📘 Capstone Portfolio: Bike-Share Behavior Analysis

Project Title: *Understanding Rider Behavior in Urban Bike-Share Systems*

Certificate: Google Data Analytics Professional Certificate

Analyst: **Ali Muhammad**

I'm a data analyst with a master's degree in Urdu and a passion for clear, impactful communication. I specialize in Tableau, R, Excel, and SQL, and I'm expanding my toolkit with Python and Power BI. I design dashboards that balance usability, aesthetics, and business relevance. My work blends analytical rigor with creative strategy to make data accessible and actionable. I've explored rider behavior, fraud detection, and gaming analytics in my portfolio projects. I'm methodical in cleaning, transforming, and validating data at every step. Accessibility matters to me—I use high-contrast visuals and direct labeling for clarity. I thrive on stakeholder engagement and tailor insights to drive real-world decisions. I'm focused on automation, predictive modeling, and multi-platform dashboard delivery. Above all, I bring resilience, curiosity, and strategic thinking to every data challenge.

🔍 Ask Phase

Guiding Question:

How do casual and member riders differ in their usage patterns across time, location, and ride duration? This question supports business decisions around station placement, inventory management, and targeted marketing.

📦 Prepare Phase

Data Source:

- Public bike-share dataset from Divvy (Chicago), covering ride-level data including timestamps, station names, rider type, and rideable type.

Tools Used:

- **Excel**: Initial inspection, column profiling, and basic filtering
- **R**: Data cleaning, transformation, and binning
- **Tableau**: Visualization and dashboard creation

🖌️ Process Phase

Steps Taken:

- Removed nulls and duplicates
- Standardized column names and formats
- Converted timestamps to day-of-week and duration metrics
- Created new variables: ride duration (in minutes), rider type flags, and station usage counts
- Validated geographic coordinates for mapping

📊 Analyze Phase

Exploratory Analysis Included:

- Weekly ride volume trends
- Ride duration distribution by rider type
- Top start and end stations
- Geographic clustering of ride starts

📈 Visualize Phase

Four key visualizations were developed:

1. **Weekly Ride Volume by Rider Type**
    - Casual riders peak on weekends
    - Members ride consistently across weekdays
2. **Ride Duration Distribution**
    - Casual rides are longer and more variable
    - Member rides are shorter and more uniform
3. **Top 10 Start Stations by Rider Type**
    - High-traffic stations identified for both rider types
    - Strategic overlap and divergence observed
4. **Top 10 End Stations by Rider Type**
    - Reinforces start station trends
    - Useful for rebalancing and predictive stocking

🧠 Model Phase (Lightweight)

While no predictive model was built, the analysis included:

- **Rule-based segmentation** of rider types
- **Binning** of ride durations for histogram analysis
- **Station ranking logic** based on ride counts

📤 Share Phase

The dashboard was designed for clarity, accessibility, and stakeholder impact:

- Color-coded by rider type
- Labeled with exact ride counts
- Structured for minimal cognitive load
- Ready for sharing via Tableau Public or PDF export

🧠 Reflection

This project demonstrates my ability to:

- Ask meaningful business questions
- Clean and structure raw data
- Analyze patterns using R and Tableau
- Communicate insights through compelling visuals

It also reflects my commitment to accessibility, user-centered design, and strategic storytelling—skills I'll carry forward into future analytics roles.