**DATA ANALYSIS CASE STUDY- CYCLISTIC BIKE SHARING PROGRAM USING EXCEL**

In 2016, Cyclistic introduced a fruitful bike-sharing scheme. Since then, the program has expanded to include over 5000 bicycles that are locked into a network of 692 stations throughout Chicago and geotracked. The goal of the marketing strategy has always been to appeal to a wide range of consumer categories and increase overall awareness. This process has always been made easier by the price plans' flexibility, which include single-ride, full-day, and annual membership options. It has been determined that the high annual membership count would yield higher profits than the other two pricing schemes. The data analysis team must comprehend how annual members and casual riders utilize the Cyclistic bikes differently in order to create a marketing strategy that aims to convert casual riders to annual members.

Understanding the different ways annual members and casual members use the bikes would help see if the difference in the uses would make more sense for casual riders to become annual members allowing for an easy transition.

The key stakeholders involved in this project include; The Director of marketing and Cyclistic executive team.

To answer this question, I would be leveraging on the historical trip data in order to analyze and identify trends.

**DATA ANALYSIS PROCESS**

Firstly, the data which was for 12 months of the year 2023 was downloaded and observed using Excel. These data was merged to a single worksheet using Query tool on Excel. I proceeded to clean and prepare the data for analysis. The steps done for the cleaning process were; checking for duplicates which none was found, observing for missing rows which a lot was observed. This missing rows were chosen to be deleted instead of filled as removing this missing data was not going to affect the result of this data as a result of the good amount of complete data still present. I then proceeded to format some columns of the data to be aligned and then started by performing descriptive analysis of the data.

I started by creating another column called ride\_length that calculated the length of each ride by substracting the column started\_at from ended\_at. This Column was formatted as Hours:Minutes:Seconds

Another Column was created called day\_of\_the\_week to determine what day of the week, the users rode the bikes.

Another Column called the distance in miles was created which calculated the miles covered by the riders based on the longitude and latitude provided in the data.

**RESULTS**

This is the result of the descriptive analysis performed on the ride\_length column in the data

|  |  |
| --- | --- |
| Mean | 00:16:57 |
| Standard Error | 00:03:07 |
| Median | 00:11:01 |
| Range | 01:34:06 |
| Minimum | 00:01:46 |
| Maximum | 01:35:52 |
| Sum | 12:25:39 |
| Count | 44 |

From the table above, it can be determined that the Minimum ride length was 1minute 46 seconds and the Maximum is 1hour 35 minutes 52 seconds, a mean of 16minutes 57seconds.

The mode of the day\_of\_the\_week was also calculated to be 3 being Tuesday.

Further analysis was performed to get a deeper understanding of how Casual differ from Members.

These are the insights gathered;

**Distance Covered:**

* Members cover a significantly greater total distance (592,297 miles) compared to casual users (214,712 miles).
* For each bike type, members cover more distance except for docked bikes, which are exclusively used by casual users.

**Average Ride Length:**

* Casual users have a higher average ride length (0.013590222 miles) compared to members (0.007395567 miles).
* This indicates that while members take more trips, casual users tend to ride for longer distances on average per trip. This also justified by the result of the maximum ride length with Casual users being at the top

**Analysis by the days of the week**

**Member Rides by Day:**

* + The highest number of rides by members is on day 3 (102,165 rides).
  + The lowest number of rides by members is on day 7 (65,656 rides).

**Casual Rides by Day:**

* + The highest number of rides by casual users is on day 7 (39,102 rides).
  + The lowest number of rides by casual users is on day 2 (23,083 rides).

**Overall Rides by Day:**

* + The highest overall number of rides is on day 3 (130,374 rides).
  + The lowest overall number of rides is on day 1 (94,757 rides).

1. **Members:**
   * Members consistently have higher ride counts throughout the week compared to casual users.
   * The ride count peaks on day 3 and gradually declines towards the end of the week.
2. **Casual Users:**
   * Casual users show a different pattern, with the highest number of rides on day 7, possibly indicating more leisure or weekend activity.
   * The ride count for casual users is lowest on day 2.
3. **Overall Trends:**
   * The overall pattern shows higher ride counts during the mid-week (days 2 to 5) and lower counts towards the start and end of the week.

**DISCUSSION AND RECOMMENDATIONS**

From the analysis conducted, Casual riders have a high average ride length compared to Members despite members having a greater distance covered. Members tend to use bikes more for consistent, possibly commute-related activities throughout the week, with a mid-week peak.

Casual users show a significant increase in rides towards the weekend, indicating a potential preference for leisure riding. The data also provided that casual riders uses the docked bikes unlike members and leveraging on this could be beneficial to the company

My recommendations to convert Casual riders to members would include providing more incentives or programs for casual riders that could benefit these users especially when it is used during the weekends.