

---

# CASE STUDY - LEADS SCORING

AYE NYEIN KHAING



---

# CONTENTS

Introduction

About Datasets

Data Manipulation

Exploratory Data Analysis (EDA)

Build a Model

Conclusion



---

## DATA MANIPULATION

---

## Dataset Observations

- Number of variables: 37
  - Number of observations: 9240
  - Missing cells: 41039
  - Missing cells (%): 41039
  - Two type of variables Numeric 30 and Categorical 7
  - According to Problem Statement, need to handle 'Select' level in many of the categorical variables because it is as good as a 'null' value. So, those values are needed to convert 'null' values.
-



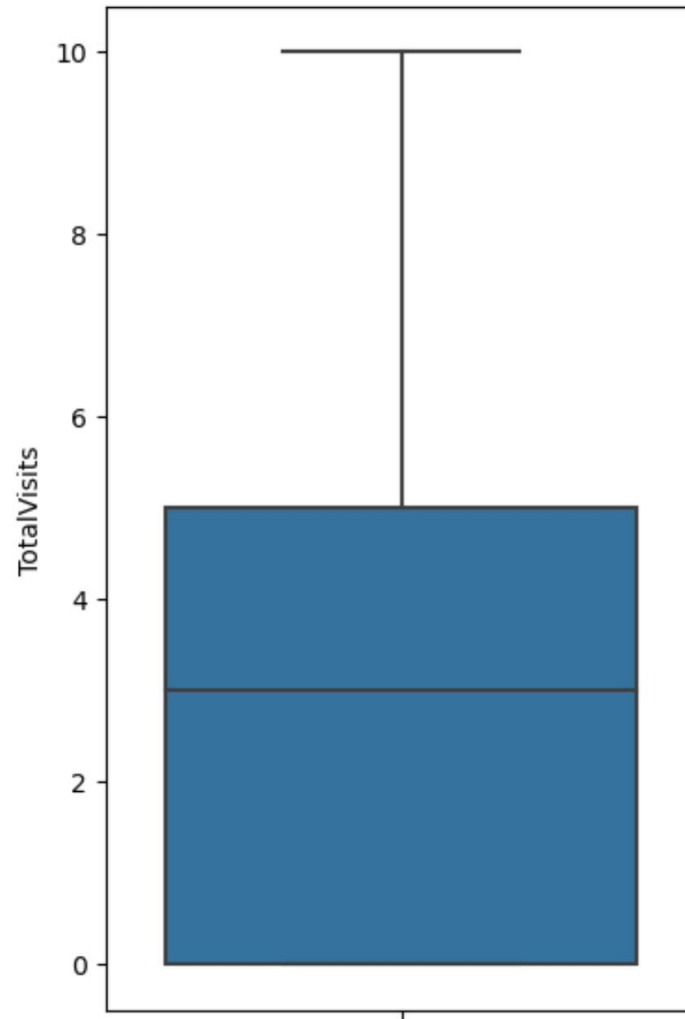
## EXPLORATORY DATA ANALYSIS (EDA)

---

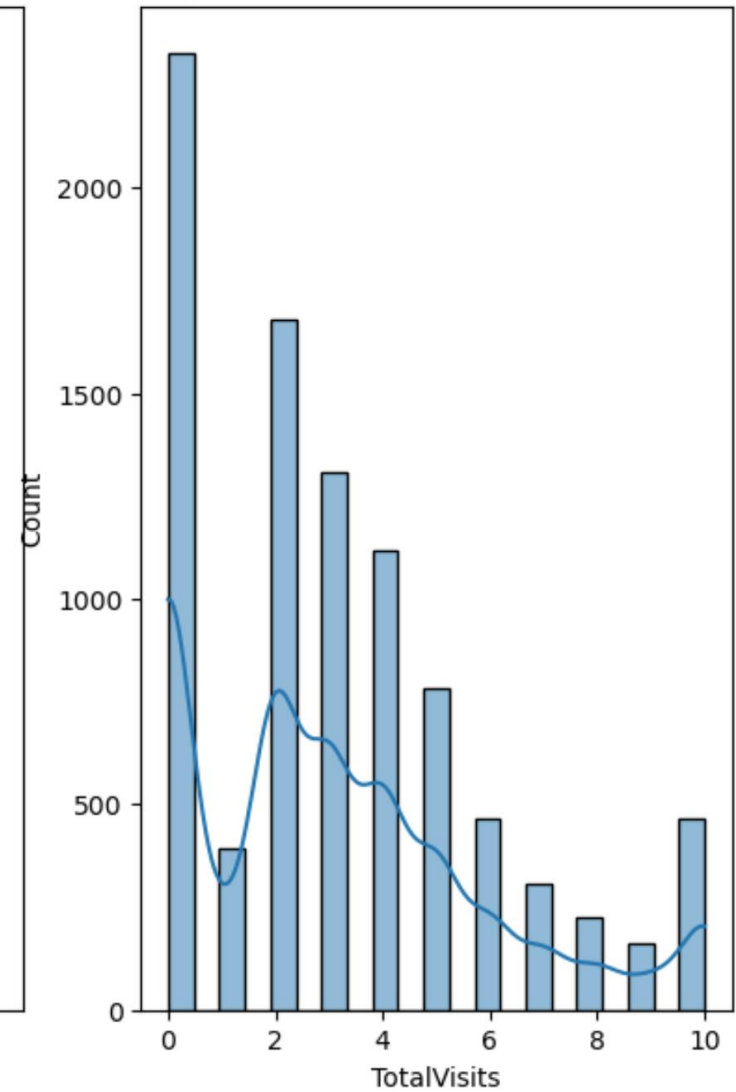
---

# TOTALVISITS VARIABLE DISTRIBUTION

Boxplot of TotalVisits Variable Distribution

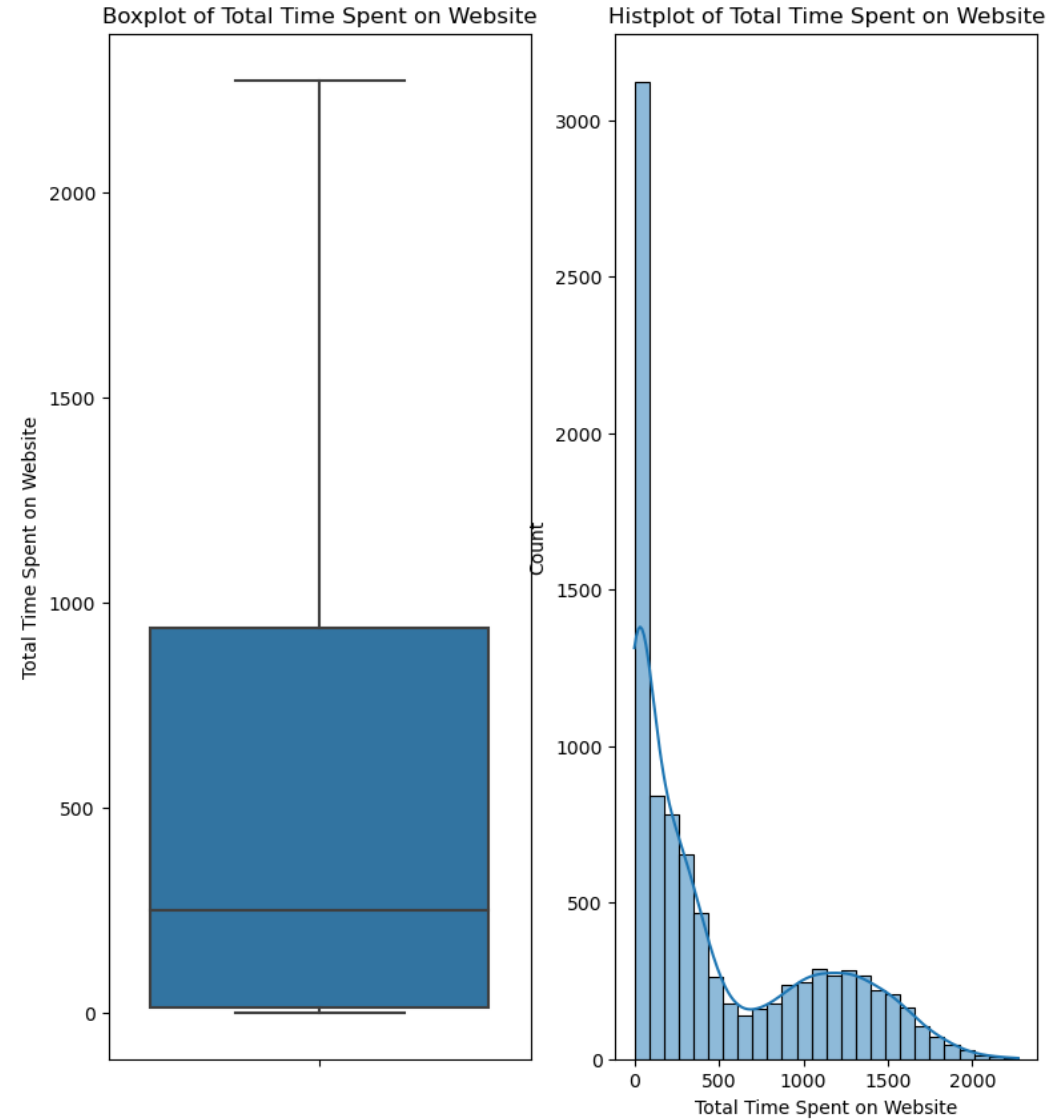


Histplot of TotalVisits Variable Distribution



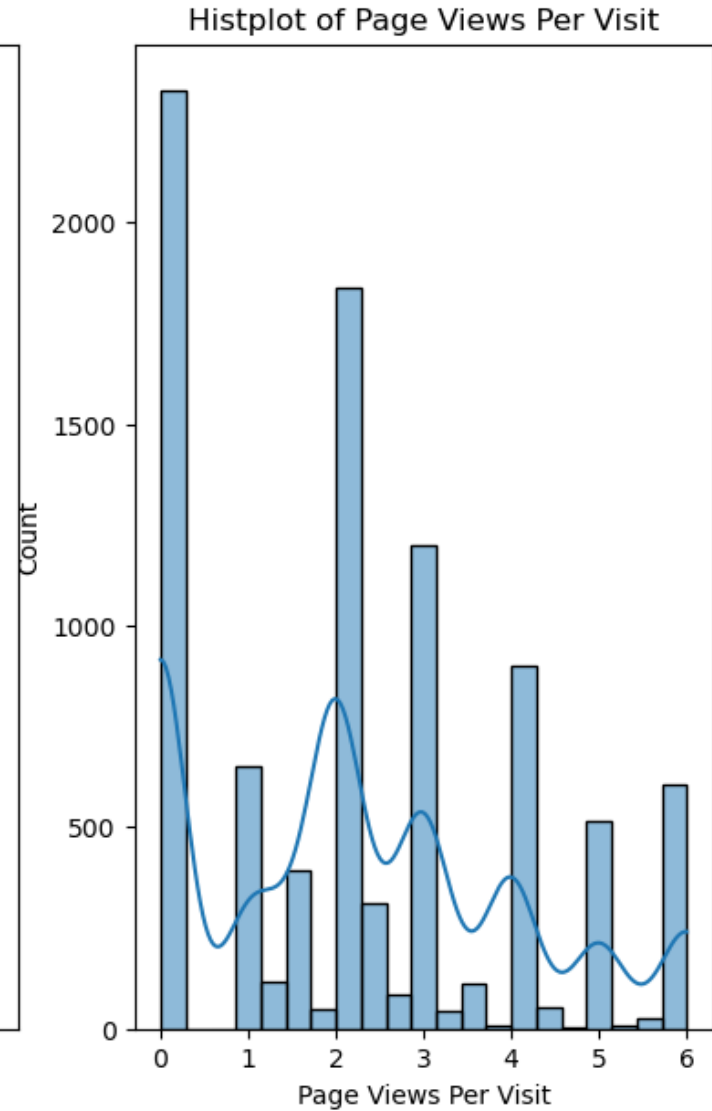
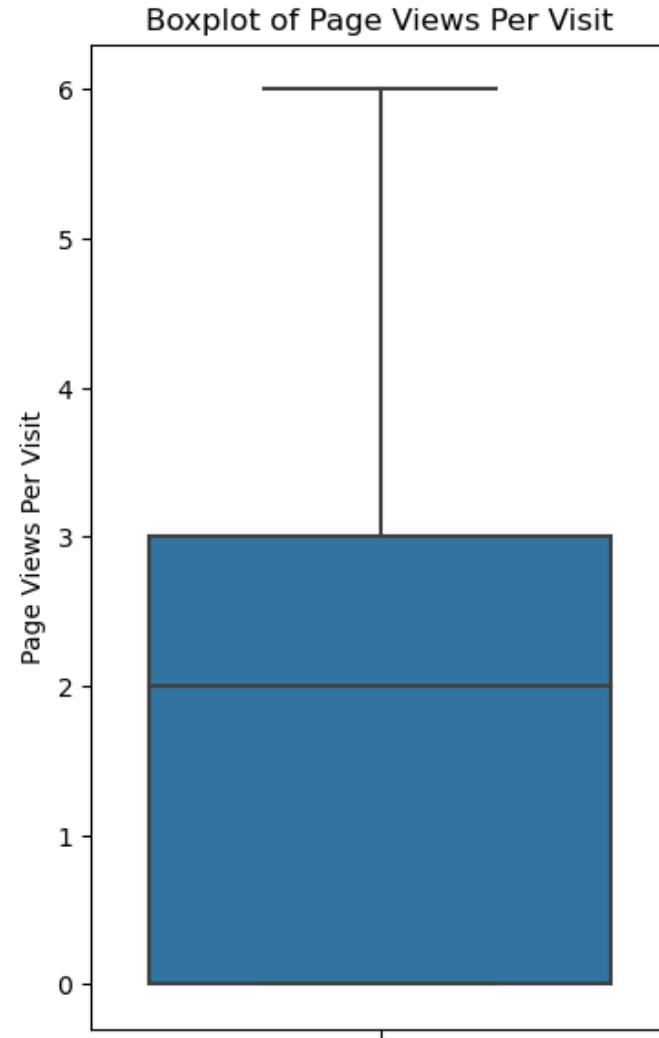
---

# TOTAL TIME SPENT ON WEBSITE VARIABLE DISTRIBUTION



---

# PAGE VIEWS PER VISIT VARIABLE DISTRIBUTION

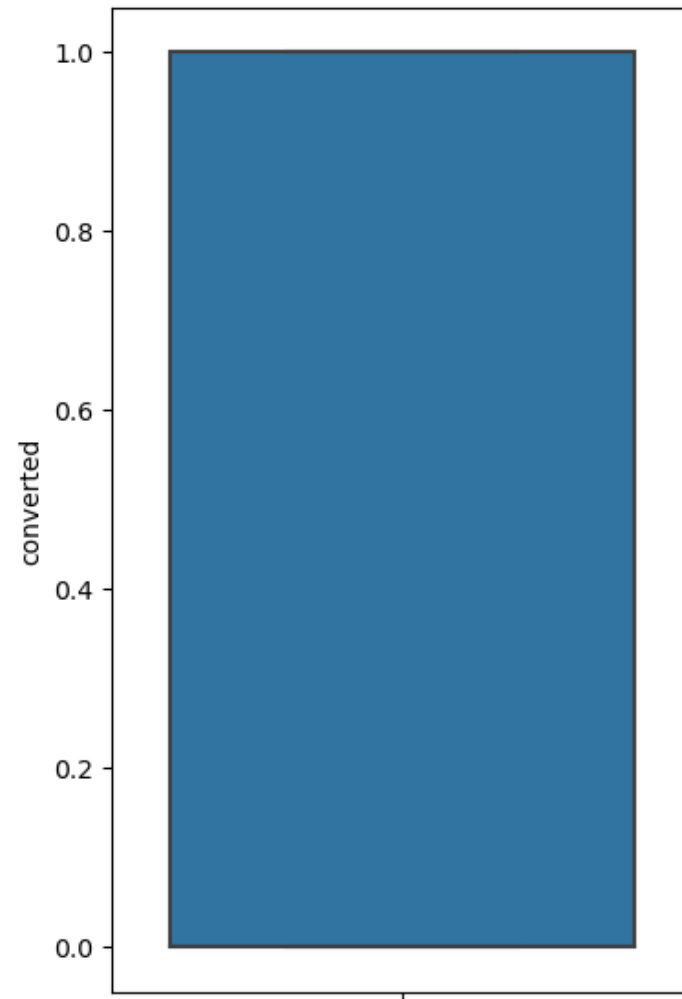




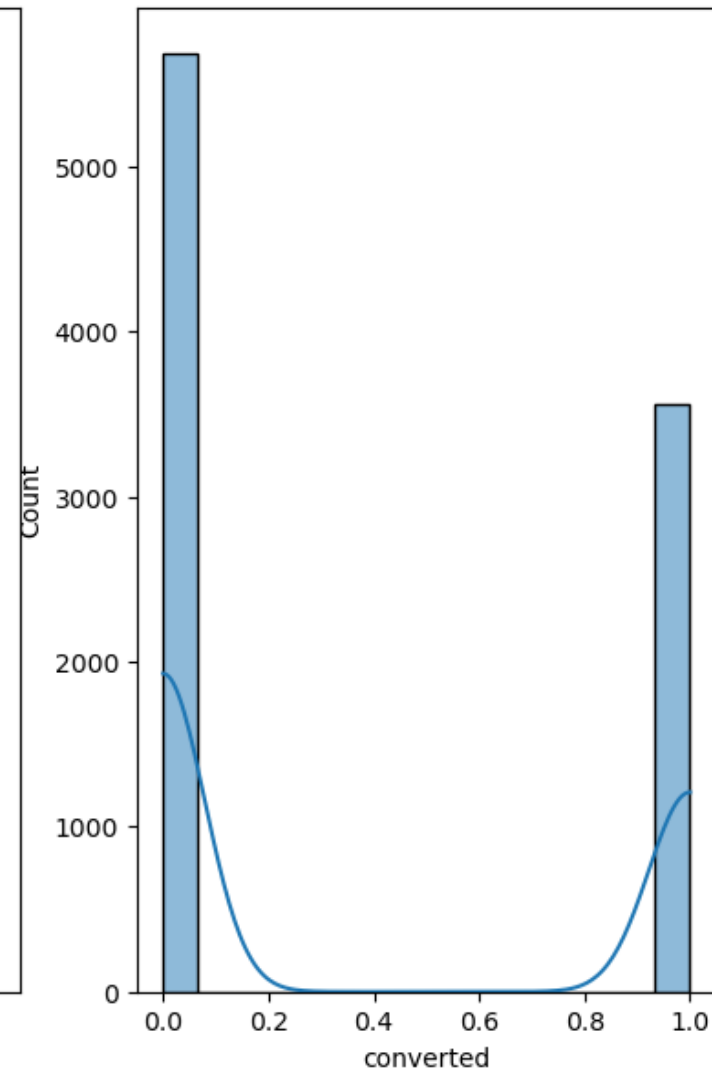
---

# TARGET VARIABLE DISTRIBUTION

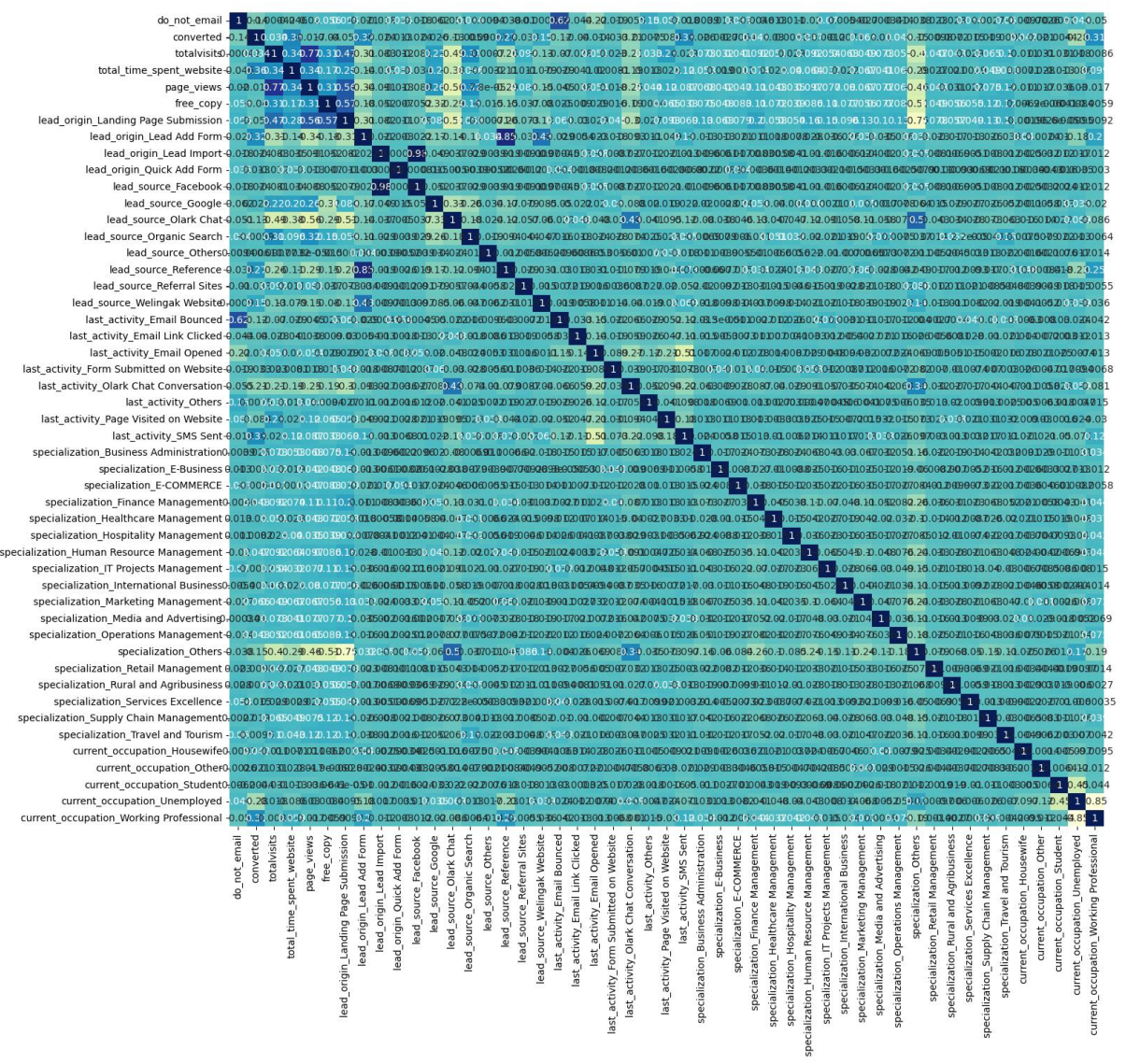
Boxplot of Target Variable Distribution



Histplot of Target Variable Distribution



# HEATMAP





**BUILD A MODEL**

---

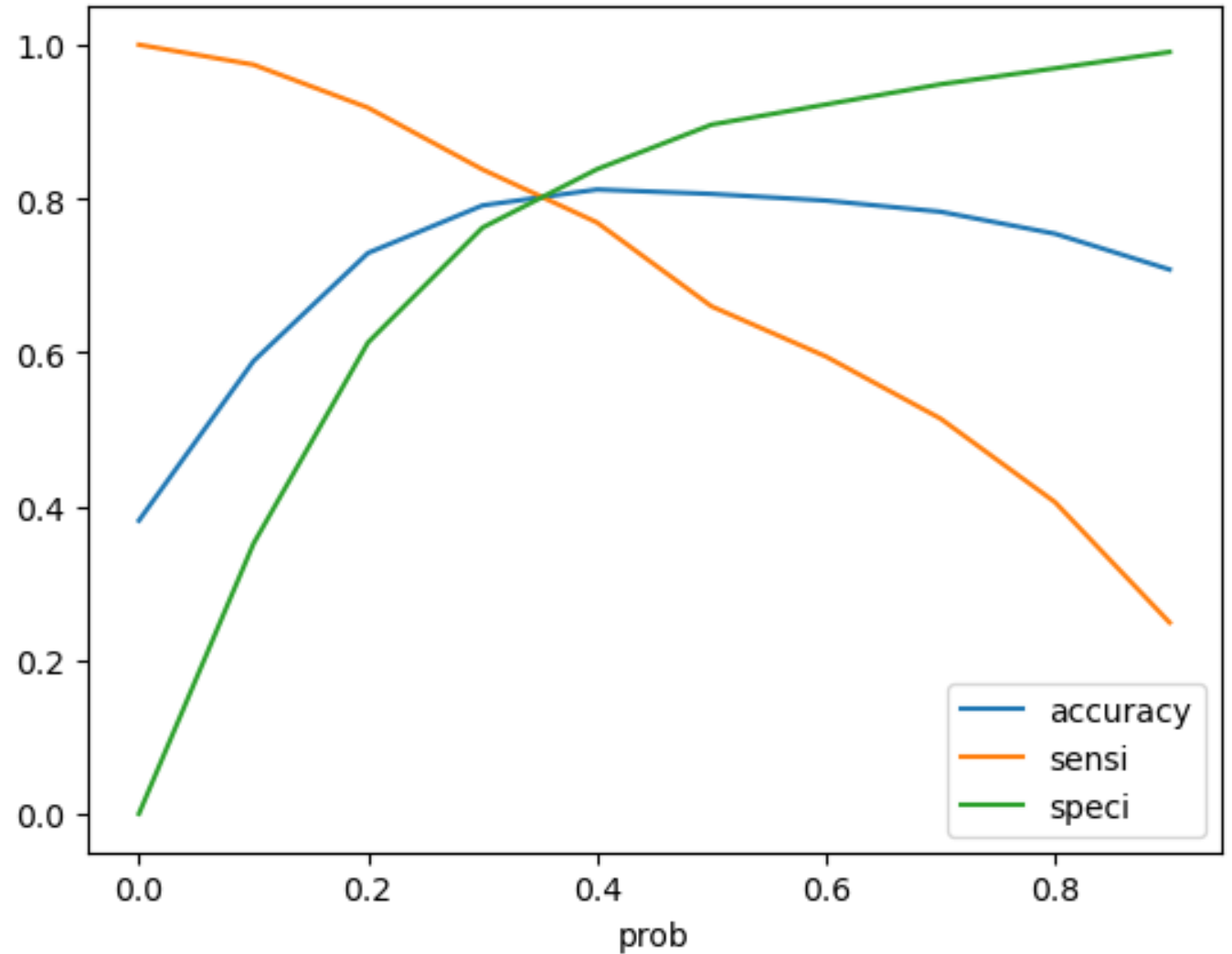
---

## 1. Explain the linear regression algorithm in detail.

- Data Split ( Train 70%, Test 30%)
  - Use RFE for Feature Selection
  - Run RFE 15 variables to select as output
  - Choose model if good VIF values less than 5 and all variables of p-values is less than 0.05.
  - Accuracy of Train and Test Dataset are same, 0.81%
-

---

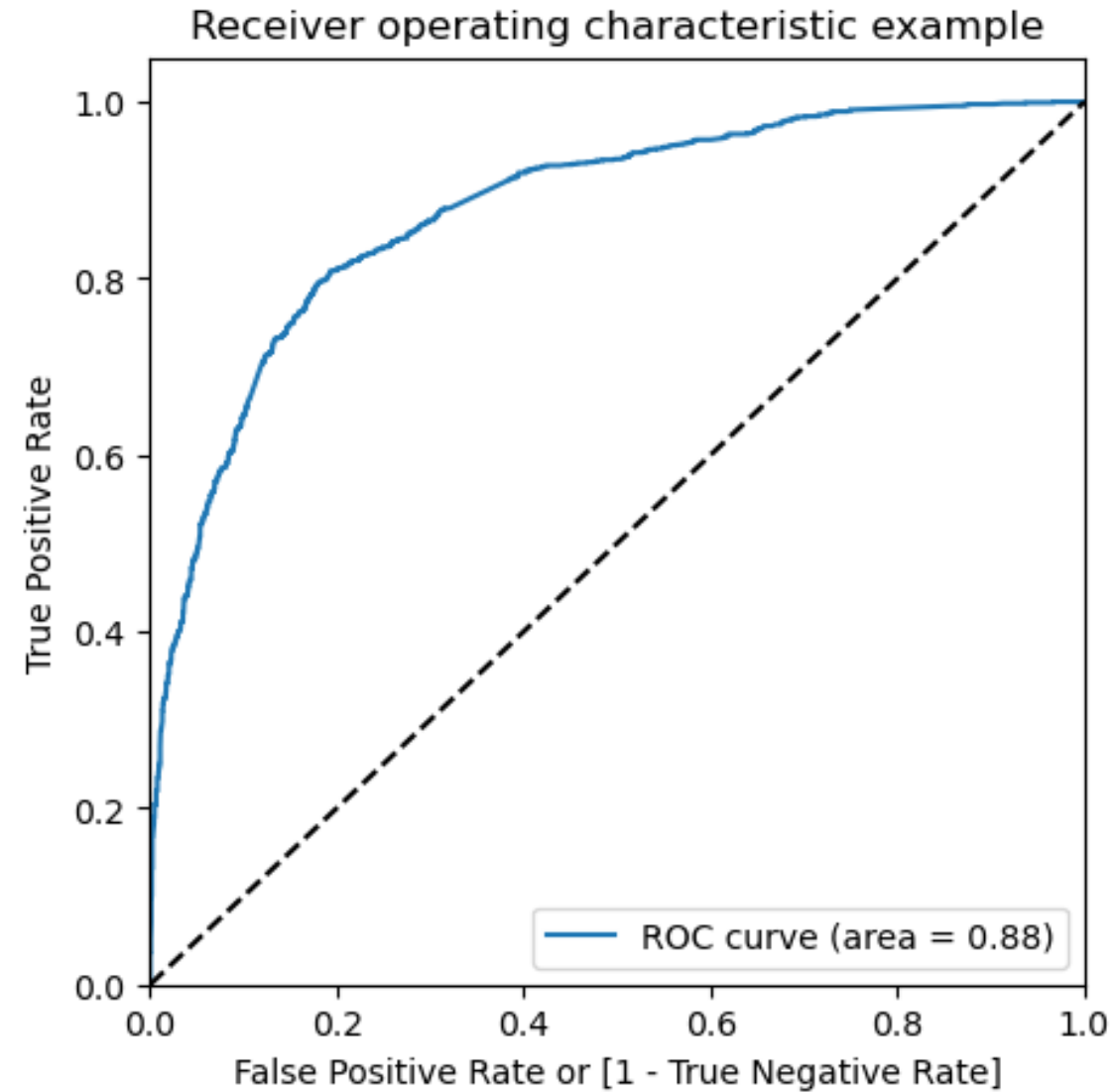
# OPTIMAL CUTOFF POINT



- 0.35 is the optimum point to take it as a cutoff probability

---

# ROC



ROC curve is  $0.88 < 1$  which indicates a good predictive model

---

# CONCLUSION

- Train Dataset
    - Accuracy: 0.81%
    - Sensitivity: 0.66%
    - Specificity: 0.89%
  - Test Dataset
    - Accuracy: 0.81%
    - Sensitivity: 0.80%
    - Specificity: 0.82%
  - Top three Features:
    - 1. lead\_source\_Welingak Website
    - 2. lead\_source\_Reference
    - 3. current\_occupation\_Working Professional
-

---

Thank you