



Analysing the Data of a Bike sharing company

Capstone project

Author: ANES BOURARA

Institute: Google Data Analytics Certificate

Contents

1	INTRODUCTION	1
2	DATA COLLECTION	1
3	ABOUT THE DATA	1
4	DATA PROCESSING	2
5	RESULTS	2
6	Conclusion	7

1 INTRODUCTION

This data analysis is performed as a part of the last course of the *Google Data Analytics* Certificate, our case study will focus on a fictional bike sharing company named Cyclistic.

Cyclistic was created in 2016, and its a bike-share company, since 2016, their offer has grown to over 5824 bicycles that are tracked and locked in 692 stations located all over the city of Chicago. A full trip is calculated every time a bike is unlocked from one station and returned to any other station in the system.

Until now, the companys marketing strategy relied on building awareness and appealing to broad consumer segments. One strategy that made this possible was proposing flexible pricing plans : single rides passes, full-day passes and yearly memberships. Customers who purchase annual memberships are Cyclistic members, and riders that use single-rides and full-day passes are referred to as casual riders.

Lily Moreno, the finance analyst of Cyclistic believes that maximizing the number of annual memberships is the key to future growth. And rather than targeting all new costumers, she thinks that there is a very good chance to convert casual riders into members.

The goal of this analysis will be to find solutions and design marketing strategies aimed at converting casual rider into annual members, To achieve this goal, we are going to try to find answers to three key questions :

How do annual members and casual riders use Cyclistic bikes differently ?

Why would casual riders buy Cyclistic annual memberships ?

How can Cyclistic use digital media to influence casual riders to become members ?

2 DATA COLLECTION

For this study we are going to use the historical trip data, the datasets have been made available by Motivate International Inc. (under this license)

The data includes the history of trips since the creation of the company. For our study we chose to use the data of the entire last year (March-2021 To February 2022). ¹

This is public data that we can use to explore how different customer types are using Cyclistic bikes.

3 ABOUT THE DATA

The data is organized in a csv files format and can be retrieved via this link², the data was made available every four months for the first years, but for the last years it became a monthly data.

¹<https://ride.divvybikes.com/data-license-agreement>

²<https://divvy-tripdata.s3.amazonaws.com/index.html>

The data respects the ROCCC principle since it is a new first party data that includes information about all the trips completed with Cyclistic bikes.

To check the integrity of the data we checked all the trips had a start date that corresponded to the period during which the data were recorded, and we checked that the trips were indeed made by members (casual or member) from Cyclistic.

We have chosen to use data from the last year because it is the most recent data, and because we believe that one year is a very sufficient period to cover all types of use and avoid bias.

Also, to verify the data integrity we calculated the duration of all the trips and we found that a lot of trips started and ended at the same time, we also found some trips had an ending time that preceded the starting time. Another problem we had with the trips duration was that some trips had a too long duration which exceeded a day and which could reach a few months in certain cases.

Since it was impossible to verify if these problems were due to a system bug or human error, or even if it was possible to keep the bike in use during all this period, we decided to drop these trips and use only trips that have a duration between 1s and 1 day. We believe that this will not affect the results of our study, especially since these trips represent less than 0.2% of the total number of trips.

4 DATA PROCESSING

Since it is a training case study, we chose to use two different tools for the data processing, we used Microsoft excel and BigQuery.

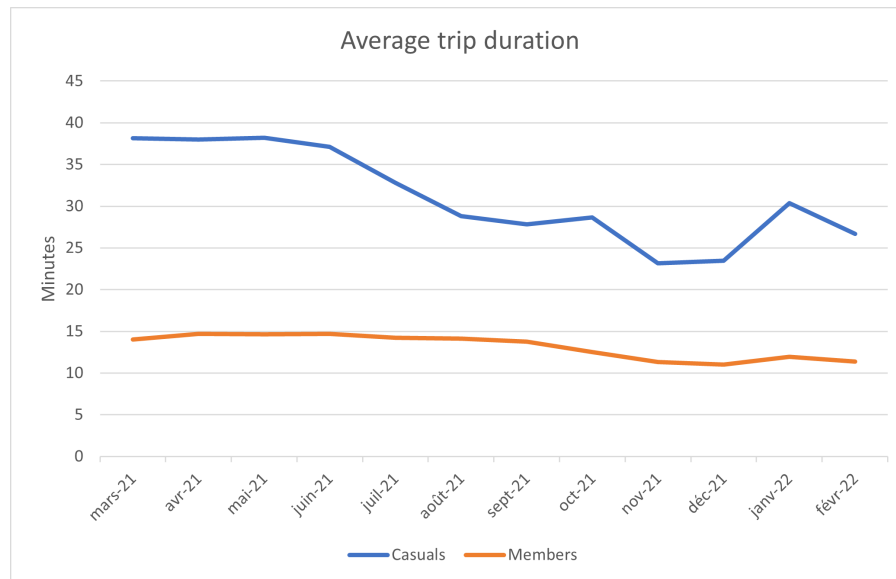
For our analysis we added two new columns, the first one contained all the trip durations by using the formula (ended_at - started_at), and we also retrieved the day the trip started by using the function weekdays (), and finally we used the functions count and count if to retrieve the details of trips made by members and casuals.

5 RESULTS

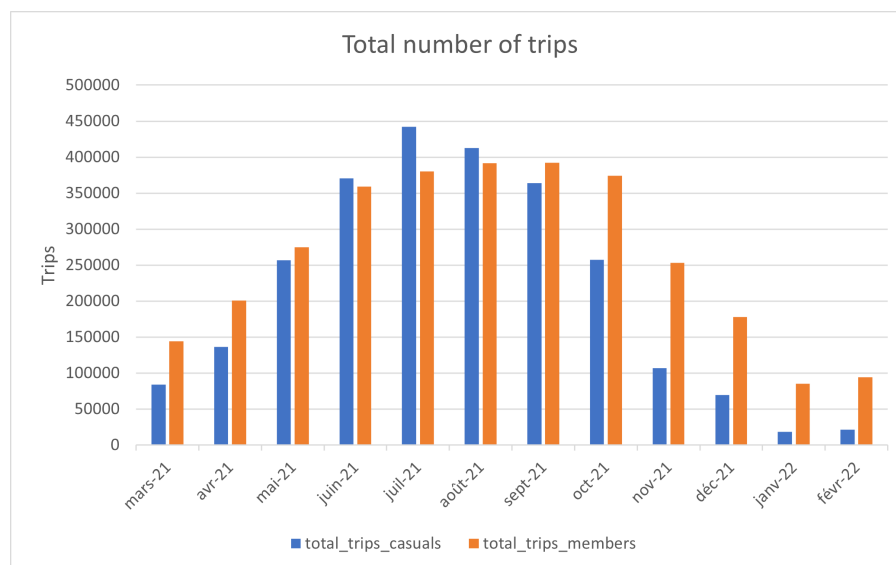
The results of our data analysis will be illustrated and explained with the following series of graphs :

5.1 Average trip duration

The graph of the average trip durations clearly shows us that the trip durations of casuals are significantly higher than those of members. We also note that the trip durations of members are almost stable around 14min throughout the year, while the trip durations of casuals vary from one period to another with a very long trips in the months of (March-April May) and shorter trips during the winter period.

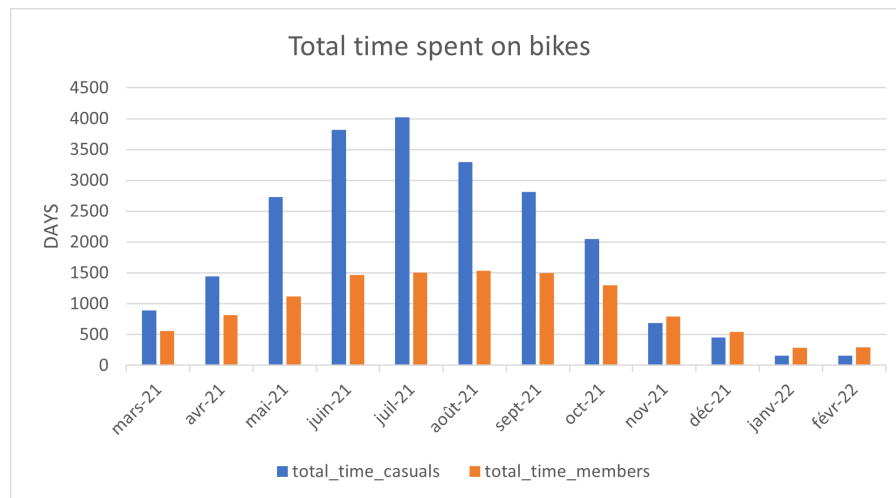


5.2 Total number of trips



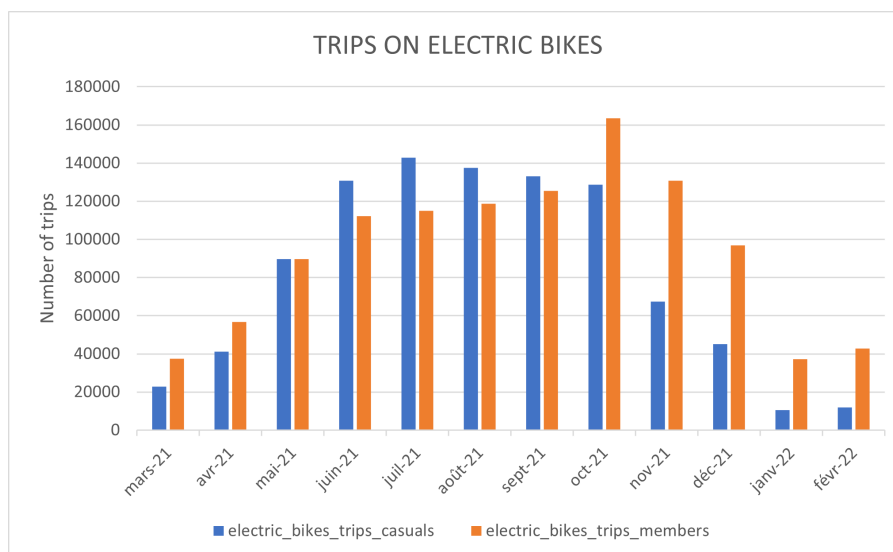
This graph shows us that the total number of trips made by casuals and members is pretty close throughout the year except for the period (april-octobre), at the same time we notice that during the rest of the year members number of rides is much higher and the casual members trips become rarer.

5.3 Total time spent on bikes

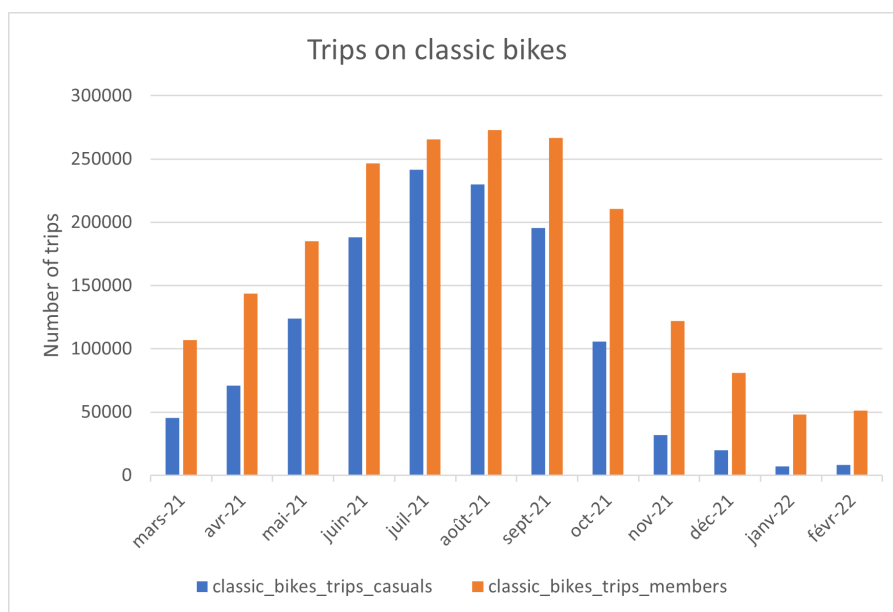


This graph shows that the total time spent by casual is significantly higher for casual members, the difference is so high during the summer period where casual members trips can be 2 times higher than yearly members trips.

5.4 Total number of trips on electric and classic bikes



In this graph we notice that the number of trips on electric bikes is slightly higher for casual members compared to annual members. During the periods where casual members trips become rarer electric bikes are used more often by annual members.



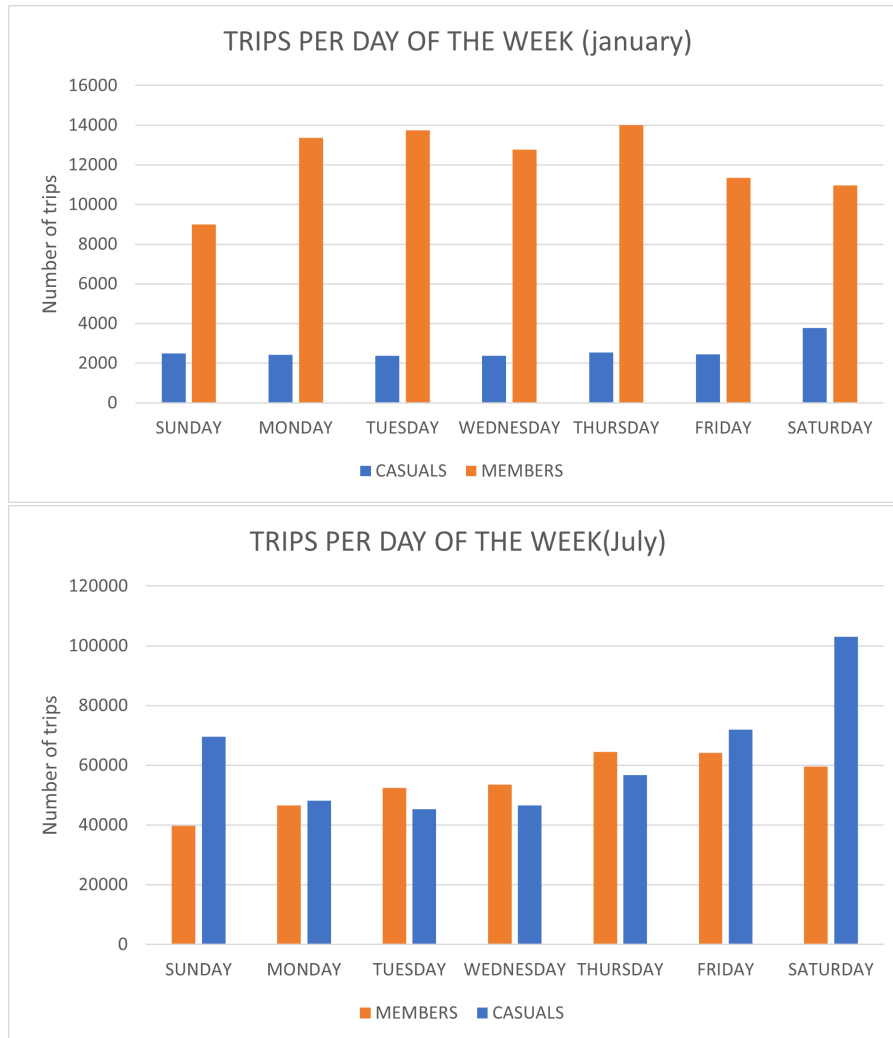
This second graph about classic bikes trips, shows that annual members use classic bikes more often throughout the year.

If we combine what we saw in the two graphs we can assume that a large part of annual members prefer using classic bikes, But in the same time electric bikes also remain very used by this type of members.

On the other hand, the usage of casual members remains divided between classic and electric bikes, but we can conclude that casual members prefer electric bikes more than annual members who prefer classic bikes.

5.5 Trips per day of the week

In this part we have chosen two months to represent the weekly trends (January from the winter period and July from the summer period) , since the trends are pretty much the same during both periods (winter and summer) it would be useless to represent the 12 months.



In these two graphs we can see that the use of Cyclistic bikes by annual members is almost the same throughout the week with a slight diminution during the weekends. The casual members instead, have an almost stable number of rides during the winter, but in the summer the number of rides is significantly higher during the weekends.

6 Conclusion

6.1 Summarizing the results

Even if the total number of trips done by casual and annual members is very close, The average trip duration and the total time spent by riders on bike is significantly higher for casual users. Classic bikes are used more often by annual members, while the use of electric bikes is a little higher among casuals.

The use of Cyclistic bikes by members is often for daily tasks and therefore is the same during all days of the week, while casual members use the bikes for daily tasks during the winter but a large part of this type of members uses bikes during the summer weekends.

* Why would casual members by annual memberships ?

- 1) Since casuals spend more time on bikes than annual members, we could change the pricing method to encourage casuals to take out an annual subscription. an example would be to set up a pricing that increases in relation to the time spent on the bike for casuals (second hour more expensive than the first one...) this will make an annual membership a better option for them.
- 2) Considering the fact that a large part of casual users uses Cyclistic bikes only during the weekends, We could add a new membership package for weekend usage only.
- 3) Another idea would be to make promotions on annual subscriptions during the winter period during which the casuals are very rare to use the Cyclistic bikes.

* How can Cyclistic use digital media to influence casual riders to become members ?

- 1) To influence casual riders, we could send emails to them with their time spent on bike and explain that taking annual memberships would not necessarily mean paying more, and that they could potentially save money.
- 2) We could also send notifications through a Cyclistic app to inform about promotions.