Airline Popularity at SFO

Alpha Release

CS 360-01
By Darren Dunn
ddunn@dons.usfca.edu
Repository
Website

Background and Motivation

I am an avid aviation geek! I love airplanes and I love flying. I would even park near the airport and just watch planes take off and land. Seeing how a large piece of aluminum just effortlessly takes off into the sky amazes me. If I were not a CS major, I would be working towards becoming a commercial pilot.

Growing up, I have traveled across the world with my parents. I would always be fascinated seeing all the different types of planes sitting at the gates of San Francisco International Airport (SFO). I slowly realized that there was a pattern of certain airlines being more popular than others. When there are more airlines present at an airport, it is because the airline is trying to meet a higher demand of passengers. Therefore, the more passengers an airline serves at a particular airport, the more popular that airline is at that specific airport. I cannot help but notice how many United Airlines planes there are at SFO at any given time.

Now that I am taking Data Visualization, I want to take SFO's data on airlines that have visited the airport over the past several years and create visuals to help people see which airlines have visited SFO the most and have served the most passengers.

Project Objectives

Provide the user with an overview of SFO. This will include:

1. Where the airport is on a map of the Bay Area, background information, and key statistics such as number of passengers served every year.
2. Historical facts such as the opening of new terminals or renovation of existing terminals

3. 4 types of visualizations representing the popularity (based on number of passengers airline served every month) of airlines at the airport
4. Give a broad overview of passenger activity at SFO

Data

The data will be obtained from the city and county of San Francisco's public data resource, DataSF. The data in the dataset is provided by SFO. This data will be obtained in the form of csv files.
https://www.flysfo.com/media/facts-statistics/air-traffic-statistics

Data Processing

The csv files will be processed using Microsoft Excel. Using my current knowledge of Excel, I will extract the following data:

- Airline name
- Airlines IATA code (ex: United Airlines = UA)
- Activity period (ex: 202106 = 6/2021 = month/year)
- Terminal type
- Boarding area
- Passenger count
- Activity (Enplaned or Deplaned)

Based on my initial analysis of the data, there will not be an extensive need for data clean up. The data is neat and organized as is.

Visualization Design

Tree map and Circle Packing
Provide the user with an intuitive visualization of data from SFO. To make the visualization appealing and relevant to the topic, an outline of the current SFO airport will be used and data will be mapped to areas of the airport based on boarding area in the form of a tree map and circle packing.

There will be layers on the silhouette of SFO airport where the first layer will be general area markers for each boarding zone. Then, the user can click on each area where the map will zoom in on that specific area, revealing the individual tree map or circle packing details.

The boarding areas will be colored according to which airline has the most of its airplanes parked at the gates.

How to interact:
Two views: **airport view** || **boarding area view**.

- Each boarding area will have its own tree map or circle packing
- User will select whether they want to see tree maps or circle packing
- In the **airport view**. when the boarding area is clicked on, the map will zoom in on that specific boarding area and it will reveal the tree map or circle packing
    - The default **airport view** will have boarding areas fit to the contours of the airport
    - Once the user clicks on a boarding area, the single **boarding area view** will have a tree map or a circle packing (depending on the user selection)
- To return to the default airport view, there will be a back button.

Color will be used to represent the airline's most identifiable theme color. Interactivity in the form of hovering over a specific point with the mouse cursor will present the airline title and the passenger count.

A slide bar at the bottom of the tree map and circle packing will scroll through time in increments of months to show the differences in airline passenger count over time (from 2005-2021)

Small snippets of SFO history will be intertwined directly in the maps in the form of pop up text boxes.

Bar Chart
A bar chart will be used to compare airline passenger totals

Scatter Plot
Lastly, the scatter plot will be used to see overall SFO passenger statistics to support the main visualizations and provide a broader sense of SFO's passenger activity

Please see sketches at the end of this proposal.

Must-Have Features

My project will be in the form of a website primarily using scrolly telling where the user will scroll through the different pages of the website to display all the information and visualizations.

Page layout concept:

Page 1 (Fulfills PO (Project Objective) #1)
- Photo of SFO with project title, class title, and author information

Page 2 (Fulfills PO #2)
- Background information on SFO including photos of construction
- Personal motivation for choosing this topic
- Explanation of dataset
- Brief introduction to the visualizations I have made and how to navigate the website

Page 3 (Fulfills PO #3)
- **Airport view map with boarding areas highlighted in color of the most frequent airline in that area.**
- Tree Map
- Goal: visually represent concentration of airlines at all boarding areas of airport
- Interactivity:
    - Hover over squares on the map to see airline names, number of passengers, and more.

    - Slide bar at bottom to cycle through the years in increments of months

Page 4 (Fulfills PO #3)

- Circle Packing
- Goal: visually represent airlines that have the most presence at a boarding area
- Interactivity:
    - Hover over circle proportions to see airline name, number of passengers, and more
    - Slide bar at bottom to cycle through the years in increments of months

Page 5 (Fulfills PO #3)

- Bar chart
- Goal: compare the airline passenger statistics
- Interactivity:
    - Hover over individual bars to see specific passenger numbers

Page 6 (Fulfills PO #3 and PO #4)

- Scatter plot
    - Give a broader sense of SFO's passenger activity
- Interactivity:
    - Drop down menu to view significant travel times (ex: holidays)

Optional Features

I will try to implement these only if I am ahead of my schedule. I want to ensure I implement my must-have features first.

Optional:

- Fancy animation screen with clouds and plane silhouette everytime the user scrolls to a new page
- Interactive sounds with animations

Project Schedule

Major Deliverable deadlines

- Revised Proposal, related work, and website
  - **3/23/22**
- Alpha Release
  - **4/6/22**
- Beta Release
  - **4/20/22**
- Final Project Presentations
  - **5/9/22 - 5/11/22**

Project progression deadlines

Progression color legend:
Start = start the task (15% completion)
Check progress = check if halfway through completion (50%)
Complete = successfully completed task (100%)

Notes:
Due date means the date the ***task*** is due at 11:59pm, not the deliverable
Milestones throughout project progression to celebrate

Preparation phase

- Collect and aggregate data/gather research on related work
  - **DUE: 3/21/22**
- Start website
  - **DUE: 3/22/22**
- Create D3 templates for all visualizations to make coding easier
  - **DUE: 3/23/22**

Creation phase

Bundle 1 0% [----------] 100%

- *D3 tasks*
  - Complete basic (no interaction) scatter plot
    - **DUE: 3/28/22**
  - Complete basic (no interaction) bar chart
    - **DUE: 4/1/22**

MILESTONE 1 / 6:
Completed scatter plot and bar chart
(4/1/22)

Bundle 2 0% [----------] 100%

- *D3 tasks*
  - Start tree map
    - **DUE: 4/2/22**
  - Start circle packing
    - **DUE: 4/2/22**
- *Website tasks*
  - Complete page 1, 2, 5, and 6 of website
    - **DUE :4/4/22**

MILESTONE 2 / 6:
Submit alpha release
(4/6/22)

Bundle 3 0% [----------] 100%

- D3 tasks
  - Check progress on tree map
    - **DUE: 4/8/22**
  - Check progress on circle packing
    - **DUE: 4/8/22**

MILESTONE 3 / 6:
Halfway through tree map and circle packing
(4/12/22)

Bundle 4 0% [----------] 100%

- D3 tasks
  - Complete basic tree map
    - **DUE: 4/15/22**
  - Complete basic circle packing
    - **DUE: 4/15/22**
- Website tasks
  - Complete page 3 and 4
    - **DUE: 4/17/22**

MILESTONE 4 / 6:
Completed basic tree map and circle packing
(4/15/22)

Completed basic version of all pages of website
(4/18/22)

Submit Beta Release
(4/18/22)

Bundle 5 0% [----------] 100%

- D3 tasks
  - Complete all visualization interactions (See Visualization Design and Must-Have Features for interaction specifics)
    - **DUE: 4/27/22**
- Website tasks
  - Refine style of website (ex: font sizes, color, transitions)
    - **DUE: 4/29/22**

MILESTONE 5 / 6:
Completed late beta project
(4/30/22)

Bundle 6 0% [----------] 100%

- D3 tasks
  - Refine and finalize all visualizations for release
    - **DUE: 5/2/22**
- Website tasks
  - Refine style of website (ex: font sizes, color, transitions)
    - **DUE: 5/2/22**

Bundle 7 0% [ ] 100%

- Present project and celebrate!!
  - **DUE: by 5/11/22**

MILESTONE 6 / 6:
Completed project and presented
(5/11/22)

Related Work

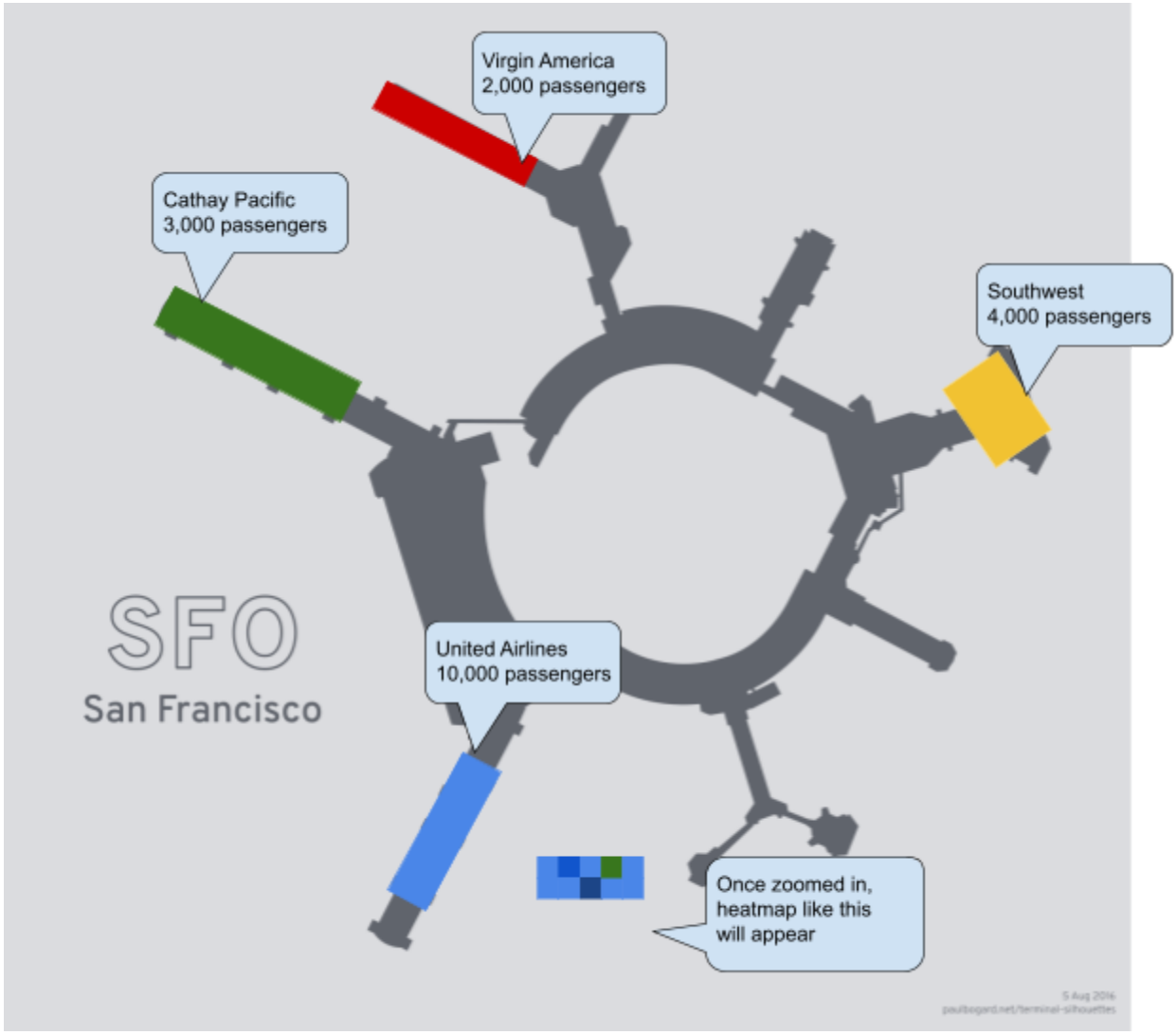
- Wilke, Claus O. *Fundamentals of data visualization: a primer on making informative and compelling figures*. O'Reilly Media, 2019.
- Thorvaldsdóttir, Helga, James T. Robinson, and Jill P. Mesirov. "Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration." *Briefings in bioinformatics* 14.2 (2013): 178-192.
- Javed, Waqas, and Niklas Elmqvist. "Stack zooming for multi-focus interaction in time-series data visualization." *2010 IEEE Pacific Visualization Symposium (PacificVis)*. IEEE, 2010.
- Kanafani, Adib, and Atef A. Ghobrial. "Airline hubbing—some implications for airport economics." *Transportation Research Part A: General* 19.1 (1985): 15-27.
- Bush, Harry, and David Starkie. "Competitive drivers towards improved airport/airline relationships." *Journal of Air Transport Management* 41 (2014): 45-49.
- Hashimoto, Yasuhiro, and Ryo Matsushita. "Heat map scope technique for stacked time-series data visualization." *2012 16th International Conference on Information Visualisation*. IEEE, 2012.
- Fisher, Danyel. "Animation for Visualization: Opportunities and Drawbacks." *Beautiful visualization* 19 (2010): 329-352.
- Gershon, Nahum D. "Visualization of fuzzy data using generalized animation." *Proceedings Visualization'92*. IEEE Computer Society, 1992.
- Francis, Graham, Ian Humphreys, and Stephen Ison. "Airports' perspectives on the growth of low-cost airlines and the remodeling of the airport–airline relationship." *Tourism Management* 25.4 (2004): 507-514.

- Buschmann, Stefan, Matthias Trapp, and Jürgen Döllner. "Animated visualization of spatial–temporal trajectory data for air-traffic analysis." *The Visual Computer* 32.3 (2016): 371-381.
- Lemer, Andrew C. "Measuring performance of airport passenger terminals." *Transportation Research Part A: Policy and Practice* 26.1 (1992): 37-45.
- Jadeja, Mahipal, and Kesha Shah. "Tree-Map: A Visualization Tool for Large Data." *GSB@ SIGIR*. 2015.
- Gorodov, Evgeniy Yur'evich, and Vasiliy Vasil'evich Gubarev. "Analytical review of data visualization methods in application to big data." *Journal of Electrical and Computer Engineering* 2013 (2013).

Visualization Sketches

(Disclaimer: data not real for example purposes)
(SFO drawing source: https://paulbogard.net/terminal-silhouettes/sfo/)
(All credit goes to Paul Bogard for SFO silhouette drawing)
(I (Darren Dunn) have adapted and edited the silhouette drawing for example purposes)

Heat Map Sketch (originally planned; see orange notes below)



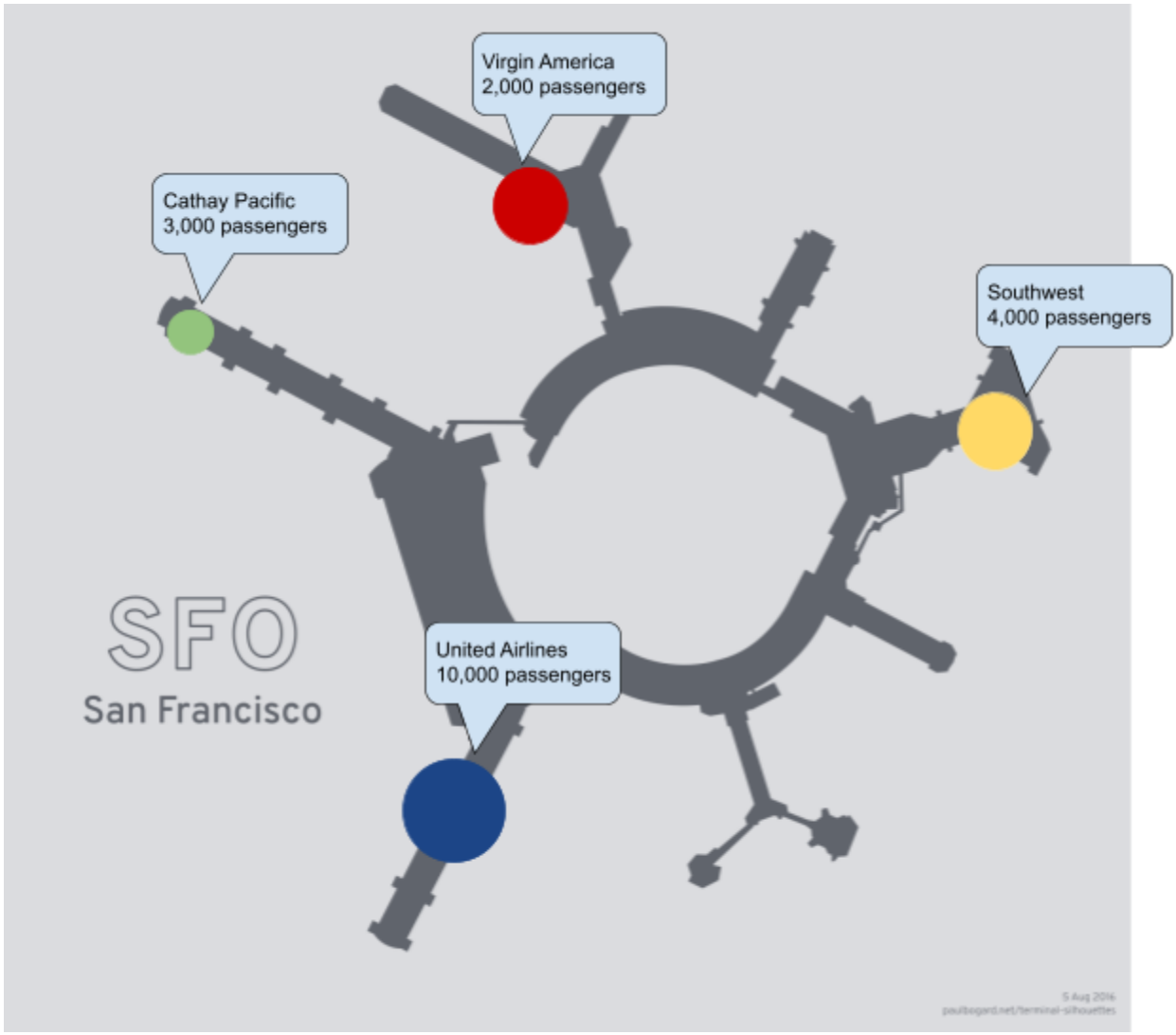
(Slide bar to cycle through months here)
United Airlines is shown as a big blue area denoting their dominance in that boarding area.

Once user clicks on that big blue area, the map will zoom in and look like a heatmap, like the other airline examples.

Changing heat map to tree map:

- Will keep this map view with zooming features.
- However, heat maps will be replaced with tree maps.

Proportional Map Sketch (originally planned; see orange notes below)

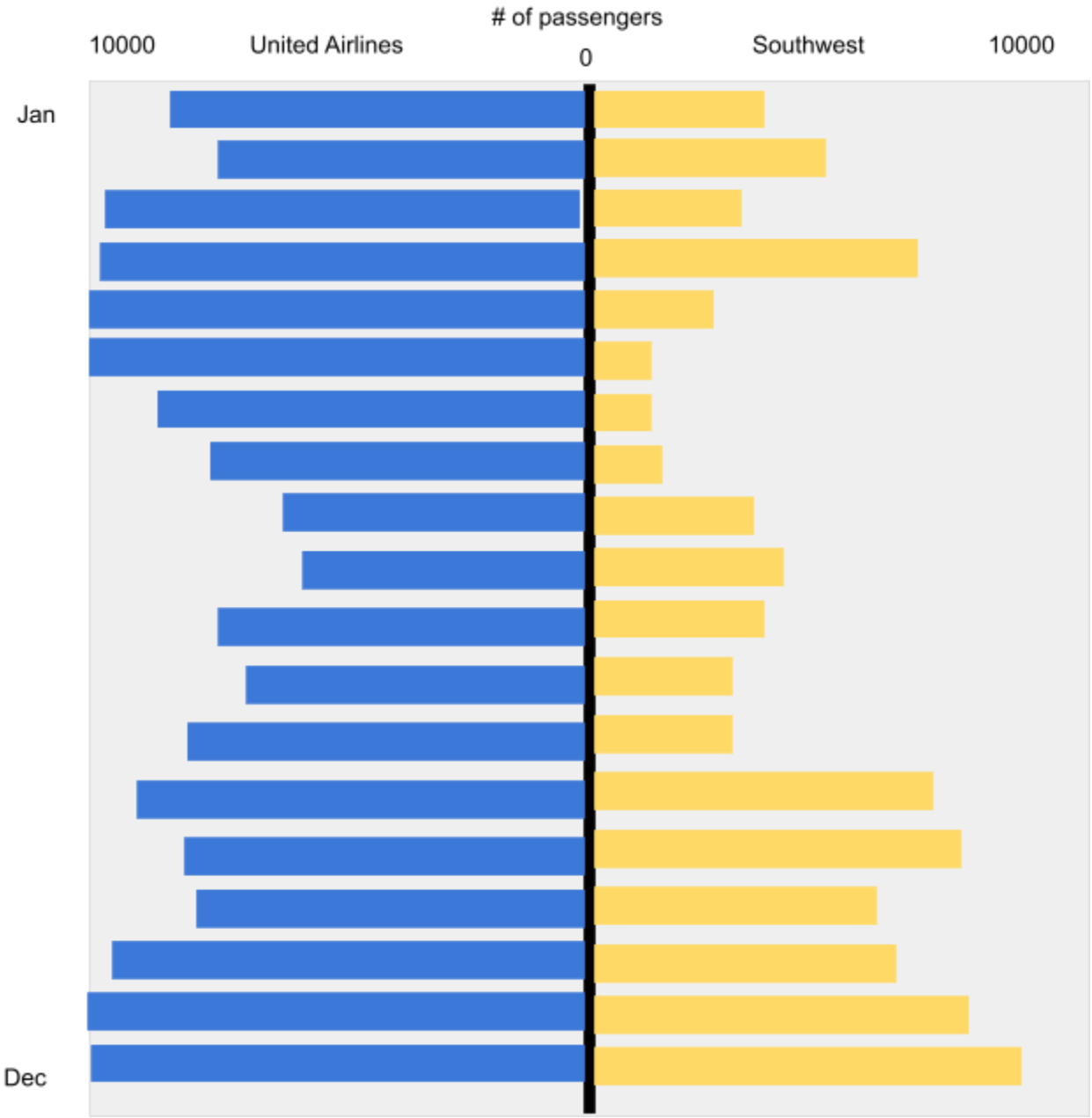


(Slide bar to cycle through months here)

Changing proportional map to circle packing:

- Will keep this map view with zooming features
- However, proportional maps will be replaced with circle packing

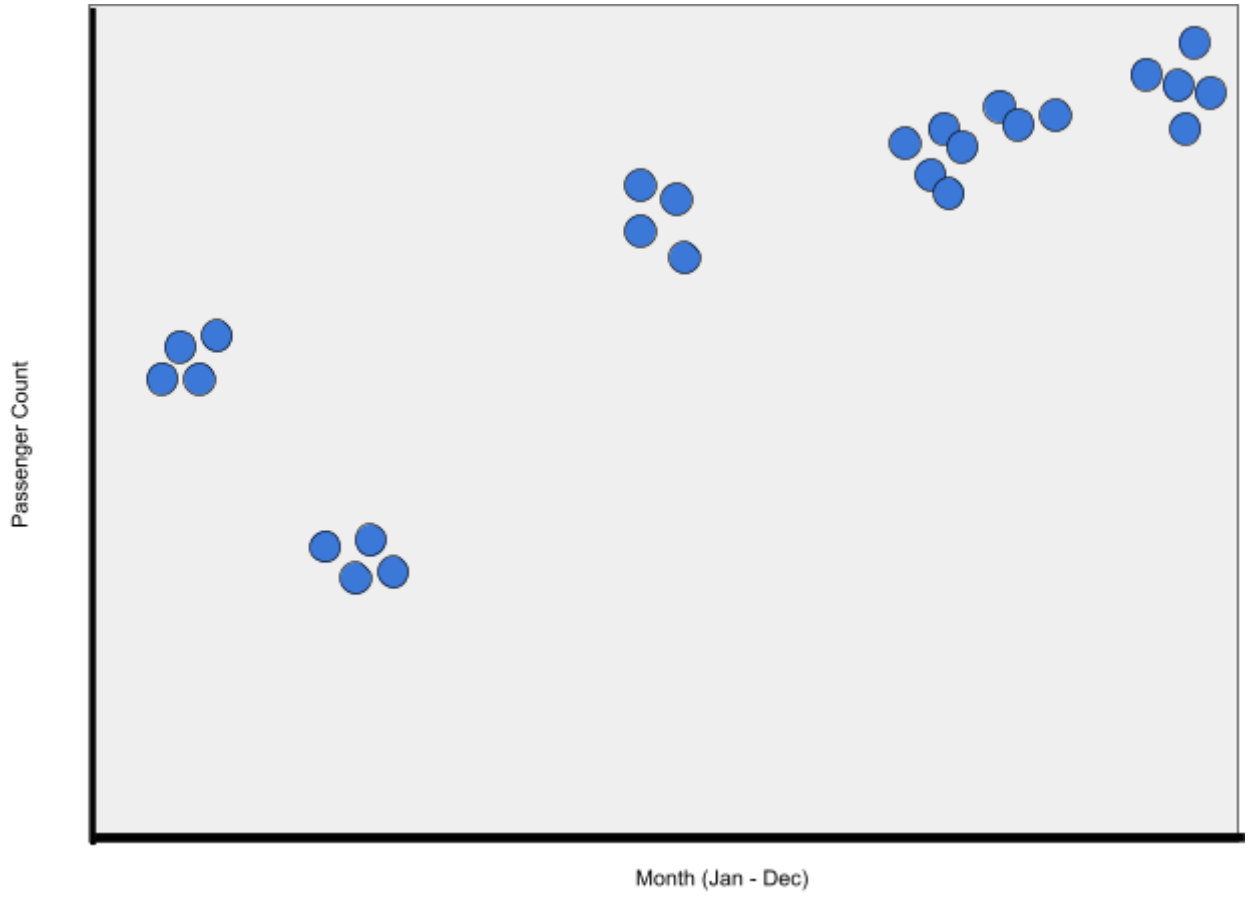
Diverging Bar Chart Sketch (originally planned; see orange notes below)



Changing to basic bar chart

- Change concept to a bar chart representing each airline's average passenger count for each year

Scatter Plot Sketch



(example showing United Airlines)

Changing context of scatter plot

- Changing the context of scatter plot to reflect SFO's average passenger count for every month per year.
- Purpose: To give broader sense of SFO's passenger activity

Alpha Release Report

Completed Features

- Research required for introduction to SFO and background information (ready to integrate into website)
- D3 Implementation of bar chart and scatter plot

Upcoming immediate milestones

- Create page 1 and 2 of website
- Create digital template of SFO silhouette
- Create more visually pleasing bar chart and scatter plots
- Tree Map D3 template

Roadblocks

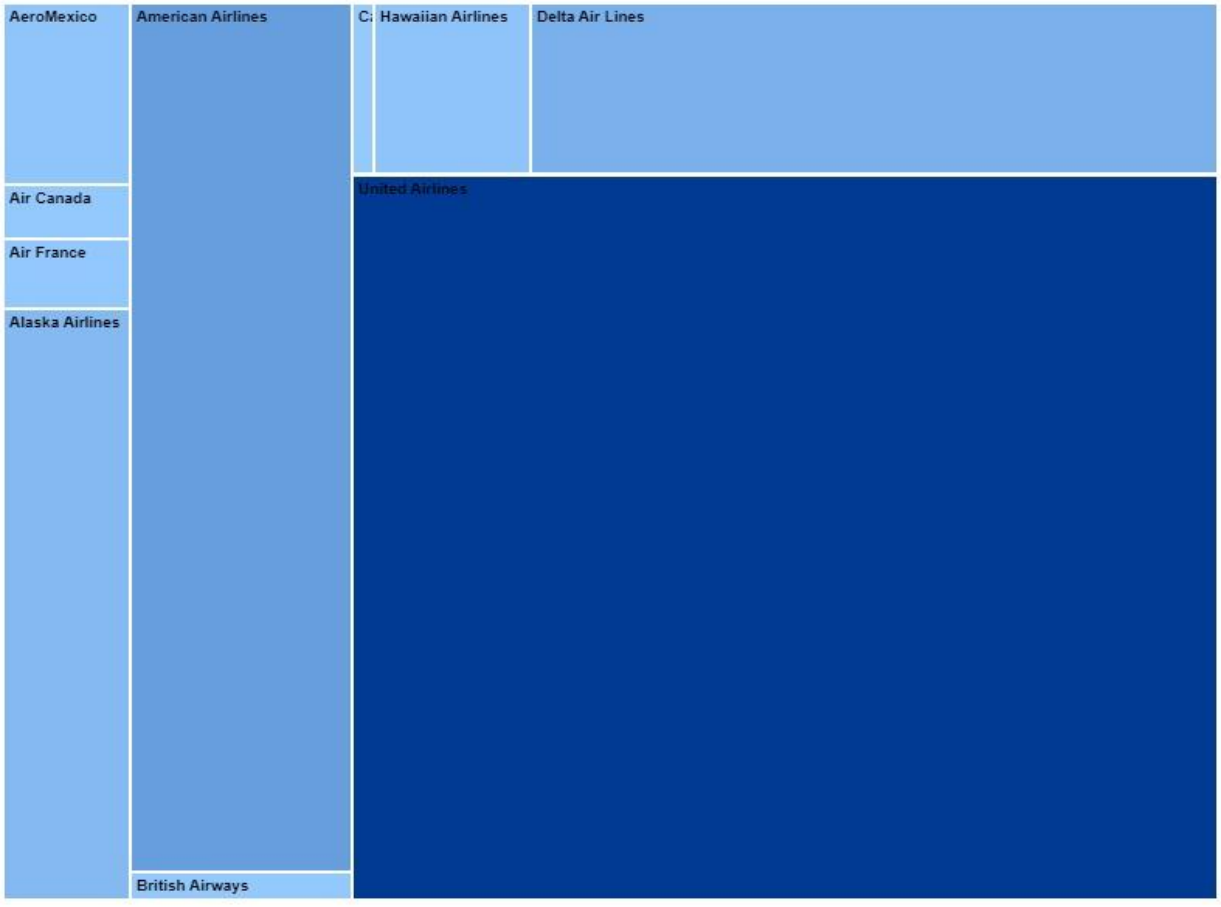
- Unable to realistically make heat and proportional map without use of professional software
- Reconsider heat and proportional map, replacing these with a treemap and circle packing (diagram)
    - Tree Map
        - Tree map made according to boarding area
        - Click on the boarding area on SFO map, and zoom into the respective treemap for that boarding area
        - Integrate tool tips when hovering over treemap to see detailed info about passenger statistics, airline, and more
    - Circle Packing
        - Similar to tree map where user will click on a boarding area and they will be zoomed in to the circle packing visualization
- New Visualization hierarchy:
    - Tree Map and Circle Packing visualization (SFO map)
    - Bar Chart
    - Scatter Plot

Implemented Visualizations in D3

Bar Chart (temporarily hosted on Vizhub, ready to integrate into website)
https://vizhub.com/darrenisdunn/9ea250f06fb34d4189594179eb4127c1

Scatter Plot (temporarily hosted on Vizhub, ready to integrate into website)
https://vizhub.com/darrenisdunn/49eee31aed184aef90057276dde026db?edit=files&file=index.html

Example Tree Map of Airline Average Passenger Count Rankings using RAWGraphs



Example Circle Packing Visualization of Airline Average Passenger Count Rankings using RAWGraphs