

CYCLISTIC BIKE-SHARE COMPANY, CHICAGO

A CASE STUDY

Introduction

You are a junior data analyst working in the marketing analyst team at Cyclistic, a bike-share company in Chicago. The director of marketing believes the company's future success depends on maximizing the number of annual memberships. Therefore, your team wants to understand how casual riders and annual members use Cyclistic bikes differently. From these insights, your team will design a new marketing strategy to convert casual riders into annual members. But first, Cyclistic executives must approve your recommendations, so they must be backed up with compelling data insights and professional data visualizations.

Until now, Cyclistic's marketing strategy relied on building general awareness and appealing to broad consumer segments. One approach that helped make these things possible was the flexibility of its pricing plans: single-ride passes, full-day passes, and annual memberships. Customers who purchase single-ride or full-day passes are referred to as casual riders. Customers who purchase annual memberships are Cyclistic members.

Cyclistic's finance analysts have concluded that annual members are much more profitable than casual riders. Although the pricing flexibility helps Cyclistic attract more customers, your marketing manager, Moreno believes that maximizing the number of annual members will be key to future growth. Rather than creating a marketing campaign that targets all-new customers, Moreno believes there is a very good chance to convert casual riders into members. She notes that casual riders are already aware of the Cyclistic program and have chosen Cyclistic for their mobility needs.

Moreno has set a clear goal: Design marketing strategies aimed at converting casual riders into annual members. In order to do that, however, the marketing analyst team needs to better understand how annual members and casual riders differ, why casual riders would buy a membership, and how digital media could affect their marketing

tactics. Moreno and her team are interested in analyzing the Cyclistic historical bike trip data to identify trends.

To solve this business problem, I'll be applying the six phases of data analytics learnt from the Google Data Analytics Certificate. I'll also be making use of the following tools: Google Sheets, MySQL, Tableau and Google Slides.

THE ASK PHASE

Cyclistic has two major types of customers, the casual riders and the annual membership subscription riders. Cyclistic wants to maximize the number of annual memberships as it is clear that they generate the most revenue for the company. Part of this initiative is to convert casual members into annual members. I then proceeded to ask "how do members and casual riders use bikes differently". This will be critical to guiding the business task.

The key stakeholders in this project include:

- The executive team: The directors who are charged with running the company. They are to decide whether to approve the recommended marketing program.
- Lily Moreno: The director of marketing and senior analytics manager.
- The marketing analytics team: A team of data analysts who are responsible for collecting, analyzing, and reporting data that helps guide Cyclistic marketing strategy.

THE PREPARE PHASE

The data consists of four different datasets; April 2020 to July 2020 datasets which I wrangled and combined into a single file. The data is an open data provided by Motivate International Inc. After evaluating the data, it turns out that the data source is reliable, original, comprehensive, current, and cited. It can thus be used for further analysis. Since the data takes into consideration the details and doesn't show any form of subconscious preference against a certain group or set of members, it can be

interpreted to be an unbiased one. For this analysis, the latitude and longitudes wouldn't be used as it contains missing values and mapping them to their stations is intricate.

THE PROCESS PHASE

To process the data, I used spreadsheets firstly, (Google sheets to be precise). Firstly, I removed null and duplicate values and then proceeded to add some new column to the datasets, including "length of ride" (which indicates the time difference between the start of each biker's journey and the end journey), hour, week, year, month and date. Also, to have an understanding of the day members ride, I created a new column, "day of the week" to do that.

THE ANALYZE PHASE

In this phase, I used SQL to format and transform the data to try to identify patterns and perform descriptive analysis. This write-up only summarizes the data. The full analysis can be found in the next phase where I also made use of visualizations. For clarity purposes, I changed the member_casual column name to 'user'. I also made use of pivot table to summarize the data in spreadsheets.

Data Info

The dataset consists of 1177593 rows and 19 different fields.

User details

There are 124197 annual members and 40200 casual riders.

Summary of ride time.

The maximum ride time for a member is *10122942966* sec and *70944465005* sec for a casual rider. The minimum ride time for a member is *2* sec and *1*sec for a casual rider. On average, members clock *15719* seconds while casual riders clock *132956* seconds.

Bike types

There were two bike types available for this data. The docked bike was used 643,028 times by members, 532,914 times by riders. The electric bike was used 971 times by members while casual riders used it 680 times.

Start stations

For casual riders, the most trips was 10510 rides which started at the *Streeter Dr & Grand* station while for members, 6,216 rides started at *Miles van der Rohe Wa* station, the highest of any.

End stations

10967 rides was ended at the *Streeter Dr & Grand station*, the highest among casual riders while it was 6, 548 at the *Clark St & Elm St* among annual members.

These tasks here are just the summary of the statistics I did. I used SQL to pull the data to perform cleaning, wrangling and other analysis. Further analysis and visualizations were made during the “Share” phase to make the insights and findings much clearer.

THE SHARE PHASE

In this phase I created effective visuals to better understand the data. I also used presentation software for data storytelling and then communicating the results to the stakeholders. I made use of the popular viz tool, Tableau to create visualizations. The link to the Tableau Visualization is [bike share for 2020\(April-July\) | Tableau Public](#).

THE ACT PHASE

After I was done with the visualizations, I moved on to the final step where I acted on the findings and insights to help the company solve the business objective; maximize the number of annual memberships as it is clear that they generate the most revenue thereby maximizing profits for the company.

Key Insights

Based on the analysis, the travel duration of casual riders is much higher than that of the annual members. Meanwhile, members take more rides than the casual riders. This indicates casual riders are just individual riders who ride at any time be it for fun, hobby, etc while annual members who are possibly dedicated members who ride for a purpose which can be for work.

Streeter Dr & Grand station hosted the highest amount of starting rides among both the annual members and casual riders, meaning it was the most popular station among both group of riders.

Annual members are more active during the weekdays.

Recommendations

- If the longer riders amongst the casual riders become members, give them some discount
- A time-limited subscription should be introduced to the casual riders that use a certain bike than the rest to entice them to become annual members.