LEAD SCORING CASE STUDY

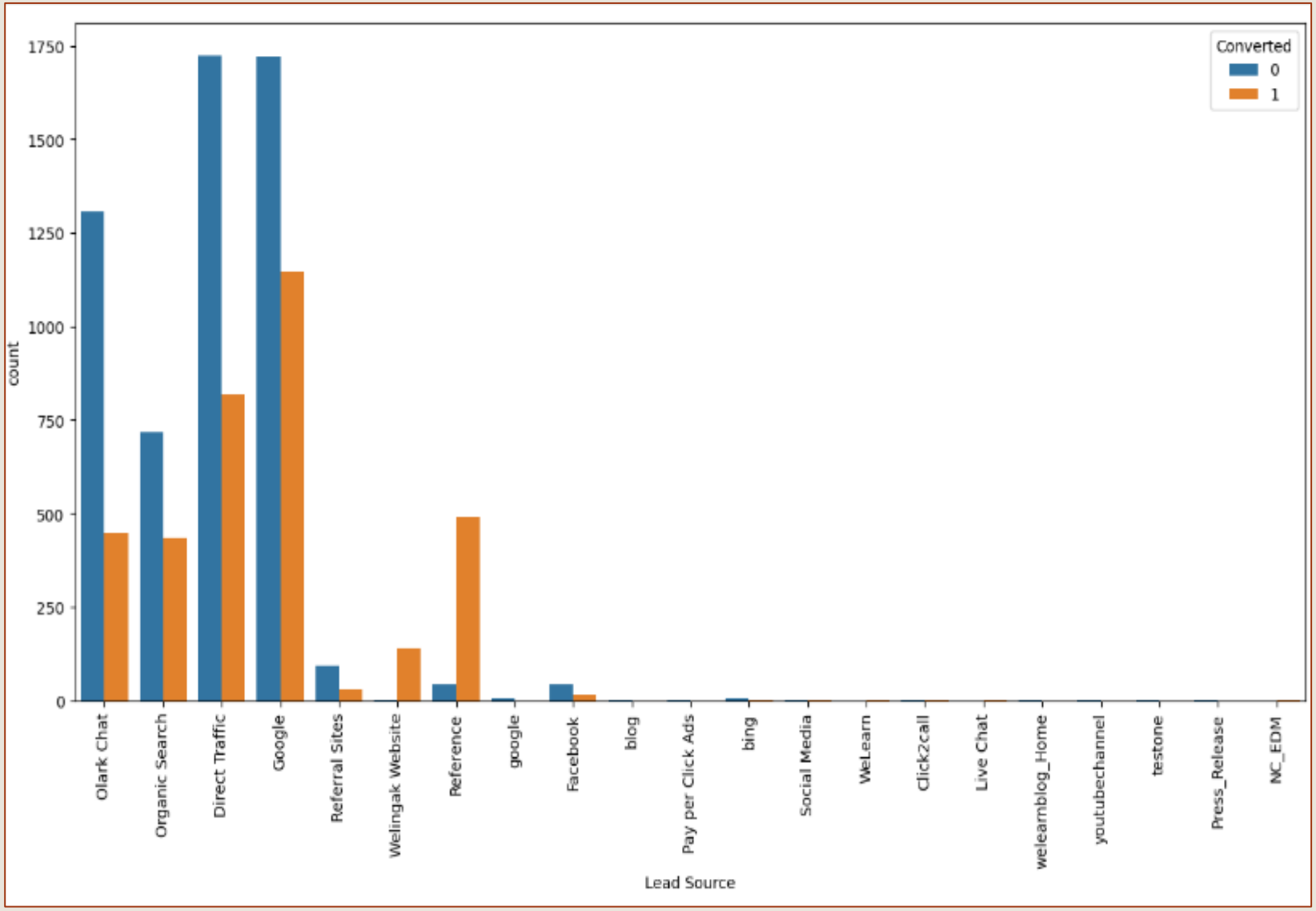
Submitted by:

- Lalit Kalbhor
- Shubham Kharde
- Shahida ParveenSyed

BUSINESS PROBLEM

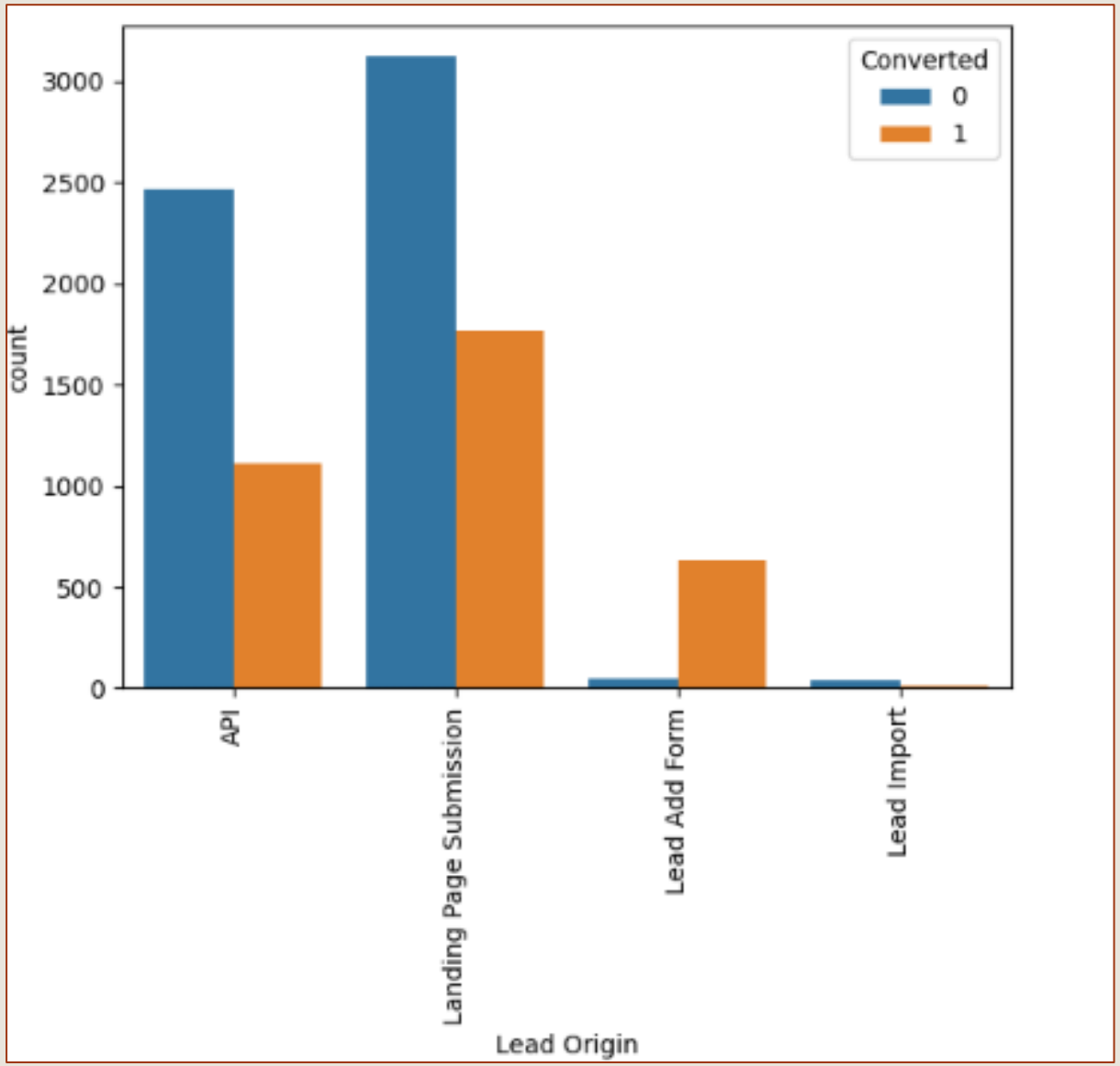
- X education is an organization which provides online courses for industry professional. The company markets its courses on several websites and search engines like Google.
- X education wants to select most promising leads that can be converted into paying customers.
- Although the company generates a lot of leads only a few are converted into paying customers, wherein the company wants a higher leads conversion
- 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:



Graph's conclusion

- The majority of leads come from Google and Direct traffic.
- Leads from references and the Welingak website have a higher conversion rate.

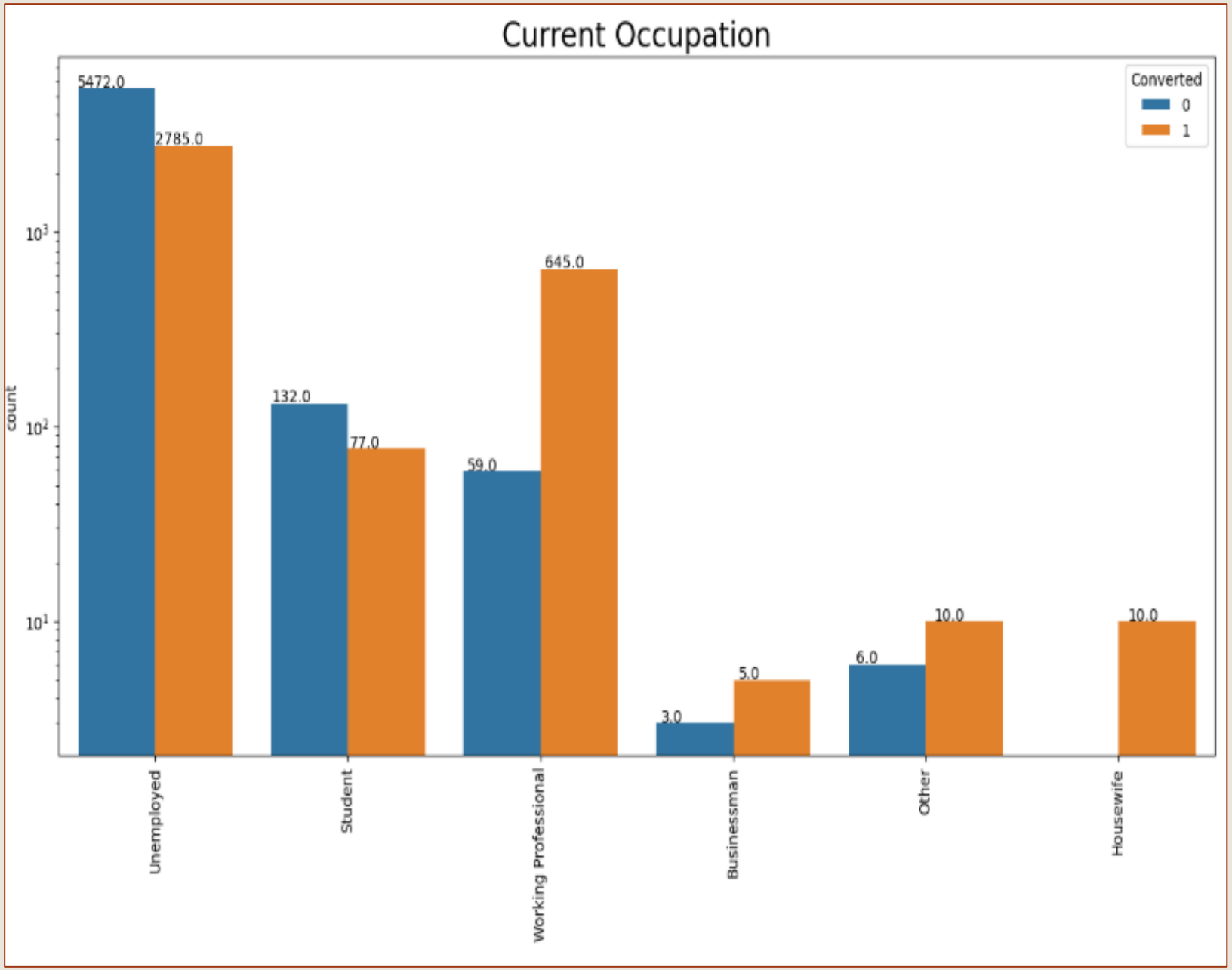


Graph's conclusion

•API and Landing Page Submissions have a conversion rate of 30-35%, but the number of leads is high.

•The Lead Add Form has a conversion rate of over 90%, though the number of leads is relatively low.

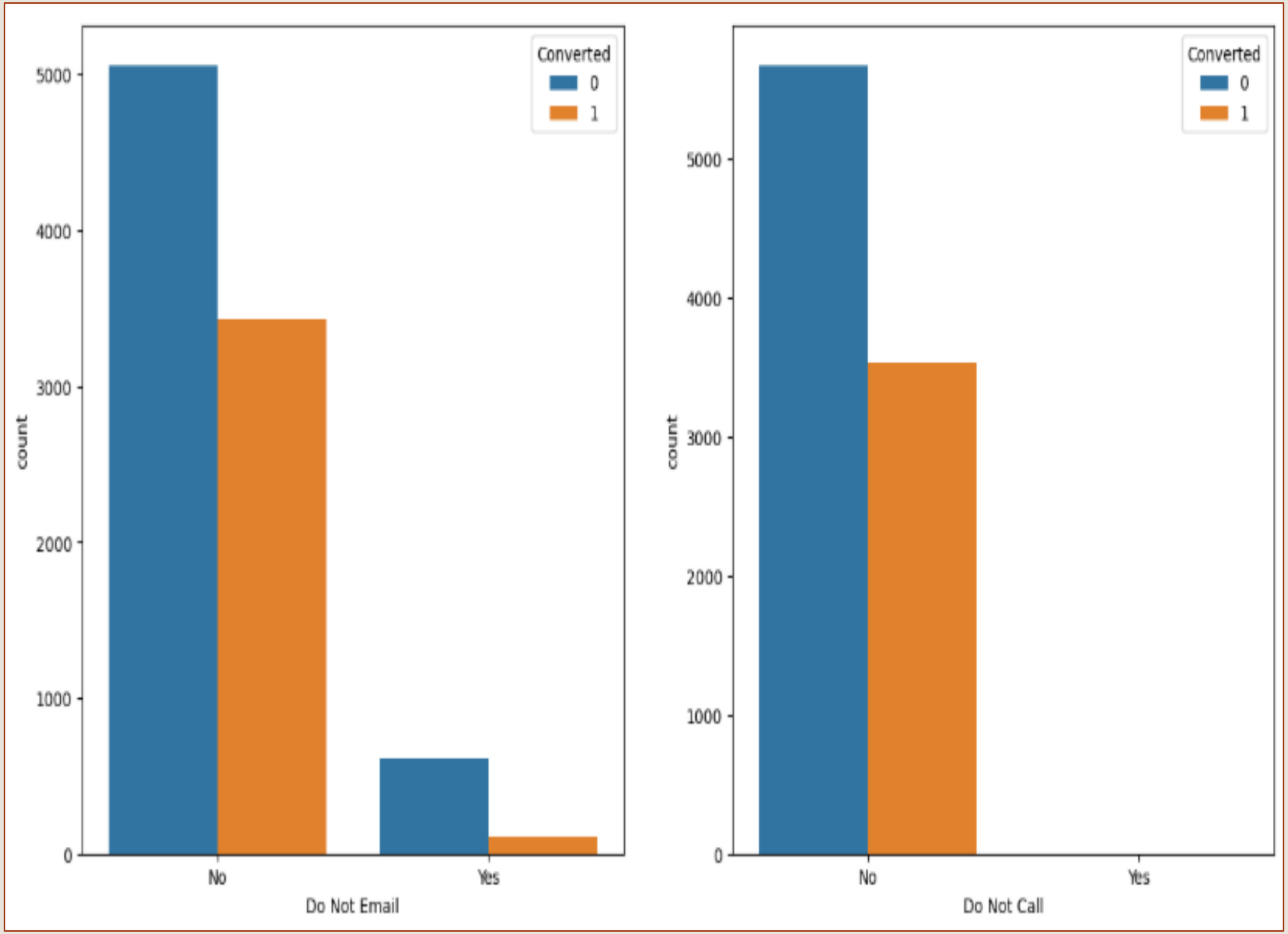
•Lead Imports have a very low count.



Graph’s conclusion

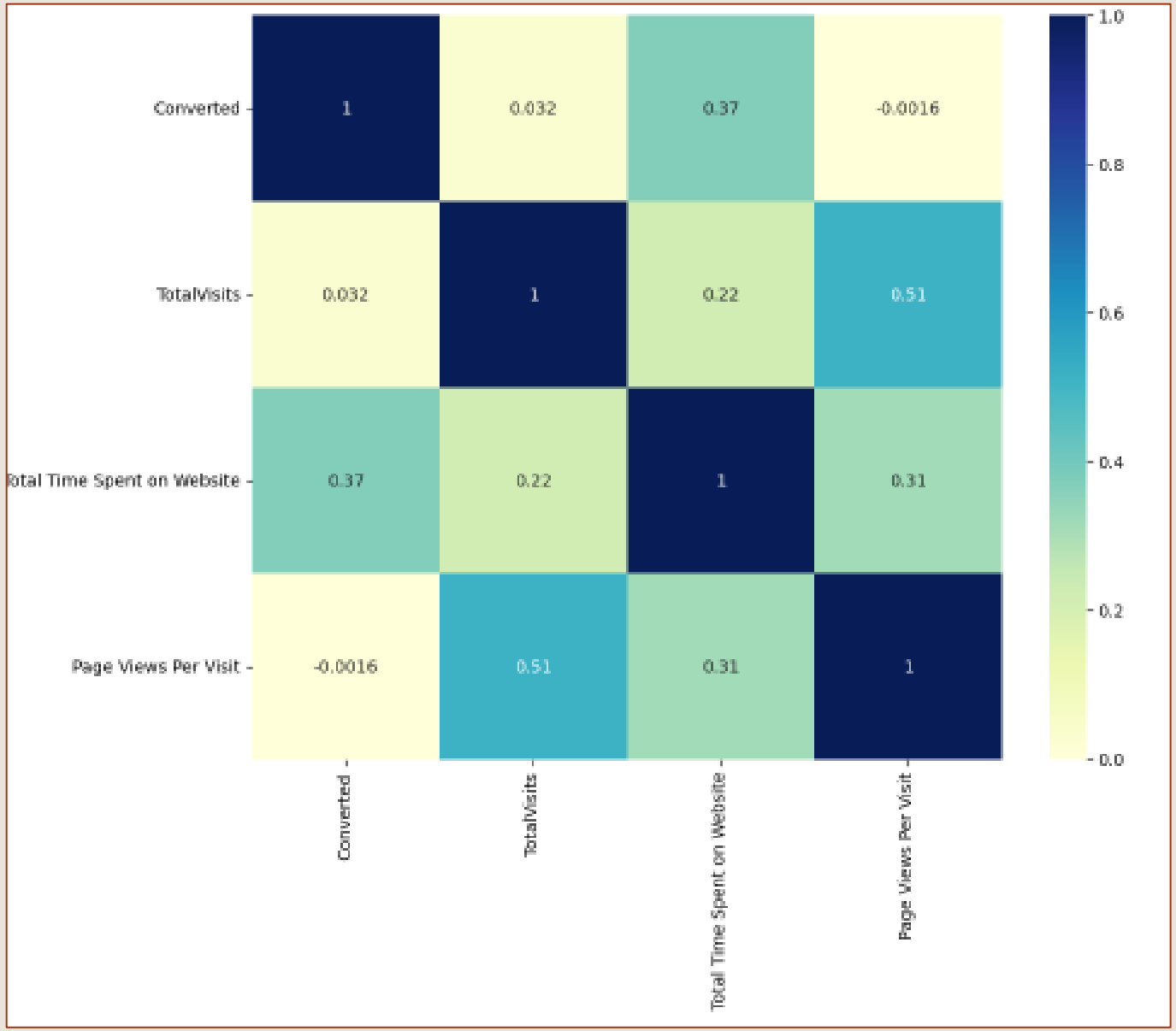
1) "Unemployed" leads generate a higher number of leads with a conversion rate of approximately 34%.

2) "Working Professionals" have a higher conversion rate



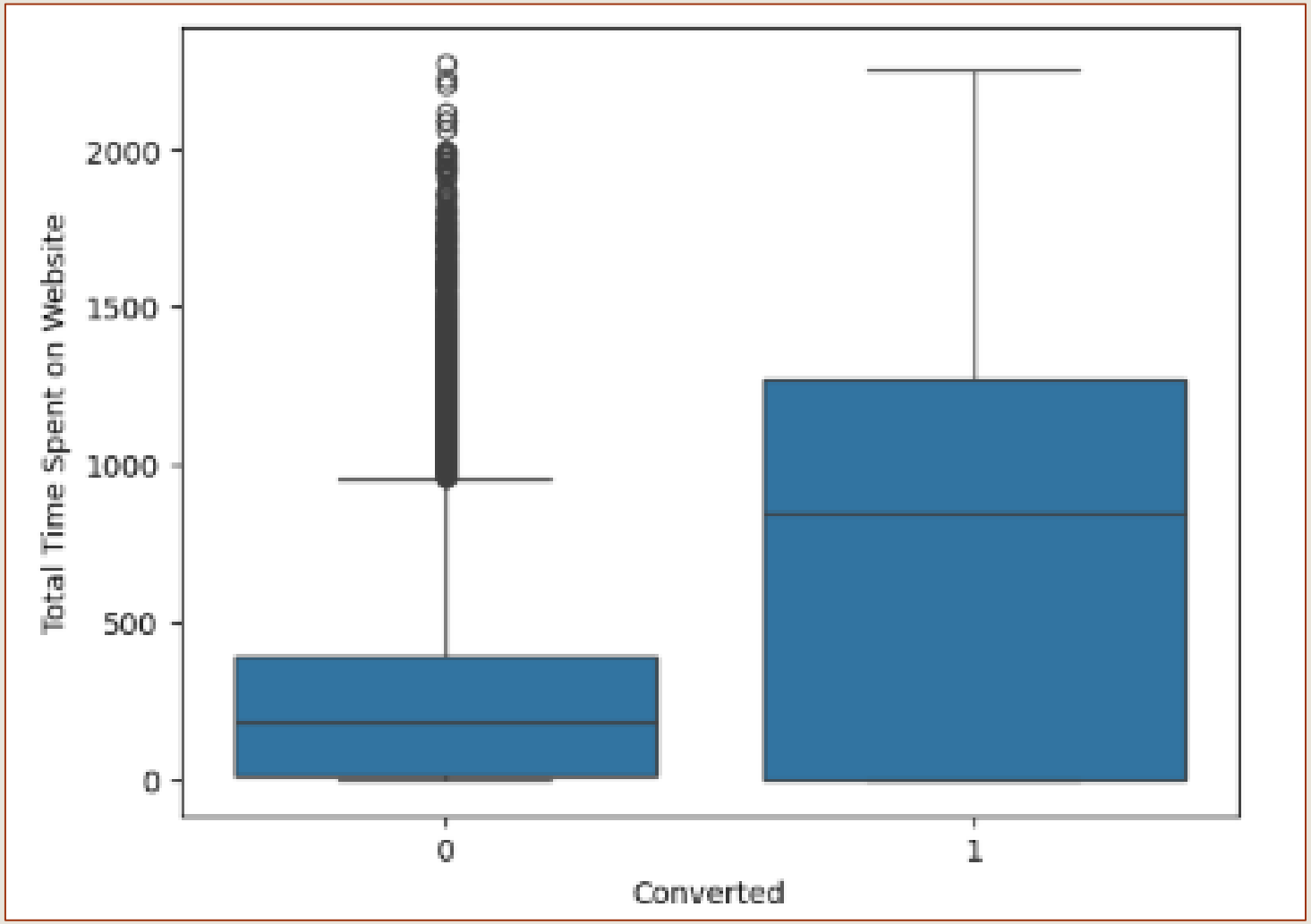
Graph’s conclusion

Since many customers select the "Do Not Email" and "Do Not Call" options, it indicates that most leads prefer not to be contacted via phone or email.



Graph's conclusion

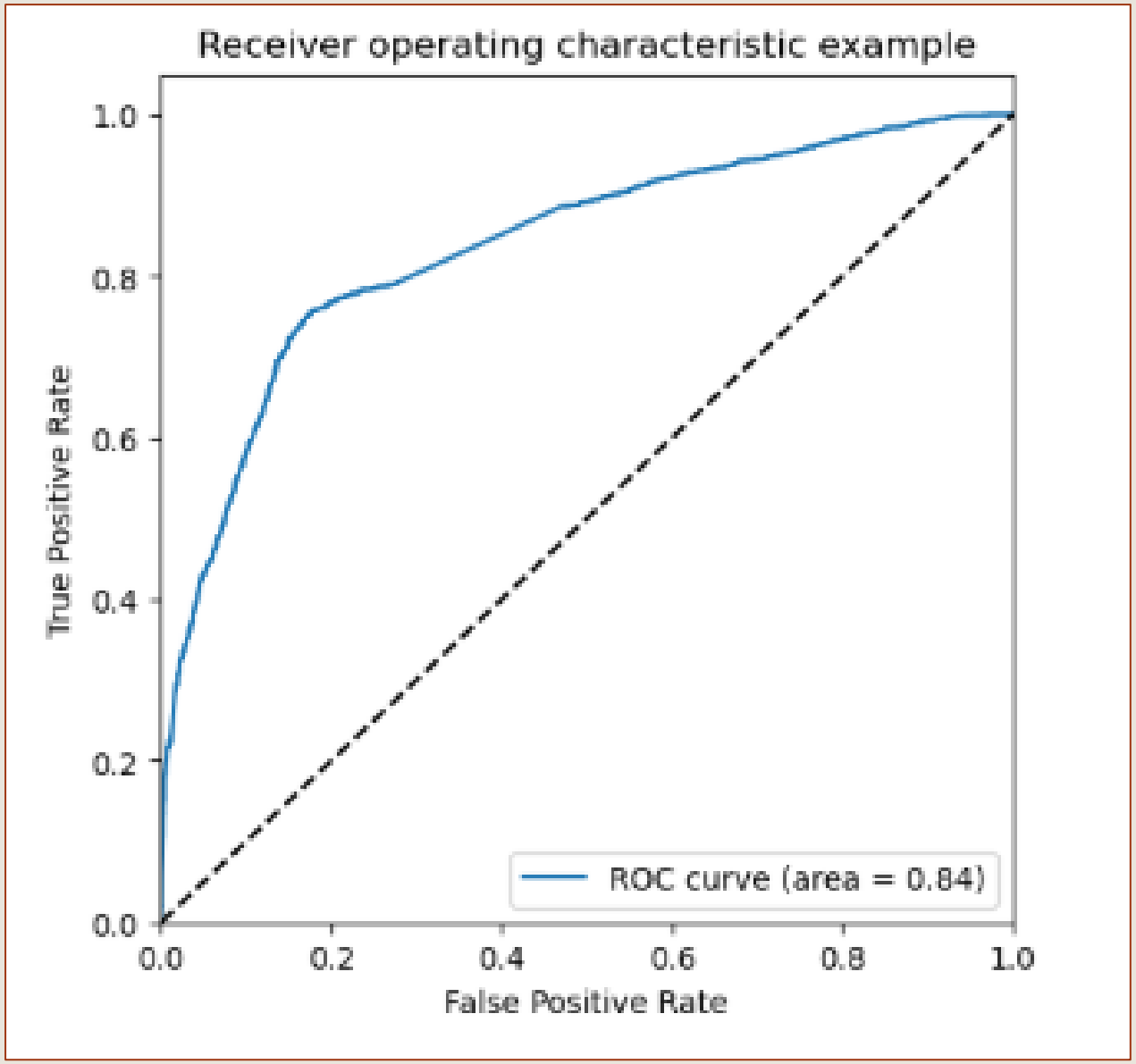
- Total Time Spent on the Website has a positive correlation with Conversion.
- Page Views Per Visit and Total Visits have minimal or no correlation with Conversion.



Graph's conclusion

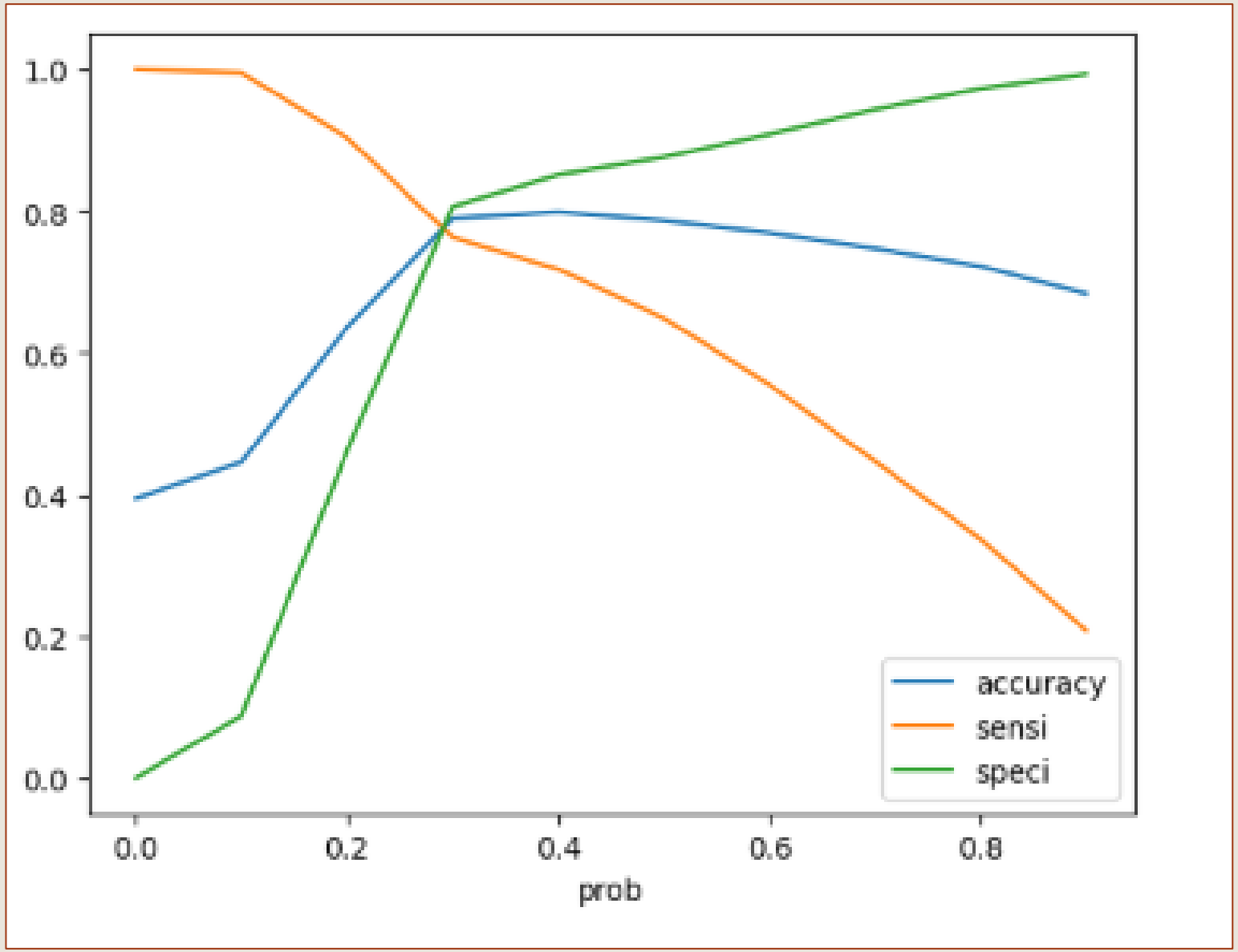
•Leads who spend more time on the website have a higher likelihood of conversion.

•The website should be made more engaging to encourage leads to spend more time.



Graph's conclusion

The area under the ROC curve (AUC-ROC) is 0.84, which indicates a good model performance.



Graph's conclusion

Since the optimal values for the three metrics are achieved at 0.3, the chosen cutoff value is 0.3.

CONCLUSION

•We have evaluated both Sensitivity-Specificity and Precision-Recall metrics, the final prediction cutoff was chosen based on Sensitivity and Specificity.

•The test set results show Accuracy, Sensitivity, and Specificity values of approximately 79%, 76%, and 81%, which are close to the corresponding values from the training set.

•The lead score analysis indicates a conversion rate of around 72% precision and 79% recall in the final predicted model.

•The top three factors influencing lead conversion in the model are:
1.Total time spent on the website.
2.Lead source reference.
3.If the lead is already a student, they are less likely to enroll in another course designed for working professionals.

•Overall, this model performs well.

Thank You