



# ML - LEAD SCORE CASESTUDY

ASSIGNMENT BY

1. HARITHA MUNDRA

2. SIDDAPPA K

# Problem Statement

X Education Company sells online courses and many professionals who are interested browse their and other website for courses. Leads are generated once they fill in the form with email id and/or mobile number or by past referrals.

The typical lead conversion rate is around 30% only. They need leads such that the conversion rate should increase, and the sales team will now be focusing more on communicating with the potential/hot leads rather than making calls to everyone.

# Objectives

X Education Company want us to build a model to rate every lead by a score 0 to 100 based on the predictive conversion factors.

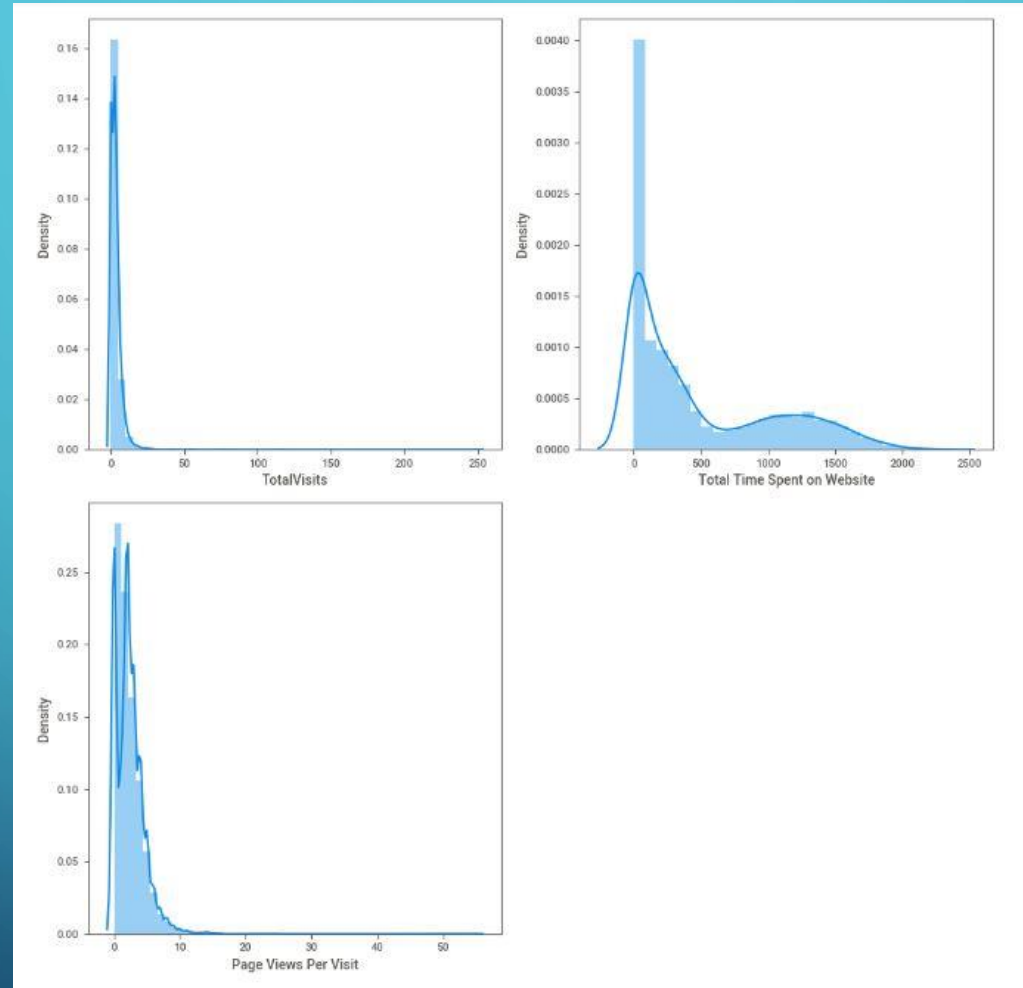
The CEO of the company wants us to give such a model which can predict the lead conversion more than 80%.

They also expect us to minimize the sales team manpower cost and to make calls only to the potential leads.

# Approach

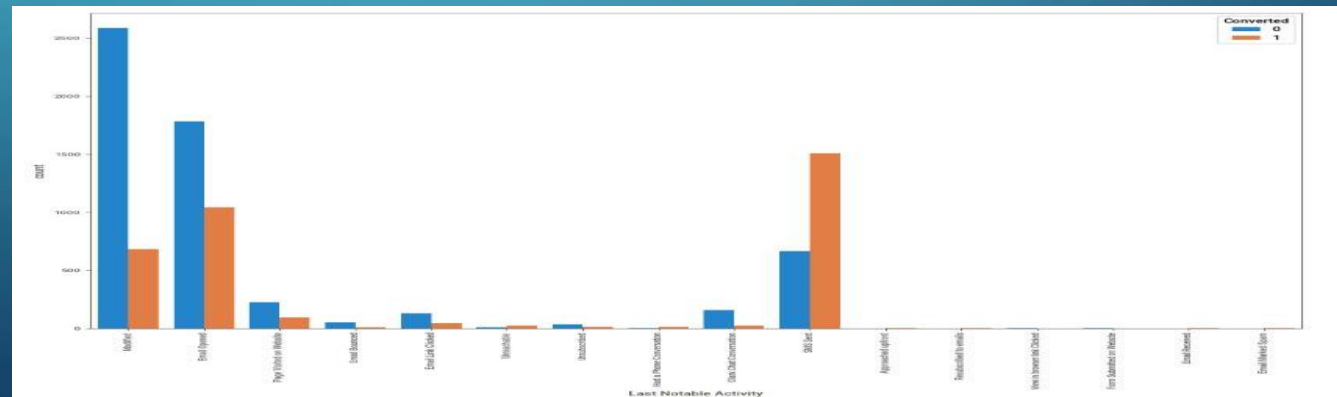
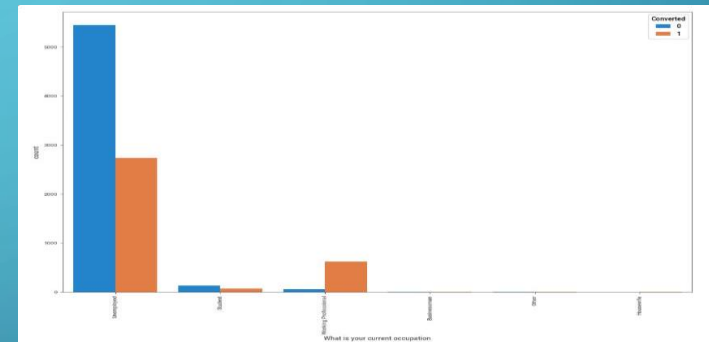
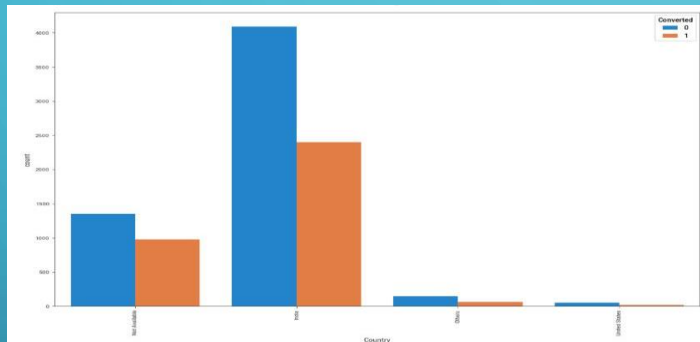
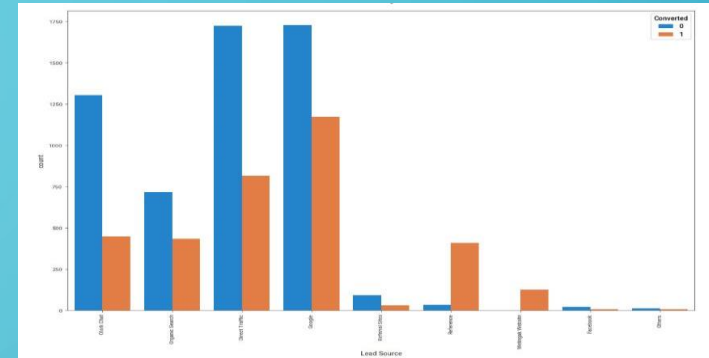
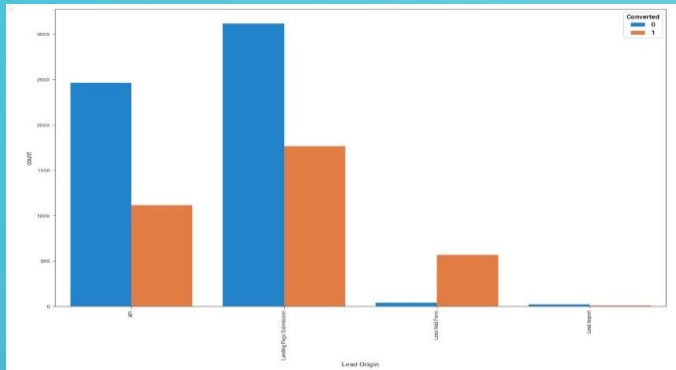
- Read and Understand Data
- Run SweetViz an automated EDA for quick analysis
- Data Cleaning
- Data Visualization and Engineering
- Prepare data for Prediction
- Scaling
- Build Model to map Prediction to Probability Matrix
- RFE and VIF
- Model Evaluation
- Confusion Matrix
- Precision and Recall
- Conclusions

# Visualising Continuous variables

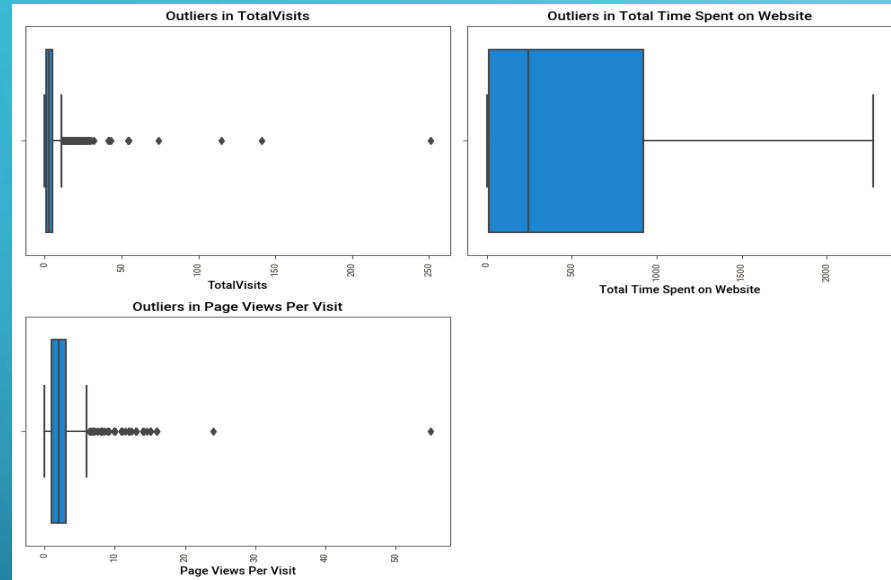


All the three have right skewness

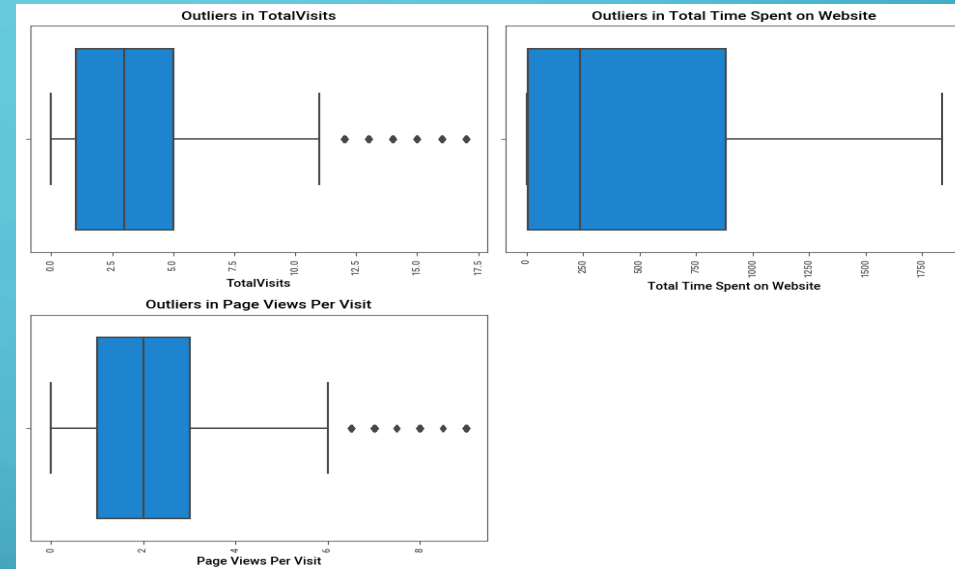
# Visualising Categorical Columns



# Outlier Analysis



Before Treatment of Outliers



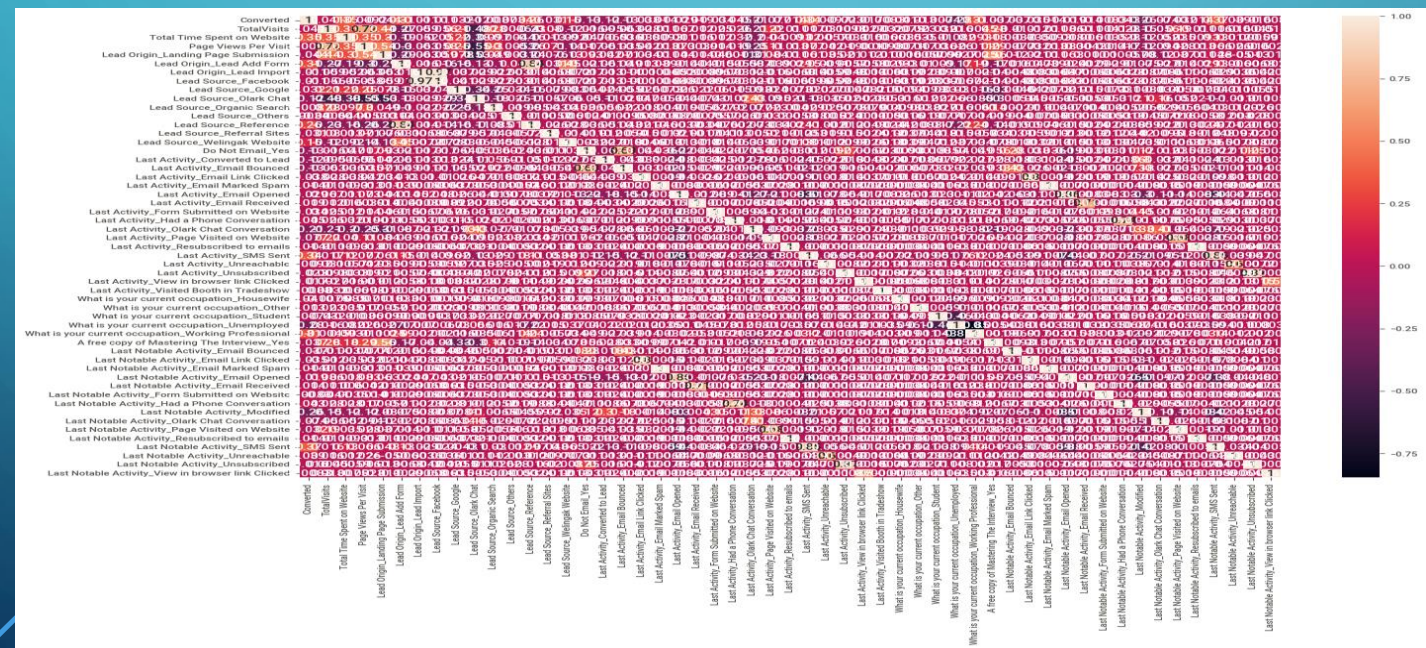
After Treatment of Outliers



# Correlation Analysis

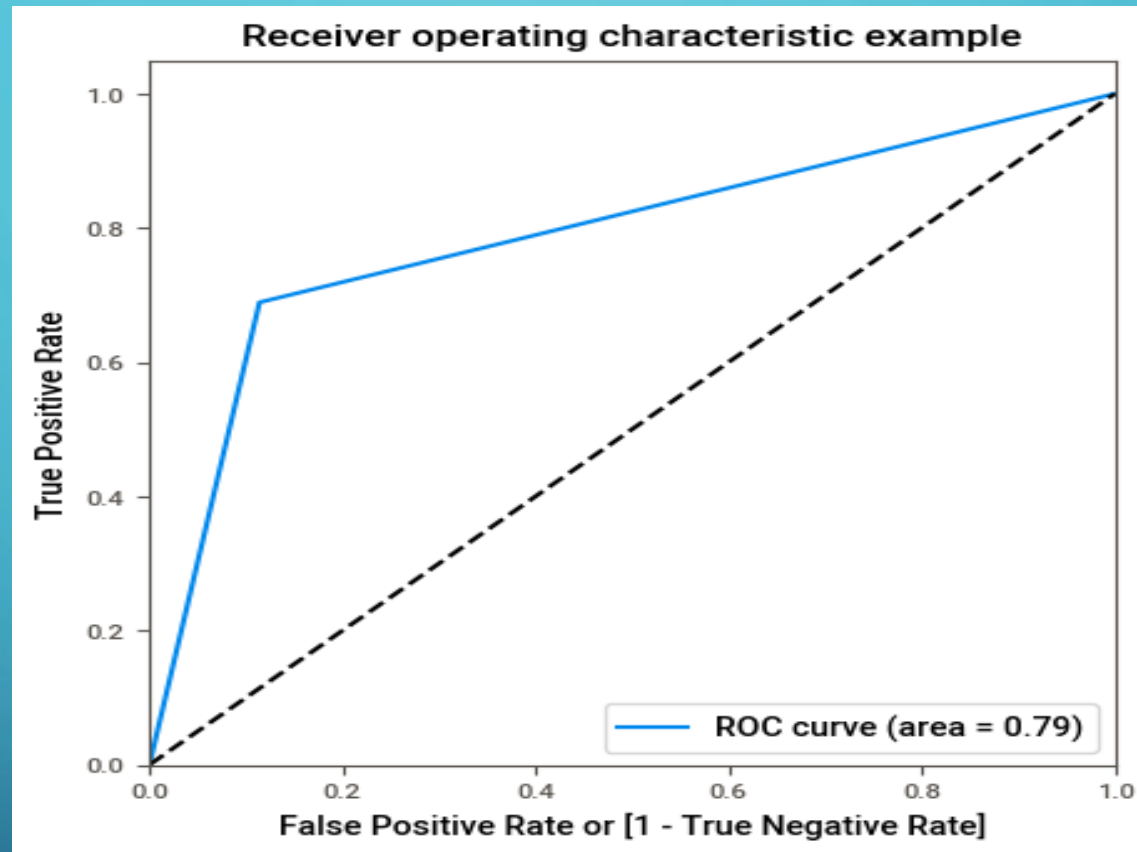


Total Time Spent on Website and Total Visits has good significance and corelated

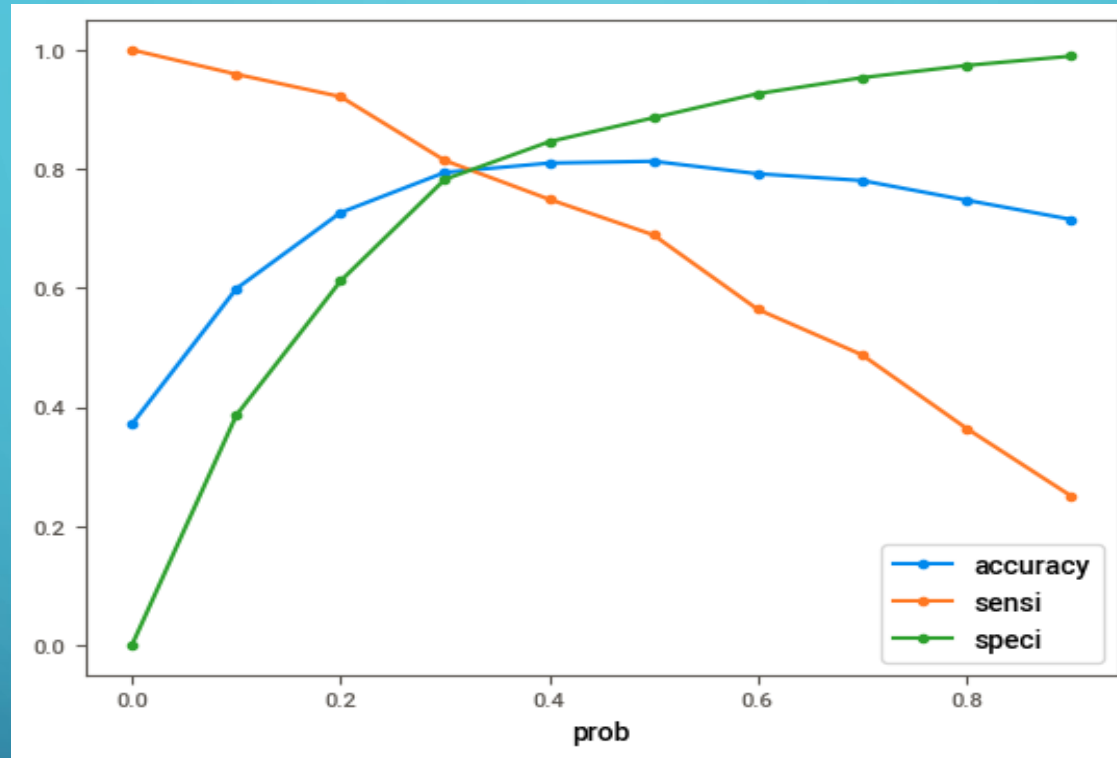




# ROC Curve



# Accuracy, Sensitivity and Specificity



0.3 is the optimal cut-off probability



# Summary

## Conclusions :

*The company has to set sales teams to focus on the below features to convert maximum leads.*

❖ Lead Origin_Lead Add Form	3.985567
❖ Total Time Spent on Website	3.850192
❖ What is your current occupation_Working Professional	2.901561
❖ Last Activity_Had a Phone Conversation	2.259259
❖ Lead Source_Welingak Website	2.084406
❖ TotalVisits	1.706243
❖ Lead Source_Olark Chat	1.482175
❖ Last Activity_SMS Sent	1.377043
❖ Page Views Per Visit	-0.885033
❖ Do Not Email_Yes	-1.388751
❖ Last Activity_Olark Chat Conversation	-1.566924

The background is a blue gradient with decorative white circuit-like lines in the corners. The lines consist of straight segments and small circles, resembling a stylized electronic circuit.

THANK YOU