Group Name: Data Detective

Name: Syed Ahmad Sohail

Email: syedahmad.sohail@ucalgary.ca

Country: Canada

College/Company: Data Glacier

Specialization: Data Analyst

Week 9

Problem description:

XYZ Credit Union is a banking institution located in Latin America, and it's doing well in selling different types of banking products such as credit cards, deposit accounts, retirement accounts, and safe deposit boxes to its customers. However, despite their success in selling these products, the bank is facing a challenge of not being able to sell more than one product to their existing customers. This means that the bank is not performing well in cross-selling, which is an important strategy to increase revenue and customer loyalty. Cross-selling is the practice of encouraging customers to purchase additional products or services beyond their initial purchase. In this case, the bank is missing out on the opportunity to sell more of their products to their existing customers, which could lead to increased revenue and stronger customer relationships. As a result, the bank has approached ABC analytics to help them solve this problem. The goal is to identify ways to increase cross-selling and provide actionable insights that the bank can implement to improve its cross-selling strategy. By doing so, the bank can improve its overall performance and maintain its competitive edge in the market.

Github Repo Link:

https://github.com/syedahmadsohail96/Cross-Selling-Analysis

Data Cleansing and transformation done on data:

I'm happy to share that I have successfully performed data cleansing and transformation on a dataset using a combination of techniques and tools. With the help of a dictionary provided, I was able to convert Spanish columns to English, making the data more accessible and easier to analyze.

To ensure that the data were suitable for analysis, I used the .dtypes function to identify the type of data in each column. Next, I transformed the data in the 'Gross_Income' column to a numeric format using pd.to_numeric, with the errors parameter set to 'coerce'. This allowed me to handle any non-numeric data and convert them to NaN values, making it easier to analyze the dataset.

Finally, I used df1.fillna(df1.mean(), inplace=True) to fill any remaining NaN values in the dataset with the mean value of the column. This ensures that the dataset is complete and suitable for further analysis.

Overall, I'm proud to have successfully navigated the challenges of data cleansing and transformation, and I'm confident that these steps will help to unlock valuable insights from the dataset.

NLP:

- Text in data was converted to lowercase for consistency.
- Stop words were removed from the data frame

All the analysis was done in Cross_selling.py