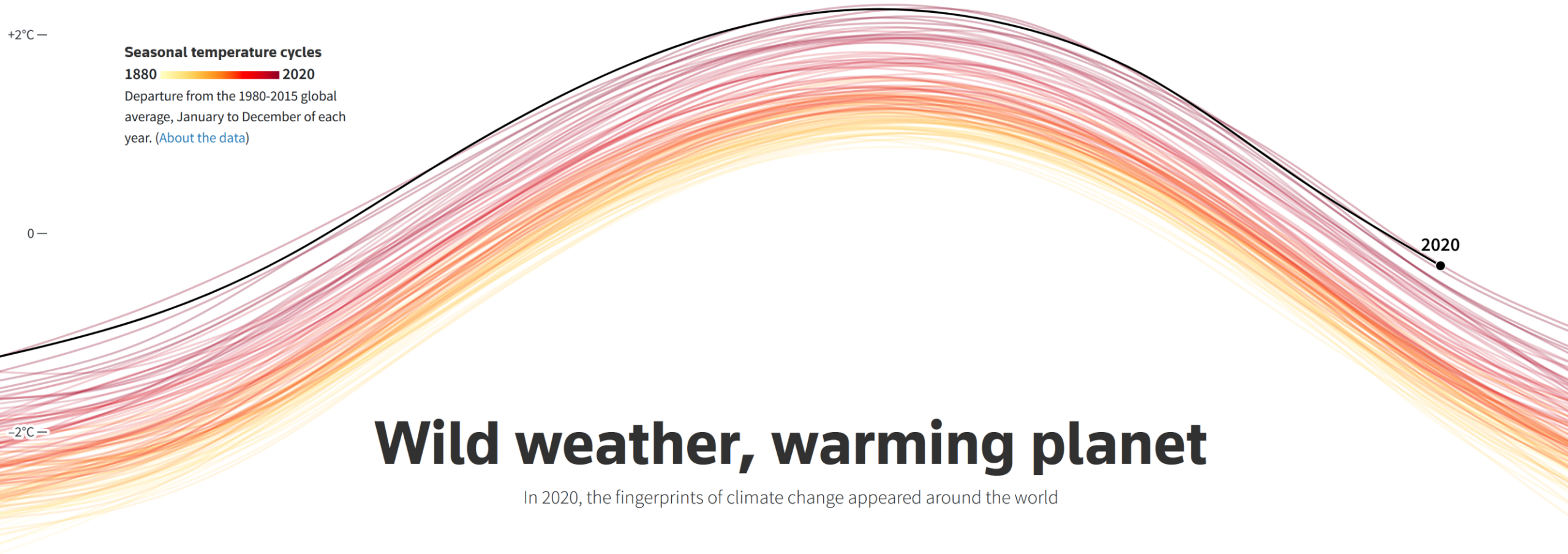
Visualization
Summary

Prof. Tobias Günther
FAU Erlangen-Nürnberg

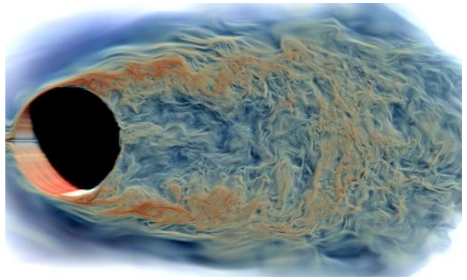
Visualization



Reuters Graphics, https://graphics.reuters.com/ENVIRONMENT-2020/WARMING/qzjpqdadnvx/

Introduction

Scientific Visualization

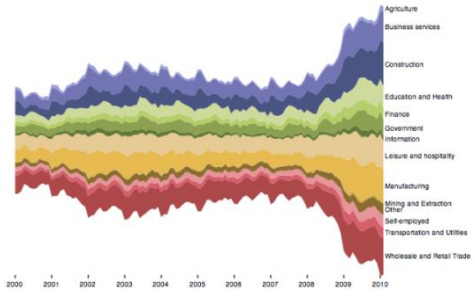


- Users:
  - Primarily other **scientists** (assume certain knowledge)
- Data:
  - **Spatial** and/or **temporal** context
  - Fields (scalar, vector, tensor)
- Concepts:
  - Use of **differential calculus**
  - **Numerically robust** algorithms
  - **Feature** definition, extraction and visualization

Baeza Rojo et al., Accelerated Monte Carlo Rendering of Finite-Time Lyapunov Exponents, IEEE TVCG 26(1), 708-718, 2019.

Information Visualization

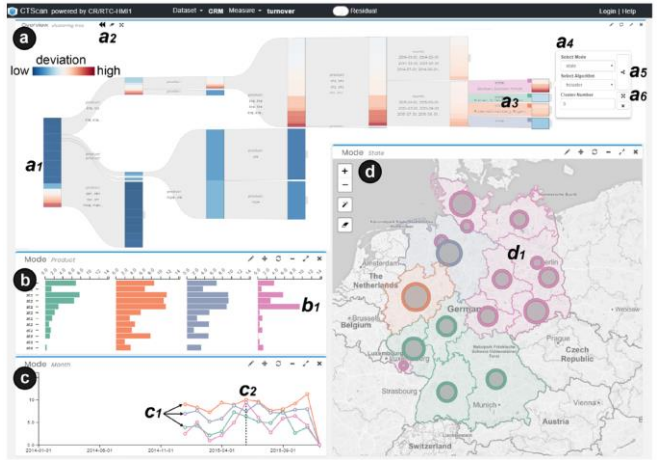
- Users:
  - **Any person** (no prior assumptions about users)
- Data:
  - **No spatial** context
  - **Abstract**, unstructured, graphs
- Concepts:
  - **Statistical** analysis methods
  - **Perceptual** foundations
  - Artistic elements



Jeffrey Heer et al., A Tour Through the Visualization Zoo, https://homes.cs.washington.edu/~jheer/files/zoo/

Visual Analytics

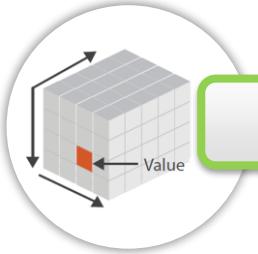
- Intersection of SciVis and InfoVis
- Sophisticated interactive systems tailored for specific use case
- Strong focus on tasks to be solved



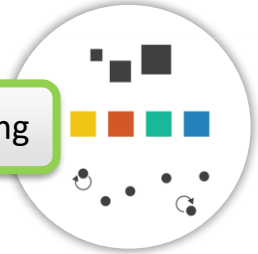
Liu et al., TPFlow: Progressive Partition and Multidimensional Pattern Extraction for Large-Scale Spatio-Temporal Data Analysis, IEEE TVCG 25(1), 1-11, 2019.

Course Overview

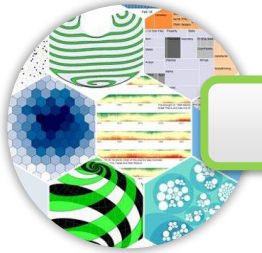
Visualization Design



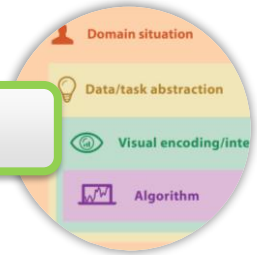
Data Abstraction



Perception and Mapping



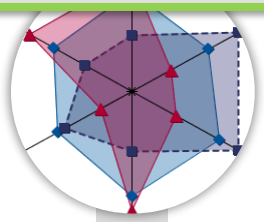
Web-based Systems



Tasks and Validation

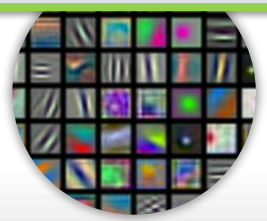
Abstract Data

Tabular Data



Networks and Trees

Deep Learning

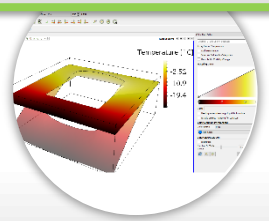


Spatial Data

Scalar and Vector Fields

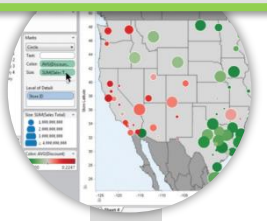
Medical Data

SciVis Tools

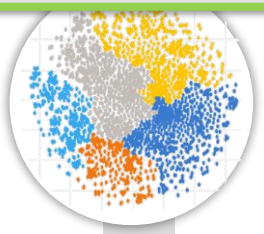


Interaction Concepts

View Manipulation



Data Reduction



Rules of Thumb

Visualization Design

Abstract Data

Spatial Data

Interaction Concepts

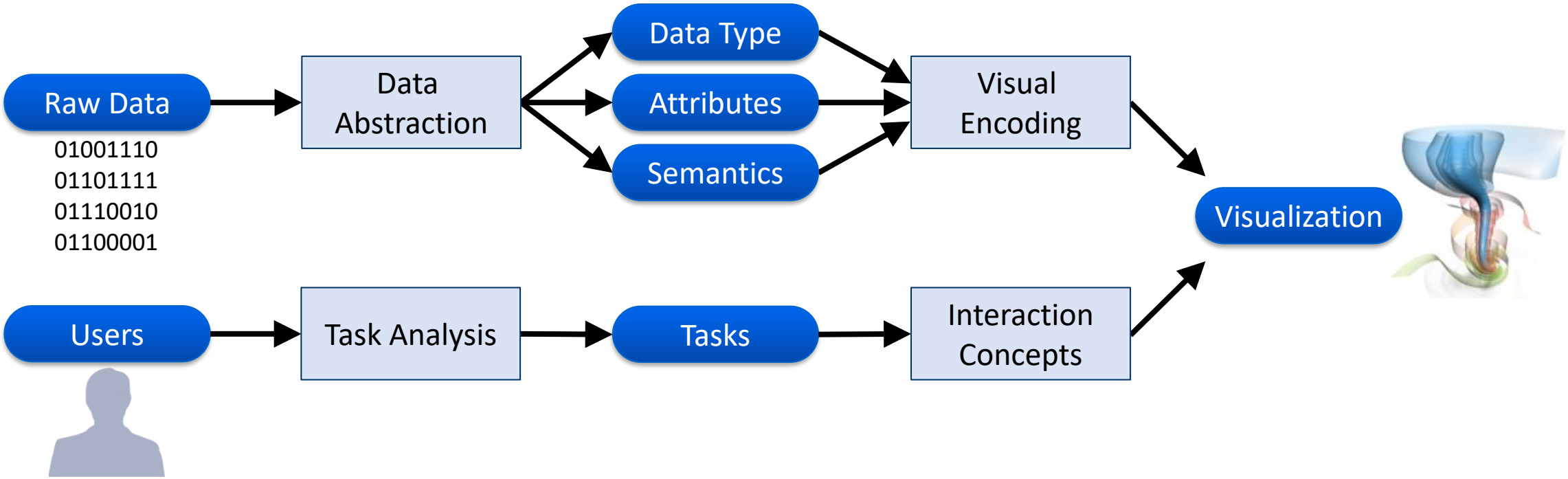
Data Abstraction

Perception & Mapping

Web-based Systems

Tasks & Validation

Simplied Visualization Design Workflow



Data Types

Attribute

- Property that can be measured, or logged, e.g., salary, price, number of sales, protein expression level, temperature, etc.

Item

- Individual discrete entity, e.g., row in a table, node in a network

Link

- Relationship between items, typically within a network

Position

- Spatial (and/or temporal) location, e.g., lat/lon on sphere, 3d coordinate in space, etc.

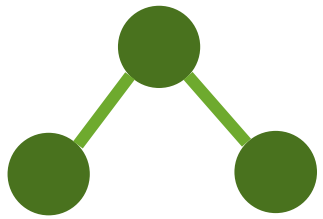
Grid

- Defines connectivity of positions for interpolation

Data Types

Abstract data

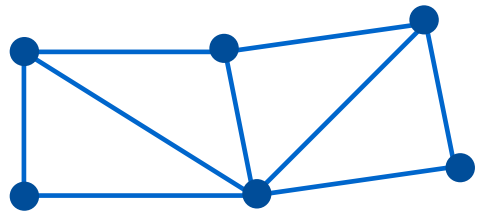
- Item
- Link
- Attribute
  - Per item
  - Per link



„variable“ is a synonym for „attribute“

Spatial data

- Position (Geometry)
- Grid (Topology)
  - Formed from cells
- Attribute
  - Per position
  - Per cell



„dimension“: number of position coordinates

Example of Data Types

- **Abstract or Spatial?**
  - Abstract
- **Item?**
  - Countries
- **Links?**
  - None!
- **Attributes?**
  - Income per person, CO2 emission, number of inhabitants, region

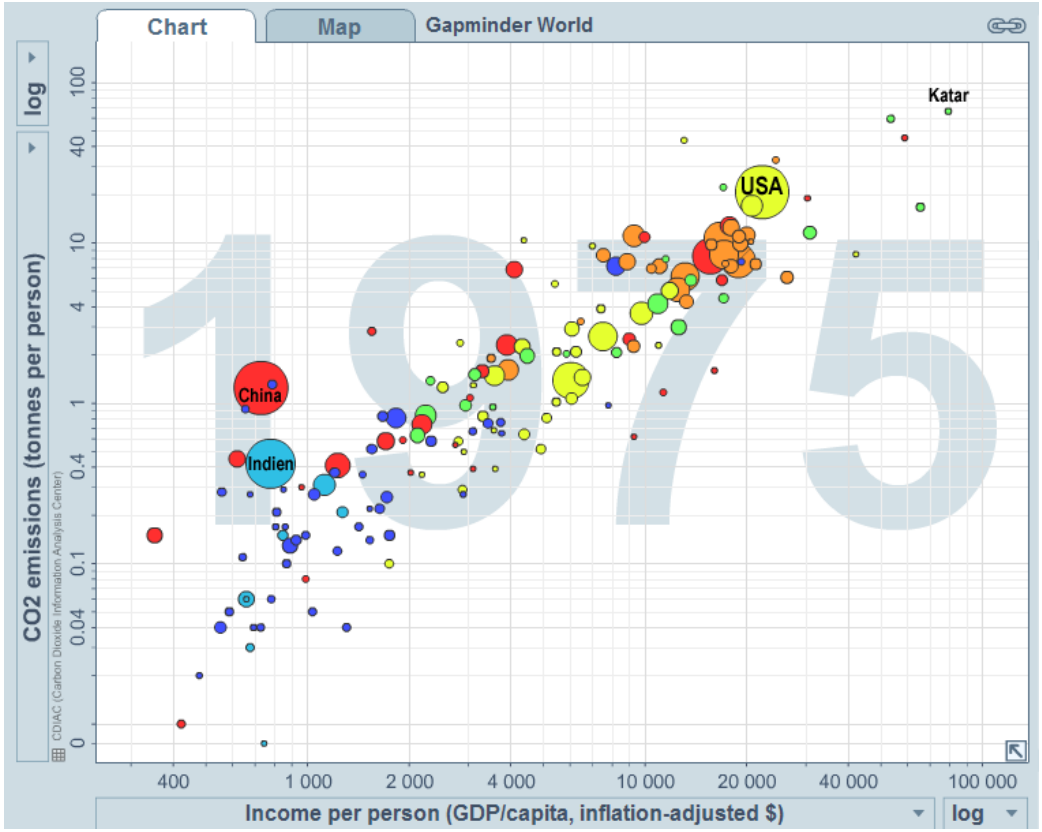
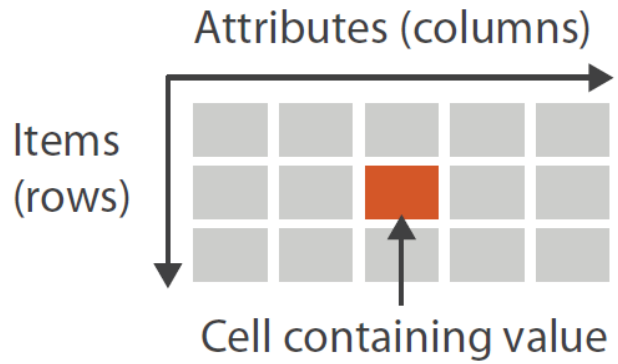


Image created with: www.gapminder.org,

Table

- Consists of **items** (rows) and **attributes** (columns)
- Value stored in the cell
- Sometimes referred to as *flat table*



A	B	C	S	T	U
Order ID	Order Date	Order Priority	Product Container	Product Base Margin	Ship Date
3	10/14/06	5-Low	Large Box	0.8	10/21/06
6	2/21/08	4-Not Specified	Small Pack	0.55	2/22/08
32	7/16/07	2-High	Small Pack	0.79	7/17/07
32	7/16/07	2-High	Jumbo Box		7/17/07
32	7/16/07	2-High	Medium Box		7/18/07
32	7/16/07	2-High	Medium Box		7/18/07
35	10/23/07	4-Not Specified	Wrap Bag	0.52	10/24/07
35	10/23/07	4-Not Specified	Small Box	0.58	10/25/07
36	11/3/07	1-Urgent	Small Box	0.55	11/3/07
65	3/18/07	1-Urgent	Small Pack	0.49	3/19/07
66	[illegible]	5-Low	Wrap Bag	0.56	1/20/05
69	[illegible]5	4-Not Specified	Small Pack	0.44	6/6/05
69	[illegible]5	4-Not Specified	Wrap Bag	0.6	6/6/05
70	12/18/06	5-Low	Small Box	0.59	12/23/06
70	12/18/06	5-Low	Wrap Bag	0.82	12/23/06
96	4/17/05	2-High	Small Box	0.55	4/19/05

attribute

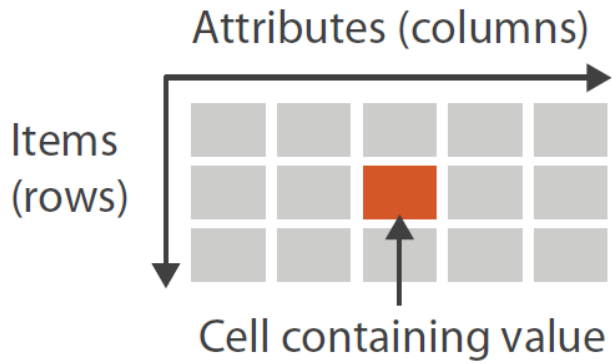
item

cell

From Munzner, Visualization Analysis & Design, 2014

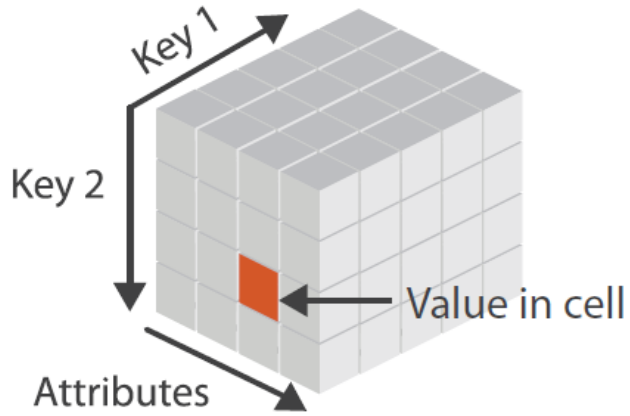
Table

- Consists of **items** (rows) and **attributes** (columns)
- Value stored in the cell
- Sometimes referred to as *flat table*



Multi-dimensional Table

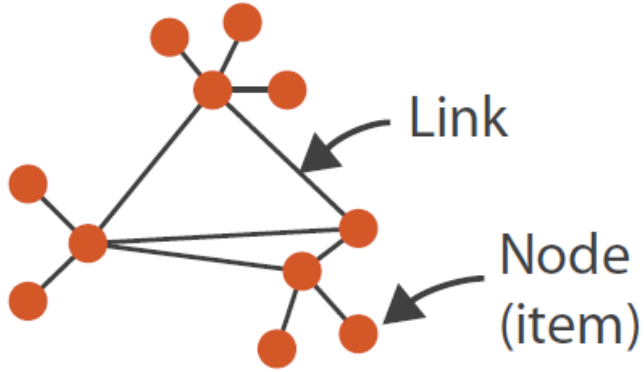
- Multiple keys for indexing an item



From Munzner, Visualization Analysis & Design, 2014

Network

- Consists of **items** (nodes), **links**, and **attributes** (on node or link)
- Used to define relationships between items, e.g., social network



Tree

- Network with **hierarchical** structure, i.e., no cycles, one parent, e.g., biological tree of life, organization chart in company, …

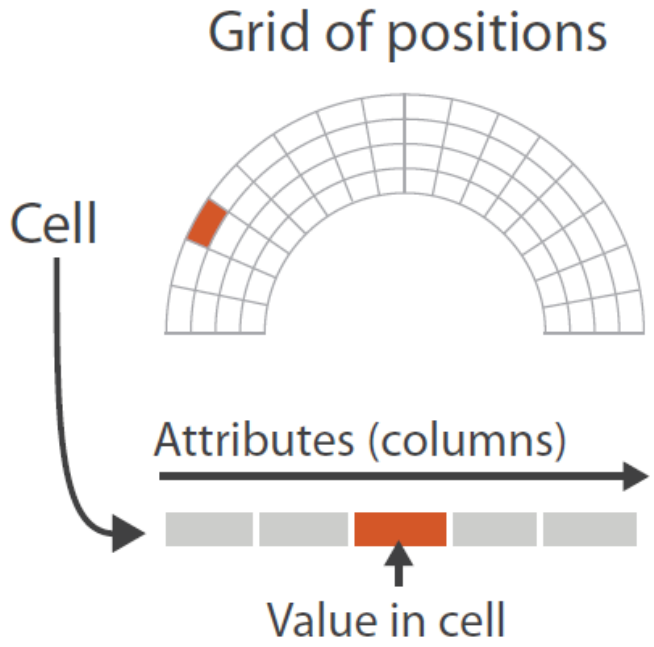
From Munzner, Visualization Analysis & Design, 2014

Fields

- Discretized onto **grids**
  - Structured grids
  - Unstructured grids
- Sampling requires **interpolation**
  - Interpolation order (linear, cubic, …)
- Map spatial location to a value

$$f(x_1, \ldots, x_n) \rightarrow (y_1, \ldots y_m)$$

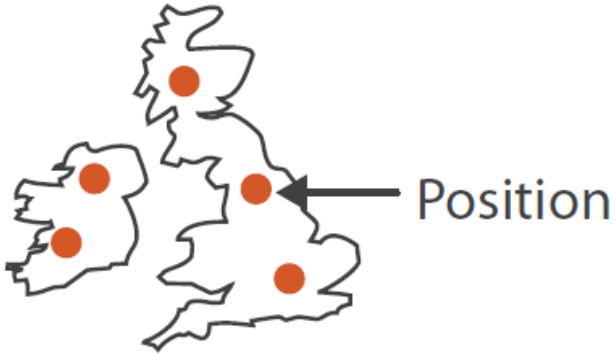
- Scalar field, vector field, tensor field, …

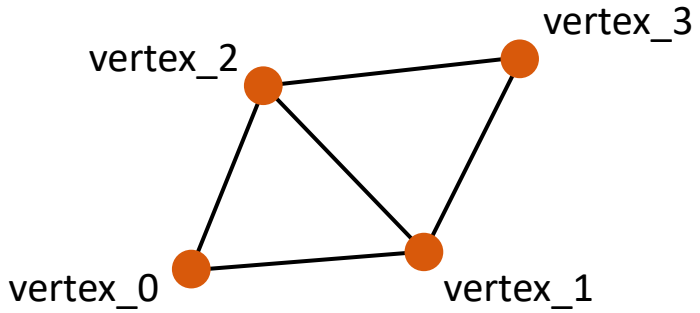


From Munzner, Visualization Analysis & Design, 2014

Geometry

- **Examples**
  - Point clouds
  - Line strips
  - Triangle meshes





vertices

vertex_0.x	vertex_0.y	vertex_0.z
vertex_1.x	vertex_1.y	vertex_1.z
vertex_2.x	vertex_2.y	vertex_2.z
vertex_3.x	vertex_3.y	vertex_3.z

coordinates

primitives

0	1	2
1	2	3

indices

- Can be stored in matrices
- Does not necessarily store attributes

From Munzner, Visualization Analysis & Design, 2014

Attribute Types

Visualization of Nominal/Categorical Data

- no order, data represent categories (e.g., car manufacturers and countries)
- Vis. goal: categorization

categorical

Visualization of Ordinal Data

- There is an order among values (often finite and discrete number of values)
- Differences between values cannot be quantified, e.g., S,M,L,XL
- Vis. goal: (qualitative) comparison

Quantitative Data

- Measured/simulated values represent numbers (land prices, temperature)
- The “distance” between two values can be measured.
- Can be integer or real numbers
- Vis. goal: comparison of numbers

ordered

[Bertin, 1983]

Example of Attribute Types

- Income per person
  - Quantitative, sequential
- CO2 emission
  - Quantitative, sequential
- Number of inhabitants
  - Quantitative, sequential
- Region (Europe, North America, …)
  - Categorical

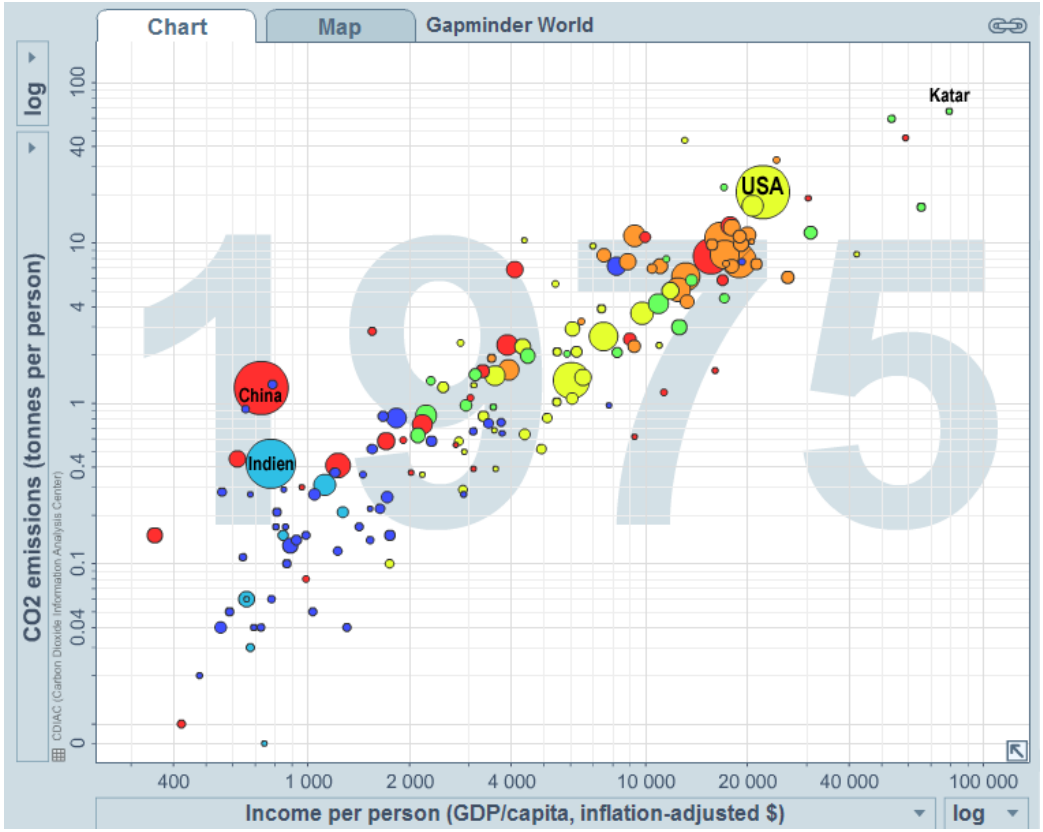
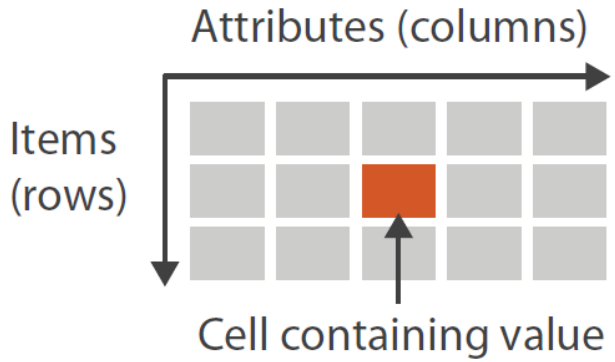


Image created with: www.gapminder.org,

Key vs. Value

- A **key** is an index to look up **value** of attributes
  - Key = independent attribute(s)
  - Value = data of dependent attribute(s)

- **Flat table**
  - A key (row) is mapped to the list of attribute values (columns)



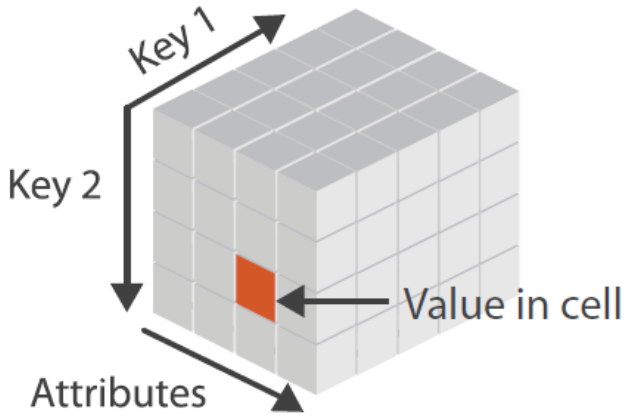
Flat table

Tamara Munzner, Visualization Analysis and Design,
AK Peters Visualization Series, CRC Press, page 36, 2014.

Key vs. Value

- A **key** is an index to look up **value** of attributes
  - Key = independent attribute(s)
  - Value = data of dependent attribute(s)

- **Multi-dimensional table**
  - The combination of keys must be unique
  - Example from biology: gene expression level (quantitative attribute) for a certain gene (key 1) at a certain time (key 2)



Multi-dimensional table

Tamara Munzner, Visualization Analysis and Design, AK Peters Visualization Series, CRC Press, page 36, 2014.

Visualization Design

Abstract Data

Spatial Data

Interaction Concepts

Data Abstraction

Perception & Mapping

Web-based Systems

Tasks & Validation

Mark

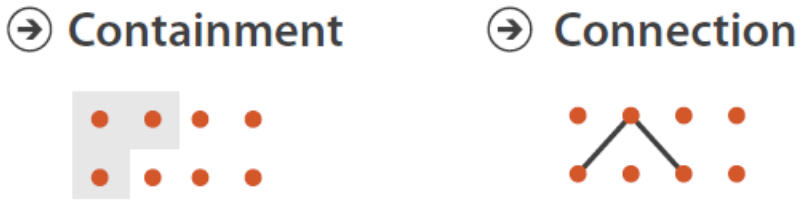
- Basic graphical element in an image that represents items or links
- Geometric objects that are classified according to spatial dimensions
  - 0-d mark: **point**
  - 1-d mark: **line**
  - 2-d mark: **area**
  - (3-d mark: **volume**)

From Munzner, Visualization Analysis & Design, 2014

Mark Types

- **Marks as item/node**
  - E.g., table datasets

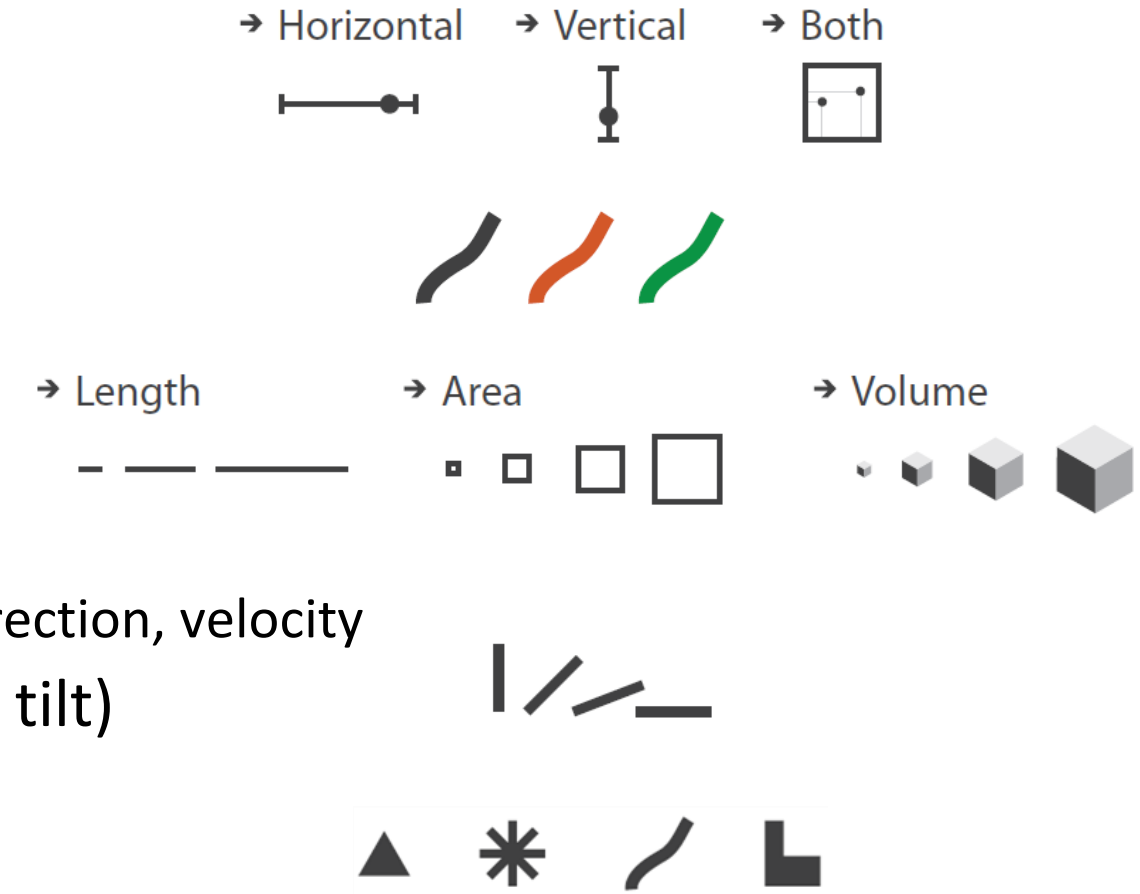
- **Marks as link**
  - E.g., network data sets
  - Types of link marks:
    - Connection mark: pairwise relationship between two items (1D line)
    - Containment mark: hierarchical relationship between areas



From Munzner, Visualization Analysis & Design, 2014

Channel (appearance of mark)

- Spatial position
  - Aligned/unaligned, depth, region
- Color
  - Hue, saturation, luminance
- Size
  - Length (1D), area (2D), volume (3D)
- Motion
  - Motion pattern (oscillation, jumps), direction, velocity
- Orientation (also known as angle or tilt)
- Curvature
- Shape



From Munzner, Visualization Analysis & Design, 2014

Channel Types

- **Identity channels** tell us *what* something is or *where* it is.
  - Shape, hue of a color, motion pattern…

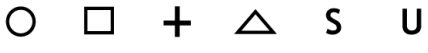
    → Categorical data

- **Magnitude channels** tell us *how much* of something there is.
  - Size, luminance, saturation, orientation, …

    → Ordered data

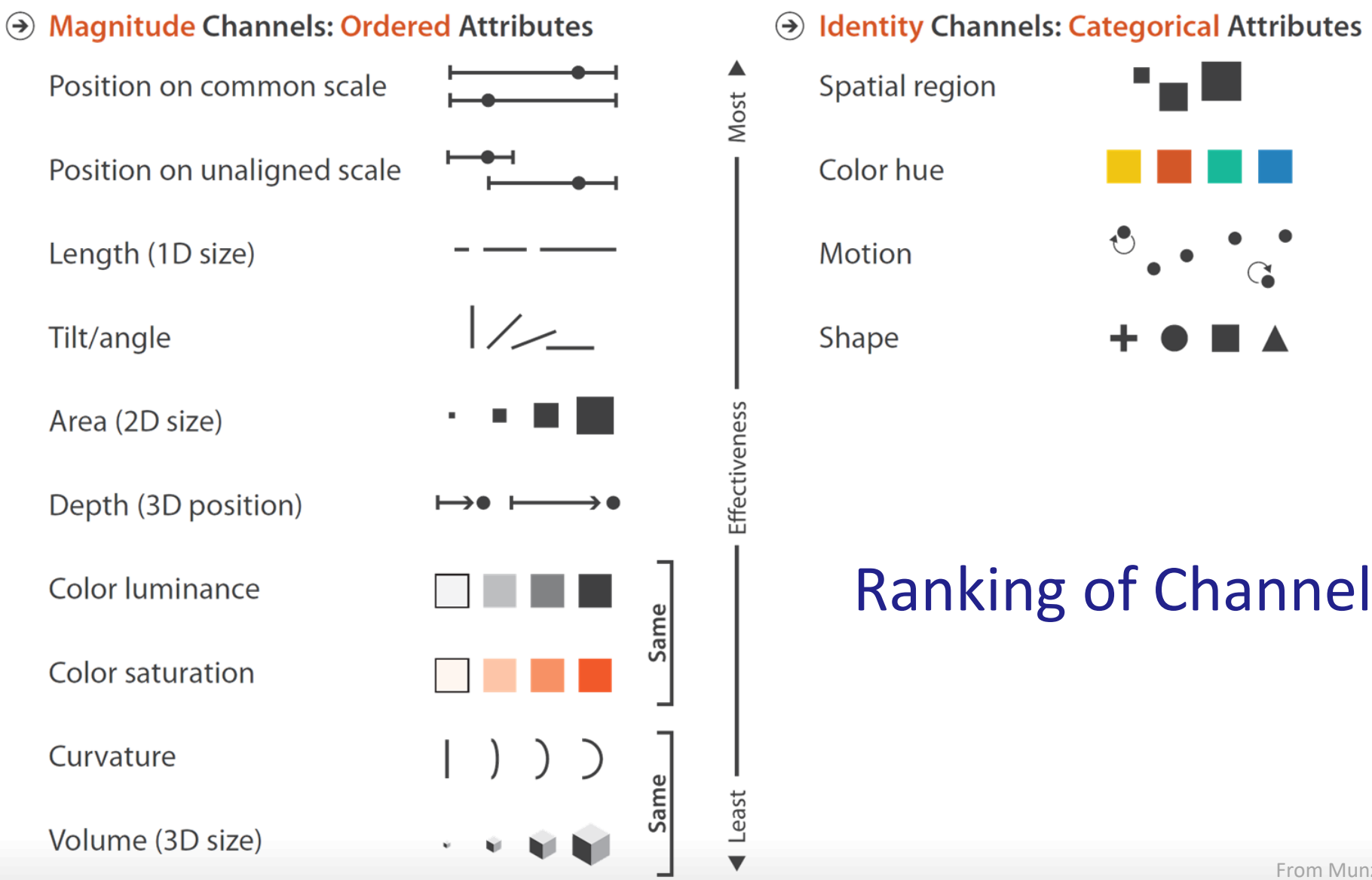
Variables for the visualization of information (Stolte, 2002).

Categorical, ordinal and quantitative data

Property	Categorical	Ordinal	Quantitative
Shape			
Size			
Orientation			
Color			

Chris Stolte et al., Polaris: A System for Query, Analysis,and Visualization of Multidimensional Relational Databases, IEEE Transactions on Visualization and Computer Graphics 8(1), 2002.

Appropriate Encodings



Ranking of Channel Effectivess

From Munzner, Visualization Analysis & Design, 2014

Appropriate Encodings

Sublinear phenomena

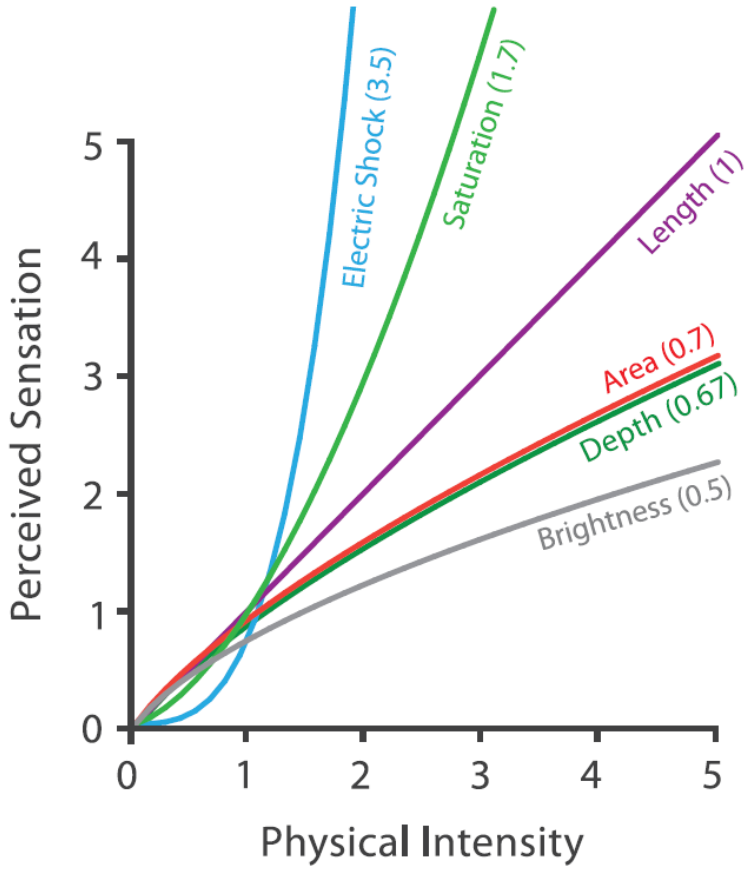
- Doubleing **brightness** results in smaller perceived difference

Linear phenomena

- **Length** of line segment on a 2D plane perpendicular to the viewer

Superlinear phenomena

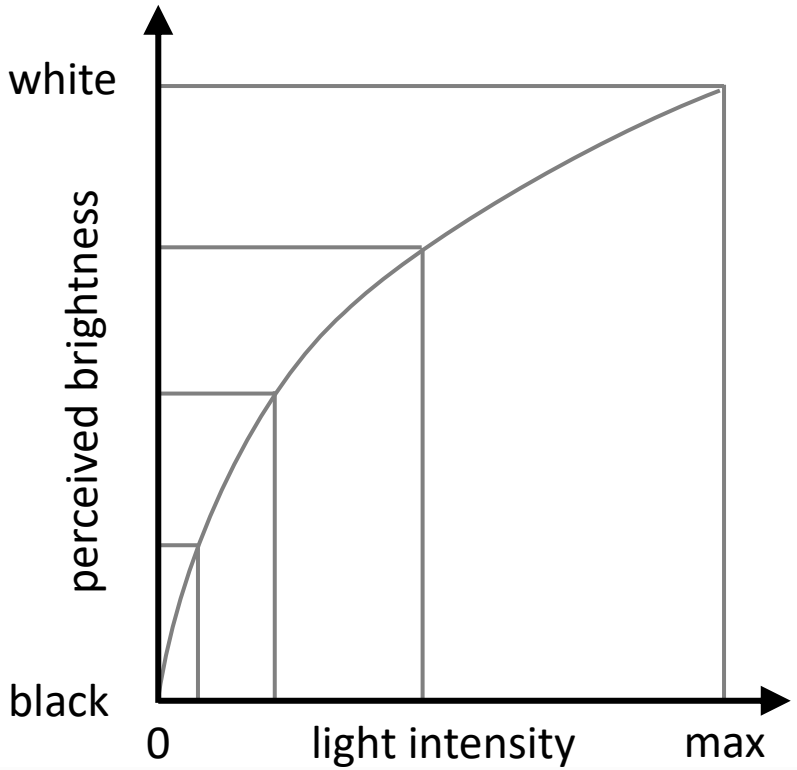
- Red-gray **saturation** is magnified

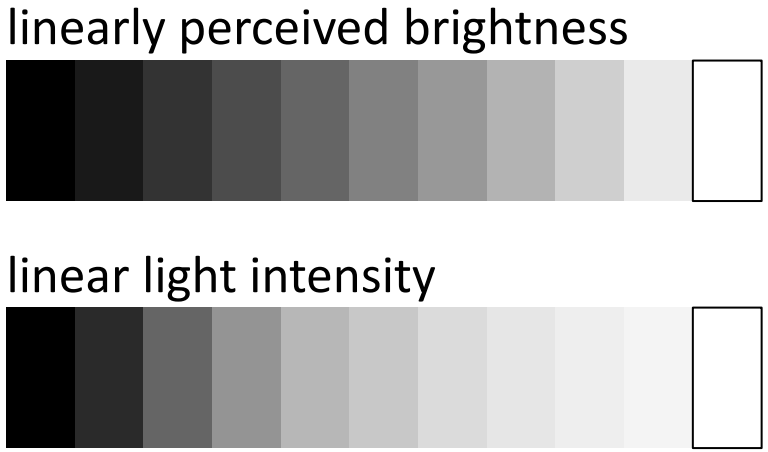


From Munzner, Visualization Analysis & Design, 2014

Luminance Perception

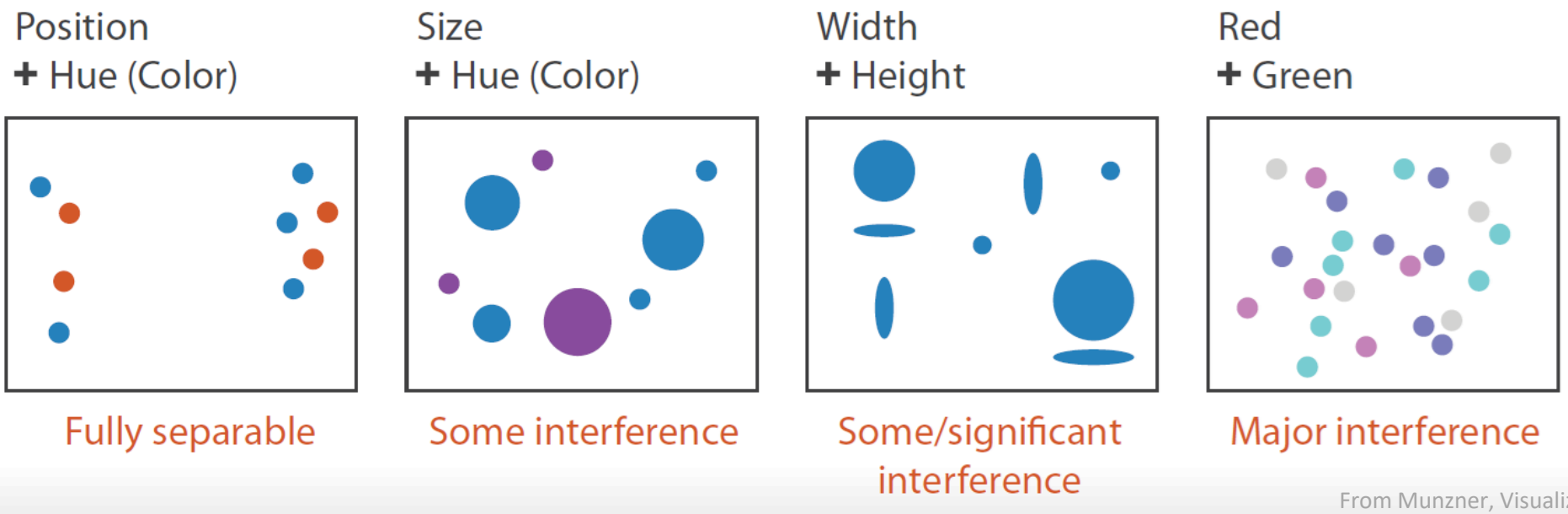
Non-linear relation between light intensity and perceived brightness





Separability

- Visual channels are **not independent** from each other!
- Combinations are on a spectrum from separable to integral
- Optimal choice depends on what we want to show

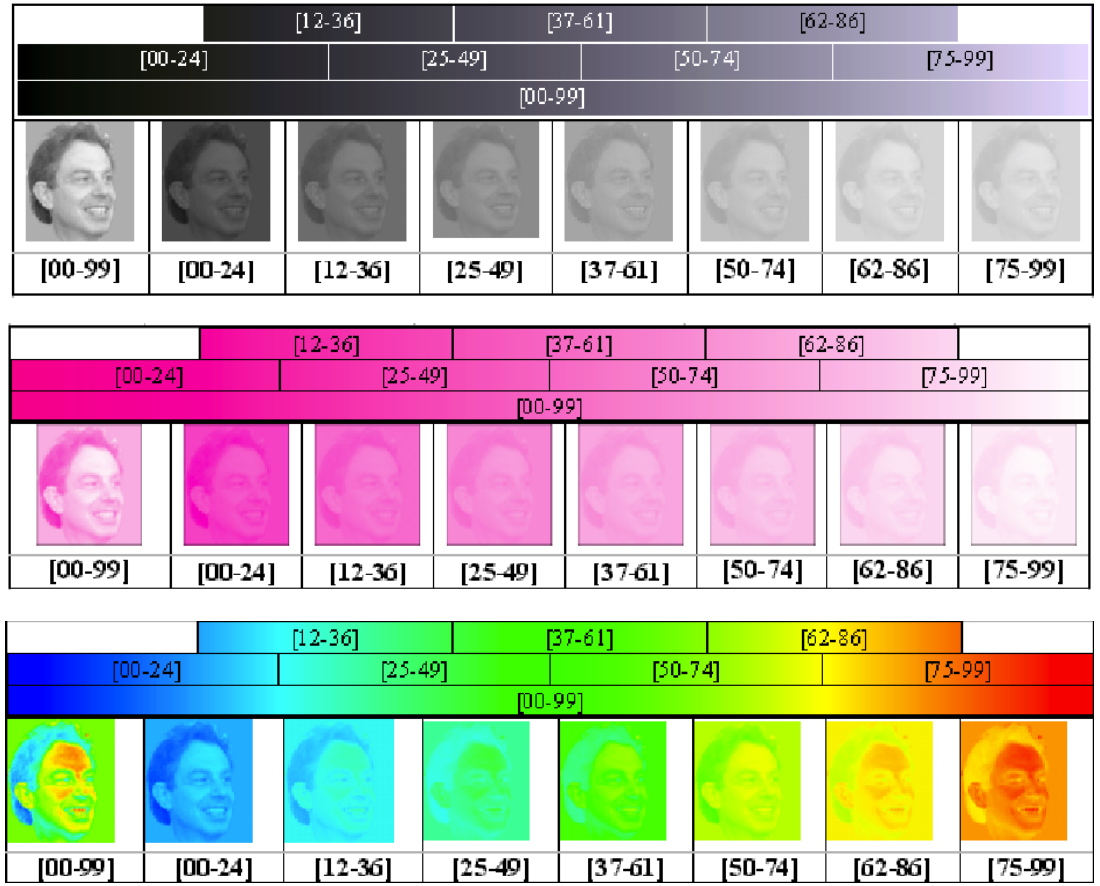


From Munzner, Visualization Analysis & Design, 2014

Grouping

- Perceptual grouping is achieved by (ordered from best to worst)
  1. **Link marks** (e.g., area of containment)
     - Containment is the strongest grouping cue
     - Connection is second-strongest grouping cue
  2. **Proximity**
     - position channel: same spatial region
  3. **Similarity**
     - Identity channels: hue, motion, shape

Color Mapping: Evaluation



Grey scale:
relatively good differentiation in all regions

Scale with decreasing saturation:
well in upper areas

Rainbow scale:
only in the lowest ¼ portion well. In several regions unnatural. Poor contrast.

Bernice E. Rogowitz, Alan D. Kalvin: The "Which Blair Project": A Quick Visual Method for Evaluating Perceptual Color Maps. Proc. of IEEE Visualization 2001, pp. 183-190

Two Stages of Visual Perception

Stage 1: Parallel processing to extract low-level properties

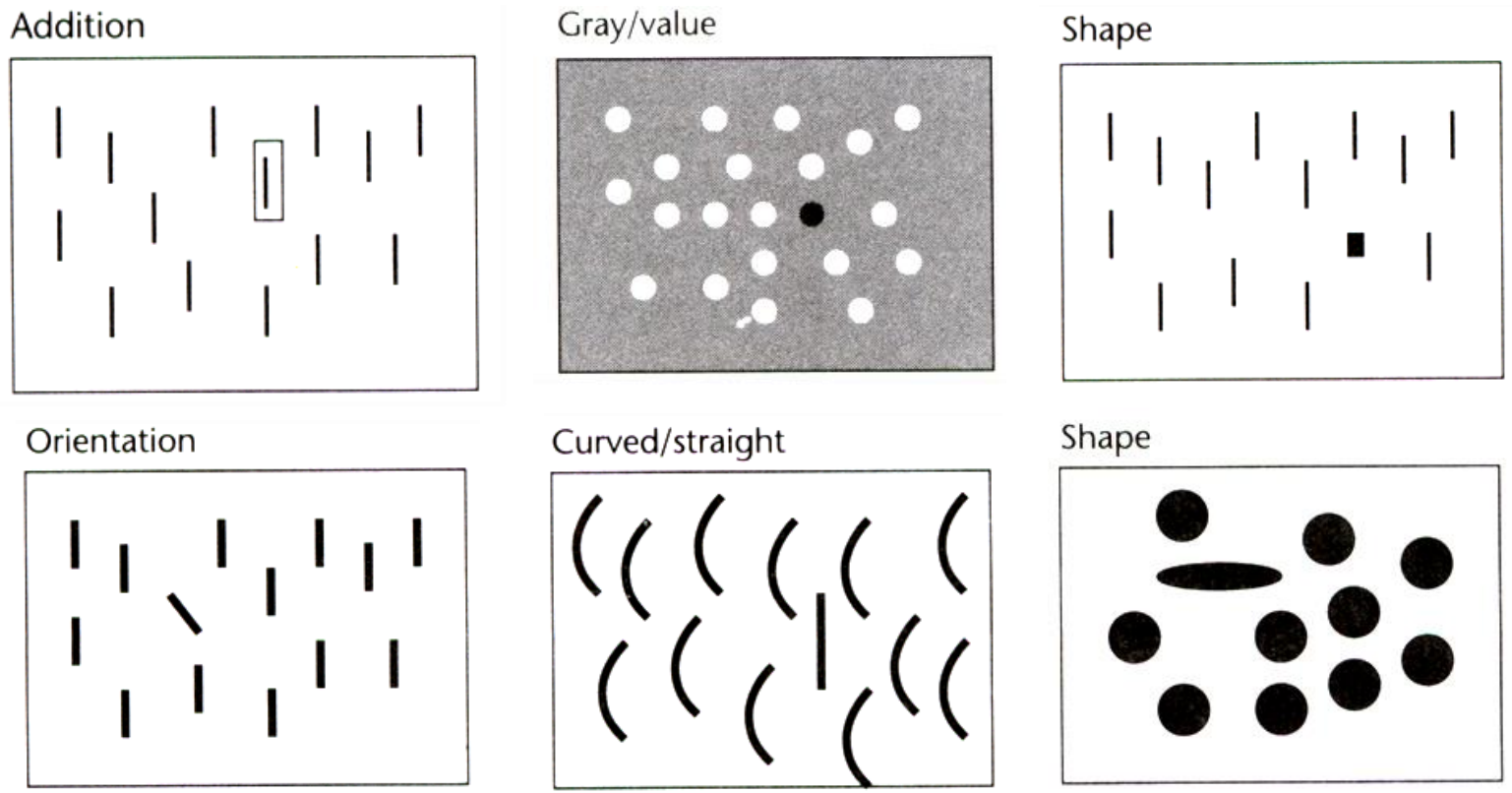
- rapid parallel processing at the retina (less than 200-250 ms)
- extraction of features, orientation, color, texture and movement
- transitory nature of information (held briefly)
- bottom-up processing without special attention

Stage 2: Sequential goal-directed processing

- slow serial processing
- involvement of short term and long term memory
- more emphasis on symbol recognition
- top-down processing (directed by goals)

If we know what is perceived in stage 1, we can encode important information such that it is perceived without cognitive effort.

These differences are perceived preattentively (stage 1).



Colin Ware, Information Visualization: Perception for Design (3rd), Elsevier, p.154, 2013.

Visualization Design

Abstract Data

Spatial Data

Interaction Concepts

Data Abstraction

Perception & Mapping

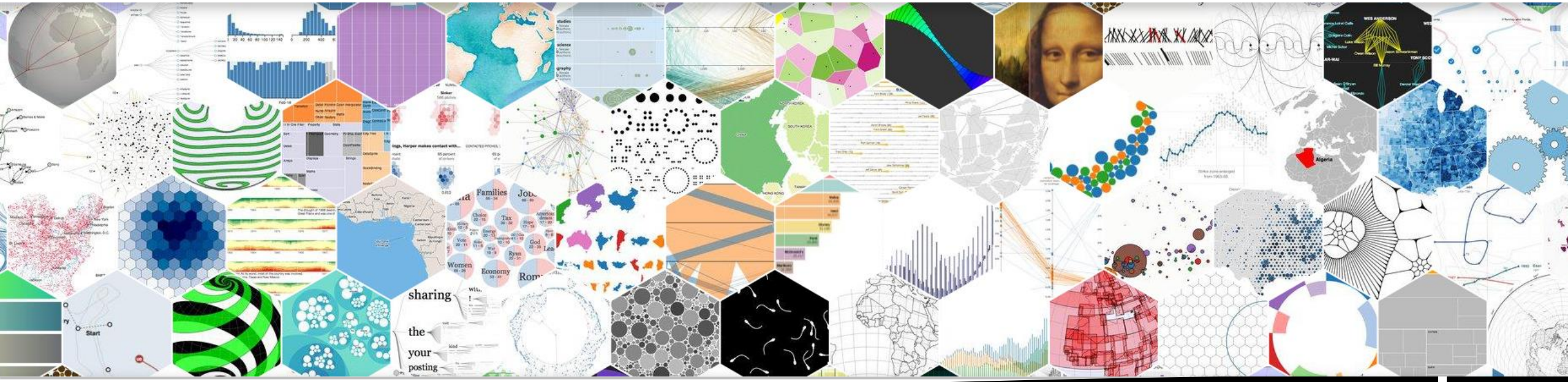
Web-based Systems

Tasks & Validation

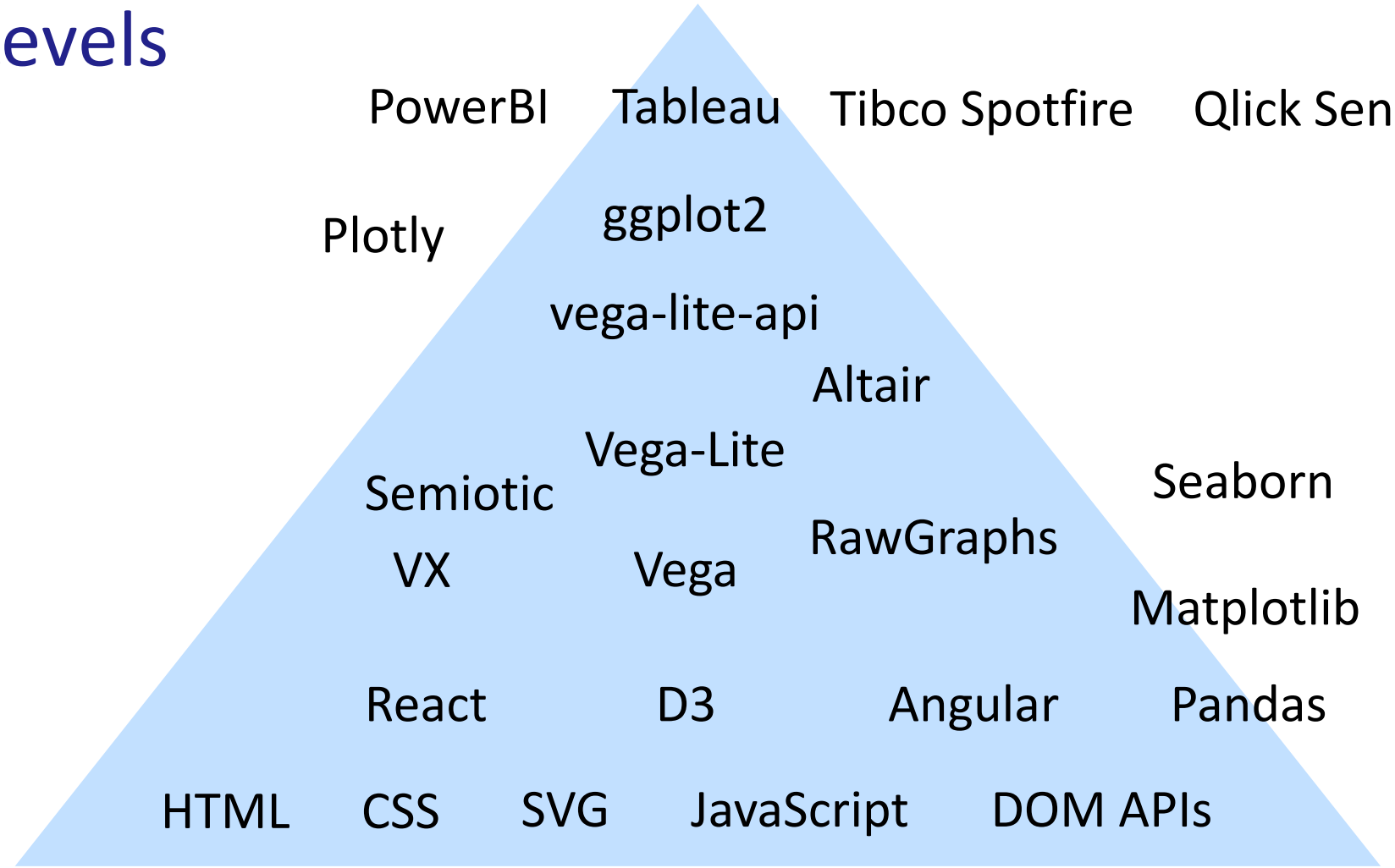
D3 (Data-Driven Documents)

- Commonly-used JavaScript library for visualization

https://d3js.org/



Abstraction Levels



Visualization Design

Abstract Data
Spatial Data
Interaction Concepts

Data Abstraction
Perception & Mapping
Web-based Systems
Tasks & Validation

Example

- **Task of Epidemiologist**: "*contrast* the prognosis of patients who were intubated in the ICU more than one month after exposure to patients hospitalized within one week."

- **Task of Biologist**: "see if the results for the tissue samples treated with LL-37 *match up* with the ones without the peptide.

- **Abstract description**: "*compare* values between two groups"

Actions

- How do we analyze the data?
- Is data consumed or produced?

→ Consume

→ *Discover*

→ *Present*

→ *Enjoy*

→ Produce

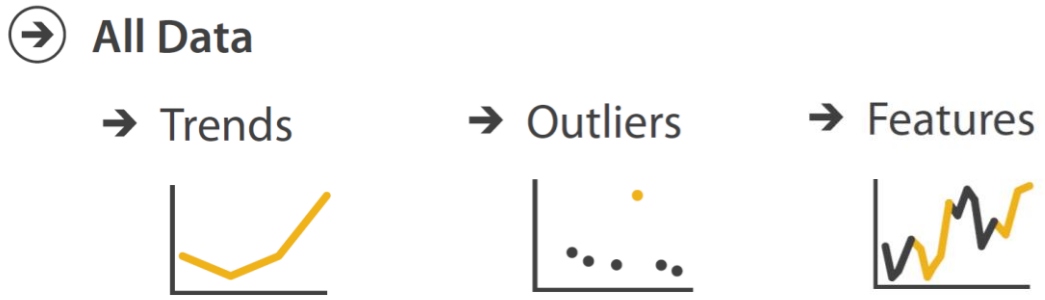
→ *Annotate*

→ *Record*

→ *Derive*

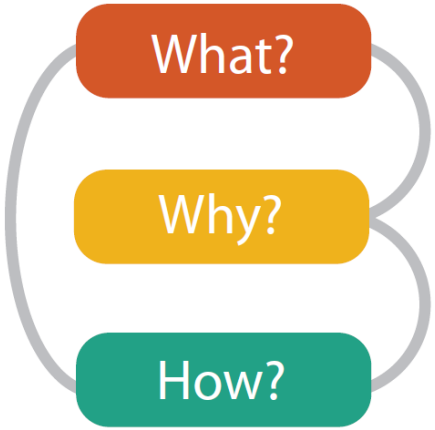
High-Level Targets

- *Trends*
  - High-level characterization of pattern in the data
  - E.g., increases, decreases, peaks, troughs, plateaus
- *Outliers*
  - Elements that stand out from a trend
- *Features*
  - Task-dependent structure of interest



From Munzner, Visualization Analysis & Design, 2014

Analyzing Visualizations



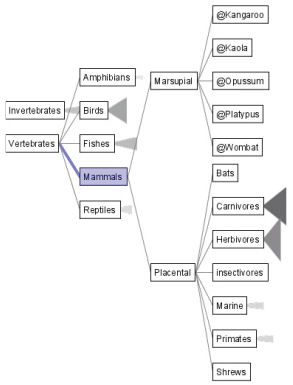
Data Abstraction

Task Abstraction

Idioms

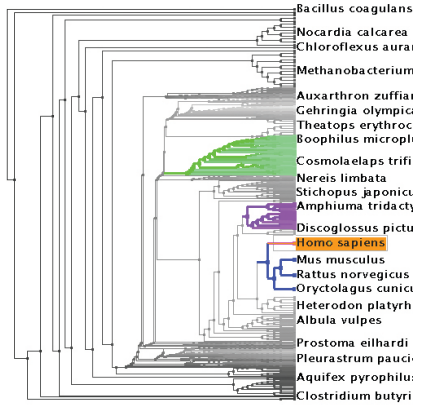
From Munzner, Visualization Analysis & Design, 2014

Comparing Two Idioms



SpaceTree
[Paisant et al. 2002]

TreeJuxtaposer
[Munzner et al. 2003]



What?

- Tree

Why?

- Actions
  - Present
  - Locate
  - Identify

- Targets
  - Path between two nodes

How?

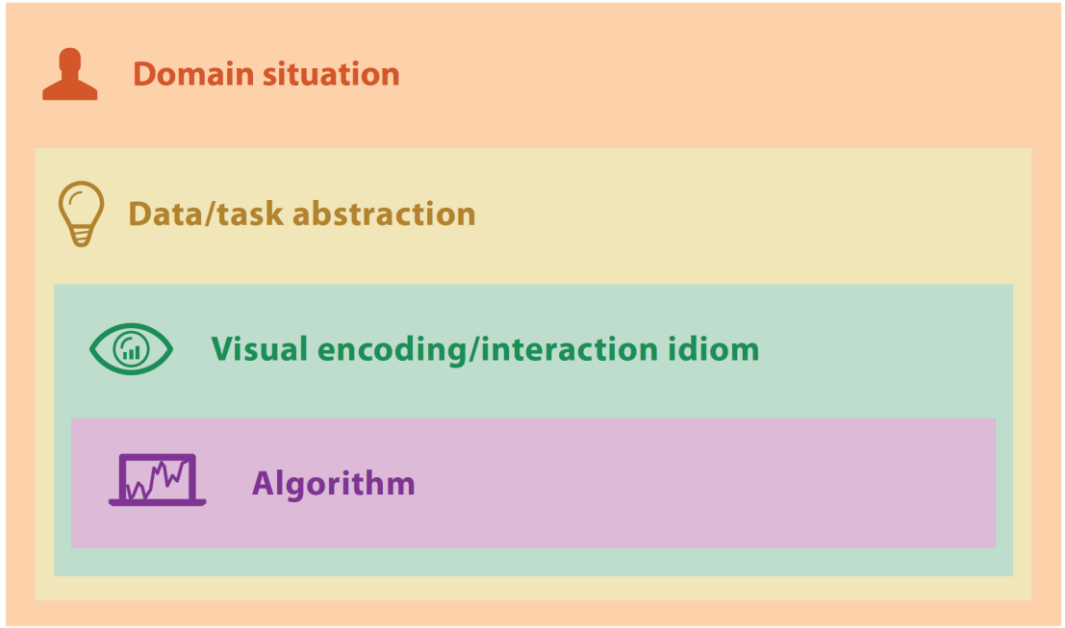
- SpaceTree
  - Encode
  - Navigate
  - Select
  - Filter
  - Aggregate

- TreeJuxtaposer
  - Encode
  - Navigate
  - Select
  - Arrange

From Munzner, Visualization Analysis & Design, 2014

Four nested levels of visualization design

- Breaking it down allows us to address different aspects
- Levels are nested and depend on previous level



From Munzner, Visualization Analysis & Design, 2014

Different threats to validity on each level

From Munzner, Visualization Analysis & Design, 2014

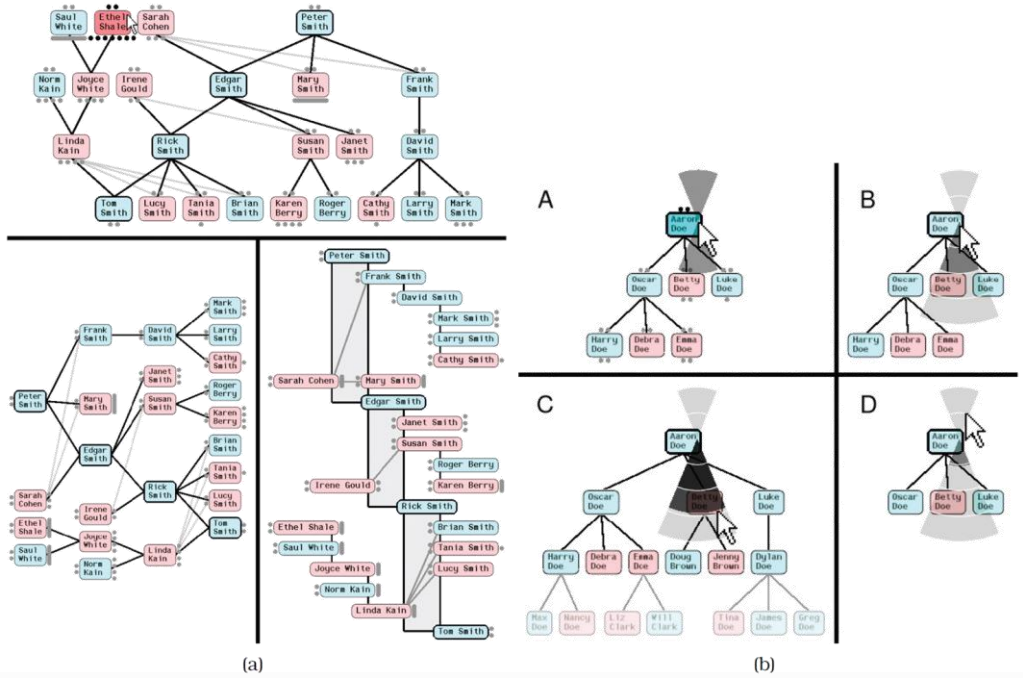
Visualization of Genealogical Graphs

- Paper proposed new visual encoding idiom called „dual tree“, which is a subgraph formed by the union of two trees

Justify encoding/interaction design

Qualitative result image analysis

Test on target users, collect anecdotal evidence of utility



From Munzner, Visualization Analysis & Design, 2014

Visualization Design

Abstract Data

Spatial Data

Interaction Concepts

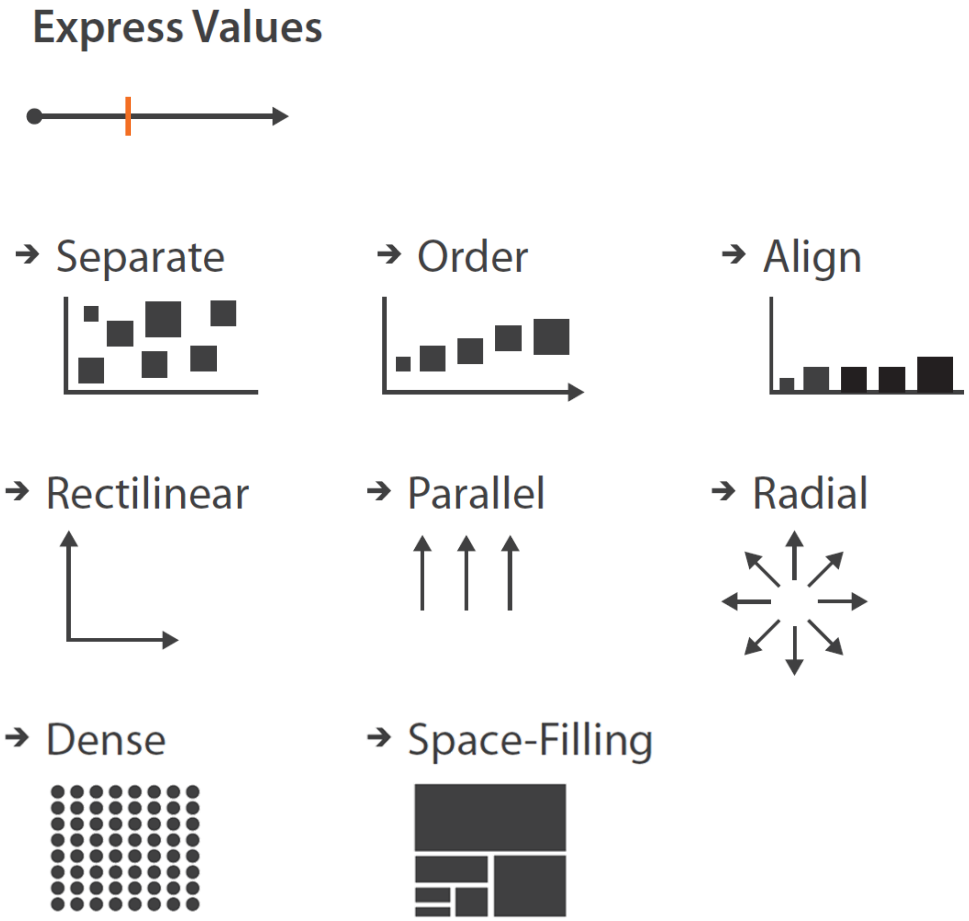
Tabular Data

Networks and Trees

Deep Learning

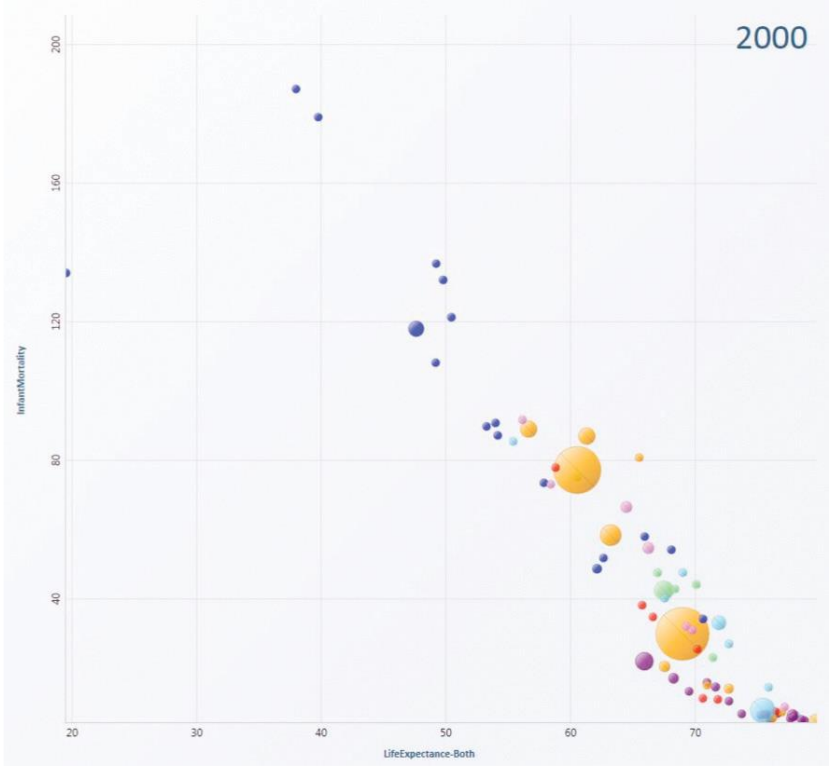
Design Choices for Spatial Arrangement of Tabular Data

- **Quantitative data**
  - Express values
- **Categorical data**
  - Separate, order, align regions
- **Axis orientation**
  - Rectilinear, parallel, radial
- **Layout density**
  - Dense, space-filling



From Munzner, Visualization Analysis & Design, 2014

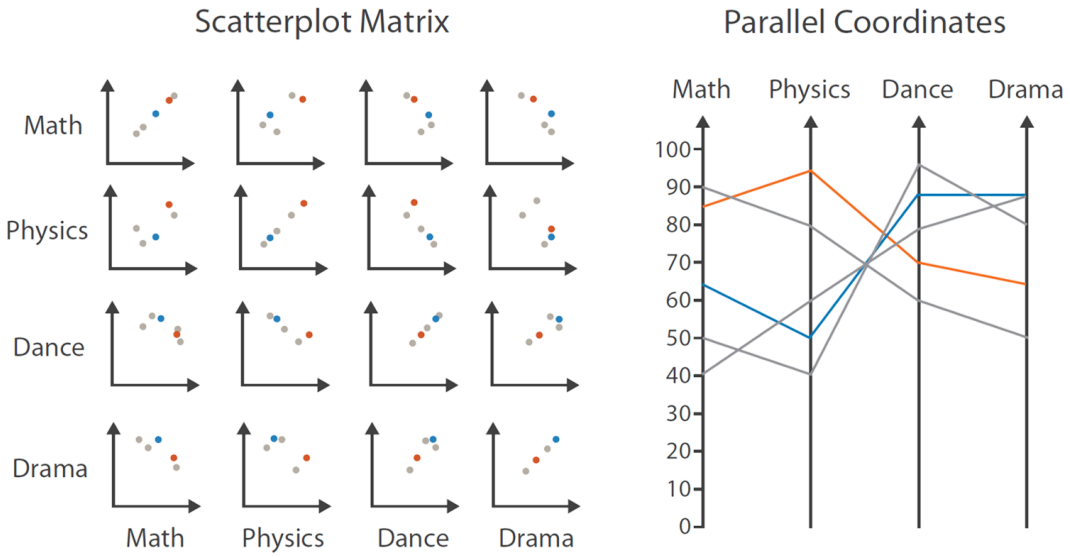
Express Values Quantitatively / Qualitatively



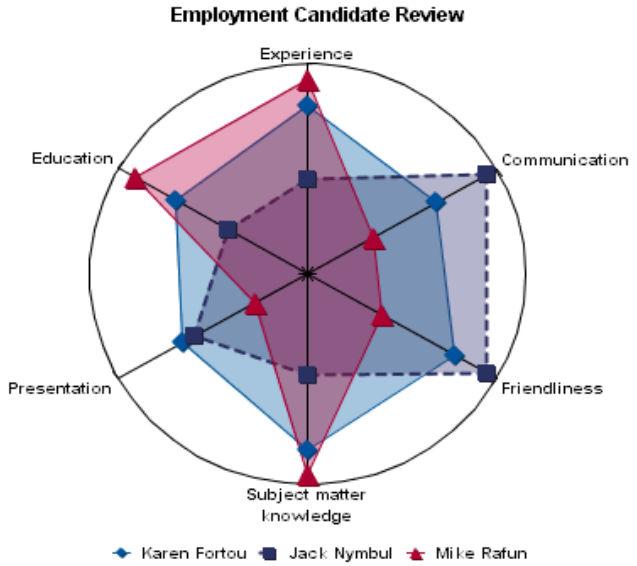
Life expectancy and infant mortality for different countries. Color encodes the country categories and size is population.
Robertson et al., 2008

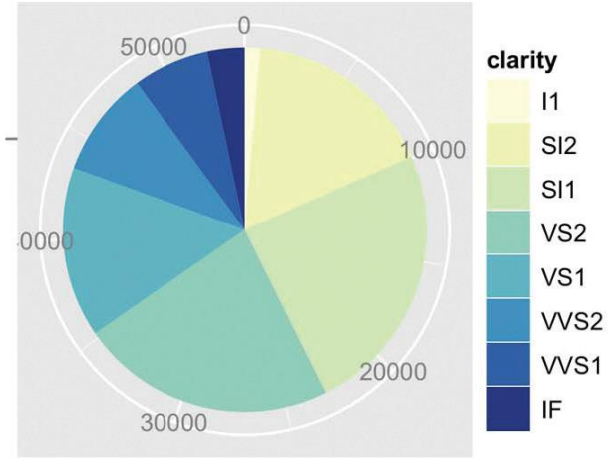
After Zacks and Tversky, 1999, by Munzner et al. 2014

Axis Orientation



From Munzner, Visualization Analysis & Design, 2014





Wickham 2010

Source of images and extended discussion: Link

Visualization Design

Abstract Data

Spatial Data

Interaction Concepts

Tabular Data

Networks and Trees

Deep Learning

Arranging Networks and Trees

Node-Link Diagram

- Marks represent links

Adjacency Matrix

- Shows adjacency relationships

Containment Channel

- Useful for tree structures

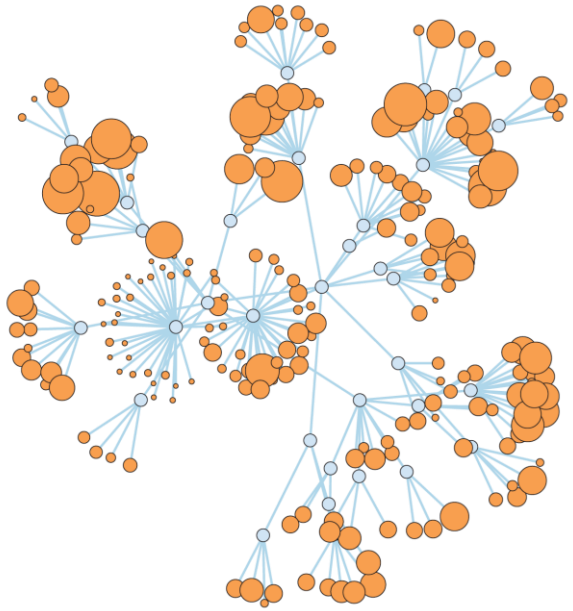
Node–Link Diagrams
Connection Marks
✔ NETWORKS ✔ TREES

Adjacency Matrix
Derived Table
✔ NETWORKS ✔ TREES

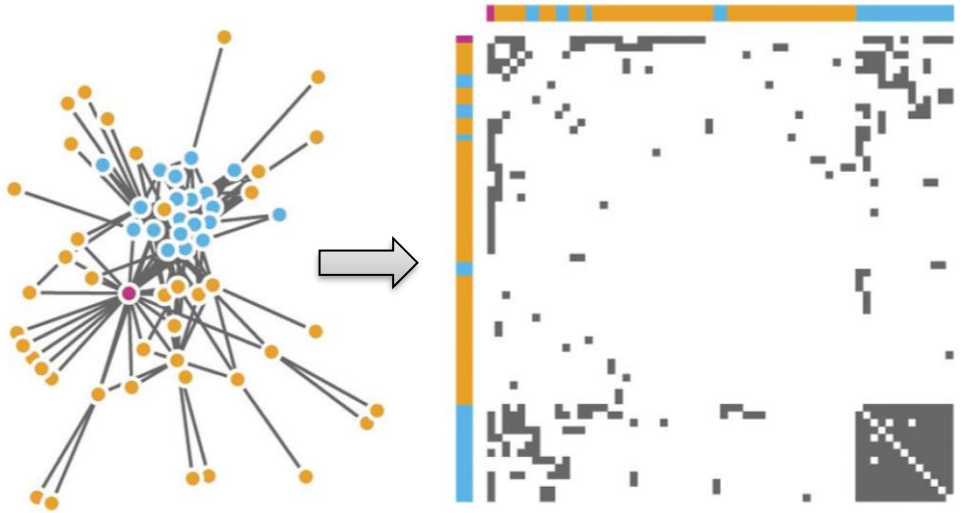
Enclosure
Containment Marks
✘ NETWORKS ✔ TREES

From Munzner, Visualization Analysis & Design, 2014

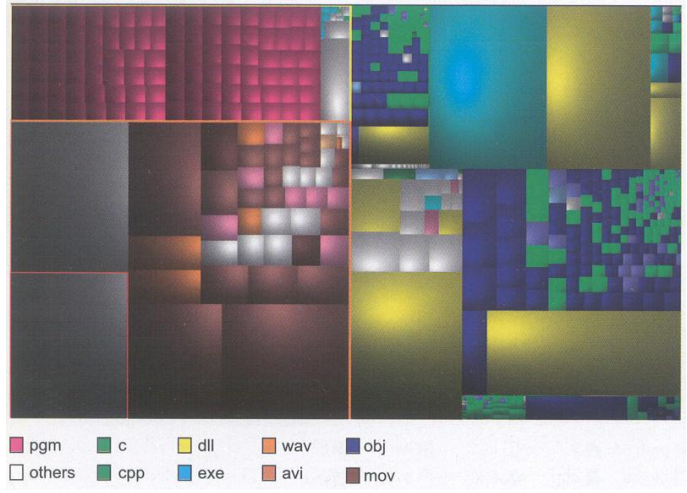
Networks and Trees



http://bl.ocks.org/1062288



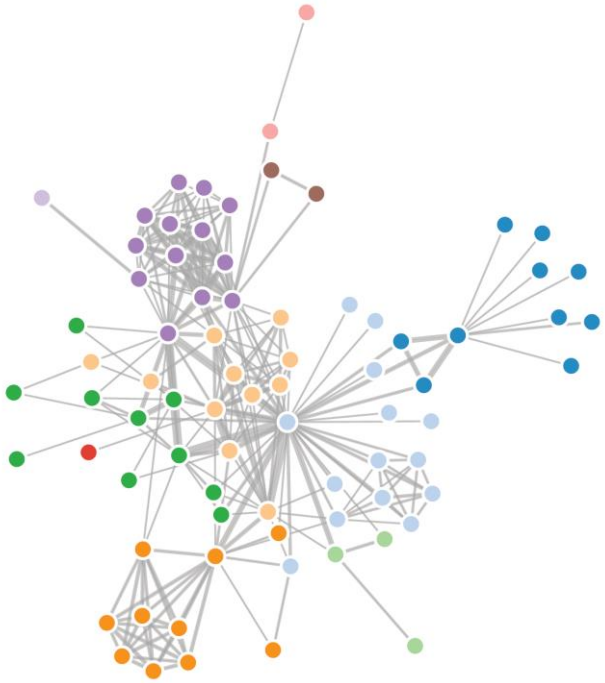
From Munzner, Visualization Analysis & Design, 2014



A. Telea, Data Visualization: Principles and Practice, AK Peters, 2008

General Layout Problem

- **Input**: Graph $G = (V, E)$
- **Output**: Clear and readable straight-line drawing of G



http://bl.ocks.org/mbostock/4062045

Drawing Aesthetics

- Adjacent vertices are close
- Non-adjacent vertices are far apart
- Edges are short, straight, and have similar length
- Densely connected parts (clusters) form communities
- As few crossings as possible
- Nodes distributed evenly

Optimization criteria are partially contradicting!

Algorithm

initial layout

residual threshold

max iterations

ForceDirected$(G = (V, E), p = (p_v)_{v \in V}, \epsilon > 0, K \in \mathbb{N})$

$t \leftarrow 1$

while $t < K$ **and** $\max_{v \in V} \|F_v(t)\| > \epsilon$ **do**

  foreach $u \in V$ **do**

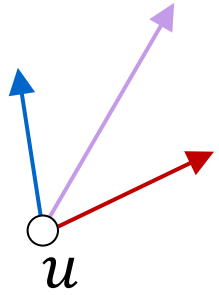
    $F_u(t) \leftarrow \sum_{v \in V} f_{\text{rep}}(u, v) + \sum_{uv \in E} f_{\text{attr}}(u, v)$

  foreach $u \in V$ **do**

    $p_u \leftarrow p_u + \delta(t) \cdot F_u(t)$

  $t \leftarrow t + 1$

return p



cooling factor

end layout

Visualization Design

Abstract Data

Spatial Data

Interaction Concepts

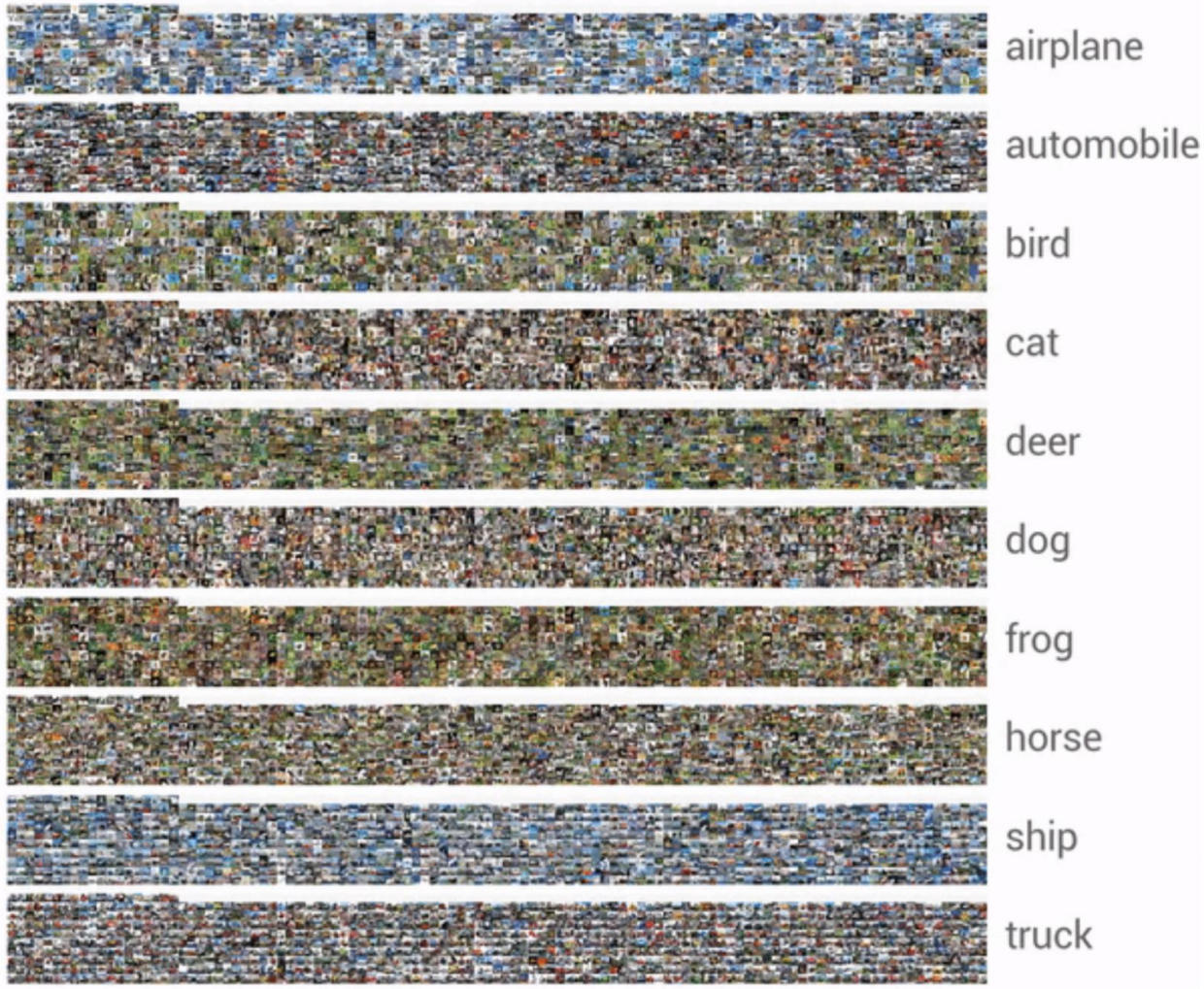
Tabular Data

Networks and Trees

Deep Learning

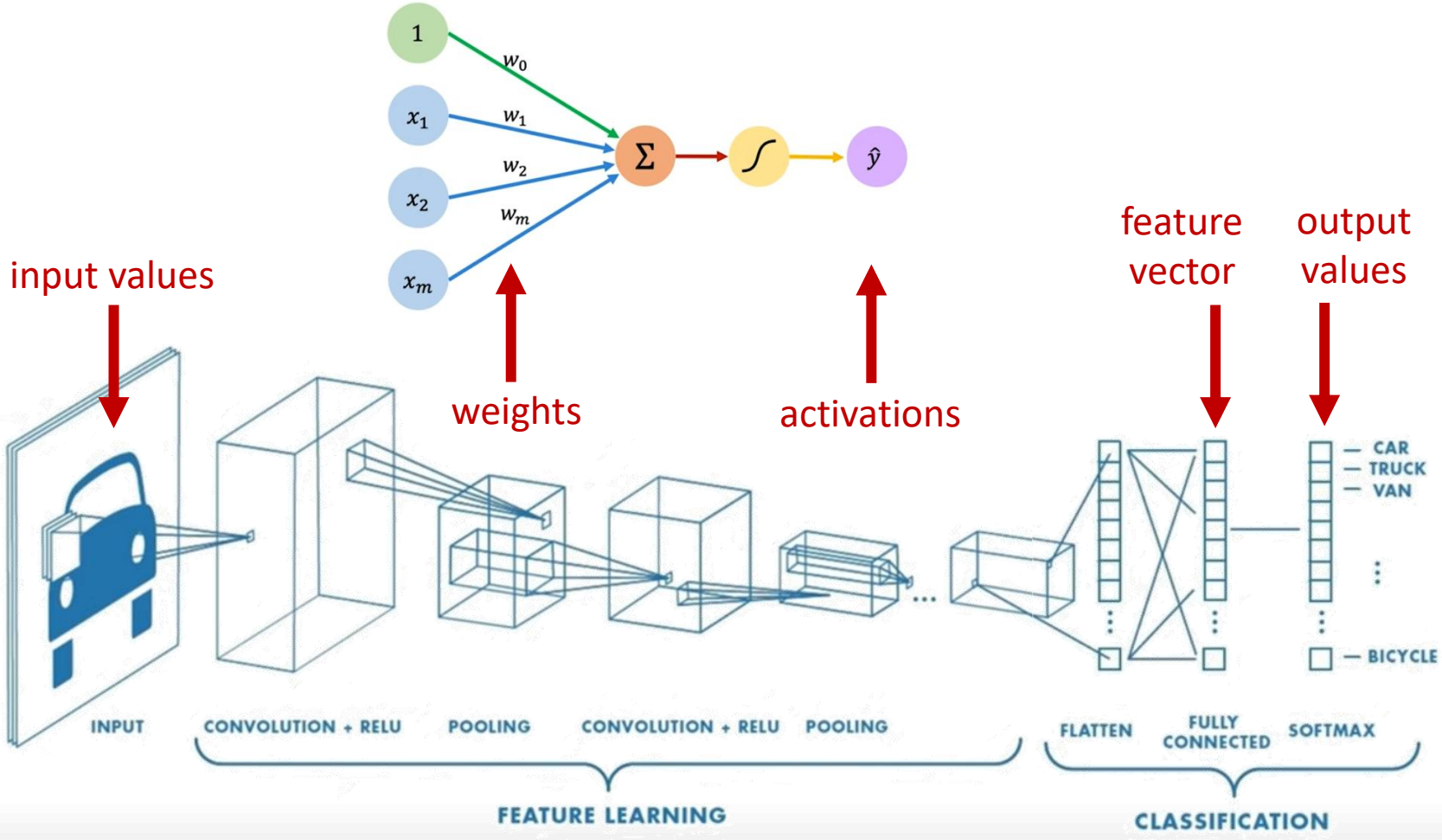
„Facets Dive“ by Google

- Overview visualization of all training data via small multiples
- Images are grouped by class
  - Overview of color schemes (hue) and similarities
  - Allows geometric zooming to see the detailed images
- Useful for visual inspection



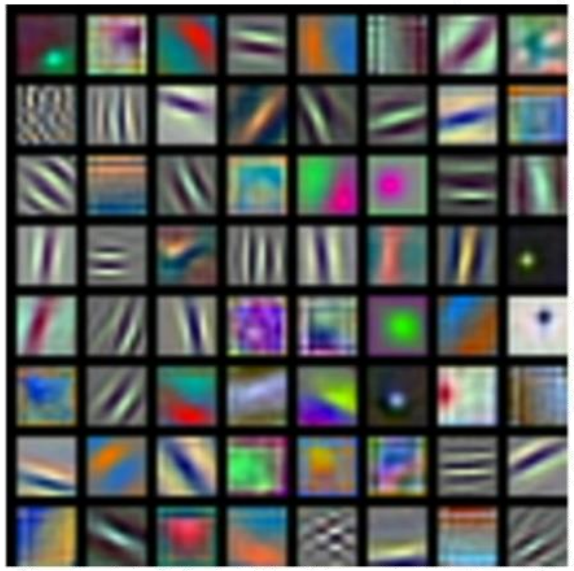
https://pair-code.github.io/facets/

Convolutional Neural Networks

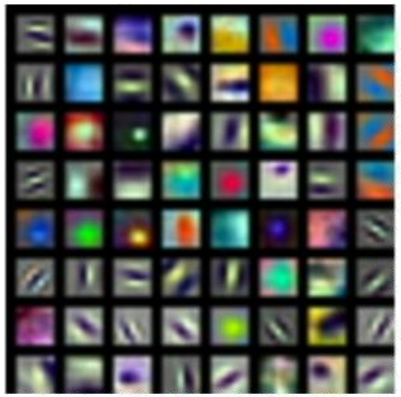


First Layer: Visualize the Learnt Weights of the Filters

- Here shown for a couple of pre-trained neural networks on Pytorch
- Weights are normalized to the $[0,1]$ range for coloring

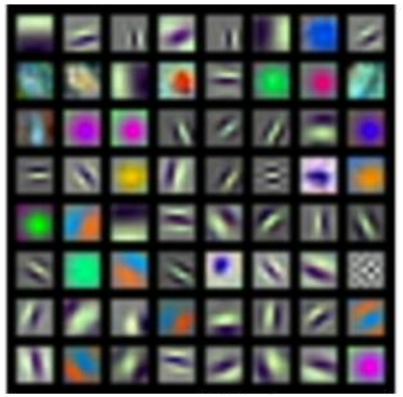


AlexNet:
64 x 3 x 11 x 11



ResNet-18:
64 x 3 x 7 x 7

ResNet-101:
64 x 3 x 7 x 7

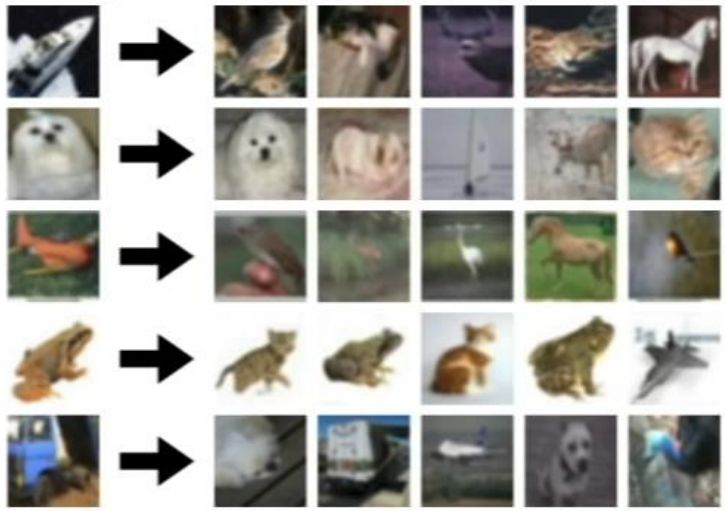


DenseNet-121:
64 x 3 x 7 x 7

Nearest Neighbors

- Find the L2 nearest neighbor in feature space

Nearest neighbors in pixel space



Test image L2 Nearest neighbors in feature space

Krizhevsky et al. „ImageNet Classification with Deep Convolutional Neural Networks“, NIPS 2012.

Visualization Design

Abstract Data

Spatial Data

Interaction Concepts

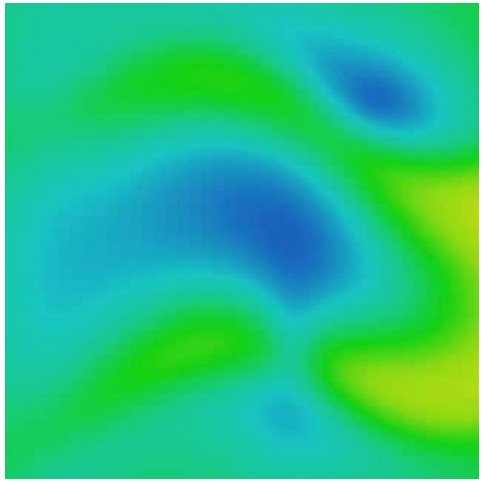
Scalar and Vector Fields

Medical Data

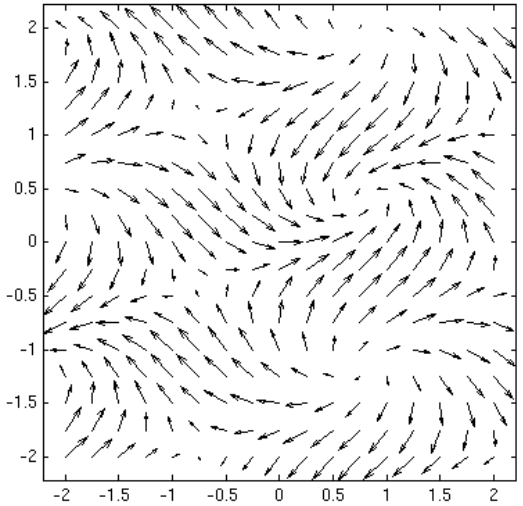
SciVis Tools

Fields (continuous data)

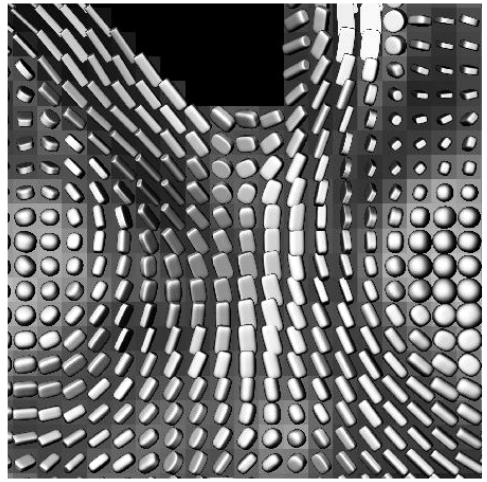
- Categorization by the structure of the output



Scalar field



Vector field



Tensor field

Image Courtesy by Tino Weinkauf.
Image courtesy by Tobias von Petersdorff, University of Maryland, Math 241 course notes.

Overview

Scalar Field Visualization

- Direct Volume Rendering
- Indirect Volume Rendering

} Volume visualization

Vector Field Visualization

- Elementary Methods / Direct Methods
- Geometry-based Methods
- Feature-based Methods (incl. topology)
- Dense / Texture-based Methods

} Flow visualization

Tensor Field Visualization

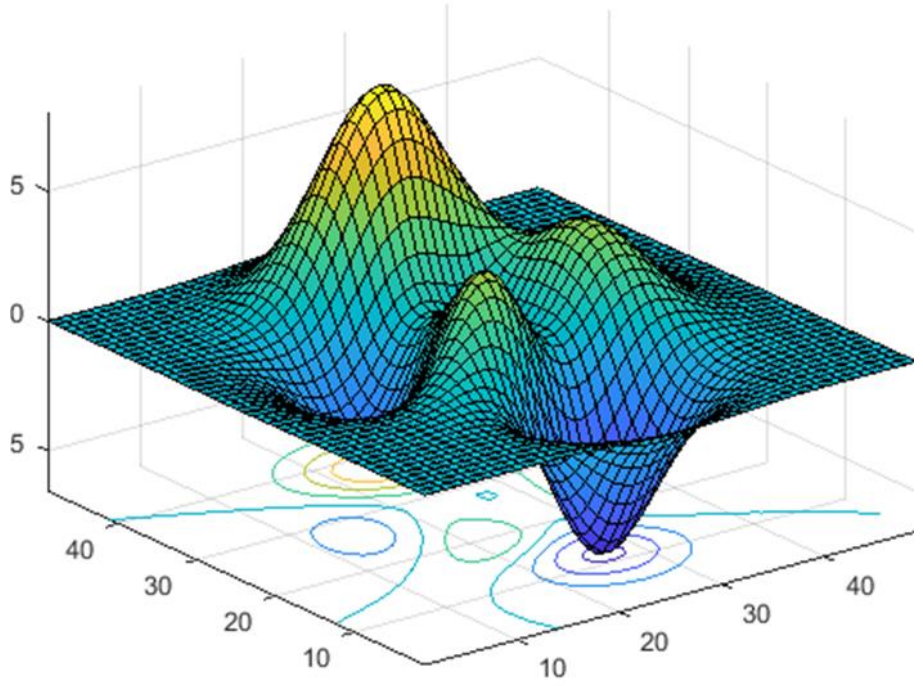
- Glyph-based methods
- Tractography

Isolines of scalar field $f(x, y)$

- Implicitly defined as curve fulfilling

$$f(x, y) = c$$

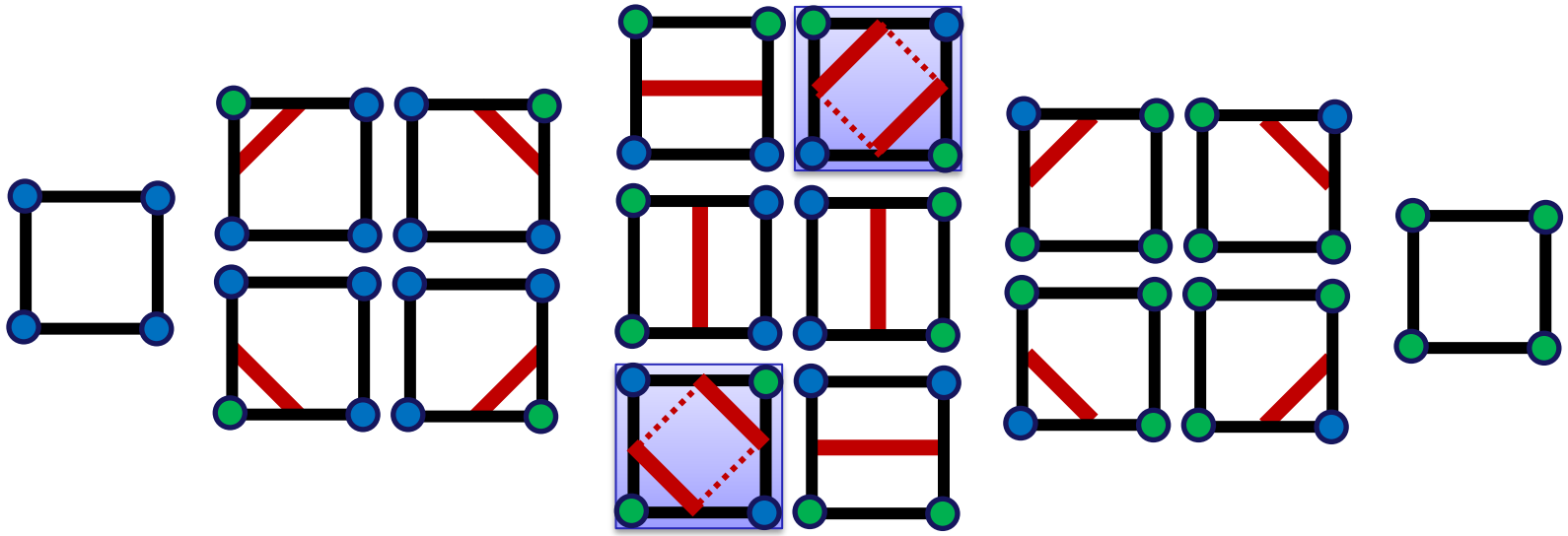
c: isovalue (iso = “equal” (Greek))



- Here, shown with
  - Height surface
  - Color-coding
- Also known as *level set*

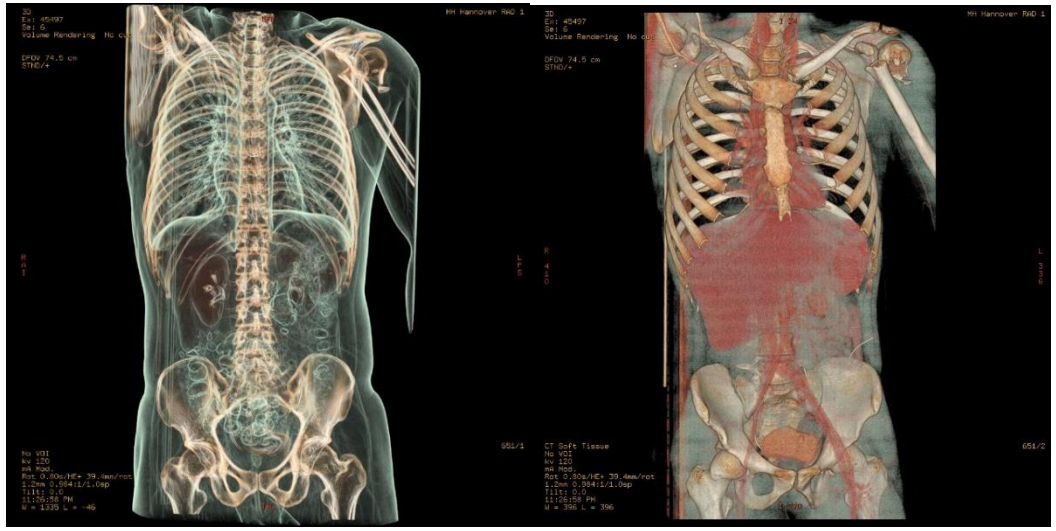
Matlab „surfc“ documentation
https://de.mathworks.com/help/matlab/ref/surfc.html

Combinations



Ray Marching

- Sample scalar field along ray
- Map via transfer function to color
- Blend colors together

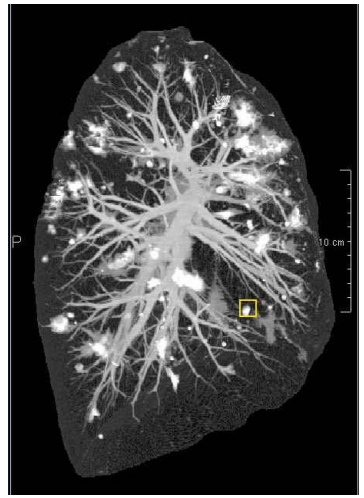


Courtesy of Hoen Oh Shin, Medical School Hannover, 2013

Alternative:

- Monte Carlo Rendering

Günther et al., MCFTLE: Monte Carlo Rendering of Finite Time Lyapunov Exponent Fields, CGF, 2016.



V. Dicken, et al. „Projektionsansichten zur Vereinfachung der Diagnose von multiplen Lungenrundherden in CT Thorax Aufnahmen“. Bildverarbeitung für die Medizin 2003: 244 248

Color-coding of scalar field behind arrow plot

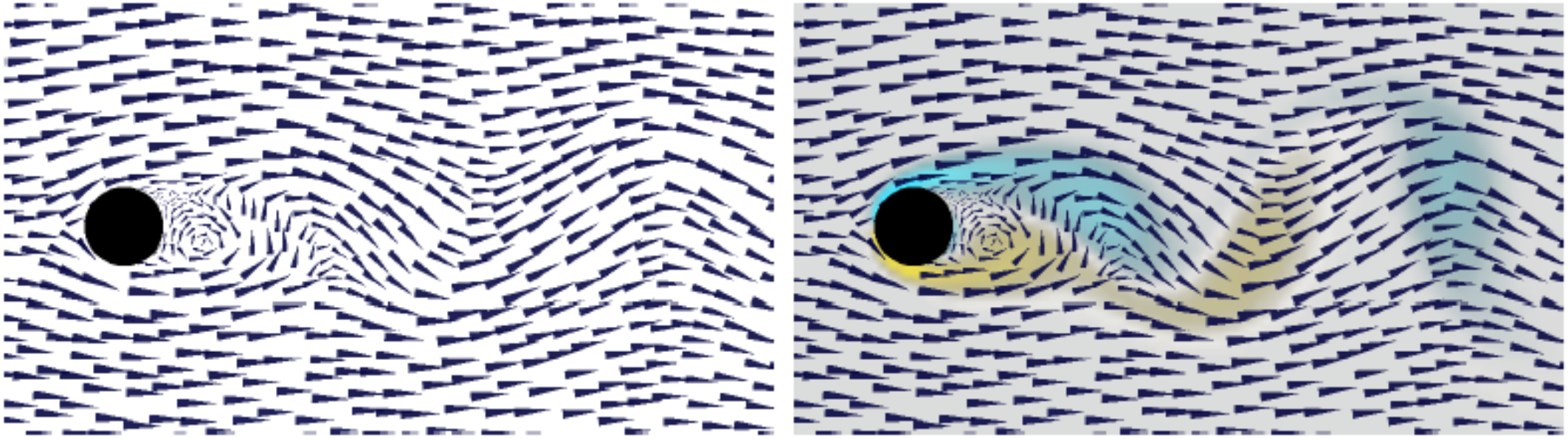


Figure 1: Typical visualization methods for 2D flow past a cylinder at Reynolds number 100. On the left, we show only the velocity field. On the right, we simultaneously show velocity and vorticity. Vorticity represents the rotational component of the flow. Clockwise vorticity is blue, counterclockwise yellow.

Kirby et al., Visualizing Multivalued Data from 2D Incompressible Flows using Concepts from Painting, Proc. IEEE Visualization, 1999.

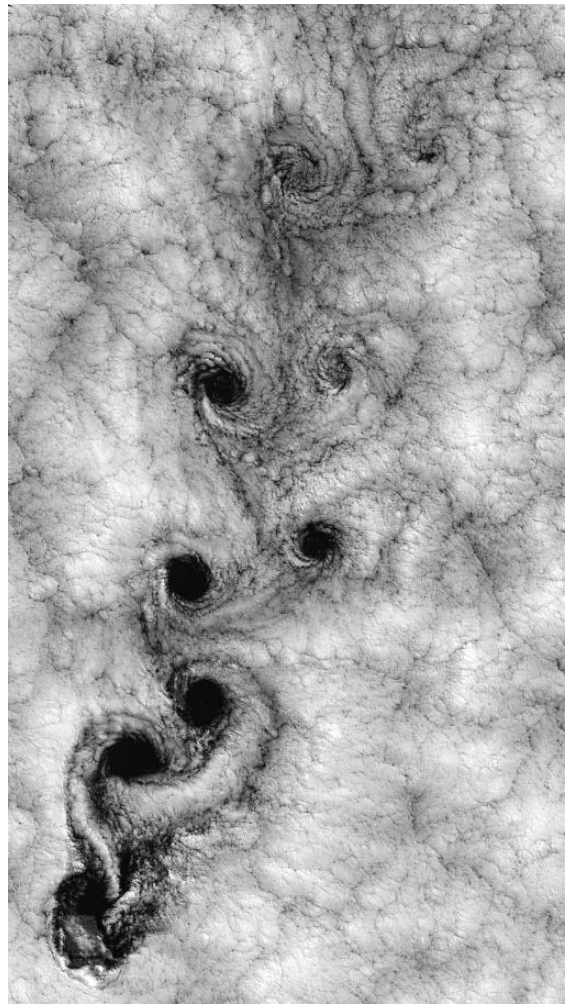
Streamlines

Streamline $\mathbf{x}(t)$ is tangential to vector field $\mathbf{u}(\mathbf{x})$

$$\frac{d\mathbf{x}(t)}{dt} = \mathbf{u}\big(\mathbf{x}(t)\big)$$

fluid flow

Particles in fluid [Tino Weinkauf]



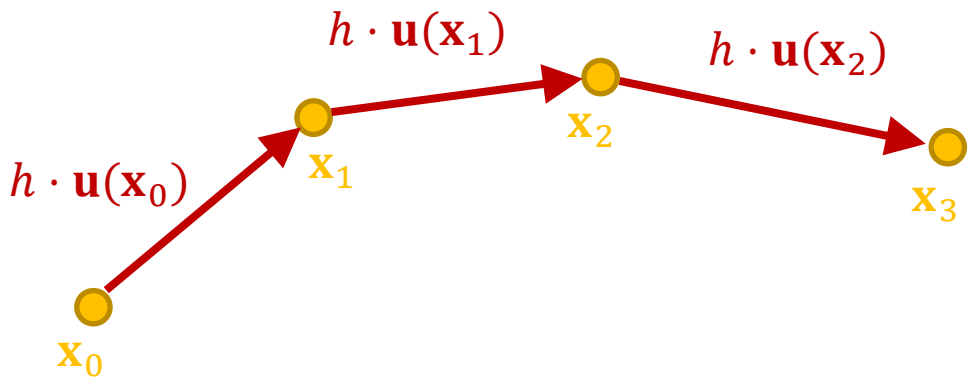
Juan Fernández Islands off the Chilean coast

Bob Cahalan, NASA GSFC, Public domain, via Wikimedia Commons

Euler integration:

$$\mathbf{x}_{i+1} = \mathbf{x}_i + h \cdot \mathbf{u}(\mathbf{x}_i)$$

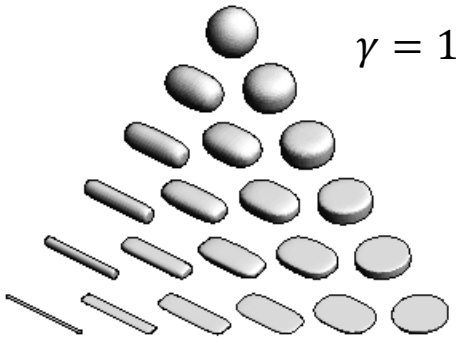
Step size: h

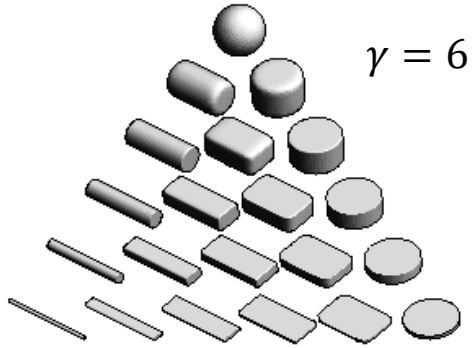


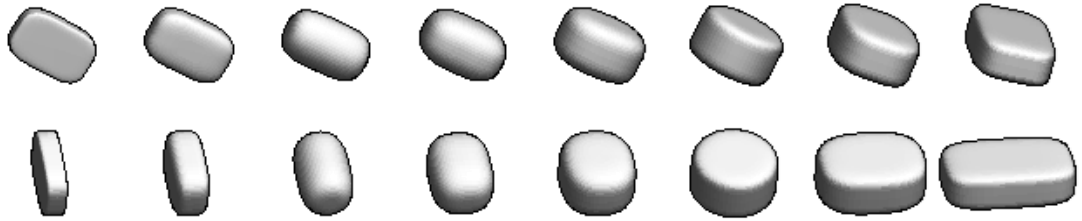

```
void Step(State& x, float h)
    x += h * Sample(x);
}
```

Kindlmann (2004) proposes to use superquadrics as glyphs

- Shape of the glyph depends on c_l and c_p
- Combines advantages of boxes and ellipsoids
  - Edges give strong visual cues for orientation
  - Round shapes are used when orientation is not clearly defined
- "Sharpness" parameter can be set depending on the level of noise in the data
- Leaves coloring and texture for encoding of additional information







Visualization Design

Abstract Data

Spatial Data

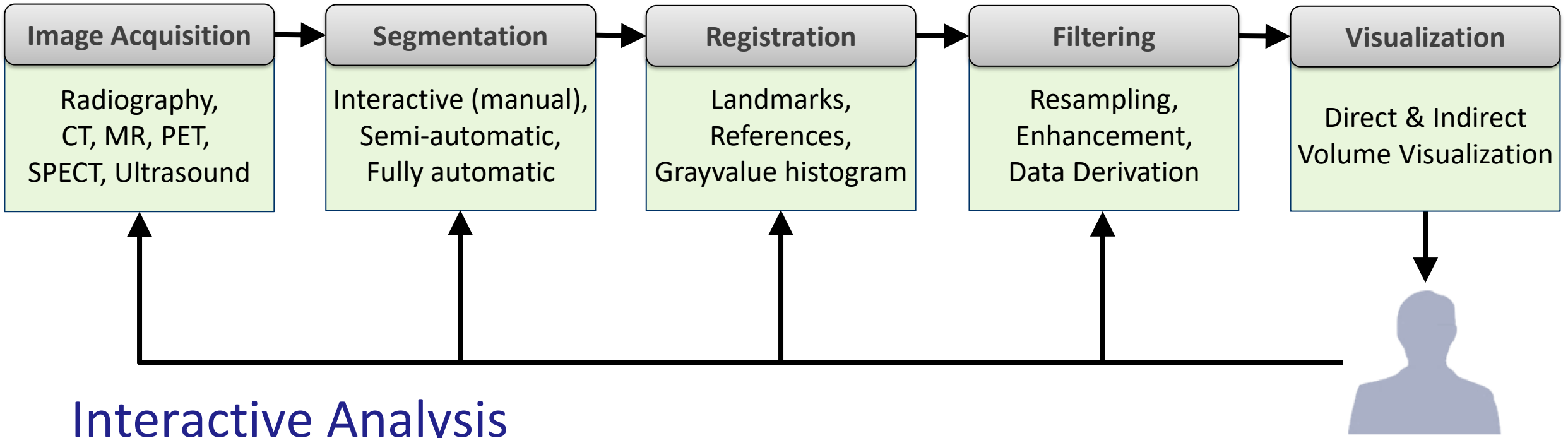
Interaction Concepts

Scalar and Vector Fields

Medical Data

SciVis Tools

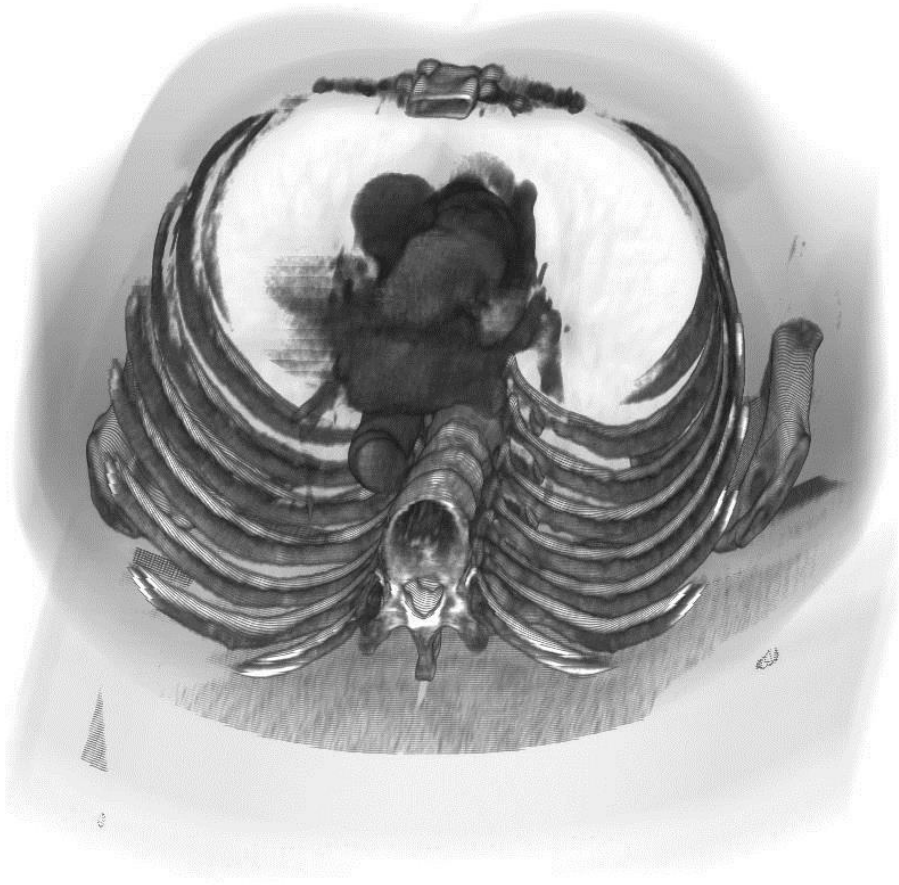
Medical Visualization Pipeline



Interactive Analysis

- Maybe image not good enough (collect more images)
- Segment further structures (or refine segmentation)
- Registration not good enough (try other algorithm, other parameters)
- …

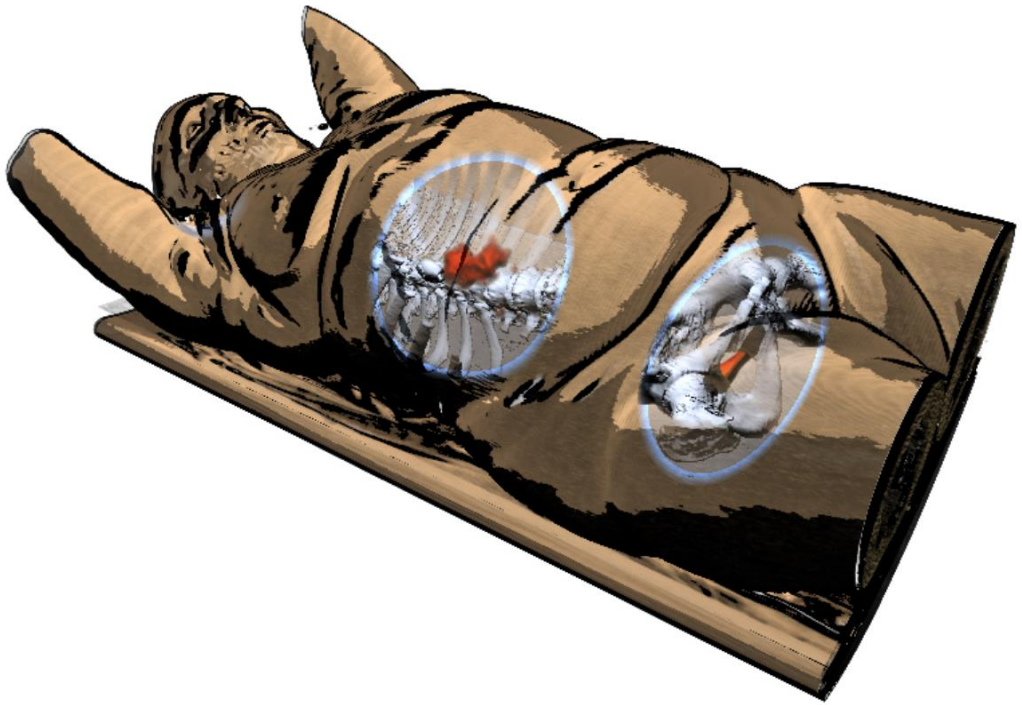
Volume Visualization

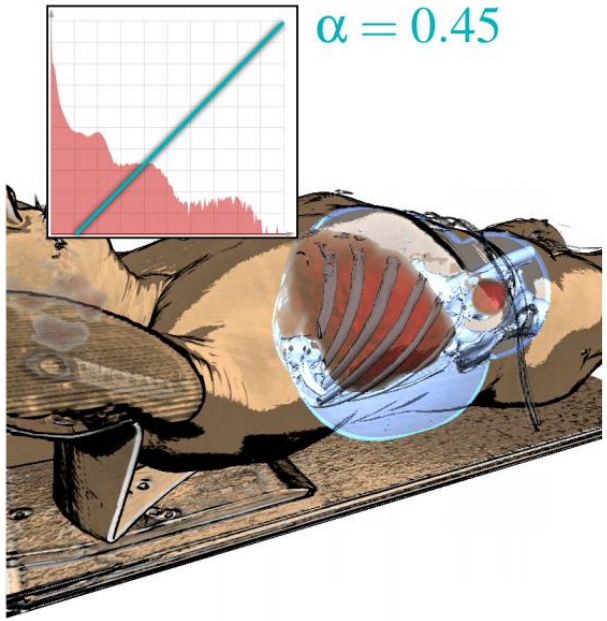


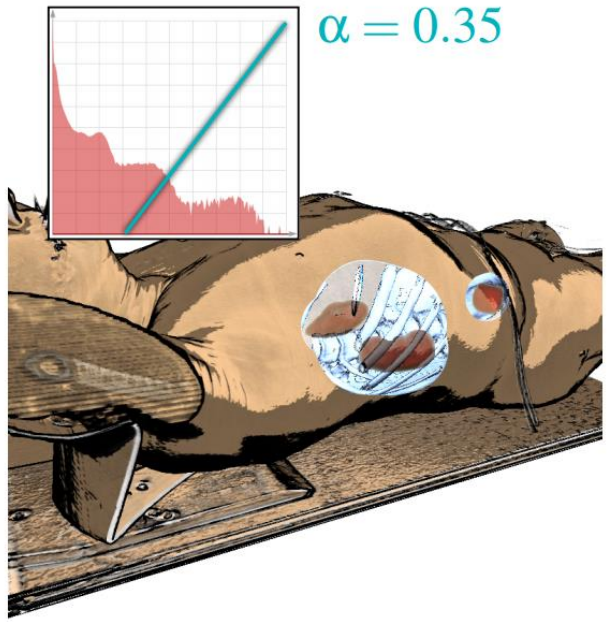
CT of the lung (250 layers, 1024x1024 pixel each)

Focus and context visualization combining PET/CT

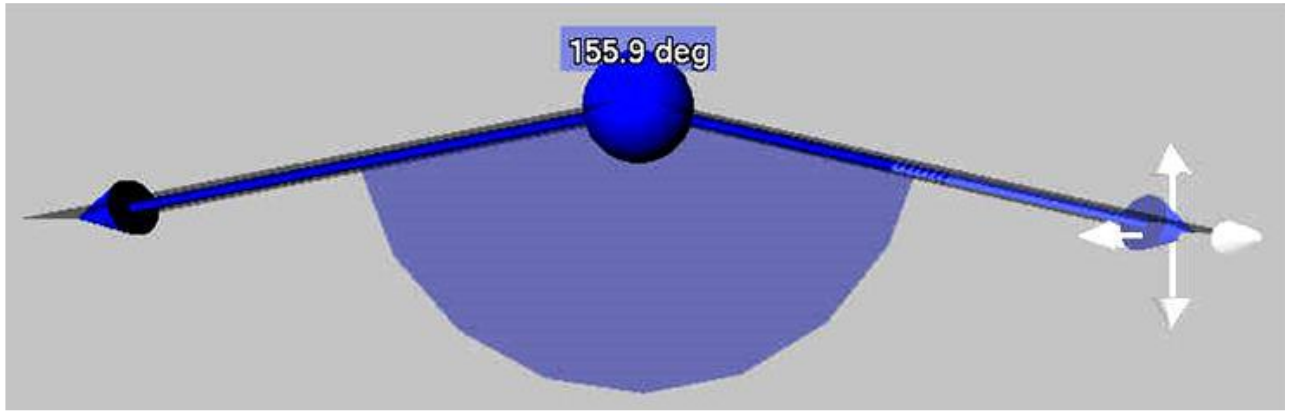
- PET shows abdominal lymphoma

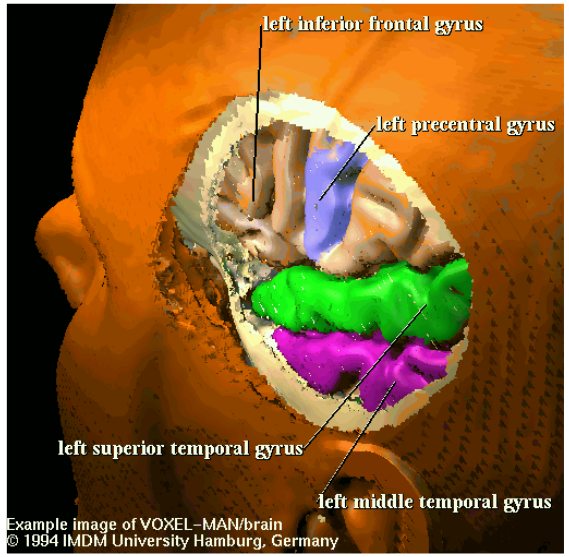


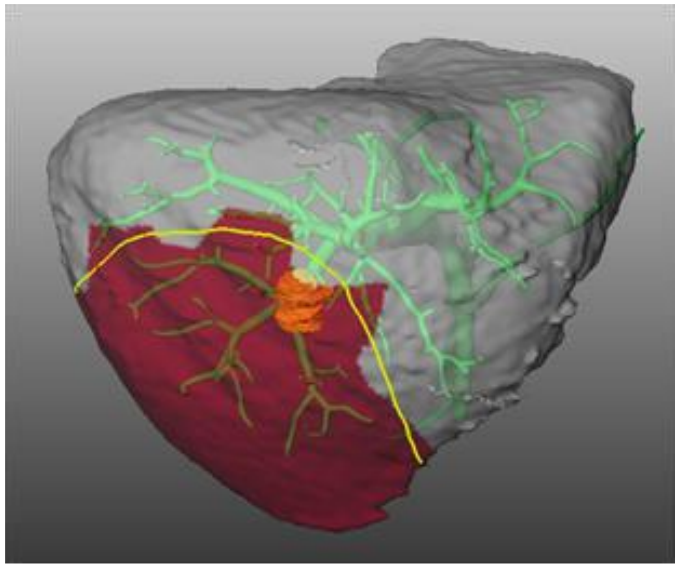


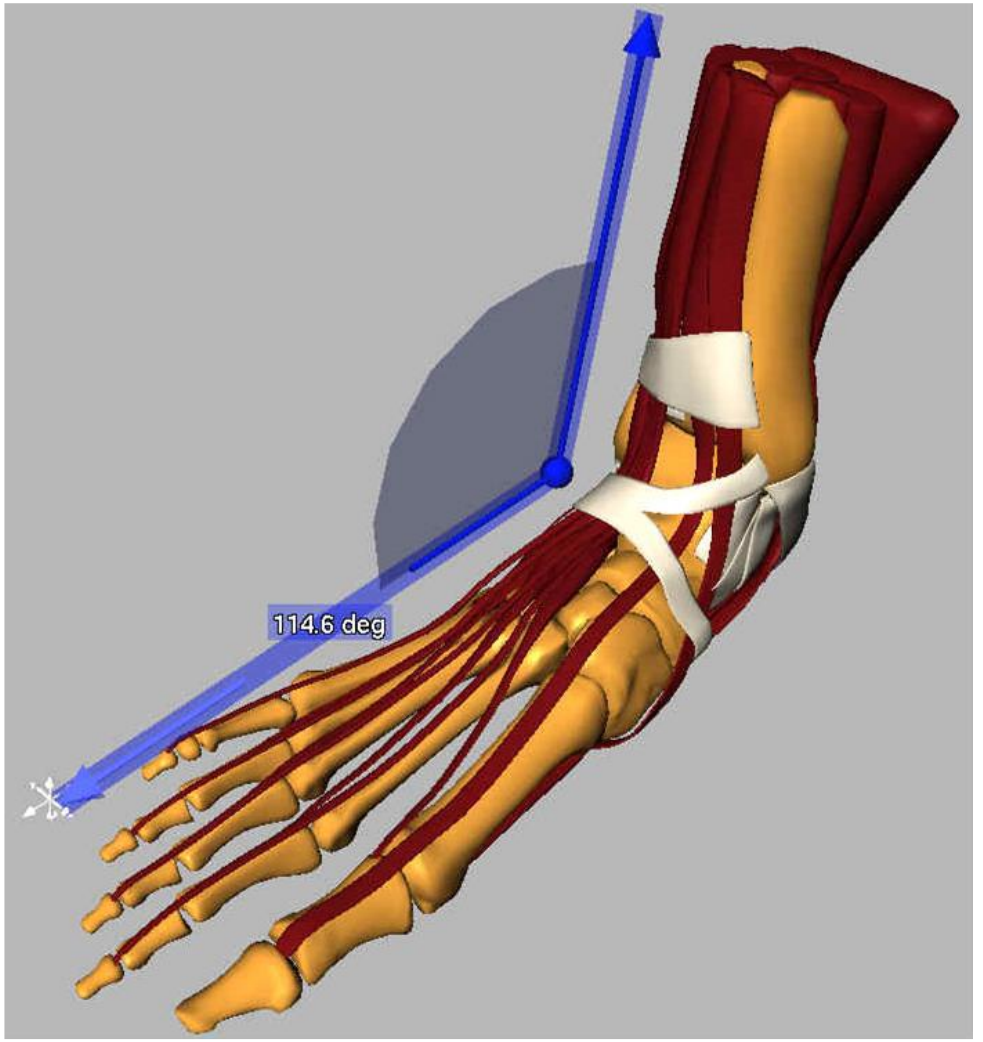


Interaction Techniques

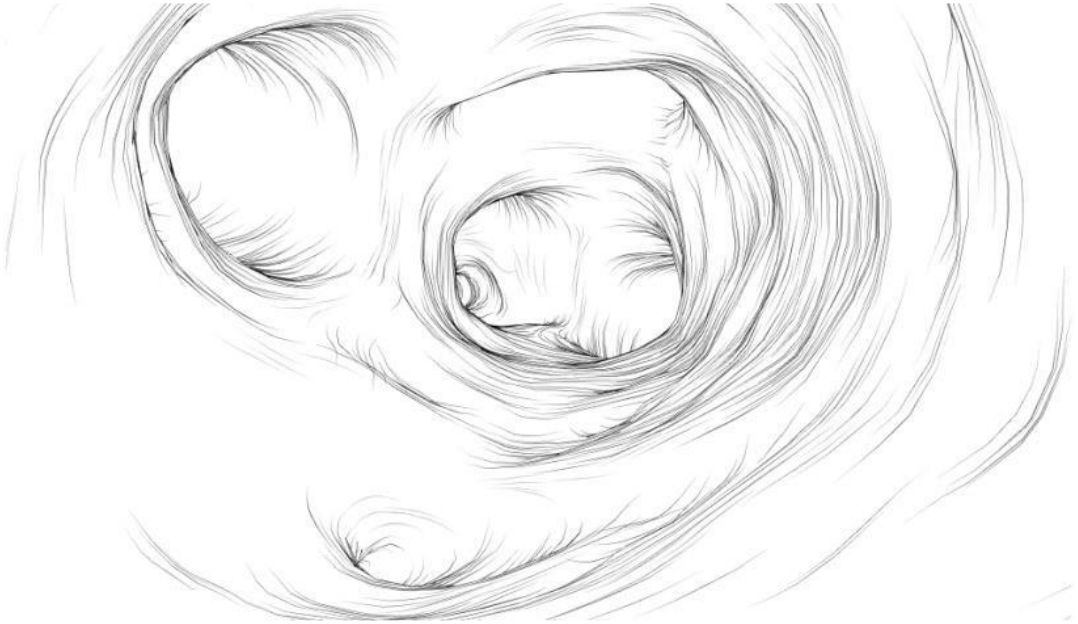




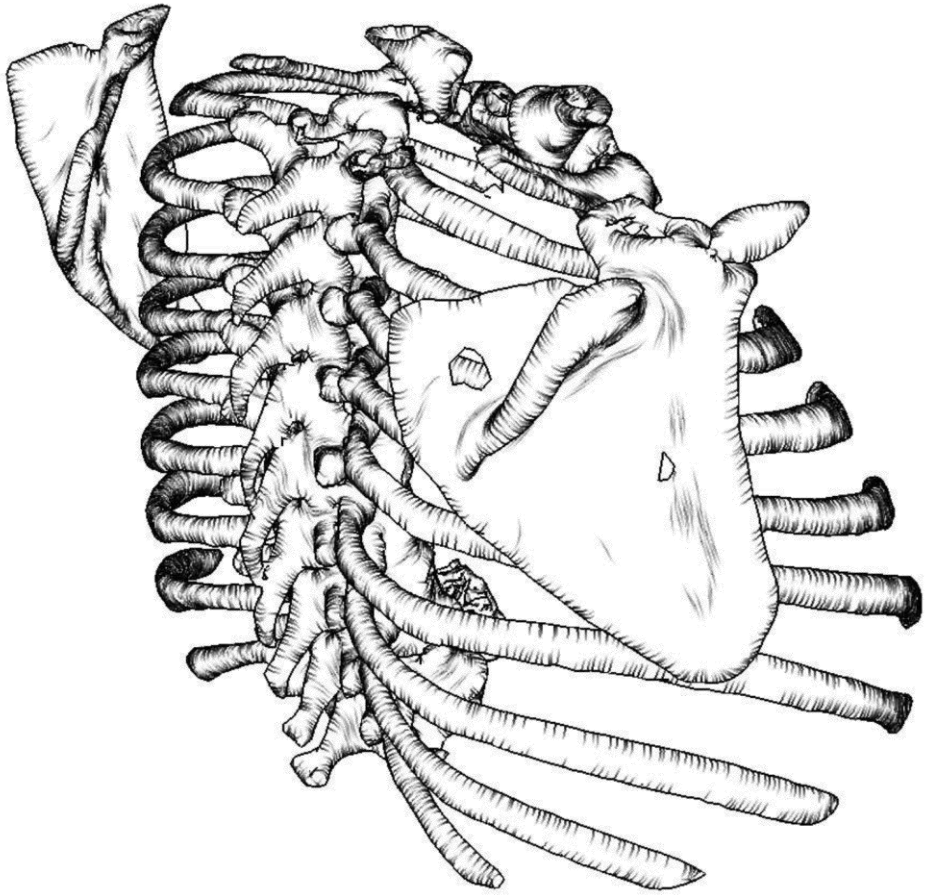




Combination of silhouette rendering and hatching



Bronchoscopic view



Lawonn et al., EuroVis, 2013.

Visualization Design

Abstract Data

Spatial Data

Interaction Concepts

Scalar and Vector Fields

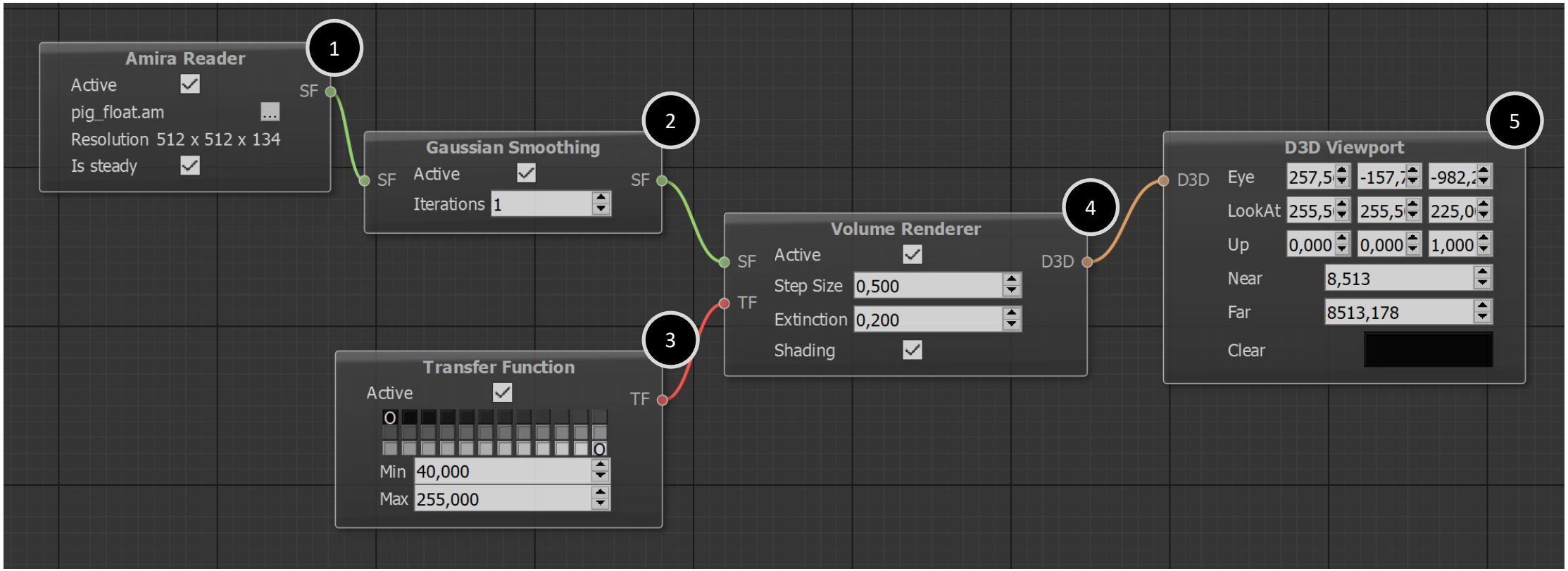
Medical Data

SciVis Tools

Scientific Visualization Tools

- How is scientific visualization software implemented?
  - Design patterns, data structures, …
- Which tools/frameworks exist?
- VTK/ParaView introduction
- File formats


Update Order: Topological Sorting



Screenshot of VisLab by Tobias Günther

Visualization Design
Abstract Data
Spatial Data

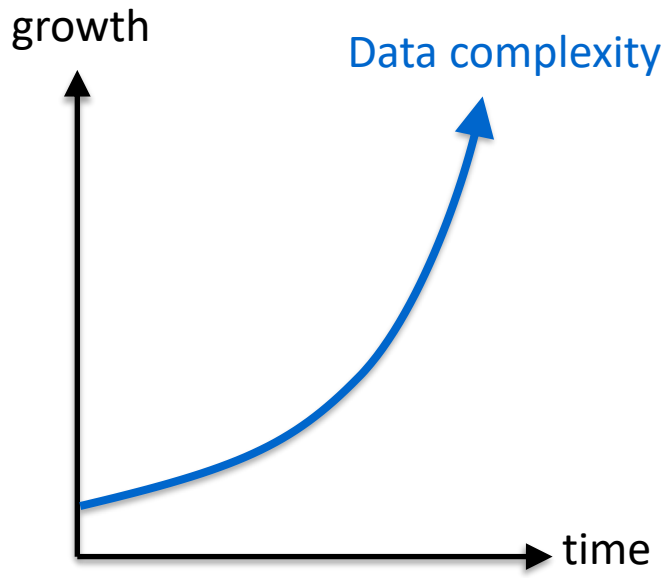
Interaction Concepts

View Manipulation
Data Reduction
Rules of Thumb

Scalability Problem

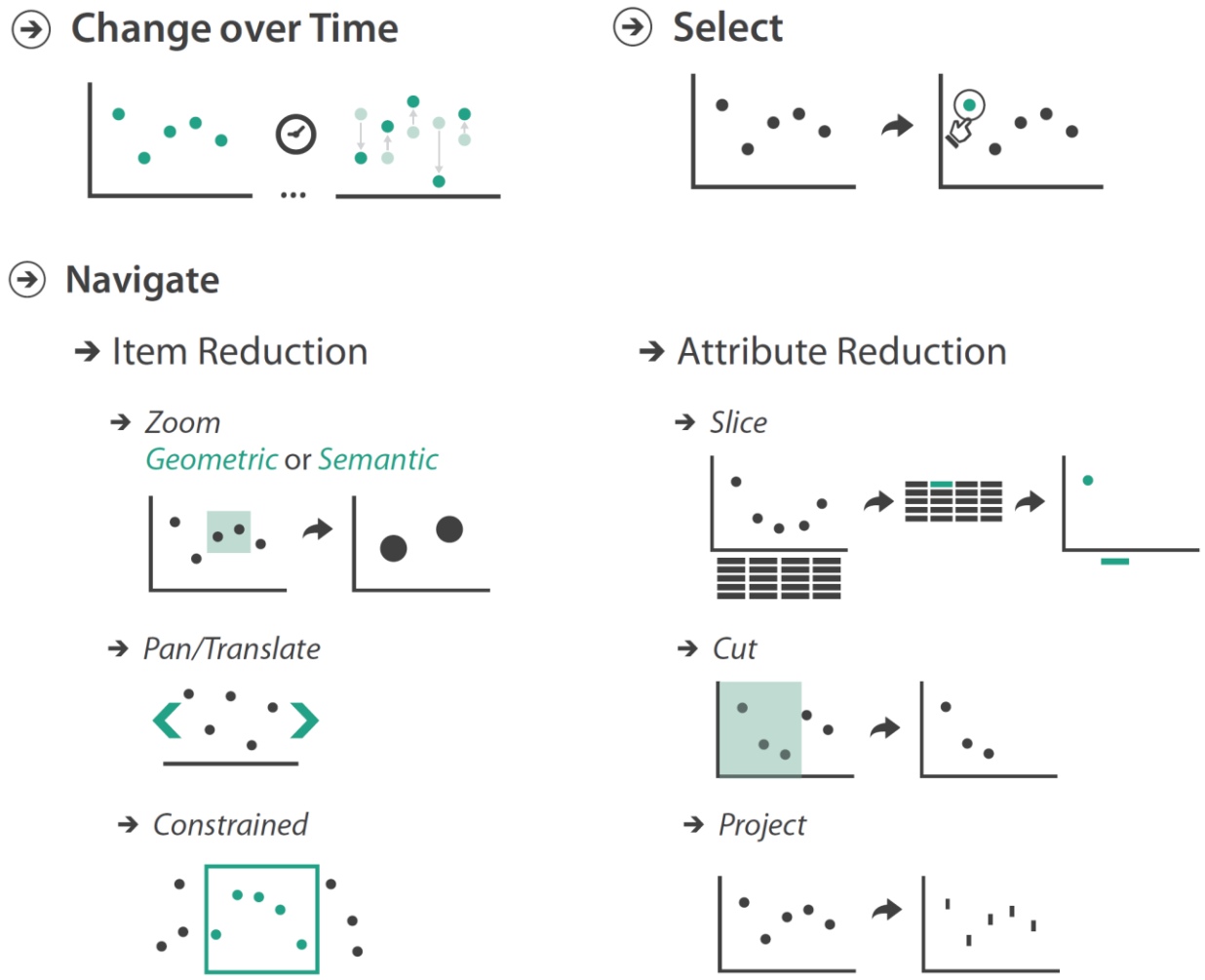
- We need to address the **scalability problem**:
  - Data is often too large to be shown at once

- **Five options** to address increasing complexity
  1. Derive new data
  2. Change the view
  3. Facet the data into multiple views
  4. Reduce the data
  5. Embedding of focus and context in single view



Design Choices

- **Change over Time**
  - Change of encoding
  - Reordering
- **Selection**
  - How many selectable?
  - Empty sets allowed?
- **Navigation**
  - Zooming / panning
  - Slicing

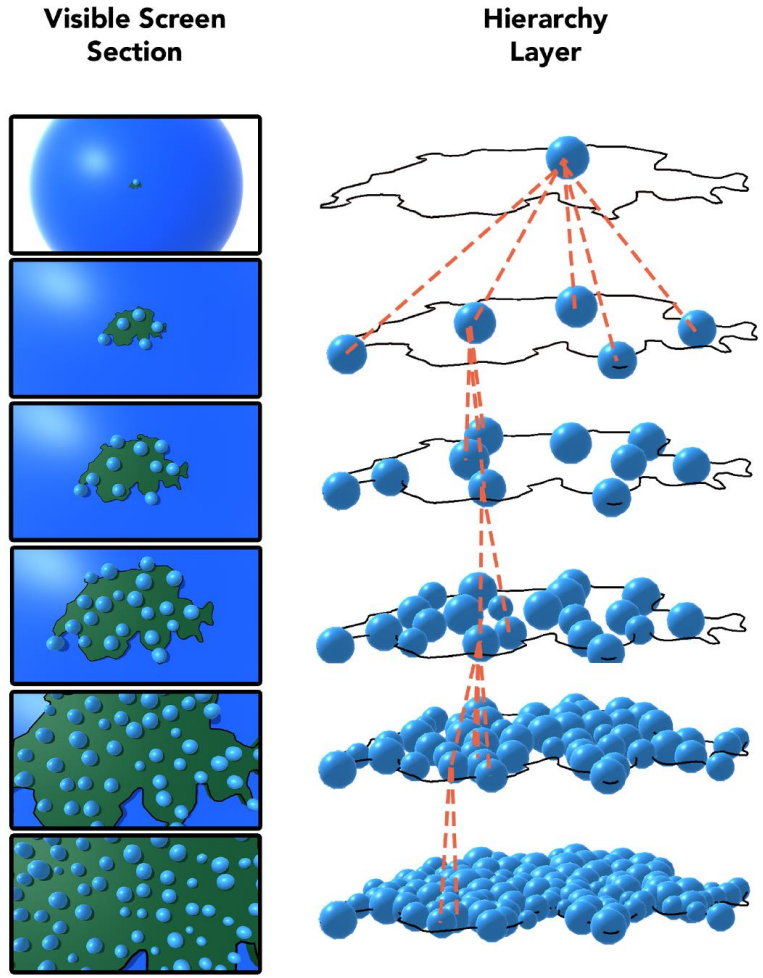


From Munzner, Visualization Analysis & Design, 2014

Navigation

- **Changing Viewpoint**
  - *Geometric Zooming:*
    - Move object closer to camera
  - *Semantic Zooming:*
    - Representation of object adapts to the available screen space

→ *Zoom*
Geometric or *Semantic*



Level of detail

Ingulfsen et al., „News Globe: Visualization of Geolocalized News Articles“. IEEE Compututer Graphics & Applications, 2021

Design Choices

- **Juxtapose and Coordinate Multiple Side-by-Side Views**
- **Partition Data**
  - Show subset of data in multiple views, also called *small multiples*
- **Superimpose Layers**
  - Blend multiple layers on top of each other

Juxtapose and Coordinate Multiple Side-by-Side Views

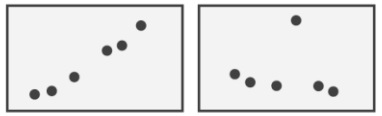
→ Share Encoding: Same/Different

→ *Linked Highlighting*

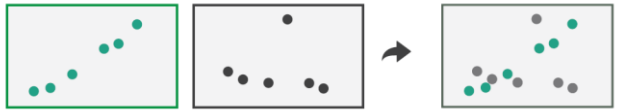
→ Share Data: All/Subset/None

→ Share Navigation

Partition into Side-by-Side Views



Superimpose Layers

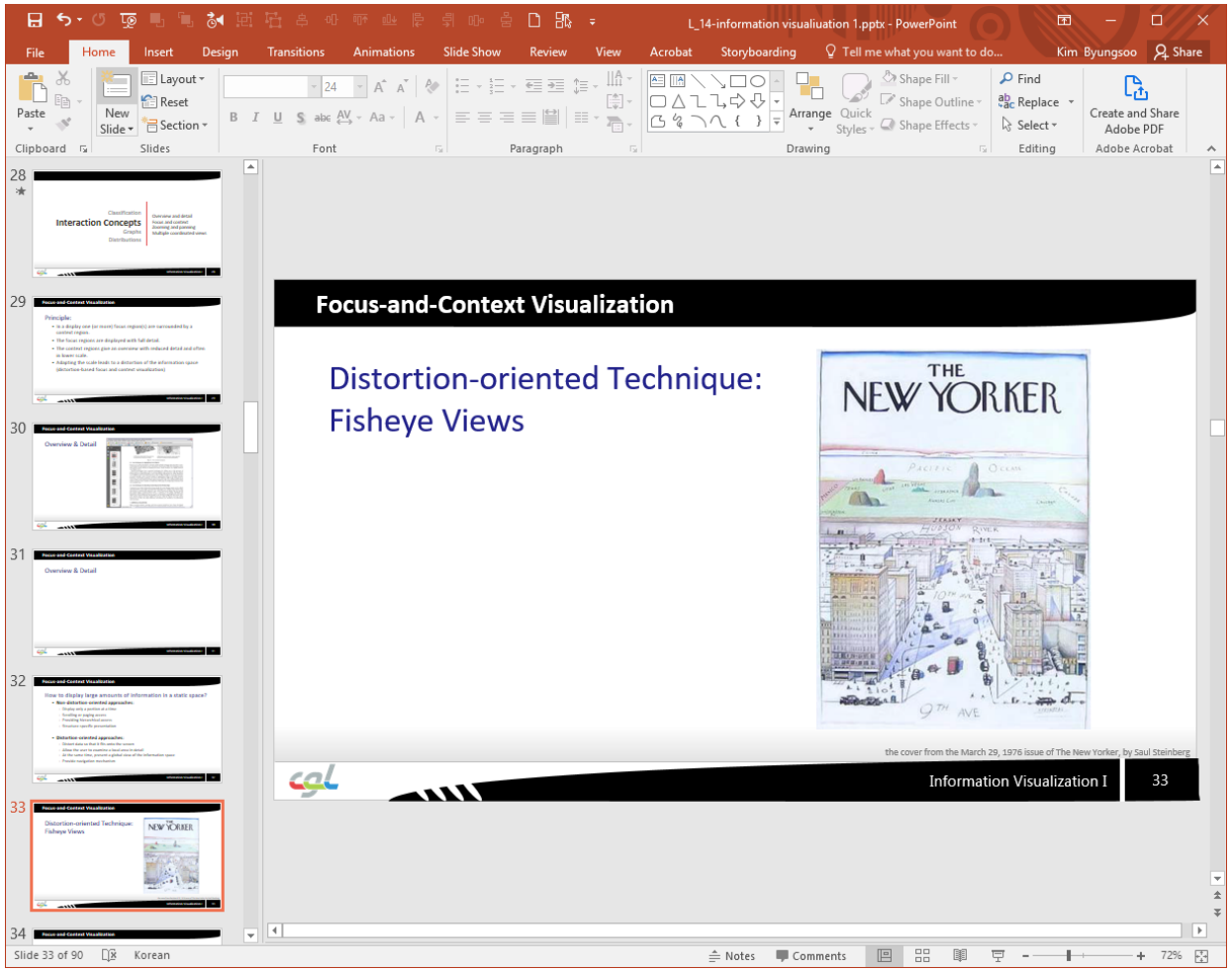


From Munzner, Visualization Analysis & Design, 2014

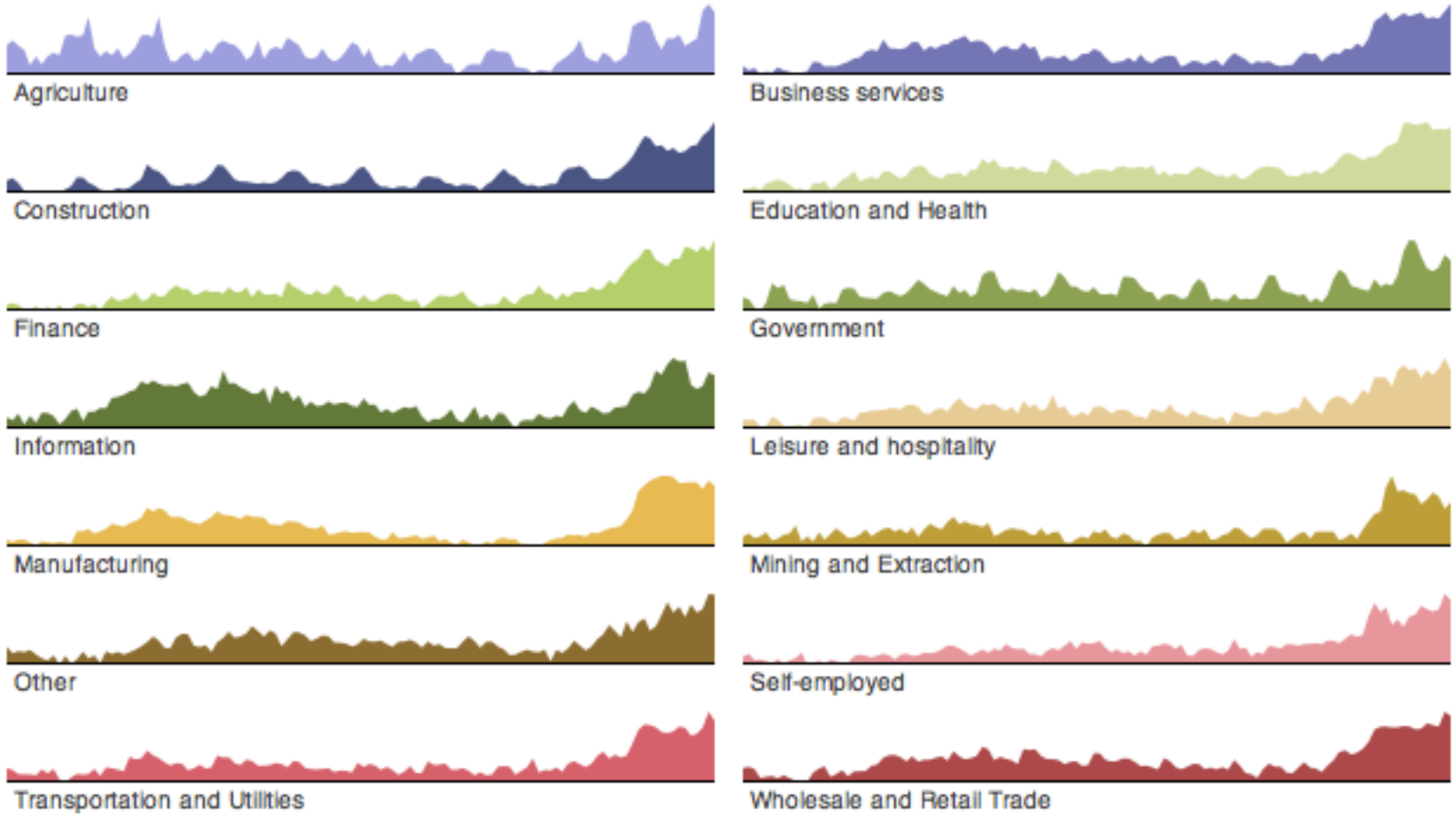
Facet into Multiple Views

Overview + Detail:

- A complete visualization serves as overview and is the “entry point” of the visualization
- One or more partial visualization then provide more detail for a selected region of interest



Facet into Multiple Views

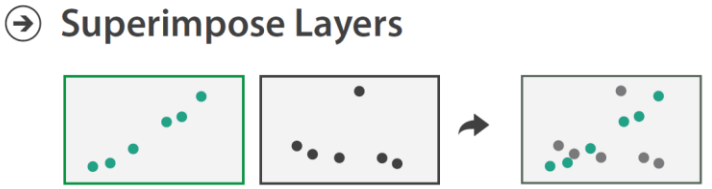


The term „Small multiples“ goes back to [Tufte, 1983]. A synonym is „Trellis display“ [Cleveland, 1997]

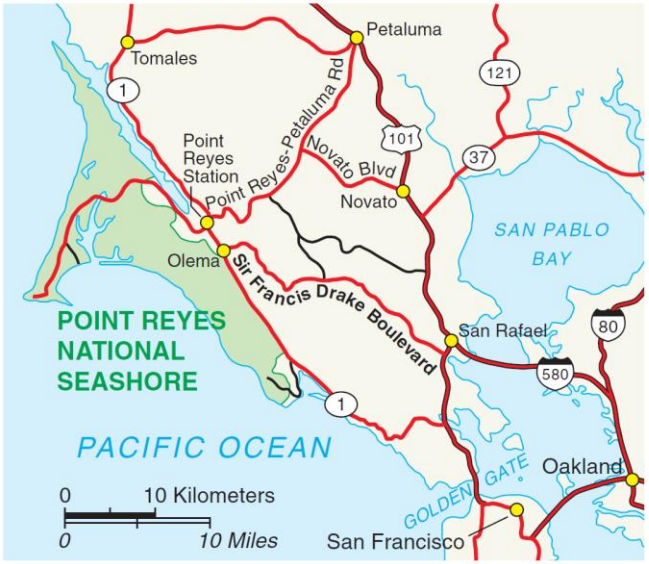
Superimpose Layers

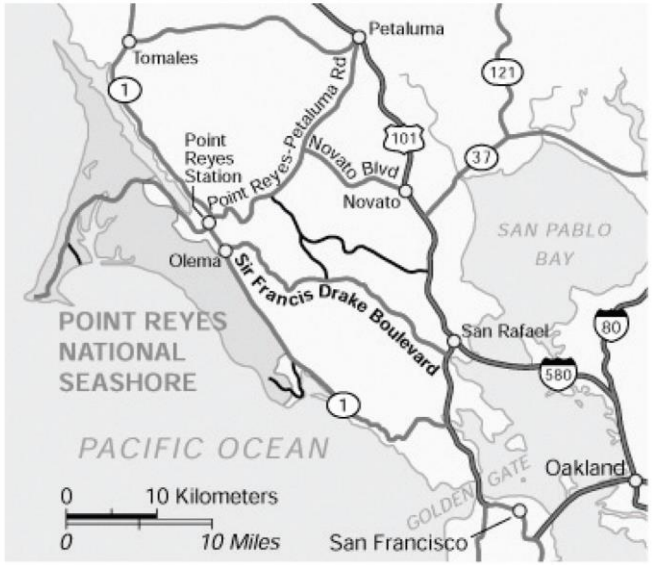
- **Static Layers**
  - All layers are always visible
  - Guide attention with visual channels

- **Example: geographic map**
  - Gray value image shows that each layer uses different luminance range



Idiom	Cartographic Layering
What: Data	Geographic
How: Encode	Area marks for regions (water, parks, other land), line marks for roads, categorical colormap.
How: Facet	Superimpose: static layers distinguished with color saturation, color luminance, and size channels.





Stone 2010

Visualization Design
Abstract Data
Spatial Data

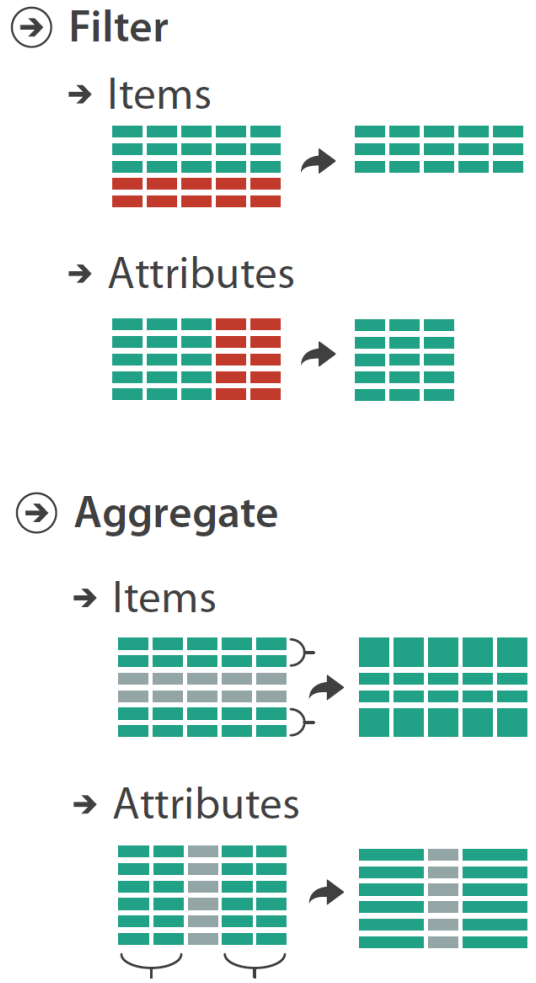
Interaction Concepts

View Manipulation
Data Reduction
Rules of Thumb

Design Choices

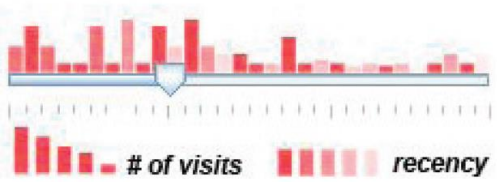
- **Filter**
  - Remove elements
  - Careful because „out of sight, out of mind“

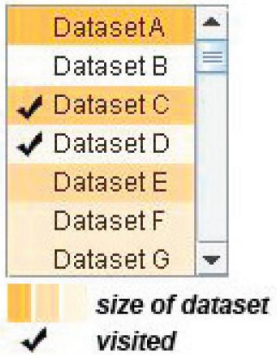
- **Aggregate**
  - Create new element that stands in for multiple elements
  - Stand-in information should convey a summary

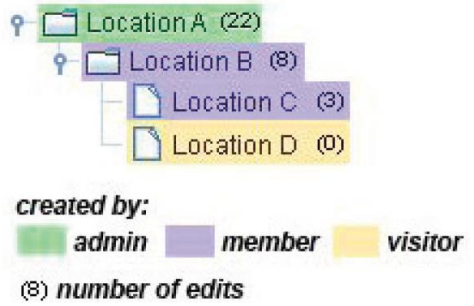


Scented Widgets

- **Add information** about the data distribution to the widgets
- Give a sense whether it makes sense to look deeper or go elsewhere
- Options:
  - Add statistical information
  - Add Icons/glyphs
  - Treat items as marks



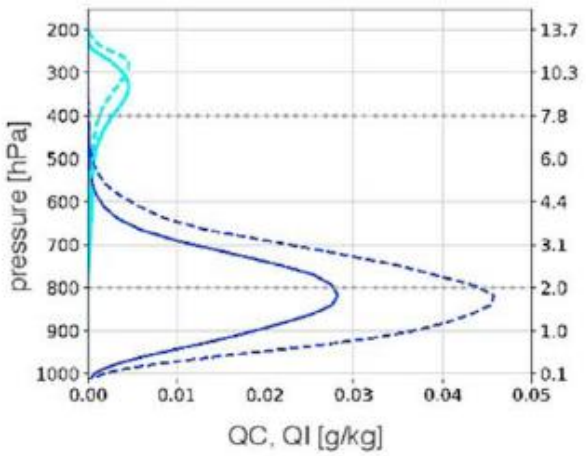




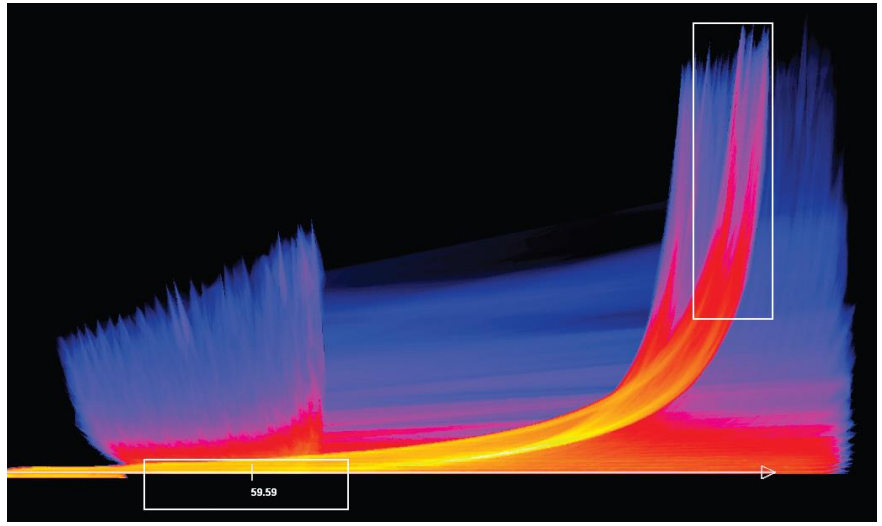
Willett et al. 2007, Scented Widgets

Item Aggregation

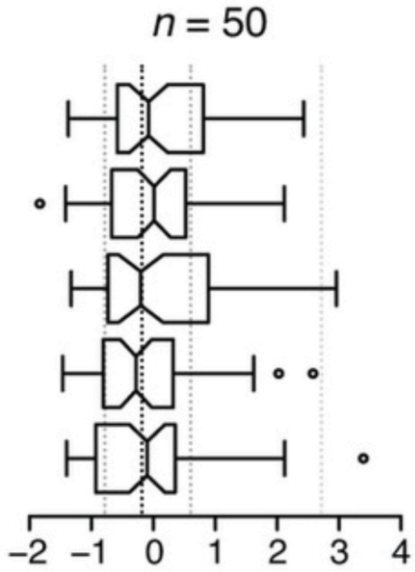
- Histograms
- Continuous scatterplots
- Box plots
- …



Hentgen et al. 19



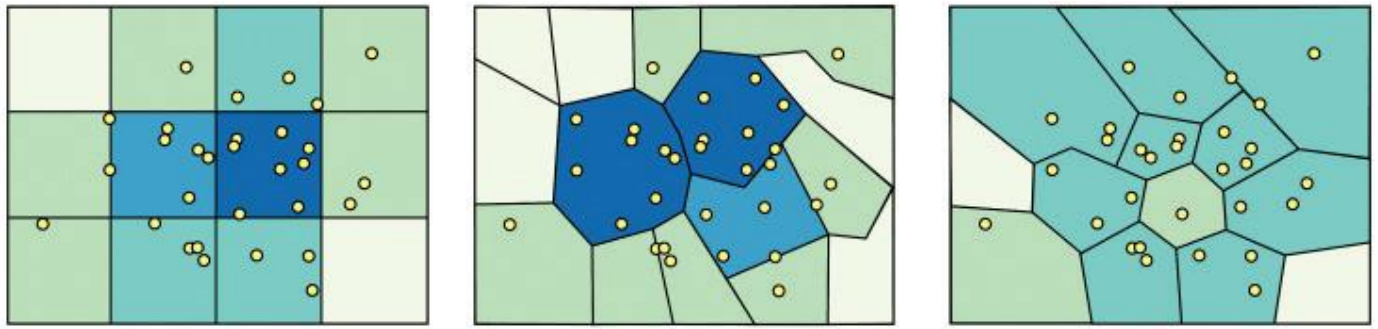
Bachthaler and Weiskopf, 2008



Wickham and Stryjewski, 2012

Spatial Aggregation

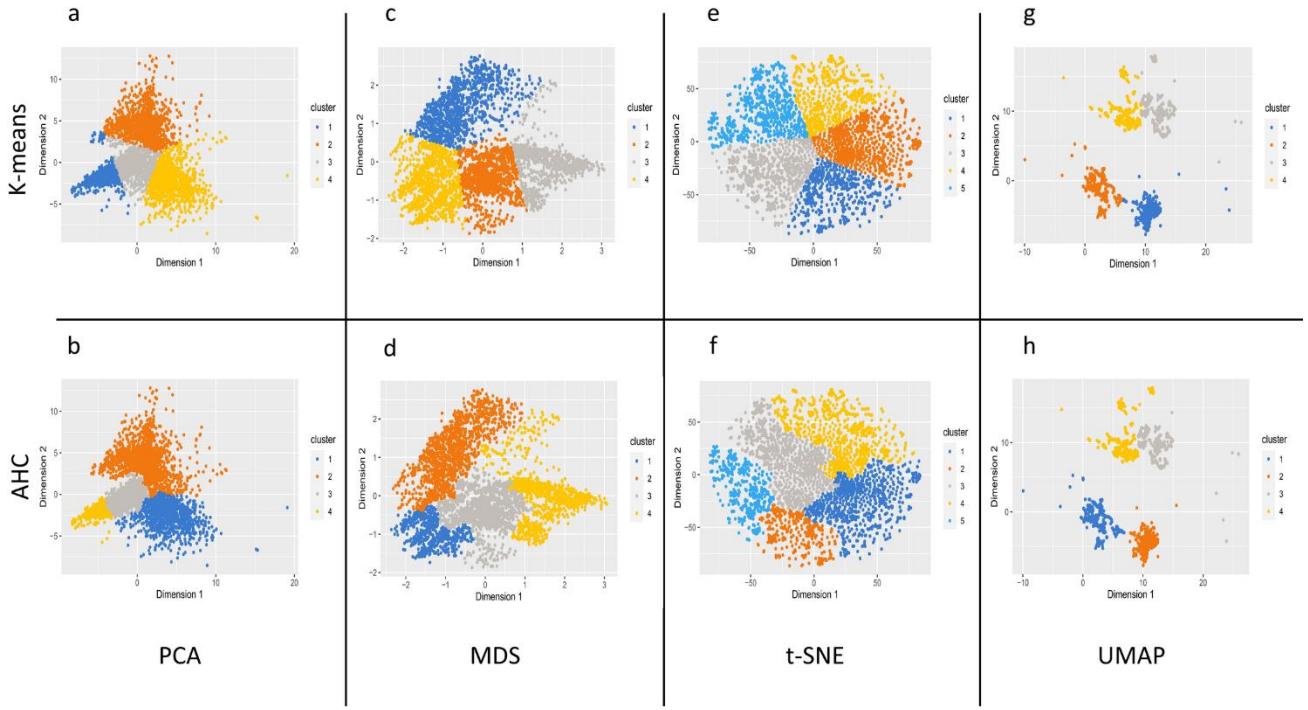
- **Modifiable Areal Unit Problem (MAUP)**
  - Changing the boundaries of spatial regions strongly influences the result!
  - Used for political gain in the US (gerrymandering)



http://www.e-education.psu.edu/geog486/l4 p7.html

Attribute Aggregation

- **Dimensionality Reduction**
- **Goal:** Preserve meaningful structure of the data
- Many algorithms available
  - Principal comp. analysis (**PCA**),
  - multi-dimensional scaling (**MDS**),
  - t-distributed stochastic neighborhood embedding (**t-SNE**)
  - Uniform manifold approximation and projection (**UMAP**)



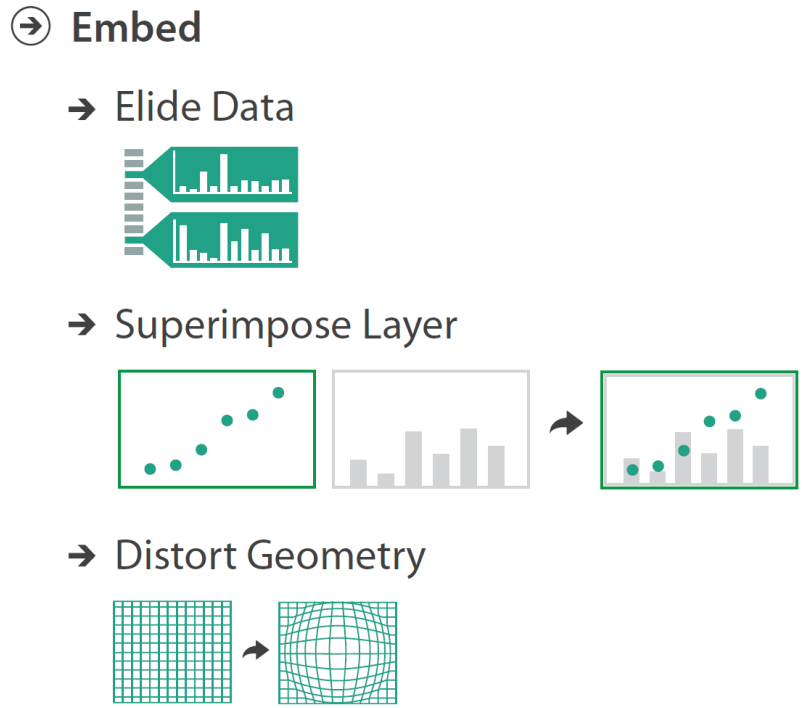
Rugard et al., Smell compounds classification using UMAP to increase knowledge of odors and molecular structures linkages, 2021

Focus + Context

- Embed detailed information of a selected set into an overview that contains the context

Design Choices

- **Elide**
  - Some items removed, others summarized
- **Superimpose**
  - Local focus region
- **Distort**
  - Context is compressed to create space for focus



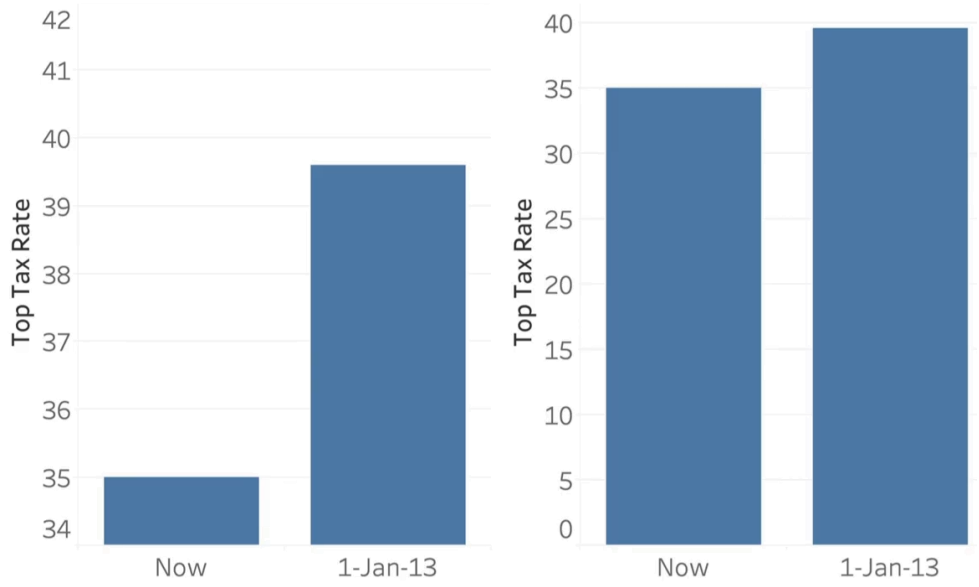
Visualization Design
Abstract Data
Spatial Data

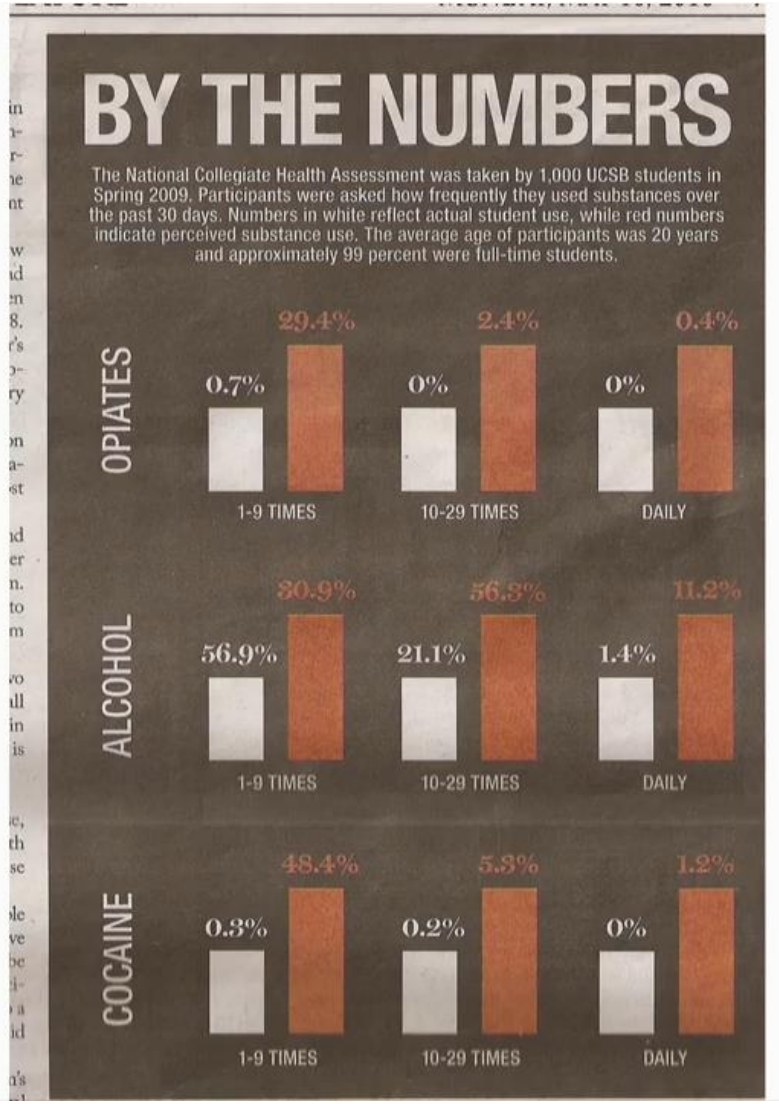
Interaction Concepts

View Manipulation
Data Reduction
Rules of Thumb

Problems with Bar Charts

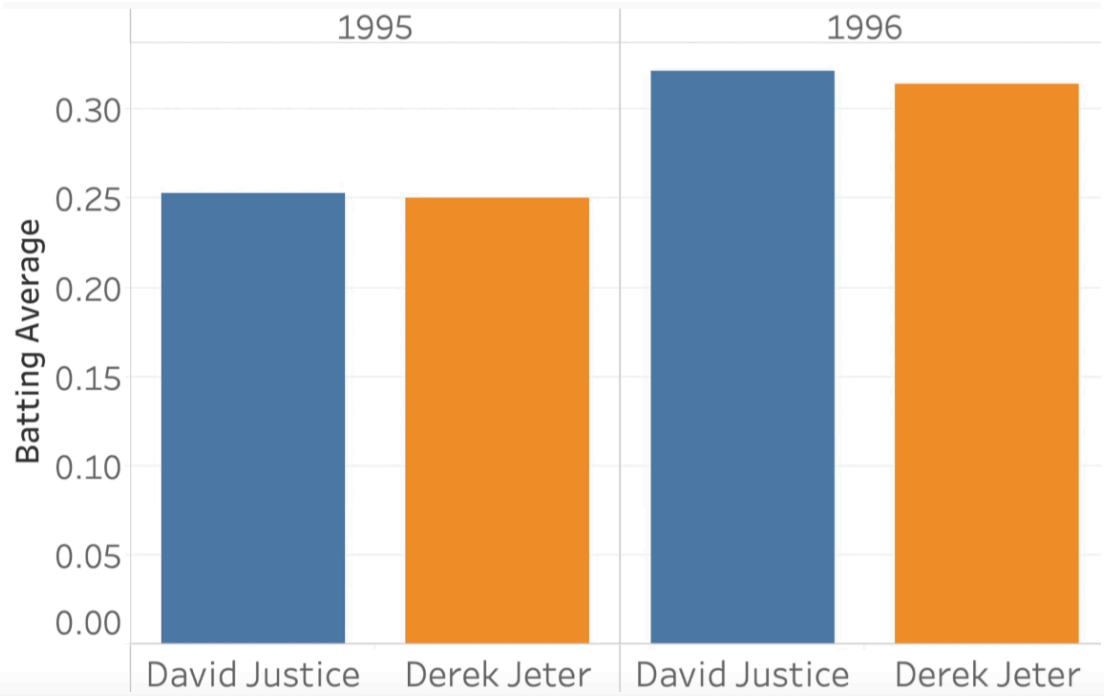
- Non-sequitors
- Y-Axis truncation

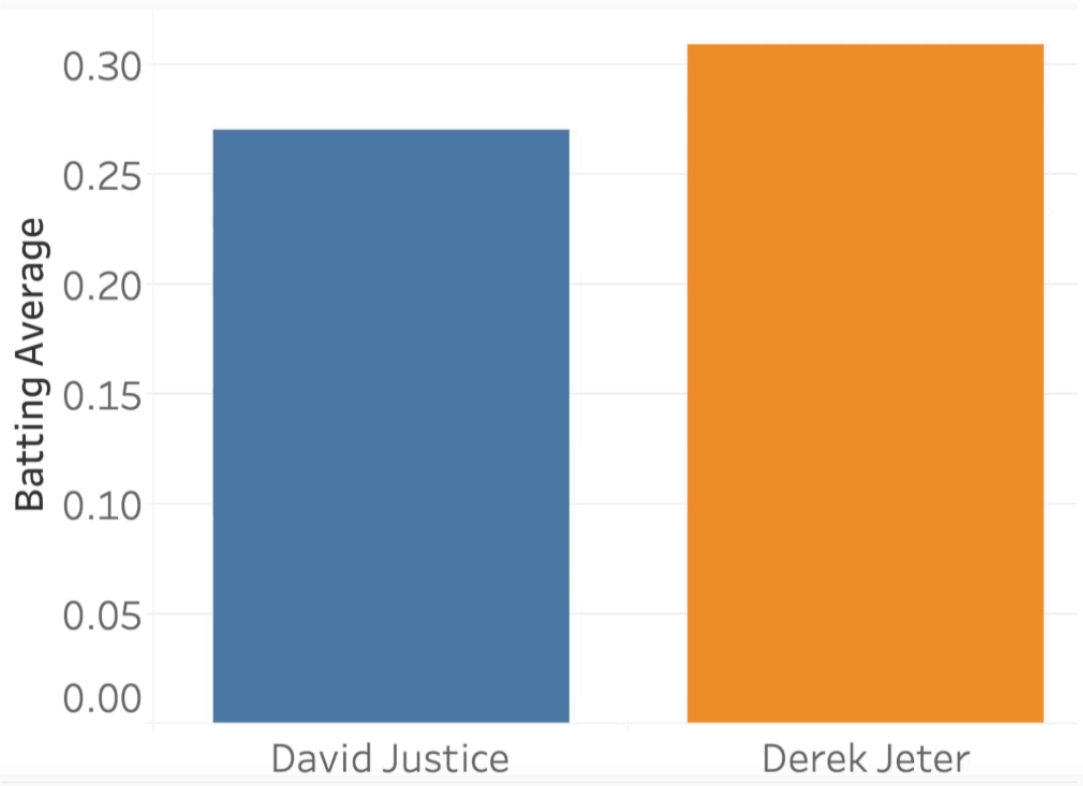




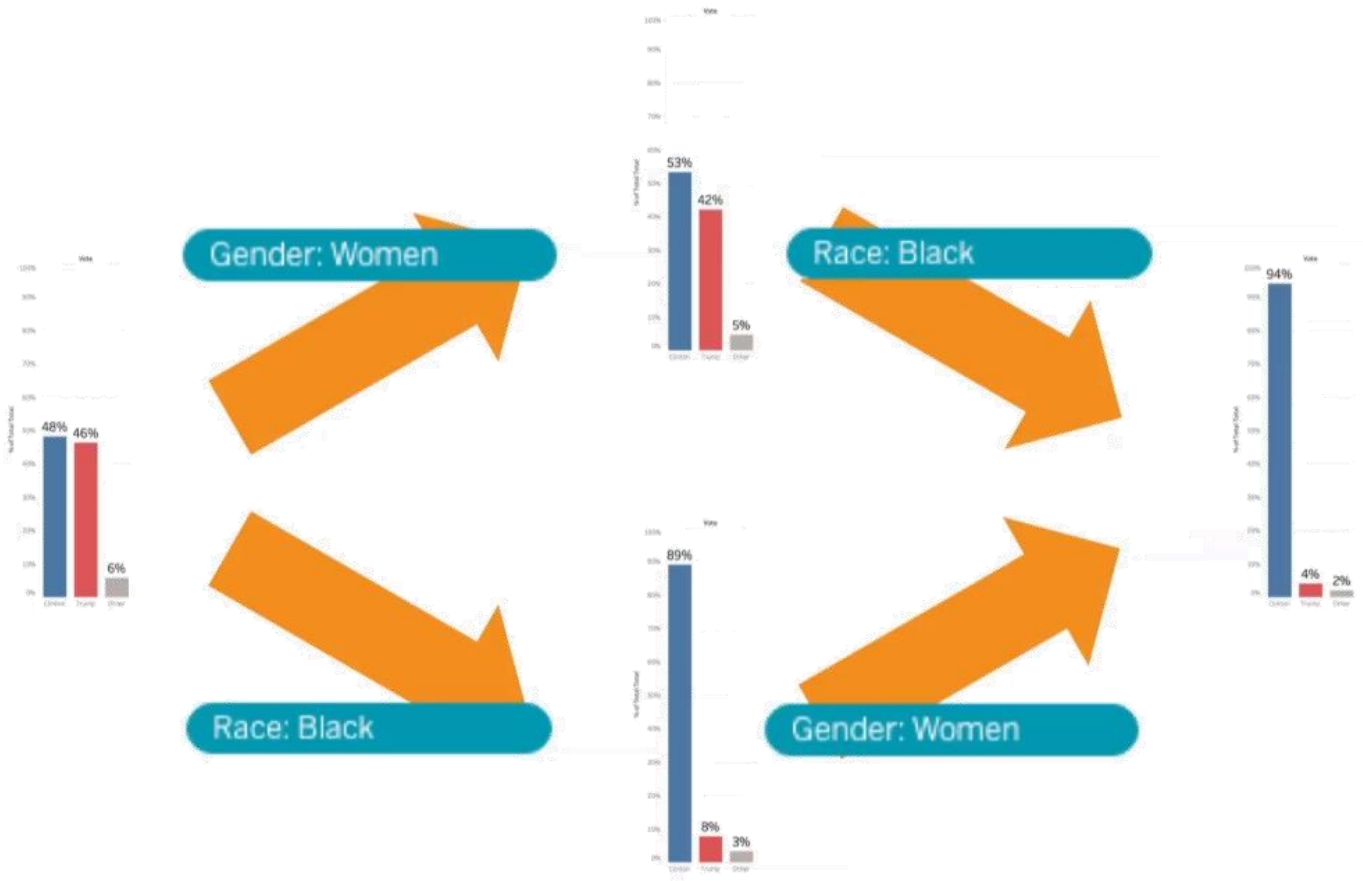
Simpson‘s Paradox

The level of aggregation determines what you see in the charts!



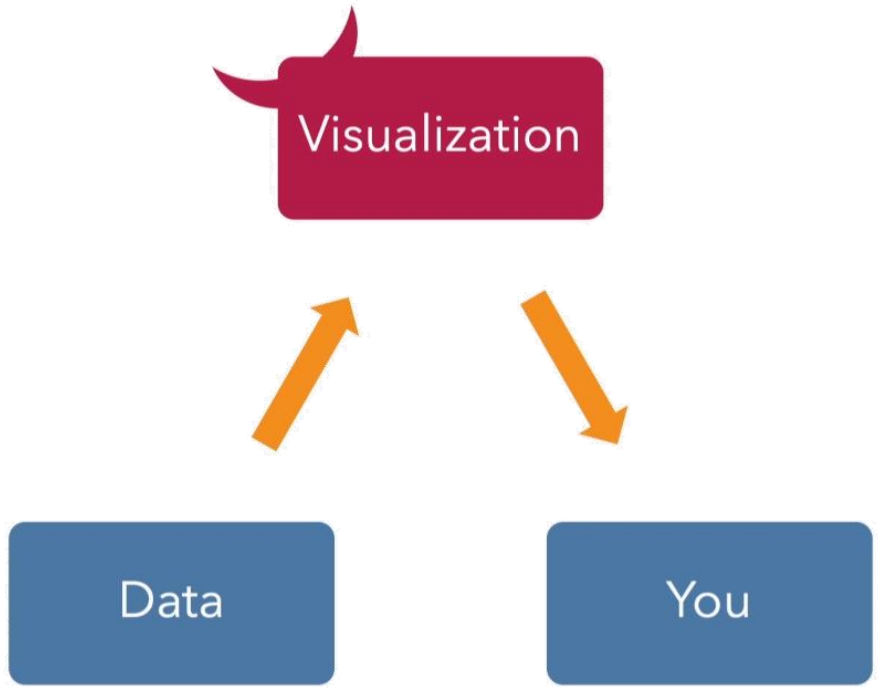


Drill Down Bias



Summary

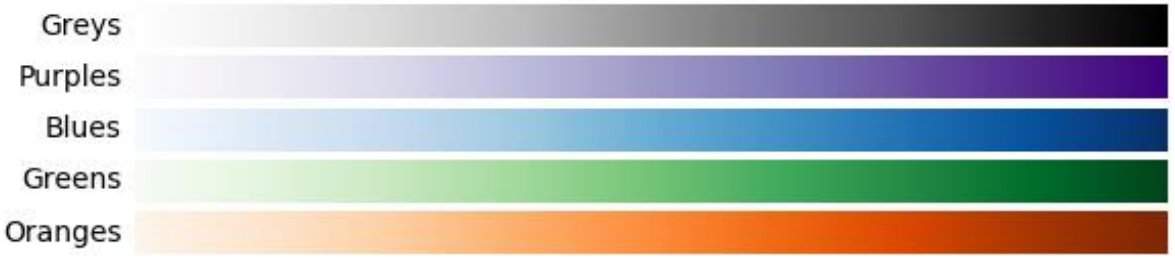
- Manipulations in visualization do not happen always with bad intention
- They just happen because the designer was not careful
- Mind the potential issues that can occur due to visualizations!
- Explore beyond the default settings



We never just show the data!

Rules of Thumb

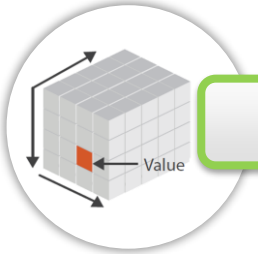
- No unjustified 3D
- No unjustified 2D
- Eyes Beat Memory
- Resolution over Immersion
- Overview First, Zoom and Filter, Detail on Demand
- Responsiveness is Required
- Get it Right in Black and White
- Function First, form Next



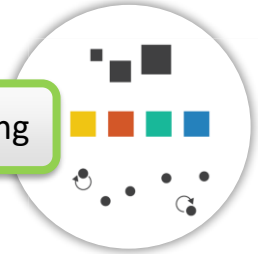
Colormaps from Matplotlib,
https://matplotlib.org/3.3.3/tutorials/colors/colormaps.html

Course Overview

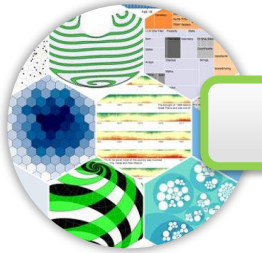
Visualization Design



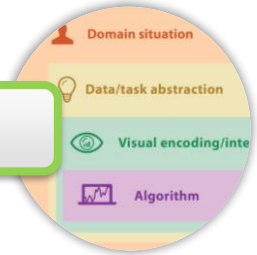
Data Abstraction



Perception and Mapping



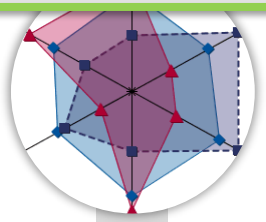
Web-based Systems



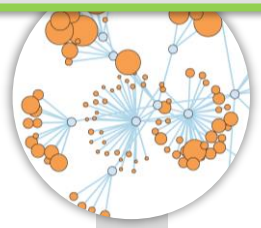
Tasks and Validation

Abstract Data

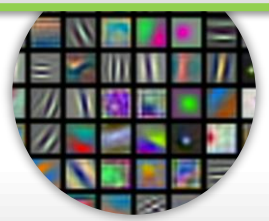
Tabular Data



Networks and Trees



Deep Learning

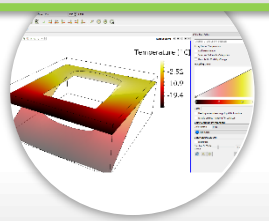


Spatial Data

Scalar and Vector Fields

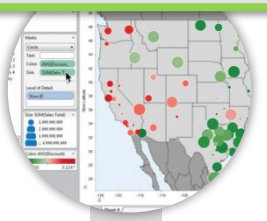
Medical Data

SciVis Tools

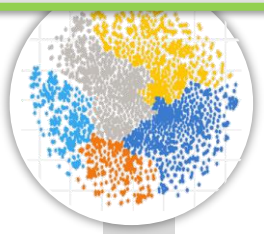


Interaction Concepts

View Manipulation

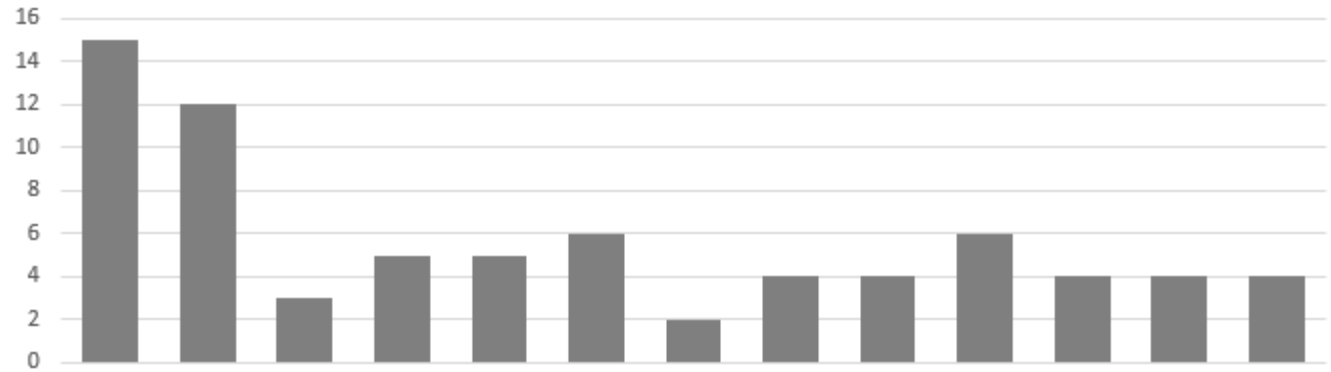


Data Reduction

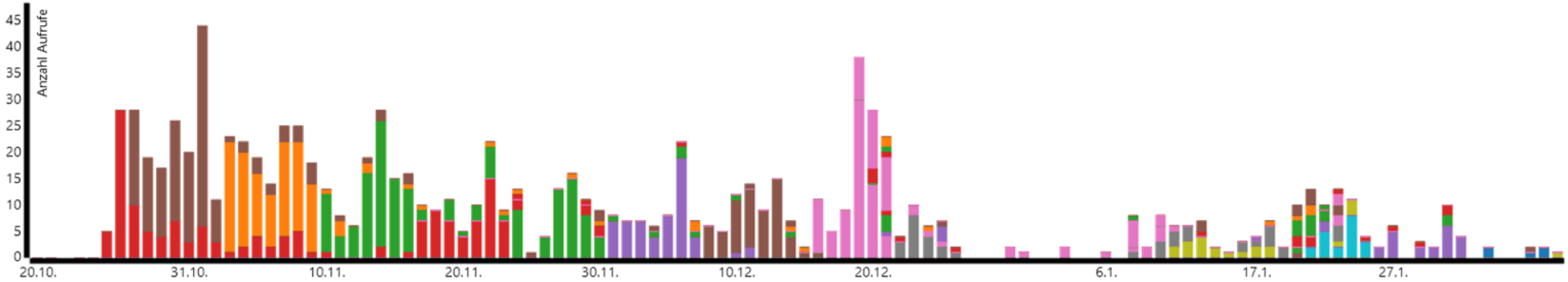


Rules of Thumb

Statistics WS 2023/24



Live stream



Recordings

Summer Semester

Scientific Visualization

- Lecture: 4 SWS, 5.0 ECTS

Information Visualization

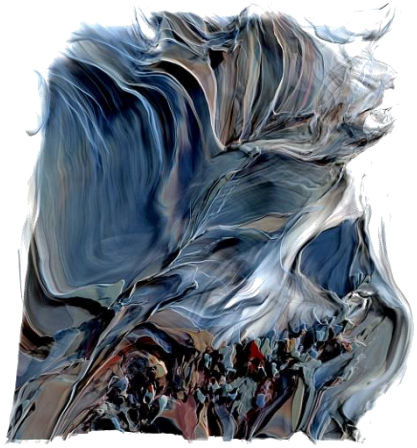
- Lecture: 4 SWS, 5.0 ECTS

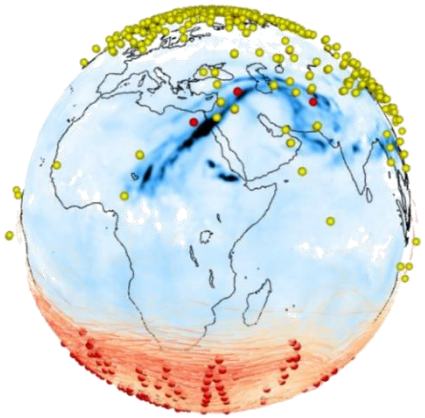
Physically-based Simulation in Computer Graphics

- Lecture: 4 SWS, 5.0 ECTS

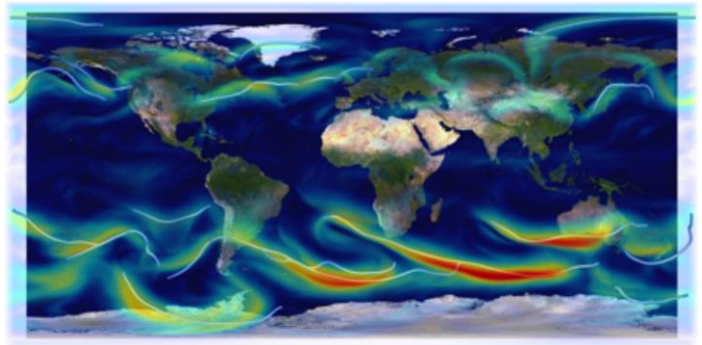
Thesis Topics

- Bachelor/Master Thesis





Images by Günther et al.



Jets stream extraction by Nathan Lerzer

Thuerey et al. 2010: A multiscale approach to mesh-based surface tension flows

Thank you