**Clear Statement of the business task**

**My Statement**

We have to analyze the customers data to find how different types of the riders use bikes differently And how can we adjust our membership plan so that the casual riders buy our annual membership and provide this data to our stake holders so that they can use this data effectively In marketing campaigns.

Stakeholders

Marketing director

Analyst team

Executive Team

**GPT Statement**

The primary business task is to understand how annual members and casual riders use Cyclistic bikes differently. By analyzing these differences, the goal is to provide actionable insights that can inform marketing strategies aimed at converting casual riders into annual members. This involves examining historical bike trip data to identify usage patterns and trends, which will help Cyclistic design targeted campaigns to increase annual memberships and, consequently, enhance profitability.

Key Stakeholders

* Lily Moreno (Director of Marketing)
* Cyclistic Marketing Analytics Team
* Cyclistic Executive Team

Insights and Business Decisions

* Targeted Marketing Campaigns
* Pricing Strategy
* Digital Media Tactics

**Dataset description**

The data set is downloaded from the divvy-trip database from the given [link](https://divvy-tripdata.s3.amazonaws.com/index.html). Database has the data from 2020 to 2024 but we have downloaded data for last 12 months from 07-2023 to 07-2024 , dataset contain 13 columns description for each is given below.

1. **ride\_id**: A unique identifier for each ride.
2. **rideable\_type**: The type of bike used for the ride (e.g., electric bike, classic bike).
3. **started\_at**: The date and time when the ride started.
4. **ended\_at**: The date and time when the ride ended.
5. **start\_station\_name**: The name of the station where the ride started.
6. **start\_station\_id**: The unique identifier for the start station.
7. **end\_station\_name**: The name of the station where the ride ended.
8. **end\_station\_id**: The unique identifier for the end station.
9. **start\_lat**: The latitude coordinate of the start station.
10. **start\_lng**: The longitude coordinate of the start station.
11. **end\_lat**: The latitude coordinate of the end station.
12. **end\_lng**: The longitude coordinate of the end station.
13. **member\_casual**: Indicates whether the rider is a member or a casual user.

**Process**

I have selected R to work on my data as dataset is very large to work on using spreadsheets,

I have combined all the datasets into one single datasets and then deleted the null values form the dataset , then I created new columns which are give below

* **ride\_length:** calculated by subtracting started\_at from ended\_at
* **weed\_day:** extracted from started\_at
* **remove:** rides with negative ride\_length

**Analyze**

* Mean ride length is almost 10 min
* Min. 1st Qu. Median Mean 3rd Qu. Max.
* "00:00:00" "00:05:52" "00:10:14" "00:16:48" "00:18:21" "114:51:13"
* Members take 64% of total rides while causal customers take only 37% rides

|  |
| --- |
|  |
| * **member\_casual** * <chr> | * **number\_of\_rides** * <int> | * **percentage** * <dbl> |  |  |
| * member | * 3079584 | * 63.95534 |  |  |
| * casual | * 1735626 | * 36.04466 |  |  |

* Most people ride on Saturday
* Casual customers have more ride time on average which is 24 mins and 14 secs while

| * **member\_casual** * <chr> | * **average\_ride\_length** * <dbl> | * **average\_ride\_length\_hms** * <chr> |  |  |
| --- | --- | --- | --- | --- |
| * casual | * 1454.3627 | * 00:24:14 |  |  |
| * member | * 756.0397 | * 00:12:36 |  |  |

* member have almost have the riding time which is about 12 mins and 36 secs

| * **week\_day** * <ord> | * **member\_casual** * <chr> | * **average\_ride\_length** * <dbl> | * **average\_ride\_length\_hms** * <chr> |  |
| --- | --- | --- | --- | --- |
| * Sun | * casual | * 1664.5027 | * 00:27:45 |  |
| * Sat | * casual | * 1630.6650 | * 00:27:11 |  |
| * Mon | * casual | * 1416.5239 | * 00:23:37 |  |
| * Fri | * casual | * 1413.5553 | * 00:23:34 |  |
| * Tue | * casual | * 1291.3136 | * 00:21:31 |  |
| * Wed | * casual | * 1279.2838 | * 00:21:19 |  |
| * Thu | * casual | * 1269.9521 | * 00:21:10 |  |
| * Sun | * member | * 848.8532 | * 00:14:09 |  |
| * Sat | * member | * 844.9760 | * 00:14:05 |  |
| * Fri | * member | * 742.0458 | * 00:12:22 |  |

* Both members and casual riders have longer rides during weekends
* Although members have low ride length on average but have a lot more rides than casual riders.
* Members take most rides during weekdays while most casual riders take more rides during weekends.
* Most casula riders use bikes for fun on weekends
* While it looks like member use bike for go to work or school that’s why use bikes for shorter and period but more regularlry
* Rides are low during winter seasons

|  |
| --- |
|  |
| **week\_day**  <ord> | **member\_casual**  <chr> | **number\_of\_rides**  <int> |  |  |
| Wed | member | 504612 |  |  |
| Tue | member | 490007 |  |  |
| Thu | member | 482499 |  |  |
| Mon | member | 444258 |  |  |
| Fri | member | 426586 |  |  |
| Sat | member | 388473 |  |  |
| Sat | casual | 360791 |  |  |
| Sun | member | 343149 |  |  |
| Sun | casual | 299400 |  |  |
| Fri | casual | 249252 |  |  |

* Most people use Classic bikes , usage of electric bikes is almost half.
* Docked bikes are only used by the casual riders

| * **rideable\_type** * <chr> | * **member\_casual** * <chr> | * **number\_of\_rides** * <int> |  |  |
| --- | --- | --- | --- | --- |
| * classic\_bike | * member | * 2097986 |  |  |
| * classic\_bike | * casual | * 1117974 |  |  |
| * electric\_bike | * member | * 981598 |  |  |
| * electric\_bike | * casual | * 584349 |  |  |
| * docked\_bike | * casual | * 33303 |  |  |

**Suggestions**

Should offer new membership package that caters for weekends rides because most casual riders take rides on weekends

**Business Task**

The goal is to analyze the differences in how annual members and casual riders use Cyclistic bikes to identify patterns that can inform marketing strategies. The objective is to convert casual riders into annual members by adjusting the membership plans based on their riding behavior. This analysis will provide insights to the stakeholders for targeted marketing campaigns to increase annual memberships.

**Key Stakeholders**

* Lily Moreno: Director of Marketing
* Cyclistic Marketing Analytics Team
* Cyclistic Executive Team

**Insights and Business Decisions**

1. Targeted Marketing Campaigns: Design campaigns that appeal to casual riders, especially those who ride on weekends.
2. Pricing Strategy: Consider introducing a membership package tailored for weekend riders.
3. Digital Media Tactics: Use data-driven insights to target casual riders with personalized offers.

**Dataset Description**

The dataset includes 12 months of bike trip data from 07-2023 to 07-2024, containing the following key columns:

* **ride\_id:** Unique identifier for each ride.
* **rideable\_type:** Type of bike (e.g., electric, classic).
* **started\_at:** Start date and time of the ride.
* **ended\_at:** End date and time of the ride.
* **member\_casual:** Indicates whether the rider is a member or a casual user.
* Plus additional columns related to station details and coordinates.

**Data Analysis Process**

* **Data Preparation:** Combined datasets, removed null values, created new columns for ride length and day of the week, and filtered out invalid data.
* **Ride Length Analysis:**
  + Mean Ride Length: Approximately 10 minutes.
  + Min/Max Ride Length: Ranges from 0 to 114 hours.
  + **Members vs. Casual Riders:**
    - Members account for 64% of total rides, while casual riders account for 36%.
    - Casual riders have a longer average ride time (24 minutes 14 seconds) compared to members (12 minutes 36 seconds).
* **Weekly Patterns:**
  + Most Popular Day: Saturday, especially among casual riders.
  + Usage Trends: Casual riders tend to use bikes more on weekends for leisure, while members use them more during weekdays, possibly for commuting.
  + Seasonality: Ride numbers drop during winter.

**Key Findings**

1. Ride Distribution: Members have shorter but more frequent rides, while casual riders have longer rides, especially on weekends.
2. Bike Type Preference: Classic bikes are the most popular, with electric bikes being used by about half as many riders. Docked bikes are used exclusively by casual riders.
3. Seasonal Influence: Ride frequency decreases significantly during winter.

**Recommendations**

1. Introduce a Weekend-Focused Membership Package: To attract casual riders who primarily ride on weekends.
2. Targeted Marketing Campaigns: Focus on weekends to convert casual riders to members.
3. Promote Electric Bikes: Increase the promotion of electric bikes to boost their usage among both casual and member riders.