

Project 2 Final Report

TA Mentor: James Chen (jzc24)

Team Members: Jiangjie Man (jm2559), Rou Qu (rq44), Xiuyan Xin (xx254)

Work Distribution

Each team member is scheduled to share the works equally. Because of the nature of our project, members have participated in all tasks together. All of us have brainstormed the project idea, coded and contributed to aesthetic quality. In terms of project management, Jiangjie is in charge of integrating codes, Rou is in charge of visual design, and Xiuyan is in charge of organizing data. Specifically,

- Rou Qu: Implemented the info webpage and real-time information (weather and visual elements)

- Xiuyan Xin: Data Processing and implemented real-time map (Info board) bar chart

- Jiangjie Man: Data reformatting and implemented real-time map visualization (added routes for each trip, total number of trips showing on map)

Project Description

Our project focuses on presenting San Francisco City bike share system. Through the real time data visualization using d3, we hope to provide people a sense of how the bike share system works. Every blue dot on the map represents a single bike station, and the small green dots represent bike trips scattered throughout the city. Our data visualization gives people an overview how bike share system works daily in 2014. Using

the date filter, you can see the real time bicycle move from day to day. Also, you can see how weather the total number of trip will change along with the time of the day.

Data Description

We found four different datasets from Kaggle. These datasets show bike trip from 2013 to 2015 in bay area. There are total of 4 files used and 3 generated files:

- Trip.csv focuses on the trip route information which includes start date, end date, duration, start station and end station. It helps us get each bike route from the start point to the end point in different times
- Weather.csv presents the weather information which includes the temperature, humidity, wind, precipitation, cloud cover, etc.
- Station.csv shows the station information which includes the city name, latitude and longitude.
- San_francisco_precinct_20120702.topojson: it helps draw the geographical map of San Francisco, The data file for SF map is produced by Roger, which is part of his work "SF precincts".
- Generated Files: SortedStartMap.csv (top 15), SortedEndMap.csv (top 15) and Top_35_Station_start(25 record)

Because of the complexity of data, we decided to focus on the year of 2014. Also, we figured out about 50% of the stations and docks are located in San Francisco City, and it makes up most of the system use. Therefore, we decided to focus on the San Francisco City Bike System.

We brainstormed a lot of ideas, and determined to use real time animation to present how the system works because it would give intuitive impression of the day's volume of traffic flow. The form of presenting the data in real time animation is also creative.

To better use the dataset in different ways, we parsed the trip.csv into "trips.json". Also, we modify the weather.csv by filling the empty events with "no rain". Secondly, the station dataset helps us locate each station on the map by using the latitude and longitude. Apart from above, we process the 2014 whole year data to calculate the number of visits for each station using Java and return the top 15 stations with their frequency, namely SortedStartMap.csv and SortedEndMap.csv. And the two files are used to draw the bar chart. When sorting the top stations we used both station.csv and trip.csv to create hashmap that only contains the SF station and count the total number.

A description of the mapping from data to visual elements

We use `scaleTime()` to scale class Date into number, `scaleLinear()` to map the length of number into another. Finally, we scale the color from light to dark according to the time of the day.

For the design part, we use the Roboto font because it is one of the best font for data visualization. We also center everything to convey structure. We choose soft colors which could make people feel comfortable.

In terms of color, we use light blue, grey for background, yellow, white and black for text in order to make it easier visually follow. We have used circles for demonstrating stations (blue) and routes (green). The grey lines between stations are trip routes for rides.

We use blue and green colors because we want to adopt the same color as the Bay Area bike share system uses for bicycles. The bright orange color for SF map increases the map's recognition in contrast with stations and routes. To explore the SF map better, we have used **mouse over function**. When you mouseover each blue circle (station), there would be a round rectangle on the top showing the name of each station. The real time information is shown in another round rectangle on the right of the map. The real time information includes time, total number of rides, weather info and **a drop down box for sorting the date** (month and day). To make the information more legible for viewers, we use the black for text and white for the rectangle's background. To make the our data visualization interesting and intuitive, we also input many visual elements (icon, picture).

The story

Our data visualization tells a story of how the bike share system works around the city of San Francisco in the year of 2014. Usually rainy and foggy days will decrease the number of bike trips. Temperature, however does not seem to have a big influence on riders. San Francisco's temperatures are always moderate.

For the data analysis part, we found that the top three stations are San Francisco Caltrain (Townsend at 4th), Harry Bridges Plaza (Ferry Building) and San Francisco Caltrain 2 (330 Townsend) both as starting and ending stations. 2nd at Townsend is also among the most popular stations, rating 4th in the starting station and 5th in the destination station. Market at Sansome, Market at 4th, and Steuart at Market, which are located within walking distance of BART stops, are popular as well. We predict that people would like to use bike share by visiting a number of attractions in downtown San Francisco and commuting by BART.