LEAD SCORE CASE STUDY

Presented by
Somee Gupta

Problem Statements:
X is education company which sell online courses They have leads but there conversion rate is poor they 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.

Business Objective:

- X education wants to know most promising leads.
- For that they want to build a Model which identifies the hot leads.
- Deployment of the model for the future use

Solution:-

Data cleaning and data manipulation.

1. Check and handle duplicate data.

2. Check and handle NA values and missing values.

3. Drop columns, if it contains large amount of missing values and not useful for the analysis.

4. Imputation of the values, if necessary.

5. Check and handle outliers in data

Univariate and Bivariate data analysis:

1. Value count, distribution of variable etc.

2. Correlation coefficients and pattern between the variables etc.

Classification technique:

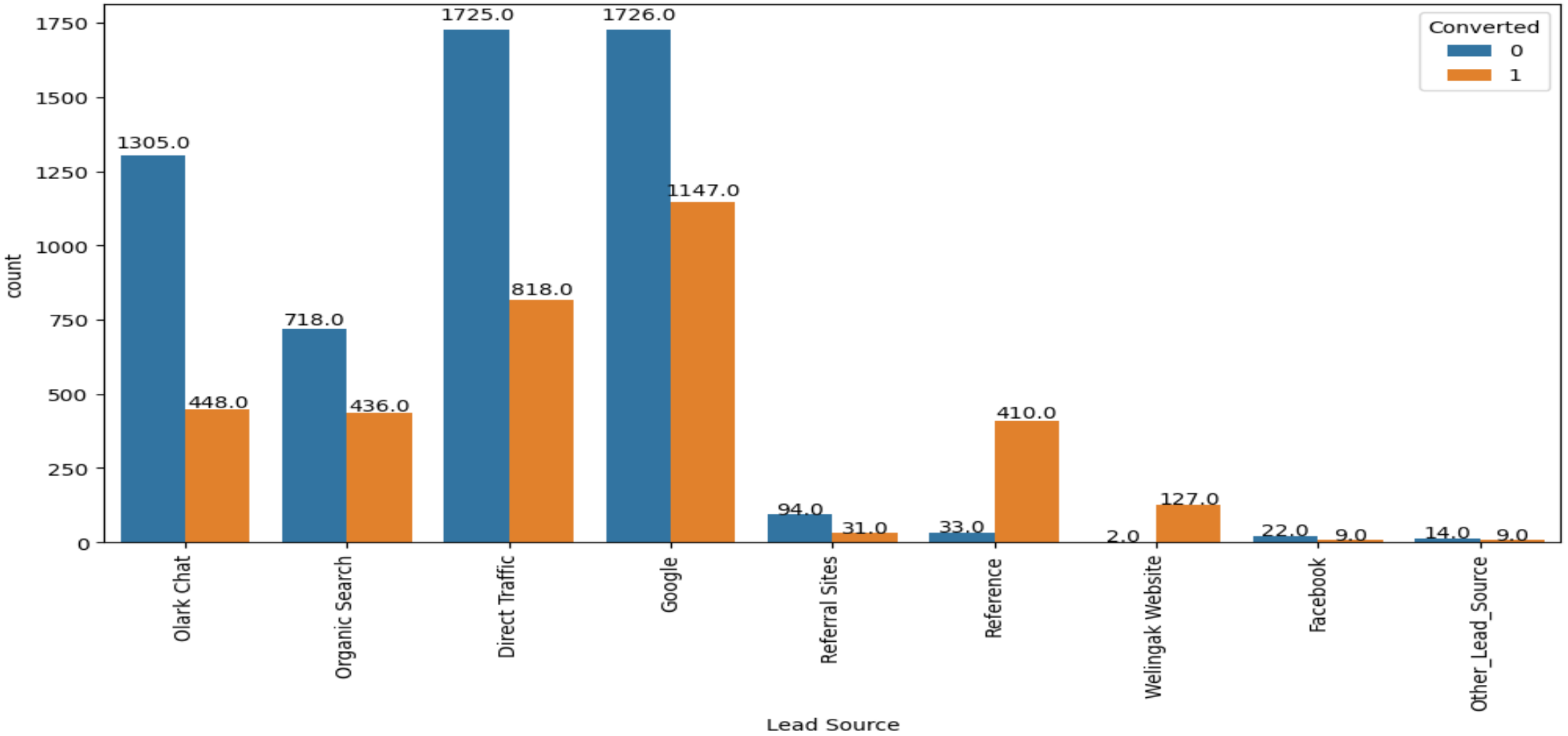
1.Logistic regression used for the model making and prediction.

2. Feature Scaling & Dummy Variables and encoding of the data.

3. Validation of the model. Model presentation.

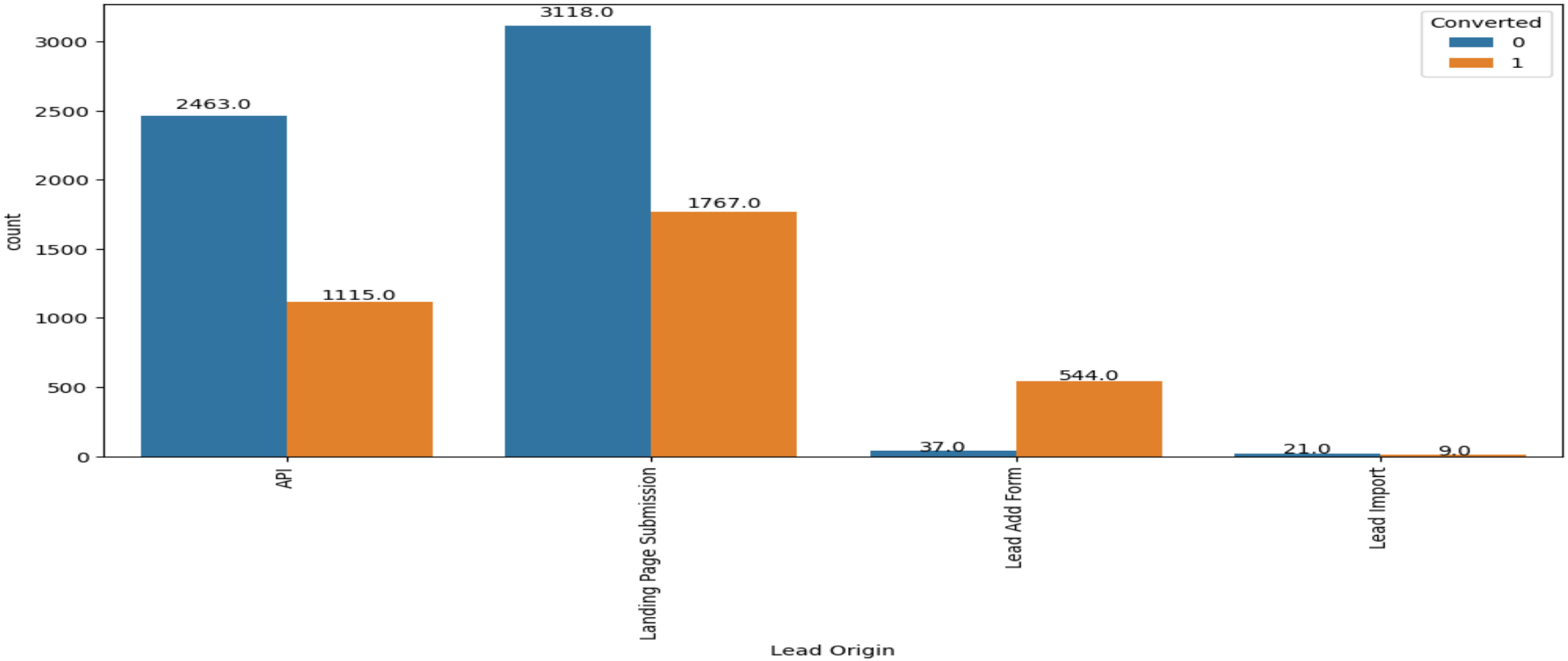
Conclusions and recommendations.

•Google and Direct traffic generates maximum number of leads.
•Conversion Rate of reference leads and leads through welingak website is high.

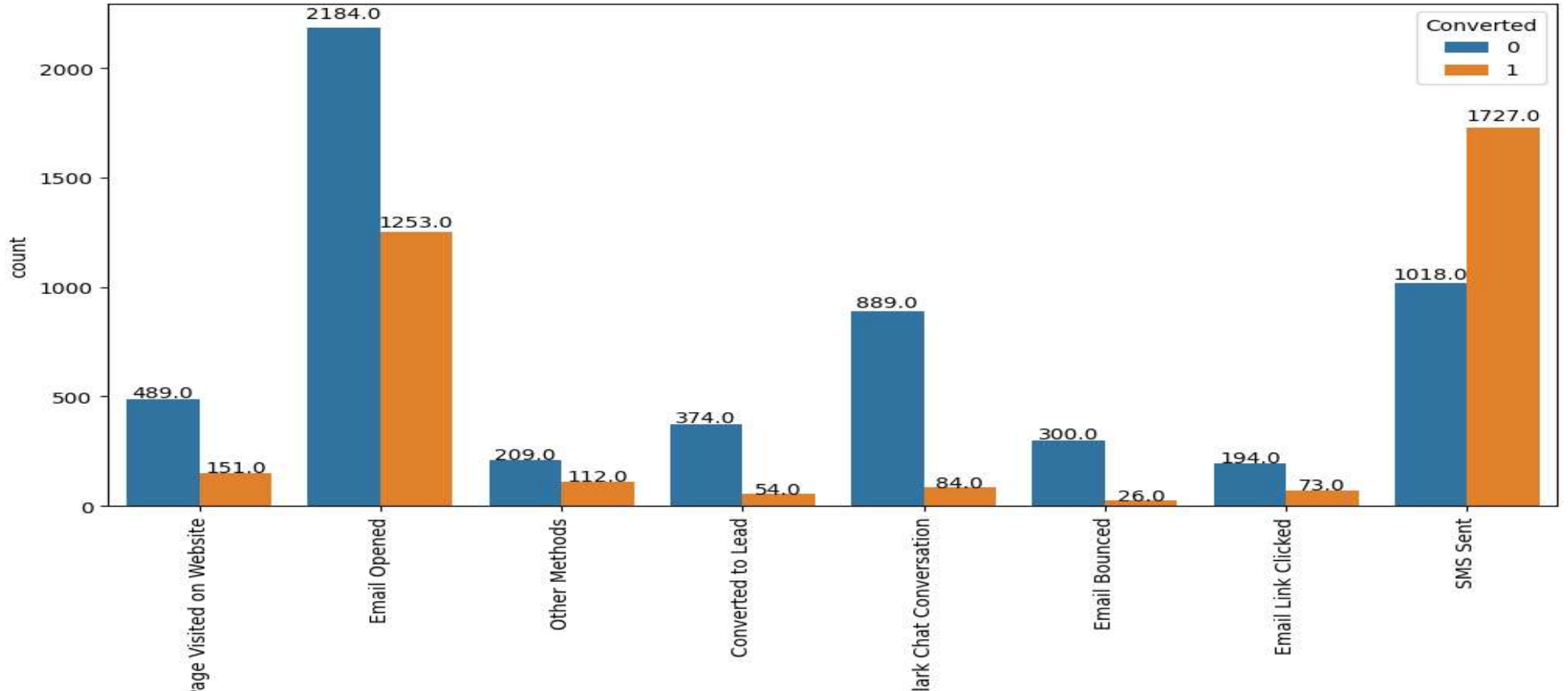


- API and Landing Page Submission have 30-35% conversion rate but count of lead originated from them are considerable.
- Lead Add Form has more than 90% conversion rate but count of lead are not very high.
- Lead Import are very less in count.

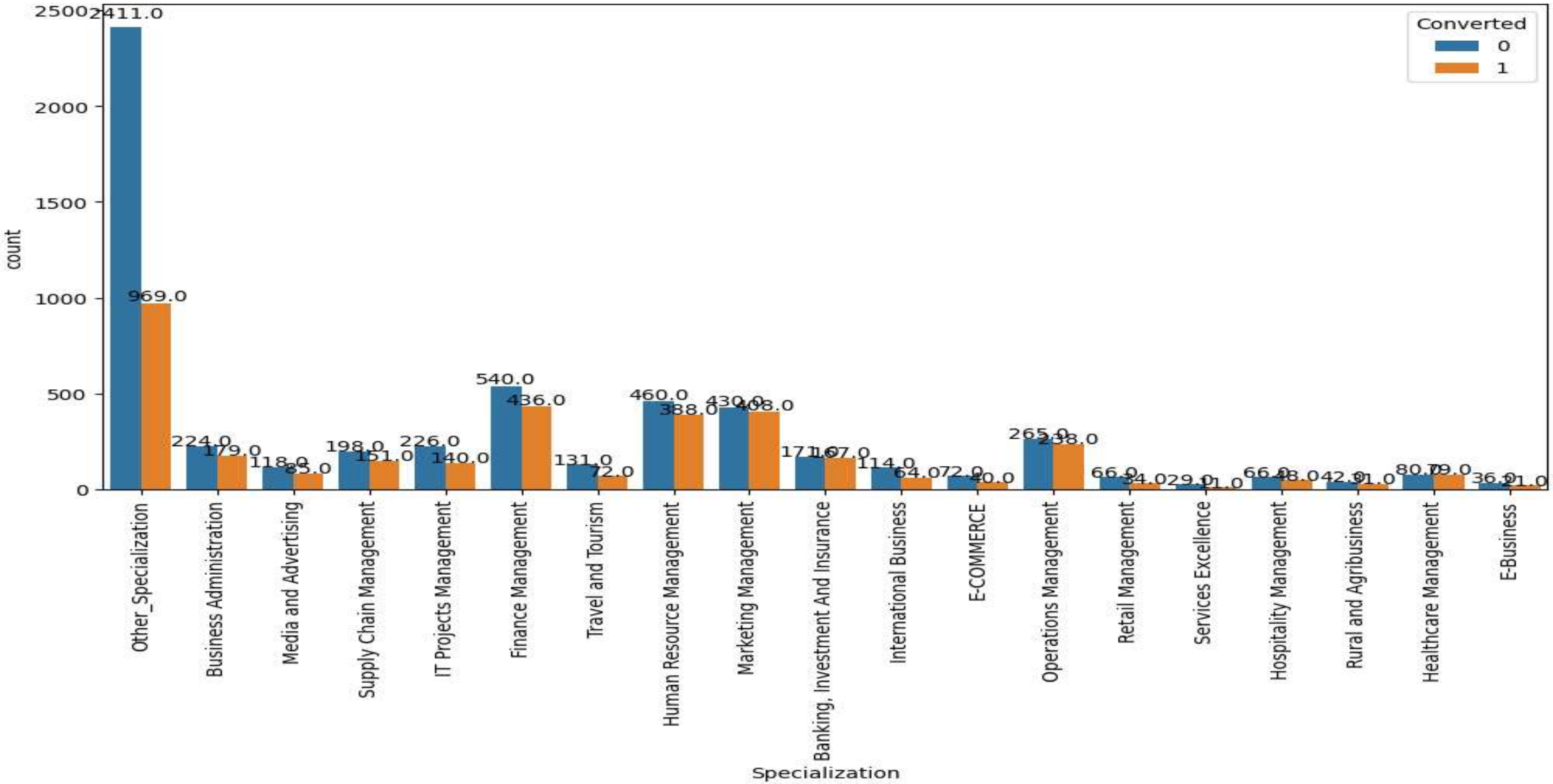
To improve overall lead conversion rate, we need to focus more on improving lead converion of API and Landing Page Submission origin and generate more leads from Lead Add Form.



The count of last activity as "Email Opened" is maximum-
The conversion rate of SMS sent as last activity is maximum
We should focus on increasing the conversion rate of those having last activity as Email Opened by making a call to those leads and also try to increase the count of the ones having last activity as SMS sent



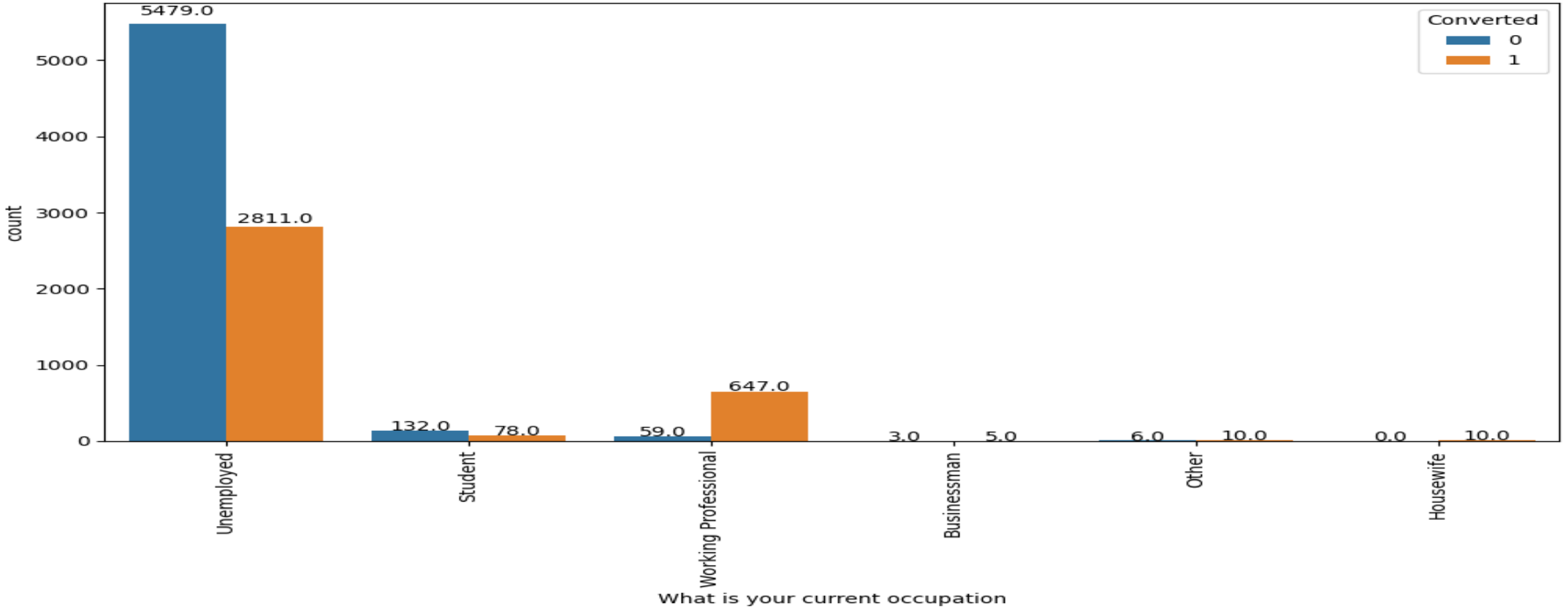
No particular inference can be made for Specialization

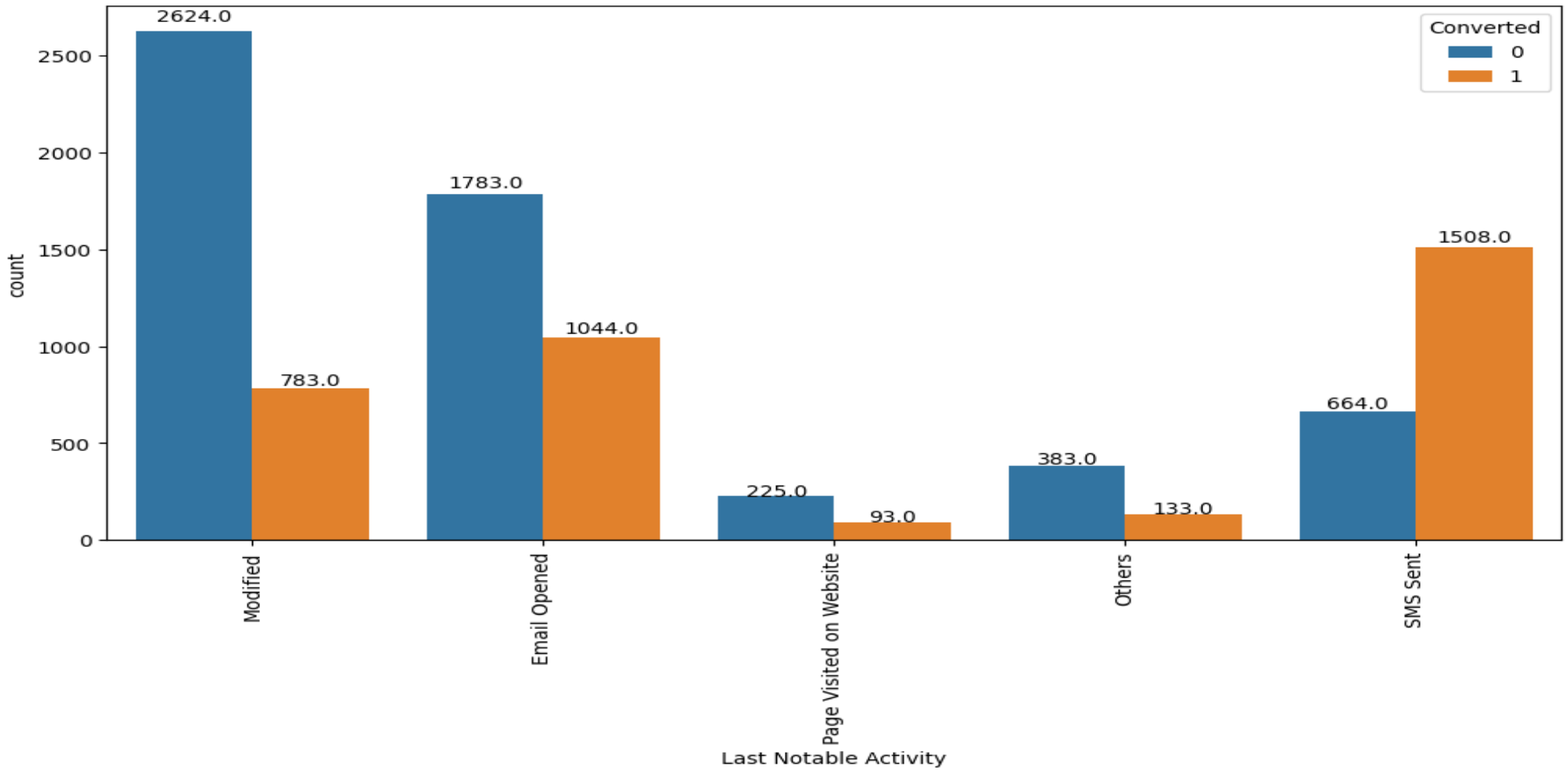


we can say that working professionals have high conversion rate - Number of Unemployed leads are more than any other category

To increase overall conversion rate, we need to increase the number of Working Professional leads by reaching out to them through different social sites such as LinkedIn etc. and also on increasing the conversion rate of Unemployed leads

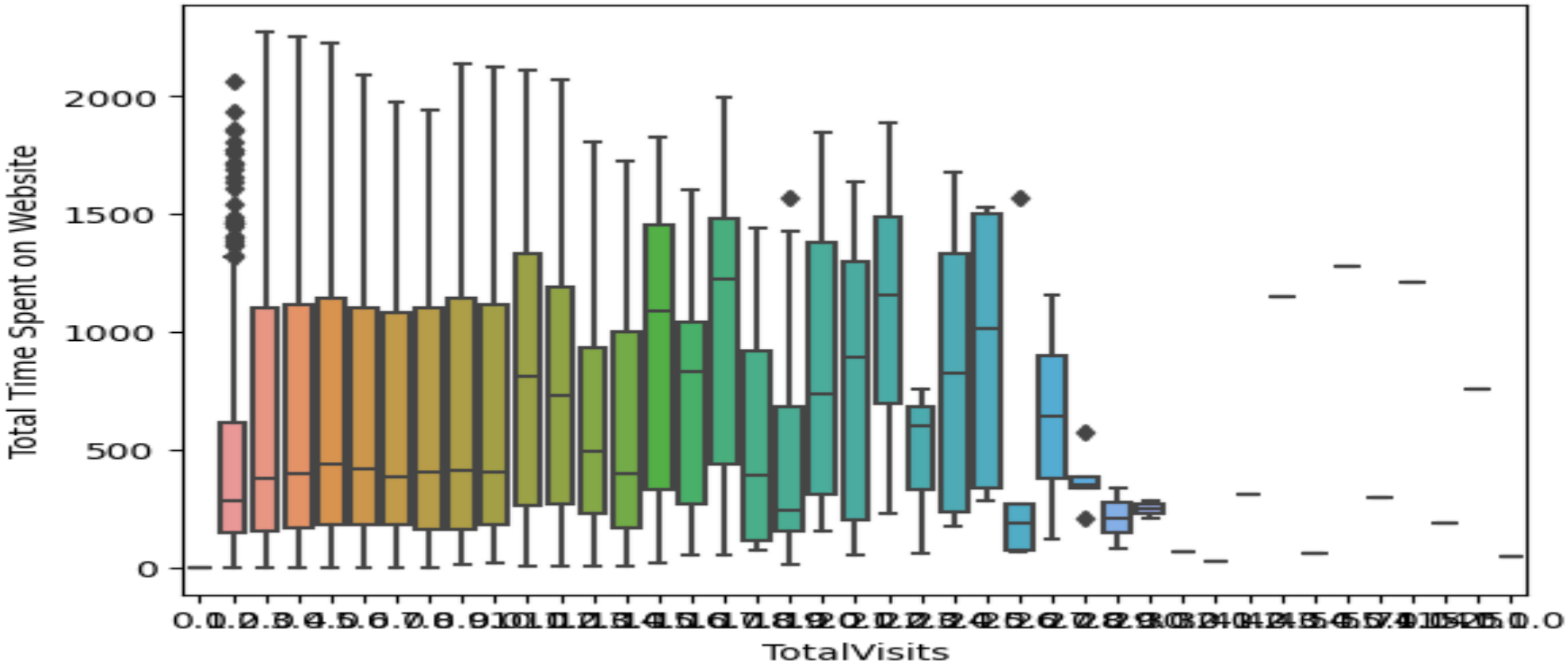
Country, What matters most to you in choosing a course, City columns have most values corresponding to one value such as India for Country, Mumbai for city and hence there is no particular insights for these columns




Converted
0
1
count
2500
2000
1500
1000
500
0
2624.0
783.0
1783.0
1044.0
225.0
93.0
383.0
133.0
664.0
1508.0
Modified
Email Opened
Page Visited on Website
Others
SMS Sent
Last Notable Activity

- Leads spending more time on the weblise are more likely to be converted.

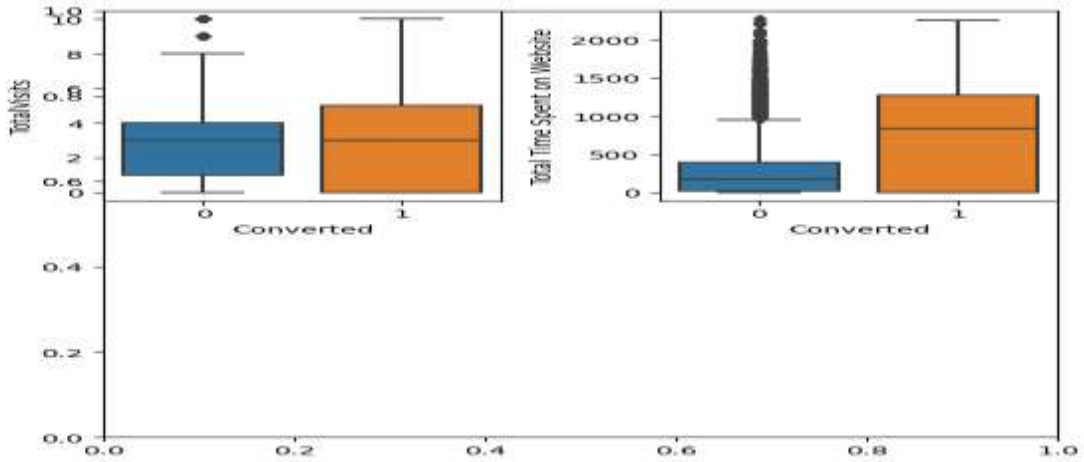
Website should be made more engaging to make leads spend more time.

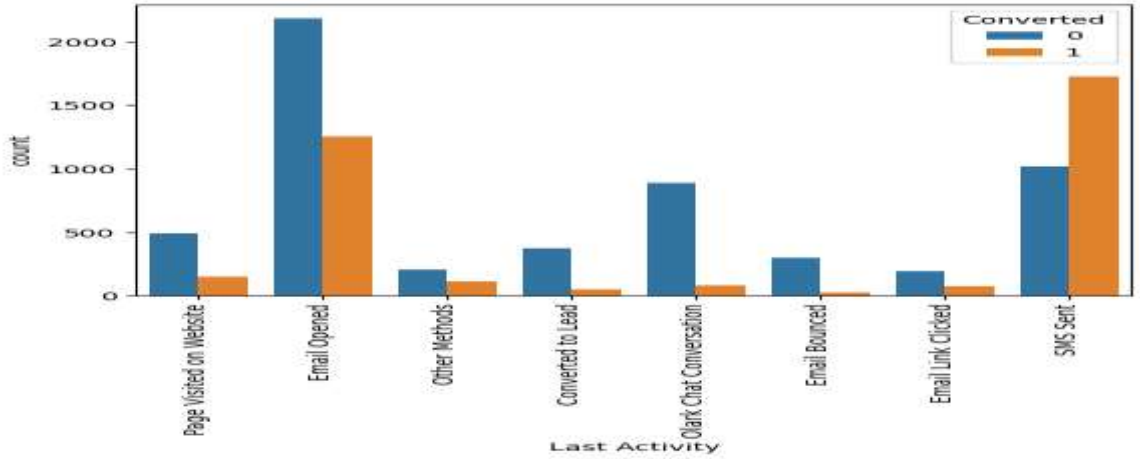


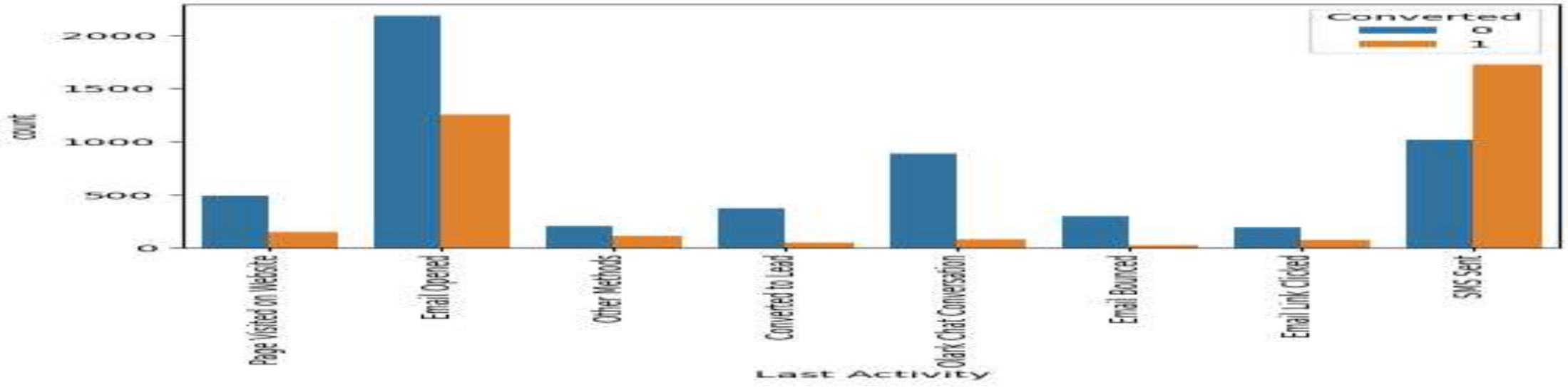
The median of both the conversion and non-conversion are same and hence nothing conclusive can be said using this information - Users spending more time on the website are more likely to get converted

The count of last activity as "Email Opened" is max

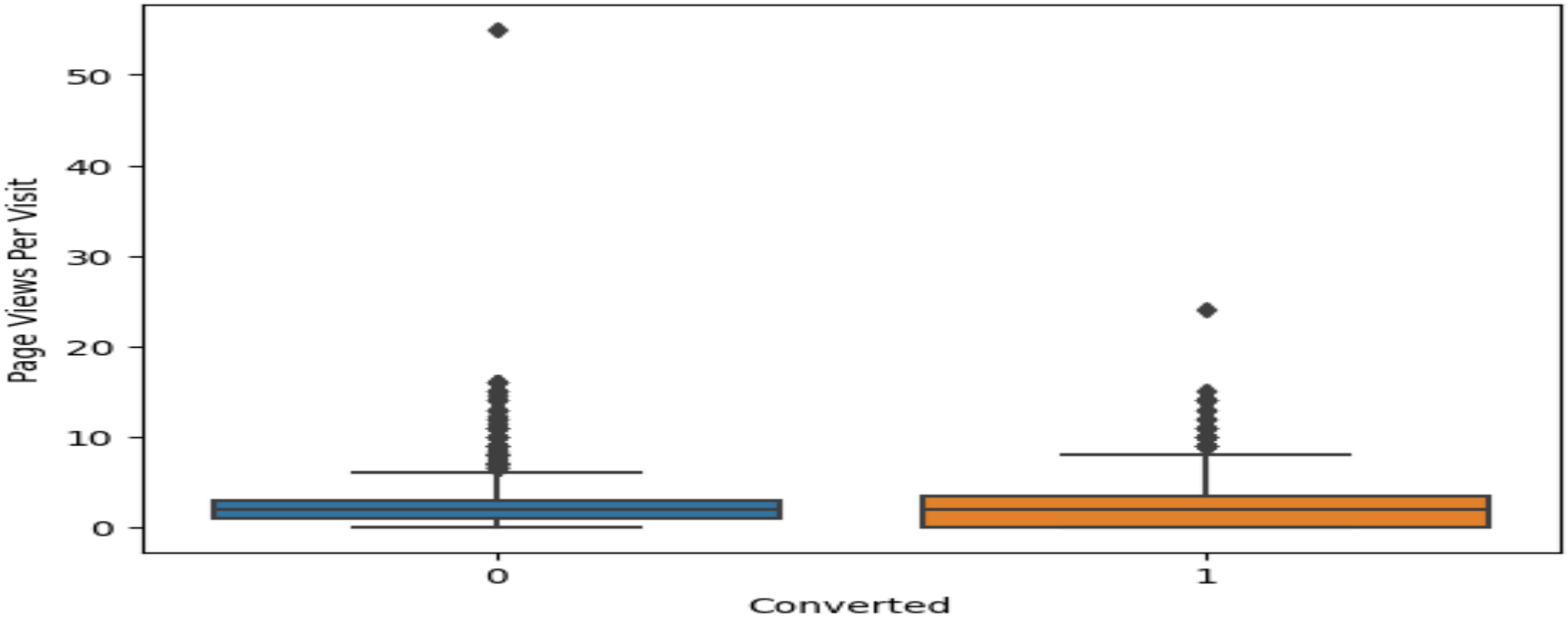
The conversion rate of SMS sent as last activity is maximum



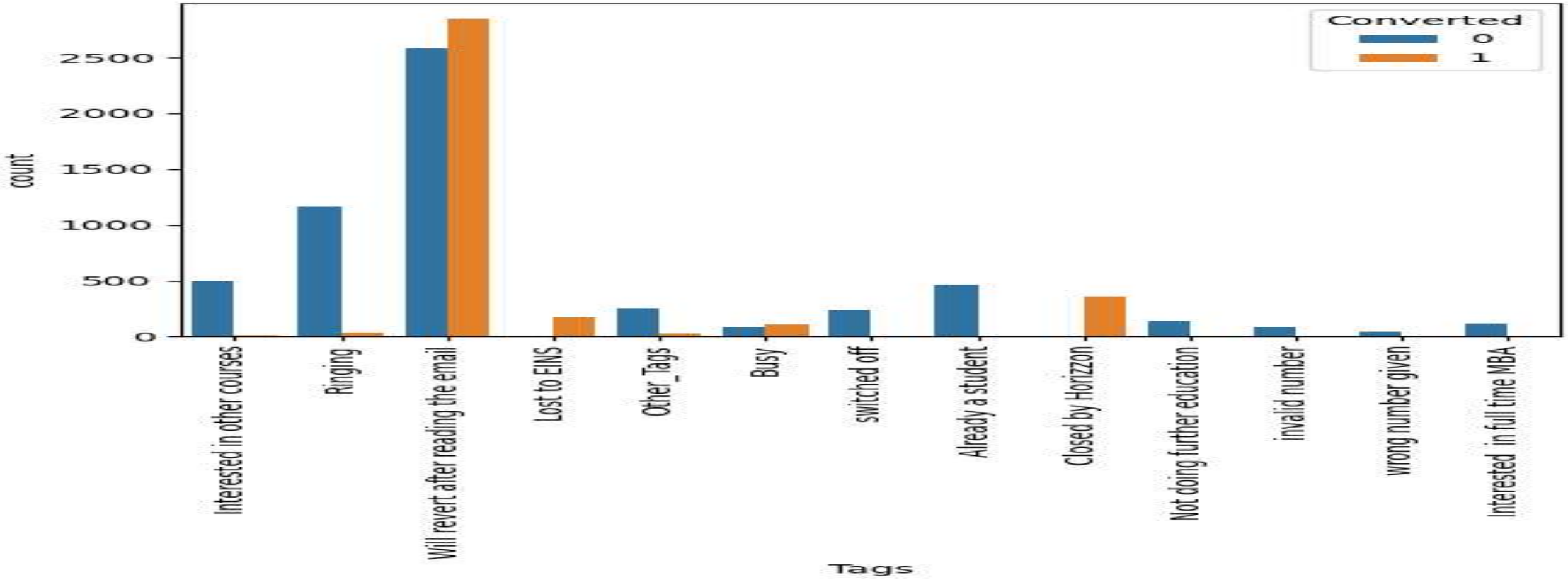




As we can see there are a number of outliers in the data. We will cap the outliers to 95% value for analysis.

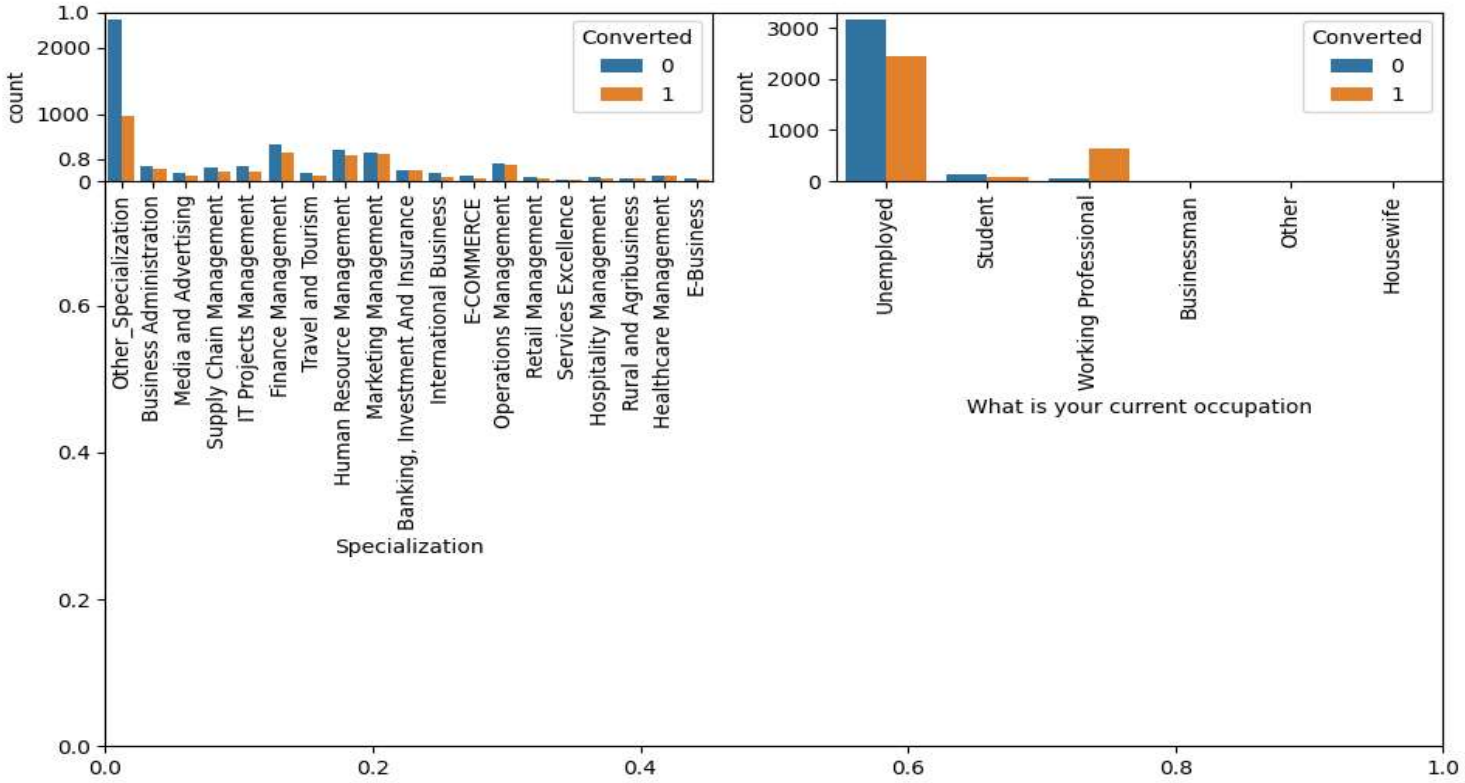


Will revert after reading the email' and 'Closed by Horizzon' have high conversion rate

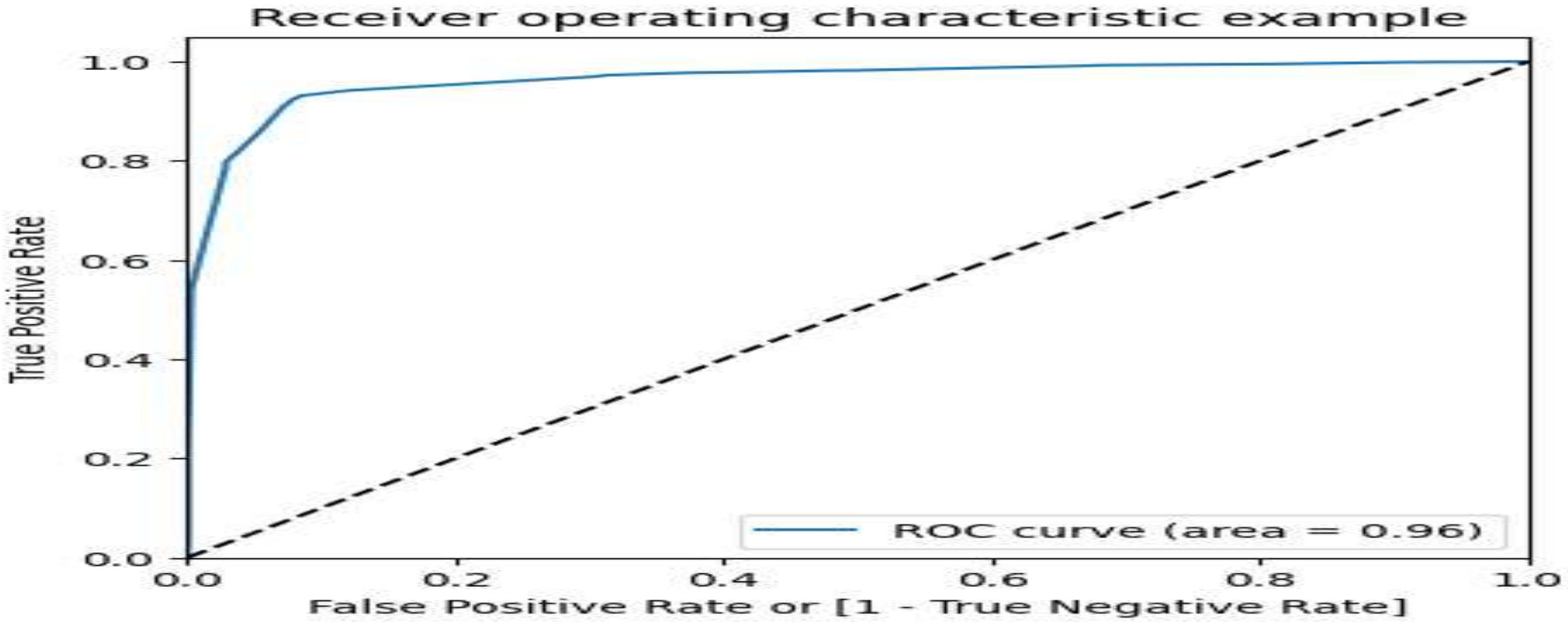


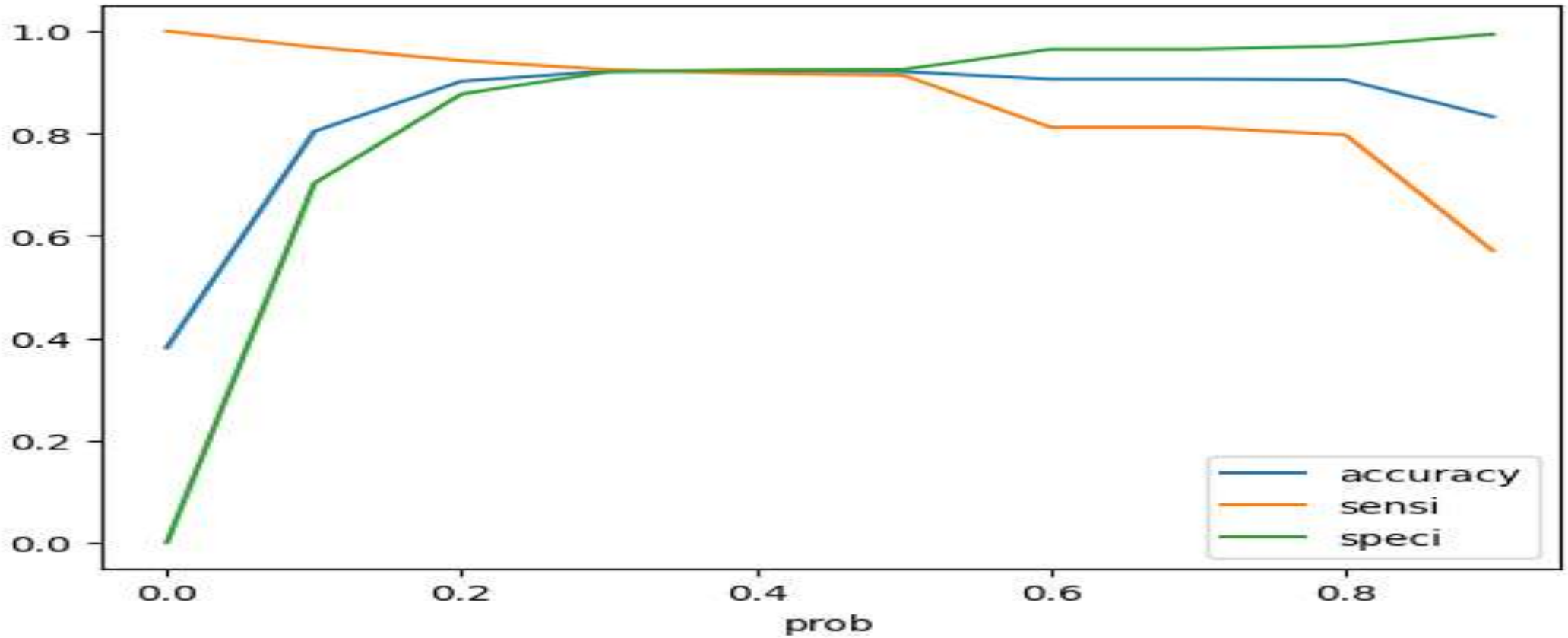
* Looking at plot, no particular inference can be made for Specialization - Looking at plot, we can say that working professionals have high conversion rate - Number of Unemployed leads are more than any other category **To increase overall conversion rate, we need to increase the number of Working Professional leads by reaching out to them through different social sites such as LinkedIn etc. and also on increasing the conversion rate of Unemployed leads** - Country, What matters most to you in choosing a course, City columns have most values corresponding to one value such as India for Country, Mumbai for

city and hence there is no particular insights for these columns



ROC-value-is-096-which-is-very-good




1.0
0.8
0.6
0.4
0.2
0.0
0.0
0.2
0.4
0.6
0.8
prob
accuracy
sensi
speci