InfoVis Systems & Toolkits

CS 7450 - Information Visualization
September 21, 2016
John Stasko

Learning Objectives

- Gain familiarity with history of visualization toolkits
  - Describe what each's new contribution was
- Understand approaches taken by systems seeking to support visualization creation without programming
- Explain what Many Eyes was, what it provided, and what its contribution was
- Describe a spectrum of approaches for creating visualizations (ranging from automatic creation given data to low-level graphics libraries) and identify representative systems that occupy different places along that spectrum

1

Background

- In previous classes, we have examined different techniques for presenting multivariate data
    - We'll continue to show more later too

- Today we look at systems that implement these ideas and provide some of their own new visualization techniques

Agenda

- Toolkits that can be used to build systems
    - D3, Processing, …
- Systems providing a view or views
    - Improvise, Many Eyes, Polaris, …
- Commercial systems
    - Spotfire, InfoZoom, Tableau, QlikView…

Toolkits & Infrastructures

- Set of components or capabilities that allow others to put together visualization systems

- Growing trend

Toolkit Design

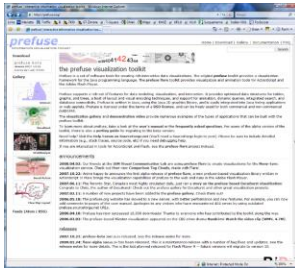- What would you include in a toolkit like this if you designed it?

Heer++ Series

Series of toolkits from Jeff Heer and his research group

Prefuse

Java2D

Heer et al, CHI '05

Heer++ Series

Series of toolkits from Jeff Heer and his research group

Prefuse

Java2D

Heer et al, CHI '05

What happened???

Web!!!

Heer++ Series

Series of toolkits from Jeff Heer and his research group

Prefuse	Flare	Protovis
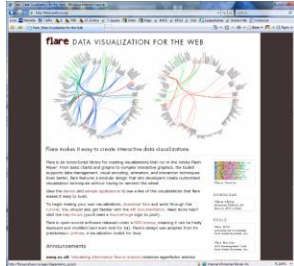
Java2D   ActionScript & Flash   Declarative spec.

Heer et al, CHI '05

Bostock & Heer, *TVCG* (InfoVis) '09

Heer & Bostock, *TVCG* (InfoVis) '10

D3: Data-Driven Documents

- Newest entry in the Heer-Bostock line of toolkits
- "Not just an infovis toolkit"
- Javascript-based
- Very similar to Protovis...
  - Except makes use explicitly of web standards such as Scalable Vector Graphics (SVG) rather than a proprietary "marks" graphics set

D3 Design Pattern

- Declarative Syntax like Protovis
- Creating/Modifying selections of the HTML DOM
- "An elegant for-loop with a bunch of useful helper functions"
- Excellent support for changing data
  – Taking advantage of CSS3 Transformations and Transitions
- Integrates seamlessly into any webpage

D3 Website
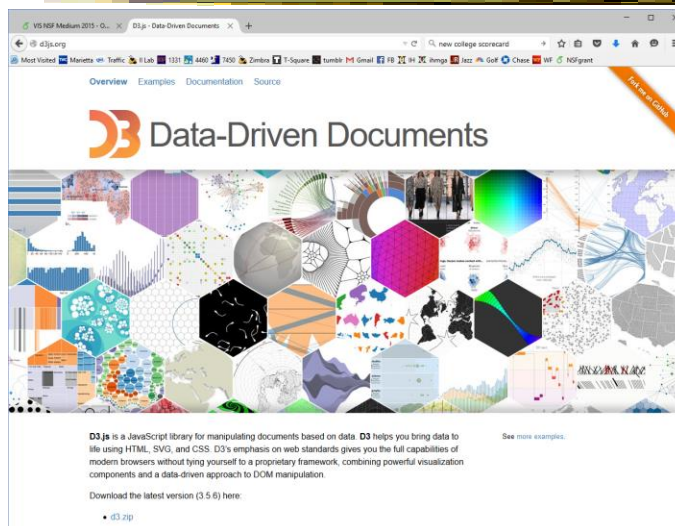
D3 Support

- Active community online
  - https://github.com/mbostock/d3/wiki
  - Including Mike Bostock often answering questions

- Tutorial from John in our last class

https://trifacta.github.io/vega/

Vega
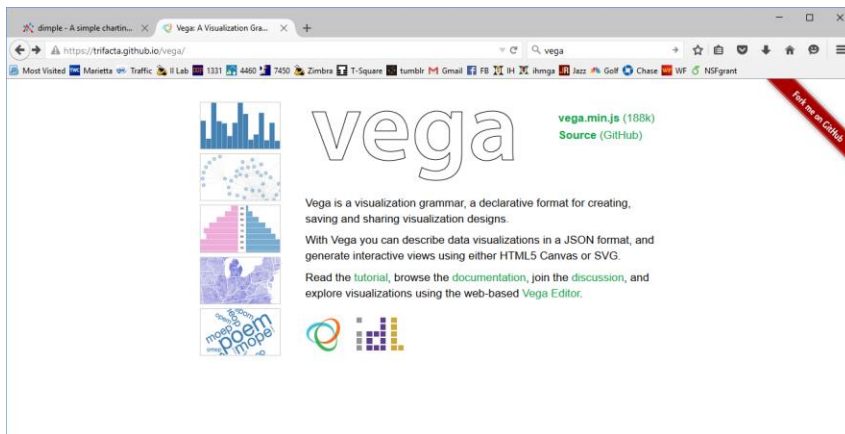
Characteristics

- Declarative visualization grammar like D3
- No CSS, DOM, etc., though
- Describe visualizations in JSON format
- Generate interactive views via HTML5 Canvas or SVG

Example Code
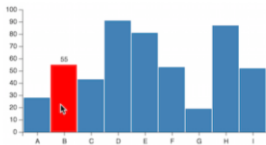

```json
{
  "width": 400,
  "height": 200,
  "padding": {"top": 10, "left": 30, "bottom": 20, "right": 10},

  "data": [
    {
      "name": "table",
      "values": [
        {"category":"A", "amount":28},
        {"category":"B", "amount":55},
        {"category":"C", "amount":43},
        {"category":"D", "amount":91},
        {"category":"E", "amount":81},
        {"category":"F", "amount":53},
        {"category":"G", "amount":19},
        {"category":"H", "amount":87},
        {"category":"I", "amount":52}
      ]
    }
  ],

  "signals": [
    {
      "name": "tooltip",
      "init": {},
      "streams": [
        {"type": "rect:mouseover", "expr": "datum"},
        {"type": "rect:mouseout", "expr": "{}"}
      ]
    }
  ],                      and more...
```

8

Premise

- Not really for humans to write
- Intended for tools to generate & exchange it
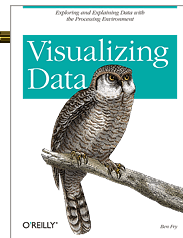
Processing

- Java based
- Unlike protovis & D3, not specifically designed for InfoVis
  - Data Reader? Layout algorithm?
  - But can definitely be used to build visualizations!
- Well documented, lots of tutorials with contributions from many people and even books

Ben Fry

9

Processing: the idea
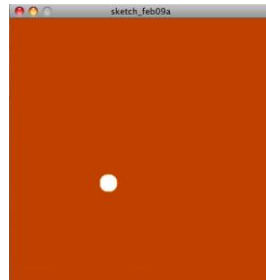
- Programming as scripting
  - PDE: processing development environment
  - A program is called a *sketch*
  - written as a list of statements

Processing: Useful Functions

```
void setup() {
    //your own code here
}

void draw() {
    //your own code here
}
```

- These are built-in functions that are called automatically.
  - The setup() block runs once.
  - The draw() block runs repeatedly: good for animation

P5.js

- Interpretation of Processing in Javascript

Piccolo

11
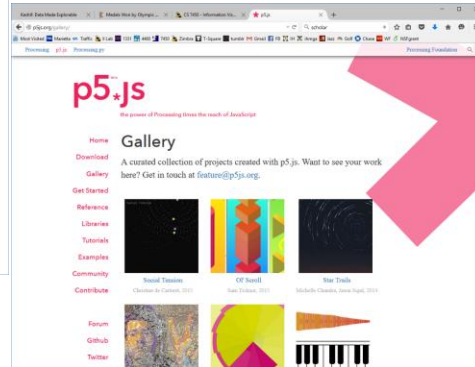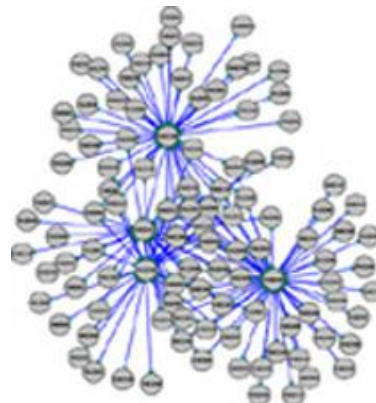
Characteristics

- Graphics toolkit with very nice built-in zooming and panning support
- Useful for implementing infovis too
- Will discuss more later in course…

https://developer.mozilla.org/en-US/docs/Web/API/WebGL_API

WebGL

When you need 3D or really sophisticated graphics

More toolkits

(Which do you know?)
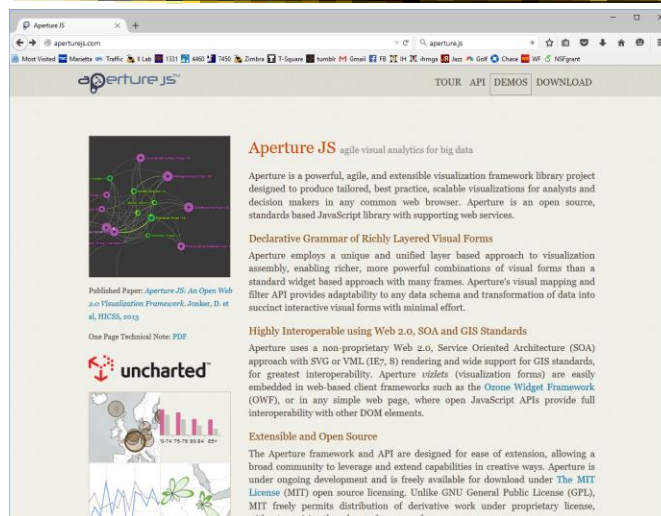
Aperture.js

`http://aperturejs.com/`

13

Bokeh
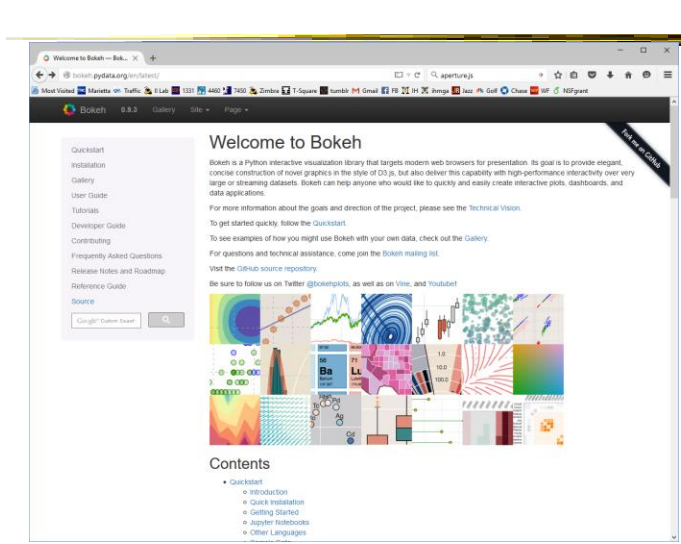
Python

HighCharts
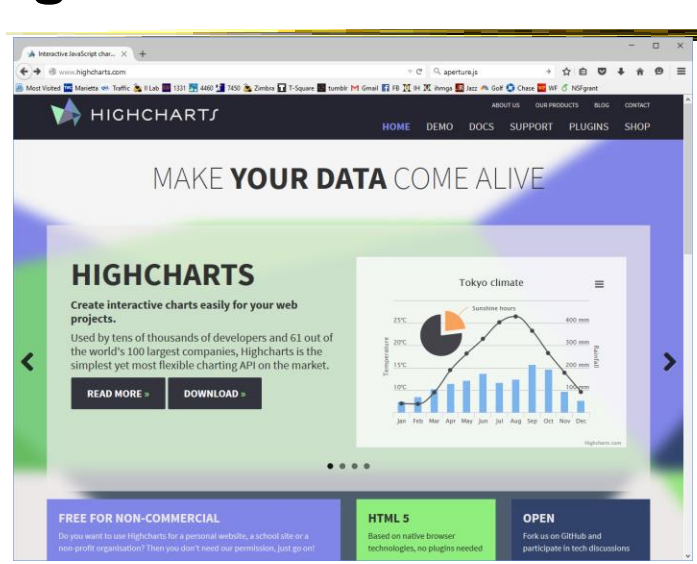
dimple
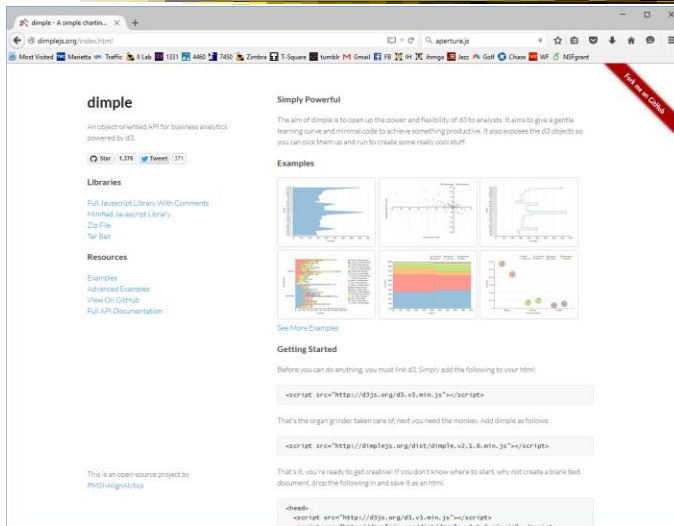
On top of D3

Fall 2016 CS 7450 29

JavaScript InfoVis Toolkit

http://philogb.github.io/jit/

Fall 2016 CS 7450 30

15

react-vis

https://www.npmjs.com/package/react-vis

http://code.google.com/apis/visualization/documentation/
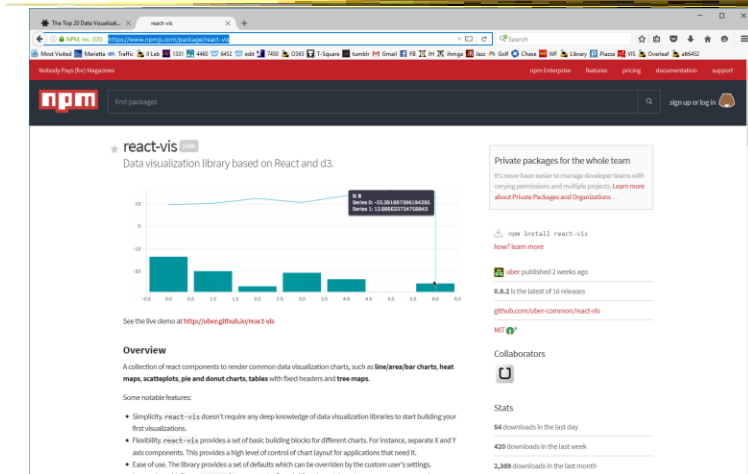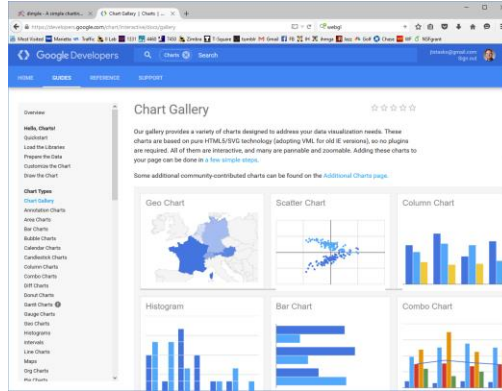
Google Chart Tools
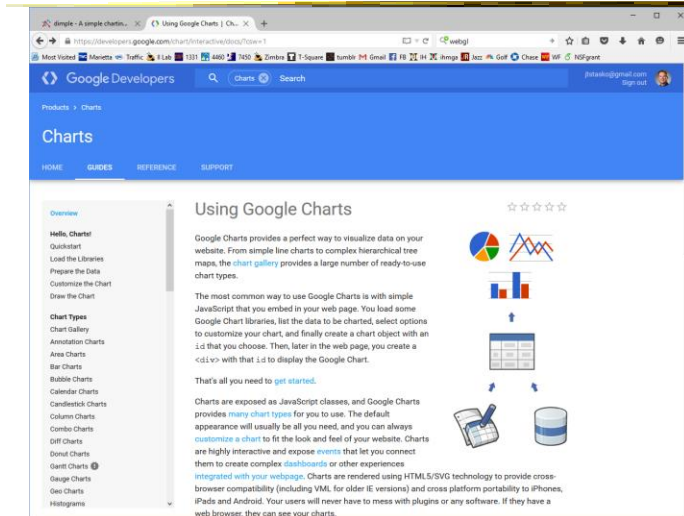
Characteristics

- Javascript-based
- Gallery of contributed code segments
- Visualizations are interactive
- Evolving API

Reflection

- What would you seek in a good infovis system or toolkit?

Toolkit Tradeoffs

- InfoVis-focused
  - Many fundamental techniques built-in
  - Can be faster to get something going
  - Often more difficult to implement something "different"
  - Documentation?

- Generic graphics
  - More flexible
  - Can customize better
  - Big learning curve
  - Doc is often better
  - Can take a long time to (re)implement basic techniques

Writing Code is Hard

- Why not just show what you want the visualization to look like?
  - What's the challenge?

Lyra

- Interactive vis builder tool without needing to program
- Graphical "marks" are bound to data fields
- User shows what vis is to look like, the mapping from data
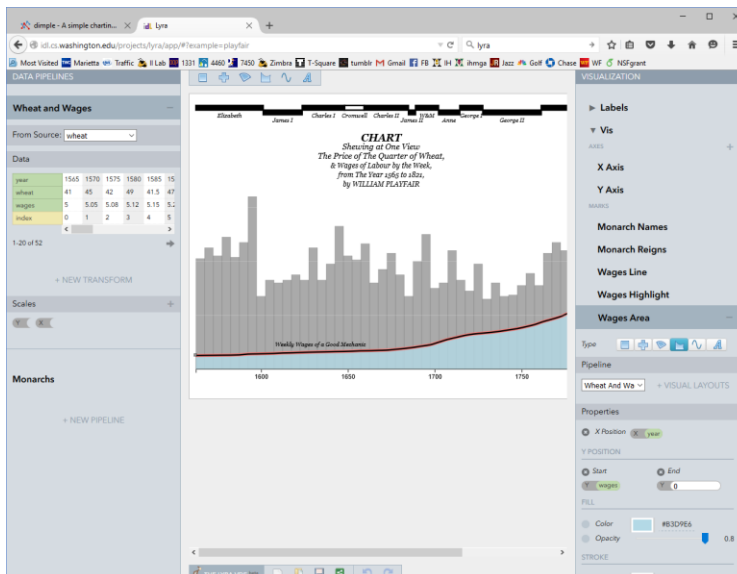- Generates code (Vega) that can be run on the web

Satyanarayan & Heer
Computer Graphics Forum (EuroVis) '14
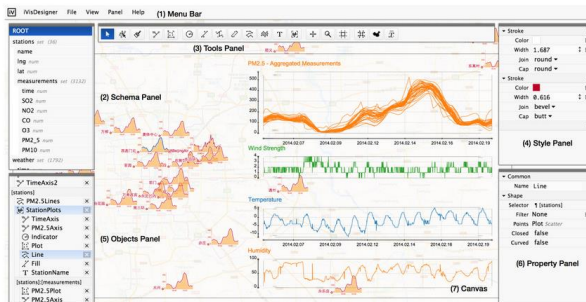
http://idl.cs.washington.edu/projects/lyra/

Video

iVisDesigner

- Interactively create mappings from data elements to graphical elements
- Has scatterplot, timeline, graph templates

Video

Ren, Hollerer, & Yuan
TVCG (InfoVIs) '14

Journalism-driven systems

Datawrapper
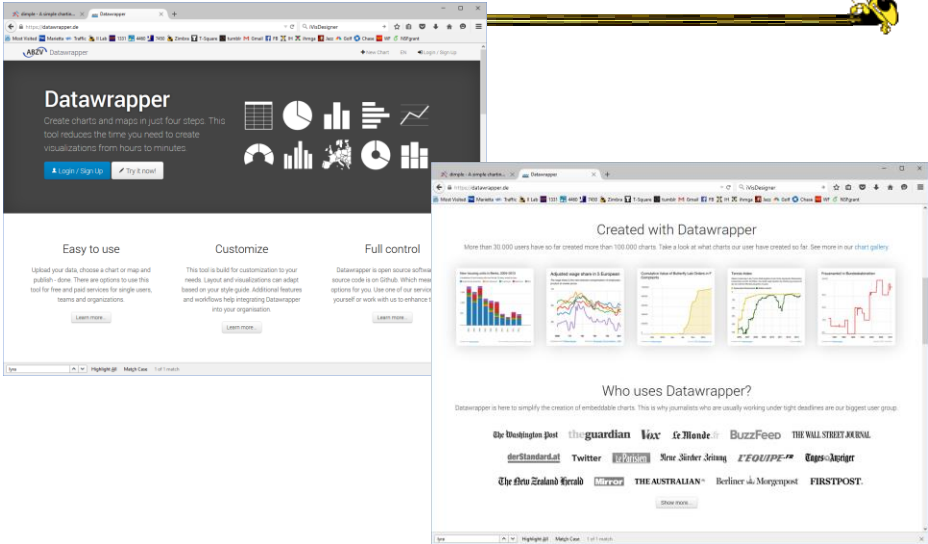
https://datawrapper.de/

Chartbuilder

http://quartz.github.io/Chartbuilder/
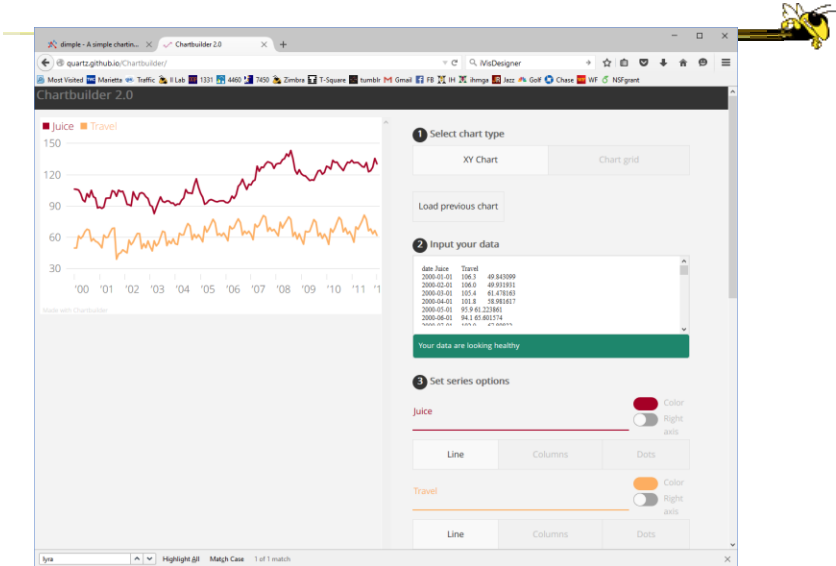
Systems/Tools

- Academic/research systems that provide preexisting views
- Commercial tools with suite of well-known visualizations

Polaris

http://www.graphics.stanford.edu/projects//polaris/

http://www.tableausoftware.com

Video

Stolte et al
TVCG `02

22

Basis

- Relational databases
- Pivot tables from spreadsheets
- N-dimensional data cubes

- Analytic approach is fundamental
- Provides visual representations of these concepts

Visualization

- Table of data (rows, columns)
- Each axis may have nested dimensions
- Table entry is a pane, and has visual marks to represent data

- Analyst drags and drops fields from database schema onto shelves of display
- Much interaction supported

Graphics

- Formal table algebra provided to describe data
- Visual mappings established from data types to appropriate (good) markings and encodings

Impact: Became Tableau

Many Eyes

Viégas et al
TVCG (InfoVis) '07

- InfoVis on the web
- Website developed from IBM's infovis group
- Motivating infovis challenges:
  - Difficulty of creating new visualizations
  - How do you discuss the visualizations?
- Project goals:
  - Enable end-user creation of visualizations
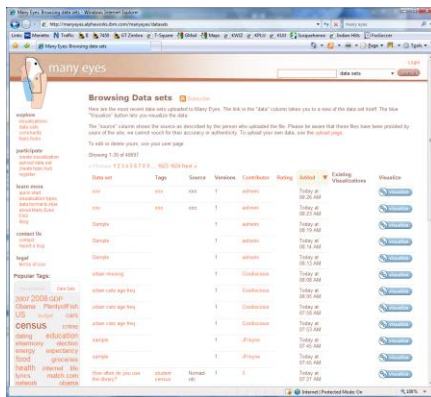 - Foster large-scale collaborative usage

Features

- Provides data upload capabilities and choice from library of visualizations
- Includes
  - Gallery of recently uploaded visualizations for browsing
  - Chosen highlighted visualizations
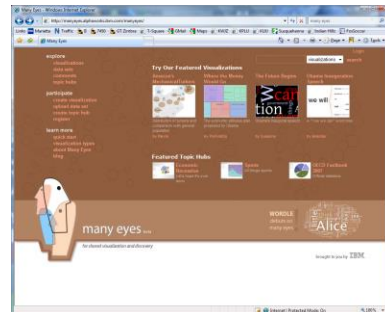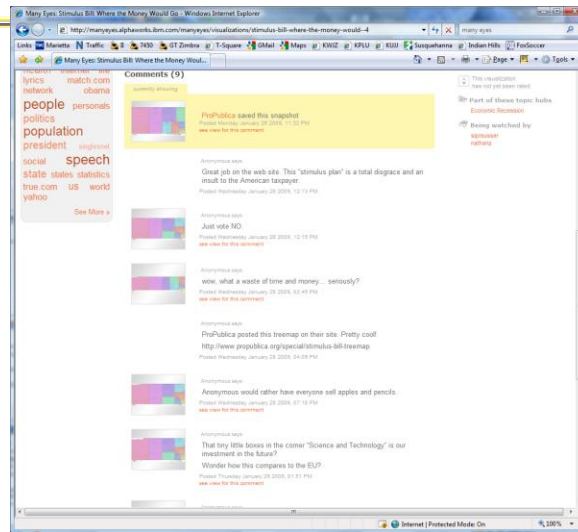 - Attached discussion forums for each vis

Recent gallery

Featured visualizations

Discussion
forums

Data

- Users upload their own data sets
  - All become public
- Format: table or unstructured text
  - Metadata allowed
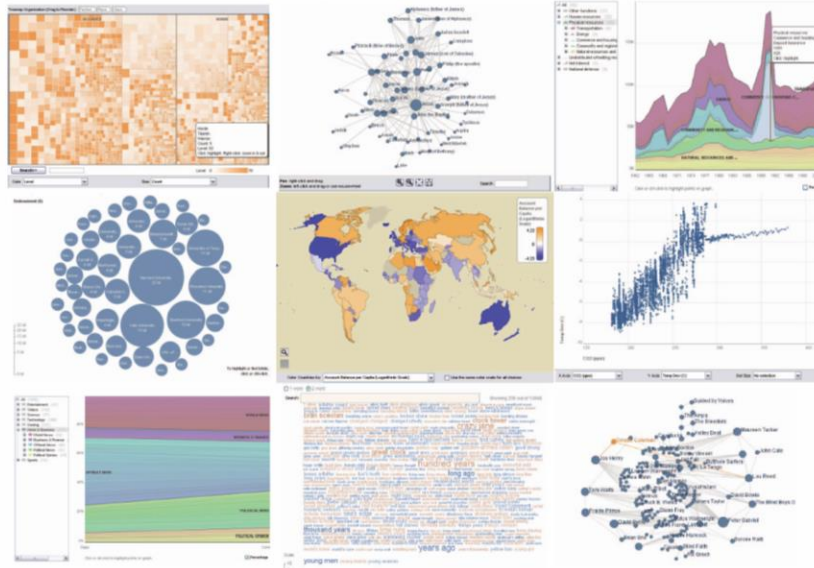- Immutable once uploaded

Visualizations

- Preloaded visualization types
  - Has grown over time
- User chooses one to combine with their data
- Provides named, typed slots that the user maps particular pieces of data to
  - System makes some reasonable guesses too

Originals

WordTree

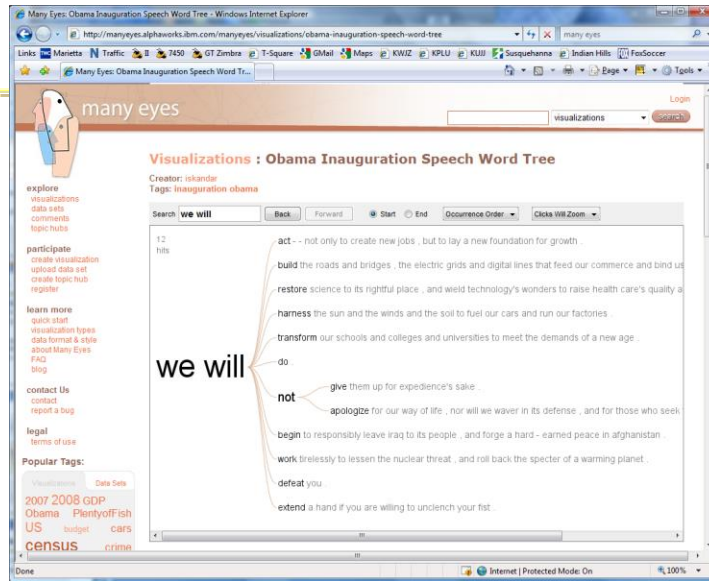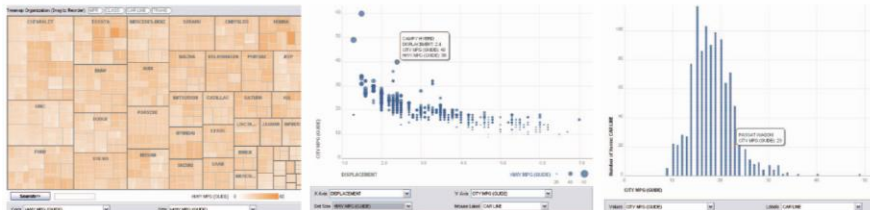Fig. 3. Three user generated visualizations offering different perspectives on the same dataset on car fuel economy. The grey areas on the top and bottom are automatically generated by the application and allow the user to browse through different dimensions in the data.

Allows the user to control the mapping from data to image

28

Social Aspects

- Users identified by login ID
- Can leave comments about different visualizations
- Can take snapshot of visualization state
  - Unique URL
- "Blog this" button

Evaluation

- Quantitative, objective
- 1895 posts of March '07
- Wide variety of topics of visualizations and motivations for creating visualizations
- Does seem to be fostering discussion

Viégas et al
HICSS '08

Evaluation

- Qualitative, subjective
- In-depth interviews with some ME users
- Visualizations used largely for communication and collaboration (not necessarily analysis)
  - Privacy and audience management a concern
- Highlights a number of interesting, non-expected uses of the technology

Danis et al
CHI '08

https://www.ibm.com/blogs/analytics-zone/six-years-of-many-eyes-a-personal-retrospective/

Status

- Shut down by IBM on June 12, 2015

Thoughts?

- What do you think of the design choices they made?

Sense.us

- Related, follow-up to Many Eyes
- Discussion and visualization of US Census data
- Go beyond Many Eyes in terms of annotation, collaboration, and discussion

Heer et al
CHI '07

Components

- Doubly-linked discussion
  - Can go from visualization to threaded discussion items or vice-versa
- Graphical annotation
  - Simple graphics editor and comments (like transparent layer)
- Bookmark trail
  - Small strip of views
- Comment listings and social navigation
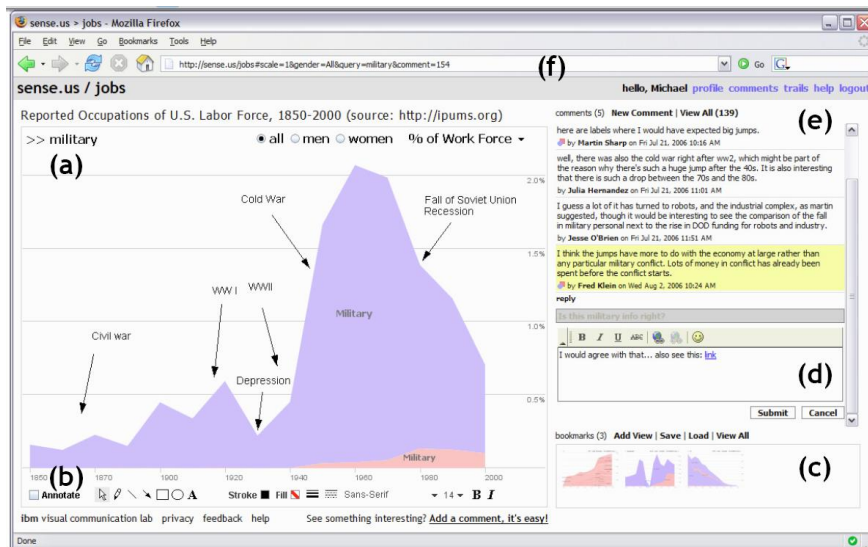 - Searchable, sortable indices and links

Figure 1. The sense.us collaborative visualization system. (a) An interactive visualization applet, with a graphical annotation for the currently selected comment. The visualization is a stacked time-series visualization of the U.S. labor force, broken down by gender. Here the percentage of the work force in military jobs is shown. **(b)** A set of graphical annotation tools. **(c)** A bookmark trail of saved views. **(d)** Text-entry field for adding comments. Bookmarks can be dragged onto the text field to add a link to that view in the comment. **(e)** Threaded comments attached to the current view. **(f)** URL for the current state of the application. The URL is updated automatically as the visualization state changes.

Figure 5. Scatterplot of U.S. states showing median household income (x-axis) vs. retail sales per capita (y-axis). New Hampshire and Delaware have the highest retail sales.
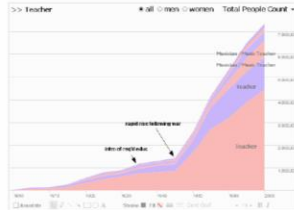
Figure 6. Visualization of the number of teachers. Annotations indicate the start of compulsory education and the rise of teachers in the post World War II era.
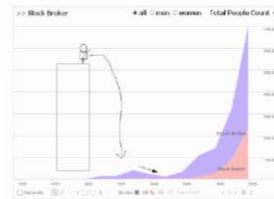
Sample annotations

Figure 7. Annotated view of stock brokers. The attached comment reads "Great depression 'killed' a lot of brokers".
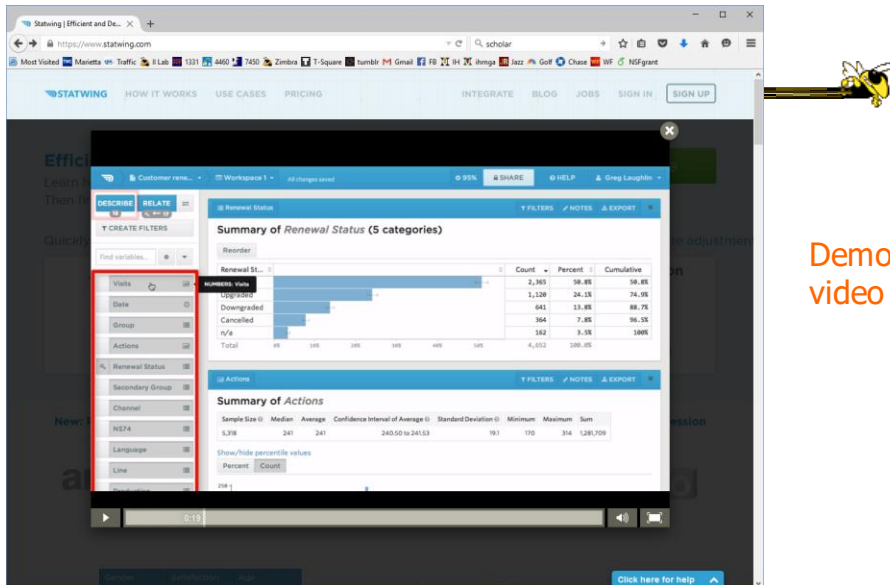

Statwing

https://www.statwing.com/

- Data analysis tool
- Upload spreadsheet or dataset, pick relationships you want to explore
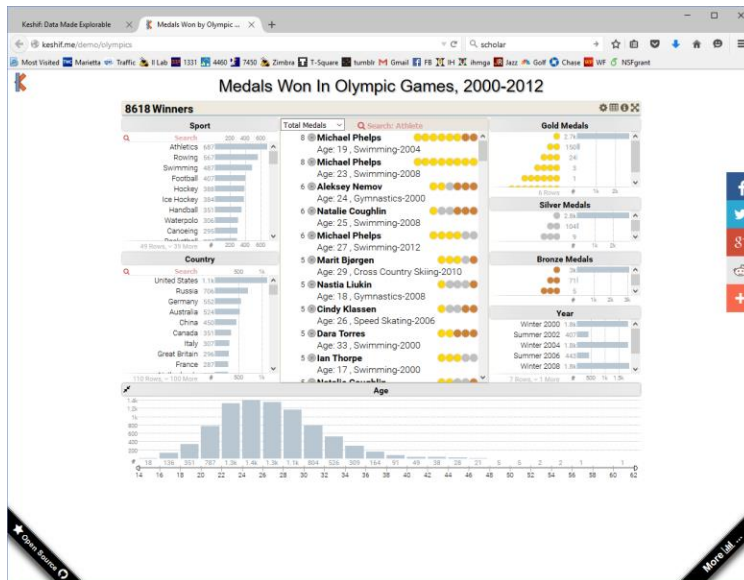- Basic graph types generated

33

Demo video

Keshif

http://keshif.me/

- Import data from spreadsheets and csv/text files
- Provides multiple coordinated bar charts and histograms
- Large set of examples on website

Demo

Commercial Systems

- Designed to handle wide variety of data types and sets
- Typically provide suite of well-known visualizations
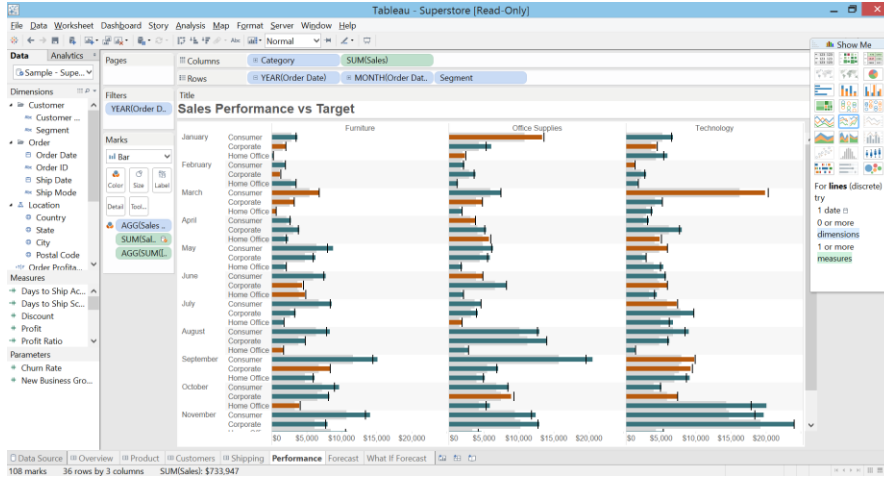
- (Preview of upcoming HW)

Tableau

Spotfire

36

Qlik Sense

www.qlik.com

Some web collections

http://www.visualisingdata.com/resources/

http://www.creativebloq.com/design-tools/data-visualization-712402
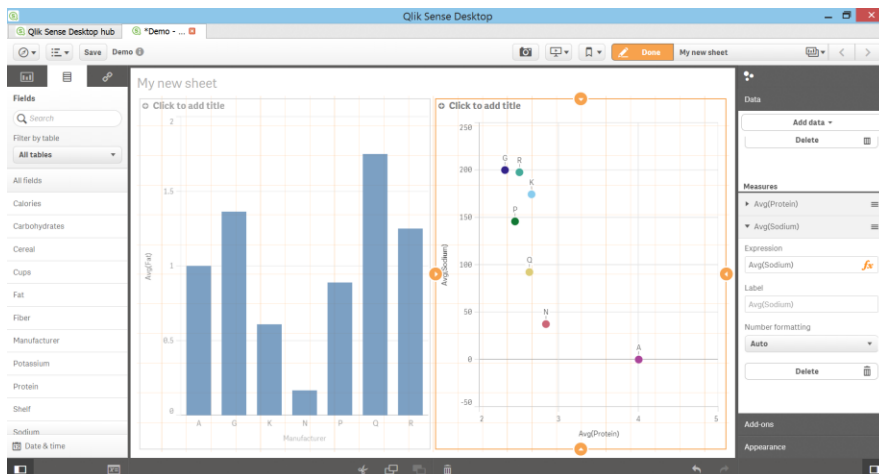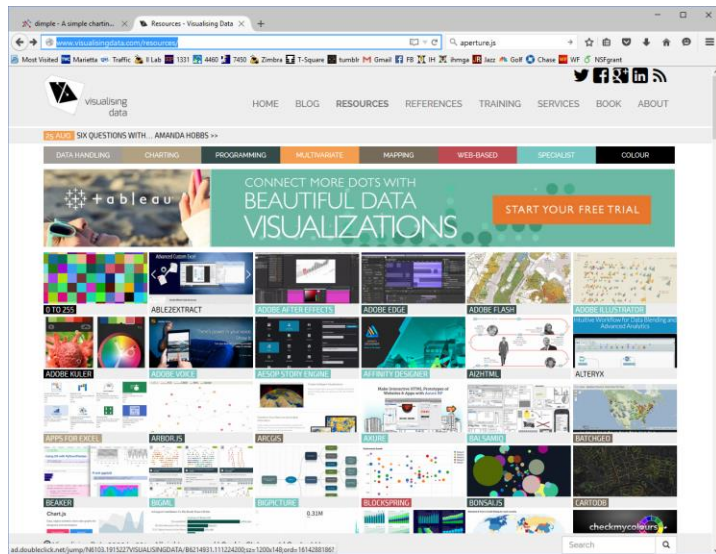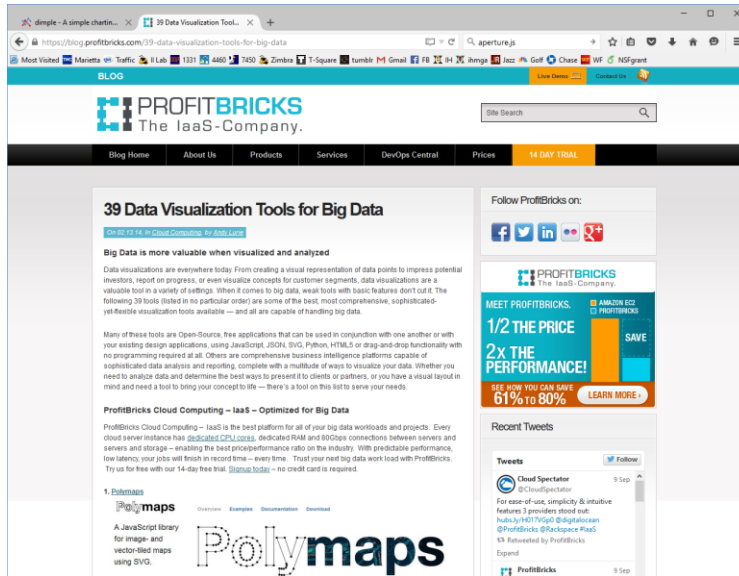
https://blog.profitbricks.com/39-data-visualization-tools-for-big-data

https://github.com/showcases/data-visualization
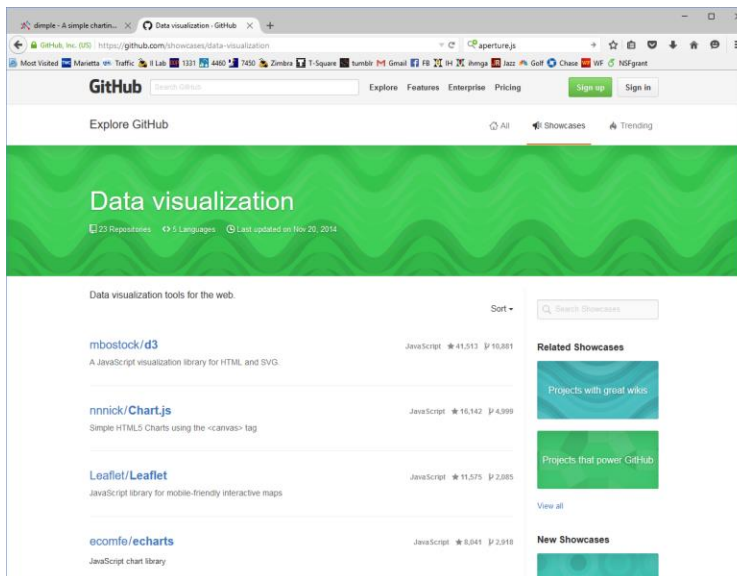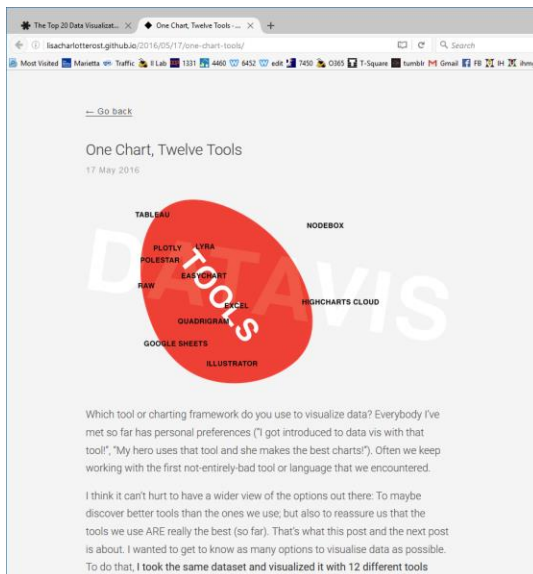
Neat comparison
of 12 different
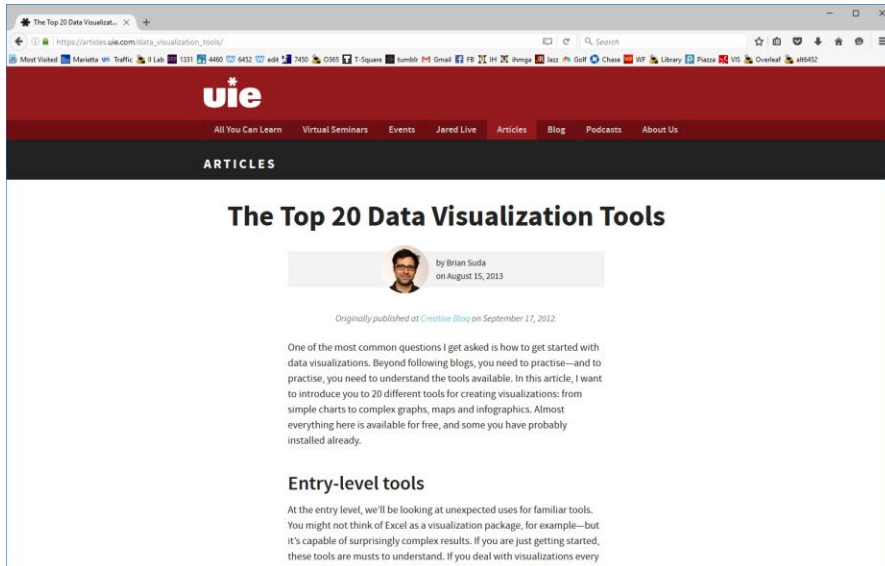tools using the
same data set

Learning Objectives

- Gain familiarity with history of visualization toolkits
  - Describe what each's new contribution was
- Understand approaches taken by systems seeking to support visualization creation without programming
- Explain what Many Eyes was, what it provided, and what its contribution was
- Describe a spectrum of approaches for creating visualizations (ranging from automatic creation given data to low-level graphics libraries) and identify representative systems that occupy different places along that spectrum

Project Topics Feedback

- Returned today
  - Comments on your document
  - Score 1->10
- Can resubmit a revision or propose a new topic
  - Do it for Monday
- Get to work on design – Poster session coming in 2 weeks!

Reading

- Viegas & Wattenberg, "Many Eyes", '07
- Check out videos/demos of systems not shown

Upcoming

- Interaction

- Overview & Detail