**Section - Strategic Thinking:**

**Goal – Optimize speed of delivery and cost**

1. **Why/how/when/where are they both important? How shall we go about balancing those two objectives? How would you find the “sweet spot”?**

* Cost is important as it determines the revenue gained by the company. It is important to minimize the cost incurred and maximize the profit in order to run the business. Similarly, delivery time is also a useful and important metric since it helps to avoid customer churn.
* To deliver faster, we have to allocate riders specific to a region and assign these riders only the orders to restaurants in the region they operate. In this way, the cost incurred due by customer compensation due to delayed delivery can be avoided.

1. **What could we do to be more cost effective? What could we do to deliver faster?**

Answers:

|  |
| --- |
| **Cost Effective ideas** |
| **Idea 1:** Cost and Delivery time can be optimized by introducing a new feature called “Urgency Level”, with categories – “Urgent”, ”Can be delayed by 30 mins”, ”Specify Time”. Based on this priority can be given the urgent ones and the restaurants can be notified of future orders and in this way the delivery time metric can be met without incurring customer compensation. This would reduce the customer compensation and also the wait time for the rider to pick the order can be reduced thus delivering faster and the rider time is also saved which in turn is effective for business. |
| **Idea 2:** Find the locations/areas where the number of orders is always high and consistent throughout the month. Split riders into two groups – high frequency order areas and low frequency order areas. Riders assigned to those high frequency order areas can be paid a fixed monthly salary instead of being paid hourly/based on number of orders. Riders assigned to low frequency order areas can be paid hourly or based on number of orders since it would not be cost effective to pay them monthly fixed rate for lesser number of deliveries. |

1. **Is there any other metric you would like to have in the mix apart from speed and costs?**

Answer: The restaurant rating is also important since the quality of the food determines if the customer would place an order again or not.

**Section - Data Analysis:**

**Effects of Stacking:**

1. **Problems with chosen approach:**

From the given data, the average revenue per order when the orders are stacked is only 0.98 GBP more than the average revenue from the unstacked orders. Customer Compensations is higher in case of stacked orders and rider cost is higher is case of unstacked orders. Not clear if revenue includes rider cost and customer compensation cost from the data provided.

In case of stacked orders, the customer places the next order at an average of 3.8 days whereas in case of unstacked orders, the customer places the next order only at an average of 4.43 days.

There is no evidence that revenue is drastically increasing when the orders are stacked since the reduction in riders cost is compensated by the increase in customer compensation.

Finally, effect of stacking cannot be properly seen unless and until we have equal set of observations for both stacked and unstacked orders. In the dataset provided, there were unequal number of observations for stacked and unstacked orders. To overcome bias, analysis was done by taking mean and still there were no drastic difference in revenue, rider cost and customer compensations between the two types.

1. **What additional data would you require to answer further questions:**

|  |  |  |
| --- | --- | --- |
| **S.No** | **Description** | **Data Provided?** |
| 1 | Order ID | Yes |
| 2 | Customer ID | Yes |
| 3 | Time Customer places the order | No |
| 4 | Time the order is accepted by restaurant | Yes |
| 5 | Time the order was assigned to Rider | No |
| 6 | Distance between the restaurant and the rider when the order was assigned to him/her | No |
| 7 | Time the rider arrived at restaurant | No |
| 8 | Wait time to pick the delivery | No |
| 9 | Time the order was picked by the rider | No |
| 10 | Distance between the restaurant and the address of delivery | No |
| 11 | Estimated Time of arrival at delivery address | No |
| 12 | Actual Time of arrival at delivery address | No |
| 13 | How big the parcel is to determine if it can be stacked or not | No |
| 14 | Is the order stacked or not | Yes |
| 15 | How many stacks? | No |
| 16 | If stacked, is the order from same restaurant or from different restaurant? | No |
| 17 | Distance between the restaurants from stacked orders | No |
| 18 | Distance between the stacked order address of delivery | No |
| 19 | Which order to be delivered first in case of stacked order? | No |
| 20 | Next Customer order date | Yes |
| 21 | Customer Compensation Rate | Yes |
| 22 | Rider Cost | Yes |
| 23 | Revenue | Yes |
| 24 | Is the order placed at Peak hour? | No |
| 25 | Restaurant Rating | No |
| 26 | Rider Rating | No |
| 27 | Cuisine Category | No |
| 28 | Regular Customer or Occasional Customer | No |
| 29 | Cost of order | No |

**Section – Experiment Design:**

**Hypothesis:**

Assumption of Null Hypothesis- Stacking has no effect. It means whether orders are stacked or not there is not much change in revenue.

**Metrics:**

Following are the metrics considered to conduct power analysis:

1. Significance Level
2. Effect Size
3. Power Level
4. Sample Size

**Power Analysis:**

Power Analysis is conducted by analyzing the effect of a metric based on the changes in other metrics to find the optimum parameters.

Initially power analysis is done to find the initial sample size. The metrics to find sample size are as follows:

1. Significance Level (alpha) - p-value – 0.05
2. Effect Size – 0.8
3. Power Level – 80%

This is done by using “TTestIndPower()” function. For the above metrics, the sample size is 25.

**Observation vs Rest of metrics:**

Keeping Effect size and significance level constant and varying the power level shows that increase in power level also increases the sample size.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| S.No | Effect Size | Significance Level | Power Level | Sample Size |
| 1 | 0.8 | 0.05 | 0.5 | 13 |
| 2 | 0.8 | 0.05 | 0.7 | 20 |
| 3 | 0.8 | 0.05 | 0.8 | 25 |
| 4 | 0.8 | 0.05 | 0.9 | 34 |

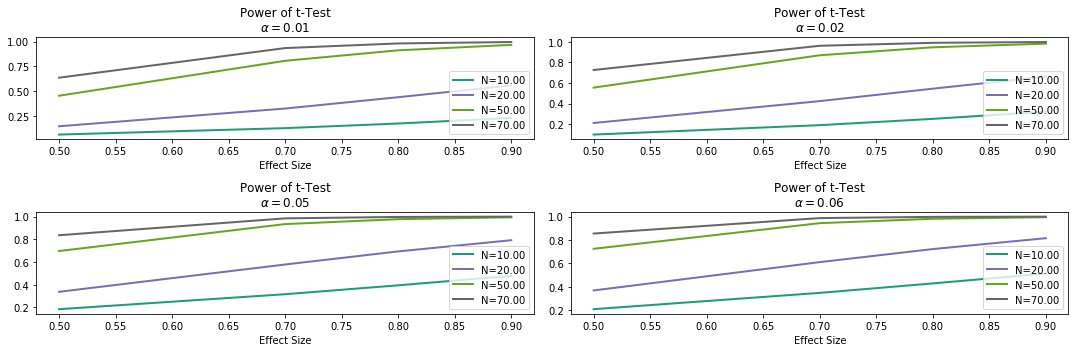
Keeping Power Level and significance level constant and varying the Effect Size shows that increase in Effect size decreases the sample size.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| S.No | Effect Size | Significance Level | Power Level | Sample Size |
| 1 | 0.5 | 0.05 | 0.8 | 63 |
| 2 | 0.7 | 0.05 | 0.8 | 33 |
| 3 | 0.8 | 0.05 | 0.8 | 25 |
| 4 | 0.9 | 0.05 | 0.8 | 20 |

Keeping Effect size and Power level constant and varying the significance level shows that increase in significance level decreases the sample size.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| S.No | Effect Size | Significance Level | Power Level | Sample Size |
| 1 | 0.8 | 0.01 | 0.8 | 38 |
| 2 | 0.8 | 0.02 | 0.8 | 32 |
| 3 | 0.8 | 0.05 | 0.8 | 25 |
| 4 | 0.8 | 0.07 | 0.8 | 22 |

**Power vs Rest of Metrics:**

****

We can see for varying significance values, the power level increases as number of observations decreases.

**Exploratory Analysis:**

1. Look for missing data, outliers
2. Plot the numerical variables individually if not many in number
3. Plot Revenue vs Rider Cost, Revenue vs Customer Compensation
4. Split the data into stacked and unstack orders and analyze them individually
5. Find the peak hours and analyze revenue, rider cost, compensation from stacked and unstacked orders
6. Plot the data for customers who placed next order and see how frequent they placed the order

Please refer to the python notebook for the detailed level of explanatory analysis.