DATA ANALYSIS ON CYCLISTIC BIKE SHARING SYSTEM DATA TO INCREASE MEMBERSHIPS

Amogh Borikar

Keywords: Data Analysis, Tableau, SQL, SQlite, CSV, Spreadsheets, problem solving.

1.Introduction and Problem Statement

Cyclistic is a bike-share company in Chicago. The director of marketing believes the company's future success depends on maximizing the number of annual memberships. The need is to understand how casual riders and annual members use Cyclistic bikes differently. From these insights, the team will design a new marketing strategy to convert casual riders into annual members. A bike-share program that features more than 5,800 bicycles and 600 docking stations. Cyclistic sets itself apart by also offering reclining bikes, hand tricycles, and cargo bikes, making bike-share more inclusive to people with disabilities and riders who can't use a standard two-wheeled bike. The majority of riders opt for traditional bikes and about 8% of riders use the assistive options. Cyclistic users are more likely to ride for leisure, but about 30% use them to commute to work each day.

2.Objective

The objective of this study is to find out the answers for following questions through data:
1. How do annual members and casual riders use Cyclistic bikes differently?
2. Why would casual riders buy Cyclistic annual memberships?
3. How can Cyclistic use digital media to influence casual riders to become members?
The need is to generate a business strategy that maximizes the number of memberships for the service.

3. Dataset Description

For this analysis, data is taken for 12 months starting from June 2021 to May 2022. Data for all the months is combined into a consolidated database structure. Database schema is shown below:
ride_id (TEXT): Unique Ride ID which is a primary key to the dataset
rideable_type (TEXT): Type of the bike used for the particular ride
started_at (INTEGER): Ride start date and time
ended_at (INTEGER): Ride end date and time
start_station_name (TEXT): Station name where the ride started
start_station_id (TEXT): Unique ID assigned to the start station
end_station_name (TEXT): Station name where the ride ended
end_station_id (TEXT): Unique ID assigned to the end station
start_lat (REAL): Geographical latitude of the start station location
start_lng (REAL): Geographical longitude of the start station location
end_lat (REAL): Geographical latitude of the end station location

end_lng (REAL): Geographical longitude of the end station location
member_casual (TEXT): Member type

4. Approach

1) The database is imported month by month in CSV format into DB Browser for SQLite.
2) Empty table is created with the database schema matching to the database.
3) Data from CSV files is appended to the empty table by using SQL commands.
4) Consolidated CSV file is exported from the DB Browser for SQLite.
5) CSV data is imported, cleaned and viewed into Tableau Public.
6) Calculated fields are created for the ride duration and analysis is performed based in the requirements.

5. Analysis and Interpretations

The ride comparison of Casual riders and members is performed by total number of rides taken and average trip duration. Results from figure 1 show that members take more rides than casual riders however, average trip duration for casual riders is more. This shows that the members use bikes for shorter durations for nearby locations, whereas casual riders take longer bike rides for long distances.

Comparison done throughout the timeline of 12 months as shown in figure 2 shows that the users take more rides from start of May to end of October having roughly similar average trip duration. Average trip duration throughout the year remains low for members as compared to casual riders. From the data insights, as we observe more number of rides and less average trip duration, members use the service for shorter distances and for regular use.

From figure 3, it is observed that members take nearly twice the rides in weekdays and casual riders take more rides on weekends. Members are predicted to take the rides for work commute, whereas casual riders are taking their trips for leisure. Average trip duration of casual riders is more than twice throughout the week, which supports the prediction of leisure rides taken by casual riders.

6. Key Takeaways:

1) Casual members use bikes for longer duration on each trip than members.
2) Number of rides taken throughout 12 months was roughly comparable for casual riders and members.
3) Casual members rent bikes more on weekend and members use the service more on weekdays.
4) Average trip duration throughout the week is more than twice for casual riders compared to members.

7. Business Recommendations:

1) As the trip duration is more for casual riders, it is proposed to give discounts for longer bike rides and offer memberships for longer rides.
2) As the number of riders are same in both categories, quarterly marketing strategy is recommended as there are variations observed with number or rides taken in different months.
3) Weekend discounts for longer bike trips are predicted to attract casual riders to purchase annual or monthly memberships as it fits their needs.
4) Incentives for longer bike duration and programs for leisure bike rides are recommended for better conversion of casual riders into members.

8. References:

1) Data reference:
    a. Data source : https://divvy-tripdata.s3.amazonaws.com/index.html
    b. License agreement (Motivate International Inc.): https://ride.divvybikes.com/data-license-agreement
2) Tableau Documentations : https://help.tableau.com/current/pro/desktop/en-us/gettingstarted_overview.htm
3) SQLite Documentations: https://www.sqlite.org/docs.html

9. Sources:

1) Tableau Visualization: https://public.tableau.com/views/CyclisticBikesharingAnalysis/Daysoftheweek?:language=en-US&:display_count=n&:origin=viz_share_link
2) SQL file and Tableau documentations: https://github.com/Amogh-Borikar/Cyclistic_Bike_Sharing_Data_Analysis

Figure 1. Ride Comparison (Casual riders vs Members)

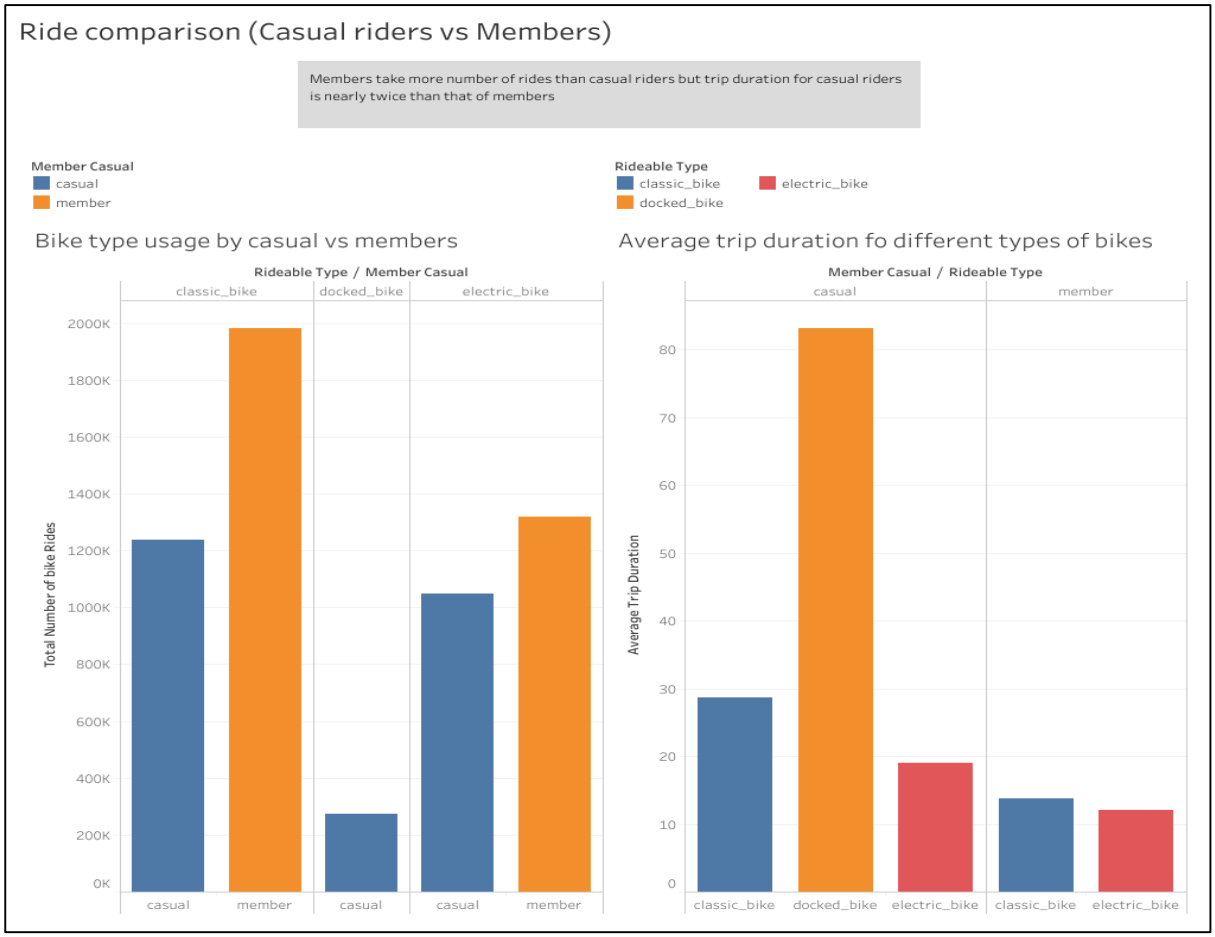

Figure 2. Timeline Comparison (Casual riders vs Members)

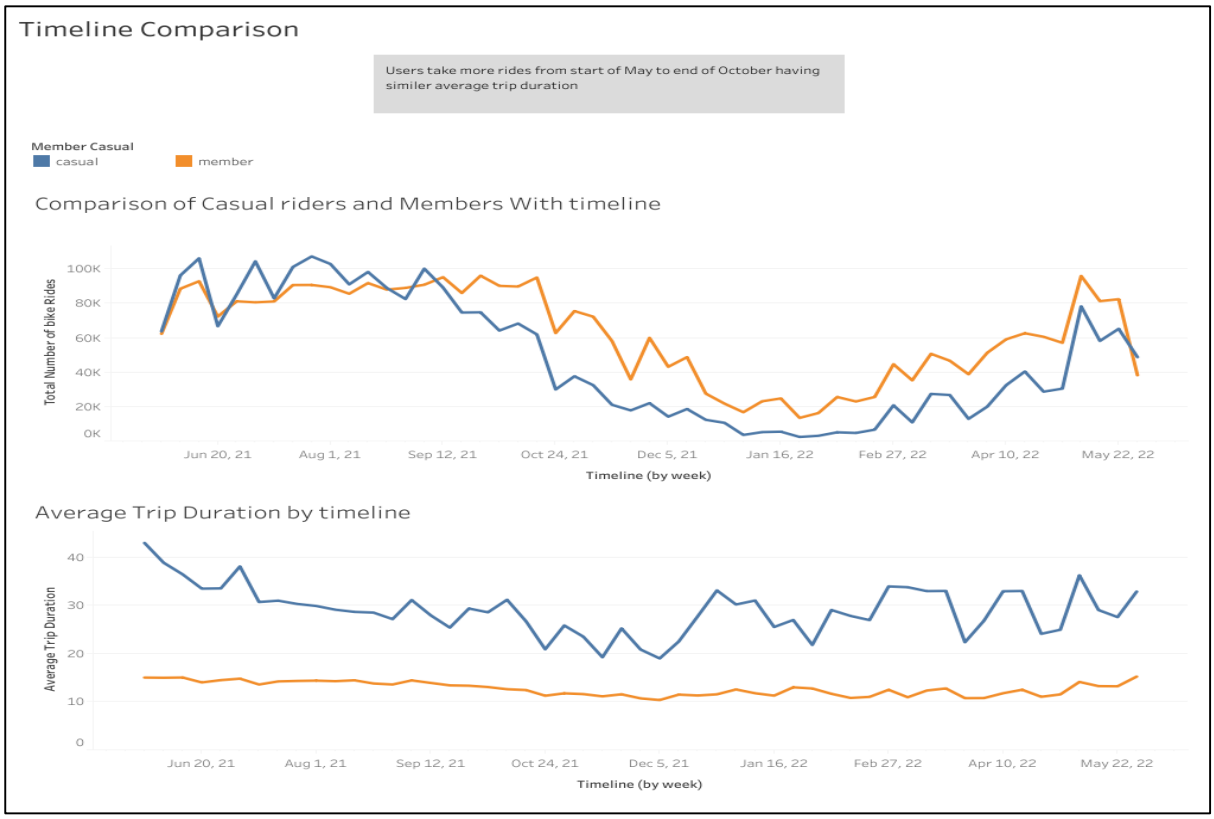

Figure 3. Comparison by Days of the week (Casual riders vs Members)