Interactive Visual Data Analysis

02 – Criteria and influencing factors

Objectives

- What makes a good interactive visual data analysis solution: Learn the criteria to be fulfilled
- What do I need to take into account when designing, developing, or selecting interactive visual data analysis solutions: Learn about the influencing factors based on which decisions can be made

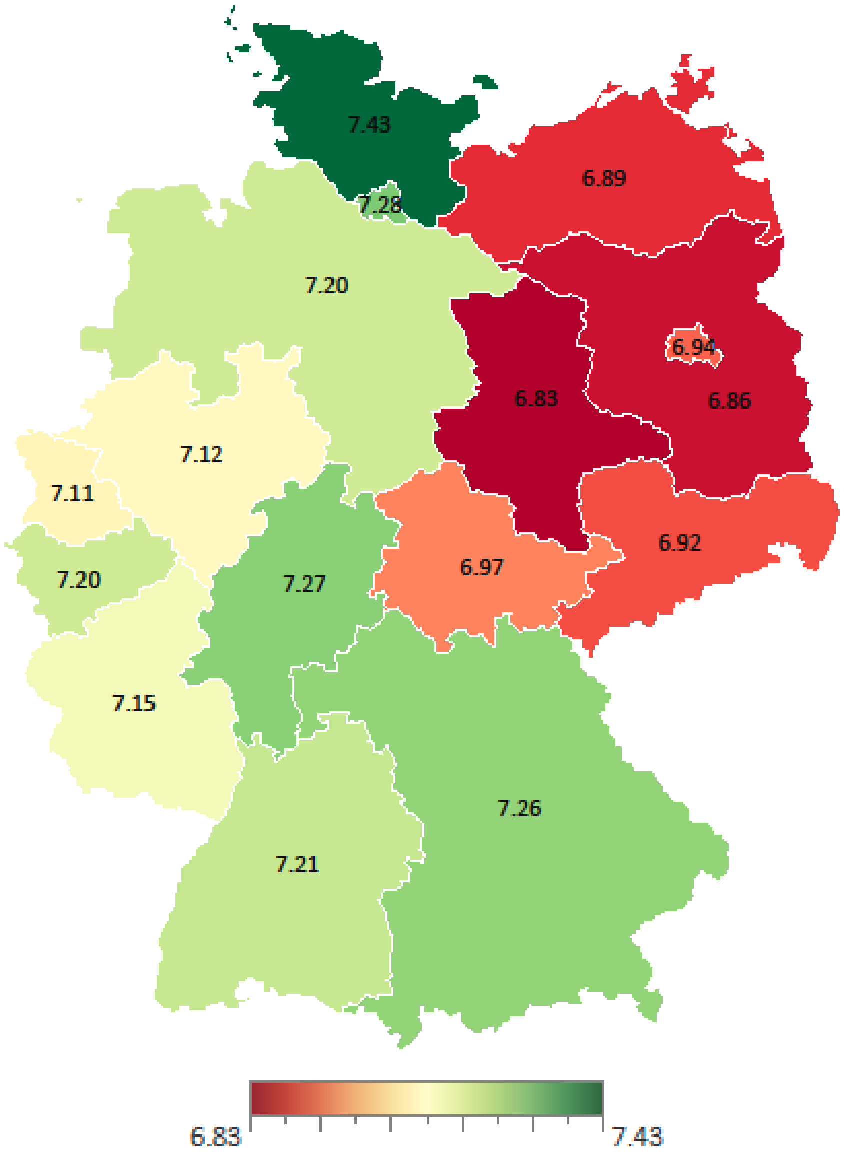
Overview

- Criteria
  - Expressiveness
  - Effectiveness
  - Efficiency
- Influencing factors
  - The subject: Data
  - The objective: Analysis tasks
  - The context: Users and technologies

Motivation

What makes a good visualization?

- Map shows life satisfaction
  - 0 unhappy – 10 very happy
- Think about: What is the first impression conveyed by this visualization and what is truly contained in the data?

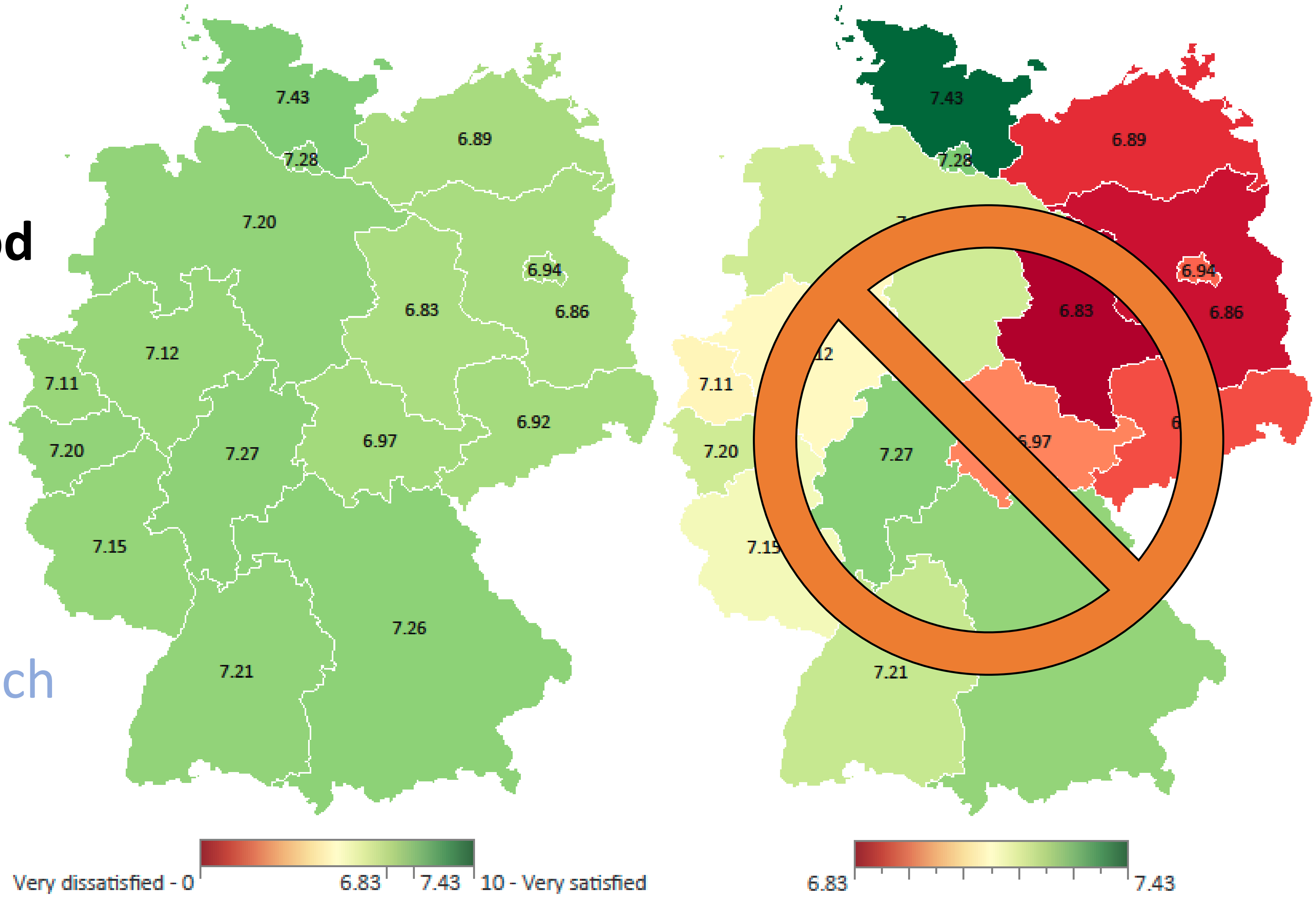


Motivation

What makes a good visualization?

- Same data, different visual encoding

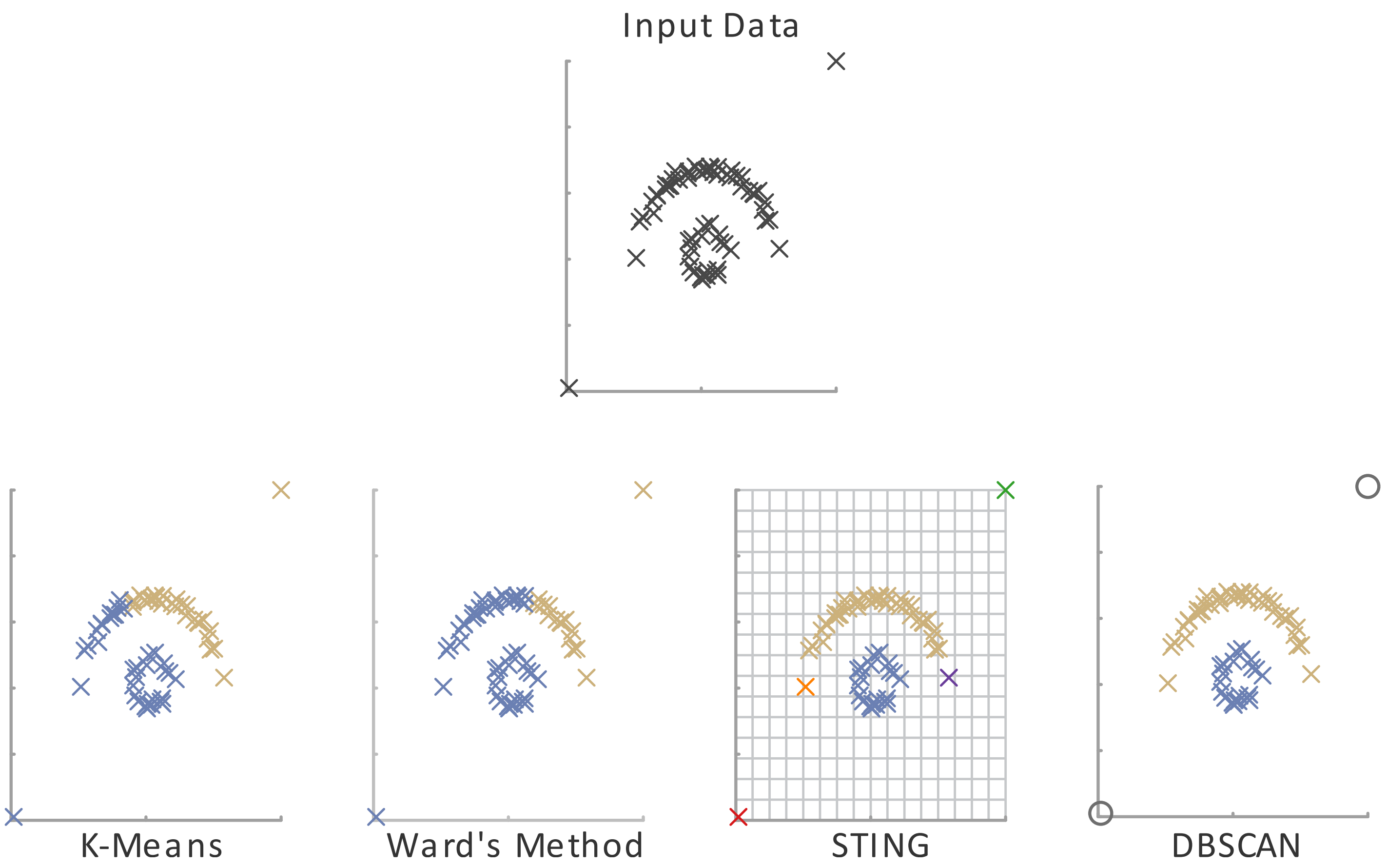
- Think about: Which version is better?



Motivation

Which computational method should be used?

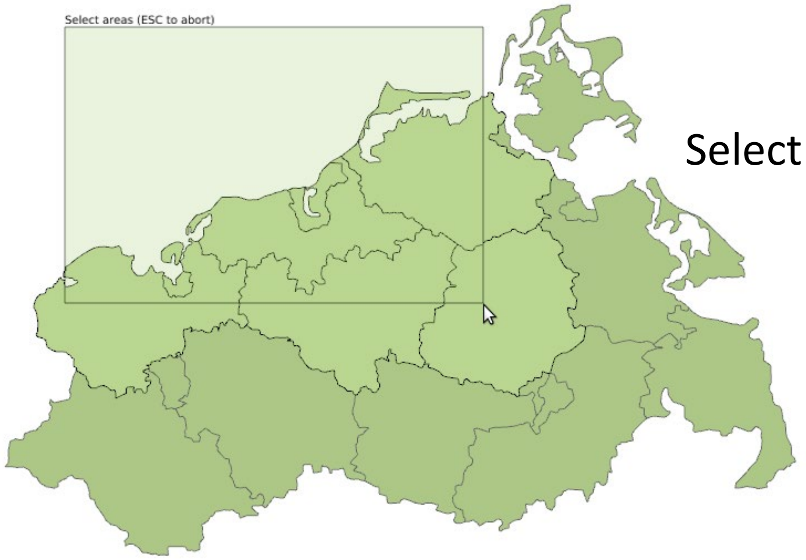
- Consider the plotted data
- Four different algorithms computed data clusters
- Think about: Which algorithm would you choose?



Motivation

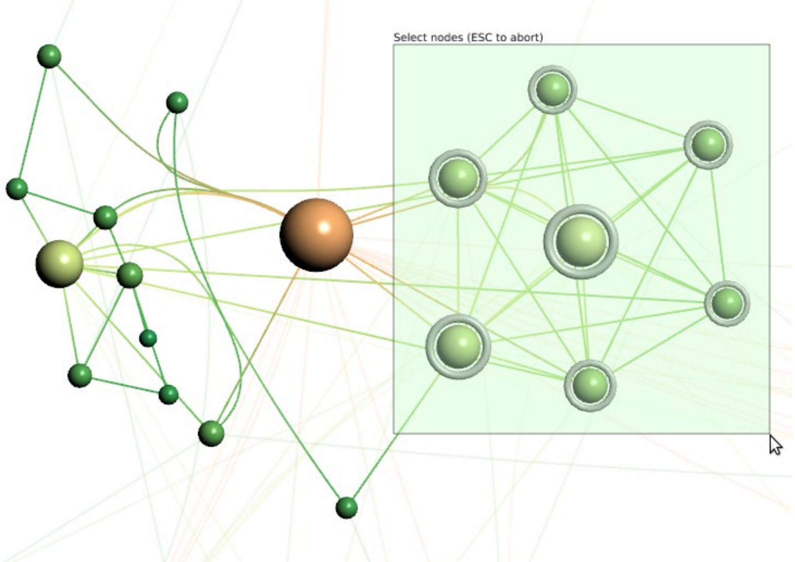
How to interact with the data?

- Two alternative interactions for selecting data items
  - Intersect: Select data items that **intersect** with rectangle
  - Include: Select data items that are fully **included** in rectangle
- Think about: Which variant would you prefer?



Select by intersecting

Select by including



Criteria

We see that some solutions are seemingly better than others. To better characterize what makes a good solution, we define **3** quality **criteria**:

- **Expressiveness**
- **Effectiveness**
- **Efficiency**

Criteria

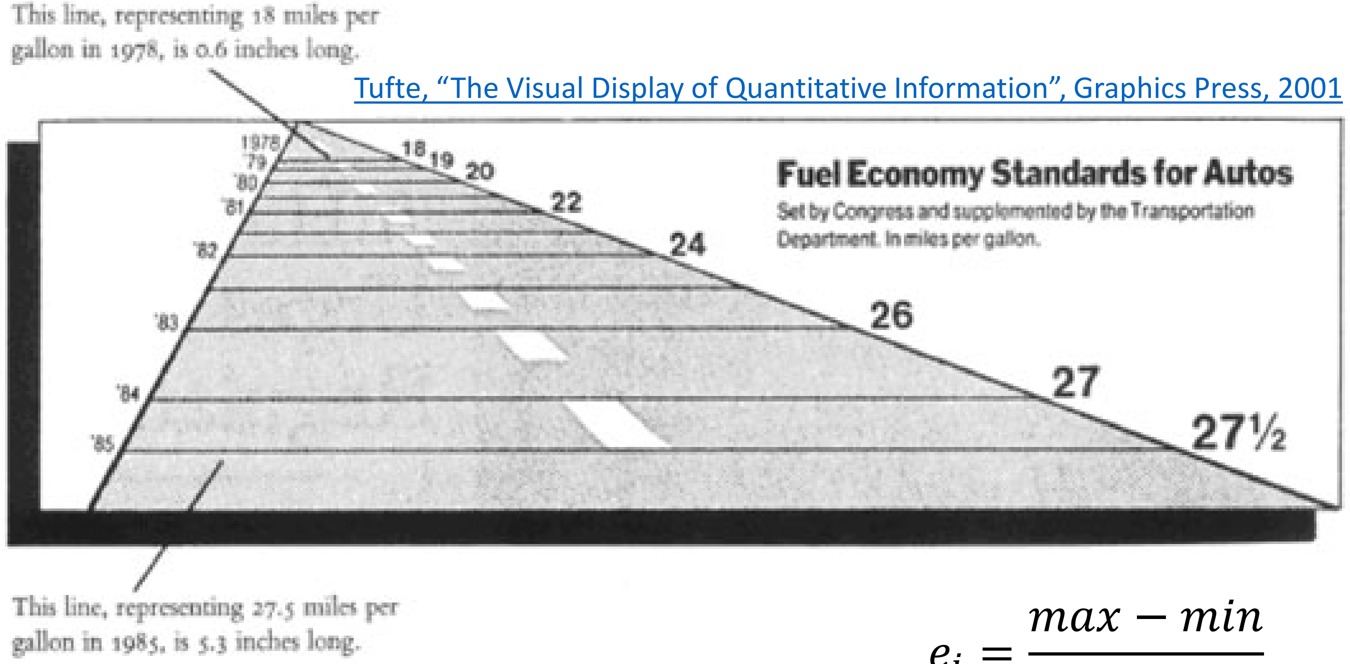
Expressiveness

- Relates to faithfulness
- An interactive visual data analysis solution is **expressive**
  - **if it communicates the relevant information** contained in the data, and only this information. No information is fabricated, no information is withheld.
  - **if it allows users to carry out the actions needed** to acquire the desired information, and only these actions. Users are enabled to do exactly what is necessary for the task at hand.

Criteria

Expressiveness: Lie Factor

- Perceived effect e_v should correspond to the effect in the data e_d
- Lie factor $l = \frac{e_v}{e_d}$
- Example:
  - $e_v = \frac{5.3-0.6}{0.6} = 7.83 \approx 783\%$
  - $e_d = \frac{27.5-18}{18} = 0.53 \approx 53\%$
  - $l = \frac{783\%}{52\%} = 14.77$



[Tufte, "The Visual Display of Quantitative Information", Graphics Press, 2001]

$$e_i = \frac{max - min}{min}$$

Also see https://www.vislies.org

Criteria

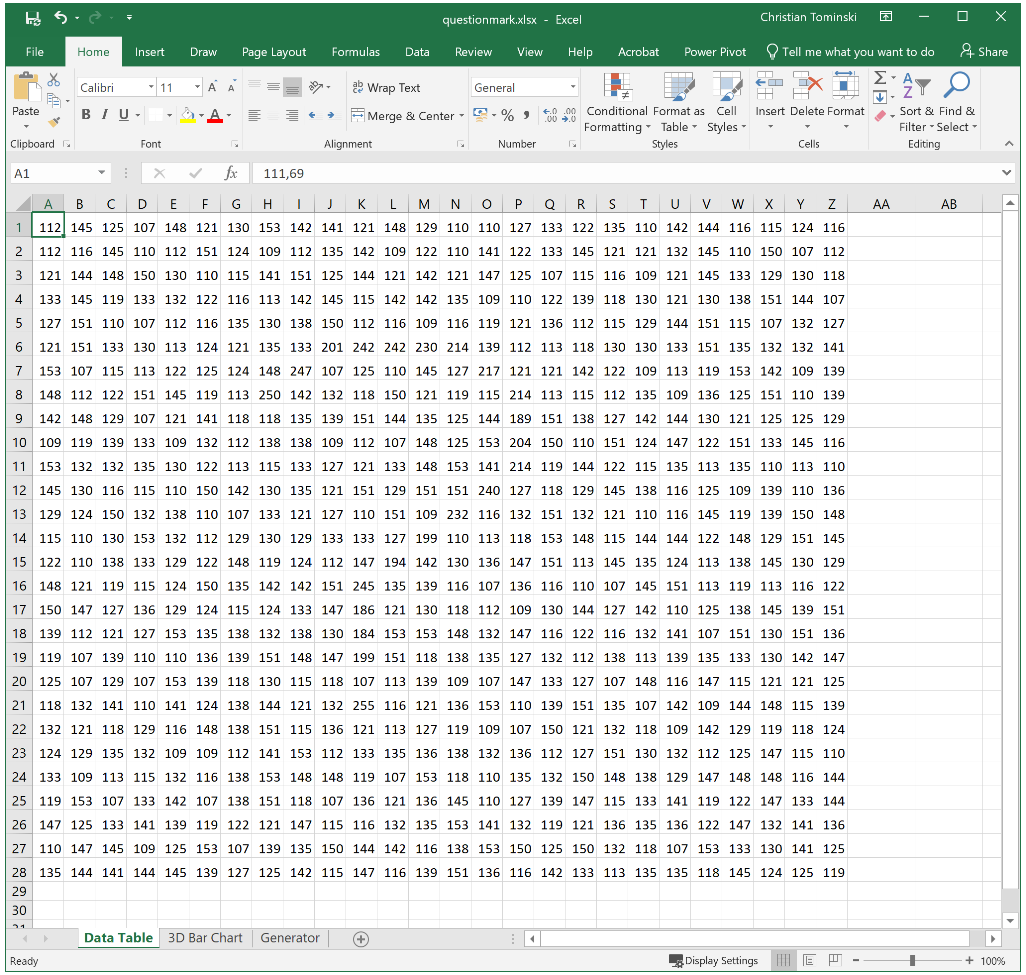
Effectiveness

- Relates to the degree to which users can achieve tasks
- An interactive visual data analysis solution is **effective if it is geared to the human sensory and motor system**
  - How well can users digest the depicted information (visual)
  - How well can users convey their intent to change (interactive)

Criteria

Effectiveness: Illustrating example

- Let's visualize a data table!



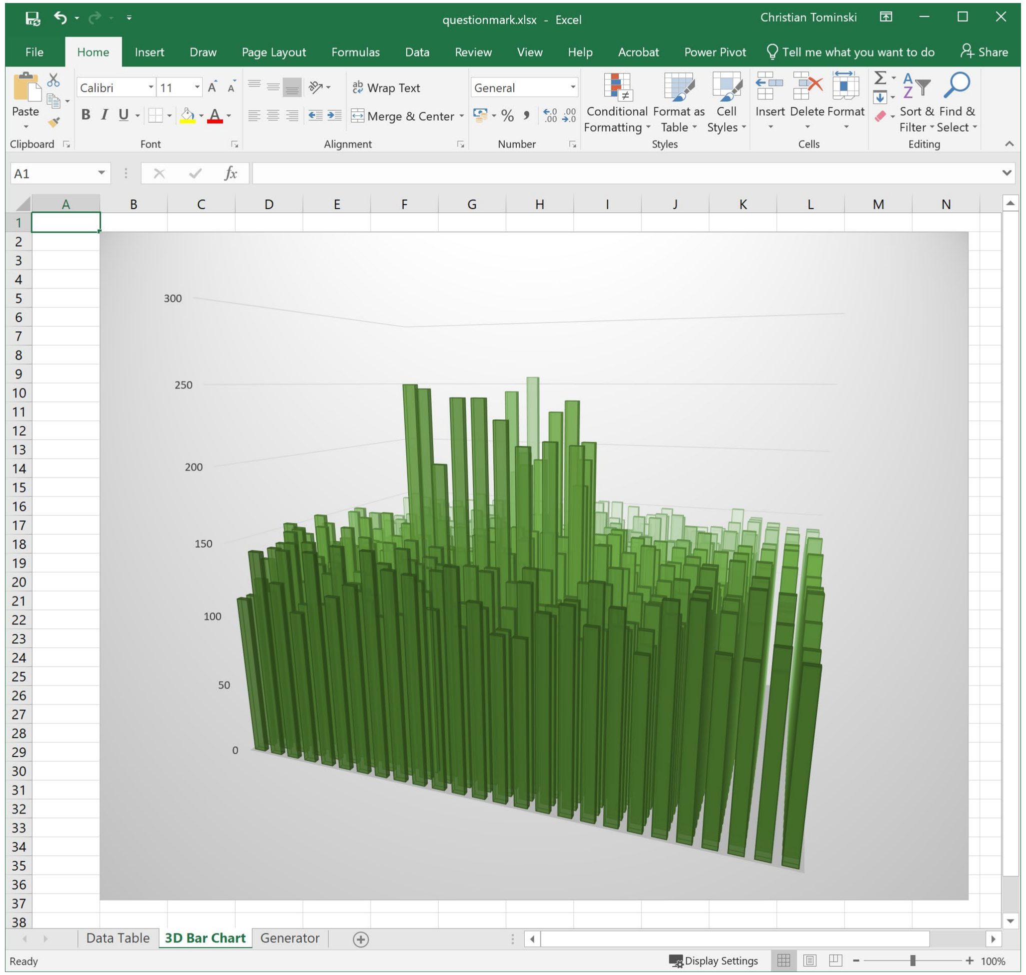
	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
1	112	145	125	107	148	121	130	153	142	141	121	148	129	110	110	127	133	122	135	110	142	144	116	115	124	116
2	112	116	145	110	112	151	124	109	112	135	142	109	122	110	141	122	133	145	121	121	132	145	110	150	107	112
3	121	144	148	150	130	110	115	141	151	125	144	121	142	121	147	125	107	115	116	109	121	145	133	129	130	118
4	133	145	119	133	132	122	116	113	142	145	115	142	142	135	109	110	122	139	118	130	121	130	138	151	144	107
5	127	151	110	107	112	116	135	130	138	150	112	116	109	116	119	121	136	112	115	129	144	151	115	107	132	127
6	121	151	133	130	113	124	121	135	133	201	242	242	230	214	139	112	113	118	130	130	133	151	135	132	132	141
7	153	107	115	113	122	125	124	148	247	107	125	110	145	127	217	121	121	142	122	109	113	119	153	142	109	139
8	148	112	122	151	145	119	113	250	142	132	118	150	121	119	115	214	113	115	112	135	109	136	125	151	110	139
9	142	148	129	107	121	141	118	118	135	139	151	144	135	125	144	189	151	138	127	142	144	130	121	125	125	129
10	109	119	139	133	109	132	112	138	138	109	112	107	148	125	153	204	150	110	151	124	147	122	151	133	145	116
11	153	132	132	135	130	122	113	115	133	127	121	133	148	153	141	214	119	144	122	115	135	113	135	110	113	110
12	145	130	116	115	110	150	142	130	135	121	151	129	151	151	240	127	118	129	145	138	116	125	109	139	110	136
13	129	124	150	132	138	110	107	133	121	127	110	151	109	232	116	132	151	132	121	110	116	145	119	139	150	148
14	115	110	130	153	132	112	129	130	129	133	133	127	199	110	113	118	153	148	115	144	144	122	148	129	151	145
15	122	110	138	133	129	122	148	119	124	112	147	194	142	130	136	147	151	113	145	135	124	113	138	145	130	129
16	148	121	119	115	124	150	135	142	142	151	245	135	139	116	107	136	116	110	107	145	151	113	119	113	116	122
17	150	147	127	136	129	124	115	124	133	147	186	121	130	118	112	109	130	144	127	142	110	125	138	145	139	151
18	139	112	121	127	153	135	138	132	138	130	184	153	153	148	132	147	116	122	116	132	141	107	151	130	151	136
19	119	107	139	110	110	136	139	151	148	147	199	151	118	138	135	127	132	112	138	113	139	135	133	130	142	147
20	125	107	129	107	153	139	118	130	115	118	107	113	139	109	107	147	133	127	107	148	116	147	115	121	121	125
21	118	132	141	110	141	124	138	144	121	132	255	116	121	136	153	110	139	151	135	107	142	109	144	148	115	139
22	132	121	118	129	116	148	138	151	115	136	121	113	127	119	109	107	150	121	132	118	109	142	129	119	118	124
23	124	129	135	132	109	109	112	141	153	112	133	135	136	138	132	136	112	127	151	130	132	112	125	147	115	110
24	133	109	113	115	132	116	138	153	148	148	119	107	153	118	110	135	132	150	148	138	129	147	148	148	116	144
25	119	153	107	133	142	107	138	151	118	107	136	121	136	145	110	127	139	147	115	133	141	119	122	147	133	144
26	147	125	133	141	139	119	122	121	147	115	116	132	135	153	141	132	119	121	136	135	136	122	147	132	141	136
27	110	147	145	109	125	153	107	139	135	150	144	142	116	138	153	150	125	150	132	118	107	153	133	130	141	125
28	135	144	141	144	145	139	127	125	142	115	147	116	139	151	136	116	142	133	113	135	135	118	145	124	125	119

Criteria

Effectiveness: Illustrating example

- Think about:
  - Is this visualization expressive?
  - Is this visualization effective?

- Arguably, yes, it is expressive, because all data are represented visually?
- But, it is not effective, because we can hardly interpret the data.

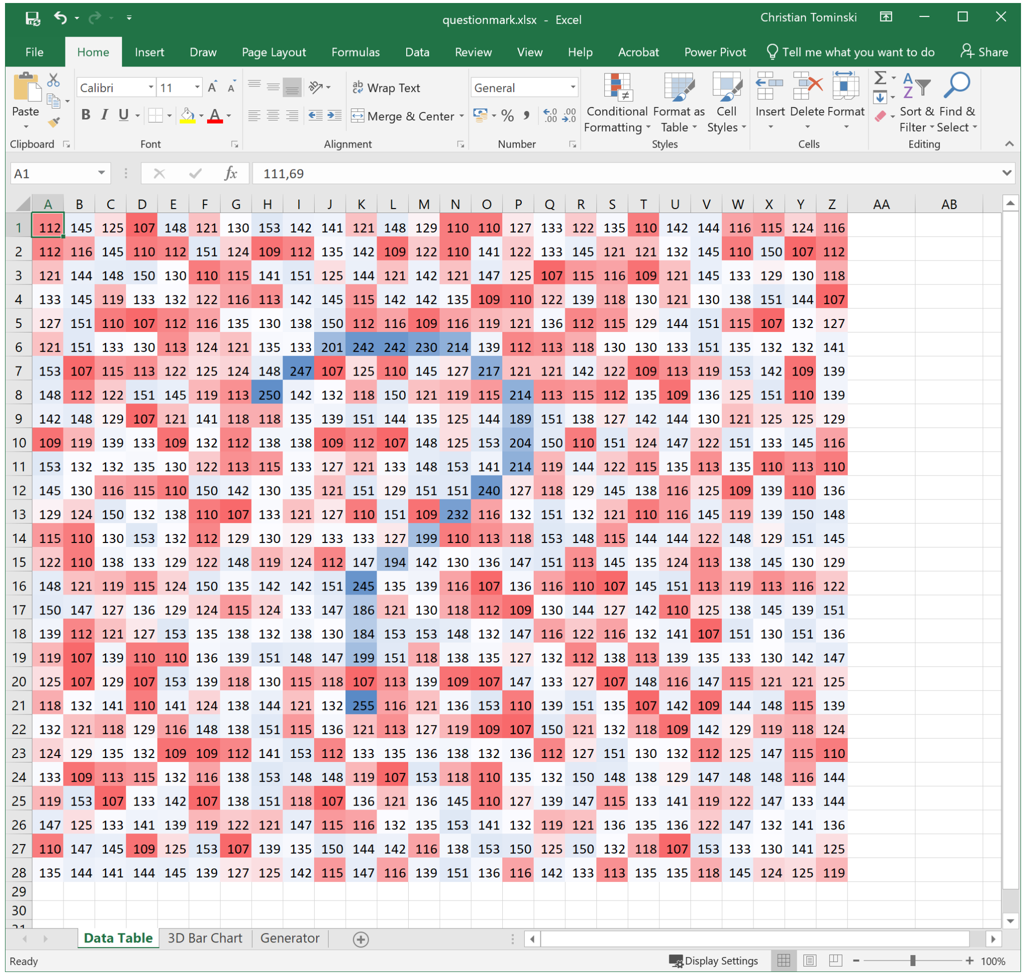


Criteria

Effectiveness: Illustrating example

- Think about:
  - Is this visualization expressive?
  - Is this visualization effective?

- Yes, expressive and effective! All data are depicted faithfully and we can easily discern a pattern from the otherwise noisy data.



	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
1	112	145	125	107	148	121	130	153	142	141	121	148	129	110	110	127	133	122	135	110	142	144	116	115	124	116
2	112	116	145	110	112	151	124	109	112	135	142	109	122	110	141	122	133	145	121	121	132	145	110	150	107	112
3	121	144	148	150	130	110	115	141	151	125	144	121	142	121	147	125	107	115	116	109	121	145	133	129	130	118
4	133	145	119	133	132	122	116	113	142	145	115	142	142	135	109	110	122	139	118	130	121	130	138	151	144	107
5	127	151	110	107	112	116	135	130	138	150	112	116	109	116	119	121	136	112	115	129	144	151	115	107	132	127
6	121	151	133	130	113	124	121	135	133	201	242	242	230	214	139	112	113	118	130	130	133	151	135	132	132	141
7	153	107	115	113	122	125	124	148	247	107	125	110	145	127	217	121	121	142	122	109	113	119	153	142	109	139
8	148	112	122	151	145	119	113	250	142	132	118	150	121	119	115	214	113	115	112	135	109	136	125	151	110	139
9	142	148	129	107	121	141	118	118	135	139	151	144	135	125	144	189	151	138	127	142	144	130	121	125	125	129
10	109	119	139	133	109	132	112	138	138	109	112	107	148	125	153	204	150	110	151	124	147	122	151	133	145	116
11	153	132	132	135	130	122	113	115	133	127	121	133	148	153	141	214	119	144	122	115	135	113	135	110	113	110
12	145	130	116	115	110	150	142	130	135	121	151	129	151	151	240	127	118	129	145	138	116	125	109	139	110	136
13	129	124	150	132	138	110	107	133	121	127	110	151	109	232	116	132	151	132	121	110	116	145	119	139	150	148
14	115	110	130	153	132	112	129	130	129	133	133	127	199	110	113	118	153	148	115	144	144	122	148	129	151	145
15	122	110	138	133	129	122	148	119	124	112	147	194	142	130	136	147	151	113	145	135	124	113	138	145	130	129
16	148	121	119	115	124	150	135	142	142	151	245	135	139	116	107	136	116	110	107	145	151	113	119	113	116	122
17	150	147	127	136	129	124	115	124	133	147	186	121	130	118	112	109	130	144	127	142	110	125	138	145	139	151
18	139	112	121	127	153	135	138	132	138	130	184	153	153	148	132	147	116	122	116	132	141	107	151	130	151	136
19	119	107	139	110	110	136	139	151	148	147	199	151	118	138	135	127	132	112	138	113	139	135	133	130	142	147
20	125	107	129	107	153	139	118	130	115	118	107	113	139	109	107	147	133	127	107	148	116	147	115	121	121	125
21	118	132	141	110	141	124	138	144	121	132	255	116	121	136	153	110	139	151	135	107	142	109	144	148	115	139
22	132	121	118	129	116	148	138	151	115	136	121	113	127	119	109	107	150	121	132	118	109	142	129	119	118	124
23	124	129	135	132	109	109	112	141	153	112	133	135	136	138	132	136	112	127	151	130	132	112	125	147	115	110
24	133	109	113	115	132	116	138	153	148	148	119	107	153	118	110	135	132	150	148	138	129	147	148	148	116	144
25	119	153	107	133	142	107	138	151	118	107	136	121	136	145	110	127	139	147	115	133	141	119	122	147	133	144
26	147	125	133	141	139	119	122	121	147	115	116	132	135	153	141	132	119	121	136	135	136	122	147	132	141	136
27	110	147	145	109	125	153	107	139	135	150	144	142	116	138	153	150	125	150	132	118	107	153	133	130	141	125
28	135	144	141	144	145	139	127	125	142	115	147	116	139	151	136	116	142	133	113	135	135	118	145	124	125	119

Criteria

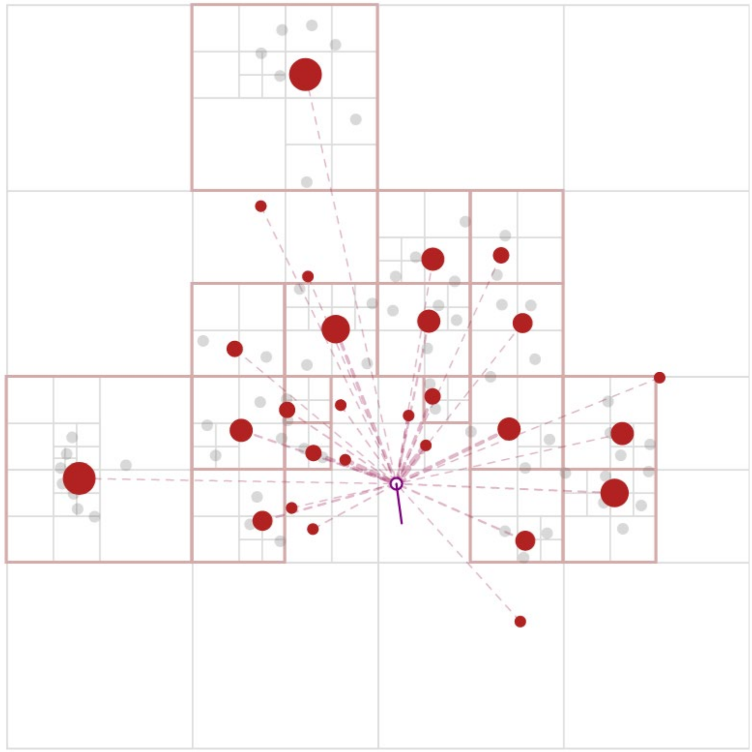
Efficiency

- Relates to the balance of benefits and costs
- An interactive visual data analysis solution is **efficient if its benefits outweigh the costs** in terms of computational and human resources
  - How much time and memory does it take?
  - How much display space is needed?
  - How much eye movement is required?
  - How long does it take to decipher the visual representation?
  - How difficult is it to interact?

Criteria

Illustrating example: Efficiency

- Force-directed layout for graph visualization
  - Attracting forces between connected nodes
  - Repelling forces between all nodes
- Computational costs:
  - Naïve $O(n^2)$
  - Barnes-Hut $O(n \log n)$



https://jheer.github.io/barnes-hut/

Criteria

Criterion	Concern
Expressiveness	Faithfully map data and tasks
Effectiveness	Enable users to accomplish task
Efficiency	Balance benefits and costs

Discussion

- Criteria capture conceptual ideas, and as such, are difficult to evaluate formally
- Some aspects easy to quantify: Computation time, display space, …
- Others hard to pin down:
  - Expressiveness: What is the “relevant” information to be depicted faithfully?
  - Effectiveness: How well can a complex data analysis system be operated?
  - Efficiency: What are the benefits of using an interactive visual solution?
- Nonetheless, the 3 criteria provide essential guidelines for designing interactive visual data analysis solutions!

Criteria

Summary

“Above all else **show the data**.

— Tufte, 1983

Influencing factors

When designing and developing interactive visual data analysis solutions, we need to take into account **4** key **influencing factors**:

- **Subject** of the data analysis: **Data** (what)
- **Objective** of the data analysis: **Analysis tasks** (why)
- **Context** of the data analysis: **Users and technologies** (who, where, when)

- **Why do they matter?**
  - Depending on the influencing factors, different design choices lead to different degrees of fulfillment of the 3 quality criteria!

Influencing factor: Data

With respect to the **data**, the following data characteristics are relevant

- Data domain (data scale and data type)
- Data structure
- Data space
- Data size
- Data scope
- Meta-data

Influencing factor: Data

Data domain

- Set of values that a datum (or data value) can assume
- Differentiate between different **data scales** and **data types**
- **Data scale**: How are the data scaled?
  - Qualitative data: Nominal and ordinal data
  - Quantitative data: Discrete and continuous data
- **Data type**: How is a datum composed of components?
  - Scalar, vector, tensor

Here, the notion of *data type* is used slightly differently from its use in programming languages.

Influencing factor: Data

Qualitative data

- Nominal data
  - Values for which only the **equality relation** = is defined
  - Also called **categorical** data
  - Example: Names *{John, Mike, Lisa, Monica}*
  - Think about: Give another example of nominal data!
- Ordinal data
  - Values for which an **order relation** < is defined in addition to equality
  - Example: Age groups *{children, youths, adults, elders}*
  - Think about: Give another example of ordinal data!

Influencing factor: Data

Quantitative data

- Discrete data
  - Numeric values whose domain can be equated to the **whole numbers**
  - Countable and **distance** between any two data values is defined
  - Example: Number of people visiting a doctor
- Continuous data
  - Numeric values whose domain can be equated to the **real numbers**
  - **Uncountable**, distance is defined, and **interpolation** is possible
  - Example: Temperature measurements
  - Think about: Why is the uncountable property of relevance here?

Influencing factor: Data

Data scale

- Summary of data scales and possible operations

Operations	Qualitative Data		Quantitative Data	
	Nominal	Ordinal	Discrete	Continuous
Equality	●	●	●	●
Order		●	●	●
Distance			●	●
Interpolation				●
(Count)	●	●	●	

Influencing factor: Data

Data type

In this lecture series, we will focus on scalar data.

- **Scalar:**
  - **Single value**
  - **Example: temperature value**

$$36$$

- Vector:
  - Magnitude and direction
  - Example: 3D vector flow vector

$$\begin{pmatrix}1\\2\\3\end{pmatrix}$$

- Tensor:
  - General multilinear transformations
  - Example: Rank-2, 3D stress tensor

$$\begin{bmatrix}4 & 5 & 6\\8 & 10 & 12\\12 & 15 & 18\end{bmatrix}$$

Influencing factor: Data

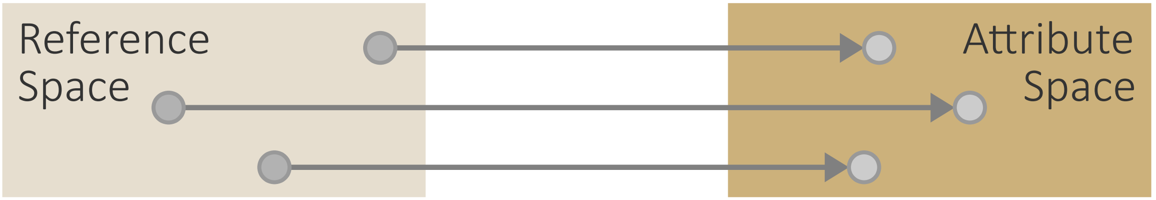
Data structure

- Raw data often “unstructured”
- Data analysis typically requires transforming raw data to a suitable structure
  - **Tabular data: Rows are tuples, columns are variables**
  - Hierarchical data: Parent-child relations, e.g., for different levels of detail
  - Graph data: General model for entities and relations between them
  - (Grid-structured data: Data aligned on different types of grids, for 3D data)

Influencing factor: Data

Data space

- Spanned by variables
  - **Independent** variables: **Dimensions** of the space where data have been collected, observed, or simulated
  - **Dependent** variables: **Attributes** of what has been collected, observed, or simulated
- Describes a functional dependency $f: (D_1 \times D_2 \times \cdots \times D_n) \rightarrow (A_1 \times A_2 \times \cdots \times A_m)$

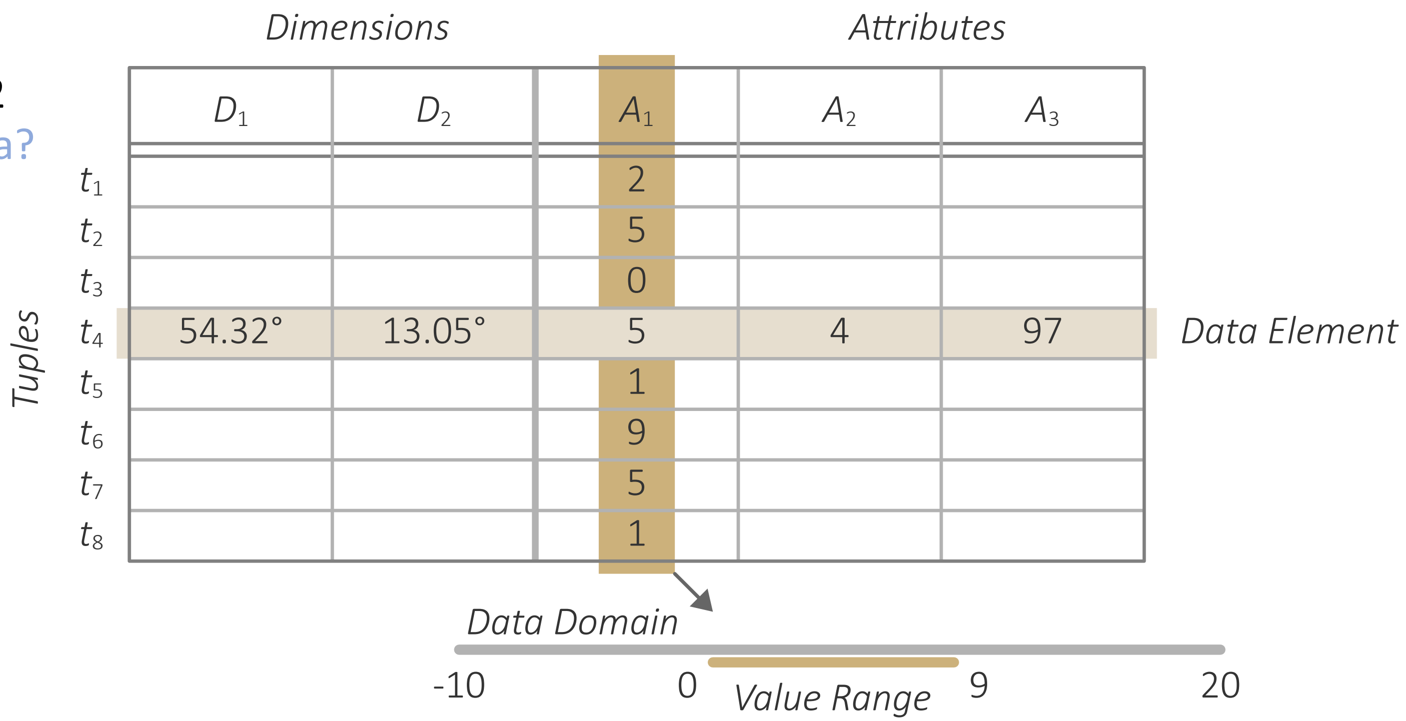


A visualization usually has the goal to make the functional dependencies in the data visible

Influencing factor: Data

Data size

- *n* - Number of dimensions
  - n-dimensional data
    - 1D data (e.g., time-series): n = 1
    - 2D data (e.g. geo-spatial data): n = 2
    - Think about: Any idea about 3D data?
    - Multidimensional data: $n > 3$
- *m* - Number of attributes
  - m-variate data
    - Univariate data: m = 1
    - Bivariate data: m = 2
    - Multivariate data: m > 2
- *k* - Number of tuples
  - Small data: k < 1000
  - Big data: k > 100.000



Influencing factor: Data

Data size

- Obviously, the larger *m*, *n*, and *k* are, the more difficult it will be to analyze the data
  - Requires additional effort for transforming, rendering, and interpreting the data
- What are considered large data, depends on the application context
  - Sometimes a graph with hundreds of nodes is large, but graphs exist with millions of nodes!
- Think about: Hardware barriers and data size

Influencing factor: Data

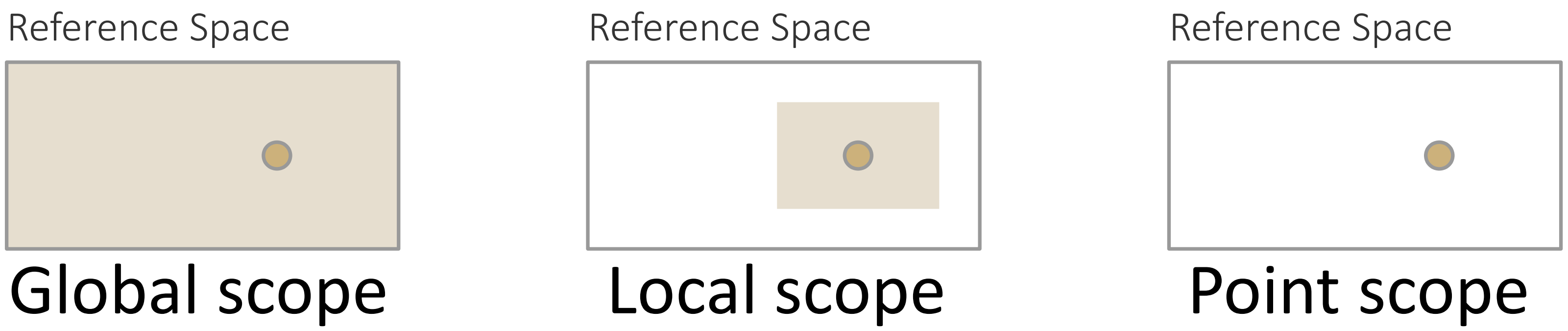
Data size

- Interactive visual data analysis solutions usually target **Big Data**
- Big Data characterized by five Vs
  - **Volume:** Many data points
  - **Velocity:** New data arrive at high speed
  - **Variety:** Many dimensions and attributes covering many different aspects
  - **Variability:** Data must be accessible to diverse user groups
  - **Value:** Monetary costs and benefits of data analysis are considerable

Influencing factor: Data

Data scope

- Characterizes the range in which the data are valid around a point of observation
- Usually cannot be inferred, must be given with the data description/context



Influencing factor: Data

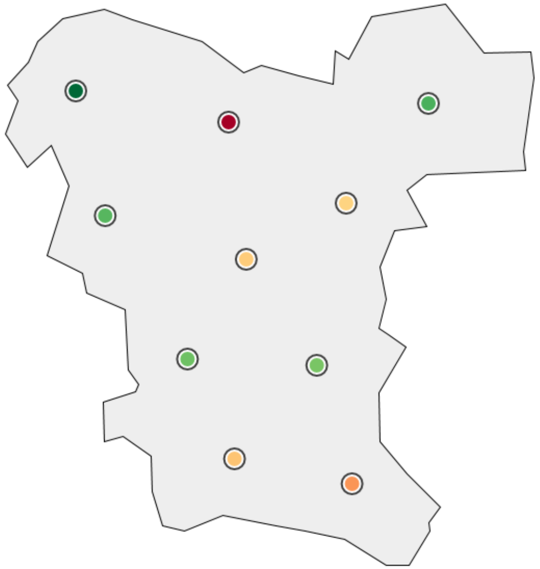
Data scope

- For geo-spatial data, according to Tobler's (1970) first law, the scope is local, because data measured at proximal locations tend to be correlated
- So, (geo-)spatial data are likely to have local scope
- Think about: How to represent local scope visually?

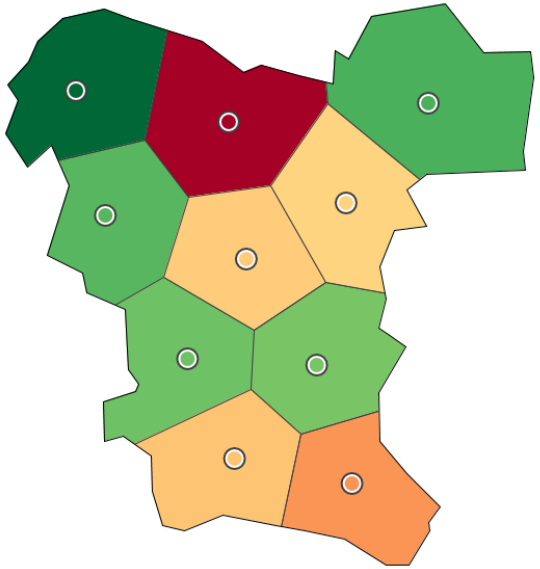
Influencing factor: Data

Data scope

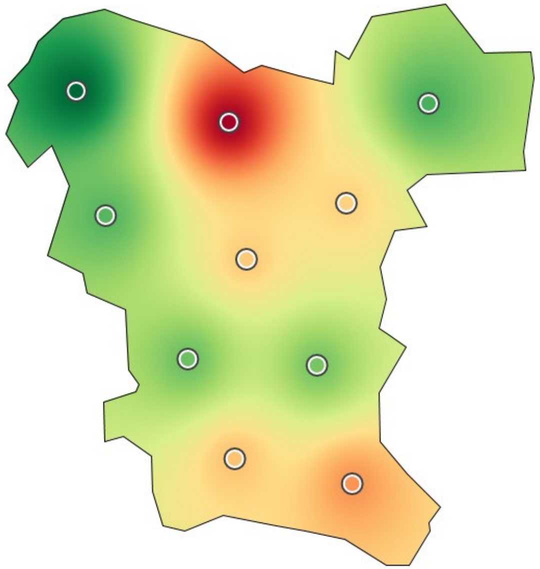
It is time for a demo!



Points only
(invisible local scope)

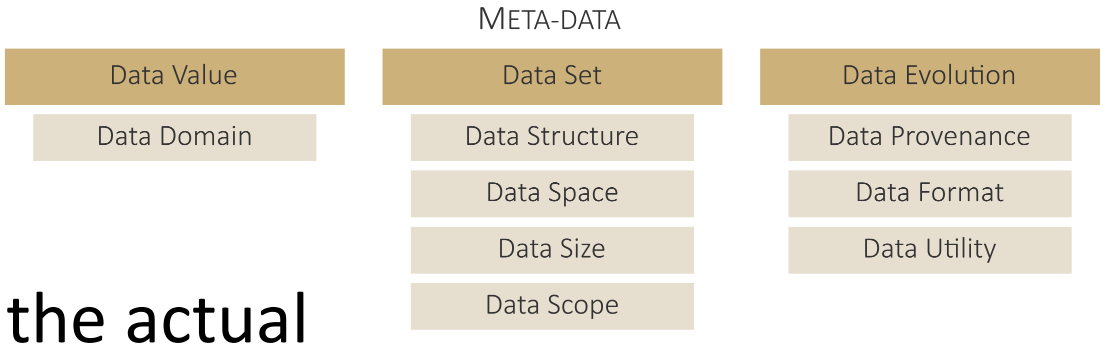


Voronoi partition
(discrete local scope)



Scattered data interpolation
(continuous local scope)

Influencing factor: Data



Meta-data

- Meta-data contain information about the actual data with respect to:
  - Data values: From which data domain are the values?
  - Data set: How is the dataset structured, how does the data space look like, how big are the data, what is their scope?
  - Data evolution:
    - Data Provenance: History of how the data were created?
    - Data Format: How are the data stored presently?
    - Data Utility: How may the data be used in the future?

Influencing factor: Data

Data classes

- Data classes pertain to different data aspects
  - **A** – data attributes
  - **T** – time
  - **S** – space
  - **R** – structural relationships
- Based on **A**, **T**, **S**, and **R**, different data classes can be defined
  - Multivariate data, temporal data, spatial data, spatio-temporal data, graph data, …

Influencing factor: Data

Data classes

- **Multivariate data (A)**
  - Data with many attributes (A)
  - Product data, player statistics, gene expressions, simulation runs, etc.
- **Temporal data (T → A)**
  - Data where attributes depend on time (T), a.k.a. time-dependent data
  - Financial data, medical treatment data, sensor data, etc.
- **Spatial data (S → A)**
  - Data being located in space (S), a.k.a. geo-spatial data
  - Election data, distribution of places, land use, etc.

Influencing factor: Data

Data classes

- **Spatio-temporal data ($S \times T \rightarrow A$)**
  - Data depends on space and time
  - Meteorological/climate data, disease spread, movement data, etc.
- **Graph data (R, $R \rightarrow A$, $R \rightarrow S \times T$, or $S \times T \rightarrow R \times A$)**
  - Entities (vertices V) and relations (edges E) between them, graph $G = (V, E)$
  - Entities and relations may be associated with further information (A, S, T)
  - Social networks, biological pathways, connectome data, etc.

Influencing factor: Data

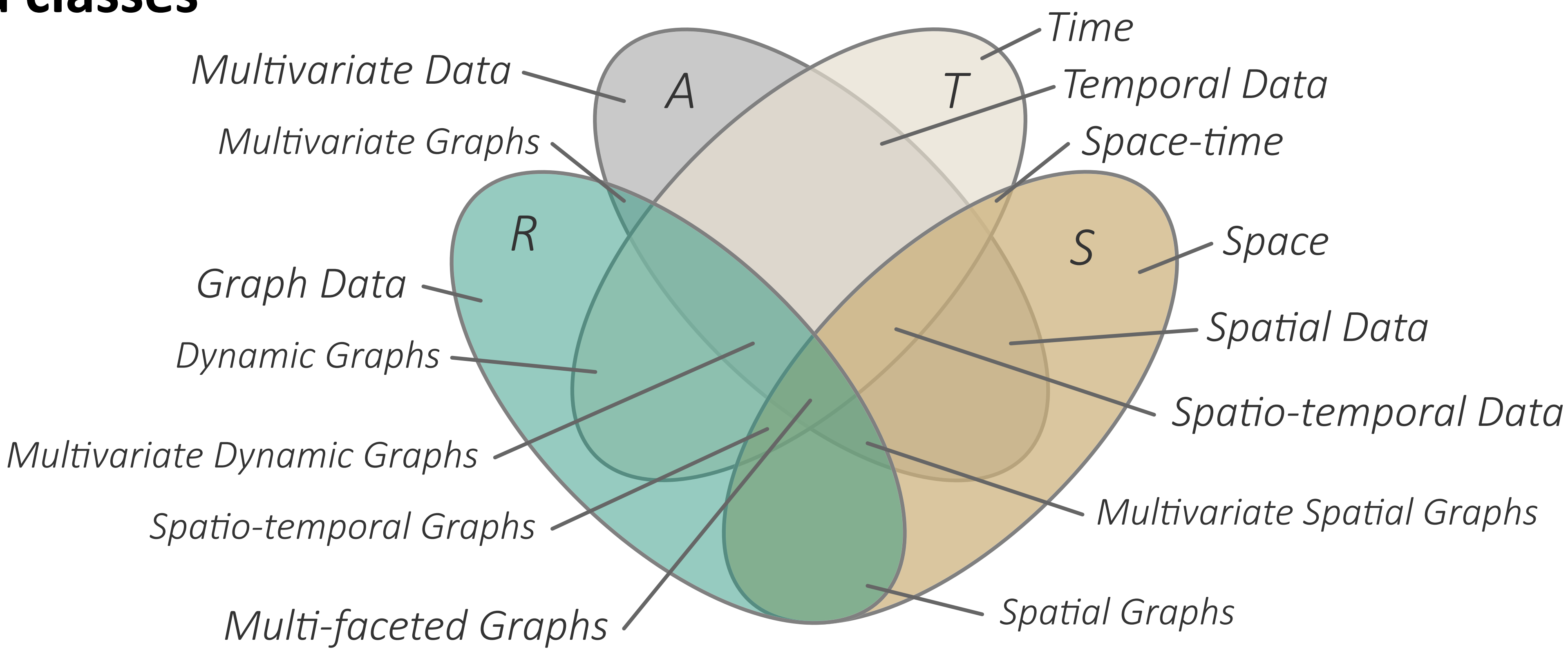
Data classes

- Further data classes
  - Text (A, R)
  - Images (S, A)
  - Flow data (S, A)
  - Volume data (S, A)

Influencing factor: Data

This type of visual representation is called a *Venn-diagram*.

Data classes



Influencing factor: Analysis tasks

With respect to the **analysis tasks**, the following aspects are relevant

- Goals: *Why* do we carry out a task?
- Analytic questions: *What* do we seek to answer?
- Targets: *Where* in the data do we operate?
- Means: *How* do we carry out the task?

Influencing factor: Analysis tasks

Goals: Why?

- **5 goals** are primarily relevant for interactive visual data analysis
  - Exploration
  - Description
  - Explanation
  - Confirmation
  - Presentation
- Goals roughly structure the analysis process
  - We start **exploring** the data and make observations. Next we **describe** findings and try to **explain** them. Then we **confirm** our hypotheses and finally **present** the confirmed analysis results.

Influencing factor: Analysis tasks

Goals: Why?

This goal very much relates to the classic idea of *Exploratory Data Analysis* as formulated by John W. Tukey, 1977.

- **Exploration**
  - Open-ended undirected search
  - "I-know-it-when-I-see-it" approach
  - Make first observations (e.g., gain overview, detect patterns and outliers, …)
- **Description**
  - Characterize observations
  - Derive specifications for observations
  - Example: Describe outliers
    - What are their characteristic values?
    - Where are they located in space and time?

Influencing factor: Analysis tasks

Goals: Why?

- **Explanation**
  - Develop deeper understanding of observations (e.g., by identifying all contributing data, finding main causes, checking re-occurrence, etc.)
  - Formulate hypotheses about the data
- **Confirmation**
  - Verify hypotheses
  - Look for concrete evidence to back up or refute hypothesis
  - Compare different subsets and different visual representations
  - Generate analysis results

Influencing factor: Analysis tasks

Goals: Why?

- **Presentation**
  - Communicate confirmed analysis results
  - Best done by telling a story
  - Convince others

  - Think about: Who are we convincing in explanation and confirmation?

Influencing factor: Analysis tasks

Analytic questions: What?

- Data analysis activities may involve a variety of analytic questions at two distinct levels
  - **Elementary questions**
    - Data elements are studied individually
    - May include one or more individual elements
  - **Synoptic questions**
    - Sets of data elements are studied
    - Consider sets as a whole, not individual elements

Influencing factor: Analysis tasks

Analytic questions: What?

- **Elementary questions**
  - Identify: What is the value?
  - Locate: Where is the value?
  - Compare: Is it less or more?
  - Rank: Is there any order?
  - Connect: Are they related?
  - Distinguish: What makes the difference?

Influencing factor: Analysis tasks

Analytic questions: What?

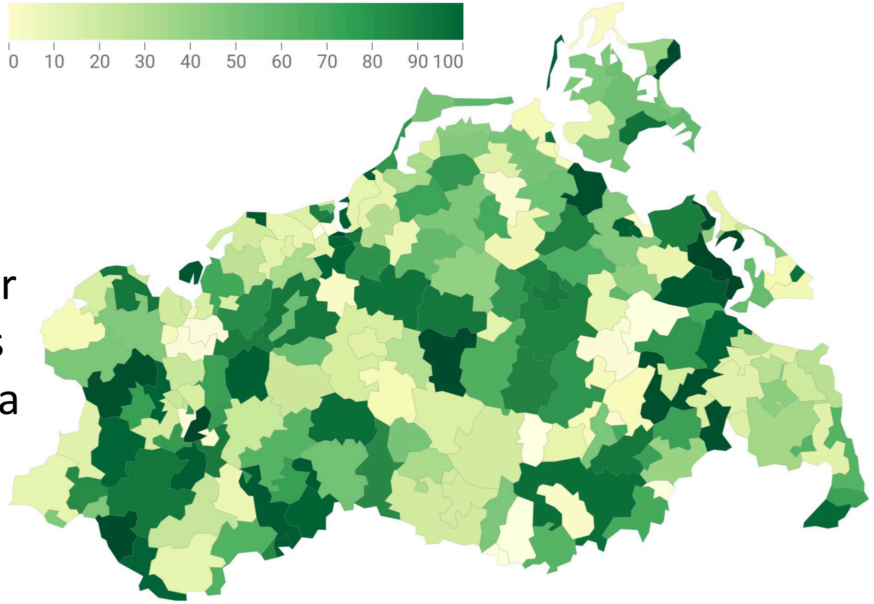
- **Synoptic questions**
  - Group: Do they belong together?
  - Correlate: Are there any dependencies?
  - Trends: Do they develop systematically?
  - Cycles: Do they re-occur periodically?
  - Outliers: Are they special with respect to the rest?
  - Features: What is characteristic for the data?

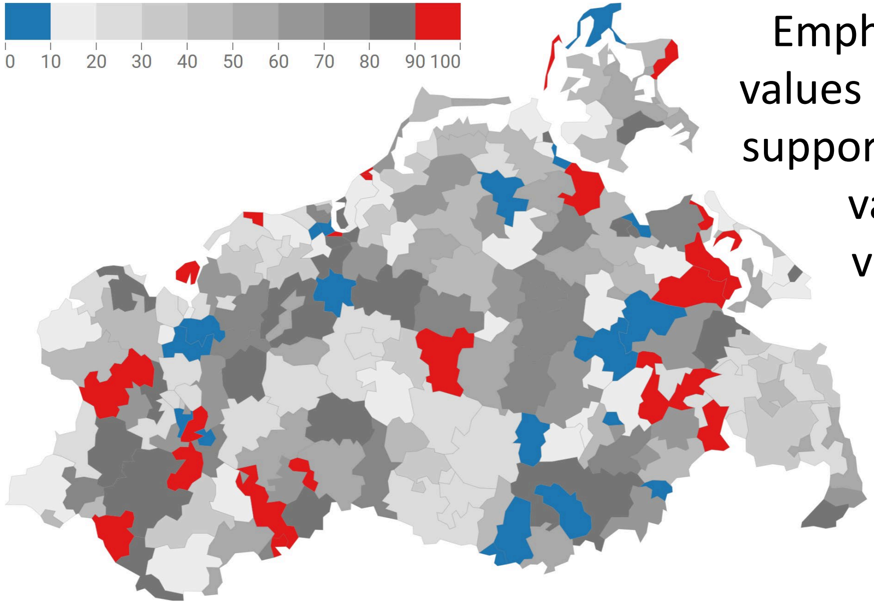
Influencing factor: Analysis tasks

Analytic questions: What?

- Different questions require different visual representations
- Example: Color-coded map for **identification** and **location** tasks

A uniform color scale supports **identifying** data values



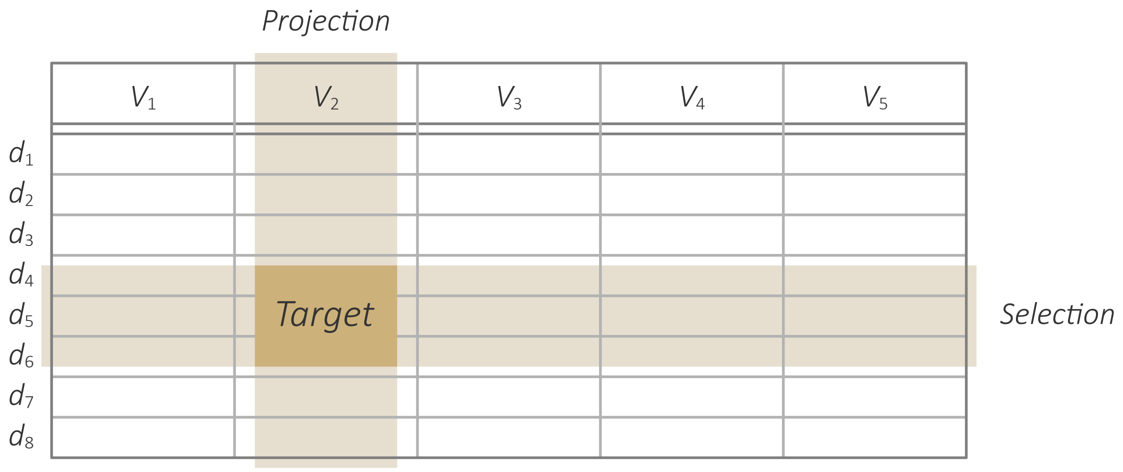


Emphasizing specific values in the color scale supports **locating** these values in the visualization

Influencing factor: Analysis tasks

Targets: Where?

- Define where in the data a task operates
- Knowing the target allows us to focus the analysis on relevant data
  - **Data of interest**
    - Particular data elements: Selection
    - Specific data attributes: Projection
  - **Data granularity**
    - Relevant for hierarchical multi-scale data
    - At what scale does a task operate
      - High level of abstraction: Overview tasks
      - High level of detail: Tasks required fine-grained information



Influencing factor: Analysis tasks

Means: How?

- We have three options for carrying out analysis tasks
  - **Visual means**
    - Tasks rely primarily on the human visual system
    - Example: Detect group of similar data elements
  - **Interactive means**
    - Tasks are carried out primarily through human action
    - Example: Select data elements for detailed inspection
  - **Computational means**
    - Tasks conducted by the machine
    - Example: Cluster similar data elements

Influencing factor: Analysis tasks

Summary

TASKS

Goals	Questions		Targets	Means
	Elementary	Synoptic		
Explore Describe Explain Confirm Present	Identify Locate …	Group Correlate …	Data of Interest Granularity	Visual Interactive Computational

Influencing factor: Analysis tasks

Examples

Goal	Question	Target	Means
Explore	locations of	maximum values	visually.
Describe	groups of	low-value elements	by marking.
Confirm	cyclic behavior	of temperature	by statistics.

Influencing factor: Context

The **context** captures the influencing factors pertaining to the users and the technologies involved in the data analysis

- **Users**
  - Human factors
  - User background and expertise
  - Application domain
  - Single-user and collaborative analysis
- **Technologies**
  - Computational resources
  - Display technologies
  - Input modalities

Summary

- **Criteria** (few)
  - Expressiveness
  - Effectiveness
  - Efficiency
- **Influencing factors** (many)
  - The subject: Data
  - The objective: Analysis tasks
  - The context: Users and technologies

*The **many influencing factors** make the design of interactive visual data analysis solutions a **challenging** endeavor!*

Summary

What can you do with criteria and influencing factors?

- Assess the quality of an interactive visual data analysis solution!
  - Is the solution expressive, effective, and efficient?
- Compare solutions based on criteria!
  - Is one solution more suited than another?
- Ask your customers the right questions!
  - Are your data qualitative or quantitative?
  - How big are your data?
  - What are your goals?
  - Think about: What else might you want to ask?

Assignments

1. Read about the context as influencing factor in detail in Chapter 2 of "Interactive Visual Data Analysis" by Tominski and Schumann!

Questions

1. Name and explain the 3 quality criteria for interactive visual data analysis!
2. What can we learn from the “lie factor”?
3. What factors influence design and selection of adequate methods?
4. How is the data domain characterized?
5. What are independent and dependent variables and how are they related to the reference space and the attribute space?
6. Name different data classes and give corresponding example data sets!
7. Name and explain the 5 primary goals of interactive visual data analysis!
8. How are elementary and synoptic analysis questions characterized? Give examples tasks!
9. How can identification tasks and location tasks be supported by appropriate color-coding?
10. What are selection and projection good for?
11. What is captured by the context as an influencing factor?
12. What would you sensibly ask a customer who approaches you with a data analysis project proposal?