Summary Report: Comparison of Annual Members and Casual Riders in Bikeshare

By: Brandon Mathews

1. Introduction

In this report we are analyzing how members and non-members differ in their use of the Divvy Bikeshare program. The goal was to analyze usage in various ways such as by weekday, by month and by season while also highlighting which bike options tend to be favored more by whom. This is my first big data analysis project so there are bound to be some errors, but I put a lot of effort into this and learned a lot during the time it took me to complete this. Also, this is the capstone project for Google’s Data Analytics Course.

2. Data and Methods

Data Source: We used the data provided by the Google’s Data Analytics Course ([Index of bucket "divvy-tripdata"](https://divvy-tripdata.s3.amazonaws.com/index.html)) This trip gives us various different types of data organized into columns such as rider\_id, bike\_type, start and end time for trip, starting and ending locations.

Tools Utilized: For tools I used R Studio Desktop Version primarily for cleaning and visualization, but I also created a Tableau Dashboard for visualization as well.

Methods: I first cleaned the data using some of R’s built-in cleaning functions and removed non-essential columns which for me is the location information except for the starting station name. I also removed all trips under 5 minutes to ensure I removed all trips that were mistakes or accidental and from there I began to create multiple separate datasets to help with visualization which I had saved in a separate folder. From there I began to do my 3 main visualizations within R and lastly, I created a Markdown file and Tableau Dashboard to test some things I learned.

3. Key Findings

Usage Patterns:

- \*\*Total Trips: There were 4,512,656 rides in the 12 recorded months from June 2023 to June 2024 with 1,725,537 being casual members and 2,787119 being members.

- \*\*Trip Duration: Casual riders have a longer trip length at 24.5 minutes versus members 15.28 minutes.

- \*\*Peak Usage Times: Weekends and During the Summer

4. Visualizations

R Programming Visualization (ggplot2)

Tableau Visualizations:

<https://public.tableau.com/app/profile/brandon.mathews/viz/BikeShareDataAnalysis_17182916162670/VisualizationPart2>

I incorporated a dashboard and a story board showing different data comparisons that will be a part of my analysis.

5. Insights and Analysis

Differences in Usage:

* There was nothing unusual from what one would expect from users of a bike-share, albeit there were a lot more members than non-members in the dataset, which is a good sign.
* An interesting result is that Saturday and Sundays are the only days of the week where the number of rides taken by non-members and members is relatively close, other days show members using the bikeshare at almost 2x the number of casual riders.
* Bike Types are unique as only casual members use docked-bikes whereas classic and electrical bikes are somewhat even between both casual and members.

Implications:

- If you are implementing a membership, you should offer a bonus for those who rent during the weekday as well as a discount price for earlier months or a flat-rate all year around.

- Figure out the reason members aren’t using the docked bikes.

- It appears most casual members could very easily be members with similar usage rates and even longer average trip length so pricing or availability in locations could help the cause.

6. Challenges and Limitations

The dataset had a lot of data which was overwhelming at first, being relatively new to working with this many data in one project and some data types when importing was not in the correct type and naming conventions could have been better. Pricing and a more overarching location column like city, township or district could help identify potential reasons casual riders are becoming members. A better definition or naming for bike types could also potentially help in terms of giving proper insights since we don’t have rider’s personal data like profession or hobbies.

7. Conclusion

Summarize the main findings and their implications for the bikeshare service. Reflect on the significance of understanding these usage differences in optimizing service delivery and customer satisfaction.

Overall, Divvy did a great job with having a greater portion of their riders being members and that over 4.5 million rides in the last year is a great sign of usage by local communities. As expected, summertime and weekends are the most popular days that see the highest usage which can be expected if we assume that riders are off work and just going for a ride. The company should investigate pricing options and plans as we weren’t given that data thus couldn’t give any insight into that, but it could be a defining factor in increasing memberships. Also, some data was corrupted as it negative trip length so make sure to check stations and ensure bikes trips are being properly collected to ensure your data is being collected correctly.

8. Recommendations

* Promote Weekday Membership Usage:

Offer incentives such as discounted rates or bonus rewards for members who use the service during weekdays. This can help balance out usage patterns and increase revenue during off-peak times.

* Investigate Docked Bike Usage among Members:

Conduct surveys or gather feedback to understand why members prefer classic and electrical bikes over docked bikes. Addressing any issues related to availability, convenience, or pricing could encourage more diverse bike usage.

* Convert Casual Riders into Members:

Introduce targeted marketing campaigns or trial membership options aimed at casual riders who demonstrate high usage patterns and longer trip durations. Highlight the benefits of membership, such as cost savings and convenience. You can reorganize your location data to get overall areas that can give you insight on where to target your marketing.

* Enhance Data Quality and Analysis:

Improve data collection methods to prevent issues like negative trip lengths. Implement data validation checks and refine naming conventions for consistent and reliable analysis. Consider adding additional location details to better understand usage patterns across different areas.

* Optimize Seasonal Pricing Strategies:

Develop flexible pricing plans that adjust rates based on seasonal demand. For example, offer lower rates during colder months to encourage year-round usage and higher rates during peak summer periods to capitalize on increased demand.