Data Visualization (CIS 468)

D3 + Marks & Channels

Dr. David Koop
Tasks

What? Why? How?

Actions

- **Analyze**
  - Consume
    - Discover
  - Present
  - Enjoy

- **Produce**
  - Annotate
  - Record
  - Derive

Search

<table>
<thead>
<tr>
<th>Target known</th>
<th>Target unknown</th>
</tr>
</thead>
<tbody>
<tr>
<td>Location known</td>
<td>Lookup</td>
</tr>
<tr>
<td>Location unknown</td>
<td>Locate</td>
</tr>
</tbody>
</table>

Query

- Identify
- Compare
- Summarize

Targets

- **All Data**
  - Trends
  - Outliers
  - Features

- **Attributes**
  - One
    - Distribution
  - Many
    - Dependency
    - Correlation
    - Similarity

- **Network Data**
  - Topology
    - Paths

- **Spatial Data**
  - Shape

[Munzner (ill. Maguire), 2014]
Analyze: Discovery
Analyze: Present

Each solid circle represents a bee species active in Carlinville, Ill., in both the late 1800s and 2010.

Hatching represents a bee species active in the 1800s but now locally extinct.

The spot where each block rests on the circle indicates one of 26 plant species frequented by these bees.

In the 1880s scientists observed the following about the bee-plant encounters:

- Present
- Frequent
- Abundant

Studies in 2009 and 2010 showed many bee-plant interactions had changed:

- Lost
- Persisted
- New

[M. Stefaner, 2013]
Analyze: Enjoy

NameVoyager: Explore baby names and name trends letter by letter
Looking for the perfect baby name? Sign up for free to receive access to our expert tools!

Baby Name: An |

Names starting with 'AN' per million babies

Annotate

[S. Lu, 2017]
Derive

Original Data

Derived Data

\[ \text{trade balance} = \text{exports} - \text{imports} \]
What does a user know?

- **Lookup**: check bearings
- **Locate**: find on a map
- **Browse**: what’s nearby
- **Explore**: where to go (patterns)

[Munzner (ill. Maguire), 2014]
Query

- Identify
- Compare
- Summarize

- Number of targets: One, Some (Often 2), or All
- Identify: characteristics or references
- Compare: similarities and differences
- Summarize: overview of everything

[Munzner (ill. Maguire), 2014]
Quiz

- Tuesday, October 2
- Only **part** of class
- Format:
  - Multiple Choice
  - Short Answer (multi-part)
- Material:
  - Ch. 1-3, 5
  - HTML, CSS, SVG, JavaScript, D3
- Examples:
  - Given dataset, which attributes are categorical, ordered, quantitative?
  - What does the following expression evaluate to?
    
    ```javascript
    [[1,2],[5,2]].map(d => d.reduce((s,t) => s+t))
    ```
D3.js is a JavaScript library for manipulating documents based on data. D3 helps you bring data to life using HTML, SVG, and CSS. D3’s emphasis on web standards gives you the full capabilities of modern browsers without tying yourself to a proprietary framework, combining powerful visualization
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()`
More D3

• Original Slide Format:
  - https://dakoop.github.io/IntroD3/

• Observable Notebook:
  - https://beta.observablehq.com/@dakoop/d3-intro
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)
      .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");
```
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
Visual Encoding
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]

D. Koop, CIS 468, Fall 2018
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)

[Gapminder, Wealth & Health of Nations]

D. Koop, CIS 468, Fall 2018
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]