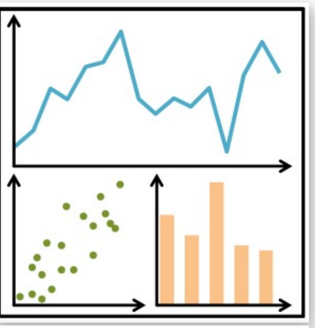
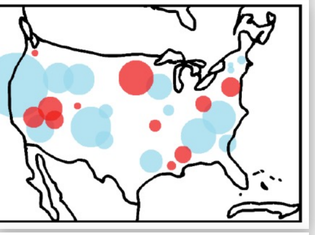
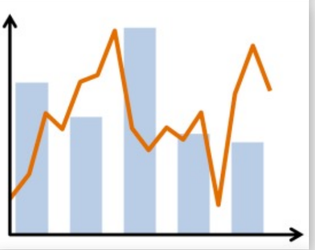
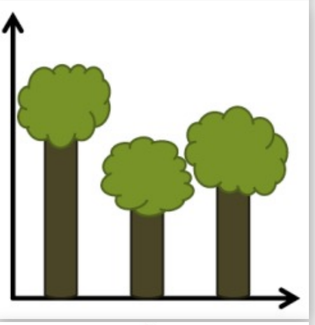
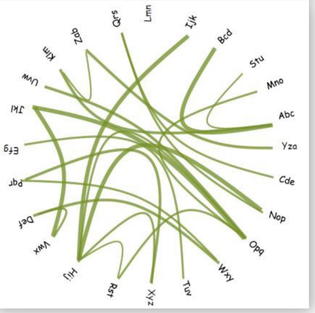






Lecture 4: Task Abstraction

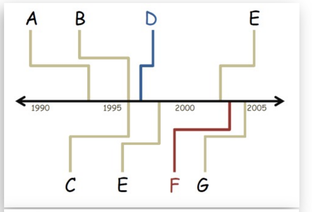
DS 4200

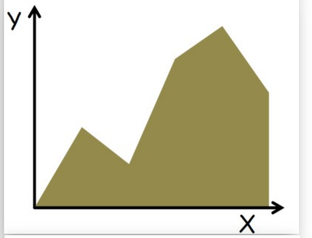
FALL 2022

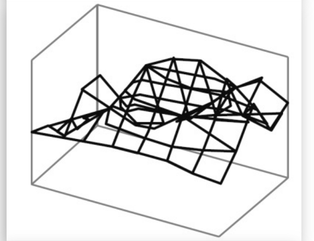
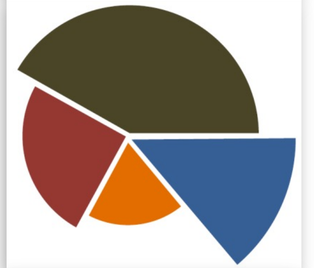
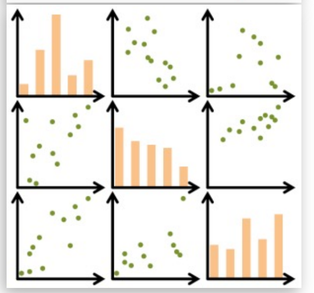
Prof. Ab Mosca (they/them)

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

What?

Why?

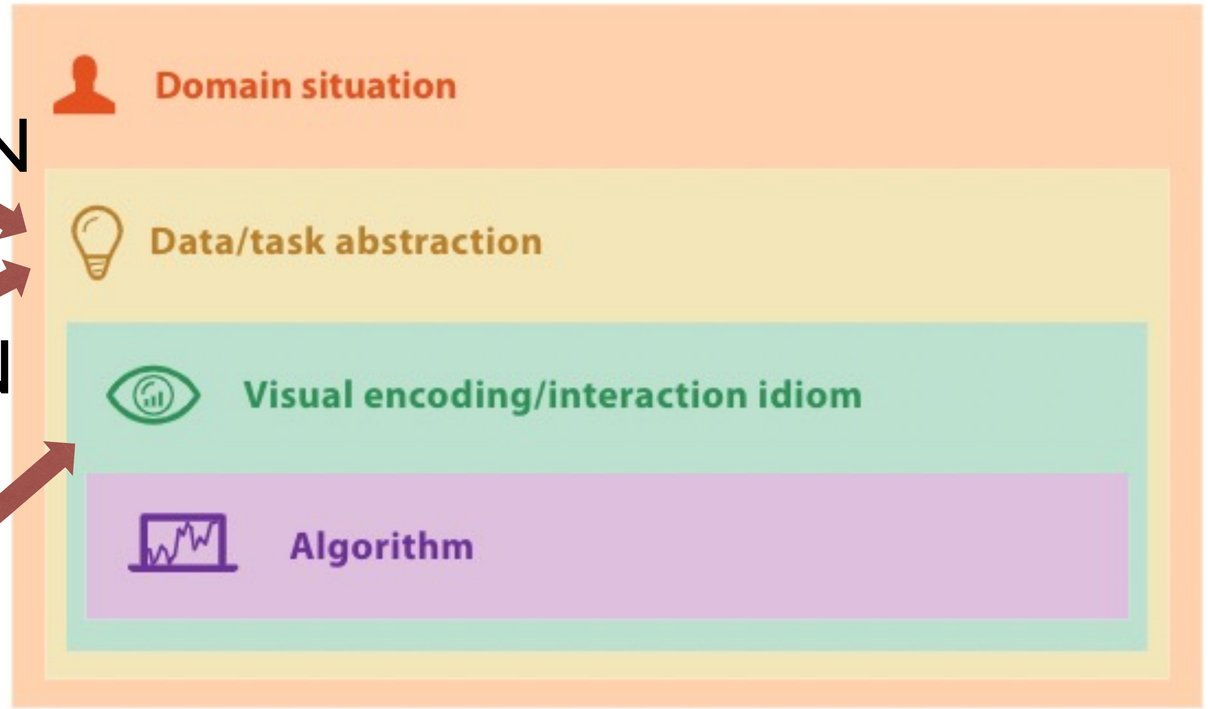
How?

DATA ABSTRACTION

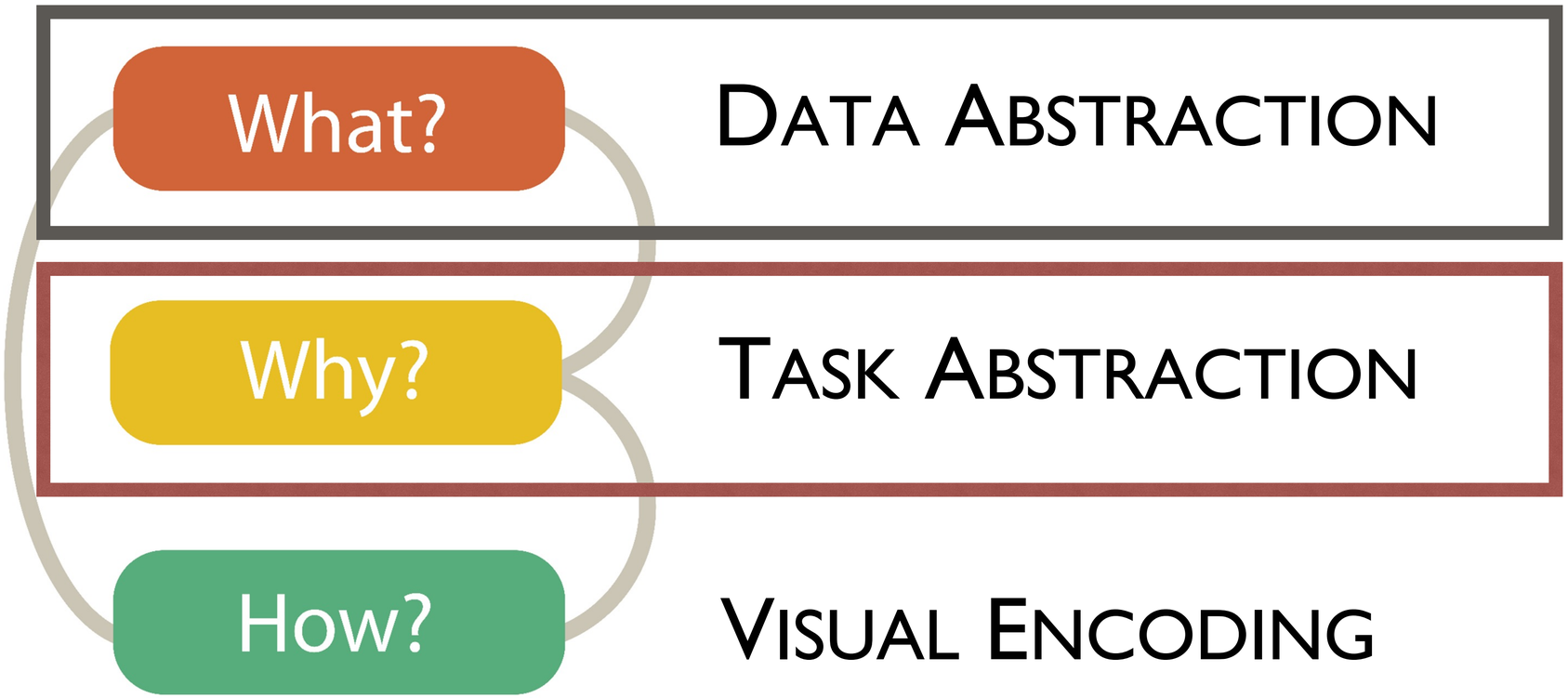
TASK ABSTRACTION

VISUAL ENCODING

Munzner's Nested Model



Visualization Building Blocks



Task Abstraction

From Munzner's book

Task Abstraction

What?

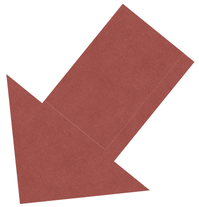
The process of taking specific **domain tasks** and thinking about them as **abstract** (modular!) pieces

Task Abstraction

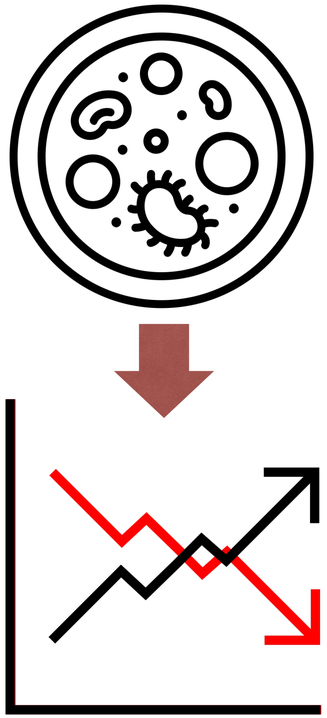
What?

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?

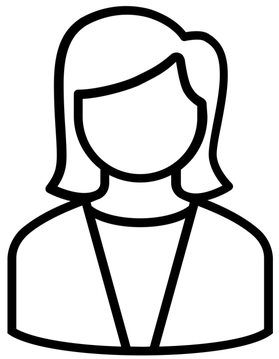
To translate domain specific terms into well known and transferable visualization tasks.

Task Abstraction

Why?

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

Ex.



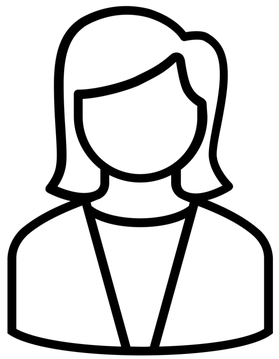
I need to show my boss which product we sold the most of last year.

Task Abstraction

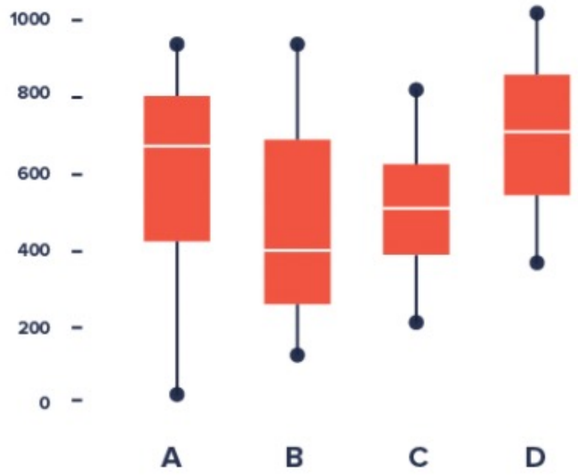
Why?

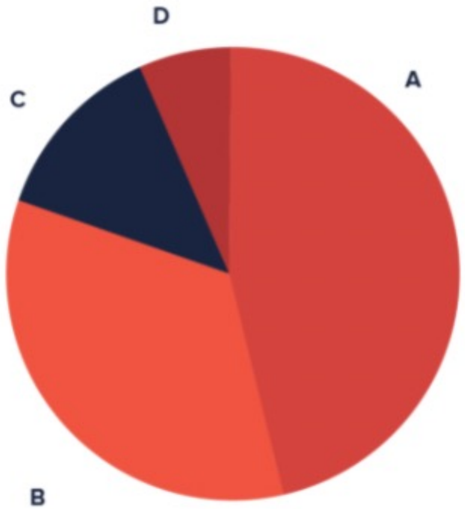
To translate domain specific terms into well known and transferable visualization tasks.

Ex.



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





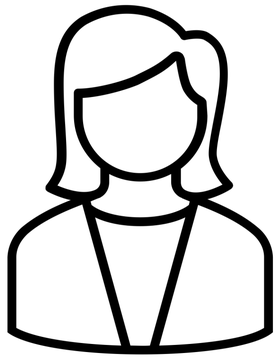
https://datavizproject.com/

Task Abstraction

Why?

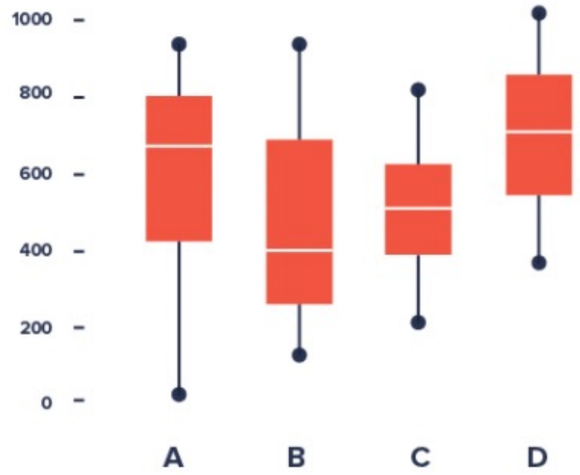
To translate domain specific terms into well known and transferable visualization tasks.

Ex.

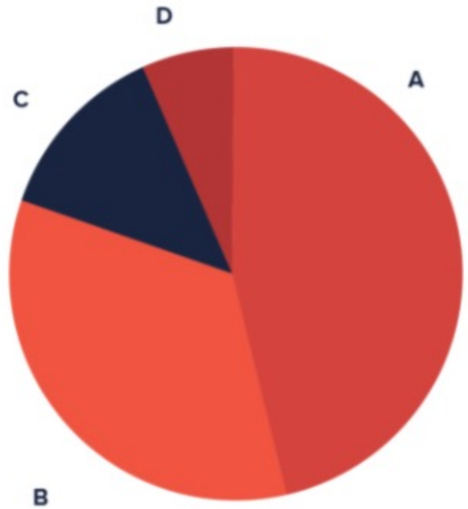


I need to show my boss which product we sold the most of last year.

Compare distributions



Compare part-to-whole



https://datavizproject.com/

Task Abstraction

Why?

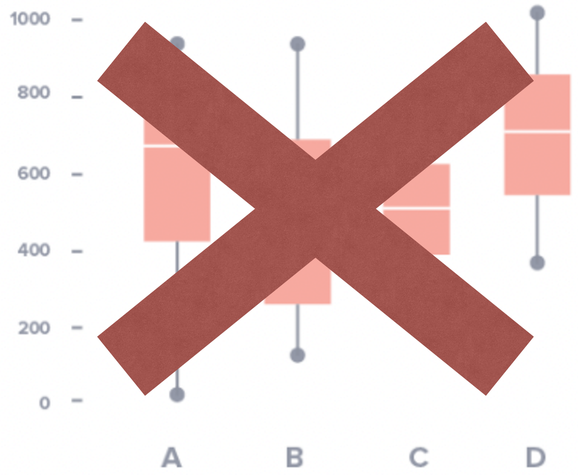
To translate domain specific terms into well known and transferable visualization tasks.

Ex.

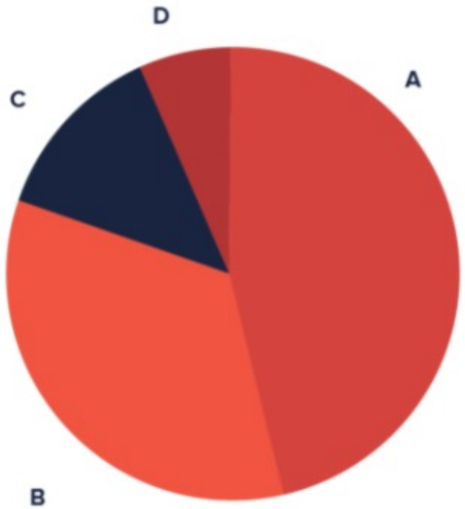
I need to show my boss which product we sold the most of last year.

Need to compare part-to-whole.

~~Compare distributions~~



Compare part-to-whole



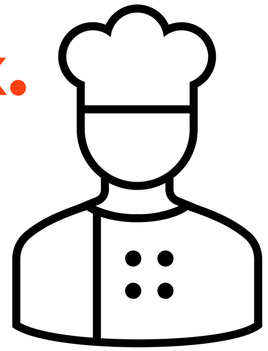
https://datavizproject.com/

Task Abstraction

Why?

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

Ex.



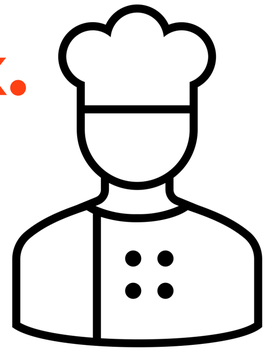
I want to see which months are best for outdoor dining.

Task Abstraction

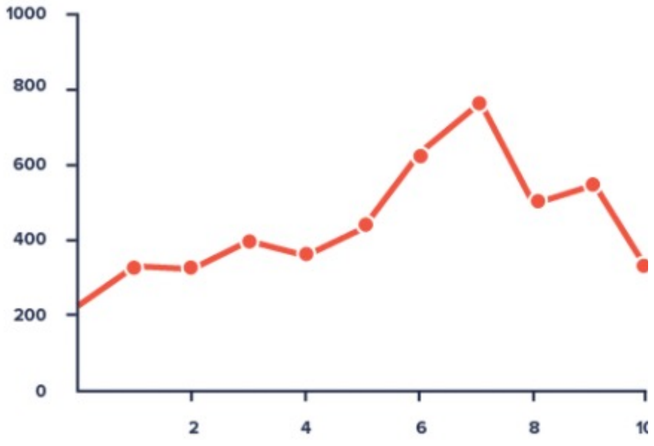
Why?

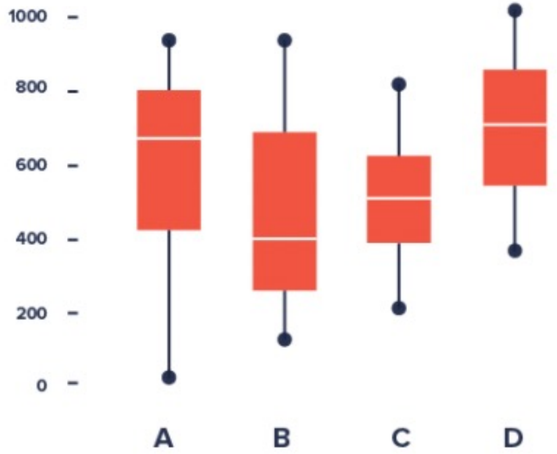
To translate domain specific terms into well known and transferable visualization tasks.

Ex.



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





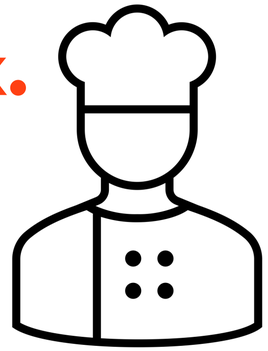
https://datavizproject.com/

Task Abstraction

Why?

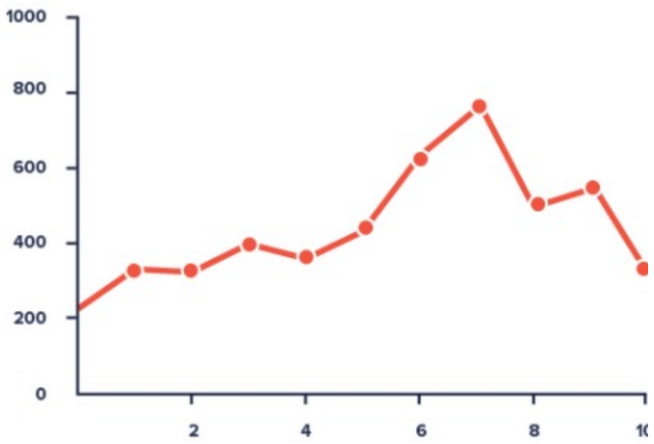
To translate domain specific terms into well known and transferable visualization tasks.

Ex.

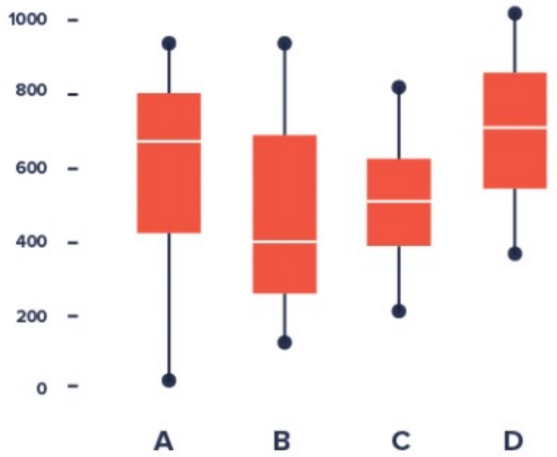


I want to see which months are best for outdoor dining.

Average trend



Compare distributions



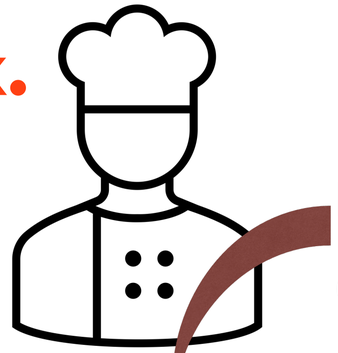
https://datavizproject.com/

Task Abstraction

Why?

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

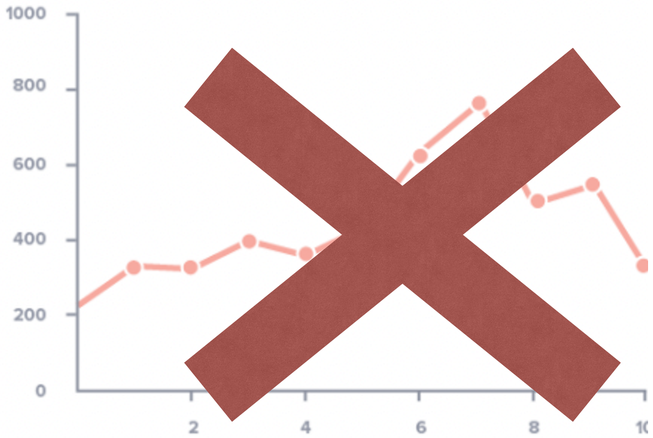
Ex.



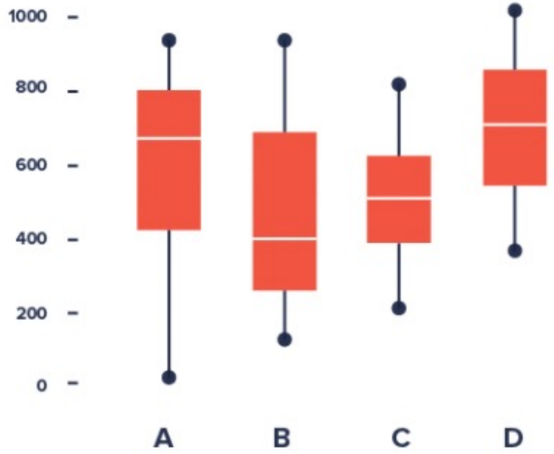
I want to see which months are best for outdoor dining.

Need to compare distribution over time.

~~Average trend~~



Compare distributions



https://datavizproject.com/

Task Abstraction

Why?

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

Task Abstraction

Why?

Actions

- **Analyze**
  - Consume
    - *Discover*
    - *Present*
    - *Enjoy*
  - Produce
    - *Annotate*
    - *Record*
    - *Derive*
- **Search**

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

- **Query**
  - Identify
  - Compare
  - Summarize

Targets

- **All Data**
  - Trends
  - Outliers
  - Features
- **Attributes**
  - One
    - *Distribution*
    - *Extremes*
  - Many
    - *Dependency*
    - *Correlation*
    - *Similarity*
- **Network Data**
  - Topology
  - *Paths*
- **Spatial Data**
  - Shape

What?
Why?
How?

Task Abstraction

Actions

3 Levels:

- High
- Medium
- Low

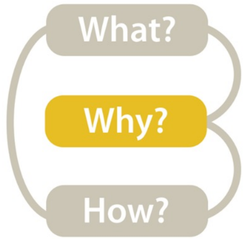
Task Abstraction

Actions

3 Levels:

- High
- Medium
- Low

Task Abstraction



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

→ Analyze

→ Consume

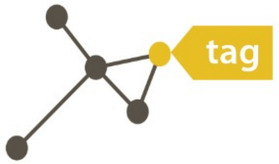
→ *Discover*

→ *Present*

→ *Enjoy*

→ Produce

→ *Annotate*



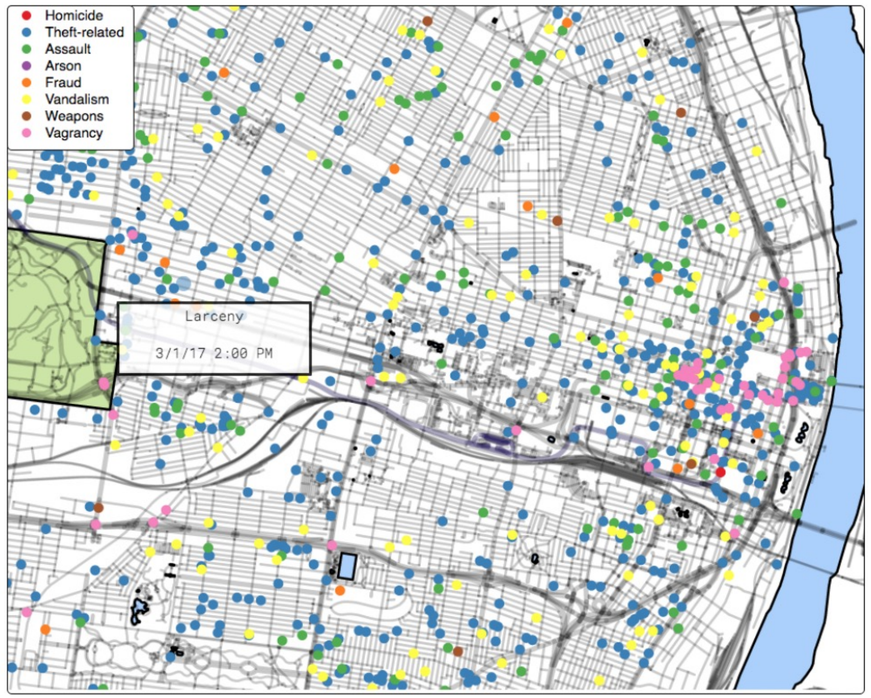
→ *Record*

→ *Derive*

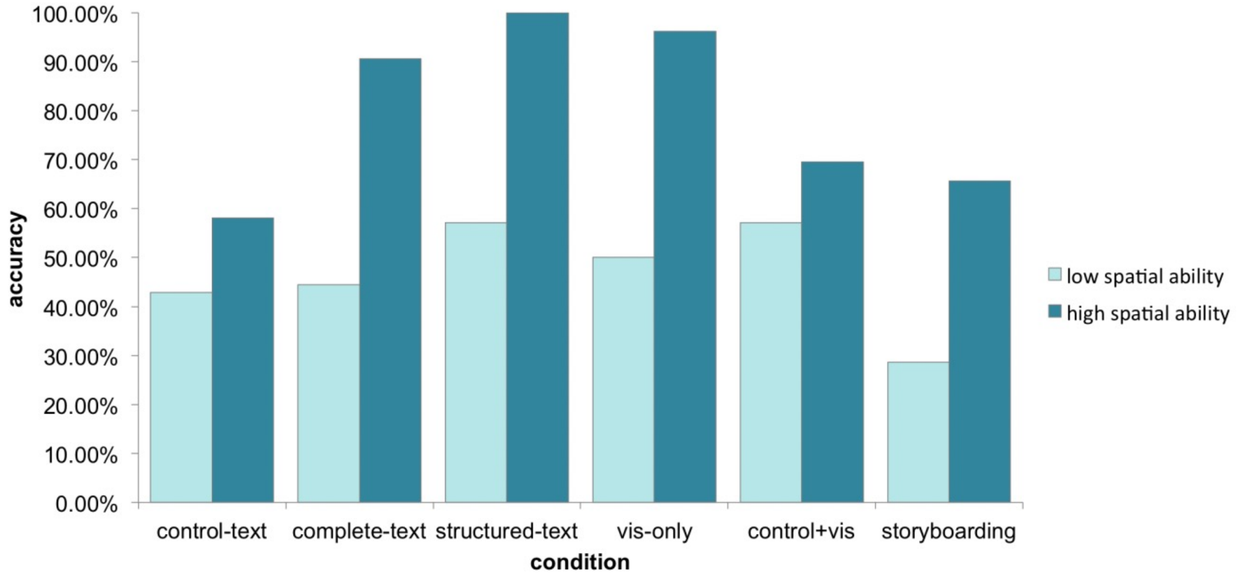
Task Abstraction

What?
Why?
How?

High-level → Consume → Discover

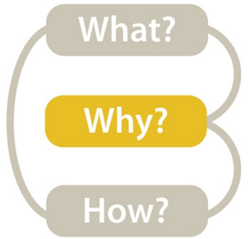


Follow The Clicks: Learning and Anticipating Mouse Interactions During Exploratory Data Analysis. Ottley et al. 2019

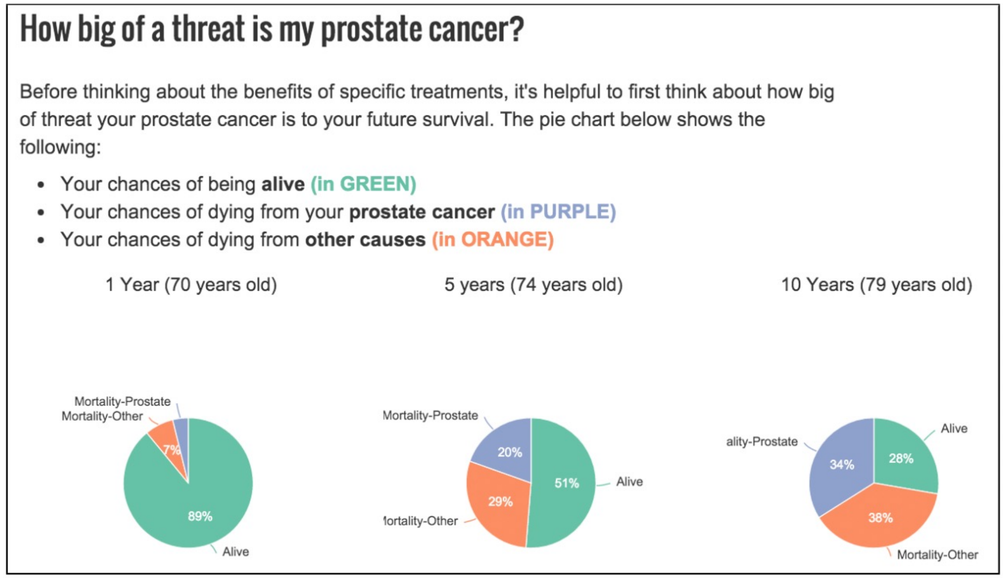


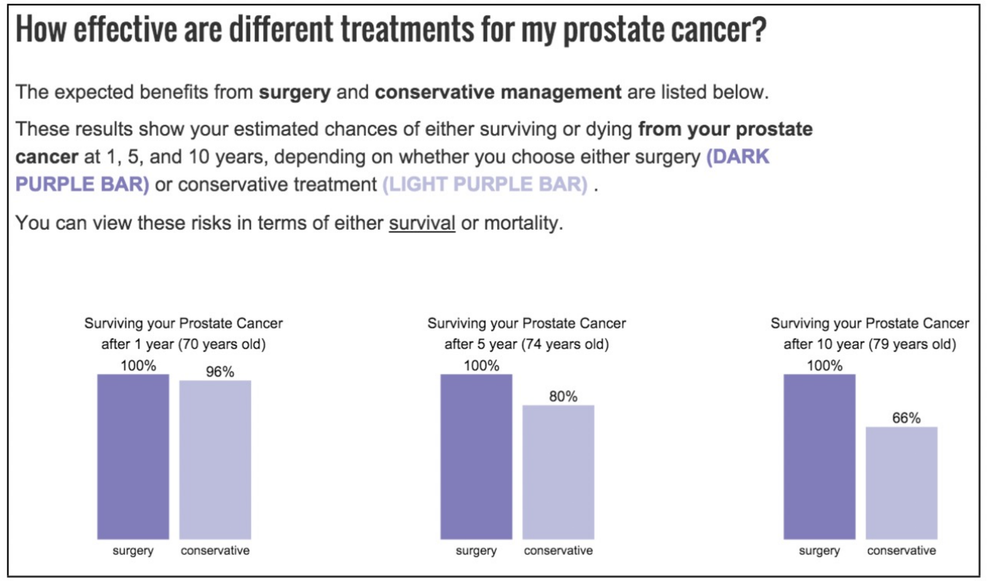
Improving Bayesian Reasoning: The Effects of Phrasing, Visualization, and Spatial Ability. Ottley et al. 2016

Task Abstraction



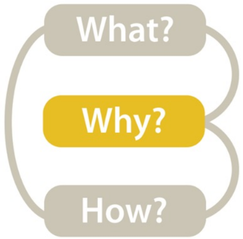
High-level → Consume → Present



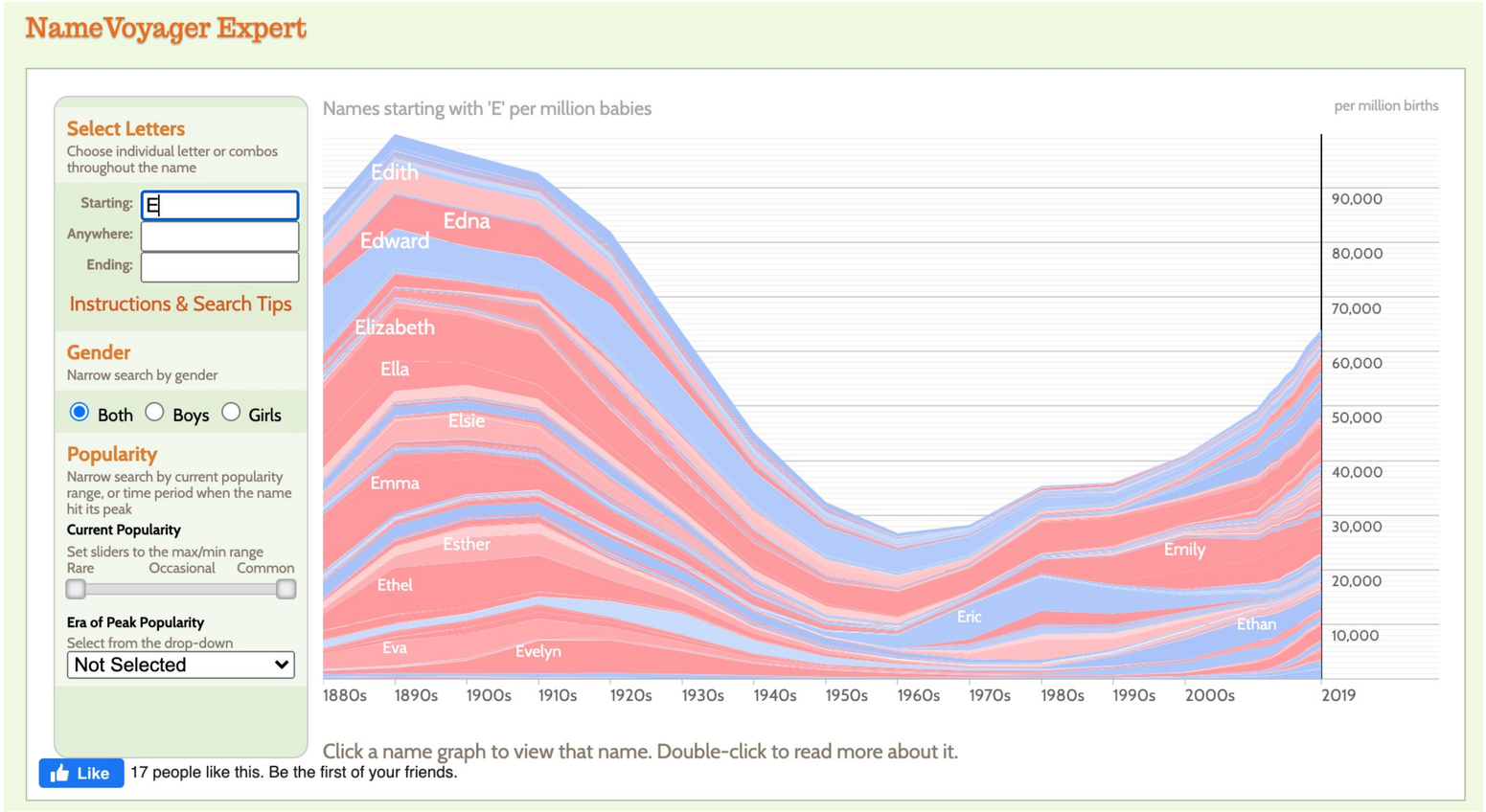


PROACT: Iterative Design of a Patient-Centered Visualization for Effective Prostate Cancer Health Risk Communication. Hakone et al. 2016.

Task Abstraction

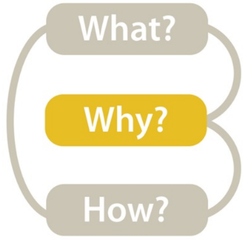


High-level → Consume → Enjoy

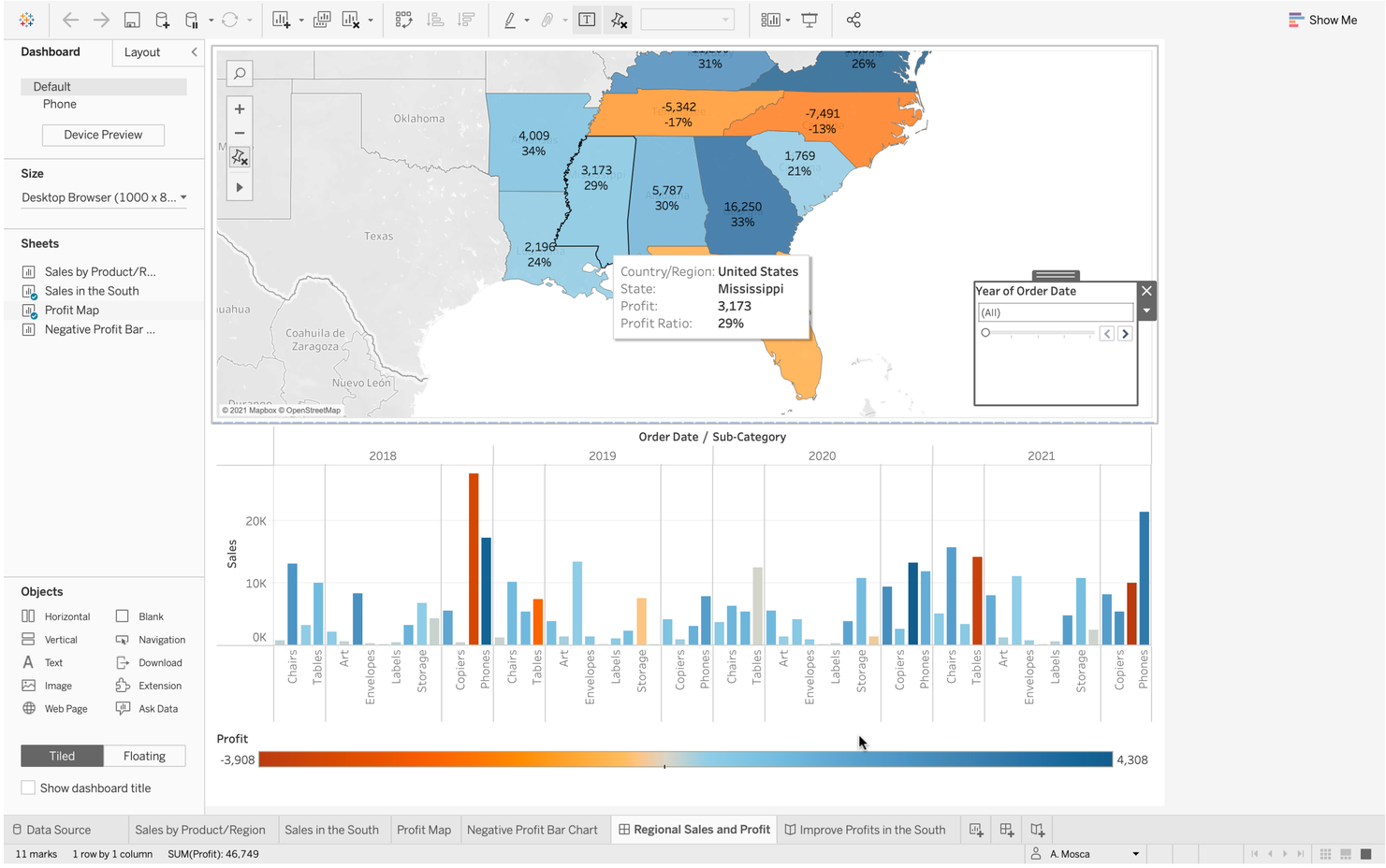


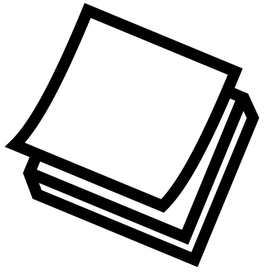
https://www.babynamewizard.com/namevoyager-expert#prefix=e&sw=both&exact=false

Task Abstraction

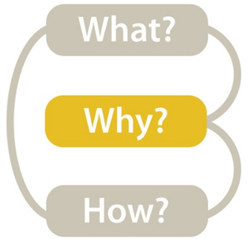


High-level → Produce → Annotate





Task Abstraction

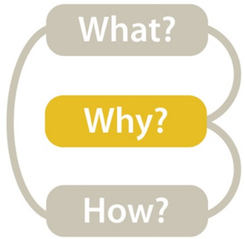


High-level → Produce → Record

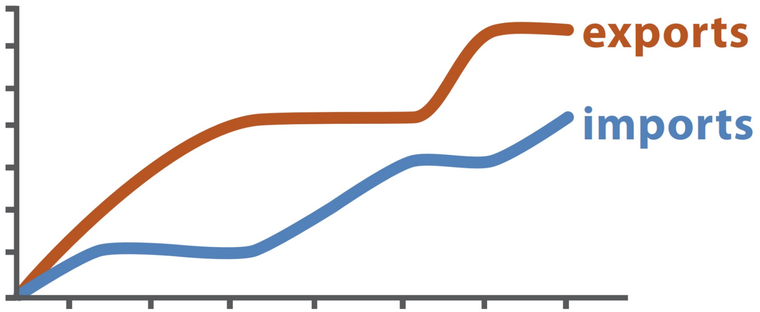
Fig. 1. A Graphical History Interface. Thumbnails show previous visualization states and labels describe the actions performed.

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

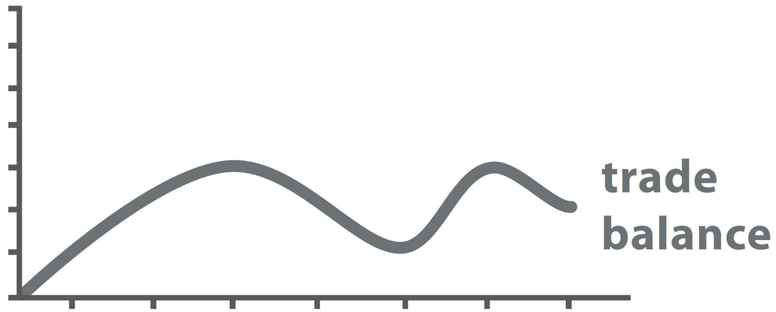
Task Abstraction



High-level → ***Produce*** → ***Derive***



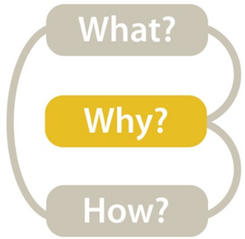
Original Data



trade balance = *exports* – *imports*

Derived Data

Task Abstraction



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

→ Analyze

→ Consume

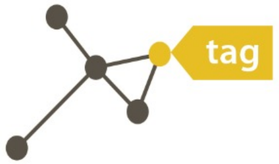
→ *Discover*

→ *Present*

→ *Enjoy*

→ Produce

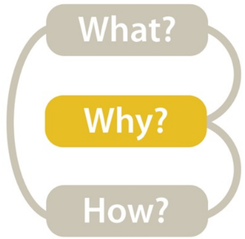
→ *Annotate*



→ *Record*

→ *Derive*

Task Abstraction



Mid-level → Search

- Analyze
  - Consume
    - Discover
    - Present
    - Enjoy

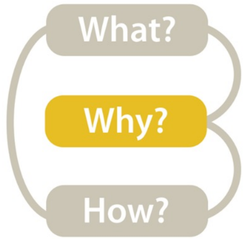


  - Produce
    - Annotate
    - Record
    - Derive

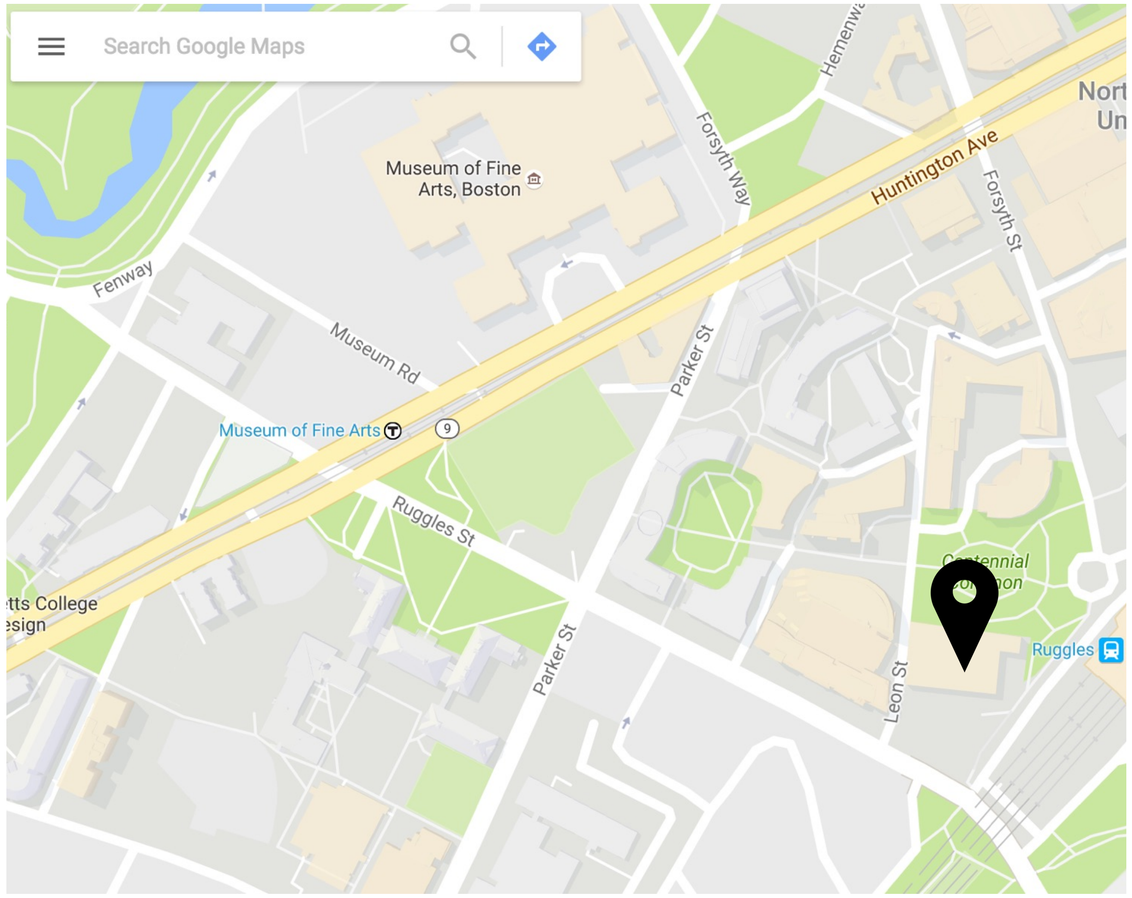

- Search

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

Task Abstraction



Mid-level → Search → Lookup

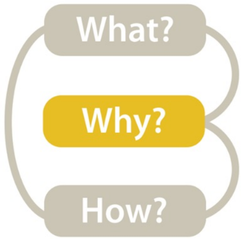


Search

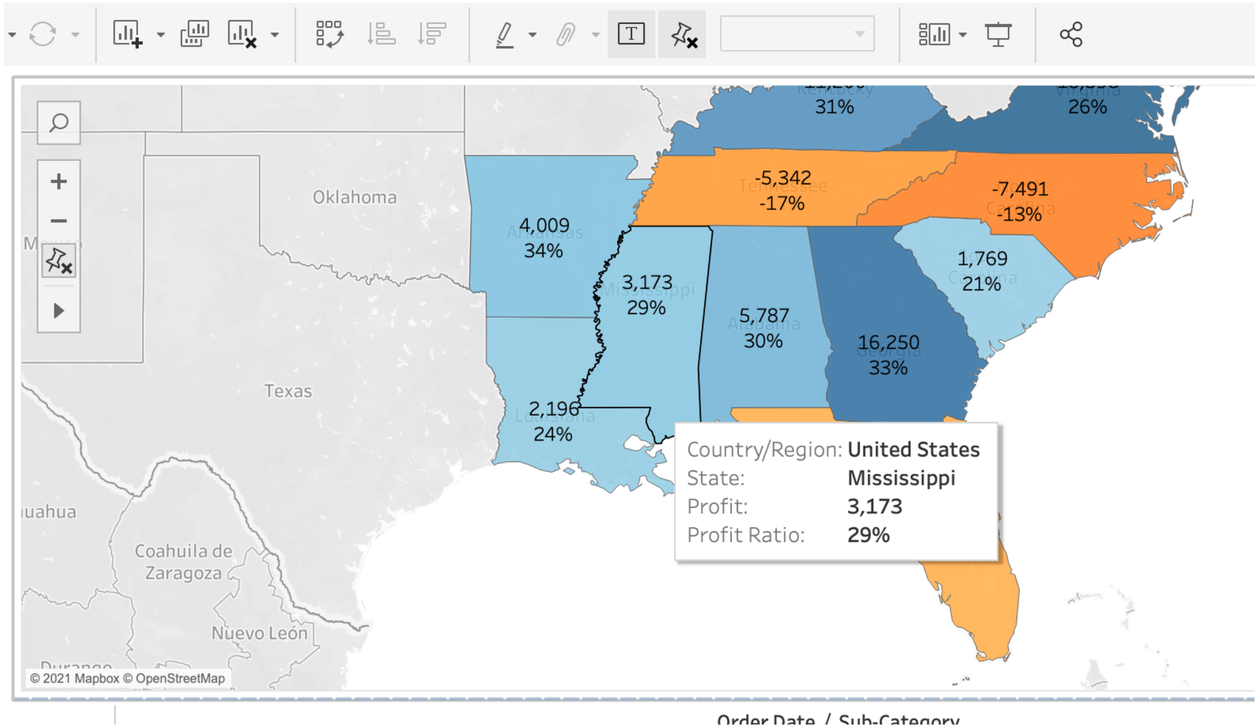
	Target known	Target unknown
Location known	*Lookup*	*Browse*
Location unknown	*Locate*	*Explore*

What is the address of Ryder Hall?

Task Abstraction



Mid-level → Search → Lookup

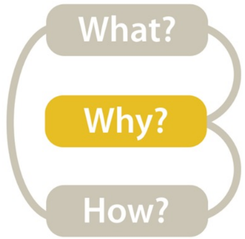


Search

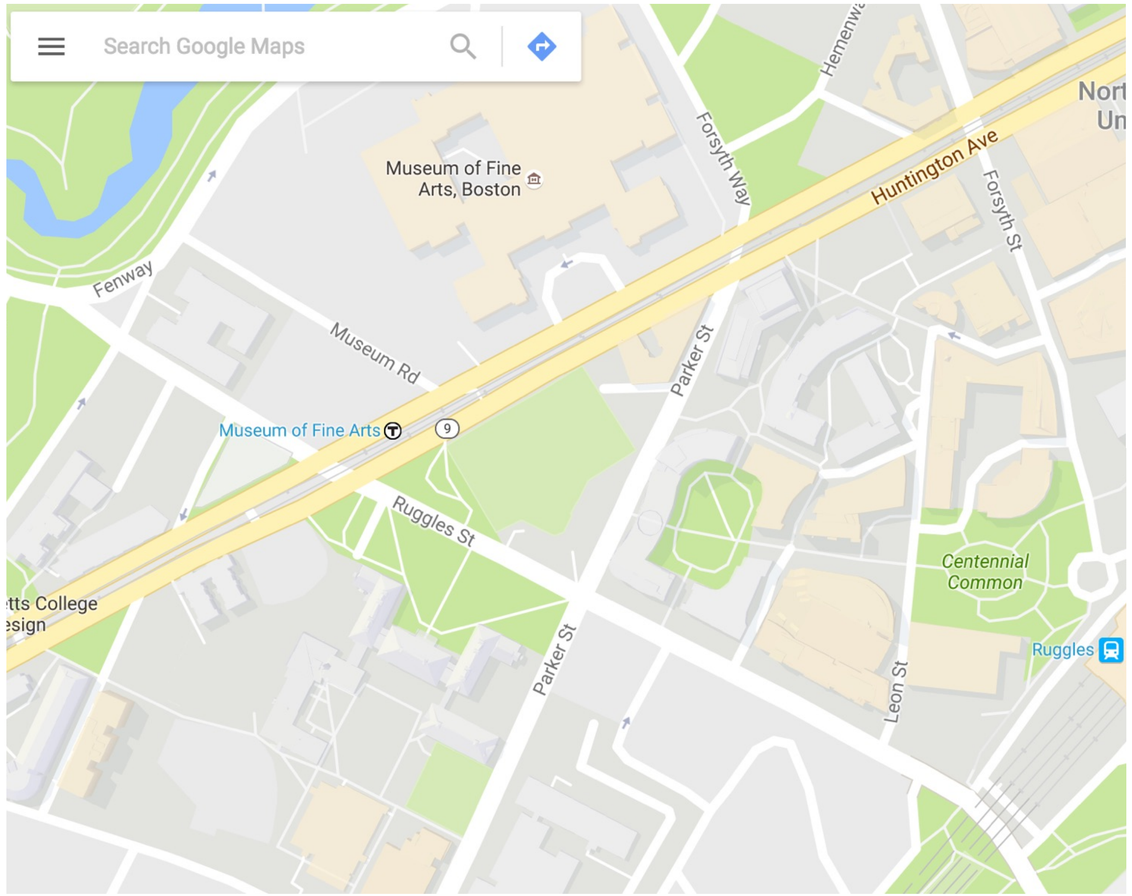
	Target known	Target unknown
Location known	*Lookup*	*Browse*
Location unknown	*Locate*	*Explore*

How much profit was there for Mississippi?

Task Abstraction



Mid-level → Search → Locate

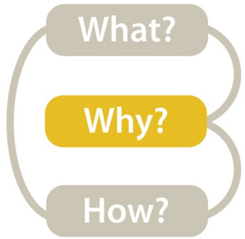


Search

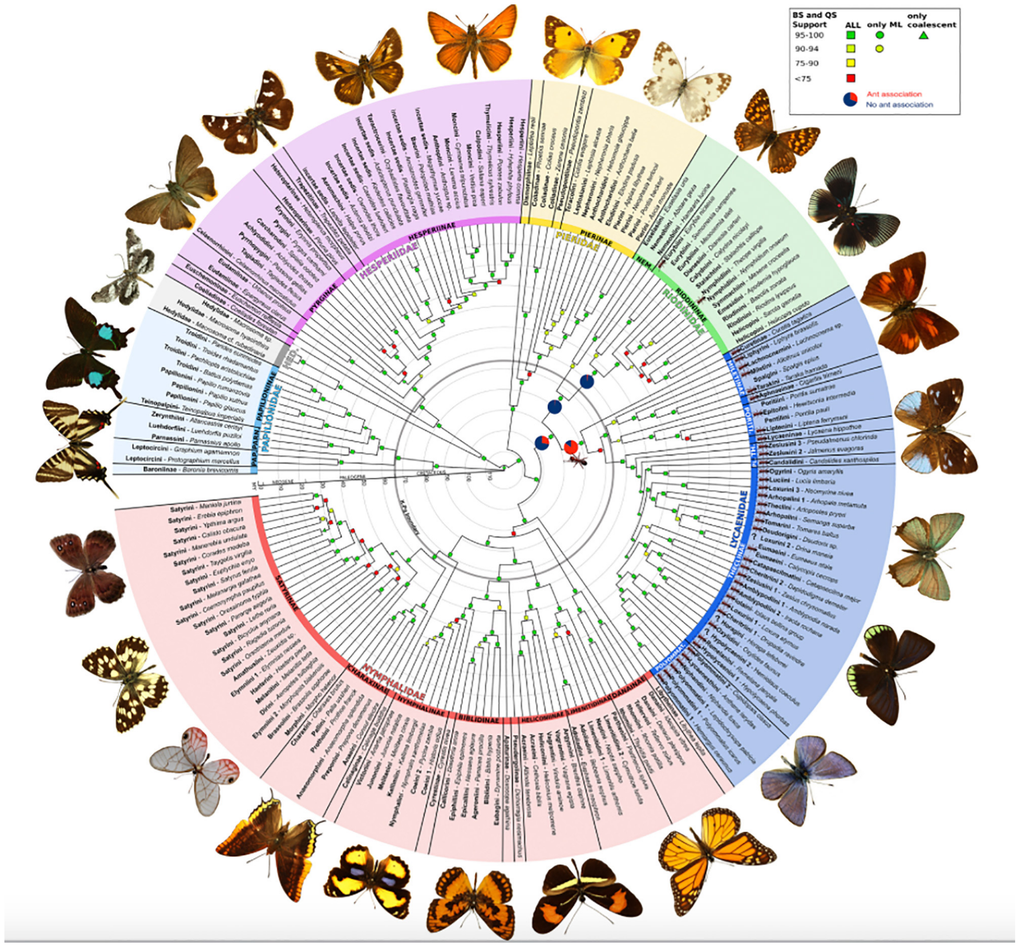
	Target known	Target unknown
Location known	*Lookup*	*Browse*
Location unknown	*Locate*	*Explore*

Where is Ryder Hall?

Task Abstraction



Mid-level → Search → Locate



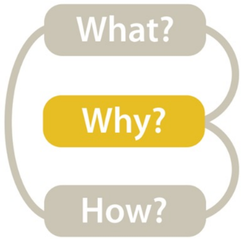
→ Search

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

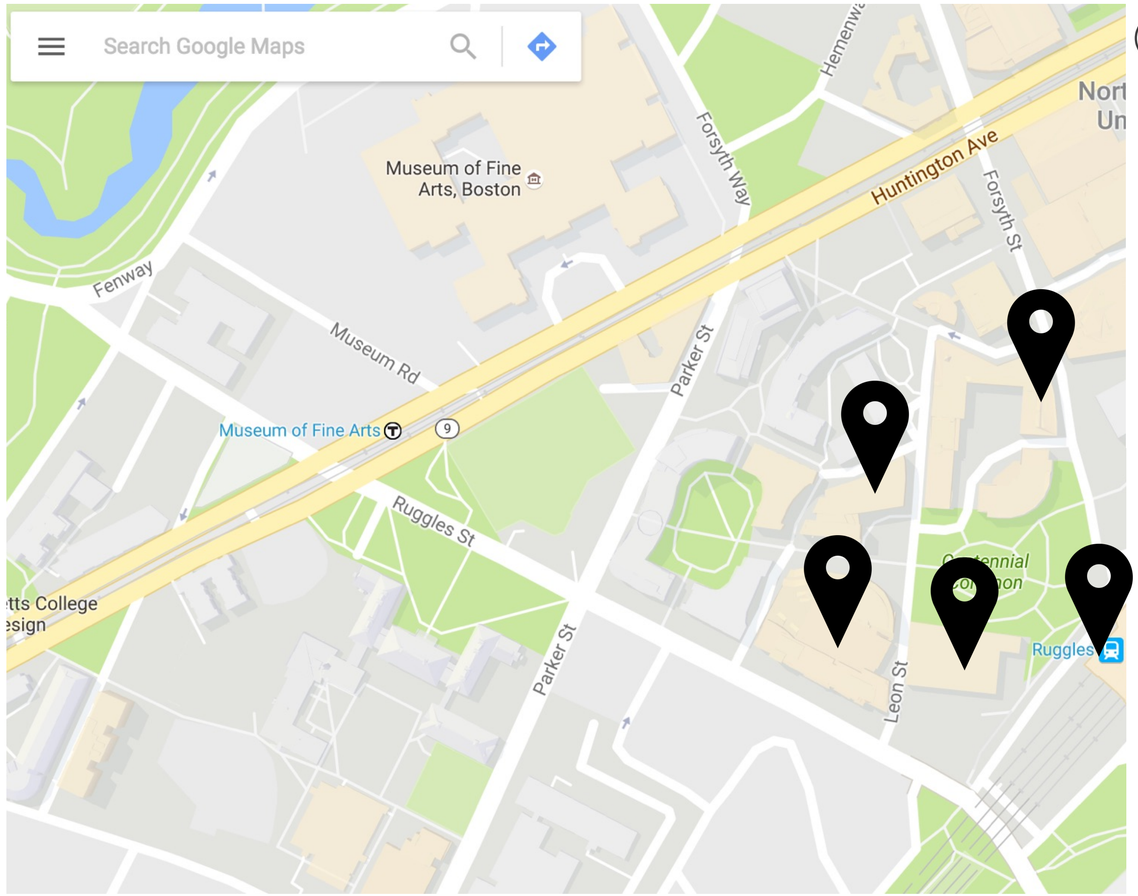
Where is the Monarch butterfly?

https://www.floridamuseum.ufl.edu/science/at-last-butterflies-get-a-bigger-better-evolutionary-tree/

Task Abstraction



Mid-level → Search → Browse

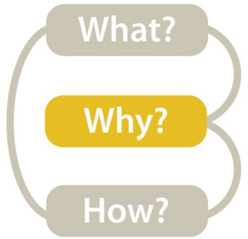


Search

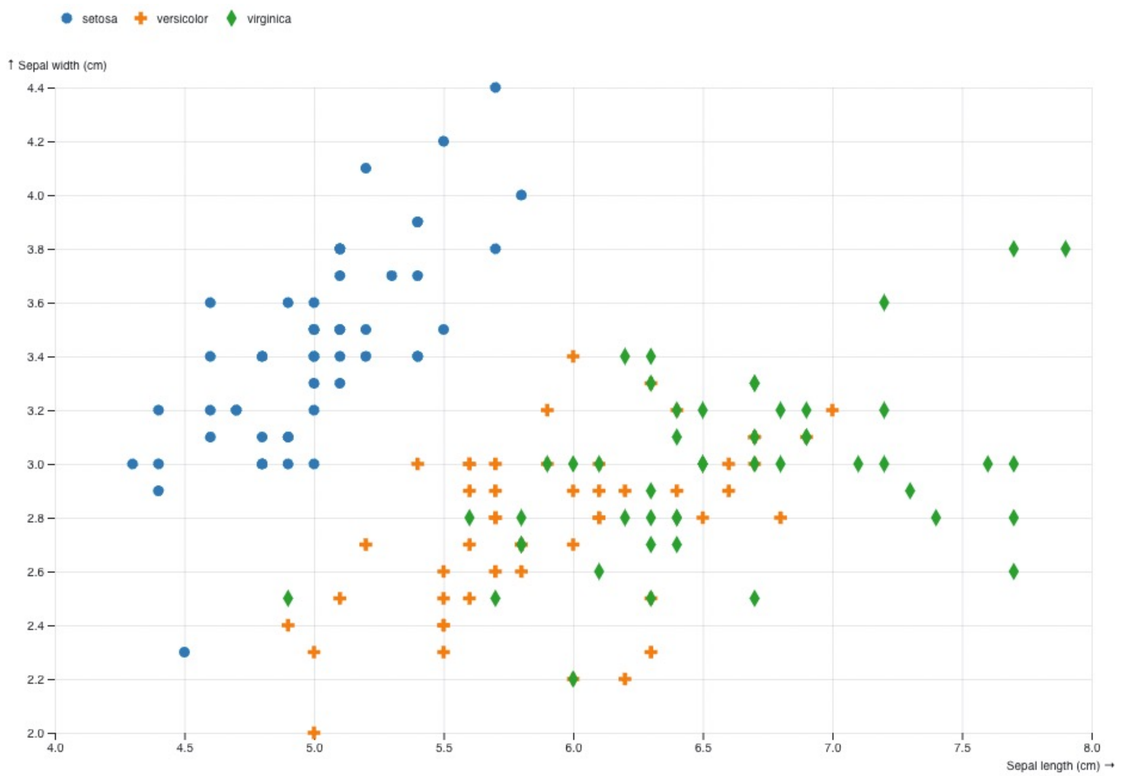
	Target known	Target unknown
Location known	*Lookup*	*Browse*
Location unknown	*Locate*	*Explore*

What buildings are near Ryder Hall?

Task Abstraction



Mid-level → Search → Browse



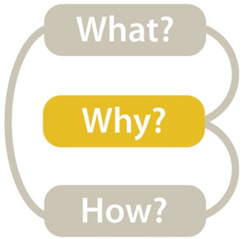
Search

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

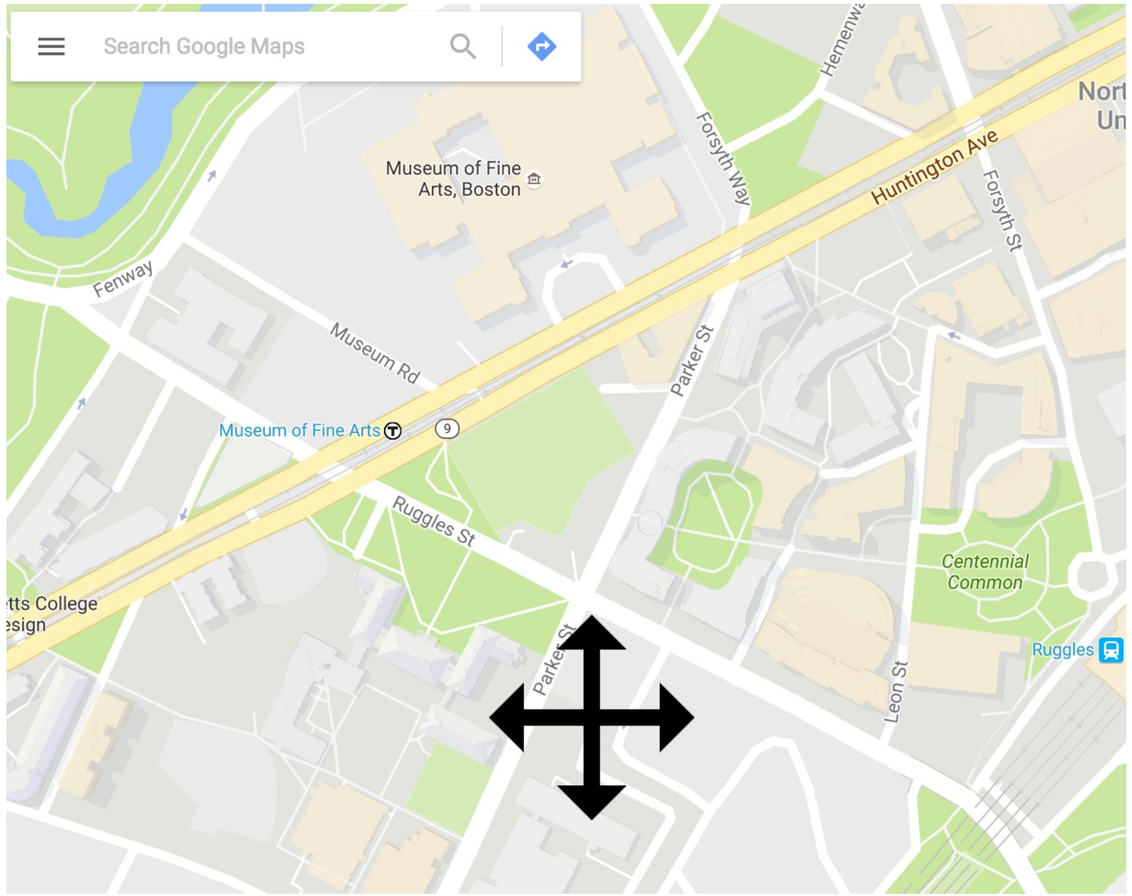
Which Irises have sepal length greater than 7.5?

https://observablehq.com/@d3/scatterplot-with-shapes

Task Abstraction



Mid-level → Search → Explore

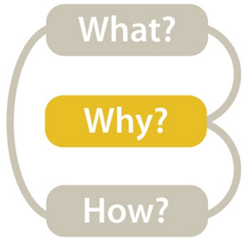


Search

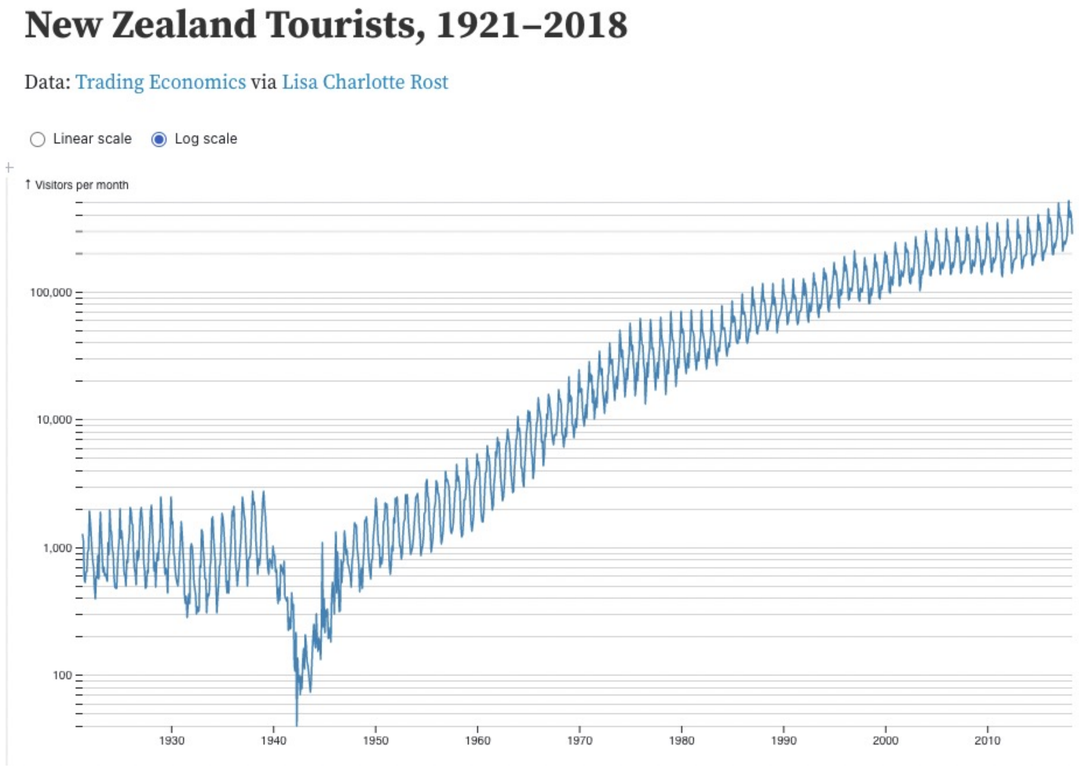
	Target known	Target unknown
Location known	*Lookup*	*Browse*
Location unknown	*Locate*	*Explore*

Where can I study?

Task Abstraction



Mid-level → Search → Explore



Search

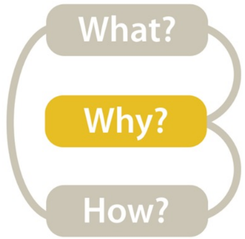
	Target known	Target unknown
Location known	*Lookup*	*Browse*
Location unknown	*Locate*	*Explore*

Is there a seasonal trend is tourism?

Are there any unusual years?

https://observablehq.com/@mbostock/new-zealand-tourists-1921-2018

Task Abstraction

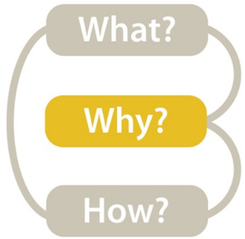


Mid-level → Search

Search

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

Task Abstraction



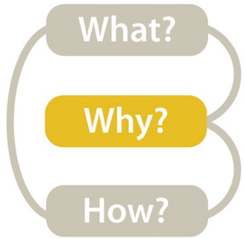
Low-level → After a target(s) is identified, query

- Analyze
  - Consume
    - Discover
    - Present
    - Enjoy
  - Produce
    - Annotate
    - Record
    - Derive
- Search

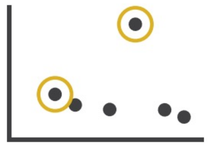
	Target known	Target unknown
Location known	Lookup	Browse
Location unknown	Locate	Explore

- Query
  - Identify
  - Compare
  - Summarize

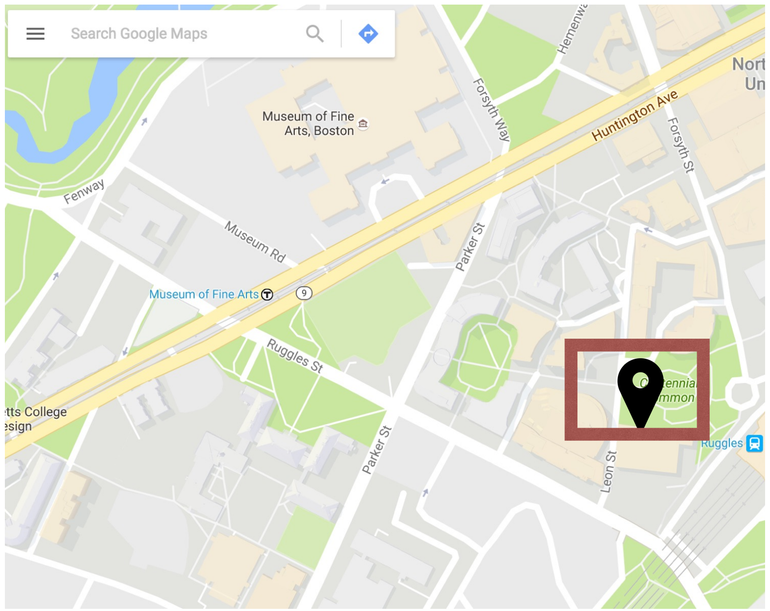
Task Abstraction

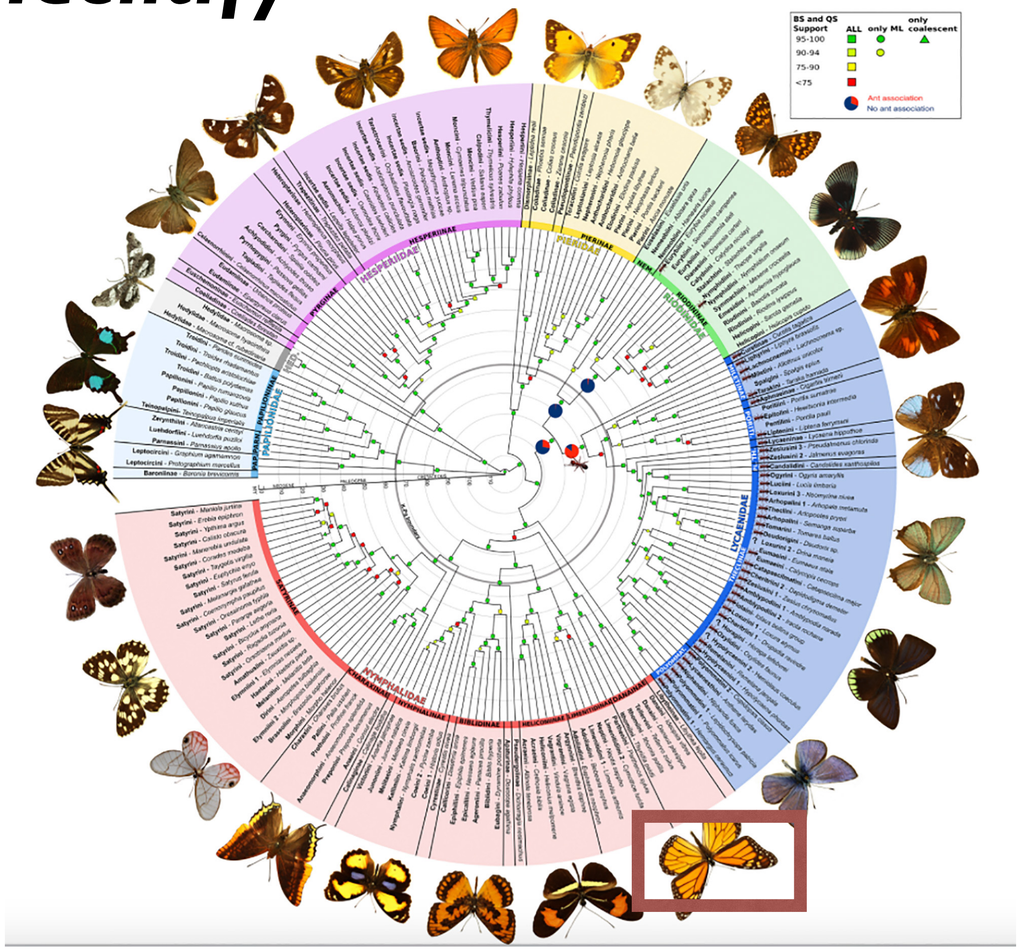


Low-level → ***Query*** → ***Identify***



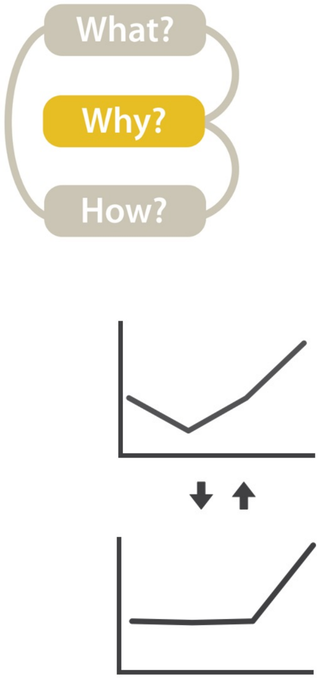
single target





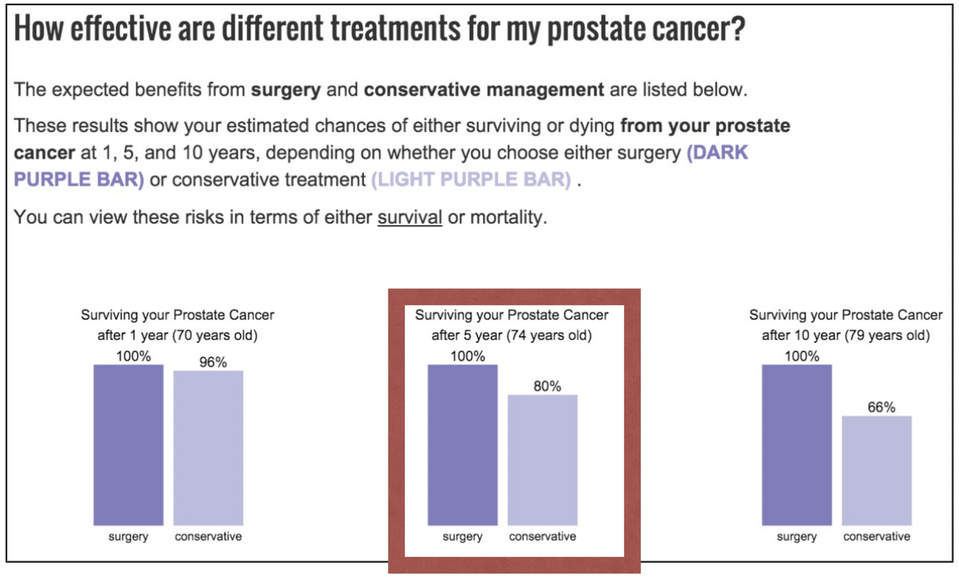
https://www.floridamuseum.ufl.edu/science/at-last-butterflies-get-a-bigger-better-evolutionary-tree/

Task Abstraction

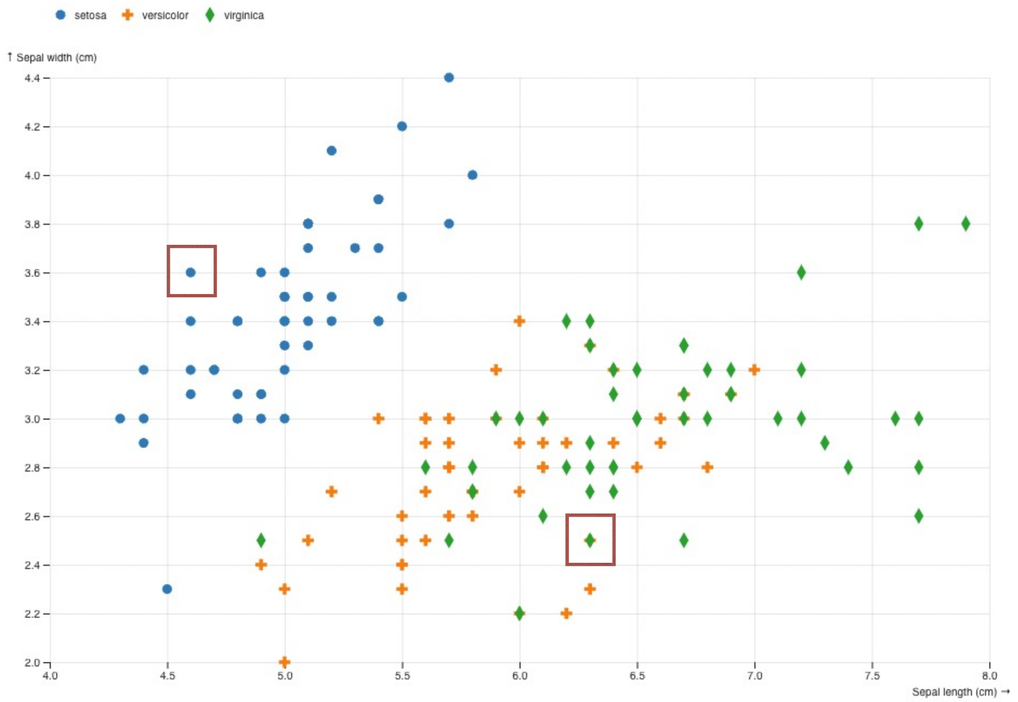


multiple targets

Low-level → ***Query*** → ***Compare***

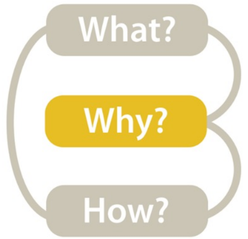


PROACT: Iterative Design of a Patient-Centered Visualization for Effective Prostate Cancer Health Risk Communication. Hakone et al. 2016.



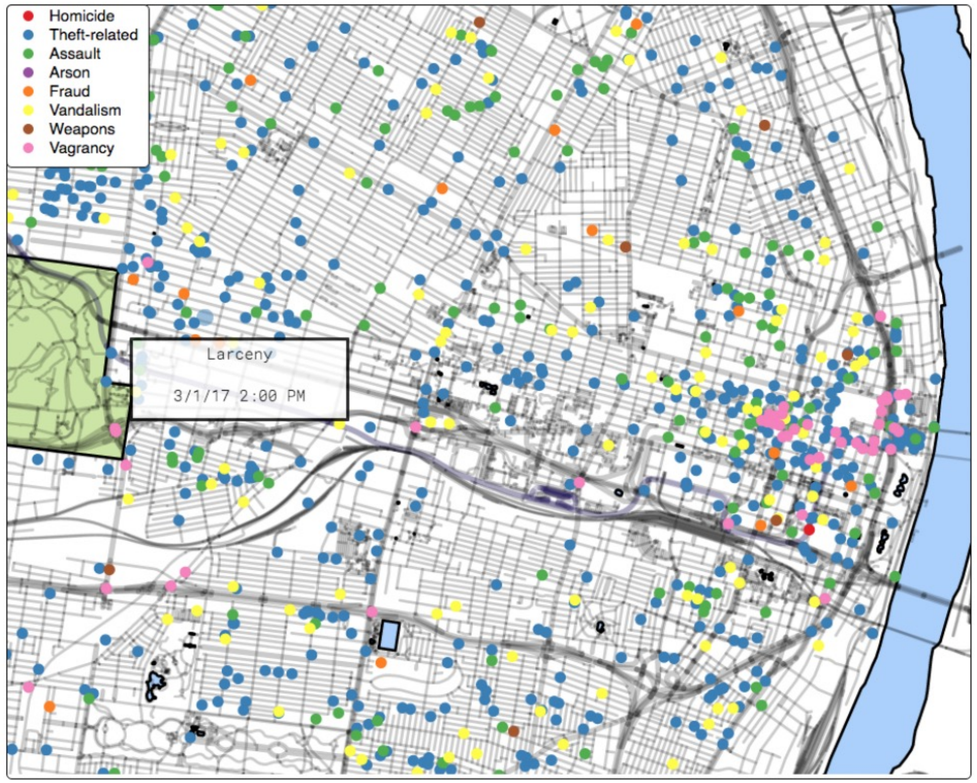
https://observablehq.com/@d3/scatterplot-with-shapes

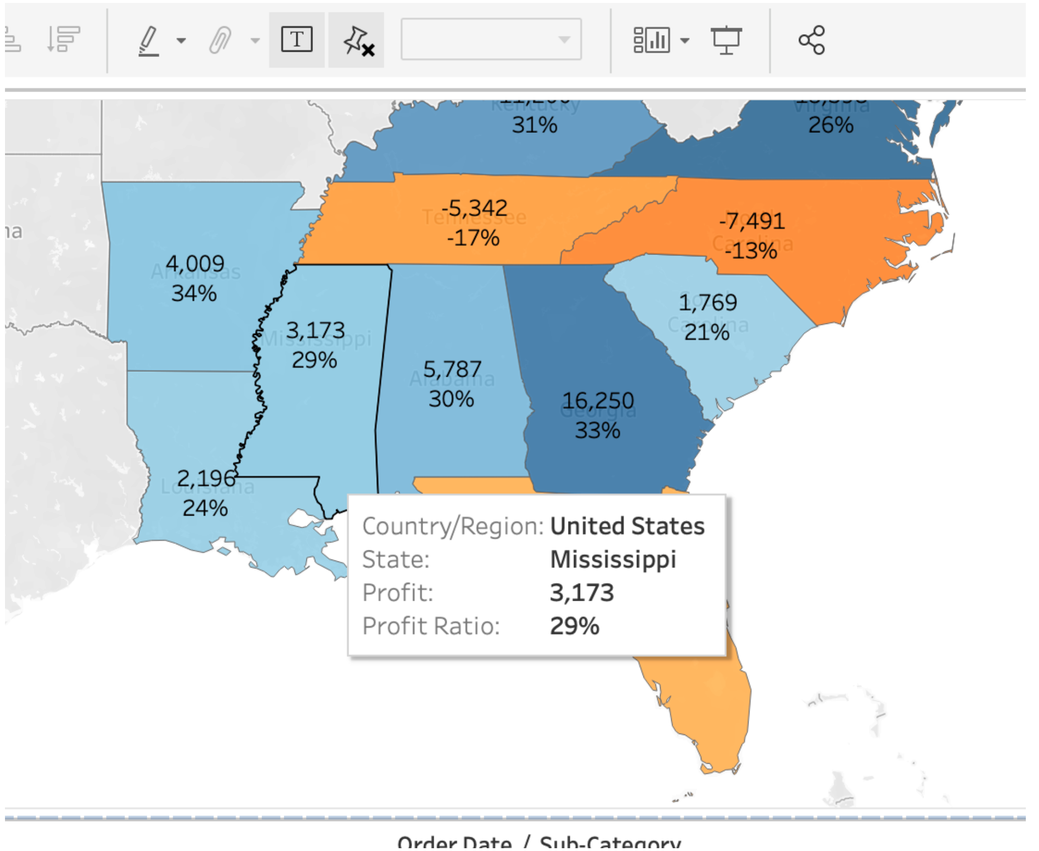
Task Abstraction



all targets

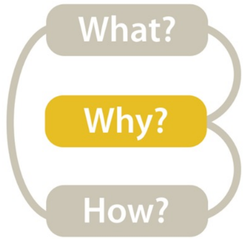
Low-level → ***Query*** → ***Summarize***





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

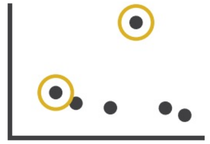
Task Abstraction



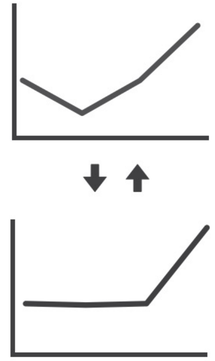
Low-level **→ *After a target(s) is identified, query***

→ Query

→ Identify



→ Compare



→ Summarize

single target

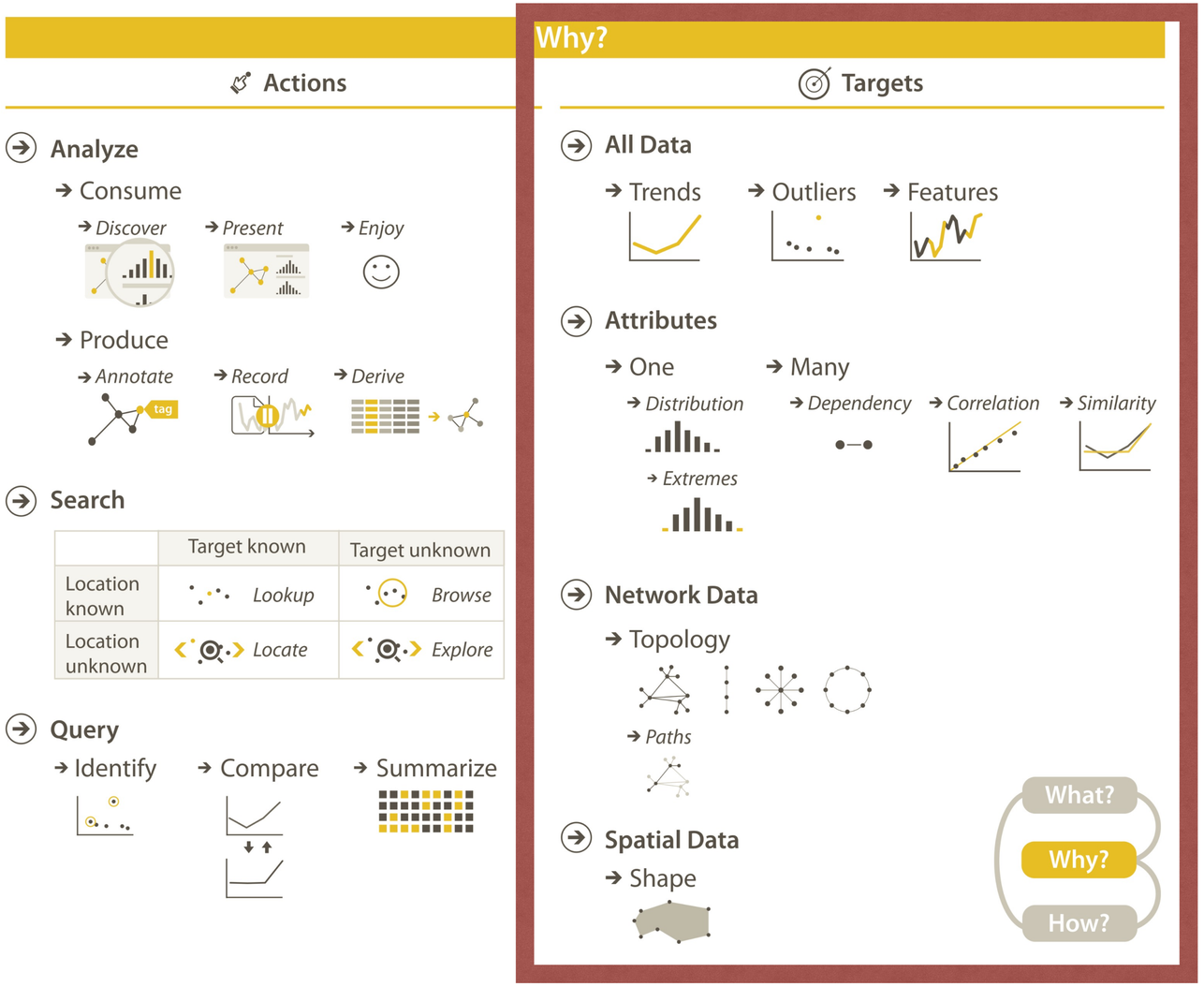
multiple targets

all targets

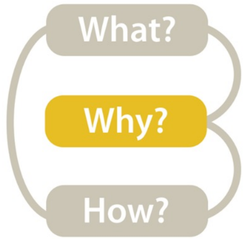
Task Abstraction

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

Attributes

- → One
  - → *Distribution*
  - → *Extremes*
- → Many
  - → *Dependency*
  - → *Correlation*
  - → *Similarity*

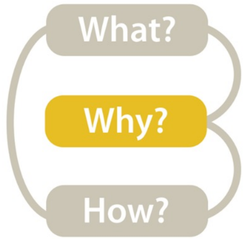
Network Data

- → Topology
- → *Paths*

Spatial Data

- → Shape

Task Abstraction



Targets → some aspect of data of interest to the user

All Data

- → Trends
- → Outliers
- → Features

Attributes

- → One
  - → *Distribution*
  - → *Extremes*
- → Many
  - → *Dependency*
  - → *Correlation*
  - → *Similarity*

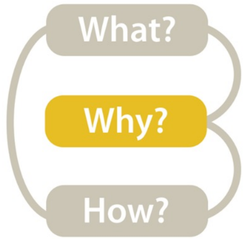
Network Data

- → Topology
- → *Paths*

Spatial Data

- → Shape

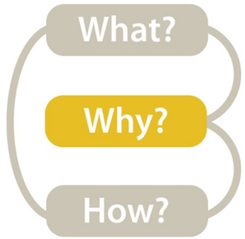
Task Abstraction



Targets* → *some aspect of data of interest to the user

- **All Data**
  - → Trends
  - → Outliers
  - → Features
- **Attributes**
  - → One
    - → *Distribution*
    - → *Extremes*
  - → Many
    - → *Dependency*
    - → *Correlation*
    - → *Similarity*
- **Network Data**
  - → Topology
  - → *Paths*
- **Spatial Data**
  - → Shape

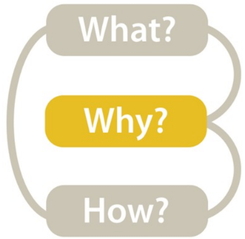
Task Abstraction



Targets* → *some aspect of data of interest to the user

- **All Data**
  - → Trends
  - → Outliers
  - → Features
- **Attributes**
  - → One
    - → *Distribution*
    - → *Extremes*
  - → 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

Attributes

- → One
  - → *Distribution*
  - → *Extremes*
- → Many
  - → *Dependency*
  - → *Correlation*
  - → *Similarity*

Network Data

- → Topology
- → *Paths*

Spatial Data

- → Shape

Example Task Abstraction

Task Abstraction

I need a visualization for performing **cellular analysis**!

Need to **compare** measure A to B over time.

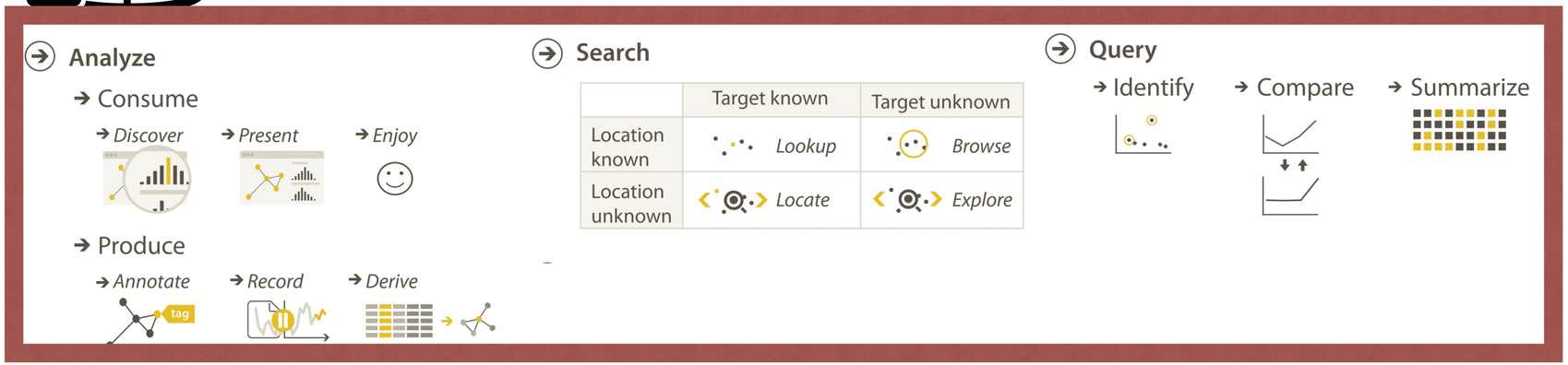
High-level →
? →
?

Medium-level →
Search →
?

Low-level →
Query →
?

Task Abstraction

I need a visualization for performing **cellular analysis**!



Need to **compare** measure A to B over time.

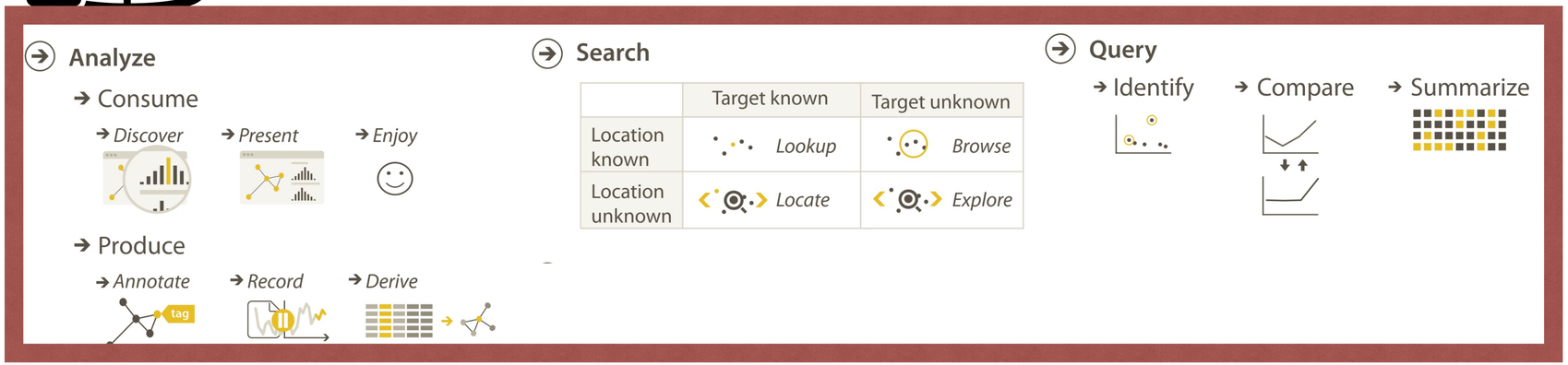
High-level →
? →
?

Medium-level →
Search →
?

Low-level →
Query →
?

Task Abstraction

I need a visualization for performing **cellular analysis**!



Need to **compare** measure A to B over time.

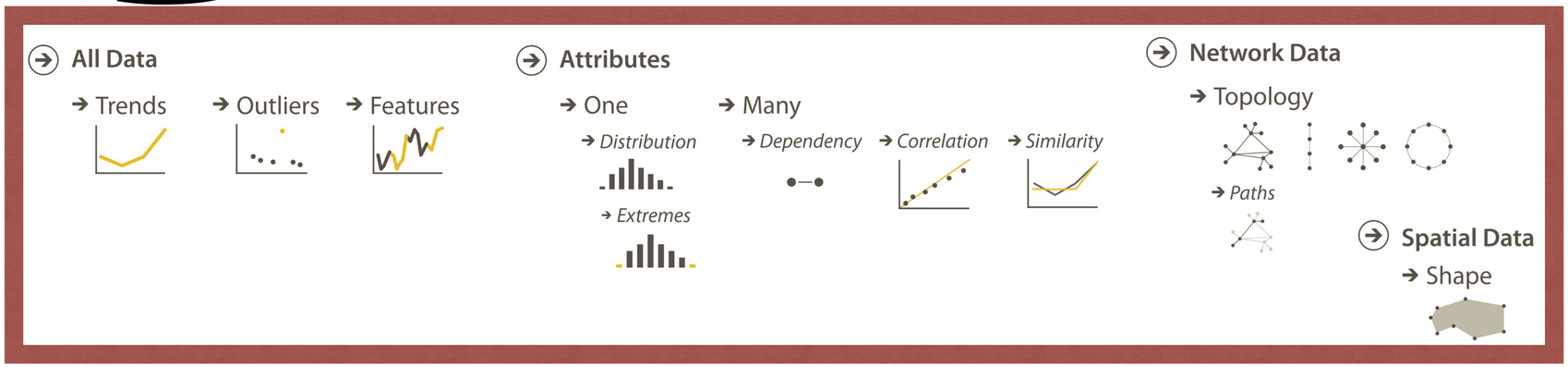
High-level →
Produce →
Derive

Medium-level →
Search →
Lookup (or Locate)

Low-level →
Query →
Compare

Task Abstraction

I need a visualization for performing **cellular analysis**!



Need to **compare** measure A to B over time.

High-level →
Produce →
Derive

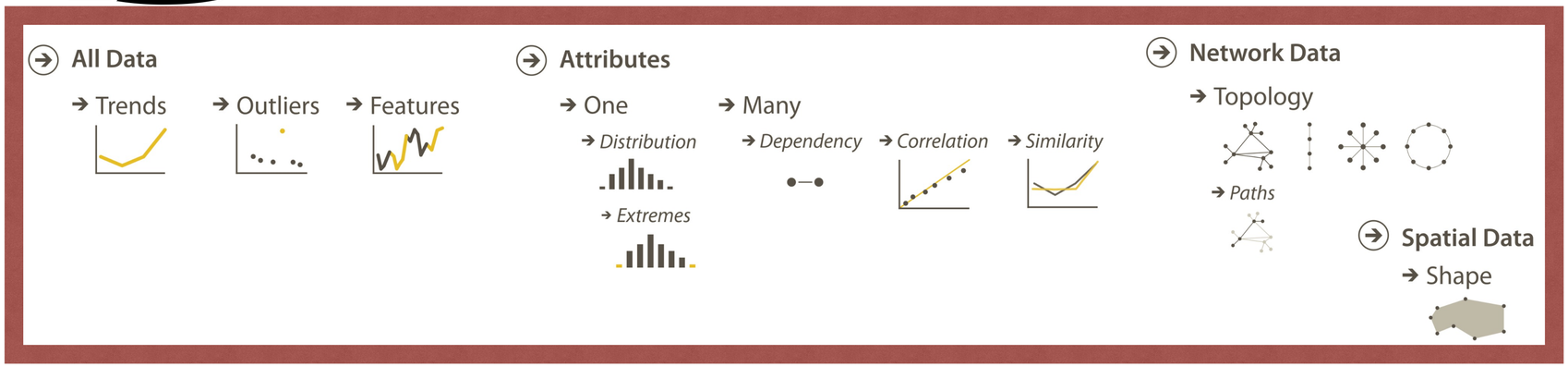
Medium-level →
Search →
Lookup

Low-level →
Query →
Compare

Target(s)

Task Abstraction

I need a visualization for performing **cellular analysis**!



Need to **compare** measure A to B over time.

High-level → **Produce** → **Derive**

Medium-level → **Search** → **Lookup**

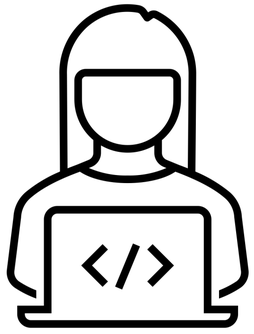
Low-level → **Query** → **Compare**

Target(s) → **Attributes** → **Similarity**

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

Interviewing Domain Experts

Human-Centered Design



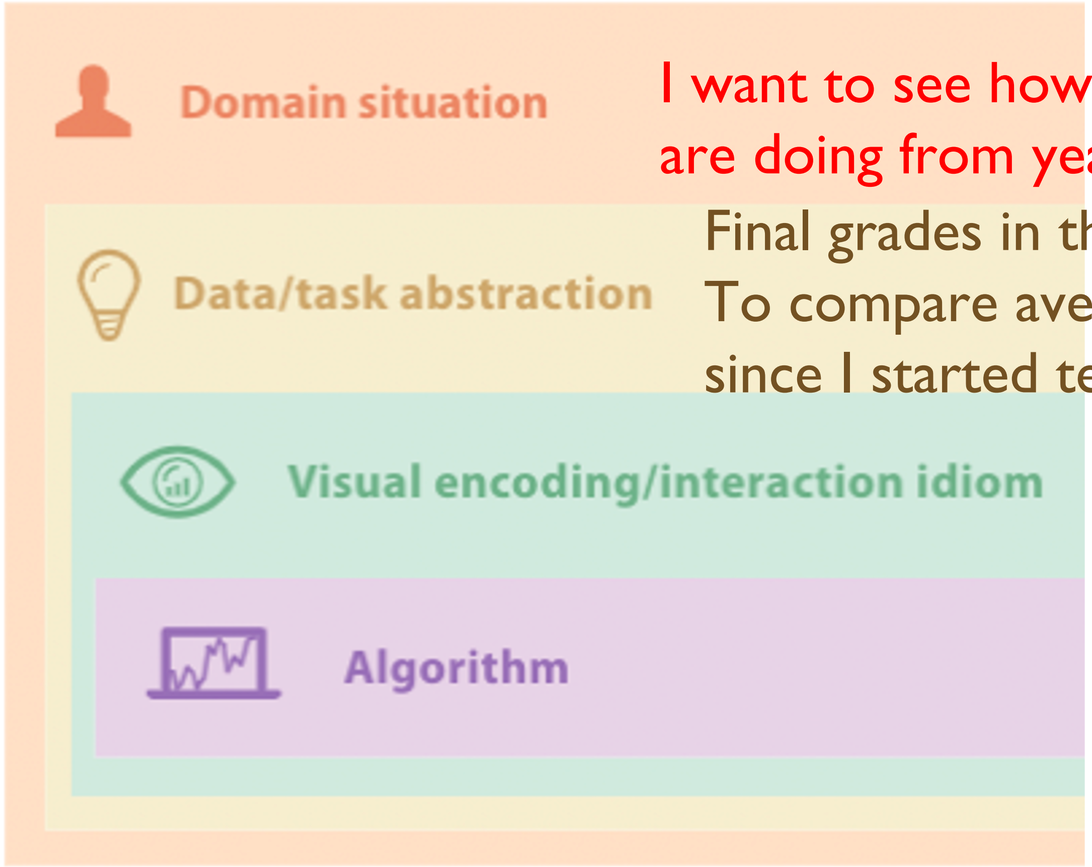
What do you need a visualization for?

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.



I want to see how my students are doing from year to year.

Final grades in the class.

To compare average grade of each class since I started teaching.

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 OUT today.
ic-03 is DUE today.