Ultimate Inc. Data Challenge – Experiments and Metric Design

1. Metrics for success of experiment

There are two metrics that can be used for the success of experiment:

1. Increase in no. of intercity trips
2. Increase in revenue and profit due to increase intercity trips

The reason for choosing the above metrics can be understood if we describe the experiment and the way to test the significance of the experiment.

2. Experiment Design

We will first divide the total no. of driver partners in two groups:

Group-1: Driver Partners to whom we will reimburse the toll costs

Group-2: Driver Partners to whom we will not reimburse the toll costs

Then the experiment will be conducted for a specified duration and daily two things are recorded.

1. No of intercity trips of both the groups for each driver partner

2. Total revenue and profit generated per day for both the groups for each driver partner

Now we have the complete record of the specified duration of experiment.

Now let's do the hypothesis testing by formulating the null and alternative hypothesis.

H_o: There is no difference in the intercity trips for Group-1 and Group-2 i.e. $\mu_{group1} = \mu_{group2}$

H_a: There is significant difference in the intercity trips for Group-1 and Group-2 i.e. $\mu_{group1} \neq \mu_{group2}$

The significance of the above experiment can be found out based on two sample t-test:

$$t = \frac{\mu_{group1} - \mu_{group2}}{\sqrt{\frac{S_p^2}{n_1} + \frac{S_p^2}{n_2}}}$$

S_p = Pooled Sample Variance

The significance of experiment can be determined by calculating t value at p=0.05 i.e. at 95% confidence interval.

If $t < t_{0.95}$ H_o cannot be rejected otherwise reject the null hypothesis and accept the alternative hypothesis.

Similar process can be followed for revenue and profit.

H_o: There is no difference in the revenue and profit for Group-1 and Group-2

H_a: There is significant difference in the revenue and profit for Group-1 and Group-2

Now comes the question why those two metrics:

The first metrics of intercity trips will confirm our basic assumption that driver partners are confined to intracity trips because of toll costs and not of any other reason. If we find that there is no significant increase in intercity trips then there is some other reason for not taking the trips from Gotham to Metropolis.

The second metrics will indicate that even after reimbursing the toll cost the increase in trips should be sufficient to increase the revenue to such an extent that final profit should increase for the Ultimate Inc. It may be possible that the increase in trips are not sufficient and because of reimbursement the total profit actually decreases.

Caveat – The duration of experiment should not be too long so that Group-2 driver partners got to know about the benefits of Group-1 driver partners and deregister themselves from Ultimate. The experiment must be designed in such a manner that there should not be any biasness between the two groups. The two groups must be identical depending upon the driver rating, no. of trips, years of association with Ultimate etc.