**ASK**

The business problem: Annual member is more profitable. Therefore, Cyclistic wants to convert casual riders into annual members.

The task: to understand how annual members and casual riders differ

Result: to backup strategies for marketing campaign in order to convert casual riders into annual members

Type of analysis: Identify Trends

**PREPARE**

The data is stored within csv files. There are 12 files with each containing ride data for each month.

|  |  |  |
| --- | --- | --- |
| **Month** | **Initial name** | **Changed Name** |
| January 2022 | 202201-divvy-tripdata.csv | 202201-divvy-tripdata.csv |
| February 2022 | 202202-divvy-tripdata.csv | 202202-divvy-tripdata.csv |
| March 2022 | 202203-divvy-tripdata.csv | 202203-divvy-tripdata.csv |
| April 2022 | 202204-divvy-tripdata.csv | 202204-divvy-tripdata.csv |
| May 2022 | 202205-divvy-tripdata.csv | 202205-divvy-tripdata.csv |
| June 2022 | 202206-divvy-tripdata.csv | 202206-divvy-tripdata.csv |
| July 2022 | 202207-divvy-tripdata.csv | 202207-divvy-tripdata.csv |
| August 2022 | 202208-divvy-tripdata.csv | 202208-divvy-tripdata.csv |
| September 2022 | 202209-divvy-publictripdata.csv | 202209-divvy-tripdata.csv |
| October 2022 | 202210-divvy-tripdata.csv | 202210-divvy-tripdata.csv |
| November 2022 | 202211-divvy-tripdata.csv | 202211-divvy-tripdata.csv |
| December 2022 | 202212-divvy-tripdata.csv | 202212-divvy-tripdata.csv |

Constraint: Due to privacy issue, the data does not contain personally identifiable information. Therefore, we won’t be able to determine if casual riders live in the Cyclistic service area or if they have purchased multiple single passes.

There are 13 columns for each file. The names and sequence are consistent throughout all files.

|  |  |  |
| --- | --- | --- |
| Columns | Type | Note |
| ride\_id | String | Unique for each ride |
| rideable\_type | String | Options: electric\_bike, classic\_bike, docked\_bike |
| started\_at | Datetime |  |
| ended\_at | Datetime |  |
| start\_station\_name | String |  |
| start\_station\_id | String |  |
| end\_station\_name | String |  |
| end\_station\_id | String |  |
| start\_lat | Decimal |  |
| start\_lng | Decimal |  |
| end\_lat | Decimal |  |
| end\_lng | Decimal |  |
| member\_casual | String; categorical | Options: casual or member |

**PROCESS**

Import every file into SQL Server and begin the cleaning process by

* Removing rows that does not contain end\_station AND end\_lat/lng
* Removing rows that have start\_date >= end\_date
* Cleaning the station\_id and station\_name in both start and end station
* Removing rows that does not contain station\_id and cannot be populated using lat/lng data
* Removing rows that have end\_lat and end\_lng equal to 0
* Removing rows with rides under 1 minute.

After the cleaning process is done, 2 rows are being added to the data to help the analysis process.

|  |  |  |
| --- | --- | --- |
| Columns | Type | Note |
| day\_of\_week | Integer | 1-7 ; 1=Monday, 2=Tuesday, … |
| ride\_length | Integer | The difference between start and end time in second |

**ANALYZE**

Conduct summary statistics and Exploratory Data Analysis using SQL. Then, export the cleaned data into CSVs. Combine all monthly data using R by importing the separate monthly data into one single row (saved to one CSV).

**SHARE**

The insights are shared using Tableau as the visualization tool to showcase the difference in pattern between casual and member riders.

There are a few insights from the analysis:

* There are more rides in comparison to casual rides. This is consistent throughout the year.
* The lowest number of rides occur in January & February, while most rides occur in May to September.
* Casual rides occur more on the weekend while member rides, despite being quite consistent throughout the week, occur more throughout week days.
* There is a significant decrease in average ride\_length for casual rides on November and December.
* There are 5 popular station for casual rides. This is the case for start\_station and end\_station. The stations are Streeter Dr & Grand Ave, DuSable Lake Shore Dr & Monroe St, Millennium Park, Michigan Ave & Oak St, DuSable Lake Shore Dr & North Blvd. However this is not the case for member rides.

**ACT**

Based on analysis, there are a few recommendations from the pattern to help convert casual riders into member riders.

* Special promotions such as discounted price for first member purchase especially during May to September.
* Creating a special price for member during the weekend which will attract casual riders since casual riders ride more on the weekend.
* Casual riders tend to have higher average ride length compared to member. Create a special price for longer ride for member.
* Set up counter and representatives on the 5 top stations for casual riders to introduce our member program and the benefit on becoming member.