Data Visualization (CSCI 627/490)

Marks and Channels

Dr. David Koop
### Visual Encoding

- How should we visualize this data?

<table>
<thead>
<tr>
<th>Name</th>
<th>Region</th>
<th>Population</th>
<th>Life Expectancy</th>
<th>Income</th>
</tr>
</thead>
<tbody>
<tr>
<td>China</td>
<td>East Asia &amp; Pacific</td>
<td>1335029250</td>
<td>73.28</td>
<td>7226.07</td>
</tr>
<tr>
<td>India</td>
<td>South Asia</td>
<td>1140340245</td>
<td>64.01</td>
<td>2731</td>
</tr>
<tr>
<td>United States</td>
<td>America</td>
<td>306509345</td>
<td>79.43</td>
<td>41256.08</td>
</tr>
<tr>
<td>Indonesia</td>
<td>East Asia &amp; Pacific</td>
<td>228721000</td>
<td>71.17</td>
<td>3818.08</td>
</tr>
<tr>
<td>Brazil</td>
<td>America</td>
<td>193806549</td>
<td>72.68</td>
<td>9569.78</td>
</tr>
<tr>
<td>Pakistan</td>
<td>South Asia</td>
<td>176191165</td>
<td>66.84</td>
<td>2603</td>
</tr>
<tr>
<td>Bangladesh</td>
<td>South Asia</td>
<td>156645463</td>
<td>66.56</td>
<td>1492</td>
</tr>
<tr>
<td>Nigeria</td>
<td>Sub-Saharan Africa</td>
<td>141535316</td>
<td>48.17</td>
<td>2158.98</td>
</tr>
<tr>
<td>Japan</td>
<td>East Asia &amp; Pacific</td>
<td>127383472</td>
<td>82.98</td>
<td>29680.68</td>
</tr>
<tr>
<td>Mexico</td>
<td>America</td>
<td>111209909</td>
<td>76.47</td>
<td>11250.37</td>
</tr>
<tr>
<td>Philippines</td>
<td>East Asia &amp; Pacific</td>
<td>94285619</td>
<td>72.1</td>
<td>3203.97</td>
</tr>
<tr>
<td>Vietnam</td>
<td>East Asia &amp; Pacific</td>
<td>86970762</td>
<td>74.7</td>
<td>2679.34</td>
</tr>
<tr>
<td>Germany</td>
<td>Europe &amp; Central Asia</td>
<td>82338100</td>
<td>80.88</td>
<td>31191.15</td>
</tr>
<tr>
<td>Ethiopia</td>
<td>Sub-Saharan Africa</td>
<td>79996293</td>
<td>55.69</td>
<td>812.16</td>
</tr>
<tr>
<td>Turkey</td>
<td>Europe &amp; Central Asia</td>
<td>72626967</td>
<td>72.06</td>
<td>8040.78</td>
</tr>
</tbody>
</table>
Potential Solution

[Gapminder, Wealth & Health of Nations]
Visual Encoding

• How do we encode data visually?
  - **Marks** are the basic graphical elements in a visualization
  - **Channels** are ways to control the appearance of the marks

• Marks classified by dimensionality:
  - Points
  - Lines
  - Areas

• Also can have surfaces, volumes

• Think of marks as a mathematical definition, or if familiar with tools like Adobe Illustrator or Inkscape, the path & point definitions
Visual Channels

- **Position**
  - Horizontal
  - Vertical
  - Both

- **Color**

- **Shape**

- **Tilt**

- **Size**
  - Length
  - Area

- **Volume**

[Munzner (ill. Maguire), 2014]
Channel Types

- **Identity** => what or where, **Magnitude** => how much

**Magnitude Channels: Ordered Attributes**
- Position on common scale
- Position on unaligned scale
- Length (1D size)
- Tilt/angle
- Area (2D size)
- Depth (3D position)
- Color luminance
- Color saturation
- Curvature
- Volume (3D size)

**Identity Channels: Categorical Attributes**
- Spatial region
- Color hue
- Motion
- Shape

[Munzner (ill. Maguire), 2014]
Assignment 3

• Same stacked bar chart visualization
• Three tools
  - Tableau (free academic license)
  - Vega-Lite
  - D3
• For Vega-Lite, use the online editor
• For D3, use template files so the data is properly loaded
• [CS 490] Only need to do a standard bar chart in D3
• Three parts: set mini-deadlines
Tableau Example
Data In Tableau

- Categorical data = Dimension
- Quantitative data = Measures
Vega-Lite Example
Expressiveness and Effectiveness

• Expressiveness Principle: all data from the dataset and nothing more should be shown
  - Do encode ordered data in an ordered fashion
  - Don’t encode categorical data in a way that implies an ordering

• Effectiveness Principle: the most important attributes should be the most salient
  - Saliency: how noticeable something is
  - How do the channels we have discussed measure up?
Mackinlay's Ranking of Perceptual Tasks

Quantitative
- Position
- Length
- Angle
- Slope
- Area
- Volume
- Density
- Color Saturation
- Color Hue

Ordinal
- Position
- Density
- Color Saturation
- Color Hue
- Texture
- Connection
- Containment
- Length
- Angle
- Slope
- Area
- Volume

Nominal
- Position
- Color Hue
- Texture
- Connection
- Containment
- Density
- Color Saturation
- Shape
- Length
- Angle
- Slope
- Area
- Volume

[Mackinlay, 1986]
### Iliinsky's Best Uses, +Ordering, +NumValues

<table>
<thead>
<tr>
<th>Example</th>
<th>Encoding</th>
<th>Ordered</th>
<th>Useful values</th>
<th>Quantitative</th>
<th>Ordinal</th>
<th>Categorical</th>
<th>Relational</th>
</tr>
</thead>
<tbody>
<tr>
<td>position, placement</td>
<td></td>
<td>yes</td>
<td>infinite</td>
<td>Good</td>
<td>Good</td>
<td>Good</td>
<td>Good</td>
</tr>
<tr>
<td>1, 2, 3; A, B, C</td>
<td>text labels</td>
<td>optional (alphabetical or numbered)</td>
<td>infinite</td>
<td>Good</td>
<td>Good</td>
<td>Good</td>
<td>Good</td>
</tr>
<tr>
<td></td>
<td>length</td>
<td>yes</td>
<td>many</td>
<td>Good</td>
<td>Good</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>size, area</td>
<td>yes</td>
<td>many</td>
<td>Good</td>
<td>Good</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>angle</td>
<td>yes</td>
<td>medium/few</td>
<td>Good</td>
<td>Good</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>pattern density</td>
<td>yes</td>
<td>few</td>
<td>Good</td>
<td>Good</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>weight, boldness</td>
<td>yes</td>
<td>few</td>
<td>Good</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>saturation, brightness</td>
<td>yes</td>
<td>few</td>
<td>Good</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>color</td>
<td>no</td>
<td>few (&lt; 20)</td>
<td>Good</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>shape, icon</td>
<td>no</td>
<td>medium</td>
<td>Good</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>pattern texture</td>
<td>no</td>
<td>medium</td>
<td>Good</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>enclosure, connection</td>
<td>no</td>
<td>infinite</td>
<td>Good</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>line pattern</td>
<td>no</td>
<td>few</td>
<td>Good</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>line endings</td>
<td>no</td>
<td>few</td>
<td>Good</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>line weight</td>
<td>yes</td>
<td>few</td>
<td>Good</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
How do we get these rankings?
Test % difference in length between elements

[Heer & Bostock, 2010]
Test % difference in **length** between elements

Answer: Left is \( \sim 5.6x \) longer than Right

[Heer & Bostock, 2010]
Test % difference in **length** between elements

![Diagram showing the test of length difference between elements A and B with a y-axis ranging from 0 to 100.](image)

[Heer & Bostock, 2010]
Test % difference in length between elements

[Heer & Bostock, 2010]
Test % difference in length between elements

\[ \text{judged percent - true percent} \]

[Modified from Heer & Bostock, 2010]
Test % difference in length between elements

Answer: Right is 4x larger than Left

[Modified from Heer & Bostock, 2010]
Test % difference in area between elements

[Heer & Bostock, 2010]
Test % difference in area between elements

Answer: A is ~2.25x larger (in area) than B

[Heer & Bostock, 2010]
Test % difference in area between elements

[Heer & Bostock, 2010]
Test % difference in **area** between elements

Answer: B is ~6.1x larger (in area) than A

![Diagram](image)

[Heer & Bostock, 2010]
Test % difference in area between elements

[Heer & Bostock, 2010]
Test % difference in area between elements

Answer: B is ~2.5 larger (in area) than A
Cleveland & McGill Experiments

Figure 4. Graphs from position–length experiment.

Figure 3. Graphs from position–angle experiment.

[Cleveland & McGill, 1984]
Heer & Bostock Experiments

- Rerun Cleveland & McGill’s experiment using Mechanical Turk
- … with more tests

Figure 2: Area judgment stimuli. Top left: Bubble chart (T7), Bottom left: Center-aligned rectangles (T8), Right: Treemap (T9).

[Heer & Bostock, 2010]
Results Summary

Cleveland & McGill’s Results

Crowdsourced Results

[Muñzner (ill. Maguire) based on Heer & Bostock, 2014]
Psychophysics

- How do we perceive changes in stimuli
- The Psychophysical Power Law [Stevens, 1975]: All sensory channels follow a power function based on stimulus intensity ($S = I^n$)
- Length is fairly accurate
- Magnified vs. compressed sensations
Ranking Channels by Effectiveness

**Magnitude Channels: Ordered Attributes**
- Position on common scale
- Position on unaligned scale
- Length (1D size)
- Tilt/angle
- Area (2D size)
- Depth (3D position)
- Color luminance
- Color saturation
- Curvature
- Volume (3D size)

**Identity Channels: Categorical Attributes**
- Spatial region
- Color hue
- Motion
- Shape

[Muñzner (ill. Maguire), 2014]
Discriminability

• Width encodes count of number of networks with a particular link.

• What is problematic here?
Discriminability

- Can someone tell the difference?
- How many values (bins) can be used so that a person can tell the difference?
- Example: Line width
  - Matching a particular width with a legend
  - Comparing two widths
Separability

- Cannot treat all channels as independent!
- **Separable** means each individual channel can be distinguished
- **Integral** means the channels are perceived together

![Diagram showing separability](image)

[Munzner (ill. Maguire) based on Ware, 2014]
Separable or Integral?
Separable or Integral?

The map at right is a product of overlaying the three sets of data. The variation in hue and value has been produced from the data shown above. In general, darker counties represent a more educated, better paid population while lighter areas represent communities with fewer graduates and lower incomes.
Visual Popout: Parallel Lines Require Search…

[Munzner (ill. Maguire), 2014]
Relative vs. Absolute Judgments

• Weber’s Law:
  - We judge based on relative not absolute differences
  - The amount of perceived difference is relative to the object’s magnitude!

[Unframed Unaligned]

[Munzner (ill. Maguire), 2014]
Luminance Perception

Edward H. Adelson

E. H. Adelson, 1995
Luminance Perception

[Edward H. Adelson, 1995]