

# Cyclistic Bike Share

DAVID NOGUEIRA

# Index

1 - Introduction.....	2
1.1 - Objective.....	2
1.2 – Analysis Process .....	2
2 - Ask .....	3
2.1 – Guiding Questions.....	3
2.2 – Stakeholders.....	3
3 - Prepare .....	4
3.1 – Data Source and Privacy.....	4
3.2 – Data Organization and Limitations.....	4
3.3 – Tools Used .....	4
4 - Process.....	5
4.1 – Data Import and Combination .....	5
4.2 – Null Values and Duplicates Analysis.....	5
4.3 – Creation of New Columns .....	5
4.4 – Outlier Identification and Treatment .....	5
5 - Analyze .....	6
5.1 – Summary of Trips and User Types.....	6
5.2 – Analysis of Trip Numbers .....	7
5.3 – Analysis of Average Trip Duration .....	8
6 - Share.....	9
7 - Act.....	10
7.1 – Introduction of Monthly and Seasonal Plan .....	10
7.2 – Targeted Engagement Strategies .....	10
7.3 – Incentives and Loyalty Programs .....	10
7.4 – Location-Based Marketing Campaigns.....	10
7.5 – Strategies to Maximize Bike Usage .....	10
8 – Conclusion.....	11

# 1 - Introduction

This case study is part of the **final project for the Google Data Analytics Professional Certificate**. Unlike the course, which uses tools like R and Tableau, **I used exclusively Power BI** for the entire analysis, from data preparation to insight visualization.

## 1.1 - Objective

**Cyclistic** is a bike-sharing service in **Chicago**, with two types of customers:

- **Casual riders** (single-ride or day passes);
- **Annual members** (subscription-based).

The goal of this project is to **analyze user behavior** and provide insights to **convert more casual riders into annual members**, maximizing the company's revenue.

## 1.2 – Analysis Process

- **Ask** – Define the business problem
- **Prepare** – Collect and organize data
- **Process** – Clean and transform data
- **Analyze** – Explore patterns and trends
- **Share** – Create visualizations and reports
- **Act** – Present strategic recommendations

## 2 - Ask

This phase involves **defining the issue to be solved**, identifying stakeholders, and understanding their expectations for the project. The goal is to analyze how Cyclistic users behave and provide insights that will help shape marketing strategies aimed at increasing membership conversions.

### 2.1 – Guiding Questions

Three key questions will guide the analysis:

- **How do annual members and casual riders use Cyclistic bikes differently?**
- **Why would casual riders buy Cyclistic annual memberships?**
- **How can Cyclistic use digital media to influence casual riders to become members?**

### 2.2 – Stakeholders

- **Lily Moreno:** Director of Marketing, responsible for developing campaigns and initiatives to promote the bike-share program.
- **Marketing Analytics Team:** Collects, analyzes, and reports data to guide Cyclistic's marketing strategies.
- **Executive Team:** Evaluates and approves recommendations based on data-driven insights.

## 3 - Prepare

### 3.1 – Data Source and Privacy

The data was provided by **Motivate International Inc.** and can be accessed [here](#) under this [license](#). Riders' personal information has been removed to ensure privacy.

### 3.2 – Data Organization and Limitations

The dataset includes records from **January to December 2024**, containing thousands of monthly trip details. During data processing, missing values were identified and handled accordingly.

### 3.3 – Tools Used

- **Data Processing & Cleaning:** Power BI
- **Visualization:** Power BI

## 4 - Process

### 4.1 – Data Import and Combination

- The **CSV files** containing the **monthly trip data for 2024** were imported into Power BI.
- The **end\_station\_id** column was incorrectly recognized as a **number** instead of **text**, requiring adjustments.
- Tables were renamed for readability, such as **2024\_01\_January**, **2024\_02\_February**, etc.
- All **12 tables** were merged into a single dataset: **2024\_All\_Months\_No\_Clean**, totaling **5860538 rows**.

### 4.2 – Null Values and Duplicates Analysis

- A **summary table** was created to identify missing and blank values.
- Columns with excessive missing data were removed:
  - **start\_station\_name**, **start\_station\_id**, **end\_station\_name**, **end\_station\_id**, **end\_lat**, **end\_lng**
- This resulted in a refined dataset: **2024\_All\_Months\_First\_Clean**.

### 4.3 – Creation of New Columns

- Several new columns were created to enhance the dataset:
  - **Duration (Minutes)**: Difference between *ended\_at* and *started\_at*.
  - **Duration (Minutes:Seconds)**: Conversion of duration into minutes and seconds.
  - **Day of Week (Number)**: Numeric representation of the day.
  - **Day of Week (Name)**: Name of the weekday.
  - **Month (Number)**: Numeric representation of the month.
  - **Month (Name)**: Name of the month.
- This led to a new dataset: **2024\_All\_Months\_First\_Calculations**.

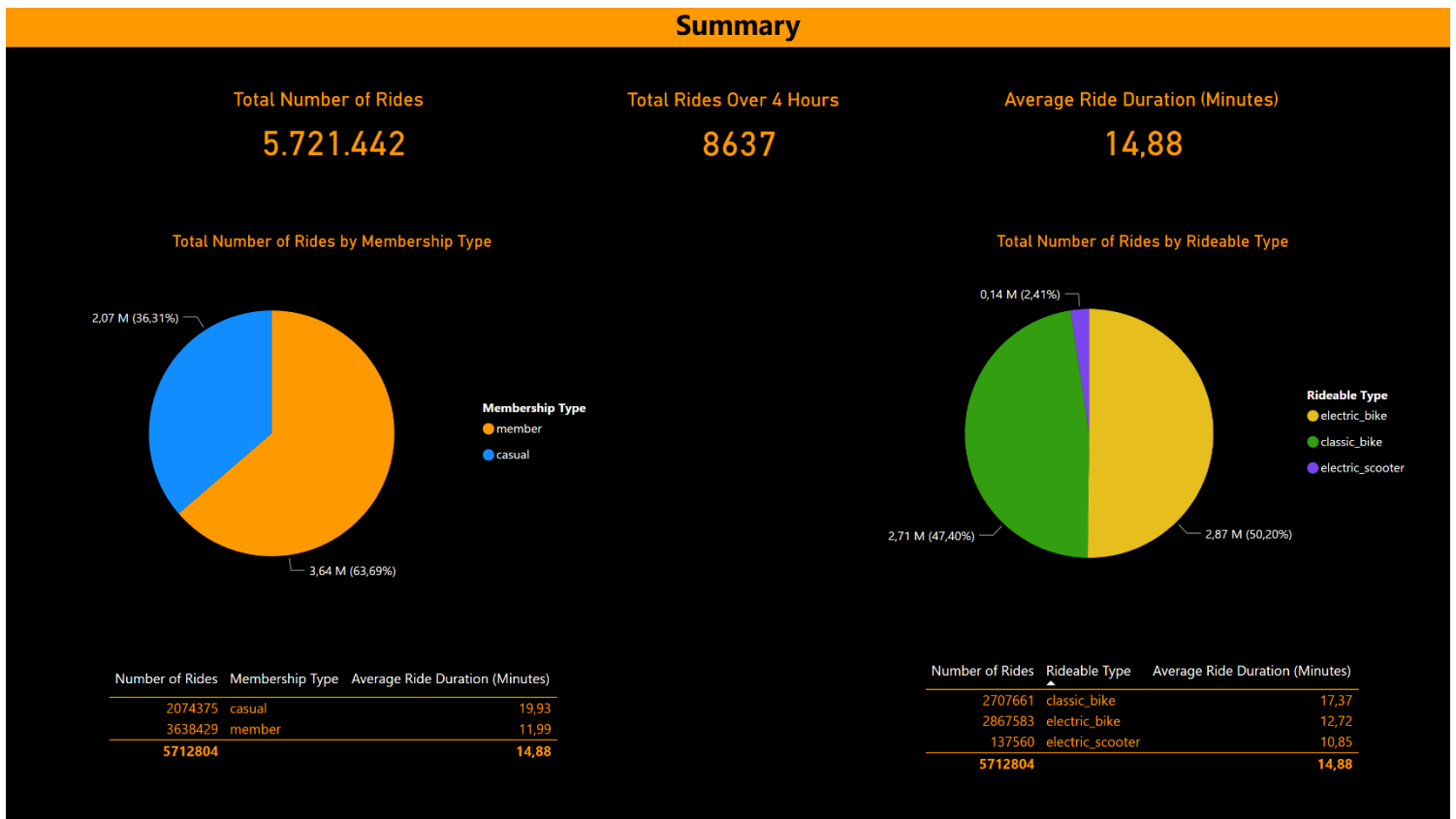
### 4.4 – Outlier Identification and Treatment

- A **Maximum\_Minimum\_Duration** table was created to detect extreme values.
- Some trips had **negative durations** or lasted **over 24 hours**, indicating errors.
- The dataset was filtered to retain only trips lasting **between 1 minute and 24 hours**, resulting in the final dataset: **2024\_All\_Months\_Final**.
- The final dataset contains **5721442 rows**.

At this stage, the data is **cleaned, structured, and ready** for further analysis and visualization in Power BI.

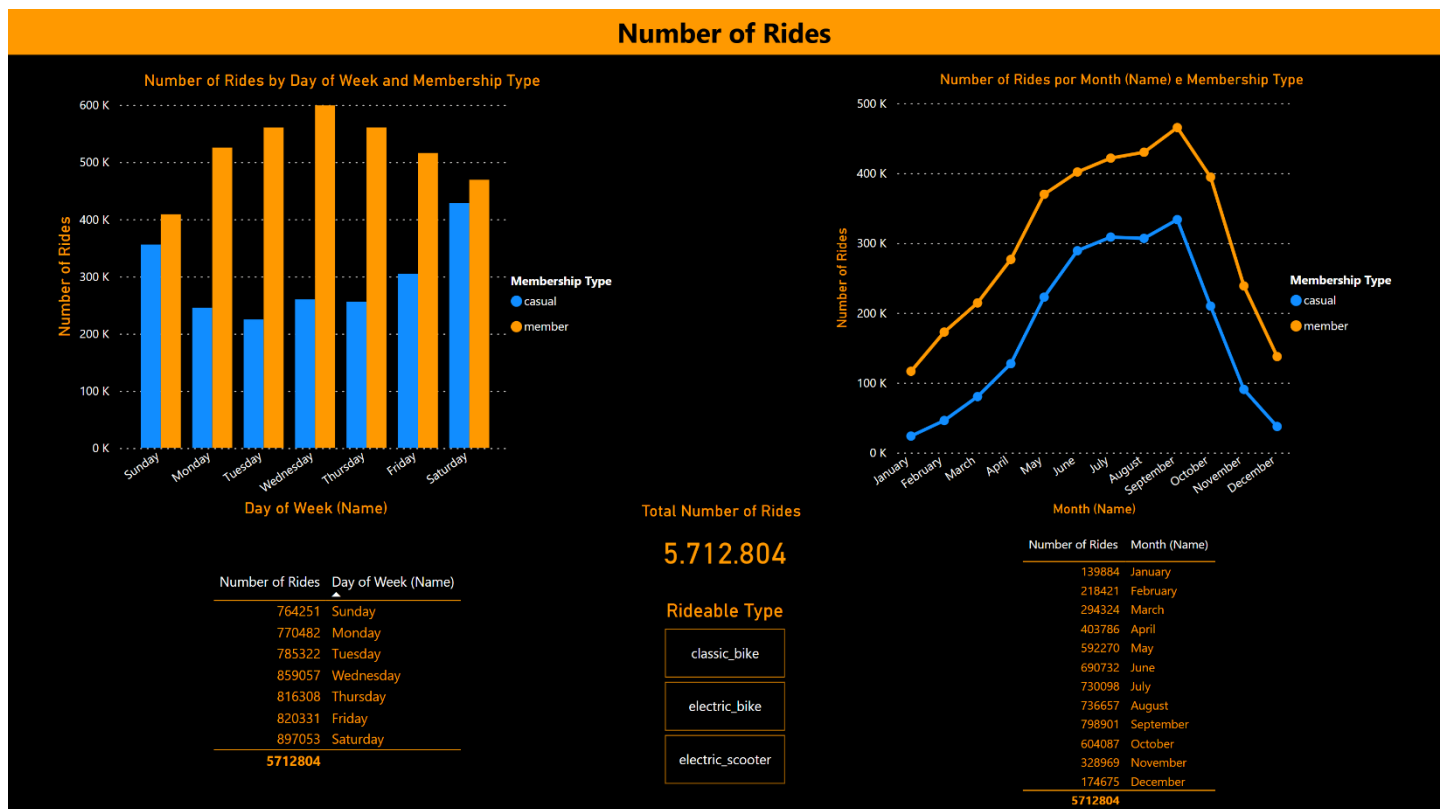
## 5 - Analyze

### 5.1 – Summary of Trips and User Types



- **Objective:** Analyze the behavior of casual passengers and annual members, and understand their vehicle preferences.
- **Data:**
  - Total Trips: 5721442 trips with a duration between 1 minute and 24 hours.
  - Long Trips (4h-24h): 8637 trips, which are rare events.
  - After this observation, a filter was applied to consider only trips with a duration of 1 minute to 4 hours.
  - Average Trip Duration: 14.88 minutes.
- **Key Insights:**
  - **Membership Type:**
    - 64% of users are annual members.
    - Members take shorter trips, with an average duration of 11,99 minutes.
    - Casuals have a longer average duration: 19,93 minutes.
  - **Vehicle Type:**
    - 50,2% of trips are made using electric bikes.
    - Classic bikes have the highest average trip duration, with 17,37 minutes.

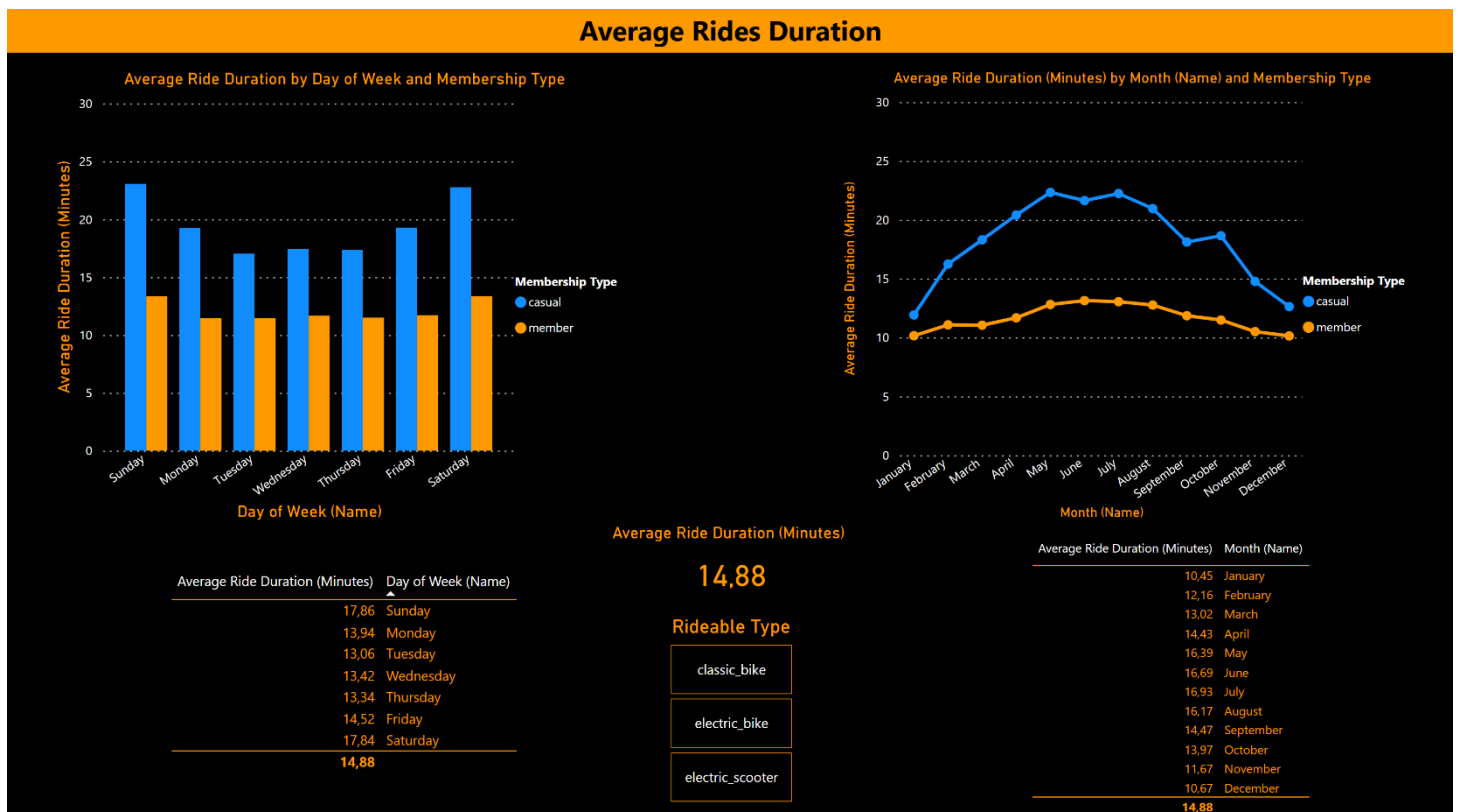
## 5.2 – Analysis of Trip Numbers



- **Objective:** Identify patterns in the number of trips throughout the week and year, segmented by user type.
- **Data:**
  - Total Trips: 5712804 trips with a duration between 1 minute and 4 hours.
- **Insights:**
  - **Weekly Distribution:**
    - Members make more trips than casuals every day of the week.
    - Saturday is the day with the highest number of trips for casuals.
    - Wednesday is the day with the highest number of trips for members.
    - In total, Saturday is the day with the highest number of trips (897053 trips), while Tuesday has the lowest (764251 trips).
  - **Monthly Distribution:**
    - The number of trips increases from January to September, reaching a peak in September (798901 trips in total).
    - After September, there is a significant drop in trips, with the lowest number recorded in January (139884 trips).



## 5.3 – Analysis of Average Trip Duration



- **Objective:** Analyze the average trip duration by day of the week and month, segmented by user type.
- **Data:**
  - Average Trip Duration: 14,88 minutes (considering trips from 1 minute to 4 hours).
- **Insights:**
  - **Weekly Distribution:**
    - Casuals take longer trips, with peaks on the weekend.
    - Members have more stable trip durations, with a slight increase on weekends.
    - Highest average duration: Sunday (17,86 minutes).
    - Lowest average duration: Tuesday (13,06 minutes).
  - **Monthly Distribution:**
    - Casuals show strong variation, with trip duration peaking in May. From April to August, they travel with an average of over 20 minutes.
    - Members maintain a stable duration between 10 and 13 minutes.
    - Month with highest average duration: July.
    - Month with lowest average duration: January.

## 6 - Share

The findings were shared and communicated with the relevant stakeholders.

The sharing phase of this project was completed using Power BI for visualizations and a detailed PDF report containing all the analyses conducted.

- [Click here to download the PDF report.](#)

This ensures that the key dashboards and insights are accessible for review and decision-making.

## 7 - Act

After analyzing the data and identifying key user behavior patterns, it is essential to define strategies to optimize service usage and increase the conversion of casual riders into annual members.

### 7.1 – Introduction of Monthly and Seasonal Plan

- Currently, the company offers daily and annual passes, but data shows that casual riders use the service more frequently during warmer months (April to September).
- A highly effective strategy would be to introduce a monthly or seasonal plan, encouraging these users to subscribe more consistently by offering progressive discounts that make the annual membership more attractive.

### 7.2 – Targeted Engagement Strategies

- Launch marketing campaigns highlighting the benefits of the annual membership, such as cost-effectiveness and access to exclusive promotions.
- Offer a one-month free trial during winter, when fewer trips are made, to encourage casual riders to experience the benefits of membership.

### 7.3 – Incentives and Loyalty Programs

- Implement a loyalty points system, where each ride accumulates points that can be redeemed for discounts on an annual subscription.
- Introduce a user ranking system, rewarding frequent riders with benefits such as free rides or exclusive perks.

### 7.4 – Location-Based Marketing Campaigns

- Use data from the most popular stations among casual riders to install billboards and posters, promoting annual memberships.
- Launch geo-targeted digital marketing campaigns to ensure advertisements reach users who frequently use the service in these locations.

### 7.5 – Strategies to Maximize Bike Usage

- Since classic bikes have the highest average trip duration (17.37 minutes), offering exclusive benefits for using them could help balance fleet utilization.
- Adjust bike distribution based on usage patterns throughout the week, ensuring sufficient availability during peak periods.

## 8 – Conclusion

- The analysis of Cyclistic's data revealed clear patterns in user behavior: members use bikes consistently throughout the week, while casual riders concentrate their usage on weekends and warmer months. The preference for electric bikes and the differences in trip duration between the two groups highlight the need for targeted strategies.
- To maximize the conversion of casual riders into members, we recommend seasonal promotional campaigns, weekend-specific incentives, and improvements to the experience of classic bikes. These initiatives can drive membership growth and optimize company revenue.