Cyclistic Bike Share Analysis

Welcome to Cyclistic Bike Share Analysis. This analysis is conducted as the final project of Google Data Analytics Profesional Certificate. In this analysis i will be analyzing a fictional company called Cyclistic Bike Share using **Rstudio** and analysis phase **ask, prepare, process, analyze, share, and act**.

Company Background

Cyclistic, a bike-sharing program, was launched in 2016 and has since grown to include 5,824 bicycles and 692 stations in Chicago. The bikes can be unlocked and returned at any station within the network. The marketing strategy has focused on raising awareness and attracting a broad consumer base by offering flexible pricing plans, including single-ride passes, full-day passes, and annual memberships. Cyclistic's finance analysts have determined that annual members are more profitable than casual riders. To drive future growth, the company aims to convert casual riders into annual members, leveraging their existing awareness of the program. Moreno, the marketing lead, has set a clear goal to design marketing strategies for this conversion. To achieve this, the team plans to analyze historical bike trip data to understand the differences between annual members and casual riders, explore the motivations behind purchasing memberships, and assess the impact of digital media on their marketing tactics.

Business Task (Ask Phase)

Three questions will guide the future marketing program:

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?

Moreno has assigned you the first question to answer: **How do annual members and casual riders use Cyclistic bikes differently**?

Preparing the Data (Prepare Phase)

I will be using Cyclistic's historical trip data to analyze and identify trends. You can download the file here **here**.

(Note: The datasets have a different name because Cyclistic is a fictional company. For the purposes of this case study, we use divvy trip data. The datasets are appropriate and will enable you to answer the business questions. The data has been made available by Motivate International Inc. under this license.) This is public data that you can use to explore how different customer types are using Cyclistic bikes. But note that data-privacy issues prohibit you from using riders' personally identifiable information. This means that you won't be able to connect pass purchases to credit card numbers to determine if casual riders live in the Cyclistic service area or if they have purchased multiple single passes.

Data Cleaning and Manipulating (Process Phase)

I will not be showing the Rstudion code here. You can see the full code of my analysis here.

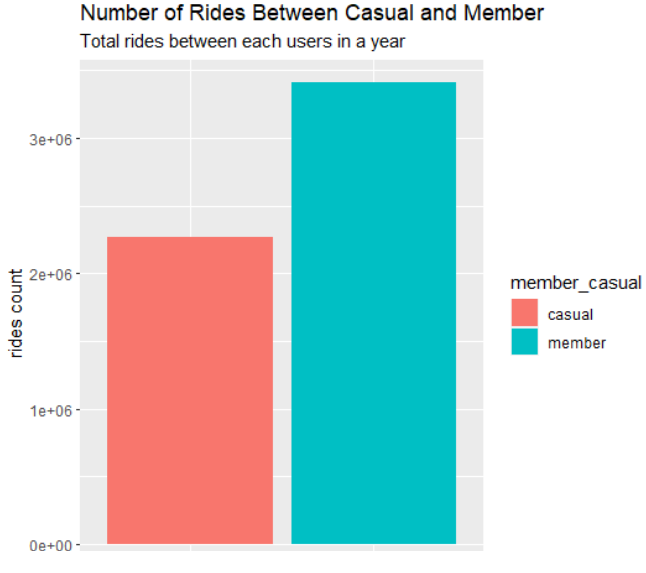
Data Analysis and Visualization (Analyze Phase)

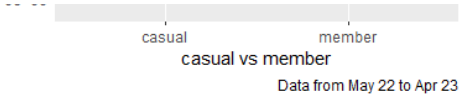
Following the business task **How do annual members and casual riders use Cyclistic bike differently.** I'll be focusing the analysis around how member and casual riders differ in using the Cyclistic bike, such as number of ride, ride duration, when the ride occur based on time, day, and seasons also busiest station between users where they rent the bike.

- **Casual and Member Users Comparison**

First thing first, here are the comparison between the number of casual and member riders across a year.

	member_casual	count_ride	percentage
1	casual	2271843	40.01157
2	member	3406122	59.98843

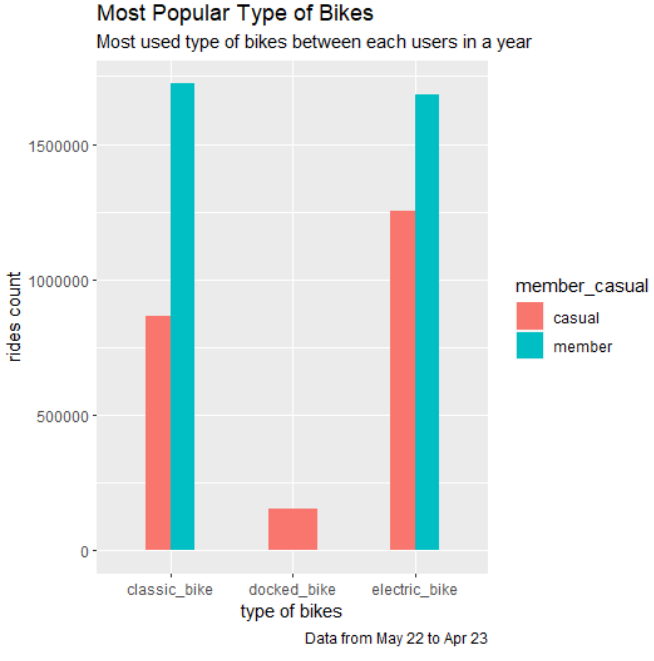




From the table above we can see that there are about 20% more member riders than the casual riders. Assuming that member riders are people who use the bike often maybe to go to work and the casual rider just use it for leisure ride, and sightseeing the city.

- **Most Popular Type of Bikes**

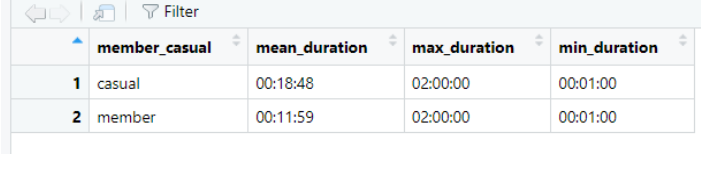
There are three types of bike that Cyclistic Bike Chare Company provide. Which are a classic bike, electric bike, and a docked bike.



From the chart above we can see that classic bike and electric bike are the most popular between users. Meanwhile only casual riders used docked bike.

- **Average Ride Duration Analysis**

Before we get on the count and ride duration analysis. Here are the average, max, and min duration of ride session between users.



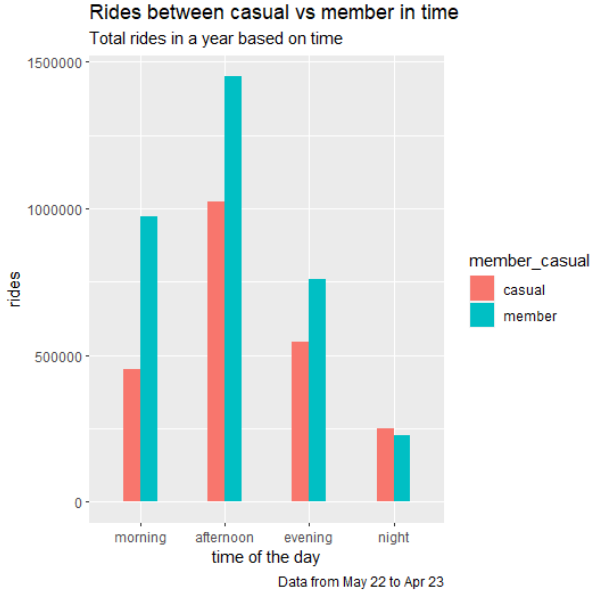
Filter

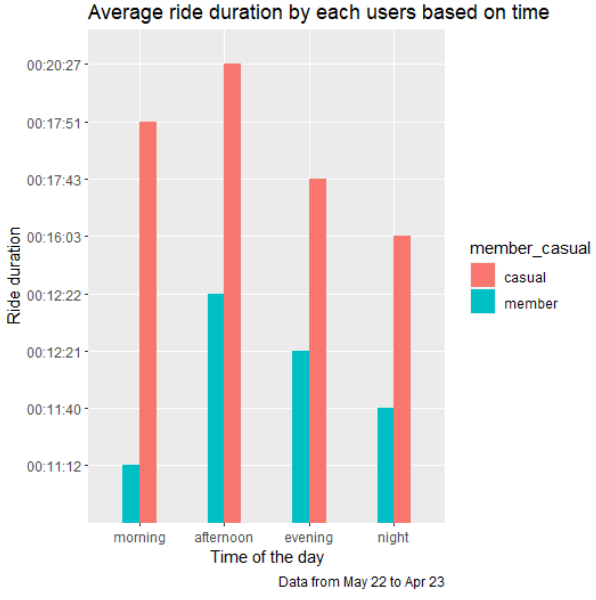
	member_casual	mean_duration	max_duration	min_duration
1	casual	00:18:48	02:00:00	00:01:00
2	member	00:11:59	02:00:00	00:01:00

The table above conclude that casual riders spent more time riding the bikes than the member users. Meanwhile as max_duration and min_duration shows the maximum rides is 2 hour and the minimum rides is 1 minute that's because there is a lot of outliers in the data that exceed above 2 hour to 23 hour of ride duration time and 0 minute ride duration time. That is a possible false data and I removed the outliers in the cleaning process. That's why it only shows the max and min ride are 2 hours and 1 minutes.

- **Ride Count and Ride Duration Based on Time of the Day**

There are 24 hour in a day. So i divide the time to four category. Morning, afternoon, evening, and night time. The morning time begin at 4 am to before 12 pm. Afternoon time begin at 12 pm to before 6 pm. Evening time begin at 6 pm to 10 pm. And the rest is night time.

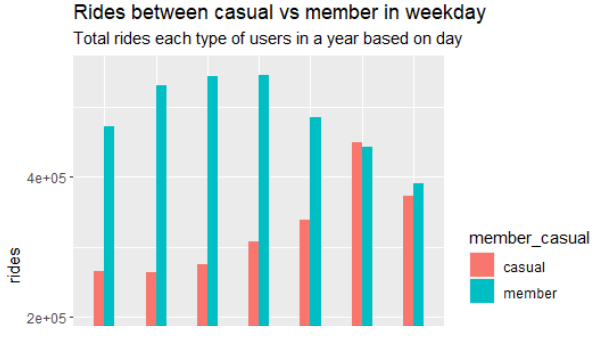


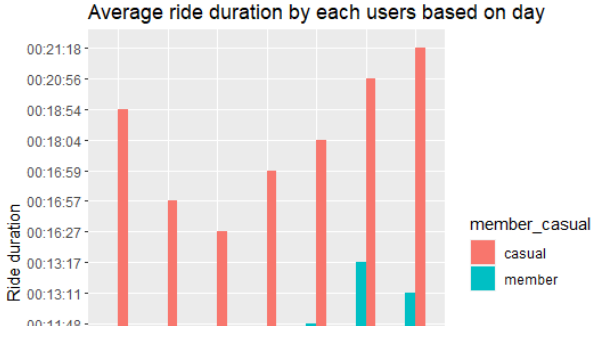


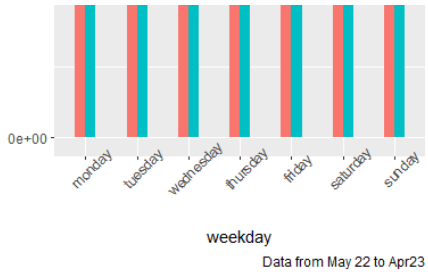
From the chart above we can see that most of the users are using the bike service around morning and afternoon. The highest number of rides is in the afternoon between two users. And the lowest is at night. This could be due to rush hour people using bikes to go to work or to do their activities during the day. Meanwhile the casual riders almost double the time in duration of using the bikes. This could be due to a lot of possibility, but mostly casual riders using it to do leisure rides around the city.

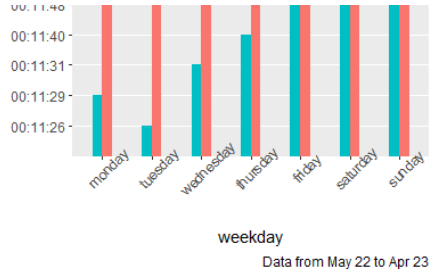
- **Ride Count and Ride Duration Based on Day of the Week**

This shows number of rides and how long each users use the bike depending on what day it is in a week.





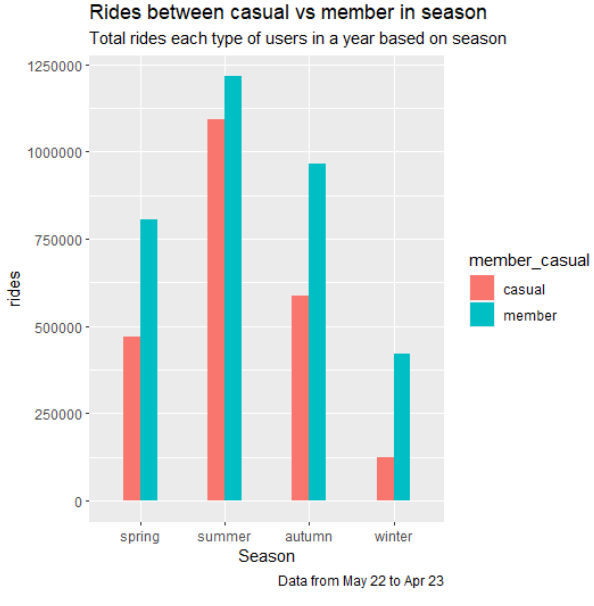


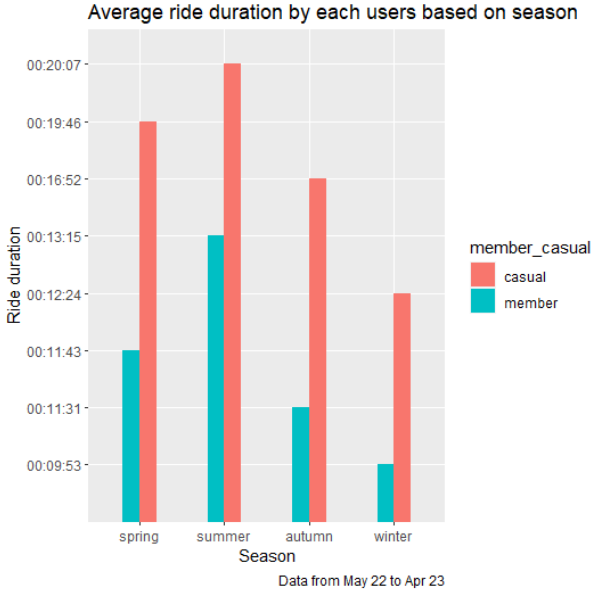


From the chart above we can see the different in bike use between users. Member users mostly use the bike between monday to friday and then started to slightly decline towards the weekend. But the casual riders is the opposite. In terms of duration of the ride it's slightly incline towards the weekend.

- **Ride Count and Ride Duration Based on Season**

I'll divide twelve month into four season. Starting with Spring, begin at the month of March to May, then Summer begin at June to August, then Autumn begin at September to November, and lastly Winter at December to February.





From the chart above we can see that the rides mostly occurs between spring to autumn, and declining towards winter due to cold weather. The duration of the rides also show similar tren.

- **Top Ten Stations Between Users**

Here are the top ten list of the most visited station that users start using the bike from.

	start_station_name	member_casual	count
1	Streeter Dr & Grand Ave	casual	54711
2	DuSable Lake Shore Dr & Monroe St	casual	30667
3	Michigan Ave & Oak St	casual	24243
4	Millennium Park	casual	23721
5	DuSable Lake Shore Dr & North Blvd	casual	22721
6	Shedd Aquarium	casual	19675
7	Theater on the Lake	casual	17864
8	Wells St & Concord Ln	casual	16119
9	Dusable Harbor	casual	13626
10	Clark St & Armitage Ave	casual	13404

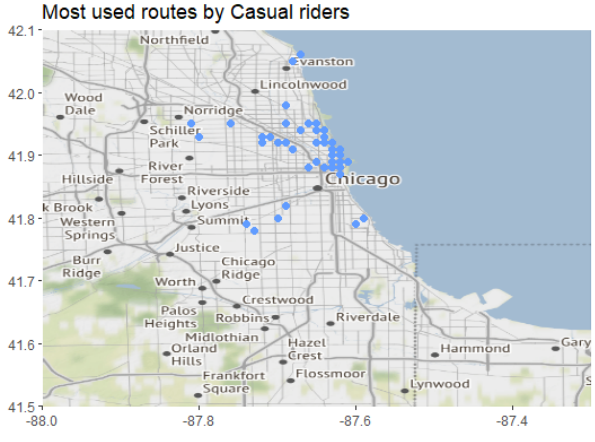
Casual Users

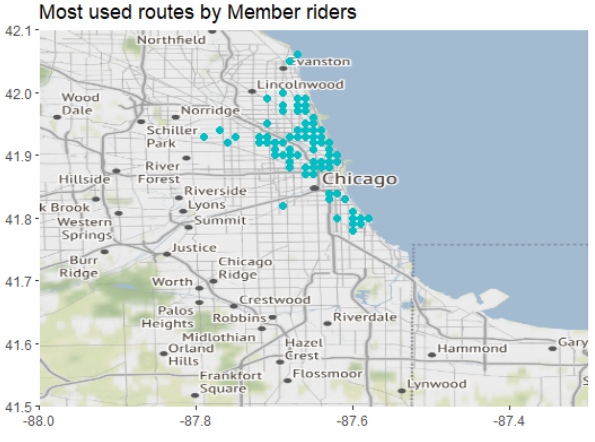
	start_station_name	member_casual	count
1	Kingsbury St & Kinzie St	member	24855
2	Clark St & Elm St	member	22739
3	Wells St & Concord Ln	member	21634
4	University Ave & 57th St	member	21620
5	Clinton St & Washington Blvd	member	21085
6	Ellis Ave & 60th St	member	20871
7	Loomis St & Lexington St	member	20269
8	Wells St & Elm St	member	19477
9	Clinton St & Madison St	member	18940
10	Broadway & Barry Ave	member	18256

Member Users

- **Map Visualization Between Users**

Here are the comparison between casual and member rides around the Chicago city.





From the map above we see that casual users spent more time riding around the downtown Chicago and the coast line. And the member users more spread around the city.

Summary and Recomendation (Share Phase)

This phase should be a presentation to the stakeholder of the Cyclistic Bike Share Company. But since this is an online project, i'll just summarize the analysis and show some recomendation.

- **Summary**

So answering how different between users use Cyclistic bike ? Here is the summary insight from the analysis :

    - Casual users spent more time riding the bikes than member users
    - Member users prefer classic and electric bikes
    - Member users use bike more on weekdays and casual users use bike more on weekends
    - Morning and afternoon time got the most rides between users
    - Casual users spent more time in downtown Chicago and the coast line, and member users more spread around the city
    - There is a significant increase in casual users in the Summer, while member users not so much a significant changes across the seasons except the Winter, both users decrease significantly.
- **Recomendation**

Here are some of my recomendation from this analysis :

- Offer limited time membership for casual users with coupons for attraction around the city since a lot of casual users probably not from the city
- Offer a weekend only membership for casual users
- Offer gift cards and discount for a full membership
- Create a marketing campaign about the benefit of being an annual member. This can be done by digital marketing or conventional in the busiest docking station and at peak month or season.

Act Phase

This phase is done by the Cyclistic's executive team based on the analysis (data driven decision).