Data Glacier

Your Deep Learning Partner

Bank Marketing (campaign)

Final Report

Name	Email	Country	Specialization
ILYAS NAYLE	Ilyasnayle5@gmail.com	Turkey	**Data Science**

Virtual Internship LISUM03

15-Nov-2021

Table of Content

A. Abstract

B. Problem Description

C. Data Description

D. EDA

E. EDA Summary

F. Anaylsis Of Model Performance

G. Feature Importance

H. Final Recommendation

Abstract

A. ABC Bank wants to sell its deposit term product to the customers

B. By evaluating and implying the EDA and afterwards analyzing it , we tried to understand the data and its pattern on the opt policy.

C. Different attributes were considered for analysis

D. The bank data has information of 45221 customers with 17 columns

E. First we focused on the customers who bought the Policy

Problem Description

I. Business Understanding, Data understanding, Exploratory data Analysis, Data Preparation

II. Identifying age group, marital status, educational level, and job description of the clients who purchased the policy.

III. Is contacting clients before of after the campaign beneficial ?

IV. What is the time taken for the communication of those clients who purchased the policy?

Data Description

- Data set Characteristics: Multivariate
- Number of Instance: 45221
- Date of data: 2012-02-14
- Attribute Characteristics: Real
- Null Values: No

EDA (EXPLORATORY DATA ANALYSIS)

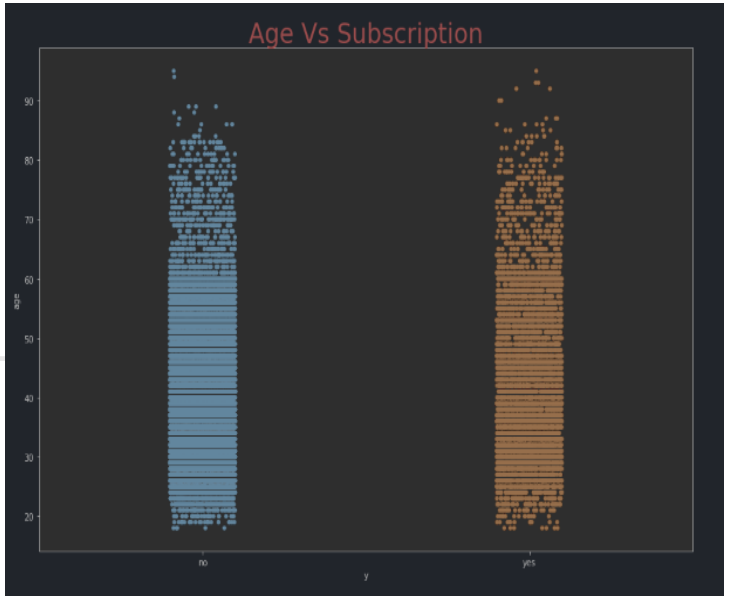
	age	balance	day	duration	campaign	pdays	previous
count	45211.000000	45211.000000	45211.000000	45211.000000	45211.000000	45211.000000	45211.000000
mean	40.936210	1362.272058	15.806419	258.163080	2.763841	40.197828	0.580323
std	10.618762	3044.765829	8.322476	257.527812	3.098021	100.128746	2.303441
min	18.000000	-8019.000000	1.000000	0.000000	1.000000	-1.000000	0.000000
25%	33.000000	72.000000	8.000000	103.000000	1.000000	-1.000000	0.000000
50%	39.000000	448.000000	16.000000	180.000000	2.000000	-1.000000	0.000000
75%	48.000000	1428.000000	21.000000	319.000000	3.000000	-1.000000	0.000000
max	95.000000	102127.000000	31.000000	4918.000000	63.000000	871.000000	275.000000

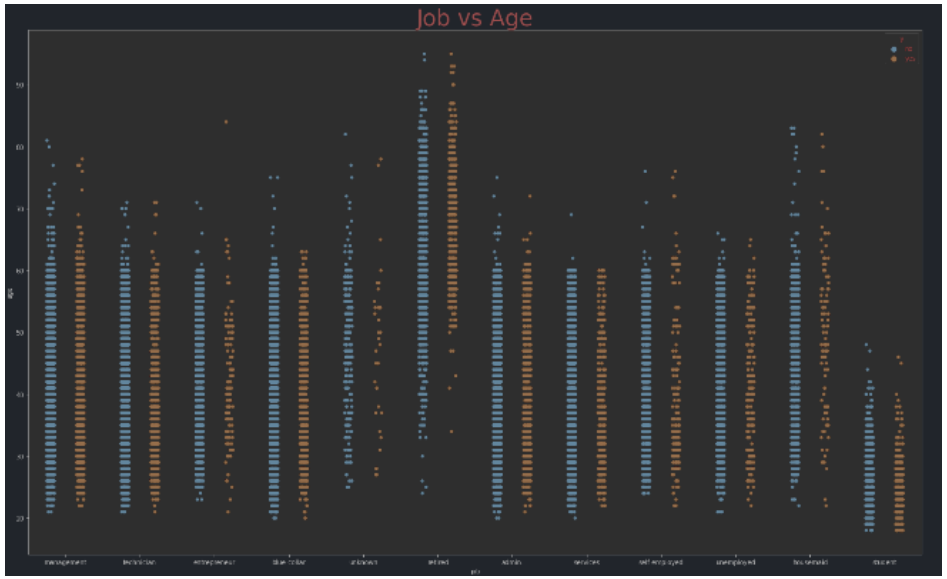
8 rows × 7 columns

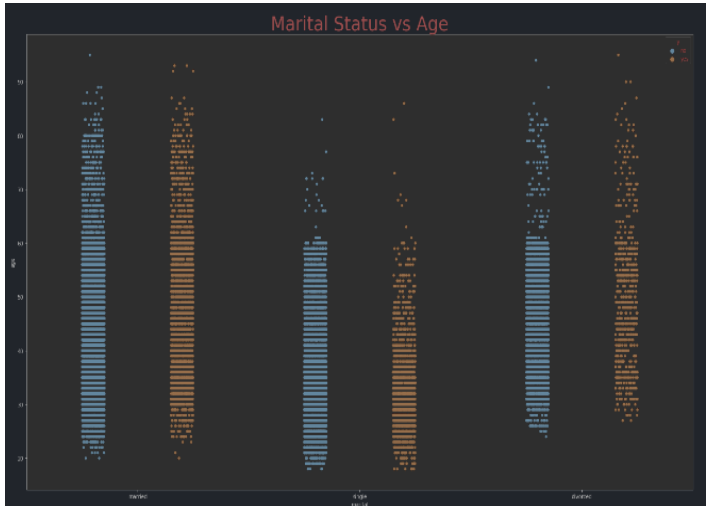
- Understanding the Data
- Knowing the purchase policy
- Anlayzing the Duration
- Finalizing the Recommendation

Customer Vs Age

- Those Who bought the policy are more likely to be between 20-60 and on the other hand those who are above the 70 have already bought the policy.

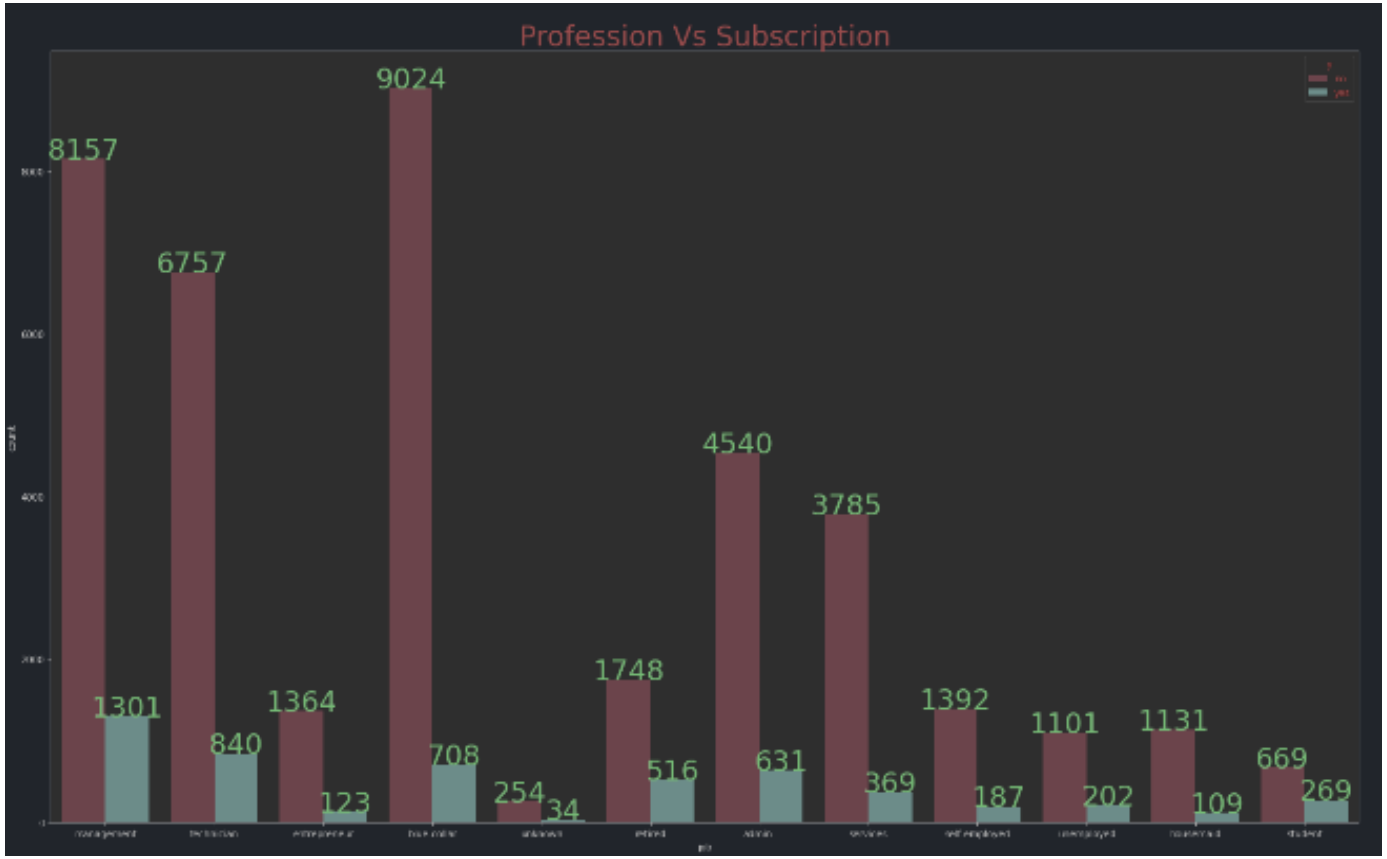






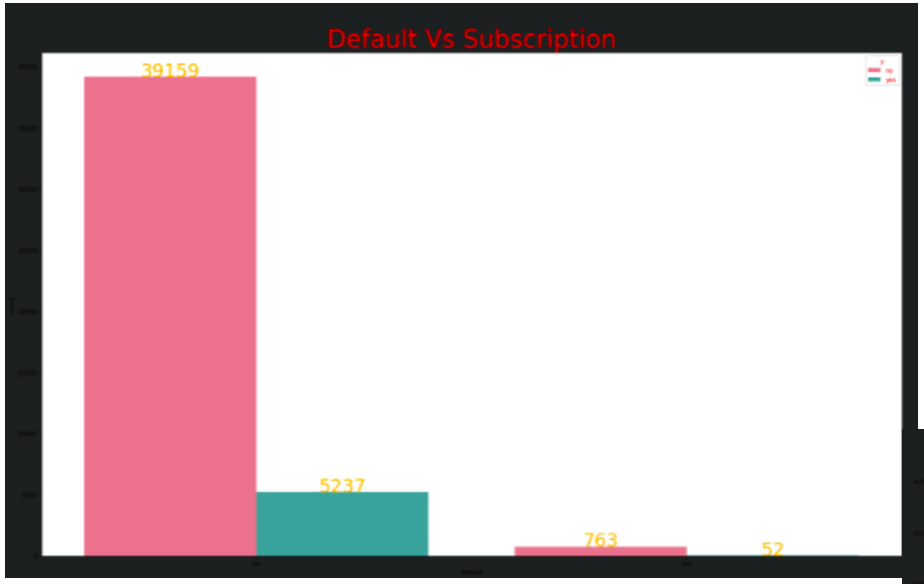
Job type vs Policy Purchase

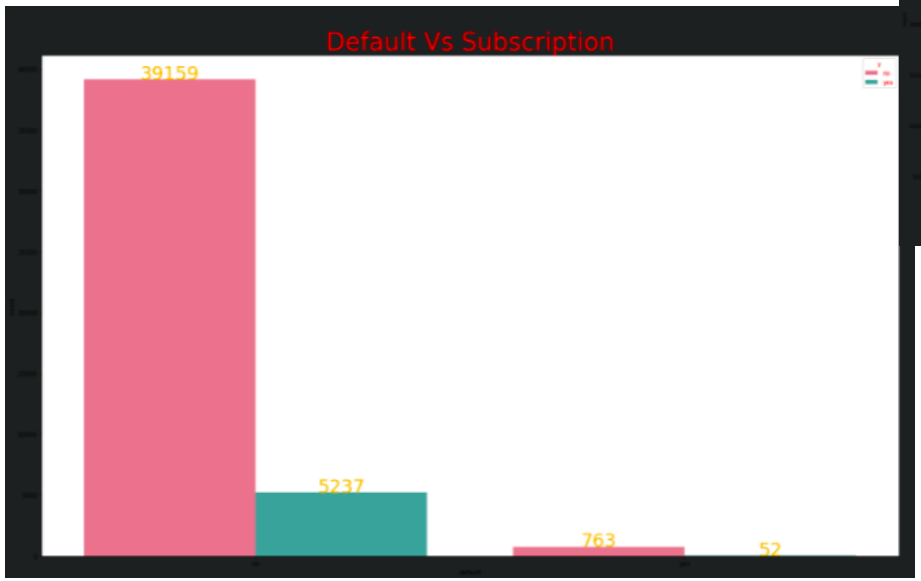
- Clients who are retired have the highest customer purchase rate

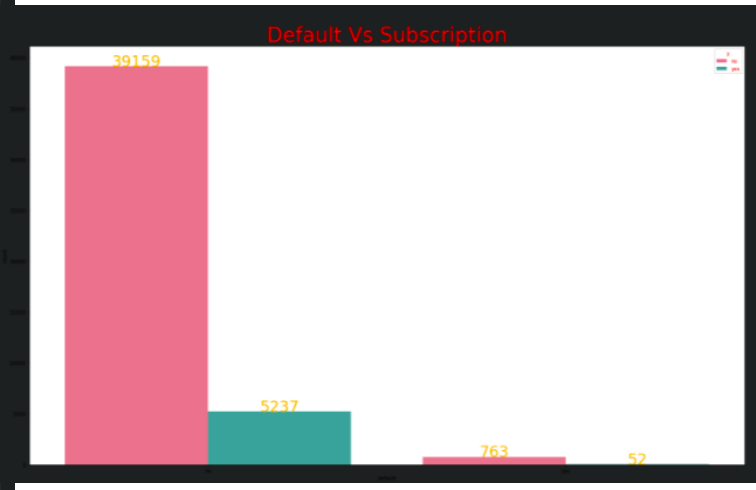


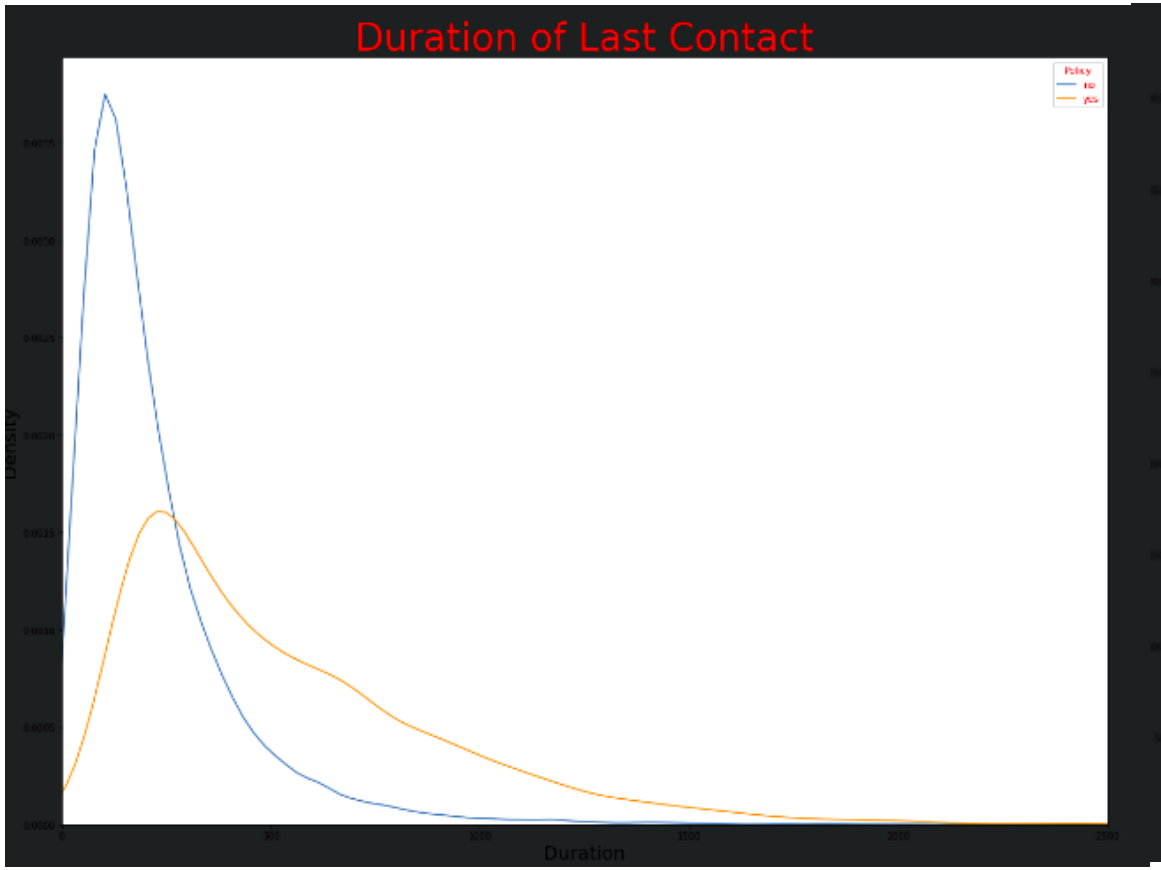
Clients Financial status

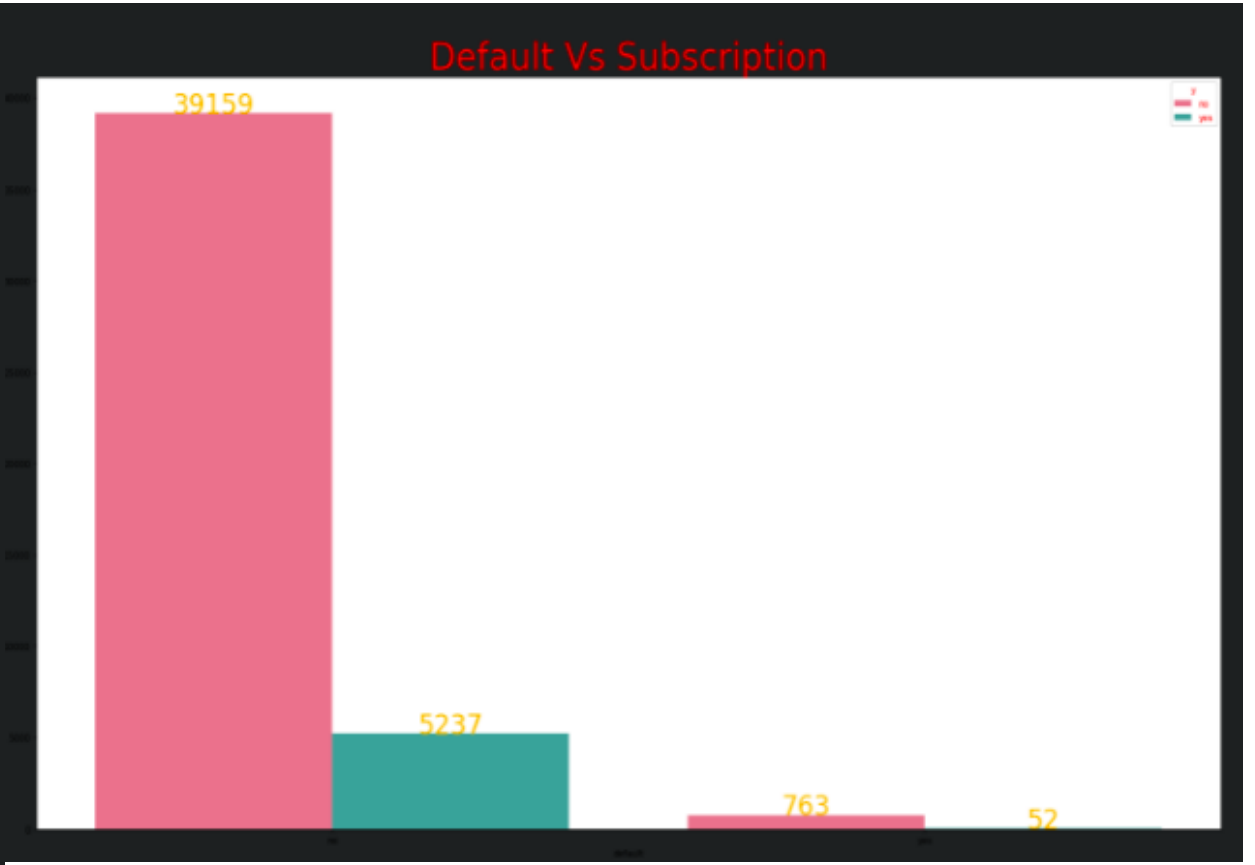
- Clients who have 'no' in their default status are more likely to buy the policy and also housing status those who have and don't have housing loan still apply for the policy





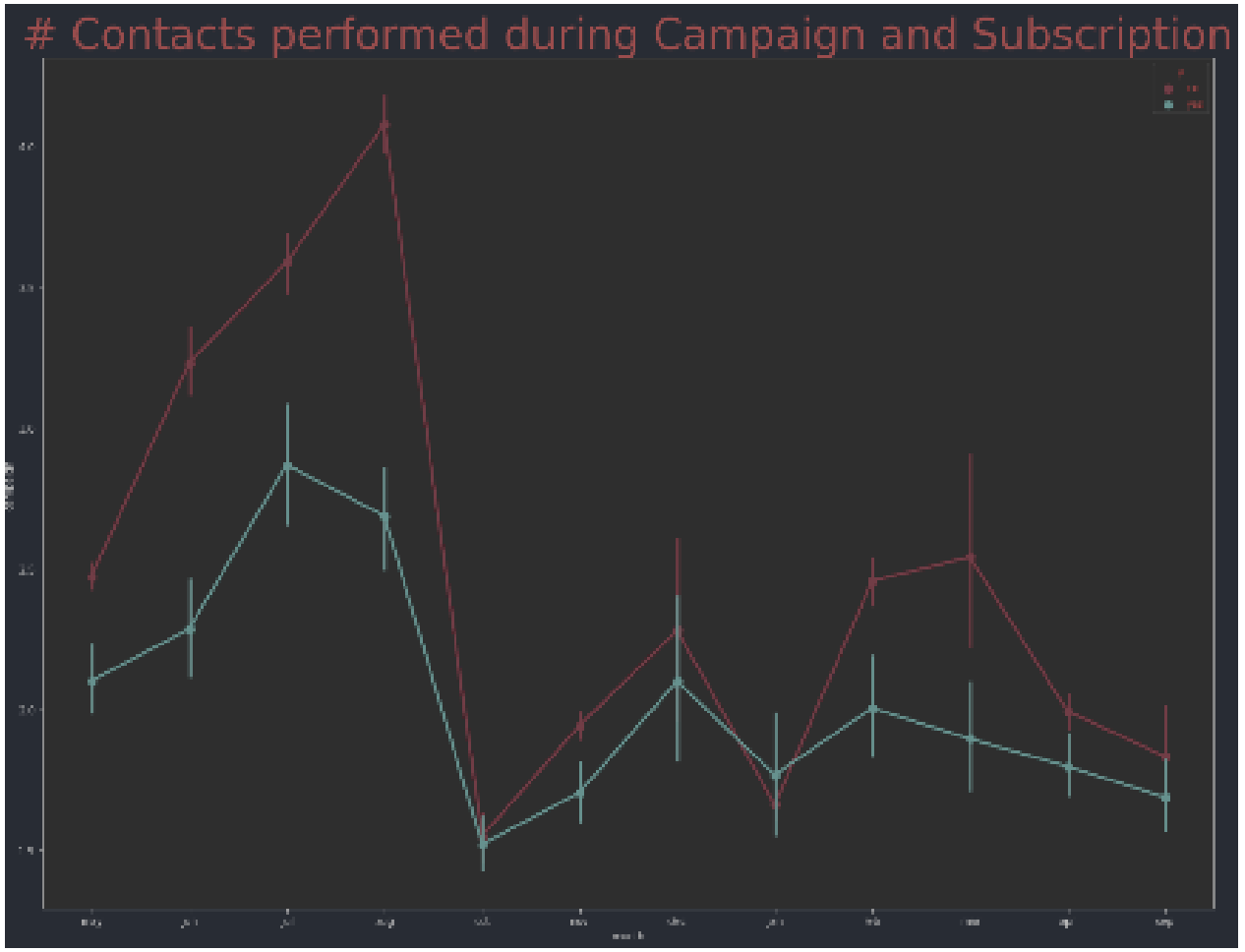


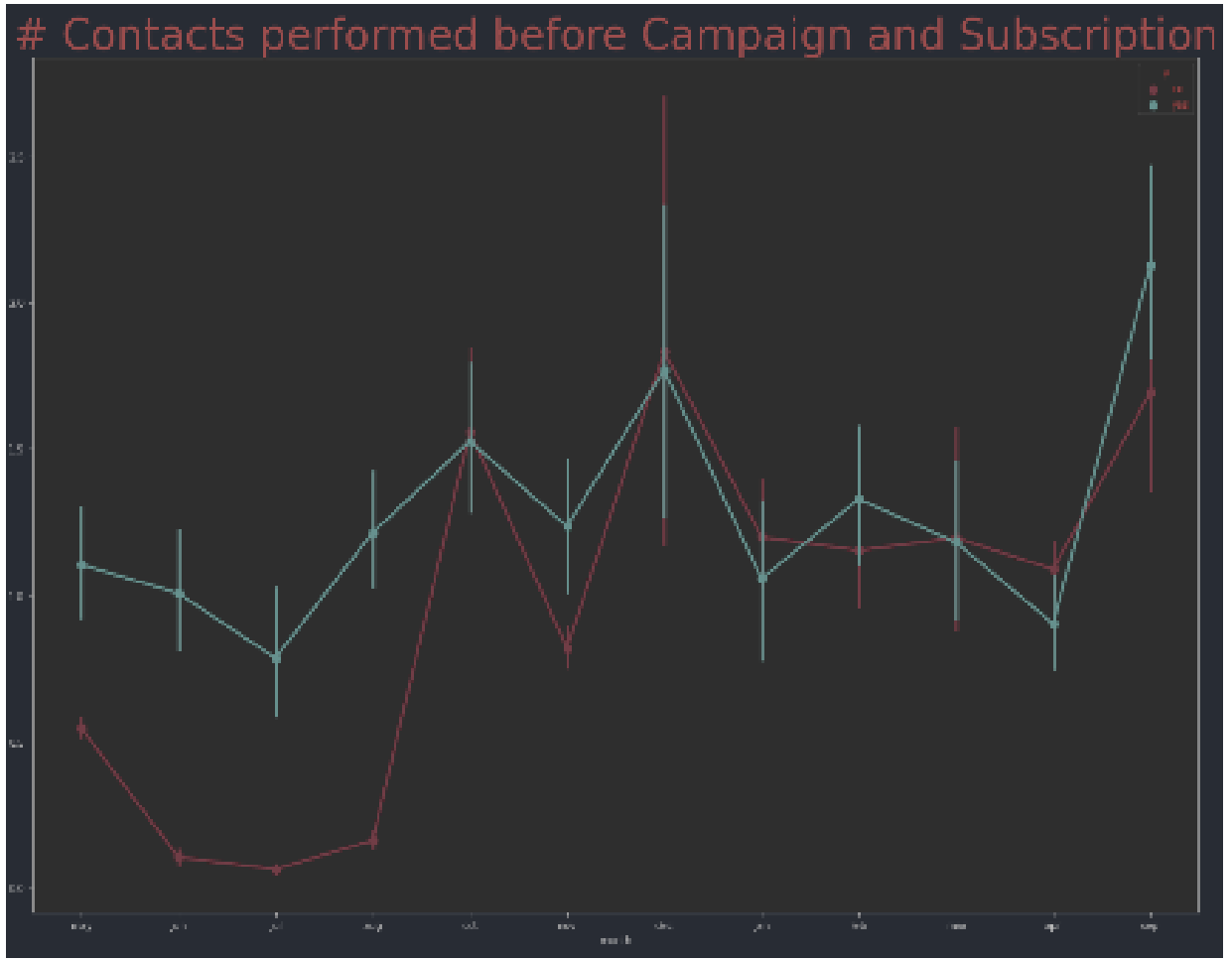




Customer Vs Contact

- On average the call lasted 426.00 seconds for those who opted the policy and those who did not the call lasted 164.00 seconds. And they spend more time on the communication.





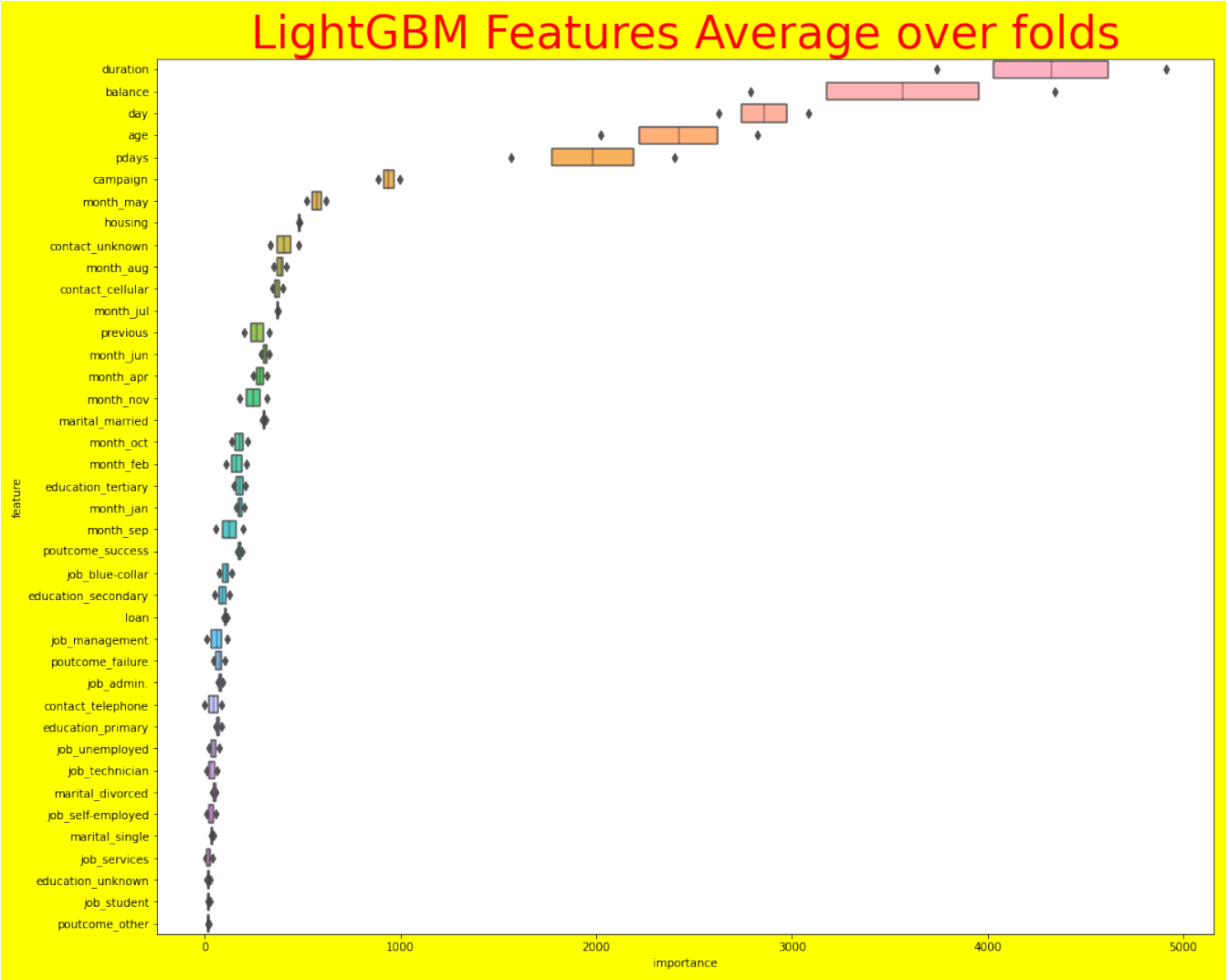
Contacts before campaign for the policy brings more clients during the may – oct, but however positive subscription to the policy subscription during the campaign doesn't outnumber the failure to subscription And the month of December seems to be the favorable month.

Analysis of the Model Performance

MODEL	Accuracy Score	Precision		Recall		F1- Score	
		Don't buy the policy	Customer who buy the policy	Don't buy the policy	Customer who buy the policy	Don't buy the policy	Customer who buy the policy
Logistic Regresssion	0.8508	0.97	0.43	0.86	0.79	0.91	0.56
Random Forest Calssifier	0.8972	0.94	0.58	0.95	0.54	0.94	0.56
LightGM	0.9054	0.93	0.63	0.96	0.47	0.95	0.54

As we can observe from the Accuracy score the best performance has the LIGHTGM model. The Precision And F1-Score is good in LightGM and therefore the recommended Model is LIGHTGM

Feature Importance



Final Recommendations

The following are the final recommendation that is finalized after implying
the EDA on the Bank data set file;

Note: The word " Those" refers to clients.

- Those Retired are more likely to buy policy than others
- Those who are between age 20-50 are more likely those who bought the policy and there are those above 70 also have opted the policy.
- Those Married are more likely to buy between the age of 60 – 80.
- Those without "no" as their option in default column are more likely to buy the policy.
- Those who are in management section are more likely to get the policy
- Those in the Secondary are more than others in term of subscription.

Final Recommendations

- Those who will most likely opt for policy are between the month of May – October as well as September. Moreover, the contacts before the campaign brings more clients.
- Those with the cell-phones have higher rate of subscription than those with none and telephone.
- There is an average of 426 seconds for those who opt the policy and for those who did not where the call lasted for 146 seconds which is fewer than those who opt for the policy. Although they spend more time on the communication when they opt for the policy.
- Those with “no” in their default and more likely between the age of 20- 60 years old.
- The Recommendation is to use the LightGM model for the prediction of who will buy the policy

Thank You