

Strategizing Cyclistic's Campaign to Convert Casual Riders to Valuable Member Riders

Julia Luong, Aug 2023 – Sep 2023
julialuongw@gmail.com

Table of Contents

Purpose	1
Preparing the Dataset	2
Descriptive Analysis	2
Insights Summary	7
Recommendations	7
Possible Further Study	8
Appendix	8

Purpose

This project aims to uncover customer trends of Cyclistic and to give data-informed recommendations to the financial analyst team's latest marketing campaign. That is, their effort to convert casual riders to member riders. In this report, we particularly provide descriptive distinctions between casual riders and member riders, and what to focus advertising efforts on. The large-scale conversion of casual riders to members has been designated by the financial team as the key to spearhead growth for the company, given that it now has a sufficiently large customer base owed to their flexible price advertising campaigns.

As a reminder, casual riders are those who have either purchased a single-ride pass or full-day pass, and member riders are those who have purchased an annual pass.

Below are **three questions central to the campaign** which we will answer in this report:

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?

Key stakeholders include:

- Lily Moreno: Director of Marketing, responsible for developing campaigns to promote Cyclistic
- Cyclistic's Marketing Analytics Team: Analyze Cyclistic data and report on marketing strategy
- Cyclistic Executive Team: The team that will decide whether to approve the recommended marketing program.

Preparing the Dataset

Data Location & Organization. Our Cyclistic trip dataset is located hosted in an AWS (Amazon Web Services) server, in the form of .csv files. The data spans 2014 to July 2023 with monthly and/or quarterly breakdowns. We choose to work with data from the entire 2022 year as they are more current. Each row belongs to one ride and contains its bike type, start and end time, start and end location, and member type.

Checking ROCCC Requirements.

Reliability/Integrity: The data has been checked for typos, missing values, and duplicates. We find no typos or duplicates. About 22% of the rows were missing either the start or end location, but these are non-critical. The numeric values (dates, duration) were numerically validated, with some rows removed where they weren't, namely rides lasting 0 or negative seconds. There is no sampling bias as these datasets are collected from the entire population of Cyclistic users. Given all this, we conclude the dataset integrity is high. The cleaning process was done and documented in a Jupyter Notebook; see Appendix.

Original: The data is first-party data directly from the Company.

Comprehensive/Completeness: The attributes are sufficient for gaining useful insights to the two biker populations. For example, the start times of rides can yield insight on when each rider group tends to ride; the start times of rides combined with duration can shed further light on rider behavior; and the start/end locations of rides can show where hot/cold spots are. The size and timespan of the data will allow any monthly or daily trends to be identified.

Current: We analyze the full 2022 dataset and make comparisons to the current trends in 2023.

Cited/Vetted: The parts of the dataset from previous years have been vetted and used to guide past business decisions, such as the previous advertising campaign to generate broad appeal.

Security, Privacy, & Accessibility. The data was provided by Motivate International Inc. under this [license](#). The datasets contain no personal identifiable information, including credit card numbers or addresses, and hence each row/observation is anonymous. The data is representative of the entire Cyclistic user population since this data logs all rides taken by users.

Descriptive Analysis

Our analysis is done in a Jupyter Notebook and is narrated in this section. To focus our analysis on the distinctions between casual and member riders, any statistics we computed we computed for each. Following this section is a list of insights we discovered that would be useful to the conversion marketing campaign.

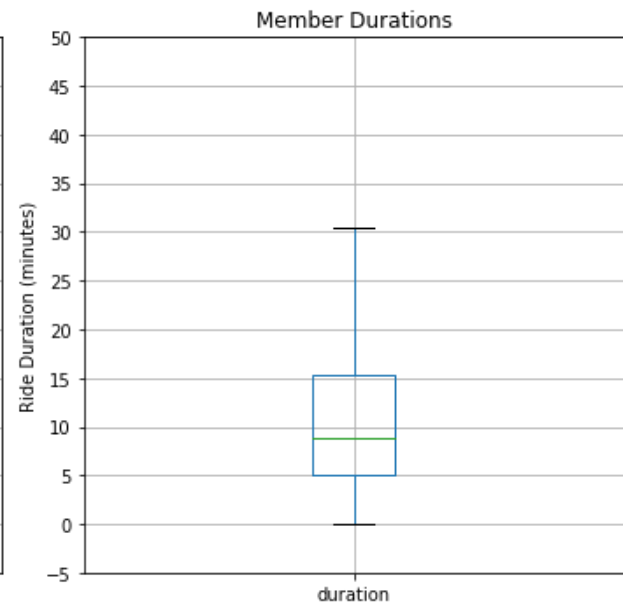
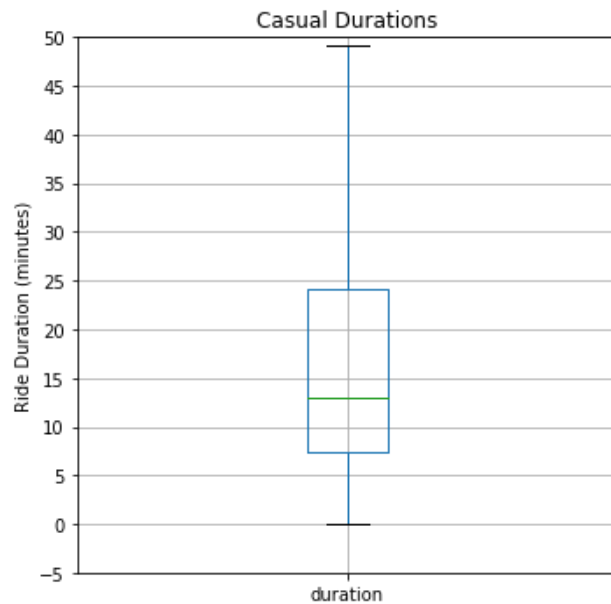
Preliminaries

Number of Rides by Casuals = 2.32 million
Number of Rides Members = 3.34 million
Time Span: January 1, 2022 – Dec 31, 2022

1. Average and Median of Duration ([hh:]mm:ss)

The table below details the [mean & standard division] and [percentiles including the median], and the following boxplots depict the latter.

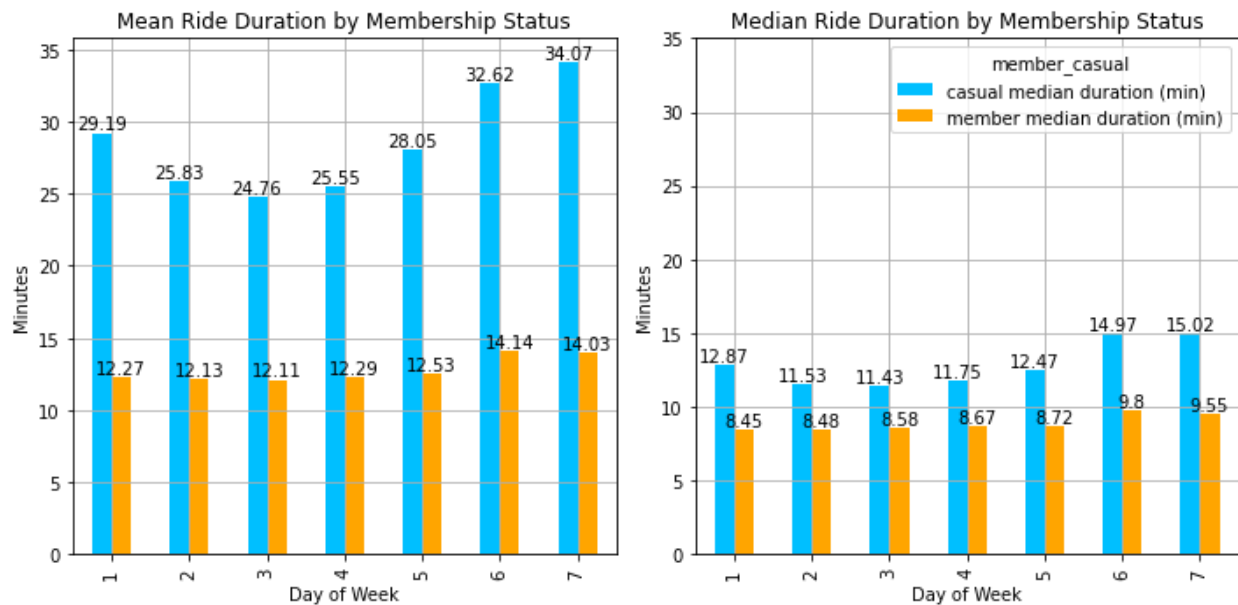
	mean	std	min	25%	50%	75%	max
Casual	29:09	04:32:42	00:01	07:20	13:00	24:06	29 days
Member	12:43	28:48	00:01	05:07	08:50	15:16	25:59:54



It is evident from the values and sample size that casual riders are significantly more likely to have longer rides than member riders. However, there is still significant overlap between the two rider types.

2. Ride Duration by Day of Week

Duration by Day of Week



(1=Monday, ..., 7=Sunday)

The trend from Section 1 is apparent across all weekdays.

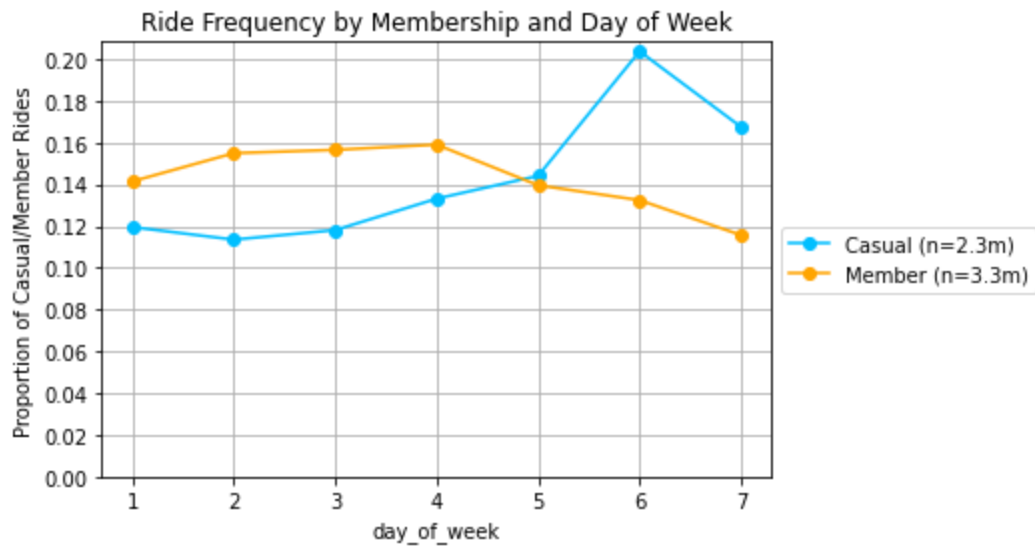
More interestingly, however:

- Both types of riders tend to have longer rides on the weekends and Friday than on the weekdays Monday through Thursday.
- The durations for members on the weekdays are very consistent, while those of casuals are comparatively more variable.

We hypothesize these trends are due to many casual riders riding for recreation or getting to a variety of places, on Fridays and the weekends; and members riding to commute to a stationary location such as work or school, consistently on the weekdays. These hypotheses are further bolstered by the findings in the following Sections 3 – 5.

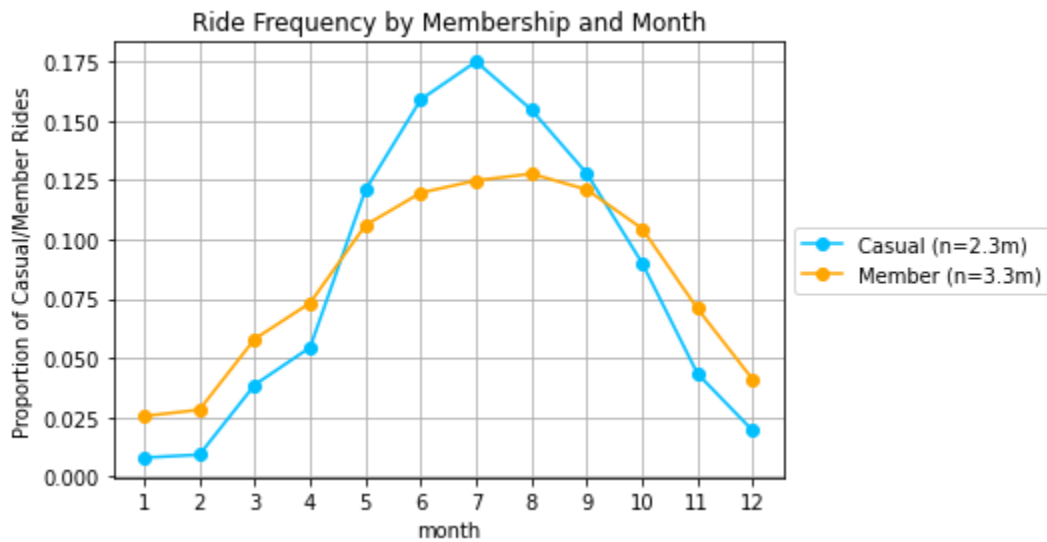
Administering a survey to both groups for their reasons for using the bikes would help clarify the exact reasons for the trends, but we believe it is safe to assume casual riders are more likely to ride for occasional recreation than a consistent means of transportation as members do.

3. Bike Usage Frequency by Day of the Week



Casual riders are more likely to ride as the weekend approaches, and member riders are more likely to ride on weekdays than on weekends.

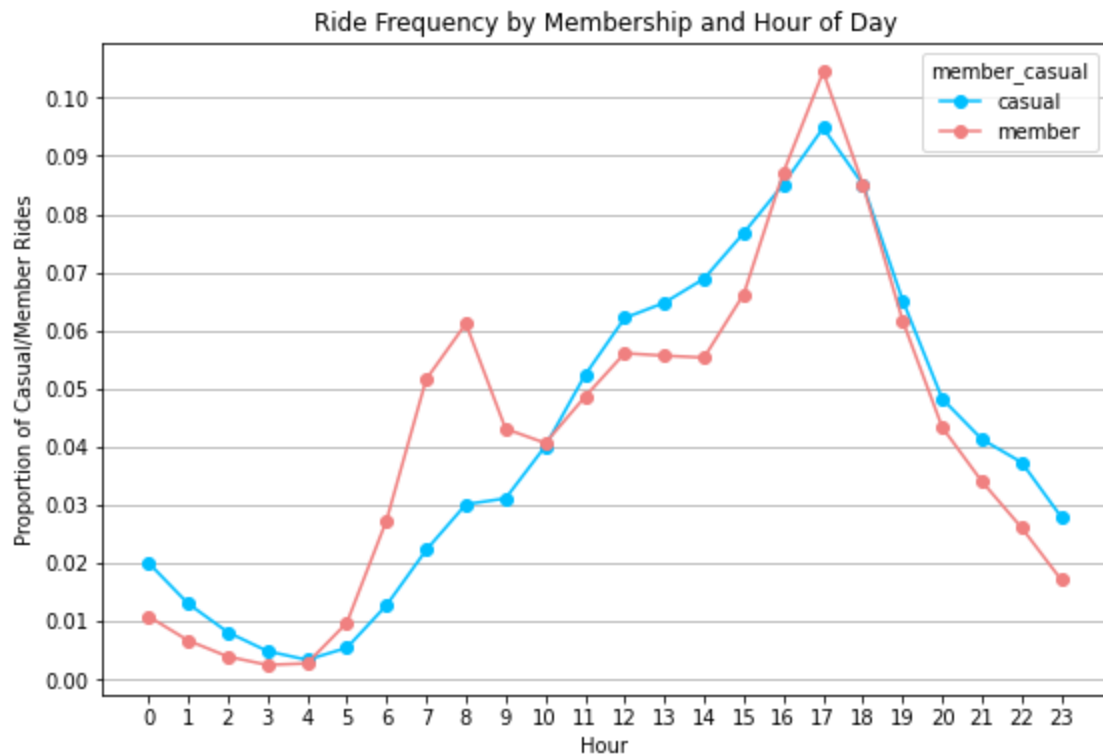
4. Bike Usage Frequency by Month



Casual riders are more likely to ride during late spring, summer, and autumn. There is a spike of usage in May, peaking in July, and dropping back to pre-May levels by November.

Member riders follow a similar trend, though their usage during the winter months are substantially higher than that of casual riders.

5. Bike Usage Frequency by Membership and Time of Day

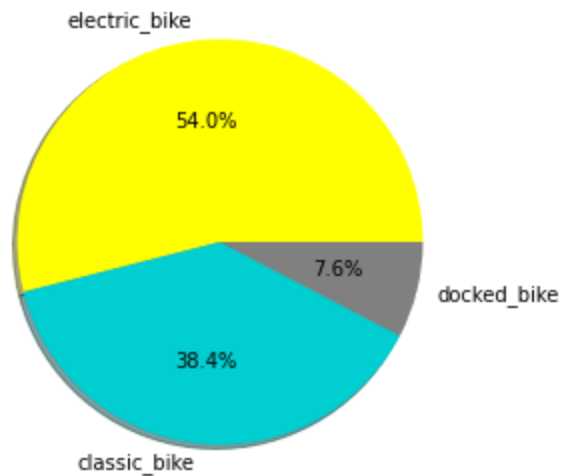


We notice that the ride frequency for member riders significantly surpass that of casual riders from 5am to 9am, and similarly during 4pm to 5pm. These are the usual times when people get off work, further validating the hypothesis that members are more likely to ride to commute to and from work.

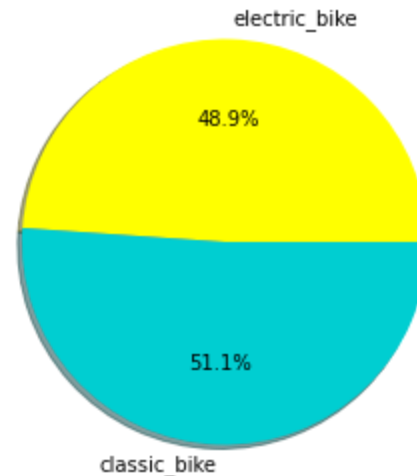
What is also interesting is that the casual line does not differ much from the member line -- having a steady rise from the morning, peaking at 5pm, and dropping off. We are unsure of why this trend is as such, but this similarity indicates our casual audience might be more primed than previously inferred to purchase a membership to commute to and from work.

6. Ride Type Composition by membership status

Composition of Bike Types of Casual Riders



Composition of Bike Types of Member Riders



(We note that docked bikes are only available to casual riders.)

The compositions are roughly the same, with casual members having a slight preference for electric bikes.

Insights Summary

We extrapolate from the analyses the following distinctions between casual and member riders. Note that these are generalizations and are not applicable to every rider.

Casual Riders	Member Riders
Ride for recreation on weekends and Fridays	Ride to work/school on weekdays
Enjoy riding in and around the summer	Have a more level usage throughout the year, with a less dramatic peak in and around summer
Do longer rides than members by 3-5 minutes	Do shorter rides than casual riders by 3-5 minutes
Follow steady rise and drop in usage throughout any given day	Follows steady rise and drop in usage throughout any given day, but notably with peaks during the morning and evening hours.

Recommendations

Given these insights, we curate the following recommendations to Cyclistic's Executive Team for review:

(1) Advertise the following to casual riders:

- The opportunity to use Cyclistic as an economical, efficient, and healthy option to commute to work or school year-round.
- The opportunity to routinely ride during the weekend for either exercise or recreation, especially in and around the summer season.

(2) Distribute free, discounted, or refundable trials for the perks of an annual pass to casual riders, especially during the summer months when usage is highest. For budget considerations, these trials could be limited to casual riders who have already purchased full-day passes in the past.

(3) As briefly mentioned in the analysis, administer surveys to collect more data on casual and member riders to deliver more personalized advertisement. We suggest towards casual riders a survey to collect their interests and needs, and towards annual members a survey to collect positive testimonials and any other features they want from their membership.

Possible Further Study

For further study, we can consider the geographical data for the rides, which is readily available. This could entail uncovering possible hot or cold spots for ridership, informing potential spots to advertise or spots to install new docks, respectively.

The type of bike could be more thoroughly investigated to uncover trends between it and the other attributes: time of ride, duration, and location.

For a more advanced task, a cluster analysis might be useful to further uncover trends and groups among the riderbase.

Appendix

Data Cleaning & Analysis Log (Jupyter Notebook) <[link](#)>
