

Null_Data_Challenge

Roger Yuan

9/3/2019

```
## [1] 13.09894
## Time difference of 55.75609 days
## $stats
## [1] 1.74000 42.84226 57.96610 73.11388 90.78852
##
## $n
## [1] 836
##
## $conf
## [1] 56.31189 59.62031
##
## $out
## numeric(0)
```

A driver's lifetime value would be calculated using a linear combination of projected lifetime revenue, average prime time bonus per ride, and proportion of rides driven during rush hour or late nights. After normalizing each component, we multiply them together to get a lifetime value.

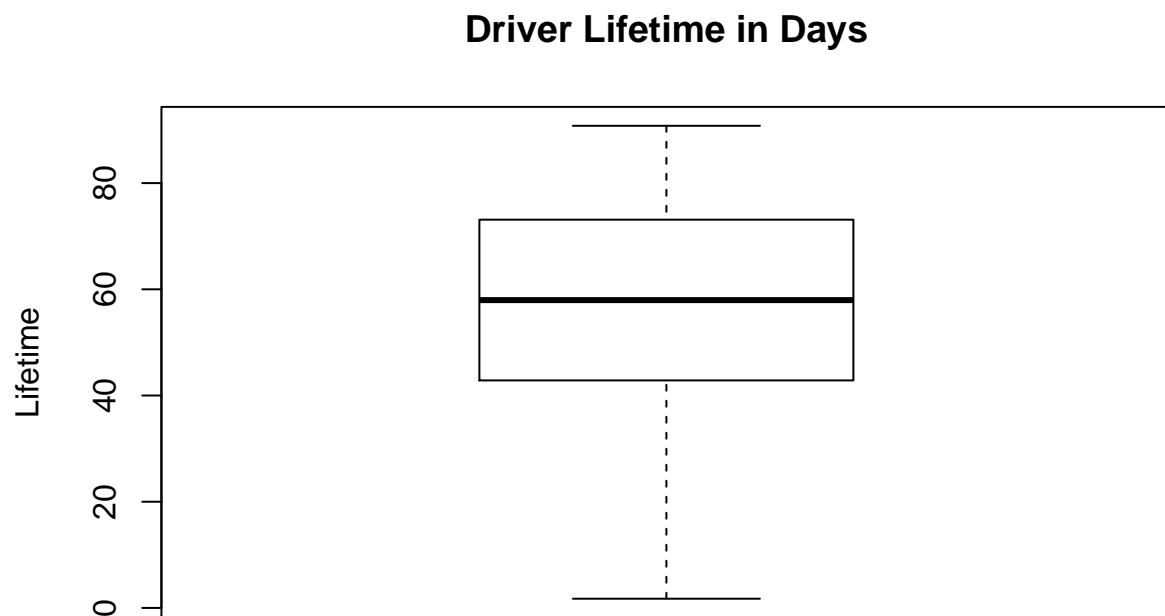
Revenue is often an important indicator of an employee's worth to a company. At the end of the day, an employee's contribution to their company's bottom line determines how important they are to the company, at least financially. To determine a driver's financial value, we found their average daily revenue by totaling the revenue of their rides during their lifetime and dividing by their lifetime before multiplying by the average driver lifetime. This revenue calculation accounts for prime time bonuses as well.

However, while revenue is important, it is not all a driver has to offer to Lyft. We believe that a driver who works for Lyft during times of high demand, such as holidays or rush hours, would be of much value to the company. Not only would they bring in more revenue from the prime time bonuses that have taken effect, but they also bolster Lyft's consistency and dependency during times when people rely on them the most. Moreover, they increase Lyft's ability to contend against competitors when there are more potential riders unfamiliar with the service or are not loyal to a specific one. By allowing Lyft to provide its services to more riders during especially busy, critical times, these drivers have increased worth towards Lyft that cannot be described by revenue alone.

In order to represent this, we summed up all the prime time bonuses for each driver as an approximate measure of how often they drive during high demands. Larger total prime time bonuses indicate a tendency to work during busy times and lower totals indicate otherwise. We once again normalize this factor.

Another factor we wish to consider is how often a driver works during rush hours. This is somewhat similar to the previously mentioned total prime time, but the number of rides during rush hour or late nights represents a more consistent work schedule. Because riders who hail rides during these times often need to, often due to work, we believe Lyft drivers who work within these hours are not only fulfilling demand but also building customer loyalty. Because many regularly require rides during these times,

```
boxplot(as.numeric(driver_lifetime$life_time), main="Driver Lifetime in Days", ylab="Lifetime")
```



As expected, the lifetime of a driver has no relation to the total revenue of a driver. The correlation between the two is 0.52, which can be seen by plotting the lifetime vs the total revenue. This makes sense, as a part-time driver working the same amount of time as a full-time driver will make much less. Thus, we can rule out lifetime as a factor in a driver's lifetime value.

Outlier Drivers

There are 8 drivers whose total revenue is about three to four times greater than the median total revenue, making \$10,000-\$12,000 while the median is about \$3,000. These drivers are considered outliers, and we will be analyzing them alongside median and lower end drivers to observe how they set themselves so far apart from others.

We were first interested in the percent of rides that did not earn revenue for each group of drivers, namely the time between the request and passenger pick up. We first calculated for each ride the total ride time between request and drop off and subtracted the ride duration. We divided this by the total ride time to get the proportion of the ride that did not earn revenue. Surprisingly, the distribution of percentages of times with no revenue were not significantly different between the three groups, which is evident from the p-values of

