Project Proposal: Statistical Modeling and Analysis of Telemarketing Campaign for Bank Term Deposits

I. Project Overview

This project leverages a dataset from a telemarketing campaign by a Portuguese bank focused on selling term deposits. A term deposit is a fixed-interest investment product that provides consistent income for the bank, making it essential to optimize outreach efforts for client conversion. This dataset is designed to help predict which customers are most likely to subscribe to a term deposit, providing valuable insights for efficient targeting in future campaigns and potentially reducing telemarketing costs.

II. Dataset

This project utilizes a publicly available dataset, widely referenced in data-driven marketing research and studies using the CRISP-DM methodology for predictive modeling. It provides detailed customer demographic, financial, and engagement information, making it a valuable resource for understanding customer behavior in direct marketing. Developed by Paulo Cortez (Univ. Minho) and Sérgio Moro (ISCTE-IUL) in 2012, this dataset was sourced from the UCI Machine Learning Repository, with added columns and random sampling to enhance research flexibility. *[S. Moro, P. Cortez, and P. Rita, A Data-Driven Approach to Predict the Success of Bank Telemarketing, Decision Support Systems, Elsevier, 62:22-31, June 2014]*

III. Research questions and proposed methods

The project will explore the following research questions using statistical models:

a) **What factors significantly affect a customer's likelihood of subscribing to a term deposit?**
   Model: Logistic Regression.
   Method: Logistic regression will assess how predictor variables (e.g., age, job, balance) impact the subscription probability. This model is suitable for binary outcomes, providing odds ratios for each variable and showing its influence on the likelihood of subscription.

b) **Is there a statistically significant difference in subscription rates across demographic groups?**
   Model: Chi-Square Test for Independence.
   Method: By analyzing categorical variables like job, education, and marital status, the Chi-square test will reveal if certain groups are more likely to subscribe. This will help identify demographic segments more receptive to term deposit offerings.

c) **How do previous campaign results impact current subscription rates?**
   Model: Survival Analysis (e.g., Cox Proportional Hazards Model).
   Method: Survival analysis will investigate how variables like pdays (days since last contact) and previous (number of prior contacts) affect subscription rates, helping determine if a client's responsiveness changes with time or frequency of contact.

d) **What is the optimal contact strategy in terms of frequency and timing?**
   Model: Generalized Linear Model (GLM).
   Method: A GLM will model how the number and timing of contacts (variables like campaign, day, and month) influence conversion likelihood, guiding optimal scheduling and frequency for future campaigns.

e) **Customer Segmentation Based on Subscription Likelihood**
   Model: Cluster Analysis (K-means or Hierarchical Clustering).
   Method: Clustering demographic and interaction data will segment clients into groups with varying subscription likelihoods. Each cluster's characteristics can be used to design targeted marketing strategies.

IV. Significance

This analysis will provide a data-driven basis for optimizing the bank's telemarketing strategy, enabling targeted, cost-effective outreach. Additionally, using proven statistical methods, this project offers a replicable approach for analyzing customer behavior in marketing contexts, contributing to the broader field of data-driven marketing strategies.

V. Team Details

Team Member Name	**NetID**
Hardik Lad	hlad2
Praveena Acharya	ppa3
Omkar Chalke	ochalke2
Rujul Khatavkar	rujulsk2
Radha Jawale	rjawale2

VI. Contribution Statement

Our team maintained a fair distribution of tasks, with each member taking on research, dataset exploration, and method selection equally. We collaborated closely, reviewed each other's work, and provided feedback, leading to a well-rounded proposal.