Citi Bike Analysis

CitiBank has a New York City bike service where many bike stations are available to the general public. Any user may rent out a bike for a time to ride around the city. There is also a subscription service for these bikes where users can pay a monthly fee for an unlimited use per year. Many different people use this bike service.

Are there unifying variables, which connect all these users? What users is this service more popular with and how can a new market of people be convinced to use this bike service? How can Citibank continue to appeal to its current customers? How can CitiBank market using deals or ads?

Data Sources:

..

Screenshot of raw data:

A	B	C	D	E	F	G	H	I
trip_id	bike_id	weekday	start_hou	start_time	start_station_id	start_station_name	start_station_latitude	start_station_longitude
LnQzQk	16013	Mon	18	9/9/2013 18:18	523	W 38 St & 8 Ave	40.75466591	-73.991381
IL9boN	15230	Thu	18	9/12/2013 18:38	257	Lispenard St & Broadway	40.71939226	-74.002472
46clGB	17942	Wed	19	9/18/2013 19:44	479	9 Ave & W 45 St	40.76019252	-73.99125
v7vdFt	19683	Sat	11	9/28/2013 11:54	527	E 33 St & 1 Ave	40.74315566	-73.974347
VGBsb5	18024	Sat	18	9/7/2013 18:08	521	8 Ave & W 31 St	40.75044999	-73.994810
HbKJBz	14581	Fri	22	9/27/2013 22:01	293	Lafayette St & E 8 St	40.73028666	-73.99076
aOKk3s	15022	Sun	13	9/29/2013 13:13	380	W 4 St & 7 Ave S	40.73401143	-74.002938
kpcluW	16253	Sat	10	9/21/2013 10:48	346	Bank St & Hudson St	40.73652889	-74.006180
GC8mJC	16745	Fri	12	9/20/2013 12:29	510	W 51 St & 6 Ave	40.7606597	-73.980420
qyXq7X	19556	Tue	16	9/3/2013 16:56	446	W 24 St & 7 Ave	40.74487634	-73.995298
aZsCdK	15732	Sat	20	9/7/2013 20:01	457	Broadway & W 58 St	40.76695317	-73.981693
1tS1Gm	15316	Fri	23	9/6/2013 23:02	251	Mott St & Prince St	40.72317958	-73.994800
5ey1zk	15522	Sun	2	9/22/2013 2:02	393	E 5 St & Avenue C	40.72299208	-73.979954
AvwgC6	15503	Fri	9	9/13/2013 9:09	388	W 26 St & 10 Ave	40.74971775	-74.002950
6IMkxy	17172	Tue	15	9/24/2013 15:06	546	E 30 St & Park Ave S	40.74444921	-73.983035
p9S46q	15973	Mon	18	9/16/2013 18:45	167	E 39 St & 3 Ave	40.7489006	-73.976048

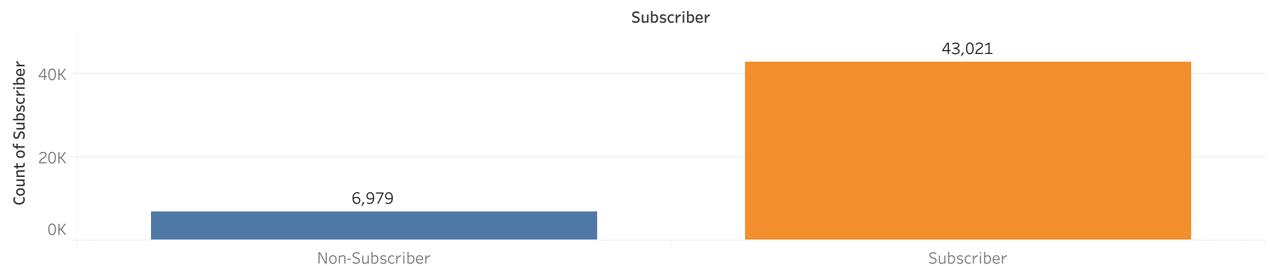
Description of each variable:

Variable	Description
trip_id	Bike trip's individual id
bike_id	Bike's individual id
weekday	Day of the week the bike trip occurred
start_hour	Hour of the day the bike trip occurred
start_time	Date and time the bike trip started
start_station_id	Station id of the station where the bike trip started
start_station_name	Station name where the bike trip started
start_station_latitude	Station latitude where the bike trip started
start_station_longitude	Station longitude where the bike trip started
end_time	Date and time the bike trip ended
end_station_id	Station id of the station where the bike trip ended
end_station_name	Station name where the bike trip ended
end_station_latitude	Station latitude where the bike trip ended
end_station_longitude	Station longitude where the bike trip ended
trip_duration	Trip's duration measured in seconds
subscriber	User's status on whether or not they are a CitiBike subscriber
birth_year	User's birth year, non-subscribers do not enter their birth year
gender	User's gender (0 = unknown, 1 = male, 2 = female)

CitiBike Analysis

Subscriber vs Non-Subscriber Count

Subscriber Count



Non-Subcribers are approxiamtely around 14% of riders and Subscribers are 86% of riders.

Both groups have similar start and end hours, but non-subscribers ride for longer. The groups have similar schedules, but ride for seemingly different purposes.

Subscriber Stats:

	start_hour	birth_year	end_hour	trip_duration_min
count	43021.000000	43021.000000	43021.000000	43021.000000
mean	14.084563	1975.627786	14.226703	12.843723
std	4.932676	11.089001	4.976645	8.724368
min	0.000000	1899.000000	0.000000	1.000000
25%	10.000000	1968.000000	10.000000	6.583333
50%	15.000000	1978.000000	15.000000	10.350000
75%	18.000000	1984.000000	18.000000	16.650000
max	23.000000	1997.000000	23.000000	44.950000

Non-Subscriber Stats:

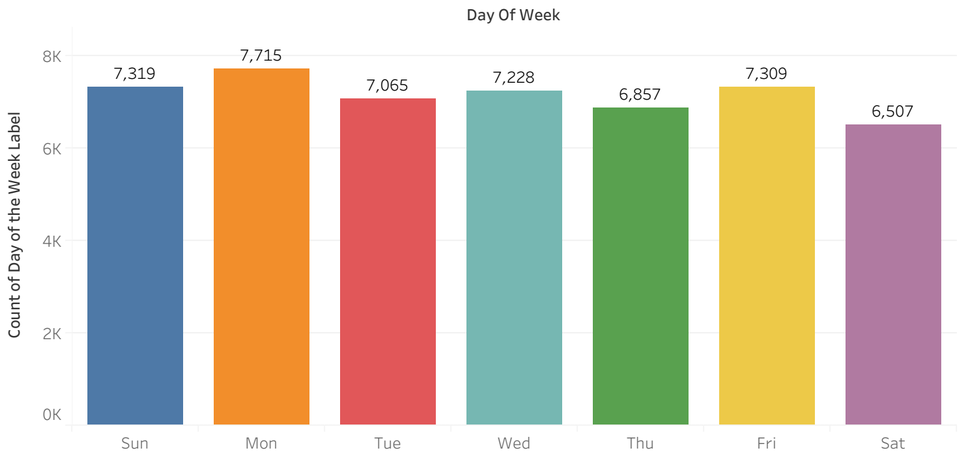
Non-subscribers do not have to enter information such as their birth year or gender, therefore they do not have statistics for the aforementioned varia..

	start_hour	birth_year	end_hour	trip_duration_min
count	6979.000000	0.0	6979.000000	6979.000000
mean	14.519272	NaN	14.807852	21.006250
std	4.371541	NaN	4.440933	11.331987
min	0.000000	NaN	0.000000	1.000000
25%	12.000000	NaN	12.000000	12.283333
50%	15.000000	NaN	15.000000	19.616667
75%	17.000000	NaN	18.000000	27.166667
max	23.000000	NaN	23.000000	44.950000

CitiBike Analysis

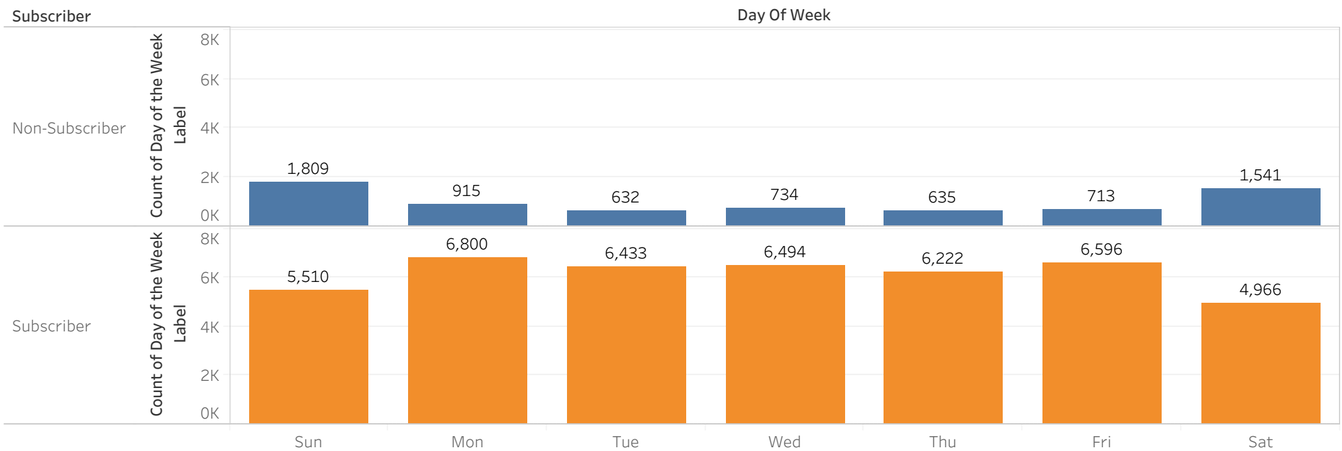
Day of the Week and Riders

Day of the Week Count: Subscribers and Non-Subscribers



Most days fall with in a similar range for number of rides, but Thursday and Saturday fall below 7,000 total rides. An sharp increase in the number of rides from Sunday to Monday and sharp decrease from Friday to Saturday, show the weekend has an impact on whether or not the average rider chooses to use the service.

Day of the Week Count: Subscribers vs Non-Subscribers

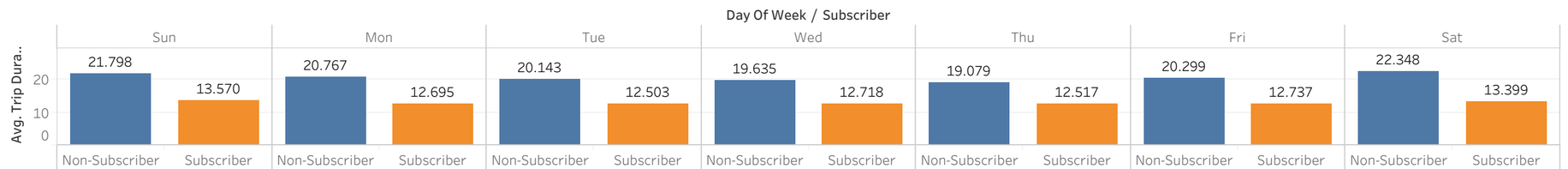


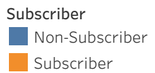
Subscribers and non-subscribers do not follow similar trends. Subscribers mostly use CitiBike on the weekdays, possibly to get to work. Non-subscribers have a high use for the weekends. This higher weekend use along with the longer average ride time shown in the previous slide points to non-subscribers using CitiBike for leisure instead of work.

CitiBike Analysis

Trip Length:

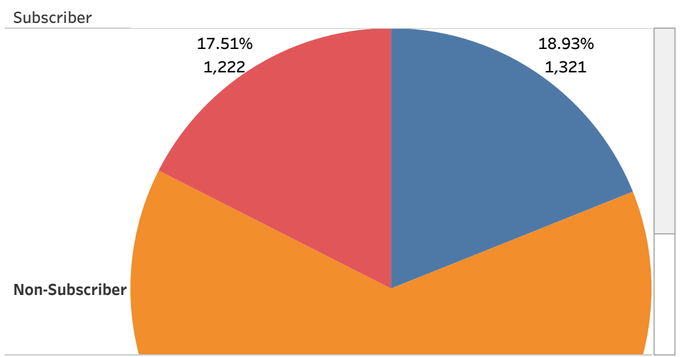
Average Trip Length Per Day Sub vs Non



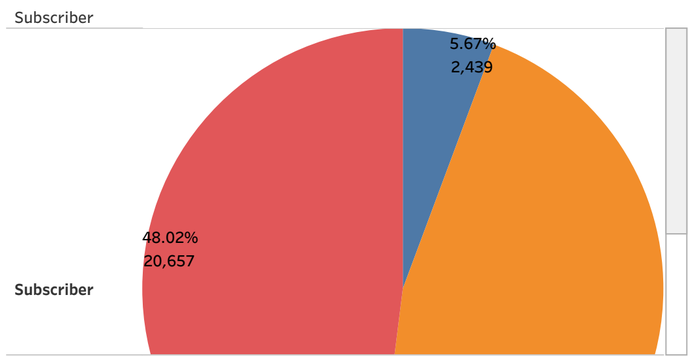


Through out every single day, non subscribers have longer rides than subscribers. Saturday and Sunday have the biggest difference between non-subscriber and subscriber times. These differences in the weekend ride times again show non-subscribers use CitiBike for more leisurely activities.

Trip Length Non-Sub



Trip Length Sub



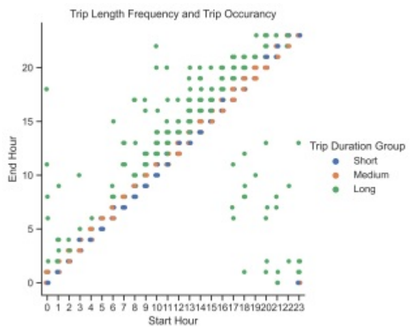
Ride times have been divided into three different groups.

Short: 10 minutes or less
Medium: More than 10 minutes, but less than 30 minutes
Long: 30 minutes or more

Subscribers have much shorter ride times. Their short category takes up al..

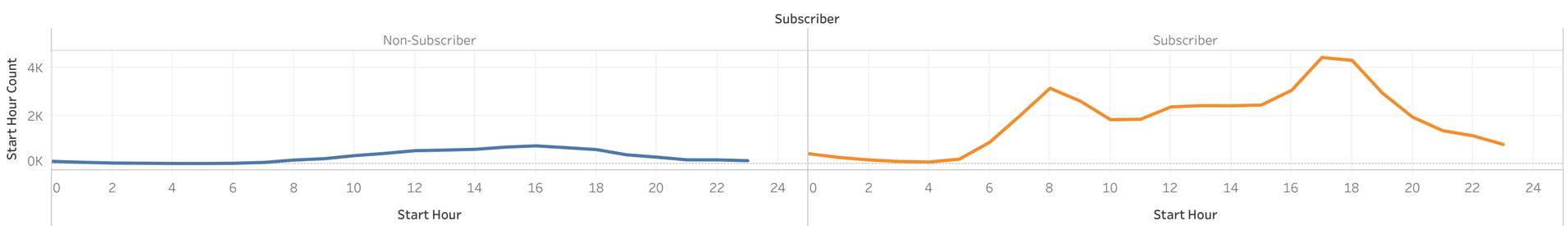
CitiBike Analysis

Start Hour:



Riders have various trip lengths at any time through out the day. Start hour is not an important variable in trip length.

Start Hour Count Subscriber vs Non-Subscriber



Subscriber
Non-Subscriber
Subscriber

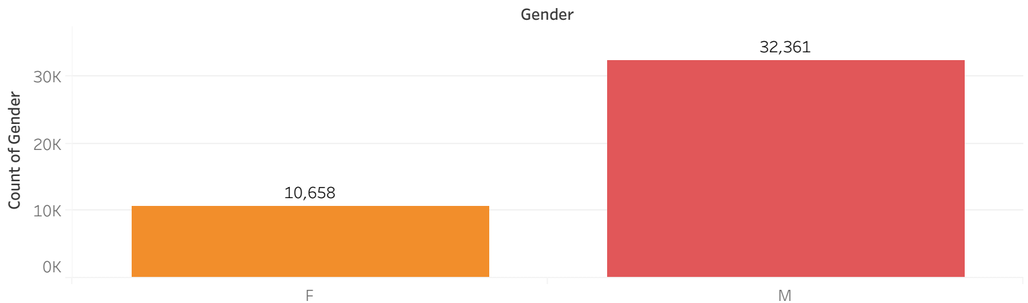
Non-subscribers' start hour is usually between 8:00 am to around 6:00 pm. There is a gradual increase from 8:00 am until it peaks at 4:00 pm and declines afterwards. After 6:00 pm there is a noteable drop where there is not much usage until 8:00am again.

Subscribers have two different peak times at 8:00 am and 6:00 pm. The amount of trips is inconsistent afterwards. These tim..

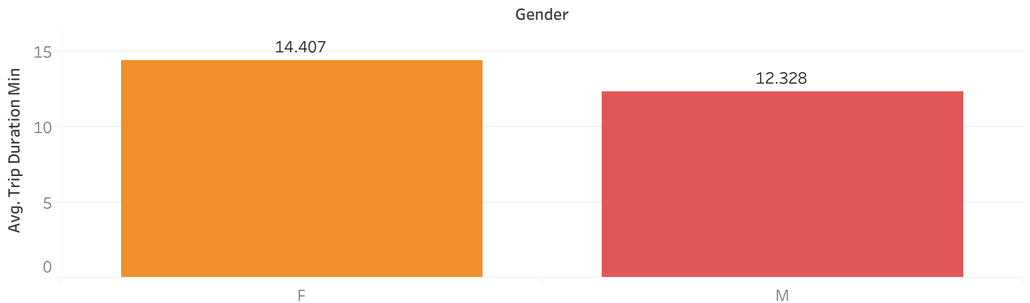
CitiBike Analysis

Subscriber Statistics

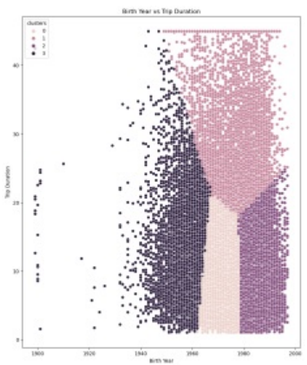
Gender Count



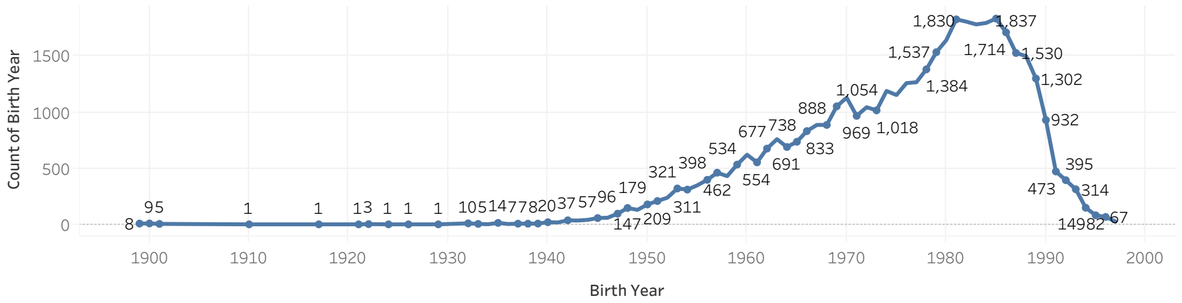
Gender vs Trip Length



Gender
F
M



Birth Year Count



The majority of CitiBike users are males, around 64%. The marketing department could develop marketing campaigns, which targets the female demographic.

Women are shown to have higher on average trips than their male counterparts. This trip length suggests t..

A large portion of riders are between the ages of 28 to 43, born between 1970 to 1985.

There should be marketing targeted towards both seniors and the younger crowds. A senior or student discount could draw more people from these demographics.

Additionally, the cluster map shows riders in similar age groups have similar trip lengths, which could possibly mean age group is a factor on why a rider is using CitiBike.

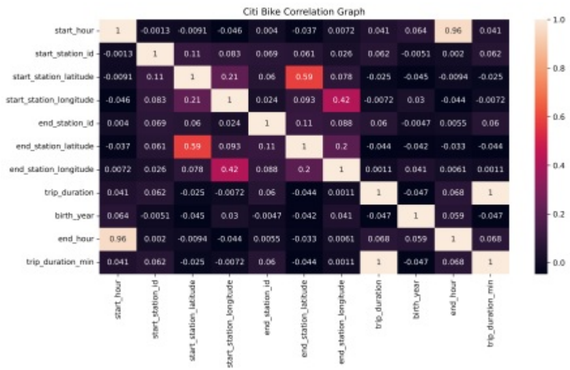
CitiBike Analysis

Riders and Travel Locations

Choropleth Map Showing Riders Where They Start Their Trips:

Manhattan and The Bronx has the most rides. Brooklyn follows the first two boroughs while Staten Island has the least besides Queens. Queens does not have any CitiBike service.

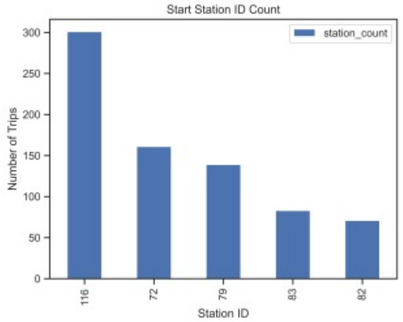
CitiBike should set up stations in Queens to capitalize on this untapped area. Marketing should create ads which appeal to people living in Brooklyn or Staten Island. Discounts could be given to people using s..

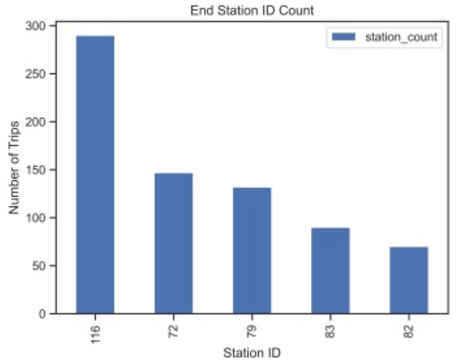


This correlation graph shows the start station latitude and the end station latitude have a 0.59 correlation. The end station longitude and the end station longitude have a 0.42 correlation.

Many riders use the start at the same locations and end up at the same locations as well.

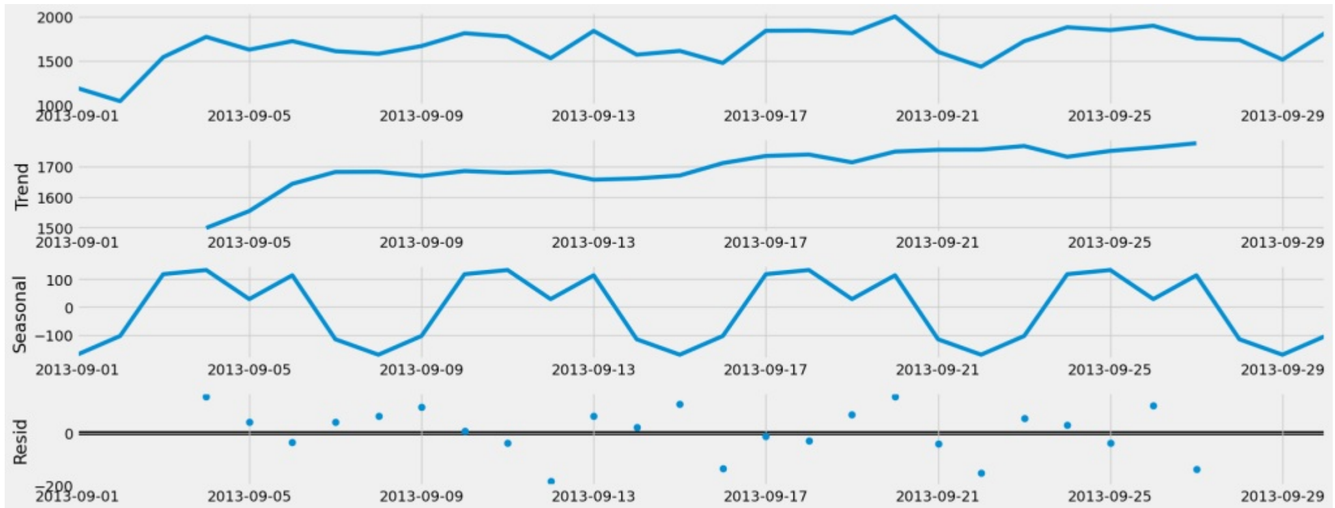
These station ID count graphs show many of the same riders frequent the same routes. Many trips start and end at the same stations.





CitiBike Analysis

CitiBike Trends



CitiBike is shown to be successful as its trend line shows an increase in rides throughout the month. Users, subscribers and non-subscribers, are interested in using the service for one reason for another.

The seasonal line confirms riders use CitiBike less during the weekends as opposed to the increase of rides during the weekdays.

CitiBike Analysis

Conclusions

Summary and Recommendations:

Subscriber vs Non-subscribers:

Subscribers make up a higher percentages of riders than non-subscribers, but both have similar statistics in start and end hours.

Subscribers mostly use the CitiBike service on the weekdays, possibly for work. Non-subscribers use their rides mostly on the weekend. Start Hours are similar for both types of riders, since ride count increases for both at 8:00 am. Futhermore, ride count declines after 6:00 pm for both types of riders as well. Both types of riders have similar schedules, but use CitiBike for different purposes.

The two rider types have different average trip lengths as well. The average non-subscriber ride is longer than its subscriber counterpart during any day of the week.

CitiBike should do several things to either gain new subscribers from their non-subscriber users or encourage non-subscribers to ride for longer. CitiBike should offer a sign on bonus which includes free rides during the weekend for a period of time, providing a discount for the first month based on the free rides is a recommendation. In order to still capitalize on people who do not wish to subscribe, CitiBike should offer a discounted rate after 25 min of ride time. Many of the average ride times throughout the month are almost around 20 min, almost at the discounted time range.

Subscriber Stats:

The majority of CitiBike users are males by a considerable margin. Women are shown to have larger average trips than men, which suggests stations many not be located in places convenient for women. Furthermore, the average age for a CitiBike user between their late 20s to early 40s. CitiBike could provide a discount to subscribers who are students and seniors, encouraging people outside the core age demographic to continue to use CitiBike. Additionally, CitiBike could send out surveys to females and people outside the core age demographic to see how they would improve the service. The additional data could inform further changes.

citibank