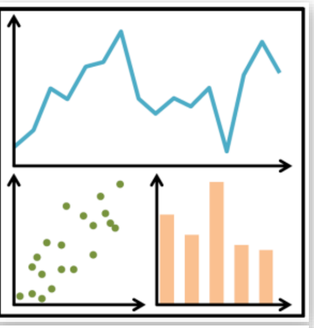
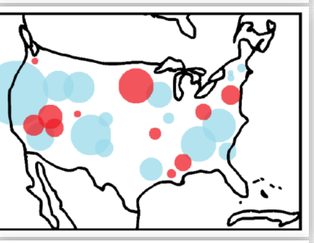
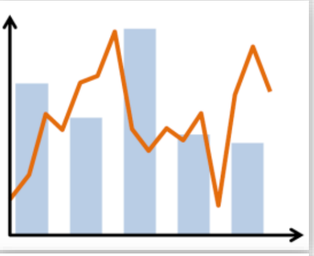
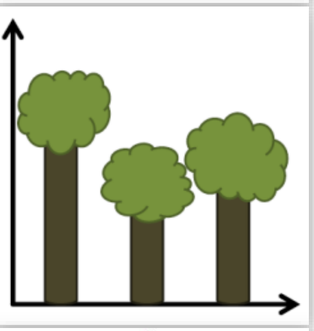
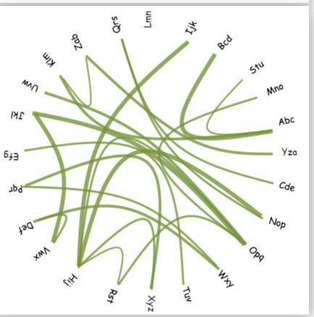






Lecture 5: Design Rules of Thumb, Marks and Channels, Data Types

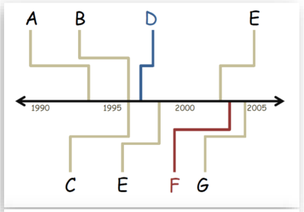
CS 7250

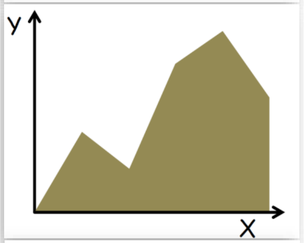
SPRING 2021

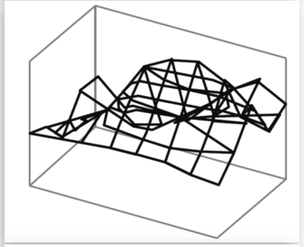
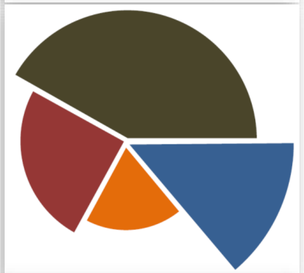
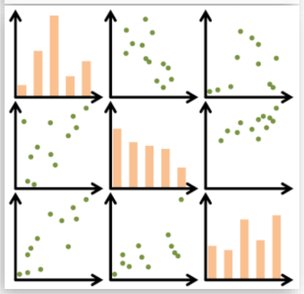
Prof. Cody Dunne

NORTHEASTERN UNIVERSITY

Slides and inspiration from Michelle Borkin, Krzysztof Gajos, Hanspeter Pfister, Miriah Meyer, Jonathan Schwabish, and David Sprague





Checking In

READING QUIZ

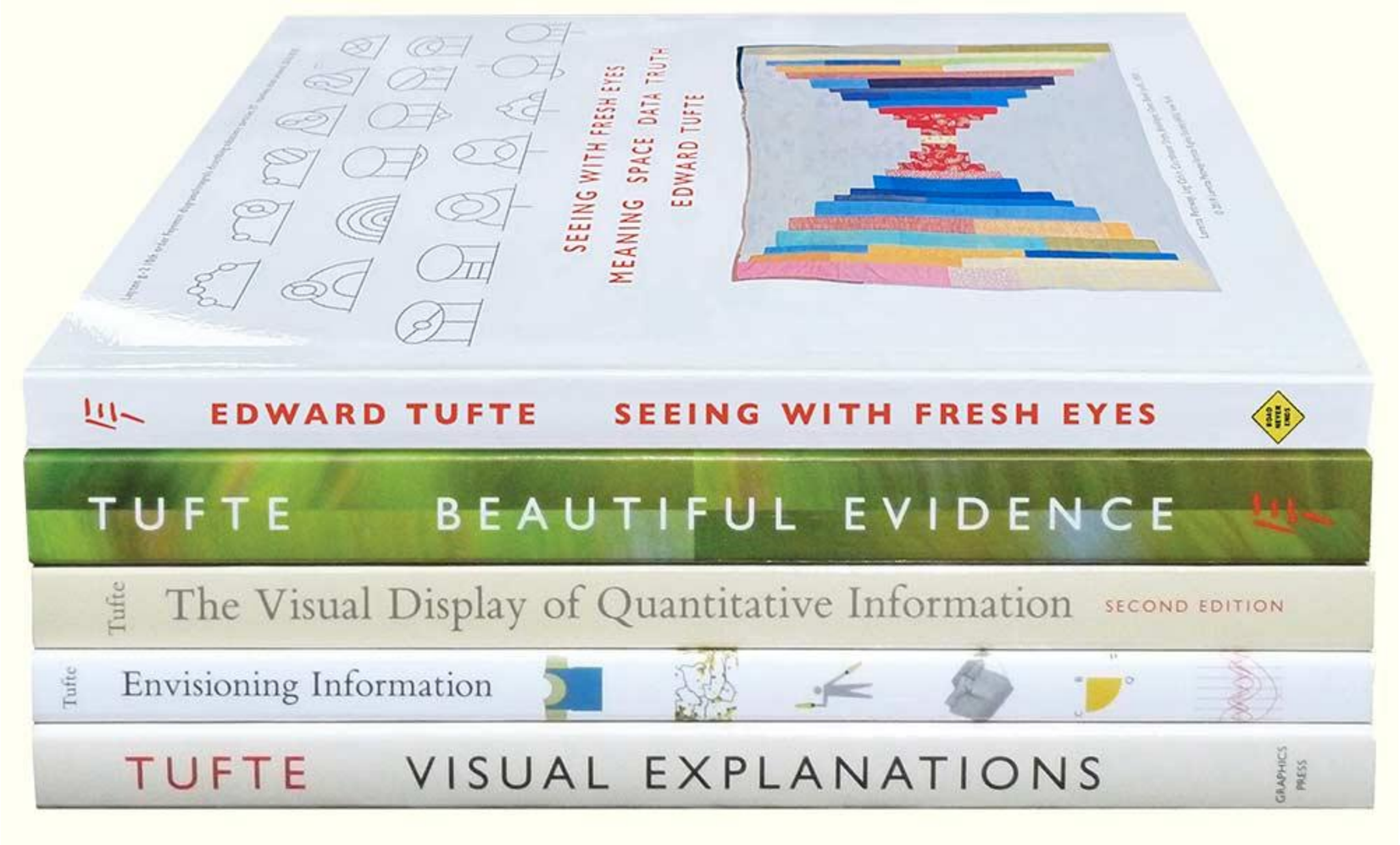
Quiz — Data Types & Tasks

Password: “not_first_quiz”

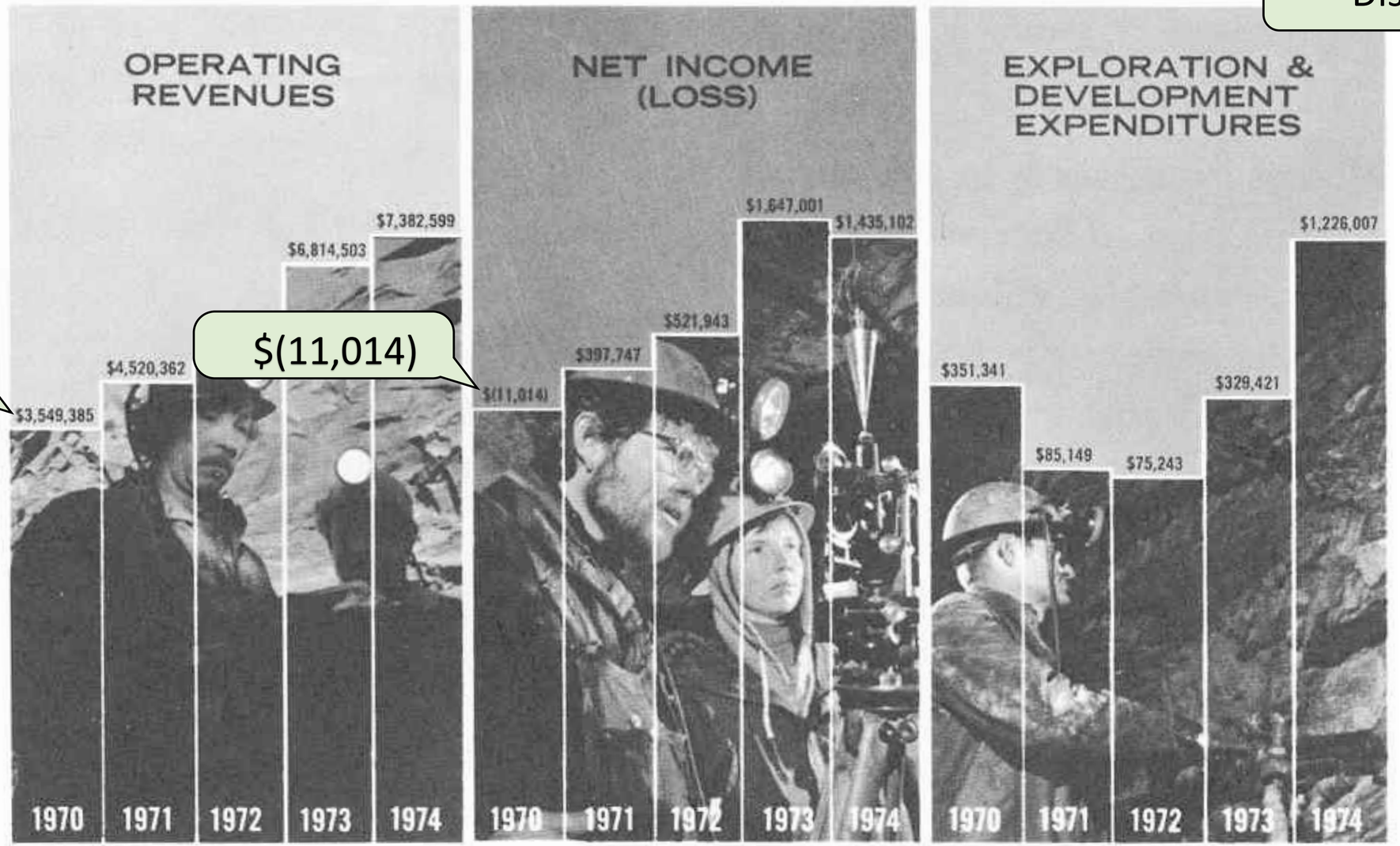
Previously, on CS 7250…

Design Rules of Thumb

Edward Tufte



"Distorted Scales"



$3,549,385

$(11,014)

y-axis baseline?!

"Clear, detailed, and thorough labeling should be used to defeat graphical distortion and ambiguity. Write out explanations of the data on the graphic itself. Label important events in the data."

Tufte, "Visual Display of Quantitative Information" (1983)

Lie Factor

Lie Factor = $\frac{\text{(Size of effect in graphic)}}{\text{(Size of effect in data)}}$

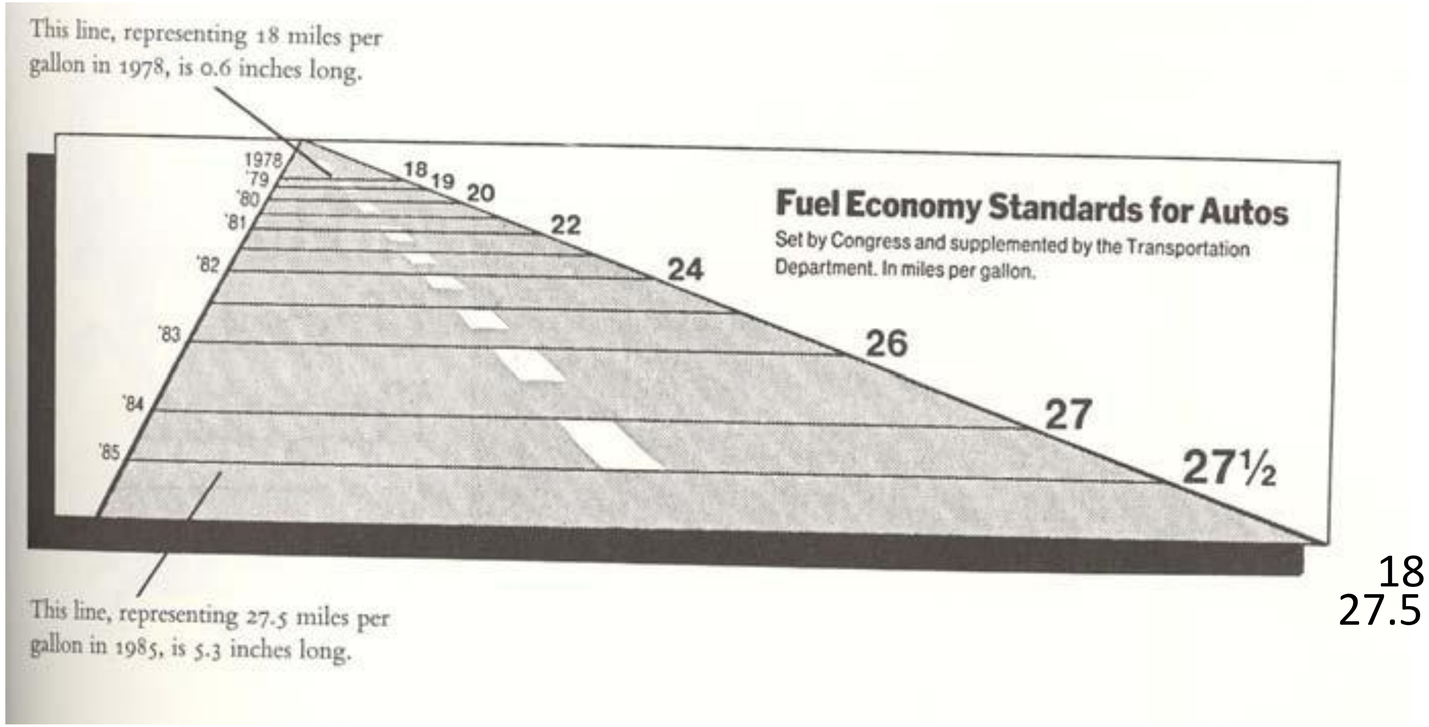


Image = $\frac{5.3" - 0.6"}{0.6"}$ = 7.83 = 783%

Data = $\frac{27.5 - 18}{18}$ = 0.53 = 53%

Lie Factor = $\frac{783\%}{53\%}$ = 14.8

Lie Factor = >1, overstating

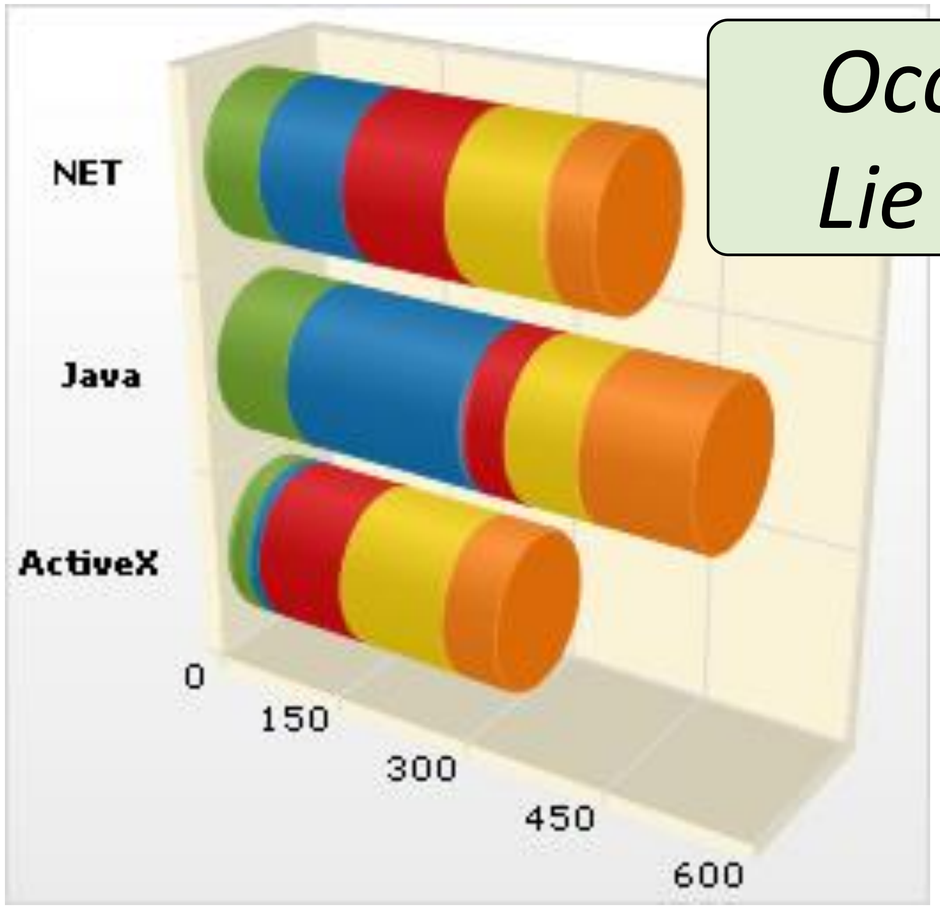
18
27.5

"The representation of numbers, as physically measured on the surface of the graphic itself, should be directly proportional to the numerical quantities measured."

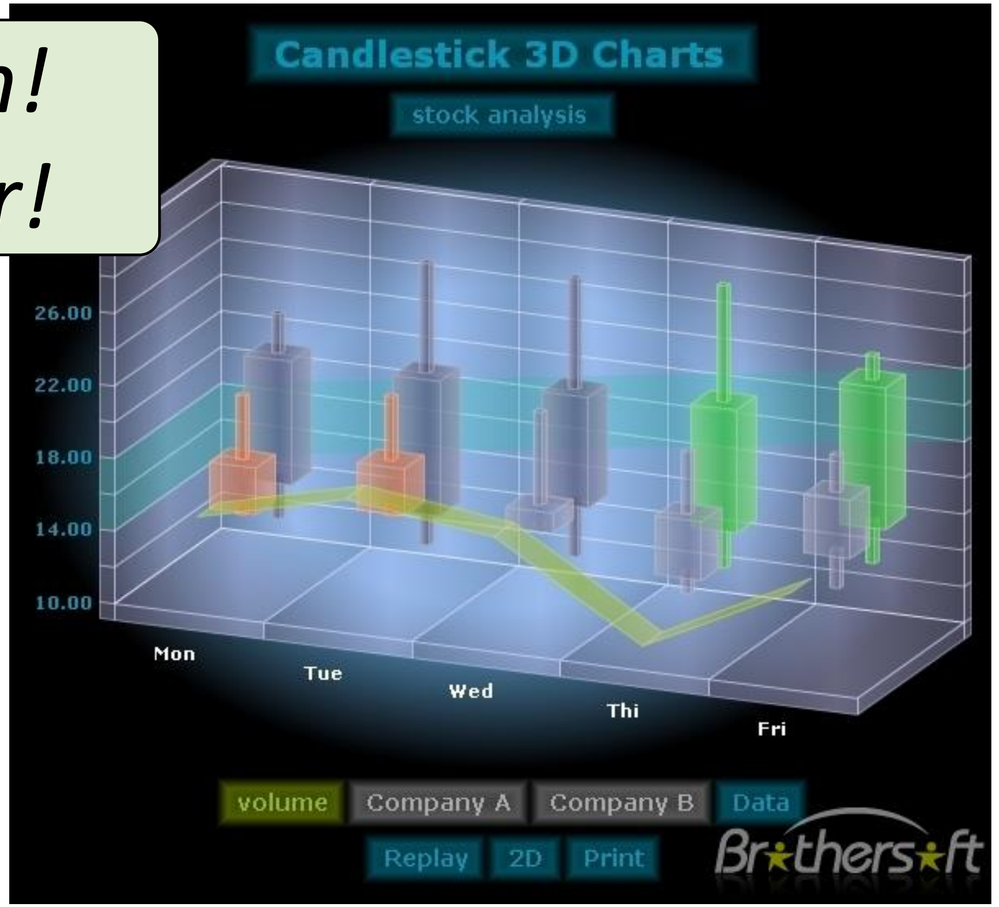
Tufte, "Visual Display of Quantitative Information" (1983)

“No Unjustified 3D”

Occlusion!
Lie Factor!



http://help.infragistics.com/Help/Doc/WinForms/2014.2/CLR4.0/html/Images/Chart_Bar_Chart_03.png



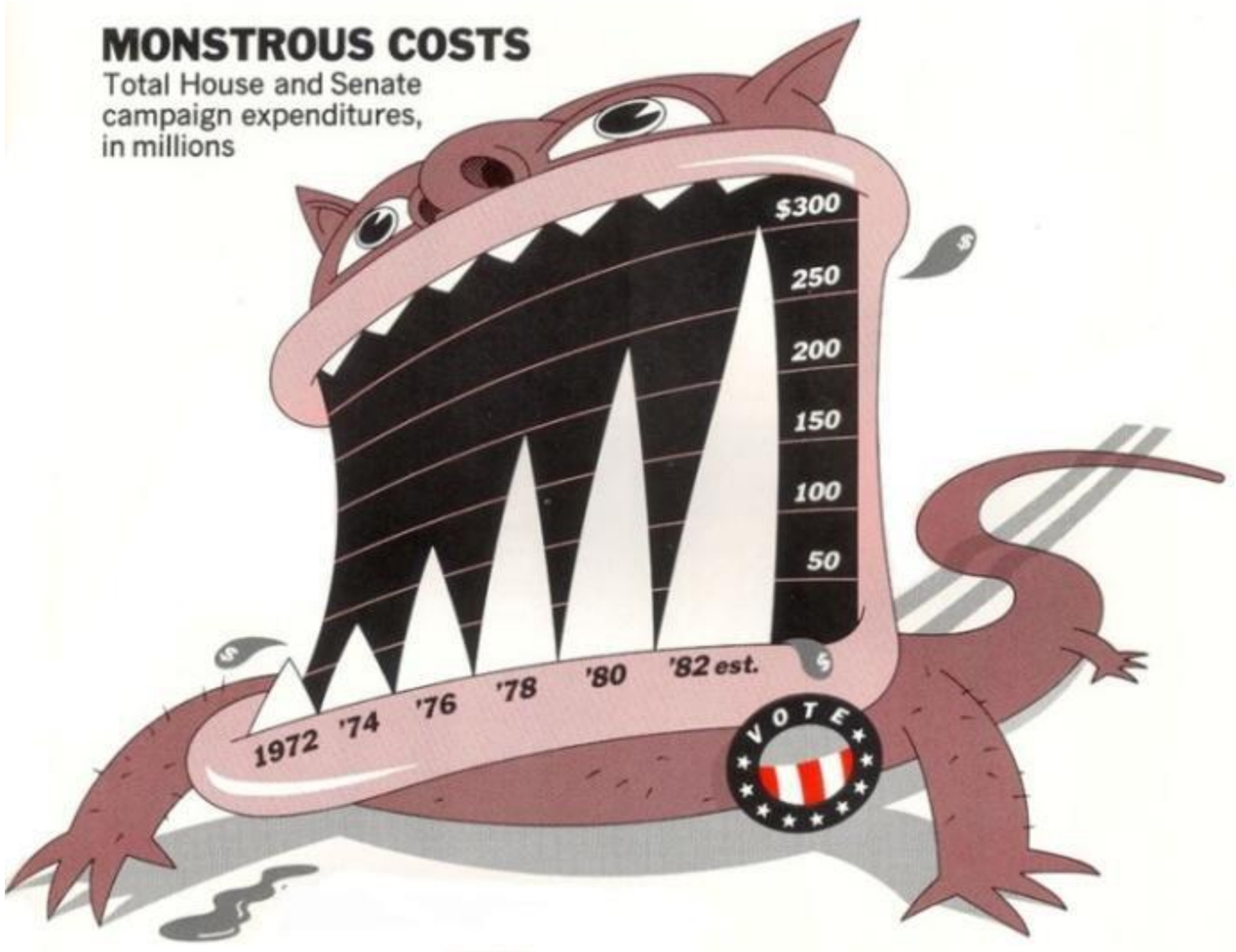
http://img.brothersoft.com/screenshots/softimage/0/3d_charts-171418-1269568478.jpeg

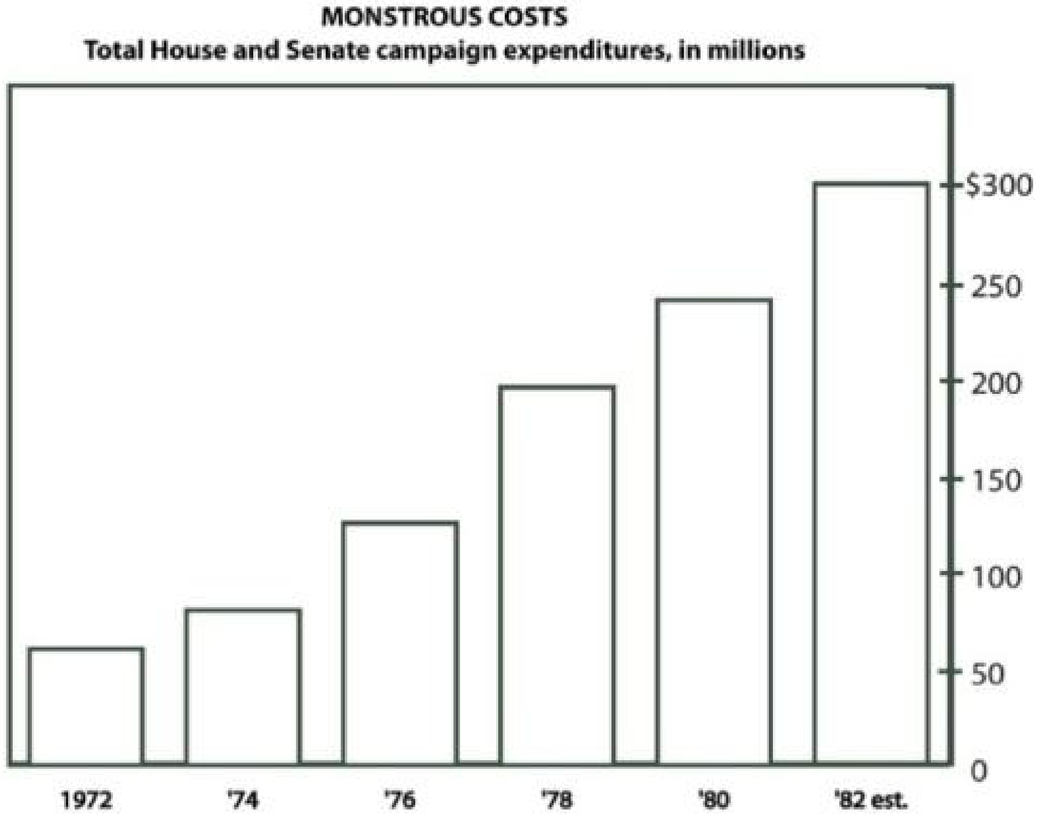
“The number of information-carrying (variable) dimensions depicted should not exceed the number of dimensions in the data.”

Now, on CS 7250…

“CHART JUNK”

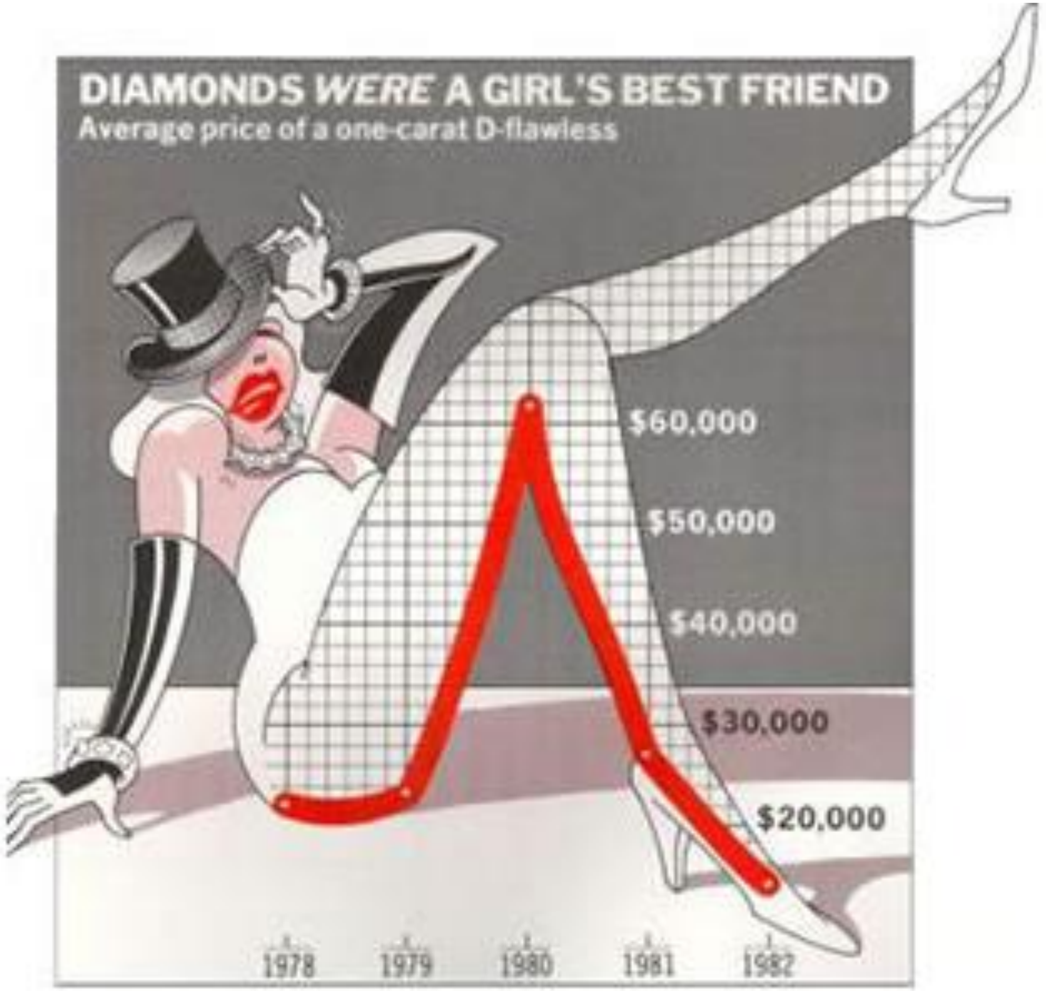
“Chart Junk”

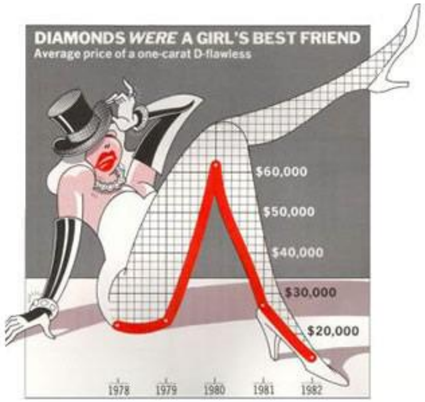


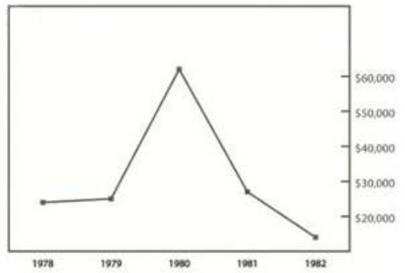


Bateman, et al. (2010)

“Chart Junk”



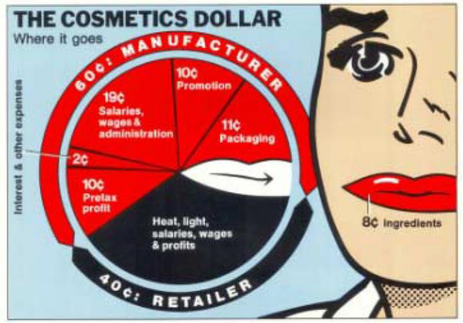


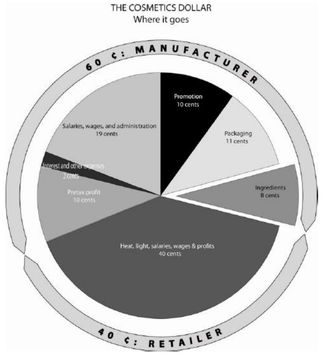


Bateman, et al. (2005)

“Chart Junk Debate”

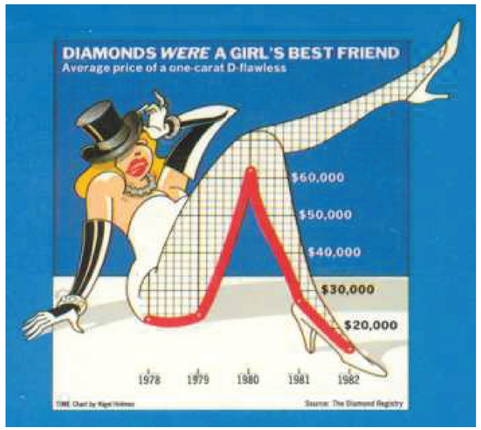
Useful Junk? The Effects of Visual Embellishment on Comprehension and Memorability of Charts





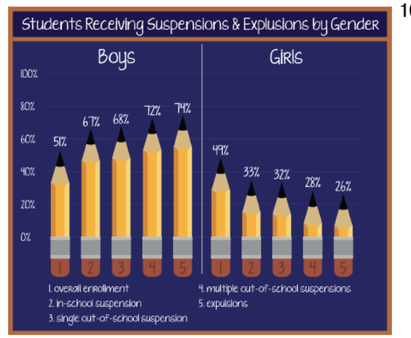
Bateman, et al. (2010)

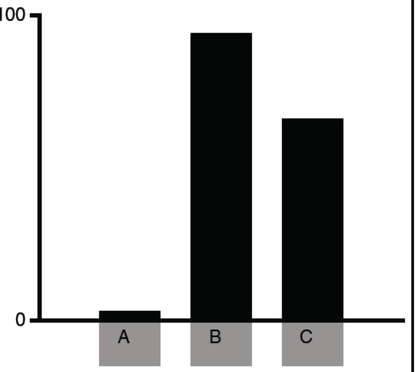
Benefitting InfoVis with Visual Difficulties



Hullman, et al. (2011)

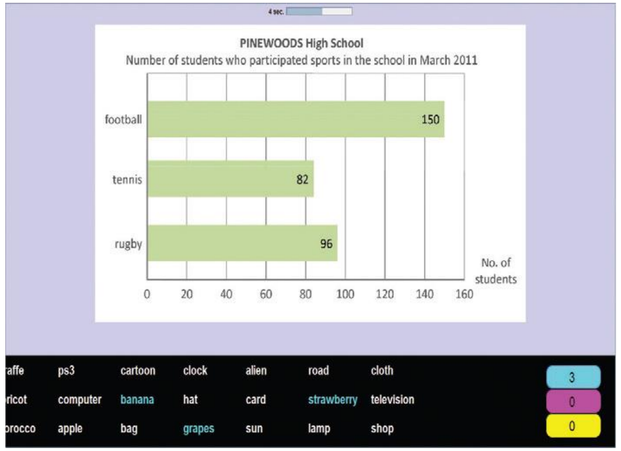
An Evaluation of the Impact of Visual Embellishments in Bar Charts





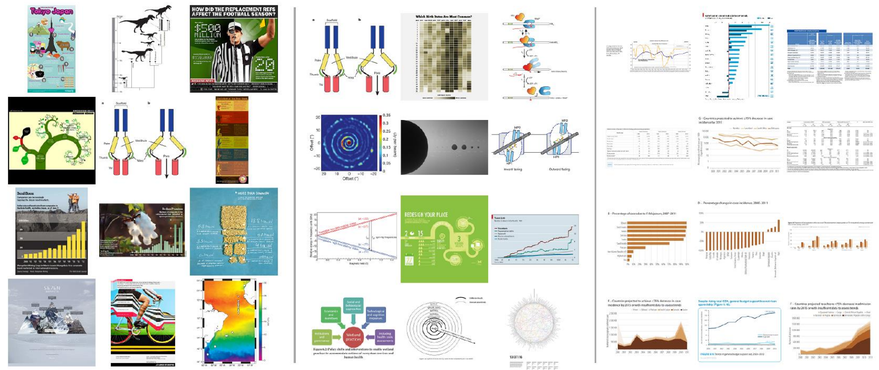
Skau, et al. (2015)

An Empirical Study on Using Visual Embellishments in Visualization



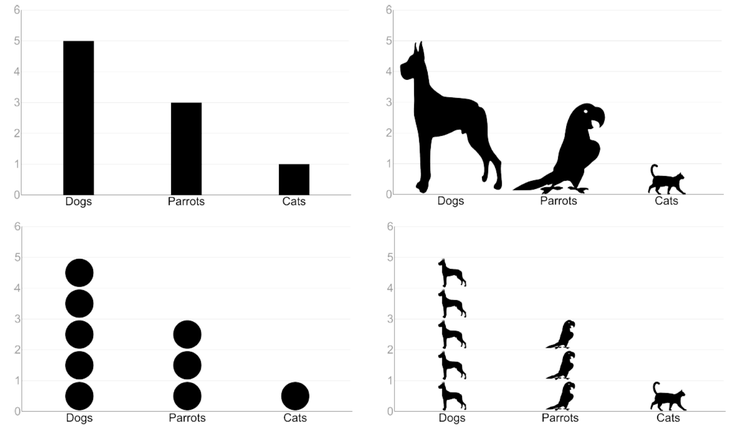
Borgo, et al. (2012)

What makes a visualization memorable?



Borkin, et al. (2013)
Borkin, et al. (2015)

ISOTYPE Visualization – Working Memory, Performance, and Engagement with Pictographs



Haroz, et al. (2015)

“Chart Junk”

Chart junk can... persuade, help with memorability, engage

... bias, limit data-ink ratio, clutter, lower trust

<u>Take-away</u>: it depends on your audience, task, and context...

MARKS AND CHANNELS

Goals for Today

- Learn the basic visual primitives of visualizations (marks and channels)
- Understand how marks and channels are assembled to make visualizations
- Learn which marks and channels are most effective for a given task (“perceptual ordering”)

Visualization Building Blocks

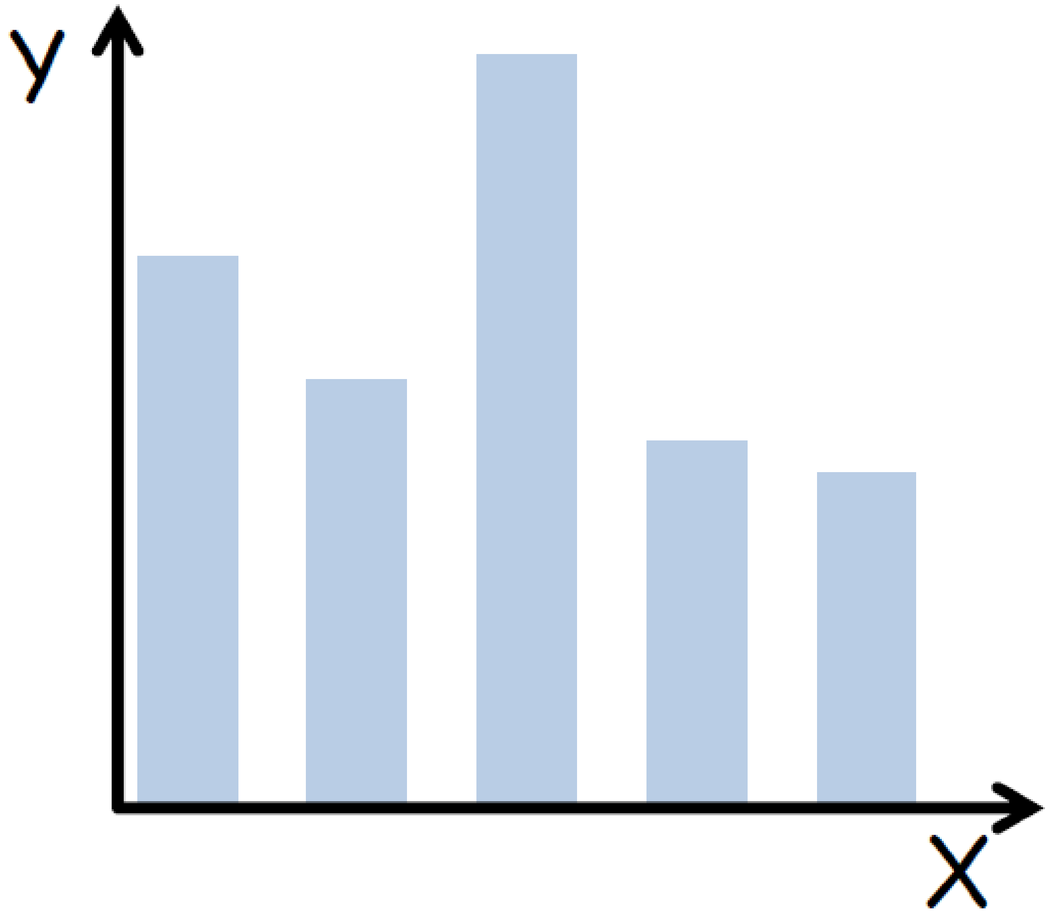
MARK = basic graphical element in an image

Visualization Building Blocks

CHANNEL = way to control the appearance of marks, independent of the dimensionality of the geometric primitive

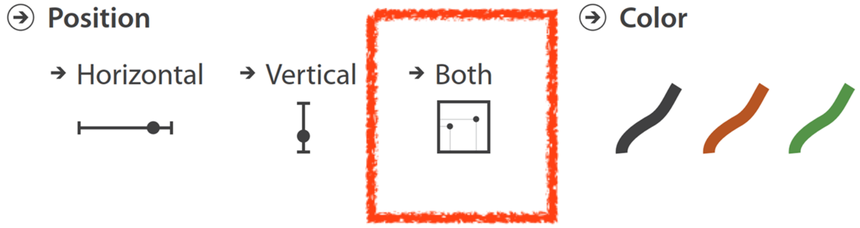
Visualization Building Blocks

of attributes encoded: 2



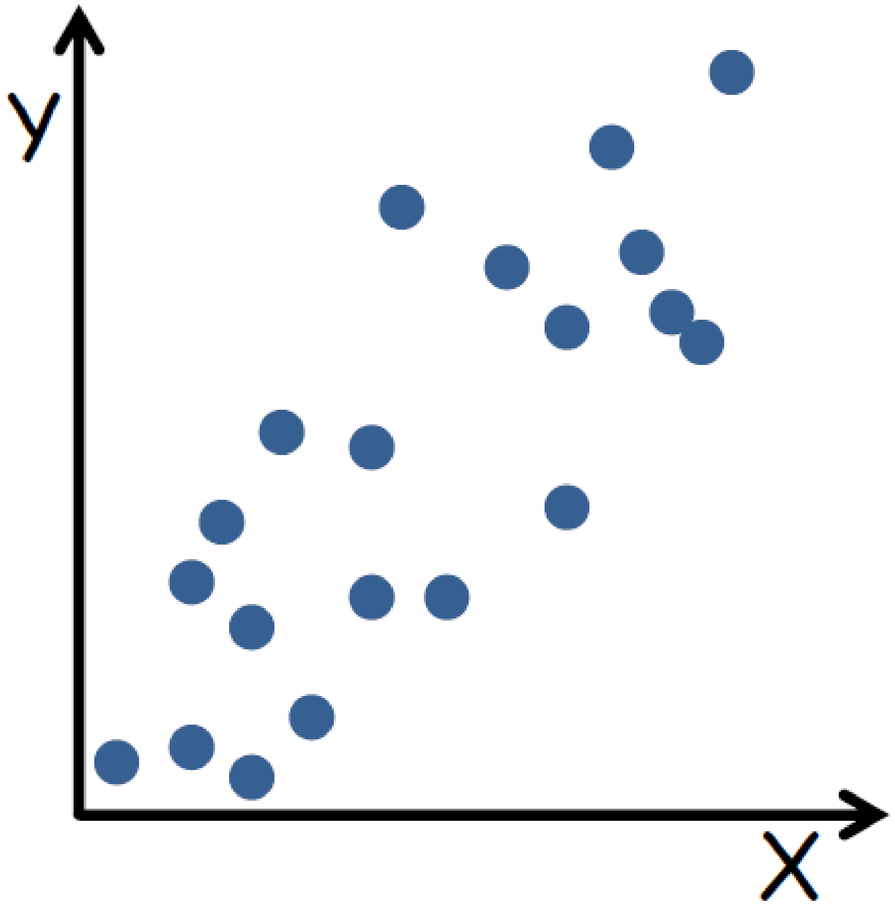
MARK:

CHANNEL :

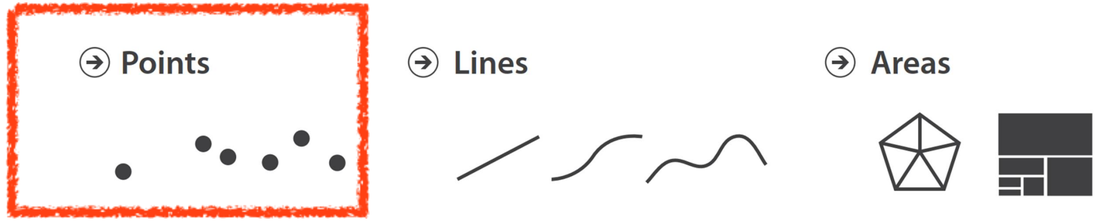


Visualization Building Blocks

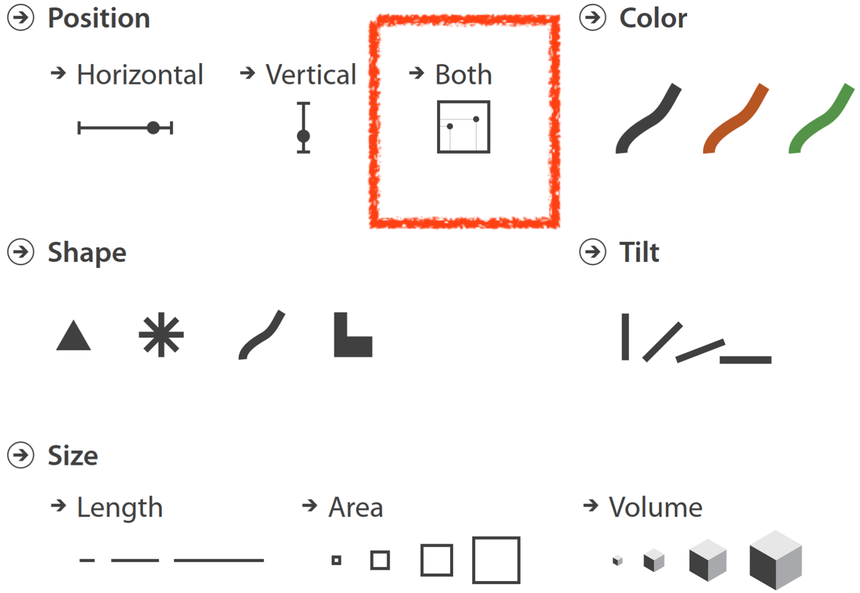
of attributes encoded: 2



MARK:

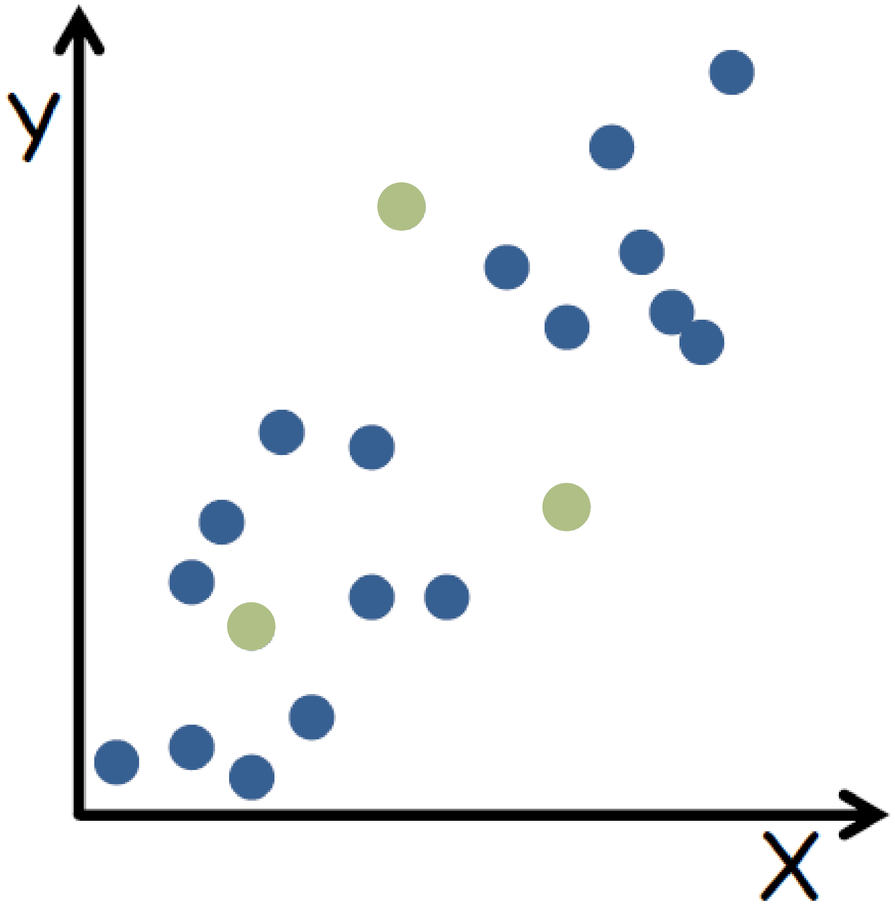


CHANNEL :

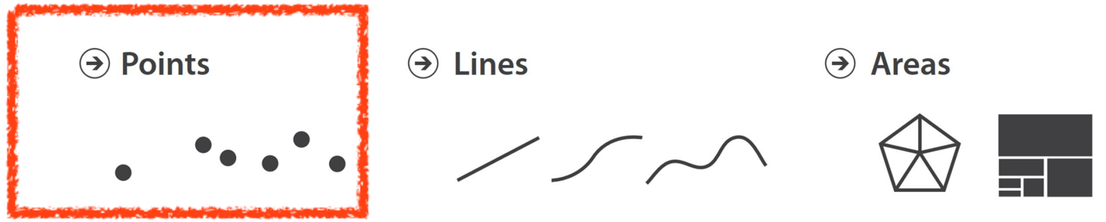


Visualization Building Blocks

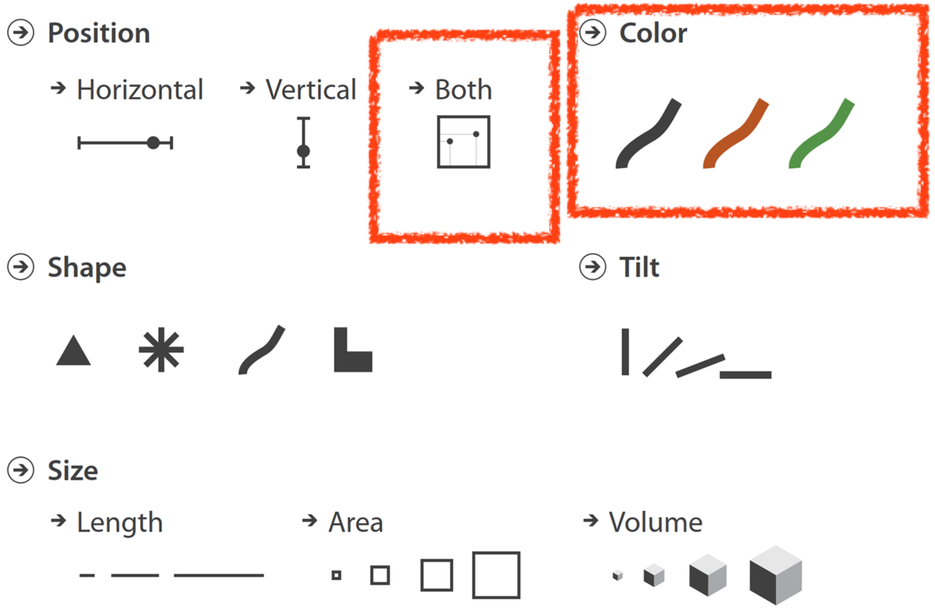
of attributes encoded: 3



MARK:

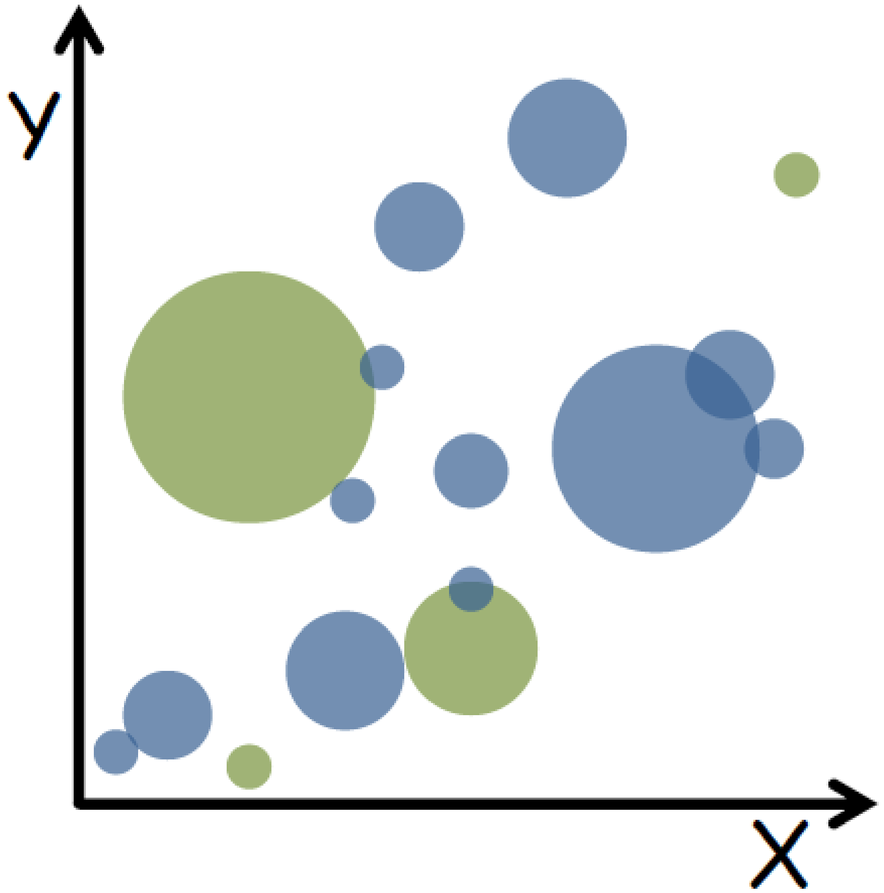


CHANNEL :

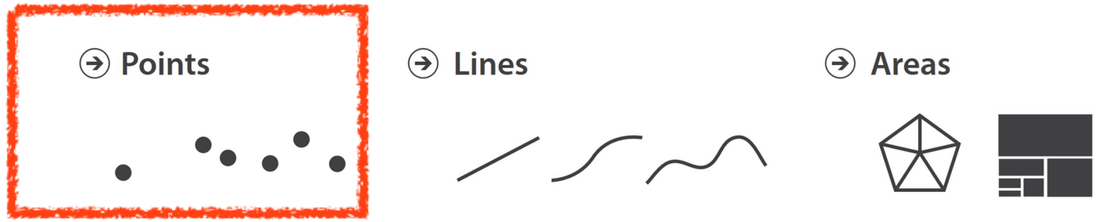


Visualization Building Blocks

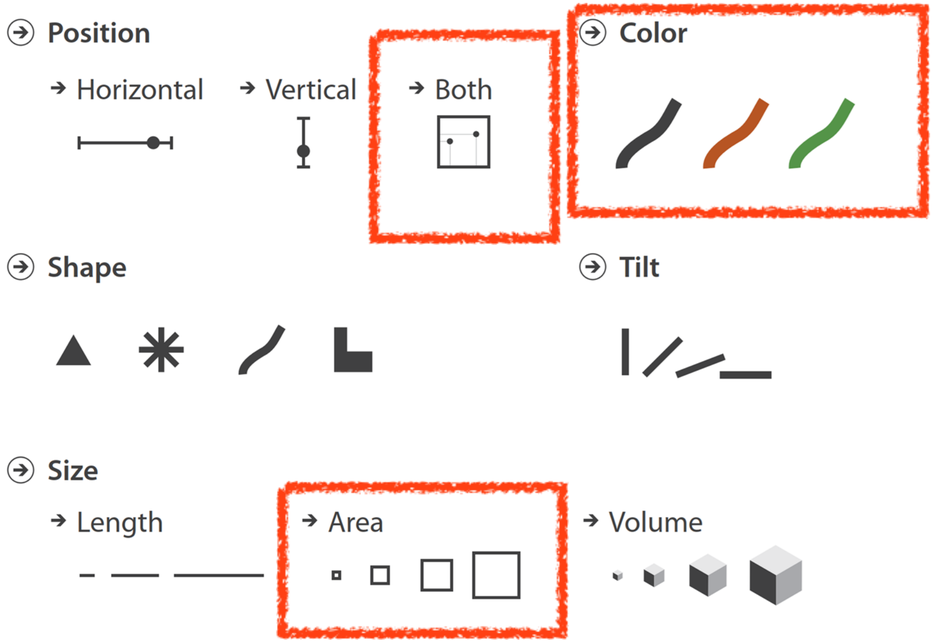
of attributes encoded: 4



MARK:

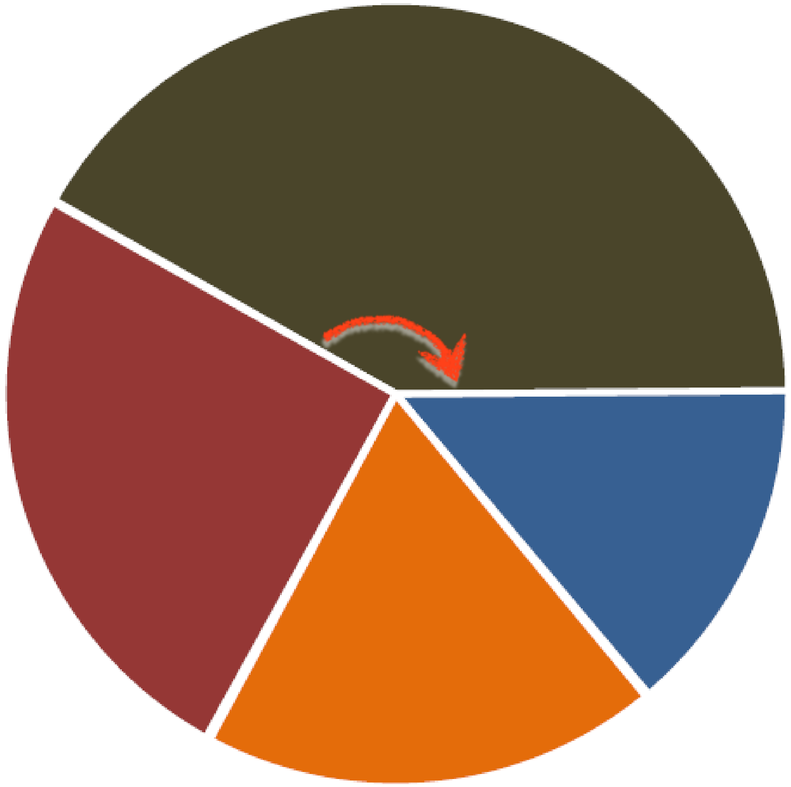


CHANNEL :

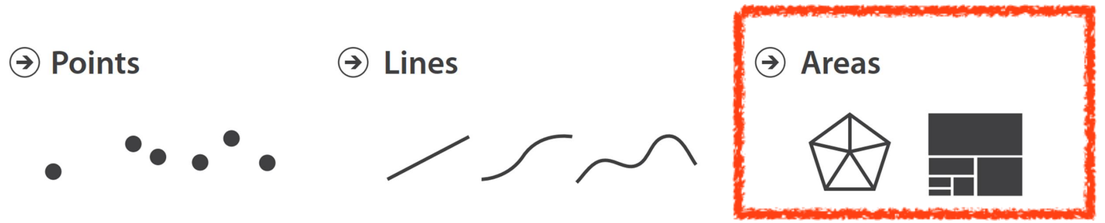


Visualization Building Blocks

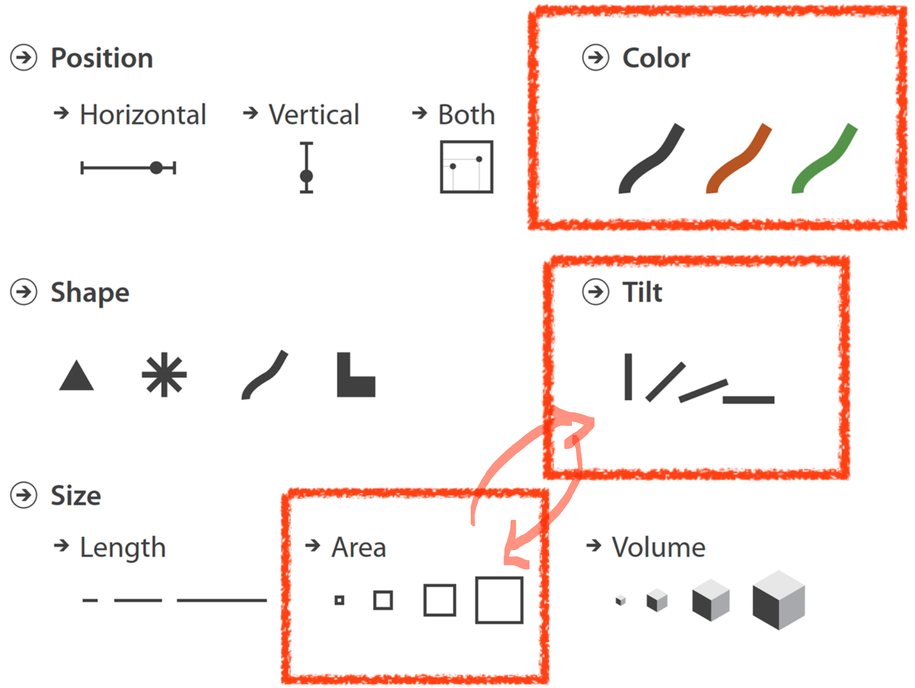
of attributes encoded: 2



MARK:



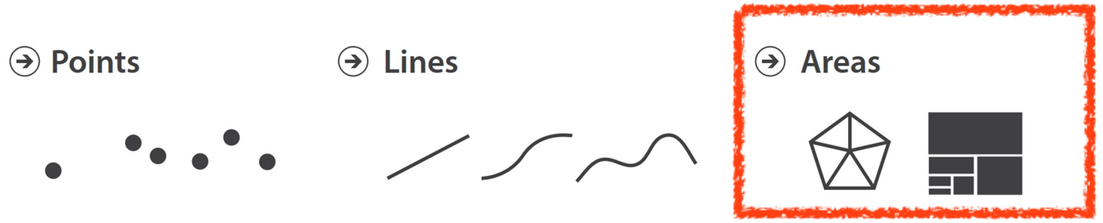
CHANNEL :



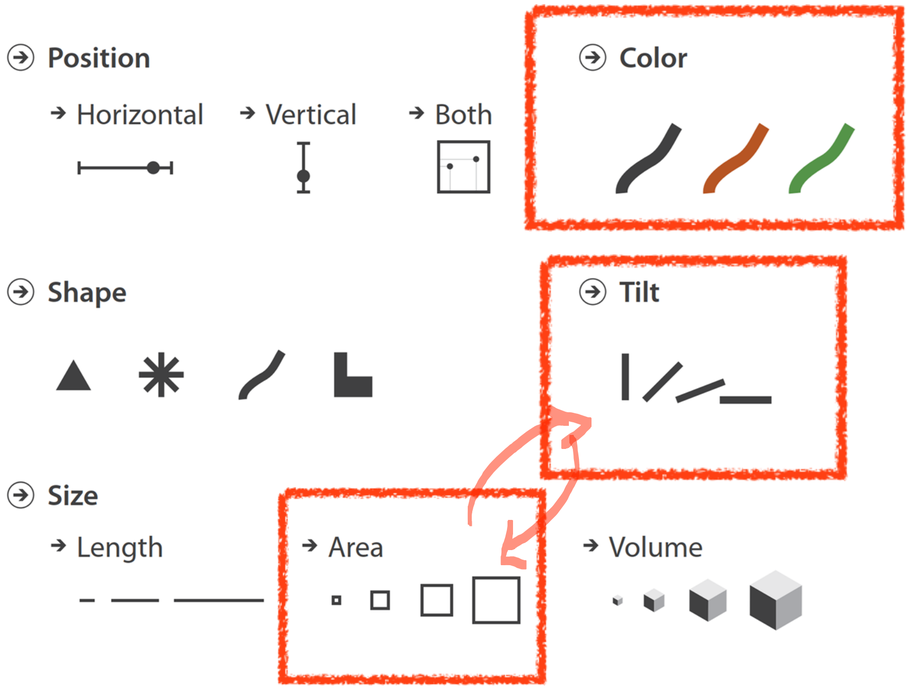
Visualization Building Blocks

of attributes encoded: 2

MARK:

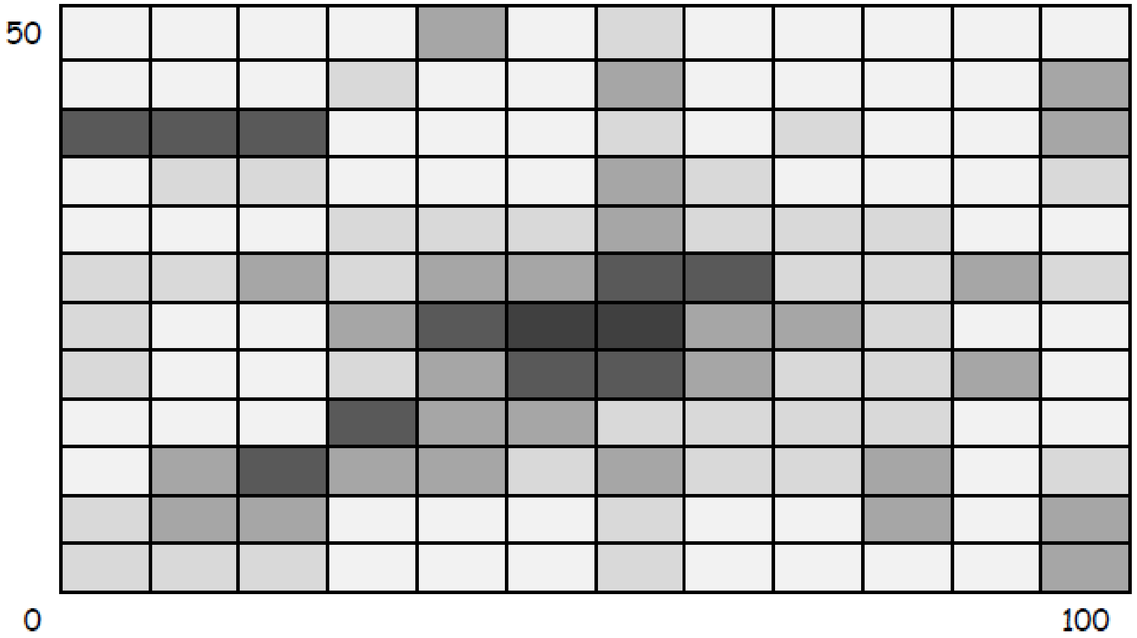


CHANNEL :

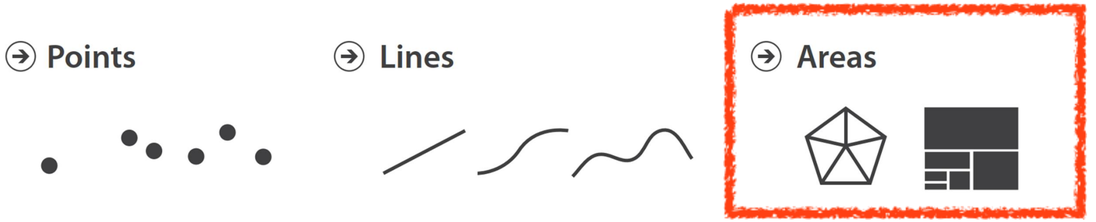


Visualization Building Blocks

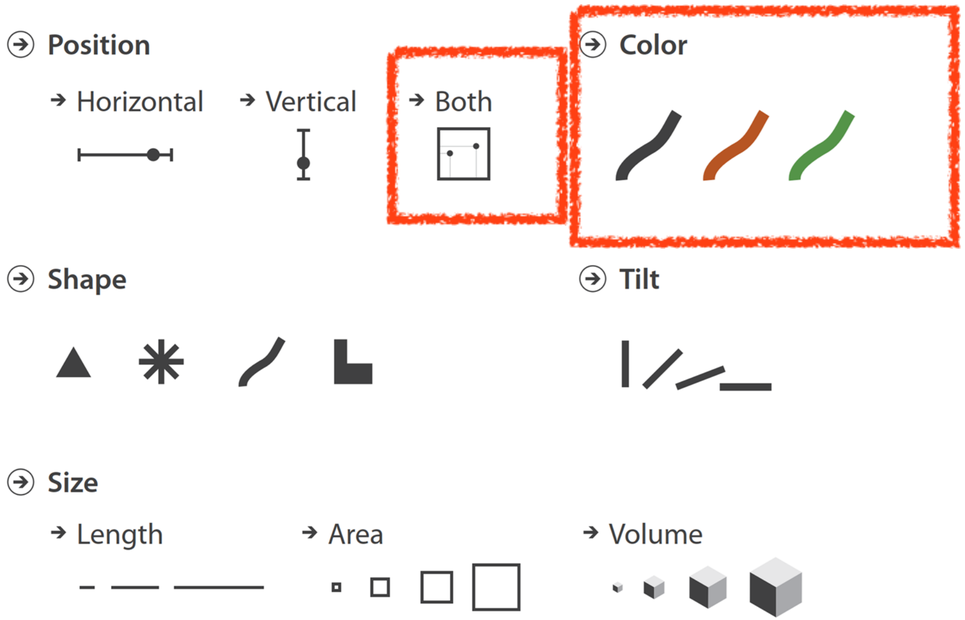
of attributes encoded:



MARK:



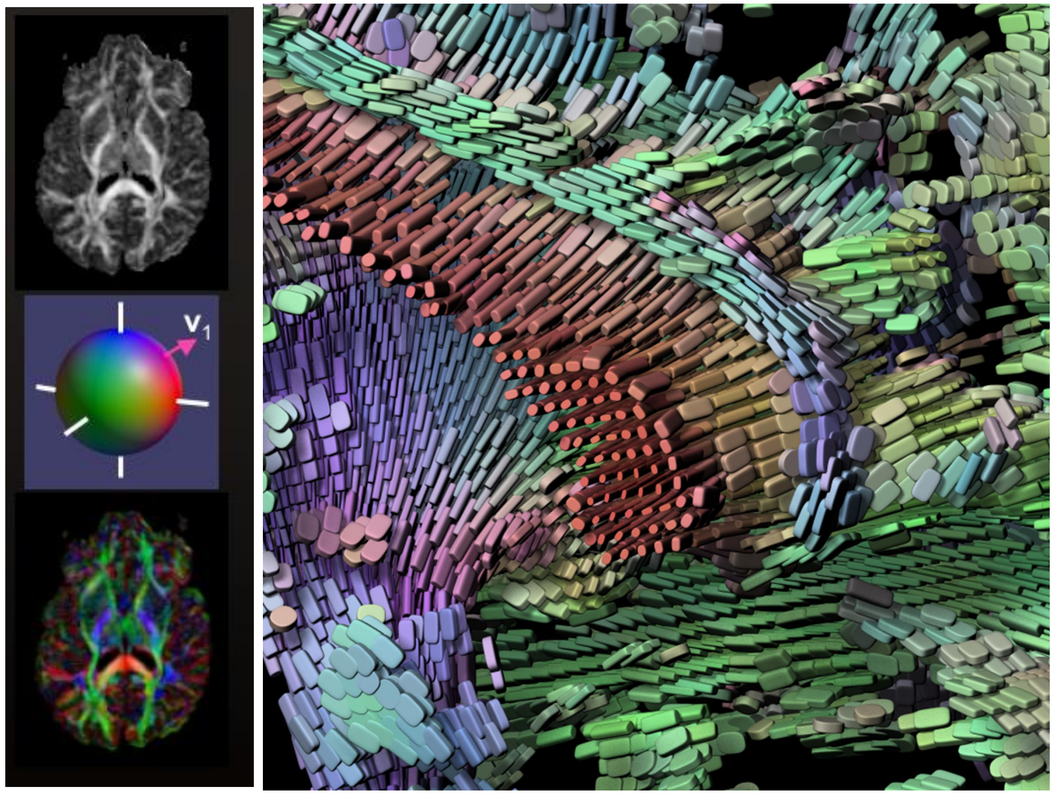
CHANNEL :



Visualization Building Blo

Don't overload the user with encodings!

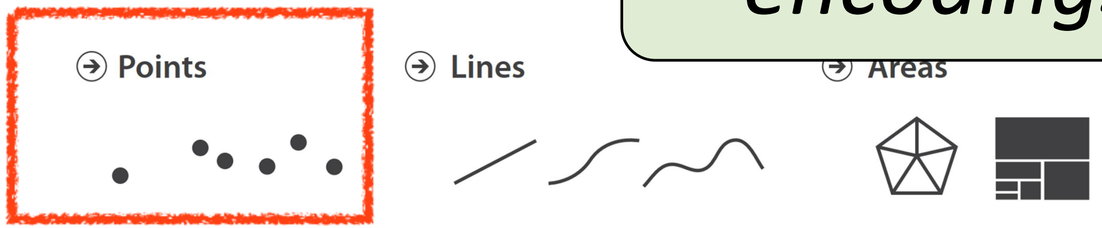
# of attributes encoded: ?



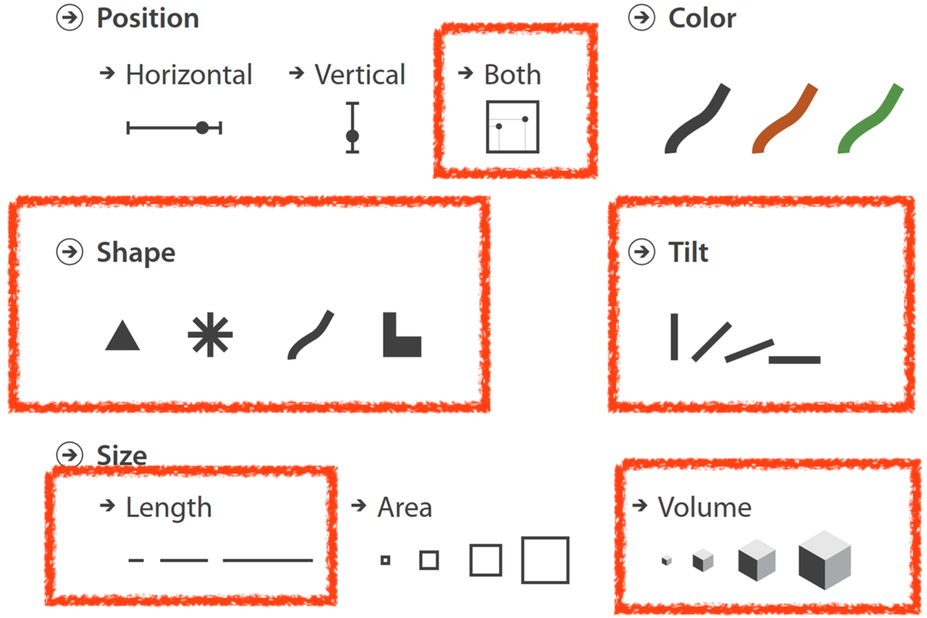
Kindlmann (2004)

+ position in 3D space

MARK:



CHANNEL :



Visualization Building Blocks

Marks as Items/Nodes

- Points

- Lines

- Areas

Marks as Links

- Containment

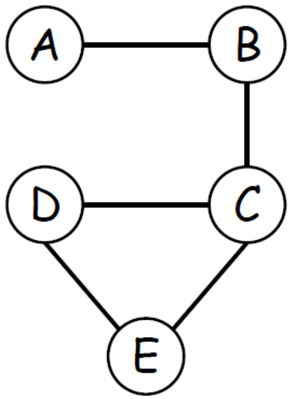
- Connection

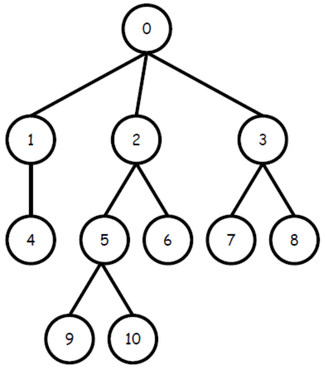
Visualization Building Blocks

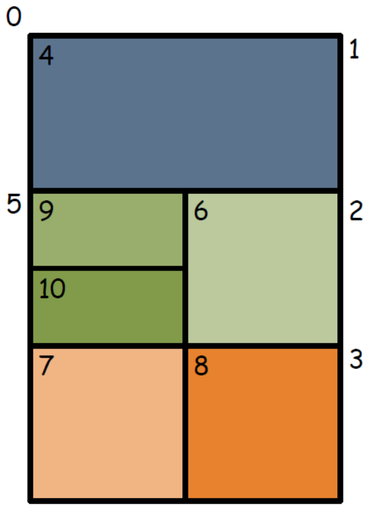
Marks as Links

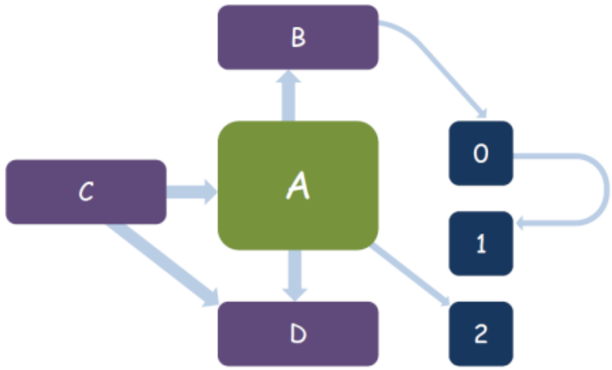
- Containment

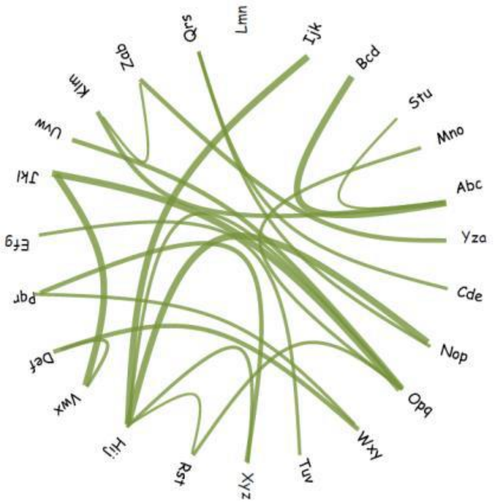
- Connection











Visualization Building Blocks

Marks as Items/Nodes

- Points
- Lines
- Areas

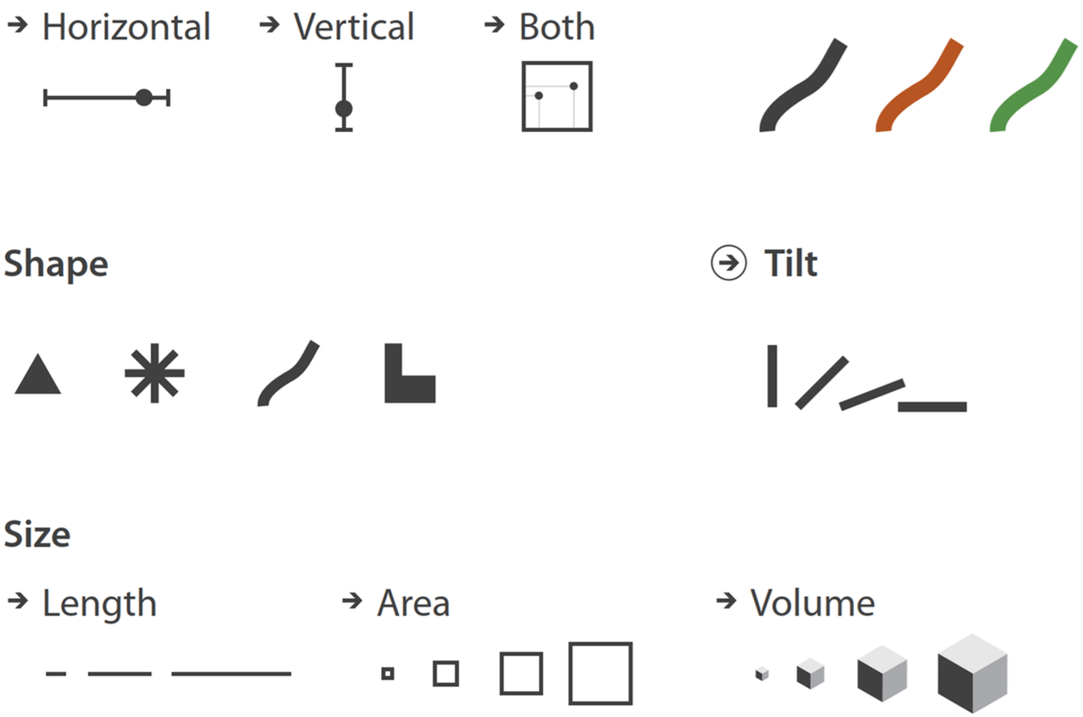
Marks as Links

- Containment
- Connection

Note: these are all really important concepts when it comes time to coding your visualizations...!

Channels :

- Position
  - Horizontal
  - Vertical
  - Both
- Color
- Shape
- Tilt
- Size
  - Length
  - Area
  - Volume



How do I pick *which* marks or channels to use?

“Ordering of Elemental Perceptual Tasks”

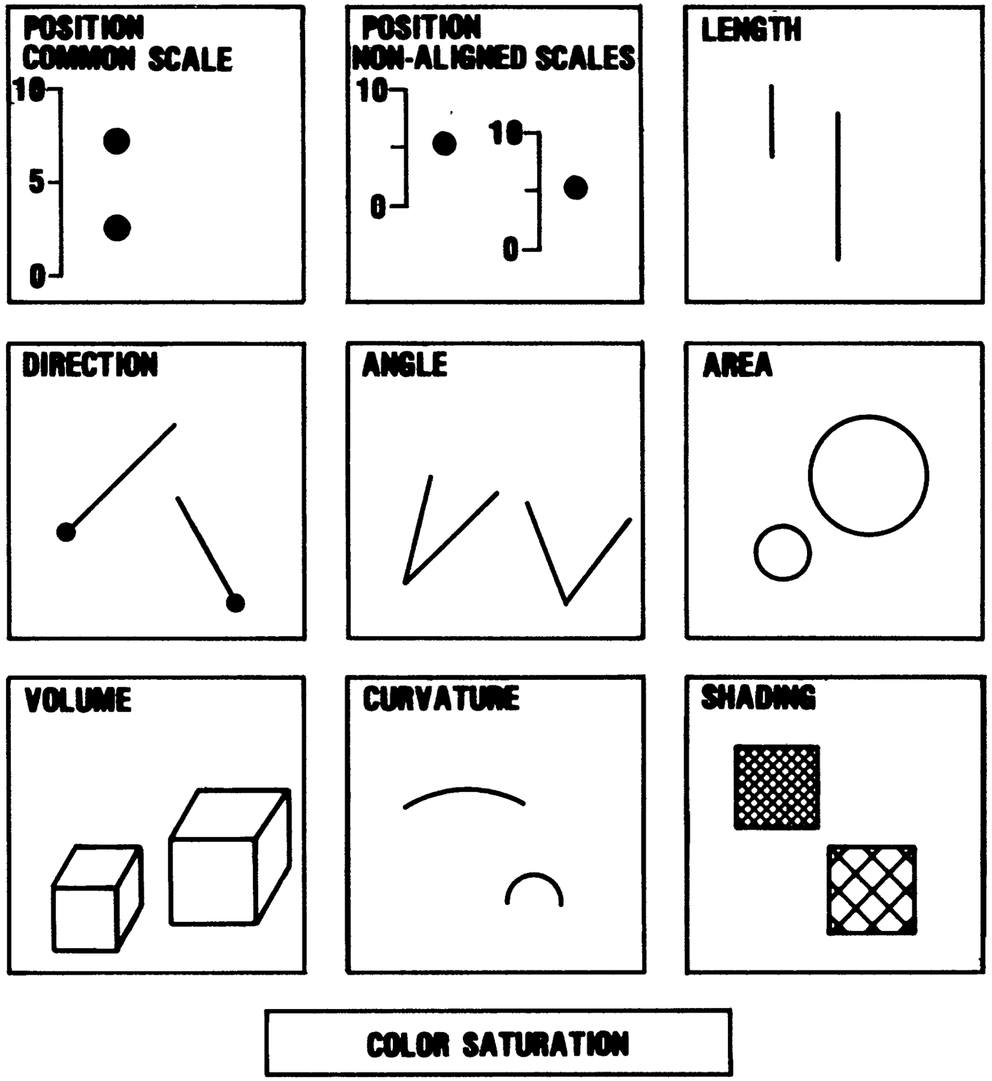


Figure 1. Elementary perceptual tasks.

[Cleveland & McGill (1984)]

“Ordering of Elemental Perceptual Tasks”

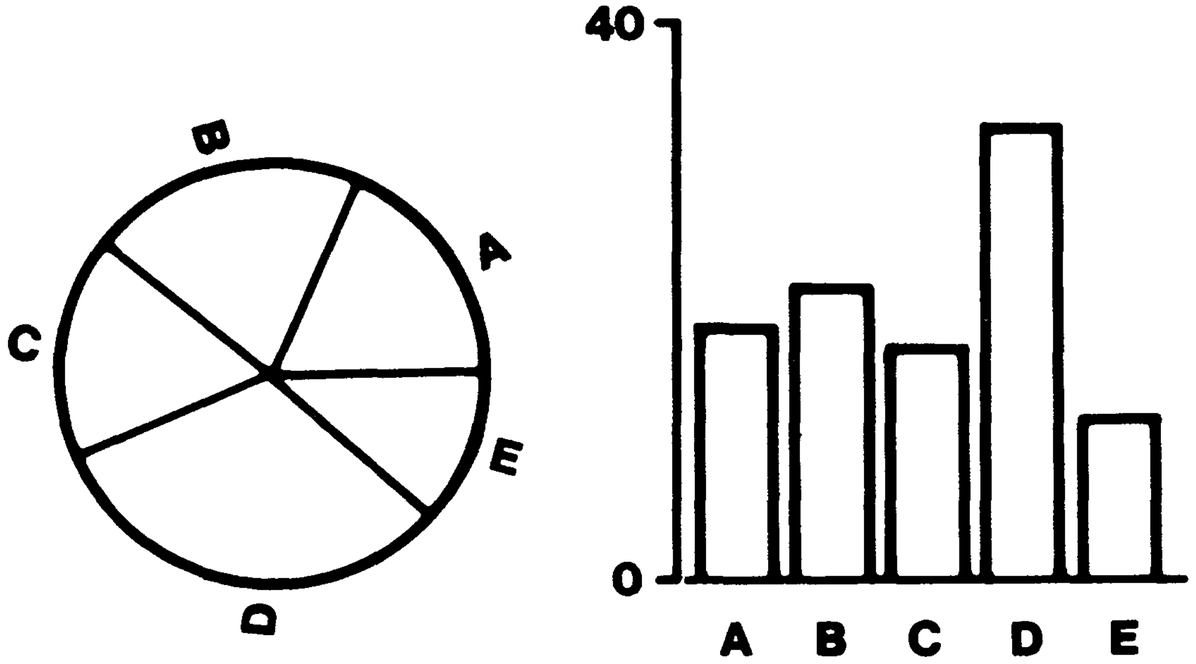


Figure 3. Graphs from position–angle experiment.

TASK: Which segment/bar is the maximum, and what is its percentage/value?

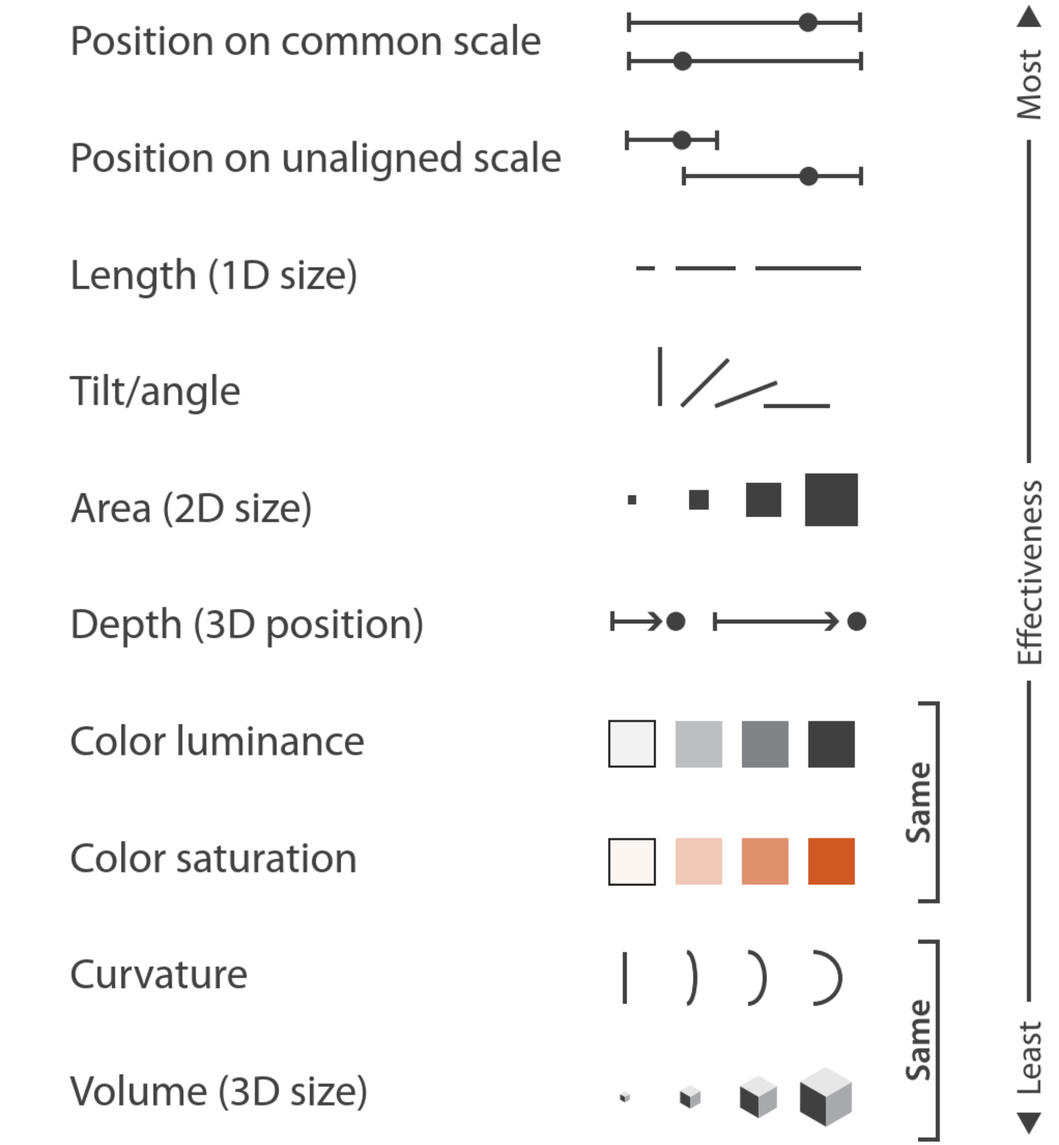
Cleveland & McGill (1984),
larger replication on AMT by Heer & Bostock (2010)

Channels: Expressiveness Types and Effectiveness Ranks

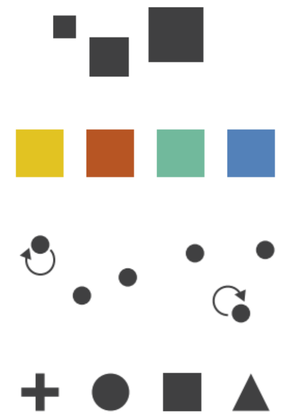
- **Magnitude Channels: Ordered Attributes**

Channels: Expressiveness Types and Effectiveness Ranks

Magnitude Channels: Ordered Attributes



Identity Channels: Categorical Attributes



Summarizes results from [Cleveland & McGill (1984)](), [Heer & Bostock (2010)]()

Expressiveness and Effectiveness

Effectiveness principle: the importance of the attribute should match the salience of the channel; that is, its noticeability.
(i.e., encode most important attributes with highest ranked channels)

Expressiveness principle: the visual encoding should express all of, and only, the information in the dataset attributes.
(i.e., data characteristics should match the channel)

Mackinlay (1986)

My Summary: Prioritize choosing the most appropriate channel for each attribute

Expressiveness and Effectiveness

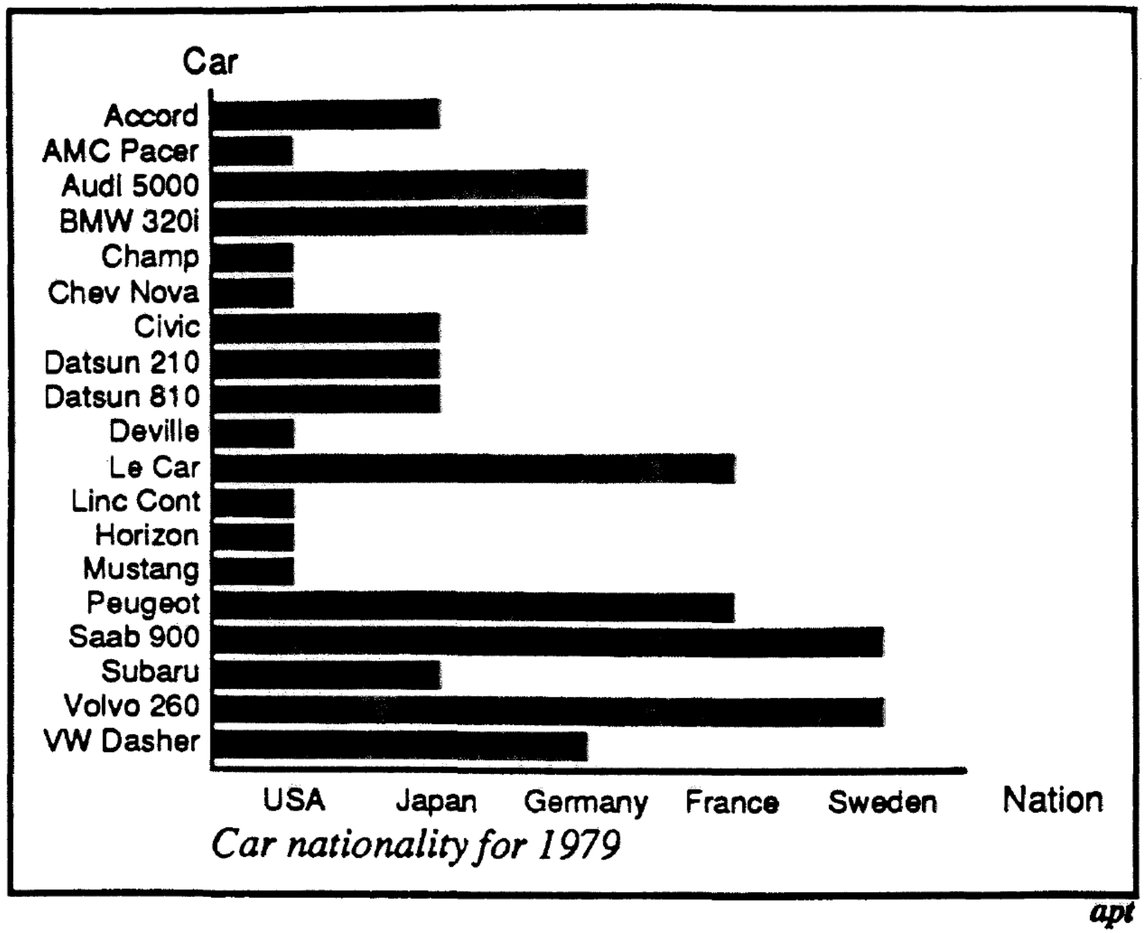


Figure 11: Incorrect Use of a Bar Chart for the Nation Relation. *The lengths of the bars suggest an ordering on the vertical axis, as if the* USA *cars were longer or better than the other cars, which is not true for the Nation relation.*

Mackinlay (1986)

Expressiveness and Effectiveness

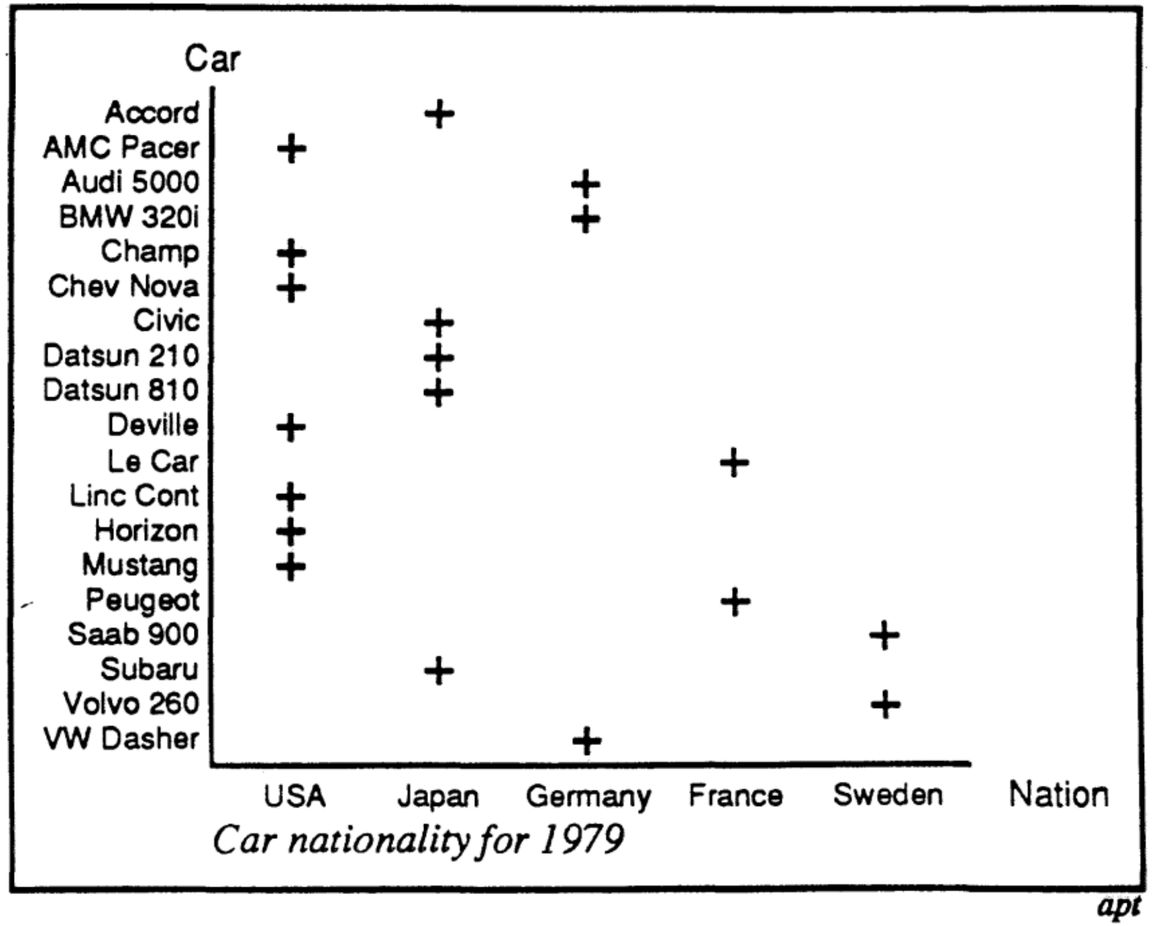
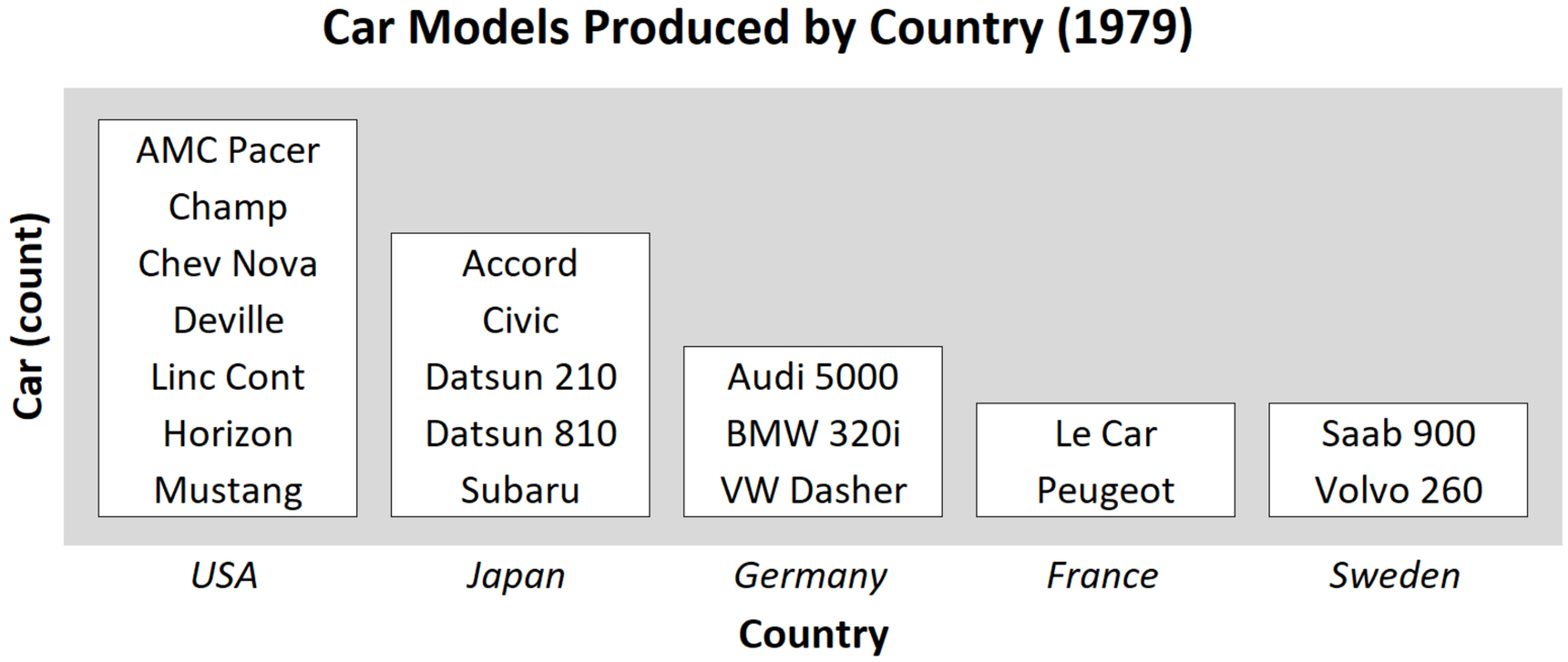


Figure 12: Correct Use of a Plot Chart for the Nation Relation. *Since bar charts encode ordered domain sets, plot charts are conventionally used to encode nominal domain sets. The ordering of the labels on the axes is ignored.*

Mackinlay (1986)

Expressiveness and Effectiveness



IN-CLASS EXERCISE

3, 12, 42

3, 12, 42

In-class Sketching: “Three numbers”

20m

1. Break-out into groups of ~3 students.
2. **Together** *(15m)* use paper and pens/pencils to sketch as many possible visualizations as you can of these three numbers.
3. No upload required
4. **As a class** *(5m)* some groups will report on key designs and themes.

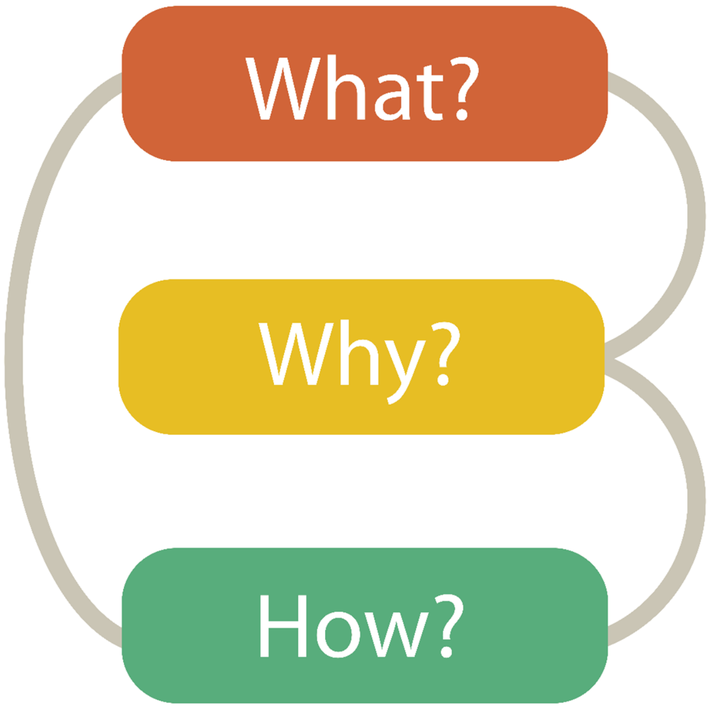
Jonathan Schwabish

Data Types

Goals for Today

- Learn what are data types and dataset types
- Learn what are attribute types
- Learn how to pick appropriate visual representations based on attribute type and perceptual properties

Analysis

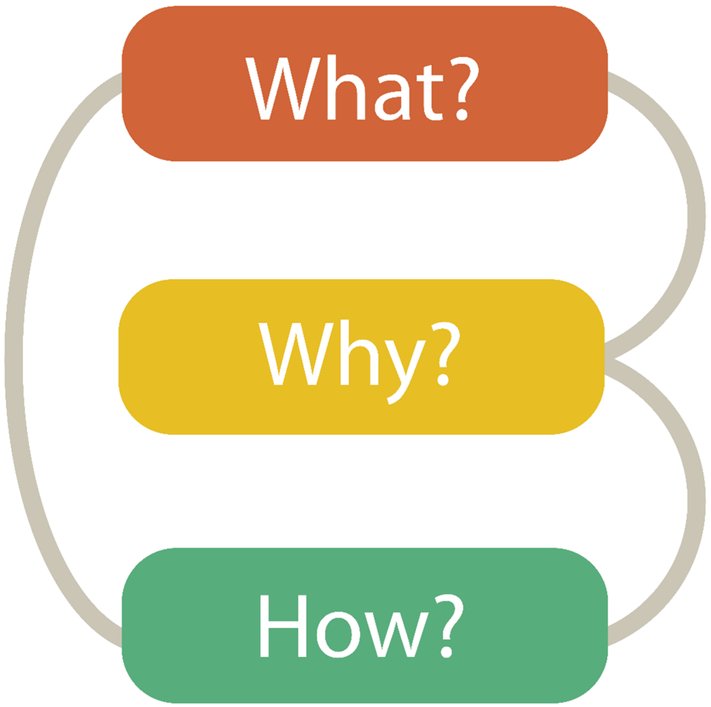


What data is shown?

Why is the user analyzing / viewing it?

How is the data presented?

Analysis

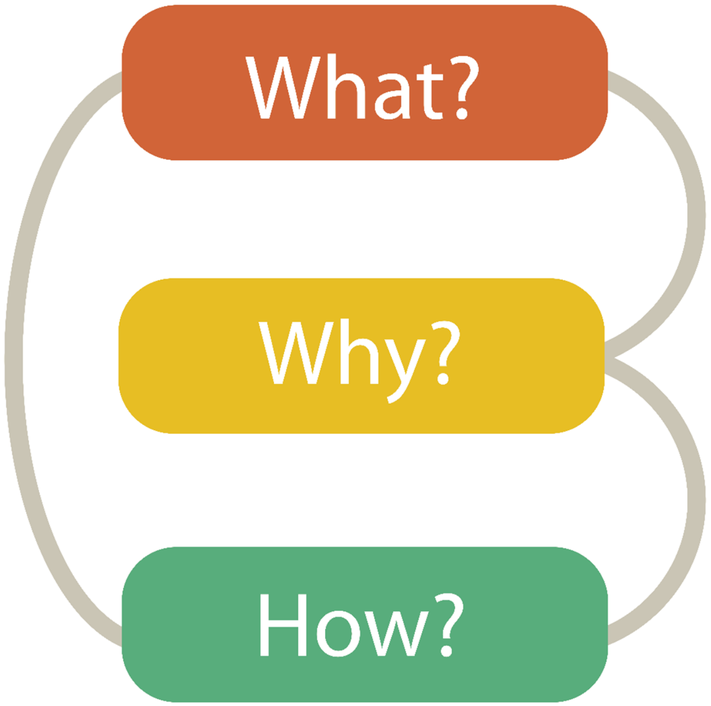


DATA ABSTRACTION

TASK ABSTRACTION

VISUAL ENCODING

Analysis



DATA ABSTRACTION

TASK ABSTRACTION

VISUAL ENCODING

Data Types

TYPE = structural or mathematical interpretation of the data

Data Types

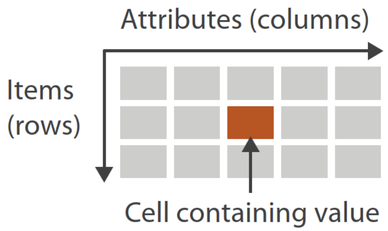
➔ Items	➔ Attributes	➔ Links	➔ Positions	➔ Grids
(row, node)	*(variable, data dimension)*	*(relationship)*	*(spatial location)*	*(sampling)*

Data Types

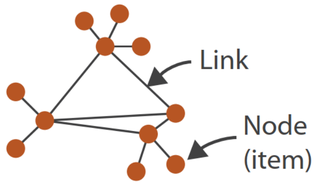
DATASET = collection of information that is the target of analysis

Dataset Types

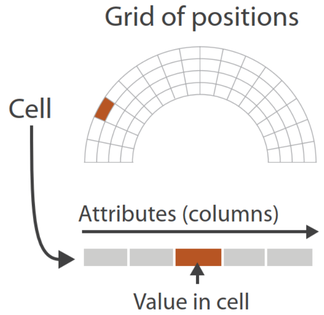
Tables



Networks

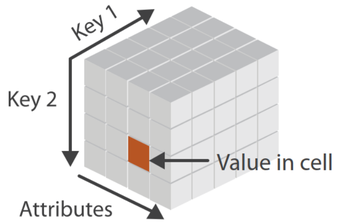


Fields (Continuous)



Geometry (Spatial)

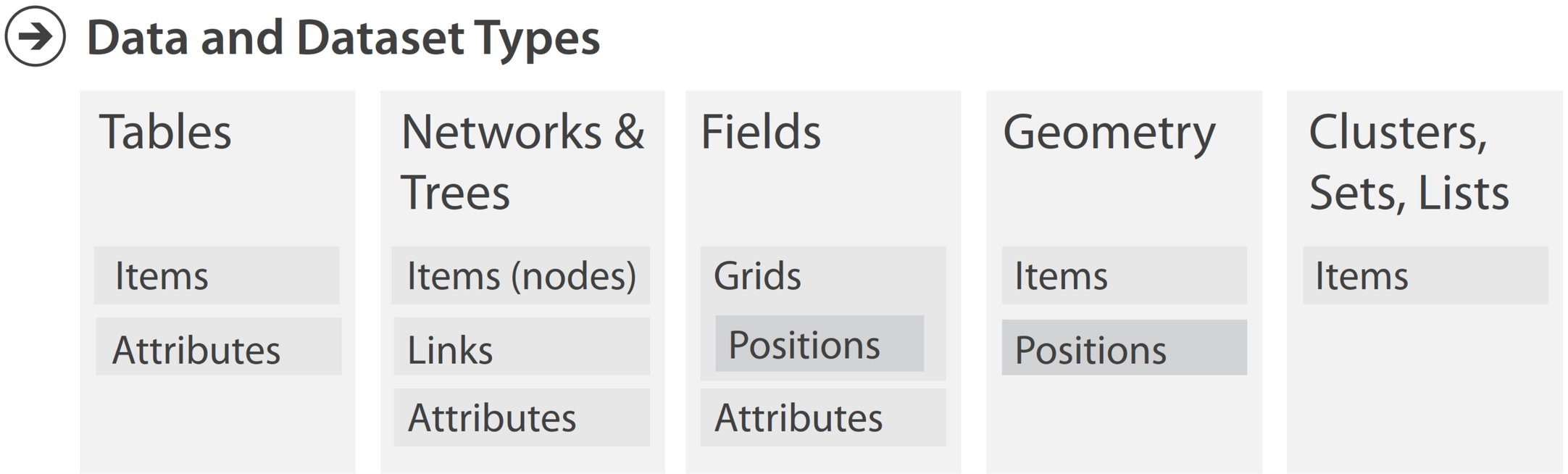
Multidimensional Table



Trees

Data Types

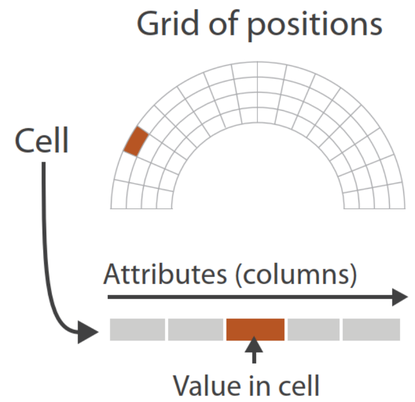
DATASET = collection of information that is the target of analysis



grid types

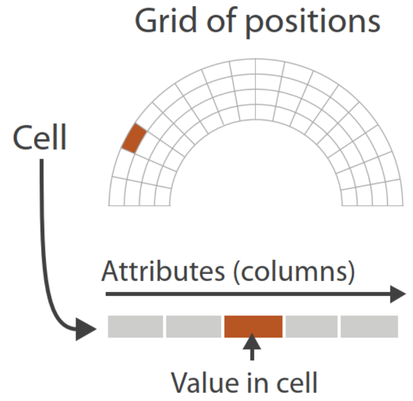
Relevant to anyone in the sciences!

➔ Fields (Continuous)



grid choices impact how continuous data is interpreted

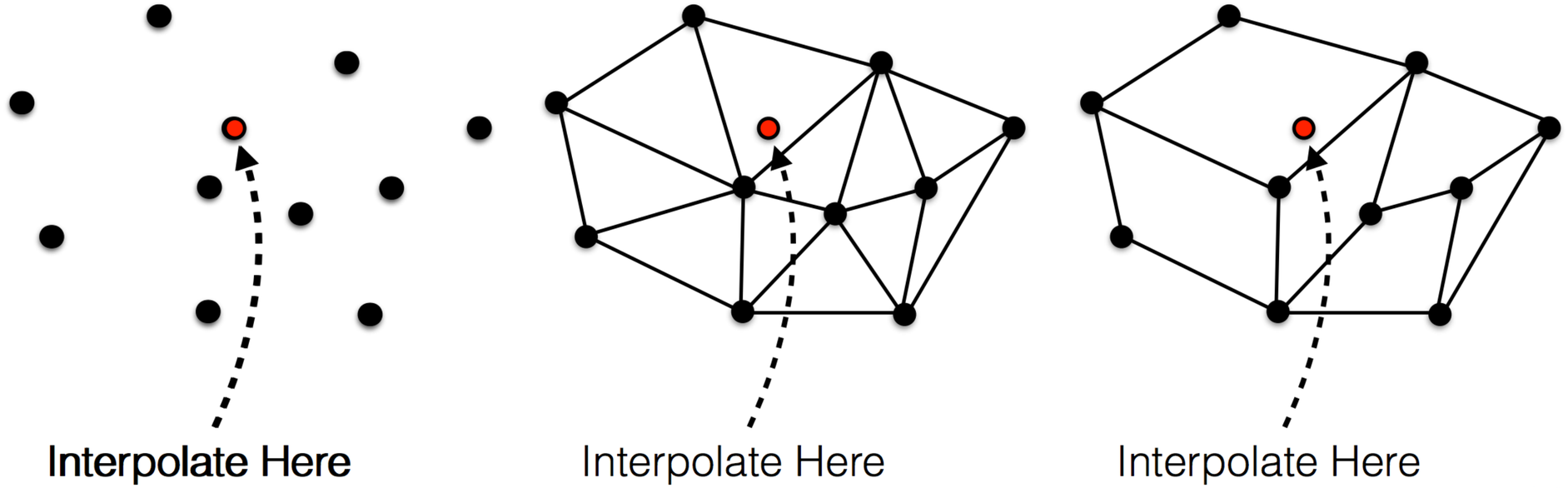
➔ Fields (Continuous)



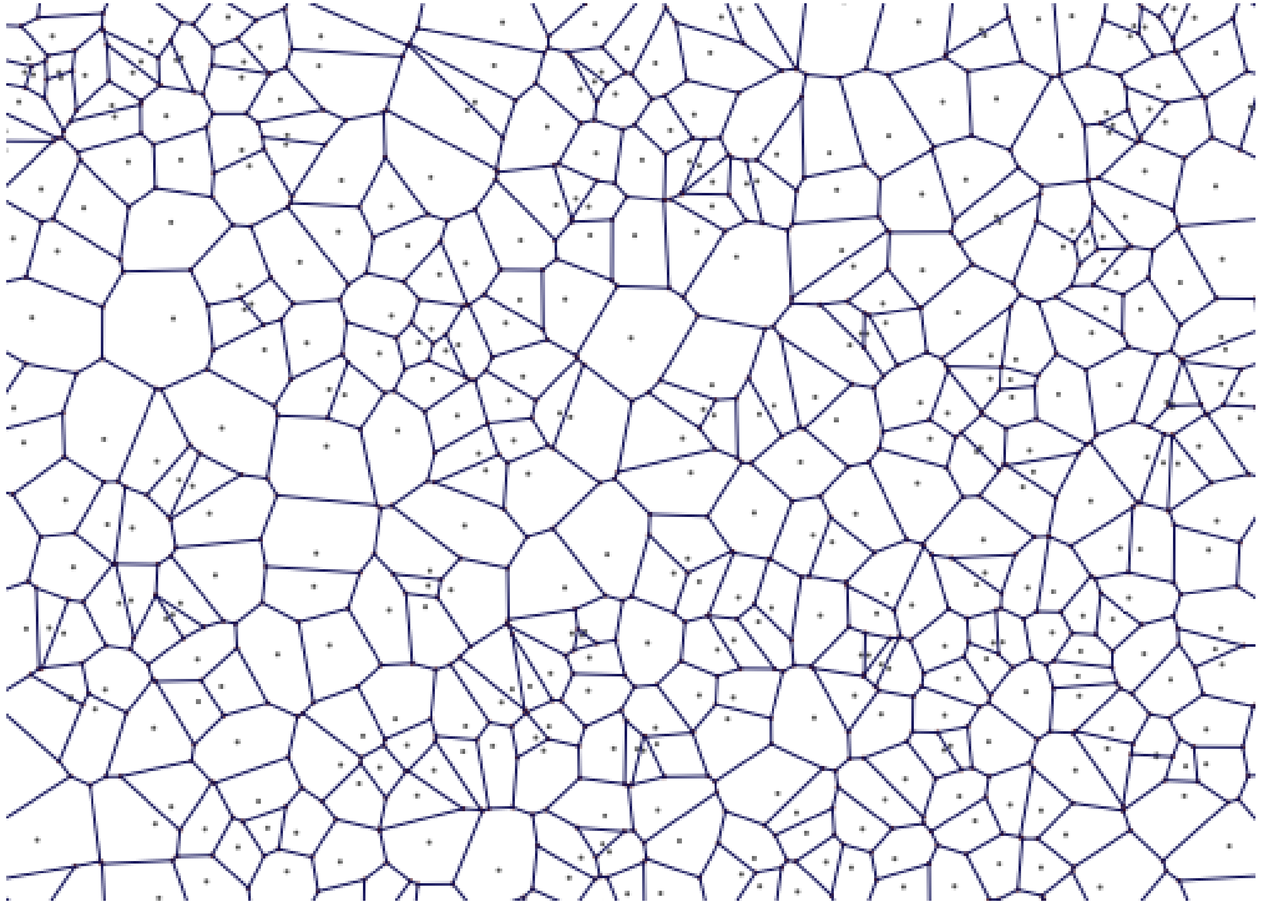
two key considerations:

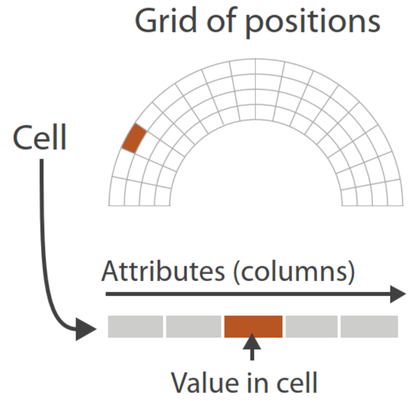
sampling, or the choice of where attributes are measured

interpolation, or how to model the attributes in the rest of space



“Voronoi Tessellation”

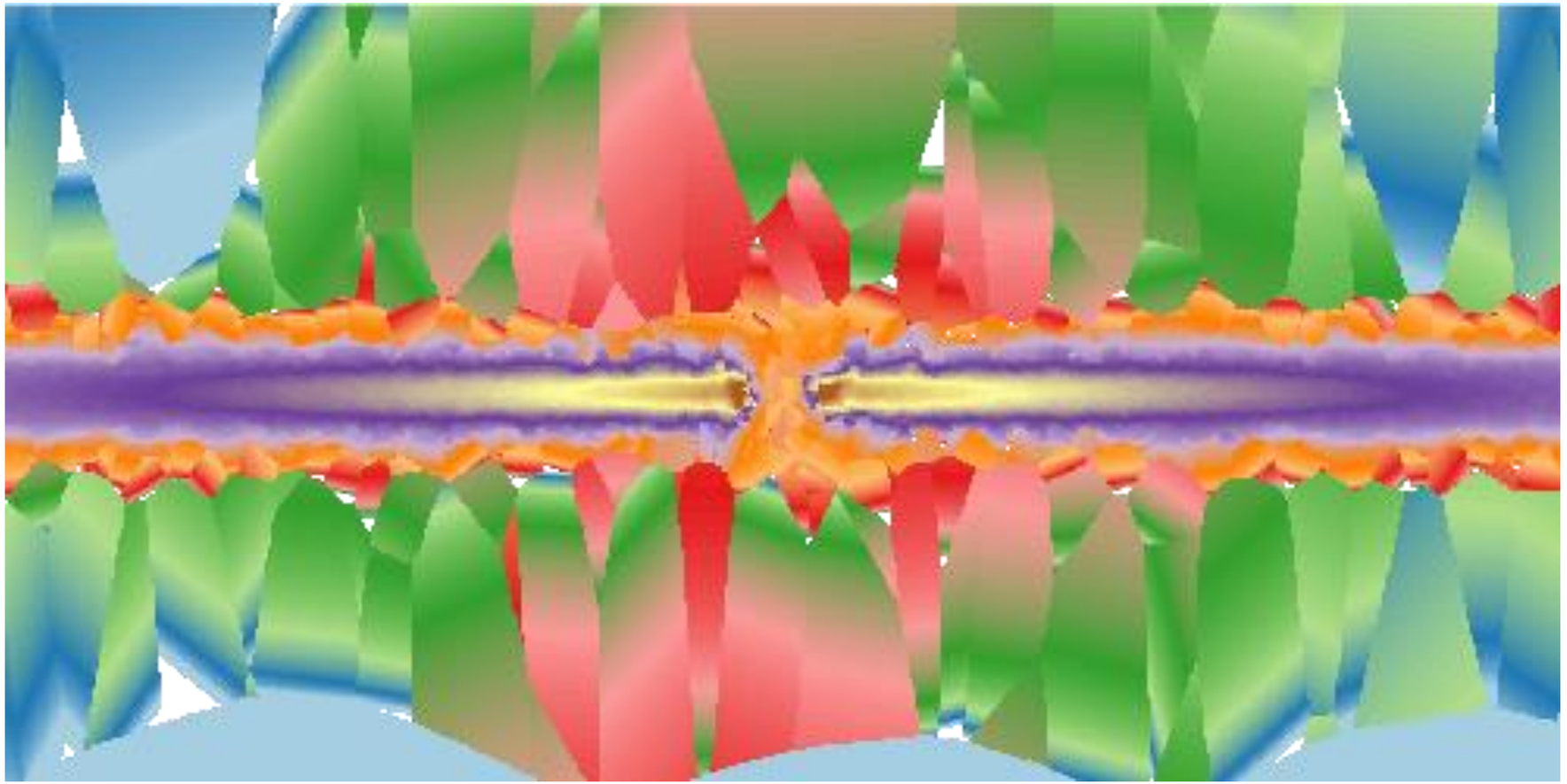




https://en.wikipedia.org/wiki/Voronoi_diagram

➔ Fields (Continuous)

Voronoi Tessellation for Galaxy Evolution Simulation



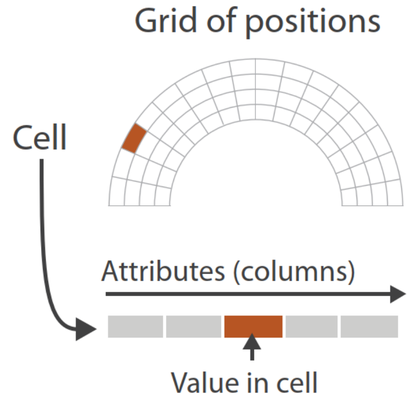
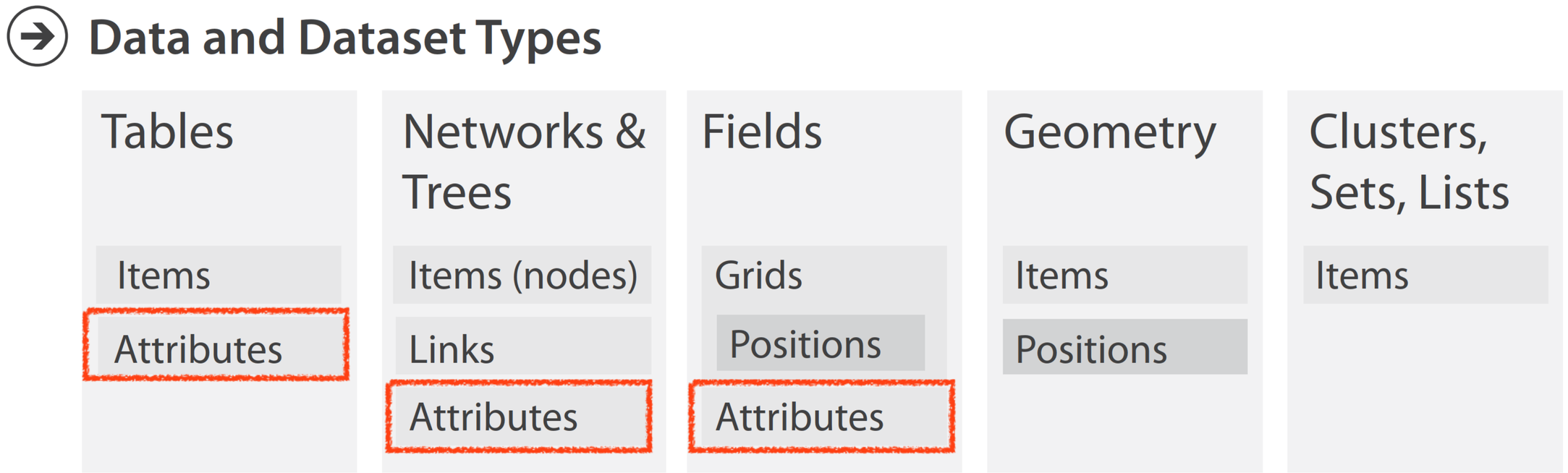


Image courtesy of Patrik Jonsson

Data Types

DATASET = collection of information that is the target of analysis



Attribute Types

➔ Categorical

e.g.,
fruit (apple, pear, grape),
colleges (CAMD, Khoury, COE)

➔ Ordered

➔ *Ordinal*

e.g.,
sizes (xs, s, m, l, xl),
months (J, F, M)

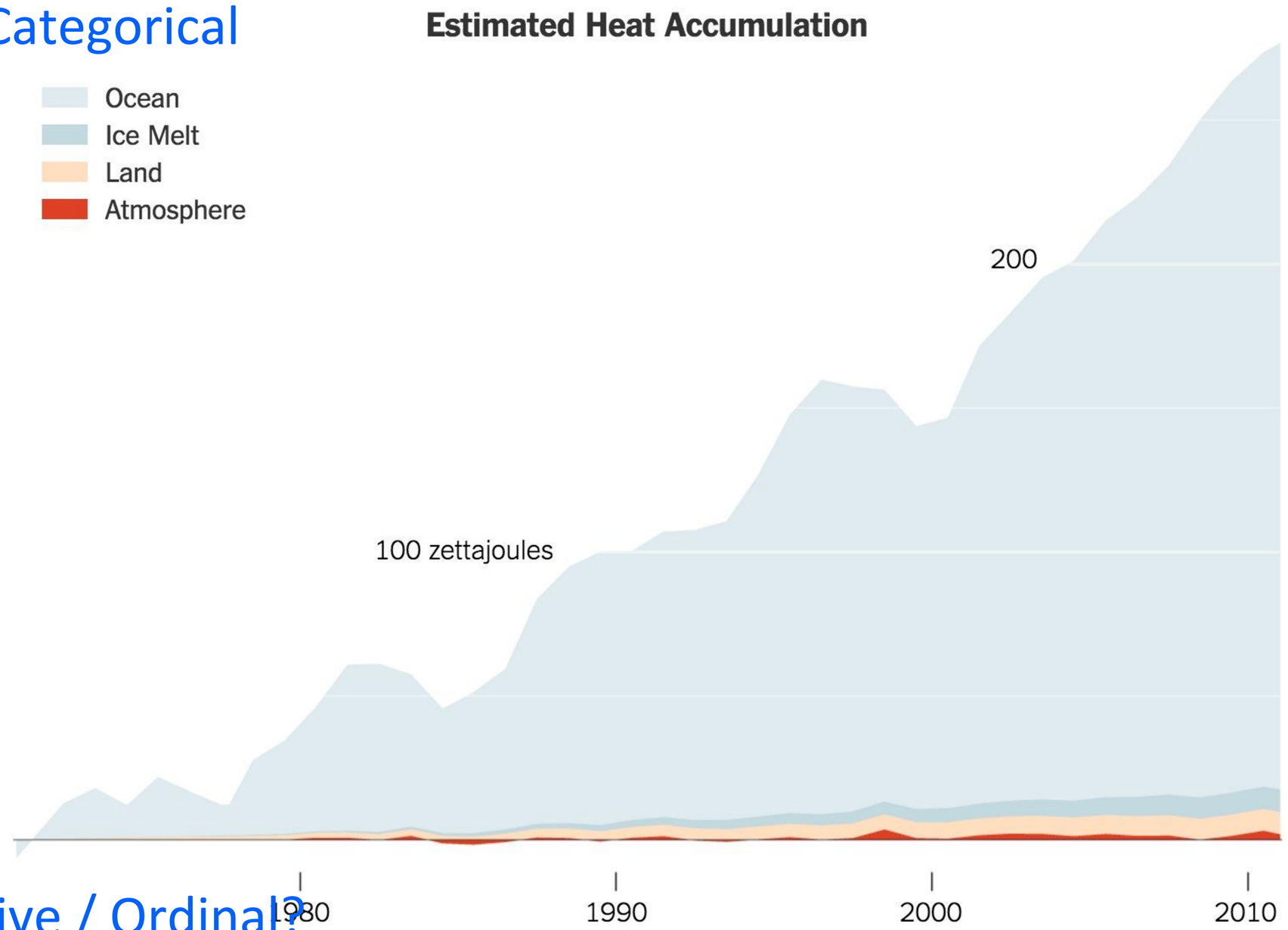
➔ *Quantitative* *(continuous)*

e.g.,
lengths (1’, 2.5’, 5’),
population

Categorical

Quantitative

?Quantitative / Ordinal?



http://www.nytimes.com/interactive/2016/09/12/science/earth/ocean-warming-climate-change.html

Upcoming Assignments & Communication

https://northeastern.instructure.com/courses/63405/assignments/syllabus

Look at the upcoming assignments and deadlines regularly!

- Textbook, Readings, & Reading Quizzes — Variable days
- In-Class Activities — 11:59pm same day as class
  - F: Lecture & in-class activity on D3 (1/2)
  - Next T: Lecture & in-class activity on D3 (2/2)
- Assignments & Projects— Generally due R 11:59pm
  - This R (2 days): Assignments 3a, 3b due
  - Next R (9 days): Project 1 (pitches) due
  - Next-next R (16 days): Project 2 (proposals) due
- Project Overview

Everyday Required Supplies:

- 5+ colors of pen/pencil
- White paper
- Laptop and charger

Use Canvas Discussions for general questions, email the TAs/S-LTA/instructor for questions specific to you: codydunne-and-tas@ccs.neu.edu. Include links!