

Part 2 - Experiment and metrics design

The neighboring cities of Gotham and Metropolis have complementary circadian rhythms: on weekdays, Ultimate Gotham is most active at night, and Ultimate Metropolis is most active during the day. On weekends, there is reasonable activity in both cities.

However, a toll bridge, with a two way toll, between the two cities causes driver partners to tend to be exclusive to each city. The Ultimate managers of city operations for the two cities have proposed an experiment to encourage driver partners to be available in both cities, by reimbursing all toll costs.

1. *What would you choose as the key measure of success of this experiment in encouraging driver partners to serve both cities, and why would you choose this metric?*
2. *Describe a practical experiment you would design to compare the effectiveness of the proposed change in relation to the key measure of success. Please provide details on:*
 - a. *How you will implement the experiment*
 - b. *What statistical test(s) you will conduct to verify the significance of the observation*
 - c. *How you would interpret the results and provide recommendations to the city operations team along with any caveats.*

1. This experiment's key measure of success would be the number of drivers that serve both cities per week. I would choose this as it directly measures success for our goal of making more drivers non-exclusive to each city. We can also see how trends move from when tolls were on the drivers to when tolls were reimbursed to see how much of an impact this has.
2. Experiment Design
 - a. We would design an A/B test to determine the likelihood that an increase in drivers serving both cities is correlated with our proposed change.
 - i. Group A: drivers before reimbursement for tolls
 - ii. Group B: drivers after reimbursement
 - iii. Duration: 3 months before and after reimbursement was implemented
 - iv. Key data: Driver, origin & destination of the trip, time/date of trip, number of trips, toll crossing, trip across cities (boolean)
 - v. Key Metric: Drivers serving both cities on weekly basis (boolean 0 or 1)
 - b. To verify the significance of the observation, we would conduct a Chi-Square Test of Independence.
 - i. Null Hypothesis: There is no significant difference in the proportion of drivers serving both cities before and after reimbursement was implemented.
 - ii. Alternative Hypothesis: There is a significant difference in the proportion of drivers serving both cities before and after reimbursement was implemented.
 - iii. Significance Level: Set a significance level at 0.05.

- c. We would interpret the results based on the p-value. If the p-value is <0.05 , we would reject the null hypothesis and determine that the reimbursement initiative had a positive impact. If it is > 0.05 , we would fail to reject the null hypothesis and determine that any increase in drivers serving both cities is random chance.
 - i. Recommendations: If we reject the null hypothesis, we could expand this initiative to other areas/cities and ensure that we keep this initiative as long as it remains cost effective. In the other scenario, we could examine this over a longer time period or try other tests. Other tests could include incentives for cross city drivers or customers, or having customers take on the toll cost.
 - ii. Caveats: Excitement/curiosity about the initiative at first may increase drivers serving both cities but might not be sustainable long-term. Other factors could affect cross city driving but may not be sustainable. Road closures/construction, major events, seasonality. As mentioned in the recommendations, we should also pay attention to the cost/benefit of this. If this doesn't lead to profitability, it won't be sustainable long term.