



# Data Visualization

---

Somayeh Fatahi, Ph.D.

Winter 2026



# Review of Previous Session

- Course overview: structure, tools (Python), assessments, and expectations
- What data visualization is and why it matters
- Visualization vs. summary statistics: limits of numbers alone
- Key lesson from Anscombe's Quartet: identical statistics  $\neq$  identical data
- Importance of visualization for detecting patterns, outliers, and misleading relationships
- Visualization as a tool for faster understanding and better decision-making



# Learning Objectives

- See how data visualization evolved over time
- Understand why theory matters in visualization design
- Learn what abstraction means in visualizations
- Understand how cognitive load affects learning
- Recognize what makes visualizations clear and trustworthy



# History

- People have been using graphics to represent knowledge for a long time

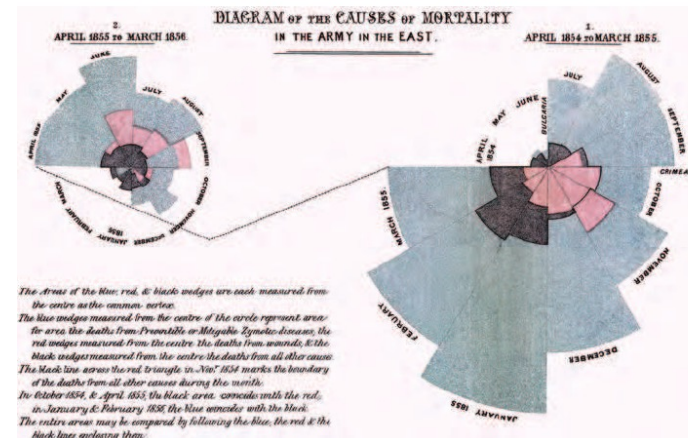
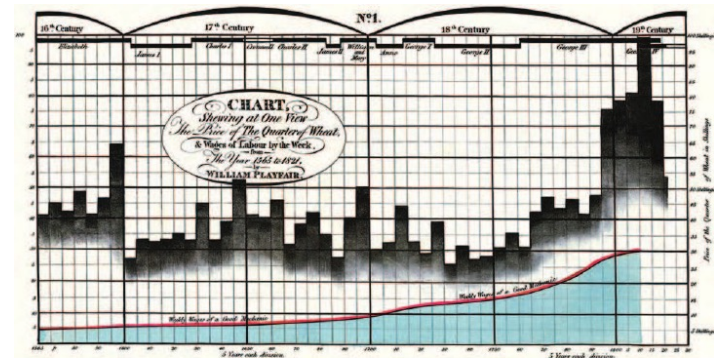


A Babylonian world map (6<sup>th</sup> century BC)

- Starting from 18<sup>th</sup> century, massive tables of numbers are collected.

- William Playfair (1759 – 1823) devised a number of now-familiar graphical devices such as pie, bar, and line charts

- Florence Nightingale expand on the pie charts to create rose diagrams (now known as coxcombs)

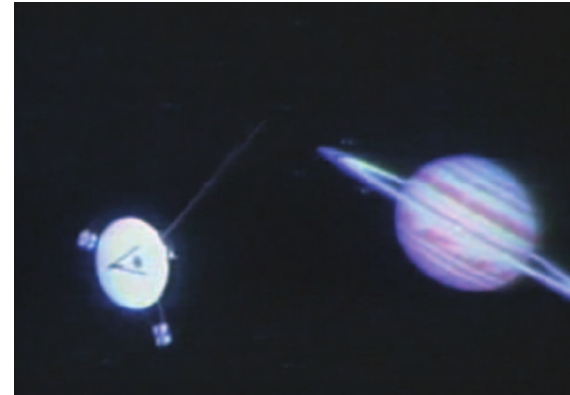


- In 1815, William Smith created the first published geological map of Britain.

*A map that changed the world.*



- The impact of computers
  - complex graphics based on large sets of data to create data-rich graphical visualizations
  - realistic models of complex scientific phenomena in three dimensions
  - interact with the graphics



- The Impact of the Internet
  - Free accessibility have made data more widely available for instantaneous download
  - Internet companies have added simple API to their popular online services such as maps

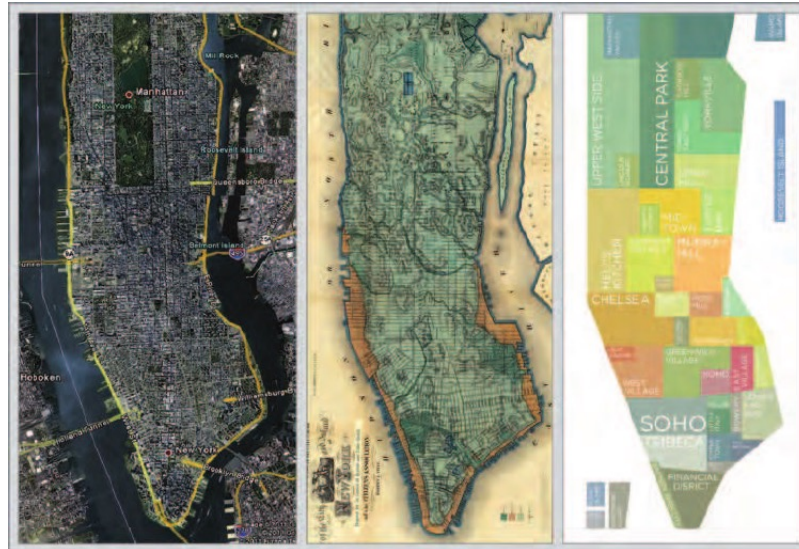


# Models for Information Visualization

- *Theoretical models* that try to offer *insight* as to the fundamental issues at stake.
- *Descriptive models* that create *taxonomies* of the elements that characterize their fundamental attributes.
- *Prescriptive models* that provide a framework to scaffold scholars through the *process* of creating and evaluating information visualizations.

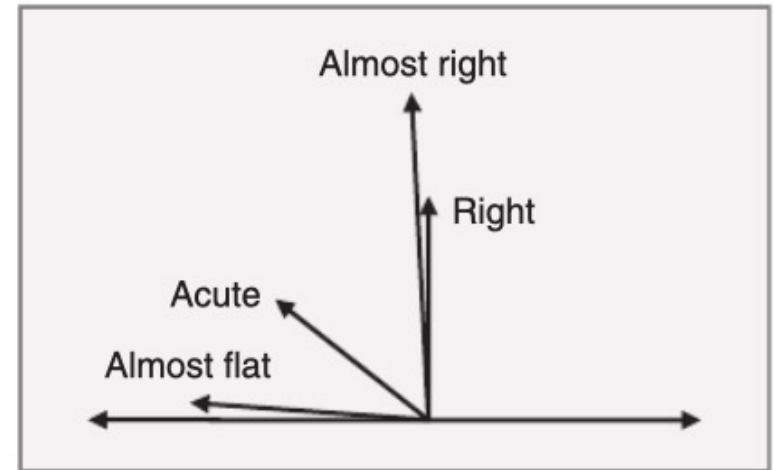
# Theoretical Models

- Abstraction

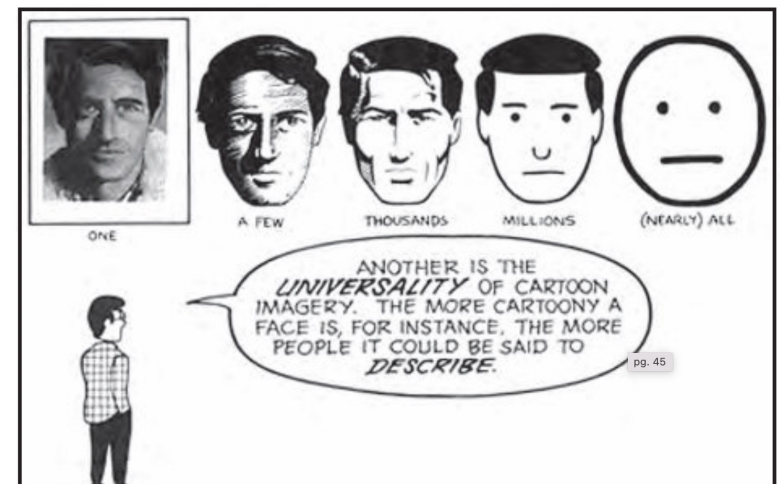


Degrees of abstraction in lower Manhattan

- Continuity
  - Although most measurements taken in the world are continuous in nature, our perceptions tend to divide them into discrete parts.
  - How we make sense of the world and are important to keep in mind when we design visualizations of continuous phenomena



Mileposts along the continuum



Part of the figures included are from the following sources:

- <https://www.slideshare.net/jmborda86/the-split-attention-principle-in-multimedia-learning>
- [https://www.researchgate.net/figure/Modality-Effect\\_fig1\\_269112838](https://www.researchgate.net/figure/Modality-Effect_fig1_269112838)
- <https://3starlearningexperiences.wordpress.com/tag/modality-principle/>
- <http://ggregor.weebly.com/cognitive-load.html>
- <https://slideplayer.com/slide/5720365/>
- <https://sites.google.com/site/cognitivetheorymmllearning/contiguity-principle>
- <https://www.slideshare.net/Baynard/personalization-principle-by-group-a>
- <https://www.youtube.com/watch?v=BcWSUnXz8kw>

- Cognitive Load Theory (CLT)
  - CLT has emerged as one of the most important factors in understanding, learning, and overall attention
  - Memory is comprised of two primary structures — short term and long term
  - The aim of all instruction is to alter long-term memory, but information must first pass through short-term memory
  - Short-term memory can hold up to 7 ( $\pm 2$ ) items (which will be lost within 20 seconds without rehearsal).
  - CLT in data visualization
    - present information in a comprehensible manner.
    - respect the limitations and make use of the affordances suggested by CLT about people's ability to accept information

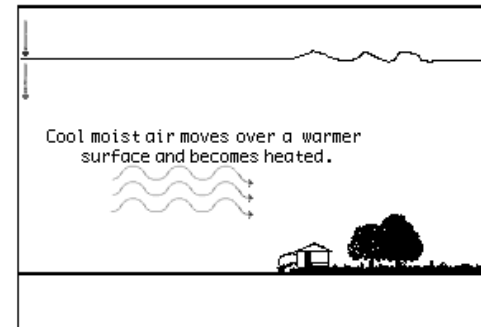
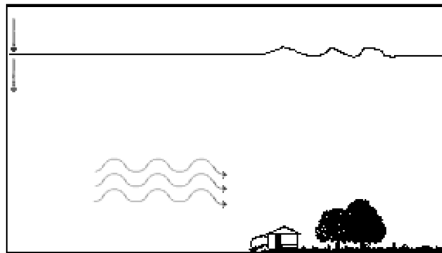
- Multiple Representation/ Multimedia

- Different types of information representations are relatively independent to absorb
- Multiple-channel learning Principles (Richard Mayer, 2005)

1. *Multimedia principle*: People learn better from words and pictures than from words alone.

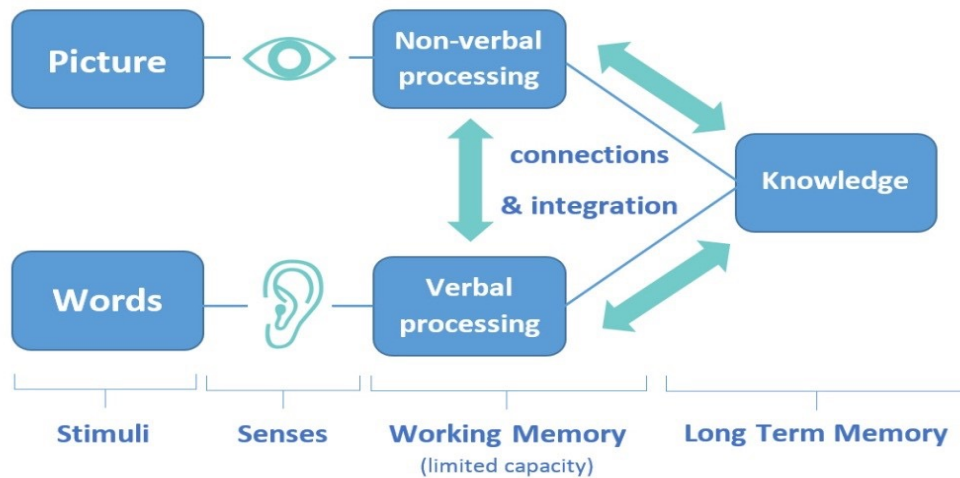
2. *Split attention principle*: it is important to avoid formats that require learners to split their attention between, and mentally integrate, multiple sources of information

Cool moist air moves over a warmer surface and becomes heated.



3. *Modality principle*: People learn better from graphics and narration(audio) than from graphics and text.

### Allan Paivio's Dual-Coding Theory



4. *Redundancy principle*



Redundant Instruction

Bears enjoy eating Honey.

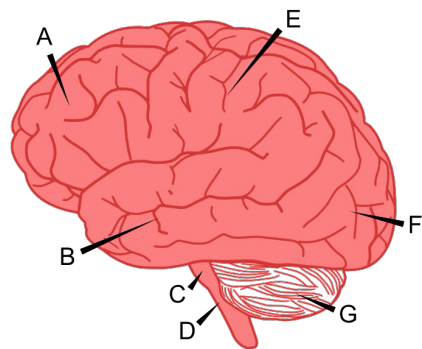
-Bears enjoy eating Honey

Non-Redundant

Bears enjoy eating Honey.

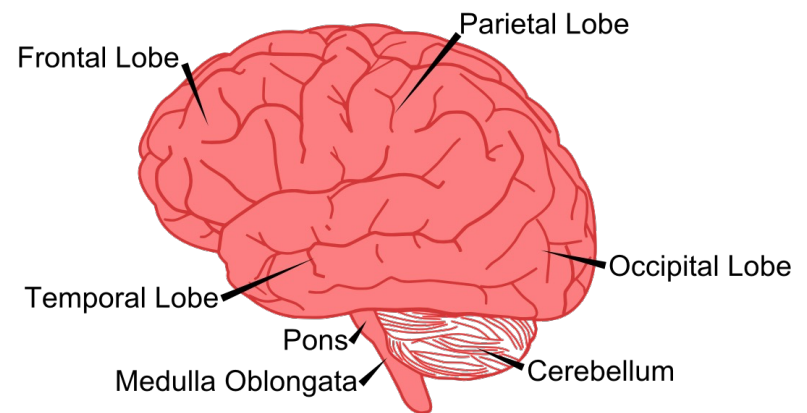
6

5. *Spatial Contiguity Principle*: People learn more deeply from a multimedia message when corresponding words and pictures are presented near rather than far from each other on the page or screen



A - Frontal Lobe  
B - Temporal Lobe  
C - Pons  
D - Medulla Oblongata

E - Parietal Lobe  
F - Occipital Lobe  
G - Cerebellum



6. *Temporal contiguity principle*: People learn better when corresponding words and pictures are presented simultaneously rather than successively

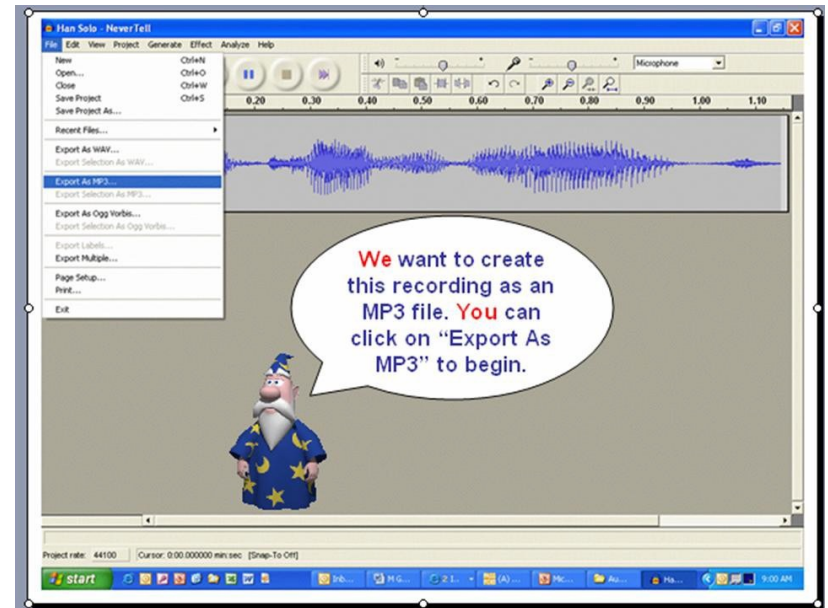
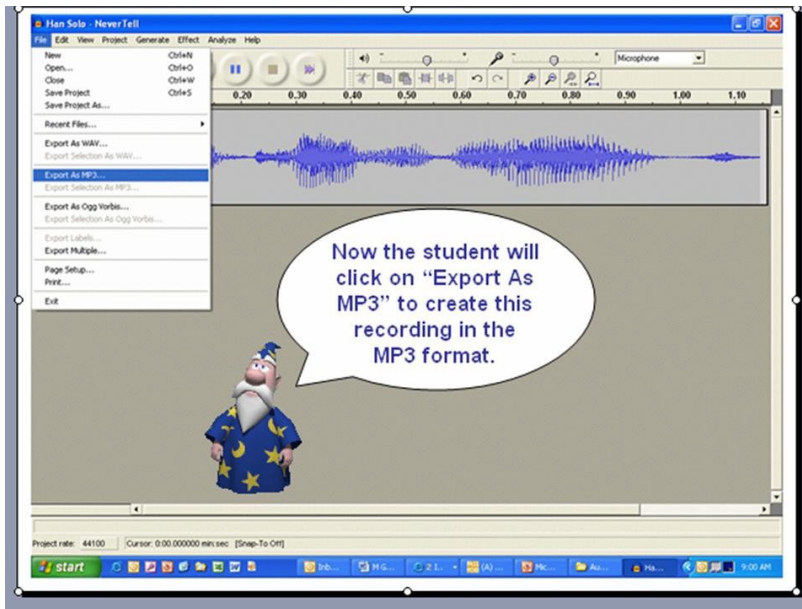
7. *Coherence Principle: People learn more deeply from a multimedia message when extraneous material is excluded than included.*

8. *Signaling Principle: People learn more deeply from a multimedia message when cues are added that highlight the critical aspects of the presented information.*

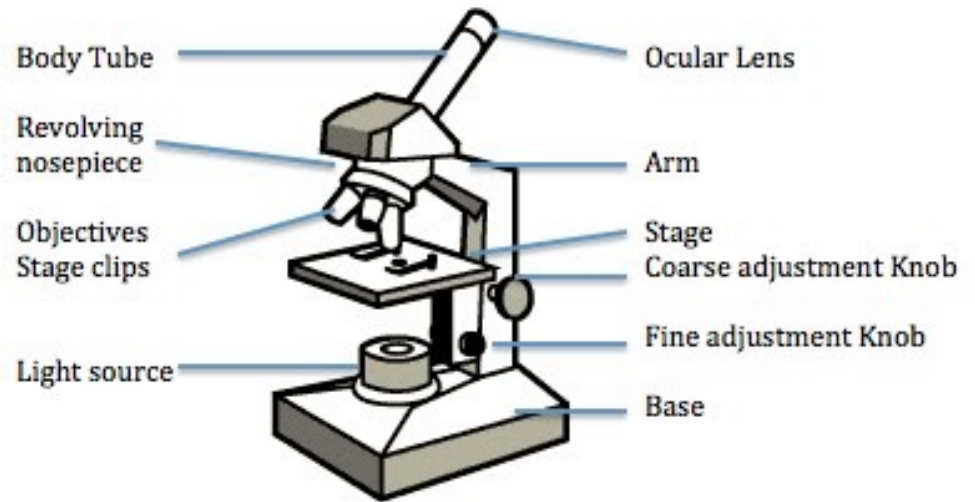
e.g. headings, highlighting, emphasis...



9. *Personalization Principle*: People learn more deeply when the words in a multimedia presentation are in conversational style rather than formal style.



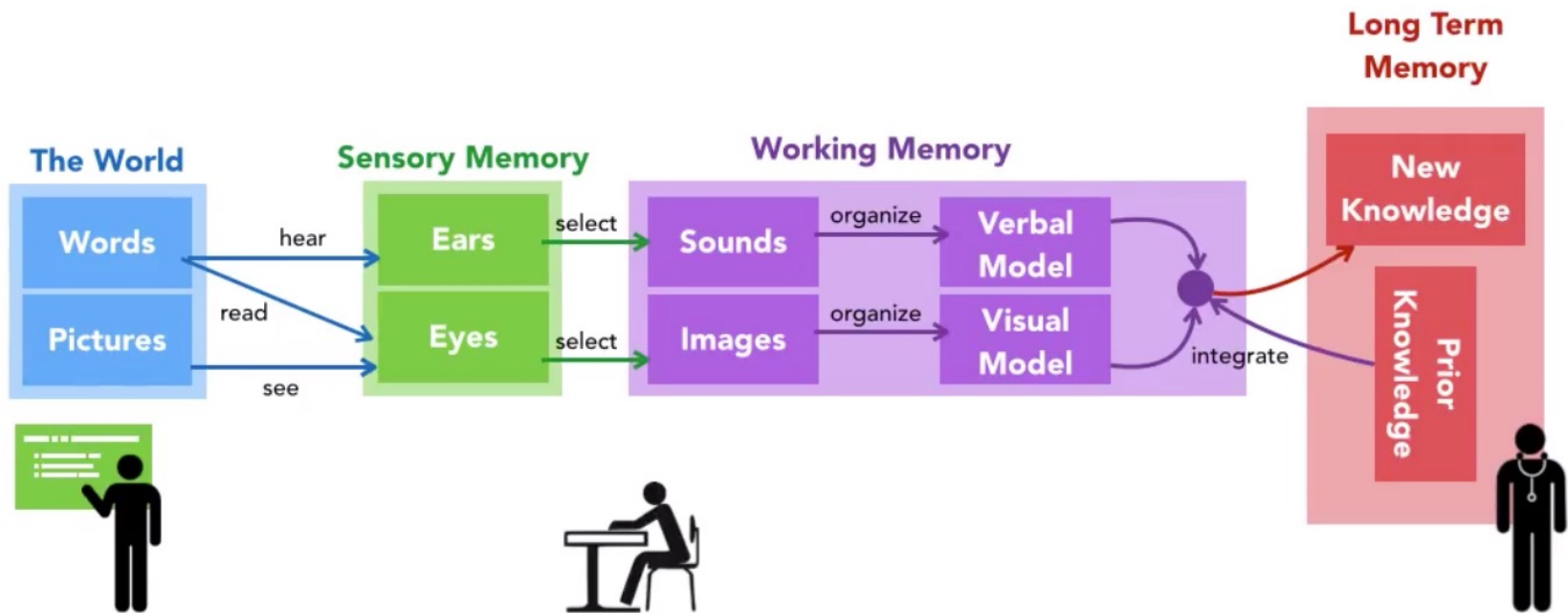
10. *Pre-training Principle:* People learn more deeply from a multimedia message when they know the names and characteristics of the main



Place the Slide on the Microscope

- Stage Clips are not necessary
  - Click Nosepiece to the lowest (shortest) setting – Scanning Objective
  - Look into the Eyepiece
  - Use the Coarse Focus
  - Once the slide is focused, rotate the nosepiece to the low power objective (medium sized)
  - Refocus using the coarse (large) knob
- Move slide to get a centered view

# Walk through Cognitive Model



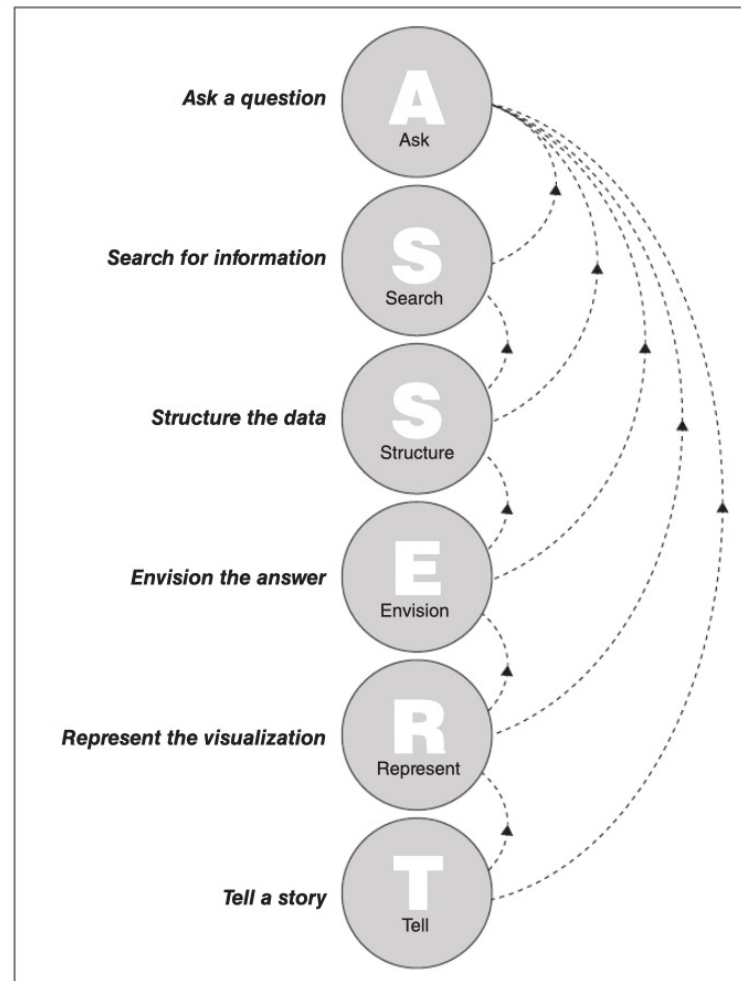
# Models for Information Visualization

- *Theoretical models*
  - Abstraction
  - Continua
  - *Cognitive Load Theory and Multimedia Principles*
- *Descriptive models*
- *Prescriptive models*

# Prescriptive Models

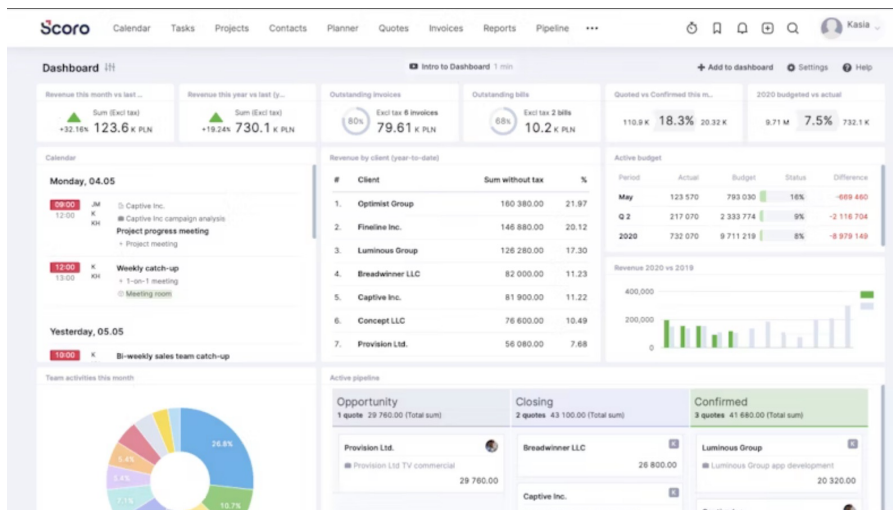
- *Prescriptive models* seek to provide support in the creation, understanding, and evaluation of visualizations.
  - *Shneiderman's Visual Information-Seeking Mantra*
  - *Van Ham and Perer's Search and Expand*
  - *Chi's Data State Reference: analytical abstraction, visualization abstraction, and the view*
  - *Fry's Acquire, Parse, Filter, Mine, Represent, and Interact*
  - *The ASSERT model*

- The ASSERT model
  - Support the creation of visualizations that are accessible, insightful, educational, compelling
  - Includes the complete developmental life cycle of visualizations



# Question 1:

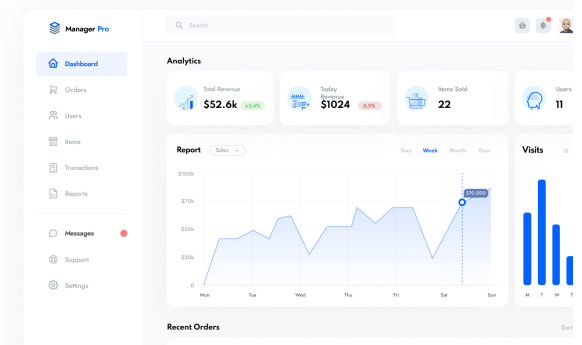
Which dashboard would you trust more?  
Why does the crowded one feel stressful?



## Simple Dashboard UI

Analytics  
Report  
Recent order  
Visits

Design Space  
@jabastin



# Answer

- Based on Cognitive Load Theory, which we learned today, our working memory is limited.
- When a visualization shows too many elements at once, viewers spend mental effort deciding what to focus on instead of understanding the data.
  - Too many colors → higher cognitive load
  - Too many charts → split attention
  - No visual hierarchy → confusion

# Question 2:

## Which map would you use to drive? Which one for transit?



1



2

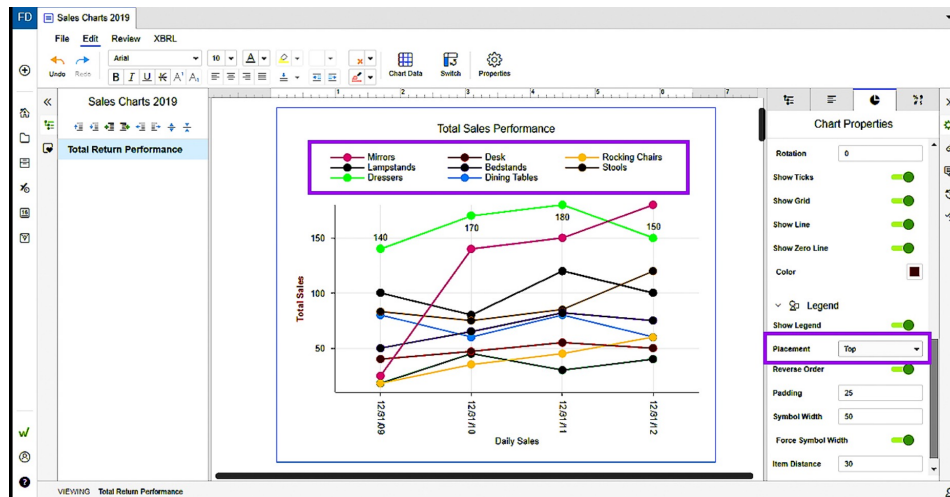


3

# Answer

- Abstraction removes details based on the task.
- For driving
  - the street map because it clearly shows roads, intersections, and navigation details.
- For public transit
  - the transit map because it emphasizes subway or bus lines and connections, even though it removes many geographic details.
- All of these maps show the same city.  
The data did not change—the task did.

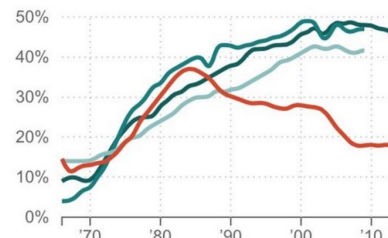
# Question 3: Which chart requires less effort?



## What Happened To Women In Computer Science?

% Of Women Majors, By Field

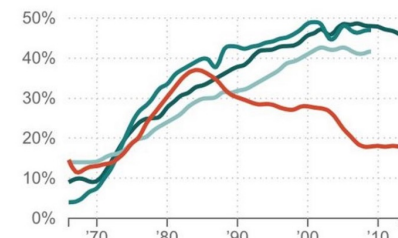
■ Medical School    ■ Law School  
■ Physical Sciences    ■ Computer science



## What Happened To Women In Computer Science?

% Of Women Majors, By Field

■ Medical School  
■ Law School  
■ Physical Sciences  
■ Computer Sciences



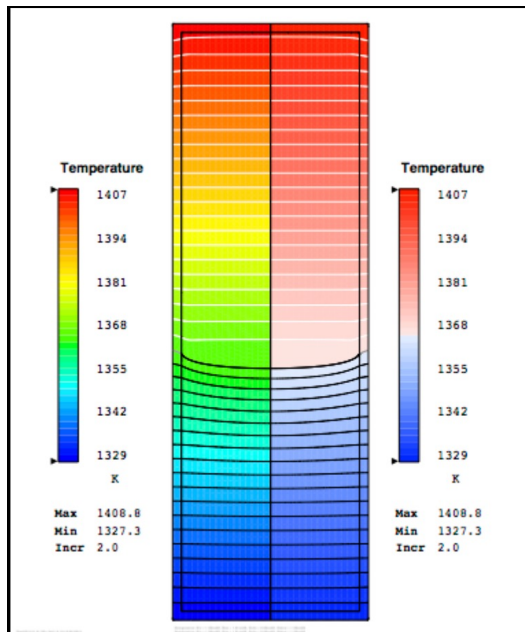
# Answer

- In the left chart, the legend is far from the data, so your eyes move back and forth.
- In the right chart, the labels are on the data itself, which makes it easier to understand.
- When attention is split between multiple locations, understanding becomes harder and slower.
- Reducing eye movement reduces cognitive load.

# Question 4:

Where would you draw the boundary between cold, warm, and hot?

Would everyone choose the same boundaries?



**Disclaimer: Always check with local officials for appropriate actions and activity levels. Experienced heat stress will depend upon duration and intensity of activity and personal health and vulnerability.**

WBGT by Region (°F)			Threat Level WBGT at these values increasing heat stress.	Risk of heat illness
Region 1	Region 2	Region 3		
< 72.3	< 75.9	< 78.3	Low Threat	Increased risk for heat illness 
72.3 - 76.1	75.9 - 78.7	78.3 - 82.0	Elevated Threat	
76.2 - 80.1	78.8 - 83.7	82.1 - 86.0	Moderate Threat	
80.1 - 84.0	83.8 - 87.6	86.1 - 90.0	High Threat	
>84.0	>87.6	>90.0	Extreme Threat	

Regions are from Grundstein, A., Williams, C., Phan, M and Cooper, E., 2015. Regional heat safety thresholds for athletics in the contiguous United States. *Applied Geography*, 56, pp.55-60. 10.1016/j.apgeog.2014.10.014.

# Answer

- Those decisions influence how users interpret the data.
- **Continuous picture:**  
Good for precision and showing smooth variation
- **Categorized picture:**  
Good for communication and fast interpretation
-