How does a Bike-Share Navigate Speedy Success?

Oscar Leyva

2023-02-15

Introduction

Since Cyclistic's successful launch in 2016, they have provided 5824 bicycles into a network of 692 stations across Chicago. Customers unlock a bike from one station and return it to another at their journey's end. The easy-to-use system coupled with flexible pricing plans has made the business appealing. A distinction is made between customers based on what pricing plan they follow. Single-ride and full-day pass holders are classified as **Casual riders**. On the contrary, customers who purchase annual memberships are considered **Cyclistic members**. Now, Cyclistic wants to focus its next marketing campaign on converting Casual riders into Cyclistic members. Casual riders are familiar with how the Bike-Share operates and thus seems like an important focal point. To accomplish this we need to understand what differentiates these two customer types from each other. After understanding how their behaviors differ we can design a marketing campaign based on data. This approach will help maximize the potential growth a successful conversion could attain.

To help us with the data analysis Cyclistic has provided us with historical trip data to identify trends. After cleaning and transforming the data we can begin our search for critical differences. After that, we present our top three recommendations for a successful marketing program.

Data

Cyclistic data is located at https://divvy-tripdata.s3.amazonaws.com/index.html. For this analysis, we looked at all bike trips spanning the time period of February 2022 to January 2023. Since the data provided comes from a first party it boosts its credibility. The data consists of each observation being a completed bike ride. As stated by Cyclic it denotes all bike rides that were longer than 60 seconds to ensure no false starts are included. They also removed trips taken by staff as they cannot be classified as either of the two customer types. The data includes the location of the start and end of a trip as well as the day and time it started/ended. The rider's type is also collected as Member or Casual. With this, we can infer key differences between the customer types.

Data manipulation

Cyclistic stores monthly data in a separate file. Thus we had to download each zip file individually and then merge them together to get the full data set for our time period. Before merging we dropped columns that contained longitudinal and latitudinal for the start and end locations of each bike trip. We also dropped columns containing the start and end locations of the trips. A glimpse at the data shows that no NA values are present.

To prepare for the upcoming analysis we created four new columns. One of them is **trip_duration** which denotes the duration of the trip based on when it started and ended. This was done by subtracting the end time with the start time. The duration is in minutes. When checking for data validation it was found that 100 trips resulted in a negative trip duration. Investigating further shows that 40 of these occurred on November 6th. This could imply that a system error made the start and end times swap and thus calculating

the duration yields negative times. However, due to it likely being a system error and the possibility of other trip information also being wrong, we remove these trips. Since the data consists of over five million trips, removing 100 of them will not impact the business task in question.

The second column created is **weekday_started_at** which extracts what day of the week the bike was rented on. For the column **hour_started_at** we extract the hour the trip began on. In particular, this aided in the computation of the last column **part_of_day**. The column denotes if the trip started in the morning, afternoon, evening, or night. We used the time conventions found in the Britannica Dictionary. Morning is defined as the time between 5:00 and 11:59. Afternoon lasts until 16:59. Evening is until 21:59. Night hours are then from 22:00 to 4:59.

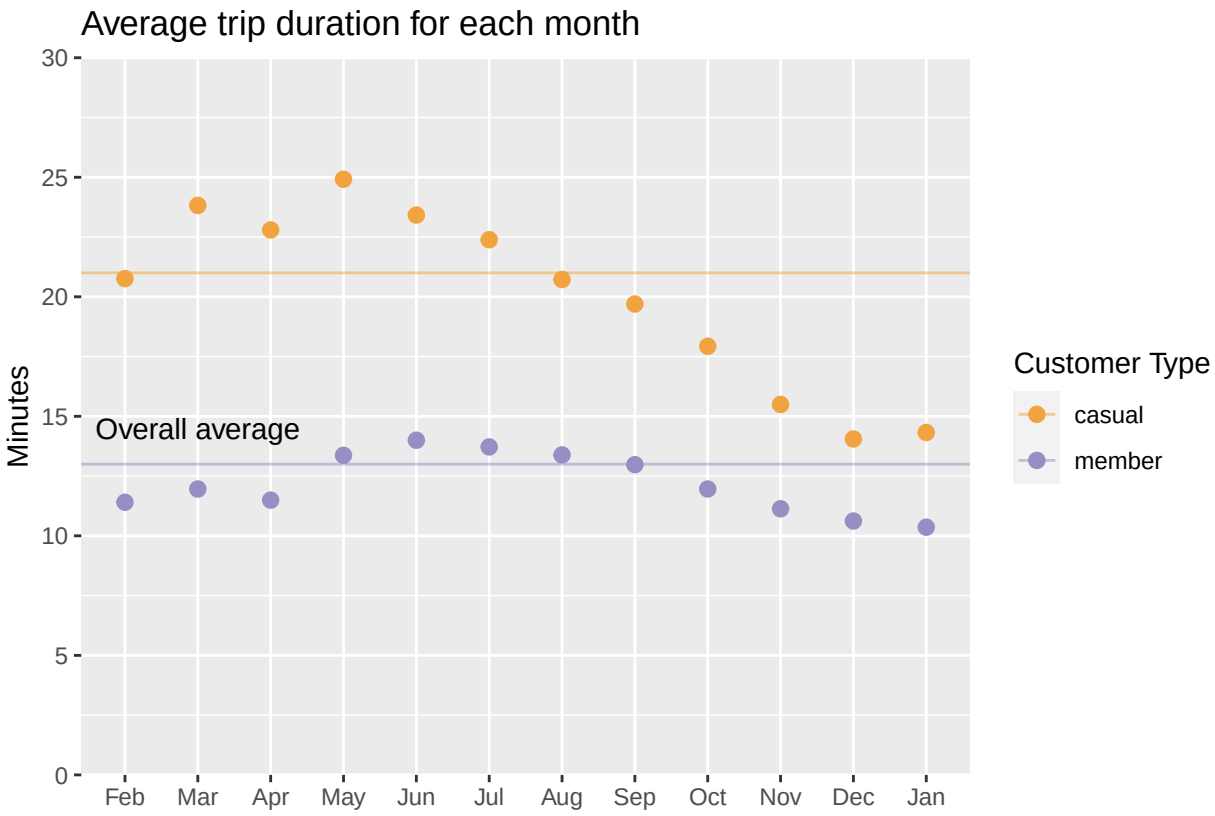
Finally, we converted the **rideable_type** and **member_casual** columns to categorical columns. In the data, there are three bike types. Docked, electric, and classic bikes. There are no trips recorded for members riding a docked bike. Casuals who rode a docked bike had an average trip duration of 123 minutes. The second-highest average trip duration of 29 minutes was for casuals who rode a classic bike. Thus, docked bike trips skew the data. Since there is no way in comparing the same bike type for members it would be unfair to include casuals riding docked bikes. Another argument for its removal is looking at the highest trip duration based on bike type and customer type. The highest trip duration for a casual riding a docked bike was 41400 minutes. The second and third highest trip duration was by a casual and member riding a classic bike, clocking in at 1560 minutes. Both of them being at the same time indicates that the trip duration hit a ceiling. For all the reasons above, we chose to remove the docked bike in the data which means that 178245 rides were removed.

Results

Trip duration

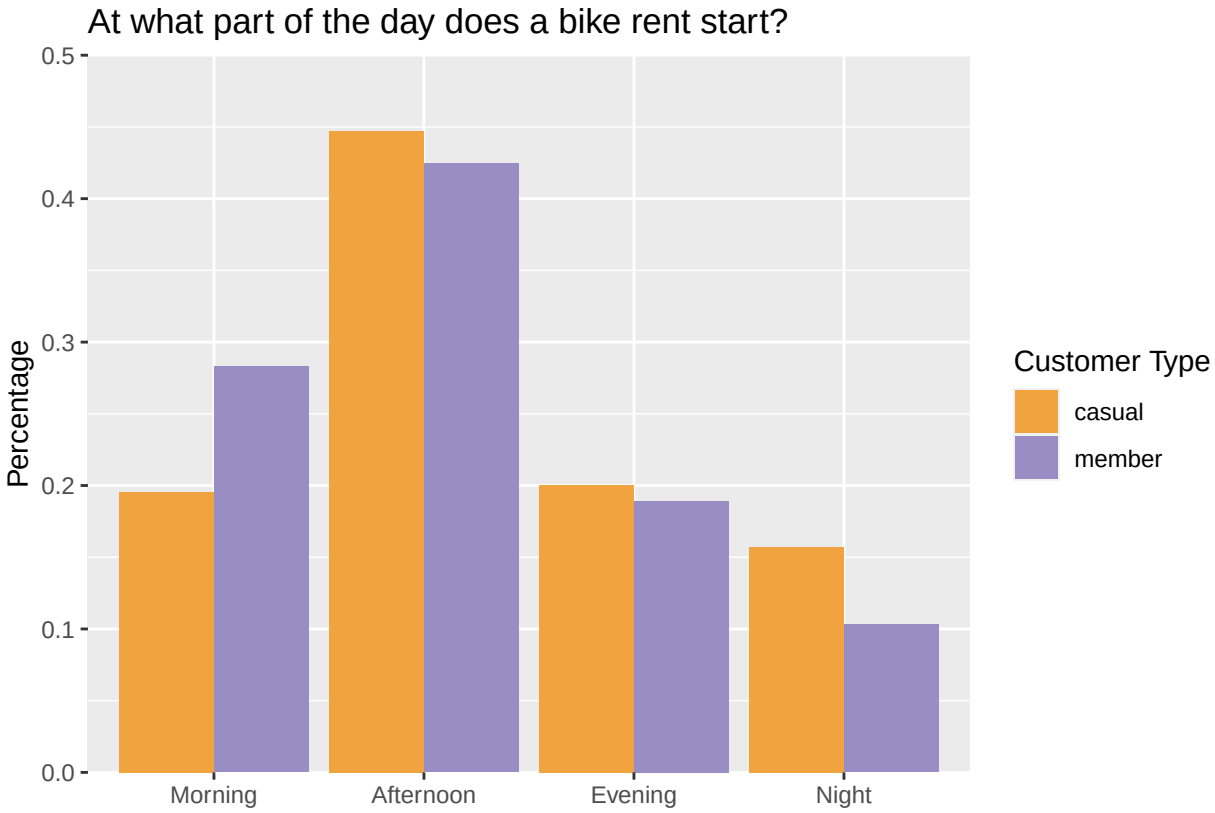
member_casual	mean	sd	median
casual	21	60	12
member	13	28	9

Trip duration seems to differ between the two customer types. Casual customers have an average trip duration of 21 minutes whereas members bicycle for an average of 13 minutes. The standard deviation implies that casual customers vary greatly in how long they rent a bike. Performing a statistical test shows that the difference in average between the customer types is not due to randomness.



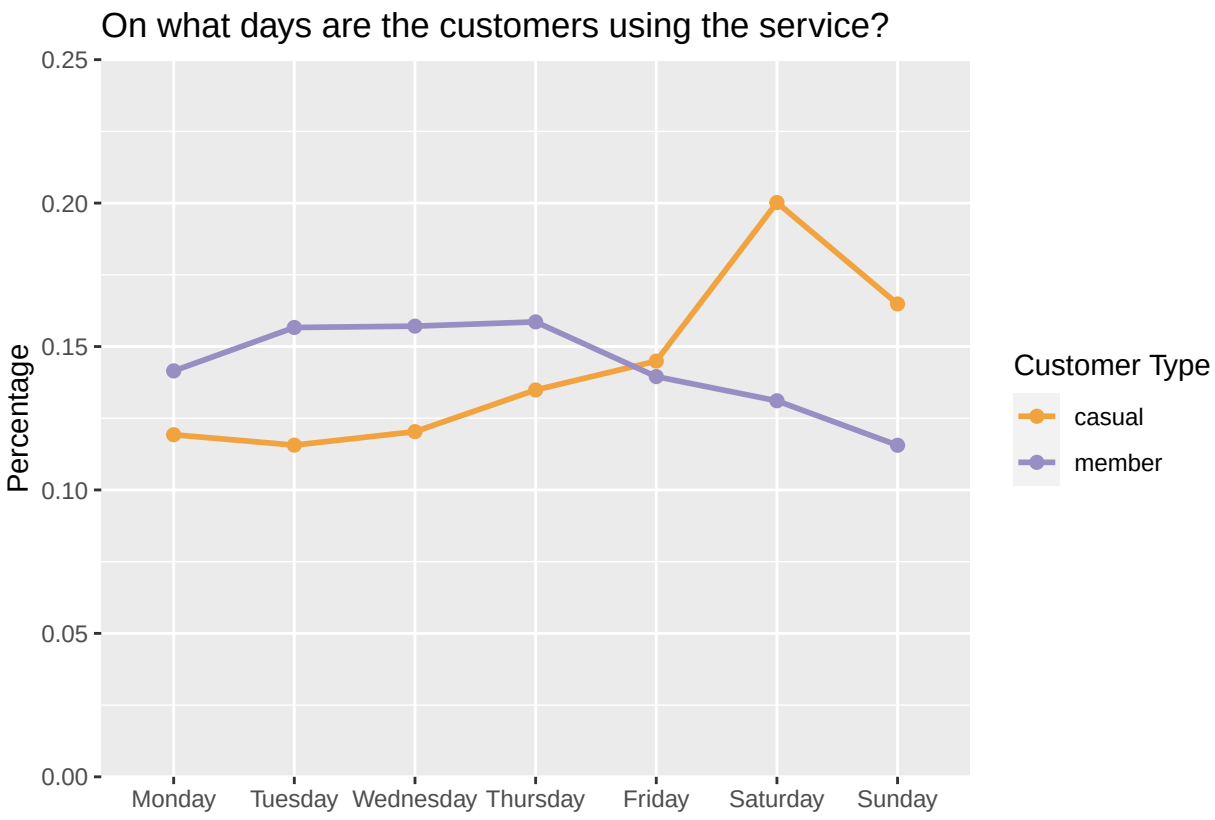
The average trip duration for casuals is greater than members during the months investigated. Member's trip duration remains consistent throughout the year. On the other hand, fluctuations can be seen for the trip duration among casuals. For example, in May the average trip duration is 25 minutes for casuals. This decreases to 14 in December which means that the trip duration is lowered by 11 minutes. For members, their highest average trip duration was in June, and their lowest was in January. The difference between the two is 3 minutes.

Part of the day



Members utilize the service in the mornings at a higher rate than casual customers do. On the contrary, casuals use the service more frequently at night compared to members. During the afternoon and evening, the usage is similar between casuals and members. The day with the largest difference between the two being on Saturday.

Day of the week the trips start in.



Members use the service more during working days. The only exception to this is on Friday when casuals use it more. Casuals use Cyclistic during the weekend at a higher rate than members.

Summary

Let us summarize our findings.

- Casuals have on average a higher trip duration than members. Their trip duration varies among them as their standard deviation was 60 minutes. Members are found to use the system for short trips, having an average trip duration of 13 minutes.
- The average trip duration depends on the month. Casuals have a higher-than-average trip duration between the months of March and July. For members, this applies to the months from May to August.
- Members use Cyclistic more often than casuals during the morning. During the afternoon and evening usage does not differ between the two customer types. At night casuals use the service at a higher rate than members.
- Compared to casuals, members use the service more between Monday to Thursday. On Fridays, both customer types have similar usage rates. On weekends, casuals rent bikes more often.

Recommendations

The marketing program should take an approach that breathes adventure. Since casuals rent a bike for longer the system should make that option more appealing. This can be done by **expanding the number of**

stations found in Chicago. The expansion should focus on covering a wider area rather than introducing new stations in the most central avenues.

Electric bikes are not convenient for long trips due to them needing to charge at certain intervals. **Increasing the supply of classic bikes** would help out casuals who prefer longer trips or ones that have few stations in between.

During the night and on the weekends when there are a lot of people walking on the streets and visibility is reduced. To help out casuals who use Cyclistic at these time periods **a front lamp could be attached to bikes**. This would encourage the use of the system rather than taking a taxi or public transportation.

With these changes, casuals might use the system more frequently. By using Cyclistic more an annual membership would be a more lucrative option than buying single-ride tickets or day passes multiple times.