Data Visualization

(Ch 1,2,3)

Darren Yu

What is Data Visualization

Computer based visualization systems provide visual representation of data sets designed to help people carry out tasks more out tasks more effectively.

When?

- Sometimes there are circumstances when human capabilities need to be augmented or supported
- This is excludes trusted and fully automated processes

Why Use External Representation? Ch.1

Why have a human in the decision making loop?

Visualization is suitable when there is a need to augment human capabilities rather than replacing them as a whole with computation decision-making methods.

Why have a human in the decision making loop?

- Provides visual means to analyze patterns, information, and validity of statistical models
- Establishes a logical structure on design space for:
  - Decision making
  - Designing new systems
  - Comparing strong and weak task/data combinations
- Presenting findings and results (e.g. hidden details and unexpected patterns)
- Data verification
- Refining and providings insight towards building better systems

Why depend on vision?

Computer based visualization systems provide visual representation of data sets designed to help people carry out tasks more out tasks more effectively.

Vision versus the other senses:

- Human visual system is high-bandwidth channel to brain
  - Subjective experience of seeing everything simultaneously
  - Visualization allows our brains to process the visual information as background tasks, which allows us to more intuitively digest the given information
- Other senses: hearing, touch, and smell/taste
  - Actualized in low bandwidth processing or can only be applied in extremely niche cases

What are the limits of data visualization

- Computation time and system memory
- Display limits
  - Pixels may be too condensed and may show a poor representation
  - Information density: ratio of space for information over unused space
  - Information can be too cluttered and dense or too sparse
- Human limits
  - Time
  - Memory
  - Attention

Why analyze visual data?

Data Types

Items

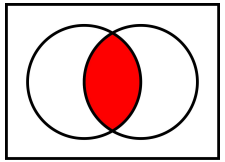
- Individual identities
- E.g. a row (labels) of a data table

Attributes

- Property that is typically measured/observed
- E.g. the column of a data table

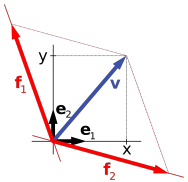
Links

- Relationships between two items (instances)



Position

- Locations in 2D/3D space
- E.g. spatial data



Grid

- Strategy for sampling continuous data

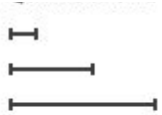
Attribute Types

Categorical

Can compare equivalence

Quantitative

Magnitude and quantity matters. Can be used in arithmetic calculations



Ordinal

Can be compared using size (less than or greater than)

What is Task Abstraction

Task Abstraction

"Task abstraction is the transformation of a task description from domain-specific contexts into an abstract form."

Target

Certain aspect of the data that the user is interested in.

-Features-

-Outliers-

-Trends-

Methods of Abstraction

Identifying the dataset type(s) and attribute types

Identifying the cardinality at each level

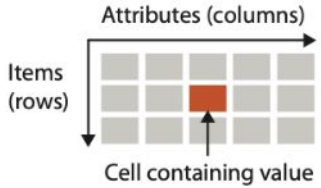
Whether to transform data

Dataset Types

Tables

- Items
- attributes

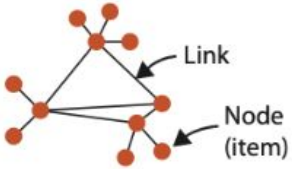
➔ Tables



Networks

- Items
- Attributes
- Links

➔ Networks



Spatial

- Fields (continuous, vectors, tensors)
- Positions
- attributes

Geometry

- Items
- Attributes

By Abstracting data in generic form:

We can compare values between two groups

Types of Actions

Action

Defining Goals:

Analyze

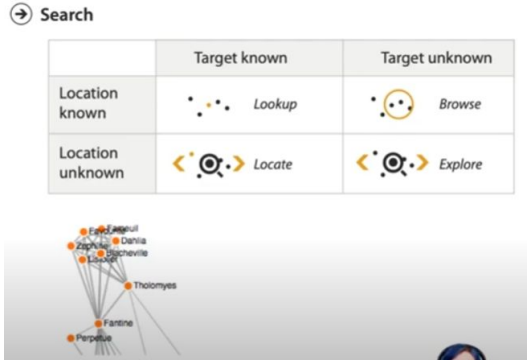
High-level choices

Consume vs. Produce data

Search

Search

Looking, locating, browsing, and exploring for an unknown item



Query

Identify

Looking (1 target), compare (2+ target), summarize (full set of targets)