FINAL REPORT

Bank Marketing Campaign

14-August-2021

Agenda

Executive Summary

Problem Statement

EDA

EDA Summary

Model Performance Analysis

Final Recommendations

Executive Summary

- ABC Bank wants to sell it's term deposit product to customers.
- By analyzing the clients data we try to understand what are the patterns on opting for the policy.
- 17 different attributes where considered to perform the analysis.
- The data had information of 45211 customers who purchased and who did not purchase for the policy.
- Analysis was primarily focused on the customers who purchased the policy.

Problem Statement(Research)

- Identify the age group, marital status, education level, job description of customers who purchase the policy.
- What is the purchase pattern for the policy throughout the year.?
- Does contacting the customer before or after the campaign beneficial for the company?
- What is average time taken for communication by a client who purchases the policy?

Approach

1. Understanding the Data.
2. Purchase of policy over time.
3. Duration Analysis
4. Recommendation

Assumption: The data sample that which is given is a random sample and it truly represents the population.

Exploratory Data Analysis (EDA)

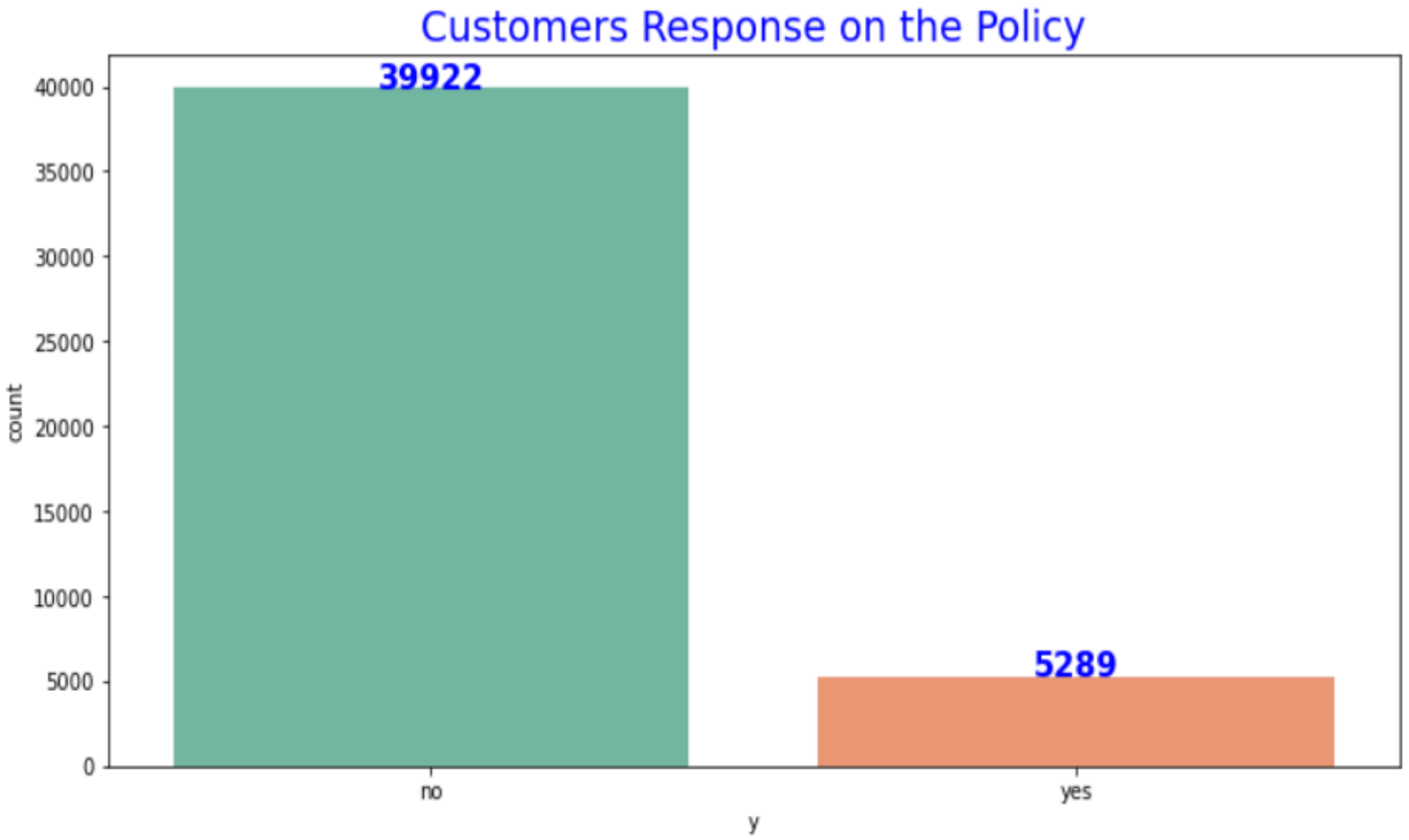
Summary of some Key Attributes

	Age	Duration
mean	40.936210	258.163080
std	10.618762	257.527812
min	18	0
50%	39	180
max	95	4918

Key Insights

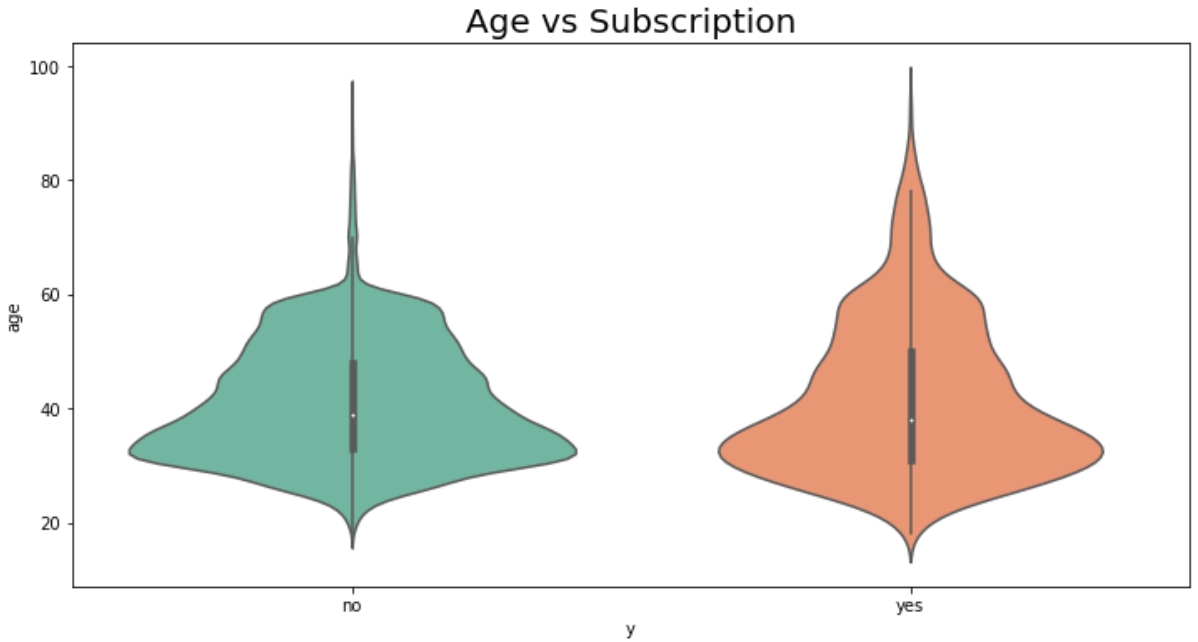
1. On average the age of the clients is 40.
2. The age group selected for this campaign is from 18-95.
3. The maximum communication time spent with a customer is 4918 seconds.

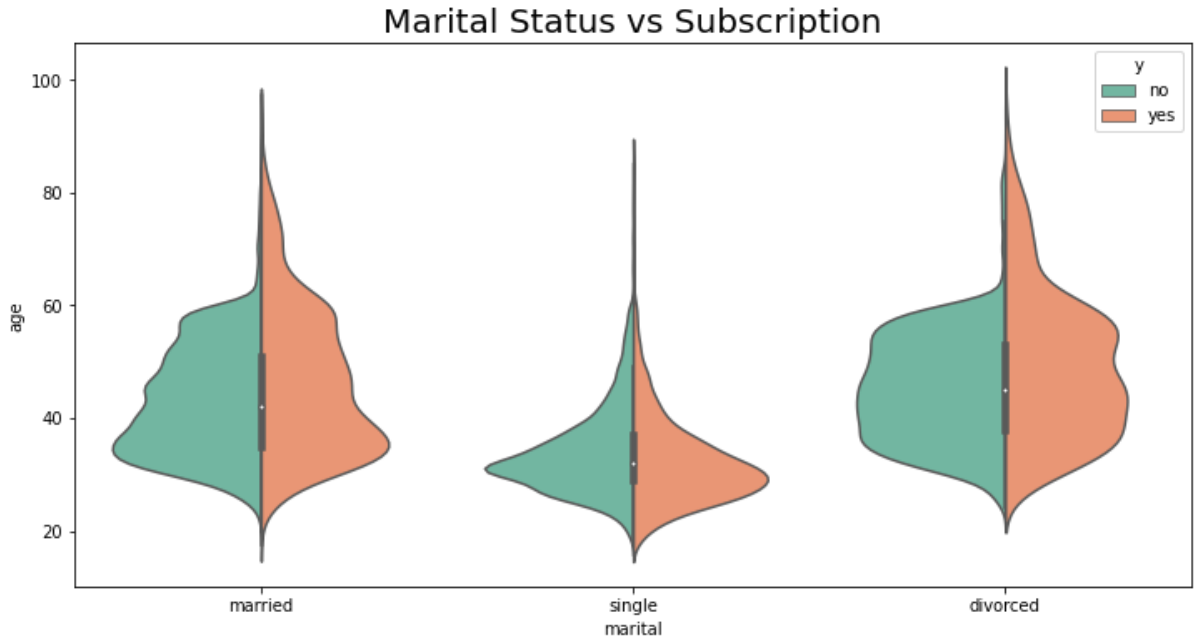
Customer Analysis

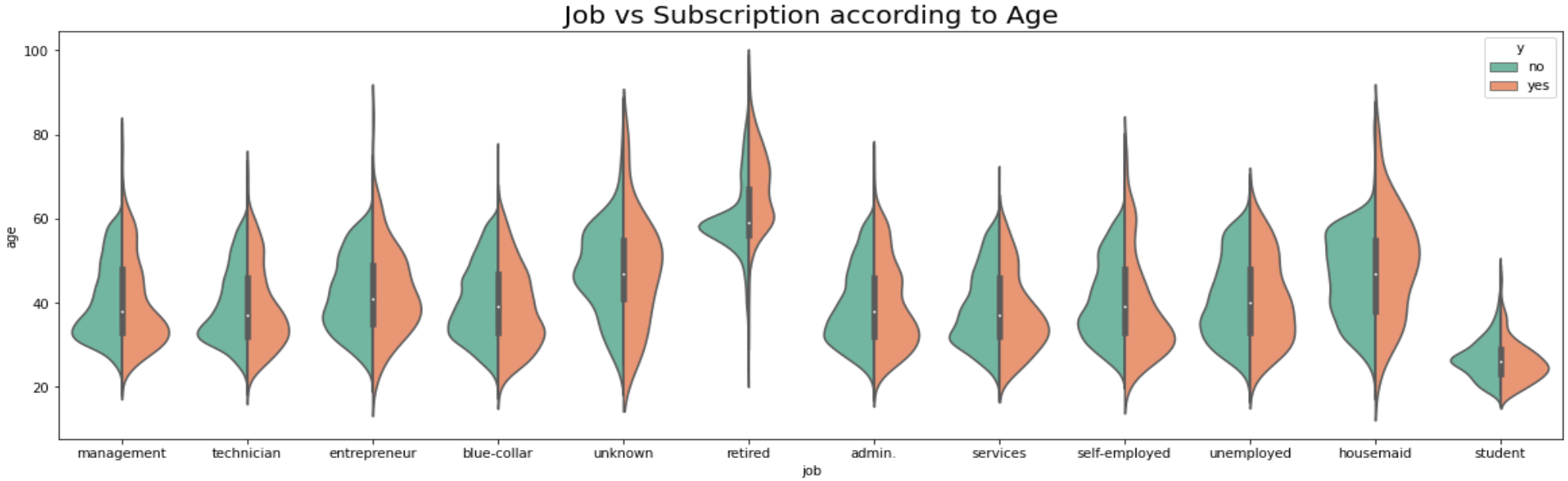


There are totally 5,289 customers who have purchased the policy from the span of January to December. Also 39,922 customers have rejected the policy.

Customer Age Analysis

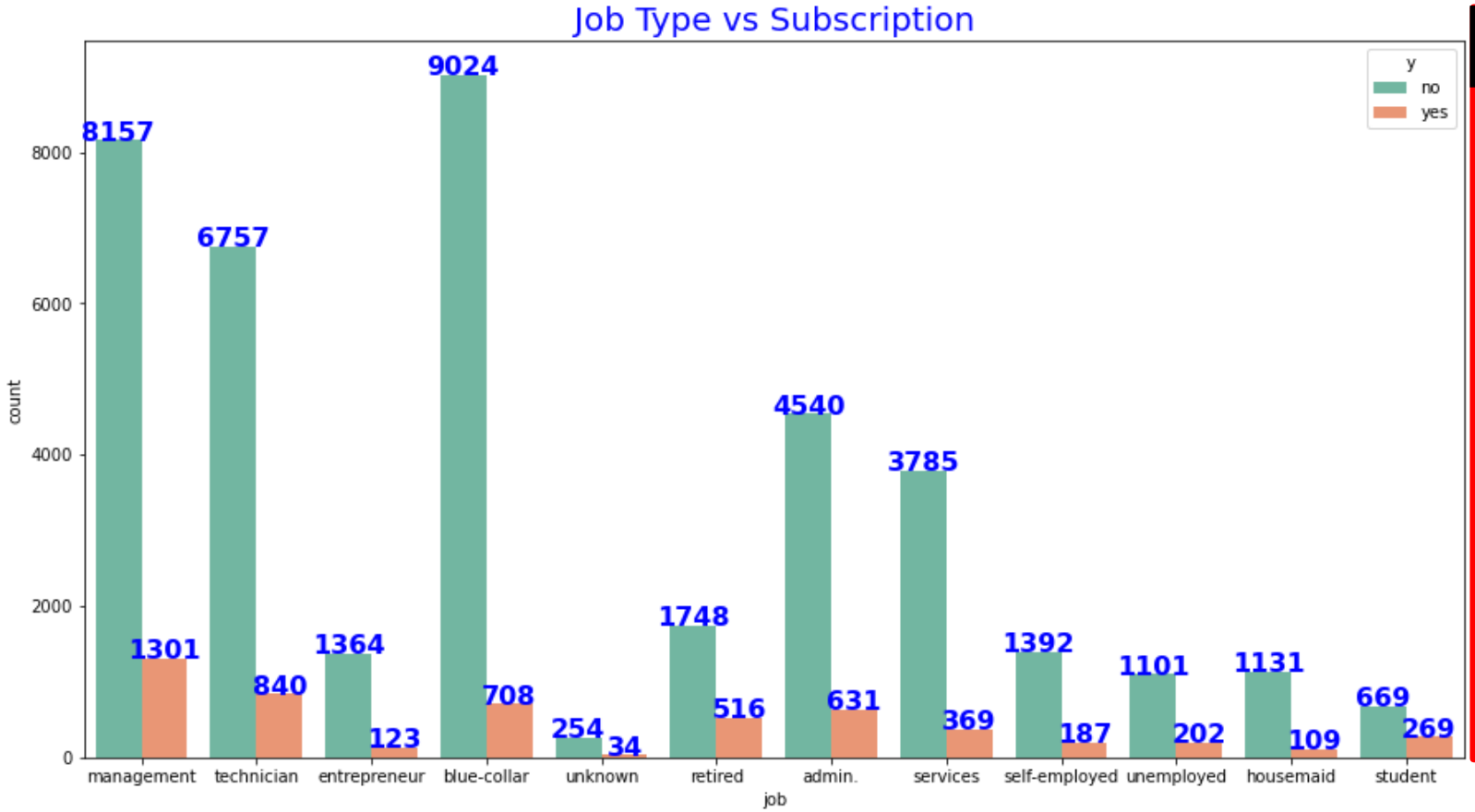






Clients who buy the policy are more likely to fall between the age of 20-40. Also there are more number of clients beyond age of 60 who have purchased the policy.

Policy Purchase with Job Type

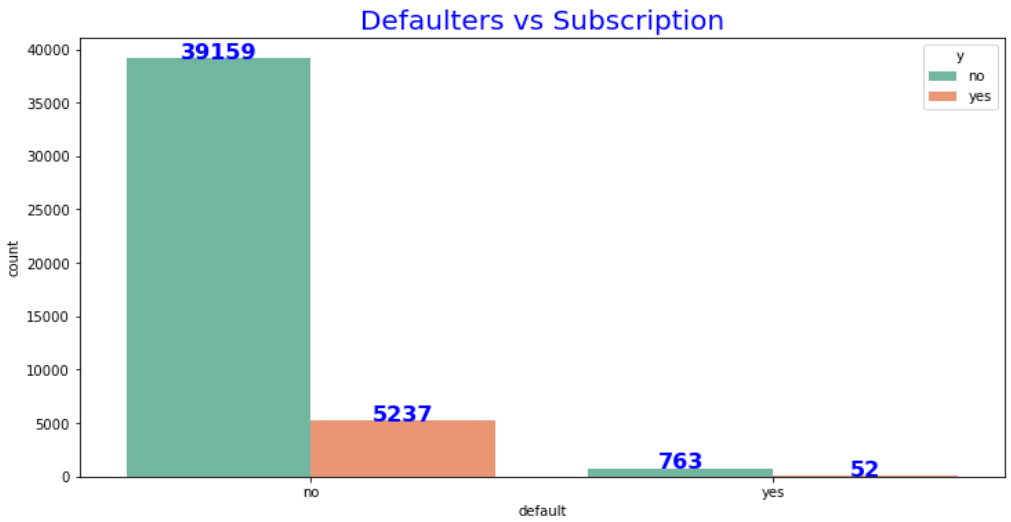


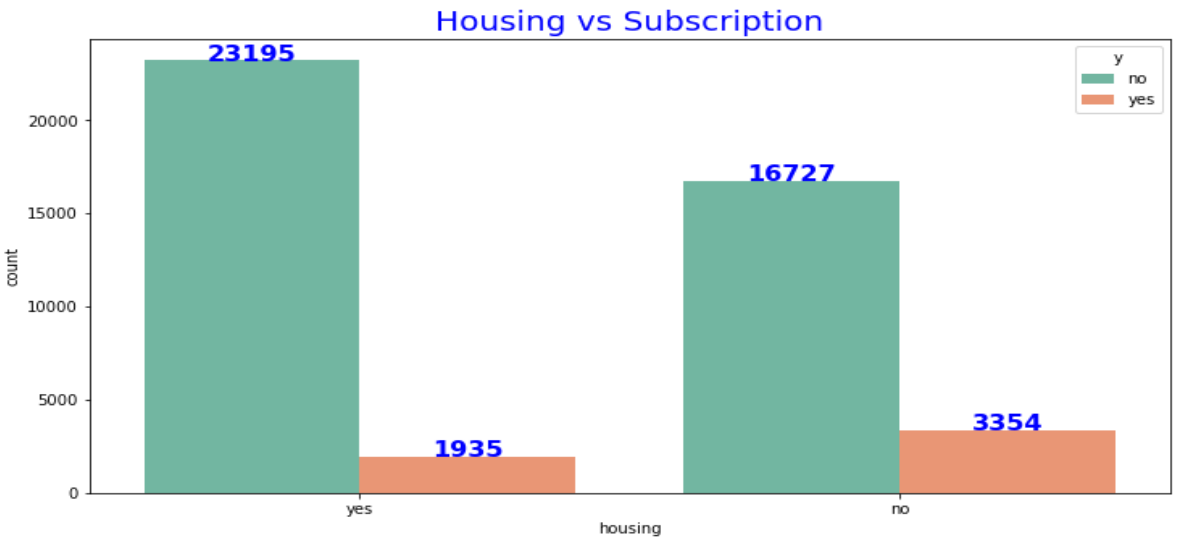
Purchase Percentage

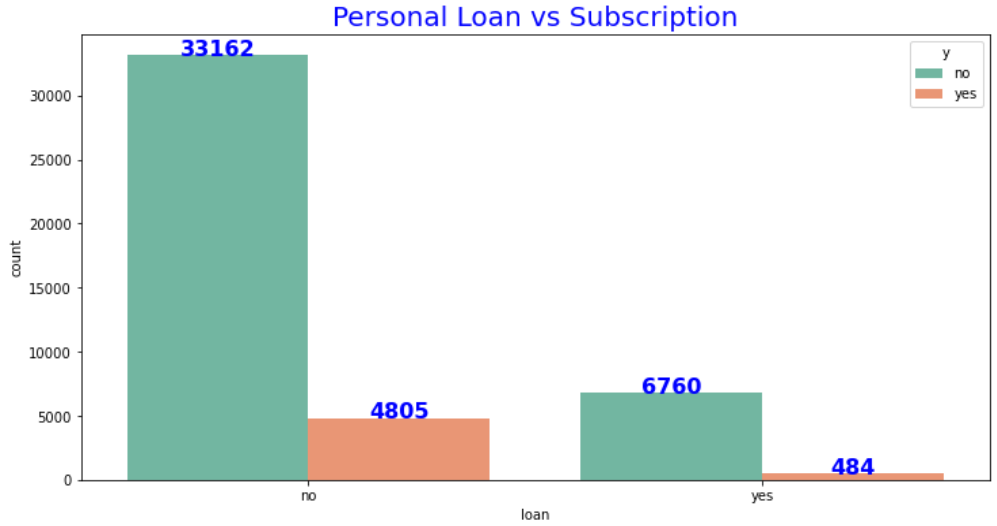
1. Management = 13%
2. Technician = 11%
3. Entrepreneur = 8%
4. Blue Collar = 7%
5. Unknown = 11%
6. Retired = 23%
7. Admin = 12%
8. Services = 9%
9. Self Employed = 12%
10. Unemployed = 15%
11. Housemaid = 9%
12. Student = 29%

Students have the highest purchase percentage, though the reach to students is not very high. The retired clients have the second highest customer purchase rate.

Financial Status of Clients

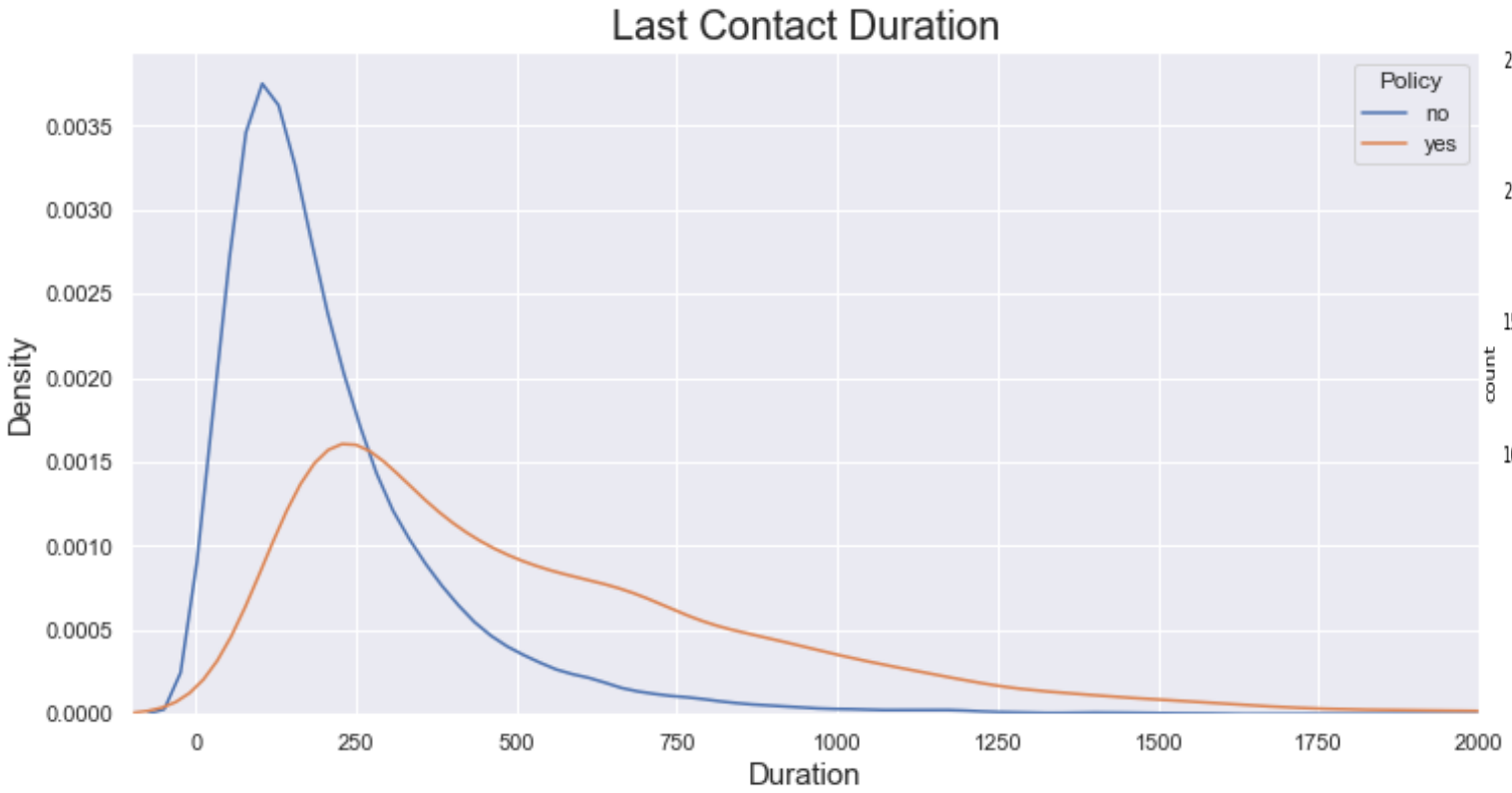


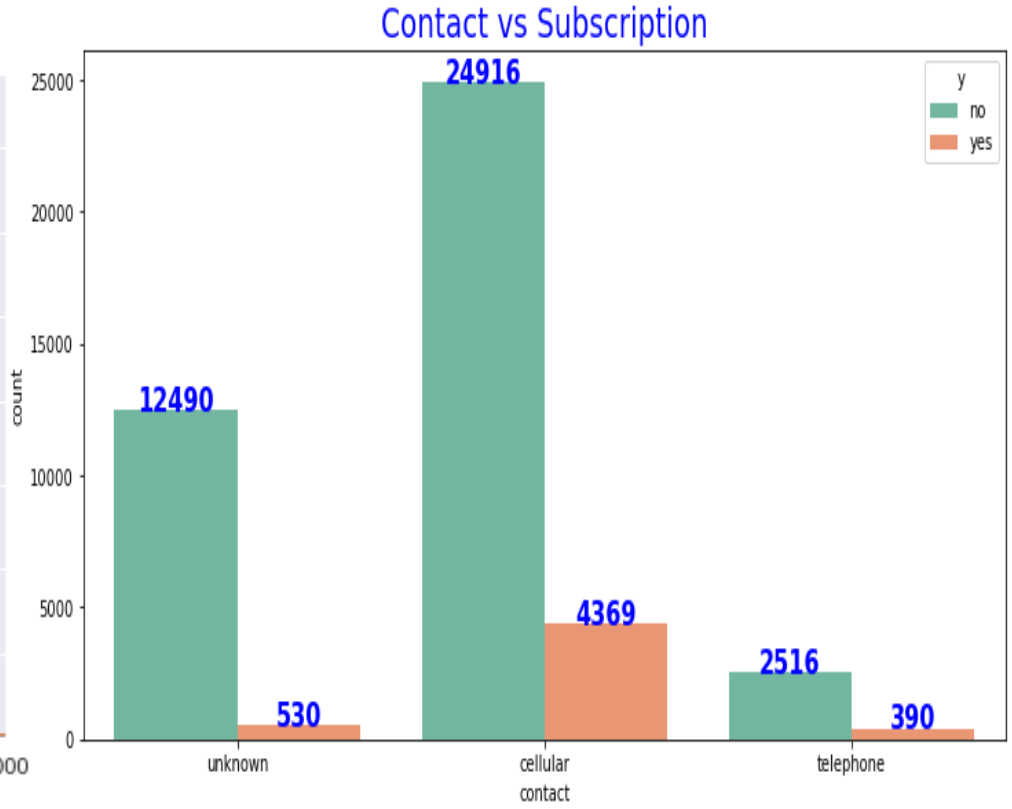




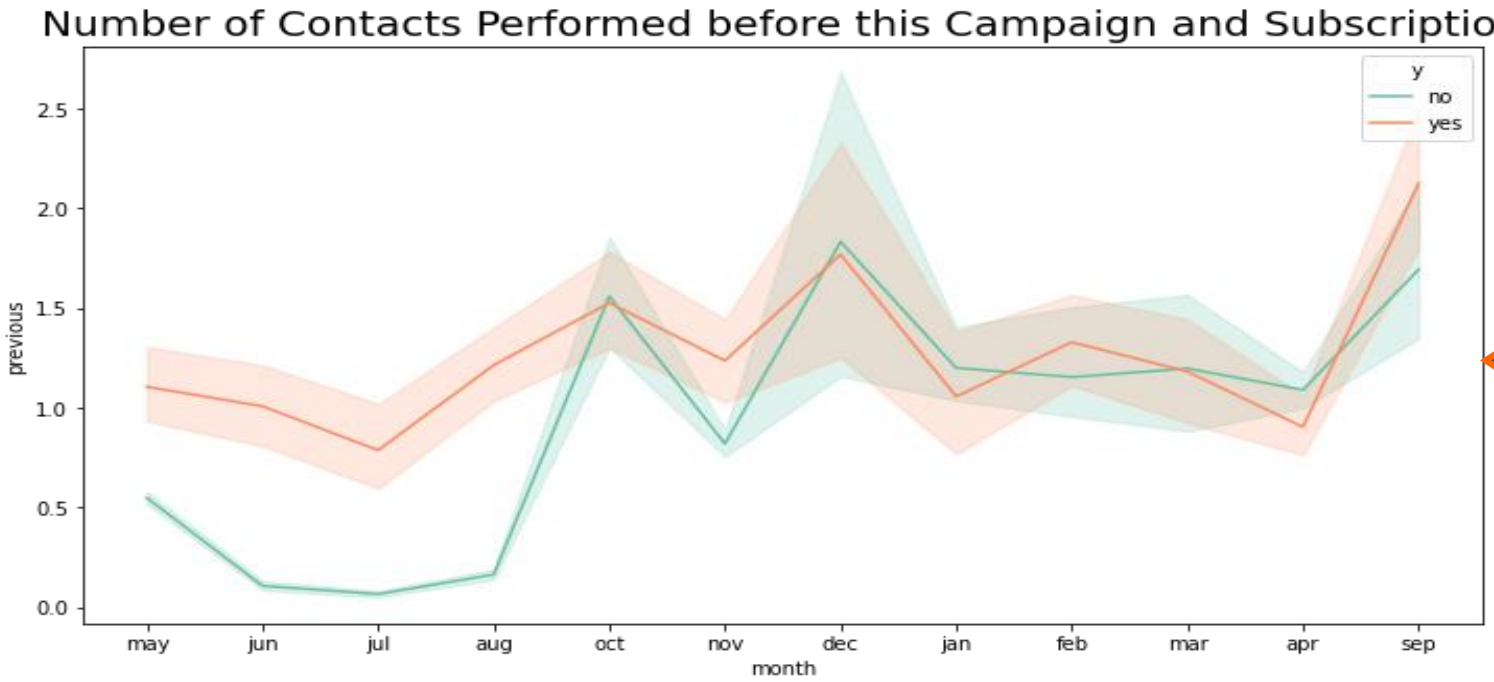
1. Irrespective of clients housing loan status there is a small difference(26%) between clients who have and not have housing loan and yet apply for the policy.
2. Clients who do not have default as their status are more likely to buy the policy.

Contact with Customer Analysis



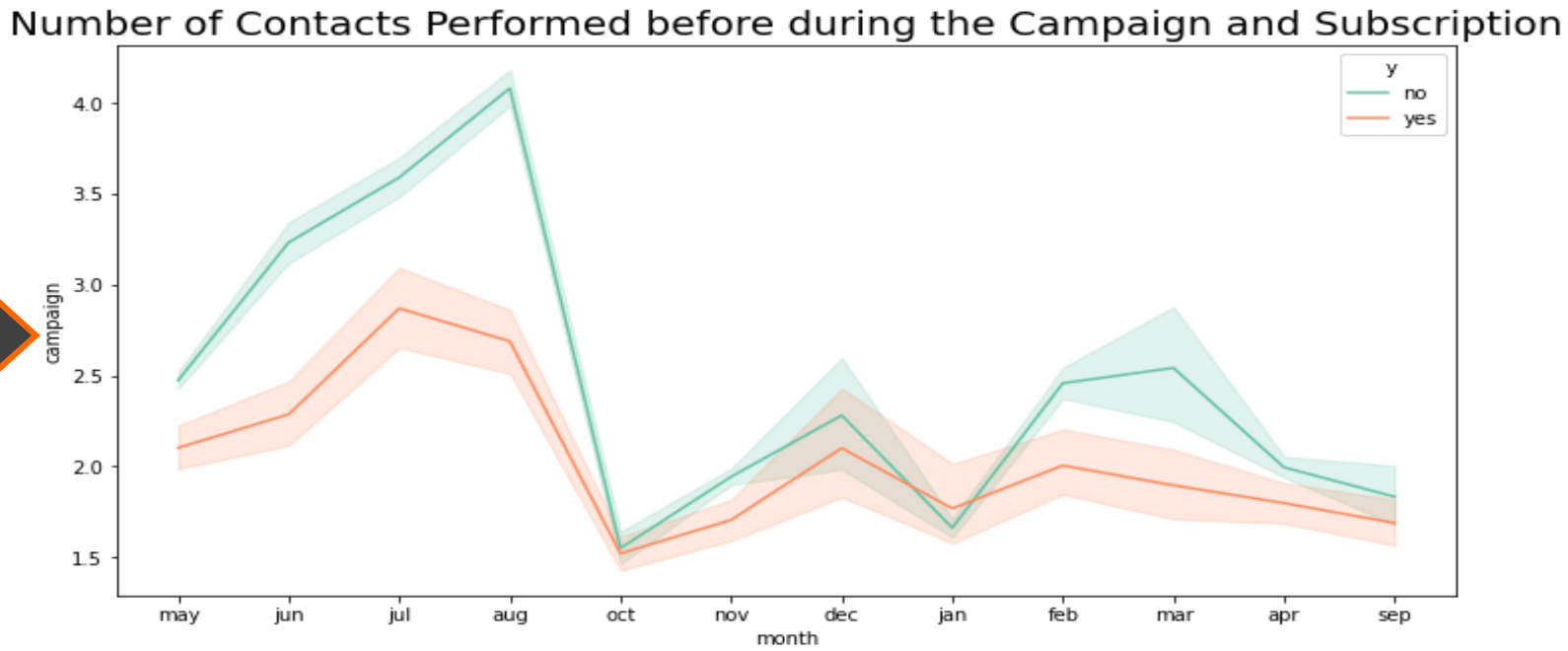


On average the call lasts for 426 seconds for the clients who opted for the policy and for clients who did not opt for policy the call lasted for 164 seconds which is lesser than clients who opt for policy. Thus clients spend more time on communication when they opt for the policy.



Clients are more likely to opt for policy from the month of May to October also in the month of September. Thus contact before the campaign brings more clients to opt for the policy.

Positive subscription to the policy during the campaign does not outnumber the failure to subscription. However the month of December seems to be favorable for the outcome of our interest.

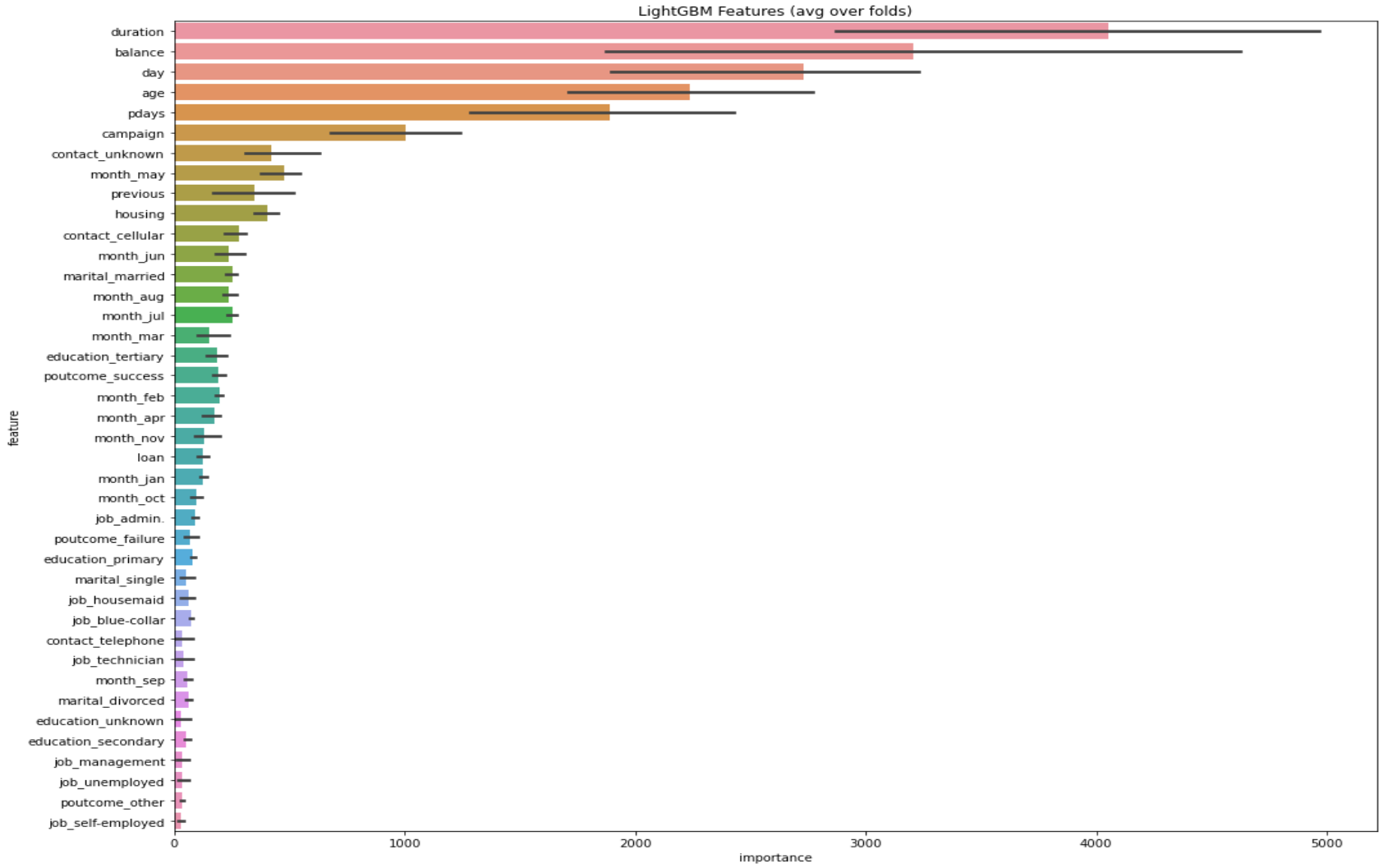


Model Performance Analysis

Model	Accuracy	Precision		Recall		F1 - Score	
		0(Don't buy the Policy)	1(Customer who buy the policy)	0(Don't buy the Policy)	1(Customer who buy the policy)	0(Don't buy the Policy)	1(Customer who buy the policy)
Logistic Regression	0.85	0.97	0.44	0.86	0.79	0.91	0.56
Random Forest Classifier	0.89	0.94	0.58	0.95	0.54	0.94	0.56
LightGM	0.91	0.93	0.63	0.96	0.47	0.95	0.54

The best performing model is LightGM as it gives the best accuracy. The precision score is good for LightGM model. However the recall score is good for Random Forest Classifier. Thus the recommended model is **LightGM**.

Feature Importance



It can be seen that by the best five features are Duration, Balance, Day, Age, Previous Day

Final Recommendations

1. Clients who buy the policy are more likely to fall between the age of 20-40. Also there are more number of clients beyond age of 60 who have opted for the policy.
2. Retired clients are more likely to buy the policy.
3. Clients who have their job description as management are more likely to opt for the policy.
4. Clients whose do not have default as their status are more likely to buy the policy.
5. Irrespective of the housing loan status there is a small difference(26%) between clients who have and not have housing loan and yet apply for the policy.
6. Clients with secondary and tertiary education are more likely to opt for the policy.

Note: All analysis was performed strictly to help the company understand their customers who opt for the policy thus help them approach the right customers.

Final Recommendations(conti..)

7. Cellular contact has a higher rate of subscription.

8. On average the call lasts for 426 seconds for the clients who opted for the policy and for clients who did not opt for policy the call lasted for 164 seconds which is lesser than clients who opt for policy. Thus clients spend more time on communication when they opt for the policy.

9. Clients are more likely to opt for policy from the month of May to October also in the month of September. Thus contact before the campaign brings more clients to opt for the policy.

10. Its recommended to use LightGM model for prediction of clients who will buy the policy.

11. The top five features which can be used are Duration, Balance, Day, Age, Previous Day

- Runa Veigas

Data Glacier

Your Deep Learning Partner

Thank You