LEAD SCORING CASE STUDY

BY PRAJWAL N, PRANIT ROY
, ANURAG SHARMA

PROBLEM STATEMENT

Industry professionals may purchase online courses from X Education. The business advertises its classes via a number of websites and search engines, including Google.

Upon accessing the website, these individuals may peruse the available courses, complete the course registration form, or view some videos. These folks are categorized as leads when they complete a form with their phone number or email address. Additionally, the business receives leads from previous recommendations.

After obtaining these leads, sales team members begin calling, emailing, and so on. Some of the leads convert during this process, but the majority do not. At X education, the lead conversion rate is typically 30%.

BUSINESS GOAL:

X Education need assistance in identifying the most promising prospects—that is, the leads with the highest likelihood of becoming paying clients.

The business requires a model in which each lead is given a lead score, giving consumers with higher lead scores a better probability of converting, and customers with lower lead scores a lesser chance.

The intended lead conversion rate, as stated by the CEO in particular, is about 80%.

STRATEGIC STEPS

➢ Source the data for analysis
➢ Clean and prepare the data
➢ Exploratory Data Analysis.
➢ Feature Scaling
➢ Splitting the data into Test and Train dataset.
➢ Building a logistic Regression model and calculate Lead Score.
➢ Evaluating the model by using different metrics -Specificity and Sensitivity or Precision and Recall.
➢ Applying the best model in Test data based on the Sensitivity and Specificity Metrics.
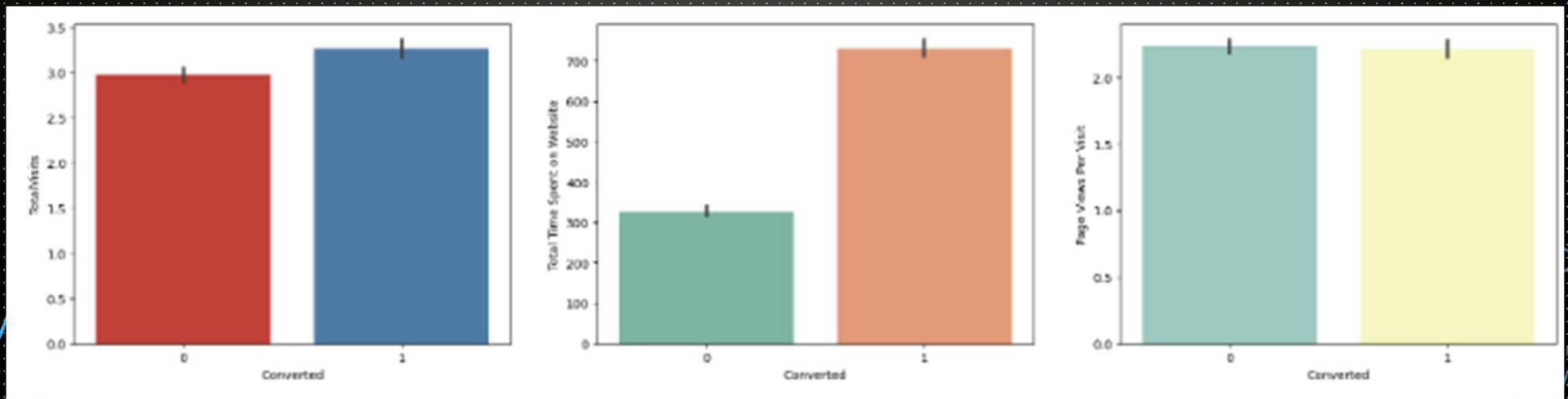
EXPLORATORY DATA ANALYSIS

We have around 39%
Conversion rate in Total.
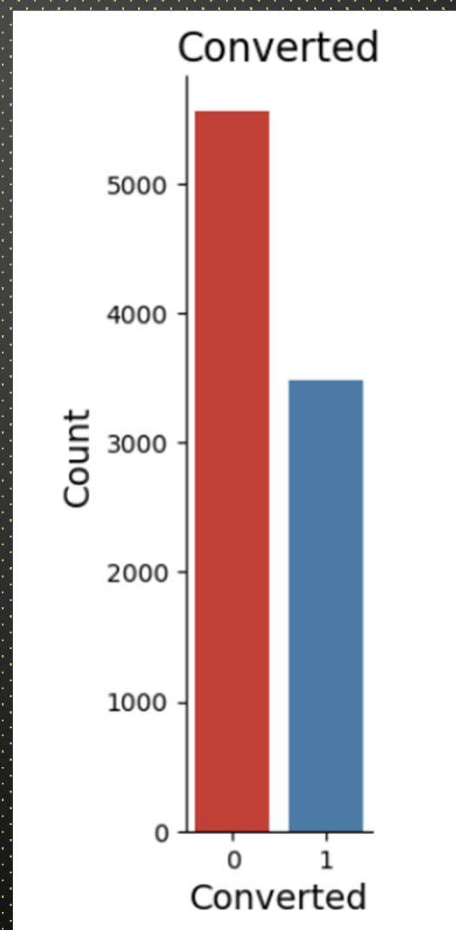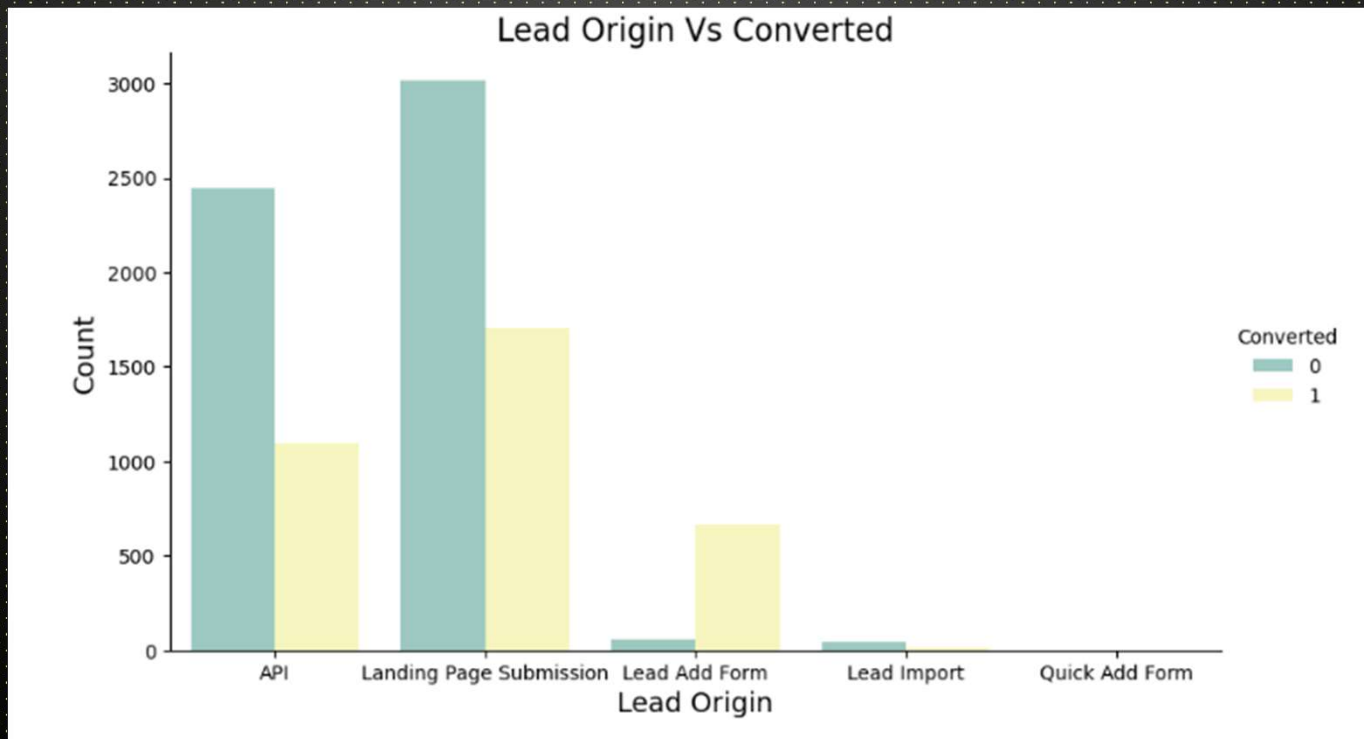
EXPLORATORY DATA ANALYSIS

For Total Visits, Total Time Spent on Website, and Page Views Per Visit, the conversion rates were high.
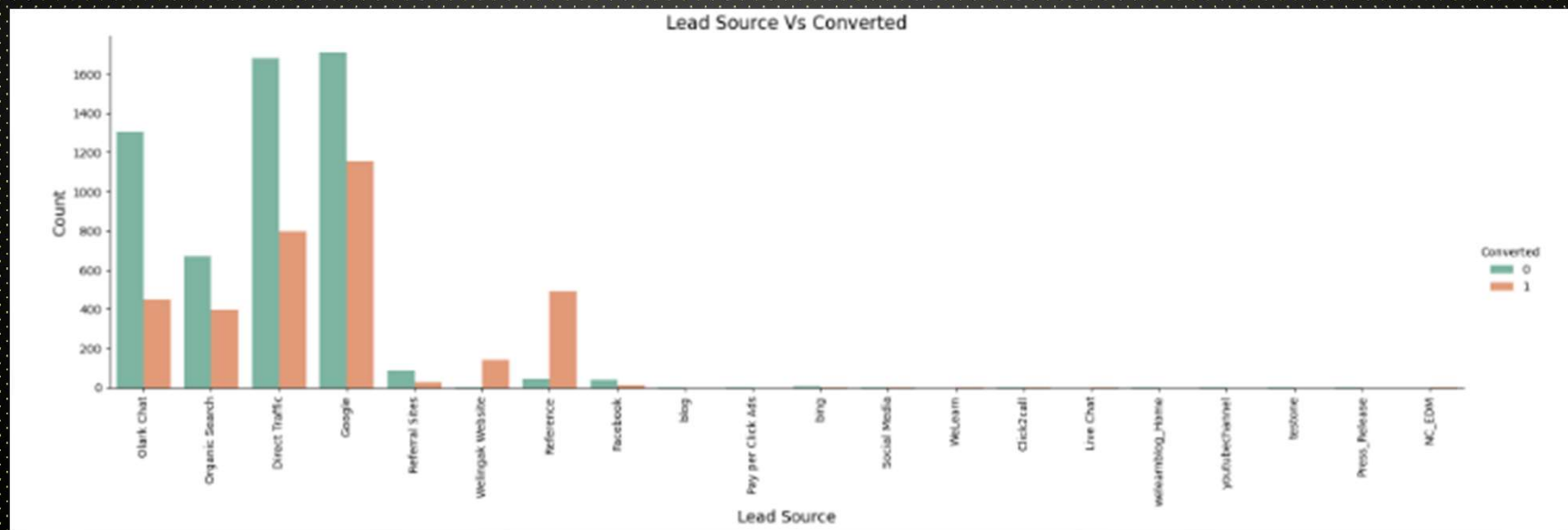
EXPLORATORY DATA ANALYSIS

The greatest conversion rate in Lead Origin occurred from the submission of landing pages.
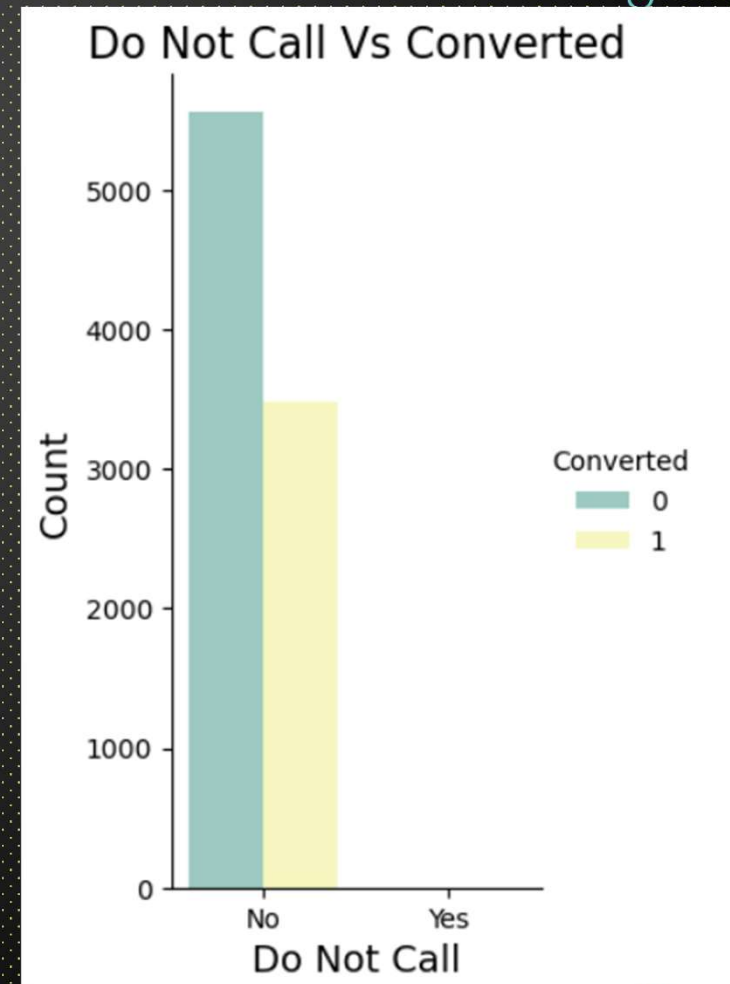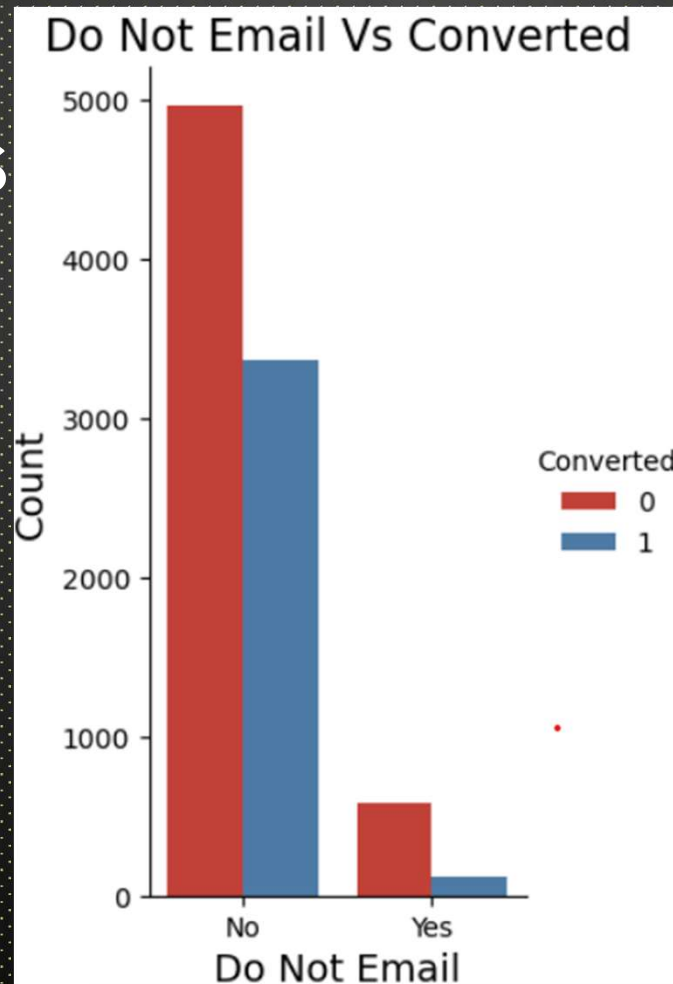
EXPLORATORY DATA ANALYSIS

Google is the primary source with the highest conversion rate.
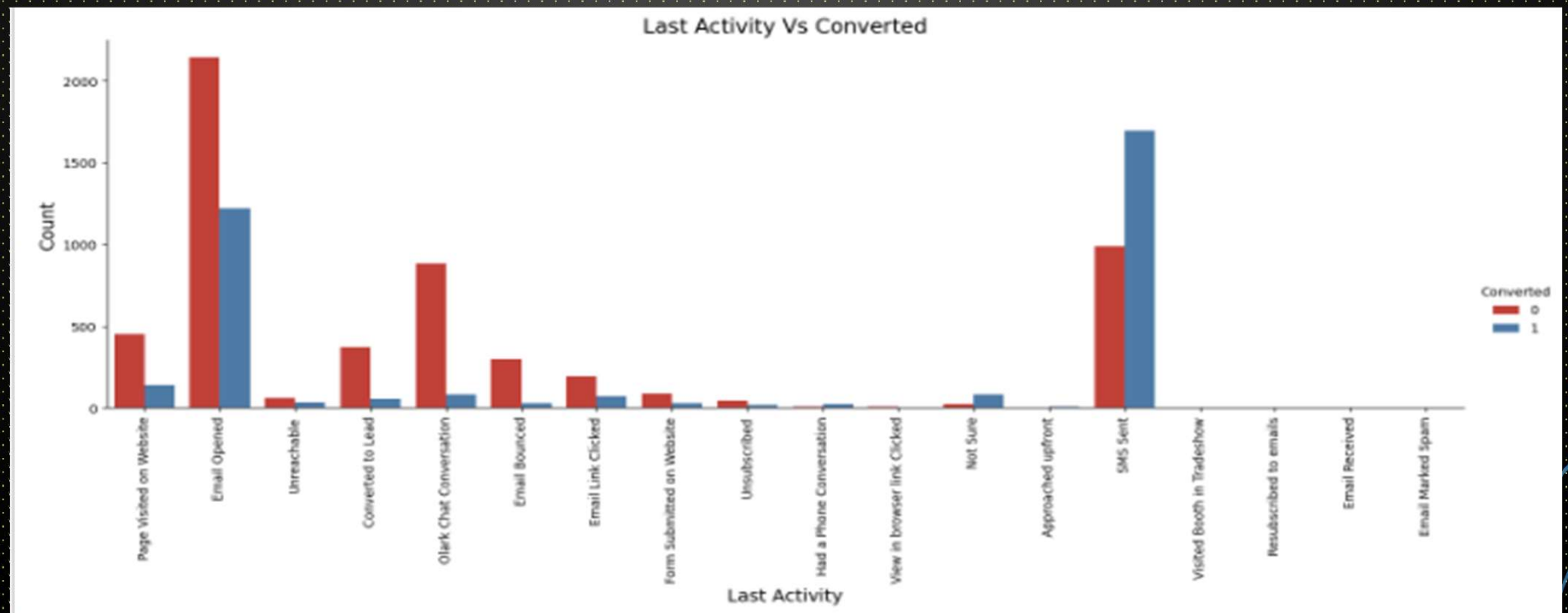
EXPLORATORY DATA ANALYSIS

Significant conversion has occurred from calls and emails exchanged.
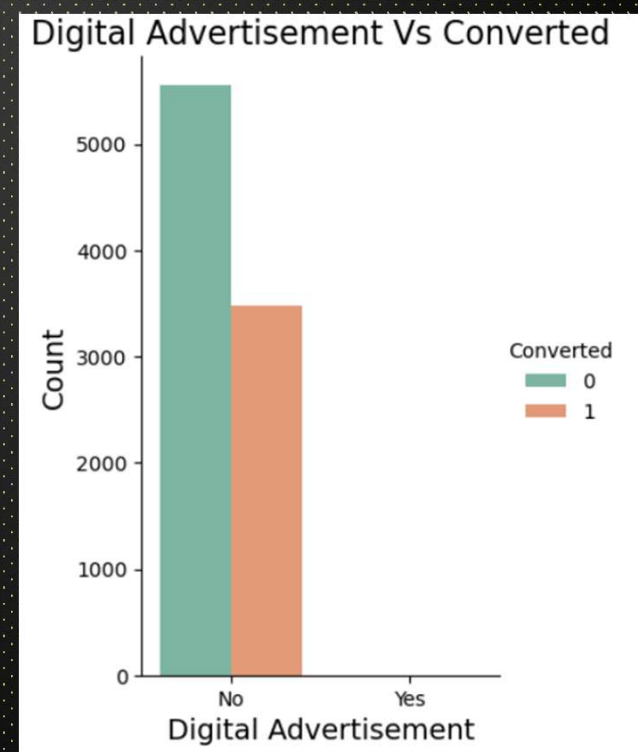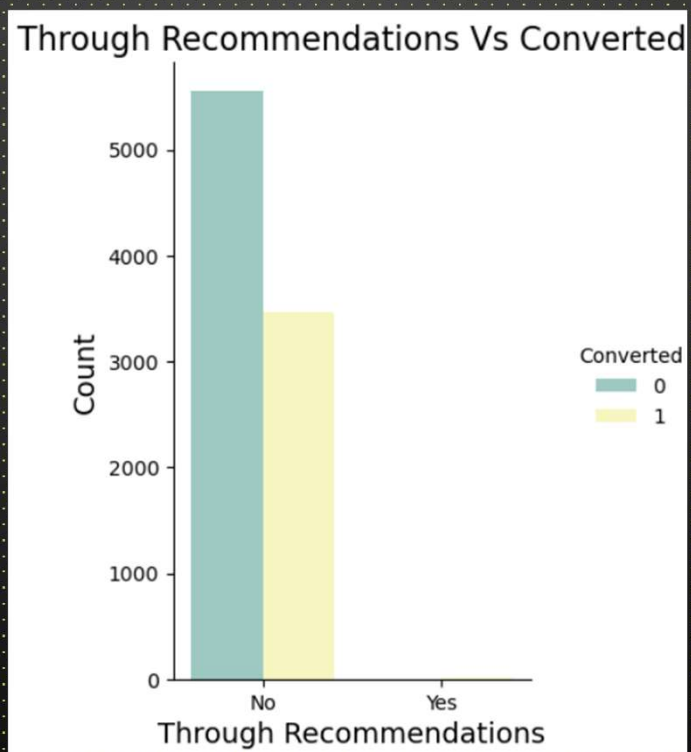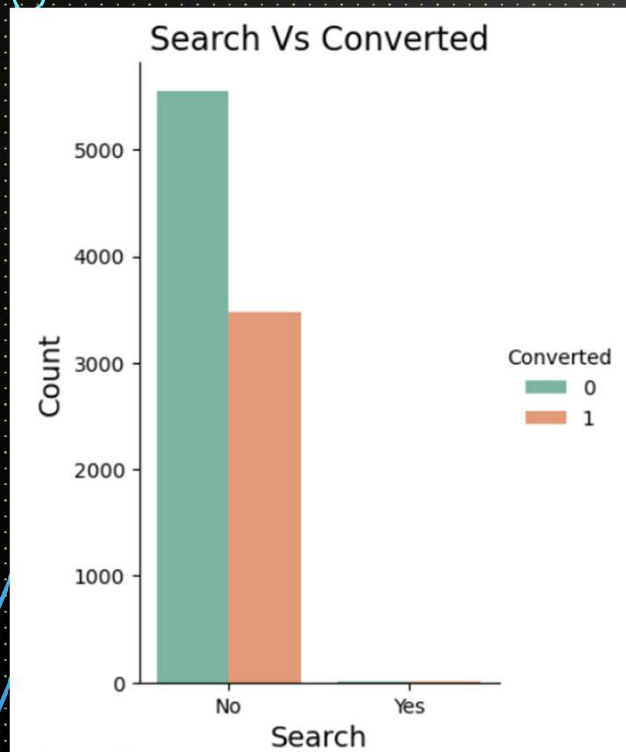
EXPLORATORY DATA ANALYSIS

The SMS Sent's Last Activity value has a higher conversion rate.
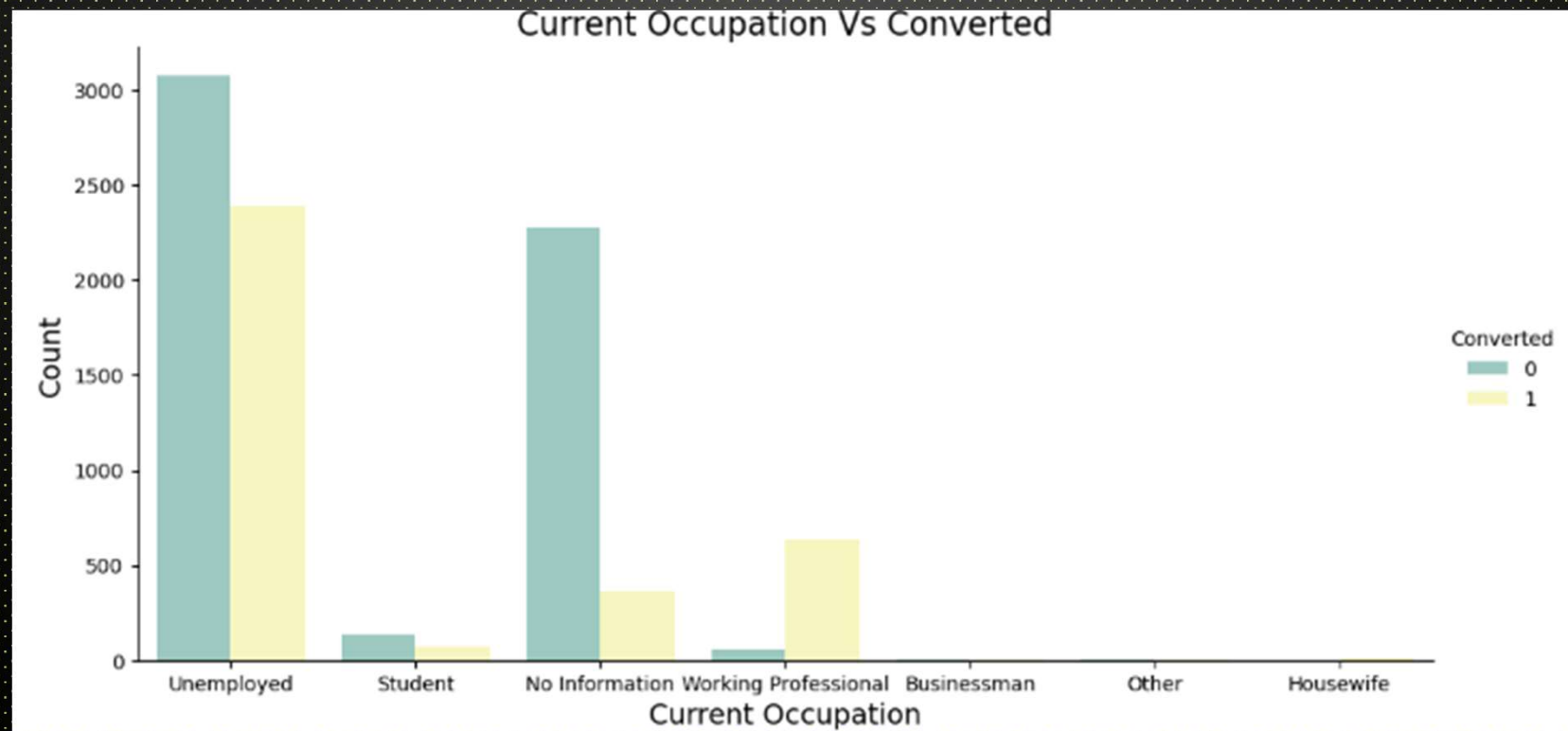
EXPLORATORY DATA ANALYSIS

Little influence on conversion rates from digital ads, search, and recommendations

EXPLORATORY DATA ANALYSIS

Increased conversion occurred with jobless individuals.

CORRELATION FOR TYPE 1

➢ Do Not Email

➢ Total Visits

➢ Total Time Spent on Website

➢ Lead Origin_ Lead Add Form

➢ Lead Source_ Olark Chat

➢ Lead Source_ Welingak Website

➢ Last Activity_ Olark Chat Conversation

➢ Last Activity_ SMS Sent

➢ Current Occupation_ No Information

➢ Current Occupation_ Working Professional

➢ Last Notable Activity_ Had a Phone Conversation

➢ Last Notable Activity_ Unreachable

MODEL EVALUATION

An ideal cut off of 0.42 based on precision and recall is shown in the graph.
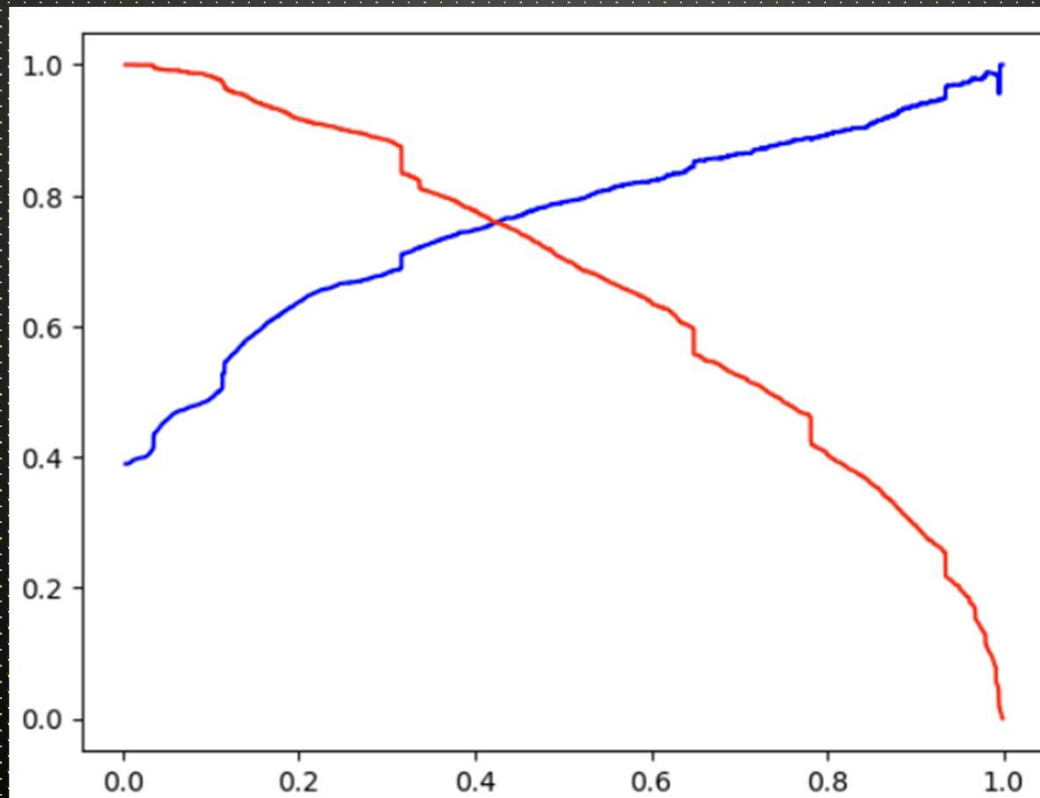
CONCLUSION

➢ The ideal cut off for determining the final prediction was determined by taking into account Sensitivity and Specificity, after checking both Sensitivity-Specificity and Precision and Recall metrics.

➢ The test set's accuracy, sensitivity, and specificity values are around 81%, 79%, and 82%, respectively, which is closer to the corresponding values determined using the training set.

➢ Additionally, the lead score computed in the training data set indicates that the final projected model's conversion rate is about 80%.

THANK YOU