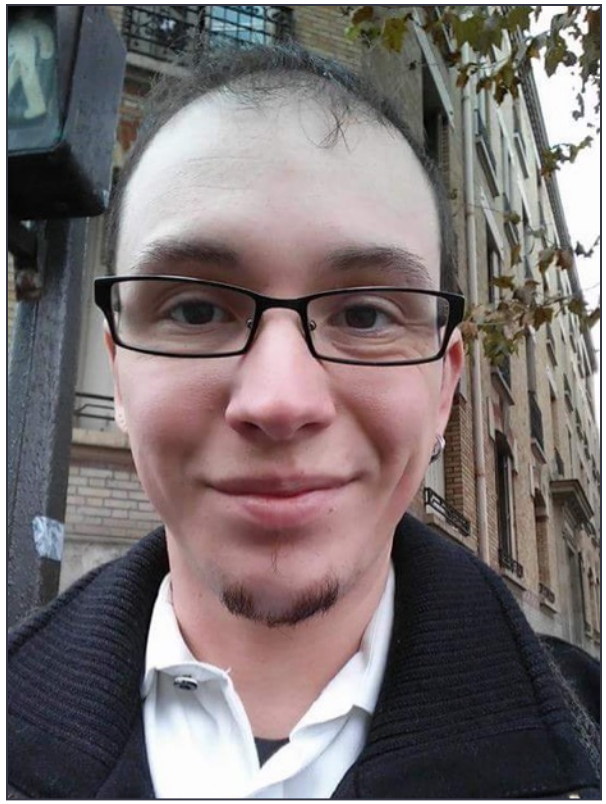
INTRODUCTION TO

DATA VISUALIZATION

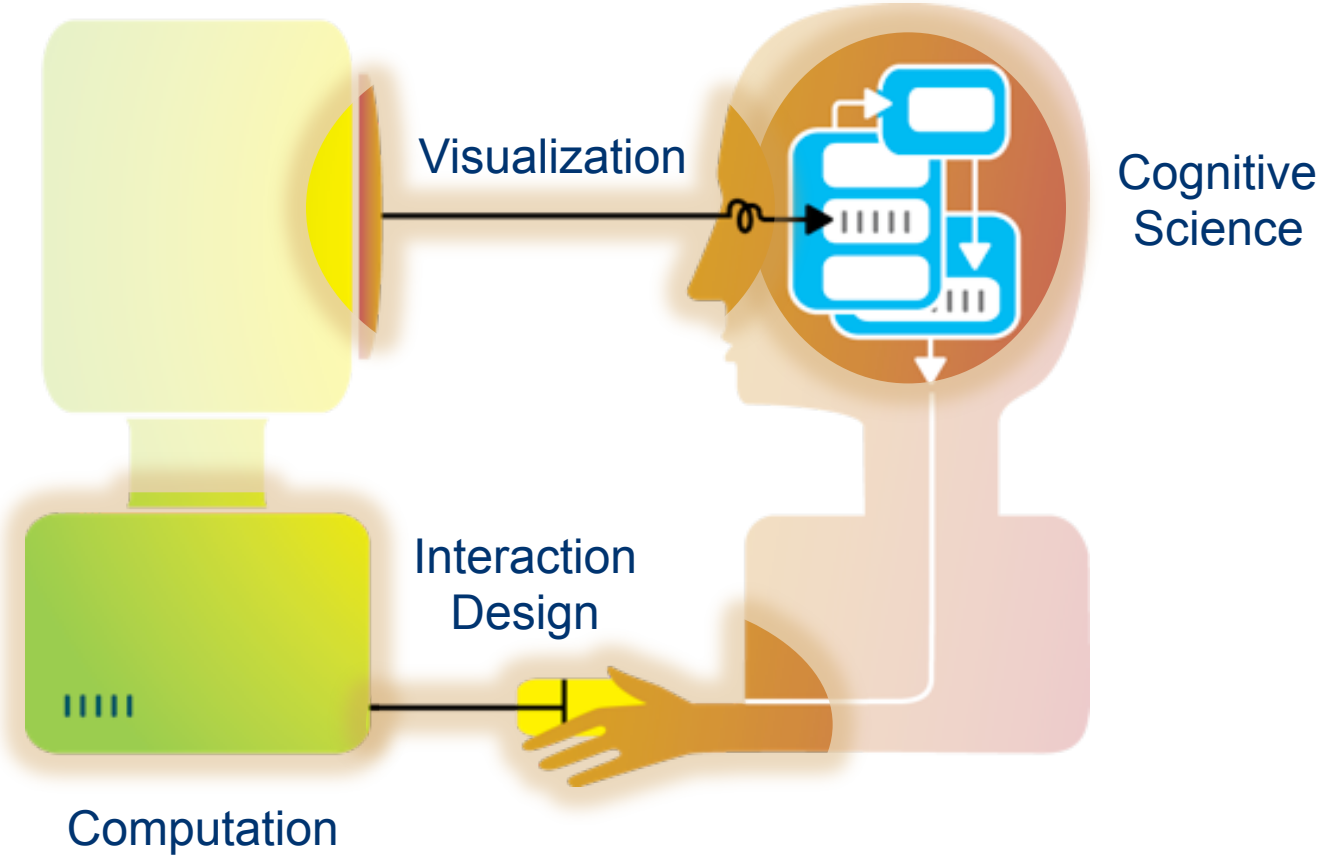
R. Jordan Crouser

Introductions & Background

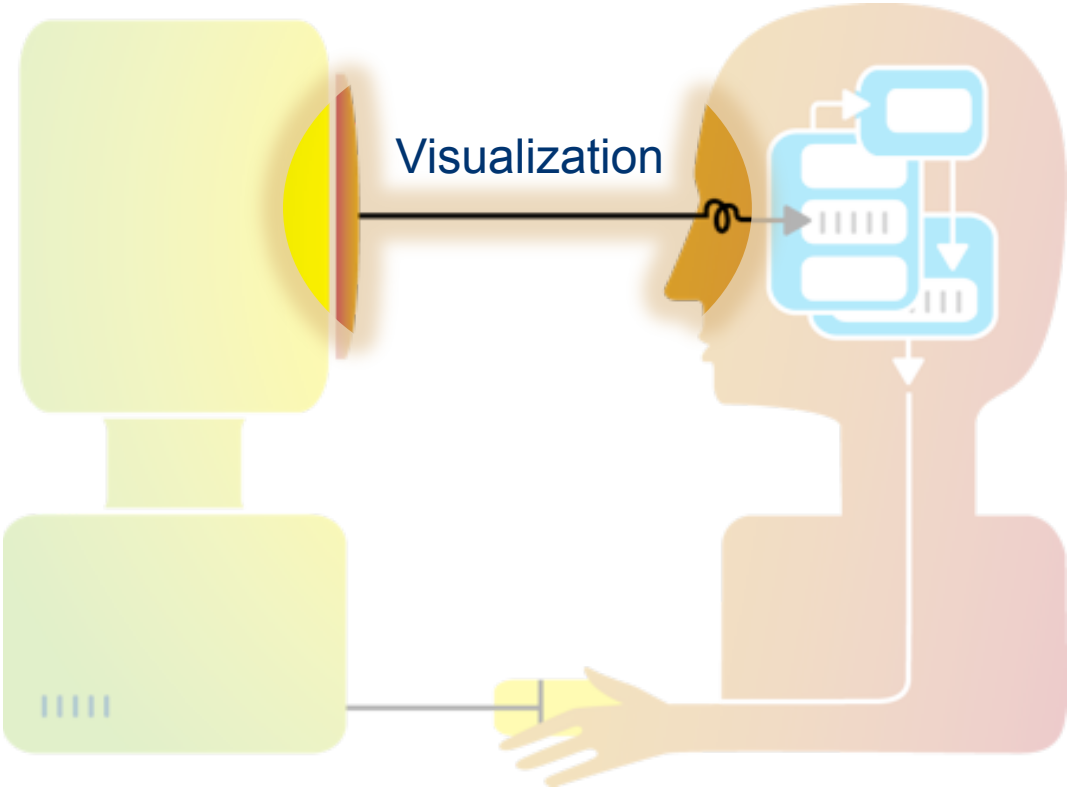


Jordan
(data scientist)

My research (generally)



About this lecture / demo



Tons of useful tools…

Crossfilter

Fast Multidimensional Filtering for Coordinated Views

tableau

visual.ly

Data-Driven Documents

ggplot2

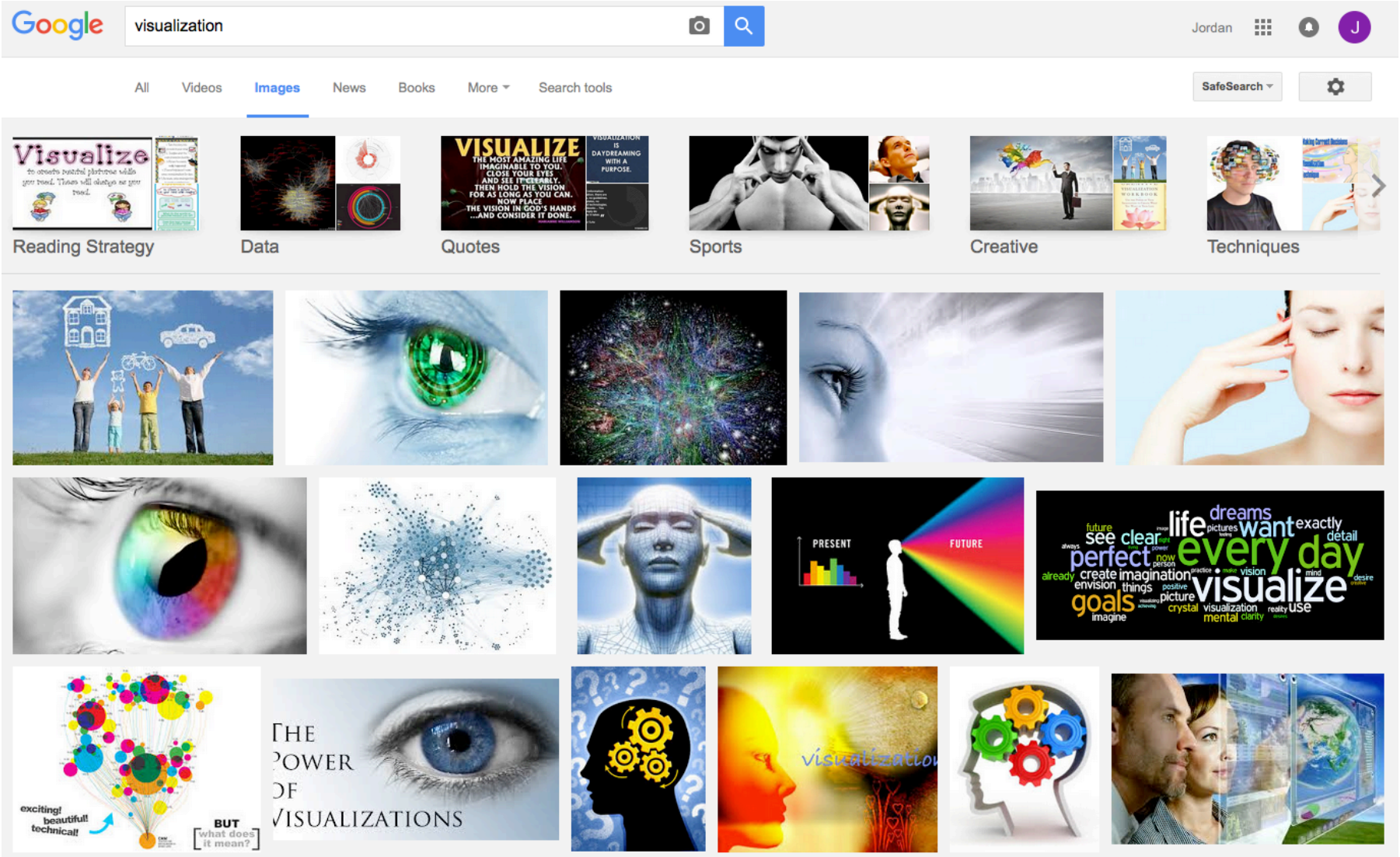
node js

Gephi

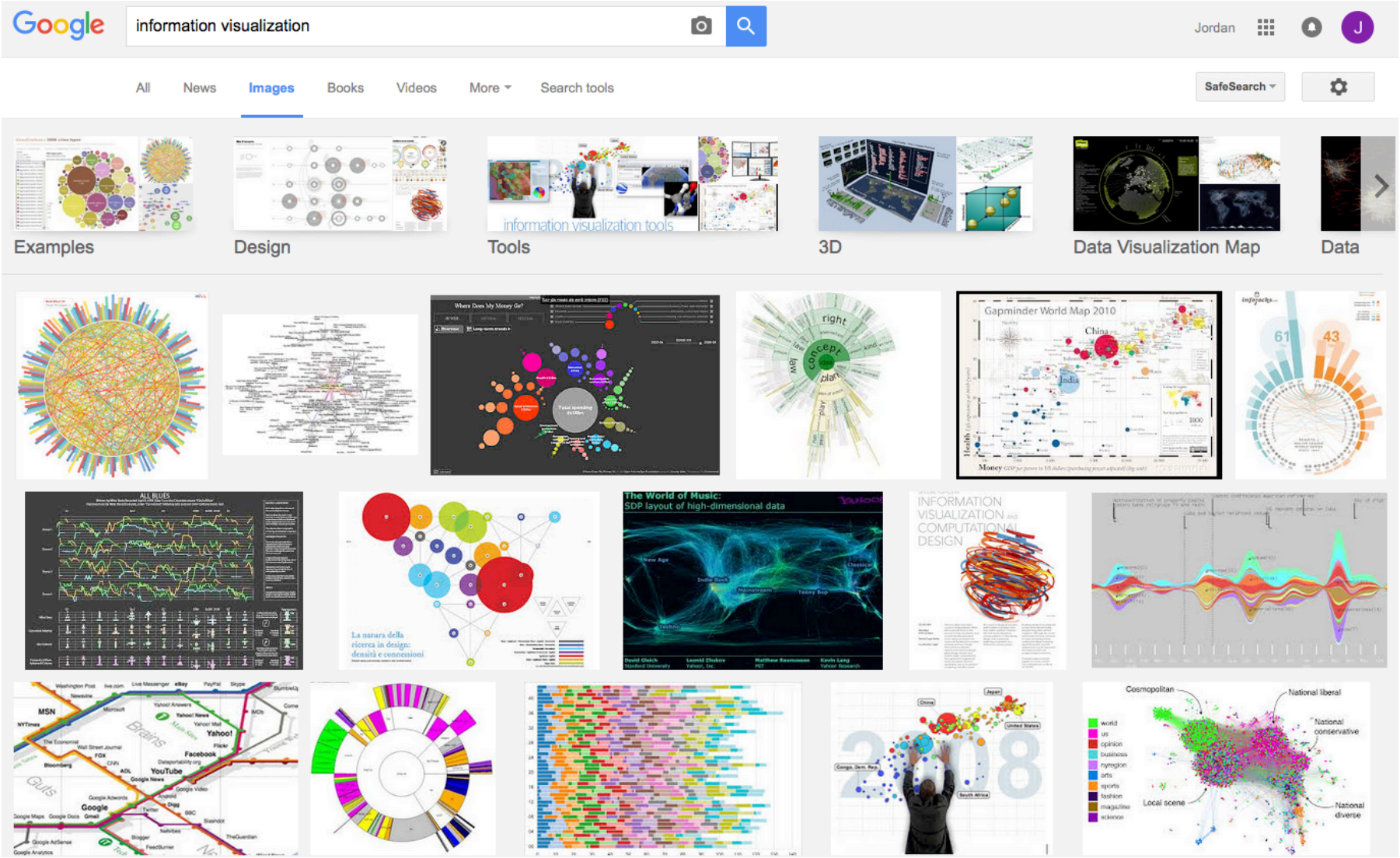
Leaflet

plotly

Part 1: what is **visualization**?



Part 1: what is **visualization**??



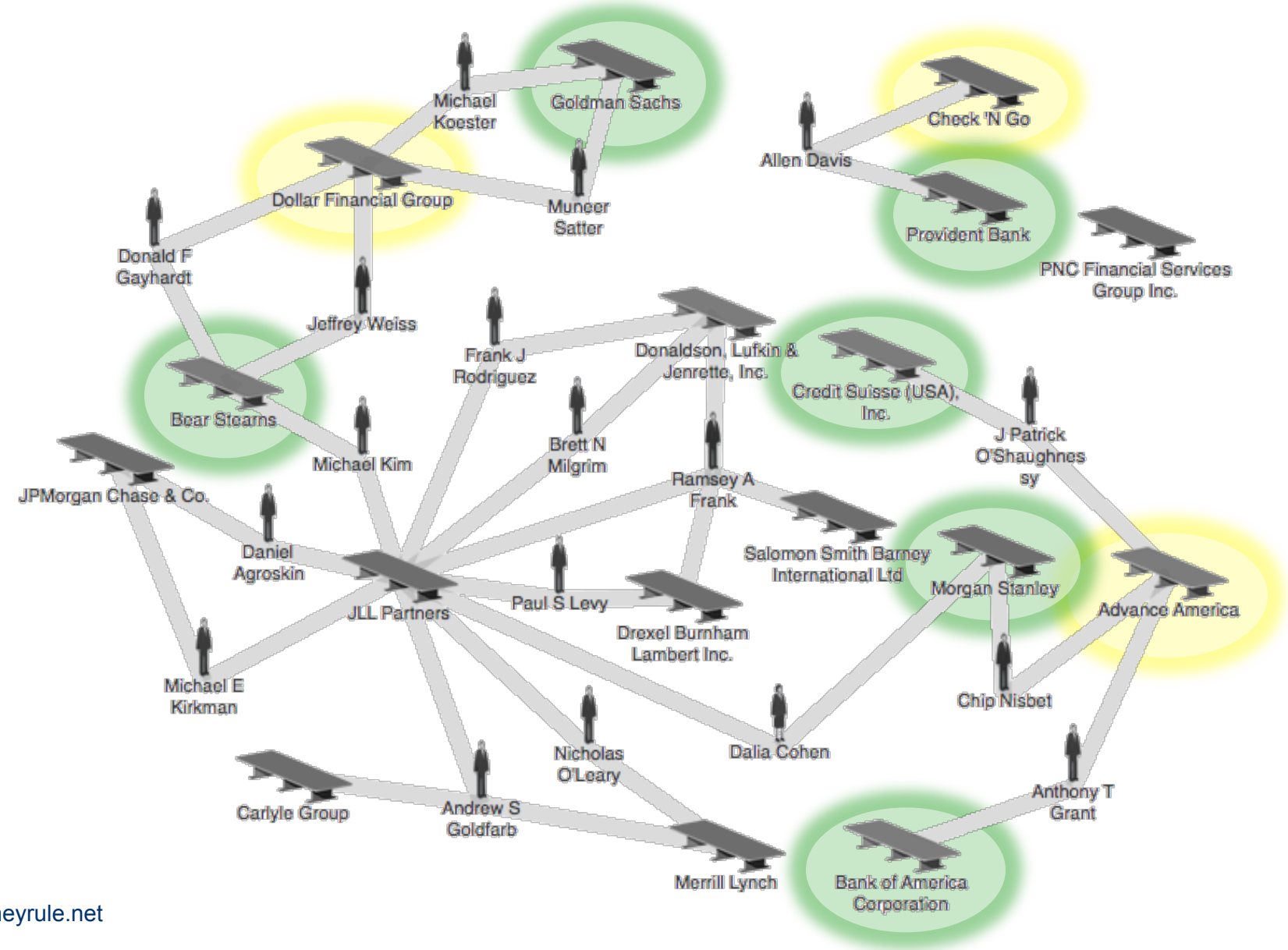
Perhaps a more helpful question

What are some ways
a “visualization” can be **useful**?

Does it help you spot trends?

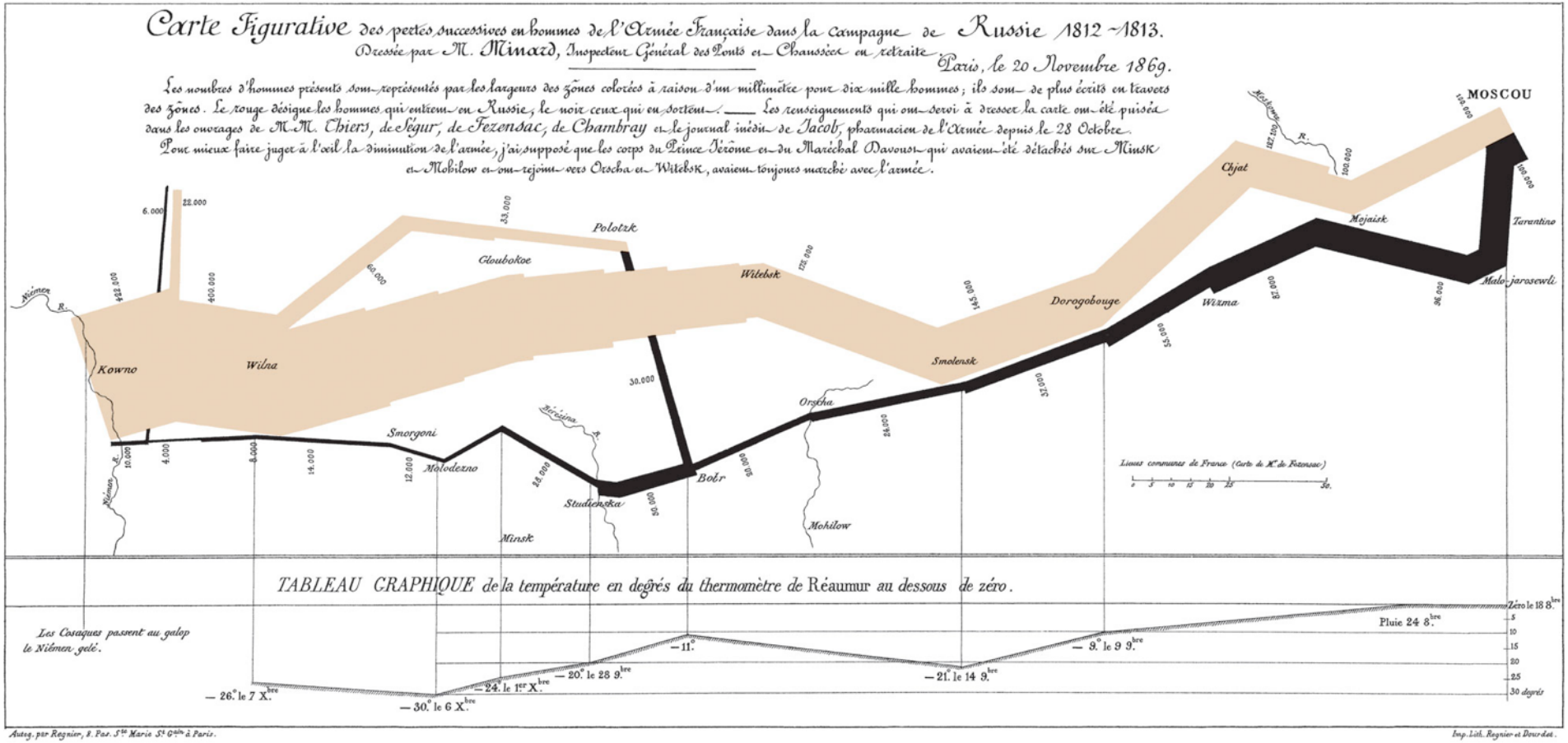
More info here: http://en.wikipedia.org/wiki/1854_Broad_Street_cholera_outbreak

Does it help you explore?



Credit: theyrule.net

Does it tell a story?

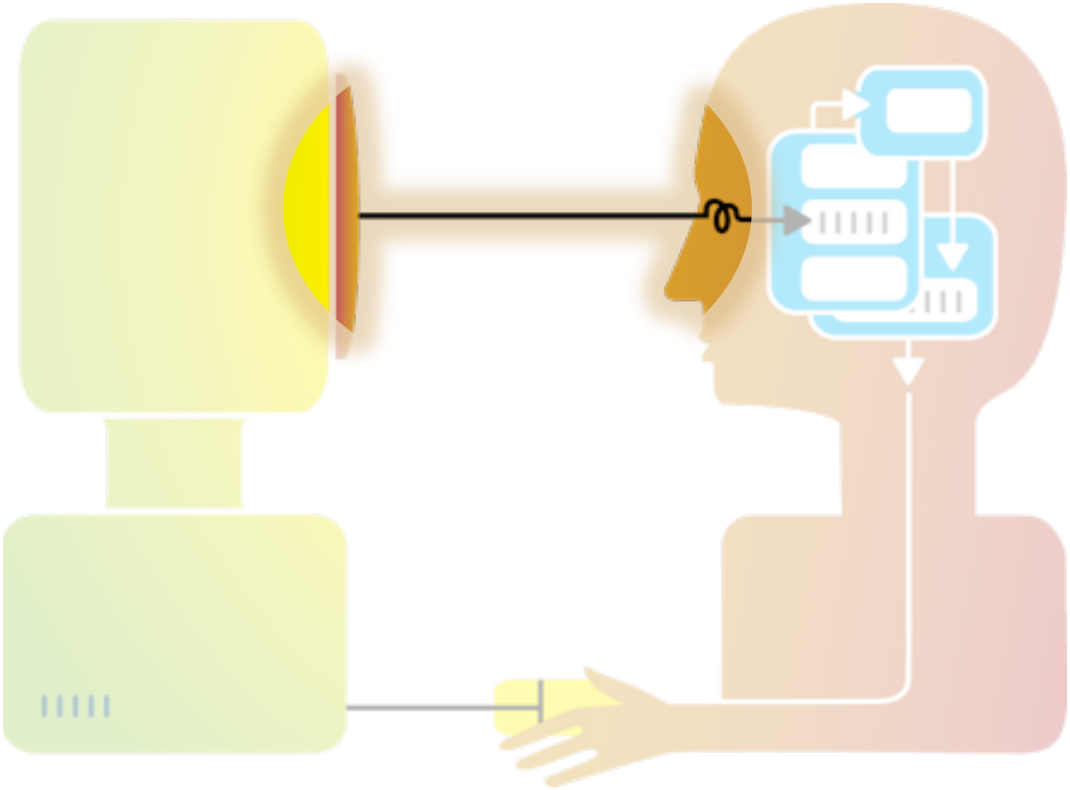


"Figurative Map of the successive losses in men of the French Army in the Russian campaign 1812-1813."
Drawn by Mr. Charles Minard, Inspector General of Bridges and Roads in retirement. Paris, 20 November 1869.

Visualization (def.)

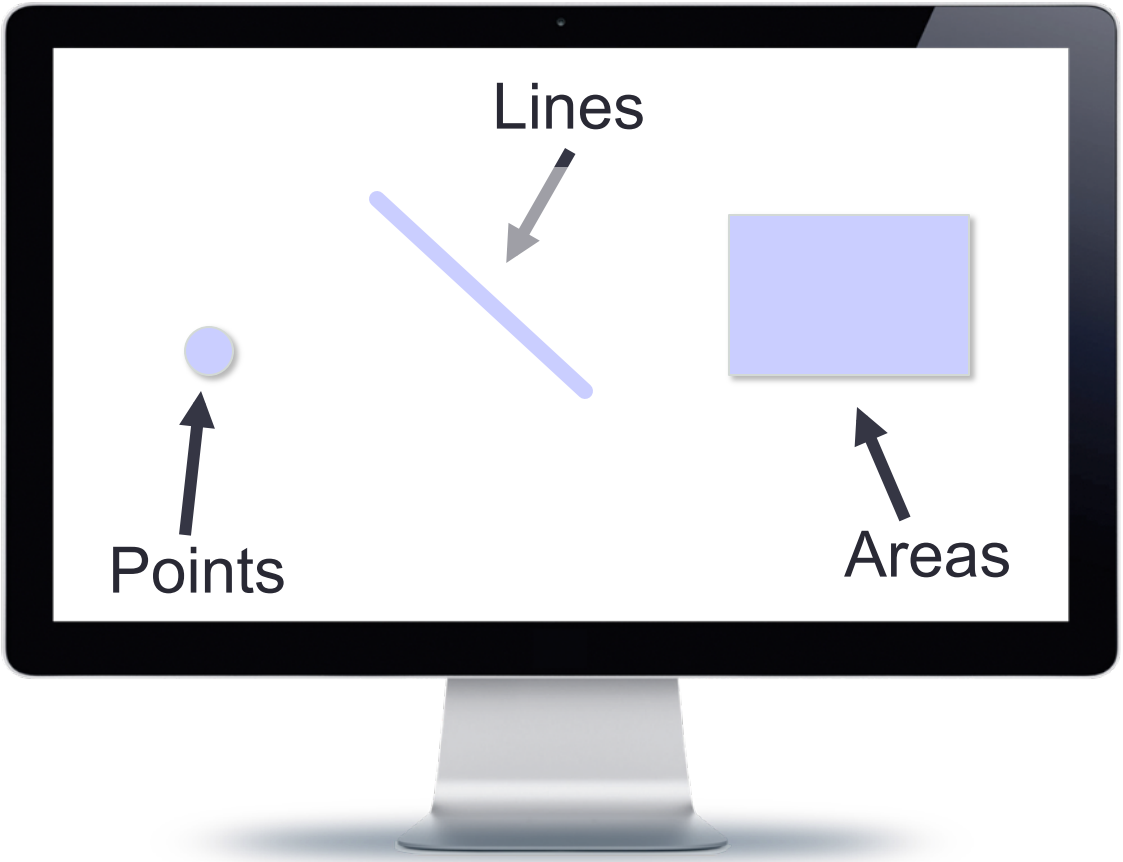
Visual **representations** of data that reinforce human **cognition**

Part 2: how do we **build** them?



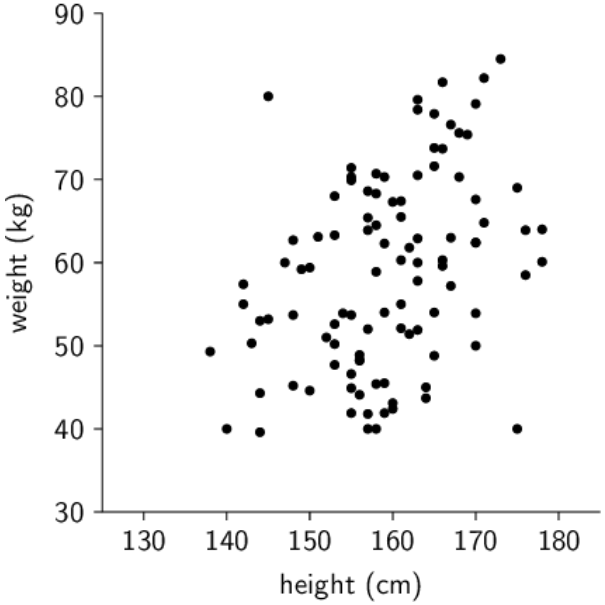
Graphical primitives

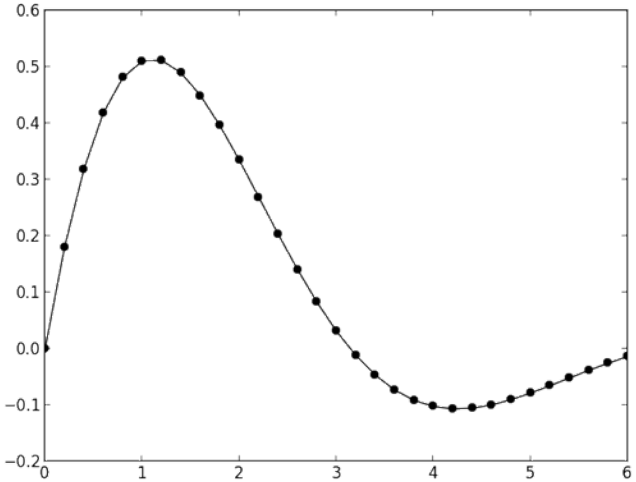
The images we draw are composed of marks: like ink



Visual dimension: position

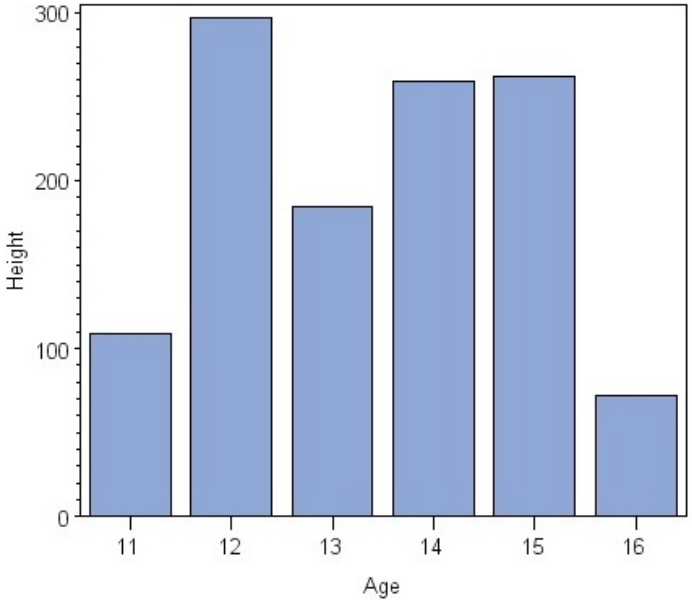
- Encode information using **where** the mark is drawn
- Some examples:

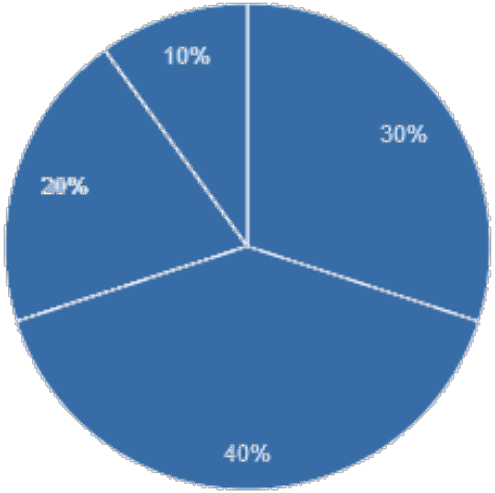




Visual dimension: size

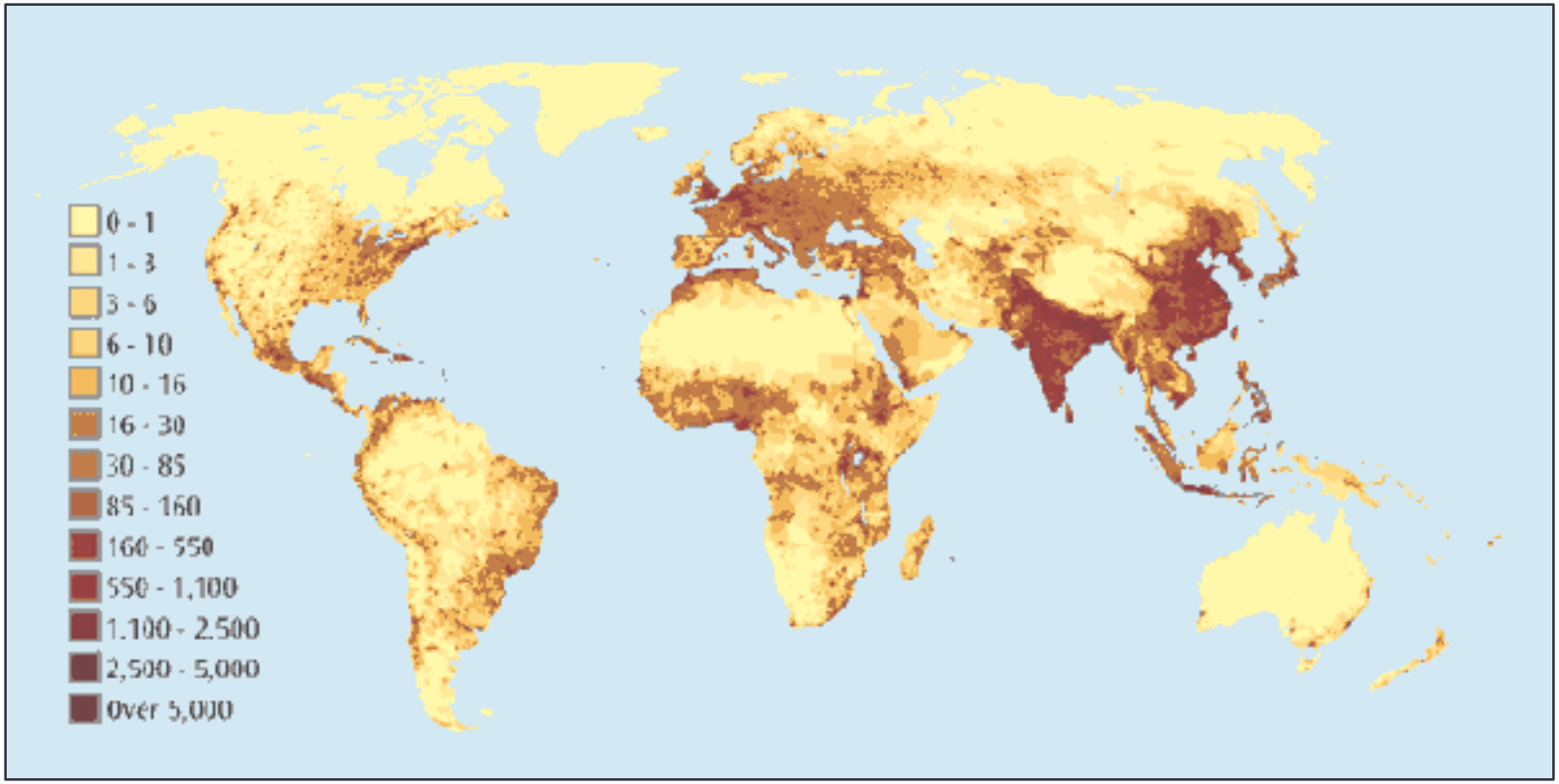
- Encode information using **how big** the mark is drawn
- Examples:





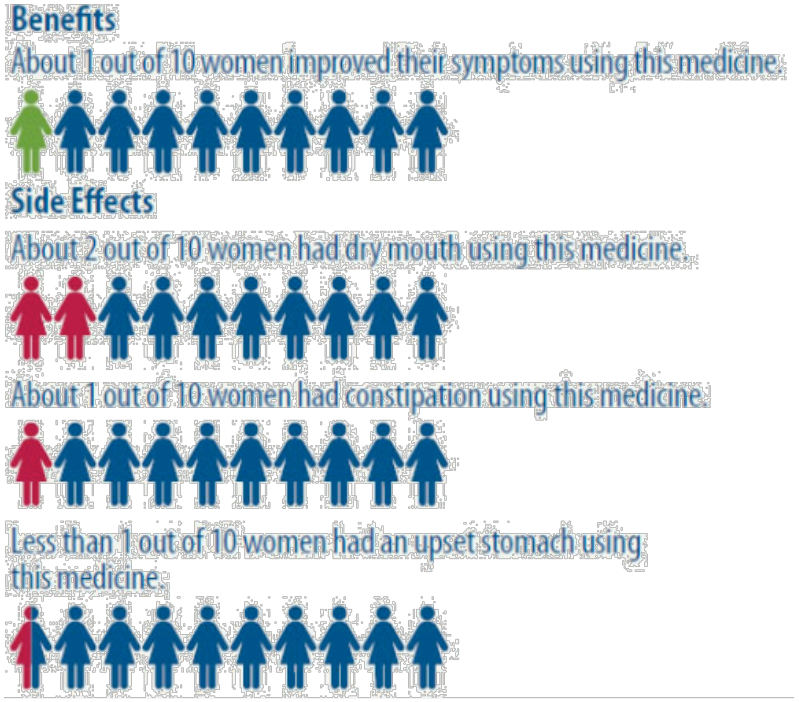
Visual dimension: value

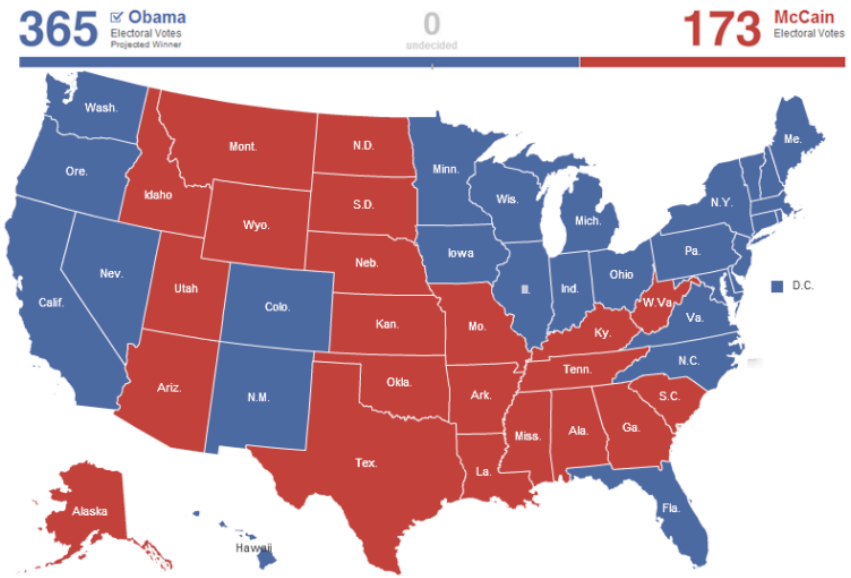
- Encode information using **how dark** the mark is drawn
- Example:



Visual dimension: color

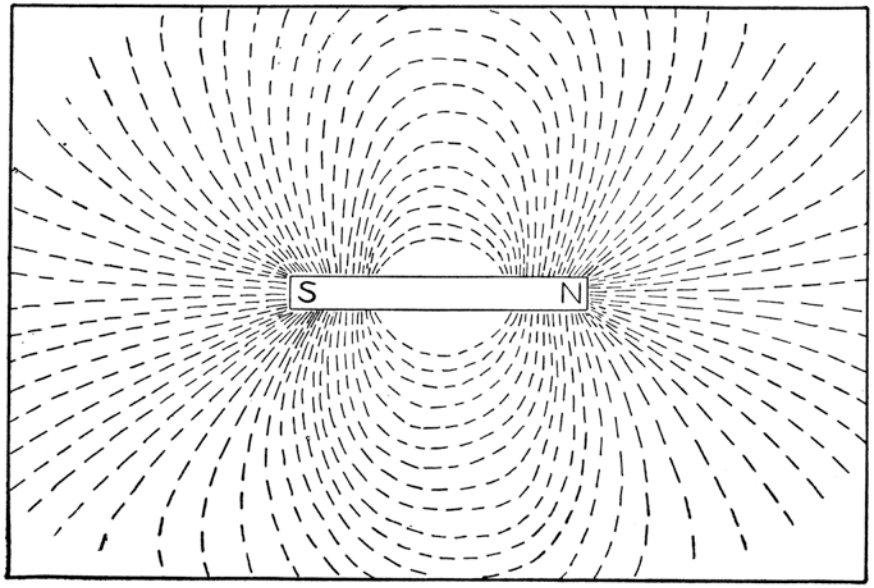
- Encode information using the **hue** of the mark
- Examples:





Visual dimension: orientation

- Encode information using how the mark is **rotated**
- Examples:

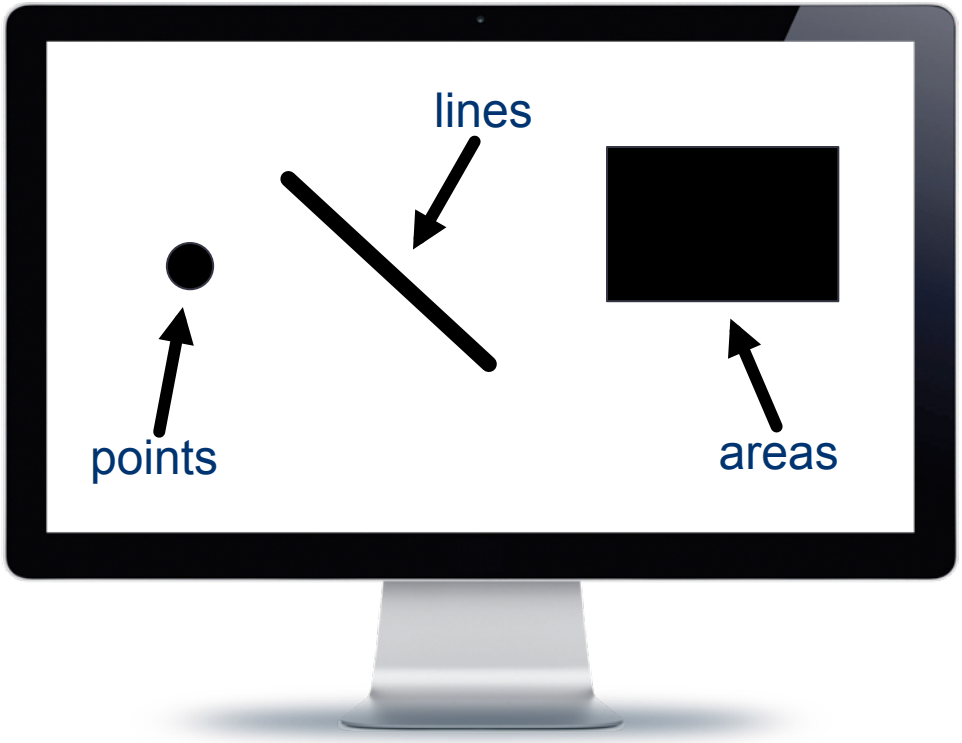


Visual dimension: shape

- Encode information using how the mark is **shaped**
- Examples:

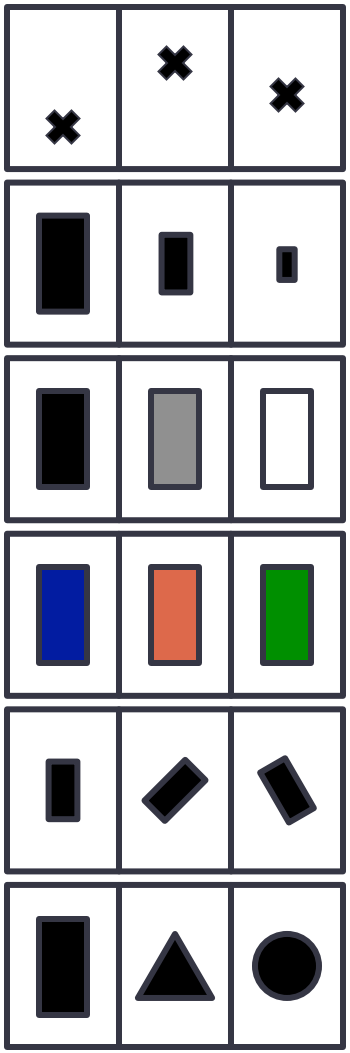


Quick recap: what we have to work with



graphical primitives

dimensions



Part 3: data visualization in python

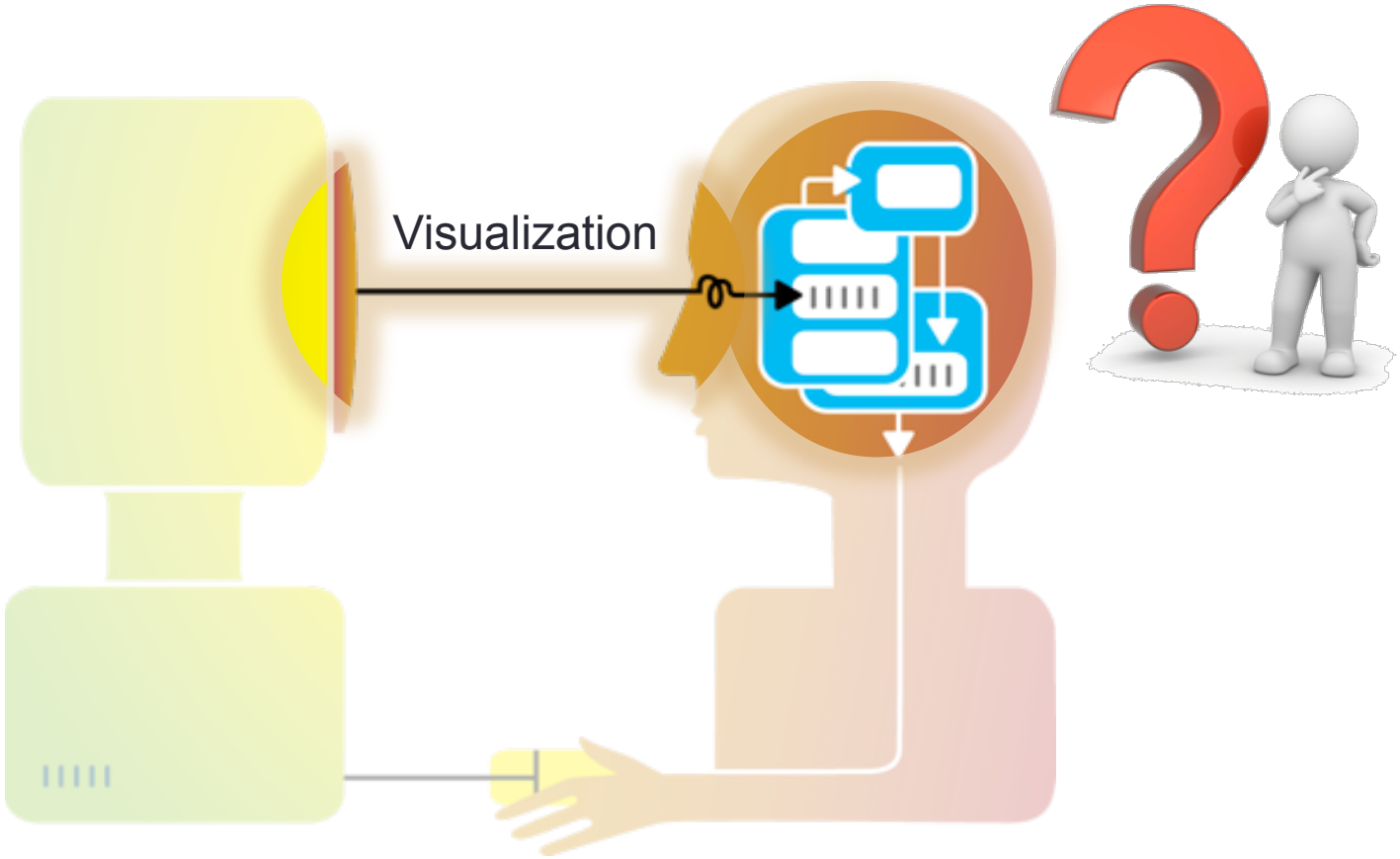
plotly

Part 3: data visualization in python

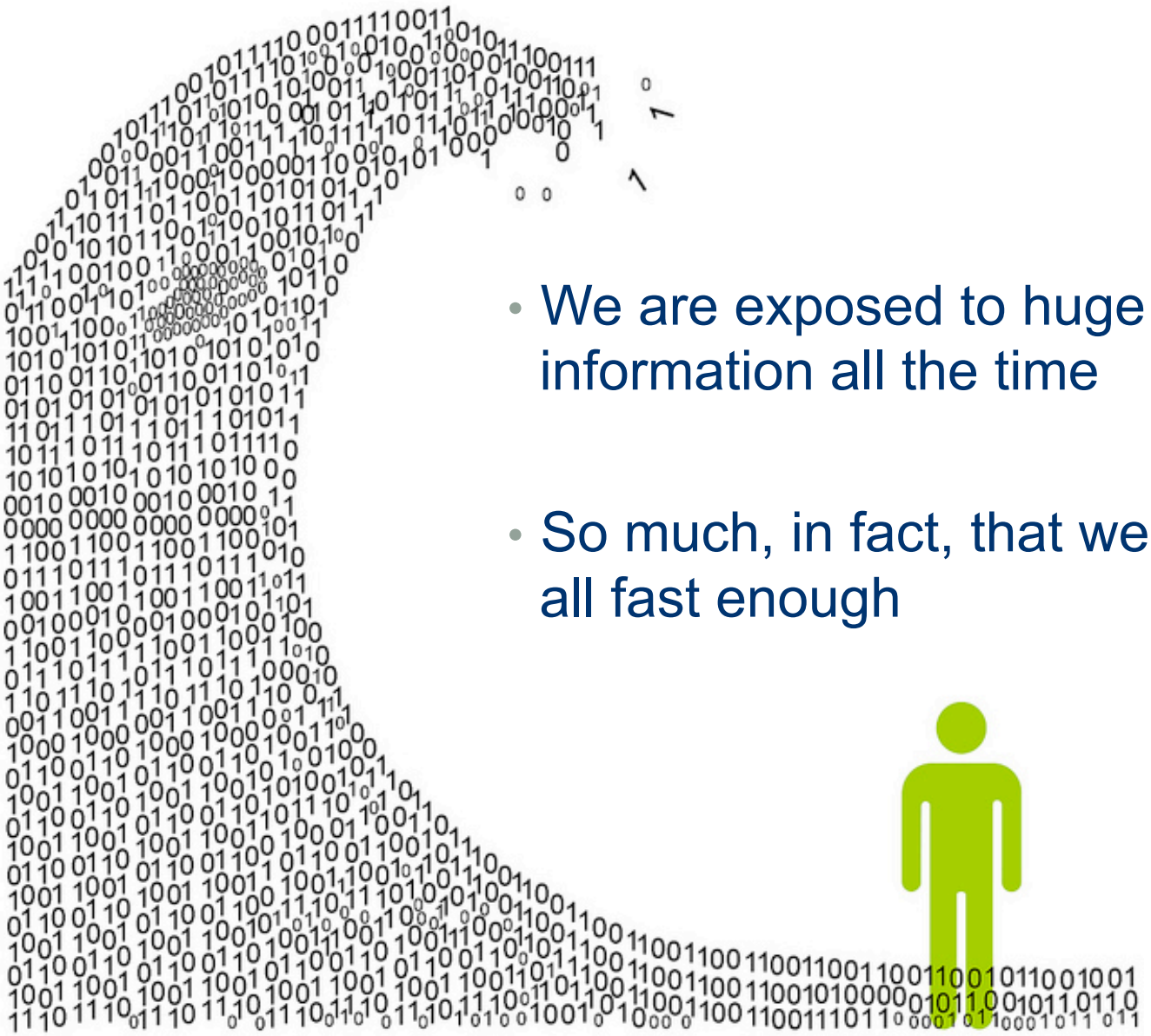
Bottom line: it's just like using any other functions!

Part 3: why **visualization** works (or not)

Visualization helps shape *mental models*



Information overload



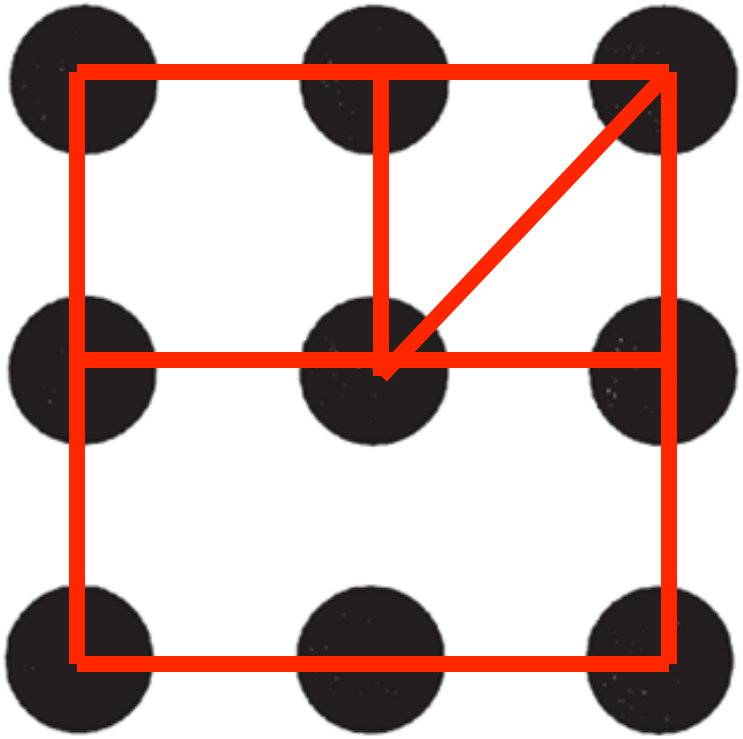
- We are exposed to huge amounts of information all the time
- So much, in fact, that we can’t process it all fast enough

Mental models

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

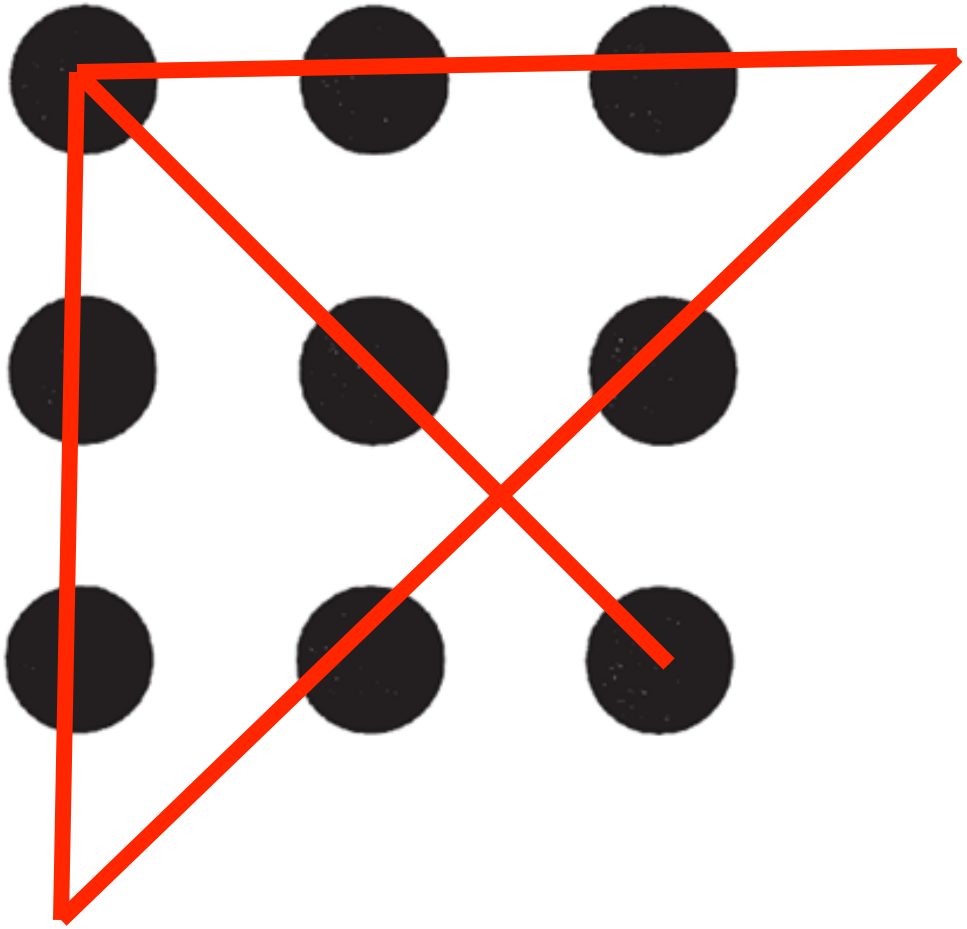
The 9-dot problem

Task 1: Connect all 9 dots using only straight lines



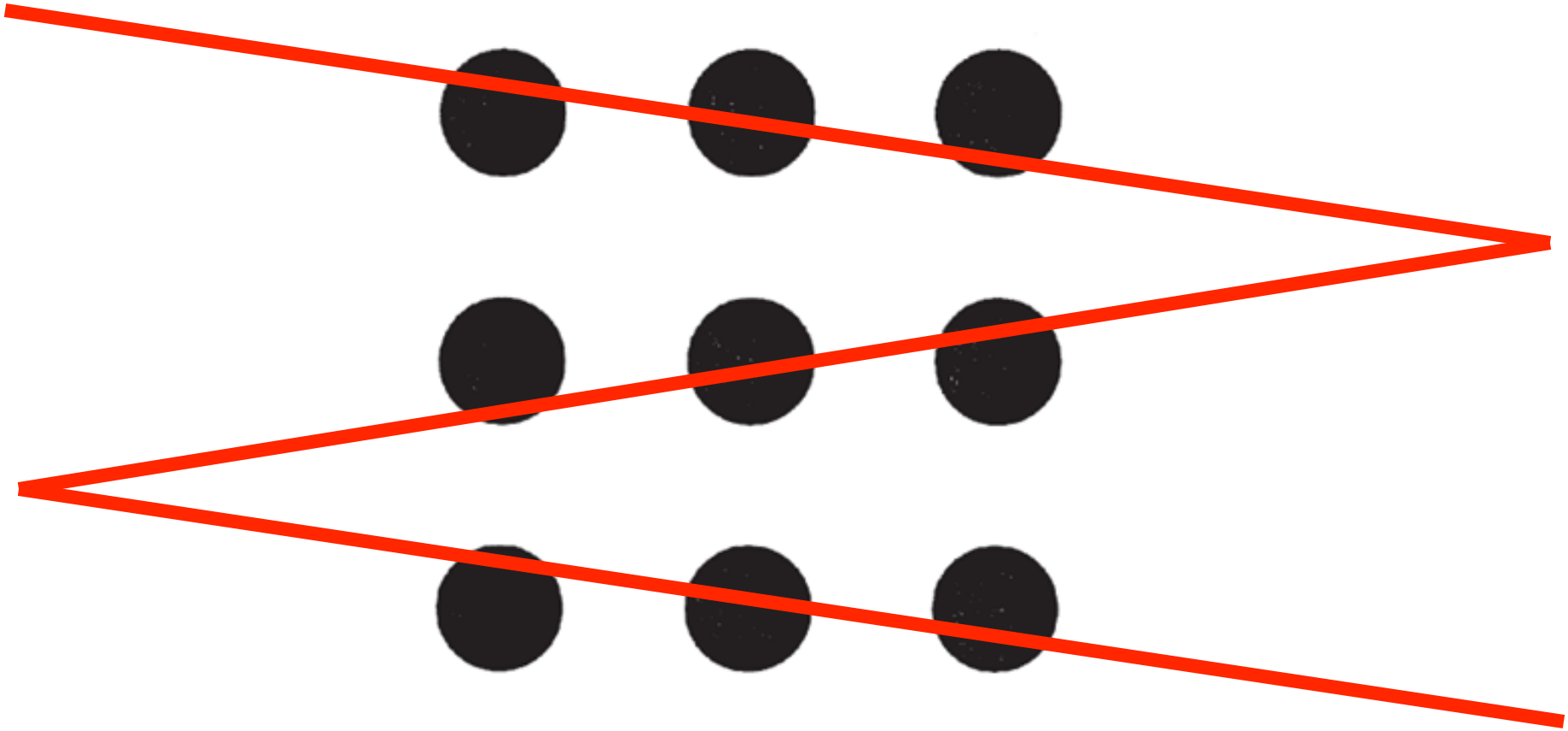
The 9-dot Problem

Task 2: Connect all 9 dots using 4 straight lines



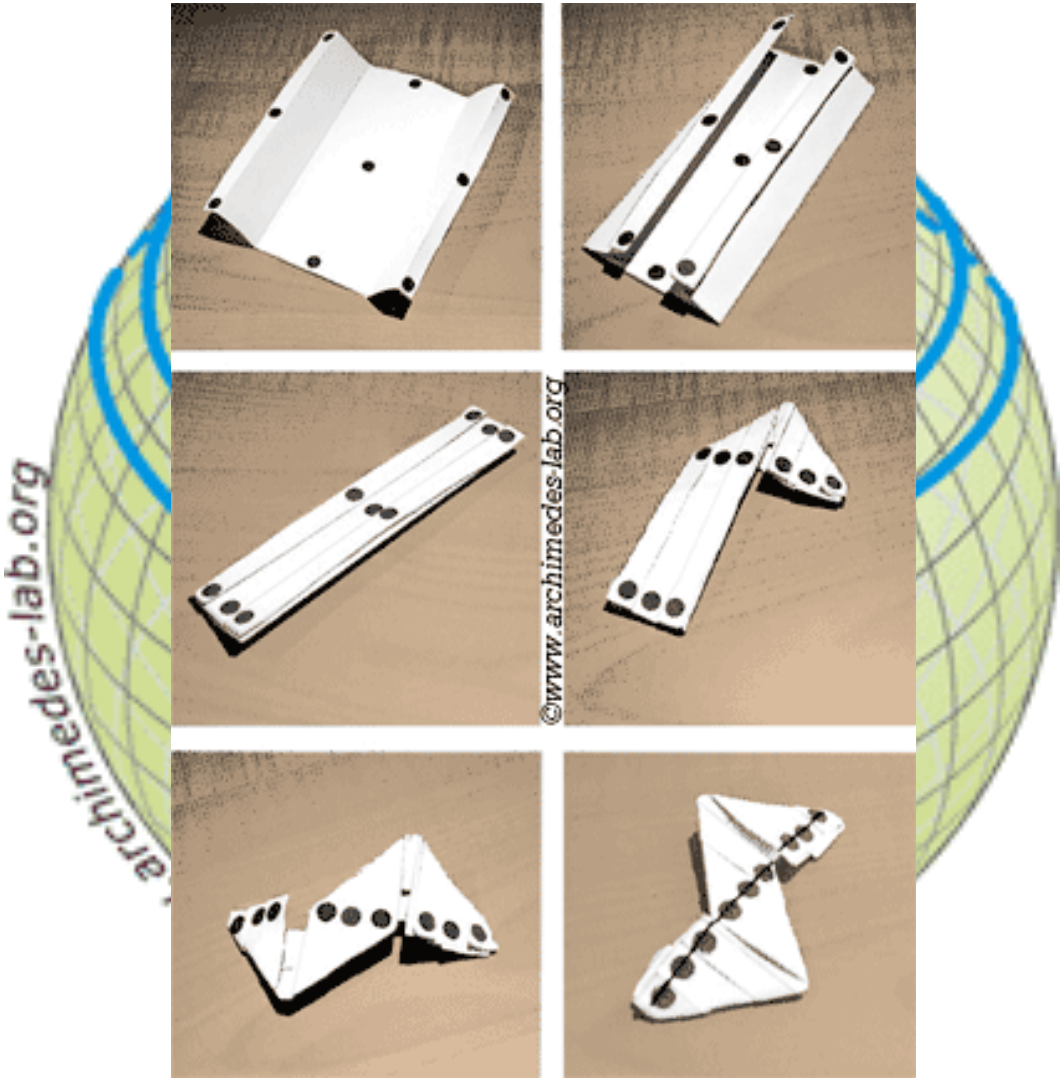
The 9-dot Problem

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

1. We tend to see what we expect to see

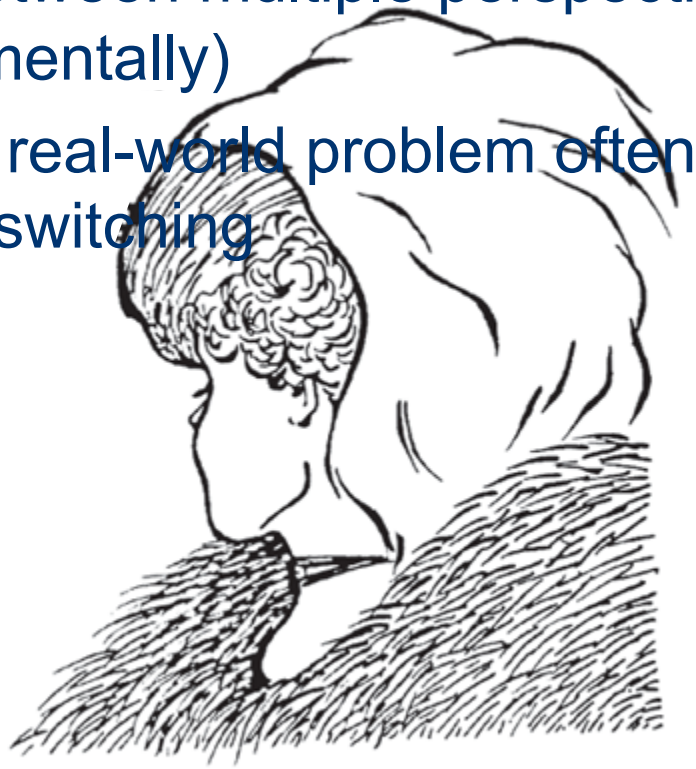
- Mental models are built from prior experience
- We expect new input to “fit” the existing model
- Updates are **expensive**: given input that almost fits, we’ll distort information to avoid re-fitting the model
- **Expectation** is at least as strong as perception

2. Mental models form quickly, and update slowly:

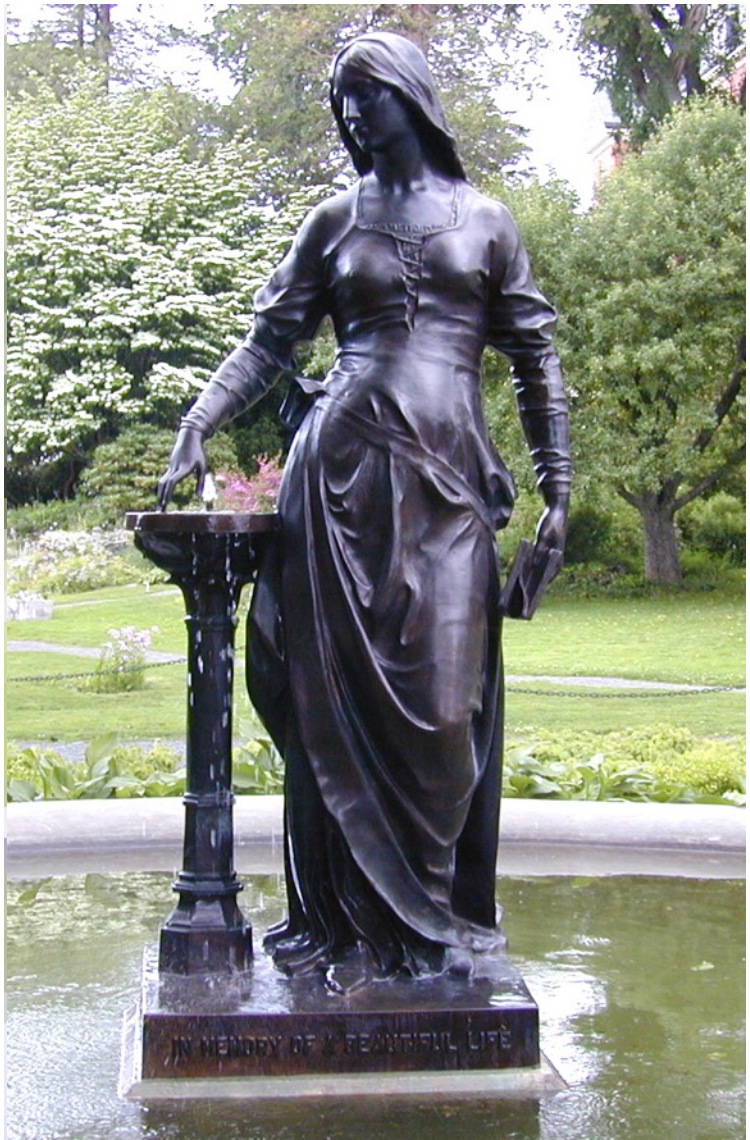
- “First impressions matter”
- Early information can have the highest impact
- The order in which we present pieces of information can shape how a person understands the whole
- Once a mental model is formed, it takes effort to alter it

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



4. Initial exposure interferes with accurate perception



Blur size

128px
64px
32px
16px
8px
None

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**

The good, the bad, and the ugly…

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

Takeaways

- Visualization is about more than just **aesthetics**
- Tools to help us build a **mental model** of the data without actually having to look at all of it
- There are compelling **cognitive reasons** why some visualization techniques are helpful and others aren't

Want to learn more?
Take SDS136 with
Prof. McNamara next fall!

Thanks!

jcrouser@smith.edu

@jordancrouser

www.science.smith.edu/~jcrouser

HUMAN COMPUTATION
& VISUALIZATION
LABORATORY