Data Visualization 101

> *“Learn all you can about your data before anything else, and your analysis and visualization will be better for it. You can then pass what you know on to readers.”*

- Nathan Yau

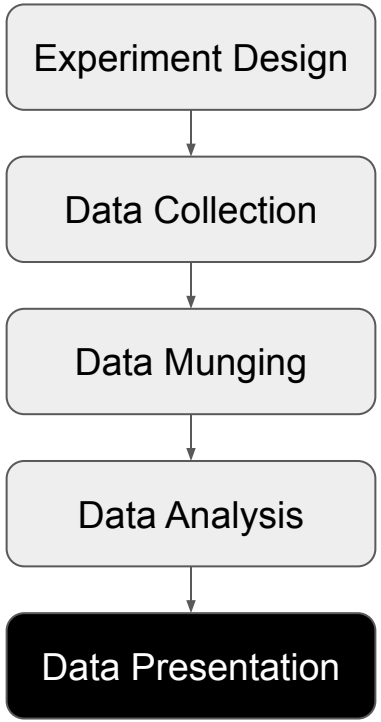
Data Visualization 101: Agenda

1. References
2. What is Data Visualization?
3. Designing for an Audience
4. Graphical Perception
5. Data Type Taxonomy

References

1. *Data Points: Visualization that Means Something* by Nathan Yau
2. *Graphical Perception: Theory, Experimentation and Application to the Development of Graphical Methods* by William S. Cleveland and Robert McGill
3. *The Eyes Have It: A Task by Data Type Taxonomy for Information Visualizations* by Ben Shneiderman
4. *Presenting Data to Non-Analysts: How to Make an Impact on All Kinds of Audiences* by Allison Sliter

What is Data Visualization?



What is Data Visualization?

> *“Data is more than numbers, and to visualize it, you must know what it represents. Data represents real life. It’s a snapshot of the world in the same way that a photograph captures a small moment in time.”*

- Nathan Yau

Data Visualization 101: Agenda

1. References
2. What is Data Visualization?
3. Designing for an Audience
4. Graphical Perception
5. Data Type Taxonomy

> *I am fully reliant on my audience – I need the questions my audience brings to me. Step one is reminding ourselves how dependent we are on our audience. Step two is making sure you can get what you need from your audience to do your job well.*

- Allison Sliter

Designing for an Audience

Who is my audience?

Real World

Data

Shapes and Colors

Designing for an Audience

What do we already know?

Real World

Data

Shapes and Colors

Designing for an Audience

What are the questions that need to be asked?

Real World

Data

Shapes and Colors

Designing for an Audience

AWARENESS	INFORMATION: Known	INFORMATION: Unknown
Known	Things we know we know	Things we know we don’t know (conscious ignorance)
Unknown	Things we don’t know we know (tacit knowledge)	Things we don’t know we don’t know (meta-ignorance)

Designing for an Audience

AWARENESS	INFORMATION: Known	INFORMATION: Unknown
Known	Things we know we know	Things we know we don’t know (conscious ignorance)
Unknown	Things we don’t know we know (tacit knowledge)	Things we don’t know we don’t know (meta-ignorance)

Designing for an Audience

AWARENESS	INFORMATION: Known	INFORMATION: Unknown
Known	Things we know we know	Things we know we don’t know (conscious ignorance)
Unknown	Things we don’t know we know (tacit knowledge)	Things we don’t know we don’t know (meta-ignorance)

Designing for an Audience

AWARENESS	INFORMATION: Known	INFORMATION: Unknown
Known	Things we know we know	Things we know we don't know (conscious ignorance)
Unknown	Things we don't know we know (tacit knowledge)	Things we don't know we don't know (meta-ignorance)

Designing for an Audience

AWARENESS	INFORMATION: Known	INFORMATION: Unknown
Known	Things we know we know	Things we know we don’t know (conscious ignorance)
Unknown	Things we don’t know we know (tacit knowledge)	Things we don’t know we don’t know (meta-ignorance)

Designing for an Audience

What is the context that needs to be encoded into our abstractions?

Real World

Data

Shapes and Colors

Designing for an Audience

What resources do I have access to?

Designing for an Audience

Who collected the data?

Designing for an Audience

What does the data represent in the world? Does it makes sense? How does it relate to other data?

Designing for an Audience

When, where, and why was the data collected?

Designing for an Audience

How do I know if it’s any good?

Designing for an Audience

The goal is for your audience to understand the data

Designing for an Audience

Follow guidelines but not as rules

Designing for an Audience

Try out different visualizations and compare them

Designing for an Audience

Balance functionality and uniqueness

Data Visualization 101: Agenda

1. References
2. What is Data Visualization?
3. Designing for an Audience
4. Graphical Perception
5. Data Type Taxonomy

> *The subject of graphical methods for data analysis and for data presentation needs a scientific foundation. In this article, we take a few steps in the direction of establishing such a foundation.*

- William S. Cleveland, Robert McGill

Graphical Perception

The visual **decoding** of information **encoded** on graphs

Graphical Perception

The visual **decoding** of information **encoded** on graphs

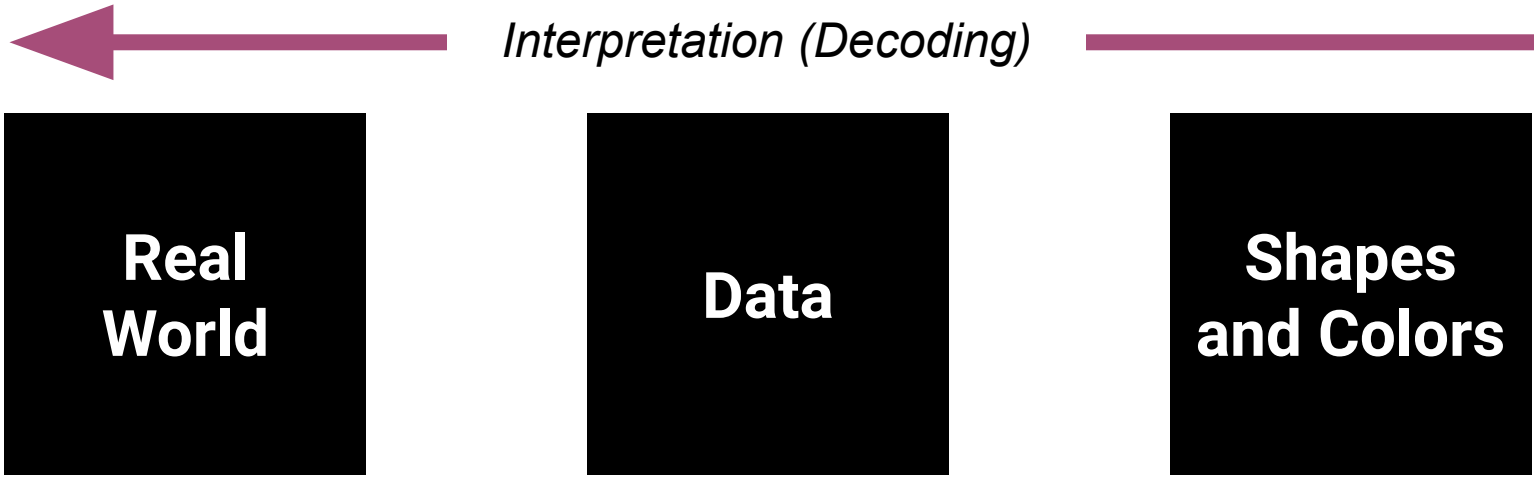
Real World

Data

Shapes and Colors

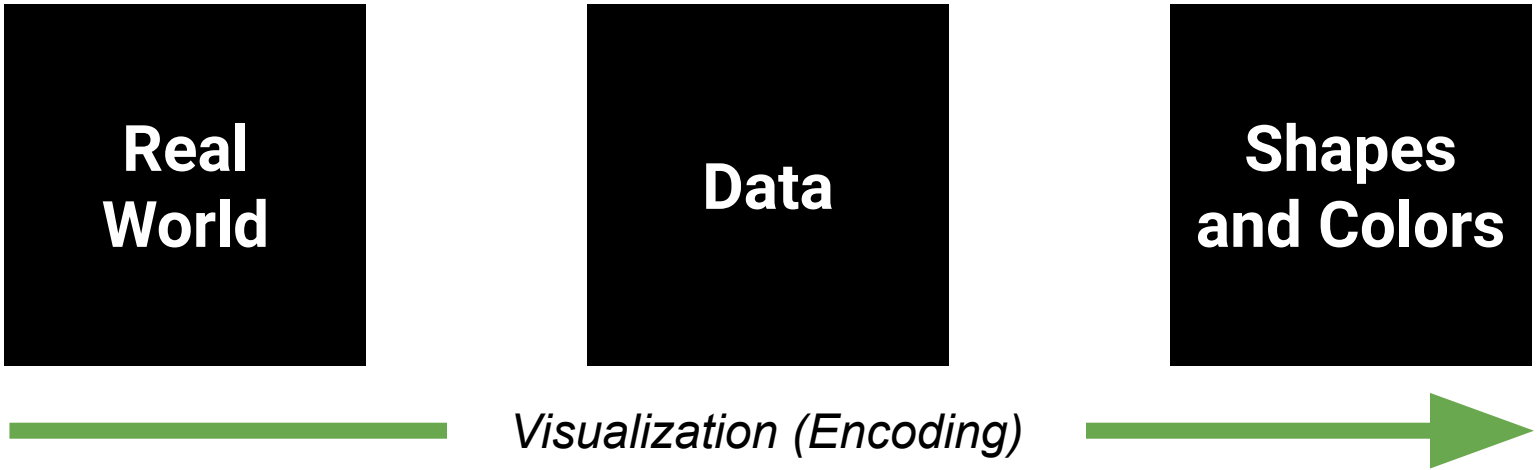
Graphical Perception

The visual **decoding** of information **encoded** on graphs



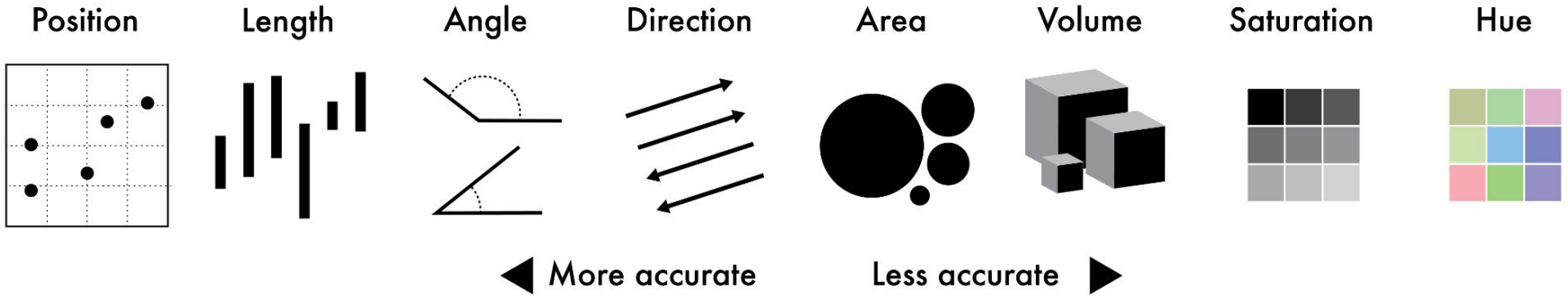
Graphical Perception

The visual **decoding** of information **encoded** on graphs



Graphical Perception

Perception of visual cues



Graphical Perception

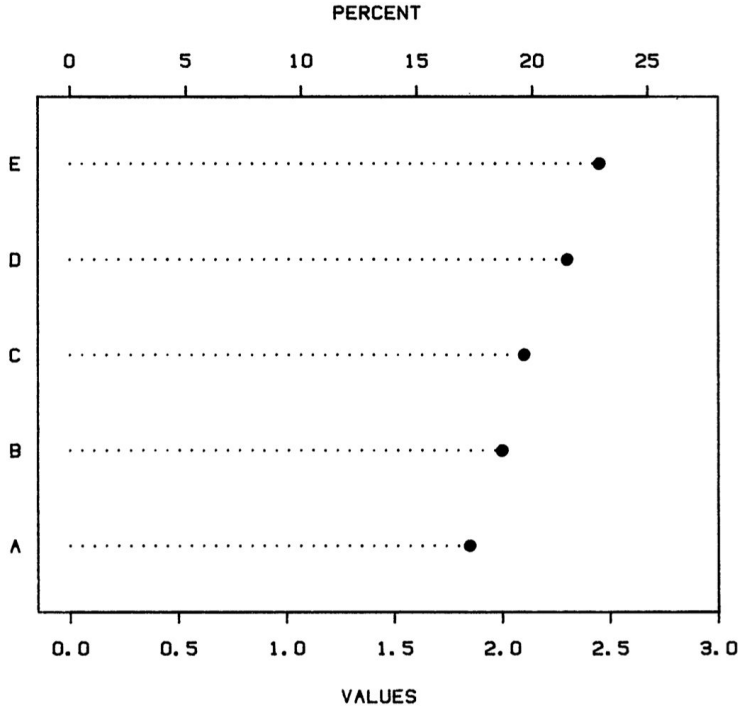


Figure 23. Dot chart.

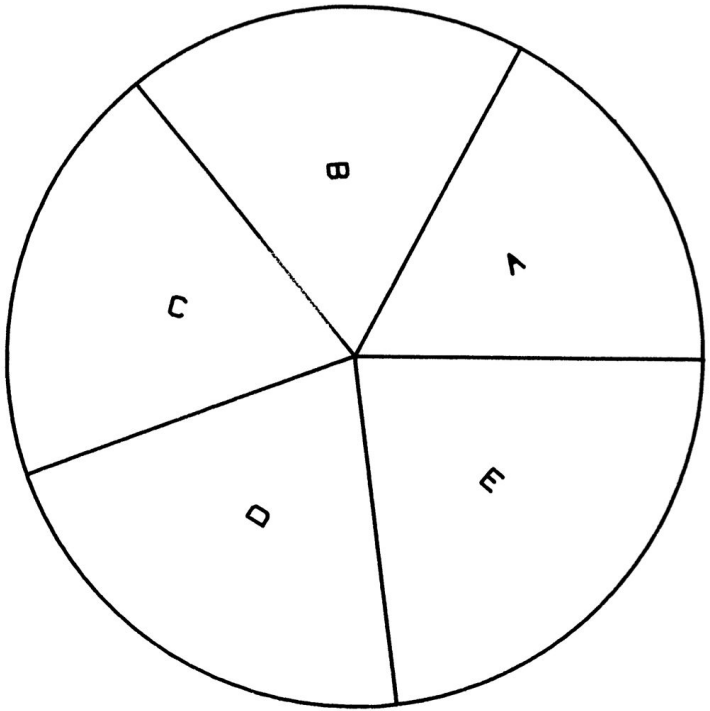


Figure 22. Pie chart.

Data Visualization 101: Agenda

1. References
2. What is Data Visualization?
3. Designing for an Audience
4. Graphical Perception
5. Data Type Taxonomy

Data Type Taxonomy

How can we categorize the types of **data** that we need to present?

Data Type Taxonomy

Seven Data Types

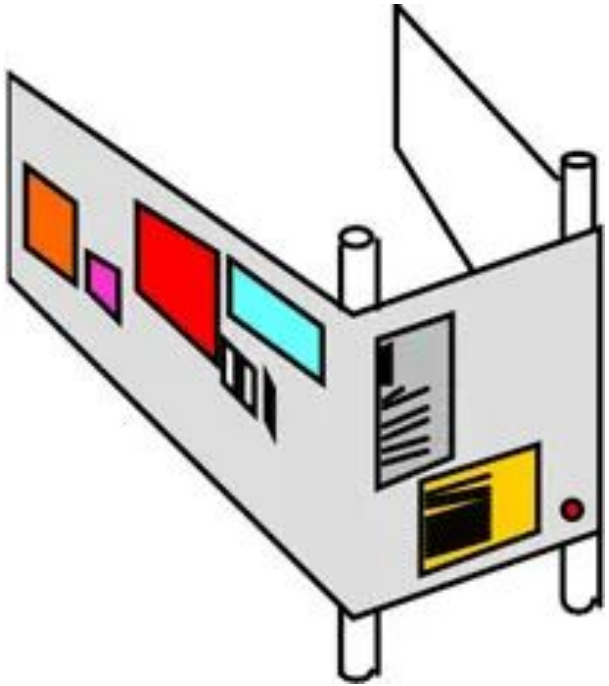
- 1D
- 2D
- 3D
- Temporal
- Multi-dimensional
- Tree
- Network

Data Type Taxonomy

1-dimensional data

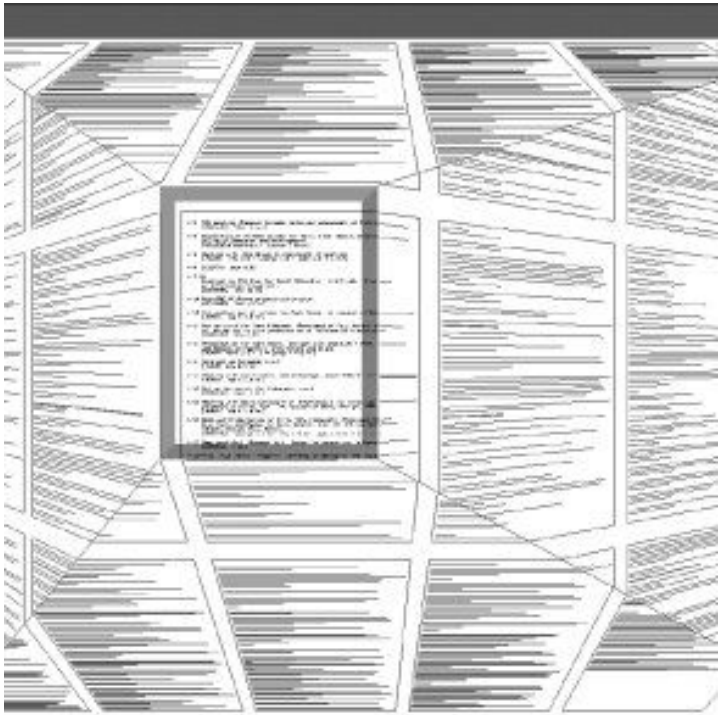
- Each item in the collection:
  - is a line of text containing a string of characters
  - may have different attributes

Data Type Taxonomy



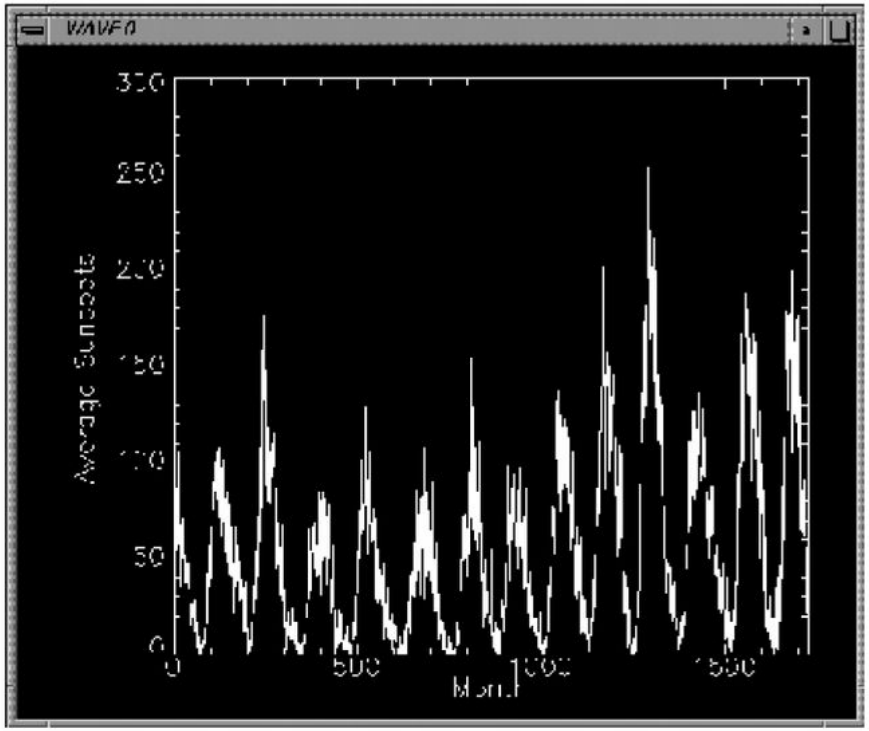
Bifocal Display (Spence and Apperley, 1982)

Data Type Taxonomy



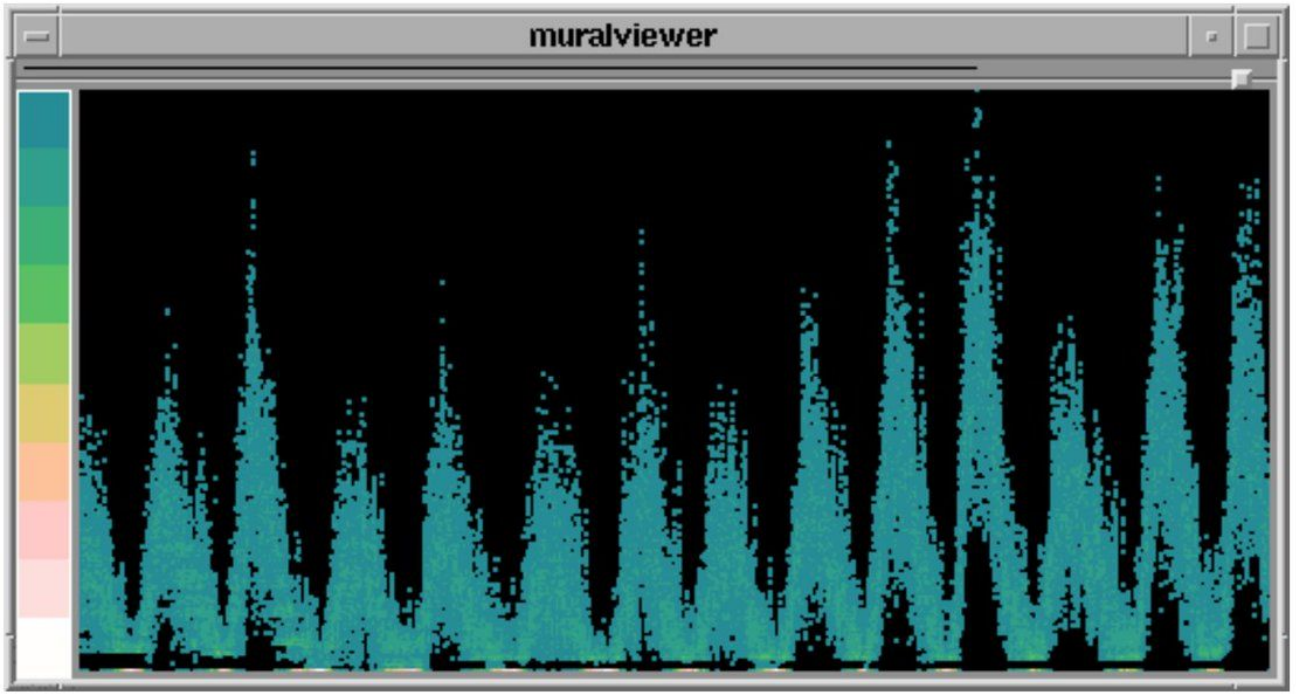
Document Lens (Robertson and Mackinley, 1993)

Data Type Taxonomy



Information Mural (Jerding and Stasko, 1995)

Data Type Taxonomy



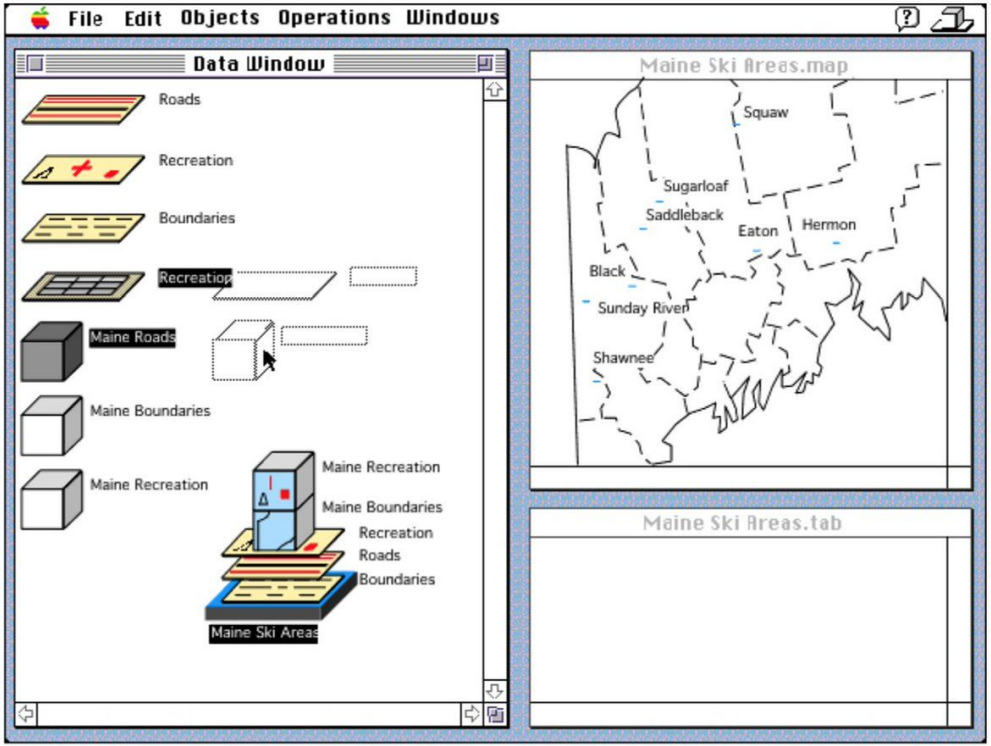
Information Mural (Jerding and Stasko, 1995)

Data Type Taxonomy

2-dimensional data

- Each item in the collection:
  - covers some part of the total area
  - has task-domain attributes (name, owner, value, etc)
  - has interface-domain features (size, color, opacity, etc)

Data Type Taxonomy



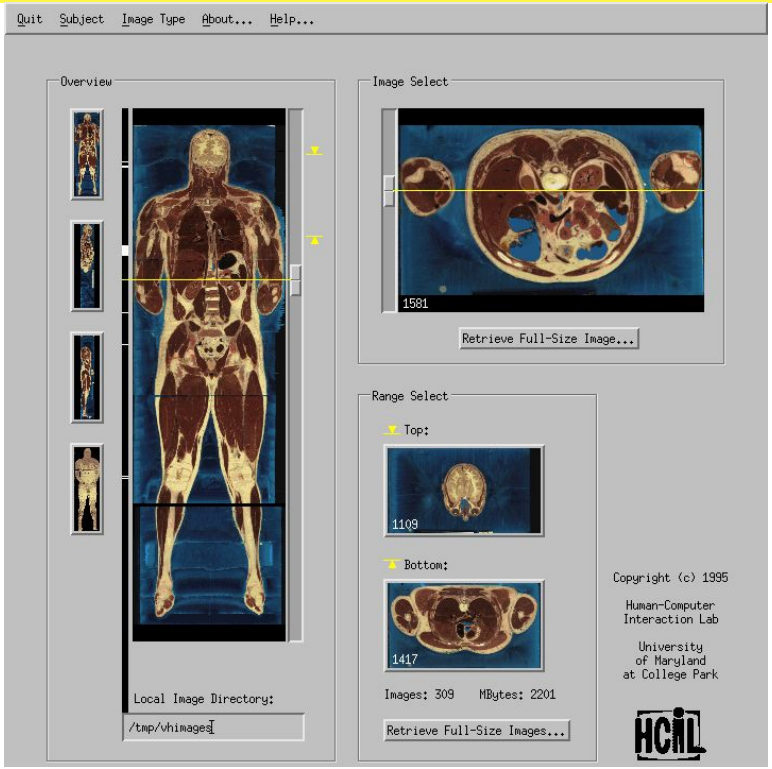
GIS (Laurini and Thompson, 1992)

Data Type Taxonomy

3-dimensional data

- Each item in the collection:
  - has volume
  - has potentially complex relationships with other items

Data Type Taxonomy



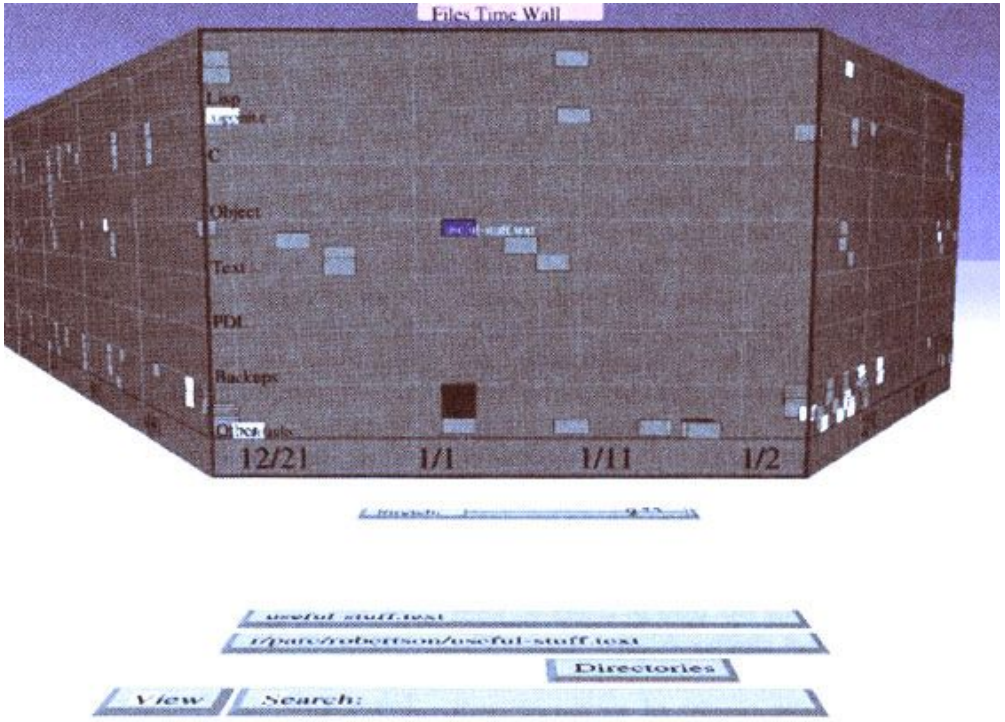
The Visible Human Project (North, 1996)

Data Type Taxonomy

Temporal data

- Each item in the collection:
  - has a start and finish time
  - may overlap with other items

Data Type Taxonomy



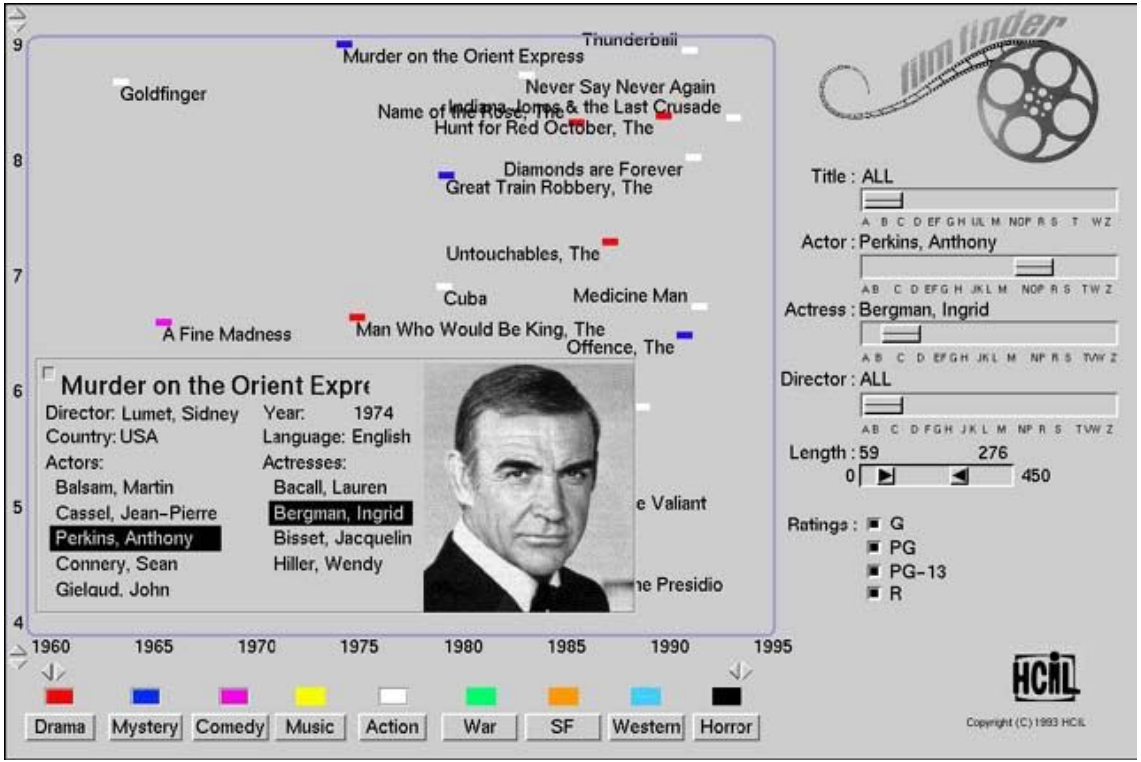
Perspective Wall (Robertson, 1993)

Data Type Taxonomy

Multi-dimensional data

- Each item in the collection:
  - has n-attributes
  - can be a part of a pattern or cluster
  - may have correlations with other items
  - may be an outlier
  - may be missing

Data Type Taxonomy



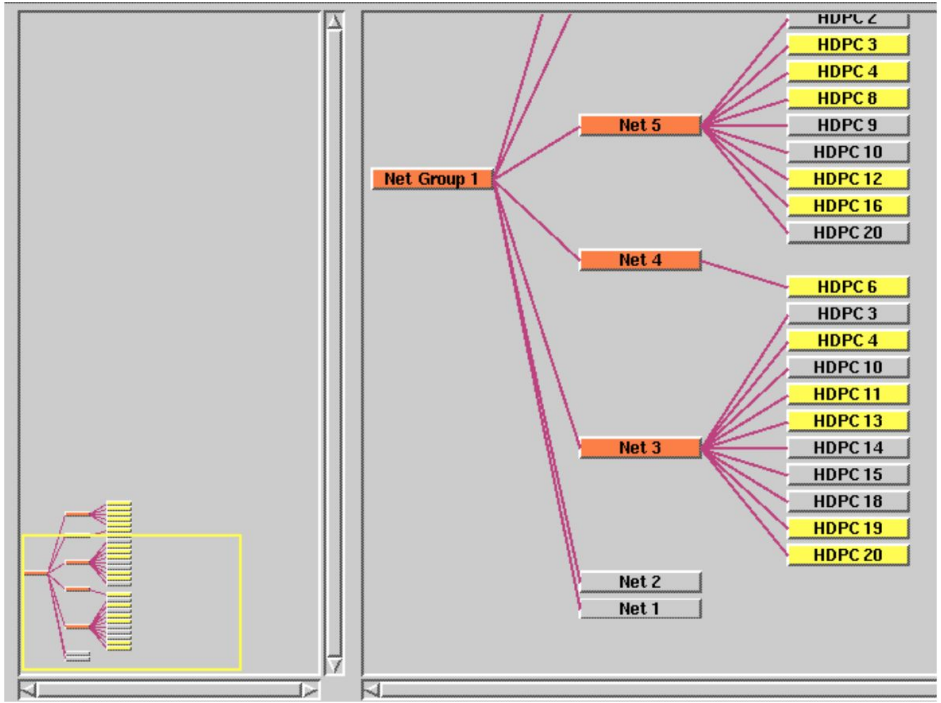
FilmFinder (Ahlberg and Shneiderman, 1994)

Data Type Taxonomy

Tree data

- Each item in the collection:
  - has relationships to other items (parent or child)
  - can have multiple attributes

Data Type Taxonomy



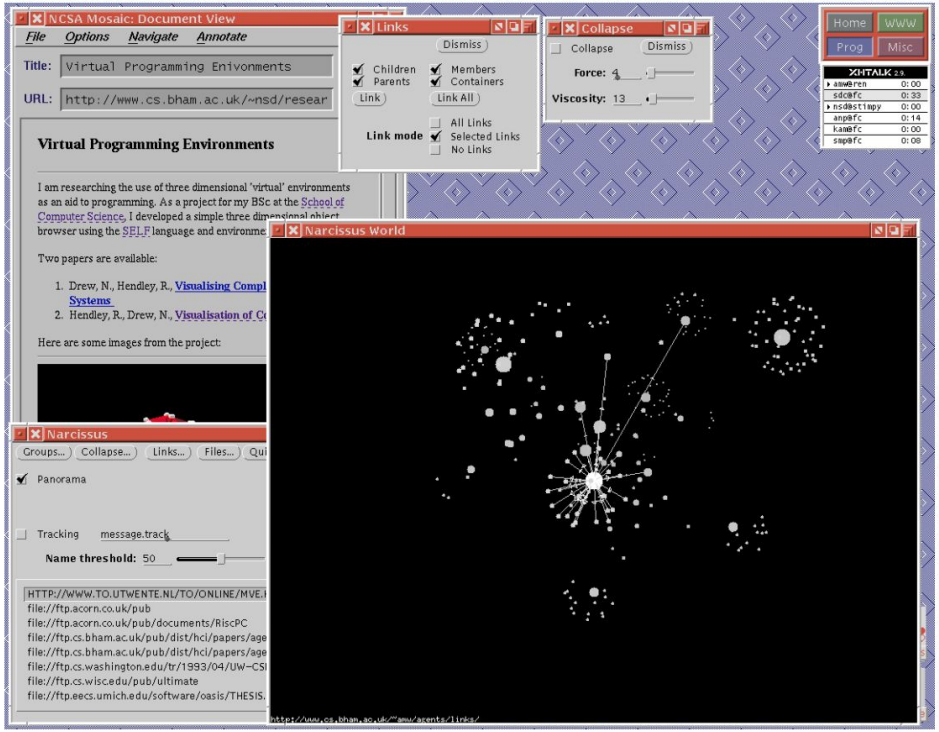
TreeBrowser (Kumar, 1995)

Data Taxonomy

Network data

- Each item in the collection:
  - can be linked to an arbitrary number of other items

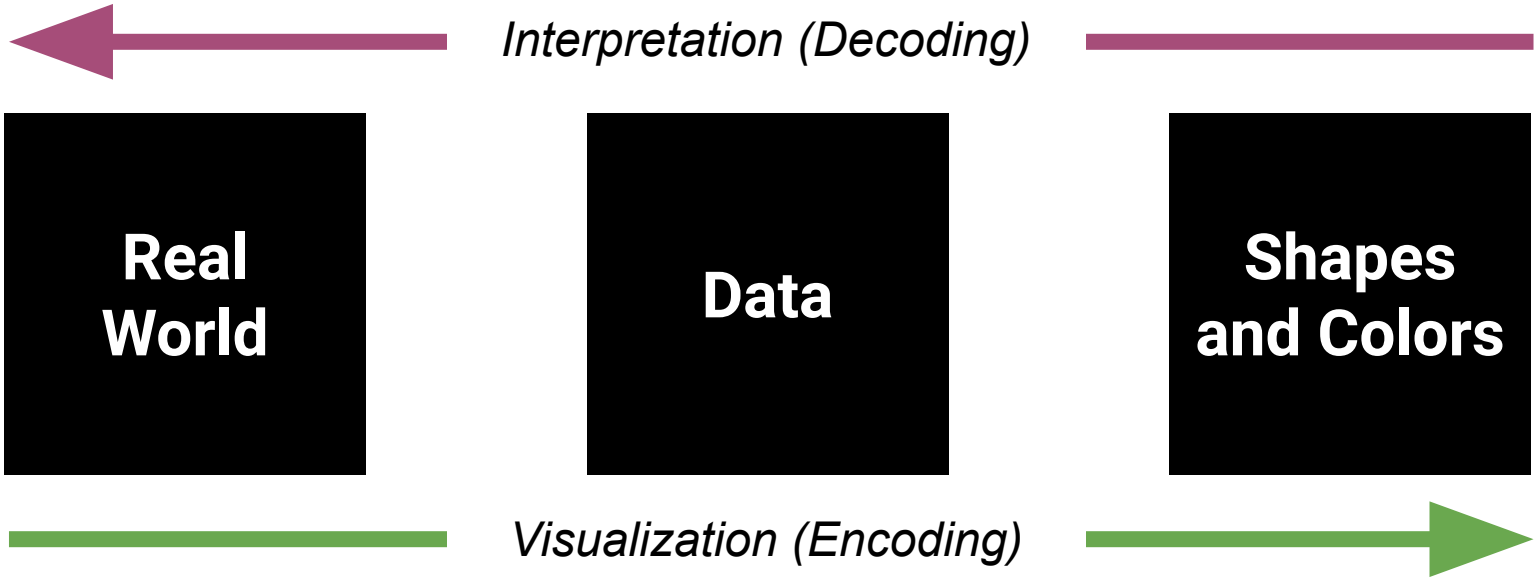
Data Taxonomy



Hendley, 1995

> *Good visualization is a winding process that requires **statistics** and **design knowledge**. Without the former, the visualization becomes an exercise only in illustration and aesthetics, and without the latter, one of only analyses. On their own, these are fine skills, but they make for incomplete data graphics. Having skills in both provides you with the luxury—which is growing into a necessity—to jump back and forth between data exploration and storytelling.*

- Nathan Yau


Interpretation (Decoding)
Real World
Data
Shapes and Colors
Visualization (Encoding)