*Sam Gutentag*

# Bay Area Bike Share Inferential Statistics

Identify Key User Groups

There are two key user groups in the data. Subscribers who purchase an annual membership and Customers who purchases rides in 24 hour or 3 day increments. Subscriber Trips account for 844,570 trips which is 88.39% of all trips and Customer trips account for 110,987 trips which is 11.61% of all trips.
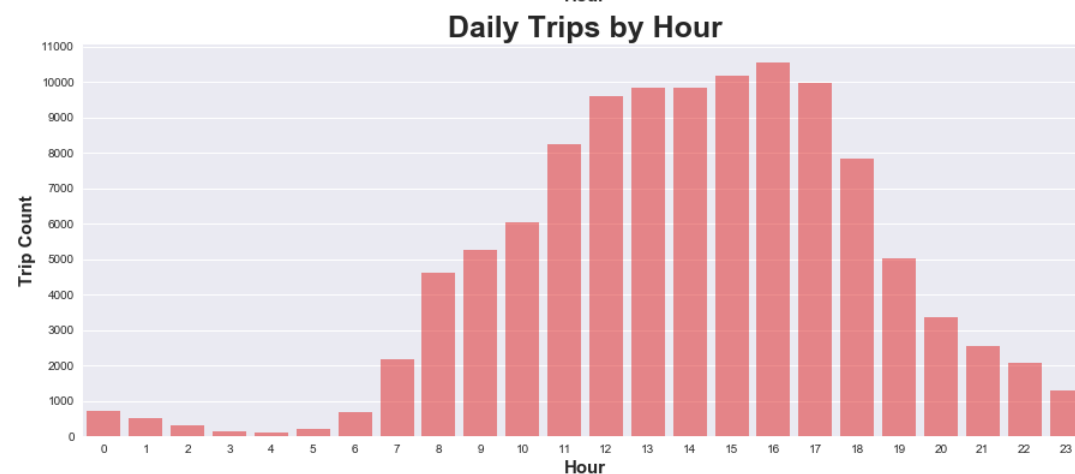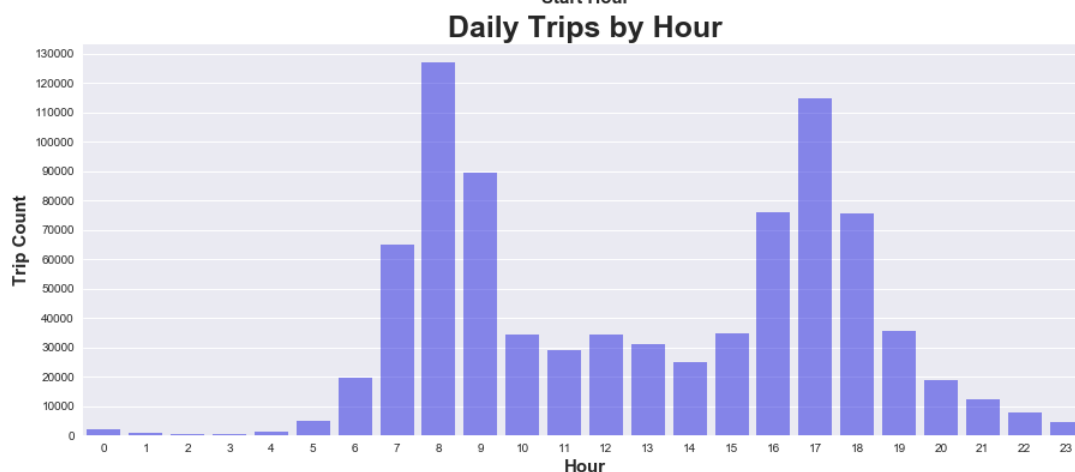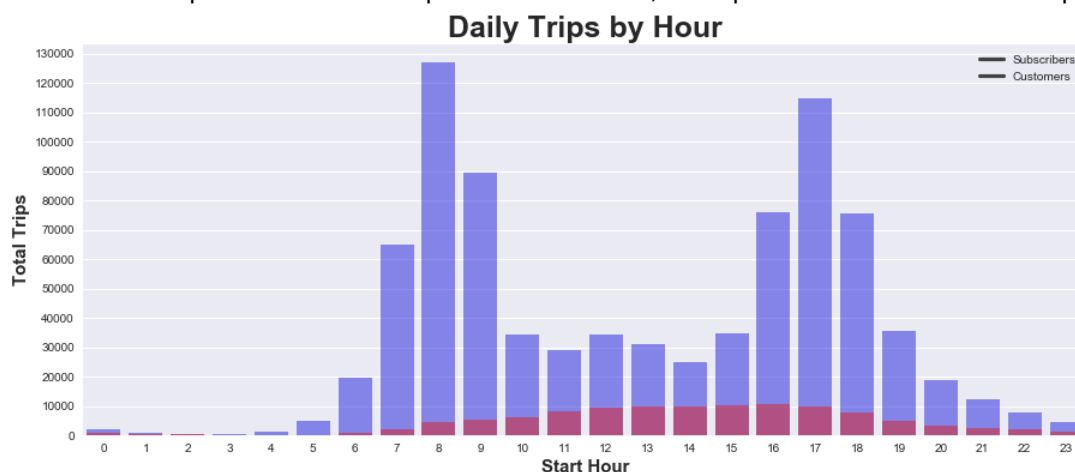
Identify Key Service Areas

The Bay Area Bike Share has stations in 5 service areas, San Francisco, San Jose, Redwood City, Mountain View, and Palo Alto. Very rarely does a single ride cross from one service area to another. These Non Regionalized Trips account for just 0.08% of all subscriber trips, and 0.42% of all Customer trips. From the quick overview, we can see that the bulk of Subscriber trips take place at specific stations in each region, while Customers appear to take a large number of round trips, that is starting and ending a ride at the same station.

Subscriber - Terminal Usage

Customer - Terminal Usage
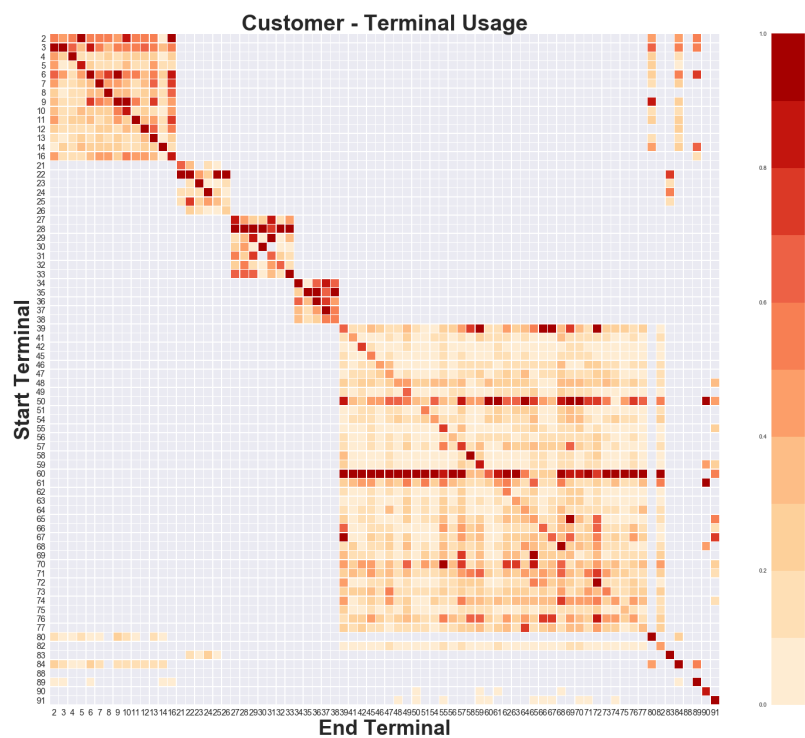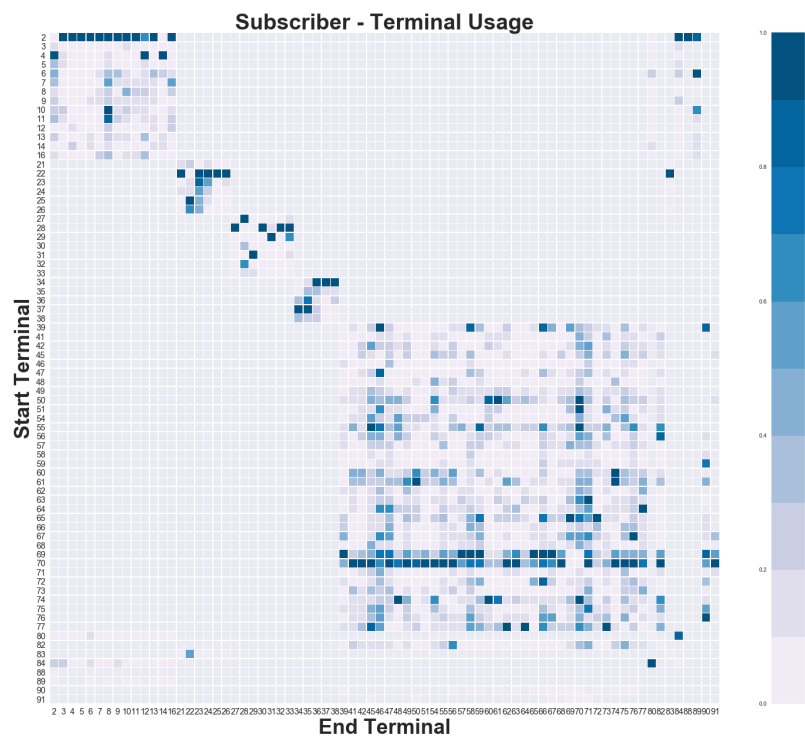
Stations Usage for Subscribers by Region

      Zooming in a bit, Subscribers heavily favor a distinct group of stations to start and to end their trips in each service region. San Francisco trips account for 91.21% of trips, San Jose 5.21%, Redwood City 0.48%, Mountain View 2.44%, and Palo Alto 0.66%. As such, we will focus on the San Francisco and San Jose Trips. In San Francisco, terminal 69 and 70 are located across the street from each other just outside of a Caltrain station.   Similarly, Station 2 in San Jose is also at a Caltrain station.
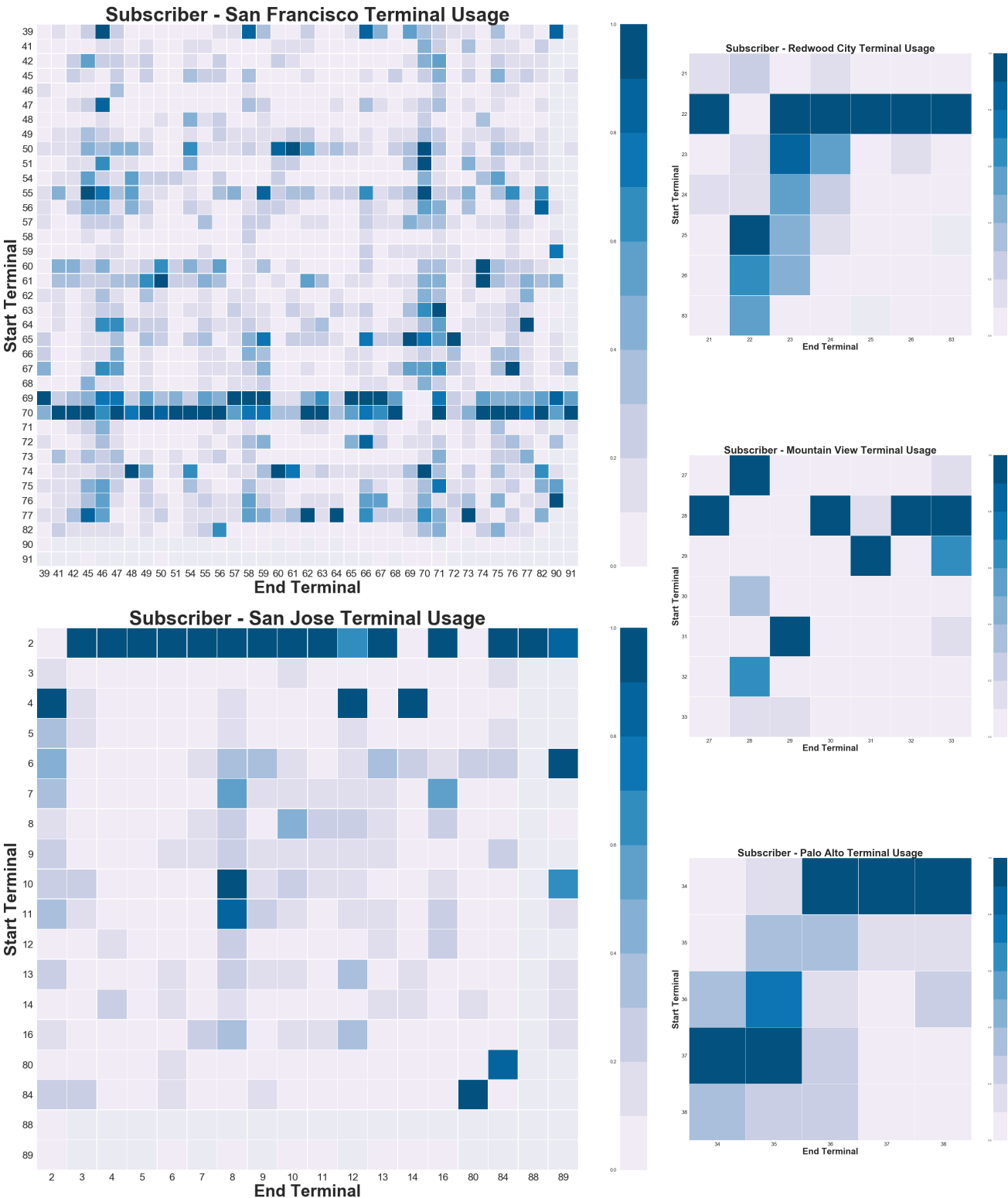
Stations Usage for Customers by Region

      Customers heavily favor a distinct group of stations to start and to end their trips, and the data indicates that there are a large number of trips that start and end at the same station.   San Francisco trips account for 89.15% of trips, San Jose 6.36%, Redwood City 0.46%, Mountain View 2.16%, and Palo Alto 1.86%. As such, we will focus on the San Francisco and San Jose Trips. In San Francisco, terminal 60 and 50 are located along the Embarcadero, a long flat street with large pathways for bike riders and tourists alike to view the city's waterfront.
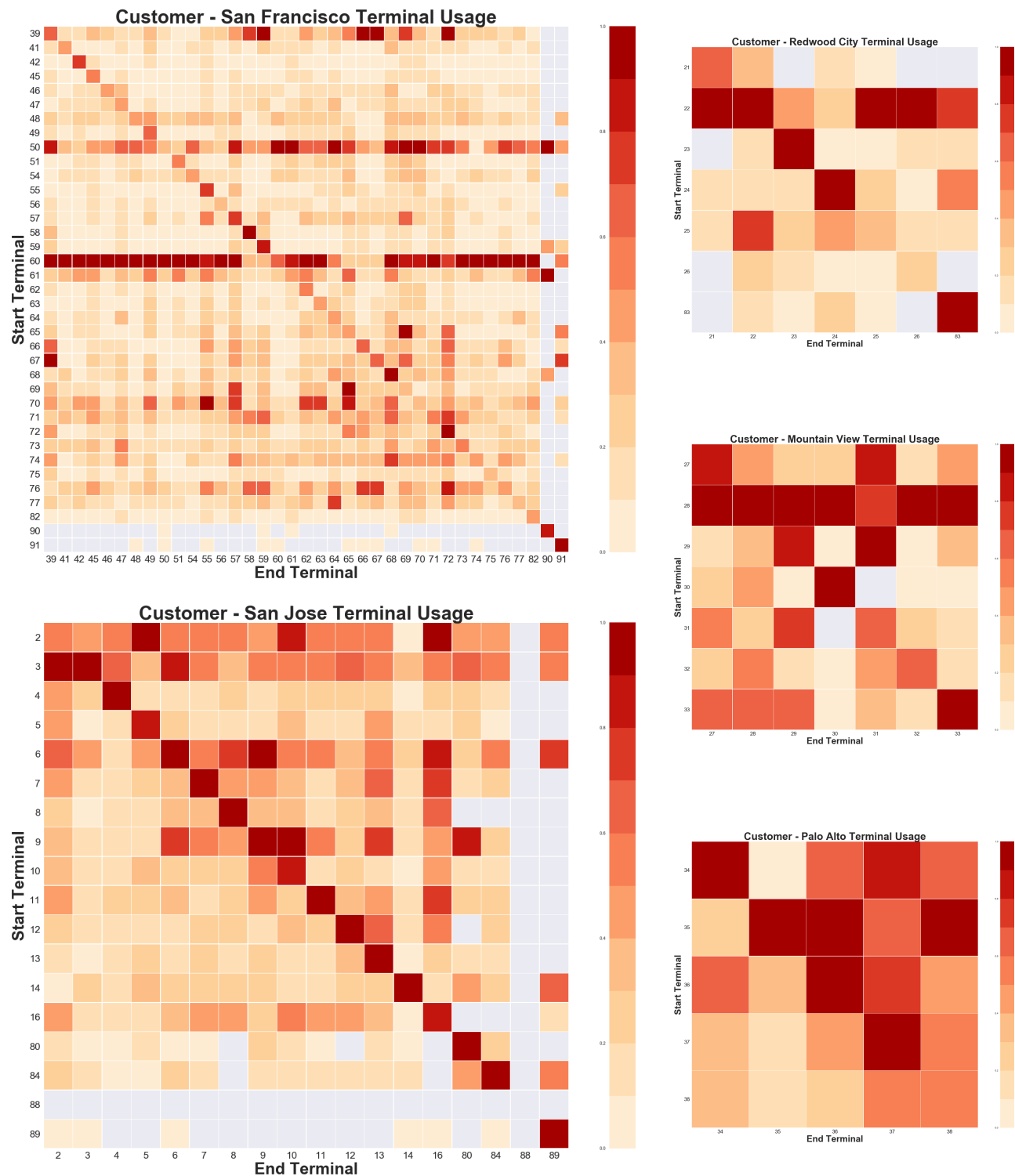
Making Sense of Subscriber Usage Patterns in San Francisco

Subscribers are most active in the San Francisco Service area, accounting for 91.21% of all trips taken by riders. From here we can focus effort on seeing that month to month usage is stable, with slight dips during November and December likely due to Holiday Travel or colder weather.
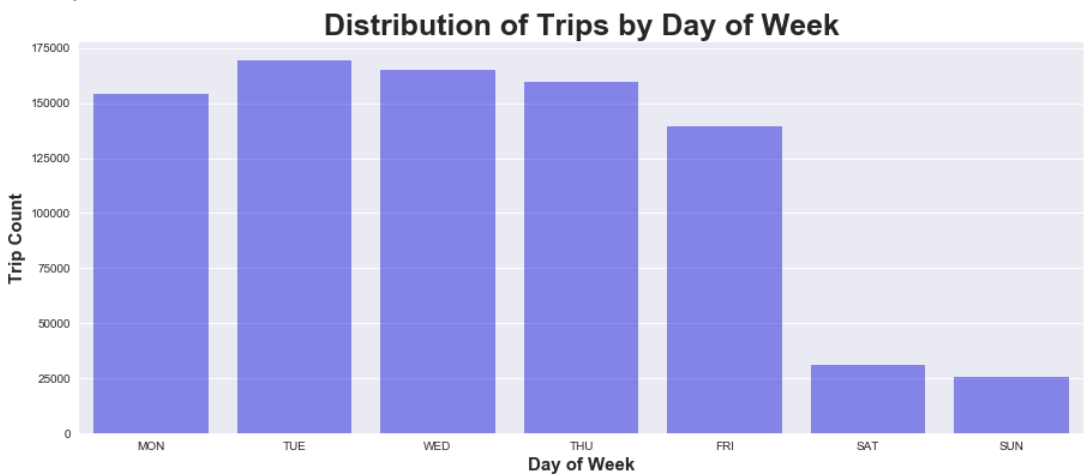
Distribution of Trips by Month of Year
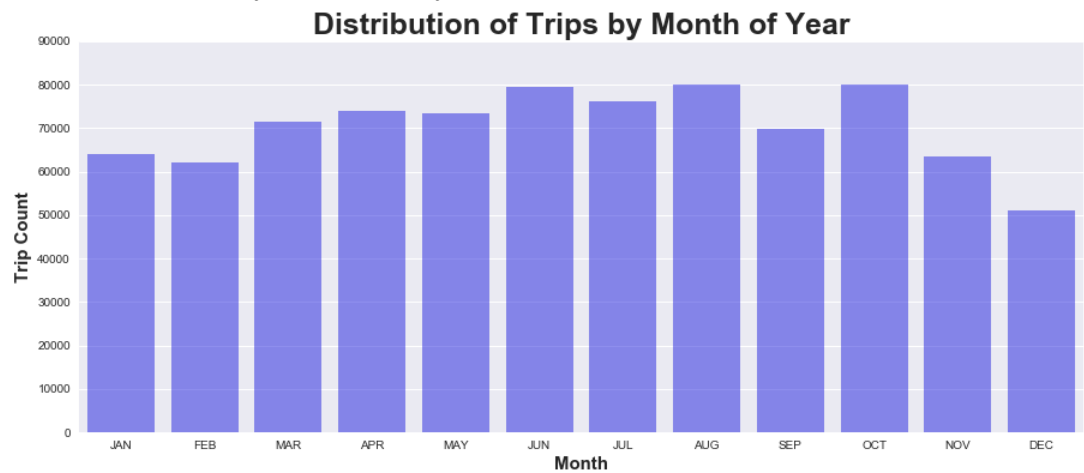
With daily stable monthly ridership rates, we continue diving in to see weekly subscriber usage and find that the vast majority, of trips are taken on weekdays, indicating that the most common Subscriber Trips are during work days.

Distribution of Trips by Day of Week

Finally, zooming in one last time to see that Subscriber usage clearly has visible peaks during morning (7a-10a) and evening (4p-7p) commute hours. These Morning Commuters account for 340,345 trips which is 36.61% of all trips, while Evening Commuters account for 318,936 trips which is 33.38% of all trips.

Subscriber Daily Trips by Hour - Weekdays Only
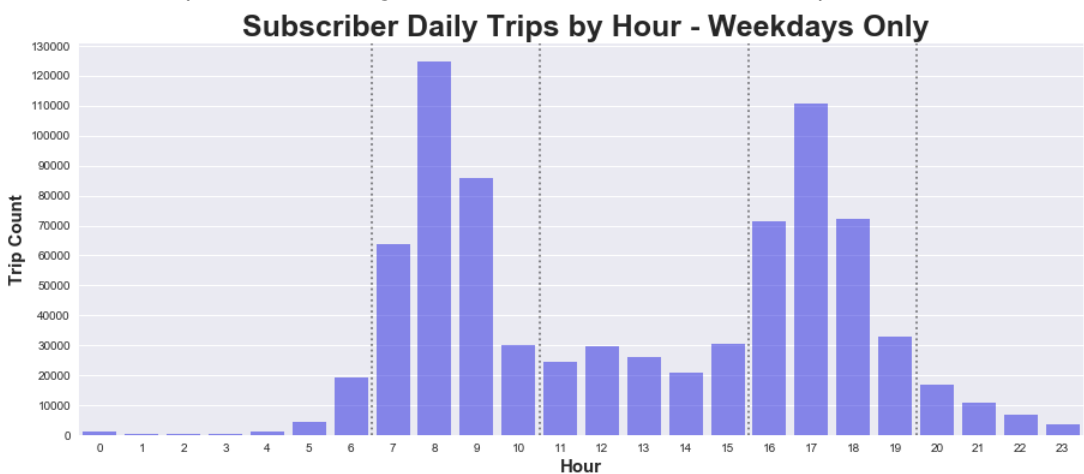
Making Sense of Customer Usage Patterns in San Francisco

       Customer usage is a smaller number of the overall usage, but has distinct trends of its own. Starting with a broad view, we see that Customer usage peaks in August and September, with nearly double the number of trips taken in Winter months.
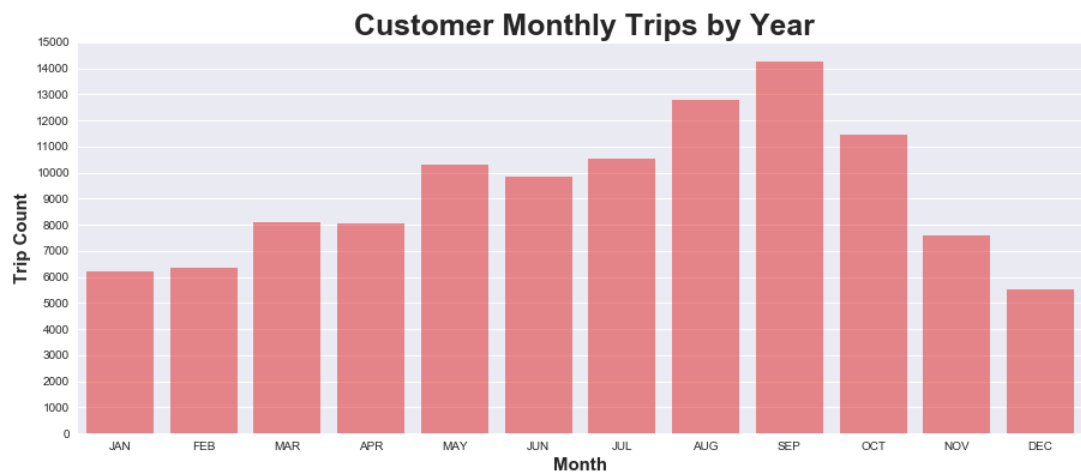
       Taking a look at daily trips, we note that the number of trips peaks on Weekends, with a bump of roughly 40% more trips on week end days than on week days.
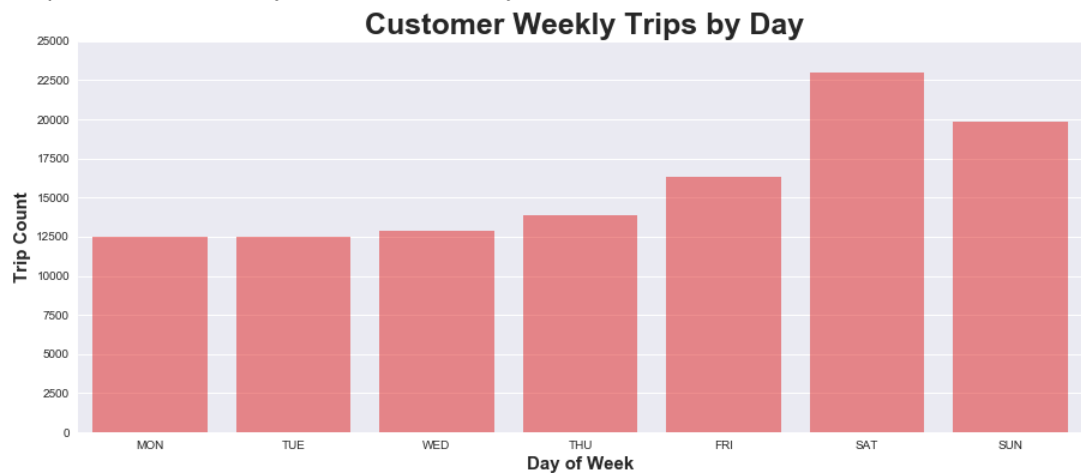
Customer Usage Weekday vs Weekend Patterns in San Francisco

Finally, weekend and weekday usage is quite different for Customers, they key difference being a spike in rides during the evening hours which coincide with the evening spike observed in the Subscriber data. This could be the result of Customers joining friends who are Subscribers on rides or for more sporadic commuters not committing to a Subscription Plan.

Most Important User Group - Commuters

Initial exploration shows that the largest user group of the Bay Area Bike Share Program are trips within San Francisco, taken by Subscribers accounting for 75.288% of all trips taken as part of the Bike Share Program, and 80.55% of all San Francisco local trips.

Within that group there are two distinct spikes in ridership, one during morning commute hours (7am-11am) and another during evening commute hours (4pm-8pm) on weekdays. Morning Commuter trips account for 30.27% of all trips in San Francisco while Evening Commuter trips account for 28.51% of all trips in San Francisco. This assumption that these riders are Commuting is further corroborated by looking at the top stations used in Morning Commutes and in Evening Commutes.

The top Commuter start and end terminals are all within three blocks of major public mass transit locations. Most notably, stations 70 and 69 are located across the street from each other, and are immediately outside the final Caltrain station in the city. Stations 61 and 65 are also two blocks East and two blocks West respectively of this same Caltrain station. The Embarcadero and Ferry Building stations are all just across the street from the Major ferry terminal into the city, and the Market Street stations are all within a block of BART stations, the local subway system.

	Top Morning Commute Start Terminal			Top Morning Commute End Terminal		
Rank	Station ID	Terminal Name	Share	Station ID	Terminal Name	Share
1	70	San Francisco Caltrain	14.58	70	San Francisco Caltrain	7.60
2	69	San Francisco Caltrain 2	11.64	65	Townsend at 7th	6.55
3	50	Harry Bridges Plaza Ferry Building	7.22	61	2nd at Townsend	6.54
4	55	Temporary Transbay Terminal	7.00	77	Market at Sansome	5.05
5	74	Steuart at Market	5.60	69	San Francisco Caltrain 2	4.28
6	61	2nd at Townsend	4.11	60	Embarcadero at Sansome	4.18
7	73	Grant Avenue at Columbus Avenue	3.82	55	Temporary Transbay Terminal	4.14
8	77	Market at Sansome	3.21	51	Embarcadero at Folsom	4.06
9	54	Embarcadero at Bryant	2.96	63	Howard at 2nd	4.02
10	67	Market and 10th	2.93	74	Steuart at Market	3.96
	Top Evening Commute Start Terminal			Top Evening Commute End Terminal		
Rank	Station ID	Terminal Name	Share	Station ID	Terminal Name	Share
1	70	San Francisco Caltrain	6.18	70	San Francisco Caltrain	17.29
2	65	Townsend at 7th	5.86	69	San Francisco Caltrain 2	12.09
3	61	2nd at Townsend	5.17	50	Harry Bridges Plaza Ferry Building	6.66
4	69	San Francisco Caltrain 2	4.68	55	Temporary Transbay Terminal	5.85
5	77	Market at Sansome	4.65	74	Steuart at Market	5.77
6	64	2nd at Southpark	4.36	77	Market at Sansome	4.31
7	60	Embarcadero at Sansome	4.16	61	2nd at Townsend	4.11
8	67	Market and 10th	4.00	65	Townsend at 7th	3.39
9	55	Temporary Transbay Terminal	3.97	39	Powell Street BART	3.05
10	74	Steuart at Market	3.97	60	Embarcadero at Sansome	2.95

Resiliency of Commuters

As a refresher, a 'Commuter Trip' is a trip taken by a Subscriber that starts and ends at a San Francisco terminal. Additionally it must start between 7am and 11am or between 4pm and 8pm on weekdays only. This subset accounts for 58.04% of all trips taken and is the largest of such category.

Dat Subset	Total Trips	Share of All Trips
All Trip Data	955557	100.00%
Subscriber Trip Data	769677	80.55%
Commuter Trip Data	554588	58.04%
Morning Commuter Trip Data	279169	29.22%
Evening Commuter Trip Data	275419	28.82%

TEST 1 - DOES TEMPERATURE AFFECT COMMUTER RIDERSHIP

Commuter trip numbers **are** meaningfully affected by hotter temperatures. Mean trip count increases 8.76%, or average <u>62.68 more trips</u>, on days with a Maximum Temperature above 75.86 degrees Fahrenheit

Commuter trip numbers **are** meaningfully affected by colder temperatures. Mean trip count drops 14.70%, or average <u>105.17 fewer trips</u>, on days with a Minimum Temperature below 45.41 degrees Fahrenheit

Temperature is split into three categories, 'Hot Days', 'Cold Days' and 'Normal Days'. Hot Days are days in which the maximum temperature is greater than one standard deviation above average. Cold Days are days in which the minimum temperature is one standard deviation below average. Leaving Normal days to be the cross section of days not included in neither Hot nor Cold Days.

<u>Test 1a - Hot Days vs Normal Days</u>

 Ho : Mean Commuter Trips on Hot Days = Mean Commuter Trips on Normal Days
 Ha : Mean Commuter Trips on Hot Days ≠ Mean Commuter Trips on Normal Days

<u>Results 1a - Mean Commuter Trips increase 8.76% on Hot Days</u>

 Margin of Error 35.48664
 Mean Difference 62.68656
 T Statistic   3.32146
 P Value   0.00094

<u>Test 1a - Cold Days vs Normal Days</u>

 Ho : Mean Commuter Trips on Hot Days = Mean Commuter Trips on Normal Days
 Ha : Mean Commuter Trips on Hot Days ≠ Mean Commuter Trips on Normal Days

<u>Results 1a - Mean Commuter Trips decrease 14.70% on Cold Days</u>

 Margin of Error 48.13000
 Mean Difference -105.17197
 T Statistic   -5.34202
 P Value   1.26081e-7

TEST 2 - DOES PRECIPITATION AFFECT COMMUTER RIDERSHIP

Commuter trip numbers **are** meaningfully affected by Rainy weather. Mean trip count decreases 39.64%, or average <u>282.48 fewer trips</u>, on days with precipitation above 0.46 inches.

This is a substantial dip in Ridership, but it is important to also note that in across all Data Records, there are only 23 'Rainy Days' across all 1099 days of recorded data.

Data is split into two categories, 'Rainy Days' and 'Dry Days'. Rainy Days are days in which the precipitation is greater than one standard deviation above average. Dry Days are days in which the precipitation is less than one standard deviation above average.

<u>Test 2 - Rainy Days vs Dry Days</u>

Ho : Mean Commuter Trips on Rainy Days = Mean Commuter Trips on Dry Days
Ha : Mean Commuter Trips on Rainy Days ≠ Mean Commuter Trips on Dry Days

<u>Results 2 - Mean Commuter Trips decrease 39.64% on Rainy Days</u>

Margin of Error	71.18552
Mean Difference	-282.48131
T Statistic	-5.72130
P Value	1.50476e-8

TEST 3 - DOES WIND AFFECT COMMUTER RIDERSHIP

Commuter trip numbers **are** meaningfully affected by Rainy weather. Mean trip count increases 7.58%, or average <u>53.99 fewer trips</u>, on days with precipitation above 0.46 inches.

Data is split into two categories, 'Windy Days' and 'Calm Days'. Windy Days are days in which the mean wind speed is greater than one standard deviation above average. Calm Days are days in which the mean wind speed is less than one standard deviation above average.

<u>Test 3 - Windy Days vs Calm Days</u>

Ho : Mean Commuter Trips on Windy Days = Mean Commuter Trips on Calm Days
Ha : Mean Commuter Trips on Windy Days ≠ Mean Commuter Trips on Calm Days

<u>Results 3 - Mean Commuter Trips decrease 7.58% on Windy Days</u>

Margin of Error	46.90469
Mean Difference	-53.99338
T Statistic	-2.36479
P Value	0.01828