



# Bank Marketing Data Visualization

## EM622 Final Project

Yuxin Cao, Mehul Kumar Jain, Sonali Pandurang Raut

### Introduction

The data is related with direct marketing campaigns of a Portuguese banking institution. The marketing campaigns were based on phone calls. Often, more than one contact to the same client was required, in order to access if the product (bank term deposit) would be ('yes') or not ('no') subscribed. By visualizing the data, this project aims at predicting the features of the client who will subscribe.

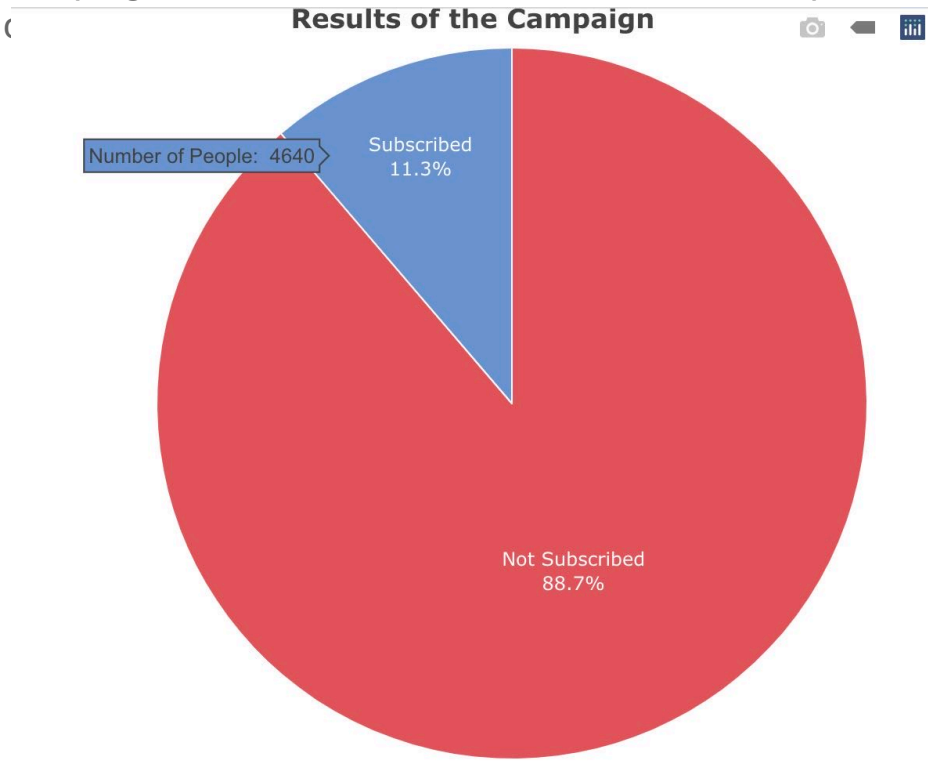
### Data Analysis

After reading and understanding the data, we selecting the important variables to work with. We converted some data types and cleaned the data for analysis. We chose several types of visualizations to present our analysis:

Pie chart, Histogram, tree map, scatter plot, interactive line chart.

### Final Results of the Campaign

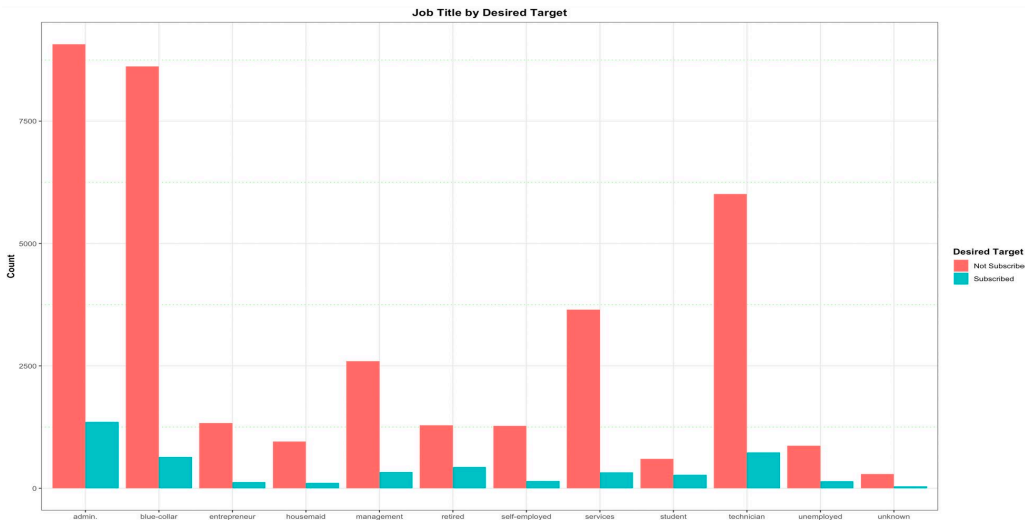
Since we wanted to analyze the share of the subscribed and not subscribed people, pie chart is the best type of visualization. It gives us details about the percentage and the total number of people who subscribed and did not subscribe to the service (term deposit). We can see that around 10% of the targeted audience opted for term deposit which seems to be relatively low. But we guess that's how campaigns work, we need to make thousands of phone



### Relationship between Desired Target and different categorical variables

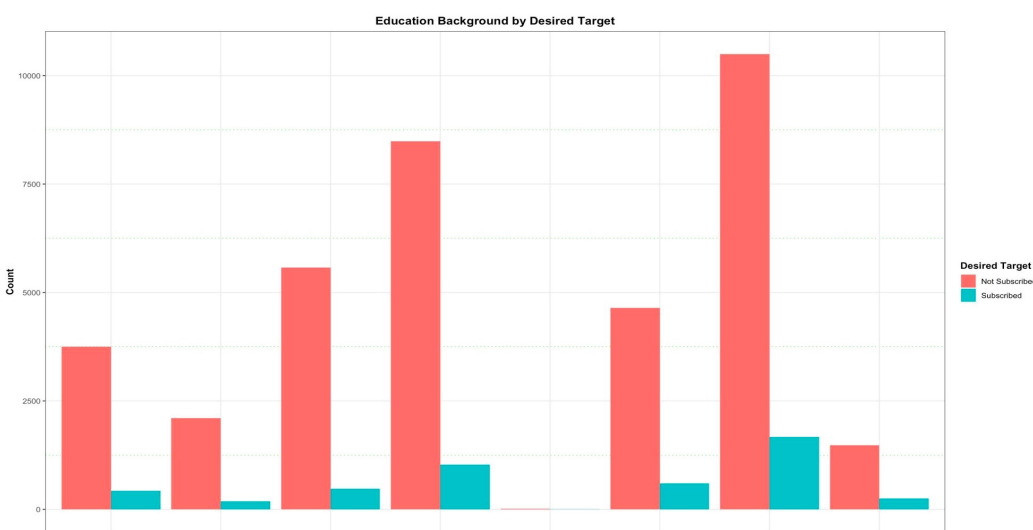
We chose histograms for the above visualizations because, we wanted to analyze the counts of the different categorical variables and histograms are the best type of visualizations to provide us with the required information.

#### 1. Job Tittle



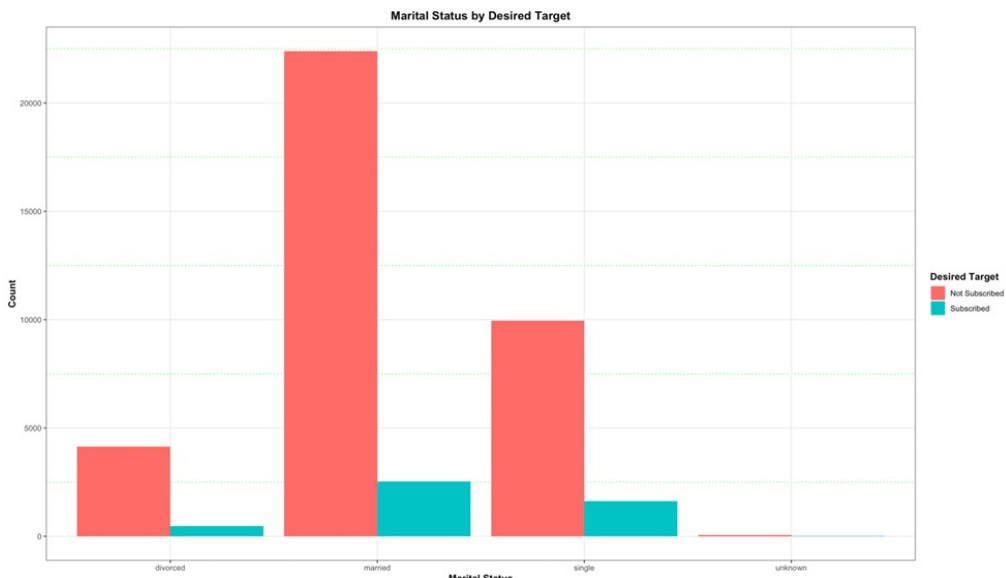
From the above graph, we can see that most of the targeted audience and the people who actually subscribed to the service (term deposit) were admins by profession. Around 1250 people who subscribed to service were admins. The campaign also targeted people with blue-collar jobs and out of the people with blue-collar jobs, around 800 of them subscribed to the service (term deposit).

#### 2. Education Background



The above visualization gives us insights about the Educational background of the targeted audience and the subscribers. We can see that, most of the targeted audience and the subscribers had a university degree. Around 1500 people with a university degree subscribed to the service. Many people with just high school education also subscribed to the service.

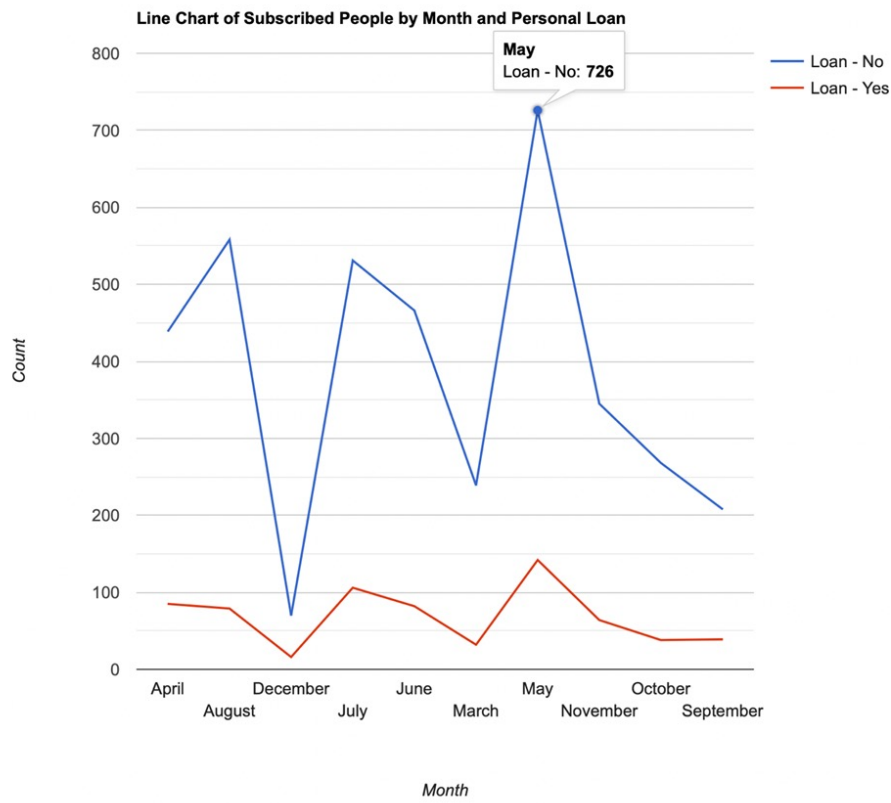
### 3. Education Background



The above visualization gives us information about the Marital status of the targeted and the subscribed audience. We can see that most of the targeted audience and the subscribers were married.

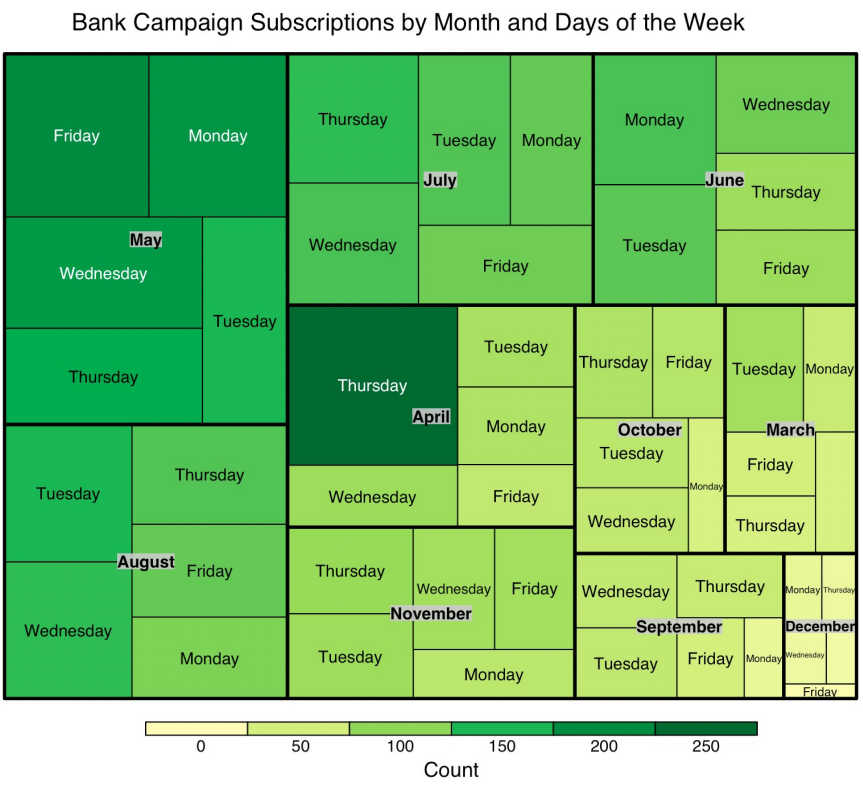
### Subscriptions based on combined categorical variables

#### 1. Month and Loan



The above visualization is an interactive line chart between the month and the count of the subscribed people who have personal loan or not. We can see that most of the people who subscribed to the service (term deposit) did not have personal loan. Also, in the month of May during which the campaign was relatively more successful many people did not have loan. The interactive visualization gives us the exact count of the subscribers who had personal loan or not in every month.

### 2. Month and Days of the Week



The above uses a tree map to get insights about the months and the days of the week in which the campaign was more successful. We can observe that, the campaign was more successful in the months of May, June, July and August. We can notice that highest number of subscriptions is on Thursdays in the month of April. We choose tree map because it is easier to understand the contributions by different categories.

### 3. Duration, Age, Outcome and Contact Type



The scatter plot between the duration of the call and the age of the subscribed audience, split by the outcomes of the previous campaign. The outcome of most of the subscribed audience was either successful or nonexistent and the highest contact type was cellular. Also, there are very less subscribers whose previous campaign was failure. the most the subscribers belonged to the ages from 20 to 50 years.