Lead Score Case Study

Problem Statement :

X Education sells online courses to industry professionals. The company markets its courses on several websites and search engines like Google.

Once these people land on the website, they might browse the courses or fill up a form for the course or watch some videos. When these people fill up a form providing their email address or phone number, they are classified to be a lead. Moreover, the company also gets leads through past referrals.

Once these leads are acquired, employees from the sales team start making calls, writing emails, etc. Through this process, some of the leads get converted while most do not. The typical lead conversion rate at X education is around 30%.

Business Goal:

X Education needs help in selecting the most promising leads, i.e. the leads that are most likely to convert into paying customers.

The company needs a model wherein you a lead score is assigned 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.

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

Solution Approach

- Obtain the information for analysis.
- Make sure the data is clean and ready for exploratory data analysis.
- Feature Scaling - dividing the data into two sets: test and train
- Calculating the Lead Score and constructing a logistic regression model.
- Assessing the model using several metrics, such as Precision and Recall or Specificity and Sensitivity.
- Using the optimal model in test data according to the metrics for sensitivity and specificity
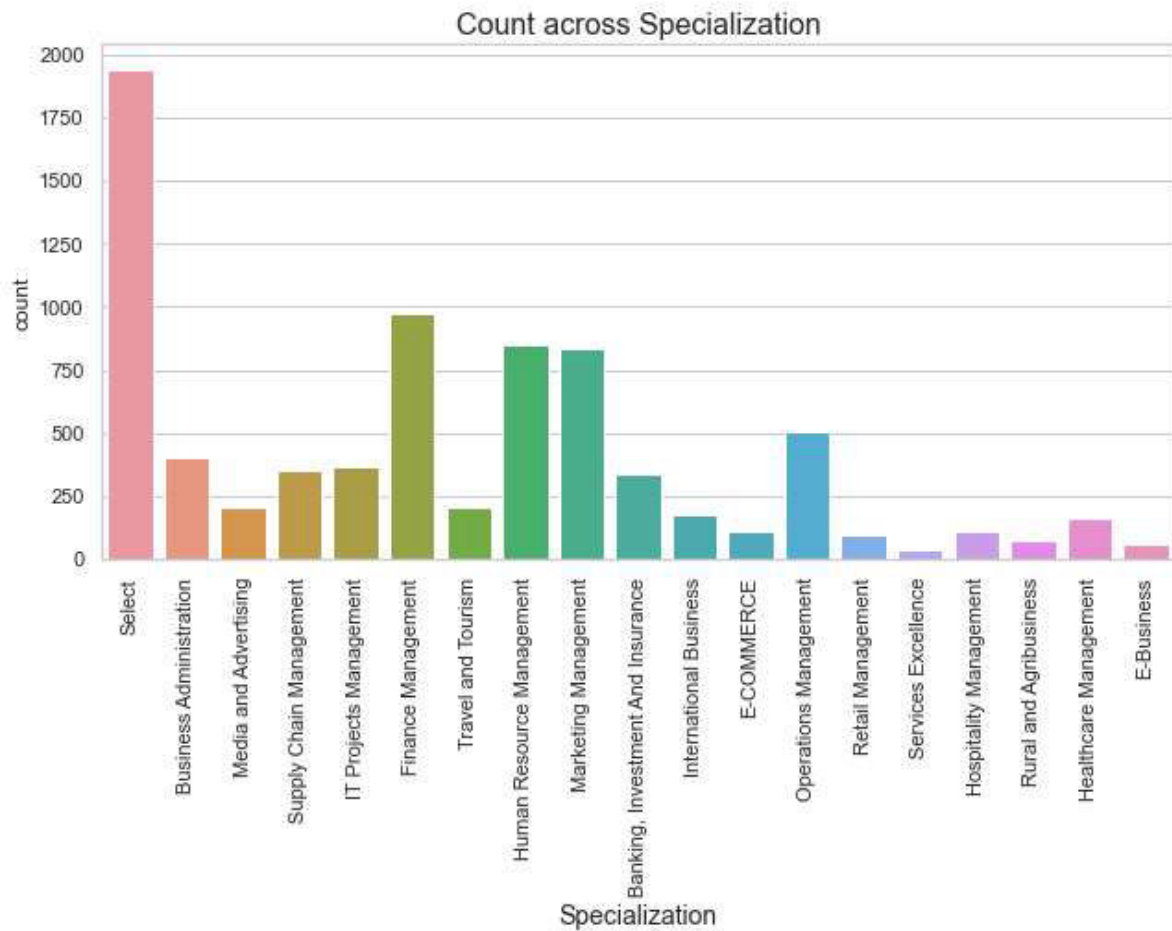
EDA – Data Cleaning

There are few columns in which there is a level called 'Select' which is taking care
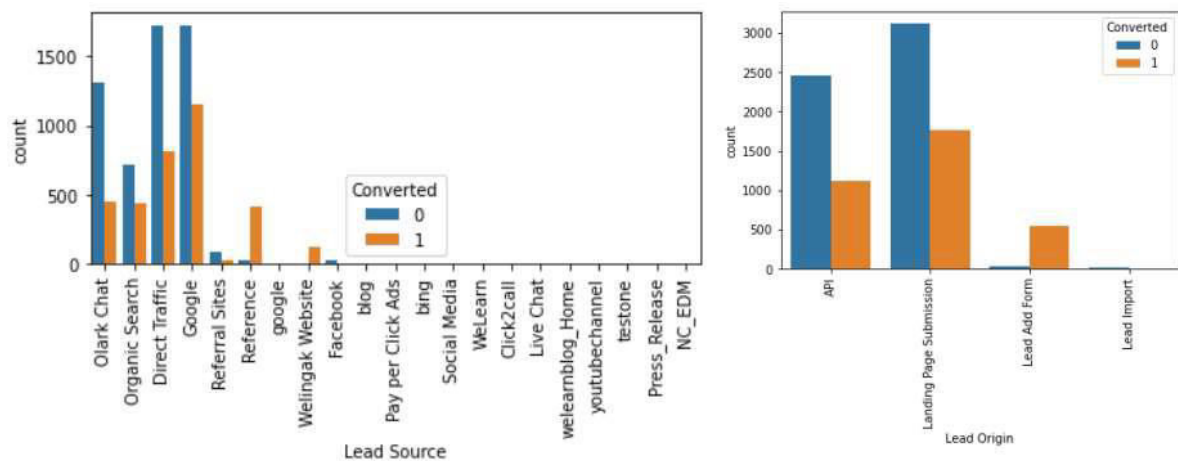
Count across Specialization
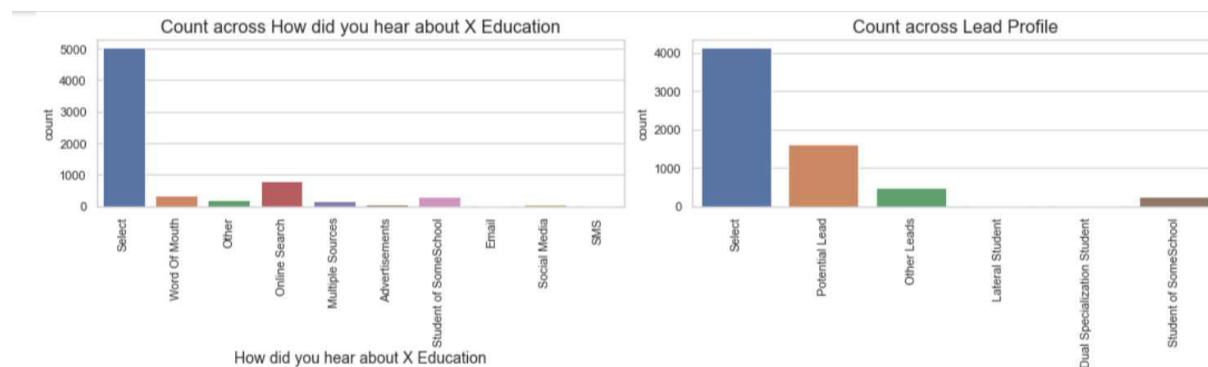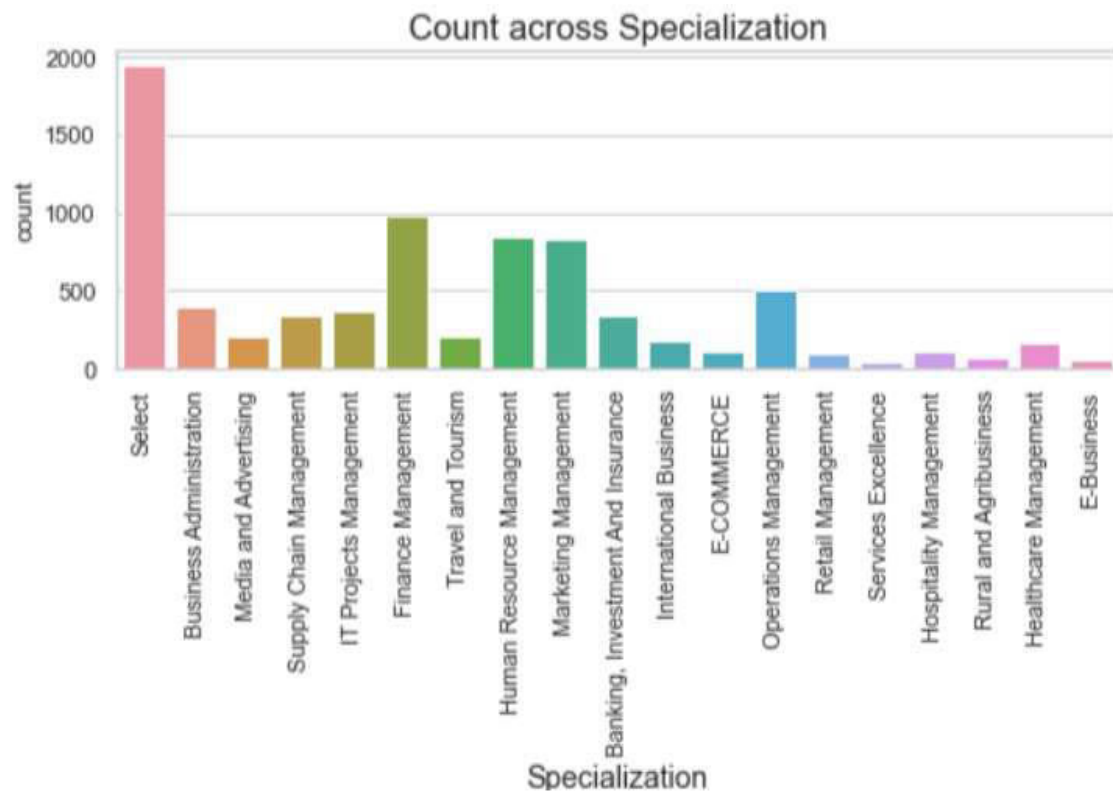
Specialization

Leads from HR, Finance & Marketing management specializations are high probablitiy to convert
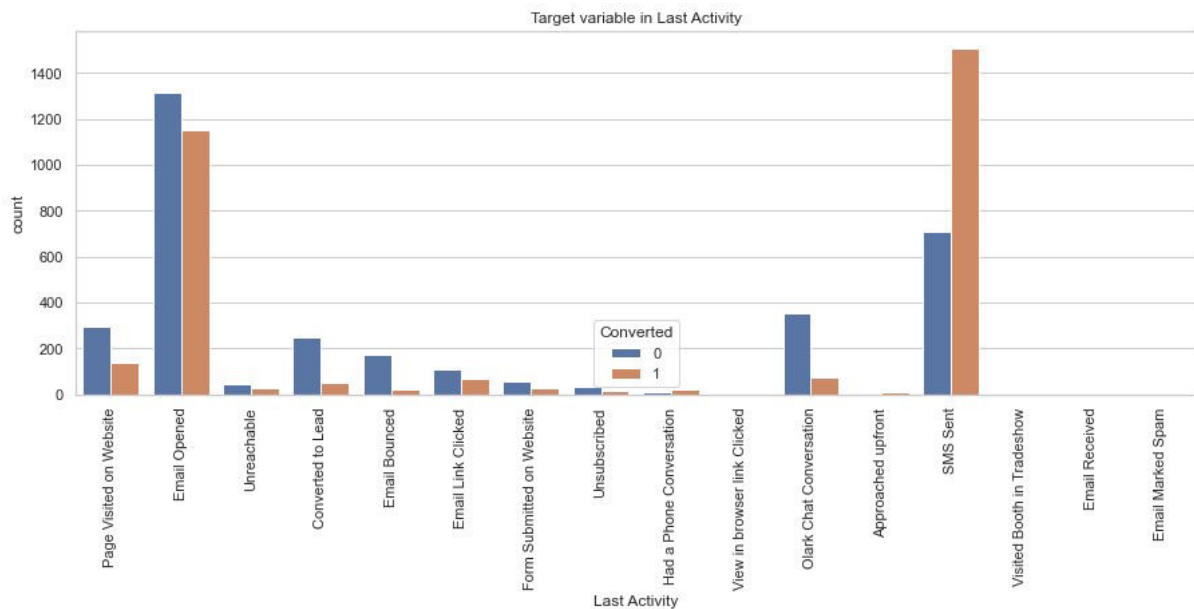
Count across Specialization

Lead Source & Lead Origin

In lead source the leads through google & direct traffic high probability to convert
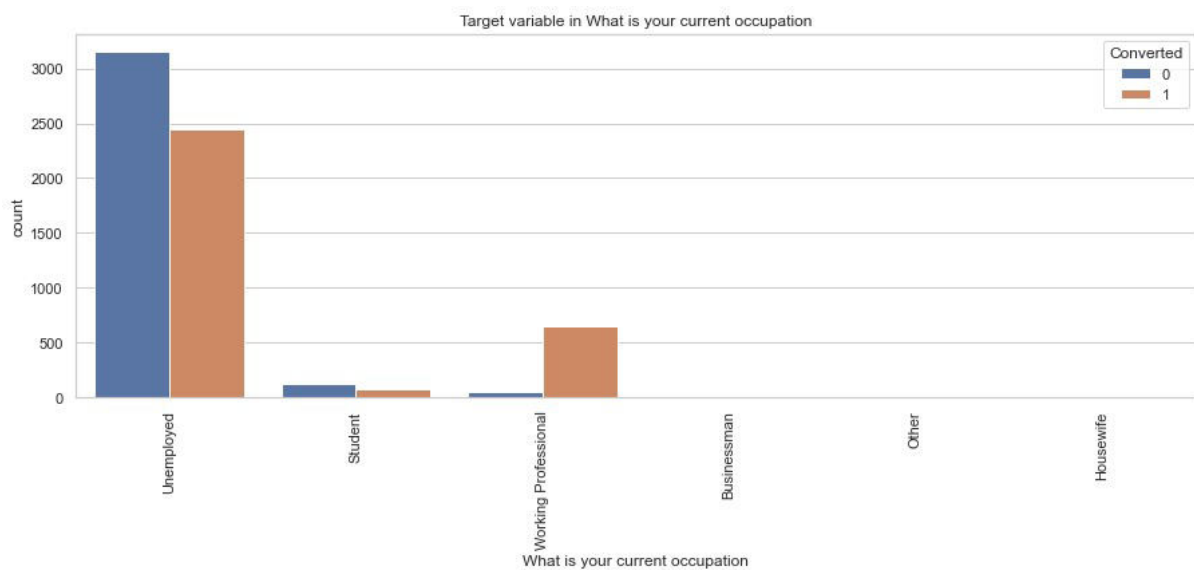
Last Lead Activity

Leads which are opening email have high probability to convert, Same as sending SMS will also benefit
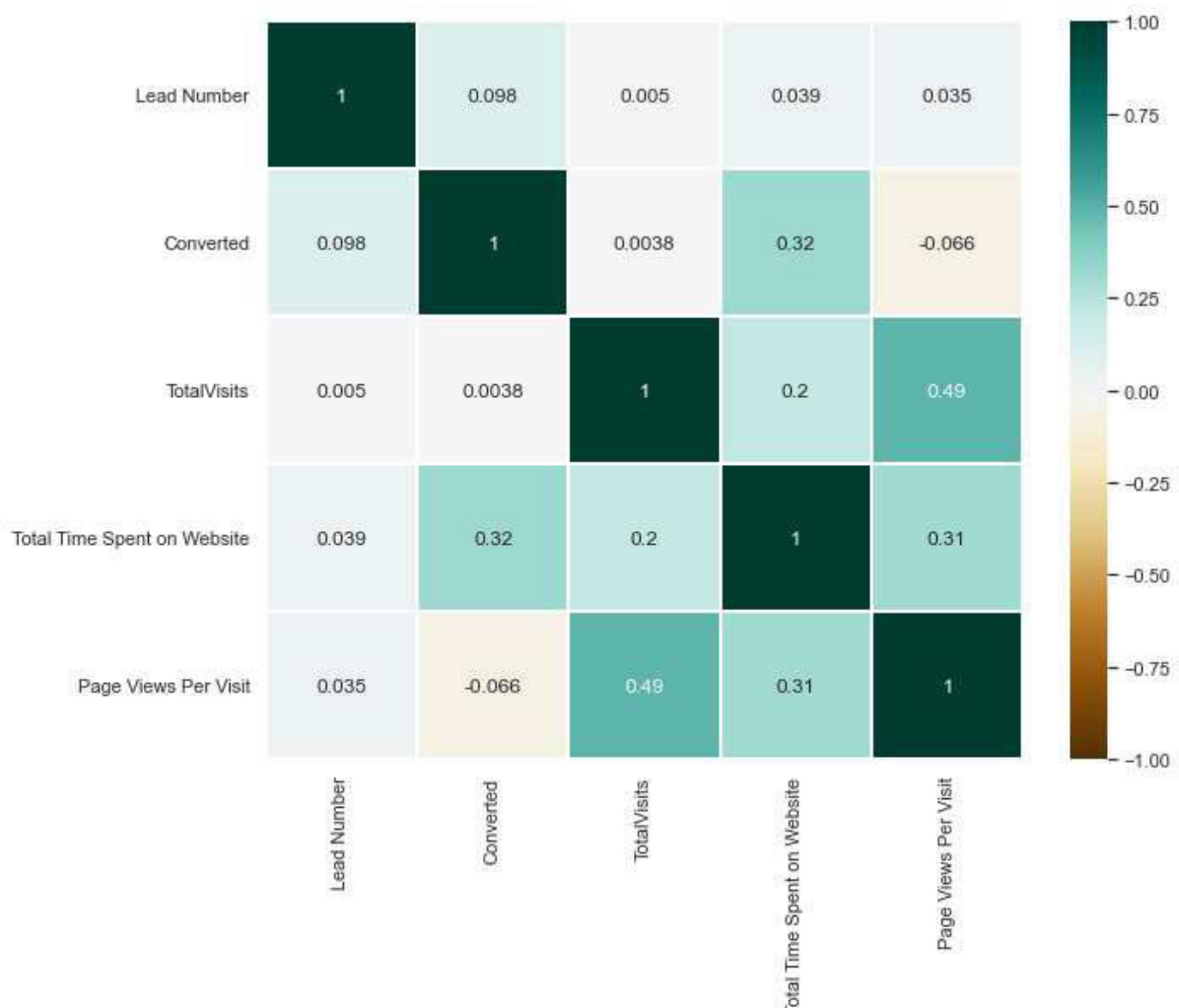

What is Your Occupation

Leads which are unemployed are more interested to join the course than others

Correlation

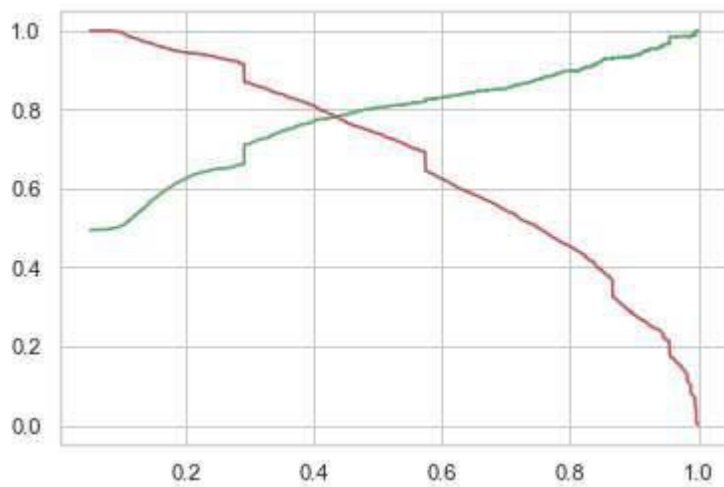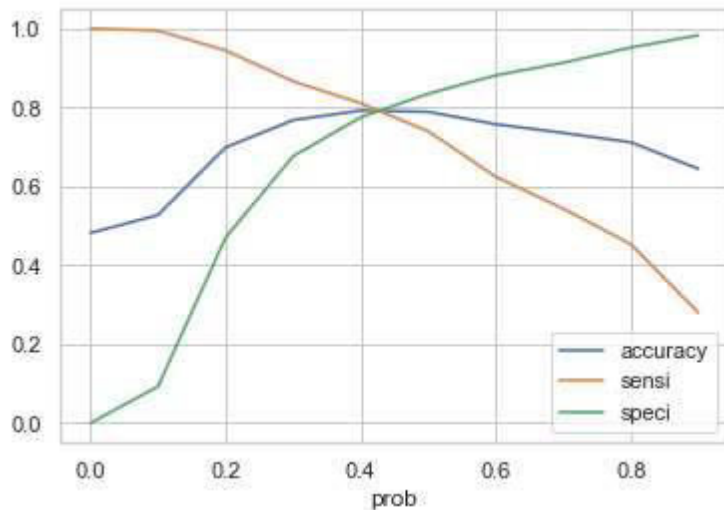No correlation between variables

Model Evaluation

ROC Curve

0.42 is the tradeoff between Precision and Recall

We can safely choose to consider any Prospect Lead with Conversion Probability higher than 42 % to be a hot Lead

Observations

Train Information:
80% accuracy
77% sensitivity

80% specificity
Test Results: 80% Accuracy
77% sensitivity
80% specificity

List of final features:

Specialization_Others
Lead Source_Olark Chat
Lead Source: Welingak Website
 Lead Origin: Lead Add Form
 Total Time Spent on Website
 Lead Origin: Landing Page Submission
 Current Job Title: Working Professionals
 Do Not Email

Conclusion

We observe that the conversion rate for landing page and API submissions is between 30 and 35 percent, which is about average. However, quite low for Lead import and Lead Add form. As a result, we can step in and emphasize the importance of the leads that came from the API and the submission of landing pages.

We observe that Google and direct traffic create the greatest number of leads. The maximum conversion rate is based on the Welingak website and references.

More time spent on the website increases the likelihood that a lead will convert. The most frequent last action is opening an email.

SMS Sent has the highest rate. Max does not have a job. maximum conversion with a professional in the workforce.