

# Cyclistic Project Final Report

**Case study:** How does a bike-share navigate speedy success?

Hi! This report is part of the **Google Data Analytics Professional Certificate** Capstone project on Coursera. Specifically, it focuses on **Case Study 1: Cyclistic Bike-Share Analysis**, where the goal is to analyze user behavior and provide data-driven recommendations to convert casual riders into annual members.

To achieve this, I will use **SQL** in the **Bigquery platform** for data cleaning, processing, and analysis. Additionally, I will use **Tableau Public** to keep analysing and to create visualizations that highlight key trends and insights. The dataset used in this project is provided by the course and contains real-world bike-sharing trip data.

This case study follows the Google Data Analytics framework: Ask, Prepare, Process, Analyze, Share, and Act, ensuring a structured approach to solving the business problem. The final insights and recommendations aim to help Cyclistic improve its membership strategy and overall service offering.

## Scenario

You are a junior data analyst working on the marketing analyst team at Cyclistic, a bike-share company in Chicago. The director of marketing believes the company's future success depends on maximizing the number of annual memberships. Therefore, your team wants to understand how casual riders and annual members use Cyclistic bikes differently. From these insights, your team will design a new marketing strategy to convert casual riders into annual members. But first, Cyclistic executives must approve your recommendations, so they must be backed up with compelling data insights and professional data visualizations.

## About the company

In 2016, Cyclistic launched a successful bike-share offering. Since then, the program has grown to a fleet of 5,824 bicycles that are geotracked and locked into a network of 692 stations across Chicago. The bikes can be unlocked from one station and returned to any other station in the system anytime.

Until now, Cyclistic's marketing strategy relied on building general awareness and appealing to broad consumer segments. One approach that helped make these things possible was the flexibility of its pricing plans: single-ride passes, full-day passes, and annual memberships. Customers who purchase single-ride or full-day passes are referred to as casual riders. Customers who purchase annual memberships are Cyclistic members.

Cyclistic's finance analysts have concluded that annual members are much more profitable than casual riders. Although the pricing flexibility helps Cyclistic attract more customers, Moreno believes that maximizing the number of annual members will be key to future growth. Rather than creating a marketing campaign that targets all-new customers, Moreno believes there is a solid

opportunity to convert casual riders into members. She notes that casual riders are already aware of the Cyclistic program and have chosen Cyclistic for their mobility needs.

Moreno has set a clear goal: Design marketing strategies aimed at converting casual riders into annual members. In order to do that, however, the team needs to better understand how annual members and casual riders differ, why casual riders would buy a membership, and how digital media could affect their marketing tactics. Moreno and her team are interested in analyzing the Cyclistic historical bike trip data to identify trends.

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

## 1. Ask - Business Task

The problem we seek to solve is: **How do annual members and casual riders use Cyclistic differently?**

Through our knowledge in data analysis, we will be in charge of preparing, cleaning and analyzing the data from the last year (from October 2023 to September 2024) to identify the differences in the use of the service between these two types of clients. With this information, we will be able to make recommendations based on solid data that will help business leaders and the marketing team make more informed business decisions. The ultimate goal is to design effective strategies that incentivize casual users to become annual members of Cyclistic.

## 2. Prepare - Data Sources

The data used in this project was sourced from Coursera's provided repository. <https://divvy-tripdata.s3.amazonaws.com/index.html>. This data adheres to the ROCCC acronym, ensuring its credibility. It is **Reliable** as it is provided under a license from Motivate International Inc. It is **Original** because the company directly collects the data based on their agreement. It is **Comprehensive** because it includes all the necessary information to address the question at hand. The data is **Current** since the most up-to-date data is being downloaded during the course of the exercise. Finally, it is **Cited** because the source is properly credited throughout the analysis.

The dataset includes monthly data files from October 2023 to September 2024, detailing ride information such as:

- **ride\_id**: Unique identifier for each ride.
- **rideable\_type**: Type of bike (classic, electric, scooter).
- **started\_at**: Start timestamp.
- **ended\_at**: End timestamp.
- **start\_station\_name/id and end\_station\_name/id**: Station details.
- **start\_lat/lng and end\_lat/lng**: Coordinates.
- **member\_casual**: User type (annual member or casual).

The dataset contains over 5,697,108 records and 13 columns, which were cleaned and processed in BigQuery to ensure readiness for analysis.

### **3. Data Cleaning and Manipulation**

In the **process** phase of the data analysis, I initially attempted to use Google Sheets to handle the data. However, due to the large size of the files and the numerous rows, the program kept freezing when trying to merge them into a single sheet. Given these limitations, I ultimately decided to switch to SQL in BigQuery. BigQuery is better suited for handling large volumes of data, allowing for more efficient processing and querying without the performance issues I encountered in Google Sheets. This tool provided the scalability and speed necessary to manage, clean and analyze the data effectively.

The cleaning process was documented extensively (other document) and included the following key steps:

#### **Handling Missing Values:**

- 27.77% of the dataset (approximately 1.6 million rows) had missing station data, primarily associated with electric bikes and scooters. These records were retained for analysis to avoid bias, as they represent a significant portion of the data.

#### **Eliminating Anomalies:**

- Removed 21,604 rides lasting over 3 hours, representing 0.37% of the total data, as these were likely errors.
- Removed 294 records where `ended_at` occurred before `started_at`, representing less than 0.01% of the total dataset.
- Deleted 49 records with out-of-range coordinates, a negligible 0.0008% of the dataset.

#### **Standardizing Data:**

- Rounded coordinates to six decimals for consistency.
- Added calculated fields such as `ride_length` (duration of each ride) and `day_of_week` (day name).
- Ensured all timestamps were sorted in ascending order.

#### **Managing Duplicates:**

- Identified and removed 149 duplicate records, ensuring data accuracy. These represented 0.0026% of the total data.

#### **Final Dataset:**

- Created monthly and consolidated tables for efficient analysis and visualization, containing 5,696,910 rows after cleaning.

### **4. Summary of Analysis**

The analysis process was documented extensively (other document). Key findings from the analysis include:

### 1. User trends:

**Annual members** completed a higher number of trips throughout the year, totaling 3,645,605, which accounts for **63.99%** of the total. This suggests a more frequent and consistent use of the service, likely as a **primary mode of transportation**.

**Casual members** recorded a significantly longer average trip duration, exceeding that of annual members by **54.45%**. This indicates a more **recreational or tourism-oriented use of the service**.

### 2. Seasonal Decline During Winter Months:

A drop in service usage was observed during the winter months. Between December and February, **casual members** completed a total of 119,623 trips, representing only **2.10%** of the annual trips. In contrast, **annual members completed** 457,053 trips in the same period, accounting for **8.02%** of **total trips for the year**.

### 3. Successful Launch of the Electric Scooter:

In its launch month, the **electric scooter** received an **excellent response**, particularly among casual members, who completed 81,406 trips, representing **22.43%** of their **total trips for the month**. In comparison, **annual members** recorded 56,130 trips, making up **12.07%** of **their total trips for the month**.

### 4. Increased Summer Usage by Casual Members:

During the summer, **casual members** significantly **increased** both the **total number of trips** and their **average ride duration**, particularly on weekends. Compared to the rest of the year, the number of trips taken by casual members rose by **85.25% for classic bikes** and **38.48% for electric bikes**. Additionally, the **average trip duration increased by 22.52% for classic bikes** and **29.06% for electric bikes** on Saturdays and Sundays, reinforcing the seasonal recreational trend.

## 5. Supporting Visualizations

Visualizations were created in Tableau, offering insights into:

- **Monthly Ride Trends:** Showcased the decline in winter and growth from spring.
- **User Comparisons:** Highlighted ride duration and frequency differences between user types.
- **Vehicle Preferences:** Revealed the adoption patterns of electric scooters on their launch month.
- **Seasonal Insights:** Illustrated the increase in casual usage during summer weekends.

[Click here to see the dashboard on Tableau](#)

## 6. Recommendations

1. **Encouraging the conversion of casual users to annual members:**

**Finding: Annual members** make more trips (**63,99% of the total**), indicating frequent and consistent use of the service.

**Recommendation:**

- Implement annual membership campaigns highlighting the cost savings for casual users who make long and frequent trips.
- Offer seasonal discounts for new subscriptions during high-demand months (summer).

**2. Summer-specific promotions**

**Finding:** Casual users increase their service usage on Saturdays and Sundays during the summer months, with longer trip durations.

**Recommendation:**

- Create unlimited weekend ride packages or one-day passes at attractive prices.

**3. Campaigns for the adoption of electric scooters**

**Finding:** During its launch month, electric scooters were well received, especially by **casual users**, exceeding **annual members** by **10.36%**.

**Recommendation:**

- Run "free trial" campaigns or discount vouchers for scooters to continue encouraging adoption.
- Promote electric scooters as a recreational and sustainable transport option for casual users.

**4. Strategy to address seasonal decline in winter**

**Finding:** A decrease in service usage was detected during the winter months, especially among casual users.

**Recommendation:**

- Introduce special winter offers, such as discounts for frequent trips or temporary memberships.
- Highlight the benefits of electric bikes in cold conditions, such as greater comfort and less physical effort.