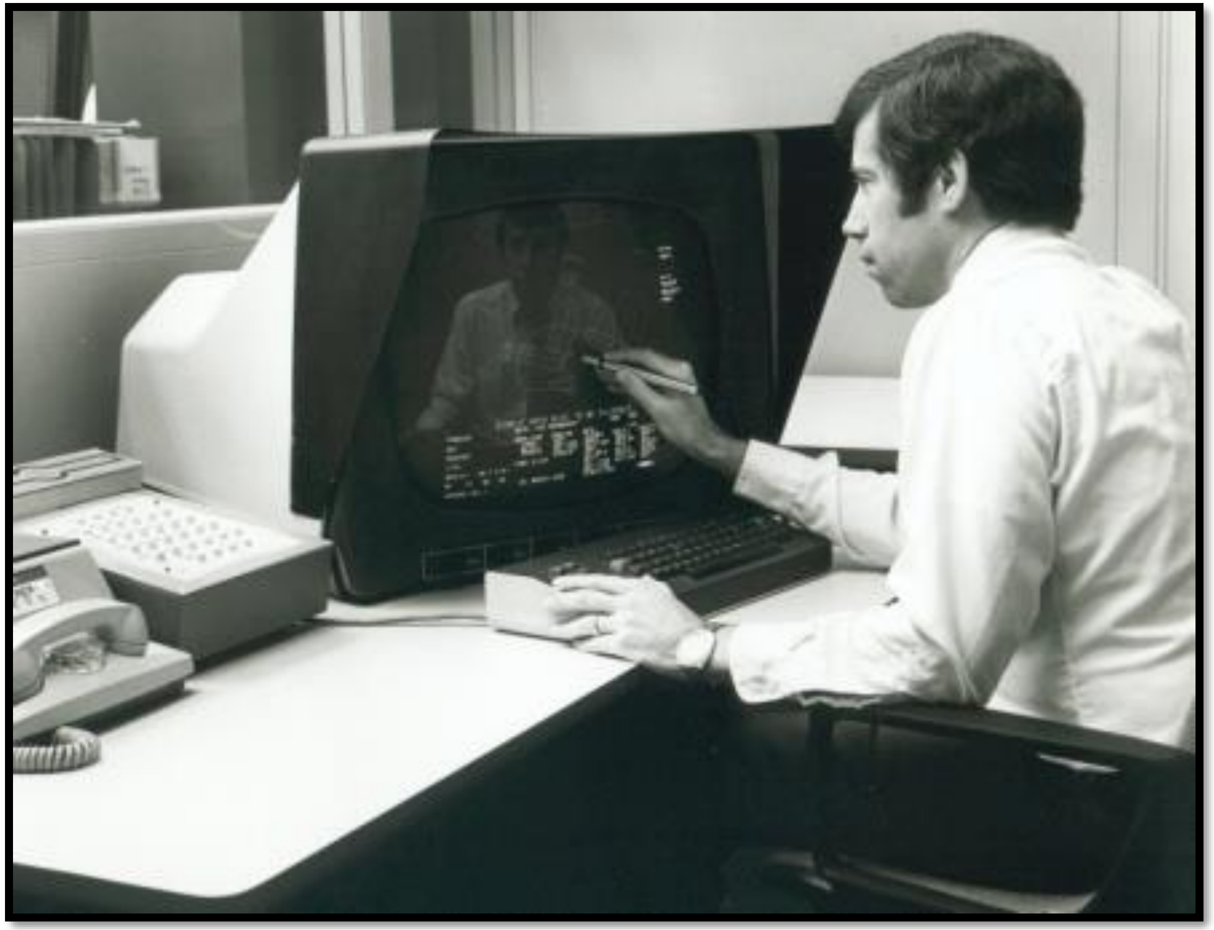




- W.E.B. Du Bois
- Sociologist
- "Data Portraits" were displayed at the Paris Exposition in 1900 to challenge norms and show how Black folks fit into American progress

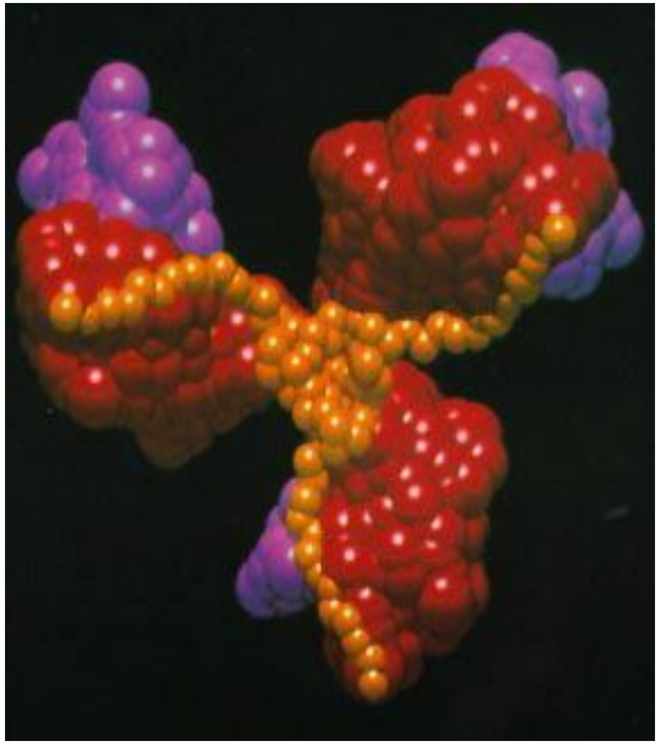
https://www.wbur.org/news/2019/02/21/web-du-bois-infographics-humanity-african-american

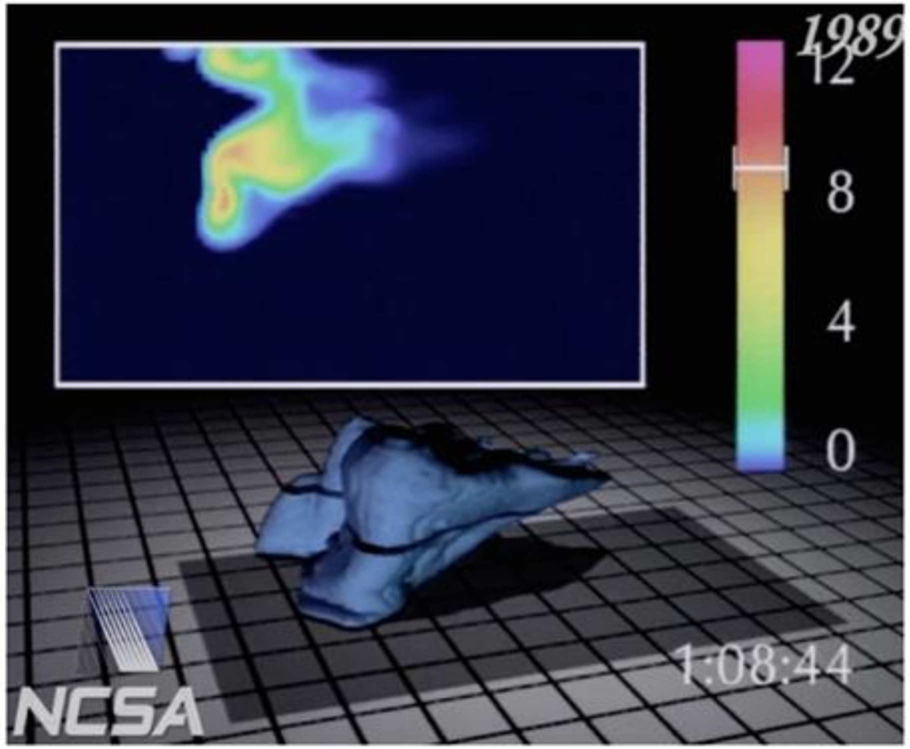
(Incomplete) History of Visualization: 1970s



- CAD/CAM, building cars, planes, chips
- Starting to think about: 3D, animation, edu, medicine

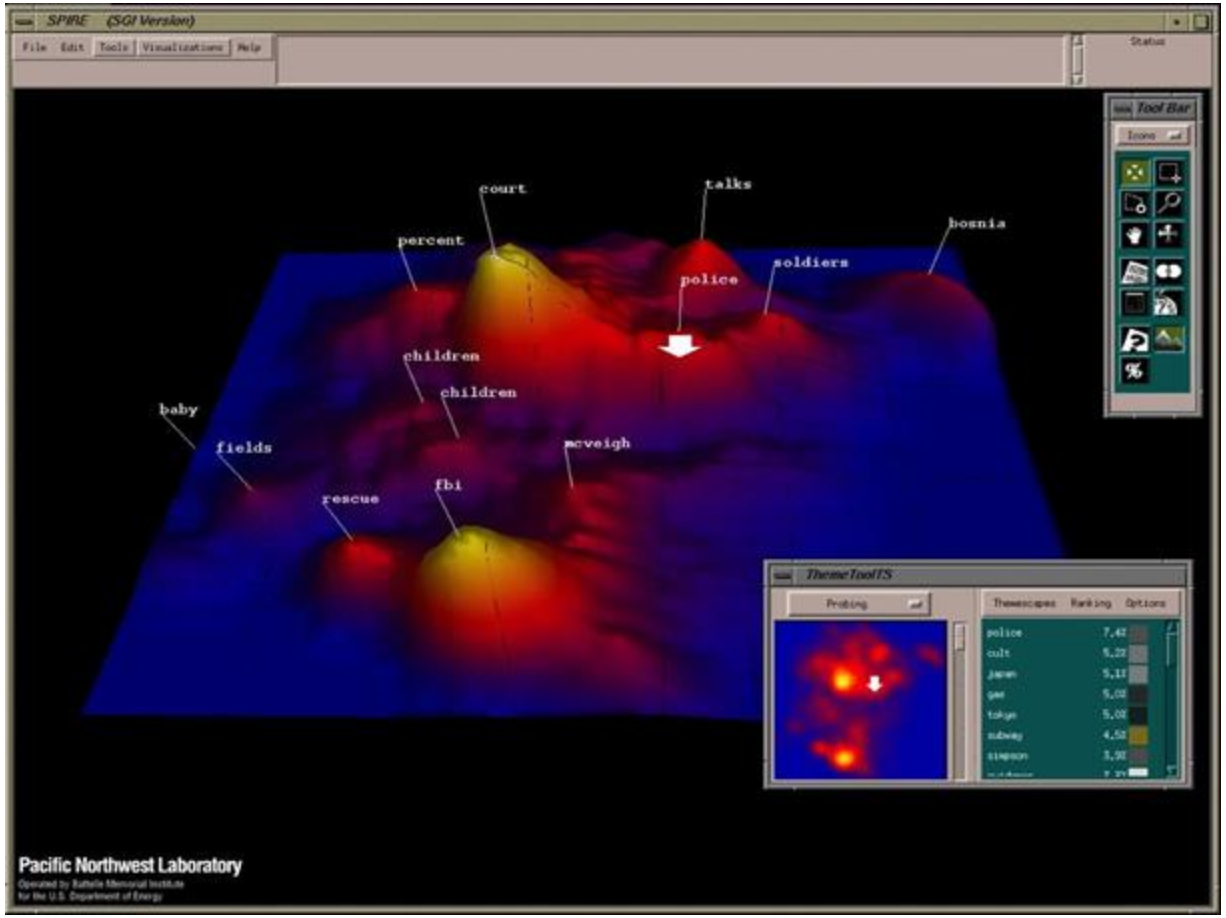
(Incomplete) History of Visualization: 1980s





- Scientific visualization, physical phenomena
- Starting to think about: photorealism, entertainment

(Incomplete) History of Visualization: 1990s

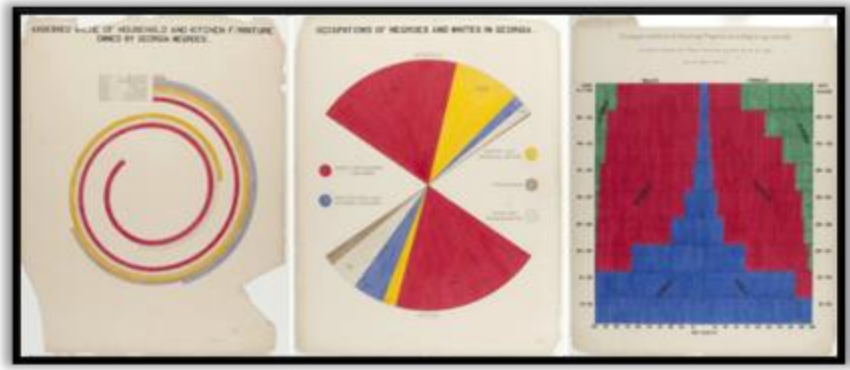


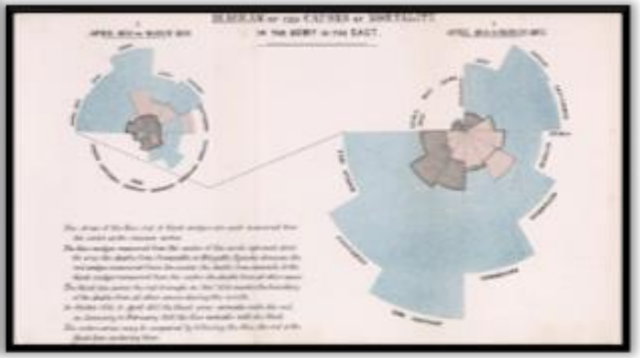
- Information visualization, storytelling
- Starting to think about: human cognition, interaction

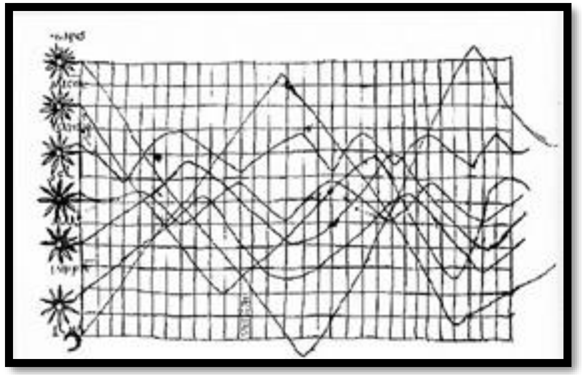
(Incomplete) History of Visualization: 2000s

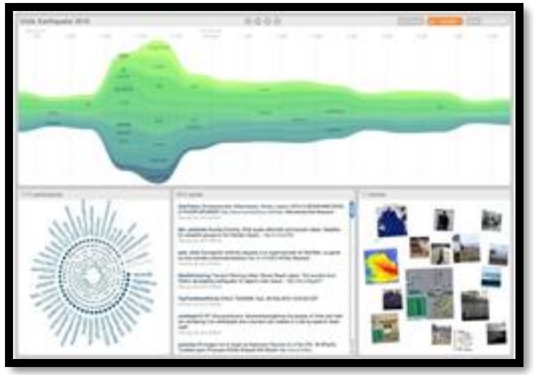
- Coordination across multiple views, interaction
- Starting to think about: sensemaking, provenance

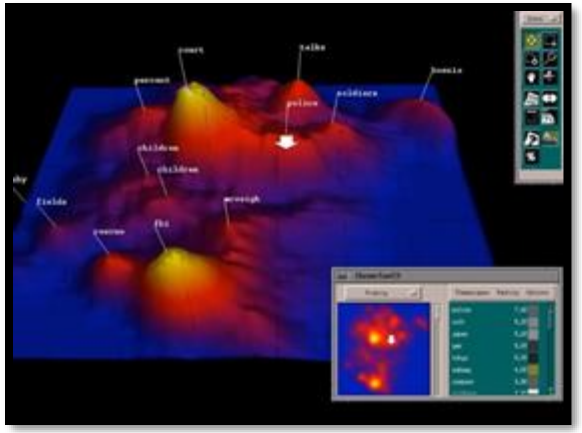
Discussion: what are they all trying to do?

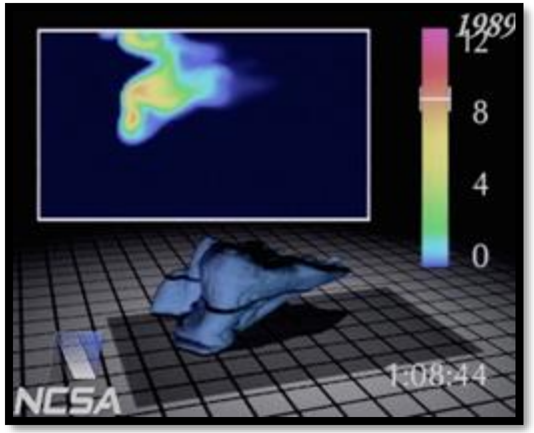




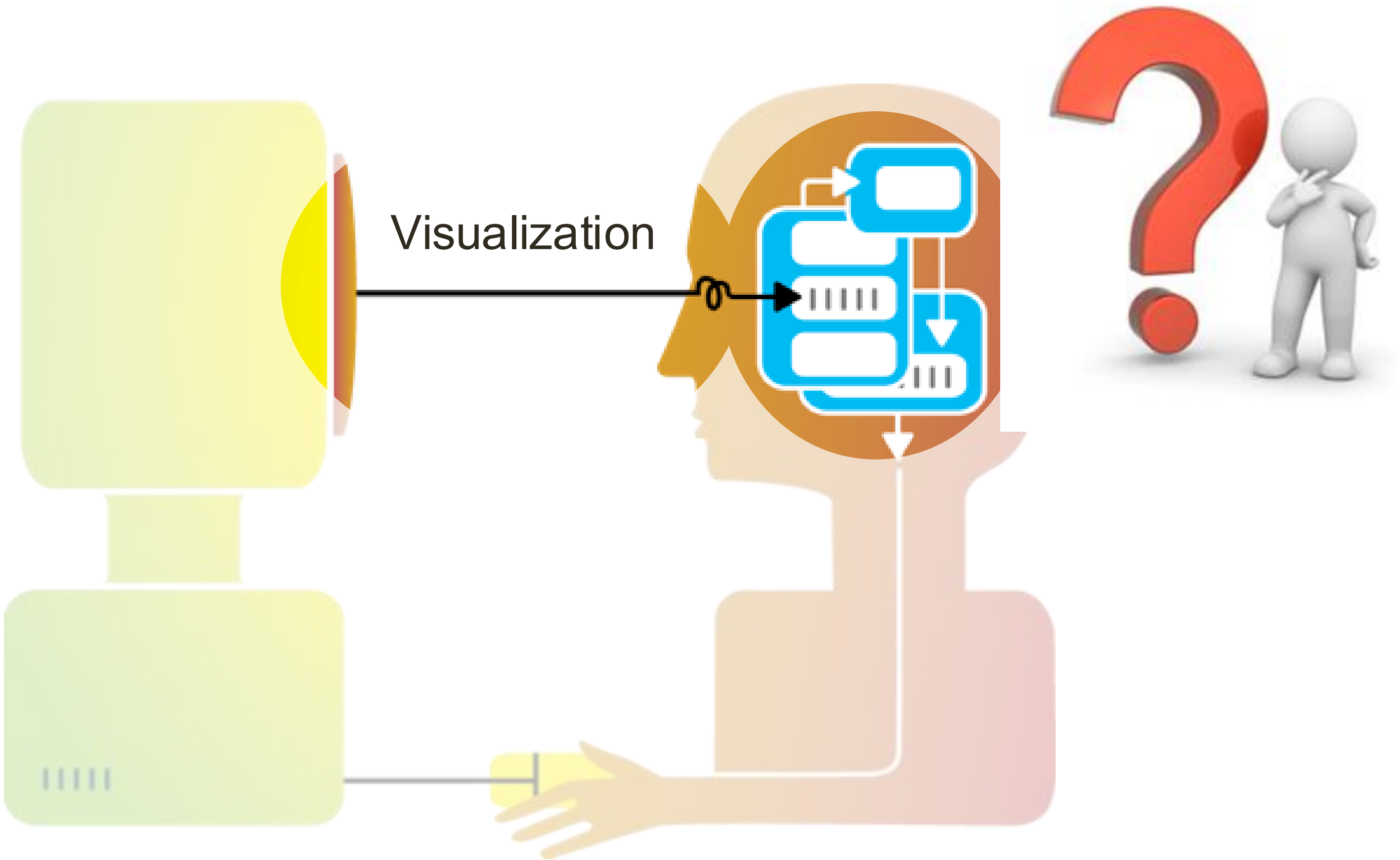




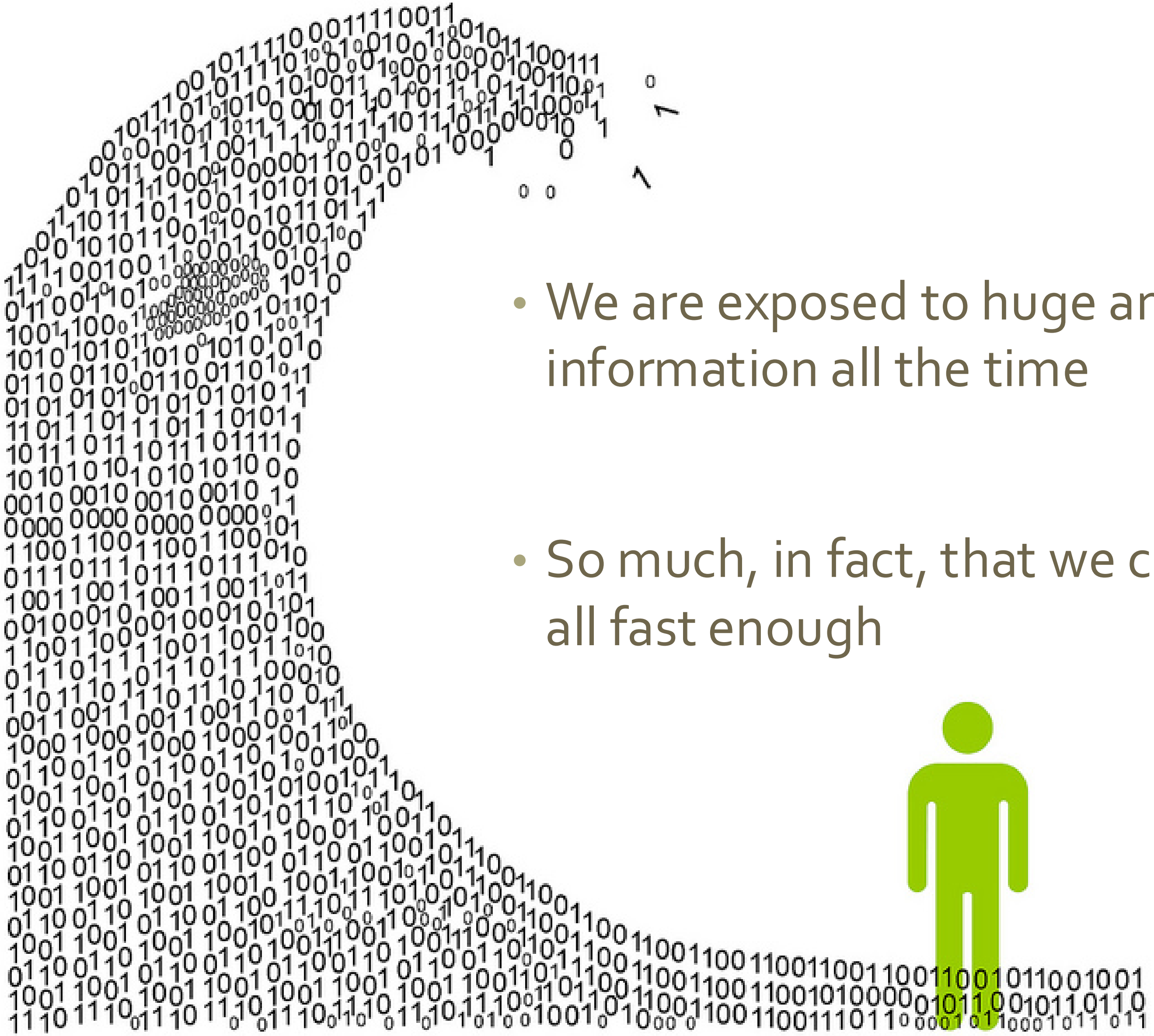




Visualization helps shape *mental models*



Information overload



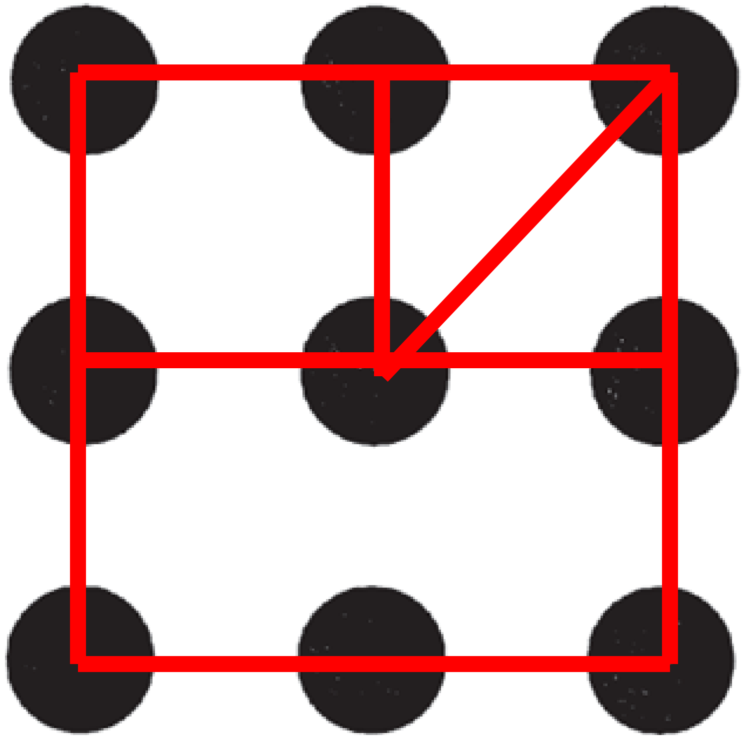
- We are exposed to huge amounts of information all the time
- So much, in fact, that we can't process it all fast enough

Mental models

To cope, we construct **mental models:** abstracted, simplified versions of the world that are more manageable

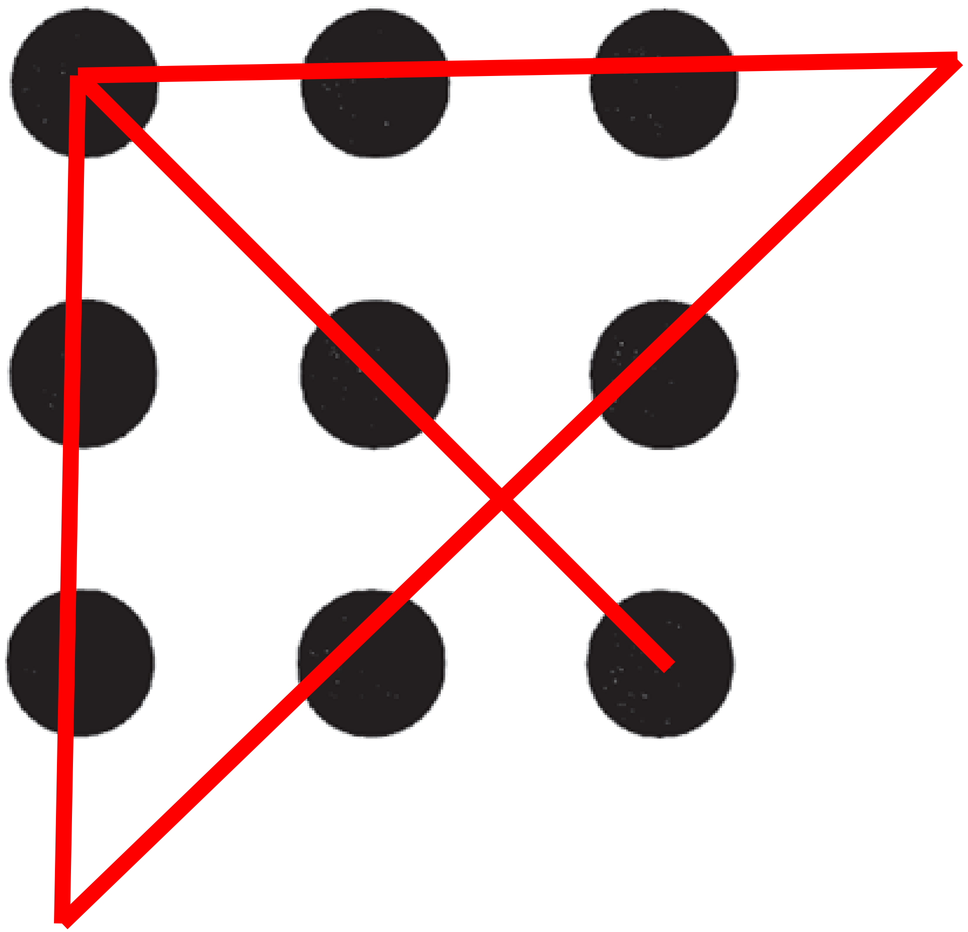
The 9-dot problem

Task 1: Connect all 9 dots using only straight lines



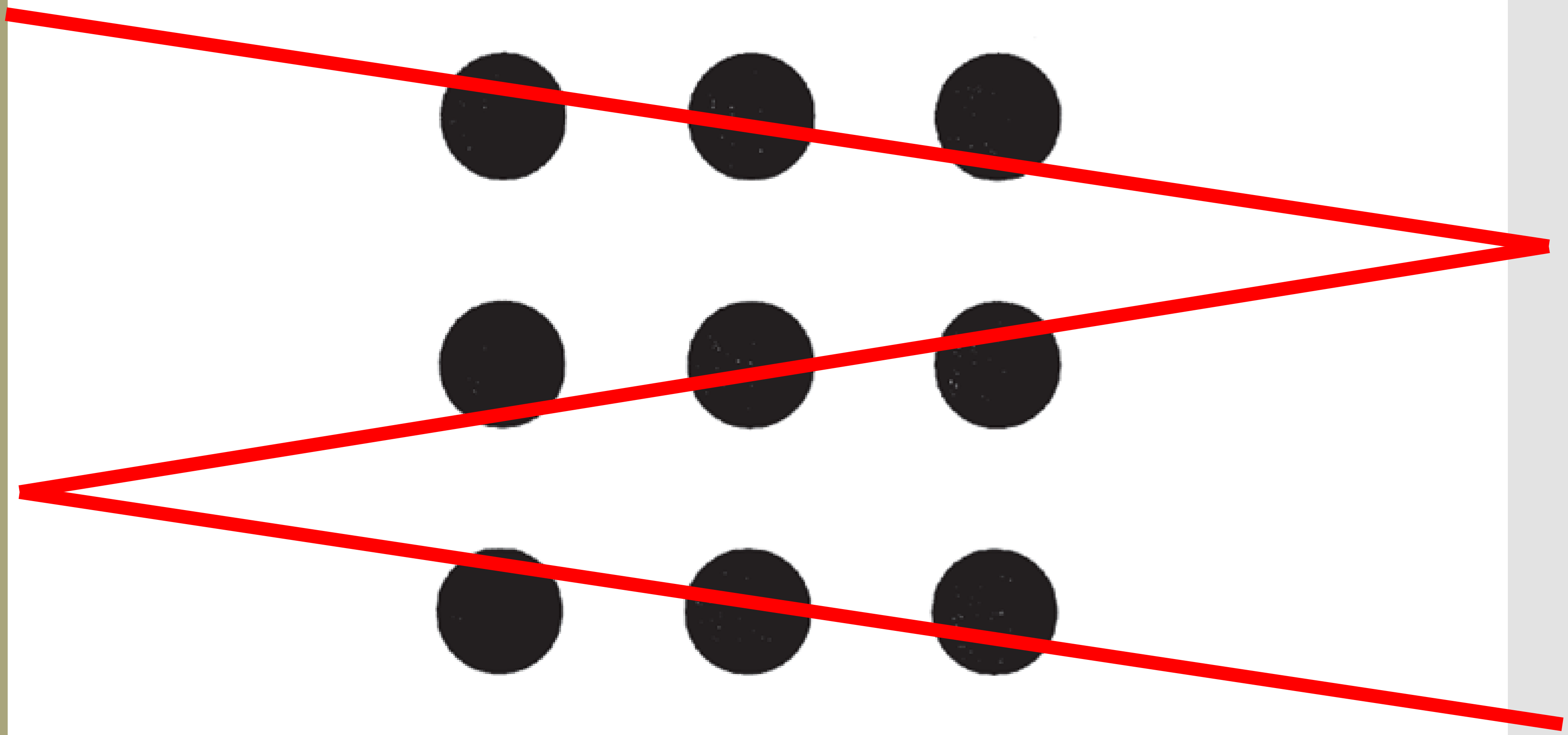
The 9-dot problem

Task 2: Connect all 9 dots using 4 straight lines



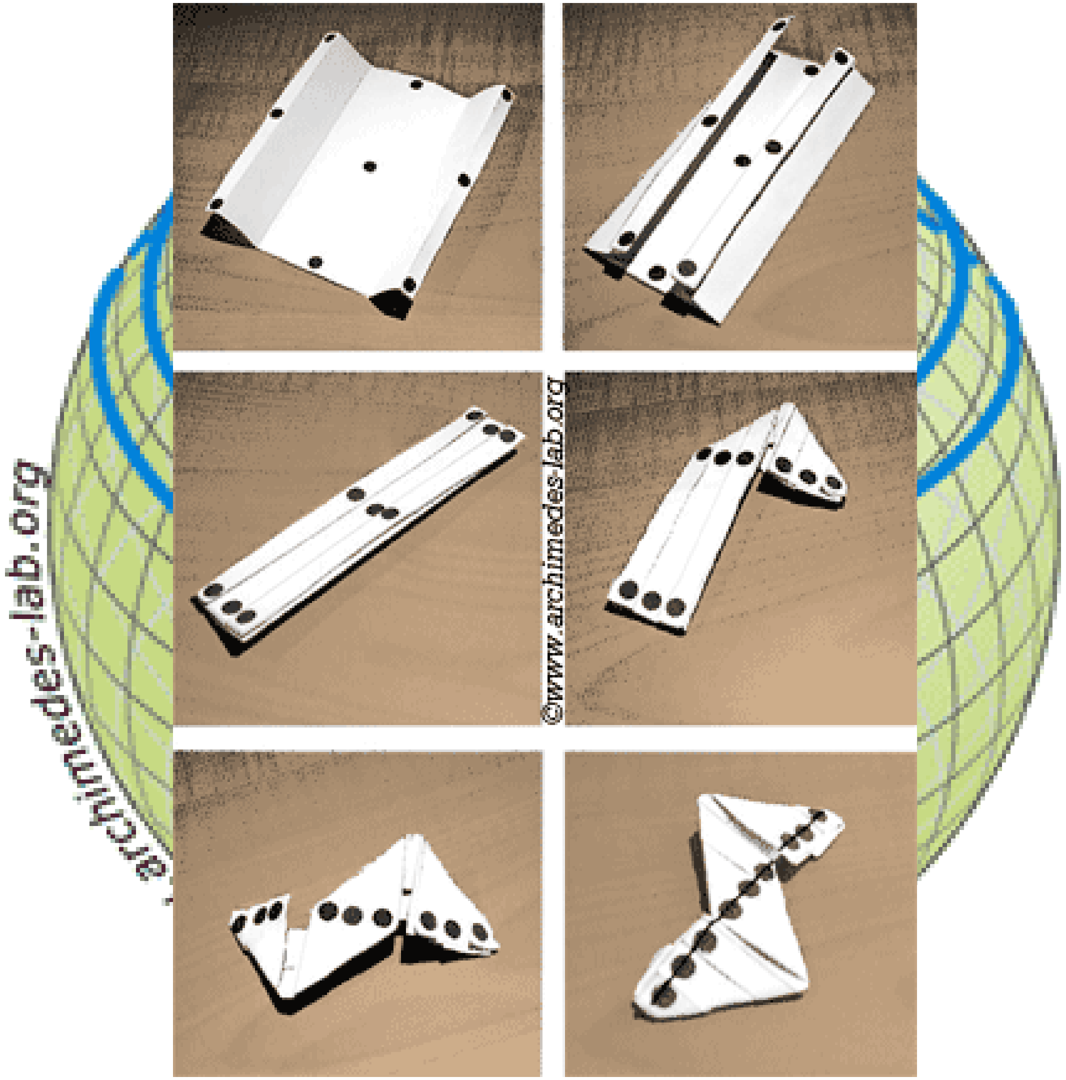
The 9-dot problem

Task 3: Connect all 9 dots using 3 straight lines



The 9-dot problem

Task 4: Connect all 9 dots using 1 straight line



Mental Models: a Sketch

Mental Models

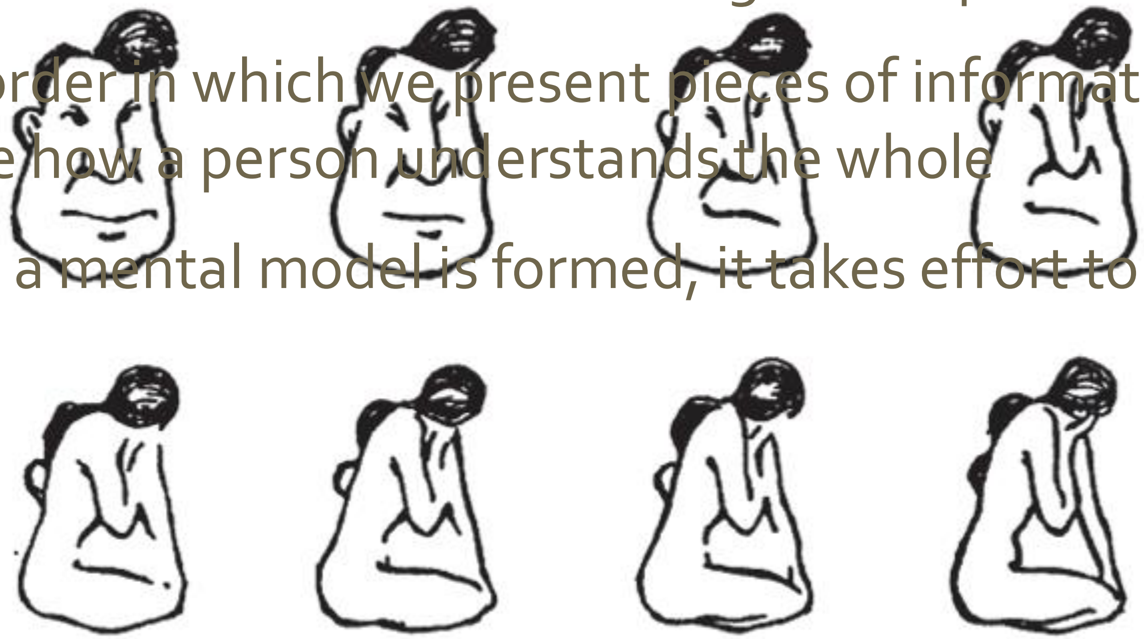
1. We tend to see what we expect to see

- Mental models are built from prior experience
- We expect new input to “fit” the existing model
- Updates are **expensive**: given input that almost fits, we’ll distort information to avoid re-fitting the model
- **Expectation** is at least as strong as perception

Mental Models

2. Mental models form quickly, & update slowly

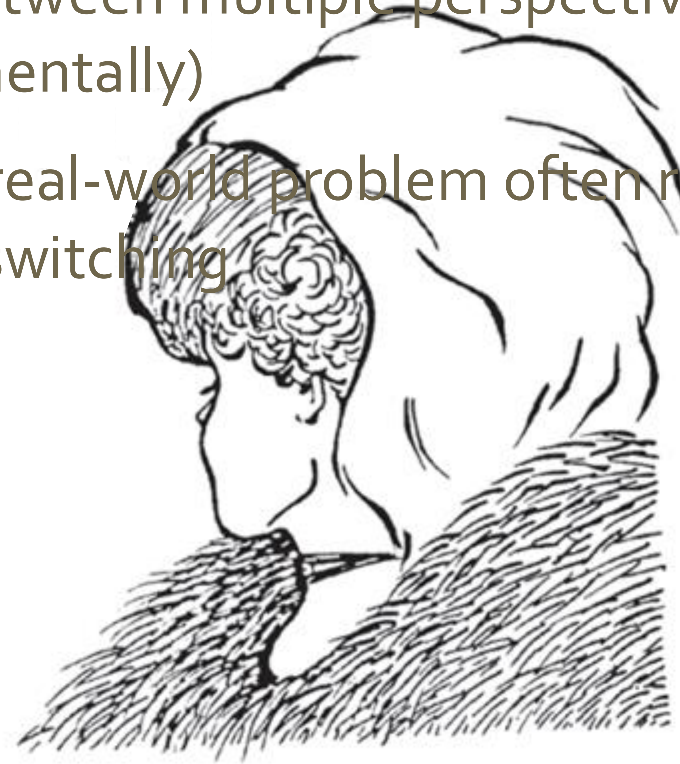
- “First impressions matter”
- Early information can have the highest impact
- The order in which we present pieces of information can shape how a person understands the whole
- Once a mental model is formed, it takes effort to alter it



Mental Models

3. New information gets incorporated into the existing model

- Integrating competing perspectives is challenging
- Switching between multiple perspectives is also difficult (visually or mentally)
- **Tricky part**: real-world problem often require such perspective switching



Mental Models

4. Initial exposure interferes with accurate perception

Blur size

128px
64px
32px
16px
8px
None

Mental Models

4. Initial exposure interferes with accurate perception

- Longer exposure to ambiguous data makes people **more confident** in their initial model
- This is true even if new data presents strong evidence that their model is **wrong**!
- Important: need to be intentional when we design, because incremental information can be **misleading**

Mental Models

The good:

- Well-tuned mental models let us process information quickly
- Frees up more processing power to synthesize information

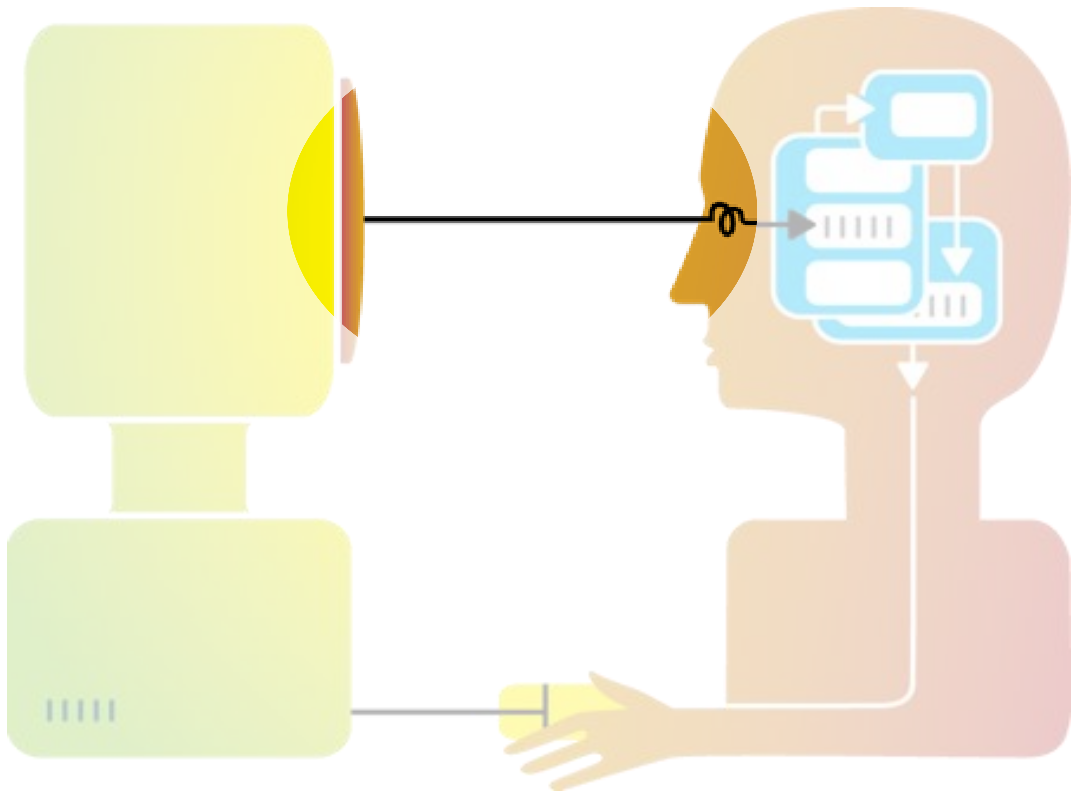
The bad:

- People (esp. experts) tend not to notice information that contradicts their mental model
- A "fresh pair of eyes" can be beneficial

The ugly:

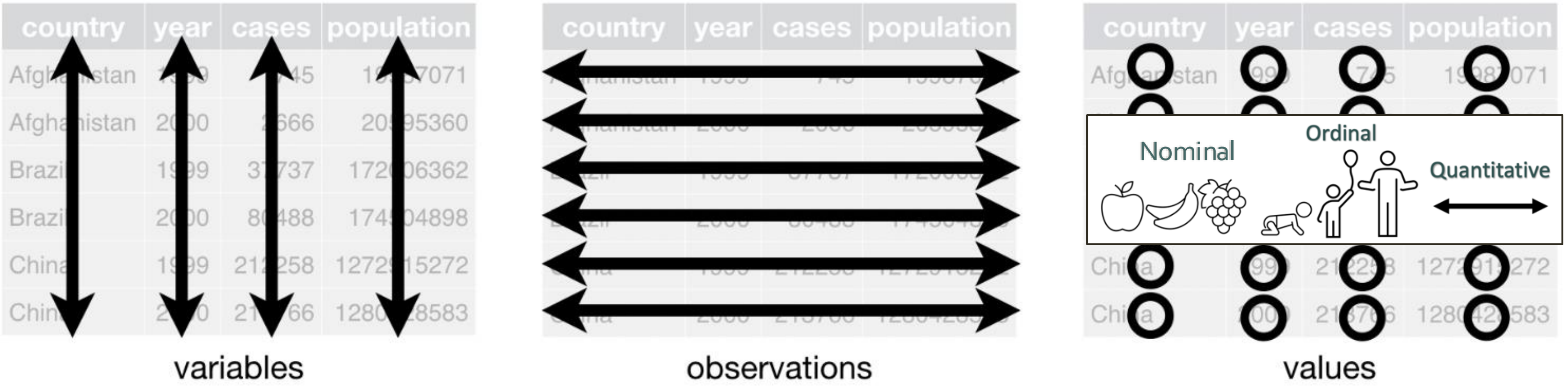
- Mental models are unavoidable: everyone has them, and they're all different
- **Key:** be aware of how mental models form, how they shape perception, and how to support (or challenge) them

So what do we have to work with?



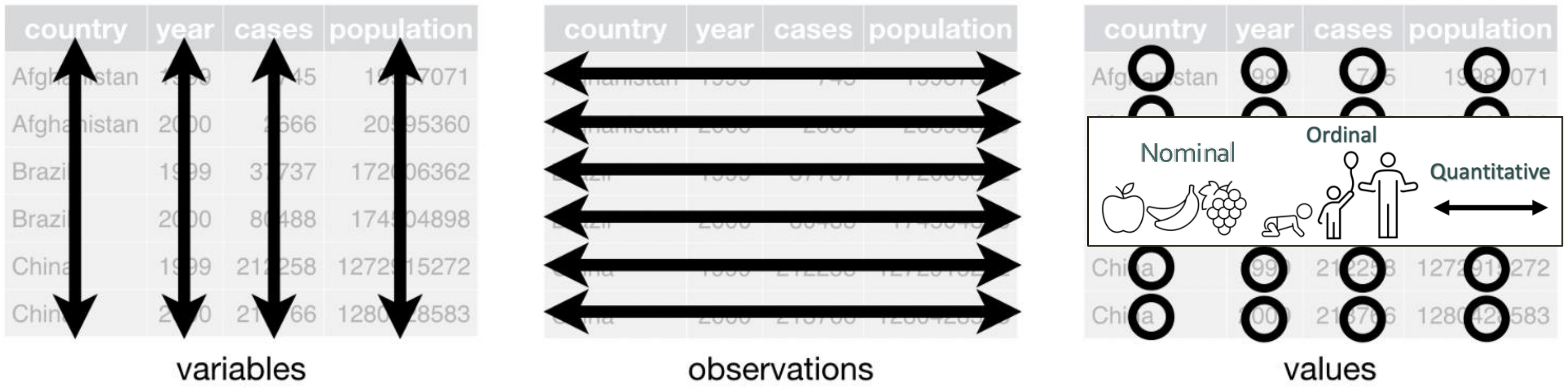
Data

- Remember…



Data → Visuals

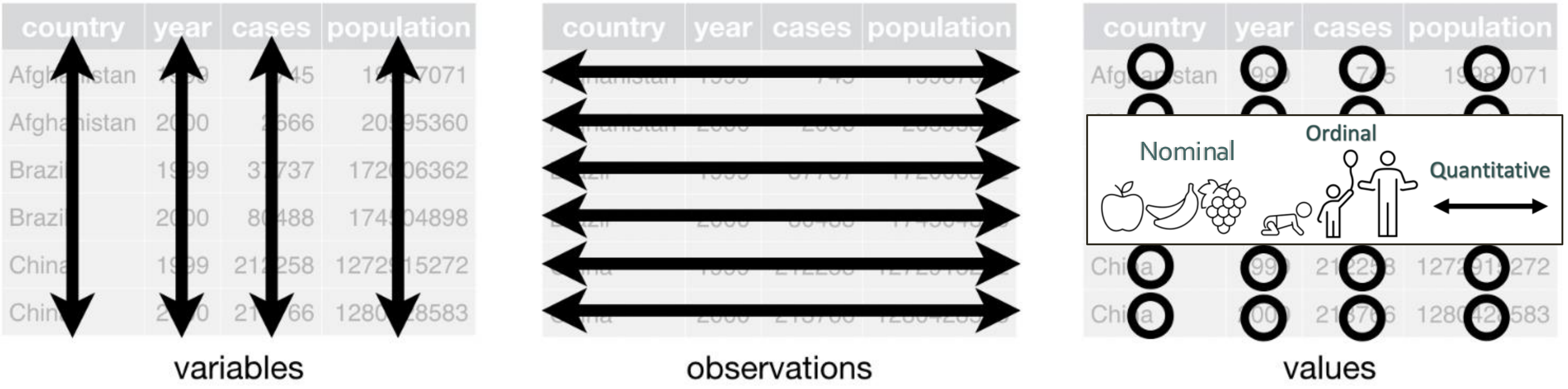
- Remember…



- **Big idea behind visualization**
  - Data have dimensions
  - Visualizations have dimensions, too
  - To build good visualizations, we need to **map data dimensions to visual dimensions** in a principled way

Data → Visuals

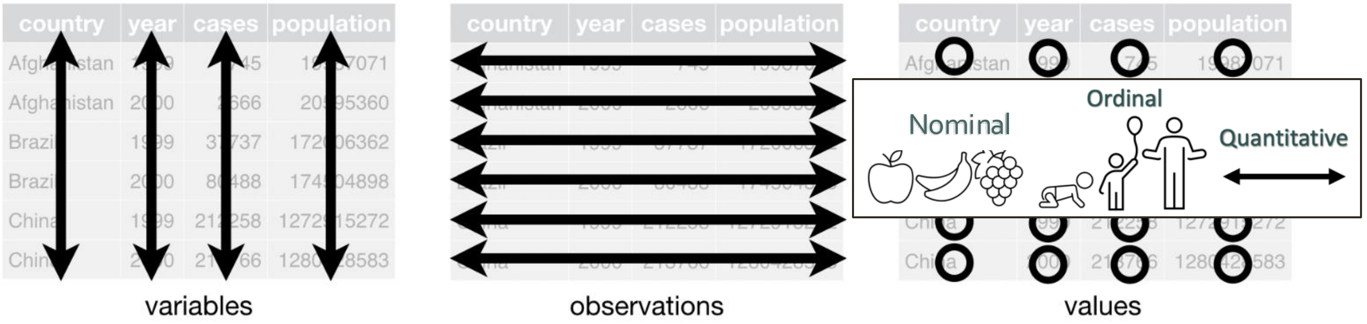
- Remember…



- **Big idea behind visualization**
  - Data have dimensions
  - Visualizations have dimensions, too
  - To build good visualizations, we need to **map data dimensions to visual dimensions** in a principled way

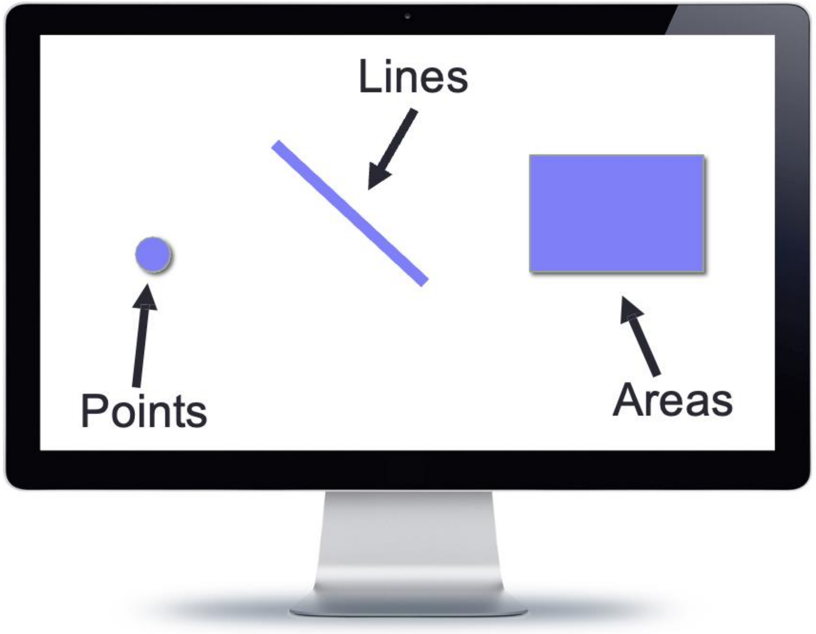
Data → Visuals

Data



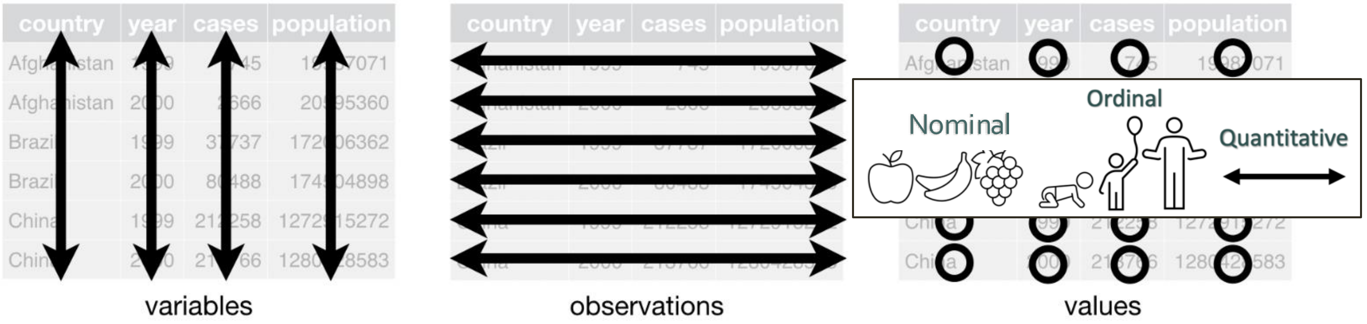
Visuals

- **Marks**
    - The “ink”



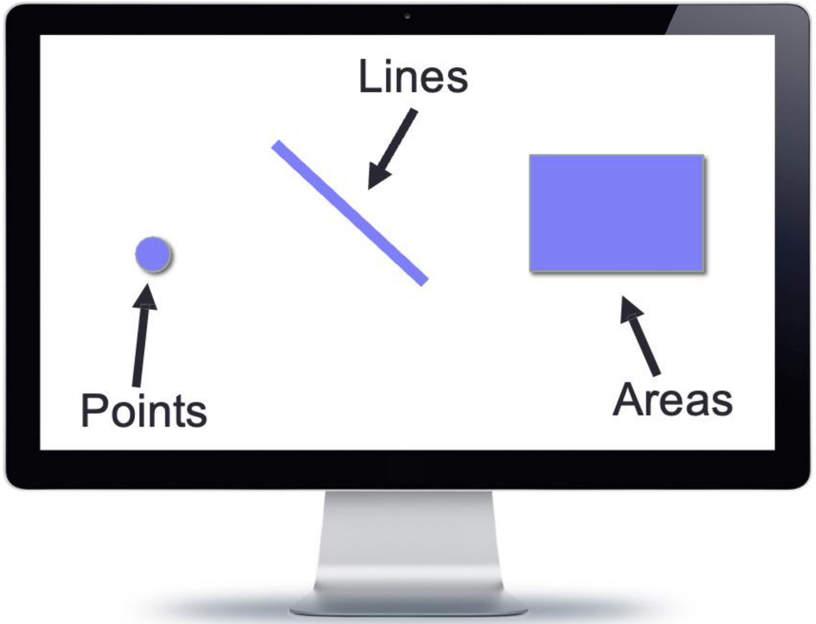
Data → Visuals

Data



Visuals

- **Marks**
  - The "ink"

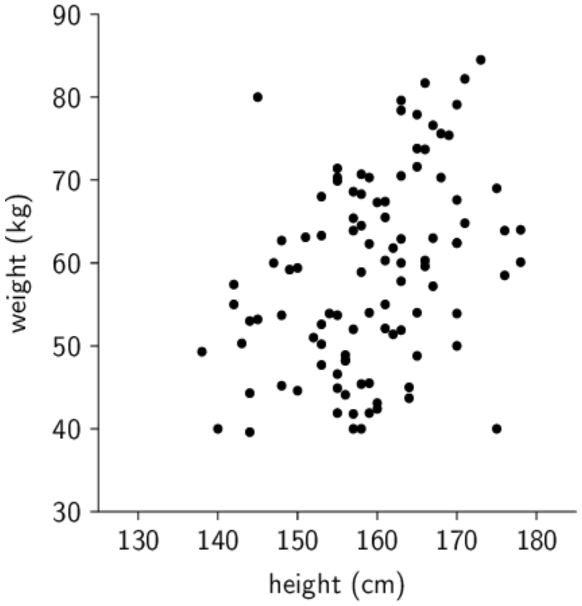


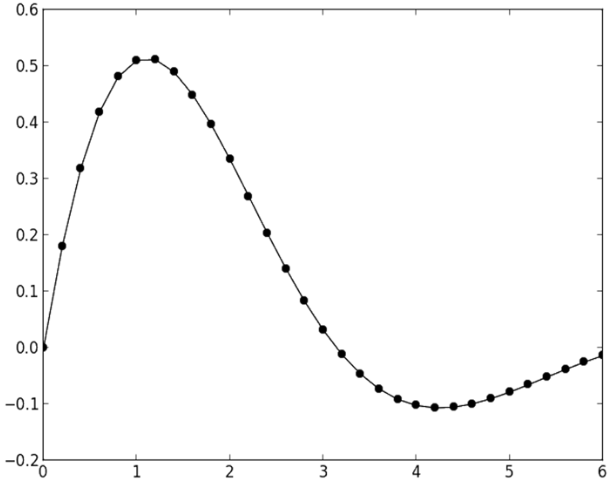
- **Channels or dimensions**
  - *How* the marks show up on the page

Visual Channels / Dimensions

Position

- Encode information using *where* mark is drawn
- Ex.





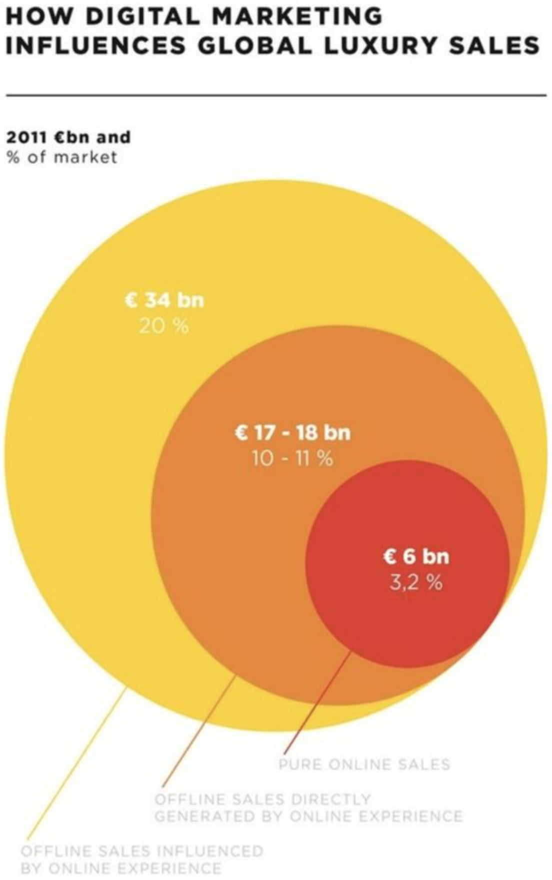
along a common scale

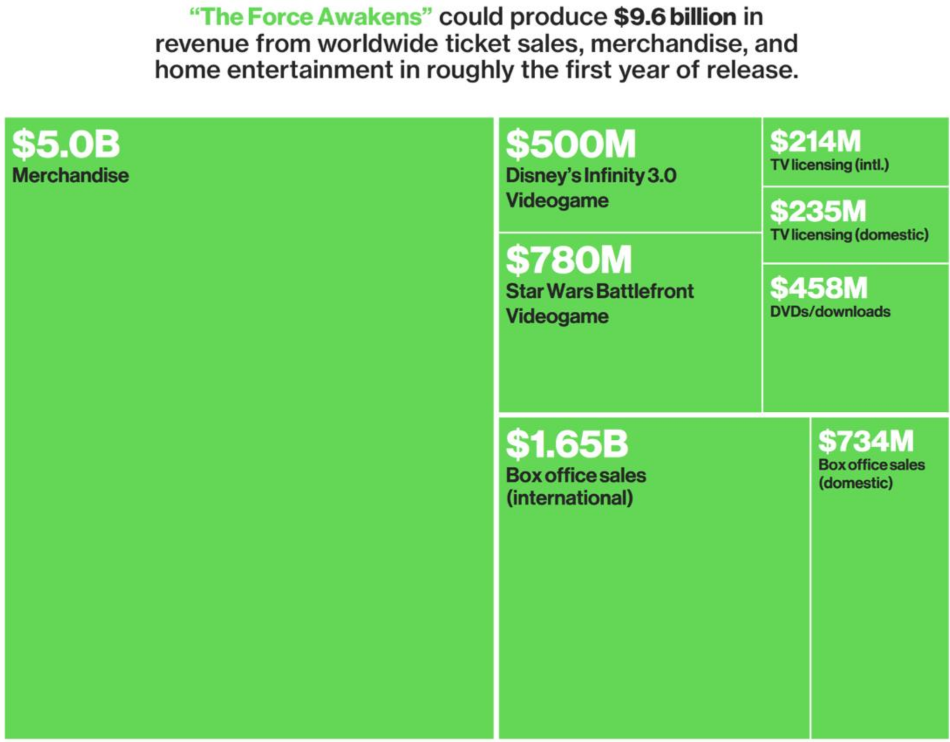
spatial region

Visual Channels / Dimensions

Size / Area

- Encode information using *how big* mark is drawn
- Ex.

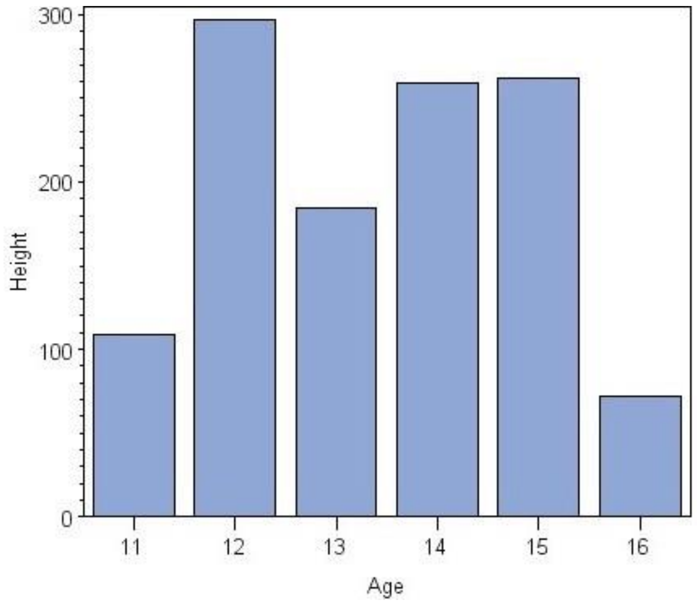


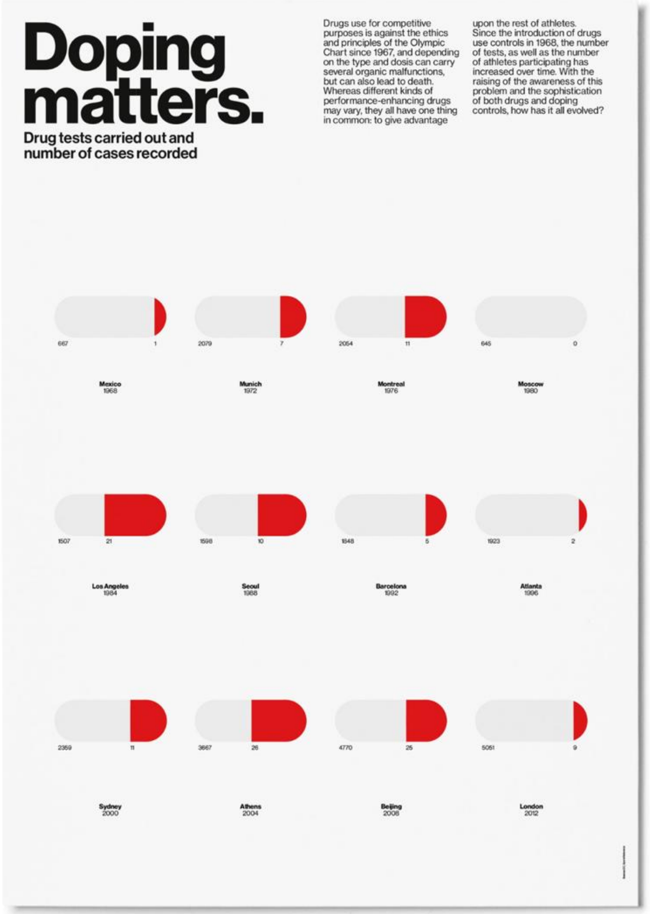


Visual Channels / Dimensions

Length

- Encode information using *how long* mark is drawn
- Ex.

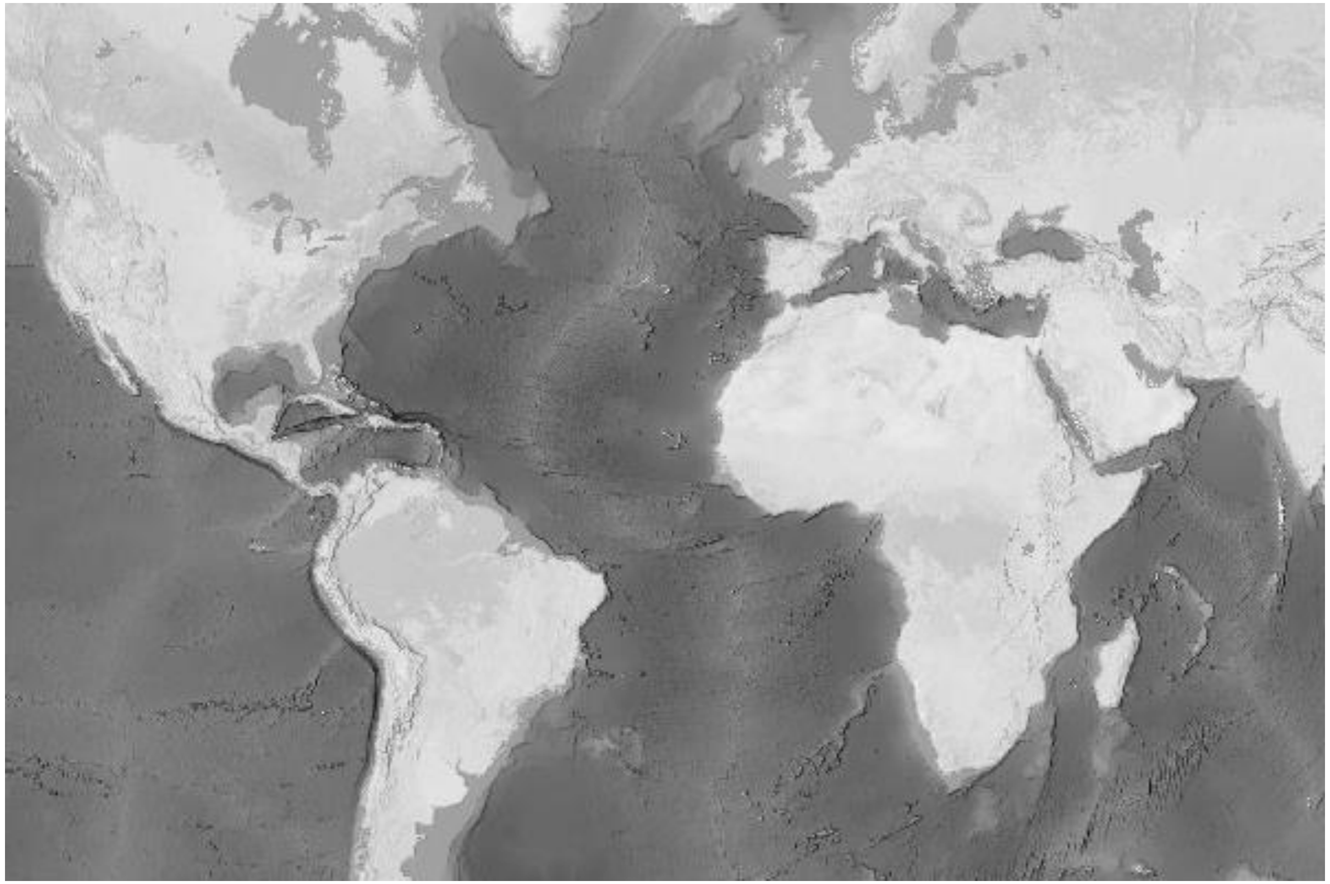




Visual Channels / Dimensions

Color: Luminance

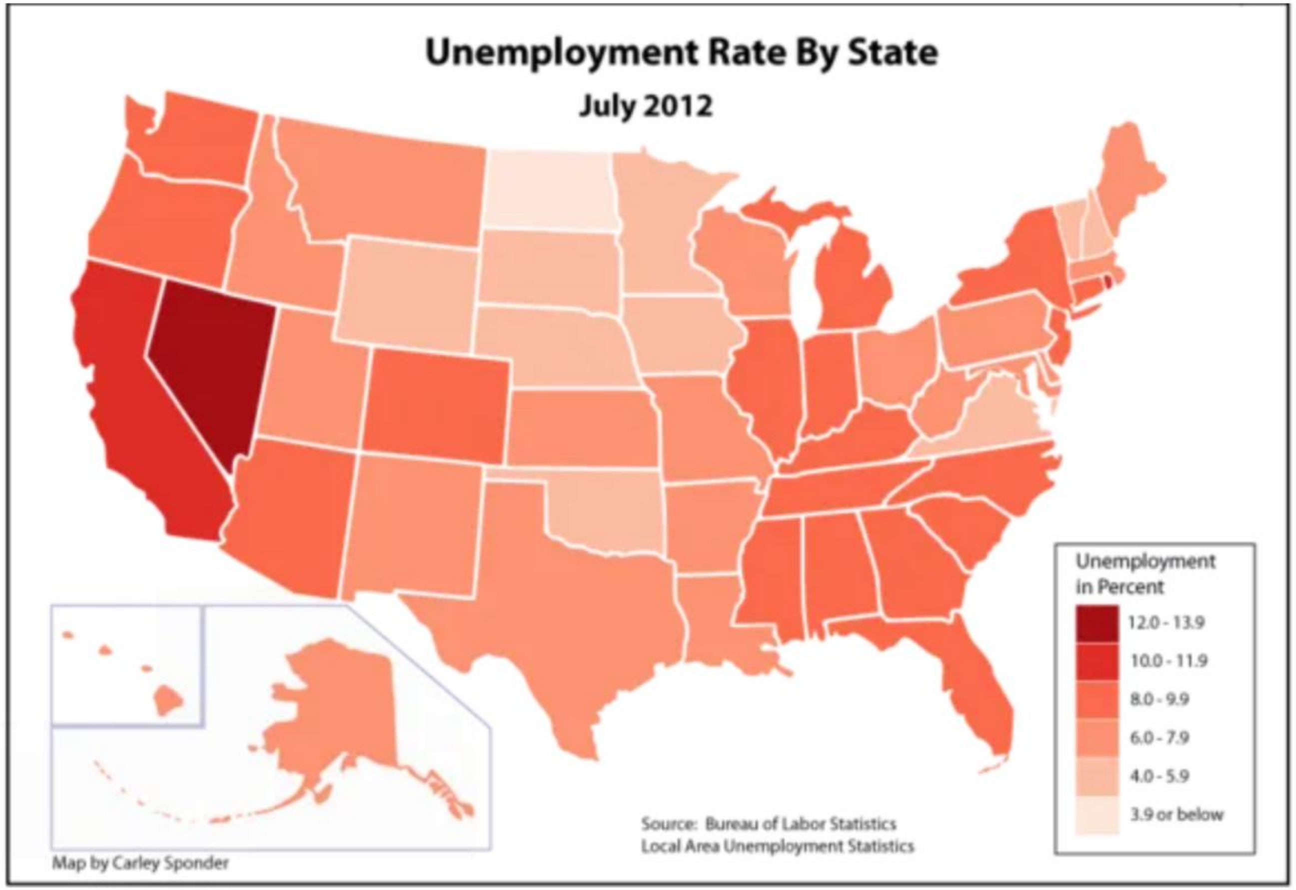
- Encode information using *how dark* mark is drawn
- Ex.



Visual Channels / Dimensions

Color: Saturation

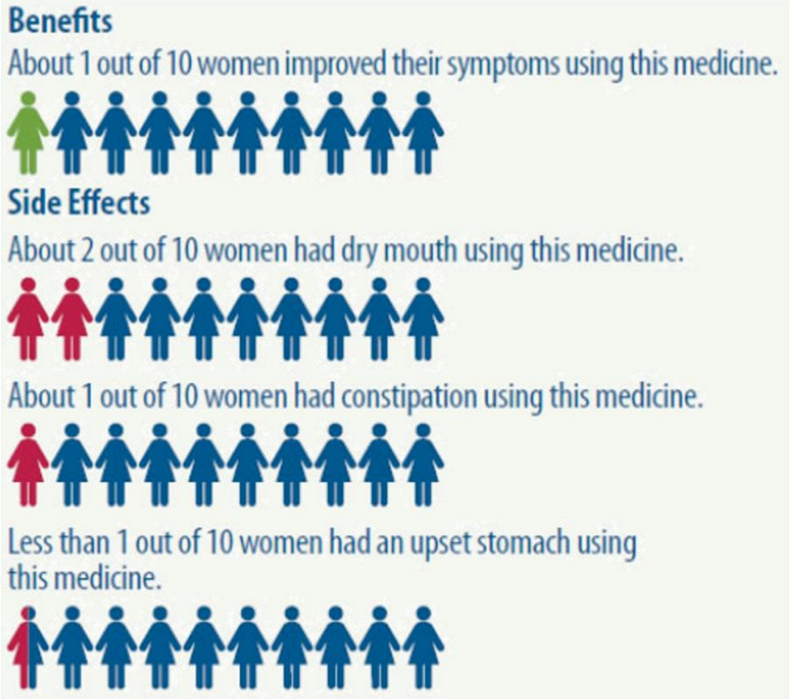
- Encode information using *how much color* mark has
- Ex.

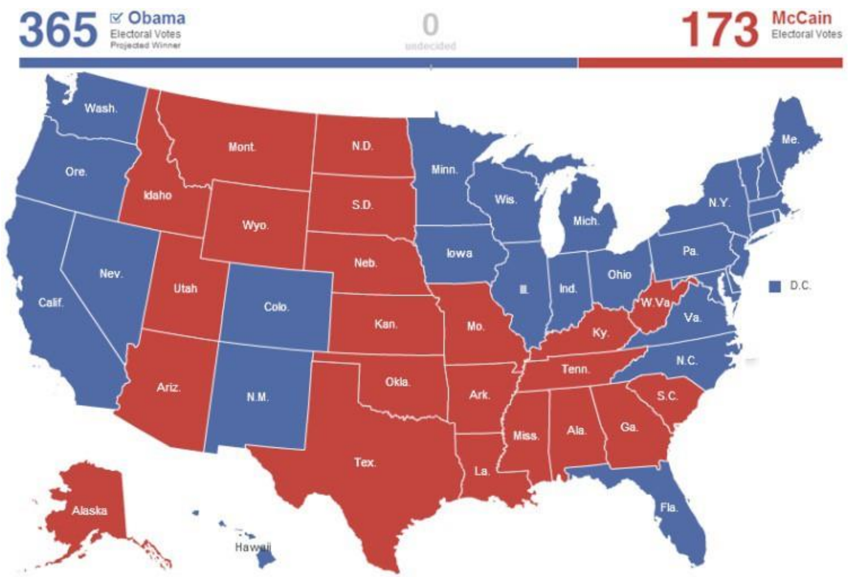


Visual Channels / Dimensions

Color: Hue

- Encode information using *hue* of mark
- Ex.

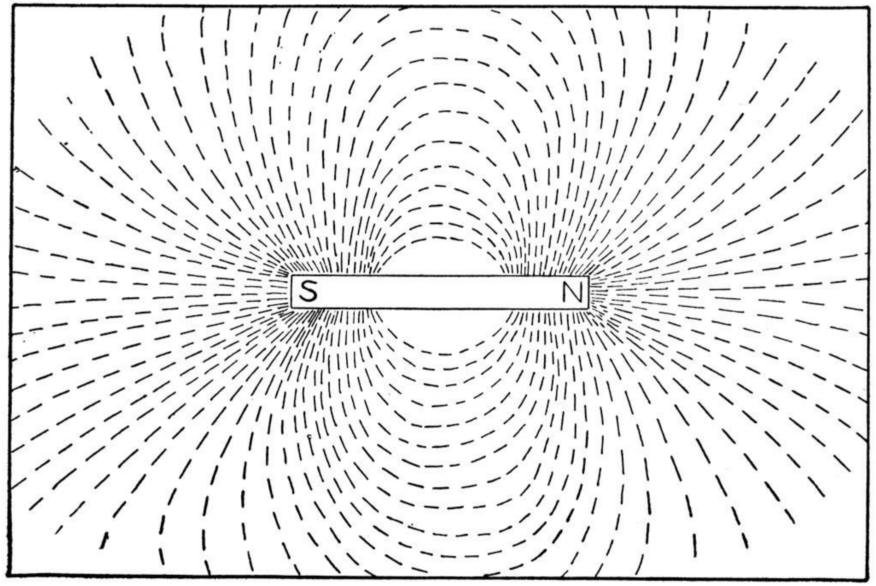




Visual Channels / Dimensions

Orientation / Tilt / Angle

- Encode information using how mark is *rotated*
- Ex.



Visual Channels / Dimensions

Shape

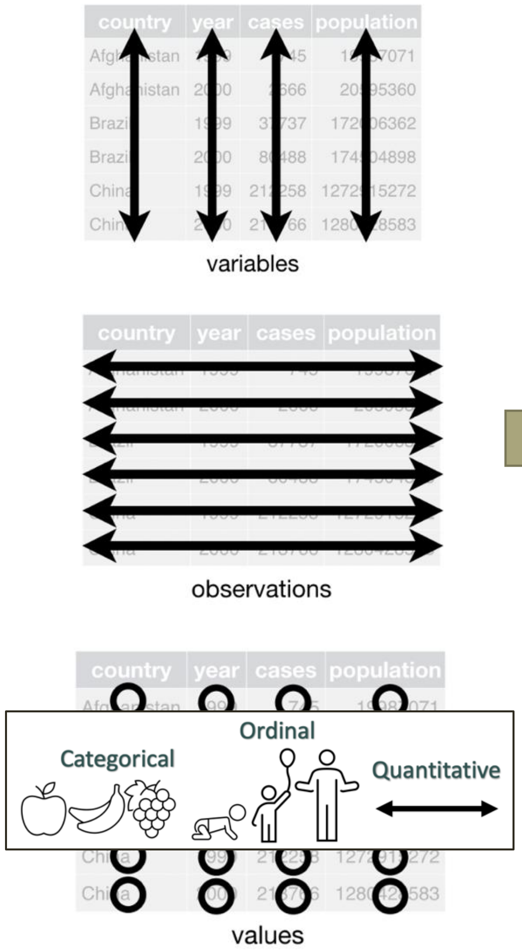
- Encode information using how mark is *shaped*
- Ex.

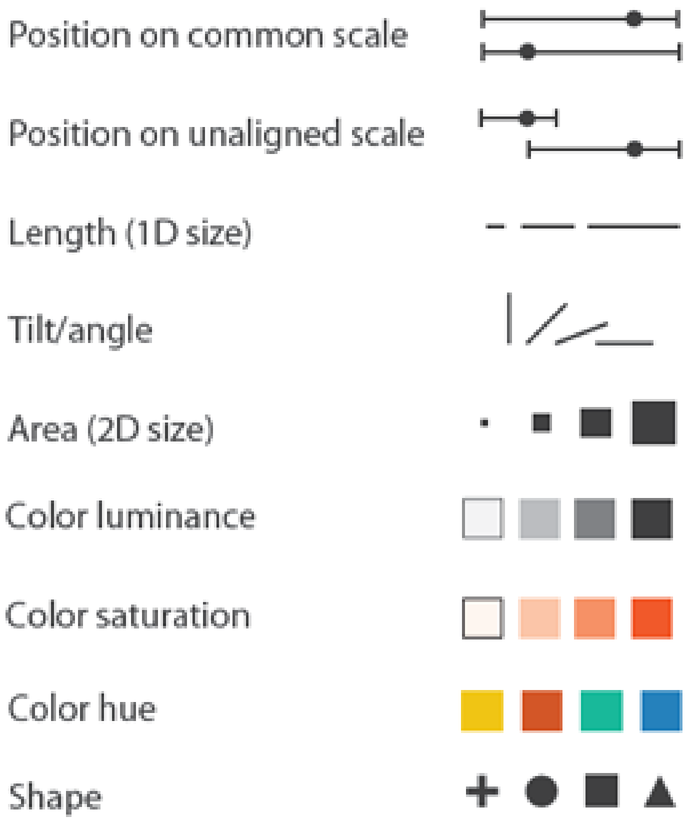
Data → Visuals

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

Data → Visuals

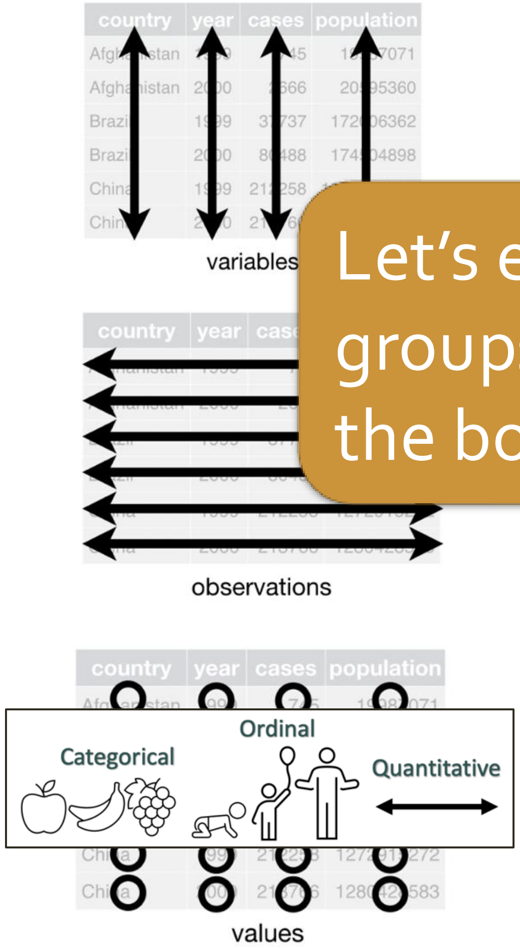
- Remember… **Big idea behind visualization**
  - Map data dimensions to visual dimensions in a principled way



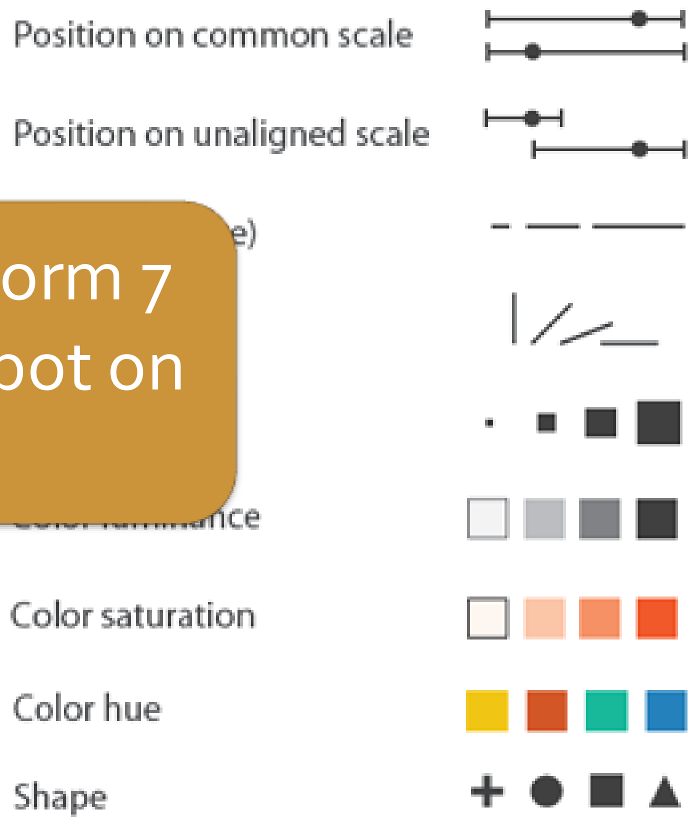


Data → Visuals

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



Let's experiment. Form 7 groups and find a spot on the board.



Work with your group to represent each observation in this dataset as a point (mark) styled only using the visual channel you were assigned.

You goal is for other groups to easily infer which point represents which observation.

Data → Visuals

Name
blueberry
asparagus
pumpkin
pea
watermelon
long bean

Work with your group to represent each observation in this dataset as a point (mark) styled using the visual channel you were assigned.

You goal is for other groups to easily infer which point represents which observation.

Data → Visuals

Peak
July
April
October
June
August
September

Work with your group to represent each observation in this dataset as a point (mark) styled using the visual channel you were assigned.

You goal is for other groups to easily infer which point represents which observation.

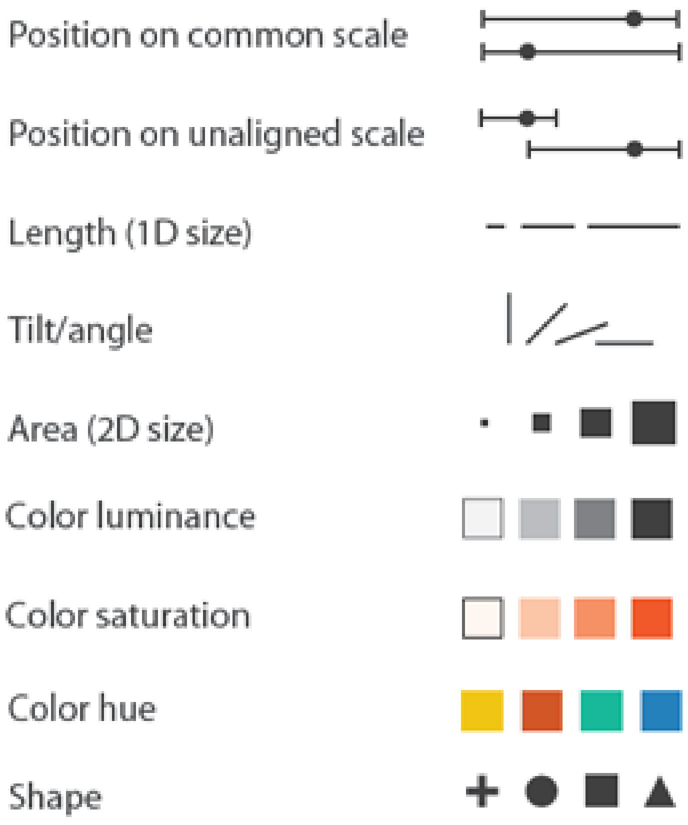
Data → Visuals

Price (per lb)
6.45
4.99
0.24
7.00
3.65
3.64

Data → Visuals

What type of variable is Name?

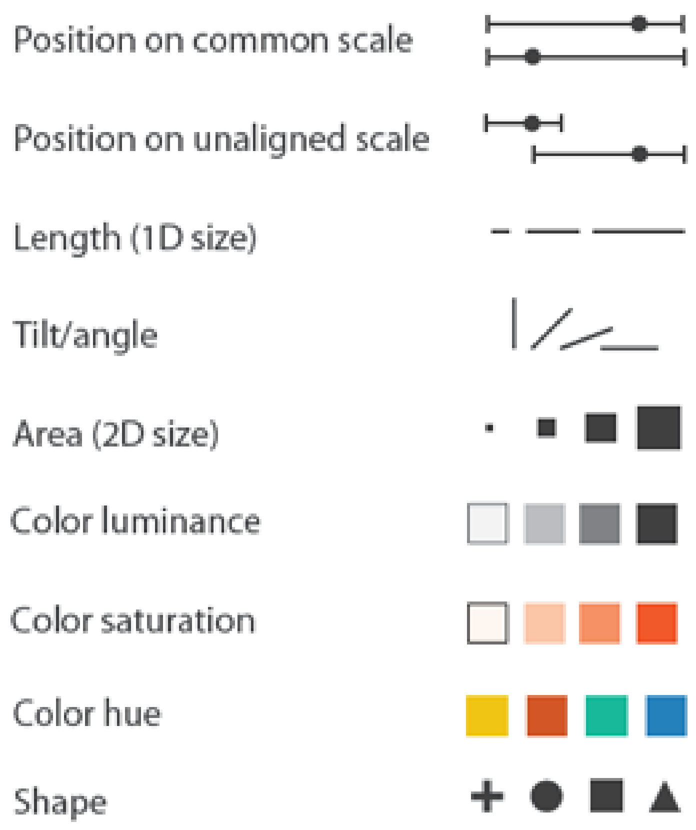
Name
blueberry
asparagus
pumpkin
pea
watermelon
long bean



Data → Visuals

What type of variable is Peak?

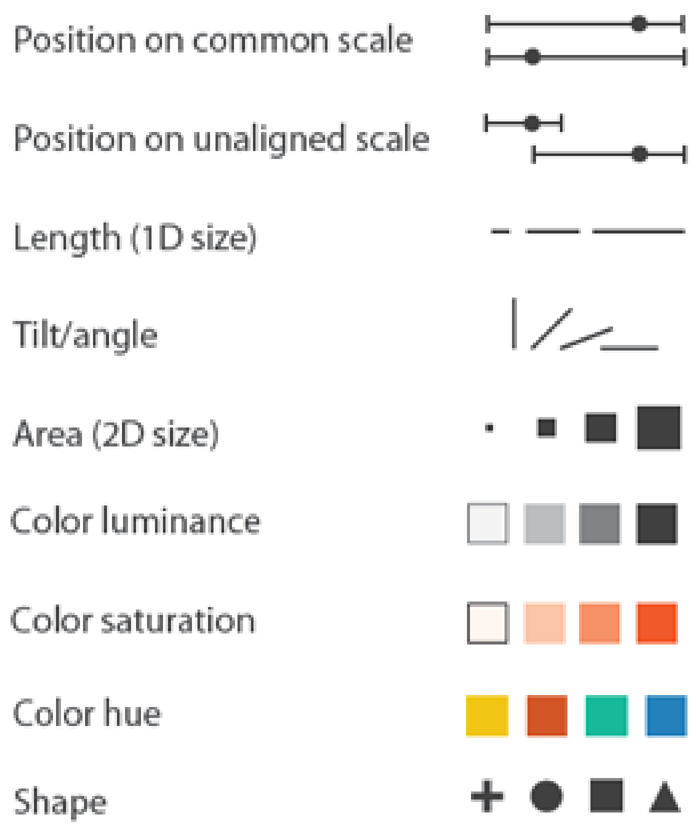
Peak
July
April
October
June
August
September



Data → Visuals

What type of variable is Price?

Price (per lb)
6.45
4.99
0.24
7.00
3.65
3.64



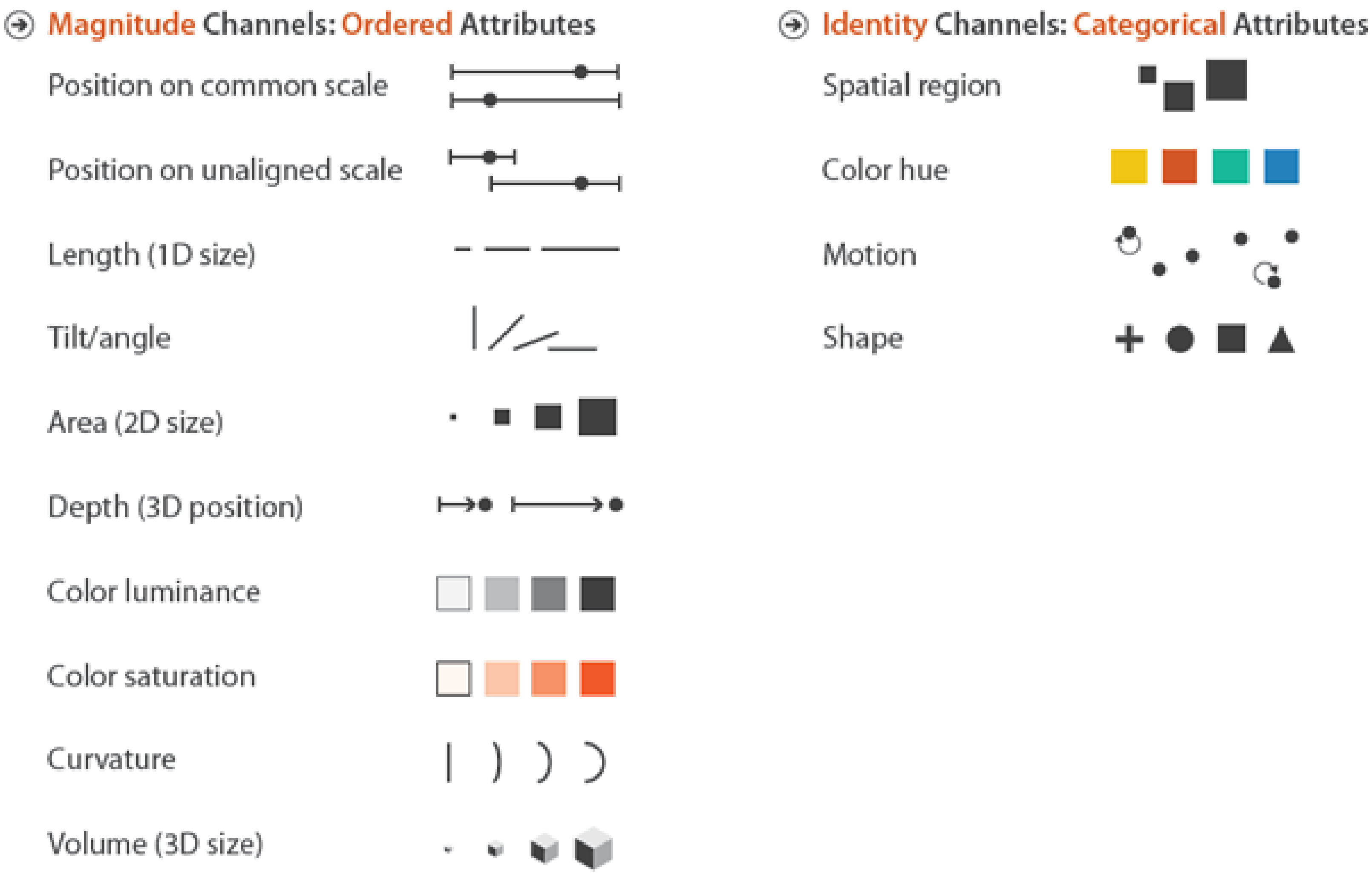
Data → Visuals

- Remember… **Big idea behind visualization**
  - Map data dimensions to visual dimensions in a principled way
  - Insight 1: Not all visual dimensions can represent all data types

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

Data → Visuals

- Remember… **Big idea behind visualization**
  - Map data dimensions to visual dimensions in a principled way
  - Insight 1: Not all visual dimensions can represent all data types



Try it out!

- Work with 2 other people. Be prepared to share your work with the class.
- Find a data visualization you think is interesting
  - Some ideas for where to look: New sites, government sites, Tableau Viz Gallery, massvis.mit.edu
- Identify the following:
  - What is the data that's being visualized?
  - Is the data source included?
  - What marks are used?
  - What is the mapping between data variables and visual variables?