

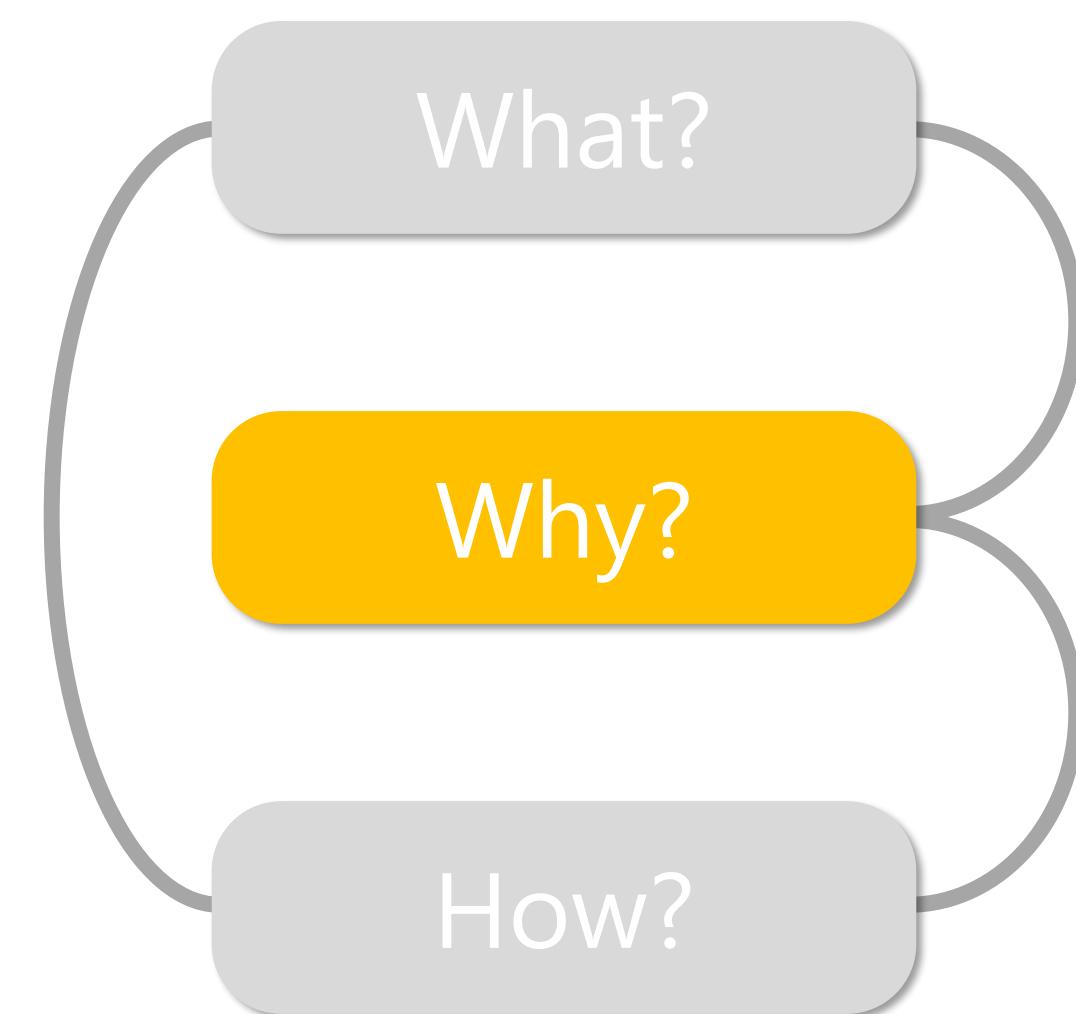


Visualization of Biological Data – Winter Term 2018/2019

Why: Task Abstraction

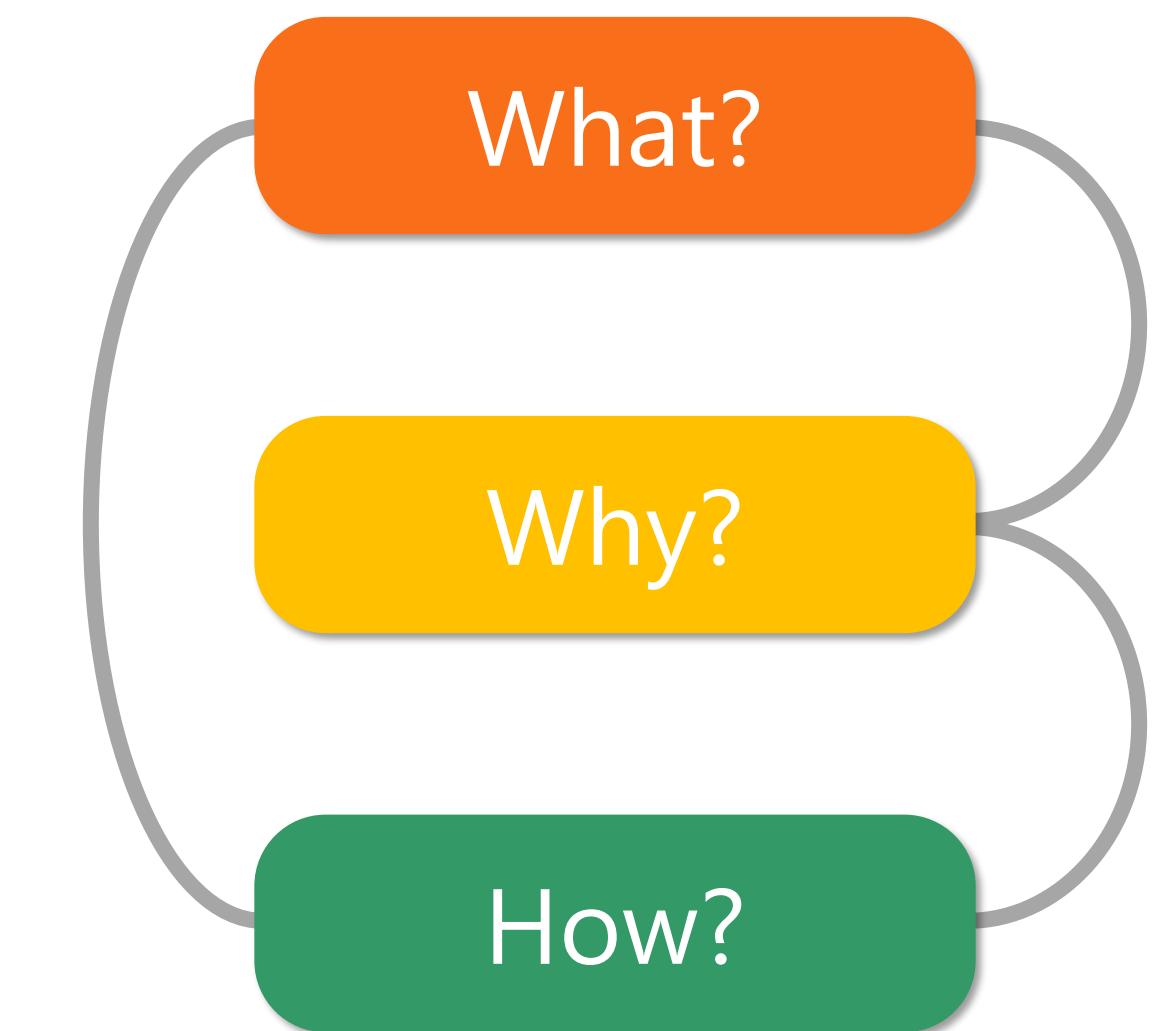
Jun.-Prof. Dr. Michael Krone

29.10.2018



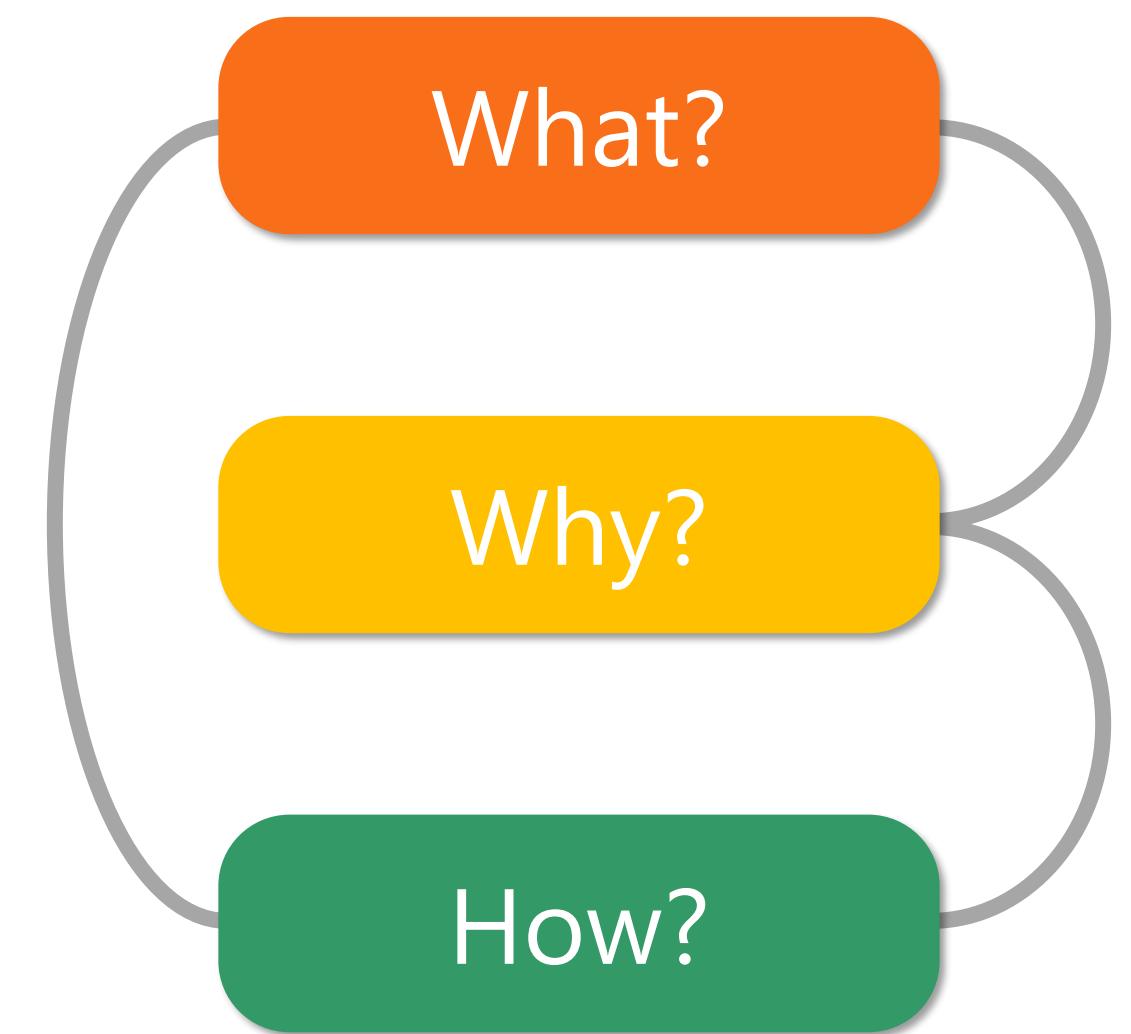
Visualisation Analysis Instance

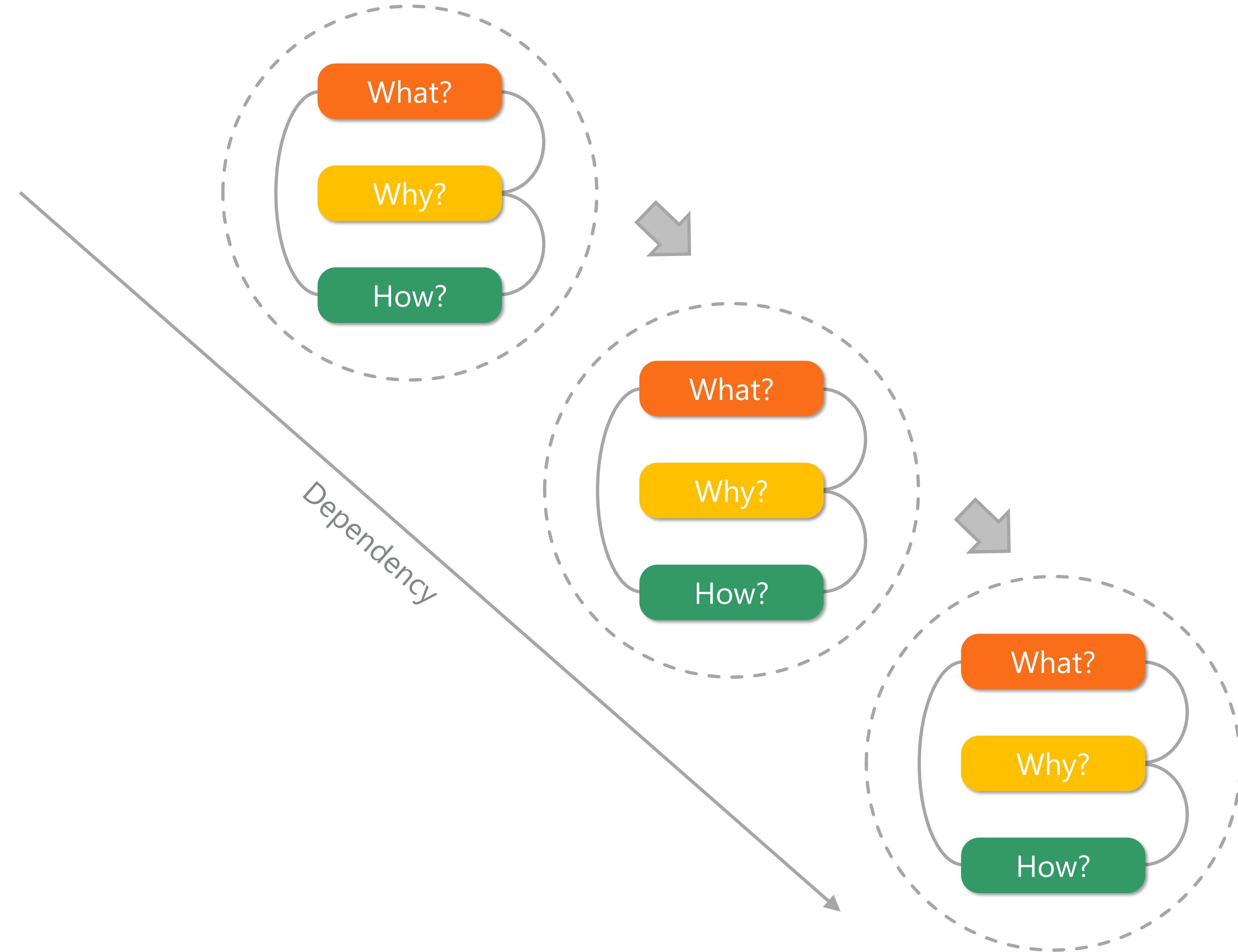
- What data user sees
- Why the user wants to use the tool
- How is it constructed in terms of the design choices



Who is the User?

- User of visualisation
- Designer of visualisation





Visualization Continuum



What?

What?

Why?

How?

Datasets

→ Data Types

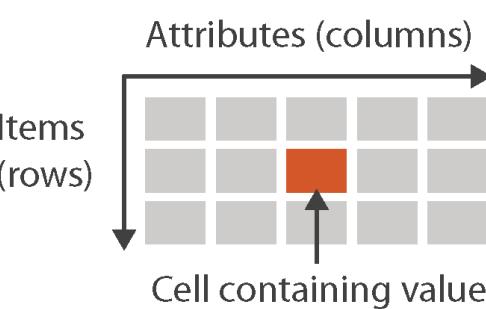
- Items
- Attributes
- Links
- Positions
- Grids

→ Data and Dataset Types

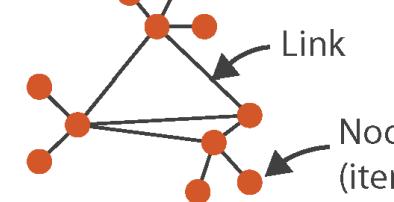
Tables	Networks & Trees	Fields	Geometry	Clusters, Sets, Lists
Items	Items (nodes)	Grids	Items	Items
Attributes	Links	Positions	Attributes	

→ Dataset Types

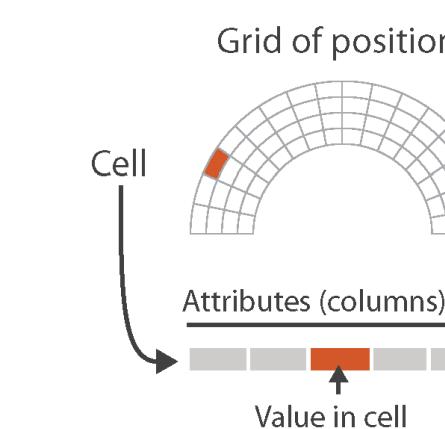
→ Tables



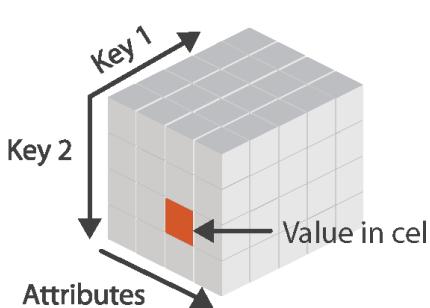
→ Networks



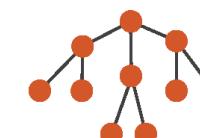
→ Fields (Continuous)



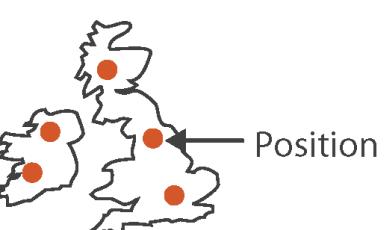
→ Multidimensional Table



→ Trees



→ Geometry (Spatial)



Attributes

→ Attribute Types

- Categorical



- Ordered

- *Ordinal*



- *Quantitative*



→ Ordering Direction

- Sequential



- Diverging



- Cyclic



→ Dataset Availability

→ Static



→ Dynamic



What?

Why?

How?

Why?

Actions

Targets

→ Analyze

→ Consume

→ Discover



→ Present



→ Enjoy

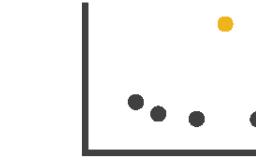


→ All Data

→ Trends



→ Outliers

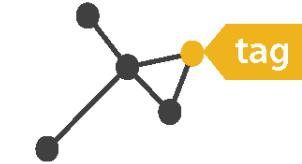


→ Features



→ Produce

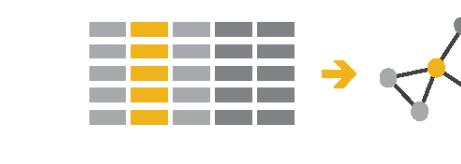
→ Annotate



→ Record



→ Derive

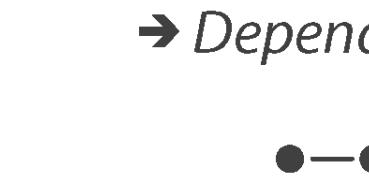


→ Attributes

→ One



→ Many



→ Distribution



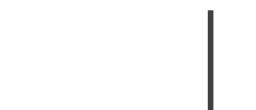
→ Dependency



→ Correlation



→ Similarity



→ Extremes

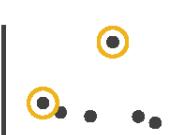


→ Search

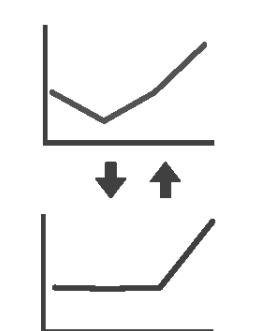
	Target known	Target unknown
Location known	••• <i>Lookup</i>	••• <i>Browse</i>
Location unknown	◁ 🔎 ▷ <i>Locate</i>	◁ 🔎 ▷ <i>Explore</i>

→ Query

→ Identify



→ Compare

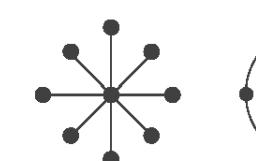
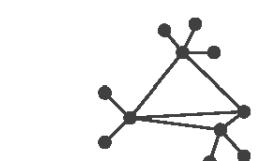


→ Summarize



→ Network Data

→ Topology

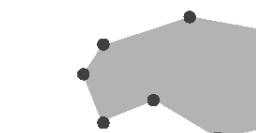


→ Paths



→ Spatial Data

→ Shape



Munzner, 2014



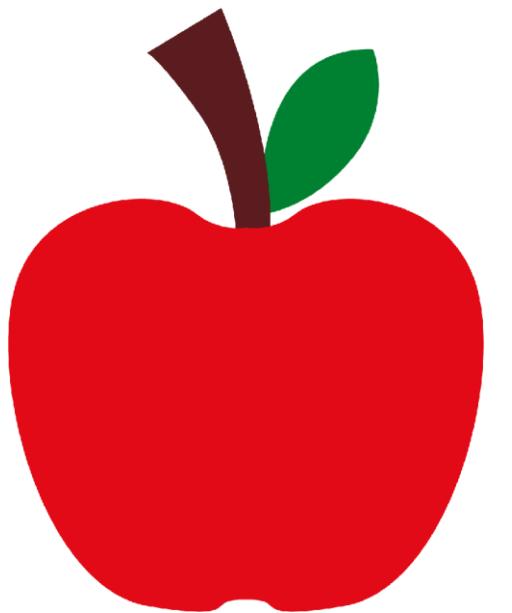
Tasks in domain specific language

- **Task 1:**
Find differences between gut microbioms of patients treated with penicillin and those treated with placebo.
- **Task 2:**
Check if the white blood cell size is the same in blood samples taken before sugary drink consumption and after.



Tasks abstracted

- **Task 1:**
Compare the two groups.
- **Task 2:**
Compare the two groups.



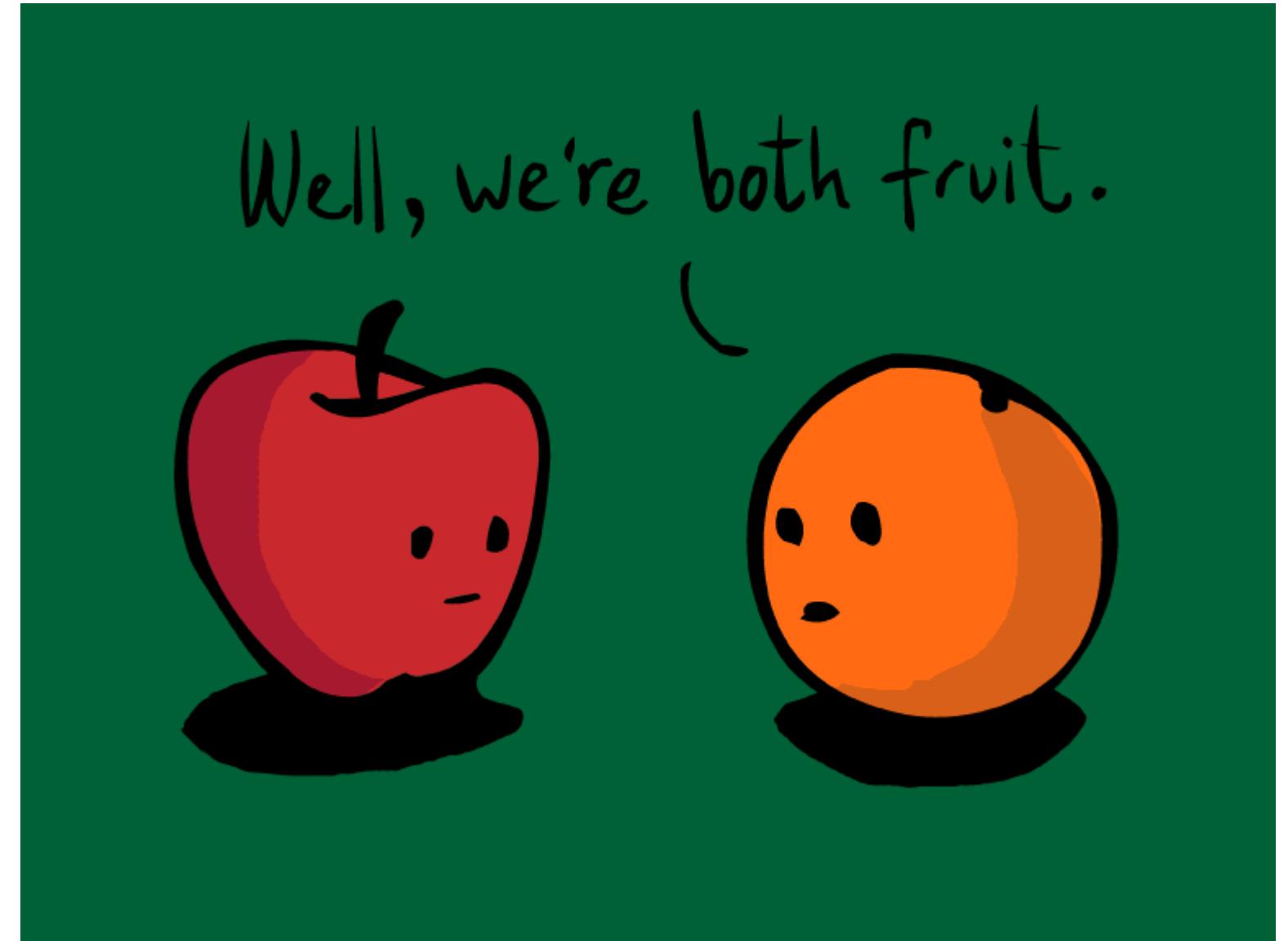
Why analyze tasks abstractly?

- To compare
- To guide data abstraction



Synonyms for Compare

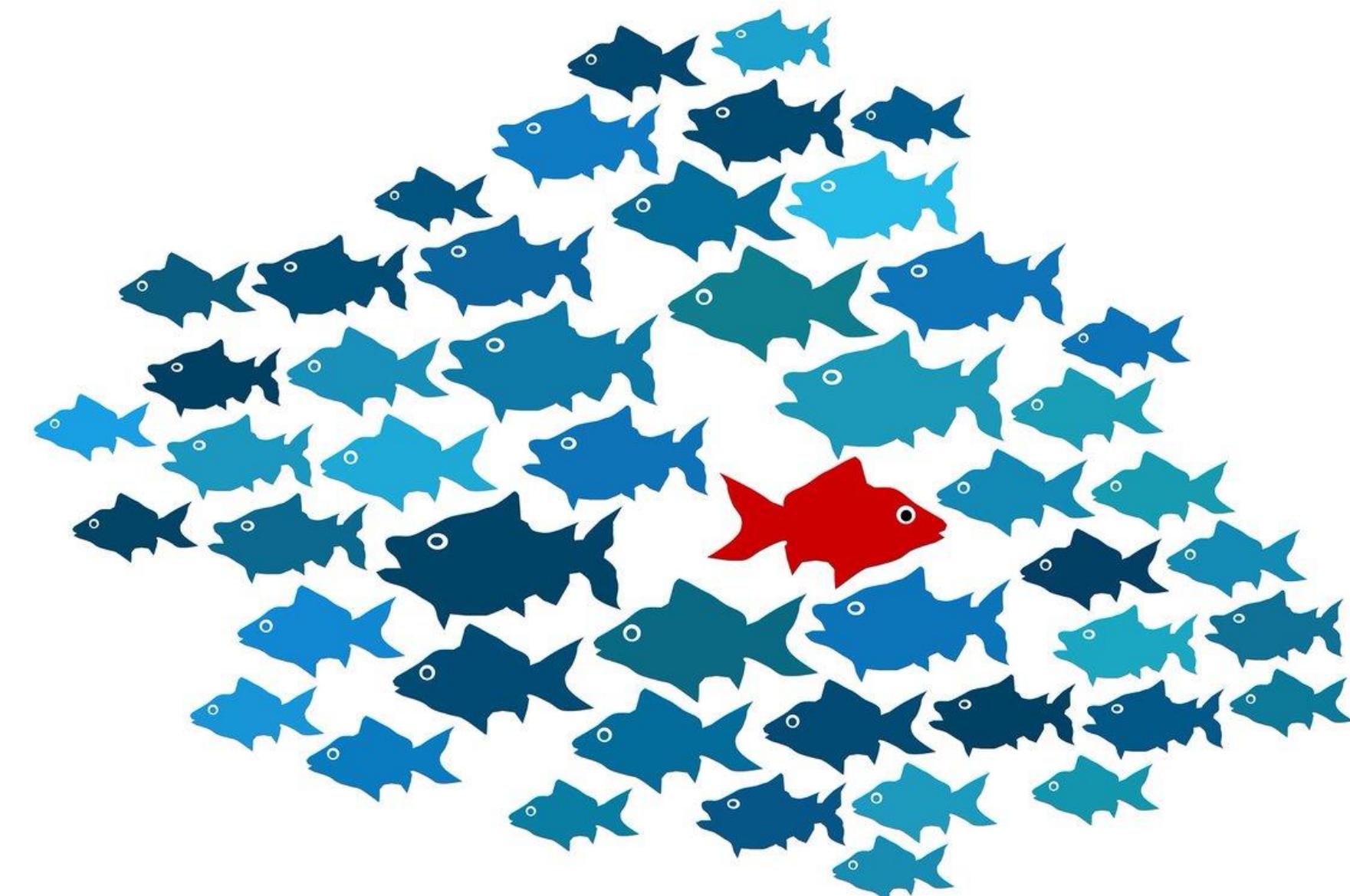
- Contrast
- Correlate
- Match (up)
- Equate
- Oppose
- Find differences/similarities
- Others?



<https://osboncapital.com/5-ways-to-compare-investment-advisors/>

Task Abstraction

- **Actions** define goals
- **Targets** are aspects of data that are of interest
- e.g., **Present** the results by **identifying outliers**.



<https://towardsdatascience.com/a-brief-overview-of-outlier-detection-techniques-1e0b2c19e561>

What?

Why?

How?

{action, target} pairs
→ *discover distribution*
→ *compare trends*
→ *locate outliers*
→ *browse topology*

Why?

Targets

Actions

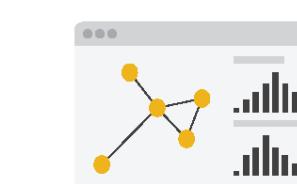
→ Analyze

→ Consume

→ Discover



→ Present

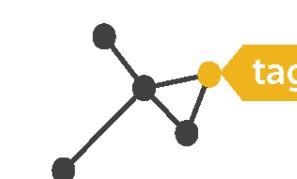


→ Enjoy



→ Produce

→ Annotate



→ Record



→ Derive

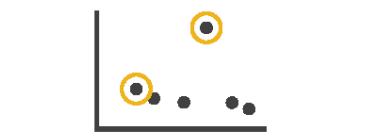


→ Search

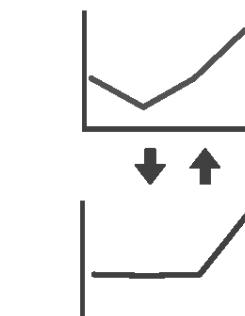
	Target known	Target unknown
Location known	••• <i>Lookup</i>	••○ <i>Browse</i>
Location unknown	◁○▷ <i>Locate</i>	◁○▷ <i>Explore</i>

→ Query

→ Identify



→ Compare



→ Summarize

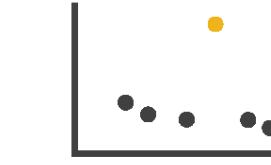


→ All Data

→ Trends



→ Outliers



→ Features



→ Attributes

→ One



→ Distribution



→ Extremes

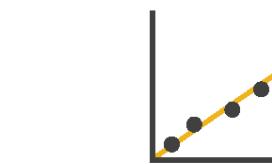


→ Many

→ Dependency



→ Correlation

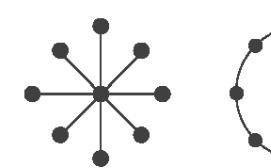
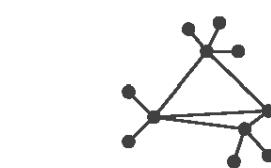


→ Similarity

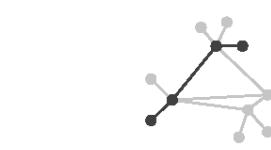


→ Network Data

→ Topology

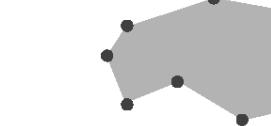


→ Paths



→ Spatial Data

→ Shape



Munzner, 2014

Actions

→ Analyze

→ Consume

→ Discover



→ Present

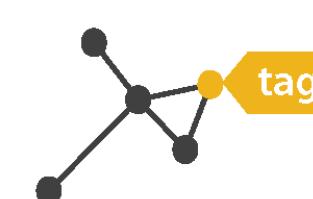


→ Enjoy

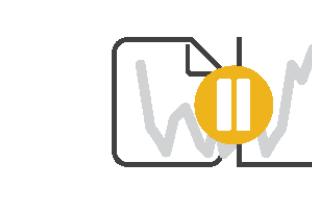


→ Produce

→ Annotate



→ Record



→ Derive

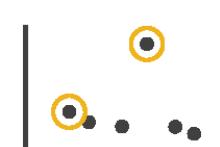


→ Search

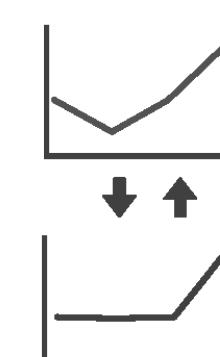
	Target known	Target unknown
Location known	•••• <i>Lookup</i>	•••• <i>Browse</i>
Location unknown	◁•••▷ <i>Locate</i>	◁•••▷ <i>Explore</i>

→ Query

→ Identify



→ Compare



→ Summarize



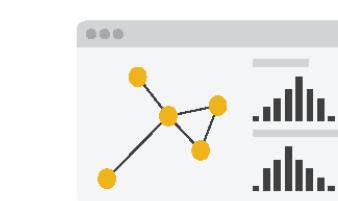
→ Analyze

→ Consume

→ Discover



→ Present

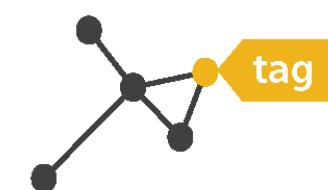


→ Enjoy



→ Produce

→ Annotate



→ Record



→ Derive



Consume → Discover

- Exploratory analysis
- Find new knowledge
- Generate new hypothesis
- Verify existing hypothesis



ABOUT F53

FUNCTION: Acts as a tumor suppressor in many tumor types; induces growth arrest... [+]

COFACTOR: Binds 1 zinc ion per sub-unit.

SUBUNIT: Interacts with AXIN1. Probably part of a complex consisting of TP53, ... [+]

INTERACTION: Self; NbExp=7; In-
tAct=EBI-366083, EBI-366083;
P03070:- (xeno); NbExp=1 ... [+]

SUBCELLULAR LOCATION:

- Cytoplasm. Nucleus. Nucleus, PML body. Endoplasmic reticulum.

[more details \(8\)](#)

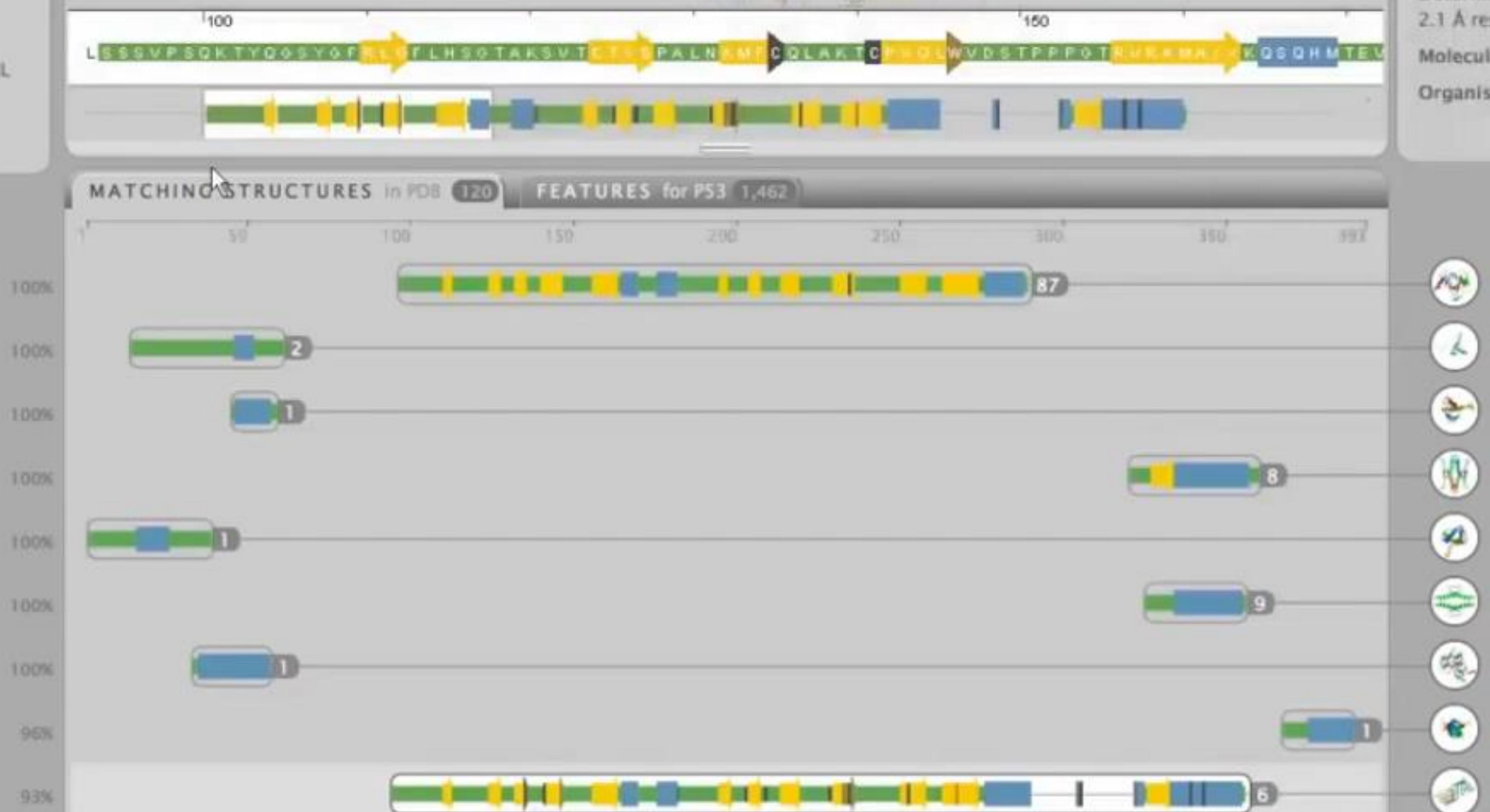


quence-specific DNA binding and homo-tetramerization domains. Interestingly, the affinities of p53 for specific and non-specific DNA sites differ by only one order of magnitude, making it hard to understand how this protein recognizes its specific DNA targets *in vivo*. We describe here the structure of a p53 polypeptide containing both the DNA binding and oligomerization domains in complex with DNA. The structure reveals that sequence-specific DNA binding... [+]

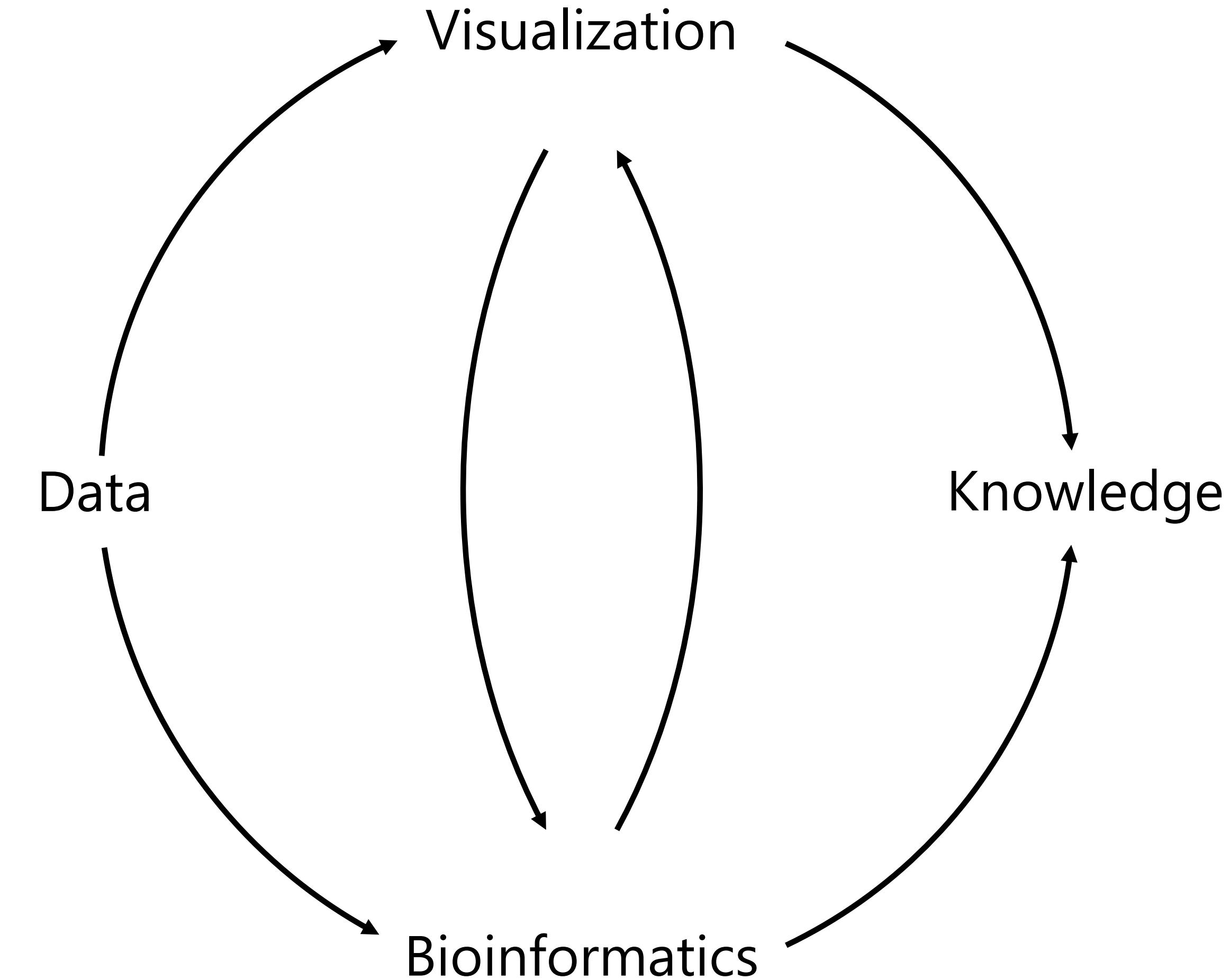
Determined by: X-ray diffraction at
2.1 Å resolution

Molecule: unavailable

Organism: *Homo sapiens*



Exploratory Data Analysis



Consume → Present

- Guiding audience through cognitive operations.
- Explain something that you know to others (communication).
- Telling a story with data.

→ *Present*



Carte Figurative des pertes successives en hommes de l'Armée Française dans la Campagne de Russie 1812-1813.

Dressée par M. Minard, Inspecteur Général des Ponts et Chaussées en retraite
Paris, le 20 Novembre 1869.

Les nombres d'hommes présents sont représentés par les largeurs des zones colorées à raison d'un millimètre pour dix mille hommes; ils sont de plus écrits en travers des zones. Le rouge désigne les hommes qui entrent en Russie, le noir ceux qui en sortent. — Les renseignements qui ont servi à dresser la carte ont été puisés dans les ouvrages de M. M. Chiers, de Ségur, de Fezensac, de Chambray et le journal inédit de Jacob, pharmacien de l'Armée depuis le 28 Octobre.

Pour mieux faire juger à l'œil la diminution de l'armée, j'ai supposé que les corps du Prince Jérôme et du Maréchal Davout qui avaient été détachés sur Minsk en Mohilow et qui rejoignirent vers Orscha et Witebsk, avaient toujours marché avec l'armée.

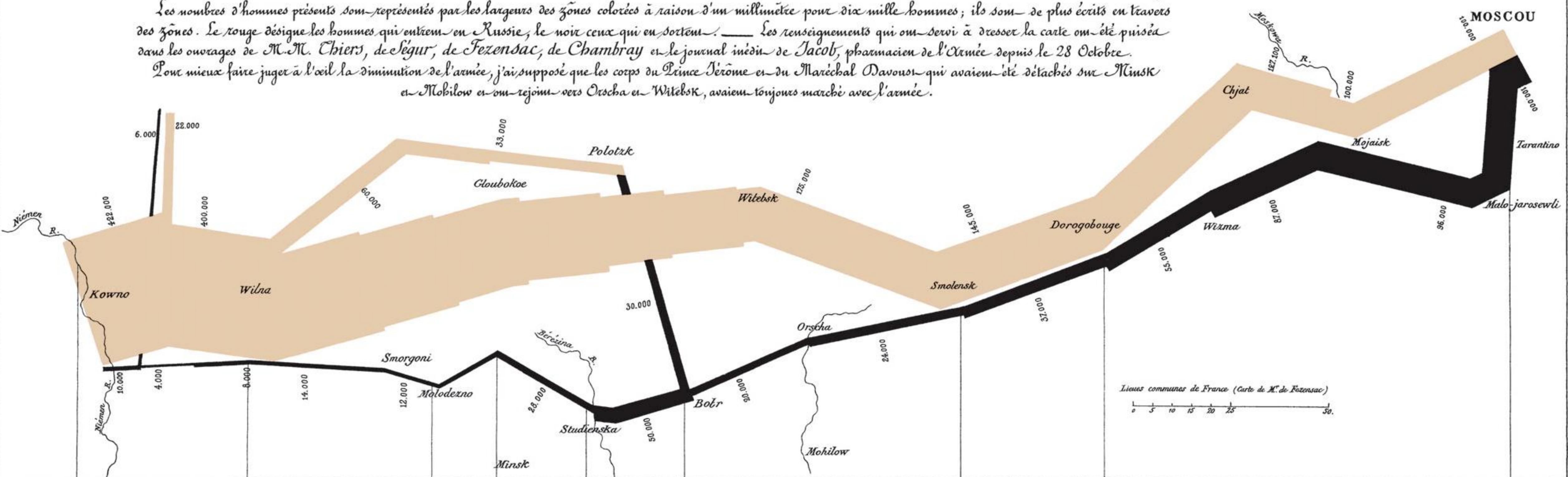
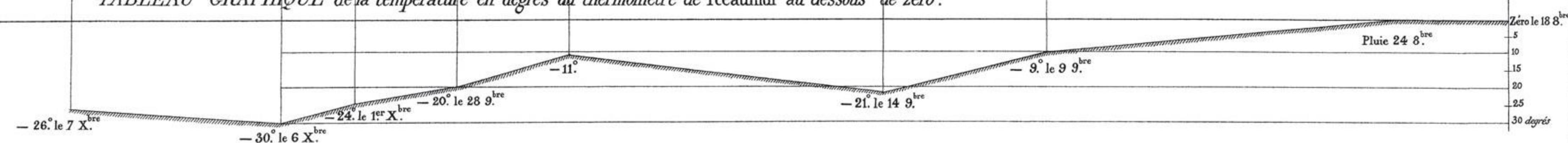


TABLEAU GRAPHIQUE de la température en degrés du thermomètre de Réaumur au dessous de zéro.

Les Cosaques passent au galop
le Niémen gelé.

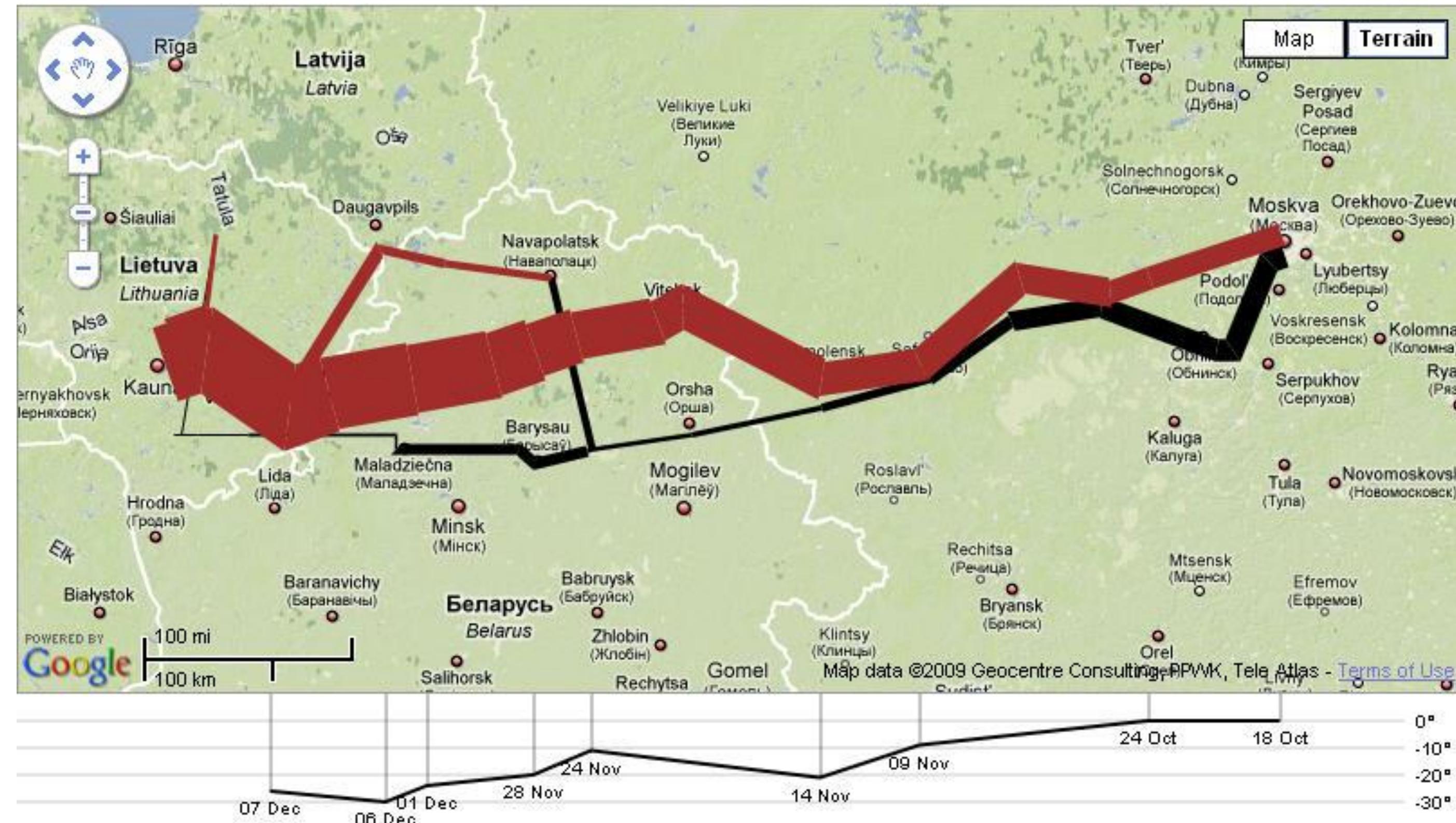


Autog. par Regnier, 8. Pas. S^ete Marie S^t G^ain à Paris.

Imp. Lith. Regnier et Dourdet.



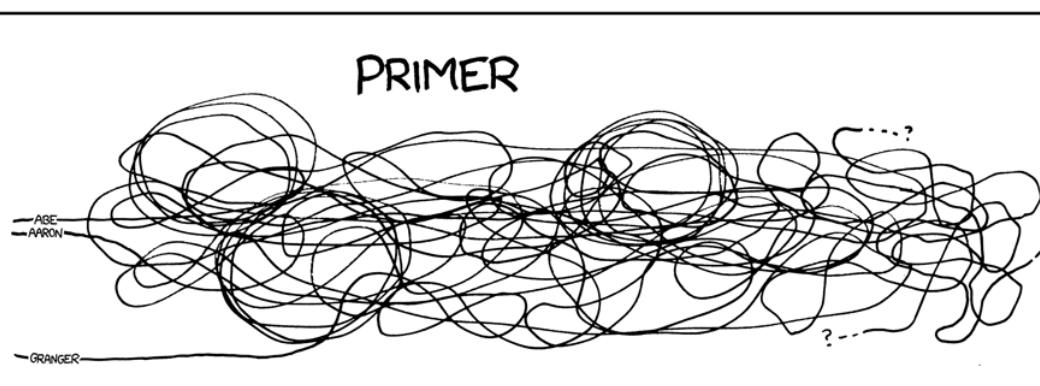
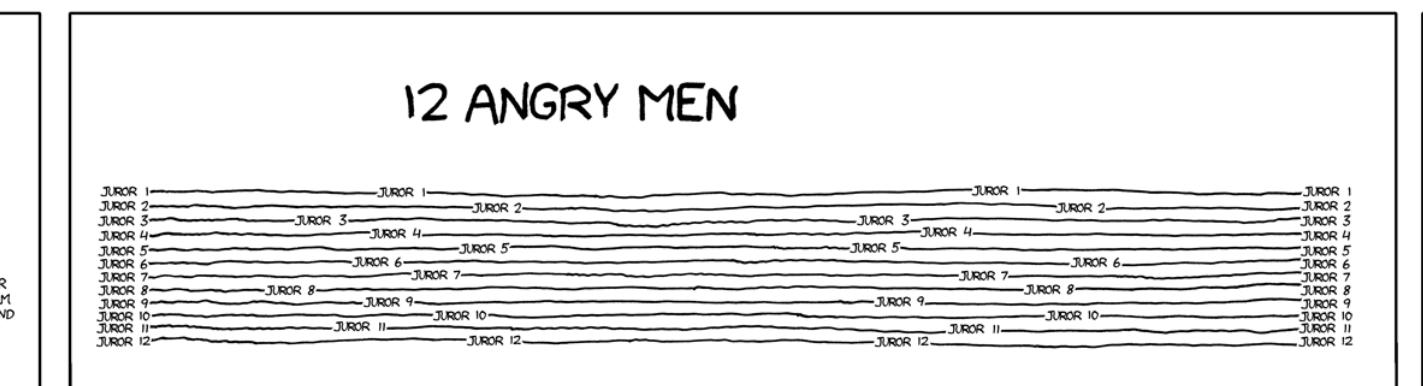
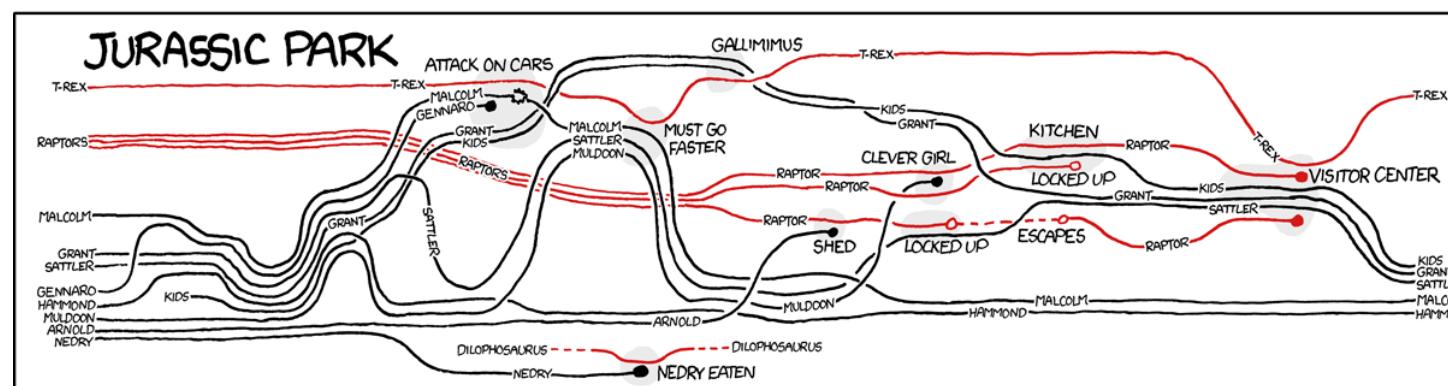
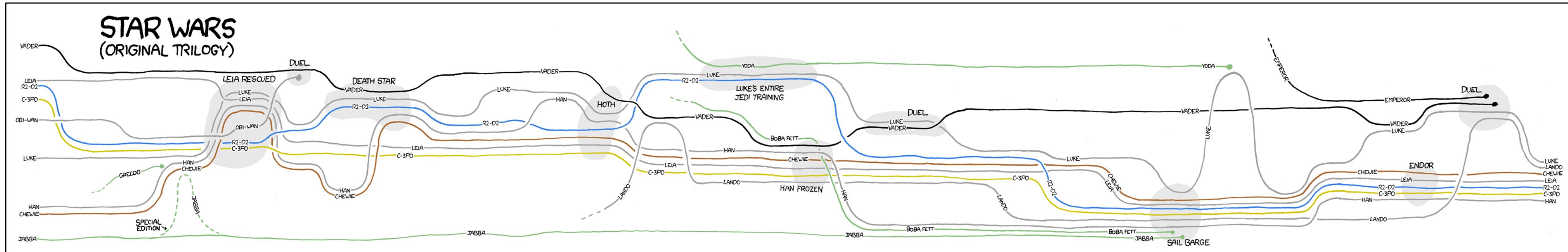
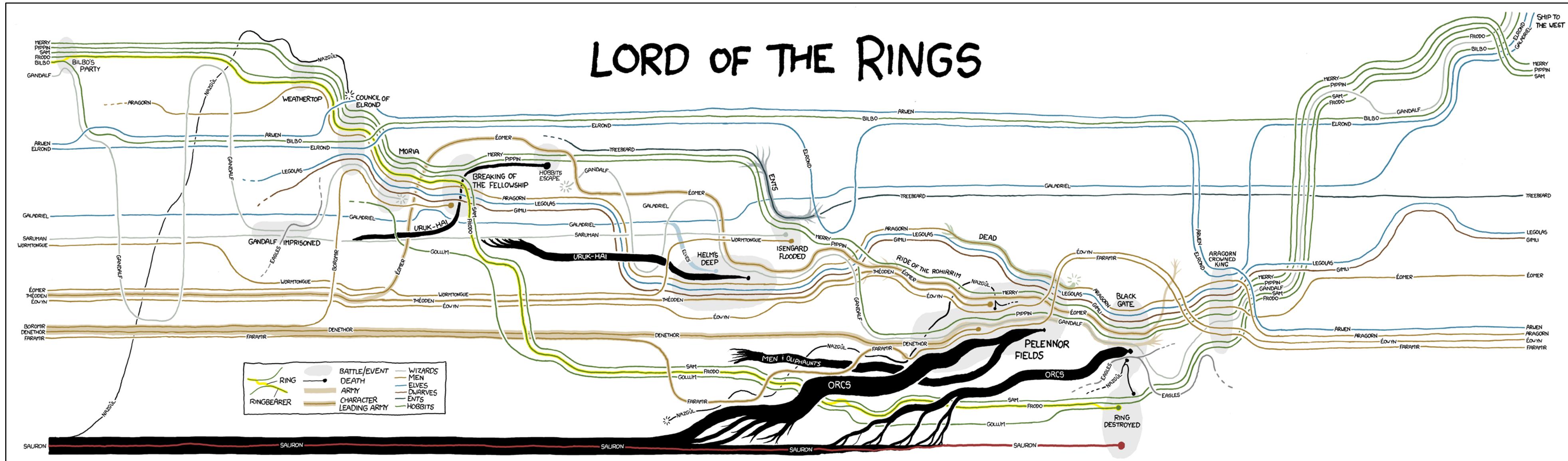
Charles J. Minard (1869): Figurative Map of the successive losses in men of the French Army in the Russian campaign 1812–1813.



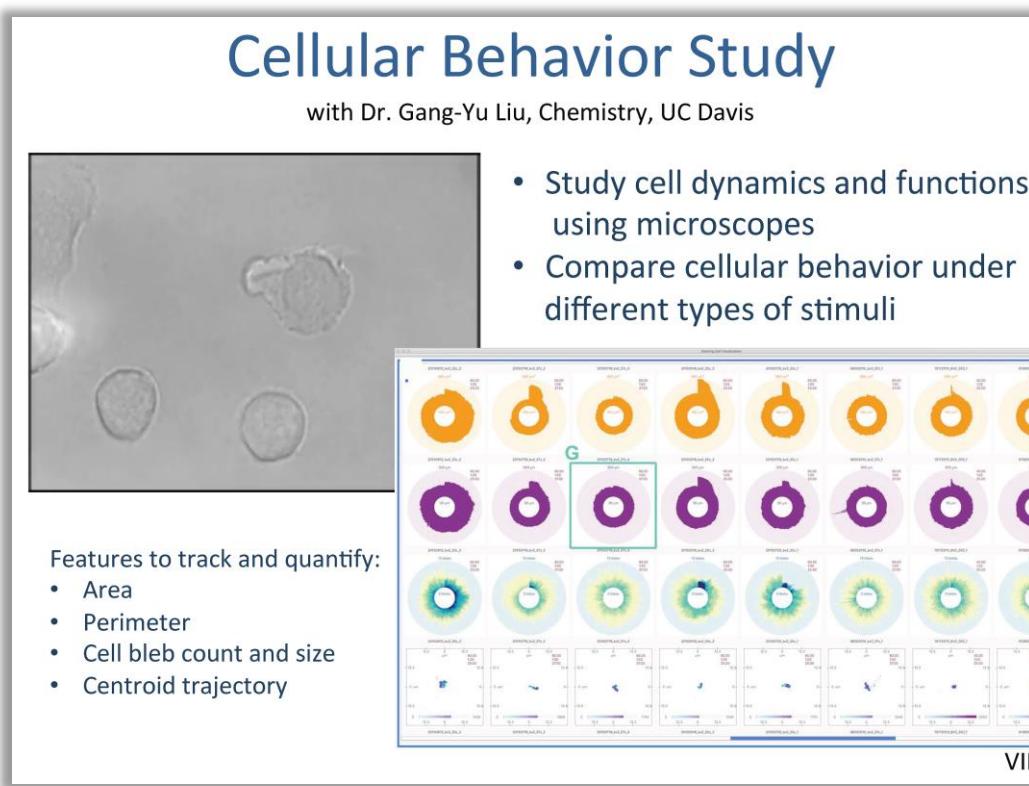
<http://euclid.psych.yorku.ca/SCS/Gallery/re-minard.html>

Storytelling: XKCD

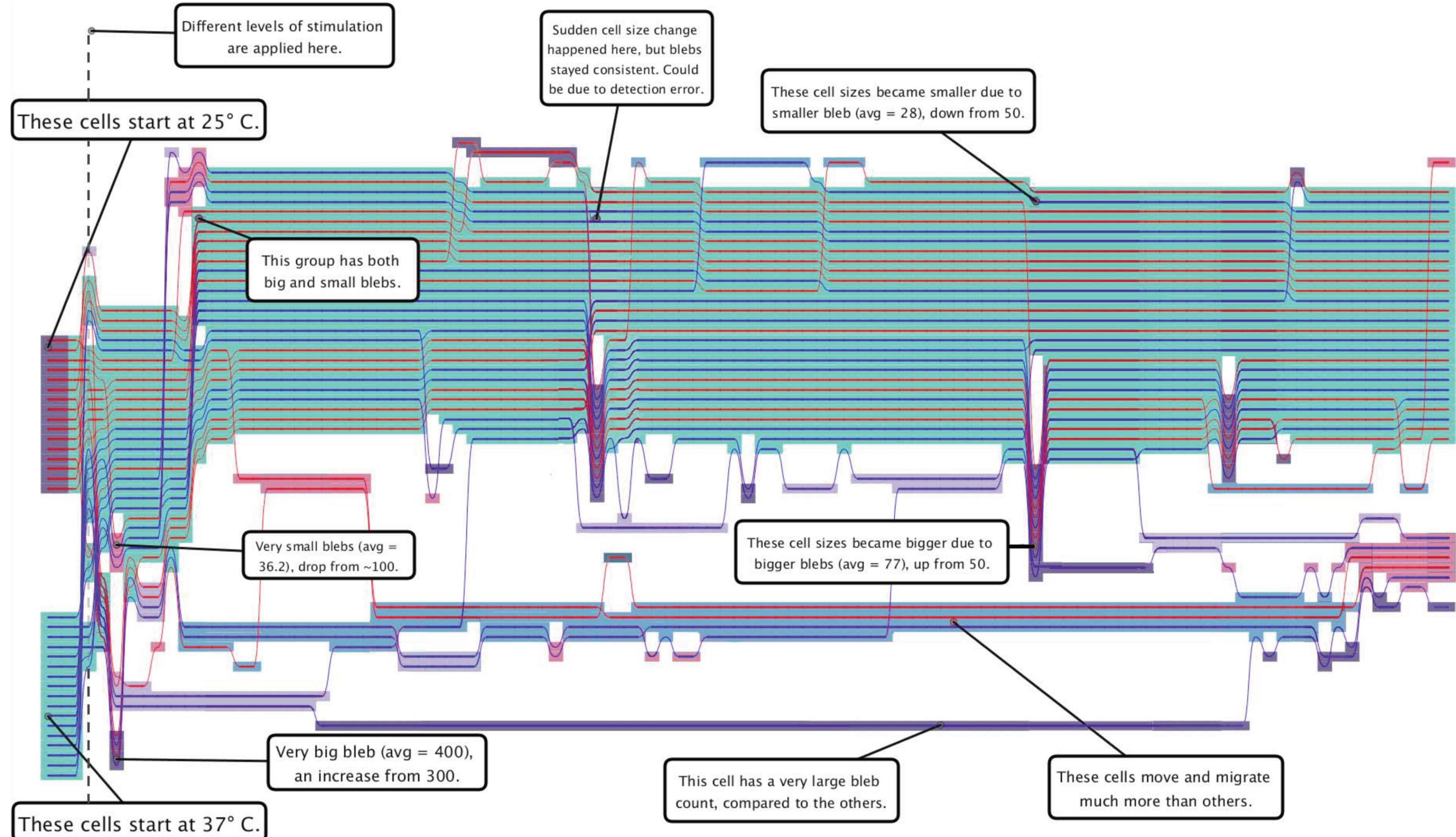
THESE CHARTS SHOW MOVIE CHARACTER INTERACTIONS.
THE HORIZONTAL AXIS IS TIME. THE VERTICAL GROUPING OF THE
LINES INDICATES WHICH CHARACTERS ARE TOGETHER AT A GIVEN TIME.



Storytelling: Cellular behaviour



Kwan-Liu Ma with Gang-Yu Liu, VIZBI 2016



Consume → Enjoy

- Novelty stimulates curiosity and exploration.
- Fleeting (short and casual) or immersive and time consuming (e.g., museum exhibitions).

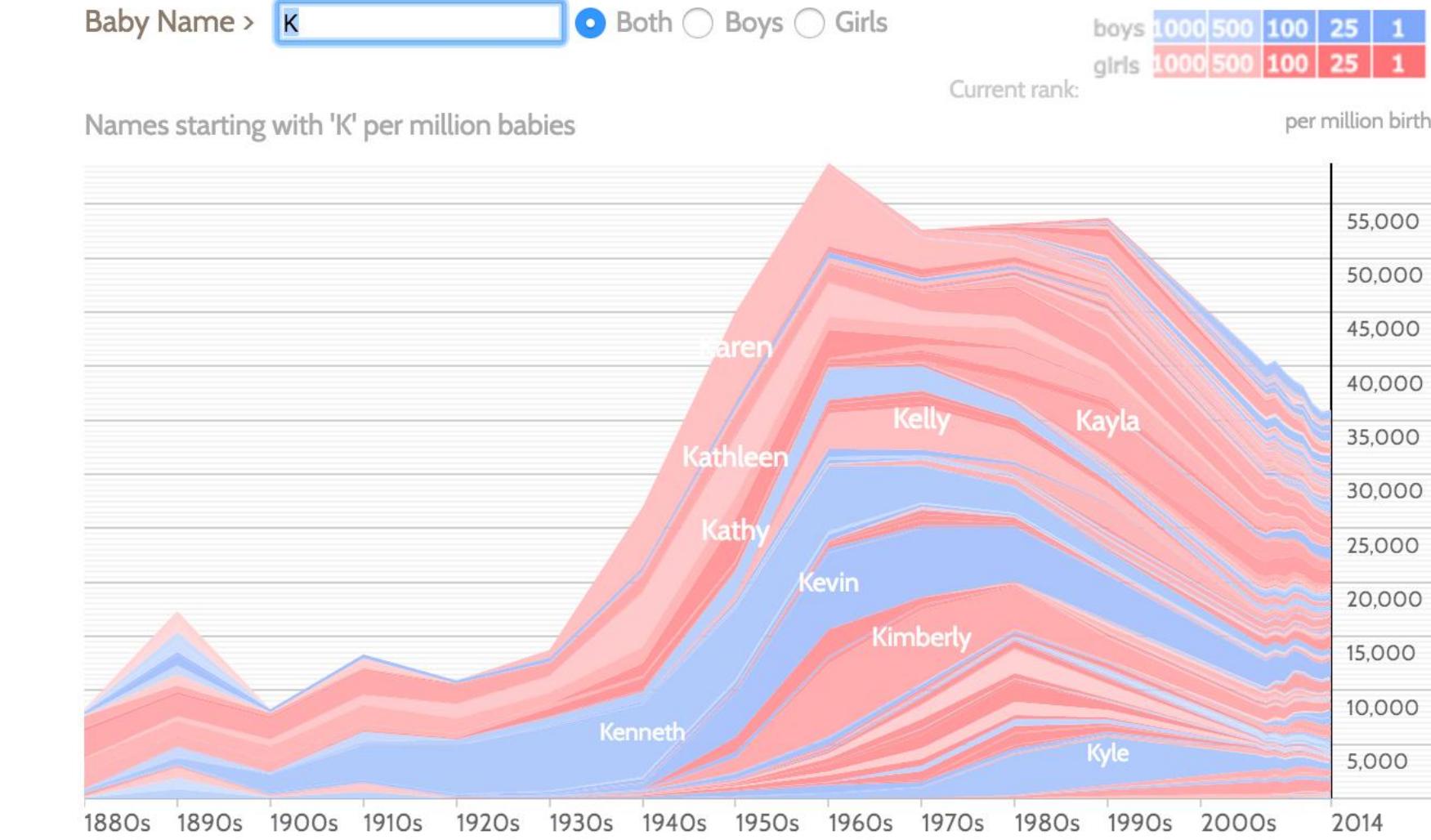
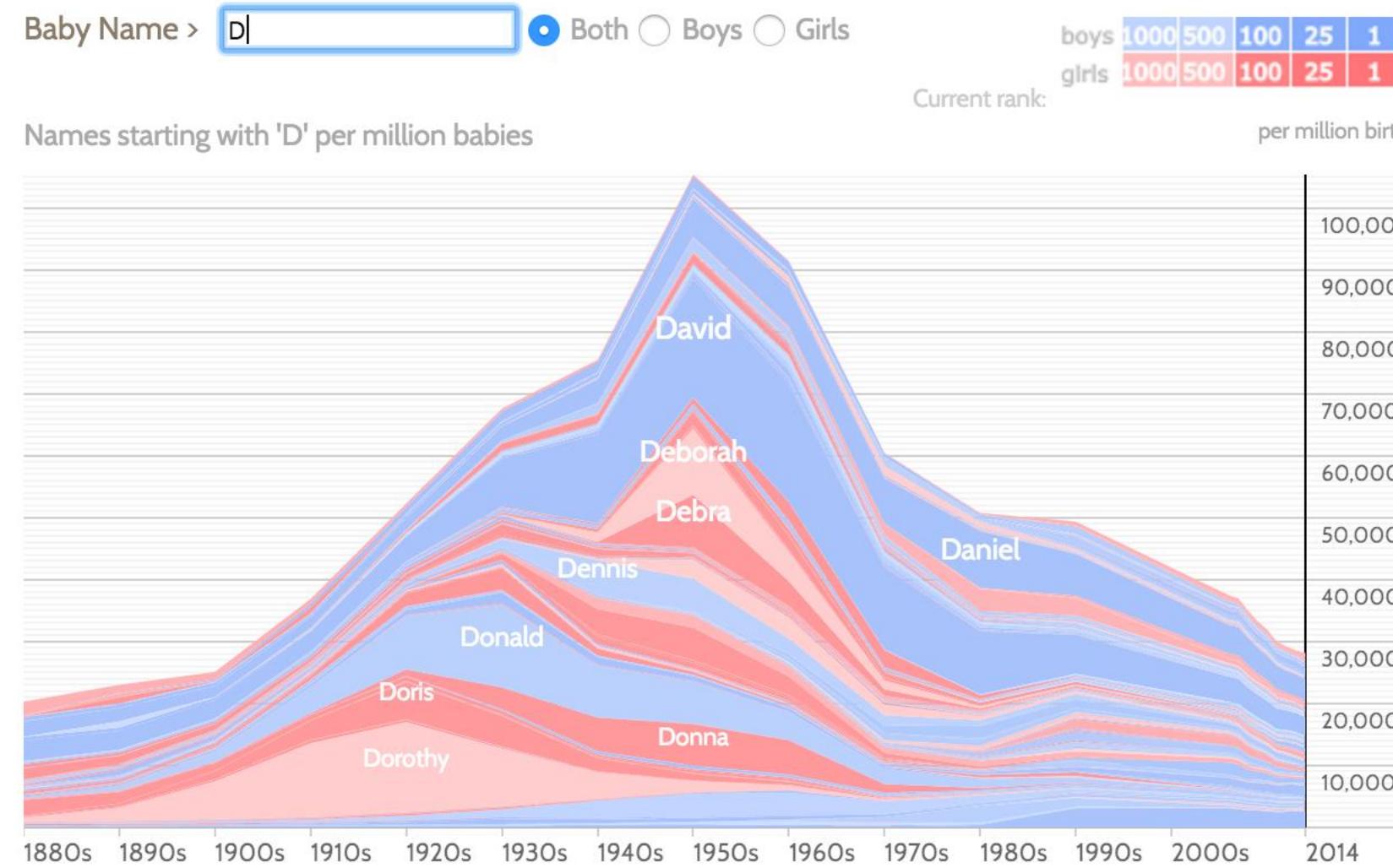
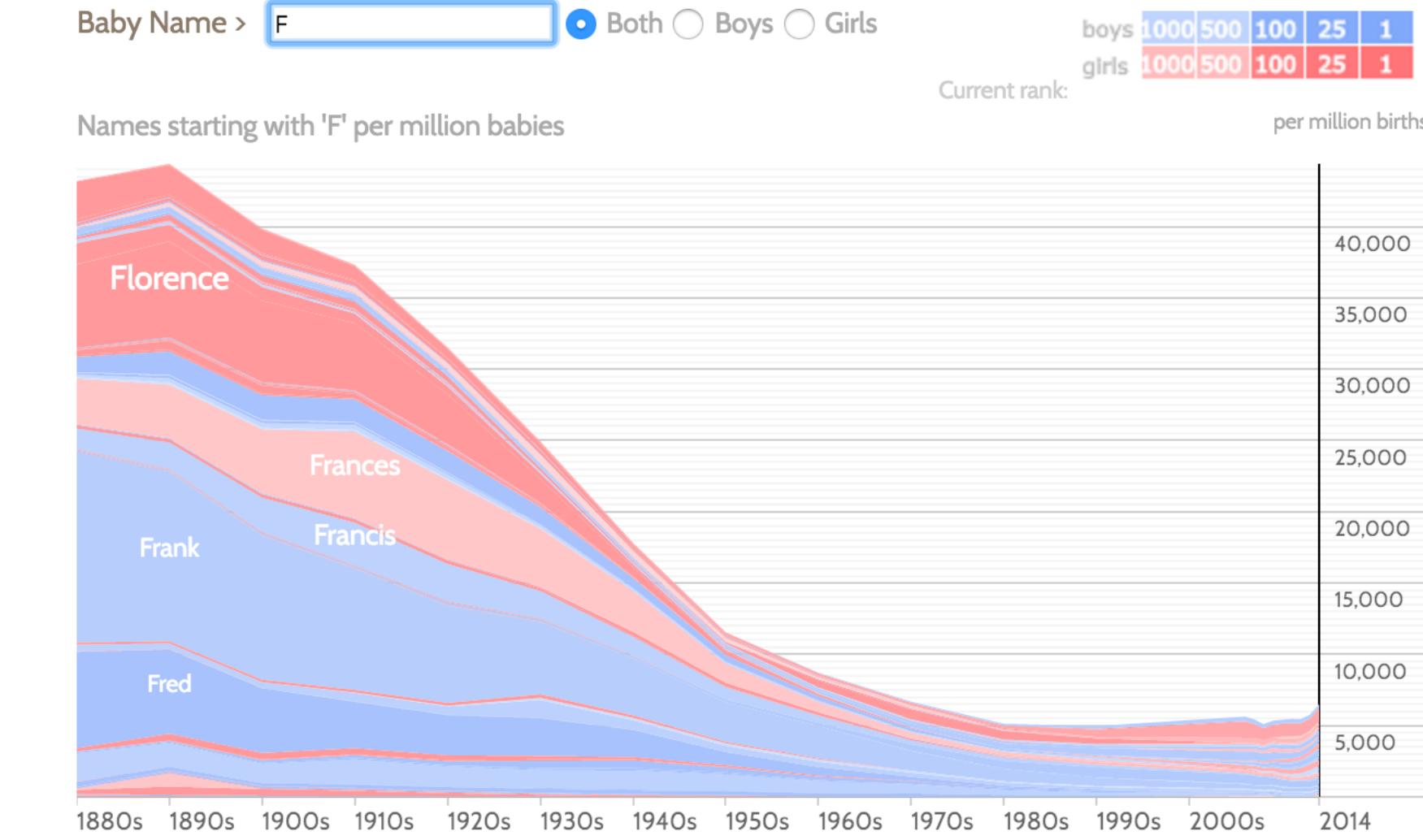
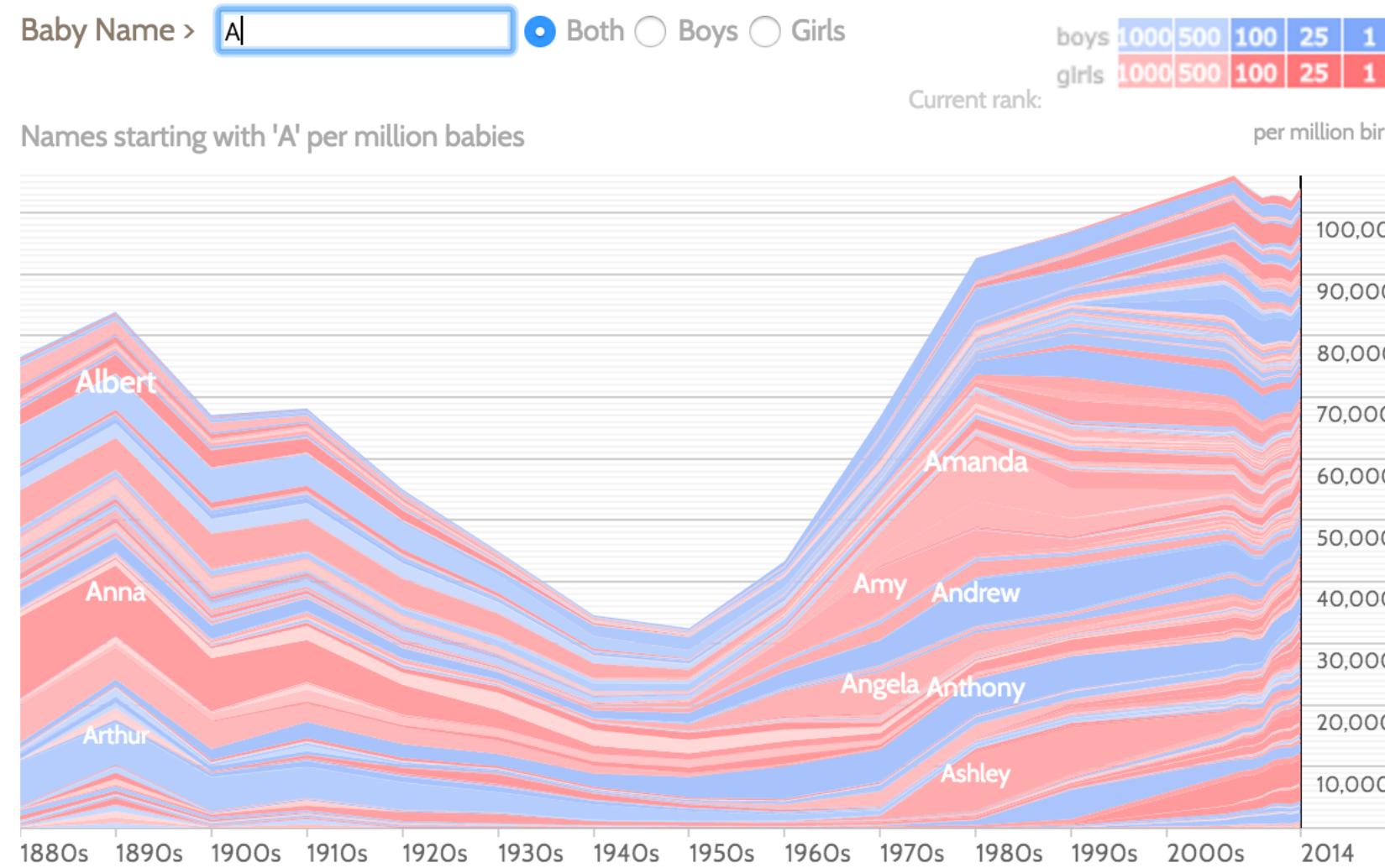
→ *Enjoy*



Munzner, 2014



Example: Name Voyager



www.babynamewizard.com



Example: Name Voyager

Target audience

“This is perfect, as baby names weigh heavily on my mind these days.”

“Useful fodder for historical fiction, too, if you’re looking for typical names for a given age and time period.”

People outside the target audience

“Surprisingly addictive”

“This rules, even though it’s about baby names”

“Cool... by the way, I don’t like babies or children.”

Wattenberg, Martin. "Baby names, visualization, and social data analysis." Information Visualization, 2005. INFOVIS 2005. IEEE Symposium on. IEEE, 2005.



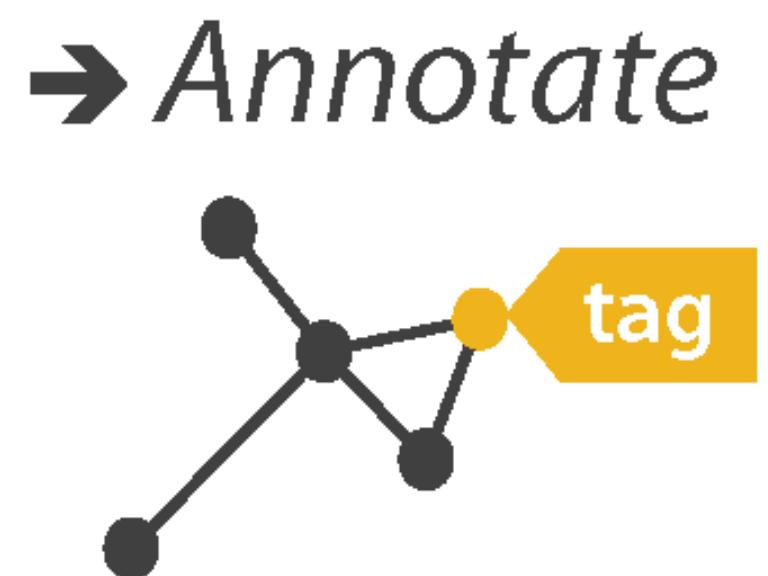
WHY: TASK ABSTRACTION

Analyze → Produce



Produce → Annotate

- Adding extra information (usually manually).
- Annotation – new attribute.



EXAMPLE:

<https://jolecule.appspot.com/pdb/1aoi>



Produce → Record

- Screen shots
- Bookmarked elements
- Parameter settings
- Interaction logs
- Annotations

→ *Record*



Example: Tableau



Worksheet History

Filters:

Graphic Type: All

Bookmarks: Show Only

Add Inventory

Show Me!

Move Market Size to Columns

Add Product to Columns

Open Sheet 1

Add Market

Add Market Size

Add Product

Add Budget Profit

Add Budget Sales

Show Me!

Undo

Redo

Add Inventory

Add Margin

A screenshot of the Tableau software interface. On the left, a 'Worksheet History' panel displays a list of recent actions: 'Open Sheet 1', 'Add Market', 'Add Market Size', 'Add Product', 'Add Budget Profit', 'Add Budget Sales', and 'Show Me!'. Below this, a row of five preview cards shows different data visualizations: a bar chart, a bar chart with a large outlier, a scatter plot, and two other charts. A second row of preview cards shows 'Undo' and 'Redo' operations, followed by 'Add Inventory' and 'Add Margin' cards. The main workspace shows a data grid with columns 'Major Ma..', 'Small Mar..', 'Central', 'East', 'South', and 'West', and values like 1,683,579, 1,563,045, etc. To the right of the grid are three bar charts. The bottom right corner contains a URL: <http://vis.stanford.edu/papers/graphical-histories>.

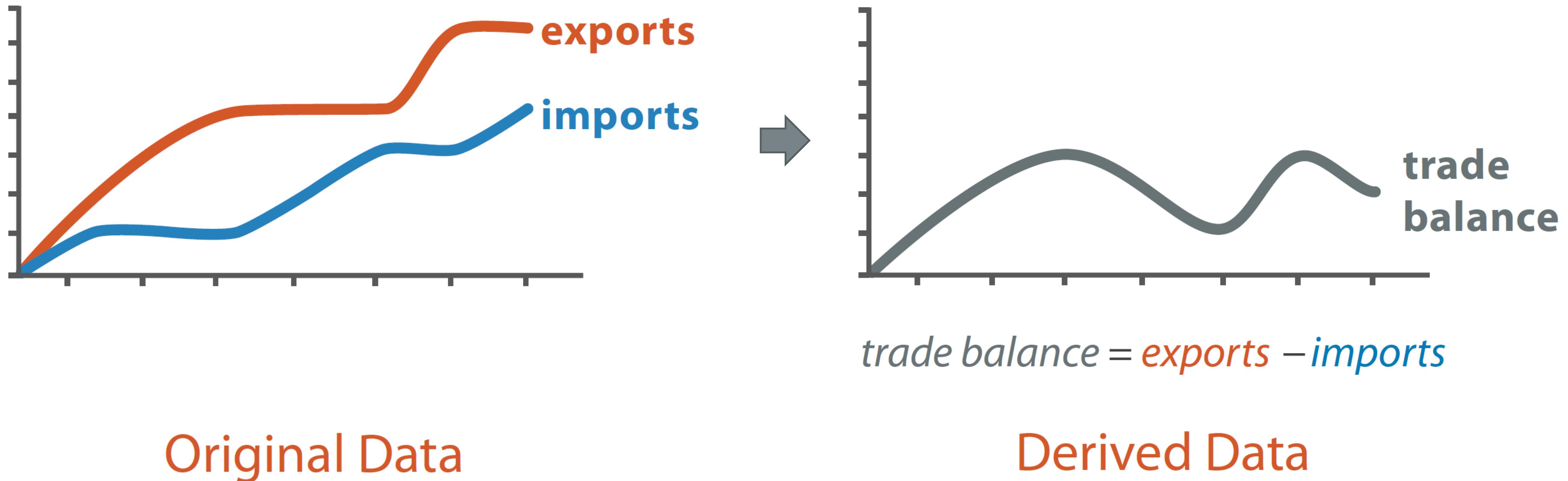
Produce → Derive

- Producing new data elements from existing.
- Dataset "as is" **or** transformed:
 - Derived attributes
 - Dataset type conversion
- Transforming data expands the design space

→ *Derive*



Derived attributes



Derived attributes

Fr, 29.04.
heute

max **14°**
min **-1°**



Vormittag



Nachmittag



Abend/Nacht

Sa, 30.04.
morgen

max **16°**
min **1°**



Vormittag



Nachmittag



Abend/Nacht

01.05.
Sonntag

max **13°**
min **5°**



Vormittag



Nachmittag



Abend/Nacht

02.05.
Montag

max **16°**
min **6°**



Vormittag



Nachmittag

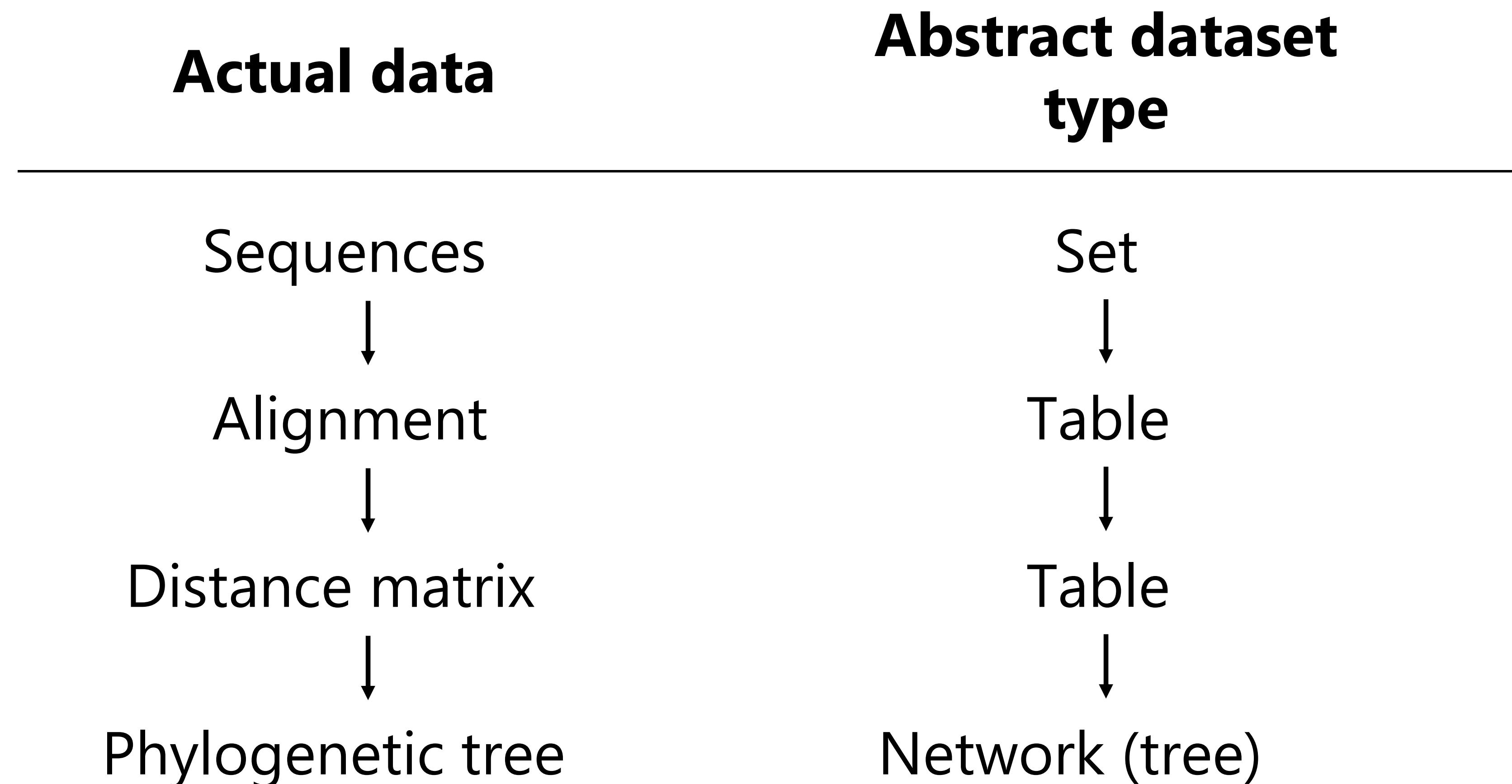


Abend/Nacht



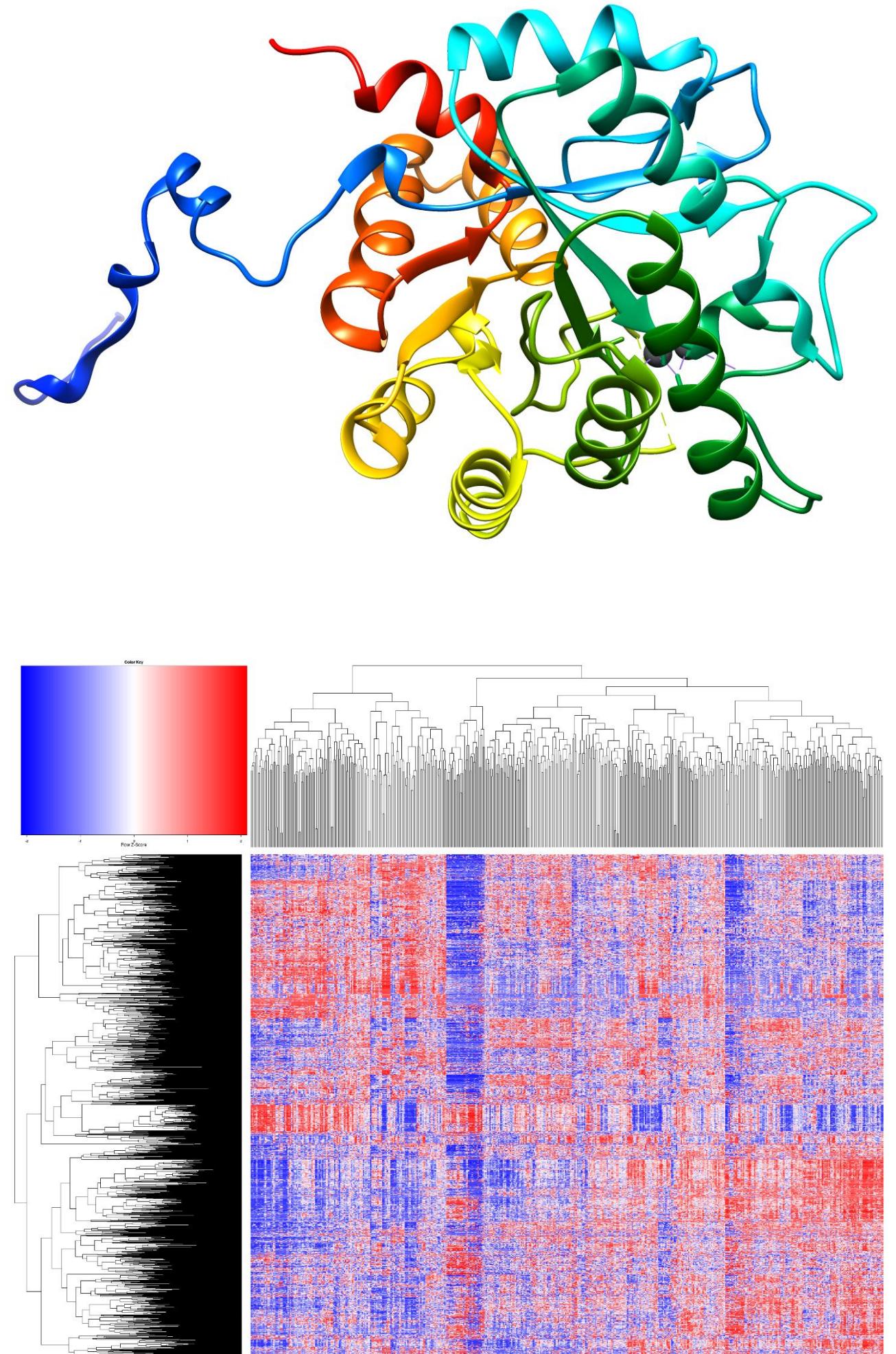
Transformation Example

Task: Show relationships among a set of sequences of a gene X in various organisms.



Examples: Transformations

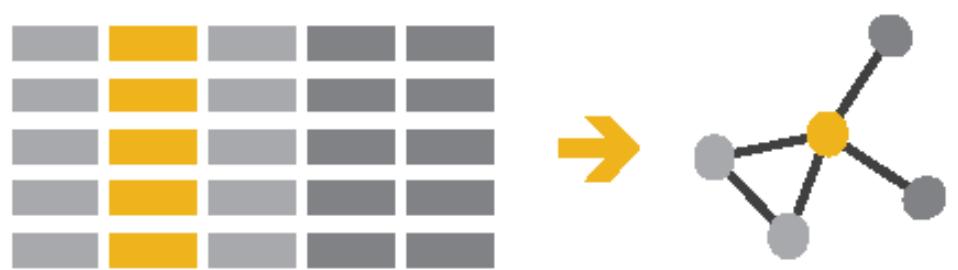
- Given a protein sequence, show its predicted three dimensional structure.
- Given gene expression data, show a heatmap with a clustering of the genes by their expression patterns.



Produce → Derive

1. Decide what needs to be shown
2. Create a series of transformations
3. Draw

→ *Derive*



Actions

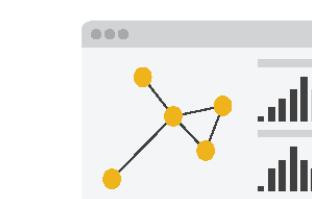
→ Analyze

→ Consume

→ Discover



→ Present

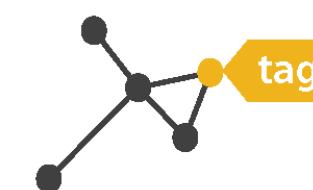


→ Enjoy



→ Produce

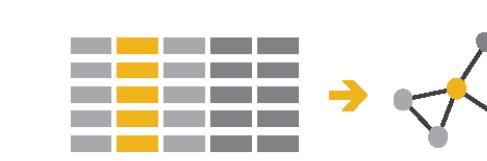
→ Annotate



→ Record



→ Derive

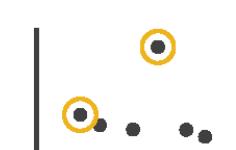


→ Search

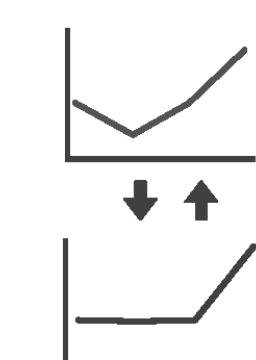
	Target known	Target unknown
Location known	•••• <i>Lookup</i>	•••• <i>Browse</i>
Location unknown	◁•••▷ <i>Locate</i>	◁•••▷ <i>Explore</i>

→ Query

→ Identify



→ Compare



→ Summarize



 Search

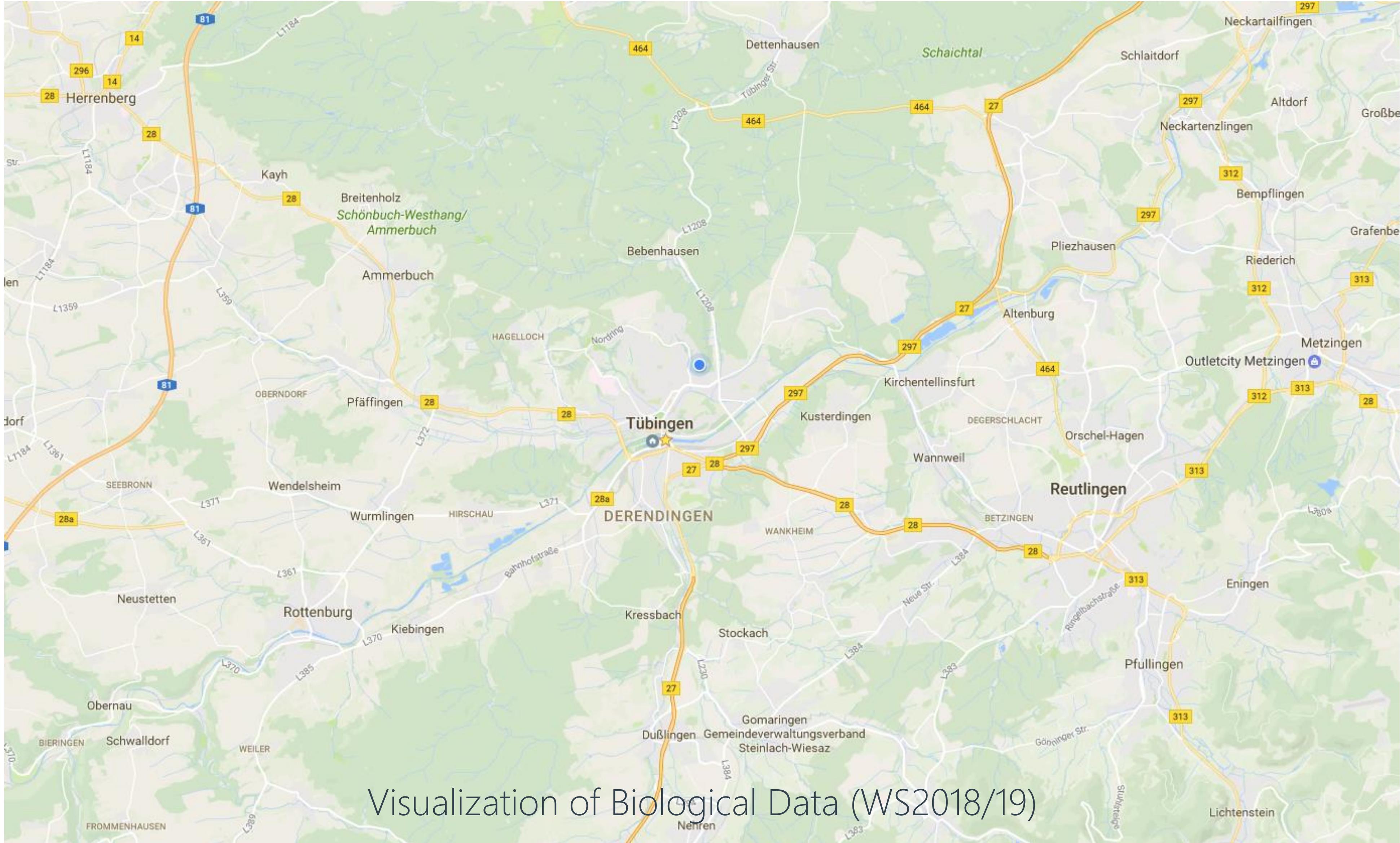
	Target known	Target unknown
Location known	 <i>Lookup</i>	 <i>Browse</i>
Location unknown	 <i>Locate</i>	 <i>Explore</i>

Search

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

Lookup: Tübingen

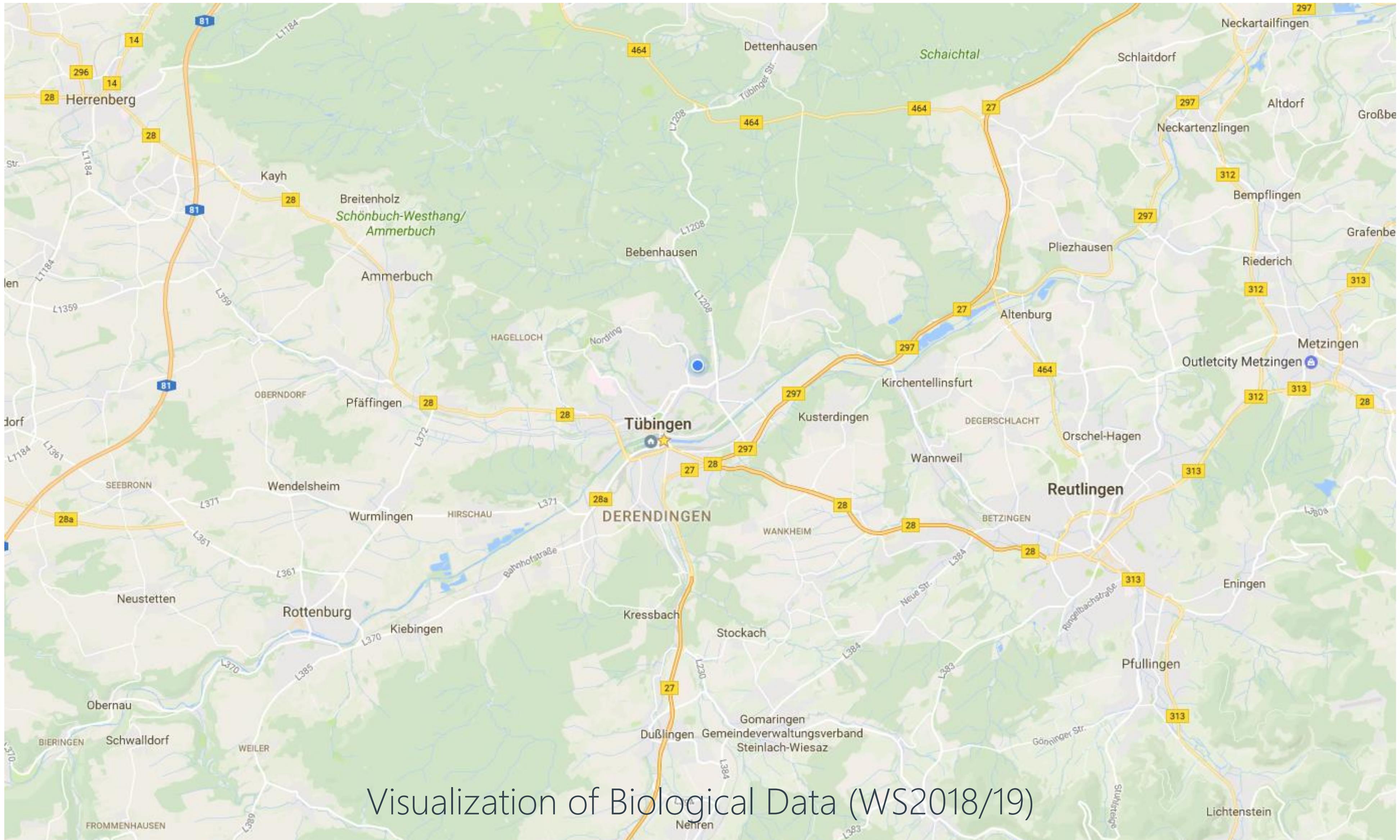
• • • Lookup



Visualization of Biological Data (WS2018/19)

Locate part of town: Kayh

◀ ⚡ ▶ Locate



Visualization of Biological Data (WS2018/19)

Browse for the largest Lake

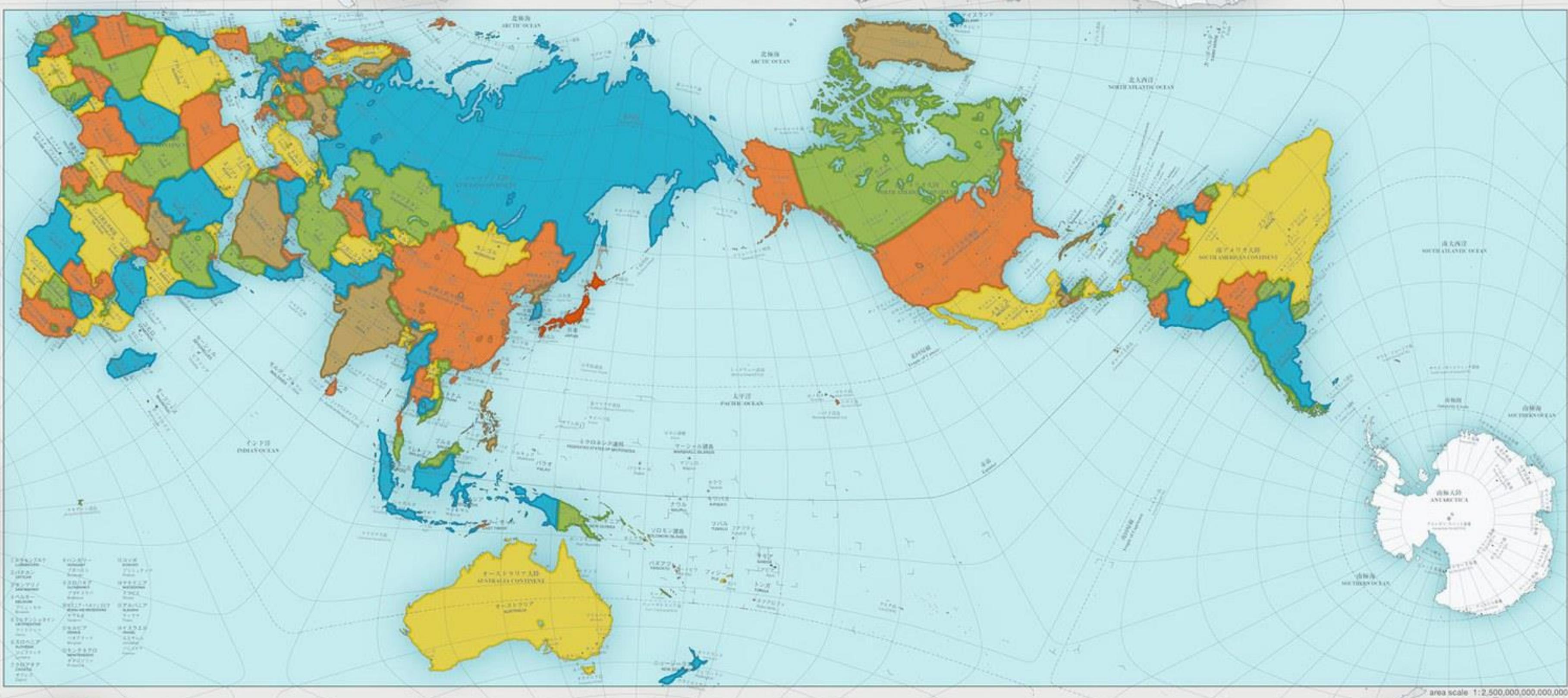


Browse



Explore

AuthaGraph World Map



Actions

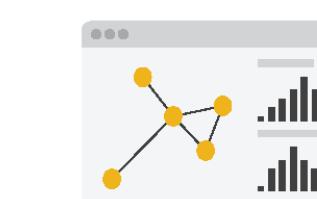
→ Analyze

→ Consume

→ Discover



→ Present

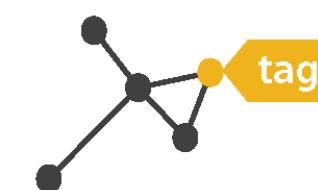


→ Enjoy



→ Produce

→ Annotate



→ Record



→ Derive

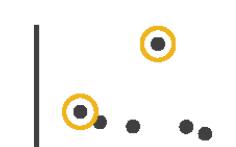


→ Search

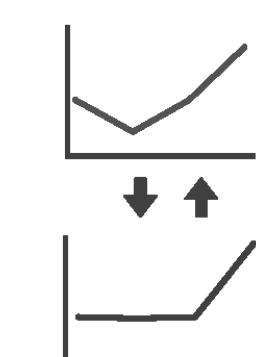
	Target known	Target unknown
Location known	•••• <i>Lookup</i>	•••• <i>Browse</i>
Location unknown	◁•••▷ <i>Locate</i>	◁•••▷ <i>Explore</i>

→ Query

→ Identify



→ Compare

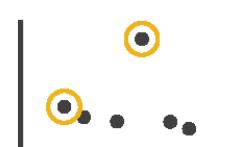


→ Summarize

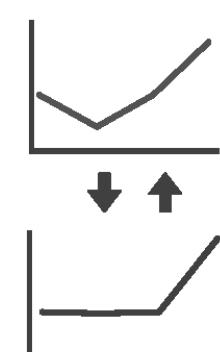


④ Query

→ Identify



→ Compare

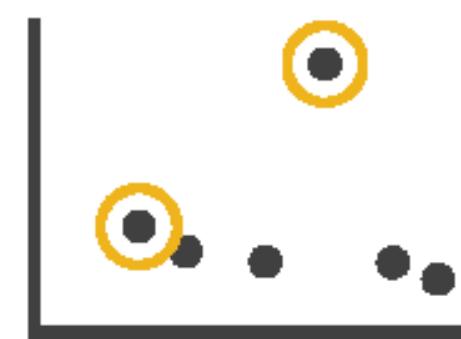


→ Summarize



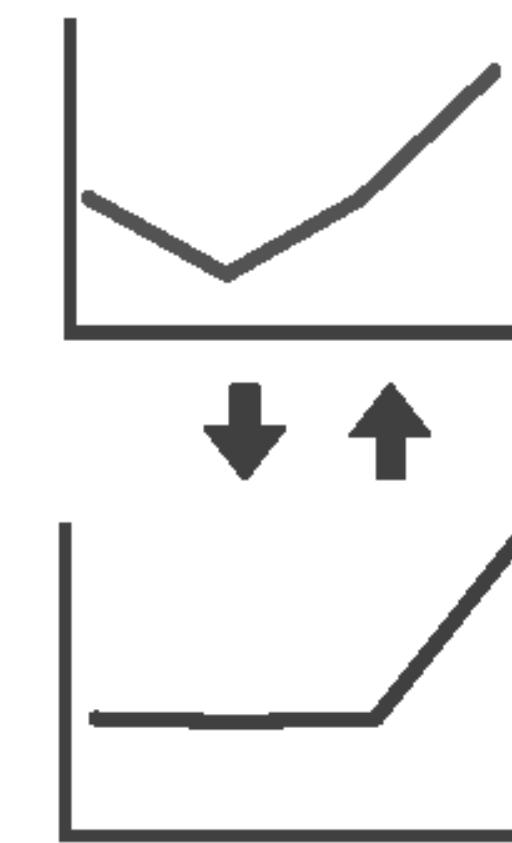
Query

→ Identify



→ Single target

→ Compare



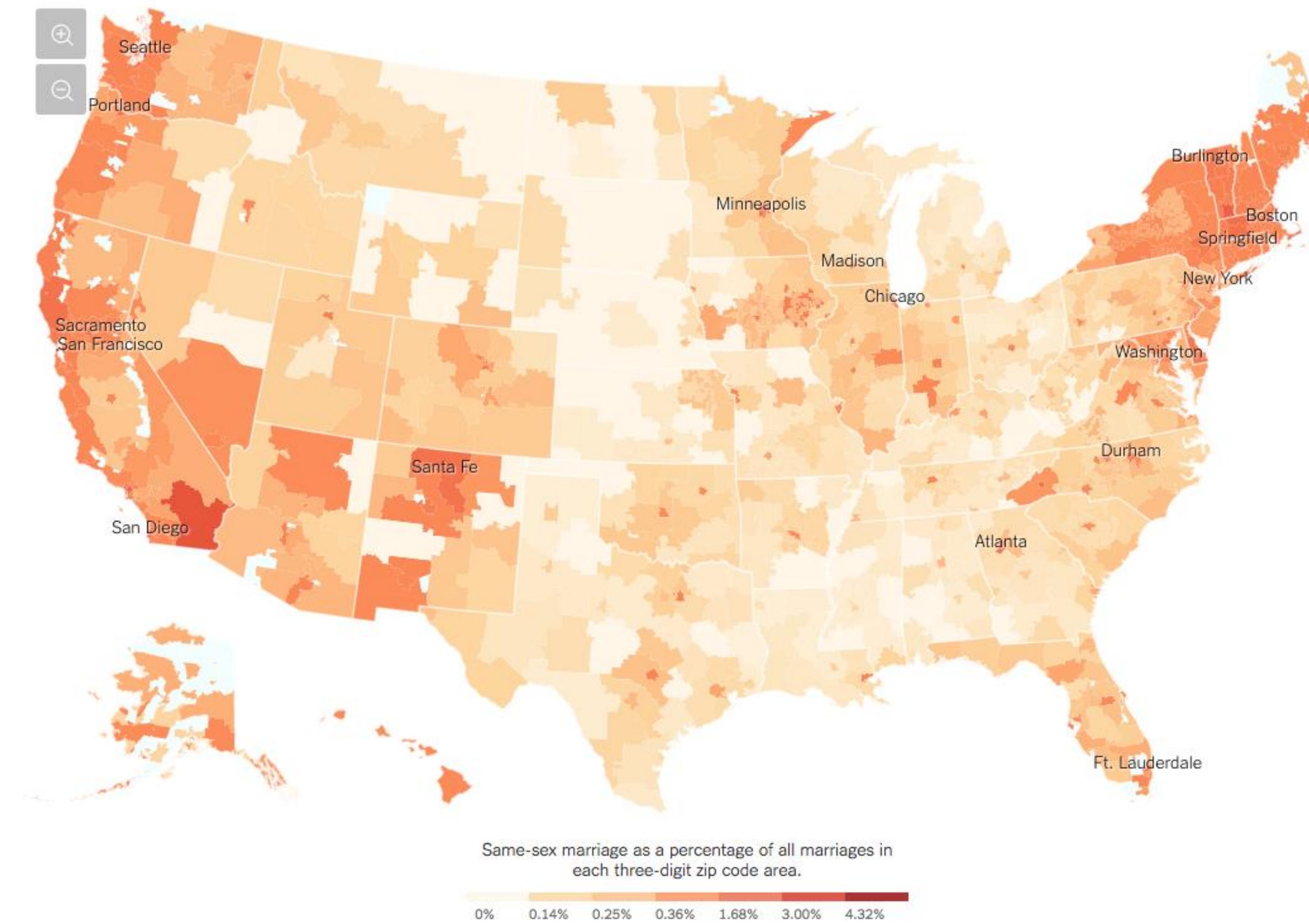
→ Multiple targets

→ Summarize

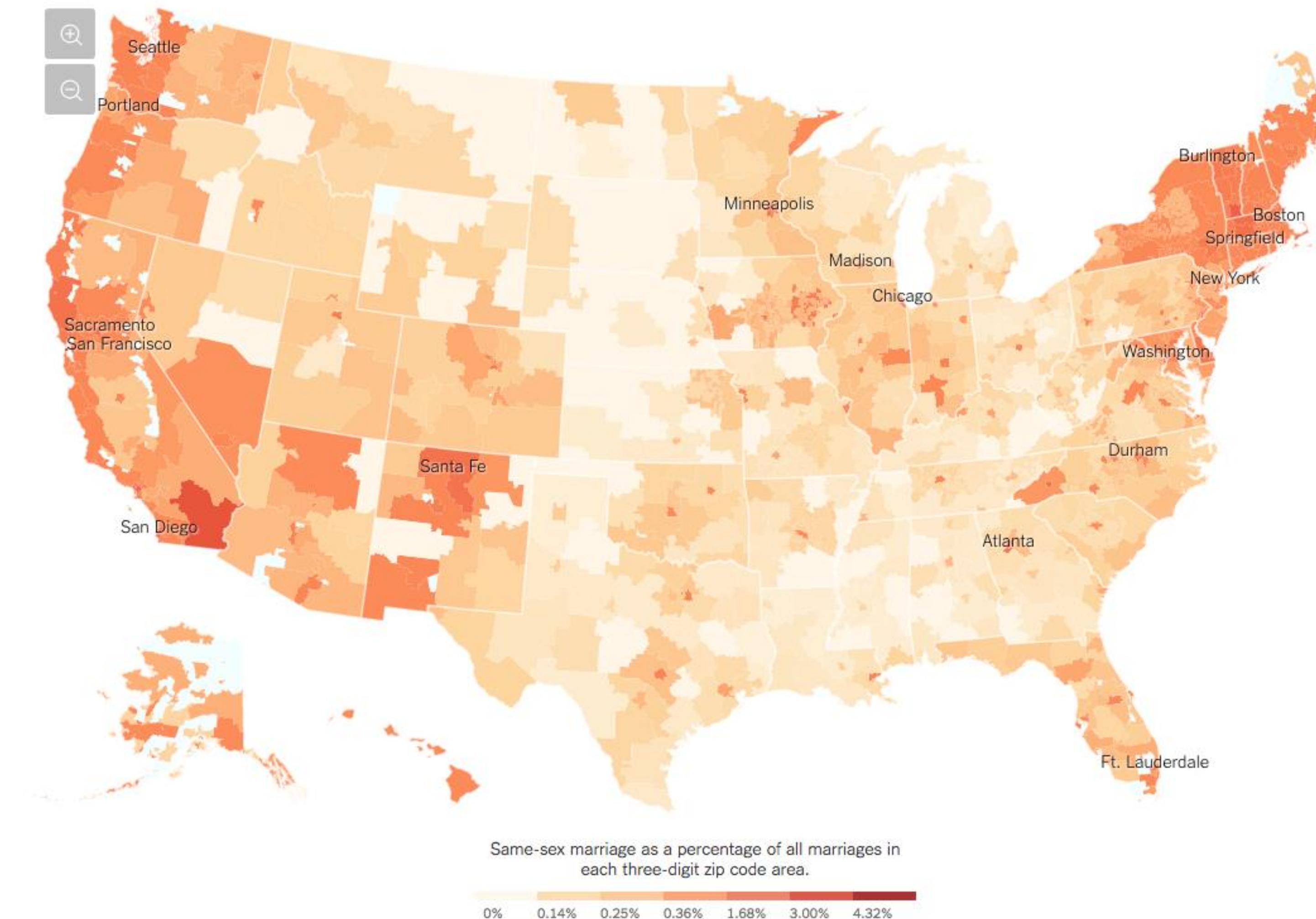


→ All available targets

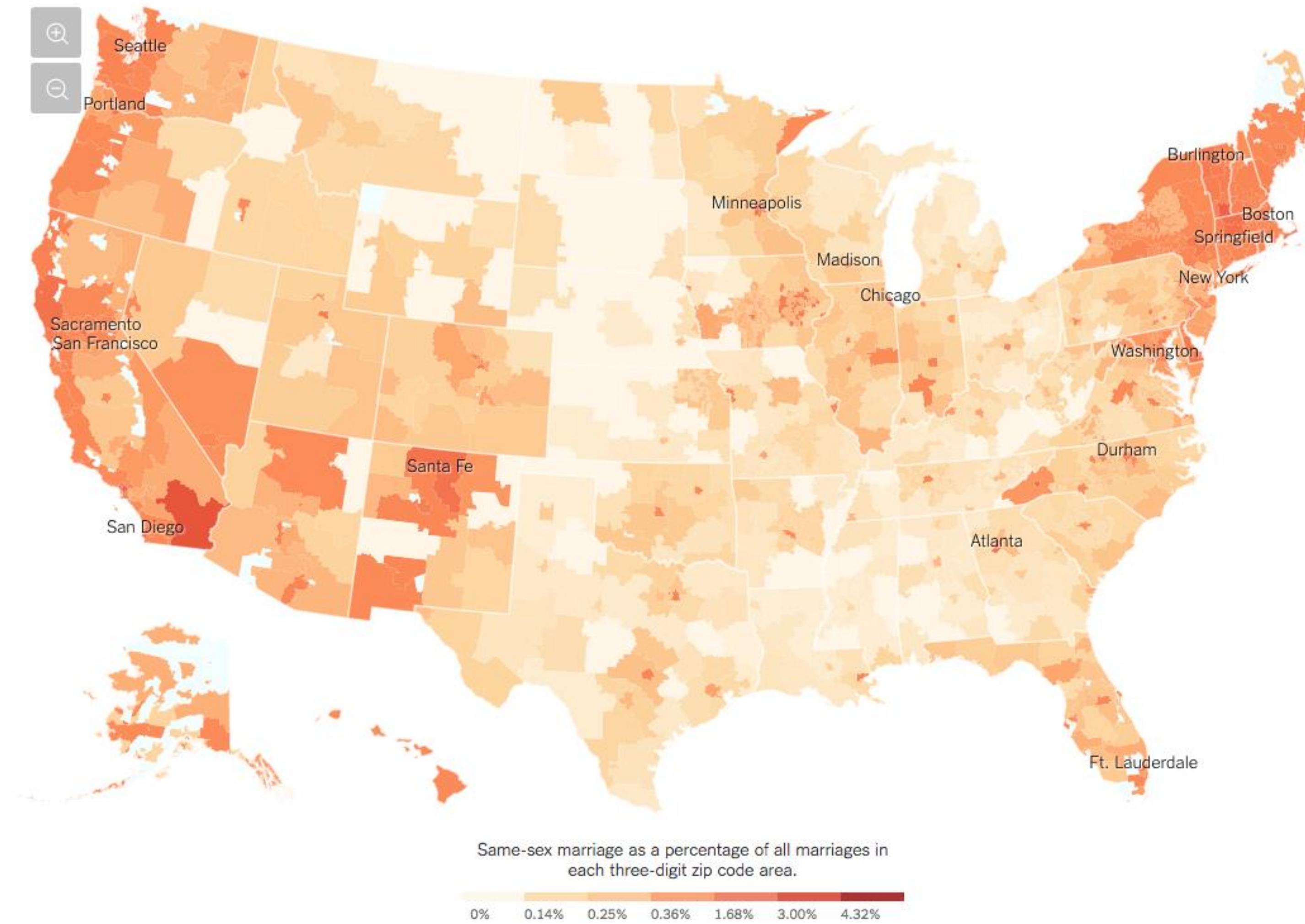
Identify Same-Sex Marriage Fraction in San Diego



Compare Same-Sex Marriage Fraction in San Diego to Texas



Summarise Same-Sex Marriage Fraction in the USA



What?

Why?

How?

Why?

Actions

Targets

→ Analyze

→ Consume

→ Discover



→ Present



→ Enjoy



→ Produce

→ Annotate



→ Record



→ Derive

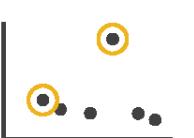


→ Search

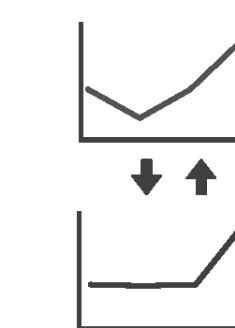
	Target known	Target unknown
Location known	••• <i>Lookup</i>	••• <i>Browse</i>
Location unknown	◁•••▷ <i>Locate</i>	◁•••▷ <i>Explore</i>

→ Query

→ Identify



→ Compare



→ Summarize

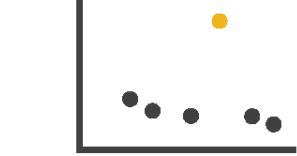


→ All Data

→ Trends



→ Outliers



→ Features



→ Attributes

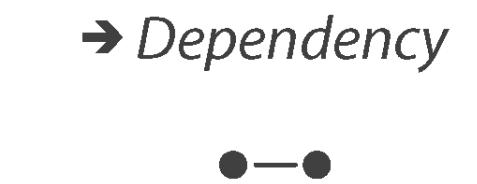
→ One



→ Extremes



→ Many



→ Dependency



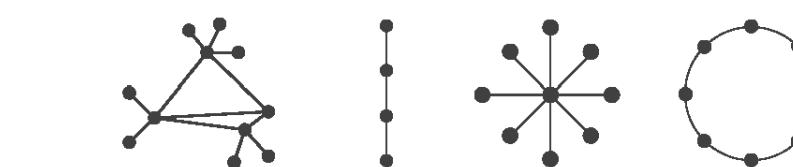
→ Correlation



→ Similarity

→ Network Data

→ Topology



→ Paths



→ Spatial Data

→ Shape



Munzner, 2014



What?

Why?

How?

Why?

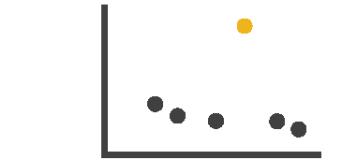
Targets

→ All Data

→ Trends



→ Outliers



→ Features



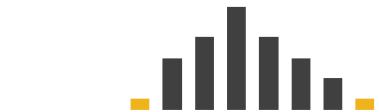
→ Attributes

→ One

→ Distribution



→ Extremes

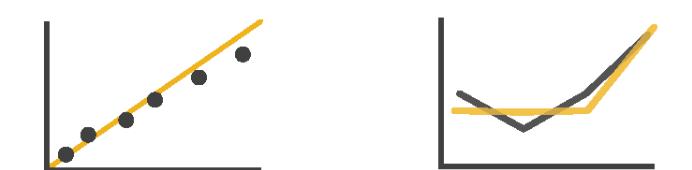
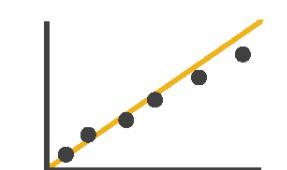


→ Many

→ Dependency



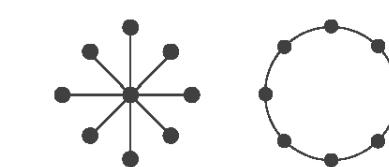
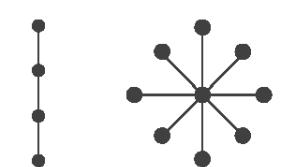
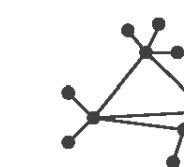
→ Correlation



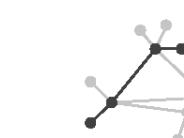
→ Similarity

→ Network Data

→ Topology

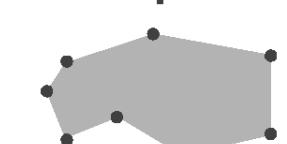


→ Paths



→ Spatial Data

→ Shape



Targets → All Data

→ Trends



Increases,
decreases,
peaks,
plateaus,
etc.

→ Outliers



Anomalies,
novelties,
deviants,
surprises,
etc.

→ Features



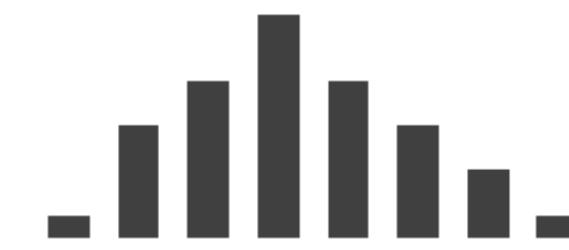
Task dependent



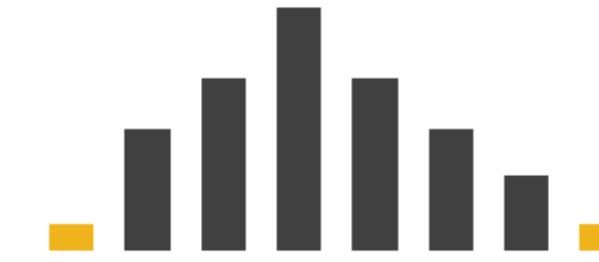
Targets → Attributes

→ One

→ *Distribution*



→ *Extremes*

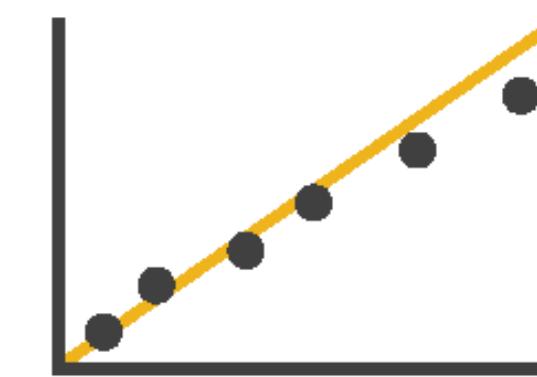


→ Many

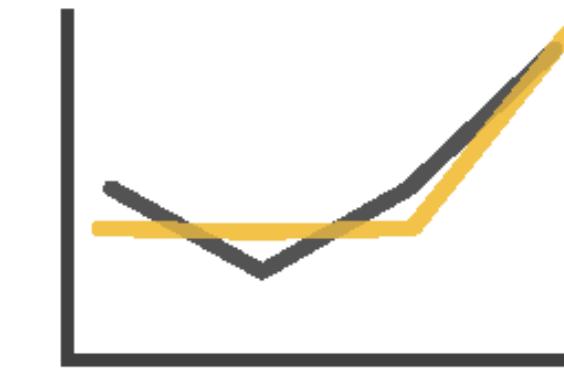
→ *Dependency*



→ *Correlation*



→ *Similarity*



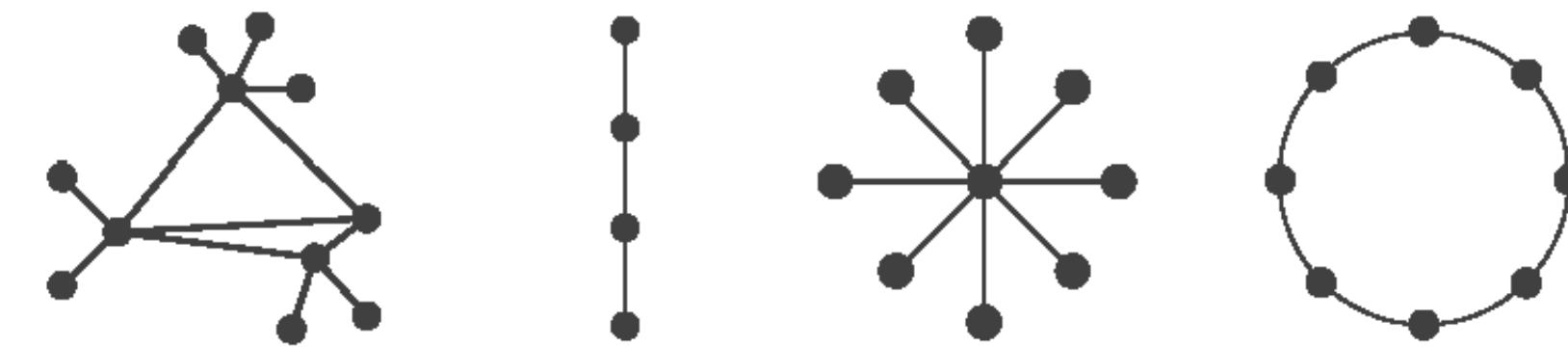
→ Very common tasks!

Munzner, 2014

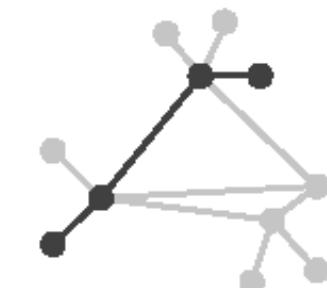
Targets → Other

→ Network Data

→ Topology

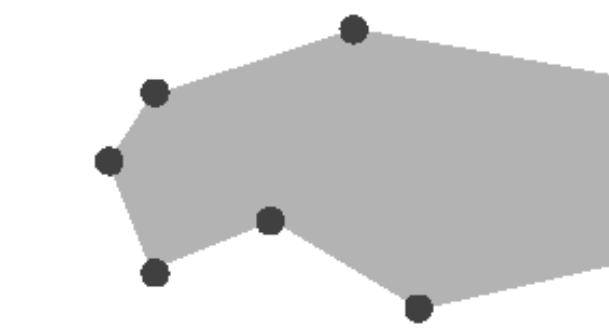


→ Paths



→ Spatial Data

→ Shape

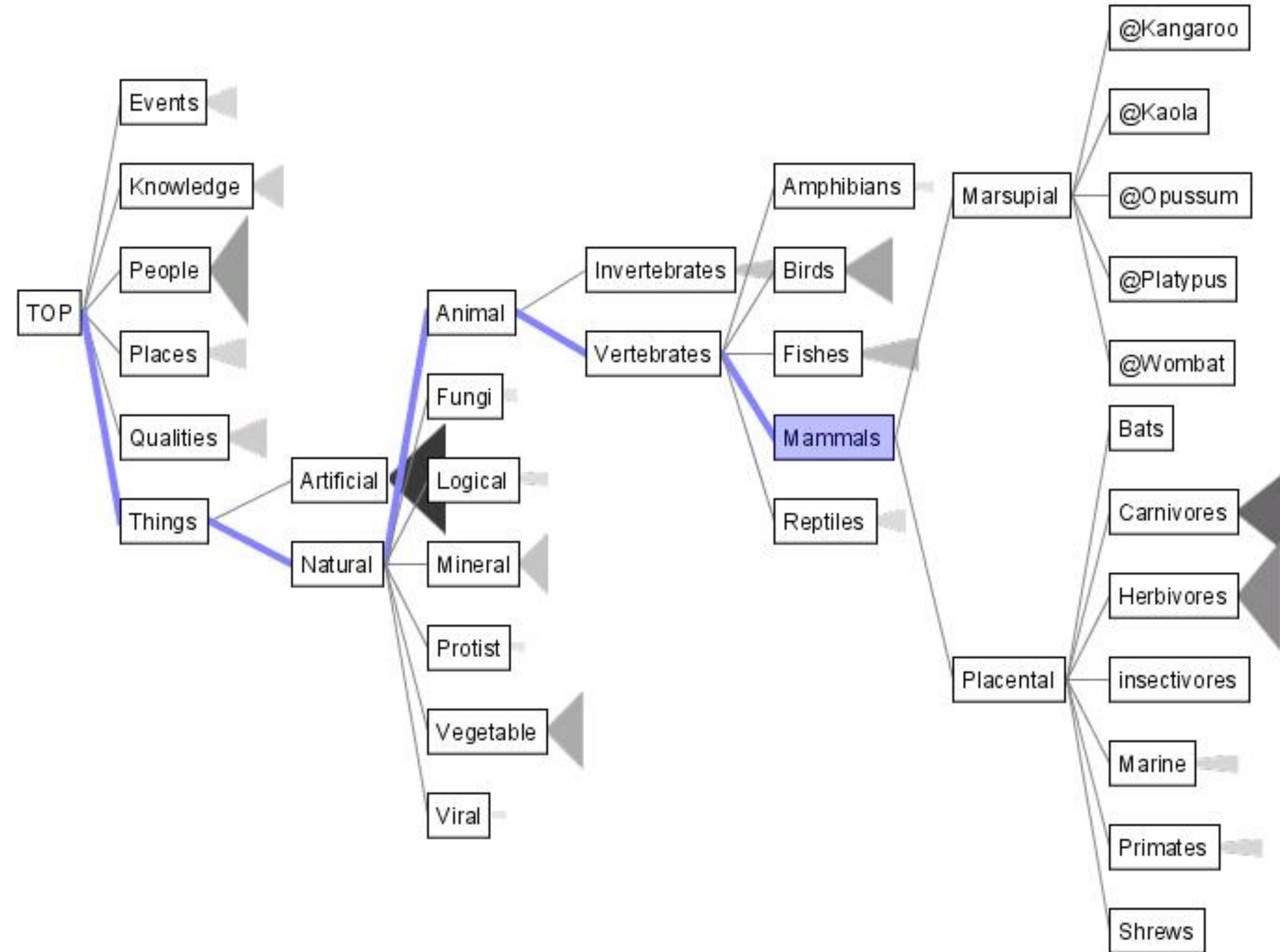


WHY: TASK ABSTRACTION

Case Study: SpaceTree



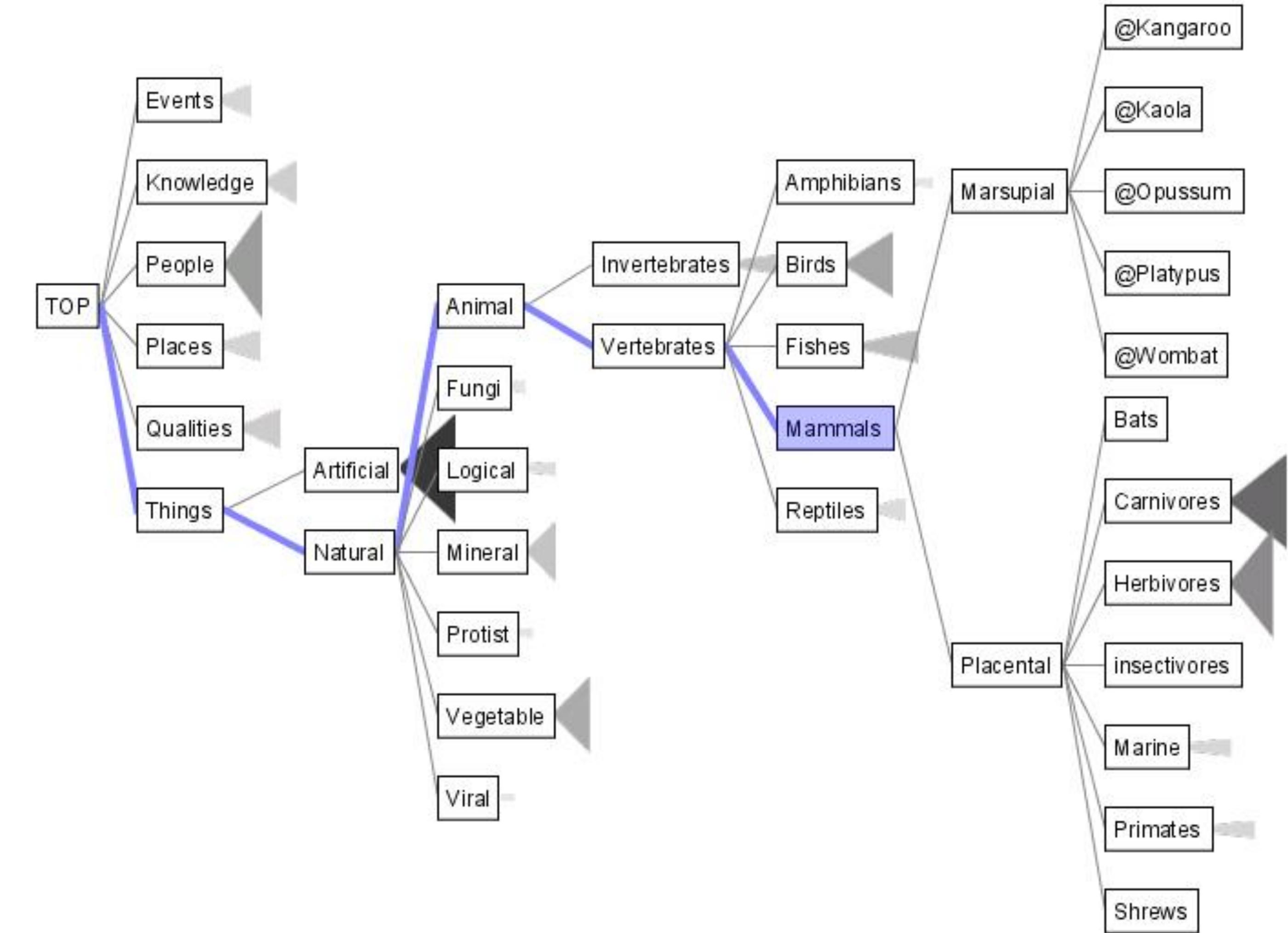
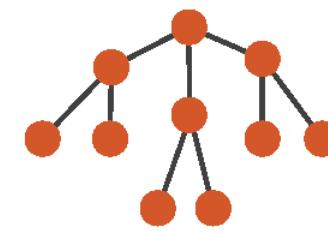
SpaceTree



SpaceTree

What?

→ Tree

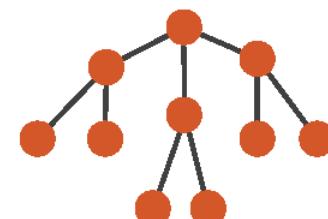


SpaceTree

What?

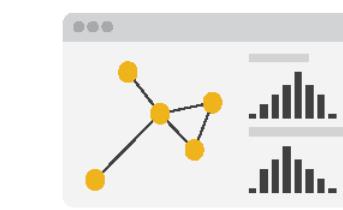
Why?

→ Tree



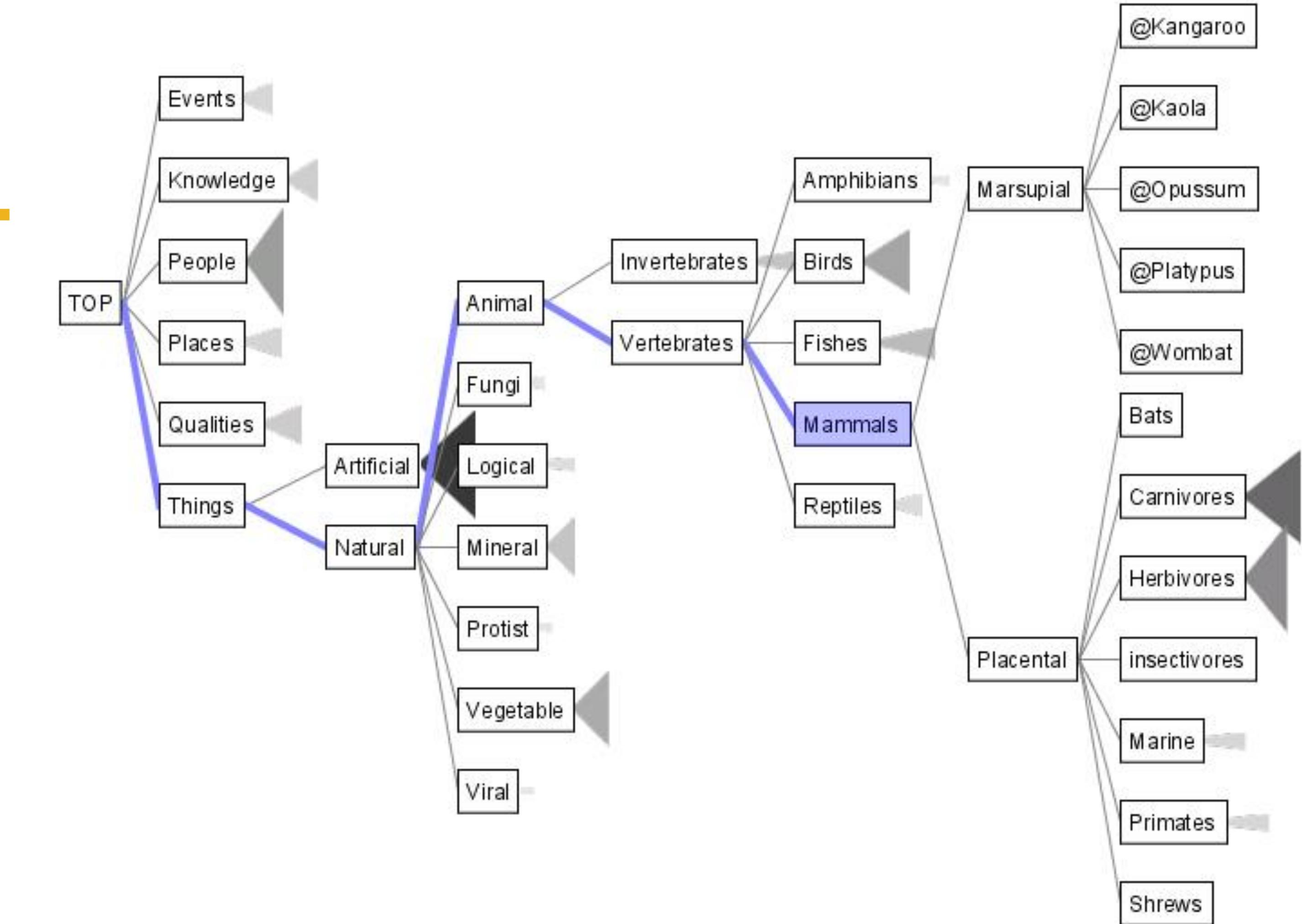
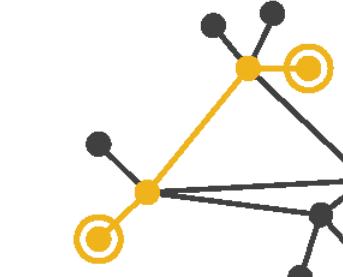
→ Actions

→ Present → Locate → Identify



→ Targets

→ Path between two nodes

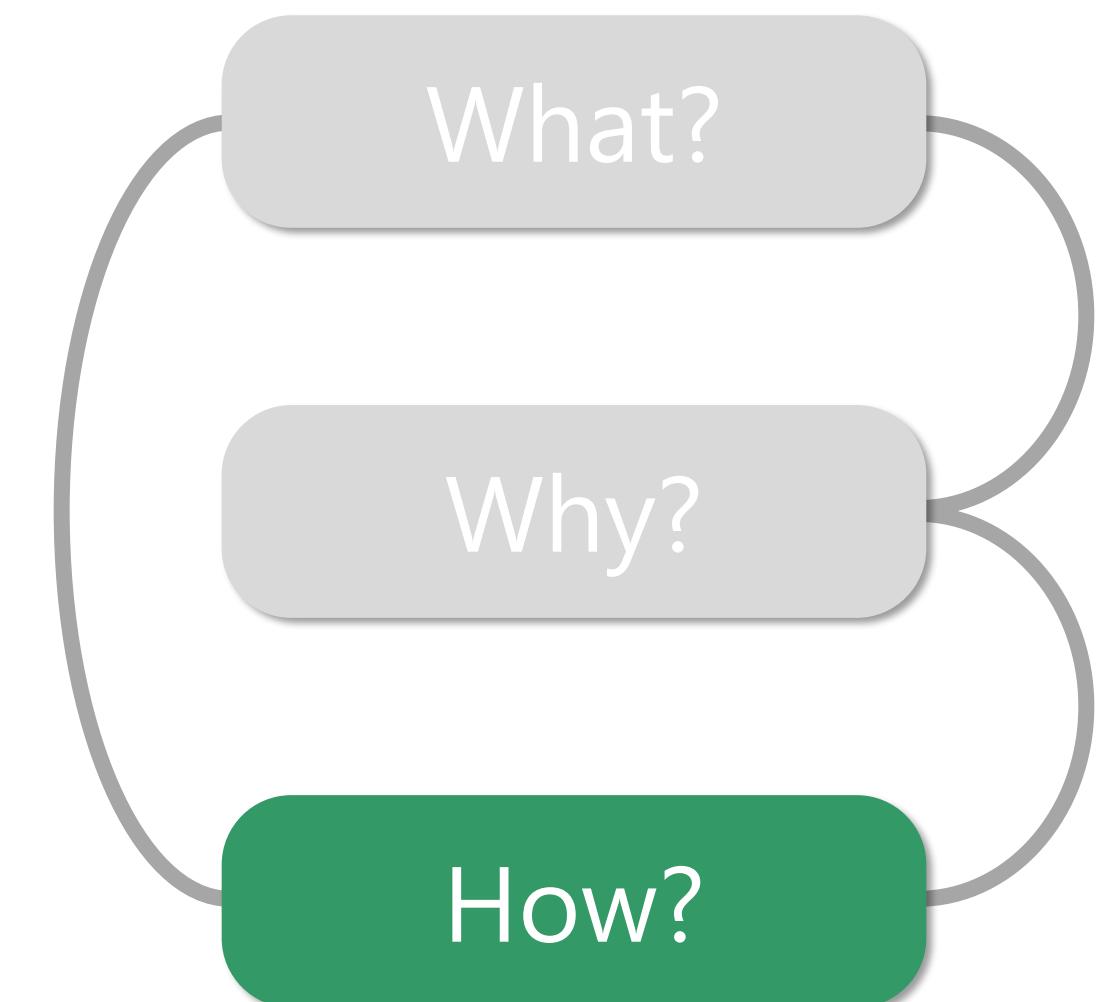




Visualization of Biological Data – Winter Term 2018/2019

How: Data Visualization Idioms

Jun.-Prof. Dr. Michael Krone
29.10.2018



What?

Why?

How?

How?

Encode

Manipulate

Facet

Reduce

④ Arrange

→ Express



→ Separate



→ Order



→ Use



④ Map

from categorical and ordered attributes

→ Color

→ Hue



→ Saturation



→ Luminance



→ Size, Angle, Curvature, ...

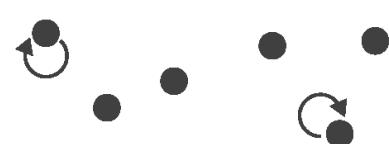


→ Shape



→ Motion

Direction, Rate, Frequency, ...



④ Change



④ Select



④ Navigate



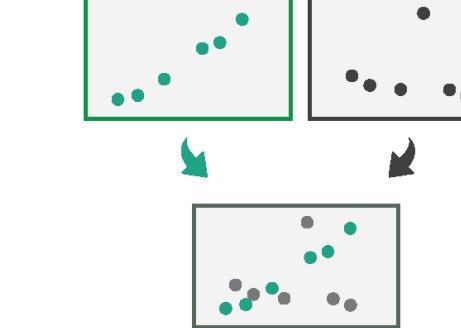
④ Partition



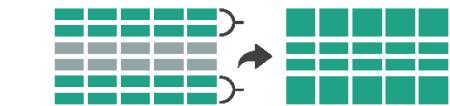
④ Filter



④ Superimpose



④ Aggregate



Encode

→ Arrange

→ Express



→ Separate



→ Order



→ Align



→ Use



→ Map

from categorical and ordered attributes

→ Color

→ Hue



→ Saturation



→ Luminance



→ Size, Angle, Curvature, ...

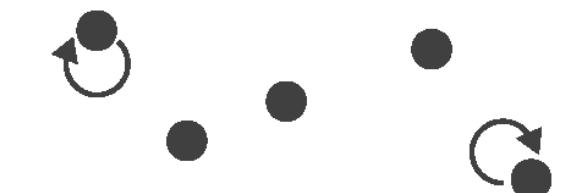


→ Shape



→ Motion

Direction, Rate, Frequency, ...

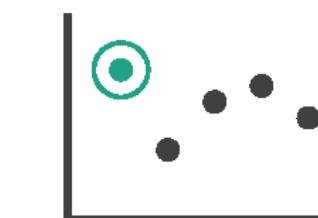


Manipulate

→ Change



→ Select



→ Navigate

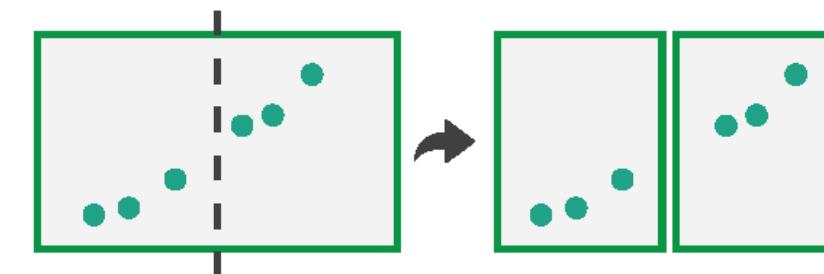


Facet

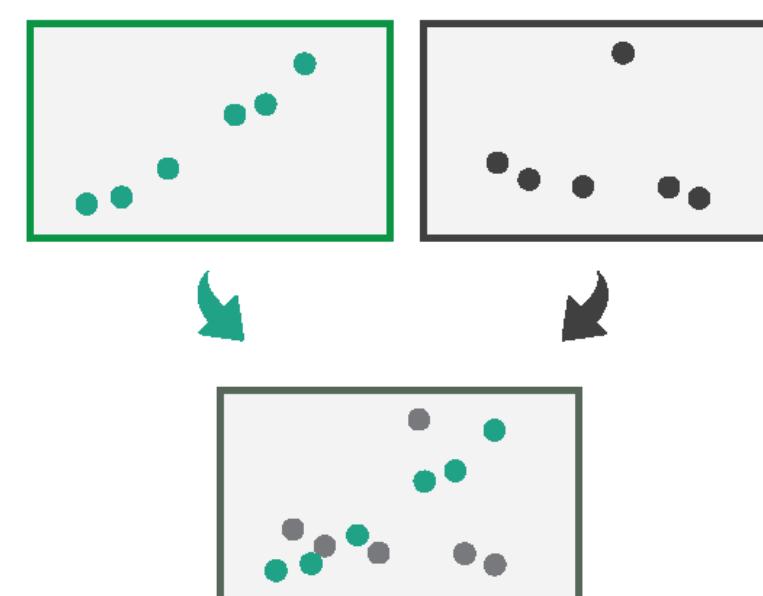
→ Juxtapose



→ Partition



→ Superimpose



Reduce

④ Filter



④ Aggregate



④ Embed

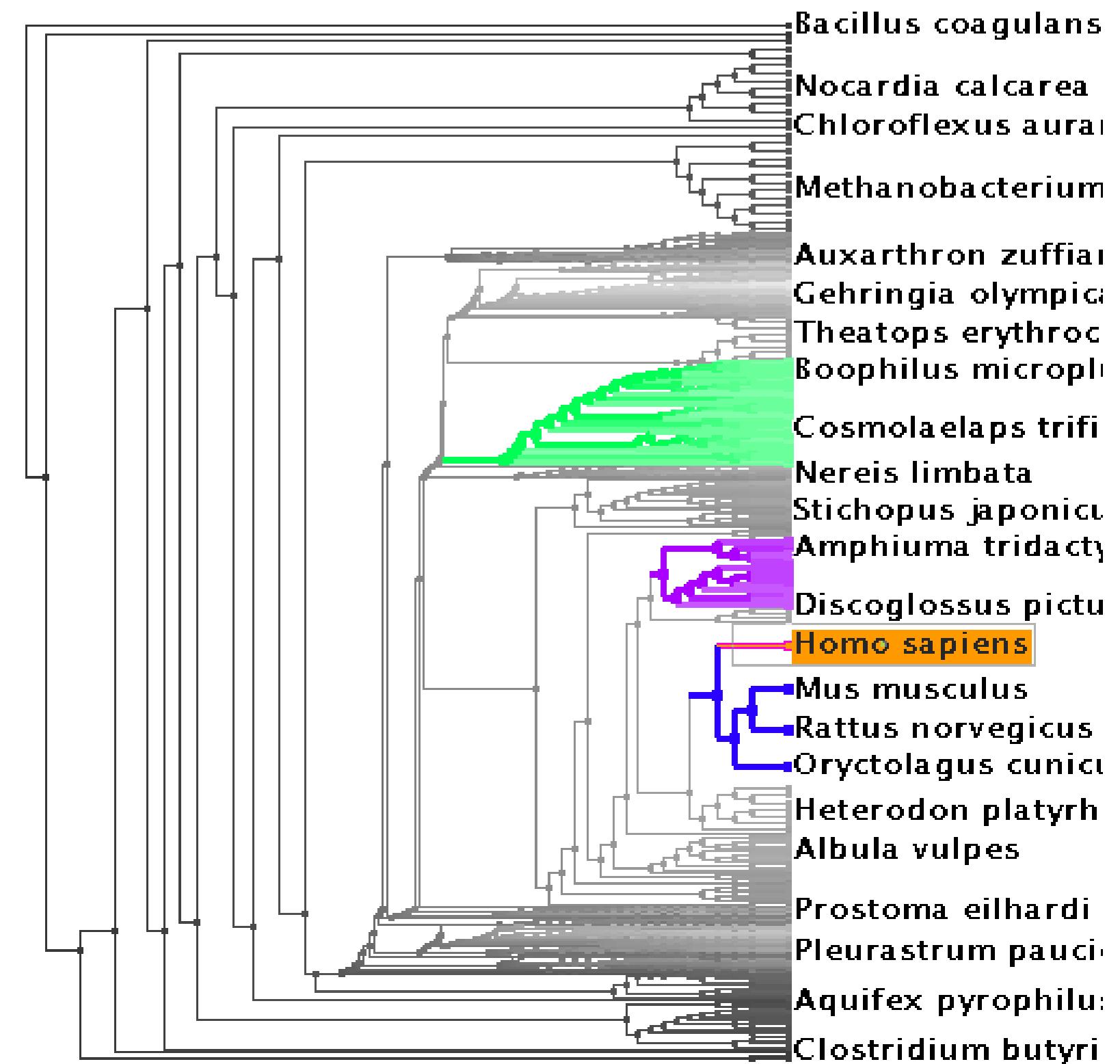
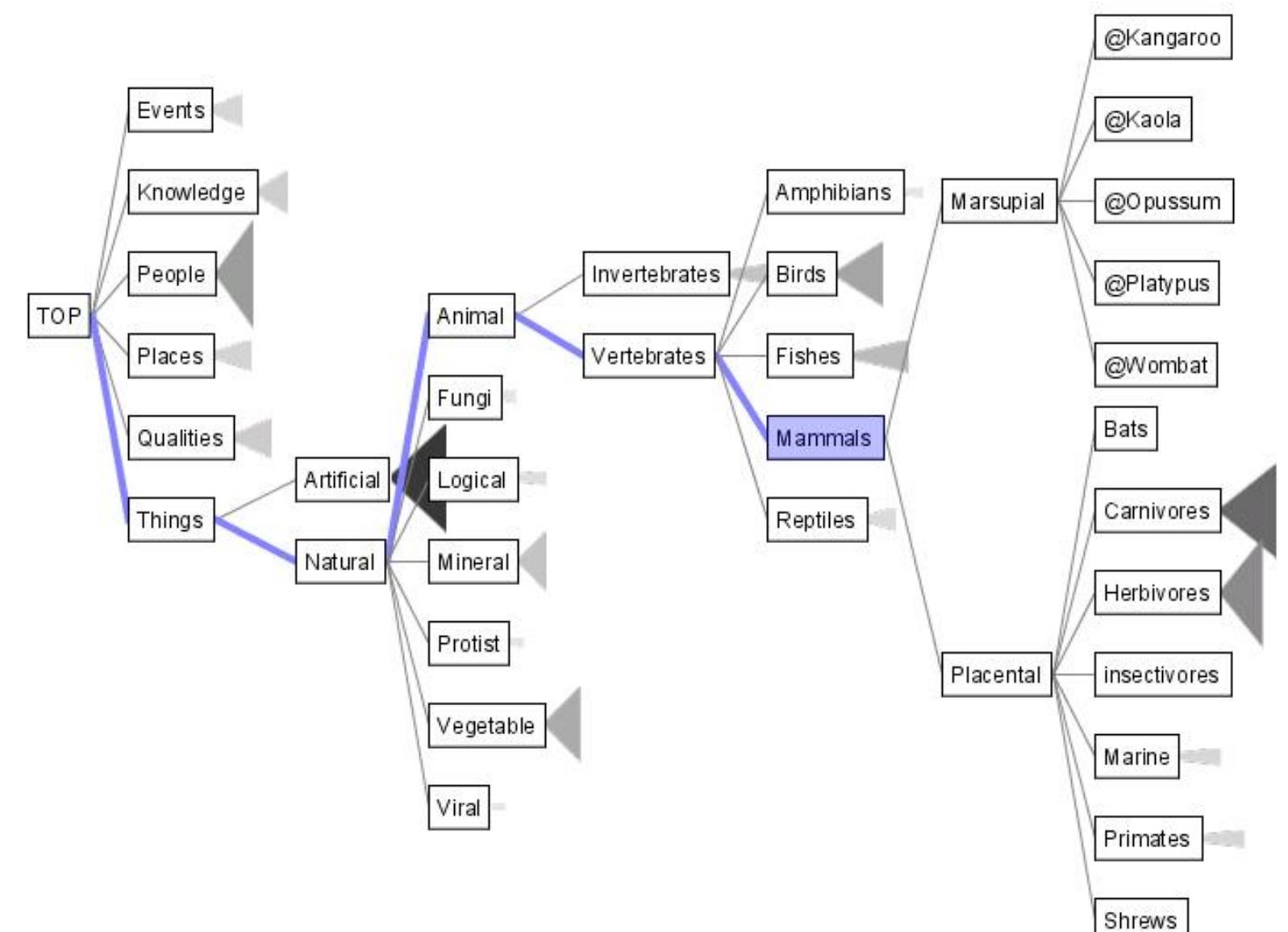


How: DATA VISUALIZATION IDIOMS

Case Study: SpaceTree vs. TreeJuxtaposer



SpaceTree vs. TreeJuxtaposer

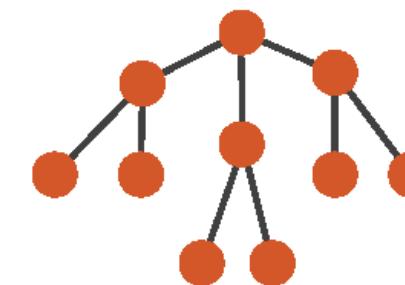


Munzner, T, et. al. Proc. SIGGRAPH 2003, published as ACM Transactions on Graphics 22(3), pages 453--462

SpaceTree vs. TreeJuxtaposer

What?

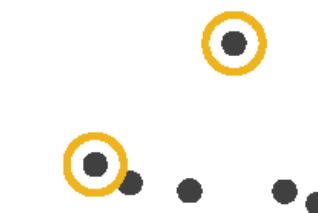
→ Tree



Why?

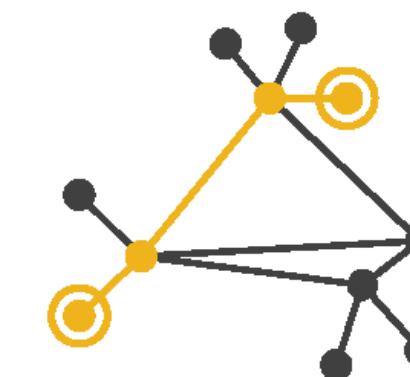
→ Actions

→ Present → Locate → Identify

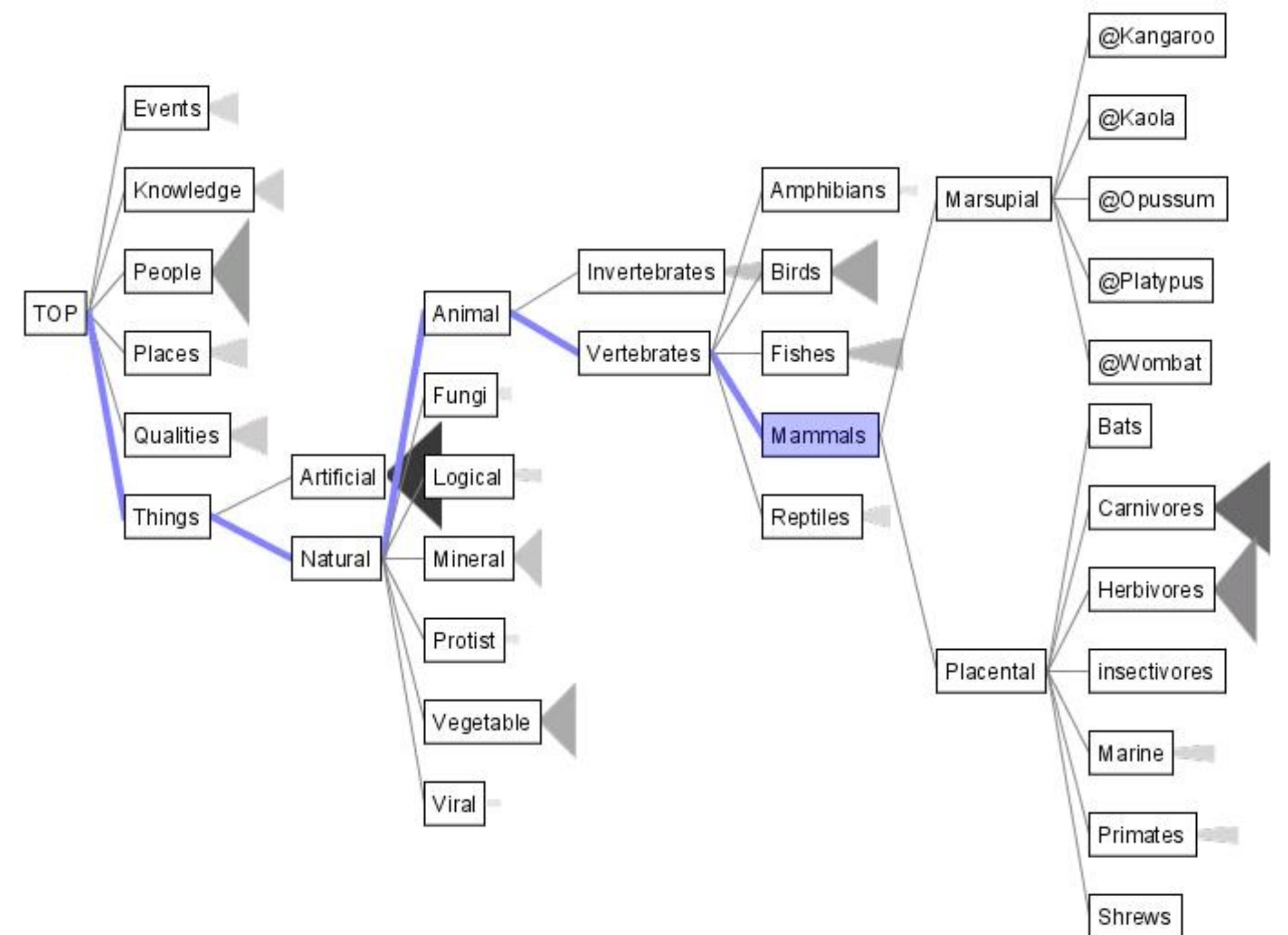


→ Targets

→ Path between two nodes

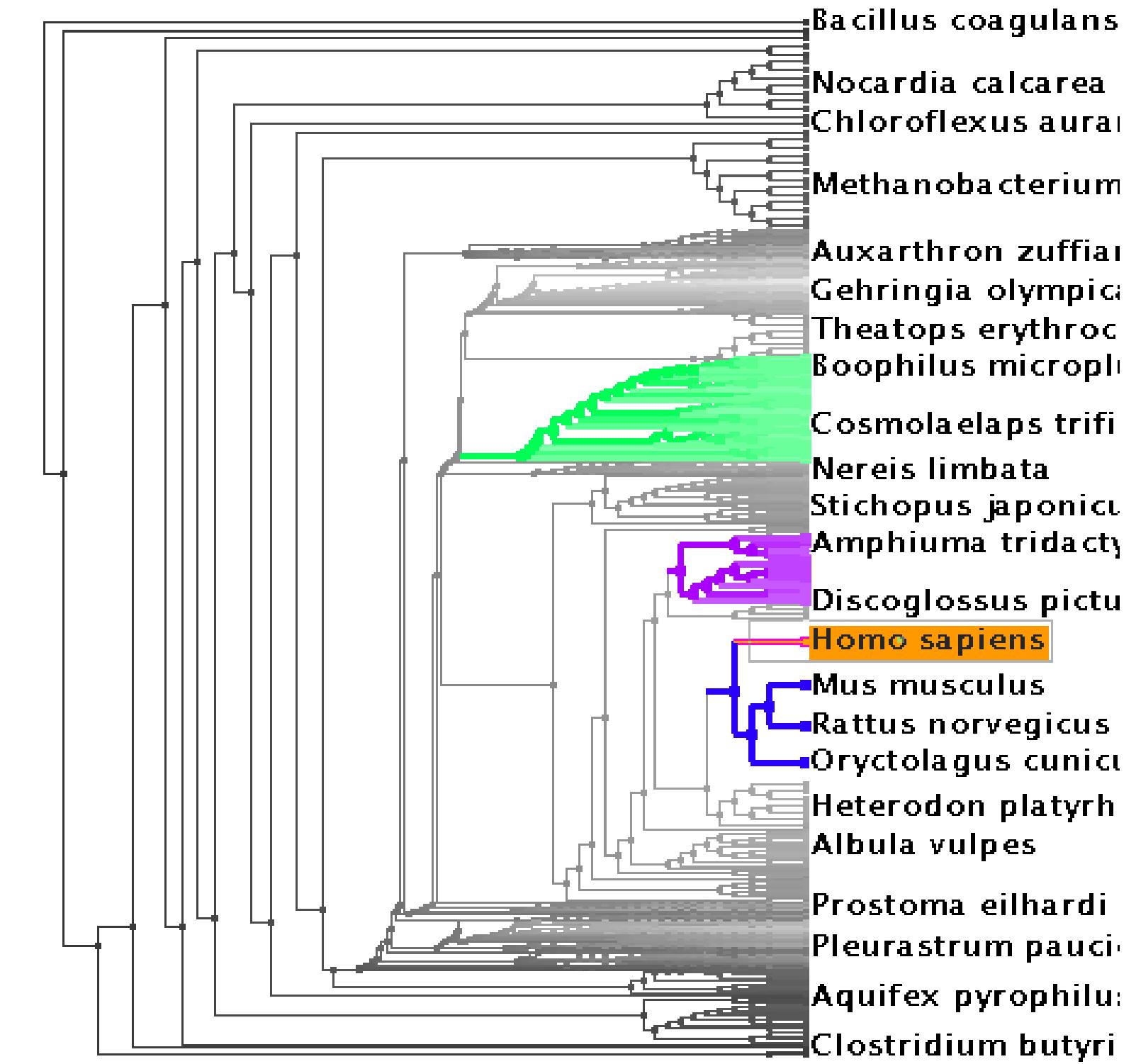
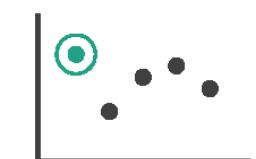
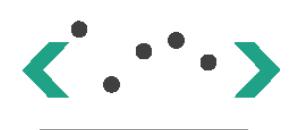


SpaceTree vs. TreeJuxtaposer



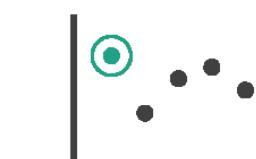
SpaceTree

→ Encode → Navigate → Select → Filter → Aggregate



TreeJuxtaposer

→ Encode → Navigate → Select → Arrange



NEXT LECTURE:

Marks and Channels

→ Map
from categorical and ordered attributes

→ Color
→ Hue → Saturation → Luminance



→ Size, Angle, Curvature, ...



→ Shape
+ ● ■ □ ▲

→ Motion
Direction, Rate, Frequency, ...

