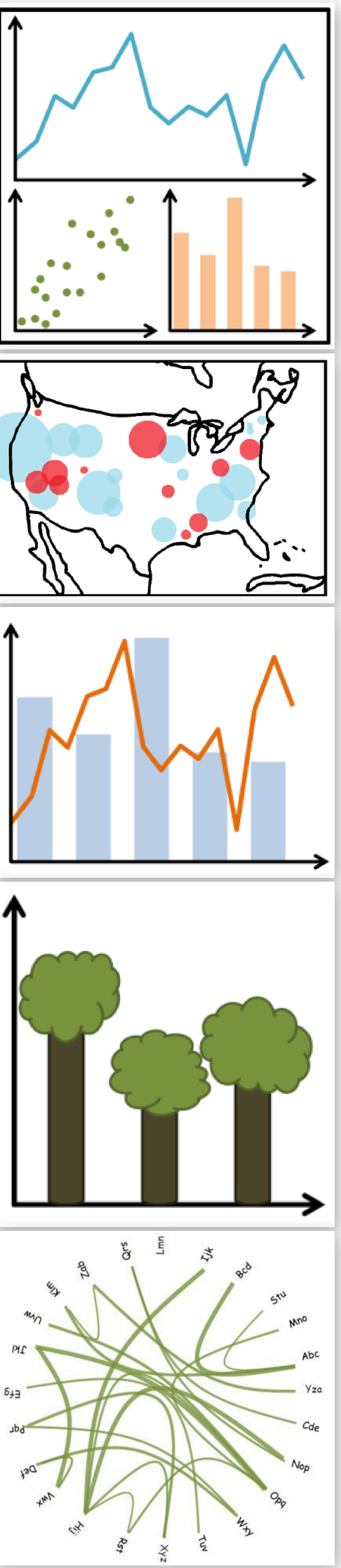


Data Types, Tasks, Visual Encodings

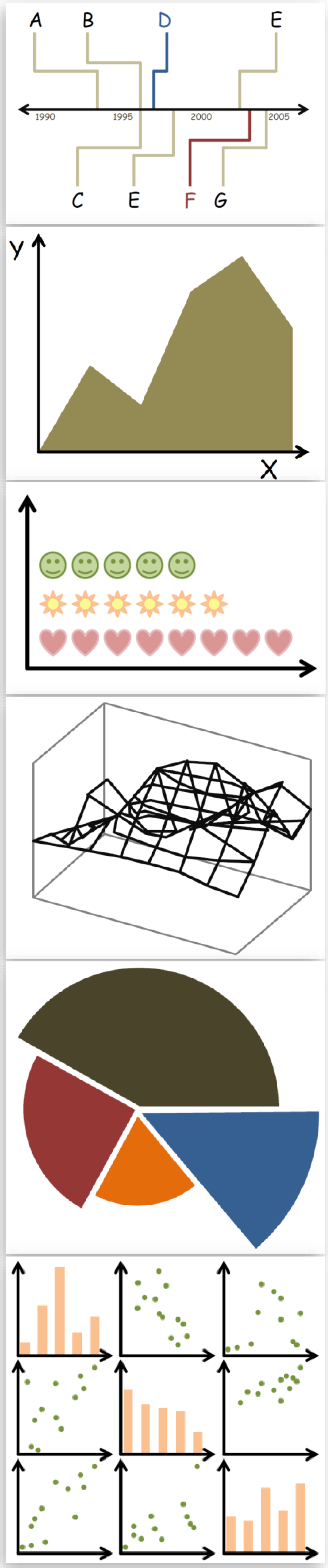
CS 7250

Spring 2020

Prof. Cody Dunne

Northeastern University

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



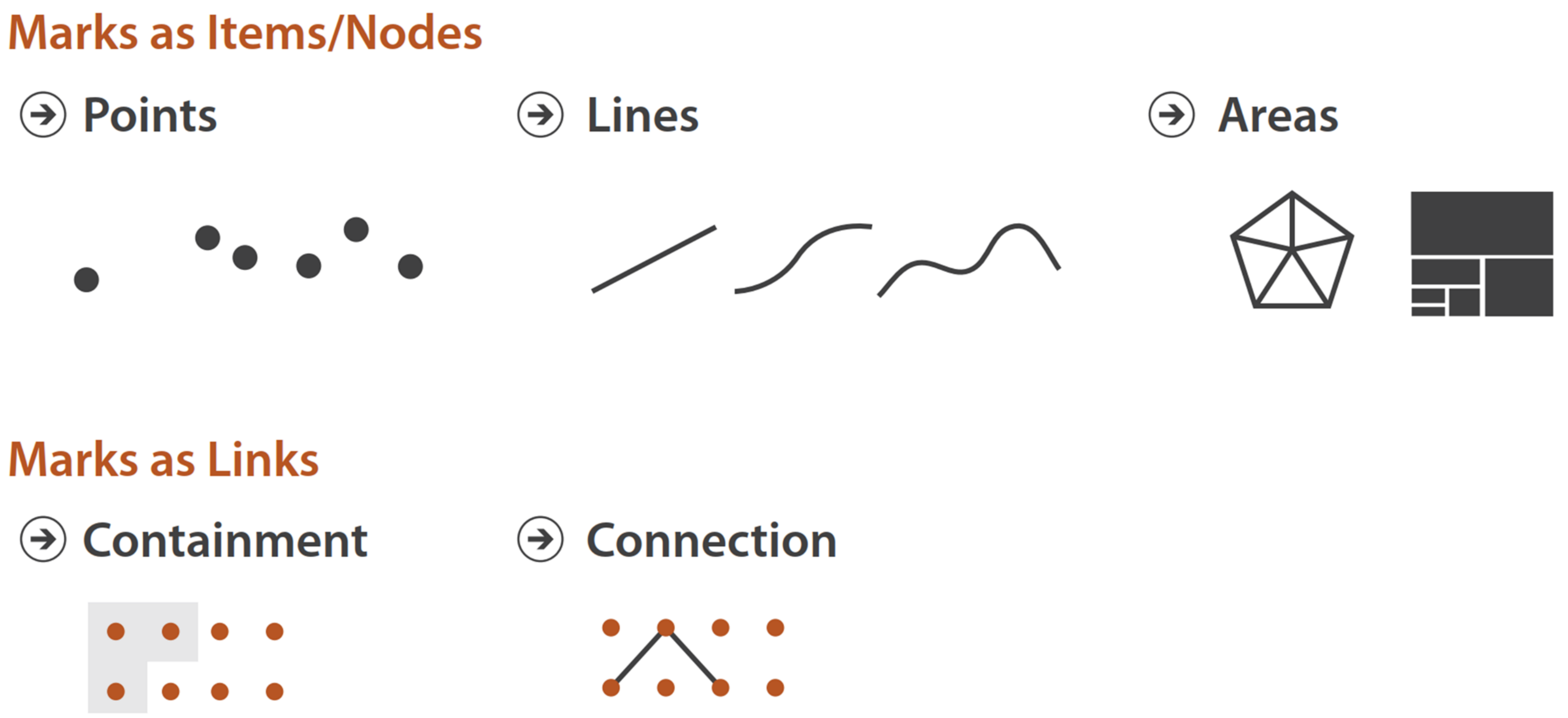
Reading Quiz

8 min

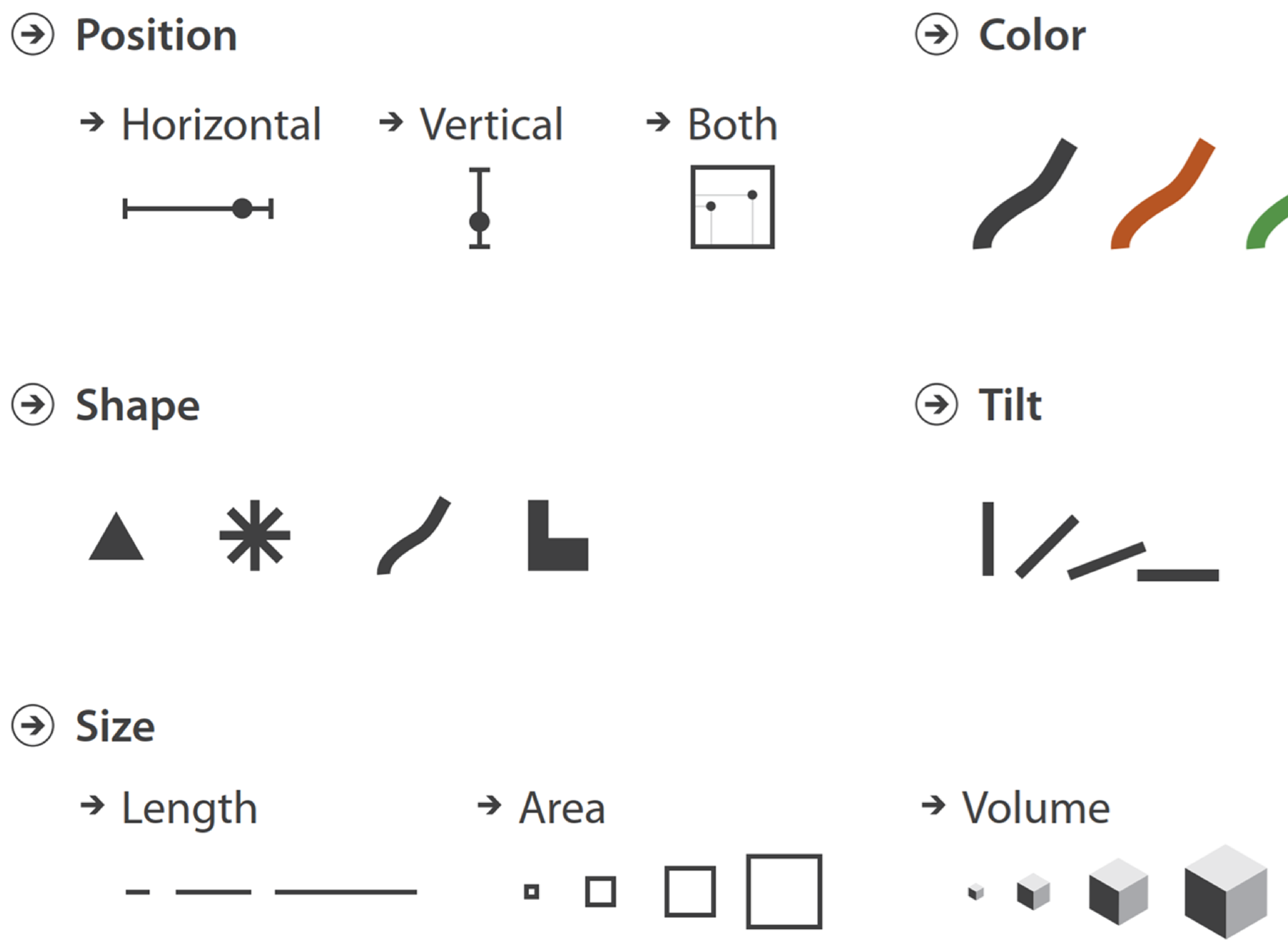
QUESTIONS?

Previously, on CS 7250…

Visualization Building Blocks

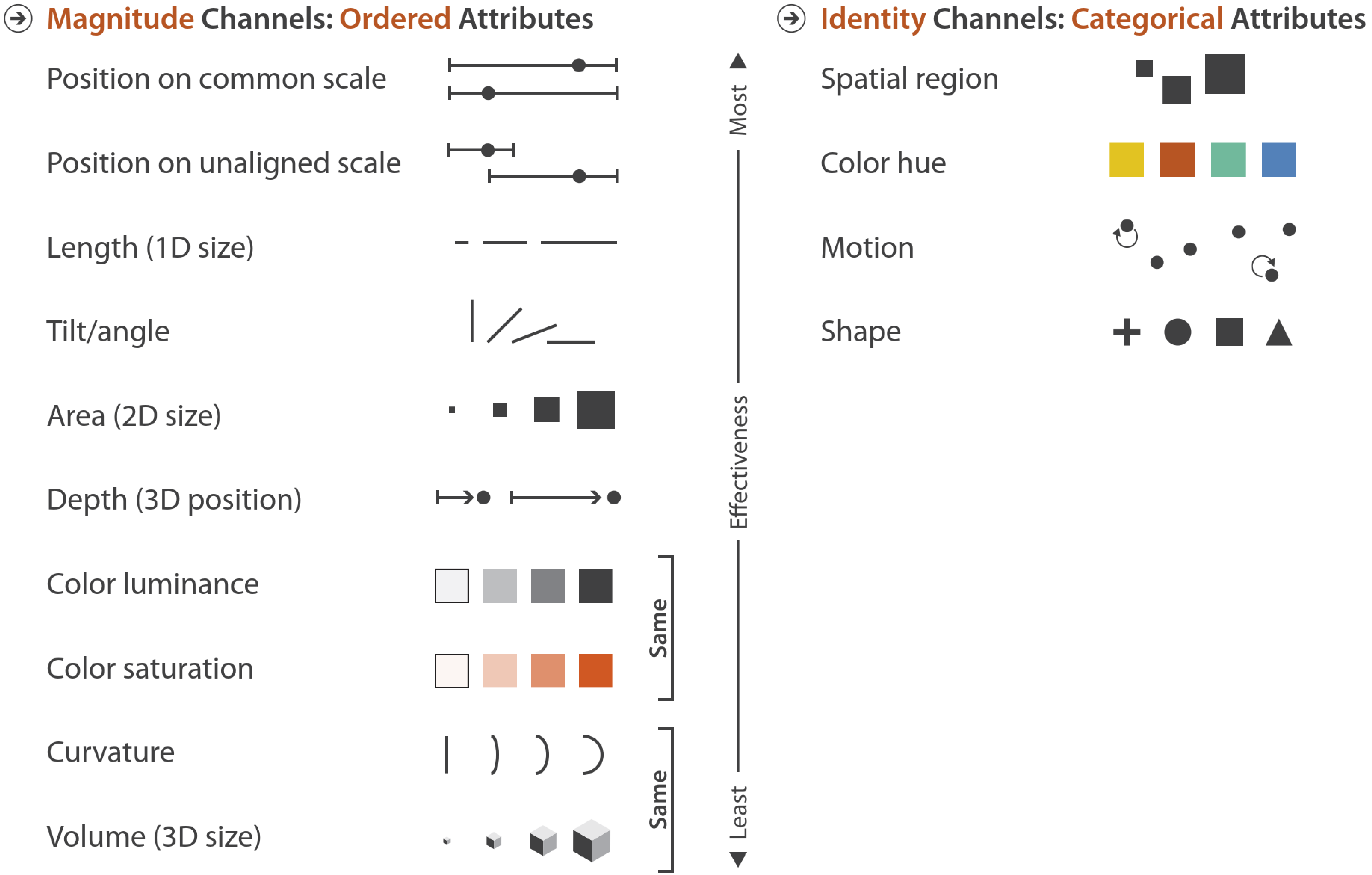


Channels :



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

Channels: Expressiveness Types and Effectiveness Ranks



Expressiveness and Effectiveness

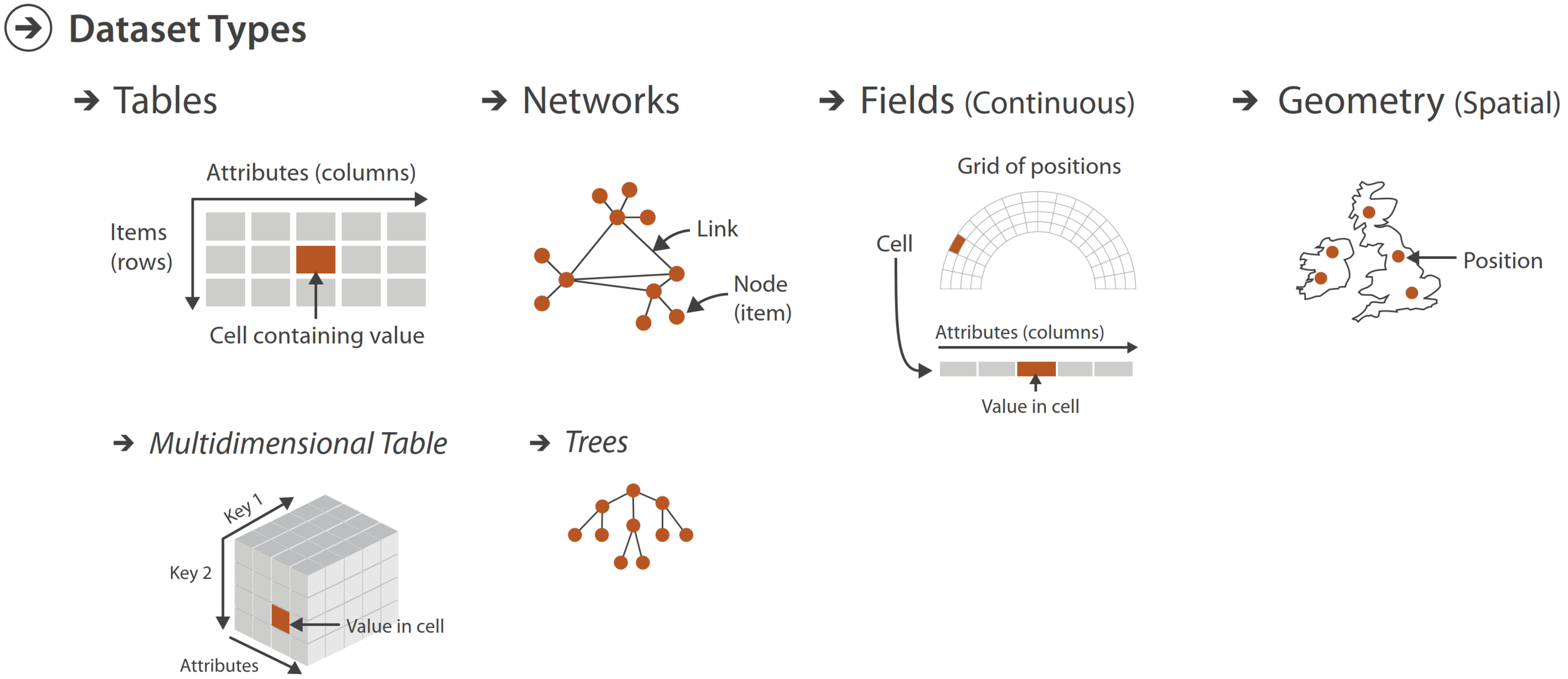
Effectiveness principle: the importance of the attribute should match the salience of the channel; that is, its noticeability.
(i.e., encode most important attributes with highest ranked channels)

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)

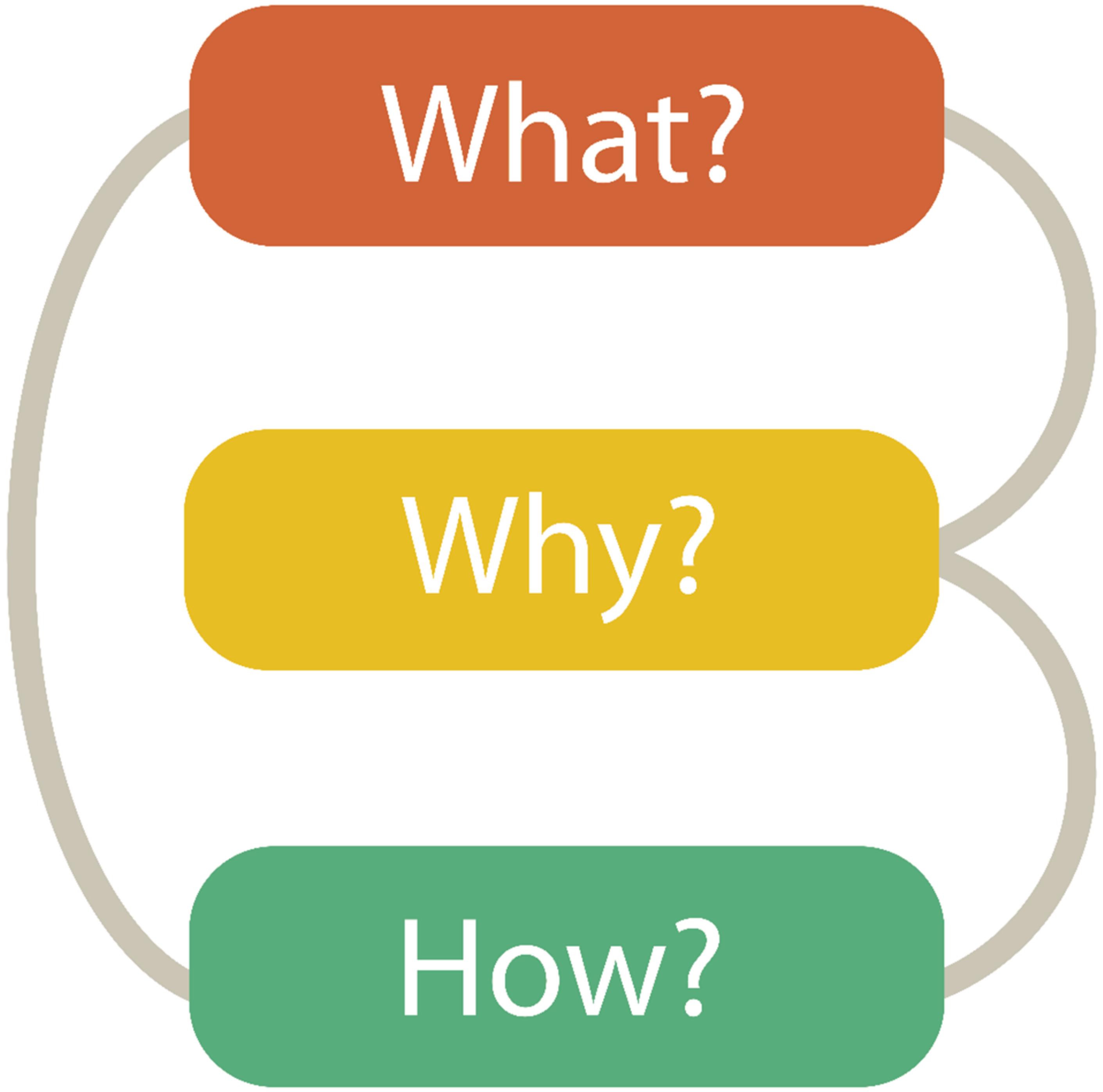
Data Types

DATASET = collection of information that is the target of analysis



Now, on CS 7250…

Analysis

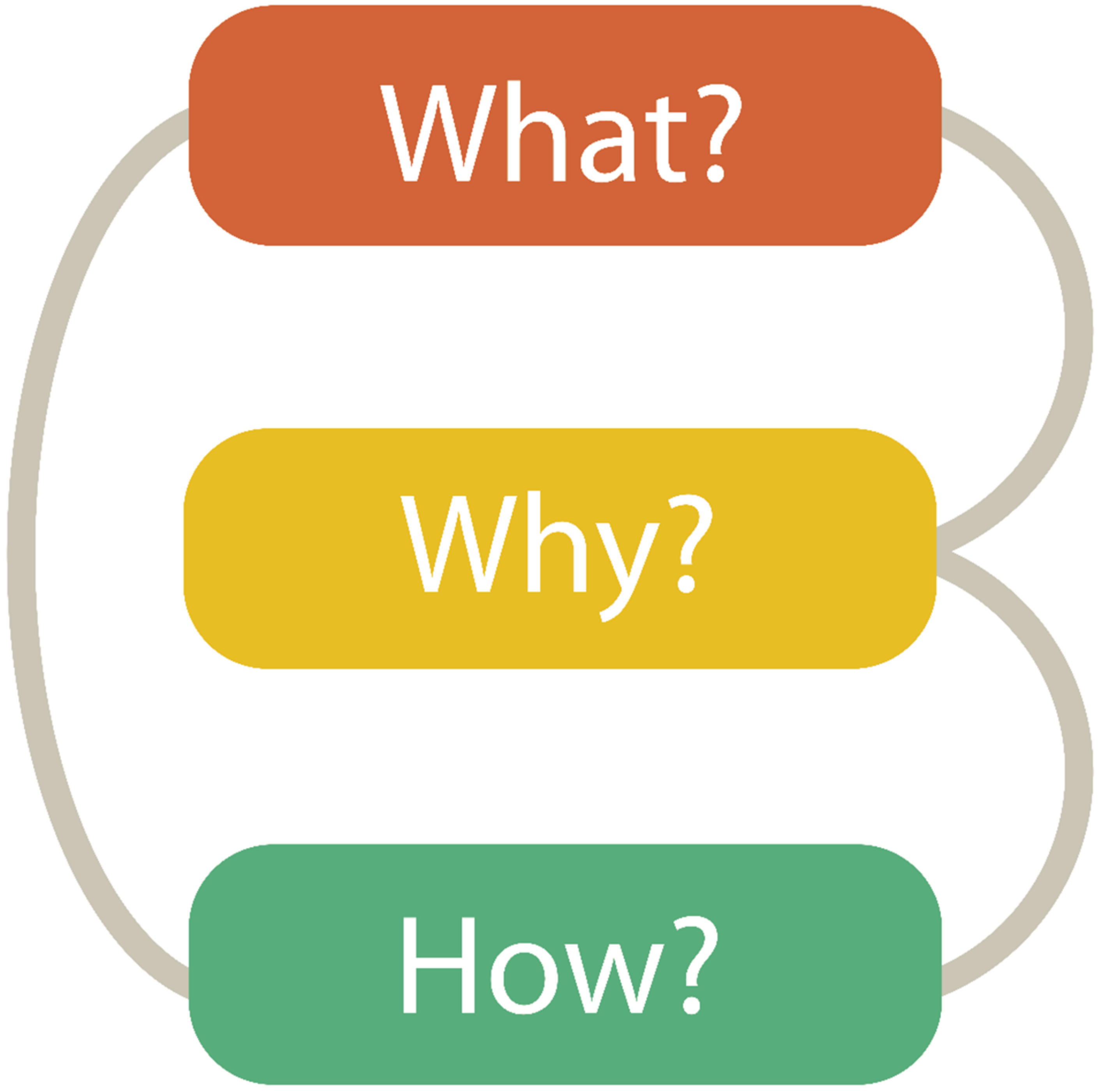


What data is shown?

Why is the user analyzing / viewing it?

How is the data presented?

Analysis

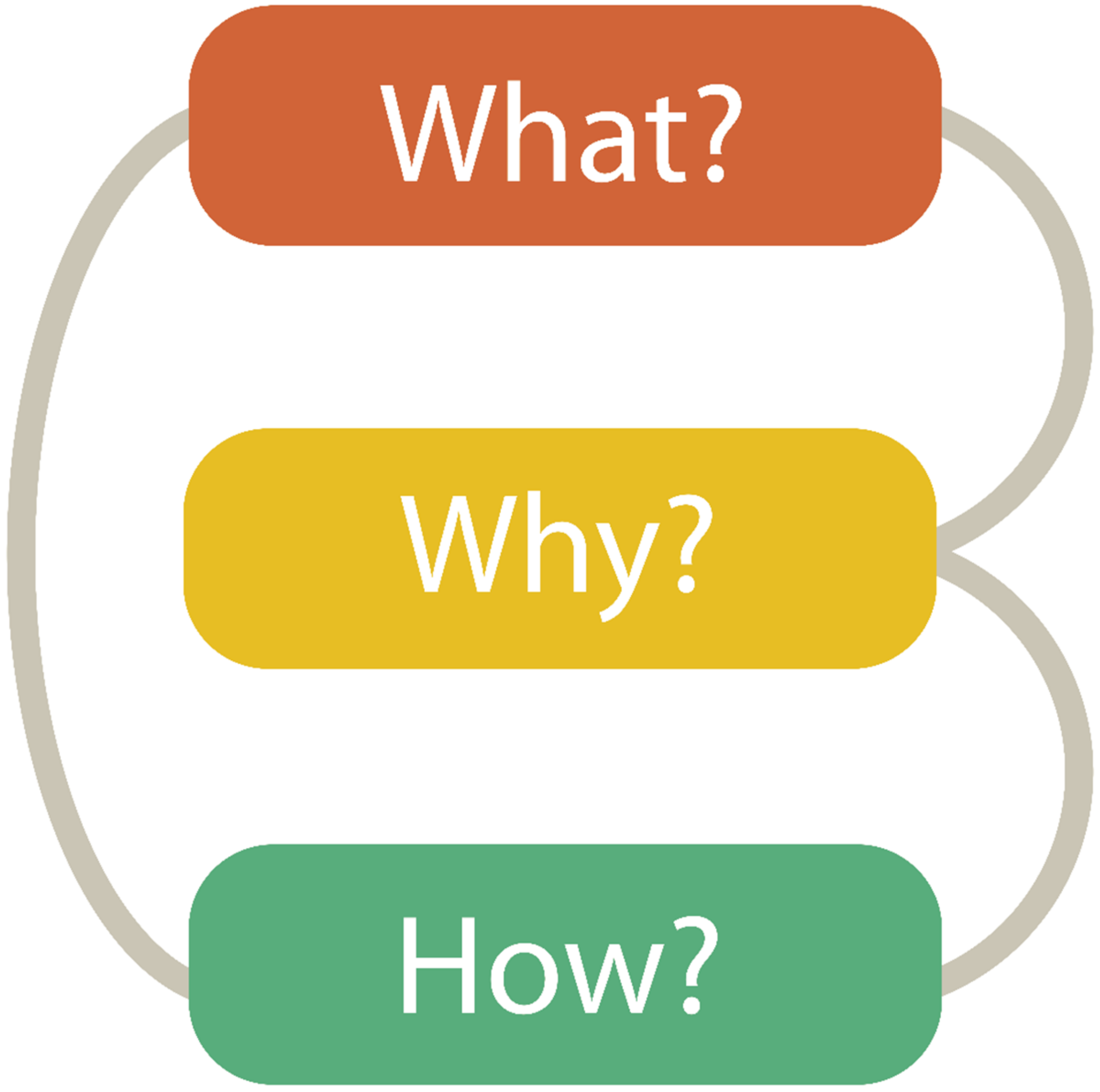


DATA ABSTRACTION

TASK ABSTRACTION

VISUAL ENCODING

Analysis



DATA ABSTRACTION

TASK ABSTRACTION

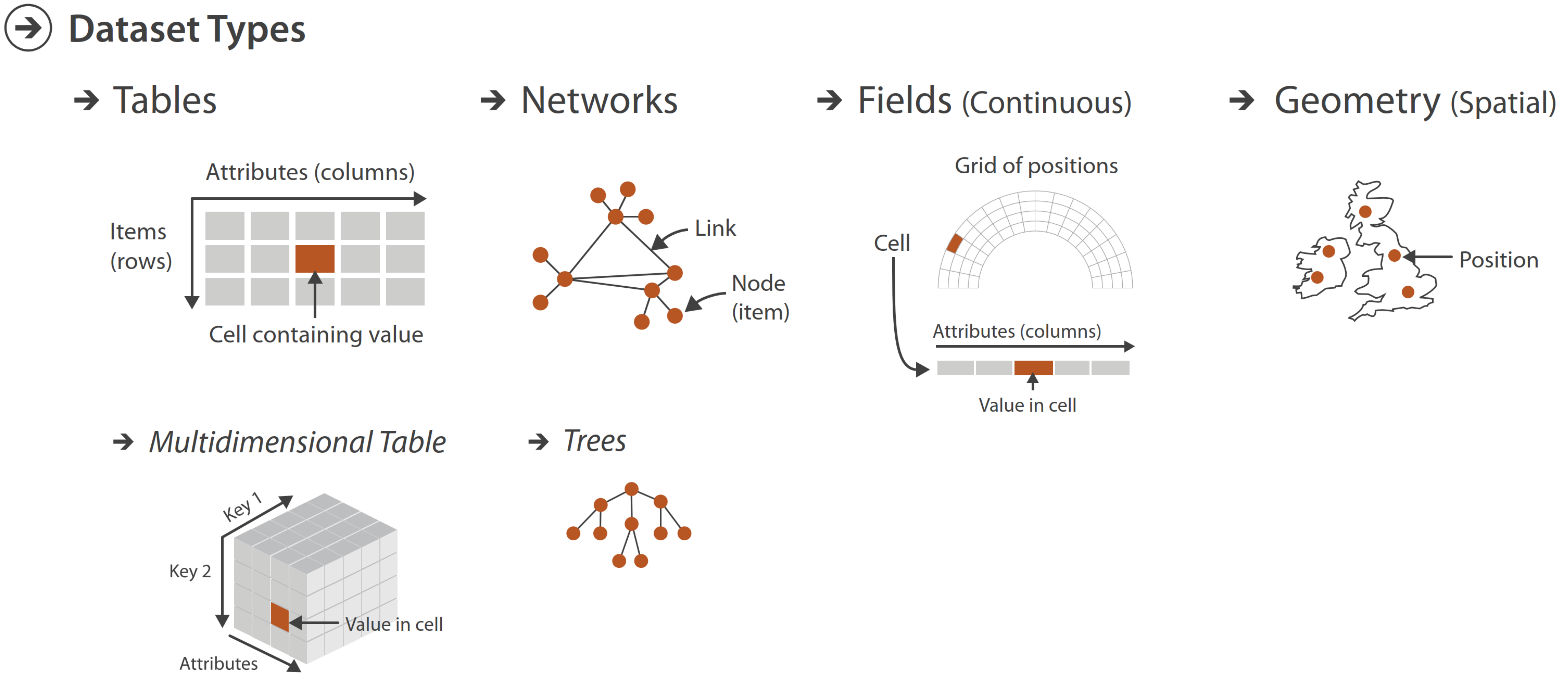
VISUAL ENCODING

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

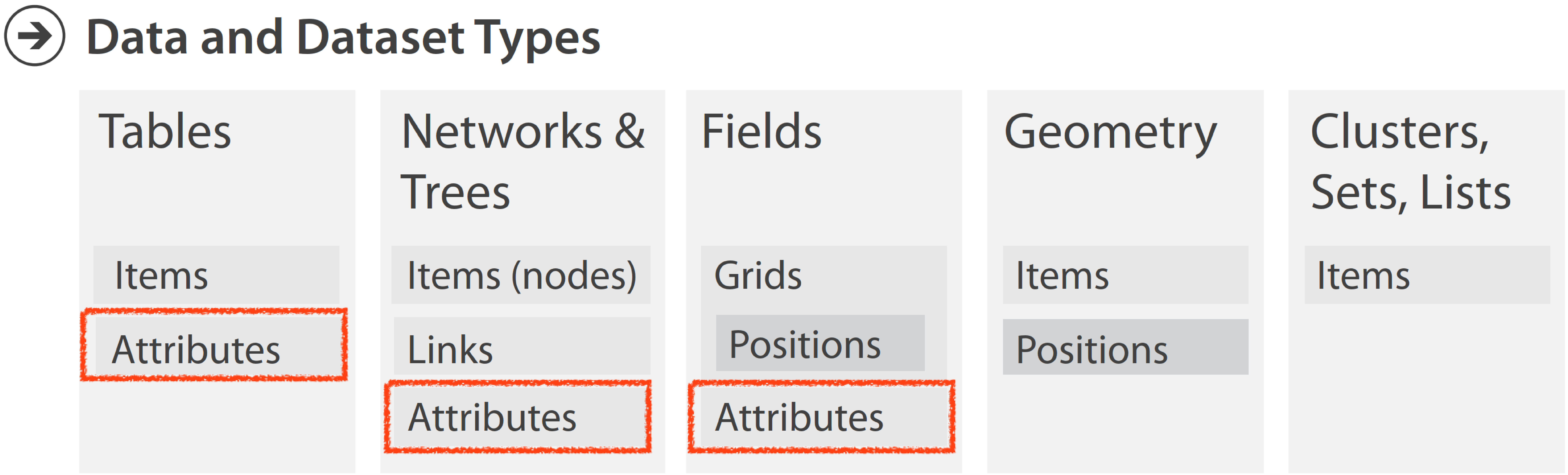
Data Types

DATASET = collection of information that is the target of analysis



Data Types

DATASET = collection of information that is the target of analysis



Attribute Types

- Categorical

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

- Ordered
  - *Ordinal*

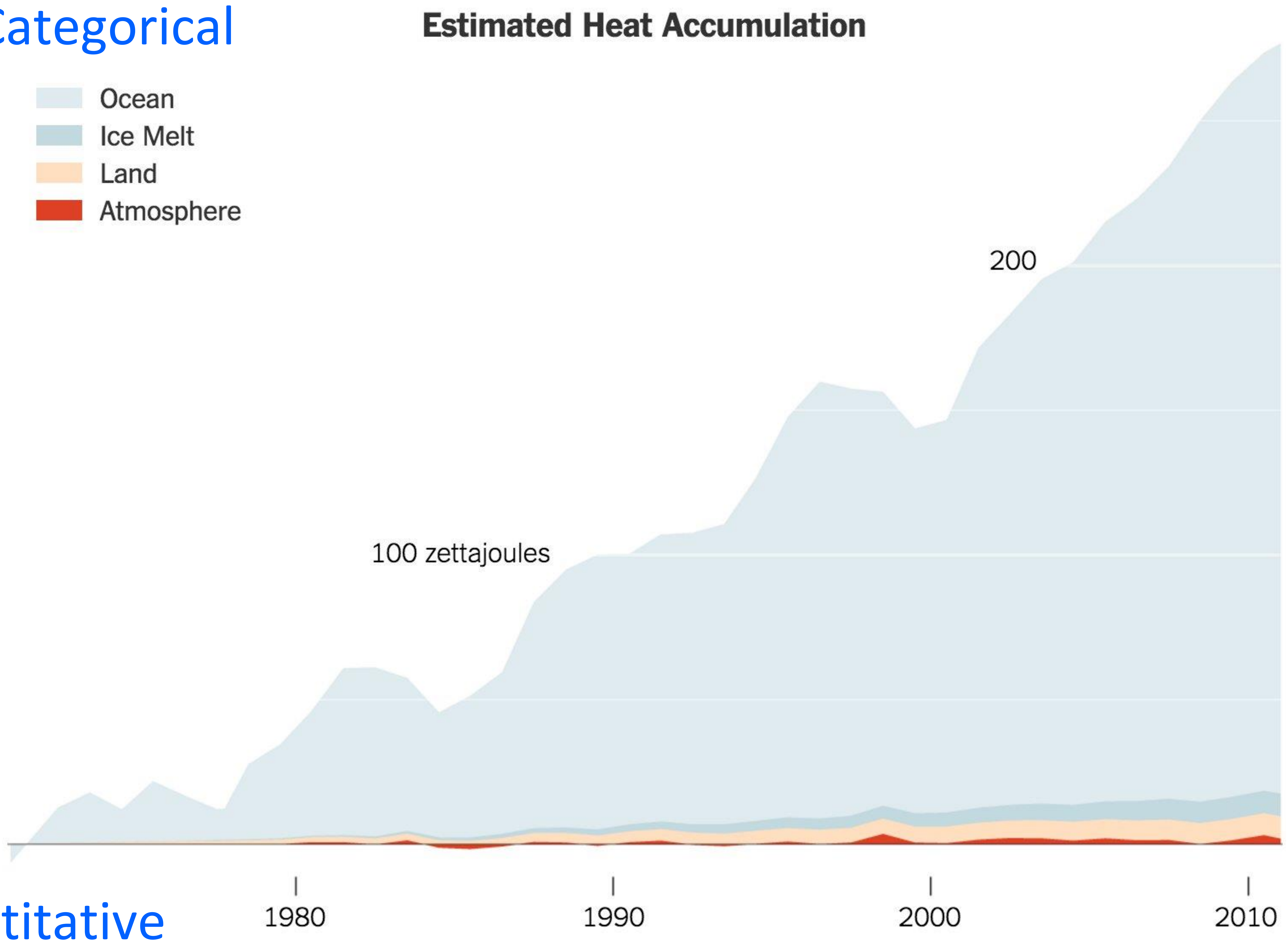
e.g.,
sizes (xs, s, m, l, xl),
months (J, F, M)

  - *Quantitative* *(continuous)*

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

Categorical

Quantitative

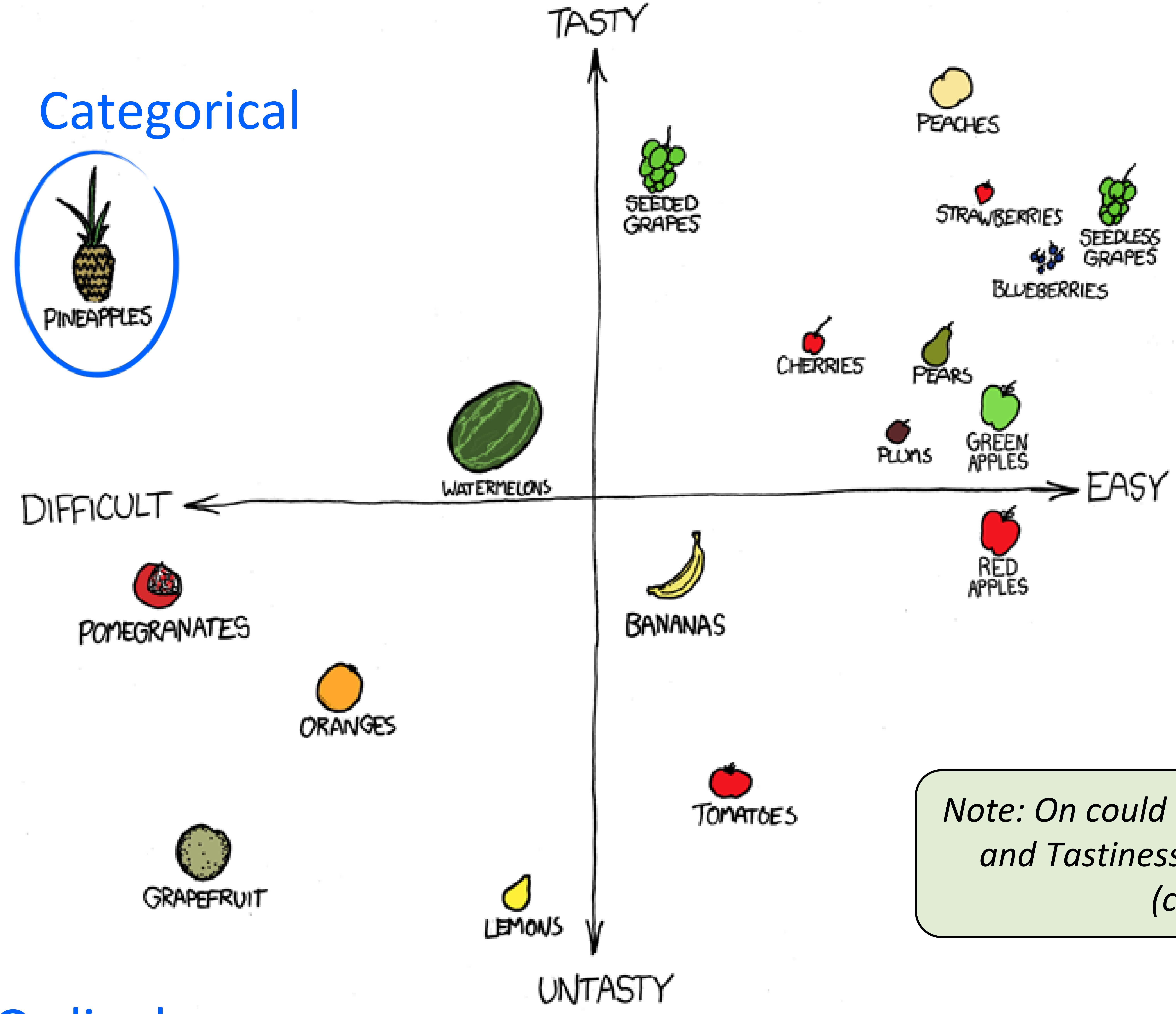


Quantitative

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

Categorical

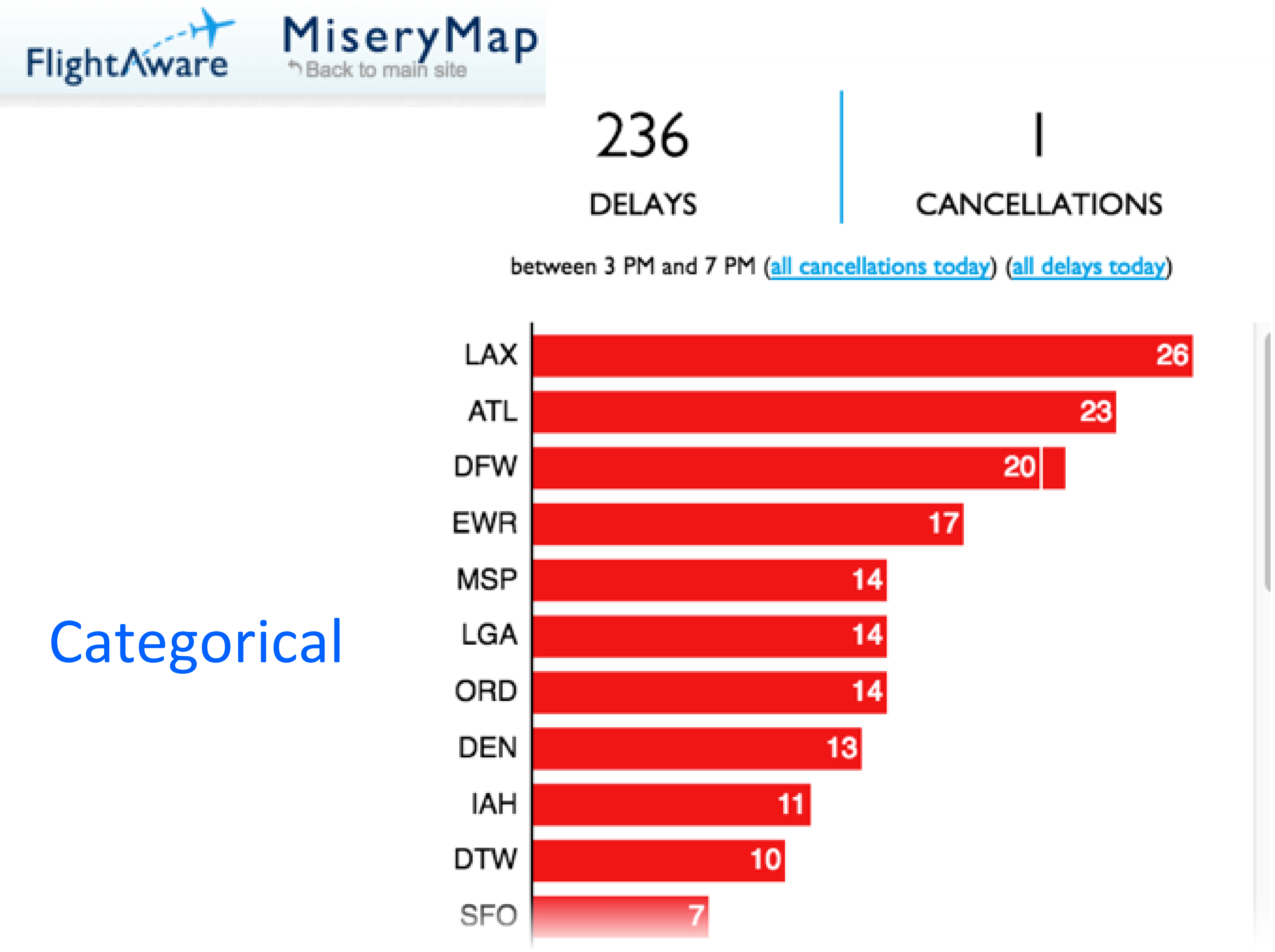
Ordinal



Ordinal

Note: On could also argue that Difficulty and Tastiness could be quantitative (continuous)

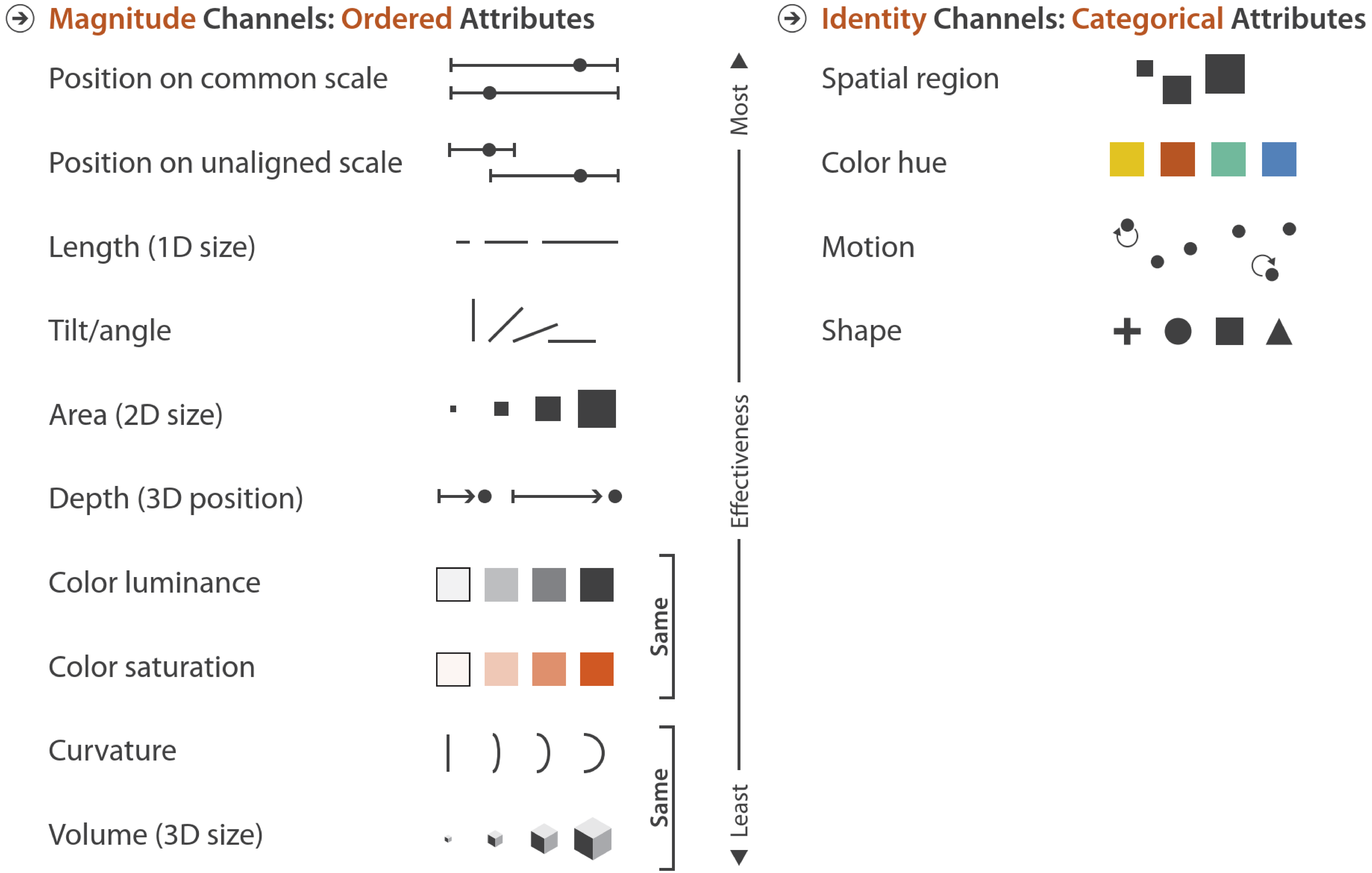
https://xkcd.com/388/



Categorical

Quantitative

Channels: Expressiveness Types and Effectiveness Ranks



Channel Ranking by Data Type

Quantitative

Position
Length
Angle
Slope
Area
Volume
Density
Color Saturation
Color Hue
Texture
Connection
Containment
Shape

Figure 15: Ranking of Perceptual Tasks. *The tasks shown in the gray boxes are not relevant to that type of data.*

Mackinlay (1986)

Channel Ranking by Data Type

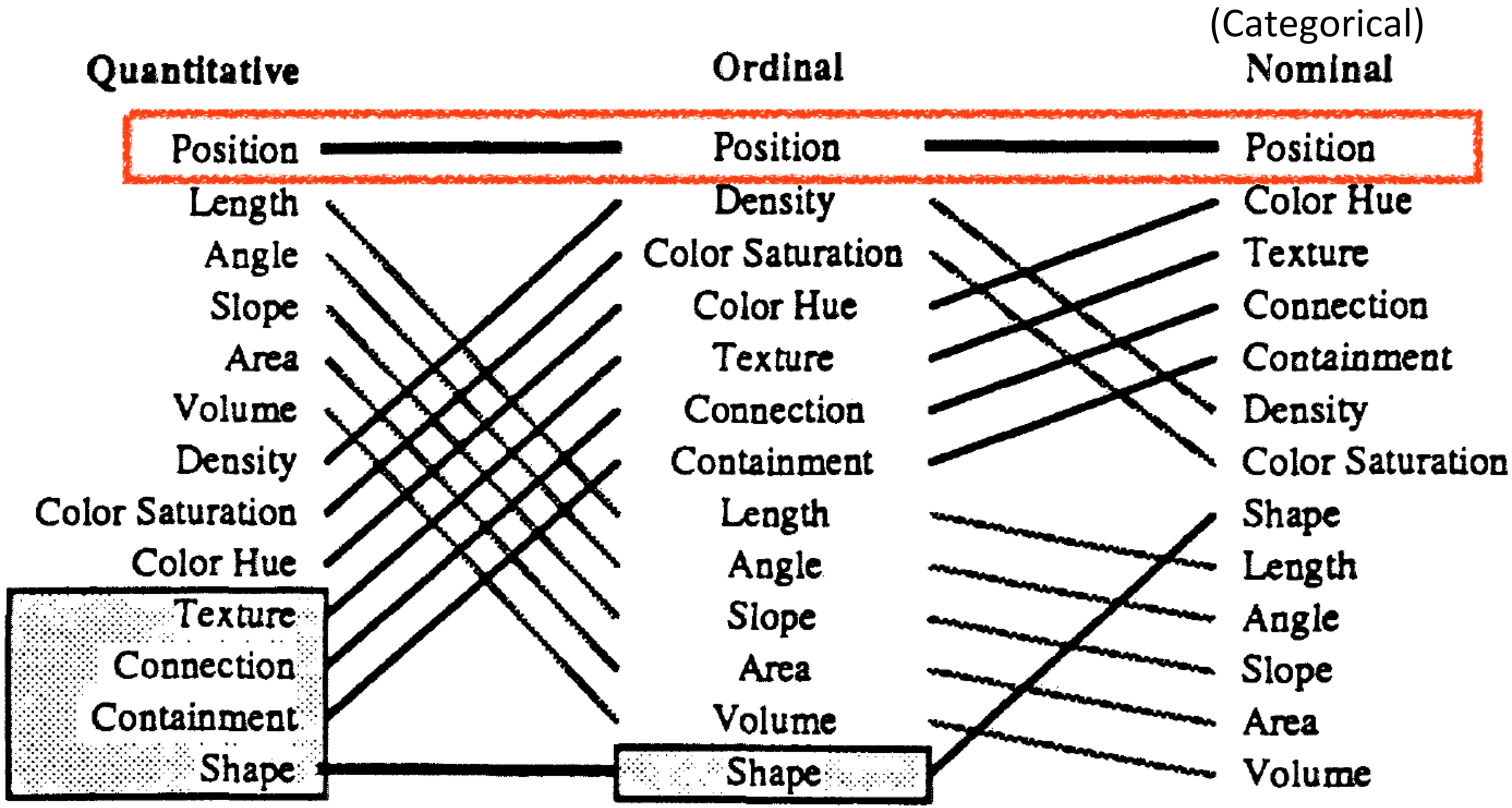


Figure 15: Ranking of Perceptual Tasks. *The tasks shown in the gray boxes are not relevant to that type of data.*

Mackinlay (1986)

Channel Ranking by Data Type

AREA

Quantitative

Ordinal

Categorical

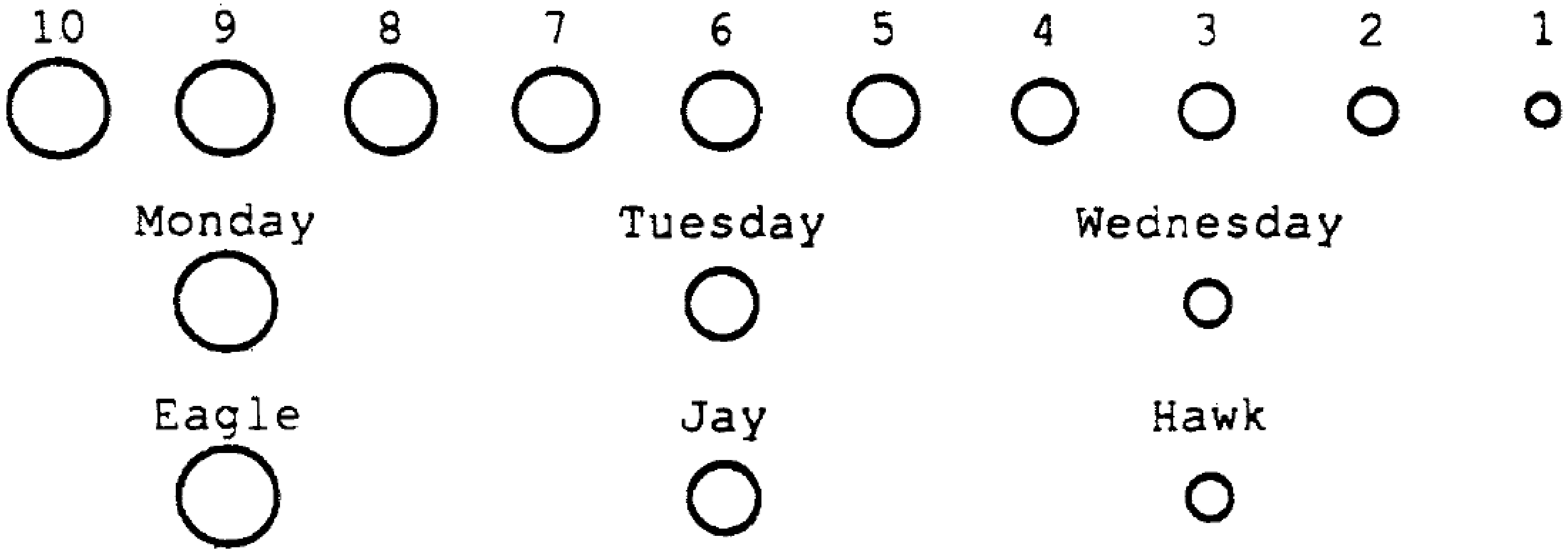


Figure 16: Analysis of the Area Task.

Channel Ranking by Data Type

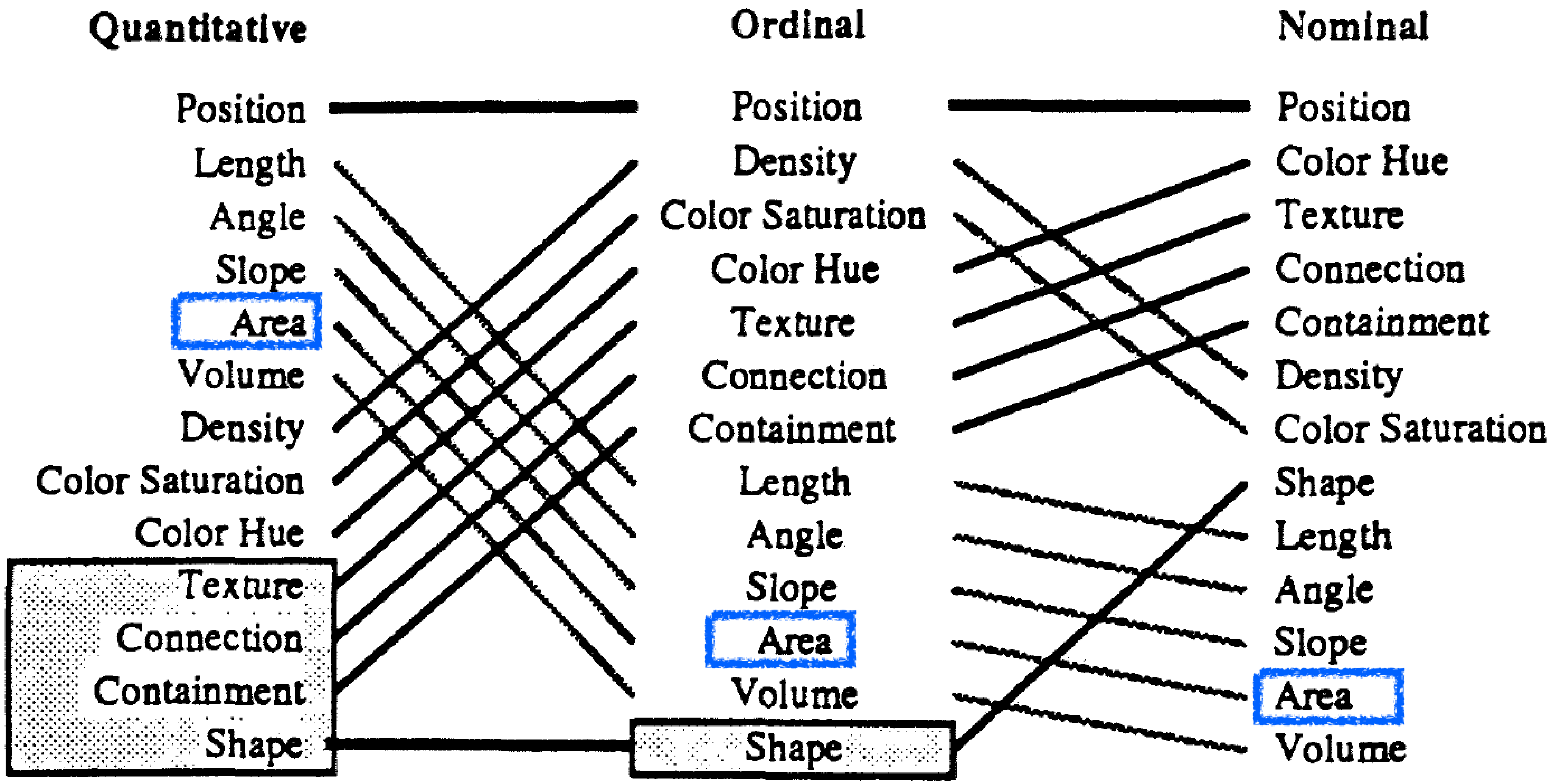
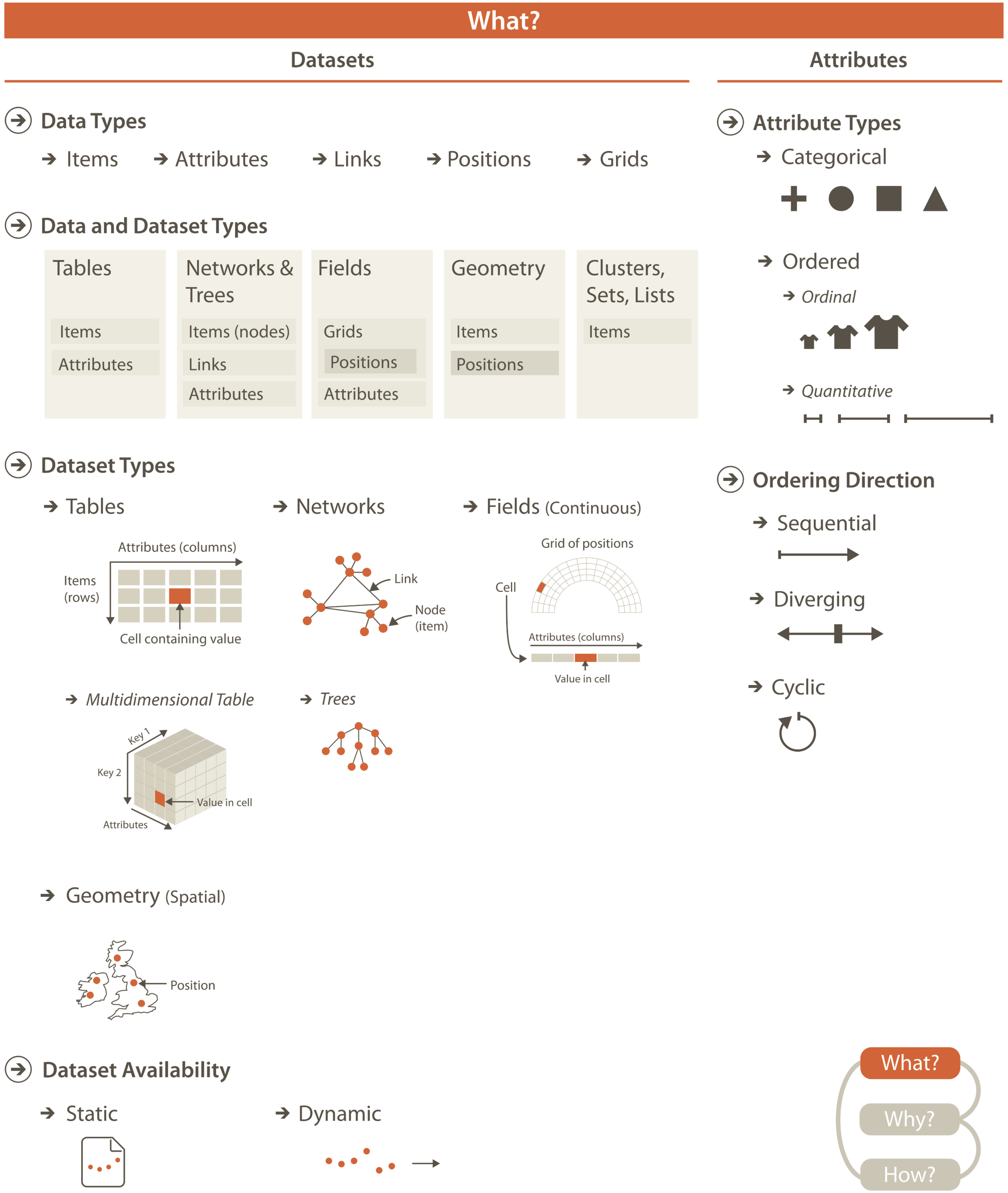
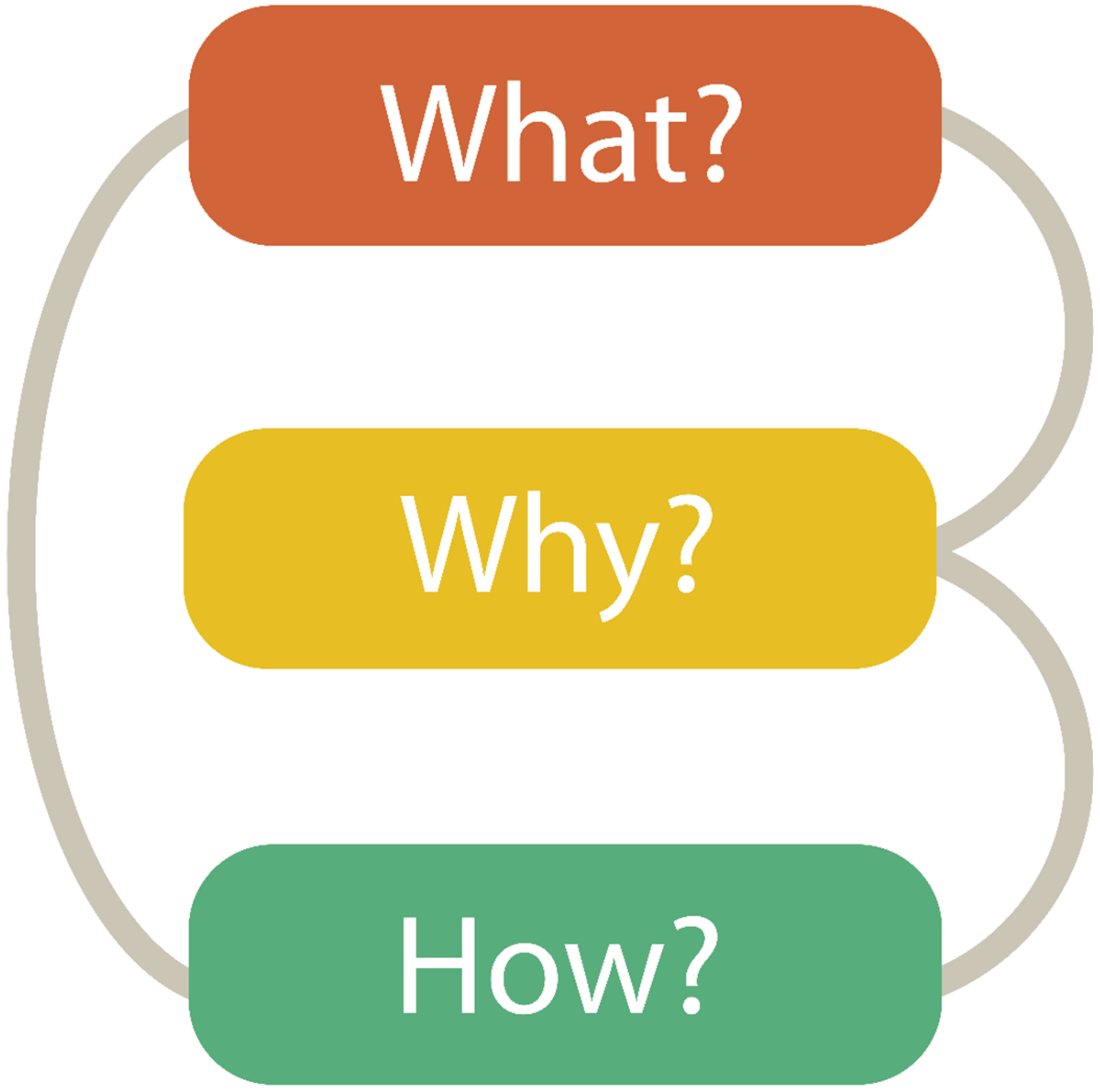


Figure 15: Ranking of Perceptual Tasks. *The tasks shown in the gray boxes are not relevant to that type of data.*

Data Abstraction



Analysis



DATA ABSTRACTION

TASK ABSTRACTION

VISUAL ENCODING

Goals for Today

- Learn what “Tasks” are and why they are so important.
- Learn the differences between high, mid, and low level task classifications.
- Begin practicing how to classify tasks (key step in visualization design process!).

Task Abstraction

Why abstract?

Avoids domain specific terms thus easier to apply to other cases (broadly applicable results).

Task Abstraction

Why abstract?

Avoids domain specific terms thus easier to apply to other cases (broadly applicable results).

Visualization Tools

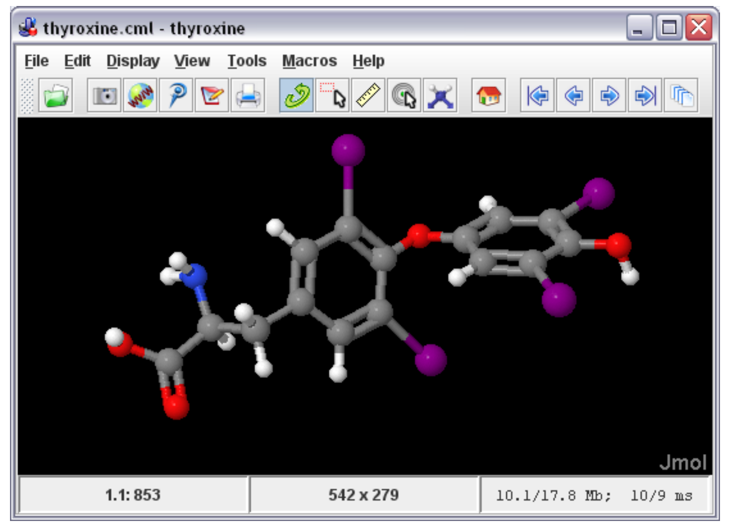
Specific ⟷ *General*

Gephi

Altair

ArcGIS

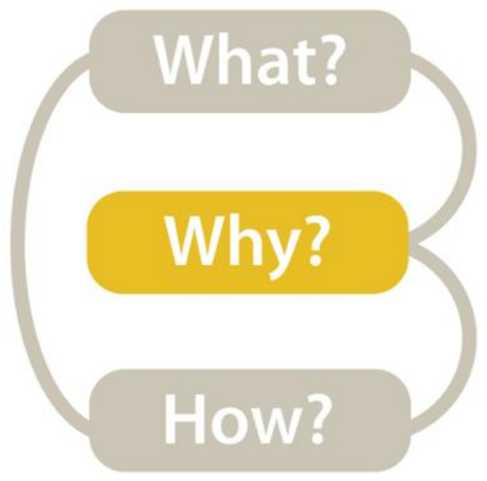
tableau



TASK ABSTRACTION

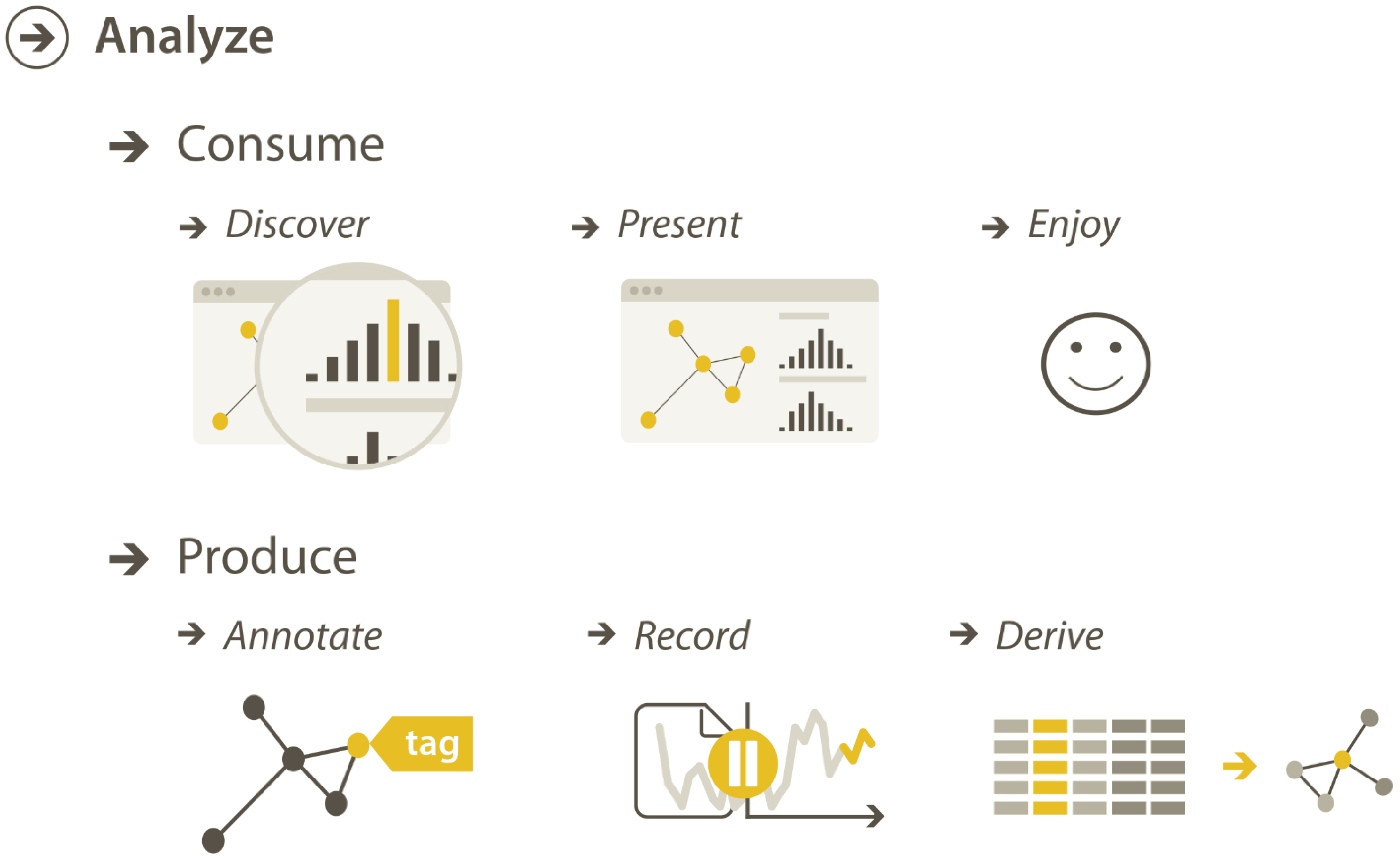
Why abstract?

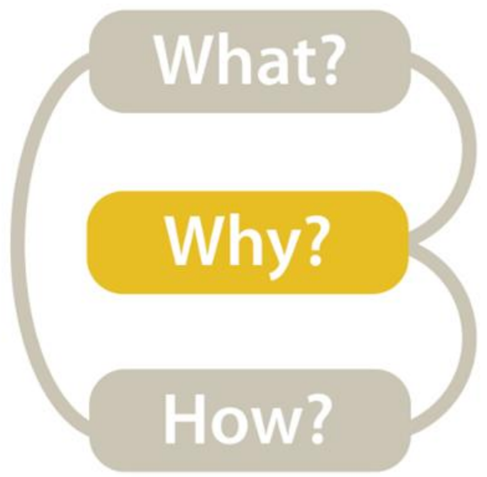
Avoids domain specific terms thus easier to apply to other cases (broadly applicable results).



High-level

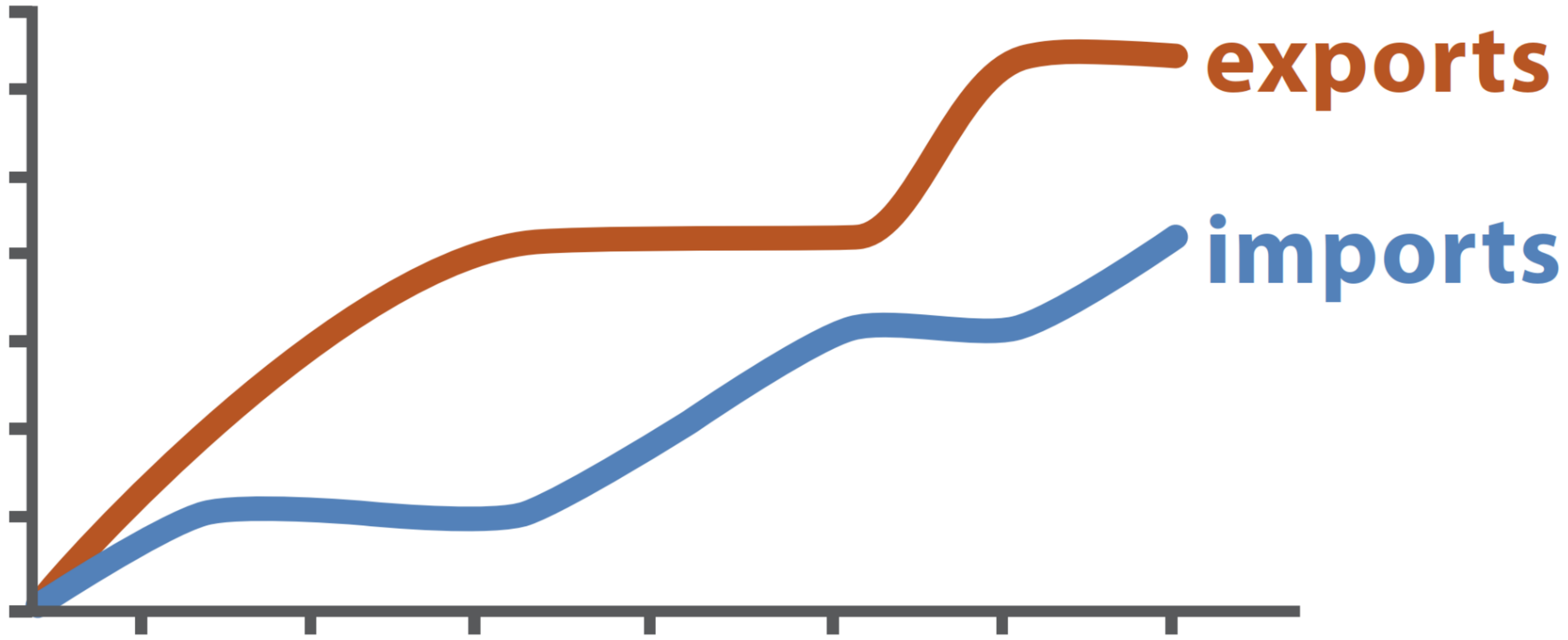
ACTIONS define user goals.



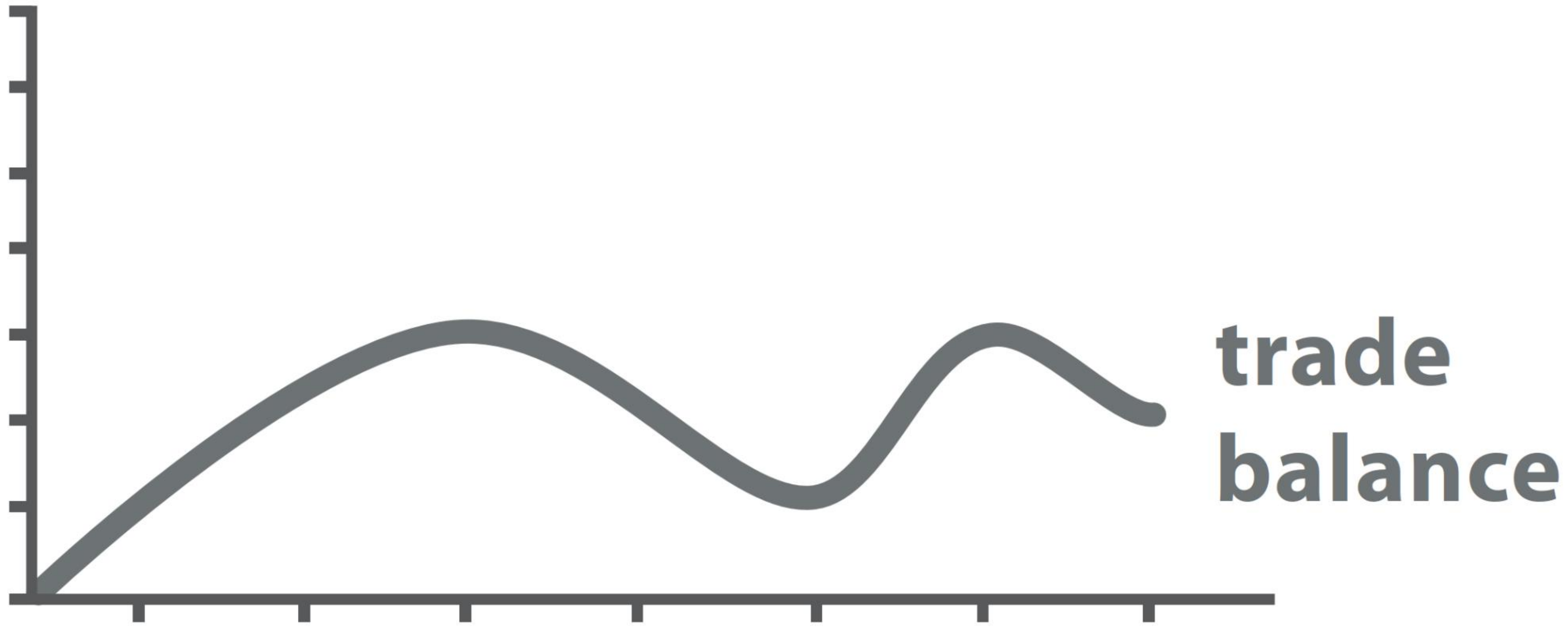


Actions define user goals.

→ *Derive*

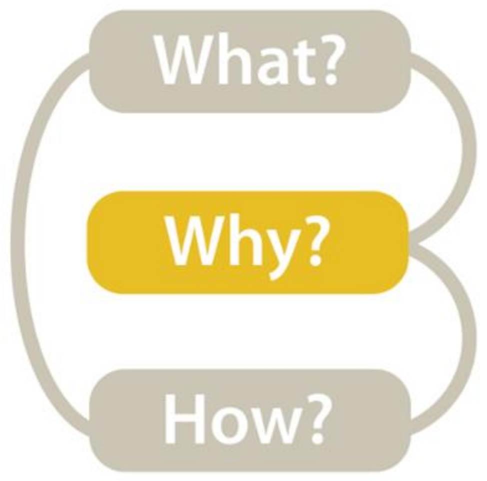


Original Data



trade balance = exports − imports

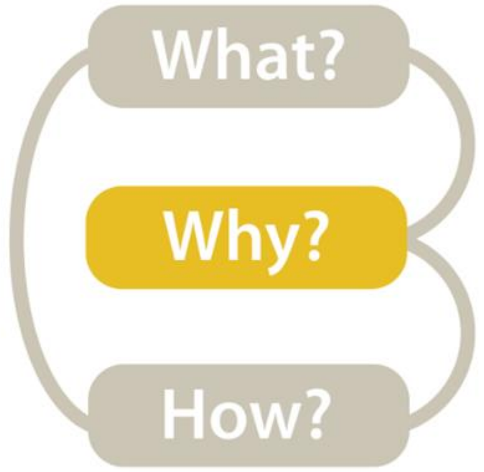
Derived Data



Actions define user goals. *Mid-level*

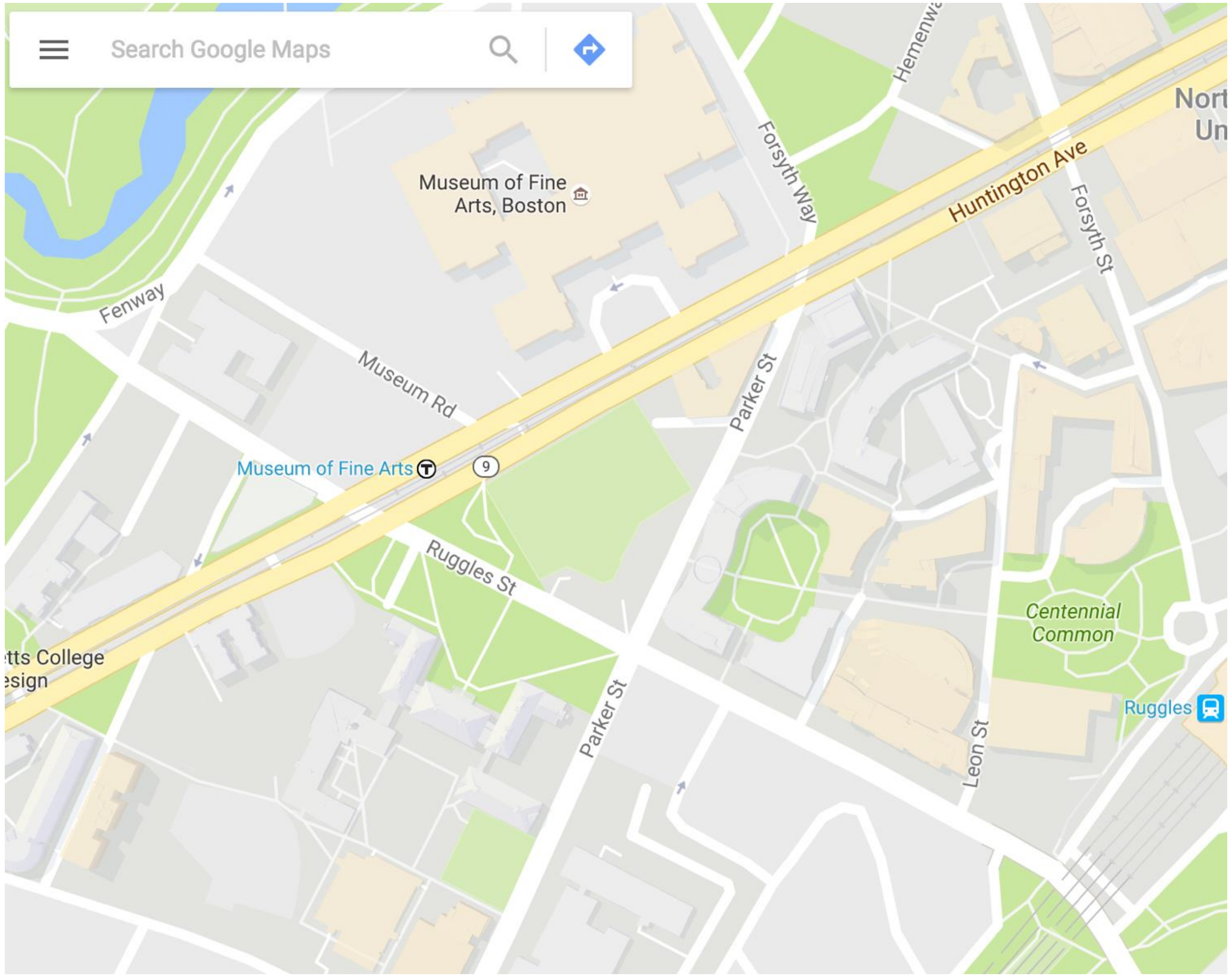
Search

	Target known	Target unknown
Location known	*Lookup*	*Browse*
Location unknown	*Locate*	*Explore*



Search

	Target known	Target unknown
Location known	*Lookup*	*Browse*
Location unknown	*Locate*	*Explore*



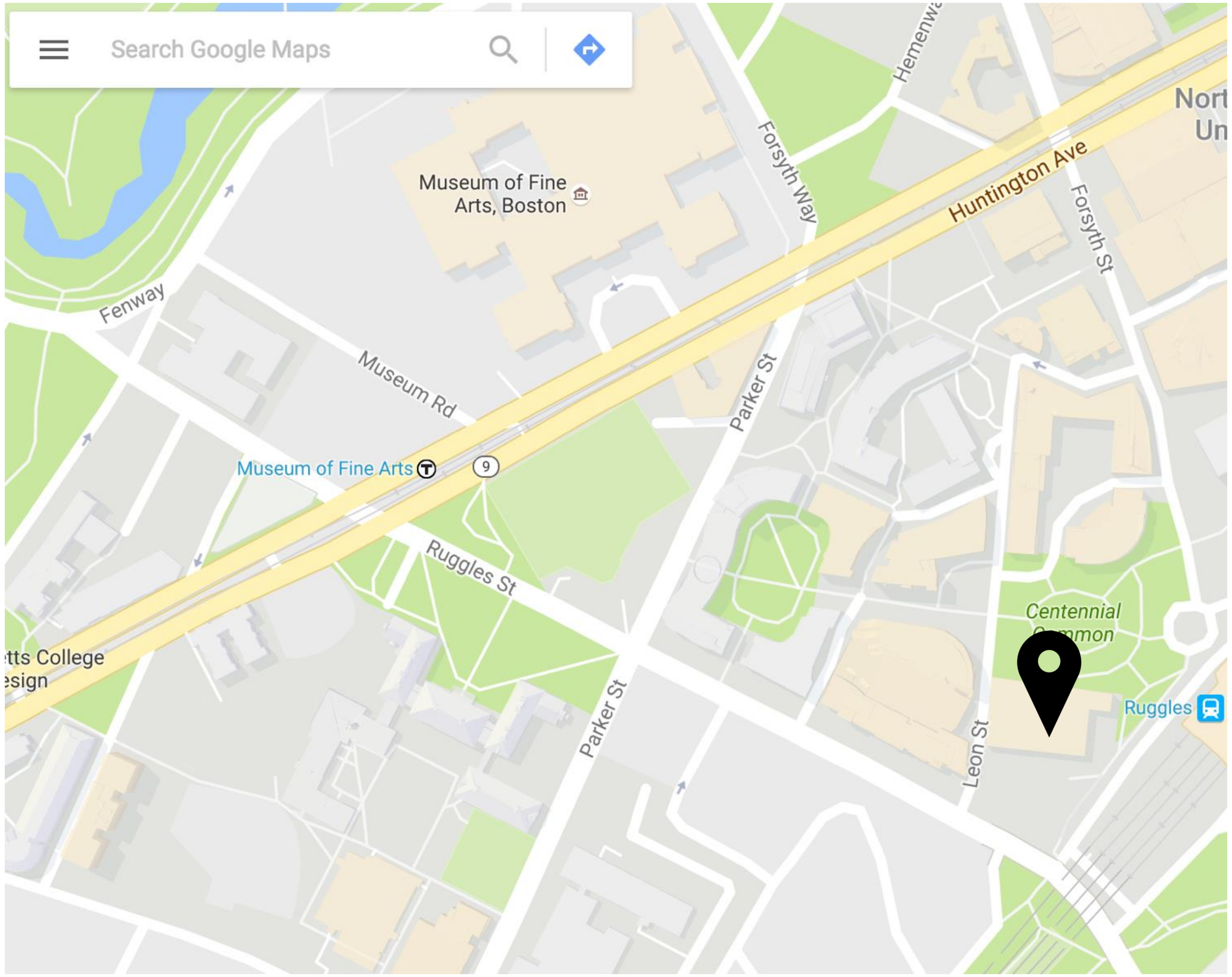
What?

Why?

How?

Search

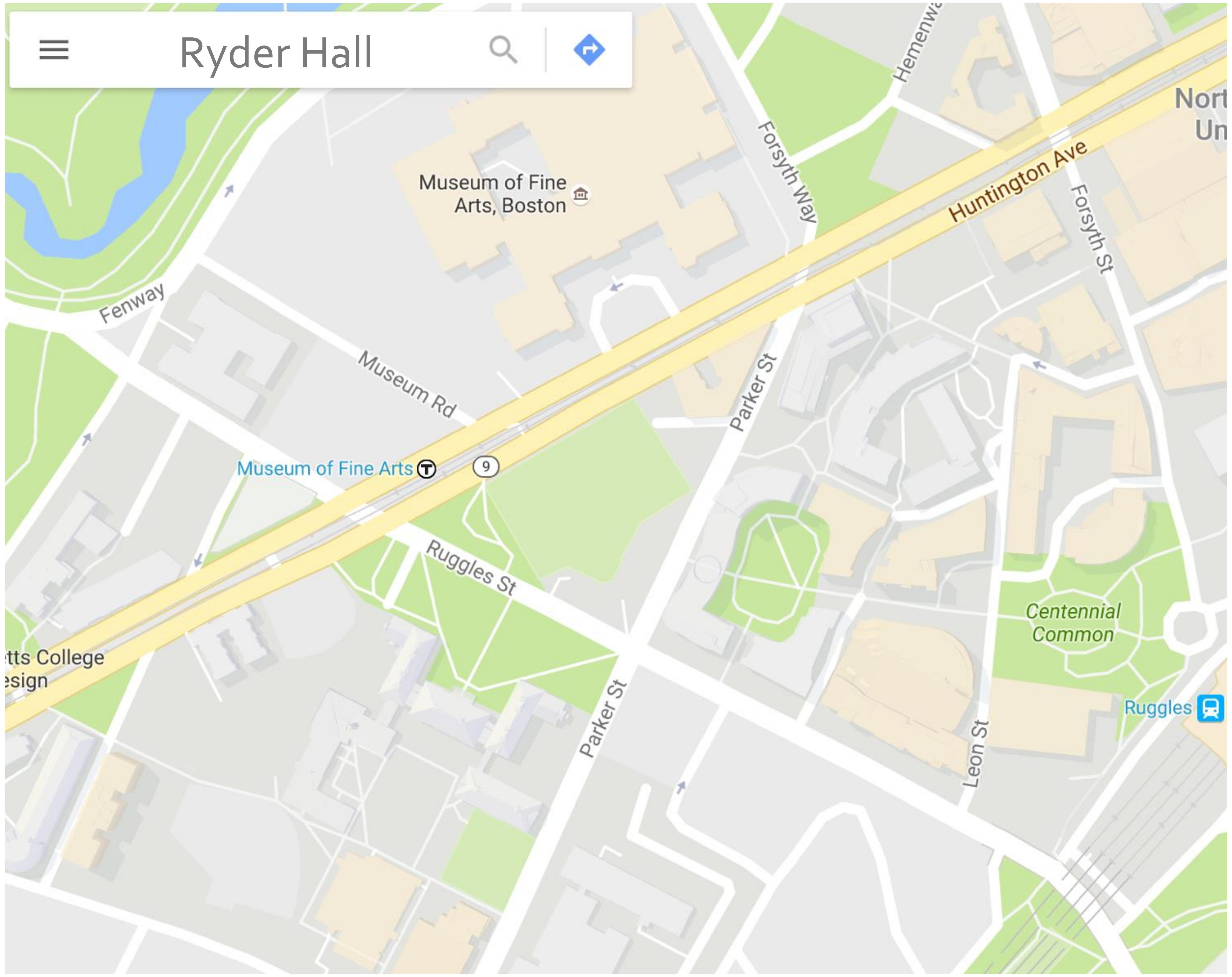
	Target known	Target unknown
Location known	*Lookup*	*Browse*
Location unknown	*Locate*	*Explore*



What is the address of Ryder hall?

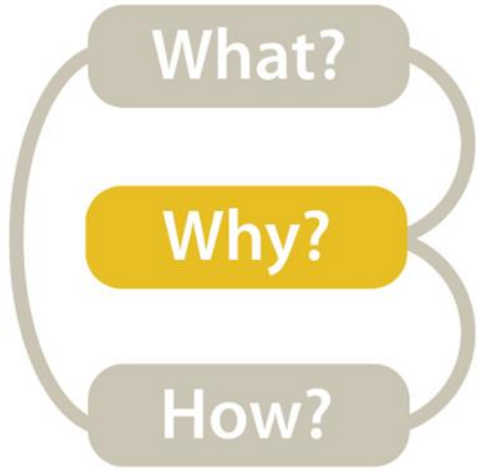
What?
Why?
How?

Search



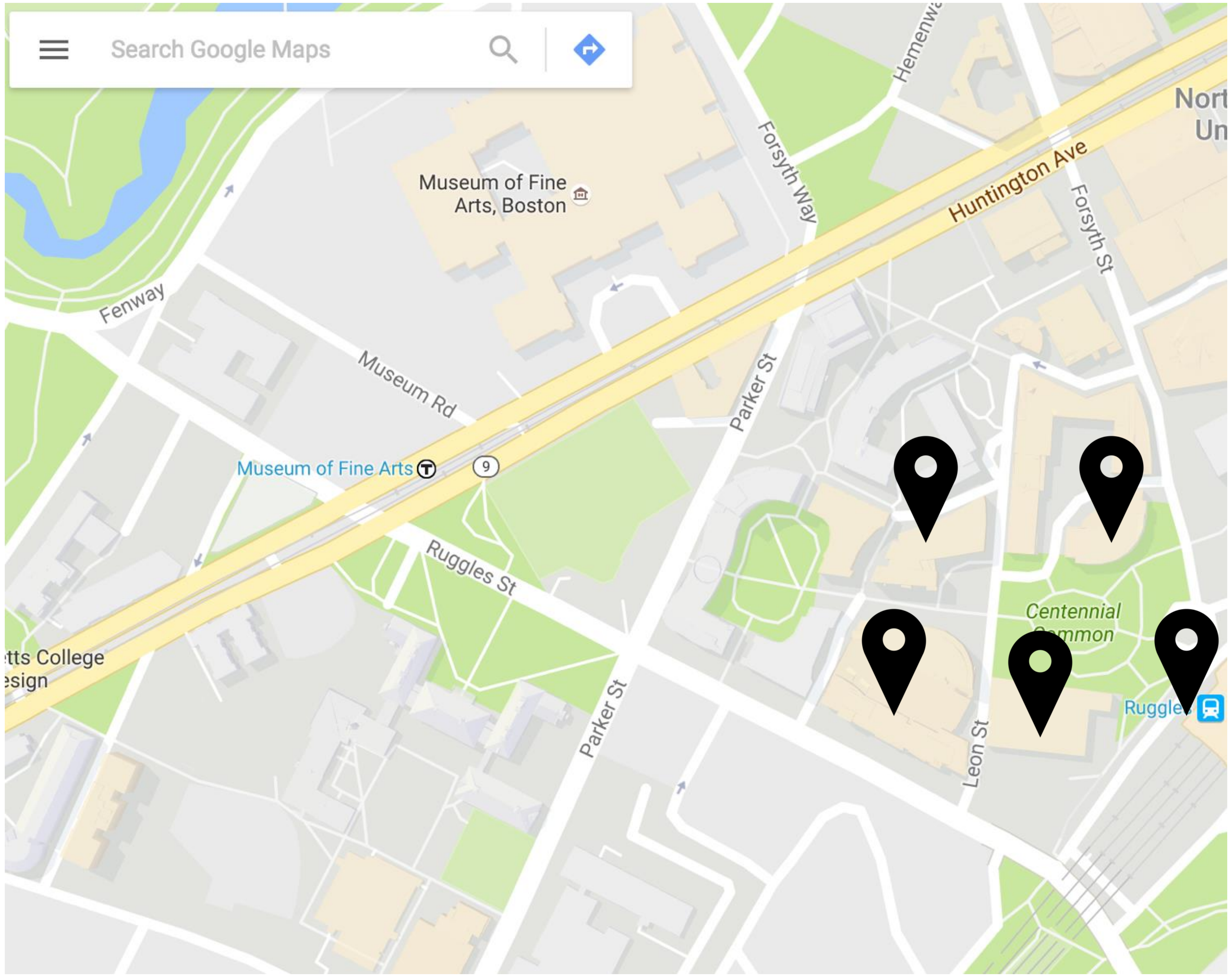
	Target known	Target unknown
Location known	*Lookup*	*Browse*
Location unknown	*Locate*	*Explore*

Where is Ryder Hall?



Search

	Target known	Target unknown
Location known	*Lookup*	*Browse*
Location unknown	*Locate*	*Explore*



What buildings are near Ryder Hall?

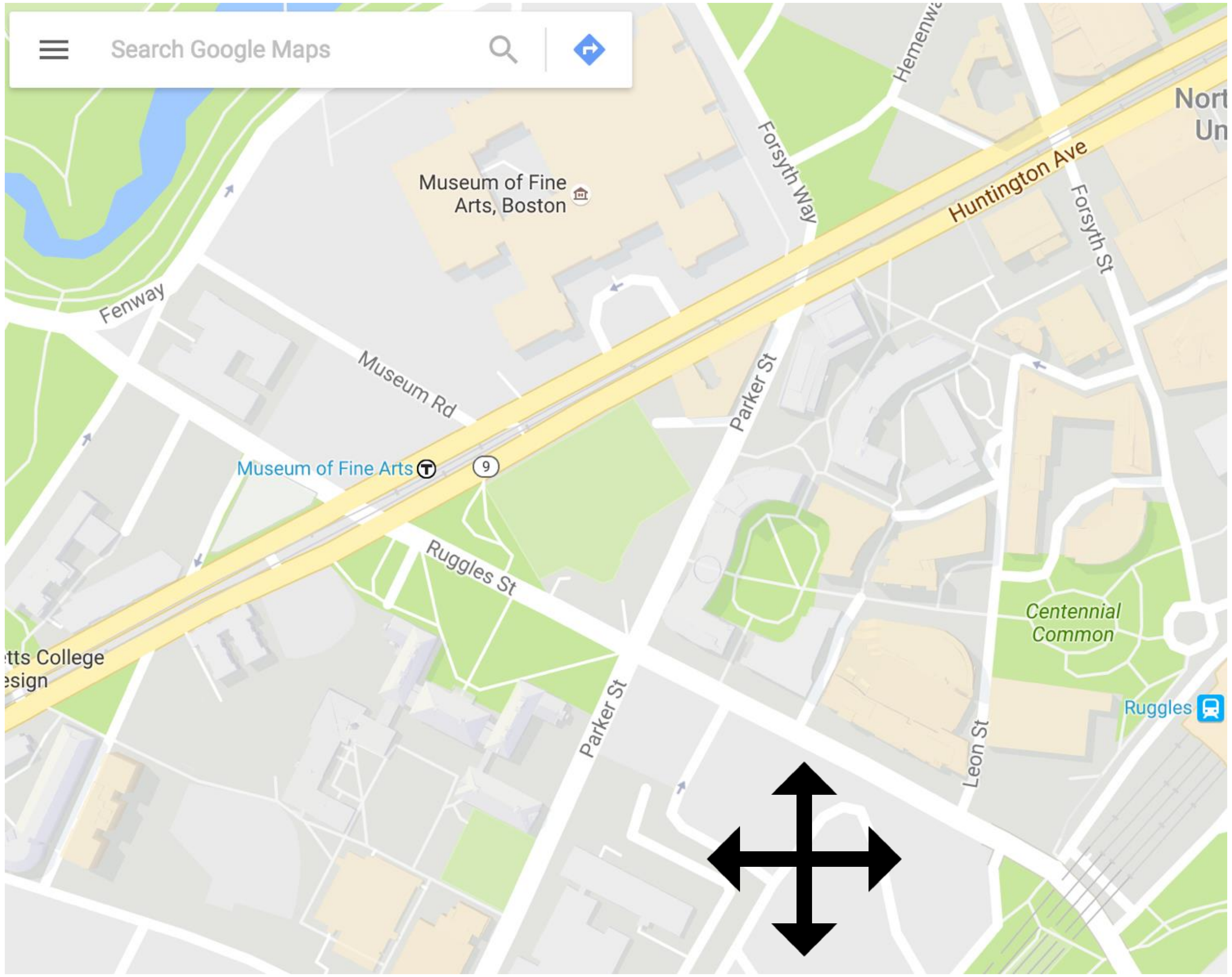
What?

Why?

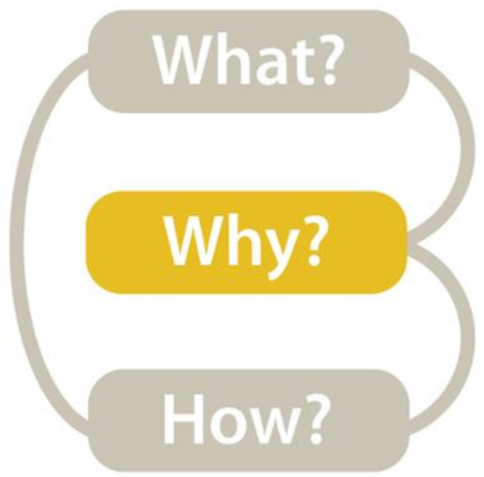
How?

Search

	Target known	Target unknown
Location known	*Lookup*	*Browse*
Location unknown	*Locate*	*Explore*



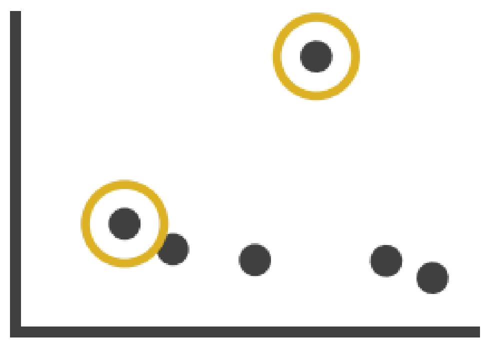
What is south of Huntington Ave?



ACTIONS define user goals. *Low-level*

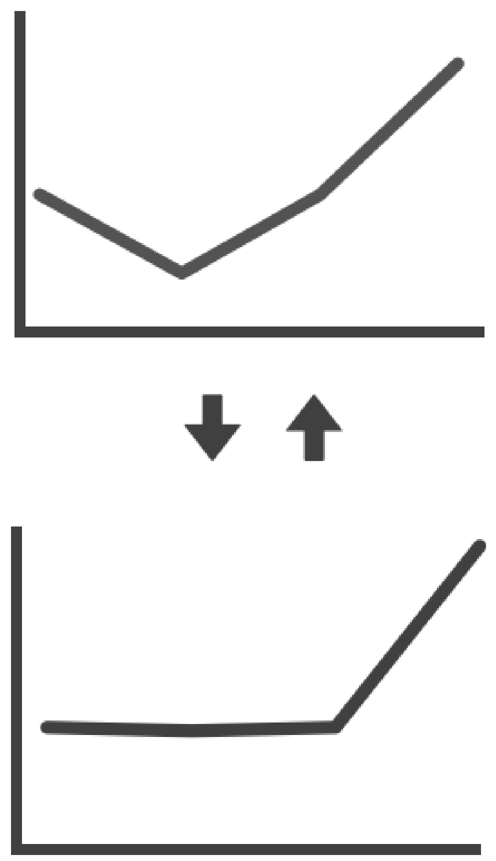
Query

Identify



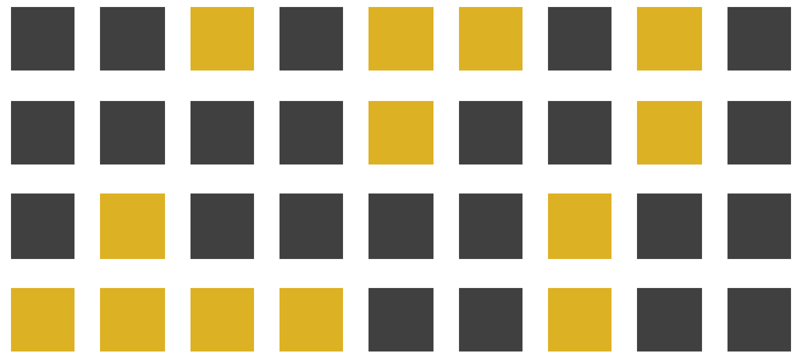
single target

Compare

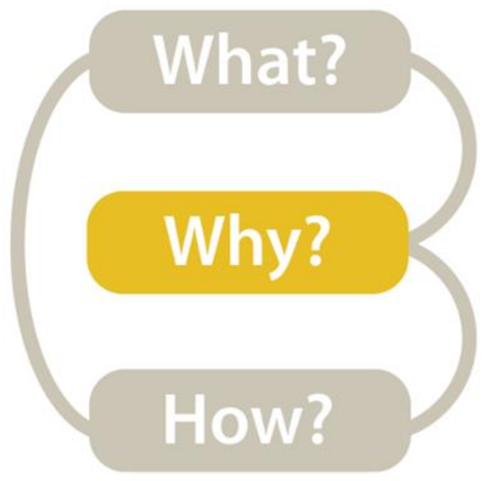


multiple targets

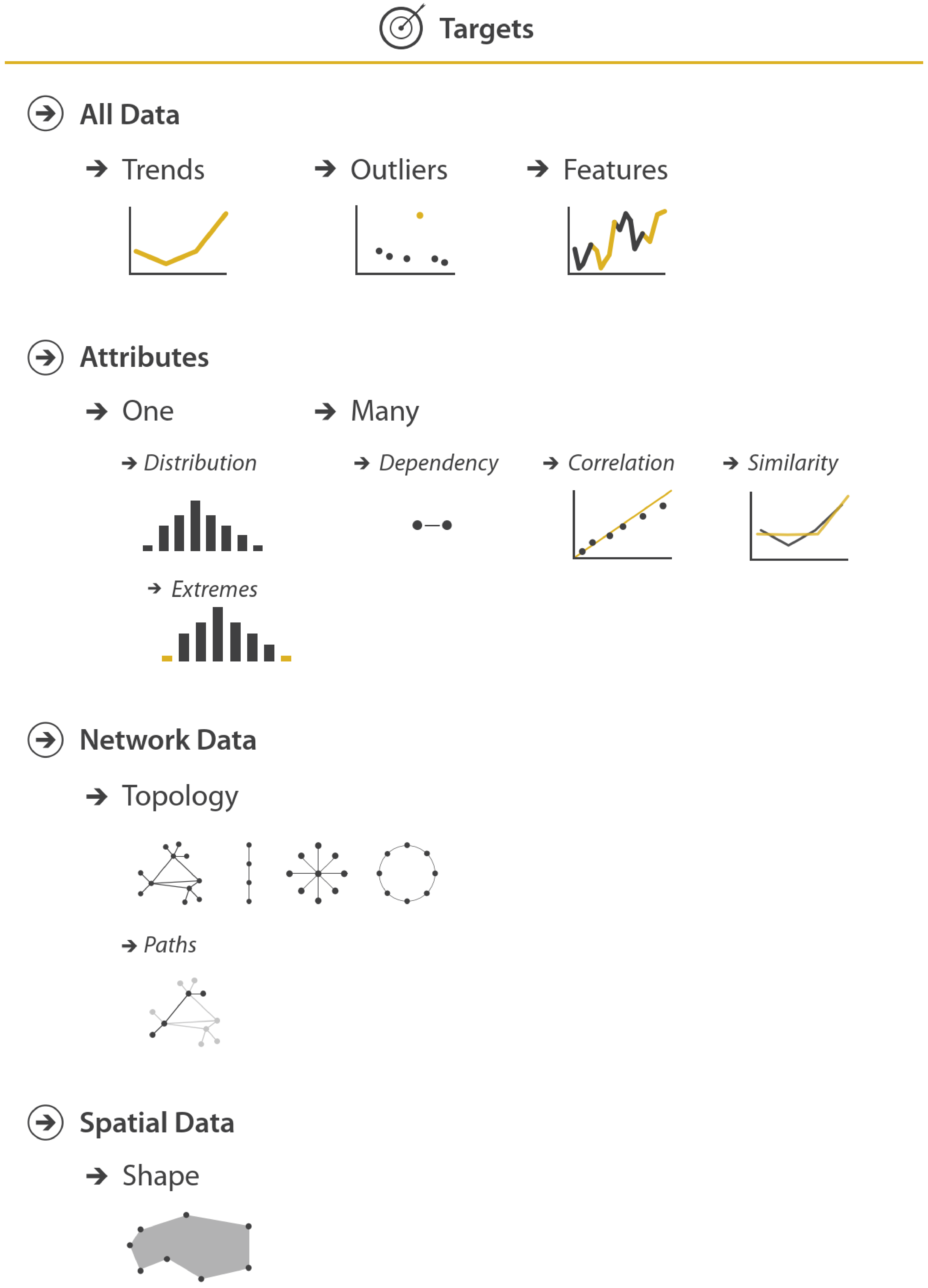
Summarize

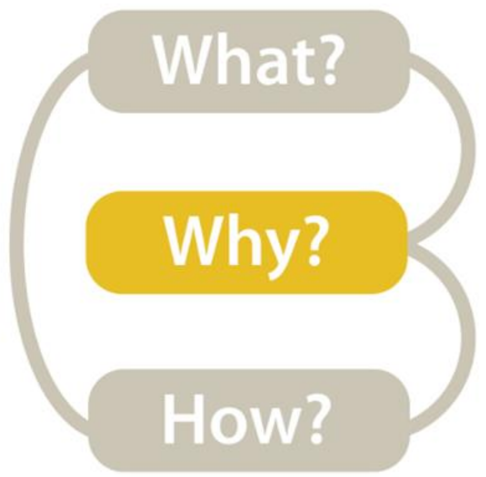


all targets

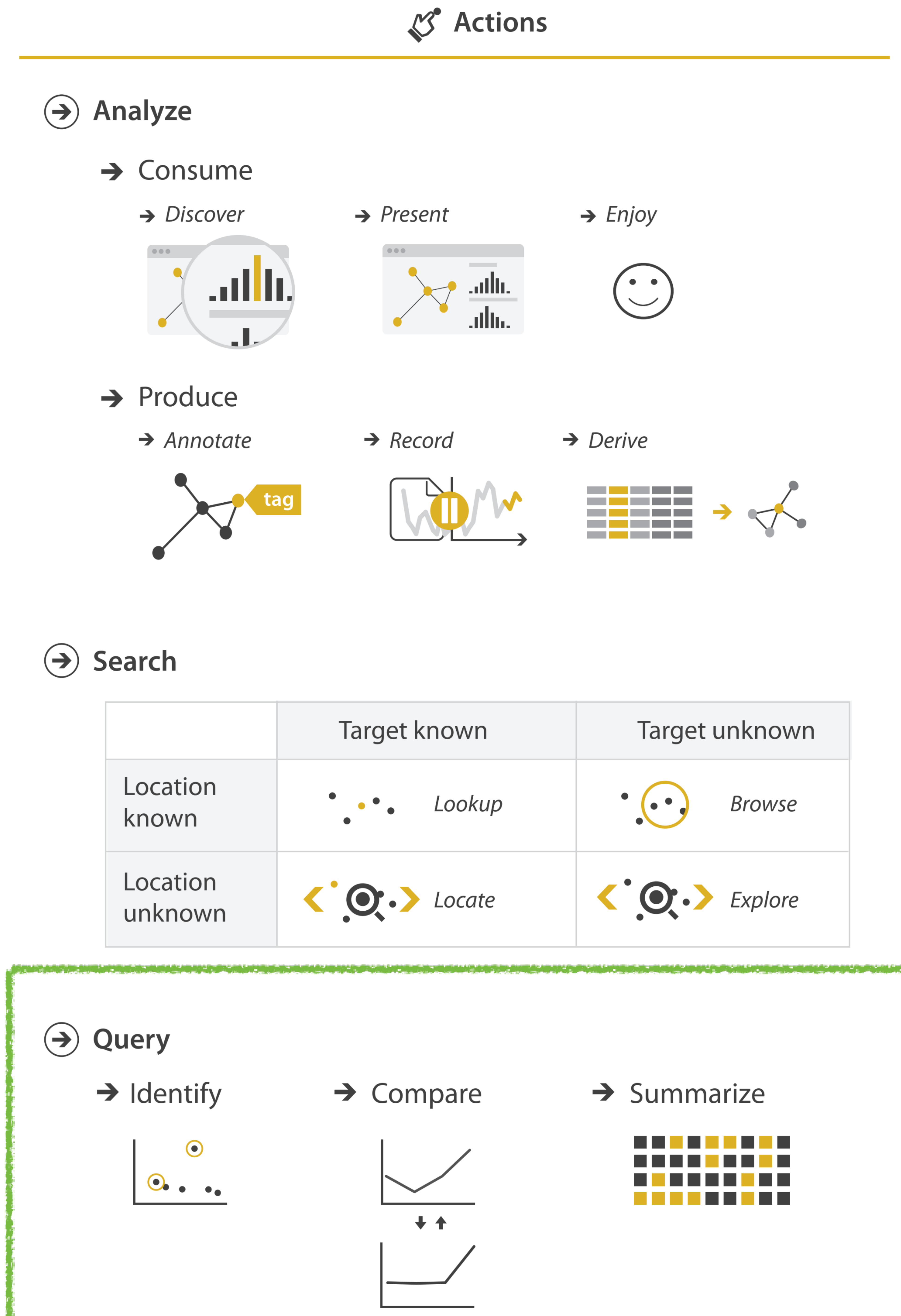


TARGETS are aspects of the data interest that are interest to the user.





ACTIONS define user goals.



High-level

Mid-level

Low-level

Lots of other task taxonomies...!

Analytic Task Taxonomy *Low-level*

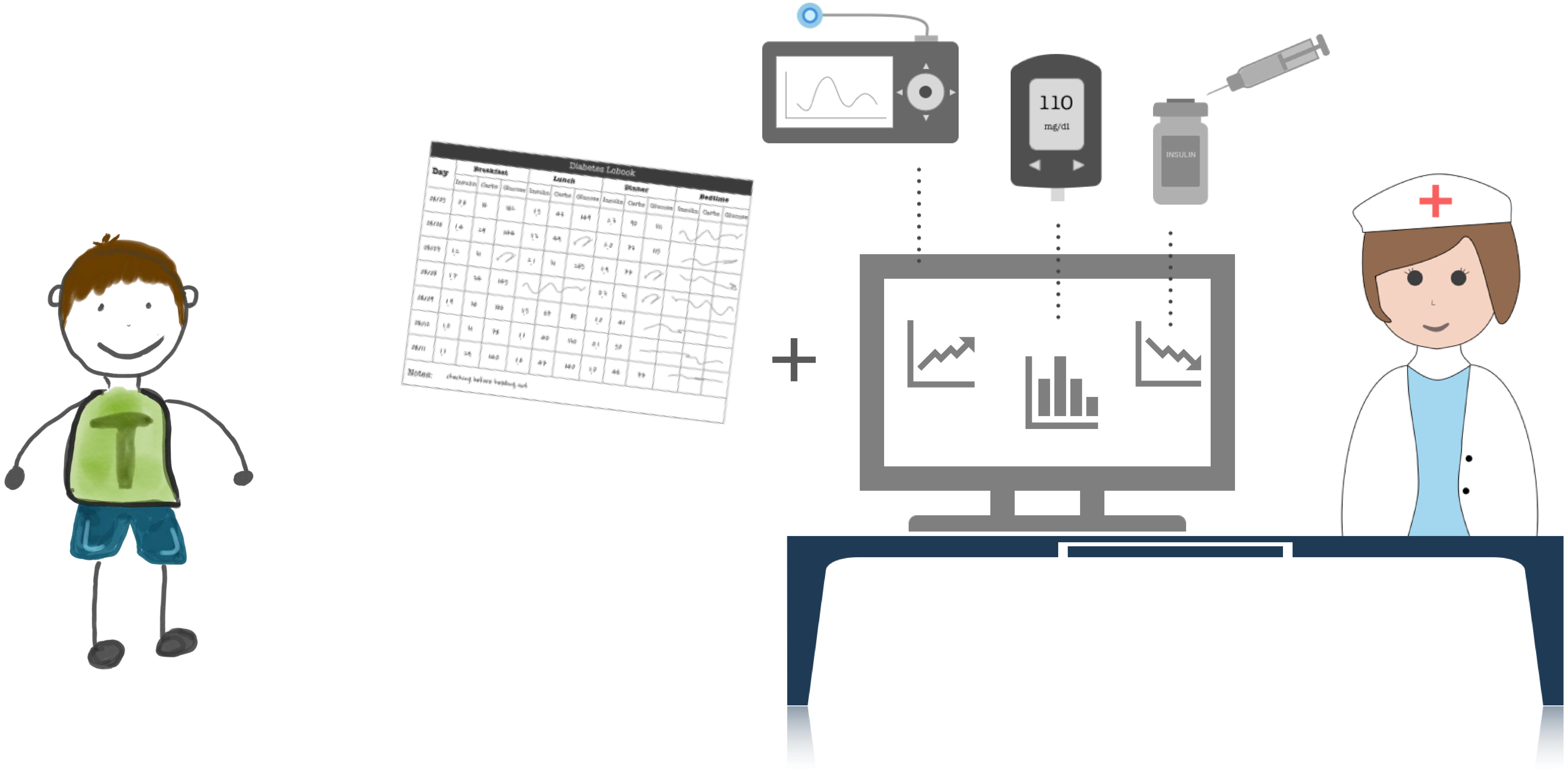
Retrieve Value	*How long is the movie Gone with the Wind?*
Filter	*What comedies have won awards?*
Compute Derived Value	*How many awards have MGM studio won in total?*
Find Extremum	*What director/film has won the most awards?*
Sort	*Rank movies by most number of awards.*
Determine Range	*What is the range of film lengths?*
Characterize Distribution	*What is the age distribution of actors?*
Find Anomalies	*Are there exceptions to the relationship between number of awards won and total movies made by an actor?*
Cluster	*Is there a cluster of typical film lengths?*
Correlate	*Is there a trend of increasing film length over the years?*

Amar et al., 2005

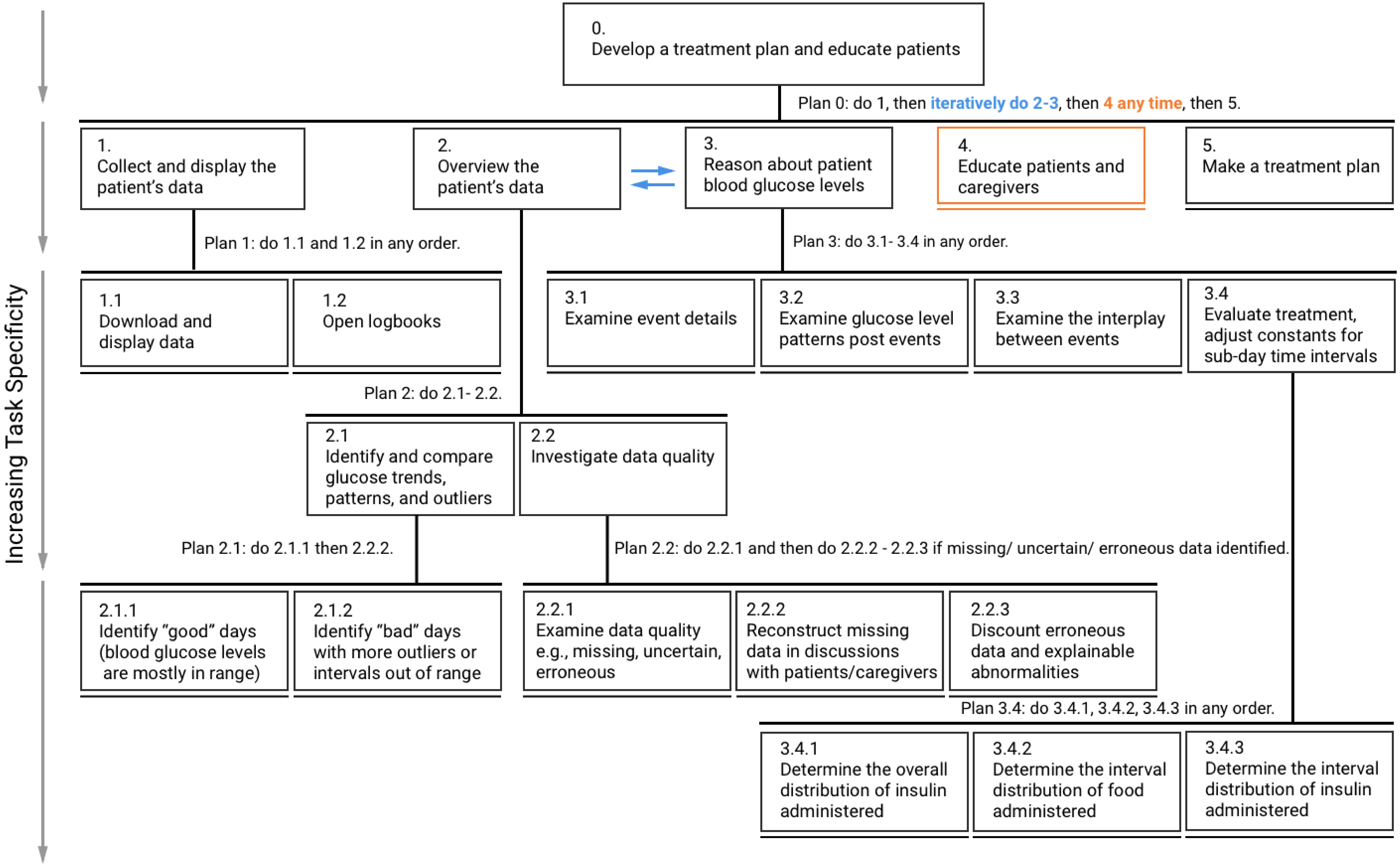
An Example of Task Analysis → Visualization Design

[Zhang et al., 2018]

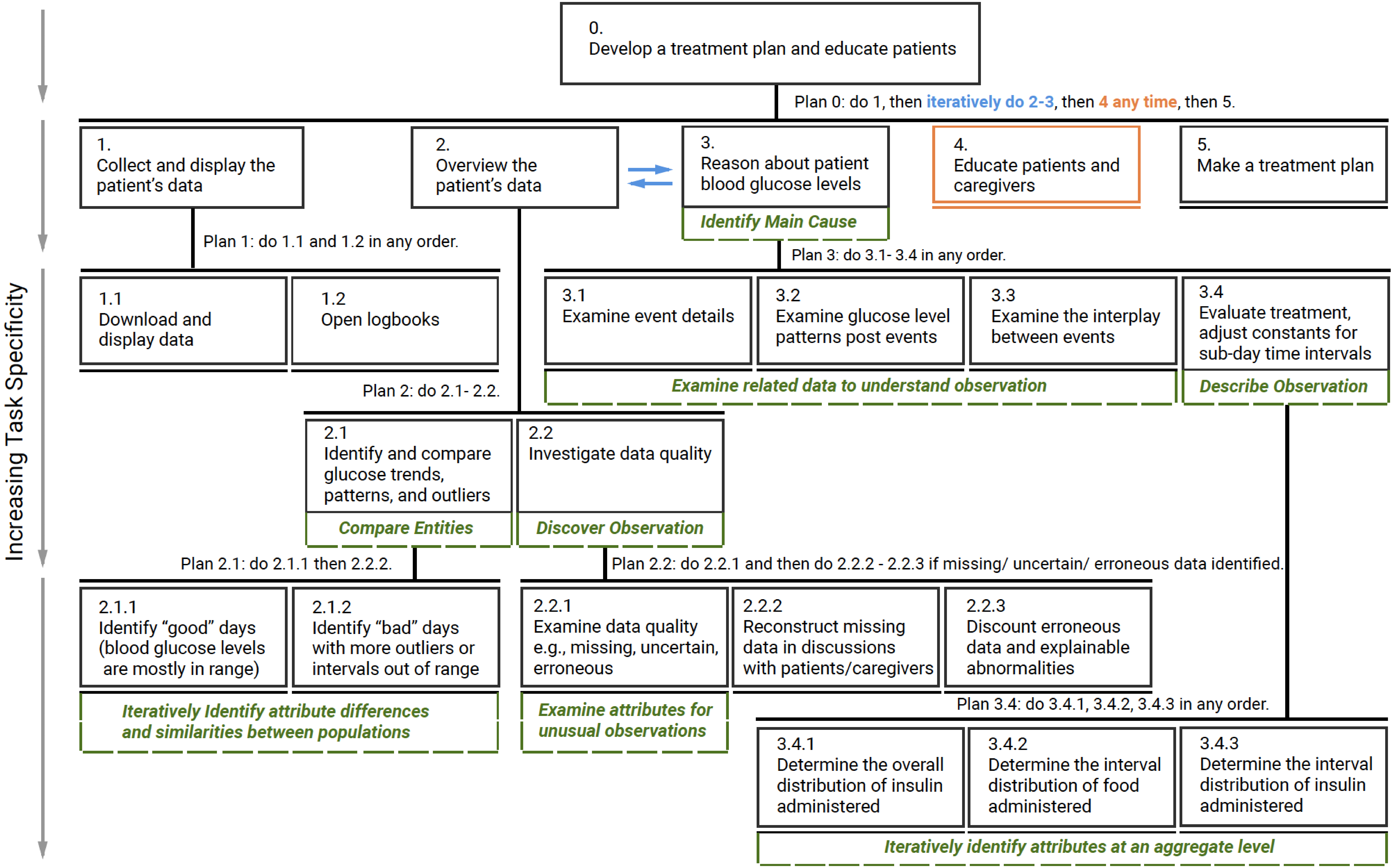
During a type 1 diabetes clinical visit with a Certified Diabetes Educator…



[Zhang et al., 2018](#)



[Zhang et al., 2018](#)



Zhang et al., 2018

Design Requirements

- DR1. Composite Visualization of **Integrated** Data
- DR2. Visualization of **Folded Temporal** Data
- DR3. **Align and Scale** Temporal Data
- DR4. **Summary** Statistics

Hierarchical Task Analysis | Task Abstraction | **Design**

14-Day Overview

Detail View

Summary Statistics Panel

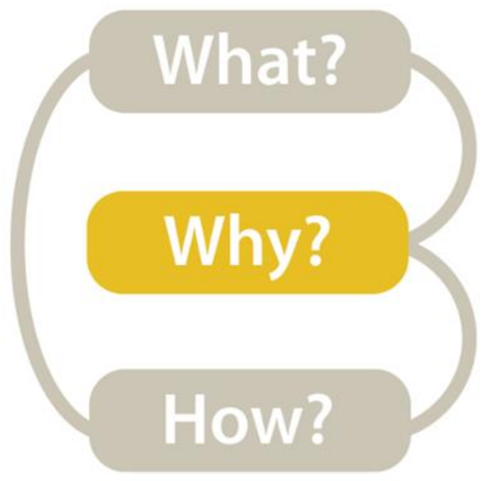
Zhang et al., 2018

In-class exercise:
Mock Interview, Task Analysis

Interview Advice

- Have a designated note-taker and designated leader
- Be prepared. (Have some questions prepared in advance.)
- Start slow, safe, and personal.
- Coax, don’t hammer.
- Make some questions open ended.
- Ask what you don’t know.
- Let the interviewees wander a bit–but be careful.
- Listen, really listen.
- For software, look for “work arounds” and hacks.
- Make sure to write down your thoughts and impressions immediately after the interview.
- You are the visualization expert – don’t ask them what vis they want, don’t think too early about what vis to build.

www.forbes.com/sites/shelisrael/2012/04/14/8-tips-on-conducting-great-interviews/+&cd=3&hl=en&ct=clnk&gl=us



Task Analysis

Visualization for Public Transit Development

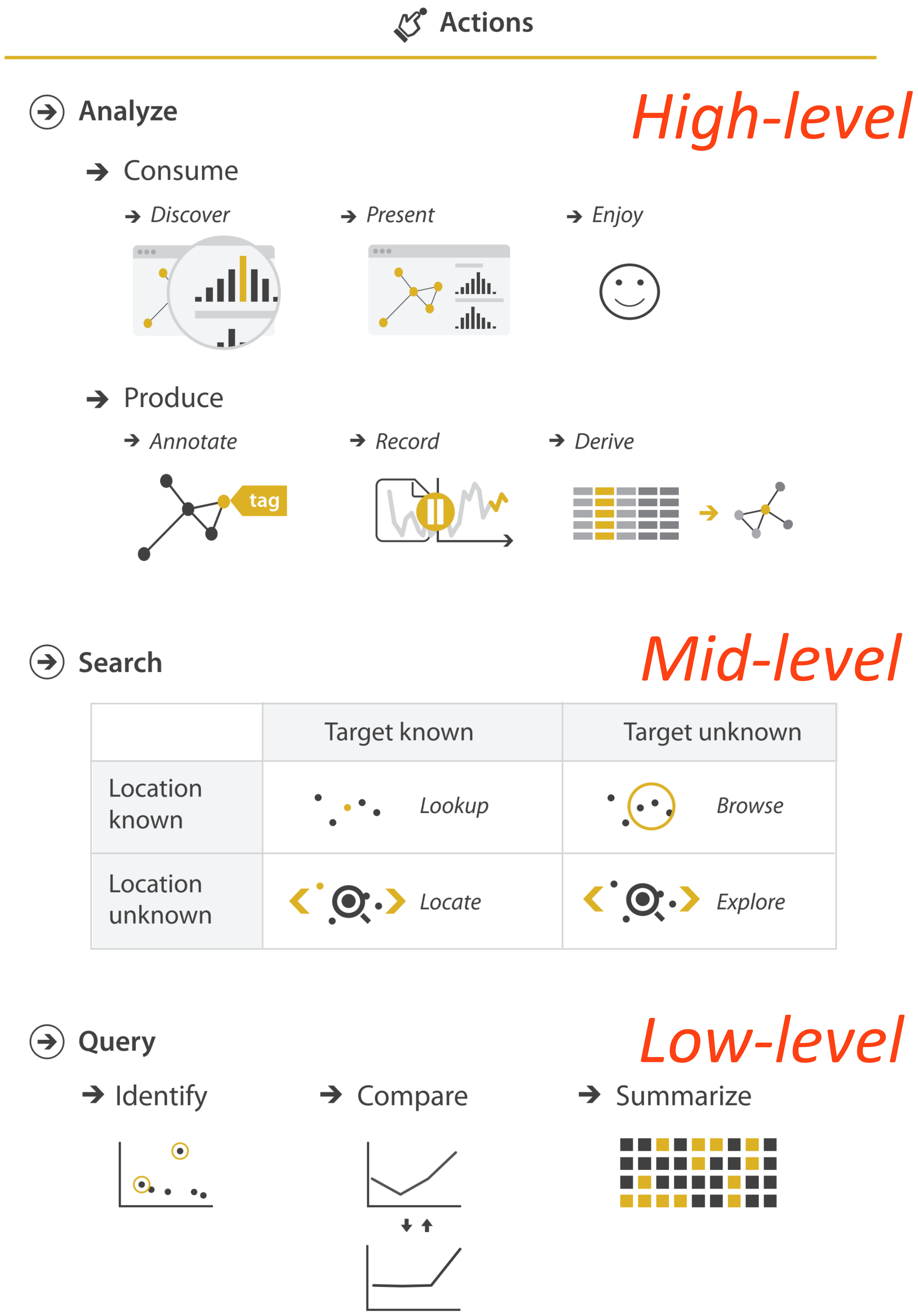
15m

INSTRUCTIONS:

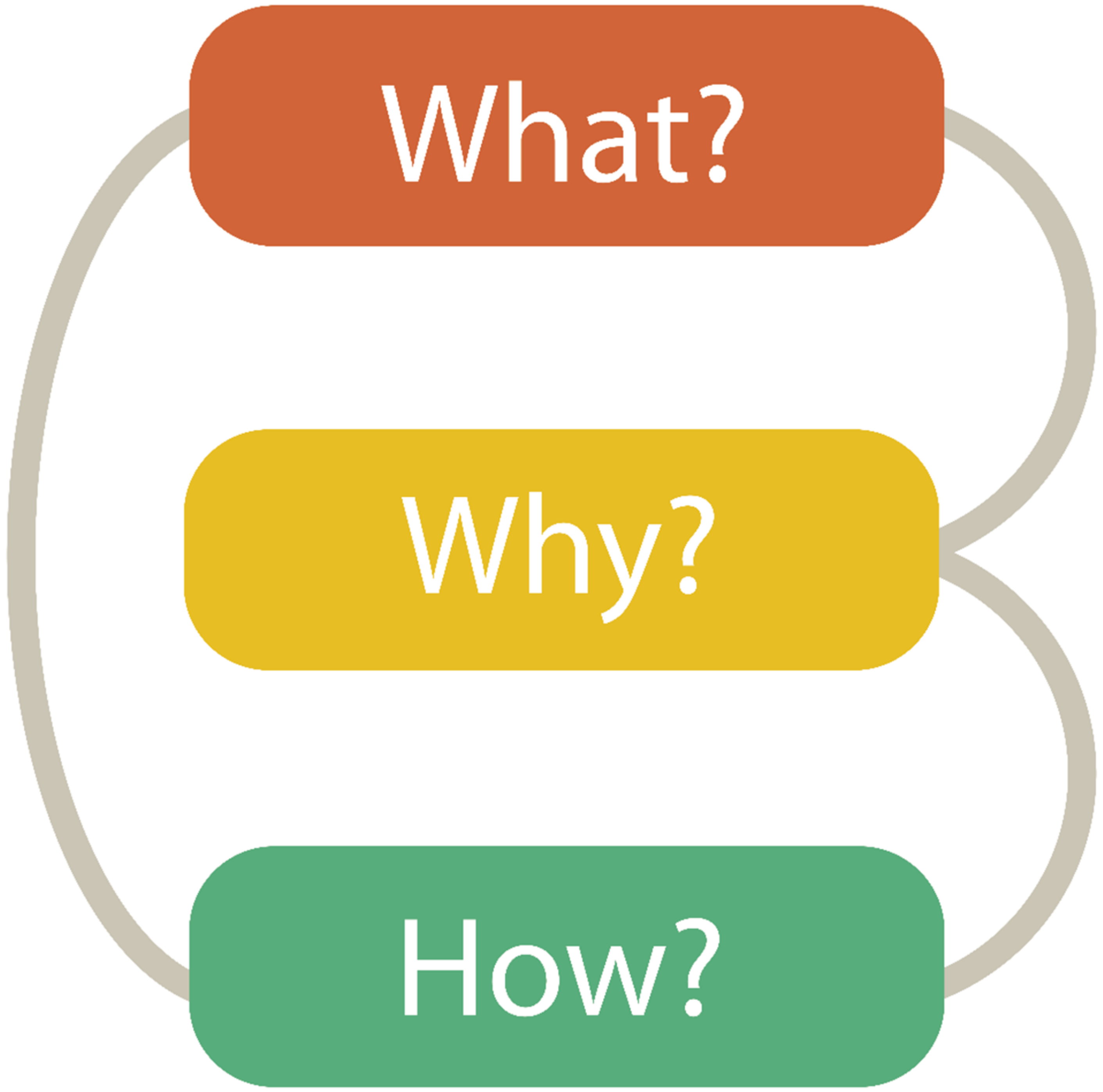
- Break-out into groups of ~3 people.
- Pretend you are transportation engineers, e.g., for the MBTA, City of Boston.
- Discuss the “domain tasks” and classify the tasks.
- Save your notes for a later exercise!!!

Low-level

Retrieve Value	*How long is the movie Gone with the Wind?*
Filter	*What comedies have won awards?*
Compute Derived Value	*How many awards have MGM studio won in total?*
Find Extremum	*What director/film has won the most awards?*
Sort	*Rank movies by most number of awards.*
Determine Range	*What is the range of film lengths?*
Characterize Distribution	*What is the age distribution of actors?*
Find Anomalies	*Are there exceptions to the relationship between number of awards won and total movies made by an actor?*
Cluster	*Is there a cluster of typical film lengths?*
Correlate	*Is there a trend of increasing film length over the years?*



Analysis



Data Abstraction

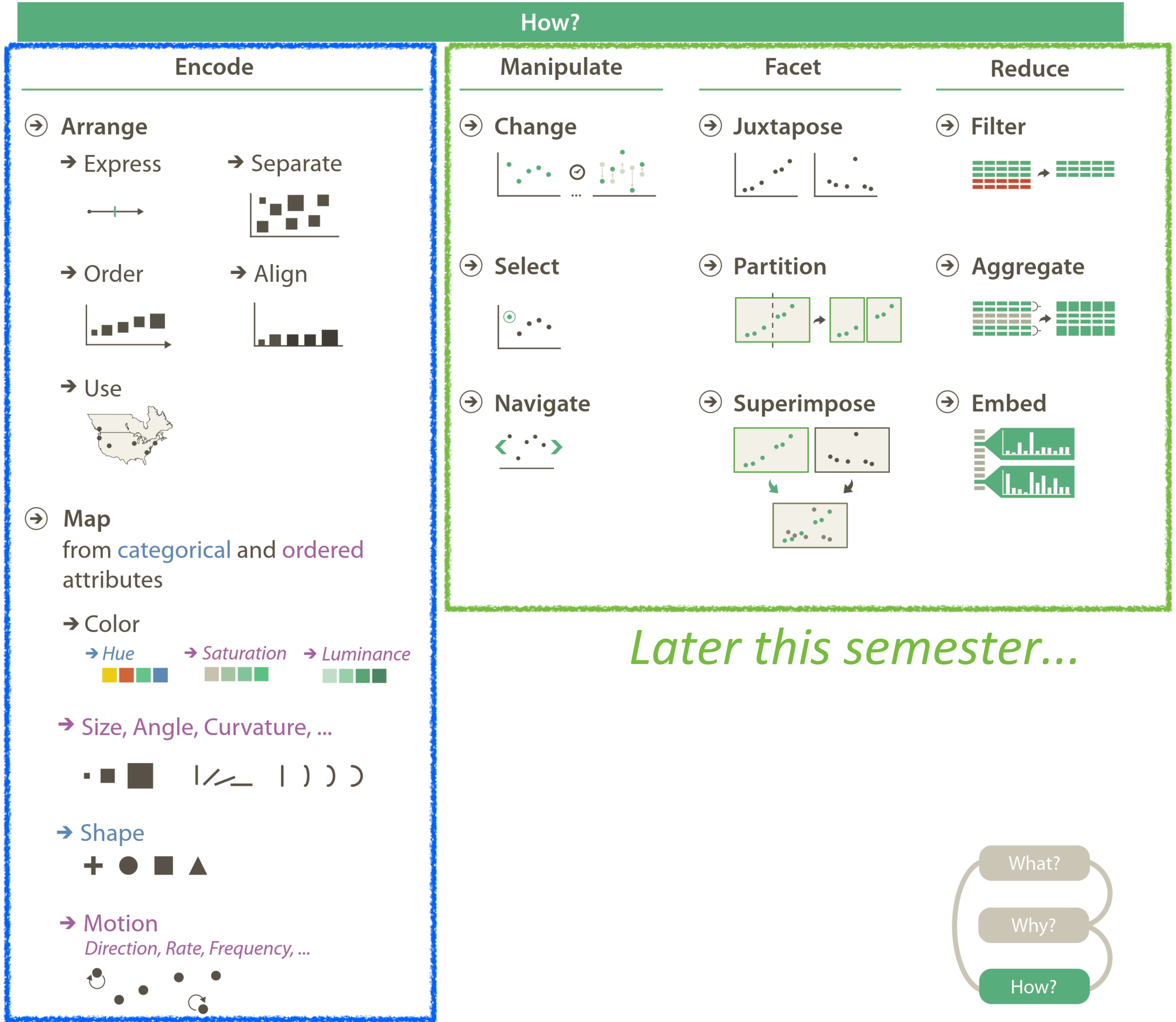
Task Abstraction

Visual Encoding

Goals for Today

- Learn about visual encodings, esp. arranging tables
- Learn how to pick appropriate visual representations based on attribute type and perceptual properties

Visual Encoding



Now…

Later this semester…

Visualization Building Blocks

Marks:

Marks as Items/Nodes

- Points

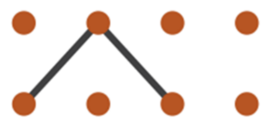
- Lines

- Areas

Marks as Links

- Containment

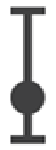
- Connection



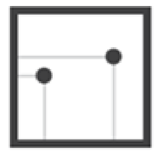
Channels:

- Position
  - Horizontal

  - Vertical



  - Both



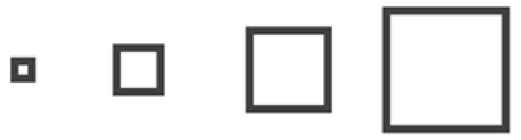
- Color

- Shape

- Tilt

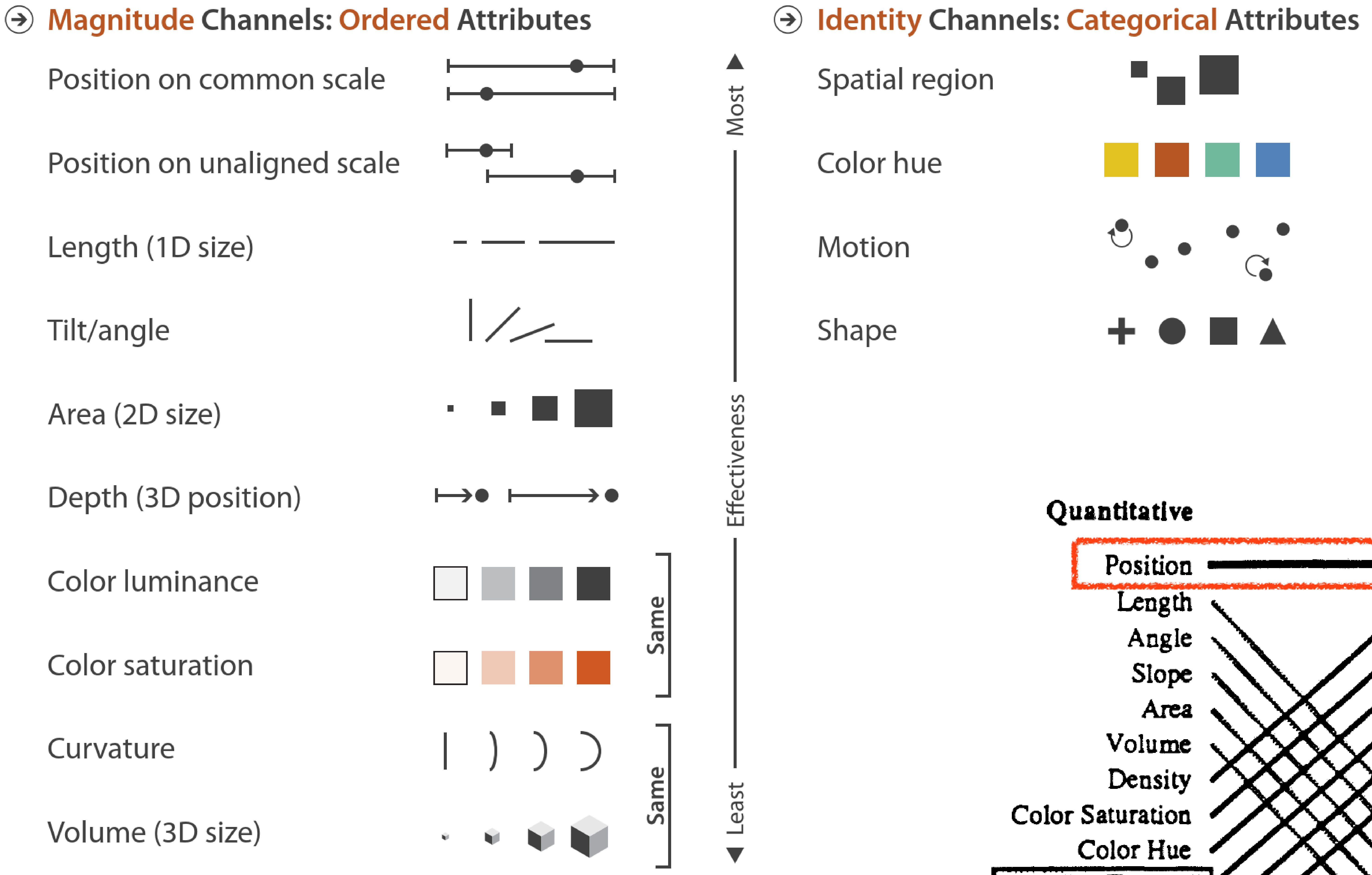
- Size
  - Length

  - Area



  - Volume

Channels: Expressiveness Types and Effectiveness Ranks



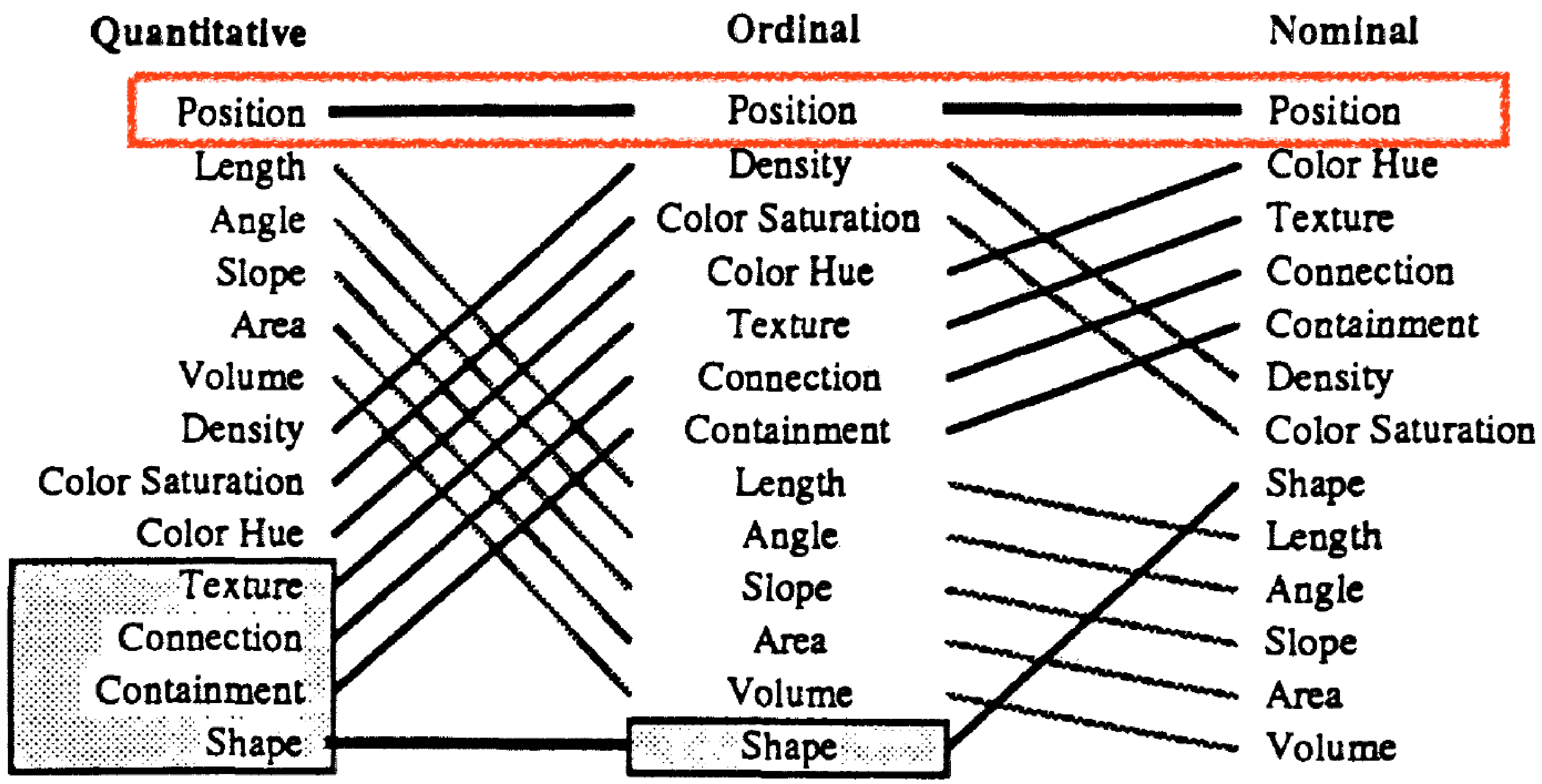


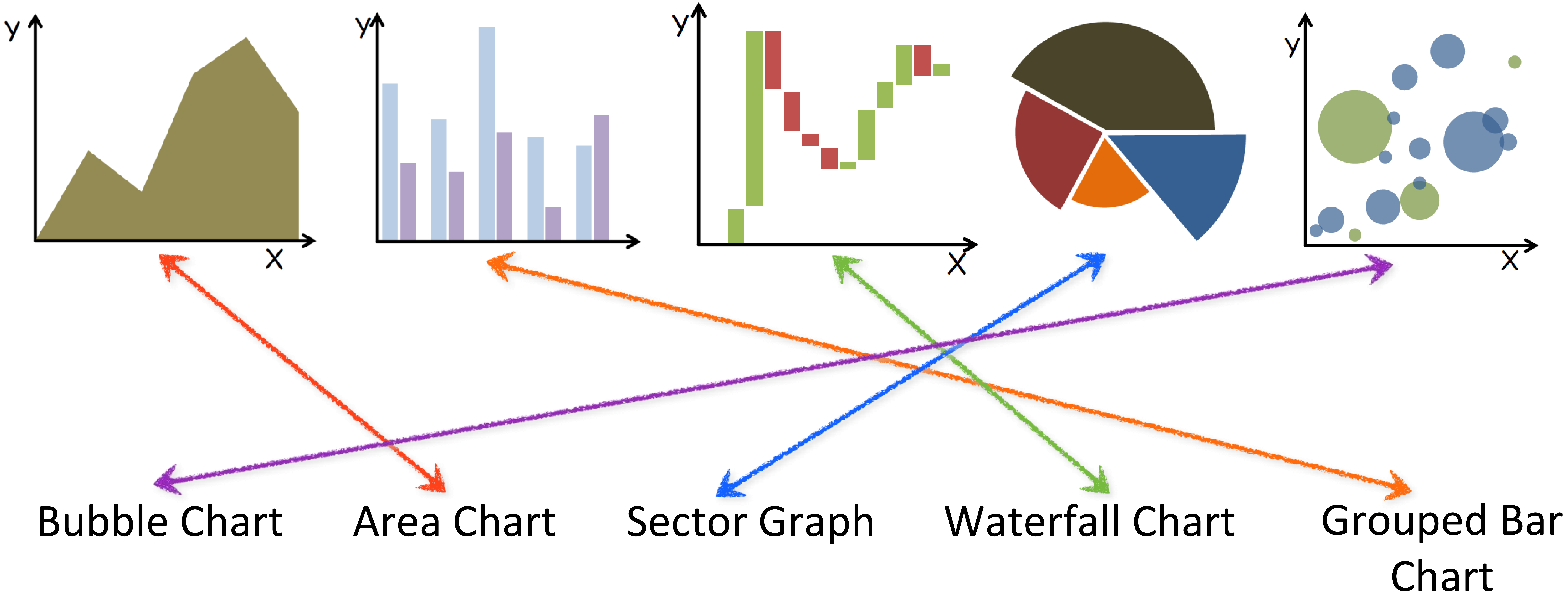
Figure 15: Ranking of Perceptual Tasks. *The tasks shown in the gray boxes are not relevant to that type of data.*

[Mackinlay (1986)]

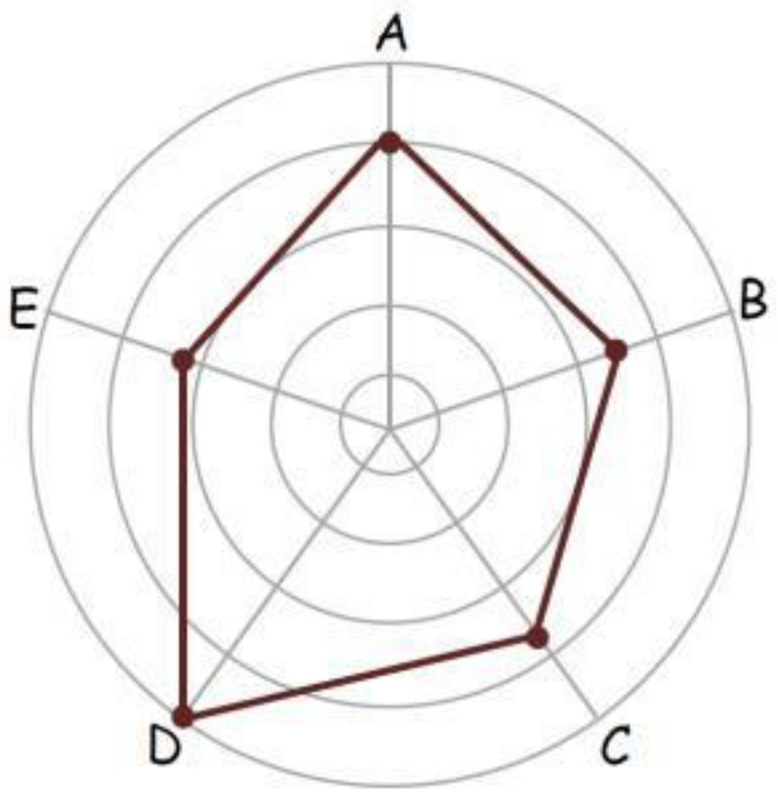
Munzner's VAD

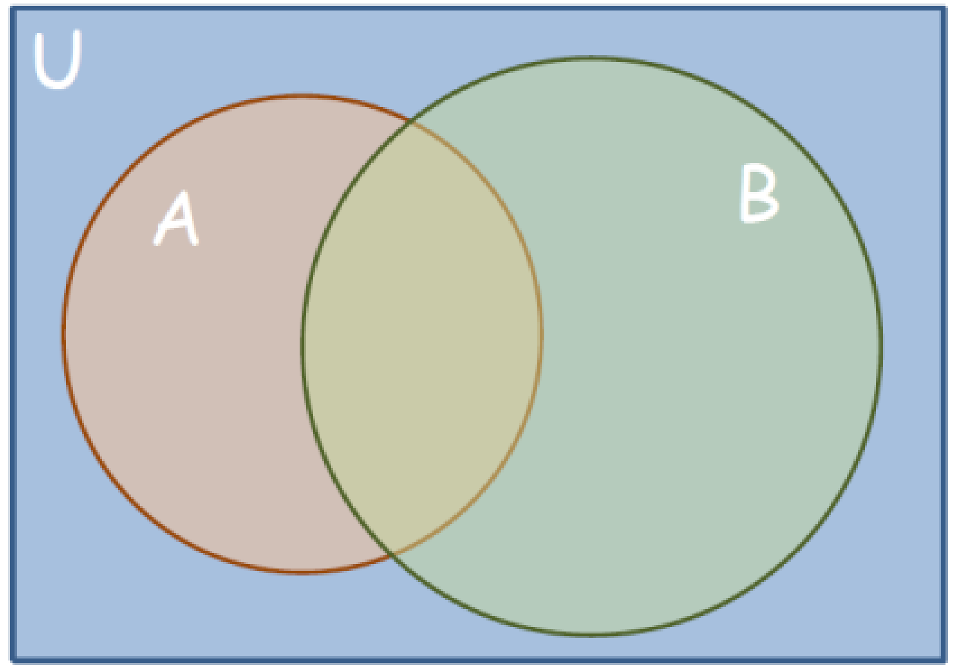
In-class exercise: Encodings Worksheet

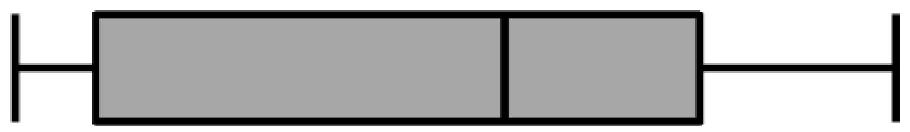
Encoding Match-up

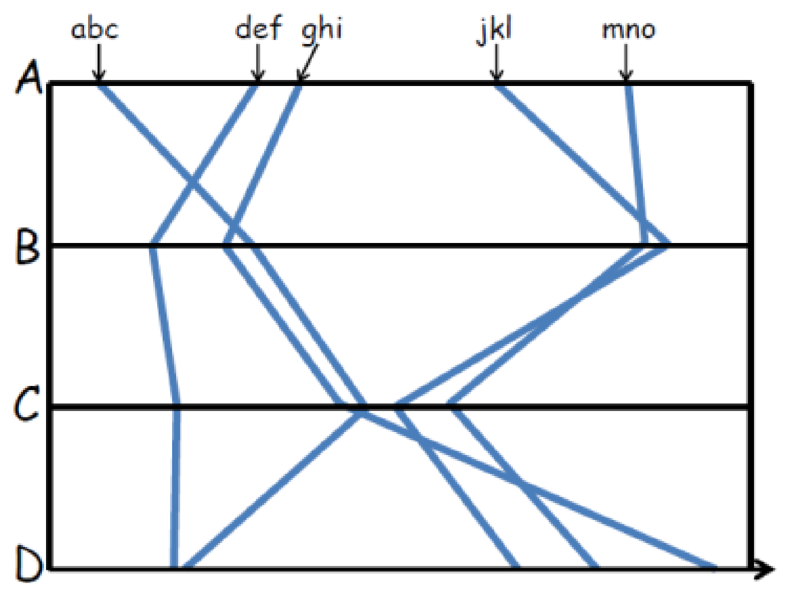


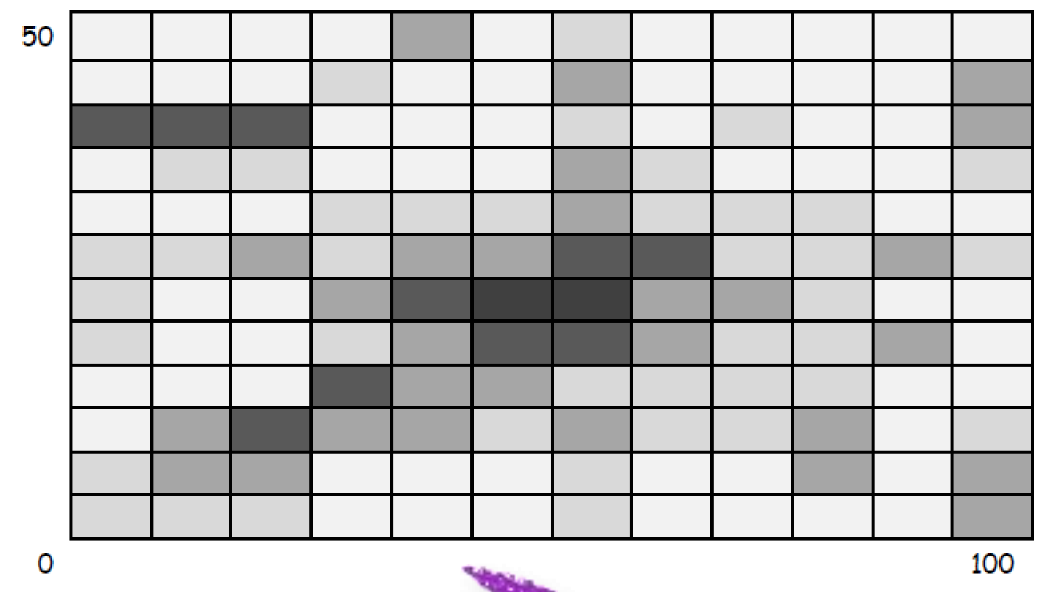
Encoding Match-up



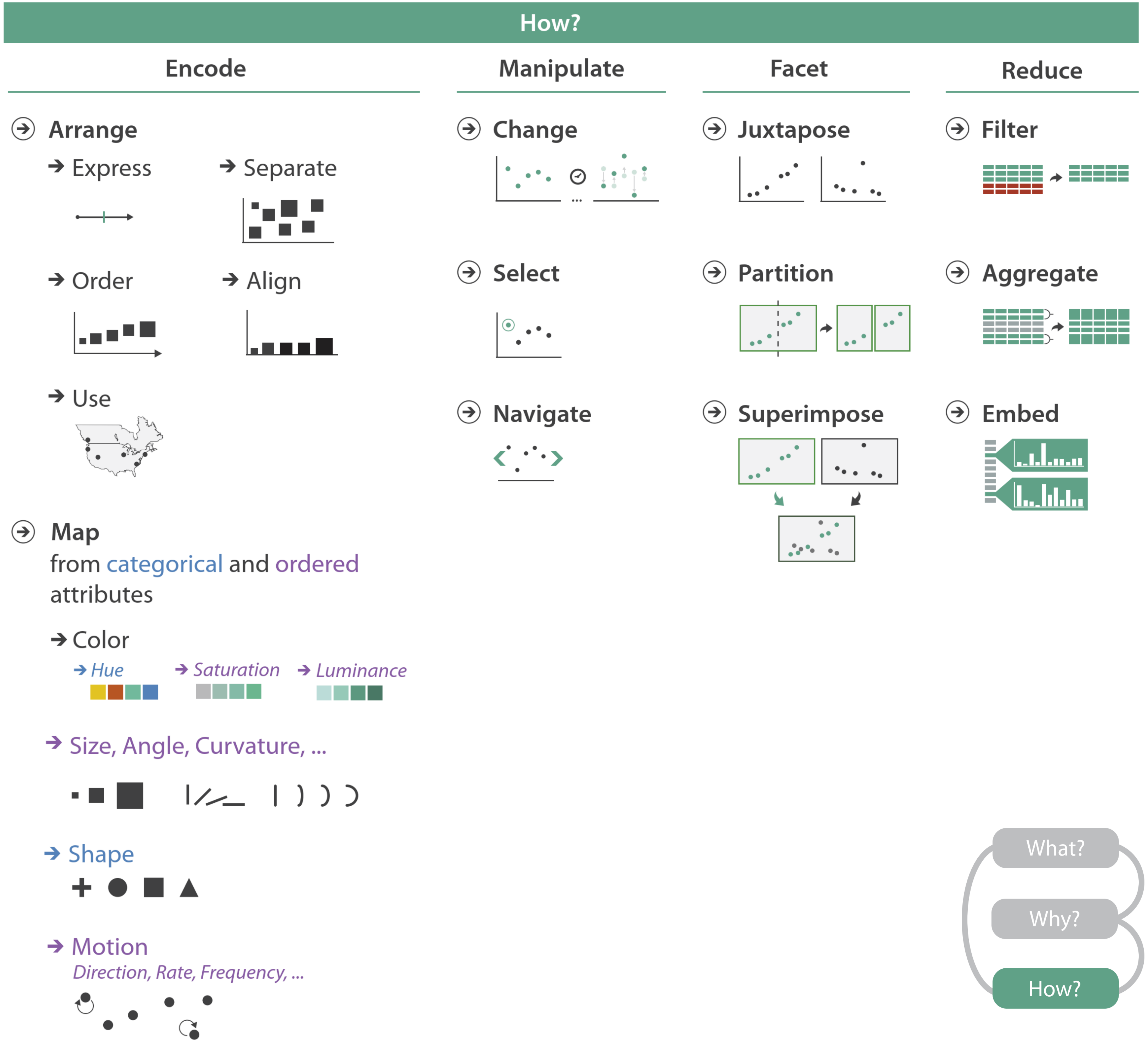








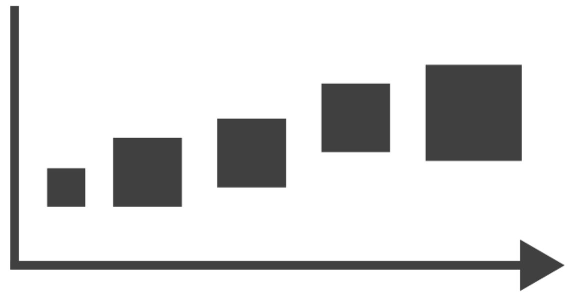
Venn Diagram   Heat Map   Star Plot   Box & Whisker Plot   Parallel Coordinates


How?
Encode
Manipulate
Facet
Reduce
Arrange
Express
Separate
Order
Align
Use
Map
from categorical and ordered attributes
Color
Hue
Saturation
Luminance
Size, Angle, Curvature, ...
Shape
Motion
Direction, Rate, Frequency, ...
Change
Select
Navigate
Juxtapose
Partition
Superimpose
Filter
Aggregate
Embed
What?
Why?
How?

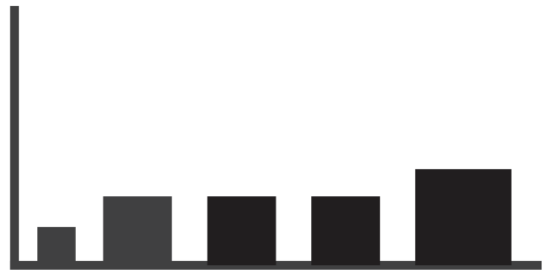
Arrange Tables

- **Separate, Order, Align Regions**
  - Separate

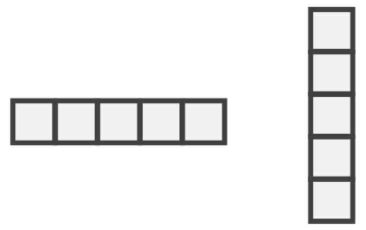
  - Order



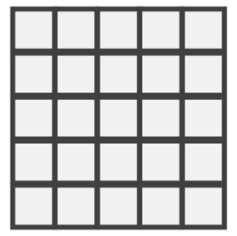
  - Align



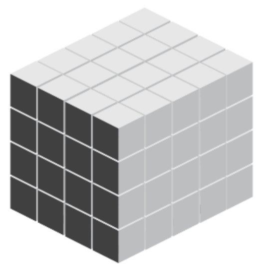
    - *1 Key*
      List



    - *2 Keys*
      Matrix



    - *3 Keys*
      Volume



    - *Many Keys*
      Recursive Subdivision

Key: an independent attribute that can be used as a unique index (Tableau Dimension)

Categorical or Ordinal

Value: a dependent attribute (i.e., cell in a table) (Tableau Measures)

Categorical Ordinal, or Quantitative

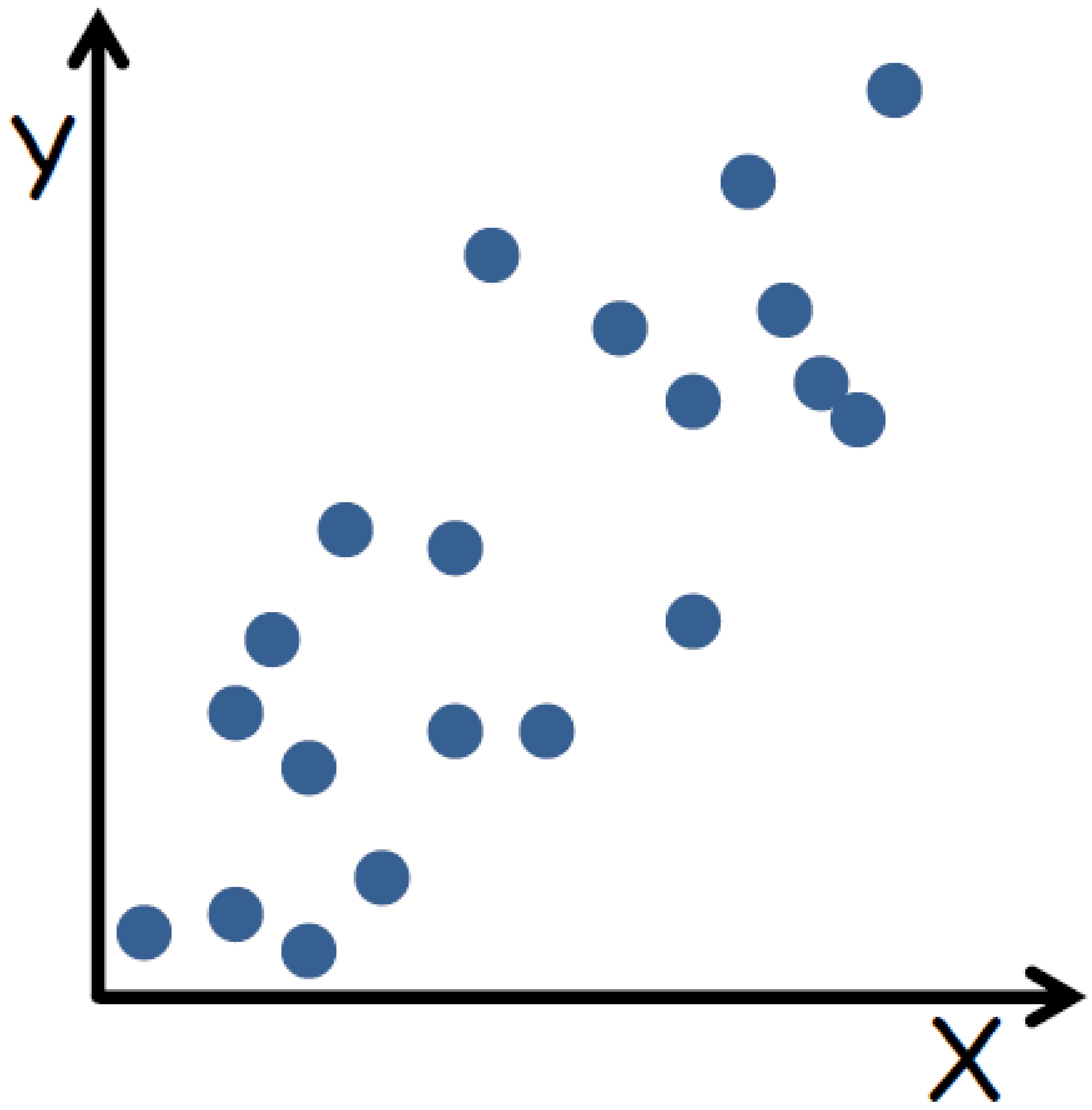
Example Keys

Key

Date	Precipitation	High Temperature
May 1, 2016	0”	60
May 2, 2016	0.3”	62
May 3, 2016	1”	55
May 4, 2016	0”	67

Student	College	HW1 grade (out of 10)
John	COS	9
Jane	Khoury	10
June	Khoury	8
Joe	Khoury	8

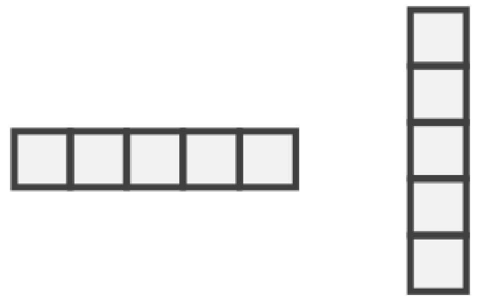
Arrange Tables - no key

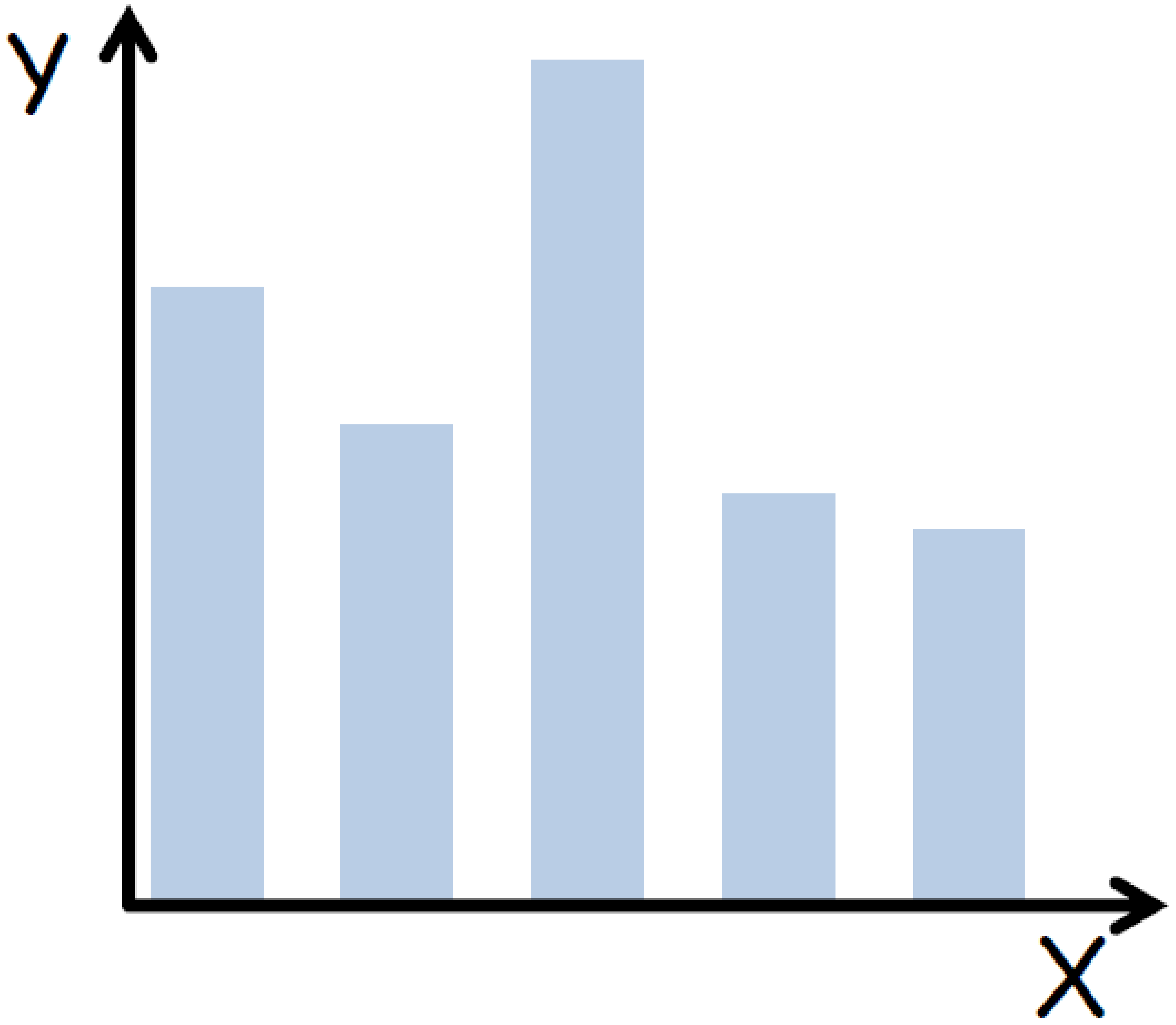


SCATTER PLOT

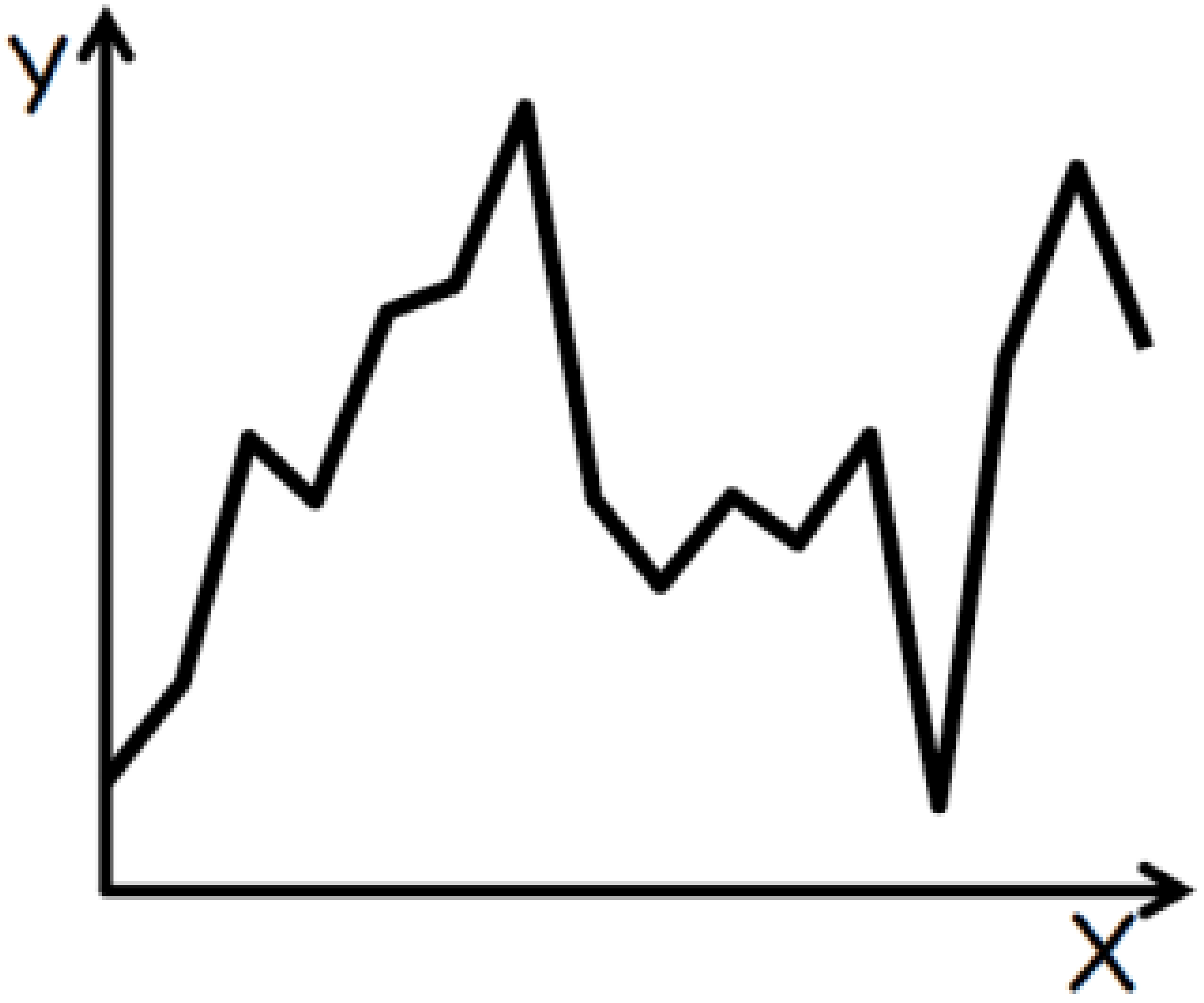
Arrange Tables - one key

➔ *1 Key*
List





BAR CHART

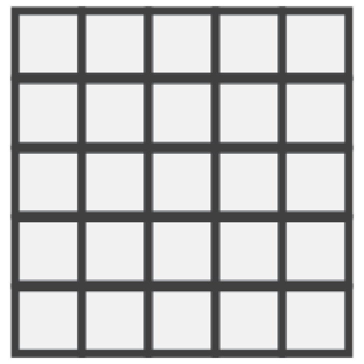


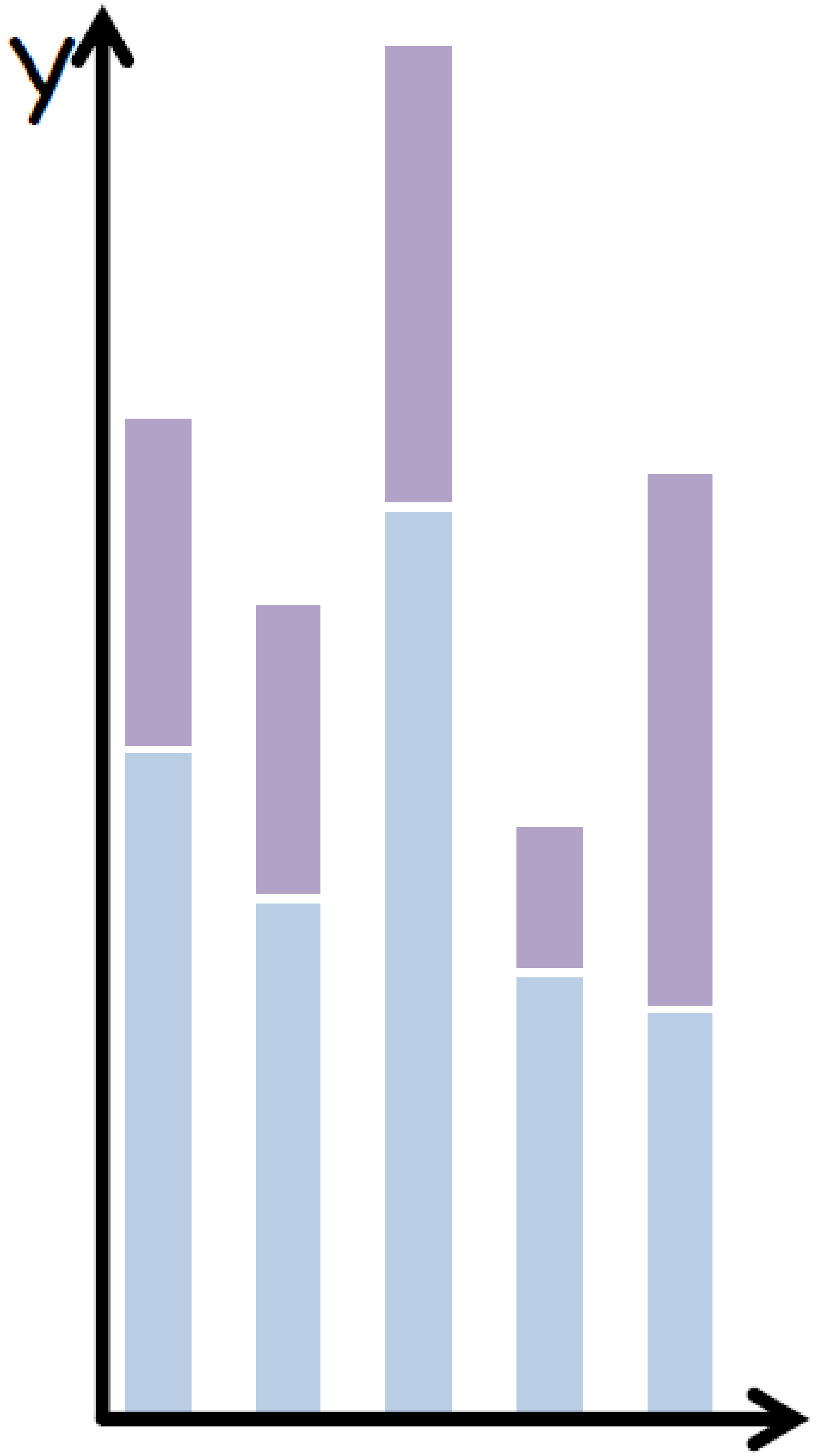
LINE GRAPH

Arrange Tables - two keys

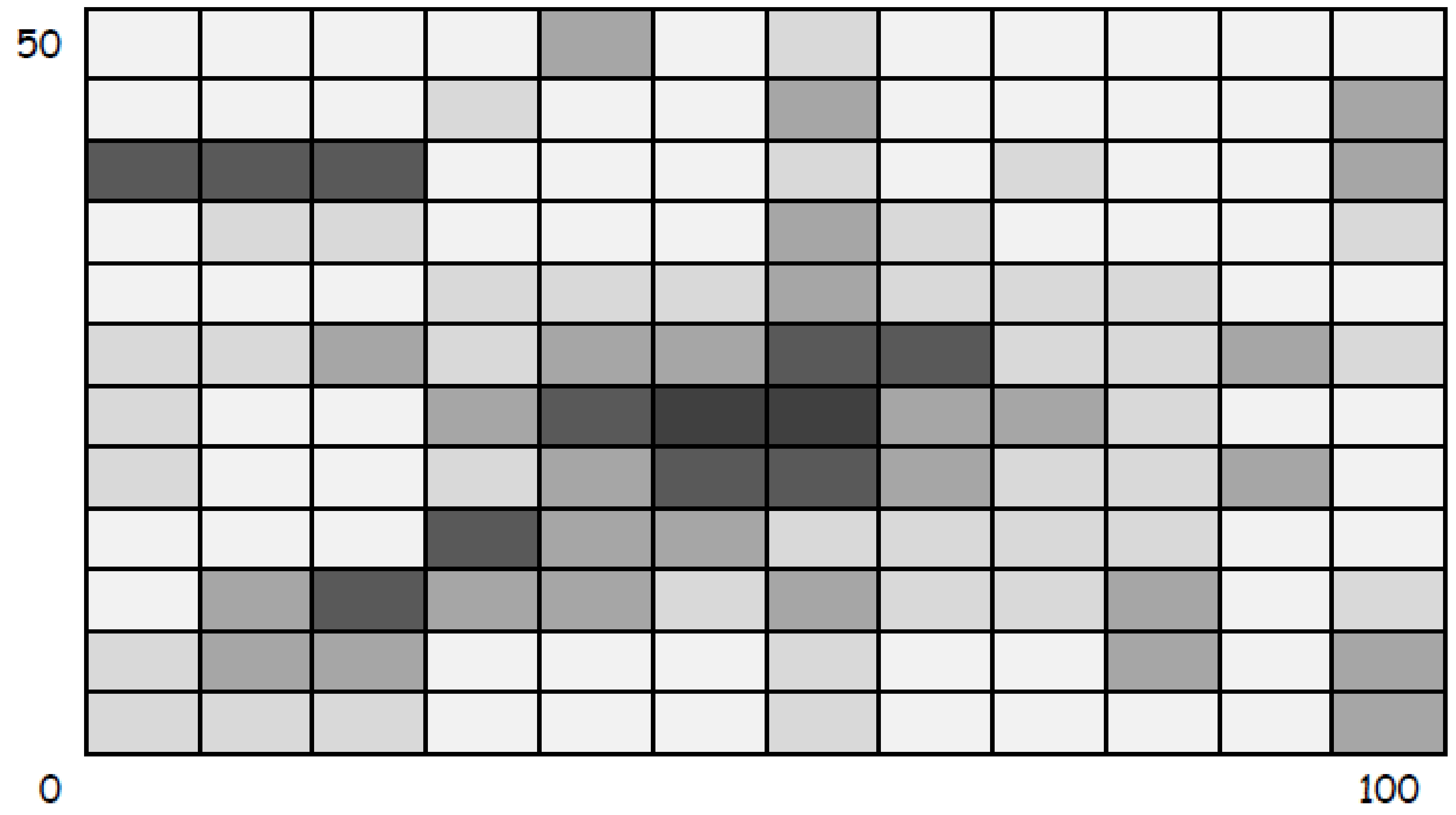
➔ *2 Keys*

Matrix





STACKED BAR CHART

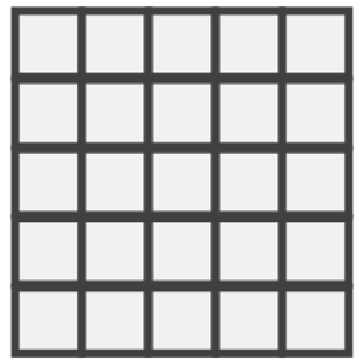


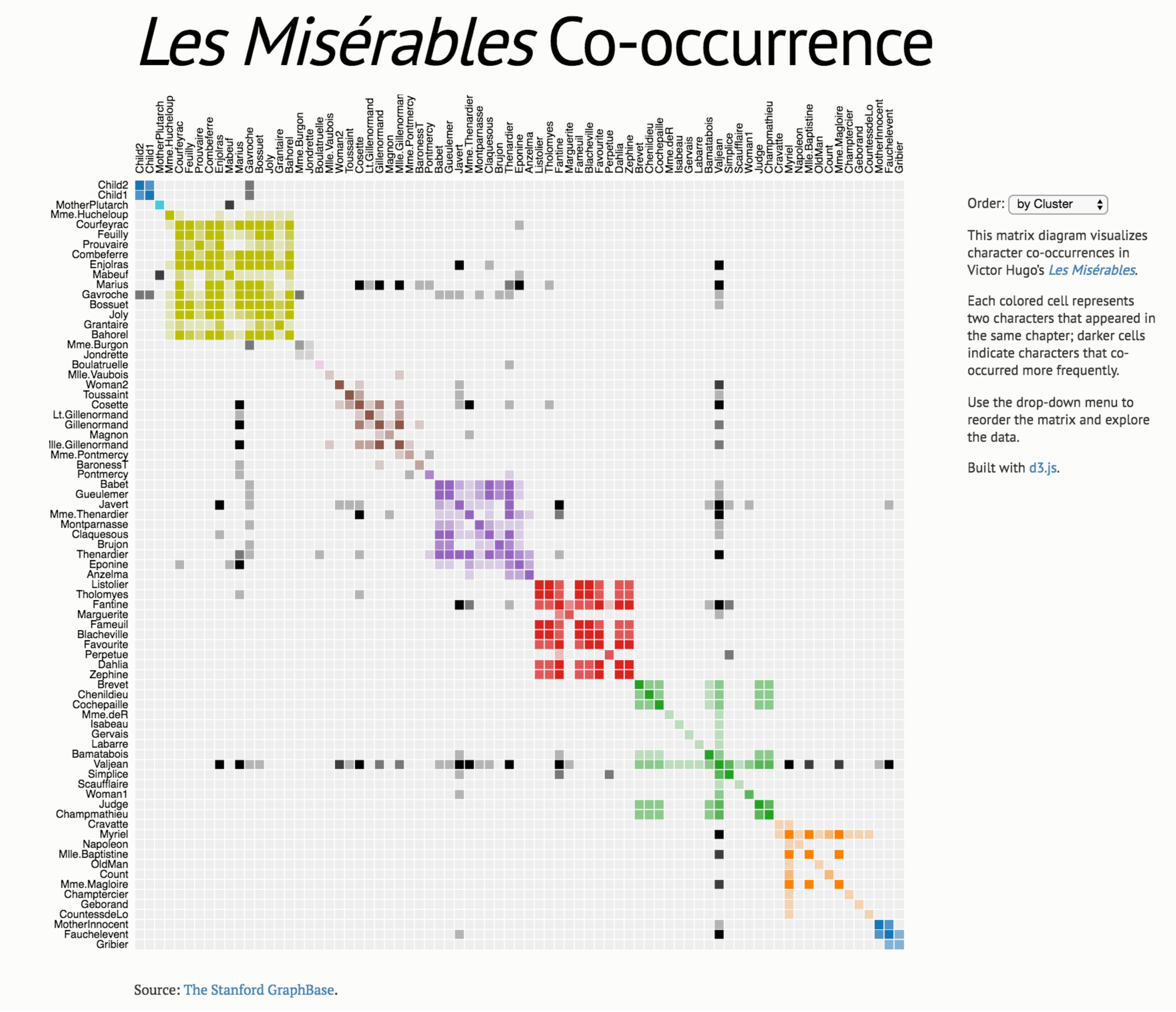
HEATMAP

Arrange Tables - Two Keys (Network)

➔ *2 Keys*

Matrix



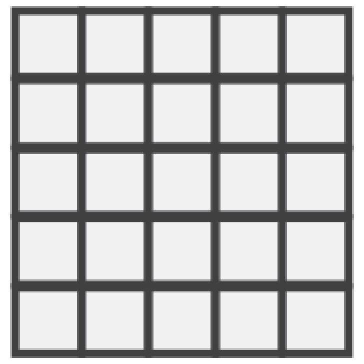


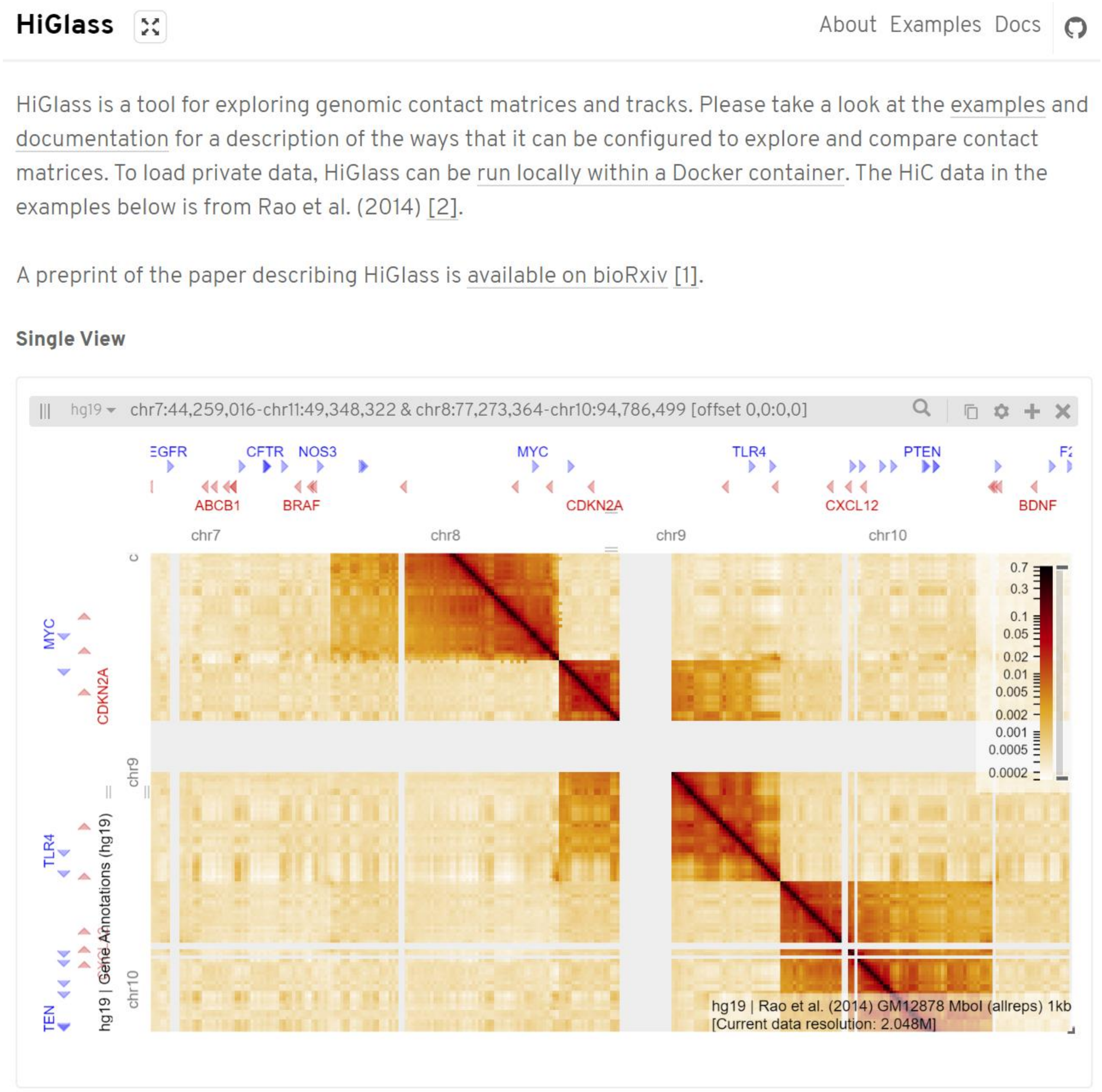
https://bost.ocks.org/mike/miserables/

Arrange Tables - Two Keys (Network)

➔ *2 Keys*

Matrix





http://higlass.io/