**Cyclistic Case Study**

Table of Contents

[1.0 ASK 1](#_Toc109813673)

[2.0 PREPARE 2](#_Toc109813674)

[3.0 PROCESS 3](#_Toc109813675)

[4.0 ANALYSIS 3](#_Toc109813676)

[5.0 SHARE 4](#_Toc109813677)

[6.0 ACT 8](#_Toc109813678)

1.0 ASK

**1.1 ABOUT THE COMPANY**

Cyclistic is a successful bike-share offering program. Since its launch in 2016, the program has grown exponentially to a fleet of 5,824 bicycles that are geotracked and locked into a network of 692 stations across Chicago. Cyclistic has made bike-sharing more inclusive by offering assistive bike options.

**1.2 BUSINESS TASK**

Cyclistic offers annual memberships for regular customers and single-ride or full-day passes for casual riders. Lily Moreno, the director of marketing, has relied on building general awareness and appealing to broad customer segments. Cyclistic finance analysts have recently concluded that the annual members are more profitable than casual riders. Hence, the marketing analyst team will be analyzing the trends in the historical bike trip data from the last 12 months to better understand

* How do annual members and casual riders use Cyclistic bikes differently?
* Why casual riders would buy a membership?
* How digital media could affect their strategy?

**1.3 SCOPE**

The scope of this case study is to answer the first question from the above list as assigned by Lily Moreno.

**1.4 KEY STAKEHOLDERS**

* **Lily Moreno,** the director of marketing.
* Cyclistic marketing analytics team.
* Cyclistic executive team.

2.0 PREPARE

**2.1 DATA SOURCE AND LICENSE**

* The dataset used in this analysis is Cyclistic’s historical bike trip data can be found [*here*](https://divvy-tripdata.s3.amazonaws.com/index.html).
* This dataset has been made available by Motivate International Inc. under this [*license*](https://ride.divvybikes.com/data-license-agreement).

**2.2 DATA CREDIBILITY**

**2.2.1 RELIABILITY:** This dataset is considered to be reliable as the data is provided by trusted stakeholders, and assured to answer the business questions.

**2.2.2 ORIGINALITY:** Since Cyclistic is a fictional company the dataset is not original to the company. The dataset is made available by Motivate International Inc. under this [*license*](https://ride.divvybikes.com/data-license-agreement).

**2.2.3 COMPREHENSIVE:** The data is collected from every rider regardless of the status of the customer(member/casual). Hence, the sample represents the population accurately.

**2.2.4 CURRENT:** Lily Moreno has asked to analyze the previous 12month historical trip data of Cyclistic. This case study was conducted in July 2022, and the data used dates from June 2021 to May 2022. The data for June 2022 was not available yet. Therefore the data is considered current.

**2.2.5 CITED:** Lyft Bikes and Scooters, LLC (“Bikeshare”) operates the City of Chicago’s (“City”) Divvy bicycle-sharing service. Bikeshare and the City are committed to supporting bicycling as an alternative transportation option. As part of that commitment, the City permits Bikeshare to make certain Divvy system data owned by the City (“Data”) available to the public, subject to the terms and conditions of this License Agreement (“Agreement”).

**2.3 IMPORTING DATA**

The datasets are downloaded from [here](https://divvy-tripdata.s3.amazonaws.com/index.html). The downloaded datasets are imported into the MS SQL Server database using the MS SQL Server Import and Export Wizard into tables for further exploration, cleaning, and analysis.

**2.4 DATA EXPLORATION**

The imported datasets can be queried now to study and familiarize the individual rows and columns. The datasets are large, which will help with the detailed analysis report. Upon further analysis, it is evident that the columns match every table imported. Some rows have null values that need to be cleaned in the next step.

3.0 PROCESS

**3.1 MERGING THE TABLES**

Since the columns of all the tables are the same, for ease of analysis, these tables are merged into a single table. The data\_types of some of the columns are altered to match the data in it and also match the other tables that need to be merged.

**3.2 CLEANING**

The merged table is named trip\_data and needs to be cleaned for the most accurate results.

The rows containing null values, duplicate values, and ride\_length with less than a minute or more than 24 hours are deleted. The values are then checked for anomalous entries, misspellings, and unnecessary trailing or preceding white spaces.

**3.4 MANIPULATION**

The DateTime columns ‘started\_at’ and ‘ended\_at’ are manipulated to separate the time into other new columns. Using the dates, the ‘day\_of\_week’ column is calculated. By subtracting the columns ‘started\_at’ and ‘ended\_at’, the length of each ride is calculated and added to the new column ‘ride\_length’.

Queries used for processing the data can be found [*here*](https://github.com/SubasiniR/cyclistic_case_study_using_sql/blob/main/Process.sql)**.**

4.0 ANALYSIS

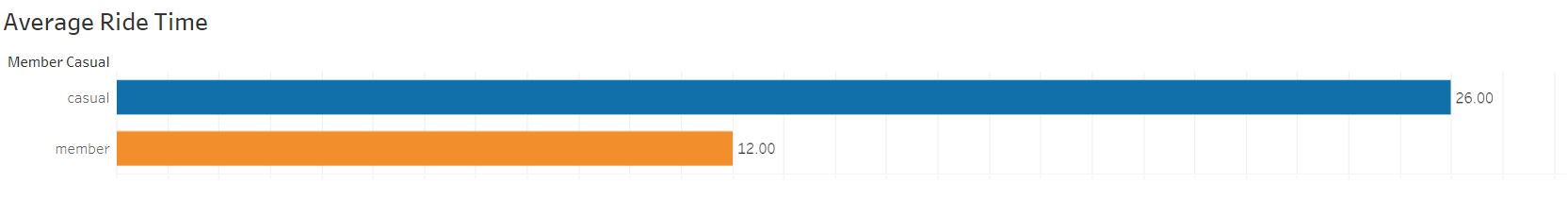
First, I ran a few calculations to get some basic information like mean, min, and max ride\_length. I have grouped the entire analysis based on the member\_casual column as it is a report on how the member rides compare to casual riders. Based on the query results the following observations are made.

* The average ride\_length of members is 12mins and the average ride\_length of casual riders is 26mins. Meanwhile, the percentage distribution of the member vs casual riders is 57.8% vs 42.2%. This shows that the riders with membership use the bikes consistently even if the period of usage is shorter. On the other hand, casual riders are mostly using the bikes for occasional, unscheduled, or recreational rides.
* The classic bikes are chosen by the riders for 68% of the total rides last year. The electric bikes are chosen for 25.5% of rides and docked bikes are the least used at 5.5%. Evidently, classic bikes are the rider's first choice. When further analyzed, 48.5% of the classic bike users are members and the remaining 25.5% are casual riders. Likewise, 15.3% of electric bike users are members and only 11.2% are casual riders.
* The usage of bikes varies according to different seasons as well. Chicago gets extremely cold in the winter and the snow makes it unsafe for bikers. So that explains the drop in the usage of bikes in winter. Rides are almost equally enjoyed well in the spring and summer by both members and casual riders. Fall is the season where we can see an exponential increase in rides.

Queries used for analysis can be found [*here*](https://github.com/SubasiniR/cyclistic_case_study_using_sql/blob/main/Analysis.sql).

5.0 SHARE

The visualizations of the results are made using Tableau. Visualization can be found [*here*](https://public.tableau.com/app/profile/suba8614/viz/CyclisticCaseStudy_16578300614860/AvgRideTime)**.**



As shown above, the average ride time of the members is less than half of the casual riders.

Chart, pie chart

Description automatically generated

The percentage of rides by members is more than the casual riders. Hence, the members are more consistent users than the casual riders.

Chart, bar chart

Description automatically generated

In the above bar chart, there are distinguishing curves for members versus casual riders. Casual riders use the bikes more on the weekends than on the weekdays and vice versa for members. This explains that the casual members are using bikes more for recreational rides and members use them for commuting to work on the weekdays. The casual members use their bikes relatively low on the weekdays compared to the weekends. Whereas, most members use the bikes even on the weekends as it is included with their membership.

Seasonal usage of the bikes chart below shows how the weather in Chicago plays a vital role in riders choosing to ride on any day. Since Chicago has harsh winters there is an evident drop in the number of rides. Surprisingly the fall weather is when the riders enjoy the bike rides the most than spring and summer.

Chart, bar chart

Description automatically generated

Chart, line chart

Description automatically generated

The above graph shows the usage of the bikes over the day. Members graph peeks at the regular office times like 6 to 8 in the mornings and 4 to 7 in the evenings. Whereas, casual riders' graph peeks at 4 to 7 in the evenings. This again proves that most members use the bikes for commuting to work and most casuals use them for recreational purposes.

Map

Description automatically generated

This map shows the popular destinations for both members and casual riders. Further study could help us understand what makes these destinations popular among the riders.

6.0 ACT

Based on the usage pattern of the members vs casual riders, the following suggestions are made to help convert the casual riders into members.

* Since most members use the bikes to commute to work and casual riders use them for recreational activities, two different types of membership plans can be curated to meet the individual needs.
* Providing free trial period offers to casual riders to aid them to understand the membership benefits.
* Targeted advertising and offers at popular destinations.
* Conducting biking events to create social awareness about the benefits of biking and attract more potential members.