**Team 12:** Aaron Rock, Anushka Vuppala, Yashwant Bhaidkar  
**DATS 6103:** Summary Report Professor Edwin Lo  
December 15, 2022

Marketing Campaign of Portuguese Bank

**Introduction**

Bank telemarketing has played a vital role in creating revenue for banking institutions across the globe. By calling clients of each bank, it allows the bank to sell products and services that can be overcomplicated for many of us to understand. A popular bank telemarketing product to sell to clients are term deposits. Term deposits are where customers of the bank deposit money into the banking institutions account so that the bank may lend the money out to other businesses and or customers. As for the client who deposited the money to the bank, after the term duration is up the client will receive the amount initially deposited and with the compensation in the form of interest on the main deposit (Chen). The Portuguese Bank is a banking institution that relies on successful telemarketing campaigns to raise the overall revenue of the bank. In the data set examined for this assignment the bank would like to sell more of their long term deposits to existing bank customers.

**Data Set Description**

The data set for the Portuguese Bank was sourced from Kaggle and as mentioned from above it will focus maximizing the total subscriptions by clients for long term deposits at the Portuguese Bank. The data set consists of 17 variables with 45,211 data point observations. The variables can be split into two subcategories one involving features relevant to the customer and the other subcategory dealing with features related to the last campaign call. The data set does not contain any incomplete rows or missing data. For the customer relevant variables, it will contain age of customer, job sector, marital status, education level, has the customer defaulted on their credit before, has the ever taken a personal loan, and has the customer taken a house loan. The variables contact, day of the week, month, duration of call, campaign (number of times contacted for this campaign), pdays (number of days passed since contacted), previous (number of contacts performed last campaign), poutcome (outcome of the last campaign), and is subscribed (if the client subscribed to the term deposit).

**Exploratory Data Analysis**

***Age Variable***

Chart

Description automatically generatedChart, box and whisker chart

Description automatically generated

The average age is 41 years, the median is right below 40 years of age. When split by if the client has subscribed to a term deposit each group’s age is evenly distributed along the boxplot with the ‘Yes’ response having a taller whisker.

***Job Variable***

***Chart, bar chart

Description automatically generatedChart

Description automatically generated***

For the count plot, retired and student job sectors have the proportion of subscribed and not subscribed. To maximize profits, it would be wise to target these job sectors. Not only should these sectors be targeted, but by plotting the boxplot to the upper right we see that the higher retired ages and lower student ages say ‘Yes’ more.

***Marital and Education Variables***

Chart, bar chart

Description automatically generatedChart, bar chart

Description automatically generated

Marriage and education have the smallest effect on the target variable, but it should be noted that single people and people who are in tertiary school have the best chance to ‘Yes’ after analyzing the count plots above.

***Default, Loan, and Housing Variables***

***Chart, bar chart

Description automatically generatedChart, bar chart

Description automatically generated***

Chart, bar chart

Description automatically generated

Default is one of the most significant variables with clients who have defaulted on their credit only answering ‘No’ to the term deposit. With Loan and Housing there are only a select few that have taken these loans and then responded ‘Yes’ to the marketing campaign. From the analysis it is transparent that only targeting clients that have not taken and loan or defaulted on their credit is beneficial.

***Balance Variable***

***Chart, scatter chart

Description automatically generated Chart, histogram

Description automatically generated***

The scatterplot above plots shows no real correlation between age and amount of money in the persons banking account. A good portion of the data points keep a balance lower than 20,000. As for the line graph, there is massive spike at the age range of 80 – 90 who have subscribed. This is also the age range of retired people and could allow us to target them again like discussed earlier.

***Is\_Subscribed Variable (Target Variable)***

***Chart, pie chart

Description automatically generated***

**References**

Chen, J. (2022, September 15). *Time Deposit: Definition, how it's used, rates, and how to invest*. Investopedia. Retrieved December 16, 2022, from https://www.investopedia.com/term