

CONTENT

Problem statement

Problem approach

EDA

Correlations

Model Evaluation

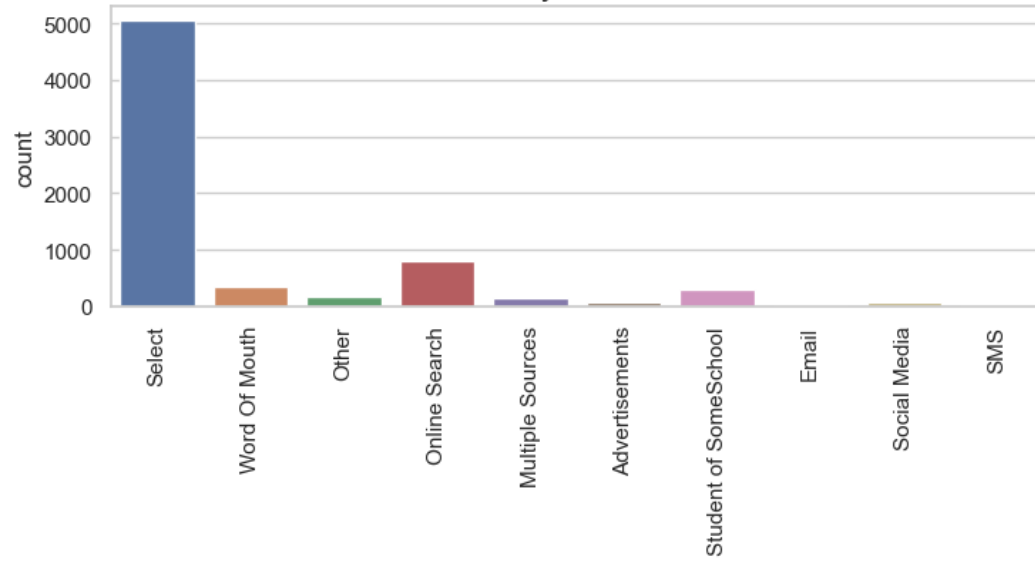
Observations

Conclusion

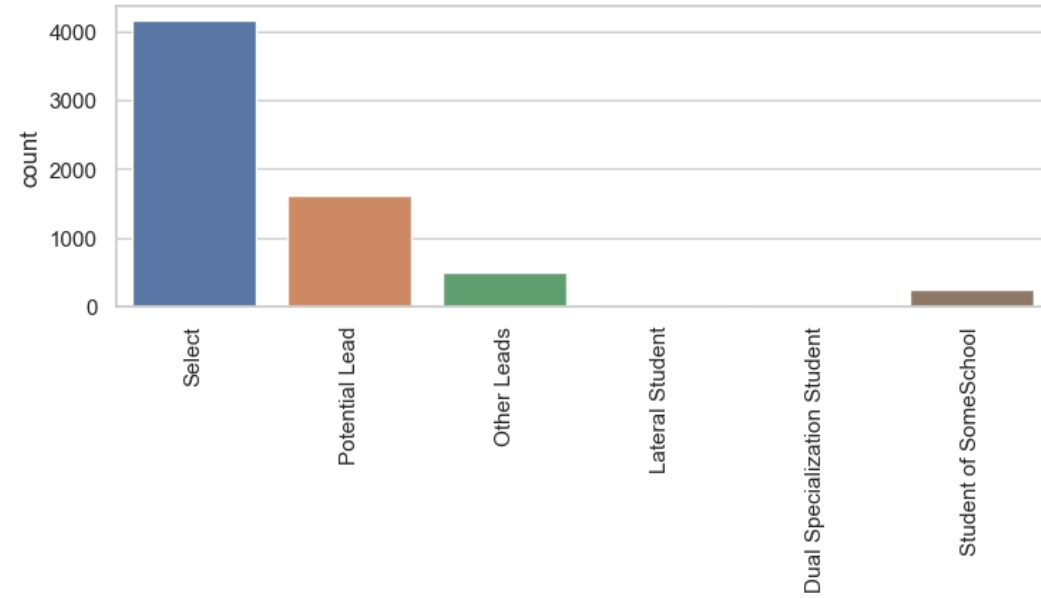
- Lead X wants us to build a model to give every lead a lead score between 0 -100 .
- So that they can identify the Hot leads and increase their conversion rate as well.
- The CEO want to achieve a lead conversion rate of 80o/o.
- They want the model to be able to handle future constraints as well like Peak time actions required, how to utilize full man power and after achieving target what should be the approaches.

- **Importing the data and inspecting the data frame**
- **Data preparation EDA**
- **Dummy variable creation Test-Train split**
- **Feature scaling Correlations**
- **Model Building (RFE Rsquared VIF and p values)**
- **Model Evaluation**
- **Making predictions on test set**

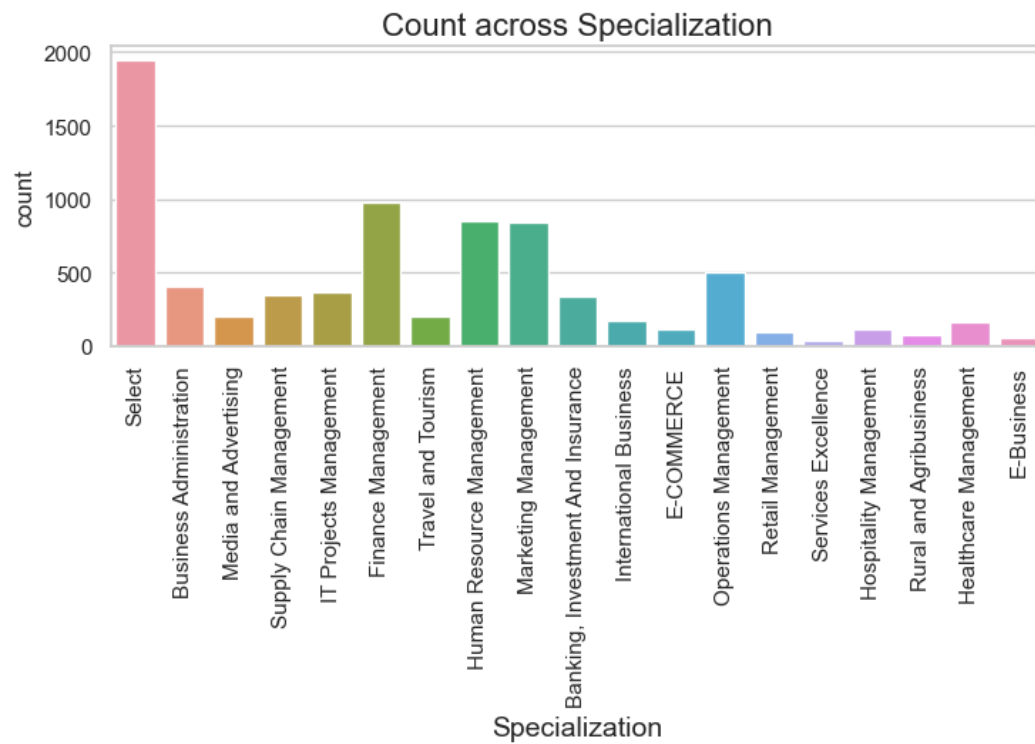
Count across How did you hear about X Education



Count across Lead Profile

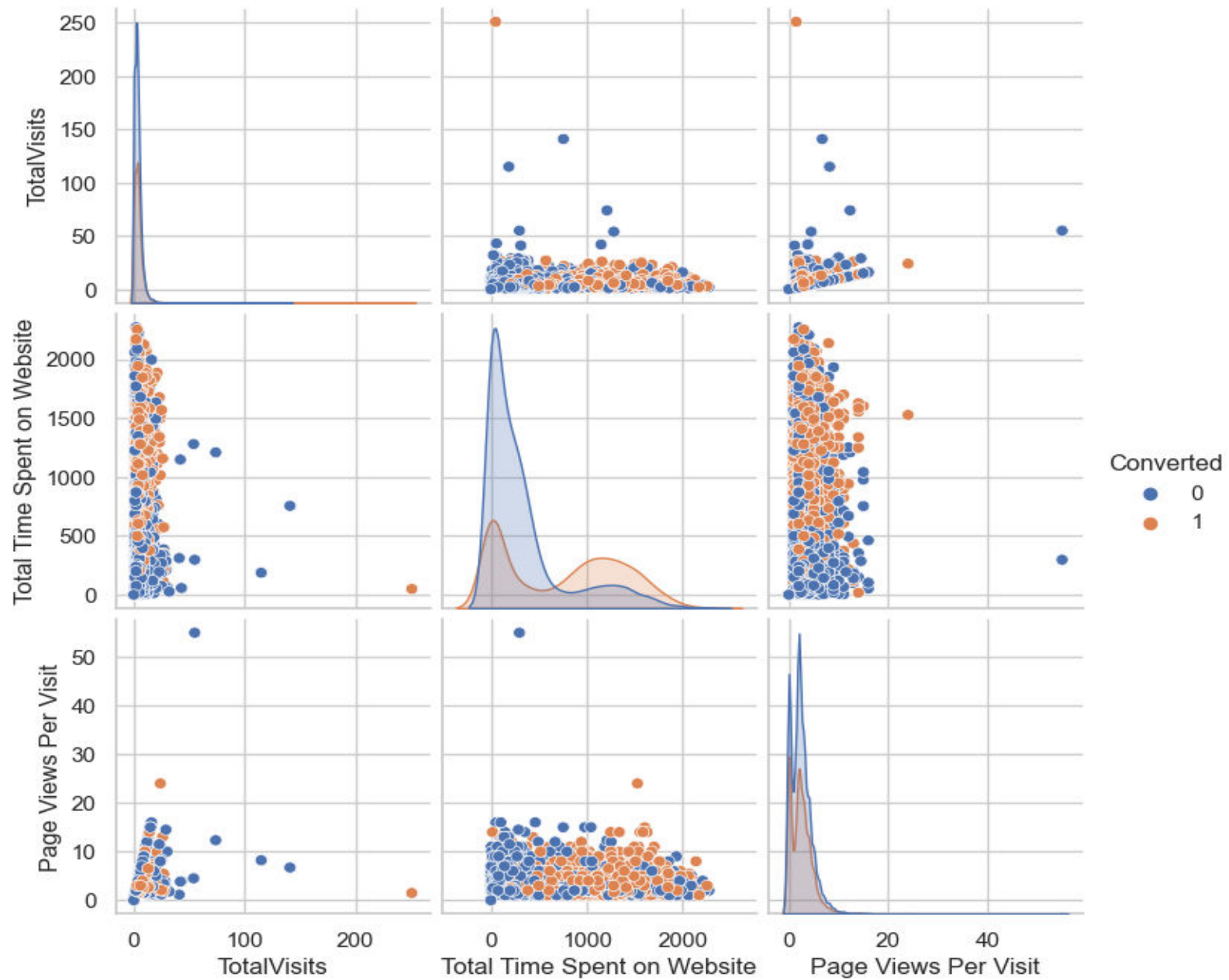


Count across Specialization



Lead Profile

Lead and Education Analysis



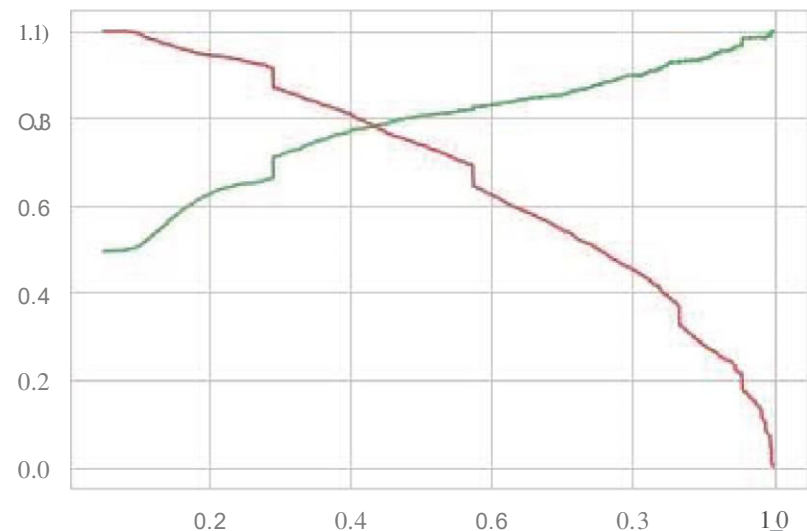
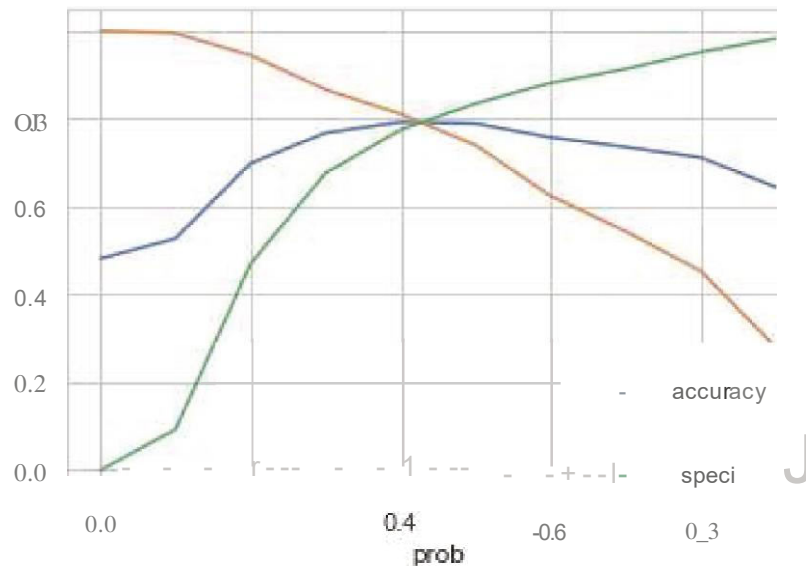
Correlation



Model Building

0.42 is the tradeoff between Precision and Recall -

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



Train Data:

Accuracy : 80%
Sensitivity : 77%
Specificity : 80%

Test Data:

Accuracy : 80%
Sensitivity : 77%
Specificity : 80%

Final Features list:

Lead Source Olark Chat
Specialization Others
Lead Origin_Lead Add Form
Lead Source_Welingak Website
Total Time Spent on Website
Lead Origin_Landing Page Submission
What is your current occupation_Working Professionals
Do Not Email

Conclusion

We see that the conversion rate is 30-35% (close to average) for API and Landing page submission. But very low for Lead Add form and Lead import. Therefore we can intervene that we need to focus more on the leads originated from API and Landing page submission.

We see max number of leads are generated by google / direct traffic. Max conversion ratio is by reference and welingak website.

Leads who spent more time on website, more likely to convert.

Most common last activity is email opened. highest rate = SMS Sent. Max are unemployed. Max conversion with working professional.