

Lead Scoring Assignment

SUBMITTED BY ->GAURAV SHARMA

Contents

- ❑ Problem statement
- ❑ Exploratory Data Analysis
- ❑ Correlations
- ❑ Model Evaluation
- ❑ Conclusion

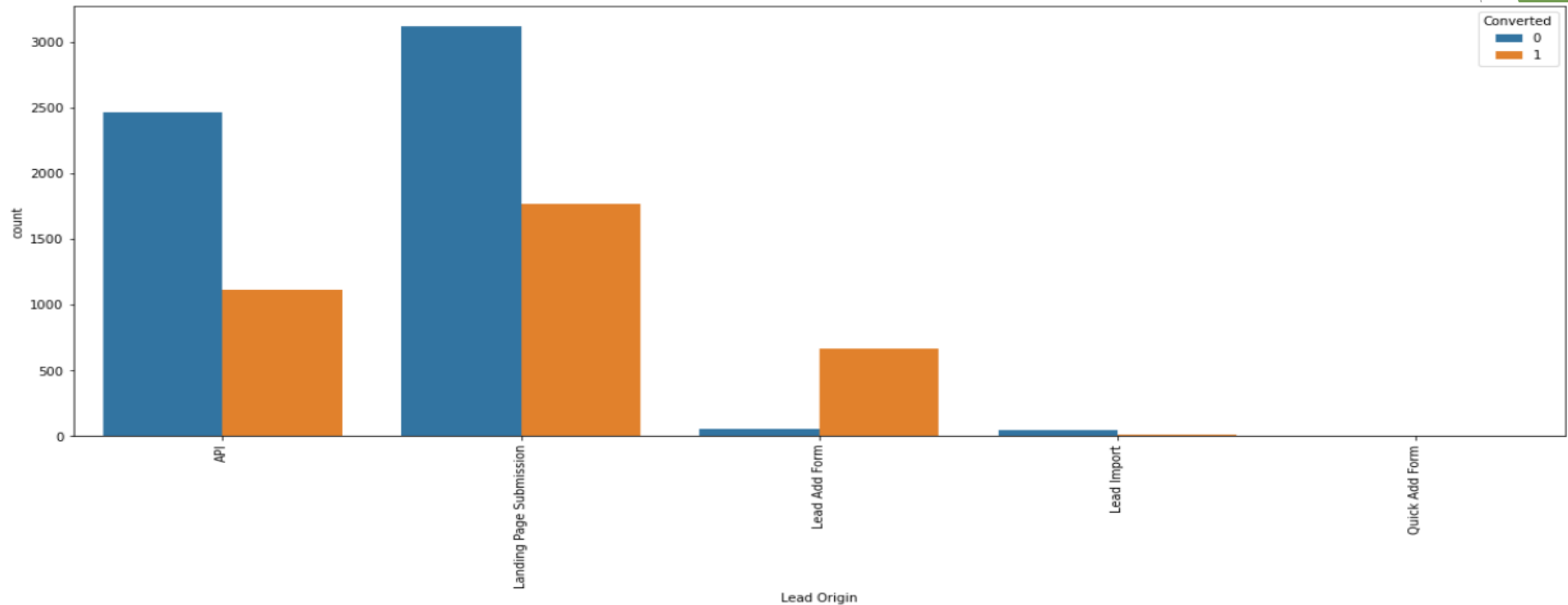
Problem statement

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%.

Now, although X Education gets a lot of leads, its lead conversion rate is very poor. For example, if, say, they acquire 100 leads in a day, only about 30 of them are converted. 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

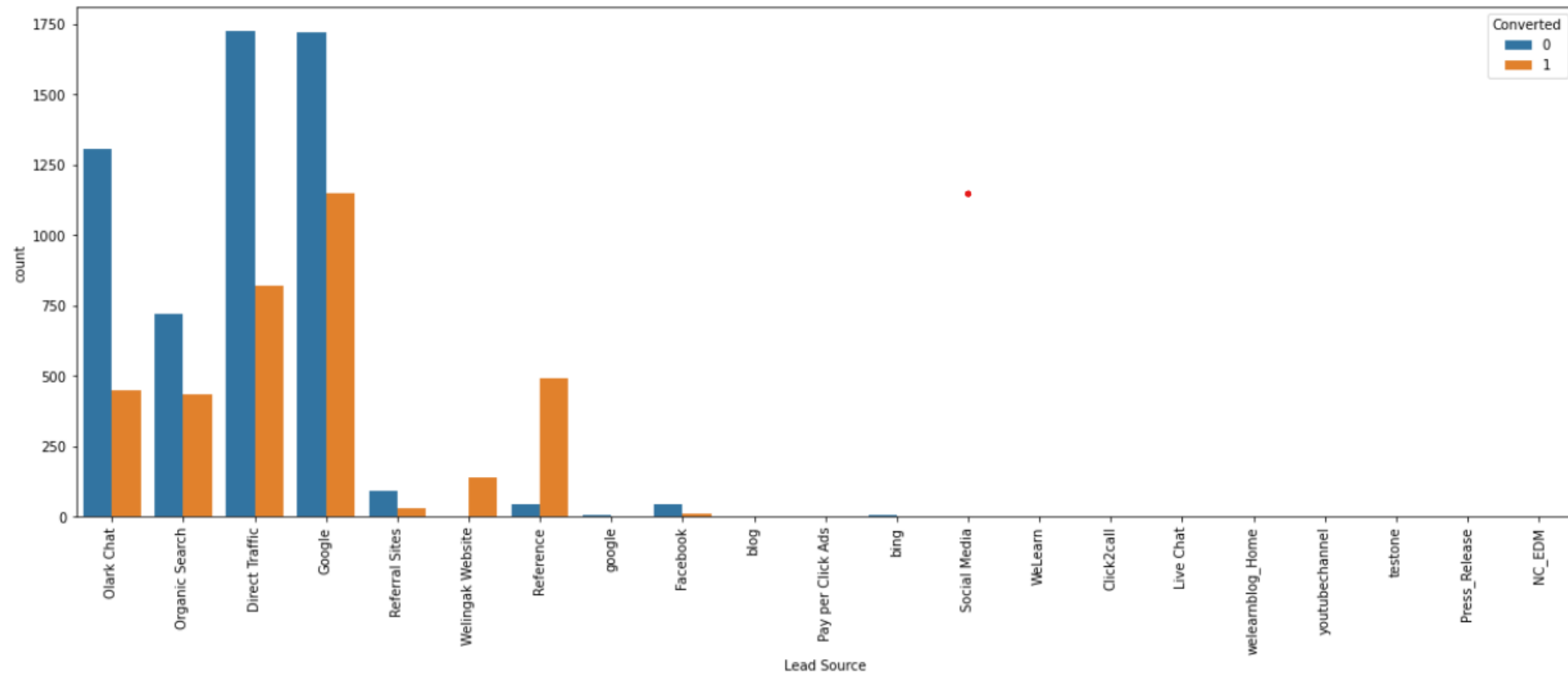
Exploratory Data Analysis



INFERENCES

Lead Add form has the highest conversion rate among all other lead origin.

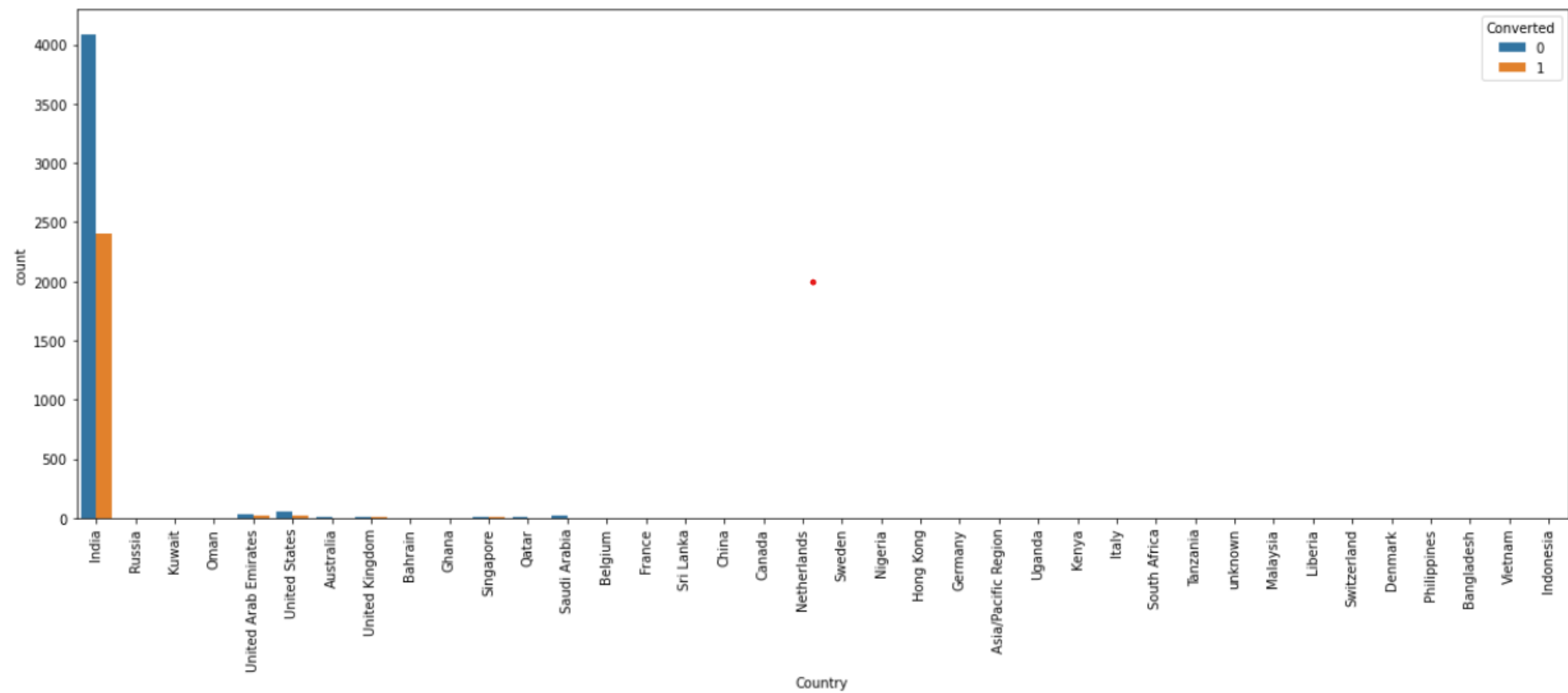
Lead Source



INFERENCES

Reference Lead source has the best rate of conversion.

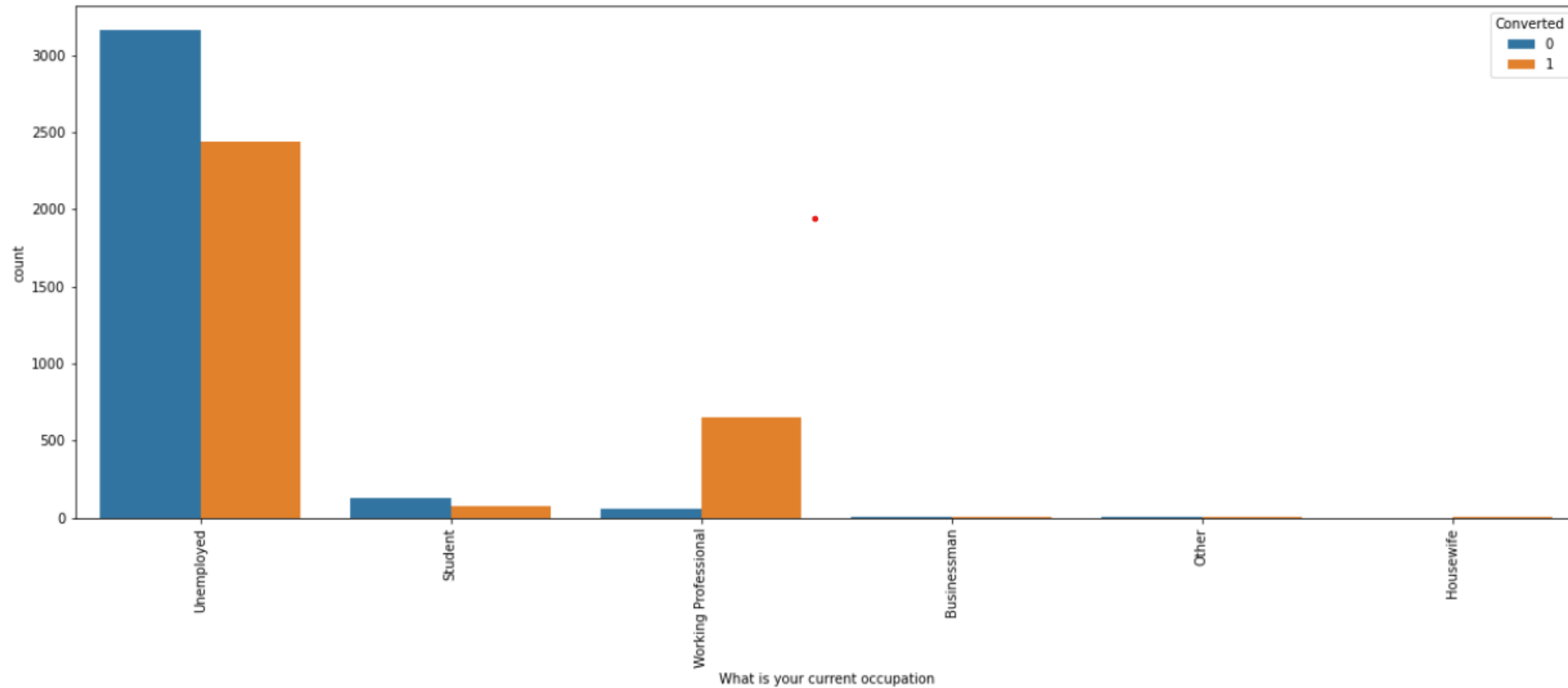
Country



INFERENCES

Approx 95% leads are from india only.

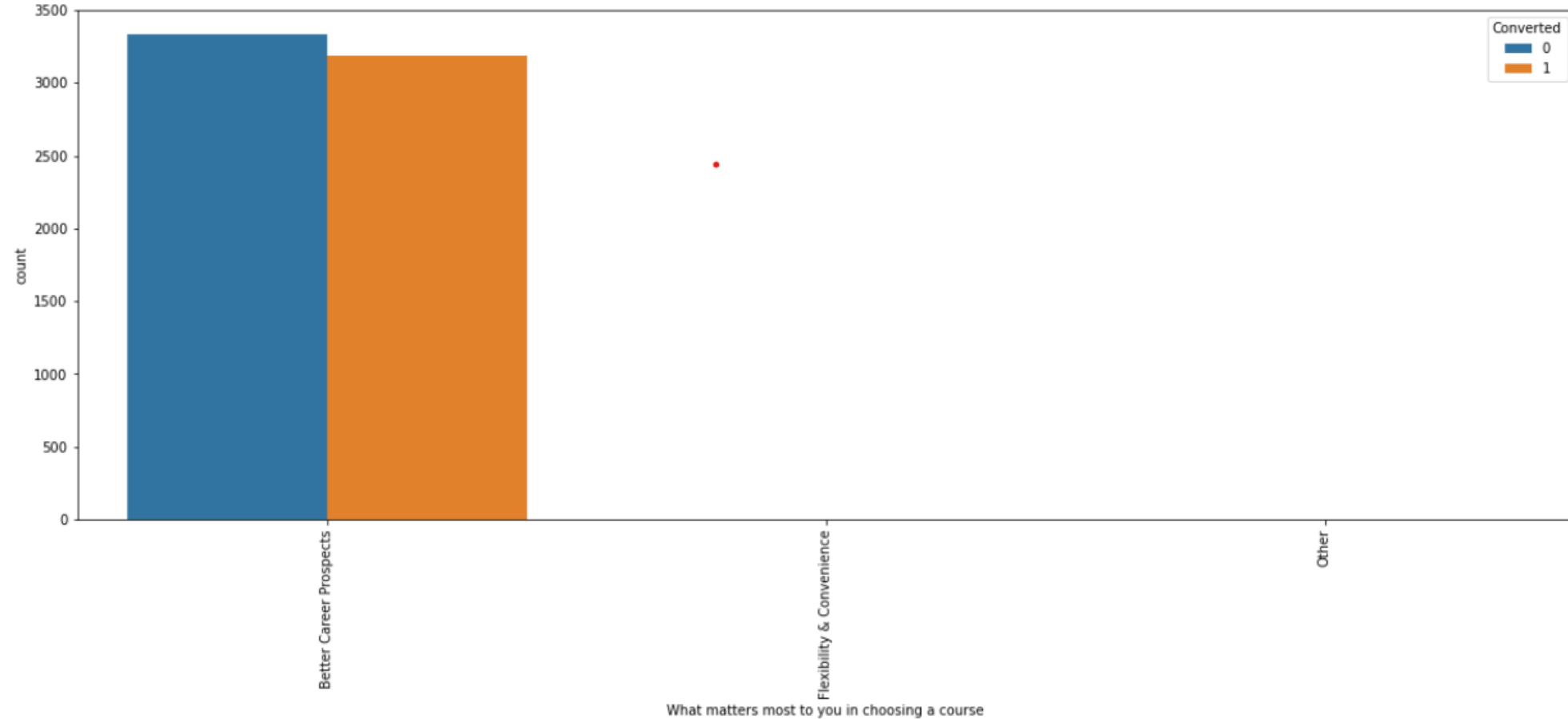
What is your current occupation



INFERENCES

Mostly leads are unemployed.

What matters most to you in choosing a course

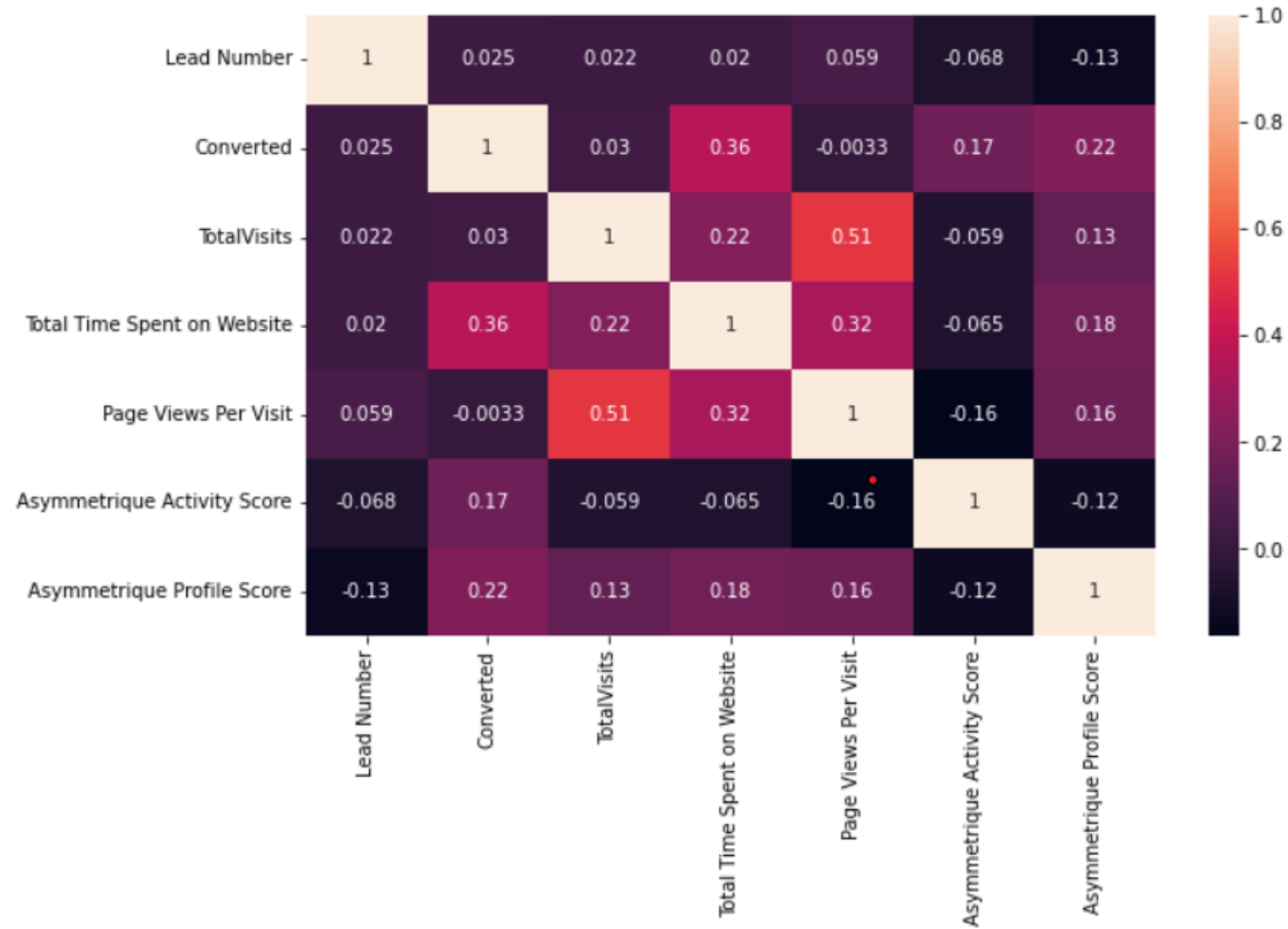


INFERENCES

Mostly leads have goal of better career Prospects .

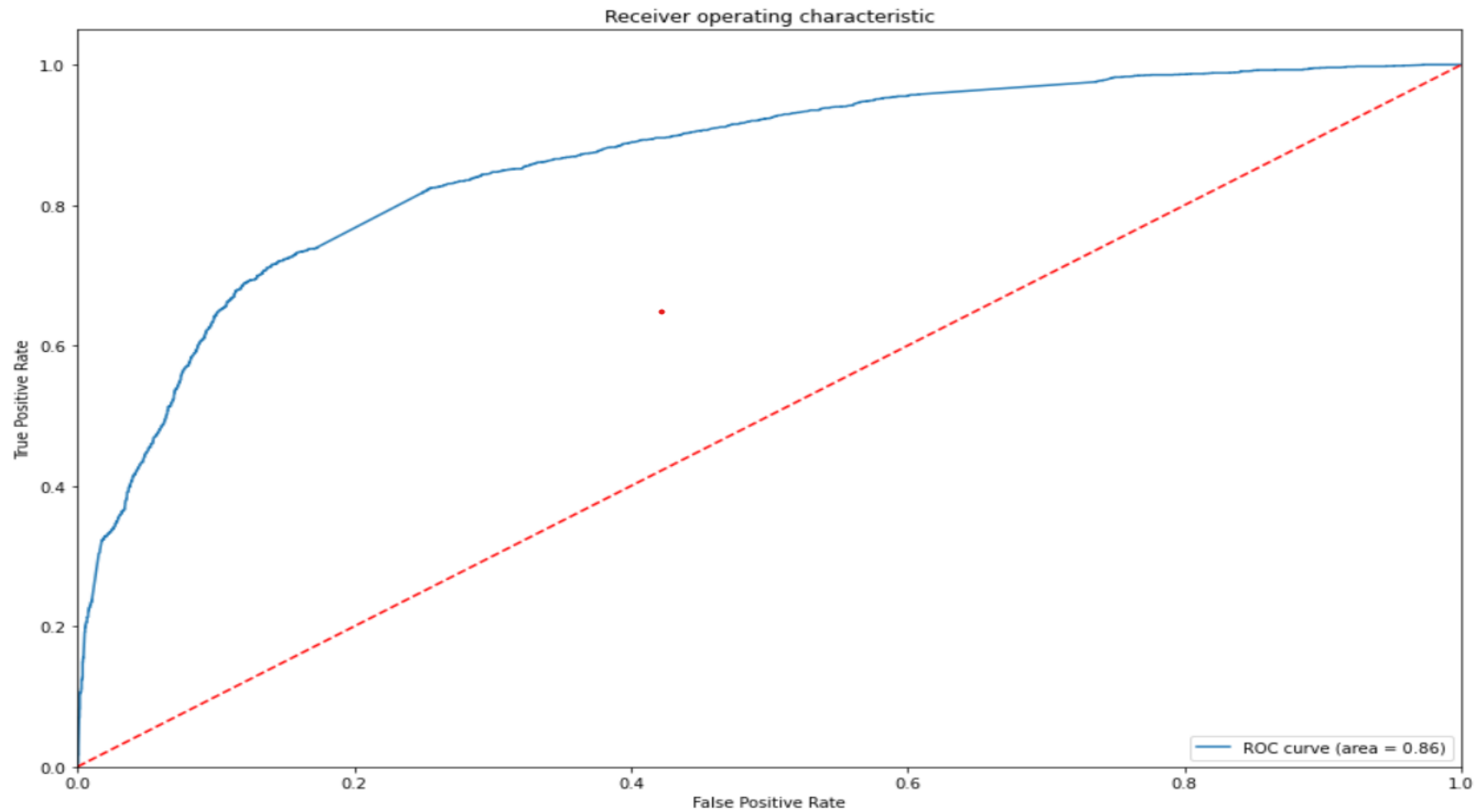
Correlation

<AxesSubplot:>

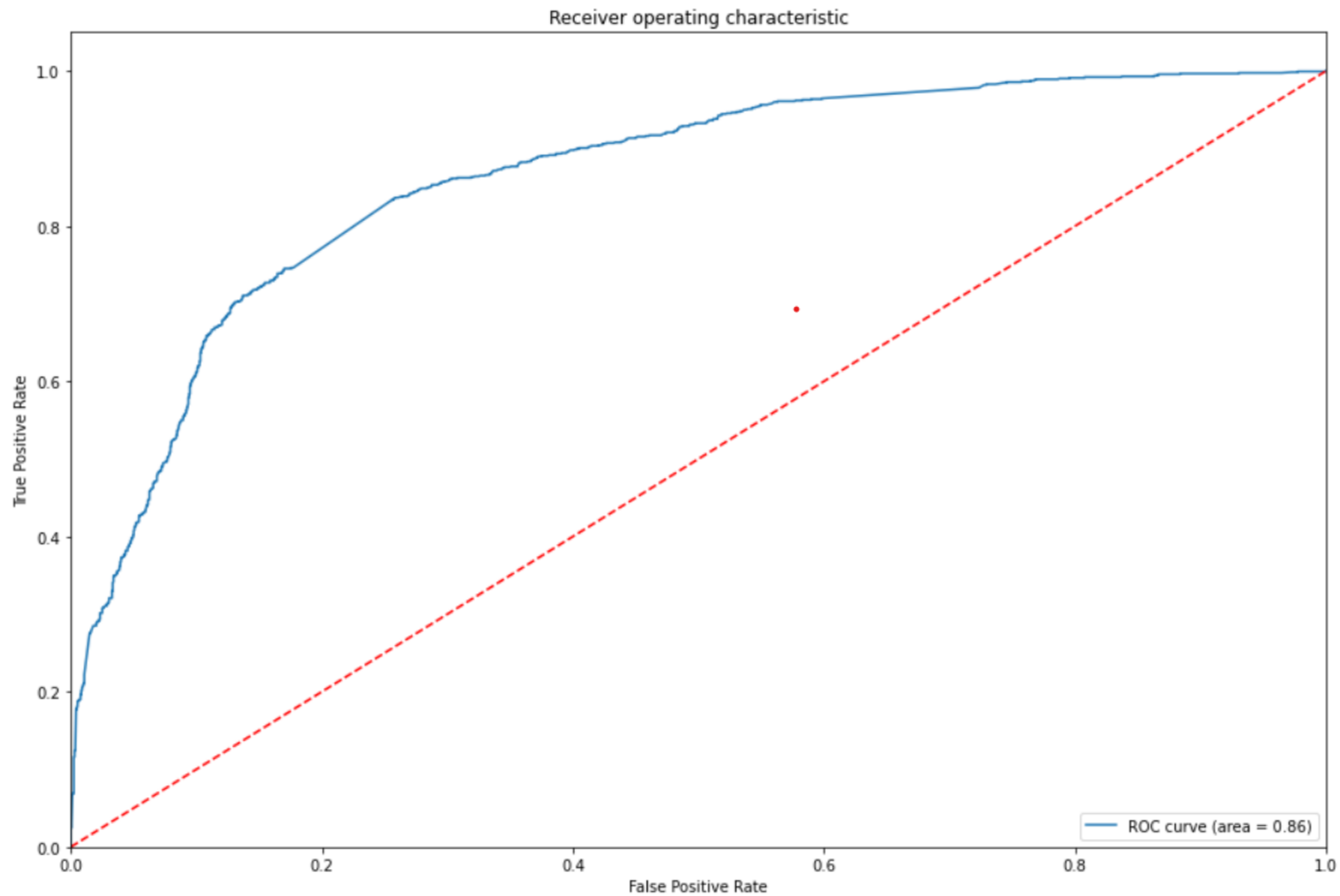


Model Evaluation

Model Evaluation on Train Dataset



Model Evaluation on Train Dataset



Conclusion

- 1)Lead Origin_Lead Add Form
- 2)Total Time Spent on Website
- 3)What is your current occupation

These all above features are the most important features in lead conversion aspect.