LEAD SCORING CASE STUDY

Group Members
1. Akshita Gupta
2. Kaivalya Kulkarni
3. Ashish Yadav

Problem Statement

- X Education sells online courses to industry professionals.
- X Education gets a lot of leads, its lead conversion rate is very poor. For example, if, say,
- they acquire 100 leads in a day, only about 30 of them are converted.
- To make this process more efficient, the company wishes to identify the most potential leads, also known as 'Hot Leads'.
- If they successfully identify this set of leads, the lead conversion rate should go up as the sales team will now be focusing more on communicating with the potential leads rather than making calls to everyone.
- X education wants to know most promising leads.
- For that they want to build a Model which identifies the hot leads.
- Deployment of the model for the future use.

STEPS INVOLVED

Data cleaning :

- Handling Duplicate Data.
- Handling NA values and missing values.
- Dropping columns, if it contains large amount of missing values and not useful for the analysis.
- Imputation of the values, if necessary.

Exploratory Data Analysis (EDA):

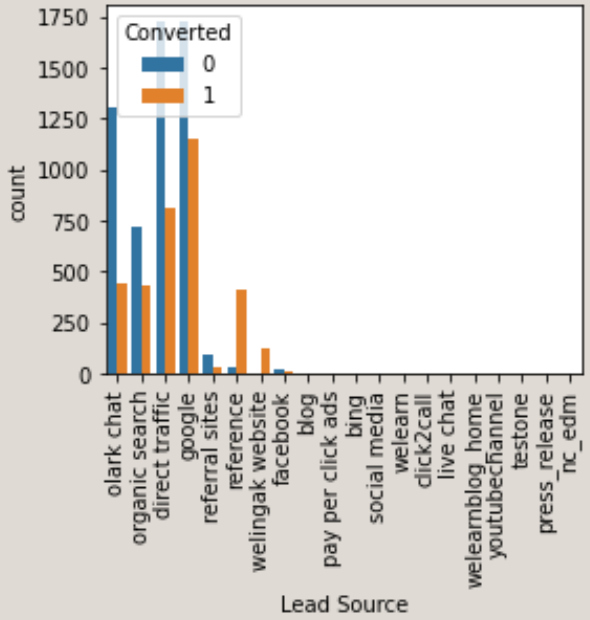
- Univariate Data Analysis: Boxplots
- Bivariate Data Analysis: correlation coefficients and pattern between the variables etc. using heatmap.

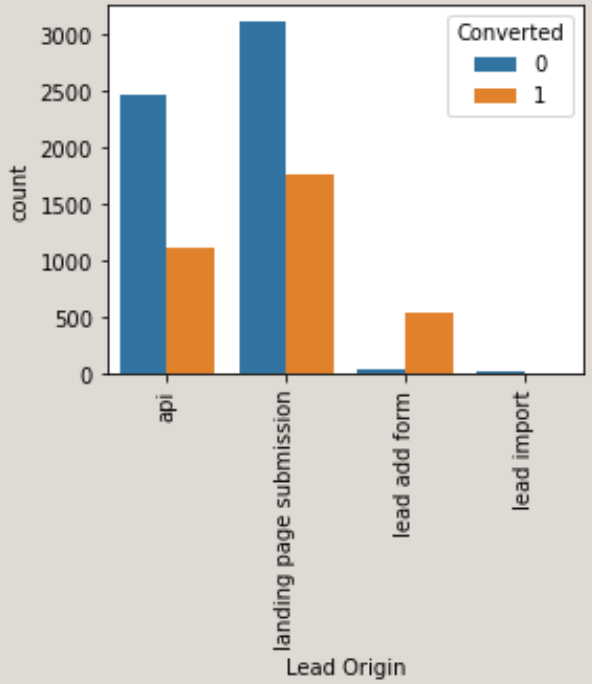
Data Preparation: Feature Scaling & Dummy Variables and encoding of the data.

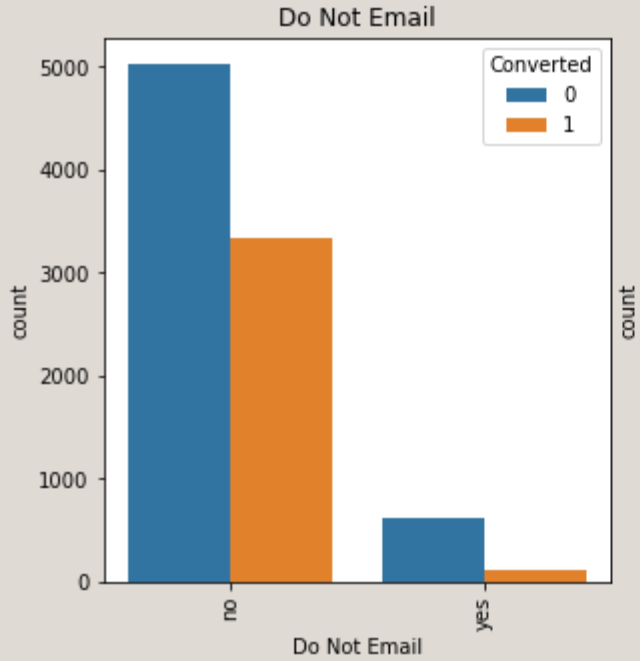
Logistic Regression for model building

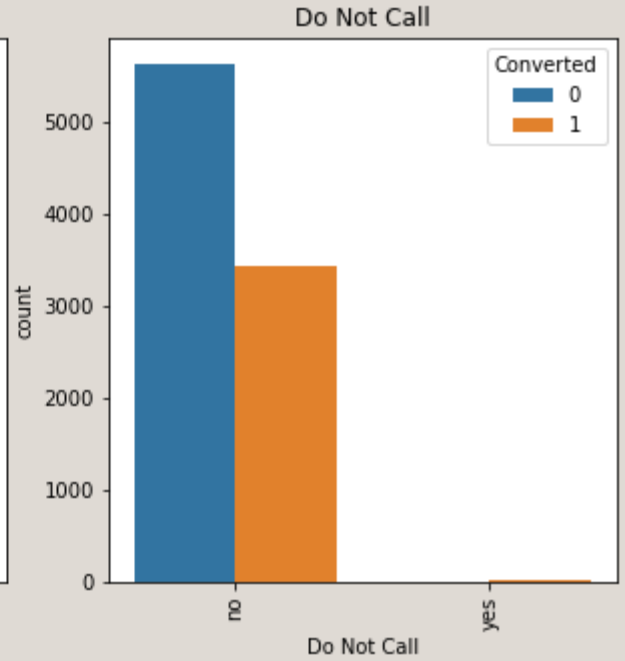
Validation of the model.

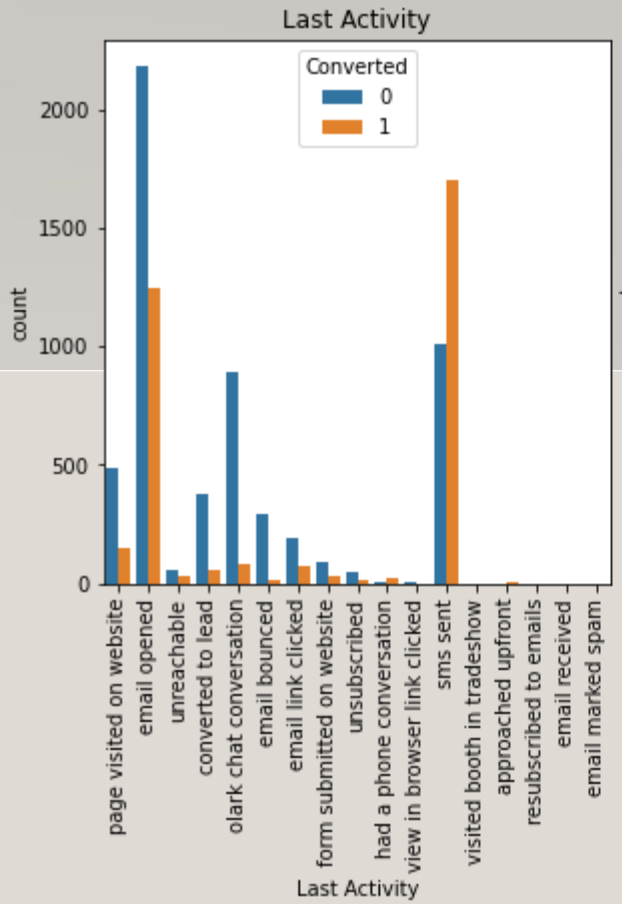
Exploratory Data Analysis (EDA):

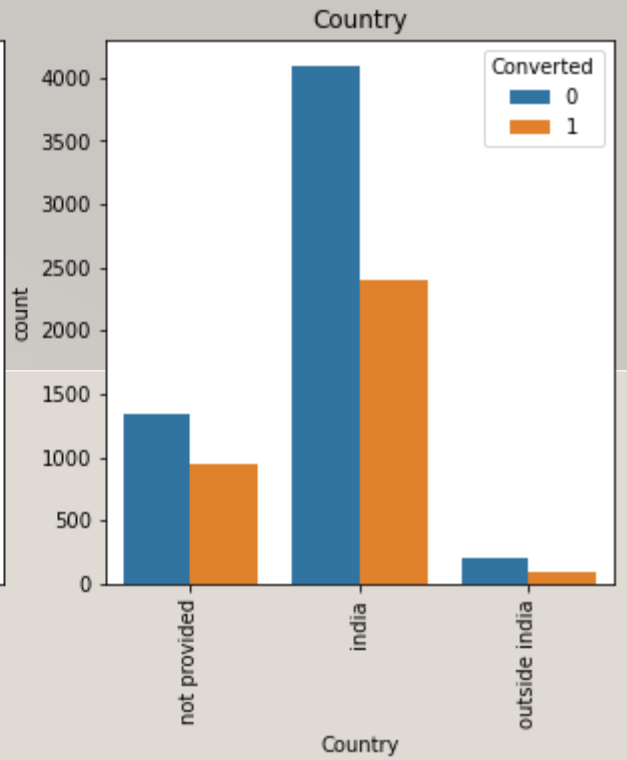


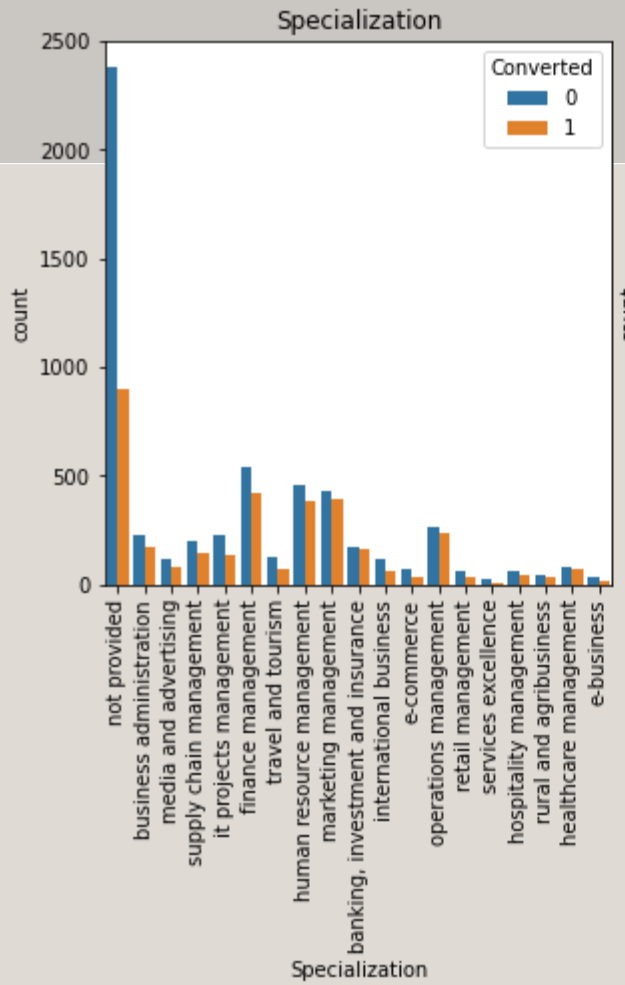


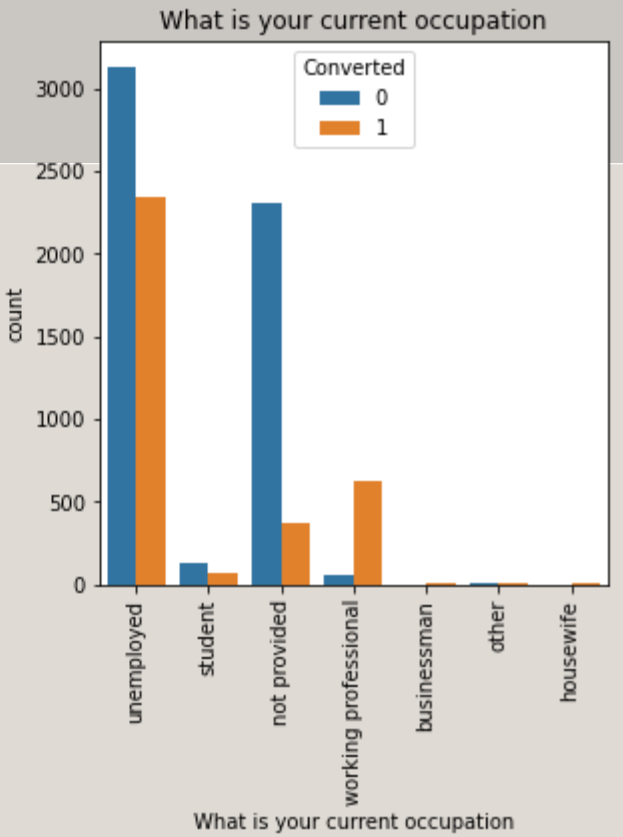


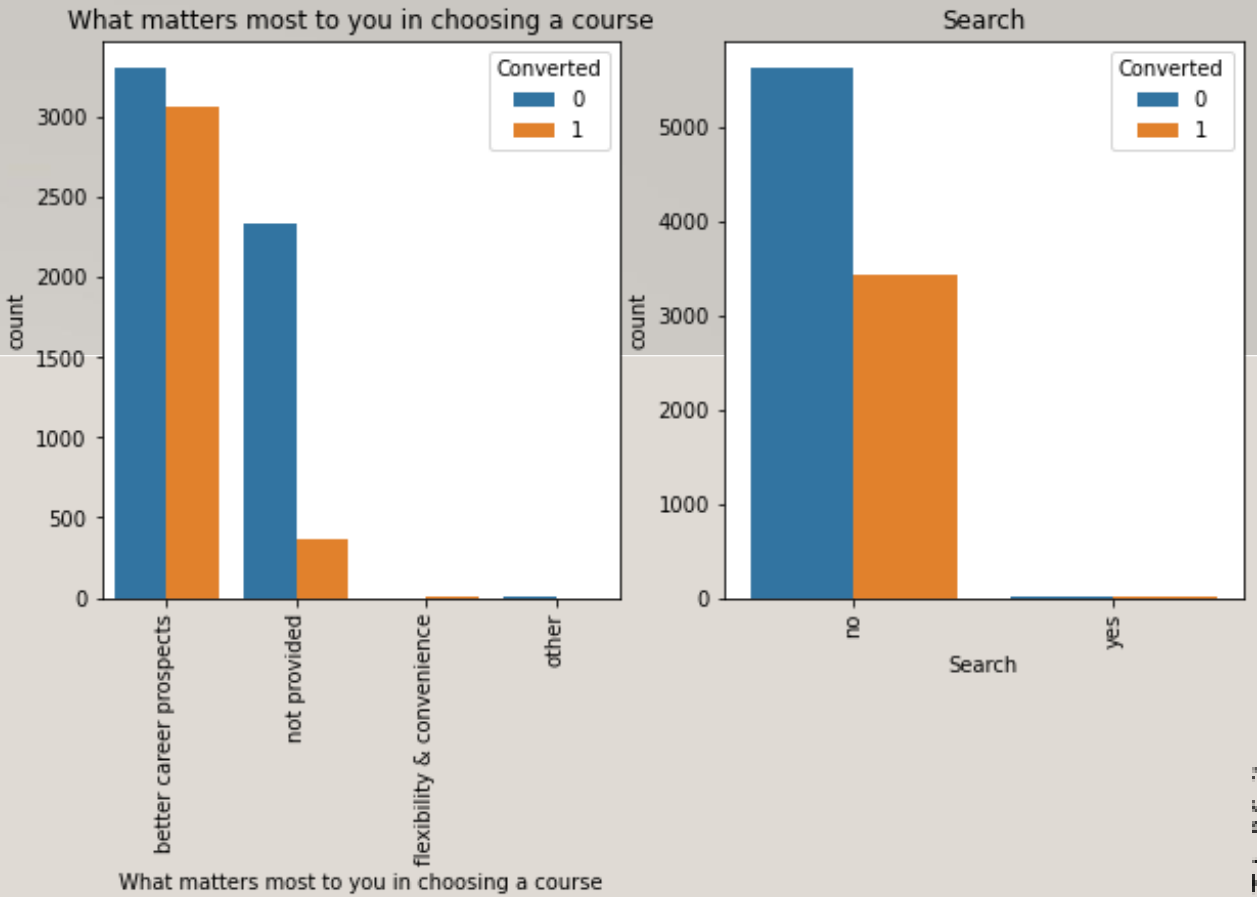



Last Activity
Converted
0
1
count
page visited on website
email opened
unreachable
converted to lead
olark chat conversation
email bounced
email link clicked
form submitted on website
unsubscribed
had a phone conversation
view in browser link clicked
sms sent
visited booth in tradeshow
approached upfront
resubscribed to emails
email received
email marked spam
Last Activity

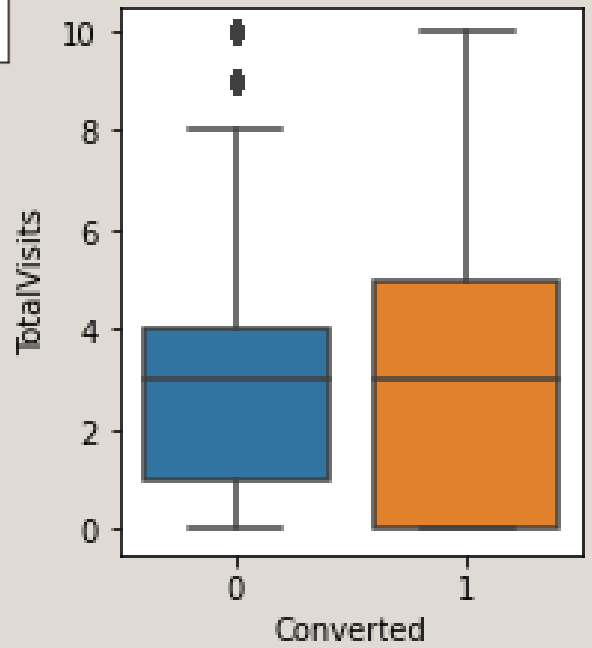

Country
Converted
0
1
count
not provided
india
outside india
Country

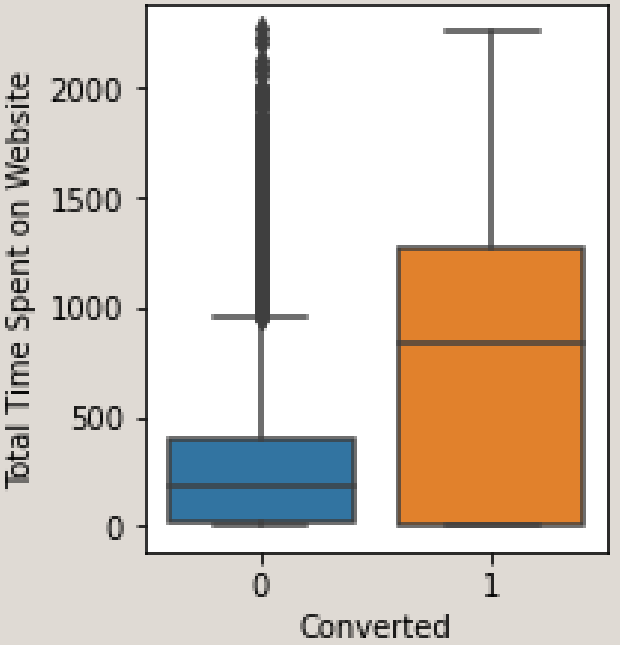

Specialization
Converted
0
1
count
not provided
business administration
media and advertising
supply chain management
it projects management
finance management
travel and tourism
human resource management
marketing management
banking, investment and insurance
international business
e-commerce
operations management
retail management
services excellence
hospitality management
rural and agribusiness
healthcare management
e-business
Specialization


What is your current occupation
Converted
0
1
count
unemployed
student
not provided
working professional
businessman
other
housewife
What is your current occupation

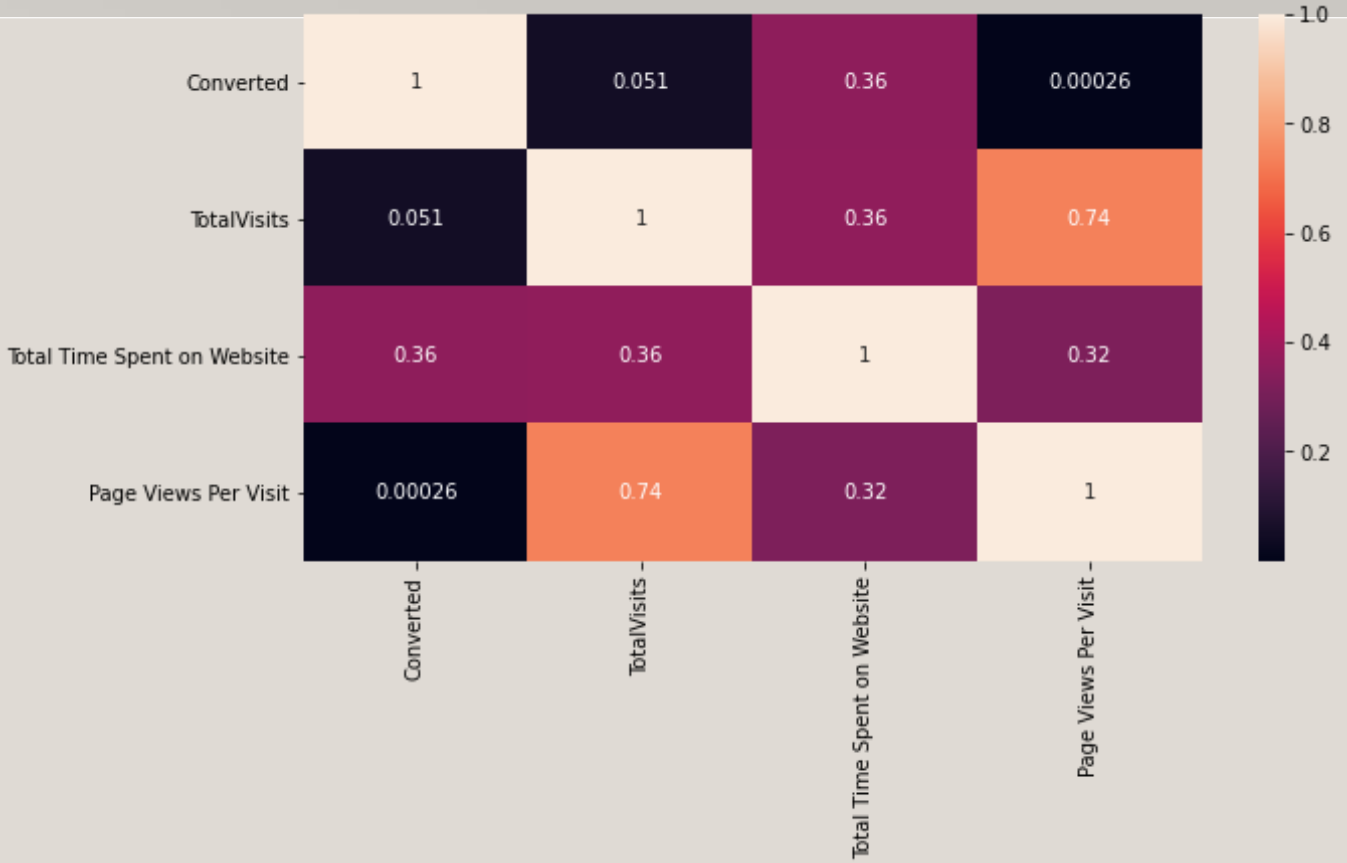

What matters most to you in choosing a course
Converted
0
1
count
3000
2500
2000
1500
1000
500
0
better career prospects
not provided
flexibility & convenience
other
What matters most to you in choosing a course

Search
Converted
0
1
count
5000
4000
3000
2000
1000
0
no
yes
Search

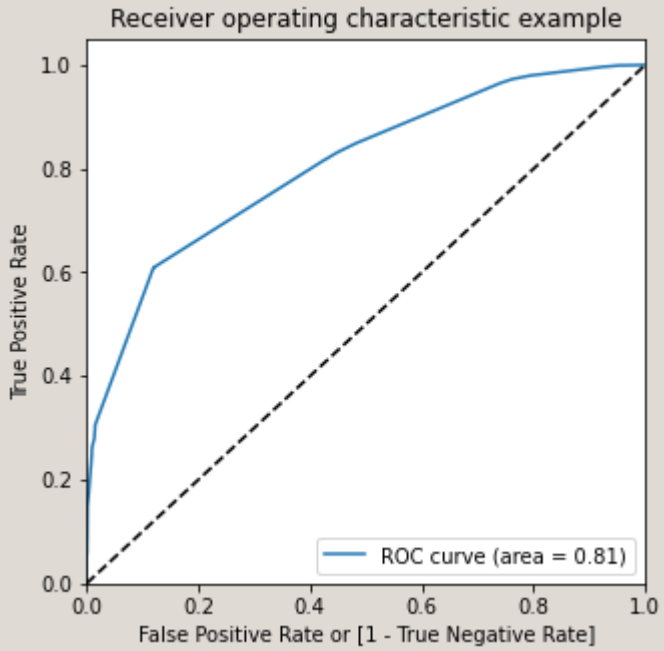

TotalVisits
10
8
6
4
2
0
0
1
Converted

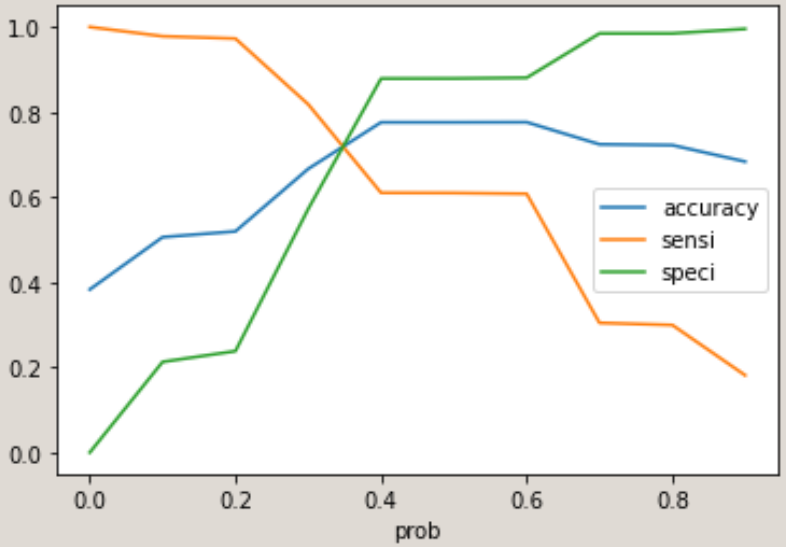

Total Time Spent on Website
2000
1500
1000
500
0
0
1
Converted

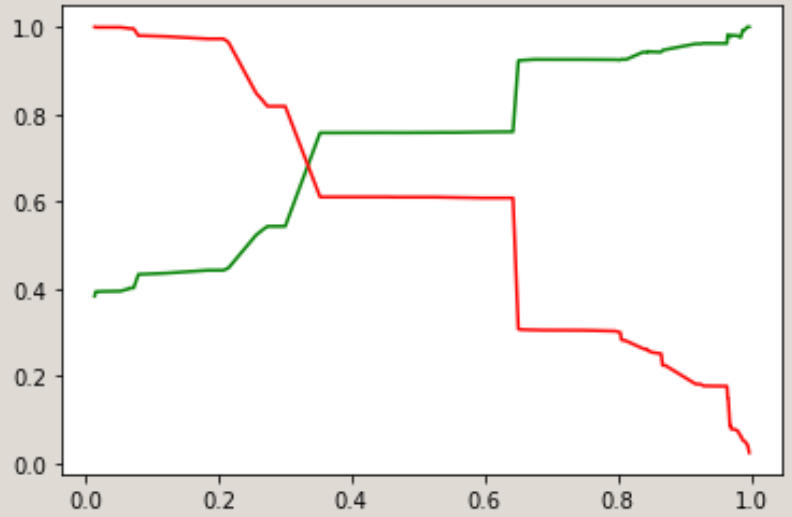
Correlation Between numerical columns (Heat Map)



Linear Regression Final Model







MODEL ANALYSIS

Overall accuracy on Test set: 0.7759
Sensitivity of our logistic regression model: 0.609
Specificity of our logistic regression model: 0.878

LINEAR REGRESSION MODEL

Our Logistic Regression Model is decent and accurate enough, when compared to the model derived using PCA, with 77.8 % Accuracy on Test Set, 56.8 % Sensitivity and 88.6 % 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

- 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