CS 235 – User Interface Design

Project 3 – Data Visualization

Team Innovative Designers

Bhavna Gurnani
Radha Kotipalli
Rakhi Poonam Verma
Tien Nguyen

Table of Contents

What is Bay Area Bike Share?

Bay Area Bike Share is the region's bike sharing system with 700 bikes and 70 stations across the region, with locations in San Francisco, Redwood City, Mountain View, Palo Alto, and San Jose.

Bay Area Bike Share is fun, easy, and affordable and is intended to provide Bay Area residents and visitors with an additional transportation option for getting around the region.

Purpose of our Application

- The main purpose of the application is to analyze the data in such a way that it can benefit the frequent users of the bike share facility.
- To explore benefits and bottlenecks of the organization and to provide solutions to those bottlenecks.

Who will use our App, and what are their Goals?

The following can be considered as the users of our application

- Bay Area Bike Share Company
- Any other companies whose business benefits from bikeshare analysis
- Customers/ Clients of Bay Area Bike Share

Different users will have different goals for using this application.
Following are some of the goals that the users will focus while using the application:

- **Goals for Bay Area Bike Share Company:**
    - The company can compare the availability and usage of bikes at a particular station and can increase the count if necessary to meet demand.
    - If a station has very less usage of bikes then they can reduce the bike count from that station and add those bikes at some other station where the availability is more. This would help them to optimize utilization of their resources

- **Goals for companies whose business gets benefitted with bikeshare analysis**
    - Analysis of bikeshare data could be useful to companies depending on what kind they are. Companies that deal directly with customers, such as stores or restaurants, might prefer being located near stations with high bike usages because that would mean there would be a greater chance for customers. A company that doesn't deal directly with random walk-in customers, such as a tech company, might prefer a location with low bike usage so that their employees would have bikes available to use.

- **Goals for Customers/ Clients of Bay Area Bike Share:**
    - Can check during what period the availability of bikes is more, so that they can rent bikes during that time.
    - Customers can also check which stations has maximum bikes rented and can avoid those stations.

How did we fulfill our users' Goals?

In order to fulfill our users' goal, we took 6 months data starting from Aug 2013 and ending at Feb 2014 and analyzed it using Tableau. We discovered different interesting patterns in data like the station with more usage of bikes, peak hours of bike usage, weather related analysis of bike share etc. We also showed all our analysis using different easy to understand visualizations so that all the intended users of this application can get their desired information of interest with minimum efforts.

Data Sources

We used data provided by Bay Area Bike Share's website. Out of the two data sets available, we chose six months data set i.e., from August 2013 till February 2014. In some cases, we omitted August due to insufficiencies in the data.

How did we create our visualizations?

We used Tableau to create our visualizations. We went through the tutorials available online and learnt how to use Tableau for our analysis purpose. Also, for creating visualizations, the most important thing was to have better understanding of data. We went through all the files available in our data set and studied all the attributes available in those files and came up with different interesting queries which could be helpful for the intended users.

Design Patterns used

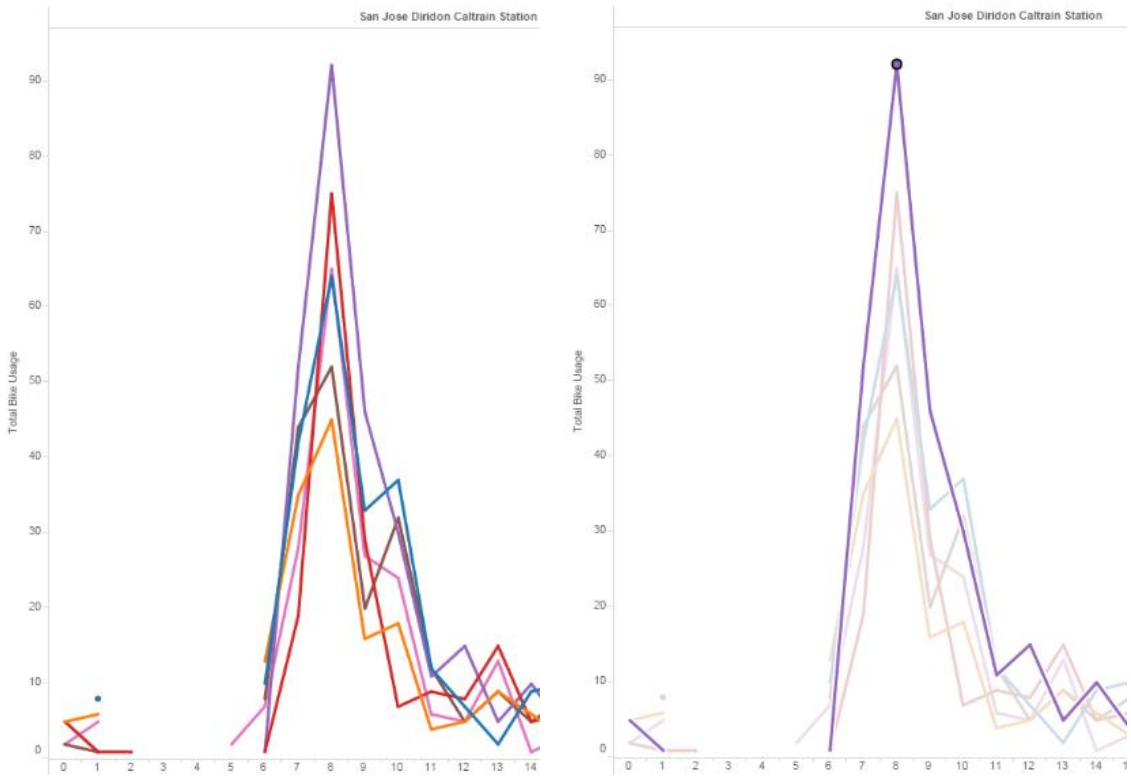
1. Data Spotlight

What?

This design pattern is used to highlight data on mouse click and dim the rest.

Why did we use this?

It allows one to focus on a specific data item while maintaining the ability to easily compare to the other items of data. Almost all graphs uses this design pattern. It untangles data threads. Hence, we used this so that user don't have to put extra efforts in understanding the data.



2. Dynamic Queries:

What?

Dynamic Queries give the standard controls to filter data dynamically.

Why did we use this?

We wanted to give some easy to learn controls to user so that they can filter the data based on their interest and see the visualization for the filtered data. Dynamic Queries suited best to our need so we included this.

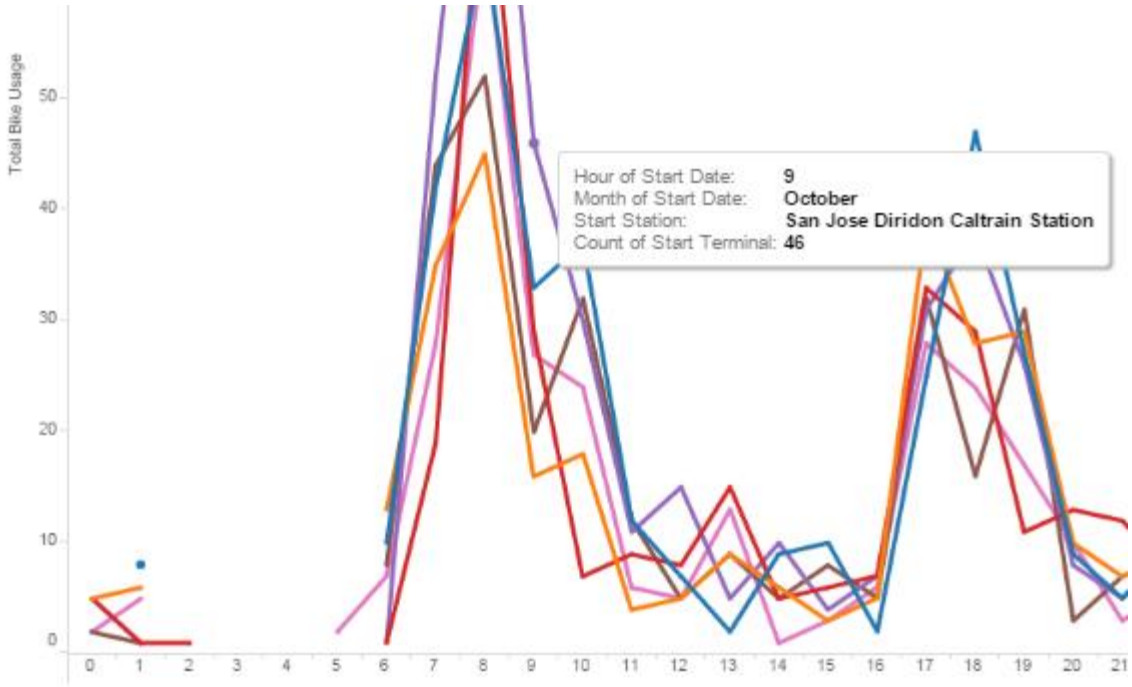
3. Data Tips:

What?

This design pattern allows the data values at a particular point on graph to pop up when the mouse rolls over it.

Why did we use this?

It shows details about specific points on an overview and helps the user to find data values for further analysis.



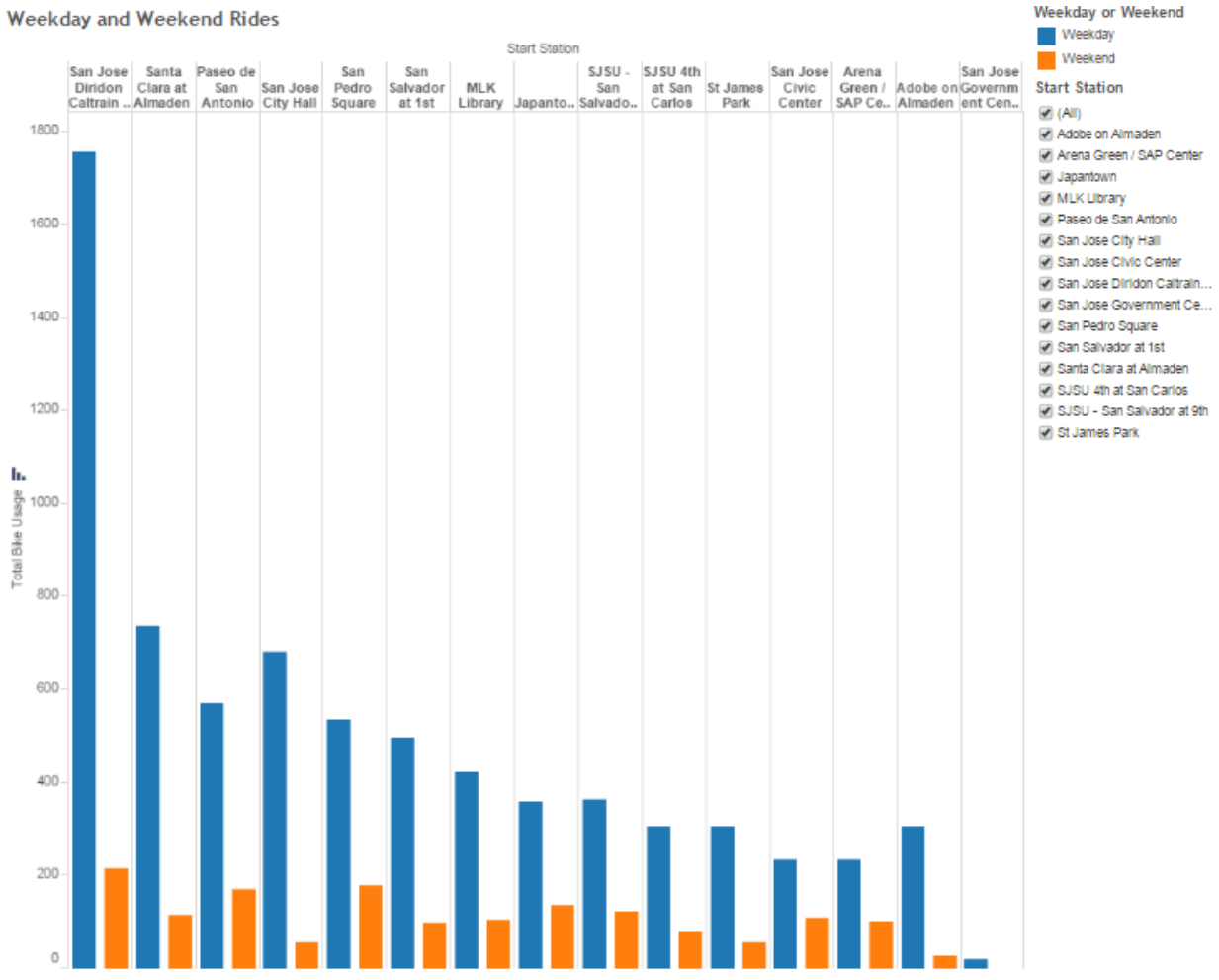
4. Bar Chart Guidelines:

What?

For comparing quantities and for showing the differences, bars are very effective and popular.

Why did we use this?

Our eyes perceive differences, not absolute values and if we show these differences in terms of bar graph then that would be more effective as compared to other design patterns. So, we found this design pattern apt to fulfill this requirement.



5. Multi Y-Graph:

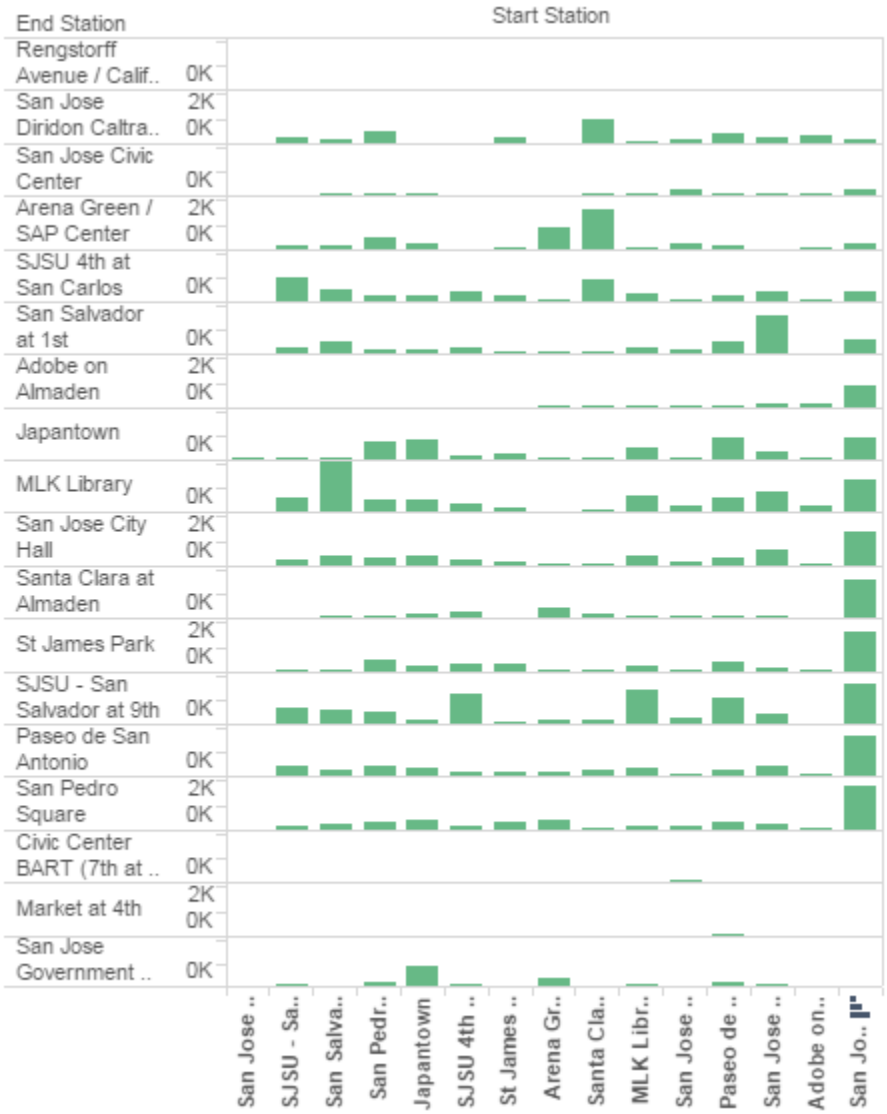
What?

In this design pattern, the graph shares same X axis and the bars are stacked vertically.

So, it shows multiple graphs with different Y axis but same X axis.

Why did we use this?

We found this design pattern suitable to show the number of trips with same start station and different end stations.



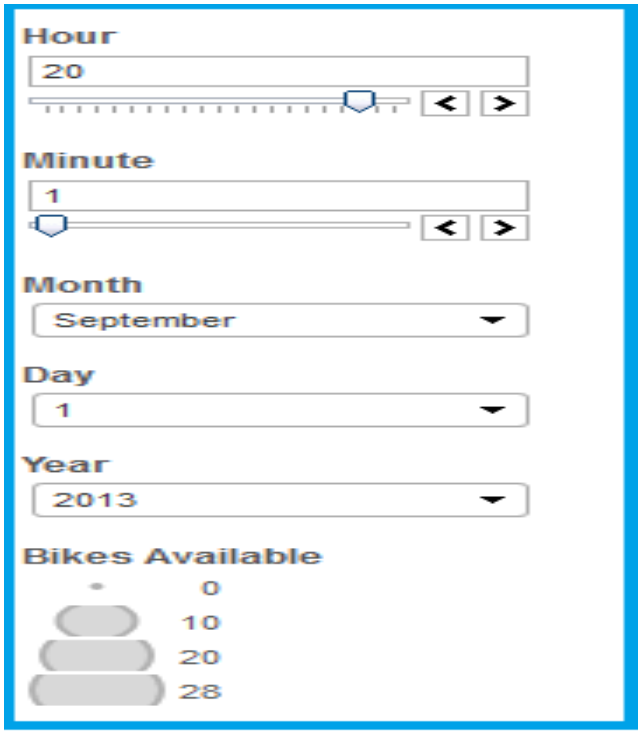
6. Choice of Control:

What?

This design pattern gives the interface to the user so that user can see the visualization by filtering the data using sliders and drop down

Why did we use this?

It allows users to filter their data with an easy to use GUI so best suited as per our requirement.



7. Other Design Patterns:

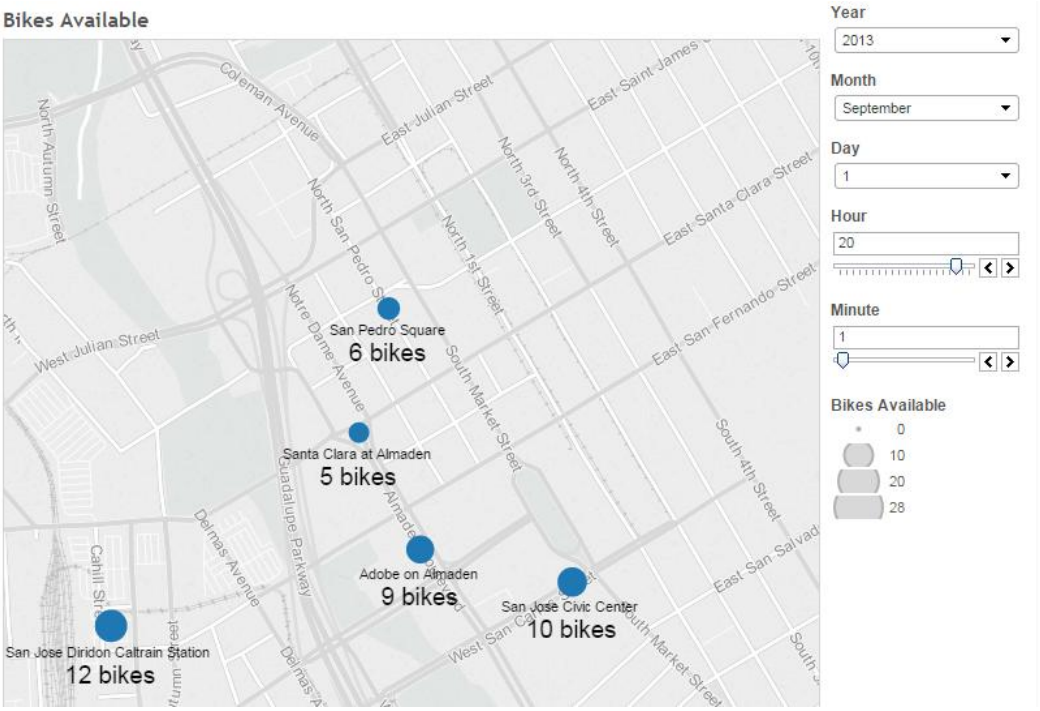
- Visual Framework
- Module Tabs
- Grid of Equals
- Hub and Spoke

What Insights did our visualizations reveal?

We produced the following visualizations with our data:

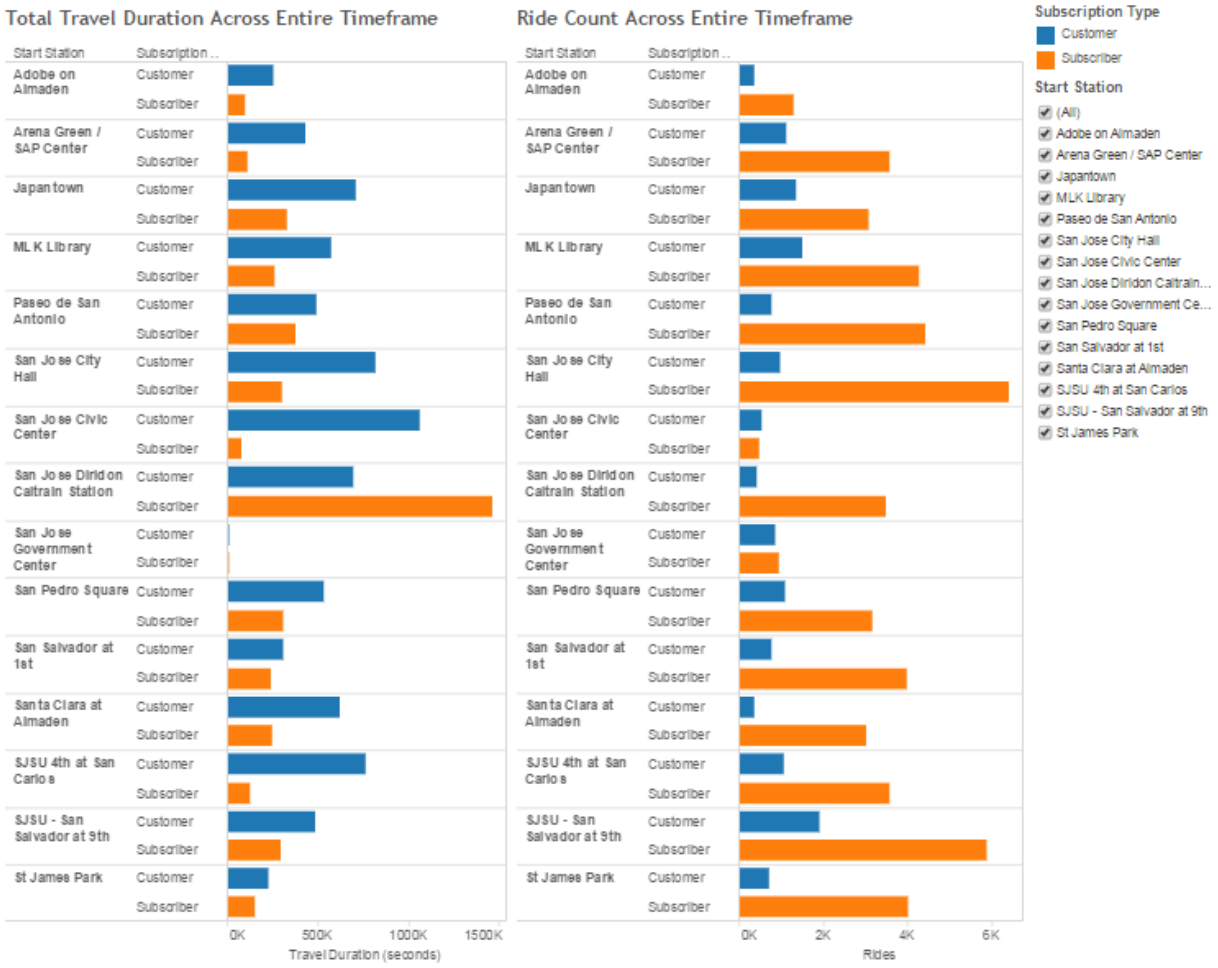
- **Map of Bike Availability**

  The map allows one to quickly determine whether a particular station has bikes available or not at a particular date/time. The user uses the dynamic queries to select the year, month, day, and time, which causes the map to update the data based on those parameters. Each station is marked on the map with a bubble that is sized based on the number of bikes available. We only showed bike availability for 5 of the stations because the size of the data file containing data for all of the stations was too big.



- **Comparison of Customer and Subscriber**

  The two bar charts show a comparison between the total duration of seconds spent on bikes and the amount of trips made by subscribers and customers. A subscriber is someone who pays the annual membership fee. A customer is someone who is not an annual member; they pay for either the 24-hour membership or the three-day membership. For the most part, a high bike duration correlates to a high amount of trips. However, in the row corresponding to the station at San Jose City Hall, customers were making longer trips in shorter amounts; subscribers were making shorter trips in larger amounts. What we suspect from this insight is that people who regularly do business in City Hall are subscribers, and that they regularly make trips in short bursts of time. Since fewer customers were making trips from City Hall, we suspect that customers who borrow bikes from the City Hall station tend to not be regulars. These customers might not usually have business in the City Hall area, or they might not use bikes frequently when in the area.

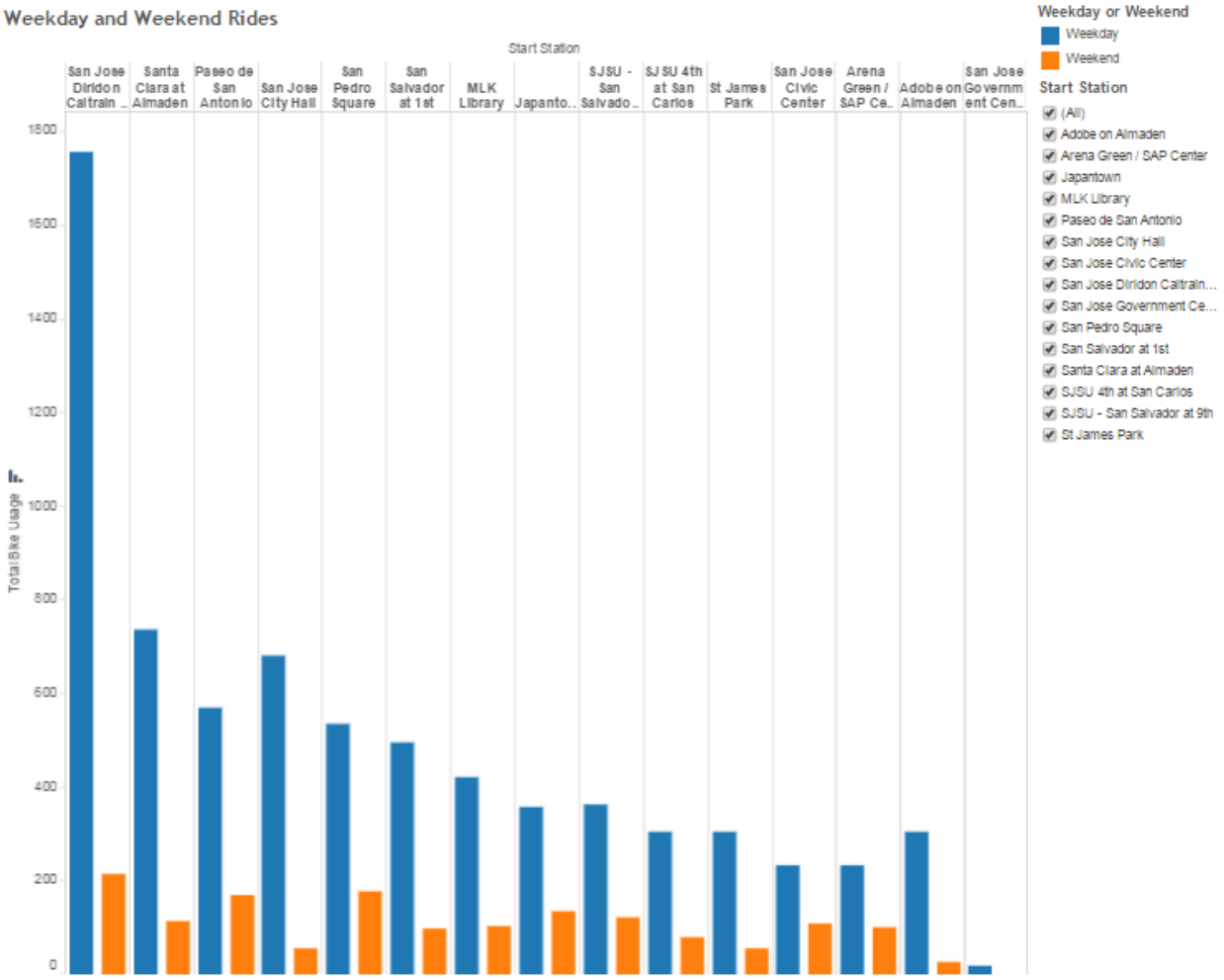


- **Weekend Rides Compared to Weekday Rides**

  We made a bar chart to compare the total weekend rides to the total of weekday rides. Rides are counted as individual bike rentals. Not surprisingly the amount of weekday rides exceeds the amount of weekend rides, likely since there are more weekdays than weekends. However, it is interesting to note that although weekend bike rentals at the San Jose Diridon Caltrain Station was close to the amount at other stations, the amount of weekday bike rentals at the station was about two times as much as the next highest amount of rentals at another station. We suspect this discrepancy is due to the fact that a significant amount of people use Caltrains on weekdays due to work or school; however, we do not have statistics to support this suspicion.

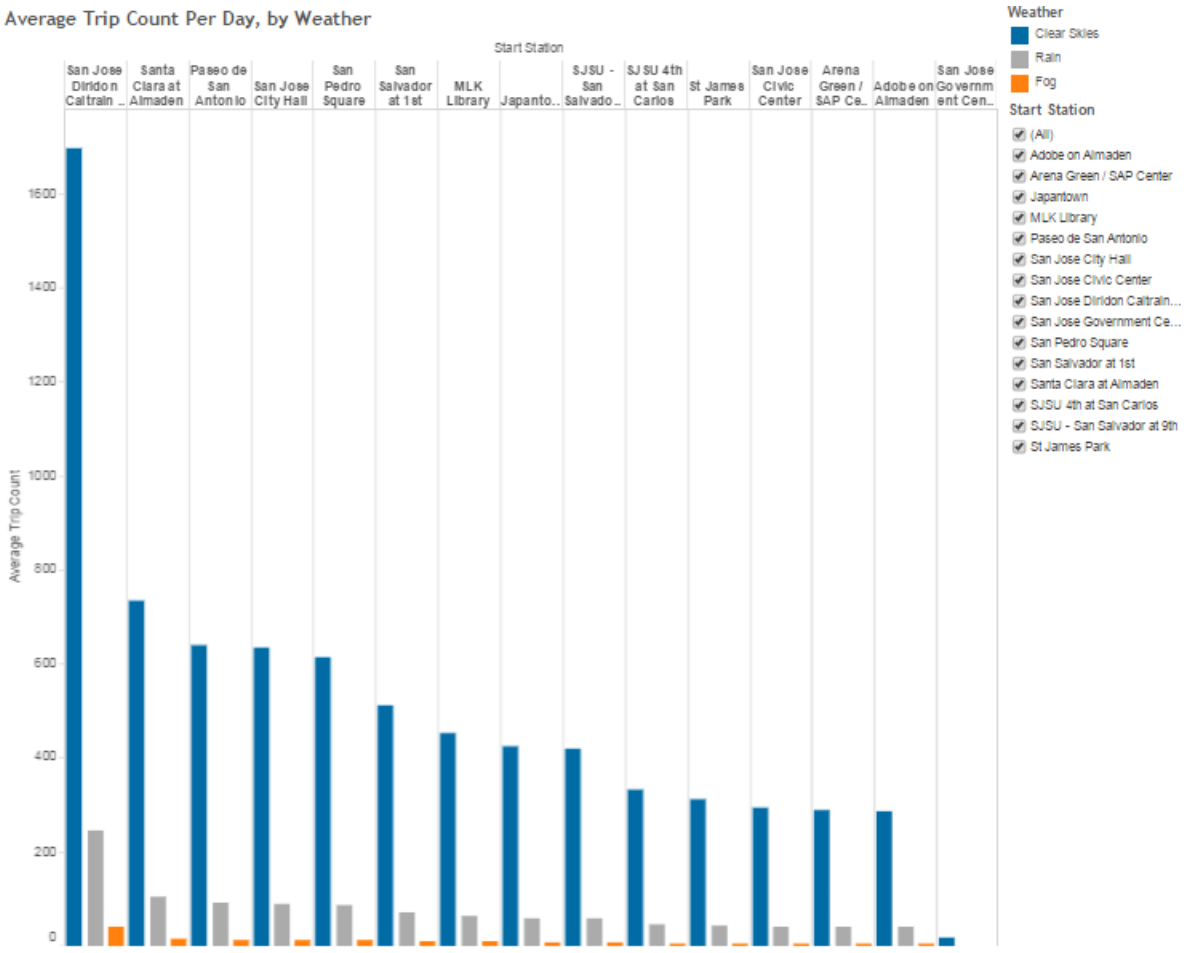
  The Average rides / day: (Aug 2013 - Feb 2014)

  Weekends: 1592 rides / 52 days = 31 rides
  Weekdays: 7343 rides / 132 days = 56 rides



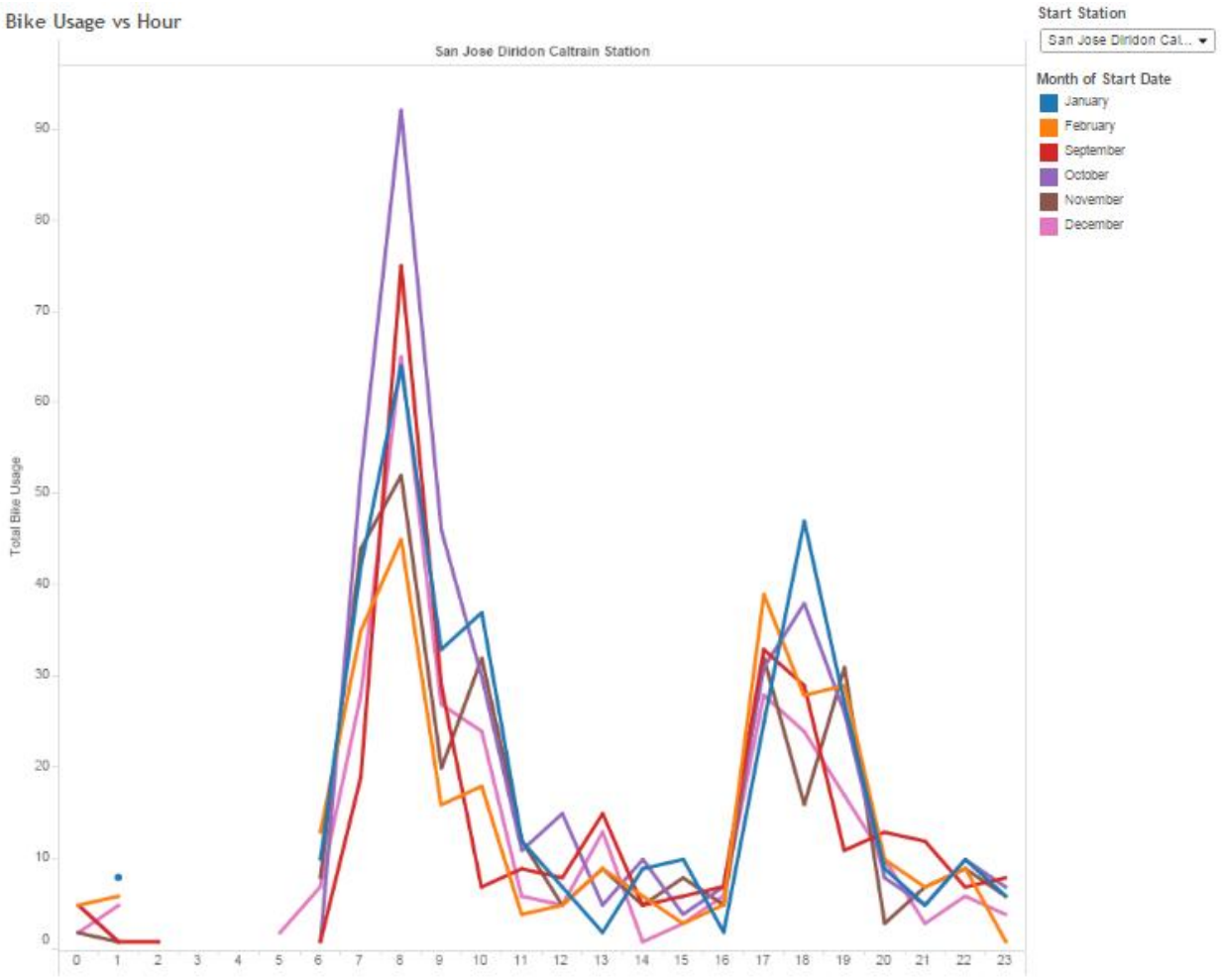
- **Bike Usage and Weather**

  We produced a bar graph that shows the average number of rides per day, separated by weather conditions. This graph allows one to see how weather was affecting average bike usage. One could easily see that bad weather was negatively affecting bike usage across all bike stations. Interestingly, the average number of bike rides during rainy weather was mostly scaled consistent to the number of rides during clear skies; for example, the Diridon Caltrain Station had the greatest average number of bike trips during all weather conditions.



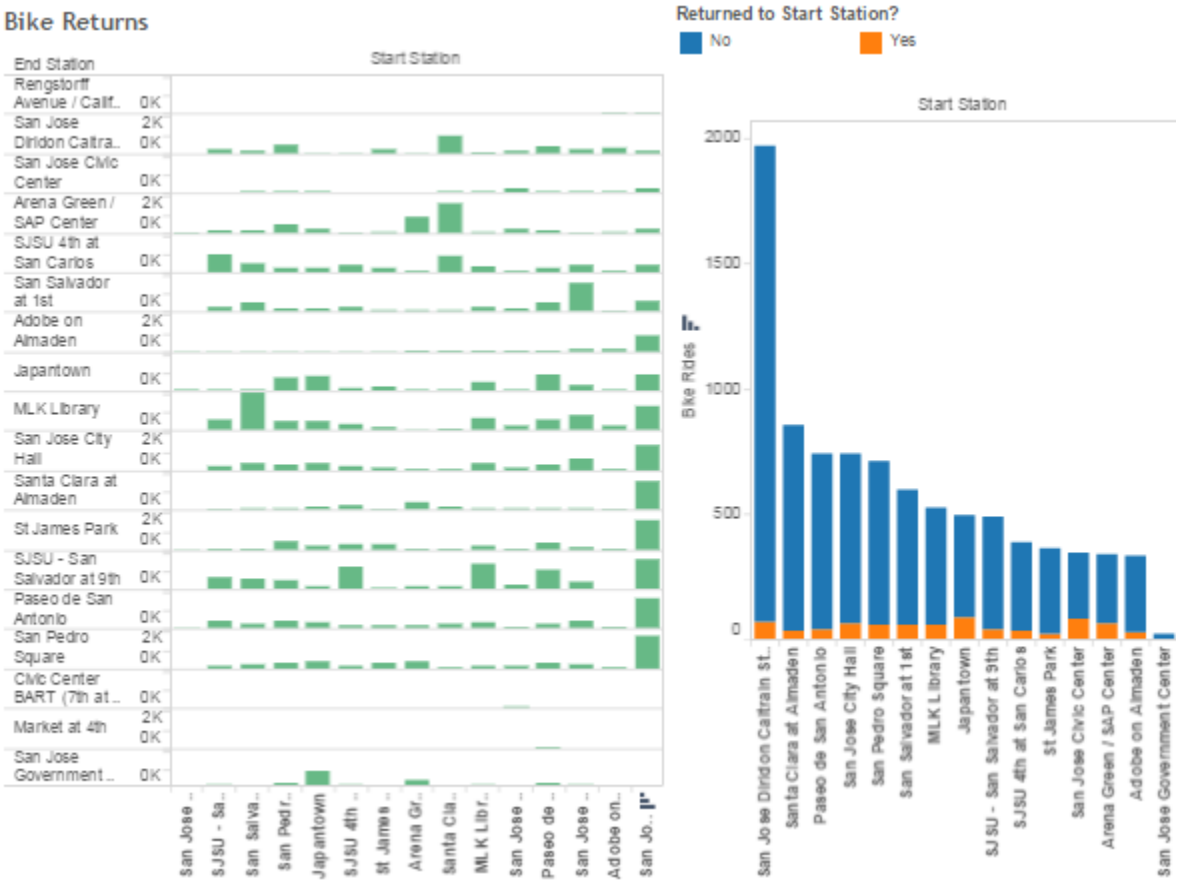
- **Bike Usage vs. Hour**

  We produced overlapping line graphs to compare total bike usage over the course of a day, separated by month. There are breaks in the lines due to data not existing for certain hours. This visualization allows one to easily see the peak times for bike rental. For example, by filtering so that the graph shows bike usage for the San Jose Diridon Caltrain Station, one can see that the peak times are around 8 AM and 4 PM. These correspond to the rush hour times, which we suspect means that many workers or students use the bike station. Since our visualization shows total bike usage instead of average bike usage, we are not able to determine how many bike rentals occur on average at a particular hour. If we showed such an average, we may be able to gain more insight on bike scarcity at specific hours. Such information might be useful for people who want to use bikes.



- **Bike Returns**

  To study the trends of bike returns, we made a dashboard containing a multi-Y bar graph and a bar chart showing the number of bikes returned to a particular station. The multi-Y graph puts the amount of bikes returned to a station vs. the station the bike was taken from. This allows one to quickly see where bikes from a particular station tended to end up. The bar graph on the right allows one to compare bike returns to the same start station to bike returns to different stations. For ease of use, clicking on a station on a graph will cause the other graph to filter based on that station. One quickly sees that in the majority of trips the bikes are being returned to different stations from where they came from. This insight makes sense, since bike rentals are timed and people who use bikes for travel (rather than leisure) would tend to return the bike near where they have business at. It is possible that some people might be making return trips from the same end station to the same start station, but we do not have the data necessary for making this connection, since bike riders aren't identified.



- **Busiest and Calmest Days**

  Out of the available data (Aug 2013 - Feb 2014) :

  Busiest Day: October 24, 2013 with 106 rides

  Calmest Day: December 25, 2103 with 7 rides

How to run our Application

Download the source from https://github.com/rakhirpv24/bay_area_bikeshare and open index.html in a web browser.

Works cited

Bay Area Bike Share. 2015. *First 6 months of data* [Data file] *August 2013 - February 2014*

. Available from http://www.bayareabikeshare.com/datachallenge