**GENERAL QUESTIONS**

1. **Mention the palette of tools which you would recommend to be used for analysis of our data, and write for each of them how you would be using each tool**.

Power BI – I would use this tool to create interactive dashboards.

Numpy and pandas – I would use these tools to manipulate data as I am trying to make sense out of it.

Matplolib.pyplot, seaborn, plotly – I would use these tools for visualizations.

Scipy and statsmodels – I would use these tools for statistical tests

1. **What metrics would you track on a regular basis, and how do you use the information to adjust your approach?**

I would keep an eye on features that had missing values , look for a reason why there were missing values and adjust accordingly still observing how the changes affect the model.

I would also track all the features that resulted from feature engineering and observing their effects on the final model.

1. **How would you go about identifying the unique needs for new products from our customers? You have all departments at your disposal, and you need those data, telling management what our customers needs are - so that we can start delivering! What would you do?**

I would listen to the products customers like talking about, what products make them happy and also be keen on the products that receive high ratings.

Other than that, I would use internal capabilities of the departments. For example use of a call center to know whether customers are satisfied and also get more opinions of what would make them feel happier.

This can also be achieved by use of an app or a link to a website to gather more opinions about a product from the customers.

**4 .** **Analysis: The good old joke persists; “20% of all traffic accidents are caused by people who have drunk alcohol. This means that 80% of all traffic accidents are caused by drivers who are sober. Conclusion: Drive drunk!”. Now from a data analytics perspective, explain step by step with arguments presented in good order, whether this above conclusion is right or wrong, and how you do the analysis to reach to the conclusion**.

The good old joke follows a pareto principle which states that , many outcomes, roughly 80% of consequences come from 20% of causes.

There are other causes that cause traffic accidents other driving drunk only that driving drunk has a significant grouping in the causes of traffic accidents. A driver might be sober but he/she is driving carelessly, roads might be having potholes, it might be foggy yet the driver is sober.

The conclusion drive drunk is absolutely wrong since the pareto principle, it doesn’t rule out the possibility of an event occurring.