Visualisation in Data Science

VU Visual Data Science
Johanna Schmidt

WS 2018/19
Visualisation in Data Science

• Visualisation techniques to support data science tasks
• Usage of charts
• Visualisation principles
Visualisation in Data Science
Visualisation in Data Science
Visualisation in Data Science

• Development of new visualisation techniques
• Guidelines for usage of techniques / design elements
• Solutions for overview generation / big data handling
• Solutions for interaction (filtering, details selection)
Visualisation in Data Science

• New research field, mostly related to information visualisation
• Further development needed for visualisation techniques specifically for data science tasks
• Visualisation techniques to be used for data science (1st lecture),
• Lecture today:
  • Usage of charts
  • Principles from human perception / visualisation principles which are important when designing visual representations
  • Challenges in visual data science
Usage of Charts
Usage of Charts

• Guidelines available from many sources
Usage of Charts

• Usage by task
  • Distribution
    • How values are distributed within value range
  • Relationship
    • How do elements relate to each other
  • Comparison
    • Comparison of different values
  • Composition
    • To show the composition of an element
Usage of Charts

• **Distribution**
  - One variable: Histogram (column- or line-based)
  - Two variables: Scatter plot
  - Three variables: 3D area chart
Usage of Charts

• Relationship
  • Scatter plot
  • Bubble chart
Principles of Visual Encodings

• Gestalt Laws
  • Define how humans visually interpret structural information
  • We typically try to detect patterns in what we see
  • Which objects are perceived as a group? -> Principles of
    • Proximity
    • Similarity
    • Connectedness
    • Good continuation
    • Common fate
    • Symmetry
Principles of Visual Encodings

• Gestalt Laws
  • Define how humans visually interpret structural information
  • We typically try to detect patterns in what we see
  • Which objects are perceived as a group? -> Principles of
    • Proximity
    • Similarity
    • Connectedness
    • Good continuation
    • Common fate
    • Symmetry
Proximity

- Things that are near to each other seem to be grouped together
Usage of Charts

• Relationship

• Parallel coordinates
Usage of Charts

• Comparison

Among items

Over time
Usage of Charts

- Comparison

Among items

Over time
Comparisons

- **Length** (bar chart) is better for comparisons than **area** (pie chart)
Usage of Charts

- Comparison

Among items:
- Bar Chart: Many Items, Few Categories
- Column Chart: Few Items

Over time:
- Circular Area Chart: Cyclical Data
- Line Chart: Non-Cyclical Data
- Column Chart: Single or Few Categories
- Line Chart: Many Categories

Many Periods to Few Periods
Usage of Charts

- **Comparison**

  - **Bar Chart**
    - Many Items
    - Few Categories

  - **Column Chart**
    - Few Items
    - Few Categories

  - **Circular Area Chart**
    - Cyclical Data
    - Many Periods

  - **Line Chart**
    - Non-Cyclical Data
    - Few Periods

  - **Column Chart**
    - Single or Few Categories

  - **Line Chart**
    - Many Categories

Among items

Over time
Usage of Charts

• Comparison
Usage of Charts

• Comparison
Usage of Charts

• **Composition**
  
  • Share of total: Pie chart
  
  • Components: Stacked bar chart
  
  • Changes: Stacked bar chart / Stacked area chart
Usage of Charts

• Composition
  • Stream graphs
Usage of Charts

• Consider data and task
Usage of Charts

• Consider **data and task**
• **2D** usually better than 3D
Usage of Charts

• Consider **data and task**
• **2D** usually better than **3D**
• **Axes scales**
Usage of Charts

- Consider **data and task**
- **2D** usually better than **3D**
- **Axes scales**
Usage of Charts

- Consider **data and task**
- **2D** usually better than **3D**
- **Axes scales**
Usage of Charts

• Consider **data and task**
• **2D** usually better than 3D
• **Axes scales**
• Carefully choose **chart elements**
Visualisation Principles

• Visual mappings
• Principles of visual encodings
• Usage of color
• Usage of shapes
• Consistency
• Interaction
Visualisation Principles

• Visual mappings
• Principles of visual encodings
• Usage of color
• Usage of shapes
• Consistency
• Interaction
Visual Mappings

- Position
- Length
- Angle
- Slope
- Area
- Volume
- Density
- Colour
Visual Mappings

- position
- length
- angle
- slope
- area
- volume
- density
- colour

More accurate

Less accurate
Preattentive Visual Features

• Rapid identification of visual features (low-level human visual system)
Preattentive Visual Features

• Rapid identification of visual features (low-level human visual system)

- closure
- curvature
- length
Preattentive Visual Features

- Target detection
- Counting
- Boundary detection
Preattentive Visual Features

- Target detection
- Counting
- Boundary detection

[2]
Visualisation principles

• Visual mappings
• **Principles of visual encodings**
• Usage of color
• Usage of shapes
• Consistency
• Interaction
Principles of Visual Encodings

• Gestalt Laws
  • Define how humans visually interpret structural information
  • We typically try to detect patterns in what we see
  • Which objects are perceived as a group? -> Principles of
    • Proximity
    • Similarity
    • Connectedness
    • Good continuation
    • Common fate
    • Symmetry
Proximity

• Things that are **near to each other** seem to be grouped together
Similarity

• **Similar objects** seem to be grouped together
Connectedness

- Objects that are **physically connected** seem to be grouped together.
Good Continuation

- Connected lines seem to follow the smoothest path
Symmetry

• **Symmetrical objects** seem to be seen as a whole
Visualisation principles

• Visual mappings
• Principles of visual encodings
• **Usage of color**
• Usage of shapes
• Consistency
• Interaction
Color

• Strong **visual channel** to transport information
Color

- Strong **visual channel** to transport information
- **Several aspects** have to be considered
  - Similar/dissimilar colors for similar/dissimilar attributes
  - Psychological effects (e.g., red vs. blue for alerts)
  - Consider color blindness
Color

- Do not use **too many colors** in one chart
Color

• Consider **relation and size** of elements
Color

• Do not use gradient colours for **categorical data**
Color

- Use **intuitive** colours
Color

• Carefully design color maps
Color

• **Suggestions** for color maps
  - Colorbrewer ([http://colorbrewer2.org](http://colorbrewer2.org))
Color

- **Rainbow color maps**
  - Very prominent, but **should not be used** in data visualisation
  - Produce hard borders and may mislead interpretation
Color

• Rainbow color maps
Color

- **Rainbow color maps**
  - Very prominent, but **should not be used** in data visualisation
  - Produce hard borders and may mislead interpretation
  - Many other (perceptually uniform) color maps available
Colour

• **Fewer colours** better than many
• **Contrast vs. size/relati**on of elements
• Consider **intuitive** colours
• Consider **alternative color maps** to the rainbow color map
• **Colour blindness**
Visualisation principles

• Visual mappings
• Principles of visual encodings
• Usage of color
• **Usage of shapes**
• Consistency
• Interaction
Shapes

- Mapping **geometric elements** to data attributes
  - Color/position/size for additional information
Shapes

- Mapping **geometric elements** to data attributes
  - Color/position/size for additional information

- When using shapes to encode quantitative attributes, it is important to remember that the **area** of the shape has to be adjusted, not the **diameter** (size)
Shapes

• Mapping geometric features to data attributes

• More complex geometric objects allow to map multivariate attributes
Complex shapes
Shapes

- **Fewer** different shapes is better than many
- **Shape parameters** (size, angle, aspect ratio) can be mapped to data attributes for multivariate data
- Always adjust **area**, not diameter/size when visualising quantitative attributes
Visualisation principles

• Visual mappings
• Principles of visual encodings
• Usage of color
• Usage of shapes
• **Consistency**
• Interaction
Consistency

• Use the **same visual encoding** for similar values
Consistency

- Use the **same visual encoding** for similar values
Consistency

• Do not use the same encoding for different values
Visualisation Principles

• Visual mappings
• Principles of visual encodings
• Usage of color
• Usage of shapes
• Consistency
• Interaction
Interaction

*Overview first, details on demand*
[Shneiderman, 1986]

- **Overview** needed, before being able to explore the data
- Exploration can then lead to **details**
Interaction

- **Details** by interacting with the data
  - Selection techniques
  - Filtering algorithms
  - Data processing

- **Focus+context**
  - While exploring details, context should be preserved

- **Linking and brushing**
  - Apply filtering to all views
Visualisation Challenges in Data Science

- Large data sets
- Complex data
- Web-based vs. desktop applications
- Exploration vs. presentation
Visualisation Challenges in Data Science

- Large data sets
- Complex data
- Web-based vs. desktop applications
- Exploration vs. presentation
Visualisation Challenges in Data Science

- Large data sets
- Complex data
- Web-based vs. desktop applications
- Exploration vs. presentation
Visualisation Challenges in Data Science

• Large data sets
• Complex data
• Web-based vs. desktop applications
• Exploration vs. presentation
Visualisation in Data Science

• Visualisation techniques to be used for data science (1st lecture),

• Lecture today:
  • Usage of charts
  • Principles from human perception / visualisation principles which are important when designing visual representations
  • Challenges in visual data science

• Next lectures:
  • Specific visualisation approaches
  • Applications / libraries for visual data science
References

[8] https://i.stack.imgur.com/lp4tr.png
[10] https://bl.ocks.org/mbostock/raw/4060954/5d12f94f35e5b86d2074ef0dbb478efe722ef54c/preview.jpg
[11] https://pbs.twimg.com/media/Cl4-PbTXAA9Wo.jpg
References

[27] https://docs.microsoft.com/de-de/power-bi/consumer/media/end-user-dashboards/power-bi-dashboard2.png