



# **Lead Scoring Case Study using Logistic Regression**

**By**

**Kavitha Choudhary**



# Problem Statement

X Education sells online courses to professionals. Many visit the website, become leads after form filling, and are contacted by the sales team. About 30% of leads convert. To boost efficiency, the company wants to pinpoint high-potential "Hot Leads." This targeted approach should increase the conversion rate as the sales team focuses on promising leads.

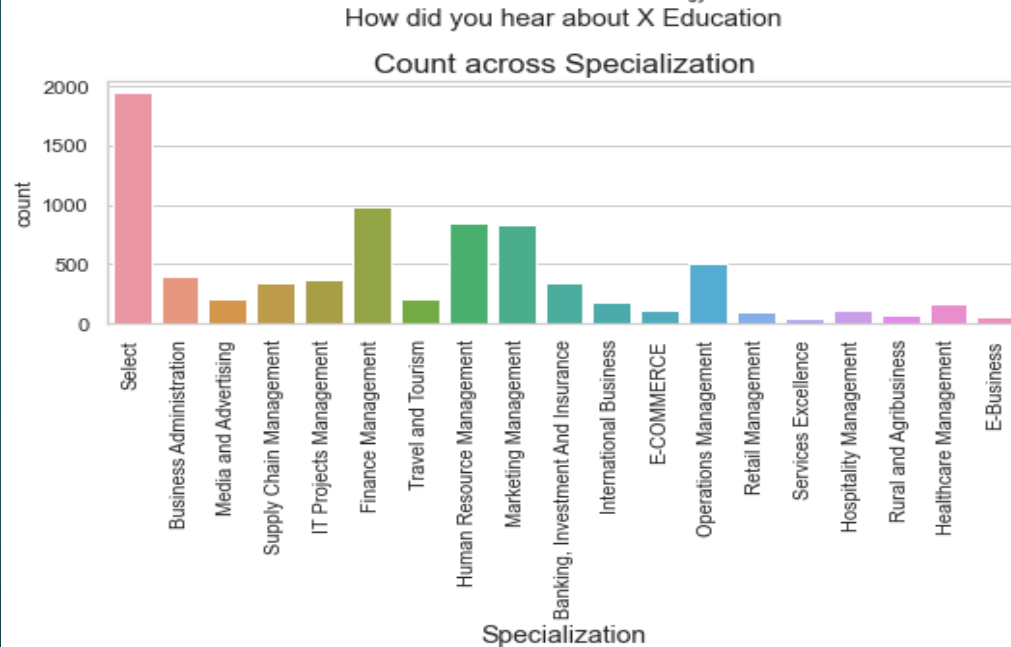
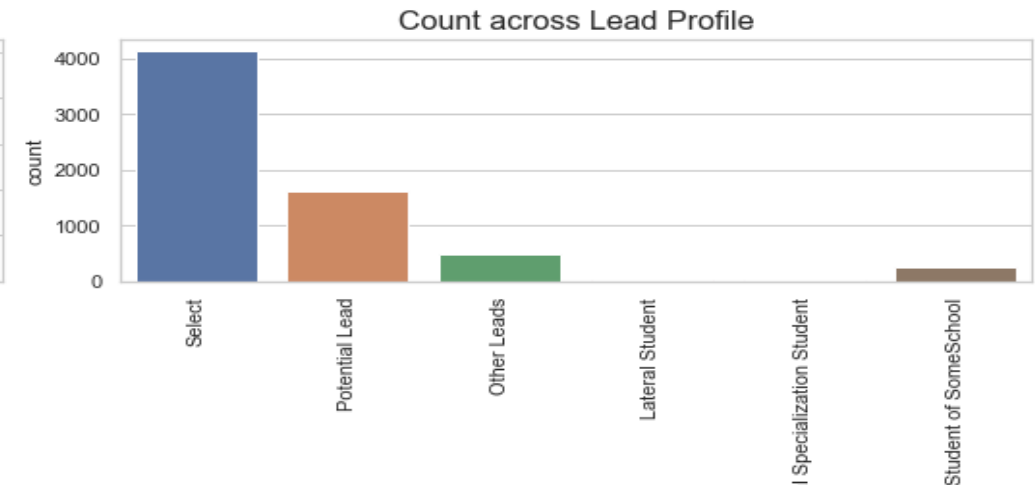
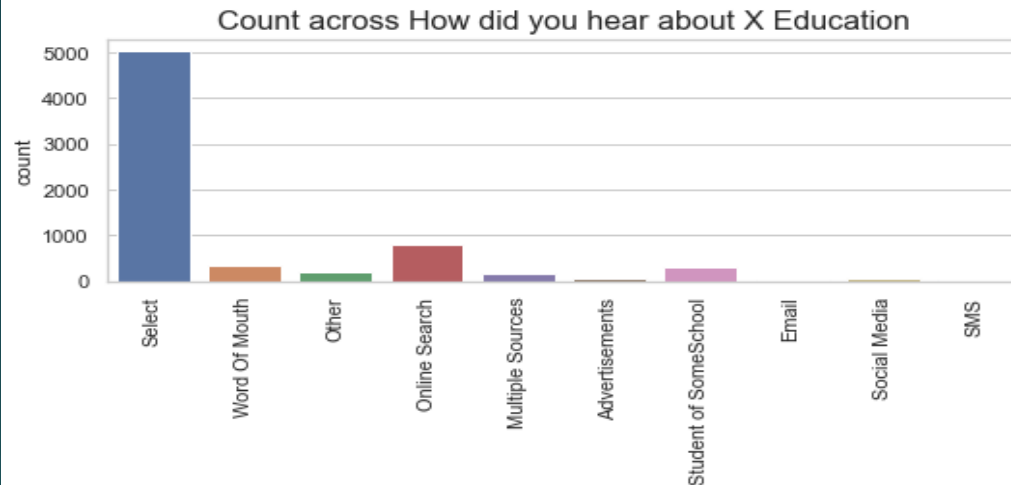


# Business Objective

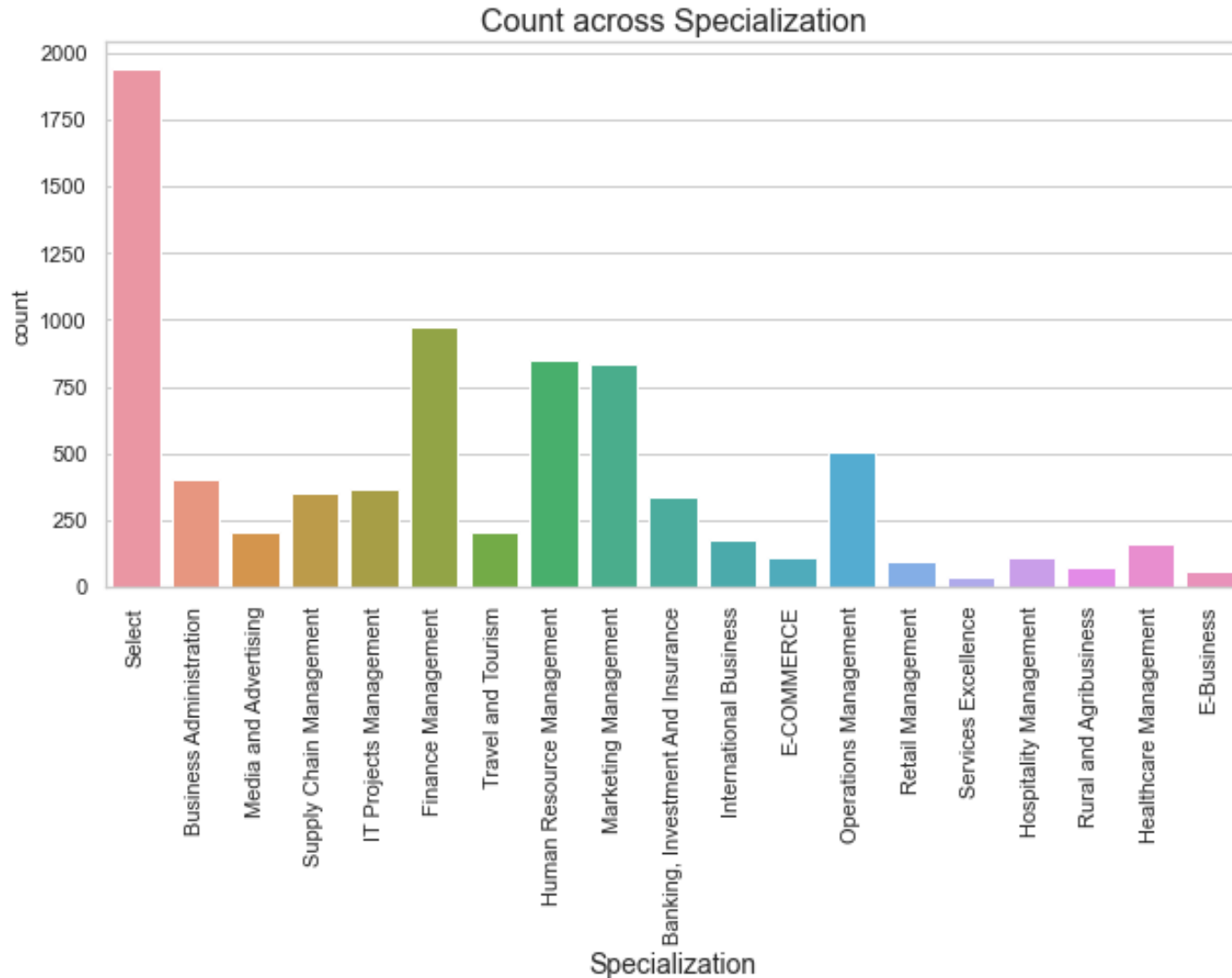
Lead X seeks a model to assign lead scores (0-100) for identifying Hot leads and boosting the conversion rate, aiming for 80%. The model should adapt to future needs, addressing peak times, resource optimization, and post-target approaches.

CEO's goal: enhance conversion and efficiency.

# EDA – Data Cleaning

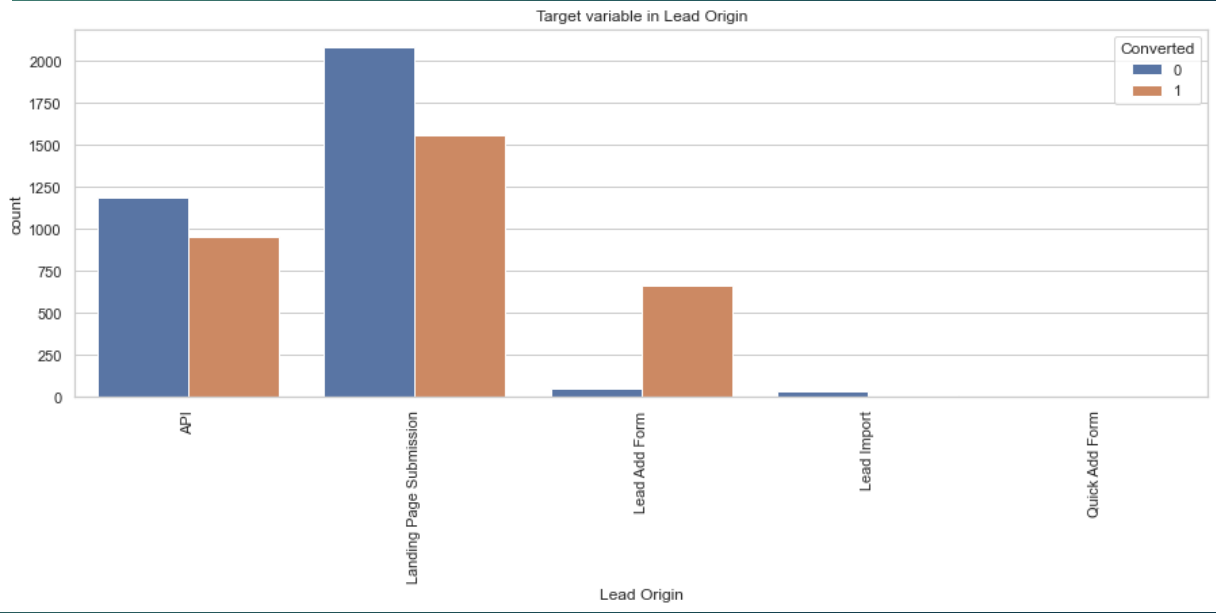
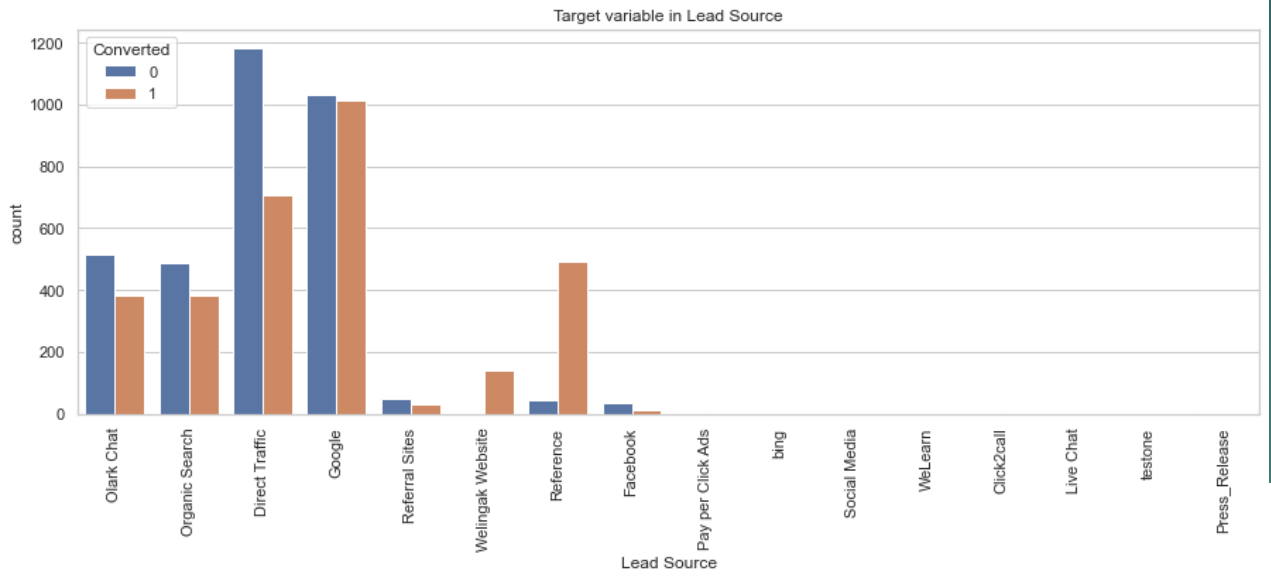


# Specialization

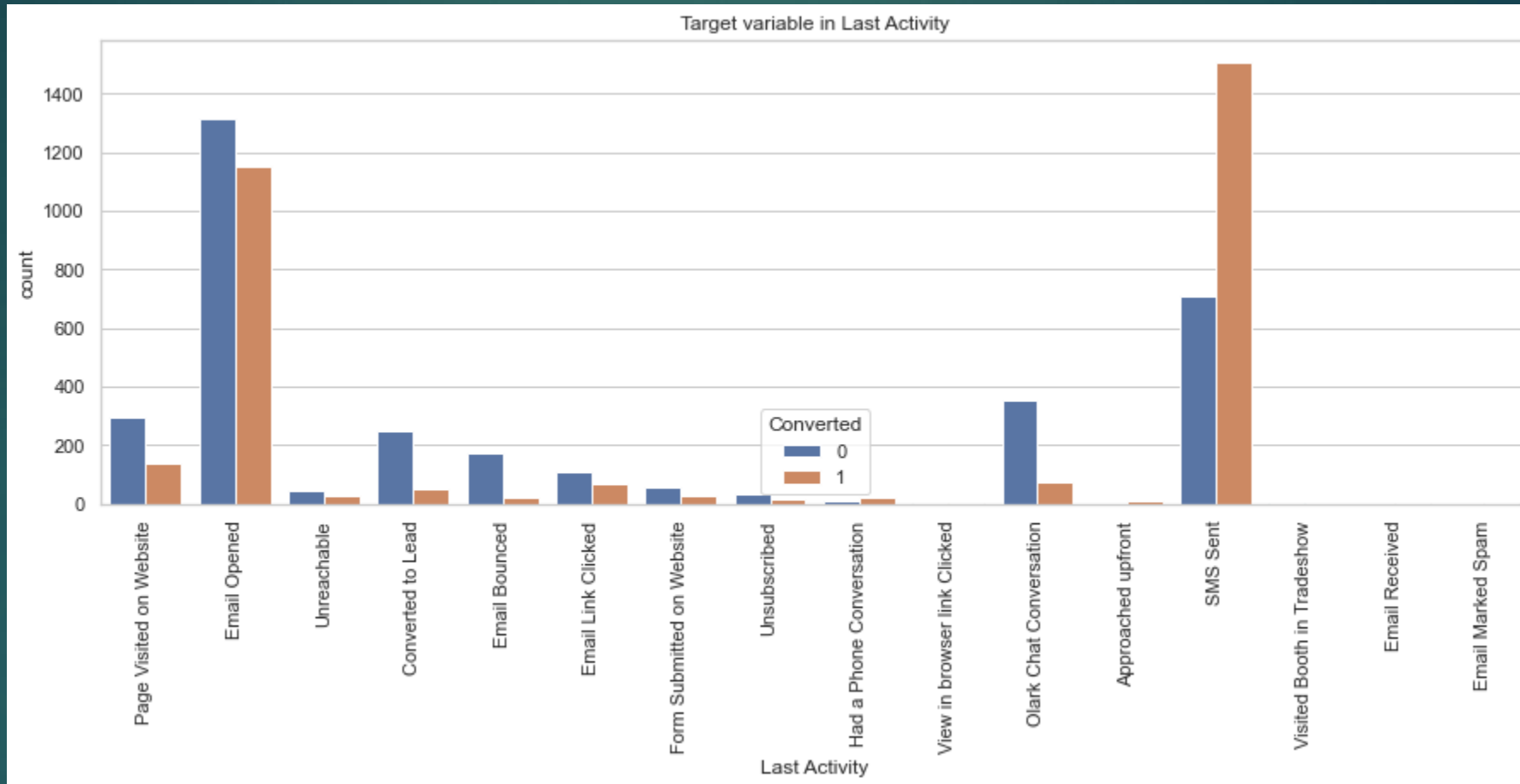


Leads with specializations in HR, Finance, and Marketing management exhibit a strong likelihood of conversion.

# Lead Source & Lead Origin



# Lead Last Activity



**Leads that engage with emails show a high likelihood of conversion, and utilizing SMS also proves beneficial.**

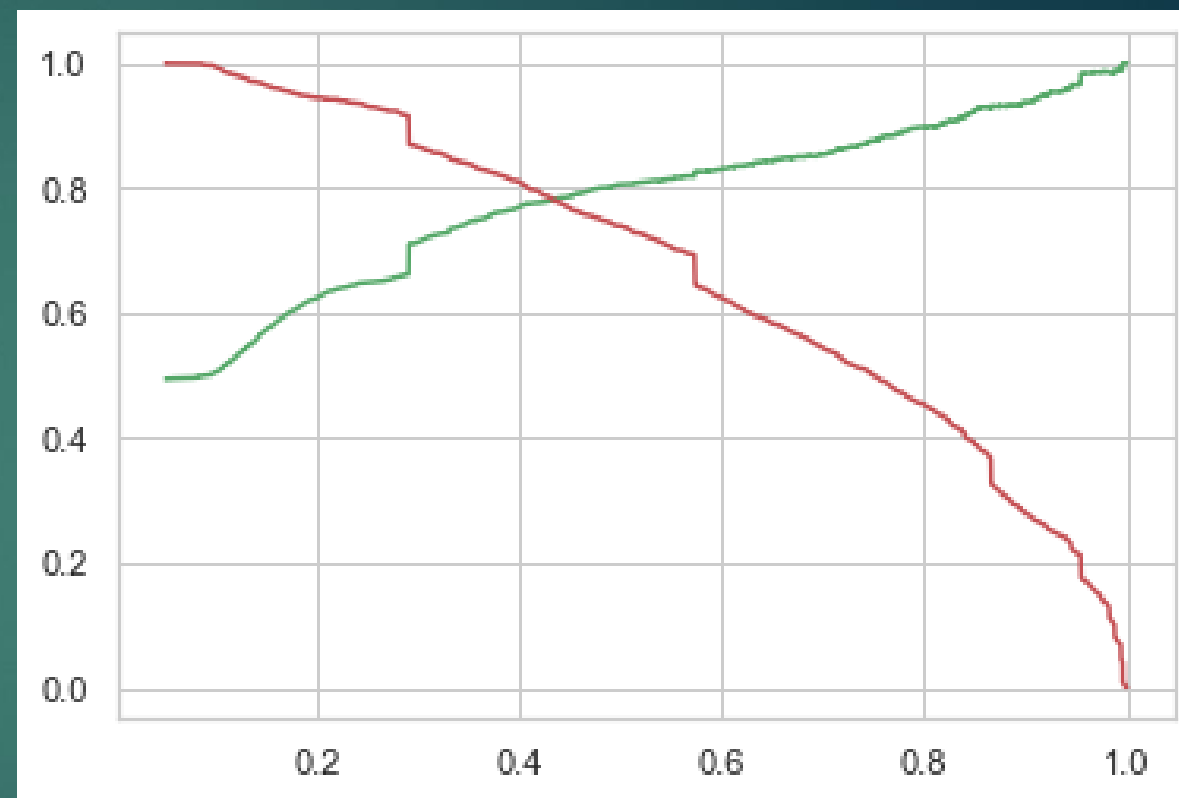
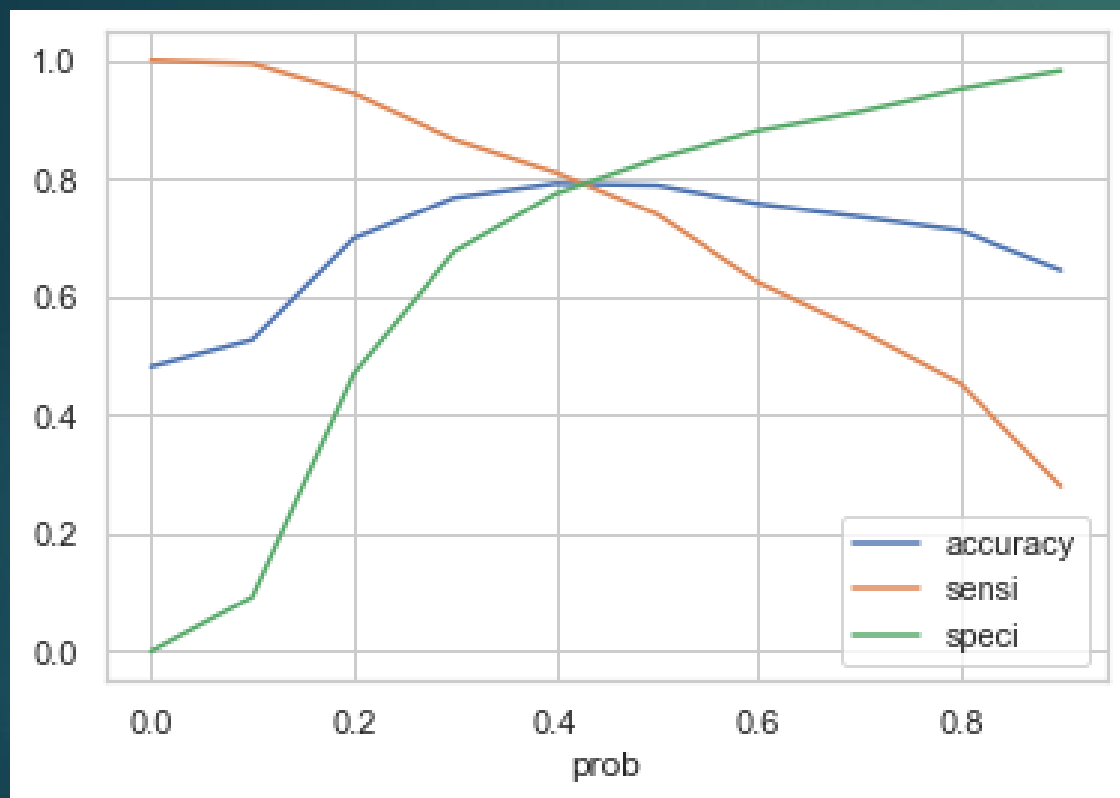
# Correlation



**There is no correlation between the variables.**



# Model Evaluation – ROC Curve



**The tradeoff between Precision and Recall is at 0.42. Hence, it's reasonable to classify any Prospect Lead with a Conversion Probability exceeding 42% as a hot Lead.**

# Observations

## **Train Data:**

Accuracy : 80%

Sensitivity : 77%

Specificity : 80%

## **Test Data:**

Accuracy : 80%

Sensitivity : 77%

Specificity : 80%

## **Final Features list:**

Lead Source\_Olark Chat

Specialization\_Others

Lead Origin\_Lead Add Form

Lead Source\_Welingak Website

Total Time Spent on Website

Lead Origin\_Landing Page Submission

What is your current occupation\_Working Professionals

Do Not Email

# Conclusion

- Conversion rates for API and Landing page submissions hover around 30-35%, aligning with the average. However, Lead Add form and Lead import exhibit notably lower conversion rates. Consequently, a strategic focus on leads originating from API and Landing page submissions is warranted.
- Google/Direct traffic yields the highest lead generation, with the greatest conversion ratio stemming from references and the Welingak website.
- Leads spending extended time on the website demonstrate an increased likelihood of conversion.
- Commonly observed last activity is email opening, with the highest conversion rate linked to SMS transmission. A significant portion of leads are unemployed, while the highest conversion is seen among working professionals.