

# **LEAD SCORING CASE STUDY USING LOGISTICS REGRESSION**

REKHA SARASWAT  
SIDDHI AGARWALA  
DINESH KUMAR

DS C53

# SCOPE

Problem Statement

Business Problem

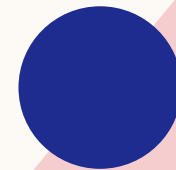
Problem Approach

Exploratory Data Analysis (EDA)

Model Evaluations

Observations

Strategy and Key Variables



## **PROBLEM STATEMENT**

An education company named X Education sells online courses to industry professionals. On a particular day, professionals who are interested in courses, visiting company's website and browse for courses. When these people fill up a form providing their email address or phone number, they are classified to be a lead. Moreover, the company also gets leads through past referrals. Once these leads are acquired, employees from the sales team start making calls, writing emails, etc.

# **BUSINESS PROBLEM**

- ✓ X Education seeking assistance in identifying profitable leads
- ✓ Company aims to develop a Lead Scoring System
- ✓ Lead to reflect likelihood of conversion
- ✓ CEO provided ballmark of the target lead conversion ~80%



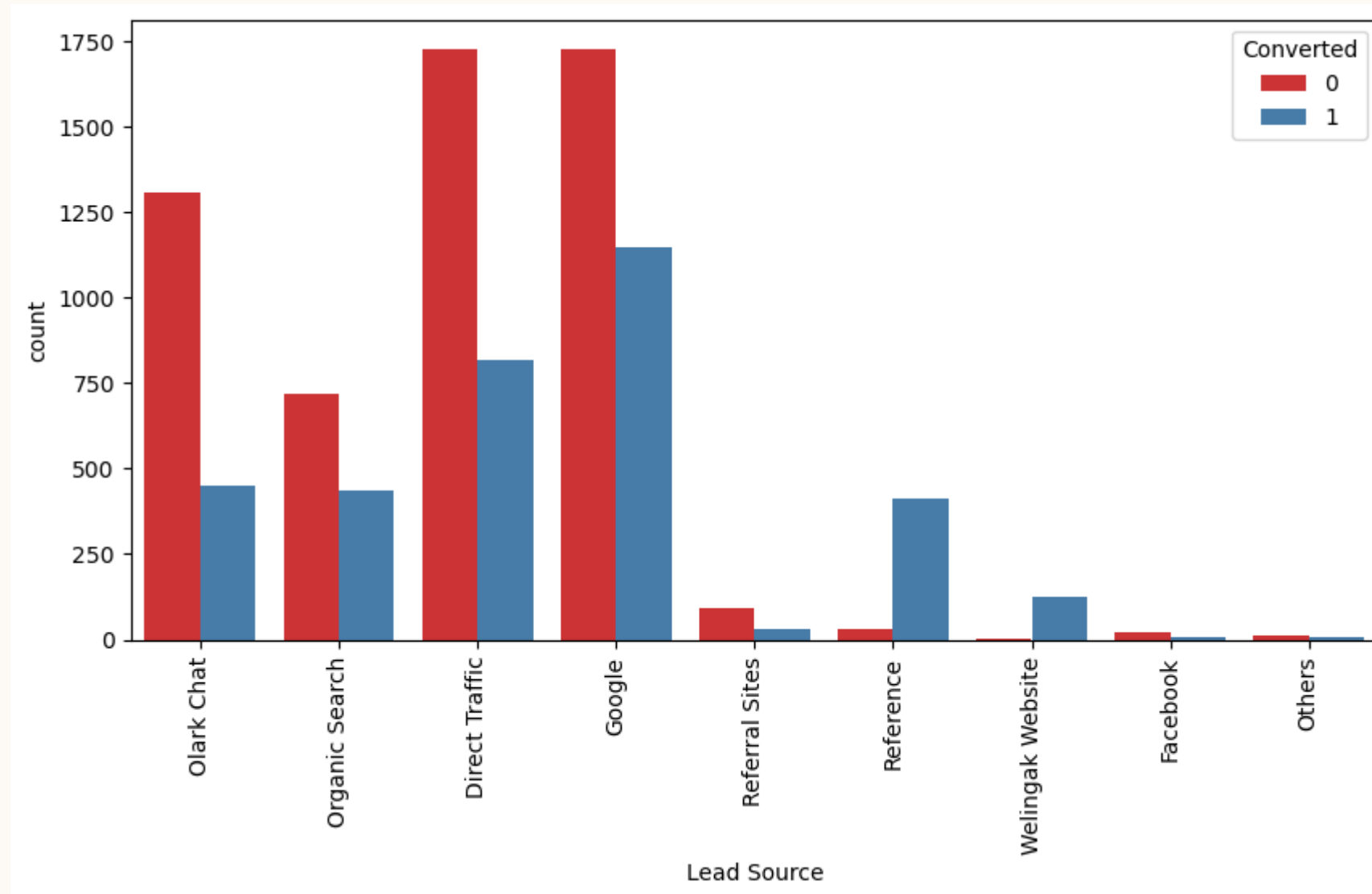
# **PROBLEM APPROACH**

Lead Scoring Data

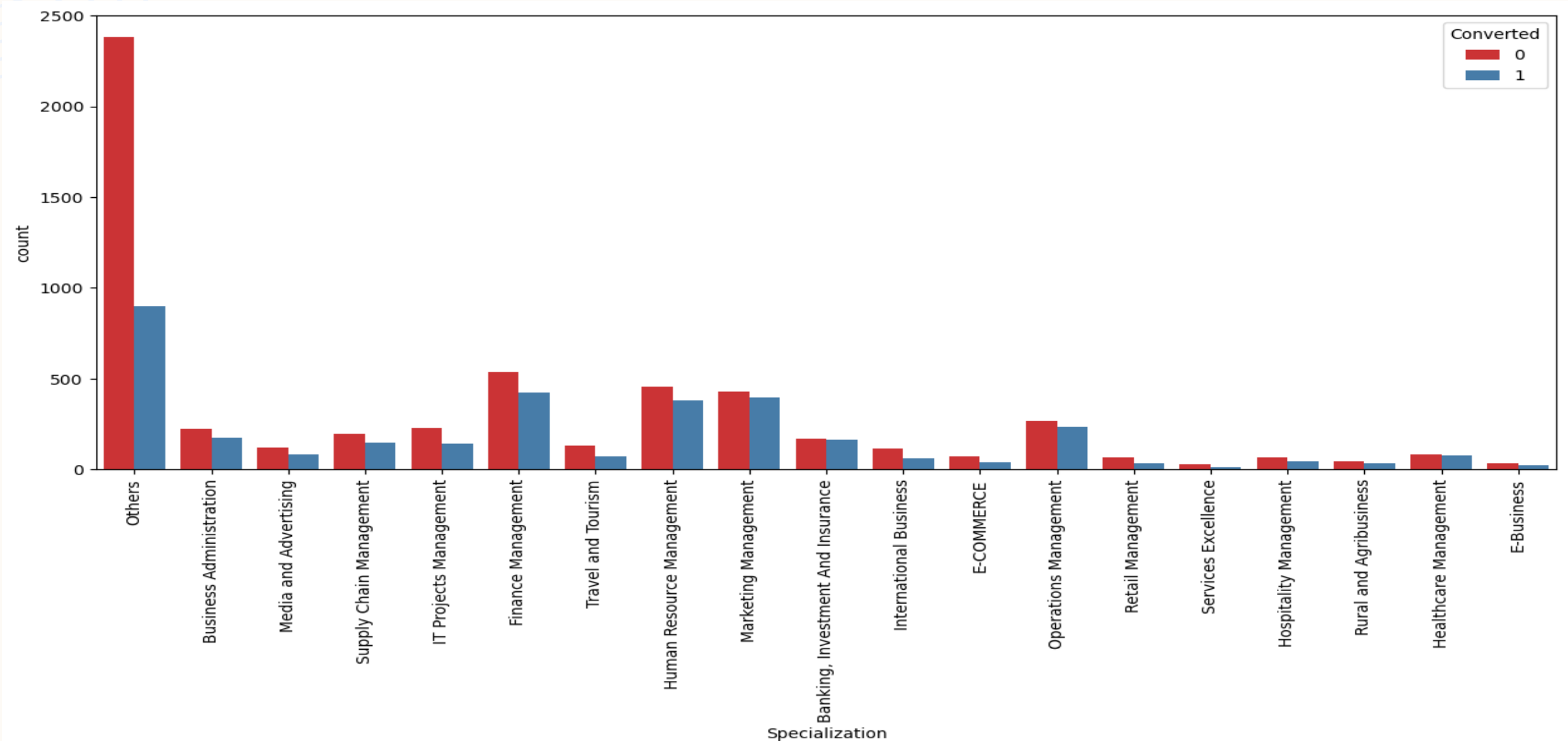
# PROBLEM APPROACH

- ✓ Importing Data
- ✓ Data Preparation
- ✓ Exploratory Analysis
- ✓ Dummy Variable Creation and Drop certain Variables
- ✓ Test-Train Split
- ✓ Scaling
- ✓ Model Building
- ✓ Model Evaluation
- ✓ Prediction on Test Set

# LEAD SOURCE

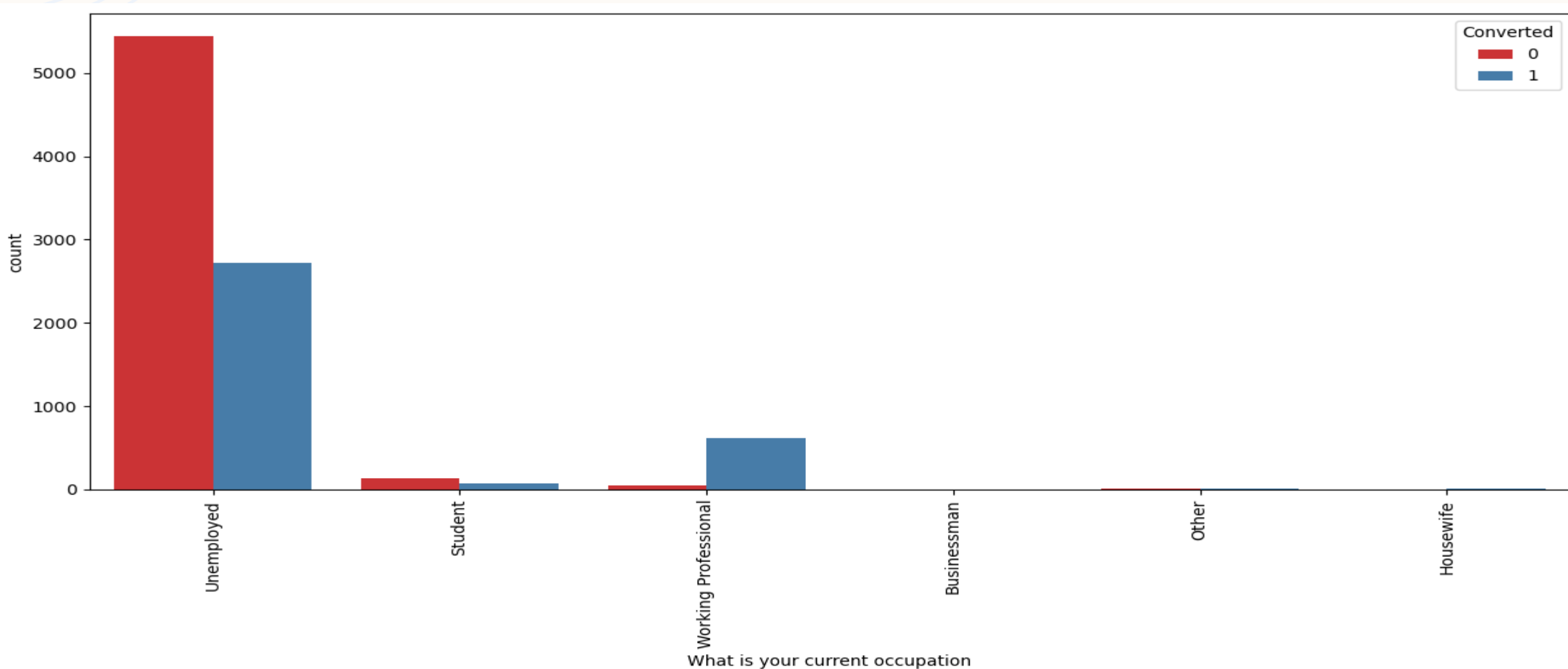


# LEAD - SPECIALIZATION

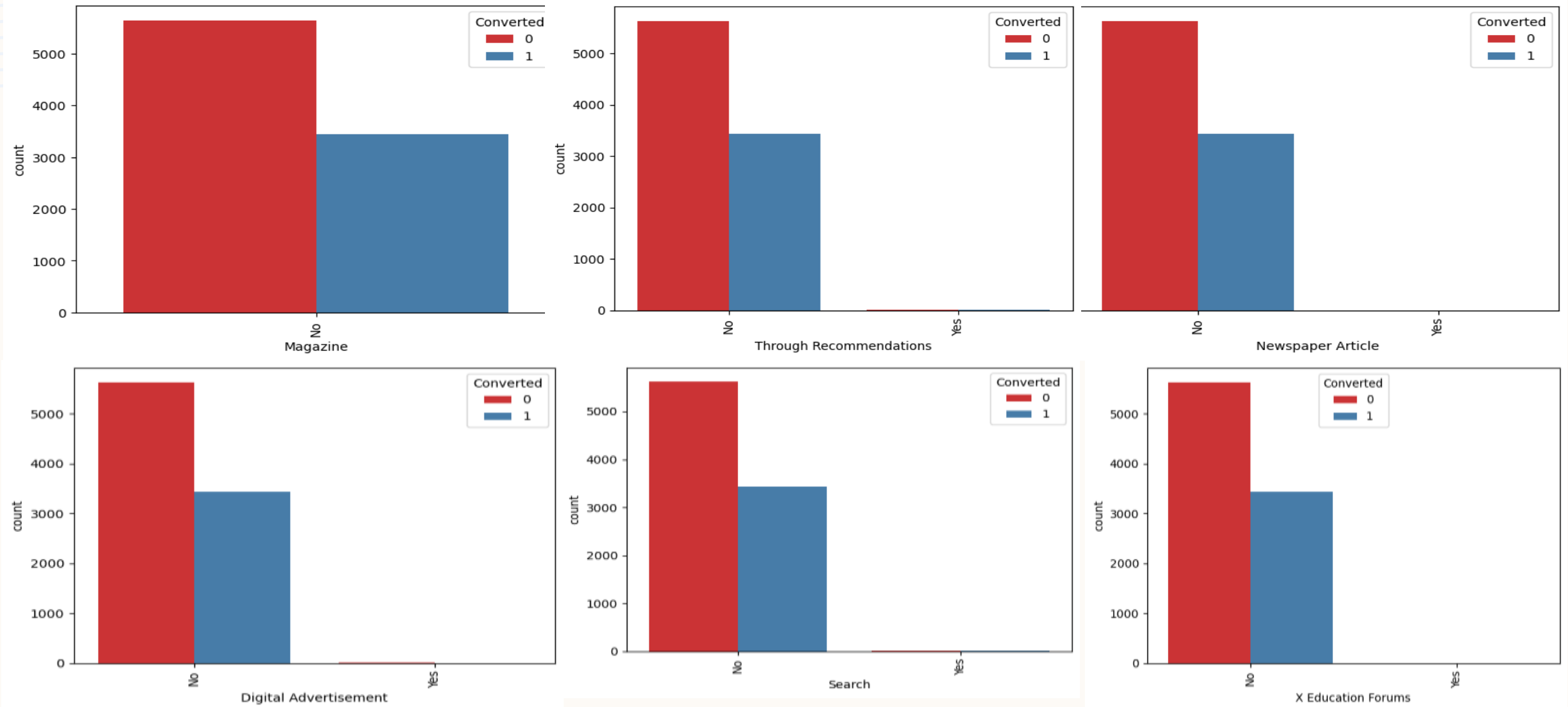




# LEAD – CURRENT OCCUPATION



# LEAD – SOURCE

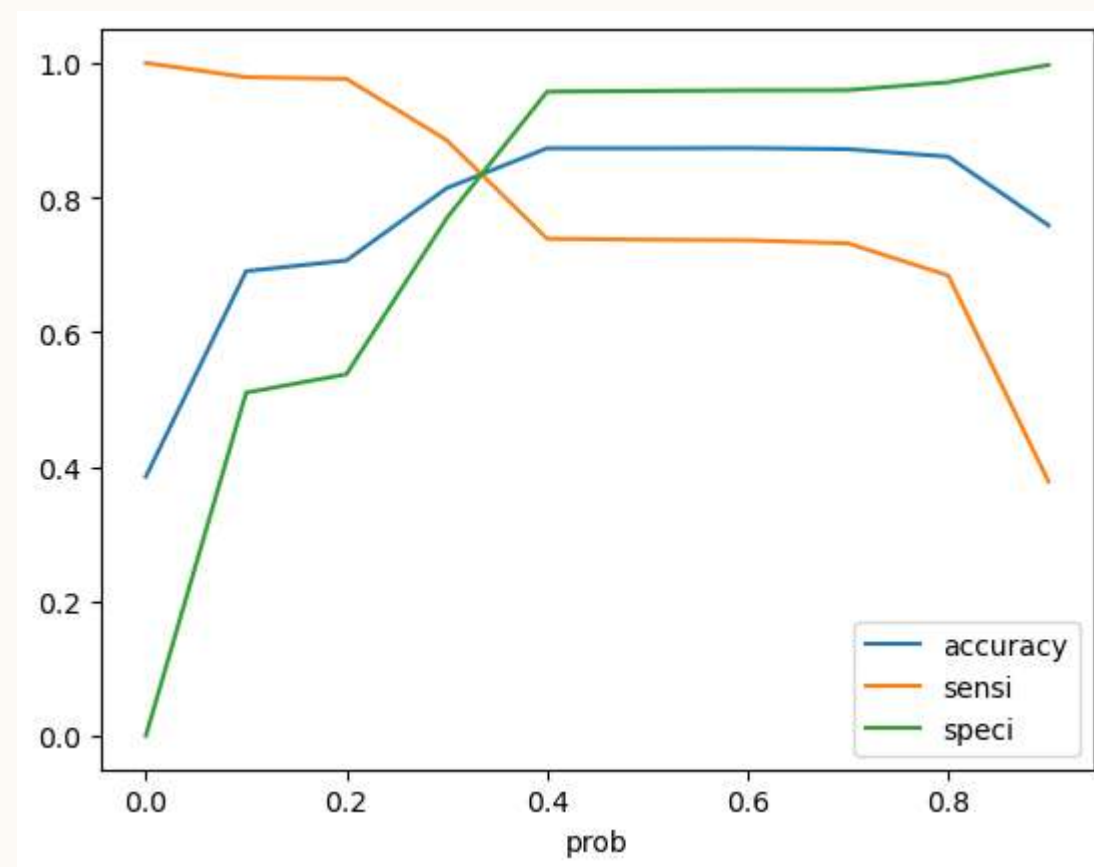
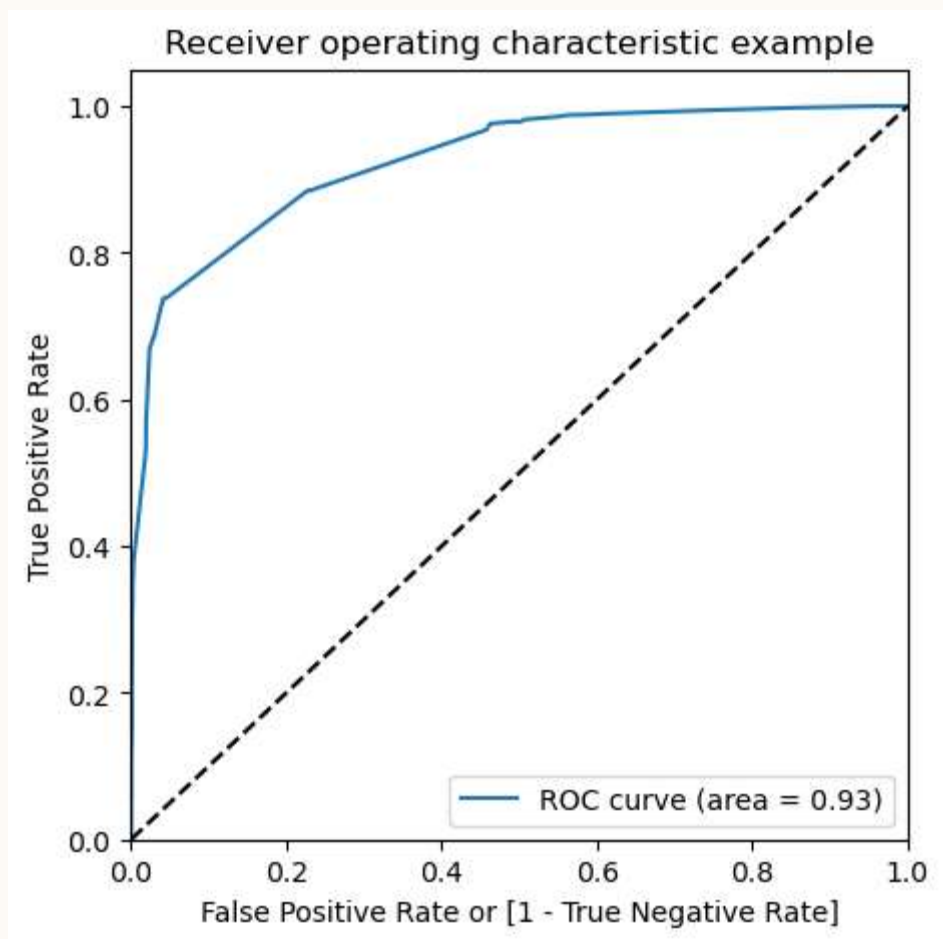


# VIF – AFTER MODEL BUILDING

Features		VIF
2	Lead Source_Welingak Website	1.34
6	Tags_Closed by Horizzon	1.23
5	Tags_Busy	1.11
7	Tags_Lost to EINS	1.07
10	Last Notable Activity_Had a Phone Conversation	1.01
1	Lead Origin_Lead Add Form	0.69
9	Tags_Will revert after reading the email	0.15
8	Tags_Ringing	0.11
0	Lead Origin_Landing Page Submission	0.10
3	current occupation_Unemployed	0.09
11	Last Notable Activity_SMS Sent	0.02
4	Specialization_Others	0.00

# MODEL EVALUATION

## ROC CURVE



# OBSERVATIONS


## TRAIN DATA SET

- Accuracy : 0.87
- Sensitivity : 0.73
- Specificity : 0.953

## TEST DATA

- Accuracy : 0.85
- Sensitivity : 0.70
- Specificity : 0.94

## FINAL FEATURE LIST



Lead Source_Welingak Website
Tags_Closed by Horizzon
Tags_Busy
Tags_Lost to EINS
Last Notable Activity_Had a Phone Conversation
Lead Origin_Lead Add Form
Tags_Will revert after reading the email
Tags_Ringing
Lead Origin_Landing Page Submission
current occupation_Unemployed
Last Notable Activity_SMS Sent
Specialization_Others

## **KEY VARIABLE INFLUENCING CONVERSIONS AND STRATEGY**

- ✓ Total time spend on the Website and total number of visits emerged as crucial factors
- ✓ Lead scores from Olark chat and Welingak Websites showed significant impact
- ✓ Last activity with SMS and Olark Chat conversation were found influential
- ✓ Engage potential buyers through personalized interactions and based on their online behaviour
- ✓ Leverage effective Last Activities such as SMS and Olark chat conversations



# **THANK YOU**

Rekha Saraswat

Siddhi Agarwala

Dinesh Kumar