

Visualization

– Information Visualization I

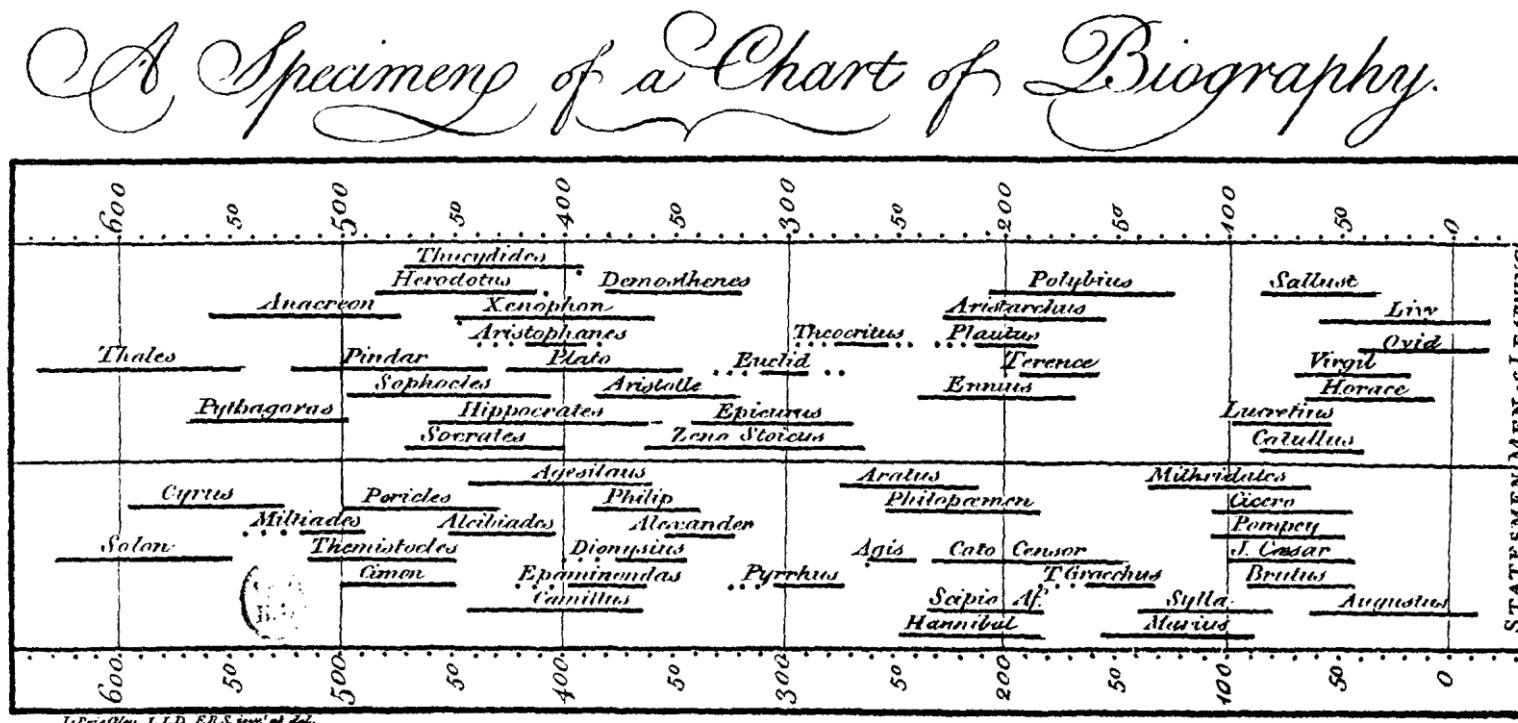
J.-Prof. Dr. habil. Kai Lawonn

Outline

- Historical Examples
- Motivation
- Visualization Pipeline
- Visualization of Multivariate Data
- Visualization of Trees and Hierarchies
- Visualization of Time-Oriented Data

Historical Examples

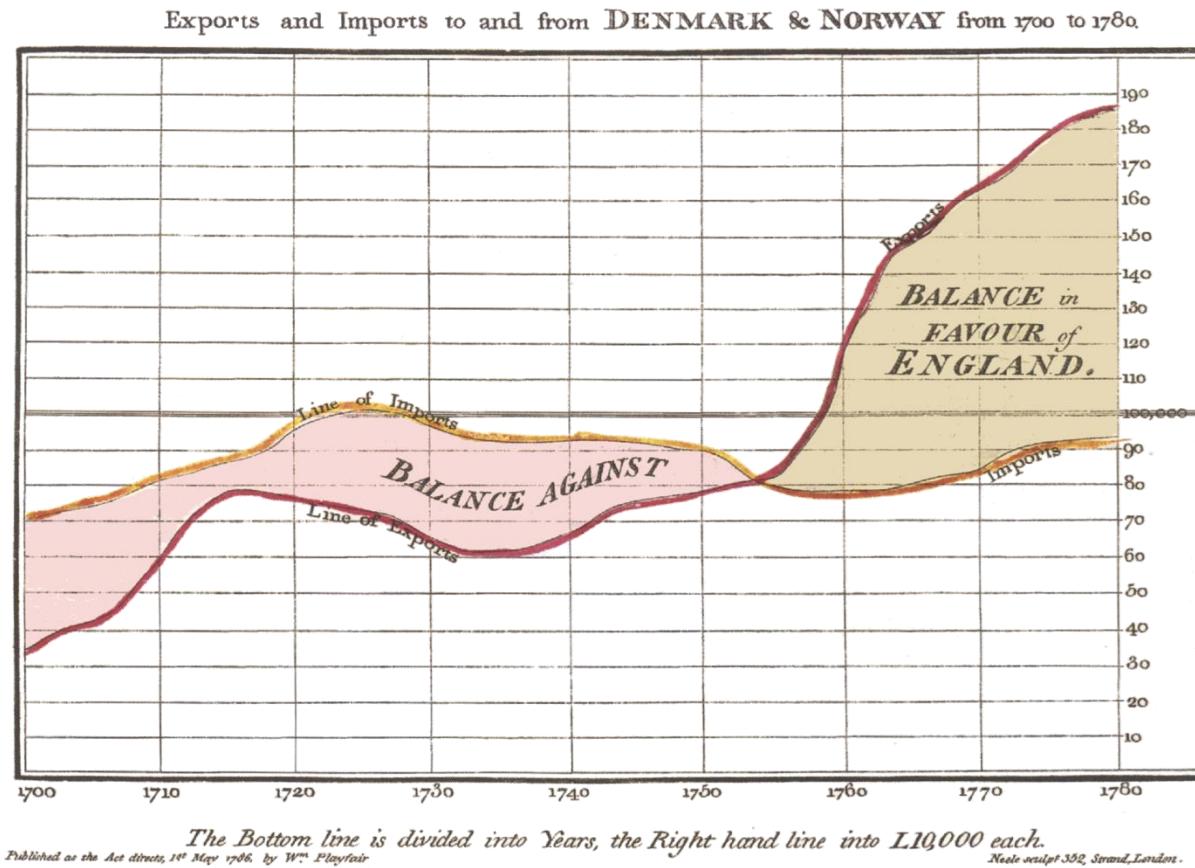
- Timeline chart by Joseph Priestley created in 1765
- Bars visualize the life spans of famous persons



Wikimedia Commons, http://en.wikipedia.org/wiki/William_Playfair

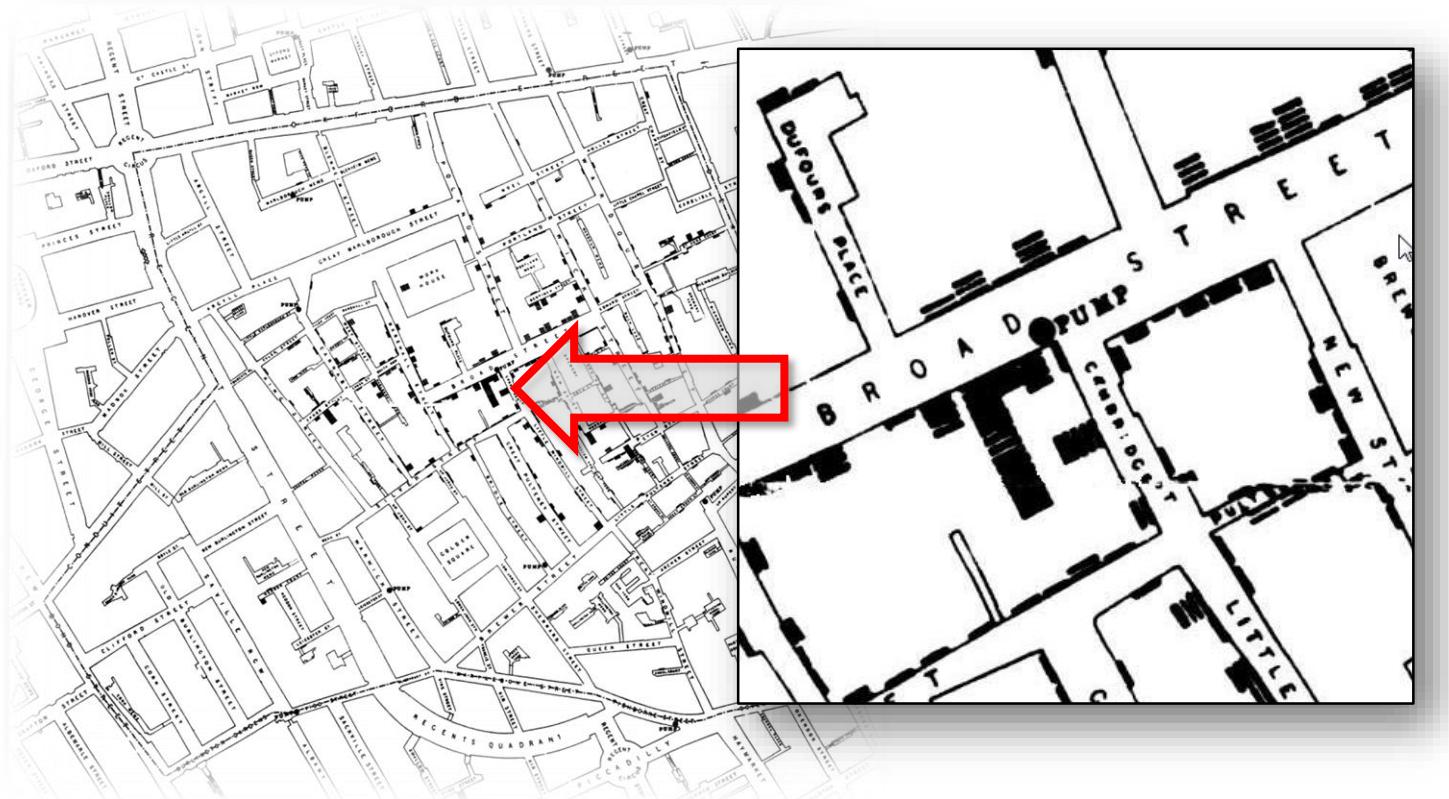
Historical Examples

- Trade-balance time-series chart by William Playfair created in 1786



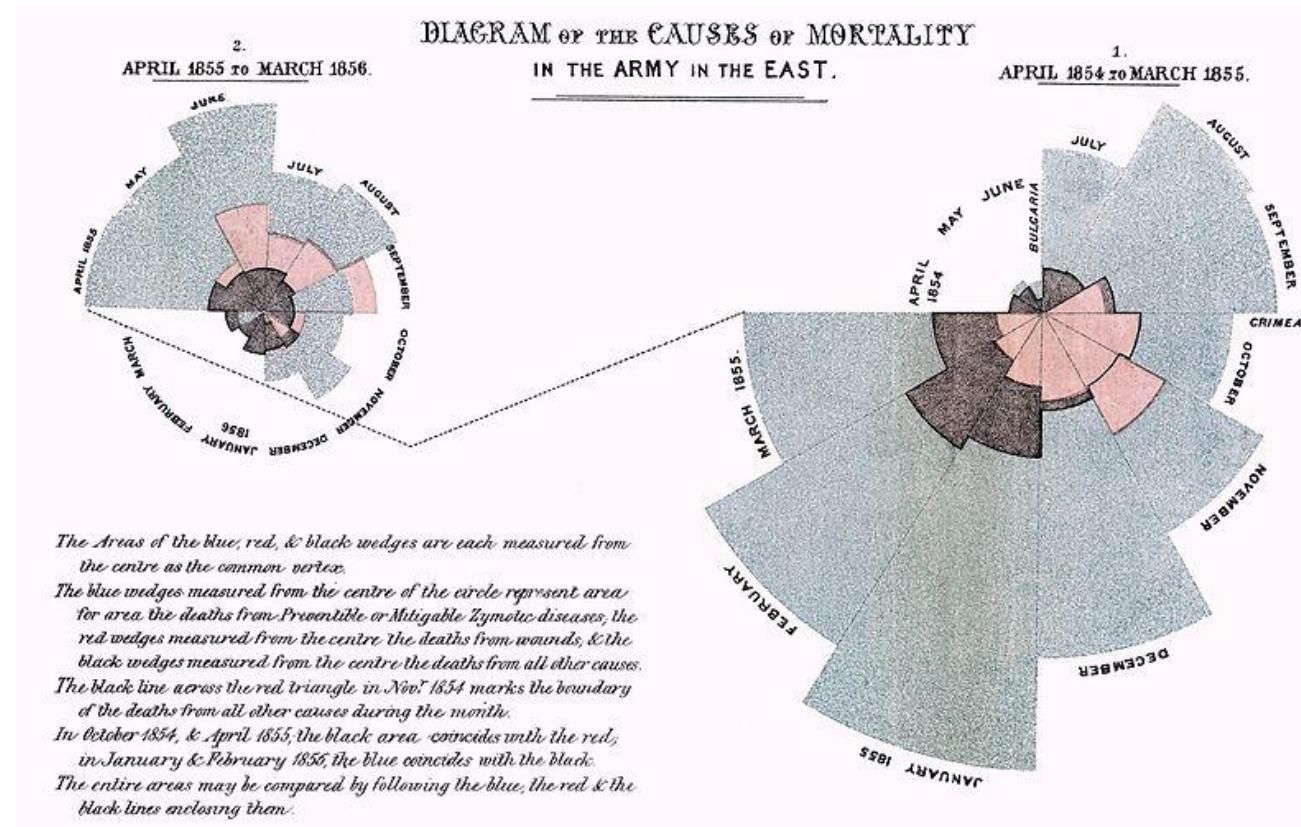
Historical Examples

- John Snow identified water pump as cause of cholera pandemic in London (1854) based on street map



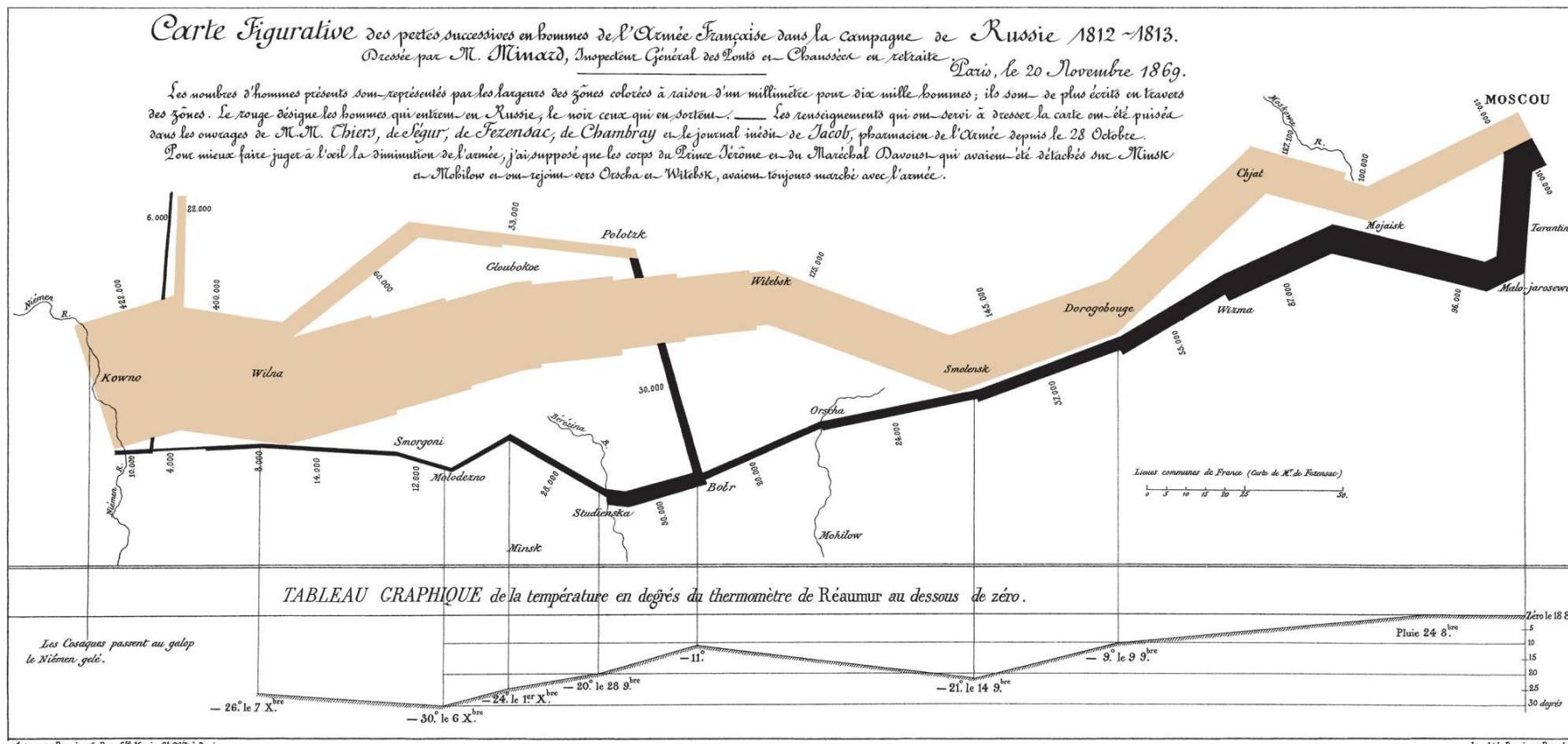
Historical Examples

- Rose plot by Florence Nightingale created in 1858
- Wedges represent different causes of death (color)



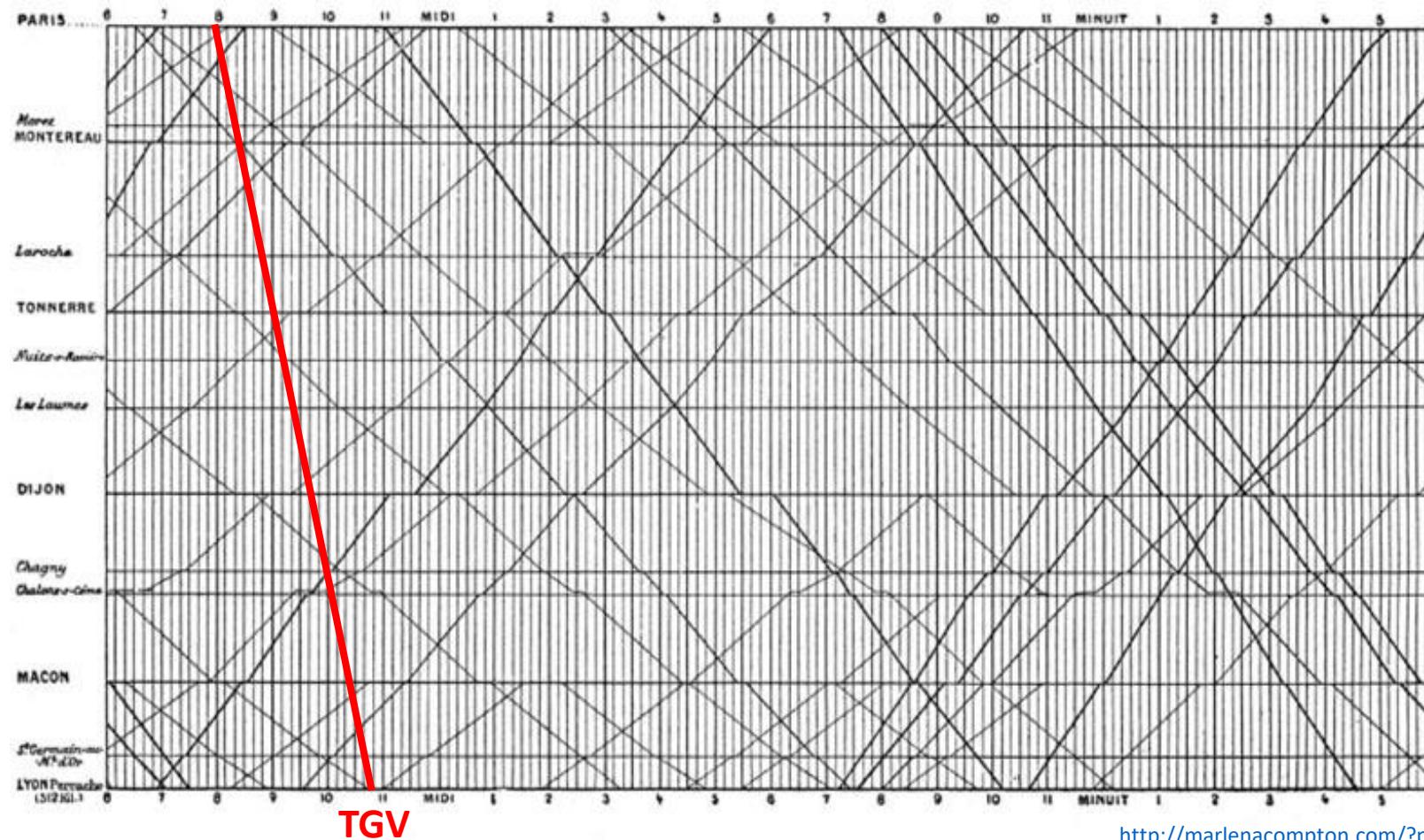
Historical Examples

- Map of Napoleon's march towards Moscow created by Charles Joseph Minard in 1869



Historical Examples

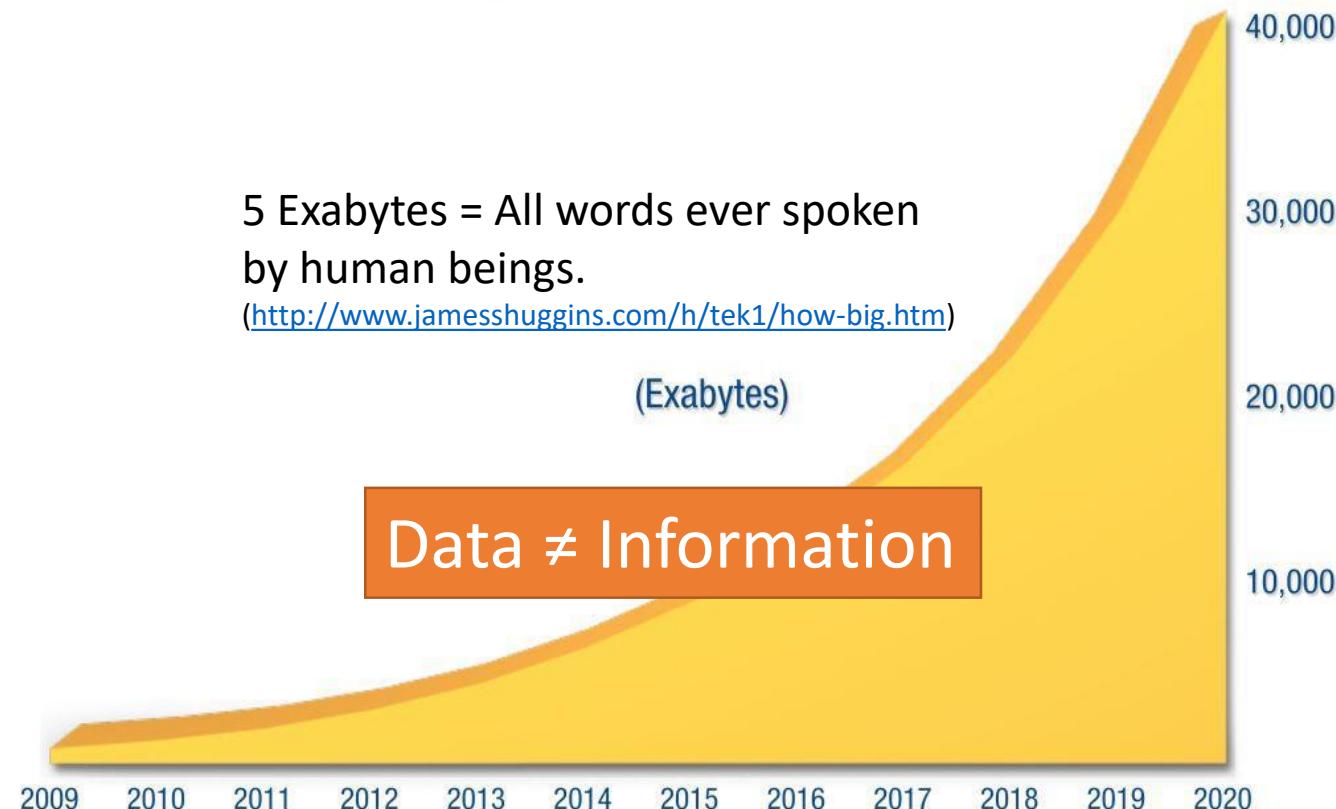
- Train schedule Paris-Lyon from 1880's created by Étienne-Jules Marey



Motivation

“Big Data”

The Digital Universe: 50-fold Growth from the Beginning of 2010 to the End of 2020



“Big Data” – Examples

- Experiments at CERN’s Large Hadron Collider generate 45TB of data per second
- Four-engine jumbo jet creates 640TB of data during transatlantic flight
- Monitoring trading activities at the New York Stock Exchange (NYSE) produces 4-5 TB of data per day
- Social network data
 - 320 million active users per month on Twitter
 - 1.55 billion monthly active users on Facebook

From Data to Information

- How to extract information from huge mass of data?
- Problem: Cognition is limited!
- Cognitive abilities:
 - Attention
 - Visual Perception
 - Memory
 - Calculating
 - Reasoning
 - Problem solving
 - Decision making ...

Cognition

- Limited attention
- Change blindness study by D. Simons and D. Levin



Cognition

Limited visual perception

NMCOWUZTZTVNMSDSJKKHGFRU
HVJBKNVKYMNMBQDQWOOTERU
MYPBCYVBGYJGZIOEWZUHFEGHJ
SKHFEHJWZQWGBSBMFBMHJEEU

Cognition

Limited visual perception

NMCOWUZTZTVNMSDSJKKHGFRU
HVJBKNVKYMNMBQDQWOOTERU
MYPBCYVBGYJGZIOEWZUHFEGHJ
SKHFEHJWZQWGBSBMFBMHJEEU

Cognition

- Limited memory demonstrated by calculation exercise

$$83 \times 24$$

$$\begin{array}{r} 83 \\ \times 24 \\ \hline \end{array}$$

$$166$$

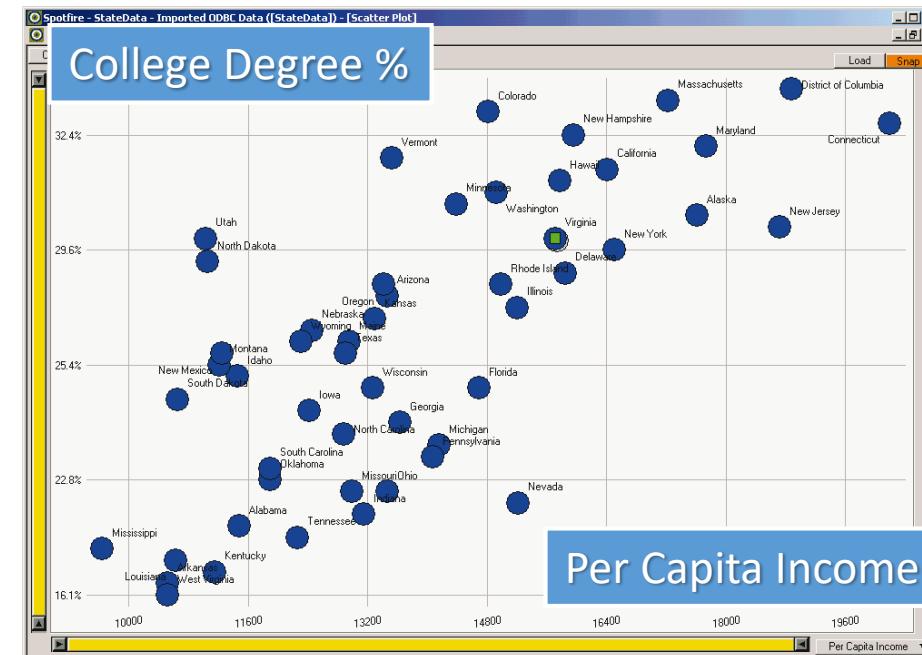
$$332$$

$$1992$$

Visualization Amplifies Cognition

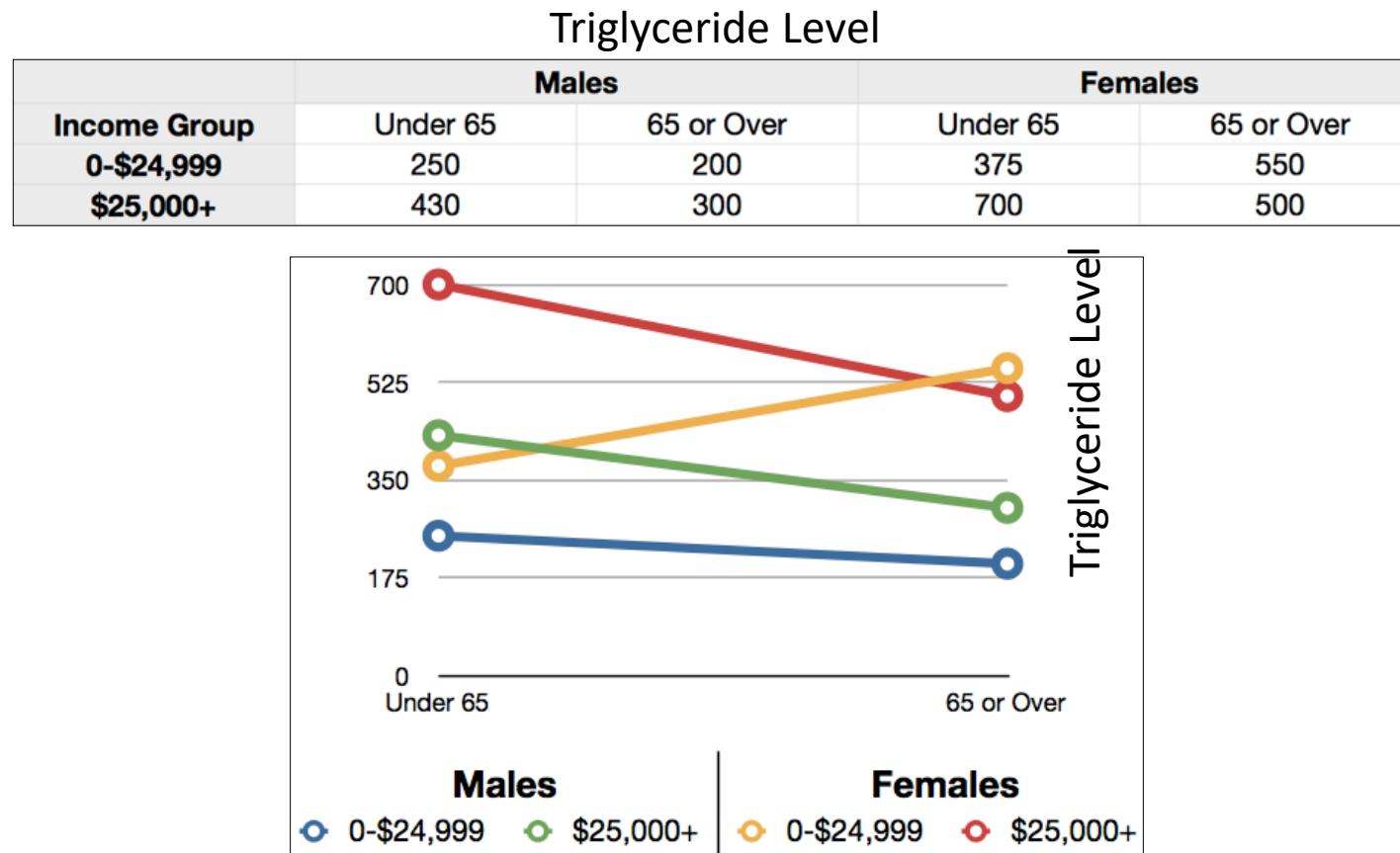
- Which federal state has the highest per capita income? What is the average? → Excel
- What is the relation between income and education? Outliers? → Visualization

| State | College Degree % | Per Capita Income |
|----------------------|------------------|-------------------|
| Alabama | 20.6% | 11486 |
| Alaska | 30.3% | 17610 |
| Arizona | 27.1% | 13461 |
| Arkansas | 17.0% | 10520 |
| California | 31.3% | 16409 |
| Colorado | 33.9% | 14821 |
| Connecticut | 33.8% | 20189 |
| Delaware | 27.9% | 15854 |
| District of Columbia | 36.4% | 18881 |
| Florida | 24.9% | 14698 |
| Georgia | 24.3% | 13631 |
| Hawaii | 31.2% | 15770 |
| Idaho | 25.2% | 11457 |
| Illinois | 26.8% | 15201 |
| Indiana | 20.9% | 13149 |
| Iowa | 24.5% | 12422 |
| Kansas | 26.5% | 13300 |
| Kentucky | 17.7% | 11153 |
| Louisiana | 19.4% | 10635 |
| Maine | 25.7% | 12957 |
| Maryland | 31.7% | 17730 |
| Massachusetts | 34.5% | 17224 |
| Michigan | 24.1% | 14154 |
| Minnesota | 30.4% | 14389 |



Visualization Amplifies Cognition

- Which gender and income level shows a different effect of age on triglyceride levels? (Example from [Meyer12])



Information Visualization

„Large complex datasets have some of the following properties, often all:

- a large number of records,
- many variables,
- complex data structures ...,
- Intricate patterns and dependencies ... that require complex methods of analysis.“ (Guha, 2009)

Information Visualization

- Deals with abstract information, e.g. resulting from data base queries
- Data is
 - network data (graphs),
 - tree data (special graphs),
 - high-dimensional data (more than 3D+time) represented in tables or
 - text
- No inherent grid structure
- Data inherently discrete (no interpolation applicable)
- Attributes are of any type (nominal, ordinal, scalar)

Visualization Amplifies Cognition

- Visualization amplifies cognition [Card1999] by
 - Reducing the search for information
 - Increasing memory and processing resources
 - Visual representations support pattern recognition
 - Enabling perceptual inference operations
 - Using perceptual attention mechanisms for monitoring
 - Encoding information in a manipulable medium
- Visual perception plays a crucial role
 - Visual sense has highest bandwidth
 - Fast and parallel (compared to sequential text processing)
 - Extremely good pattern recognition

Why do we create visualizations?

High-level [Meyer12]

- answer questions
- evaluate/generate hypotheses
- make decisions
- see data in context
- expand memory
- support computational analysis
- find patterns
- tell a story

Low-level [Dachselt11]

- detect pattern and trends
- trace changes over time
- relations, correlations
- find clusters
- detect outliers, exceptions
- missing, false data

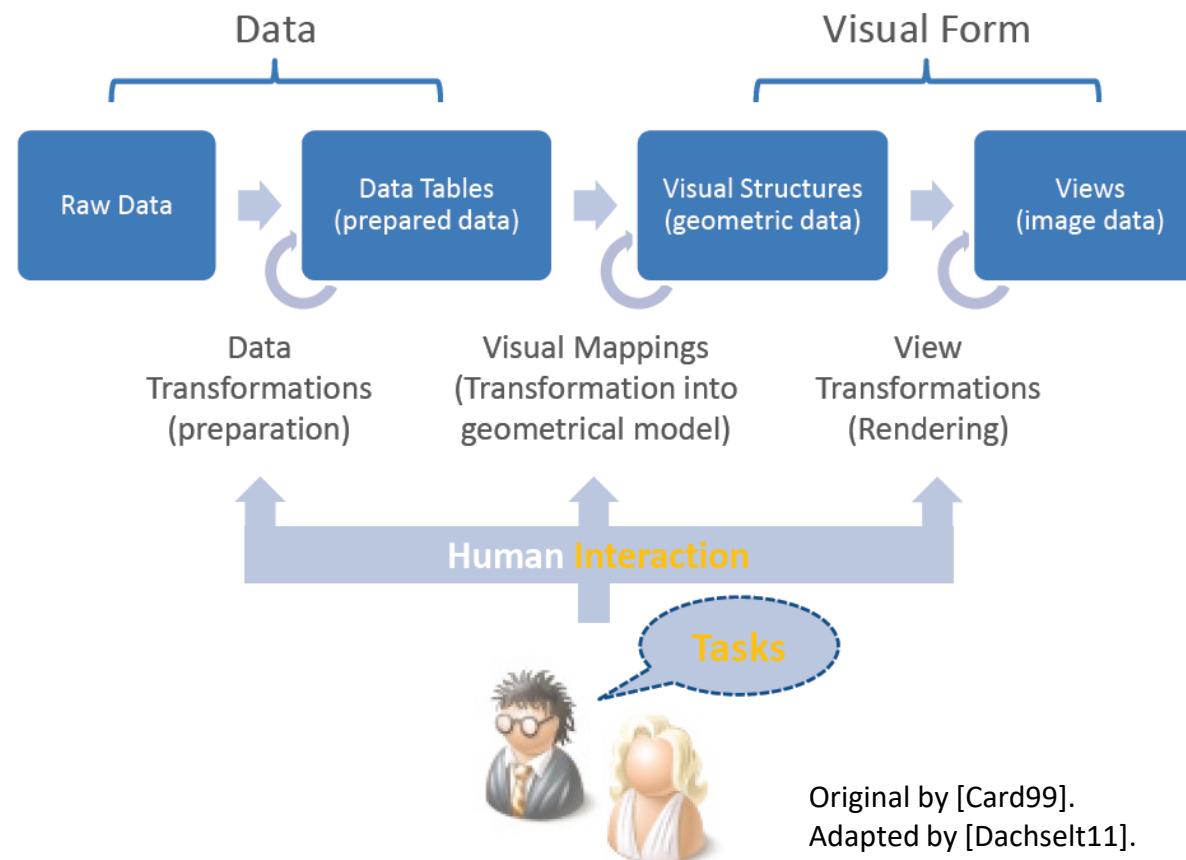
Ingredients of visualization: visual representation, human perception and cognition, and interaction.

Data Visualization Pipeline

From Data to Knowledge

Data Visualization Pipeline

- From data to knowledge along the data visualization pipeline aka *reference model of data visualization* [Card99]



Data Visualization Pipeline

Stages of the pipeline [Ward10,Mazza09]

- Data Transformations
 - Modeling: convert raw data to an organized logical structure
 - Selection: select subset of the data to be visualized
 - Pre-processing: normalization, interpolation, subsampling,...

Data Visualization Pipeline

Stages of the pipeline [Ward10,Mazza09]

- Data Transformations
 - Modeling: convert raw data to an organized logical structure
 - Selection: select subset of the data to be visualized
 - Pre-processing: normalization, interpolation, subsampling,...
- Visual Mappings
 - Mapping data to graphical entities
 - Choice of spatialization

Data Visualization Pipeline

Stages of the pipeline [Ward10,Mazza09]

- Data Transformations
 - Modeling: convert raw data to an organized logical structure
 - Selection: select subset of the data to be visualized
 - Pre-processing: normalization, interpolation, subsampling,...
- Visual Mappings
 - Mapping data to graphical entities
 - Choice of spatialization
- View Transformations
 - As in traditional computer graphics
 - Specify visualization attributes such as color map, lighting,...
 - Rendering – from graphical entities and mappings to image

Data Models

- Represent aspects of the real world
- Consist of basic elements
 - Information objects
 - Subjects of interest, e.g., humans, animals, cars, movies,...
 - Often belong to a specific problem domain
 - Different dimensionality: 1D/2D/3D/nD, temporal

Based on [Schumann2000]

Data Models

- Represent aspects of the real world
- Consist of basic elements
 - Information objects
 - Subjects of interest, e.g., humans, animals, cars, movies,...
 - Often belong to a specific problem domain
 - Different dimensionality: 1D/2D/3D/nD, temporal
 - Attributes
 - Are properties of an information object
 - Different data types: nominal, ordinal, quantitative
 - Number of attributes: uni-, bi-, tri-, multivariate
 - Suitability of visualization technique strongly depends on data type and number of attributes

Based on [Schumann2000]

Data Models

- Represent aspects of the real world
- Consist of basic elements
 - Information objects
 - Subjects of interest, e.g., humans, animals, cars, movies,...
 - Often belong to a specific problem domain
 - Different dimensionality: 1D/2D/3D/nD, temporal
 - Attributes
 - Are properties of an information object
 - Different data types: nominal, ordinal, quantitative
 - Number of attributes: uni-, bi-, tri-, multivariate
 - Suitability of visualization technique strongly depends on data type and number of attributes
 - Relations
 - Connections between ≥ 2 information objects or attributes
 - Might be hierarchical: leafs part of plant, hierarchy of enterprise,...

Based on [Schumann2000]

Data Types

Nominal (unordered set of names)

- Examples: Car manufacturers and countries
- Only test for (in)equality are possible ($=, \neq$)

Ordinal (ordered set of non-measureable data)

- Examples: days of the week and rating scales
- Tests for equality and “direction” ($<, >, =, \neq$)

Quantitative (measured or simulated data)

- Examples: physical measurements of height, weight, length
- Full set of arithmetic operations possible

Transformations:

- Nominal to ordinal data by introducing order, e.g., alphabetical
- Quantitative to ordinal data by binning (*classification*)

Data Type by Task Taxonomy [Shneiderman96]

InfoVis data types:

| | | |
|---|-------------------|--|
| 1 | 1-dimensional | alphabetic lists, source code, text/documents |
| 2 | 2-dimensional | planar or map data, photos |
| 3 | 3-dimensional | molecules, human body, buildings |
| 4 | temporal | {start, finish}, e.g., medical records, project management, historical presentations |
| 5 | multi-dimensional | n attribute => points in n-dimensional space, e.g., relational DB |
| 6 | tree | Hierarchies or tree structured, e.g., file directories, business organizations |
| 7 | network | connected as graph(s), e.g., telecommunications network, www |

Data Type by Task Taxonomy [Shneiderman96]

Interaction Tasks: (*lecture “Design- and Interaction Principles”*)

- *Overview*: gain overview of entire data space.
- *Zoom*: zoom in information objects of interest.
- *Filter*: filter out uninteresting information objects.
- *Details-on-demand*: select object and get details.
- *Relate*: view relationships between objects.
- *History*: keep history to allow undo/redo.
- *Extract*: extract objects for further processing.

Shneiderman’s Mantra: Overview first, zoom and filter, then details on demand.

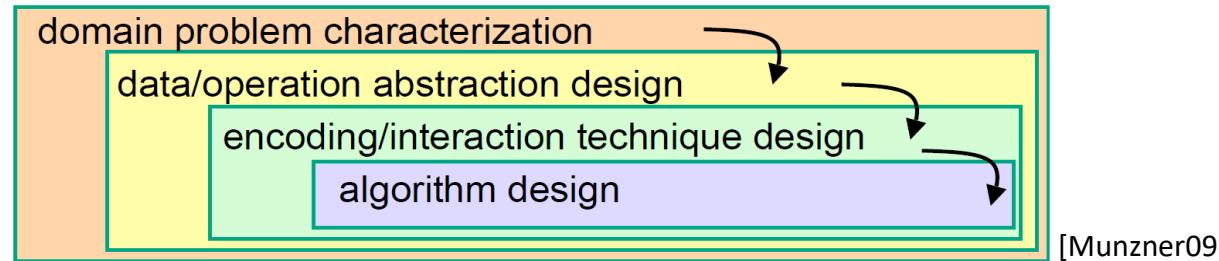
Generation of an Information Visualization

- Consider the following aspects:
 - *Target group*: Who is the user?
 - *User Tasks*: What does the user wish to accomplish?
 - *Data Source*: Where does the data come from?
 - *Representation*: How to represent the data? Which InfoVis techniques are appropriate?
 - *Output Medium*: Computer screen, paper, handheld,...?
- Be careful:
 - Always keep user tasks in mind!!!
 - Visualization emphasizes and hides certain information
 - Visualization may be misleading ([How NOT to Lie with Visualization](#))

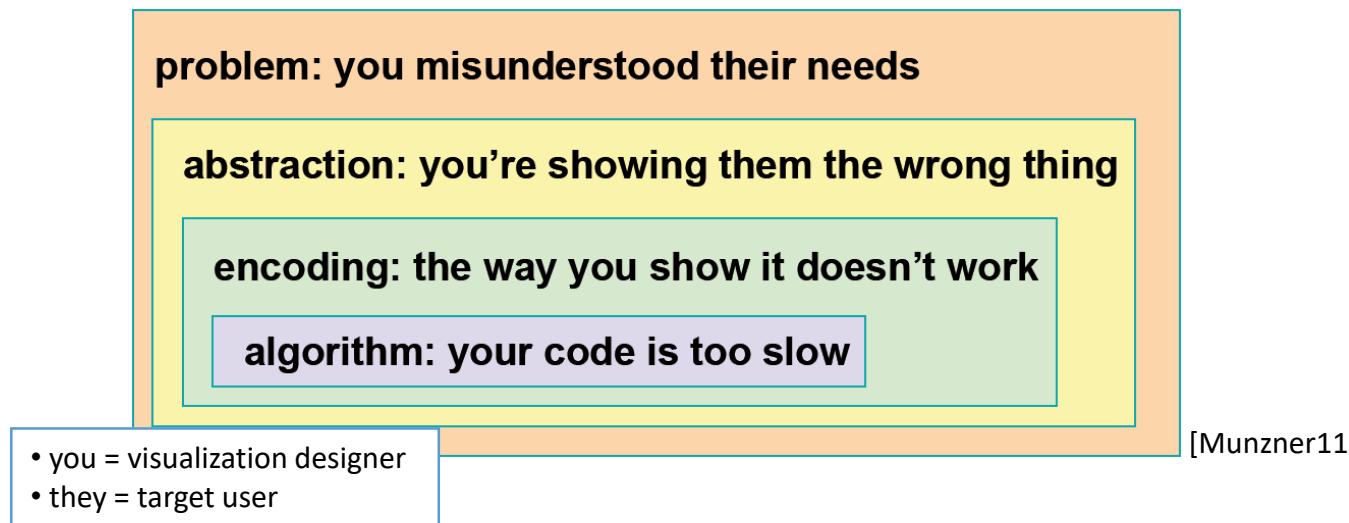
Generation of an Information Visualization

Nested Model for Vis. Design and Validation [Munzner09]

- Separating design process into four layers



- Validation against the right threat based on layer

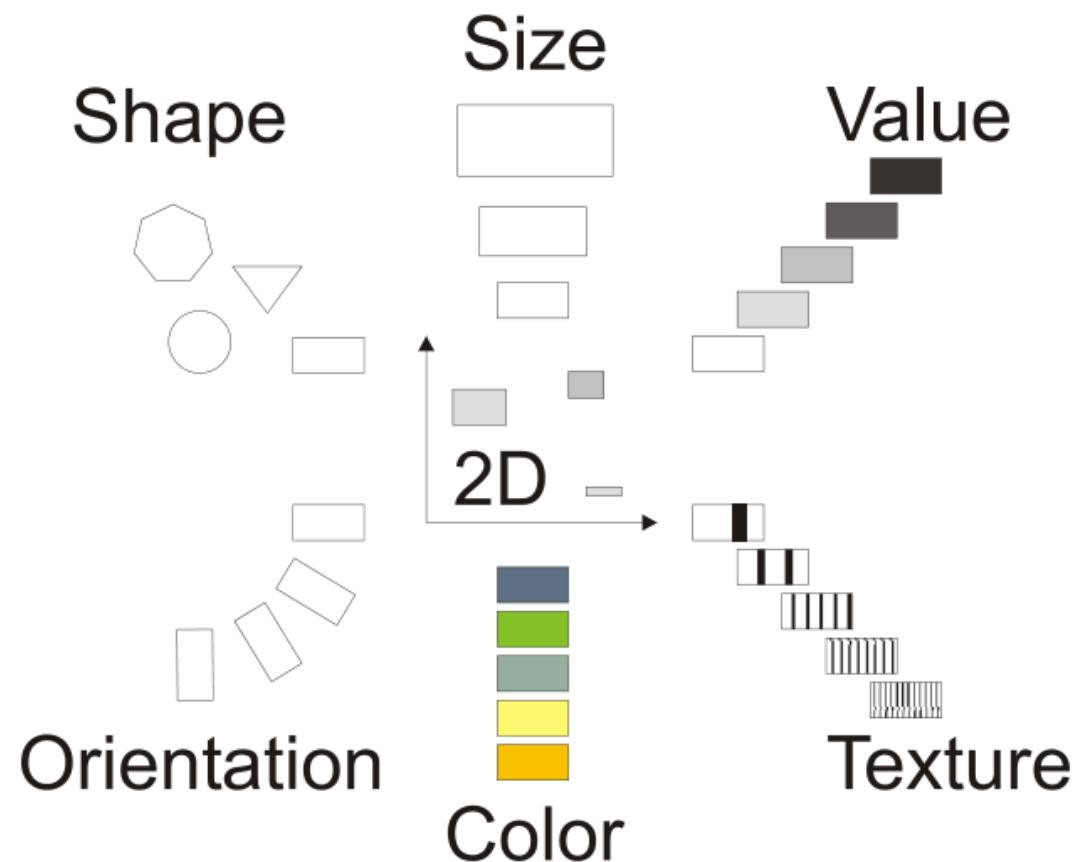


Visualization of Multivariate Data

Data Classification

- Multi-parameter data includes:
multivariate data and *multidimensional data*
- MULTIVARIATE DATA:
more than one dependent variable
(multidimensional feature space)
 - example: temperature, pressure, wind speed
- MULTIDIMENSIONAL DATA:
more than one independent variable (multidimensional observation space [1d, 2d, 3d, nD])
 - example: time varying medical volume data (4d)

Visual Variables



The Visualization Problem

- Hundreds or thousands of attributes but only a limited number of visual variables!

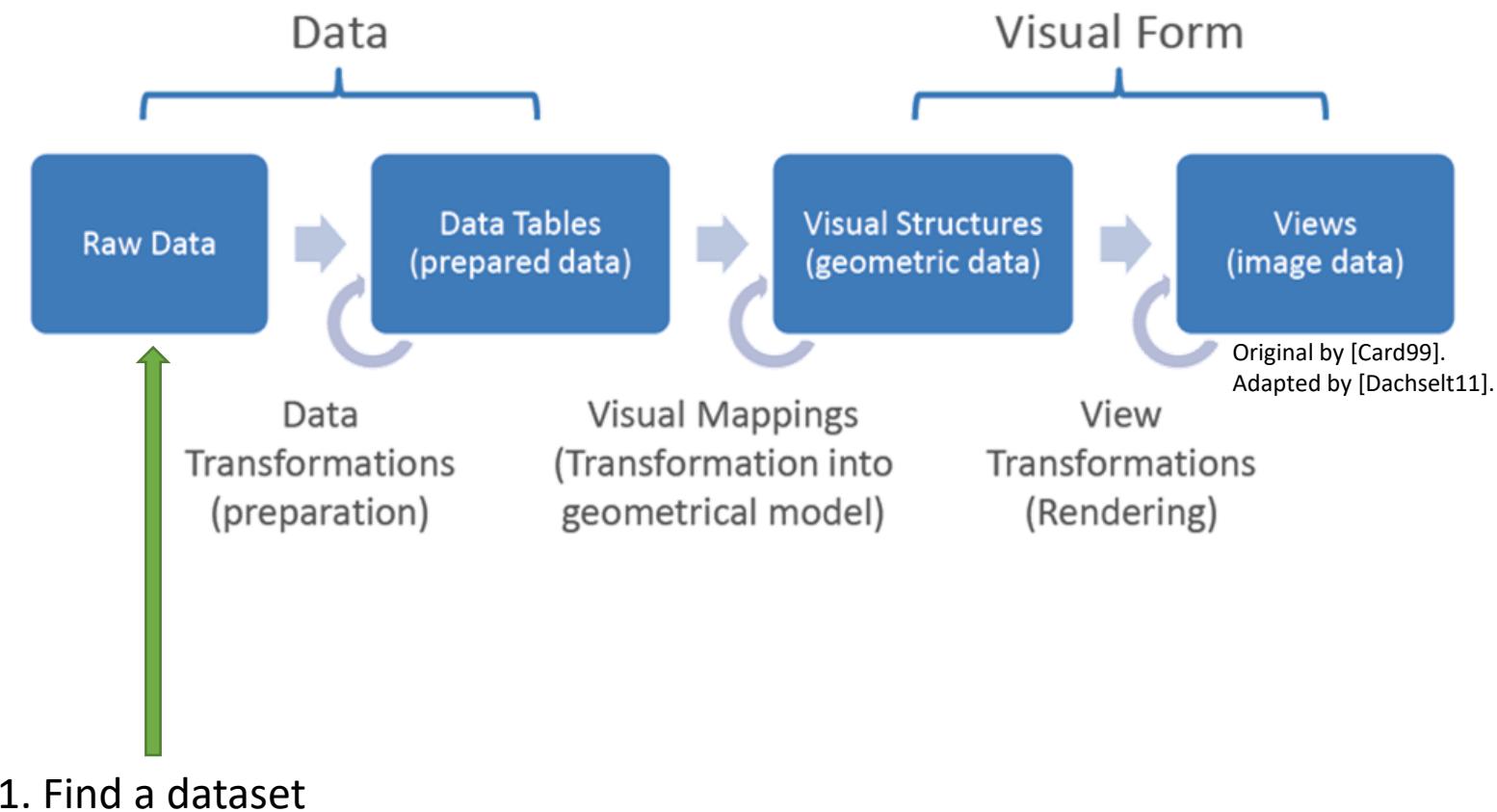
The diagram shows a screenshot of a spreadsheet application. The columns are labeled from A to O, and the rows are numbered 1 to 47. Column A is titled "Country Name" and contains a list of country names. A vertical blue arrow on the left points downwards through the country names, labeled "Countries". A horizontal blue arrow at the top points to the right across the columns, labeled "Attributes". The columns include "Country Name", "Country Code", "Education", "Environment", "Economic Policy & Debt", "Financial Sector", "Access", "Assets", "Capital Markets", "Health", "Infrastructure", "Social Protection & Labor", "Poverty", "Private Sector & Trade", and "Public Sector". The data extends from row 1 to row 47.

| A | B | C | D | E | F | G | H | I | J | K | L | M | N | O |
|-----------------------------|--------------|-----------|-------------|------------------------|------------------|--------|--------|-----------------|--------|----------------|---------------------------|---------|------------------------|---------------|
| 1 Country Name | Country Code | Education | Environment | Economic Policy & Debt | Financial Sector | Access | Assets | Capital Markets | Health | Infrastructure | Social Protection & Labor | Poverty | Private Sector & Trade | Public Sector |
| 2 Afghanistan | | | | | | | | | | | | | | |
| 3 Albania | | | | | | | | | | | | | | |
| 4 Algeria | | | | | | | | | | | | | | |
| 5 American Samoa | | | | | | | | | | | | | | |
| 6 Andorra | | | | | | | | | | | | | | |
| 7 Angola | | | | | | | | | | | | | | |
| 8 Antigua and Barbuda | | | | | | | | | | | | | | |
| 9 Argentina | | | | | | | | | | | | | | |
| 10 Armenia | | | | | | | | | | | | | | |
| 11 Aruba | | | | | | | | | | | | | | |
| 12 Australia | | | | | | | | | | | | | | |
| 13 Austria | | | | | | | | | | | | | | |
| 14 Azerbaijan | | | | | | | | | | | | | | |
| 15 Bahamas, The | | | | | | | | | | | | | | |
| 16 Bahrain | | | | | | | | | | | | | | |
| 17 Bangladesh | | | | | | | | | | | | | | |
| 18 Barbados | | | | | | | | | | | | | | |
| 19 Belarus | | | | | | | | | | | | | | |
| 20 Belgium | | | | | | | | | | | | | | |
| 21 Belize | | | | | | | | | | | | | | |
| 22 Benin | | | | | | | | | | | | | | |
| 23 Bermuda | | | | | | | | | | | | | | |
| 24 Bhutan | | | | | | | | | | | | | | |
| 25 Bolivia | | | | | | | | | | | | | | |
| 26 Bosnia and Herzegovina | | | | | | | | | | | | | | |
| 27 Botswana | | | | | | | | | | | | | | |
| 28 Brazil | | | | | | | | | | | | | | |
| 29 Brunei Darussalam | | | | | | | | | | | | | | |
| 30 Bulgaria | | | | | | | | | | | | | | |
| 31 Burkina Faso | | | | | | | | | | | | | | |
| 32 Burundi | | | | | | | | | | | | | | |
| 33 Cambodia | | | | | | | | | | | | | | |
| 34 Cameroon | | | | | | | | | | | | | | |
| 35 Canada | | | | | | | | | | | | | | |
| 36 Cabo Verde | | | | | | | | | | | | | | |
| 37 Cayman Islands | | | | | | | | | | | | | | |
| 38 Central African Republic | | | | | | | | | | | | | | |
| 39 Chad | | | | | | | | | | | | | | |
| 40 Channel Islands | | | | | | | | | | | | | | |
| 41 Chile | | | | | | | | | | | | | | |
| 42 China | | | | | | | | | | | | | | |
| 43 Colombia | | | | | | | | | | | | | | |
| 44 Comoros | | | | | | | | | | | | | | |
| 45 Congo, Dem. Rep. | | | | | | | | | | | | | | |
| 46 Congo, Rep. | | | | | | | | | | | | | | |
| 47 Costa Rica | | | | | | | | | | | | | | |

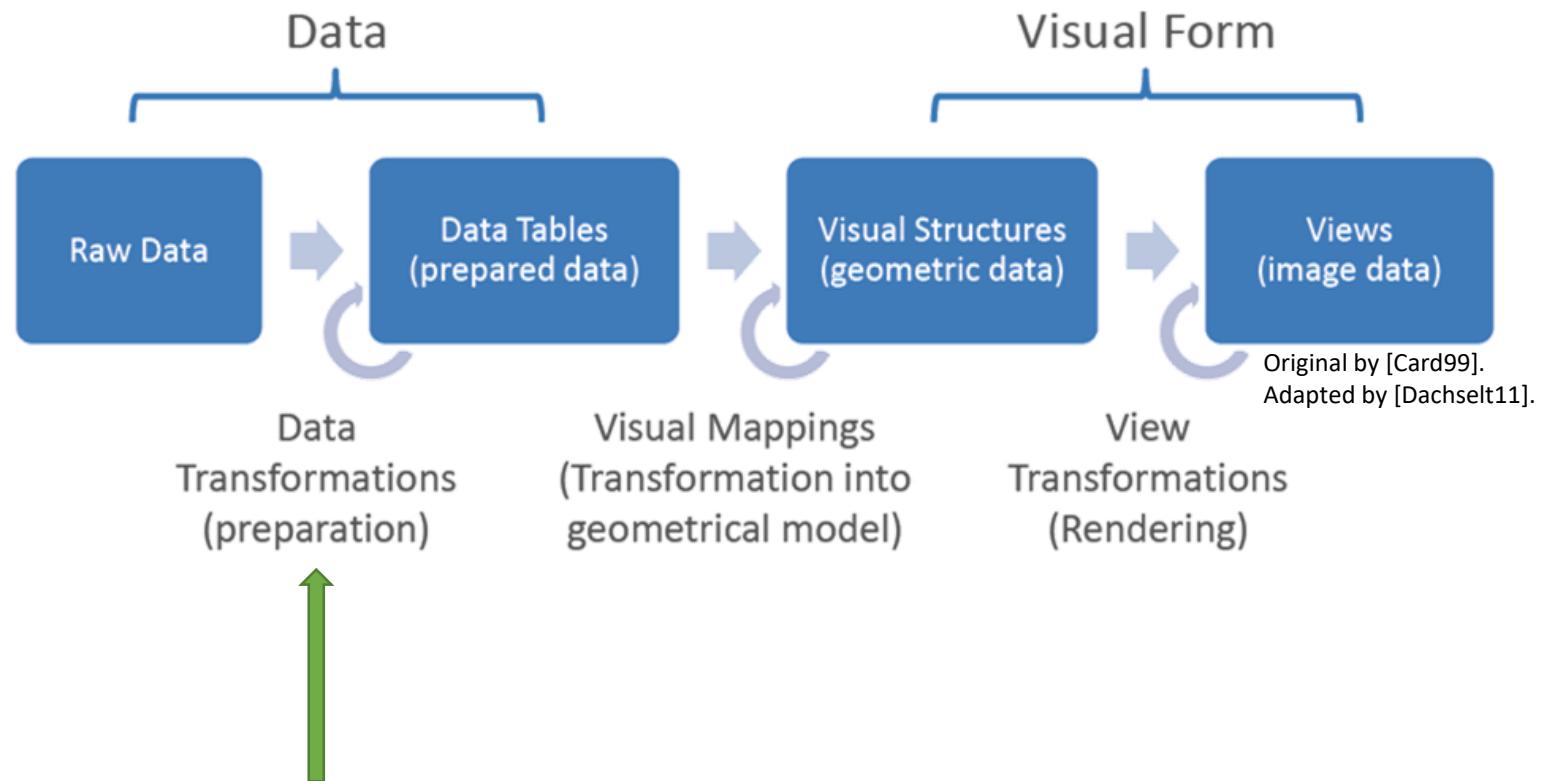
Two Approaches

1. Attempt to visualize all attributes
2. **Visualize only parts of the data**
→ Add interaction facilities

How Do We Visualize Multivariate Data?



How Do We Visualize Multivariate Data?

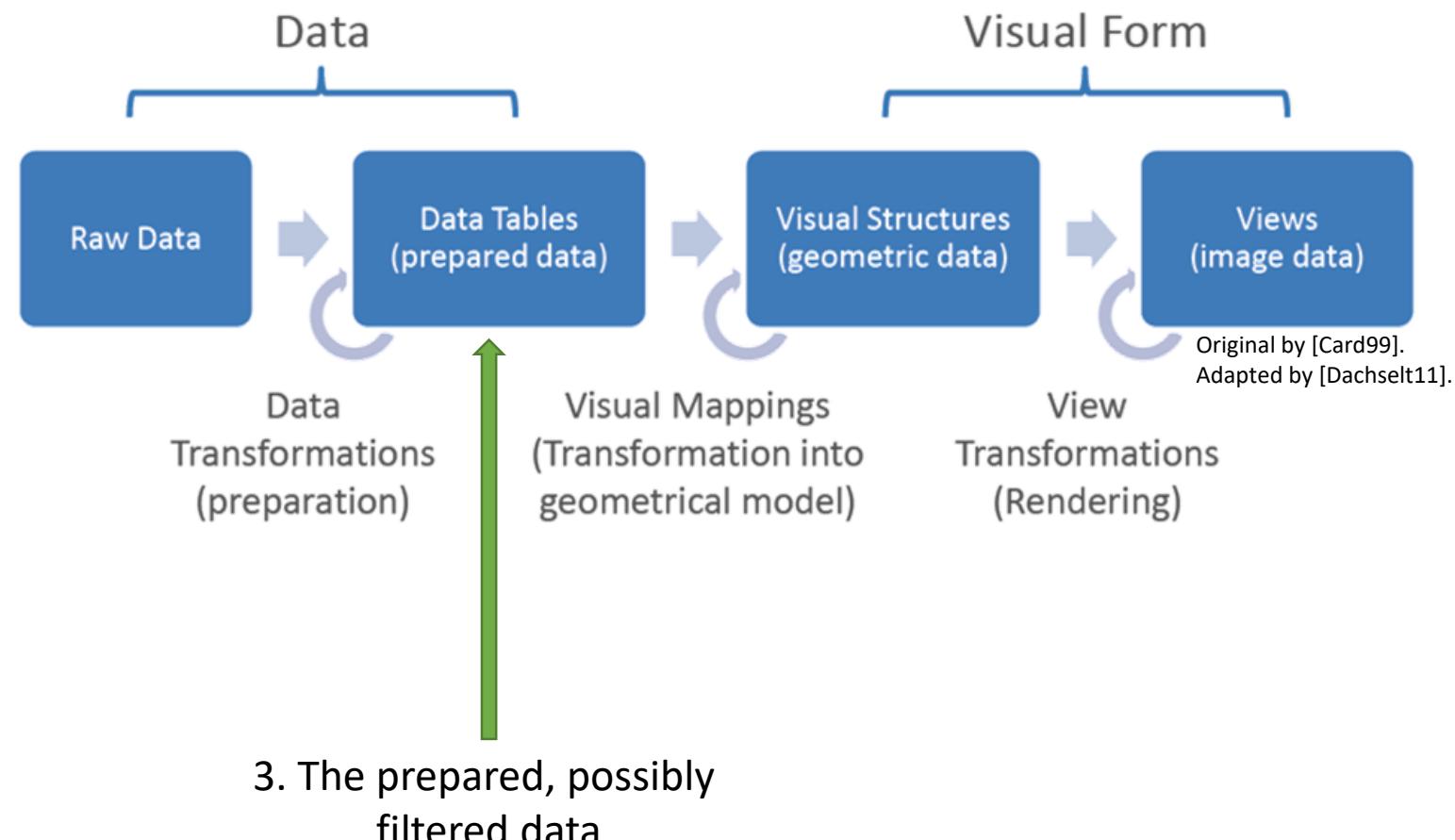


Original by [Card99].
Adapted by [Dachselt11].

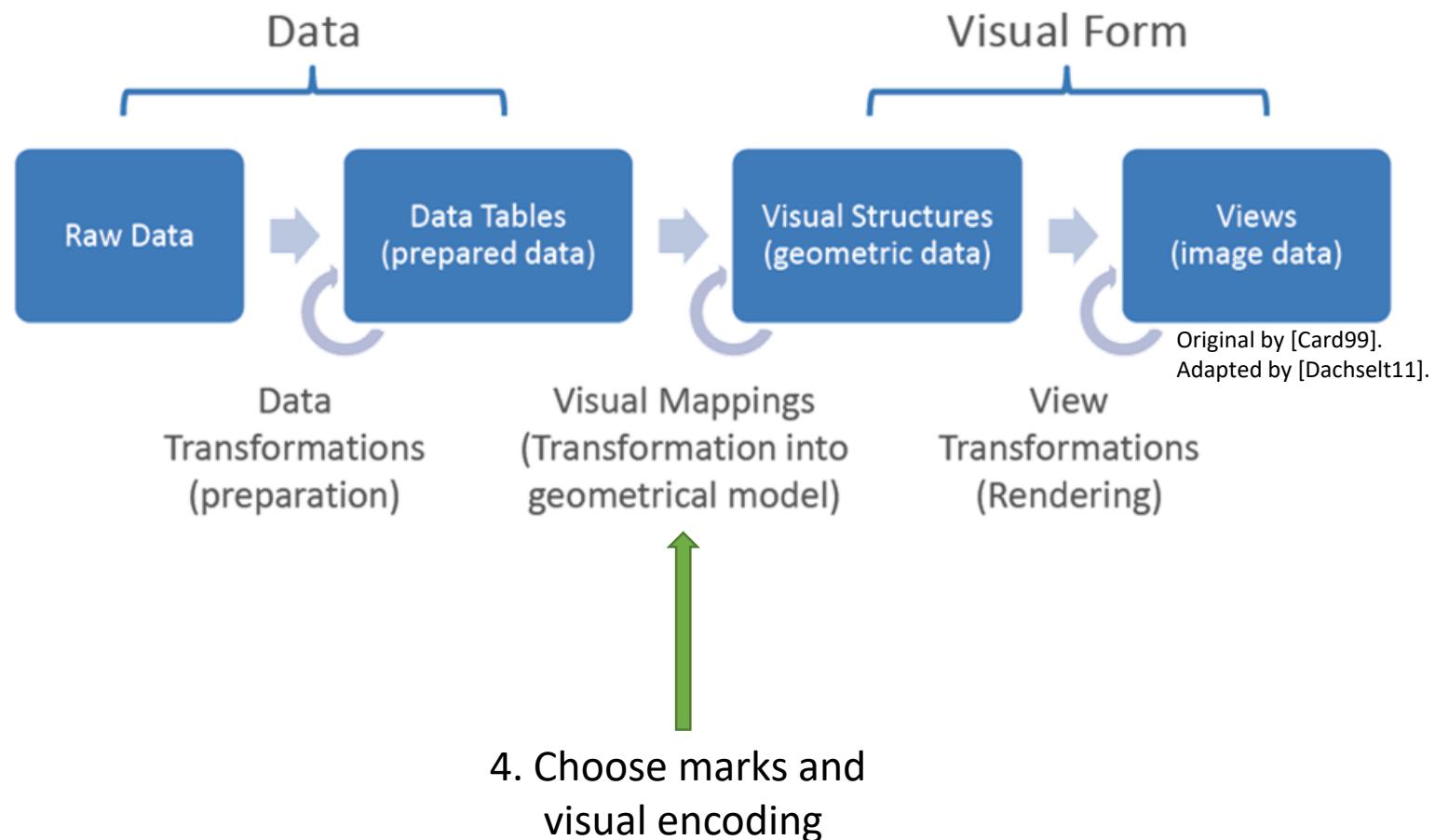
Data Transformations

- Modeling
 - Convert raw data to an organized logical structure
- Pre-processing
 - Normalization, interpolation, subsampling,...
- Filtering
 - Show subset of the data
 - Alternative: use dimensionality reduction technique
 - E.g., principal component analysis (PCA), multi-dimensional scaling (MDS)
 - Often used to explain data variance and detect trends

How Do We Visualize Multivariate Data?



How Do We Visualize Multivariate Data?

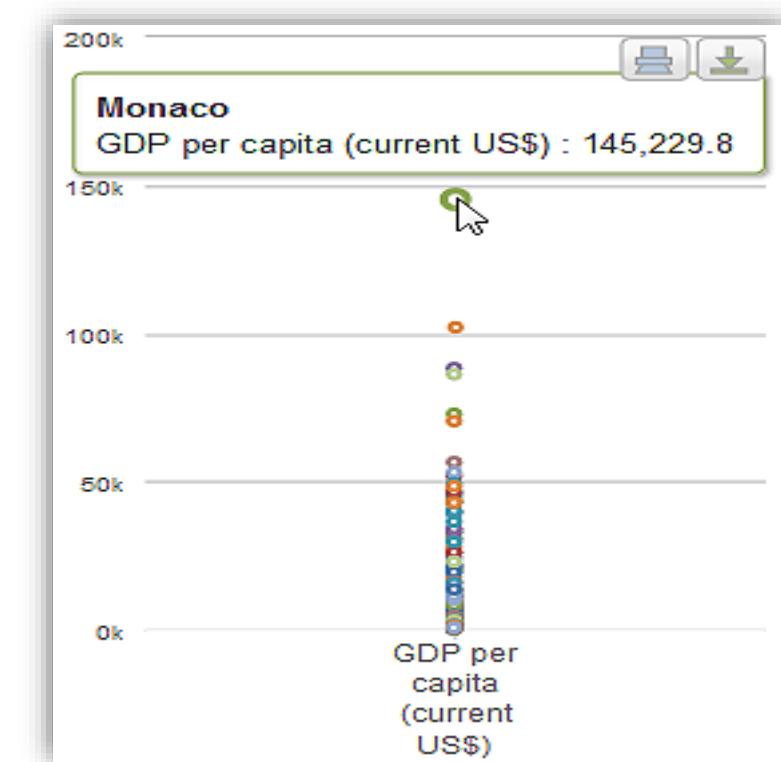


Marks and Visual Encoding

- Find a visual mapping of the data
- Let's start with one attribute and a single time step

| A | B |
|------------------------|-----------------------|
| 1 Country Name | GDP per Capita (US\$) |
| 2 Afghanistan | 561.1976175 |
| 3 Germany | 40144.50942 |
| 4 Japan | 43117.76827 |
| 5 Sierra Leone | 447.7533238 |
| 6 United States | 48357.68451 |
| 7 Albania | 3764.326348 |
| 8 Algeria | 4349.569325 |
| 9 American Samoa | |
| 10 Andorra | |
| 11 Angola | 4218.649126 |
| 12 Antigua and Barbuda | 13315.24327 |
| 13 Argentina | 9132.957259 |
| 14 Armenia | 3124.788199 |
| 15 Aruba | 24289.14152 |
| 16 Australia | 51824.79842 |
| 17 Austria | 44723.20394 |
| 18 Azerbaijan | 5843.169753 |
| 19 Bahamas, The | 21881.08394 |

position
+
circle

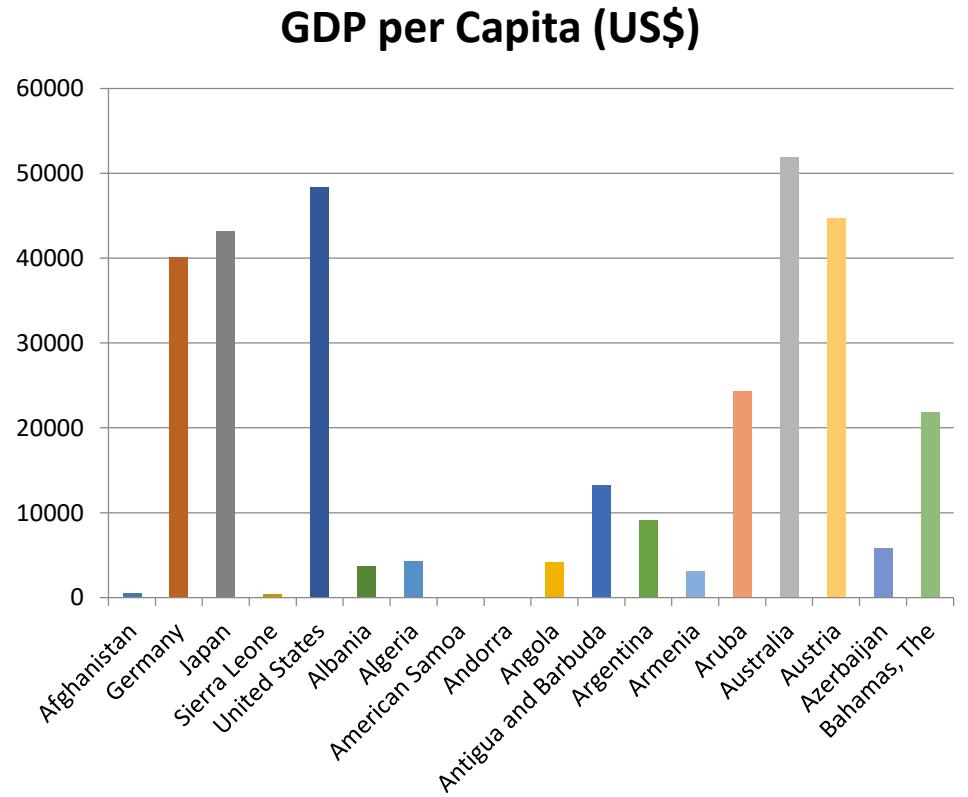


Marks and Visual Encoding

- Find a better mapping that improves visual separation of countries and comparison of attribute values

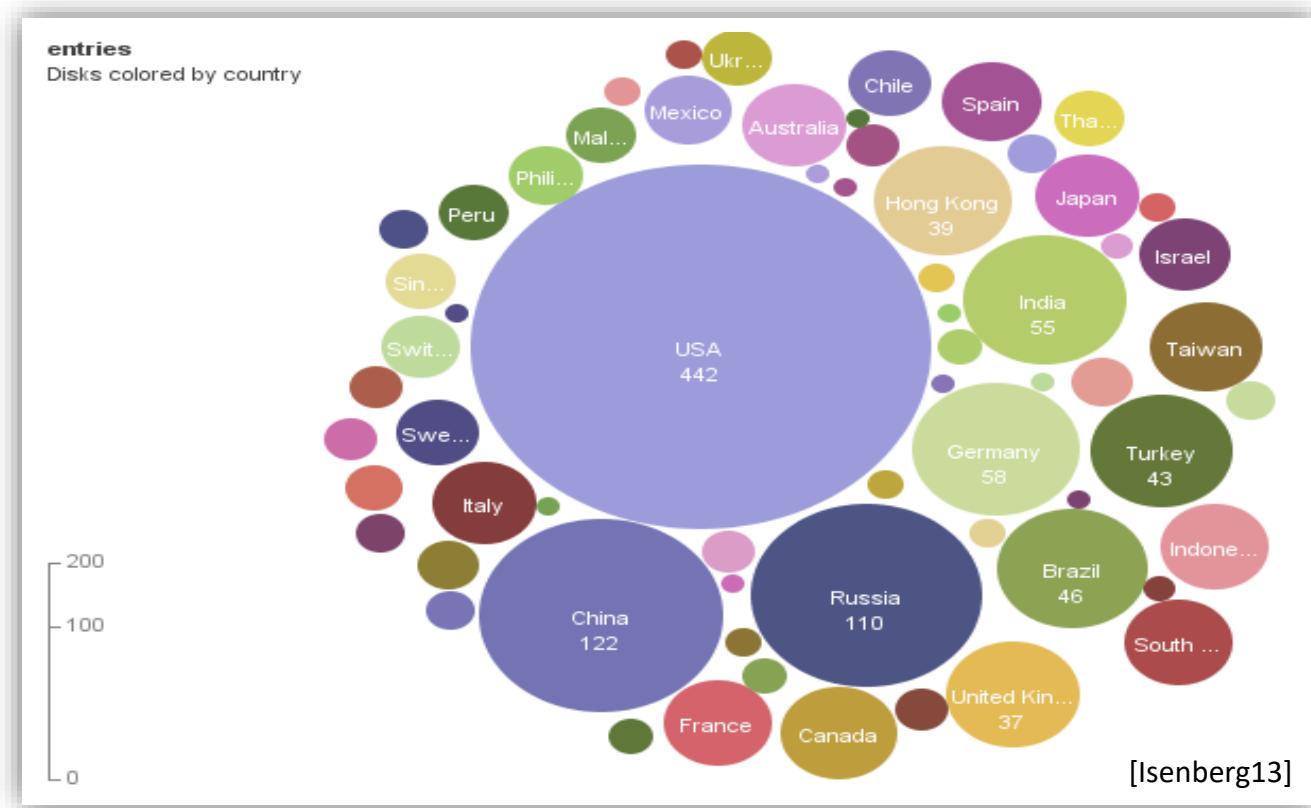
| A | B |
|---------------------|-----------------------|
| Country Name | GDP per Capita (US\$) |
| Afghanistan | 561.1976175 |
| Germany | 40144.50942 |
| Japan | 43117.76827 |
| Sierra Leone | 447.7533238 |
| United States | 48357.68451 |
| Albania | 3764.326348 |
| Algeria | 4349.569325 |
| American Samoa | |
| Andorra | |
| Angola | 4218.649126 |
| Antigua and Barbuda | 13315.24327 |
| Argentina | 9132.957259 |
| Armenia | 3124.788199 |
| Aruba | 24289.14152 |
| Australia | 51824.79842 |
| Austria | 44723.20394 |
| Azerbaijan | 5843.169753 |
| Bahamas, The | 21881.08394 |

position
&
length
+
rectangle



Marks and Visual Encoding

- Spatial layout does not have to be ordered and can be computed by a variety of algorithms



Adding a Second Attribute

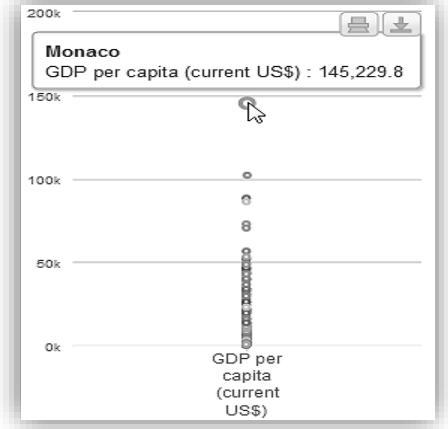
| | A | B | C |
|----|---------------------|-----------------------|-------------------------|
| 1 | Country Name | GDP per Capita (US\$) | Life Expectancy (Years) |
| 2 | Afghanistan | 561.1976175 | 59.6 |
| 3 | Germany | 40144.50942 | 80 |
| 4 | Japan | 43117.76827 | 82.8 |
| 5 | Sierra Leone | 447.7533238 | 44.8 |
| 6 | United States | 48357.68451 | 78.5 |
| 7 | Albania | 3764.326348 | 77 |
| 8 | Algeria | 4349.569325 | 70.6 |
| 9 | American Samoa | | |
| 10 | Andorra | | |
| 11 | Angola | 4218.649126 | 50.7 |
| 12 | Antigua and Barbuda | 13315.24327 | 75.3 |
| 13 | Argentina | 9132.957259 | 75.7 |
| 14 | Armenia | 3124.788199 | 74.2 |
| 15 | Aruba | 24289.14152 | 75 |
| 16 | Australia | 51824.79842 | 81.7 |
| 17 | Austria | 44723.20394 | 80.4 |
| 18 | Azerbaijan | 5843.169753 | 70.5 |
| 19 | Bahamas, The | 21881.08394 | 74.6 |

position
+
circle

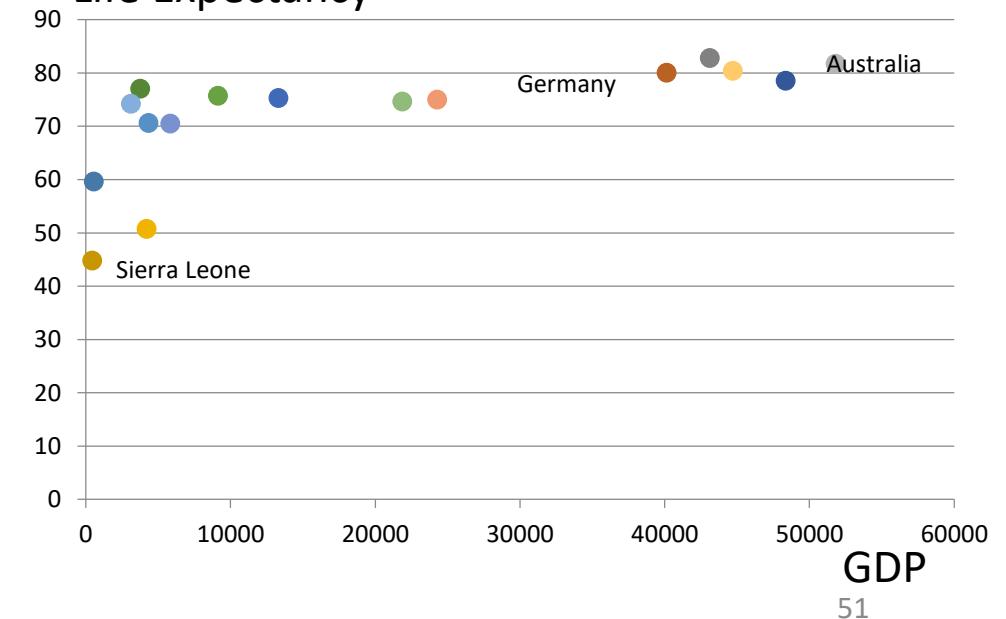


| A | B |
|------------------------|-----------------------|
| 1 Country Name | GDP per Capita (US\$) |
| 2 Afghanistan | 561.1976175 |
| 3 Germany | 40144.50942 |
| 4 Japan | 43117.76827 |
| 5 Sierra Leone | 447.7533238 |
| 6 United States | 48357.68451 |
| 7 Albania | 3764.326348 |
| 8 Algeria | 4349.569325 |
| 9 American Samoa | |
| 10 Andorra | |
| 11 Angola | 4218.649126 |
| 12 Antigua and Barbuda | 13315.24327 |
| 13 Argentina | 9132.957259 |
| 14 Armenia | 3124.788199 |
| 15 Aruba | 24289.14152 |
| 16 Australia | 51824.79842 |
| 17 Austria | 44723.20394 |
| 18 Azerbaijan | 5843.169753 |
| 19 Bahamas, The | 21881.08394 |

position
+
circle

Life Expectancy

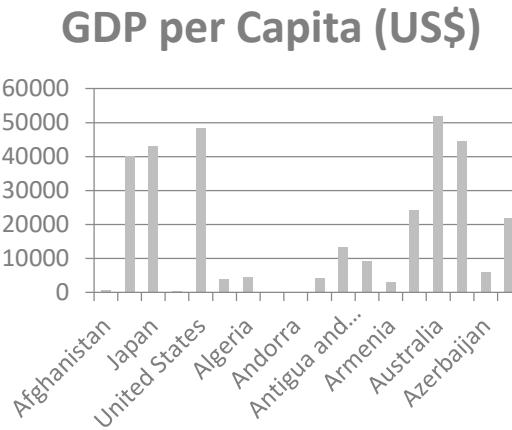


Adding a Second Attribute

| A | B | C |
|---------------------|-----------------------|-------------------------|
| Country Name | GDP per Capita (US\$) | Life Expectancy (Years) |
| Afghanistan | 561.1976175 | 59.6 |
| Germany | 40144.50942 | 80 |
| Japan | 43117.76827 | 82.8 |
| Sierra Leone | 447.7533238 | 44.8 |
| United States | 48357.68451 | 78.5 |
| Albania | 3764.326348 | 77 |
| Algeria | 4349.569325 | 70.6 |
| American Samoa | | |
| Andorra | | |
| Angola | 4218.649126 | 50.7 |
| Antigua and Barbuda | 13315.24327 | 75.3 |
| Argentina | 9132.957259 | 75.7 |
| Armenia | 3124.788199 | 74.2 |
| Aruba | 24289.14152 | 75 |
| Australia | 51824.79842 | 81.7 |
| Austria | 44723.20394 | 80.4 |
| Azerbaijan | 5843.169753 | 70.5 |
| Bahamas, The | 21881.08394 | 74.6 |

| A | B |
|------------------------|-----------------------|
| 1 Country Name | GDP per Capita (US\$) |
| 2 Afghanistan | 561.1976175 |
| 3 Germany | 40144.50942 |
| 4 Japan | 43117.76827 |
| 5 Sierra Leone | 447.7533238 |
| 6 United States | 48357.68451 |
| 7 Albania | 3764.326348 |
| 8 Algeria | 4349.569325 |
| 9 American Samoa | |
| 10 Andorra | |
| 11 Angola | 4218.649126 |
| 12 Antigua and Barbuda | 13315.24327 |
| 13 Argentina | 9132.957259 |
| 14 Armenia | 3124.788199 |
| 15 Aruba | 24289.14152 |
| 16 Australia | 51824.79842 |
| 17 Austria | 44723.20394 |
| 18 Azerbaijan | 5843.169753 |
| 19 Bahamas, The | 21881.08394 |

position
&
length
+
rectangle



position
&
length
+
rectangle

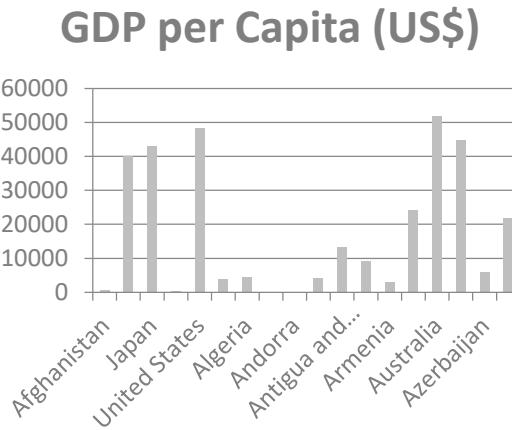


Adding a Second Attribute

| A | B | C | |
|----|---------------------|-----------------------|-------------------------|
| 1 | Country Name | GDP per Capita (US\$) | Life Expectancy (Years) |
| 2 | Afghanistan | 561.1976175 | 59.6 |
| 3 | Germany | 40144.50942 | 80 |
| 4 | Japan | 43117.76827 | 82.8 |
| 5 | Sierra Leone | 447.7533238 | 44.8 |
| 6 | United States | 48357.68451 | 78.5 |
| 7 | Albania | 3764.326348 | 77 |
| 8 | Algeria | 4349.569325 | 70.6 |
| 9 | American Samoa | | |
| 10 | Andorra | | |
| 11 | Angola | 4218.649126 | 50.7 |
| 12 | Antigua and Barbuda | 13315.24327 | 75.3 |
| 13 | Argentina | 9132.957259 | 75.7 |
| 14 | Armenia | 3124.788199 | 74.2 |
| 15 | Aruba | 24289.14152 | 75 |
| 16 | Australia | 51824.79842 | 81.7 |
| 17 | Austria | 44723.20394 | 80.4 |
| 18 | Azerbaijan | 5843.169753 | 70.5 |
| 19 | Bahamas, The | 21881.08394 | 74.6 |

| A | B |
|------------------------|-----------------------|
| 1 Country Name | GDP per Capita (US\$) |
| 2 Afghanistan | 561.1976175 |
| 3 Germany | 40144.50942 |
| 4 Japan | 43117.76827 |
| 5 Sierra Leone | 447.7533238 |
| 6 United States | 48357.68451 |
| 7 Albania | 3764.326348 |
| 8 Algeria | 4349.569325 |
| 9 American Samoa | |
| 10 Andorra | |
| 11 Angola | 4218.649126 |
| 12 Antigua and Barbuda | 13315.24327 |
| 13 Argentina | 9132.957259 |
| 14 Armenia | 3124.788199 |
| 15 Aruba | 24289.14152 |
| 16 Australia | 51824.79842 |
| 17 Austria | 44723.20394 |
| 18 Azerbaijan | 5843.169753 |
| 19 Bahamas, The | 21881.08394 |

position
&
length
+
rectangle

position
&
length
+
rectangle



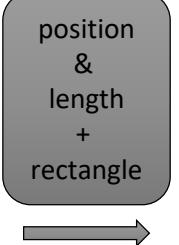
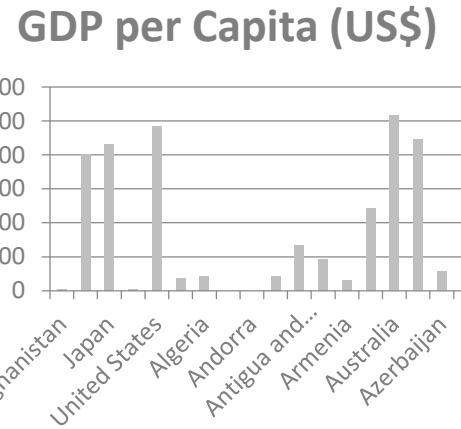
Groups of bars or stacked bars would be possible solutions. However, the different scales of the two attributes are problematic. Any Ideas?

Adding a Second Attribute

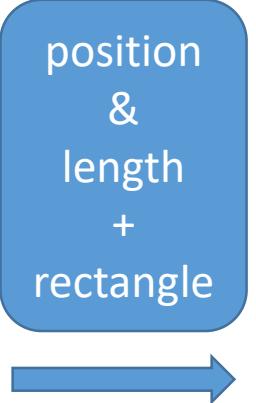
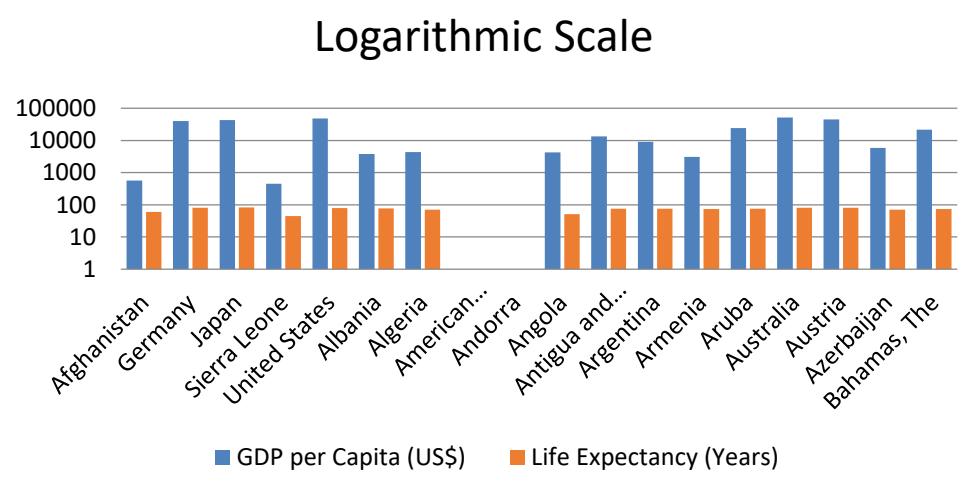
| A | B | C |
|------------------------|-----------------------|-------------------------|
| 1 Country Name | GDP per Capita (US\$) | Life Expectancy (Years) |
| 2 Afghanistan | 561.1976175 | 59.6 |
| 3 Germany | 40144.50942 | 80 |
| 4 Japan | 43117.76827 | 82.8 |
| 5 Sierra Leone | 447.7533238 | 44.8 |
| 6 United States | 48357.68451 | 78.5 |
| 7 Albania | 3764.326348 | 77 |
| 8 Algeria | 4349.569325 | 70.6 |
| 9 American Samoa | | |
| 10 Andorra | | |
| 11 Angola | 4218.649126 | 50.7 |
| 12 Antigua and Barbuda | 13315.24327 | 75.3 |
| 13 Argentina | 9132.957259 | 75.7 |
| 14 Armenia | 3124.788199 | 74.2 |
| 15 Aruba | 24289.14152 | 75 |
| 16 Australia | 51824.79842 | 81.7 |
| 17 Austria | 44723.20394 | 80.4 |
| 18 Azerbaijan | 5843.169753 | 70.5 |
| 19 Bahamas, The | 21881.08394 | 74.6 |

| A | B |
|------------------------|-----------------------|
| 1 Country Name | GDP per Capita (US\$) |
| 2 Afghanistan | 561.1976175 |
| 3 Germany | 40144.50942 |
| 4 Japan | 43117.76827 |
| 5 Sierra Leone | 447.7533238 |
| 6 United States | 48357.68451 |
| 7 Albania | 3764.326348 |
| 8 Algeria | 4349.569325 |
| 10 Andorra | |
| 11 Angola | 4218.649126 |
| 12 Antigua and Barbuda | 13315.24327 |
| 13 Argentina | 9132.957259 |
| 14 Armenia | 3124.788199 |
| 15 Aruba | 24289.14152 |
| 16 Australia | 51824.79842 |
| 17 Austria | 44723.20394 |
| 18 Azerbaijan | 5843.169753 |
| 19 Bahamas, The | 21881.08394 |

position
&
length
+
rectangle

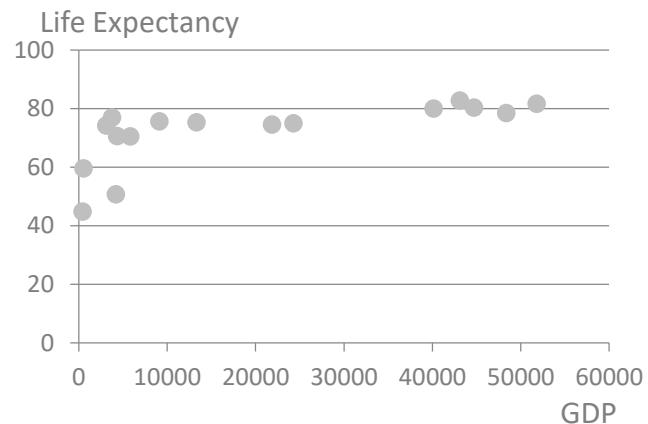
position
&
length
+
rectangle

Adding a Third Attribute

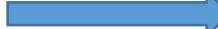
| | A | B | C |
|----|---------------------|-----------------------|-------------------------|
| 1 | Country Name | GDP per Capita (US\$) | Life Expectancy (Years) |
| 2 | Afghanistan | 561.1976175 | 59.6 |
| 3 | Germany | 40144.50942 | 80 |
| 4 | Japan | 43117.76827 | 82.8 |
| 5 | Sierra Leone | 447.7533238 | 44.8 |
| 6 | United States | 48357.68451 | 78.5 |
| 7 | Albania | 3764.326348 | 77 |
| 8 | Algeria | 4349.569325 | 70.6 |
| 9 | American Samoa | | |
| 10 | Andorra | | |
| 11 | Angola | 4218.649126 | 50.7 |
| 12 | Antigua and Barbuda | 13315.24327 | 75.3 |
| 13 | Argentina | 9132.957259 | 75.7 |
| 14 | Armenia | 3124.788199 | 74.2 |
| 15 | Aruba | 24289.14152 | 75 |
| 16 | Australia | 51824.79842 | 81.7 |
| 17 | Austria | 44723.20394 | 80.4 |
| 18 | Azerbaijan | 5843.169753 | 70.5 |
| 19 | Bahamas, The | 21881.08394 | 74.6 |

position
+
circle

| | A | B | C | D |
|----|---------------------|-----------------------|-------------------------|----------------|
| 1 | Country Name | GDP per Capita (US\$) | Life Expectancy (Years) | Fertility Rate |
| 2 | Afghanistan | 561.1976175 | 59.6 | 5.659 |
| 3 | Germany | 40144.50942 | 80 | 1.39 |
| 4 | Japan | 43117.76827 | 82.8 | 1.39 |
| 5 | Sierra Leone | 447.7533238 | 44.8 | 4.943 |
| 6 | United States | 48357.68451 | 78.5 | 1.931 |
| 7 | Albania | 3764.326348 | 77 | 1.741 |
| 8 | Algeria | 4349.569325 | 70.6 | 2.817 |
| 9 | American Samoa | | | 1.22 |
| 10 | Andorra | | | |
| 11 | Angola | 4218.649126 | 50.7 | 6.218 |
| 12 | Antigua and Barbuda | 13315.24327 | 75.3 | 2.13 |
| 13 | Argentina | 9132.957259 | 75.7 | 2.215 |
| 14 | Armenia | 3124.788199 | 74.2 | 1.738 |
| 15 | Aruba | 24289.14152 | 75 | 1.701 |
| 16 | Australia | 51824.79842 | 81.7 | 1.927 |
| 17 | Austria | 44723.20394 | 80.4 | 1.44 |
| 18 | Azerbaijan | 5843.169753 | 70.5 | 1.92 |
| 19 | Bahamas, The | 21881.08394 | 74.6 | 1.901 |

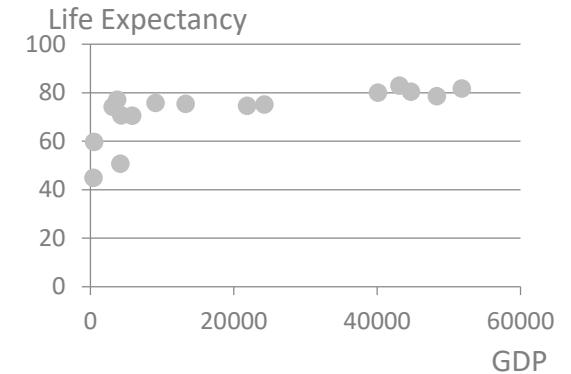
position
+
circle




Adding a Third Attribute

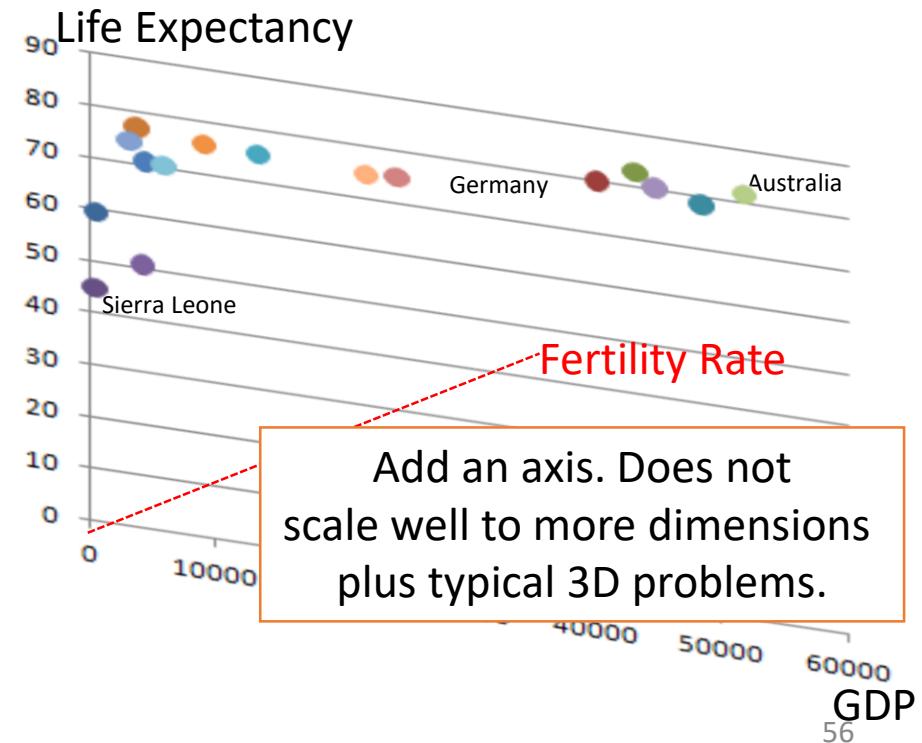
| A | B | C |
|---------------------|-----------------------|-------------------------|
| Country Name | GDP per Capita (US\$) | Life Expectancy (Years) |
| Afghanistan | 561.1976175 | 59.6 |
| Germany | 40144.50942 | 80 |
| Japan | 43117.76827 | 82.8 |
| Sierra Leone | 447.7533238 | 44.8 |
| United States | 48357.68451 | 78.5 |
| Albania | 3764.326348 | 77 |
| Algeria | 4349.569325 | 70.6 |
| American Samoa | | |
| Andorra | | |
| Angola | 4218.649126 | 50.7 |
| Antigua and Barbuda | 13315.24327 | 75.3 |
| Argentina | 9132.957259 | 75.7 |
| Armenia | 3124.788199 | 74.2 |
| Aruba | 24289.14152 | 75 |
| Australia | 51824.79842 | 81.7 |
| Austria | 44723.20394 | 80.4 |
| Azerbaijan | 5843.169753 | 70.5 |
| Bahamas, The | 21881.08394 | 74.6 |

position
+
circle

| A | B | C | D |
|---------------------|-----------------------|-------------------------|----------------|
| Country Name | GDP per Capita (US\$) | Life Expectancy (Years) | Fertility Rate |
| Afghanistan | 561.1976175 | 59.6 | 5.659 |
| Germany | 40144.50942 | 80 | 1.39 |
| Japan | 43117.76827 | 82.8 | 1.39 |
| Sierra Leone | 447.7533238 | 44.8 | 4.943 |
| United States | 48357.68451 | 78.5 | 1.931 |
| Albania | 3764.326348 | 77 | 1.741 |
| Algeria | 4349.569325 | 70.6 | 2.817 |
| American Samoa | | | |
| Andorra | | | 1.22 |
| Angola | 4218.649126 | 50.7 | 6.218 |
| Antigua and Barbuda | 13315.24327 | 75.3 | 2.13 |
| Argentina | 9132.957259 | 75.7 | 2.215 |
| Armenia | 3124.788199 | 74.2 | 1.738 |
| Aruba | 24289.14152 | 75 | 1.701 |
| Australia | 51824.79842 | 81.7 | 1.927 |
| Austria | 44723.20394 | 80.4 | 1.44 |
| Azerbaijan | 5843.169753 | 70.5 | 1.92 |
| Bahamas, The | 21881.08394 | 74.6 | 1.901 |

position
+
circle

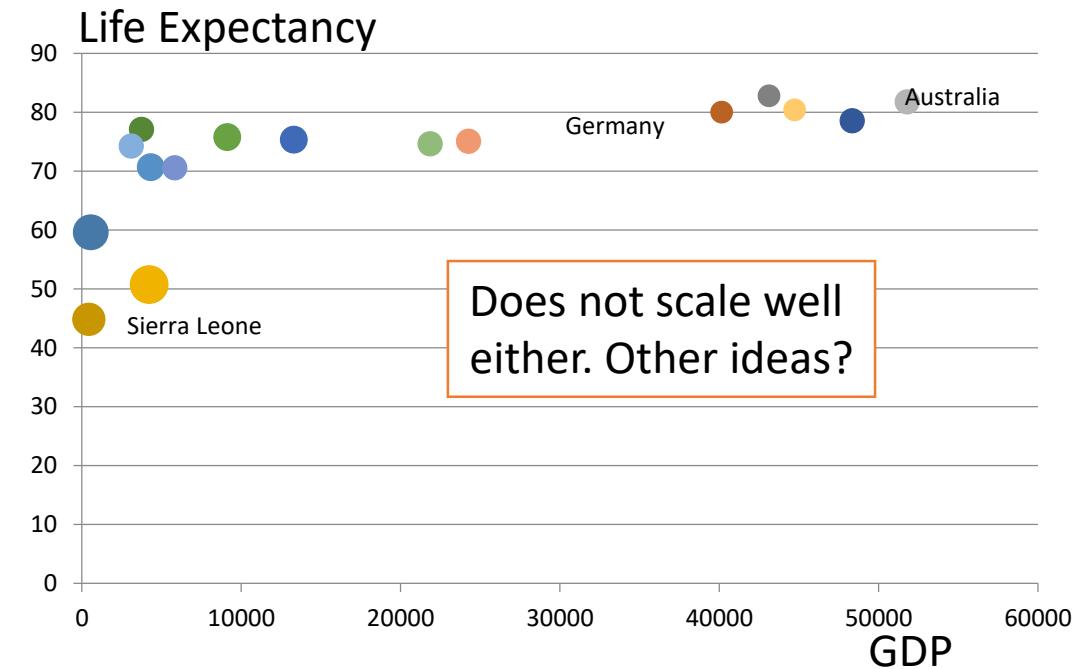



Adding a Third Attribute

| A | B | | C | D |
|----|---------------------|-----------------------|-------------------------|----------------|
| 1 | Country Name | GDP per Capita (US\$) | Life Expectancy (Years) | Fertility Rate |
| 2 | Afghanistan | 561.1976175 | 59.6 | 5.659 |
| 3 | Germany | 40144.50942 | 80 | 1.39 |
| 4 | Japan | 43117.76827 | 82.8 | 1.39 |
| 5 | Sierra Leone | 447.7533238 | 44.8 | 4.943 |
| 6 | United States | 48357.68451 | 78.5 | 1.931 |
| 7 | Albania | 3764.326348 | 77 | 1.741 |
| 8 | Algeria | 4349.569325 | 70.6 | 2.817 |
| 9 | American Samoa | | | 1.22 |
| 10 | Andorra | | | 6.218 |
| 11 | Angola | 4218.649126 | 50.7 | 2.13 |
| 12 | Antigua and Barbuda | 13315.24327 | 75.3 | 2.215 |
| 13 | Argentina | 9132.957259 | 75.7 | 1.738 |
| 14 | Armenia | 3124.788199 | 74.2 | 1.701 |
| 15 | Aruba | 24289.14152 | 75 | 1.927 |
| 16 | Australia | 51824.79842 | 81.7 | 1.44 |
| 17 | Austria | 44723.20394 | 80.4 | 1.901 |
| 18 | Azerbaijan | 5843.169753 | 70.5 | 1.92 |
| 19 | Bahamas, The | 21881.08394 | 74.6 | |

| A | B | C |
|------------------------|-----------------------|-------------------------|
| 1 Country Name | GDP per Capita (US\$) | Life Expectancy (Years) |
| 2 Afghanistan | 561.1976175 | 59.6 |
| 3 Germany | 40144.50942 | 80 |
| 4 Japan | 43117.76827 | 82.8 |
| 5 Sierra Leone | 447.7533238 | 44.8 |
| 6 United States | 48357.68451 | 78.5 |
| 7 Albania | 3764.326348 | 77 |
| 8 Algeria | 4349.569325 | 70.6 |
| 9 American Samoa | | |
| 10 Andorra | | |
| 11 Angola | 4218.649126 | 50.7 |
| 12 Antigua and Barbuda | 13315.24327 | 75.3 |
| 13 Argentina | 9132.957259 | 75.7 |
| 14 Armenia | 3124.788199 | 74.2 |
| 15 Aruba | 24289.14152 | 75 |
| 16 Australia | 51824.79842 | 81.7 |
| 17 Austria | 44723.20394 | 80.4 |
| 18 Azerbaijan | 5843.169753 | 70.5 |
| 19 Bahamas, The | 21881.08394 | 74.6 |

position
&
size
+
circle

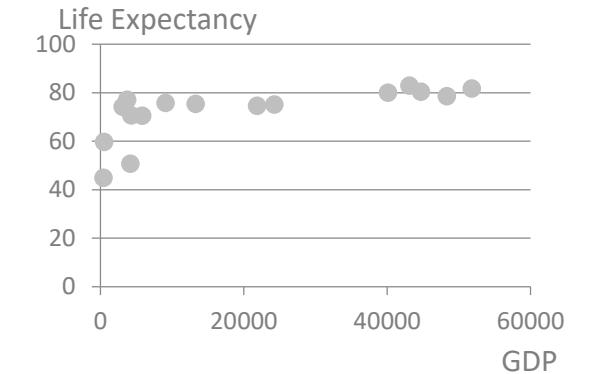
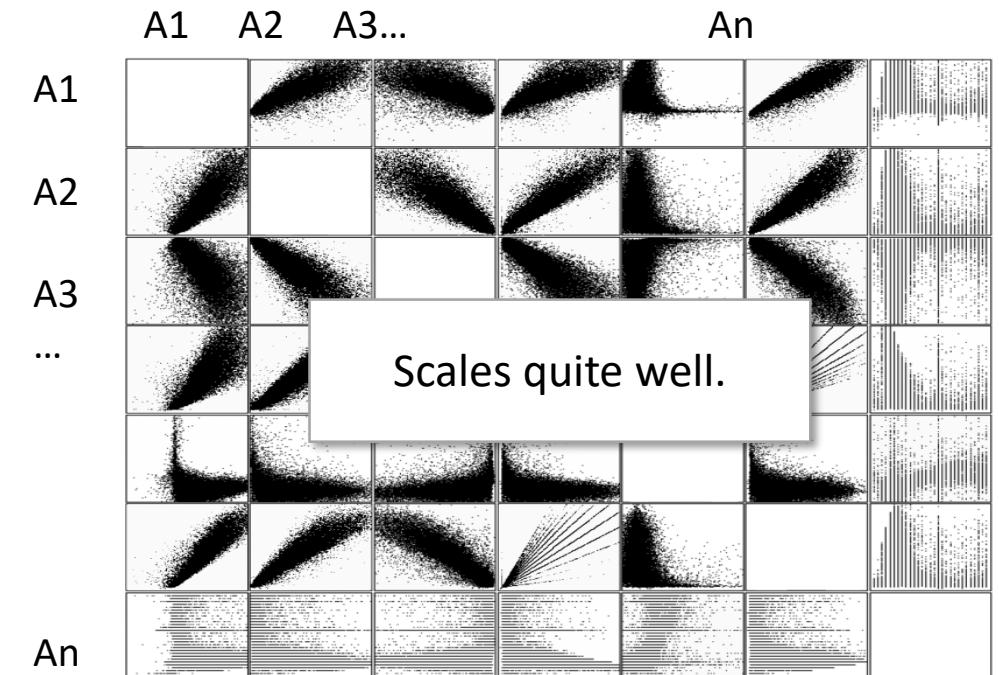


Adding a Third Attribute

| A | B | C | D | |
|----|---------------------|-----------------------|-------------------------|----------------|
| 1 | Country Name | GDP per Capita (US\$) | Life Expectancy (Years) | Fertility Rate |
| 2 | Afghanistan | 561.1976175 | 59.6 | 5.659 |
| 3 | Germany | 40144.50942 | 80 | 1.39 |
| 4 | Japan | 43117.76827 | 82.8 | 1.39 |
| 5 | Sierra Leone | 447.7533238 | 44.8 | 4.943 |
| 6 | United States | 48357.68451 | 78.5 | 1.931 |
| 7 | Albania | 3764.326348 | 77 | 1.741 |
| 8 | Algeria | 4349.569325 | 70.6 | 2.817 |
| 9 | American Samoa | | | 1.22 |
| 10 | Andorra | | | |
| 11 | Angola | 4218.649126 | 50.7 | 6.218 |
| 12 | Antigua and Barbuda | 13315.24327 | 75.3 | 2.13 |
| 13 | Argentina | 9132.957259 | 75.7 | 2.215 |
| 14 | Armenia | 3124.788199 | 74.2 | 1.738 |
| 15 | Aruba | 24289.14152 | 75 | 1.701 |
| 16 | Australia | 51824.79842 | 81.7 | 1.927 |
| 17 | Austria | 44723.20394 | 80.4 | 1.44 |
| 18 | Azerbaijan | 5843.169753 | 70.5 | 1.92 |
| 19 | Bahamas, The | 21881.08394 | 74.6 | 1.901 |

| A | B | C |
|------------------------|-----------------------|-------------------------|
| 1 Country Name | GDP per Capita (US\$) | Life Expectancy (Years) |
| 2 Afghanistan | 561.1976175 | 59.6 |
| 3 Germany | 40144.50942 | 80 |
| 4 Japan | 43117.76827 | 82.8 |
| 5 Sierra Leone | 447.7533238 | 44.8 |
| 6 United States | 48357.68451 | 78.5 |
| 7 Albania | 3764.326348 | 77 |
| 8 Algeria | 4349.569325 | 70.6 |
| 9 American Samoa | | |
| 10 Andorra | | |
| 11 Angola | 4218.649126 | 50.7 |
| 12 Antigua and Barbuda | 13315.24327 | 75.3 |
| 13 Argentina | 9132.957259 | 75.7 |
| 14 Armenia | 3124.788199 | 74.2 |
| 15 Aruba | 24289.14152 | 75 |
| 16 Australia | 51824.79842 | 81.7 |
| 17 Austria | 44723.20394 | 80.4 |
| 18 Azerbaijan | 5843.169753 | 70.5 |
| 19 Bahamas, The | 21881.08394 | 74.6 |

position
+
circle

Univariate Data

Single Number

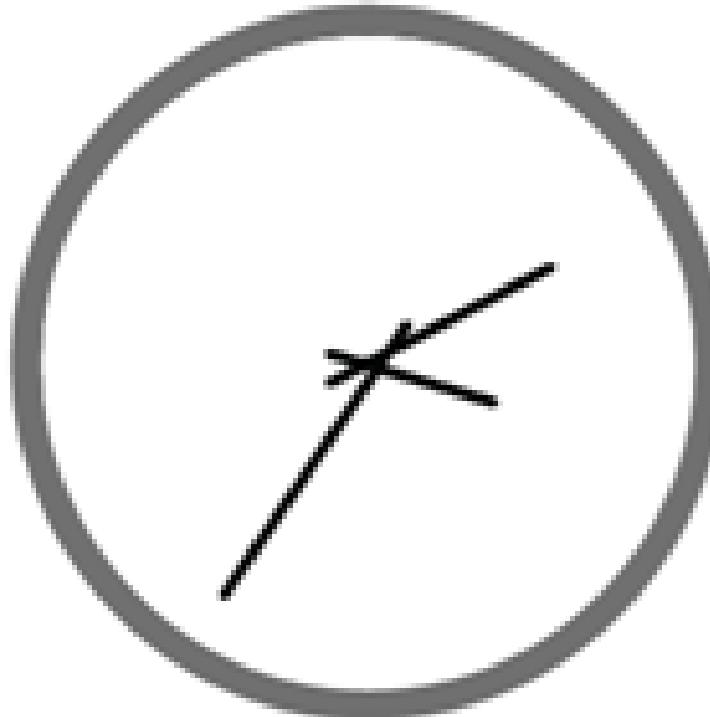
May encoding a single number pose problems?

- Yes!
- Example: airplane altimeter
 - Little hand: $l \times 10,000$ foot
 - Middle hand: $m \times 1,000$ foot
 - Big hand: $n \times 100$ foot
- Pilots are trained in reading altimeter
- But, distraction may cause reading mistakes
- Reason: change blindness



[Spence07]

Change Blindness

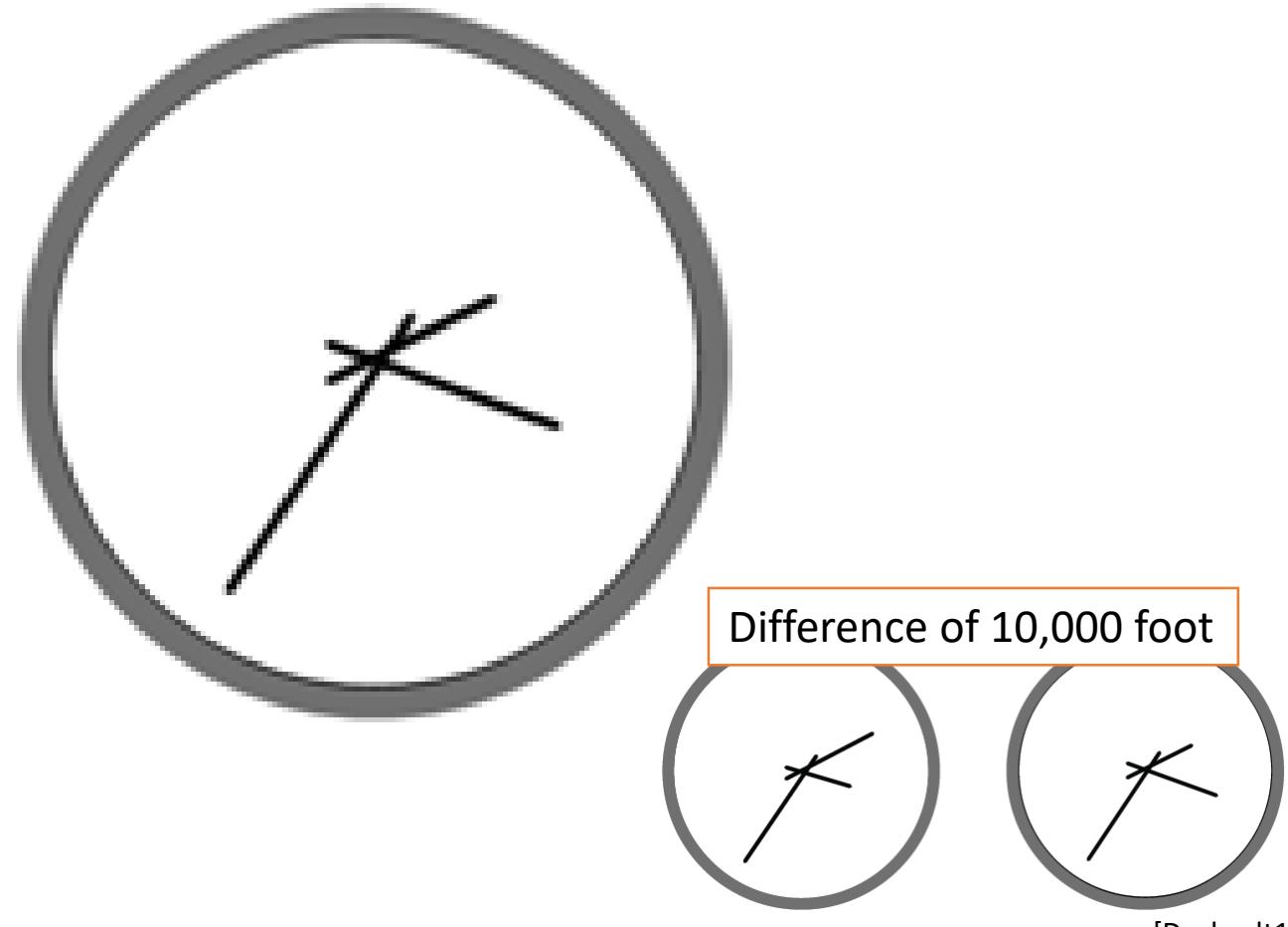


[Dachselt11]

Change Blindness – Distraction



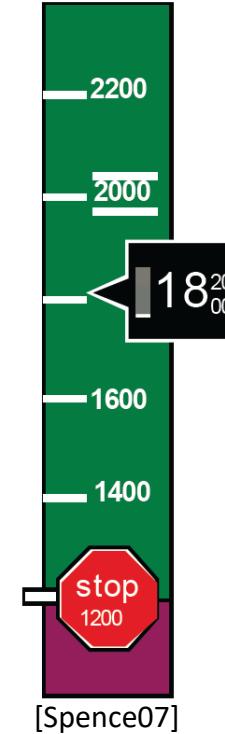
Change Blindness



Single Number

A modern altimeter

- Facilitates preattentive risk assessment
 - Critical point (1200) and range below (divided scale)
 - Black marker at plane's current altitude



Gauge Chart

- Visualization of a single number
- Inspired by speedometer
- High familiarity
- Puts the number into context, i.e. the range of possible values
- Critical values might be emphasized, similar to city speed limit at 50km/h
- Criticism:
 - Take up lots of space
 - Do not show history but only the current value

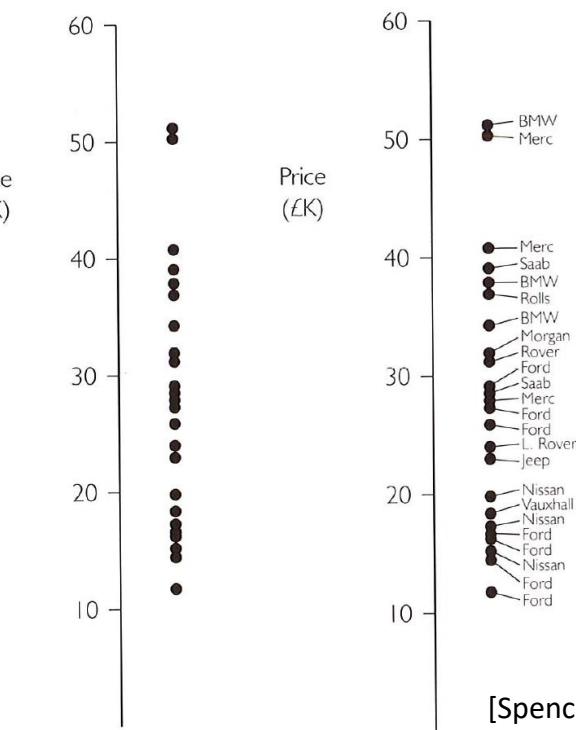


© Oracle®, 2010/12

Point Plot

- Distribution of data may be inferred
- Average may be roughly estimated
- Minimum and maximum are clearly visible

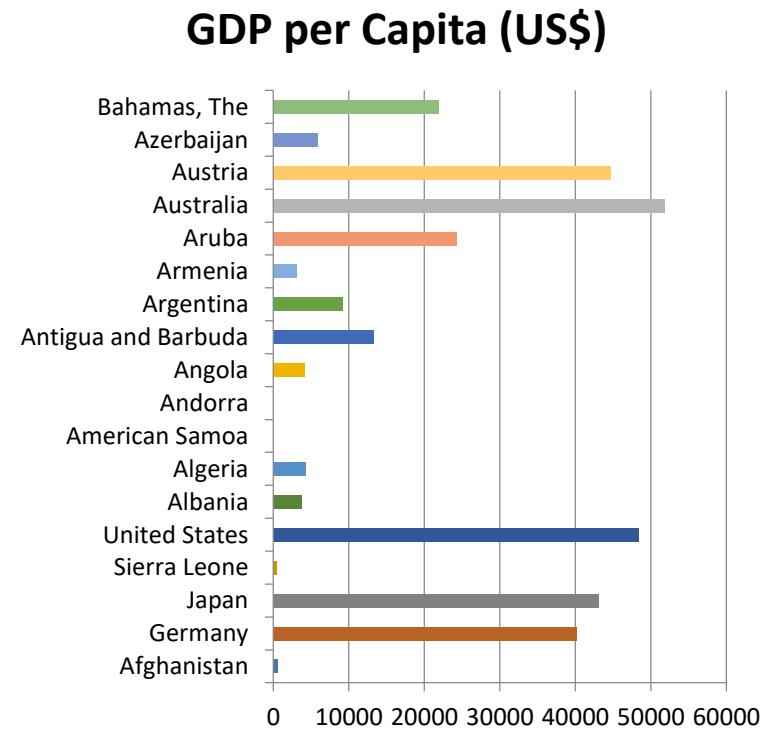
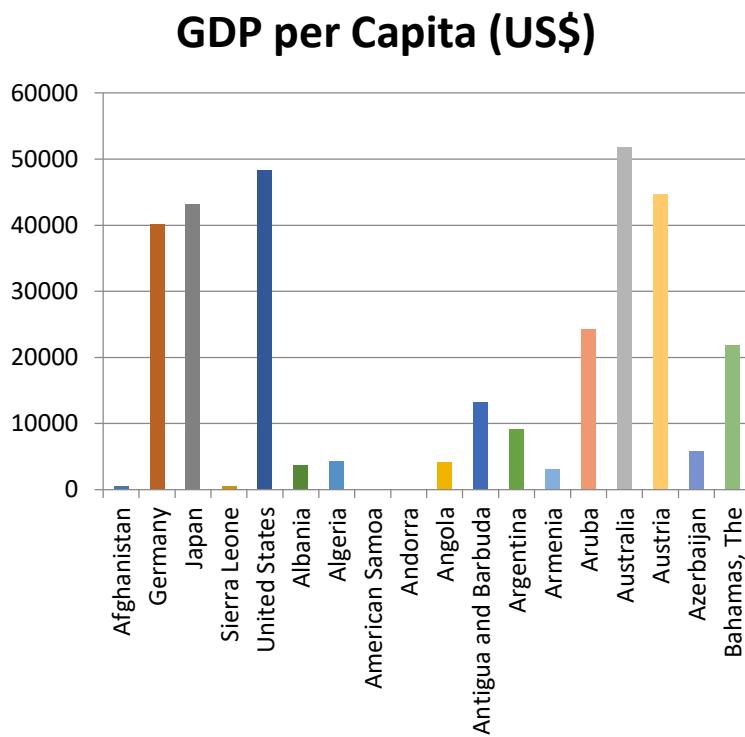
| Car | Price (£) |
|------------|-----------|
| BMW | 51,395 |
| Mercedes | 50,850 |
| Mercedes | 41,000 |
| Saab | 39,085 |
| BMW | 38,000 |
| Rolls | 36,950 |
| BMW | 34,550 |
| Morgan | 32,000 |
| Rover | 31,300 |
| Ford | 29,250 |
| Saab | 28,750 |
| Mercedes | 28,000 |
| Ford | 27,600 |
| Ford | 25,950 |
| Land Rover | 24,000 |
| Jeep | 23,200 |
| Nissan | 20,000 |
| Vauxhall | 18,500 |
| Nissan | 17,400 |
| Ford | 17,000 |
| Ford | 16,500 |
| Nissan | 15,500 |
| Ford | 14,950 |
| Ford | 12,000 |



Labeling all
points can be
challenging

Bar Chart

- Vertical or horizontal bars with lengths proportional to encoded values
- Suitable for displaying discrete values / categories



Histogram

- Binned representation of continuous data
- Each vertical bar (bin) represents a range of values
- Bar length encodes number of data objects having a value within the respective range
- In contrast to bar charts, bars CANNOT be rearranged

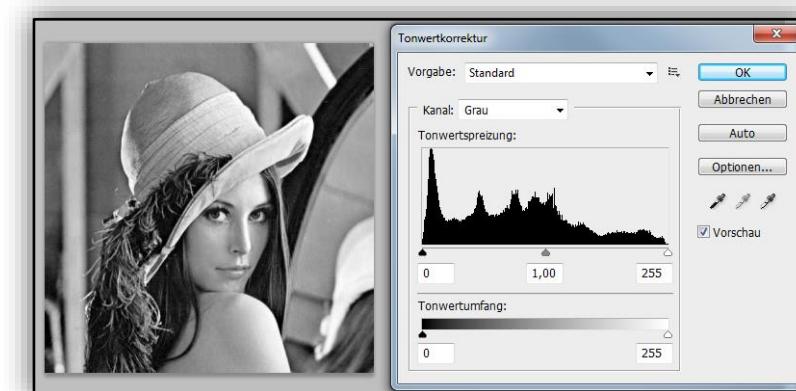
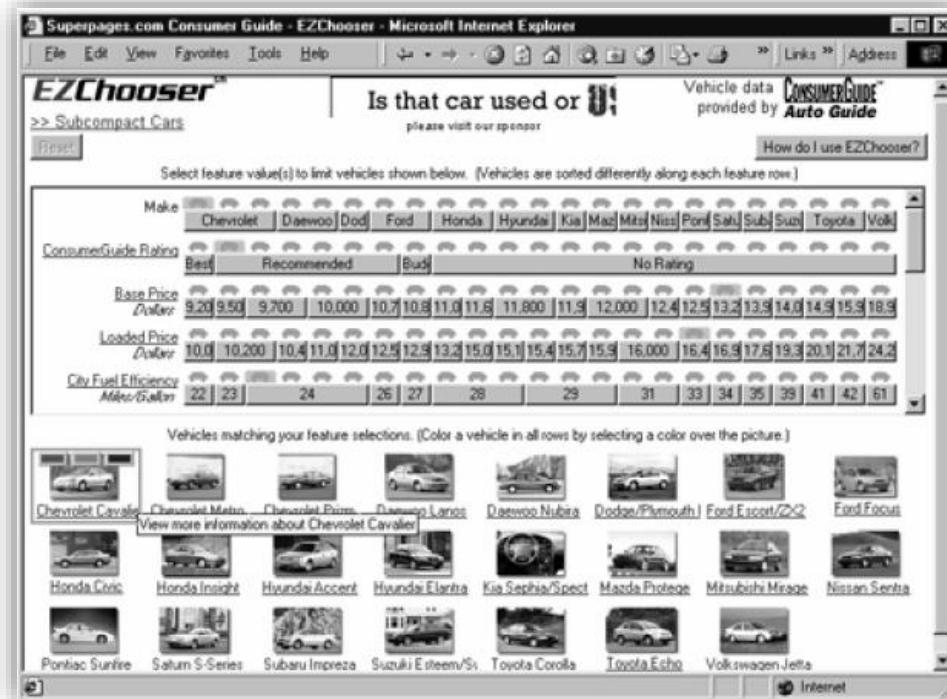
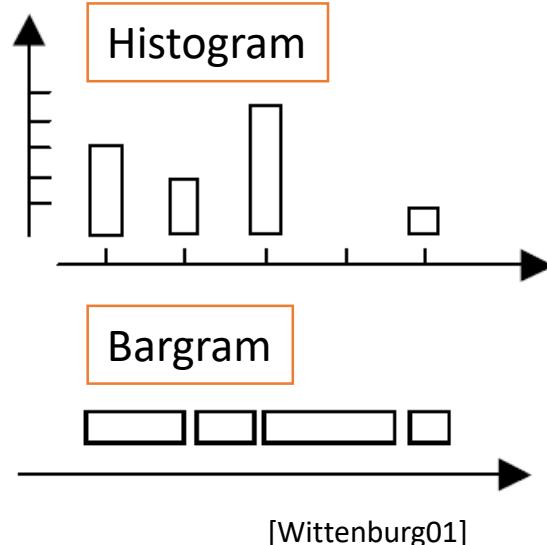


Image histograms show tonal distribution in a digital image

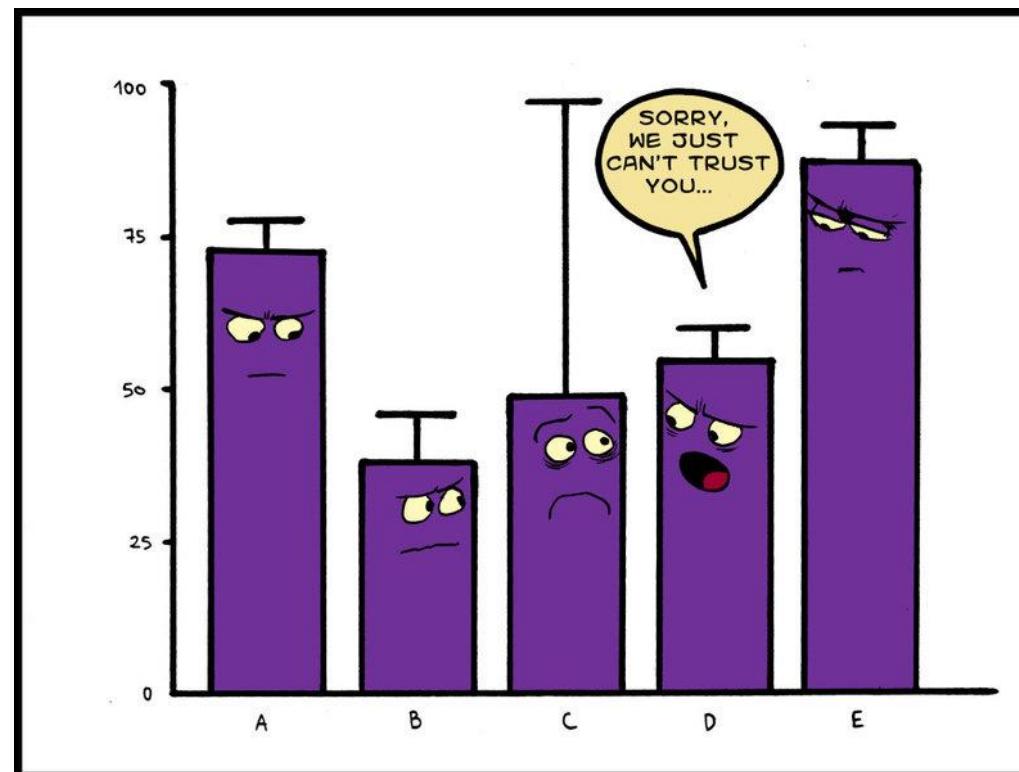
Bargram [Wittenburg01]

- Derived from histogram by “tipping over” the bars
- Layout of bars end-to-end ignoring null bins
- Reduced information but less vertical space covered



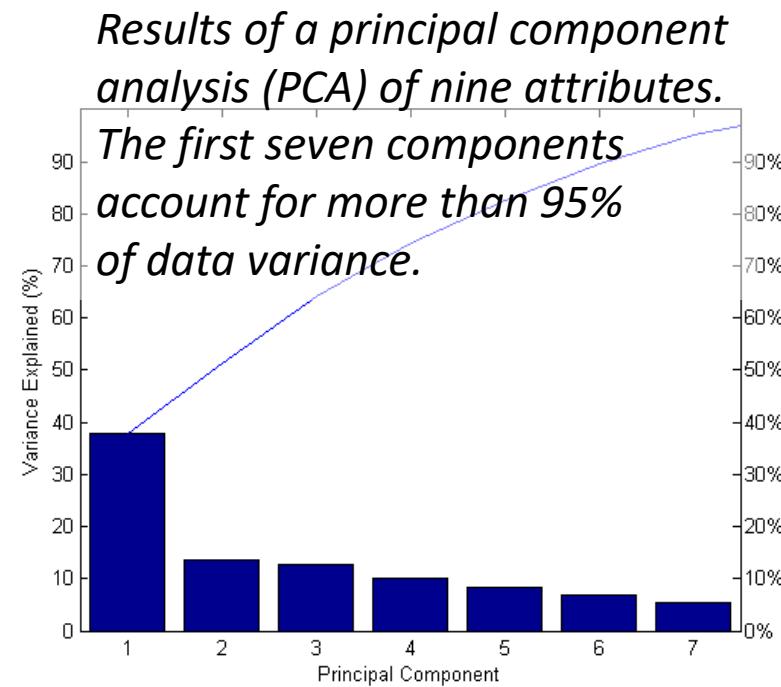
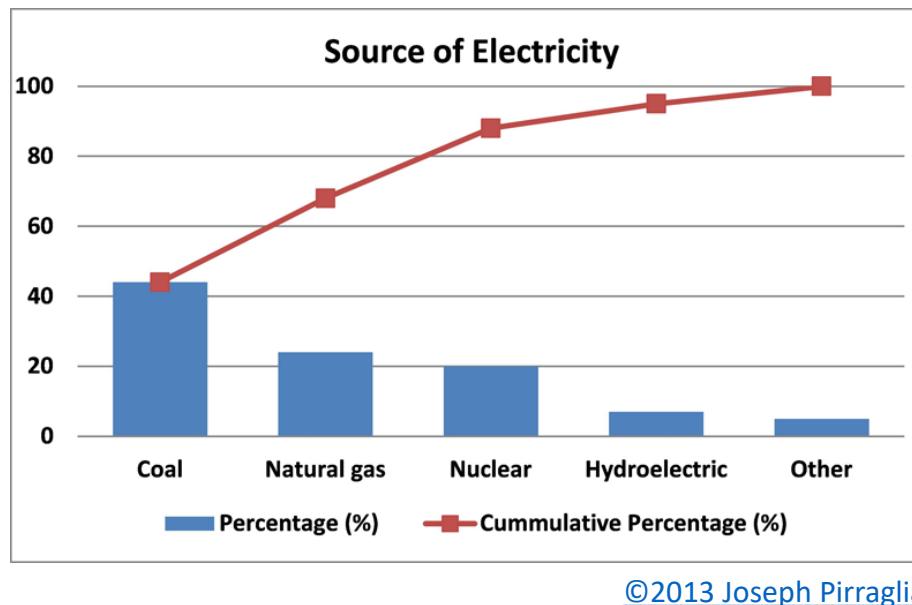
Error Bars

- Indicate variability or uncertainty of the data by encoding standard deviation, standard error or a confidence interval



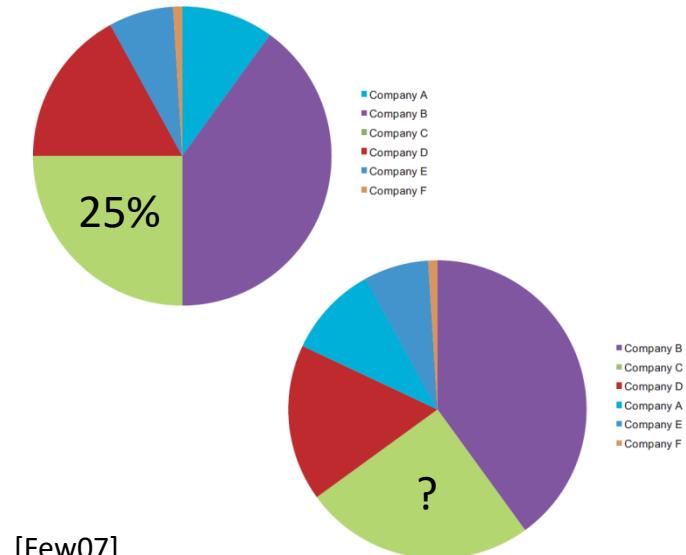
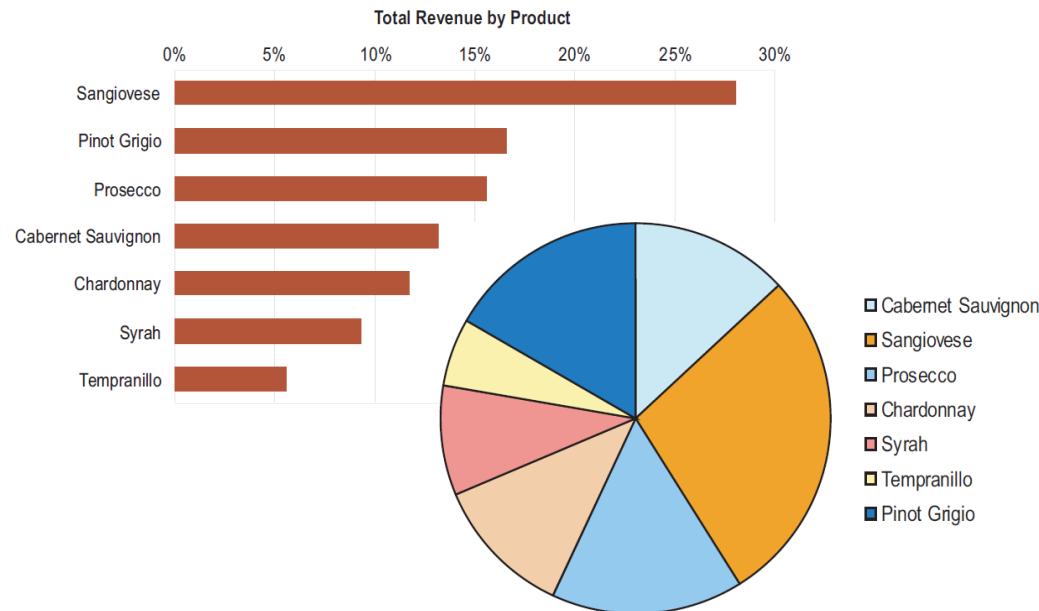
Pareto Chart

- Named after famous economist Vilfredo Pareto
- Bars encode individual percentage values while line encodes the cumulative total



Pie Chart

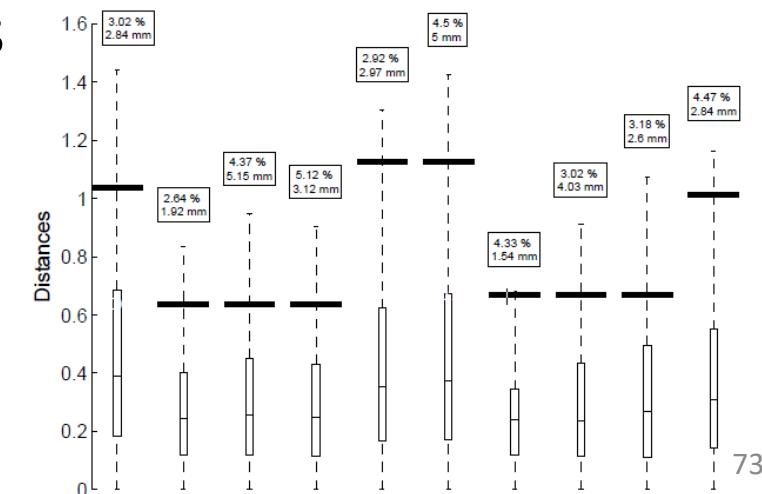
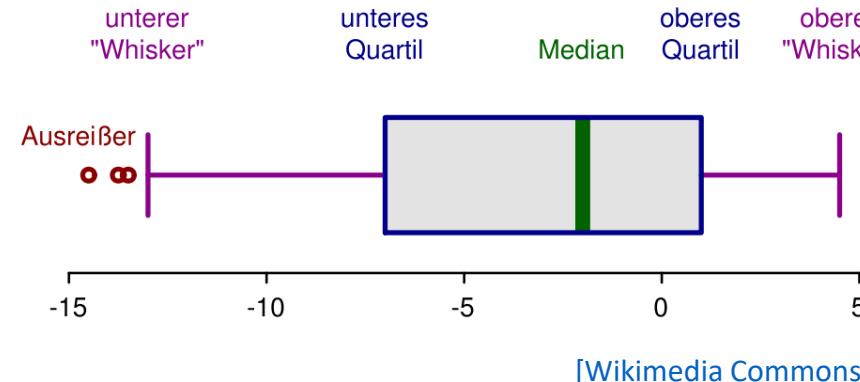
- Circular chart divided into pieces illustrating numerical proportion of a whole
- Much criticism, e.g., Few's "Save the Pies for Dessert"
 - Difficult to compare pieces in a chart or across charts
 - Readability depends on orientation of pieces



[Few07]

Box Plot

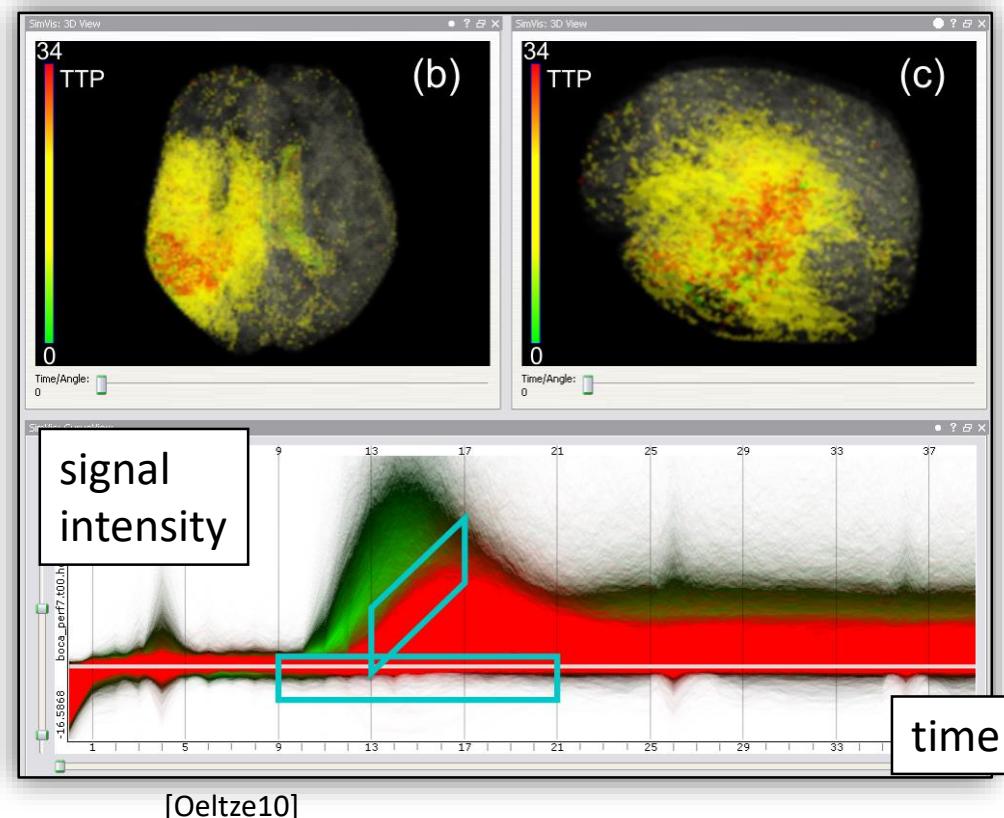
- Also known as Tukey boxplot or box-and-whisker plot
- Shows data dispersion, skewness, and outliers
- Box represents three quartiles (25%, 50%, 75%)
- 2 extensions (*whiskers*) at $1.5 \times \text{IQR}$ (possibly clamped at existing data values)
- Outliers located beyond ends of the whiskers



Bivariate Data

Line Chart / Time Series Chart

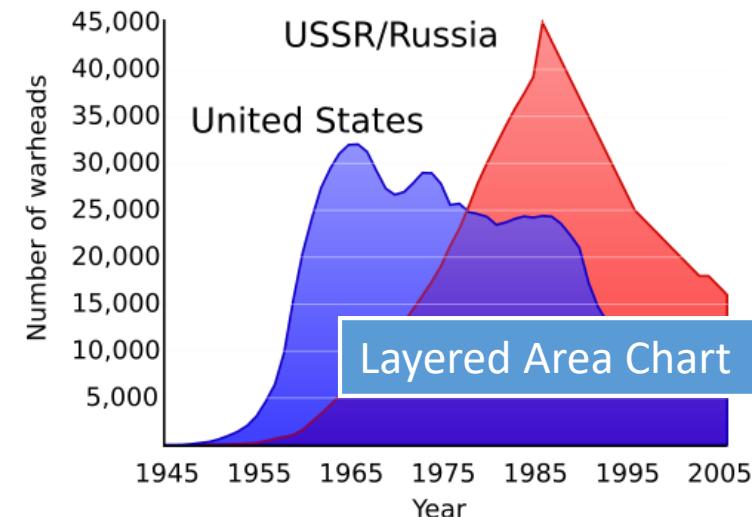
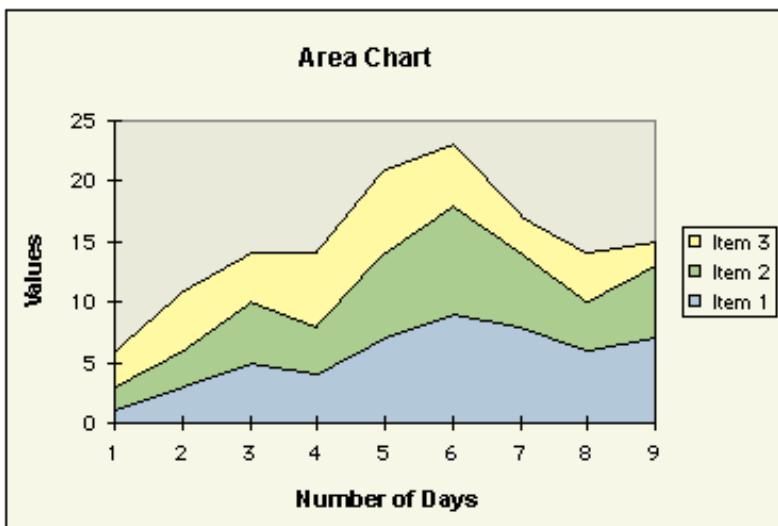
- Displays a series of data points connected by straight line segments
- Well suited for showing trends over time (x-axis)



Signal curves of all voxels in a MR cerebral perfusion dataset (a). Interesting curves are brushed (quads) and corresponding voxels are highlighted (color; b,c).

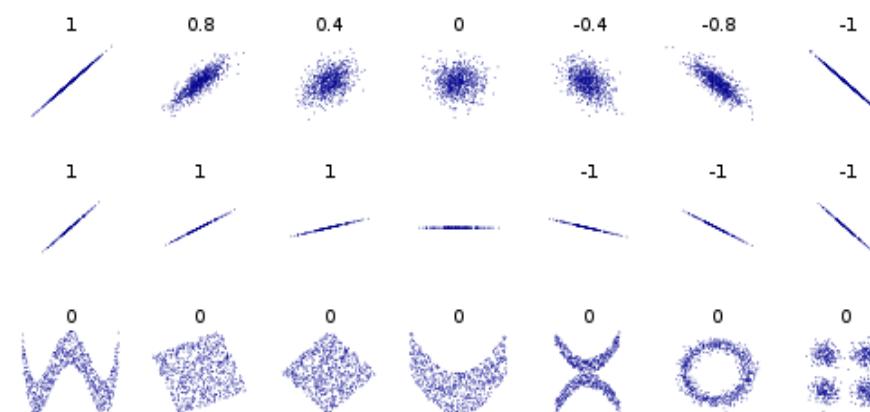
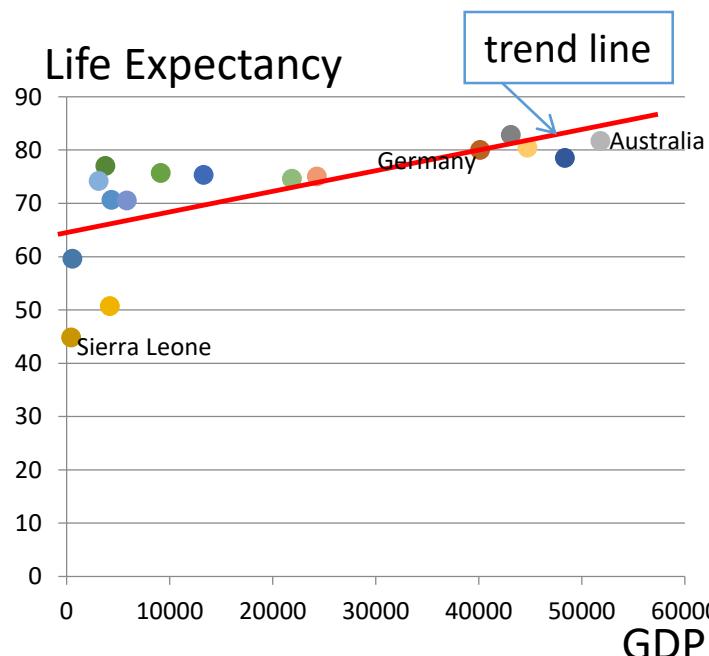
Area Chart

- Area below line chart is filled to indicate volume
- Displays over nominal or interval scale, e.g., time, how values of data items add up to a sum
- Sum represents 100% or absolute values



Scatterplot

- Data is represented by a collection of points
- First and second attribute determine the points' position on the x- and y-coordinate axis, respectively
- Shows correlation of the two attributes

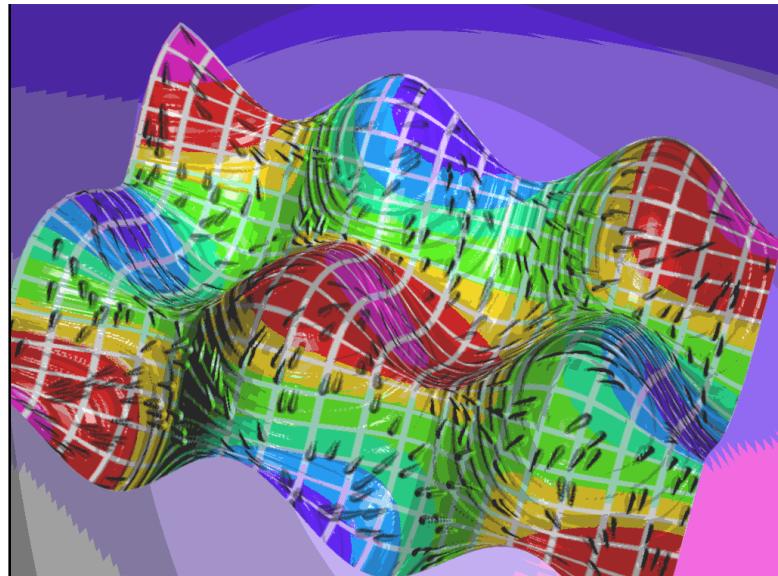


Correlation patterns. Numbers represent the Pearson correlation coefficient.

Colored Height Fields

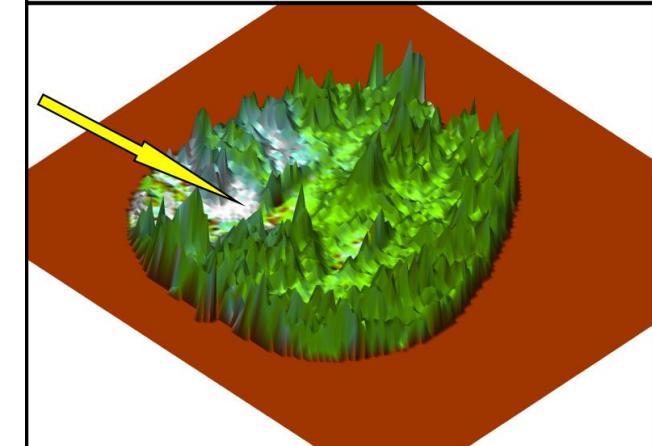
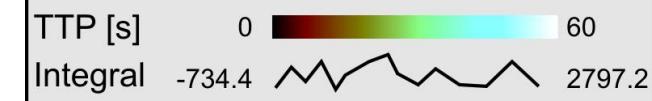
- One attribute is mapped to height and the resulting height field is colored w.r.t. another attribute
- Criticism: occlusion problems, shadowing

Visualization of three scalar attributes (height, color, isocontours) and a vector field (glyphs)



[Crawfis&Allison91]

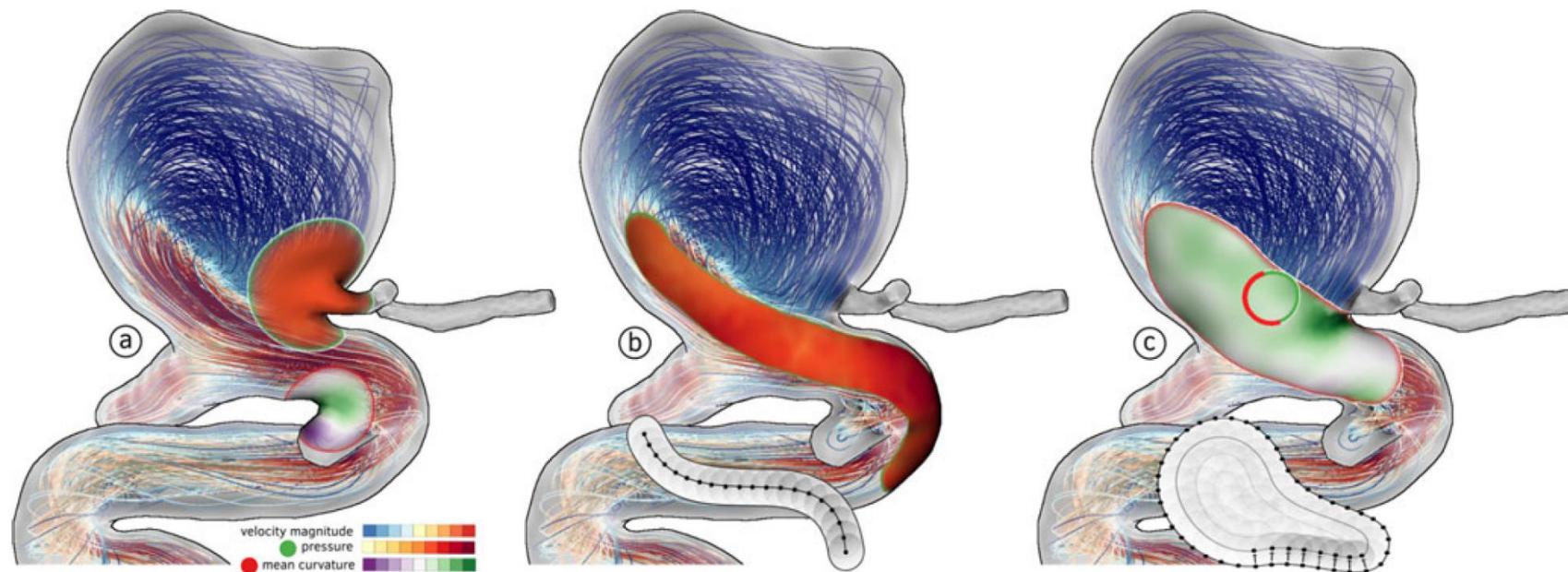
Visualization of two perfusion parameters (height and color)



[Oeltze10]

Flexible / Magic Lenses

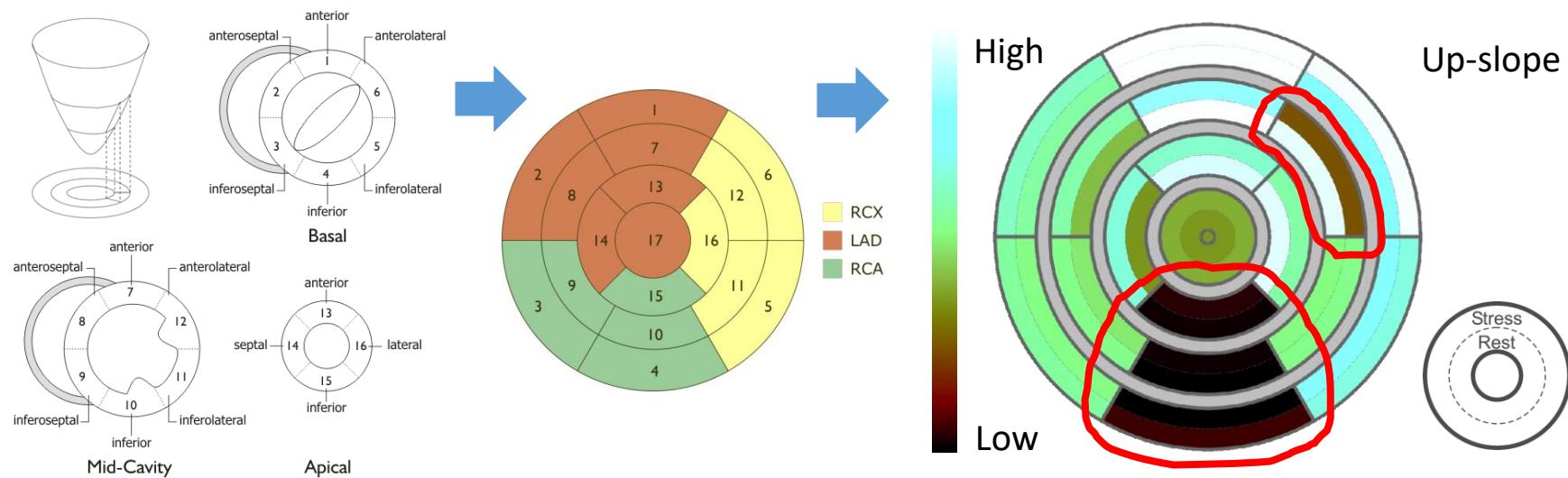
- Extend the function of a lens beyond magnification, e.g., by superimposing additional information
- Example: aneurysm flow data shown in context of anatomical data or/and of another parameter



[Rocha 2019]

Bivariate Bull's Eye Plot (BiBEP) [Oeltze10]

- BEP used for reporting diagnosis results in cardiology
- Very abstract representation of the heart muscle
- Encodes one parameter of the muscle's physiology
- Bivariate plot facilitates rest/stress comparison of the parameter or joint inspection of two parameters

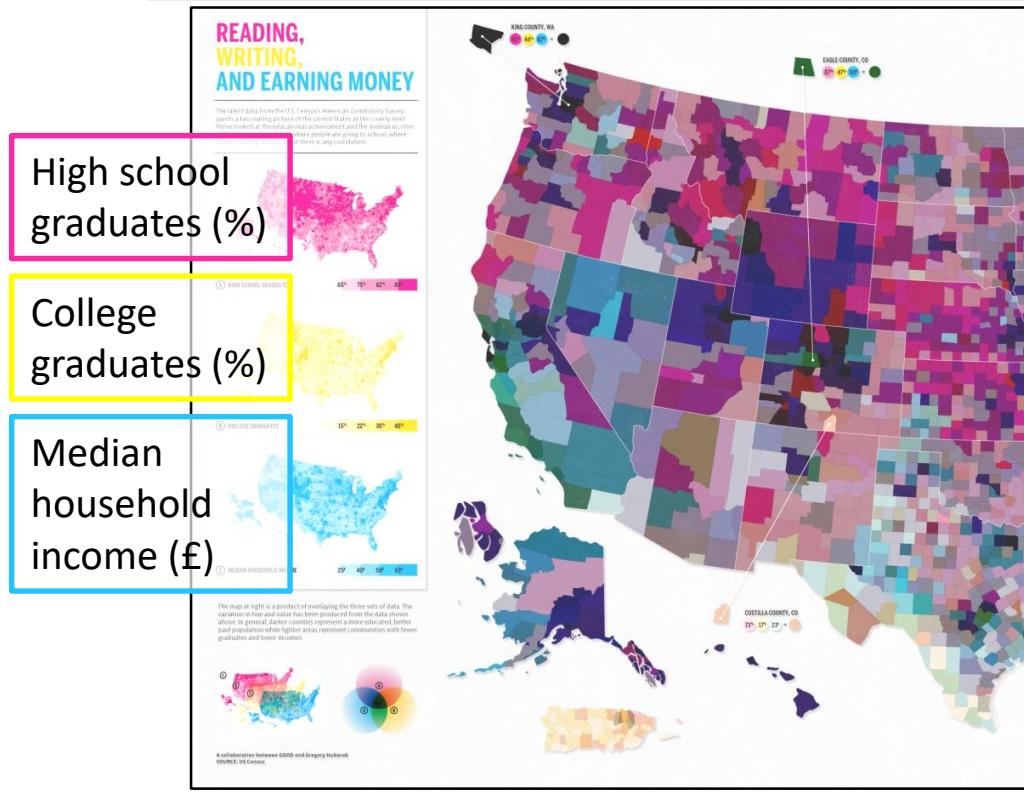


Trivariate Data

Color Channel Encoding

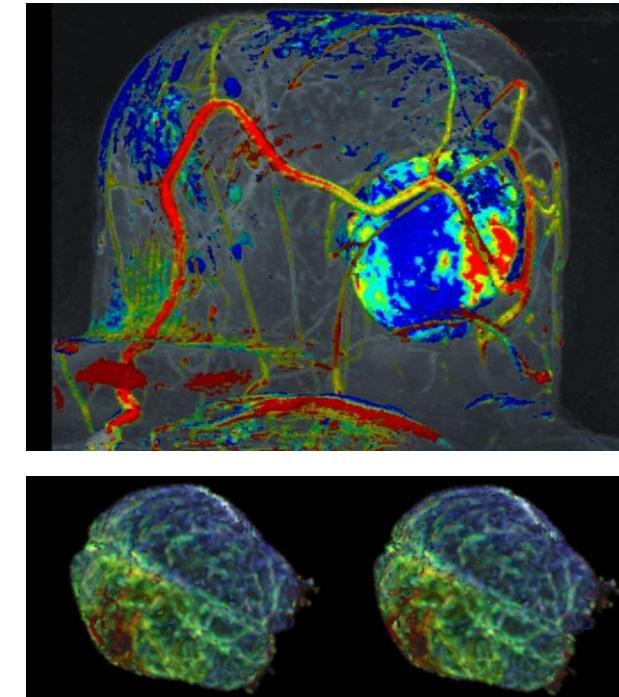
- Max. three attributes are mapped to RGB/HSV channels

Choropleth map: “Are the richest American’s also the best educated?” Can you decode it?



[GOOD&Hubacek]

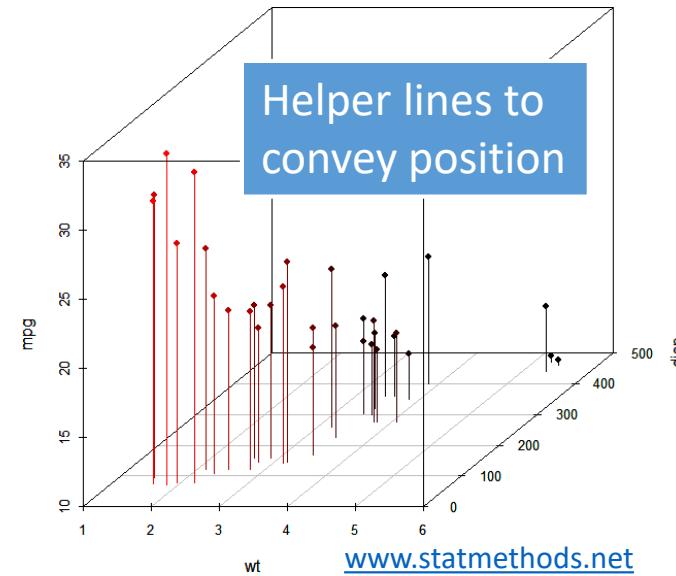
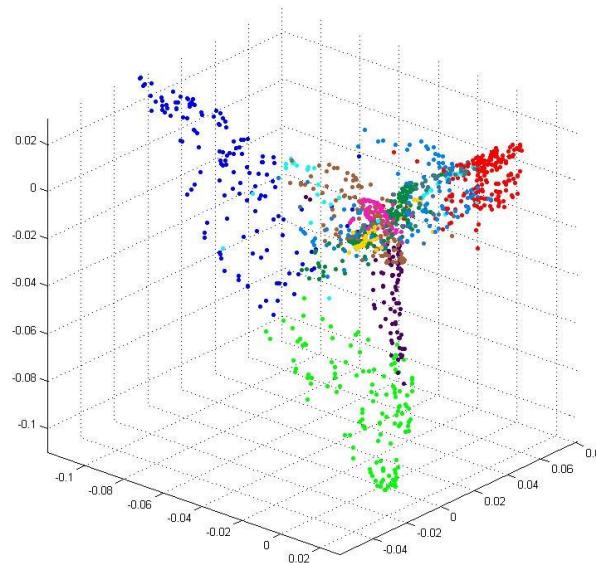
One perfusion parameter mapped to Hue (H) and another one to Value (V)



[Kohle02]

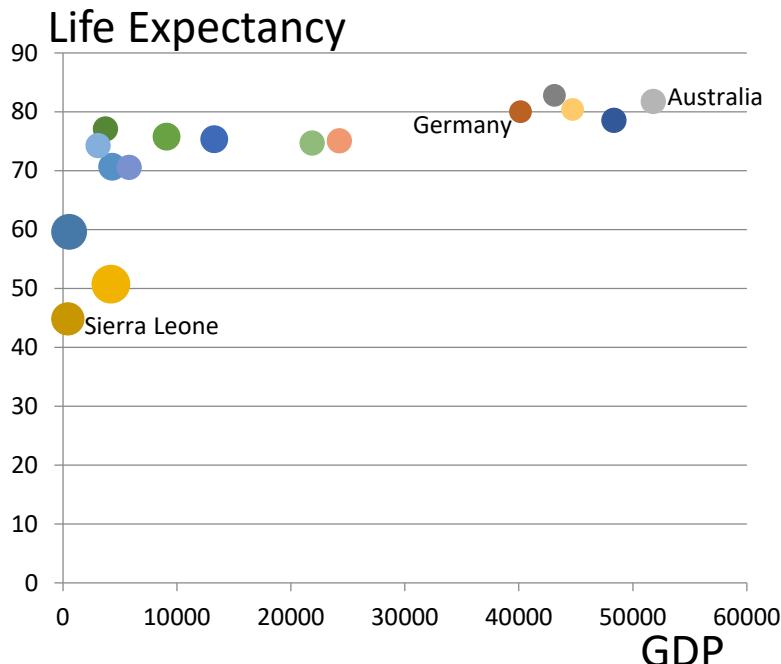
3D Scatterplot and Spinplot

- 3D scatterplot: static 3D extension of 2D version
- Spinplot: interactive 3D scatterplot
- Interaction mitigates occlusion problems
- Structure, e.g., clusters, and point positions still difficult to infer



Bubble Chart

- 2D scatterplot with point size modified according to quantity or another third attribute
- Color may be used to encode a fourth attribute and motion to illustrate change over time (motion chart)



www.gapminder.org/world/

Multivariate Data Visualizations

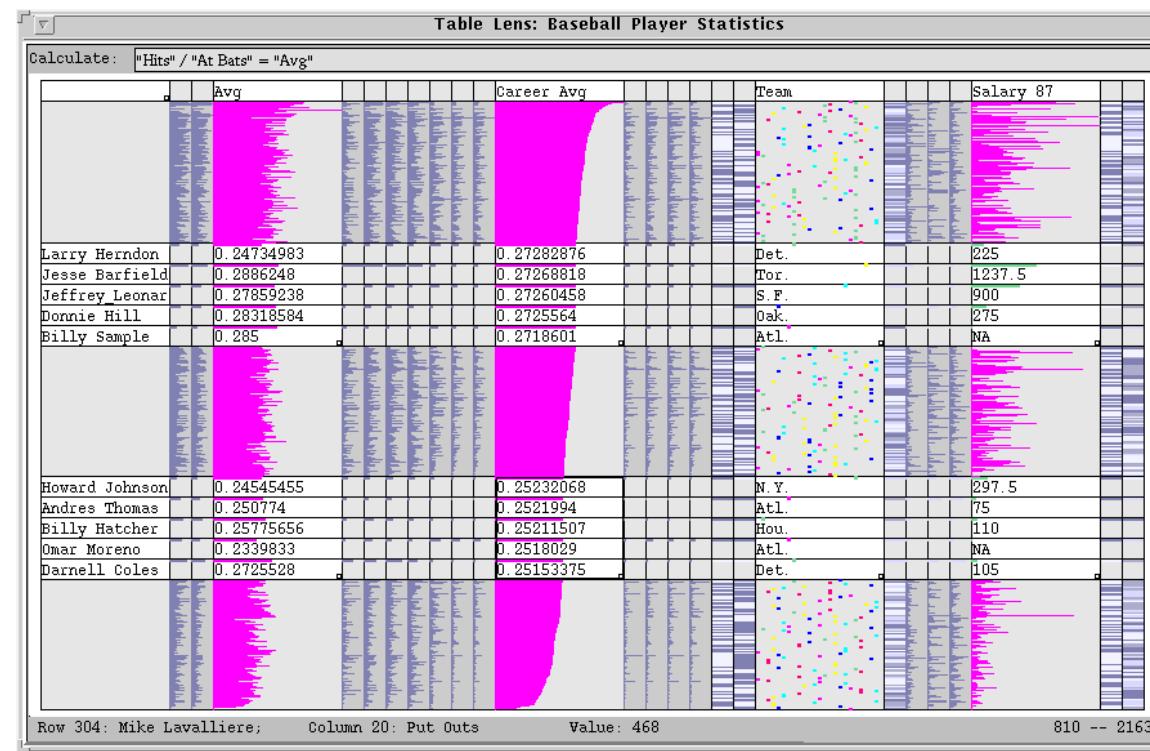
Display the Data Table Itself

- Conditional formatting of individual cells in Excel
- Color scales, data bars, symbols

| Country Name | GDP per Capita (US\$) | Life Expectancy (Years) | Fertility Rate |
|---------------------|-----------------------|-------------------------|----------------|
| Afghanistan | 561.1976175 | 59.6 | 5.659 |
| Germany | 40144.50942 | 80 | 1.39 |
| Japan | 43117.76827 | 82.8 | 1.39 |
| Sierra Leone | 447.7533238 | 44.8 | 4.943 |
| United States | 48357.68451 | 78.5 | 1.931 |
| Albania | 3764.326348 | 77 | 1.741 |
| Algeria | 4349.569325 | 70.6 | 2.817 |
| American Samoa | | | |
| Andorra | | | 1.22 |
| Angola | 4218.649126 | 50.7 | 6.218 |
| Antigua and Barbuda | 13315.24327 | 75.3 | 2.13 |
| Argentina | 9132.957259 | 75.7 | 2.215 |
| Armenia | 3124.788199 | 74.2 | 1.738 |
| Aruba | 24289.14152 | 75 | 1.701 |
| Australia | 51824.79842 | 81.7 | 1.927 |
| Austria | 44723.20394 | 80.4 | 1.44 |
| Azerbaijan | 5843.169753 | 70.5 | 1.92 |
| Bahamas, The | 21881.08394 | 74.6 | 1.901 |

Table Lens [Rao&Card94]

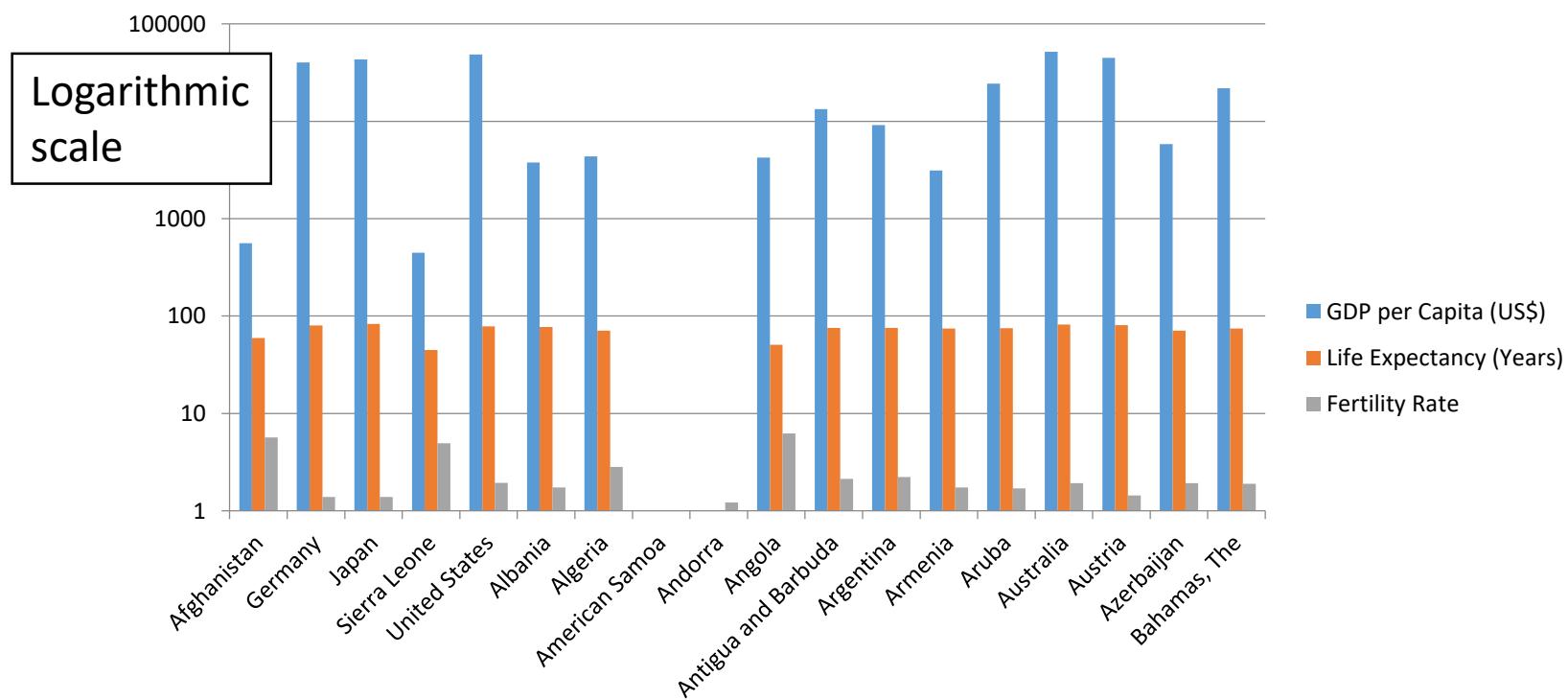
- Focus+Context display of the table
- Context data visualized by graphical representations
- Space advantage of 30 to 100 times



[Rao&Card94]

Multiple Bar Chart

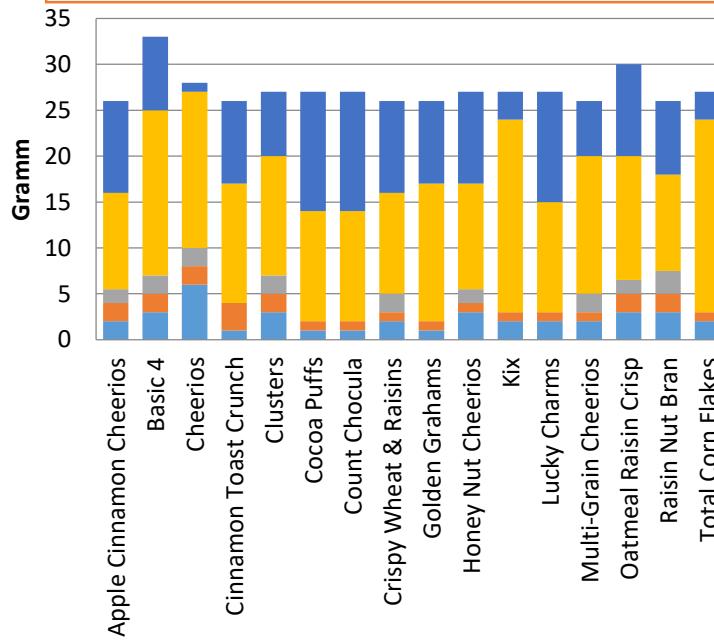
- Groups of bars represent multiple attributes
- Does not scale well but ok for 3-5 attributes
- Problems in case of strongly varying scales



Stacked Charts

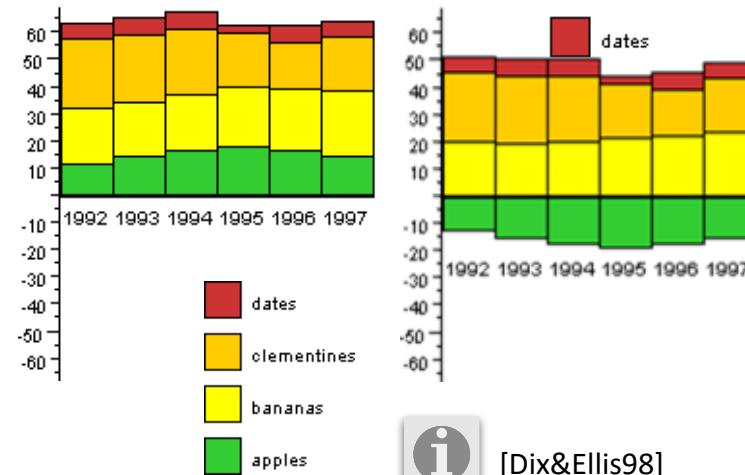
- Display how values add up to a sum (100 percent or absolute value → try yourself at [hci.stanford.edu/...](http://hci.stanford.edu/))
- Feasible if units of values are equal across attributes

Problem: assessing trends for attributes located off the baseline



Solution: interactive stacked bar charts

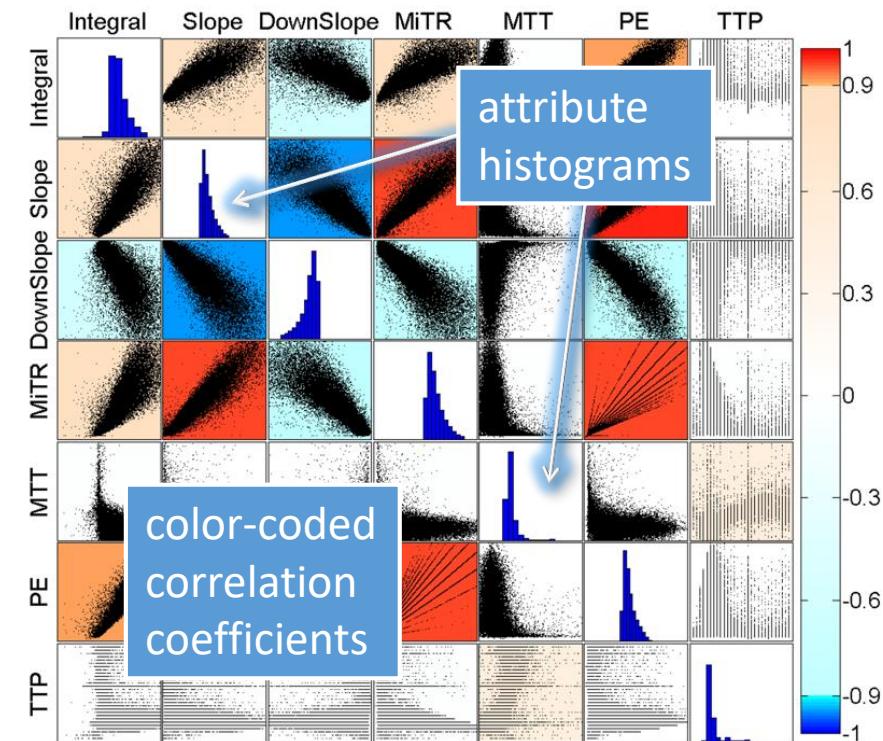
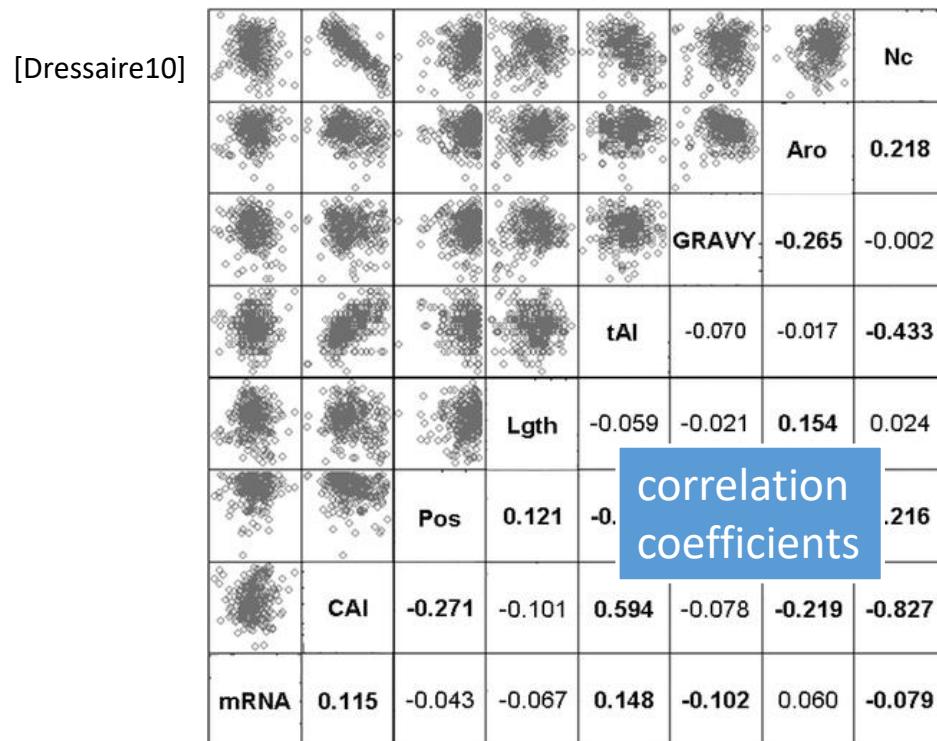
Fruit Sales 1992-1997



[Dix&Ellis98]

Scatterplot Matrix

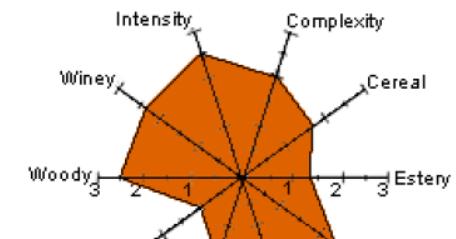
- Arranges scatterplots of all possible attribute pairs
- Main diagonal and one triangular part of symmetric matrix often show additional information



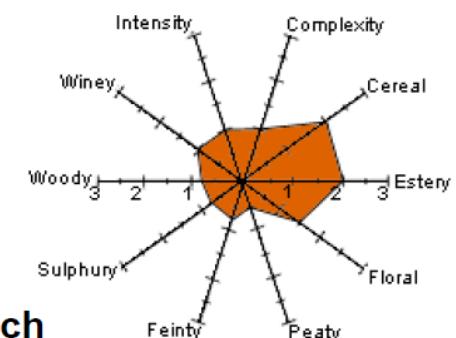
Star Plot aka Radar Chart aka Kiviat Graph

- Star-shaped polygon: vertices are defined by several axes that all have the same origin
- Each attribute is mapped to length of each vertex
- Problems:
 - Scales not well about 6 to 8 attributes and up to 5 data entries
 - Qualitative, and not the quantitative, aspect of attributes is to be explored

Tasting Whisky:
www.scotchwhisky.com



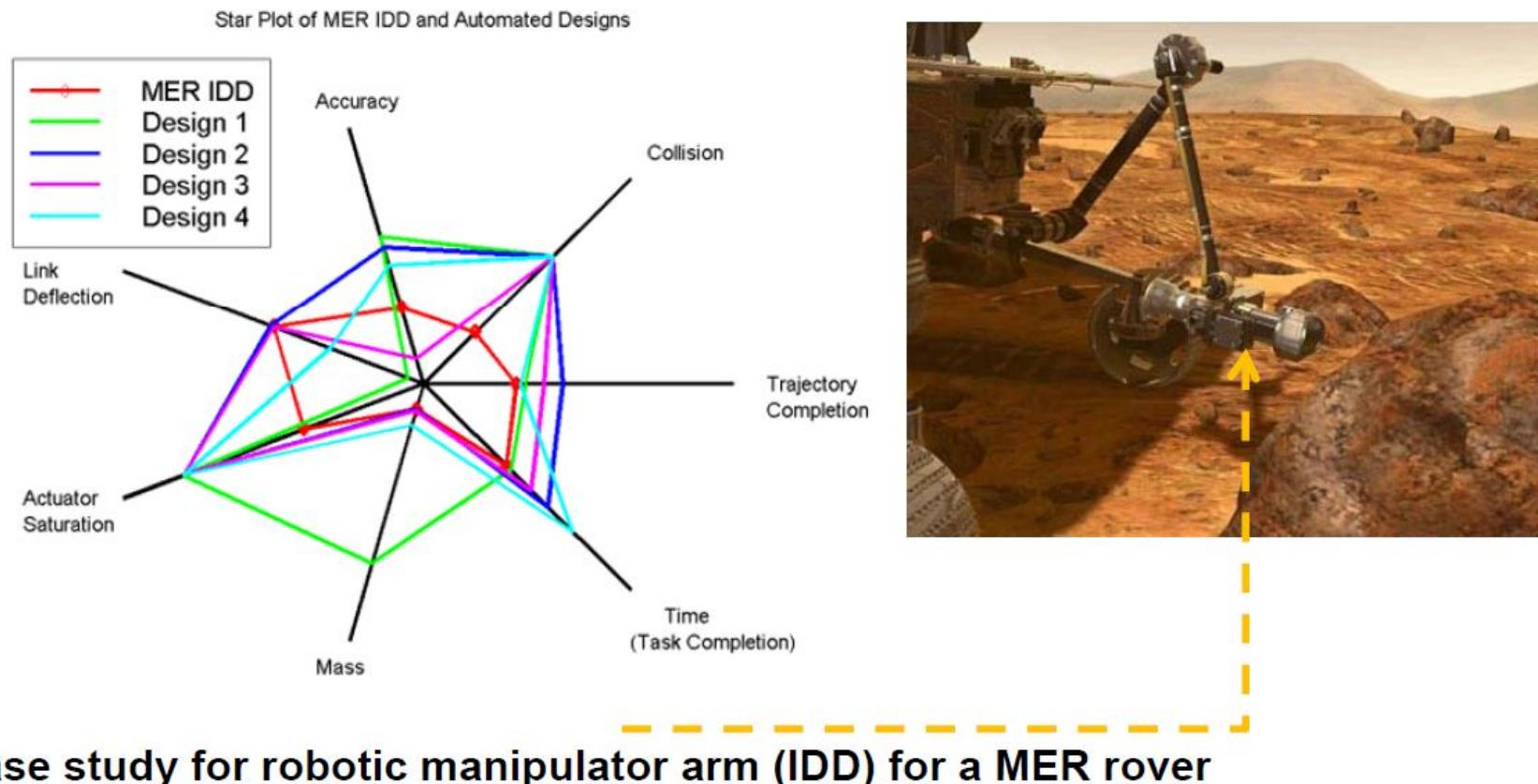
Ardbeg
(17 years old)



Glenfiddich

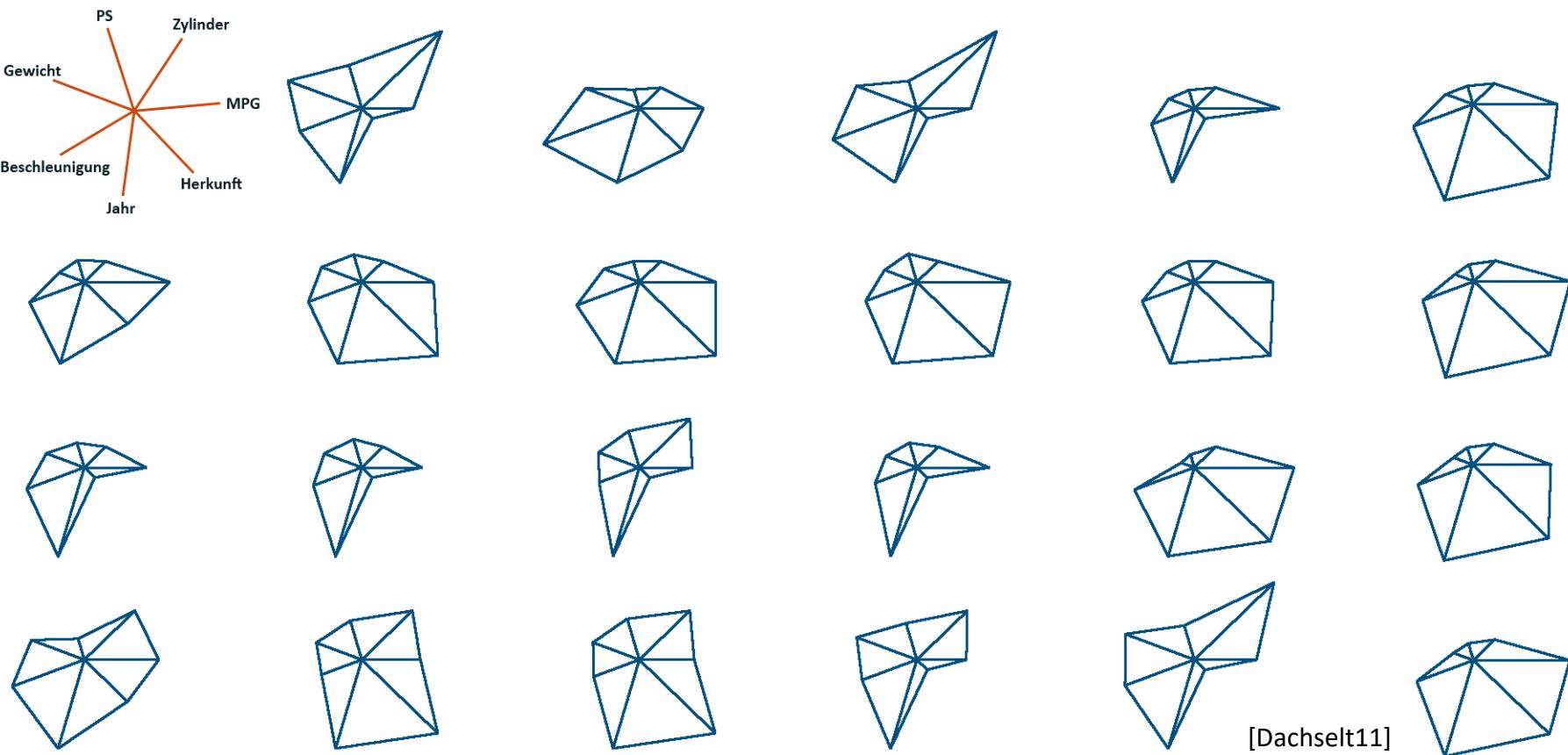
Star Plot aka Radar Chart aka Kiviat Graph

Designing spacecraft and planetary exploration devices
(NASA's Jet Propulsion Lab at Caltech)



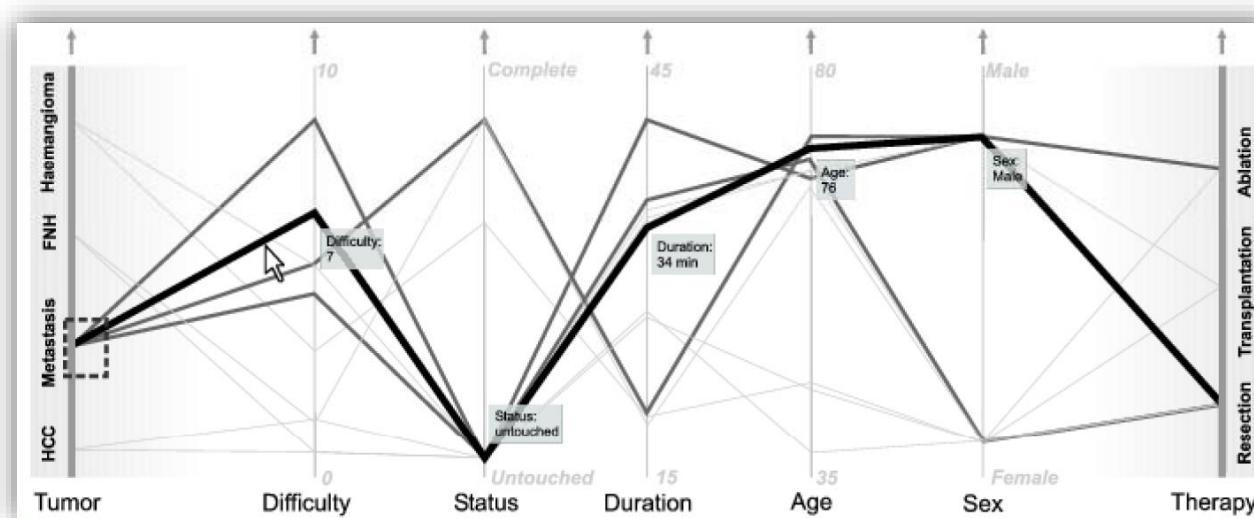
Star Plot aka Radar Chart aka Kiviat Graph

Star plots as small multiples for overview visualization



Parallel Coordinates [Inselberg&Dimsdale90]

- Parallel layout of attribute axes
- Suitable for:
 - Moderate number of variables and hundreds of data entries
 - Variables with possibly different scales and units, and of different data types (nominal/ordinal/quantitative)
- Example: visualization of patient data base



Important Aspects

Arbitrary axes order:

- Only trends between neighboring axes may be observed
→ Algorithms aiming at an optimal ordering exist

Overplotting:

- Visual clutter occurs for huge numbers of data entries
→ Adjustment of alpha transparency
→ Clustering of lines and showing representatives

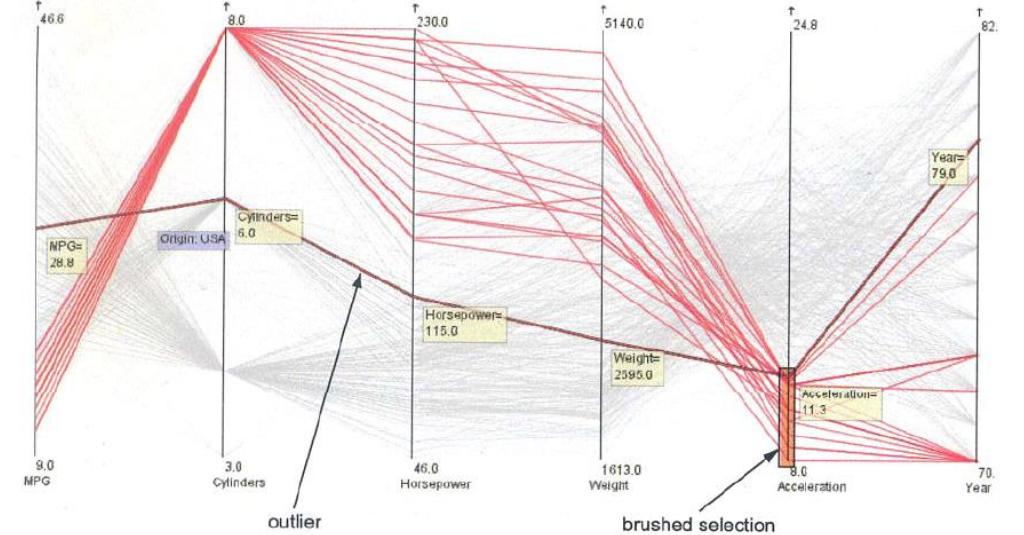
Scaling:

- Heights of data points may not be comparable across axes
→ Data normalization or alignment of axes at mean/median

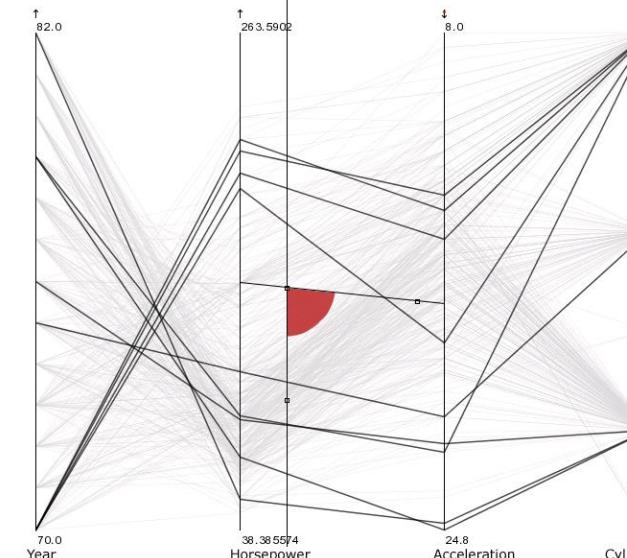
Item Selection in Parallel Coordinates

- Brushing a single or multiple axes
- → try yourself at [blocks.org/...](http://blocks.org/)

- Angular brushing to select lines having a local slope within the specified range
- Example: easier selection of lines sloping downward



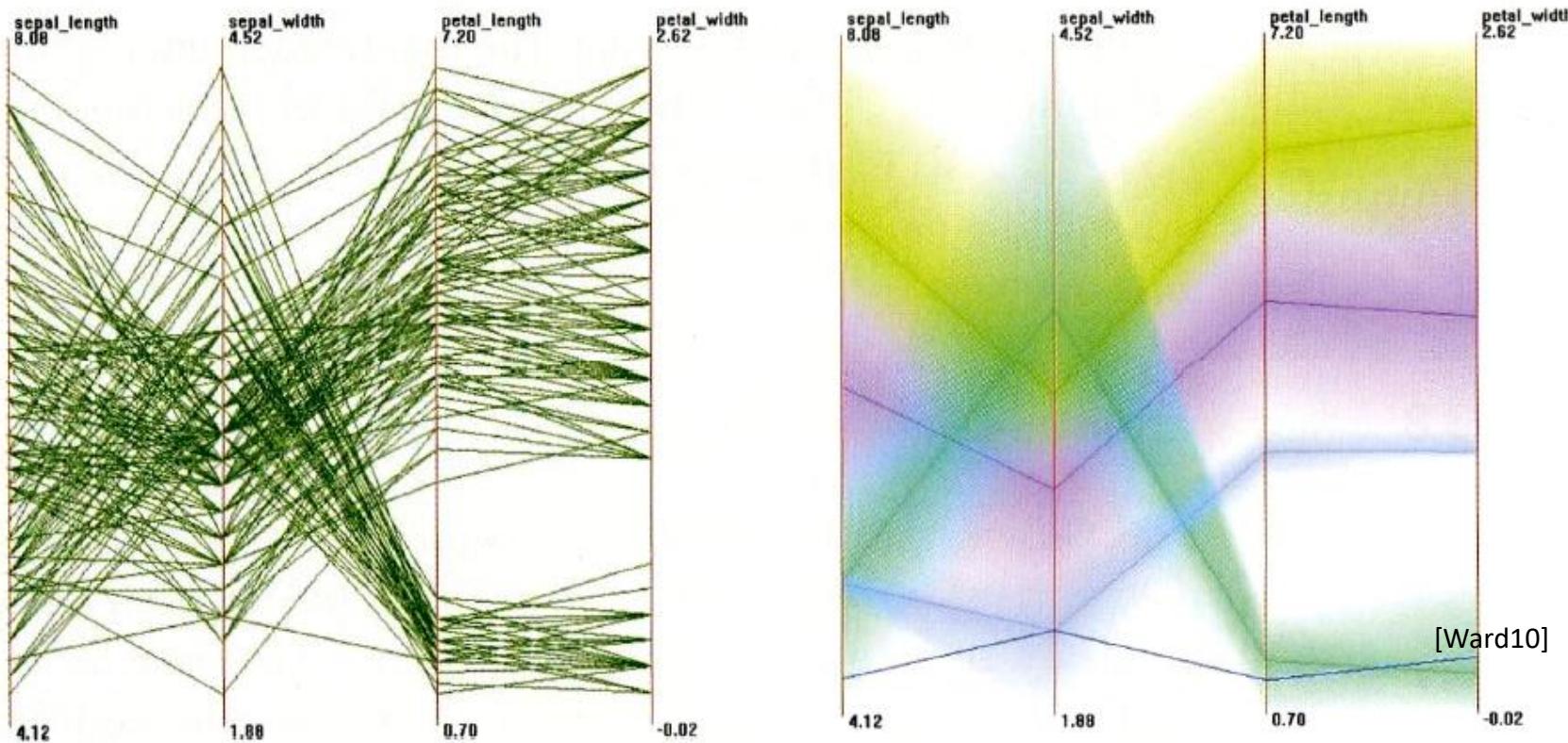
[Telea08]



96

[Hauser02]

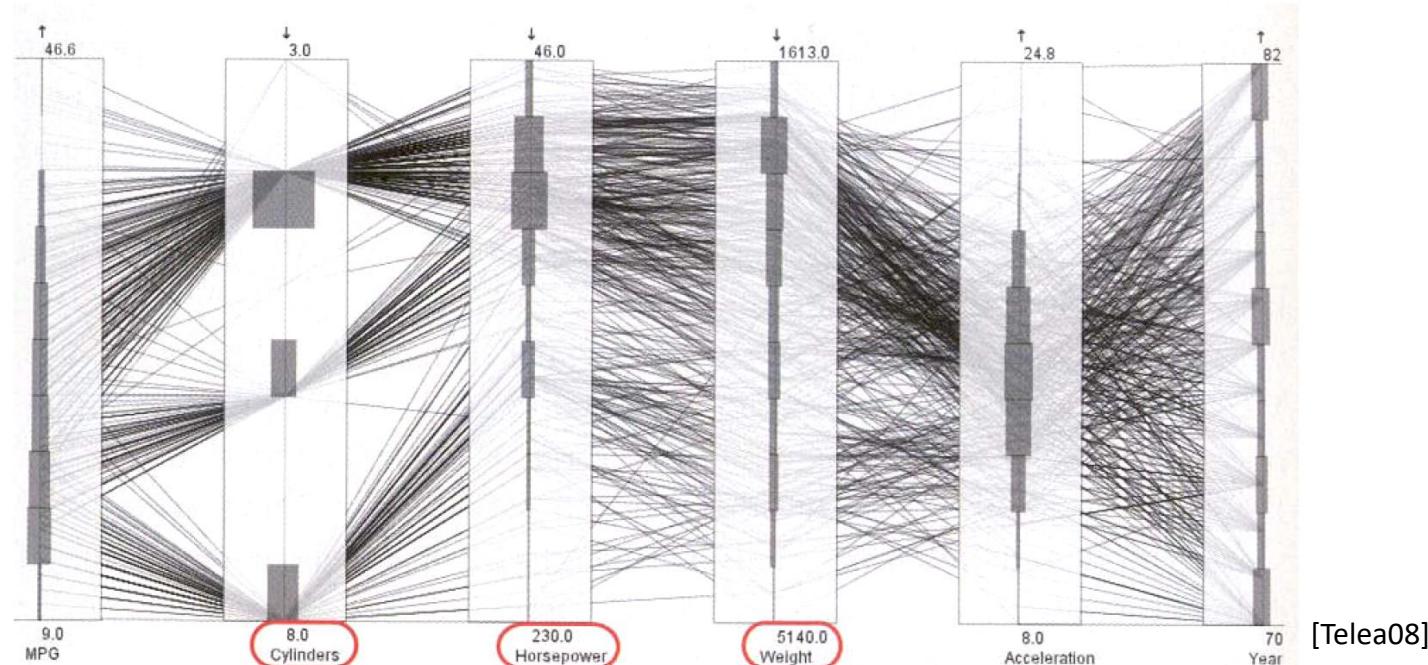
Clustering in Parallel Coordinates



Left: original visualization, Right: lines were clustered. Cluster centers and cluster extents are shown.

Parallel Histograms

- Integrated visualization of attribute value distribution and attribute combinations
- Widths of the dark gray bars represent the amount of data entries for each subrange of attribute values

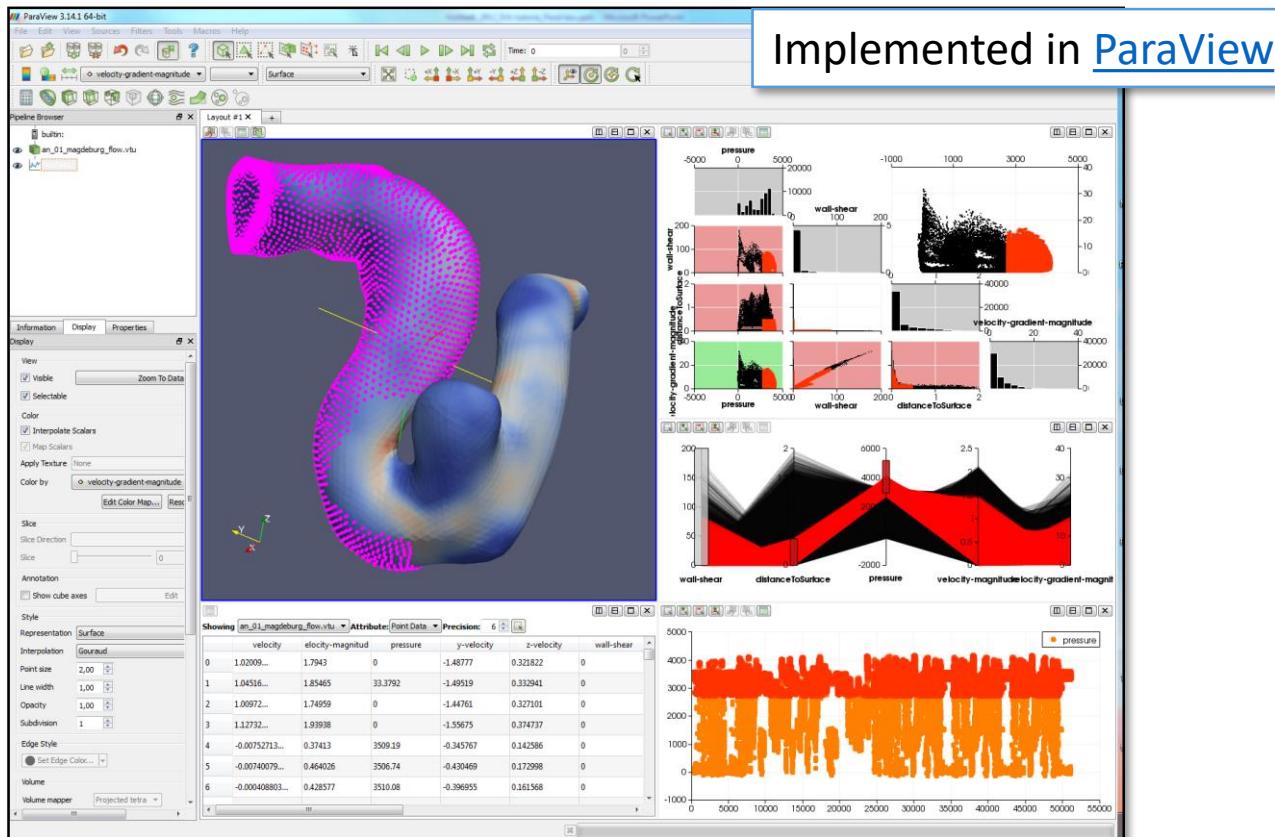


[Telea08]

98

Coordinated Multiple Views

- One dataset but multiple views on the data
- Views are linked and equipped with brushing facilities

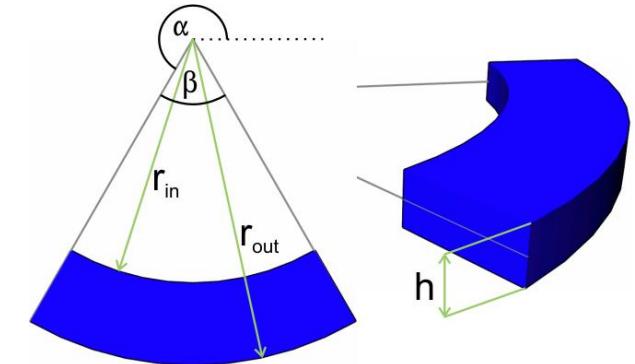


Glyph-Based Techniques

Definitions

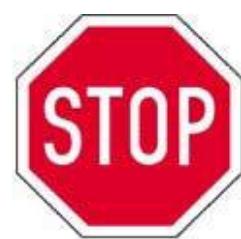
Glyph (synonym: *icon*)

- Graphical object or primitive
- Its properties are modified w.r.t. attribute values



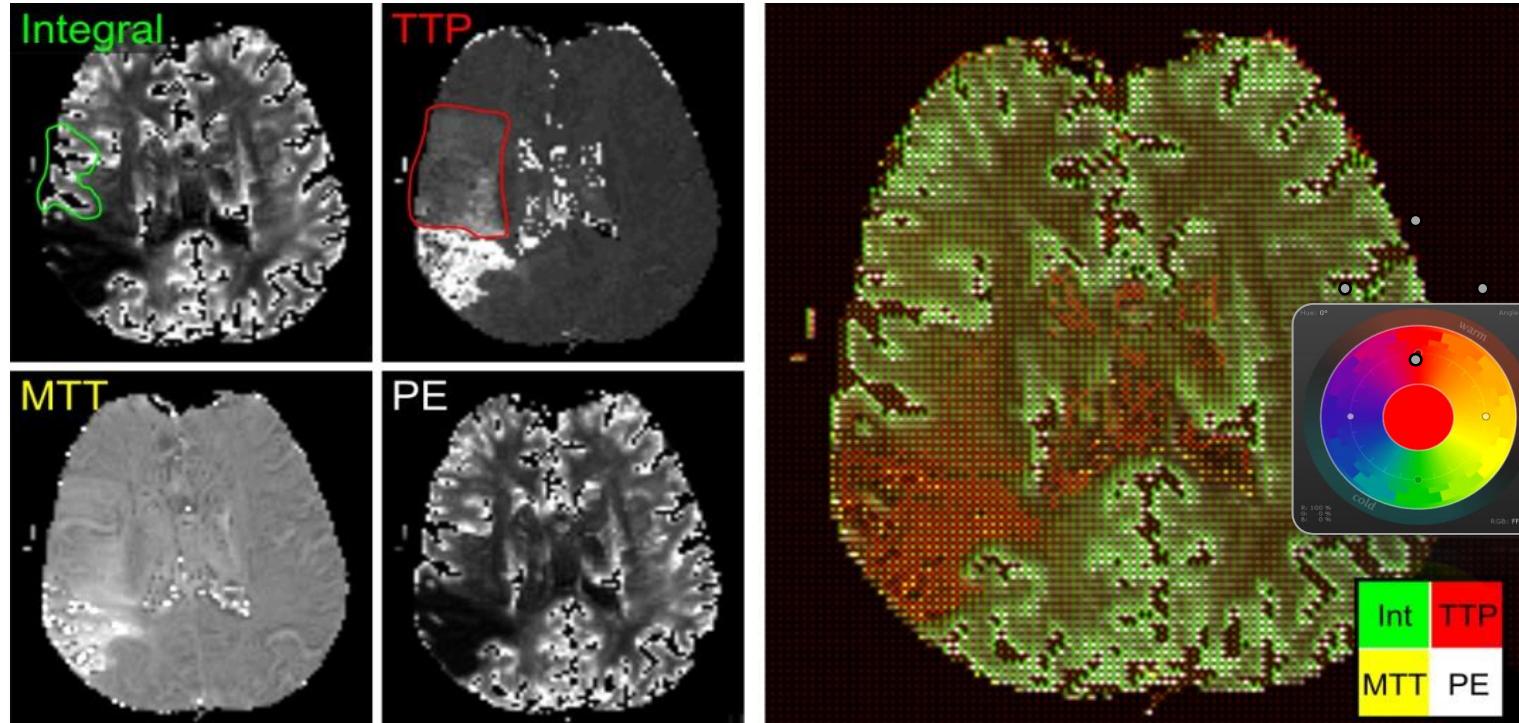
Symbol

- Has meaning as a whole



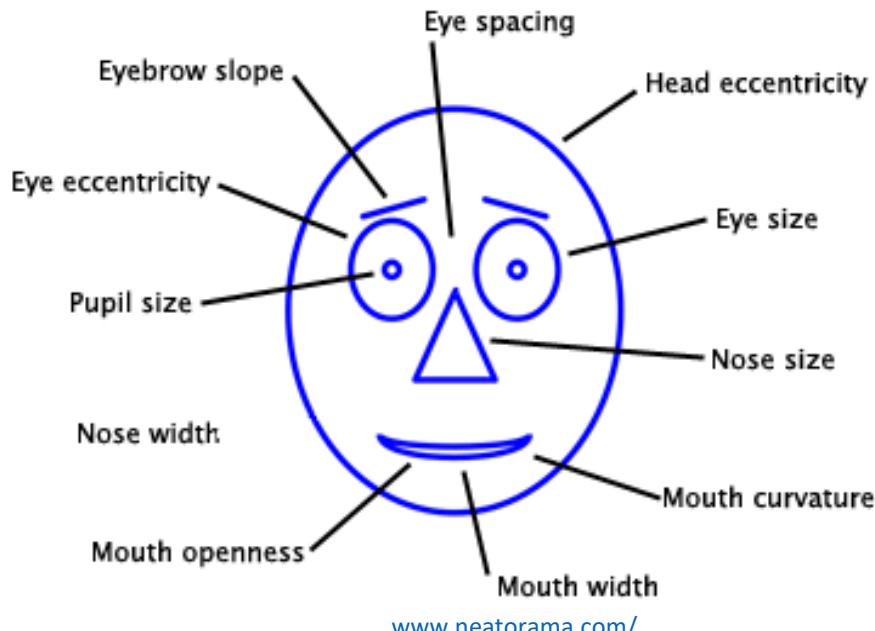
Color Icons

- Each original pixel is replaced by 4 new, quadratically arranged pixels (color icon) [Levkowitz91]
- Each new pixel is assigned a color and an attribute

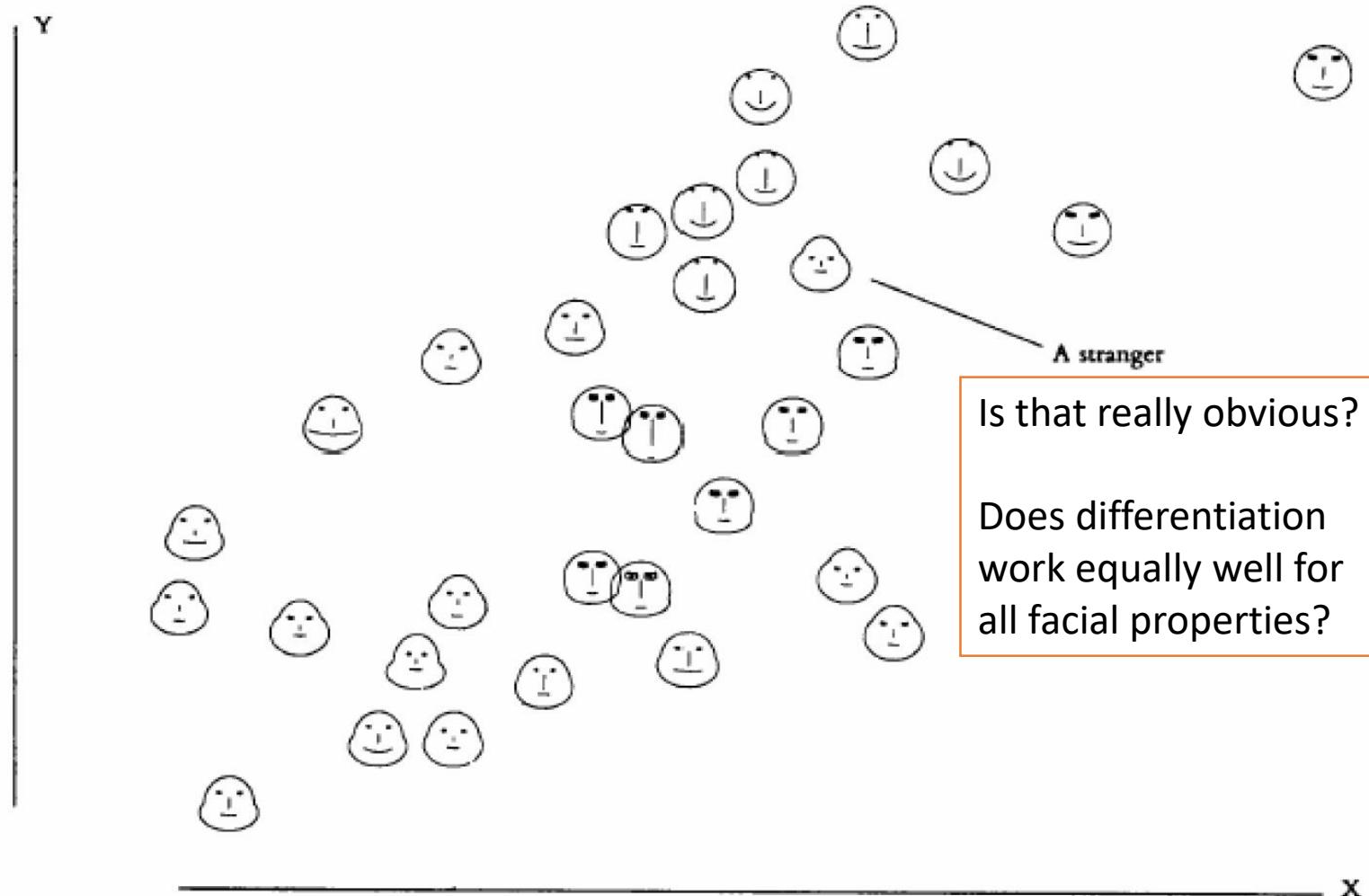


Chernoff Faces [Chernoff73]

- Up to 18 attributes can be mapped to face properties
- **Pro:** humans are familiar with interpreting faces
- **Con:** high variance in interpretation of the same face



Chernoff Faces: Examples



Categorical Data

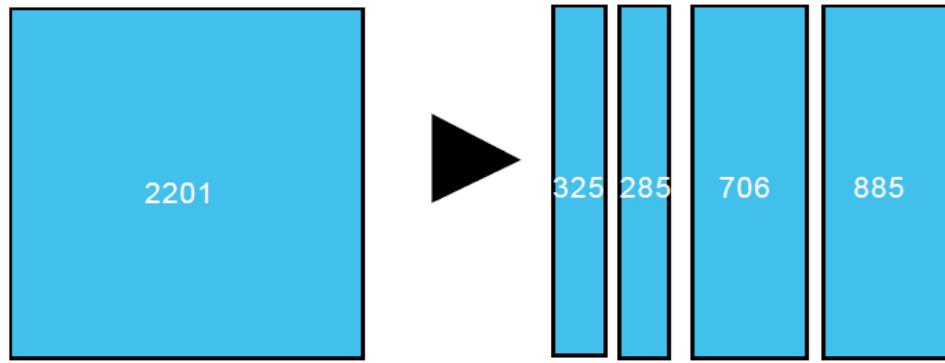
Mosaic Plot

- Iterative subdivision of an area w.r.t. numerical values and categories
- Example: passenger data of the Titanic (from [Spence14])

| Survived | Age | Gender | Class | | | |
|----------|-------|--------|-------|-----|-----|------|
| | | | 1st | 2nd | 3rd | Crew |
| No | Adult | Male | 118 | 154 | 387 | 670 |
| Yes | | | 57 | 14 | 75 | 192 |
| No | Child | | 0 | 0 | 35 | 0 |
| Yes | | | 5 | 11 | 13 | 0 |
| No | Adult | Female | 4 | 13 | 89 | 3 |
| Yes | | | 140 | 80 | 76 | 20 |
| No | Child | | 0 | 0 | 17 | 0 |
| Yes | | | 1 | 13 | 14 | 0 |

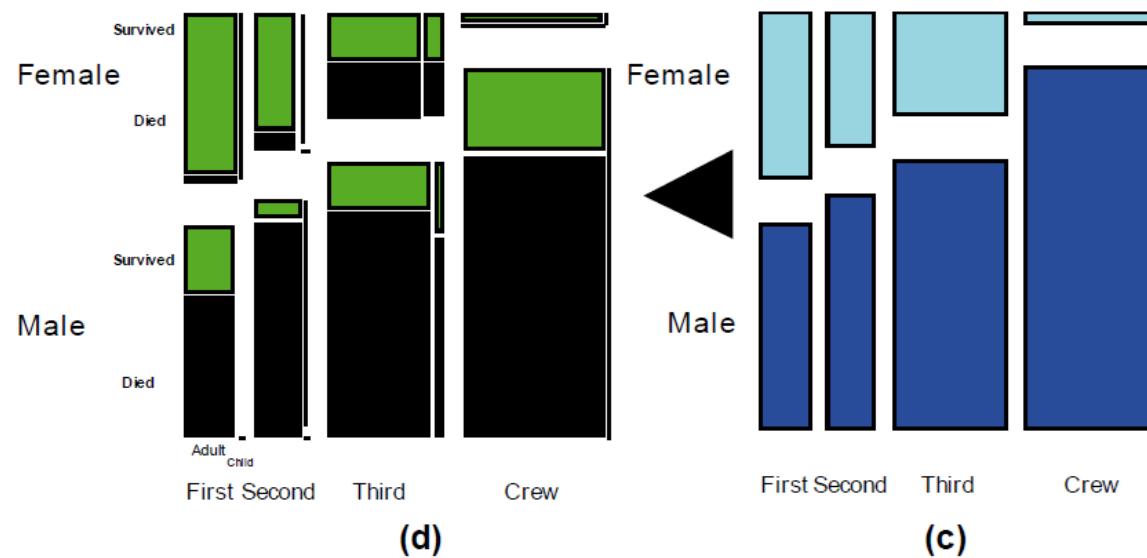
Mosaic Plot

Iterative
subdivision:



(a)

(b)



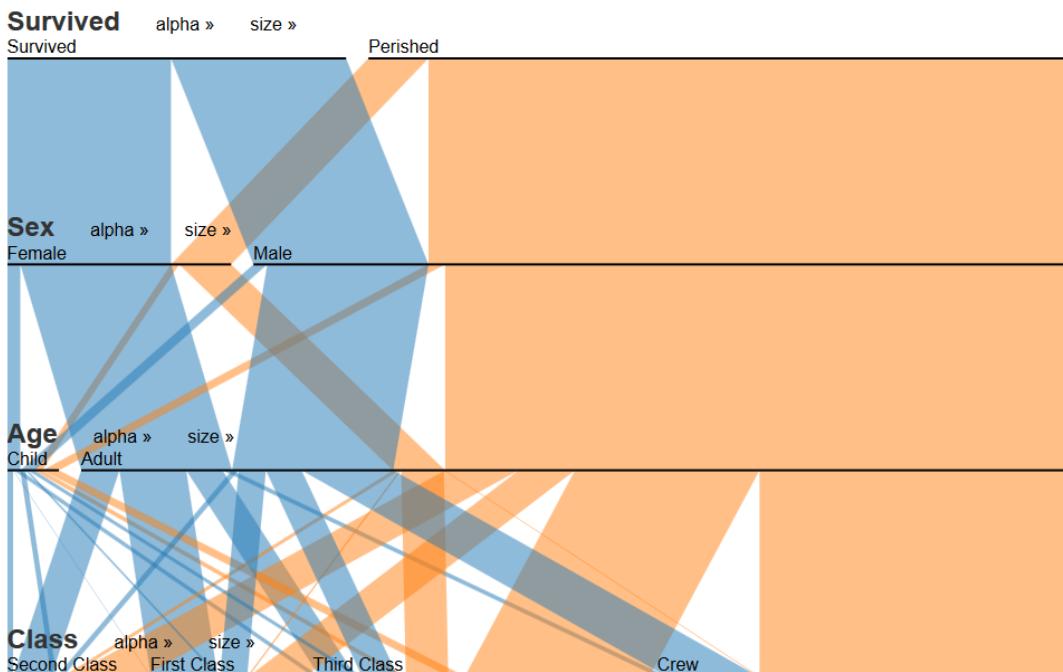
(d)

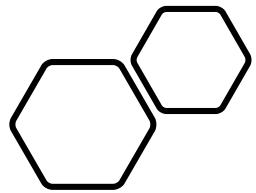
(c)

[Spence14]

Parallel Sets [Bendix05]

- Horizontal bar is drawn for each attribute and each of its possible categories
- Bar width encodes category's number of data entries
- Example: again, the Titanic data





Questions???