Siddhesh Jadhav 30-08-21

Google Data Analytics Course Capstone Project

Case Study: How Does a Bike-Share Navigate Speedy Success?

Cyclistic Company, Chicago

Background information

I am working for Cyclistic, a bike-sharing company. Bikes can be unlocked from one station and returned to any other station in the system anytime.

Cyclistic has flexible pricing plans: single-ride passes, full-day passes, and annual memberships. Customers who purchase single-ride or full-day passes are referred to as casual riders. Customers who purchase annual memberships are Cyclistic members.

Director of marketing believes the company's future success depends on maximizing the number of annual memberships as finance analysts have concluded that annual memberships are much more profitable than casual riders. She also believes that there is a good chance of converting casual riders to members as they are already aware of Cyclistic program and have chosen it for their mobility needs.

Business Task

To analyze user behaviours on how annual members and casual riders use Cyclistic bikes differently to make recommendations on how to convert casual riders into annual members.

Dataset used and its limitations

Dataset used in the analysis is the 12 month data from April-20 to Mar-21 from Index of bucket "divvy-tripdata" by Motivate International Inc. under this license (Data License Agreement | Divvy Bikes).

Data won't be able to connect pass purchases to credit card numbers to determine if casual riders live in the Cyclistic service area or if they have purchased multiple single passes.

Data Processing

1. Started with removing duplicate records

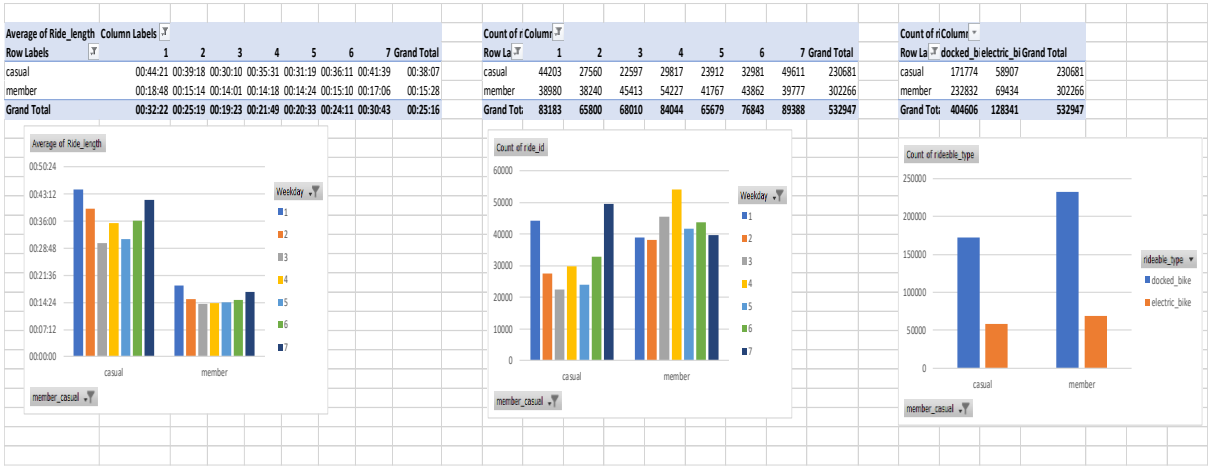
2. Removed NA values

3. Made additional fields for trip duration, weekday for which day of the week the trip took place etc.

4. Gave the data more aesthetic view.

started_at	ended_at	start_station_name	start_station_id	end_station_name	end_station_id	start_lat	start_lng	end_lat	end_lng	member_casual	Trip_duration	Weekday
17-09-2020 14:27	17-09-2020 14:44	Michigan Ave & Lake St	52	Green St & Randolph St	112	41.88669	-87.6236	41.88357	-87.6487	casual	00:17:13	5
16-09-2020 22:58	17-09-2020 00:09	Michigan Ave & 18th St	273	Michigan Ave & 18th St	273	41.8579	-87.6247	41.85795	-87.6246	casual	01:10:50	4
16-09-2020 23:05	16-09-2020 23:22	Halsted St & Polk St	108	Emerald Ave & 31st St	339	41.87188	-87.6466	41.83813	-87.645	casual	00:16:35	4
17-09-2020 17:42	17-09-2020 18:07	Albany Ave & Bloomingdale Ave	511	Albany Ave & Bloomingdale Ave	511	41.91392	-87.7052	41.91402	-87.7052	casual	00:25:33	5
16-09-2020 20:17	16-09-2020 20:28	Lake Shore Dr & Diversey Pkwy	329	Lake Shore Dr & North Blvd	268	41.93249	-87.6364	41.91174	-87.6268	casual	00:11:06	4
17-09-2020 16:43	17-09-2020 17:08	Elston Ave & Wabansia Ave	315	Damen Ave & Foster Ave	464	41.91291	-87.6643	41.97547	-87.6795	casual	00:25:33	5
17-09-2020 13:21	17-09-2020 13:36	Clybourn Ave & Division St	138	Campbell Ave & Fullerton Ave	504	41.90462	-87.6405	41.92448	-87.6893	casual	00:14:43	5
17-09-2020 06:34	17-09-2020 06:48	Campbell Ave & Fullerton Ave	504	Clybourn Ave & Division St	138	41.92462	-87.6893	41.9046	-87.6407	casual	00:14:28	5
17-09-2020 15:55	17-09-2020 16:30	Clark St & Lake St	38	Franklin St & Lake St	164	41.88525	-87.6311	41.88598	-87.6346	casual	00:35:06	5
14-09-2020 18:45	14-09-2020 19:13	Sawyer Ave & Irving Park Rd	485	Clarendon Ave & Gordon Ter	312	41.95358	-87.7093	41.95787	-87.6496	casual	00:28:12	2
14-09-2020 10:14	14-09-2020 10:23	Financial Pl & Ida B Wells Dr	89	Wells St & Huron St	53	41.87496	-87.6332	41.89455	-87.6342	casual	00:08:39	2
14-09-2020 17:16	14-09-2020 17:33	Wells St & Huron St	53	Federal St & Polk St	41	41.89482	-87.6345	41.87227	-87.6295	casual	00:16:06	2
13-09-2020 19:26	13-09-2020 19:44	Shedd Aquarium	3	McClurg Ct & Erie St	142	41.86727	-87.6155	41.89424	-87.6178	casual	00:18:10	1
14-09-2020 18:40	14-09-2020 18:51	Bissell St & Armitage Ave	113	Wells St & Elm St	182	41.91848	-87.6522	41.90313	-87.6345	casual	00:10:36	2
14-09-2020 13:11	14-09-2020 13:45	Franklin St & Lake St	164	Franklin St & Lake St	164	41.88599	-87.6349	41.88659	-87.6353	casual	00:34:08	2

5. Made Pivot Tables for each month with different categories such as Trip duration, no. of rides per day of the week, and Type of bike.



Data Analysis

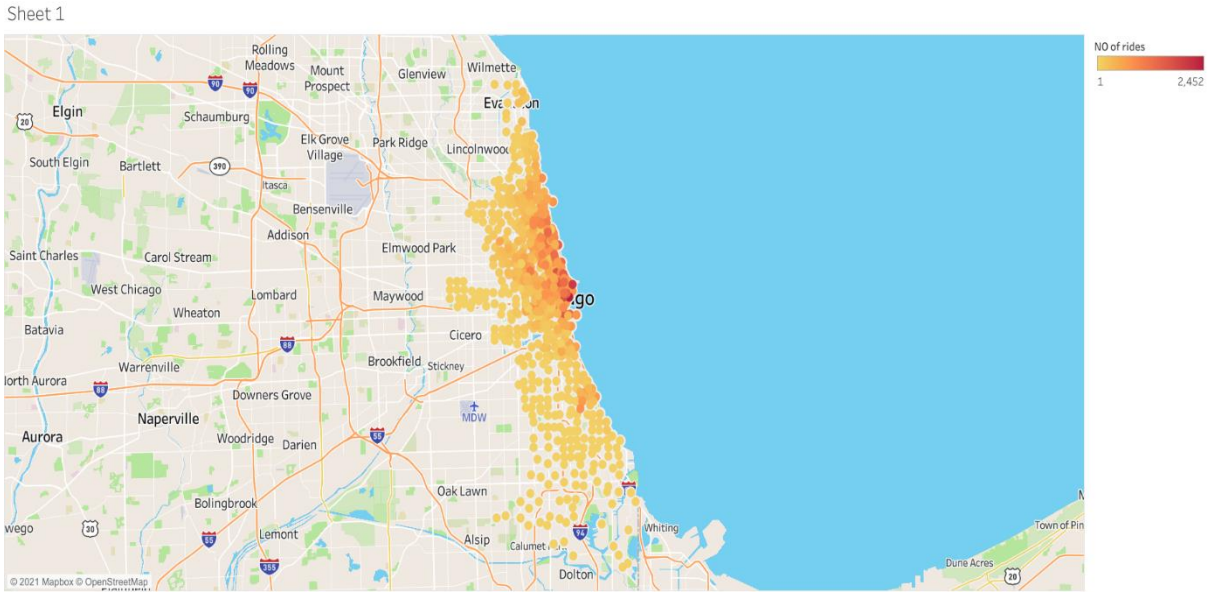
After processing, cleaning the data for each month, a separate excel sheet was made for each month and their key attributes required for the analysis such as No. of rides per month distinguished by the type of user, type of bikes used, average trip duration etc.

Month	Member					Casual					Total		
	Average_ride_time	No_of_trips	docked_bike	electric_bike	classic_bikes	Average_ride_time	No_of_trips	docked_bike	electric_bike	classic_bikes	Average_ride_time	No_of_trips	Temp
Apr-20	00:21:28	61142.00	61142.00	0	0	01:13:05	23621.00	23621.00	0	0	00:35:51	84763.00	49.7
May-20	00:19:34	113185	113185	0	0	00:50:34	86762	86762	0	0	00:33:01	199947	60.6
Jun-20	00:18:31	188020	188020	0	0	00:51:10	154505	154505	0	0	00:33:15	342525	70.6
Jul-20	00:17:30	281690	280719	971	0	00:59:09	268730	268051	679	0	00:37:50	550420	75.4
Aug-20	00:16:45	286492	294965	34143	0	00:44:58	329108	261201	25291	0	00:29:53	615600	73.8
Sep-20	00:15:28	302266	232832	69434	0	00:38:07	230681	171774	58907	0	00:25:16	532947	66.3
Oct-20	00:13:57	243641	156594	87047	0	00:30:09	145003	79883	65120	0	00:19:59	388644	54
Nov-20	00:13:32	171617	105569	66048	0	00:31:45	88096	46013	42083	0	00:19:42	259713	41.3
Dec-20	00:12:45	101137	7801	34041	59295	00:26:51	29994	4939	13736	11319	00:15:58	131131	30.5
Jan-21	00:12:30	78629	1	25273	53355	00:24:28	18094	2105	7753	8236	00:14:44	96723	25.2
Feb-21	00:15:29	39328	0	10174	29154	00:44:27	10072	1271	3165	5636	00:21:23	49400	28.8
Mar-21	00:13:58	144459	0	37445	107014	00:38:10	84030	15657	22847	45526	00:22:52	228489	39

Conclusion of the Analysis

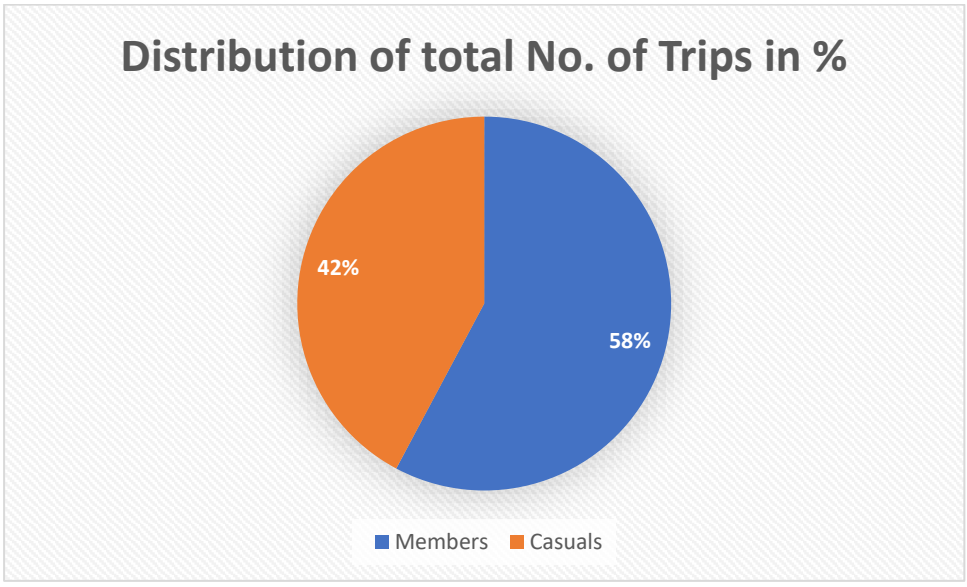
(All the visualizations are created with the help of Tableau)

The bike-share program that features more than 5,800 bicycles and 600 docking stations. Cyclistic sets itself apart by also offering reclining bikes, hand tricycles, and cargo bikes, making bike-share more inclusive to people with disabilities and riders who can't use a standard two-wheeled bike.

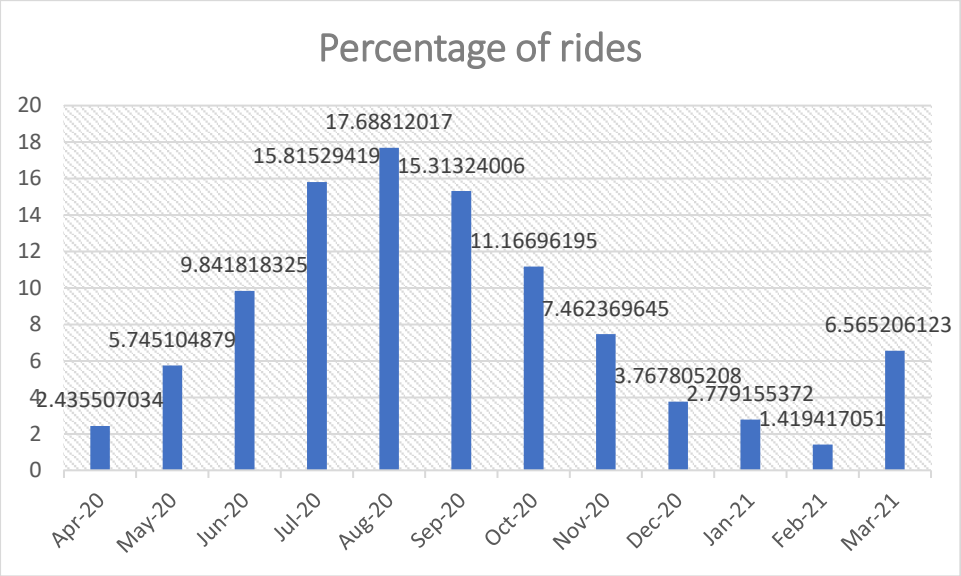


These are all the docking stations spread across the city. Maximum docking stations are mainly in the downtown region and along the Eastern Freeway.

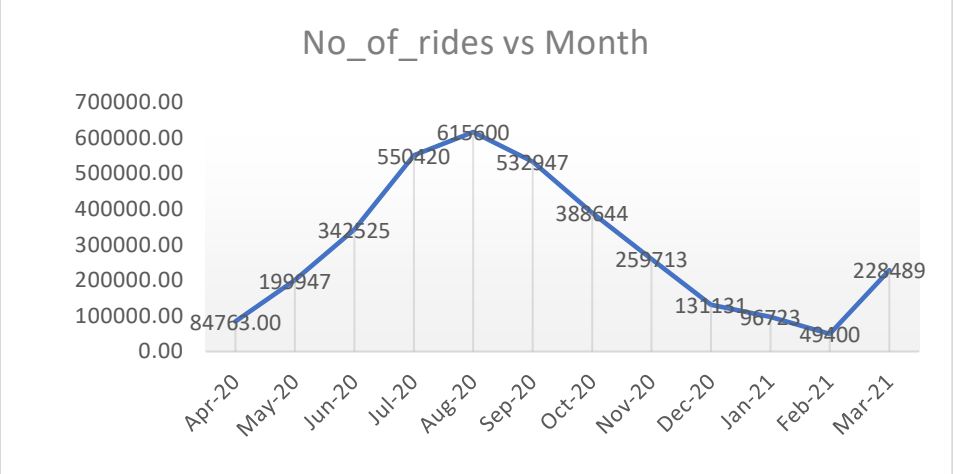
1. The total Trips of the members are 16% more than the casuals. Members have the high share of bike rides.

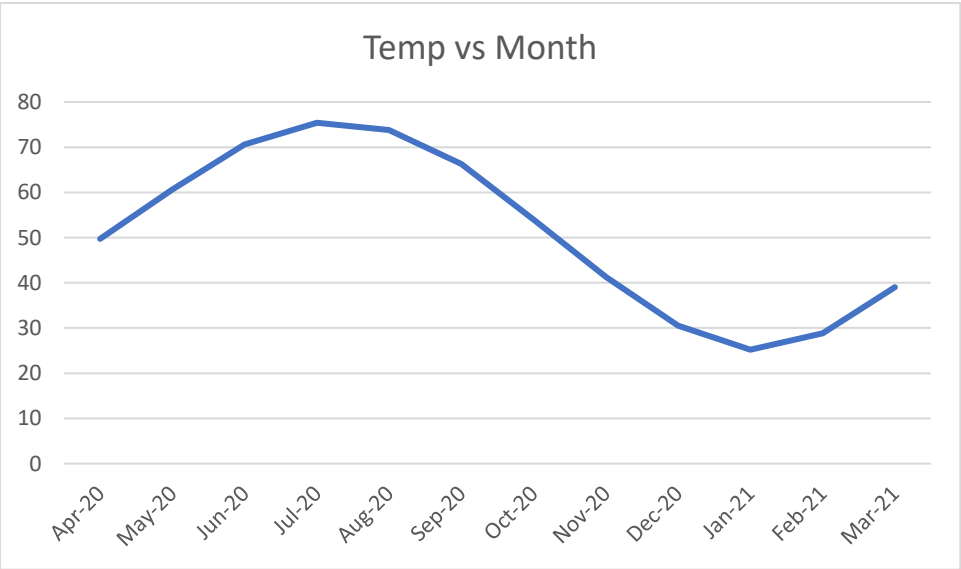


2. From April-20 to Mar-21, August-20 had highest Rides



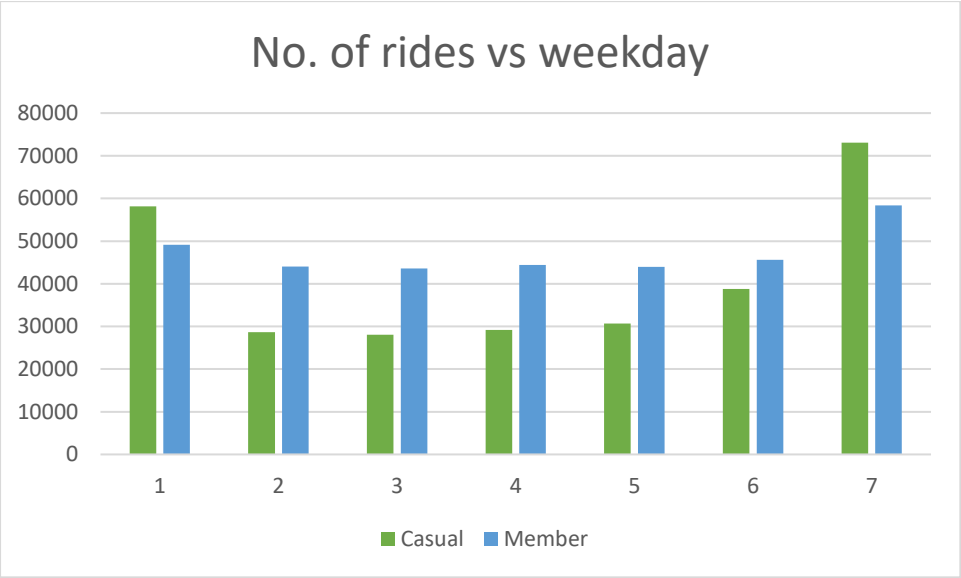
3. The No. of rides has a positive correlation with the temperature of the city





As the temperature rises from April to August, the no. of rides also increases and from thereafter decreases as temp decreases. People like to ride in warmer conditions.

4. No. of rides vs weekday



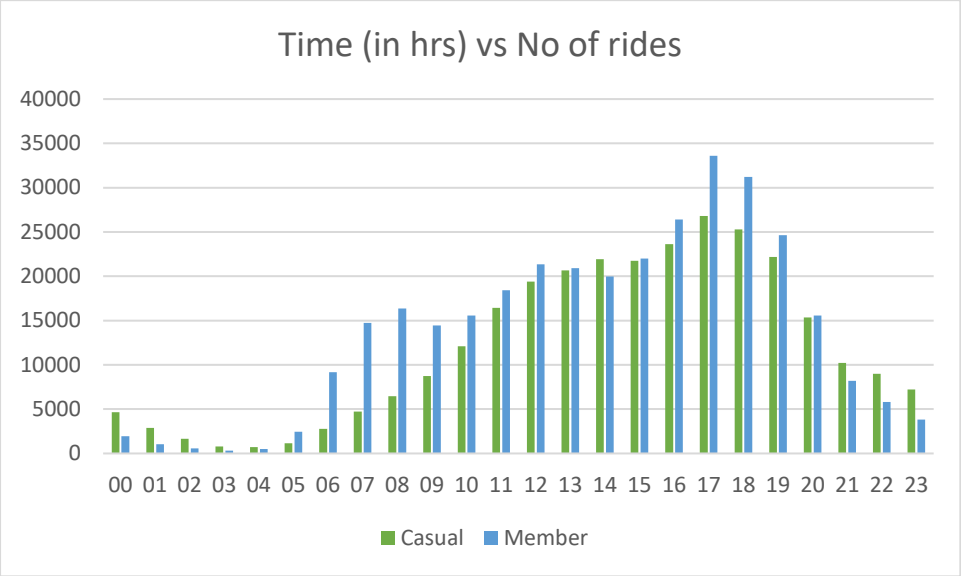
(Here, 1 = Sunday, 2 = Monday, …….)

For Members

No, of rides are evenly distributed which concludes that fixed number of members daily ride the bikes, with peak on Saturday.

For Casual riders, again the no. of rides are even and less than of the members for the weekdays but more on the weekends.

5. Time (in hrs) vs No. of rides



No. of bike rides increases from 5 am but grows exponentially after 6 am due to office tome on the city. There are more member rides in the morning compared to casuals but from 12 pm to 16 pm, casuals have slightly more rides.

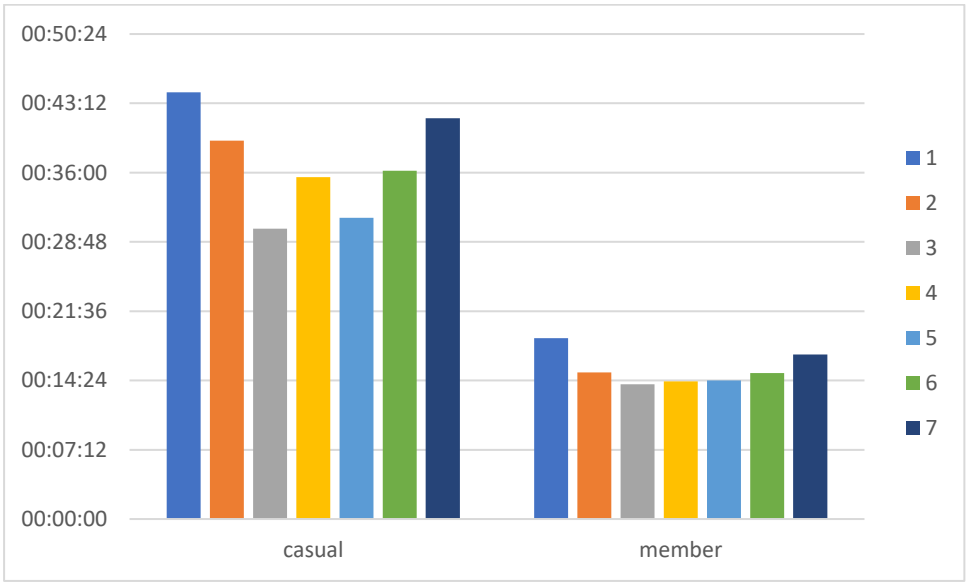
At around 5 pm, the graph is at its peak as it is usually closing times for most of the offices and business in the city.

As the rush hours past, the no. of rides decreases.

6. The average trip duration of casuals is more than the members.

	Average Trip Time(in sec)
Members	00:15:57
Casuals	00:42:44
Total	00:25:49

And, the average time throughout the week is:



(Here, 1 = Sunday, 2 = Monday, …….)

For Members,

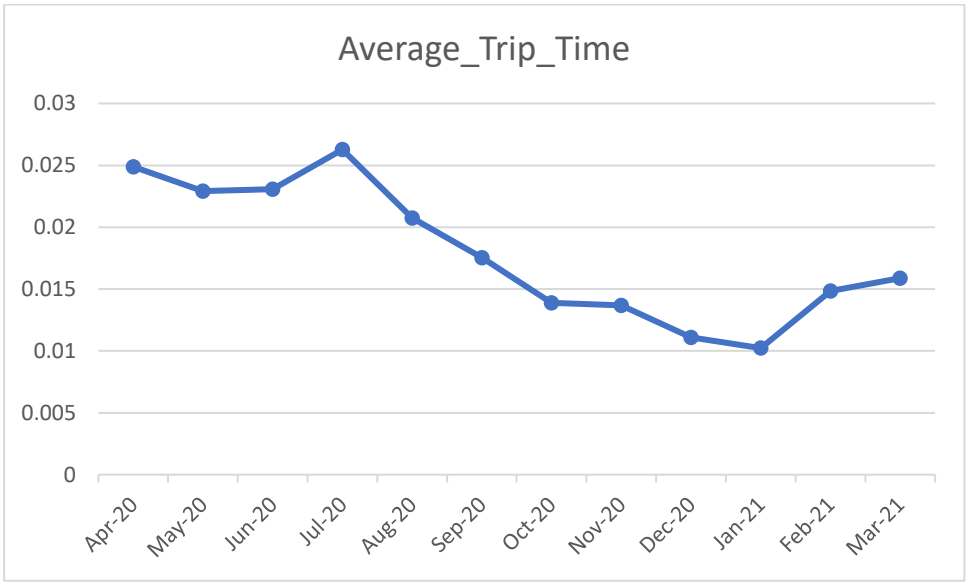
The Trip Duration is similar throughout the week with slight increase on the weekends.

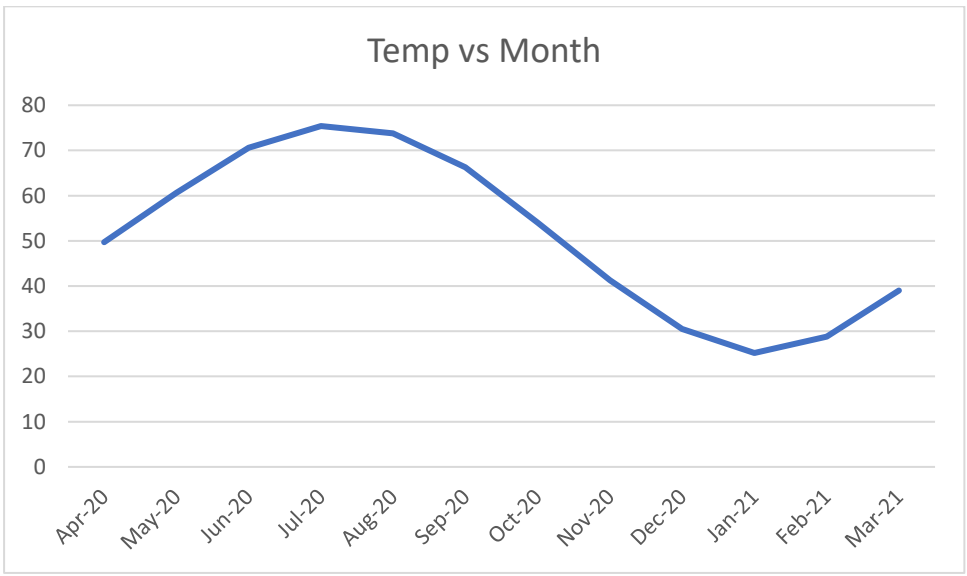
Here, members take the bike for their daily commute to the work and in weekends in the leisure time, the average time goes up slightly.

For Casuals,

The Trip Duration is more uneven throughout the week with peak on the weekends. The main reason for more trip duration can be greater distance of travelling.

7. Average Trip time vs Month

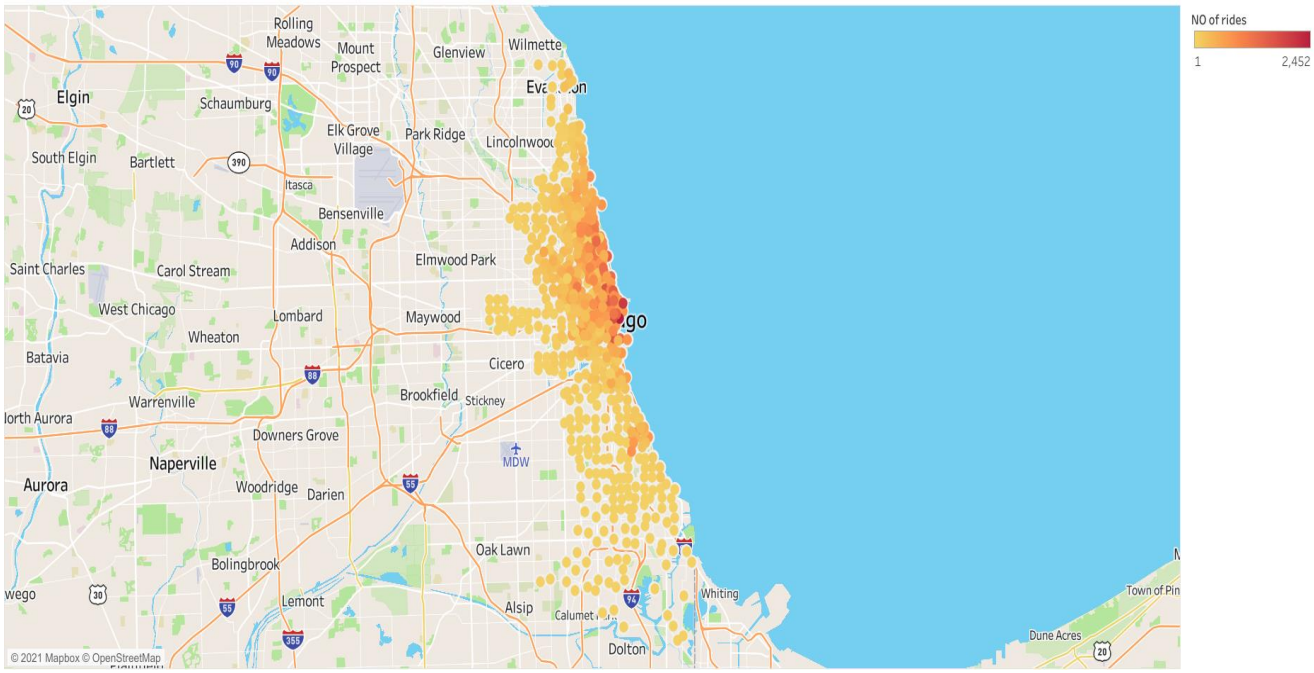




Similar to no. of rides, Average Trip time have positive correlation with temperature of the city.

8. Station vs No. of rides

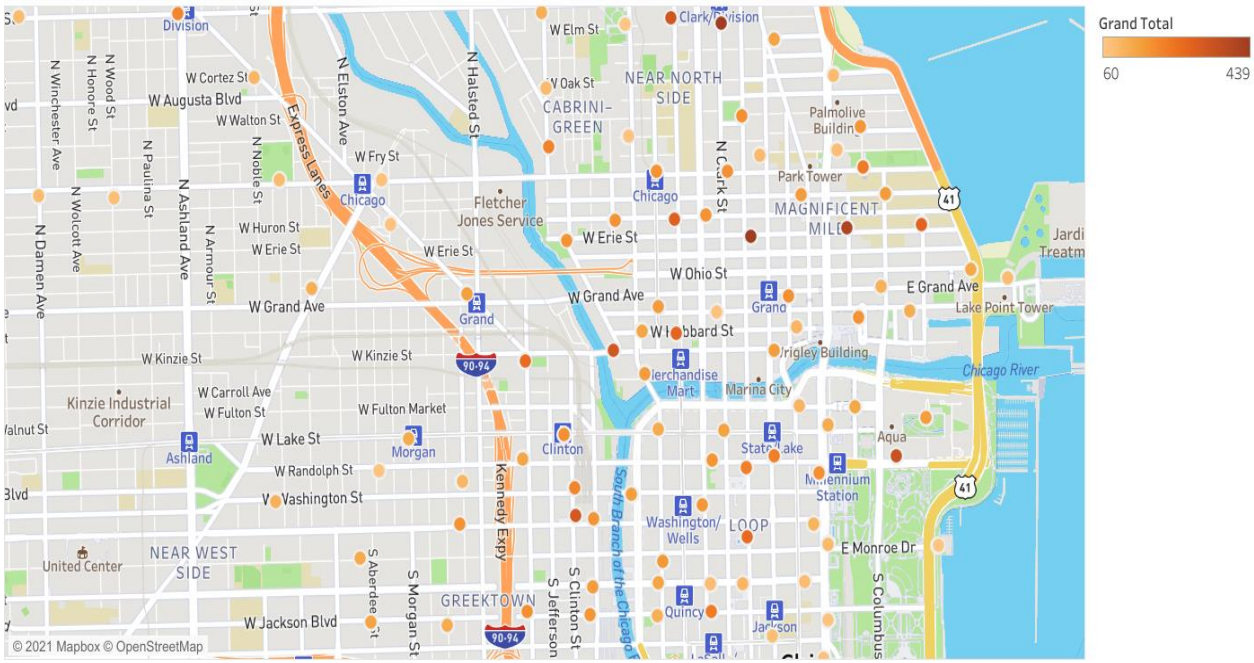
Sheet 1



The most-busiest stations are in the downtown,

1. Clark St & Chicago Ave
2. State St & Pearson St
3. Wells St & Polk St
4. Aberdeen St & Monroe St

Sheet 1

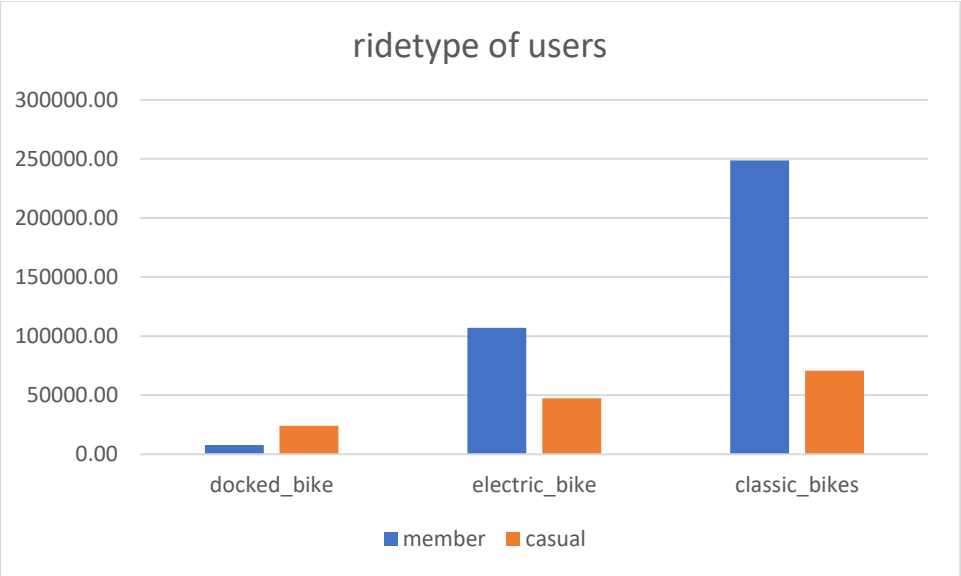


9. Ride-type of users

In Aug-20, Electric bike were introduced

In Dec-2, classic bikes were introduced

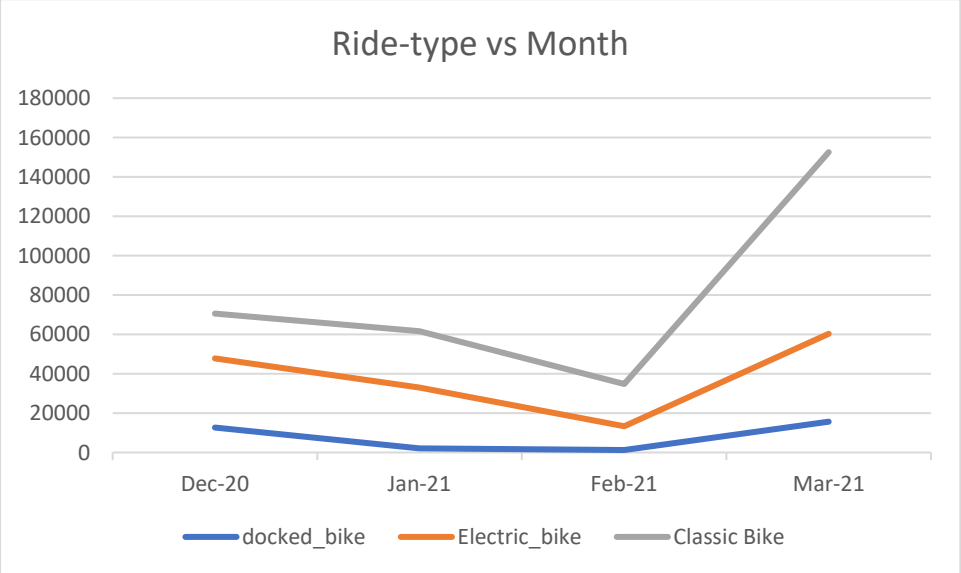
The analysis comprises of data from Dec-21



Classic bike is more preferred option of the users.

Members account for the highest share of the classic bikes, and electric bikes.

Whereas, docked bikes have more share of casual riders.

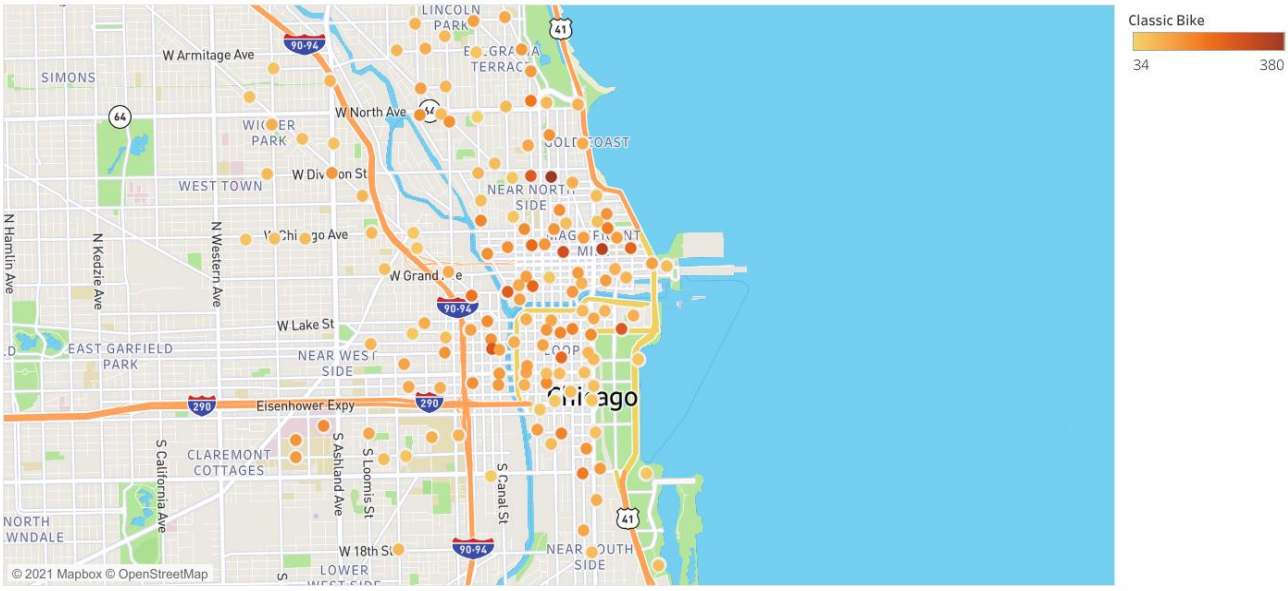


Overall, Classic Bikes have an upward trend among the users.

10. Stations with highest Classic bikes users

1. Clark St & Chicago Ave
2. State St & Pearson St
3. Aberdeen St & Monroe St

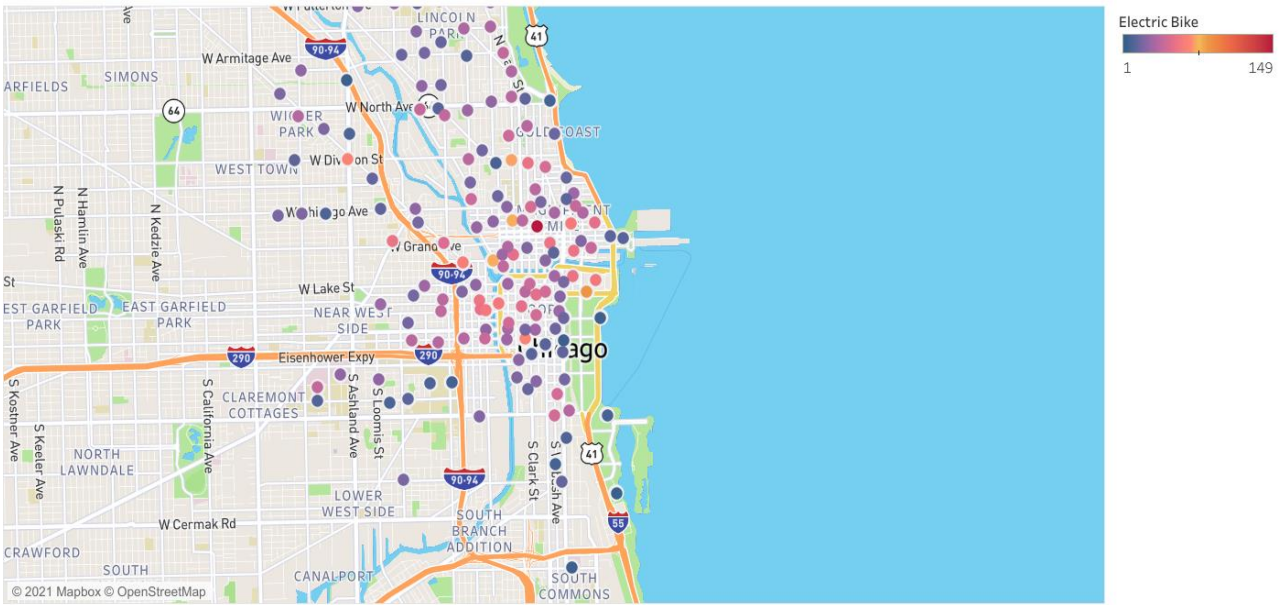
Sheet 1



11. Station with highest electric bike users

1. Clark St & Chicago Ave
2. Wells St & Polk St
3. Aberdeen St & Monroe St

Sheet 1



Act

The act phase would be done by the marketing team of the company. The main takeaway will be the top three recommendations for the marketing.

Conclusion

The Google Analytics Professional Certificate helped me a lot in developing and enhancing my analytics skill, which is very essential in every professional field as you have to deal with data everywhere. I used Excel and MySQL to conduct this analysis, but I would prefer python(numpy, pandas, matplotlib, etc) as it is more efficient with large amount of data than Excel or Google Sheets.

Tools Used : Excel, BigQuery, Tableau

Analysis By-

Siddhesh Jadhav (S.Y.B.Tech)

Mumbai University

siddheshjadhav2001@gmail.com