**Banking Business Case**

**Problem Statement**: A major bank in Middle East came to NeoStats with help in analysing its current customer base and its marketing campaigns. It wants to understand which customers are most likely to take a term deposit (fixed deposit), and then send this list to their call centre.

The data has been provided in attached excel sheet. You as NeoStats employee are tasked with carrying out this project. Please carry out the analysis and answer the questions given below.

Make suitable assumptions, define the steps clearly and document your code when required.

**Guidelines for Submission:**

**1. Data Analysis Tool:**

* Use python for all the analysis. Share the Jupyter notebook or Google Colab link for data pre-processing, analysis, visualization, and model training.

**2. Presentation:**

* Prepare a crisp PowerPoint presentation summarizing your findings, approach, and results. Keep it within a 10-minute timeframe to maintain engagement.

**3. Question & Answer Session:**

* After your presentation, be prepared for a 5-minute Q&A round. Ensure you understand every aspect of your analysis and approach, as the panel might ask in-depth or technical questions.

**Analysis Questions:**

* **Data Analysis & Visualization:**

**1. Income Insights:**

* How many customers have no annual income? Plot and present the data distribution of these customers.

**2. Loan-less Customers Profile:**

* Filter out customers who don’t have any type of loan. Plot the distribution of their Income, balance, and profession. How do these metrics differ from those with loans?

**3. Loan and Insurance Analysis:**

* Calculate the percentage of customers with a loan who have taken out insurance. Visualize this data and discuss potential implications.

**4. Communication Strategy Insights:**

* Analyse and summarize the best Contact method (with the highest success percentage) to contact people to ascertain the status of term deposit subscriptions.

**5. Age and Home Loans:**

* Determine which age group has the highest percentage of home loans. Present this data visually and discuss possible reasons.

**6. Income and Age Relationship:**

* Investigate any relationships between annual income and age group. Use appropriate plots and statistics to present the findings.

**Modeling:**

**1. Term Deposit and Related Variables:**

* Identify variables strongly related to Term Deposit. Discuss your approach when the variable is categorical. Which tests or metrics will you employ?

**2. Predictive Model Building:**

* Train a prediction model of your choice to estimate the probability that a customer will opt for a term deposit. Adhere to an 80:20 train:test split. Report and present the model's performance metrics on both the train and test datasets.

**3. Model Improvement Strategies:**

* Discuss potential methods or approaches to enhance model performance. This could include feature engineering, different algorithms, or refining the data preprocessing steps.

**Evaluation Criteria:**

Candidates are expected to showcase their analytical, logical thinking, and deduction skills throughout this challenge. Your approach to the problem, as well as your ability to present and defend your findings, will be central to the evaluation process.

Please ensure the work is your original work. Any kind of plagiarism or taking help from others will have adverse consequences.

Best of Luck!

Team NeoStats.