Cyclistic Bike-Share Case Study (Google Data Analytics Capstone – 2024)

Project Overview

This project analyzes Cyclistic's 2024 bike-share data to understand how members and casual riders use the service differently — and identify strategies to convert casual users into annual members.

The analysis follows the full data-analysis process: Ask → Prepare → Process → Analyze → Share → Act using R, Tableau, and Excel.

Objectives

- 1. Identify key differences in usage patterns between **members** and **casual riders**.
- 2. Determine when, where, and how each group rides most frequently.
- 3. Generate actionable insights to increase membership conversions and improve operational efficiency.

11 Target Audience

- Cyclistic Marketing Team: to design data-driven campaigns.
- Operations Managers: to plan fleet allocation and maintenance.
- Hiring Managers / Recruiters: to assess analytical workflow and storytelling ability.

Tools & Technologies

Tool	Purpose
R (v4.5.1)	Data cleaning, transformation, EDA
Excel	Initial structure inspection
Tableau	Data visualization and dashboard design
R Markdown	Project documentation

Tool	Purpose
GitHub	Version control and portfolio presentation

Project Workflow

A. Project Setup & Planning

- Created folder structure (data raw, data clean, scripts, docs, tableau dashboards).
- Defined main business question:

How do annual members and casual riders use Cyclistic bikes differently?

Identified success metric: Increase in casual-to-member conversion rate.

B. Data Acquisition

- Downloaded 12 monthly CSV files (Jan-Dec 2024) from Cyclistic's open dataset.
- Verified consistent column names across all files.
- Noted missing or blank values in station IDs and names.

C. Data Cleaning & Preparation (R)

- Combined all monthly files into one dataset (all trips 2024.csv).
- Parsed and standardized started_at and ended_at to POSIX datetime.
- Removed duplicate and invalid entries (e.g., ride length ≤ 0).
- Calculated new variable ride_length in minutes.
- Validated rows and column counts after merging: **5.86M rows** × **14 columns**.

D. Feature Engineering

Added derived columns for analysis:

day_of_week

- start_hour
- start_month
- season (Winter, Spring, Summer, Fall)
- weekday_or_weekend
- Saved final dataset as all_trips_features.csv.

E. Exploratory Data Analysis (EDA)

Performed using **R** and dplyr.

Key metrics generated:

Metric	Casual	Member
Total Rides	1.52M	2.69M
Avg Ride Length (min)	24.0	12.4
Median Ride Length (min)	13.4	8.8

Exported multiple summary files for Tableau visualization.

F. Data Visualization (Tableau)

Created 5 main visuals for storytelling:

- 1. Total Rides & Avg Duration by User Type
- 2. Rides by Day of Week
- 3. Monthly Ride Trends
- 4. Rides by Season
- 5. Rides by Day Type (Weekday vs Weekend)

Each visualization exported as high-resolution PNG and included in presentation slides.

G. Insights Summary

- Members: Consistent weekday commuters (peaks mid-week, stable usage).
- Casual Riders: Highly seasonal and weekend-oriented (peaks in summer).
- Overall: 65–75% of total rides from members; 45% of rides happen in summer.

H. Business Recommendations

1. Marketing:

- o Launch "Weekend-to-Member" campaigns targeting casual users in peak months.
- Offer trial memberships or discounted passes.

2. Operations:

- Reallocate bikes to tourist areas during weekends.
- Schedule maintenance during low-usage periods (winter & early weekdays).

3. Customer Retention:

- Create **loyalty rewards** for frequent weekday riders.
- Develop seasonal promotions to sustain engagement.

I. Final Deliverables

- Tableau Dashboard (.twbx)
- Clean CSV datasets (data_clean/)
- I R Markdown documentation (.Rmd)
- Presentation slides (PowerPoint)
- Final project report (PDF)
- README.md (this file)

J. Key Learnings & Skills Demonstrated

- Data Wrangling & Cleaning: Using R (vroom, dplyr, lubridate).
- Exploratory Analysis: Aggregations, grouping, data validation.
- Visualization & Storytelling: Tableau charts & business interpretation.
- **Business Thinking:** Turning data patterns into marketing and operational recommendations.
- Reproducibility: Organized folder structure, clear scripts, documented workflow.

K. How to Run the Project

- 1. Clone the repository:
- 2. git clone https://[github.com/Ahmedmagdy1999/cyclistic_project.git
- 3. cd cyclistic_project](https://github.com/Ahmedmagdy1999/Cyclistic-Bike-Share-Case-Study-Google-Data-Analytics-Capstone-2024-)

4.

My LinkedIn profile: https://linkedin.com/in/ahmedmagdyyahia