Bank Marketing Campaign- Presentation

15-May- 2021

Who are we?

This project was done by the **Data Team** which comprised of the following team peers:

- Fawzi El Khatib, fawzi_khatib@hotmail.com, Lebanon, Data Glacier, Data Science.
- Nigel Chitere, tawanda.nigel@gmail.com, Zimbabwe, Rocapply, Data Analysis.
- Malipalema Khang, malipalema@gmail.com, South Africa, University of the Witwatersrand, Data Science.
- Arthur Mupfumira, artmupf@gmail.com, Zimbabwe, Data Glacier, Data Science.

Github repository link:

https://github.com/Fawzikh/DG-BankMarketing-Campaign.git

Project Problem Description:

In this case, from the given company proposition, the problem identified and to be addressed was:

The goal of the project is to develop a model that will aid our client ABC Bank **identify** which clients are **likely to be interested** in their **term deposit product** before actually introducing this product into their customer base without prior research and client targeting.

Project Breakdown Approach:

In response to tackling the problem brought forward the following four step strategy was used:

1. Data understanding.
2. Data insight analysis.
3. Recommendations
4. Modelling Technique.

Project Breakdown Approach: (1) Data Understanding.

The given data involved namely two data sets

1. Bank-full.csv- mapped out all the information regarding the age, marital status, education as well as job type of the clients just to mention a few. **The dataset had 45211 rows of data and over 16 attributes.**

2. Bank-additional.csv- mapped a list of additional information regarding the duration, campaigns, month, housing, defaults, loans ,etc that were related to the clients. **The dataset had 41188 rows of data and over 20 attributes.**

Project Breakdown Approach: (1) Data understanding (cont)

Quick pointers regarding the data:

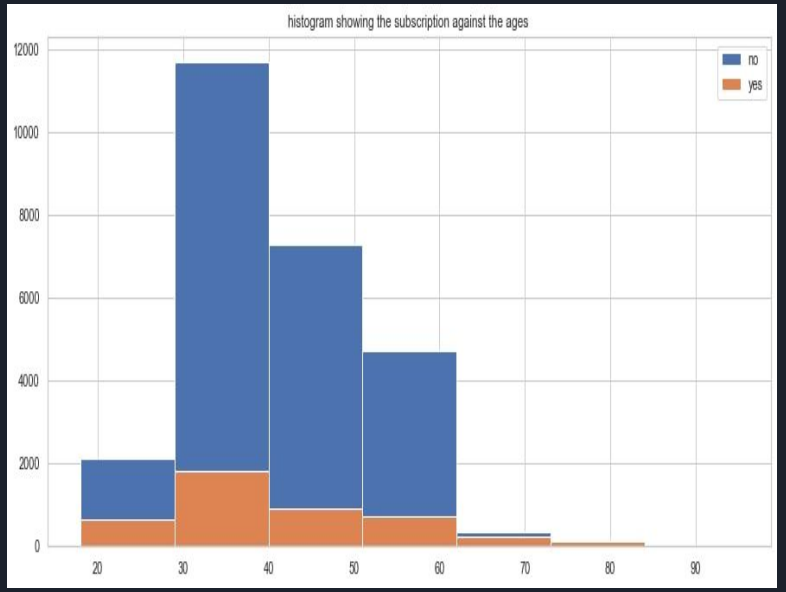
- Total of 31 data features from the overall data sets
- Timeframe of the data: from May 2008 to November 2010
- The data had a significant number of missing values, and was taken from a random sample of 10% of the actual real life bank data.
- In Spite of the data sets having a combined number of 31 attributes if considered together, they were a couple of features which our team examined and determined these as being key inputs for the building of a predictive model.
- These include:
- ‘age, default, marital status, balance, loans, education, job.’

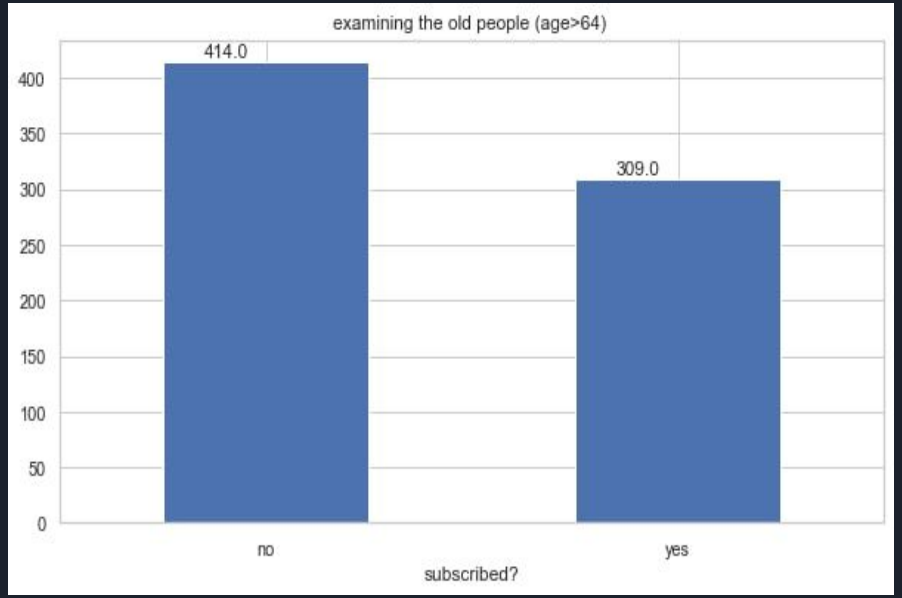
Project Breakdown Approach: (2) Data insight analysis.

In understanding each given dataset it was imperative to also establish the relations between the data features and how these were actually of any use in helping to answer or solve the problem at hand.

Hence this resulted in a number of visual graphics being deduced (as will be shown in the following slides) in order to highlight the significance of some of the features within the given data sets.

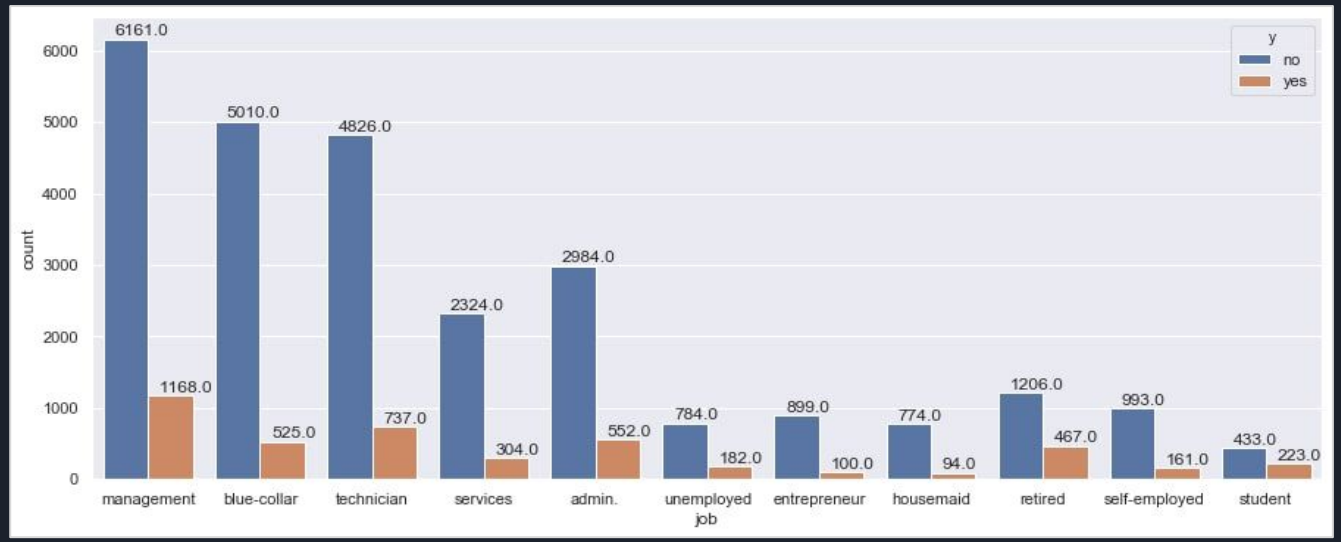
Project Breakdown Approach: (2) Data insight analysis (Age Feature Analysis).





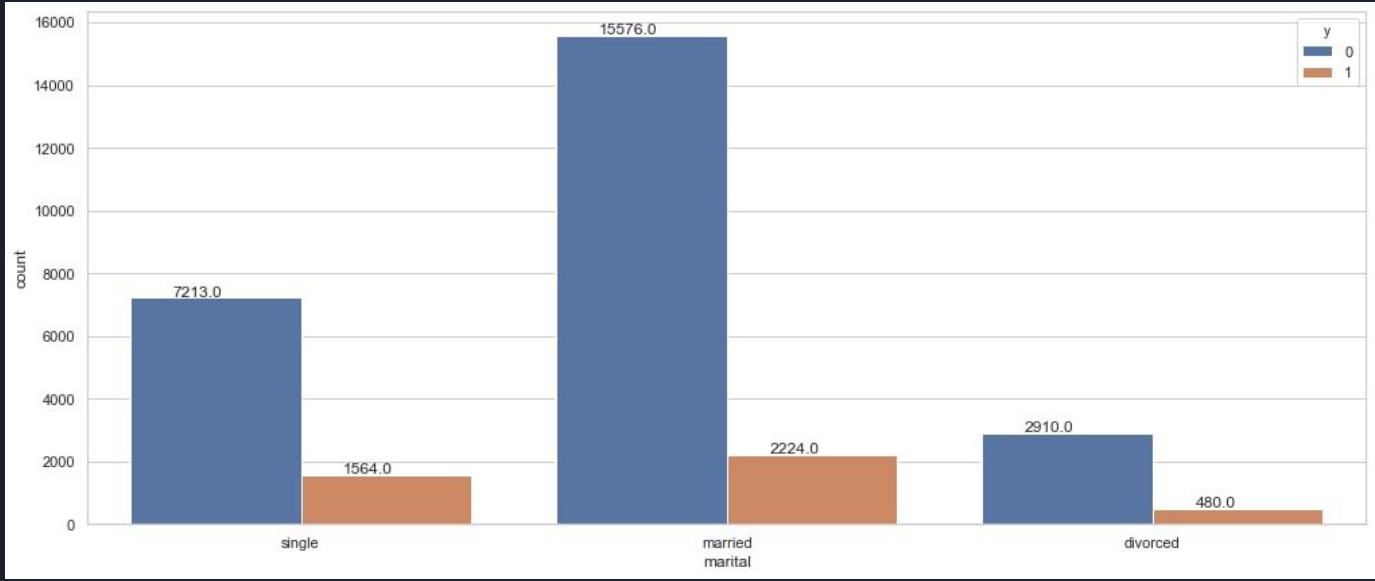
The age feature actually presented interesting insights as to which age range could afford the product. Given that the term deposit is for people with extra savings and willing to put their money away for a prolonged period of time the first group to be considered was the **'retired and working class' population**. 43% over 64 agreed to the new product, while only 13% of the working age population said yes. Hence this allows us to narrow in our target group to the retired age group.

Project Breakdown Approach: (2) Data insight analysis (Job Feature Analysis).



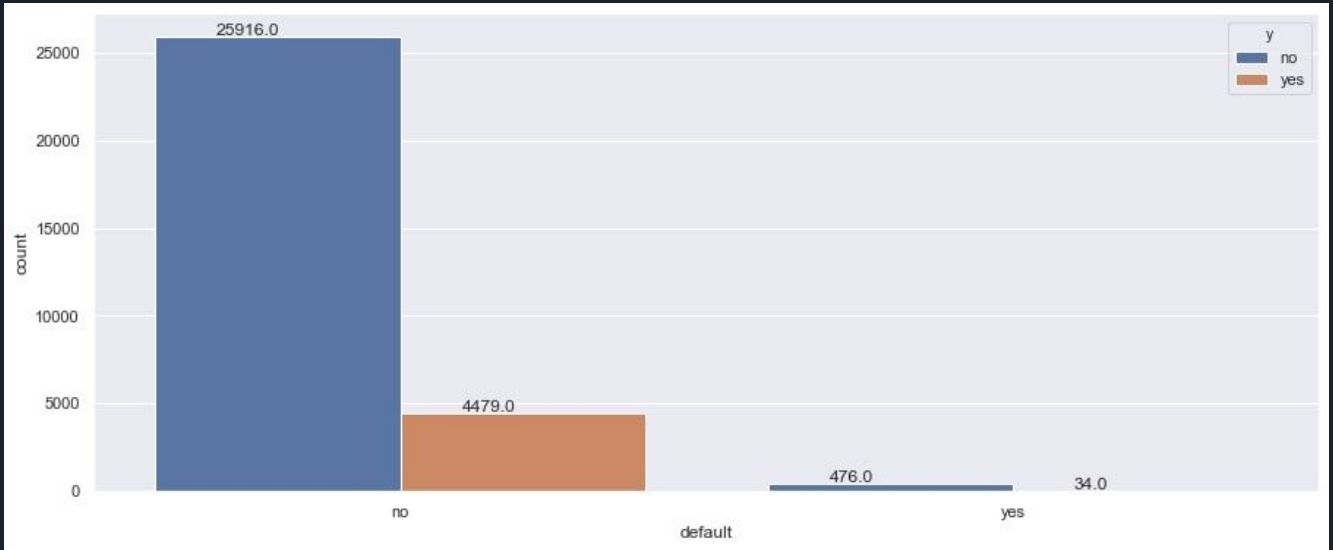
As is visible from this graph over 80% of the working population said No however what is interesting is the ratio of yes to no for students as 34% of the students asked agreed try out for this new campaign meaning they are a great target. Followed closely by the retired at 28% and lastly admin at 16%. This supports the previous slide that the age group over 64 years have a higher propensity to be interested in this product.

Project Breakdown Approach: (2) Data insight analysis (Marital Feature Analysis).



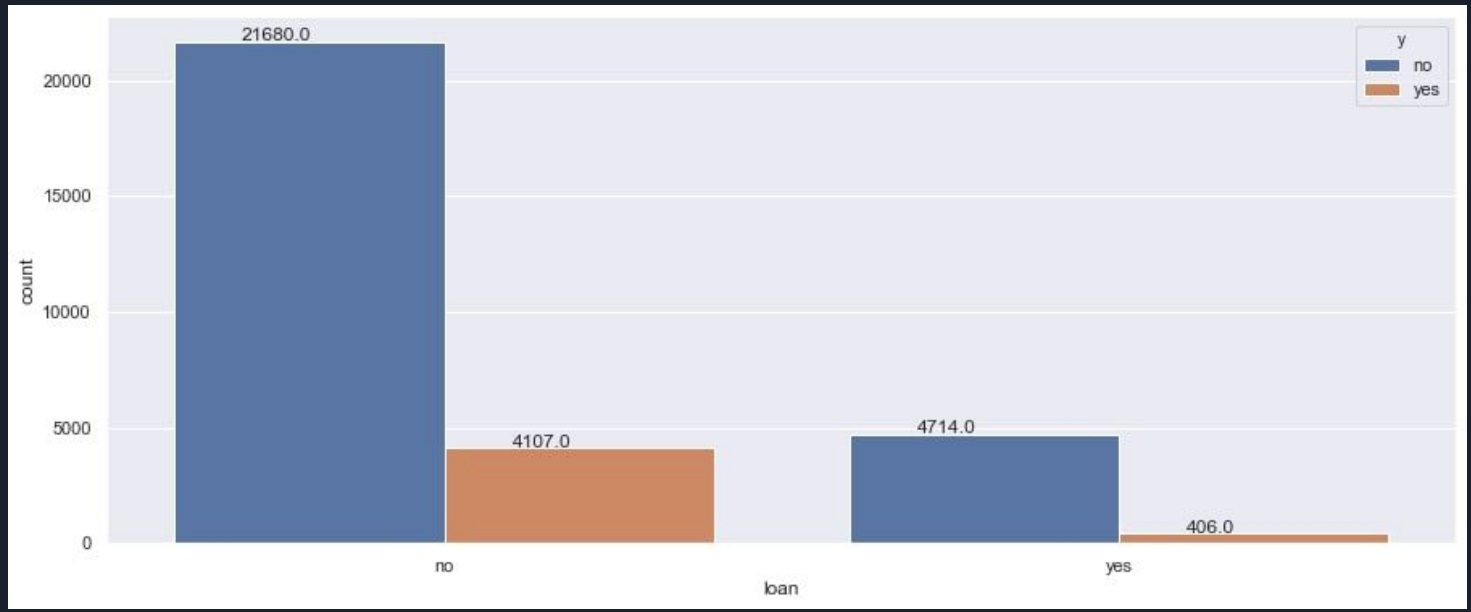
The analysis showed that the highest ratio of those who agreed were single at 22% followed by divorced at 16%. This perfectly aligns with the assumption that these social groups are likely to be interested in the term deposit product as they may not have several responsibilities such as children, school fees, joint loans or mortgages, etc. Hence these groups would be a perfect target for marketing campaigns.

Project Breakdown Approach: (2) Data insight analysis (Default Feature example).



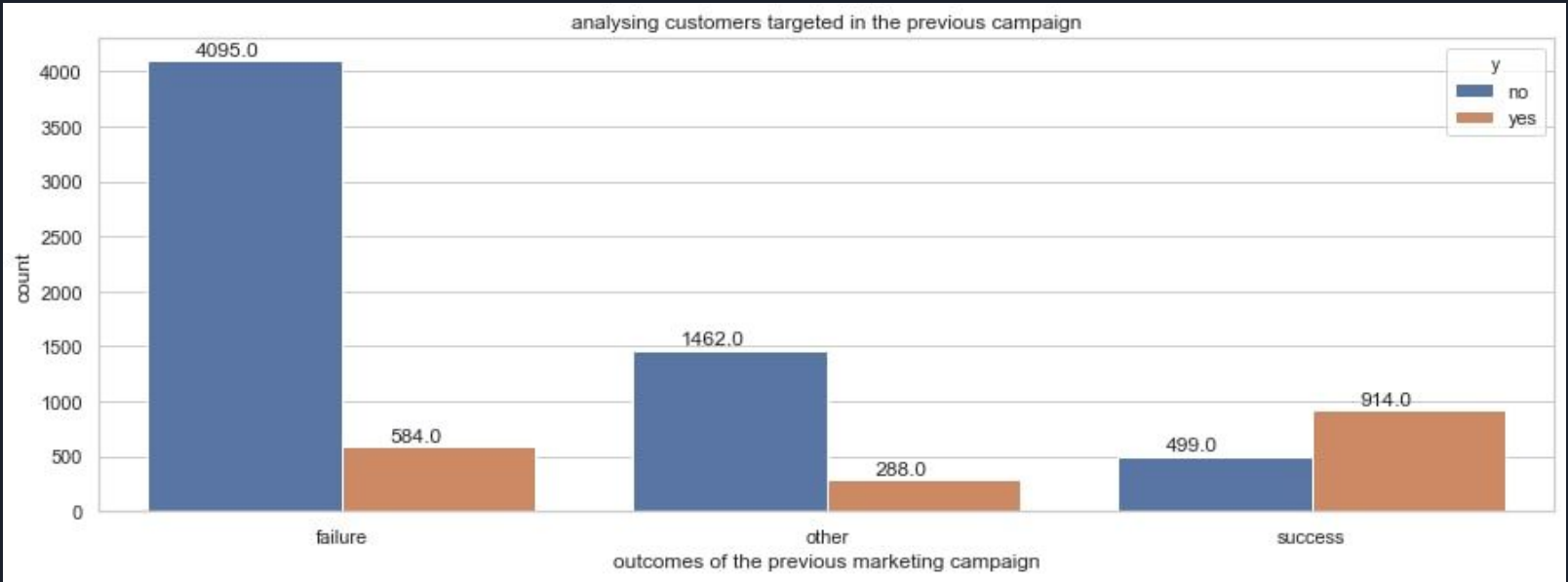
The graphical image shows the two categories of clients interest on the term deposit from those who defaulted and did not default. 19% from those who have not defaulted on previous bank commitments agreements agreed where as only 8% of those who have defaulted said yes. This will allow us to focus on those who have **not defaulted** as their previous commitment history counts in their favour.

Project Breakdown Approach: (2) Data insight analysis (Loan Feature example).



A point of interest from this visualization is that the majority of those interested in the term deposit were clients who currently don't have any loan commitments. From a business perspective this is a positive insight as such individuals are likely not to be affected by bad previous credit score and are less like likely to default and have more money for savings as they dont have any current financial commitments with the ank.

Project Breakdown Approach: (2) Data insight analysis (Previous Campaign Feature example).



From the illustration, the largest ratio 65% of those interested in the new term fixed deposit, are clients who were successfully targeted in the previous campaigns. This could be to a variety of factors e.g. good customer retention, satisfactory terms and conditions etc, but more importantly it allow the marketing team to work with clients who are already keen to try new products and have a history with the bank of previously

Data Insight Recommendations

After having analyzed the data above, the following recommendations were deduced:

1. The marketing campaign has to target retired people as 43% of them advised they would be interested in the new product.

2. From a marital status point, over 30% of those interested were the single and divorced hence these two bins would make attractive target market points.

3. The students, retired and admin groups in the employed attribute were more keen to try out the feature therefore narrowing down the clients to focus on based on employment status.

4. Lastly with regards to previous client history consideration, it came as no brainer that clients who had not previously defaulted, did not have an current loan commitments and also those who were successfully subscribed for the previous campaigns made the majority of those interested in the new term deposit feature. Consequent backed by a good hostical standing such clients should be first for marketing consideration in this project.

Modelling Technique

- We will be working on the Bank-Full data set.
- We will be implementing three **logistic regression models** , all with dummies for categorical nominal features.
- In the data set, we will drop the **attribute ‘pdays’** as it didn't not bear much impact on the model prediction.
- Of the logistical regression models, 2 of them were s**tandardized for the numerical regression features**. The last model will be built without standardization and duration feature so as to test the general accuracy.
- We will also make one **random forest model**, but without dropping the first dummy features as this model is not affected by multicollinearity.

Acknowledgements

Special appreciation to

Fawzi and Nigel for this presentation.

The end..Thank you!