

The background of the image is a close-up, slightly blurred photograph of a spiral-bound notebook. The notebook has white pages with horizontal lines. A silver spiral binding is visible on the left side. A pencil with a white eraser and a dark body lies diagonally across the middle of the notebook. The text "Company X" is overlaid in a large, white, serif font, centered horizontally. Below it, the tagline "Making learning interesting" is written in a smaller, white, serif font. The overall lighting is soft and even, creating a professional and educational feel.

Company X

Making learning interesting

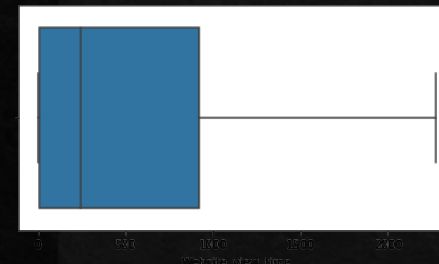
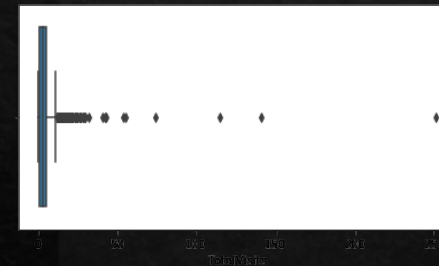
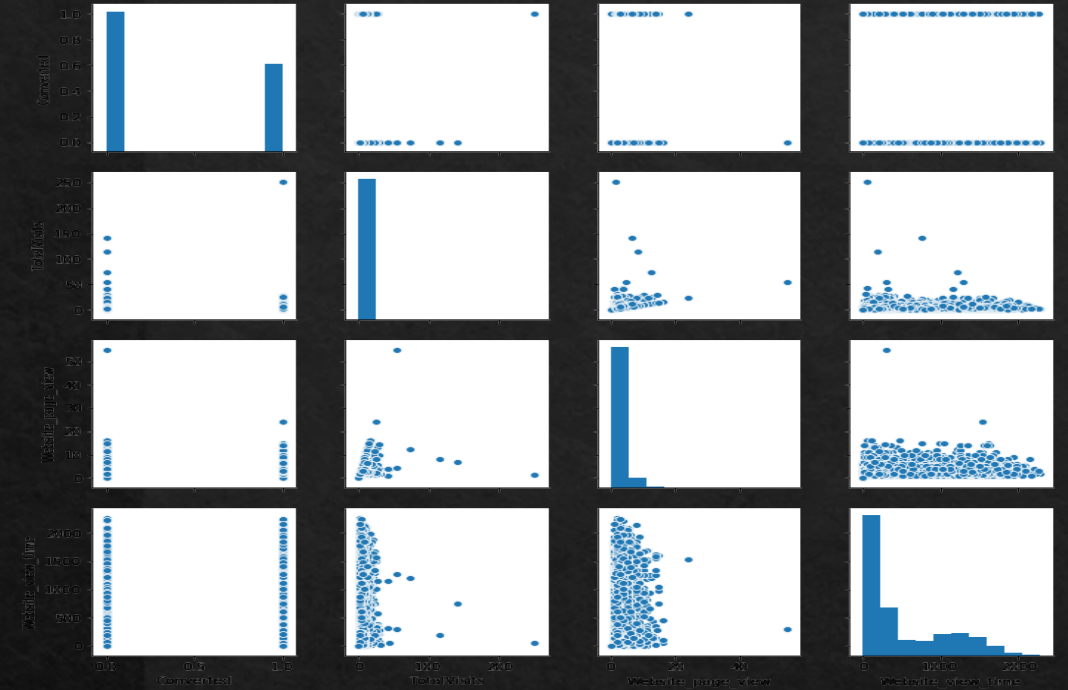
Problem Statement:
To Identify Hot Leads

Analysis Approach

EDA	Drop columns with more than 40% Nulls Consider categorical variables which has more variability
Outlier Treatment	Delete records which has null values for most of the columns Cap 1% of data for numerical values
Data Preparation	Create dummy variables, Split the data into Test and Train. Finally apply feature scaling
Logistic Regression	Drop variables with high correlation values, high VIF. Find optimal cut-off point using ROC curve

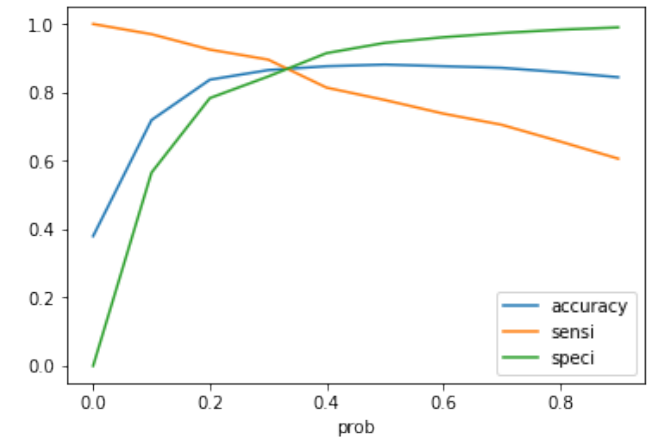
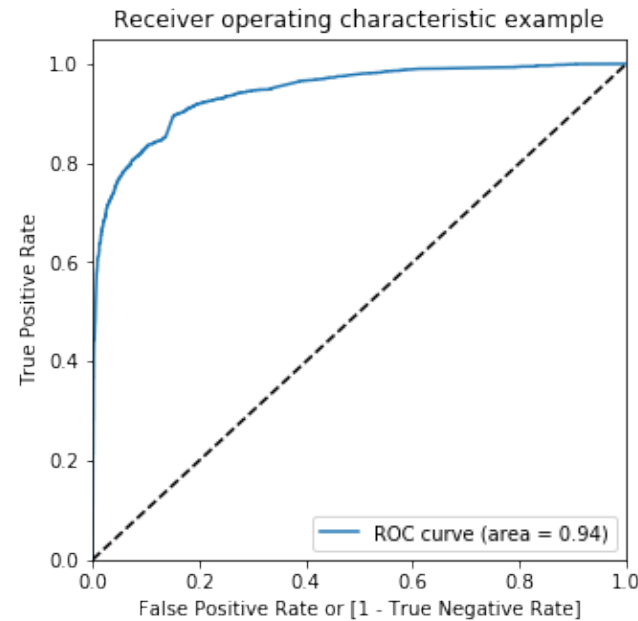
Data Preparation

- ◇ Instead of Dropping Country column we can group these into Asian and Non-Asian countries
- ◇ We can approach several categorical columns in the same way, so analysis of these groups becomes easy.
- ◇ In this way we can avoid dropping these because of variability
- ◇ Also, since we have few outliers, we can cap them



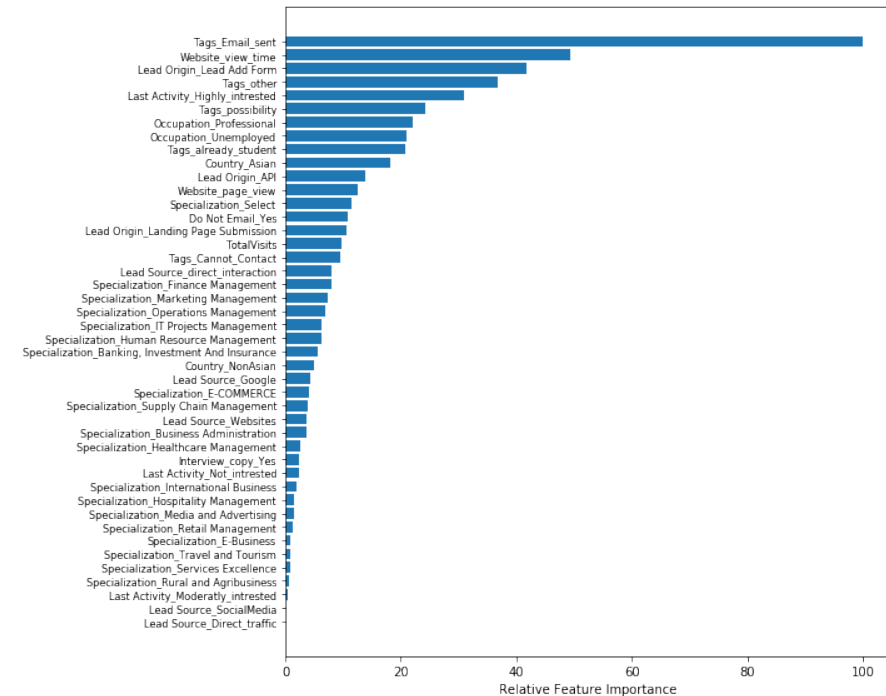
Logistic Regression

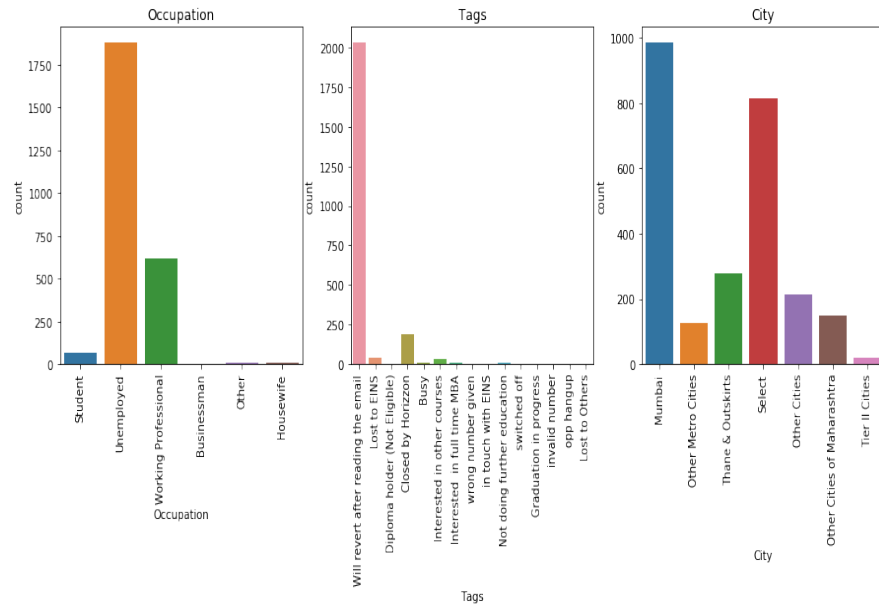
- ◇ ROC shows curve area is 0.94 and away from diagonal saying the model is good
- ◇ Also, looking at these 3 metrics we can say 0.32 as optimal cut-off point



Important factors to consider

- ◇ As we can see some of the important factors are
 - ◇ Email_Sent
 - ◇ Website_view_time
 - ◇ Lead_Add_form
 - ◇ Professional/Unemployed/Student
 - ◇ Asian Countries





Business Aspects

- ◇ To have high Conversion we need to consider population having below characteristics
- ◇ Unemployed/Working Professional
- ◇ Metropolitan cities
- ◇ Source from Google

