Presentation Order

Samuel Crow, Wesley Lee, Sachin Mehta
Aaron Bauer, Donny Huang, Zuoming Shi
Dylan Babbs, Jordan Starkey
David Caldwell, James Wu
Yan Jin
Zhizhen Lin, Parmita Mehta, Adam Richie-Halford, Joshua Smith
Xiaochuan Du, Mingwei Tang
Erin Peach, Wenhuan Yang
Lucy Simko, Eric Zeng
Gabriella Gorsky, Younghoon Kim, Kyle Thayer
Youying Lin, Vivek Paramasivam
Ayush Saraf
William Strimling
John C Walker, Yitao Wang, Ricki Si Xie, Li Zeng
Haoran Cai, Harsh Keswani, Abhishek Pratap, Xinglu Yao

Examining the CSE Admission Pipeline

Sachin Mehta Sam Crow Wesley Lee

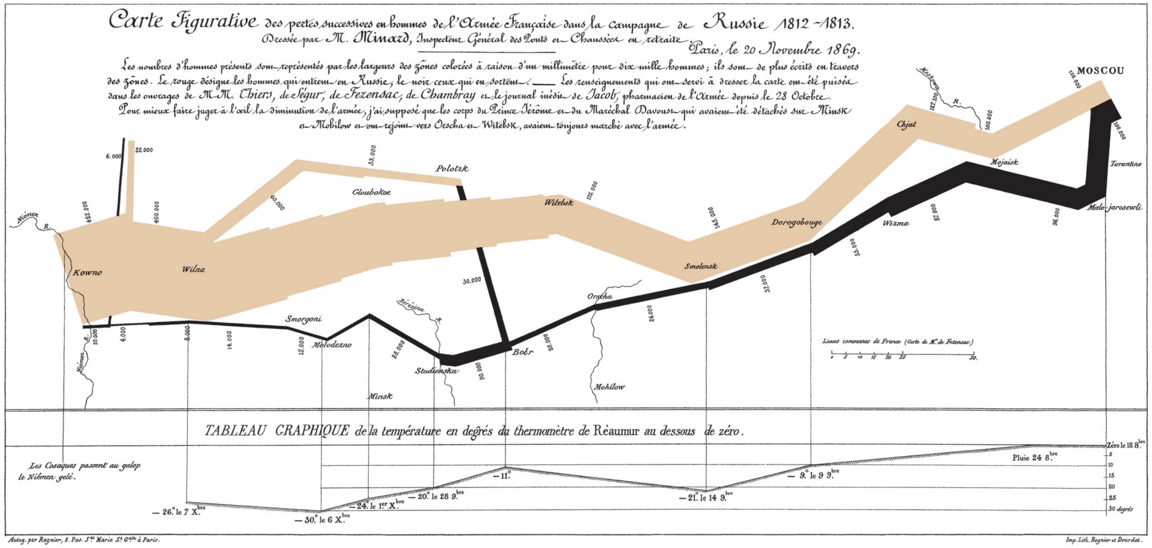
CSE 512

Problem Description

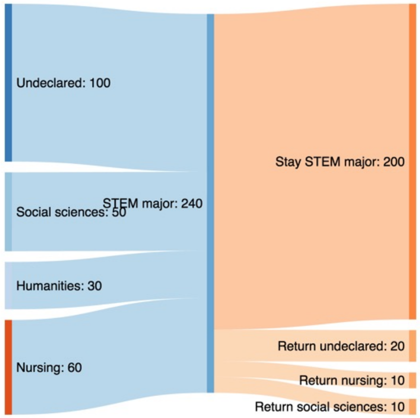
- **Data:** A cohort of CSE 142 students tracked from 142 to 143 to applying for the CSE major
- **Goal:** Understand why some students drop out and how these students differ from those who continue
- **Potential Questions**
    - Are gender disparities present at every stage of the pipeline?
    - What kind of students drop out after taking 142-143?

Sankey diagrams

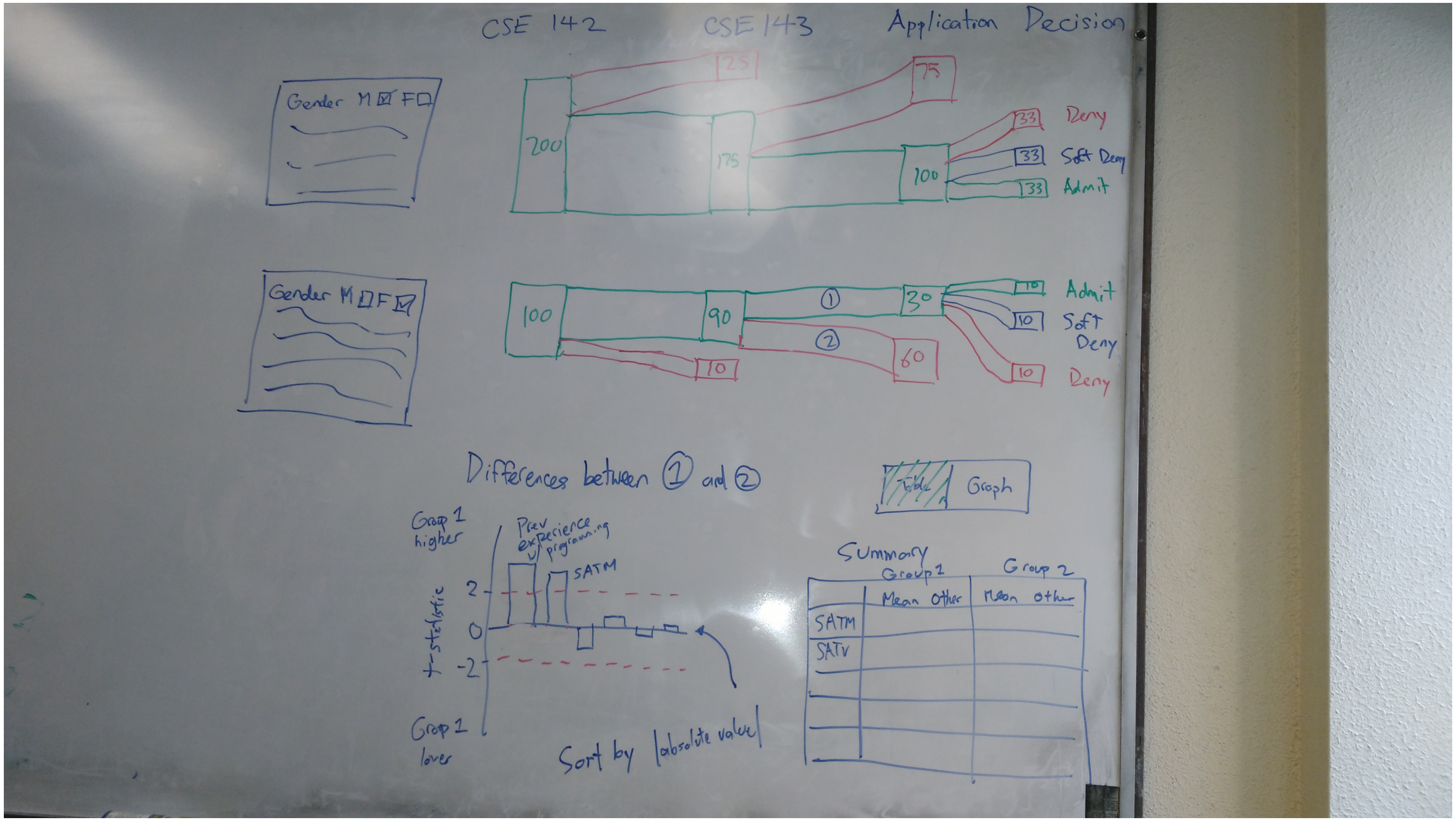
- Minard's map of Napoleon's Russian campaign



- "Fake" example of a Sankey diagram



Storyboard



CSE 142
CSE 143
Application
Decision
Gender M ☒ F ☐
200
25
175
75
100
33
Deny
33
Soft Deny
33
Admit
Gender M ☐ F ☒
100
90
10
1
30
2
60
Admit
10
Soft Deny
10
Deny
Differences between 1 and 2
Table
Graph
Group 1 higher
Prev experience w/ programming
SATM
t-statistic
2
0
-2
Group 1 lower
Sort by |absolute value|
Summary
Group 1
Group 2
Mean
Other
Mean
Other
SATM
SATV

Feedback

- How to handle multiple attempts of 142/143/applications?
- How to handle a high number of covariates when comparing groups?
- Should we restrict filter options (simplicity vs complexity)?
- Is it useful to allow users to follow (anonymized) individuals?

Impact of Competitive CSE Admissions

Zuoming Shi, Donny Huang, Aaron Bauer

Motivation

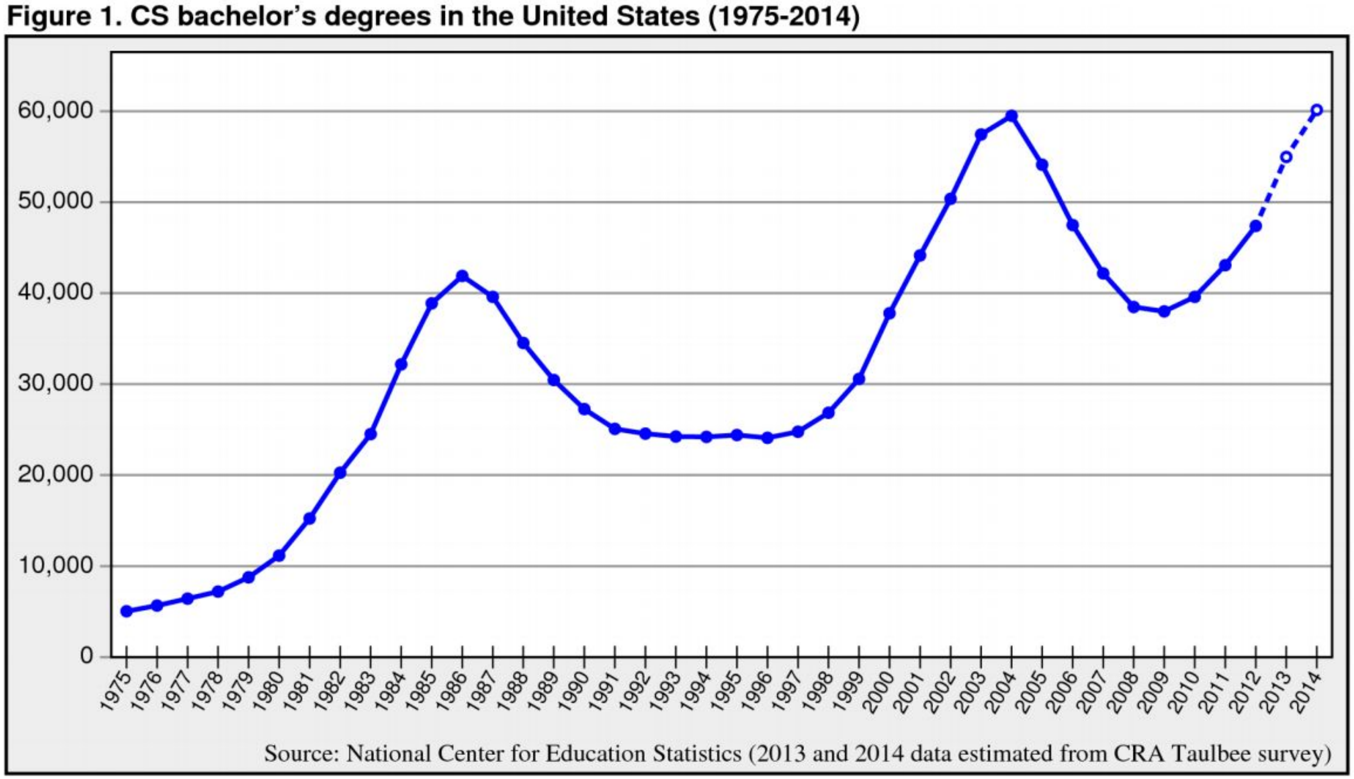
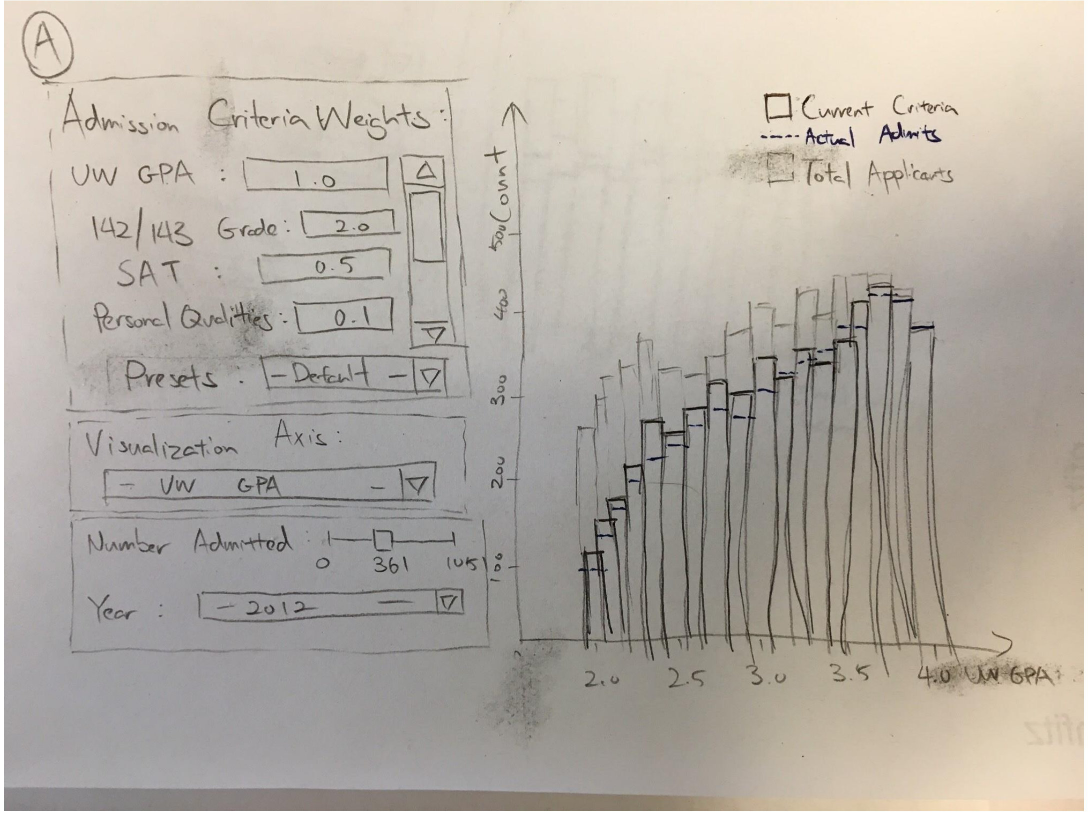


Figure 1. CS bachelor's degrees in the United States (1975-2014)
60,000
50,000
40,000
30,000
20,000
10,000
0
1975 1976 1977 1978 1979 1980 1981 1982 1983 1984 1985 1986 1987 1988 1989 1990 1991 1992 1993 1994 1995 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011 2012 2013 2014
Source: National Center for Education Statistics (2013 and 2014 data estimated from CRA Taulbee survey)

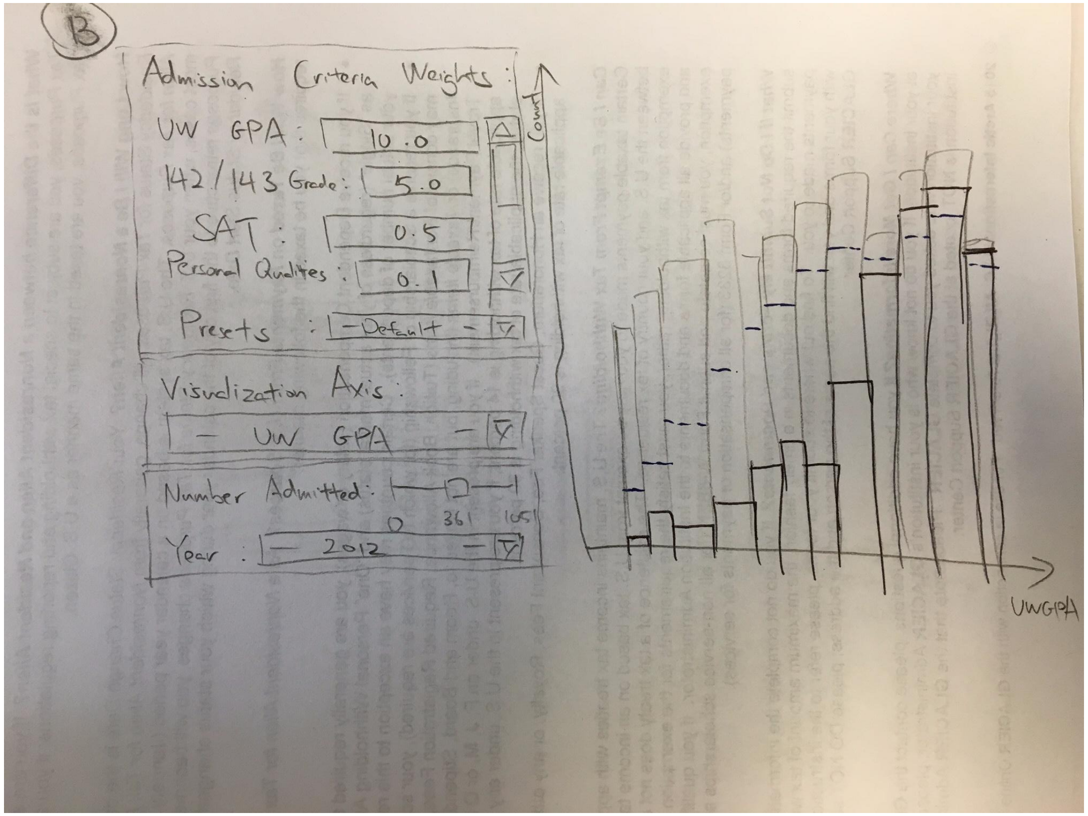
Related Work

- *Deciding to major in computer science: a grounded theory of students' self-assessment of ability* (Lewis et al., 2011)
  - Competitive culture has a significant impact
- *Perceptions of non-CS majors in intro programming: The rise of the conversational programmer* (Chilana et al., 2015)
  - Students what computational fluency
- *A Study of Factors Promoting Success in Computer Science Including Gender Differences* (Wilson, 2002)
  - Most predictive: comfort level, math background, luck
  - No gender difference in these factors

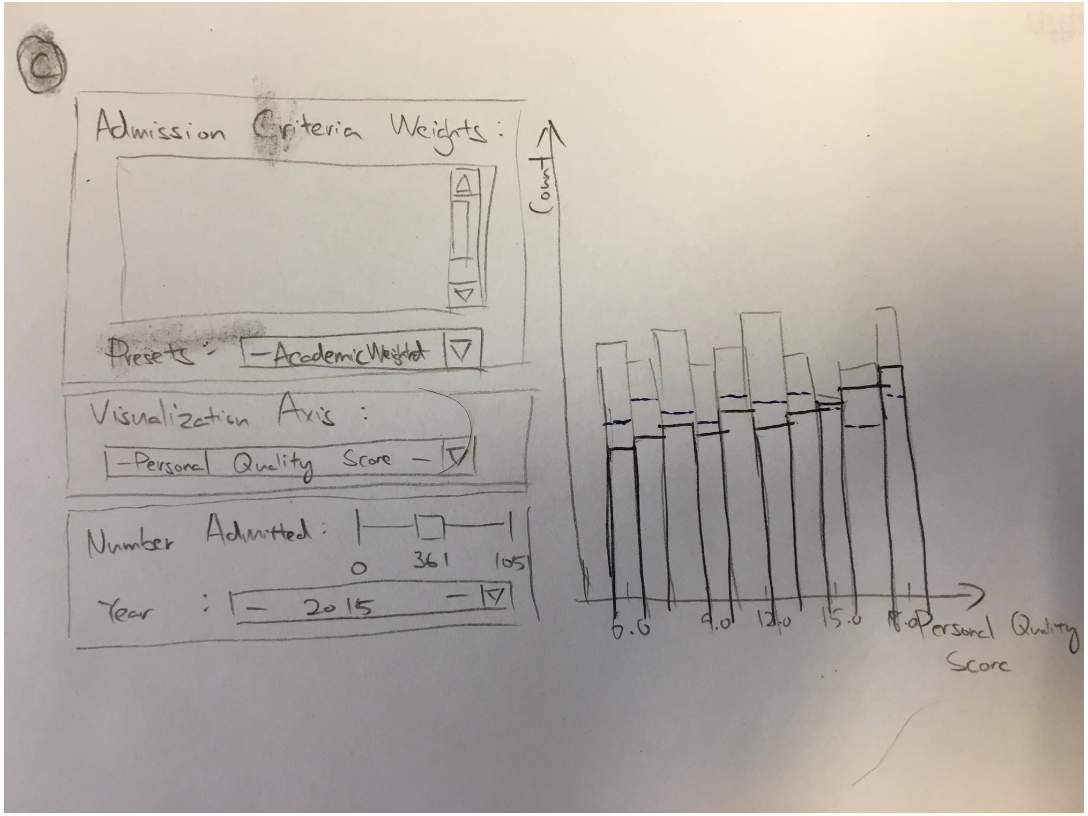
Progress Report



Progress Report



Progress Report



Feedback Questions

- Is the single bar-chart representation effective understanding the current and custom admission criteria? Might small multiples be more effective?
- What is an elegant way to present the weighing UI to the reader?
- How can we allow the audience to adjust weighing in a fast and intuitive way?
- What insights from this project might feel particularly relevant to you?
- Are there anything you wish we included in our project?
- Do you think we are showing appropriate amount of detail?
- How might things be done differently?
- Do you know of any similar projects from which we could get inspiration from?

Visualizing UW Prerequisite Course Sequences

CSE 512: Final Project

Dylan Babbs
Jordan Starkey

19 May 2016

Unclear prerequisite info in course catalog

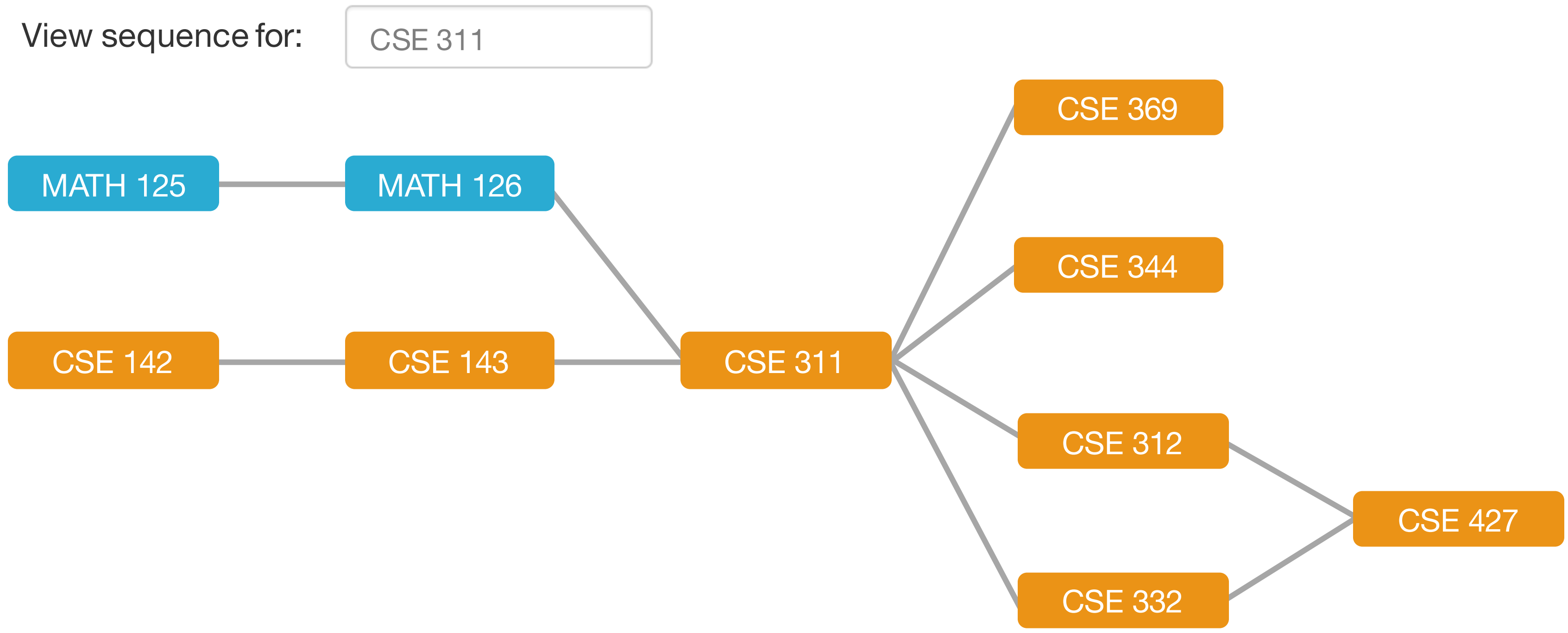
- Course catalog only shows latest prerequisite requirement, not all coursework required
- Degree planning information or course sequences only available on departmental websites.
- **Solution**: visualize course sequences with trees!

CSE 373 Data Structures and Algorithms (3)
Fundamental algorithms and data structures for implementation. Techniques for solving problems by programming.
Prerequisite: CSE 143

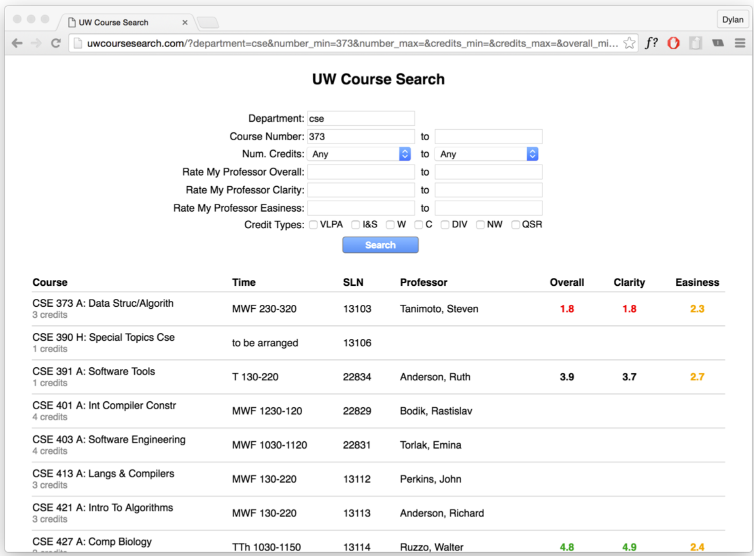
CSE 373, an elective course, lists CSE 143 as the only prerequisite. However, CSE 143 requires CSE 142 as a prerequisite.

Therefore, the prerequisites for CSE 373 are actually CSE 143 and CSE 142.

Anticipated storyboard

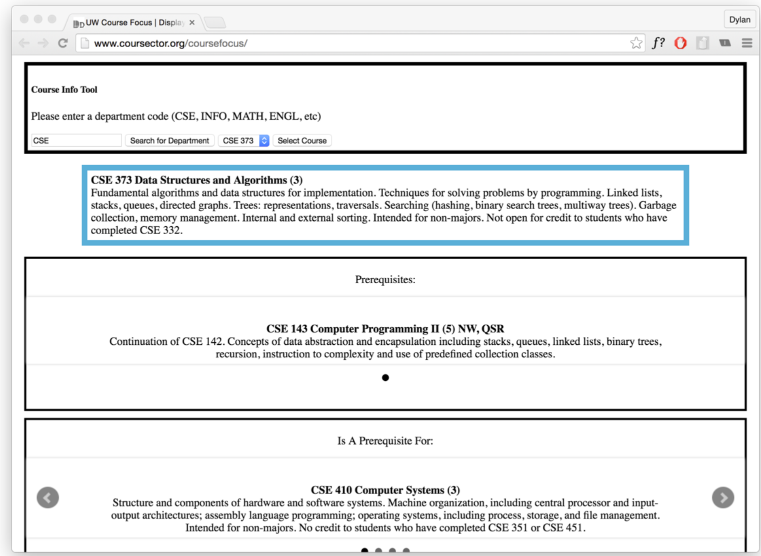


Relevant prior work



UW Course Search

Brice Hulse, 2016



Course Focus

iSchool's Course Sector, 2015

Questions/Issues we are addressing

- How do we handle "soft prerequisites" in course description (*i.e.*: "by permission of instructor")?
- What's the best tool for writing web scrapers?
- What's the best data schema for trees and networks? Two column CSV or JSON?

Interactive Brain Connectivity Visualization

David Caldwell and Jing/James Wu

(working with Emily Fox and Nick Foti)

Brain connectivity

Goal:

Interactive visualization exploration tool

Tailored to large numbers of pairwise neural connectivity dynamics

Why?

Analyze connections from various regions of the brain to one another

Underexplored metrics of magnetoencephalography (and electrocorticography) data.

Challenge:

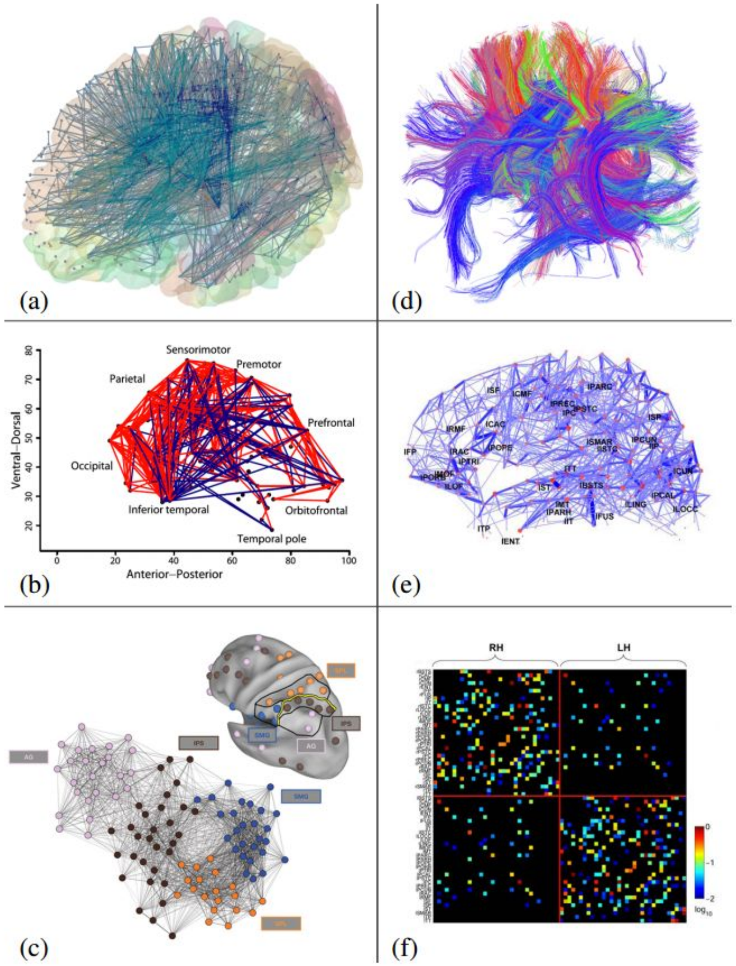
Visualizations are largely exploratory:

We don’t know the importance of metrics.

Encoding of a large number of multivariate attributes that exist for each pairwise edge.

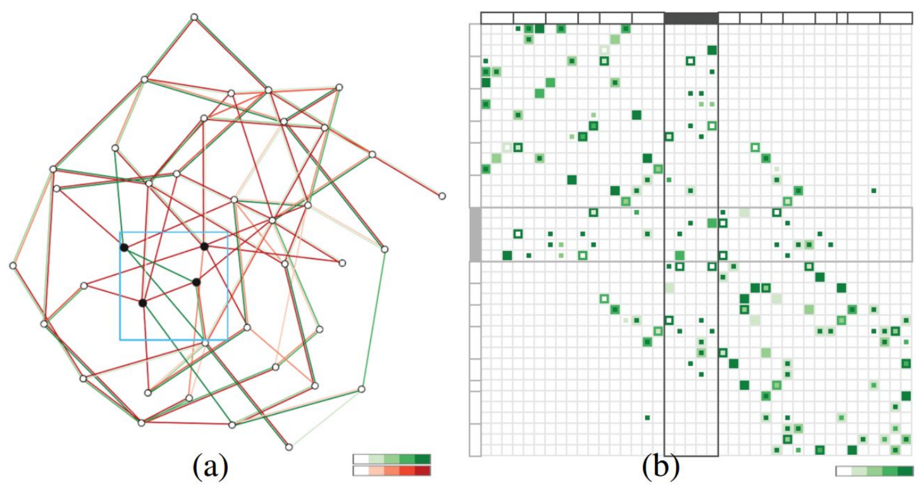
Related work

Functional Connectivity (mostly fMRI data) | Anatomic Connectivity (mostly diffusion tensor imaging)



Comparing Different Weighted Graphs

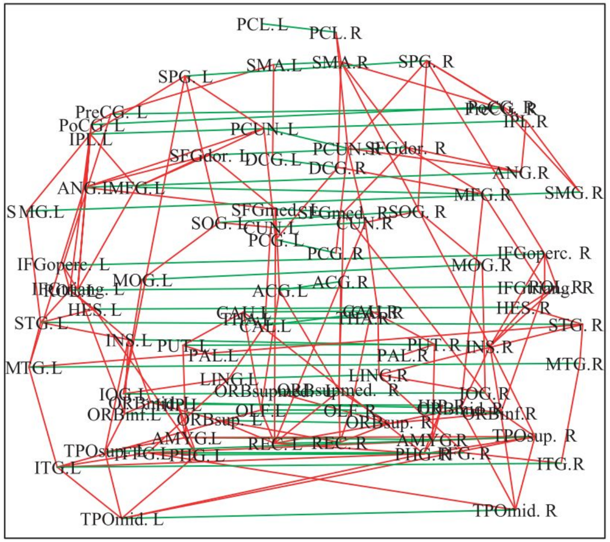
Node - link **Matrix representation**



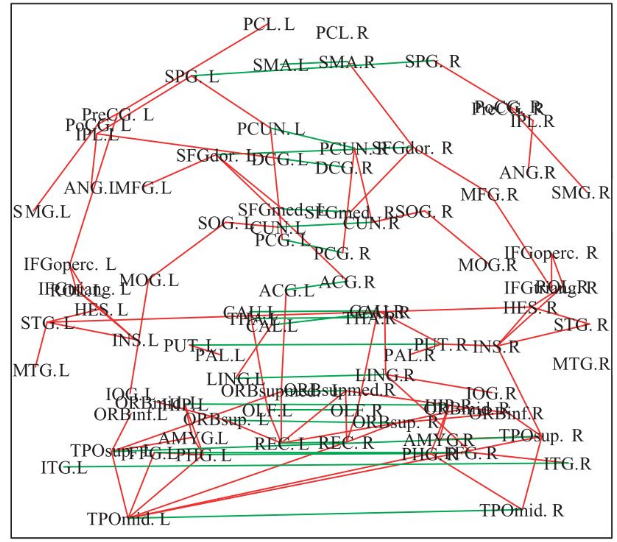
Alper, B, et al. “Weighted Graph Comparison Techniques for Brain Connectivity Analysis”, CHI 2013

Undirected fMRI based connectivity graph

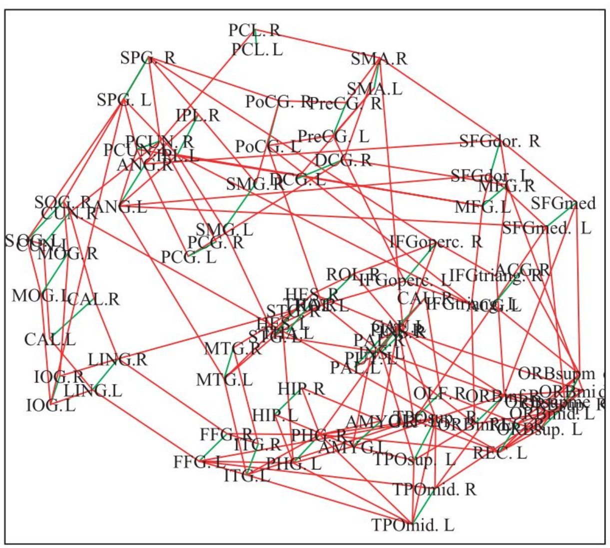
low frequencies, coronal view



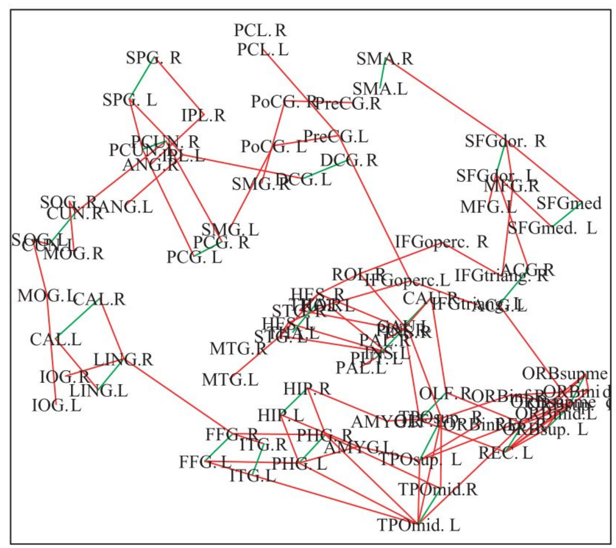
high frequencies, coronal view



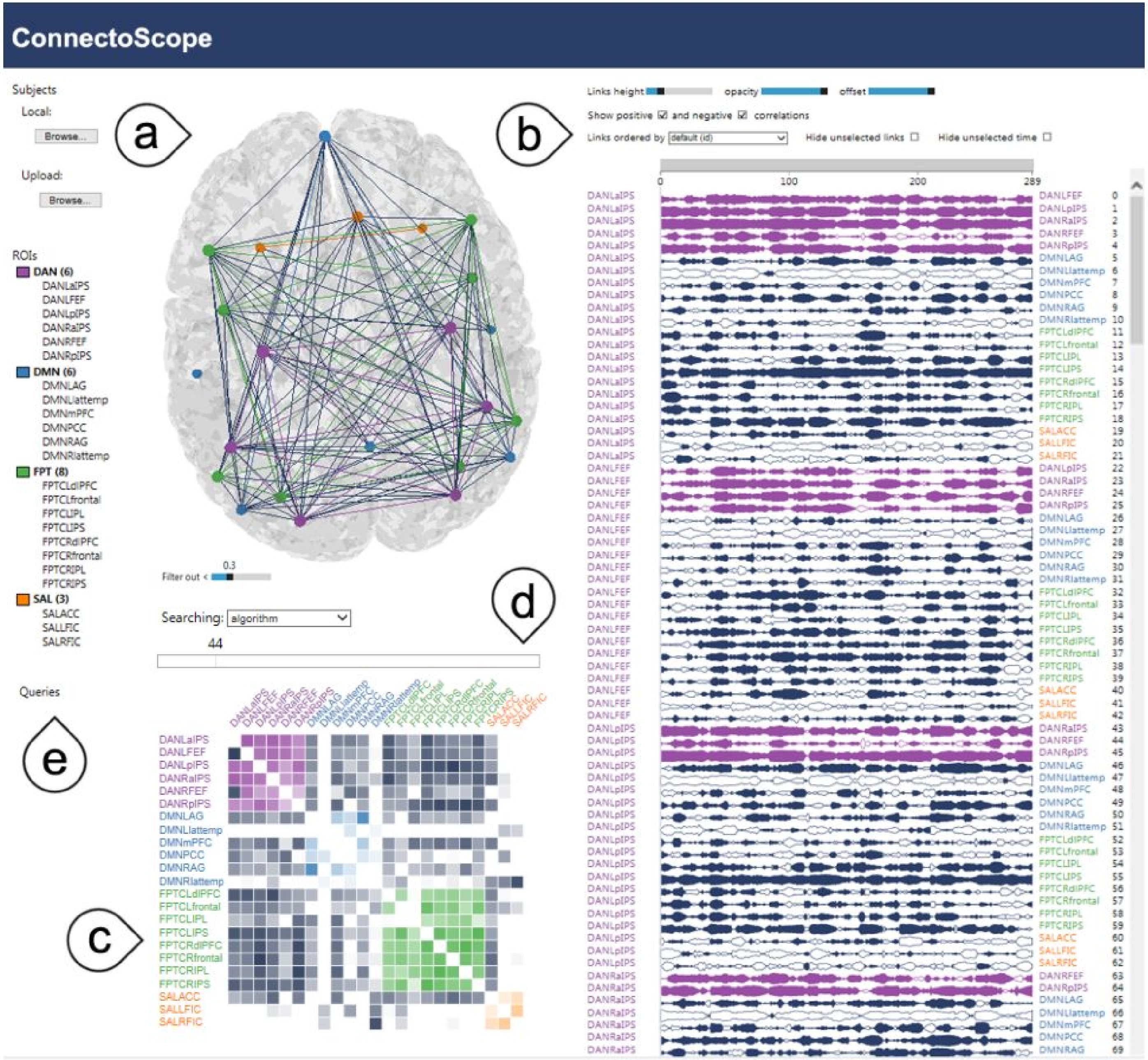
low frequencies, sagittal view



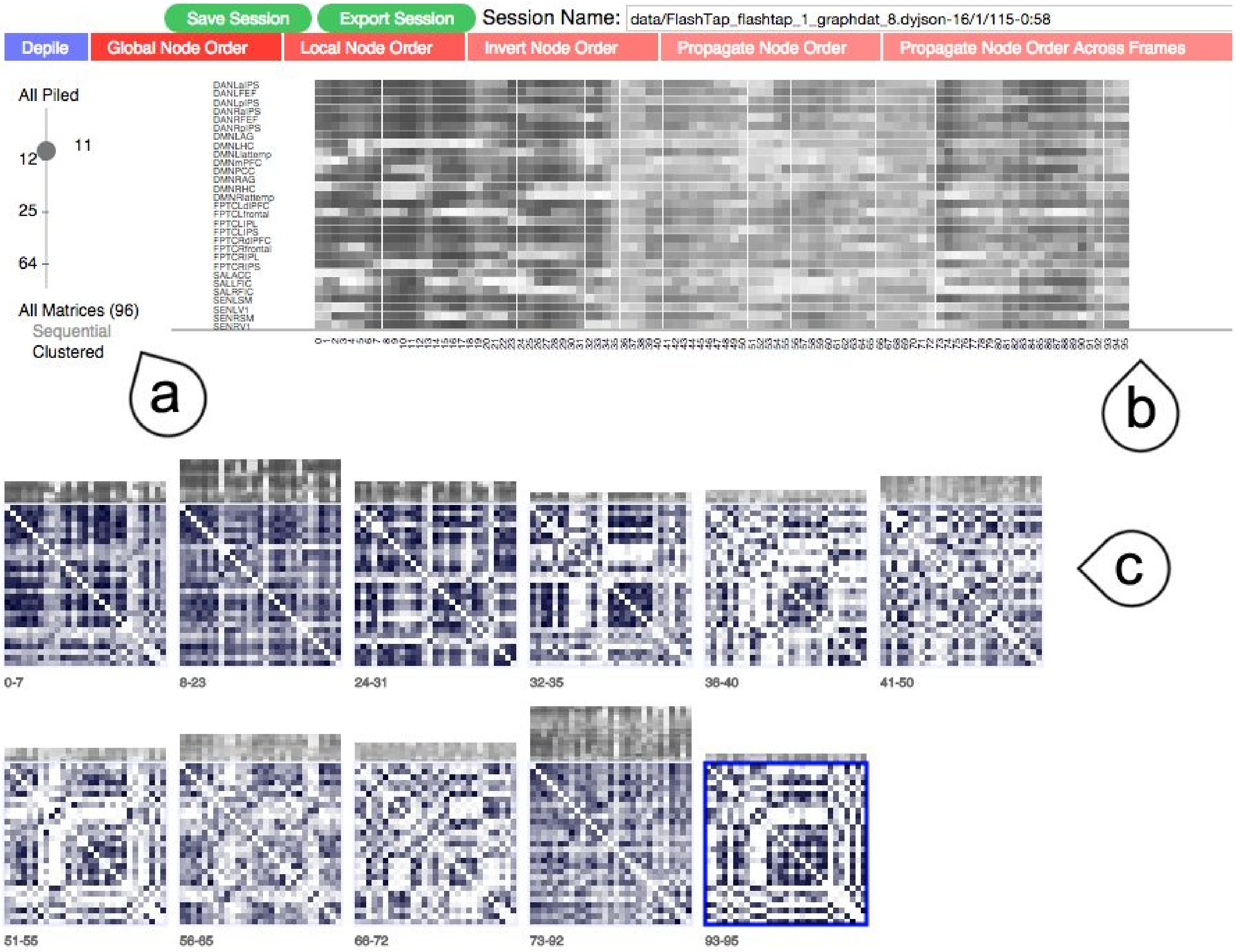
high frequencies, sagittal view



ConnectoScope - from Riche et al, OHBM 2015



SmallMultiples- from Bach et al, OHBM 2015



Where are we different?

Domain specific knowledge:

- MEG (magnetoencephalography) + ECoG (electrocorticography)
- Cortical surface only, frequency-based metrics

Aggressive visual decluttering and de-occlusion:

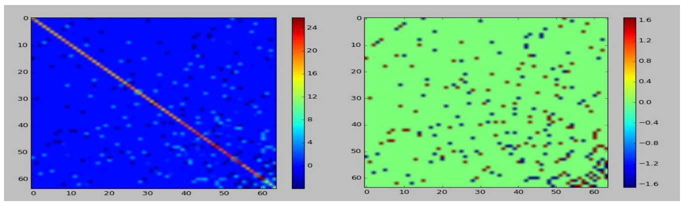
- Use of chord diagrams to show overall functional connectivity
- Threshold pruning: showing few unnecessary elements by thresholding
- Bundling - similar to hierarchical bundling

Client-side computation of different statistics for:

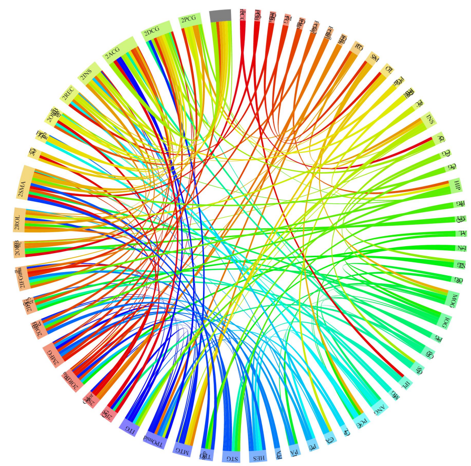
- frequency bands of interest, threshold connectivity strengths for subselection, and anatomical locations of connectivity for a region of interest.

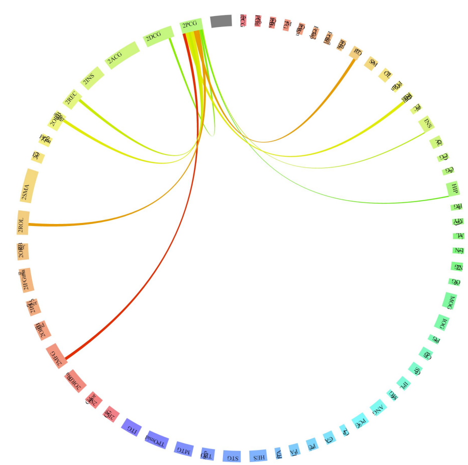
Structure of our data and program - current progress

Data currently exists as **f** x **l** x **l** matrix, where f is frequency, l represents locations, and the values are complex numbers, comes in .npy file format.



Graphs of real component, phase angle for a single frequency





Input a lower and upper frequency range of interest

Select Real, Imaginary, or Absolute Value from the dropdown box

Once you have made your choices, hit the button to submit and draw

Lower Frequency = 0

Upper Frequency = 65

Real ▼

Show self connections   Exclude self connections

Click here to submit selections from above and update the graph

Future directions / workflow

User runs data processing script on .npy file, JSON format .txt file is output and saved to the same folder as the index.html file

→

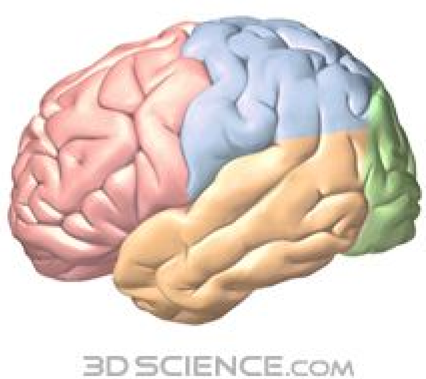
User interacts with web browser to perform interactive visualization - all subselections done in web

→

Chord diagram dynamically updates

→ Anatomic plot of brain updates, color coded for region (subselected by hover)

→ Bar chart demonstrating connections for a region of interest (subselected by hover)



Uncertainties - feedback?

Useful statistics?

Partial coherence, partial mutual information?

Chord diagram - is it optimal for pairwise graphs?

Hierarchical edge bundling?

Most useful features?

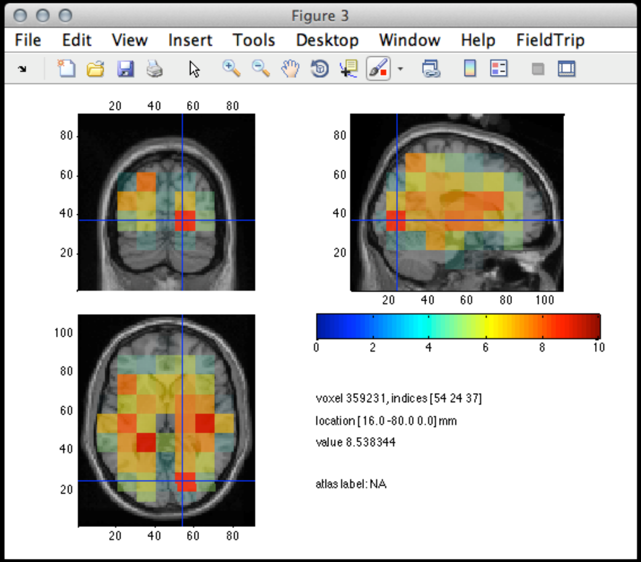
As a user, is it ok to have to run a python file on your own PC, within a folder of interest, and subsequently upload it into the folder for visualization?

Understand Brain as an Interactive Network

Yan Jin, PhD Student
Industrial Engineering Department

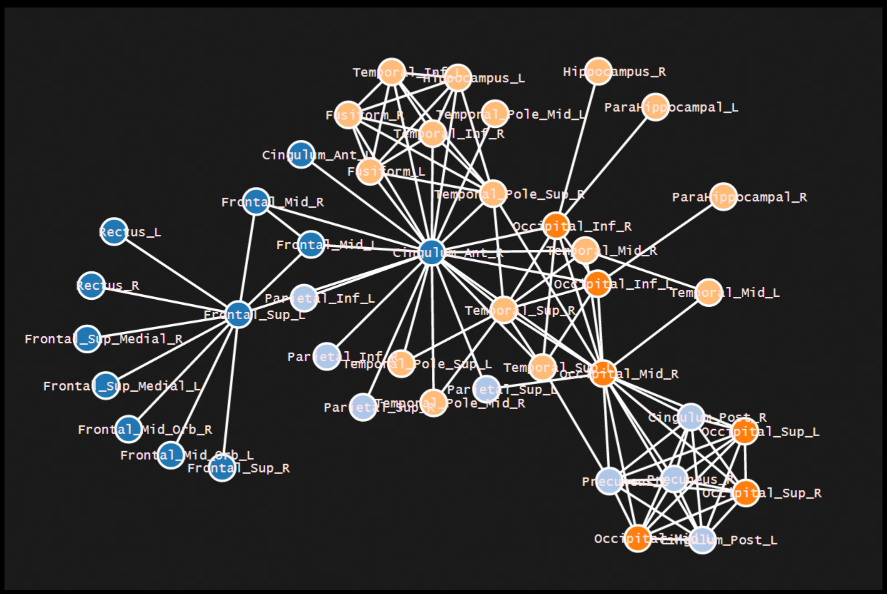
Prior Work

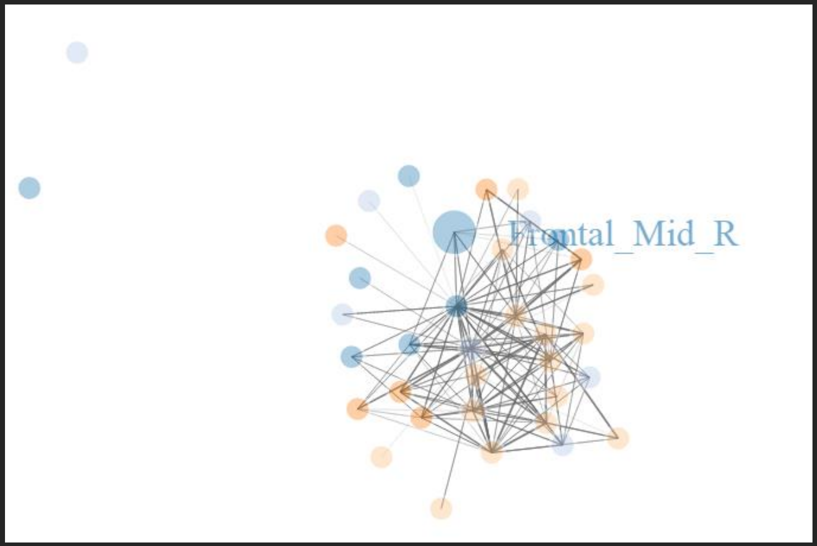
- Always integrating network analysis functions besides visualizing
- Cannot zoom in to check the details especially for dense network
- Lack interaction.



Name	Feature	Website
BrainNet Viewer [2]	3D graph-based brain network demonstration with nodes and edges; 3D brain surface view	www.nitrc.org/projects/bnv/
Connectome Viewer [1]	3D graph-based brain network demonstration with nodes and edges; 3D brain surface view	cmtk.org/viewer/
Caret	3D surface view and nodes	brainvis.wustl.edu/wiki/index.php/Caret:About
NetworkX	Graph-based network analysis; 2D demonstration with dots and lines	networkx.github.io
Pajek	Graph-based network analysis; 2D demonstration with dots and lines	vlado.fmf.uni-lj.si/pub/networks/pajek/

Current Progress





Task	Progress
Label each node	Finished
Zoom in to watch specific area of nodes	Finished
Width of edges represents the strength of connection	Late May
Use 3D coordinates to fix the layout	Late May
Improve the color encoding	Early June
Add legend	Early June
Allow user to input its own network data (optional)	Early June

Questions To Deal With

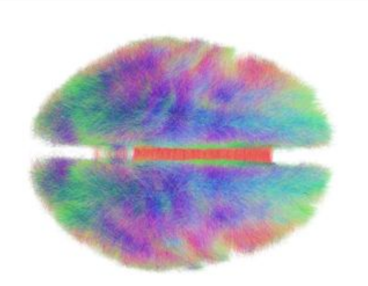
- The most difficult ongoing task is how to make us of 3D coordinate information to fix the layout of connectivity network in D3.

Visualizing the Human Brain Spaghettome

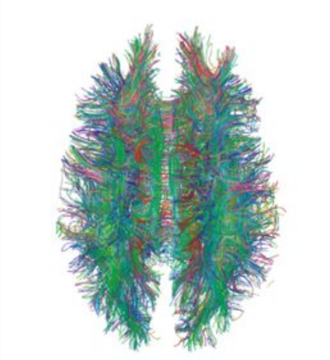
Adam Richie-Halford, Josh Smith, Parmita Mehta, Zhizhen Lin

Research Problem

- Difficulty in analyzing individual fiber tracts in 3D brain models



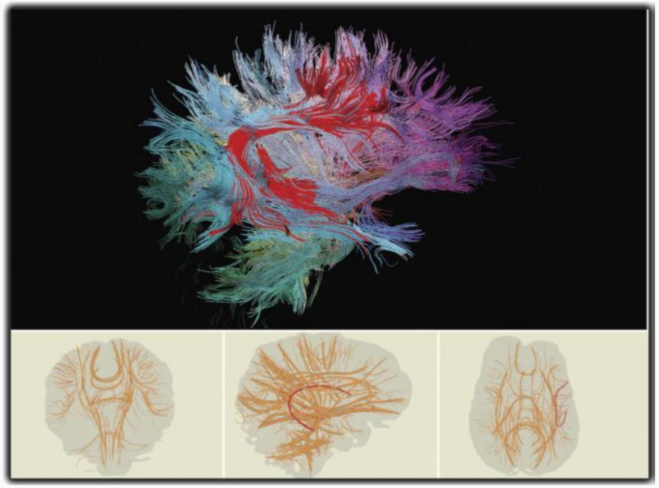
- Need to represent both:
  - Anatomical geometry of the brain
  - Volumetric data along individual neural tracts



Literature Review

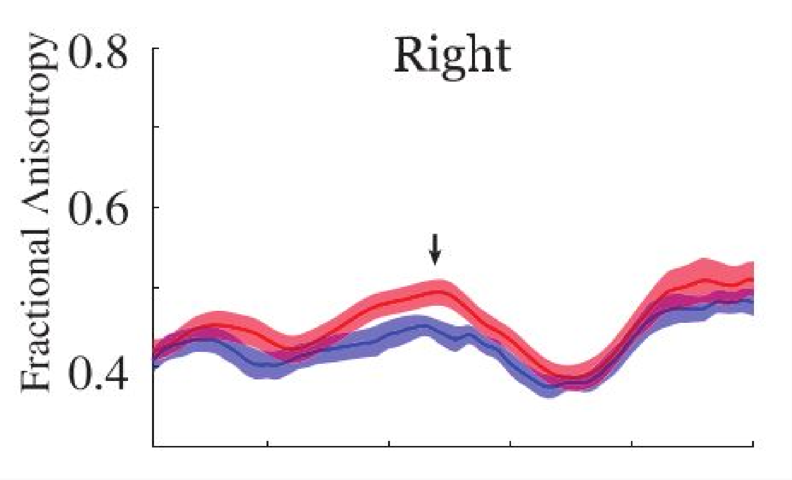
Visual Challenge

3D spaghettome -> 2D slice



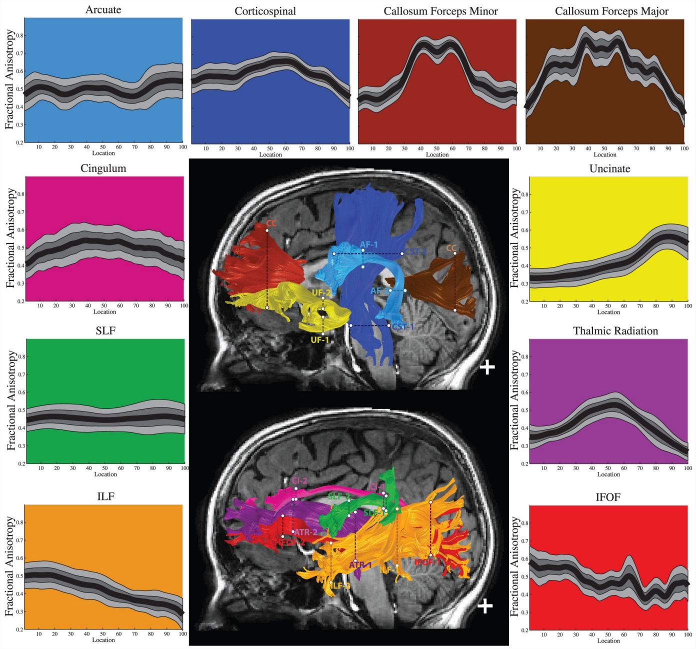
Functional Challenge

Variation between brains

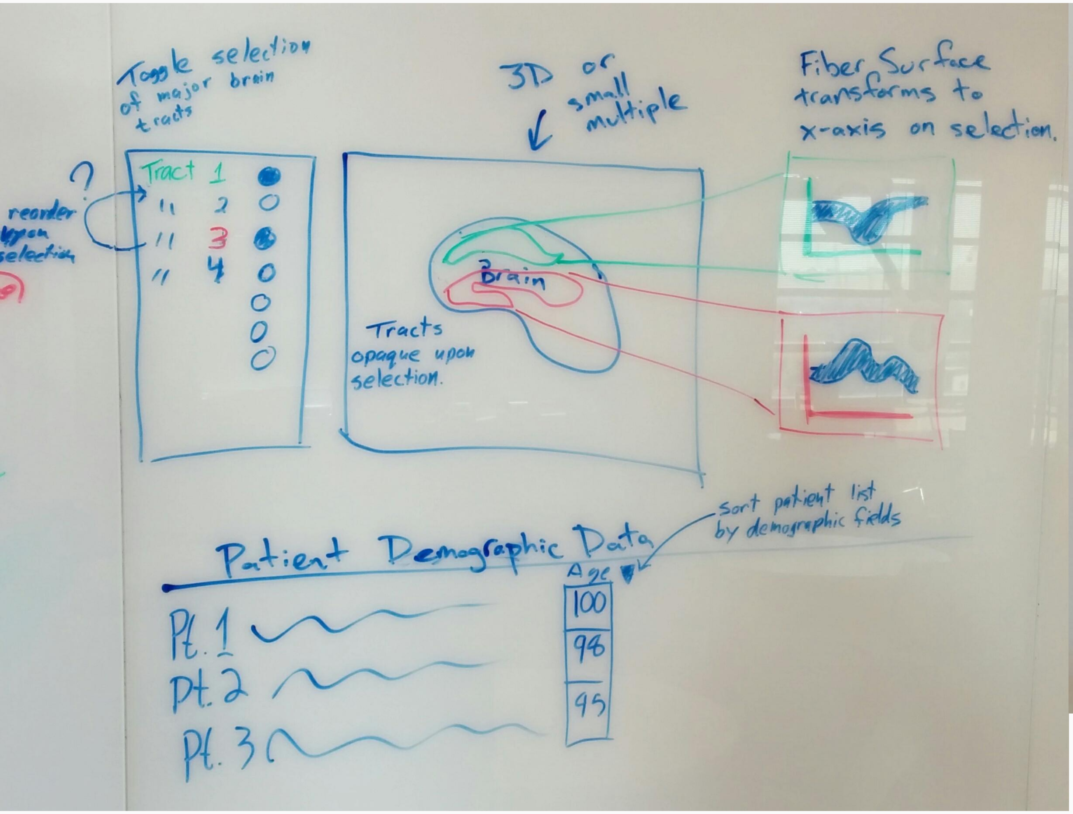


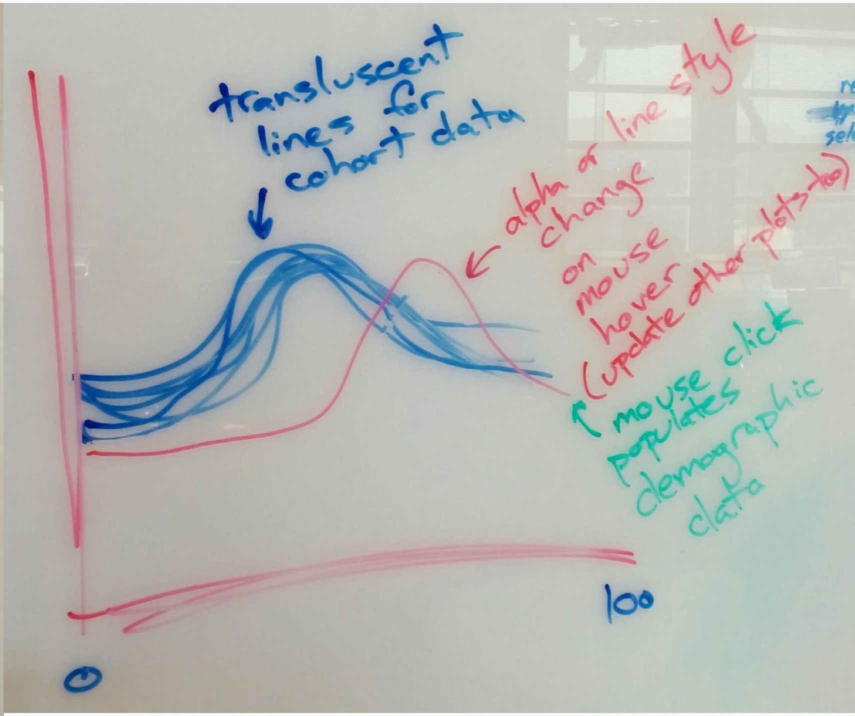
Our Approach

- Don't reinvent the wheel
- Abandon 3D for representation of volumetric data
- Focus on visualizing bundles and data along these bundles
- Use animation and interactivity to orient user to anatomical brain
- Inspired by this static visualization approach



Current Progress: Storyboard





Feedback Questions

1. Are we using appropriate visual encodings?
2. Any other libraries we can use?
3. Would small multiples be better than 3D for anatomical brain view?
4. ~2.7 millions points in 3D. Current approach is to plot an enclosing surface (i.e. alpha-shape) to reduce computational load. Any better ideas?

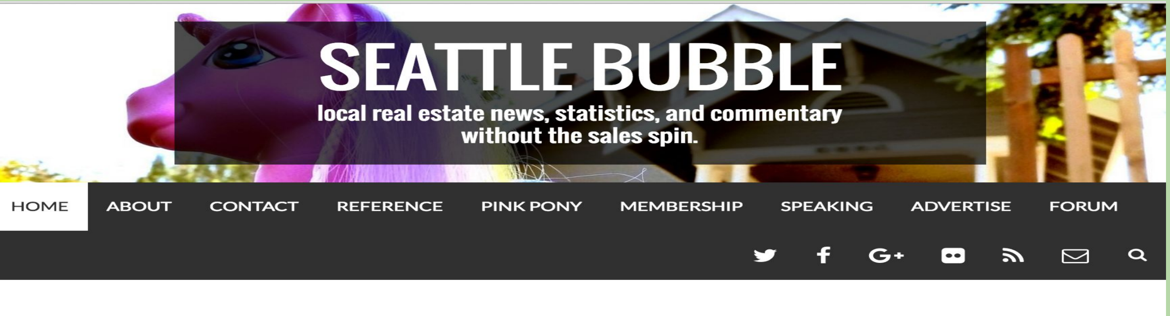
Current Progress: Development

- Design Decisions
- Component Development
  - 3D viewer
  - Selection + Highlighting
  - Sorting and Comparing

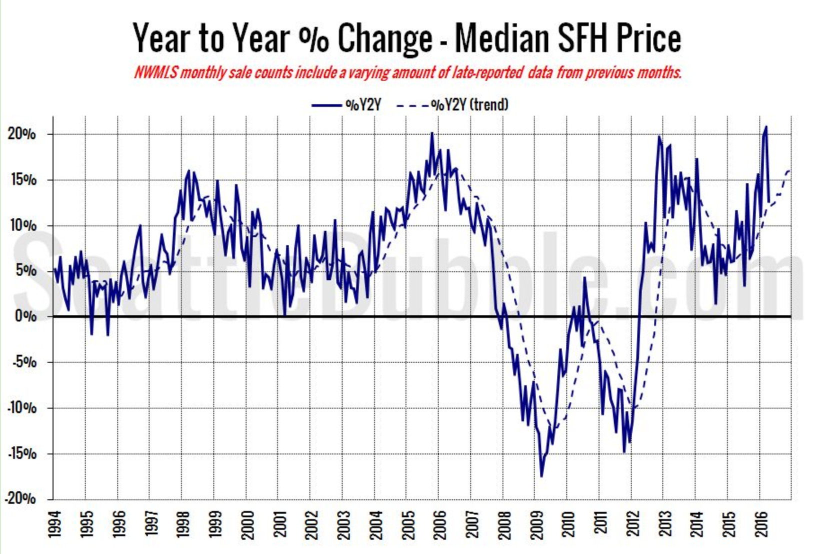
Visualization of housing sales and rental price in the great Seattle area

Mingwei Tang
Xiaochuan Du

Prior work

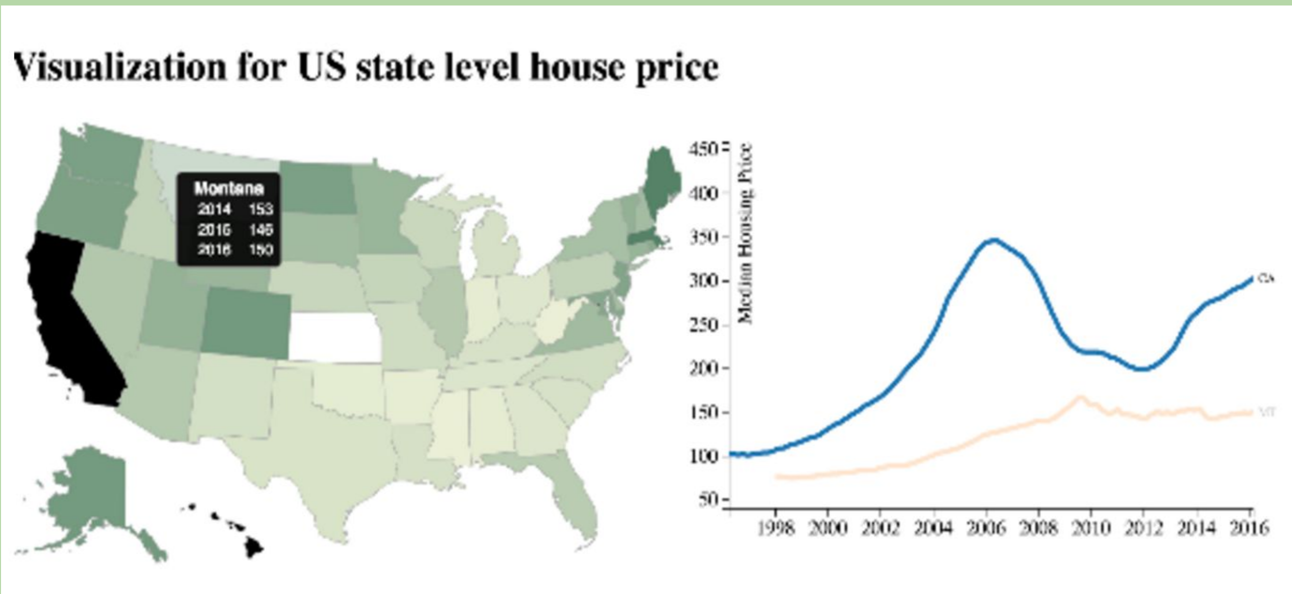


Seattle bubble
http://seattlebubble.com/blog/
A blog discussing Seattle housing price bubble



The changing rate of housing price.

Prior work

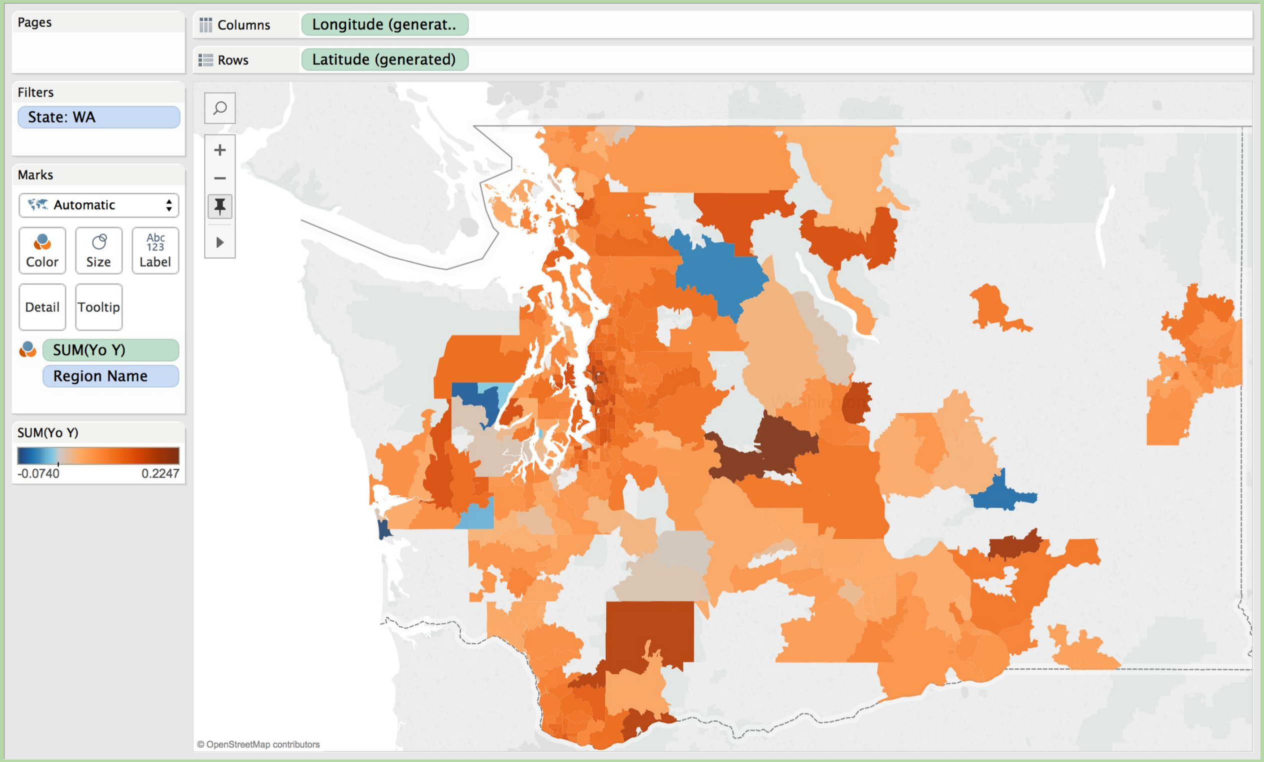


A3: U.S state level housing price

Using Zillow data

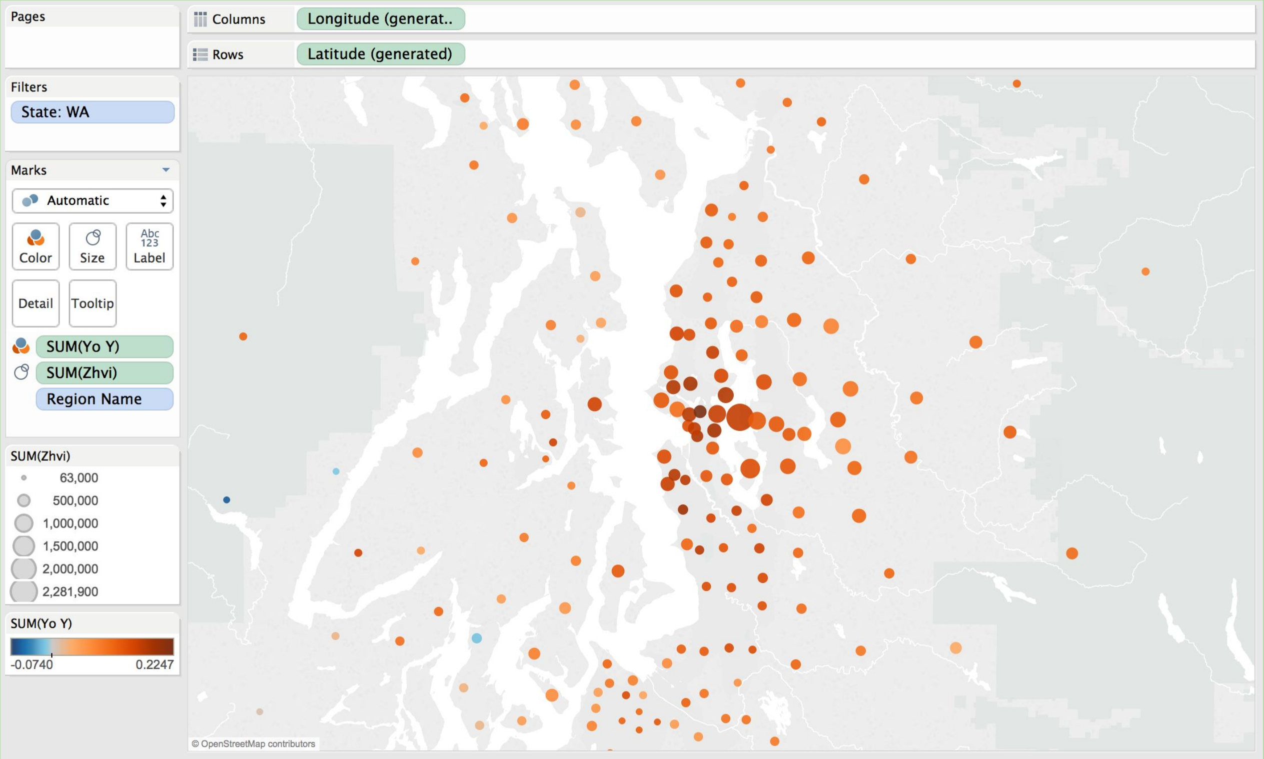
Click the state and show line chart

Current Progress



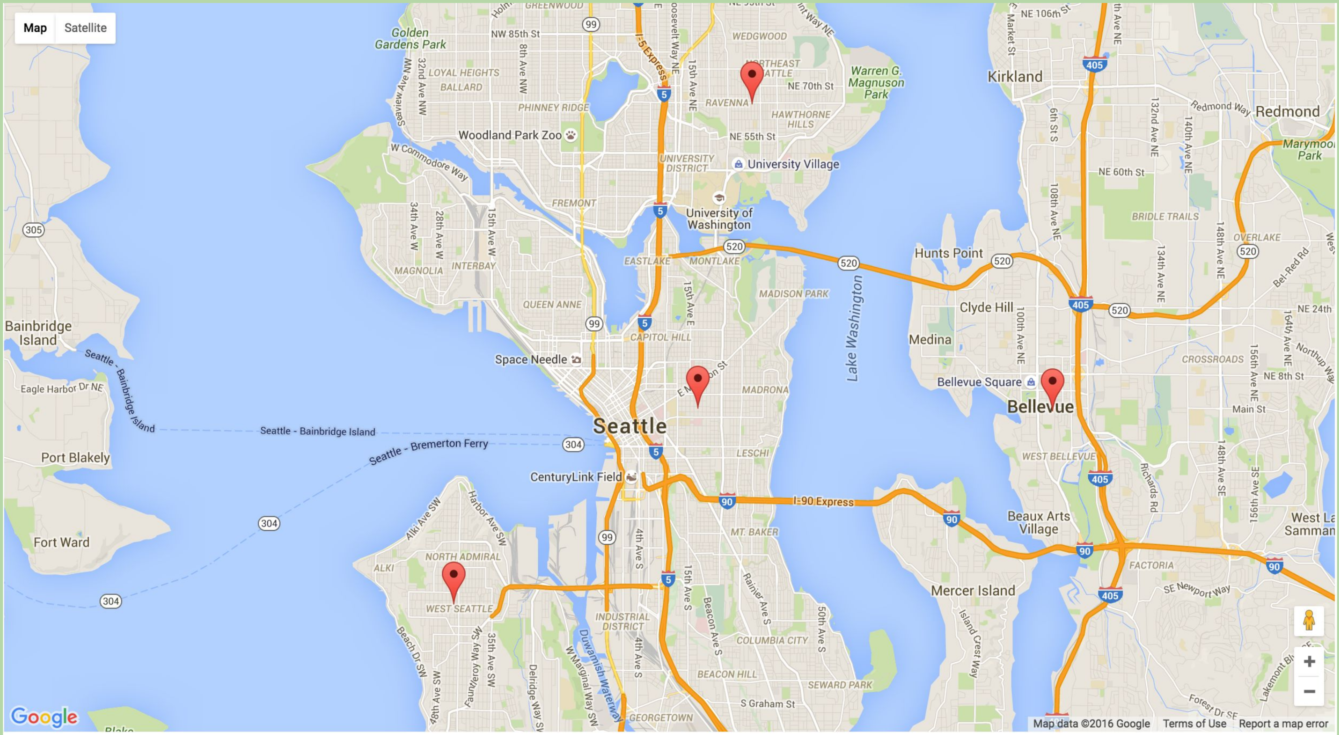
In Washington state, we have housing data for the colored area.

Current Progress



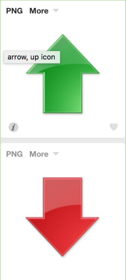
Using a dot-representation, we are able to show the map with more detail.

Current Progress

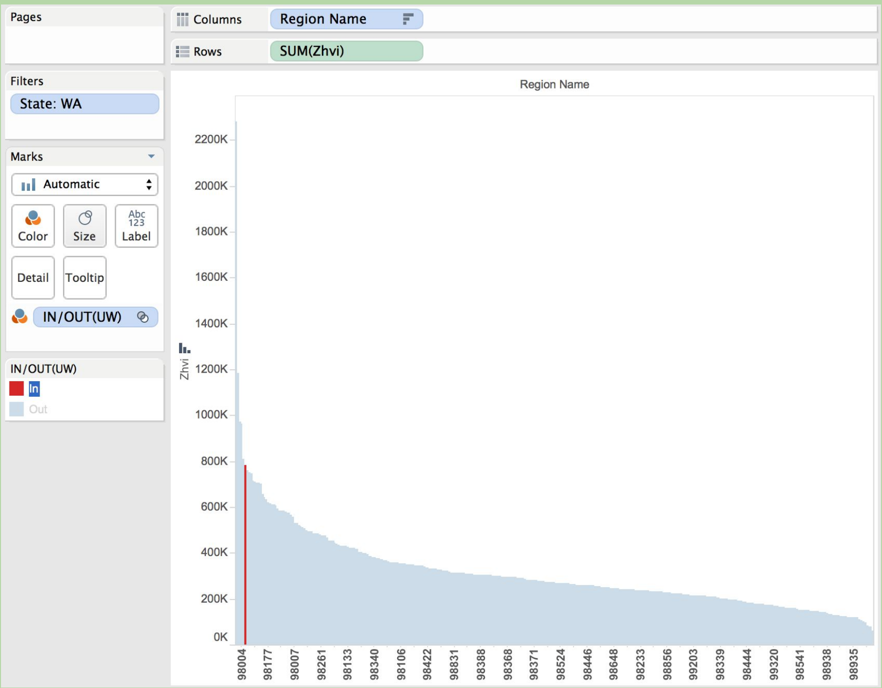


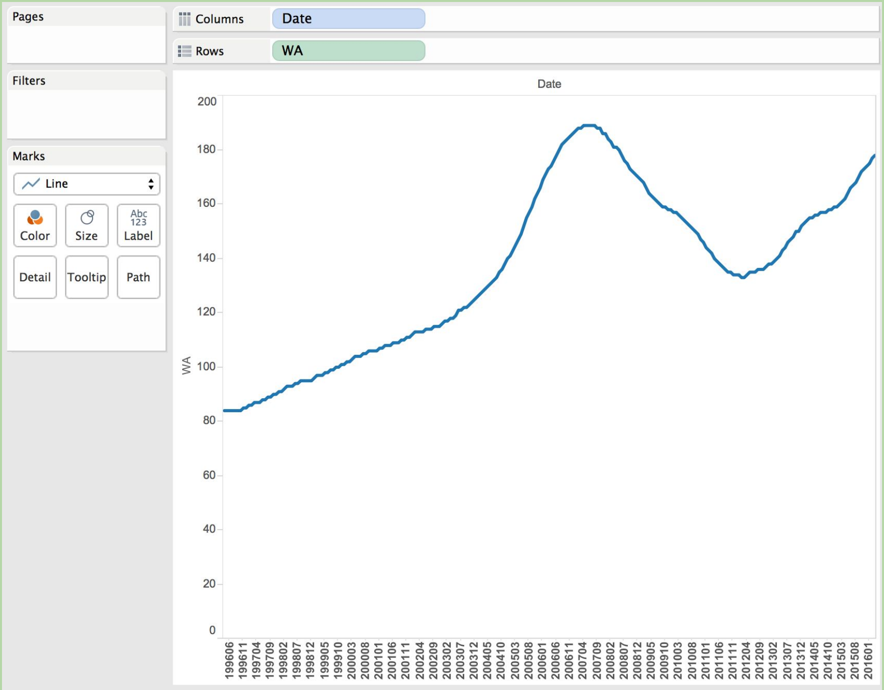
Mapping price features on top of the google map API

The arrows denote the price changing

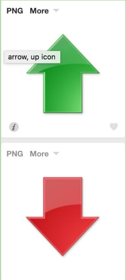


Current Progress





Feedbacks



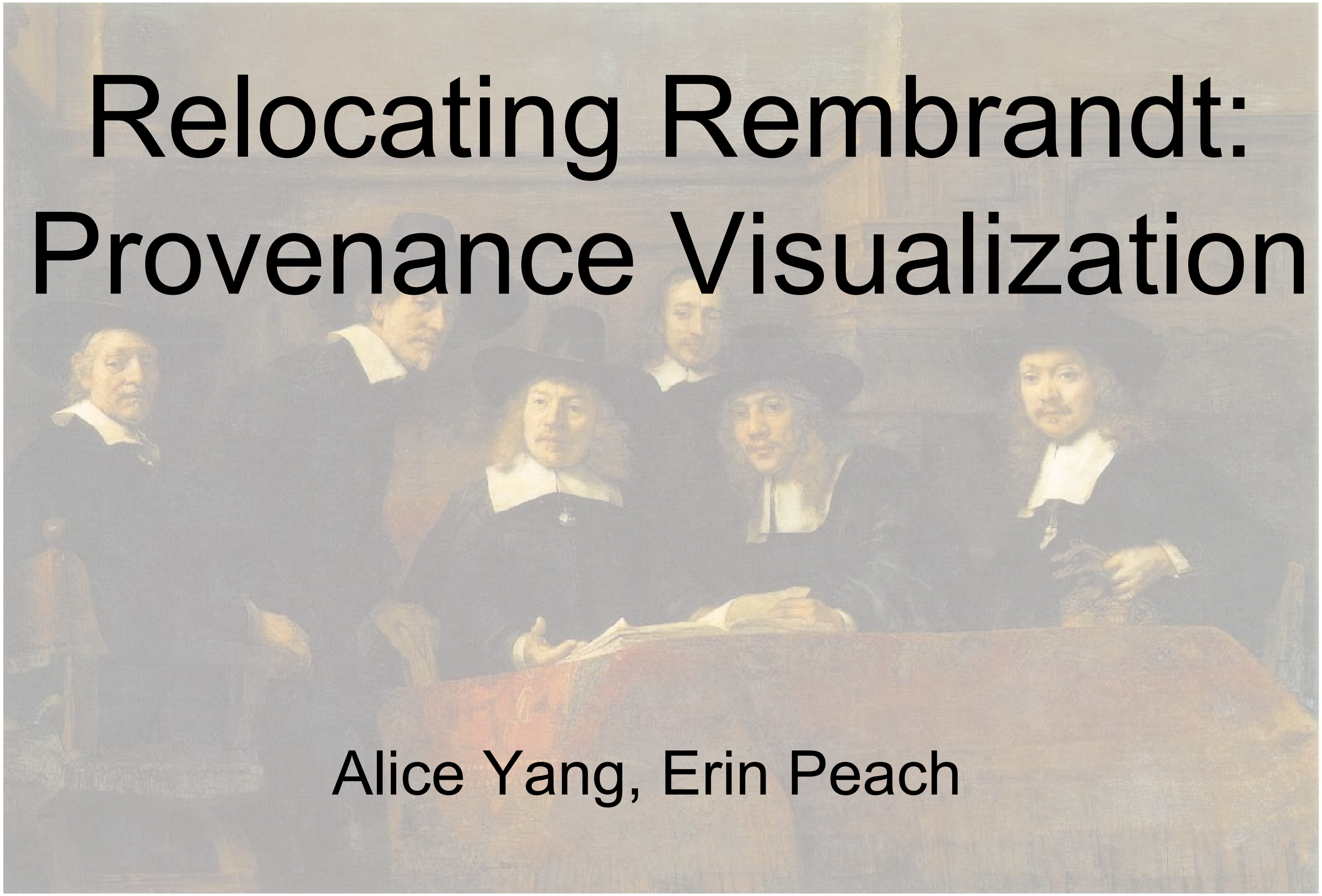
- Which icon do you like?
- Do you like the movable pop-up window idea?
- Other representation you'd like to see besides line chart and bar graph?

Literature Review

Chapter 11: The Cartogram: Value-by-Area Mapping. In Cartography: Thematic Map Design. Dent

Adaptive Composite Map Projections. Bernhard Jenny. InfoVis 2012

SEATTLE HOUSING MARKET, Skylar Olsen, 2016



Relocating Rembrandt: Provenance Visualization

Alice Yang, Erin Peach

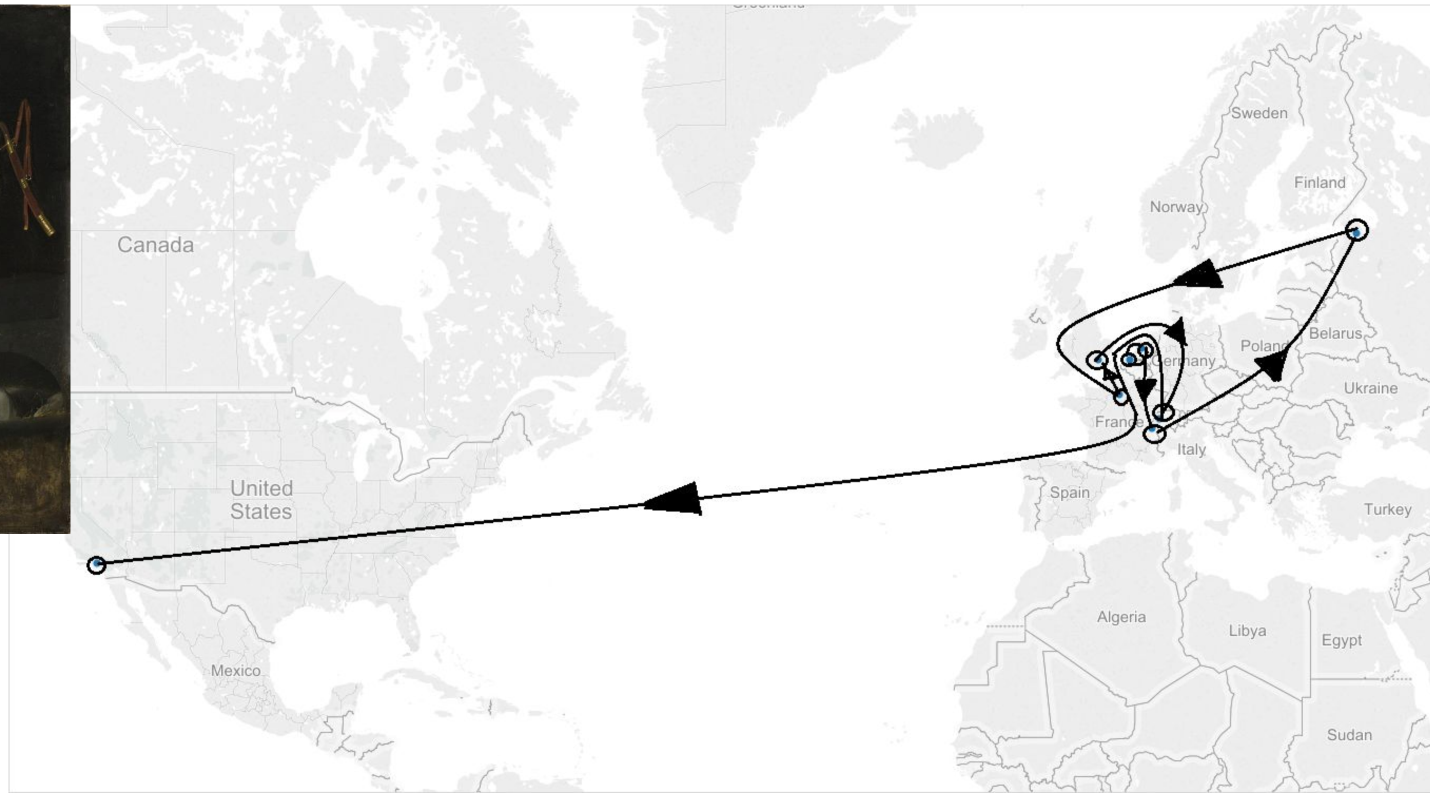
Background

- Provenance is the chronology of ownership/location of an item
- Association of Art Museum Directors and American Association of Museums
  - Guidelines for looted art in WW2
  - Museums research and publish the provenance of pieces
- Carnegie Museum of Art
  - ArtTracks -
    - Standardize provenance data
    - In-museum visualizations
    - Expand to other museums

Data Source

- Getty Provenance Index
  - Sales Catalogs, Public Records, Private Collections
- Data is difficult to parse
  - Handwritten notes, in a variety of languages
  - Narrow focus of project

Mock-Up (The Raising of Lazarus)



Questions

Art can change ownership, but remain in the same location. How to visualize?

What are the best ways to explore through time: draggable slider through time, requesting a specific year, requesting a max-min year, focusing on major events (Industrial Revolution, World Wars, etc)?

AwesomeBus

a public transit visualization focused on rider usability

Eric Zeng and Lucy Simko

Design/Usability Goals

- Provide a novel resource to Seattle-ites for exploring their city via public transit
- Increase awareness of locations connected by transit by making all routes visible and promoting exploration of available routes
- Related Work
    - Trip planning - Google Maps
    - Real time arrival data - OneBusAway
    - Real time vehicle locations - TRAVIC (tracker.geops.ch)
    - Urban planning and transit optimization
        - Transit Quality & Equity (ray-mon.com/urbandatachallenge)
        - Visualizing MBTA Data (mbtaviz.github.io)

Users: Someone who …

- has recently moved neighborhoods
- is new to Seattle
- is looking to change up their commute to use public transit
- wants to get between two neighborhoods but doesn’t have an exact destination and is willing to walk (ie a tourist)

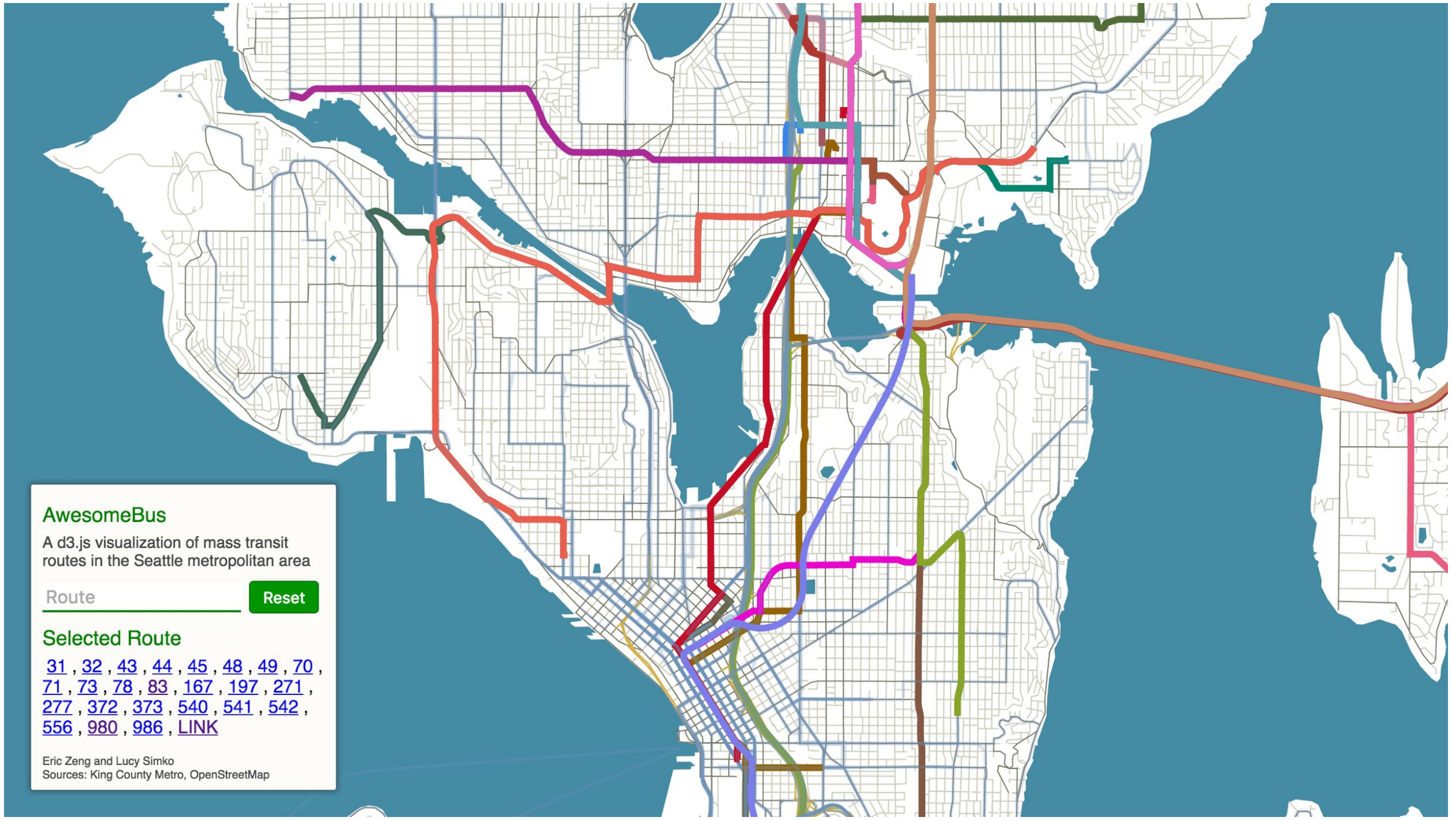
NOT Google Maps or OneBusAway

- No exact destination required
- Not for directions.
- **Exploration of transit routes encouraged!**

NOT Metro/Sound Transit's website

- Geographic detail in maps
- Can see multiple routes at once
- Can see intersecting routes
- **Exploration of transit routes encouraged!**

Example: Use a selection box to figure out what neighborhoods are best connected to UW



Planned Features

- Map panning/zooming
- Bus timing
  - Show which routes are regularly late
  - Filter buses by when they run - peak only, night owl
- Show stops on the map
- Handle overlapping routes

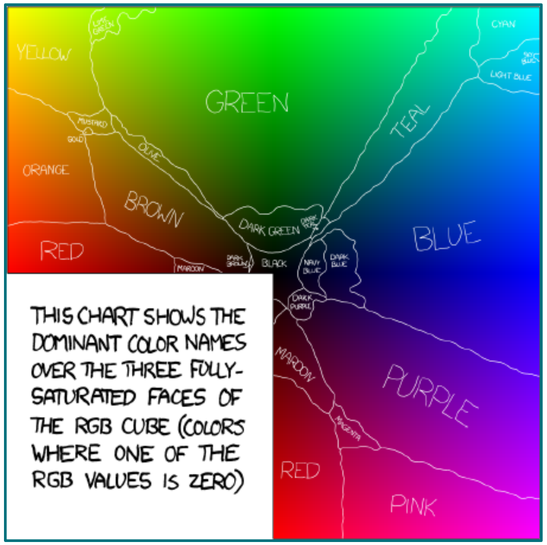
Multilingual Color Survey

Kyle Thayer, Younghoon Kim, and Gabriella Gorsky

Research Rationale

Motive

- The aim of the project is to extend previous work on mapping different patterns of color naming across languages
  - Data was sparse and certain languages were over-represented (especially English)
  - Participants were not all native speakers of the language they used to name colors
- This study will provide a more fine-grained exploration
  - We will emphasize gathering denser data from a few focal languages
  - Collecting more demographic information will allow us to filter out non-native speaker
- The new data will allow for more sophisticated mathematical modeling
  - Where are the boundaries between color groups?
  - Are there statistically significant differences in color naming trends across languages?

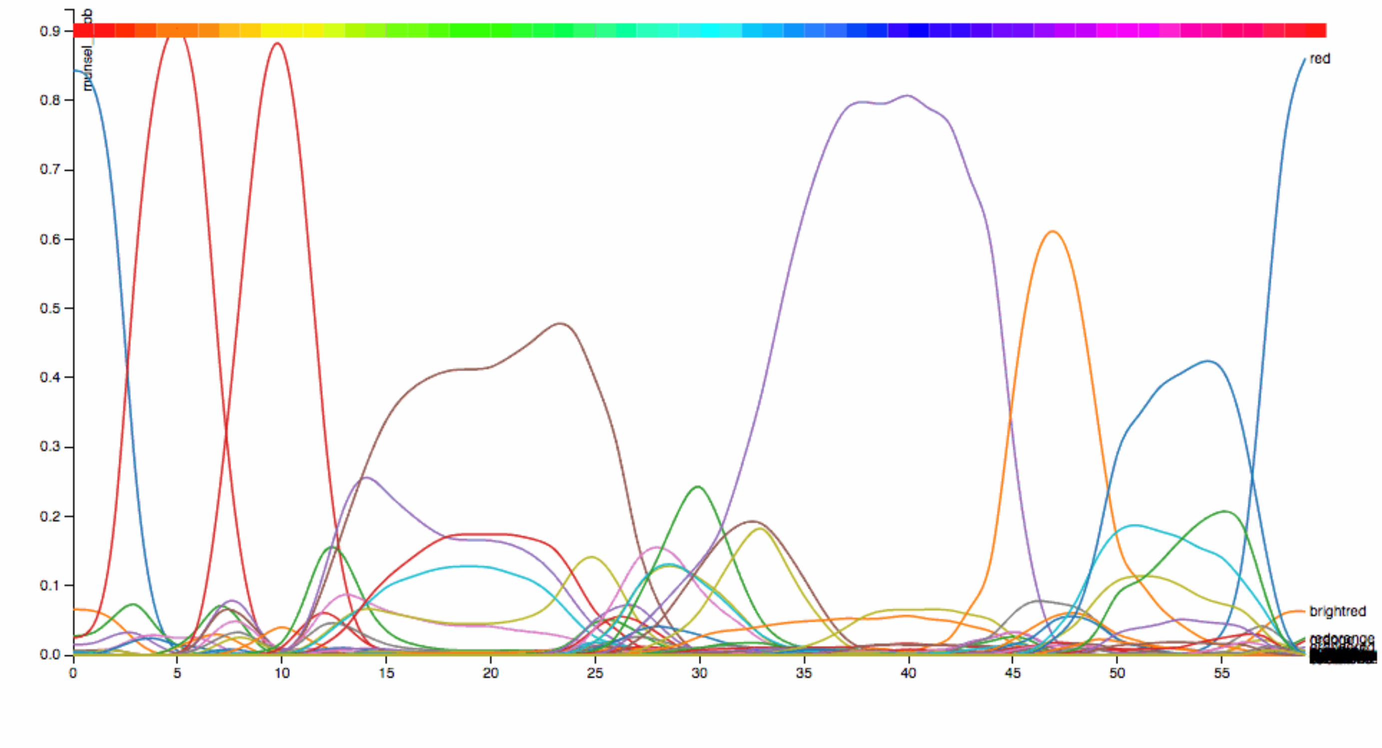


Backround

- Previous research explores how language affects color perceptions
  (“Russian Blues,” Winawer, et al., 2007; “XKCD Color Survey,” Munroe, 2010; “World Color Survey,” Berlin & Kay, 1969, Kay, et al., 2010)/

- Researchers have also explored probabilistic modeling approaches for color data
  (Chuang, Stone, & Hanrahan, 2008)

Preliminary Data Exploration



XKCD Color Survey

Study Design

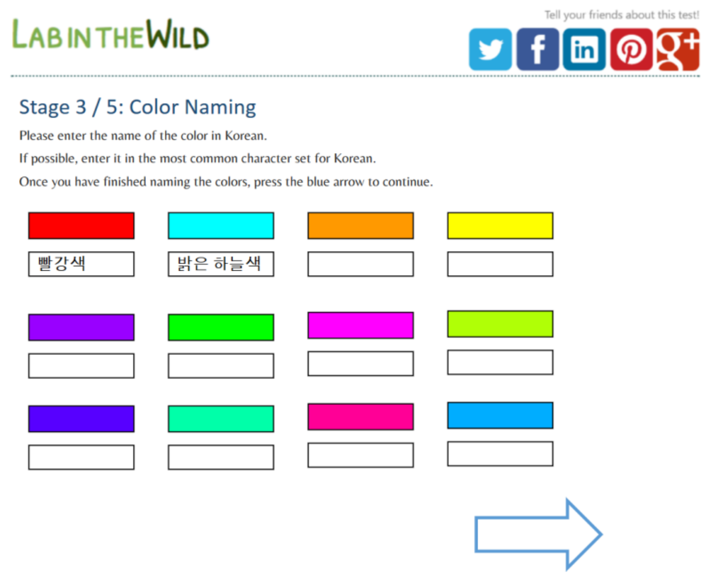
Ask Demographics

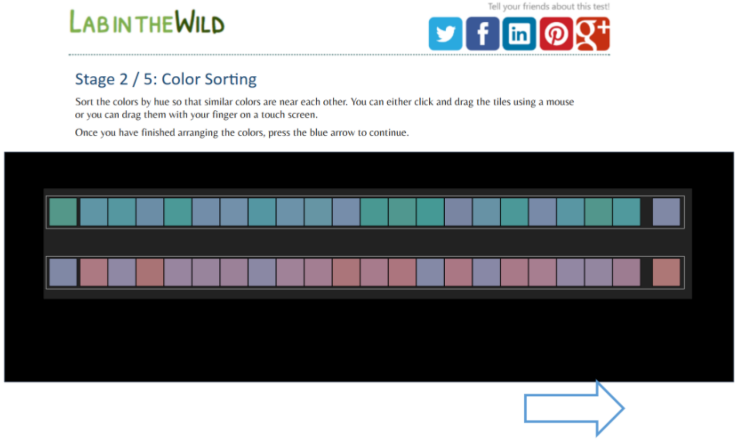
- Color Naming Task
  - Language they know
  - Colors of maximum hue

  - We collect color names

- Color Sorting
  - Colors chosen from LAB space
  - Fun Task!
  - We collect color vision score

- Results page
  - Give Color Sorting Result
  - Give Color Naming Oddness Rating
  - Share Results Option





Areas for improvement-

- Changes to data collection?
- Suggestions for encouraging participation?
- Alternative approaches to data analysis?
- Will you help us by taking our survey?
    - accessible early next week

Geographic Support for Vega-lite

Youying Lin, Vivek Paramasivam

Problem Description / Project Goals

Support projection of Lat/Long data onto a map in **Vega-Lite**

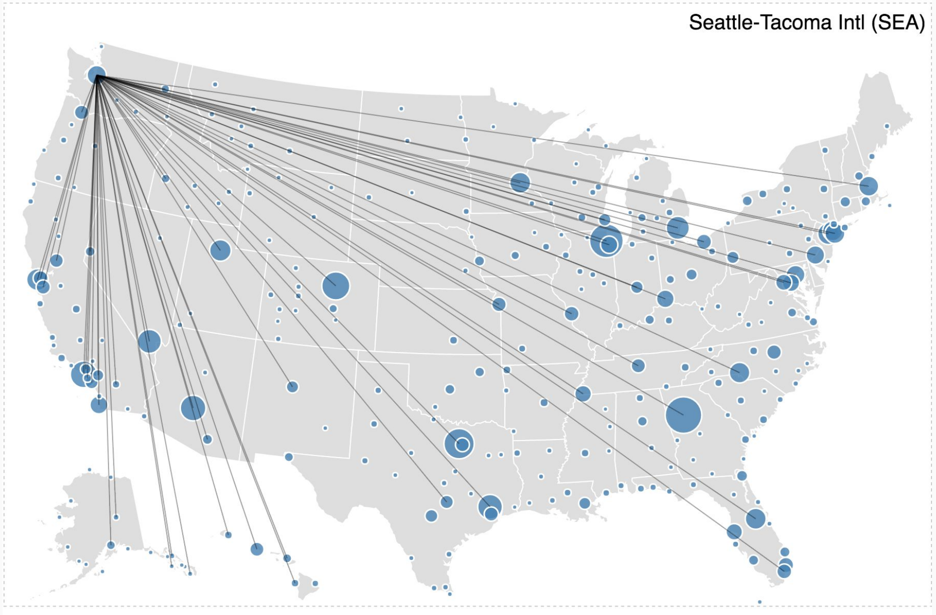
Support tiles layer for **Vega/Vega-Embed**

Prior Work

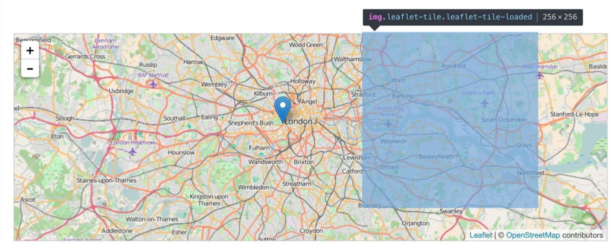
Vega, **Vega-Lite**, and **Vega-Embed** are developed by the UW Interactive Data Lab (https://github.com/vega)

Leaflet - open source JS library for interactive maps primarily developed by Vladimir Agafonkin to Mapbox (http://leafletjs.com/)

Vision


Seattle-Tacoma Intl (SEA)

+


img.leaflet-tile.leaflet-tile-loaded | 256 × 256
London
Leaflet | © OpenStreetMap contributors

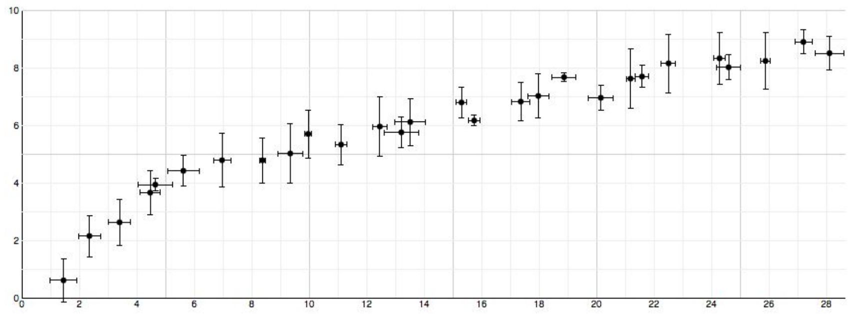
Questions for Feedback

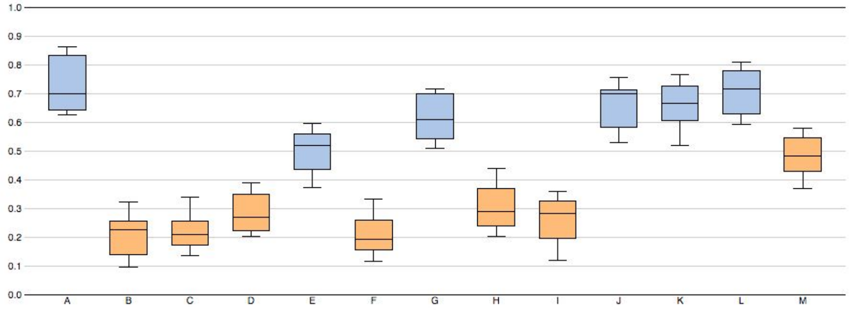
What do you like about our project? What do you think could use improvement or be more clear?

Is there a feature or idea you think is missing?

Status Report: Statistical Graphics In Vega-Lite

Ayush Saraf • 05.18.2016





Overview

Goal

Add support for statistical graphics in vega-lite grammar of graphics

- Error Bars
- Error Bands
- Box Plots

Progress - Designing Syntax

Error Bars

- Create Error Bars layering tick and rule mark.

Error Bands

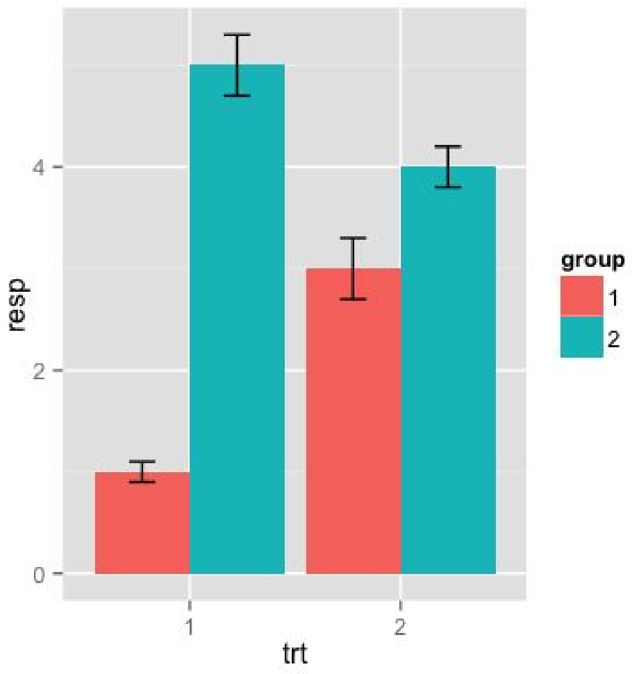
- Create Error Bars using area mark

Box Plots

- Create Box Plots By Layering 2 ticks (hinges), 2 rules (whiskers), 2 bars (boxes)

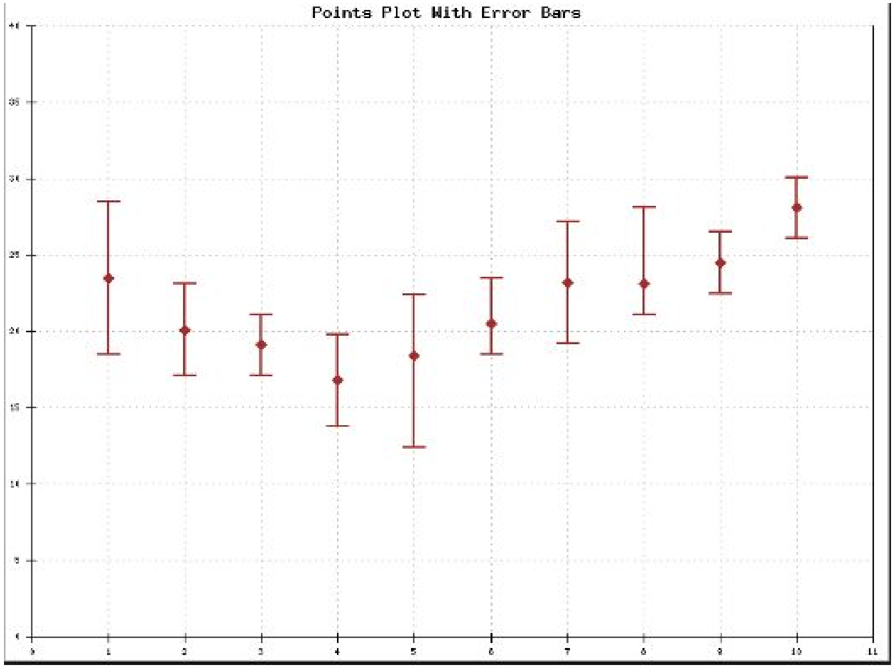
Symmetric Error Bars

```
{
  "data": { “url” : “data.csv” },
  "mark": "rule",
  "encoding": {
    "x": {"field": "a","type": "ordinal"},
    "y": {"field": "b","type": "quantitative",
"aggregate": "mean"},
    "spanY": {"field": "b", "type": "quantitative",
"aggregate" : "stdev"}
}
```



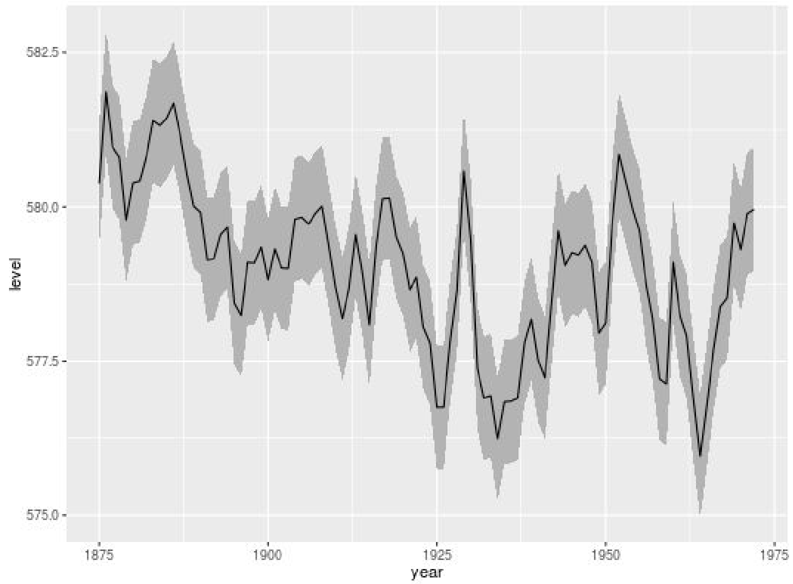
Asymmetric Error Bars

```
{
  "data": { “url” : “data.csv” },
  "mark": "rule",
  "encoding": {
    "x": {"field": "a","type": "ordinal"},
    "y": {"field": "b","type": "quantitative",
"aggregate": "q1"},
    "y2": {"field": "b", "type":
"quantitative", "aggregate" : "q3"}
}
```



Error Bands

```
{
  "data": { “url” : “data.csv” },
  "mark": "area",
  "encoding": {
    "x": {"field": "a","type": "temporal"},
    "y": {"field": "b","type": "quantitative",
"aggregate": "q1"},
    "y2": {"field": "b", "type": "quantitative",
"aggregate" : "q3"}
}
```

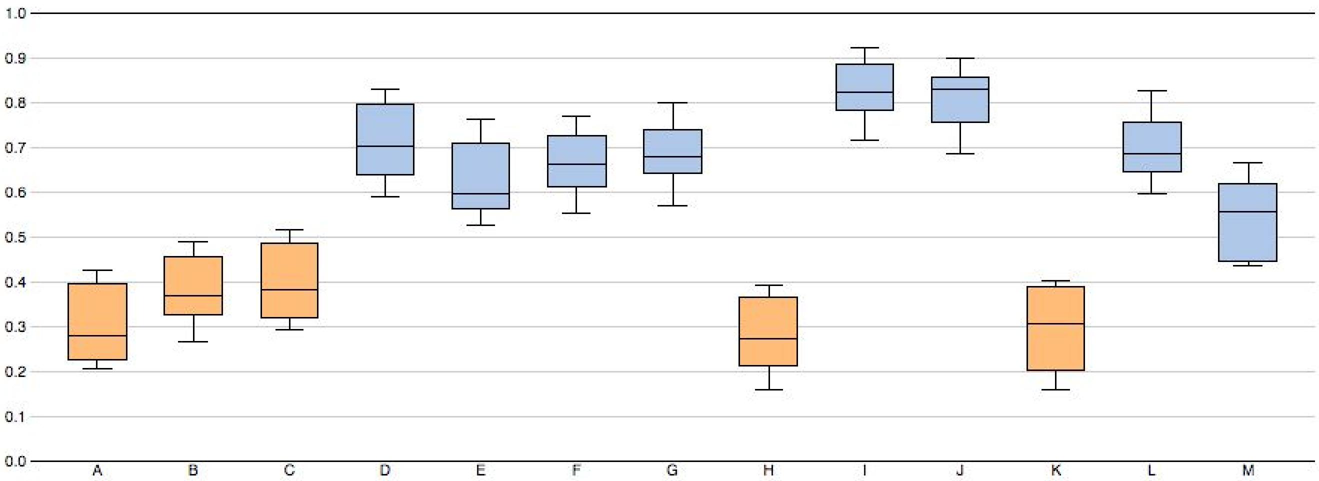


Box Plots

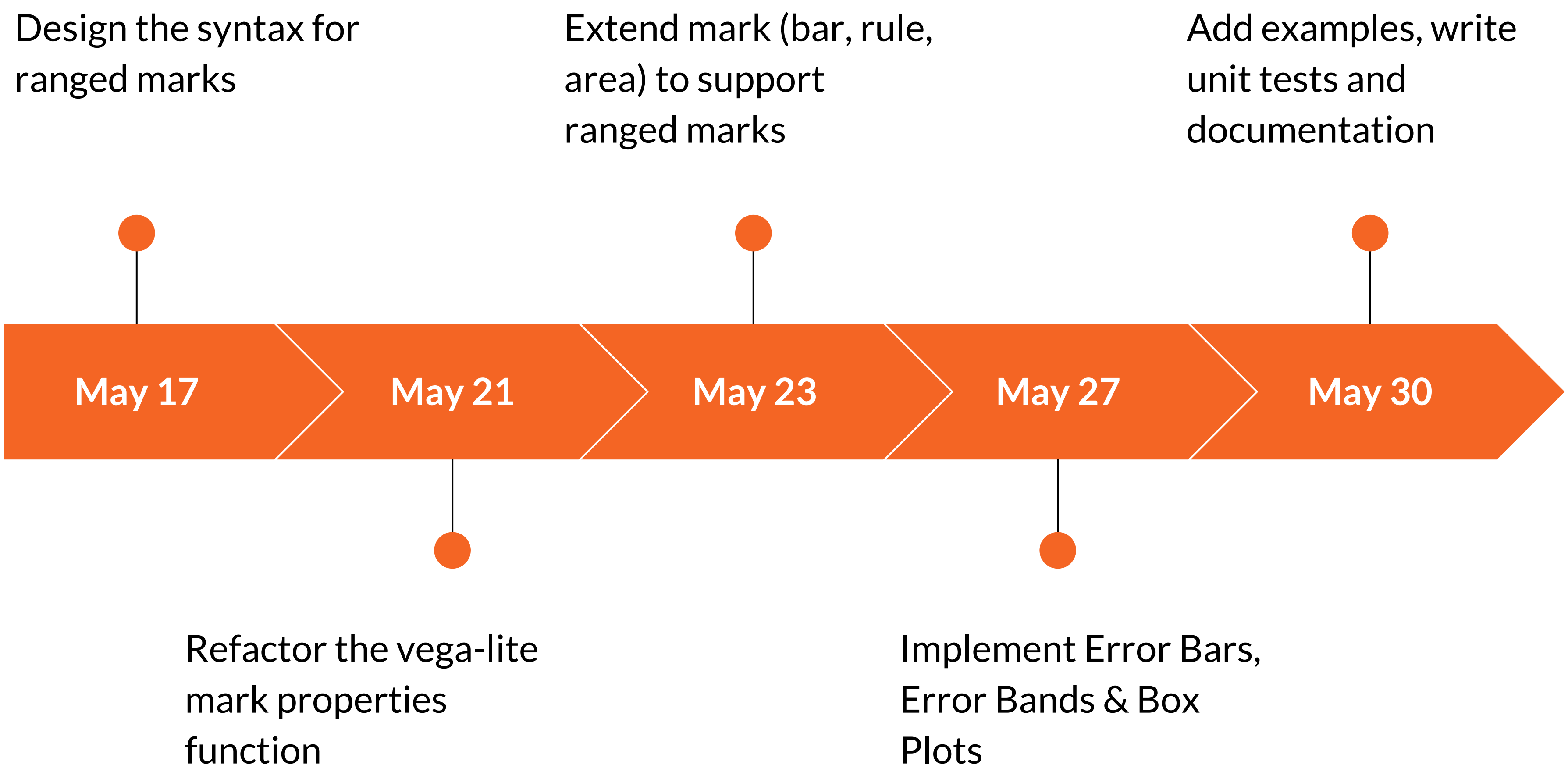
```
{
  "data" : { "url" : "data.csv" },
  "layers": [{
    "mark": "tick",
    "encoding": {
      "x": {"field": "a", "type": "ordinal"},
      "y": {"field": "b", "type": "quantitative", "aggregate":
"min"}
    }
  },
  {
    "mark": "tick",
    "encoding": {
      "x": {"field": "a", "type": "ordinal"},
      "y": {"field": "b", "type": "quantitative", "aggregate":
"max"}
    }
  },
  {
    "mark": "rule",
    "encoding": {
      "x": {"field": "a", "type": "ordinal"},
      "y": {"field": "b", "type": "quantitative", "aggregate":
"min"},
      "y2": {"field": "b", "type": "quantitative", "aggregate":
"q1"}
    }
  },
```

```
  {
    "mark": "rule",
    "encoding": {
      "x": {"field": "a", "type": "ordinal"},
      "y": {"field": "b", "type": "quantitative", "aggregate": "q3"},
      "y2": {"field": "b", "type": "quantitative", "aggregate": "max"}
    }
  },
  {
    "mark": "bar",
    "encoding": {
      "x": {"field": "a", "type": "ordinal"},
      "y": {"field": "b", "type": "quantitative", "aggregate": "q1"},
      "y2": {"field": "b", "type": "quantitative", "aggregate": "median"}
    }
  },
  {
    "mark": "rule",
    "encoding": {
      "x": {"field": "a", "type": "ordinal"},
      "y": {"field": "b", "type": "quantitative", "aggregate": "median"},
      "y2": {"field": "b", "type": "quantitative", "aggregate": "q3"}
    }
  }]
}
```

Box Plots



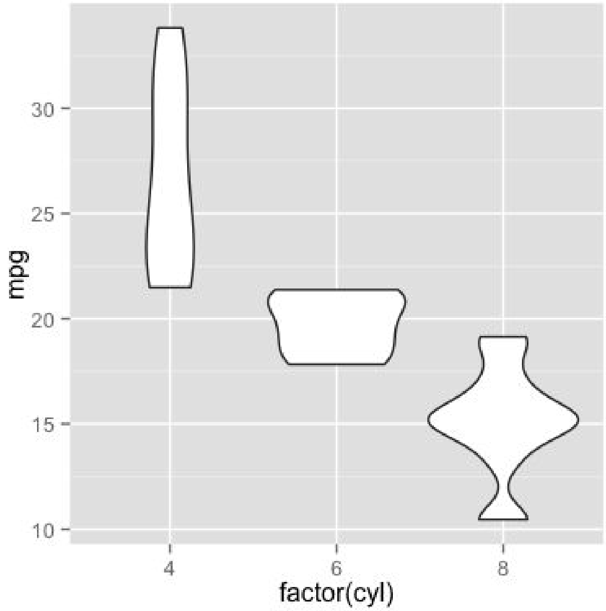
Schedule


Design the syntax for ranged marks
Extend mark (bar, rule, area) to support ranged marks
Add examples, write unit tests and documentation
May 17
May 21
May 23
May 27
May 30
Refactor the vega-lite mark properties function
Implement Error Bars, Error Bands & Box Plots

Nice To Have

Density & Violin Plots

Similar process as ranged marks but requires additional vega transforms.



Thank You

Any Suggestions?

Email: ayush29f@uw.edu

Labeling in Vega and Vega-lite

Will Strimling

Vega

“Vega is a visualization grammar, a declarative format for creating, saving and sharing interactive visualization designs.”

- The Vega documentation

Think: rendering engine

Vega

- Writing Vega is difficult and verbose
- Has reactive geometry where marks can get their data from other marks
- Has data transforms to manipulate data before charting (using expressions)

Vega-lite

“Vega-Lite specifications describe visualizations as mappings from data to properties of graphical marks (e.g., points or bars). It automatically produces visualization components including axes, legends, and scales. It then determines properties of these components based on a set of carefully designed rules.”

- The Vega-lite website

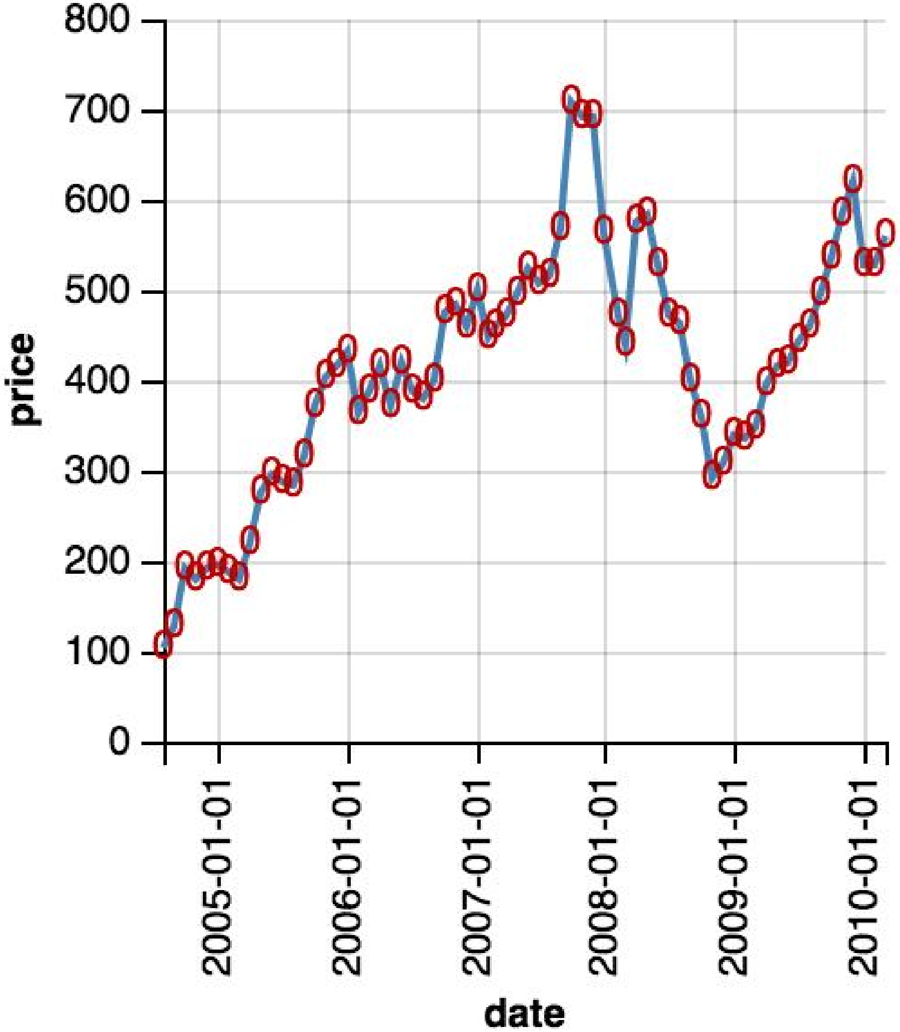
Vega-lite

- Writing Vega-lite is simple and succinct
- Compiles to Vega
- Doesn’t have access to the data
- Recently introduced “layering”

The Problem: Labeling

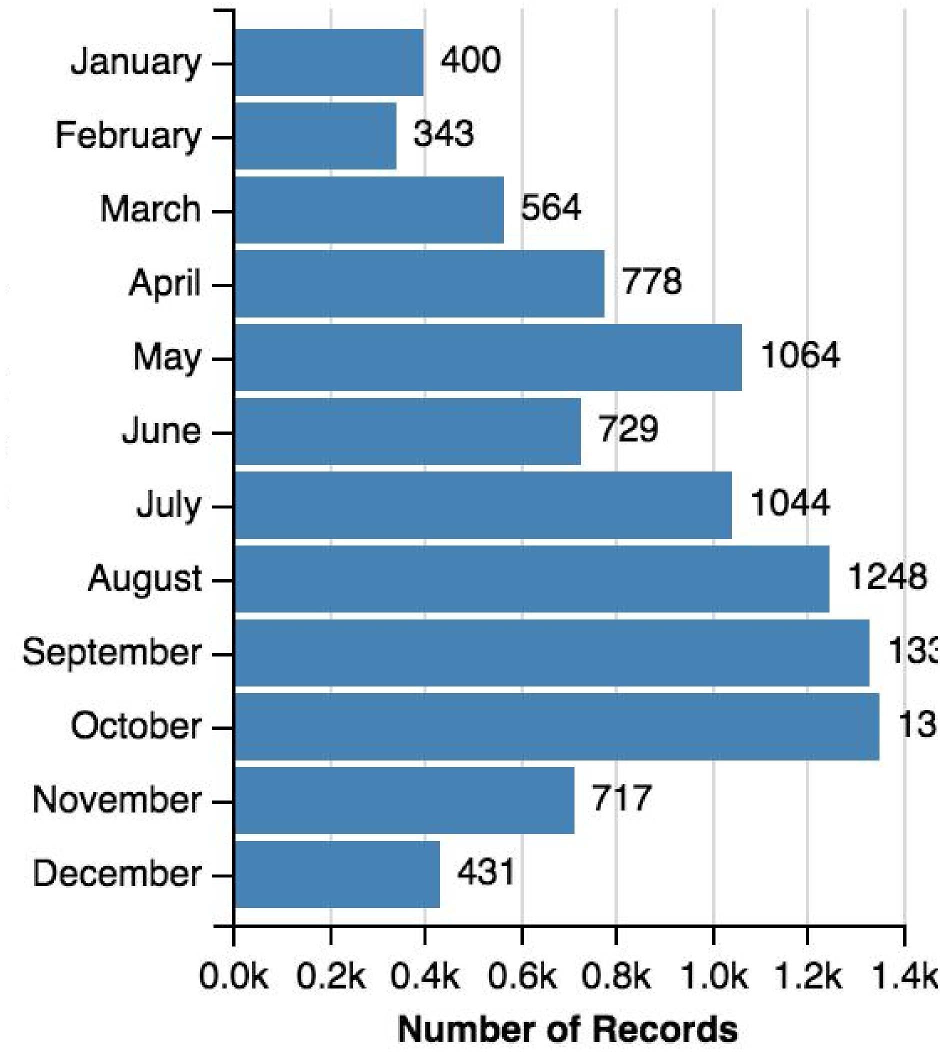
- Writing Vega is **Hard** and doing labeling in it requires a working knowledge of reactive geometry (undocumented)
- Lack of Vega transforms for deterministic labeling (minimizing occlusion, deciding direction, setting offset)
- No Vega-lite implementation of labeling
    - No grammar for this (Is it encoding? Is it a layer?)
    - No compilation steps completed
- Labeling needs vary by mark type

Wanted



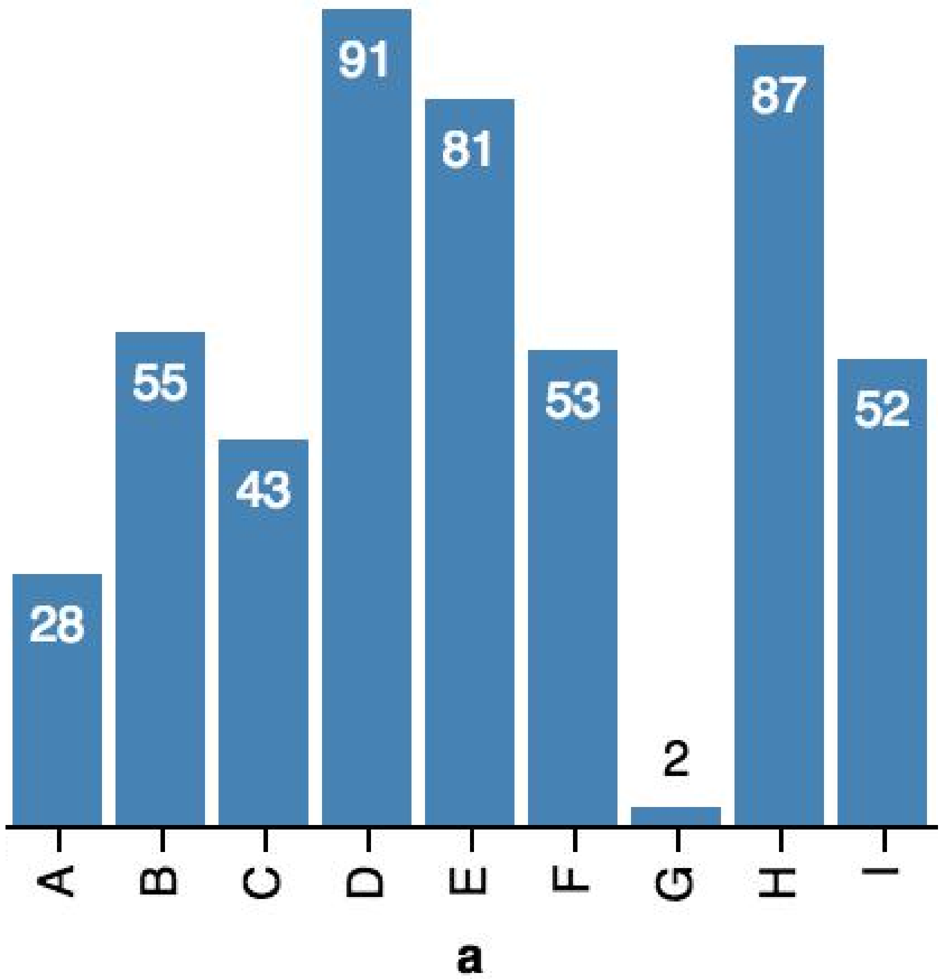
800
700
600
500
400
300
200
100
0
price
2005-01-01
2006-01-01
2007-01-01
2008-01-01
2009-01-01
2010-01-01
date

Wanted



January
February
March
April
May
June
July
August
September
October
November
December
400
343
564
778
1064
729
1044
1248
717
431
0.0k
0.2k
0.4k
0.6k
0.8k
1.0k
1.2k
Number of Records

Wanted



Implementation

Vega transforms

- Deterministic positions for labels
  - Intelligent anchor choice
  - automatic offset
  - Minimize occlusion

Implementation

Vega-lite Complication

- Design and implement new features in Vega-lite to allow for easy labeling of points

Implementation

Layering

```
{
  "layers": [{
    "mark": "label",
    "data": {"from": "ham"},
    "encoding": {
      "anchor": {"value": "top-left"},
      "offset": 10,
      "text": {"field": "price"},
      "color": {"field": "ham.color"}
    }
  }, {
    "description": "Google's stock price over time.",
    "data": {"url": "data/stocks.csv", "formatType":"csv"},
    "transform": {"filter": "datum.symbol==='GOOG'"},
    "name": "ham",
    "mark": "line",
    "encoding": {
      "x": {"field": "date", "type": "temporal"},
      "y": {"field": "price", "type": "quantitative"},
    }
  }]
}
```

Implementation

Per Spec

```
{
  "description": "Google's stock price over
  "data": {"url": "data/stocks.csv", "format
  "transform": {"filter": "datum.symbol==='G
  "mark": "line",
  "encoding": {
    "x": {"field": "date", "type": "temporal
    "y": {"field": "price", "type": "quantit
  }
  "label": {
    "field": "price",
    "anchor":  "top-left",
    "offset": 10,
    "color": "#sdfsdf/match/auto"
  }
}
```

Prior works

- Reactive geometry in Vega

  Different because this implementation will be more about expressiveness and will incorporate transforms to do positioning well

- D3 Labeling (especially paper and library by Evan Wang)
  - Currently lacks useful features
    - doesn’t hide occluding labels
    - only implemented for scatter plots
  - Not implemented in Vega/Vega-lite

Progress

1. Preparing
    a. Develop understanding of Vega and reactive geometry
    b. Create Vega specs of labeling, understand transforms
    c. Design specs, get approval from Vega/Vega-lite team
2. Vega
    a. Decide on algorithm for intelligently placing labels (in progress)
    b. Design transforms to implement these algorithms
    c. Implement transforms in Vega, extend as needed (for access to Scenegraph)
3. Vega-lite
    a. Implement layering ‘text’ mark version
    b. Implement in-spec version

Questions

When is appropriate to hide/show label?

How far can a label be from a mark?

Are tails useful?

What features would you want in labeling?

Explore my Strava Club

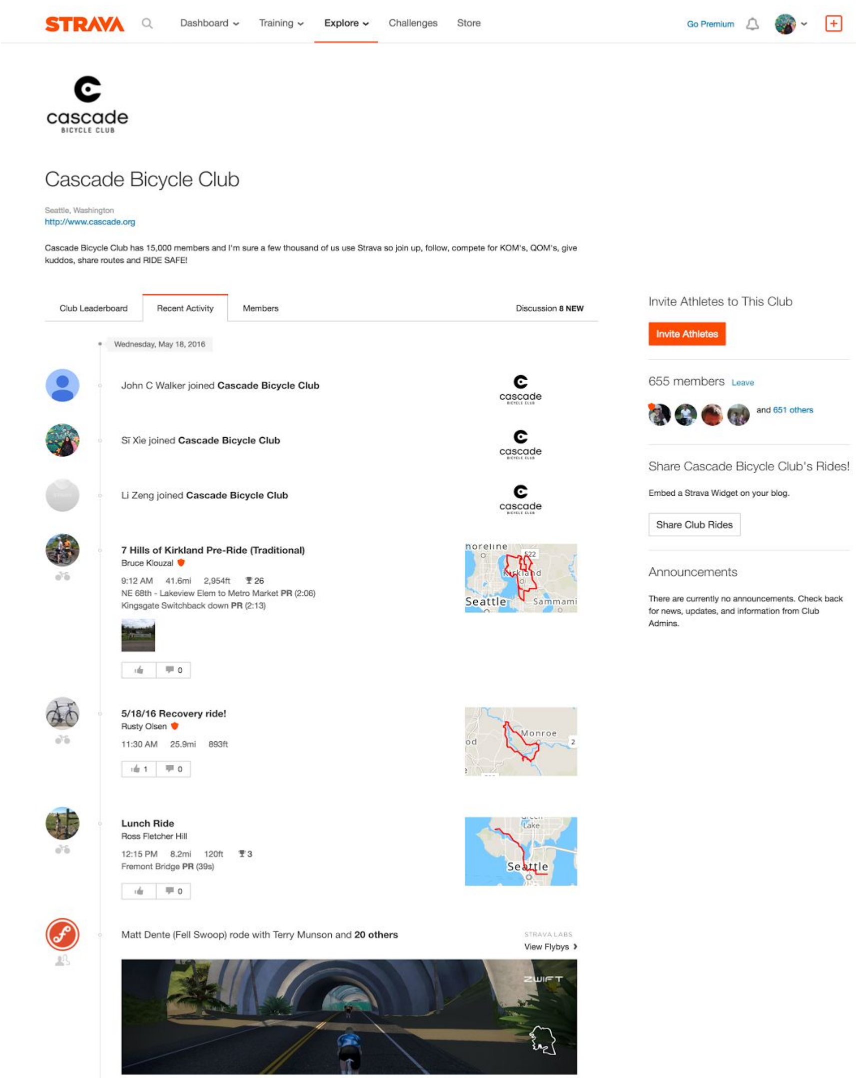
CSE 512 Progress Report

About Strava Club

Strava is the most popular social network and application for tracking cycling and running.

Strava club is a social community for people who are interested in connecting with similar athletes and participating in group activities.

Additionally, Strava club members can schedule events and create group rides and runs.



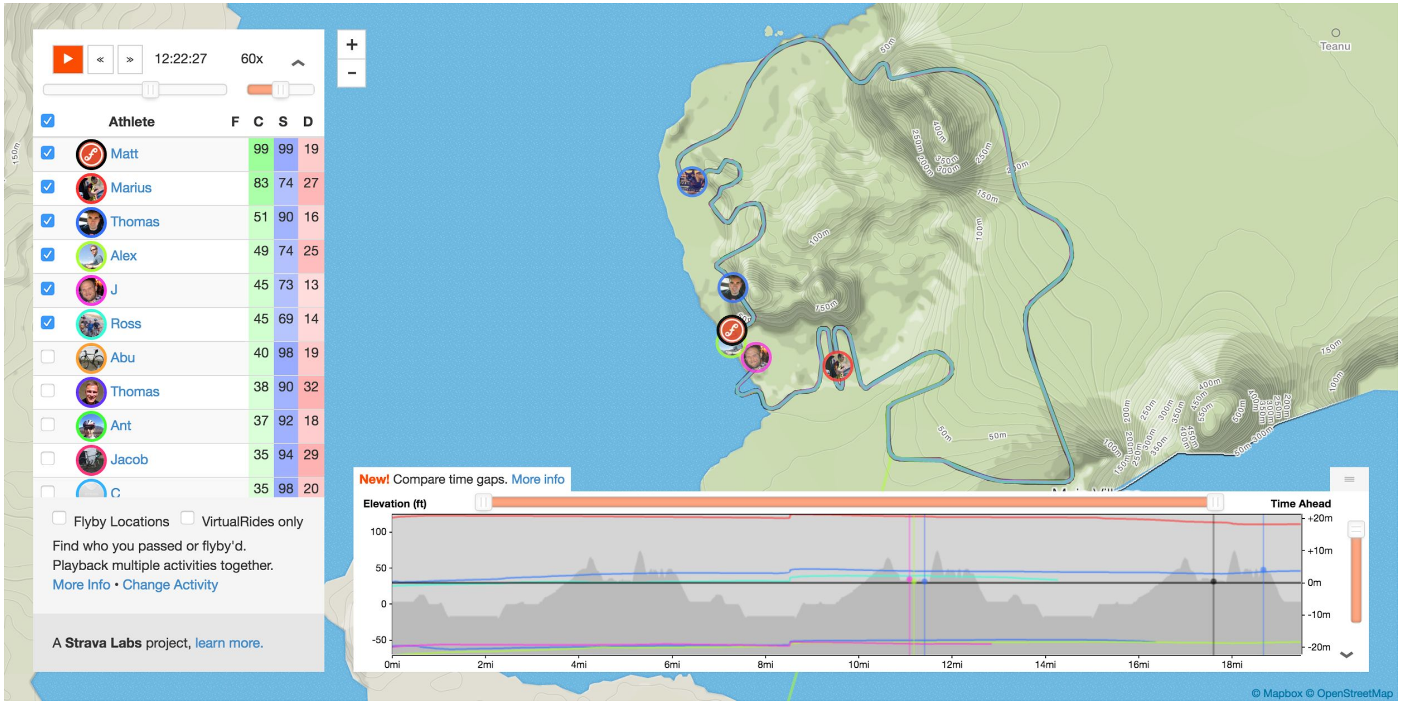
The Problem and Motivation

Currently, Strava lacks of providing meaningful aggregated club members' activity information that could have been beneficial for the three scenarios:

1. **Club owners** can recruit members and expand club memberships easily
2. **Prospective members** can see if the club has members active in their area
3. **Current members**
   a. can discover new partners in the club that are active along similar routes
   b. can discover new routes based on other club members' activities

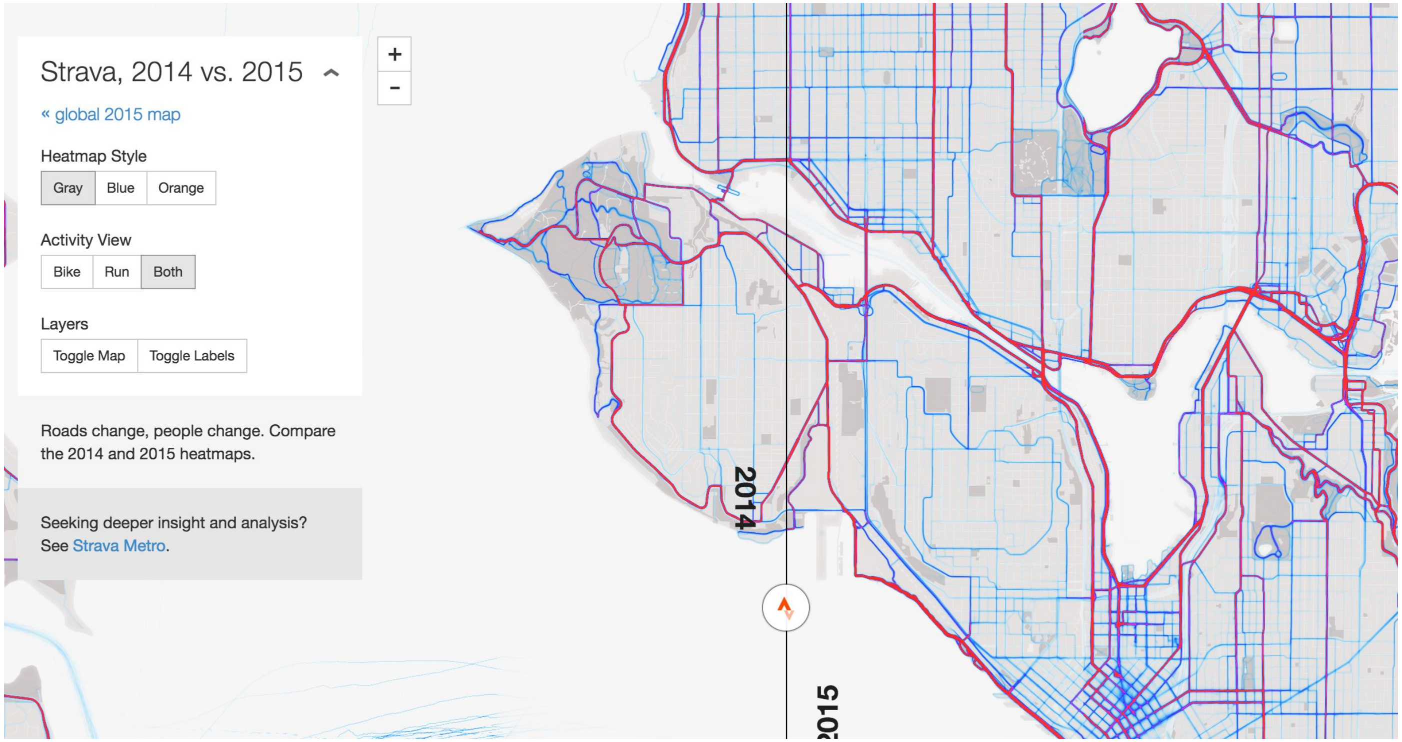
Activity Review: Flyby

Visualizing individual and group activity performance across time and distance.



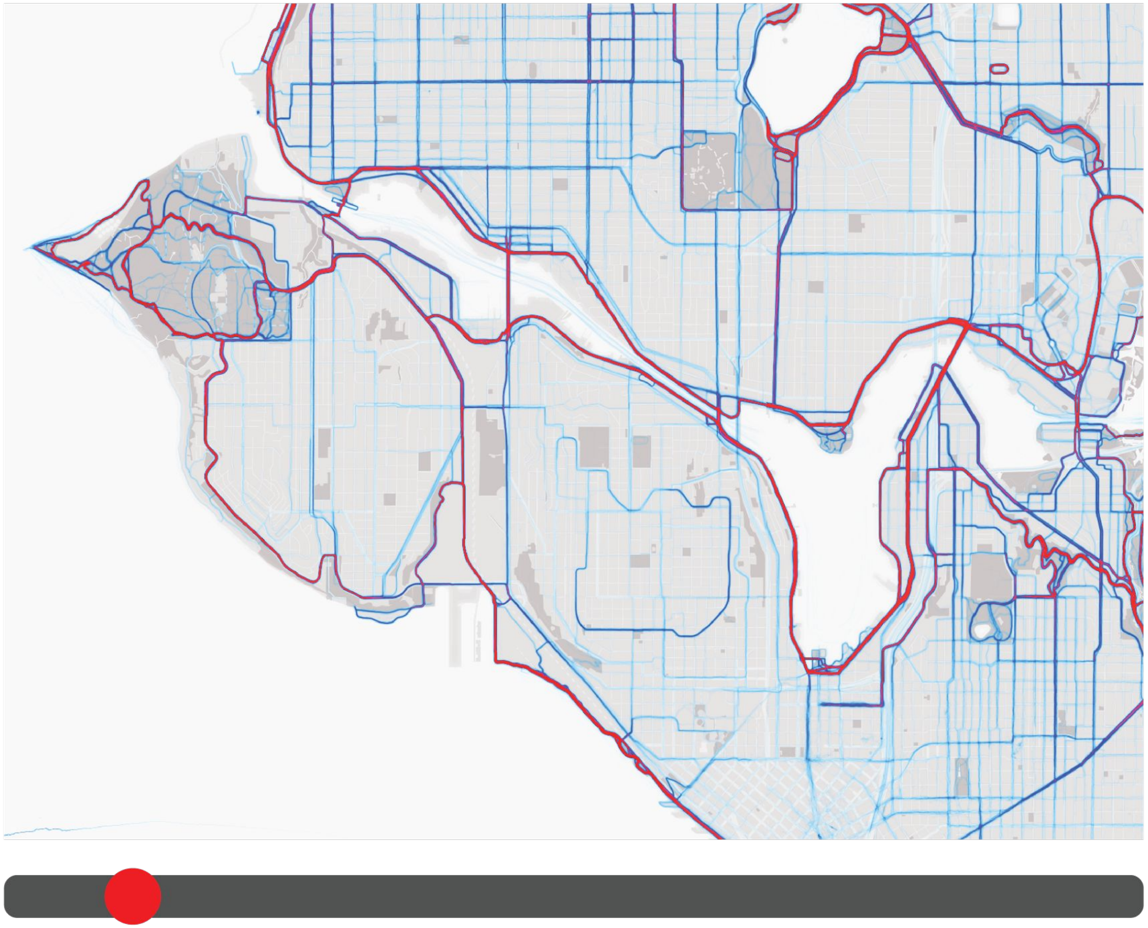
Activity Review: Strava Heat Map

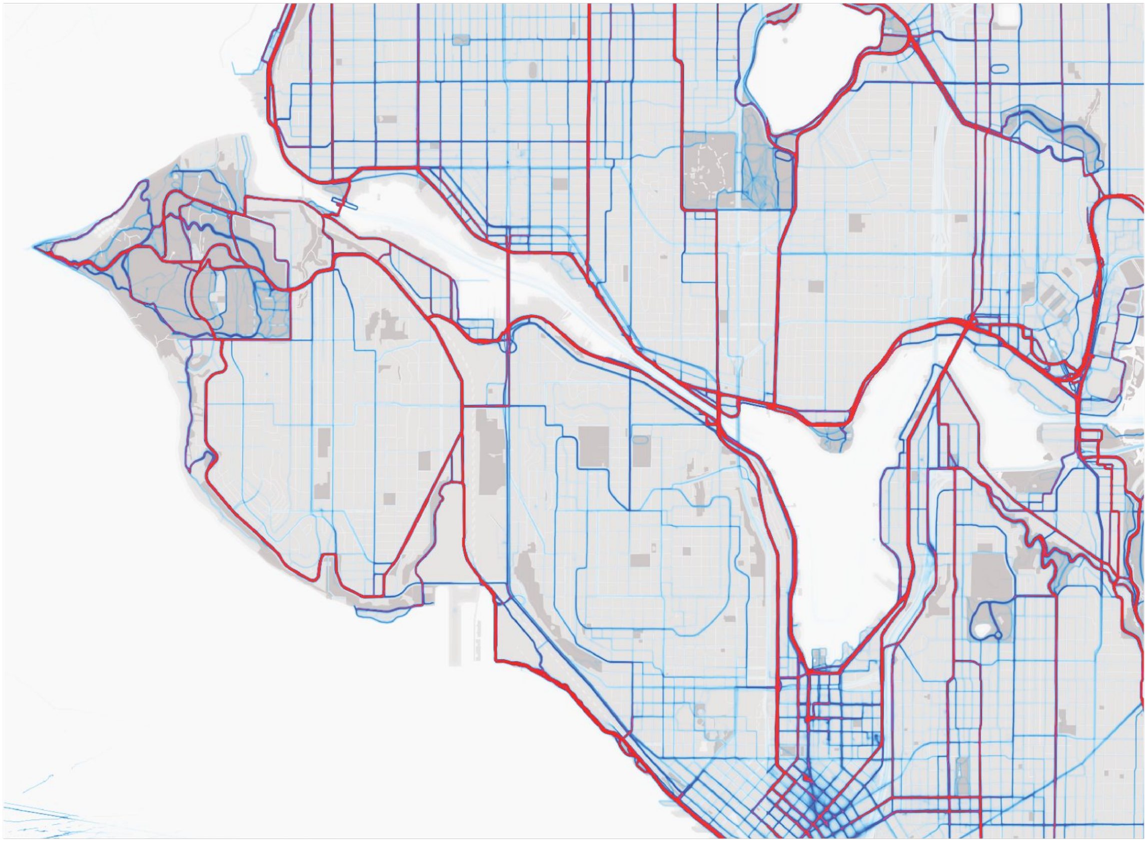
Displays the different heat maps of routes between 2014 and 2015.



Current Progress

- Explored small sample (3,000) of users data
- Pivoted from original idea
  - Geo-Social network visualization
  - After discussing with Jeff, we decided to refocus our audience to be the Strava users
- Joined Strava and an active cycling club based in Seattle, WA
- Collecting data from API and scraping data
- Identified Strava users to interview
- Began designing basic interaction model





10

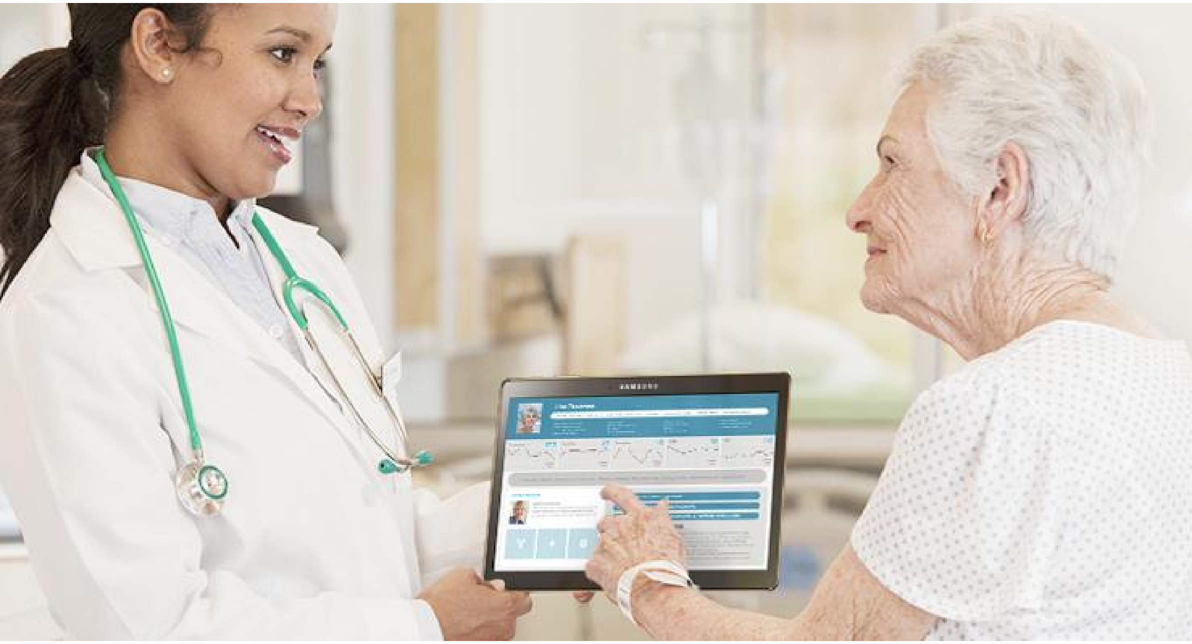
Questions and Feedback

Considering Data and Privacy

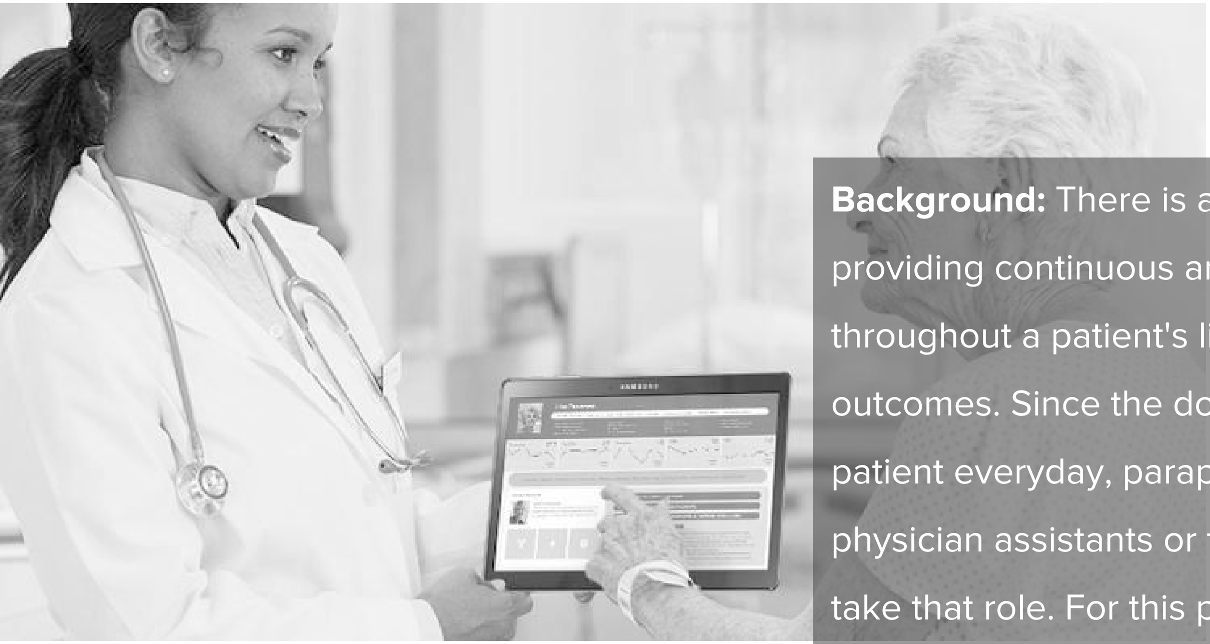
- Protecting private data: people who commute to work
- Different visualizations and resolutions to non-club vs. club members?

Applying case study of a bicycle community data to other online health communities

- Recommendation for visualization literatures?
- Things to look out for?

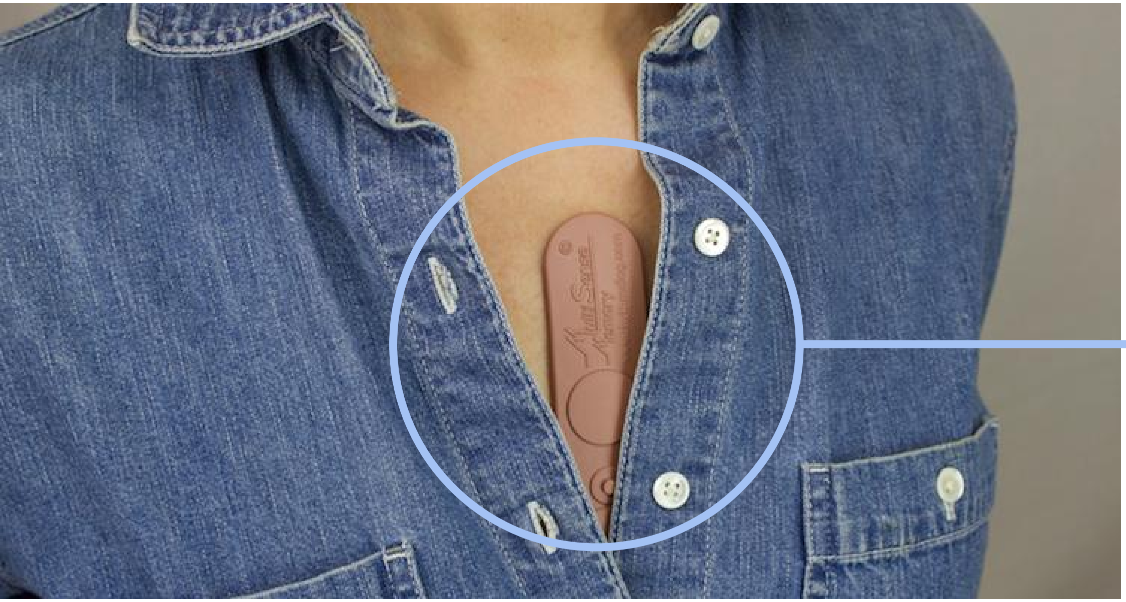


Our team addresses a growing yet unmet need to effectively visualizing sensor and body monitor data for clinical insights.



Background: There is an increasing need in providing continuous and coordinated care throughout a patient's lifetime to maximize health outcomes. Since the doctor could not monitor the patient everyday, paraprofessionals, such as physician assistants or family members, will likely take that role. For this population, they will need easy-to-understand visual tools.

OUR MISSION: We are designing a data visualization interface that both clinicians and patient/paraprofessionals could easily understand.



MultiSense™ Strips: Clinical Quality Data in a Small, Wearable Strip

- Blood oxygen saturation
- ECG
- Heart rate and heart rate variability
- Respiration rate and relative depth of respiration
- Body position
- Relative physical activity
- Temperature

Scientific and accurate data presentation.

Hard to understand for non professionals.

Easy to understand for all population.

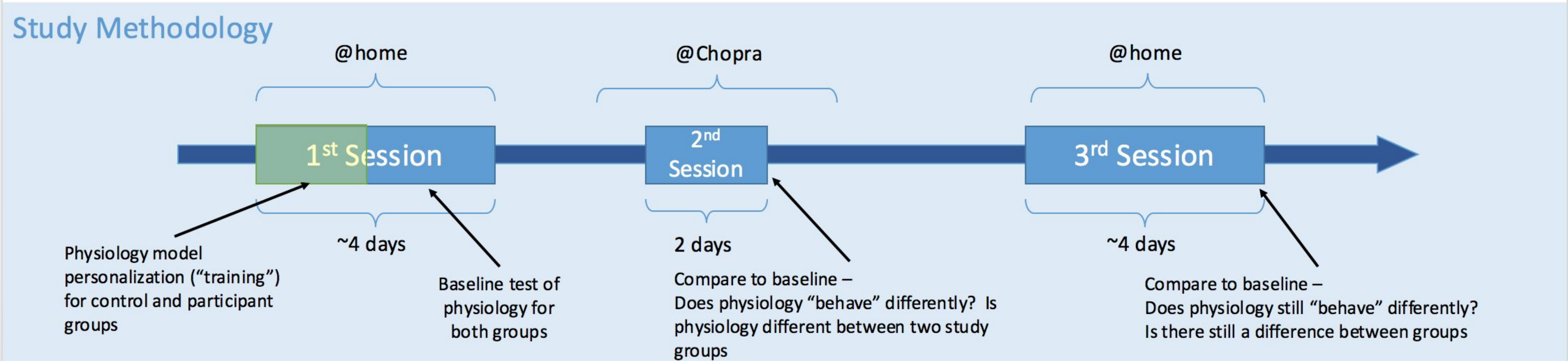
Lack scientific and accurate data presentation.

OUR DATA VISUALIZATION:

Easy to understand for all population.

Scientific and accurate data presentation.

OUR DATA



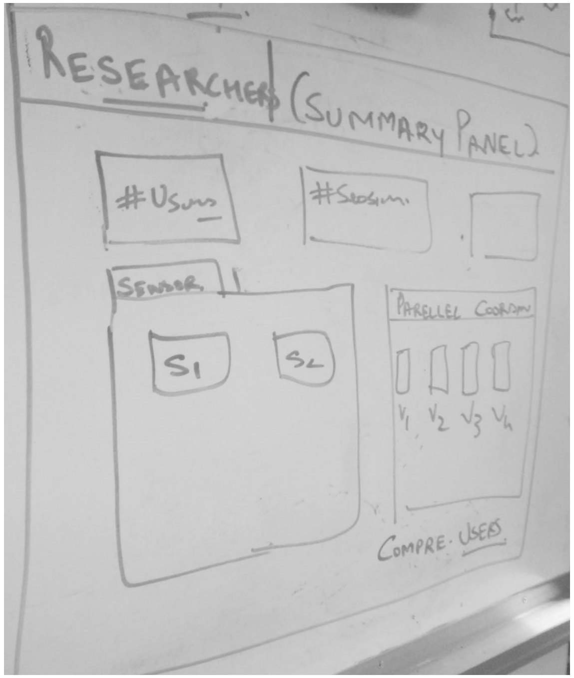
Due to privacy concerns, we could not use real patient data. Hence, we obtained our data from a research study, which simulated a clinical intervention.
In this study, researchers used MultiSense™ Strips on 94 people. Half of the participants enrolled in a yoga session (2nd session) while another half served as the control condition.
During the study, 23 columns of sensor data were generated. We will visualize those sensor data in our project.

CURRENT PROGRESS

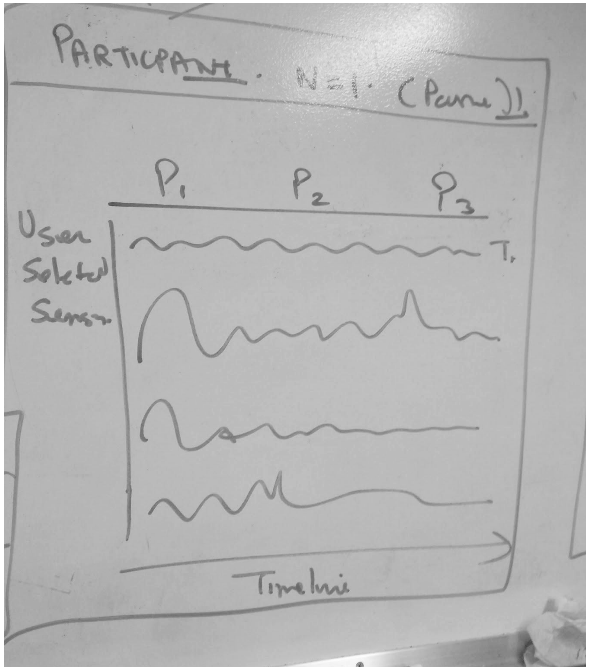
SKETCHES

Our visualization would serve two populations: professionals(i.e. researchers) and non-professionals (i.e. patient).

We generated initial dashboard designs.

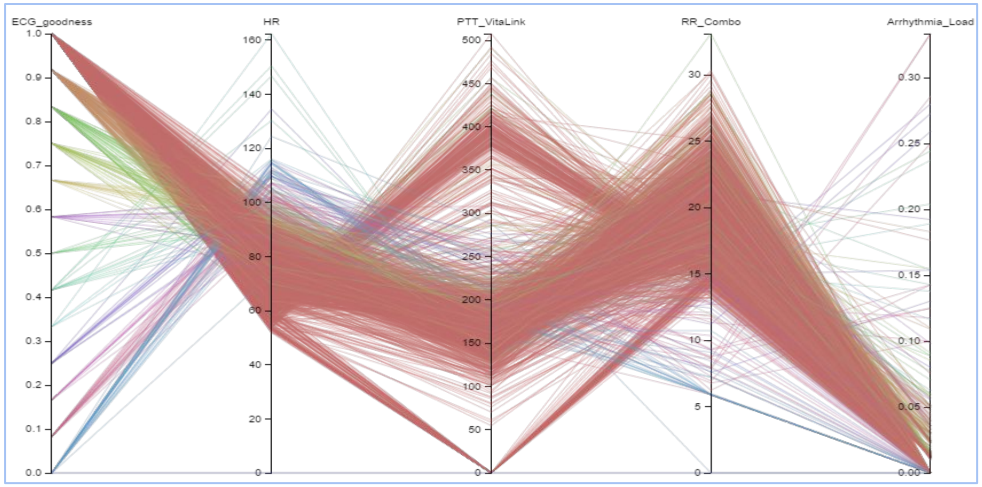


Audience: Researcher/clinician
Here all the data would be represented in a such a way that researcher can easily analyze variables. For example we would show the parallel co-ordinates visualization of different variables ,single variable histograms etc.

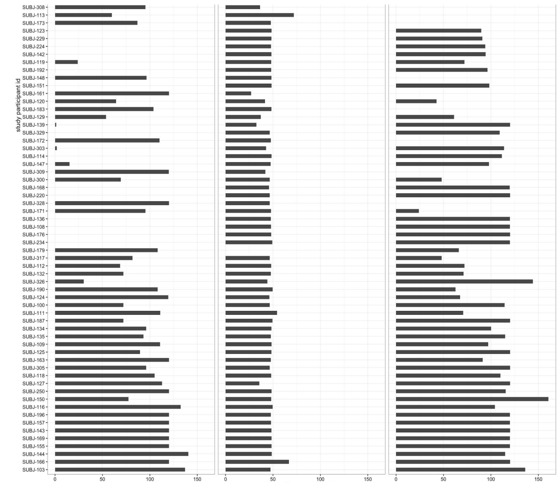


Audience: Patient/para-professionals

We would be creating a time series visualization where user can select which variable he/she wants see and compares with other variables in same time series.



This is the the sample parallel coordinates visualization showing data for one patient with various variables.



This shows the number of hrs participants used the MultiSense device in three phases. Based on which we can select data of particular participants.

CURRENT PROGRESS

We conducted some initial analysis to explore the limitation and correlation of the data.
We experimented some possible visualizations to explore the relationship among different variables.