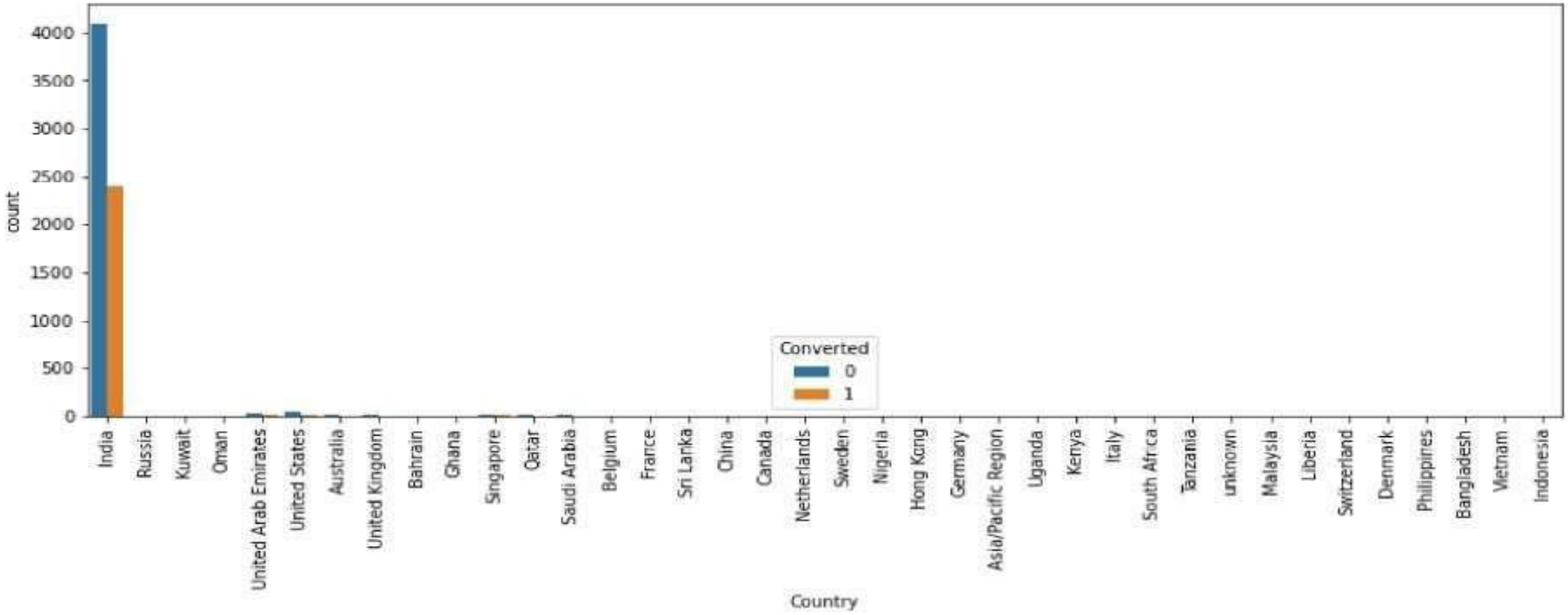
Lead Score

DHARMENDRA SINGH NEGI

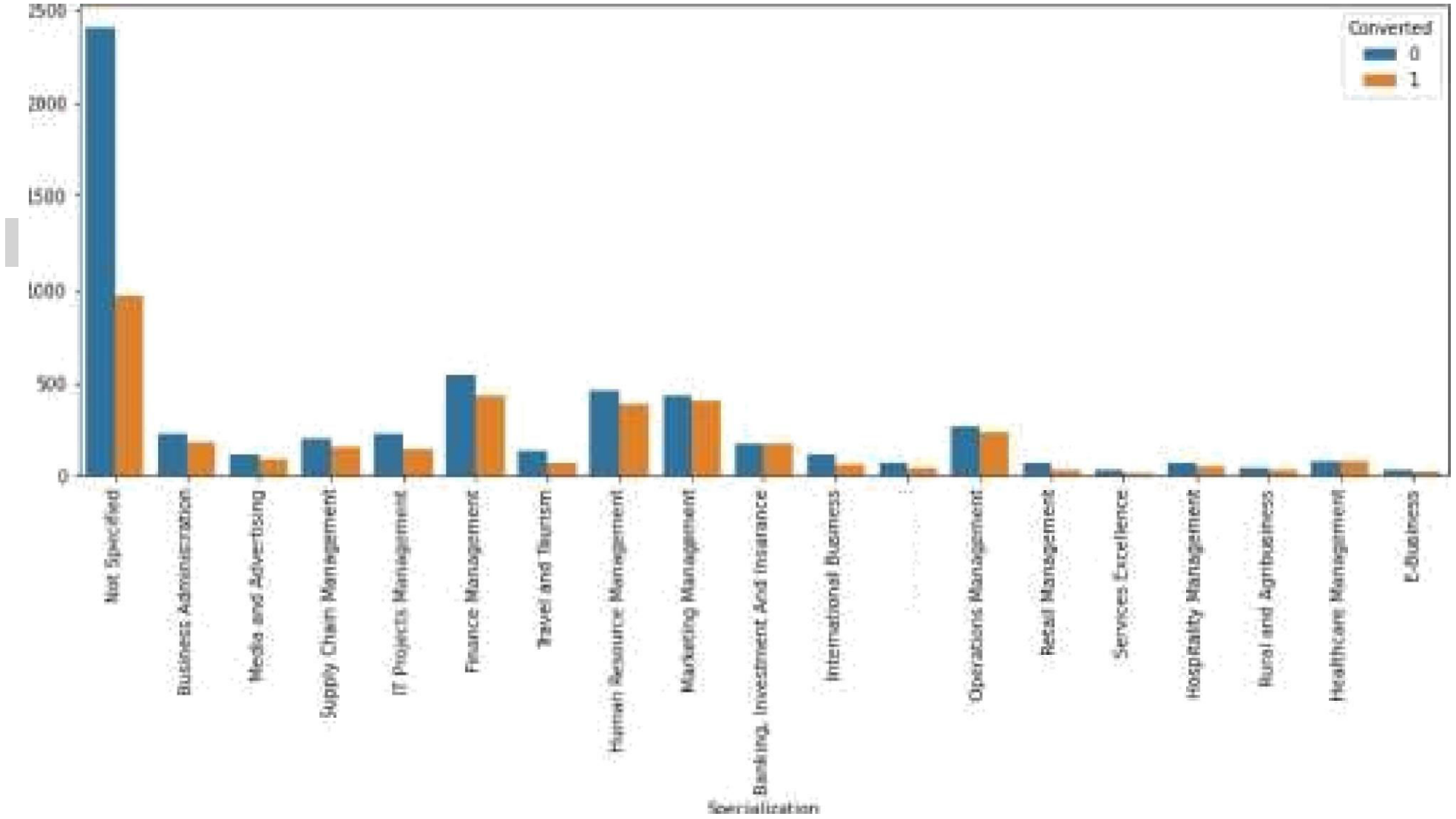
DIPSHIKHA PURKAYASTHA

DIPALI Pawar

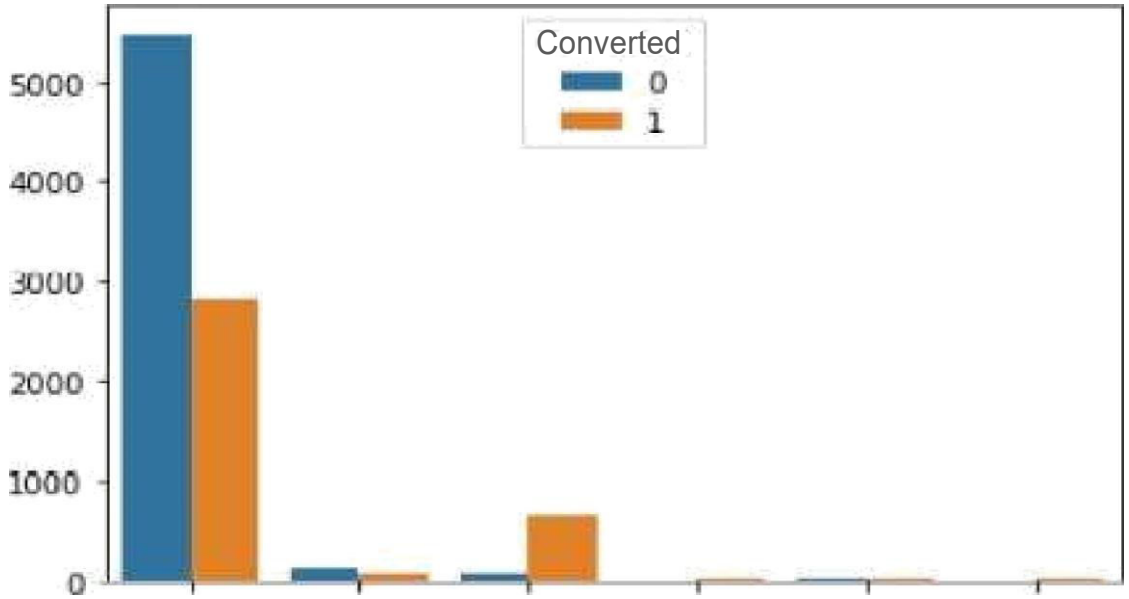
Exploratory Data Analysis



As we can see the Number of Values for India are quite high (about 97% of the Data), this column can be dropped

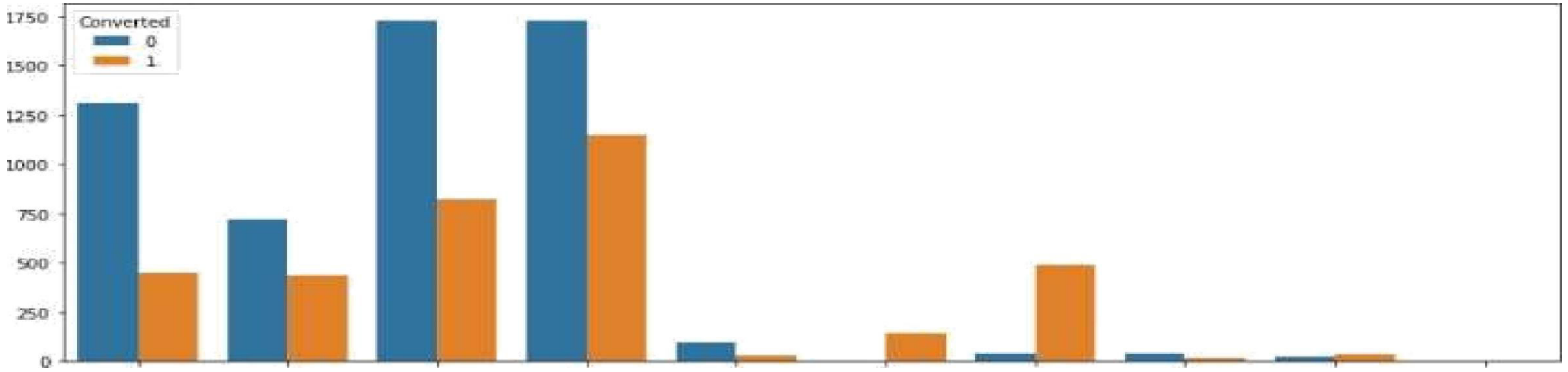


We can see that the specialization with higher values will definitely be a significant variable and should not be dropped.



Wbat is your current occupation

Working Professionals gDing for the course have high chances of joining it. Unemployed leads are the most in terms of Absolute numbers.



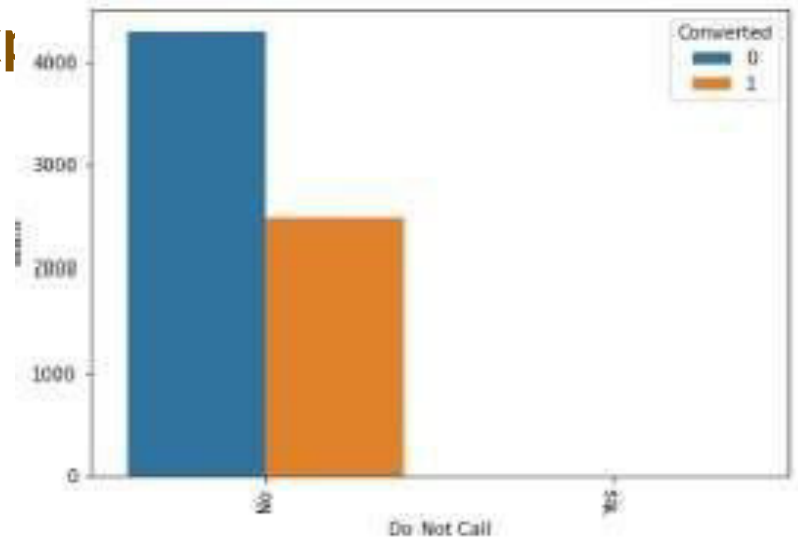
Inference

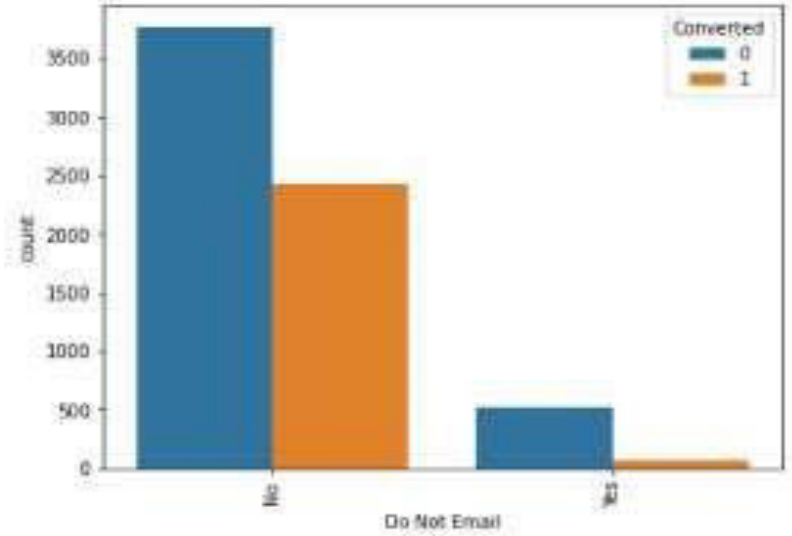
1. Maximum number of lea4s are generated by Google and Direct traffic.
2. Conversion Rate of reference leads ancJ leads through welingak website is high.

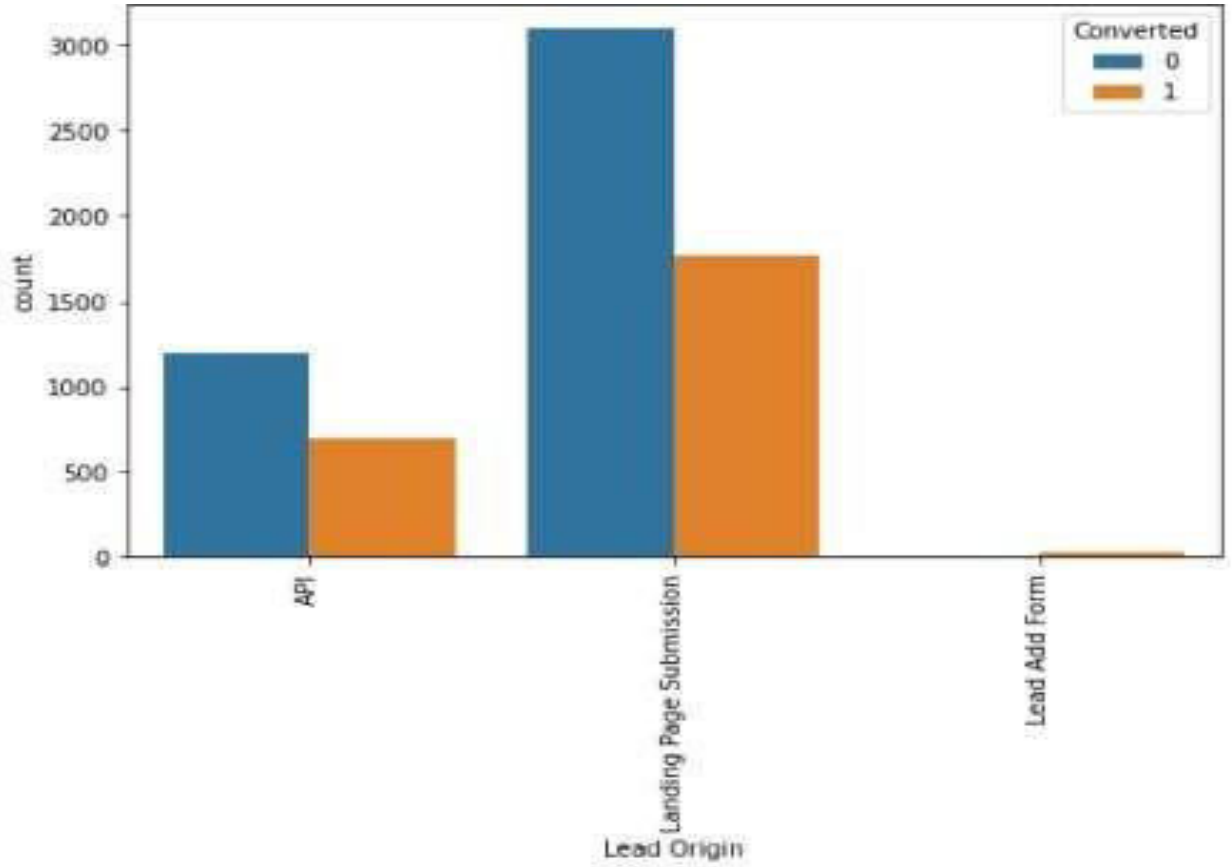
E. Cm improve overall lead conversion rate. should be on improving lead cc›nverion of olark chat. organic search, direct traffic, and google leads and generate mc›re leads reference and welin;gak website.

focus

Exp

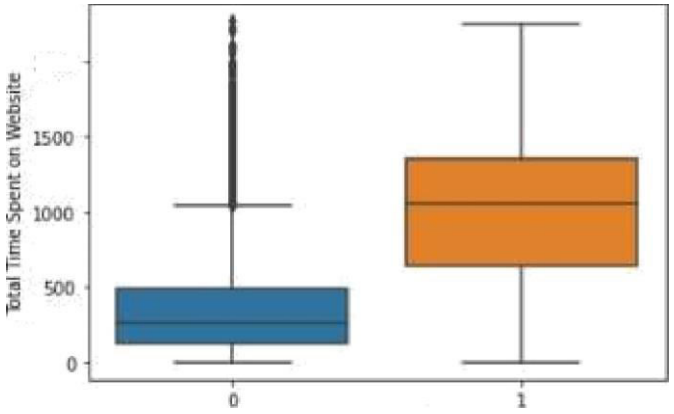






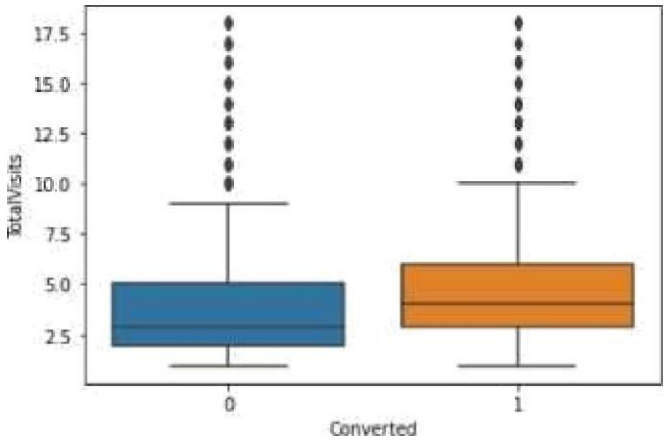
Inference

1. API and Landing Page Submission bring higher number of leads as well as conversion.
2. Lead Add Form has a very high conversion rate but count of leads are not very high.
3. Lead Import and Quick Add Form get very few leads.
4. In order to improve overall lead conversion rate, we have to improve lead converion of API and Landing Page Submission origin and generate more leads from Lead Add Form.



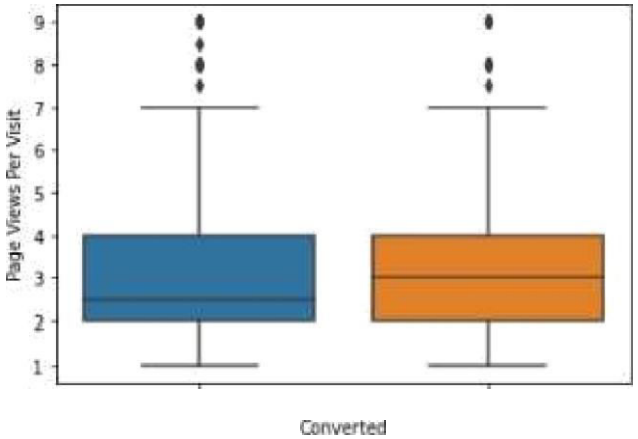
Inference

1. Ledds Spending more time on the website are more likely to be converted.
?. bsite should be made more enpagin g to make leads spend more time.



Inference

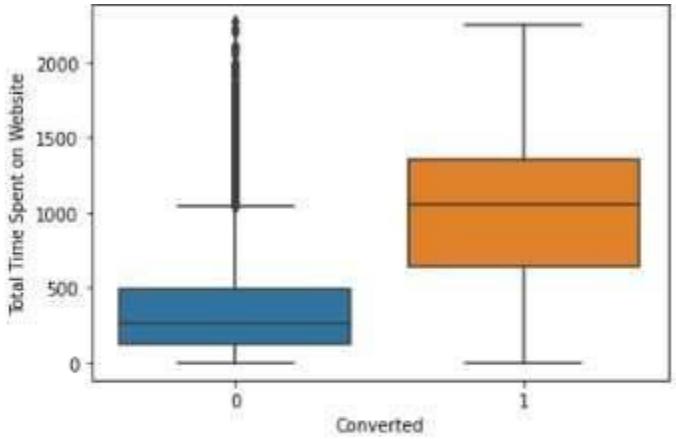
1. I‘dedian for converted and not convened leads are the close.
Z, Nothng conclusive can be said on the basis of Total Visits



Inference

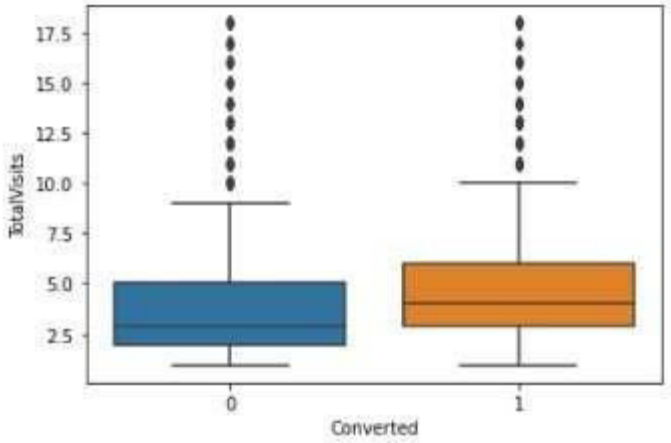
1. Median for conve:rted and unconverred leads is the Anne.
7. Nothing can be said specifically for Sea d conversion from Paqe Vi0ws Per Visit

Exploratory Data Analysis



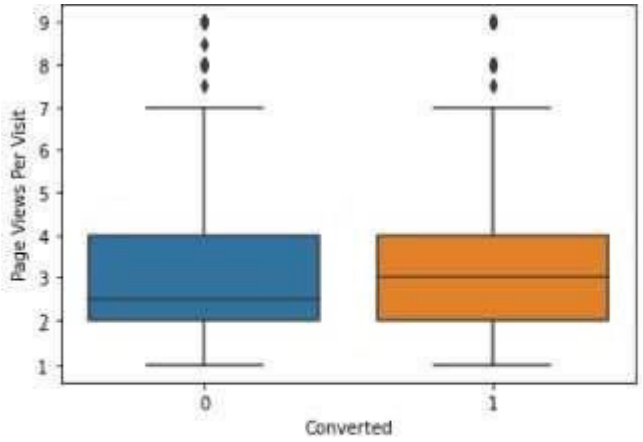
Inference

1. Leads spending more time on the website are more likely to be converted.
2. Website should be made more engaging to make leads spend more time.



Inference

1. Median for converted and not converted leads are the close.
2. Nothng conclusive can be said on the basis of Total Visits

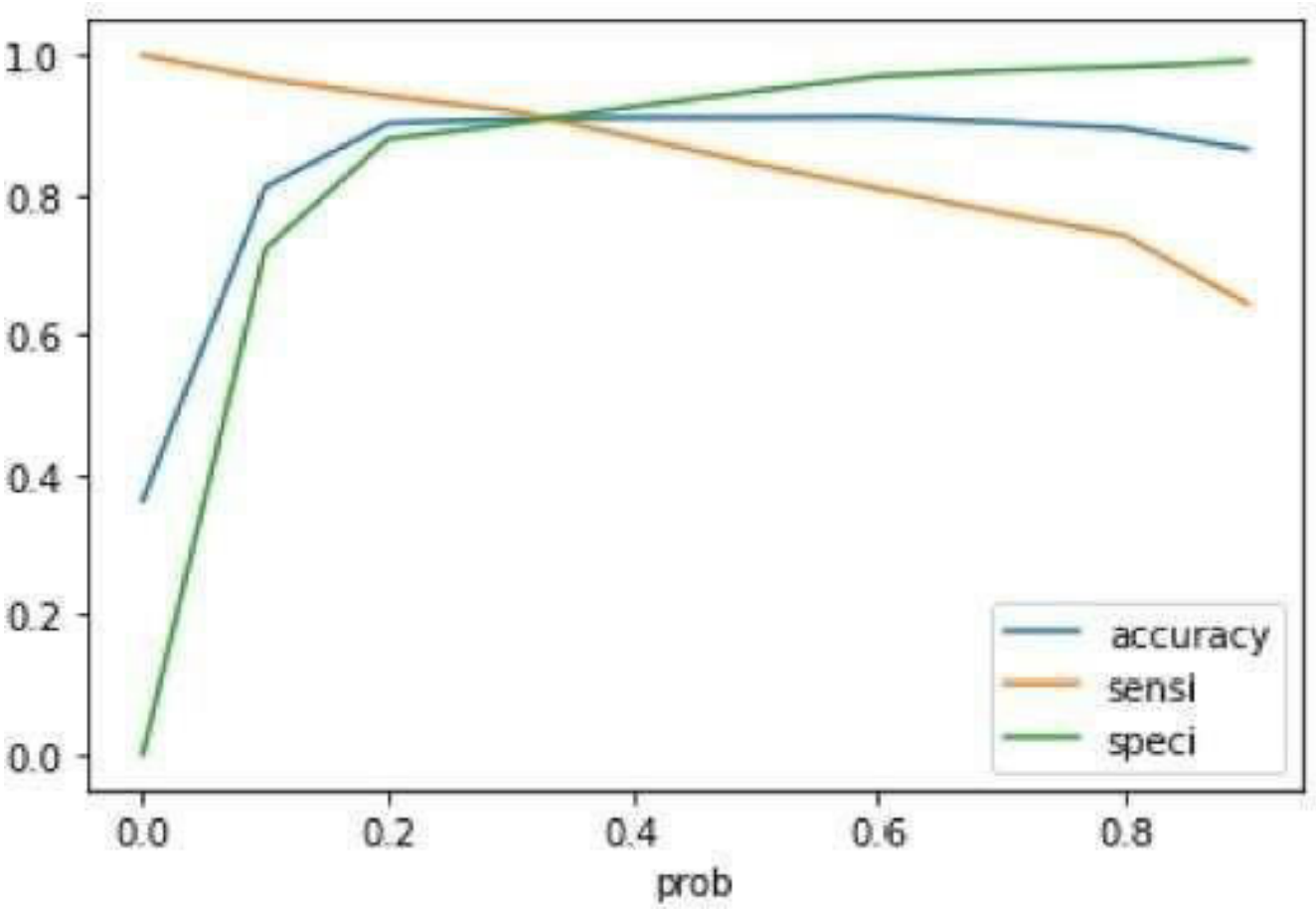


Inference

1. Median for converted and unconverted leads is the same.
2. Nothing can be said specifically for lead conversion from Page Views Per Visit

Variables Impacting the Conversion Rate

Model Evaluation -Sensitivity and Specificity on Train Data Set



Observation:

So as we can see above the model seems to be performing well. The ROC curve has a value of 0.97, which is very good. We have the following values for the Train Data:

Accuracy : 90.81%

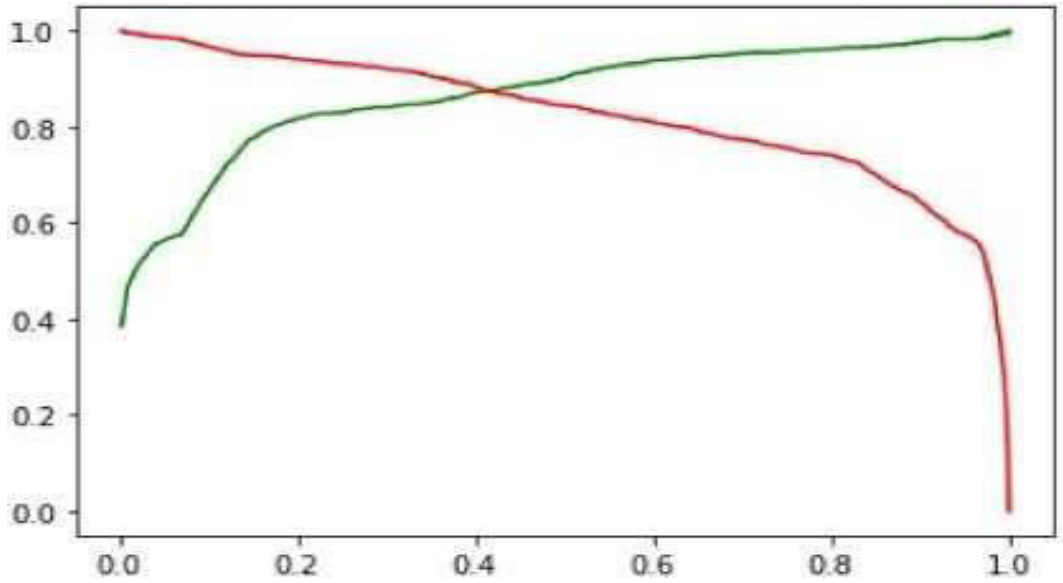
Sensitivity : 92.05%

Specificity : 90.10%

Some of the other Stats are derived below, indicating the False Positive Rate, Positive Predictive Value,Negative Predictive Values, Precision & Recall.

The graph depicts an optimal cut off 0.37 based on Accuracy, Sensitivity and Specificity

Model Evaluation -Precision and Recall on Train Data Set



The graph depicts an optimal cut off 0.42 based on Precision and Recall

- Precision-84.12%
- Recall-92.05 %

Observation:

After running the model on the Test Data these are the figures we obtain:

Accuracy : 90.92%

Sensitivity : 91.41%

Specificity : 90.62%

Conclusion

- While we have checked both Sensitivity-Specificity as well as Precision and Recall Metrics, we have considered the optimal cut off based on Sensitivity and Specificity for calculating the final prediction.

- Accuracy, Sensitivity and Specificity values of test set are around 91%, 91.41% and 90.62% which are approximately closer to the respective values calculated using trained set.

- lead score calculated shows the conversion rate on the final predicted model is around 92.05% (in train set) and 91.41% in test set

- The top variables that contribute for lead getting converted in the model are

1. Total time spent on website
2. W hat is your current occupation
3. LeadAdd Form from Lead O rigin
4. Had a Phone Conversation from Last NotableActivity

- Hence overall this model seems to be good.