Name: Amaan Khan

Email: amaan58231@gmail.com

Country: Canada

College: York University – Schulich School of Business

Specialization: Data Analyst

Problem: Cross selling recommendation - Group Project - How to increase cross selling of Banking Product

Problem Description: The problem here is that they need to increase credit union's cross selling. The bank wants to sell their customers more than one product.

Data Understanding: The data given here contains a combination of numerical and analytical data, as it includes product details, demographic data, and account details.

Problems with data:

- There are some columns where it has fields with "NA"

Type of Data:

- This is tabular data, also often referred to as structured dataset.
- Specific types of data in the set are nominal data, ordinary data, discrete data, time data, continuous data, and binary data.

What approaches you are trying to apply on your data set to overcome problems like NA value, outlier etc and why?

- For missing values I could use imputation. This is because the other option of deletion seems risky and could mess up my data analysis. Although, I will try to keep and open mind here for more solutions that could be better.
- For handling outliers, I could use deletion if there is something that is clearly incorrect, or I could use transformation to reduce the effects the outliers may have on the case. Winsorization also seems useful here as it would limit the very extreme values of the statistical data, which would also reduce the effect of the outliers.