Lead Scoring Case Study

Group Members
Dr. Reshma S
Ashwin Nadig
Shubra Rath

Problem Definition

- Education Company sells the course.

- The people who visits the webpage, watches video and questions about the courses are more likely to purchase the course.

- All the viewers are not purchasing the course.

- So, the problem here is to identify the lead viewers who will purchase the course.

- This is decided based on lead score. The maximum lead score viewer purchase the course and termed as hot and the smaller lead score viewers are not most interested in purchasing and it termed as cold

Steps Followed

- Imported Packages

- Imported data set

- Identified and Cleaned up missing/null values

- Perform univariate analysis

- Established relationship between target variable with all other categorical variable

- Model building

- Evaluated model using cut off to evaluate sensitivity, simplicity, precision and recall
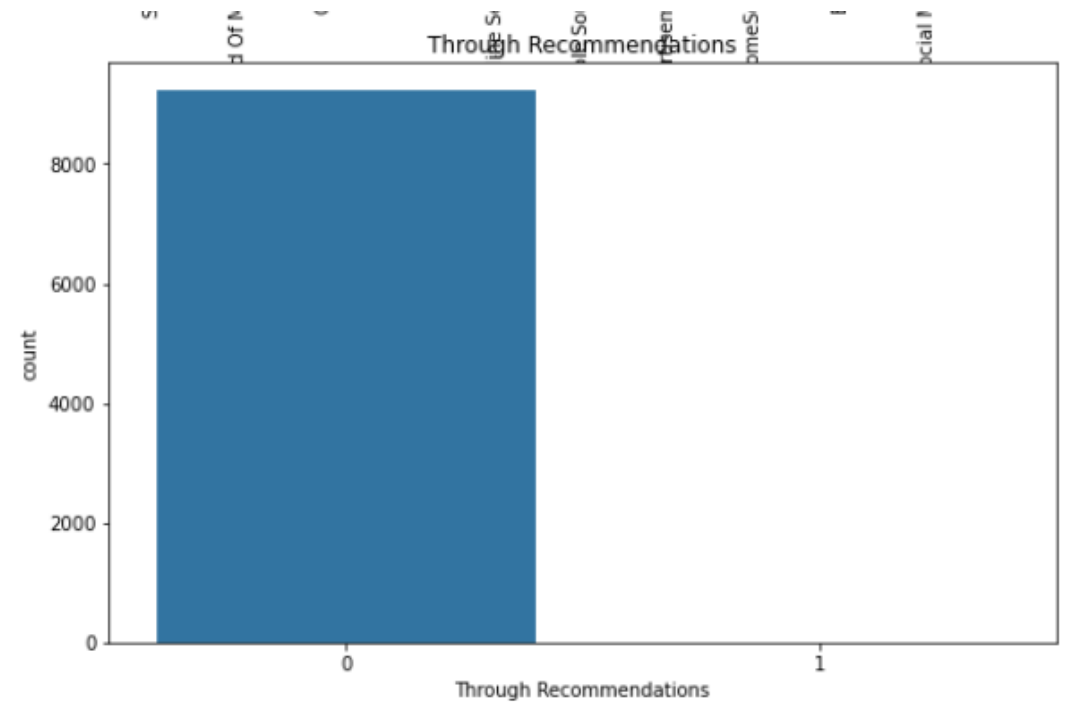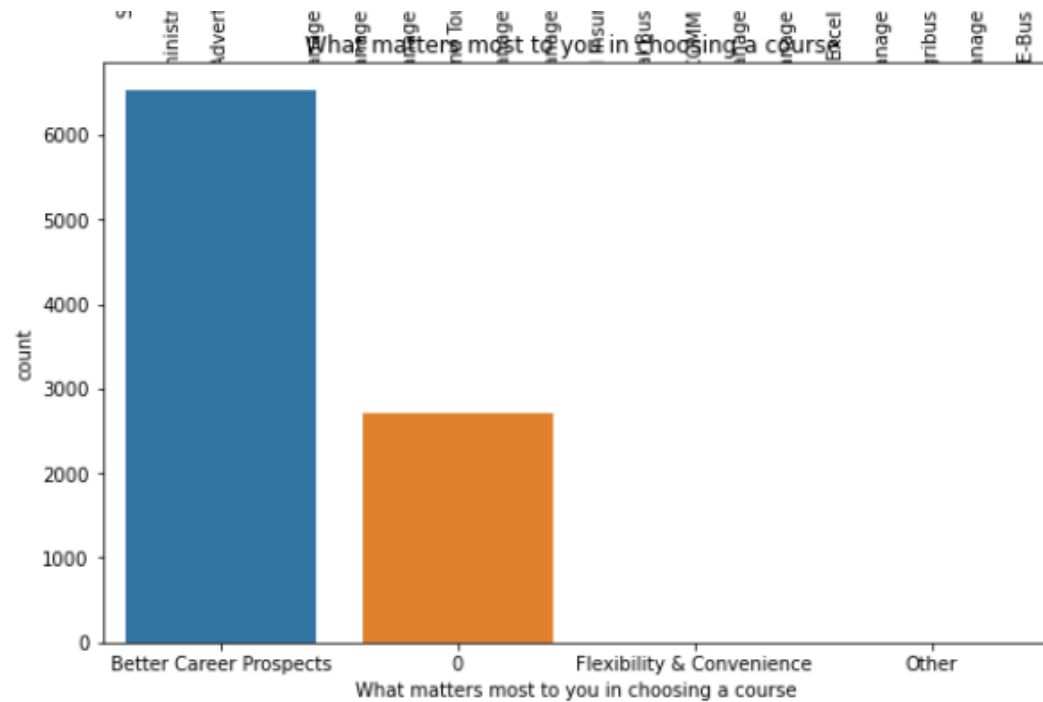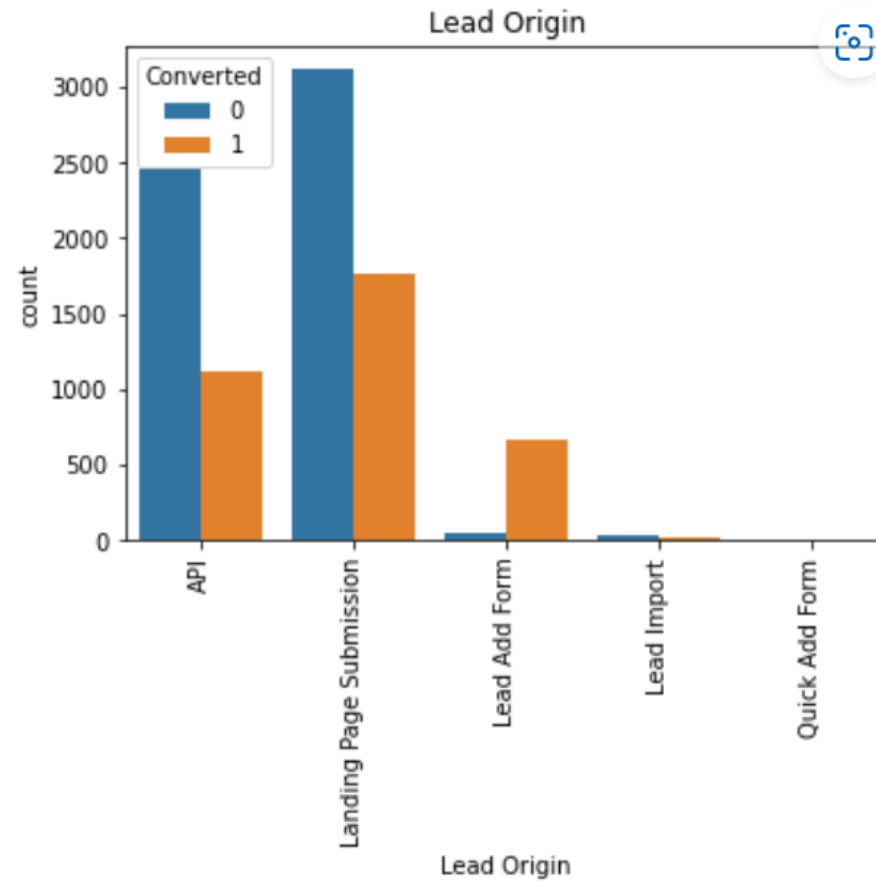
Univariate Analysis
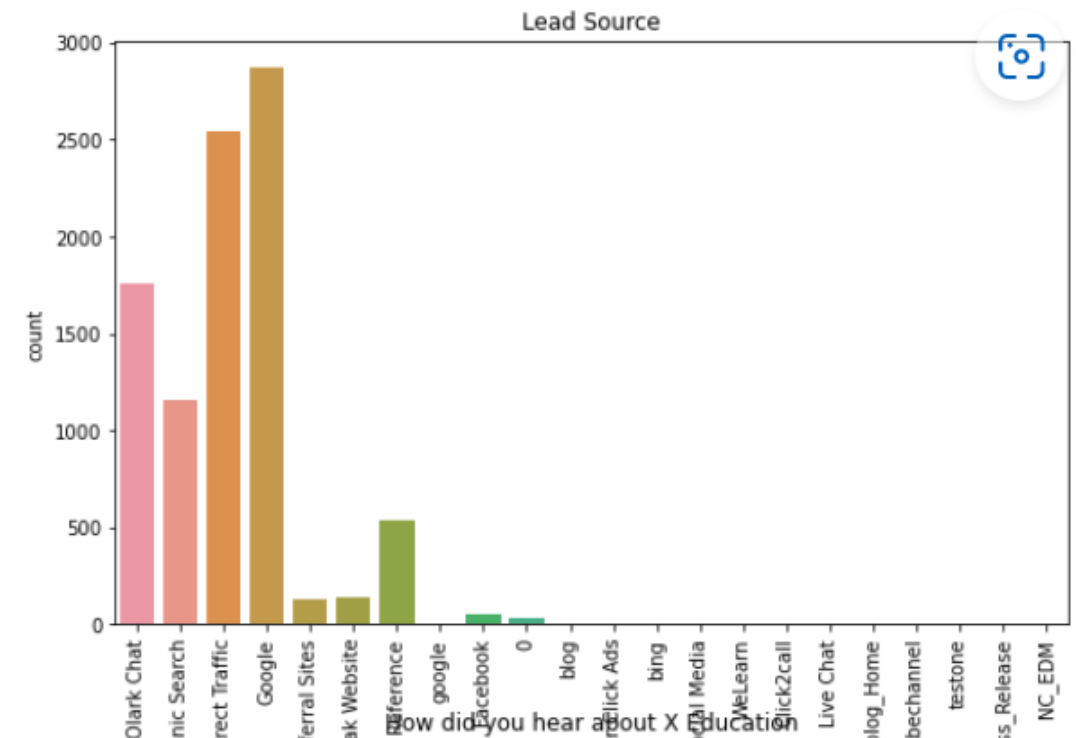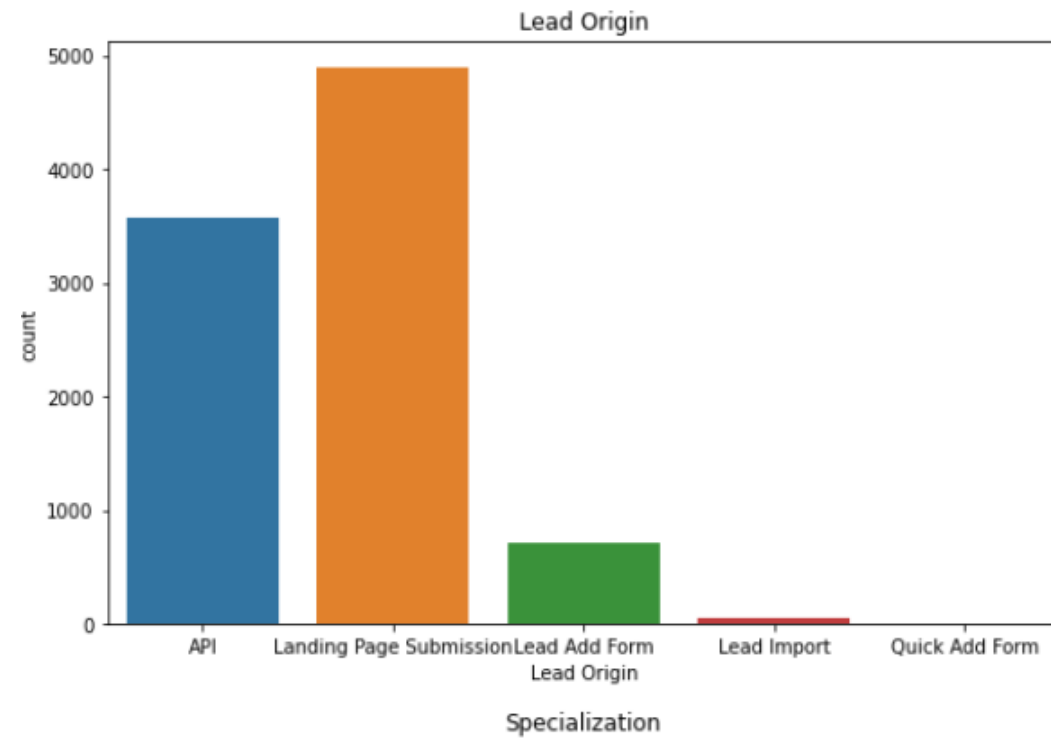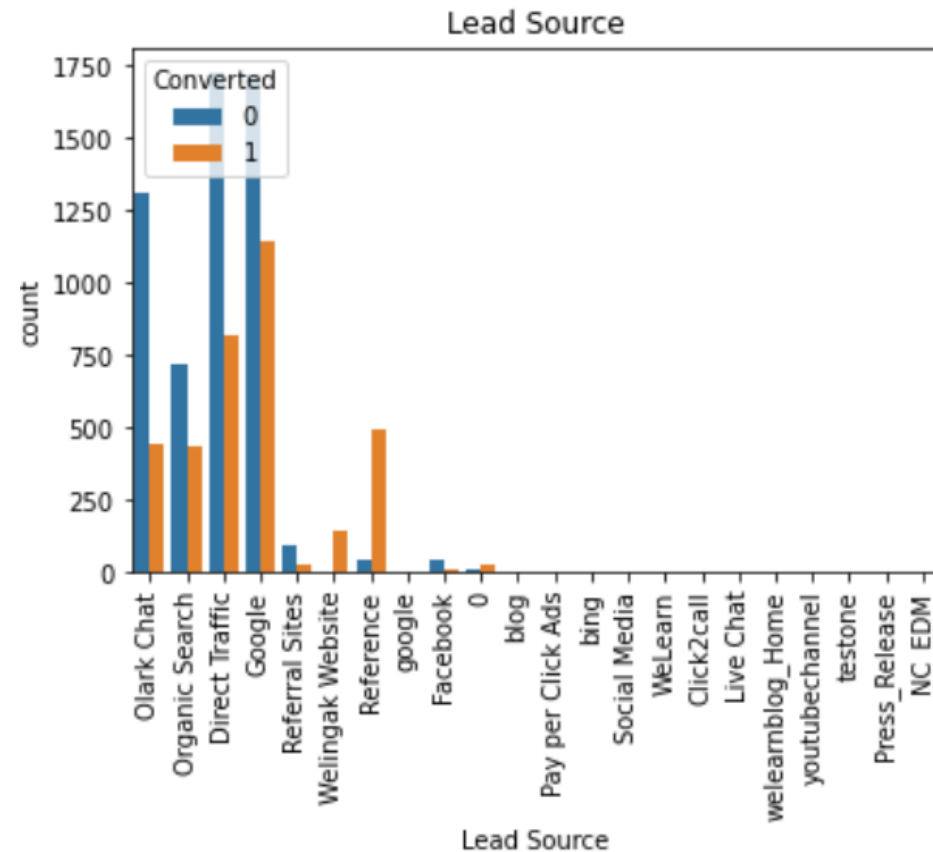
Univariate Analysis (Contd..)
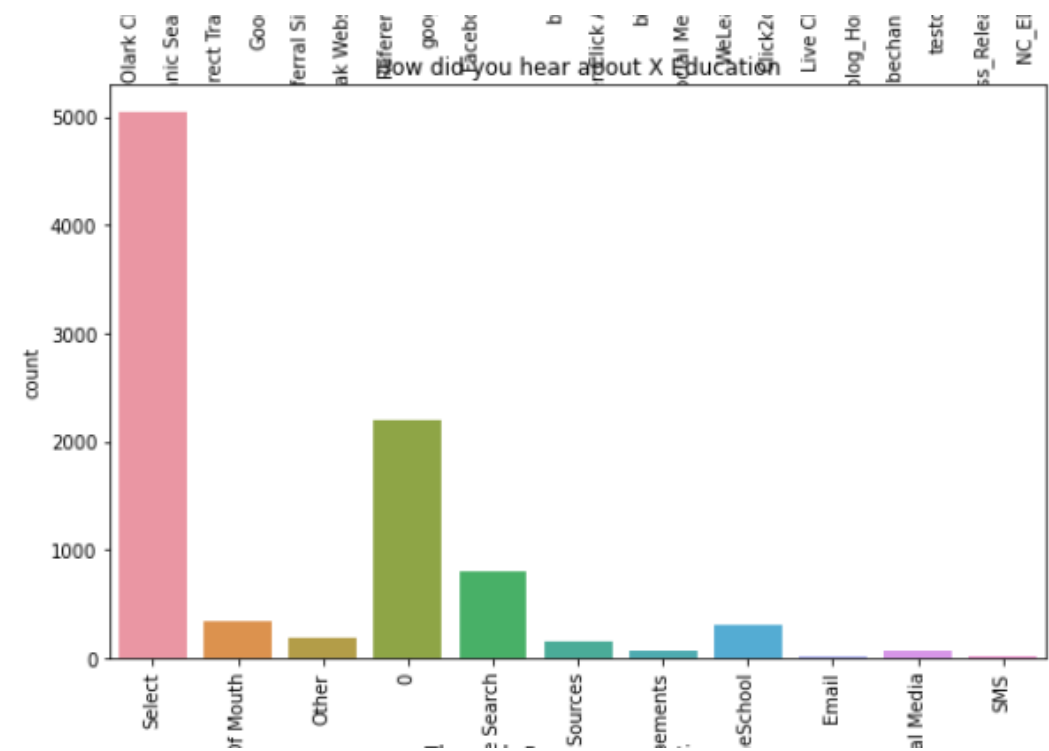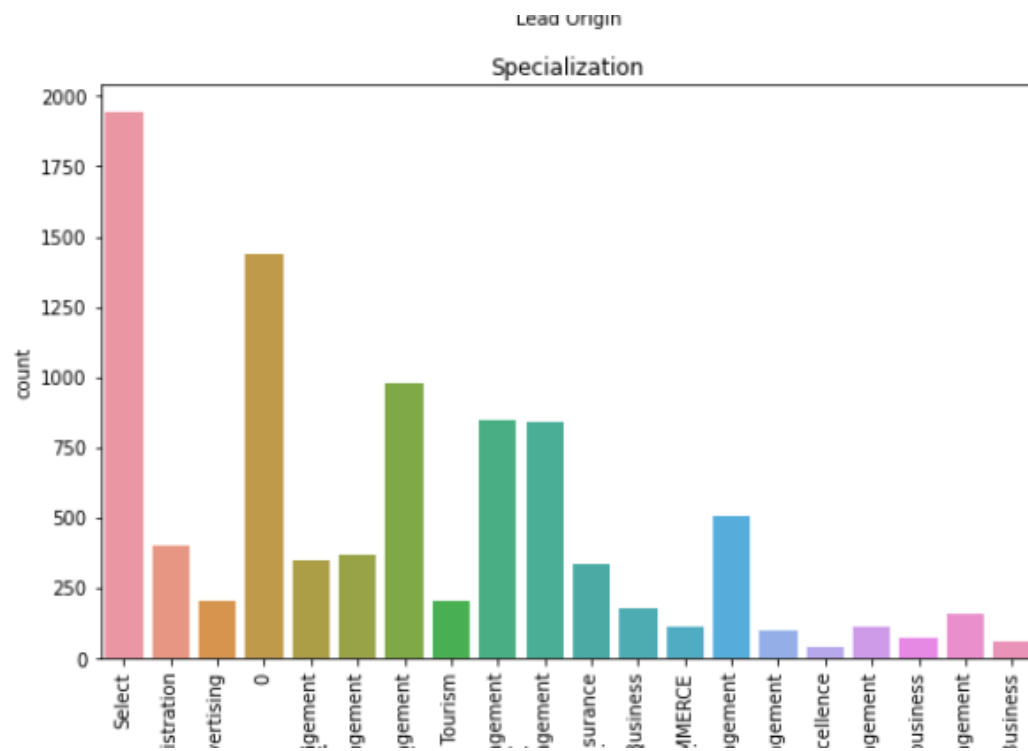
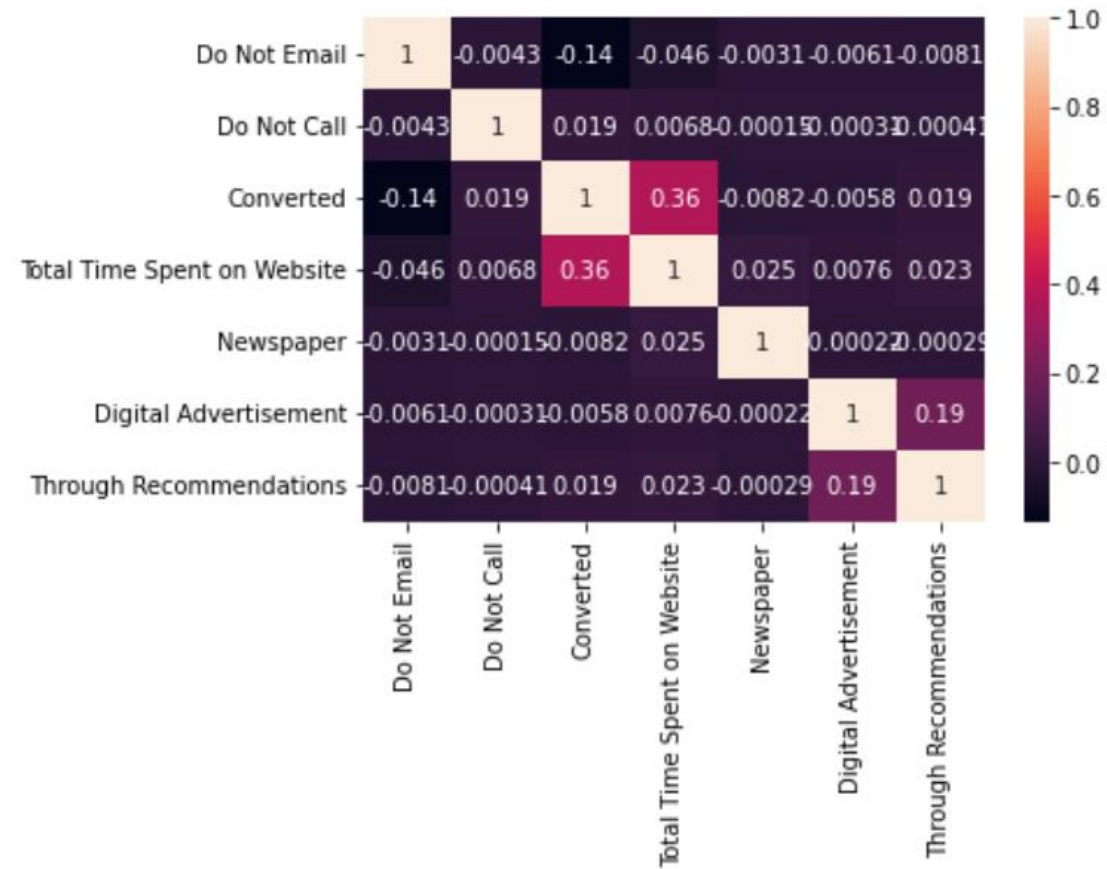# Univariate Analysis (Contd..)

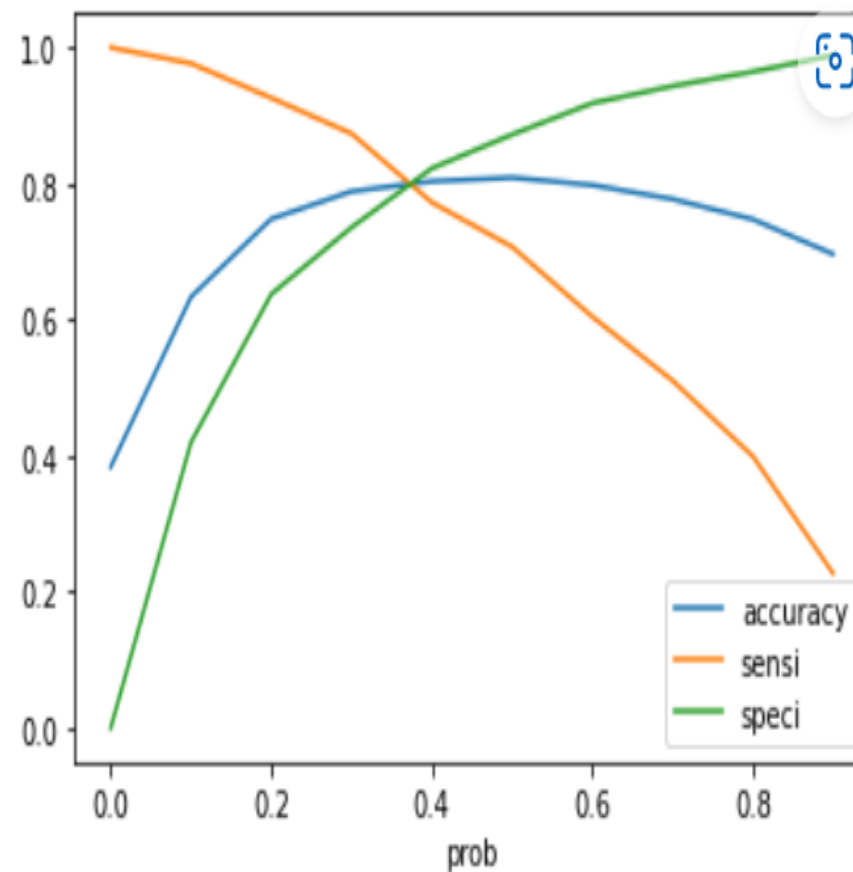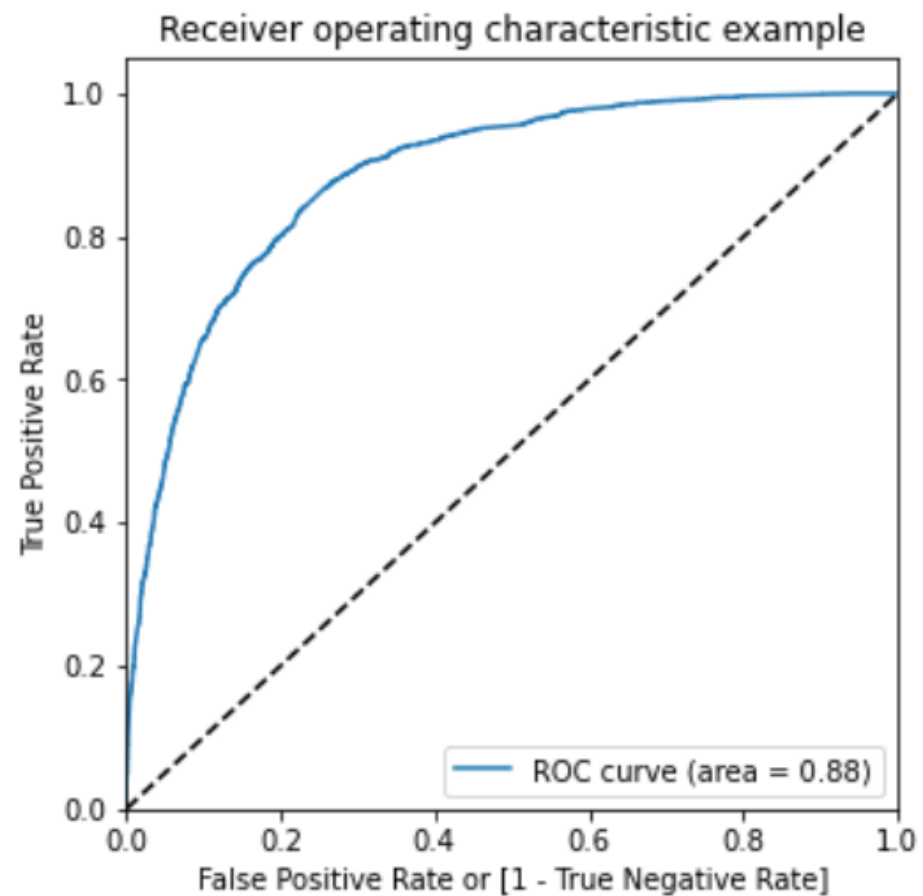# Categorical Variable Relationship Analysis

Categorical Variable Relationship Analysis (Contd ..)

Multivariate Analysis

ROC Curve

Summary

- Its observed that the maximum time spend on the website, total number of visits, the lead source (like Google, Direct traffic, Organic search, Welingak website), last activity (*viz.,* SMS, Olark chat conversation), lead origin and when their current occupation is as a working professional are the key factors to predict the course buyers.

- Keeping these in mind the X Education can flourish as they have a very high chance to get almost all the potential buyers to change their mind and buy their courses.