



Lead Scoring Case Study

Presentation by Kartikey Kumar

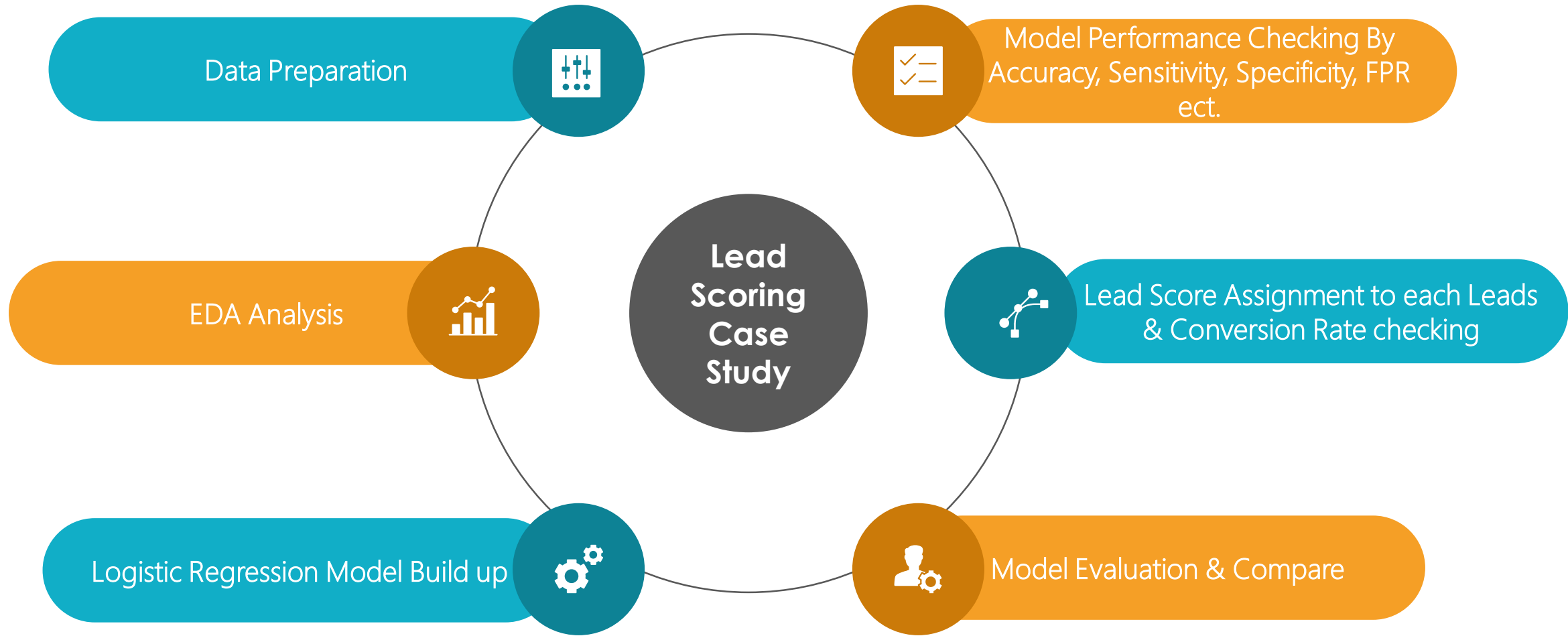
PROBLEM STATEMENT

- X Education offers online courses for purchase by business professionals.
- Leads arrive through website exploration, form submissions, and video views.
- Conversion rate for leads is around 30%.
- Need to identify leads with the best chance of becoming paying clients.

BUSINESS OBJECTIVE:

- Develop a Lead Score model to predict conversion probability.
- Higher Lead Score for better conversion chances, lower for lower chances.
- Targeting desired lead conversion rate of around 80%.

Project Strategy



Project Methodology

DATA PREPARATION

Clean and analyze data by removing duplicates, treating outliers, and conducting exploratory analysis. Standardize features for further analysis.

FEATURE SCALING AND SPLITTING TRAIN AND TEST SETS

"Perform feature scaling on numeric data and then split the dataset into training and test sets."

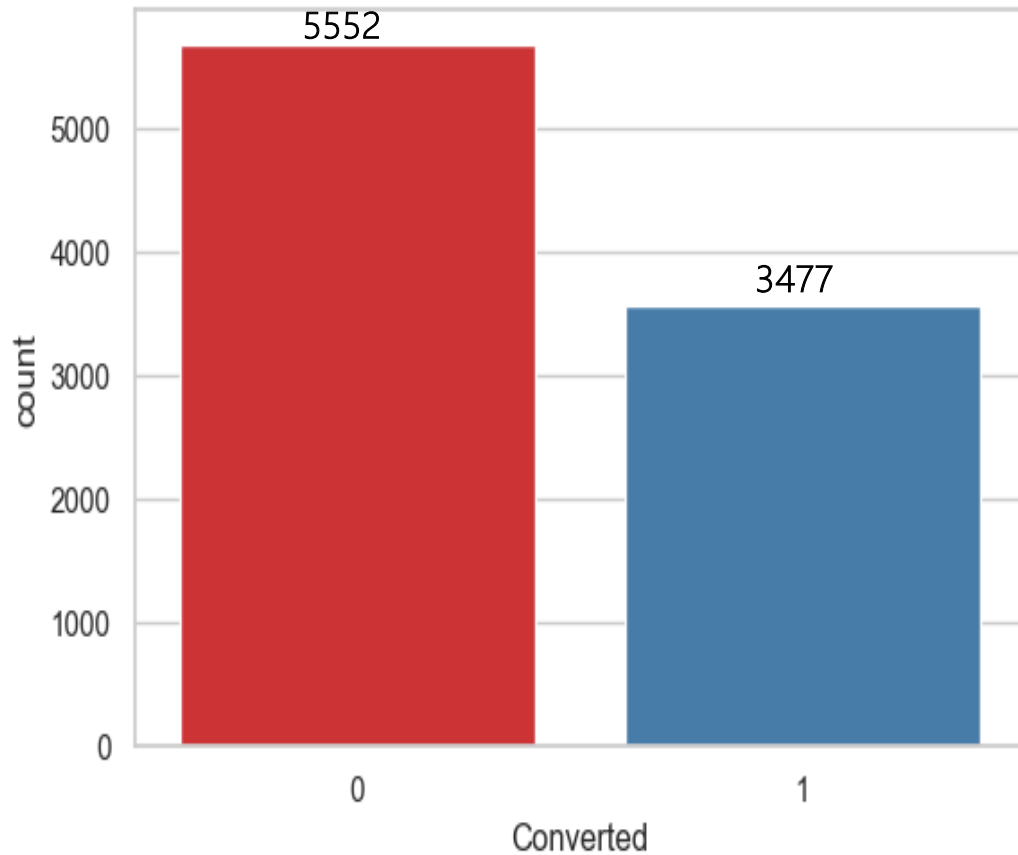
MODEL BUILDING

"Apply RFE for feature selection, determine optimal model using Logistic Regression, and calculate key metrics for model evaluation."

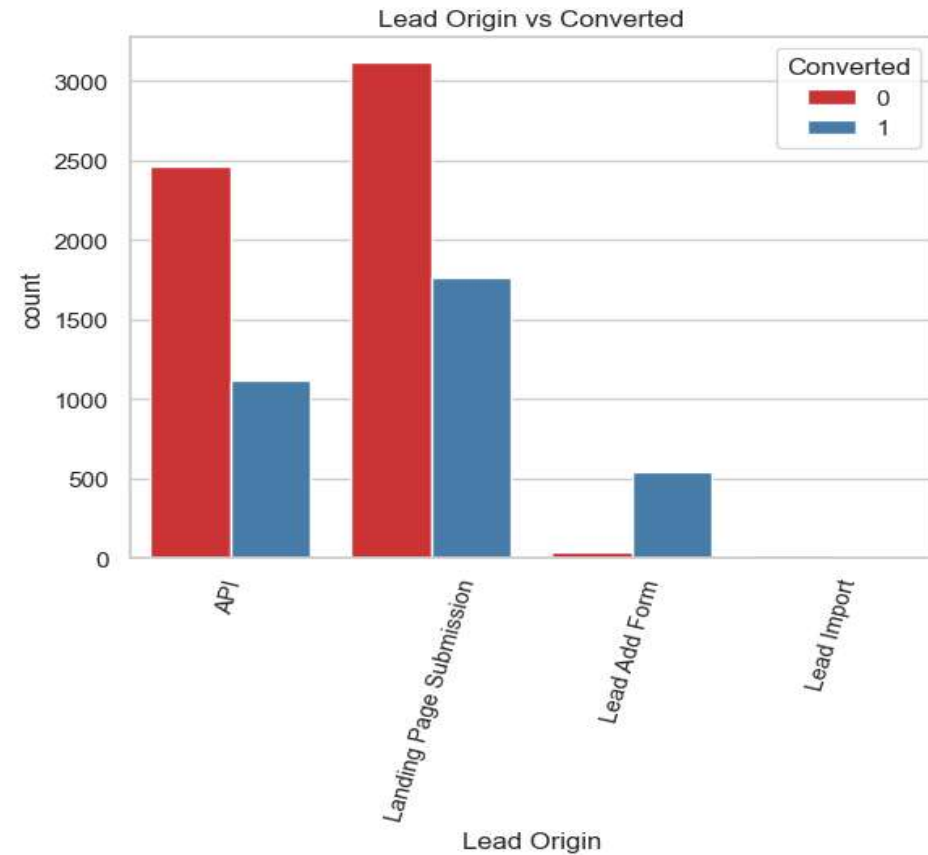
RESULT

"Calculate lead score, check for 80% conversion rate, and evaluate test predictions using sensitivity specificity cutoff."

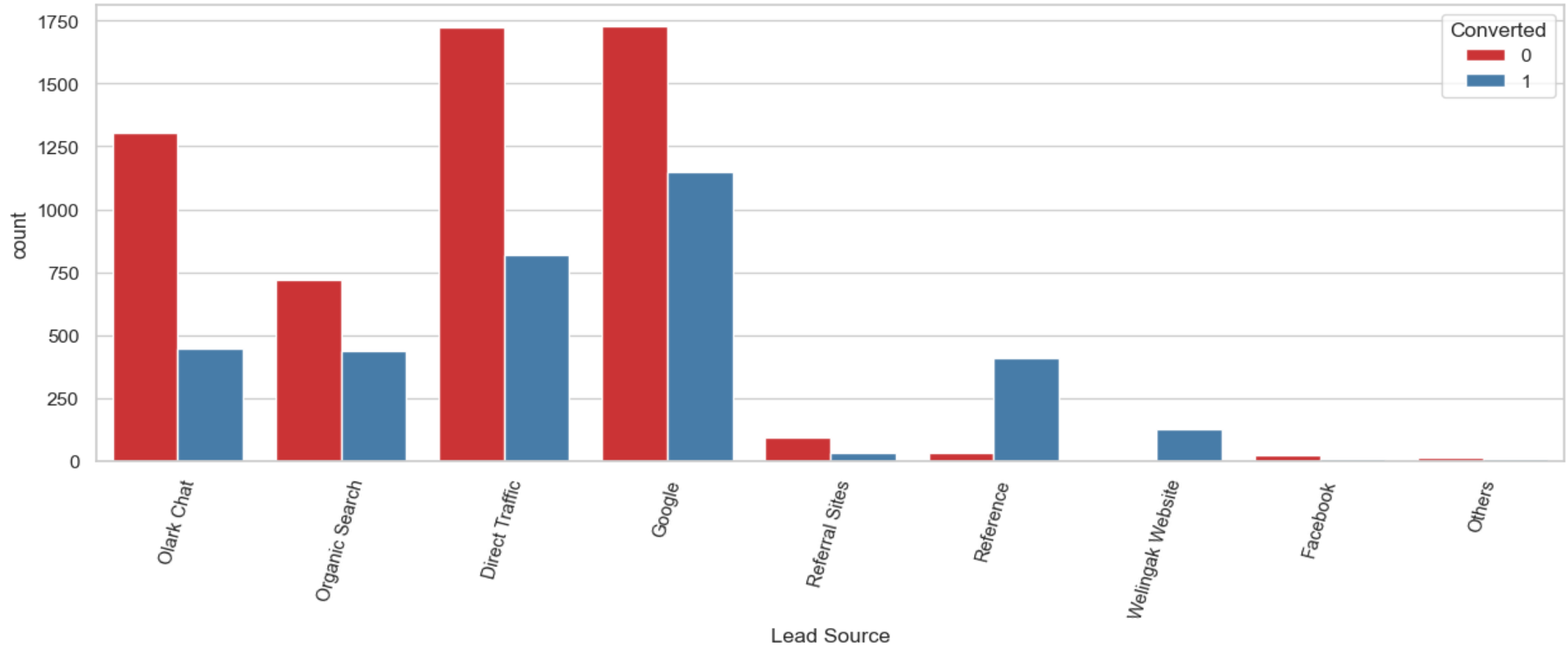
Exploratory Data Analysis



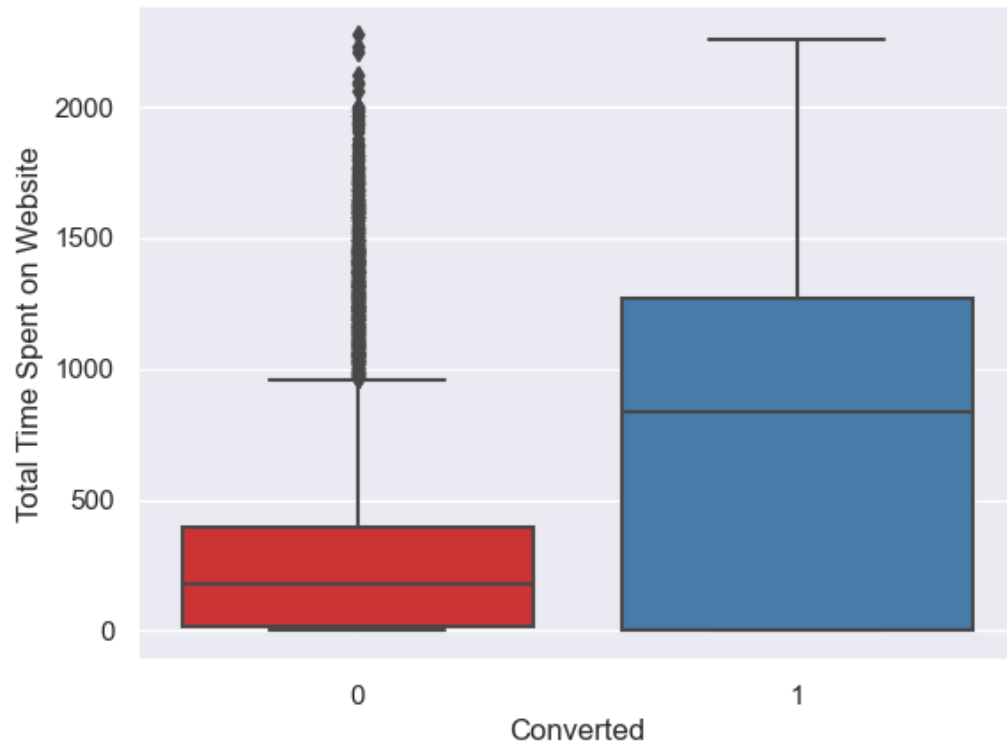
The conversion rate is currently 38.5%



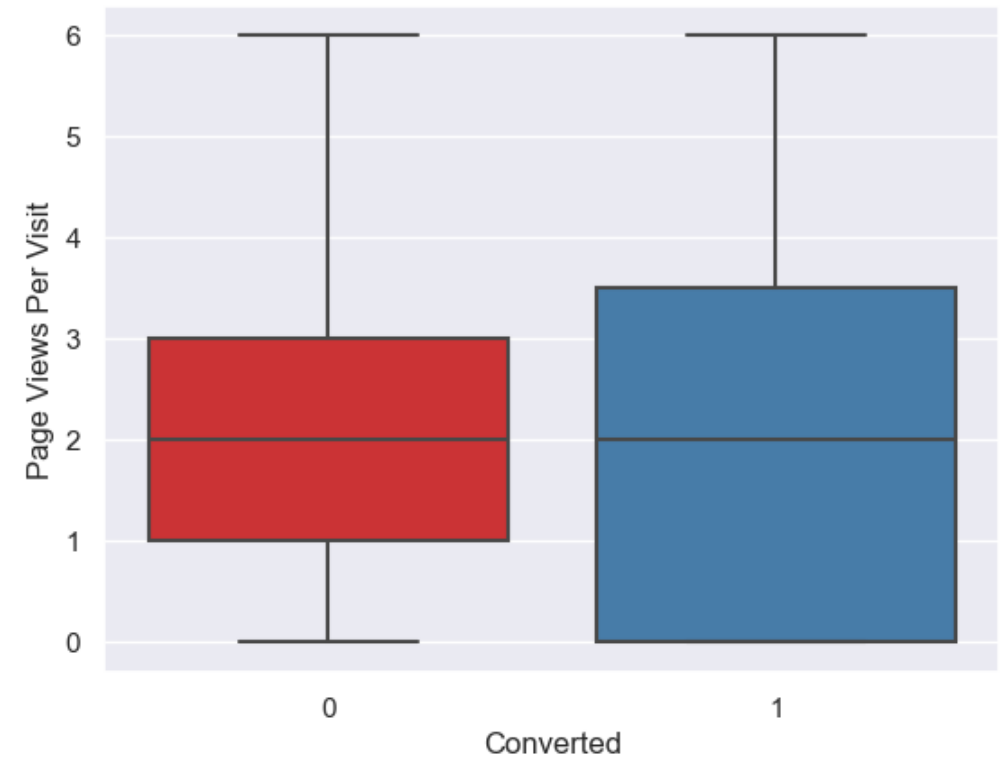
To improve overall lead conversion, focus on enhancing conversion rates from API and Landing Page Submission sources despite their lower rates, and also work on increasing the lead volume from the highly converting Lead Add Form.



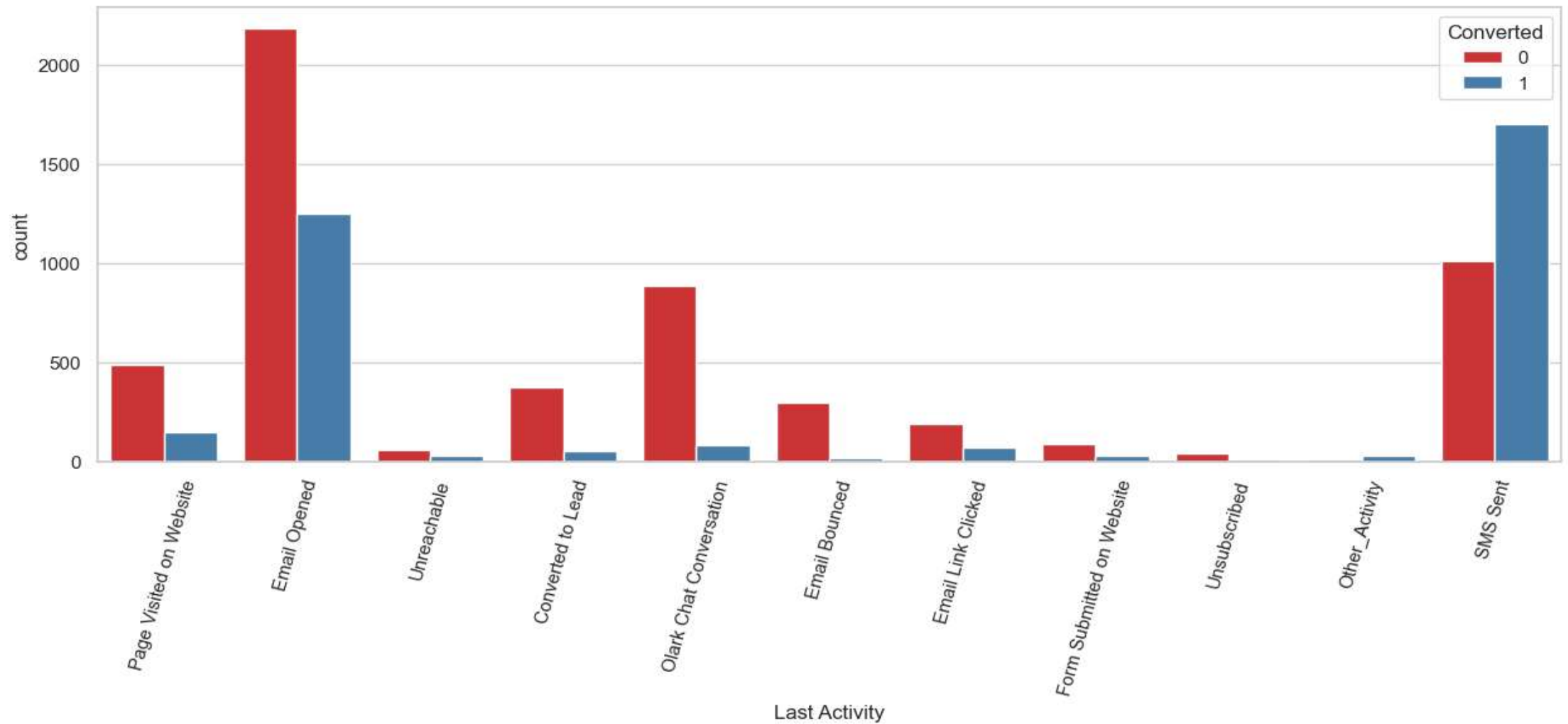
The greatest amount of leads are generated by Google and Direct traffic. Reference leads and leads obtained from the Welingak website have a good conversion rate. Focus should be placed on enhancing lead conversion via olark chat, organic search, direct traffic, and Google leads in order to increase lead conversion overall. More leads should also be generated through the Welingak website and references.



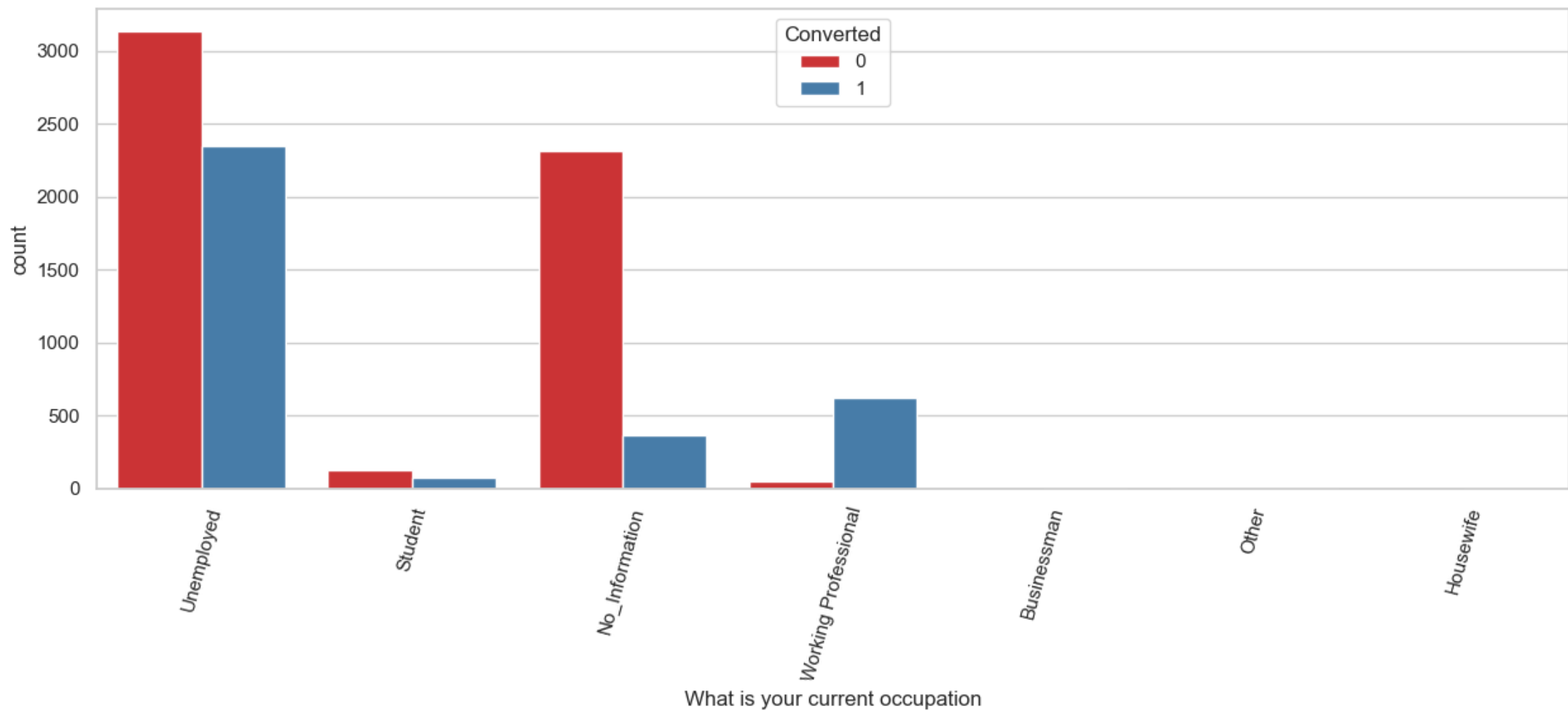
Leads who spend more time on the website have a higher chance of becoming customers. To get visitors to stay on the website longer, it should be made more interesting.



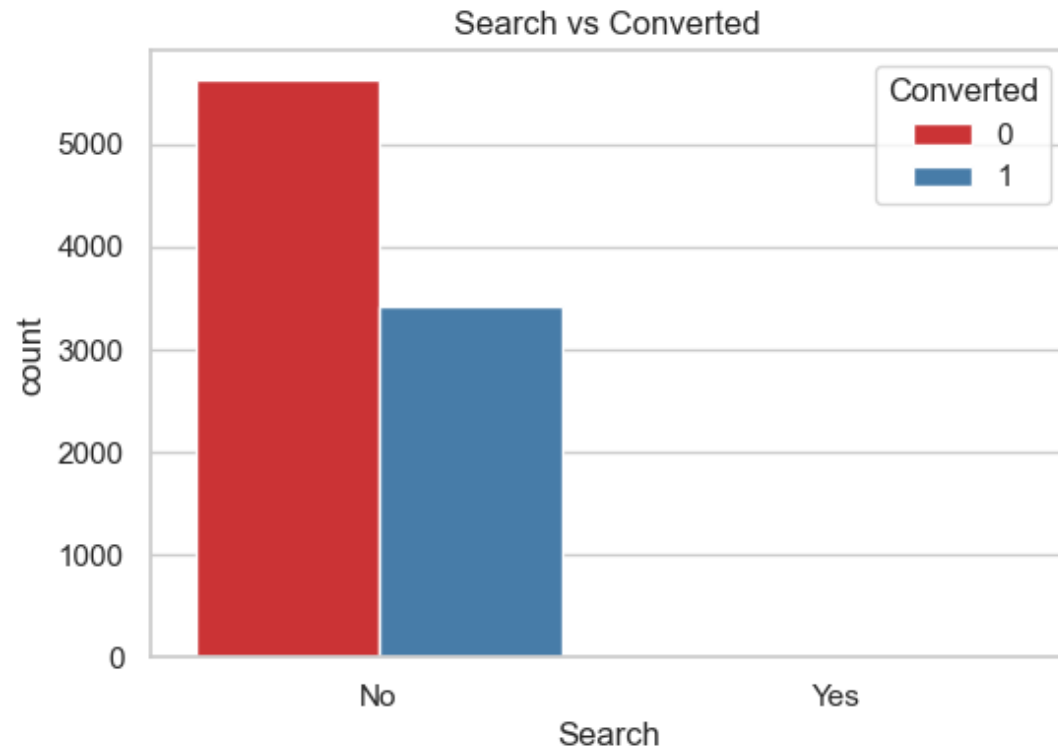
The value of the median for both converted and unconverted leads is the same, inference. Nothing specific about lead conversion from page views per visit can be mentioned.



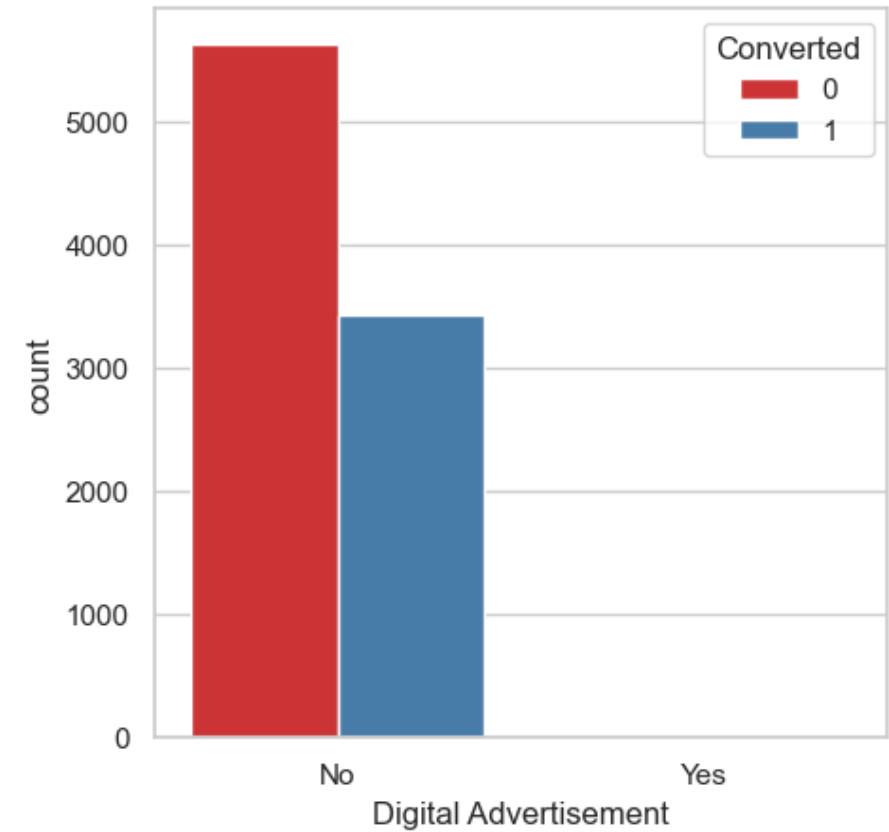
The majority of the leads' most recent action was to open an email. Nearly 60% of leads with an SMS sent as their most recent action convert.



Working Professionals going for the course have a high chance of joining it.
Unemployed leads are the most in numbers but have around 3035% conversion rate.



For the graph of Search vs. Converted, it can be observed that the majority response is NO which leads to a higher chance of nonconversion and lower chances of conversion



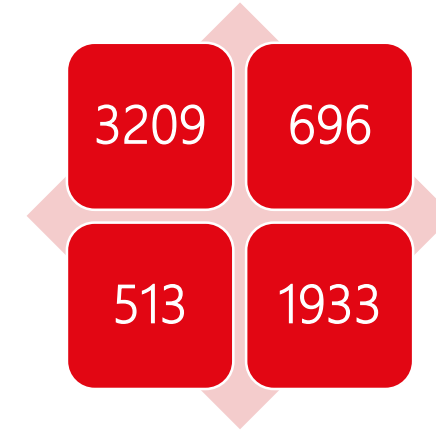
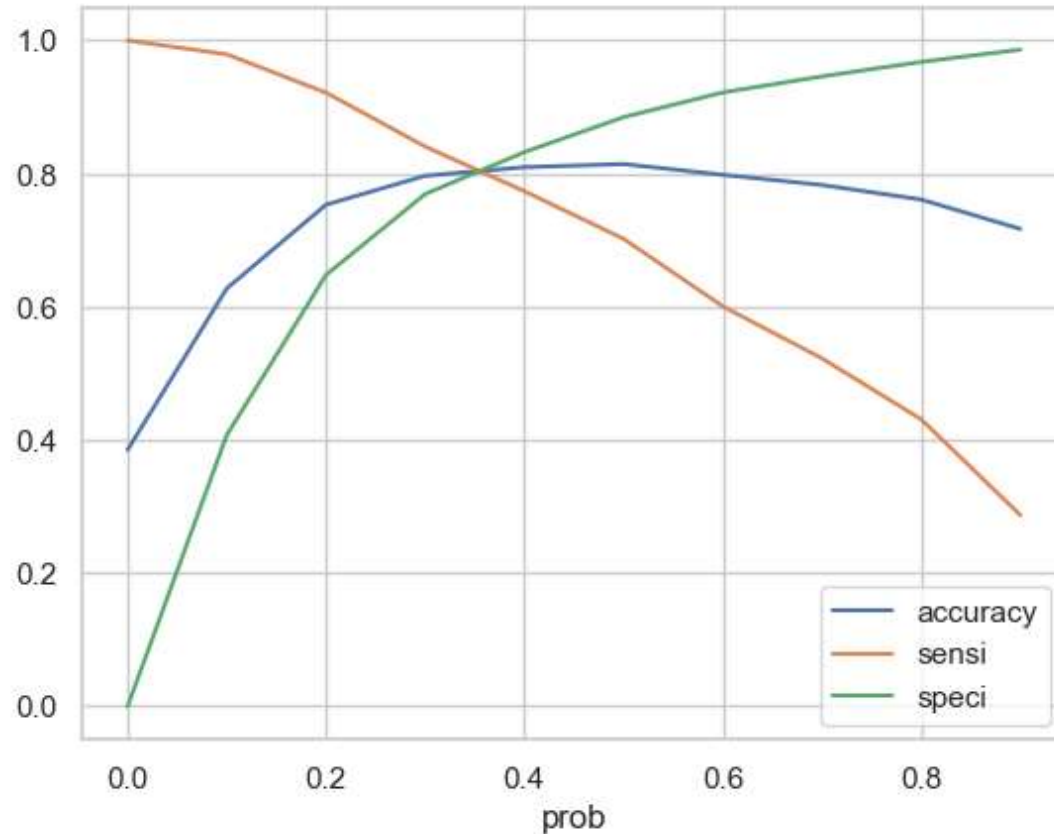
The majority of replies are No, however, there is a significant conversion rate among those responses. We thus cannot disregard this reality.

Variables Impacting the Conversion Rate

- Do Not Email
- Total Time Spent On Websites
- Lead Origin Landing Page Submission
- Lead Origin Lead Add Form
- Lead Source Olark Chat
- Lead Source_Welingak Website
- Lead Activity_Other Activity
- SMS Sent
- Unsubscribed
- Specialization Other Specialization
- What is your current Occupation No Information
- What is your current Occupation Working Professional
- Last Notable Activity Olark Chat Conversation
- Last Notable Activity Unreachable

X Education seeks an 80% lead conversion rate by leveraging variables like 'Lead Origin,' 'Lead Source,' engagement activities, and occupation preferences. Through personalized communication and a Lead Score model, they aim to prioritize high-conversion leads and optimize lead generation sources for improved overall conversion success.

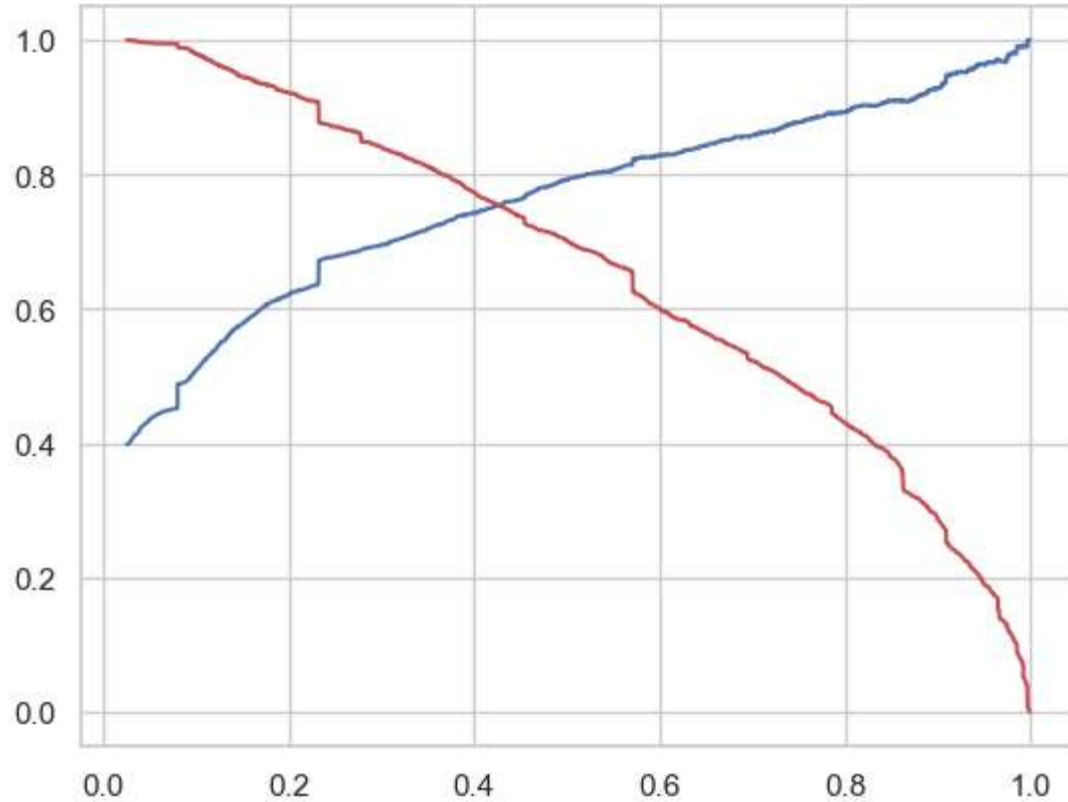
Model Evaluation -Sensitivity and Specificity on Train Data Set



The graph illustrates an optimal cut-off of 0.38 based on Accuracy, Sensitivity, and Specificity.

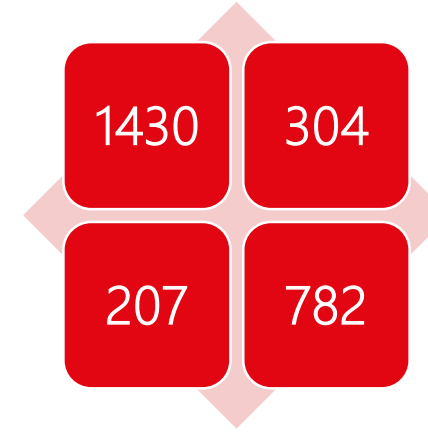
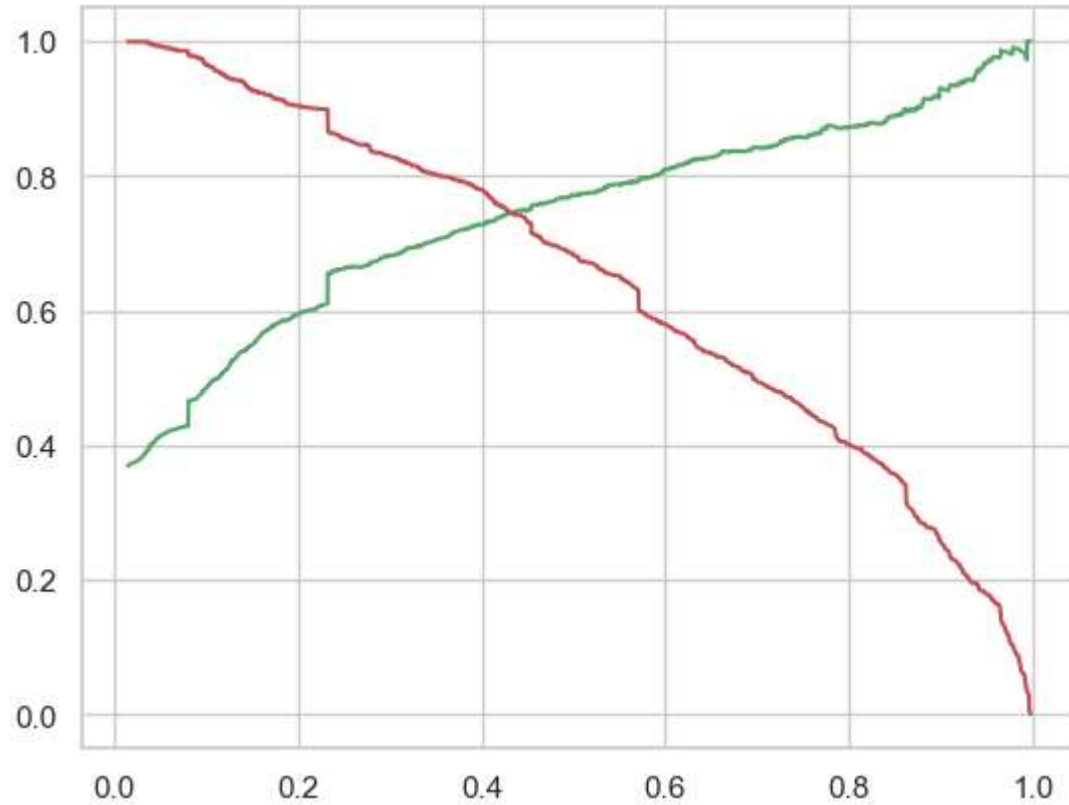
Based on the results, the accuracy is 81%, sensitivity is 79%, and specificity is 82%. The false positive rate stands at 17%. Additionally, the positive predictive value is 73.5%, while the other one is 86.2%.

Model Evaluation -Sensitivity and Specificity on Train Data Set



With a precision rate of 79%, the model accurately identifies a significant portion of positive predictions. However, the recall rate of 70% suggests that it may miss identifying some true positive instances. Balancing precision and recall is crucial to achieve a well-rounded model performance.

Model Evaluation-Precision and Recall on Test Dataset



The graph illustrates an optimal cut-off of 0.42 based on Precision and Recall

The precision rate is 72%, while the recall rate is 79%.

Key Observations and Model Performance:

- Accuracy of approximately 81.43% with 0.5 cutoff.
- Sensitivity, Specificity, FPR all at 70.19%.
- Train set: Sensitivity 79%, Specificity 82%, Accuracy 81%.
- The ROC curve's AUC is likely around 0.89.
- Similar results were observed on the Test Set.

Conversion Rate Analysis:

- Lead Score allocation resulted in an 80% conversion rate.
- Comparable conversion rate of about 80% on the Test Set.

Here are some ways to improve your conversion rates:

1. Focus on sources that convert well, such as the 'Lead Add Form'.
2. Personalize your outreach by taking into account the recipient's profession, engagement activities, and specialization.
3. Experiment with different cutoff points to strike a balance between precision and recall, which can improve your model's performance.
4. Continuously gather feedback from users and use it to refine your predictions.
5. Engage with leads based on their activity to nurture warmer leads.
6. Use insights to optimize your lead generation strategies by targeting sources with high conversion potential.

Overall, your model is performing well. By strategically optimizing your efforts and personalizing your engagement, you can further increase your conversion rates and drive business growth.



Thank You