| **Company** | **Cyclistic** |
| --- | --- |
| **Author** | **Usborn Ocampo** |
| **Date** | **12/22/2022** |
| Business Task | This case study will analyze the difference between casual rider and annual member rider habits in order to convert casual riders into annual members. |
| Data Description | The data was obtained from by Motivate International Inc. and includes data of each ride from the last 12 months, including start and end times/locations, bike information, and rider information. The data was split into several files, with each file containing the data for one month. The data is reliable as it is from the company directly, and it is current as it contains ride information from the past 12 months. |
| Data Processing Notes | I examined the data and deemed the data to be well-formatted with no errors. There were some data points missing, most notably that some of the sheets had the start and end station data missing. However, this data point is not used for my analysis.  I uploaded all files into Bigquery and used the “UNION ALL” command to combine the data into one table.  Then, using “ALTER TABLE” and “UPDATE” commands, I created a column to calculate the ride duration based on the start and end times, and another column to note the day of the week for each ride using the start date data.    I also used the TRIM function to trim whitespace off of relevant string data.    Finally, I dropped columns that were not relevant (start/end station name and id data). |
| Analysis | I used SQL to extract several subsets and summaries of the data. For example, this query gave me a table of average ride length per hour of day separated by rider type.    I analyzed data to look at the differences in average ride times between casual and member riders, total rides by day and by month for casual and member riders, and average ride times by day and by month for casual and member riders. Below are some key points I found in my analysis:   * Members ride more during the week, but both groups ride equally on weekends. This could be because members use the bikes to commute, while casual riders ride recreationally. * Casual riders tend to ride longer than members. * Members ride more throughout the year, but both groups ride less in the winter and more in the summer. The exception is that members rode the most during the month of November. There were no duplicate ride IDs in either November dataset, so this looks to be accurate. * Casual riders consistently rode longer than members at any hour of the day and at any month of the year, riding longest in the early hours of the morning and in the first half of the year. * Members consistently nearly the same amount of time no matter the time of day or the time of year.   Visualizations are included in the presentation slide deck. Visualizations were made by Tableau, Google Sheets, and R. |
| Recommendations | 1. Consider additional pricing tiers for an annual plan that provides value for riders who ride recreationally. 2. Focus marketing efforts on spring and summer months to reach the most casual riders. 3. Consider qualitative data collection to find out why casual riders don’t opt for a membership. |
| Resources | Data:   * [Divvy Trip Data](https://divvy-tripdata.s3.amazonaws.com/index.html)   Tools:   * [Google Sheets](https://sheets.google.com/) * [Bigquery](https://cloud.google.com/bigquery) * [Tableau Public](https://public.tableau.com/app/discover) * [RStudio](https://rstudio.cloud/) |