

X Education - Lead Scoring Case Study

QUES. - IDENTIFICATION OF HOT LEADS TO FOCUS MORE ON THEM AND THUS ENHANCING THE CONVERSION RATIO FOR X EDUCATION

BY –
VINIBA D PONRAJ
SAI VENKATESH
ANKIT KUMAR SINGH

X Education Company – (History)

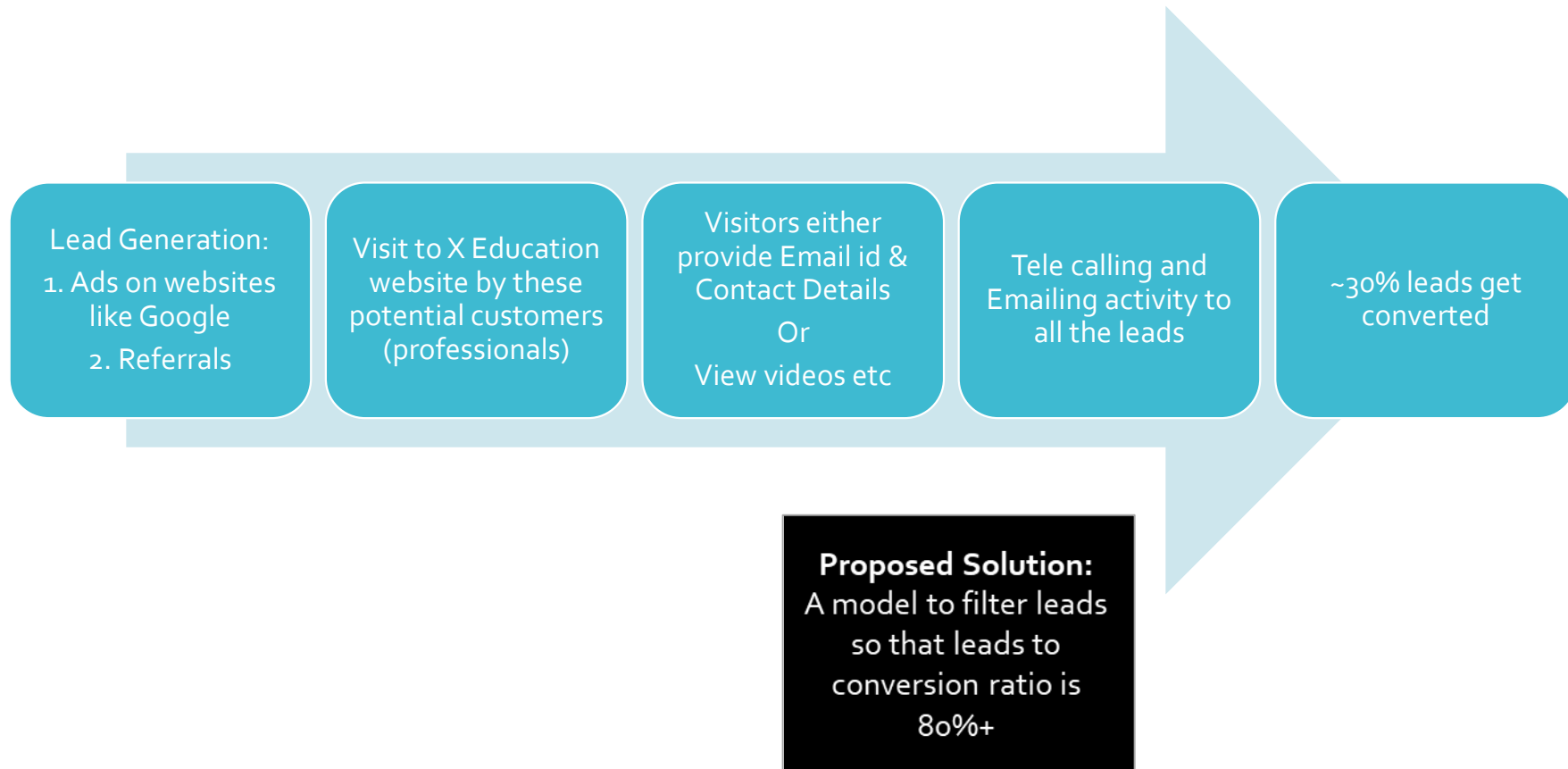
- X Education , An education company named sells online courses to industry professionals
- Many interested professionals land on their website
- The company markets its courses on several websites like Google. Once these people land on the website, they might browse the courses or fill up a form for 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
- 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 Statement

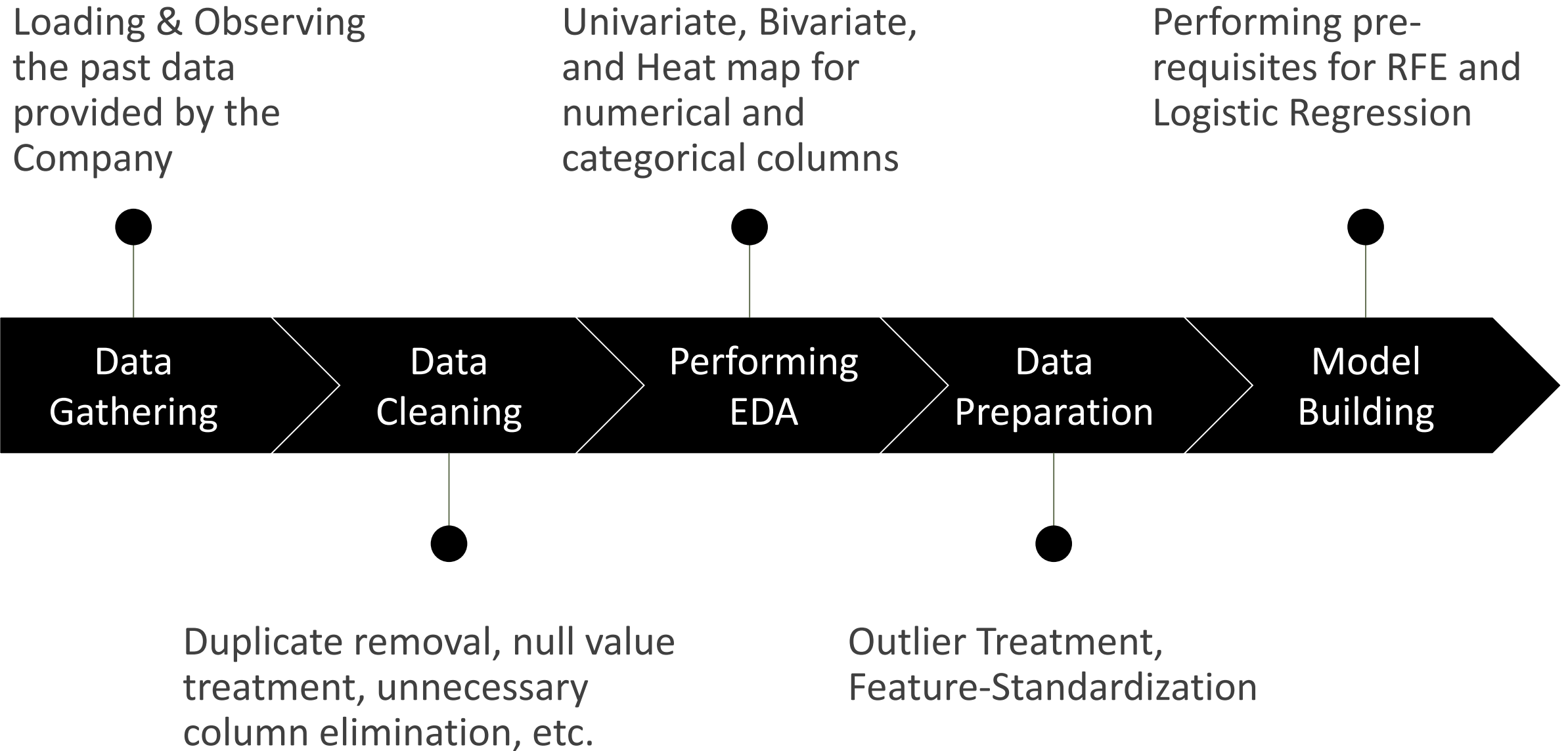
- X Education gets a lot of leads but its lead conversion rate is very poor
- To make this process more efficient, the company wishes to identify the most potential leads, also known as 'Hot Leads'
- If they successfully identify this set of leads, the lead conversion rate should go up as the sales team will now be focusing more on communicating with the potential leads rather than making calls to everyone
- We will help them to select the most promising leads, i.e. the leads that are most likely to convert into paying customers.
- We are required to build a model wherein we need to assign a lead score to each of the leads such that the customers with higher lead score have a higher conversion chance
- The CEO, in particular, has given a ballpark of the target lead conversion rate to be 80%.

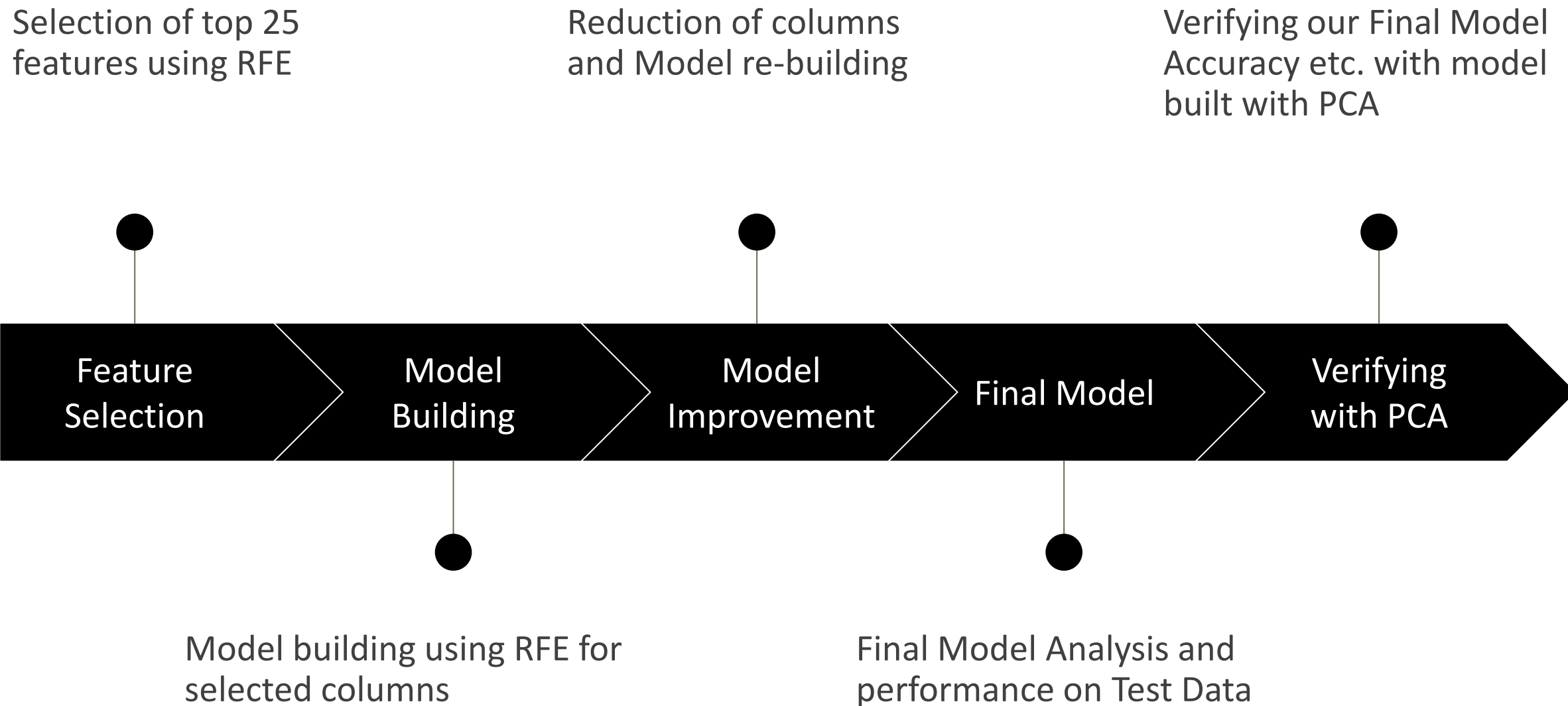
Lead – Conversion Process

Lead to Conversion
process



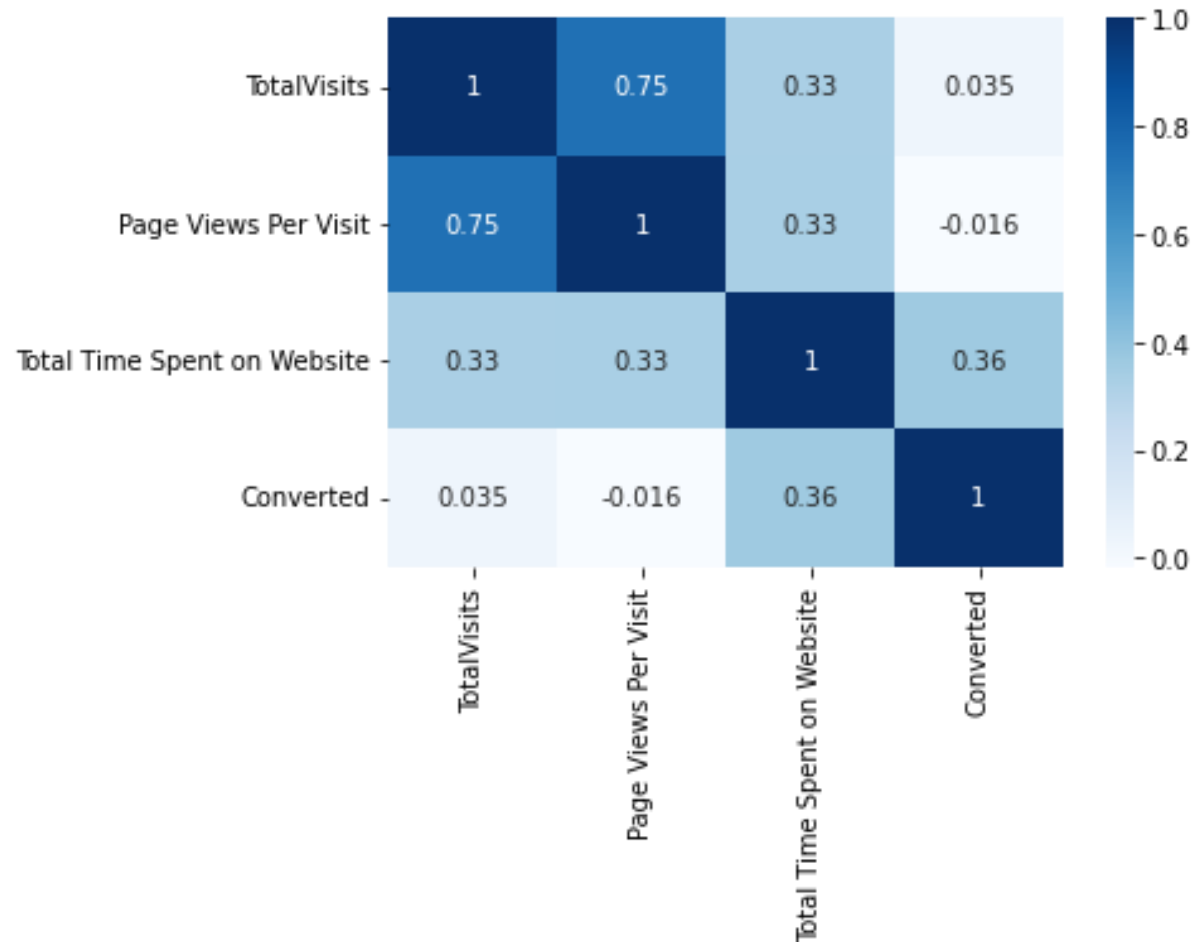
Execution

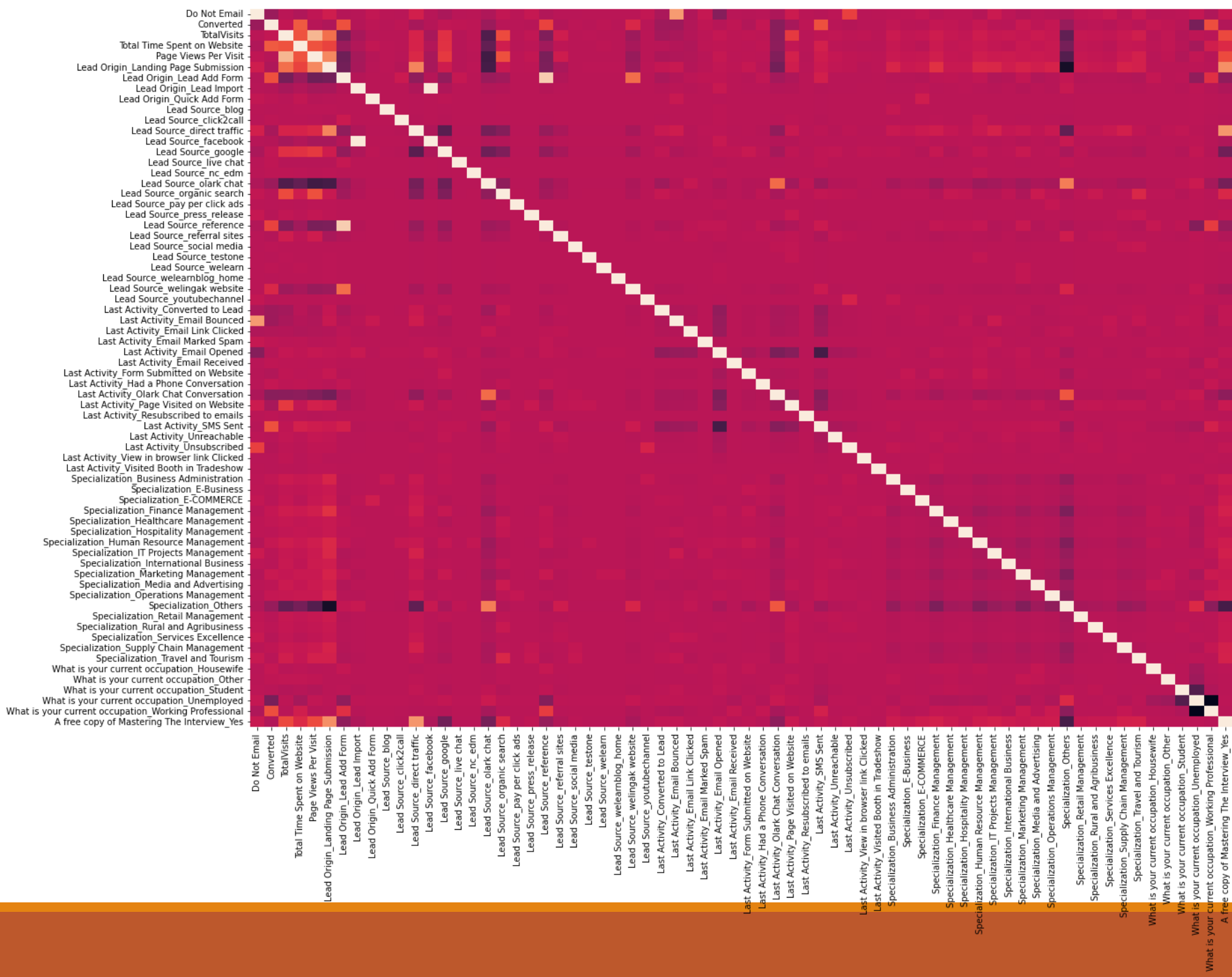




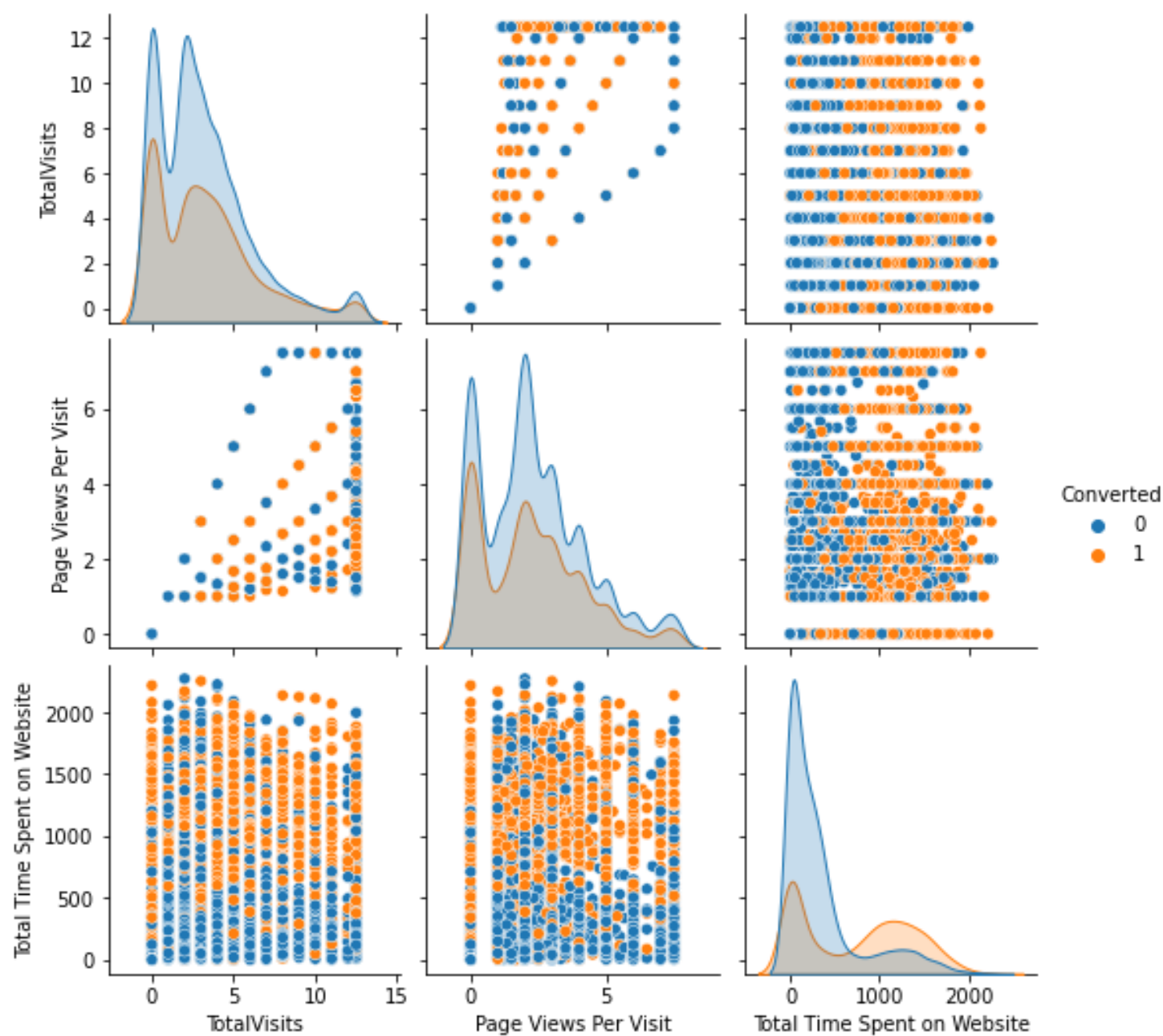
Plots (Visualization)

Heatmap to show correlation between numerical variables





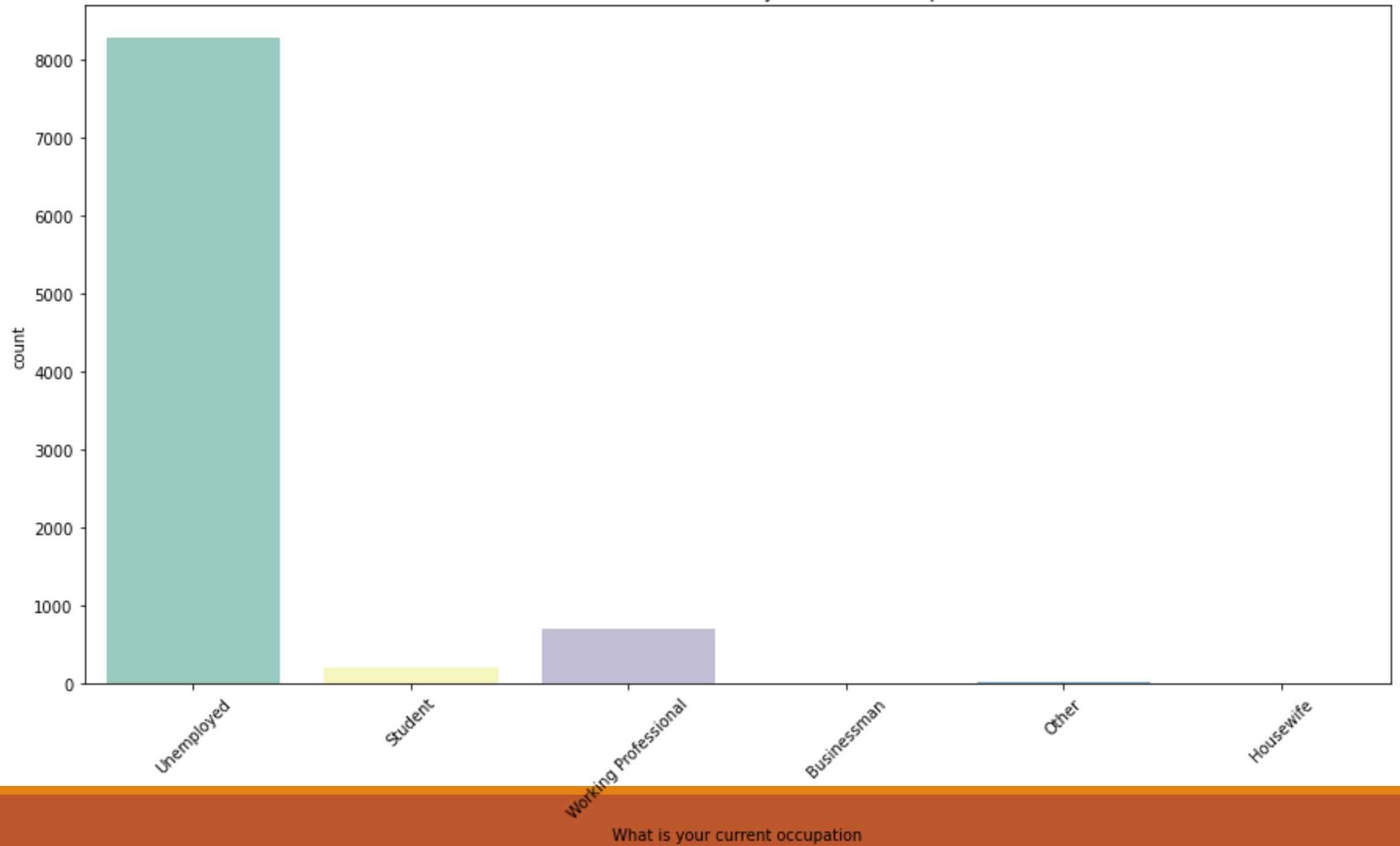
EDA plots depicting correlation (Heat Map) of all selected columns (numerical columns and dummy columns).



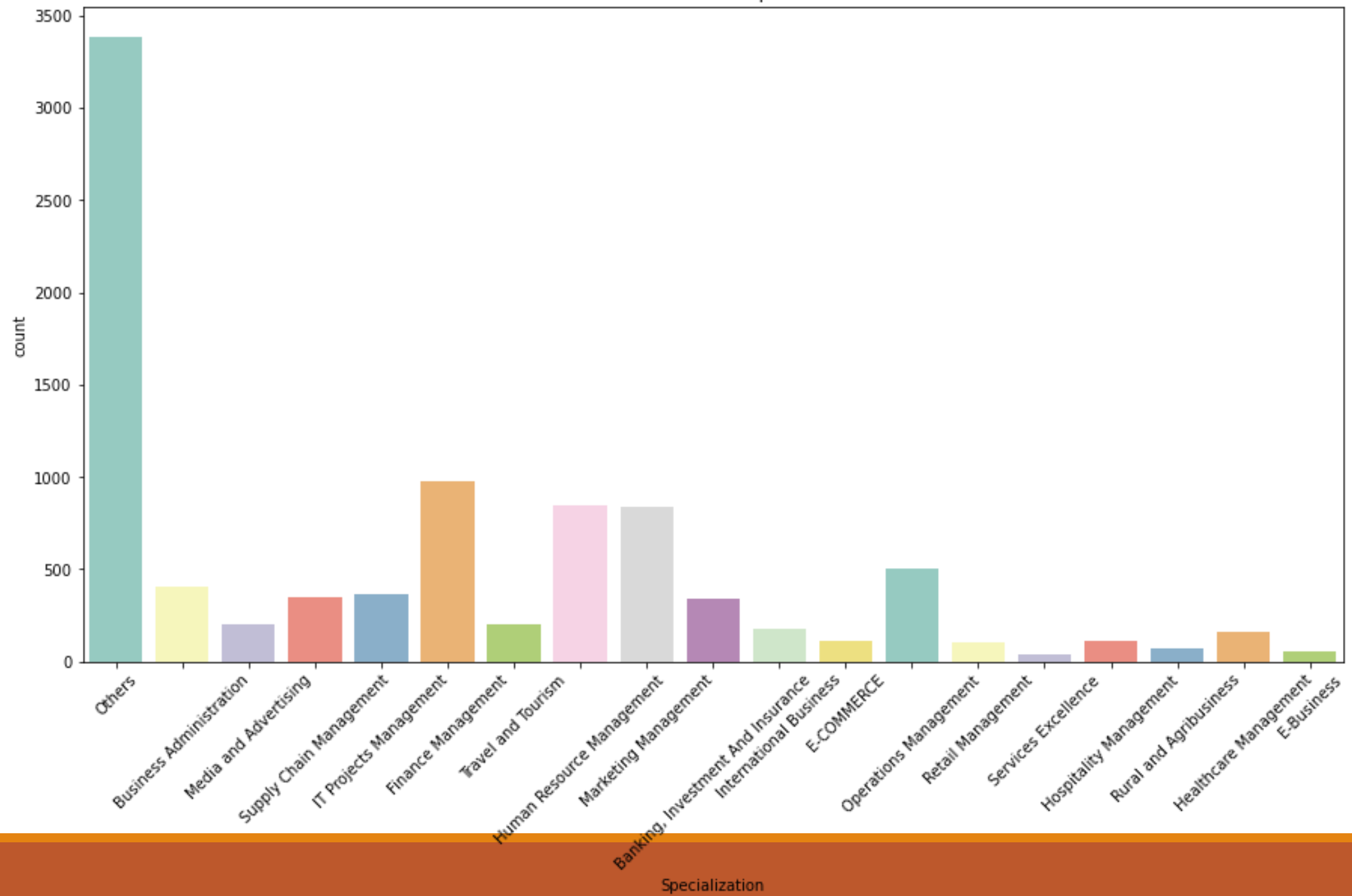
Bivariate analysis

Univariate analysis on categorical variable

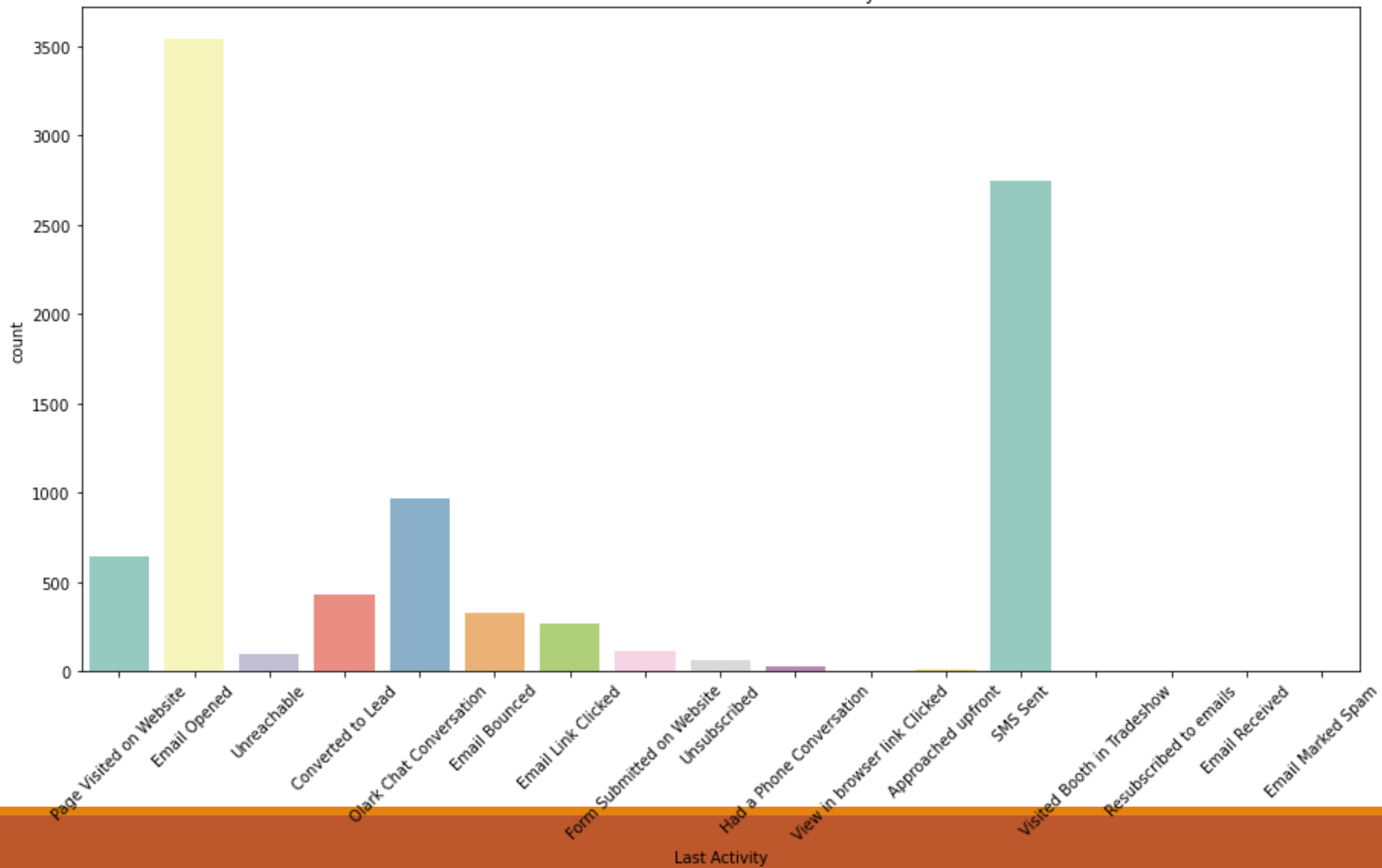
Count Distribution of What is your current occupation



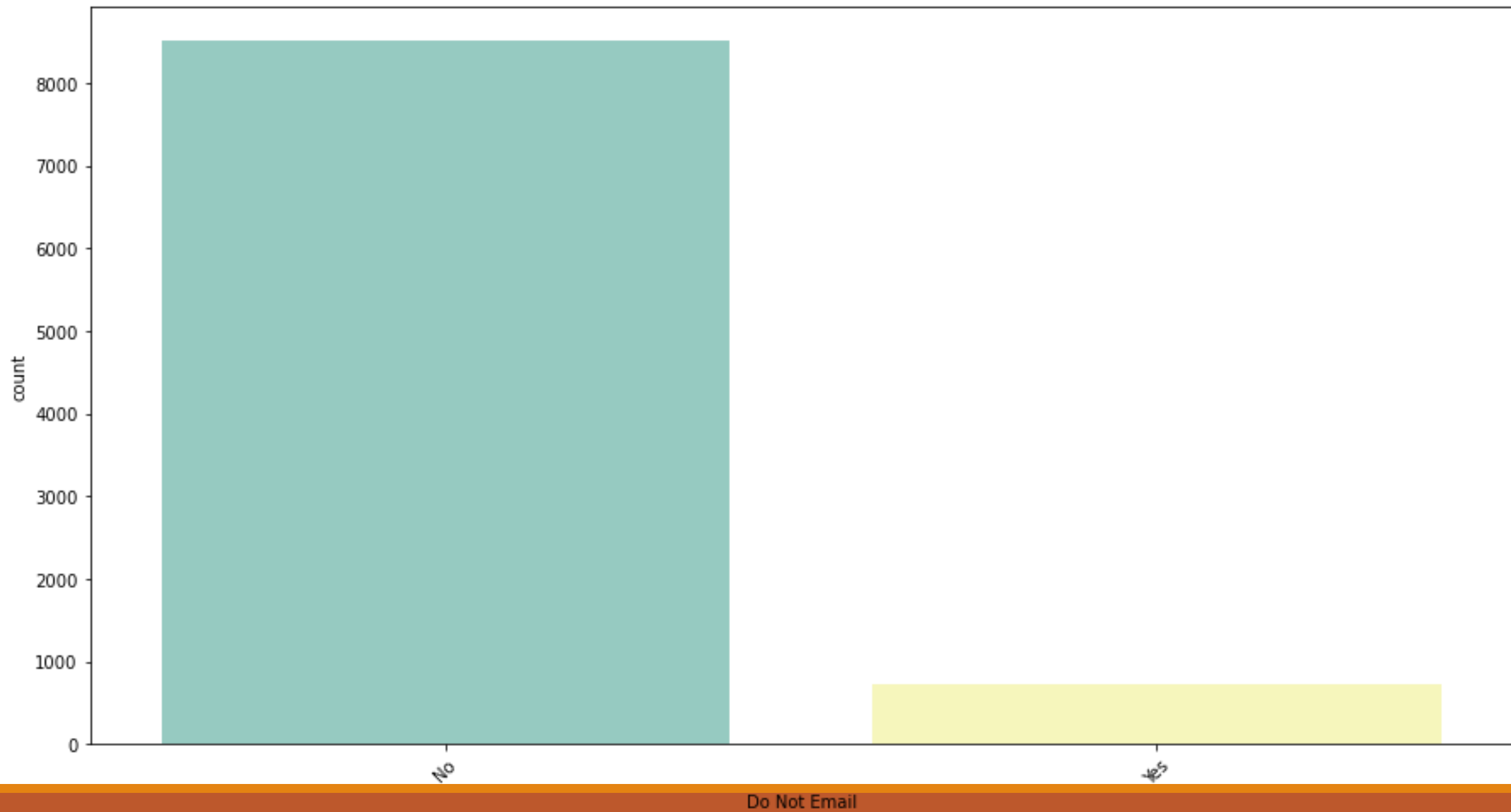
Count Distribution of Specialization



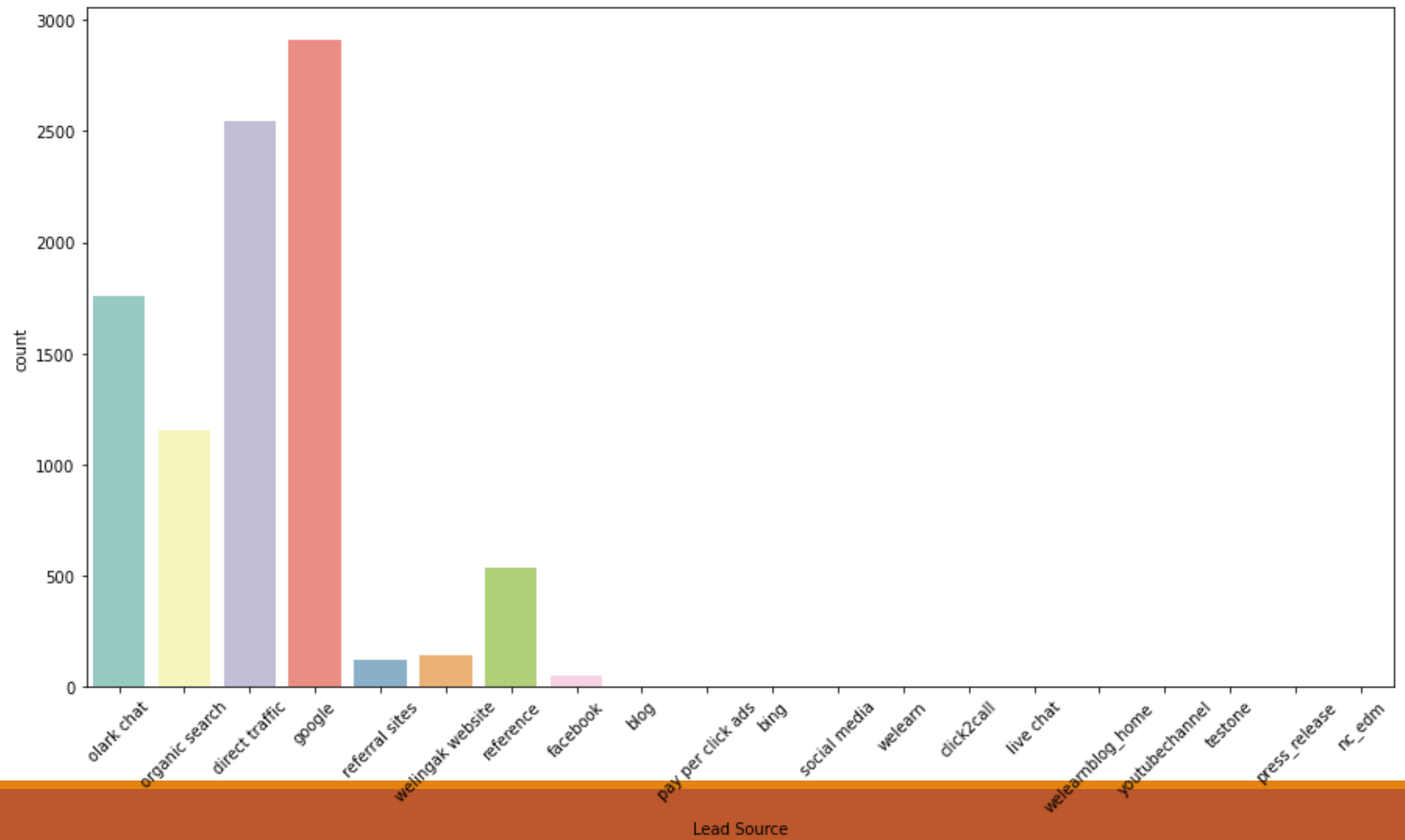
Count Distribution of Last Activity



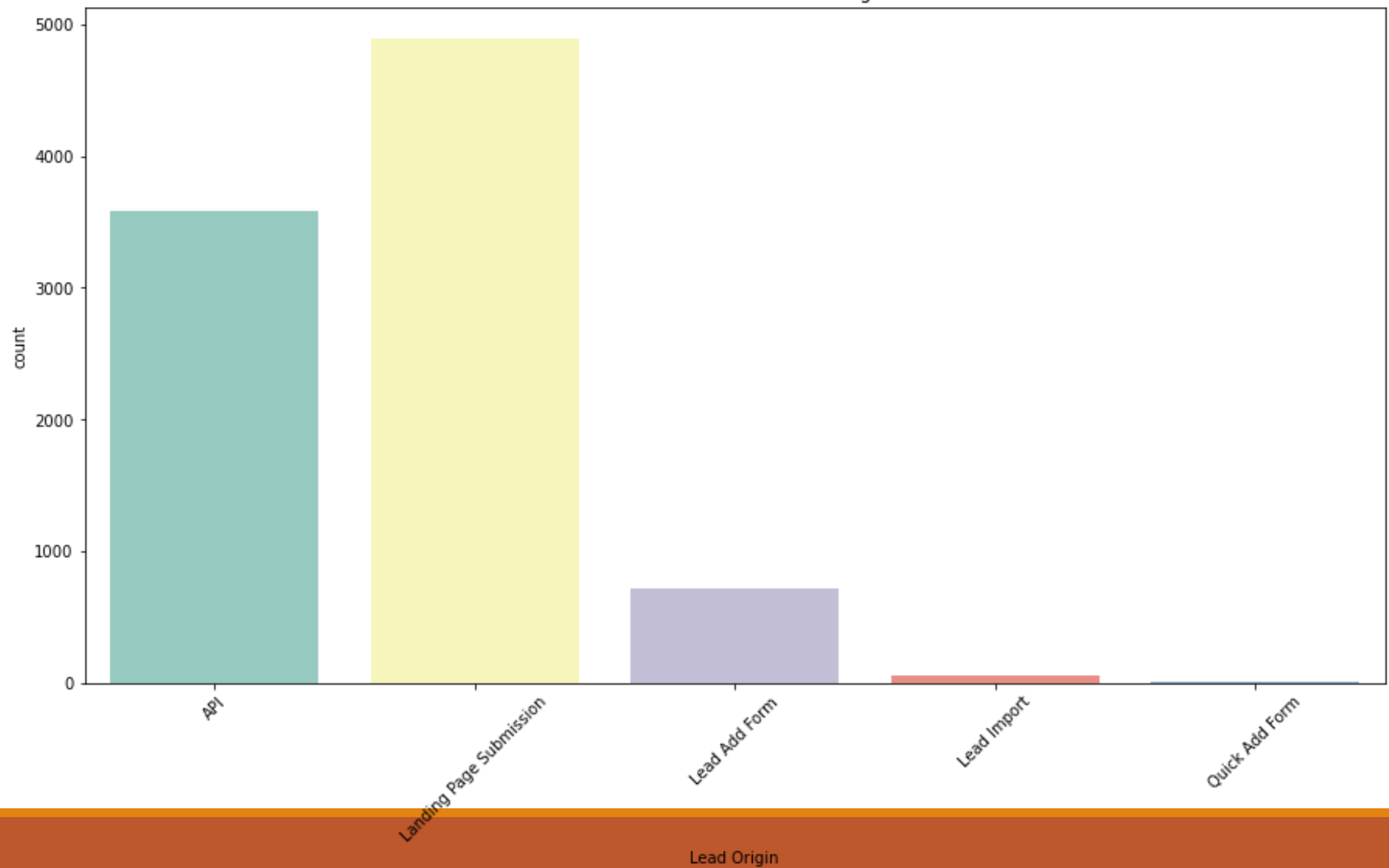
Count Distribution of Do Not Email



Count Distribution of Lead Source



Count Distribution of Lead Origin



Inference / Conclusion

Model Analysis (Performance of our Final Model)

- Overall accuracy on Test set: 0.781
- Sensitivity of our logistic regression model: 0.768
- Specificity of our logistic regression model: 0.790

Inferences from Model (Business Insights Derived from our Model)

Top 3 variables in model, that contribute towards lead conversion are:

- Total Time Spent on Website
- Last Notable Activity_SMS Sent
- TotalVisits

Top 3 variables in our model, that should be focused are:

- Last Activity_SMS Sent (positively impacting)
- Last Activity_Olark Chat Conversation (negatively impacting)
- Lead Source_Olark Chat (negatively impacting)

Conclusion

Our Logistic Regression Model is decent and accurate enough, when compared to the model derived using PCA, with 78.1 % Accuracy on Test Set, 76.8 % Sensitivity and 79.01 % 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.

Recommendation

It's good to collect data often and run the model and get updated with the potential leads. There is a belief that the best time to call your potential leads is within few hours after the lead shows interest in the courses.

- It's good to understand about lead from the available profile information and group similar kind of leads.
- Along with phone calls, it's good to mail the leads also to keep them reminding as email is as powerful as cold calling.
- While mailing, it is good to send personalized message for different set of leads as it will have positive impact on leads.
- Generally, it take 5 - 6 calls to successfully convert a lead, so if we focus on reducing it to 3 to 4 attempts by taking appointments and calling, sending emails and providing right information and keep the leads in touch, there will be more time available which can be used to convert still more leads