X Education : Lead Scoring Case Study

Submitted By:

T.Karan Sai Reddy

Chitra Morwal

Lead Conversion Process

Problem Statement

- X Education gets lots of leads, but conversion rate is poor

- Company wants to efficiently identify most potential leads which are known as 'Hot Leads'

- Successful identification of Hot Leads will increase the conversion rate as sales team will now focus more on communicating with potential leads rather than making calls to everyone

- We are required to build a model where we need to assign lead score to each of the leads such that customers with higher lead score have a higher conversion chance.

- The CEO, in particular, has given a ballpark of the target lead conversion rate to be around 80%.

Proposed Solution

Selection of Hot Leads

Need to identify hot leads based on their tendency to convert, thus we have smaller section of hot leads to focus more.

Communication with Hot Leads

Sales Team can focus on smaller section of Hot leads to have effective conversation with them

Conversion of Hot Leads

Focus on Hot Leads which are more probable to convert will lead to increase in conversion rate to achieve target of 80%

SOLUTION :
Selection of Hot Leads

- Identifying hot leads accurately is major part & very crucial for our problem solution

- Higher accuracy will lead to higher conversion rate

- We need to **build a model** wherein we need to assign a **lead score** to each of the leads such that the customers with higher lead score have a higher conversion chance and the customers with lower lead score have a lower conversion chance

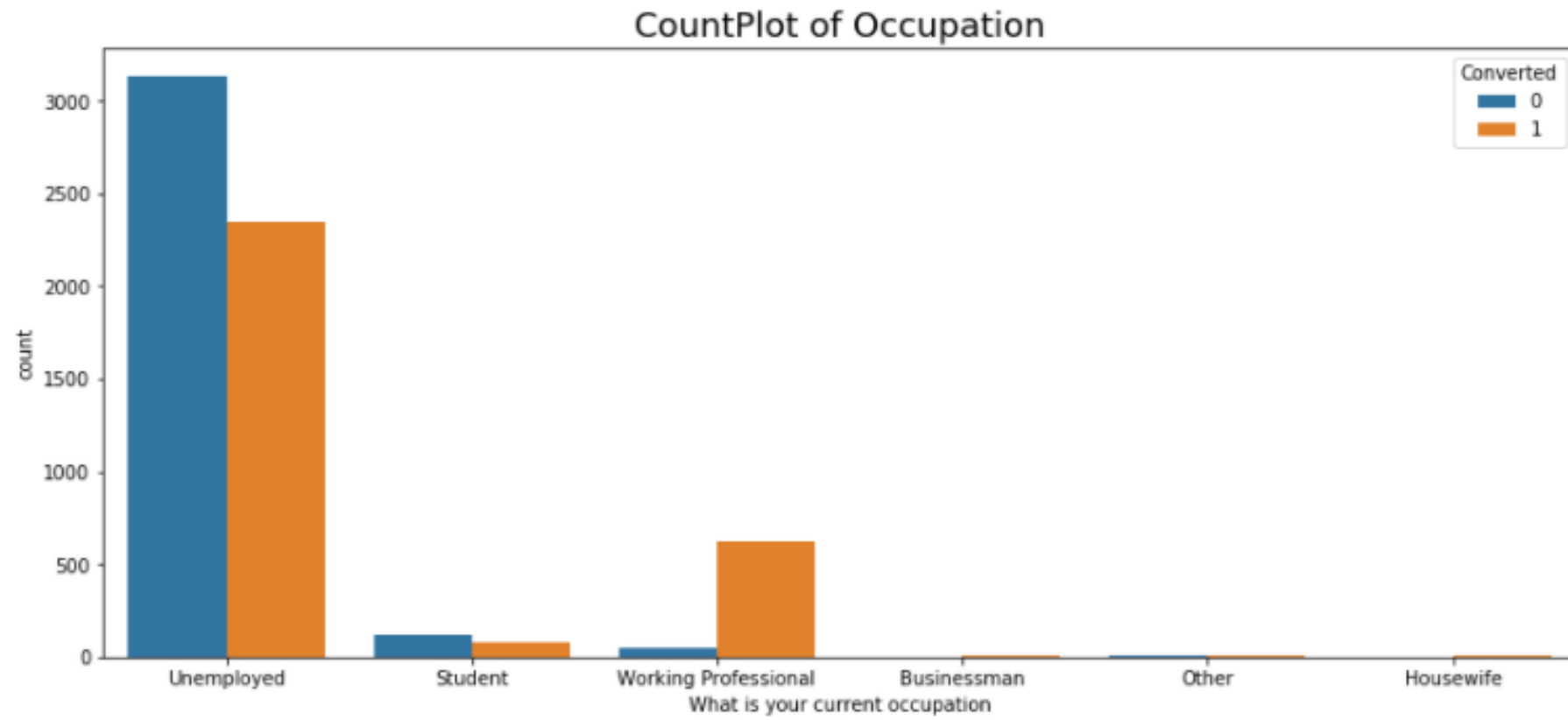- As per CEO, we are given a ballpark of the target lead conversion rate to be around 80%
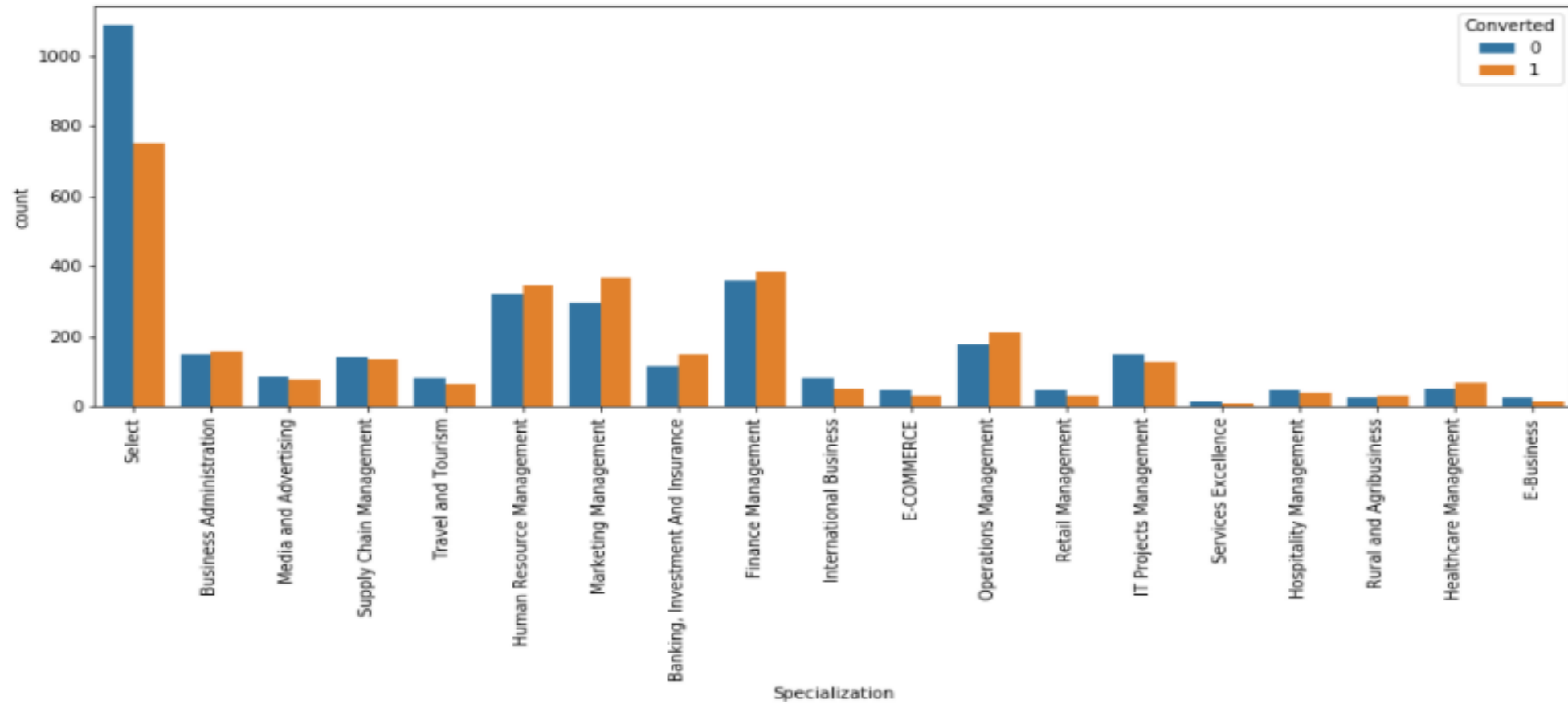
Implementation
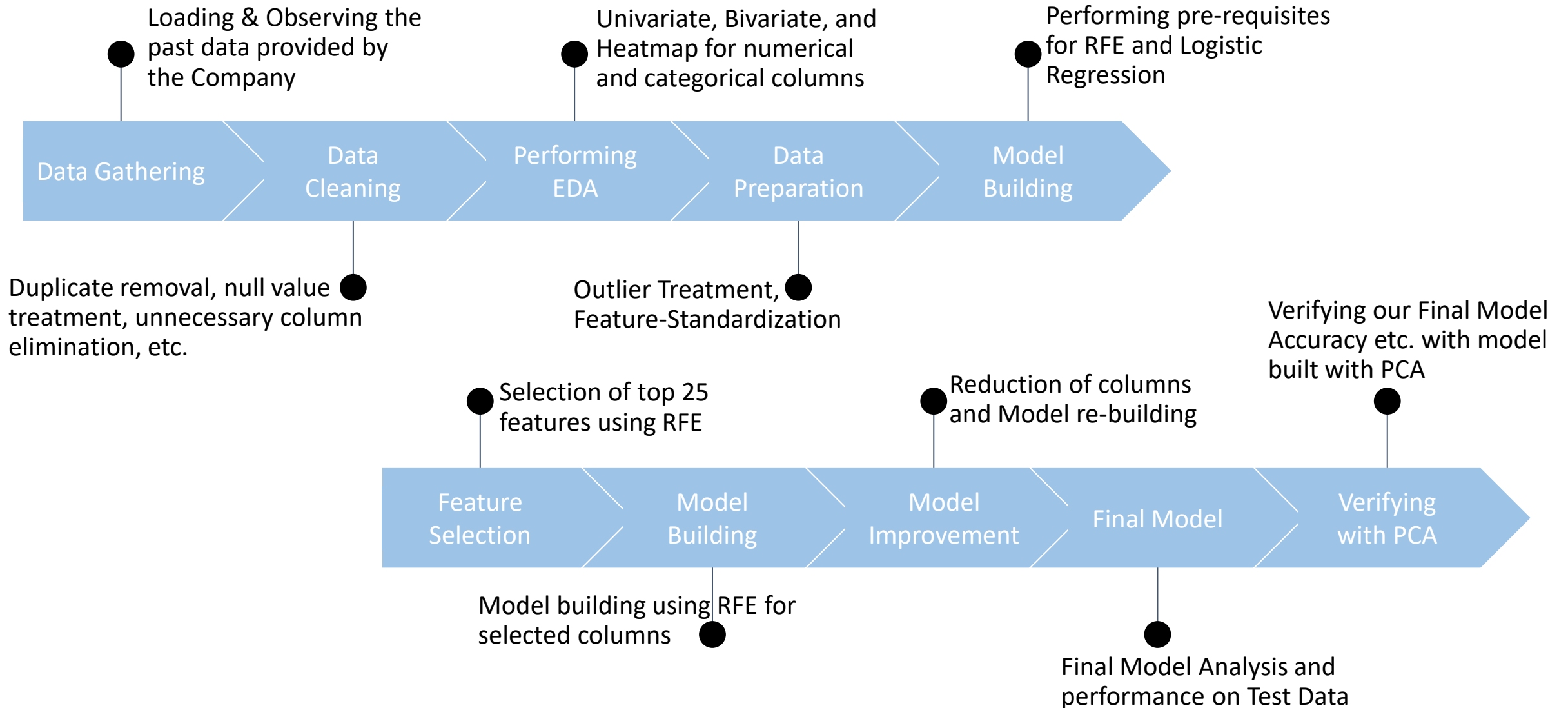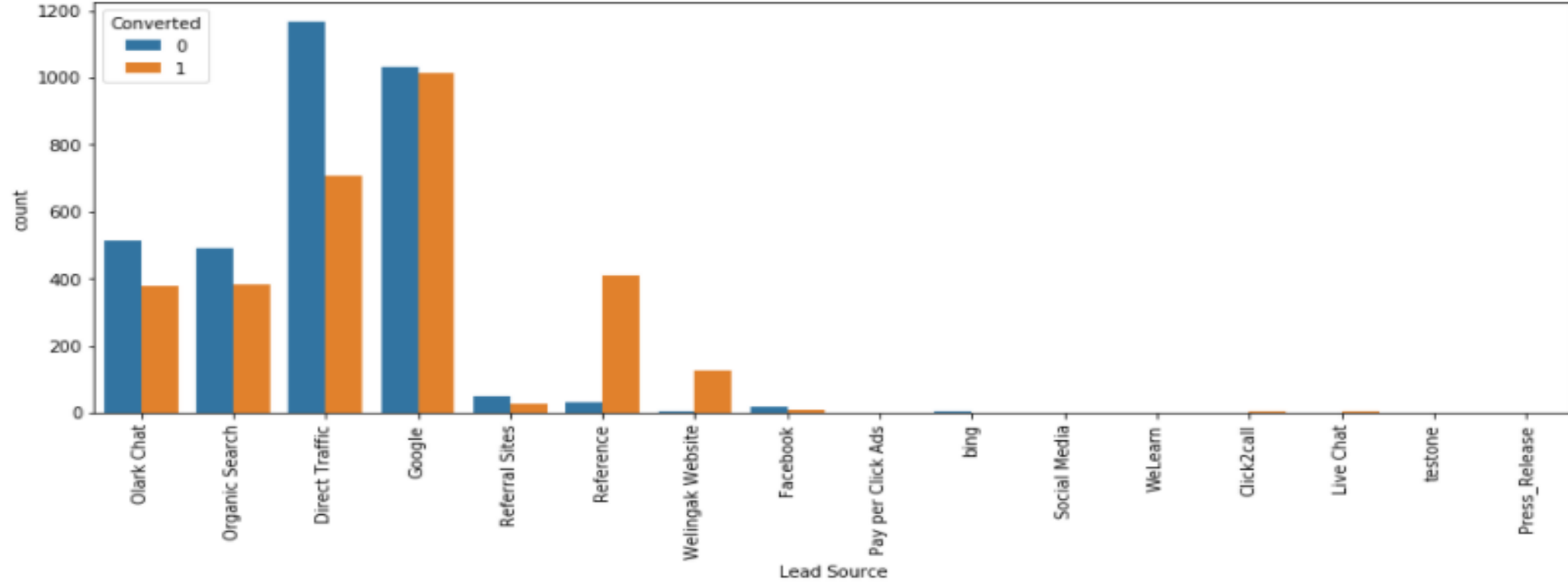
Data Analysis of Categorical variables
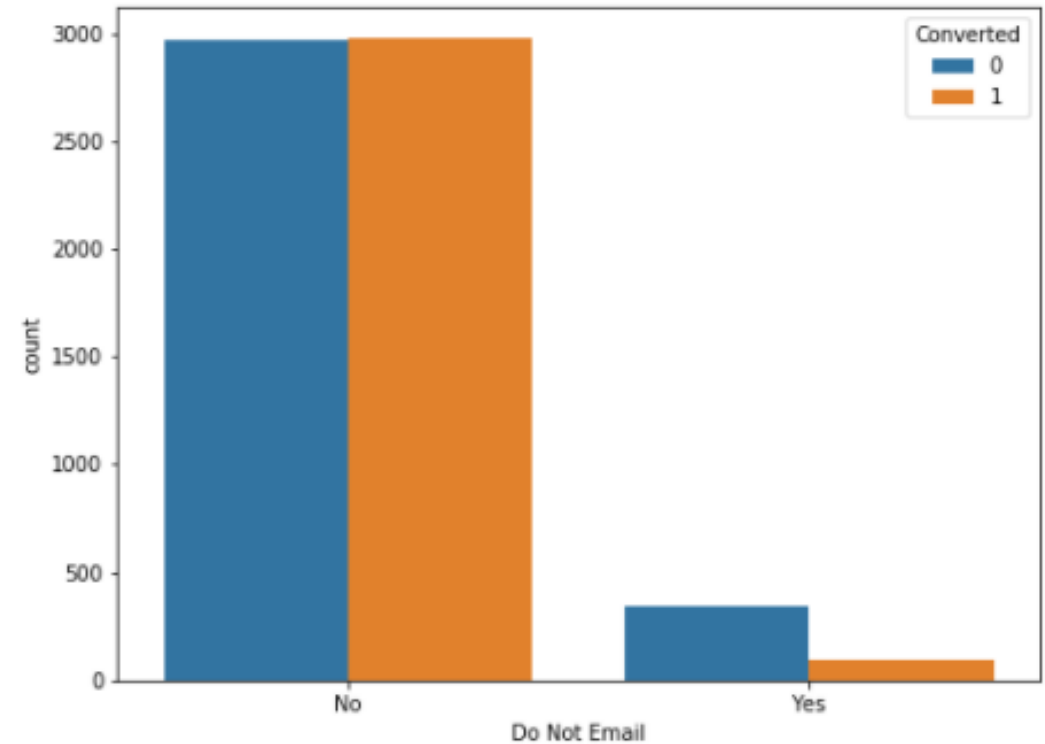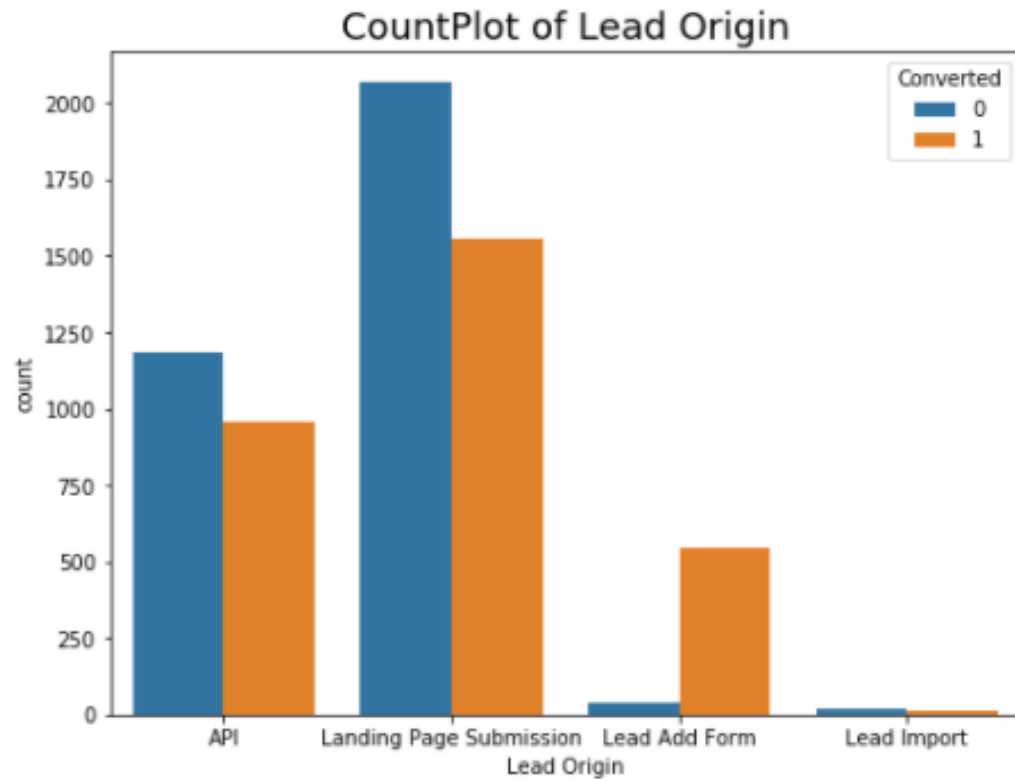
"Landing Page Submission" has higher count of Converted
"Do Not Email" don't have significant impact on conversion

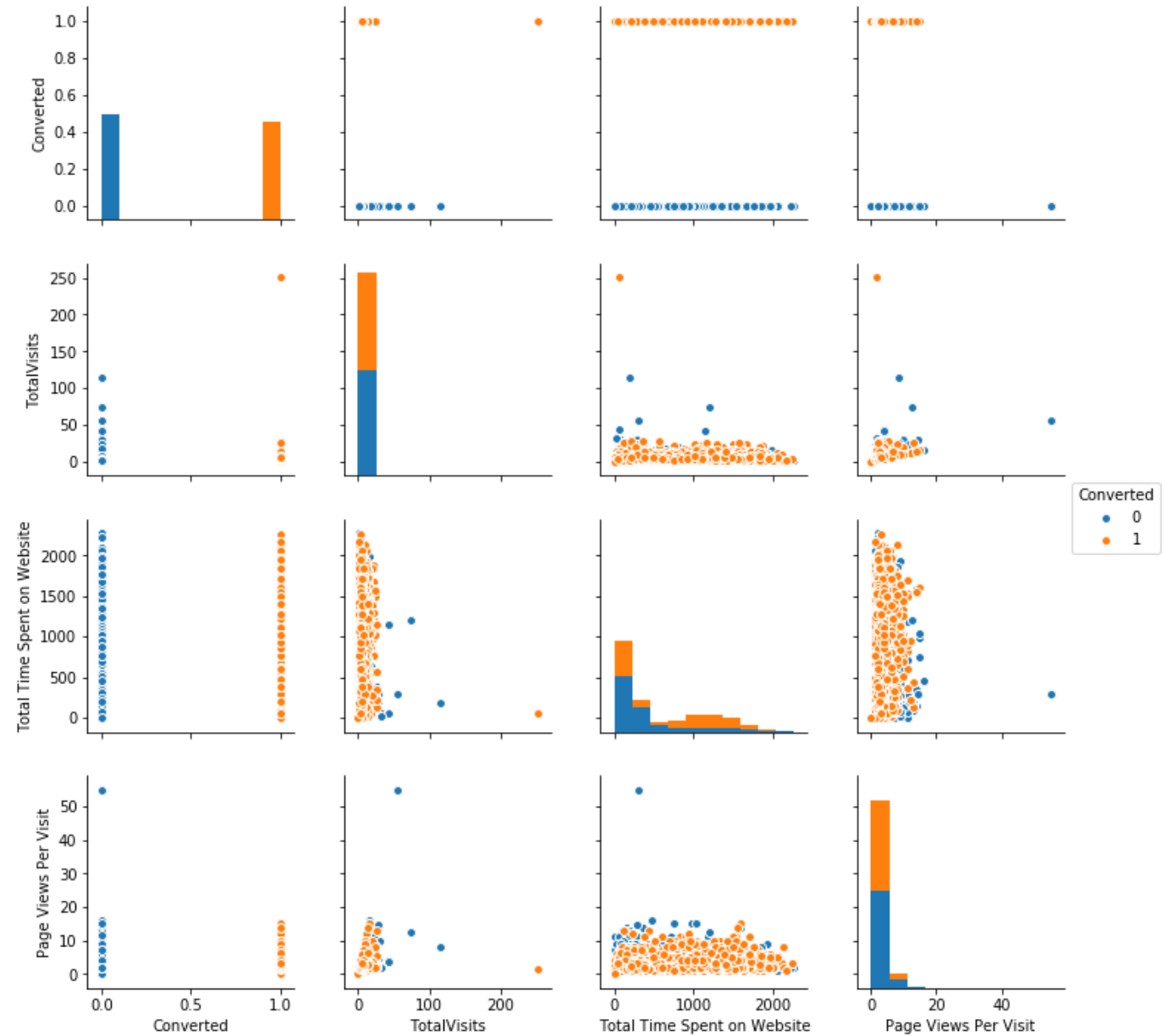"Working Professional" has comparatively higher conversion rate

"Finance Management" , "Marketing Management" , "Human Resource Management" :
Customer from these specialization are more probable to be converted

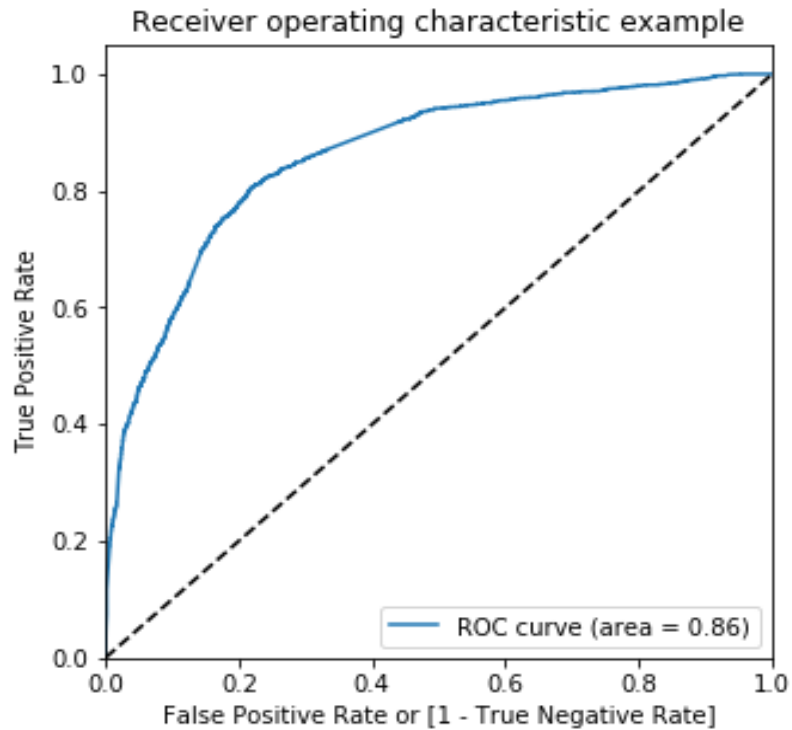"Google" , "Direct Traffic" has higher count of converted

Analysis of Numerical Variables :

Plot depicting correlation of selected Numerical columns
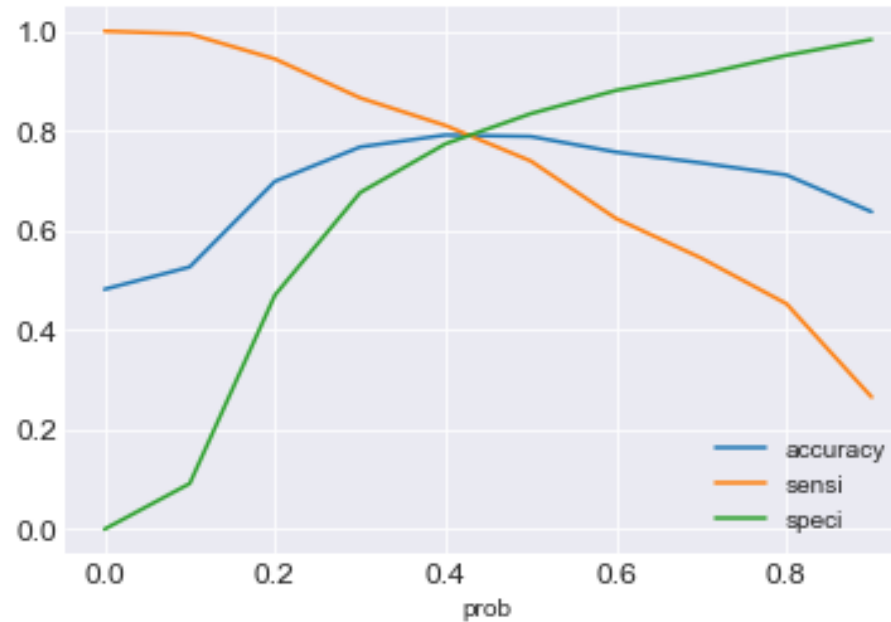
Correlation heatmap of all selected columns including numerical and dummy variables
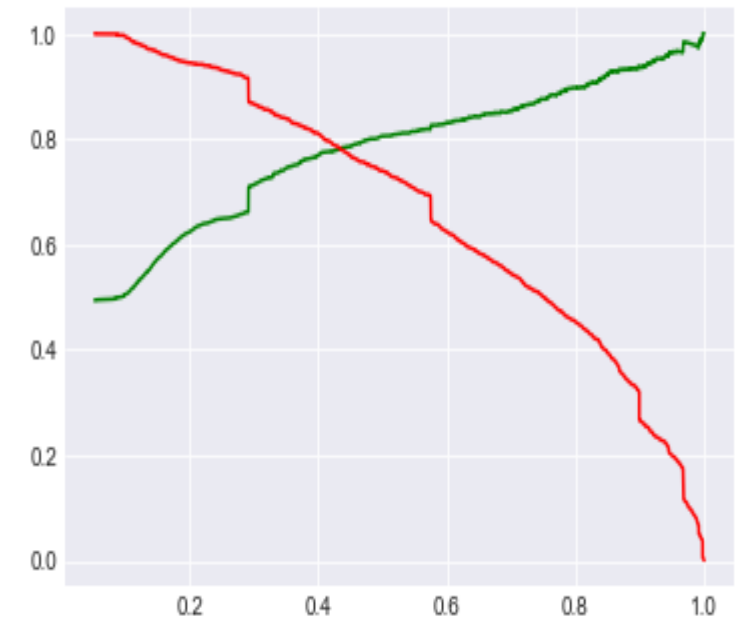
Final Model Parameters

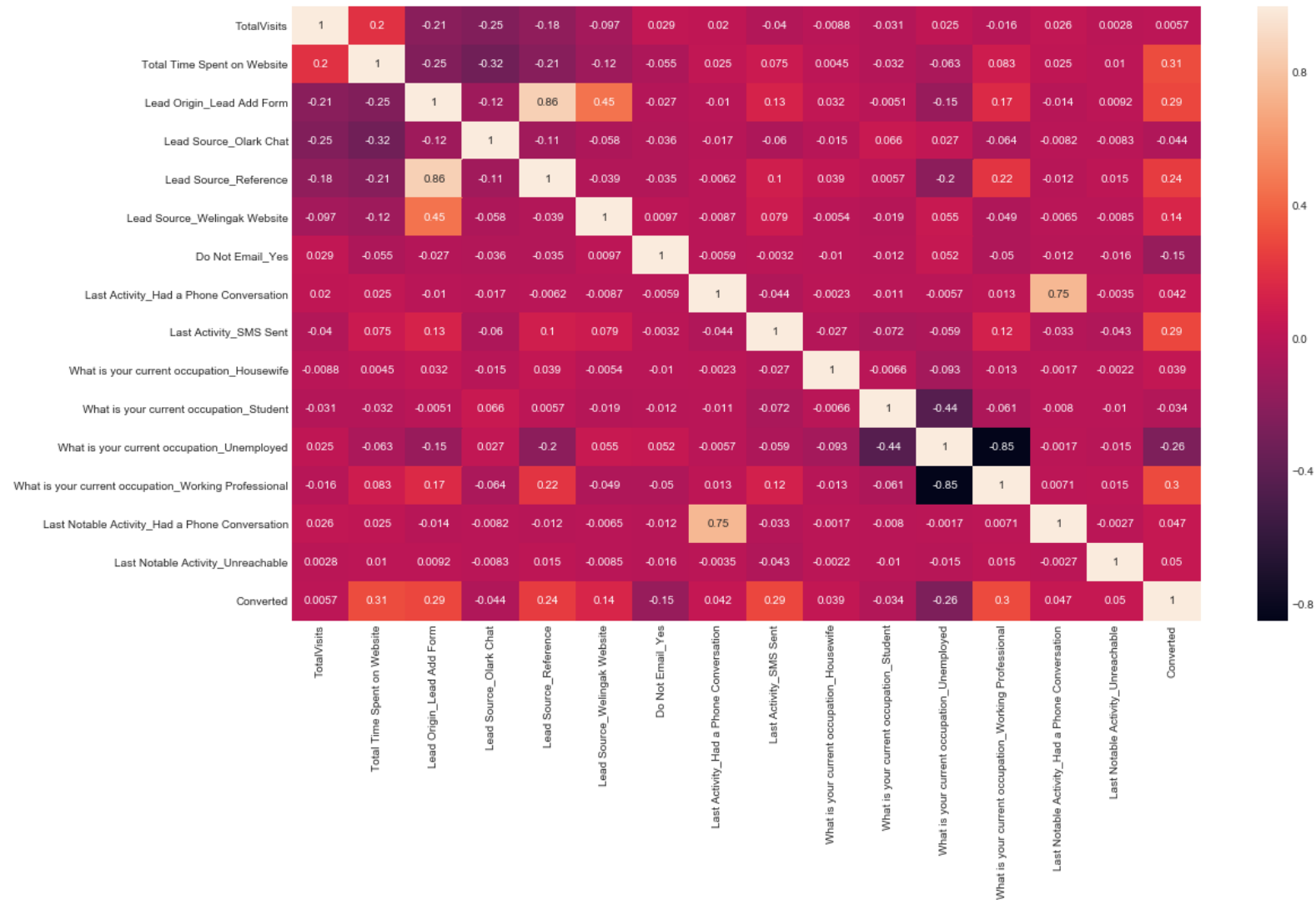**Area under ROC = 0.86**   **Finalizing Cut Off = 0.41**   **Precision Recall Trade Off**

Correlation Heat Map of all selected columns in Final Model

Model Analysis

Performance of our Final Model

- Overall accuracy on Test Set : 78.39%

- Sensitivity of our Logistic Regression Model : 78.82%

- Specificity of our Logistic Regression Model : 78.01%

- Precision : 76.72%

- Recall : 78.82 %

Inferences from Model

Business Insight derived from our Model

- The top variables that contributed towards leads being converted are:
  - Lead Origin_Lead Add Form
  - What is your current occupation_Working Professional
  - Total Time Spent on Website

- The top 3 categorical/dummy variables which should be focused the most in order to increase the probability of lead conversion are:
  - Lead Origin_Lead Add Form
  - Lead Source_Olark Chat
  - Do Not Email

Conclusion [LR Model]

- Our Logistic Regression Model is decent and accurate enough, with 78.39 % Accuracy on Test Set, 78.82 % Sensitivity and 78.01 % Specificity.

- We can vary these parameters by varying the cut-off value and thus predict Hot leads based on scenarios like availability of extra resources and vice-versa.

Conclusion [Recommendation]

- X Education Company needs to focus on following key aspects to improve the overall conversion rate:

- Increase user engagement on their website since this helps in higher conversion

- Increase on sending SMS notifications since this helps in higher conversion

- Get Total visits increased by advertising etc. since this helps in higher conversion

- Improve the Olark Chat service since this is affecting the conversion negatively

Thank You !!