

The background features several abstract, organic shapes in shades of purple and blue. A large, irregular shape dominates the right side, with a smaller circular shape above it and another smaller shape at the bottom right. The colors transition from a deep purple in the center to a lighter blue at the edges.

Bank Marketing Campaign Analysis

By: Kevin Agustinus L - Final Project JCDS



Contents

- Understand Problem Statement
- Exploratory Data Analysis (EDA)
- Model Building & Hyperparameter Tuning
- Recommendation for Bank Marketing Team

01. Problem Statement

Abstract:

Data from a marketing campaign run by Banco de Portugal is examined. The campaign's aim was to increase customers' subscription rates to fixed-term deposit products. Using knowledge from the course, a number of machine learning algorithms are implemented to answer the question: How can bank successfully market term deposit product?

Goal:

- To predict if the client will subscribe (yes/no) a term deposit (variable y).
- To find marketing strategies by analyzing bank marketing dataset and recommend the right customers to be targeted

02. Data Preprocessing

	dataFeatures	dataType	null	nullPct	unique	uniqueSample
0	age	int64	0	0.0	76	[29, 68]
1	job	object	0	0.0	12	[entrepreneur, admin.]
2	marital	object	0	0.0	3	[married, single]
3	education	object	0	0.0	4	[tertiary, unknown]
4	default	object	0	0.0	2	[yes, no]
5	balance	int64	0	0.0	3805	[1435, 1220]
6	housing	object	0	0.0	2	[yes, no]
7	loan	object	0	0.0	2	[yes, no]
8	contact	object	0	0.0	3	[cellular, unknown]
9	day	int64	0	0.0	31	[15, 30]
10	month	object	0	0.0	12	[dec, nov]
11	duration	int64	0	0.0	1428	[1074, 1204]
12	campaign	int64	0	0.0	36	[18, 19]
13	pdays	int64	0	0.0	472	[233, 202]
14	previous	int64	0	0.0	34	[0, 19]
15	poutcome	object	0	0.0	4	[failure, unknown]
16	deposit	object	0	0.0	2	[yes, no]

Data Source

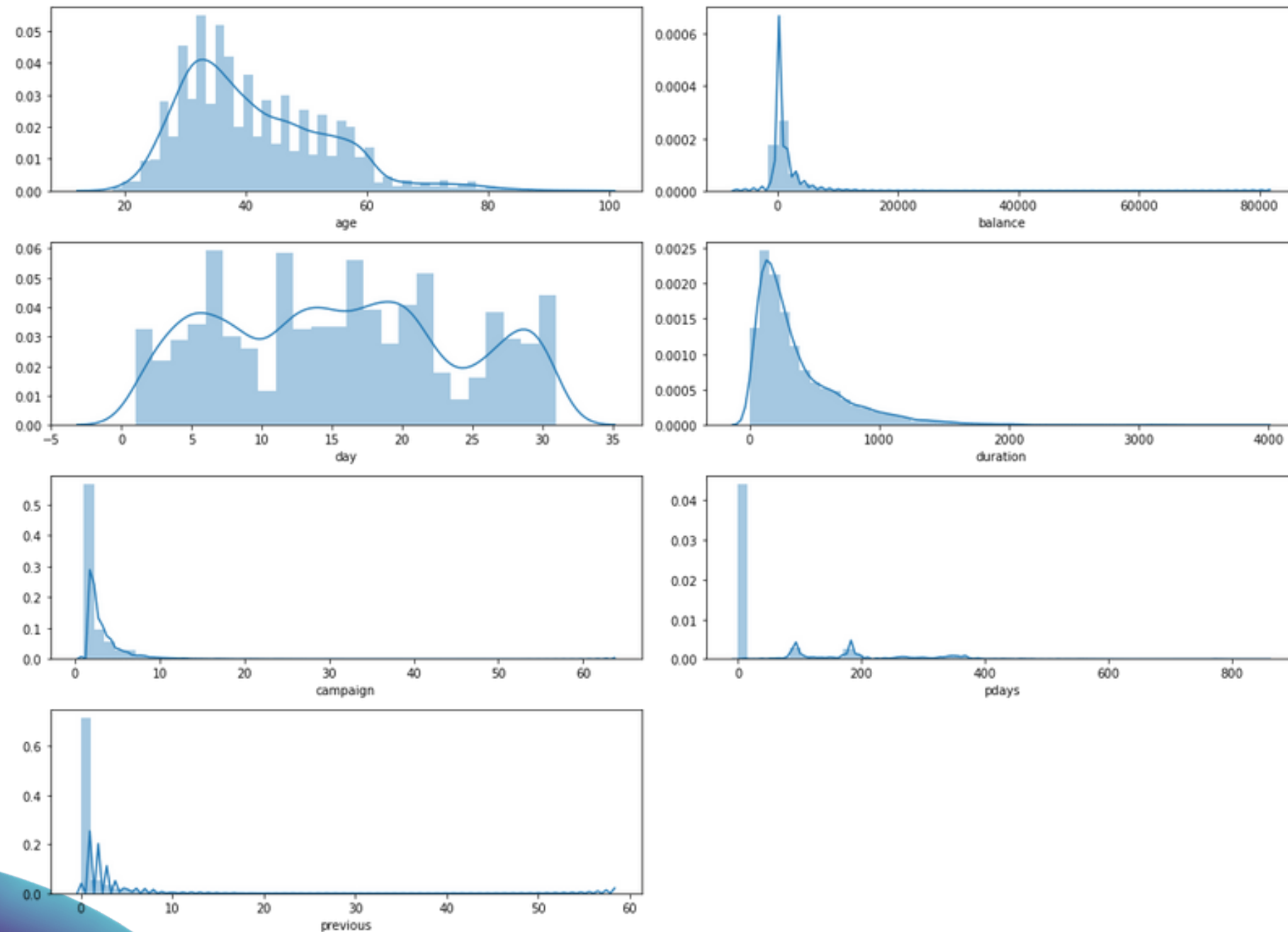
- [Moro et al., 2014] S. Moro, P. Cortez and P. Rita. A Data-Driven Approach to Predict the Success of Bank Telemarketing. Decision Support Systems, Elsevier, 62:22-31, June 2014

Data Info

- There is no missing values
- 16 attributes, with deposit as a target
- There is 11,162 rows data: integer and object/ categorical data
Object data will change into numerical data with LabelEncoding method before modeling

03. Exploratory Data Analysis

A. Numerical Data Exploration



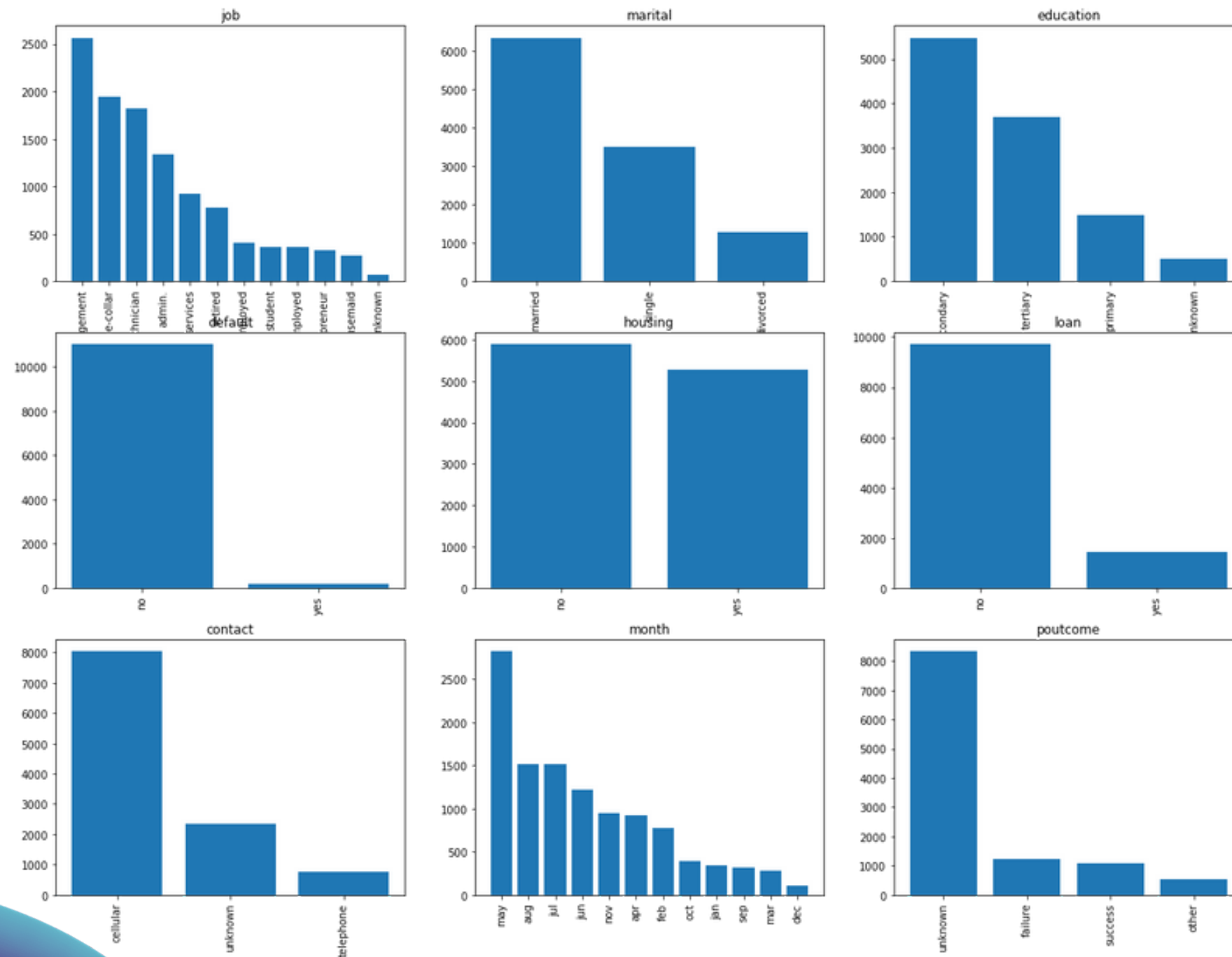
	age	balance	day	duration	campaign	pdays	previous
count	11162.000000	11162.000000	11162.000000	11162.000000	11162.000000	11162.000000	11162.000000
mean	41.231948	1528.538524	15.658036	371.993818	2.508421	51.330407	0.832557
std	11.913369	3225.413326	8.420740	347.128386	2.722077	108.758282	2.292007
min	18.000000	-6847.000000	1.000000	2.000000	1.000000	-1.000000	0.000000
25%	32.000000	122.000000	8.000000	138.000000	1.000000	-1.000000	0.000000
50%	39.000000	550.000000	15.000000	255.000000	2.000000	-1.000000	0.000000
75%	49.000000	1708.000000	22.000000	496.000000	3.000000	20.750000	1.000000
max	95.000000	81204.000000	31.000000	3881.000000	63.000000	854.000000	58.000000

Statistically we can tell that :

- We have clients of the youngest 18 years old and the oldest clients 95 years old
- Average clients balance is 1,528, the lowest balance amount is minus 6,847 and the highest amount balance is 81,204
- Average last contact is 15 days, with minimum 1 day last contact and maximum 31 days last contact
- Average number of contact (campaign) is 2 - 3 times, with minimum 1 time number of contact and maximum 63 times number of contact
- We still need to explore what is -1 pdays represent in number of day passed by after the client was last contacted from a previous campaign

03. Exploratory Data Analysis

B. Categorical Data Exploration



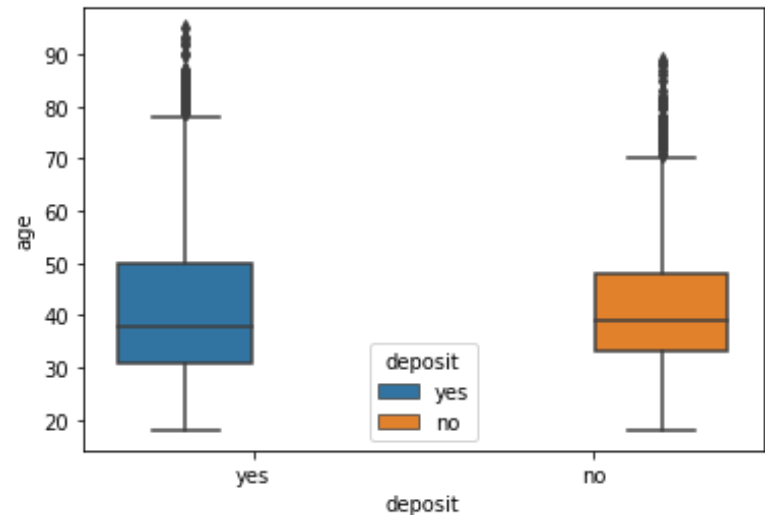
Insight from diagram besides are:

- Most clients work as Management, Blue-collar, technician, and admin
- Client population is most likely married
- Majority of client population are come from Secondary Education Level
- Contacted clients are mostly from people who work as management, blue-collar, and technician. We also can tell that housemaid is less likely to be contacted to offer term deposit
- Most population tend to have housing loan but not have personal loan
- Clients population is most likely being contacted using cellular
- May - Aug is the bussiest month to contact the customers to offer term deposit

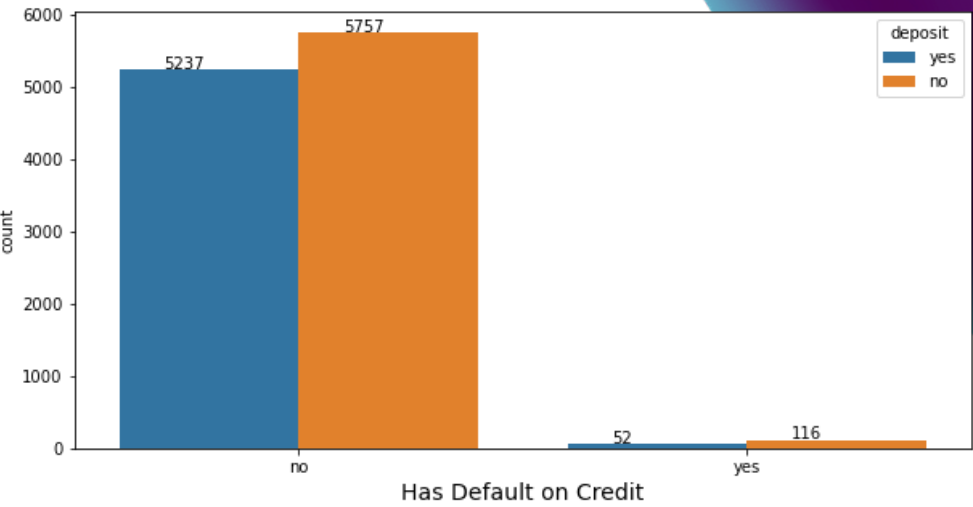
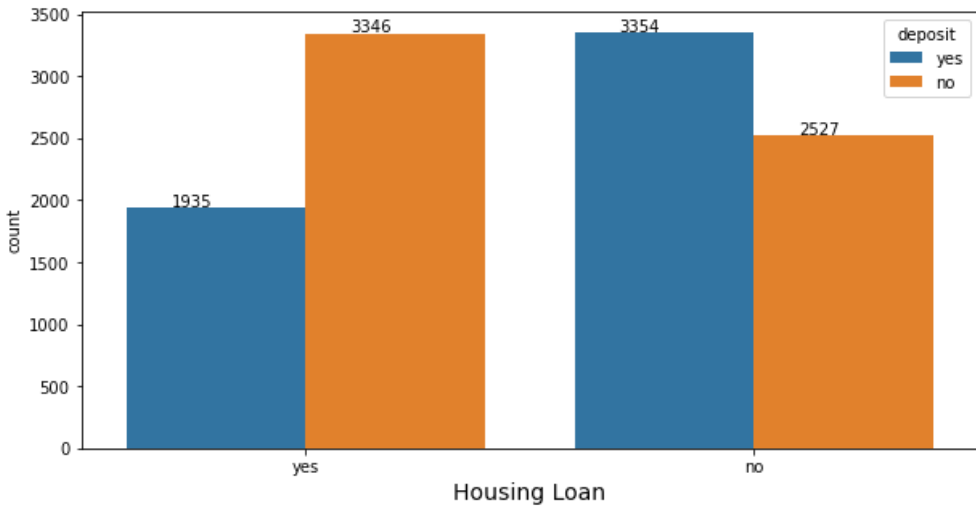
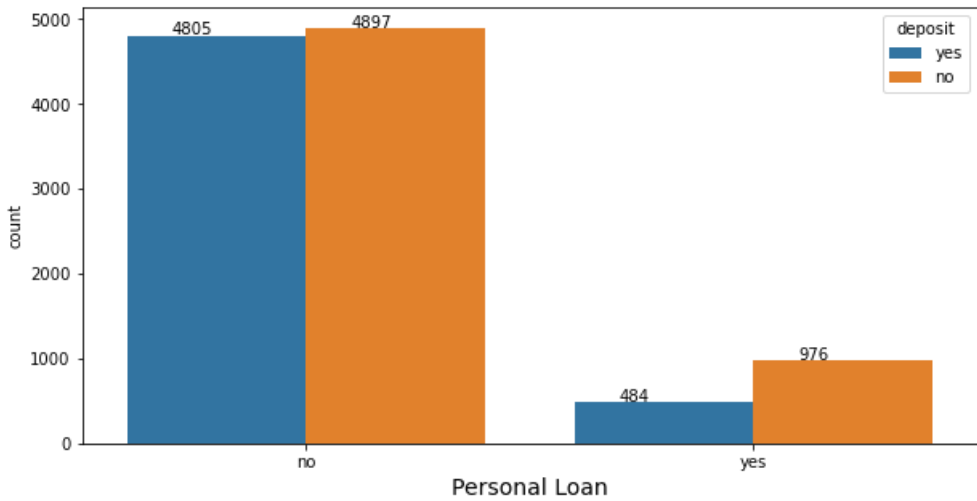
03. Exploratory Data Analysis

Client Information vs. Yes/ No Subscribe Term Deposit

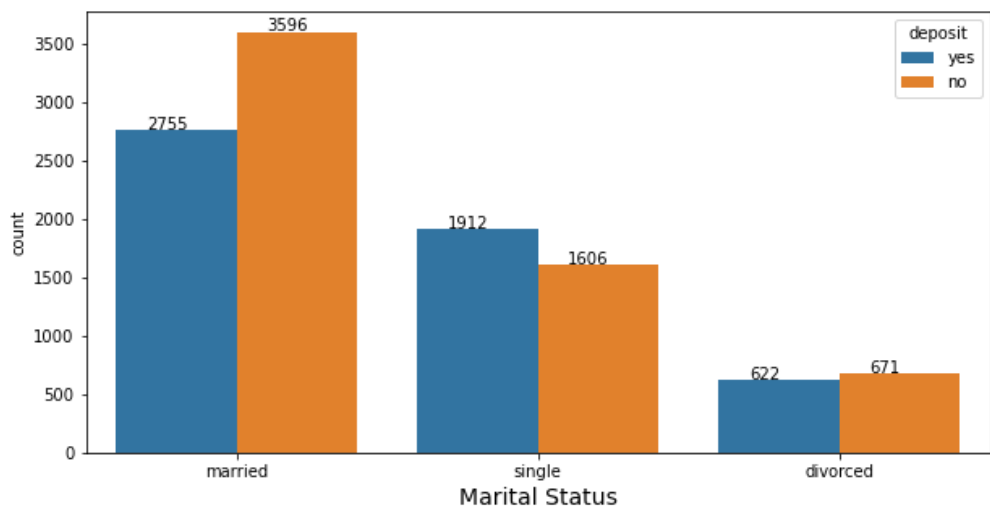
Age



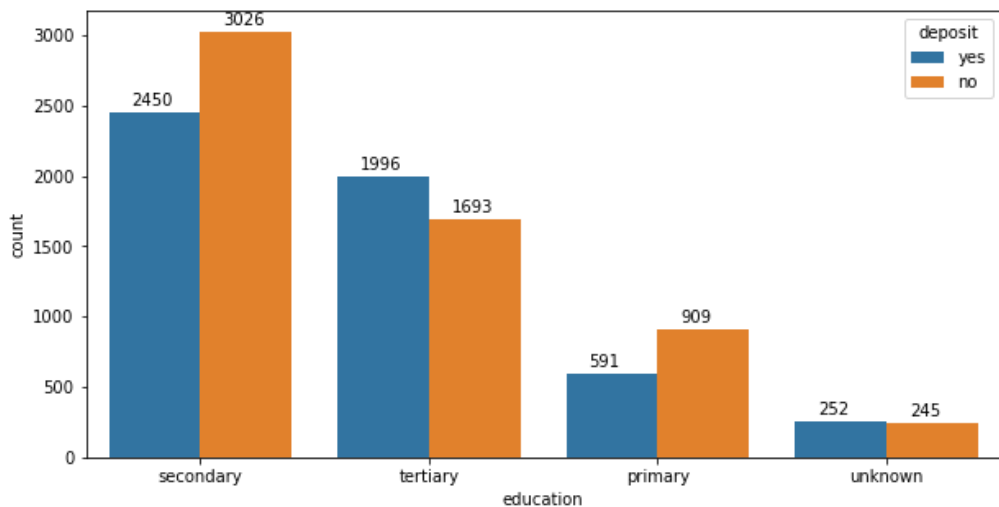
Personal Loan, Housing Loan, Default on Credit



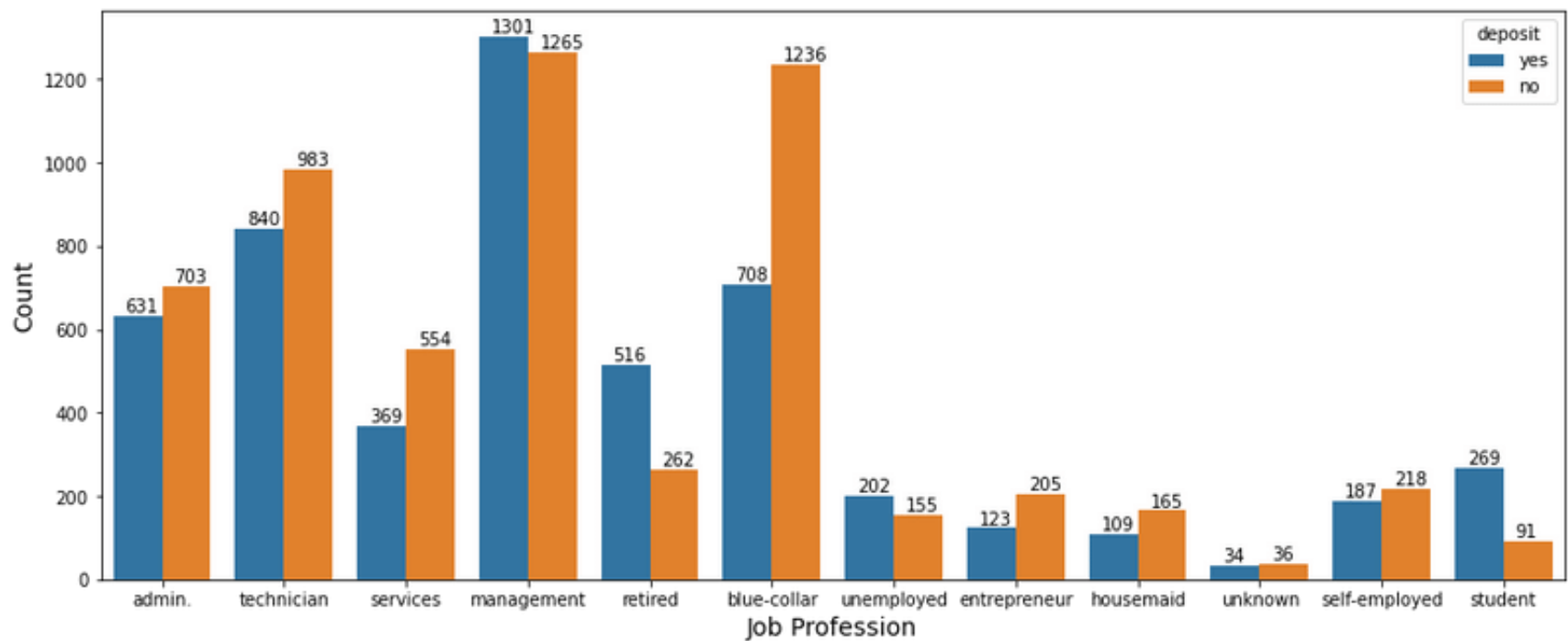
Marital Status



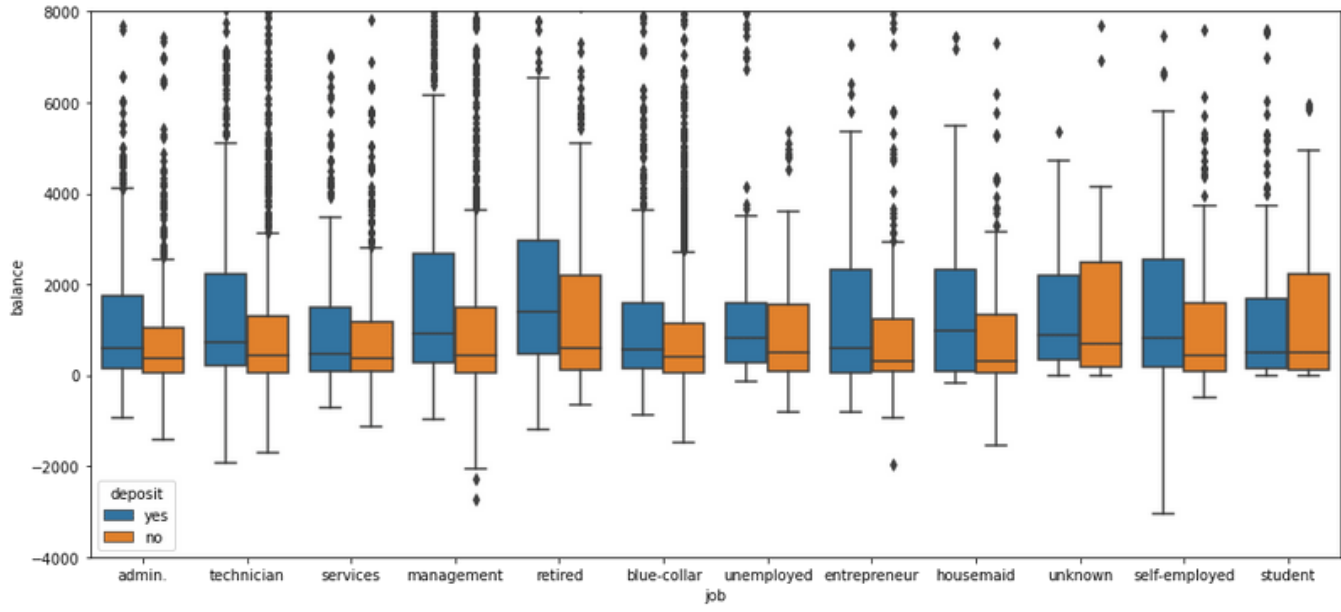
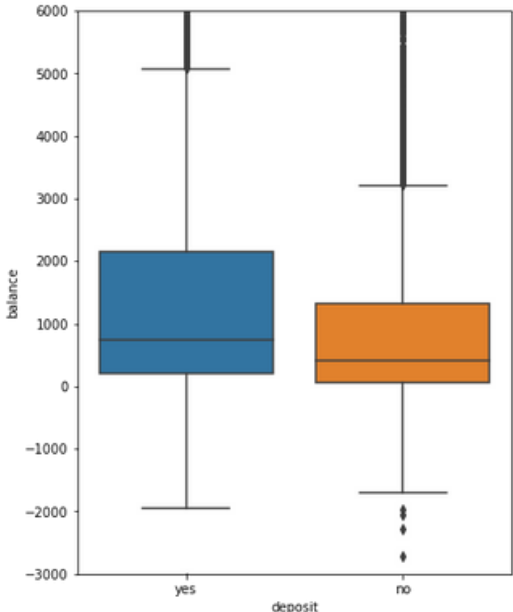
Education Level



Job Profession



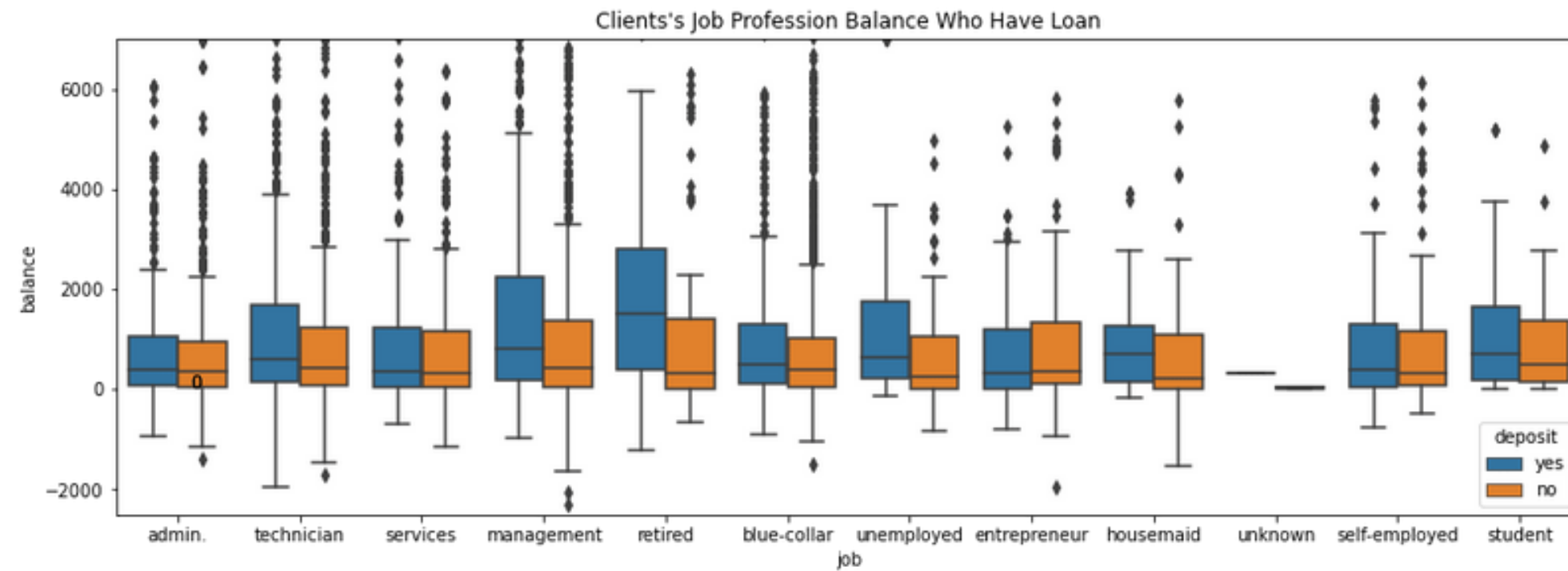
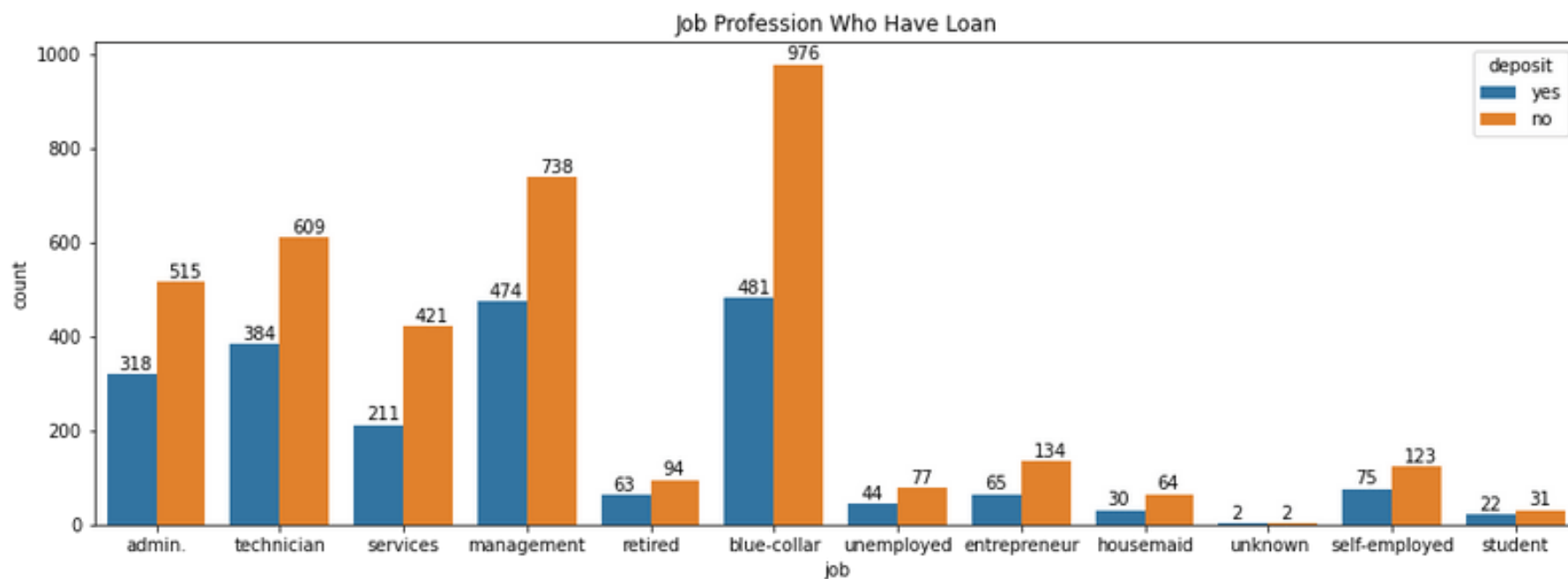
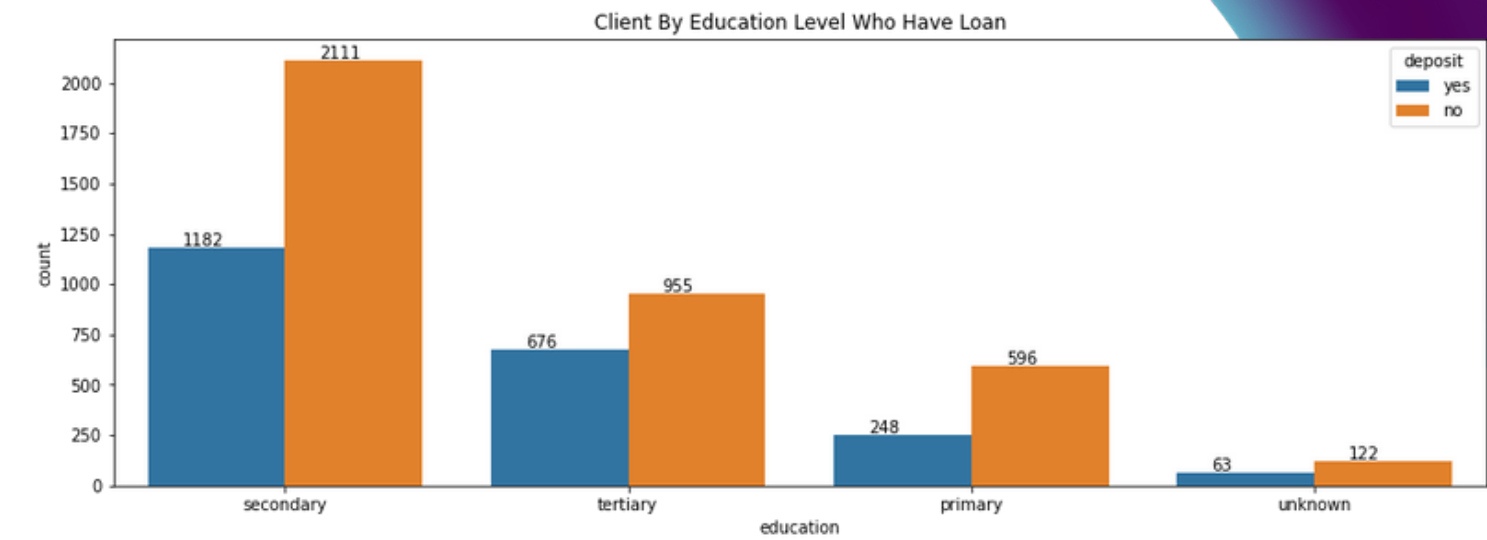
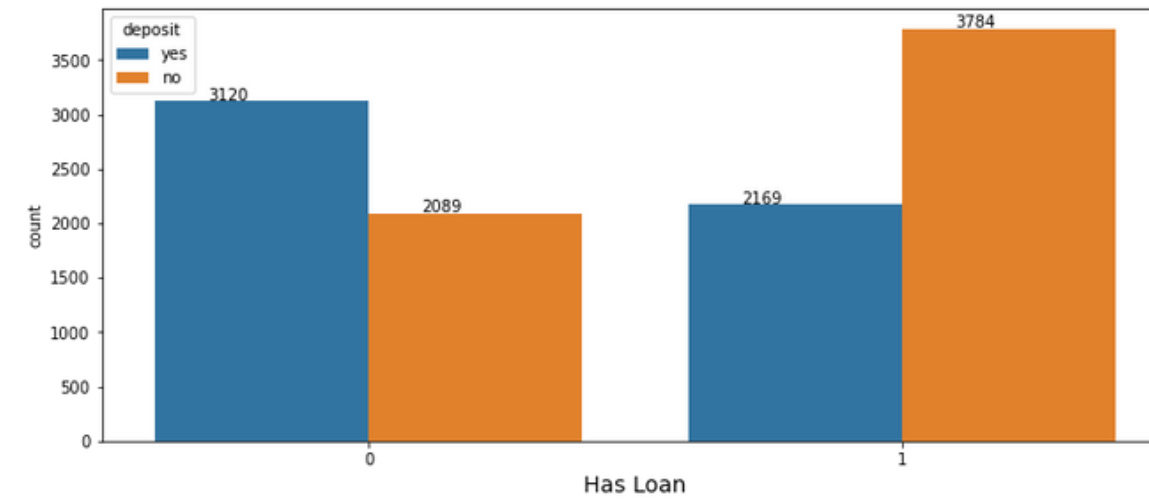
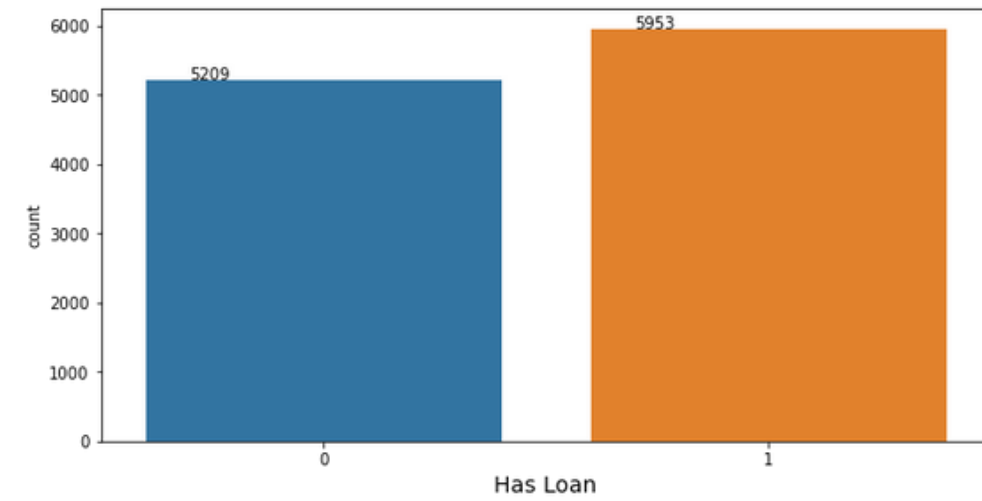
Balance vs. Job Profession



- Client's age who subscribe term deposit and not have many outliers numbers
- For client who have some debt that need to be paid, tend to not subscribe term deposit
- Married clients tend to subscribe term deposit
- Client from secondary education level have highest numbers as term deposit subscribers
- Median balance of client who subscribe term deposit is higher than those client who do not subscribe in all job profession categories

03. Exploratory Data Analysis

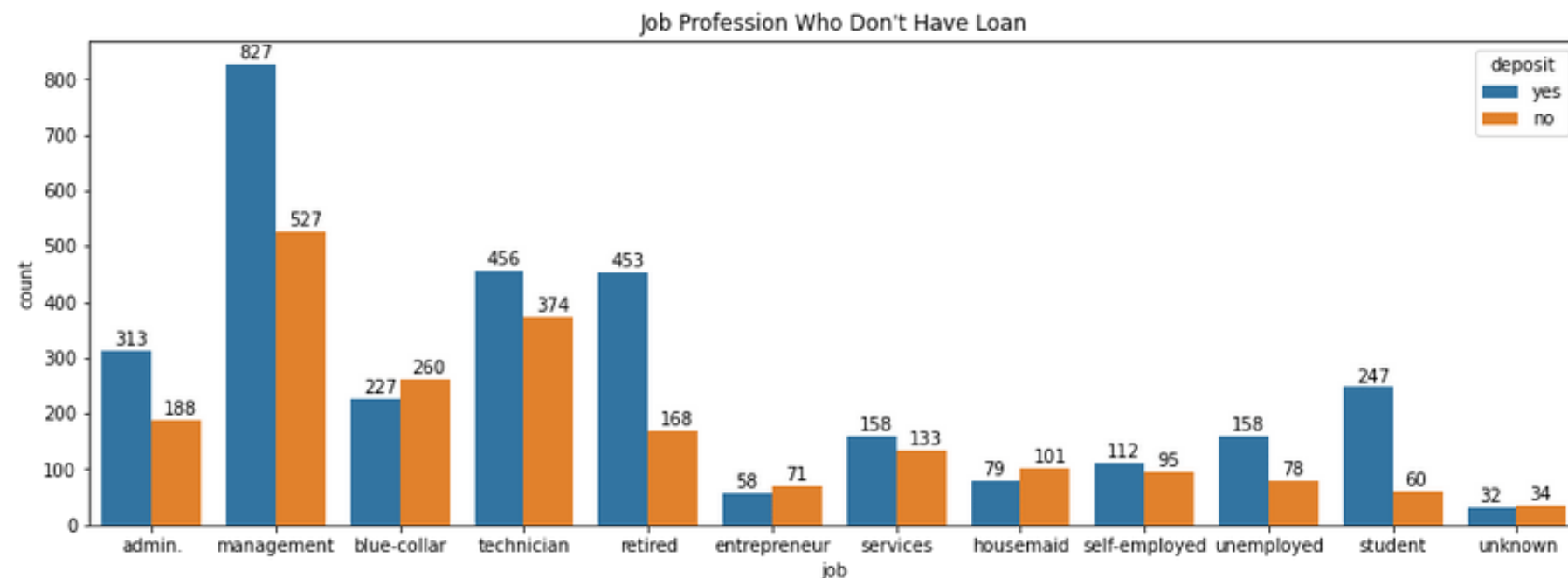
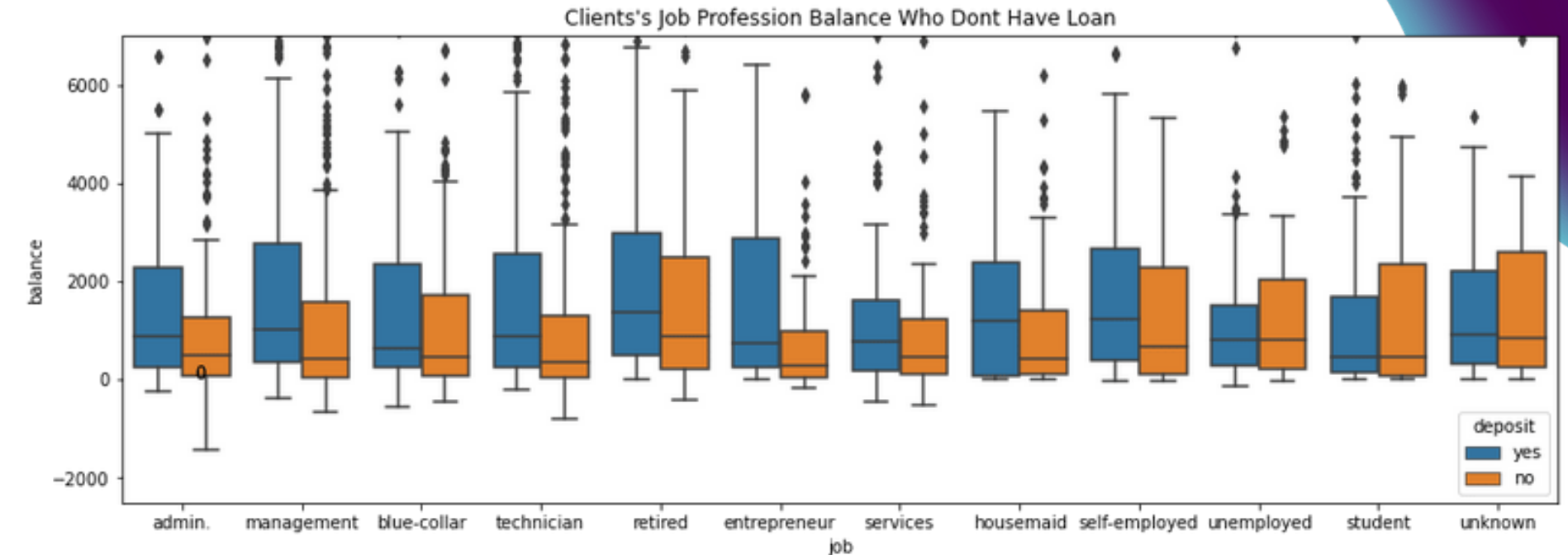
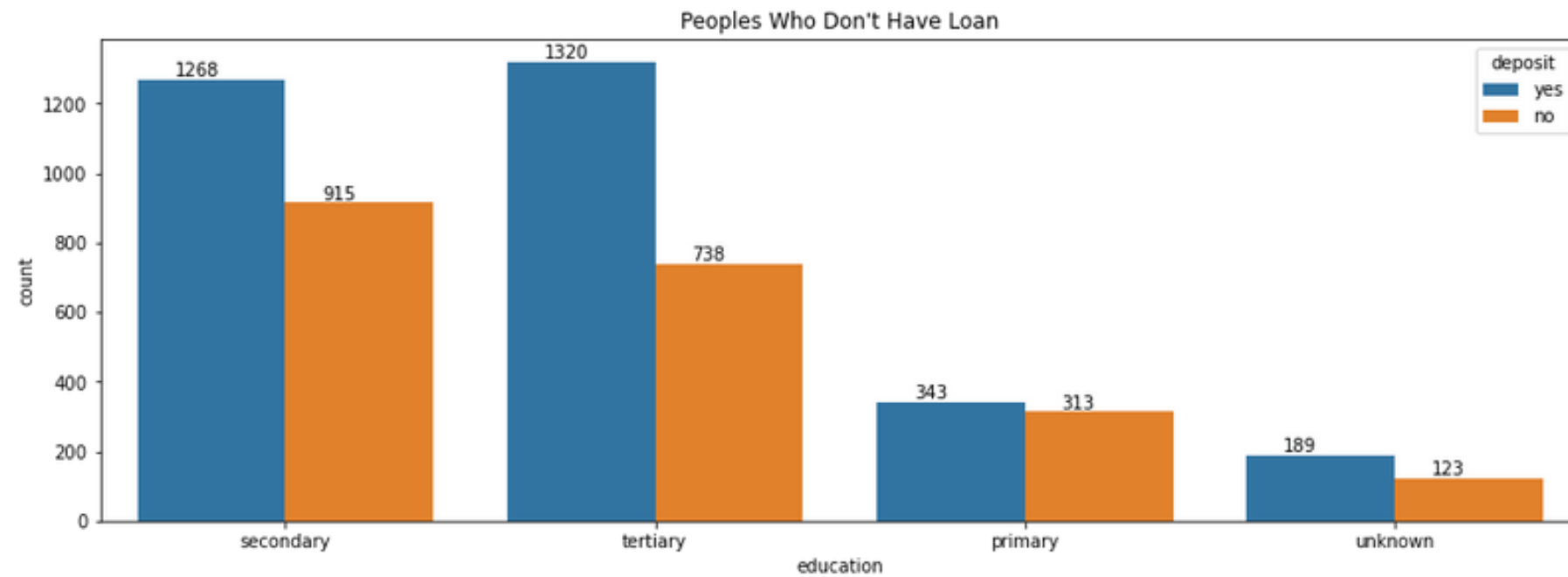
Explore Client Who Have Loan



- Loan-Free and having loan with sufficient good balance, good education background, clients who work as admin, management, technician, retired, and student are the important attributes of opening term deposit

03. Exploratory Data Analysis

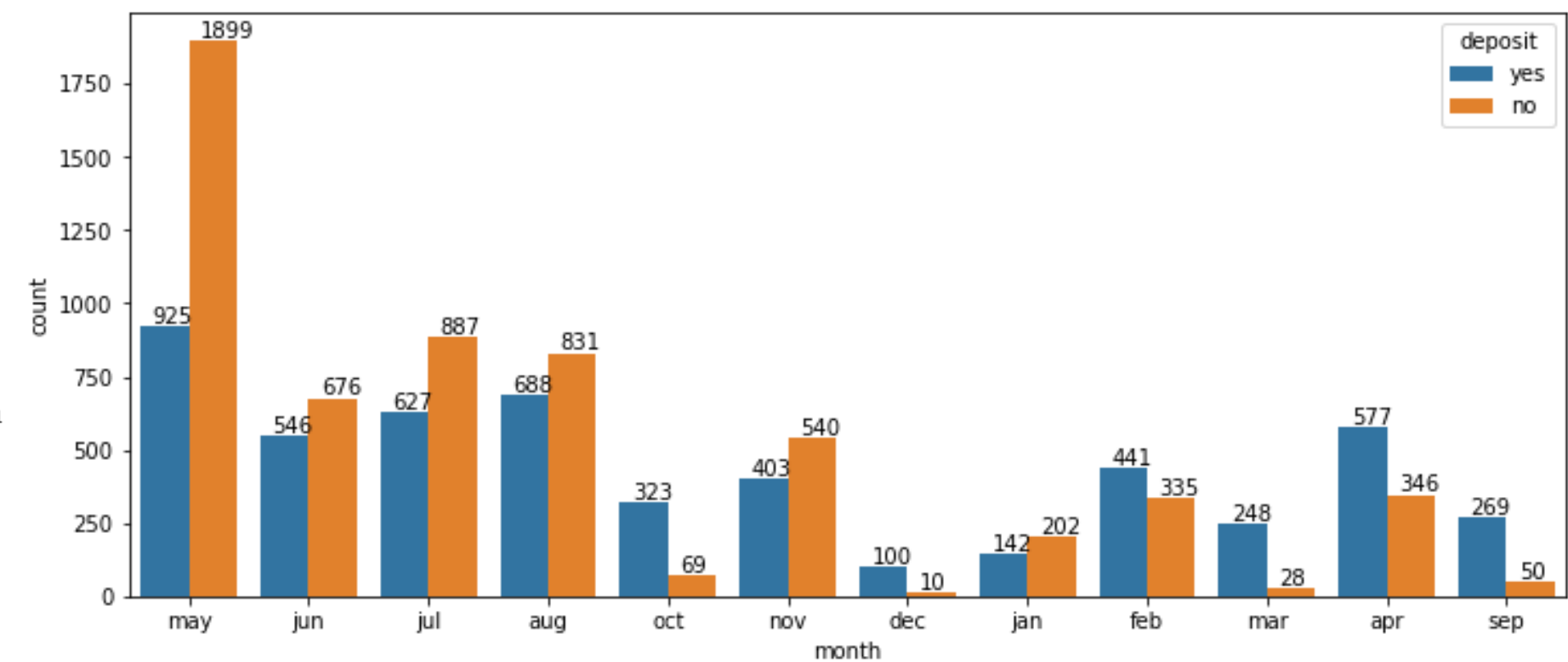
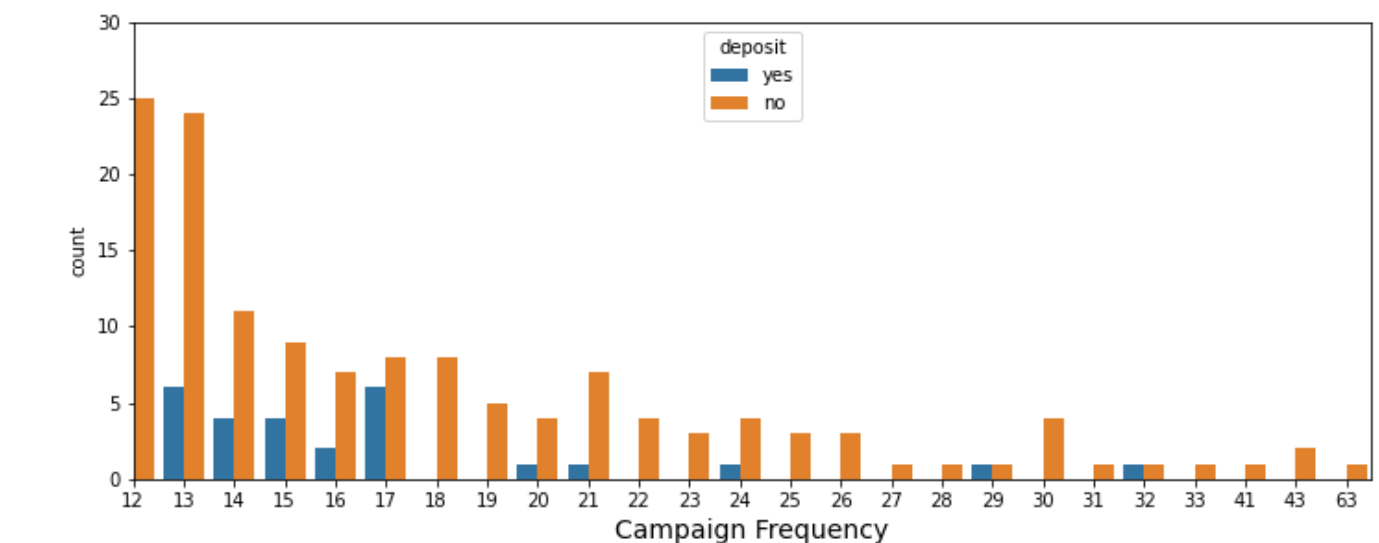
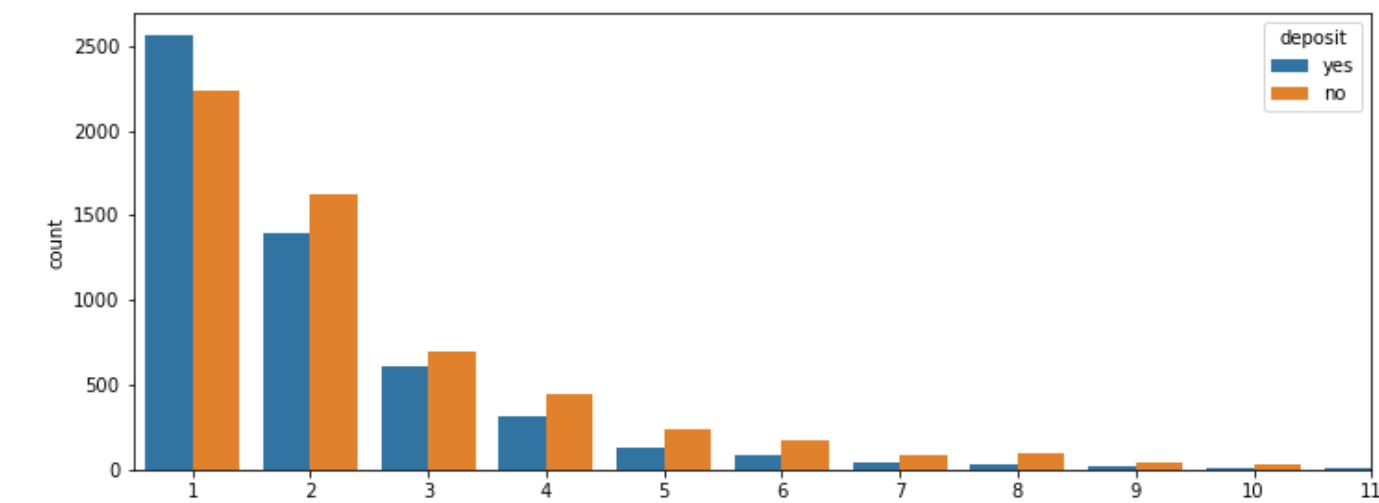
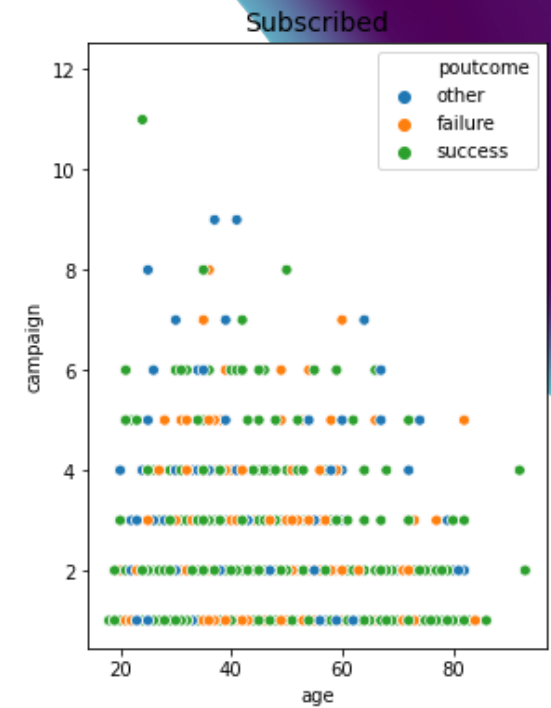
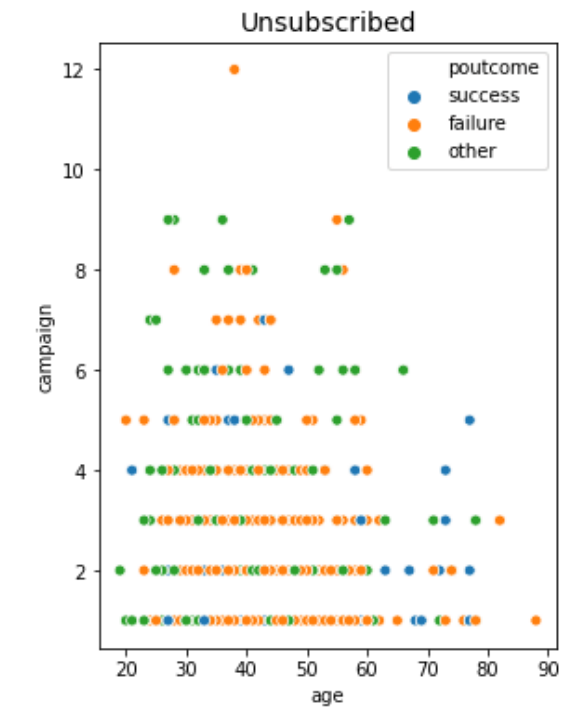
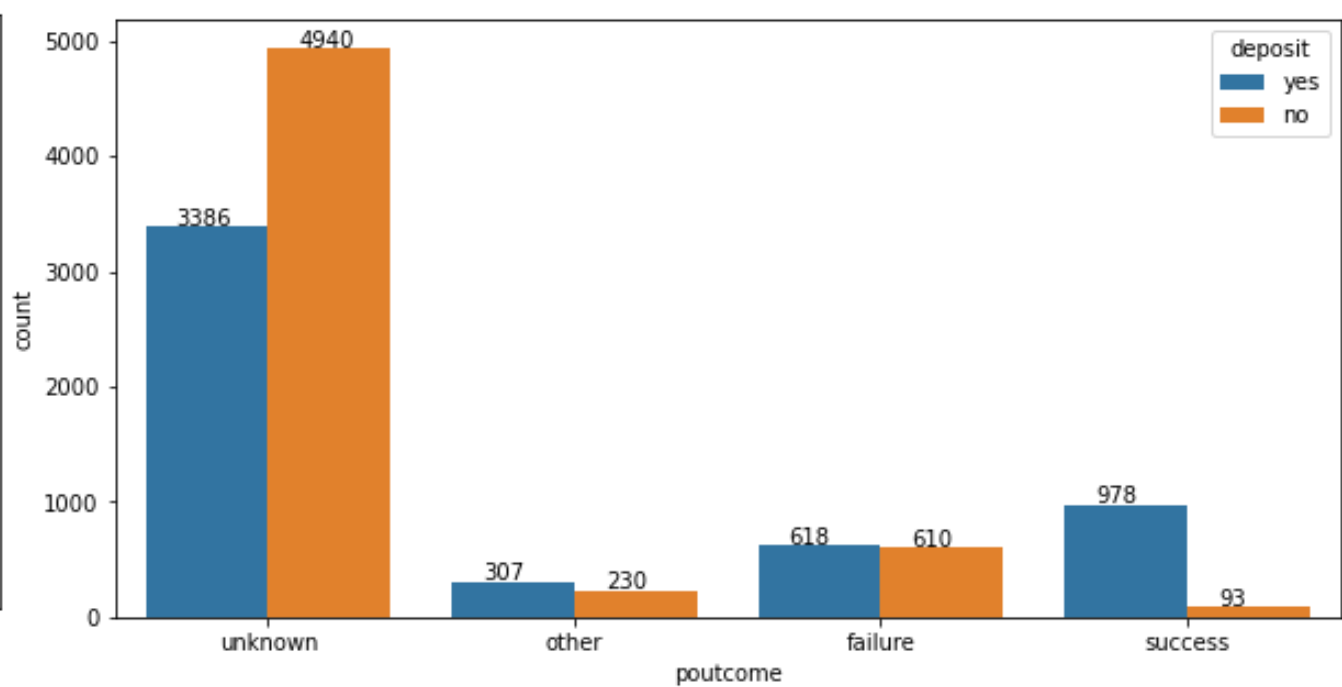
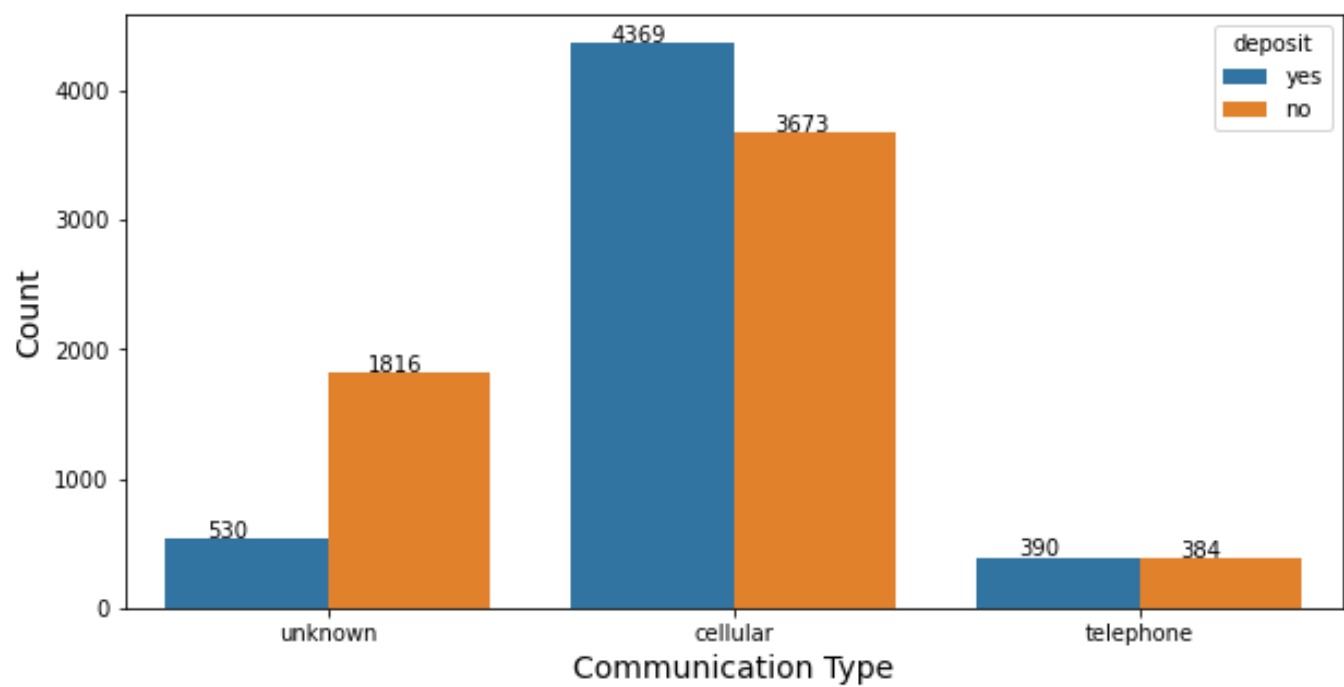
Explore Client Who Don't Have Loan



- Loan-Free and having loan with sufficient good balance, good education background, clients who work as admin, management, technician, retired, and student are the important attributes of opening term deposit

03. Exploratory Data Analysis

Deposit Relations with Marketing Communication Type



- Average contact duration with client who has subscribed: 8.95
Average contact duration with client who hasn't subscribed: 3.71
- The more often client is contacted, the lesser number of subscribed deposit is
- we can get information that if client reject to subscribe term deposit, there is chance that for them to subscribe.
- Client with many clients with 60 years old and above subscribed term deposit

04

Model Building & Hyperparameter Tuning

- Logistic Linear
- XGBoost
- Random Forest

04. Machine Learning Model

Train Model	Precision	Recall	F-1 Score
Logistic Linear			
Without Tuning	0.64	0.61	0.63
With Tuning	0.64	0.62	0.63
XGBoost			
Without Tuning	0.74	0.8	0.76
With Tuning	0.75	0.62	0.68
Random Forest			
Without Tuning	0.73	0.63	0.68
With Tuning	0.76	0.58	0.65

Recommendation

In order to increase client subscribers of term deposit, here are some strategies need to do:

- Data showed that client with job profession from blue-collar, entrepreneur, housemaid, etc are somewhat beneficial but before contact to them, marketing team should check their balance, loan, and education as consideration. It would be more successful if marketing team approach client who work as management, technician, or retired people especially for them that high balance amount in their bank account
- Prepare marketing material to be more interesting to attract client, the more engaging communication will increase chances for client to open term deposit
- Most subscribers are the ones who are contacted in months of February to May and September and October. So cutting down the activities in other months and focusing more in these periods will be beneficial.
- There are many loan-free clients who subscribed to the term deposit and also there are some clients who do have loan but still subscribed to the deposit and there are loan-free ones who didn't subscribed to the deposit. The main reason for this is the balance. The ones with sufficiently good balance are tend to subscribe to the term deposit.
- Based on data, the number of contacts made with most of the subscribers is less than 4. Therefore, contacting person for more than 3-4 time should be avoided.

