Presentation

CAPSTONE PROJECT

IMARTICUS LEARNING

USING ML AND POWERBI

By Naveen B

Presentation

BANK MARKETING ANALYSIS AND PREDICTION

MACHINE LEARNING

TABLE OF CONTENTS

INTRODUCTION

- Today organizations, which hire marketing management are analysis of organization's marketing data

- Analysisng the Customer data is one of the most typical applications of data science and machine learning.

- Prediction of the results of the marketing campaign for each customer and clarification of factors

Machine Learning

OBJECTIVE

- Predicting the future results of marketing companies based on available statistics and, accordingly, formulating recommendations for such companies in the future.
- Building a profile of a consumer of banking services (deposits).
- Make recommendations for future campaigns

PROJECT FLOW

- Importing the Libraries
- Collection of data
- Data Pre-Processing
- Visualizing the data
- Model Selection
- Evaluating the Model

DATA PRE-PROCESSING

- Data Pre-processing is a process of cleaning the acquired dataset, by removing duplicates, checking outliers, creating dummies, label encoding, finding and filling or dropping out the null values.

EDA

EXPLORATORY DATA ANALYSIS

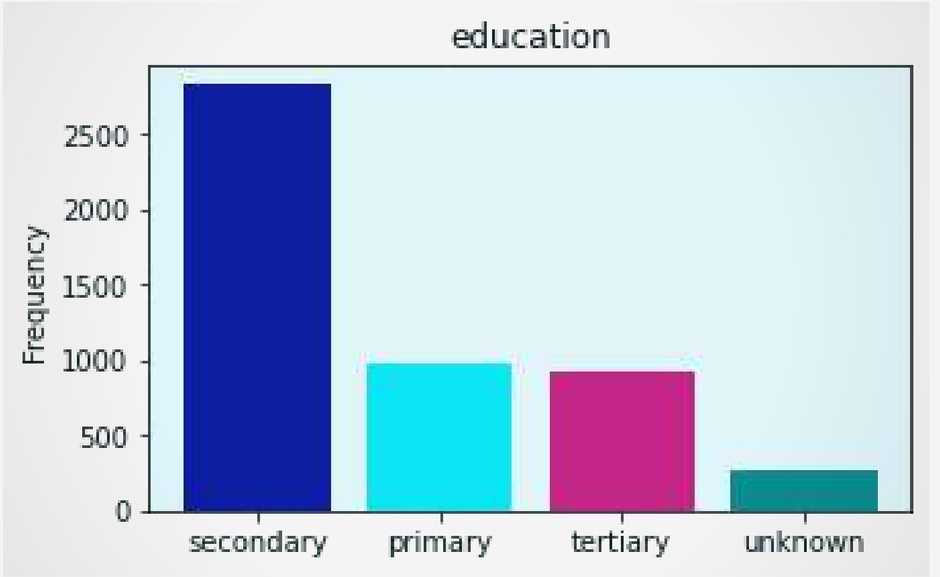
- In Exploratory Data Analysis, consists of Inspecting and cleansing, transforming, and modeling data.
- To visualize the data shows the range between the rows and columns that are taken in the data.
- To analysis the age shows the differential formats according to their ages

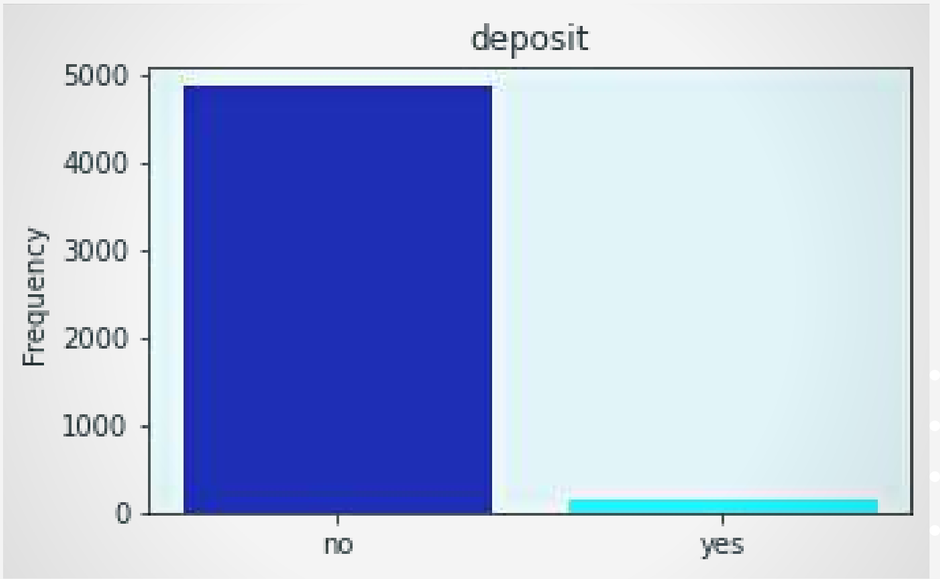
EDA

VISUALIZING CATAGORICAL VARIABLE

Bar Plot

- To represents the category of data with rectangular bars with lengths and heights that is proportional to the values which they represent.
- This graph shows the number of observations within each given interval.

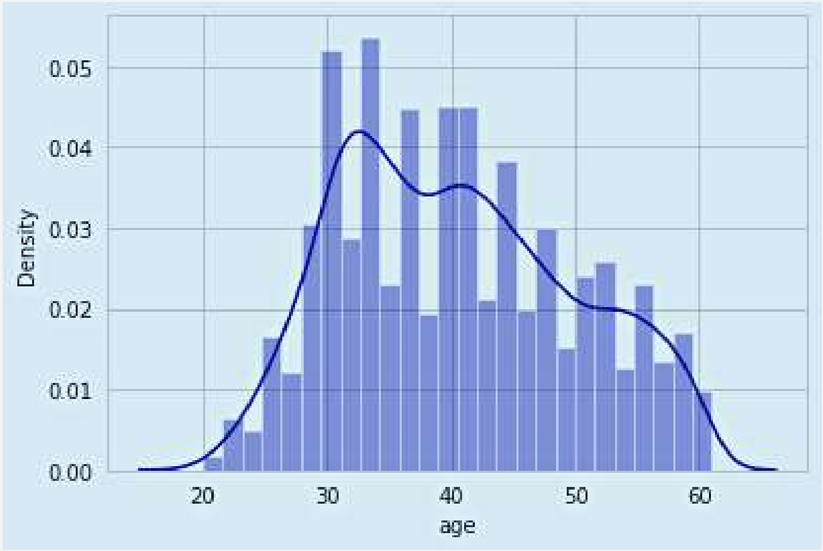




VISUALIZING NUMERICAL VALUE

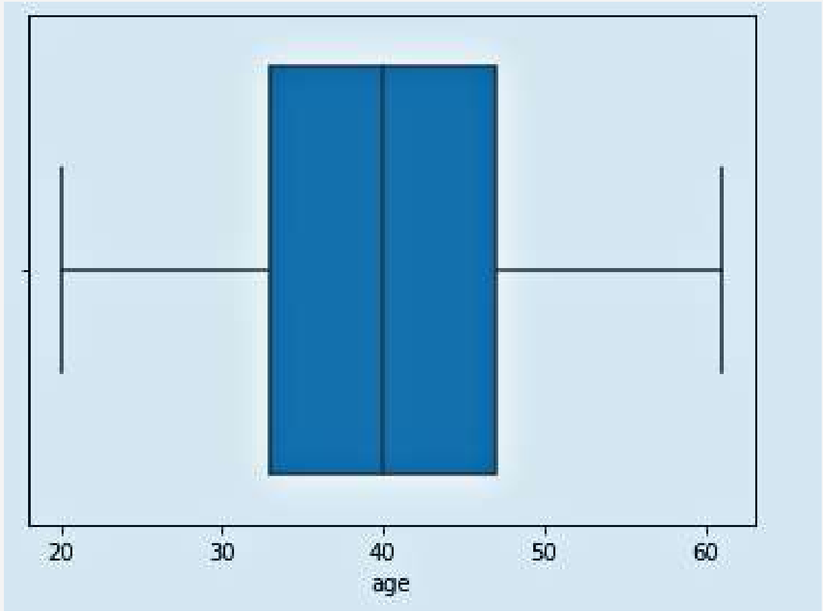
Dist Plot

- The distplot represents the univariate distribution of data i.e. data distribution of a variable against the density distribution.

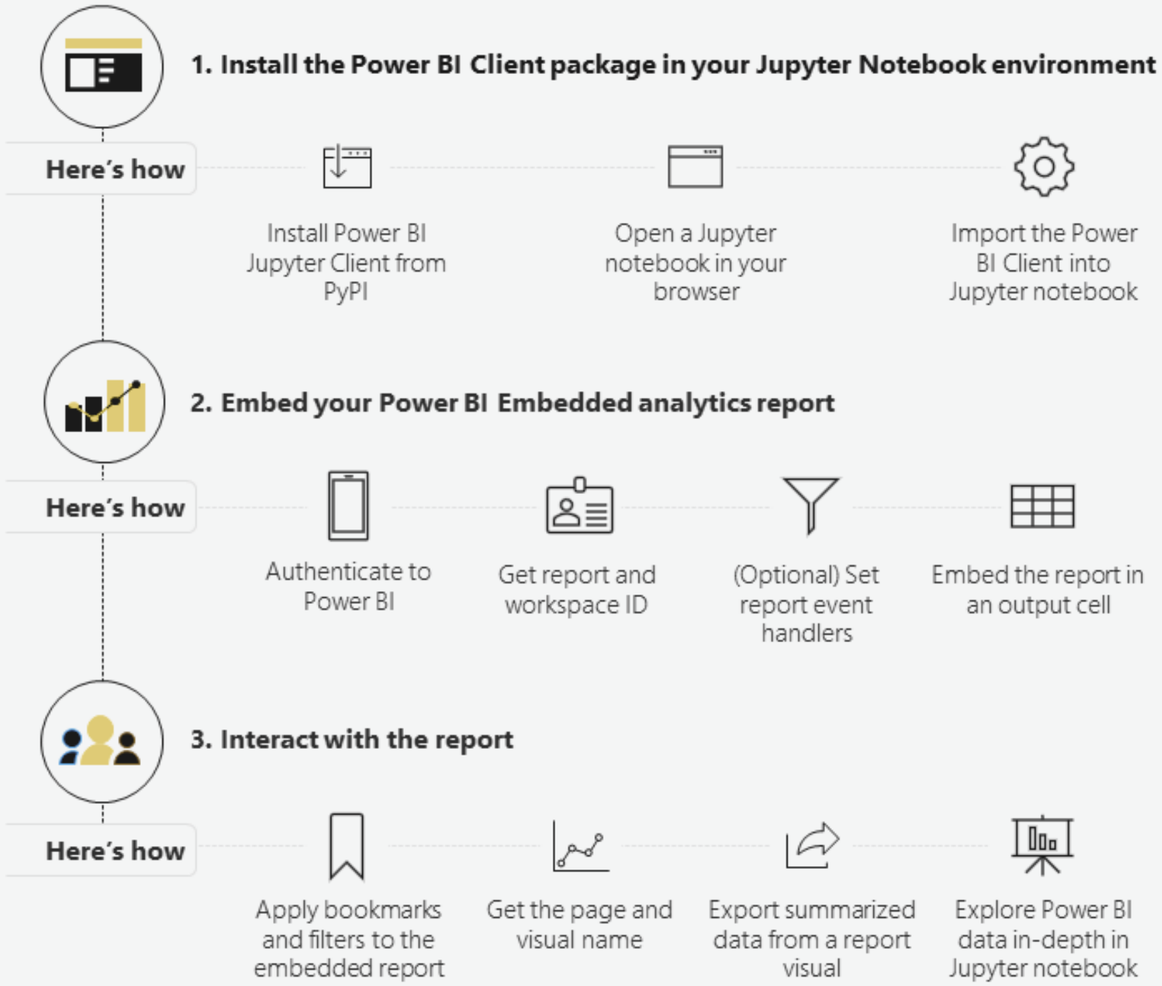


Box Plot

- To display the summary of the set of data values having properties like minimum, first quartile, median, third quartile and maximum.



POWER BI

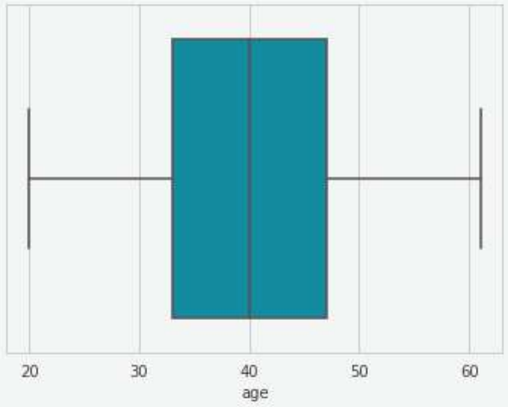


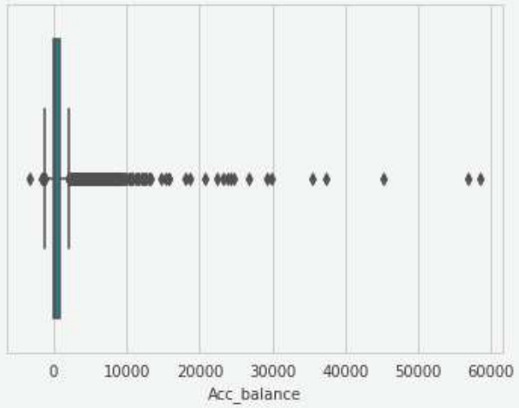
- Get your Power BI analytics in a Jupyter notebook with the new powerbiclient Python package.
- The new package lets you embed Power BI reports in Jupyter notebooks easily
- You'll be able to export data from visuals in a Power BI report to the Jupyter notebook for in-depth data exploration.
- You can also filter the report for quick analysis or use bookmarks to apply a saved view.

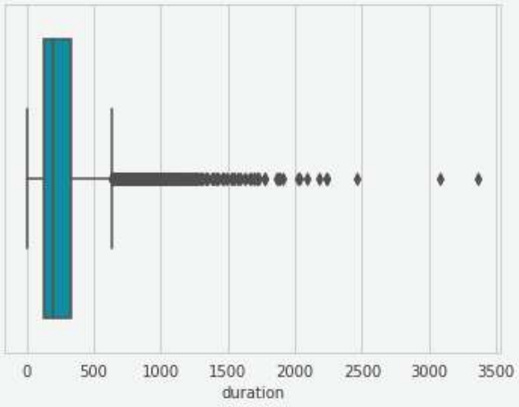
CHECKING OUTLIERS

- It is the process of identifying data that have extreme values compared to rest of the data distribution.
- We are checking the outliers for only the required columns in our dataset.
- After the execution of IQR method we are supposed to find that outliers are present in our dataset.

Outliers







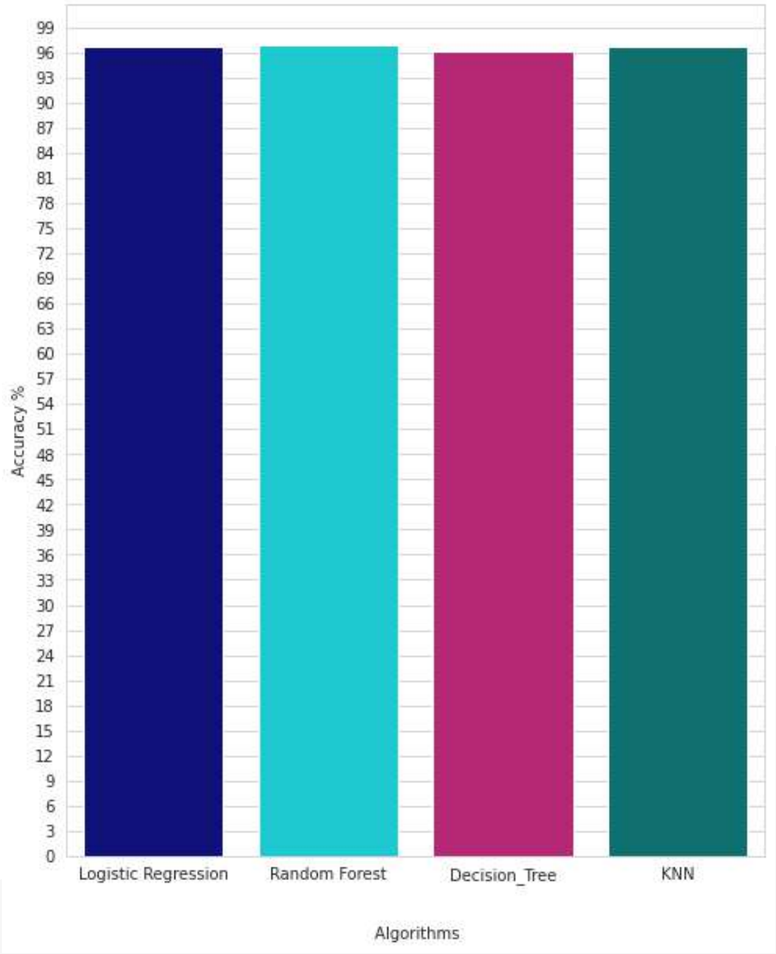
MODEL SELECTION

- Model selection is the process of choosing one among many candidate models for a predictive modeling problem.

CLASSIFICATION

- Logistic Regression
- Random Forest
- Decision Tree
- K-Nearest Neighbor

RESULT OF MODELS



COMPARE ACCURACY

- Logistic_acc = 96.82666666666667
- Random_Forest_acc = 96.96000000000001
- DecisionTree_acc = 96.18666666666667
- KNN_acc = 96.8

CONCLUSION

- The customer's account balance has a huge influence on the campaign's outcome.
- Number of contacts with the customer during the campaign is also very important.
- Future campains should concentrate on customers from age categories below 30 years old and above 50 years old.

Thank You