Cyclistic Data Analysis Report by Francis Bautista

Cyclistic is a 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. The majority of 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 Director of marketing believes that the company's success depends on maximizing the number of annual memberships, therefore understanding how casual riders and annual members use Cyclistic differently must be performed. To be approved with the proposed recommendations, these recommendations will have to be backed with compelling data insights and professional data visualisations.

Cyclistic's finance analysts have concluded that annual members are much more profitable than casual riders. Although the pricing flexibility helps Cyclistic attract more customers, it's believed that maximising annual members will be key to future growth. Rather than basing the marketing campaign towards targeting new customers, there's a good chance that shifting casual members to annual members will be most optimal.

Case Study

The below will focus on one simple business question, "What will be the most effective marketing strategy to shift casual members to annual subscription-based members?"

We can analyse this through:

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

Stakeholders

- Cyclistic Users
- Director of Marketing
- Marketing Team
- Executive Team

Data Source

Monthly user data has been made available where each data set is in csv format where all rides are logged within. The data has been made publicly available via a license by Motivate International Inc. and the city of Chicago. It's further noted that all user details have been removed for privacy purposes. Data for the financial year ended 30 June 2022 is used for this analysis.

Cleaning and Preparation of Data

It was noted that each monthly data set ranged from 18MB to 190MB in which aggregately would've totalled to approved 1.2GB worth of data. Therefore, manually merging the data into one excel file wouldn't have been an option. The method of combining and cleansing the data using RStudio was used(noted: whilst combining data is possible using SQL, using R to merge datasets I found is to be more time efficient where less lines of functions are to be written)

Please refer to:

https://github.com/francisbautista15/Cyclistic-Project/blob/main/Data%20Cleansing.R

Once a combined CSV file was created, I then proceeded with importing the data into SQL Server management studio to run queries over the data for analysis
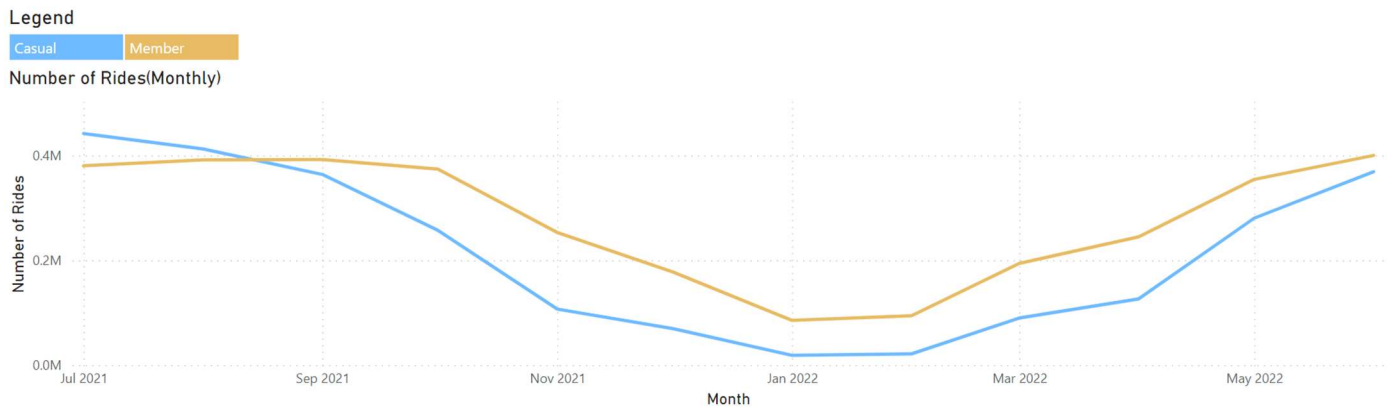

Please refer to:

https://github.com/francisbautista15/Cyclistic-Project/blob/main/Cyclistic.sql

Once I had generated my desired tables using SQL, I then transferred them over to individual Excel files where I then proceeded to use PowerBI to produce visualisations.

Total Number of Rides (Monthly)

- Casual riders tend to ride more frequently during the warmer months within Chicago (around the July and June period). In addition, Casual riders tend to drop around the winter period (around July and February).
    - Recommendation
      Cyclistic should use the winter downtime to prepare marketing for the summer period. The most optimal time for Cyclistic to launch their marketing is just before the peak period for casual users (around March, April and May).

Total Rides (per Rideable Type)
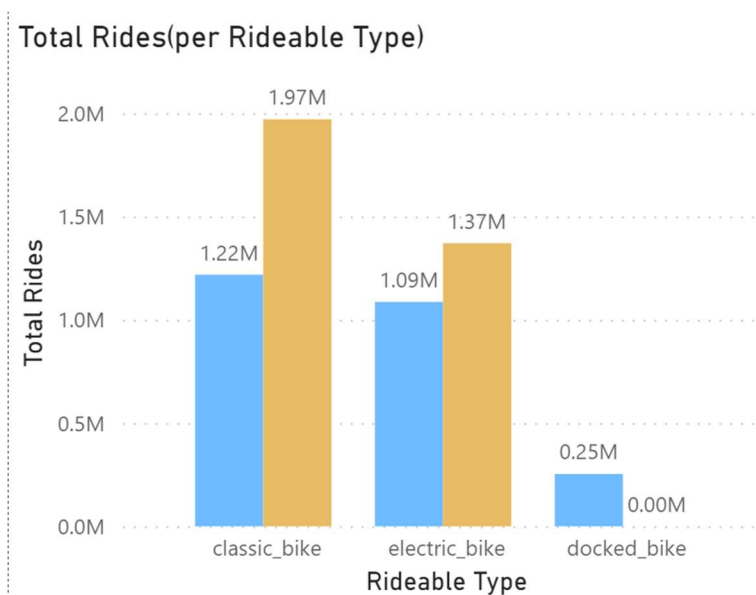
- Classic bike and electric bike are two out of the three most popular rideable options.
  - o Recommendation:
    Suppose to draw more shift in promotions and marketing towards electric and docked rideables to increase the ratio between annual members and casual members for these two ridable types. Suppose Cyclistic should consider removing their docked bike service due to the limited utilisation of this service and revaluate whether by shifting funding from docked bikes to marketing will provide a greater improvement to Cyclistic.

Average Ride Time (per Day of Week)

Average Ride Time(per Day of Week)
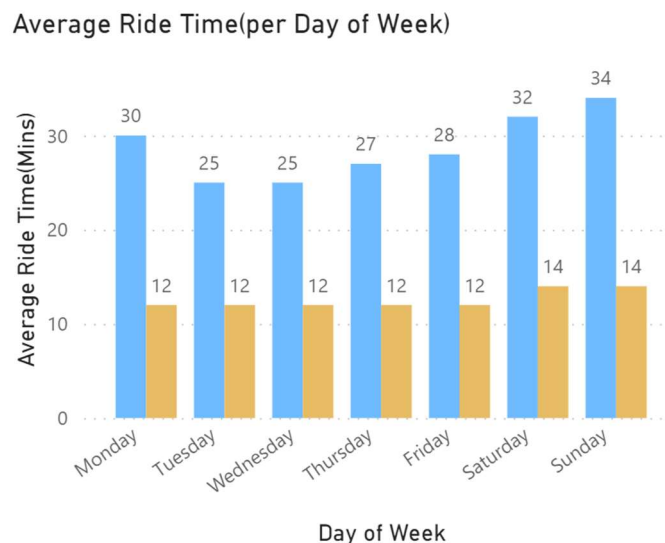
- We can see that casual riders have an overall higher average ride time of about 110%. Using this information, we can predict that annual members primarily consists of users who use Cyclistic for commute purposes where casual members primarily use the service for leisure.
  - o Recommendation:
    Recommend using this information to push marketing towards leisure purposes rather than commute purposes.

Total Rides (per Day of Week)

Total Rides(per Day of Week)
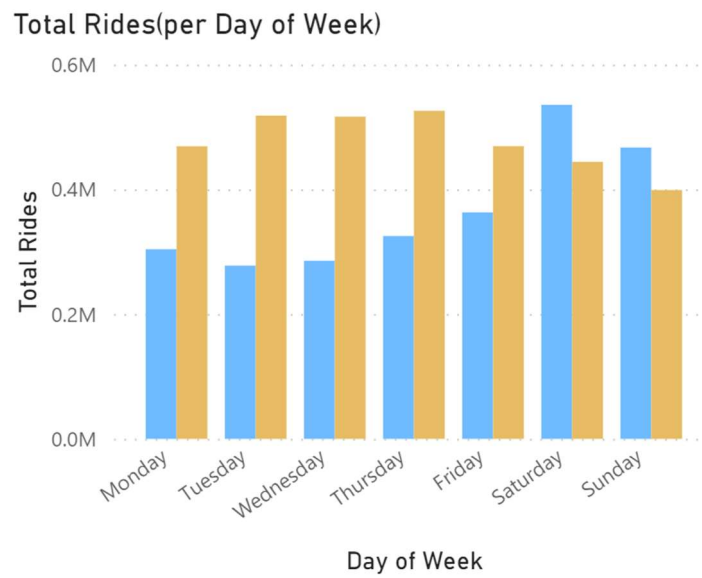
- Casual members increase during the weekends. This strengthens the idea in which is how casual riders use the cyclistic service for leisure rather than commute purposes.
    - o Recommendation:
    Push marketing towards signing up to annual memberships for leisure purposes. Possible even have a promotion where annual memberships are on discount on holiday weekends.

Factors not Considered Within This Analysis

- Age and gender

  This would add a dynamic to whether customers are being targeted across demographic lines. Is the existing marketing effective? Is there potential for more inclusive targeting?

- Pricing structure:

  Data on pricing plans were not provided which would've given further insight to which plans are the most popular (by price) when contrasting the plans. It would also be effective to understand the spending behaviour of casual users.

- Household income data:

  Analysing the average income of long-term members as compared to the casual members would allow for further analysis to what is the typical economic standing of each type of member, as well as providing the ability to analyse overall price sensitivity between the two different membership types.