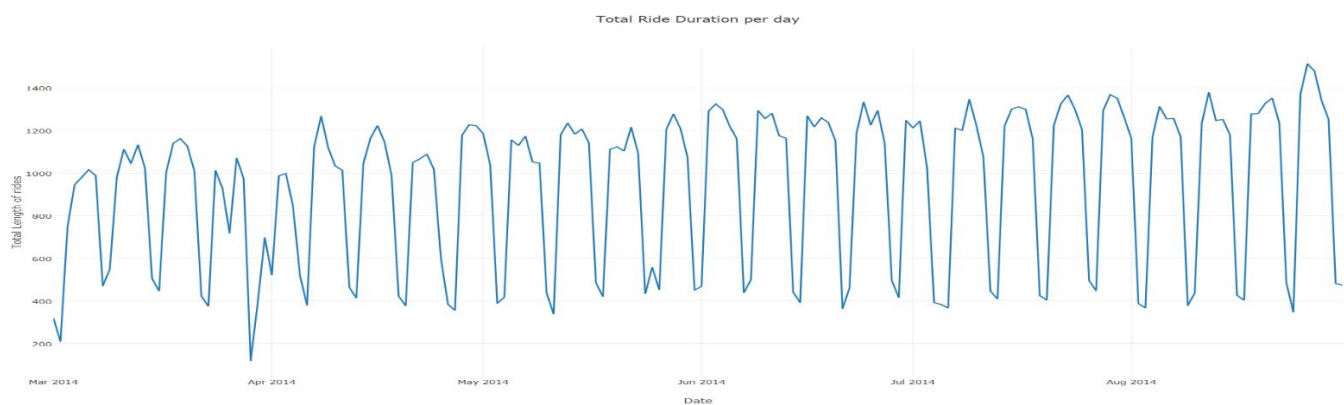


Bike Share Program

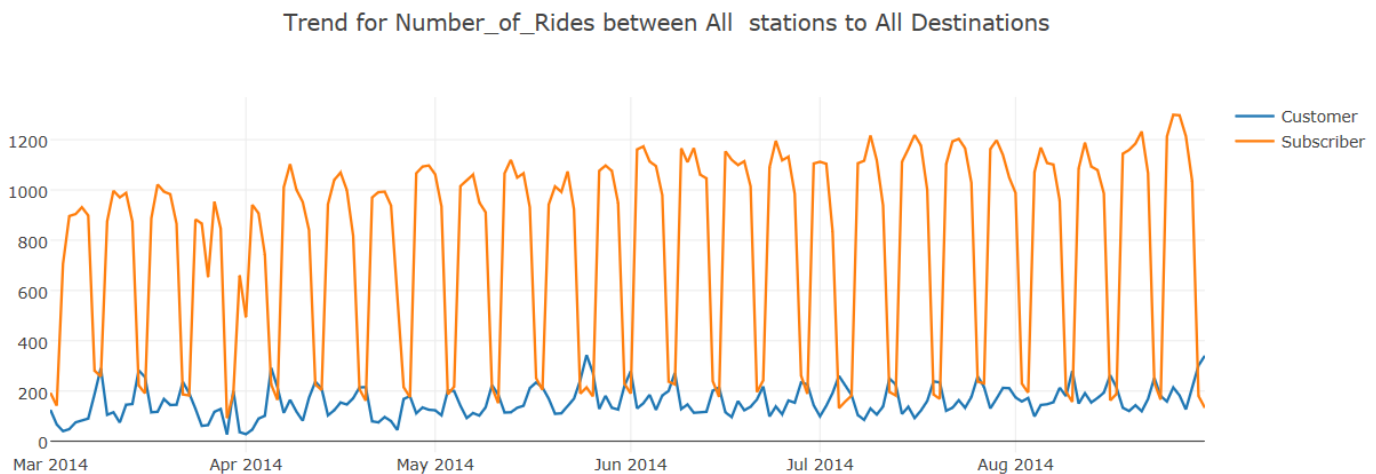
The Bike share program provides bicycles for travel in city. In given data set there are 70 stations with average terminals of 24 per station (median = 26) and 35 stations having terminals more than average.

The data for our analysis is from March 1 2014 to Aug 31 2014. The total Number of bikes up to Aug 31 2014 are 681. The total number of trips occurred between these days is 171792. So, in these 184 days' total duration of rides is 3242407 minutes. So, average ride duration for the whole period is 19 minutes. So, Total Duration of Rides offered by program is equal to 2,251' ½ days.

Bike Share usage throughout year:



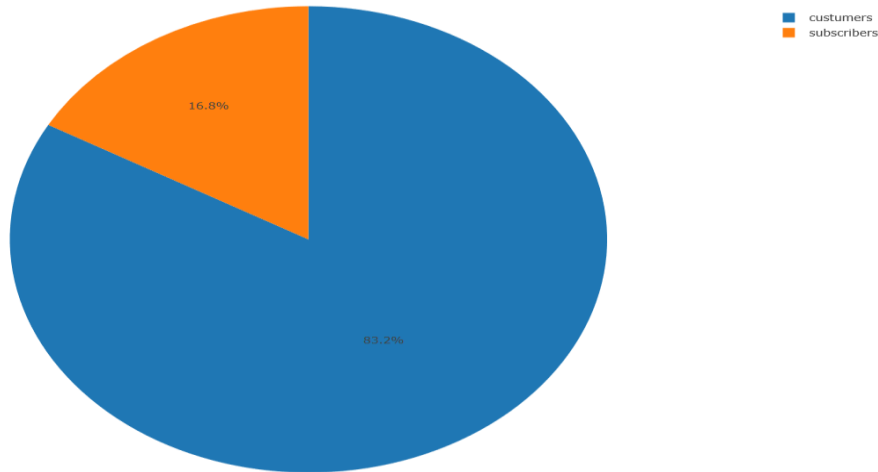
Interactive visual: <images/Trend Number of rides per day.html>



The above graphs show that type of Customer Segment Subscribers choose to take rides more than customers who take ride each time. Busiest month is July with total rides of 31278 followed by August 31212.

Interactive visual: <images/trend.html>

Customer Segment of Bike Share:



Interactive
visual:

[images/basic_pie_chart-line.png.html](#)

83.2% of riders are subscribers who pay subscription fees at once. Remaining 16.8% are customers who take instant ride pass.

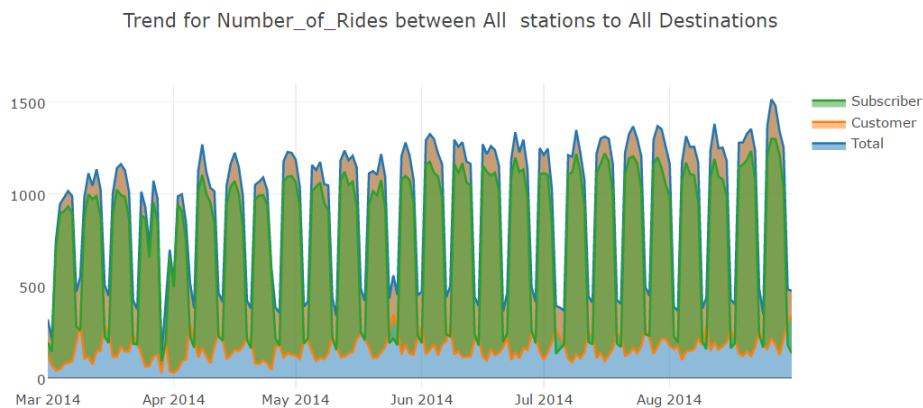
Average Number of Rides per day

Subscriber	777
Customer	157
Total	934

Usage of Bike Share:

Monthly Analysis:

Interactive visual: [images/monthly.html](#)

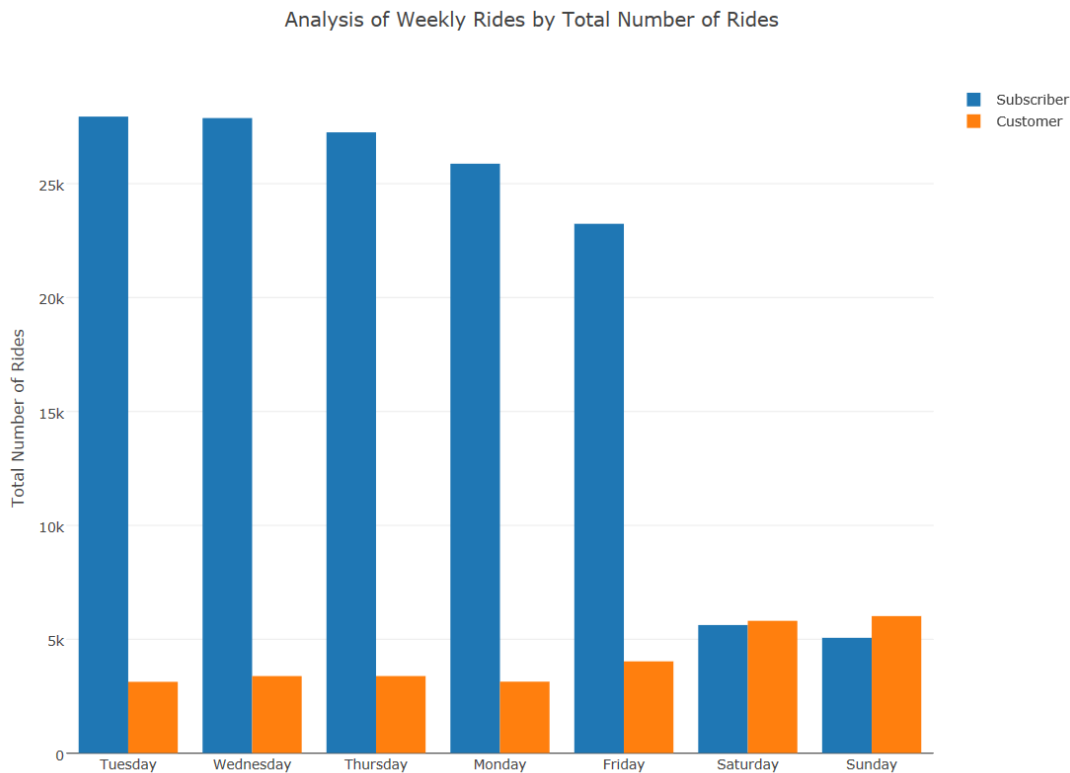


The graph shows the total number of rides plotted against total period of data. There is increase trend in number of rides from March to Aug.

Total Rides Grouped by Each Month

	March	April	May	June	July	August
Subscriber	20584	22137	23377	25151	26125	25508
Customer	4008	4082	5121	4836	5153	5705

Let's look at which days of all months are busiest



[images\weekly.html](#)

Weekdays Subscribers usage is enormously larger than customer's usage. On weekends usage suddenly drops very low with Customers and subscribers have equal share of usage.

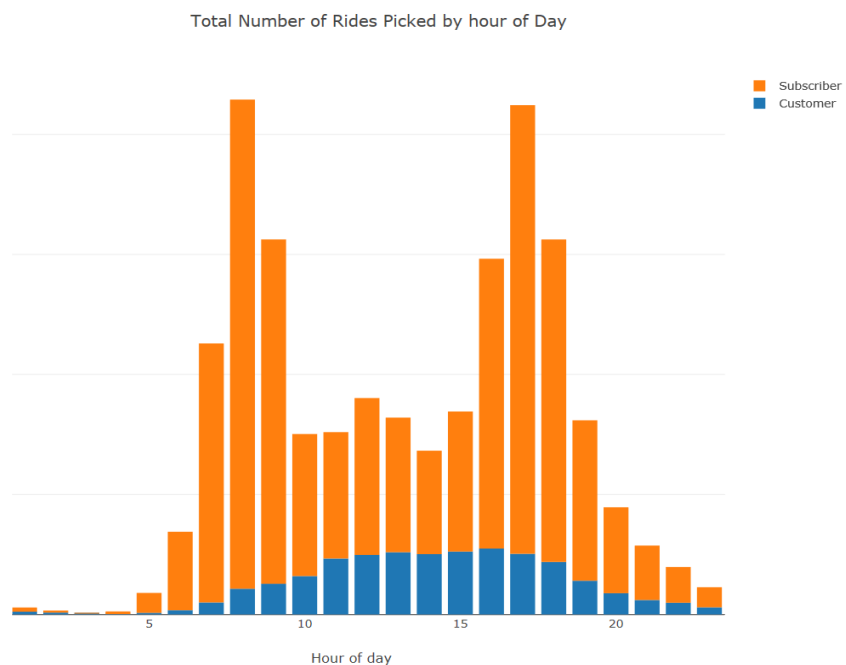
Average Number of Rides on Weekdays are 1135, on Weekends 429.

Interactive plot: [images\basic-hour.html](#)

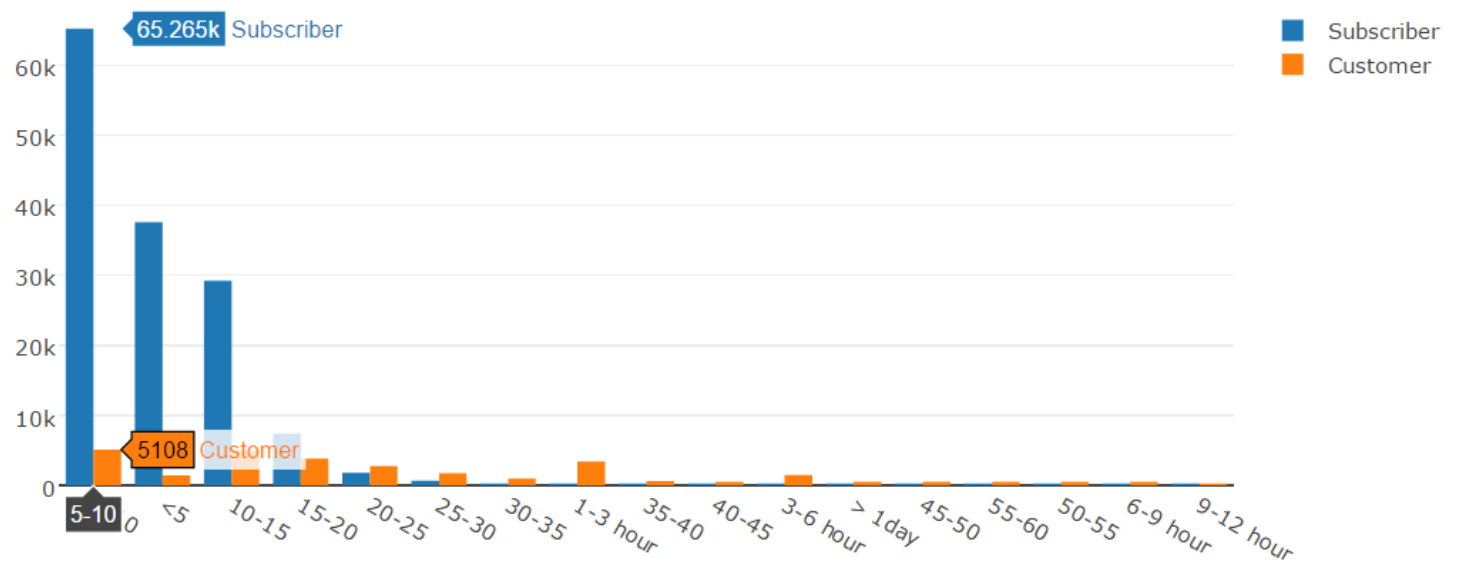
Hour of the day

From above graphs it is sure that most of the customers are subscribers and busiest period of week is weekdays. The Total Number of Rides by hour of the day graph shows that busiest period of the day is from 8:00 AM to 10 AM and 4:00 PM to 7:00 PM.

This indicates that bike share may be mostly used during work days and commute hours of the day. There is an increase in Customers usage during 10:00Am to 3:00PM. These users are one who may get time out during these periods and likes to take short rides.



Analysis for length of trips taken:



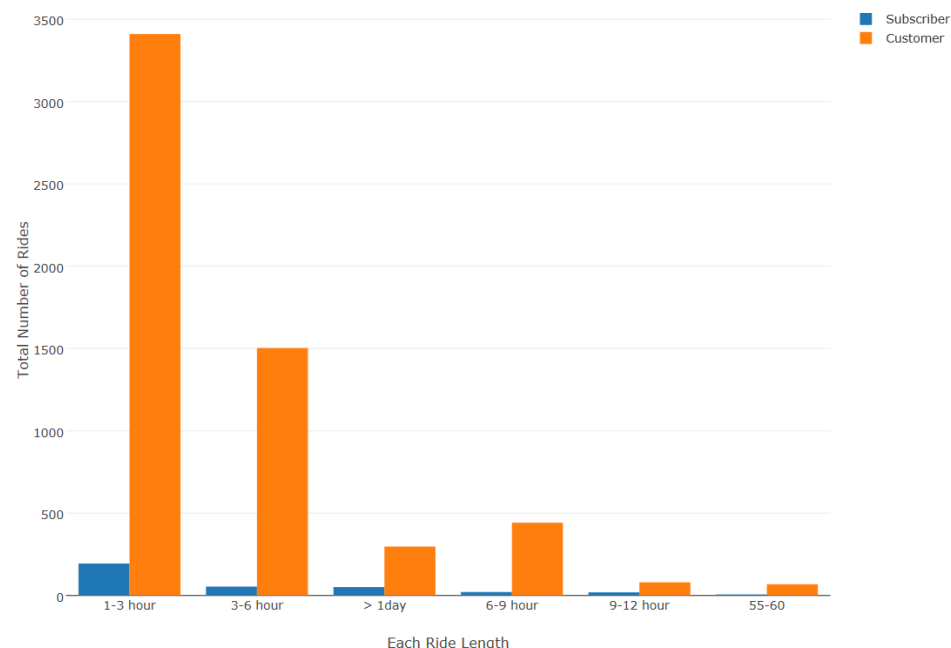
interactive visual:[images\d_bucket.html](#)

Graph shows total number of rides to the ride length in minutes and hours. Most of the rides are below 35minutes duration. Graph shows 70,373 rides (65265 Subscriber + 5108 Customer rides) falls into 5-10-minute ride interval.

In general bike share is for short trips. So, if pricing is considered, there can be a penalty charge applied who takes trips longer than 1 hour. If data is considered only for rides whose length is greater than 1 hour. 94% of penalty riders fall under customers and subscribers are only accountable for 6% penalty rides.

	<5	5-10min	10-15min	15-20min	20-25min	25-30min	30-35min	35-40min	40-45min	45-50min	50-55min	55-60min
Customer	1450	5108	4992	3846	2787	1772	996	641	471	414	372	329
Subscriber	37624	65265	29273	7392	1787	674	255	110	63	47	24	29

Analysis for Ride length more than 1 hour by Total Number of Rides



Interactive
visual:[images\morethan hour.html](#)

	1-3 hour	3-6 hour	6-9 hour	9-12 hour	>1day
Customer	3410	1503	442	80	297
Subscriber	194	54	21	19	51

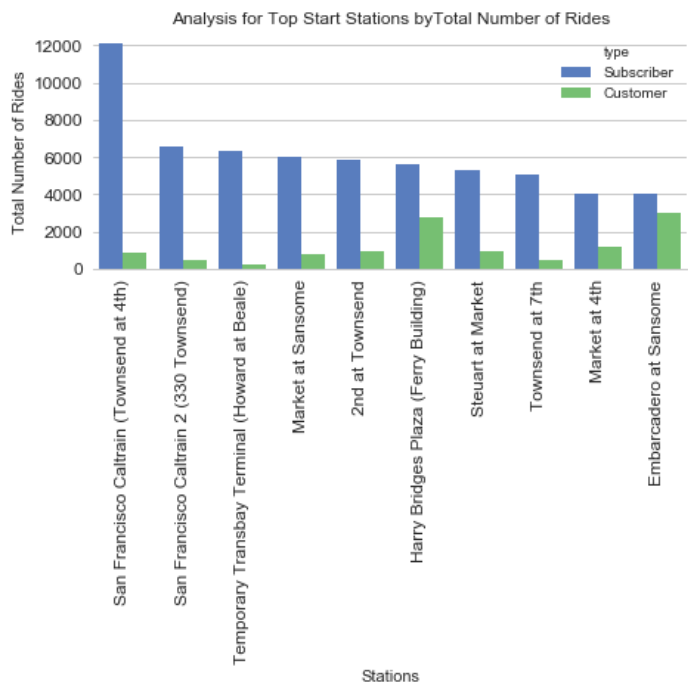
About 57% of rides longer than 1 hour falls under 1-3-hour group. Total Number of rides longer than 1hour are 6145. 347 Rides are more than 1 day. Rides falling under categories >12-hour and >1 day period can be false alarm (wrongly noted data) or unpredicted circumstances for customer or stolen. A pricing system can be developed depending time of usage and delay period.

Analysis for Top Starting Stations and Destinations

Analyzing starting stations and destinations or routes which are most common i.e., having most number of rides helps to analyze bike share company capacity. By understanding these metric number of terminals or bikes can be increased or decreased in routes. We understood that or customers are mostly from Subscribers analyzing their behavior in the top routes can lead to efficient tracking of system.

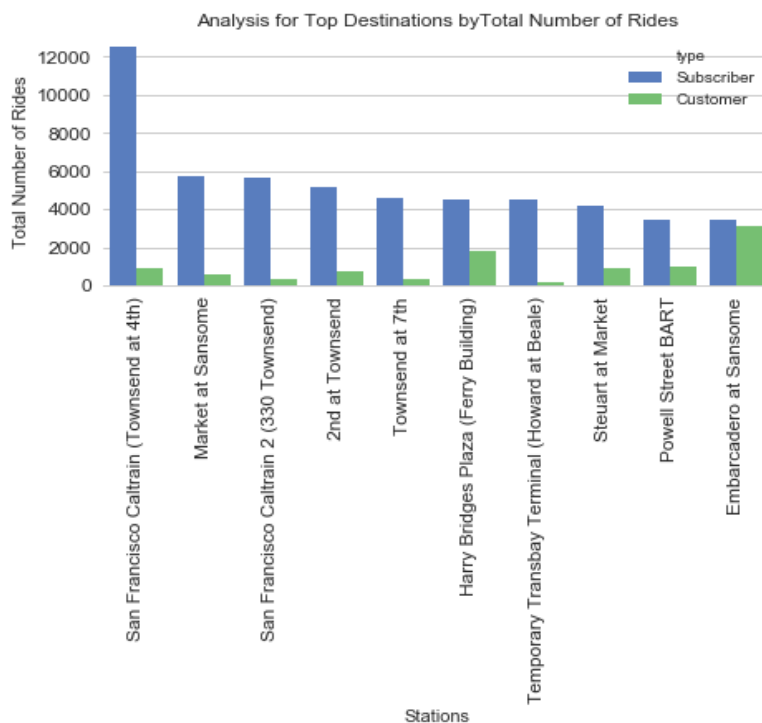
Top Starting Station	Number of Rides
San Francisco Cal train (Townsend at 4th)	12950
Harry Bridges Plaza (Ferry Building)	8336
Embarcadero at Sansome	7010
San Francisco Caltrain 2 (330 Townsend)	7008
2nd at Townsend	6824
Market at Sansome	6819
Temporary Transbay Terminal (Howard at Beale)	6540
Steuart at Market	6238
Townsend at 7th	5479
Market at 4th	5241

Interactive visual: <images\Top-start-station.html>



Top Destination	Number of Rides
San Francisco Cal train (Townsend at 4th)	16732
Embarcadero at Sansome	8141
Harry Bridges Plaza (Ferry Building)	7908
Market at Sansome	7879
San Francisco Caltrain 2 (330 Townsend)	7482
2nd at Townsend	7411
Steuart at Market	6328
Townsend at 7th	6142
Temporary Transbay Terminal (Howard at Beale)	5758
Powell Street BART	5448

Interactive visual:<images\top-end-station.html>



Graphs for top 10 stations to start and destination to reach in the data are plotted to number of rides. San Francisco Cal train (Townsend at 4th) top of list both as starting station with 12950 rides and as destination with 16732 rides. The common stations that are used are near Townsend, market place and Ferry Building. May be this is due to many corporate companies present near the location where people use bike daily for commute. Commuters living near these places are mainly using bike sharing system. With more data on nearest landmarks or locations or housing places, population wise analysis can be done.

From the graphs above Embarcadero at Sansome station is most popular among Customers compared to subscribers followed by Harry Bridges Plaza (Ferry Building). Powell Street BART has not have much traffic as others but customers with pass has good usage compared to other place. This may be because customers are using BART as transition for transportation and using rail transport.

Top 4 Popular stations among Customers

	Embarcadero at Sansome	Harry Bridges Plaza (Ferry Building)	Powell Street BART	San Francisco Cal train (Townsend at 4th)
Start Station Rides	3002	2735	1083	873
Destination Rides	3883	2276	1184	1136

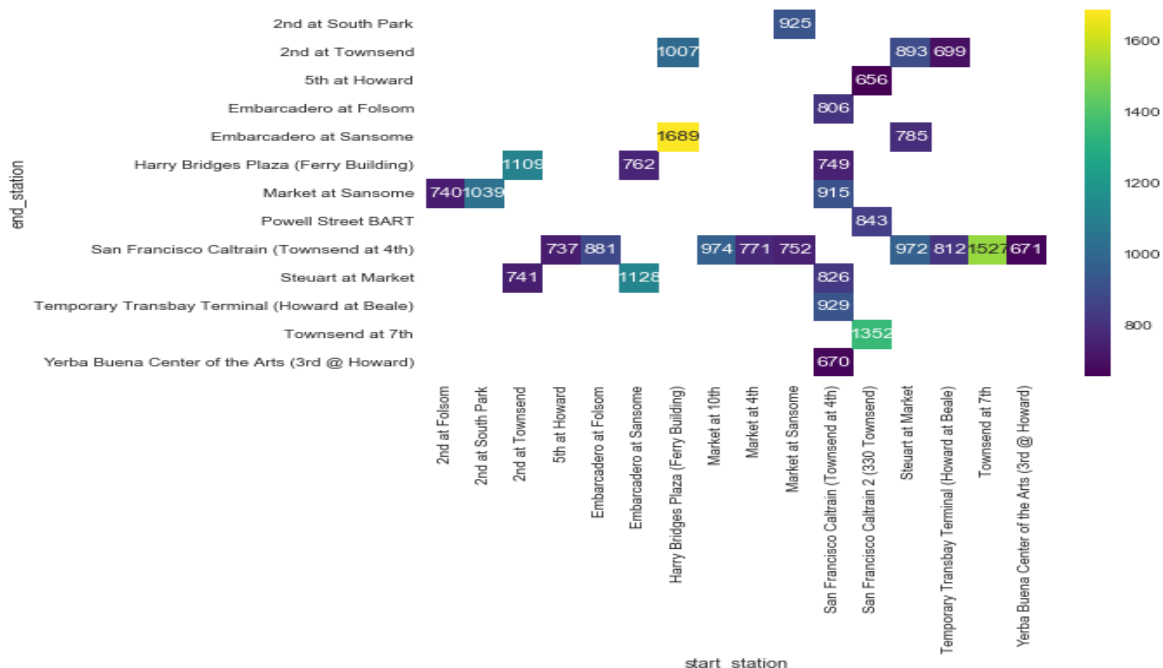
Top 4 Popular Stations among Subscribers:

	San Francisco Caltrain (Townsend at 4th)	Market at Sansome	San Francisco Caltrain 2 (330 Townsend)	2 nd at Townsend
Start Station Rides	12077	5996	6573	5889
Destination Rides	15596	7185	7100	6462

Analyzing individual Route:

In Bike share analysis each trip can be analyzed and route popularity can be visualized using heatmap. Starting Stations are on x-axis and ending stations are on y-axis. The count on heatmap shows total number of times the route is used. The larger the number and near to yellow shade the route is more popular. The most popular route is between Harry Bridges Plaza (Ferry Building) and Embarcadero at Sansome with 1689 rides. But the return route is not that popular. May be this is due that Embarcadero is a connecting station or people use other means of transport from there. The most popular two way used route is between Townsend at 7th and San Francisco Caltrain (Townsend at 4th) with 1527 rides and vice versa with 1352 rides. Next popular route is Harry Bridges Plaza (Ferry Building) and 2nd at Townsend with 1007 rides and vice versa 1109. The Harry Bridges Plaza attracts more tourist and high financial market place. It is closed to Embarcadero. Near the Townsend at 7th there are lot of technology companies which attracts commuters for shortest distances by various memberships.

Starting Station	Destination	Rides
Harry Bridges Plaza (Ferry Building)	Embarcadero at Sansome	1689
Townsend at 7th	San Francisco Caltrain (Townsend at 4th)	1527
San Francisco Caltrain 2 (330 Townsend)	Townsend at 7th	1352
Embarcadero at Sansome	Steuart at Market	1128
2nd at Townsend	Harry Bridges Plaza (Ferry Building)	1109
2nd at South Park	Market at Sansome	1039
Harry Bridges Plaza (Ferry Building)	2nd at Townsend	1007
Market at 10th	San Francisco Caltrain (Townsend at 4th)	974
Steuart at Market	San Francisco Caltrain (Townsend at 4th)	972
San Francisco Caltrain (Townsend at 4th)	Temporary Transbay Terminal (Howard at Beale)	929



Interactive
visual:

<images/labelled-heatmap.html>

Summary:

Bike share is mostly used by subscribers for rides mostly under 30 minutes. Major part of analysis consisted of analyzing bike share logistics in respective to customer type. Throughout the time line number of rides are low on weekends and high on weekdays. If additional data like weather, holidays or events near by or games nearby is considered more interpretation can be obtained from data.

Analysis of customer route patterns, Ride Metrics can be done on holidays, events, special occasions.

Data regarding nearby landmarks, distance to landmarks or historical sights, or nearest railway station can be used to analyze usage patterns among tourists and subscribers.

Monetary metrics like Costs, Capital, Profits, Risk can be used to generate predictive models on risks or profit generation models.

If provided with longitude, latitude and geospatial data wide variety of visual maps can be developed to analyze the logistics. Like Real time tracking of routes. Interactive visuals on routes.

How to update tool:

1. Adding new data to existing one is major task in analysis. The following codes checks for duplicate values and copies only the new values of data.

From more_itertools import unique_everseen

With open('GDA_DATA.csv','r') as old_file,open("new.csv","w") as out_file:

```
Out_file.writelines(unique_everseen(old_file))
```

2. Estimation of demand of bike share using time series analysis.
3. Geo special maps can be created for interactive visualization.