Cyclistic Case Study Report

Business task

Analyse how annual members and casual riders use Cyclistic bikes differently, to determine marketing strategies that can be used to convert casual riders to annual members.

Data sources used

The data is publicly licenced (licence: divvybikes.com/data-license-agreement) from Divvy divvybikes-marketing-staging.lyft.net/system-data which gives details of every ride from Cyclistic since 2013. Each entry contains information such as the start and times of each ride, bike type, start and end locations, and the rider's membership status. Data is kept private through anonymisation. This anonymisation presents the limitation of not being able to attribute multiple trips to the one user, so we're limited in that way of how the Cyclistic service is used. Due to time/computing power constraints, most analysis is of the data from Q1 (January-March) of 2019 and 2020. Further investigation could be made as to how different times of the year affect the

Data cleaning and manipulation

Several cleaning steps were required to make data consistent, such as converting time data to recognisable datetime formats. There was also cleaning required to make numeric values for the longitude and latitude in some of the data, though these were not used in the final analysis.

Example clean log of 202309-divvy-tripdata:
By looking at maximum and minimum values in each column, it was found that some rows (838) have 0 as their latitude and longitude values. None of these rows had end station values so couldn't be filled in that way. Most of these rides had ride times of just under an hour, so there must be some bug causing this for riders around this mark. Rows subsequently removed.
50 rows removed where start time is after end time.
14310 rows were removed where ride time is less that 1 minute.

The ride length of each ride was calculated, a crucial feature for analysing how riders use the bikes. A feature was also created for the day of the week, which may have important links to the behaviours of the different types of users.

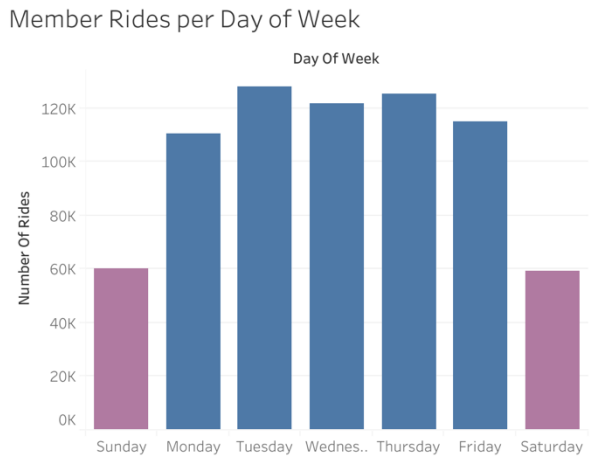
Analysis Summary

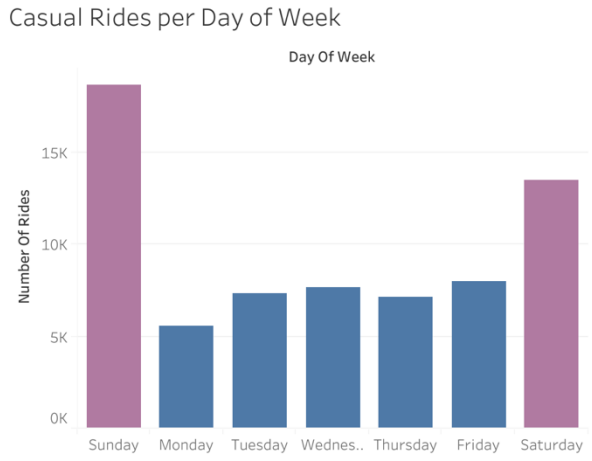
Data was organised by day of the week, and the number of trips taken by riders, grouped by rider type (member vs casual) were calculated, as well as the average ride times for each type of rider.

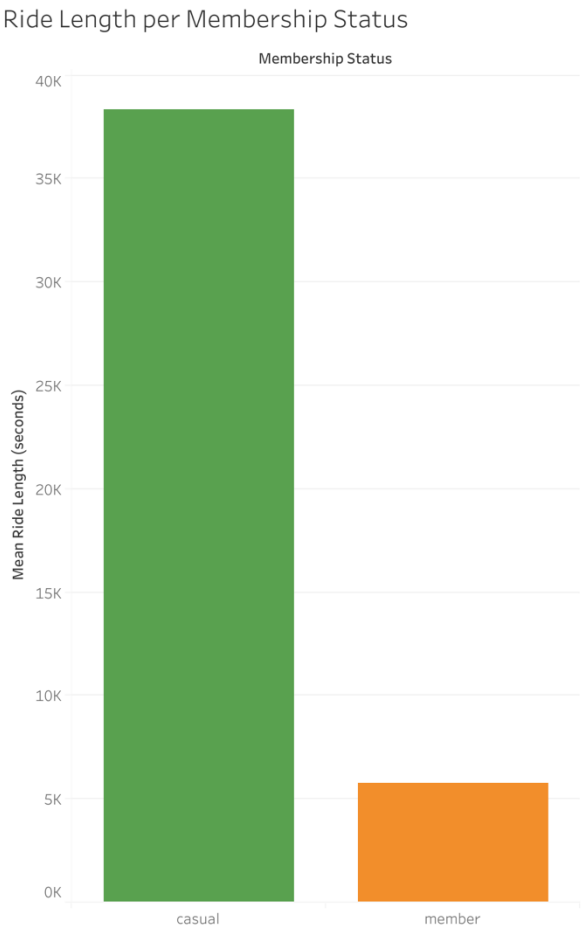
Analysis shows that trips from casual riders are on average longer than members. From the data in Q1 2019 and Q1 2020, the average ride time for casual riders was 26 minutes 4 seconds, almost double the 13 minutes 25 seconds for members.

Analysis also showed that members ride more often on weekdays, whereas casual riders ride more on the weekends. Around 60k trips were taken per weekend day by members in total over the six months analysed, whereas double that, around 120k trips were taken per weekday. Around 15k trips were taken per weekend day by members in total over the six months analysed, whereas half that, around 7.5k trips were taken per weekday.

Visualisations and key findings







As shown in the above figures:

- (Left) Cyclistic members tend to ride about twice as much on weekdays as on weekends, whereas casual riders are the opposite. This would indicate that members are taking advantage of Cyclistic bikes for their daily commutes, whereas casual riders are more likely to use the bikes for leisure.
- (Right) Members are more likely to use the bikes for quick, trips (likely the daily commute), whereas casual riders take longer, more leisurely rides.

Recommendations

Based on these findings, it is recommended to the executive team that new marketing campaigns:

1. Highlight the convenience of having a membership for the daily commute, on top of the less regular, leisurely trips already being taken.
2. Highlight the convenience of having a membership for short trips in general, on top of the longer trips already being taken.
3. Emphasise the time efficiency, cost savings and health benefits of relying on Cyclistic to make these short, regular trips on a membership.