Cyclistic Bike-Share Case Study – Data Workflow

1. Introduction

This project focuses on analyzing one year of Cyclistic bike-share trip data to uncover differences between **casual riders** and **annual members**.
The goal was to clean, transform, and visualize the data to generate insights that can help improve marketing and user engagement strategies.

---

2. Dataset Overview

- **Files used**: 12 CSV files (Jan–Dec)

- **Source**: Cyclistic's open data (Kaggle or official repo)

- **Key columns**:
  ride_id, rideable_type, started_at, ended_at,
  start_station_name, end_station_name, member_casual

All files had consistent columns but required transformation before analysis.

---

3. Data Cleaning Steps (Detailed)

Combining Files in Excel:

- Opened Power Query in Excel

- Used **"Get Data from Folder"** to load all 12 CSVs

- Combined them using "Combine & Transform"

Cleaning in Power Query:

- Removed nulls in important columns like ride_id, started_at, ended_at

- Removed duplicates on ride_id

- Standardized datetime format for start and end time

Added New Columns:

- ride_length_minutes = ride duration in minutes

- day_of_week_name = weekday of started_at

- hour_of_day = hour from started_at

- month_name = full month from started_at

Filtered Out:

- Negative durations

- Rides longer than 24 hours (outliers)

---

4. Difficulties & How I Solved Them

Excel Row Limit:

- The combined file had nearly **4 million rows**, so I loaded it into the **Data Model**, not a worksheet.

Power BI Import Issue:

- Power BI couldn't detect any tables from Excel

- Got errors like: "The import 'Transform File' matches no exports"

Real Fix:

1. Copied the **M code** from Excel's Power Query → Advanced Editor

2. Cleaned it by removing references to Transform File / Sample File

3. Replaced with a direct folder path load

4. Pasted into Power BI using **Blank Query > Advanced Editor**

5. Clicked **Close & Apply** → All 3.96M rows loaded!

---

5. Dashboard Development Process (2 Pages)

Page 1 – Executive Summary

- Total Rides (KPI)

- Casual vs Member distribution

- Bike Type Usage

- Weekly Usage Pattern

- Filters: Month, Day, Rider Type, Bike Type

Page 2 – Behavioural Insights

- Avg Ride Duration by Member Type

- Monthly Ride Trends

- Hourly Ride Frequency

- Peak Hour Pattern

- Insight box with rider behavior notes

---

6. Final Insights

Visuals revealed key patterns like:

- Casual riders take longer trips but ride less often

- Members ride more on weekdays, especially commute hours

- Electric bikes were more common among casual users

- Activity spikes in warmer months for casuals