

Lead Scoring Case Study

Problem Statement

- X education sells online courses to industry professionals
- Their lead conversion rate is very low
- To make the process effective, they wish to identify potentially convertible leads, also known as 'Hot Leads'
- Tangible output/ target to be achieved: increase in lead conversion rate for X Education

Methodology

Data Cleaning

- ✓ Removal of Duplicate Data
- ✓ Handling NaN and missing values
- ✓ Imputing values, if necessary
- ✓ Drop columns with lot of missing values

EDA

- ✓ Univariate Analysis: Normalized Value Counts, Bar Charts
- ✓ Bivariate Analysis: Correlation coefficients

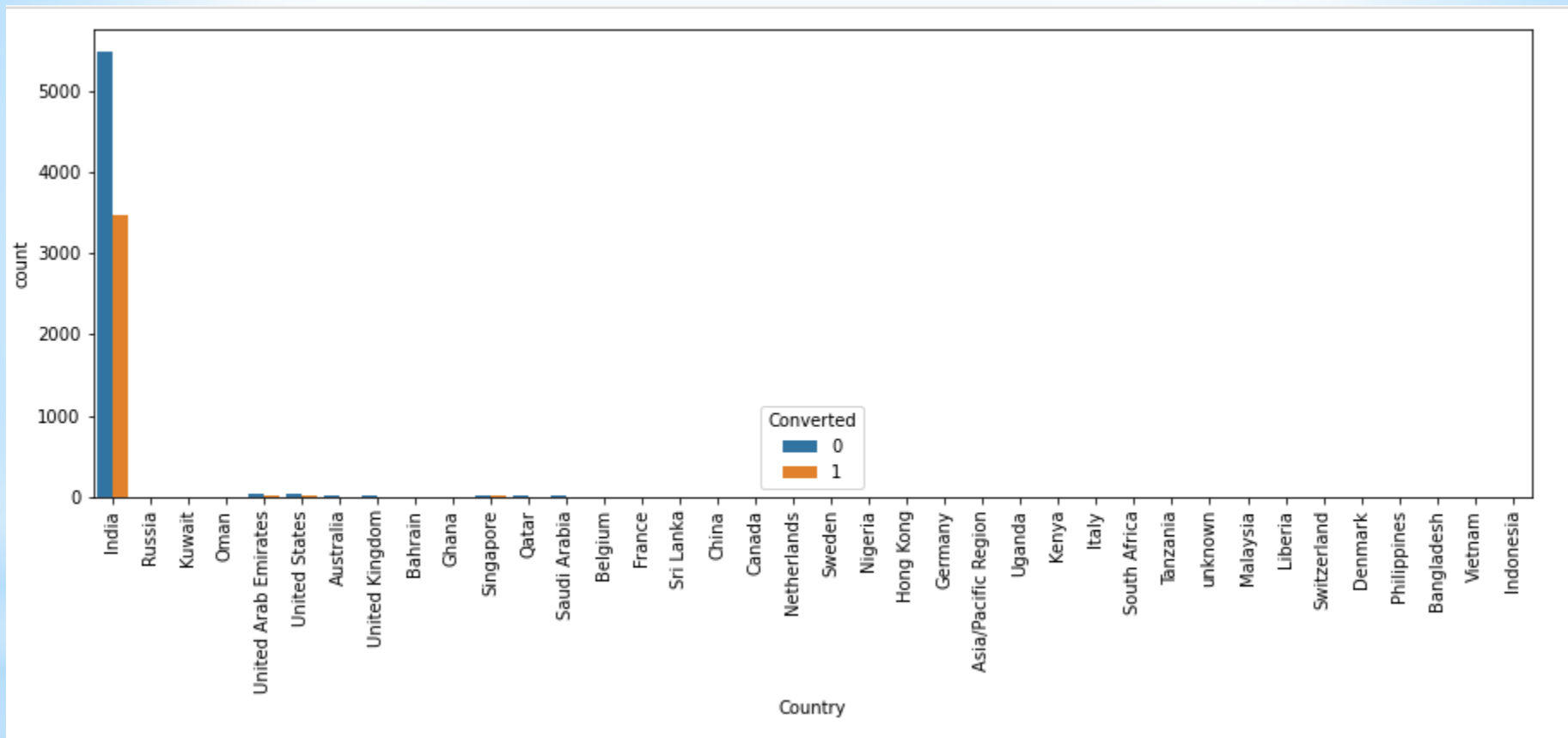
Contt.

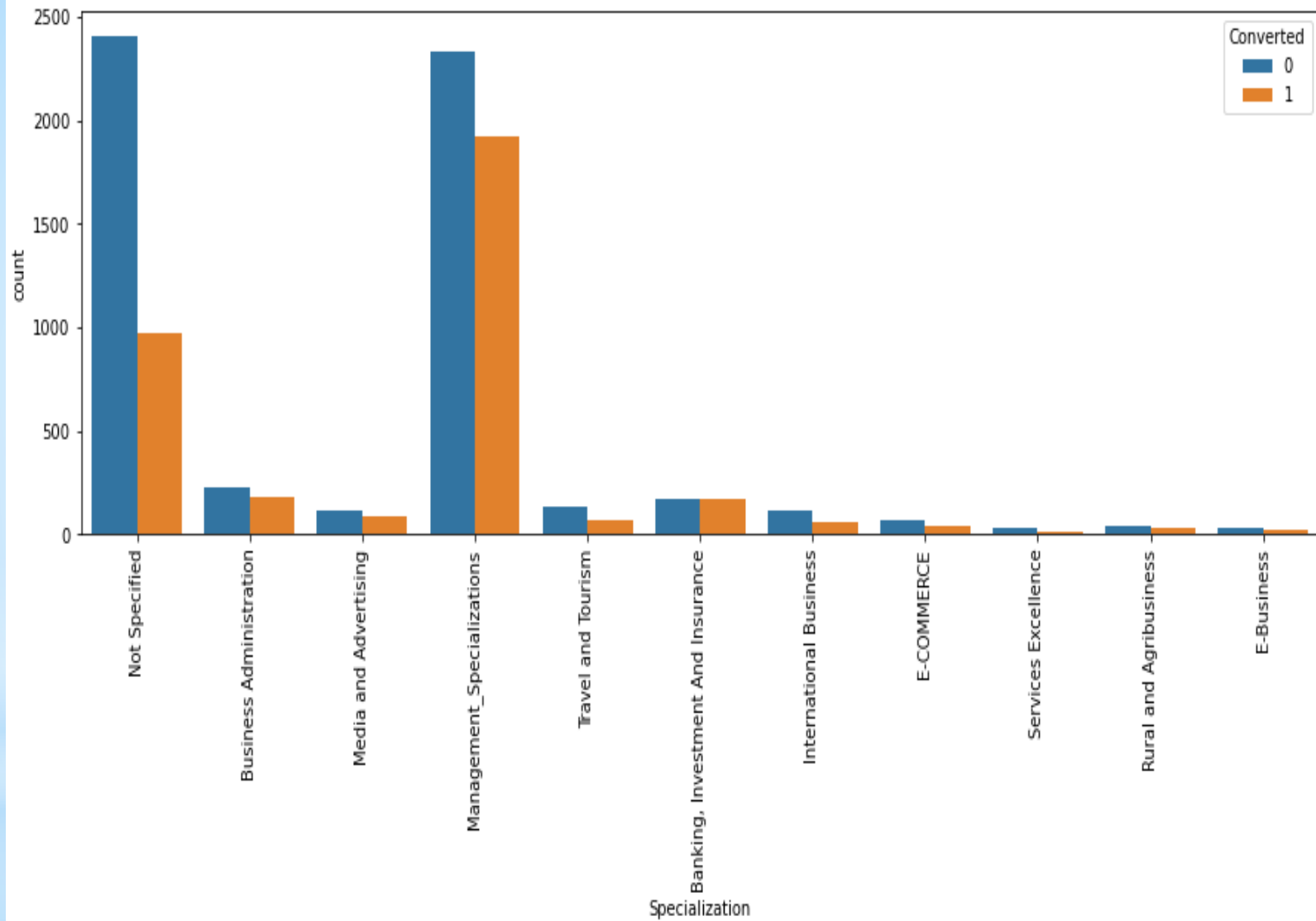
- ✓ Scaling and Dummy Encoding
- ✓ Feature selection using RFE and manual method combined
- ✓ Building Logistic Regression Model
- ✓ Model Evaluation and Predictions

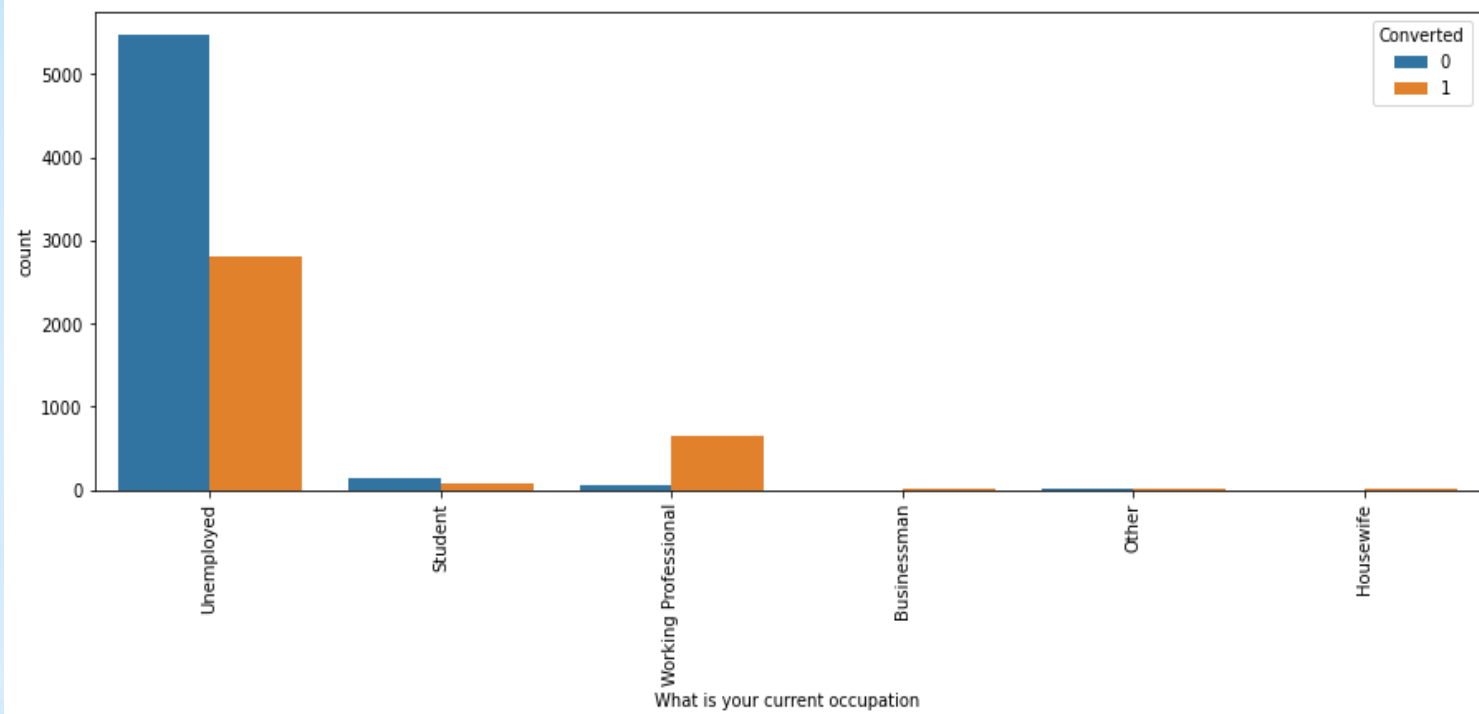
Data Understanding

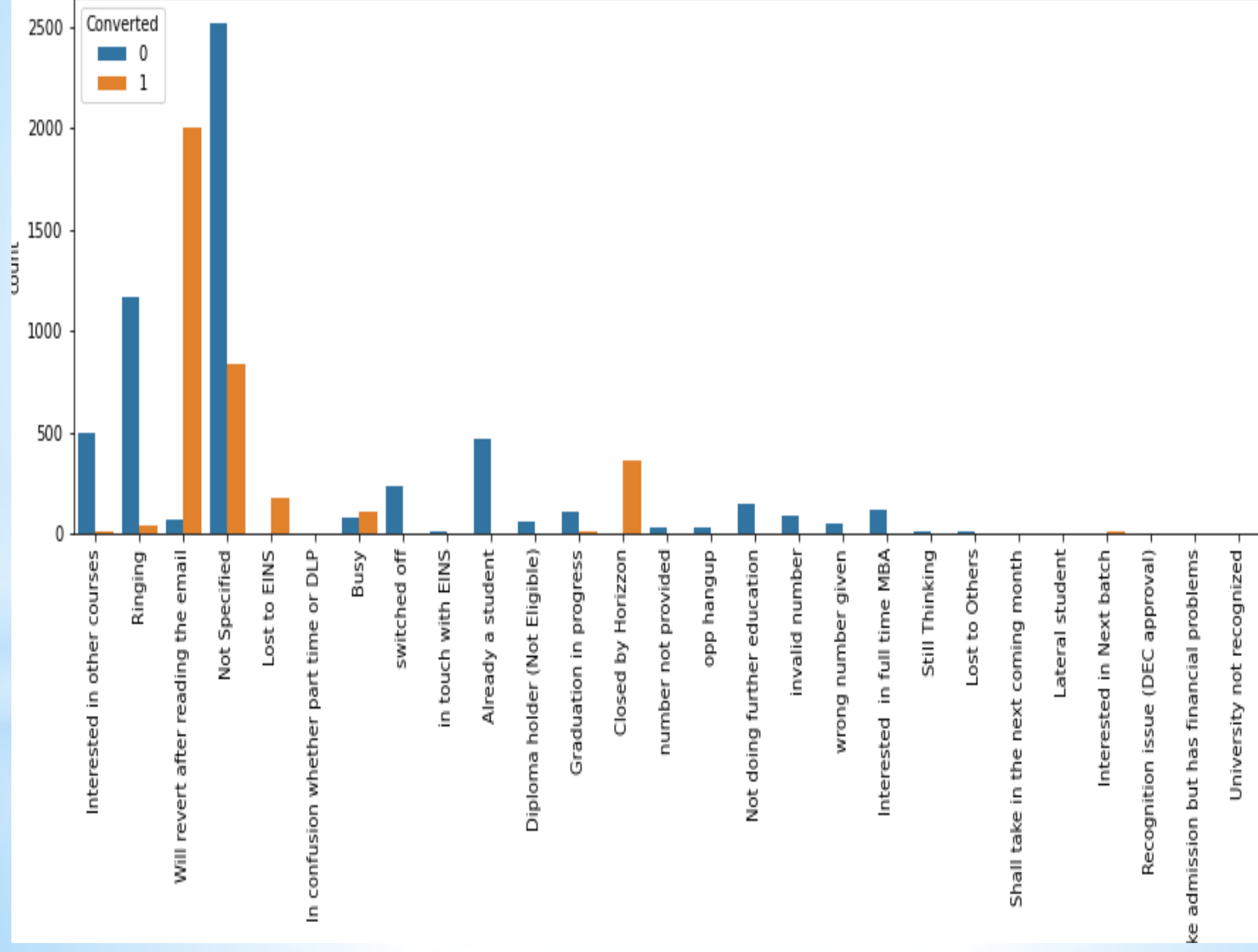
- ✓ Shape of Data 9240 X 37
- ✓ Single Value Features like “Magazine”, “Update me on supply chain content”, “Get Updates on DM Content”etc. were dropped
- ✓ “Prospect ID” and “Lead Number” were dropped for they were not useful in analysis.
- ✓ Variables having large skew were dropped because of class imbalance ex. Country

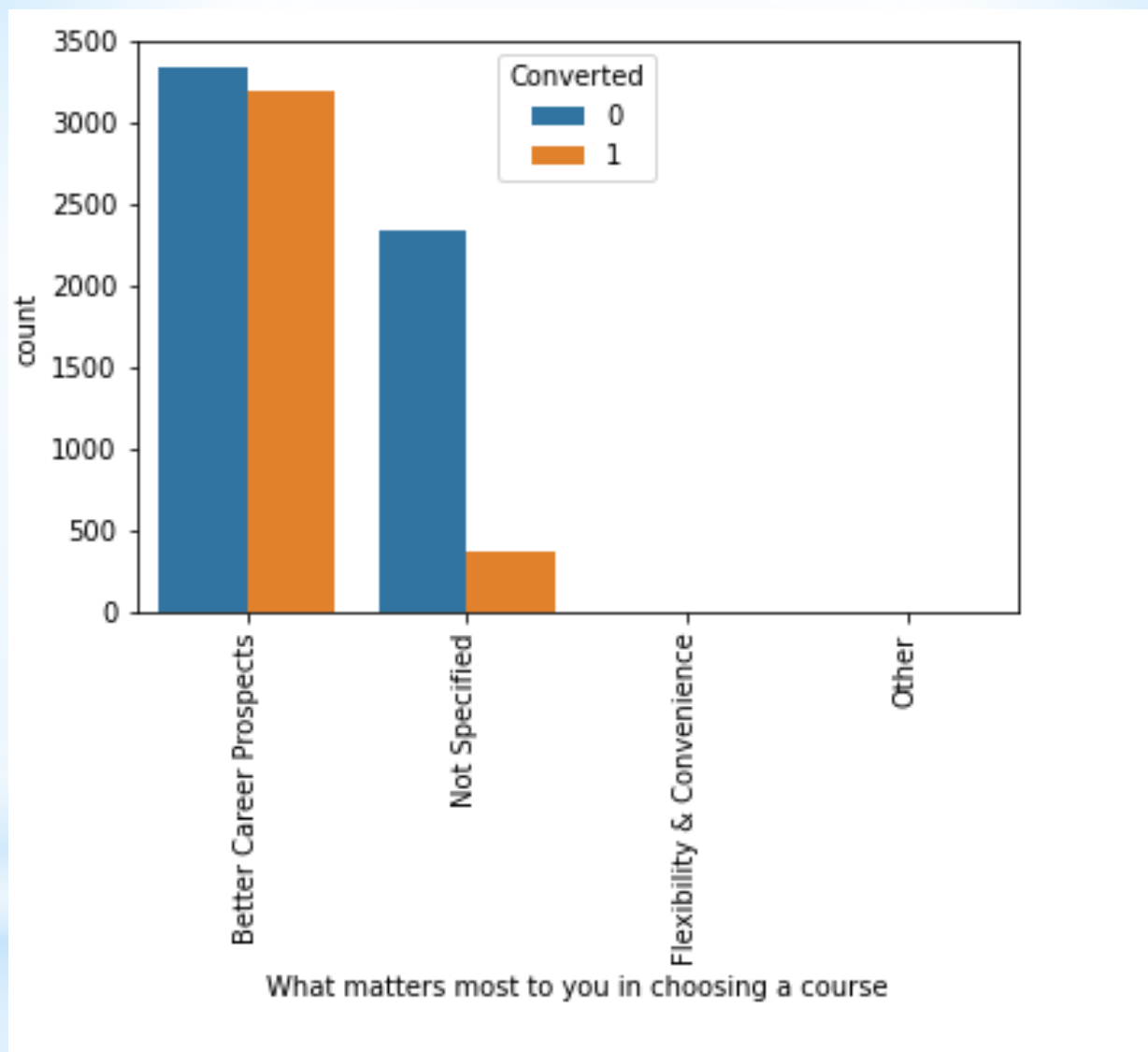
EDA

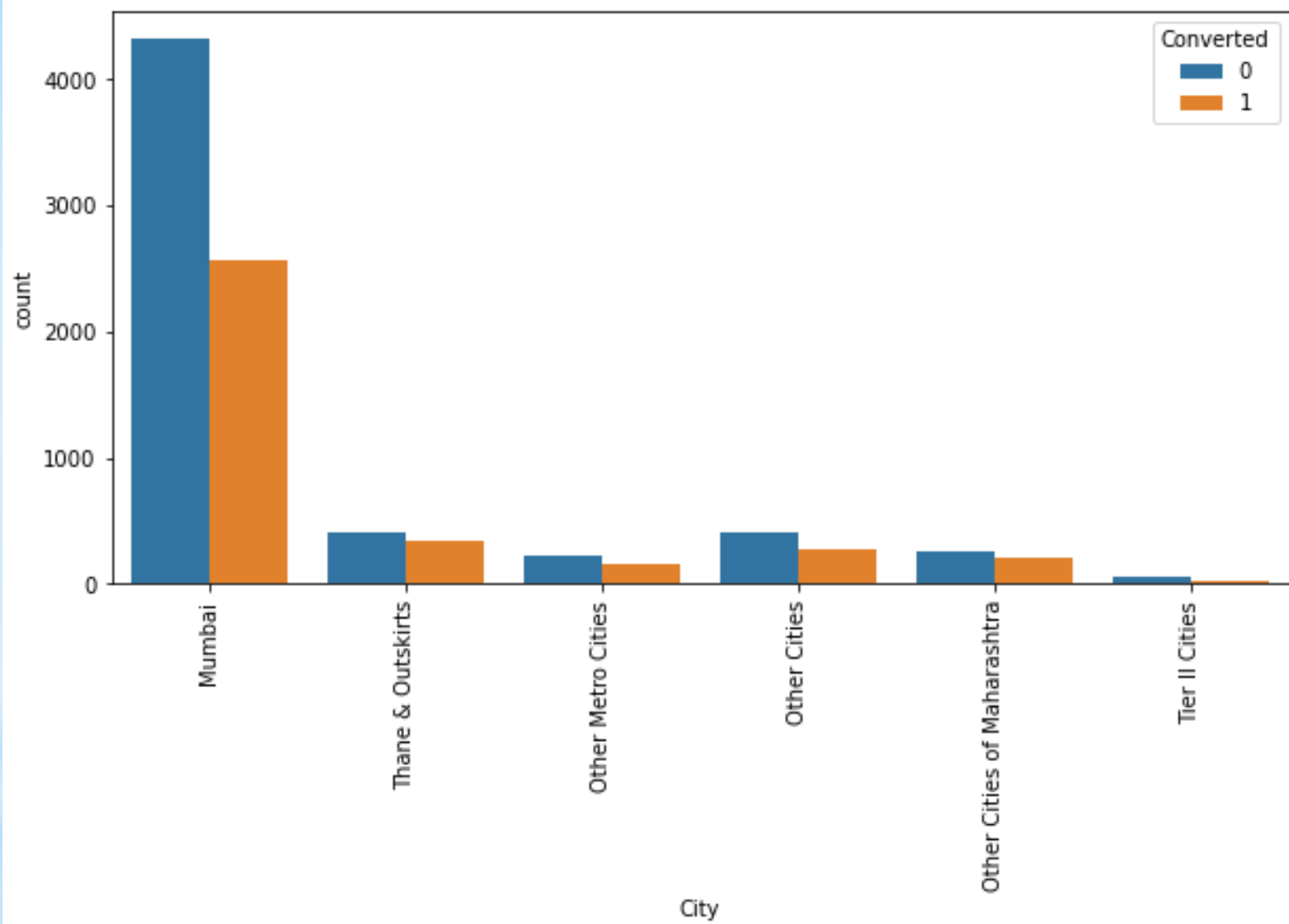


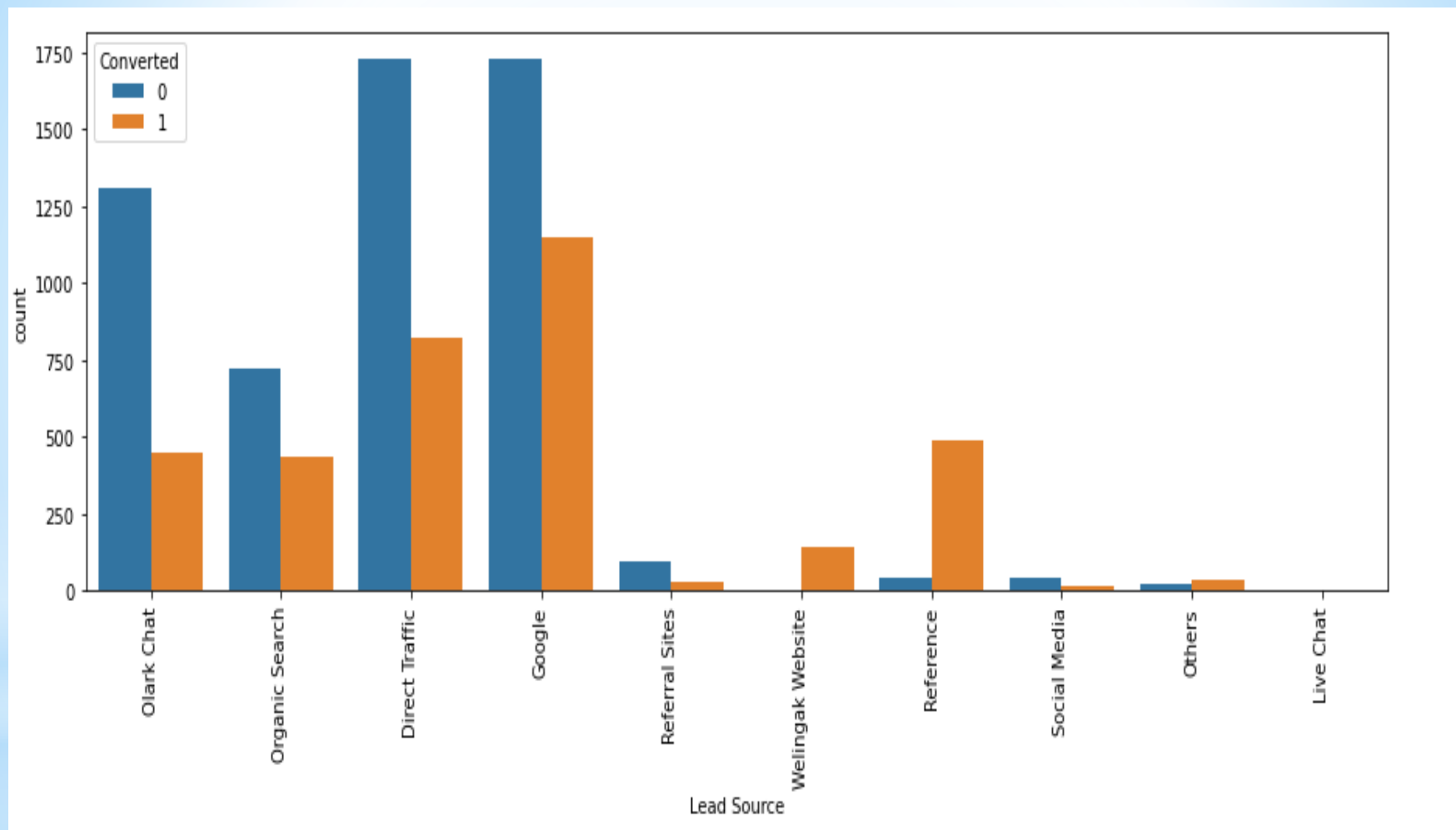


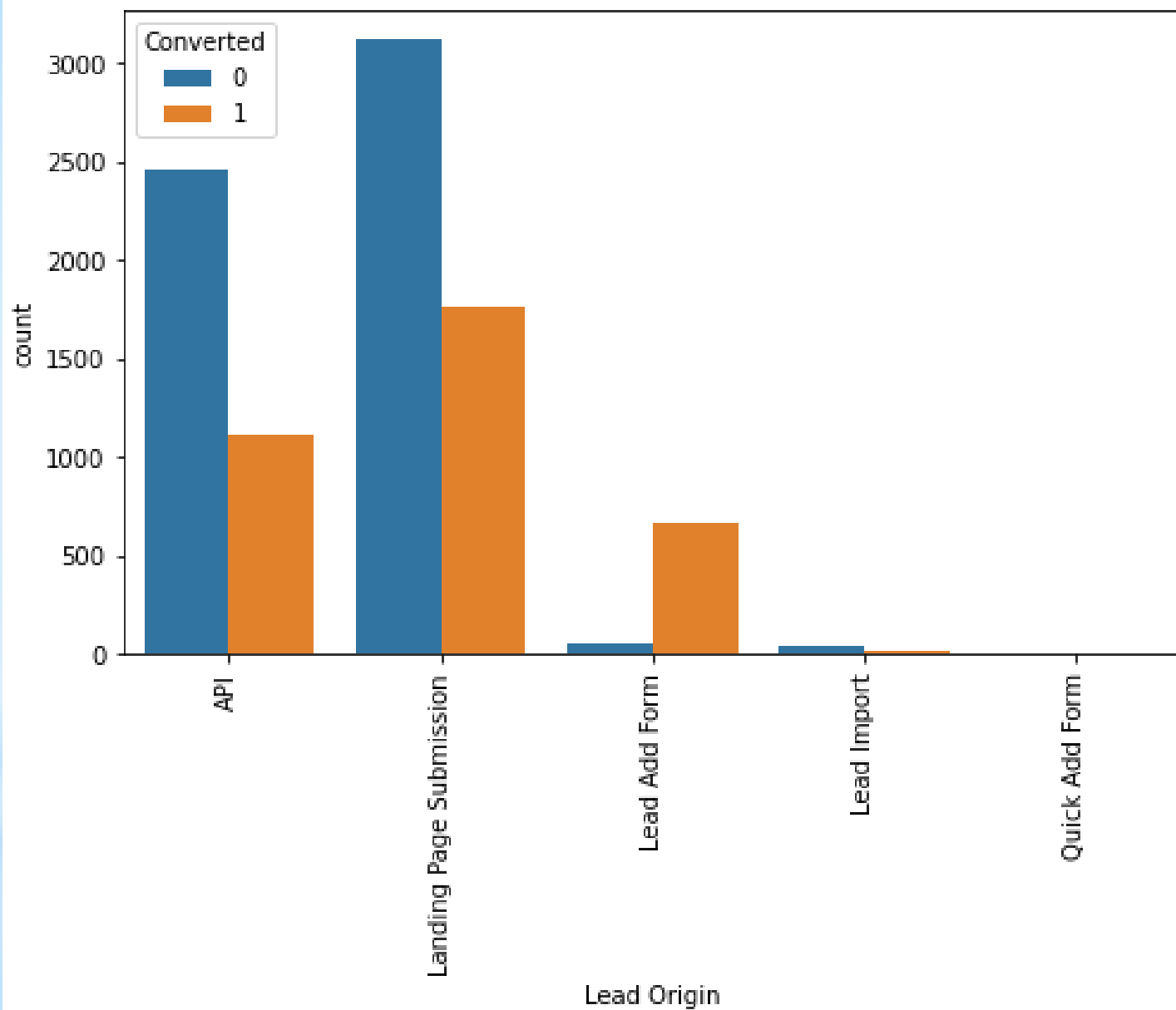




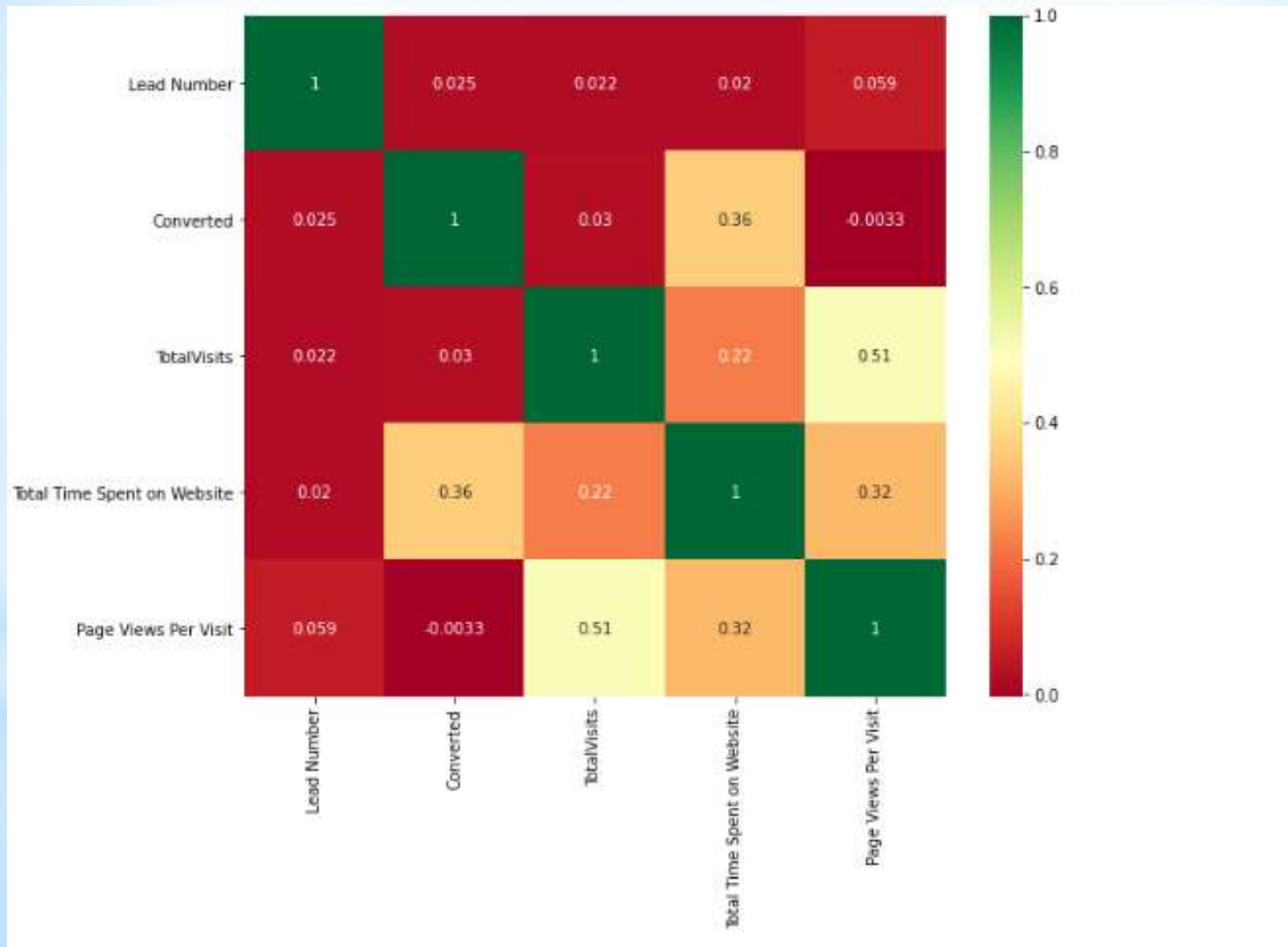








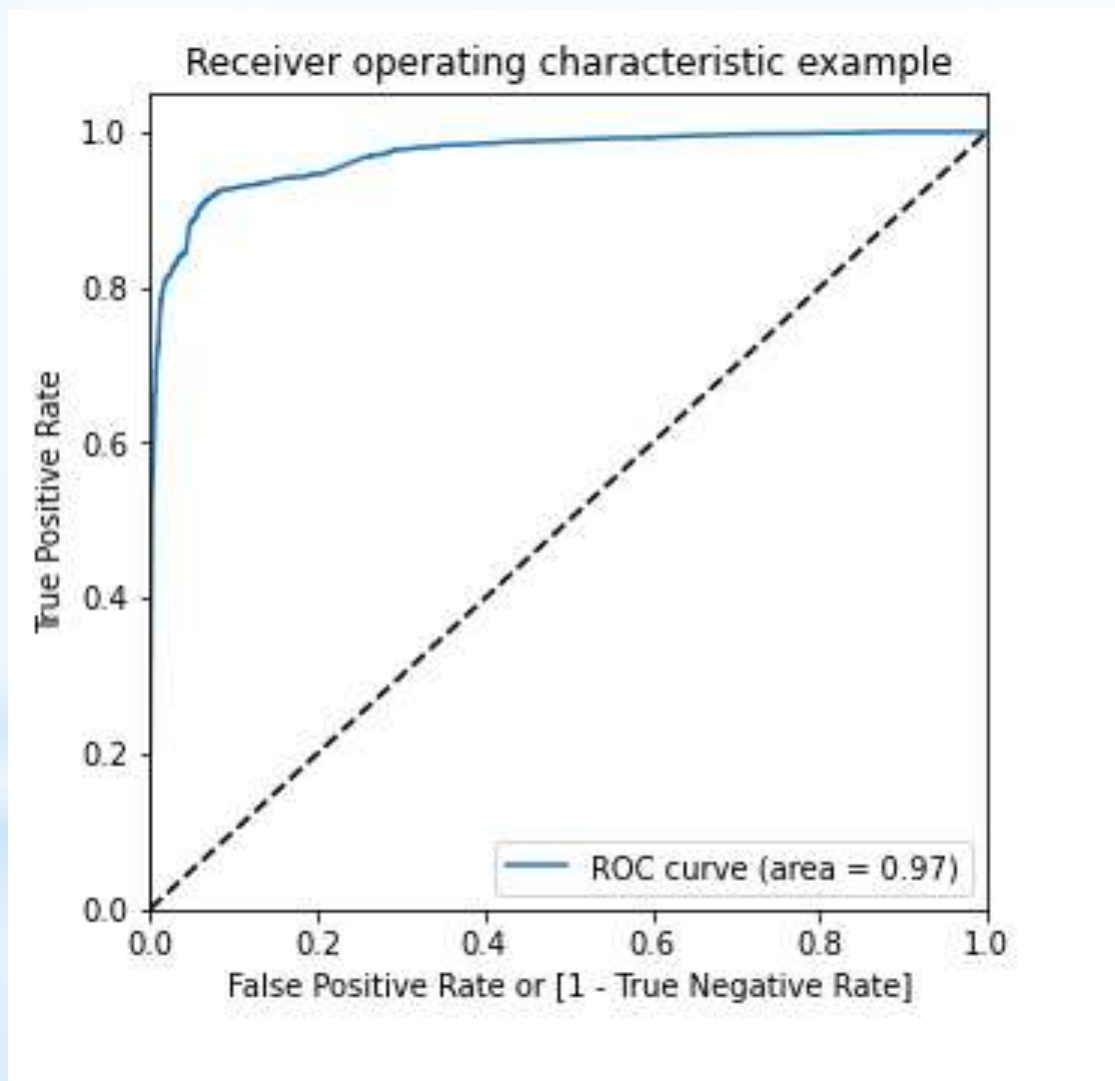
Heat map of numerical variables



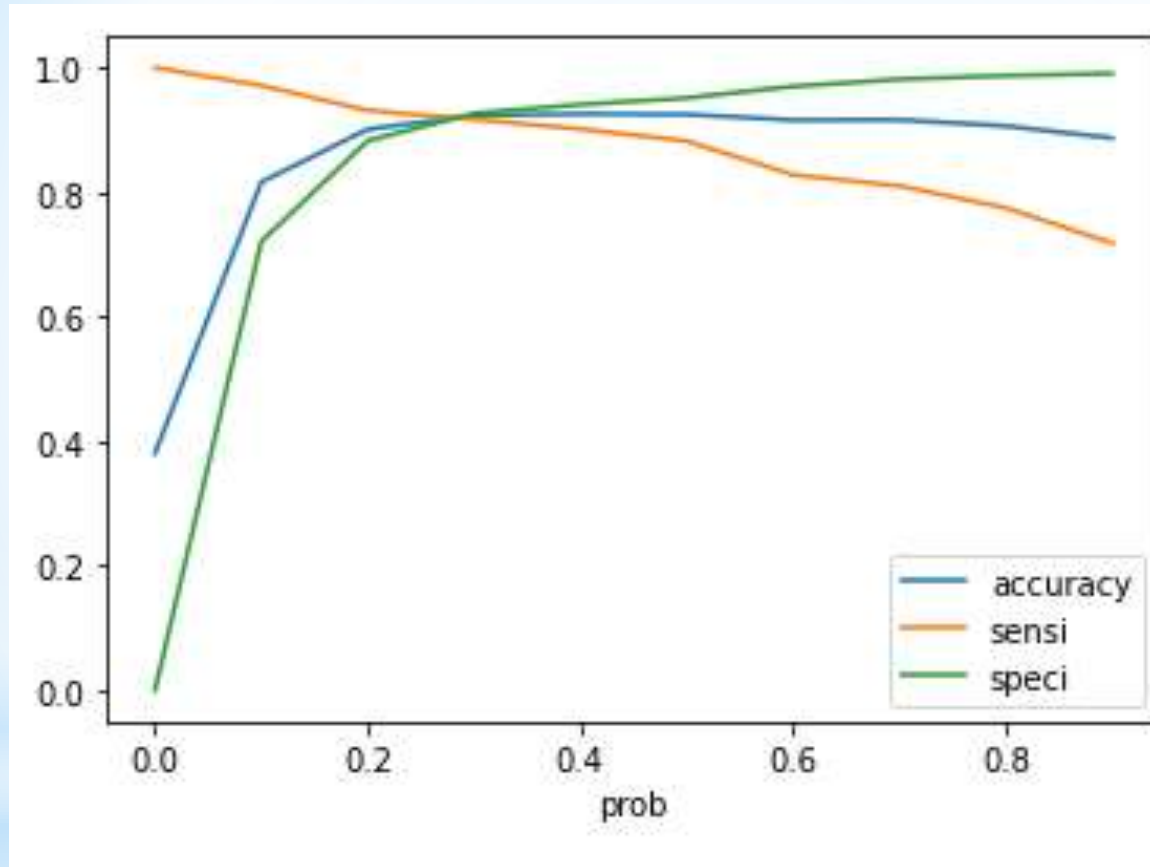
Model Building

- ✓ Splitting data into test & train set
- ✓ Feature scaling for better comparability
- ✓ Feature Selection using RFE with 15 features
- ✓ Model building by eliminating variables with p value > 0.05 and $VIF > 5$
- ✓ Predictions on test data
- ✓ Accuracy attained : 92.55%

ROC Curve



Optimum Cut Off Point



From the curve, optimum cutoff point is the one that balances sensitivity and specificity , which is 0.3 in the above case

Conclusion

We observed that the following variables were the most important for lead conversion

- ✓ Total Time Spent on the Website
- ✓ Lead Origin
 - Lead Add Form
- ✓ Lead Sources
 - Direct Traffic
 - Welingak Website
- ✓ Last Activity
 - SMS Sent
- ✓ Last Notable Activity
 - Modified
 - Olark Chat Conversation

Conttd.

✓ Tags

Closed by Horizzon

Interested in other courses

Lost to EINS

Others

Ringling

Will revert after reading the email