

Bank Marketing

Campaign

Aly Medhat



Data Glacier

Your Deep Learning Partner

TABLE OF CONTENTS

01

Introduction

02

EDA & Recommendations

03

Model Recommendations

04

Evaluation
Recommendations

05

GIT Repo.

Data Glacier

Introduction

BACKGROUND

ABC Bank wants to sell its term deposit product to customers and needs a model that can predict whether a particular customer would buy a product or no.

OBJECTIVE

Build a ML model for the bank that shortlist customer whose chance of buying the product is more.

Introduction

BUSINESS UNDERSTANDING

The data is related with direct marketing campaigns of a Portuguese banking institution. The marketing campaigns were based on phone calls.

Project Lifecycle



TABLE OF CONTENTS

01

Introduction

02

EDA & Recommendations

03

Model Recommendations

04

Evaluation
Recommendations

05

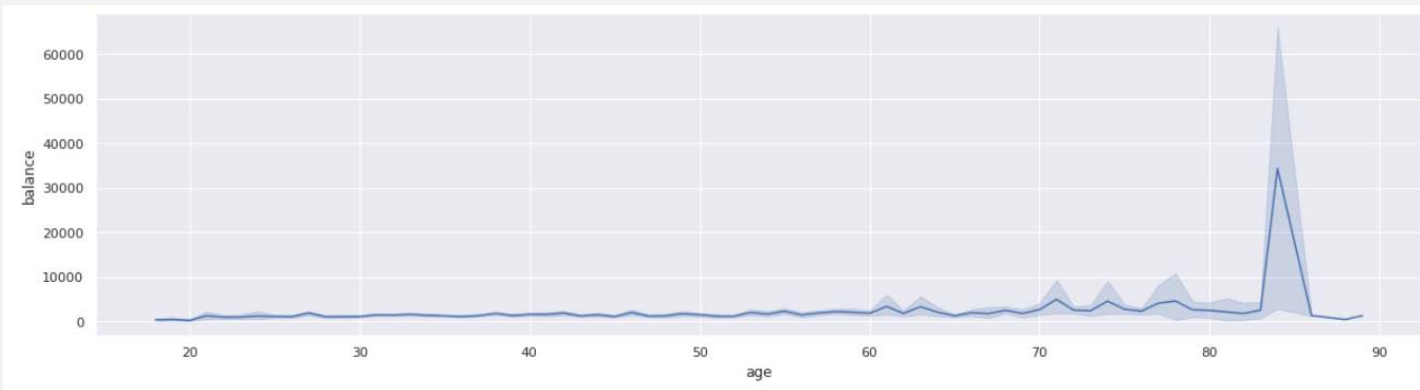
GIT Repo.

Data Glacier

ANALYZING THE DATA

“Exploratory Data Analysis”

The relation between the bank balance and the client age

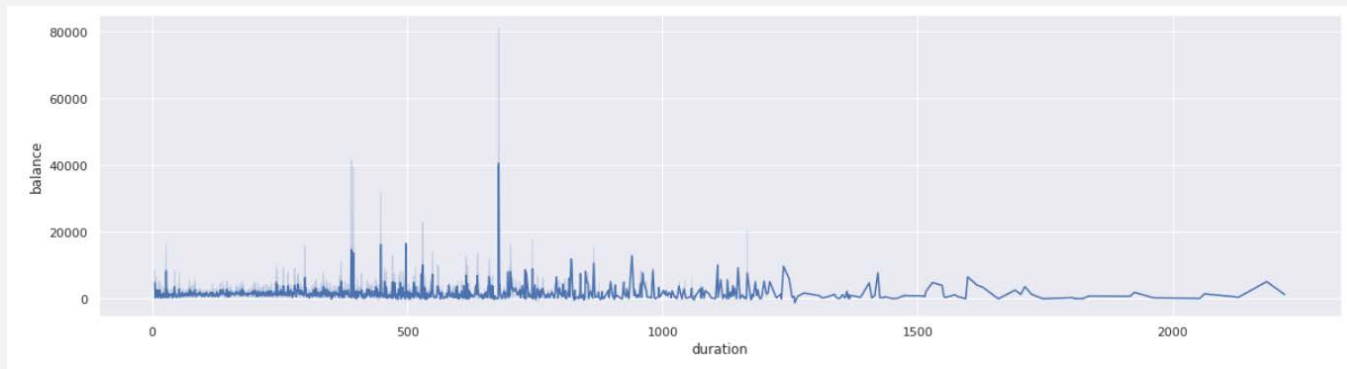


- The graph shows the relation between the account balance and the age of the client. This may not help in the final prediction but gives an insight of the data

ANALYXING THE DATA

“Exploratory Data Analysis”

The relation between the bank balance and
the client age



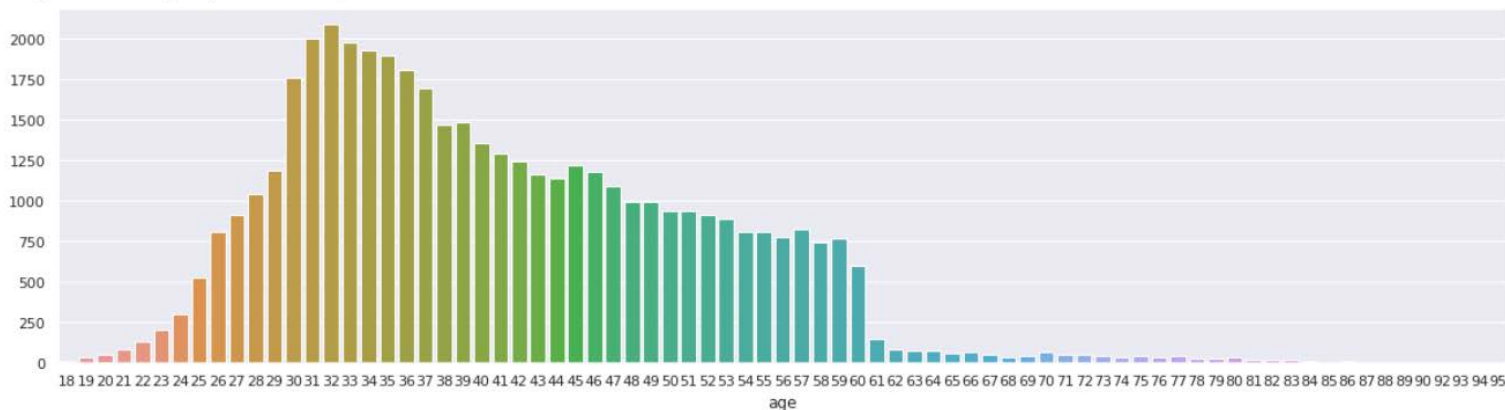
- The graph shows the relation between the account balance and the duration. This mayn't help in the final prediction but gives an insight of the data

ANALYXING THE DATA

“Exploratory Data Analysis”

The relation between the total number of applications and the client's age

<matplotlib.axes._subplots.AxesSubplot at 0x7f69d27dbc50>



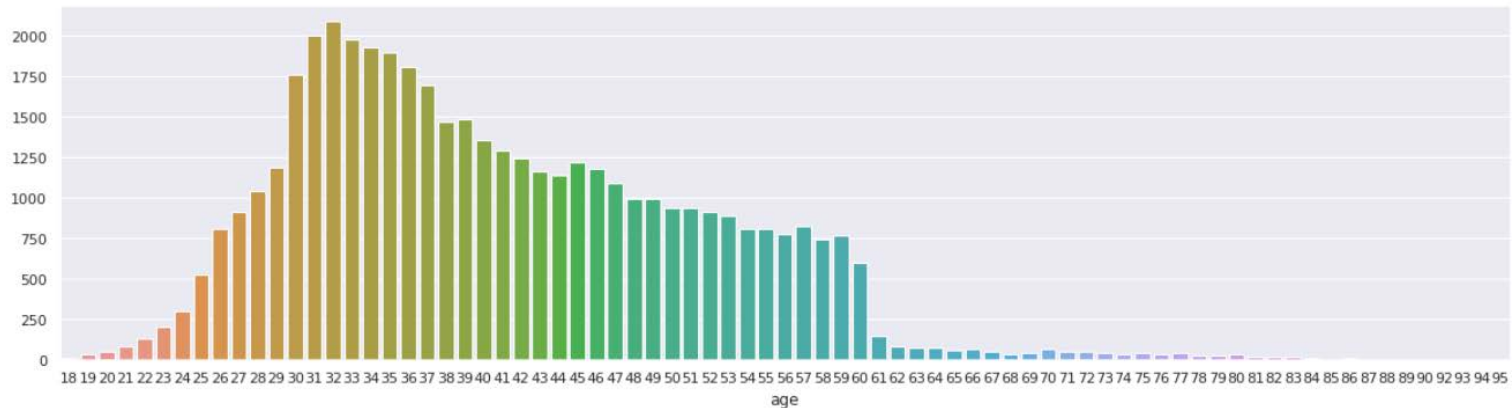
- The graph shows the relation between the total number of applications the bank received and the age of the applicant. It can be seen that most applicants age varies from 25-60 years old.

ANALYZING THE DATA

“Exploratory Data Analysis”

The relation between the accepted applications and the client's age

<matplotlib.axes._subplots.AxesSubplot at 0x7f69d27dbc50>

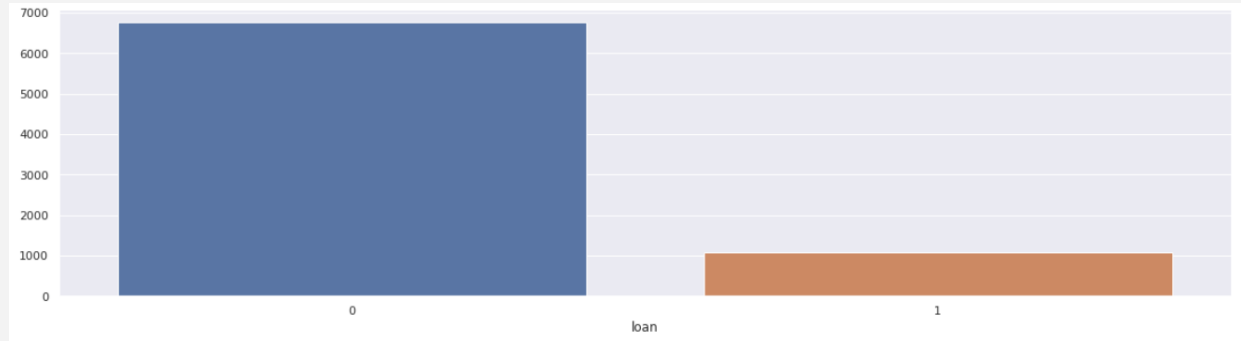


- The graph shows the relation between the number of accepted applications and the age of the applicant. The graph is similar to the previous graph meaning that there is no dependency of accepting the application and the applicant age.

ANALYZING THE DATA

“Exploratory Data Analysis”

The relation between the accepted applications and if they had a load

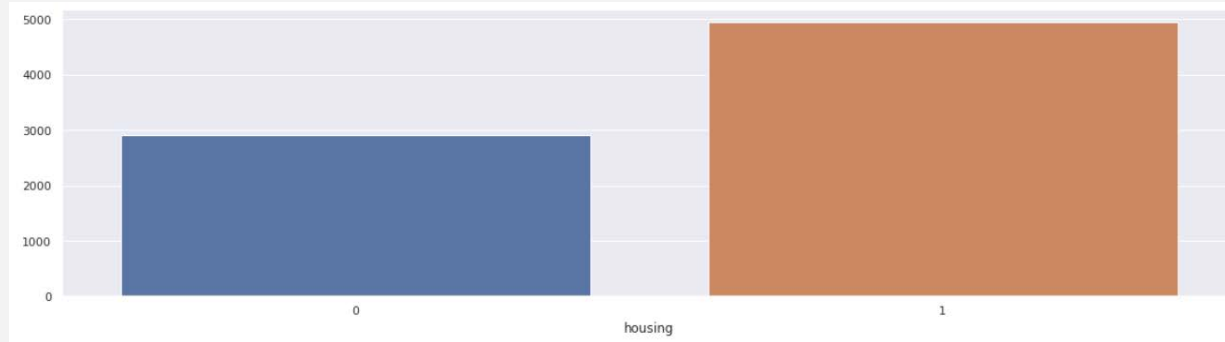


- It can be seen that most of the accepted applications didn't have a previous loan. Meaning that it can be taken into consideration in decision making

ANALYZING THE DATA

“Exploratory Data Analysis”

The relation between the accepted applications and if they have a house

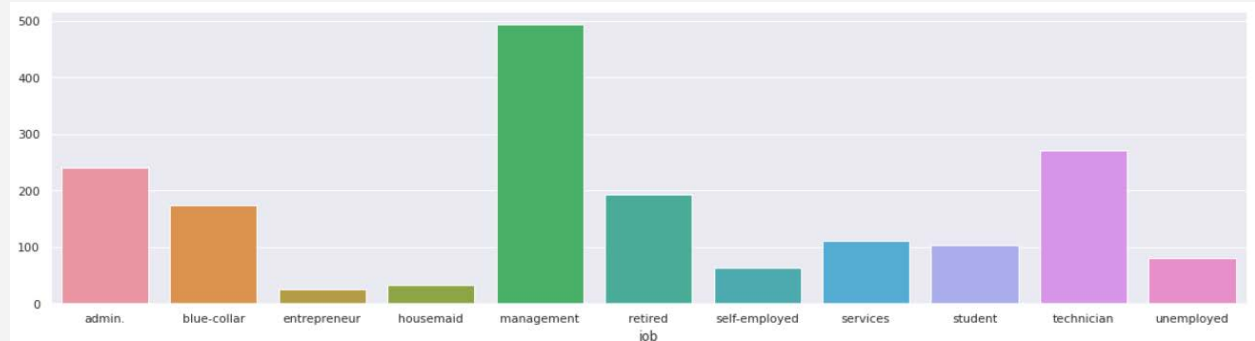


- It can be seen that most of the accepted applications have house. The relation is not that high still it should be taken in consideration in decision making

ANALYZING THE DATA

“Exploratory Data Analysis”

The relation between the accepted applications and the applicant's job

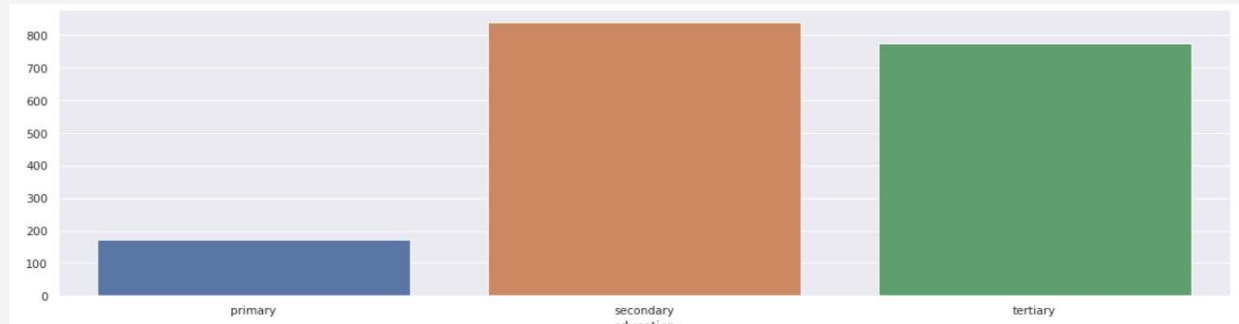


- It can be seen some jobs have higher probability of getting accepted. This means it should be taken into consideration.

ANALYZING THE DATA

“Exploratory Data Analysis”

The relation between the accepted applications and the applicant's education

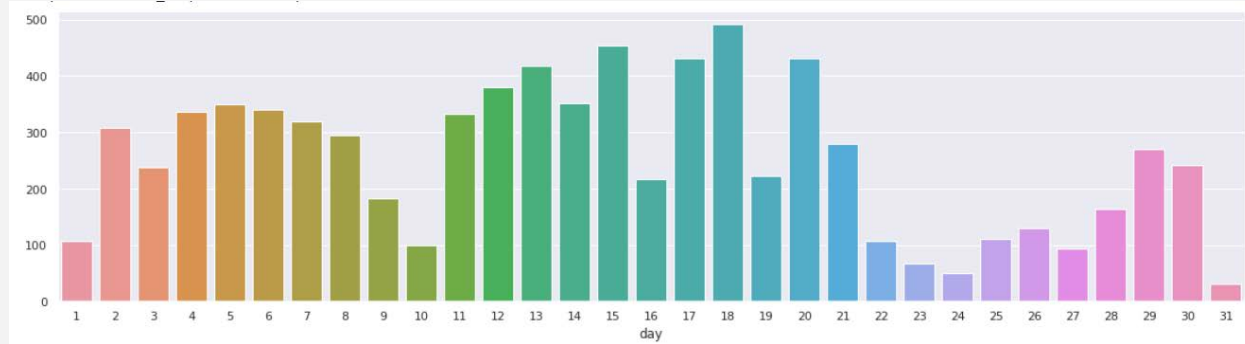


- It can be seen education influences the probability of getting accepted. This means it should be taken into consideration.

ANALYZING THE DATA

“Exploratory Data Analysis”

The relation between the accepted applications and the application day

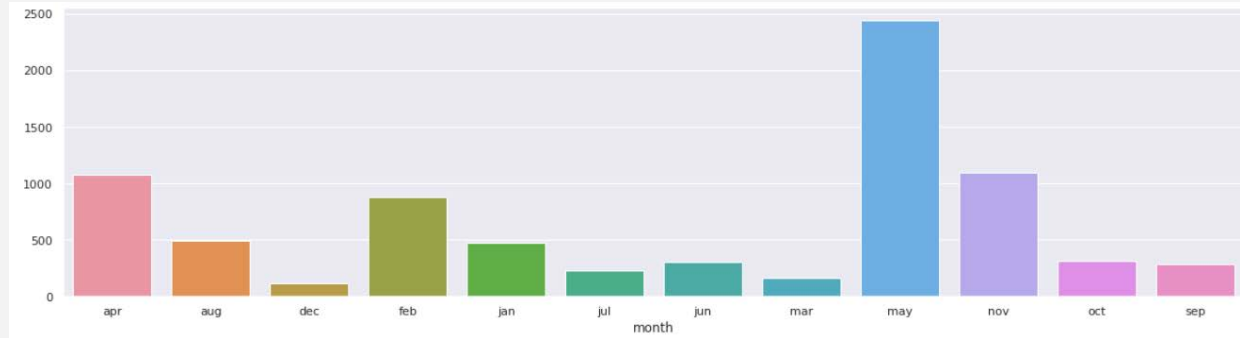


- It can be seen that the day of application does not influence the probability of getting accepted. This means it should not be taken into consideration.

ANALYZING THE DATA

“Exploratory Data Analysis”

The relation between the accepted applications and the application month



- It can be seen that the month of application does not influence the probability of getting accepted. This means it should not be taken into consideration.



Final Recommendations

1. Outliers are not to be removed.
2. There is a direct relation between the loan and the acceptance of the application (Important feature for classification).
3. Some jobs have higher priority for acceptance (important feature for classification).
4. Some features do not have effect on the output as day and month (Not to consider).



Data Glacier

Your Deep Learning Partner

TABLE OF CONTENTS

01

Introduction

02

EDA & Recommendations

03

Model Recommendations

04

Evaluation
Recommendations

05

GIT Repo.

Data Glacier

Model Recommendation

- The first model will be **Naïve Base** because:
 - It is very fast at classifying unknown records.
 - It is simple and easy to implement.
 - It doesn't require as much training data.
 - It handles both continuous and discrete data.
 - It is highly scalable with the number of predictors and data points.
- The second model will be **Linear regression** because:
 - It is easier to implement, interpret, and very efficient to train.
 - It can interpret model coefficients as indicators of feature importance.
 - It can interpret model coefficients as indicators of feature importance.



Data Glacier

Your Deep Learning Partner

Model Recommendation

- The Third model will be **Ensemble Models (logistic regression, xgboost, and random forest) using max voting** because:
 - Reduce the spread in the average skill of a predictive model.
 - Improve the average prediction performance over any contributing member in the ensemble.
 - Has better performance than singles model.
- The Fourth model will be **Ensemble Models (Bagging)** because:
 - Allows many weak learners to combine efforts to outdo a single strong learner
 - Higher predictive accuracy, compared to the individual models



Data Glacier

Your Deep Learning Partner

Model Recommendation

- The Third model will be **Ensemble Models (logistic regression, xgboost, and random forest) using max voting** because:
 - Reduce the spread in the average skill of a predictive model.
 - Improve the average prediction performance over any contributing member in the ensemble.
 - Has better performance than singles model.
- The Fourth model will be **Ensemble Models (Bagging)** because:
 - Allows many weak learners to combine efforts to outdo a single strong learner
 - Higher predictive accuracy, compared to the individual models



Data Glacier

Your Deep Learning Partner

Model Recommendation

- The Fifth model will be **Stacking classifier (RF, KNN - DT, RF)** because:
 - Stacking can yield improvements in model performance.
 - Stacking reduces variance and creates a more robust model by combining the predictions of multiple models.



Data Glacier

Your Deep Learning Partner

TABLE OF CONTENTS

01

Introduction

02

EDA & Recommendations

03

Model Recommendations

04

Evaluation
Recommendations

05

GIT Repo.

Data Glacier

MODEL EVALUATION

The models can be evaluated using different methods, here three evaluation methods shall be used.

1. **Accuracy:** The number of classifications a model correctly predicts divided by the total number of predictions made.
2. **Confusion Matrix:** is a summary of prediction results on a classification problem.
3. **F1 Score:** The F-score is a way of combining the precision and recall of the model.

TABLE OF CONTENTS

01

Introduction

02

EDA & Recommendations

03

Model Recommendations

04

Evaluation
Recommendations

05

GIT Repo.

Data Glacier

GitHub Repository link

<https://github.com/alymedhat10/Bank-Marketing-Campaign-.git>



Data Glacier

Your Deep Learning Partner

The background features several orange geometric lines. One line starts from the top left, goes down and to the right, then up and to the right. Another line starts from the top right, goes down and to the left, then down and to the right. A third line starts from the middle right, goes down and to the left, then down and to the right. These lines create a modern, abstract design.

THANK YOU

Data Glacier