Summary

Given the data set that was provided, we concluded that a driver's lifetime value is composed of their projected lifetime with respect to the average lifetime of a Lyft driver, a distance calculation based on the total ride distance and a prime time calculation based off of the frequency that each driver drives during prime time hours. With the first part of this calculation, we predicted that the average lifetime of a driver is approximately 15 weeks. The two main segments of drivers that we identified were the ones that drove on a consistent basis and ones that did not. Consistency was defined in this case as a relatively steady cumulative time driven per week. The main recommendation that we have is to focus on different types of incentives for the different types of distribution that a driver can have. Details of our approach are outlined in the report below.

Driver's Lifetime Value

The following equation was used to calculate the lifetime value for each driver:

$$\frac{projected\ driver\ lifetime}{AVG(projected\ driver\ lifetimes)}$$

$$\times$$

$$\frac{\frac{\Sigma\ ride\_distance}{today - onboard\_date}}{AVG(\frac{\Sigma\ ride\_distance}{today - onboard\_date})} + \frac{\frac{\Sigma(prime\_time/100) \times COUNT(prime\_time)}{SUM(total\ rides)}}{AVG(\frac{\Sigma(prime\_time/100) \times COUNT(prime\_time)}{SUM(total\ rides)})}$$

*Where $today$ is the last ride date provided in the dataset

The first term calculates the distance travelled by each driver divided by the average distance for all drivers. Drivers with a higher ratio drive more frequently and thus add more value to Lyft. The second term calculates a value that represents how often a driver drives during prime time divided by the average of all drivers. Drivers that take rides more frequently during times of high prime time would have a higher ratio for the term, increasing their lifetime value. The last term uses the projected lifetime of a driver as calculated in question 3. This gives us a lifetime measurement for the driver compared to others.

Once computed for every driver, this provides us with a comparative measure to analyze the drivers and quantify their lifetime value.

Factors Affecting Driver's Lifetime Value

The three factors that we determined affect a driver's lifetime value is the total distance that they've traveled to complete trips, the amount and frequency of prime time rides and the amount of time they've been with the company.

Rather than simply using the number of rides a driver takes, we chose to use distance values to compensate for a driver consistently getting shorter or longer rides. Since all data is collected in San Francisco, the distance values serve as a good measurement of how much a driver works. The prime time frequency is also an important factor that adds to the value of the driver. High prime time values indicate the driver took a ride during hours of peak demand, reducing customer wait times and increasing customer satisfaction. Lastly we included the amount of time within the dataset that a driver is projected to stay with Lyft.

We omitted the dataset on ride_timestamps because we didn't believe that it was a decision that a driver could control (eg. arrival_time can be affected by many external factors), and thus should not be factored into driver value.

Average Projected Lifetime of a Driver

By analyzing the cumulative time per week that drivers spent driving for Lyft, we generalized the trends that we saw into two categories. The first being drivers who drove with a uniform distribution (see Figure 1) and drivers who drove with a normal distribution (see Figure 2).

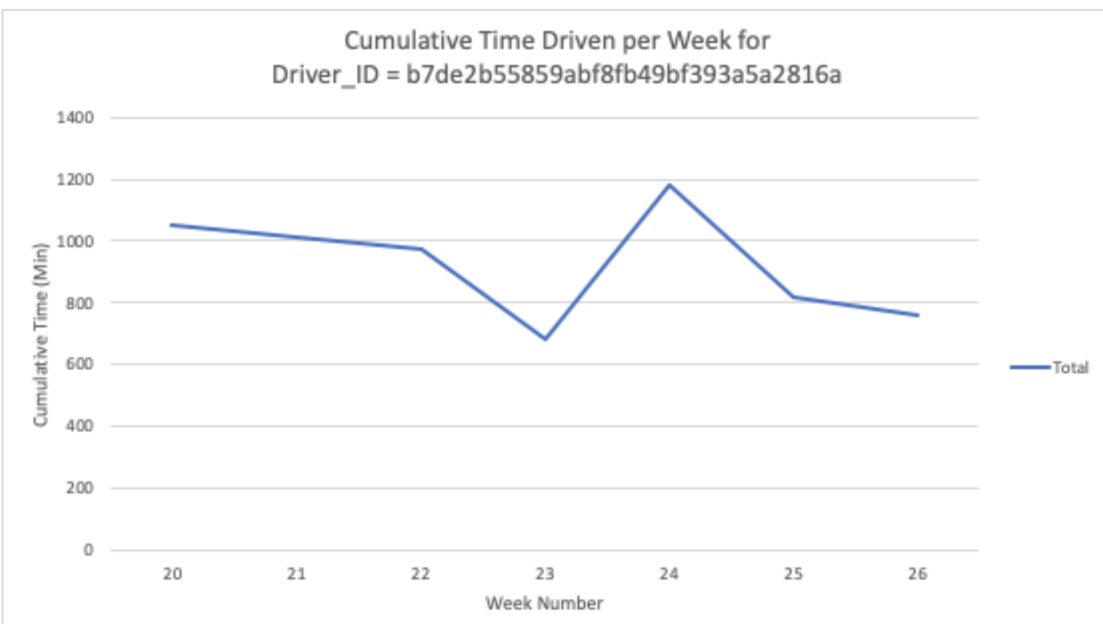

Figure 1: Driver that is showing an approximate uniform distribution

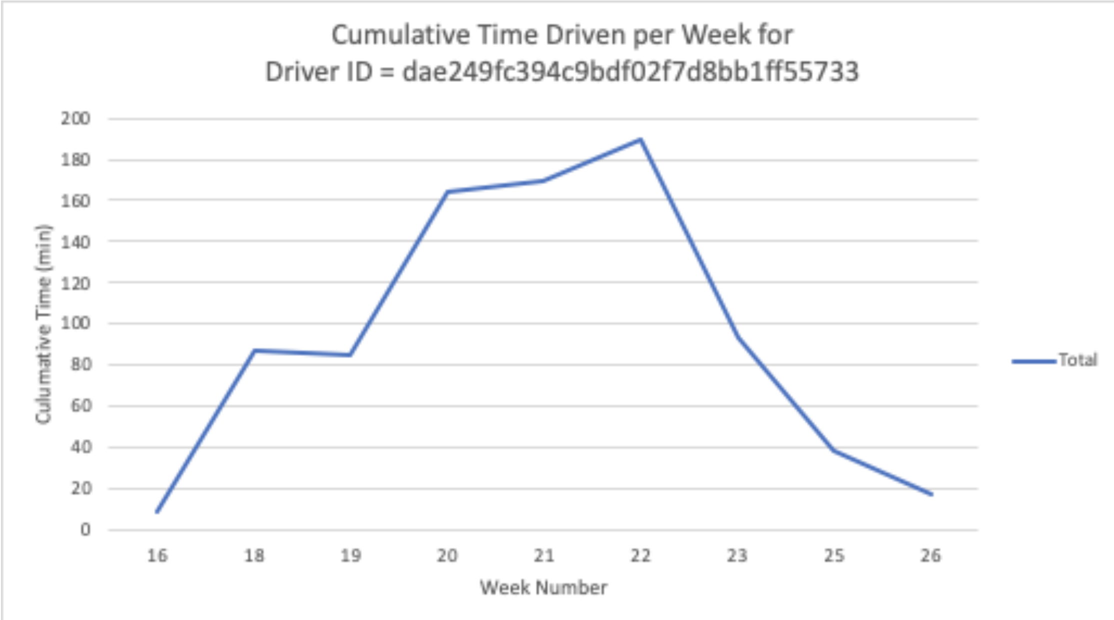

Figure 2: Driver that is showing an approximate normal distribution

With drivers that drove with a uniform distribution, we assumed that they would consistently drive at those rates for two years. We assumed a maximum lifetime of two years, or 104 weeks, for a driver based off of external factors such as the time frame that a car would need to be replaced due to constant increase in mileage over that period of time. With drivers that drove with a normal distribution, we assumed that their lifetime was the number of weeks that they took to get to their maximum cumulative time per week, multiplied by two. Given that the data set that was provided was only for thirteen weeks, we recognize that the maximum cumulative time found could be a local maximum and that would need more data to ensure that the global maximum was captured.

Using this method, we had a final table with the unique Driver IDs and their projected lifetime of either 104 weeks for uniformly distributed drivers, or two times the time that it took for drivers to reach their local maximum for normally distributed drivers. We found that the average lifetime of a driver based off of the given data set is approximately fifteen weeks.

Driver Segments

From our analysis of the projected lifetime of a driver, two distinguished sets of drivers can be specified: those that drive consistently on a per week basis and those that peak and then decline in the weekly distance they drive. The average revenue for each segment over the timeframe of the dataset was calculated, the average for consistent drivers was $1436.17, whereas the average for drivers that followed a normal distribution was $2738.76. It is apparent that normally distributed drivers, ones that drive with Lyft in the short term, are of more value.

Recommendations

Based off of our analysis, we recommend to focus on quantity incentives for consistent drivers and time incentives for drivers with a normal distribution. For example, consistent drivers can be incentivised by bonuses to drive incrementally more hours and normal distribution drivers can be incentivised by meeting a minimum number of hours driven per week X weeks in a row. This would aim to create consistency while maximizing revenue.