

Bank Marketing Analysis

I. Introduction

One of the Portuguese banking institutions managed a marketing campaign based on phone calls. The output of the campaign is a dataset of customers' data who were included in the program. The target of this analysis is assessing the efficiency of this campaign, comparing with the previous one and defining the factors which influence the result of the campaign. Overall, the duration, number and time of contacts and the age of clients are the most important features affecting the outcome of this campaign.

2. Analysis

2.1 Data Understanding

The dataset contains various numerical and categorical features. It has 45211 rows of data and 17 features. None of the rows contain a "null" or "duplicated" value. The explanation of the concept of each feature is as follows:

Age: This is the age of a client in this campaign ranging from 18 to 95 years old.

Job: The job title is inserted into this column. It has 12 categories and 288 clients marked as "unknown".

Marital: This is a categorical feature which consists of three categories: single, married, and divorced.

Education: It is the educational state of clients. Four categories are defined in this feature.

Default: It expresses whether the client has credit in the bank or not? This column is filled with "yes" or "no".

Balance: It specifies the actual balance in the client's account. It is a numerical feature and covers positive and negative figures.

Housing: Used to define if the clients have a housing loan from a bank or not.

Loan: This column is related to personal loan. It contains "yes" or "no" attributes.

Contact: This categorical feature expresses the means of communications in the campaign. It can be cellular or via telephone. About 28 percent of clients have "unknown" registered for this variable.

Day: Which day of the month the client was contacted.

Month: The month in which the client was contacted.

Duration: The duration of the call in seconds for the last time they were contacted in the campaign. The duration of calls of many clients were less than 5 seconds.

Campaign: This numerical feature specifies the number of times that a client was contacted.

Pdays: Number of days that passed from the last contact in the previous campaign. It is shown as "-1" if there were not any contacts in the previous campaign.

Previous: Number of contacts with a client in the previous campaign.

Poutcome: The outcome of the previous marketing campaign is specified in this column. There are many clients who did not contribute to that who are defined as “unknown”.

Y: To mark the result of clients contributing and subscribing to the deposit. This categorical feature contains a “yes” or “no” attribute.

First of all, a process of data cleaning was performed including removing outliers, categorizing age and mapping categorical columns to 0 and 1. Consequently, the dataset was converted to 44691 rows and 18 features which were ready to be analyzed. Then, a suitable KPI to assess the efficiency of the campaign was defined. Finally, the items that affect the result of the campaign were taken into account and a method to grow the outcome in the future was suggested.

2.2 KPI

○ Conversion Rate:

This KPI illustrates the rate of clients who accepted the campaign among all clients who contributed. It is noteworthy that while the conversion rate for the new campaign is **11.6 percent**, this rate for the previous one was **18.4 percent**. It shows that the conversion rate has declined. Obviously, the process should be optimized for the next campaign as per the suggestions which will later follow.

2.3 Effective Factors:

1. **The Number and Duration of Contact:** One of the most significant features in this campaign was the duration of calls. The shorter the phone call, the less possibility of success there was... In contrast, the more the clients were called, the less successful it was in the campaign.
The related charts illustrate while the most of clients who were unsuccessful in this campaign had contacts of less than four seconds, this figure for successful clients was 13 seconds. Additionally, the scatter plot demonstrates there is a negative relationship between the number of contacts and success rate. Calling clients less than four times made a significant difference in the outcome.
As a result, the **quality of contacts** is more important than their quantity. It is necessary that bank officers try to explain more about the campaign and answer clients' questions in order to establish trust and create enthusiasm in clients.
2. **Effect of Time Contact:** Although the most contacts were in May and on the 20th of the month, the most successful rate happened in March followed by September, October and December and on the first day of the month.
It is noteworthy that the age of clients affected the results as well. For instance, the best outcome for young clients (18-24 years old) occurred in January and for seniors (older than 64 years old) was in December. On the other hand, the day of contacts for the young reached a pick on the 22nd of the month, while for young adult and adult groups hit the highest point on the first day of the month. It is evident that the line graph for young and senior groups shows some fluctuation and for young adult and adult groups they were highly similar in these two aspects.

3. **Effect of Age:** While the Young Adults group (25-44 years old) has the greatest number of clients who participated successfully in the campaign with 64 percent, the senior group (+64 years old) has the greatest percentage of success. These figures demonstrate that Young Adults group is the most significant client of the bank so they should receive more attention to their preferences and requirements. On the other hand, senior clients are the appropriate groups of clients that can be easier to engage.
4. **Previous clients:** 64.7 percent of the clients who participated successfully in the previous campaign, took part successfully in the new campaign as well. At the same time, participation rate for new clients was just 9 percent. Although the number of new clients was much more than previous clients, the percentage of success rate for previous ones was more. It shows that positively previous clients have more recognition and trust for the bank. Thus, contacting previous positive ones imposes less cost on the bank with a higher success rate.

2.4 Further Factors:

There are five other factors that affected the outcome of the campaign but are not as important as those mentioned above.

- The **distance of contact** between previous and new campaign was important. Clients who were contacted more than two years or less than six months distance, had a better success rate.
- Clients with middle **balance** (between 3000 and 13000 dollars) had a better conversion rate.
- Cellular **contact method** had a greater percentage of success rate than contact via telephone.
- About **job title**, clients who worked as a retiree or students had better conversion rate.
- Clients who had tertiary or secondary **education** had more tendency to participate in the campaign.

3. Conclusion

To sum up, based on what has been said above, the most important features affecting the outcome of this campaign are: the duration, number and time of contacts and the age of clients. There are some factors which with less importance such as: distance of contacts, balance, contact method, job title and education.

It is noteworthy that the bank has a large group of clients with specific attributes such as their ages and job titles whose preferences and requirements were not paid attention to, so their successful rate of contributing was not satisfactory. In contrast, there is a smaller group of clients with more enthusiasm to participate in the bank campaign with more success rate.

It seems that to increase the achievement of a similar future campaign, the bank should change the target of the campaign to suit large groups of clients or try to absorb more clients with similar situations that have more successful conversion rate.