Google Data Analytics Capstone Project

Cyclistic Bikeshare Dataset

Solène Catella

December 2024

PART 1: ASK

The business task of this project is to analyze and optimize the usage patterns of both casual and member riders in order to provide actionable insights for enhancing Divvy's bike-sharing service. The focus will be on:

- **Increasing User Satisfaction:** Identifying ways to improve the rider experience for both casual and member users.
- **Improving Fleet Distribution:** Ensuring the bikes are distributed effectively across stations based on demand from different rider segments.
- **Boosting Rider Engagement:** Identifying strategies to increase rider usage and interaction with the service, particularly among casual riders.
- **Converting Casual Riders into Annual Members:** Converting casual riders into annual members, as annual memberships offer a more predictable revenue stream compared to casual rides, which generate less revenue per trip. This can be achieved by analyzing casual rider behavior and implementing targeted marketing strategies that encourage long-term commitment.

PART 2: PREPARE

Data Source

The data is first-party data, publicly available and provided by Motivate International Inc., which operates the City of Chicago's bicycle-sharing service. The bike-sharing data was collected from 2013 until now. However, for this project, only the data from July 2023 to June 2024 will be used.

Data Credibility Test

The data is:

- **Reliable:** The data used is accurate, complete, and unbiased, ensuring it is suitable for analysis.
- **Original:** This is first-party data provided by the company requesting the analysis.
- **Comprehensive:** The dataset includes all necessary information to address the questions from various stakeholders.
- **Current:** The data is continuously collected, with the analysis focusing on a specific time frame (July 2023 to June 2024).
- **Cited:** The data is properly licensed and credited to the company itself.

In other words... the data **ROCCC!**

PART 3: PROCESS

Data Cleaning

- **Handling Missing Values:** If any rows with missing essential values (e.g., *ride_id*) were found, they were excluded from analysis to ensure the accuracy of the results.
- **Outlier Detection & Removal:** Outliers were detected using measures of central tendency (mean, median) and measures of dispersion (standard deviation). Outlier filtering was based on IQR.
- **Fix Negative Ride Durations:** For negative *ride_length values, start_at* and *end_at times* were swapped.

Data Transformation

The start_time and end_time columns were converted from character format to datetime format to enable analysis based on ride durations and start/end times. Columns for ride duration, month, and day of ride were extracted. A column for ride distance using the distGeo() function was added. These actions transformed the dataset into a cleaner, more reliable form for further analysis.

PART 4: ANALYZE

Here is a summary of the analysis:

- **Peak Season:** Summer is the peak season for riders, with the top three months being August, July, and June.

- **Peak Usage Times:** Casual riders tend to use bikes more on weekends, while member riders predominantly use bikes on weekdays, with a peak on Wednesdays, suggesting a commuting pattern.
- **Riders' Bike Usage Patterns:** Casual riders and member riders travel similar distances, but casual riders tend to have longer trips, indicating a more leisure-focused use, while member riders likely use the bikes for practical commuting purposes.
- **Member Riders' Consistent Return Behavior:** Member riders are more likely to return bikes to the same station, supporting the idea of practical, commuting-focused use.
- **Bike Type Preferences:** Casual riders prefer electric bikes, while member riders are more inclined to choose classic bikes.

PART 5: SHARE

Refer to the graphs in the `viz` folder.

PART 6: ACT

This is top three recommendations based on my analysis:

- **Weekend & Summer Promotions for Casual Riders:** Since summer (particularly June, July, and August) is the peak season for bike usage, marketing efforts should be intensified during this period. Offer occasional membership discounts to new riders on summer, especially on weekends.
- **Highlight the Benefits of Longer Rides:** Casual riders tend to take longer trips, suggesting they use bikes for leisure. Subscription plans could be marketed with added benefits for longer rides, such as discounts for extended trips or exclusive access to premium bikes (e.g., electric bikes).
- **Partner with Local Businesses:** Partner with local businesses near the top 10 most frequently used stations by casual riders to launch targeted digital marketing campaigns. These campaigns should aim to attract local casual riders and engage regular visitors, encouraging them to subscribe.