Bank Marketing Analysis (EDA Report)

Project : Bank Marketing Analysis

Name : Apoorv Ajay Thite

University : The Pennsylvania State University

Email : apoorvthite21@gmail.com

Country : United States of America

Specialisation : Data Science

Internship Batch : LISUM34

Date : 7th August, 2024

**Problem Description:**

The data is related to direct marketing campaigns (phone calls) of a Portuguese banking institution. The classification goal is to predict if the client will subscribe to a term deposit (variable y).

The main objective of the project is to predict customer subscriptions to time deposits based on a direct marketing campaign conducted by a Portuguese banking institution via telephone. This goal translates into optimizing marketing resources, enhancing customer engagement strategies, and ultimately increasing the effectiveness of these campaigns. By leveraging historical data for a binary classification challenge, the project aims to effectively identify potential subscribers, improving campaign ROI and customer experience through targeted and informed promotion.

**Data Intake Report:**

The dataset consists of 45,211 entries and 17 columns, encompassing both numerical and categorical variables. Key attributes include the client's age, job type, marital status, education level, account balance, and the outcome of previous marketing campaigns. The target variable is whether the client subscribed to a term deposit. The data was thoroughly checked for completeness, with no missing values found. The categorical variables were analysed for unique values, and numerical variables were inspected for outliers. The data intake process ensures that the dataset is suitable for further exploratory data analysis and modelling.

**Exploratory Data Analysis (EDA) Summary:**

Several exploratory data analysis techniques were used to understand the dataset better. The analysis included visualizing the distribution of age, account balances, job types, marital status, education levels, and the impact of housing loans on term deposit subscriptions. A correlation heatmap was also generated to identify relationships between numerical variables.

**Key Findings:**

* Age distribution shows most clients are in their 30s to 40s.
* Account balances vary widely, with a concentration on lower values.
* Most clients are in blue-collar jobs, followed by management and technician roles.
* Marital status and education levels influence the likelihood of term deposit subscription.
* Clients with no housing loans and longer contact durations tend to subscribe more.

**Final Recommendation:**

Based on the EDA performed, it is recommended to focus on clients with longer contact durations and higher balances, as they are more likely to subscribe to a term deposit. Additionally, targeting individuals with tertiary education and those without housing loans could yield better results. These insights can help tailor marketing campaigns more effectively, improving subscription rates.