We will reiterate the problem, the goal and the aim of the experiment to solve the problem for Ultimate Inc, Gotham, and Metropolis.

**Problem Statement**: Drivers partners of Ultimate Inc. tend to serve the same city that they belong to thus limiting the capacity of available rides in both cities.

**Goal**: The goal of the cities and Ultimate Inc. is to encourage drivers from a city that is active, to serve a city that is relatively less active or underserved. And thus have an even serving capacity between cities and across the time of the day.

**Solution**: With an assumption that the toll is the sole reason drivers don’t tend to go into another city. Conduct an experiment to see the effect of the absence of toll, on drivers being available for other city’s riders.

**Aim**: The aim of the experiment is to validate, if not having a toll really helps in increasing the capacity of Ultimate Inc. driver during underserved times or in underserved areas.

**Assumption**: For our experiment to perform as expected we assumed certain things.

* Drivers are willing to go into the other city given the opportunity and incentives.
* A sufficient number of drivers are available as excess capacity in a city to serve the other city.

**Success Metrics**: Since our concern here is the availability of the number of drivers and keeping this number even between cities at different times. We will use the following metrics to measure the capacity and its usage and needs.

1. Metric 1: The number of drivers accepting rides to and from or within the city they don’t belong to. This will give a direct impact of the toll on the capacity.
2. Metric 2: Number rides completed by these drivers either originating or ending at the other city. This tells us how much of it is used.
3. Metric 3: The ratio of the number of available drivers in a city to the number of unique customer logins. This will give a sense of the need for the supply or the capacity.

All of these metrics can be compared between active time and less active time of the day for a city.

**Implementation**: As a general rule of thumb for any experiment, it needs to be random enough to capture most of the variation in the data. Also, it needs to be large enough so as to nullify the effects of any random events on the results. Considering this we can largely have 3 phases of the experiments where the tolls are refunded for a period of time to see before and after effect. And the other two phases refunds tolls only for days on a schedule or on an random day. This way we ensure that if the change in driver capacity is seen and it follows closely to the pattern of toll refund. This will also ensure that any random influencer will have equal or less effect on the results. Going into the details of the experiments;

* The toll is reimbursement period is kept long and continuous enough to see an significant and constant impact on the behavior of the drivers from both cities.
* In this phase, a fixed schedule will help the driver plan their trips between cities well ahead of time. This schedule needs to be repeated for a sufficient number of times to be statistically significant. The most common schedule can be thought of repeating weekly, with special adjustment for weekends.
* A randomly selected day/s in a week would ensure random events in cities have fewer effects on the trails. This phase can be shorter compared to the other two but a necessity.

The details of these phases like how long each phase needs to be and how long it can actually be is up for debate. Even the configuration of these phases, either opening the gates suddenly and gradually reducing the frequency or slowly introducing an idea and then ramping up for a snowball effect. All these can be agreed up by city officials, statisticians, and traffic experts.

**Validating numbers**: Even after all the care is taken to make experiments as random as possible and have true results of toll reimbursements, the numbers still need to be validated. A hypothesis test that validates the statistical significance of the metric numbers is one way of doing it. Here we assume no change by the toll reimbursement and calculate the chances of getting these numbers.

**Results**: An increase in Metric 1 during the toll reimbursement days vs other days would mean that generally, drivers are moving between cities to serve rides. This can further be broken down into active time periods of each city to see if it is happening during desired times.

Metric 2 along with location data, will tell us if the capacity is used and serves the areas of under service. Here an increase during lesser active times of a city validates the need for the extra capacity.

For metric 3 an increase in the ratio is a good indicator of rising capacity but needs to stay in a certain range to be healthy. Here experience with the domain will give a good idea of a healthy range for this ratio. Thus, an increasing ratio towards this range means that there is no shortage of capacity and drivers stay busy.