Information Visualization (Part II)

EECS6414 – Data Analytics and Visualization
Agenda*

• Review
  – What is data visualization?
  – Jacques Bertin’s visual variables (semiotics)
  – Perception & cognition (pre-attentive vs attentive processing)
  – Gestalt principles
  – Tufte’s principles of graphical excellence

• Data Types

• A Taxonomy of Representation
  – A detailed listing of data representations

*Thanks to Ana Jofre for part of content in slides
Part I Review
Why visualize data? Anscombe’s Quartet

Summary statistics for all four datasets
- $\text{avg}(x) = 9$
- $\text{avg}(y) = 7.50$
- $\text{Var}(x) = 11$
- $\text{Var}(y) = 4.12$
- $\text{Correlation}(x,y) = 0.816$
- A linear regression line: $y = 0.5x + 3$

Always plot your data!

Anscombe’s Quartet
What is data visualization?

Use of **visual elements** like charts, graphs, and maps to see and understand trends, outliers, and patterns in data.
Jacques Bertin’s visual variables (vv)

Jacques Bertin proposed an original set of “retinal variables” in Semiology of Graphics (1967)
Perception & cognition

- perception is fragmented
- eyes are constantly scanning and constructing reality

The “Door Study”*
https://www.youtube.com/embed/FWSxSQsspiQ

## Pre-attentive vs attentive processing

<table>
<thead>
<tr>
<th>Pre-attentive Processing</th>
<th>Attentive Processing</th>
</tr>
</thead>
<tbody>
<tr>
<td>- bottom-up</td>
<td>- top-down</td>
</tr>
<tr>
<td>- fast, automatic</td>
<td>- slow, deliberate</td>
</tr>
<tr>
<td>- instinctive</td>
<td>- focused</td>
</tr>
<tr>
<td>- efficient</td>
<td>- singe-task</td>
</tr>
<tr>
<td>- multitasks</td>
<td></td>
</tr>
</tbody>
</table>

### Goal of information design
- help humans process information as efficiently as possible
- make as much use of pre-attentive processing as possible
Gestalt Principles (Princ. of Visual Grouping)

- Figure/Ground
- Proximity
- Similarity
- Symmetry
- Continuity
- Closure

**Gestalt Principles**

**Good Figure**
Objects grouped together tend to be perceived as a single figure. Tendency to simplify.

**Proximity**
Objects tend to be grouped together if they are close to each other.

**Similarity**
Objects tend to be grouped together if they are similar.

**Continuation**
When there is an intersection between two or more objects, people tend to perceive each object as a single uninterrupted object.

**Closure**
Visual connection or continuity between sets of elements which do not actually touch each other in a composition.

**Symmetry**
The object tend to be perceived as symmetrical shapes that form around their center.
Principles of Graphical Excellence (Tufte’ 01)

• Show the data
• Induce the viewer to think about the substance of the findings rather than the methodology, the graphical design, or other aspects
• Avoid distorting what the data have to say
• Present many numbers in a small space, i.e., efficiently
• Make large data sets coherent
• Encourage the eye to compare different pieces of data
• Reveal the data at several levels of detail, from a broad overview to the fine structure
• Serve a clear purpose: description, exploration, tabulation, decoration
• Be closely integrated with the statistical and verbal descriptions of the data set

High data to ink ratio (demo)
What makes a visualization beautiful?

https://informationisbeautiful.net/visualizations/what-makes-a-good-data-visualization/
Physical visualizations (data sculpture)

Keyboard Frequency Sculpture

A 3D bar chart on top of a keyboard which shows the frequency of each letter in the alphabet. 
*Source: Michael Knuepfel*

2011 – Tōhoku Japanese Earthquake Sculpture

A data sculpture by Luke Jerram that depicts nine minutes of seismographic readings during the 9.0 earthquake. 
*Source: Gizmodo*
Data Types
Data types

Data

Qualitative (Descriptive)
- Nominal: Data has no natural order
  - examples: gender, race, religion, sport
- Ordinal: Data can be arranged in order or rank
  - examples: sizes (s/m/l), attitudes (disagree, neutral, agree), house number.

Quantitative (Numerical)
- Continuous: Data is measured on a continuous scale
  - examples: temperature, length, height
- Discrete: Data is countable, and exists only in whole numbers
  - examples: number of people taking this class
Information Visualization Taxonomy
Information Visualization Taxonomy

Qualitative Data
- qualitative
- relational

Quantitative Data
- quantitative
- comparison
- relational

Textual

Hierarchical

Temporal

Spatial

Spatiotemporal

Distributions
Quantitative Comparison Structures

Qualitative Data
- qualitative
- relational

Quantitative Data
- quantitative
- comparison
- temporal
- spatial
- spatiotemporal

Distributions
Pie Chart

Image: https://google.github.io/charts/flutter/gallery.html
Doughnut Chart

Image: https://google.github.io/charts/flutter/gallery.html
Bar Chart

Image: https://google.github.io/charts/flutter/gallery.html
Stacked Bar Chart

Image: https://google.github.io/charts/flutter/gallery.html
Clustered/Grouped Bar Chart

Image: https://google.github.io/charts/flutter/gallery.html
Grouped-stacked Bar Chart

Image: https://google.github.io/charts/flutter/gallery.html
Bubble Chart

Image: https://github.com/UsabilityEtc/d3-country-bubble-chart
Bubble Chart (interactive)

Four Ways to Slice Obama’s 2013 Budget Proposal

Explore every nook and cranny of President Obama’s federal budget proposal.

How $3.7 Trillion Is Spent

Mr. Obama’s budget proposal includes $3.7 trillion in spending in 2013, and forecasts a $901 billion deficit.

Circles are sized according to the proposed spending.

Color shows amount of cut or increase from 2012.

The proposal forecasts a $901 billion deficit.

Pictogram Chart (for discrete data)

use of icons to give a more engaging overall view of small sets of discrete data can help overcome differences in language, culture and education

Image: https://datavizcatalogue.com/methods/pictogram.html#Vk0RGt-rSRs
Quantitative Relational Structures

Qualitative Data

- qualitative relational

Quantitative Data

- quantitative relational
- quantitative comparison
- quantitative relational
- distributions

- spatiotemporal
- spatial
- temporal
- hierarchical
- textual
Line Chart

Image: https://google.github.io/charts/flutter/gallery.html
Scatter Plot

Image: https://google.github.io/charts/flutter/gallery.html
3d Coordinate Systems

Alternative ways to define plane or 3D-space:
• Cartesian
• Cylindrical
• Spherical

Image: https://planetcalc.com/7952/
Radar Chart

Image: https://python-graph-gallery.com/390-basic-radar-chart/
Surface Plot

Image: https://google.github.io/charts/flutter/gallery.html
Heat Map

Google eye-tracking heatmap study to optimize analysis of search results
Heat Map/ Co-occurrence Map

Flows into all fund classes (all, including ETFs) – a time series

**Figure 15: Heat map** showing the flows as % of total assets into various fund classes (all including ETFs)

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Total Equity Funds</td>
<td>4%</td>
<td>3%</td>
<td>3%</td>
<td>1%</td>
<td>-3%</td>
<td>2%</td>
<td>2%</td>
<td>-1%</td>
<td>1%</td>
<td>3.4%</td>
</tr>
<tr>
<td>Total Developed Market Equity Funds</td>
<td>4%</td>
<td>2%</td>
<td>2%</td>
<td>-1%</td>
<td>-3%</td>
<td>-3%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>3.8%</td>
</tr>
<tr>
<td>International Equity Funds</td>
<td>8%</td>
<td>6%</td>
<td>7%</td>
<td>6%</td>
<td>-4%</td>
<td>4%</td>
<td>1%</td>
<td>1%</td>
<td>1%</td>
<td>3.8%</td>
</tr>
<tr>
<td>US Equity Funds</td>
<td>1%</td>
<td>-1%</td>
<td>-1%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>-1%</td>
<td>3.5%</td>
</tr>
<tr>
<td>Western Europe Equity Funds</td>
<td>1%</td>
<td>-1%</td>
<td>7%</td>
<td>-13%</td>
<td>-12%</td>
<td>1%</td>
<td>-3%</td>
<td>-2%</td>
<td>-2%</td>
<td>0.4%</td>
</tr>
<tr>
<td>Japan Equity Funds</td>
<td>56%</td>
<td>44%</td>
<td>6%</td>
<td>-27%</td>
<td>-16%</td>
<td>-16%</td>
<td>-3%</td>
<td>5%</td>
<td>10%</td>
<td>24.7%</td>
</tr>
<tr>
<td>Pacific Equity Funds</td>
<td>7%</td>
<td>3%</td>
<td>12%</td>
<td>1%</td>
<td>16%</td>
<td>17%</td>
<td>8%</td>
<td>8%</td>
<td>1%</td>
<td>7.9%</td>
</tr>
<tr>
<td>Total Emerging Market Equity Funds</td>
<td>3%</td>
<td>16%</td>
<td>11%</td>
<td>12%</td>
<td>-7%</td>
<td>27%</td>
<td>16%</td>
<td>5%</td>
<td>7%</td>
<td>0.4%</td>
</tr>
<tr>
<td>Global Emerging Market Equity Funds</td>
<td>10%</td>
<td>3%</td>
<td>4%</td>
<td>16%</td>
<td>-4%</td>
<td>22%</td>
<td>23%</td>
<td>-1%</td>
<td>12%</td>
<td>2.5%</td>
</tr>
<tr>
<td>EMMA Equity Funds</td>
<td>27%</td>
<td>40%</td>
<td>6%</td>
<td>2%</td>
<td>9%</td>
<td>11%</td>
<td>29%</td>
<td>11%</td>
<td>4%</td>
<td>7.4%</td>
</tr>
<tr>
<td>Latin America Equity Funds</td>
<td>10%</td>
<td>31%</td>
<td>27%</td>
<td>46%</td>
<td>12%</td>
<td>48%</td>
<td>4%</td>
<td>12%</td>
<td>1%</td>
<td>6.5%</td>
</tr>
<tr>
<td>Asia Pacific Ex-Japan Funds</td>
<td>21%</td>
<td>22%</td>
<td>27%</td>
<td>14%</td>
<td>9%</td>
<td>21%</td>
<td>10%</td>
<td>7%</td>
<td>3%</td>
<td>0.2%</td>
</tr>
<tr>
<td>Total Bond Funds</td>
<td>14%</td>
<td>4%</td>
<td>8%</td>
<td>2%</td>
<td>-10%</td>
<td>24%</td>
<td>16%</td>
<td>4%</td>
<td>11%</td>
<td>1.5%</td>
</tr>
<tr>
<td>International Bond Funds</td>
<td>12%</td>
<td>12%</td>
<td>16%</td>
<td>2%</td>
<td>-10%</td>
<td>25%</td>
<td>23%</td>
<td>3%</td>
<td>6%</td>
<td>1.1%</td>
</tr>
<tr>
<td>Corporate High Yield Bond Funds</td>
<td>NA</td>
<td>-18%</td>
<td>-2%</td>
<td>-4%</td>
<td>-5%</td>
<td>40%</td>
<td>15%</td>
<td>4%</td>
<td>18%</td>
<td>1.4%</td>
</tr>
<tr>
<td>US Bond Funds</td>
<td>NA</td>
<td>-17%</td>
<td>-3%</td>
<td>4%</td>
<td>-2%</td>
<td>23%</td>
<td>19%</td>
<td>6%</td>
<td>12%</td>
<td>2.2%</td>
</tr>
<tr>
<td>Western Europe Bond Funds</td>
<td>NA</td>
<td>1%</td>
<td>58%</td>
<td>8%</td>
<td>46%</td>
<td>29%</td>
<td>7%</td>
<td>26%</td>
<td>2%</td>
<td>-3.4%</td>
</tr>
<tr>
<td>Germany Bond funds</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>65%</td>
<td>-19%</td>
<td>-2%</td>
</tr>
<tr>
<td>Switzerland Bond funds</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>-65%</td>
<td>-19%</td>
<td>-2%</td>
</tr>
<tr>
<td>United Kingdom Bond funds</td>
<td>NA</td>
<td>22%</td>
<td>-17%</td>
<td>-141%</td>
<td>-26%</td>
<td>64%</td>
<td>8%</td>
<td>-3%</td>
<td>0%</td>
<td>-4.1%</td>
</tr>
<tr>
<td>Emerging Markets Debt Funds</td>
<td>12%</td>
<td>24%</td>
<td>18%</td>
<td>9%</td>
<td>21%</td>
<td>19%</td>
<td>54%</td>
<td>7%</td>
<td>25%</td>
<td>2.4%</td>
</tr>
<tr>
<td>Asia ex-Japan Bond funds</td>
<td>NA</td>
<td>4%</td>
<td>3%</td>
<td>16%</td>
<td>16%</td>
<td>2%</td>
<td>21%</td>
<td>25%</td>
<td>12%</td>
<td>2.2%</td>
</tr>
<tr>
<td>Emerging Europe Bond funds</td>
<td>NA</td>
<td>40%</td>
<td>-12%</td>
<td>-18%</td>
<td>-37%</td>
<td>-19%</td>
<td>28%</td>
<td>38%</td>
<td>6%</td>
<td>0.1%</td>
</tr>
<tr>
<td>Lat-Am Bond funds</td>
<td>NA</td>
<td>-26%</td>
<td>-22%</td>
<td>-35%</td>
<td>-10%</td>
<td>31%</td>
<td>17%</td>
<td>18%</td>
<td>4%</td>
<td>-1.2%</td>
</tr>
<tr>
<td>Money Market Funds</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
</tr>
</tbody>
</table>

*Colour Legend:*

- >10%  
- 8% to 10%  
- 0% to 5%  
- -5% to 0%  
- -10% to -5%  
- < -10%

Source: EY, Deutsche Bank analysts

Area Graph

Image: https://google.github.io/charts/flutter/gallery.html
Stacked Area Graph

Image: https://google.github.io/charts/flutter/gallery.html
Distribution Structures

Qualitative Data
- qualitative
- relational

Quantitative Data
- quantitative
- relational

Textual
Hierarchical
Temporal
Spatial
Spatiotemporal

Distributions
Box and Whisker Box

Interquartile range (IQR)

Left-Skewed  Symmetric  Right-Skewed

Image: https://www.simplypsychology.org/boxplots.html
Box and Whisker Box

Image: https://www.simplypsychology.org/boxplots.html
Histogram

Test scores

Students scores

No. of students

Image: https://support.google.com/docs/answer/9146867?hl=en
Quantitative Relational Structures

Qualitative Data

qualitative relational

textual

hierarchical

temporal

spatial

spatiotemporal

Quantitative Data

quantitative comparison

quantitative relational

distributions
Networks: Force-directed Layout

Image: https://homes.cs.washington.edu/~jheer/files/zoo/
Networks: Arc Graph

Image: https://homes.cs.washington.edu/~jheer/files/zoo/
Networks: Adjacency Matrix

Image: https://homes.cs.washington.edu/~jheer/files/zoo/
Networks: Chord Diagram

movement of refugees within Africa in 2011

Networks: Chord Diagram (interactive)

Uber rides in SF by neighborhoods

Image: https://bost.ocks.org/mike/uberdata/
Networks: Sankey Charts

Movement between countries

Image: https://www.highcharts.com/demo/sankey-diagram
Textual Structures

Qualitative Data
- qualitative
- relational

Quantitative Data
- quantitative
- relational

Overlap:
- textual
- hierarchical
- temporal
- spatial
- spatiotemporal

Distributions
Word Cloud

Image: https://blog.sharetolearn.com/classroom-resources/word-clouds-writing/
Word Tree

- cats
- are
- eat
- kibble
- mice
- better than
  - awesome
  - people too
  - family
  - evil
  - weird
- dogs
- hamsters
- kittens
- meowing
- in the cradle
- lyrics
- for adoption

Image: https://developers.google.com/chart/interactive/docs/gallery/wordtree
Hierarchical Structures

Qualitative Data
- qualitative
- relational

Quantitative Data
- quantitative
- relational

intersection:
- textual
- temporal
- spatial
- spatiotemporal

hierarchical

quantitative comparison

distributions
Tree Diagram (root, branches, nodes, leaves)

Horizontal Trees

Image: https://developers.google.com/chart/interactive/docs/gallery/wordtree
Node-link Diagram or Dendrogram

Image: https://homes.cs.washington.edu/~jheer/files/zoo/
Indented Trees & Circular Dendrogram

Image: https://homes.cs.washington.edu/~jheer/files/zoo/
Radial Trees

Hyperbolic Trees

Image: https://infovis-wiki.net/wiki/Hyperbolic_trees
Rectangular TreeMaps: World Population

Image: https://www.populationpyramid.net/population-size-per-country/2020/
Voronoi TreeMap

Voronoi treemaps are an alternative to traditional rectangular treemaps
• often more aesthetically pleasing

Image: https://homes.cs.washington.edu/~jheer/files/zoo/
Circular TreeMap

packing circles instead of subdividing rectangles can produce a different sort of enclosure diagram that has an organic appearance

- circle-packing layouts reveals the hierarchy
- node sizes can be rapidly compared using area

Image: https://homes.cs.washington.edu/~jheer/files/zoo/
Radial TreeMap (aka Sunbursts)

Image: https://homes.cs.washington.edu/~jheer/files/zoo/
Icicle TreeMap

Image: https://homes.cs.washington.edu/~jheer/files/zoo/
Temporal Structures

Qualitative Data
- qualitative
- relational
- textual
- hierarchical
- spatiotemporal

Temporal
- quantitative
- comparison
- relational
- distributions

Spatiotemporal
- spatial
Timeline of Space Exploration

- **03 Feb 1966**: First soft landing on the Moon
- **17 Apr 1976**: Closest flyby of the Sun
- **19 Feb 1986**: First inhabited space station
- **20 Nov 1998**: First multinational space station
- **04 Oct 1957**: First artificial satellite
- **20 Jul 1969**: First human on the Moon
- **04 Dec 1978**: First orbital exploration of Venus
- **08 Aug 1989**: First astrometric satellite

Info source: www.wikipedia.org

Image: https://www.highcharts.com/
Timeline of 100 yrs of Rock Music (interact.)

Image: http://www.concerthotels.com/100-years-of-rock
Spatial Structures

Qualitative Data
- qualitative
- relational
- temporal
- spatial
- spatiotemporal

Quantitative Data
- quantitative
- comparison
- relational
- distributions
Heatmaps: Fandom of Coldplay (on Youtube)

Choropleth Map

Image: https://homes.cs.washington.edu/~jheer/files/zoo/
Graduated Symbol Map

Image: https://homes.cs.washington.edu/~jheer/files/zoo/
Cartograms

Image: https://homes.cs.washington.edu/~jheer/files/zoo/
Examples: Wind Map (interactive)

**March 15, 2020**
9:42 pm EST
(time of forecast download)

top speed: **26.5 mph**
average: **7.7 mph**

Spatiotemporal Structures

Qualitative Data
- qualitative
- relational
- textual
- hierarchical
- temporal
- spatial

Spatiotemporal

Quantitative Data
- quantitative comparison
- quantitative relational
- distributions
Charles Minard’s map of Napoleon’s disastrous Russian campaign of 1812. The graphic is notable for its representation in two dimensions of six types of data: the number of Napoleon’s troops; distance; temperature; the latitude and longitude; direction of travel; and location relative to specific dates.
Hurricane Katrina Trajectory Over Time

Tracing Earthquake Discussions in Real Time

Image/video: https://www.youtube.com/watch?v=ou8L0MzGvOU
Other
Chernoff Faces

A way to display $n$ variables on a 2-D surface where each variable is assigned one of $k$ possible values.
Chernoff Faces: Example

Chernoff faces for lawyers' ratings of 12 judges
Beyond digital visualizations
Physical visualizations (data sculpture)

Keyboard Frequency Sculpture

A 3D bar chart on top of a keyboard which shows the frequency of each letter in the alphabet.  
*Source: Michael Knuepfel*

2011 – Tōhoku Japanese Earthquake Sculpture

A data sculpture by Luke Jerram that depicts nine minutes of seismographic readings during the 9.0 earthquake.  
*Source: Gizmodo*
Physical visualizations

Manifest Justice Exhibition, Los Angeles, May 2015
Participatory visualization

https://www.youtube.com/watch?v=hD5f8GuNuGQ

(what is privilege?)
Resources
Data Visualization Resources & Libraries

Data visualization catalog
http://www.datavizcatalogue.com/

Periodic table of visualization methods
http://www.visual-literacy.org/periodic_table/periodic_table.html

Interactive dynamics for visual analysis (Taxonomy of Tools)
http://queue.acm.org/detail.cfm?id=2146416

HighCharts library
https://www.highcharts.com/

D3 library
https://observablehq.com/@d3/gallery