Interactive Data Visualization … In Practice

CogSci 121 - HCI Programming Studio
Interactive Visualizations

http://www.youtube.com/watch?v=g9nvLqLM9Y0

http://gapminder.com

GapMinder D3 Demo

https://bost.ocks.org/mike/nations/
http://romsson.github.io/dragit/example/nations.html
DELPHI Data?
DELPHI

Data e-Platform to Leverage Multilevel Personal Health Information

Kevin Patrick, Chaitan Baru, Ted Chan, Sanjoy Dasgupta, Claudiu Farcas, William Griswold, Jeannie Huang, Yannis Katsis, Lucila Ohno-Machado, Yannis Papakonstantinou, Fredric Raab
Multiple sources of health data

Medical Records

Personal Health Data
(weigh-ins, run info, ...)

Environmental Data
(pollution, noise, greenspace, ...)

Genomic Data

Microbiome Data

Public Health Data
Multiple sources of health data

...of different types

Behavioral & Social Data

- Physical activity
- Sleep
- Social Networks
- Stress
- Diet

Personal Medical Data

- Medical Records
- Genomic
- Pharmaceutical
- BMI

Environmental Data

- Food & Grocery
- Pollutants
- Transportation
- Crime & Incivilities
Providing health care & population health requires reasoning across these layers.
Integrated DB: An Overview

- Contains ~3,000 indicators/variables
- From 4 datasets
  - 2013 Market Potential Data: Consumer buying patterns
  - 2012 HHSA Demographics: Demographics
  - 2012 HHSA Behavioral Health Outcomes: Hospitalizations & emergency department discharges for behavioral health conditions
  - 2012 SANDAG Healthy Communities Assessment Atlas: Information on physical & built environment (e.g., access to parks, libraries, etc.)
- With values for 41 Subregional Areas (SRAs) in San Diego County
  - An indicator may not contain values for some SRAs
- Of different types
  - Default: Absolute number of people with some characteristic
  - “Rate”: Number of people per 100,000 population
  - “Percent”: Number of people per 100 population
Raw DB: An Overview

- ~300 tables
- Each containing multiple indicators
- Contains data not contained in the integrated schema:
  - San Diego Air Pollution Control District (SDAPCD): Air quality data
  - Environmental Protection Agency (EPA): Air quality data
  - Health & Human Services (HHSA): Additional years of behavioral health data and health outcome data about other conditions (chronic diseases, injuries, etc)
  - Zillow: Housing market data
  - San Diego Association of Governments (SANDAG): Environmental point data
  - Automated Regional Justice Information System (ARJIS): Crime data
- Contains data with finer spatial granularity:
  - zip codes, points, polygons, ..
- Contains metadata
  - Table ‘metadata_tables’ contains description of each table
  - Table ‘metadata_sources’ contains the data sources
DELPHI Data Exploration Tool

Relationship Explorer

Healthcare and Administration versus Health Indicators

Color relationships:
- Similar
- Opposite

Display relationships where SRAs with measurable data exceed:
- 0
- 1

Display relationships where correlation is greater than:
- 0
- 1

Show indicators:
- Row Indicators
- Column Indicators

Carry medical/hospital/accident insurance: Buyers (Percent)
Medical insurance covers you and other HH members: Buyers (Percent)
Obtained medical insurance: through place of work: Buyers (Percent)
Obtained medical insurance: through a union: Buyers (Percent)
Obtained medical insurance: Medicaid: Buyers (Percent)
Obtained medical insurance: Medicare: Buyers (Percent)
Have dental insurance: Buyers (Percent)
Have disability insurance: Buyers (Percent)
COGS 121 - 2015
Team Thundercats

AirData

- Air Pollution
- CitiSense Data
- Interactive visualization of data collected by CitiSense sensors

http://airdata.herokuapp.com
LifeFitness

PlotSD

- Demographic-related risks in San Diego
- Crimes data, leading cause of death, HHS data,
- Visualization of personal risks in San Diego county based on age and gender

http://lifefitness.herokuapp.com
Extension-1

- Financial Crisis Analysis
- Zillow Data
- A visualization of US housing market data from 2004 to 2014 (various cities)

http://crisis-in-your-city.herokuapp.com
Siete

- San Diego Regional Crimes
- Crimes Data
- Visualization of the most common crimes in particular areas (ZIP codes)

http://siete2.herokuapp.com
Team Wayde Castillo
MyGreat

- Education, Occupation, Marital status
- San Diego Health & Human Services Agency Data
- A tool aimed to assist families and young professionals in making informed decisions about moving to San Diego!

http://mygreat.herokuapp.com
Team Nove

- Moving to San Diego
- Zillow Data
- Finding the best area in SD in terms of schools, budget and type of house

http://teamnove.herokuapp.com
DTOXX

- DUI
- Various sources of DUI Data
- Interactive visualization to inform about DUI and Ridesharing

http://dtoxx1.herokuapp.com/about
TinyElephants

PureRun

- Air Pollution and Running
- CitiSense Data
- Show pollution data around suggested running paths

http://purerun.herokuapp.com
PiedPiper

- House Market
- Zillow Data + HHS Data
- Visualization of SD regions with comparisons of age, housing costs, density, marital data

http://cogs121-piedpiper.herokuapp.com
Next Steps

• Amy Fox and Linda Weber (IBM Design)

• Technical Discussions (required), Fri 4/17
  - Troubleshooting with Assignment 2
  - Quiz on Week 4 Content

• Readings (required)
  - Johnson (Designing With the Mind in Mind) - Chapter 13, 14
  - Meirelles (Design for information: an introduction to the histories, theories, and the best practices behind effective information visualizations) - Introduction + Chapter 1

• Next Week:
  - Rapid Prototyping and Wireframing
  - Assignment 3
The BIG Idea

Amy Fox and Linda Weber