Data Visualization (CSCI 627/490)

Marks and Channels

Dr. David Koop
D3 Examples

• Observable Notebook

• Three Bar Charts:
  - Similar Solution
  - With Axes and Scales
  - With Objects and Margin Convention

• More on Margin Convention:
Toward Reusable Charts

- D3 does not provide "standard" charts
- E.g. there is no barchart method
- What is a standard chart?
  - "Should you expose the underlying scales and axes, or encapsulate them with chart-specific representations?"
  - "Should your chart support interaction and animation automatically?"
  - "Should the user be able to reach into your chart and tweak some aspect of its behavior?"
Assignment 3

- Upcoming
- Same visualization
- Different tools
### 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>228721088</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>82338108</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

D. Koop, CSCI 627/490, Fall 2022
Another Solution

Size: Population, total

[Gapminder, Wealth & Health of Nations]
What about change over years?
Another Solution showing trends over time

Income per person (GDP/capita, PPP$ inflation-adjusted)

United States
Russia
Nigeria
China

DATA DOUBTS

World Regions

Search...

World Regions

United States
Russia
Nigeria
China

Reset

Another Solution showing trends over time

Gapminder, Wealth & Health of Nations

D. Koop, CSCI 627/490, Fall 2022
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
## Bertin’s Original Visual Variables

<table>
<thead>
<tr>
<th>Visual Variable</th>
<th>Description</th>
<th>Examples</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Position</strong></td>
<td>changes in the x, y location</td>
<td><img src="image1" alt="Position Examples" /></td>
</tr>
<tr>
<td><strong>Size</strong></td>
<td>change in length, area or repetition</td>
<td><img src="image2" alt="Size Examples" /></td>
</tr>
<tr>
<td><strong>Shape</strong></td>
<td>infinite number of shapes</td>
<td><img src="image3" alt="Shape Examples" /></td>
</tr>
<tr>
<td><strong>Value</strong></td>
<td>changes from light to dark</td>
<td><img src="image4" alt="Value Examples" /></td>
</tr>
<tr>
<td><strong>Colour</strong></td>
<td>changes in hue at a given value</td>
<td><img src="image5" alt="Colour Examples" /></td>
</tr>
<tr>
<td><strong>Orientation</strong></td>
<td>changes in alignment</td>
<td><img src="image6" alt="Orientation Examples" /></td>
</tr>
<tr>
<td><strong>Texture</strong></td>
<td>variation in ‘grain’</td>
<td><img src="image7" alt="Texture Examples" /></td>
</tr>
</tbody>
</table>
Visual Channels

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

- **Color**
  - Black
  - Red
  - Green

- **Shape**
  - Triangle
  - Star
  - Line
  - Square

- **Tilt**
  - Slanted

- **Size**
  - Length
  - Area

- **Volume**
  - Small
  - Medium
  - Large

[Munzner (ill. Maguire), 2014]
### Table of Visual Attributes

**Visual Attributes Survey**

<table>
<thead>
<tr>
<th>Transform</th>
<th>Position</th>
<th>Length</th>
<th>Size (Area)</th>
<th>Orientation</th>
<th>Volume</th>
<th>Shape</th>
<th>Angle</th>
<th>Curvature</th>
<th>Mark</th>
<th>Line Ending</th>
<th>Closure</th>
<th>Local Warp</th>
<th>Edge Type</th>
<th>Corner Type</th>
<th>Icon, glyph, etc</th>
<th>Colour</th>
<th>Brightness</th>
<th>Hue</th>
<th>Saturation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Berlin 1967</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>1,2</td>
<td>3</td>
<td>4</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Cleveland 1985</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>X</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mackinlay 1986</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>X</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MacEachren 1995</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Wilkinson 1999</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ware 2000</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Maaza 2009</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Illinoisy 2012</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Chen, Floridi 2013</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Information Visualization Researchers</th>
<th>Vision Rsch</th>
<th>Shape Rsch</th>
</tr>
</thead>
<tbody>
<tr>
<td>Preattentive Perception</td>
<td>Brath 2009/2011</td>
<td></td>
</tr>
</tbody>
</table>
More Visual Attributes

<table>
<thead>
<tr>
<th>Table of Visual Attributes</th>
<th>Information Visualization Researchers</th>
<th>Vision Rsch</th>
<th>Shape Rsch</th>
</tr>
</thead>
<tbody>
<tr>
<td>Texture</td>
<td>Granularity</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td></td>
<td>Pattern</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td></td>
<td>Orientation</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Relation</td>
<td>Connection</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td></td>
<td>Containment</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Optics</td>
<td>Blur</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td></td>
<td>Transparency</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td></td>
<td>Stereo Depth</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td></td>
<td>Concavity</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td></td>
<td>Light Direction</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td></td>
<td>Shadow</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td></td>
<td>Partial occlusion</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Movement</td>
<td>Flicker</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td></td>
<td>Speed</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td></td>
<td>Direction</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Misc</td>
<td>Numerosity</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td></td>
<td>Spatial Grouping</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td></td>
<td>Arrangement</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td></td>
<td>Resolution</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td></td>
<td>Artistic Effects</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td></td>
<td>Text Labels</td>
<td>X</td>
<td>X</td>
</tr>
</tbody>
</table>

[R. Brath]
Channels

• Usually map an attribute to a single channel
  - Could use multiple channels but…
  - **Limited** number of channels

• Restrictions on size and shape
  - Points are nothing but location so size and shape are ok
  - Lines have a length, cannot easily encode attribute as length
  - Maps with boundaries have area, changing size can be problematic
Cartograms

[Election Results by Population, M. Newman, 2012]
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]
Mark Types

- Can have marks for items and **links**
  - Connection => pairwise relationship
  - Containment => hierarchical relationship

**Marks as Items/Nodes**
- Points
  ![Points Diagram]
- Lines
  ![Lines Diagram]
- Areas
  ![Areas Diagram]

**Marks as Links**
- Containment
  ![Containment Diagram]
- Connection
  ![Connection Diagram]

[Munzner (ill. Maguire), 2014]
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]
<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>yes</td>
<td>infinite</td>
<td>Good</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>length</td>
<td>yes</td>
<td>many</td>
<td>Good</td>
<td>Good</td>
<td>Good</td>
<td>Good</td>
<td>Good</td>
</tr>
<tr>
<td>size, area</td>
<td>yes</td>
<td>many</td>
<td>Good</td>
<td>Good</td>
<td>Good</td>
<td>Good</td>
<td></td>
</tr>
<tr>
<td>angle</td>
<td>yes</td>
<td>medium/few</td>
<td>Good</td>
<td>Good</td>
<td>Good</td>
<td>Good</td>
<td></td>
</tr>
<tr>
<td>pattern density</td>
<td>yes</td>
<td>few</td>
<td>Good</td>
<td>Good</td>
<td>Good</td>
<td>Good</td>
<td></td>
</tr>
<tr>
<td>weight, boldness</td>
<td>yes</td>
<td>few</td>
<td>Good</td>
<td>Good</td>
<td>Good</td>
<td>Good</td>
<td></td>
</tr>
<tr>
<td>saturation, brightness</td>
<td>yes</td>
<td>few</td>
<td>Good</td>
<td>Good</td>
<td>Good</td>
<td>Good</td>
<td></td>
</tr>
<tr>
<td>color</td>
<td>no</td>
<td>few (&lt; 20)</td>
<td>Good</td>
<td>Good</td>
<td></td>
<td>Good</td>
<td></td>
</tr>
<tr>
<td>shape, icon</td>
<td>no</td>
<td>medium</td>
<td>Good</td>
<td>Good</td>
<td>Good</td>
<td>Good</td>
<td></td>
</tr>
<tr>
<td>pattern texture</td>
<td>no</td>
<td>medium</td>
<td>Good</td>
<td>Good</td>
<td>Good</td>
<td>Good</td>
<td></td>
</tr>
<tr>
<td>enclosure, connection</td>
<td>no</td>
<td>infinite</td>
<td>Good</td>
<td>Good</td>
<td>Good</td>
<td>Good</td>
<td></td>
</tr>
<tr>
<td>line pattern</td>
<td>no</td>
<td>few</td>
<td>Good</td>
<td>Good</td>
<td>Good</td>
<td>Good</td>
<td></td>
</tr>
<tr>
<td>line endings</td>
<td>no</td>
<td>few</td>
<td>Good</td>
<td>Good</td>
<td>Good</td>
<td>Good</td>
<td></td>
</tr>
<tr>
<td>line weight</td>
<td>yes</td>
<td>few</td>
<td>Good</td>
<td>Good</td>
<td>Good</td>
<td>Good</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 ~5.6x longer than Right

![Bar chart showing percentage difference in length between elements A and B]

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

![Test % difference in length between elements](image)

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

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

[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

[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

[Heer & Bostock, 2010]
Cleveland & McGill Experiments

Figure 4. Graphs from position–length experiment.

Figure 3. Graphs from position–angle experiment.
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

[Muñozn (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

Steven’s Psychophysical Power Law: $S = I^n$
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

[Munzner (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

![Position + Hue (Color)](image1)
![Size + Hue (Color)](image2)
![Width + Height](image3)
![Red + Green](image4)

- Fully separable
- Some interference
- Some/significant interference
- Major interference

[Munzner (ill. Maguire) based on Ware, 2014]
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
Visual Popout: Parallel Lines Require Search…

[Munzner (ill. Maguire), 2014]