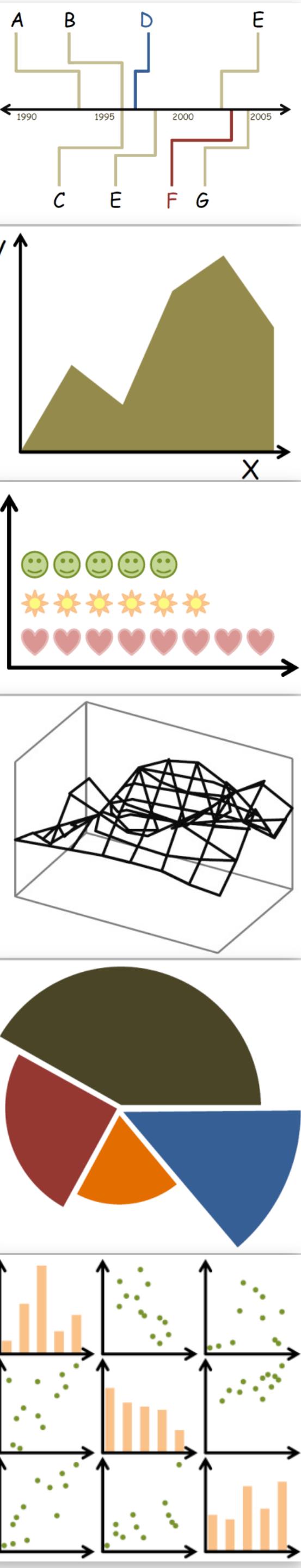


Lecture 4: Task Abstraction

DS 4200
SPRING 2023

Prof. Ab Mosca (*they/them*)
NORTHEASTERN UNIVERSITY



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

Last Class

We:

- Reviewed building blocks of visualization
- Reviewed data abstraction
- Practiced HTML & CSS (ic-02)

Any Questions?

Today

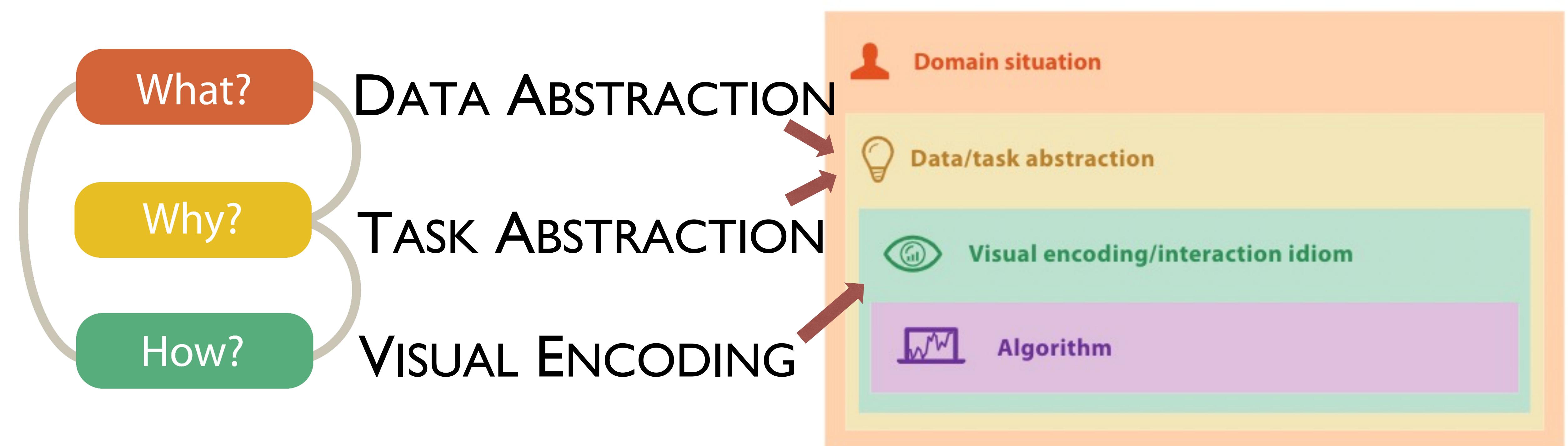
- Task Abstraction
- Practice Interviewing and Task Abstraction (ic-03)

VISUALIZATION DESIGN PROCESS

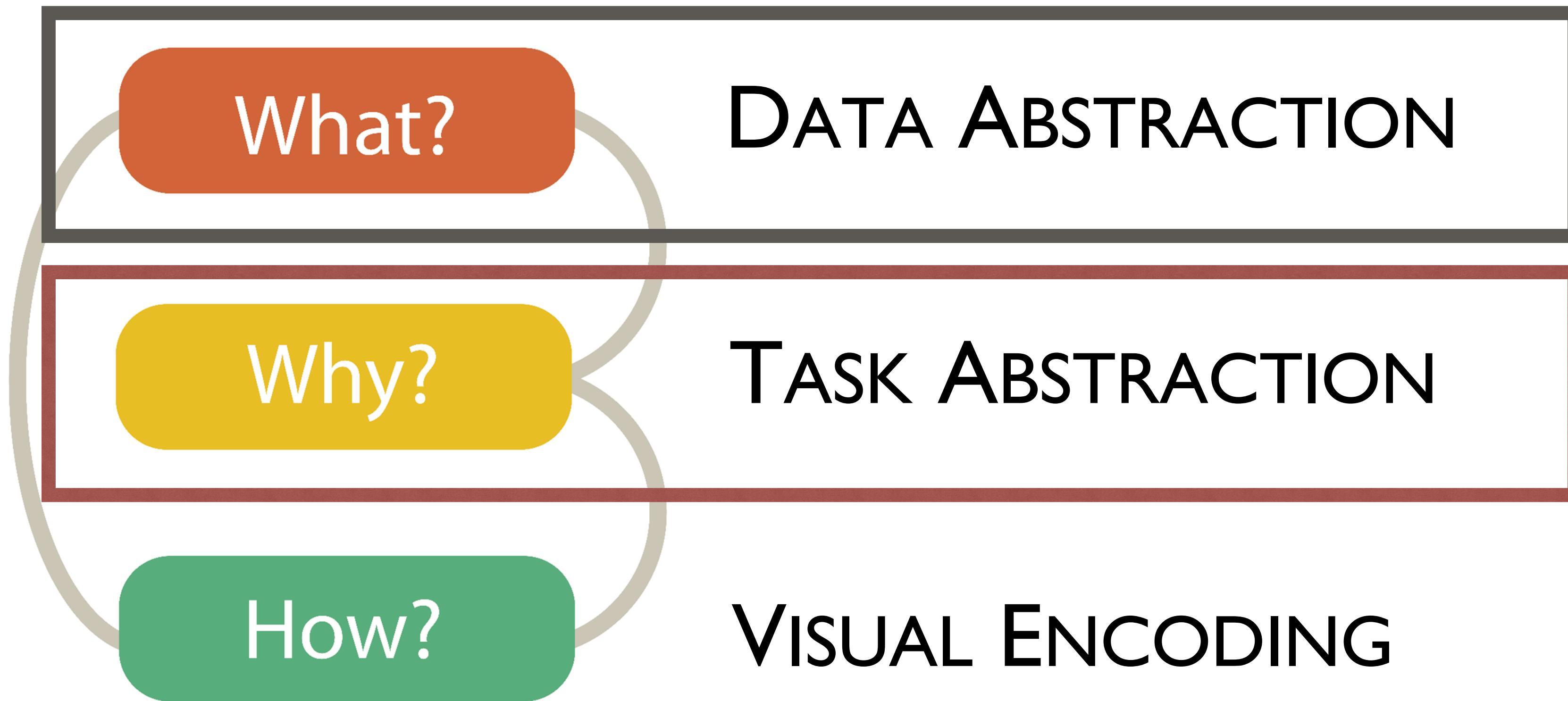
From Munzner's book

Visualization Building Blocks

Munzner's Nested Model



Visualization Building Blocks



TASK ABSTRACTION

From Munzner's book

Task Abstraction

What?

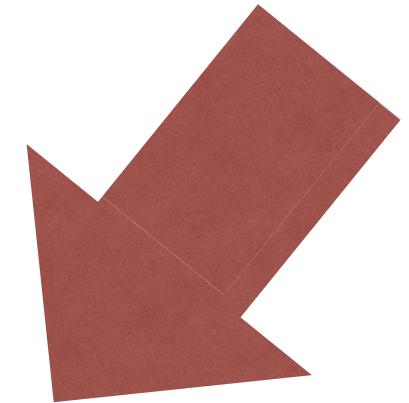
The process of taking specific **domain tasks** and thinking about them as **abstract** (modular!) pieces

Task Abstraction

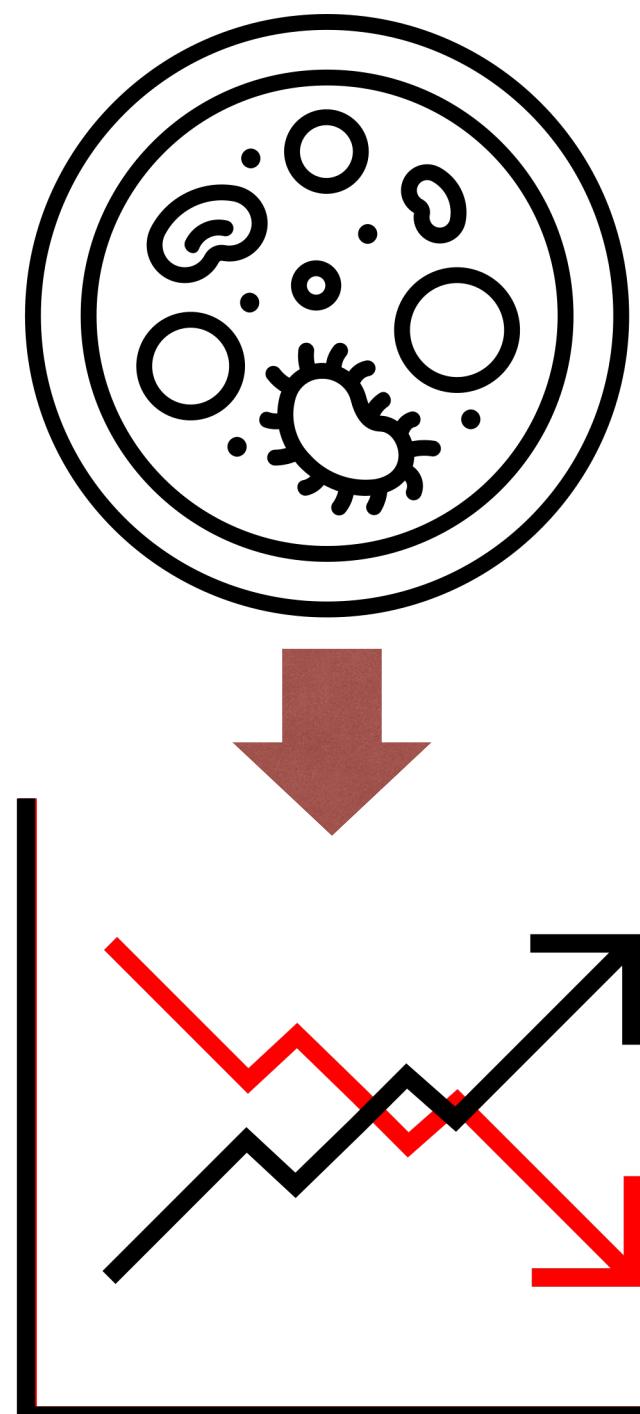
What?

The process of taking specific **domain tasks** and thinking about them as **abstract** (modular!) pieces

I need to perform **cellular analysis**.



I need to **compare** measure A to B over time.



Task Abstraction

Why?

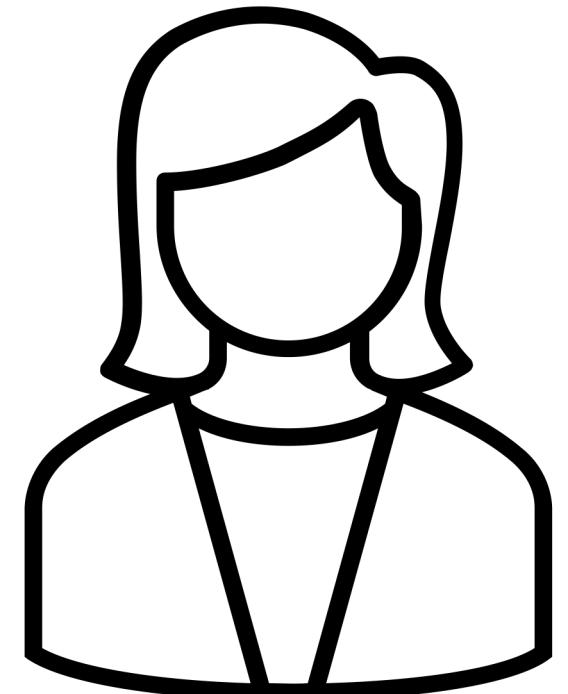
To translate domain specific terms into well known and transferable visualization tasks.

Task Abstraction

Why?

To translate domain specific terms into well known and transferable visualization tasks.

Ex.



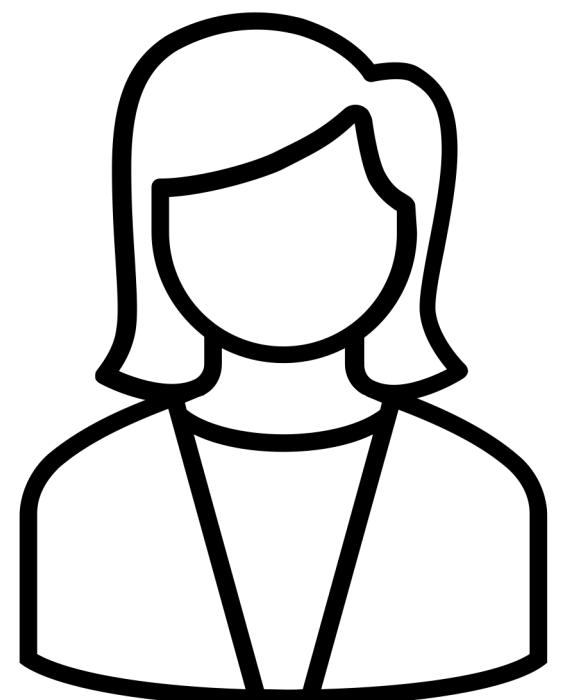
I need to show my
boss which product
we sold the most of
last year.

Task Abstraction

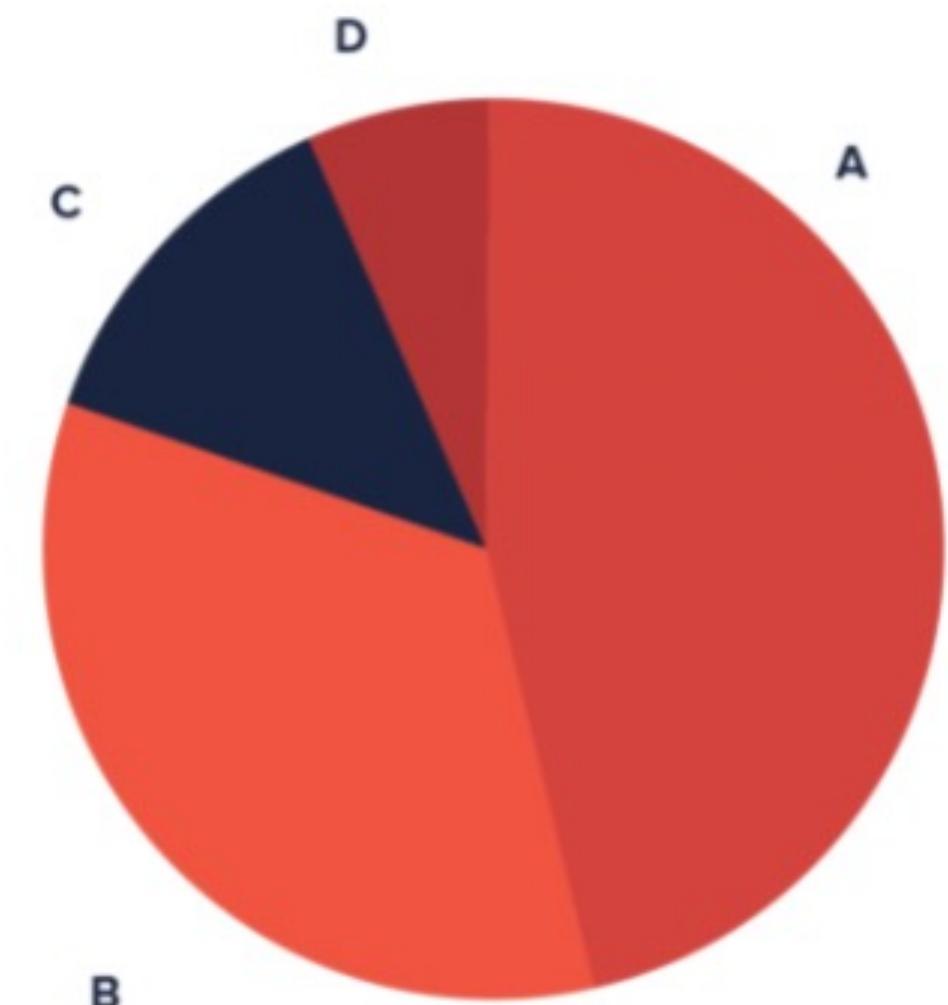
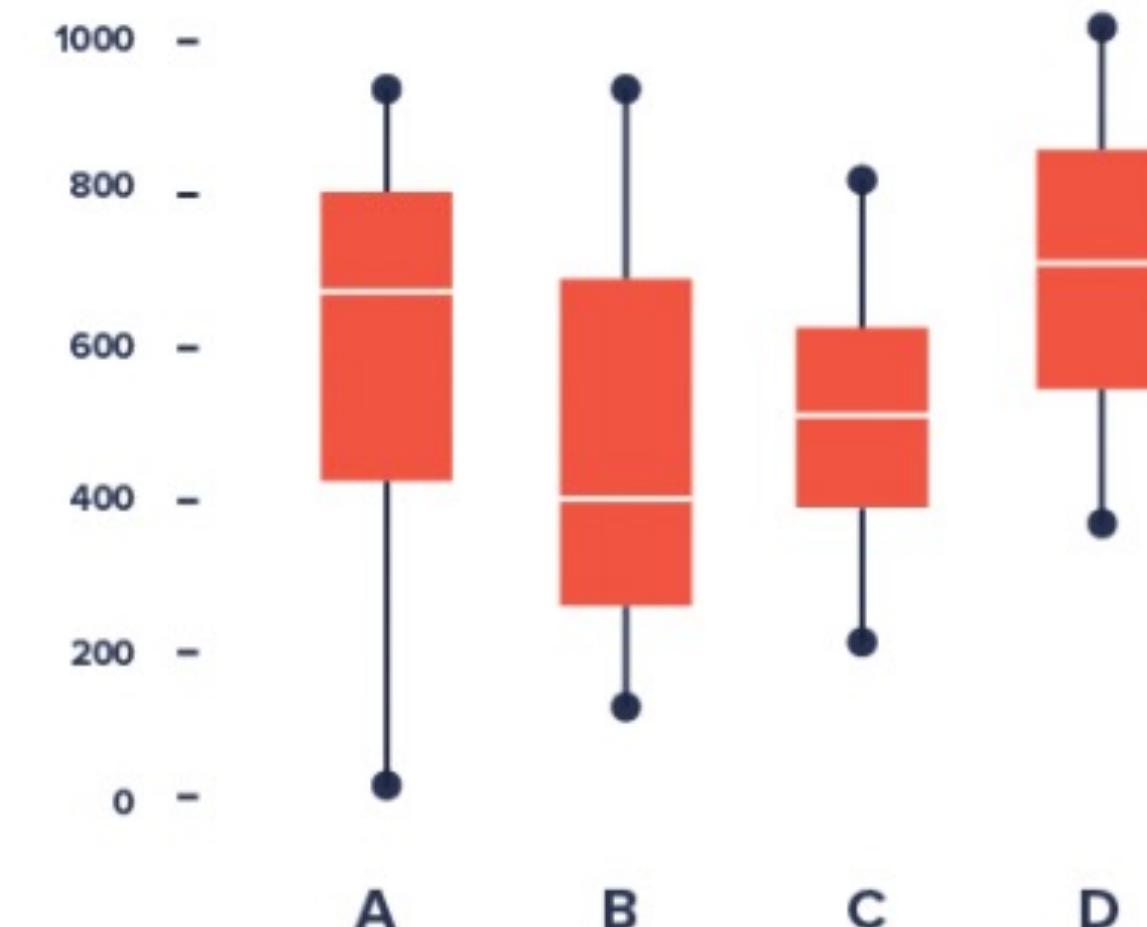
Why?

To translate domain specific terms into well known and transferable visualization tasks.

Ex.



I need to show my boss which product we sold the most of last year.



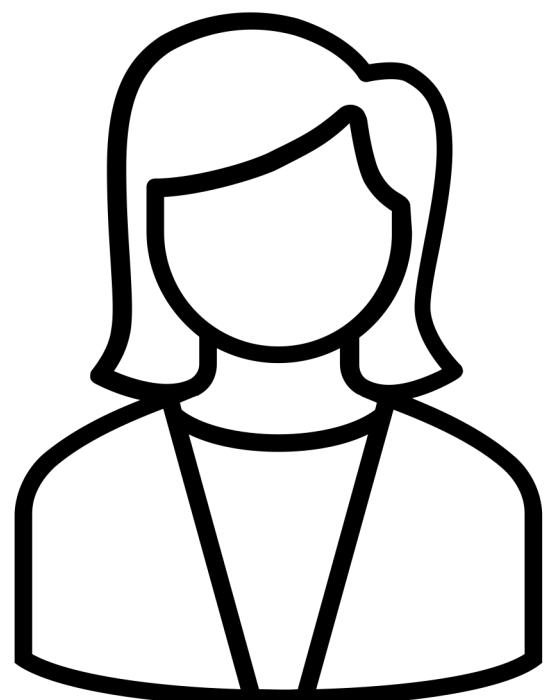
<https://datavizproject.com/>

Task Abstraction

Why?

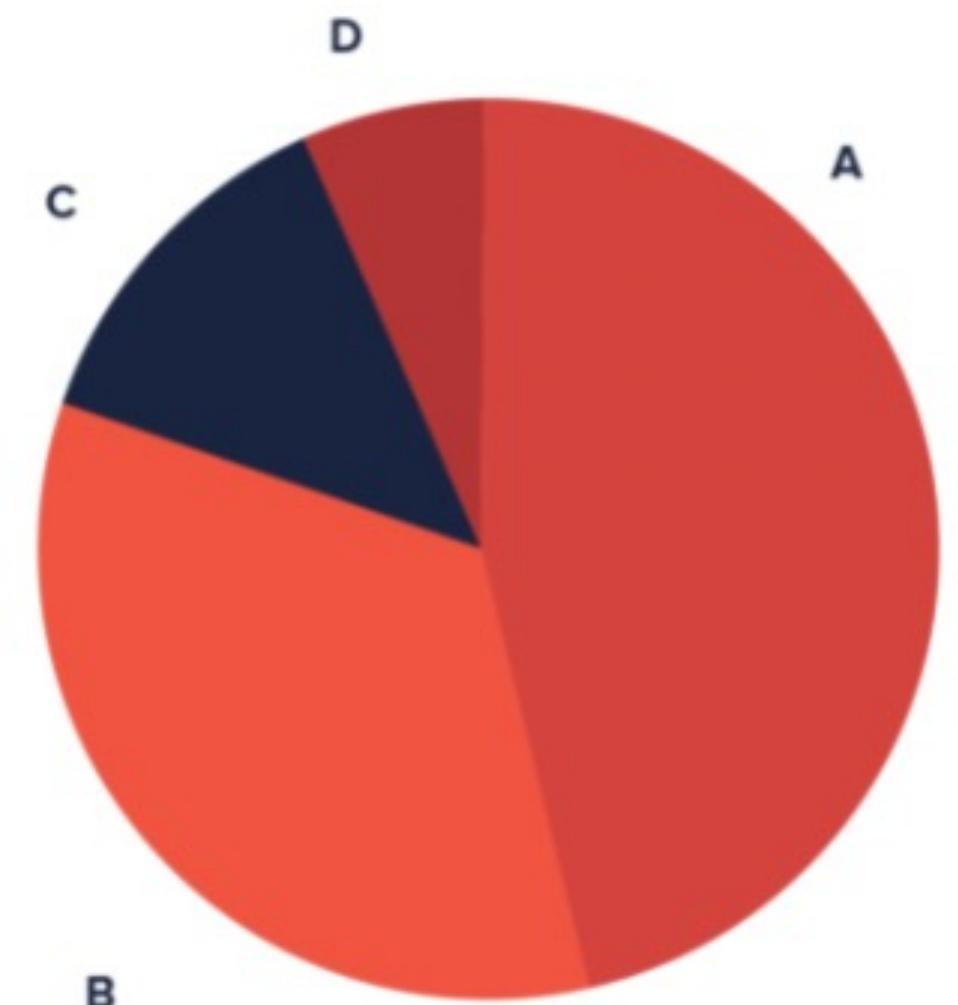
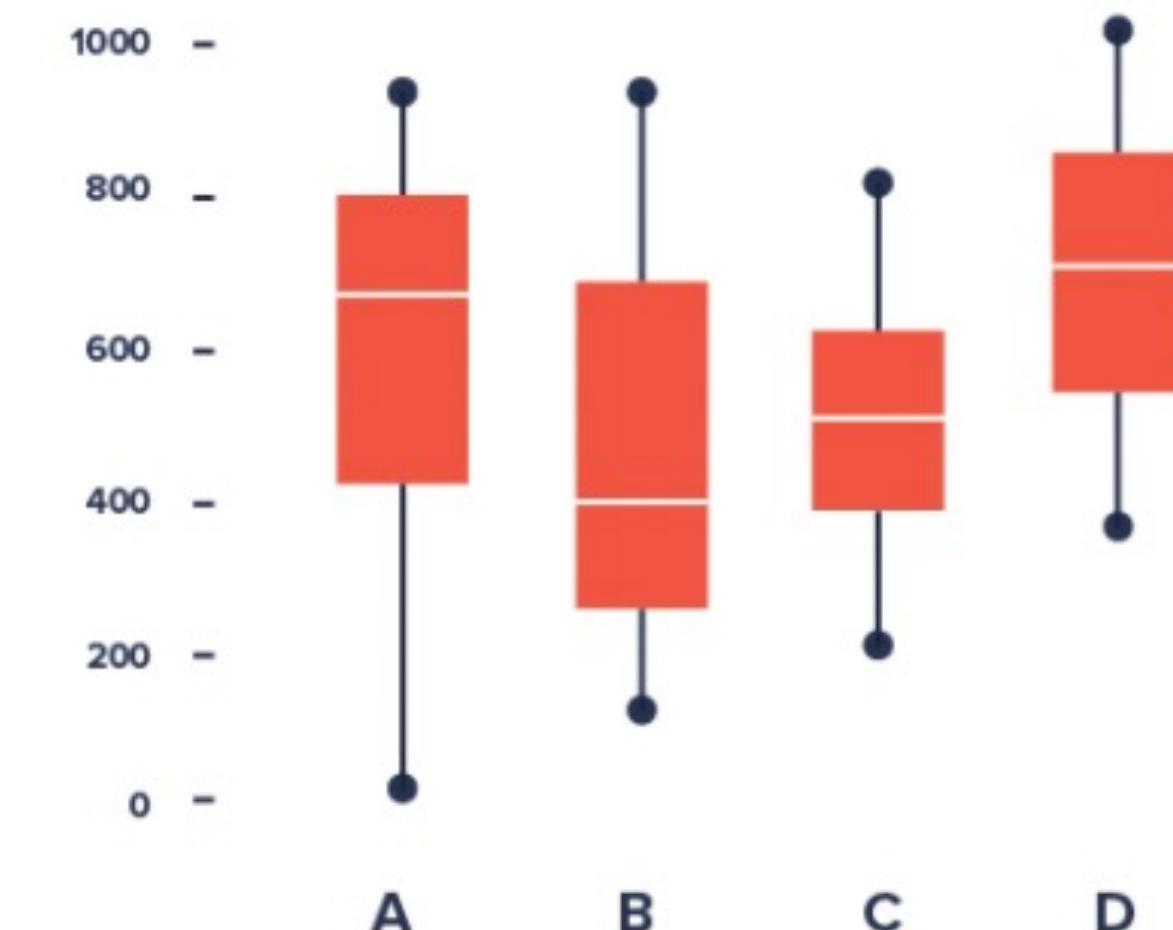
To translate domain specific terms into well known and transferable visualization tasks.

Ex.



I need to show my boss which product we sold the most of last year.

Compare distributions **Compare part-to-whole**



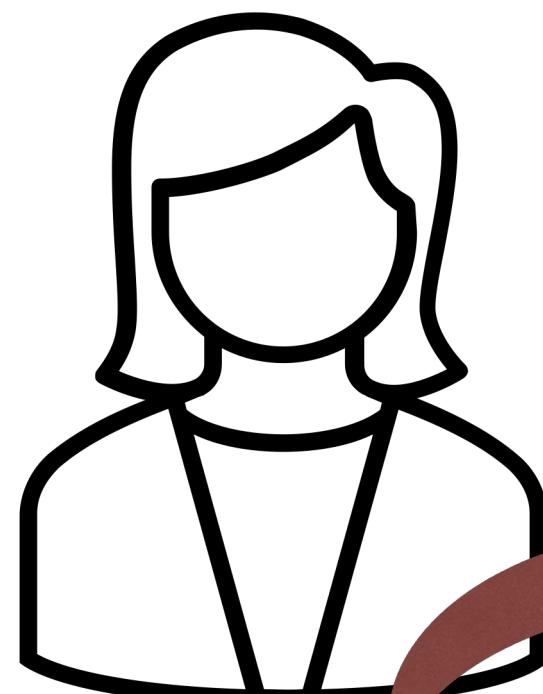
<https://datavizproject.com/>

Task Abstraction

Why?

To translate domain specific terms into well known and transferable visualization tasks.

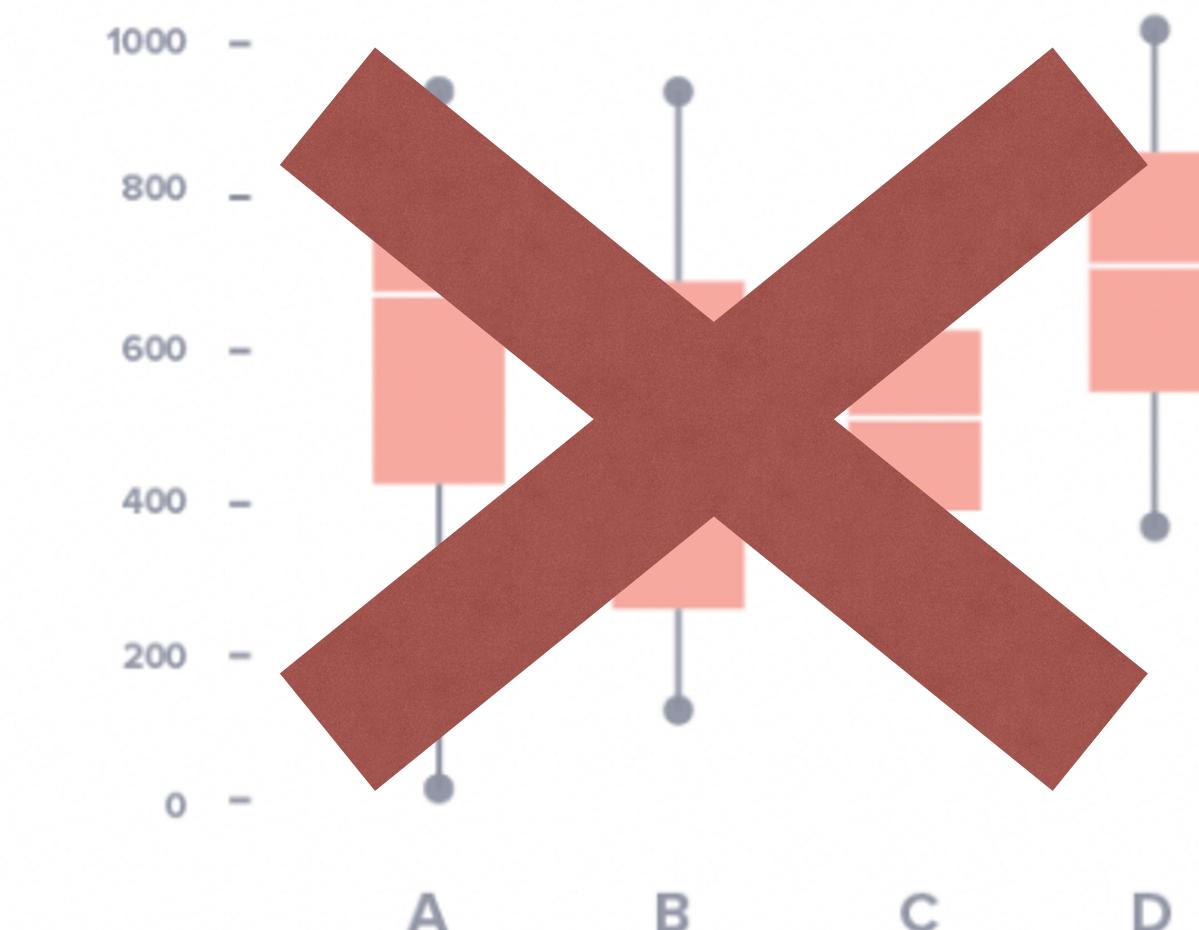
Ex.



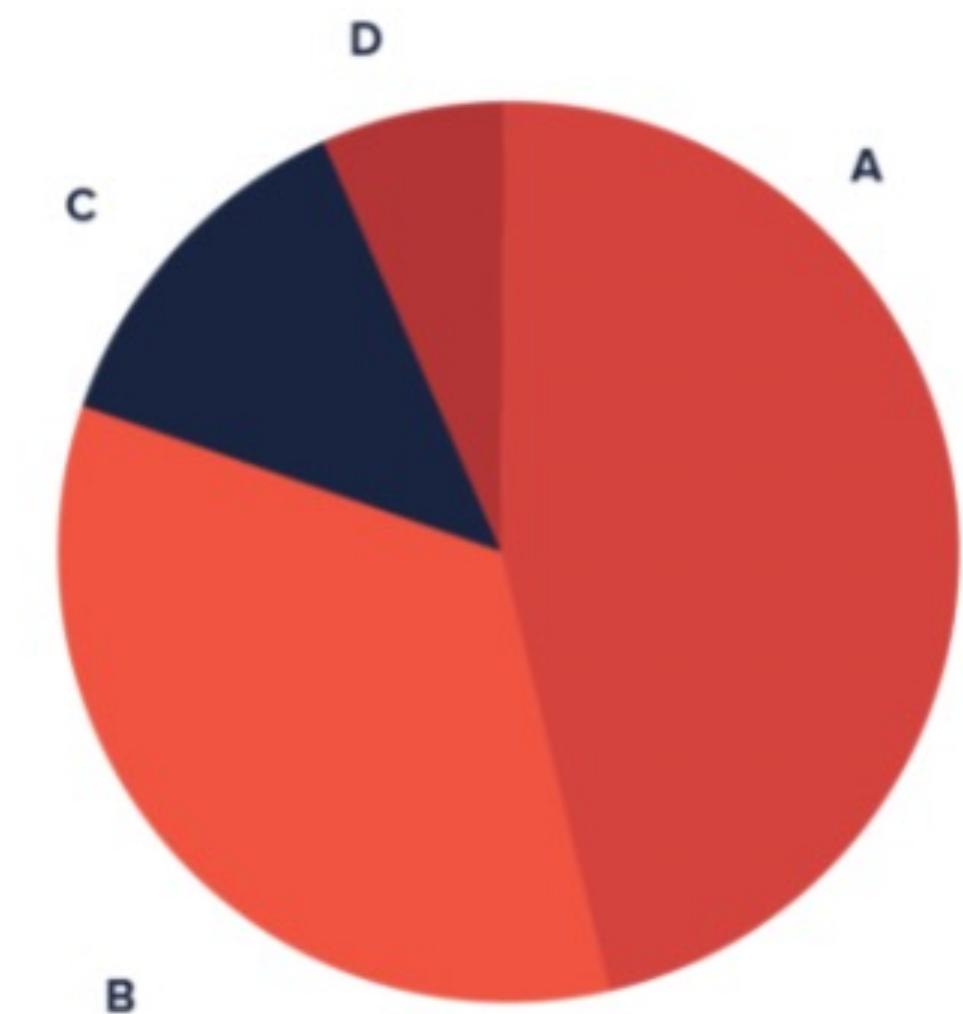
I need to show my boss which product we sold the most of last year.

Need to compare part-to-whole.

Compare distributions



Compare part-to-whole



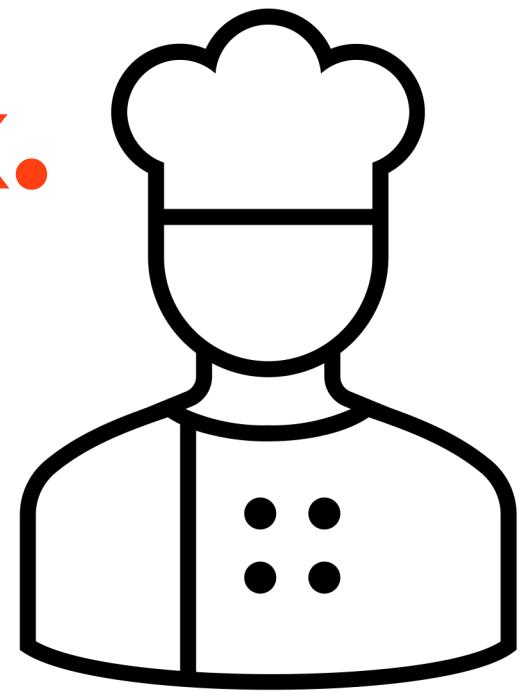
<https://datavizproject.com/>

Task Abstraction

Why?

To translate domain specific terms into well known and transferable visualization tasks.

Ex.



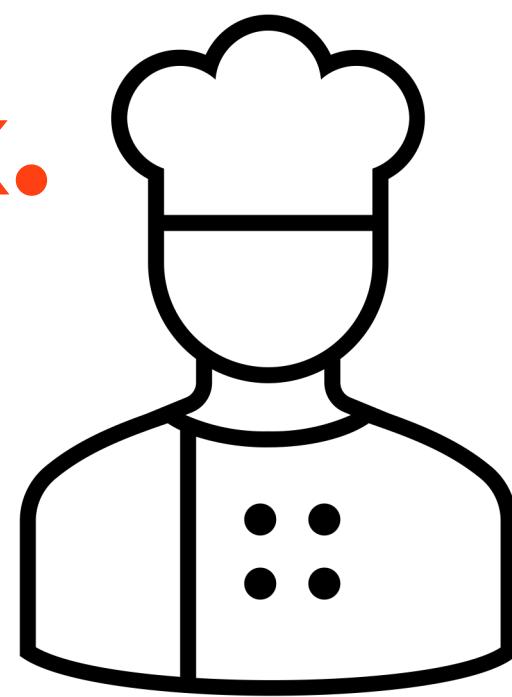
I want to see which months are best for outdoor dining.

Task Abstraction

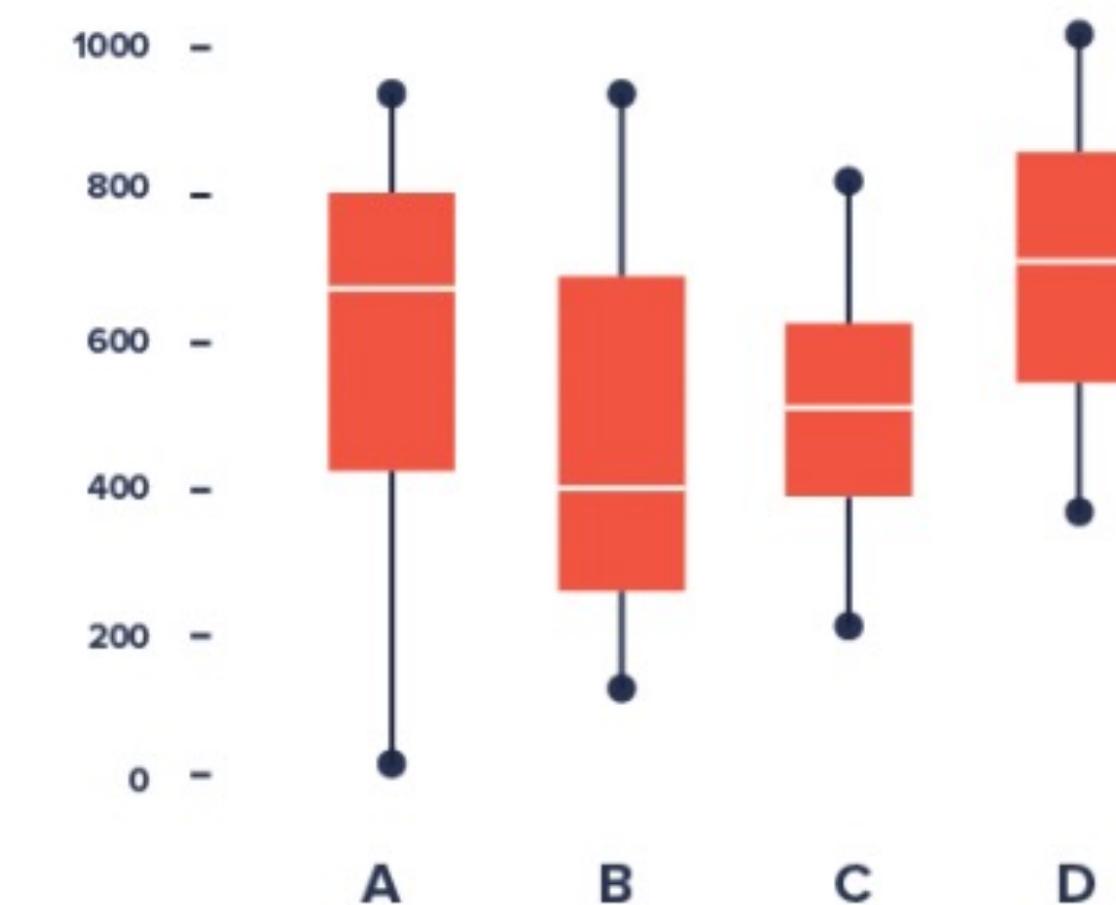
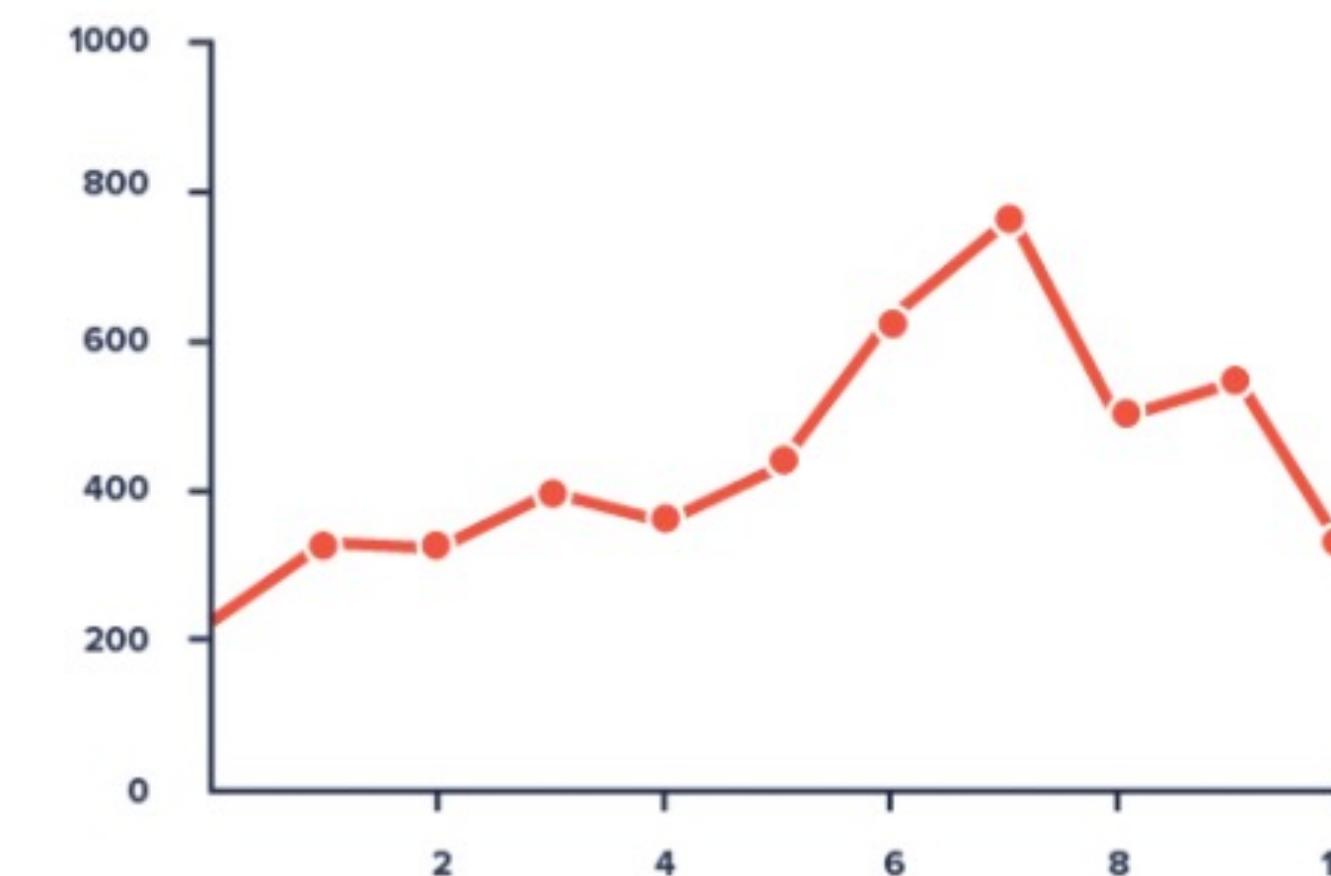
Why?

To translate domain specific terms into well known and transferable visualization tasks.

Ex.



I want to see which months are best for outdoor dining.



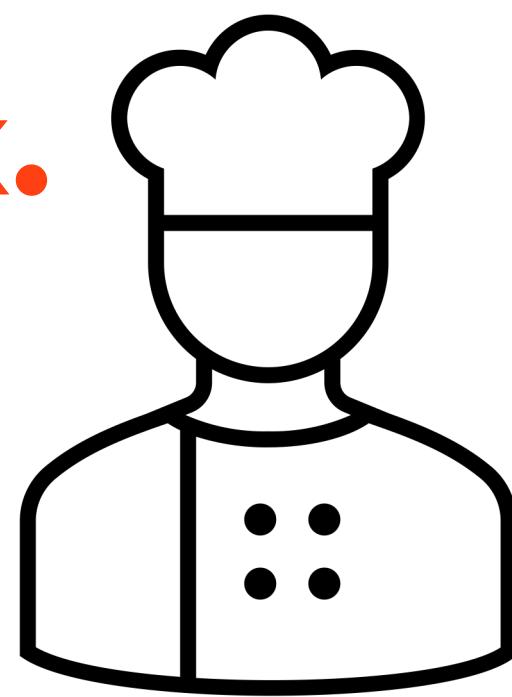
<https://datavizproject.com/>

Task Abstraction

Why?

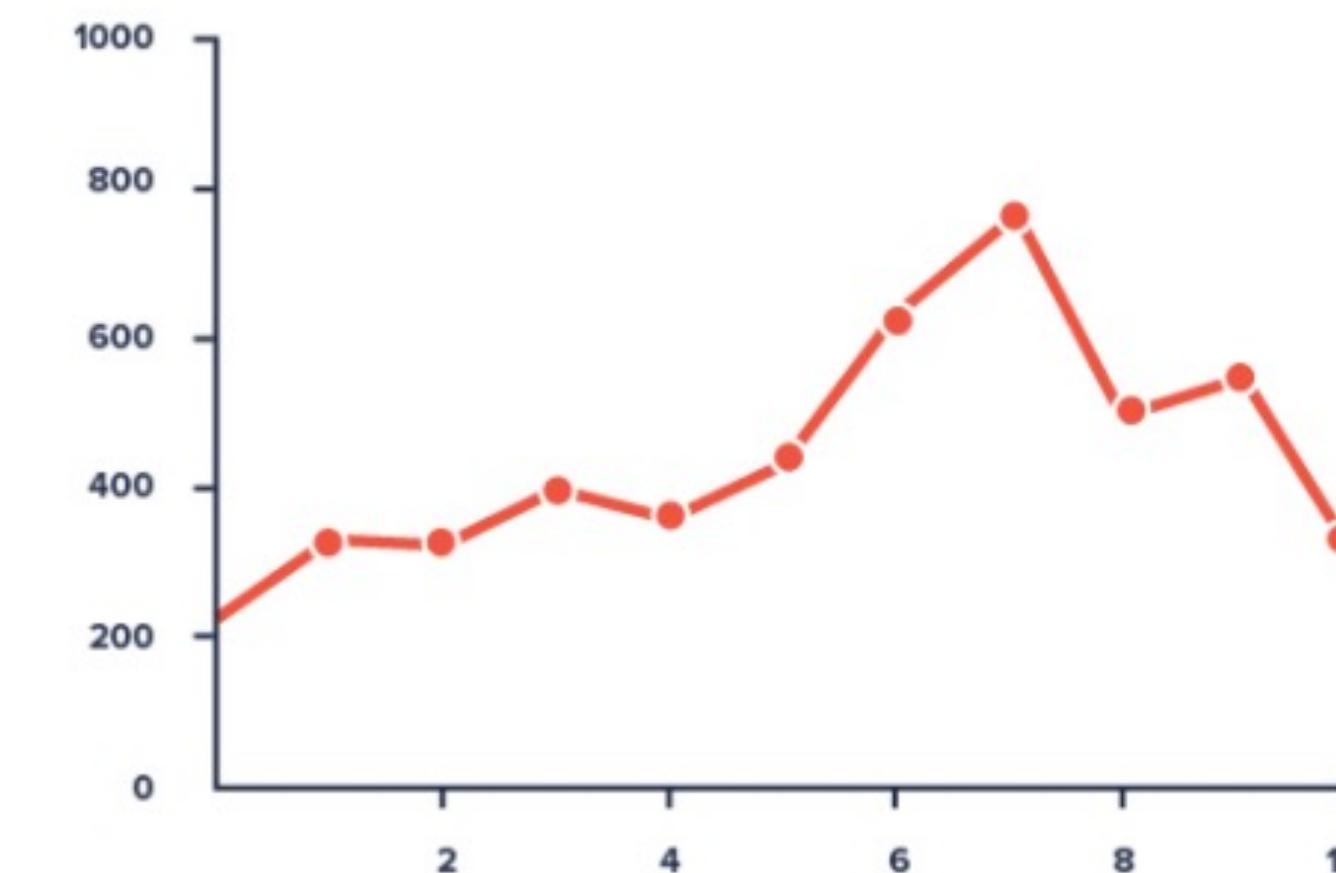
To translate domain specific terms into well known and transferable visualization tasks.

Ex.

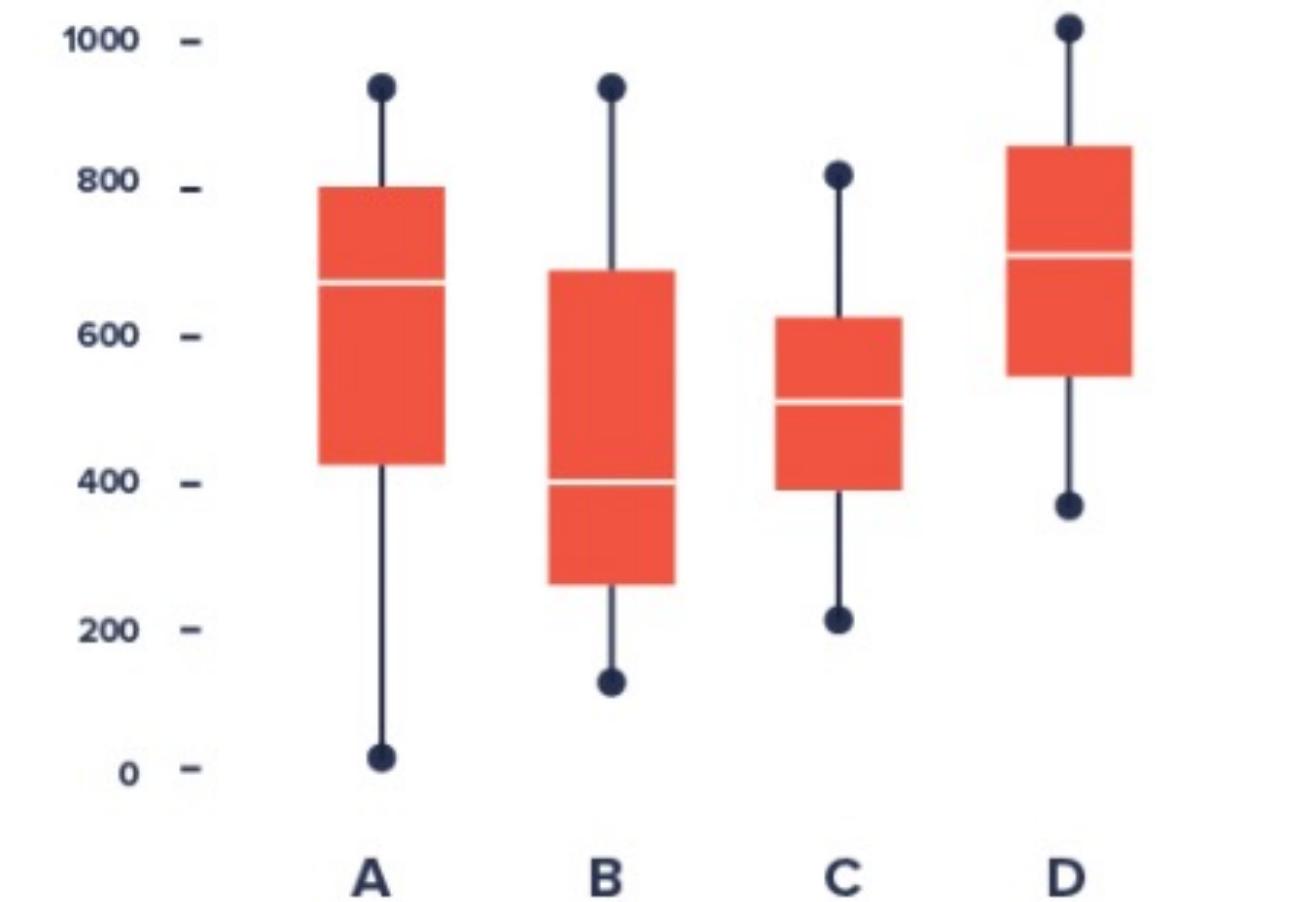


I want to see which months are best for outdoor dining.

Average trend



Compare distributions



<https://datavizproject.com/>

Task Abstraction

Why?

To translate domain specific terms into well known and transferable visualization tasks.

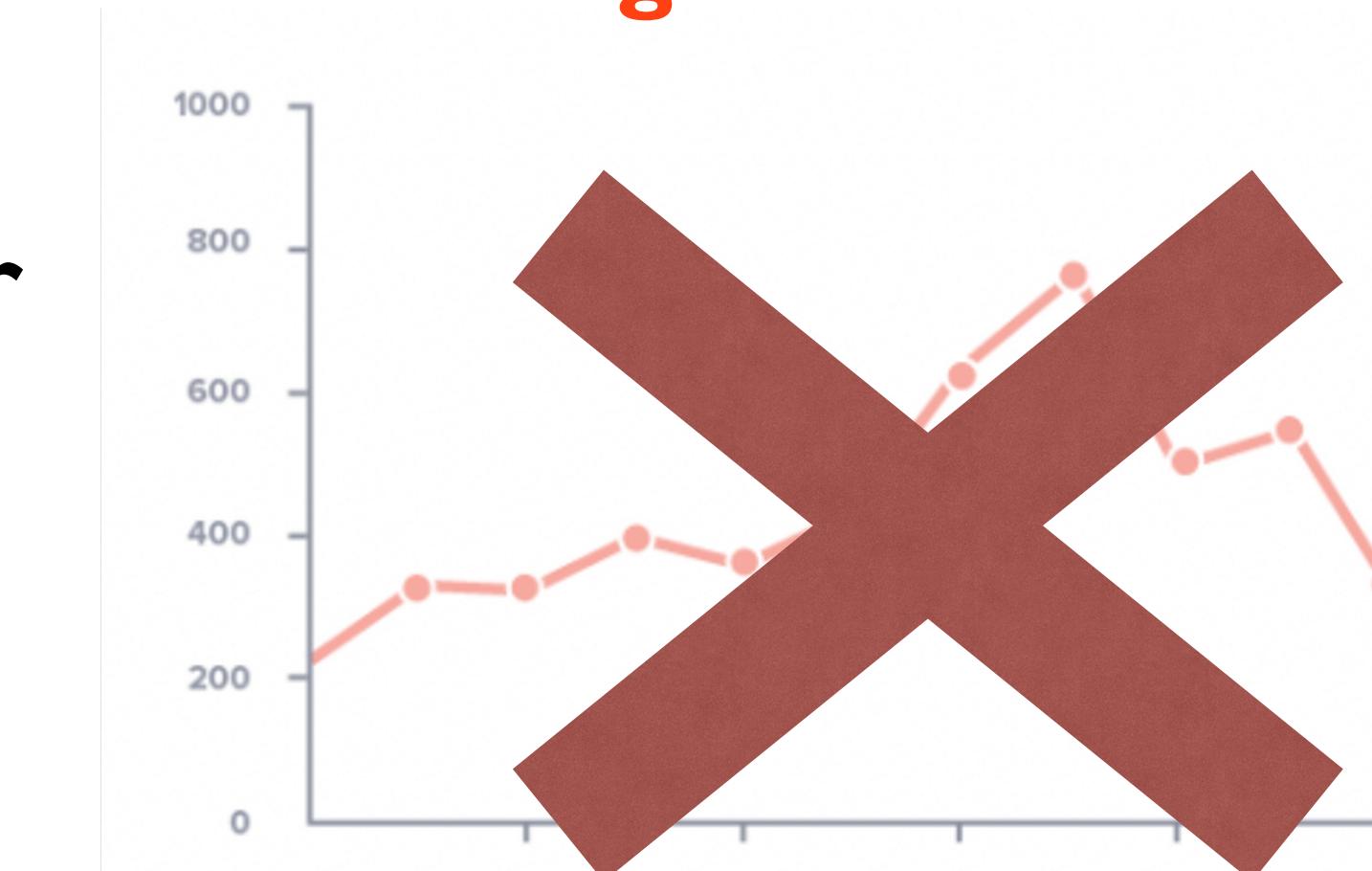
Ex.



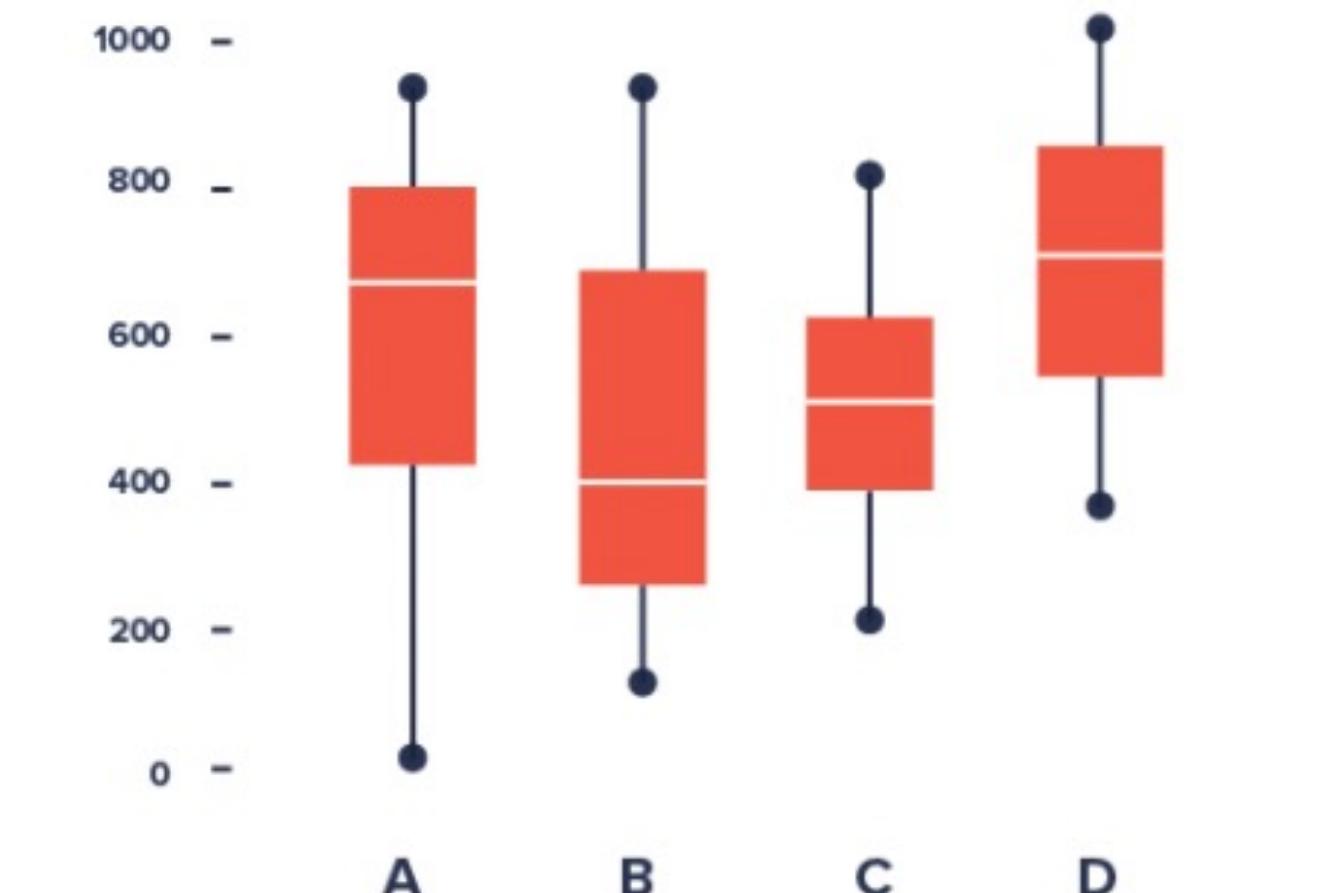
I want to see which months are best for outdoor dining.

Need to compare distribution over months.

Average trend



Compare distributions



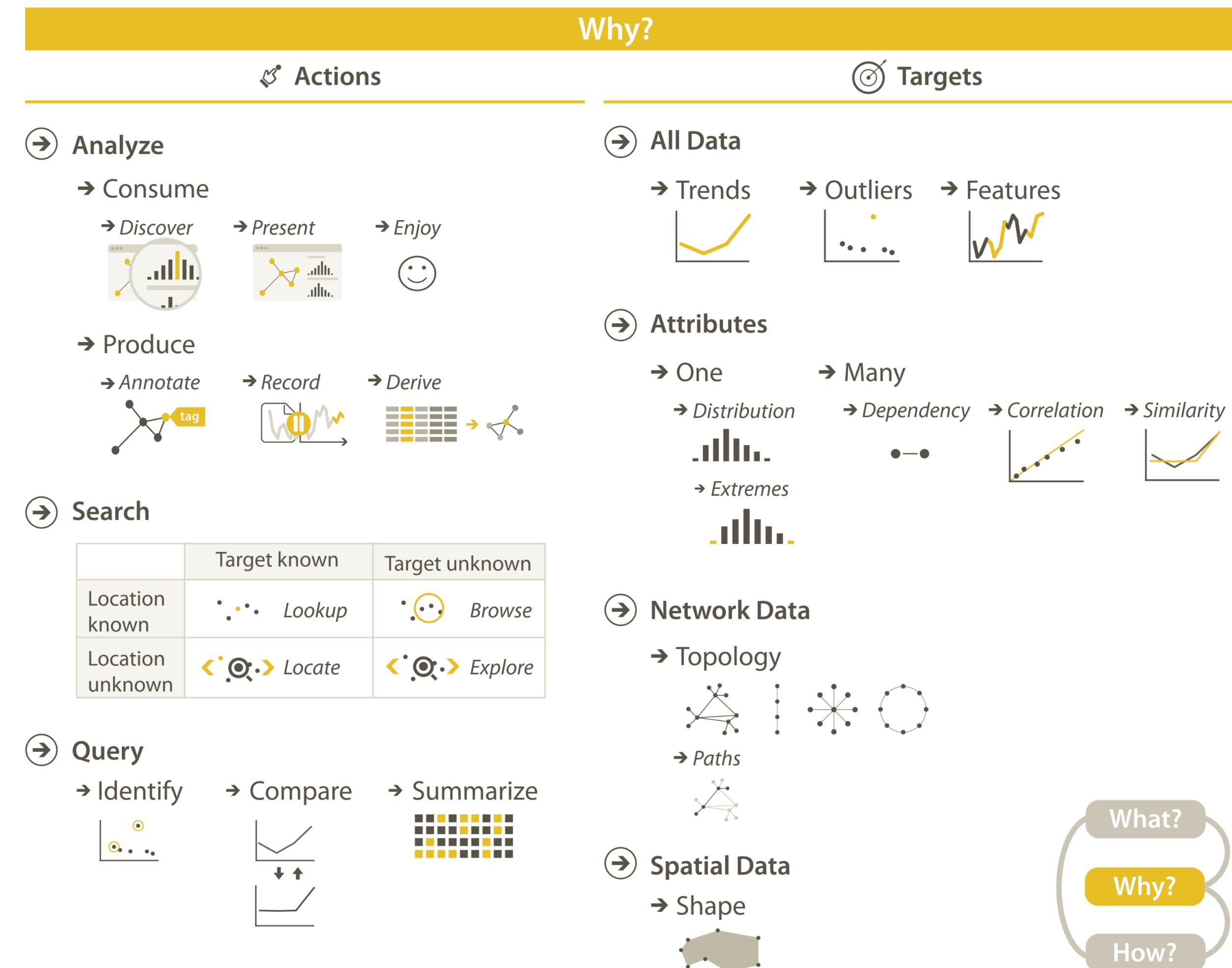
<https://datavizproject.com/>

Task Abstraction

Why?

To translate domain specific terms into **well known and transferable visualization tasks.**

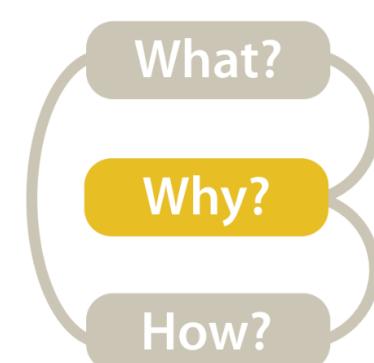
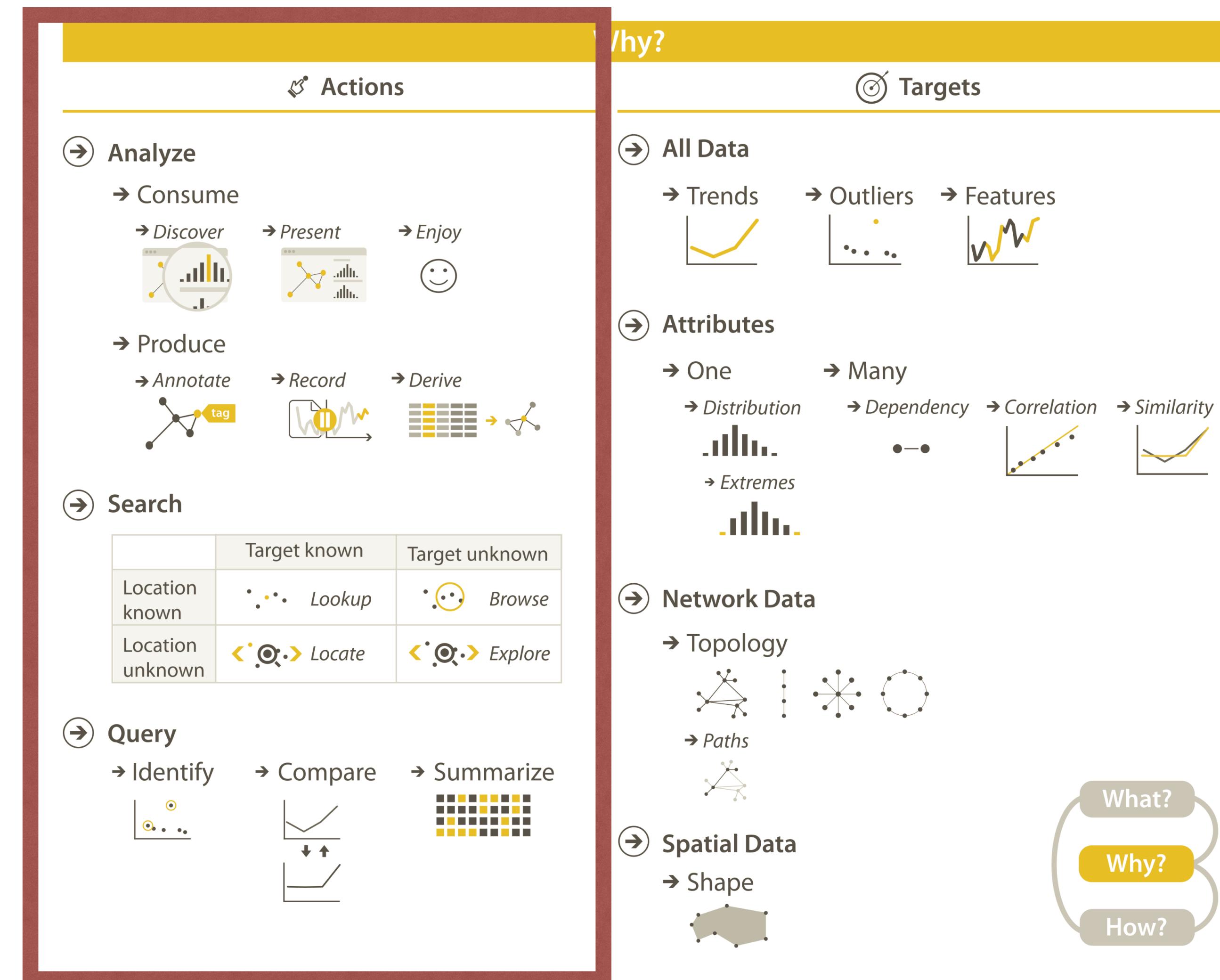
Task Abstraction



Task Abstraction

ACTIONS

- 3 LEVELS:
- High
 - Medium
 - Low



Task Abstraction

ACTIONS

- 3 LEVELS:**
- High
 - Medium
 - Low



Task Abstraction

What?
Why?
How?

High-level → How is the vis being used to analyze?

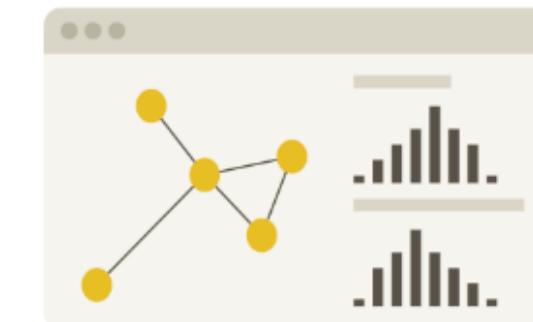
→ Analyze

→ Consume

→ Discover



→ Present

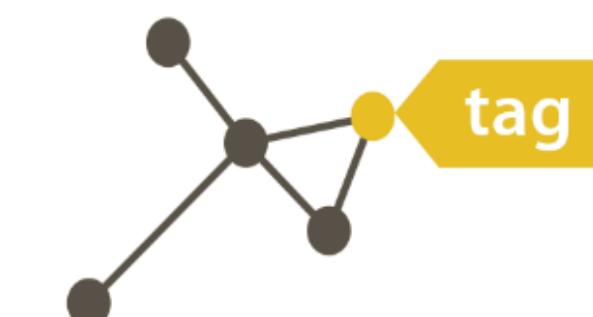


→ Enjoy

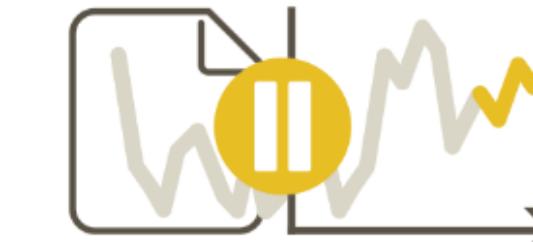


→ Produce

→ Annotate



→ Record



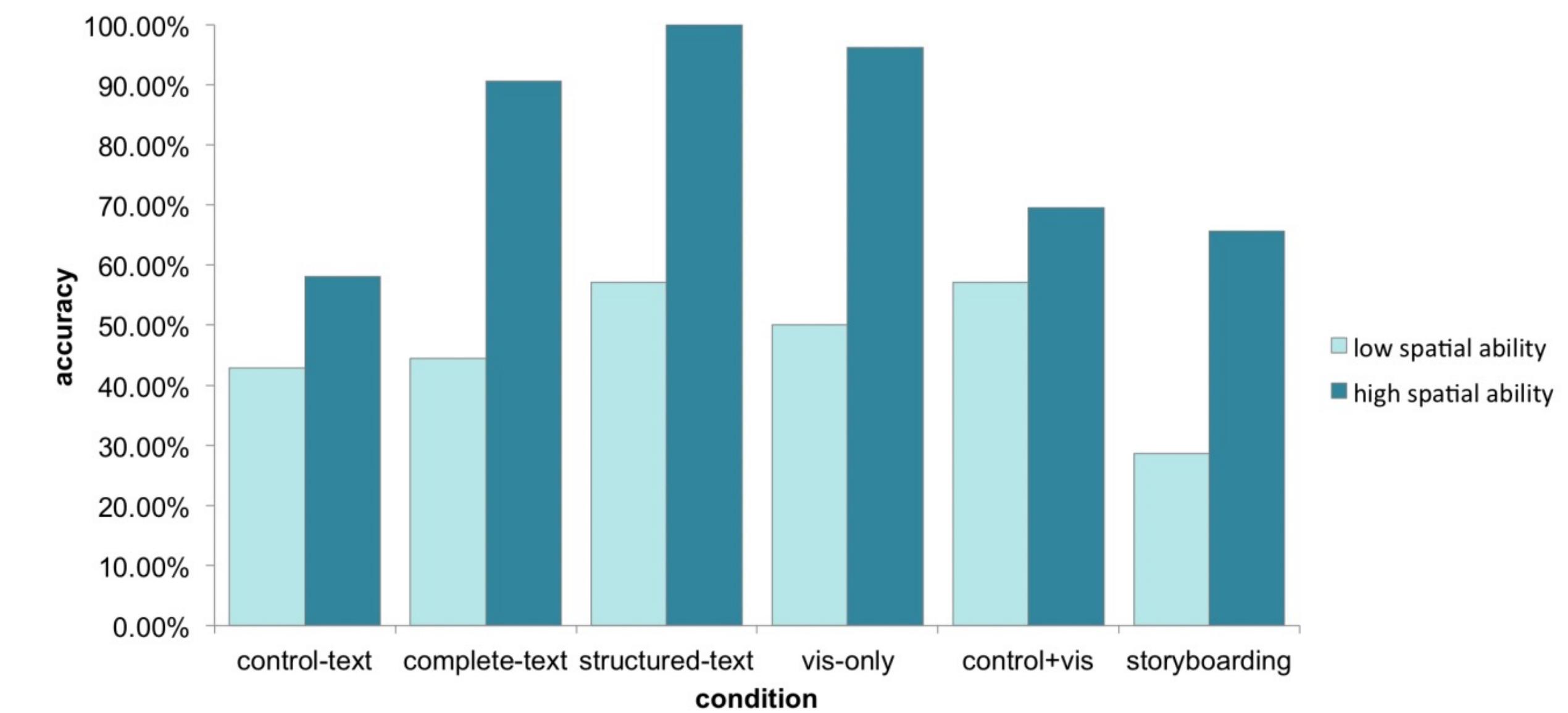
→ Derive



Task Abstraction

What?
Why?
How?

High-level → Consume → Discover



Task Abstraction

What?
Why?
How?

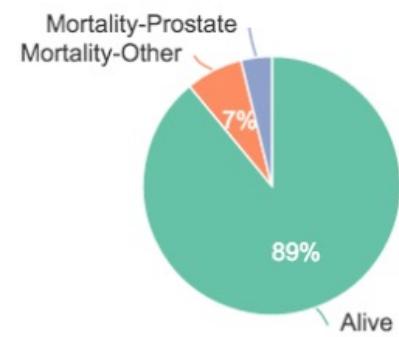
High-level → Consume → Present

How big of a threat is my prostate cancer?

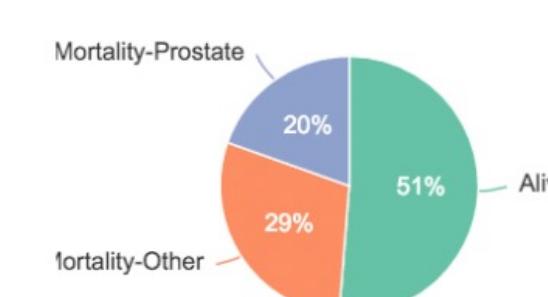
Before thinking about the benefits of specific treatments, it's helpful to first think about how big of threat your prostate cancer is to your future survival. The pie chart below shows the following:

- Your chances of being **alive** (in **GREEN**)
- Your chances of dying from your **prostate cancer** (in **PURPLE**)
- Your chances of dying from **other causes** (in **ORANGE**)

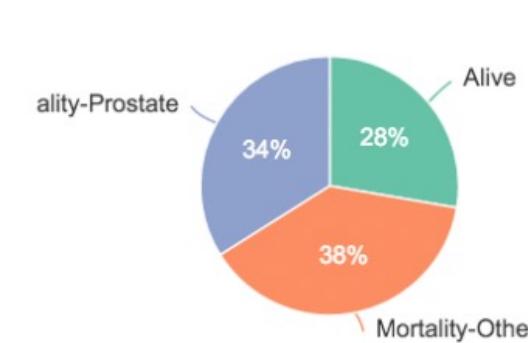
1 Year (70 years old)



5 years (74 years old)



10 Years (79 years old)

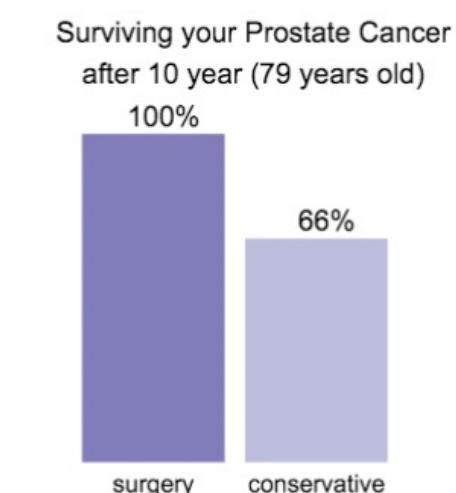
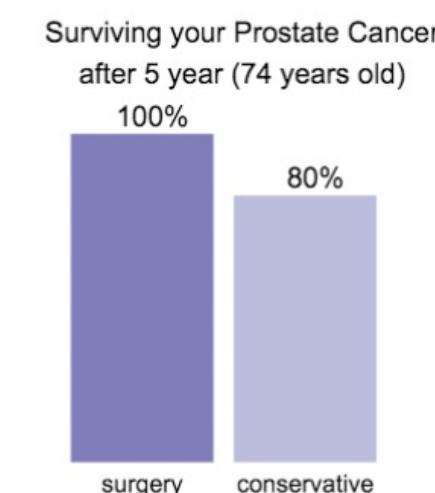
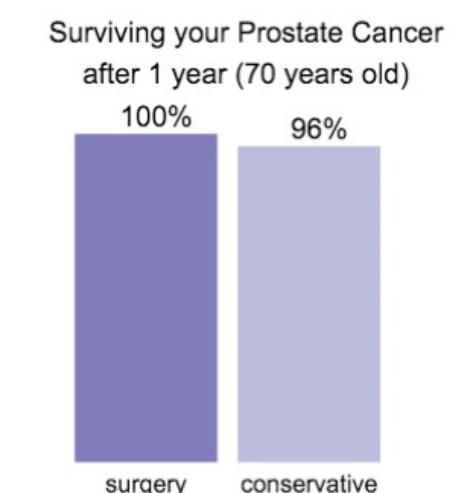


How effective are different treatments for my prostate cancer?

The expected benefits from **surgery** and **conservative management** are listed below.

These results show your estimated chances of either surviving or dying **from your prostate cancer** at 1, 5, and 10 years, depending on whether you choose either surgery (**DARK PURPLE BAR**) or conservative treatment (**LIGHT PURPLE BAR**).

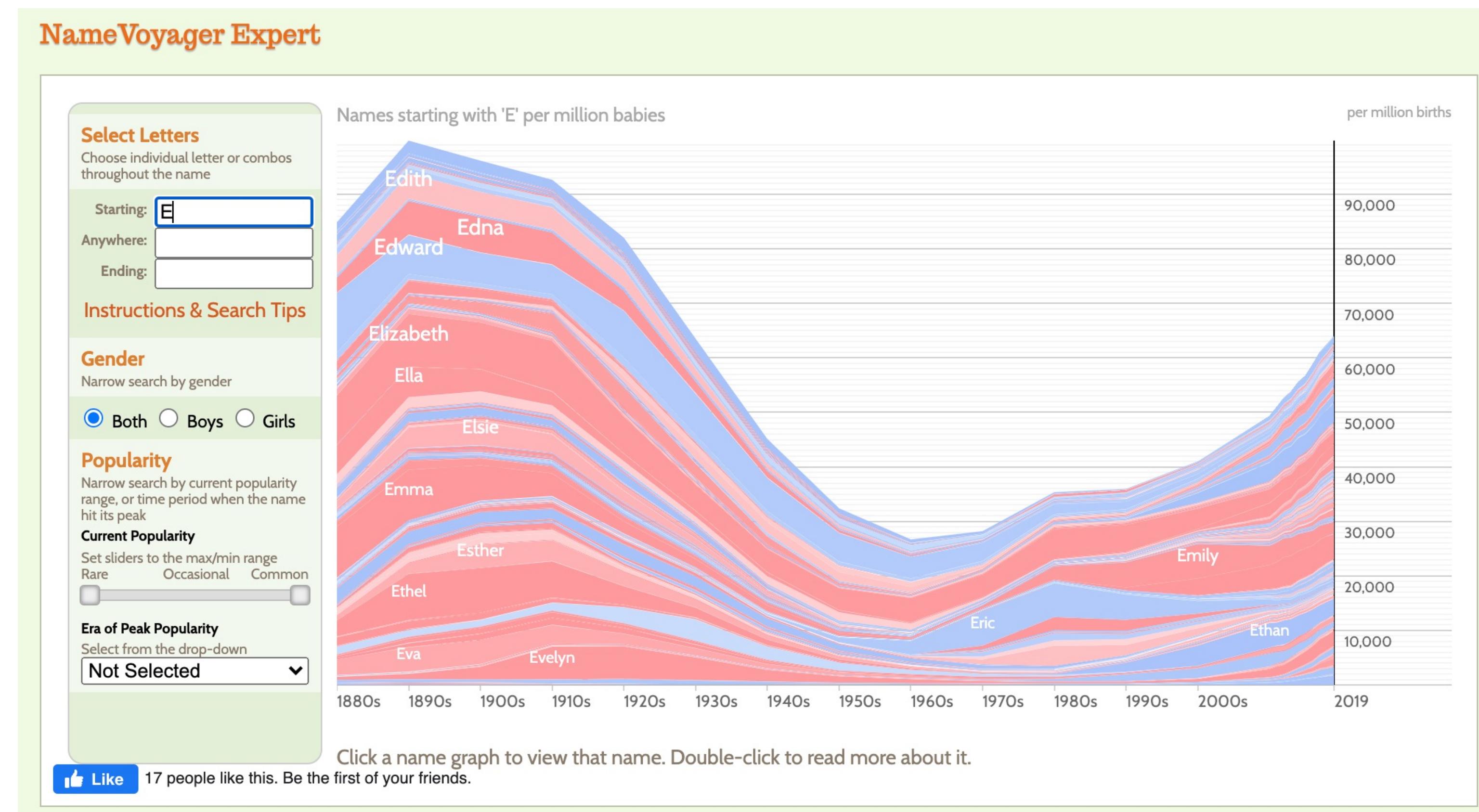
You can view these risks in terms of either survival or mortality.



Task Abstraction

What?
Why?
How?

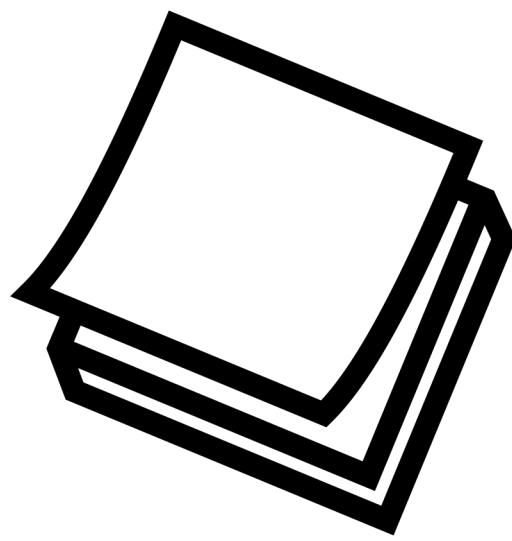
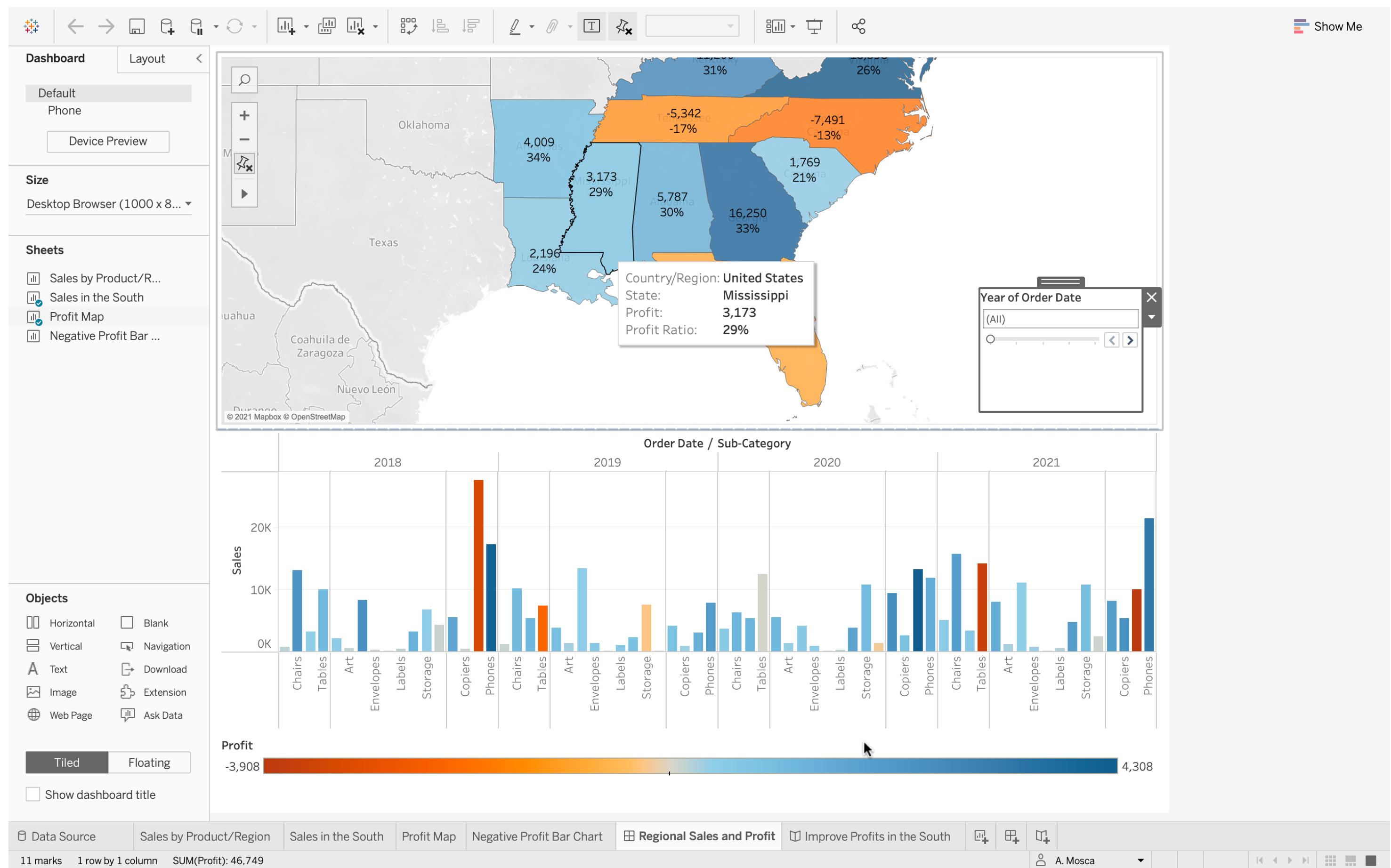
High-level → Consume → Enjoy



Task Abstraction

What?
Why?
How?

High-level → **Produce** → **Annotate**



Task Abstraction

What?
Why?
How?

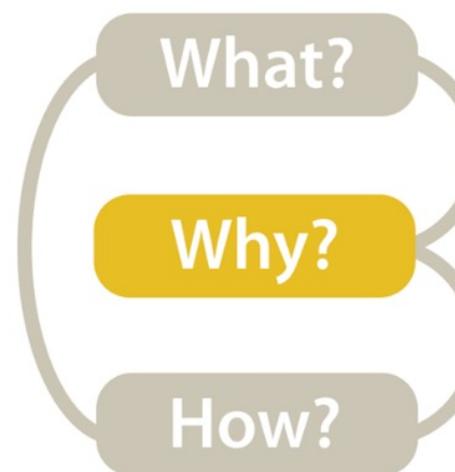
High-level → Produce → Record



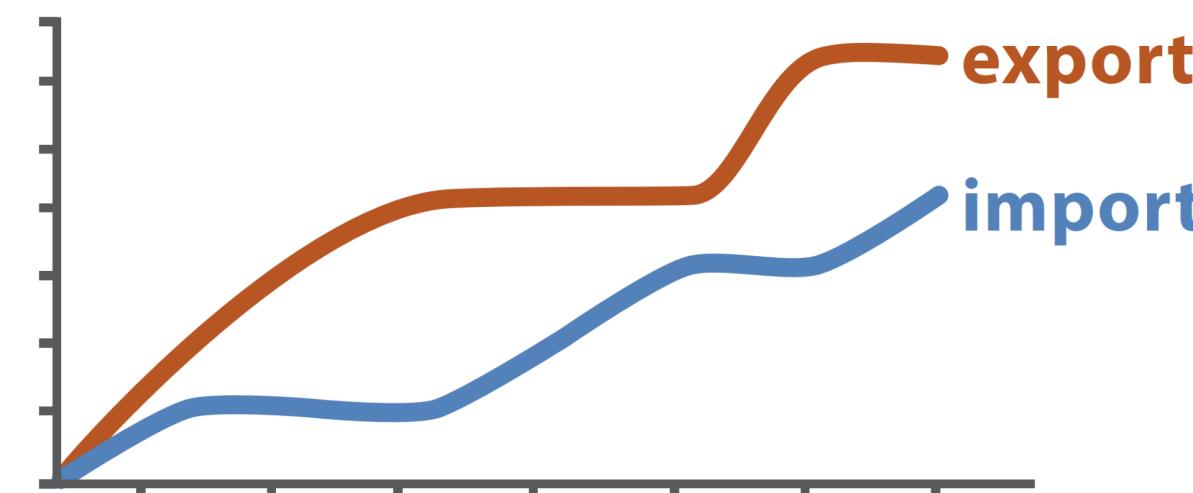
Fig. 1. A Graphical History Interface. Thumbnails show previous visualization states and labels describe the actions performed.

Graphical Histories for Visualization: Supporting Analysis, Communication, and Evaluation. Heer et al. 2008

Task Abstraction

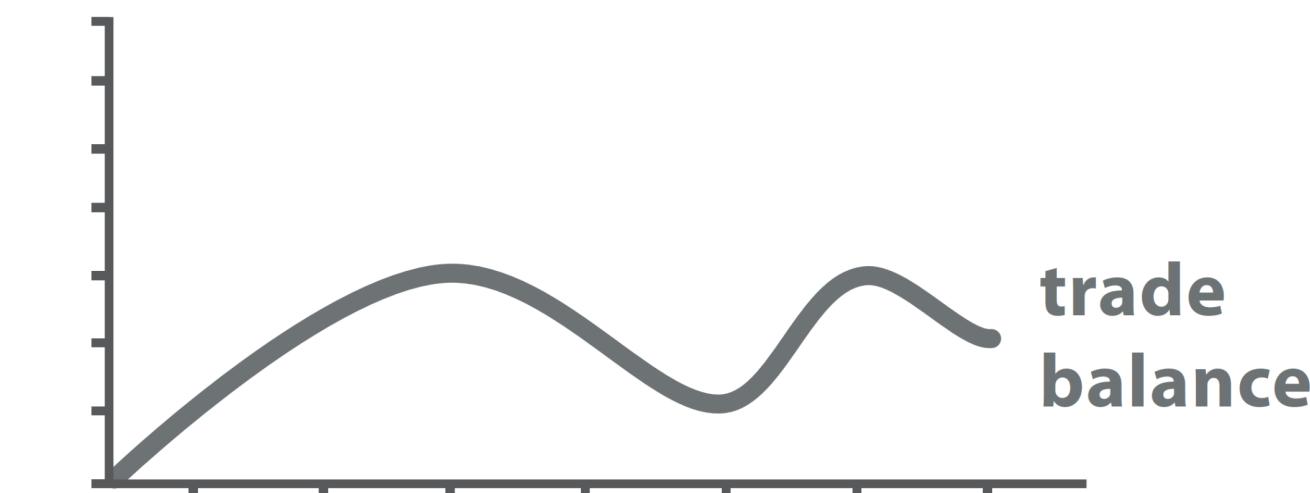


High-level → Produce → Derive



Original Data

Exports	Imports
10000	9000
15000	9050
...	...

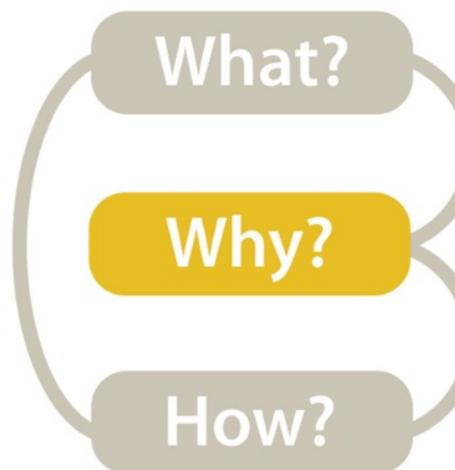


$$\text{trade balance} = \text{exports} - \text{imports}$$

Derived Data

Exports	Imports	Trade Balance
10000	9000	1000
15000	9050	5950
...

Task Abstraction



High-level → How is the vis being used to analyze?

→ Analyze

→ Consume

→ Discover



→ Present

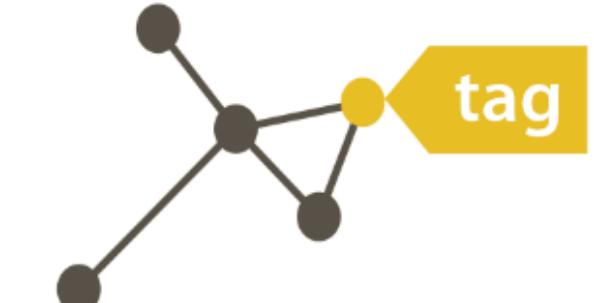


→ Enjoy

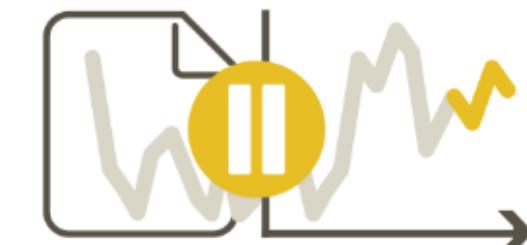


→ Produce

→ Annotate



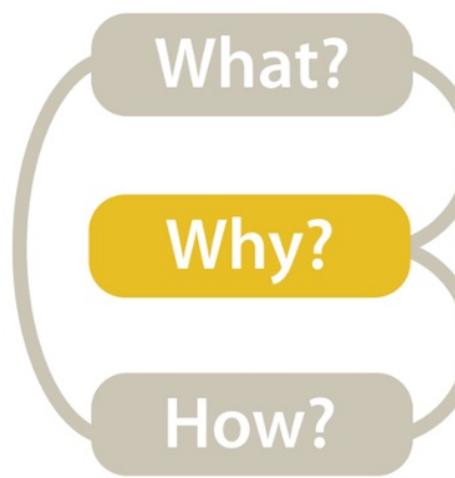
→ Record



→ Derive



Task Abstraction



Mid-level → Search

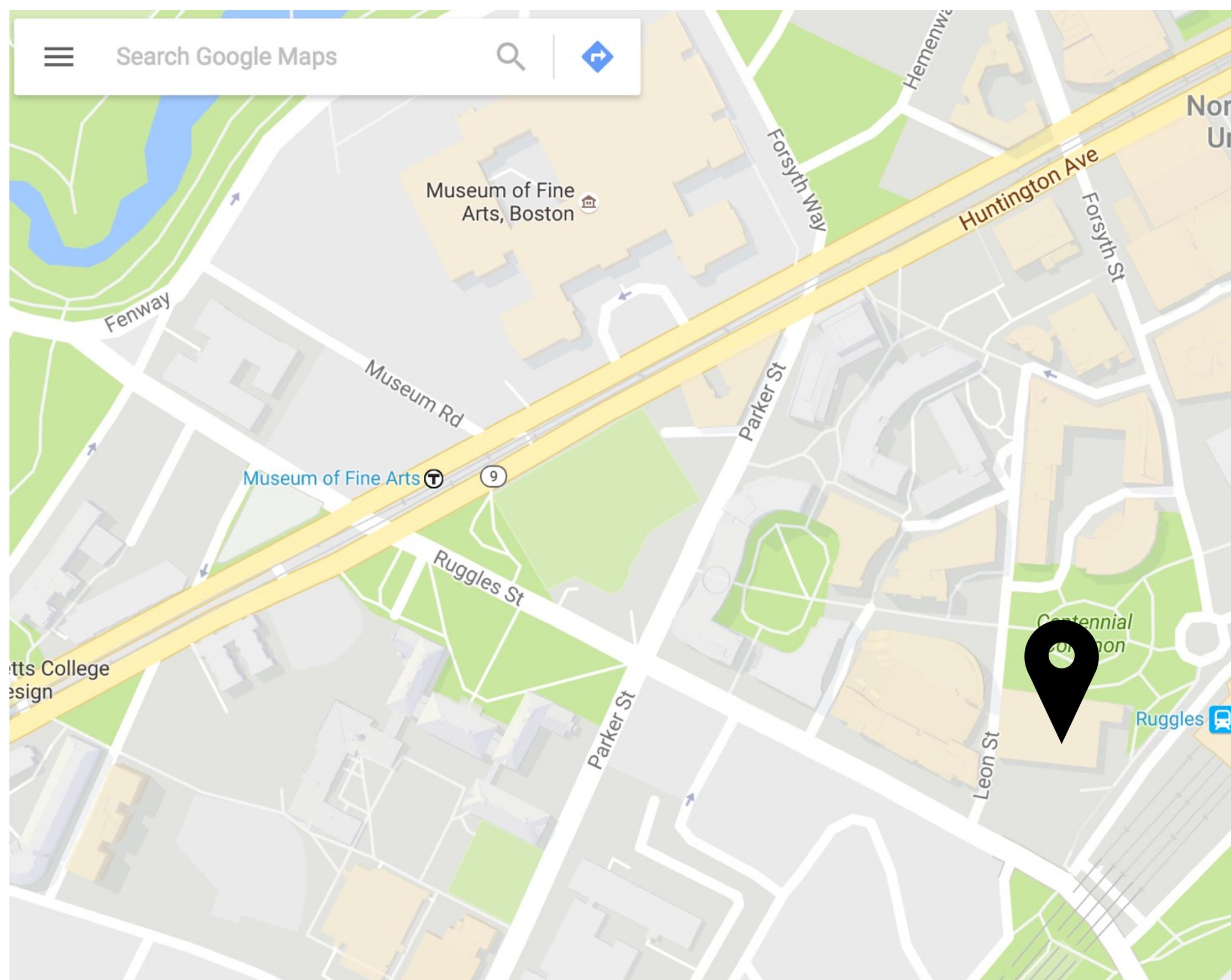
→ **Search**

	Target known	Target unknown
Location known	 <i>Lookup</i>	 <i>Browse</i>
Location unknown	 <i>Locate</i>	 <i>Explore</i>

Task Abstraction

What?
Why?
How?

Mid-level → Search → Lookup



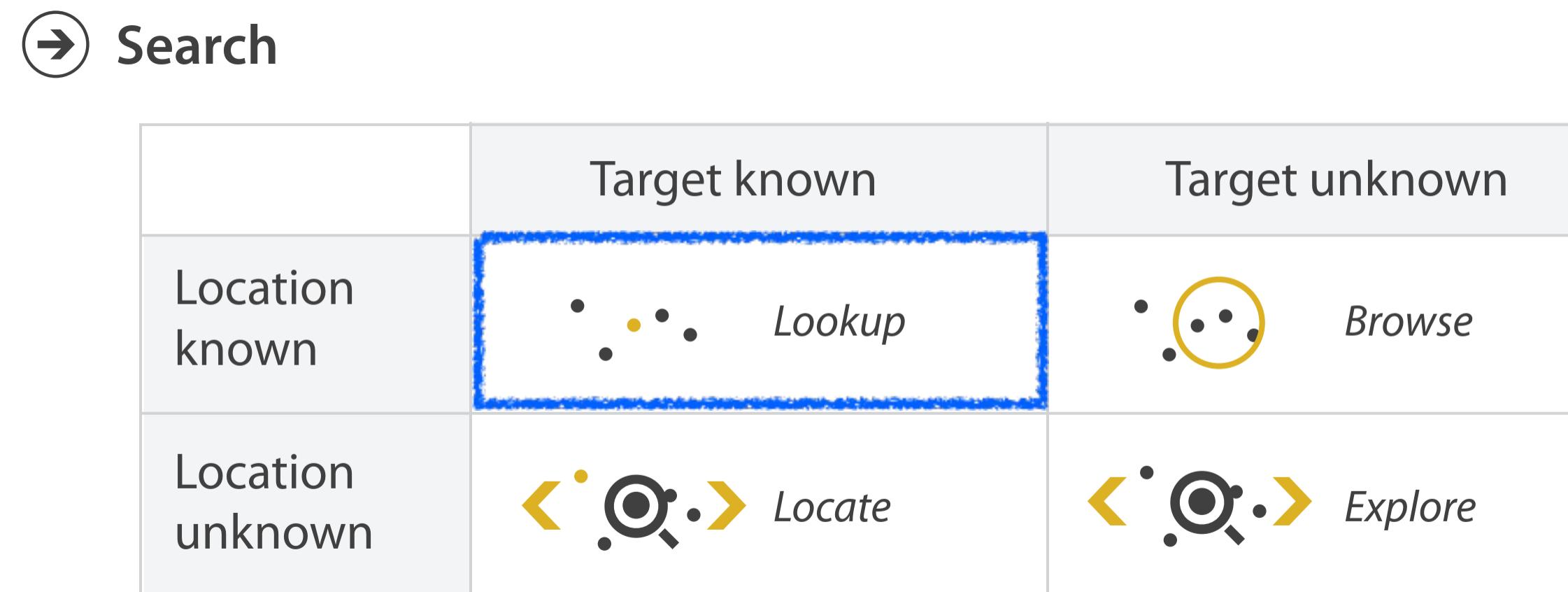
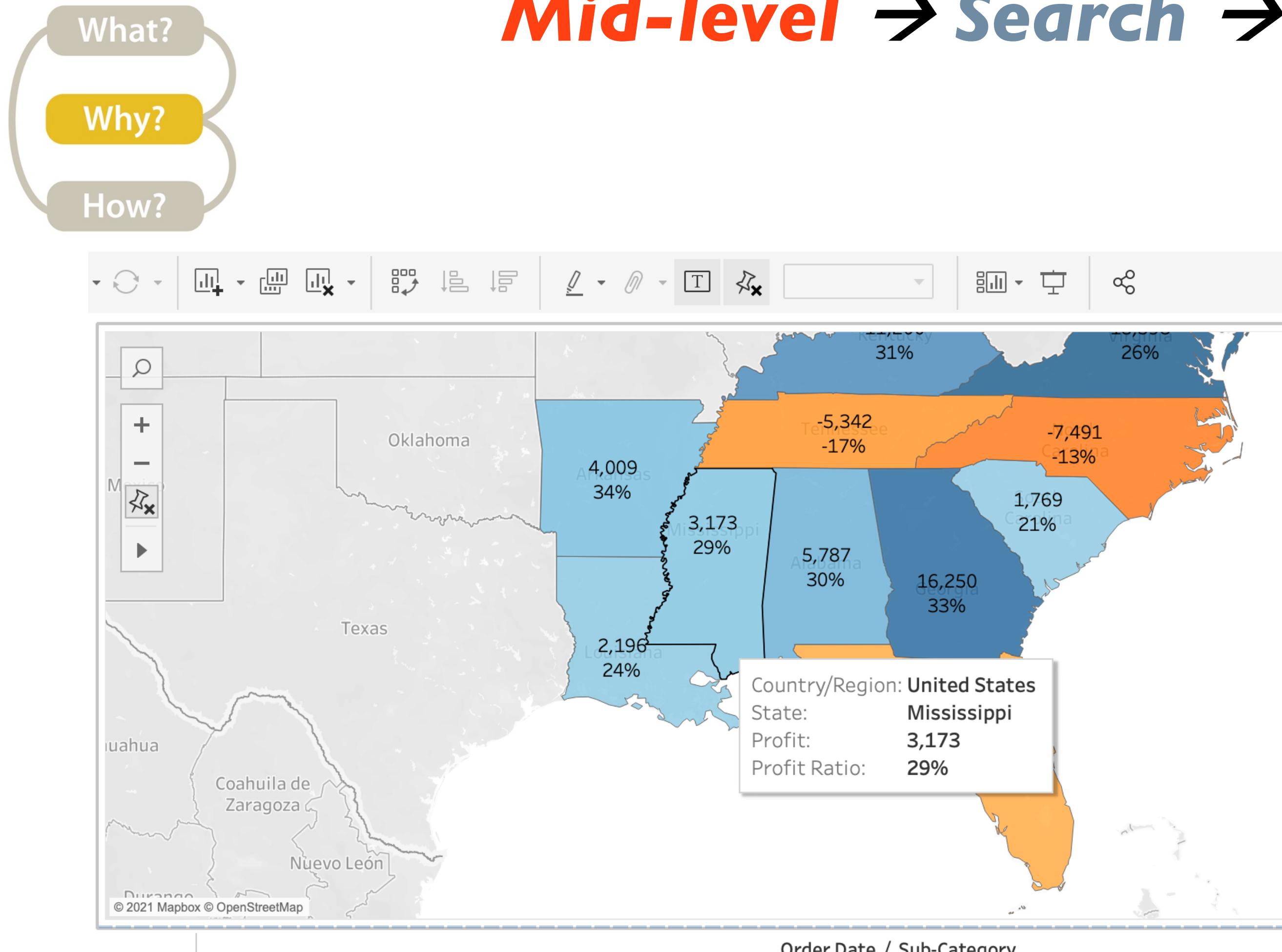
→ Search

	Target known	Target unknown
Location known	Lookup	
Location unknown		

What is the address of Ryder Hall?

Task Abstraction

Mid-level → Search → Lookup

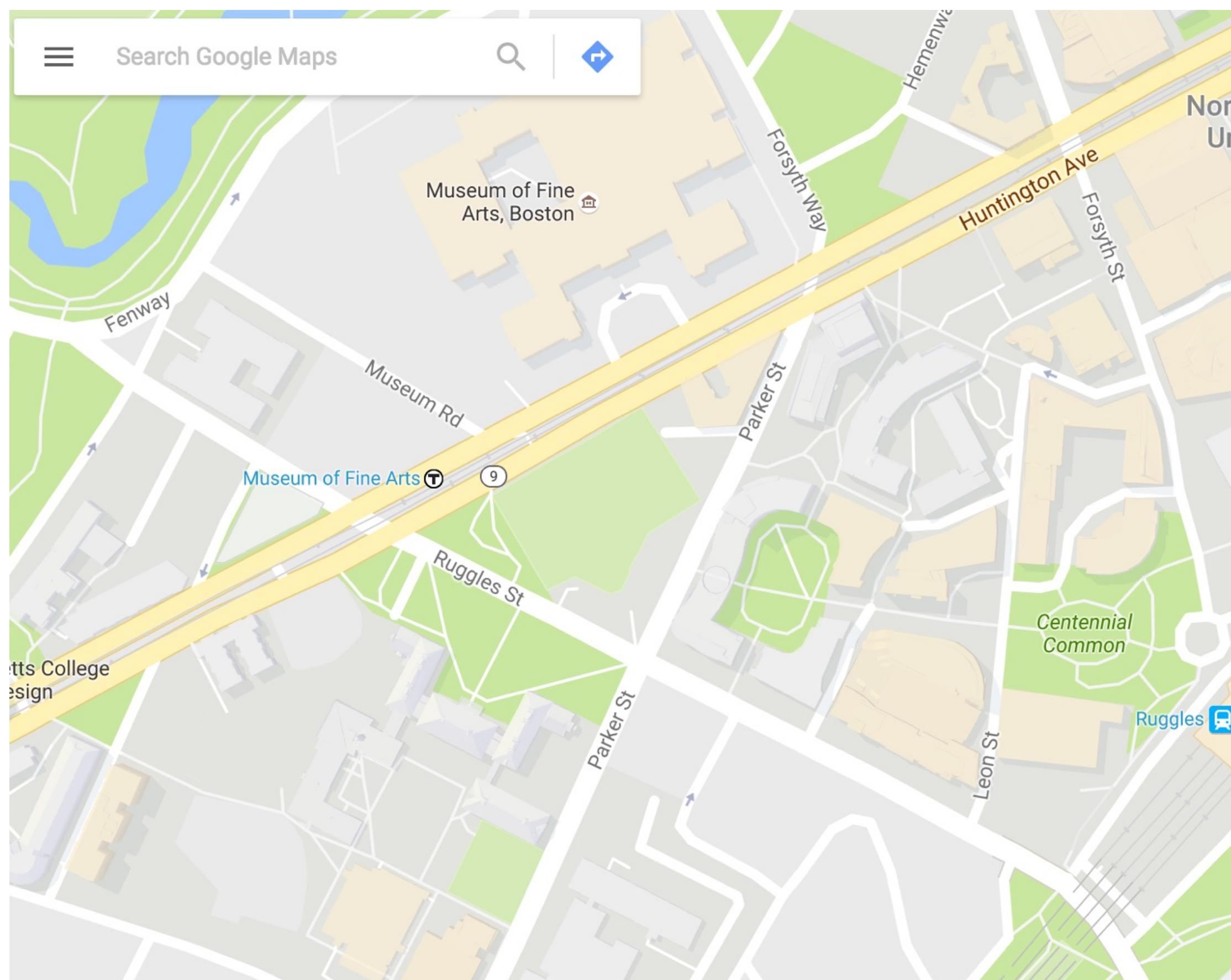


How much profit was there for Mississippi?

Task Abstraction

What?
Why?
How?

Mid-level → Search → Locate



→ Search

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

Where is Ryder Hall?

Task Abstraction

What?

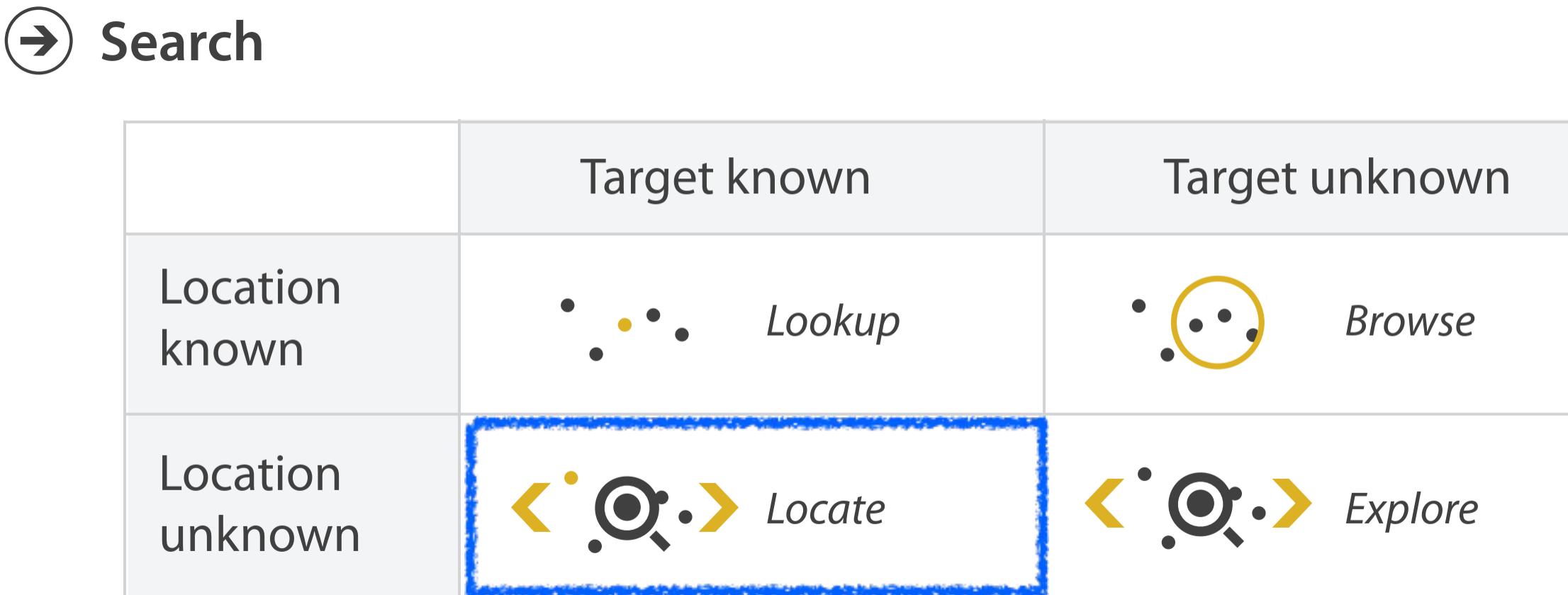
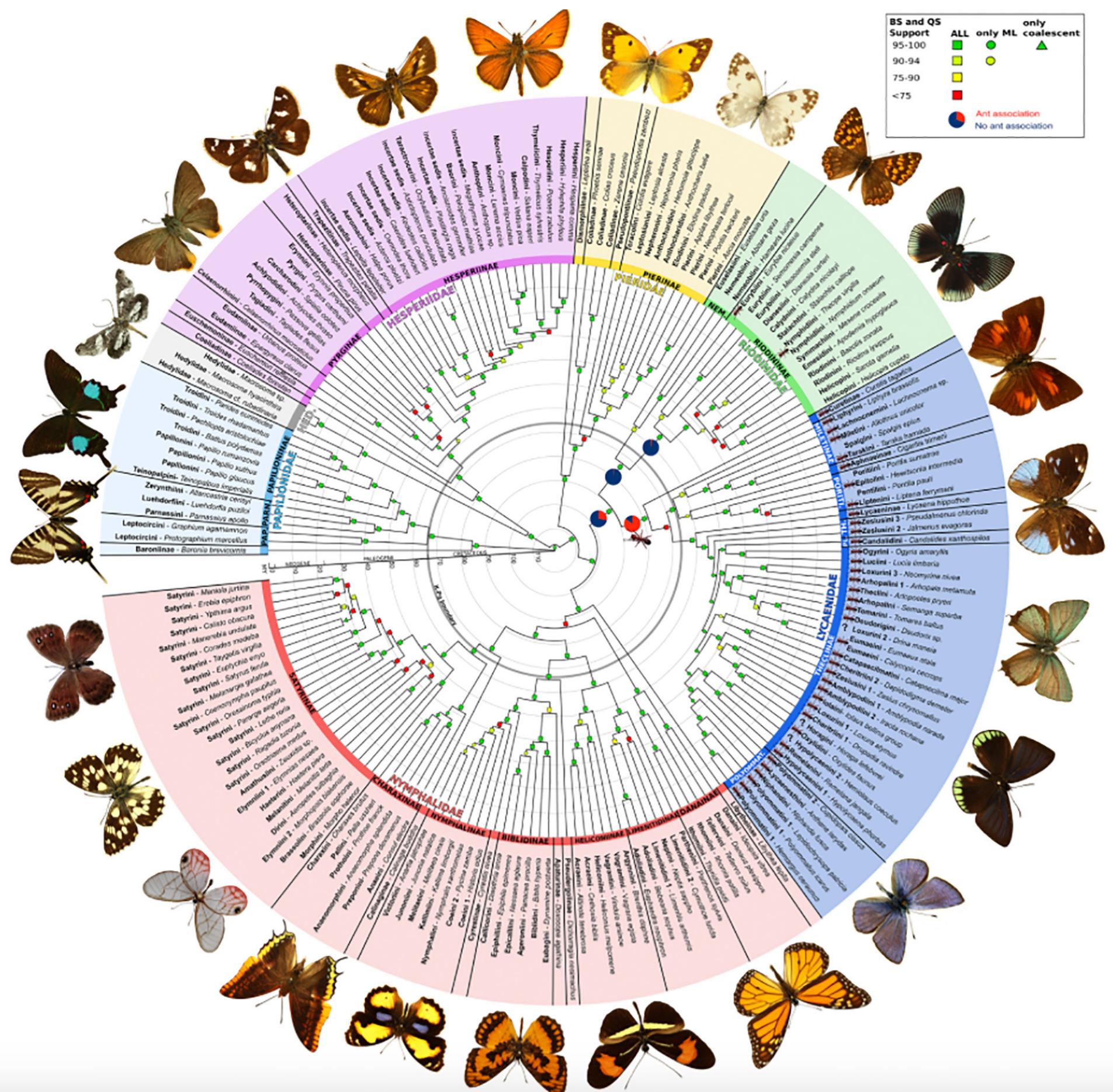
Why?

How?

Why?

How?

Mid-level* → *Search* → *Locate

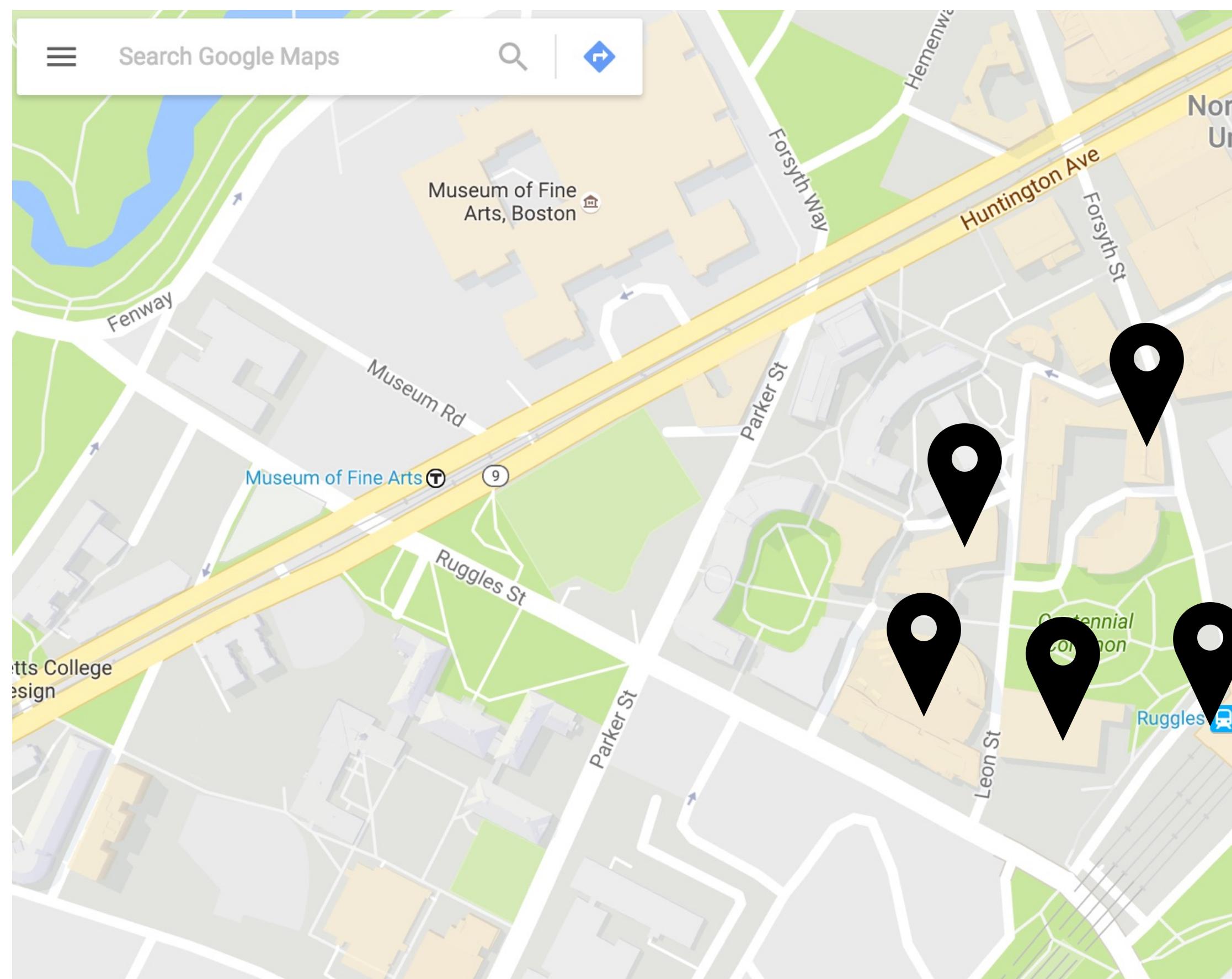


Where is the Monarch butterfly?

Task Abstraction

What?
Why?
How?

Mid-level → Search → Browse

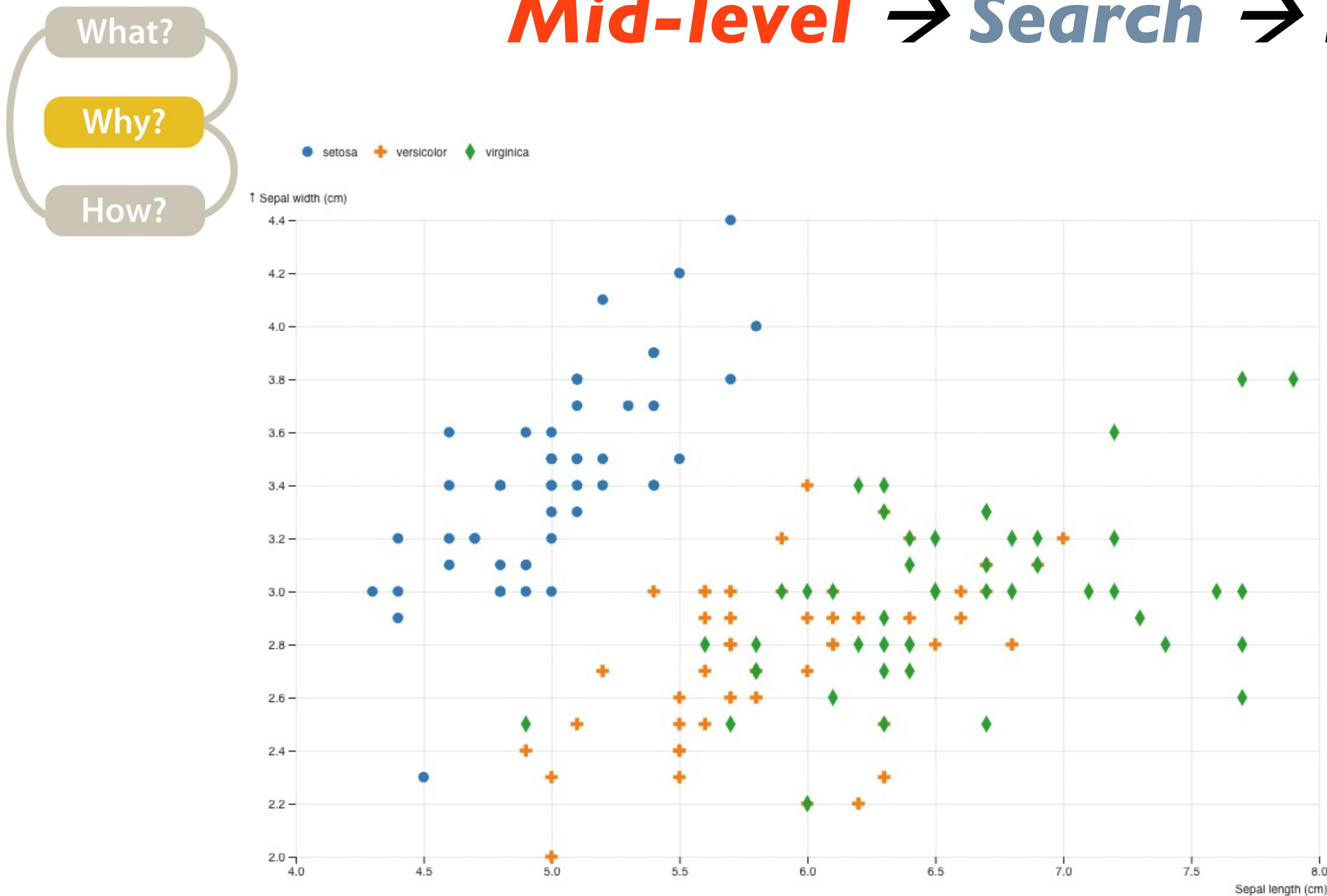


→ Search

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

What buildings are near
Ryder Hall?

Task Abstraction



Mid-level → Search → Browse

→ Search

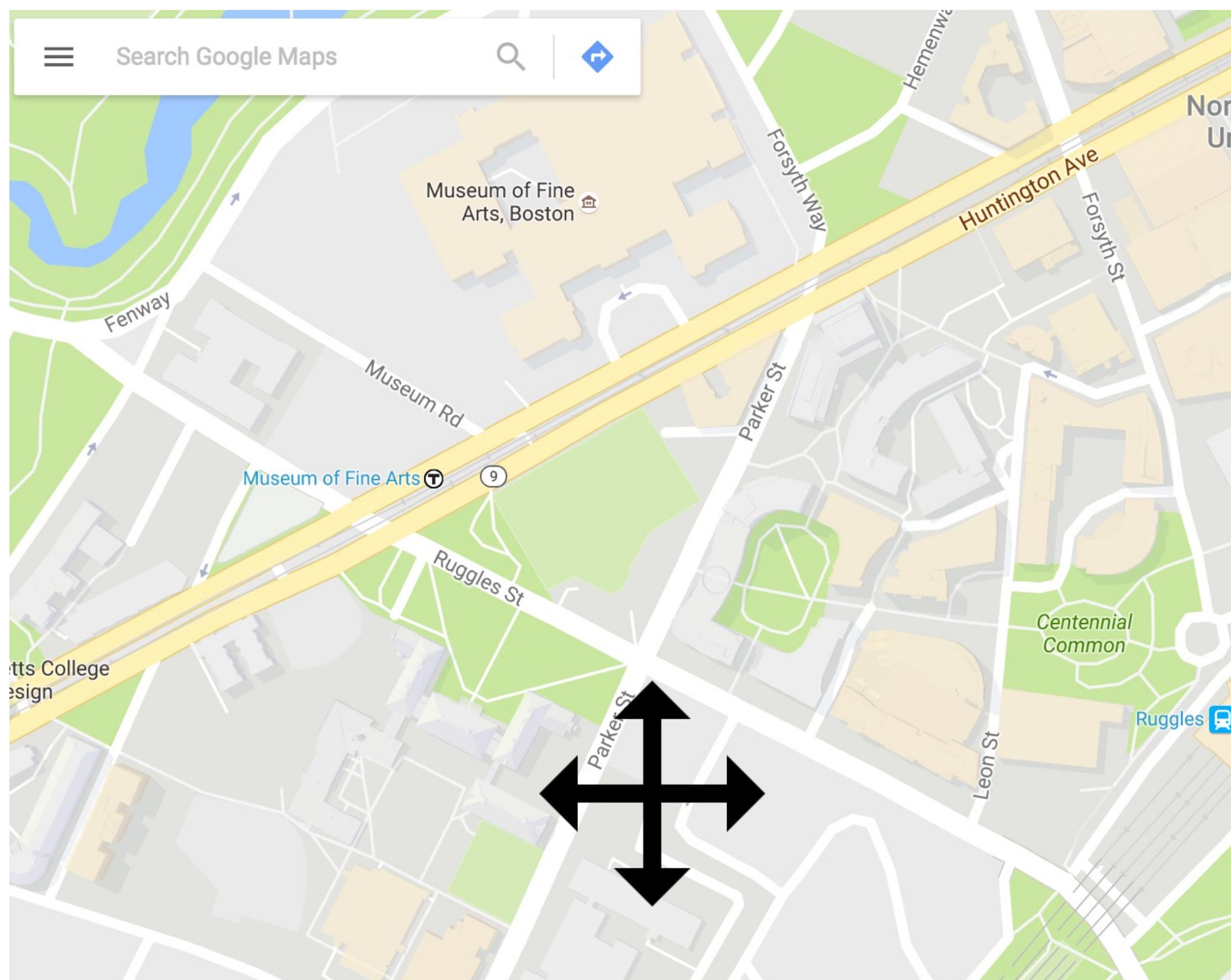
	Target known	Target unknown
Location known	Lookup	Browse
Location unknown	Locate	Explore

Which Irises have sepal length greater than 7.5?

Task Abstraction

What?
Why?
How?

Mid-level → Search → Explore

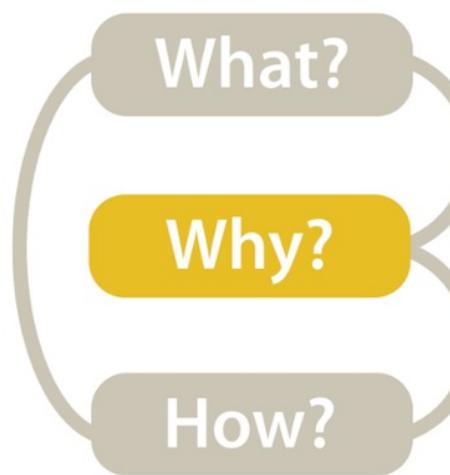


→ Search

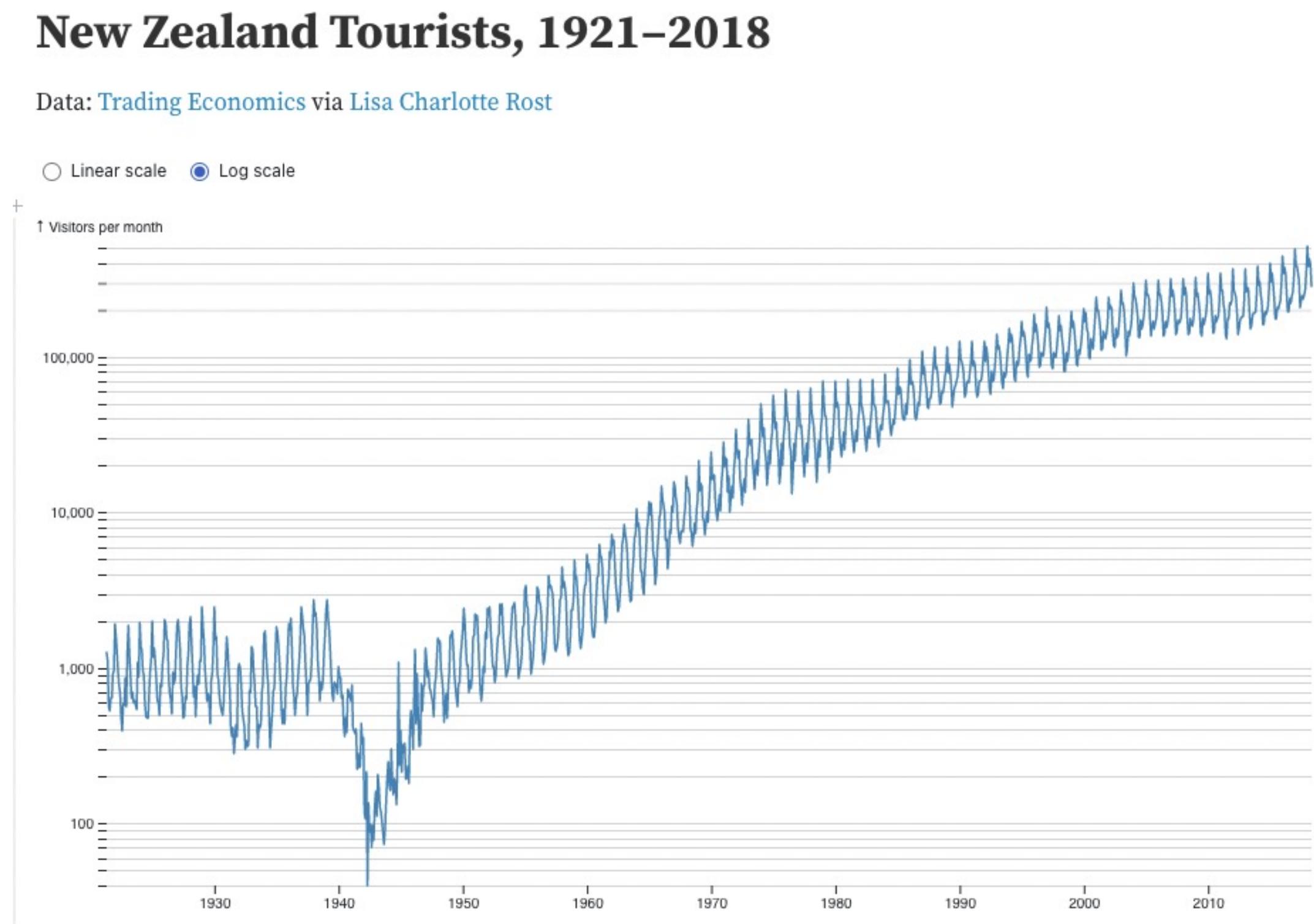
	Target known	Target unknown
Location known	Lookup	Browse
Location unknown	Locate	Explore

Where can I study?

Task Abstraction



Mid-level → Search → Explore



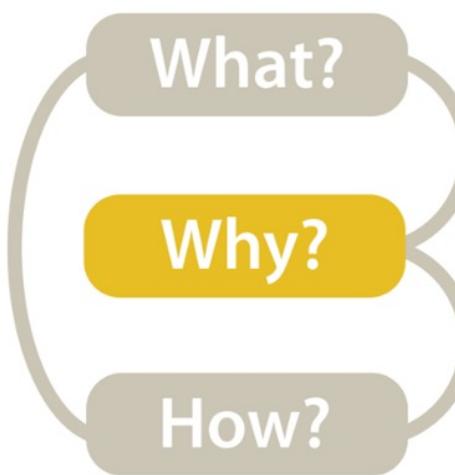
→ Search

	Target known	Target unknown
Location known	• • • <i>Lookup</i>	• • • <i>Browse</i>
Location unknown	◁ • ▷ <i>Locate</i>	◁ • ▷ <i>Explore</i>

Is there a seasonal trend in tourism?

Are there any unusual years?

Task Abstraction

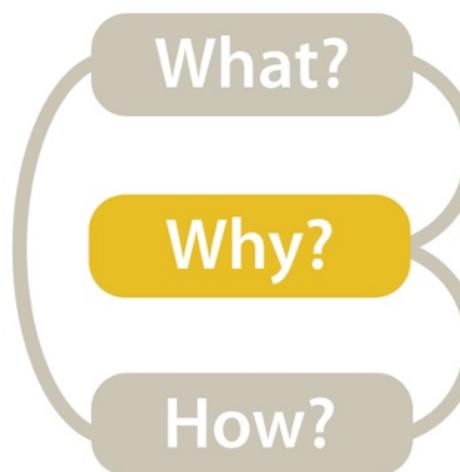


Mid-level → Search

→ **Search**

	Target known	Target unknown
Location known	 <i>Lookup</i>	 <i>Browse</i>
Location unknown	 <i>Locate</i>	 <i>Explore</i>

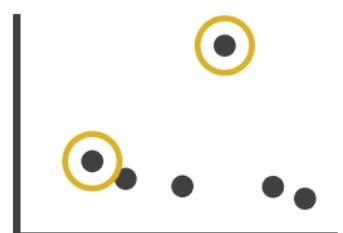
Task Abstraction



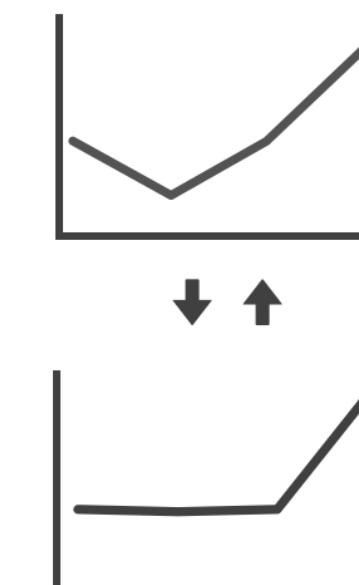
Low-level → Query

→ **Query**

→ Identify



→ Compare



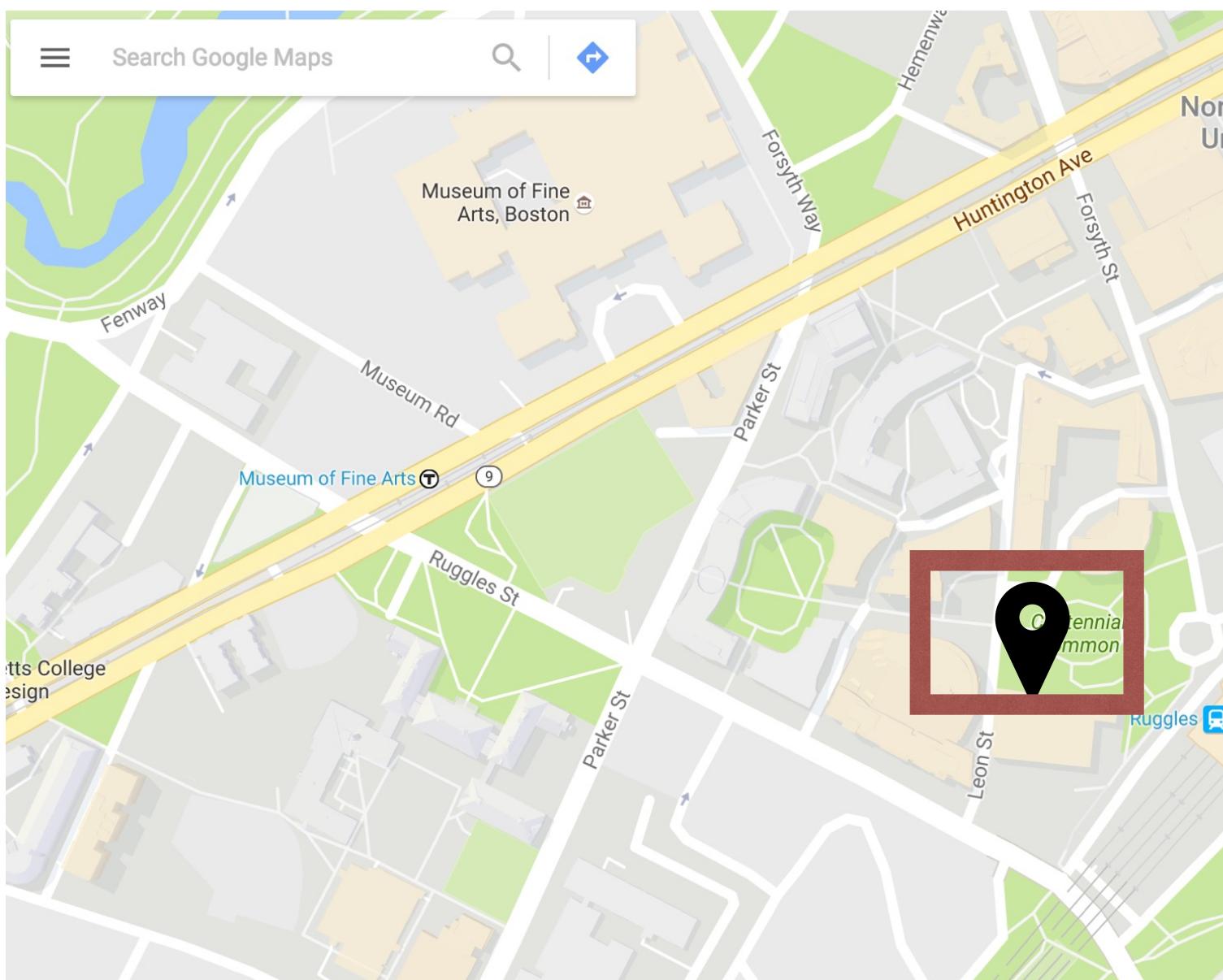
→ Summarize



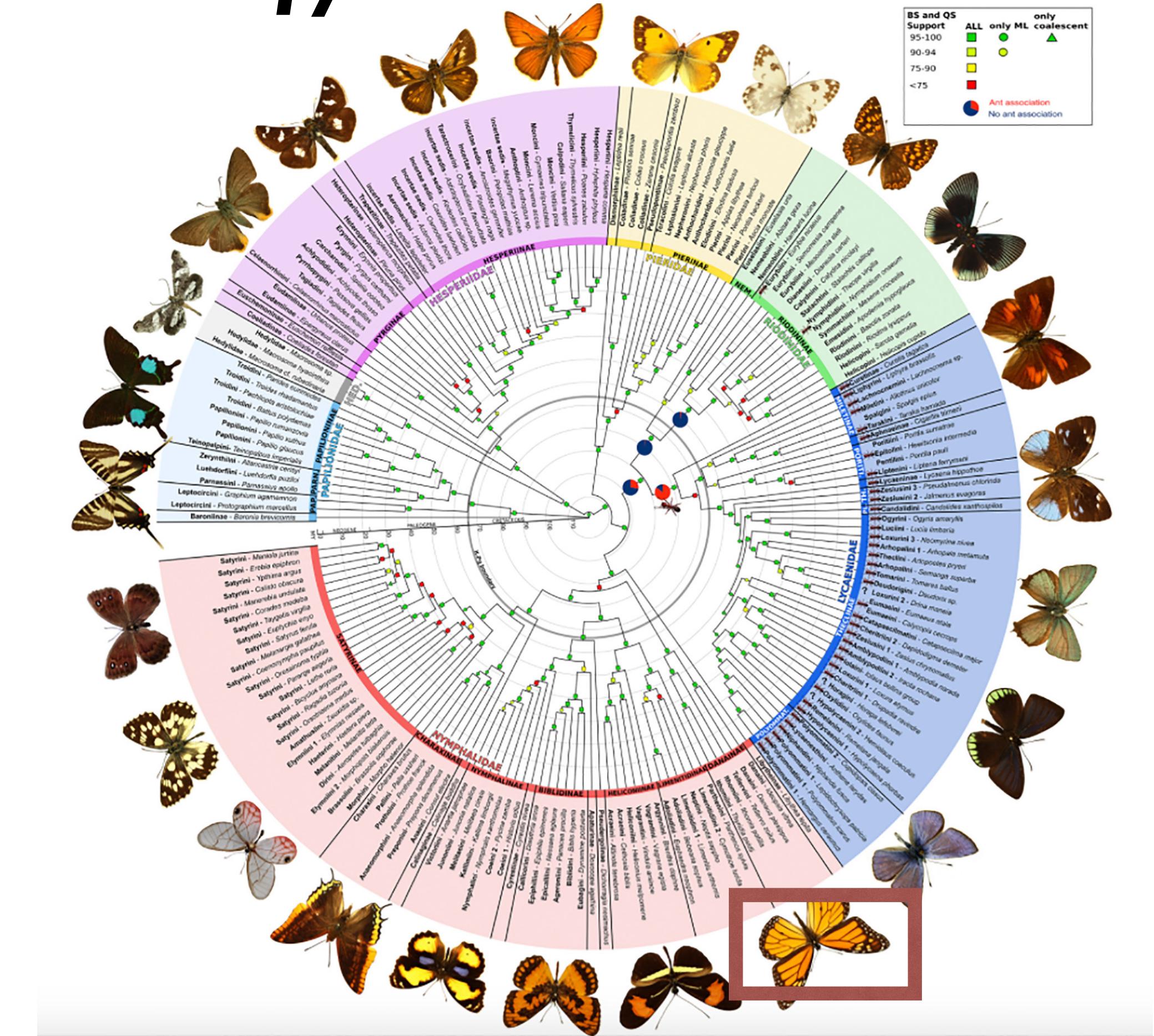
Task Abstraction

What?
Why?
How?

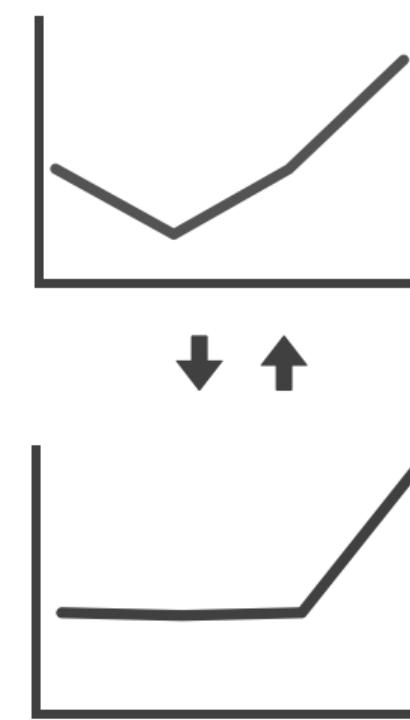
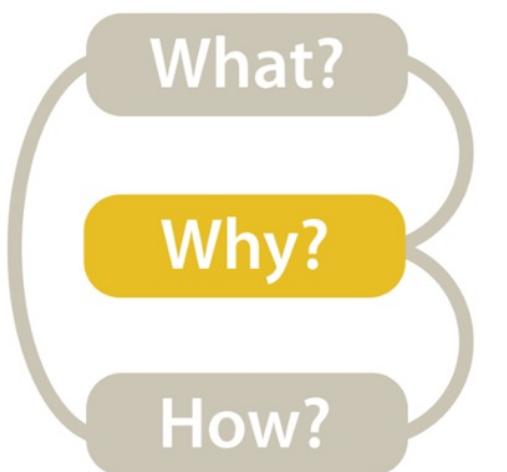
Low-level → Query → Identify



single target

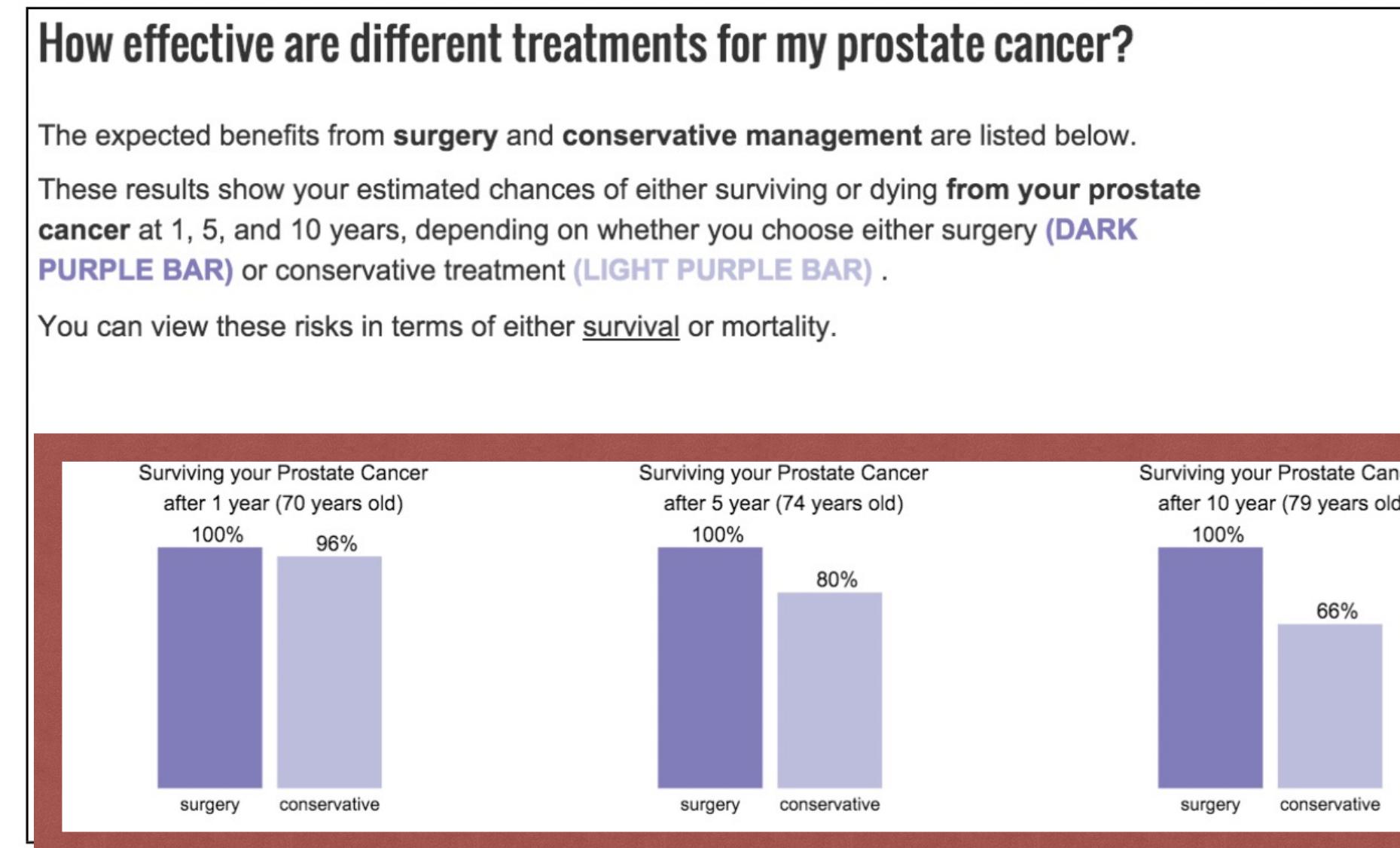


Task Abstraction



multiple targets

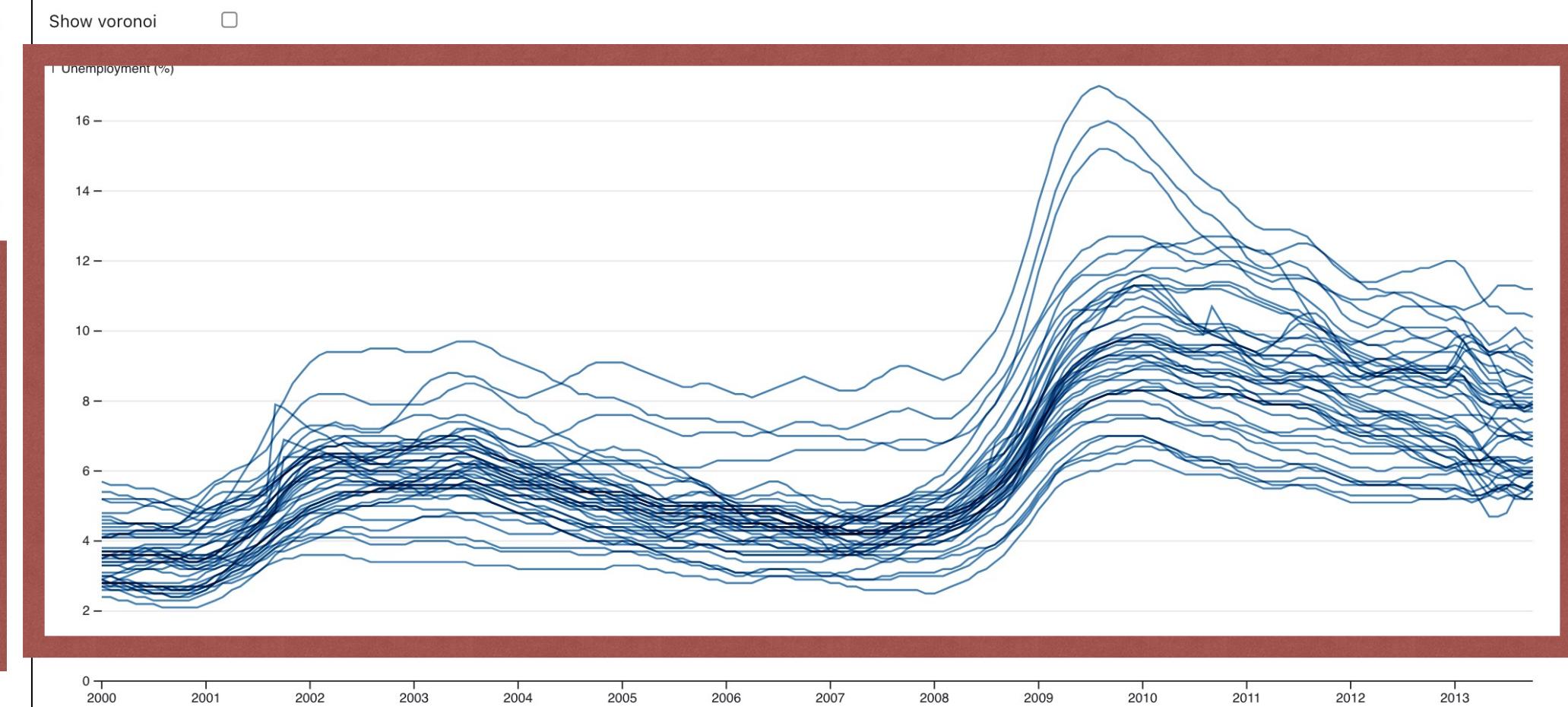
Low-level → Query → Compare



PROACT: Iterative Design of a Patient-Centered Visualization for Effective Prostate Cancer Health Risk Communication. Hakone et al. 2016.

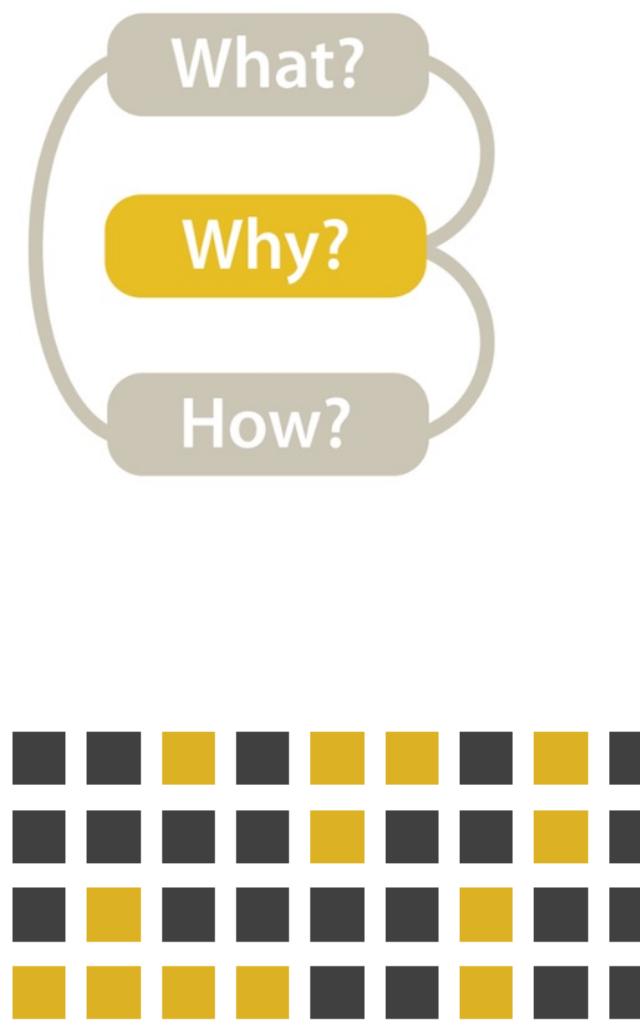
Line Chart, Multiple Series

This [line chart](#) shows the unemployment rate of various U.S. metro divisions from 2000 through 2013. On hover, the closest data point to the pointer and its associated series is highlighted. Data: [Bureau of Labor Statistics](#)

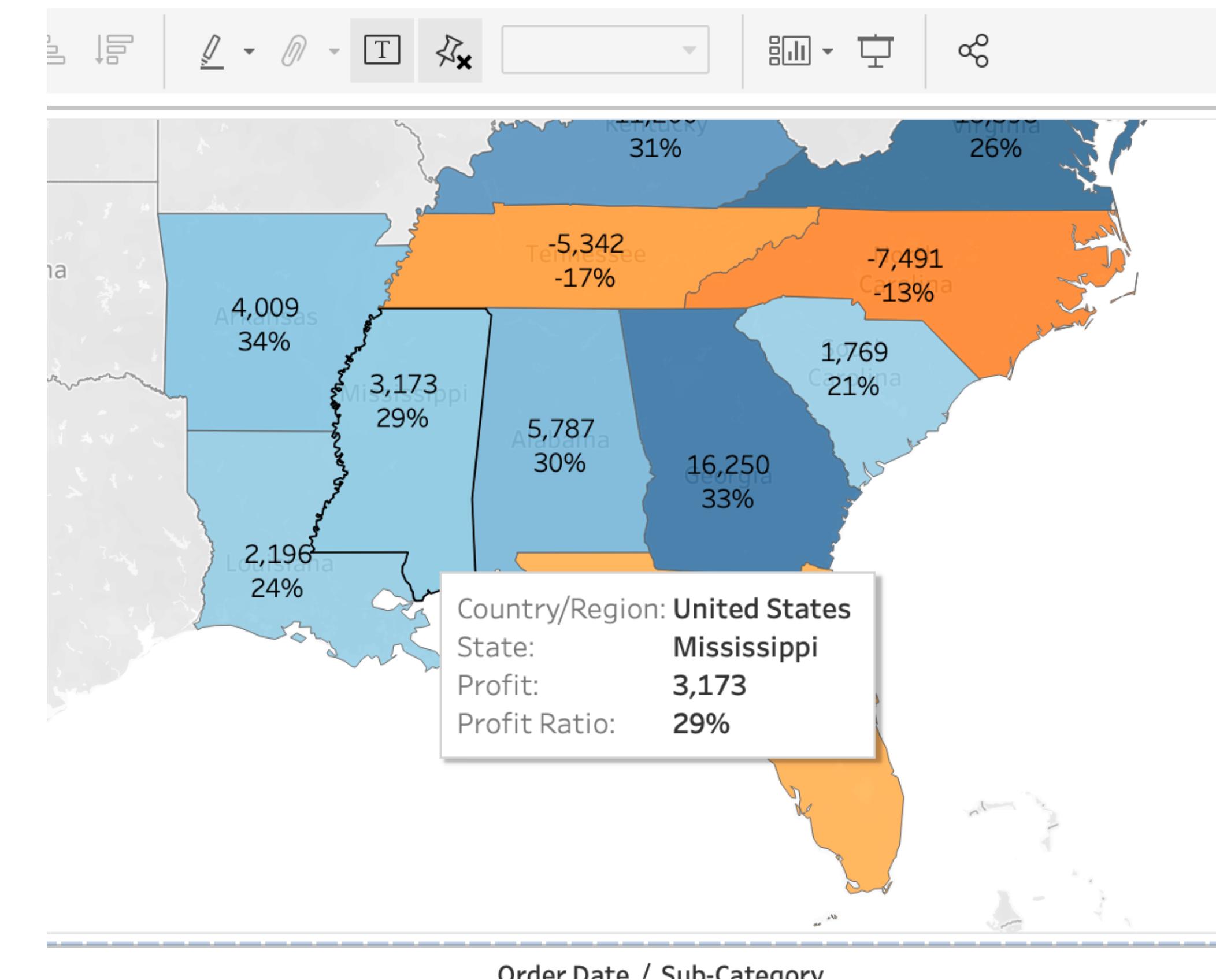
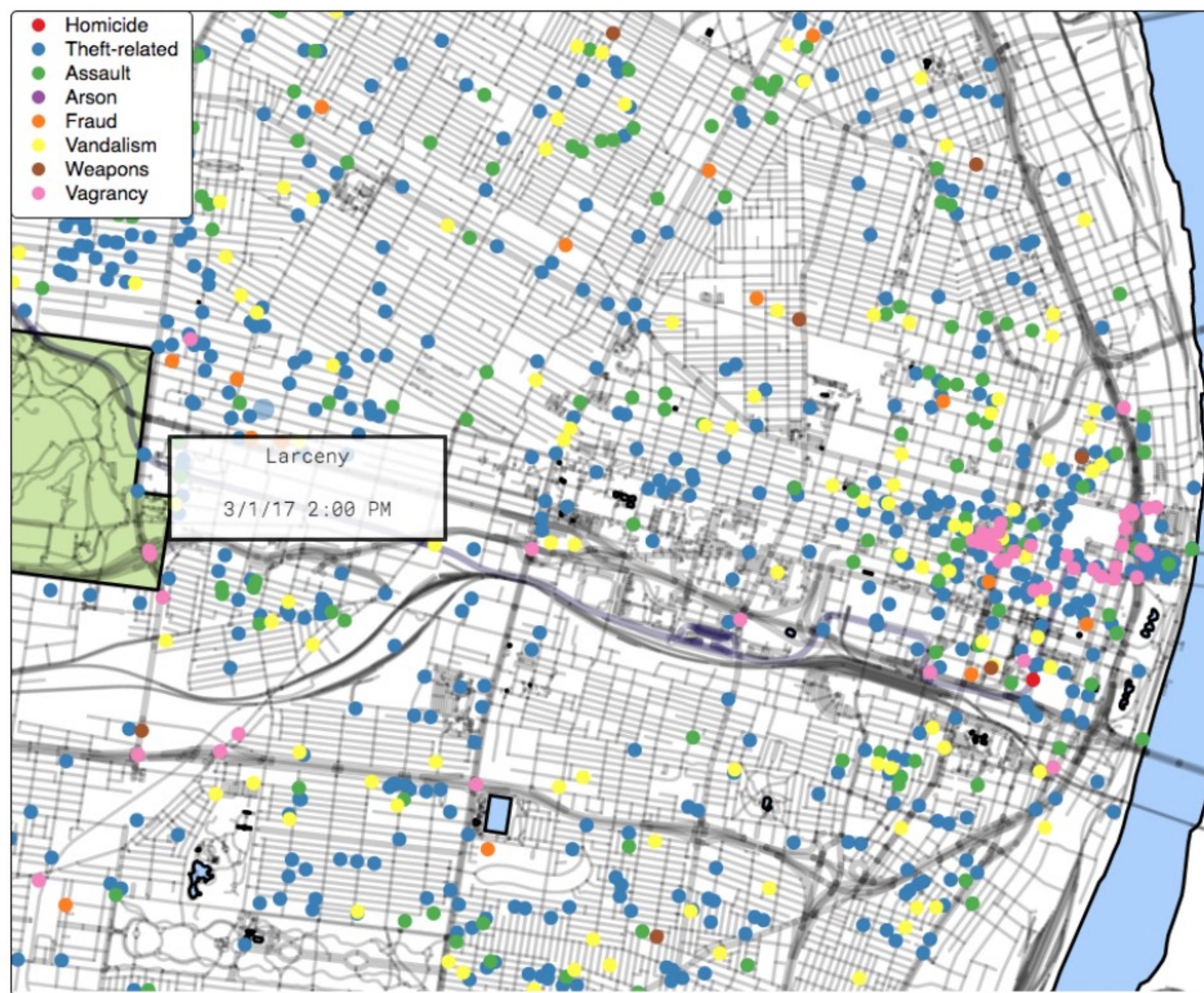


<https://observablehq.com/@d3/multi-line-chart>

Task Abstraction

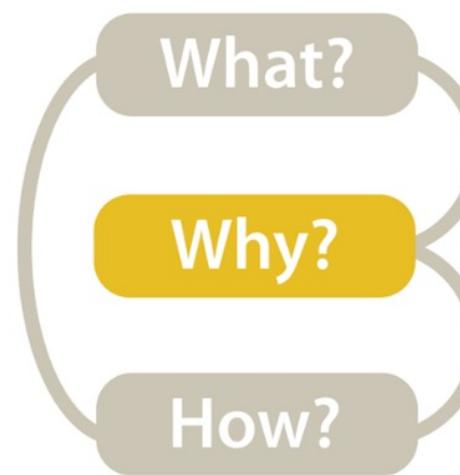


Low-level → Query → Summarize



Follow The Clicks: Learning and Anticipating Mouse Interactions During Exploratory Data Analysis. Ottley et al. 2019

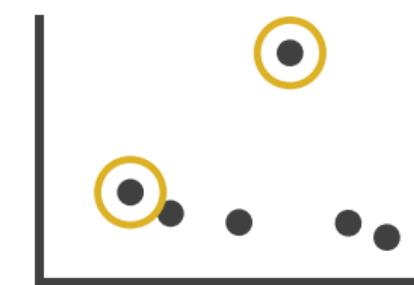
Task Abstraction



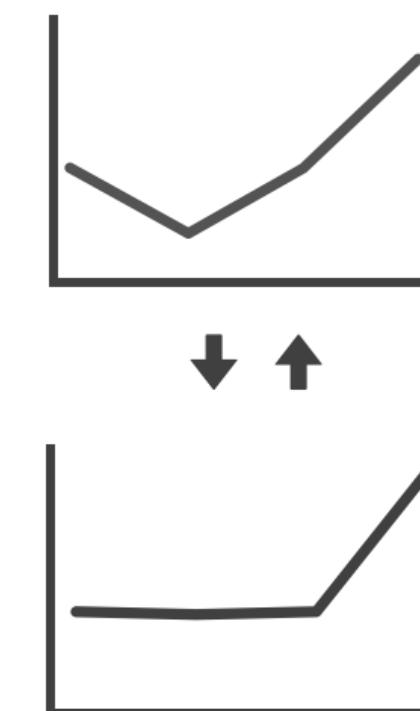
Low-level → Query

→ **Query**

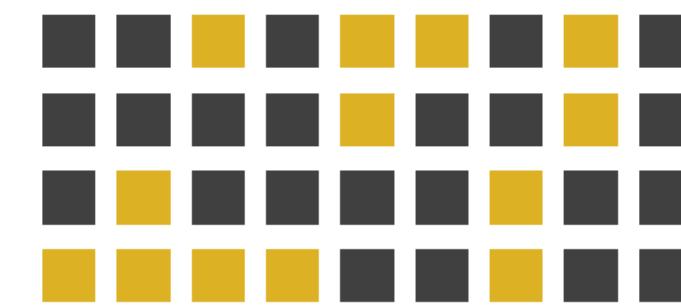
→ Identify



→ Compare



→ Summarize



single target

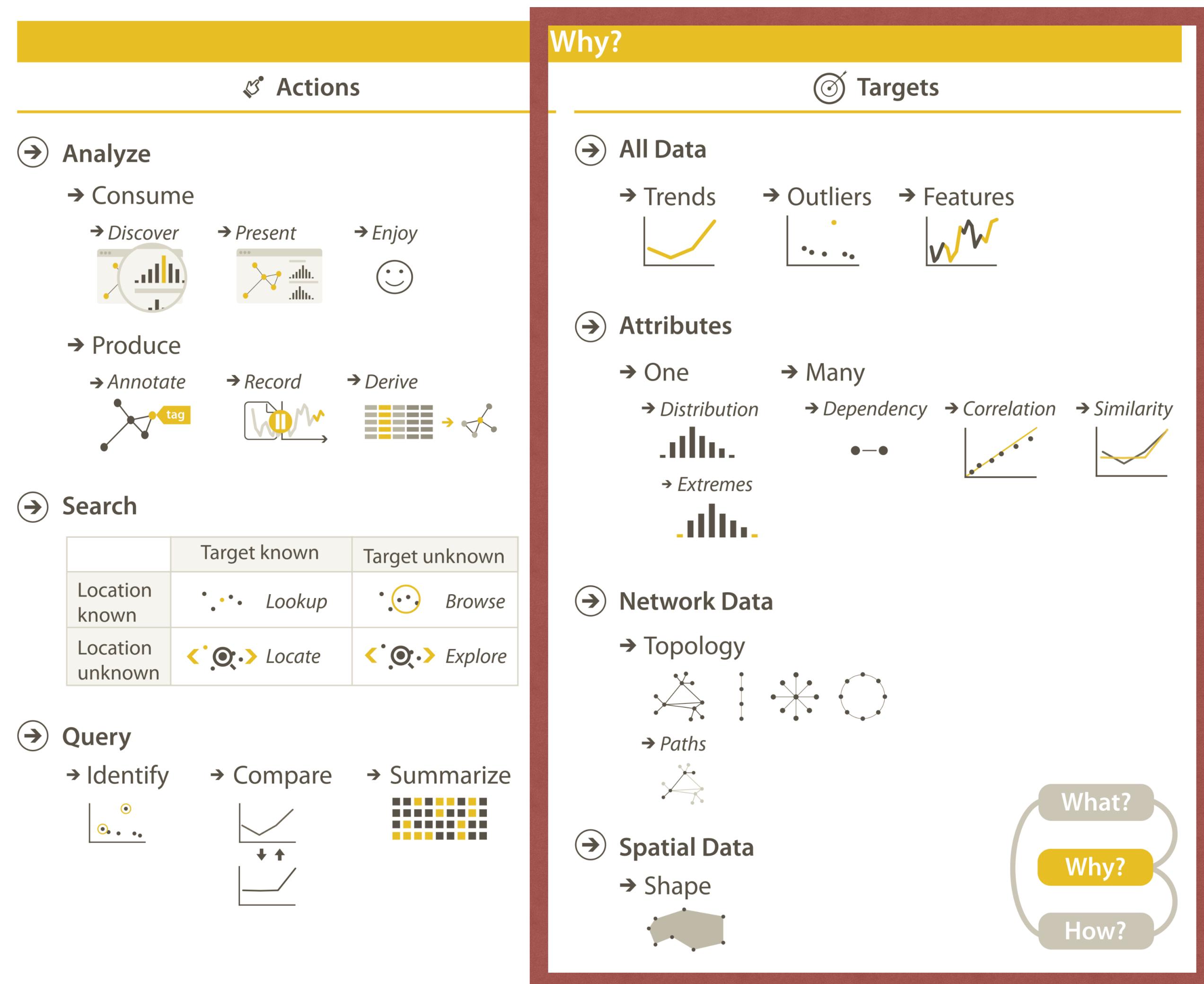
multiple targets

all targets

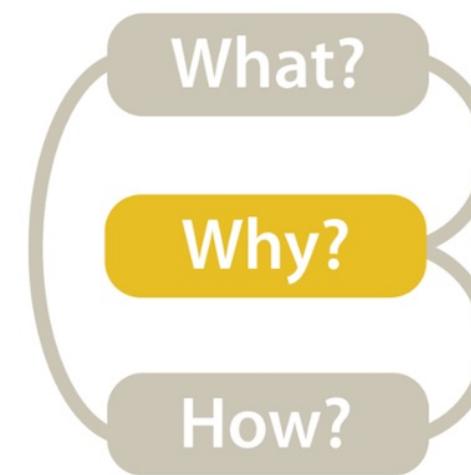
Task Abstraction

TARGETS

Some aspect of
the data of
interest to the
user



Task Abstraction



Targets → some aspect of data of interest to the user

→ All Data

→ Trends



→ Outliers

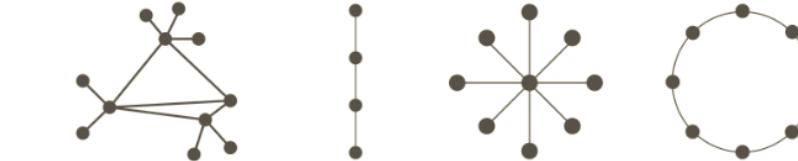


→ Features



→ Network Data

→ Topology



→ Paths

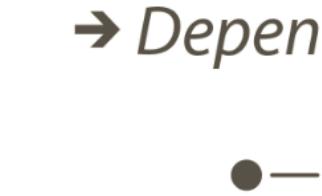


→ Attributes

→ One



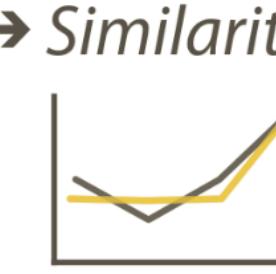
→ Many



→ Distribution

→ Dependency

→ Correlation



→ Extremes



→ Spatial Data

→ Shape



Task Abstraction

What?
Why?
How?

Targets → some aspect of data of interest to the user



→ Attributes

→ One

→ Distribution



→ Many

→ Dependency



→ Correlation



→ Similarity

→ Network Data

→ Topology



→ Paths

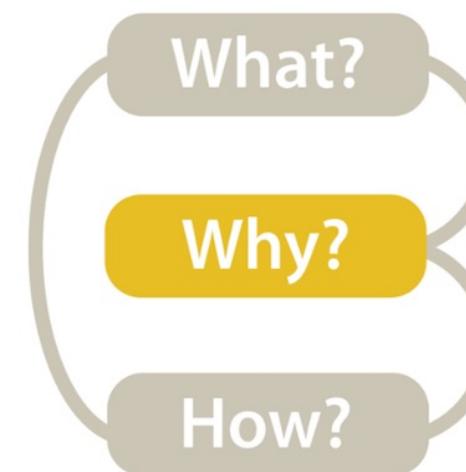


→ Spatial Data

→ Shape



Task Abstraction



Targets → some aspect of data of interest to the user

→ All Data

→ Trends



→ Outliers

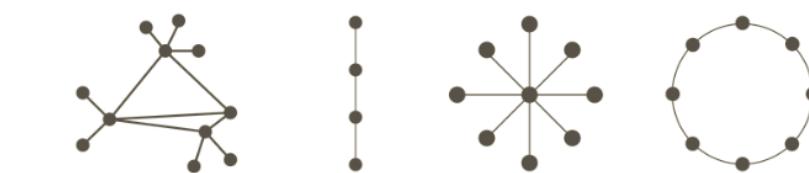


→ Features



→ Network Data

→ Topology



→ Paths



→ Spatial Data

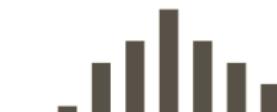
→ Shape



→ Attributes

→ One

→ Distribution



→ Many

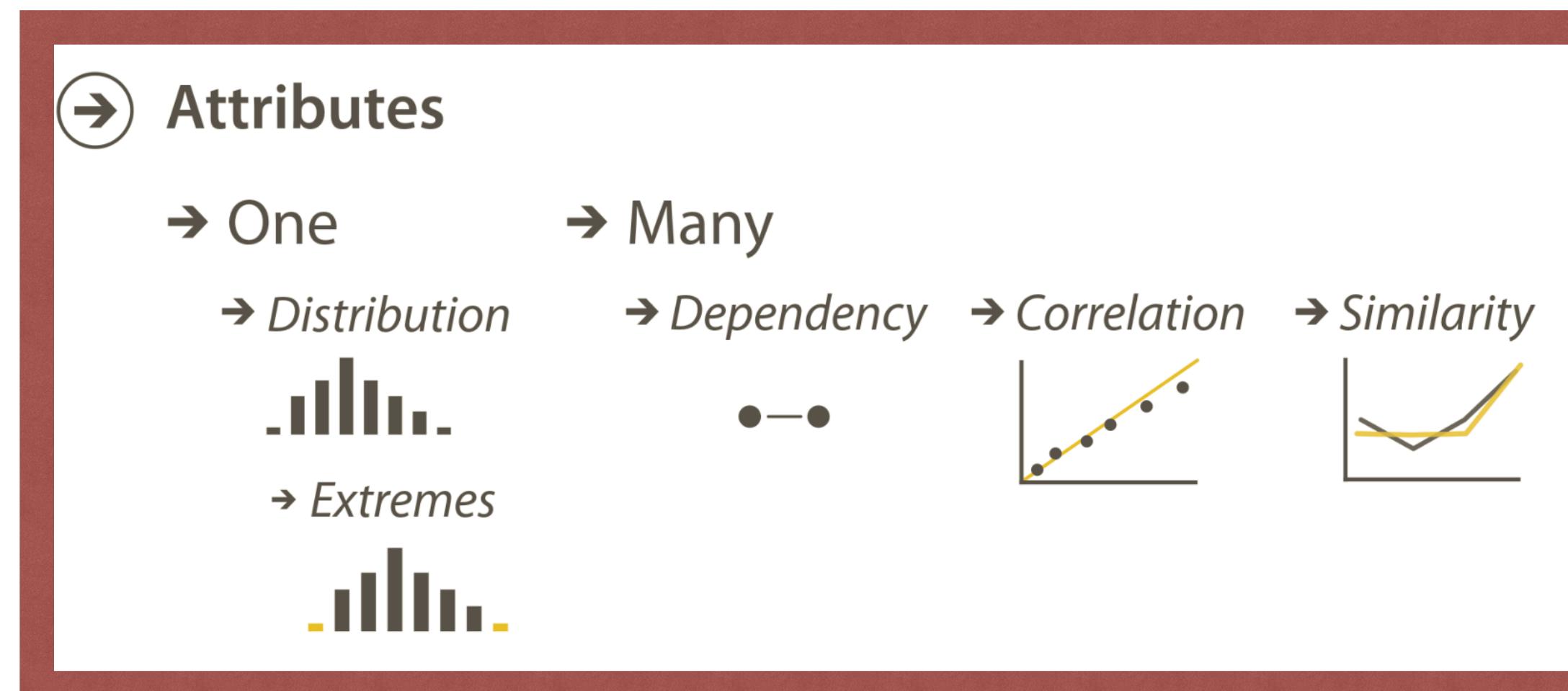
→ Dependency



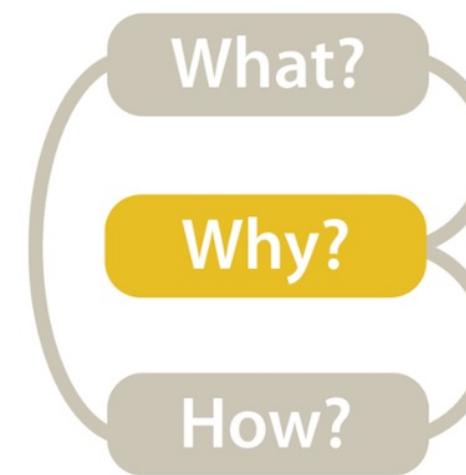
→ Correlation



→ Similarity



Task Abstraction



Targets → some aspect of data of interest to the user

→ All Data

→ Trends



→ Outliers



→ Features



→ Attributes

→ One

→ Distribution

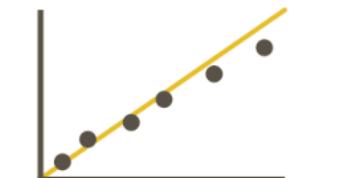


→ Many

→ Dependency



→ Correlation

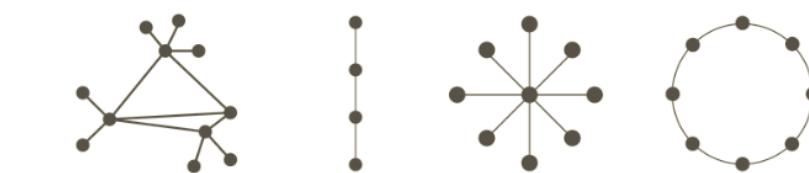


→ Similarity



→ Network Data

→ Topology

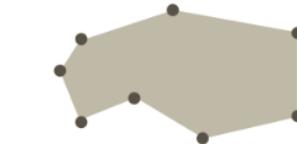


→ Paths

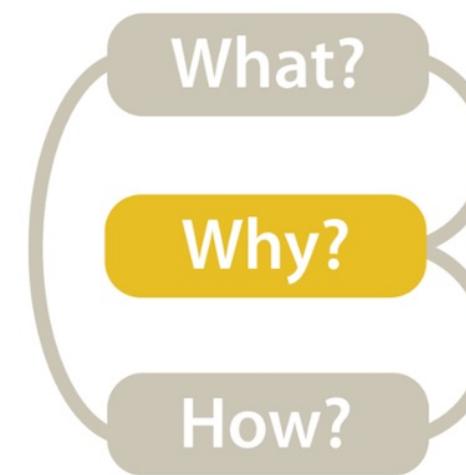


→ Spatial Data

→ Shape



Task Abstraction



Targets → some aspect of data of interest to the user

→ All Data

→ Trends



→ Outliers

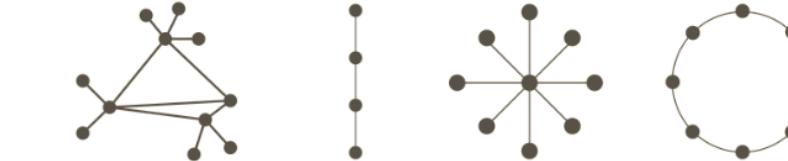


→ Features



→ Network Data

→ Topology



→ Paths



→ Attributes

→ One

→ Distribution



→ Extremes

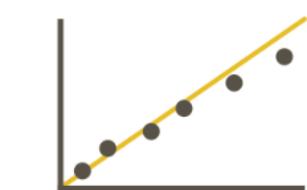


→ Many

→ Dependency



→ Correlation

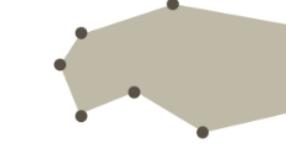


→ Similarity



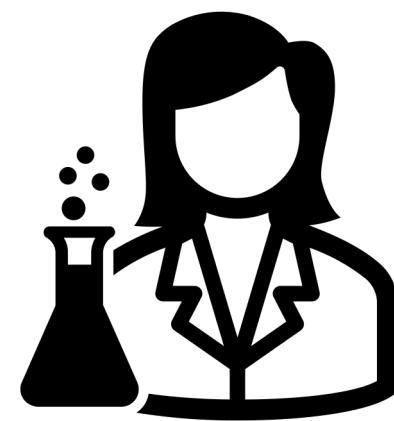
→ Spatial Data

→ Shape



EXAMPLE TASK ABSTRACTION

Task Abstraction

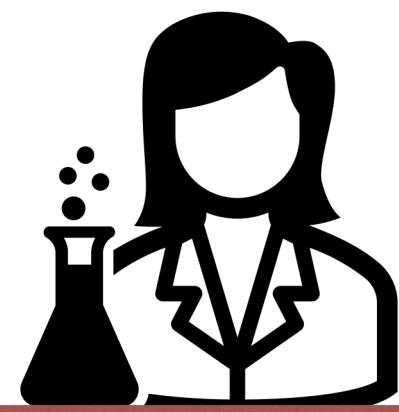
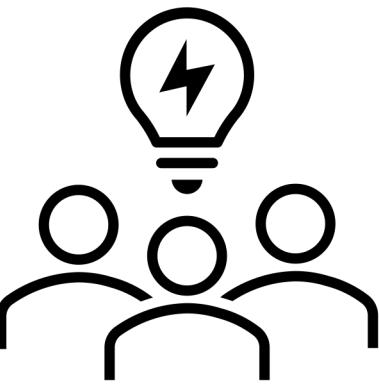


Domain Task: I need to perform cellular analysis!

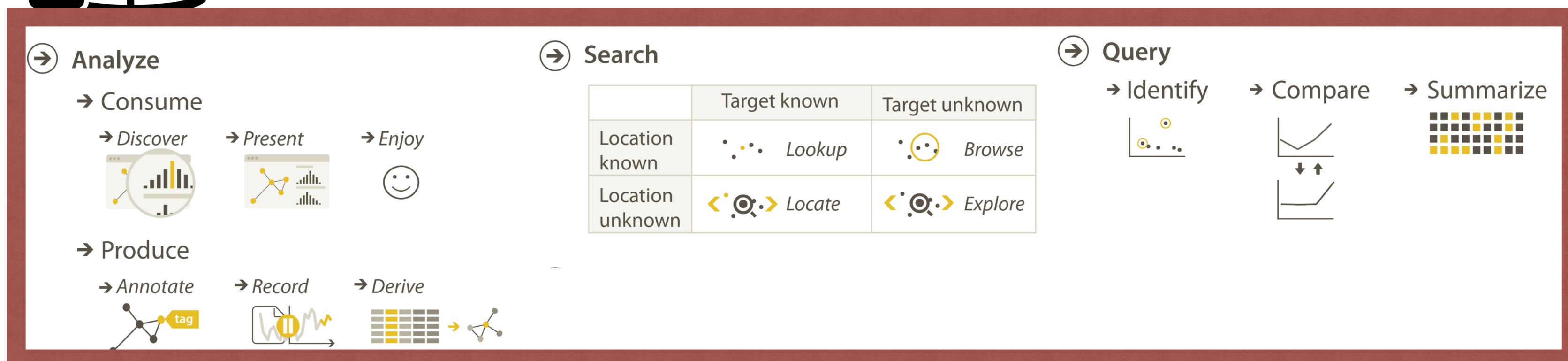


Visualization Task: compare measure A to B over time.

Task Abstraction



Visualization Task: compare measure A to B over time.



High-level →

? →
?

Medium-level →

Search →

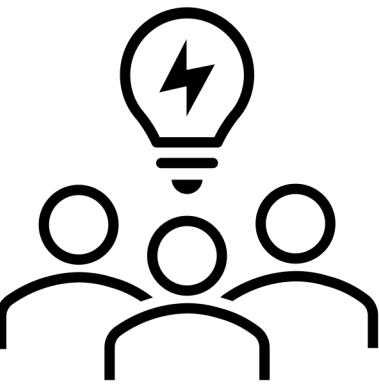
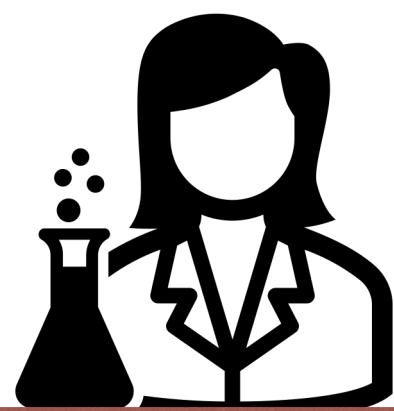
?

Low-level →

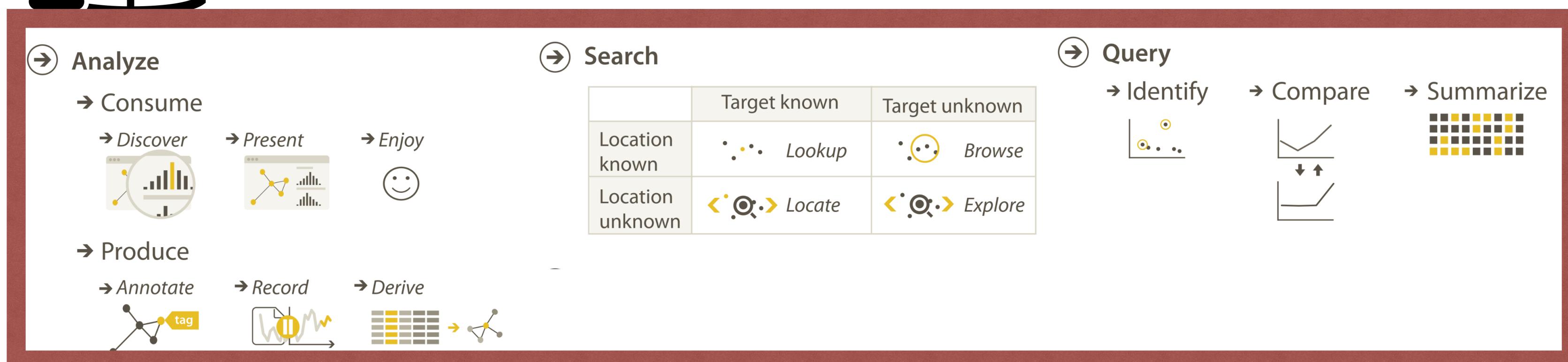
Query →

?

Task Abstraction



Visualization Task: compare measure A to B over time.

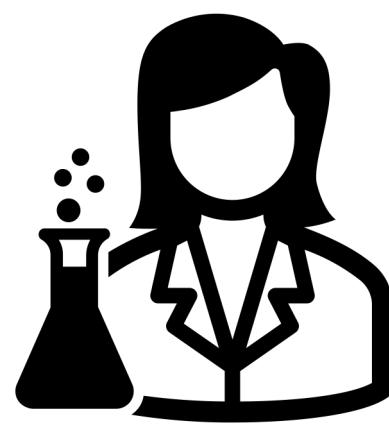


**High-level →
Produce →
Discover**

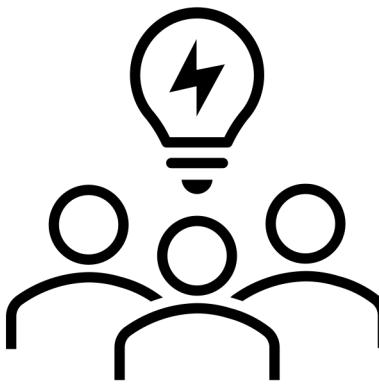
**Medium-level →
Search →
Lookup (or Locate)**

**Low-level →
Query →
Compare**

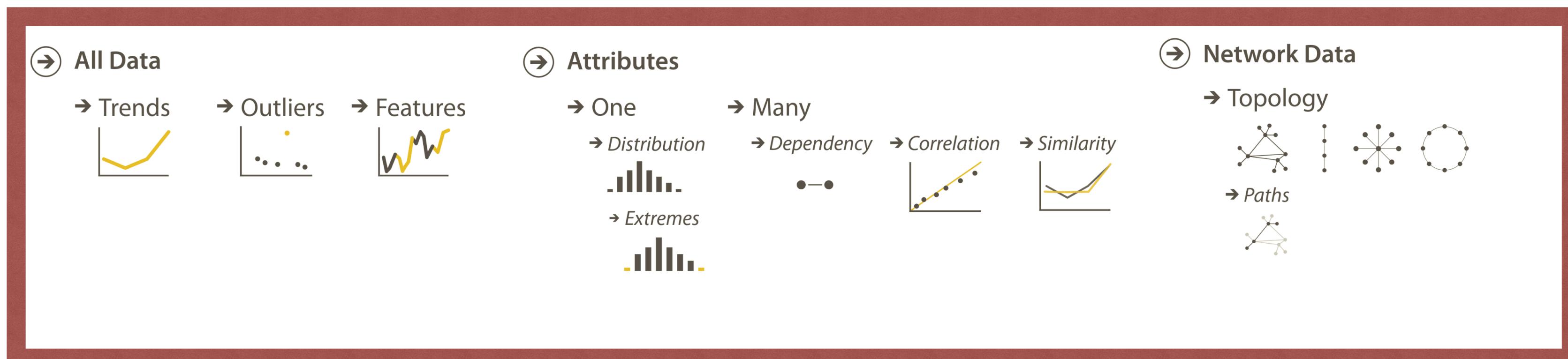
Task Abstraction



Visualization Task: compare measure A to B over time.

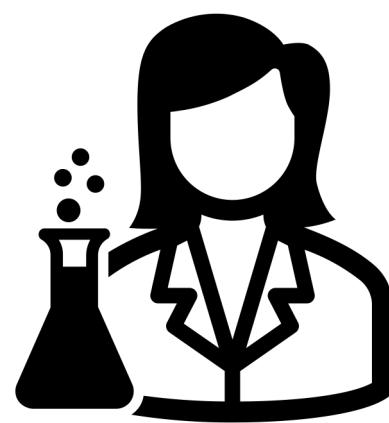


High-level → Produce → Discover
Medium-level → Search → Lookup
Low-level → Query → Compare

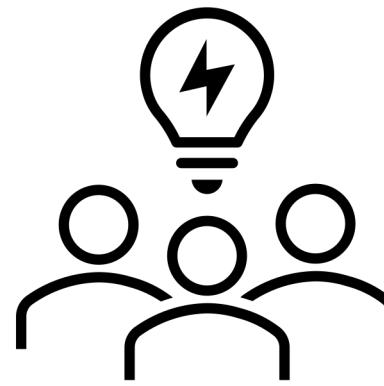


Target(s)

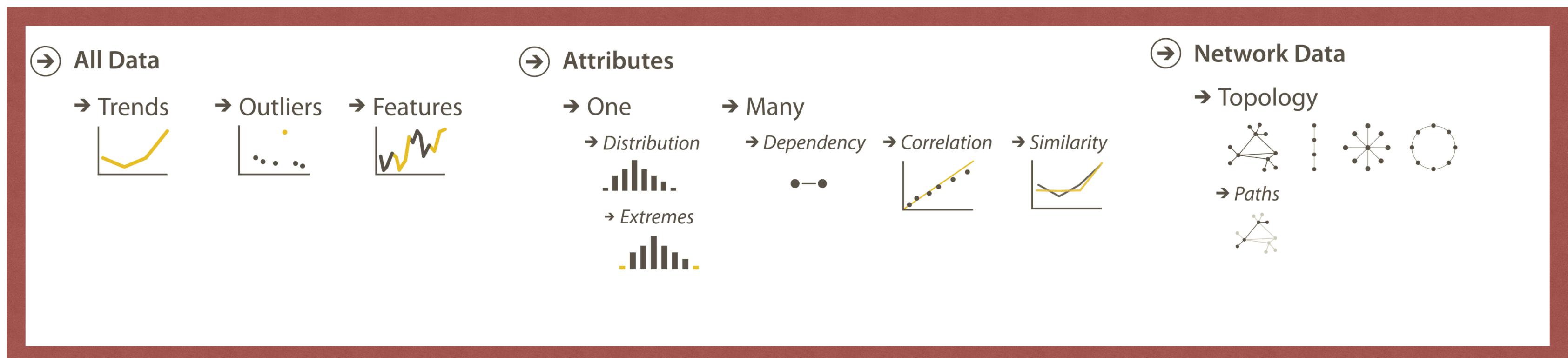
Task Abstraction



Visualization Task: compare measure A to B over time.

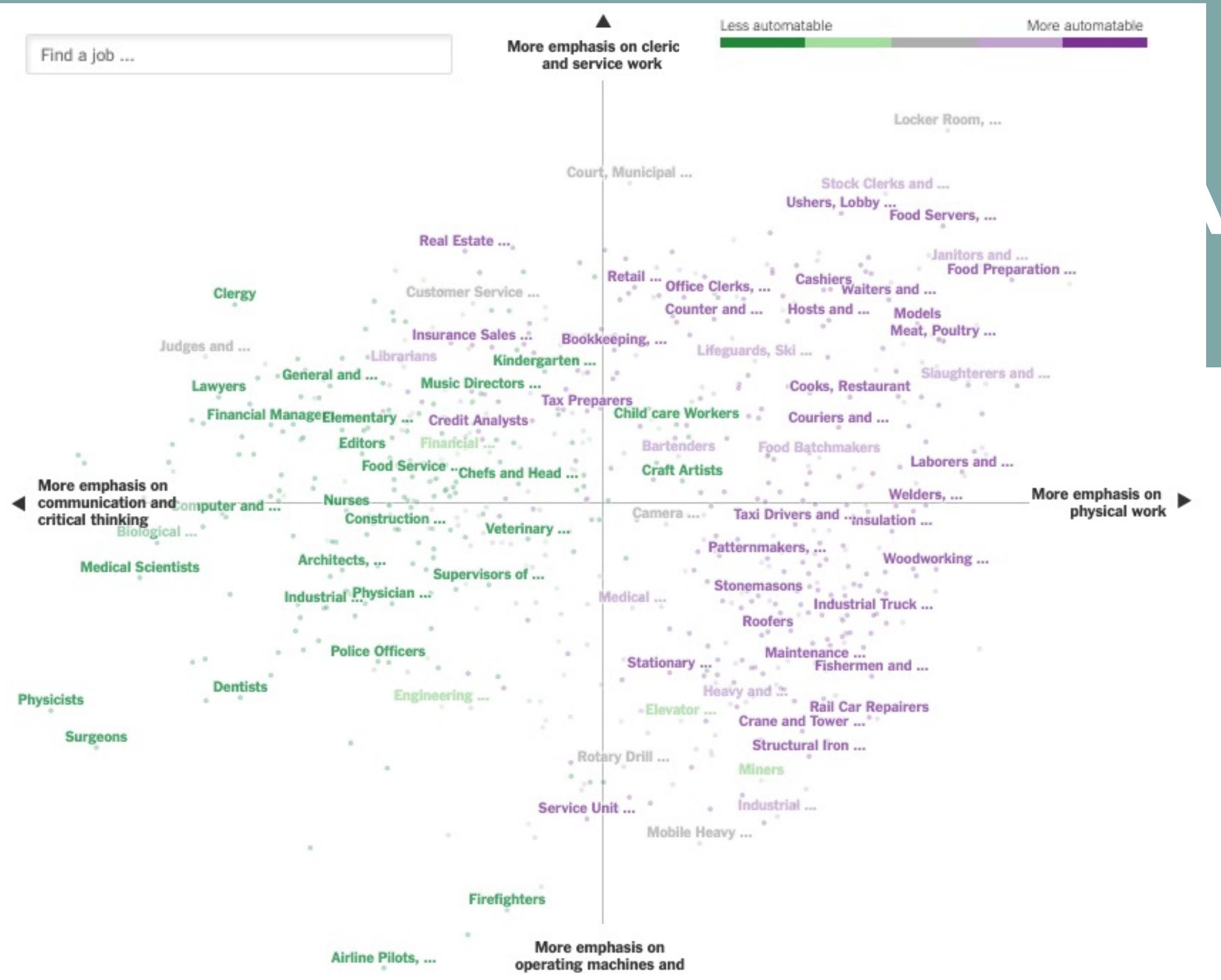


High-level →
Produce →
Discover
Medium-level →
Search →
Lookup
Low-level →
Query →
Compare



Target(s) → Attributes → Similarity

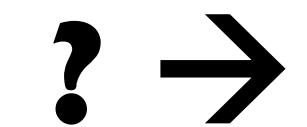
Abstraction



O*Net, Bureau of Labor Statistics, The Future of Employment: How Susceptible are Jobs to Computerisation?, Carl Benedikt Frey and Michael A. Osborne



High-level →



Medium-level →

Search →

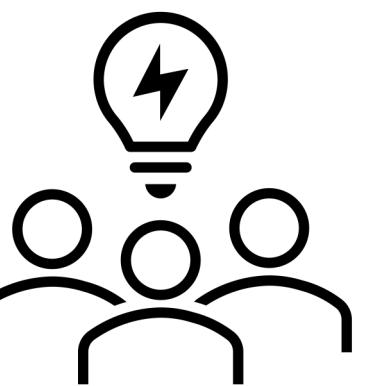
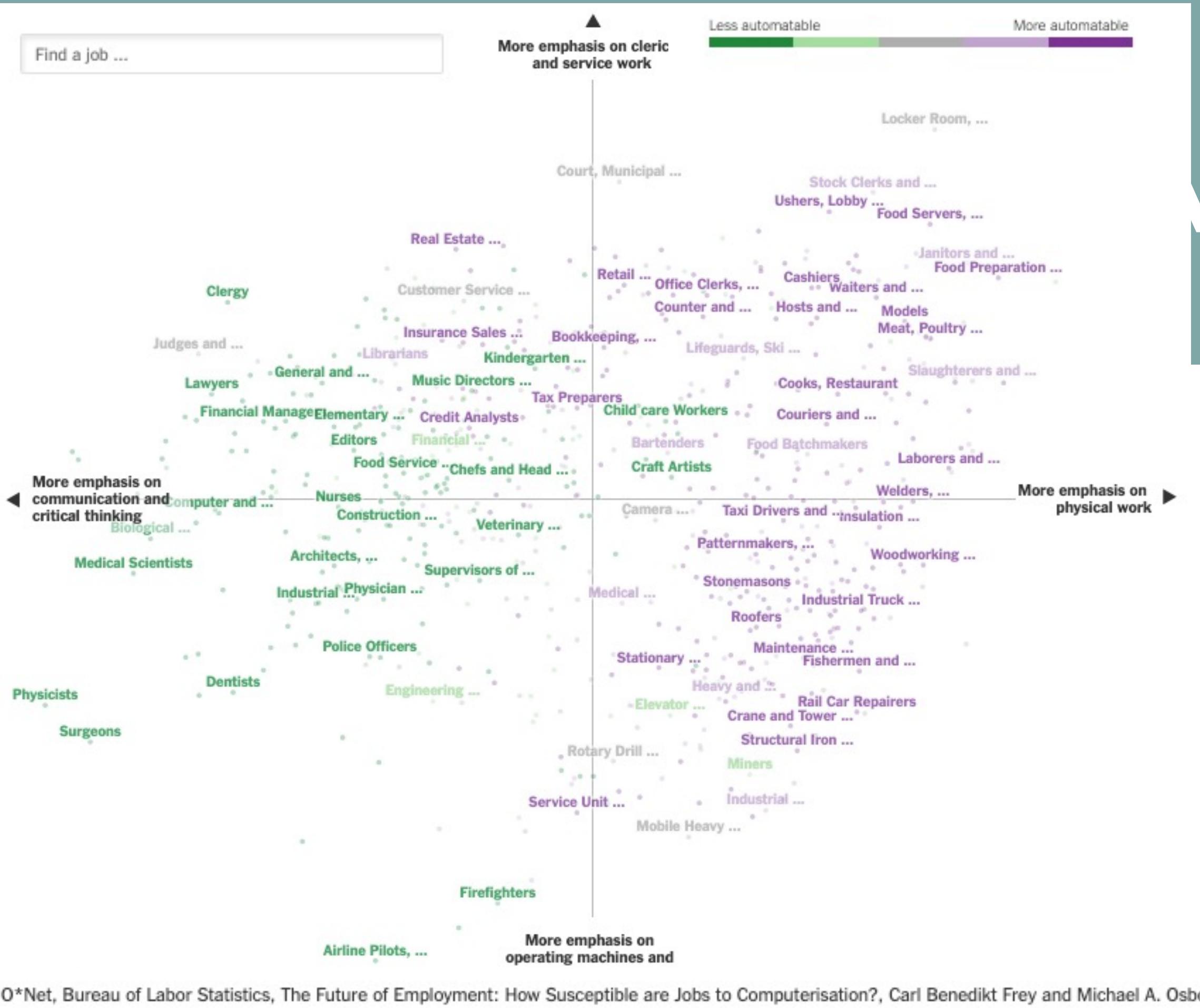


Low-level →

Query →



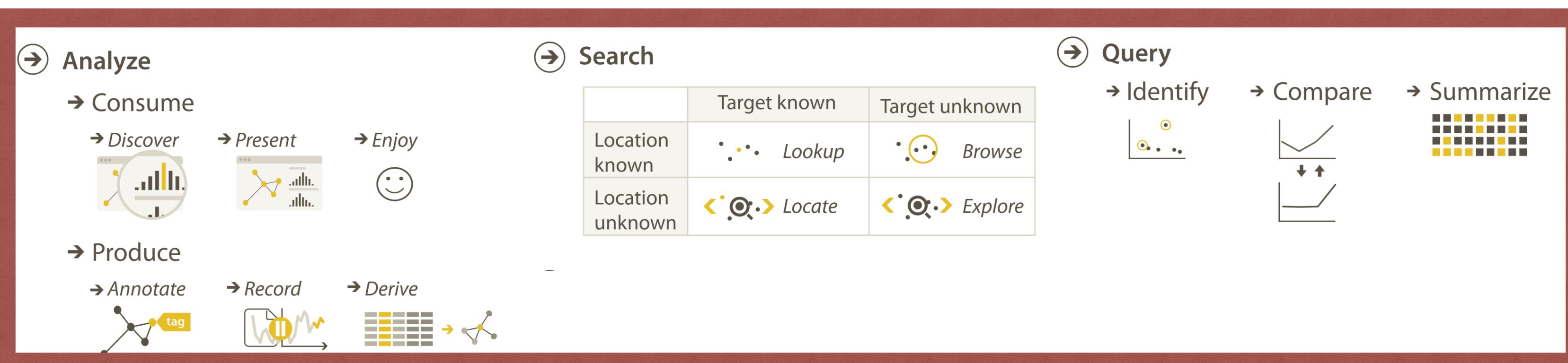
Abstraction



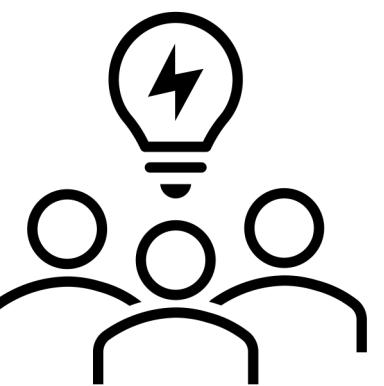
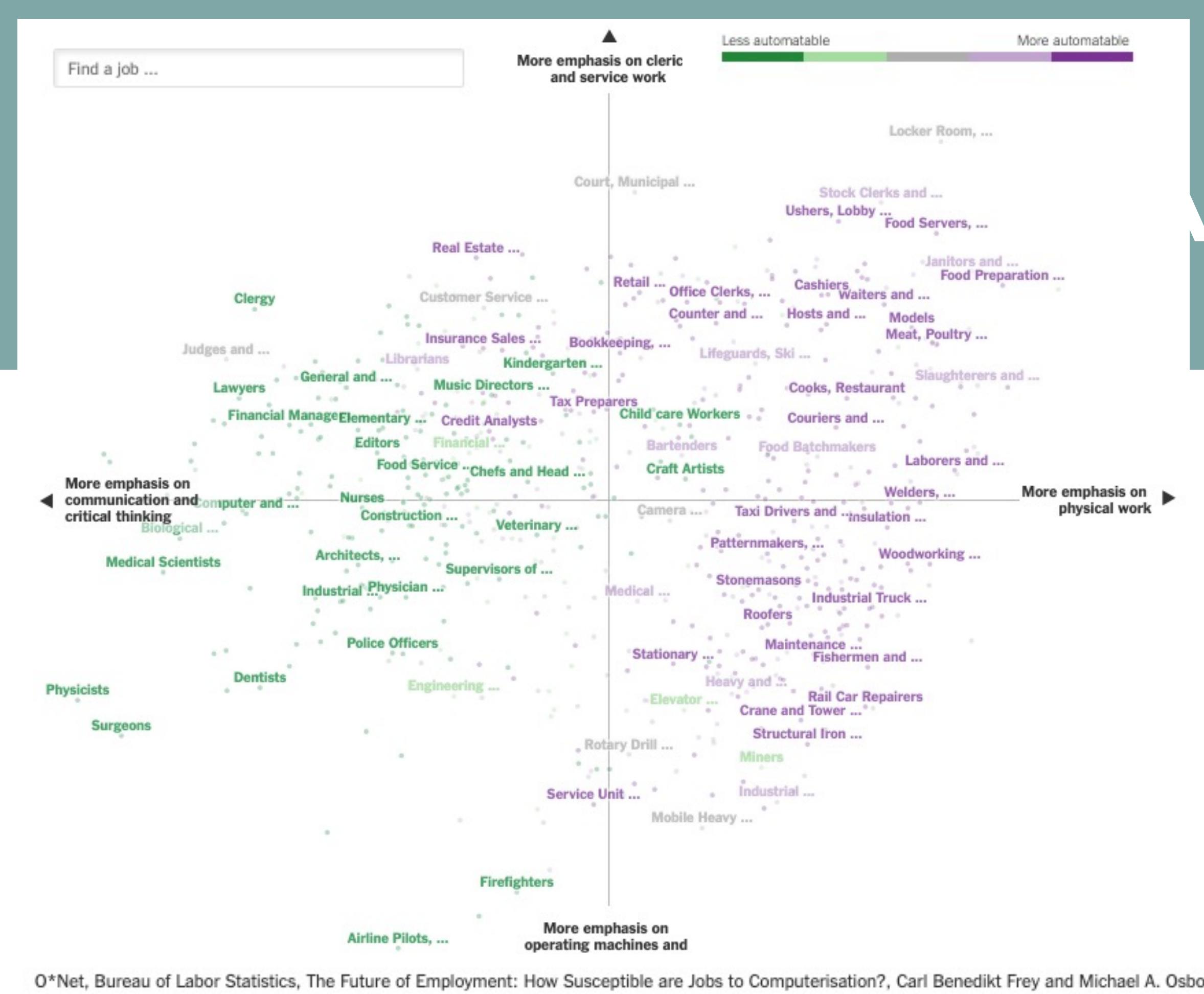
High-level → Consume → Present

Medium-level → Search → Explore (Lookup)

Low-level → Query → Summarize (Compare)



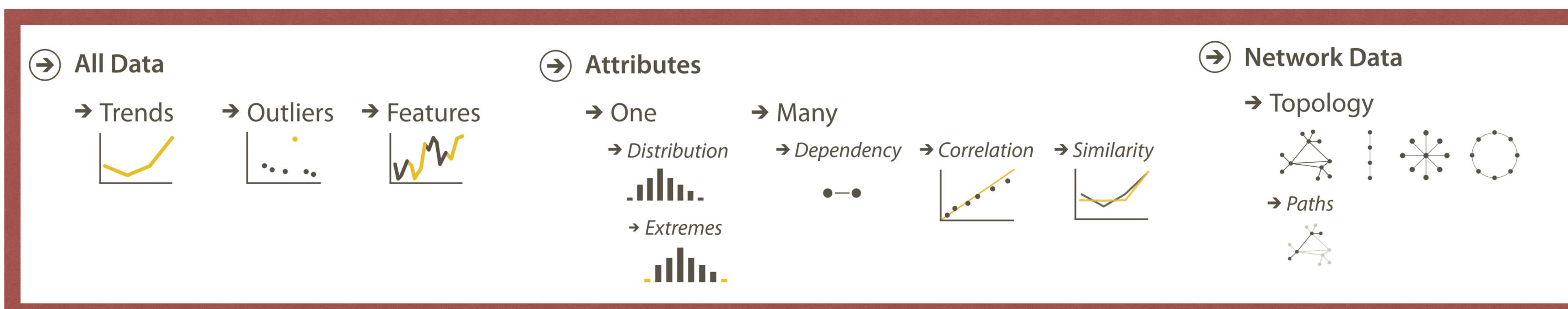
Abstraction



High-level → Consume → Present

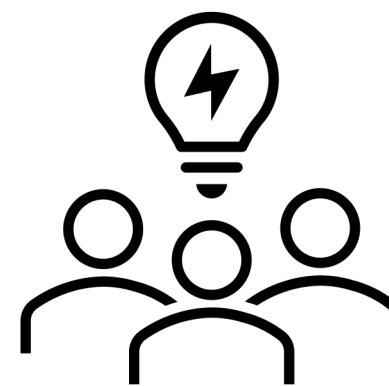
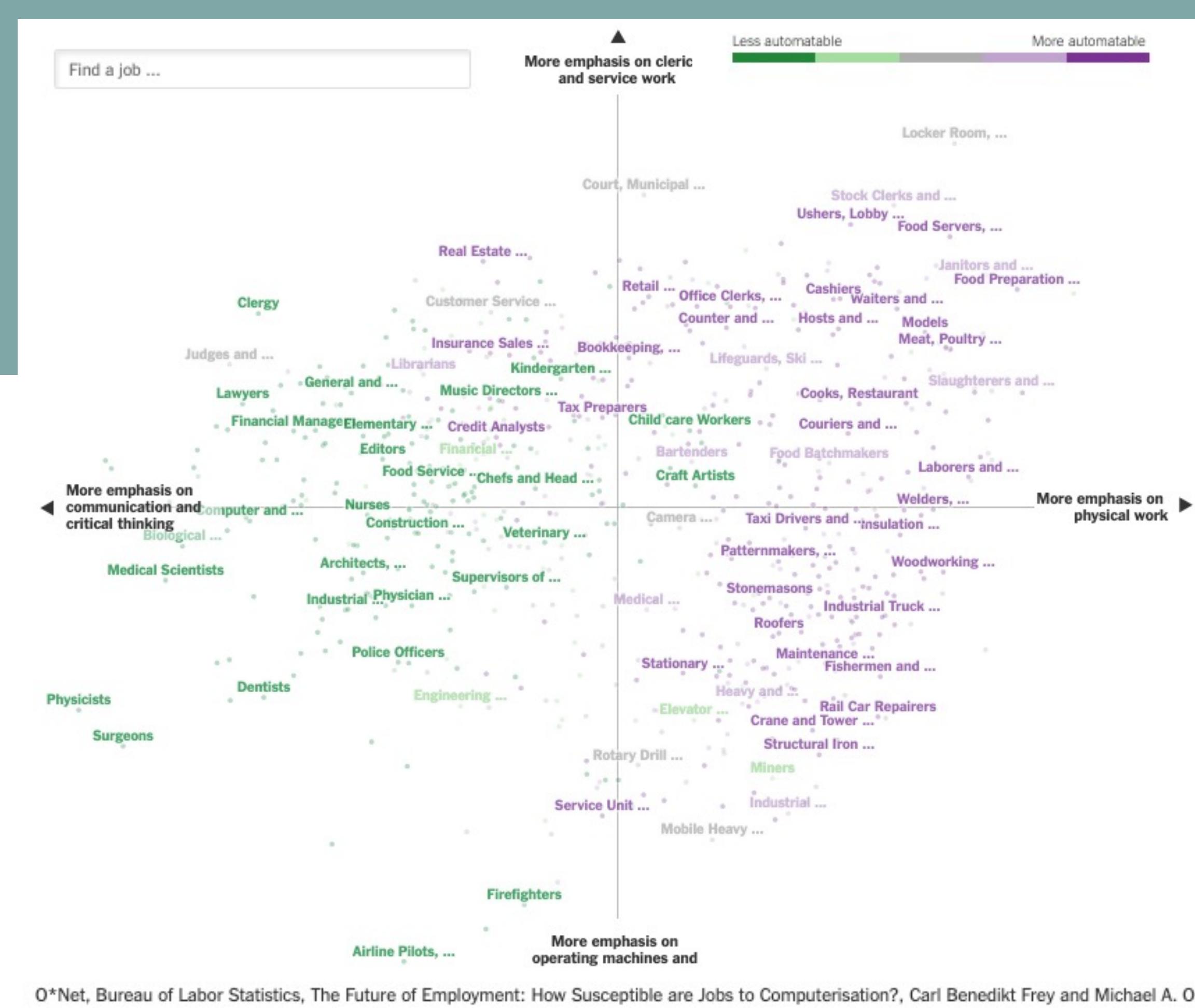
Medium-level → Search → Explore (Lookup)

Low-level → Query → Summarize (Compare)



Target(s)

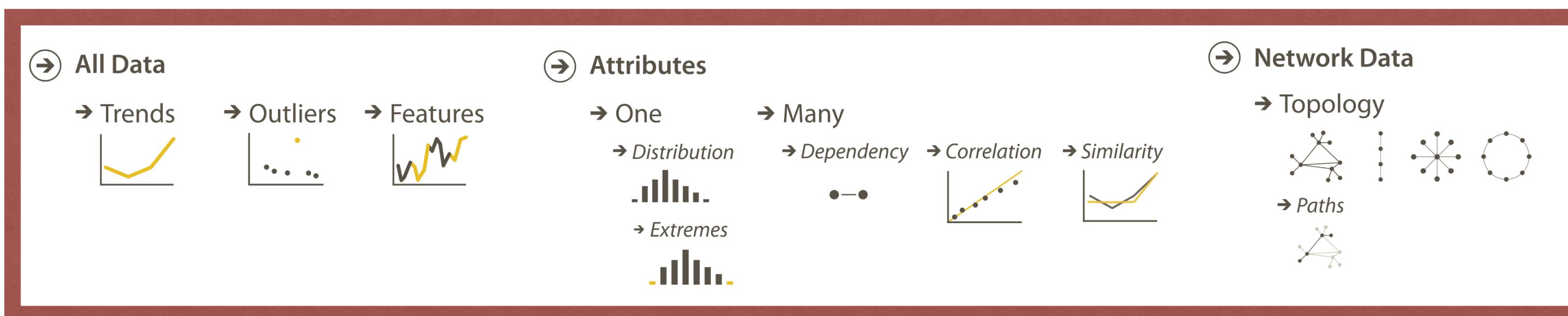
Abstraction



High-level →
Consume →
Present

Medium-level →
Search →
Explore (Lookup)

Low-level →
Query →
Summarize
(Compare)

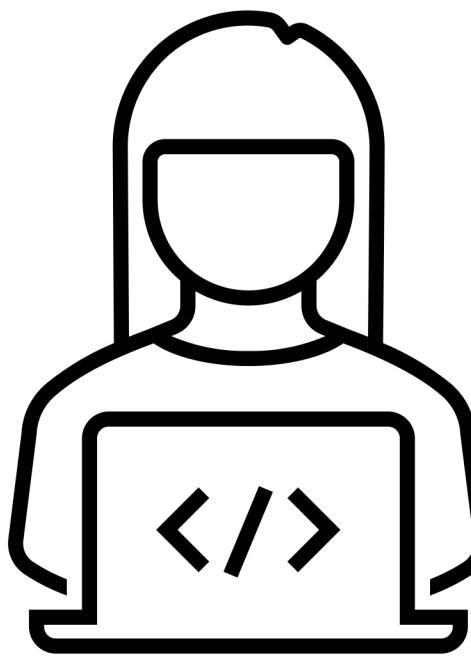


Target(s) → All Data → Trends (Attribute → Distribution)

Let's take a break! Stretch, go
for a walk, be social ☺
Be back here in 10 mins.

INTERVIEWING DOMAIN EXPERTS

Human-Centered Design



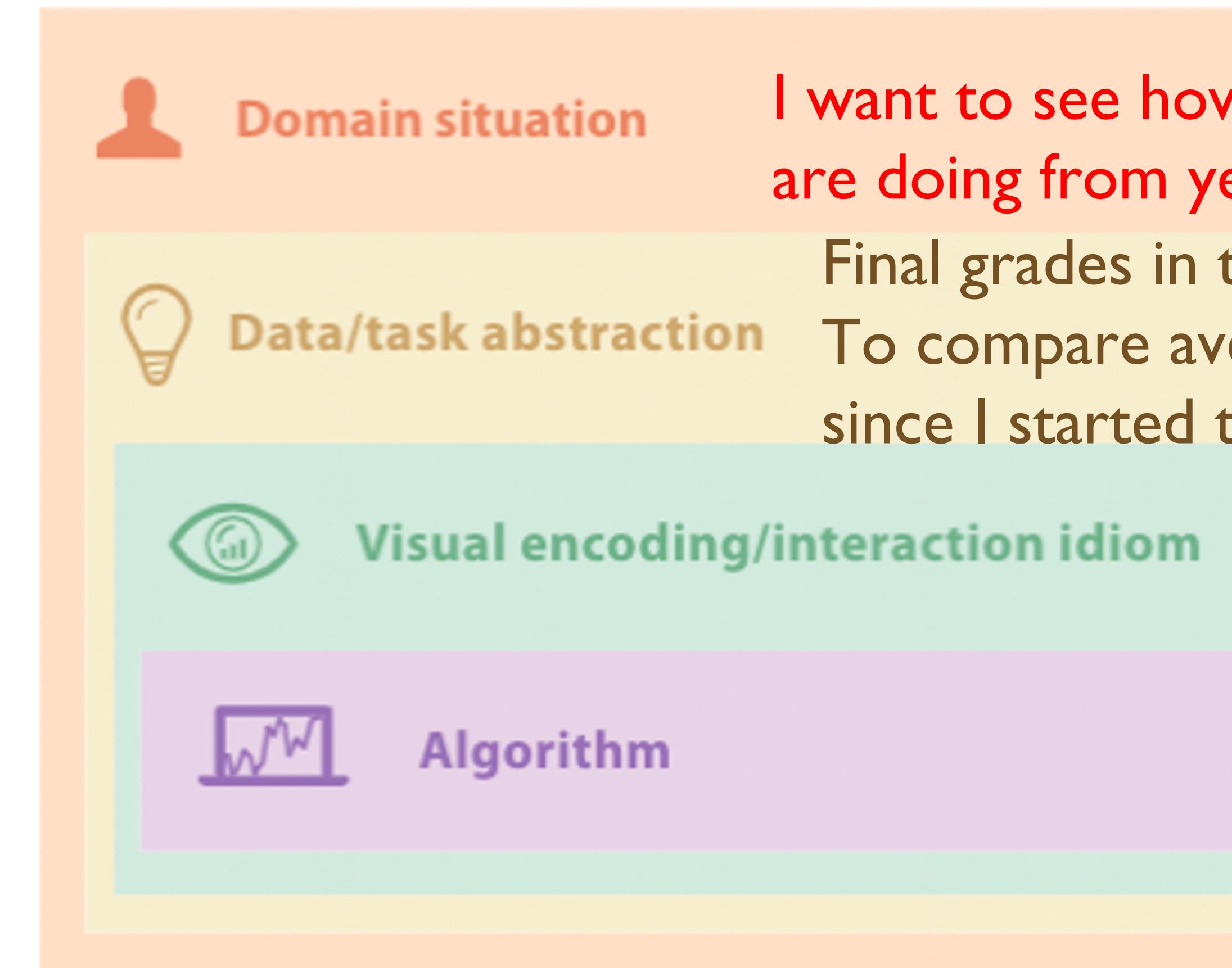
What do you need
a visualization for?

How do you
measure student
success?

What kinds of comparisons
or statistics do you need?

I'll start with a line chart
showing average grade per
year.

I'll code it in Tableau.



Interview Tips

- **Your interviewee is the expert in what they do. Let them explain it to you**
 - What does your day to day consist of?
 - How would you explain this task to a new employee?
 - What tools do you need to be successful?

Interview Tips

- **Ask specific questions**

→ Can you explain your workflow to me in five steps?

vs.

Tell me about your workflow.

→ Please describe three pain points in your current process.

vs.

Are there any pain points in your current process?

→ Please describe three positive aspects of your current process.

vs.

What parts of your process do you like or want unchanged?

Interview Tips

- **Be as interactive as possible. Ask follow-ups.**
 - Can you explain more about <X> part of your current workflow?
 - When you explained what works well for you now, you mentioned <Y>. Could you tell me about <Y> in more detail?
 - You mentioned <Z> as a current pain point. In an ideal world, how would <Z> be alleviated?

IN-CLASS EXERCISE: MOCK INTERVIEW (IC-03)

Summary

Today we:

- Reviewed Task Abstraction
- Practiced interviewing an end user to understand tasks

hw-02 is DUE before next class.

ic-03 is DUE today.