X Education - Lead Scoring Case Study

PRESENTED BY

RAMYASREE NARWA

SAI PRANAVI KOTHUR

Background

- An education company named X Education sells online courses to industry professionals.
- The company markets its courses on several websites and search engines like Google. Once people land on the website, they might browse the courses or fill up a form for the course or watch some videos.
- 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.

Problem Statement

- Although X Education gets a lot of leads, 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.

Problem Statement

- Although X Education gets a lot of leads, 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.
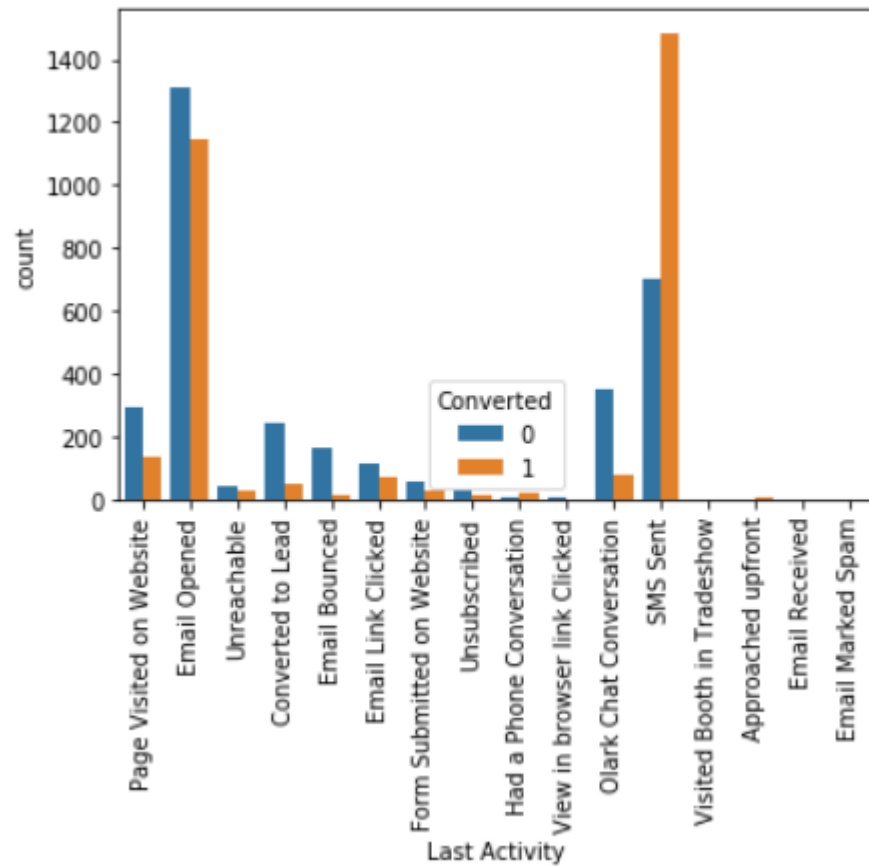
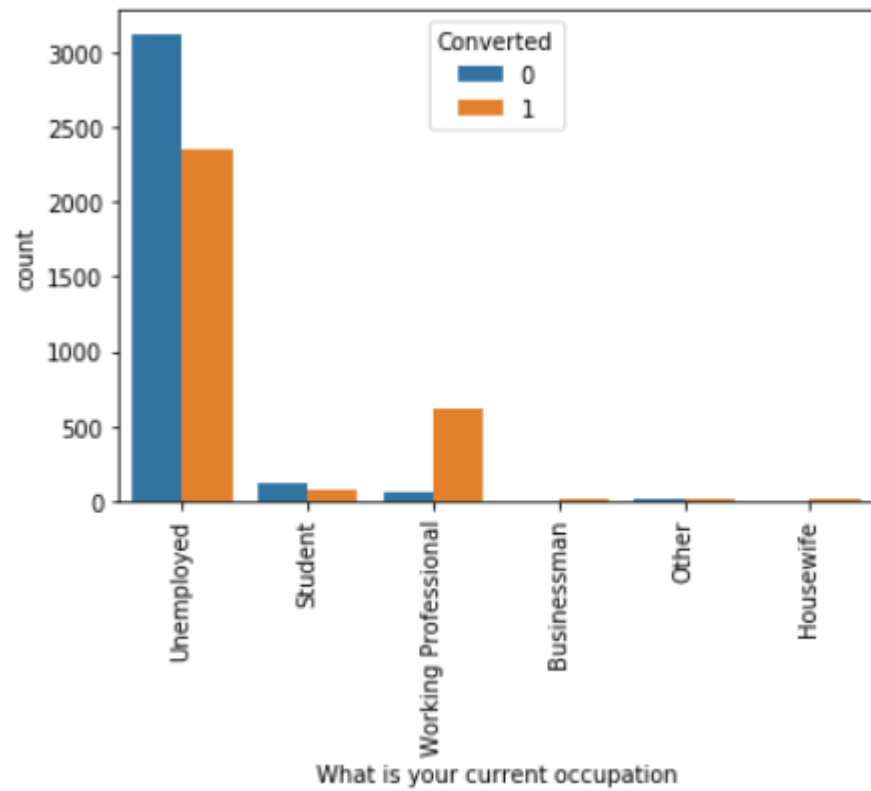# EDA(Exploratory Data Analysis)
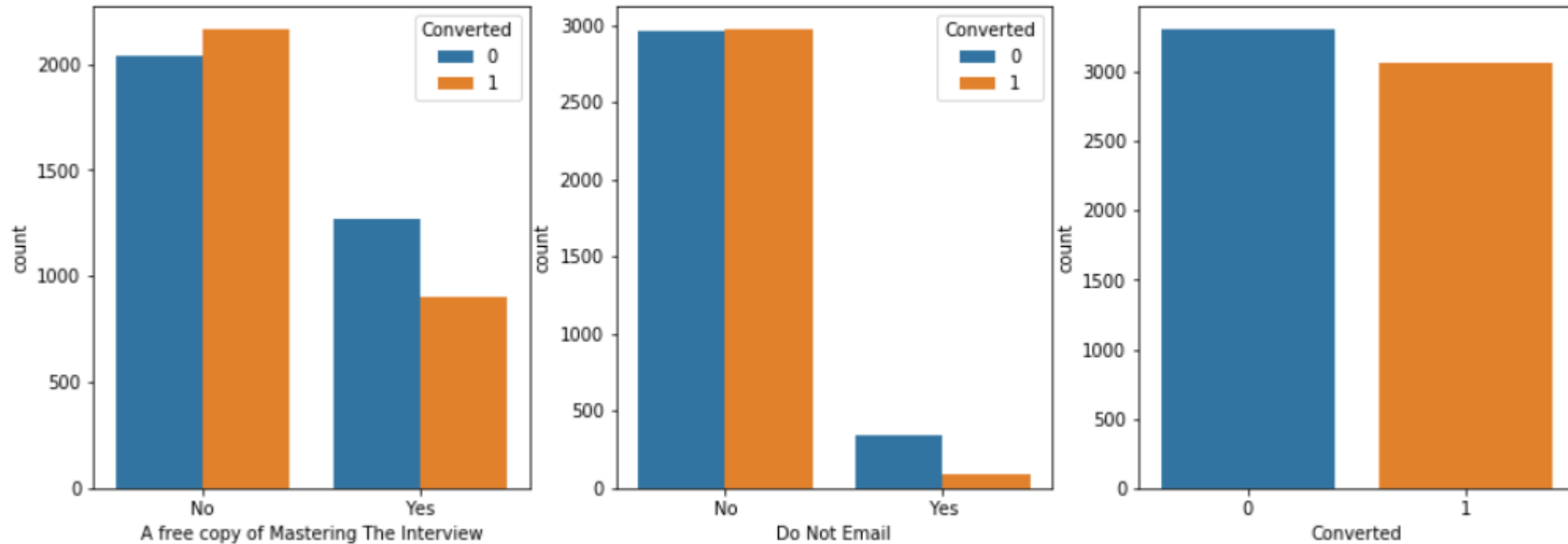
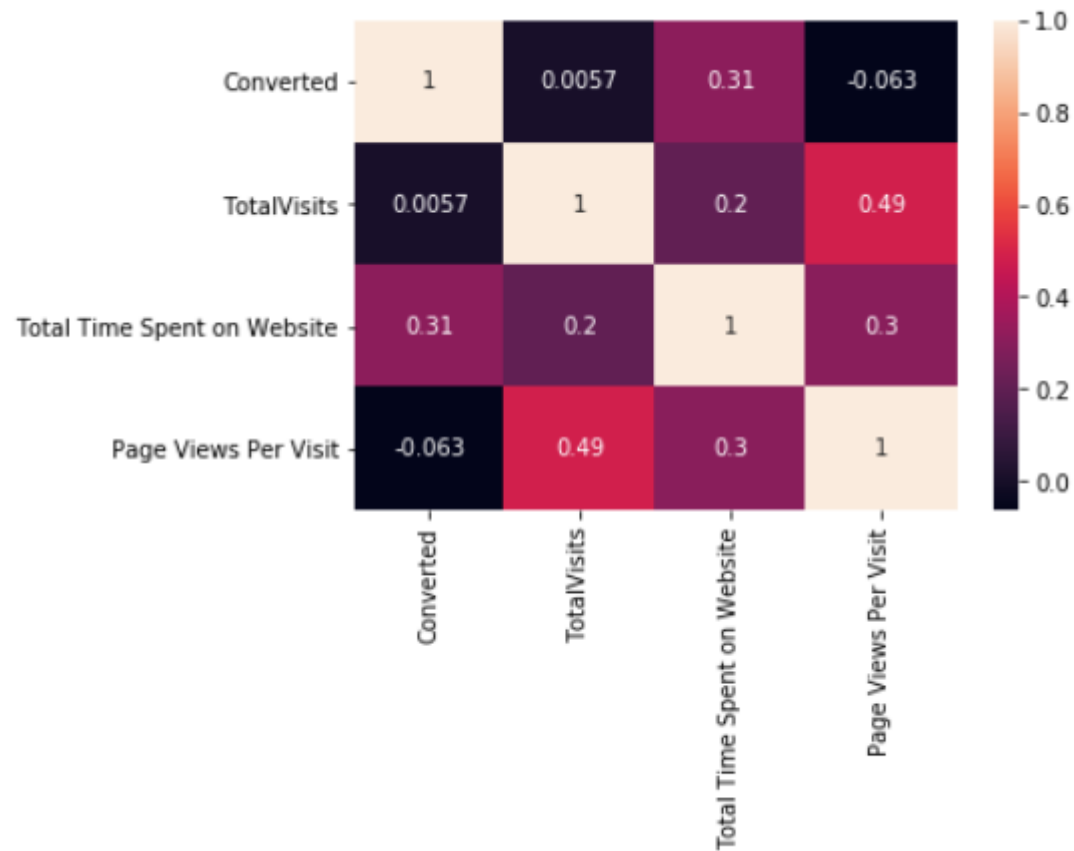- Analysis of different variables for converted parameter
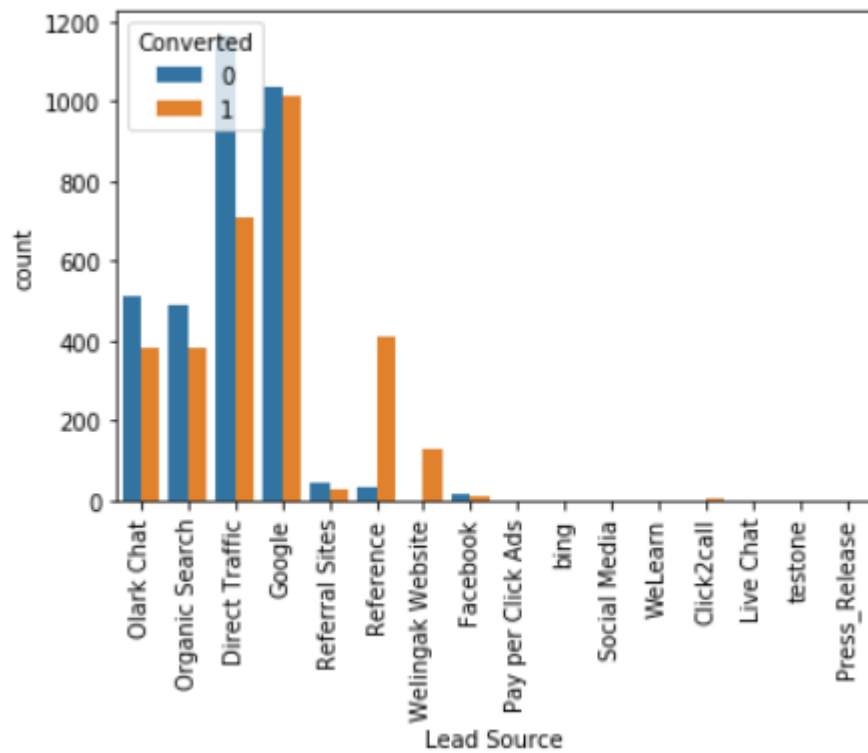
- For Lead source Parameter, analyzing the different sources and conversion rate. Google being source has high conversion rate.
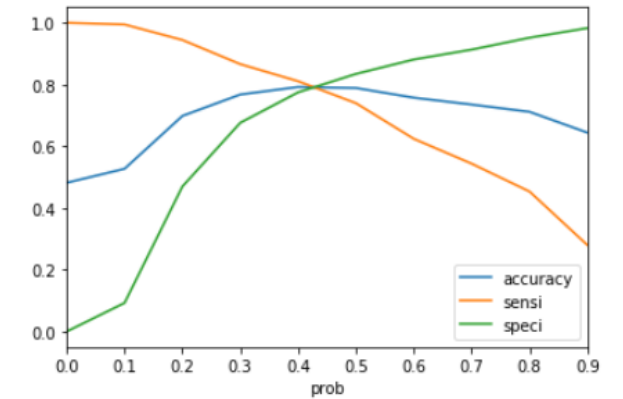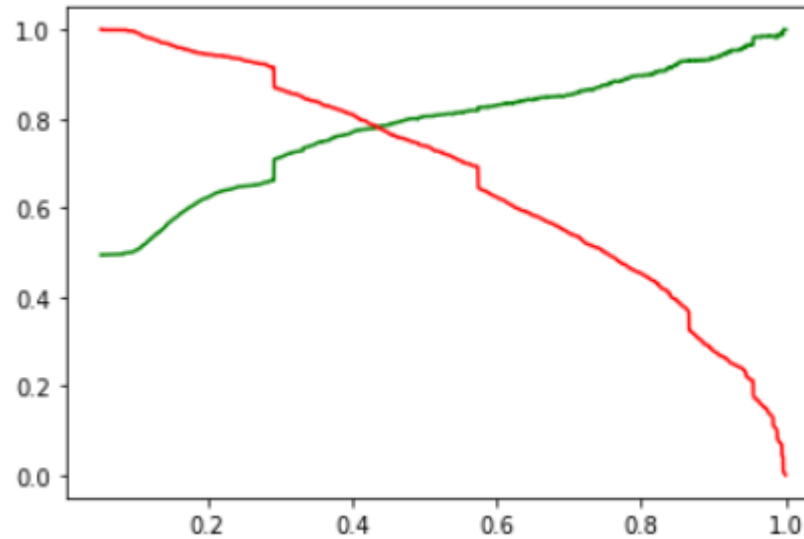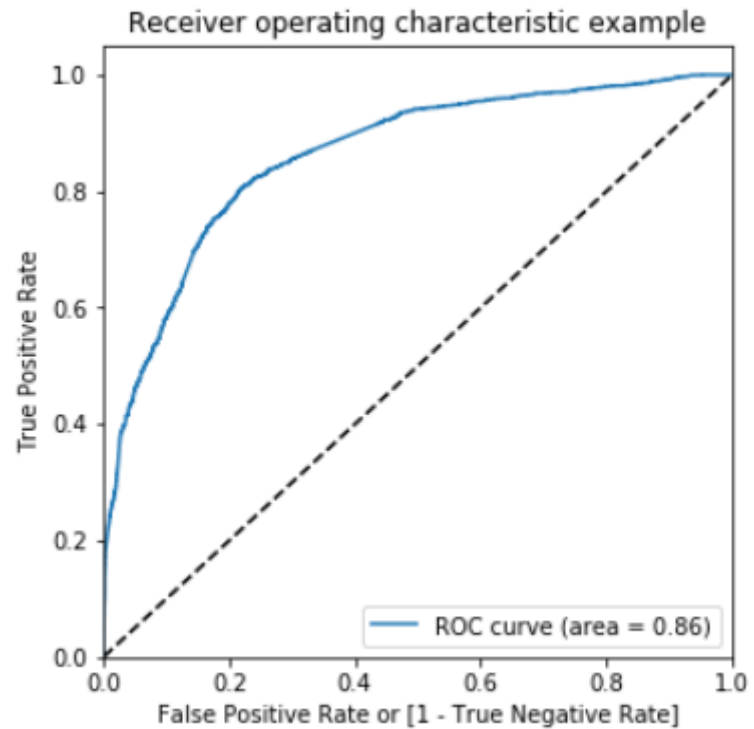
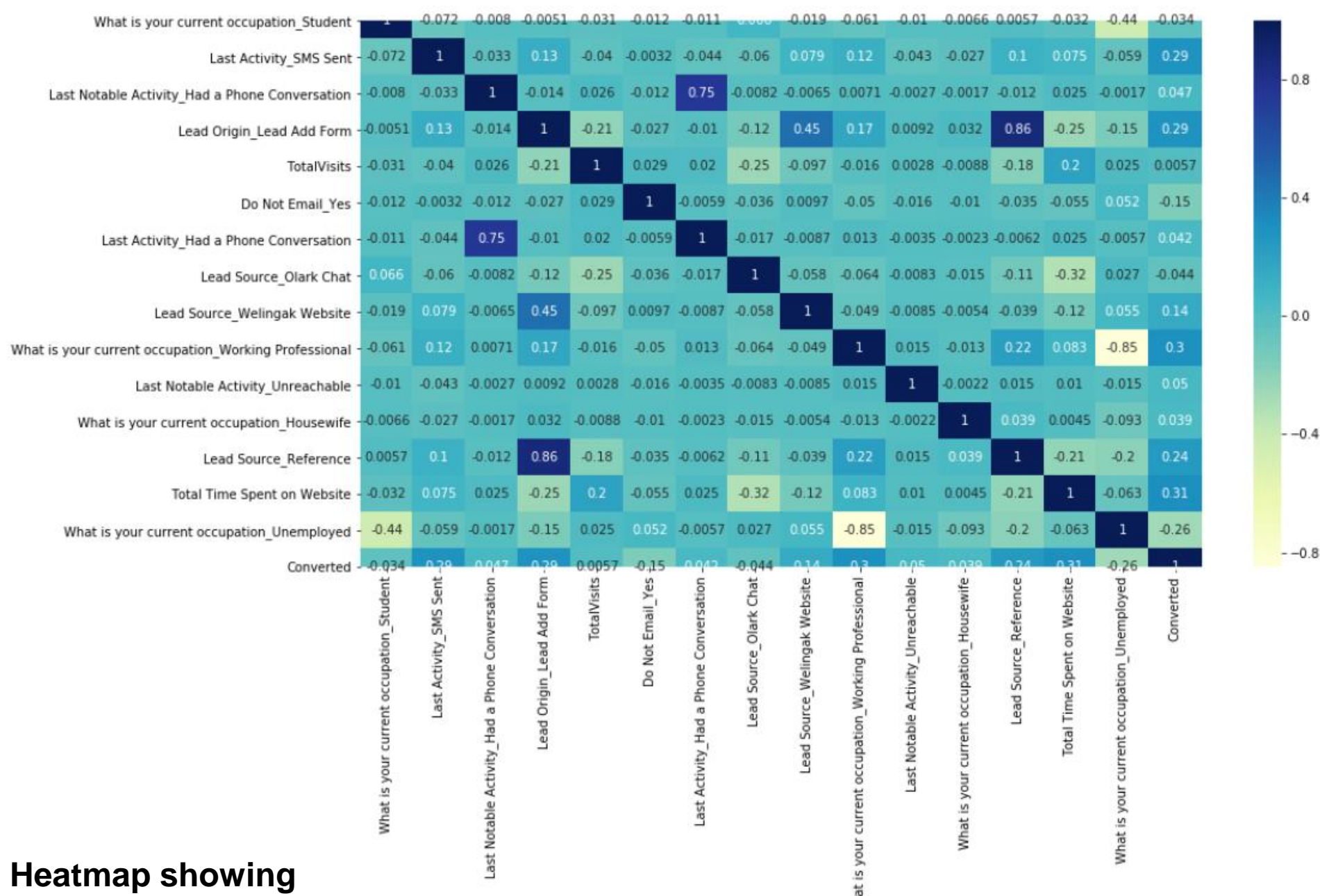- For Last Activity Parameter, analyzing the different last activities.

- Analyzing the conversion rate based on occupation.

- Analyzing the co-relation between variables.
- Highest co-relation exists between Total Visits and Page views Per visit.

Linear Regression final model parameters:
Area under ROC: 0.86
Immediate cutoff: 0.42
Final cutoff: 0.43

Heatmap showing correlation of all selected columns in our final Model.

Conclusion

Top 3 Variables in model:
1. Total Visits
2. Total Time Spent on Website
3. Lead Origin_Lead Add Form

Top 3 variables that need focus are:
1. What is your current occupation_unemployed(negatively impacting)
2. Do not email_yes(negatively impacting)
3. Total Time spent on website(positively impacting)

X Education should

Increase user engagement on website, engage the users to further increase total visits and time spent on website as this will increase the conversion rate

Concentrate on attracting the unemployed users as this is negatively impacting the cor