

Lead Scoring

GROUP CASE STUDY

Arunan V
Ashwin Virkud



AGENDA

- Problem Statement
- Business Goal
- Methodology
- Exploratory Data Analysis
- Model Building
- Model Prediction
- Conclusion



PROBLEM STATEMENT

X Education offers online courses to professionals.

Despite high lead volume, conversion rates are low.



BUSINESS GOAL

The goal is to build a model which helps us to identify the potential Hot Leads and to convert them.

The CEO, in particular, has given a ballpark of the target lead conversion rate to be around 80%.



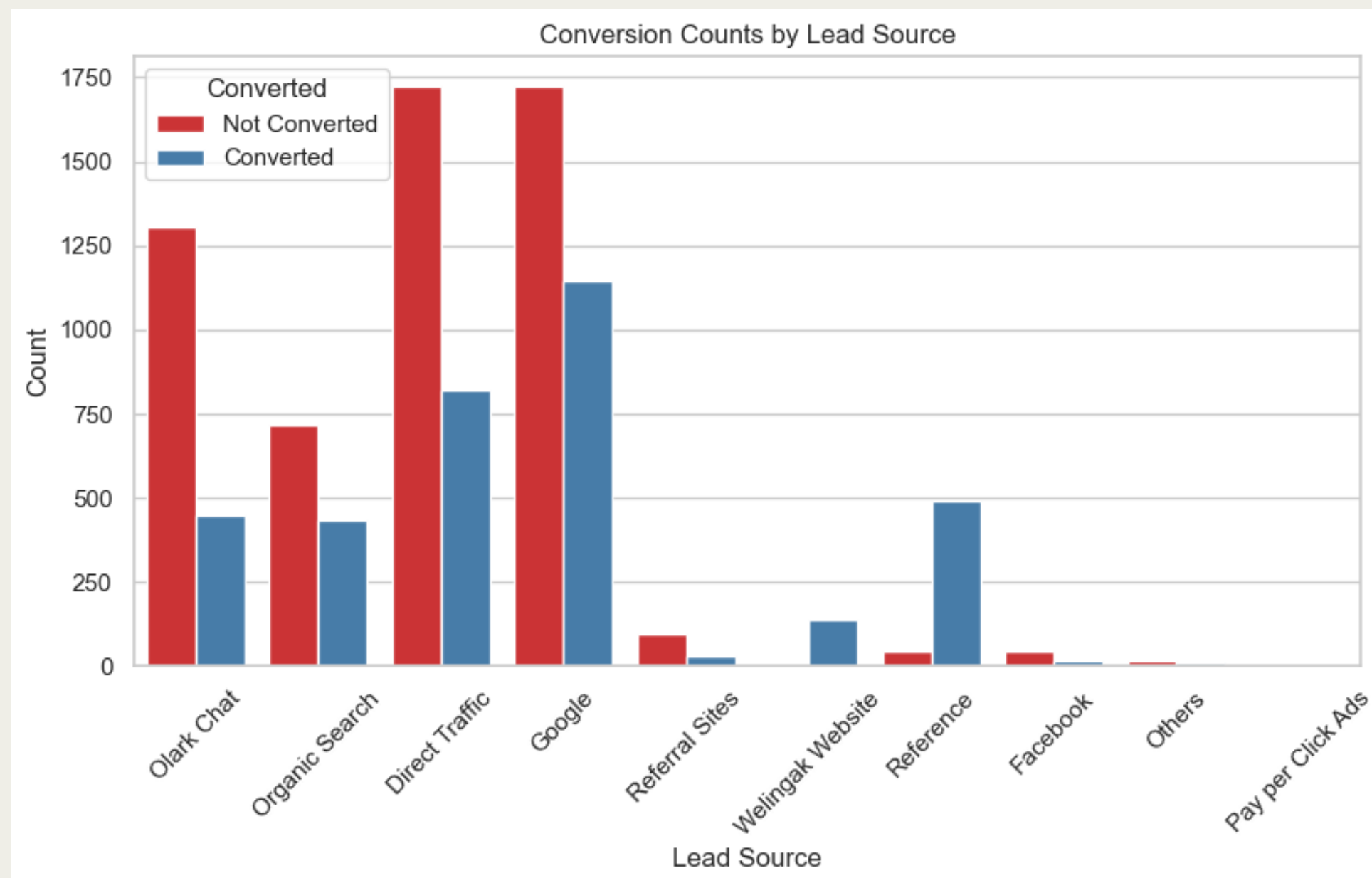
METHODOLOGY

●	●	●	●	●
<i>Data Cleaning</i>	<i>Exploratory Data Analysis</i>	<i>Creating Dummy Variable and Feature Scaling</i>	<i>Building the Logistic Regression Model</i>	<i>Evaluating the Model</i>
<i>Clean the data by taking care of the Null Values present in the Data and treating the outliers present in the Data .</i>	<i>Analysing the Data to identify different patterns in the data.</i>	<i>Splitting the Data into Train and Test Sets and Feature Scaling the Numeric Variables</i>	<i>Building the optimal Logistic Regression Model to calculate various metrics.</i>	<i>Evaluating the final model and finding out whether the percentage of conversion rate is greater than 80%</i>



EXPLORATORY DATA ANALYSIS

CONVERSION BY LEAD SOURCE

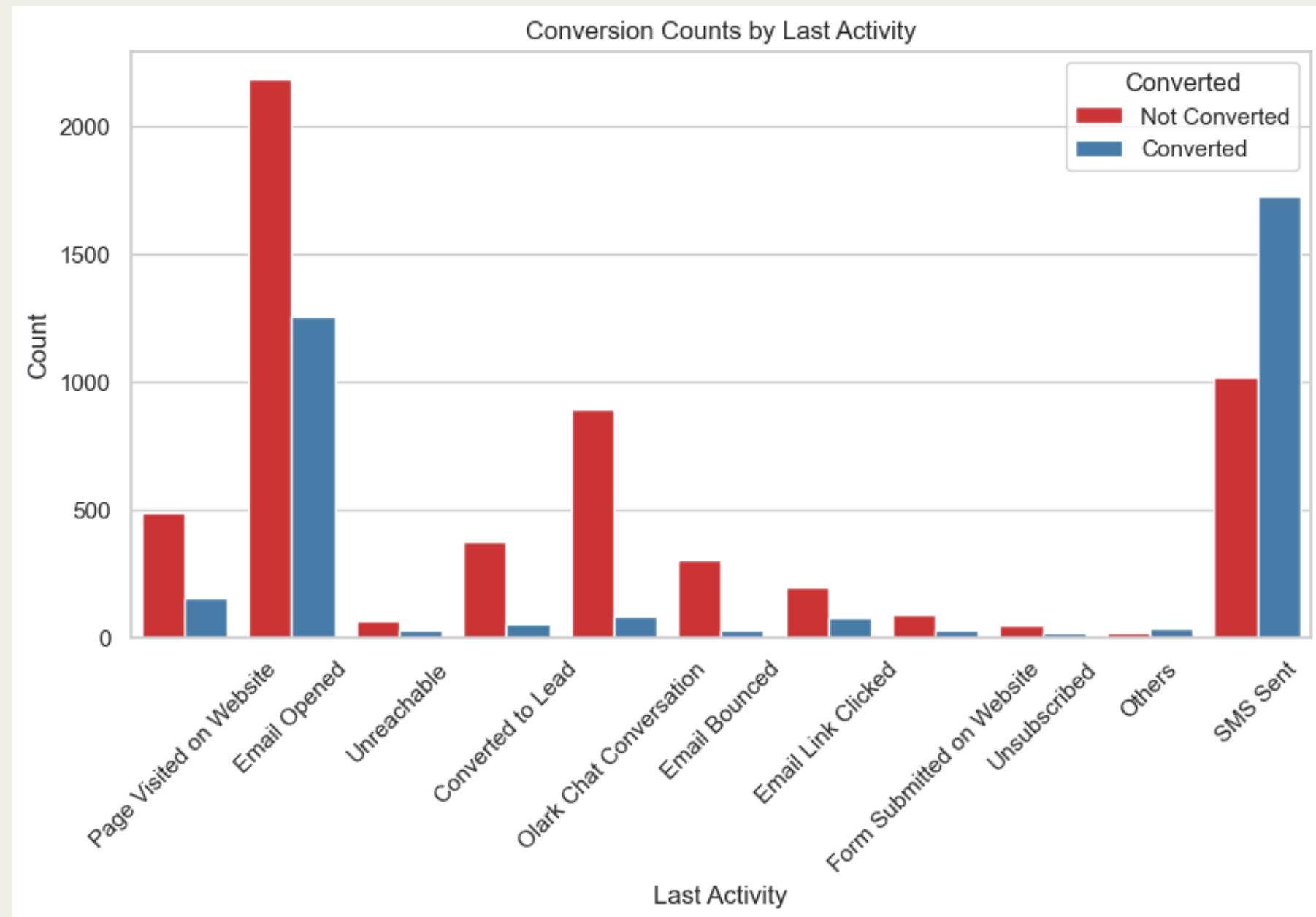


Observation:

- The leads from Google and Direct Traffic are the highest of the count.
- The highest conversion rate of the leads are from Reference and Welingak Website.
- Welingak Website has a conversion rate of 99%



LAST ACTIVITY CONVERSION COUNT

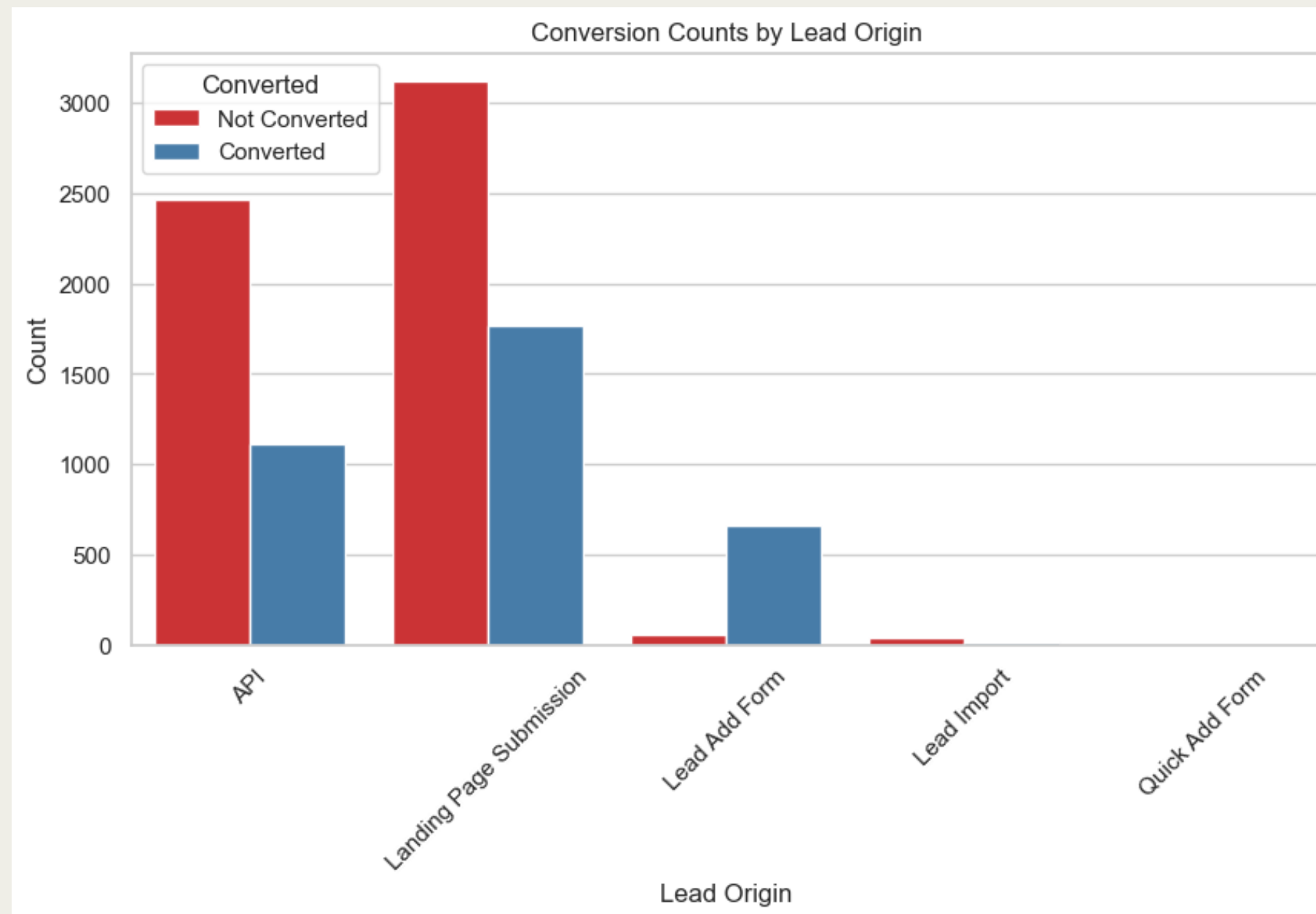


Observation:

- Email Opened has the maximum number of leads.
- SMS Sent has the highest conversion rate 63%



LEAD ORIGIN CONVERSION RATE

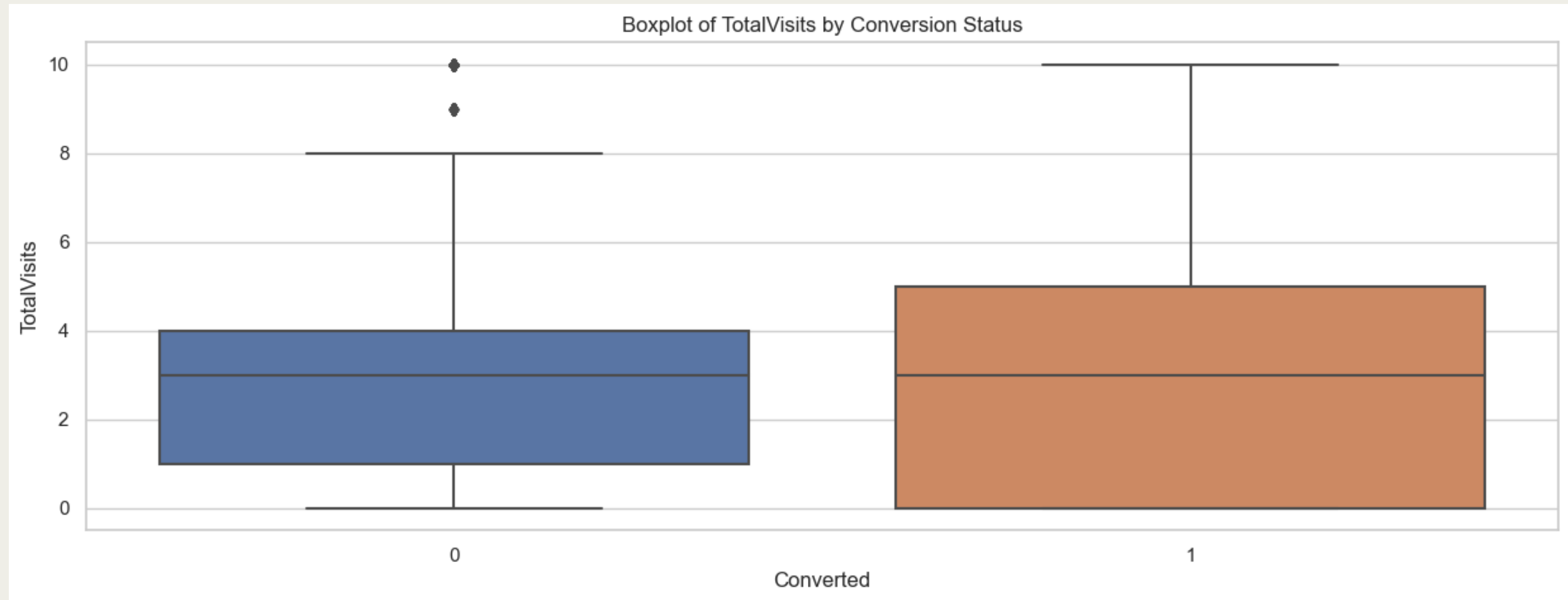


Observations:

- Lead Add Form has a conversion rate of 92%
- API and Lead Import have the least conversion rate.
- The count of leads from Lead Add Form is pretty low but the conversion rate is very high.



Time Spent on Website

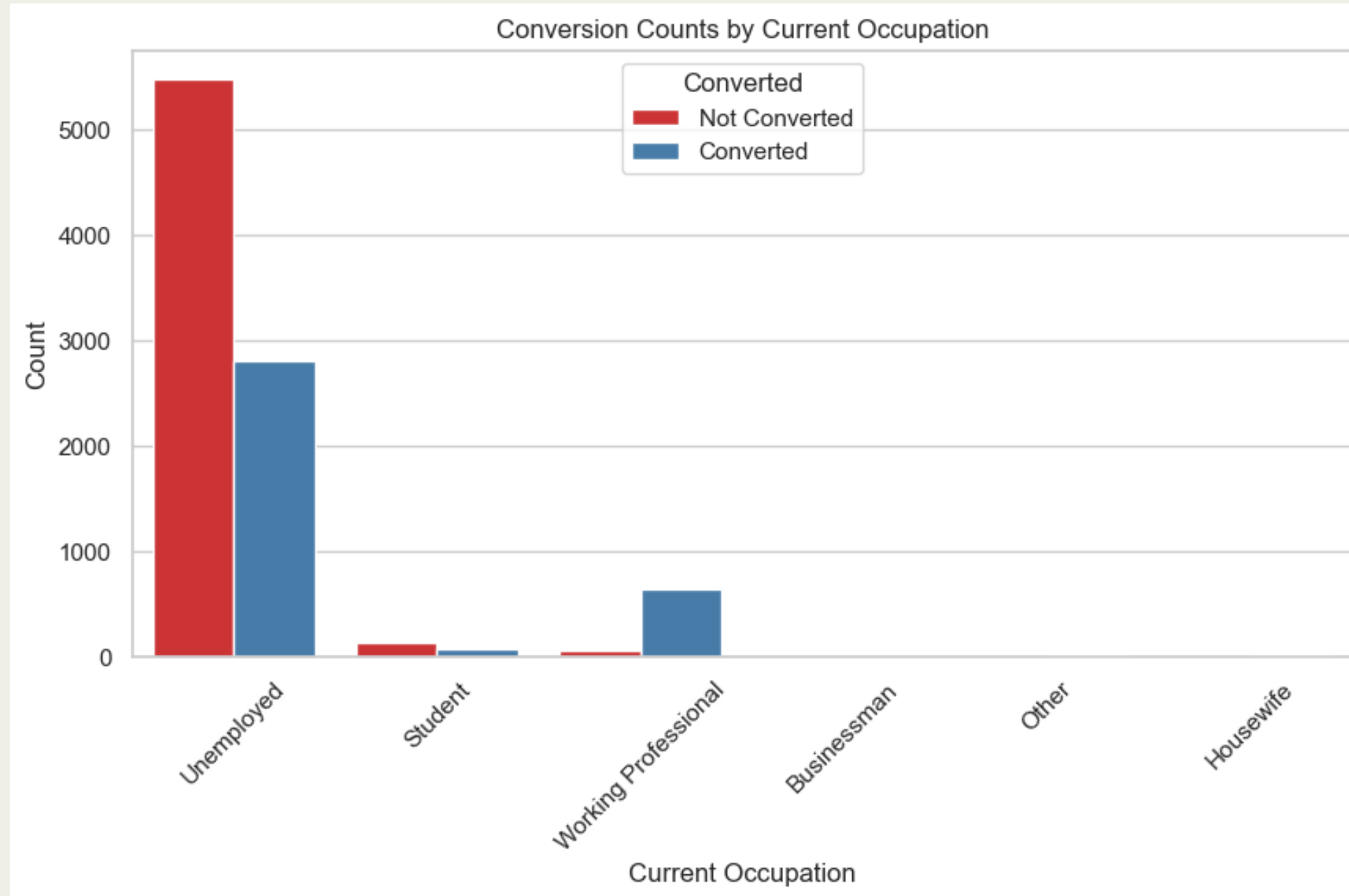


Observation:

- Leads spending more time are more likely to get converted.



CONVERSION RATE ON OCCUPATION



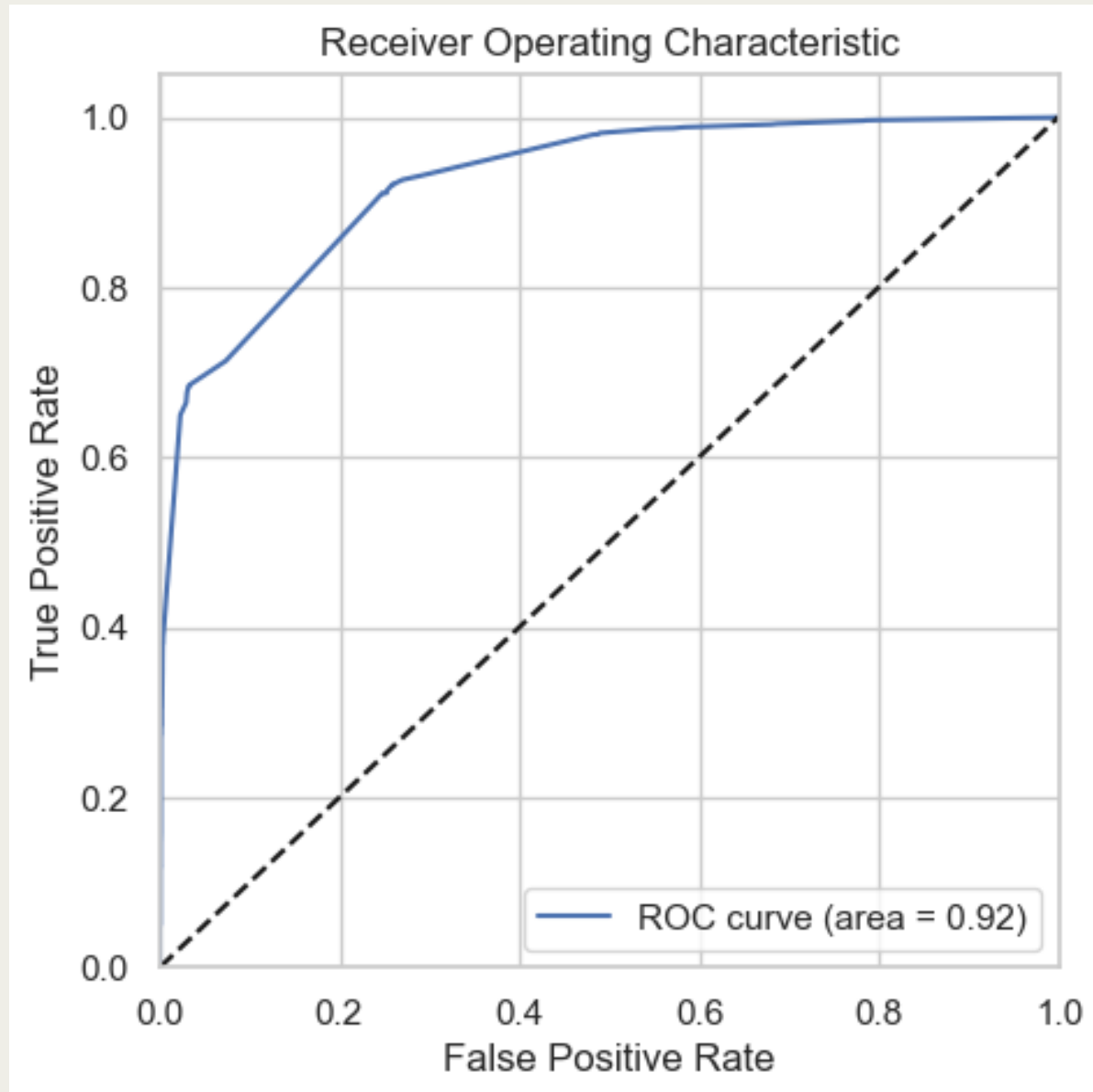
Observations:

- Working Professionals have a higher conversion rate of 92%
- Count of leads in the Unemployed category is higher but the conversion rate is only of 34%



MODEL BUILDING

ROC CURVE



- The data was split into train and test splits in a ratio of 70:30.
- 70 for train and 30 for test
- Using Recursive Feature Elimination (RFE) to select the top 15 features.
- The final model is built using these 15 features.



MODEL EVALUATION AND PREDICTION

ACCURACY	80.5%
SENSITIVITY	91.9%
SPECIFICITY	73.1%
PRECISION	69%

THE MODEL PREDICTED A 91.9% CONVERSION RATE



RECOMMENDATIONS AND CONCLUSION

- Leads spending more time on the website tend to get converted.
- Target the Working Professional leads as the conversion rate is very high.
- Concentrate more on leads who open the email as the chances of getting converted is higher.
- Increase the number of leads from Lead Add Form as they get converted easily.
- Leads from Welingak Website has a conversion rate of 99%



Thank you!

Arunan V
Ashwin Virkud

