**Cyclistic Bike Share Case Study: How Do Annual Members and Casual Riders Use Cyclistic Bikes Differently?**

**Business Task**

Cyclistic, a bike share program in Chicago, aims to increase annual memberships. The marketing team believes that converting casual riders into annual members will significantly contribute to long term profitability. To support this initiative, my task is to analyze how annual members and casual riders use Cyclistic bikes differently and deliver actionable insights through professional data analysis and visualization.

**Data Sources**

I used publicly available historical trip data from the Divvy bike share system for the full calendar year of 2023. The dataset includes trip level data such as:

* ride\_id
* rideable\_type
* started\_at and ended\_at
* start and end station
* member\_casual (user type)

The data was sourced from [Divvy Trip Data](https://divvy-tripdata.s3.amazonaws.com/index.html) and spans January 2023 to December 2023.

**Tools Used**

* **Excel**: Initial data audit, cleaning, and exploratory pivot tables
* **SQLite**: Consolidated data into a single table (all\_trips\_2023) and performed SQL queries for aggregations
* **Tableau**: Created interactive dashboards and geospatial visualizations
* **R & RStudio**: Full data cleaning, transformation, analysis, and visualization

**Data Preparation & Cleaning**

1. **Initial Audit:** Reviewed the available months and determined that 2023 had more complete and consistent data compared to 2024.
2. **Data Consolidation:** CSV files for each month were converted to Excel workbooks and initially reviewed in Excel.
3. **Cleaning Steps:**
   * Added columns for ride\_length and day\_of\_week
   * Removed rows with missing values (e.g., empty station names or times)
   * Verified consistent datetime formats and station names

**Key Analyses Conducted (SQL)**

* Created ride\_length column (difference between ended\_at and started\_at)
* Converted ride\_length to seconds and minutes
* Added day\_of\_week and hour\_of\_day columns

**SQL Queries Performed:**

* Total rides by member type
* Average ride duration by member type
* Ride counts by day of the week
* Average ride duration by day and member type
* Ride counts by hour of day
* Most popular start stations

**Tableau Visualizations**

Created multiple dashboards to support storytelling:

* Total rides by user type
* Ride length distribution by user type
* Ride volume by day of the week and hour of day
* Top start stations (with map)
* Geospatial visuals styled with color coded member/casual types

**Data Preparation & Cleaning (In R)**

* Imported raw CSV files directly into R
* Combined 12 months of trip data into a single dataframe
* Parsed datetime columns (started\_at, ended\_at)
* Calculated ride\_length, day\_of\_week, and hour\_of\_day
* Removed outliers (<1 min and >24 hrs ride length)
* Standardized day\_of\_week as an ordered factor to ensure accurate visualization

**Key Analyses Conducted (In R)**

* Summary Statistics: Average, median ride length by user type
* Ride counts by day of week and hour of day
* Rideable Type Preference Analysis
* Top Start Stations
* Ride Length Distribution Analysis (Density Plots)
* T-Test performed to compare ride lengths between members and casuals

**Visualizations Created (ggplot2)**

* Total Rides by User Type (with data labels)
* Ride Length Distribution (0-60 mins)
* Rides by Day of Week (Faceted by user type)
* Rides by Hour of Day
* Top 10 Start Stations (Horizontal Bar Chart)
* Bike Type Preference by User Type
* Ride Trend Over Time

**Key Insights**

* Casual riders tend to take longer trips, especially on weekends.
* Annual members ride more frequently, particularly on weekdays and during commuting hours.
* Casual riders are more active during midday hours and weekends, suggesting leisure oriented use.
* Classic bikes are the most used, but casuals favor docked bikes more compared to members.
* Top start stations differ slightly by user type but cluster around central Chicago and popular leisure spots.

**Recommendations**

1. Target casual riders with weekend promotions that offer incentives for trying weekday commutes.
2. Promote annual memberships at high traffic leisure locations (based on top casual rider start stations).
3. Use digital marketing during midday and weekend hours, aligning with casual rider behavior patterns.

*Prepared by: Mario Feliciano* *Google Data Analytics Capstone Project - Cyclistic Bike-Share Analysis*