

Task Abstraction

Florian Windhager

Reading

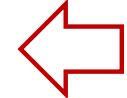
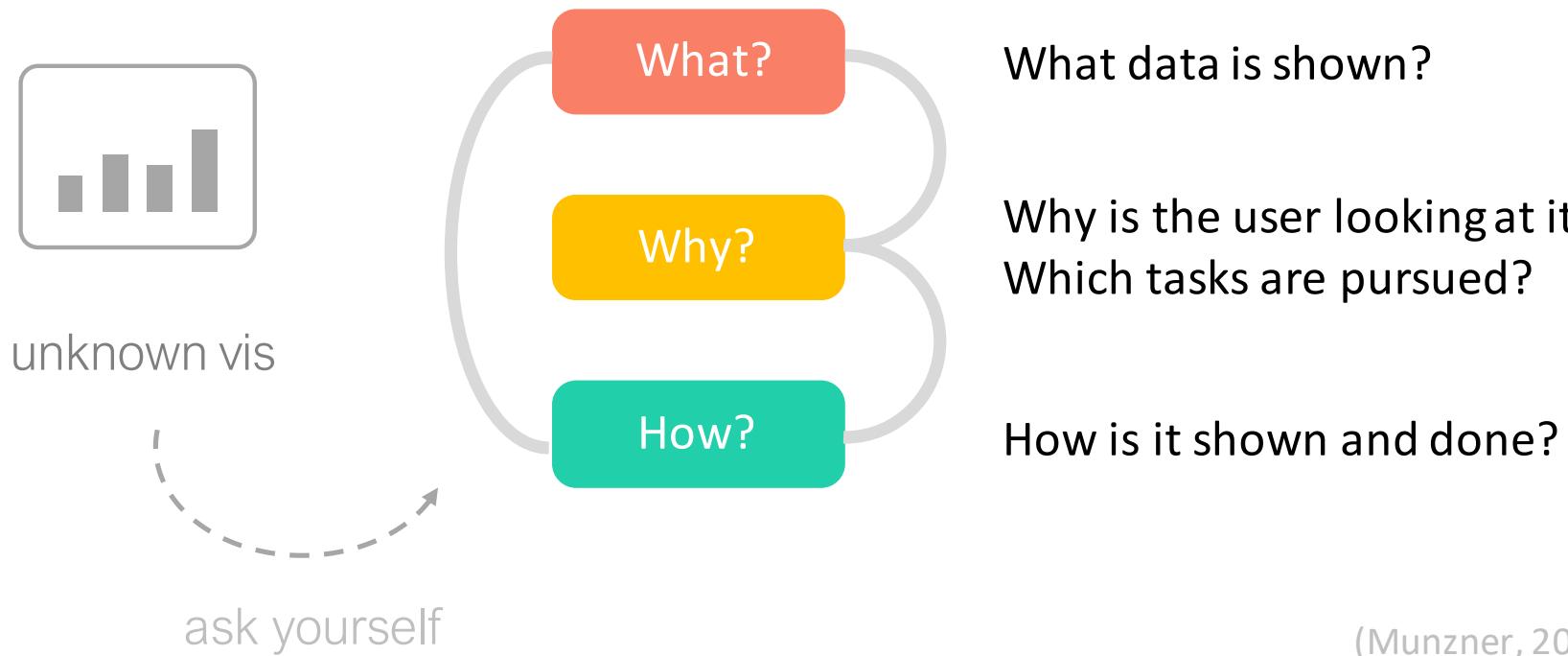
- Munzner, “Visualization Analysis and Design”: Chapter 3 (Why? – Task Abstraction)
see also <https://www.youtube.com/watch?v=pHljd-cgICY> 
- Shneiderman, “[The Eyes Have It: A Task by Data Type Taxonomy for Information Visualizations,](#)” IEEE Symposium on Visual Languages, 1996
- Brehmer+Munzer, “A Multi-Level Typology of Abstract Visualization Tasks,” InfoVis 2013. ([link](#))

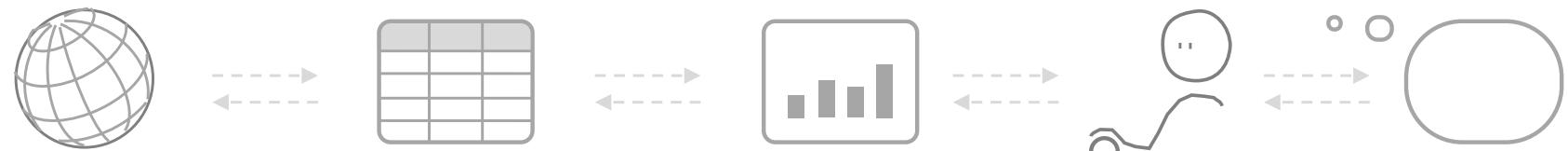
Tableau Dataset: 20w_data_happiness_development:

https://vda.univie.ac.at/Teaching/Vis/21s/data/20w_data_happiness_development.csv

Recap: A Three-part analysis framework to analyze any existing visualization



(Munzner, 2014, p. 17 ff.)



domain situation

data(set)

vis

user

task

What?

How?

Why?

Last lectures:

Data/sets Types & Semantics

- Data Types
- Dataset Types
- Attribute + Dataset Semantics

Follow-up lectures:

- Visual Encoding Principles
- Specific VIS techniques
- Complexity reduction

Today: Tasks

- The three-part Framework
- Task Taxonomies (Why?)
- VIS Design (How?) – Preview
- Tableau Examples

Schedule

Week	Date	Tuesday	Friday
1	Mar 05/08	No Lecture	Introduction (TM) pdf [Munzner Ch. 1] Rosling at TED 2006 ; Krulwich at RadioLab
2	Mar 12/15	Rекторstag	Tableau Q&A-A1 (YJ) Tutorial
3	Mar 19/22	Design Principles (TM) pdf	Data(set) Types and Semantics (TM) pdf Task Typology
	Mar 19		Due: A1 (23:55)
4	Mar 26/29	Easter	Easter
	Mar 31		Due: A2 (23:55)
5	Apr 02/05	Easter	Easter
6	Apr 09/12	Tasks (FW) pdf [Munzner Ch. 2+3] Task Taxonomy	No Lecture
7	Apr 16/19	Visual Encoding Principles (TM) pdf [Munzner Ch. 5+10] Learning_perceptual_kernels_for_vis_design Livingstone: What Art can tell us about the brain (Vis 2008 keynote),	Tableau Q&A - A3 (YJ)
8	Apr 23/26	Arrange Tables + Spatial Data (SR) pdf pdf [Munzner Ch. 7+8]	No Lecture
9	Apr 30 / May 03	Rhetorics (LK) pdf	No Lecture
	May 05		Due: A3 (23:55)
10	May 7/10	Arrange Networks / Trees (CK) pdf [Munzner Ch. 9+10]	Q&A - A4 (CK)
11	May 14/17	Arrange Networks / Trees (CK) pdf [Munzner Ch. 9+10] Example: GraphDiaries by Bach et al.	VIS in Digital Humanities I (FW) pdf
	May 19		Due: A4 (23:55)
12	May 21/24	Facet into Multiple Views (AC) pdf [Munzner Ch. 12] Improvise	VIS in Digital Humanities II (FW) pdf [Collection Visualization Survey] Survey
13	May 28/31	TextVis (TM) pdf	TextVis for DH (AC, FW) Voyant Tools Bartleby, the Scrivener Letters of Anton Chekhov
14	Jun 04/07	Design Studies (MK) pdf	Q&A - Tableau Tutorial (YJ)
15	Jun 11/14	Reduce Items & Attributes (MK) pdf [Munzner Ch. 13] DimStiller	No Lecture
	Jun 16		Due: A5 (23:55)
16	Jun 18/21	Embed: Focus+Context (MK) pdf [Munzner Ch. 11+14]	No Lecture
17	Jun 25/28	A5 Presentations SR5, 11:30-15:00	Final

Data/set Types + Semantics & Tasks

What?

- What — Data abstraction

Why?

- **Why + How — Task abstraction**

- Shneiderman's Mantra
- Empirical Study: Amar+Eagan+Stasko
- Typology: Brehmer + Munzner

- Why

- consume / produce
- search
- query

- How

(=Preview)

- encode
- manipulate
- introduce
- facets
- reduce

How?

Specific VIS techniques / idioms

Munzner's Analysis Framework : Four levels, three questions

- **domain situation**

- who are the target users?

- **abstraction**

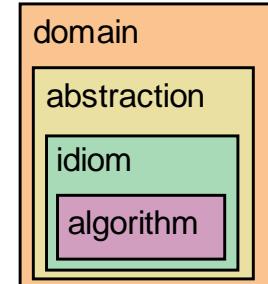
- translate from specifics of domain to vocabulary of vis
 - **what** is shown? **data abstraction**
 - **why** is the user looking at it? **task abstraction**

- **idiom**

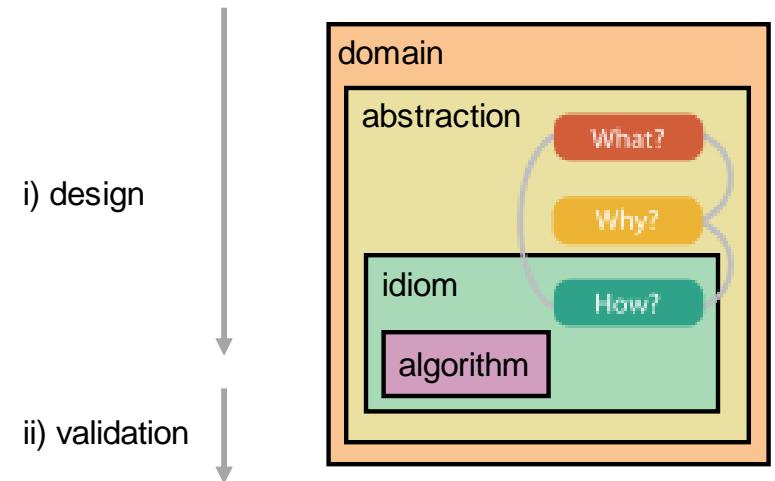
- **how** is it shown?
 - **visual encoding idiom**: how to draw
 - **interaction idiom**: how to manipulate

- **algorithm**

- efficient computation



Munzner. IEEE TVCG 15(6):921-928, 2009]



A Multi-Level Typology of Abstract Visualization Tasks
Brehmer & Munzner. IEEE TVCG 19(12):2376-2385, 2013

►► More on the whole framework in Lecture No. 13 ("Design Studies")

Q: Why should we be interested in users' „*tasks*“? If we were – how do we get a grip on them?

Def: “Visualization systems provide visual representations of datasets intended to help people carry out some **task** more effectively”

(Munzner)



task

/ta:sk/

noun

a piece of **work** to be done or undertaken.

Similar: job duty chore charge labour piece of work
piece of business assignment function commission mission engagement
occupation undertaking exercise business responsibility errand detail
endeavour enterprise venture quest problem burden ^

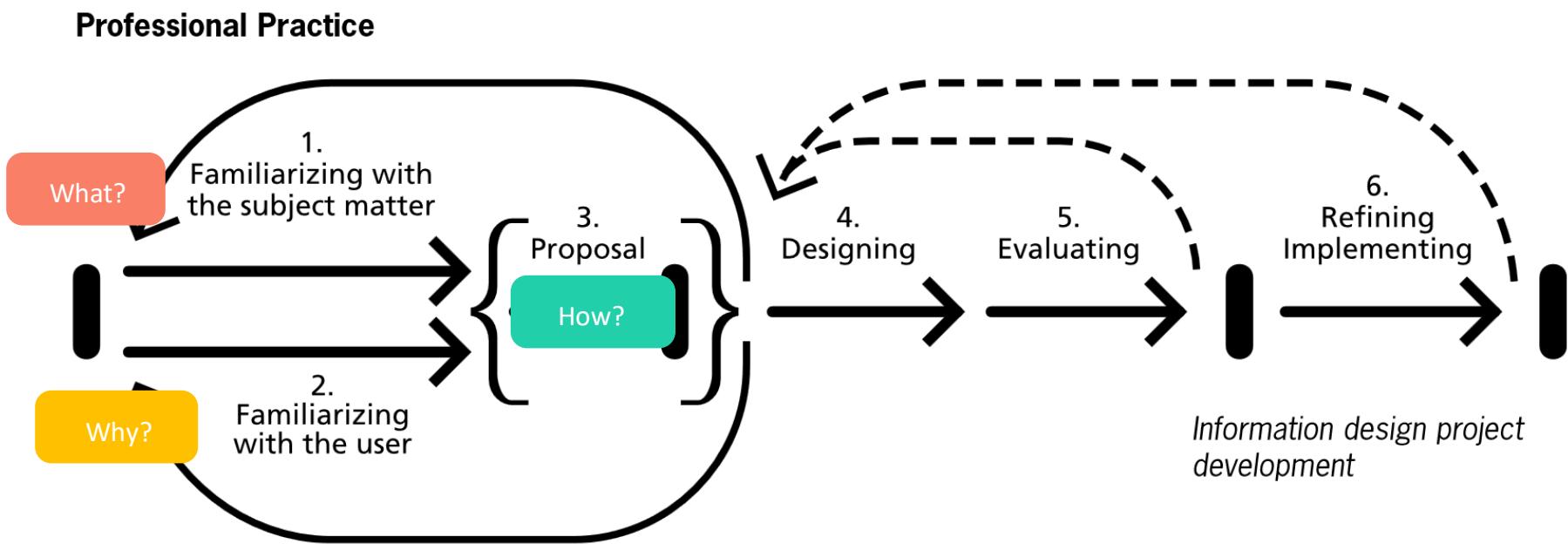


The world is full of **work** to be done. → VIS aims to „help people carry out some **work** more effectively“ →

Thus good VIS development starts with **analyzing** and **abstracting** „*tasks*“ as specific work practices and related challenges of people in real world domains.

Relevancve: For VIS designers, (knowledge about) tasks provides the actual instruction for every development job or research project!

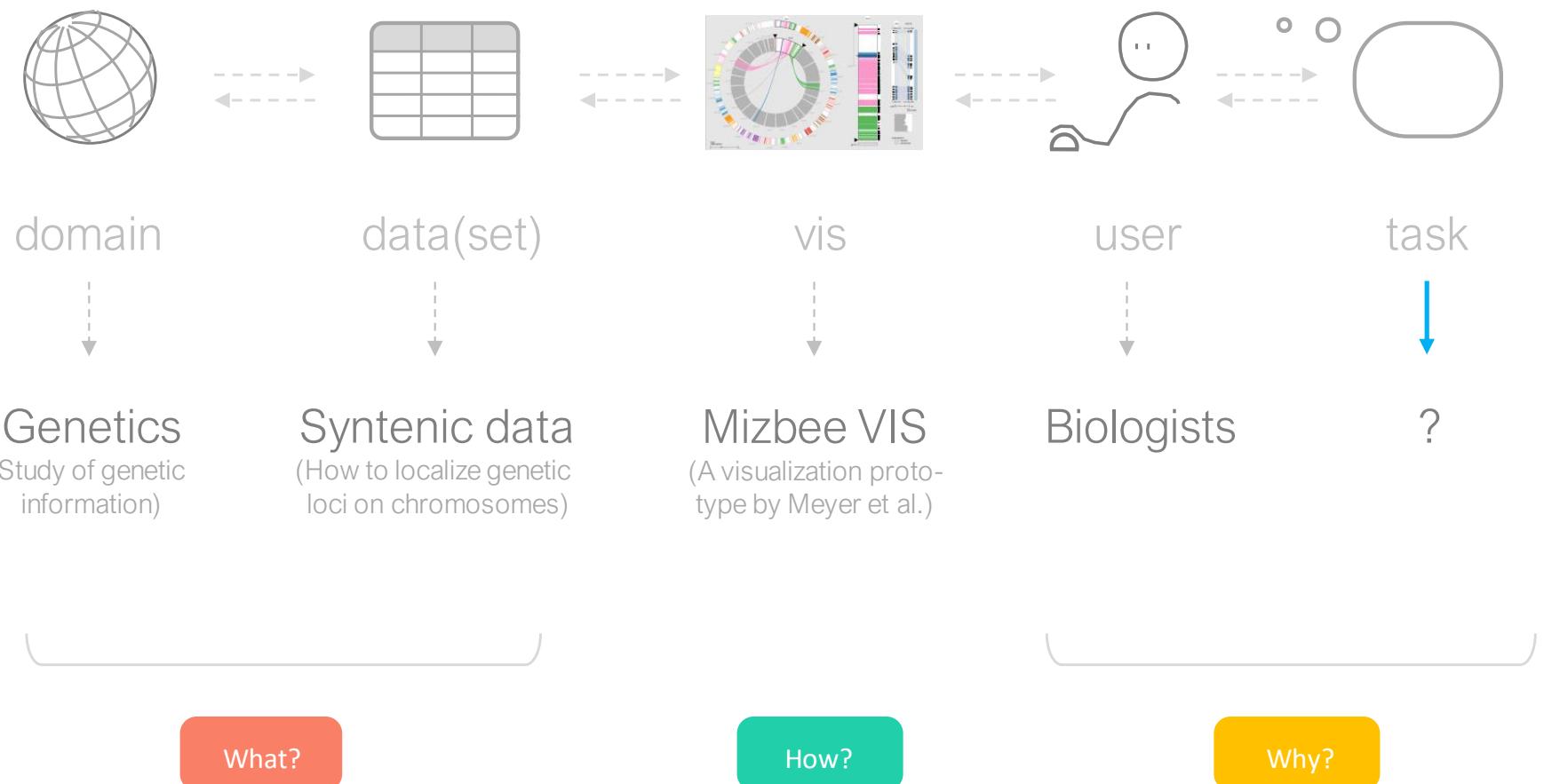
IIID Information Design - General Project Model



IIID IDX: <http://www.iid.net/PDFs/idxPublication.pdf>

Example No.1:

Meyer et al. (2009). *MizBee: A Multiscale Synteny Browser*.



Task Abstraction – Meyer et al.

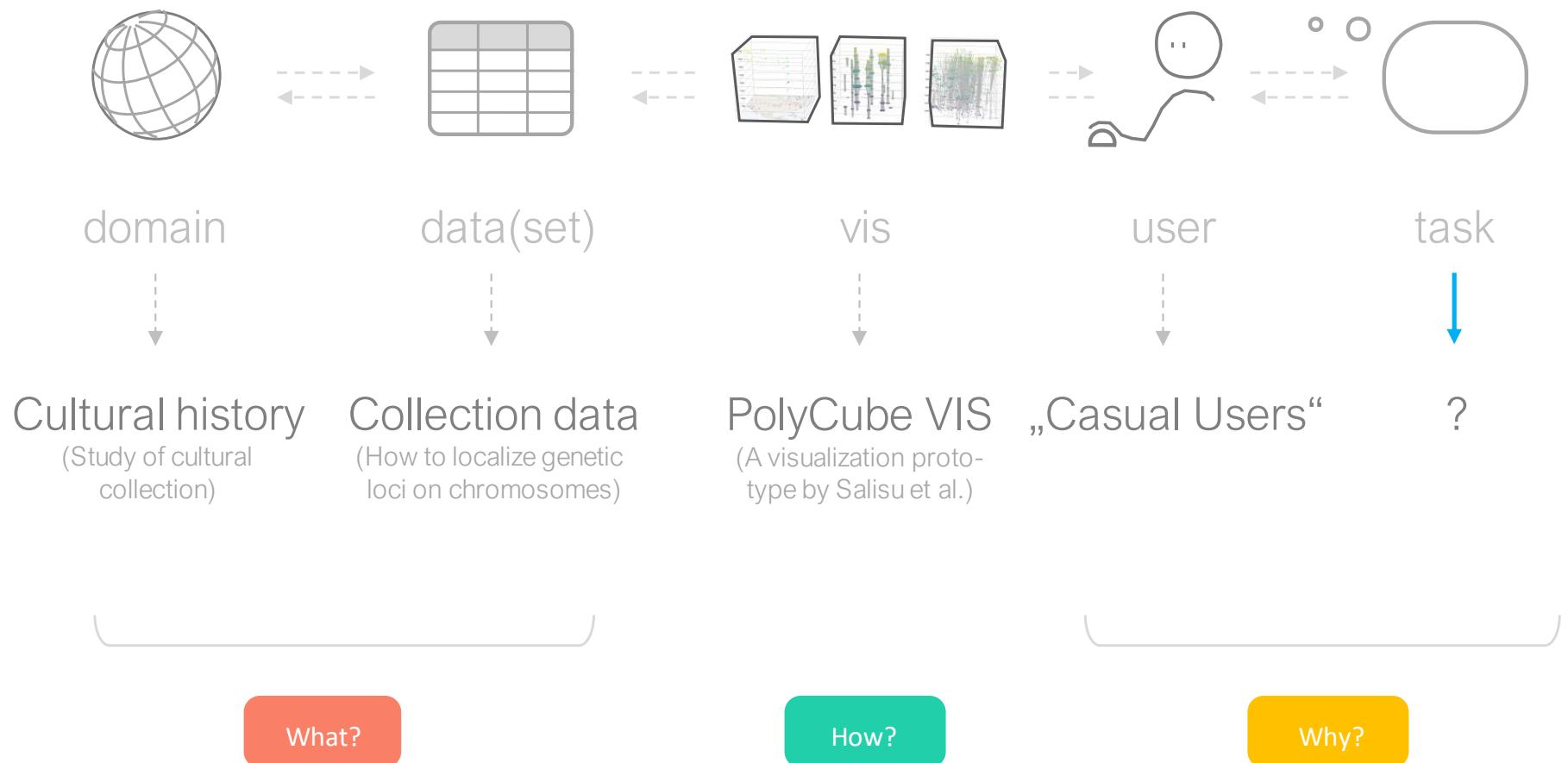
We have identified a set of 14 fundamental questions that biologists ask to gain scientific insight at different stages of the data analysis pipeline, shown in Table 1. We have organized them according to the scale at which they operate and the type of conservation relationship they address. Some of these questions pertain to the early data generation stage, probing the results of computational algorithms that de-

Table 1. Questions for the analysis of conserved syntenic data, with the scale and relationship addressed by each. The scales are: *g*, genome; *c*, chromosome; *b*, block; and *f*, feature. The relationships are: *p*, proximity/location; *z*, size; *o*, orientation; and *s*, similarity.

	question	scale				relationship			
		<i>g</i>	<i>c</i>	<i>b</i>	<i>f</i>	<i>p</i>	<i>z</i>	<i>o</i>	<i>s</i>
1	Which chromosomes share conserved blocks?	X				X			
2	For one chromosome, how many other chromosomes does it share blocks with?	X	X				X		
3	What is the density of coverage and where are the gaps on: chromosomes? blocks?	X	X	X			X		
4	Where are the blocks: on chromosomes? around a specific location on a chromosome?	X	X				X		
5	What are the sizes and locations of other genomic features near a block?		X			X	X		
6	How large are the blocks?		X				X		
7	Do neighboring blocks go to the same: chromosomes? relative location on a chromosome?	X	X				X		
8	Are the orientations matched or inverted for: block pairs? feature pairs?		X	X				X	
9	Do the orientations match for pairs of: neighboring blocks? features within a block?		X	X				X	
10	Are similarity scores alike: with respect to neighboring blocks? within a block?		X	X					X
11	Are the paired features within a block contiguous?			X		X			
12	How large is a feature relative to other genes within a block?			X			X		
13	What are the sizes, locations, and names of features within a block?			X		X	X		
14	What are the differences between individual nucleotides of feature pairs?				X				X

Example No.2:

Windhager et al. (2019). *PolyCube Collection Visualization*



Task Abstraction



Users: Regarding the users of our system, we intended to support casual users, i.e., a heterogeneous group of users with different levels of visual literacy, expert knowledge, and interest. As they explore cultural collections mainly for leisure purposes, they are not necessarily motivated to invest high amounts of cognitive load to fully process and interconnect all information and perspectives available. Therefore, casual users can benefit from visualization designs, which actively assist them in the construction of a mental model.

In addition, it is important to be aware that casual users will only persist in exploration as long as it is rewarding for them (i.e., engaging, aesthetically pleasing, interesting, or intriguing). Hermetic and complex system designs can easily terminate the interaction at an early stage.

Tasks: Casual users of cultural collections often have no specific information needs, but look around and browse for something interesting, which they can access for details on demand. They do not pursue concrete tasks, but they are keen on gaining an overview and exploring the digital collection. When designing and developing PolyCube, we attempted to clarify these rather vague tasks as follows: 1) gaining a (synoptic) overview and conceptual orientation regarding the distribution of the major data dimensions of a cultural collection (e.g., time, space, categories, relations), 2) finding single objects of personal interest and inspecting their details, and 3) browsing through objects (e.g., according to time, relations, geographic origins, or shared categories). These tasks align with the task typology proposed by Brehmer and Munzner,¹³ as the users in this case are consumers of the visualization and want to *explore, browse, and enjoy*.

The case for abstraction:

Transforming task descriptions from domain-specific language into abstract form allows you to reason about similarities and differences between them. Otherwise, it's hard to make useful comparisons between domain situations, because if you don't do this kind of translation then everything just appears to be different. That apparent difference is misleading: there are a lot of similarities in what people want to do once you strip away the surface language differences.

for example, an epidemiologist studying the spread of a new strain of influenza might initially describe her task as

“contrast the prognosis of patients who were intubated in the ICU more than one month after exposure to patients hospitalized within the first week”

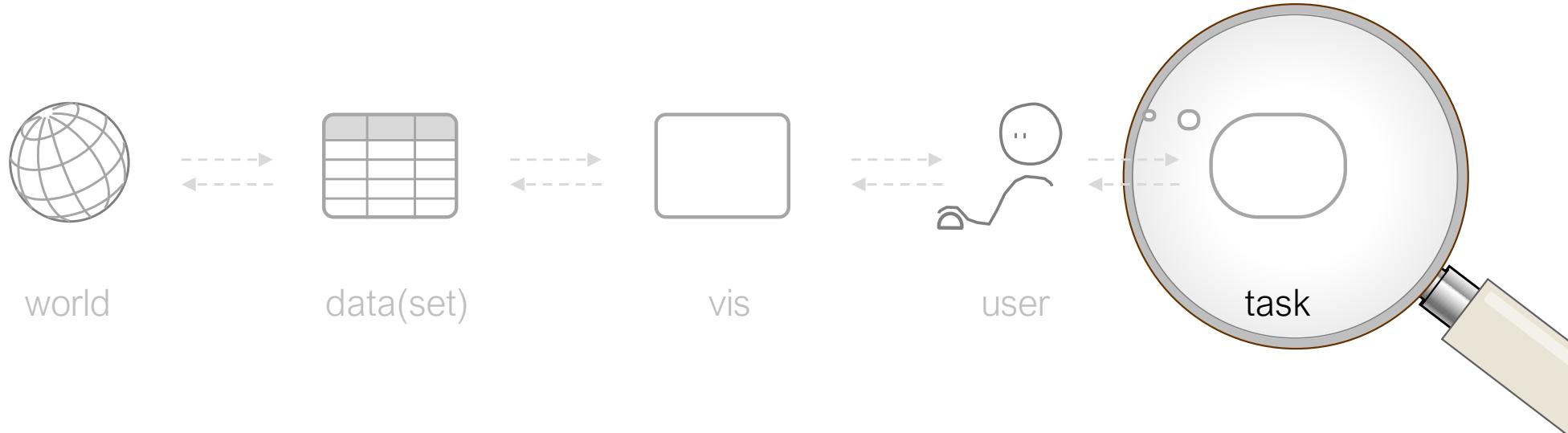
while a biologist studying immune system response might use language such as

“see if the results for the tissue samples treated with LL-37 match up with the ones without the peptide”

Even if you know what all the specialized vocabulary means, it's still hard to think about what these two descriptions have in common because they're using different words: “contrast” versus “match up”.

If you transform these into descriptions using a consistent set of generic terms, then you can spot that these two tasks are just two instances of the same thing: **“compare values between two groups”**.

Task Taxonomies



- Q: What are the most important types of **tasks or activities** users frequently (want to) pursue in complex information spaces and for various data?
- Can we describe, cluster and abstract these information activities, so that we arrive at a more general understanding of tasks (i.e. task taxonomies or typologies), and use it to guide our own VIS developments?

A: Dozens of task taxonomies and typologies has already been developed. Let's have a look at important concepts from a couple of them, including

- Shneiderman, B. (1996). ["The Eyes Have It: A Task by Data Type Taxonomy for Information Visualizations," IEEE Symposium on Visual Languages.](#)
- Amar, R., Eagan, J., & Stasko, J. (2005). ["Low-level components of analytic activity in information visualization," InfoVis.](#)
- Brehmer, M. & Munzner, T. (2013). [A Multi-Level Typology of Abstract Visualization Tasks, InfoVis.](#)
- Munzner, T. (2014). *Visualization Analysis and Design. Chapter 3* (Why? – Task Abstraction)
→ see also the video lecture: <https://www.youtube.com/watch?v=pHljd-cgICY>

Shneiderman's Task by Data Type Taxonomy

[Shneiderman, 1996, [link](#)]

Shneiderman's Basic Tasks

- **Overview:** Gain an overview of the entire collection
- **Zoom:** Zoom in on items of interest
- **Filter:** filter out uninteresting items
- **Details-on-demand:** Select an item or group and get details when needed
- **Relate:** View relationships among items
- **History:** Keep a history of actions to support undo, replay, and progressive refinement
- **Extract:** Allow extraction of sub-collections and of the query parameters

The Mantra

[Shneiderman, 1996]

Shneiderman's Visual Information Seeking Mantra

[Shneiderman, 1996]

There are many visual design guidelines but the basic principle might be summarized as the Visual Information Seeking Mantra:

***Overview first,
zoom and filter, then
details-on-demand***



e.g.: <https://galaxy.opensyllabus.org>

Please check (in groups of 2 or 3)

- How is the „overview“ done?
- How can you „zoom and filter“?
- What are the „details“ – and which can you access?

Empirical Study: Amar, Eagan & Stasko

[Amar, Eagan & Stasko, 2005, [link](#)]

- Q: What are the most important types of **analytical tasks** or users frequently (want to) pursue when using visualizations and visual analytics tools?
- Amar, Eagan & Stasko conducted an empirical study with students, who were given five different datasets - and documented their analytic activities.
- They recorded about 200 different analytical activities, and clustered them with an affinity diagram into 10 distinct types of tasks.

Task Abstraction (Amar, Eagan & Stasko, 2005)

4 AN ANALYTIC TASK TAXONOMY

The ten tasks from the affinity diagramming analysis are:

- Retrieve Value
- Filter
- Compute Derived Value
- Find Extremum
- Sort
- Determine Range
- Characterize Distribution
- Find Anomalies
- Cluster
- Correlate

Examples:

- Order the cars by weight.
- Rank the cereals by calories.

[Amar, Eagan, & Stasko, 2005]

- 1. Filter:** Find data that satisfies conditions
- 2. Find Extremum:** Find data with extreme values
- 3. Sort:** Rank data according to some metric
- 4. Determine Range:** Find span of data values
- 5. Find Anomalies:** Find unexpected / extreme values

Examples:

- Order the cars by weight.
- Rank the cereals by calories.

[Amar, Eagan, & Stasko, 2005]

- 1. Filter:** Find data that satisfies conditions
- 2. Find Extremum:** Find data with extreme values
- 3. Sort:** Rank data according to some metric
- 4. Determine Range:** Find span of data values
- 5. Find Anomalies:** Find unexpected / extreme values

Examples:

- What Kellogg's cereals have high fiber?
- What comedies have won awards?
- Which funds underperformed the SP-500?

[Amar, Eagan, & Stasko, 2005]

- 1. Filter:** Find data that satisfies conditions
- 2. Find Extremum:** Find data with extreme values
- 3. Sort:** Rank data according to some metric
- 4. Determine Range:** Find span of data values
- 5. Find Anomalies:** Find unexpected / extreme values

Examples:

- What Kellogg's cereals have high fiber?
- What comedies have won awards?
- Which funds underperformed the SP-500?

[Amar, Eagan, & Stasko, 2005]

- 1. Filter:** Find data that satisfies conditions
- 2. **Find Extremum**:** Find data with extreme values
- 3. **Sort**:** Rank data according to some metric
- 4. **Determine Range**:** Find span of data values
- 5. **Find Anomalies**:** Find unexpected / extreme values

Examples:

- Are there exceptions to the relationship between horsepower and acceleration?
- Are there any outliers in protein?

[Amar, Eagan, & Stasko, 2005]

- 1. Filter:** Find data that satisfies conditions
- 2. Find Extremum:** Find data with extreme values
- 3. Sort:** Rank data according to some metric
- 4. Determine Range:** Find span of data values
- 5. Find Anomalies:** Find unexpected / extreme values

Examples:

- Are there exceptions to the relationship between horsepower and acceleration?
- Are there any outliers in protein?

[Amar, Eagan, & Stasko, 2005]

- 1. Filter:** Find data that satisfies conditions
- 2. Find Extremum:** Find data with extreme values
- 3. Sort:** Rank data according to some metric
- 4. Determine Range:** Find span of data values
- 5. Find Anomalies:** Find unexpected / extreme values

Typology: Why?

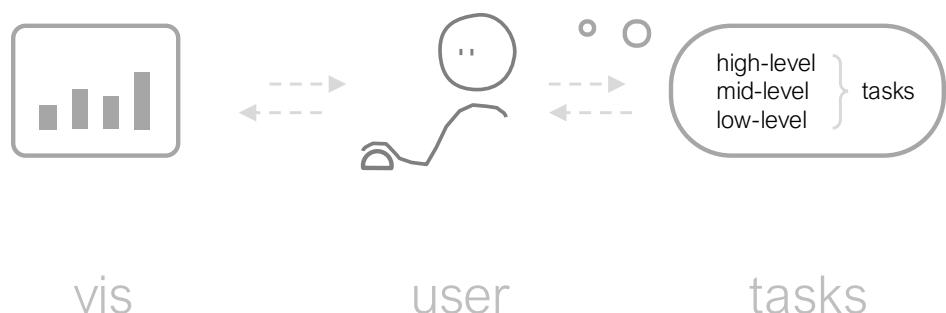
Brehmer & Munzner

[Brehmer+Munzner, 2013, [link](#)]

Why? → Task Abstraction!

Munzner provides a hierarchy of

- high-level tasks: consume / produce
- mid-level tasks: search
- low-level tasks: query



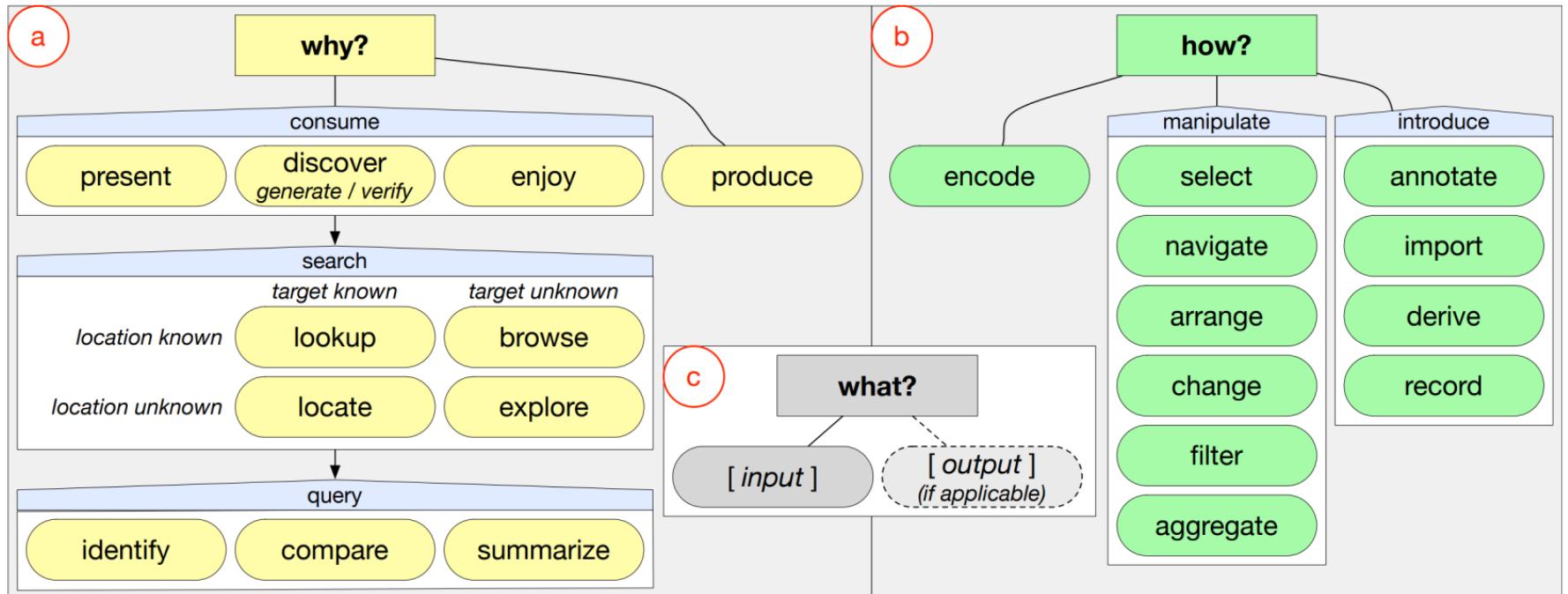
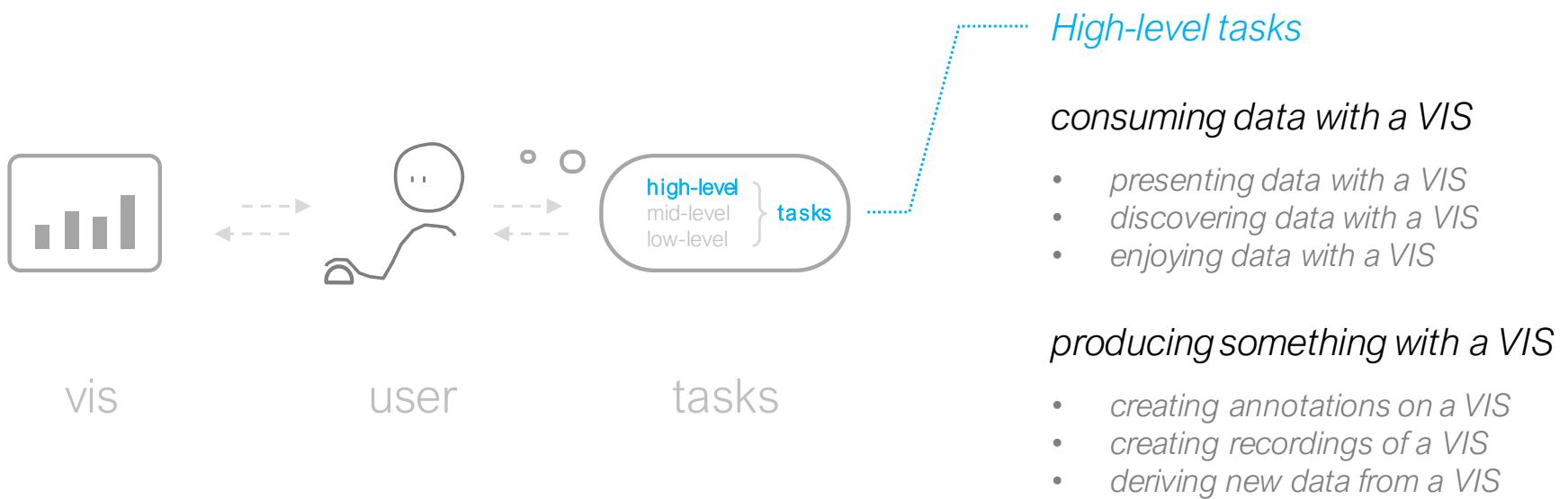


Fig. 1. Our multi-level typology of abstract visualization tasks. The typology spans *why*, *how*, and *what*; task descriptions are formed by nodes from each part: a) *why* a task is performed, from high-level (consume vs. produce) to mid-level (search) to low-level (query). b) *how* a task is executed in terms of *methods*, defined as families of related visual encoding and interaction techniques. c) *what* the task inputs and outputs are.

[Brehmer+Munzner, 2013, [link](#)]

Consume vs. Produce (High-level)



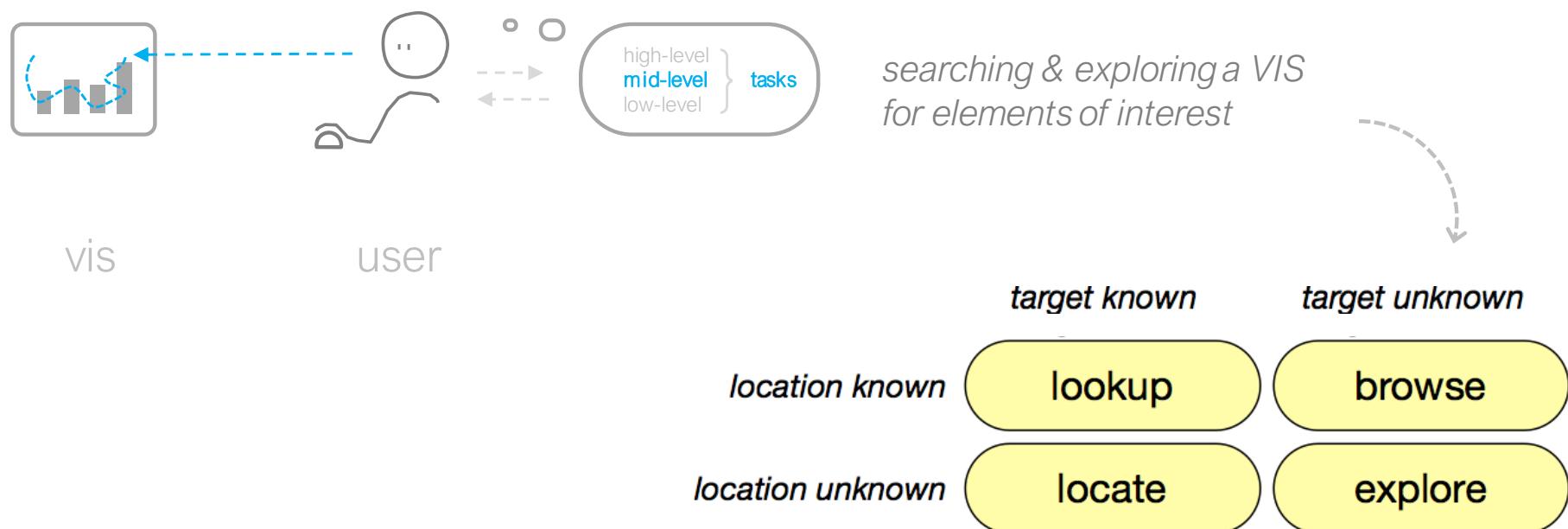
Consume vs. Produce (High-level)

- Consume (most common)
 - **present**
 - not just static (e.g. interactive graphics in newspapers / NY Times)
 - **discover**
 - generation / verification of hypothesis
 - **enjoy**
 - “casual” vis
 - e.g. Name Voyager (<http://www.babynamewizard.com/voyager>)
- Produce
 - help the user produce something based on a VIS,
e.g. annotations, recordings, derivates

Search (Mid-level)

Search: Regardless of whether the intent is to present, discover, or merely enjoy, the user must find elements of interest in the visualization.

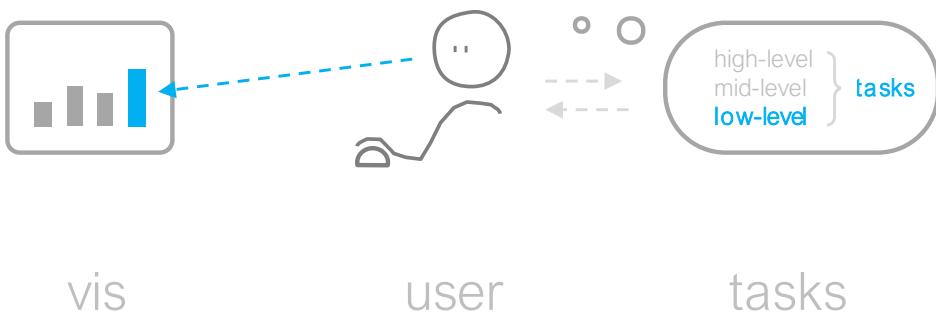
(Brehmer+Munzner, 2013)



Query (Low-level)

Query: Once a target or set of targets has been found, a user will identify, compare, or summarize these targets.

(Brehmer+Munzner, 2013)



- *identifying single targets*
- *comparing two or multiple targets*
- *summarizing a whole set of targets*

Why?

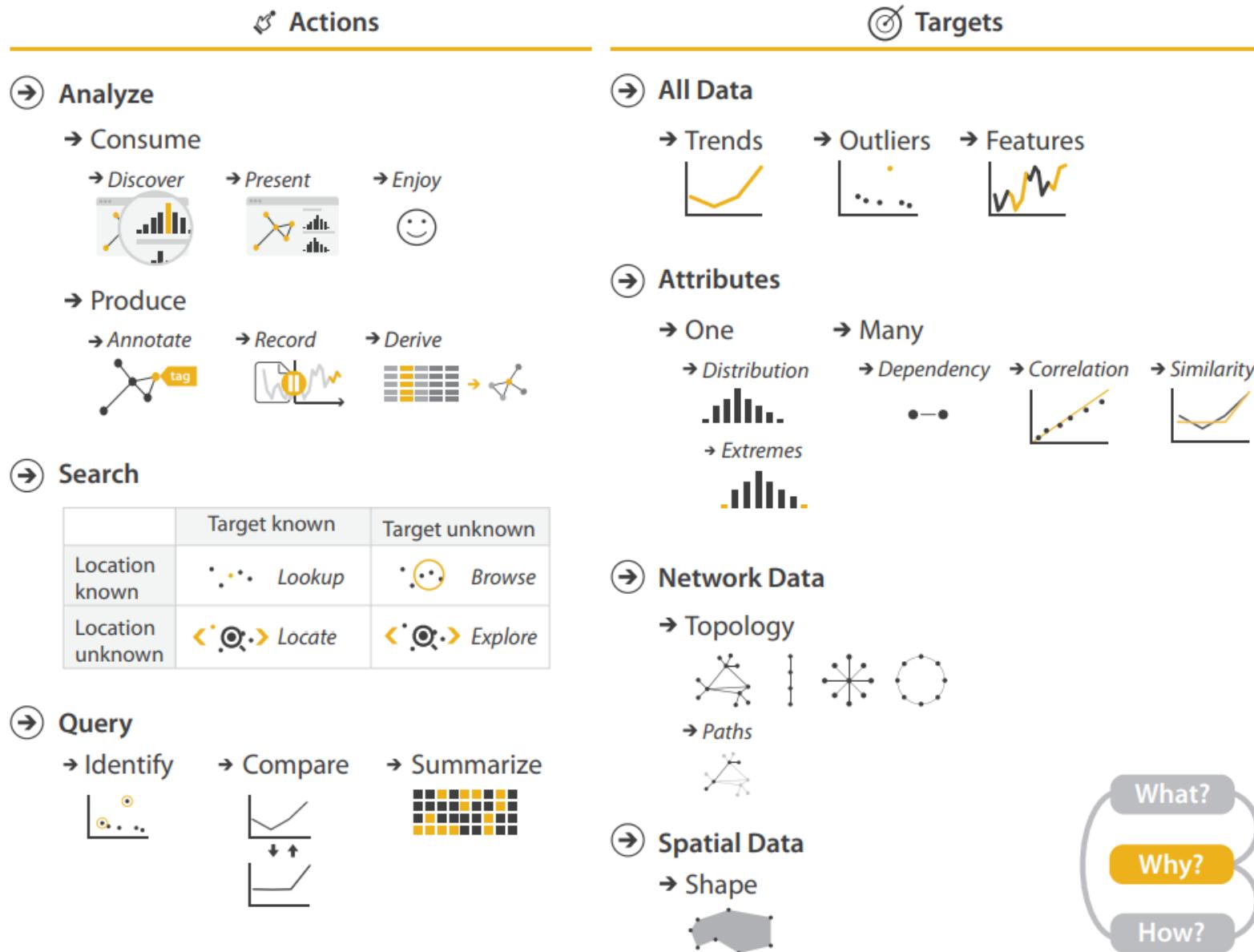


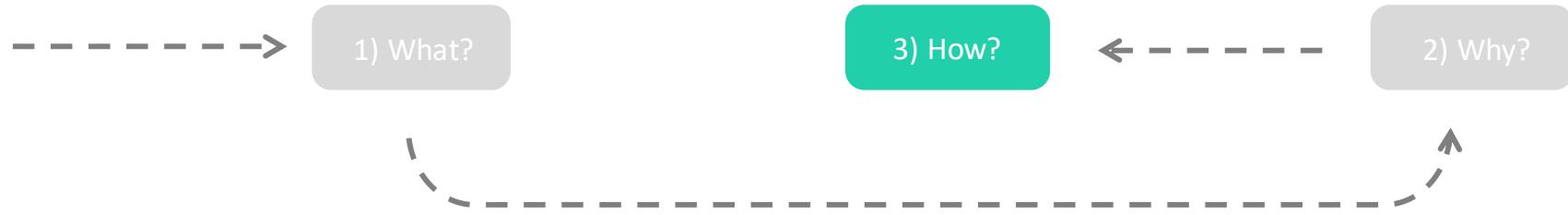
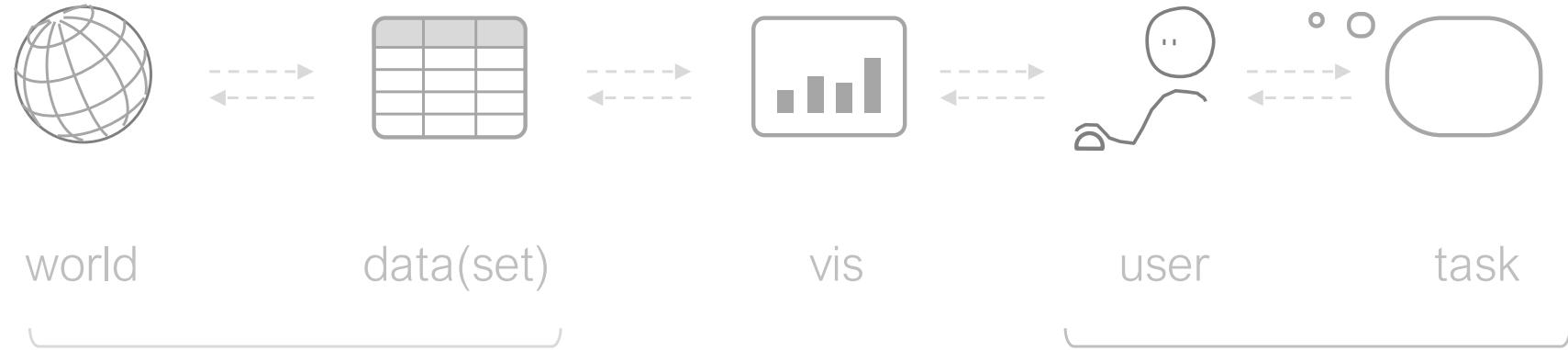
Figure 3.1. Why people are using vis in terms of actions and targets.

[Munzner, 2014]

Typology: **How?** (Preview)

Brehmer & Munzner

[Brehmer+Munzner, 2013]



[After learning about what and why, we] “now turn our consideration to the **how** part of our typology, which contains methods, defined as families of related visual encoding and interaction techniques. [...] We distinguish between three classes of methods: those for encoding data, those for manipulating existing elements in a visualization, and those for introducing new elements into a visualization.” (Brehmer & Munzner, 2013)

How can interactive visualizations be designed to support people to carry out their tasks? Munzner suggests three classes of methods to so:

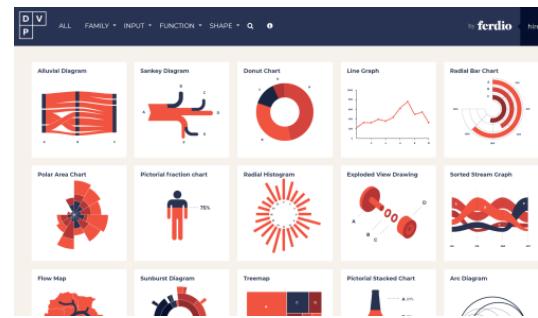
- methods to **encode** data into a VIS
- methods to **manipulate** existing elements of a VIS
- methods to **introduce** new elements into a VIS

Encode

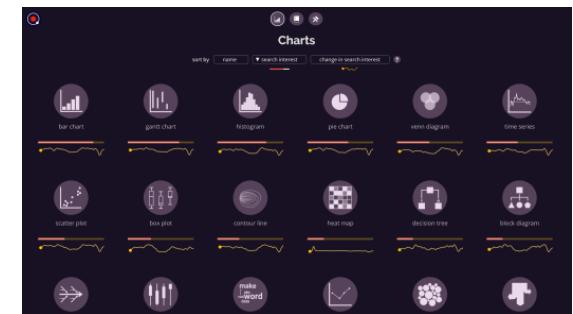
visually encoding information through marks and channels (e.g. color, shapes, size, position etc.) – to create any known or new kind of VIS



<https://datavizcatalogue.com/index.html>



<https://datavizproject.com/>



<http://visualizationuniverse.com/charts/>

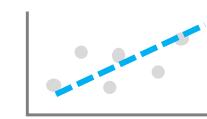
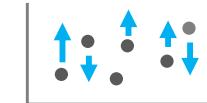
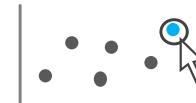
→ see last lecture for „general encoding principles“ and the following lectures for specific encoding techniques (e.g. arrangements of tables, spatial data, networks, or tree visualizations)

Manipulate

→ see specific lecture on “Reduce – Items and attributes”

The following methods affect existing elements of a visualization, modifying them to some extent. (Brehmer & Munzner, 2013)

- select
- navigate
- (re-)arrange
- change
- filter
- aggregate



Introduce

Introduce: While manipulate methods alter existing elements of the visualization, introduce methods add new elements.

(Brehmer & Munzner, 2013)

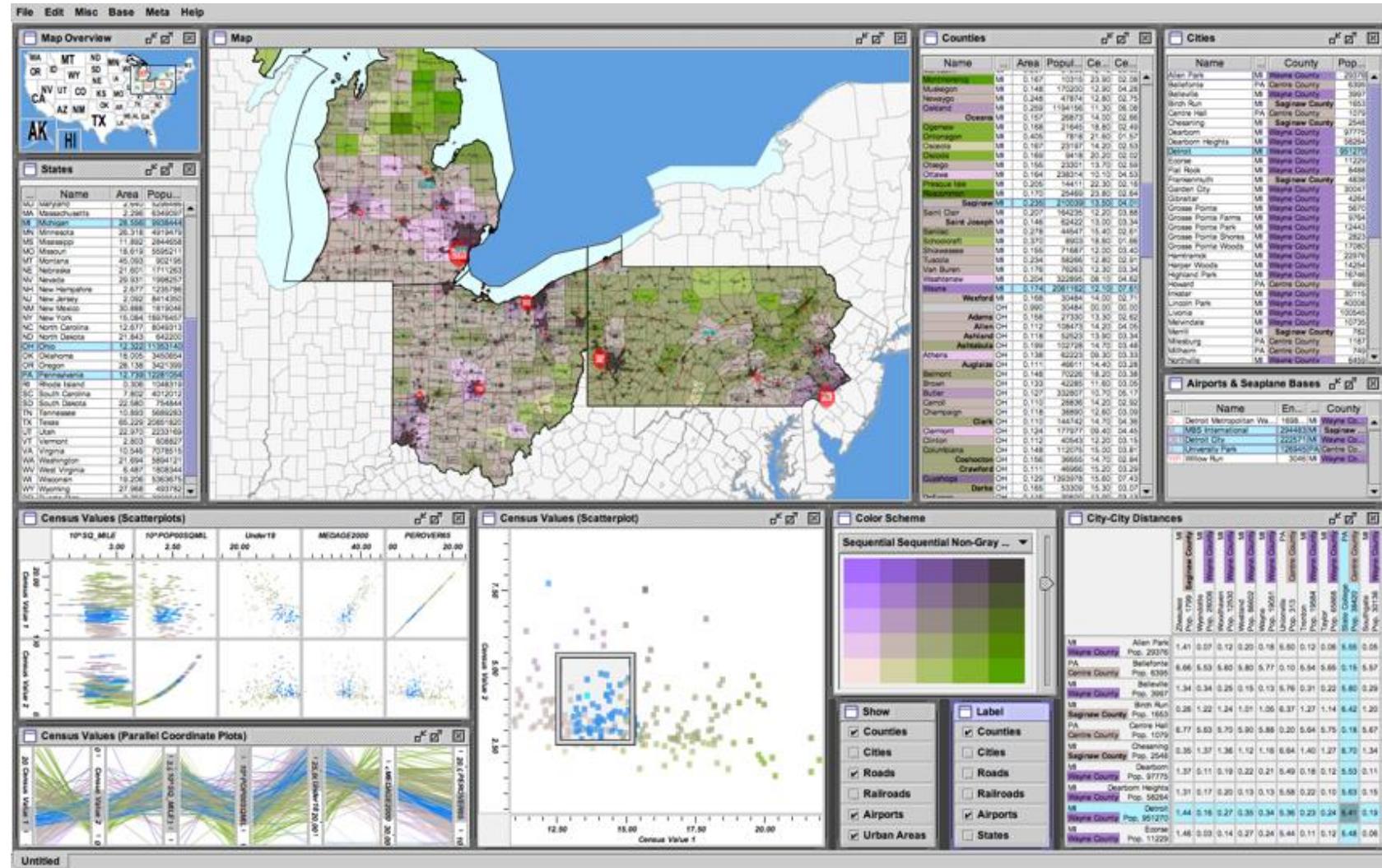
- annotate
 - with text label etc. (classification)
 - acts as a new attribute
- import
 - new data items to be loaded
- derive
 - derive new data attributes
- record
 - screenshots, bookmarks, parameter settings, logs, etc.
 - graphical / use history
 - analytical provenance!

Facets (Preview)

How to use multiple views?

- juxtapose (multiple views, side-by-side)
- superimpose (multiple layers)
- change (layout, encoding → interaction)
- select (demarcation, highlighting)
- coordinate (brushing+linking, linking views)

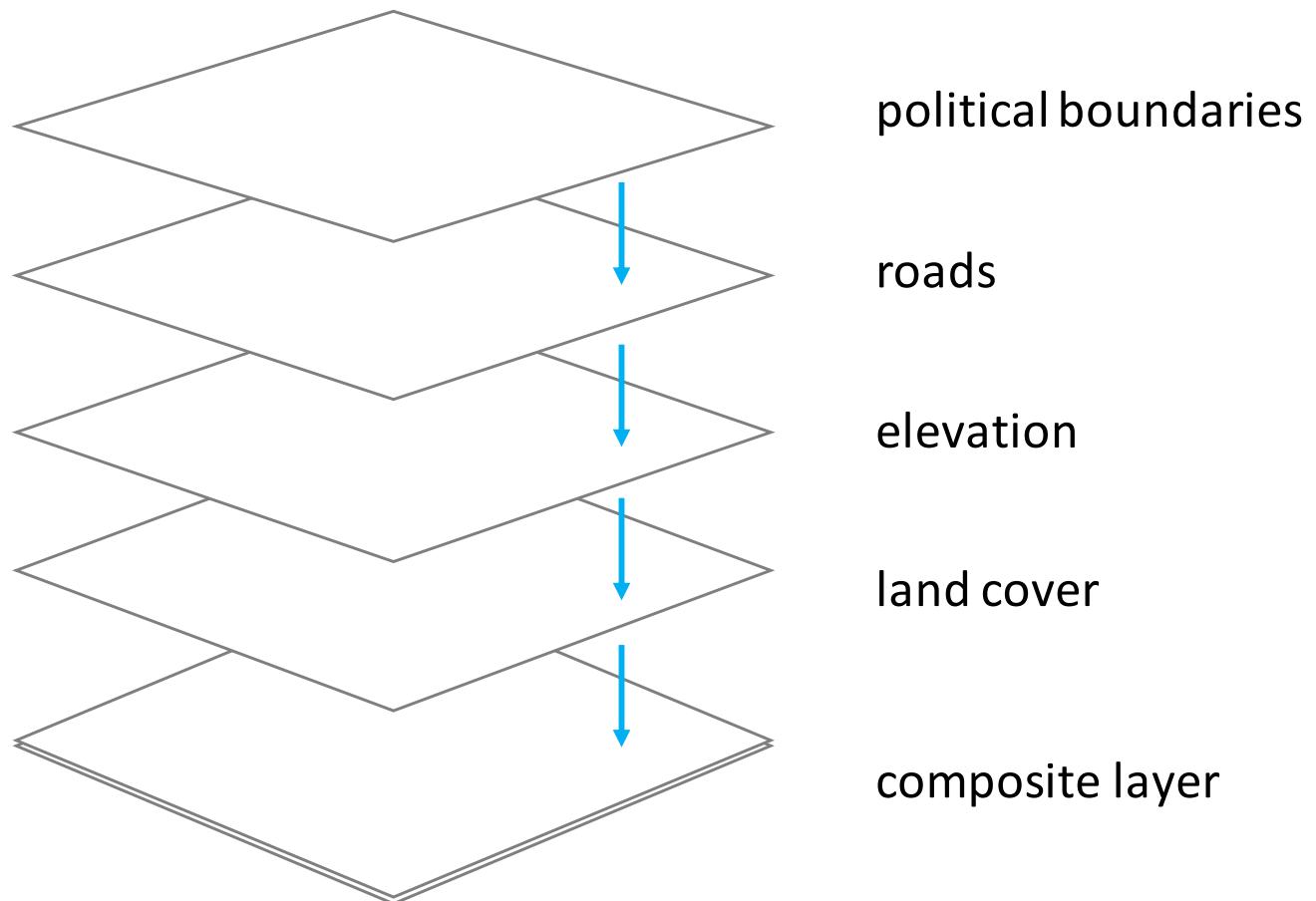
Partition & Juxtaposition



[Weaver, 2004]

Superimposition

e.g. multiple GIS layers, projected upon each other

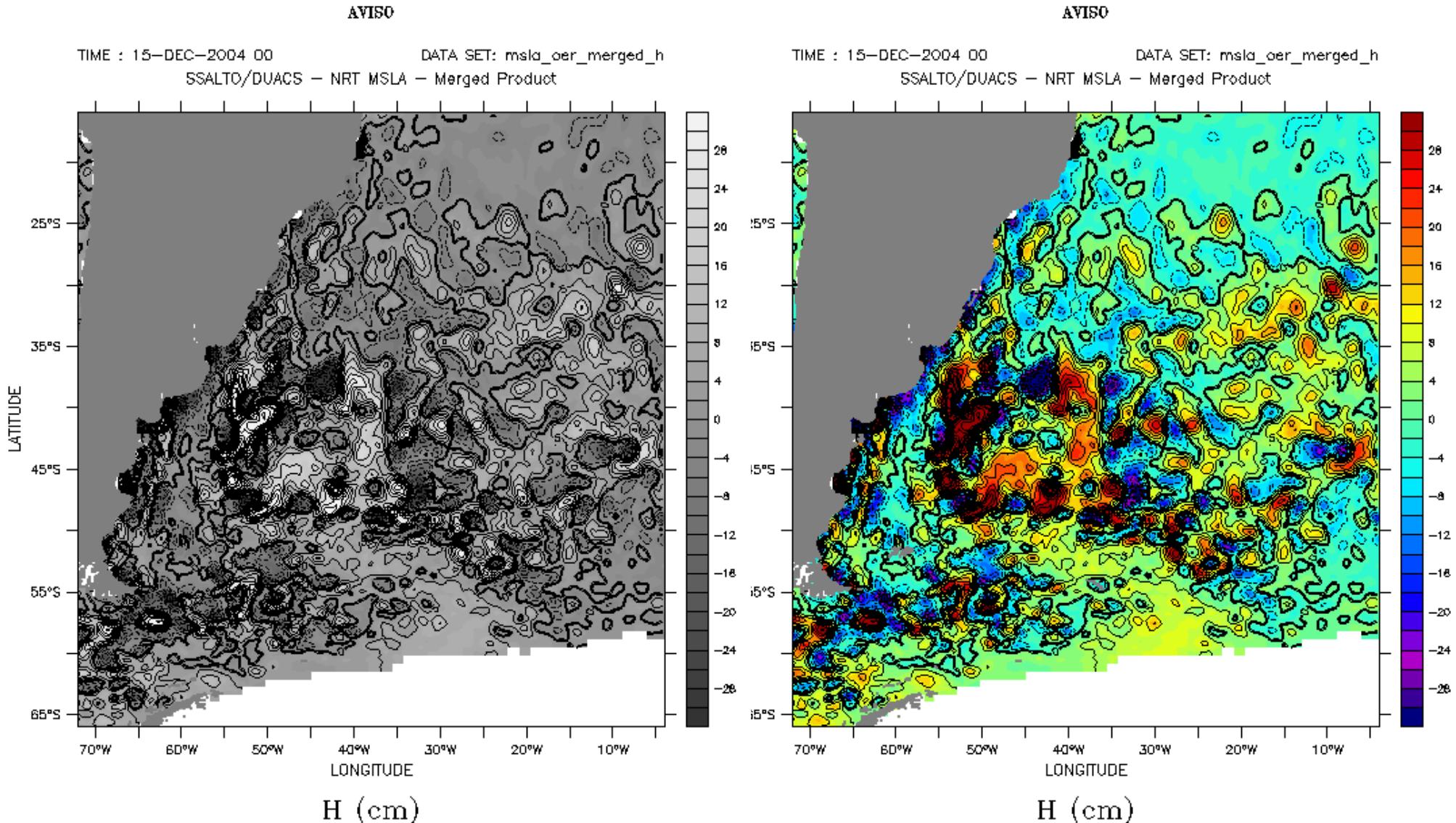


Linking (coordinate)

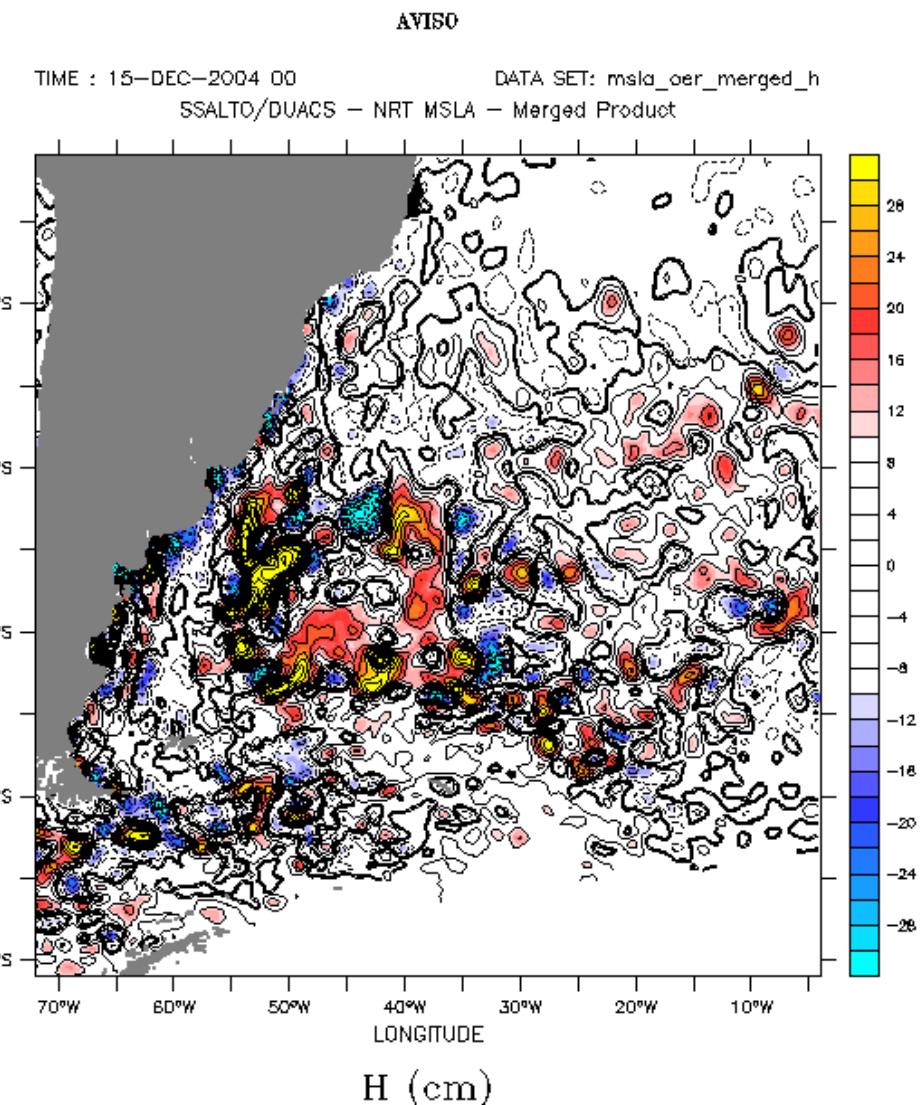
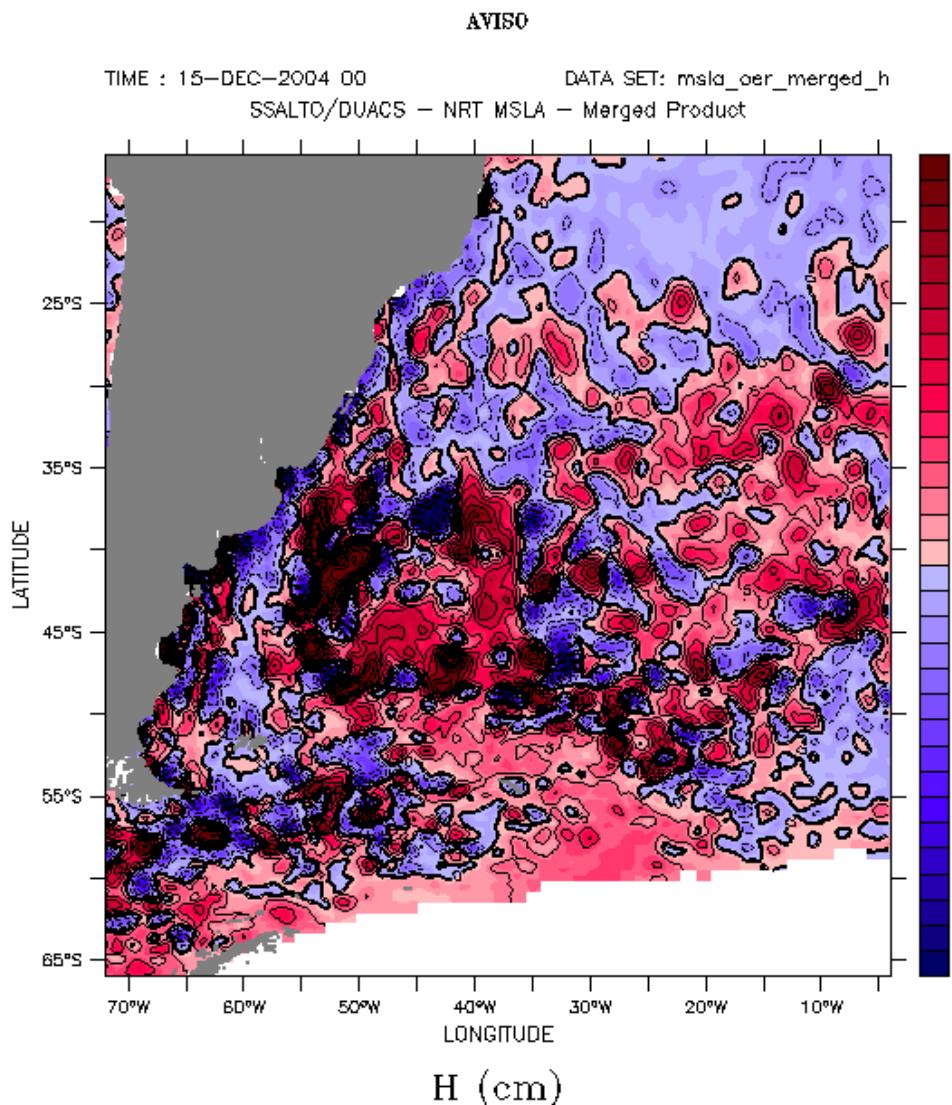


Color and Tasks

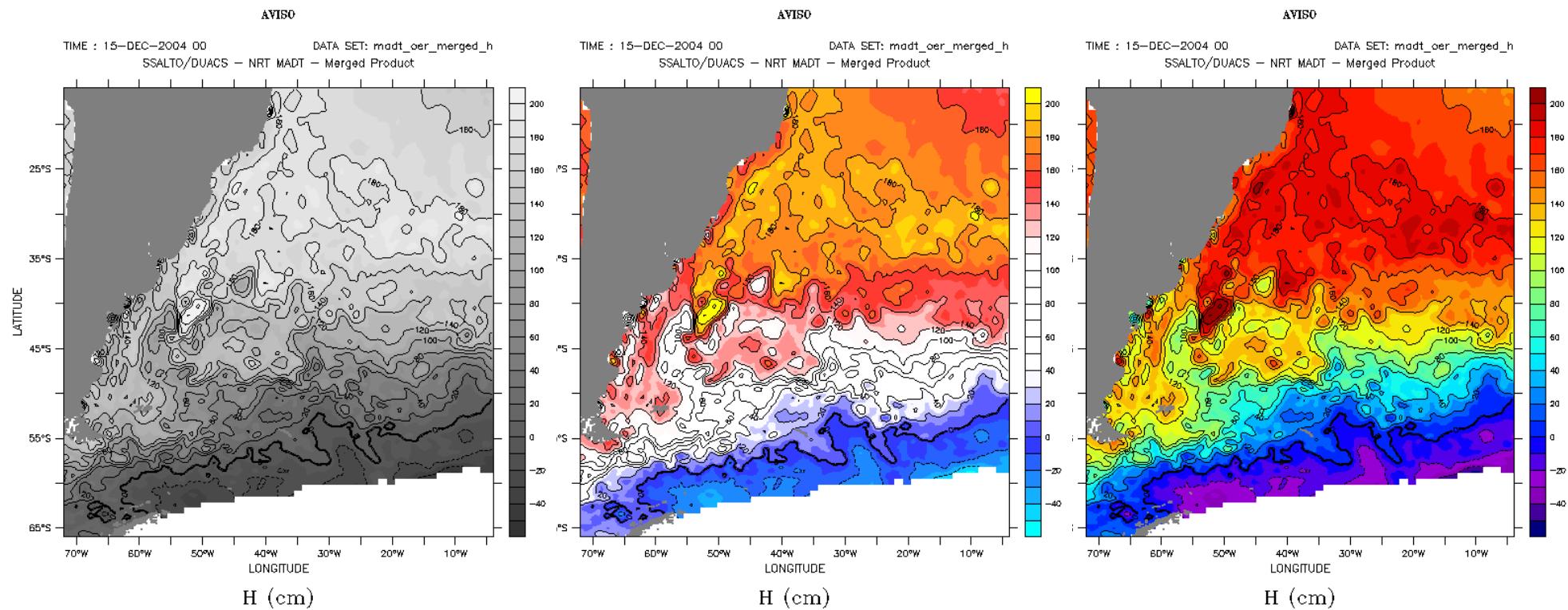
Task: Find high & low anomalies



Task: Find high & low anomalies



Task: Understand relative height



Task: Find height 120

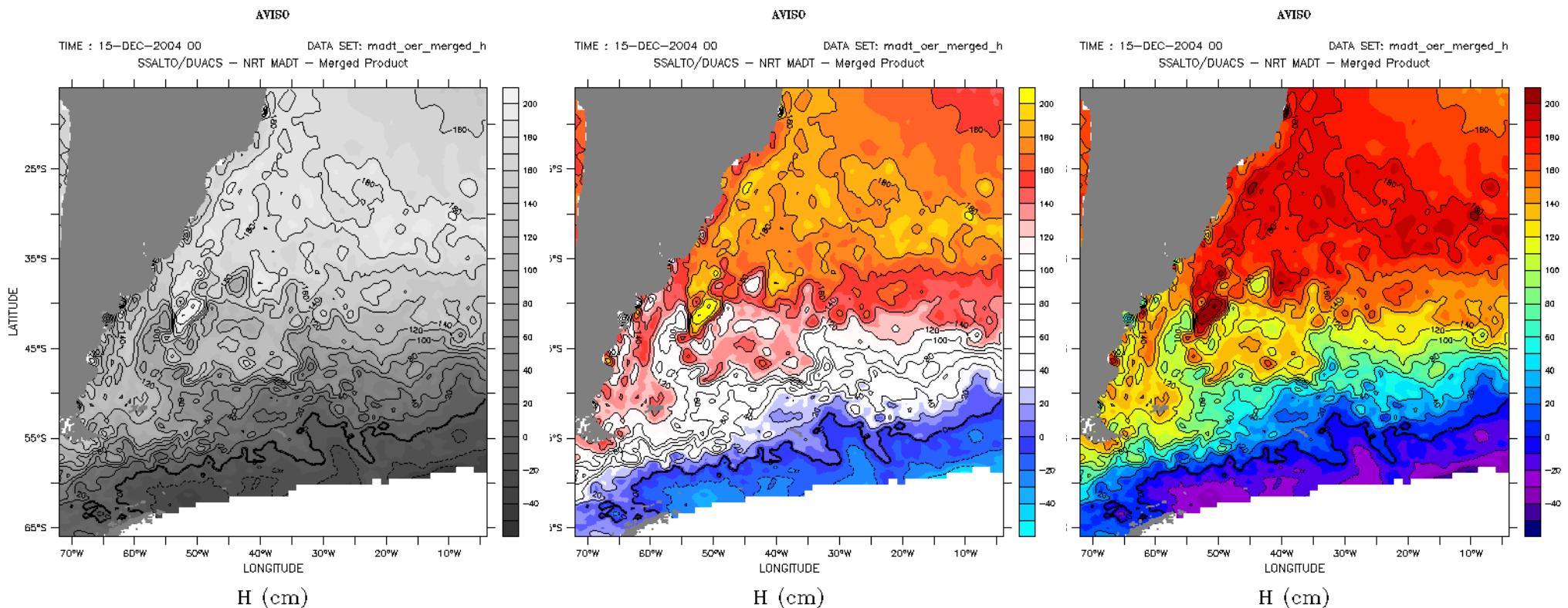
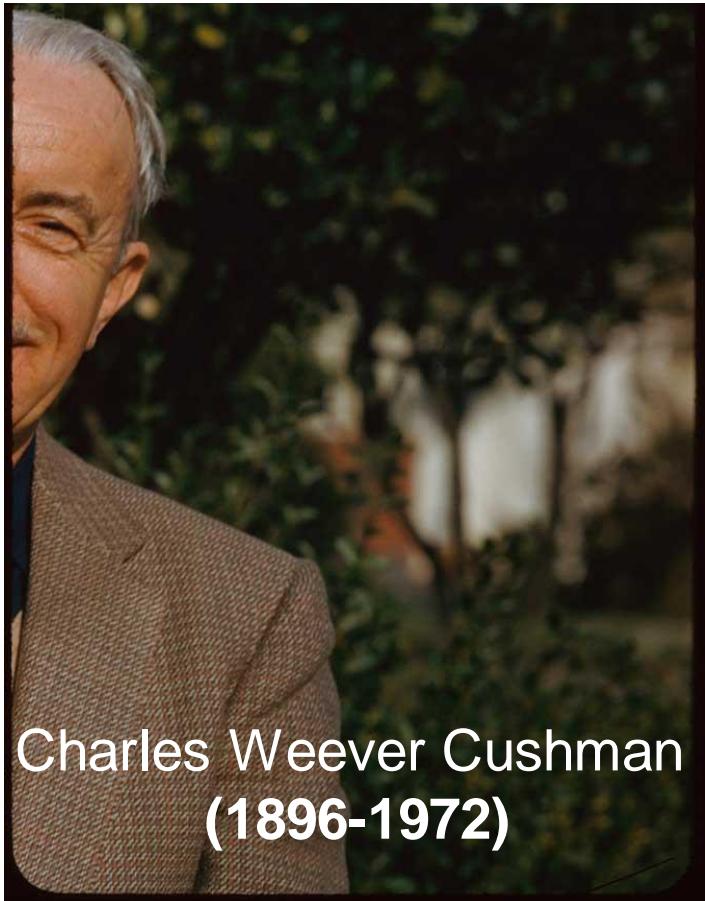


Tableau Examples

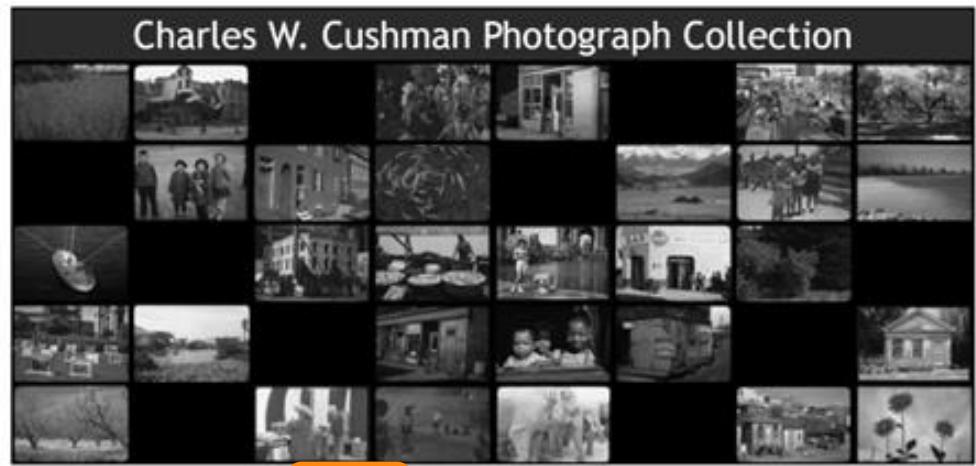
Let us explore the impact of tasks & questions on visualizations

Cushman Dataset

https://www.dropbox.com/s/2y157rrnx9551s6/cushman_collection_data_a1.csv?dl=0



Charles Weever Cushman
(1896-1972)



Charles Weever Cushman, amateur photographer and Indiana University alumnus, bequeathed approximately 14,500 Kodachrome color slides to his alma mater. The photographs in this collection bridge a thirty-two year span from 1938 to 1969, during which time he extensively documented the United States as well as other countries.

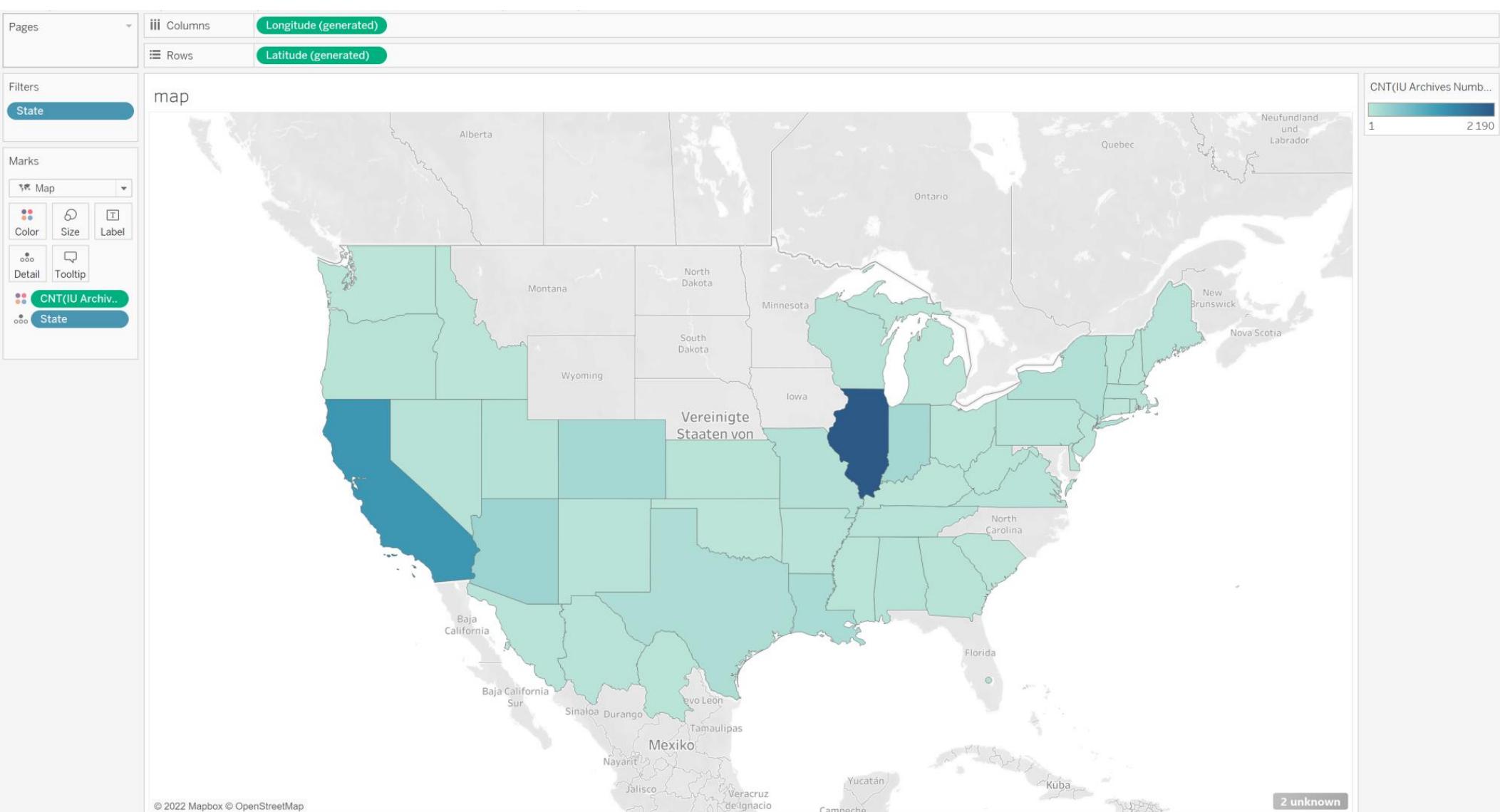
[Indiana University's Digital Library Program](#) and the [Indiana University Archives](#) invite you to explore what Cushman saw. Here you can [view his photographs](#) as well as [read contextual information](#) about Cushman's life and work.

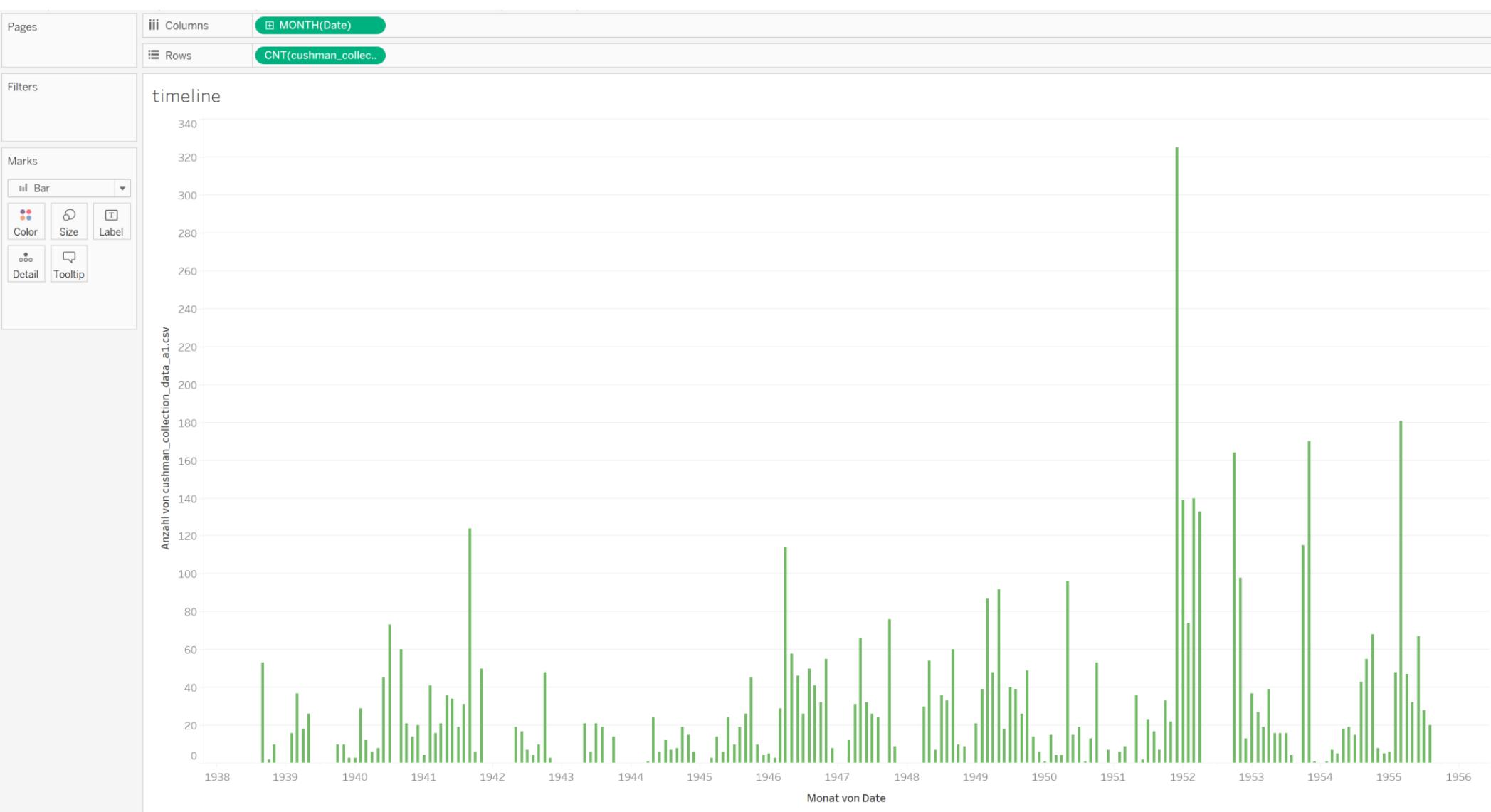
Last updated: Wednesday, May 31, 2017 02:52:48
URL: <http://webapp1.dlib.indiana.edu/cushman/index.jsp>
Collection by: [IU Archives & IU Digital Library Program](#)
Comments: djglib@indiana.edu
[Copyright](#) 1999-2004, The Trustees of [Indiana University](#)
[Accessibility Help](#)

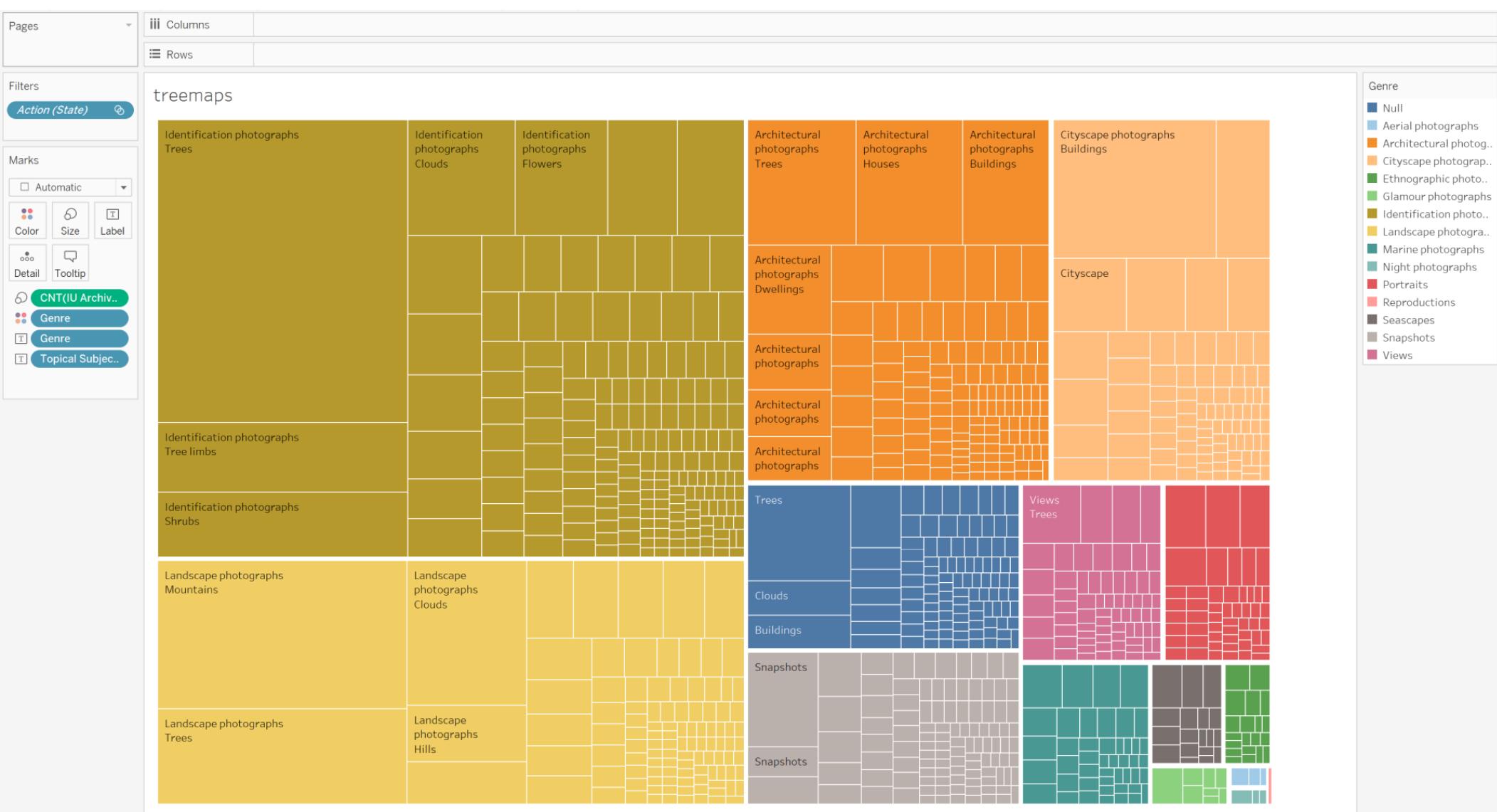


Tableau Examples

- I. What is the **geographic distribution** of these historical photographs?
- II. What is the **temporal distribution** of the photographs?
- III. What is the **distribution of genres** in the collection?







Happiness Dataset

https://vda.univie.ac.at/Teaching/Vis/21s/data/20w_data_happiness_development.csv

Tableau Dataset: 20w_data_happiness_development:

https://vda.univie.ac.at/Teaching/Vis/21s/data/20w_data_happiness_development.csv

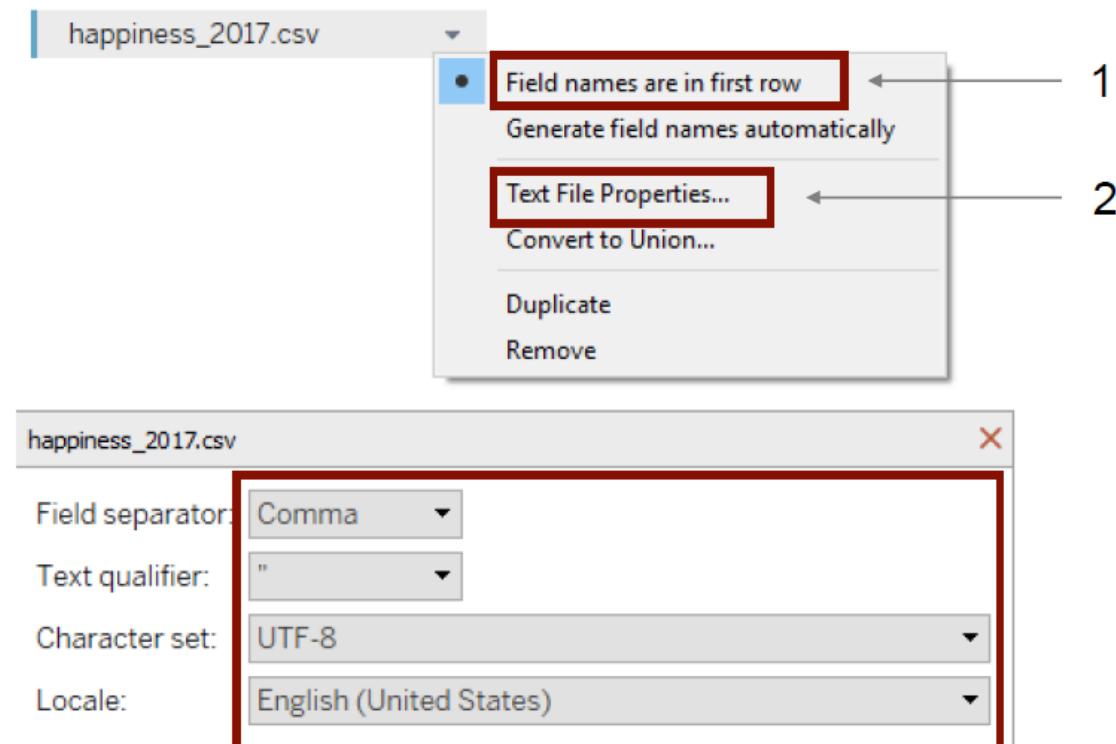


Tableau Example

- Is there a correlation between **Happiness Score** and **Inequality in Education**?
- Using the same visualization add the **Map Reference** to see if there are correlations within continents.

Tableau Example

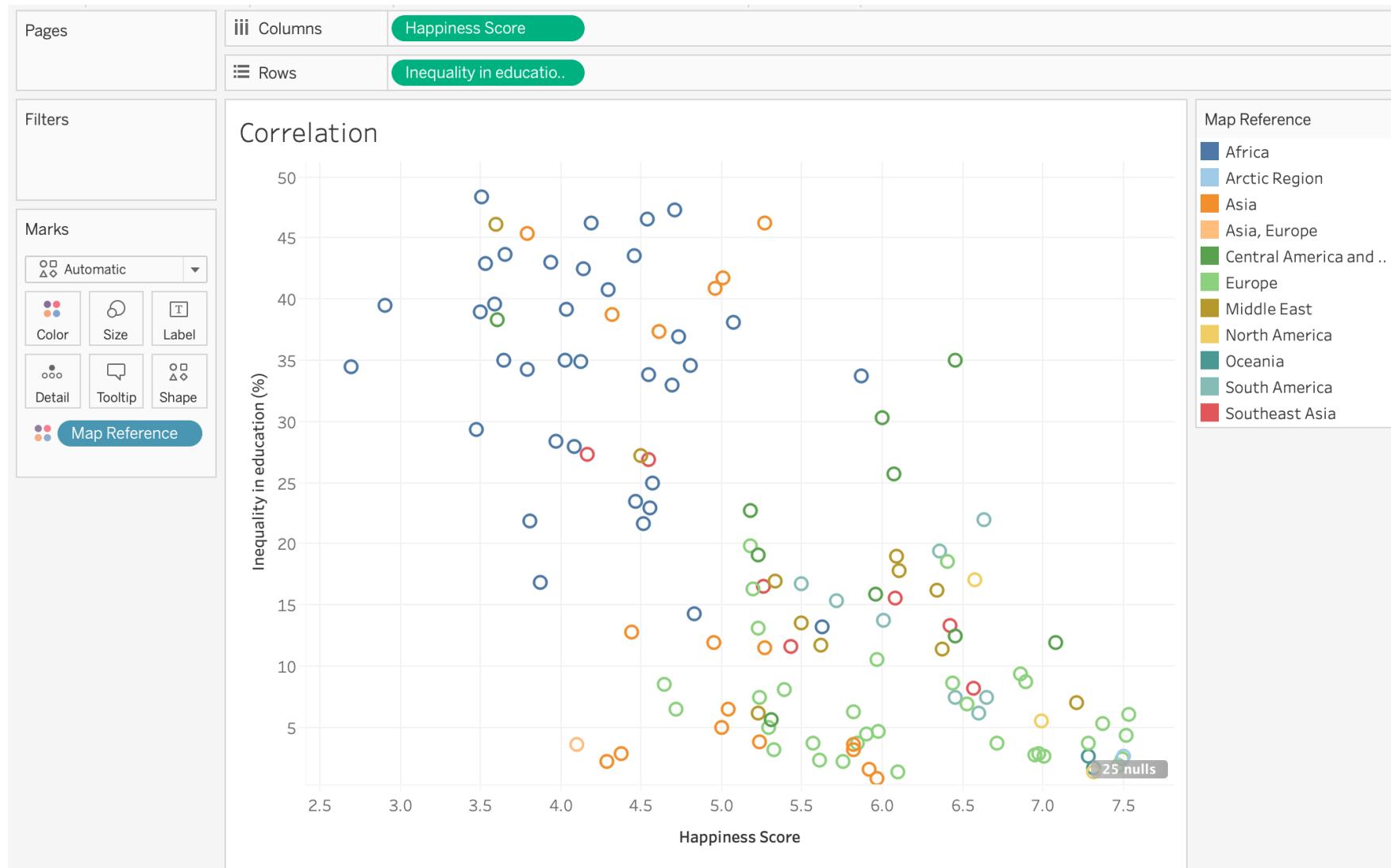


Tableau Example

- How are the two fields (**Happiness Score** and **Inequality in Education**) distributed within each continent?
- Can you see interesting patterns?

Tableau Example

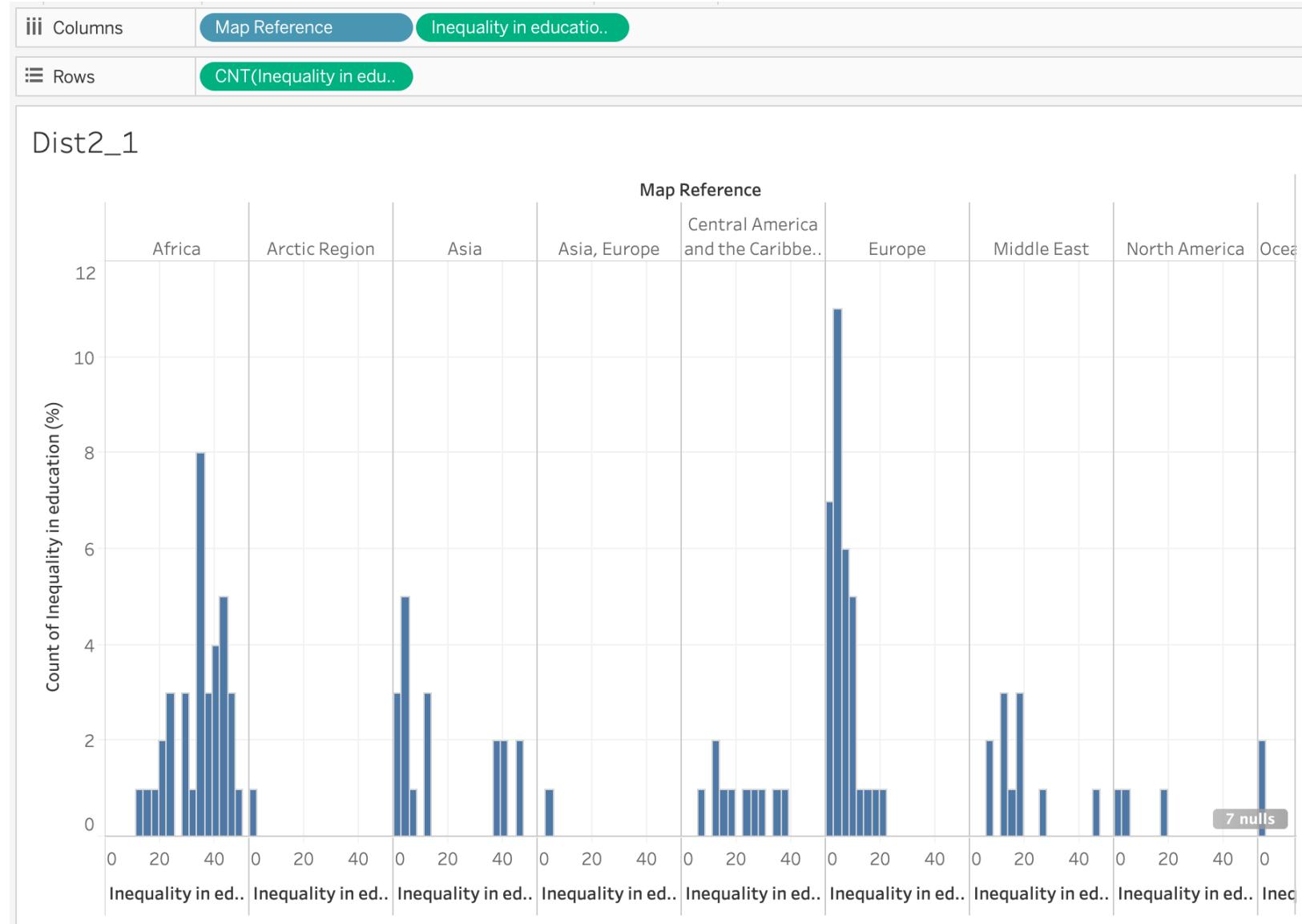


Tableau Example



Tableau Example



Tableau Example

We start with a new set of questions
focusing on language

Tableau Example

Which language is the most used (in terms of people speaking it) **official language?**

Tableau Example

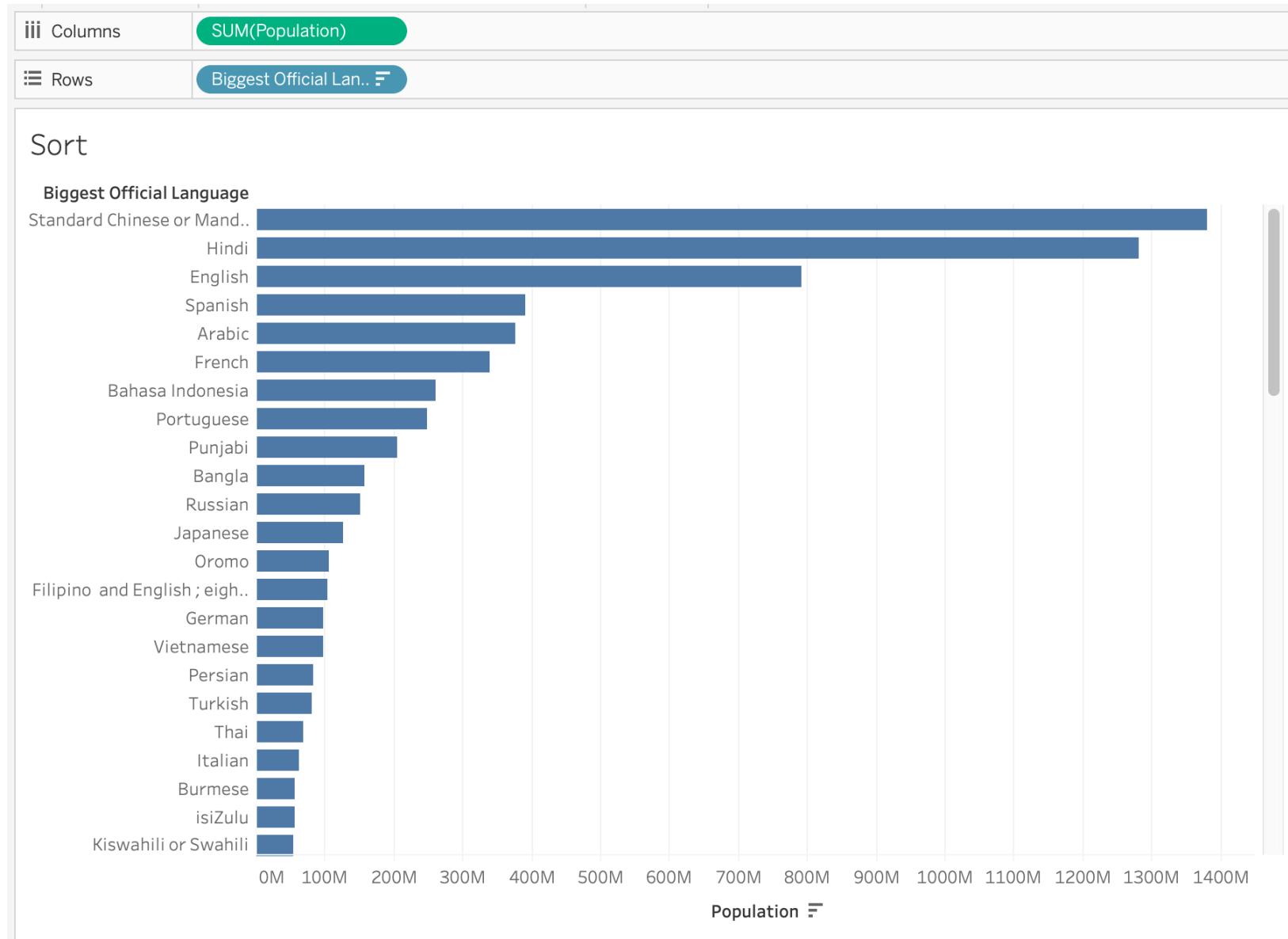


Tableau Example

As we use English in this lecture please tell me what are the top three (by looking at the number of people) countries speaking english?

Tableau Example

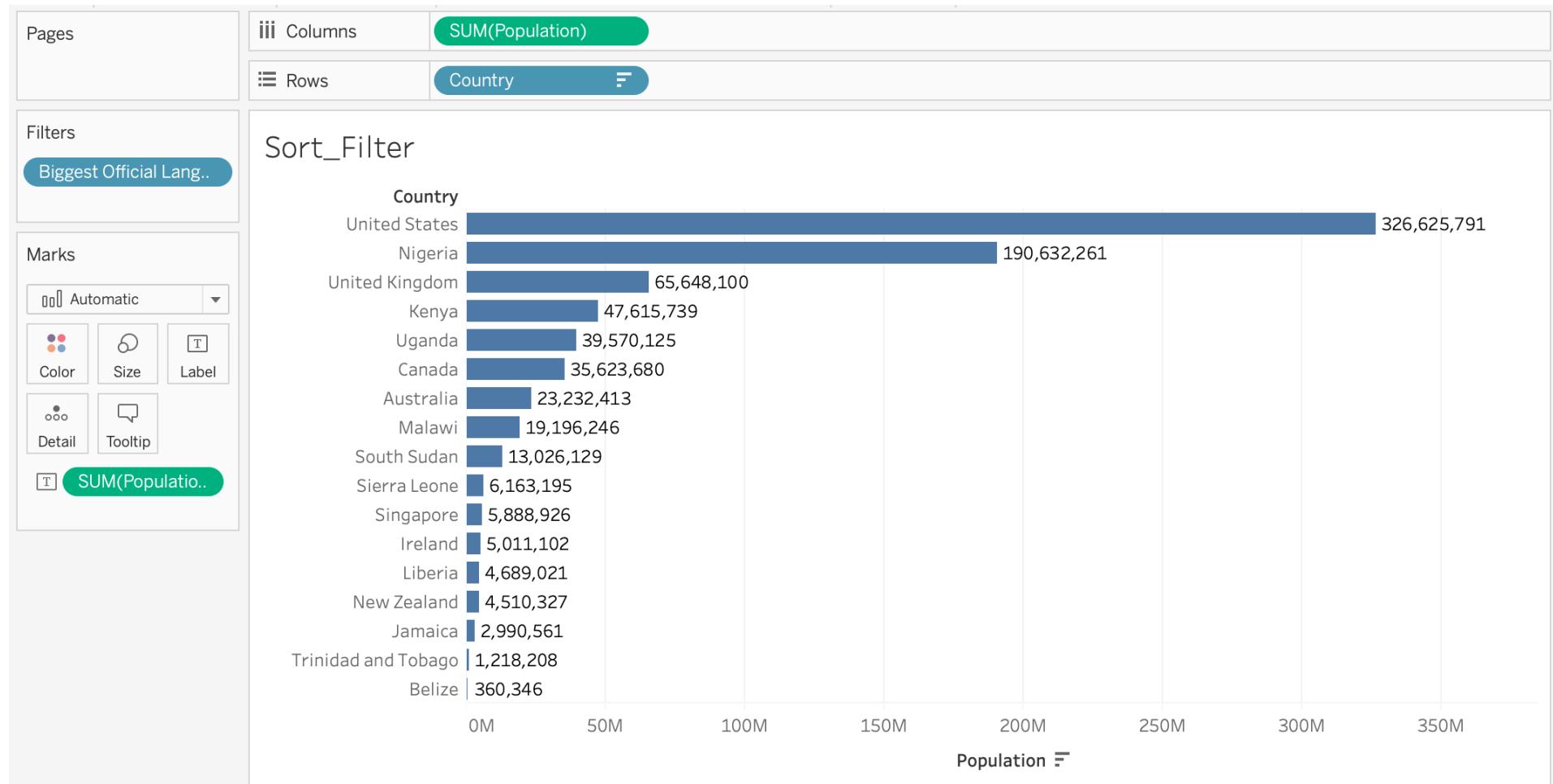


Tableau Example

Given only the top 10 most spoken languages, are there any interesting outliers (**countries**) when looking at the **inequality of education** and **inequality of life expectancy**?

Tableau Example

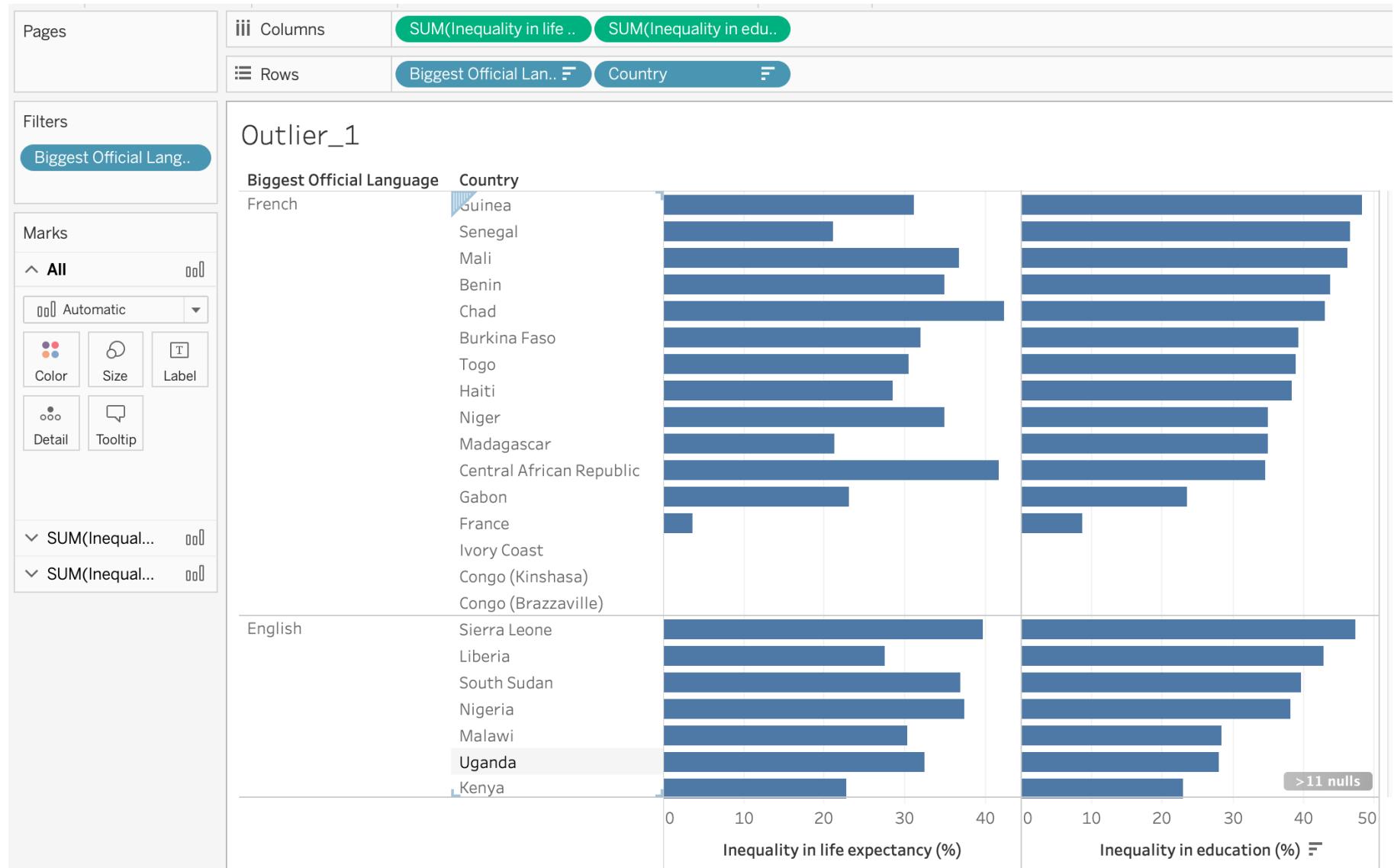


Tableau Example

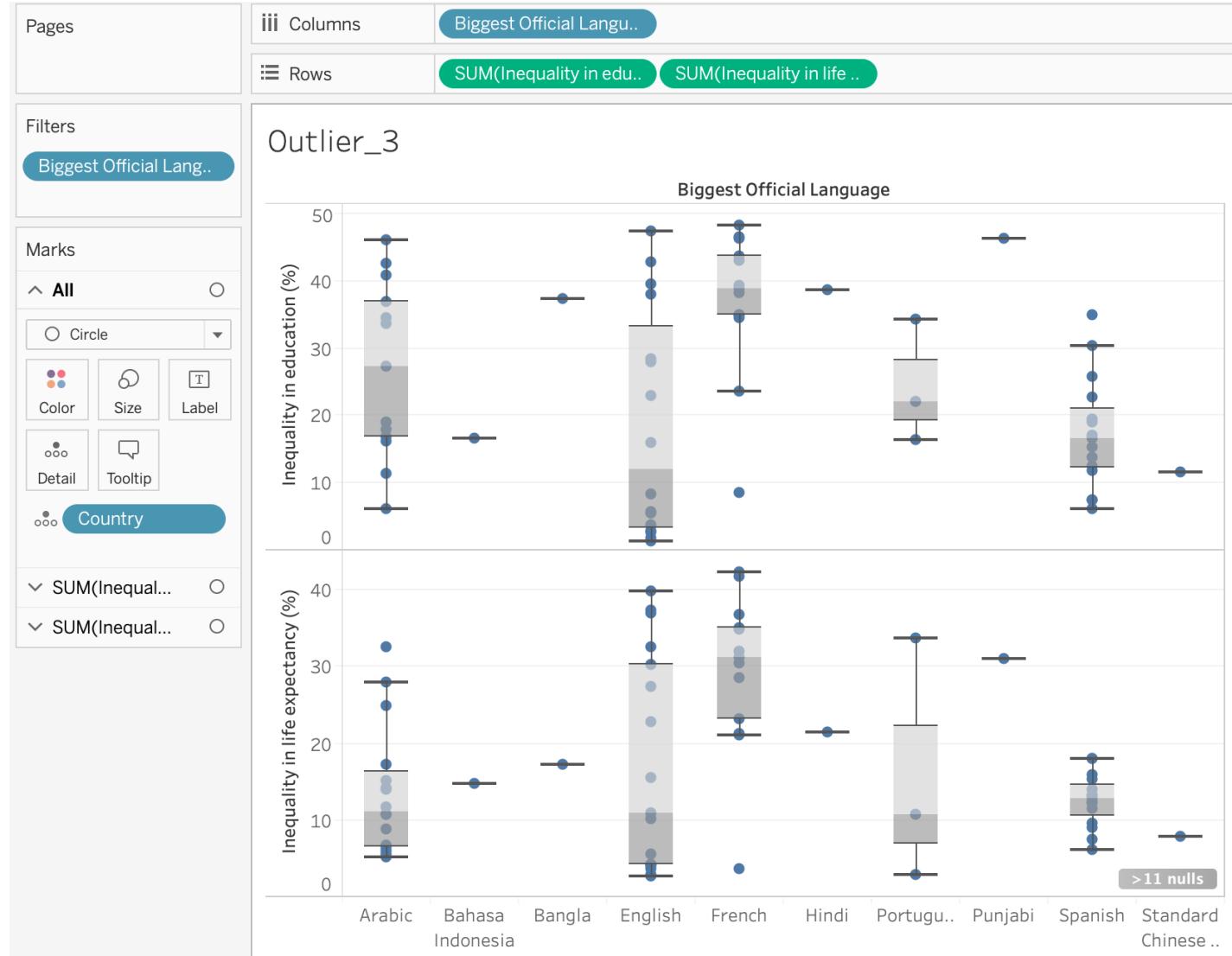


Tableau Example

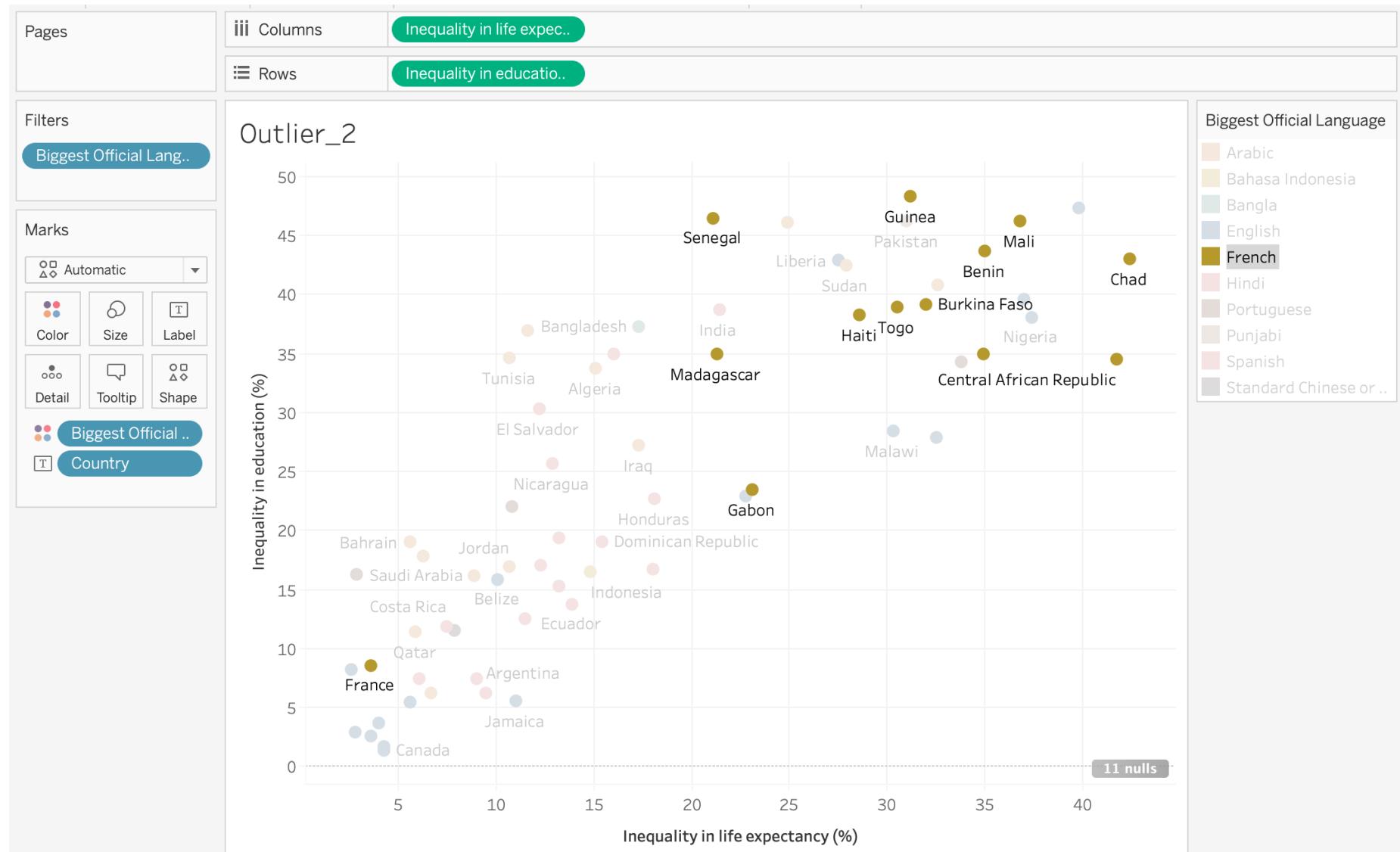
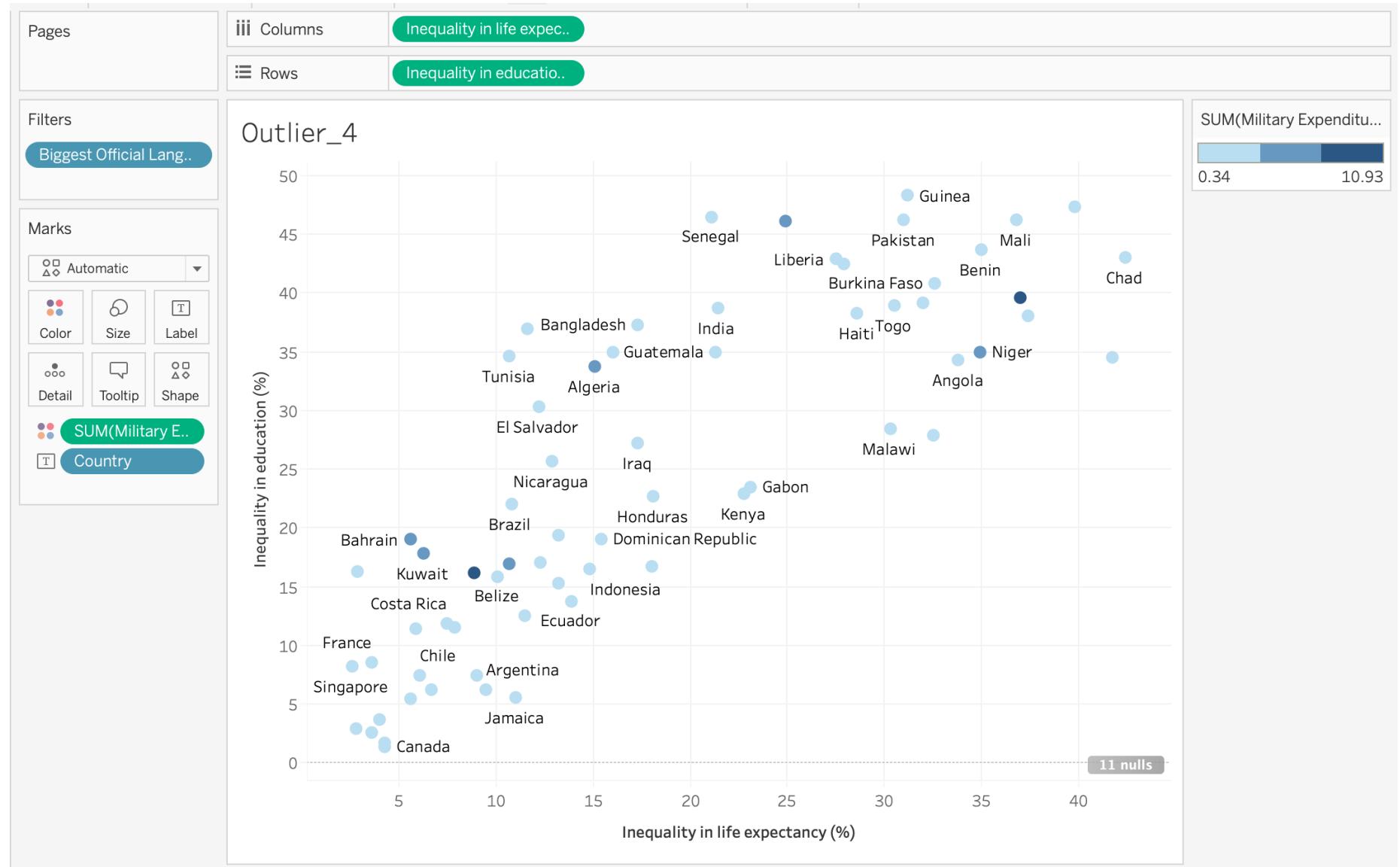


Tableau Example

Using the same fields is there a connection between the top third of the **military expenditures** field and the two inequality (education and life expectancy) fields?

Tableau Example



Apply Your Knowledge:
Reverse Engineering of Users' Tasks

- Build groups of 3
- Choose a VIS interface from the web:

100 Years of Rock <https://www.concerthotels.com/100-years-of-rock/>

Movie Galaxies <http://moviegalaxies.com/>

Inaugural Speeches <http://intuitionanalytics.com/other/inauguralSpeeches/>

Edge Maps <https://mariandoerk.de/edgemaps/demo/>

PUBVIZ <http://pubviz.fhstp.ac.at/>

- Analyze (re-engineer) what the tasks could have been, that guided the VIS development