

Cyclistic Case Study

Rebecca Gordon

6/30/2024

Introduction

Purpose: Cyclistic, a bike share company, is looking to improve profits and future growth, by converting casual riders to annual members. The overall analysis involves identifying how the two customer types differ in usage, why casual riders would buy a membership, and how digital media could affect marketing tactics. This study focuses on the first question.

Company Summary: Cyclistic launched in Chicago in 2016. It features 5824 geo-tracked bicycles that can be unlocked from and returned to any of the 692 docking stations around the city. Cyclistic also features flexible pricing plans and several assistive bicycle options.

Tools: For this analysis, I will be using Microsoft Excel, Google Slides, and Tableau.

ASK

Guiding questions:

- What is the problem you are trying to solve?
- How can your insights drive business decisions?

Business Task:

This case study is designed to assess different trends in Cyclistic bike usage between annual members and casual users in order to improve marketing and convert casual users to annual members.

Shareholders:

- Lily Moreno: Director of Marketing at Cyclistic
- Cyclistic Executive Team
- Cyclistic Marketing Analytics Team

PREPARE

About the Data: The data used in this case study is historical bike trip data (made available by Motivate International Inc. under this [License](#)). This is a vetted public dataset, and its data can be considered reasonably reliable and credible. The data has been organized by year

and quarter. This study will look at the last 6 months of data available in order to make our study as current as possible. The data is already anonymized with ride and trip ID numbers, which ensures the privacy of the bike share customers will not be impacted by this study. This data on trip and demographic information can be analyzed to assess differences in bicycle usage between casual riders and annual members.

PROCESS

Data Cleaning Process: Data cleaning was performed to:

- Investigate any errors
- Deal with Null data
- Fix misspelled words
- Fix mistyped numbers
- Eliminate any extra spaces and characters
- Remove duplicate entries
- Fix any mismatched data types
- Fix messy strings or data types
- Adjust column names for consistency
- Check the overall business logic of the cleaned data

Changelog: Please see the Changelog file under additional resources on my portfolio for documentation of cleaning and manipulation of the data.

Assessment: At this point, the data has been checked for all common issues and is clean and ready for analysis. The only significant issues found were the inaccurate data which indicated negative or zero trip durations that were removed from the analysis and the fact that gender and birth year data was not recorded for all user trips. There were a few outliers when calculating age groups with a small group of users being significantly older than seems likely, but this did not materially affect the analysis.

ANALYZE

Exploratory Analysis:

Now that the data is clean and organized, I created pivot tables for each quarter to help investigate trends. I used them to calculate AVG trip duration by day of week for all users, casual users, and annual members. Then, I calculated the number of trips by day of the week for all users, casual users, and annual members. I decided to gather any results together in a new sheet (Analysis Insights) for ease of viewing tables with aggregate data. In order to allow me to

look at the demographic data, I created separate sheets of full clean 2019_Q4 data. I named one for gender and filtered out null fields where gender was not specified; I named the other for birthyear and filtered out nulls as well. I then created another pivot table and generated the average trip duration and trip count for both user types divided by gender. Then, I added a column for age and converted the birthyear column to age for easier analysis. I added a final pivot table and generated average trip duration and trip count for both user types divided by age groups. Finally, I examined the Analysis Insights sheet with all aggregated data tables from previous steps to determine different trends by user type. I created bar charts for simple visualization of aggregate data and created a sheet to keep these graphs together for easy reference for exploratory analysis. When it came to visualization, I realized there were two issues with the current organization of data and fixed them. The 1-7 integer representation of the day of the week was cumbersome, so I changed it to text weekdays instead. I separated the age data into groups during analysis with the pivot tables, but I needed to actually create an age group column in the datasheet for proper visualization. I took the analysis conclusions from this stage and used them to inform the final data visualizations I then created in Tableau.

Observations:

Casual Users have a higher average trip duration across all days of the week. **Annual members** have consistently lower average trip duration across days of the week. This may be due to annual members utilizing the service for daily commutes and chores, while casual users utilize the service for longer occasional rides that don't reflect daily activities.

Annual Members have higher total trips across all days of the week. **Annual Members** have lower trip counts on the weekends. **Casual users** have consistently lower total trips across all days of the week. **Casual users** have higher trip counts on weekends, particularly on Sundays. This strengthens the earlier assumptions of annual members taking shorter but more frequent rides tied into daily activities, while casual users take longer rides, but less frequently.

When looking at total trip values by age group it was apparent that 21-30 and 31-40 were the biggest age groups for both types of users. Total trips by age group reflected the higher numbers of trips taken by members vs casual users. There was not an obvious trend in average trip duration per age group.

No obvious trends in average trip duration across genders. In total trips there does seem to be a trend of higher total trip values in males vs females across both user groups.

SHARE

Conclusions and Recommendations: I have prepared a presentation to communicate my findings. Please see the Presentation link under additional resources on my portfolio page for that document.

Thank You!