

TITLE: REPORT ON THE DATA ANALYSIS OF FACTORS THAT LED TO THE SUCCESS OR OTHERWISE OF THE CAMPAIGNS ON POTENTIAL CLIENTS IN SUBSCRIBING TO A BANKING TERM DEPOSIT IN PORTUGAL

PROJECT OVERVIEW

This data analysis project aims to provide insight into campaign strategy of a banking institution in Portugal who has recently introduced a term deposit to his services.

In a bid to create awareness of the service, the marketing team engaged in massive campaign, which involves making phone calls to potential clients. Multiple contacts were often made with the same client to determine whether they would eventually subscribe to the term deposit or not. The outcome was then recorded as a yes or no.

By analysing the various aspects of the dataset, using historical data provided, we carry out a descriptive, diagnostic and prescriptive analyses. This is done to identify trends, make data-driven recommendations and gain a deeper understanding on the effectiveness of the campaign of the banking institution.

DATA SOURCES

Bank Subscribed Dataset: The primary dataset used in this analysis is the "bank-subscribed_dataset.csv" file containing detailed information about various characteristics of the bank's potential client and other external features.

TOOL USED

Microsoft Excel and its Advanced Functionalities

DATA PREPROCESSING AND CLEANING

In the initial data preparation phase, the following tasks were performed:

1. Data loading and inspection
2. Removal of blank and empty spaces
3. Formatting and conversion to numeric in preparation for data modelling.

EXPLORATORY DATA ANALYSIS

Exploratory data analysis involves exploring the bank subscribed dataset to answer key questions such as:

1. How impactful is age and marital status on the likelihood of subscribing to the term deposit?
2. Does the type of job clients engage in influences their choice of subscribing to the term deposit?
3. How impactful are the duration of calls on the choices made by potential clients to subscribe to the term deposit?
4. What set of features bare more impact on the choice of clients in subscribing to the term deposits?

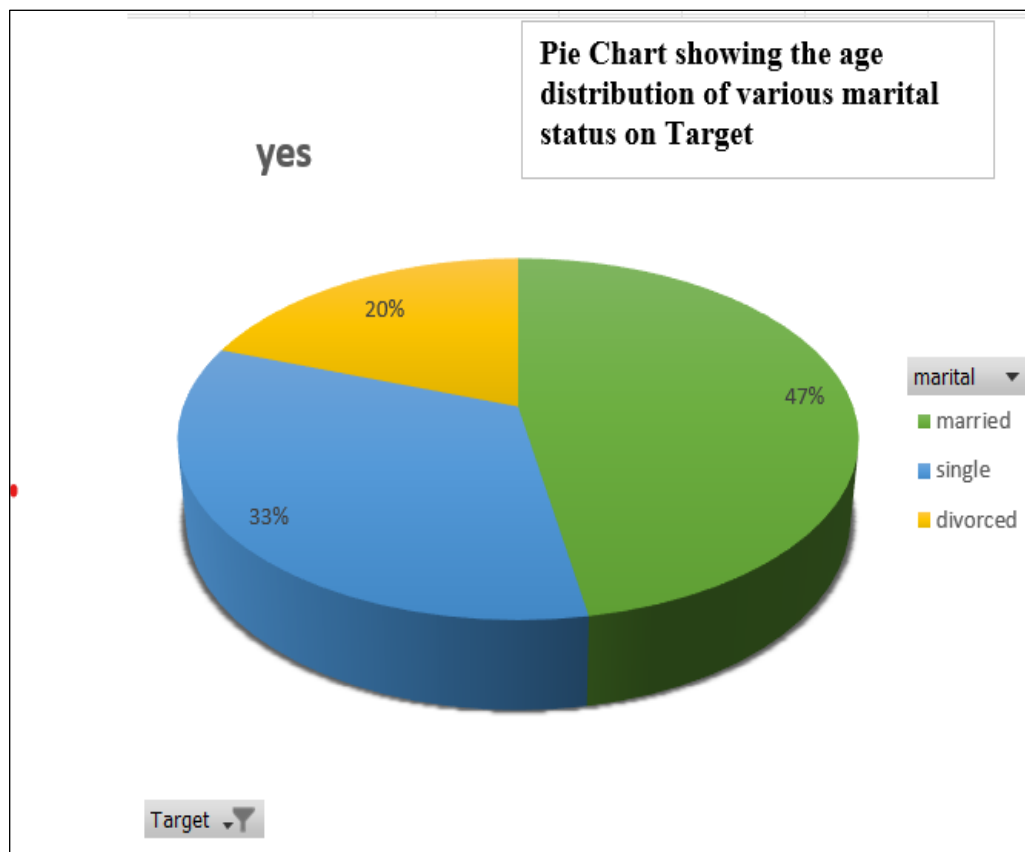
DATA ANALYSIS

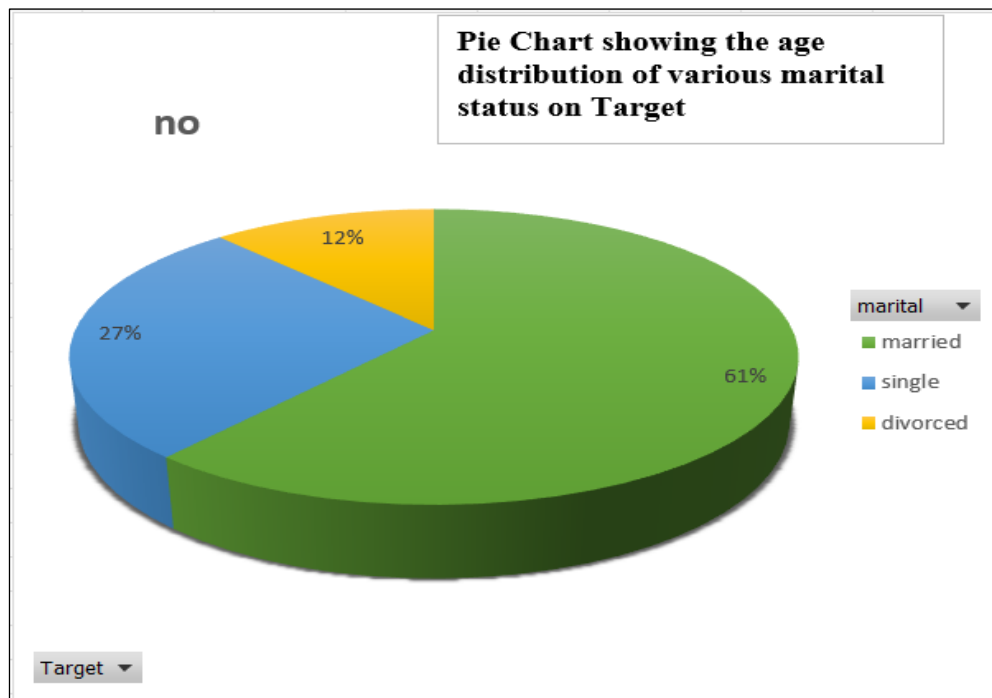
Microsoft Excel Power Query was used.

FINDINGS/RESULTS

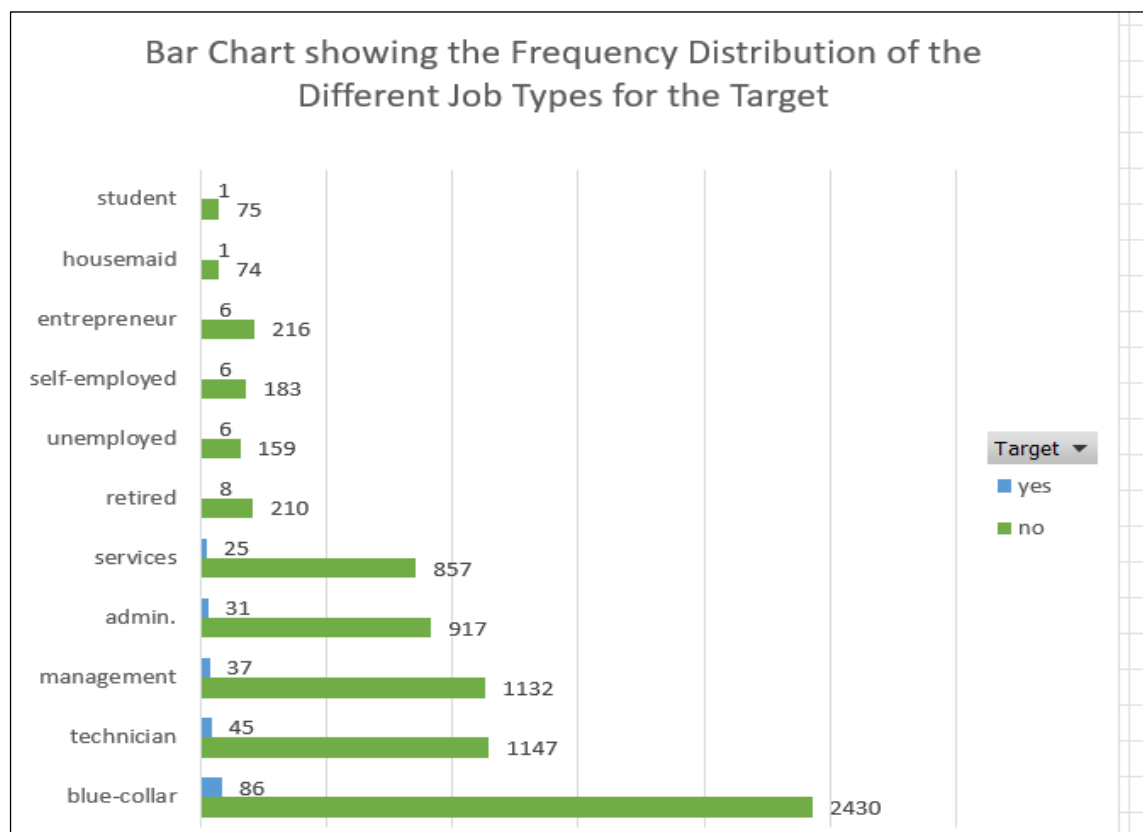
The analysis results and findings are summarized as follows:

1. From the pie chart, higher percentage of those who are married subscribed to the term deposit, with a percentage value of 47%. Of this, 51% represents those who are in the age bracket of 31 to 48 and are married.

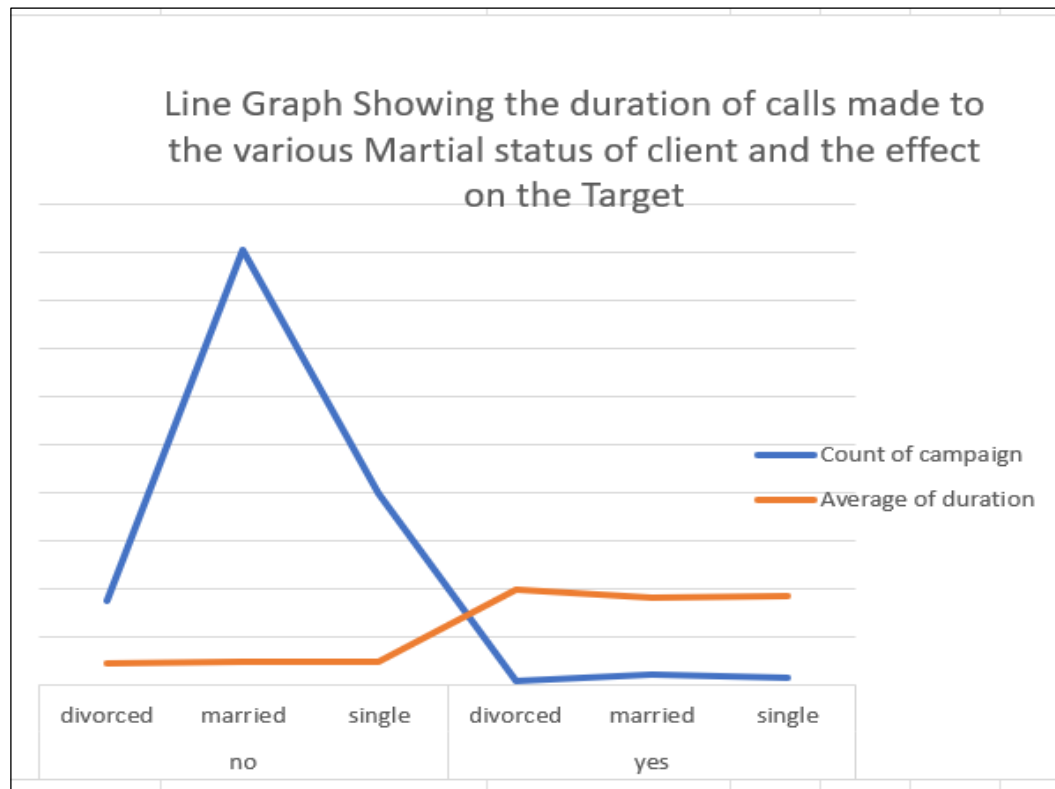




- From the chart, majority of those who accepted the new term deposits are those who are gainfully and actively employed either in the public sector or private sector, with a total number of 224, representing over 80% of the overall proportion of those who accepted the term deposit.



3. As shown by the line chart, those who accepted the term deposit had more talking time duration. This must have influenced their decision as evidenced by the average duration of calls made, which is about 3.5 times the average duration of calls made to those who rejected the new term deposit.



4. In determining the factors that played key roles in the success or otherwise of the campaign, a linear regression analysis was performed, and the results are as shown.

<i>Regression Statistics</i>	
Multiple R	0.497583882
R Square	0.24758972
Adjusted R Square	0.246407614
Standard Error	0.154940594
Observations	7651

ANOVA								
	<i>df</i>	<i>SS</i>	<i>MS</i>	<i>F</i>	<i>Significance F</i>			
Regression	12	60.33759201	5.028132667	209.4480373	0			
Residual	7638	183.3623165	0.024006588					
Total	7650	243.6999085						
	<i>Coefficients</i>	<i>Standard Error</i>	<i>t Stat</i>	<i>P-value</i>	<i>Lower 95%</i>	<i>Upper 95%</i>	<i>Lower 95.0%</i>	<i>Upper 95.0%</i>
Intercept	0.845960646	0.028895463	29.276591	1.1459E-178	0.789317604	0.902603689	0.789317604	0.902603689
age	0.00013913	0.000198269	0.701724476	0.482872387	-0.000249531	0.000527791	-0.000249531	0.000527791
job	-0.000703529	0.000804347	-0.874658798	0.381787086	-0.002280269	0.000873211	-0.002280269	0.000873211
marital	0.012073541	0.00256294	4.710817401	2.51085E-06	0.007049475	0.017097606	0.007049475	0.017097606
education	-0.002582715	0.003002286	-0.860249606	0.389678466	-0.008468019	0.003302589	-0.008468019	0.003302589
default	0.005926795	0.012628466	0.469320264	0.638854129	-0.018828467	0.030682057	-0.018828467	0.030682057
balance	-2.00161E-07	7.38121E-07	-0.271176827	0.786262363	-1.64708E-06	1.24676E-06	-1.64708E-06	1.24676E-06
housing	0.005566695	0.006018441	0.92493982	0.355026433	-0.006231101	0.017364492	-0.006231101	0.017364492
loan	-0.00579489	0.004943091	-1.172321023	0.241104757	-0.015484706	0.003894927	-0.015484706	0.003894927
day	0.000938424	0.00022338	4.20102567	2.68731E-05	0.000500538	0.001376309	0.000500538	0.001376309
month	0.049518285	0.017570406	2.818277773	0.004840615	0.015075464	0.083961106	0.015075464	0.083961106
duration	0.000351791	7.09838E-06	49.55939474	0	0.000337876	0.000365706	0.000337876	0.000365706
campaign	0.00048048	0.00064111	0.749450127	0.453609029	-0.000776272	0.001737232	-0.000776272	0.001737232

From the table, the value of multiple correlation coefficient is 0.4976, which shows that the model moderately measures a linear relationship between the target and features.

Also, the value of R – squared shows that 24.96% proportion of variance in the outcome of the target can be explained by the selected features. This low value calls for further investigation to determine the features that played dominant role, using the p-value. Thus, based on the low p-values, the four factors that are statistically significant are:

- Duration
- Month
- Marital
- Day

RECOMMENDATIONS

Based on the findings above, the following recommendations are put forward:

1. An encompassing targeted and concentrated campaign mechanism should be channelled to include more potential clients who are married as this would result to a rise in acceptance.
2. It is obvious that those who are gainfully employed favours the new term deposit. Hence, much more efforts should be reach out to those who

actively earning a living, and this will result in increase in acceptance, and which will in turn result in the success of the campaign.

3. Improve duration of calls as this has a direct and big impact on the acceptance of the term deposit. In addition to this, other means of campaign other than calls such as distribution of fliers, use of mails and so on can be helpful.
4. From the analysis, since four factors plays major role and have more impact on the choice of potential clients accepting the new term deposits, it is very important to include other features such as geographical considerations, effect of economic policies on potential clients and analyse their effectiveness on the campaign.

CHALLENGES ENCOUNTERED

Some of the challenges encountered while carrying out this project include:

- Removal of records that were blank and I had to modify the categorical variables to perform regression analysis.
- Omission of key features from the dataset.

ONGOING/FUTURE PROJECTS

- Analyzing the dominant factors responsible for Hepatitis C Virus.