Part 2 - Experiment and metrics

- Measuring how many drivers cross the bridge (one direction to another counts as one type, another direction to another counts as another type)
    - Hourly average of how many drivers cross bridge before implementation and after-
        - Test to use
            - Chi-squared test of differences
            - We might want to subset each side of the bridge, and maybe even day and night to eliminate some of the noise, and weekends
- Pick-up location
    - Difference in the number of rides given on either side of the bridge per hour where pickup location is marked as being on one side of the bridge vs. another (demand may be high in one city at night, but not enough drivers, so people use another service that is cheaper, public transportation, etc. With new implementation, rides go up. Or, more simply, the drivers are just there to pick people up, so there are more rides)
        - Test to use
            - Chi-squared test of difference
            - Again, subset each side of bridge, maybe even day and night

1. The key measure of success in this experiment should be the average number of times drivers cross the toll bridge from one side to the other each hour of the day 0-24, weekday/weekend, over a span of three months after implementing the intervention. This is an ideal metric because it accounts for natural hourly and weekend vs. weekday differences in user traffic, while providing easily reportable and viewable results. This metric is viable because it will tell us if more driving partners are crossing the toll bridge after the intervention.. As stated, these would be broken up by weekend/weekday and a running average for each hour of the day.
2. A practical experiment will attempt to observe statistical significance between observations of drivers crossing the bridge before and after the experiment. First, a specific time frame should be chosen. For example, if it is known that the number of user traffic has a distinctly different pattern between the hours of 00 and 3:00 when compared with the hours of 3:00 and 6:00, the experiment should choose average observation counts from only one of these periods, preferably a time that is more busy. If resources and time allow, observations from each distinct time period can be incorporated into the experiment and statistical analysis. Also important will be the distinction between weekend days and weekdays. For example, if there is a higher volume of user traffic from say Friday afternoon until early Saturday morning, observations gathered during this period should be analyzed separate from those on other days. Given the information that is known about activity in the two cities, it will likely be better to make observations on a weekday, as user activity levels are inversed across day and night hours. This could make differences between before and after implementation time periods more

observable, as a reasonable person would assume that a driver will cross the bridge from one city to another if the opportunity for work in the other city is greater, remembering that Gotham is busy during the day, and Ultimate is busy during the night. Given this knowledge, the experiment should count observations of driving partners crossing to Gotham during the busiest time period at night, and observations of driving partners crossing to Ultimate during the busiest time period during the day.

3.

a) The experiment will be implemented by first establishing the busiest hours for user demand in the respective cities. To give an example, let's say that Gotham's busiest hours are 11pm to 3am, and Ultimate's busiest hours are 4pm to 8pm. Data will be gathered on the daily total (only Monday-Thursday) of bridge crossings during these periods for the respective cities.

- If the company already has GPS pinging locations within a reasonable amount of time - say a few minutes - that would indicate when a driver has crossed the bridge. This data will be queried for the time period specified, for each city.
- Data can be created by automatically tracking when a driver has passed into the bridge zone, using GPS and a defined set of coordinates. In this case, data collection would begin a few months before the intervention, so as to have the two time frames to compare

b) Either a t or z difference of means test (depending on the distributions) will be used to compare the number of bridge crossings to Gotham in the 11pm-3am time period before and after the intervention, and to Ultimate in the 4pm-8pm period to measure efficacy of the toll waiver.

c) The results can be interpreted such that a difference in the mean daily number of observations of bridge crossings before and after the intervention in the respective time window, for each city, at the .05 level of significance, will indicate that the intervention has encouraged drivers to be more active in both cities. Interpreting the results might also involve calculating the return or profit increase from money spent reimbursing drivers. This might involve calculating the amount of tolls paid during the time window, vs. the difference in average revenue for the company by either the Gotham or Ultimate location during the time window. A rise in profits during this time window, however, might be harder to link back to the money spent for drivers to cross the bridge. Another caveat is that Ultimate may want to see results quicker than the 4-6 month period laid out in the experimental design. If this is the case, they could shorten the data collection periods. An assumption of the experiment that must also be considered is if it is viable given the distribution of bridge crossing observations to compare them for statistical significance using a t or z difference of means test.