Data Visualization (DSC 530/CIS 602-02)

Marks & Channels

Dr. David Koop
D3 Introduction

- Ogievetsky has put together a nice set of interactive examples that show off the major features of D3
  - (Updated from original for D3 v4)
- Other references:
  - Murray’s book on Interactive Data Visualization for the Web
  - The D3 website: d3js.org
  - Ros's Slides on v4: https://iros.github.io/d3-v4-whats-new/
D3 Data Joins

- Two groups: data and visual elements
- Three parts of the join between them: enter, update, and exit
- enter: `s.enter()`, update: `s`, exit: `s.exit()`
D3 v4 vs. v3

• v4 breaks a lot of v3 code…
• v4 is more modular, can build libraries that include only the parts you care about
  - Why worry about this?
• Result is that there is a flat namespace now
  - d3.scale.linear => d3.scaleLinear
• More important change: selections are **immutable** now
  - Used to be that enter() modified the selection to include any appended items
  - Use **merge** to explicitly merge the enter and update selections
    - s.enter().append("rect")
    .merge(s)
    ...

D3 v3 Selections

```javascript
var circleBinding = svg.selectAll("circle").data(data);

circleBinding.style("fill", "blue"); // UPDATE

circleBinding.enter()
  .append("circle") // ENTER; modifies UPDATE!
    .style("fill", "green");

circleBinding // ENTER + UPDATE
  .style("stroke", "black");
```
D3 v4 Selections

```javascript
var circleBinding = svg.selectAll("circle").data(data);

circleBinding.style("fill", "blue"); // UPDATE

circleBinding.enter()
  .append("circle") // ENTER; modifies UPDATE!
    .style("fill", "green");

  .merge(circleBinding) // ENTER + UPDATE
    .style("stroke", "black");
```

Merge

• Merge creates a new selection that includes the items from both selections

• If you want to update all elements (including those just added via enter), use merge!
Transitions

- Nested transitions (those that "hang off" of a parent transition) follow immediately after the parent transition
- In v3, they had to be delayed accordingly
Assignment 2

• Use D3

1. Repeat Part 3b of A1 using D3
2. Extend Part 1 to create a **stacked** bar chart
3. Create a line chart that shows a region's numbers that is linked to a dropdown menu allowing you to select the region. Use transitions!
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?"

[Towards Reusable Charts, M. Bostock, 2012]
### 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>80.08</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]
Another 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]
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
- Lines
- Areas

**Marks as Links**

- Containment
- Connection

[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?
  - How was this determined?
Test % difference in length between elements

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

Answer: Left is ~5.6x longer than Right

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

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

![Bar chart](image)

[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

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

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

![Diagram of area judgment stimuli](image_url)

[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. [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]
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: $S = I^n$

[D. Koop, DSC 530, Spring 2017]

[Munzner (ill. Maguire), 2014]
Ranking Channels by Effectiveness

**Channels:** Expressiveness Types and Effectiveness Ranks

**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

[Ranking Channels by Effectiveness - Munzner (ill. Maguire), 2014]
Results Summary

[Cleveland & McGill's Results]

[Crowdsourced Results]

[Log Error]

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