STorrico

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?

   My first assumption is the reason behind for this experiment is to support Ultimate's business goal of retaining its existing customers. The second assumption is that business/operations team already did a preliminary analysis and found out that there is an unmet demand at both cities and the resulting returns from servicing it will offset the cost of the suggested "we pay your toll" promotion for its drivers.

   By removing the burden of the toll cost from drivers decision-making, Ultimate also minimizes the next big consideration that drivers take into account: distance to the passenger's pickup destination and to drop-off destination. Given that drivers will most likely accept the ride, Ultimate does not need to worry about the acceptance rate.

   With that mind, I will select the success metric to be cancellation rate for this experiment. Why? Let's consider for example that after the driver arrives to the pickup location and finds out the final destination is a in neighborhood that is unlikely to get another pickup so the drivers cancels or forces a cancellation. As a result, the rider may decide to go with the competition.

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

To implement the experiment, I will capture the following historical data:

- The number of drivers that already crossed the bridge and paid the tolls for each city (going back at least 2 years)
- Average Cancellation rates for both Ultimate Metropoli drivers and Ultimate Gotham drivers
- Time periods during the day that these drivers are crossing the bridge (both ways)

STorrico

After obtaining the data, I will go ahead and plot the number of drivers vs time period to validate if these drivers are making the trips to opposite cities when their demand is high. Then, will check the distribution of the average cancelation rates.

B. what statistical test(s) you will conduct to verify the significance of the Observation

I will use Hypothesis testing of the difference between Two Population Means. The hypothesis to prove is that there will be a difference between the cancellation rates means between Ultimate Gotham drivers and Ultimate Metropolis drivers. The null hypothesis is that cancellation rates means are equal.

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

If the results from the hypothesis testing show that indeed there is a difference between the cancellation rates means, then rolling out this "We pay your toll" promotion is feasible.

My recommendations to the operations team would be that before making this promotion available to all its drivers, there should be a Soft launch available to an equal group size of drivers that were randomly selected from each city.

The duration of this launch should be 3 months, during which: monitoring of the cancellation rates takes place to check if they are decreasing as expected plus monitoring if the returns, obtained from meeting this unmet demand, continue to offset the toll cost. As the end of Soft launch approaches, a final analysis should take place to compare the Soft Launch cancellation metrics against the agreed baselines.