Case Study: How Does Bike-Share Navigate Speedy Success

Objective

The objective of the analysis is to better understand Cyclistic's customer behaviour in order to support the marketing campaign and strategy of converting casual riders to annual members.

The key questions that the analysis aims to answer are as follows.

- What are the differences between an annual member and a casual rider?
- Why would a casual rider buy an annual membership?
- How does digital media affect the company's marketing tactics?

Data Source

Data used is made available by Motivate International Inc at this link. It is public data that explores how different customer types are using Cyclistic bikes.

Scope: Data between Mar 2022 and Mar 2023 are considered for this analysis.

Data Pre-processing

Given that the data downloaded for each month are stored in separate csv files. The data are combined into a large dataset consisting of **6,087,762** records before cleaning.

A summary of the pre-processing steps are as follows. Detailed code and explanation can be found in the *cyclistic_cleaning.ipynb* notebook within the same folder.

- Converting *started_at* and *ended_at* to timestamp
- Drop records with missing values in *end_lat* and *end_lng* columns
- Added new attribute *duration_in_mins*
- Added new attribute *distance_km*
- Added new attribute *day_of_week*

After cleaning, a total of **5,694,352** (93.53%) of the records remain.

It was noted that there remain missing values for the *start_station_name*, *start_station_id*, *end_station_name*, *end_station_id* columns for older datasets. These records were not removed as they do not affect the analysis below.

Subsequently, *pandas_profiling* library is used to generate an initial report for sanity checks on the data cleaning process and provide an initial basic understanding of the final cleaned

dataset that we'll be working with. The report, *Analysis.html*, is downloadable in the same folder.

Analysis & Findings

Using the *.describe()* function, we can better understand the full data set based on the following information.

Ride duration (in minutes)

- Mean: 15.19
- Min: 1.00
- Max: 34,354.00

Ride distance (in km)

- Mean: 2.2553
- Min 0.0001
- Max: 9,814.083

Day of week

- Mode: Sat

While plotting the distribution of ride duration and distance between member and casual riders, it was noted that extreme values are observed as shown in Figure 1 and 2 below. As such, these outliers were removed from the dataset before examining the underlying distribution of ride duration and distance in Figure 3 and 4.

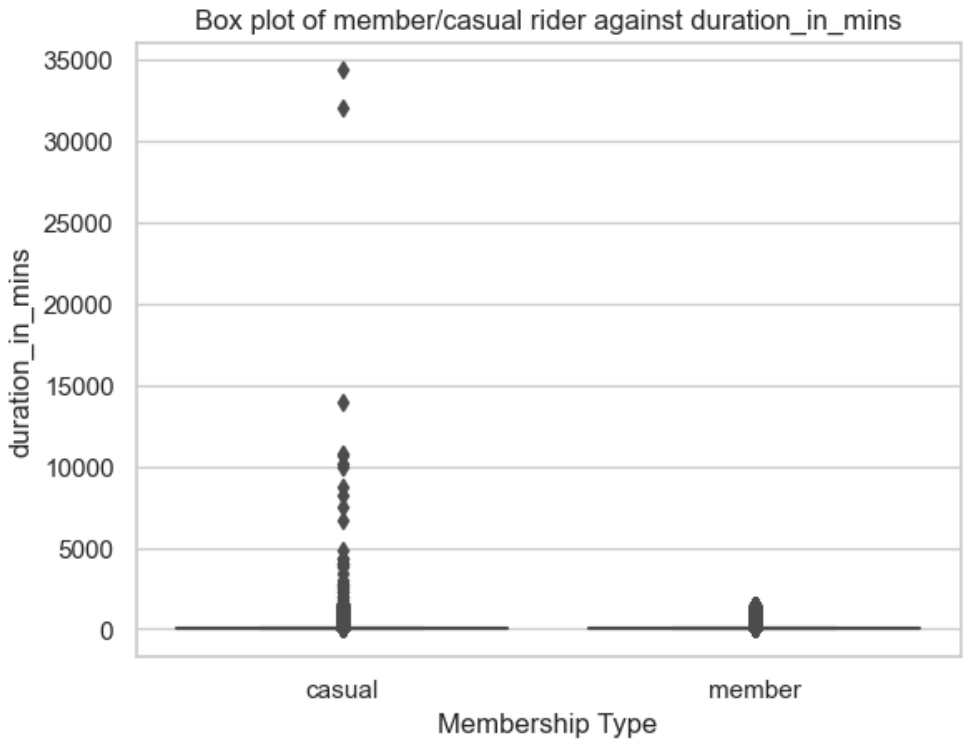

Figure 1: Distribution of ride duration among member and casual riders

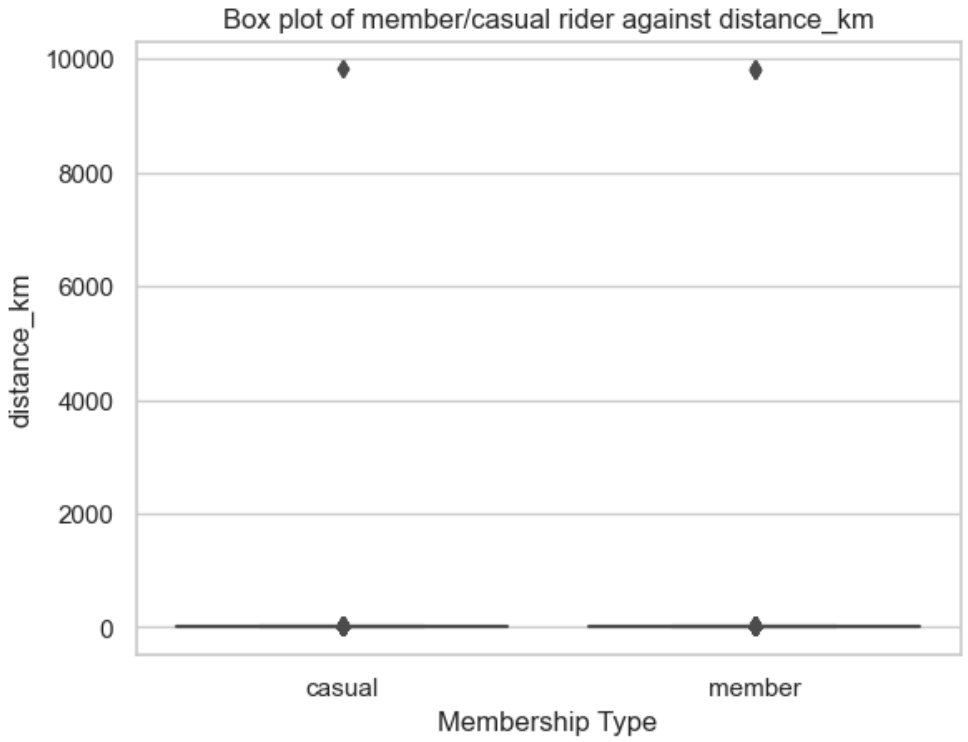

Figure 2: Distribution of ride distance among member and casual riders

After removal of outliers, we apply log computation to the ride duration due to the large range of values to improve the chart readability and increase the ease of analysis.

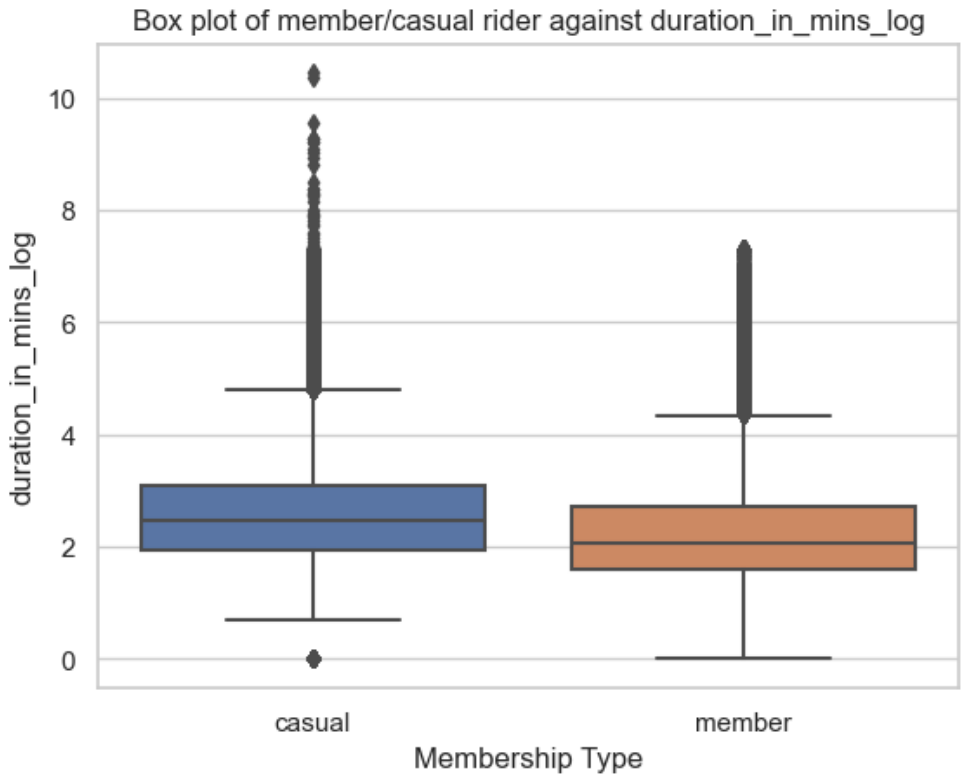

Figure 3: Distribution of ride duration among member and casual riders

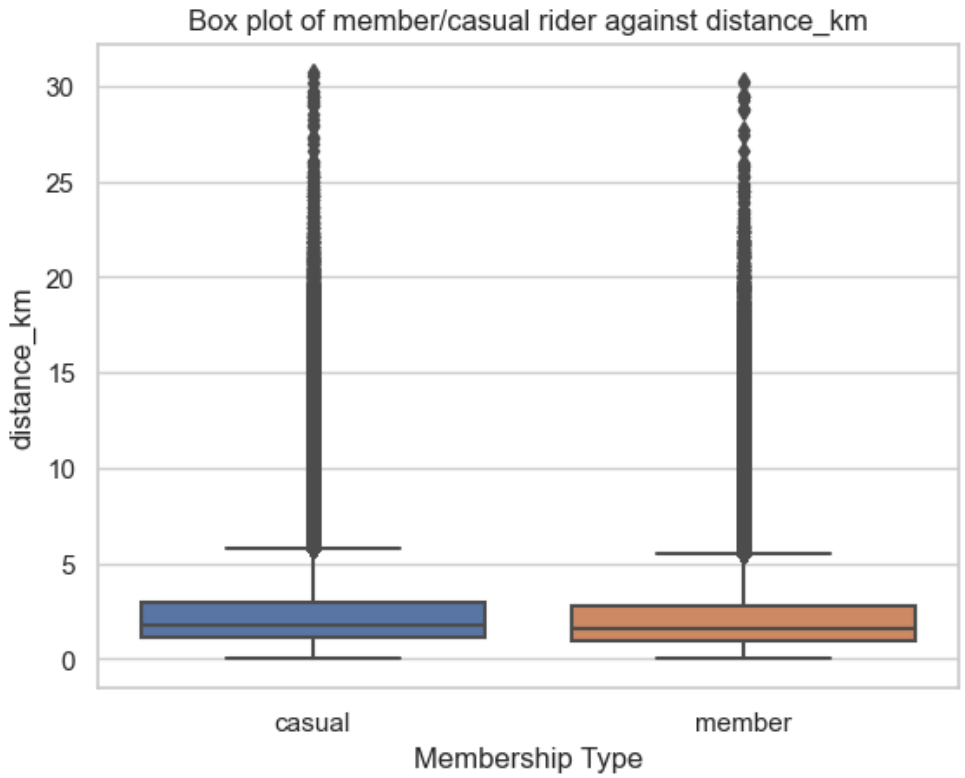

Figure 4: Distribution of ride distance among member and casual riders

From the charts above, we can infer that the distribution is quite similar in terms of the ride duration and distance between member and casual riders.

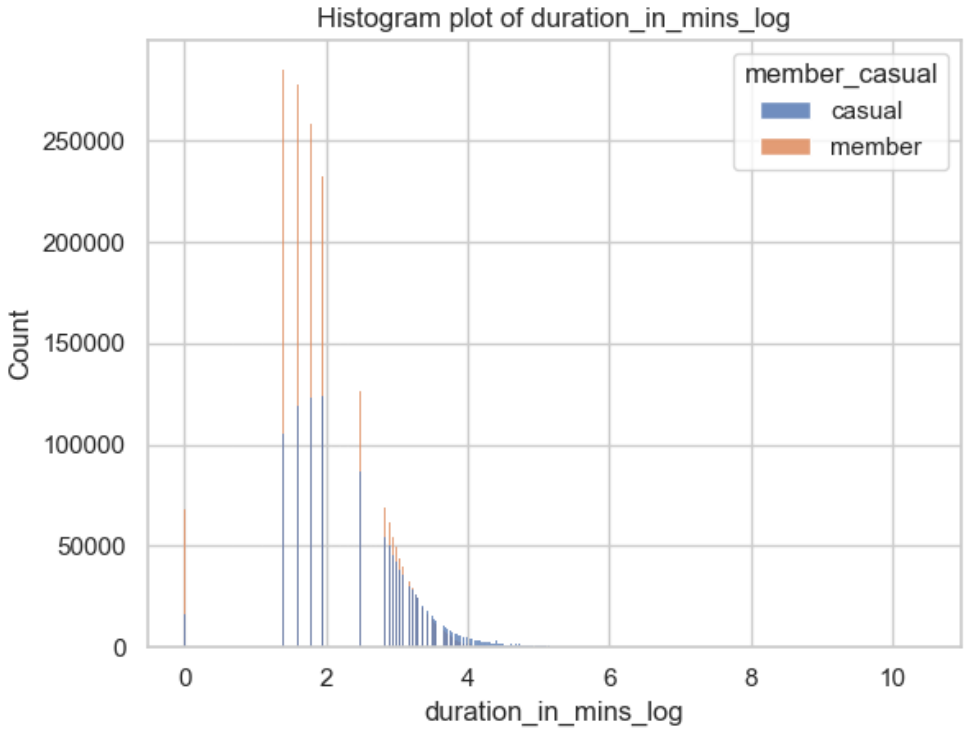

Figure 5: Histogram of ride duration computed with log

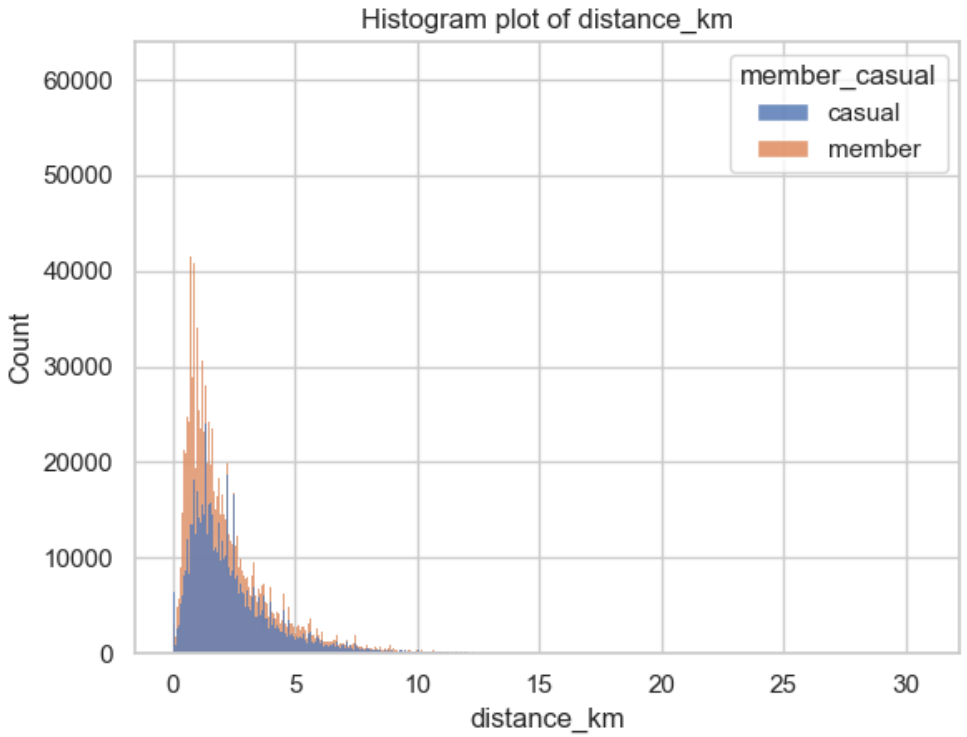

Figure 6: Histogram of ride distance

As such, the marketing campaign should not focus on ride distance and duration.

Upon diving into the number of trips taken across the days in a week, we noted a significant difference between member and casual riders: **more members ride the bike during the weekday**.

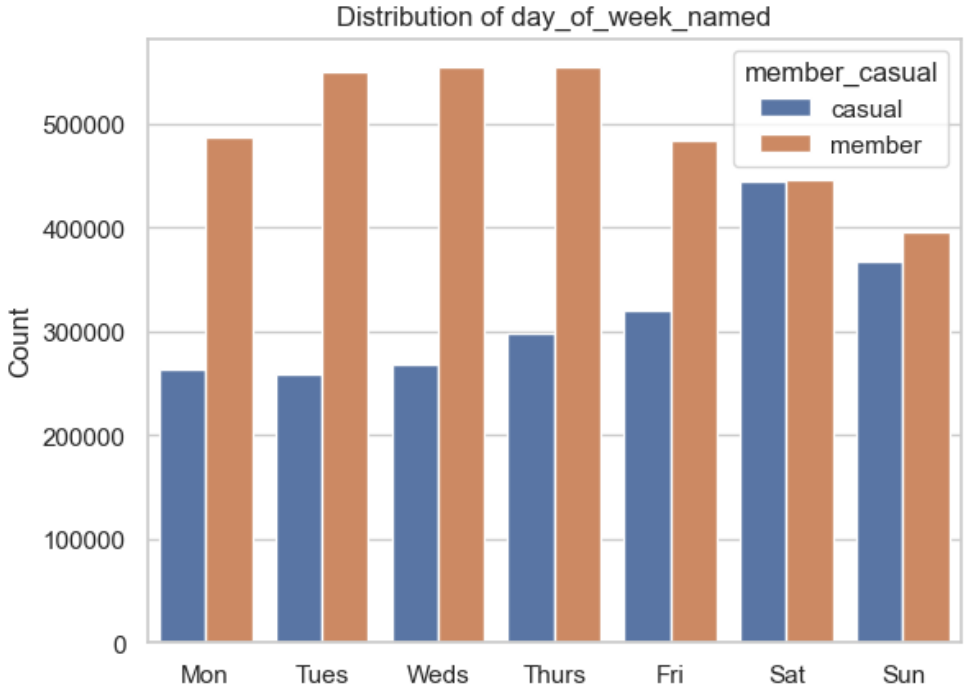

Figure 7: Count of bike trips in a week for member and casual riders

Therefore, the marketing can be targeted at casual riders who ride during the weekdays, potentially for daily commute.

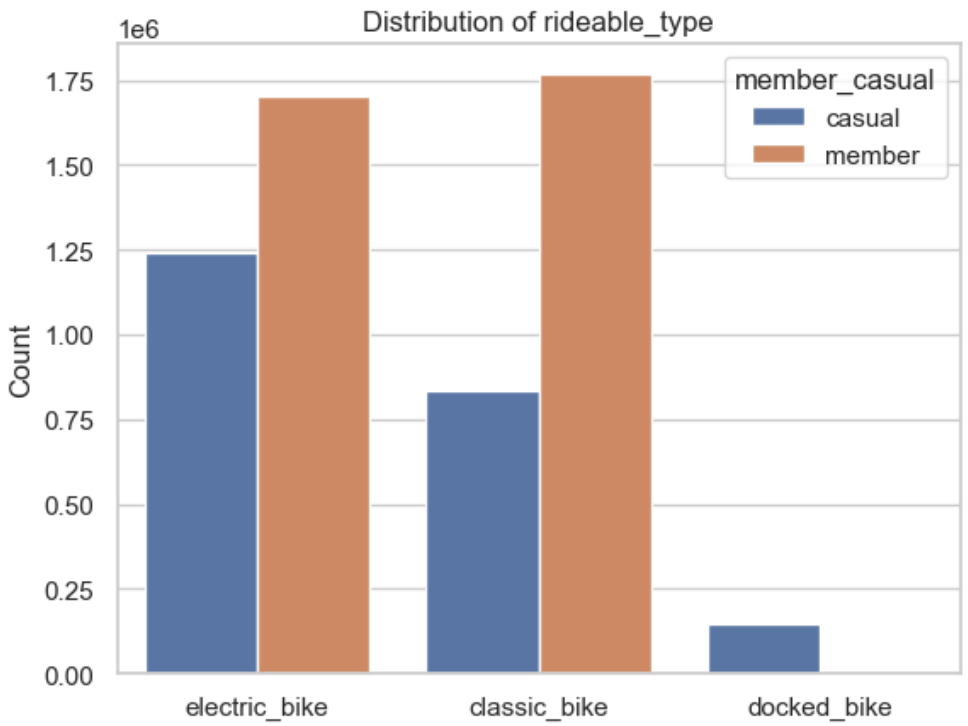

Figure 8: Count of bike type member and casual riders

Based on Figure 8, it is observed that more members prefer classic bikes as compared to casual riders. However, more information may be required to determine if the results were due to the bikes availability or members' preference.

It is possible for the marketing campaign to consider the healthy lifestyle of daily commute using classic bikes at an affordable subscription rate to target casual riders who have yet subscribed to the membership.

Summary & Recommendations

In summary, the data shows that there are more members who ride bikes during the weekdays and utilise more classic bikes as compared to casual riders.

The potential ideas which can be considered in the upcoming marketing campaign to convert casual riders into members include:

- Using bikes for daily commute as a healthy and active lifestyle.
- Showcase the different types of bike available, including classic bikes which are better for those interested in keeping fit and active.

Marketing channels that can be considered for the marketing campaign include direct electronic mailers or push notifications for casual riders who subscribed, social media platforms, official company's website, and billboards near bicycle parking space.