

Part 2: Experiment and Metrics Design

1. What would you choose as the key metric of success of this experiment in encouraging driver partners to serve both cities, and why would you choose this metric?

The metric of success of this experiment would really depend on the information that you were able to collect from the drivers and the rides. If you were able to collect the GPS locations of each ride or at least whether each ride was in Gotham or Metropolis, then a good metric of success would be looking at the number of rides that each driver gave in each city over a given period of time. The more evenly distributed the rides, or at least whether some percentage of rides were in the other city, the better. It also might be good to look specifically at weekends since there is reasonable traffic in both cities on the weekends; since the cities have complimentary circadian rhythms during the week, that might also make it more likely that each driver would stay within one city, because a driver might work only during the day (and therefore work in Ultimate Metropolis) or only during the night (and therefore work in Ultimate Gotham).

Overall, then, I think a good metric of success would be the fraction of weekend drivers who gave at least 30% of their rides in the other city. The 30% cutoff is arbitrary and could be adjusted based on how much you actually want the rides to be distributed.

2. Describe a practical experiment you would design to compare the effectiveness of the proposed change in relation to the key measure of success.

a. How you will implement the experiment

- For this experiment I would look only at weekends because there is traffic in both cities on the weekends, so we want to particularly incentivize drivers to serve both cities on the weekends.
- I would divide the drivers into two random groups for this experiment: one group would get the toll reimbursement, and one group would not.
- I would then track the number of rides that each weekend driver gave in each city for both the test group (group where you gave toll reimbursement) and the control group (group where you did not give toll reimbursement). I would then count the number of drivers in each group that had a 30-70 distribution or better of rides over a certain period of time, perhaps several weekends. If the fraction of drivers with these good distributions was statistically significantly higher for the test group than the control group, then I would draw the conclusion that the toll reimbursement is a good idea.

b. What statistical tests you will conduct to verify the significance of the observation

- I would use a hypothesis test for a difference in proportions, since we would be comparing the proportions for each group
- This is a two-proportion z-test
- Since we only want to know if the fraction of drivers with good distributions is higher, we would use a one-tailed test.

c. How you would interpret the results and provide recommendations to the city operations team along with any caveats

- If the fraction of drivers with the good distributions was statistically significantly higher for the test group than the control group, then I would draw the conclusion that the toll reimbursement is a good idea and recommend it to the city operations team
- However, statistical significance does not always mean practical significance. If the difference is statistically significant but small, then it might not be worth it to implement the change.