Lead Score Case Study

TEAM MEMBERS

ASHUTOSH GOLE

NEHA DEVATHRAJ

Problem Statement

- X Education sells online courses to industry professionals.
- X Education gets a lot of leads, its lead conversion rate is very poor. For example, if, say, they acquire 100 leads in a day, only about 30 of them are converted.
- To make this process more efficient, the company wishes to identify the most potential leads, also known as ‘Hot Leads’.
- If they successfully identify this set of leads, the lead conversion rate should go up as the sales team will now be focusing more on communicating with the potential leads rather than making calls to everyone.

Business Objective:

- X education wants to know most promising leads.
- For that they want to build a Model which identifies the hot leads.
- Deployment of the model for the future use

Goals of the Study

Our Goals of the Case Study:

To build a logistic regression model to assign a lead score between 0 and 100 to each of the leads which can be used by the company to target potential leads.

To adjust to if the company's requirement changes in the future so you will need to handle these as well.

Steps:

Read and understand the data

Clean the data

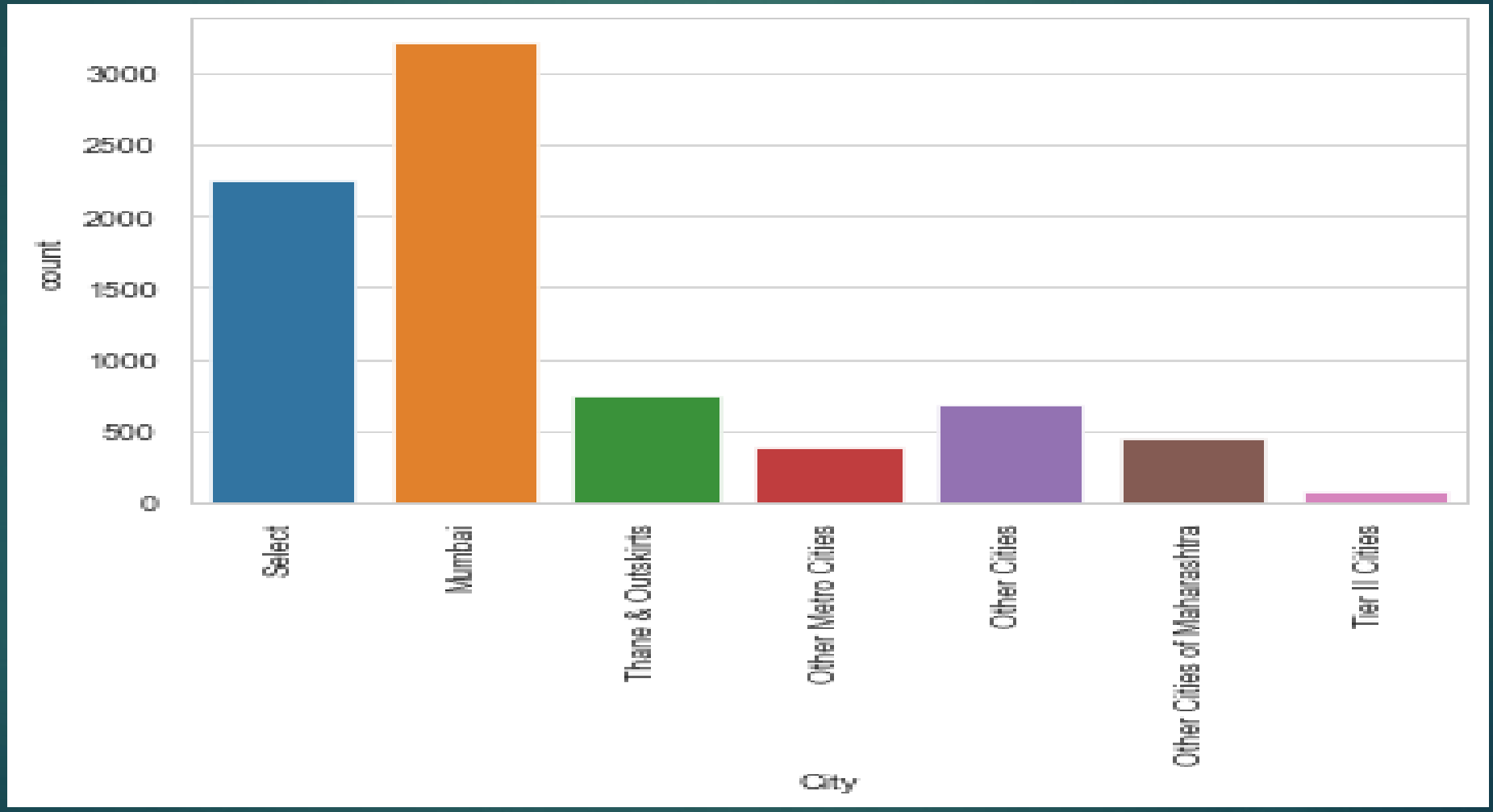
Prepare the data for Model Building

Model Building

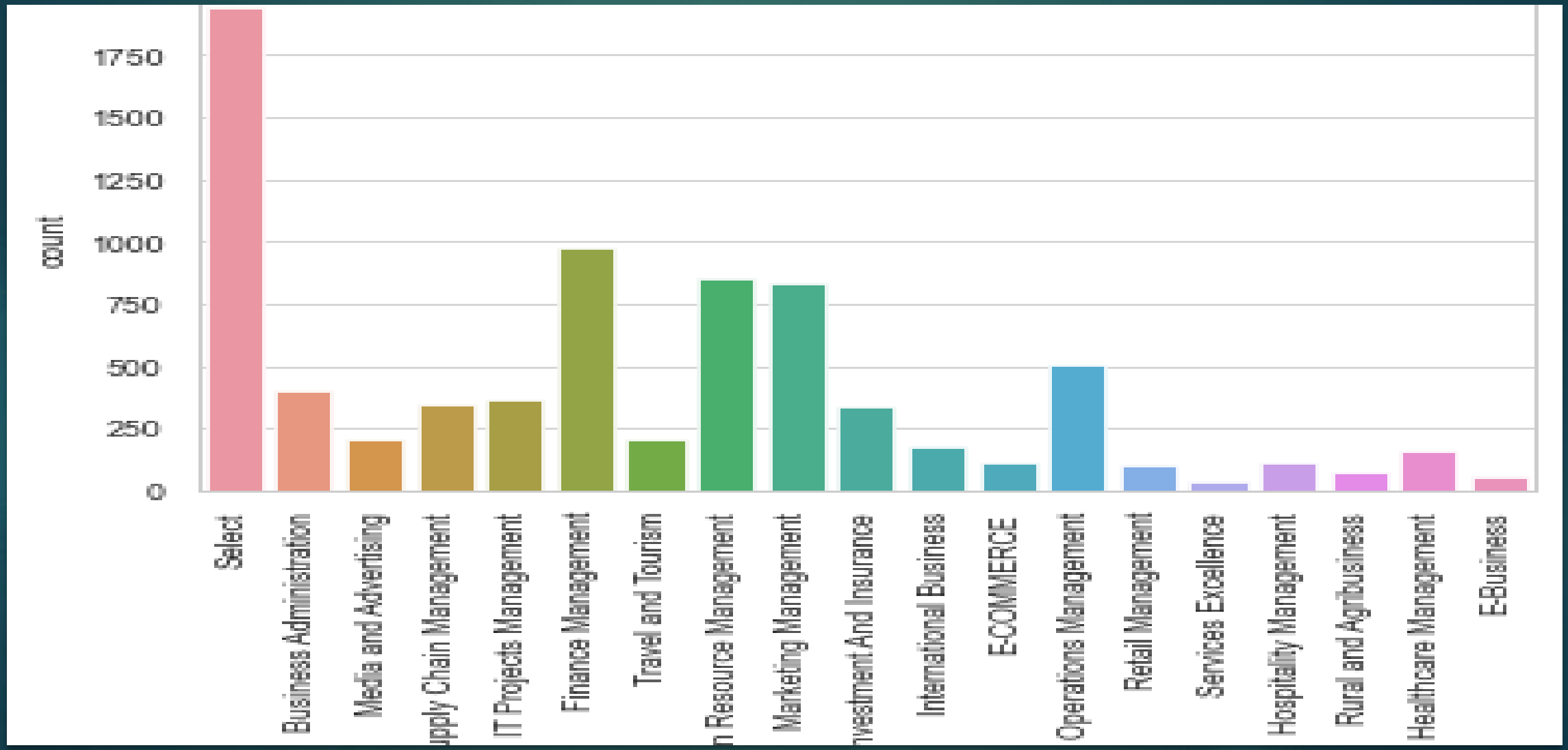
Model Evaluation

Making Predictions on the Test Set

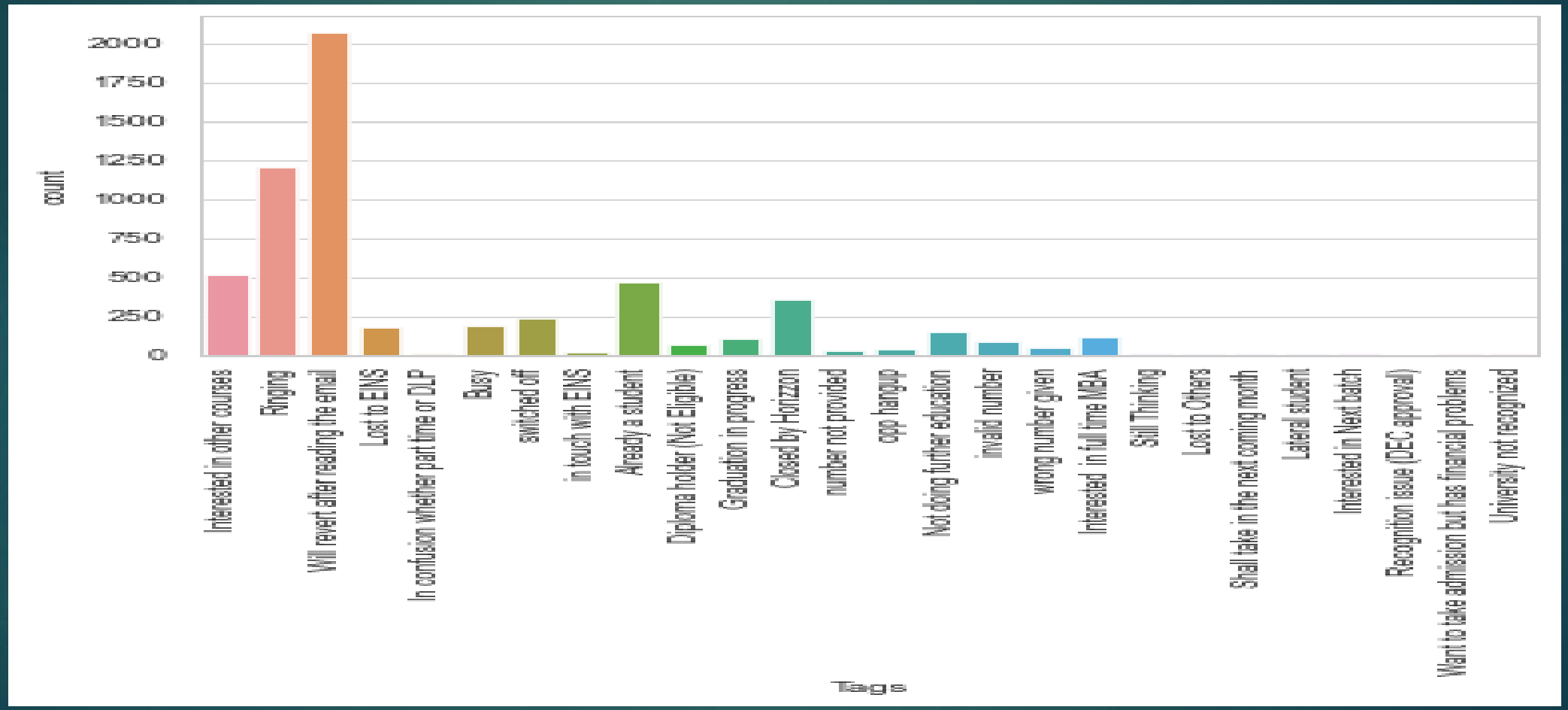
Data Cleaning and Preparation



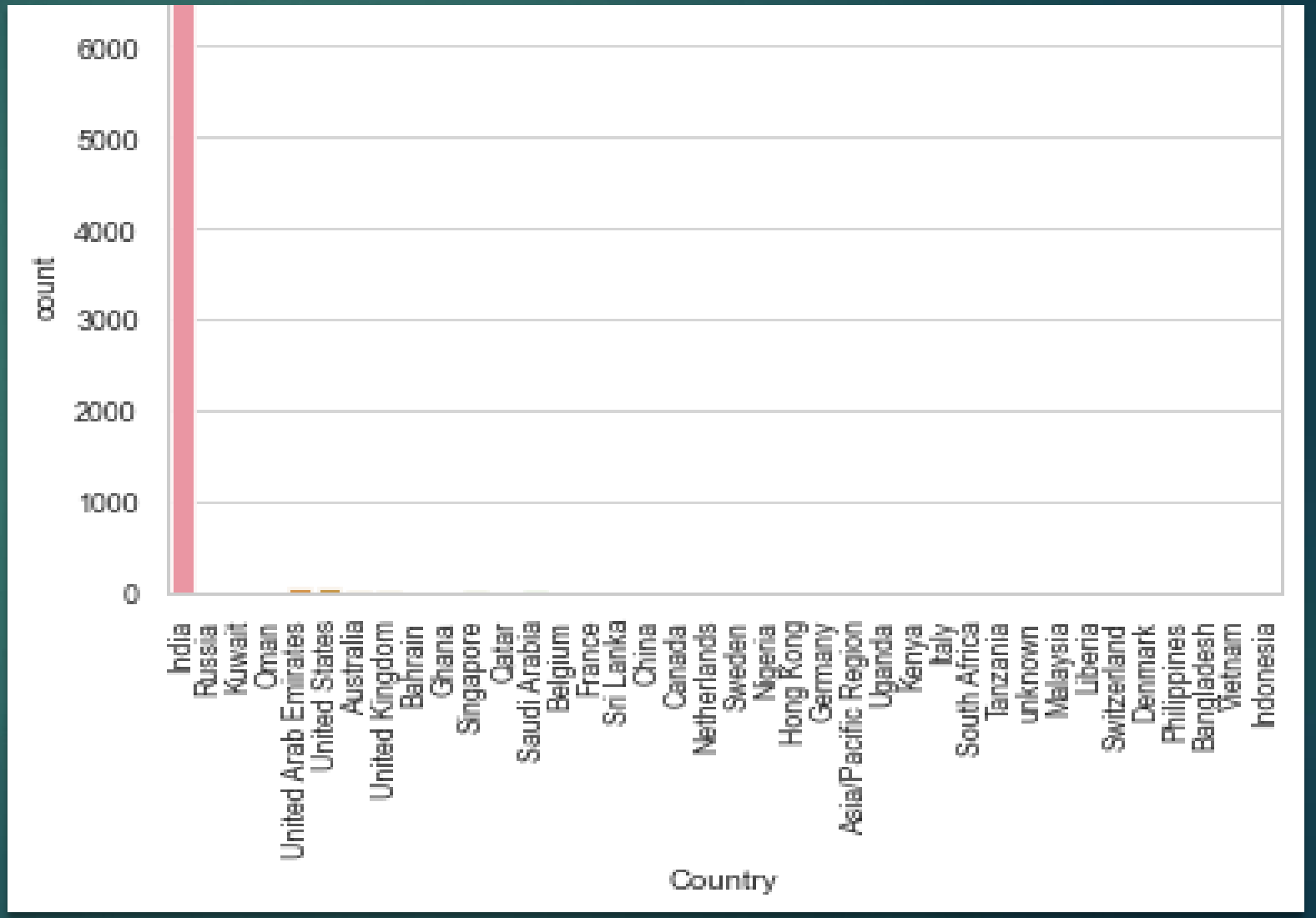
Checking specialization



Checking Tags

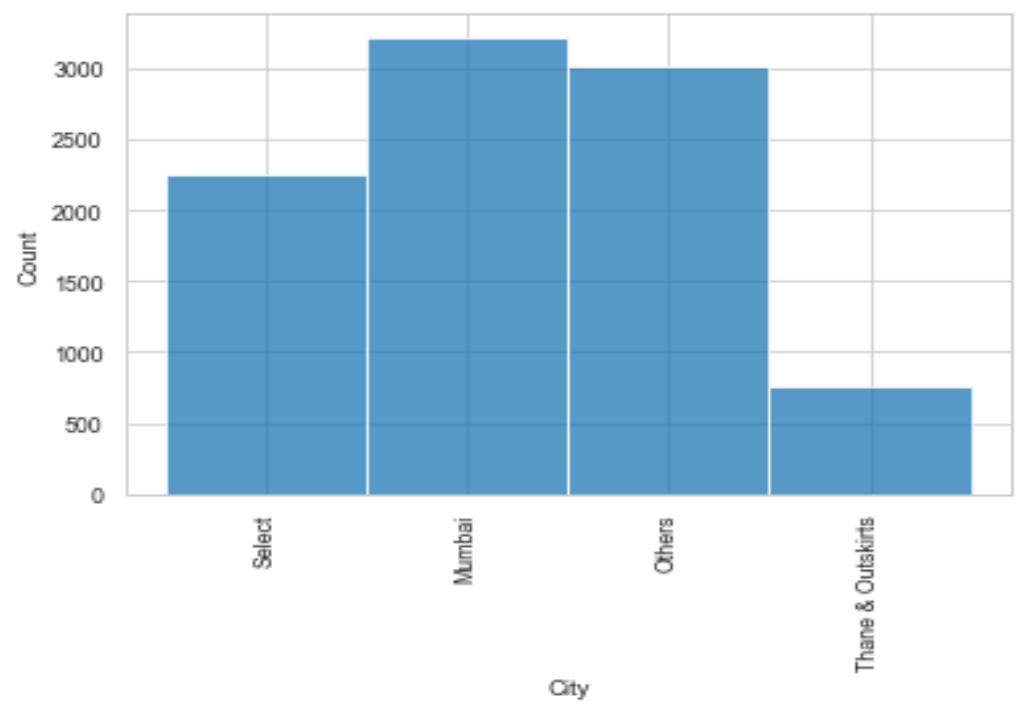


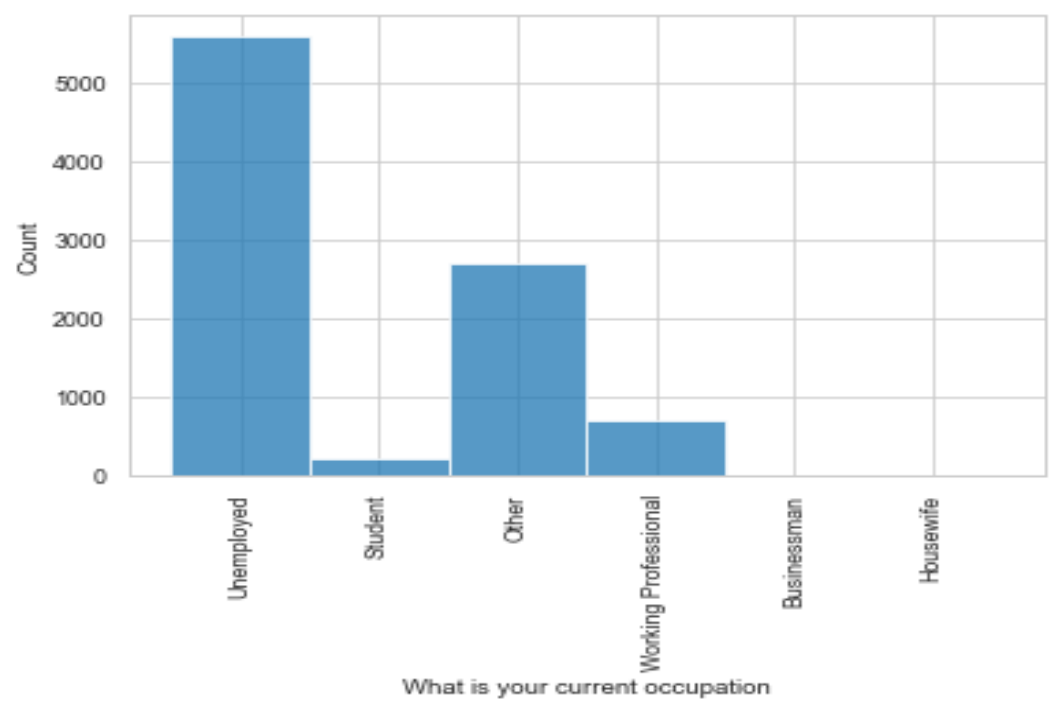
Checking Countries



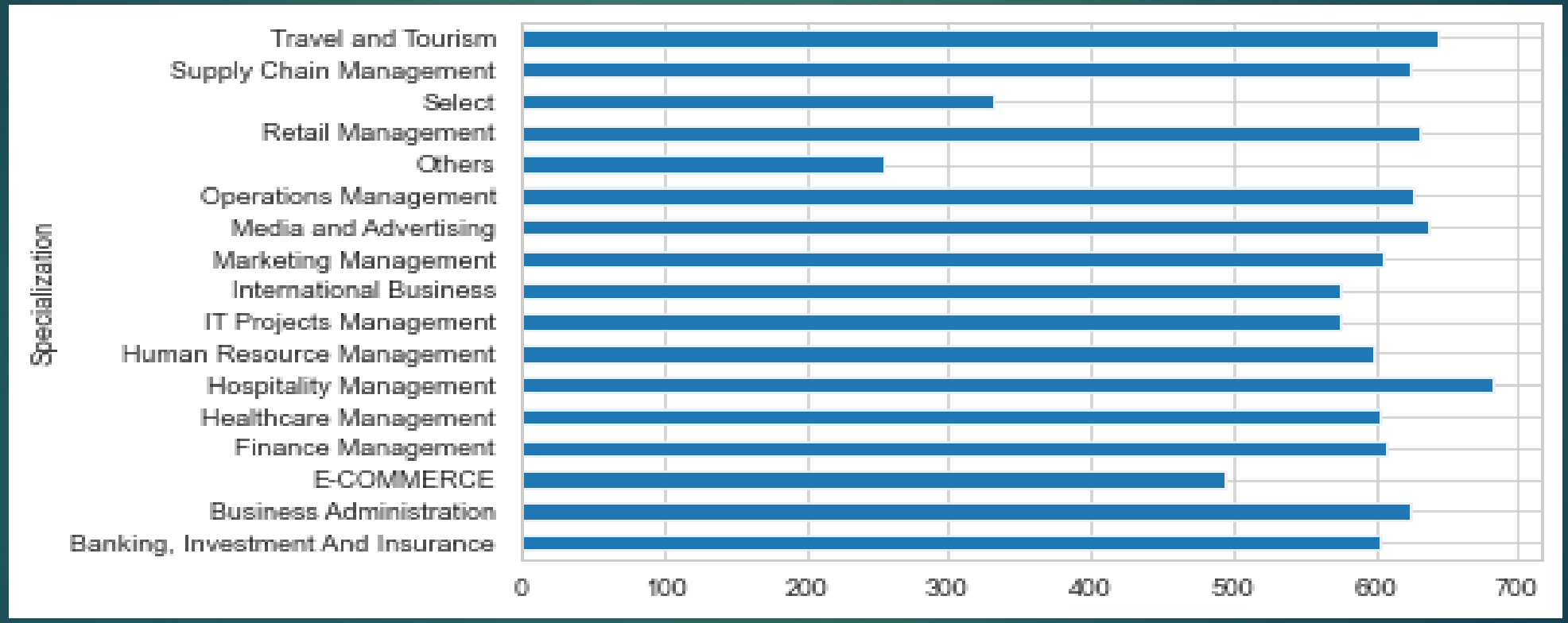
Data Visualization/ EDA

Univariate Analysis

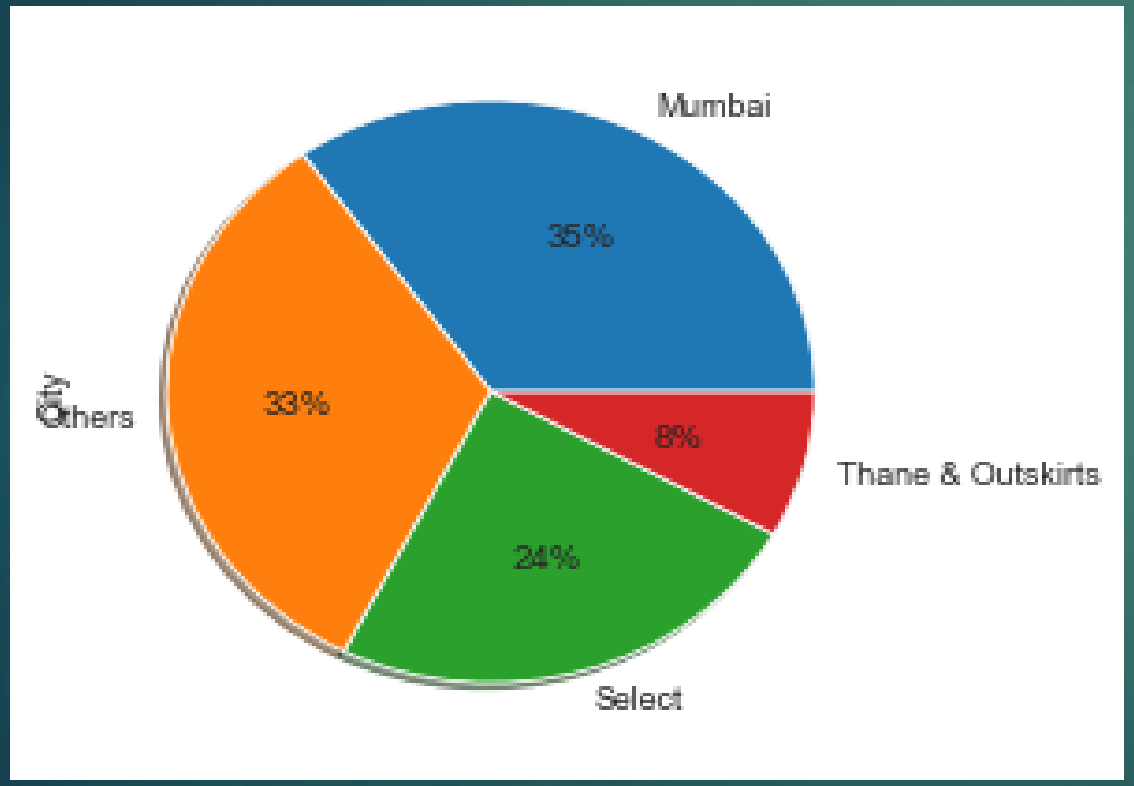


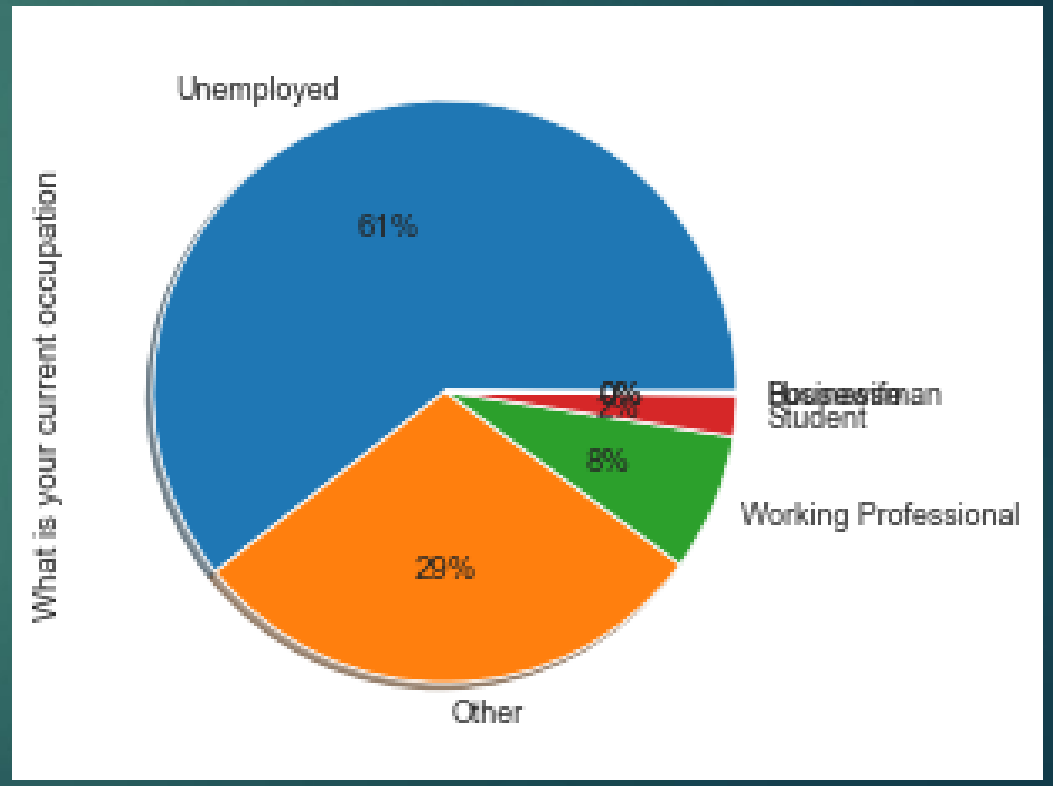


Total time spent on website

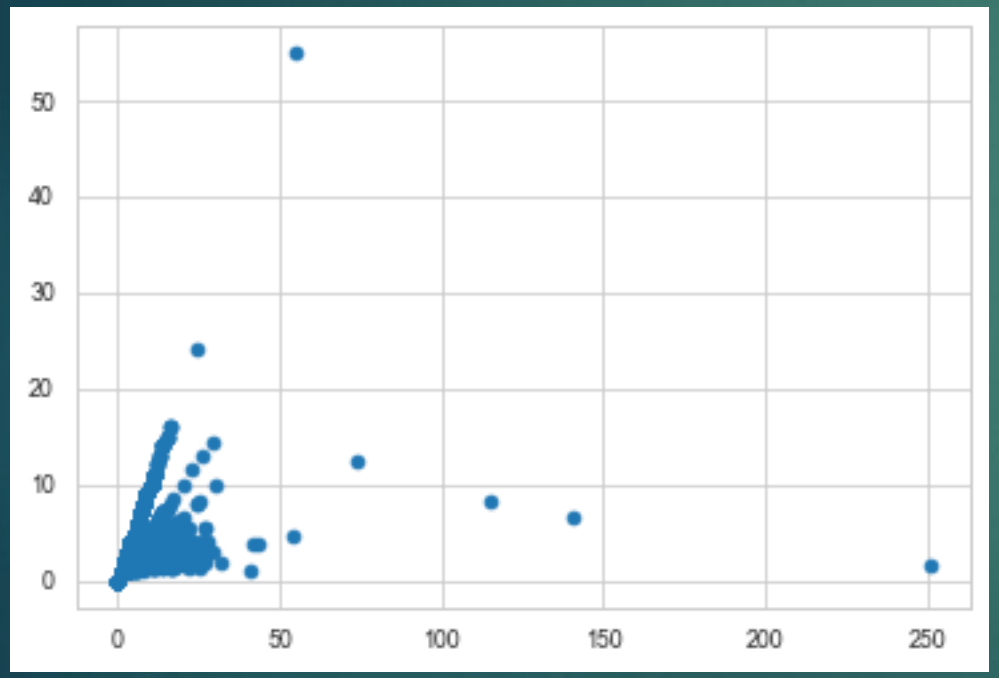


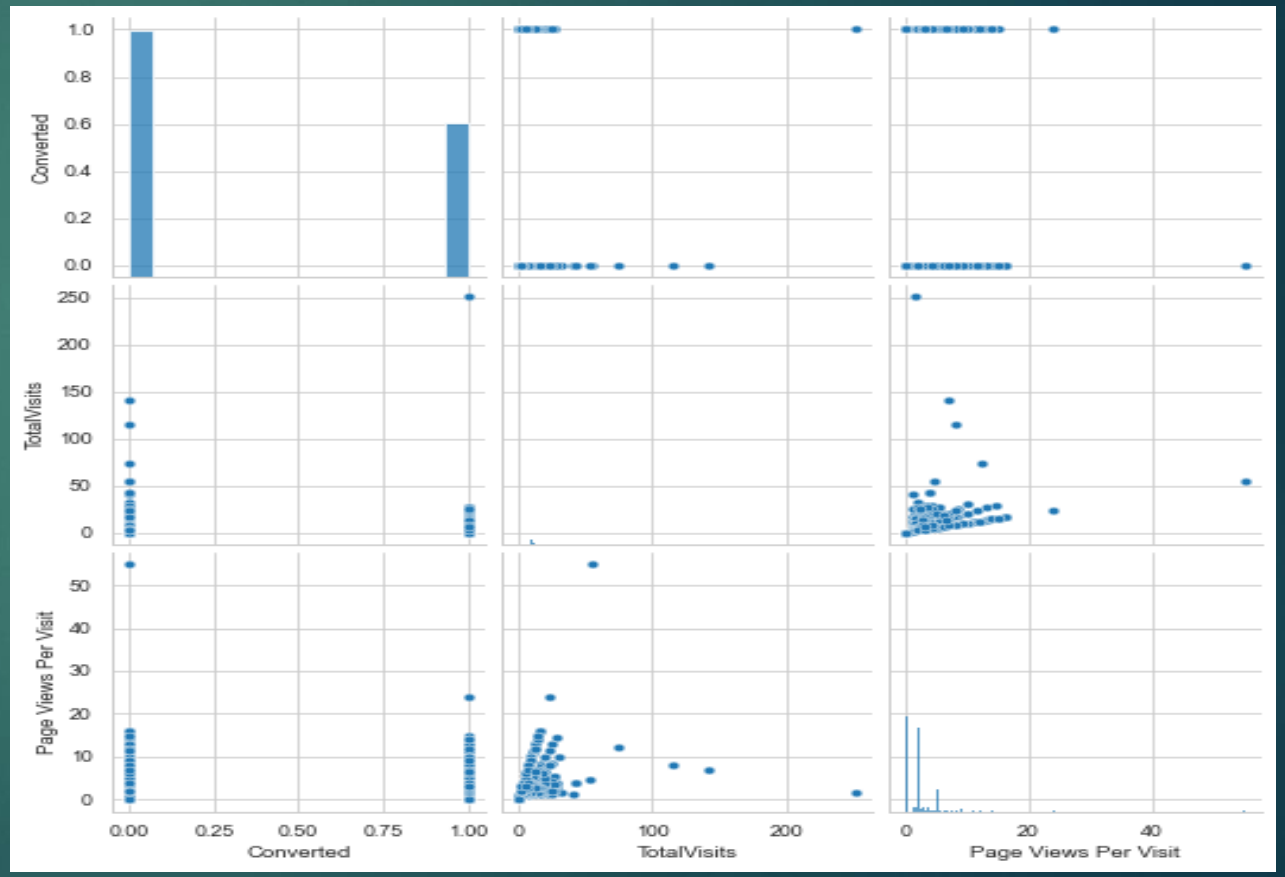
Categorical ordered univariate analysis

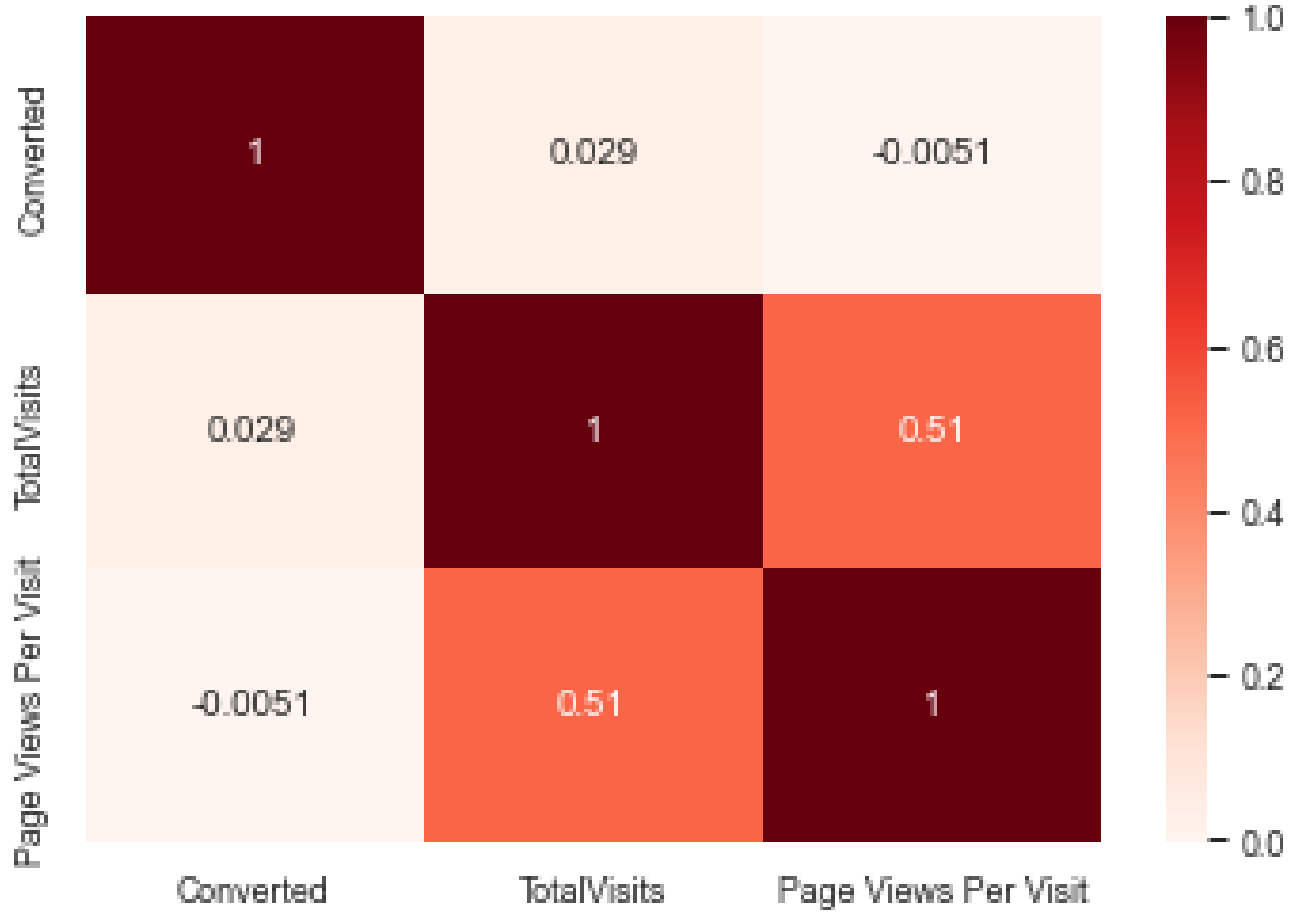




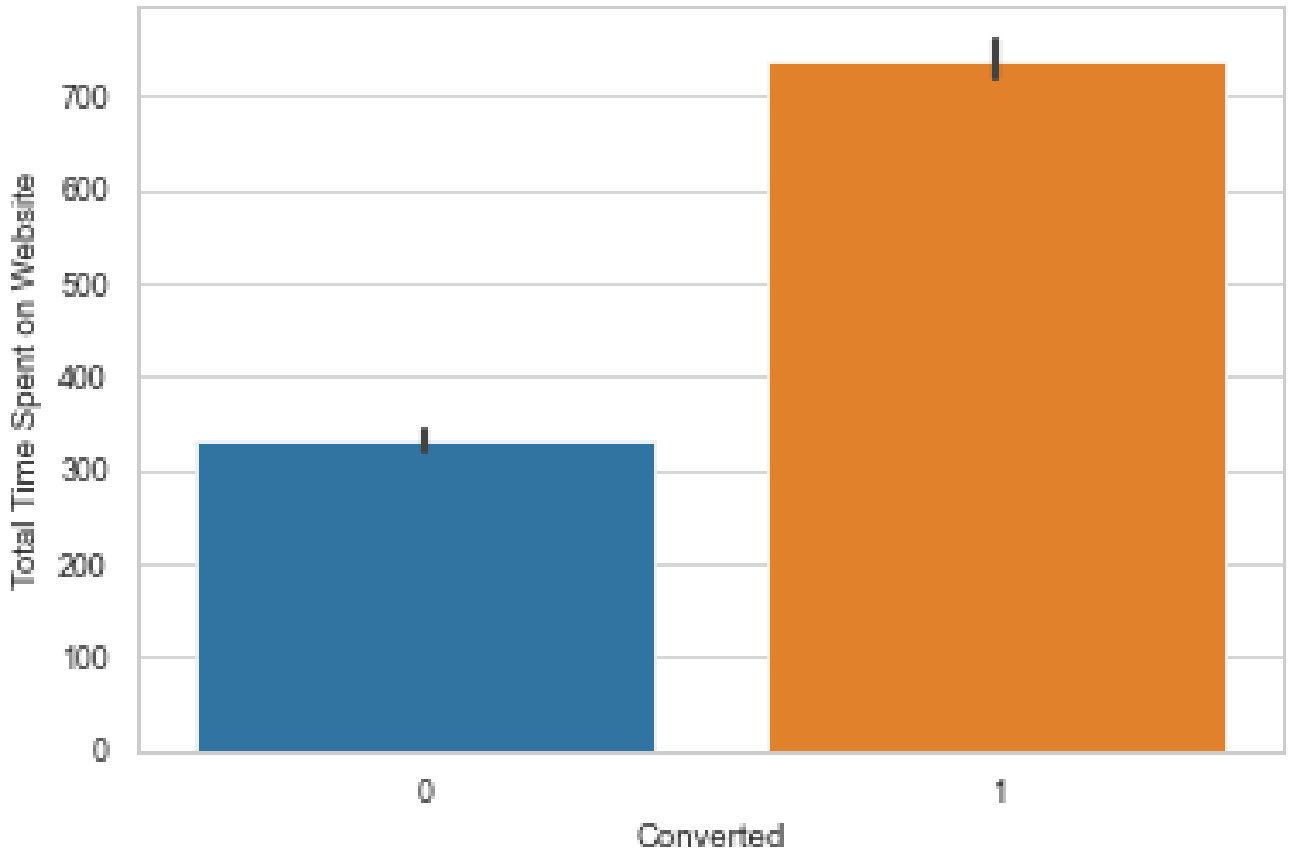
Bivariate and Multivariate Analysis







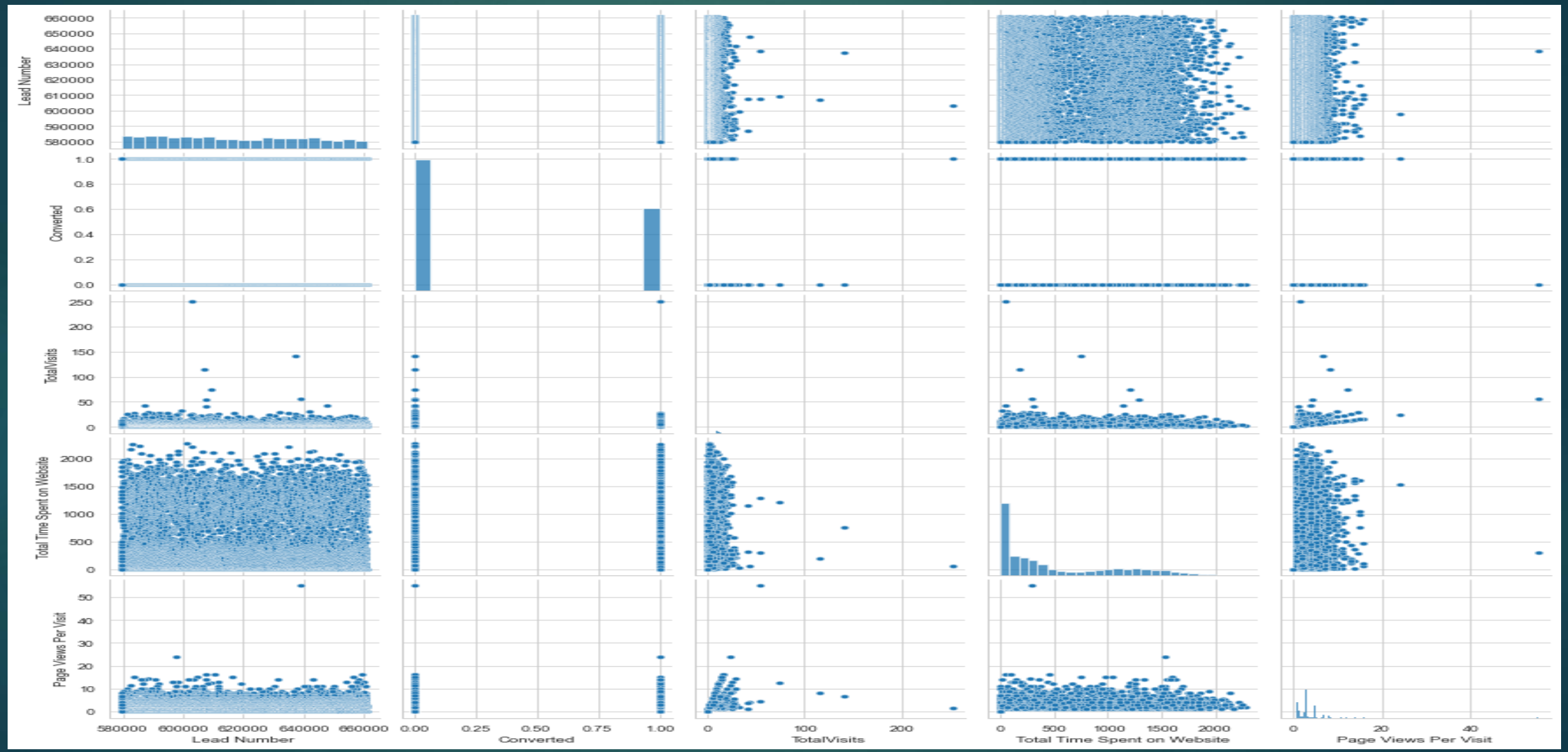
Visualization of the numeric data.



Numerical categorical variable

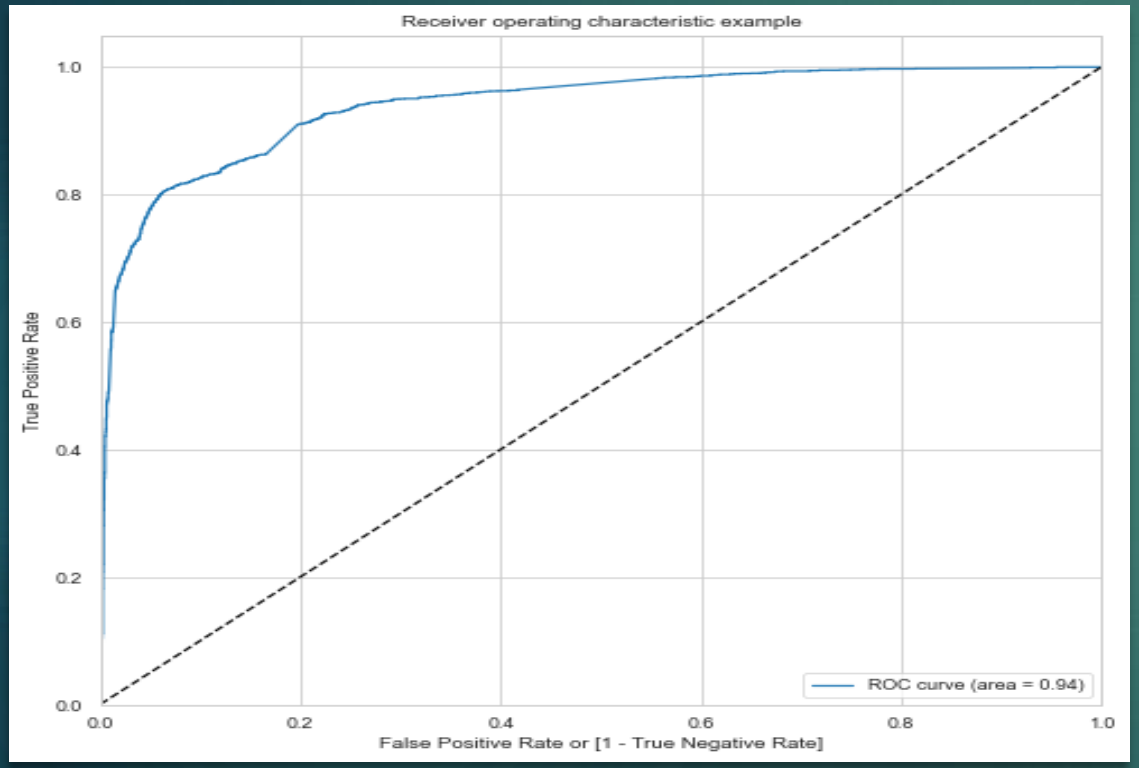
Multivariate analysis

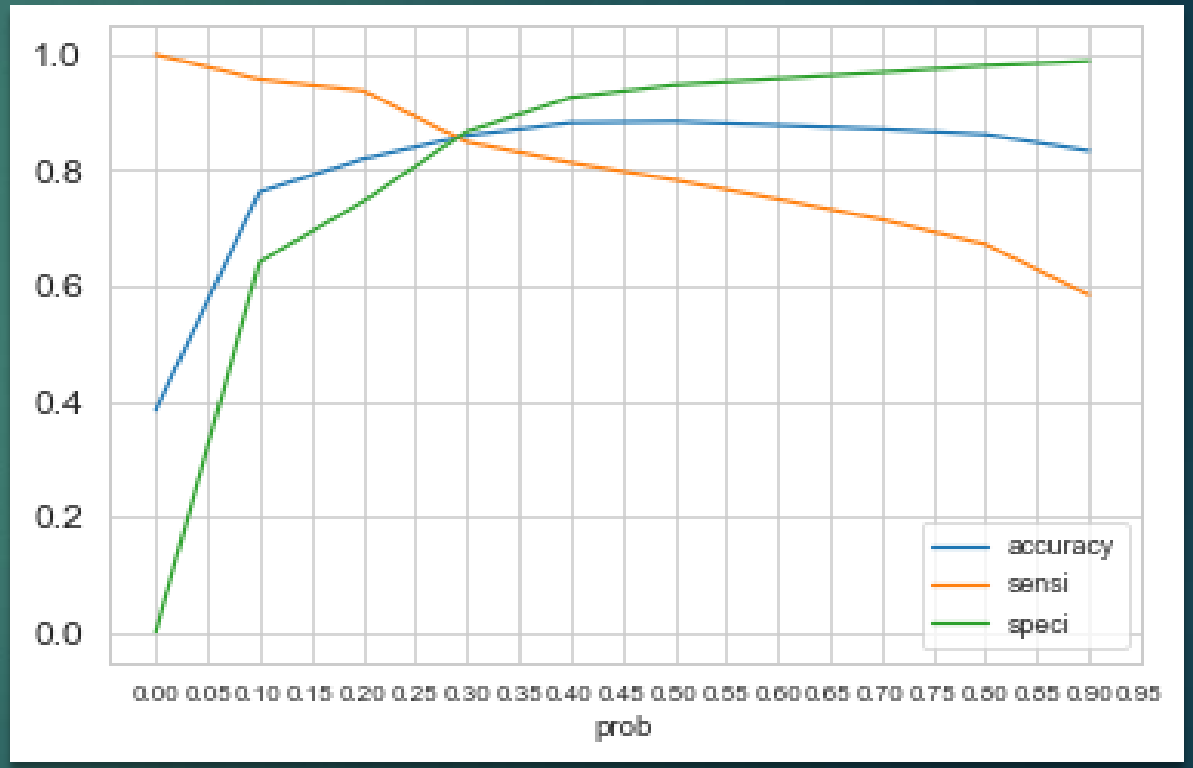
Multivariate analysis



Correlations for numeric variables

Optimize Cut off (ROC Curve)





Precision-Recall

Precision and recall tradeoff