How does a Bike-Share Navigate Speedy Success?

Report by: Farrukh Nizam Arain
Designation: Data Analyst
Date: 24th February 2022

Introduction:

Cyclistic is a cycle company operating from Chicago, USA. In 2016, it launched a successful bike-sharing offering for its residents. Since that year, it has grown to a fleet of 5,824 bicycles that a geographically tracked and locked into a network of 692 stations across the city of Chicago. The bikes can be unlocked from any of the station and then returned to any other station in the system at any time of the day.

Until now, the primary business strategy of the company was to build general awareness and appeal to the broad consumer segments within the city. One approach which was approved earlier to accomplish this strategy was to offer flexibility of its pricing plans: single-ride passes, full-day passes, and annual memberships.

Customers who purchase single-ride, and full-day passes are classified as casual riders. Customers who purchase annual memberships are termed as Cyclistic members. Additionally, most riders opt for traditional bikes; about 8% of riders use the assistive options. Cyclistic users are more likely to ride for leisure, but about 30% use them to commute to work each day.

The finance department within the company have concluded that annual members bring much more profits to the company, rather than casual riders. Although the present strategy of the company is to attract more riders, Moreno (Director of Marketing & my manager) believes that maximising the annual members will be the key for the company's growth. Rather than targeting all potential customers through the marketing campaigns, Moreno believes that it is a good chance to convert casual riders into annual members. She notes that the casual riders are already aware of the Cyclistic program and have chosen Cyclistic for their mobility needs.

Three questions will guide the future marketing program:

1. How do annual members and casual riders use Cyclistic bikes differently?
2. Why would casual riders buy Cyclistic annual memberships?
3. How can Cyclistic use digital media to influence casual riders to become members?

Problem Statement:

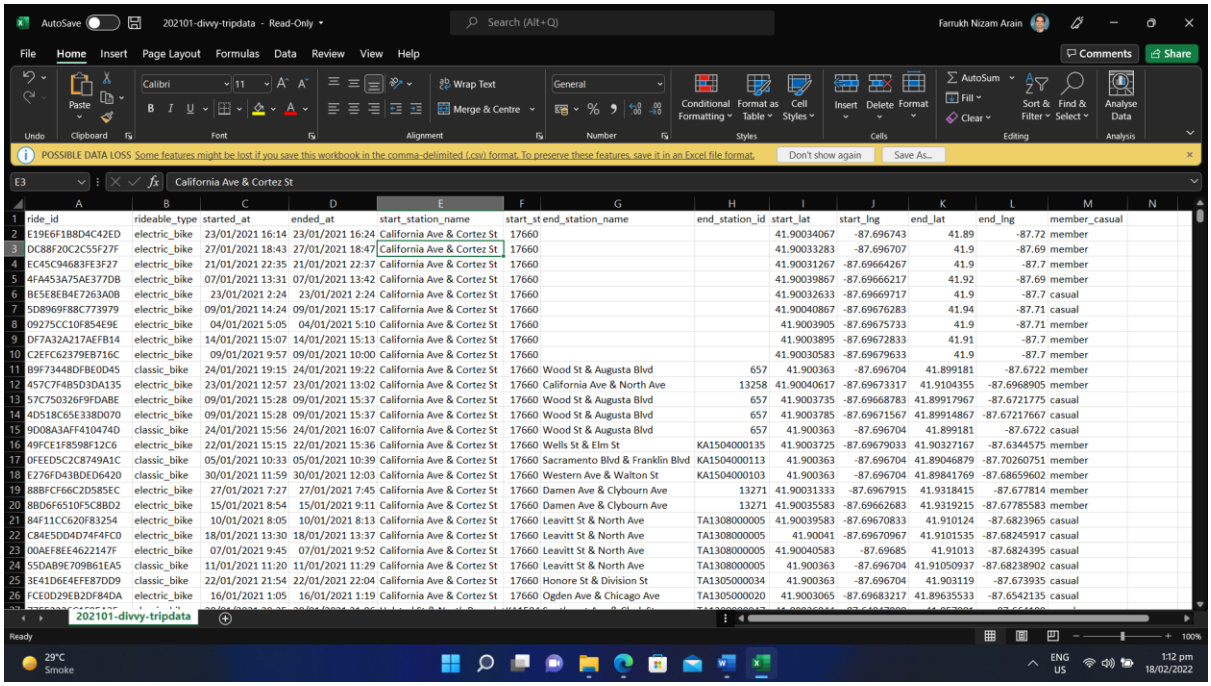
The purpose of my assessment in this project is to find out how do annual members and casual riders use Cyclistic bikes differently? The result of the analysis will be used to design a new marketing strategy to convert casual riders into annual members.

Preparation of the Datasets:

I have downloaded public datasets at the following link provided by Motivate International Inc. under this license. I have analysed and haven't found any issues with bias and credibility with the data through the methodology of ROCCC

The data is organized with the tenures of months and quarters from 2013 till 2020 in zipped format. This data will guide me in differentiating casual annual members through comparison or different variables in the dataset.

Following at the columns (variables) and first ten rows (tuples) in the dataset for a sample:



ride_id	rideable_type	started_at	ended_at	start_station_name	start_st	end_station_name	end_station_id	start_lat	start_lng	end_lat	end_lng	member_casual
E19E6F1B8D4C42ED	electric_bike	23/01/2021 16:14	23/01/2021 16:24	California Ave & Cortez St	17660			41.90034067	-87.696743	41.89	-87.72	member
DC88F20C2C55F27F	electric_bike	27/01/2021 18:43	27/01/2021 18:47	California Ave & Cortez St	17660			41.90033283	-87.696707	41.9	-87.69	member
EC45C94683FE3F27	electric_bike	21/01/2021 22:35	21/01/2021 22:37	California Ave & Cortez St	17660			41.90031267	-87.69664267	41.9	-87.7	member
4FA453A75AE377DB	electric_bike	07/01/2021 13:31	07/01/2021 13:42	California Ave & Cortez St	17660			41.90039867	-87.69666217	41.92	-87.69	member
BE5E8EB4E7263A0B	electric_bike	23/01/2021 2:24	23/01/2021 2:24	California Ave & Cortez St	17660			41.90032633	-87.69669717	41.9	-87.7	casual
5D8969F88C773979	electric_bike	09/01/2021 14:24	09/01/2021 15:17	California Ave & Cortez St	17660			41.90040867	-87.69676283	41.94	-87.71	casual
09275CC10F854E9E	electric_bike	04/01/2021 5:05	04/01/2021 5:10	California Ave & Cortez St	17660			41.9003905	-87.69675733	41.9	-87.71	member
DF7A32A217AEFB14	electric_bike	14/01/2021 15:07	14/01/2021 15:13	California Ave & Cortez St	17660			41.9003895	-87.69672833	41.91	-87.7	member
C2EFC62379EB716C	electric_bike	09/01/2021 9:57	09/01/2021 10:00	California Ave & Cortez St	17660			41.90030583	-87.69679633	41.9	-87.7	member
B9F73448DFBE0D45	classic_bike	24/01/2021 19:15	24/01/2021 19:22	California Ave & Cortez St	17660	Wood St & Augusta Blvd	657	41.900363	-87.696704	41.899181	-87.6722	member
457C7F4B5D3DA135	electric_bike	23/01/2021 12:57	23/01/2021 13:02	California Ave & Cortez St	17660	California Ave & North Ave	13258	41.90040617	-87.69673317	41.9104355	-87.6968905	member
57C750326F9FDABE	electric_bike	09/01/2021 15:28	09/01/2021 15:37	California Ave & Cortez St	17660	Wood St & Augusta Blvd	657	41.9003735	-87.69668783	41.89917967	-87.6721775	casual
4D518C65E338D070	electric_bike	09/01/2021 15:28	09/01/2021 15:37	California Ave & Cortez St	17660	Wood St & Augusta Blvd	657	41.9003785	-87.69671567	41.89914867	-87.67217667	casual
9D08A3AFF410474D	classic_bike	24/01/2021 15:56	24/01/2021 16:07	California Ave & Cortez St	17660	Wood St & Augusta Blvd	657	41.900363	-87.696704	41.899181	-87.6722	casual
49FCE1F8598F12C6	electric_bike	22/01/2021 15:15	22/01/2021 15:36	California Ave & Cortez St	17660	Wells St & Elm St	KA1504000135	41.9003725	-87.69679033	41.90327167	-87.6344575	member
0FEED5C2C87649A1C	classic_bike	05/01/2021 10:33	05/01/2021 10:39	California Ave & Cortez St	17660	Sacramento Blvd & Franklin Blvd	KA1504000113	41.900363	-87.696704	41.89046879	-87.70260751	member
E276FD43BDED6420	classic_bike	30/01/2021 11:59	30/01/2021 12:03	California Ave & Cortez St	17660	Western Ave & Walton St	KA1504000103	41.900363	-87.696704	41.89841769	-87.68659602	member
88BFCF66C2D585EC	electric_bike	27/01/2021 7:27	27/01/2021 7:45	California Ave & Cortez St	17660	Damen Ave & Clybourn Ave	13271	41.90031333	-87.6967915	41.9318415	-87.677814	member
88D6F6510F5C8BD2	electric_bike	15/01/2021 8:54	15/01/2021 9:11	California Ave & Cortez St	17660	Damen Ave & Clybourn Ave	13271	41.90035583	-87.69662683	41.9319215	-87.67785583	member
84F11CC620F83254	electric_bike	10/01/2021 8:05	10/01/2021 8:13	California Ave & Cortez St	17660	Leavitt St & North Ave	TA1308000005	41.90039583	-87.69670833	41.910124	-87.6823965	casual
C84E5DD4D74F4FC0	electric_bike	18/01/2021 13:30	18/01/2021 13:37	California Ave & Cortez St	17660	Leavitt St & North Ave	TA1308000005	41.90041	-87.69670967	41.9101535	-87.68245917	casual
00AEF8EE4622147F	electric_bike	07/01/2021 9:45	07/01/2021 9:52	California Ave & Cortez St	17660	Leavitt St & North Ave	TA1308000005	41.90040583	-87.69685	41.91013	-87.6824395	casual
55DAB9E709B61EA5	classic_bike	11/01/2021 11:20	11/01/2021 11:29	California Ave & Cortez St	17660	Leavitt St & North Ave	TA1308000005	41.900363	-87.696704	41.91050937	-87.68238902	casual
3E41D6E4EFE87DD9	classic_bike	22/01/2021 21:54	22/01/2021 22:04	California Ave & Cortez St	17660	Honore St & Division St	TA1305000034	41.900363	-87.696704	41.903119	-87.673935	casual
FCE0D29EB2DF84DA	electric_bike	16/01/2021 1:05	16/01/2021 1:19	California Ave & Cortez St	17660	Ogden Ave & Chicago Ave	TA1305000020	41.9003065	-87.69683217	41.89635533	-87.6542135	casual

Following are the key schematics of the dataset provided from Motivate International Inc:

- The dataset contains 13 columns in CSV format.
- The names and data types of the columns are:
  - ride_Id (string)
  - Rideable_type (string)
  - started_at (Date and Time)
  - ended_at (Date and Time)
  - start_station_name (string)
  - start_station_id (string)
  - end_station_name (string)
  - end_station_id (string)
  - start_lat (whole number)
  - start_lng (whole number)
  - end_lat (whole number)
  - end_lng (whole number)
  - member_casual (string)

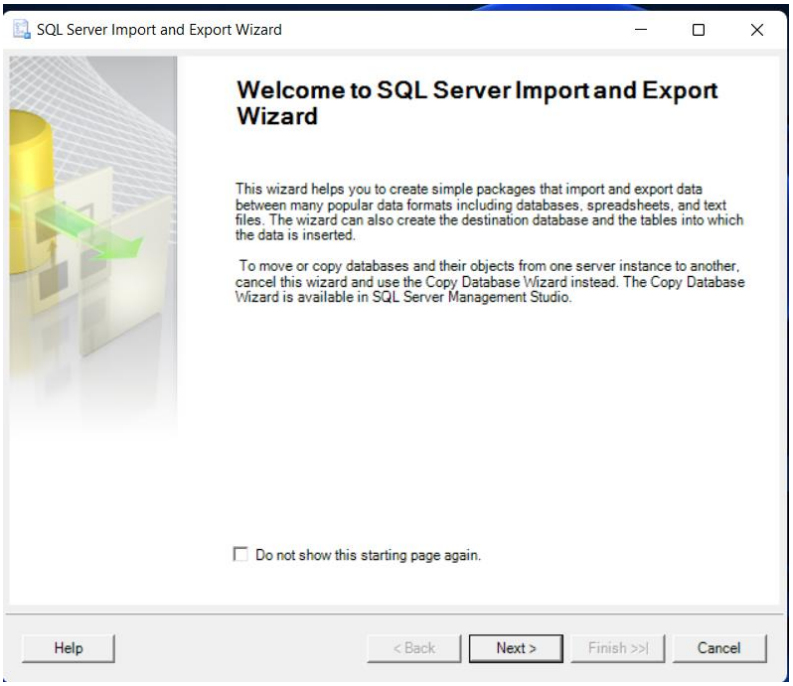
Processing of the Datasets:

The tools that will be used for the data analytical process are:

- Microsoft Excel 365
- Microsoft Power Query
- Microsoft SQL Server 2019
- Tableau Prep
- Tableau Professional

Following are the steps taken for organization and cleaning of the dataset which was provided:

- Two folders were created for the 2021's data. One is named as 'zipped' and the other is 'unzipped'. The 'zipped' folder is used as a backup of the dataset, while the 'unzipped' is used for the actual working for analysis.
- The filenames of the dataset are converted from CSV to XLS to be a bit more compatible with the tools on which I will be working on. Additionally, it will help me to familiarize with the dataset and perform some cleaning activities with the Microsoft Excel's intuitive GUI as well.
- In Microsoft Excel, a column was created called "ride_length". The length of each ride is calculated by subtracting the column "started_at" from the column "ended_at" (for example, =D2-C2) and formatting it as HH:MM: SS using Format > Cells > Time > 37:30:55.
- Another column was created, and it was called as "day_of_week". It calculated the day of the week that each ride started using the "WEEKDAY" function (for example, =WEEKDAY(C2,1)) in each row. Additionally, the column was replaced to a string value of the day of the week, noting that 1 = Sunday and 7 = Saturday.
- All the datasets were imported using inbuilt 'SQL Server Import and Export Wizard'



- After importing, I found out that the data types of the monthly data tables were different from each other. To resolve this, I corrected the data types in all the tables.

Column Name	Data Type	Allow Nulls
ride_id	nvarchar(255)	☑
rideable_type	nvarchar(255)	☑
started_at	datetime	☑
ended_at	datetime	☑
start_station_name	nvarchar(255)	☑
start_station_id	nvarchar(255)	☑
end_station_name	nvarchar(255)	☑
end_station_id	nvarchar(255)	☑
start_lat	float	☑
start_lng	float	☑
end_lat	float	☑
end_lng	float	☑
member_casual	nvarchar(255)	☑
ride_length	datetime	☑
day_of_week	int	☑
		☐

- In the next step, I used SQL queries to not only create a NEW table having structure of the existing tables, but also, I append (combined) all the monthly data tables into this new table which I called as ‘Cyclistic_M_2021’.
- For changing the column names, I used the same option which I used earlier for changing the data types of columns. Additionally, I created two new columns; ‘Country’ and ‘City’ and added relevant data for better spatial analysis in Tableau further in the project.

Column Name	Data Type	Allow Nulls
ride_id	nvarchar(255)	☑
rideable_type	nvarchar(255)	☑
started_at	datetime	☑
ended_at	datetime	☑
start_station_name	nvarchar(255)	☑
start_station_id	nvarchar(255)	☑
end_station_name	nvarchar(255)	☑
end_station_id	nvarchar(255)	☑
start_latitude	float	☑
start_longitude	float	☑
end_latitude	float	☑
end_longitude	float	☑
member_or_casual	nvarchar(255)	☑
ride_length	time(7)	☑
day_of_week	varchar(255)	☑
Country	varchar(255)	☑
City	varchar(255)	☑
		☐

- For removing the NULLS in the observations, appropriate SQL queries were executed to delete them safely to have a much more clean and accurate analysis in the next step.
- Before visualization and creation of dashboards, quick analysis was carried out to know a bit more about the dataset and how it can further help with the targeted marketing efforts. Following questions were answered using SQL query language:

    - What are the number of casual and member rides throughout the year?
    - Maximum and Minimum length of rides through our services?
    - Number of rides throughout the week in the previous year?
    - Count of type of bikes used by both members and casual riders throughout the year and the week?
    - What are the top start and end stations used throughout the week?

- In the next step, the SQL Server's table was connected via Tableau Prep for making it presentable before the actual process of visualization.

Analysing & Visualization of the Datasets:

In Tableau Desktop, a series of visualizations were created to answer all the queries pertaining to the previous year, that were asked at the start of the projects and were critically needed before the initialization of next marketing drive that the company is planning:

- Is there a route through which the overall progress can be viewed throughout the previous year?
- What are the top ten start and end stations which our riders preferred?
- What were the peak hours during which casual and members used our bike sharing service?
- During all the months, how many times both the type of riders availed our service?
- What were the average usage (in minutes) of rides by our customers?
- Among the type of bikes, which ones were most of the times opted by our riders?
- During the seven days of the week, which ones were the busiest for us?
- What were the average times for the usage of all types of our bikes?
- Is there a way to interconnect all the findings and trends?

Sharing the Findings Found in the Datasets:

After a thorough process of deep analyzing the data, following findings and trends were found:

- The overall business of Cyclistic Inc. skyrocketed during the tenure during the period of Summer till Autumn, but it did not help much because the only the number of casual riders were increased.
- Top 10 busiest start stations for casual and member riders were:

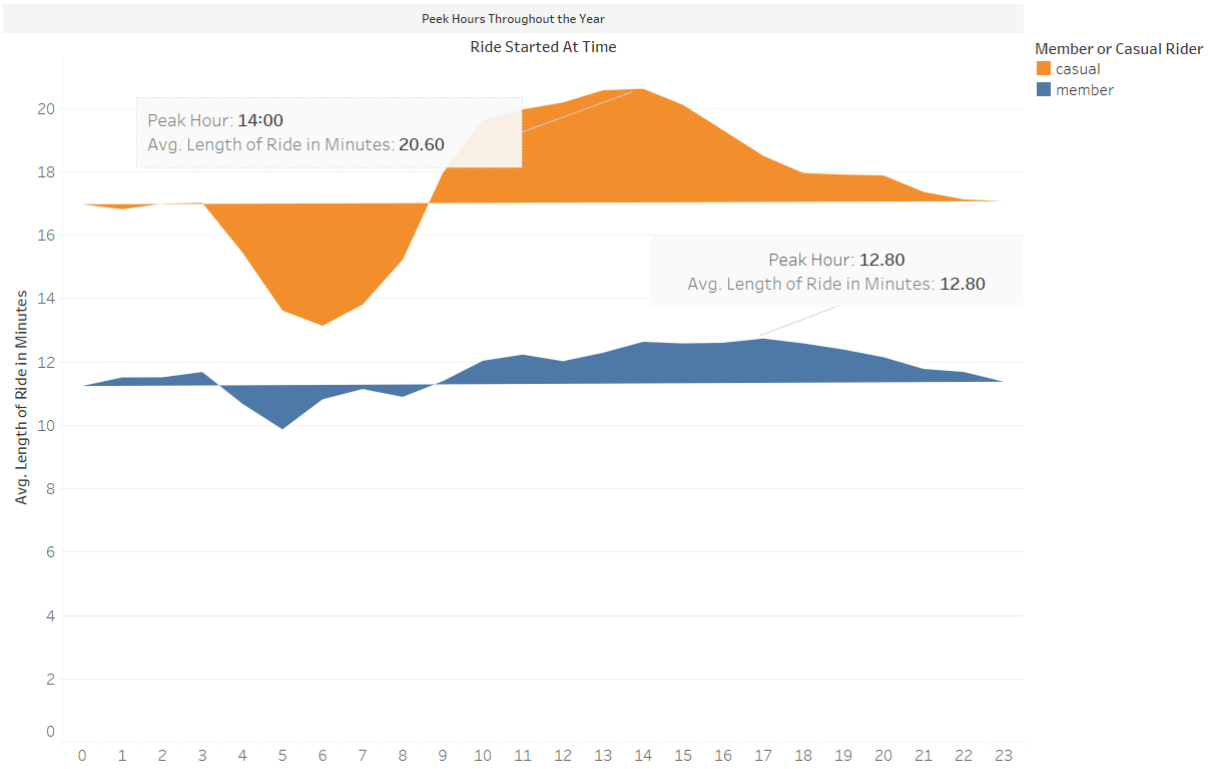
Casual	Numbers	Member	Numbers
1. Streeter Dr & Grand Ave	57,928	1. Dearborn St & Erie St	16,489
2. Millennium Park	29,199	2. St. Clair St & Erie St	16,207
3. Michigan Ave & Oak St	25,545	3. Clark St & Elm St	15,550
4. Clark St & Lincoln Ave	14,020	4. Wells St & Concord Ln	15,330
5. Wells St & Concord Ln	13,611	5. Broadway & Barry Ave	14,986
6. Clark St & Armitage Ave	13,332	6. Clark St & Armitage Ave	14,065
7. Clark St & Elm St	10,881	7. Streeter Dr & Grand Ave	13,879
8. Broadway & Barry Ave	10,574	8. Clark St & Lincoln Ave	13,646
9. Dearborn St & Erie St	10,320	9. Michigan Ave & Oak St	12,177

10. St. Clair St & Erie St	9,026	10. Millennium Park	7,383

- Top 10 Ending Stations for the casual and members are:

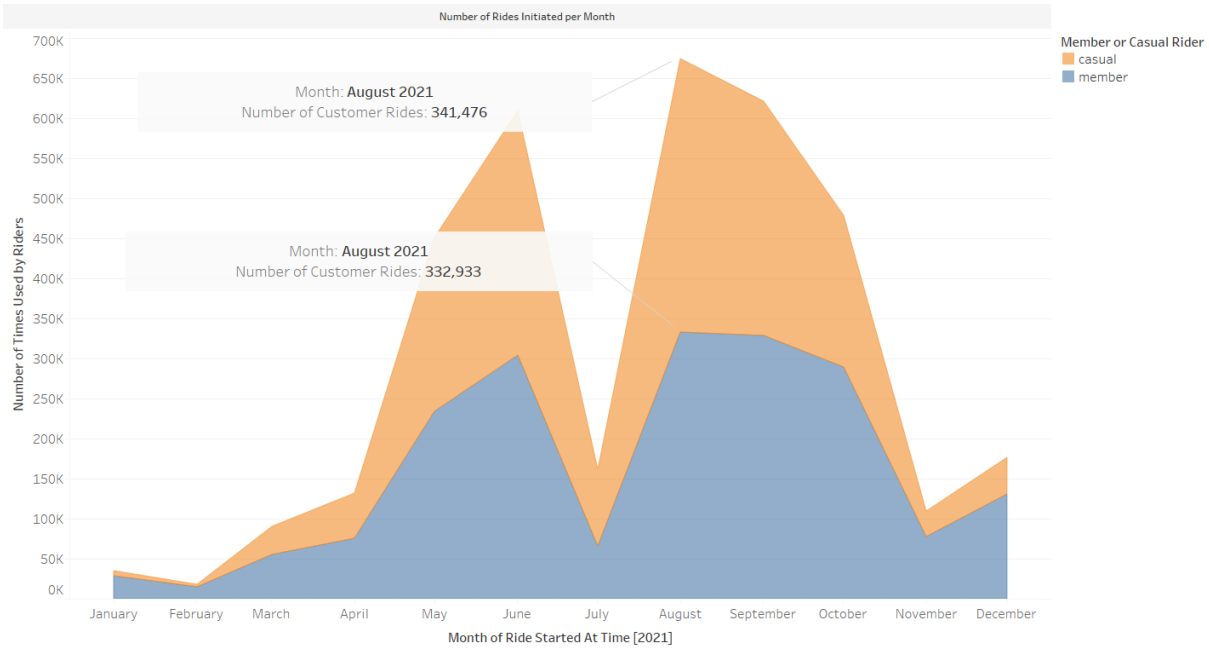
Casual	Numbers	Member	Numbers
1. Streeter Dr & Grand Ave	60,592	1. Clark St & Elm St	16,656
2. Millennium Park	30,513	2. Wells St & Concord Ln	16,321
3. Michigan Ave & Oak St	27,053	3. Dearborn St & Erie St	15,610
4. Theater on the Lake	15,618	4. Wells St & Elm St	14,484
5. Clark St & Lincoln Ave	14,130	5. Clark St & Armitage Ave	13,131
6. Wells St & Concord Ln	13,475	6. Clark St & Lincoln Ave	12,847
7. Clark St & Armitage Ave	13,169	7. Streeter Dr & Grand Ave	12,285
8. Wells St & Elm St	10,768	8. Michigan Ave & Oak St	11,180
9. Clark St & Elm St	10,595	9. Theater on the Lake	9,852
10. Dearborn St & Erie St	10,012	10. Millennium Park	6,797

- As far as the peek hours in between the customers are concerned, the peek hours for the casual members were in between **09:00 and 19:00,** with the highest average length of ride in minutes was **20.60 minutes**. As for the members of Cyclitic, the peak hours were between **12:00 and 21:00,** with the highest average length of ride in minutes was **12:80.**



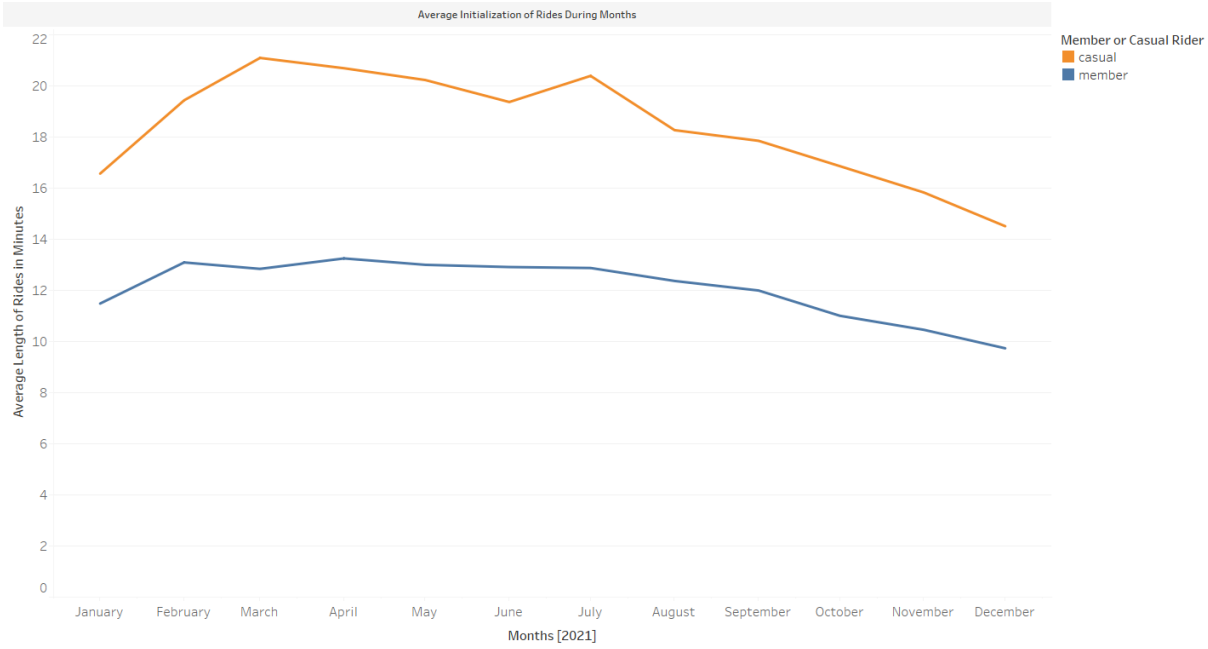
Average of Length of Ride in Minutes for each Ride Started At Time Hour. Color shows details about Member or Casual Rider. The data is filtered on Ride Started At Time Day, which ranges from January 1, 2021 to December 31, 2021. The view is filtered on Member or Casual Rider, which keeps casual and member.

- If we talk about the most rides availed throughout the last year, interestingly August was the peak month for both the type of riders. If we talk about casual and member riders, **341,476** and **332,933** rides were availed respectively.



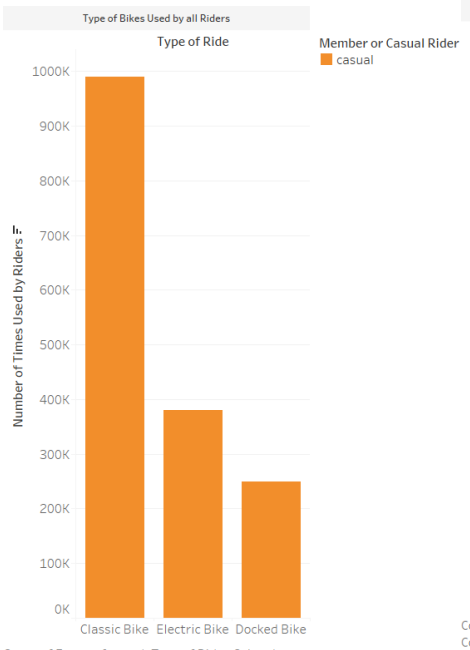
The plot of count of Extract for Ride Started At Time Month. Color shows details about Member or Casual Rider. The data is filtered on Ride Started At Time Day, which ranges from January 1, 2021 to December 31, 2021. The view is filtered on Member or Casual Rider, which keeps casual and member.

- A clear difference was noticed when we analyzed the average length of rides from the riders. The highest peak for the members (13.26 minutes) was just **62.84 percent** of the peak for the casual riders (21.10).

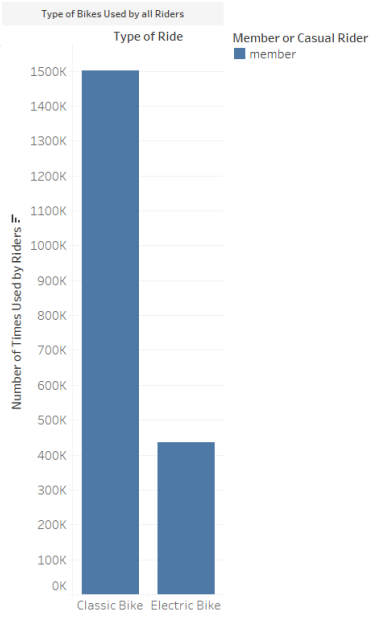


The trend of average of Length of Ride in Minutes for Ride Started At Time Month. Color shows details about Member or Casual Rider. The data is filtered on Ride Started At Time Day, which ranges from January 1, 2021 to December 31, 2021. The view is filtered on Member or Casual Rider, which keeps casual and member.

- If we talk about the type of rides preferred by casual and member riders, both of them preferred classic bikes, rather than the modern electric ones.



Count of Extract for each Type of Ride. Color shows details about Member or Casual Rider. The data is filtered on Ride Started At Time Day, which ranges from January 1, 2021 to December 31, 2021. The view is filtered on Member or Casual Rider, which keeps casual.



Count of Extract for each Type of Ride. Color shows details about Member or Casual Rider. The data is filtered on Ride Started At Time Day, which ranges from January 1, 2021 to December 31, 2021. The view is filtered on Member or Casual Rider, which keeps member.

- Usage statistics in a week differed between the casual and member riders significantly. Weekend was the preferred choice for the casual riders, while working / business days were preferred by the member riders throughout the year.

Top Recommendations Via the Analysis:

Following are the recommendation which will help the marketing department after analysis of the historical data provided:

- Marketing team needs to initiate a marketing campaign at the top 10 start and end stations to help them convert casual to member riders.
- As per the analysis of the dataset, the marketing campaign should be in between the working hours (08:00 to 18:00) for generation of new subscribers and to have a much more cost-effective and targeted use of resources.
- As per the August monthly data, almost **50 percent** of the total customers were casual members. So, it will be much more feasible to start the marketing campaign **from July till November** because during these months the number of users is higher than the rest of the months.
- The gap of average length of time for using our bike sharing service is huge in between the casual and member riders. So, some discount packages should be for the casual members during the peak times.
- Almost **7 percent** of the bikes were docked by the casual members, leading to wastage of operational costs. A time limit with a penalty should be introduced to cut down the number of docked bikes to **less than 2 percent.**
- As we know that casual members majorly use our bike sharing program during the weekends (Saturday and Sunday), so if the marketing department is short on budget or resources, then marketing campaign can also be run during those days as well.

Conclusion of the Analysis:

It was a great learning experience for me during this professional project, as it helped me not only to learn new things, but also to find interesting patterns and findings to help the stakeholders for planning the strategies.