

Google Data Analytics Capstone: Cyclistic Case Study



Author: Bindu Subramanian

Published on : <https://bindusubramanian.wordpress.com/2023/11/13/google-data-analytics-capstone-cyclistic-case-study-2/>

11/13/2023

Course: <https://www.coursera.org/learn/google-data-analytics-capstone>

Introduction

In this case study, I take on tasks of a junior data analyst at a fictional company, Cyclistic to answer key questions, by following the steps of the data analysis process: Ask, Prepare, Process, Analyze, Share and Act. Tools Used - Excel, Microsoft SQL Server, Tableau

Scenario

The Cyclistic Director of Marketing, Lily Moreno, believes the company's future success depends on maximizing the number of annual memberships. Therefore, my team wants to understand how casual riders and annual members use Cyclistic bikes differently. From these insights, my team will design a new marketing strategy to convert casual riders into annual members. But first, Cyclistic executives must approve our recommendations, so they must be backed up with compelling data insights and professional data visualizations.

About the Company

In 2016, Cyclistic launched a successful bike-share offering. Since then, the program has grown to a fleet of 5,824 bicycles that are geotracked and locked into a network of 692 stations across Chicago. The bikes can be unlocked from one station and returned to any other station in the system anytime.

Cyclistic's appeal is that they offer flexibility in their pricing plans with 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. Moreno believes there is a very good chance to convert casual riders into members.

Moreno has set a clear goal: Design marketing strategies aimed at converting casual riders into annual members. To do that, however, Moreno and her team are interested in analyzing the Cyclistic historical bike trip data to identify trends.

Ask

Analyze Cyclistic's 2022 trip data to understand how casual riders and annual members use bikes differently. Insights gained from this analysis will help the marketing team develop their strategies for the campaign.

Stakeholders

Lily Moreno: Director of Marketing. Moreno is responsible for the development of campaigns and initiatives to promote the bike-share program. These may include email, social media, and other channels.

Cyclistic marketing analytics team: A team of data analysts who are responsible for collecting, analyzing, and reporting data that helps guide Cyclistic marketing strategy.

Cyclistic executive team: The notoriously detail-oriented executive team will decide whether to approve the recommended marketing program.

Prepare

The first step is to download all the data that I will need for my analysis. We will be using the Cyclistic trip data for 2022 which needs to be download in 12 separate .csv files for each month of the year and stored in a dedicated folder. The data has been made available by Motivate International Inc. under this license - <https://ride.divvybikes.com/data-license-agreement>

Process

Data Cleaning

1. To begin the data cleaning process, I opened each .csv file in excel and did the following
 - Checked for and removed any duplicates followed by the **trim()** function to remove unneeded spaces
 - Used the **weekday()** function to create a new column labeled **day_of_week** using (1-7) to represent (Sunday-Saturday)
 - Created a new column labeled **ride_length** by subtracting the **started_at** column from the **ended_at** column
 - Changed the time format to **37:30:55** to make it more readable
 - Removed any rides under **1 minute** or longer than **24 hours** by sorting the spreadsheet.

2. After uploading each of the twelve files, I use SQL combined each file into one table labeled **combined_data** using the **union** function. In the same query, I removed each of the rows that contained null values. This is my clean data.

Analyze

The first question was - ***How do casual riders and members use bikes differently?***

To answer this question, there were a few things that I identified that I could pull from the data. I wrote queries for the following:

- Total # of trips per rider type
- Total # of trips per rider type per bike type
- Average ride length per rider type
- Total # of trips and average length of trip per rider type per month, day, and hour of day
- Most Popular Start and End Stations per rider type

All SQL queries can be found here - <https://github.com/bsbramanian/github.io/blob/main/Analyze%20Data>

These are my observations after analysis of the data:

- Summer months were more active for both casual riders and member riders.
- Members had around a million more trips in 2022 than casual riders.
- Casual riders trips on average lasted around double the length of members

Share

The dashboard I created for this project can be found on tableau here - <https://public.tableau.com/app/profile/subramanian.bt/viz/CyclisticCaseStudy-GoogleDataAnalytics/CaseStudy>

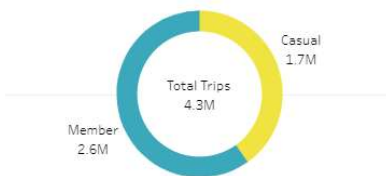
- Trips per rider type
- Trips per Bike Type
- Average ride length per rider type
- Top Start and end stations
- Number of trips per month, day, and hour



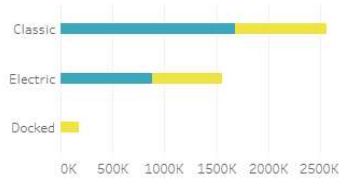
Cyclistic Case Study

How do casual riders differ from members?

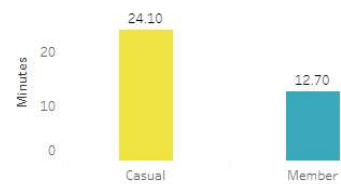
Trips per Rider



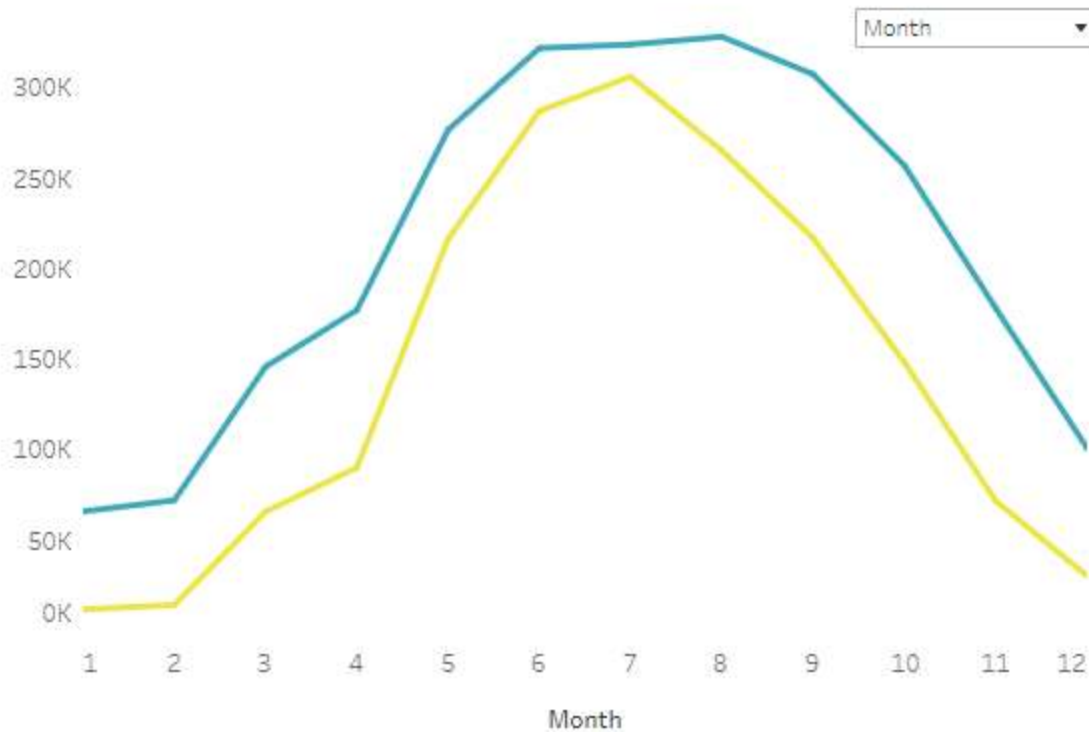
Trips per Bike Type



Avg Ride Length



Number of Rides



My key observations are –

Members:

- Members use bikes for daily commutes to work, university, etc. as evidenced by peak activity which occurs on weekdays during working and school hours.
- Average ride length is shorter and consistent, indicating routine travel to familiar destinations.
- Peak months are from May to September, coinciding with favorable weather for bike usage.

Casual Riders:

- Casual riders use bikes for leisure and entertainment purposes as we can see that they are usually near popular public attractions.
- Peak months align summer and spring seasons of the year. And weekends are the busiest for them.

Act

Based on my analysis, below are some suggestions for an effective campaign to convert casual riders into members:

1. Offer a seasonal membership for summer catering to casual riders. Discounts for renting bikes for longer durations and only for weekends can also be an effective strategy.
2. Offer Corporate Discounts to gain more customer base with the working population. Provide them with more pickup and drop off points across the routes so they can stop for grocery shopping etc on their way back home and make biking part of their day.
3. Offer Discounts to School and University students which can be a huge customer base and partnering with camping and other recreational avenues to increase usage during weekends and summer seasons.