

Lead Scoring

Case Study

An analysis by:

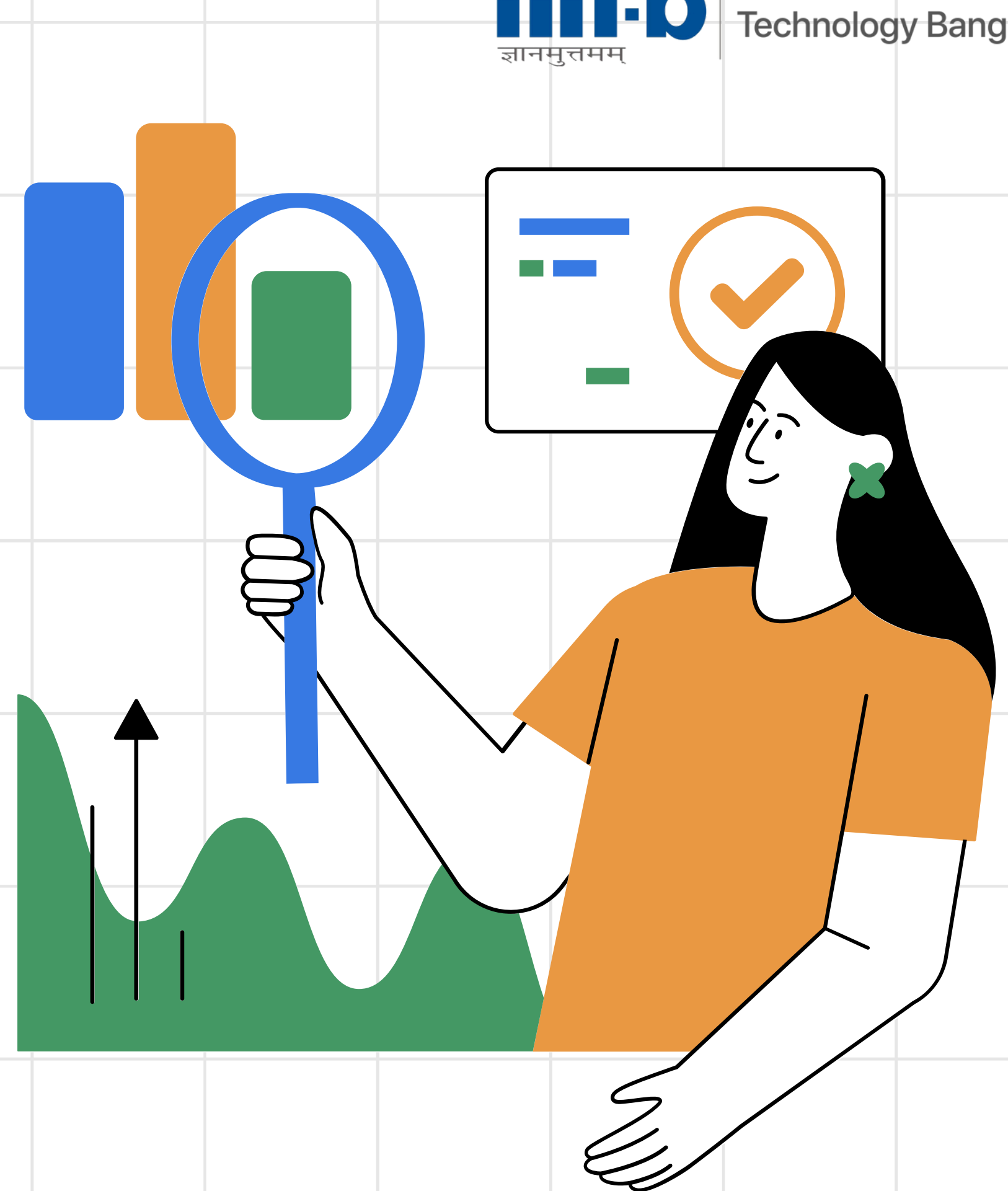
Omkar Dharkar

Shubhankar Chakravarty

Sachin Gayakwad

Shifa Syed

18th February 2025





Outline

01

EXECUTIVE SUMMARY

Summarizes the process and the report

02

INTRODUCTION

Presenting the problem statement

03

METHODOLOGY

An overview of the objectives of the study and the steps followed to reach the conclusions

04

RESULTS

Findings of EDA and visualizations

05

RECOMMENDATION

A condensed summary of ideas, viewpoints and arguments

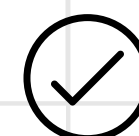
01

EXECUTIVE SUMMARY



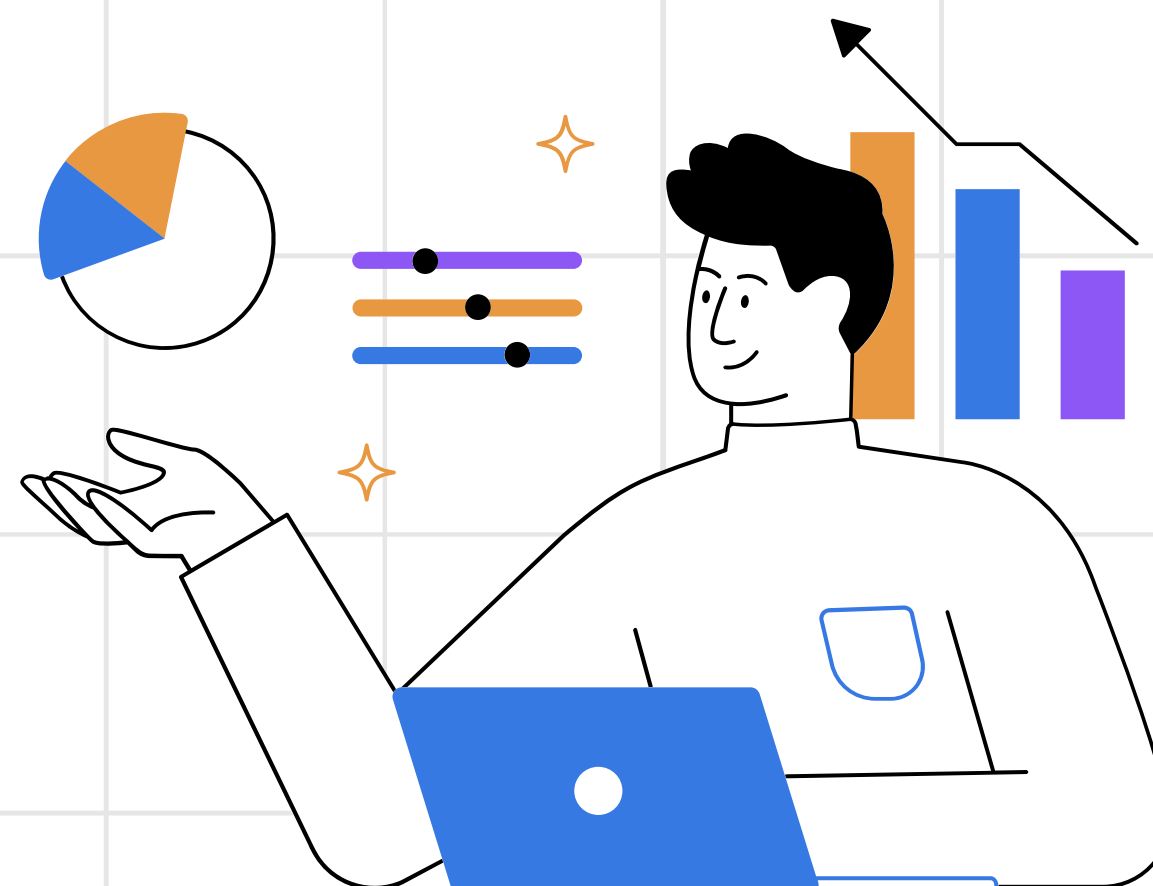
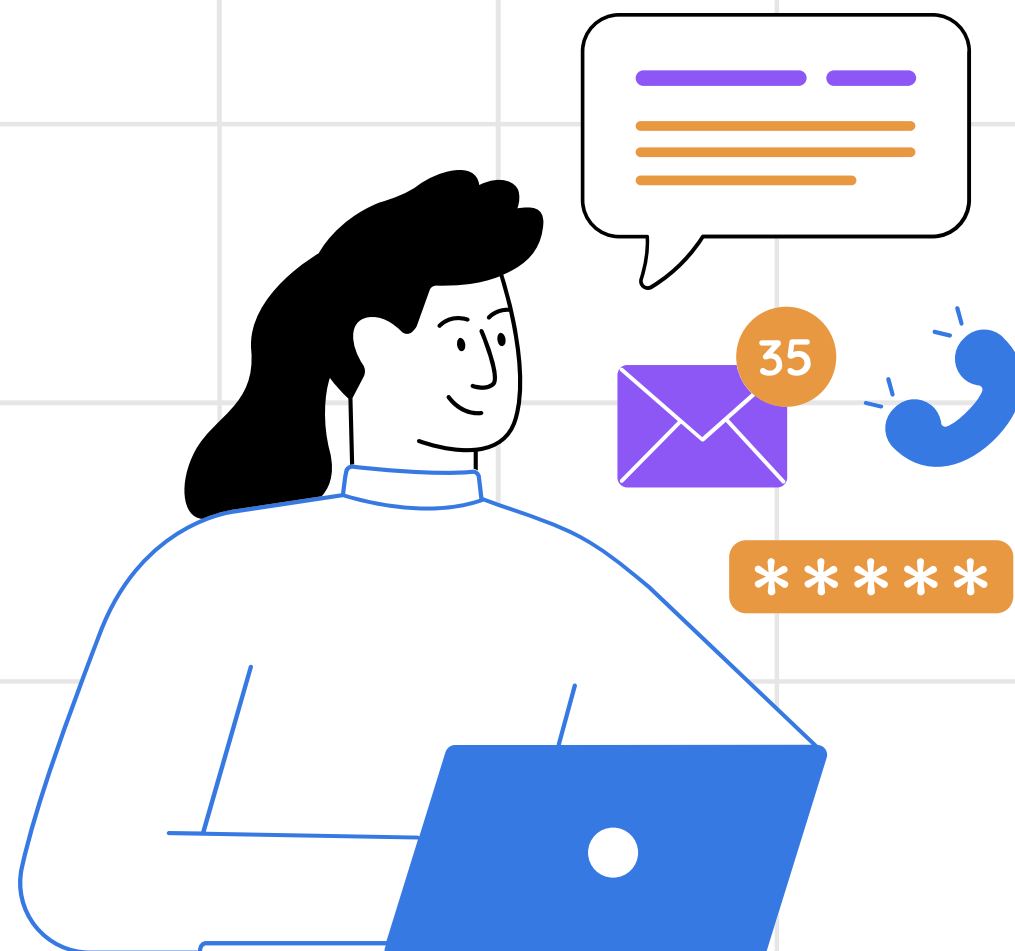
Summary of Methodologies

- Data Collection
- Quality Checks
- Exploratory Data Analysis
- Identifying Categorical Variables and creating Dummy Variables
- Model Building using Logistic Regression
- Predictions on Test Data Set
- Conclusion



Summary of Results

- Data Analysis through Visualizations
- Conclusion and Recommendations



02

INTRODUCTION

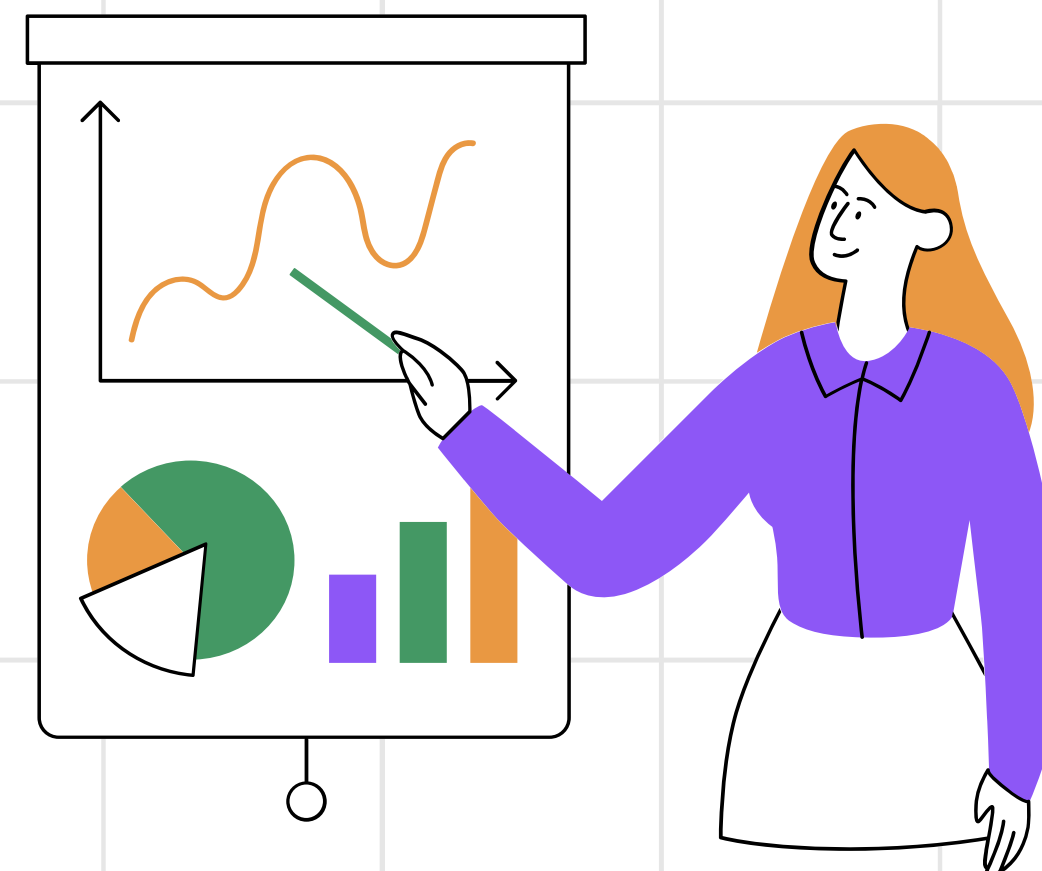
✓ Problem Statement

X Education, an online course provider for industry professionals, acquires leads through website visits, search engine marketing, and referrals. However, its lead conversion rate is only 30%, meaning most leads do not convert into paying customers. The sales team currently engages with all leads, making the process inefficient. To improve conversions, the company wants to implement a lead scoring system to identify high-potential leads (“Hot Leads”). By prioritizing these leads, the sales team can focus on the most promising prospects, increasing efficiency and overall conversion rates. A structured lead nurturing approach will further enhance engagement and drive better results.



✓ Objectives

- Exploring prospective lead
- Developing a Model for Lead Scoring
- Enhancing the rate of lead conversion
- Managing Categorical Variables effectively
- Ensuring future adaptability
- Providing actionable recommendations

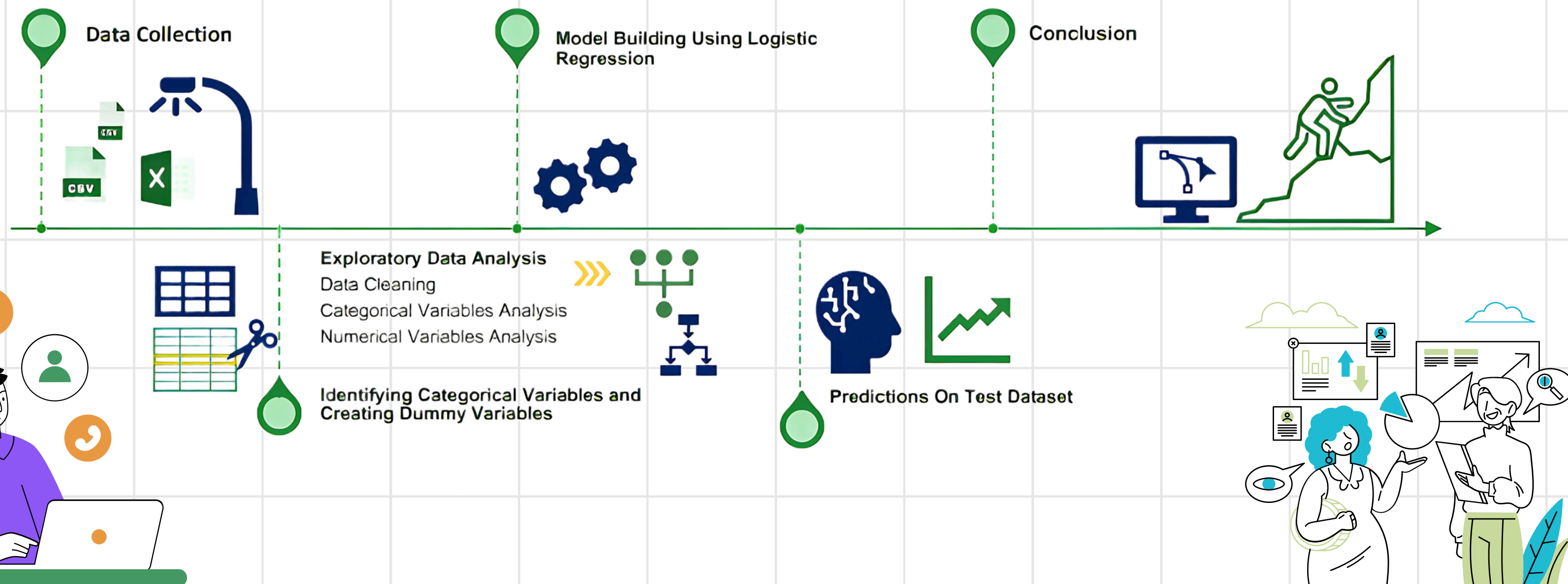


03

METHODOLOGY

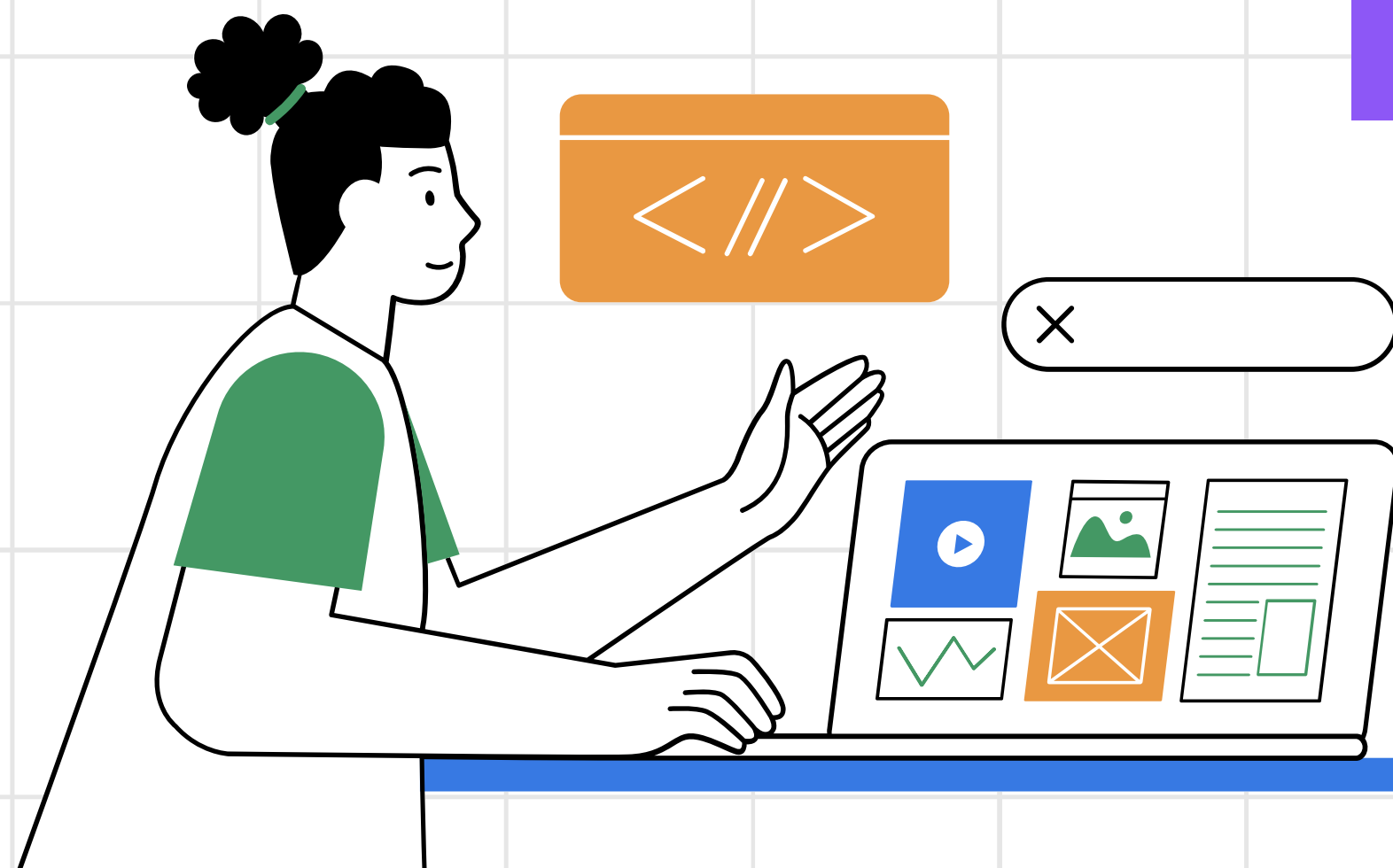
✓ Summary of Steps

The steps complete process involves gathering and measuring information on targeted variables in an established system, which then enables one to answer relevant questions and evaluate conclusions and outcomes

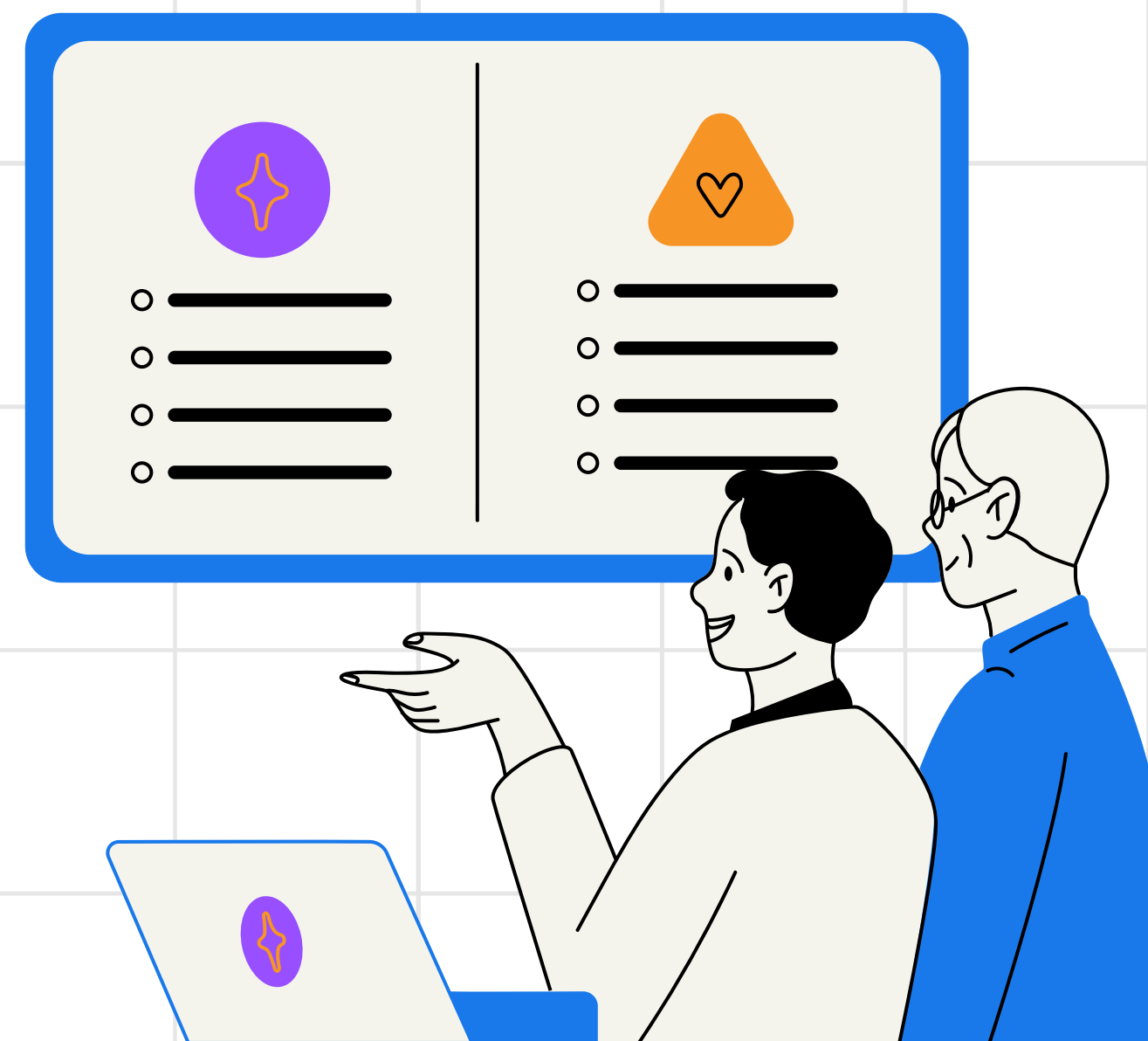


04

RESULTS

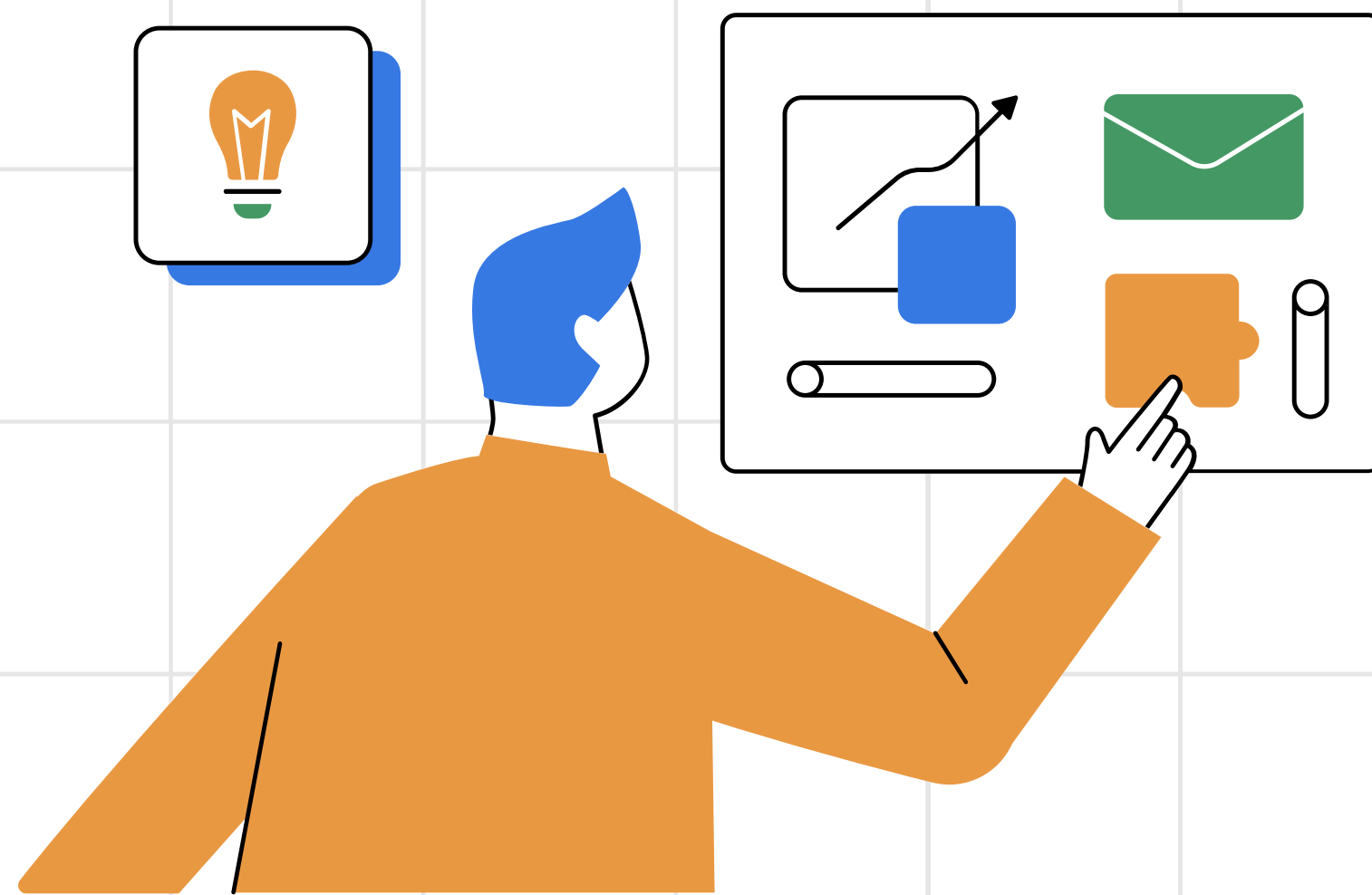


Consequence of EDA and
Visualizations



EDA with Visualizations

An analysis of significant graphs

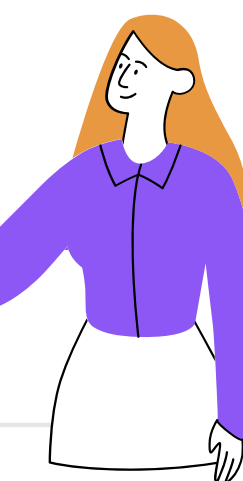
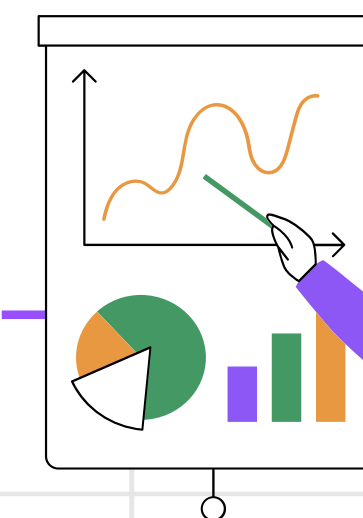
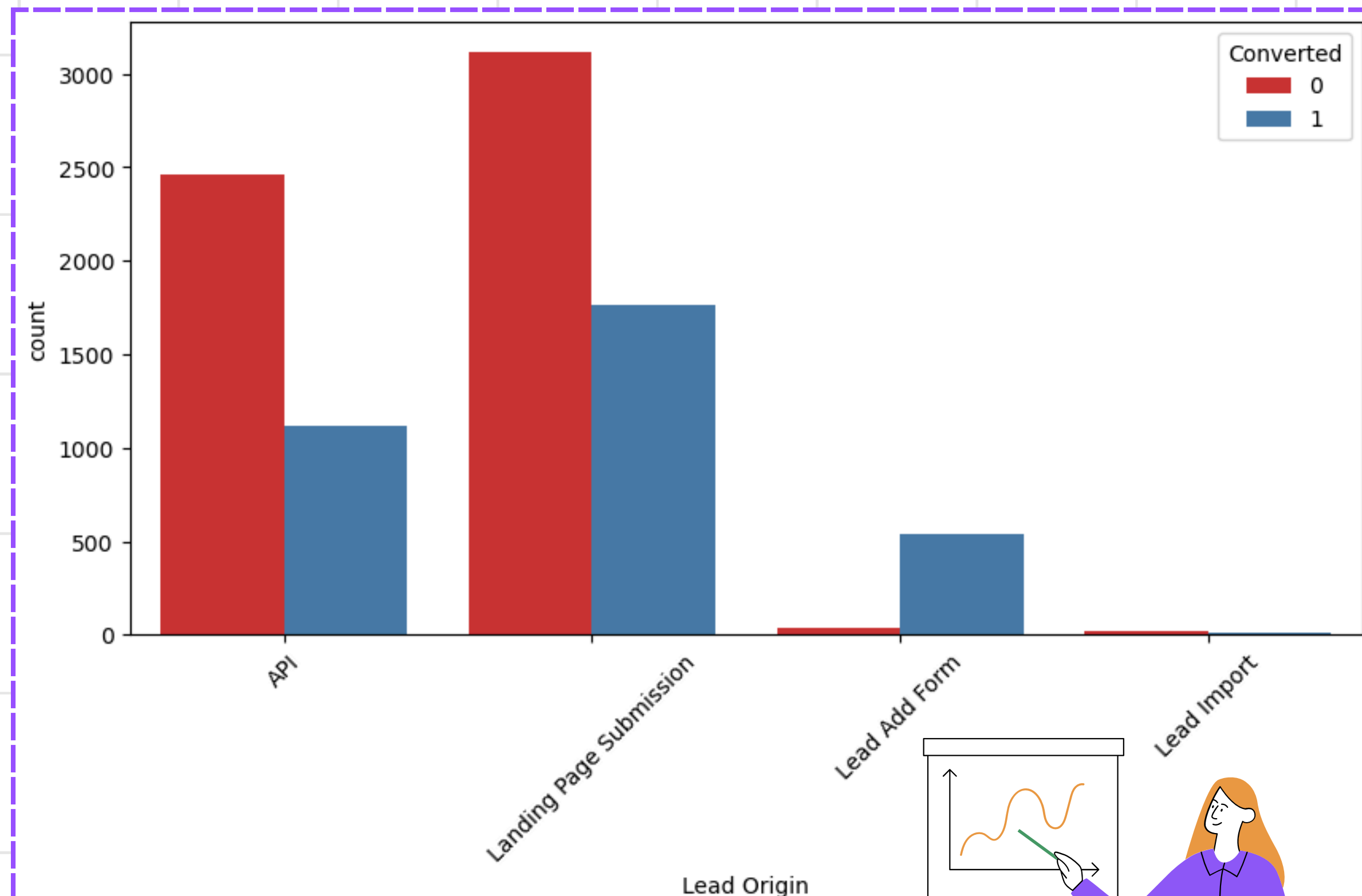


Lead Origin - Analysis



Insights

- ✓ **API & Landing Page Submissions** pull a solid 30-35% conversion, but the lead volume? Needs a boost.
- ✓ **Lead Add Form** is a powerhouse with a 90%+ conversion rate, but the inflow is low.
- ✓ **Lead Imports** are barely visible—very few in number.

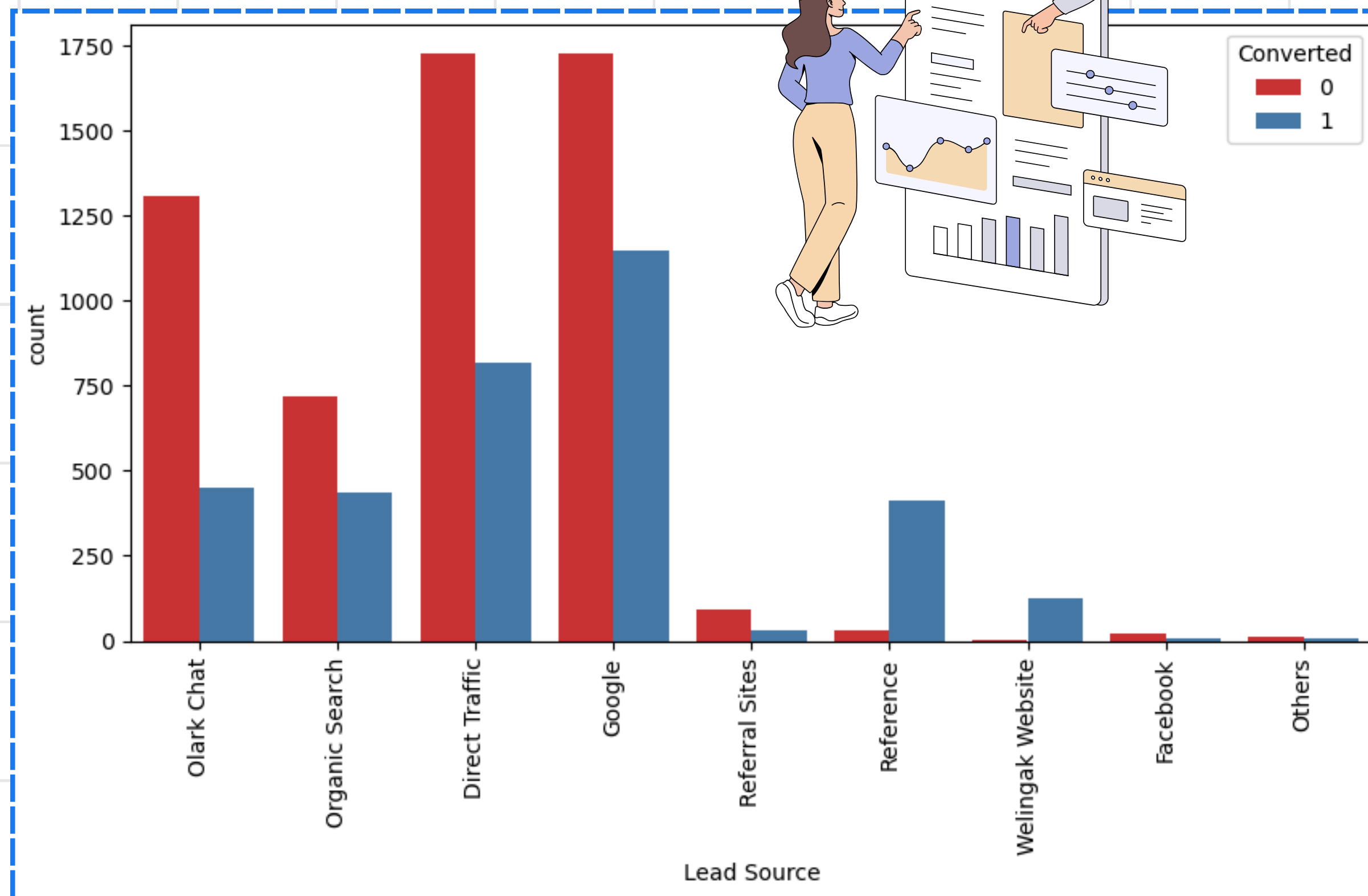


Lead Source - Analysis



Insights

- ✓ **Google & Direct traffic** are leading in volume.
- ✓ **Reference leads & Welingak website** boast high conversion rates.



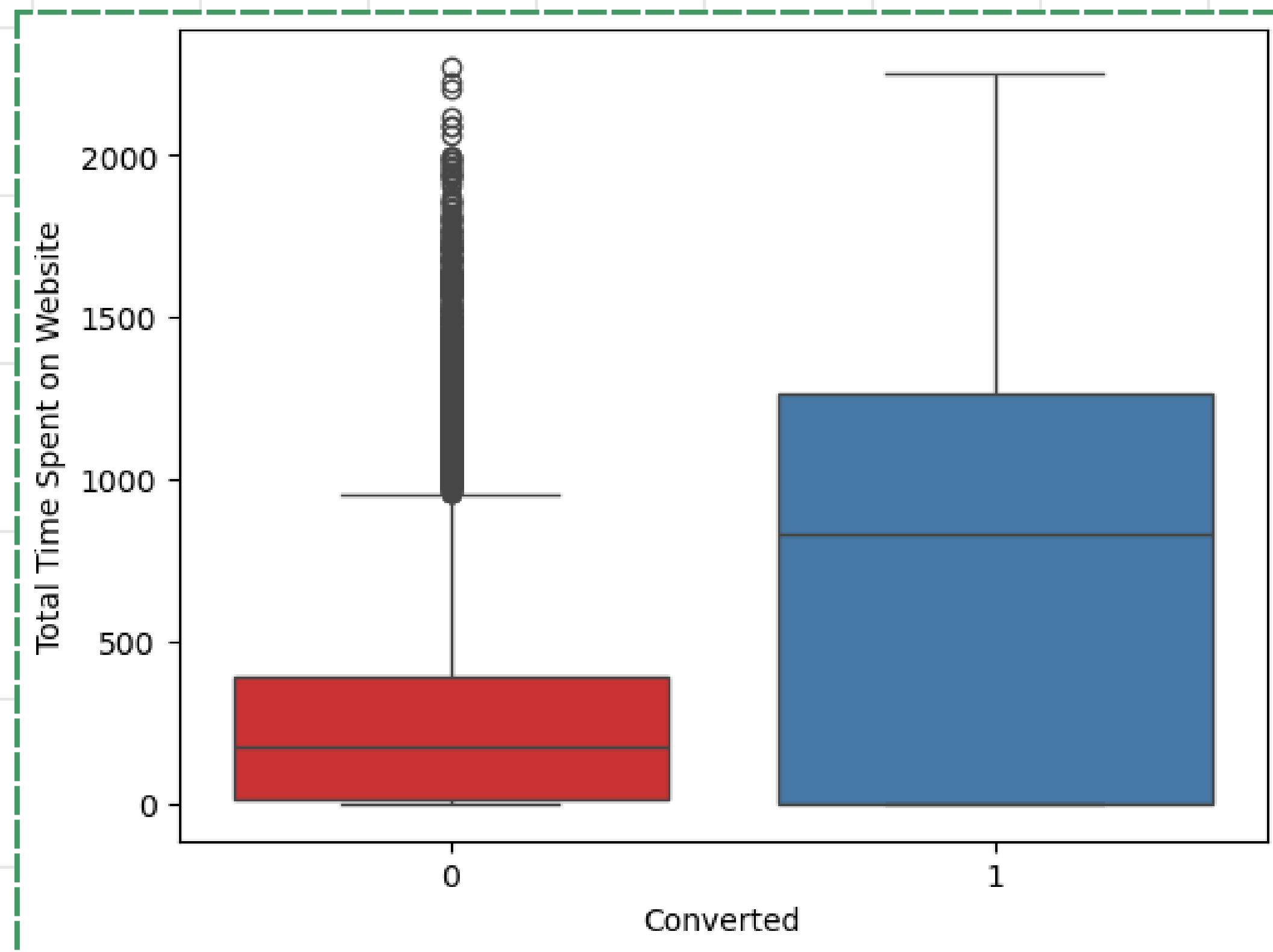
Total time Spent on Website - Analysis



Insights

Engagement & Conversion

Users who spend more time exploring the website exhibit a significantly higher conversion rate.

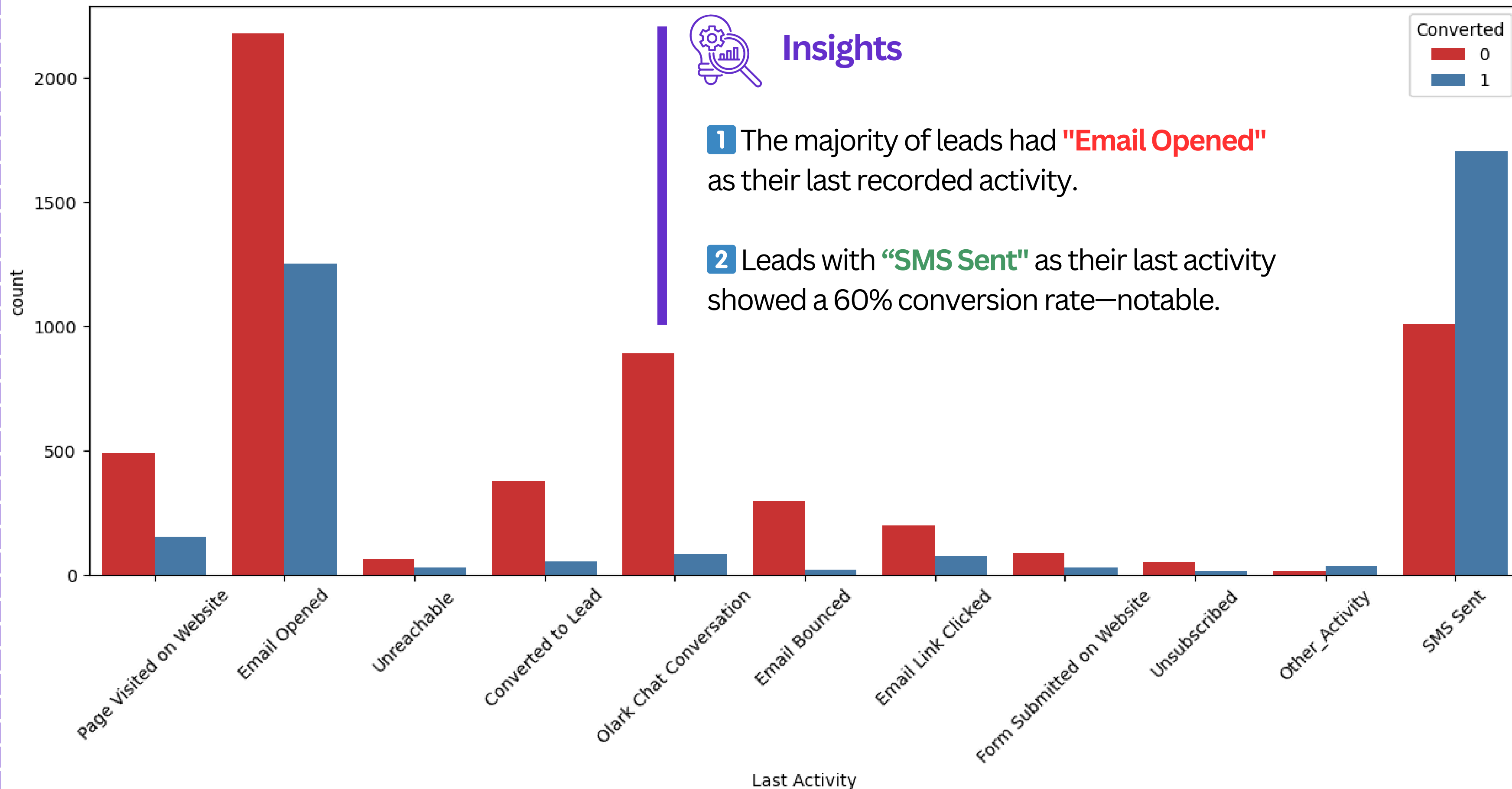


Last Activity - Analysis



Insights

- 1 The majority of leads had **"Email Opened"** as their last recorded activity.
- 2 Leads with **"SMS Sent"** as their last activity showed a 60% conversion rate—notable.



Specialization - Analysis



Insights



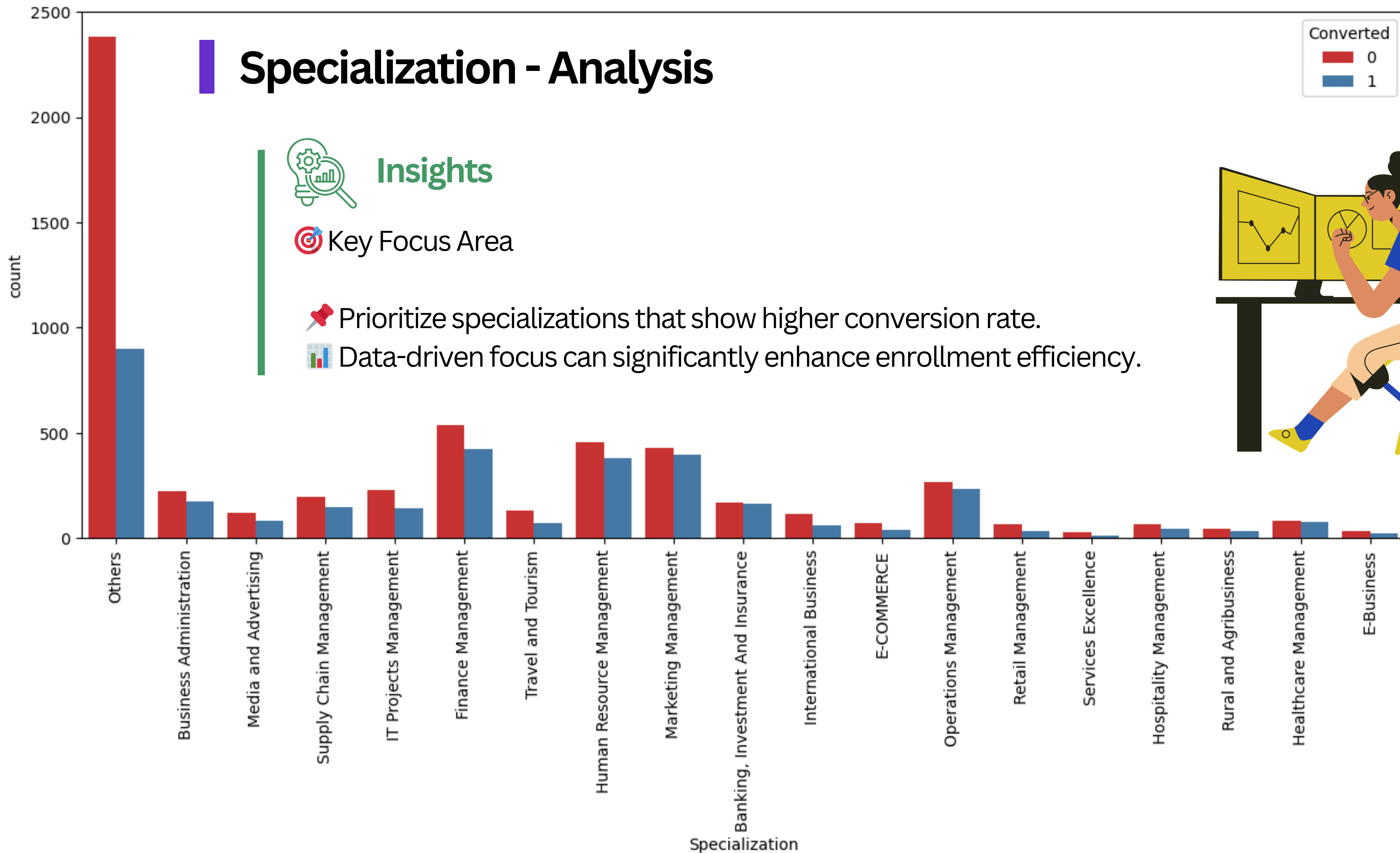
Key Focus Area



Prioritize specializations that show higher conversion rate.



Data-driven focus can significantly enhance enrollment efficiency.

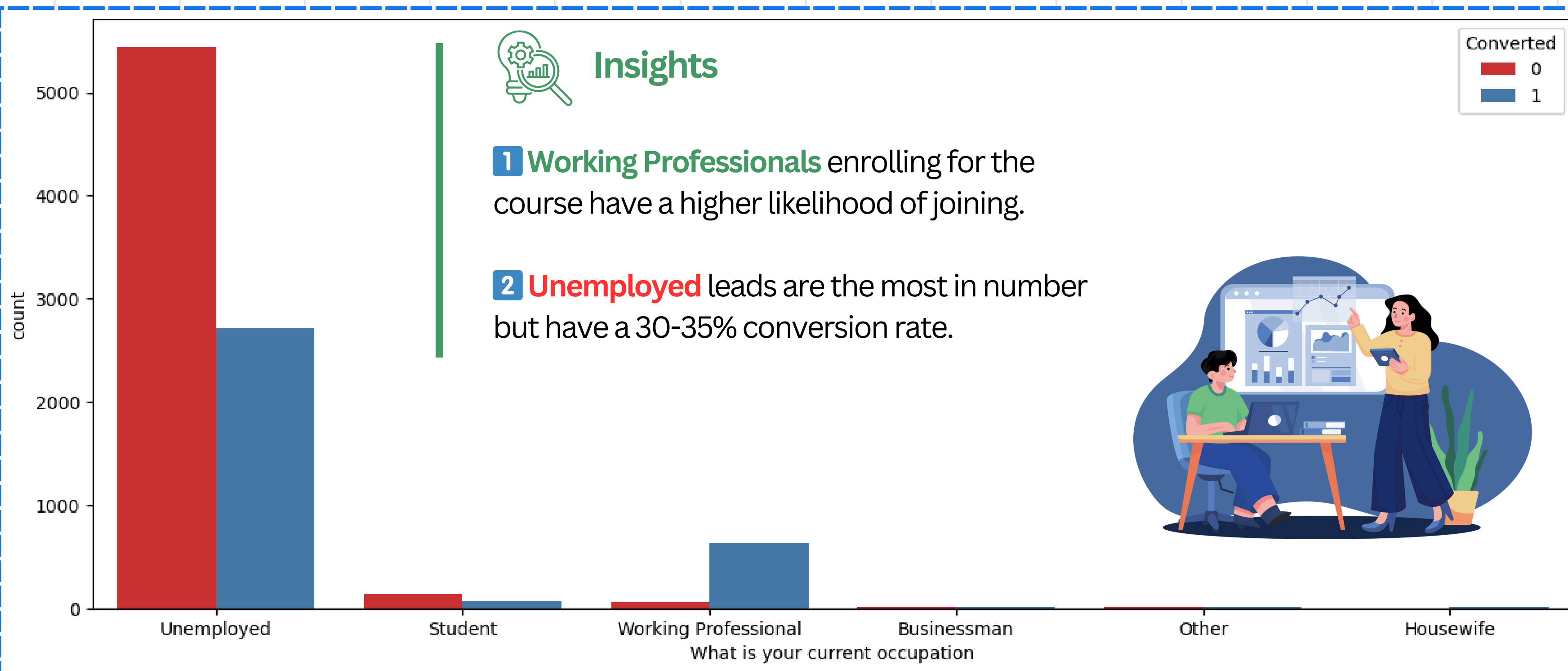


Current Occupation - Analysis



Insights

- 1 **Working Professionals** enrolling for the course have a higher likelihood of joining.
- 2 **Unemployed** leads are the most in number but have a 30-35% conversion rate.

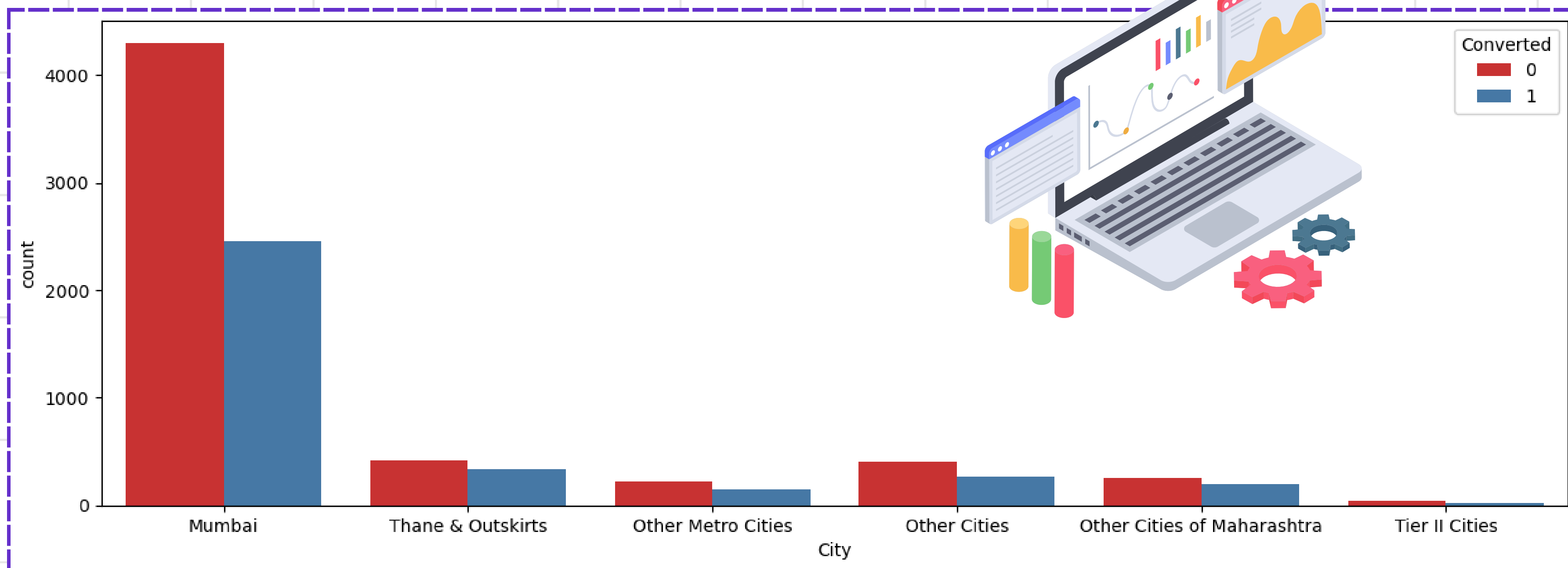


City - Analysis

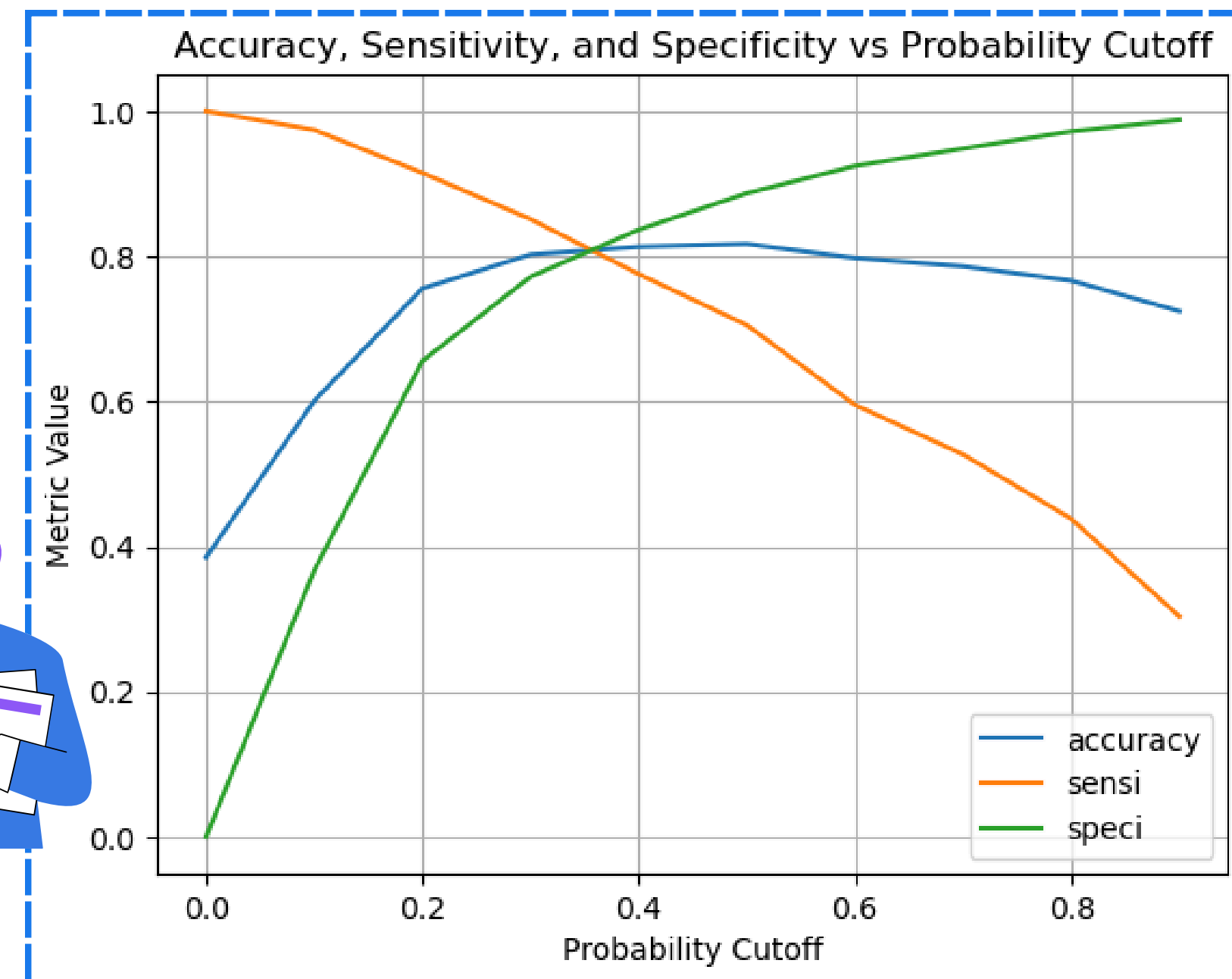
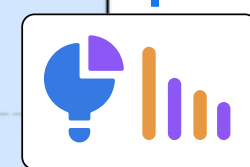
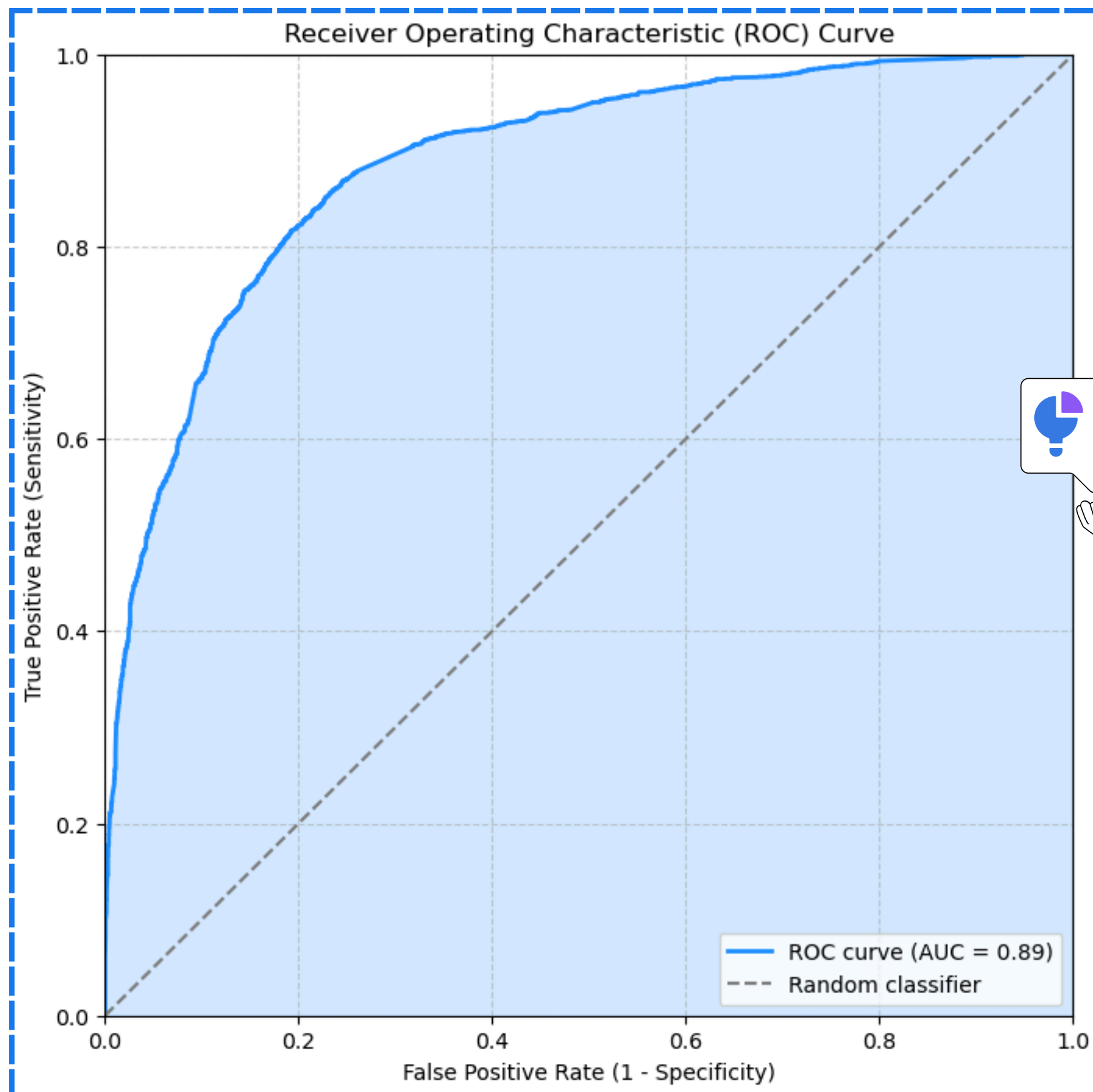


Insights

Most leads are from **Mumbai** with around a **50% conversion rate**.



ROC Curve & Optimal Cut-off Probability



Insights

Since we have a high **(0.89)** area under the **ROC** curve, **our model performs well.**

0.34 is identified as the optimum threshold for classification.

Model Evaluation



Insights

◆ Train Data:

✓ Accuracy: 80.9%

✓ Sensitivity: 82.2%

🔍 Specificity: 80%

◆ Test Data:

✓ Accuracy: 80%

✓ Sensitivity: 80.7%

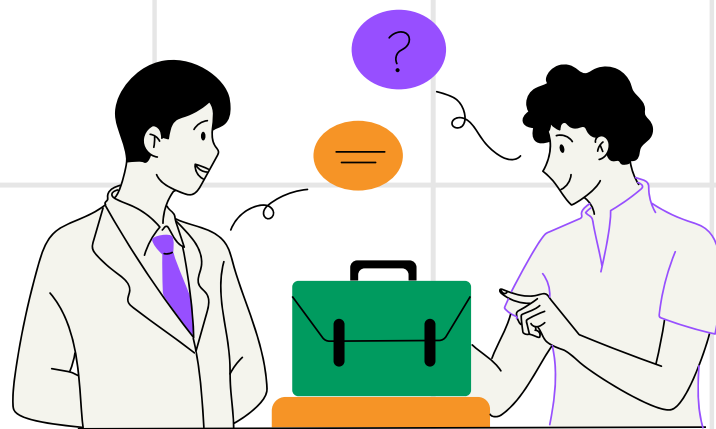
🔍 Specificity: 79.8%



Conclusion

The model shows strong performance in predicting lead conversion, with 80% test accuracy and consistent results with the training data (80.9%). This reliability offers a practical solution to X Education's 30% conversion rate.

By identifying 'Hot Leads', the sales team can prioritize outreach, saving time and resources while improving communication and lead nurturing. High sensitivity (80.7%) minimizes missed opportunities, and specificity (79.8%) filters out less promising leads. This balanced approach helps X Education boost conversions, increase revenue, and strengthen its market position. 🚀🌟



05

RECOMMENDATION



✦ Business Recommendations ✦

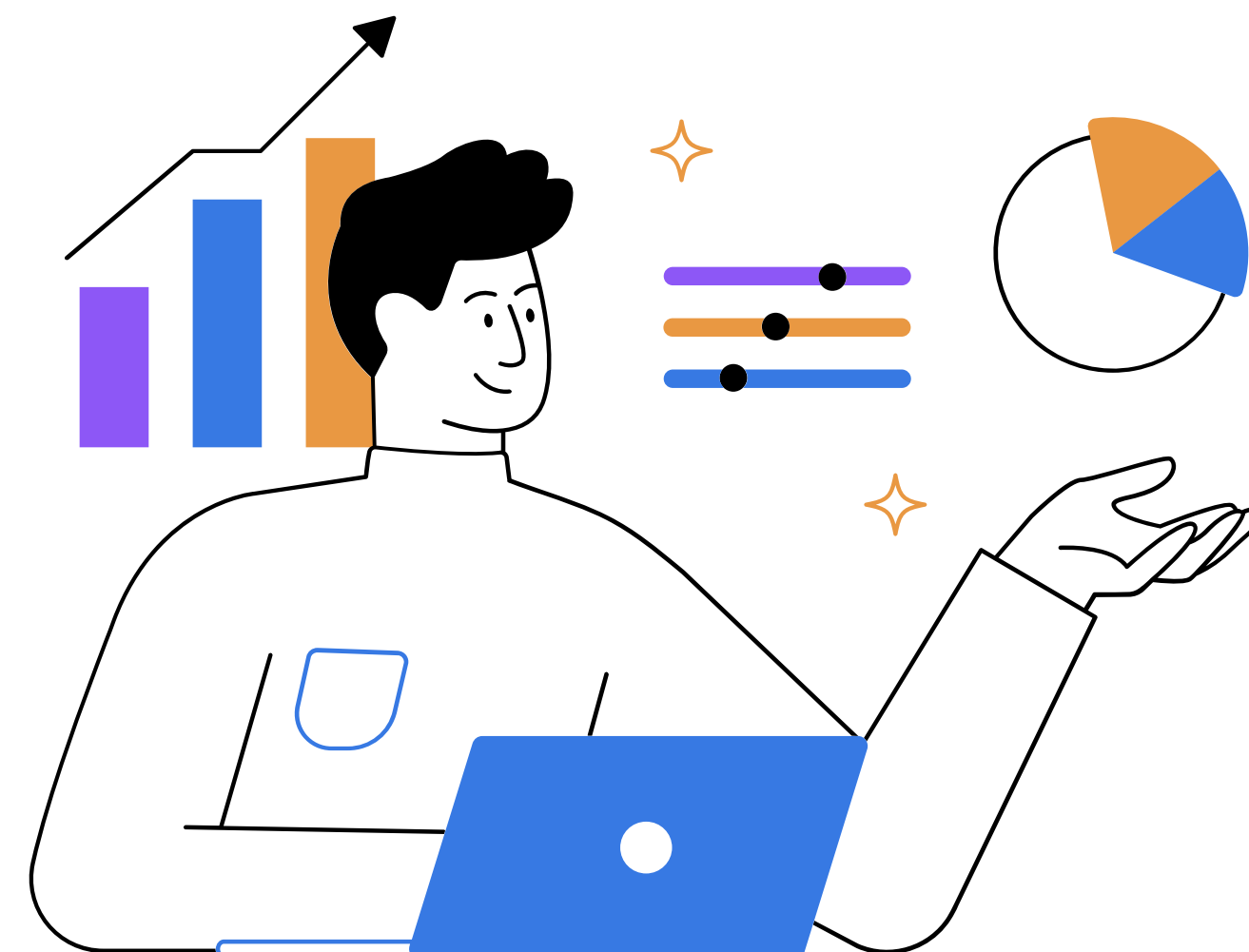
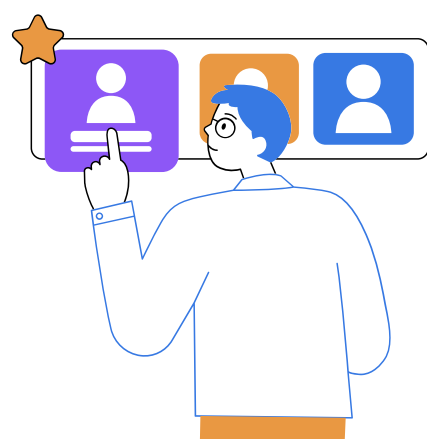
✓ Maximize Lead Conversions by reaching out to:

- Leads from "**Welingak Websites**" & "**Reference.**"
- Working professionals.
- Leads who spent more time on the website.
- Leads from "**Olark Chat.**"
- Leads whose last activity was "**SMS Sent.**"

⚠ Avoid wasting resources on:

- Leads whose last activity was "**Olark Chat Conversation.**"
- Leads from "**Landing Page Submissions.**"
- Leads whose specialization is "**Others.**"
- Leads who selected "**Do not Email**" as "**Yes.**"

🚀 Focus on high-potential leads = Higher conversions & Maximum ROI. ✨



Thank You

