



# Cyclistic Bike-Share Case Study

Larry Miller

[larrymillerprogramming@gmail.com](mailto:larrymillerprogramming@gmail.com)

Updated: Dec 25, 2021



# **Data Analysis Case Study:**

## Cyclistic Bike-Share Annual Memberships

### **Purpose of this project:**

The purpose of this case study is to demonstrate my skill in  
**six key steps of data analysis:**

- Ask meaningful, measurable questions inspired by a key business task
- Prepare and organize data for analysis
- Process data, ensuring the data is clean and reliable
- Analyze data to find insightful trends, outliers and relationships
- Share findings through visualizations and reports
- Provide recommendations for a future business plan, based on my findings



## About this case study:

Cyclistic is a fictional bike-share company based in Chicago. Users of the service are divided into two main categories: members and casual riders. Members pay an annual membership fee to access the bikes, and casual riders rent bikes for a single trip or full-day pass.

Cyclistic's director of marketing has the goal of maximizing the number of annual memberships. To do this, they want to focus on current "casual rider" customers and find strategies to encourage them to purchase an annual membership.



## **Deliverables of this case study:**

- Clear statement of the business task
- Description of all data sources used
- Change log of all data cleaning/processing
- Summary of data analysis
- Visualizations and key findings
- Top 3 recommendations for the future business plan



# **The business task and guiding question:**

## **Business Task:**

Encourage current “casual rider” users to upgrade to a yearly membership.

## **Guiding Question:**

How do members and casual riders use the Cyclistic service differently?



## Description of the data:

Motivate has supplied first-party data for this case study. Cyclistic is a fictional bike-share company, but we are using Motivate's authentic data to perform this analysis. For this project, we are using the trip data for the past 12 months (December 2020 - November 2021).

The data for each trip includes ride id, type of bike used, time of start and end, station name and station id of start and end, latitude and longitude of start and end, and the type of user (member or casual).



## Limitations of the data:

Many instances are missing data for start station and end station.

A small percentage of rows have incorrect start and end times (end time before start time), so they cannot be used in calculations.

To meet data privacy standards, customer data is limited. Therefore, we cannot see if a customer is in the Cyclistic service area.



# Questions for Analysis:

Guiding question:

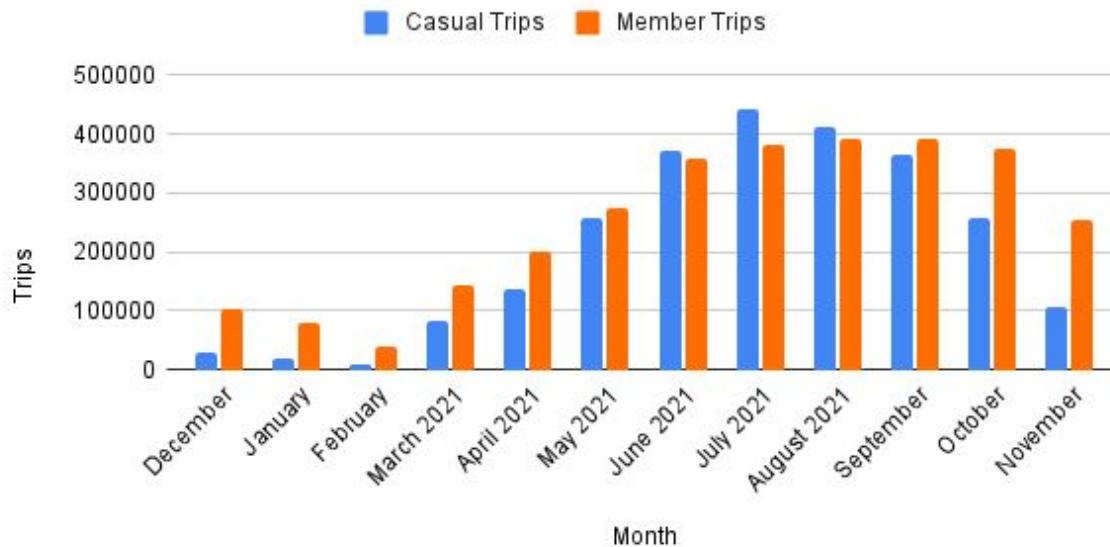
How do members and casual riders use the service differently?

Questions to consider:

- In this 12-month review, how do the months compare in total trips?
- Is there a noticeable difference in member trips vs. casual trips for each month?
- How does the average trip length compare between current members and casual riders?

## Cyclistic 12-Month Review

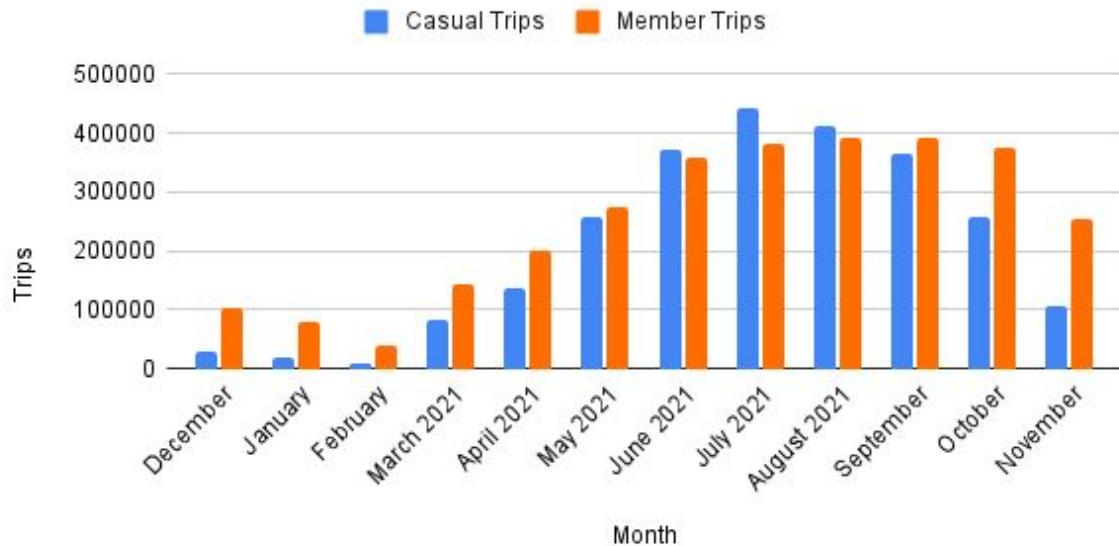
Casual Rider Trips and Member Trips by Month



Trend 1: Traffic is slow in the winter months, and it increases drastically throughout the spring.

## Cyclistic 12-Month Review

Casual Rider Trips and Member Trips by Month



Trend 2: Casual rider traffic declines faster in the fall when compared to current members.

## Average Trip Length

Casual Riders vs. Members



Trend 3: Casual riders are consistently averaging higher in trip length than current members. This could represent a higher rate of recreational use.



## Top 3 Recommendations:

- Increase advertising in spring months to encourage memberships for summer and fall.
- Create rewards programs for long trips to encourage recreational users to upgrade to a membership.
- Advertising and discounted rate new membership during slower months.

Details to follow →



# 1) Increase advertising in spring months to encourage memberships for summer and fall.

Goal	Current State	Proposal	Validation
Encourage casual riders to upgrade to a membership when their usership increases.	Traffic is increasing greatly in the spring, but casual riders are not signing up for the membership.	Increase advertising for memberships around this time.	A/B test changes new members during the spring months.



## 2) Create rewards programs for long trips to encourage recreational users to upgrade to a membership.

Goal	Current State	Proposal	Validation
Encourage casual riders to upgrade to a membership by incentivizing their longer trips.	Casual riders' trips are consistently much longer than current members.	The membership program could include rewards for long distance / time per month.	A/B test changes with membership rates and average trip length of members.



### 3) Advertising and discounted rate new membership during slower months.

Goal	Current State	Proposal	Validation
Increase business in the slower months.	Business is much slower from November to March.	Discount membership rate during slow months.  Partner with local businesses that peak in the winter months.	A/B test changes with membership rates and profit.



## Consider other ideas for growth:

Referral program for current members to recruit new members.

Survey current members to better understand relationship between casual riders and annual members.

- Had you used Cyclistic before becoming a member?
- How often do you use the Cyclistic service?
- Do you use the service for daily commute, recreation, or both?



# Documentation of Data Processing (Change Log)

Created summary spreadsheet with the following columns: month, casual\_trips, member\_trips, total\_trips, avg\_cas\_trip, avg\_mem\_trip

Used R programming language to calculate monthly casual trips, monthly member trips, total monthly trips, monthly average trip length for casual riders, and monthly average trip length for members. Populated the summary spreadsheet.

When calculating average trip time for each month, rows needed to be filtered out because their end time was before their start time. These rows have inaccurate data, but cannot be manually fixed. Original data was not manipulated, but calculations were completed with filtered data.