

X-EDUCATION LEAD SCORING CASE STUDY

Group Members:
Sharabh Das
Ashna Anoop
Yogendra S

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 conversion rates of the leads are very low for the company and they want to increase the rate of conversion.
- The correct identification of the leads which have the potential to get converted needs to be focused on and if the sales team correctly identifies those correct set of leads, they might be able to onboard business much more efficiently.

BUSINESS OBJECTIVE

- ❓ 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 80%.
- ❓ 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.

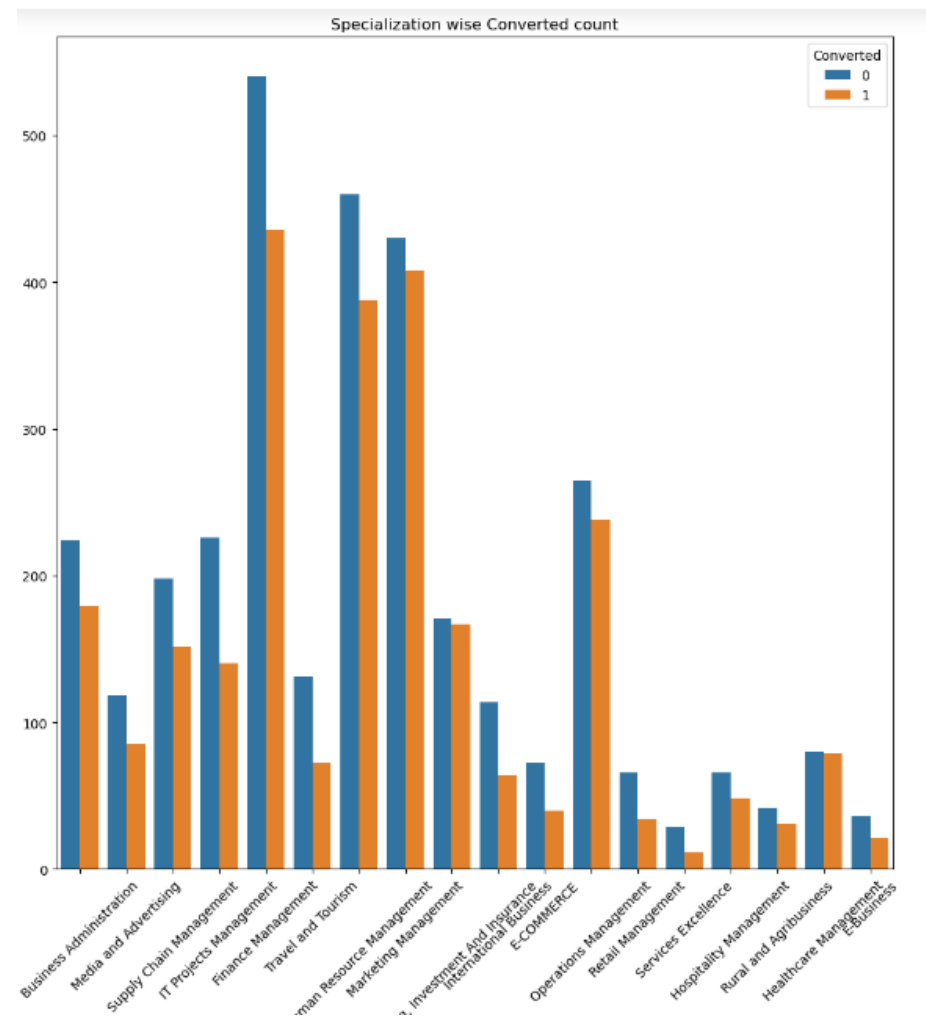
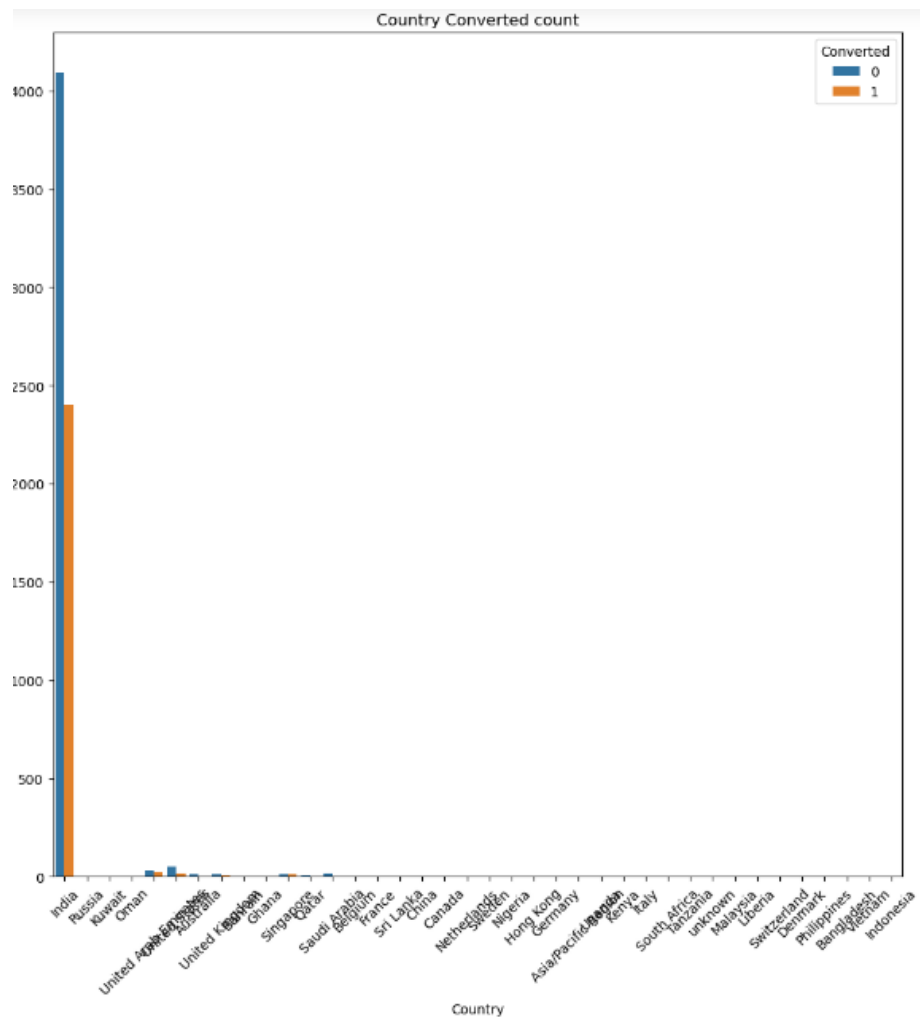
LEAD CONVERSION PROCESS

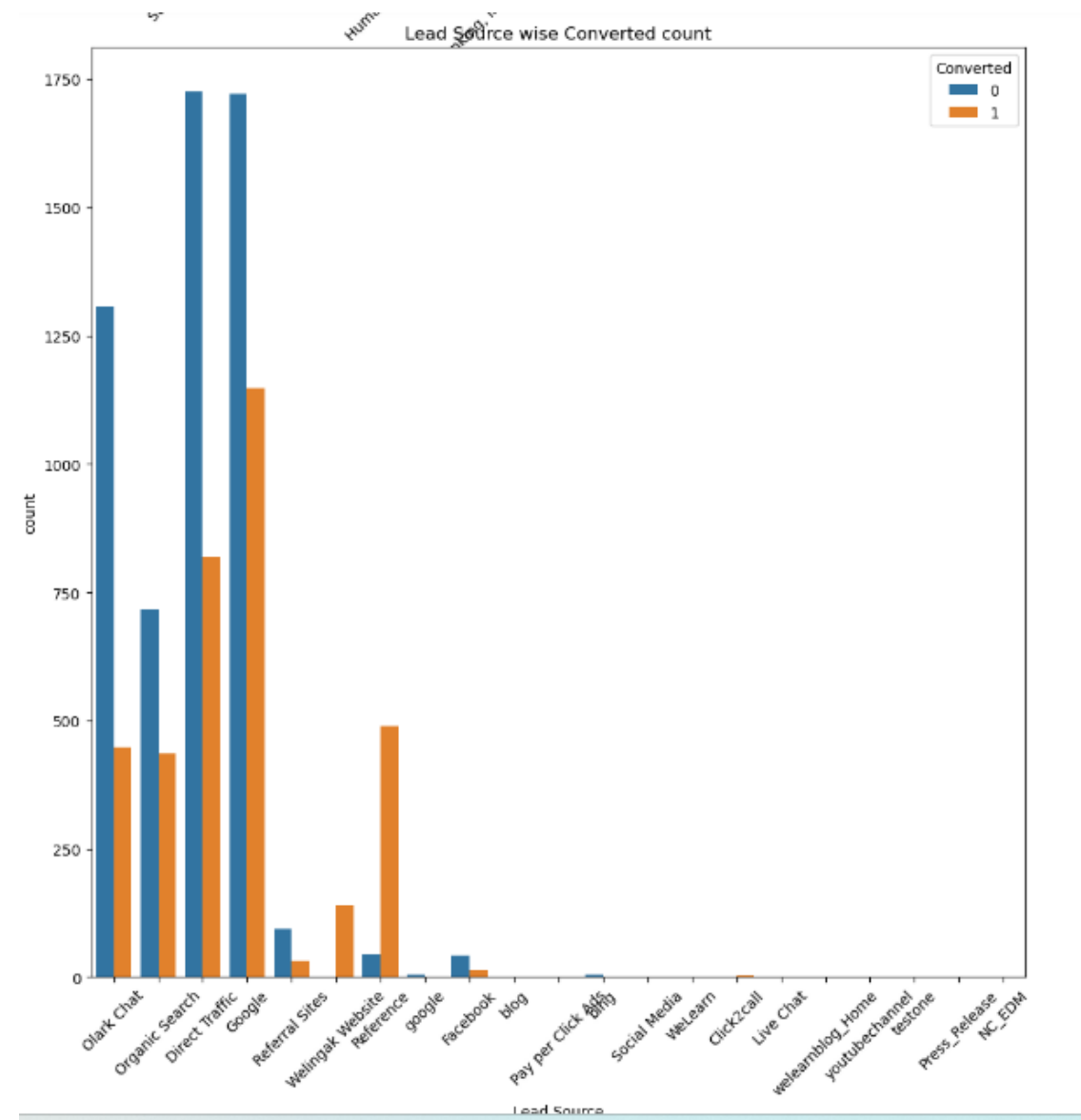
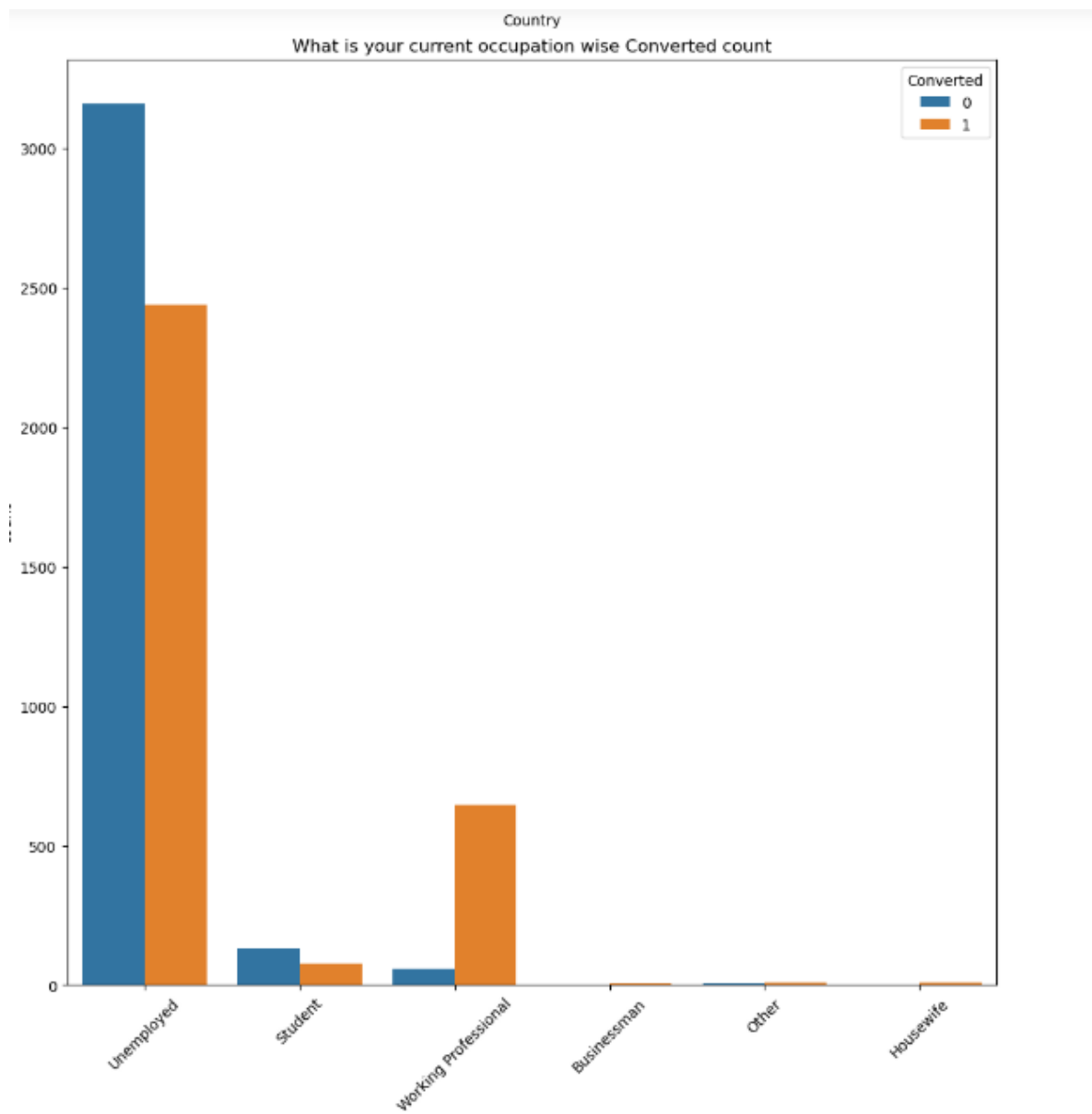


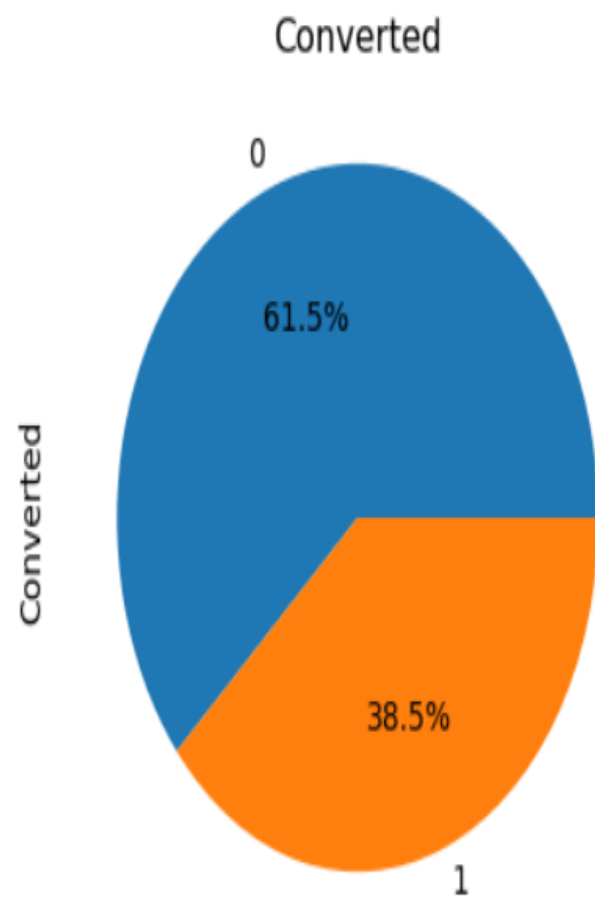
PROBLEM APPROACH

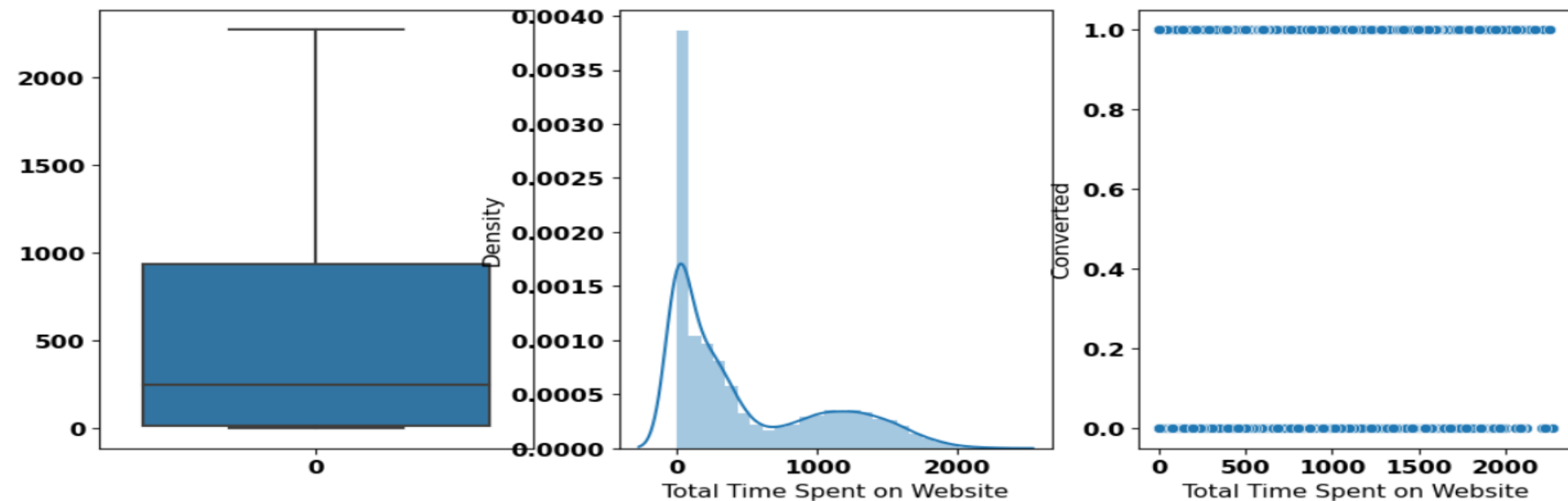
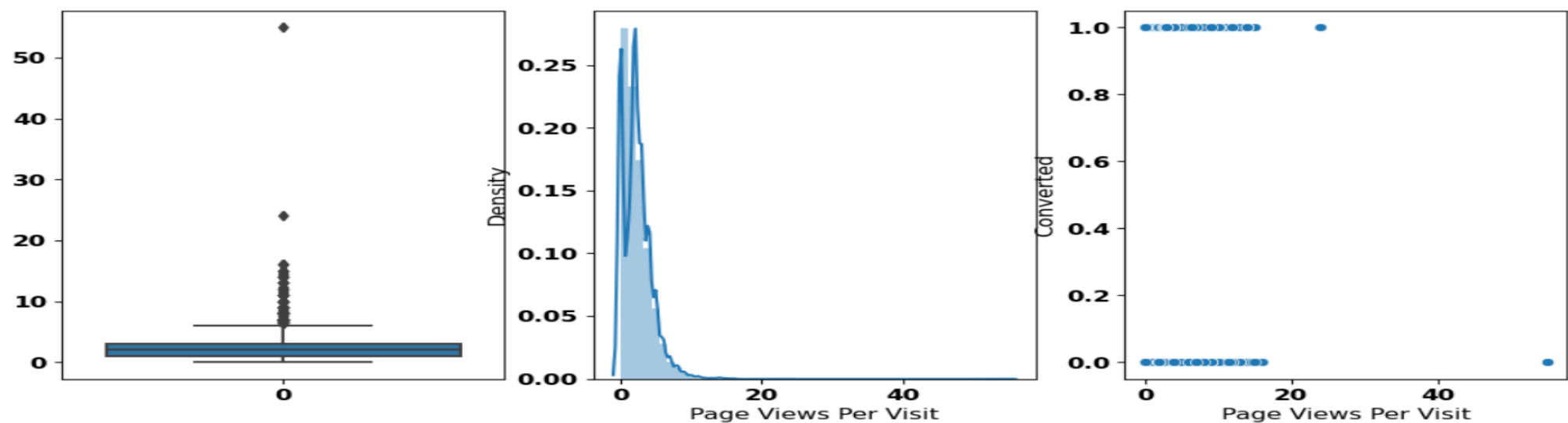
Importing the data and inspecting the data frame

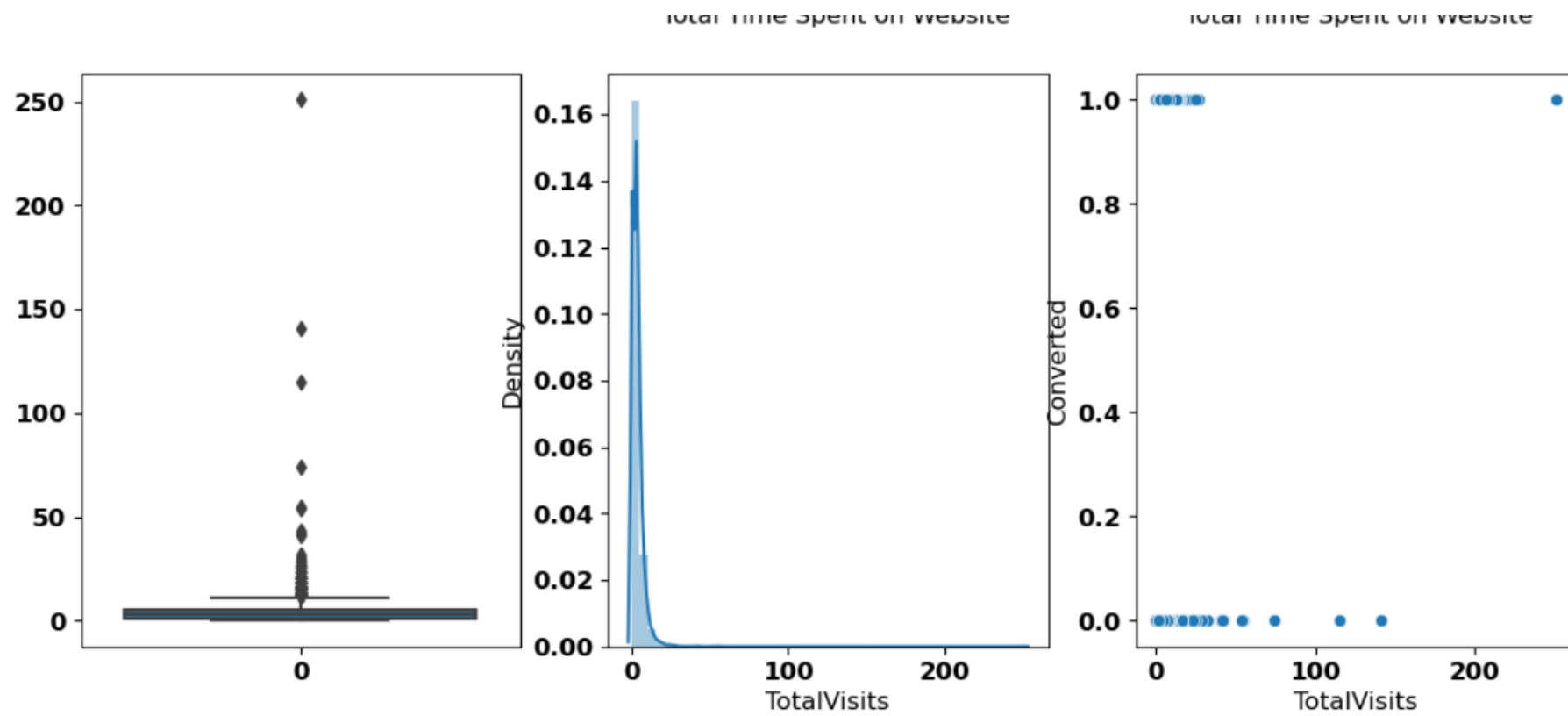
- Data preparation
- EDA
- Dummy variable creation
- Test-Train split
- Feature scaling
- Correlations
- Model Building (RFE Rsquared VIF and pvalues)
- Model Evaluation
- Making predictions on test set

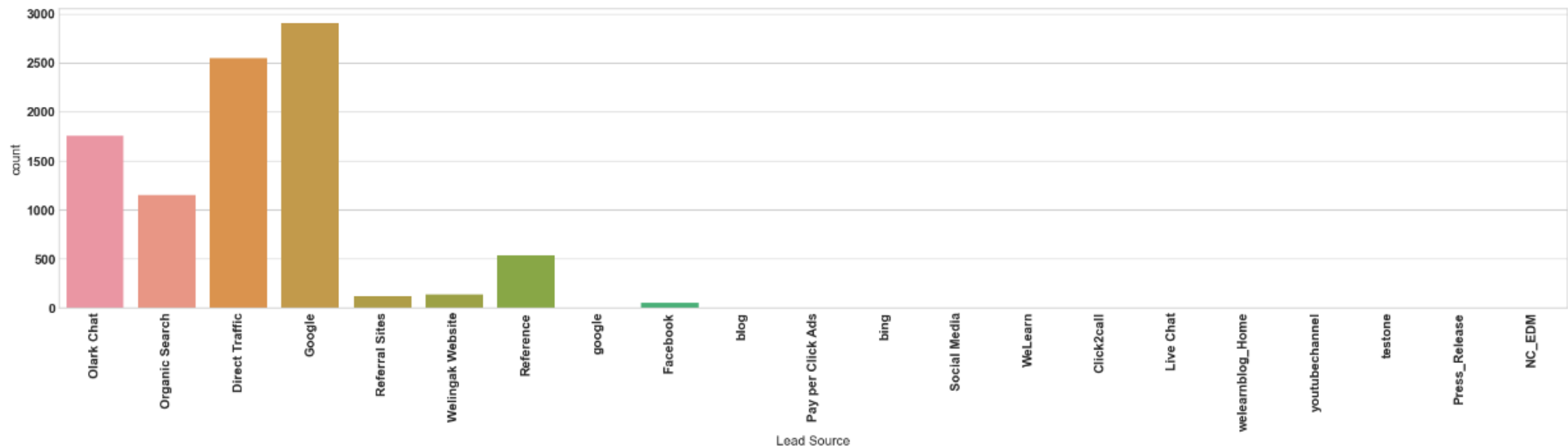
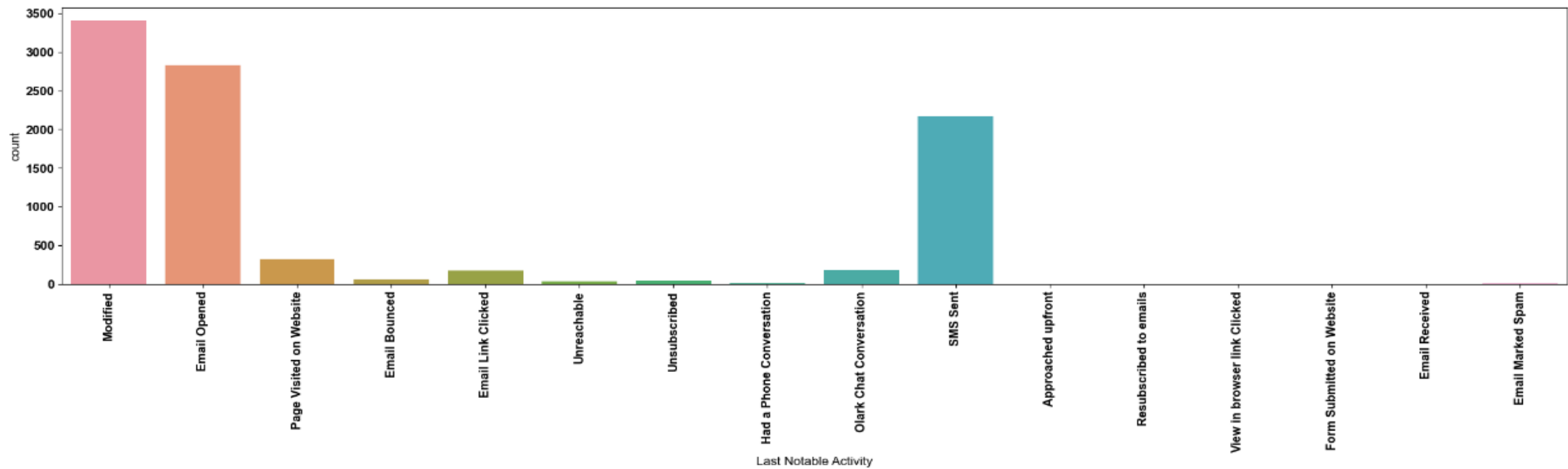


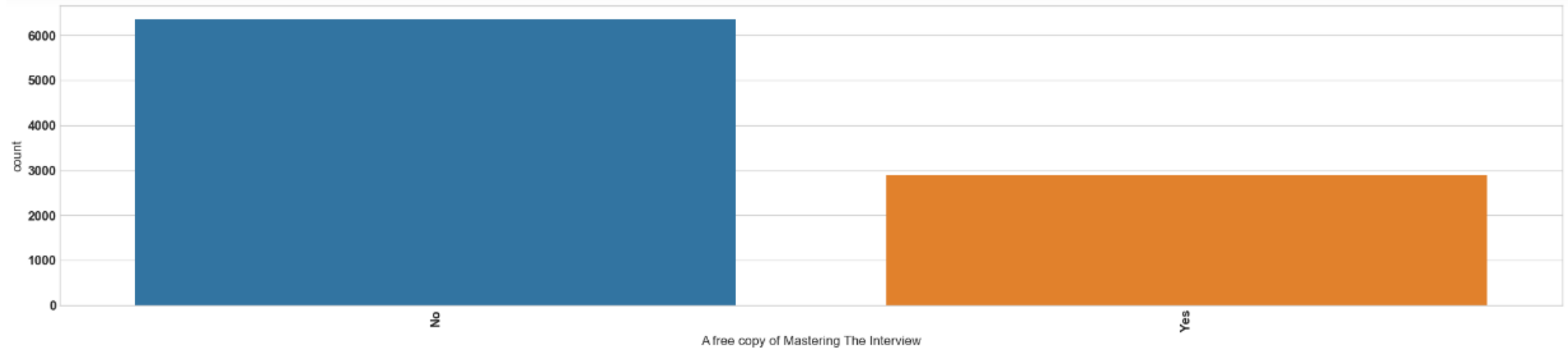
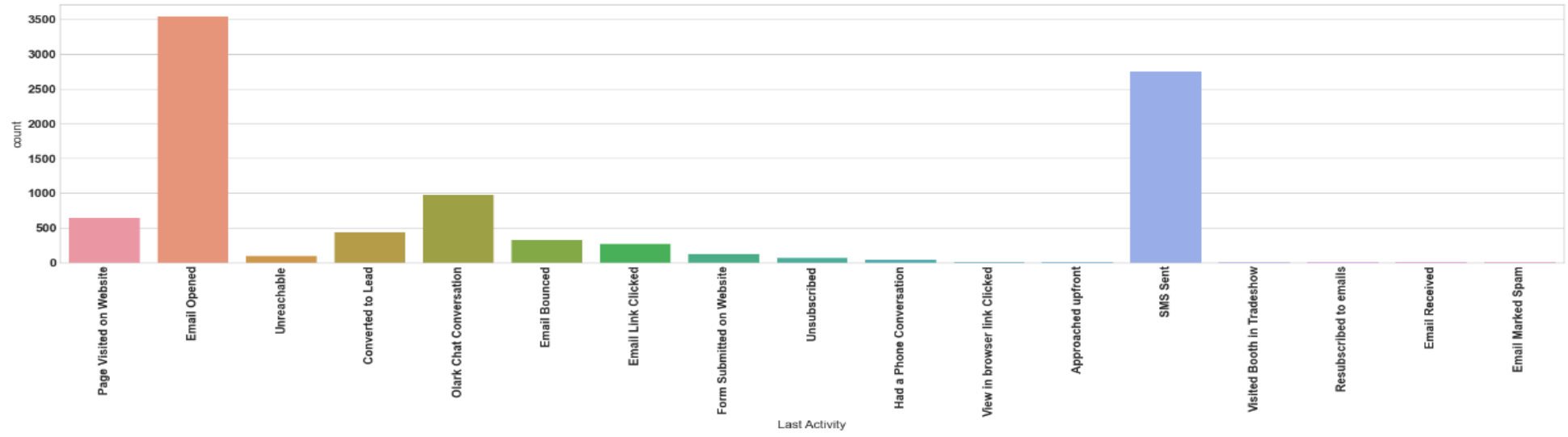


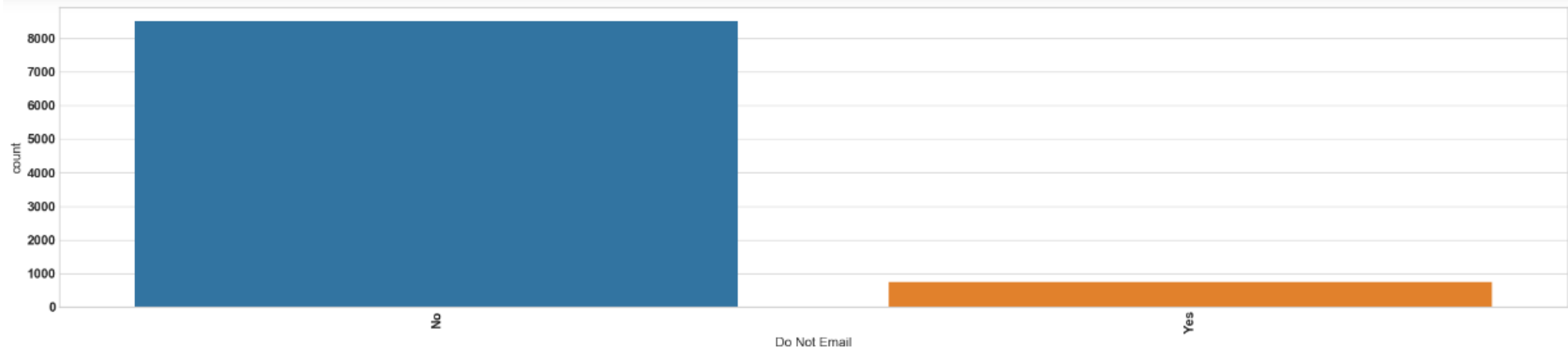
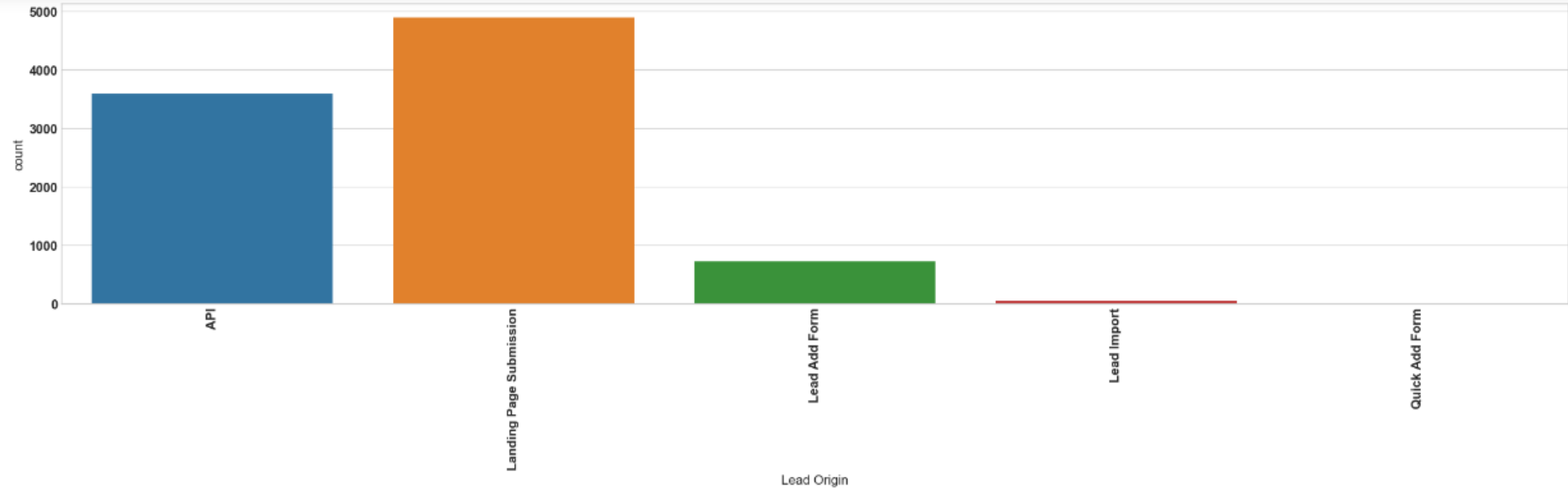


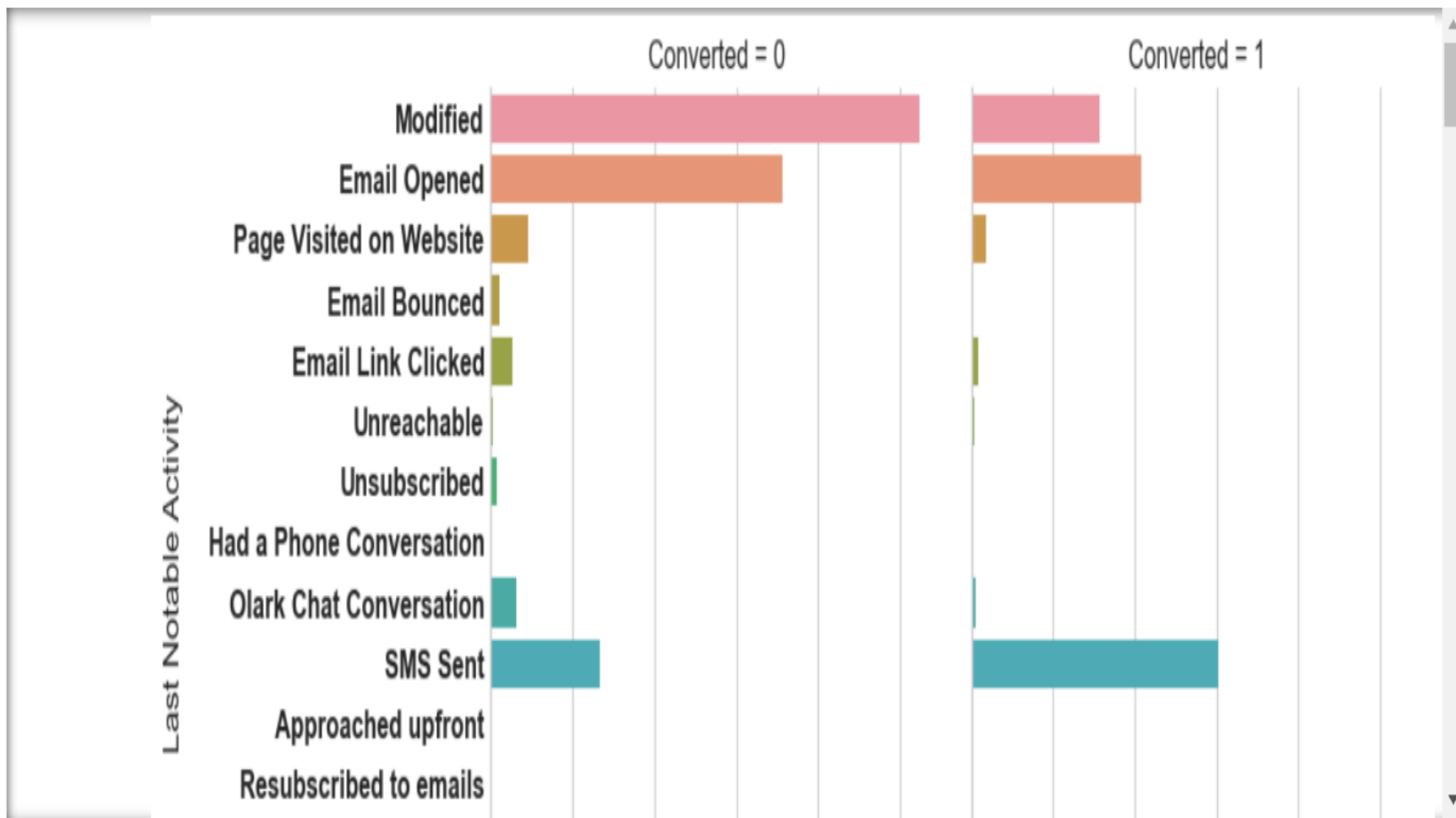








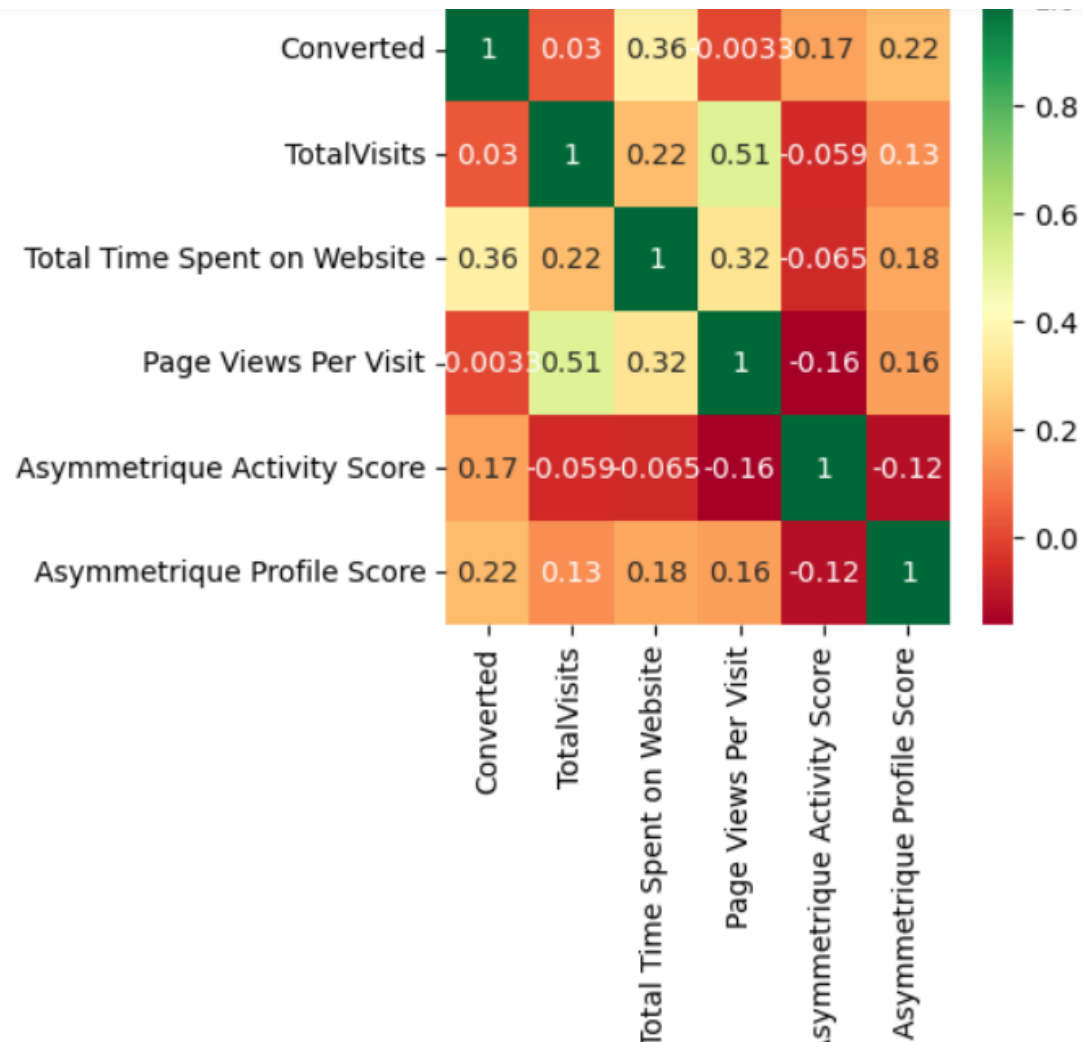




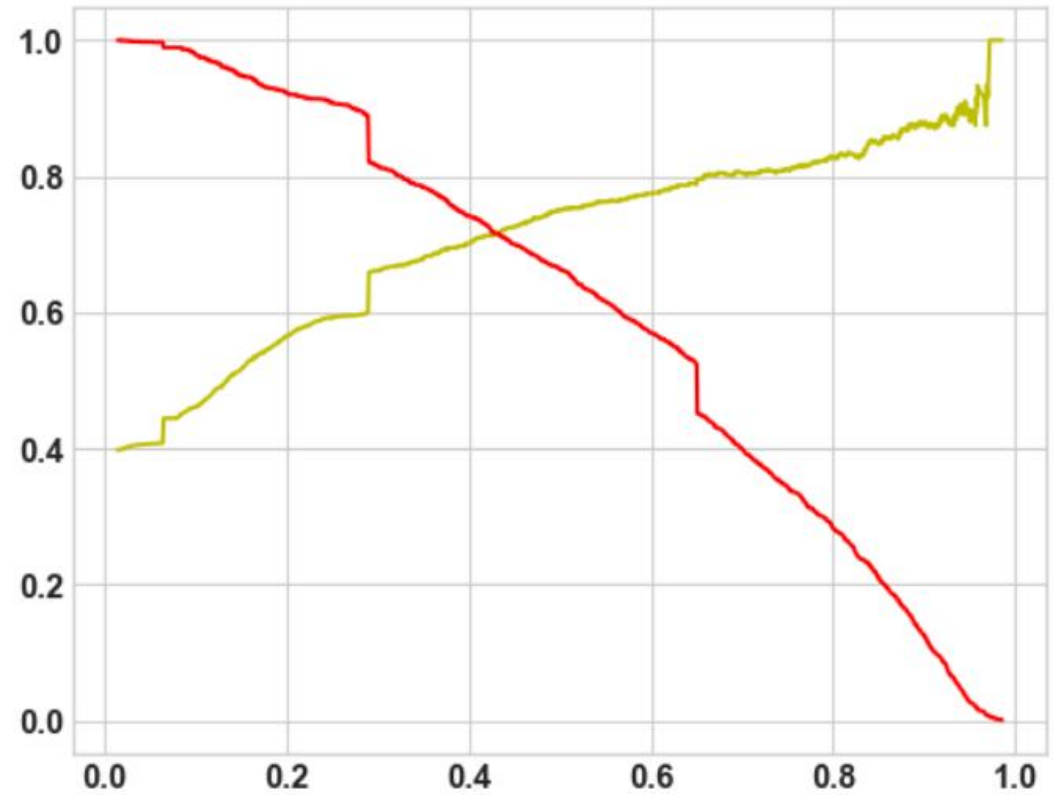
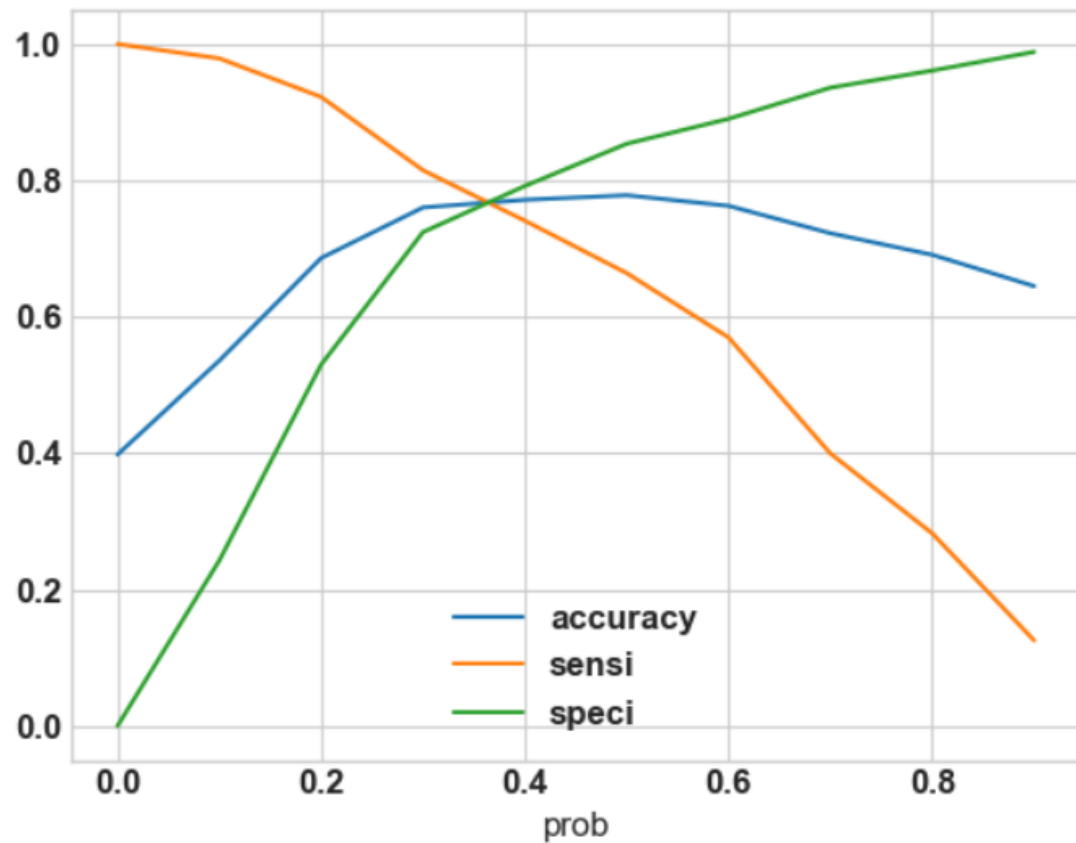
```
In [23]: plt.figure(figsize=(4,4))
```

- From the EDA we find out that India as a country is the topmost source of leads.
- Only 38.5% of the total leads were converted
- The highest amount of lead origin count is Landing page submission.
- Amongst the last notable activity the conversion rate of SMS sent activity is the highest, followed by Email opened and SMS sent

CORRELATION



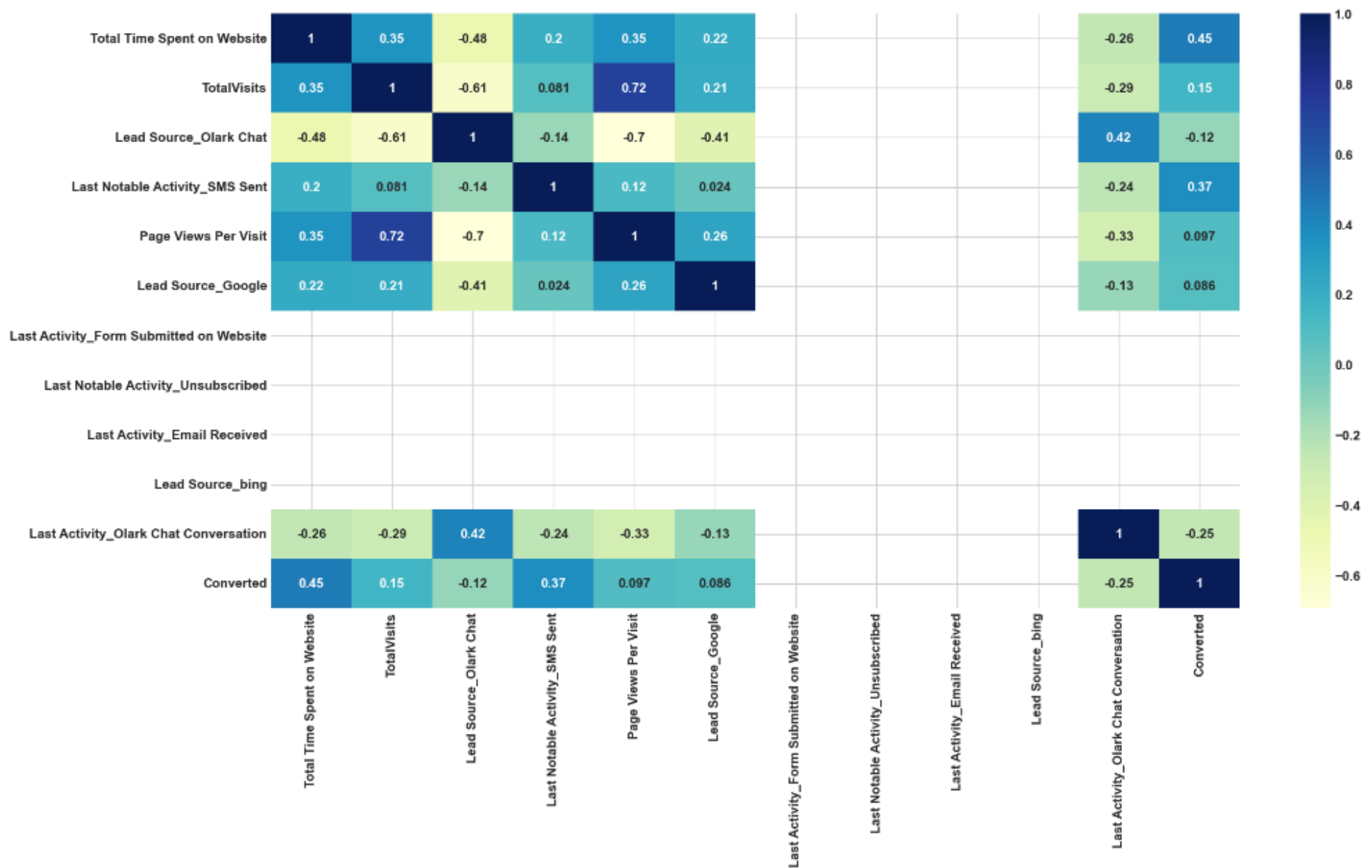
LINEAR REGRESSION PARAMETERS



The intermediate cut-off as per the graphs in the previous slide is 0.35 and the final cut-off as 0.42, we can safely say anything above the probability of 0.42 could be considered a hot lead.

MODEL ANALYSIS

- The sensitivity of the model turns out to be: 0.73
- The Specificity of the model turns out to be: 0.81
- The Precision of the model turns out to be :0.71
- The recall of the model turns out to be: 0.73
- The f1 score of the model turns out to be: 0.72



RECOMMENDATION/CONCLUSION

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

Increase user engagement on their website as this helps in higher conversion rate

Get Total visits increased by advertising, SEO, digital marketing etc.

Increase sending SMS notifications.

Improve the Olark Chat service by interacting more with the existing customers helping find leads through them.