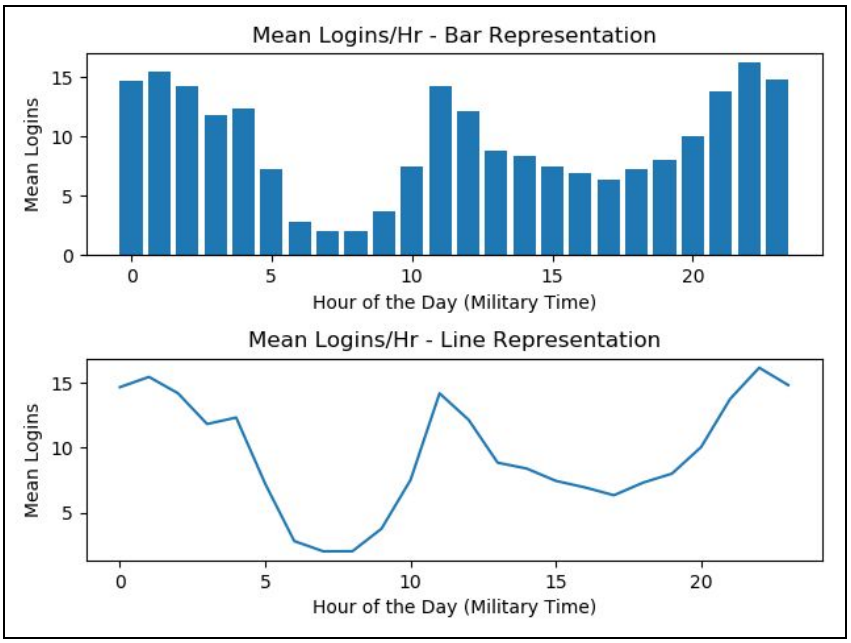
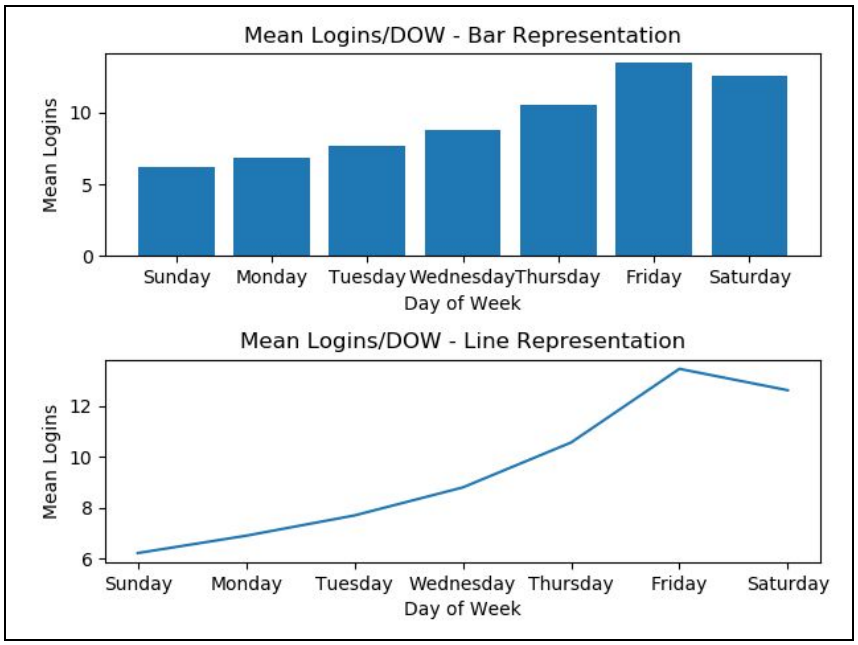
Part 1: Login Time Stamps

Login data was loaded into a pandas dataframe and resampled into 15 minute intervals with the resample method and count function. Month, day of month, day of week, hour of the day, and hour intervals were pulled out of the datetime index and used with groupby to assess trends. The two major trends in login data are logins by day of the week, and logins by hour of the day. The two seperate trends are illustrated in the 4 plots below. The first two plots illustrate that demand is cyclical day to day, with highest mean logins between the hours of 8pm and 4 am with another spike around noon. The third and fourth plots illustrate that mean logins rise throughout the week, peaking on Friday and Saturday.





Part 2: Experiment and Metrics Design

The simplest and most naive measure for evaluating the success of this experiment is a straightforward proportion of drivers who cross the toll bridge now vs. when the reimbursement goes into effect. Finding the population proportion of drivers who currently cross the toll bridge would not be difficult as presumably Ultimate maintains records of a driver's location and pick-up history. However, this is a naive measure because it assumes that the only goal of the city managers is to encourage drivers to cross the bridge, and not to maximize revenue. Given that Ultimate is a company looking to maximize revenue or utilization of the market, we'll need a more detailed methodology to gauge the success of this experiment.

Given two cities that are completely separate with no shared drivers, we would expect *two seperate supply and demand interactions*. Two cities that are adjacent with no barrier to driver movement between the *two will experience fluctuations in supply and demand that should come to equilibrium*. For example, during the day Metropolis has higher demand than Gotham. For an equilibrium in supply/demand to be reached between the two cities, drivers from Gotham will flow to Metropolis during the day. This equilibrium can be assessed by a similar price per unit distance (PPUD) for customers in each city. The current scenario lies somewhere in the middle. Gotham and Metropolis should experience an equilibrium in supply and demand given their complimentary schedules. However, due to the toll, drivers may not be inclined to cross to the other city. Therefore, demand (and therefore price) is higher in Metropolis during the day with the inverse being true of Gotham. Ultimate would like to eliminate this fluctuation in price caused by low supply with a toll reimbursement.

With the above taken into account, we can refine our metric for evaluating the success of this proposed experiment. Price per unit distance (PPUD) paid by customers is the important measure here. For maximum market utilization, Ultimate would like this value to be relatively the same in both cities at all times. This will prevent single city competitors from capitalizing on the lack of supply (and higher prices) of Ultimate. Therefore the best metric to use for evaluating the success of this experiment is the difference in PPUD paid by customers in each city. This could perhaps be sampled as a mean of all rides taken in 15 minute intervals (by city) to capture micro-swings in demand. This data already exists for rides where the toll is not reimbursed. Splitting these 15 minute interval means into two groups,night or day, allows us to see if the difference in PPUD paid by customers is the same day or night.

The experiment itself would be relatively straightforward. Given that we already have a measure for the difference in PPUD paid by customers in both cities when the toll is NOT reimbursed, we can simply inform drivers of a limited time toll reimbursement program and observe the results. For the purposes of statistical rigor this test program should last around 2 weeks. A statistical test to compare the overall means of the 15 minute PPUD intervals between pre and post toll reimbursement will let us know whether this program decreases the difference in PPUD paid by customers. As mentioned above, the two groups (day or night) will let us ascertain whether the

toll reimbursement program varies in effectiveness day or night. A two sample t-test for difference in population means is the appropriate test here.

The experiment, after a two week trial, will yield two p-values for the difference in mean difference in PPUD paid by customers in both cities. Obviously a statistically significant ($p < 0.05$) decrease in the mean difference in PPUD would indicate the program as being successful in convincing drivers to cross the toll bridge and drive in the other city. However, the effect on revenue must also be analyzed before the toll reimbursement can be officially sanctioned. To judge this, city managers will have to look at total cost of tolls vs. the increase in total revenue observed over a similar 14 day period from pre-toll reimbursement. If the city managers find that the increase in total revenue generated from rides outweighs the total cost incurred from the tolls this program will be worth keeping around.