PBO Workshop

Getting Excited About Data-Driven Documents With d3

Ben Racine 1

¹Cornerstone Systems NW

November 2, 2011



Outline

1 If you are eager to get these materials

2 Background

Outline

If you are eager to get these materials

Background

- Navigate to github.com
- Search for benracine
- This repo should be the first hit, i.e. "d3_cisnet_tutorial"

- Navigate to github.com
- Search for benracine
- This repo should be the first hit, i.e. "d3_cisnet_tutorial"

- Navigate to github.com
- Search for benracine
- This repo should be the first hit, i.e. "d3_cisnet_tutorial"

- Navigate to github.com
- Search for benracine
- This repo should be the first hit, i.e. "d3_cisnet_tutorial"

Background

- d3.js is a small, free JavaScript library for manipulating documents based on data.
- d3 allows you to bind arbitrary data to a Document Object Model (DOM), and then apply data-driven transformations to the document.
- For example:
 - You can use d3 to generate a basic HTML table from a matrix of numbers.
 - You could use the same data to create an interactive SVG bar chart with smooth transitions and interaction.



Ecosystem

- It largely targets the SVG element
- SVG is being increasingly supported:
 - IE, 9+
 - Chrome
 - Firefox
 - Safari
 - Opera, 9.5+
 - iOS
 - Android, 3.0+
- A few additional JavaScript libraries can be useful for minor data transformations and persistence
 - Underscore.js
 - Backbone.js
 - JSONSelect.js
 - Science.js
- JavaScript performance has recently increased considerably

Background (cont.)

- d3.js is loosely associated with the Stanford Visualization Group
- Supersedes the ProtoVis project
- Mike Bostock is the primary author, but it is open source on Github.com
- Only a year or two old

The Author's Tour

A Tour

- The author uses "layouts"
- Association / adjacency representations are addressed with chord layout
 - Chord produce a chord diagram from a matrix of relationships.
- A whole collection of hierarchical layouts
 - Bundle apply Holten's hierarchical bundling algorithm to edges.
 - Cluster cluster entities into a dendrogram.
 - Hierarchy derive a custom hierarchical layout implementation.
 - Histogram compute the distribution of data using quantized bins.
 - Pack produce a hierarchical layout using recursive circle-packing.
 - Partition recursively partition a node tree into a sunburst or icicle

The Author's Tour Continued

- Calendar
- Time-series (note Mike's other project: "Cube")
- Pan and zoom
- Scale elements (with tick, label, title and location options)
- Smooth transitions
- Interaction
- Animation
- Chloropleth / projections
- Force (helpful in solving the non-collision problem)

Repository Examples

- Bar / column
- Box plot quintile distribution of a single variable
- Bullet considered a best-practices replacement for gauge charts
- Histogram compute the distribution of data using quantized bins.
- Hyperbolic tree
- Pie compute the start and end angles for arcs in a pie or donut chart.
- QQ plots compare two probability distributions by graphing their quantiles against each other
- Radial -
- Stack -
- Streamgraph a generalization of stacked area graphs

My Tour

- A side project that I consulted on for the Startup Weekend
- ** My navpane task

Why d3?

- It is a small and sharp tool
- Kind of in the Unix philosophy, it does "one" thing and does it well
 - i.e. It solves the cruxes of visualization
- Plays well with others
 - i.e. doesn't pollute the global namespace

The Cruxes of Visualization

Scales

- Data access by url
 - text
 - json
 - xml
 - html
 - CSV
- Scales define the mapping between the data and pixel/color space
- A reverse function can be used to map the other direction too
- Linear, log, power, discrete/continuous, banded, ordinal and user defined

Data Binding

Reuse of Existing Standards

- Takes advantage of existing w3 standards instead of reinventing the wheel
- Authors can take their new depth of knowledge of modern web standards with them