Google Data Analytics Capstone Project: Cyclistic

Pavitra Rana

9/25/2021

Business task: Analyze the usage of Cyclistic bikes differently by annual members and casual riders.

Introduction

Welcome to the Cyclistic bike-share analysis case study! In this case study, you will perform many real-world tasks of a junior data analyst. You will work for a fictional company, Cyclistic, and meet different characters and team members. To answer the key business questions, you will follow the steps of the data analysis process: ask, prepare, process, analyze, share, and act. Along the way, the Case Study Roadmap tables — including guiding questions and key tasks — will help you stay on the right path. By the end of this lesson, you will have a portfolio-ready case study. Download the packet and reference the details of this case study anytime. Then, when you begin your job hunt, your case study will be a tangible way to demonstrate your knowledge and skills to potential employers.

Scenario

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.

Characters and teams

Cyclistic: A bike-share program that features more than 5,800 bicycles and 600 docking stations. Cyclistic sets itself apart by also offering reclining bikes, hand tricycles, and cargo bikes, making bike-share more inclusive to people with disabilities and riders who can't use a standard two-wheeled bike. The majority of riders opt for traditional bikes; about 8% of riders use the assistive options. Cyclistic users are more likely to ride for leisure, but about 30% use them to commute to work each day.

Lily Moreno: The director of marketing and your manager. 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. You joined this team six months ago and have been busy learning about Cyclistic's mission and business goals — as well as how you, as a junior data analyst, can help Cyclistic achieve them.

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

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 geo-tracked 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.

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, 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. 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.

Business task

Analyze the usage of Cyclistic bikes differently by annual members and casual riders.

Data Source

The Data used of 12 months of previous Cyclistic trip and merged or combine all data in one sheet with the help of R.

Source: **Download the previous 12 months of Cyclistic trip data here.**

For the R-code, click here >>**GitHub**

For Tableau Dashboard, click here >>**Tableau**

Tools:

- R
- Tableau
- MS PowerPoint
- MS Word

Analysis

Ask

Three questions will guide the future marketing program:

1. How do annual members and casual riders use Cyclistic bikes differently?
2. Why would casual riders buy Cyclistic annual memberships?
3. How can Cyclistic use digital media to influence casual riders to become members?

In this case study, the project followed the first question to provide a possible solution with the help of different tools and recommendations according to the analysis.

Prepare

The data is gathered from secondary sources and use Cyclistic's historical trip data to analyze and identify trends. Download the previous 12 months of Cyclistic trip data here. (Note: The datasets have a different name because Cyclistic is a fictional company. The data has been made available by Motivate International Inc. under this license.) This is public data that can be used to explore how different customer types are using Cyclistic bikes. Additionally, the data is well organized with required columns and rows with no NULL values. And by using Date, Time, and Ride duration columns in the data. It will help to solve the task problem with some calculations and analysis.

Process

The tools used in this project following are:

- R: For data wrangling and Analysis
- Tableau: For visualization
- PowerPoint: For the presentation of the project
- Word: For report and documentation

The data was verified with no errors and NULL values by reading and checking the columns of all data files to ensure identical attributes. However, for further Data clean up and the process by changing the required format and adding new columns for further analysis.

Analyze

There are various steps followed in this case study to analyze the data using R as:

STEP 1: Collect Data

STEP 2: Wrangle Data and combine it into a single file

- In this process, the data cleaning or manipulation and transformation steps followed.
- Rename Columns
- Inspect the data frames and look for incongruences
- Compare column datatype across all data frames by using compare_df_cols
- Convert end_station_id and start_station_id to the character datatype so that they can stack correctly
- Stack individual data frames into one big data frame
- Remove unused column

STEP 3: Clean up and add Data to prepare for Analysis

- Add columns that list the date, month, day, and year of each ride.
- Add a “ride_length” calculation to all_trips (in seconds).
- Convert “ride_length” from Factor to numeric so we can run.
- Calculations on the data.
- Remove “bad” data.
- The data frame includes a few hundred entries when bikes were taken out of docks and checked for quality by Divvy or ride_length was negative.
- Create a new version of the data frame (v2) since data is being removed.

STEP 4: Conduct Descriptive Analysis

- Descriptive analysis on ride_length.
- Compare members and casual users.
- See the average ride time by each day for members vs casual users.
- Fix the days of week in order.
- Run the average ride time by each day for members vs casual users.
- Analyze ridership data by type and weekday.
- Visualize the number of rides by rider type.
- Visualization for average duration.

STEP 5: Export Summary file for further Analysis

Export to CSV file for further analysis

Share

The visualization was created using the **Tableau** tool in a sophisticated and polished manner to effectively communicate to the executive team. It tells a story about the case study with the help of different visual insights. The Dashboards can answer the question of how annual members and casual riders use Cyclistic bikes differently? The presentation is accessible and sharable with the given link click here>> **pptx**

Act

This is the final process of the Cyclistic case study where the recommendations are provided according to the data analysis and visualization of the given data. The top three recommendations are given below in the **Recommendations** section.

Summary

- The duration of total trip of both user type casual and member are affected according to the month. As the data are given for 12 months and according to the months the seasonal figures depict the higher usage of the bikes in August. Due to different seasons, the duration is affected because during the winter season the temperature is very low, so fewer people willing to use the bikes or travel instead they are using public transport. This leads to a low number of users in the winter season among other seasons of trip duration.
- The riders of members use to ride at 7 AM morning that indicates that the riders are likely the office users who use bikes to ride to reach their office and at 5 pm evening, it indicates they use the rides to back to their home.
- The riders of casual use to ride in the evening with different times mostly at 4 PM and 5 PM evening that indicate that the riders are likely the locals and tourist who use to ride daily or according to their daily activities, the casual users most likely are the students, local people, family, tourists, etc.

Recommendations:

- According to the behavior of users of the casual riders based on the month the suitable time to pitch a marketing strategy or marketing campaign is between April to May.
- Additionally for an effective marketing campaign, consider different strategies such as discount coupons, seasonal coupons, and free rides to increase the conversion rate.
- For a more effective marketing strategy, the marketing team can collect more data about the types of users of casual riders. So that the strategy is designed and planned effectively for the targeted user. Therefore, the team could run a survey & questionnaire to collect the relevant information that leads to help in pitching different strategies effectively and efficiently such as their interests, likes & dislikes, and what they are looking for, considering applying for an annual subscription.