**Cyclistic Bike-Share Case Study – Google Data Analytics**

Data Analyst: Jessica Byrd Last updated: 01/02/2023

**Scenario**

This case study is on Cyclistic, a bike-share company in Chicago, IL that started in 2016. Cyclistic has since then grown and now operates 5,824 bicycles and 692 stations, which can be used to rent or return a bike at any time and location. As of now, marketing strategies have been developed to focus on building general awareness and appealing to broad customer segments. One successful approach they have used is offering flexible pricing plans: single-ride passes, full-day passes, and annual memberships. Single-ride and full-day passes are categorized as casual riders and customers with annual memberships are Cyclistic members.

Cyclistic’s financial analysts recently discovered that members are significantly more profitable that casual riders. These insights have led the Director of Marketing, Lily Moreno, to believe that the future of the company involves increasing the number of memberships and sees viability in gearing marketing campaigns towards converting current casual riders into members instead of targeting completely new customers. Lily Moreno’s goal for the future of the company has been clearly set *Design marketing strategies aimed at converting casual riders into annual members.*

There are three questions we need answered to 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?*

The question we will be answering in the following analysis is: *How do annual members and casual riders use Cyclistic bikes differently?*

**Data Collection**

*The* [*data*](https://divvy-tripdata.s3.amazonaws.com/index.html) *used in this analysis has been made available by Motive International Inc under* [*license*](https://ride.divvybikes.com/data-license-agreement)*. Cyclistic reviewed the data and has assured it to be dependable and appropriate to help answer our business question. Data privacy issues prohibit us from using riders’ personally identifiable information, however no personal information can be found in the files that can cause any data privacy issues.*

The data used in the analysis is from October 2021- September 2022. Each months’ data was individually downloaded into a csv file and extracted. Each csv file was then saved as an excel workbook file and placed into a different subfolder.

**Data Cleaning and Manipulation**

Columns provided within each file:

|  |  |
| --- | --- |
| ride\_id | unique 16-character identifier for every rental log |
| rideable\_type | type of bike rented: electric, classic, or docked |
| started\_at | time bike was rented |
| ended\_at | time bike was returned |
| start\_station\_name | street name of station where bike was rented |
| start\_station\_id | unique identifier of station where bike was rented |
| end\_station\_name | street name of station where bike was returned |
| end\_station\_id | unique identifier of station where bike was returned |
| start\_lat | latitude of start station |
| start\_lng | longitude of start station |
| end\_lat | latitude of end station |
| end\_lng | longitude of end station |
| member\_casual | states whether casual rider or member completed rental |

Columns created for each file:

|  |  |
| --- | --- |
| ride\_id\_length | created by using len() on ride\_id; counts the number of characters in string |
| month | created by using month() on started\_at; dictates what month the rental occurred; 1-12, 1 = January |
| day\_of\_week | created by using weekday() on started\_at; dictates what day of the week the rental occurred; 1-7, 1 = Sunday |
| hour\_taken | created by using hour() on started\_at; dictates that hour of the day the rental occurred; 0-23, 0 = 12AM |
| hour\_returned | created by using hour() on ended\_at; dictates the hour of the day the returned occurred; 0-23, 0 = 12AM |
| ride\_length | created by subtracting started\_at column from ended\_at column; dictates the length of the ride |

Data cleaning and manipulations completed:

|  |  |
| --- | --- |
| start\_station\_name  start\_station\_id | Both columns were removed for missing too much information. Enough data is provided to complete analysis without them. Further analysis can be completed to understand why information is not being saved. |
| end\_station\_name  end\_station\_id | Both columns were removed for missing too much information. Enough data is provided to complete analysis without them. Further analysis can be completed to understand why information is not being saved. |
| start\_lat, start\_lng  end\_lat, end\_lng | Columns were removed for missing information. Enough data is provided to complete analysis without them. Although not a lot of fields were missing data, data is irrelevant without station names and/or id. |
| ride\_length | Column was filtered and any ride entries with ride lengths giving an error were removed. Error occurred when end times logged before start times. Rides cannot end before they begin, data was not accurate therefor removed. Further analysis can be completed to understand why time logs are not accurate.  Column was filtered and any ride entries at zero ride lengths were removed. Rides cannot be at zero. There can be a possible error in the time stamps, an error occurring at time when bikes are being rented, or other. Further analysis can be completed to understand why a significant number of rides have ride lengths at zero.  Column was filtered and any ride entries greater than or equal to 24:00:01 ride lengths were removed. Analysis is comparing rider history for casual riders and members. Casual riders have the option to purchase single-ride passes and full-day passes. Analysis is assuming full-day passes are provided as an option because single-ride passes are not meant for a full day ride. Therefore, data analysis will be kept under 24 hours for both members and casual riders. |
| ride\_id | Column was checked for duplicates since ride\_id entries should be unique. |
| rideable\_type | Column was removed from analysis. Not considered relevant information based on the analysis we will be conducting. |
| started\_at  ended\_at | Columns were filtered to check if any entries were not in a time format. Any entries not showing as time formats were removed. |
| member\_casual | Column was filtered to check if anything other than casual or member. |
| ride\_id\_length | Column was filtered and any entries that were shorter or longer than 16 characters were removed. |
|  | All columns were checked for blanks. |

**Data Analysis and Findings**

All twelve excel files were imported into a single excel worksheet and transformed into a data model using Power Pivot. Pivot tables and their corresponding charts were then created to observe the distinction between casual riders’ and members’ use of Cyclistic’s biking services from October 2021- September 2022. The tables and charts break down number of rides as well as average ride lengths by year, month, day, and hour.

|  |  |  |
| --- | --- | --- |
| **Number of Rides: Year** | | |
|  | **ride count** | of total |
| casual | **2,395,924** | 41.16% |
| member | **3,425,421** | 58.84% |
| total | **5,821,345** | 100% |

**Year**

**Number of Rides: Year** table and graph demonstrate the number of rides taken by casual riders and members and the percentages of how those numbers are allocated in the yearly total.

|  |  |
| --- | --- |
| **Average Ride Length: Year** | |
|  | average time |
| casual | 0:22:26 |
| member | 0:12:29 |
| total | 0:16:35 |

**Average Ride Length**: Year table demonstrates the average ride length for casual riders, members, and their combined average for the year.

A total of **5,821,345** rides were taken with an average ride length of **0:16:35**. **2,395,924** of thoserides were taken by **casual riders** with an average ride length time of **0:22:26**. **Members took 3,435,421 rides**, and they averaged a ride length time of **0:12:29**.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Casual Riders Number of Rides: Monthly** | | | | |
|  | ride count | % of rider total | % of total | |
| **2021** |  |  | |  |
| Oct | 256,766 | 10.72% | | 4.41% |
| Nov | 106,717 | 4.45% | | 1.83% |
| Dec | 69,581 | 2.90% | | 1.20% |
| **2022** |  |  | |  |
| Jan | 18,448 | 0.77% | | 0.32% |
| Feb | 21,344 | 0.89% | | 0.37% |
| Mar | 89,634 | 3.74% | | 1.54% |
| Apr | 126,099 | 5.26% | | 2.17% |
| May | 279,769 | 11.68% | | 4.81% |
| Jun | 368,073 | 15.36% | | 6.32% |
| Jul | 405,237 | 16.91% | | 6.96% |
| Aug | 358,180 | 14.95% | | 6.15% |
| Sep | 296,076 | 12.36% | | 5.09% |
| total | 2,395,924 | 100.00% | | 41.16% |

**Month**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Members Number of Rides: Monthly** | | | | |
|  | ride count | % of rider total | | % of total |
| **2021** |  |  |  | |
| Oct | 373,801 | 10.91% | 6.42% | |
| Nov | 252,920 | 7.38% | 4.34% | |
| Dec | 177,732 | 5.19% | 3.05% | |
| **2022** |  |  |  | |
| Jan | 85,207 | 2.49% | 1.46% | |
| Feb | 94,150 | 2.75% | 1.62% | |
| Mar | 194,090 | 5.67% | 3.33% | |
| Apr | 244,749 | 7.15% | 4.20% | |
| May | 354,288 | 10.34% | 6.09% | |
| Jun | 399,937 | 11.68% | 6.87% | |
| Jul | 417,320 | 12.18% | 7.17% | |
| Aug | 426,788 | 12.46% | 7.33% | |
| Sep | 404,439 | 11.81% | 6.95% | |
| total | 3,425,421 | 100.00% | 58.84% | |

**Number of Rides: Monthly** tables demonstrate the number of rides taken by casual riders and members for each month and the percentages of how those numbers are allocated in yearly total for the specific rider type as well as the percentage for the overall yearly total.

**Number of Rides: Monthly** chart compares casual riders and members ride count for each month.

A total of **2,395,924** rides were taken by **casual riders**. The three busiest months for casual riders were **June 2022** (368,073), **July 2022** (405,237),and **August 2022** (358,180). Casual riders took **1,131,490** rides during these months. When comparing casual riders three least busiest months, **December 2021** (69,581), **January 2022** (18,448), and **February 2022** (21,334), **109,363** rides were taken during the three-month period.

**Members** took a total of **3,425,421** rides. Their three busiest months were **July 2022** (417,320), **August 2022** (426,788), and **September 2022** (404,439), totaling **1,248,547** rides. The three least busy months for members were December 2021 (177,732), January 2022 (85,207), and February 2022 (94,150), a totaling **357,089** rides. Members took more rides every month than casual riders.

|  |  |  |  |
| --- | --- | --- | --- |
| **Average Ride Length: Monthly** | | | |
|  | casual | member | total |
| **2021** |  |  |  |
| Oct | 0:22:51 | 0:12:14 | 0:16:34 |
| Nov | 0:18:44 | 0:11:06 | 0:13:22 |
| Dec | 0:18:11 | 0:10:49 | 0:12:54 |
| **2022** |  |  |  |
| Jan | 0:17:26 | 0:11:37 | 0:12:39 |
| Feb | 0:19:34 | 0:11:02 | 0:12:36 |
| Mar | 0:24:13 | 0:11:41 | 0:15:39 |
| Apr | 0:23:14 | 0:11:20 | 0:15:23 |
| May | 0:25:32 | 0:13:03 | 0:18:33 |
| Jun | 0:23:24 | 0:13:39 | 0:18:19 |
| Jul | 0:23:11 | 0:13:25 | 0:18:14 |
| Aug | 0:21:28 | 0:13:05 | 0:16:55 |
| Sep | 0:20:03 | 0:12:37 | 0:15:46 |
| total | 0:22:26 | 0:12:29 | 0:16:35 |

**Average Ride Length: Monthly** table demonstrates the average ride length for casual riders, members, and their combined average for every month.

**Average Ride Length: Monthly** chart compares the monthly ride lengths for casual riders and members.

**Casual riders** averaged a ride length time for the year of **0:22:26**. **March 2022** (0:24:13), **May 2022** (0:25:32), and **June 2022** (0:23:24) were the months where casual riders took the longest rides. **November 2021** (0:18:44), **December 2021** (0:18:11), and **January 2022** (0:17:26) were the months where casual riders took the shortest rides. Casual riders took longer riders every month over members.

**Members** averaged a ride length time for the year of **0:12:29**. **June 2022** (0:13:39), **July 2022** (0:13:25), and **August 2022** (0:13:05) were the months where members took the longest rides. They took the shortest rides during the months of **November 2021** (0:11:06), **December 2021** (0:10:49), and **February 2022** (0:11:02).

**Day**

**Number of Rides: Day of Week** tables demonstrate the number of rides taken by casual riders and members for each day of the week and the percentages of how those numbers are allocated in the yearly total for the specific rider type as well as the percentage for the overall yearly total.

**Number of Rides: Day of Week** chart compares casual riders and members ride count for each month; 1 = Sunday, 7 = Saturday.

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| |  |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | --- | | **Casual Riders Number of Rides: Day of Week** | | | | | | | |  | ride count | % of rider total | | | % of total | | | Sun | 403967 | | 16.86% |  | | 6.94% | | Mon | 279174 | | 11.65% |  | | 4.80% | | Tue | 275199 | | 11.49% |  | | 4.73% | | Wed | 281126 | | 11.73% |  | | 4.83% | | Thu | 306073 | | 12.77% |  | | 5.26% | | Fri | 351692 | | 14.68% |  | | 6.04% | | Sat | 498693 | | 20.81% |  | | 8.57% | | total | 2395924 | | 100.00% | 41.16% | | |   **Members Number of Rides: Day of Week** | | | | | | |
|  | ride count | | % of rider total | | % of total | |
| Sun | 393408 | 11.48% | |  | | 6.76% |
| Mon | 472865 | 13.80% | |  | | 8.12% |
| Tue | 541296 | 15.80% | |  | | 9.30% |
| Wed | 538279 | 15.71% | |  | | 9.25% |
| Thu | 530319 | 15.48% | |  | | 9.11% |
| Fri | 491246 | 14.34% | |  | | 8.44% |
| Sat | 458008 | 13.37% | |  | | 7.87% |
| total | 3425421 | 100.00% | | 58.84% | | |

A total of **2,395,924** rides were taken by **causal riders**. The busiest days of the week for casual riders were **Sundays** (403,967) and **Saturdays** (498,693). During these two days of the week a total of **902,660** rides were taken. **Mondays** (279,174) and **Tuesdays** (275,199) were the least busy days of the week for casual riders. A total of **554,373** rides were taken during these two days of the week. Saturdays and Sundays are the only two days of the week where casual riders took more rides than members.

**Members** took a total of **3,425,421** rides. The busiest days of the week for members were **Tuesdays** (541,296) and **Wednesdays** (538,279). **1,079,575** rides were taken during these two days of the week. **Sundays** (393,408) and **Saturdays** (458,008) were the least busy days of the week for members. **851,416** rides were taken between these two days of the week by members.

**Average Ride Length: Day of Week** table demonstrates the average ride length for casual riders, members, and their combined average for every day of the week.

|  |  |  |  |
| --- | --- | --- | --- |
| **Average Ride Length: Day of Week** | | | |
|  | casual | member | Total |
| Sun | 0:25:47 | 0:13:50 | 0:19:53 |
| Mon | 0:22:53 | 0:12:04 | 0:16:05 |
| Tue | 0:20:06 | 0:11:53 | 0:14:39 |
| Wed | 0:19:22 | 0:11:53 | 0:14:27 |
| Thu | 0:19:49 | 0:12:01 | 0:14:52 |
| Fri | 0:21:01 | 0:12:14 | 0:15:54 |
| Sat | 0:25:06 | 0:13:57 | 0:19:46 |
| total | 0:22:26 | 0:12:29 | 0:16:35 |

**Average Ride Length: Day of Week** chart compares ride lengths for casual riders and members by days of the week; 1= Sunday, 7= Saturday.

**Casual riders** averaged a ride length of **0:22:26** for the year. They took the longest rides on **Saturdays** (0:25:06) and **Sundays** (0:25:47) and the shortest rides on **Wednesdays** (0:19:22) and **Thursdays** (0:19:49). Causal riders took longer rides every day of the week compared to members.

**Members** averaged a ride length of **0:12:29** for the year. **Sundays** (0:13:50) and **Saturdays** (0:13:57) averaged the longest ride lengths for members. **Tuesday** (0:11:53) and **Wednesdays** (0:11:53) were their shortest days.

**Hour**

**Number of Rides: Hour of Day** tables demonstrate the number of rides taken by casual riders and members for every hour of the day and the percentages of how those numbers are allocated in the yearly total for the specific rider type as well as the percentage for the overall yearly total; 0 = 12:00:00 AM, 23 = 11:00:00 PM.

**Number of Rides: Hour of Day** chart compares casual riders and members ride count for each hour of the day; 0 = 12:00:00 AM, 23 = 11:00:00 PM.

A total of **2,395,924** rides were taken by **causal riders. 2:00:00 PM – 7:00:00 PM** (1,133,028) was the busiest time for them. Their least busy time of day is from **12:00:00 AM – 5:00:00 AM** (133,146). Casual riders took more rides than members from **12:00:00 AM – 3:59:59 AM** and **10:00:00 PM – 11:59:59 PM**.

**Members** took a total of **3,425,421** rides. **8:00:00AM, 3:00:00 PM – 7:00:00 PM** (1,586,862) was the busiest time for them. Their least busy time of day is from **12:00:00 AM – 5:00:00 AM** (124,093).

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Casual Riders Number of Rides: Hour of Day** | | | | | |
|  | ride count | | % of rider total | | % of total |
| 0 | 48425 | 2.02% | | 0.83% | |
| 1 | 31746 | 1.33% | | 0.55% | |
| 2 | 20136 | 0.84% | | 0.35% | |
| 3 | 11708 | 0.49% | | 0.20% | |
| 4 | 8194 | 0.34% | | 0.14% | |
| 5 | 12937 | 0.54% | | 0.22% | |
| 6 | 29379 | 1.23% | | 0.50% | |
| 7 | 52552 | 2.19% | | 0.90% | |
| 8 | 70219 | 2.93% | | 1.21% | |
| 9 | 75081 | 3.13% | | 1.29% | |
| 10 | 97736 | 4.08% | | 1.68% | |
| 11 | 126900 | 5.30% | | 2.18% | |
| 12 | 149332 | 6.23% | | 2.57% | |
| 13 | 156382 | 6.53% | | 2.69% | |
| 14 | 165244 | 6.90% | | 2.84% | |
| 15 | 182422 | 7.61% | | 3.13% | |
| 16 | 200880 | 8.38% | | 3.45% | |
| 17 | 224617 | 9.37% | | 3.86% | |
| 18 | 203161 | 8.48% | | 3.49% | |
| 19 | 156704 | 6.54% | | 2.69% | |
| 20 | 115441 | 4.82% | | 1.98% | |
| 21 | 99312 | 4.15% | | 1.71% | |
| 22 | 90043 | 3.76% | | 1.55% | |
| 23 | 67373 | 2.81% | | 1.16% | |
| total | 2395924 | 100.00% | | 41.16% | |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Members Number of Rides: Hour of Day** | | | | | |
|  | ride count | | % of rider total | | % of total |
| 0 | 36937 | 1.08% | | 0.63% | |
| 1 | 22937 | 0.67% | | 0.39% | |
| 2 | 13234 | 0.39% | | 0.23% | |
| 3 | 8121 | 0.24% | | 0.14% | |
| 4 | 9056 | 0.26% | | 0.16% | |
| 5 | 33808 | 0.99% | | 0.58% | |
| 6 | 92093 | 2.69% | | 1.58% | |
| 7 | 175765 | 5.13% | | 3.02% | |
| 8 | 207666 | 6.06% | | 3.57% | |
| 9 | 148708 | 4.34% | | 2.55% | |
| 10 | 141375 | 4.13% | | 2.43% | |
| 11 | 168543 | 4.92% | | 2.90% | |
| 12 | 193983 | 5.66% | | 3.33% | |
| 13 | 192392 | 5.62% | | 3.30% | |
| 14 | 190302 | 5.56% | | 3.27% | |
| 15 | 224719 | 6.56% | | 3.86% | |
| 16 | 295478 | 8.63% | | 5.08% | |
| 17 | 355585 | 10.38% | | 6.11% | |
| 18 | 292388 | 8.54% | | 5.02% | |
| 19 | 211026 | 6.16% | | 3.63% | |
| 20 | 148490 | 4.33% | | 2.55% | |
| 21 | 115436 | 3.37% | | 1.98% | |
| 22 | 88751 | 2.59% | | 1.52% | |
| 23 | 58628 | 1.71% | | 1.01% | |
| total | 3425421 | 100.00% | | 58.84% | |

**Average Ride Length: Day of Week** table demonstrates the average ride length for casual riders, members, and their combined average for every day of the week.

**Average Ride Length: Day of Week** chart compares ride lengths for casual riders and members by days of the week; 0 = 12:00:00 AM, 23 = 11:00:00 PM.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Average Ride Length: Hour of Day** | | | | |
|  | casual | member | Total | |
| 0 | 0:19:55 | 0:12:12 | 0:16:35 |
| 1 | 0:20:40 | 0:12:38 | 0:17:18 |
| 2 | 0:20:44 | 0:12:37 | 0:17:31 |
| 3 | 0:20:03 | 0:12:28 | 0:16:56 |
| 4 | 0:18:19 | 0:12:10 | 0:15:05 |
| 5 | 0:15:53 | 0:10:47 | 0:12:12 |
| 6 | 0:15:45 | 0:11:10 | 0:12:17 |
| 7 | 0:15:18 | 0:11:37 | 0:12:28 |
| 8 | 0:17:32 | 0:11:21 | 0:12:55 |
| 9 | 0:22:14 | 0:11:26 | 0:15:04 |
| 10 | 0:26:10 | 0:12:13 | 0:17:55 |
| 11 | 0:26:28 | 0:12:36 | 0:18:33 |
| 12 | 0:25:52 | 0:12:10 | 0:18:07 |
| 13 | 0:26:13 | 0:12:20 | 0:18:34 |
| 14 | 0:25:54 | 0:12:44 | 0:18:51 |
| 15 | 0:24:29 | 0:12:50 | 0:18:04 |
| 16 | 0:22:59 | 0:13:03 | 0:17:04 |
| 17 | 0:21:25 | 0:13:16 | 0:16:26 |
| 18 | 0:20:49 | 0:13:04 | 0:16:14 |
| 19 | 0:20:43 | 0:12:48 | 0:16:10 |
| 20 | 0:20:36 | 0:12:36 | 0:16:06 |
| 21 | 0:19:49 | 0:12:23 | 0:15:49 |
| 22 | 0:19:21 | 0:12:25 | 0:15:55 |
| 23 | 0:19:54 | 0:12:36 | 0:16:30 |
| total | 0:22:26 | 0:12:29 | 0:16:35 |

**Casual riders** averaged a ride length of **0:22:26** for the year. They averaged the longest rides from **10:00:00 AM – 3:59:59 PM** (0:25:51) and the shortest rides from **4:00:00 AM – 8:59:59 AM, 10:00:00 PM – 10:59:59 PM** (0:17:01).

**Members** averaged a ride length of **0:12:29** for the year. They averaged the longest rides from **2:00:00 PM - 7:59:59 PM** (0:12:58) and the shortest rides from **4:00:00 AM – 9:59:59 AM** (0:11:25).

**Data Insights and Recommendations**