MVP

Overview:
Predicting if a new customer of the bank will accept the deposit offer. For financial institutions to improve their future marketing campaigns, they need to find ways that can lead them to apply the best strategies. By analyzing the "Bank Marketing Dataset", banks will be able to get useful insights in which they can use to answer the project main question and so improve their marketing campaigns in the future.
In this project, the dataset used to predict whether the new customers will accept the deposit offer from the bank?

Data Description:
The "Bank Marketing Dataset" is retrieved from:
https://www.kaggle.com/janiobachmann/bank-marketing-dataset/version/1
Dataset contains 11.2k records with 10 features as the following:

Age	Customer's age 'numeric'
Job	Job type 'categorical'
Marital	Marital status 'categorical'
Education	Level of education 'categorical'
Default	Has credit in default? 'categorical'
Balance	Bank balance 'numeric'
Housing	Has housing loan? 'categorical'
Loan	Has personal loan? 'categorical'
Contact	Communication type 'categorical'
Day	Last contact day of the week 'numeric'
Month	Last contact month of the week 'categorical'
Duration	Last contact duration in seconds
Campaign	No. of contacts performed during this campaign for this client 'numeric'
Pdays	No. of days passed by after the client was last contacted from the previous campaign 'numeric'
Previous	No. of contacts performed before this campaign for this client 'numeric'
Poutcome	Outcome of the previous marketing campaign 'categorical'

Output variable "desired target", has the client subscribe to the term deposit offer? (Binary: 'yes' or 'no')

Tools:
Numpy
Pandas
Matplotlib
Seaborn
Sklearn

Algorithms:
Logistic Regression
Decision Tree Classifier
Support Vector Machine
Random Forest Classifier

We can indicate that during the call is highly correlated with the target variable "deposit". As the duration of the call is more, there is a high chance that the client shows interest in the term deposit which mean that there is a high chance that the client will subscribe to term deposit.

Target value vs other features:
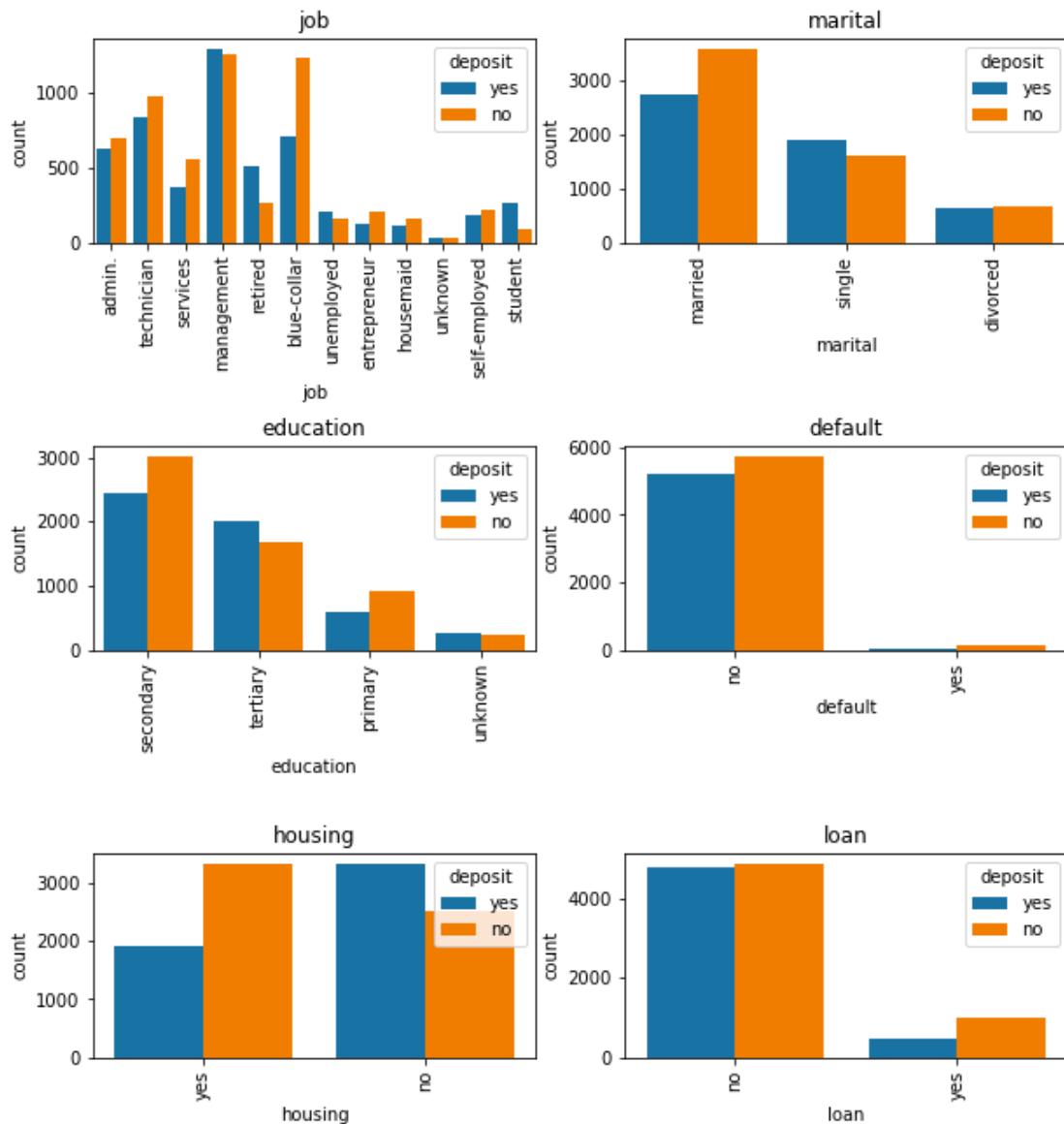
Data Processing:
Check null
Check duplicate
EDA
Convert categorical data

Conclusion:
Multiple classification algorithms have been applied on the data, and it seems that Random Forest Classifier model performed excellent with accuracy of 84% comparing to other algorithms.

Done by:
Sarah Mohammedfareed Abu Zabibah
Sarahabuzabibah@gmail.com