



LEAD SCORING CASE STUDY

A decorative geometric pattern in the top-left corner of the slide. It consists of several overlapping shapes: a dark blue triangle with white diagonal lines, a light blue circle, a dark blue square with concentric circles, a dark purple triangle, a blue square, a pink square with concentric circles, a grey square with concentric lines, and a purple square. A small dark blue circle is located at the intersection of the dark purple triangle and the pink square.

CONTENTS

Problem Statement

Approach

EDA

Correlations

Model Evaluation

Observation

Conclusion



PROBLEM STATEMENT

AN EDUCATION COMPANY NAMED X EDUCATION SELLS ONLINE COURSES TO INDUSTRY PROFESSIONALS. ON ANY GIVEN DAY, MANY PROFESSIONALS WHO ARE INTERESTED IN THE COURSES LAND ON THEIR WEBSITE AND BROWSE FOR COURSES. 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%



PROBLEM APPROACH

Importing the data and inspecting data frame

Data preparation

EDA

Dummy Variable Creation

Test Train Split

Feature Scaling

Correlations

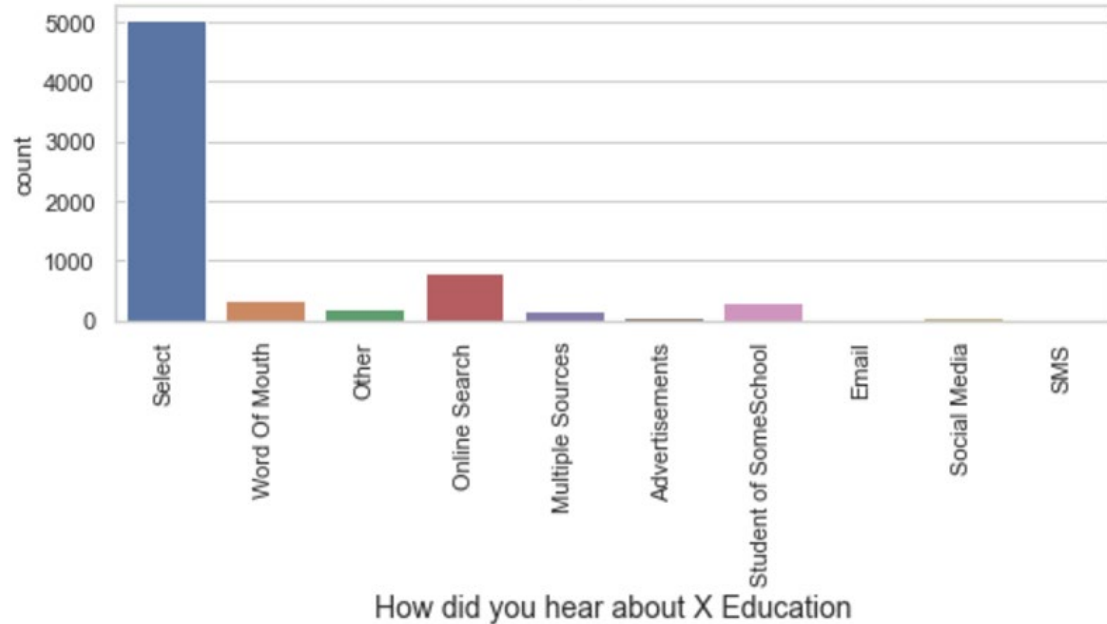
Model Building (RFE Rsquared VIF and p values)

Model Evaluation

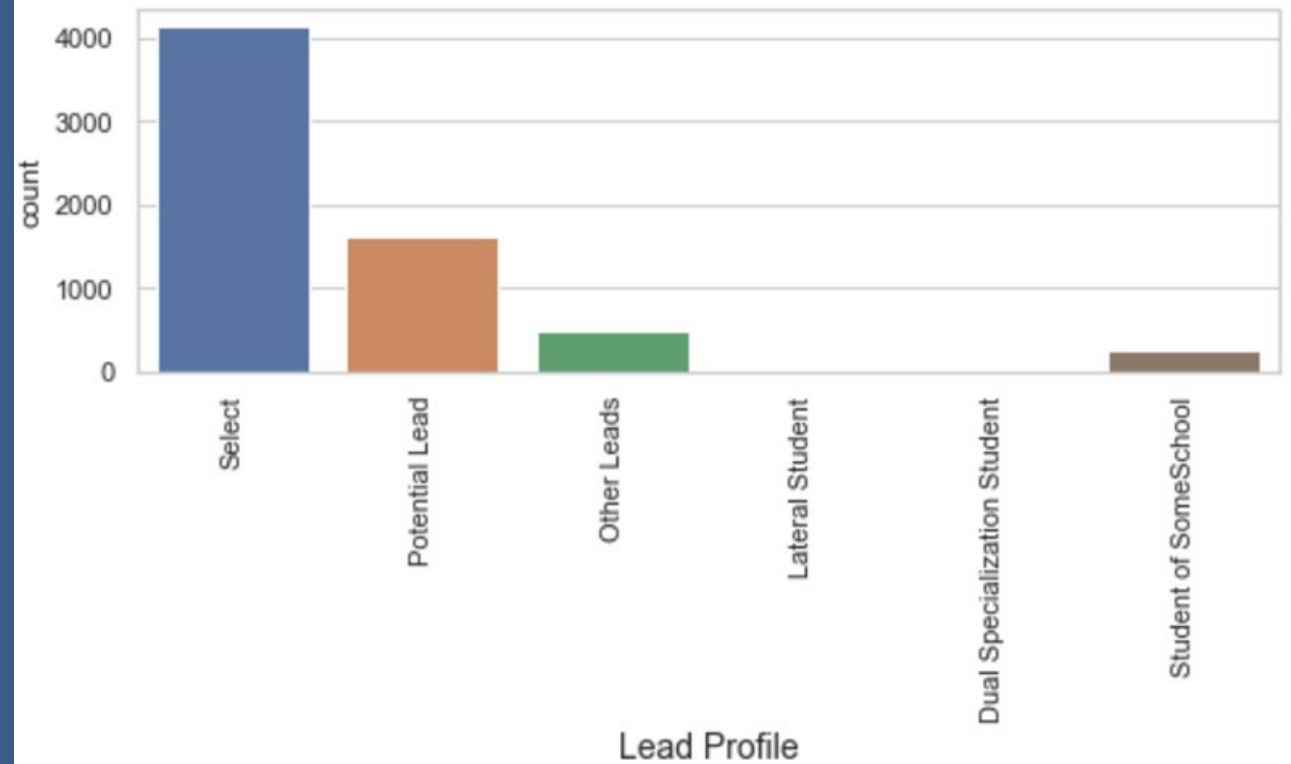
Making prediction on test set

EDA DATA CLEANING

Count across How did you hear about X Education

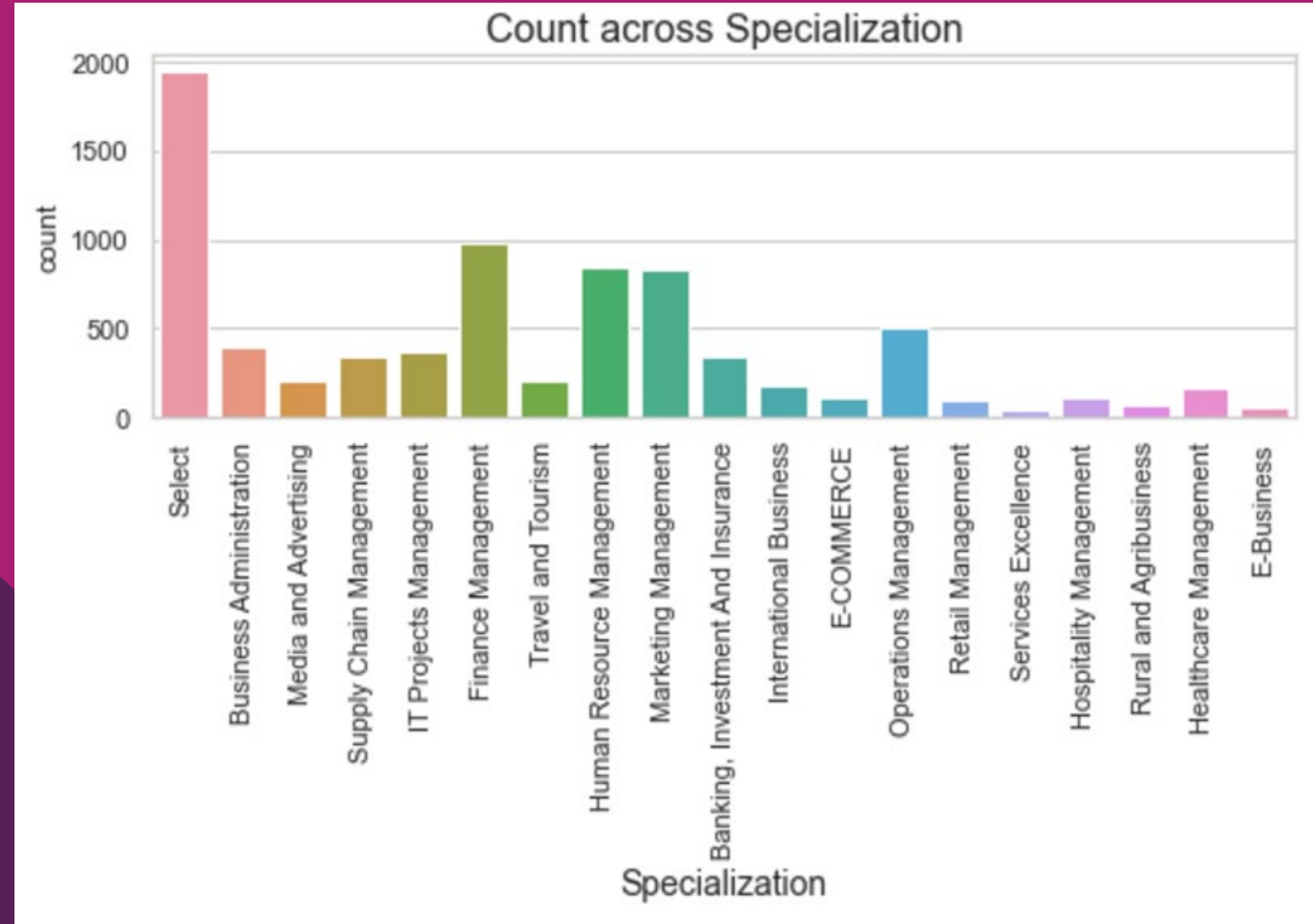


Count across Lead Profile



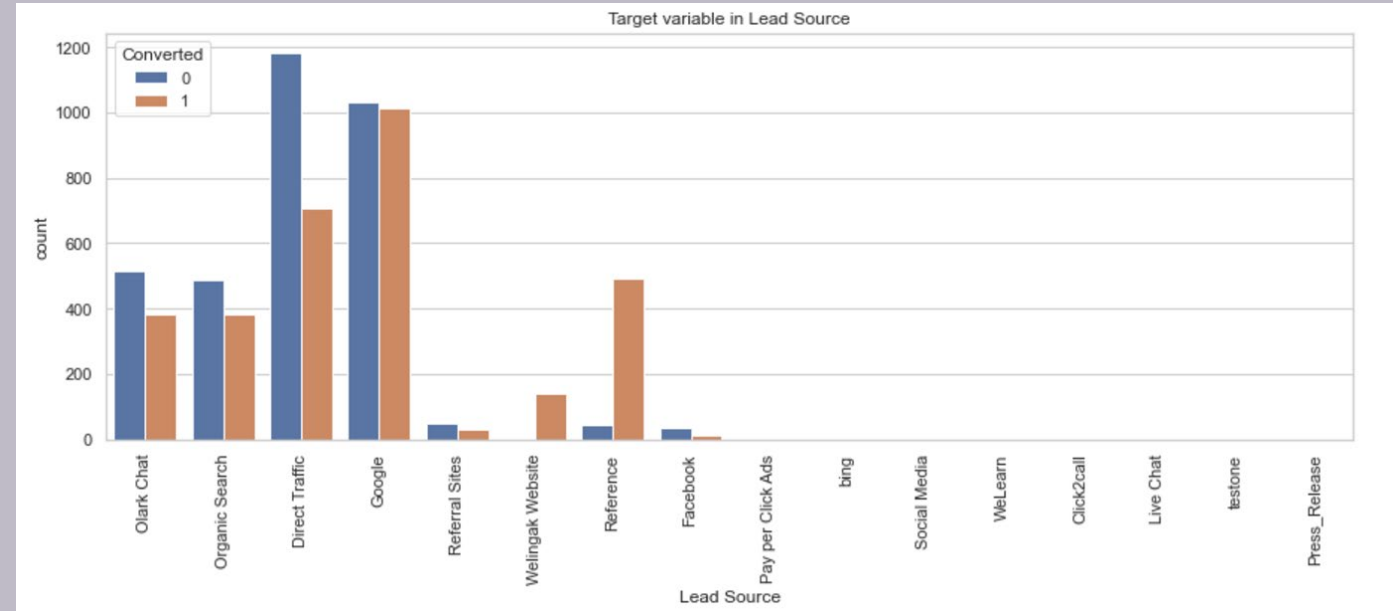
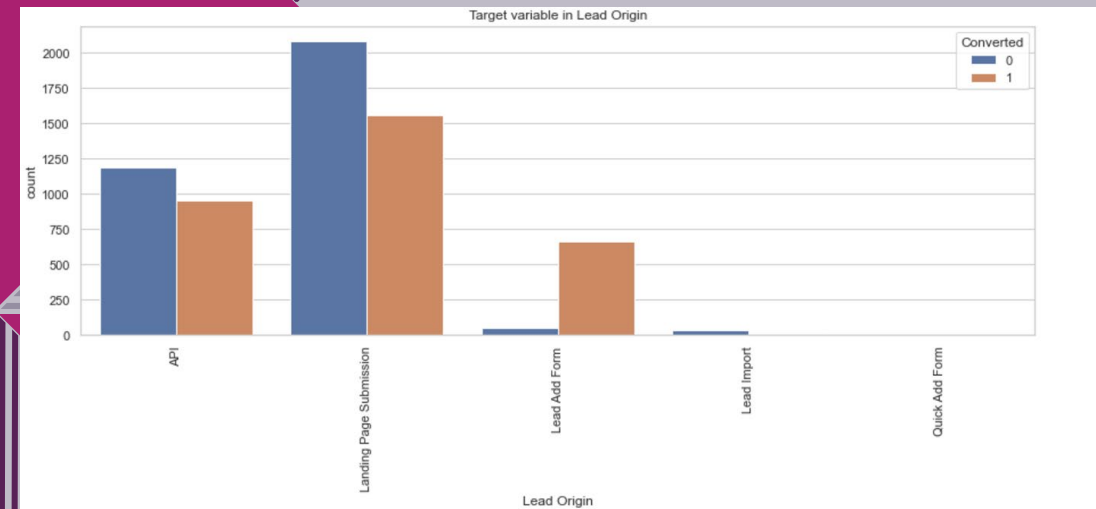
Columns with level is Select is taken care here

LEADS WITH HIGH PROBABILITY TO CONVERT ARE FROM HR , MARKETING AND FINANCE



LEAD SOURCE & ORIGIN

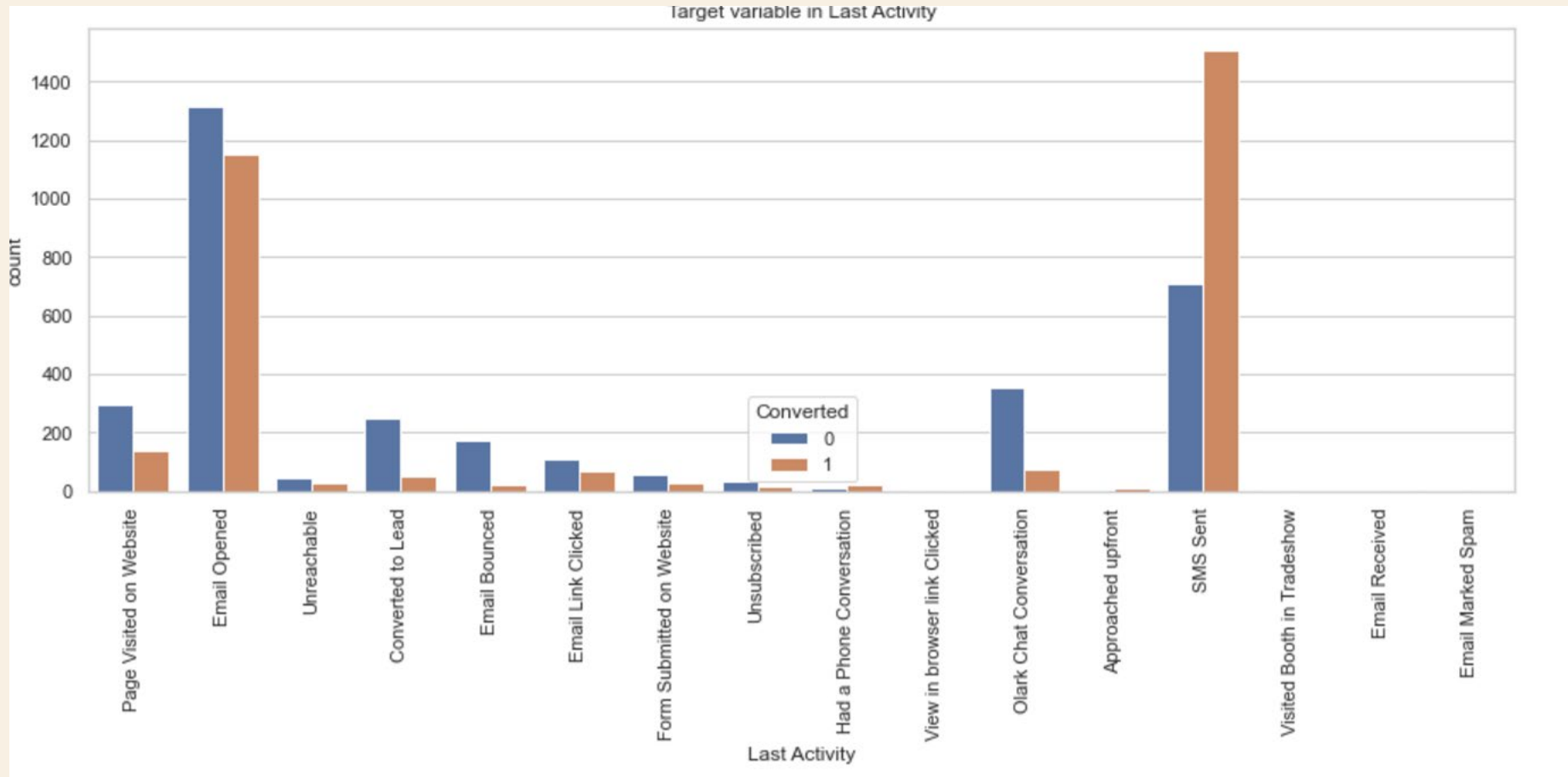
- As per lead source the leads through google & direct traffic have high probability to convert



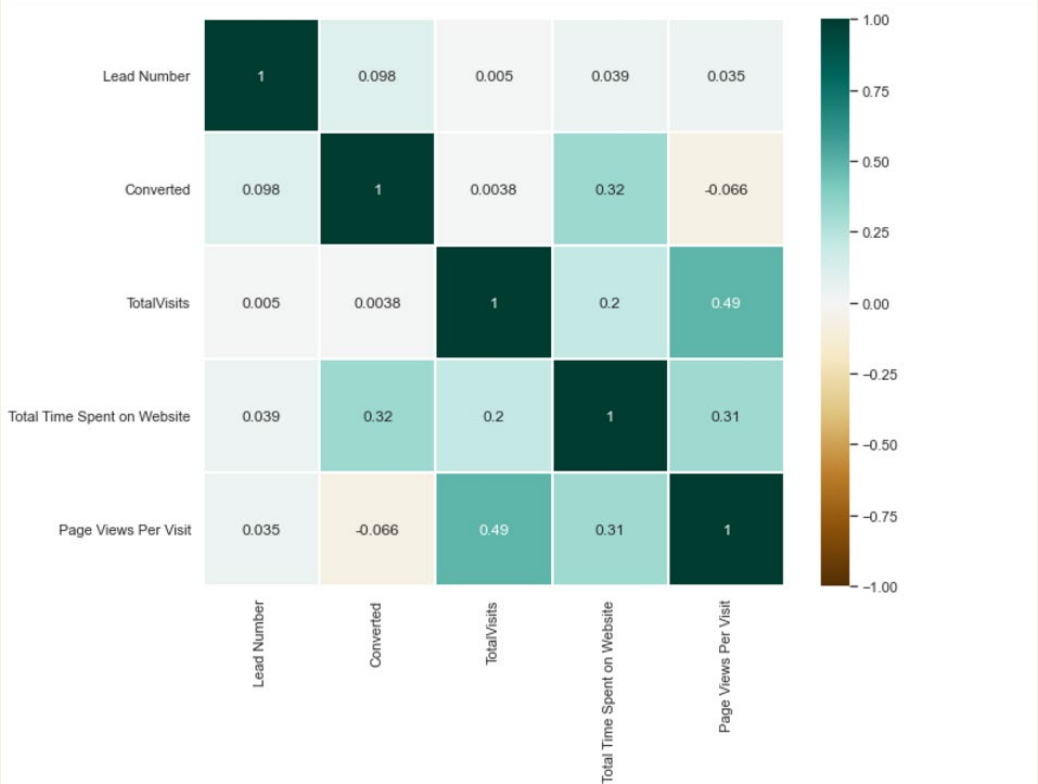
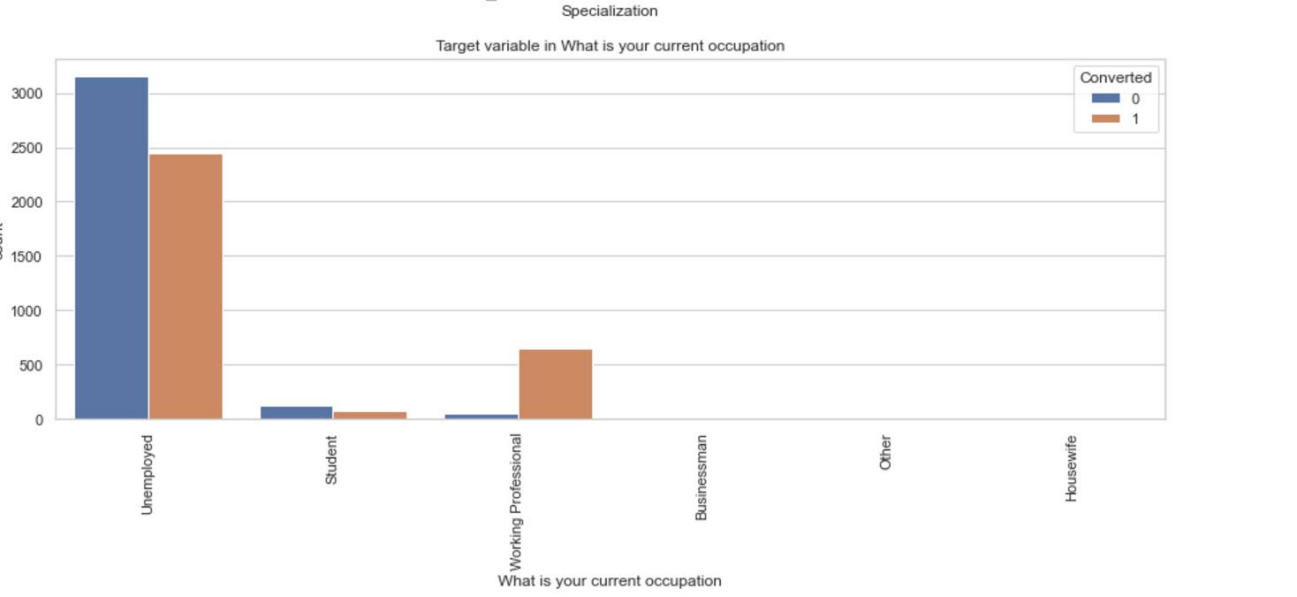
- Whereas in Lead Origin most number of leads are landing on Page Submission

LAST LEAD ACTIVITY

- Leads opening email , have higher probability to convert



LAST OCCUPATION & CORRELATION

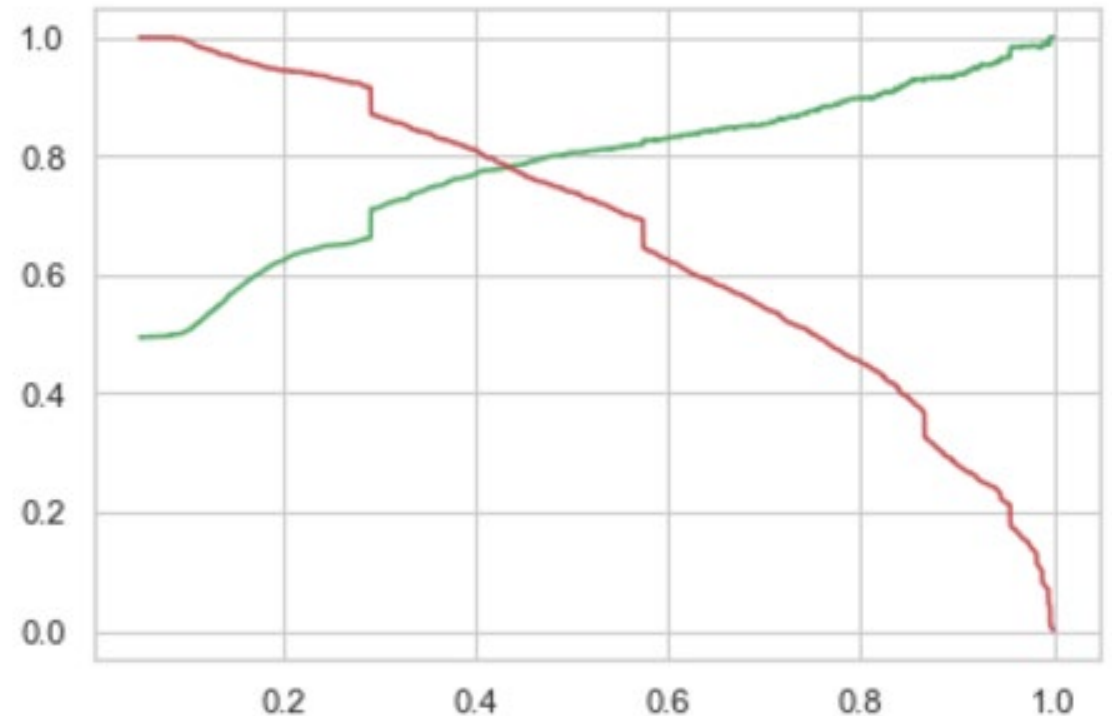
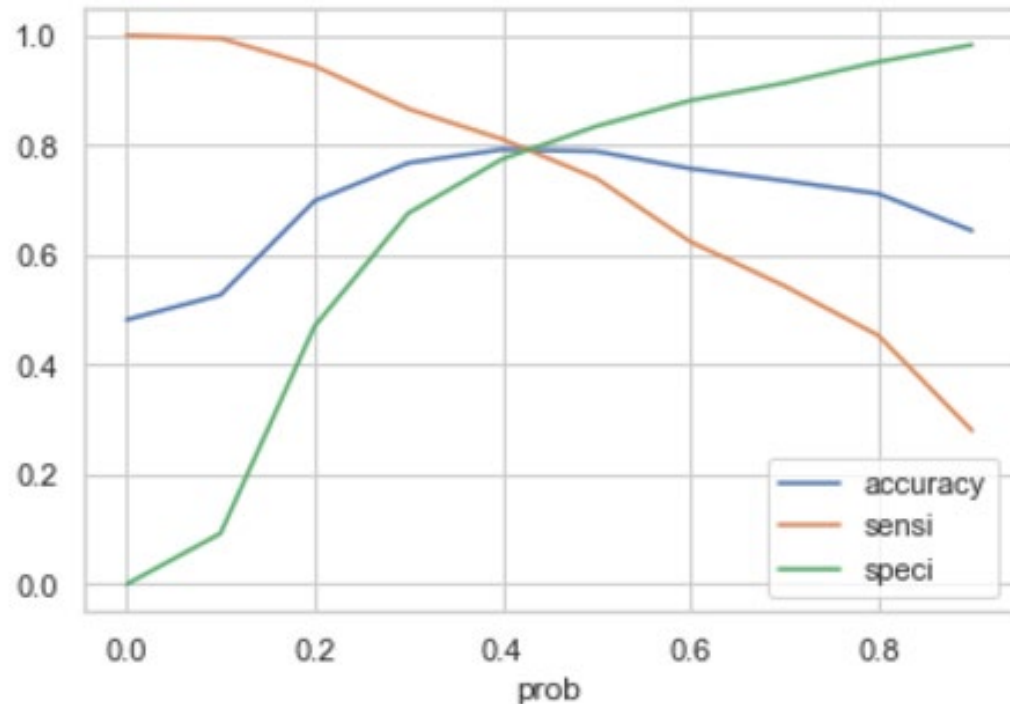


MODEL EVALUATION

- ROC Curve

0.42 is the tradeoff between Precision and Recall –

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



OBSERVATIONS

- Train Data :

Accuracy : 80%

Sensitivity : 77%

Specificity : 80%

Test Data :

Accuracy : 80%

Sensitivity : 77%

Specificity : 80%

- Final Feature 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 Occupation_Working Professionals

- Do Not Email

CONCLUSION

- We see that 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 professionals

An abstract geometric design on the left side of the slide. It features a dark blue background with various geometric shapes and patterns. A white circle is positioned near the top left. Below it, a light blue semi-circle is visible. To the right of the semi-circle, there is a pink triangle with diagonal lines. Further down, there is a pink square with a pattern of concentric lines. At the bottom, there is a pink triangle with a pattern of concentric lines. The overall design is modern and minimalist.

THANK YOU