Case Study: Bank Telemarketing Campaign

The bank aims to enhance its revenue by conducting a cost-efficient telemarketing campaign for term deposits among existing customers. Term deposits, fixed investments with predetermined interest rates, serve to foster long-term customer relationships. The objective is to conduct an end-to-end Exploratory Data Analysis (EDA) on the campaign dataset, identifying patterns and providing insights to improve the positive response rate. The analysis will involve examining customer demographics, temporal trends, and other factors influencing the success of the campaign, ultimately offering recommendations for targeted improvements in the bank's marketing strategy.

Software/Tools:

- Jupyter Notebook(Python)
- Python

Task to Perform:

Conduct a thorough Exploratory Data Analysis (EDA) on the dataset related to the bank's telemarketing campaign that should involves the following sequential steps:

1. Understanding the Dataset
    a. Load and scrutinize the dataset to comprehend its structure, encompassing columns and data types.
    b. Inspect for any instances of missing values, outliers, or data inconsistencies.
2. Descriptive Statistics
    a. Derive summary statistics (mean, median, standard deviation) for relevant columns.
    b. Examine the distribution of the target variable, indicating responses to the term deposit campaign.
3. Univariate Analysis
    a. Examine the distribution of individual key features, such as age, balance, and call duration.
    b. Employ visual aids like histograms, box plots, and kernel density plots to discern patterns and outliers.
4. Bivariate Analysis
    a. Evaluate the relationship between independent variables and the target variable.

b. Analyze how features like age, job type, education, marital status, etc., associate with the success of the term deposit campaign, using visualizations like bar charts, stacked bar charts, and heatmaps.

5. Categorical Variables Analysis
   a. Investigate the distribution of categorical variables such as job type, education, and marital status.
   b. Assess the impact of these categorical variables on the campaign's success through visualizations like bar charts.

6. Temporal Analysis
   a. Investigate temporal patterns in the success of the campaign over time.
   b. Analyze if specific months or days exhibit superior campaign performance.

7. Feature Engineering
   a. Introduce new features that may enhance prediction, such as creating age groups or income categories.
   b. Apply encoding techniques to transform categorical variables if necessary.

8. Correlation Analysis
   a. Examine correlations between independent variables to identify multicollinearity.
   b. Evaluate how correlated features may influence the target variable.

9. Outlier Detection and Handling
   a. Identify and rectify outliers that could impact the analysis and predictions.

Recommendations:

Offer actionable insights based on the analysis, proposing areas where the bank can enhance its positive response rate. Consider factors such as customer demographics, campaign timing, and communication strategies.

Artifacts to be generated (For Learners):

- Python File/Jupyter Notebook (.ipynb)
- Presentation (Slides) summarizing the project
  - Learner needs to present this to Faculty (Viva-Voice)
  - Presentation time: 10 Mins
- Artifacts generated needs to be submitted in vLearn on or before deadline
- Create a zip file of Jupyter Notebook and Presentation file
  - Zip file name: <LearnerName>_ABADS_<Batch>.zip
  - Eg: KartikMudaliar_ABADS_B10.zip

Grading Criteria (For Faculty & Learners):

Learners will be graded on below mentioned parameters(20 points for each):

1) Relevance and approach of Analysis
2) Quality of Output
3) Overall presentation
4) Response to Questions
5) Creativity & Innovation

Total Points: 100