# Targeted Marketing Proposal

Camilla McKinnon and Patric Platts

2024-10-28

Introduction

An effective marketing campaign to the right audience is important for companies trying to extend their reach or grow a product. This analysis looks at a specific bank's campaign. The bank uses a personalized marketing method to deliver individualized products to recipients. If the bank can better understand their customers, they can better target future marketing campaigns. This data set contains several variables on client characteristics (like age, job, education, loan status), method of contact, and when the contact happened. The response variable is whether or not the client opened a new bank account. One of the issues with the dataset is that the response rate is a binary variable (measured in 'yes' and 'no'). Traditional standard regression won't work since it would predict continuous values that would fall outside the 0-1 binary classification range. Additionally, many of our variables are categorical, which means they will need to be treated as factors in our model. Some of the factor levels have 'rare occurrences', meaning there are not many observations. Not accounting for that would lead to increased variance and unstable estimates. The following plot shows one of the variables, age, plotted against whether or not the client opened a bank account.
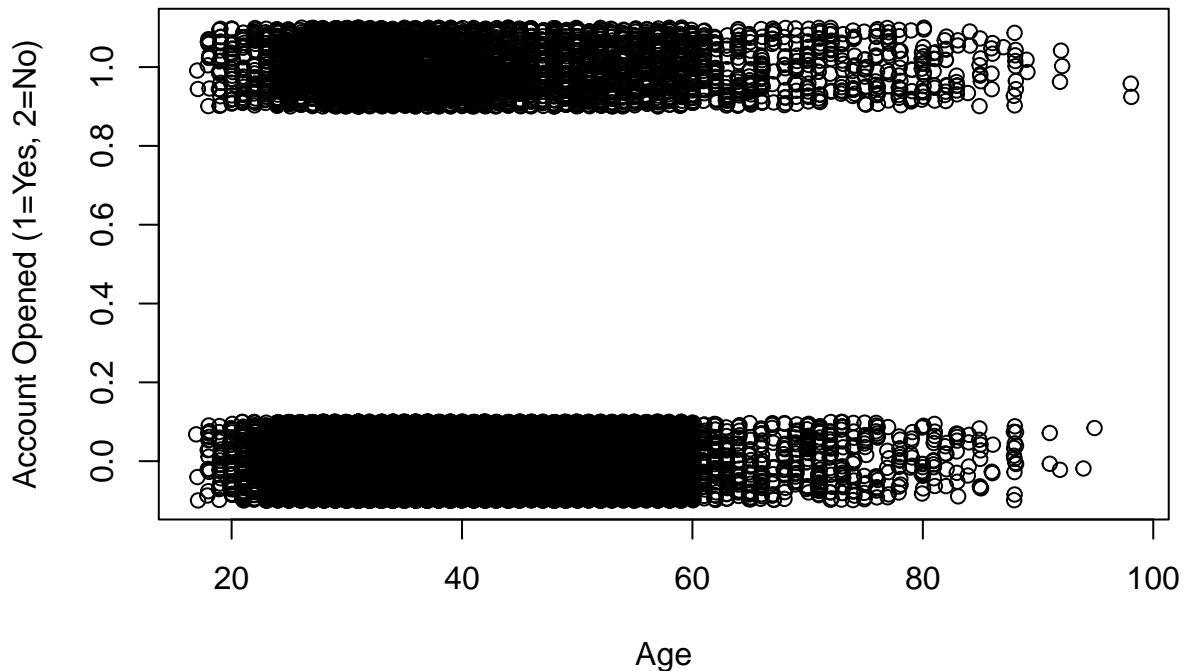
After creating a sufficient model, this analysis will explore the following questions: - What characteristics of customers are more likely to take out a new credit card? - Is there evidence that social media vs. personal contact is more effective in marketing? - Does repeated contacting seem to increase the likelihood of a person taking out an account?

Proposed Methods 1 & 2

Lasso

OLS

One of the models considered for this analysis was ordinary least squares (OLS) logistic regression. OLS is great for classification scenarios. With an OLS logistic regression model, we can look at coefficients and odds ratios to answer the research questions. A positive coefficient for a characteristic means that as that factor increases, or for a certain level, the likelihood of opening an account increases. Additionally, we can compare odds ratios of social media and direct contact to see which is more effective. The "both ways" variable selection method was used to trim down and improve the performance of the model. Assumptions required for this model are that there is independence between observations, linearity, normality and equal variance of residuals.