Springboard Chapter 26.2 - Take Home Part 2

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.

To address each of these data questions, let's break them down individually:

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?

There are three different metrics that I would consider to measure the success of the experiment:

1. **Percentage of drivers with pickups or drop-offs in both cities in the same day** (may also consider week) - this would give an understanding of how many of these drivers worked in both cities during the same time period. Considering they worked in both cities in the same day may be closer to what the managers of both cities want: where drivers work in one city during the day and the other one at night to help assist in handling the increased volume. This seemed like a good metric.

2. **Percentage of work that is handled by drivers that drive in both cities in the same day** - if drivers that work both cities are picking up a different volume of passengers, then the first statistic would be skewed. It's possible that drivers that take the extra time to go between cities would lose time and have less work in the end; it's also possible those drivers are working longer and harder hours, so they would account for a larger portion of the work.

3. **Number of drivers that work in both cities on the same day** - I prefer the first 2 statistics for how much those contribute to the overall driver pool; however, those will be smaller percentages and it would be more difficult to show that they are statistically significant. Instead the number of drivers that work in both cities would be the most valuable in testing whether a new policy worked.

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.

(a) I would refund the tolls for drivers on every Tuesday and Thursday for 2 months. For those 10 days over the 2-month period, I would compare the number of drivers that worked in both cities on those days versus the long-term average for Tuesday and Thursday. I selected these 2 midweek days, as those midweek days are when it would most benefit the transportation company and the cities if drivers worked in both cities; these days should also see less impact from weekend events. Wednesday is a control day to allow us to see any trends that develop over that time (maybe drivers were already starting to work in both cities at an increasing rate - another "control" midweek day to capture this trend would be a good idea).

(b) I would compare the mean number of drivers that work in both cities on the same day for your average Tuesday and Thursday vs. the mean number of drivers that work in both cities when the toll refund was implemented on Tuesday and Thursday. We could use a one-sided t-test on both days to make sure an improvement in that number was not just because of random chance, but actually the tolls.

(c) There are 3 possible outcomes from the test:
1. There is no discernible difference between drivers who work in both cities with or without the tolls. In this case, I would recommend considering some other policy to encourage drivers to work in both cities or perhaps in combination with this policy. It could be that drivers don't want to spend the time traveling back and forth between the cities or something else that inhibits them from working in both cities.
2. There is a slight difference, but it is not statistically significant. In this case, we can run the test over a longer period of time or including more days so that we can clarify the extent of the difference.
3. The difference between the days is statistically significant. Refunding tolls is effective at encouraging drivers to work in both cities. The governments of those cities have to decide if the revenue that they forfeit from providing these refunds is worth the predicted increased participation in work between the two cities.