CYCLISTIC BIKE-SHARING DATA ANALYSIS REPORT

Prepared by: G.R.DEVAKUMAR

Date: 3/09/2023

Executive Summary:

This report outlines the key steps and analyses conducted on the Cyclistic bike-sharing dataset to gain insights into user behavior, understand differences between annual members and casual riders, and explore potential marketing strategies. The report includes data import, cleaning, descriptive analysis, data visualization, and considerations for communication.

1. Data Import:

The analysis began with the import of the dataset "202005-divvy-tripdata.csv" from the specified location.

2. Data Cleaning and Preparation:

Date-time columns ("started_at" and "ended_at") were converted to POSIXct format for time-based analysis.

Ride durations in minutes were calculated using the difference between start and end times.

The "member_casual" column was made consistent by converting all values to lowercase.

3. Descriptive Analysis:

Summary statistics were generated to provide an overview of the dataset.

A count of member vs. casual rides was calculated to understand the distribution of user types.

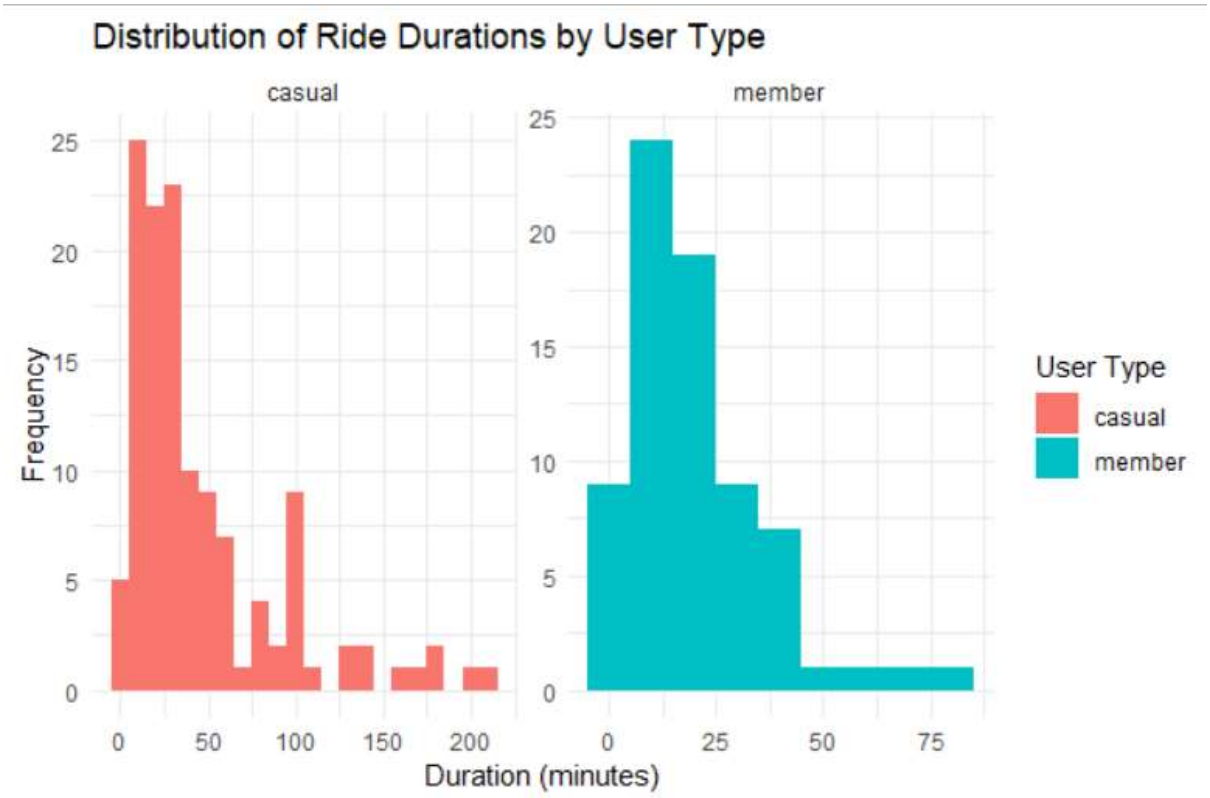
4. Data Visualization:

In this section, we present visualizations that provide a deeper understanding of the Cyclistic bike-sharing dataset and highlight the differences in user behavior between annual members and casual riders.

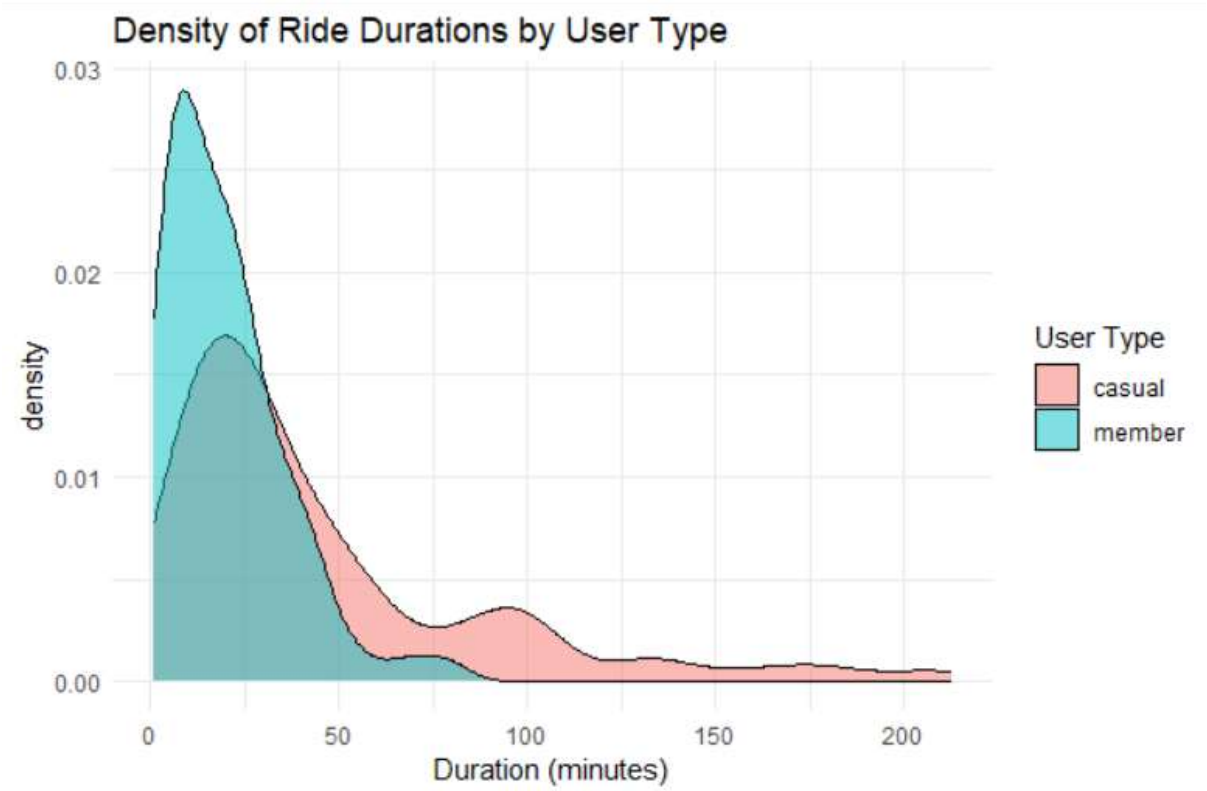
4.1 Ride Duration Distribution:

Histogram and Density Plot

These visualizations illustrate the distribution of ride durations for both user types.



Caption: Histogram of ride durations by user type.

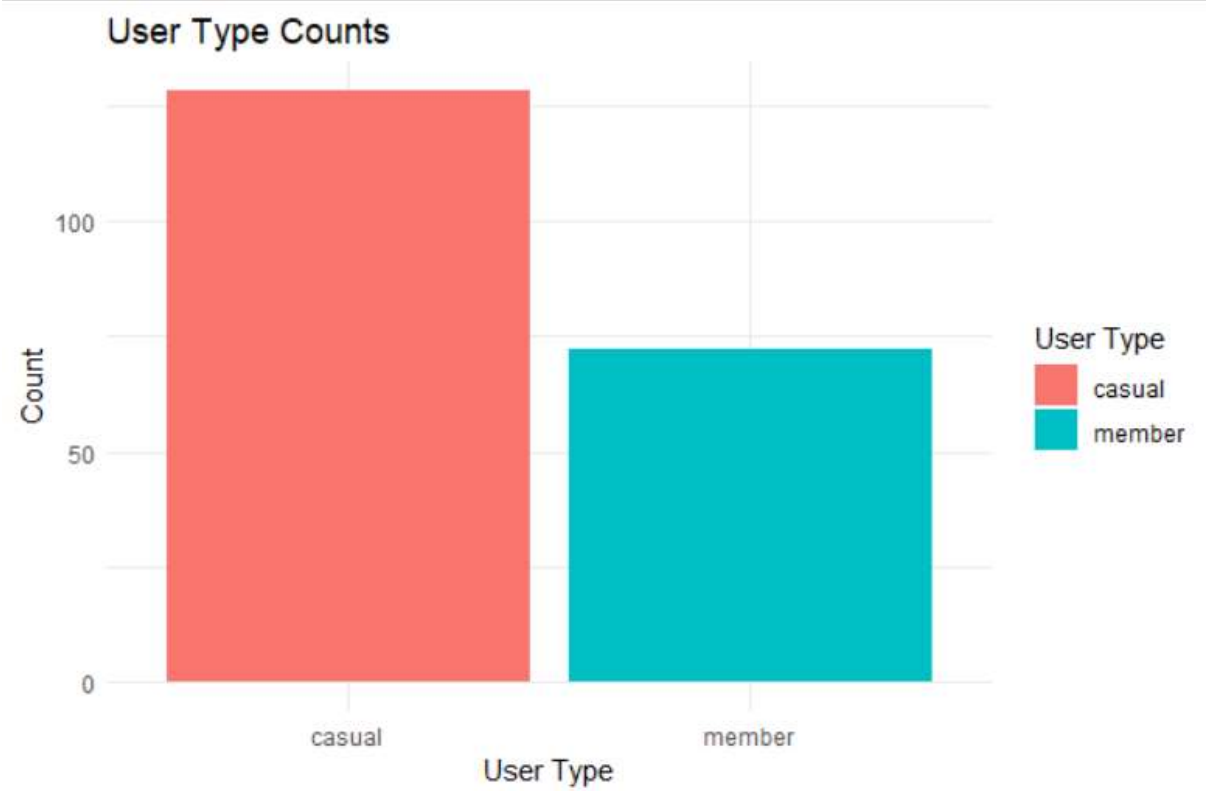


Caption: Density plot of ride durations by user type.

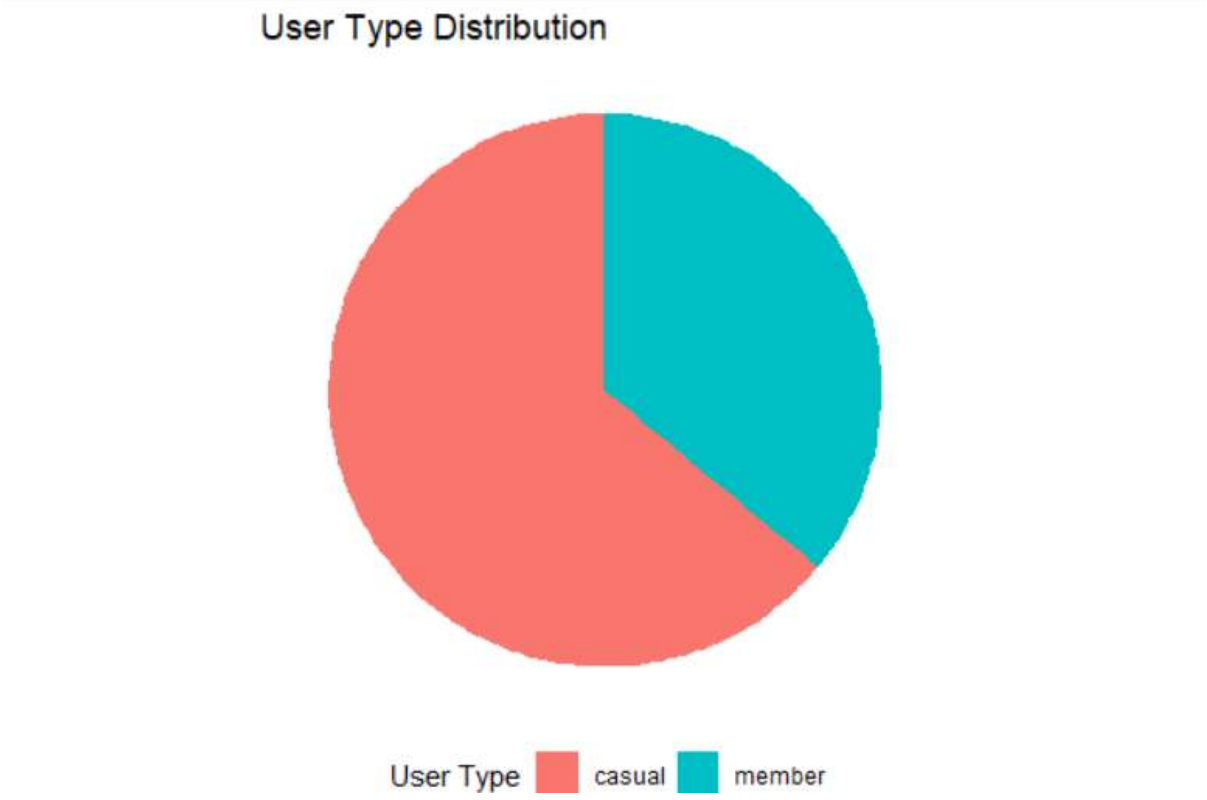
4.2 User Type Distribution:

Bar Chart and Pie Chart

These visualizations show the distribution of user types within the dataset.



Caption: Bar chart of user type counts.



Caption: Pie chart of user type distribution.

5. Answering Guiding Questions:

Were you able to answer the question of how annual members and casual riders use Cyclistic bikes differently?

The analysis suggests that annual members tend to have shorter and more consistent ride durations, while casual riders have longer and more variable rides. However, further analysis with a larger dataset is needed for conclusive findings.

What story does your data tell?

The data tells a story of varying ride behaviors among Cyclistic users, emphasizing the importance of understanding these differences for effective marketing.

How do your findings relate to your original question?

The findings align with the original question about user behavior differences between annual members and casual riders, but further research is needed.

6. Audience and Communication:

The primary audience for this analysis includes Cyclistic stakeholders, business analysts, marketing teams, and city planners. To effectively communicate the findings, visualizations, charts, and tables were used in a report or presentation format.

7. Accessibility Considerations:

Clear labels, titles, and descriptions were provided for visualizations.

Color schemes were chosen to ensure accessibility for individuals with color blindness.

Alternative text for images was included, and the presentation was made available in formats accommodating different needs.

8. Conclusion:

This report summarizes the key steps and findings from the Cyclistic bike-sharing dataset analysis. While initial insights were gained, it is essential to acknowledge the dataset's limitations and the need for more extensive data for comprehensive conclusions. The use of data visualization and accessible communication enhances the effectiveness of sharing these insights.