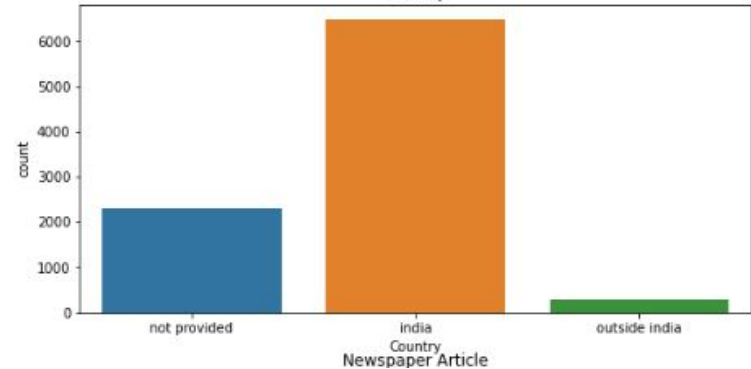
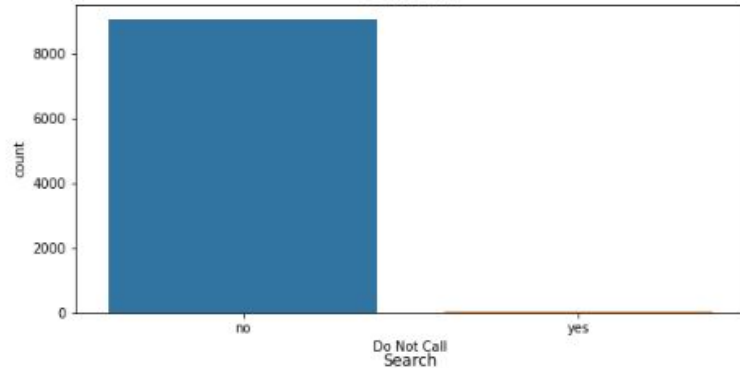
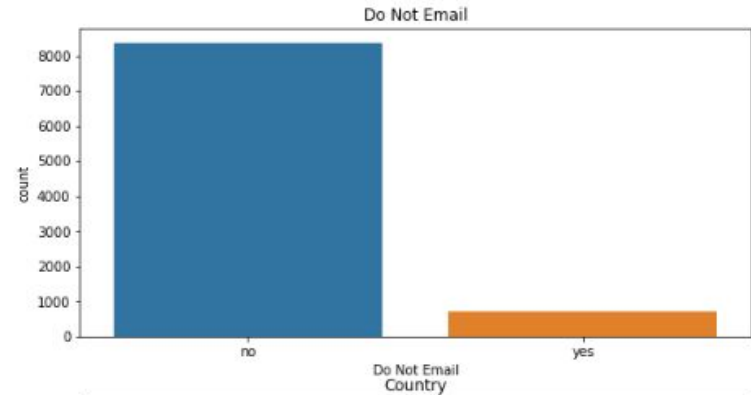
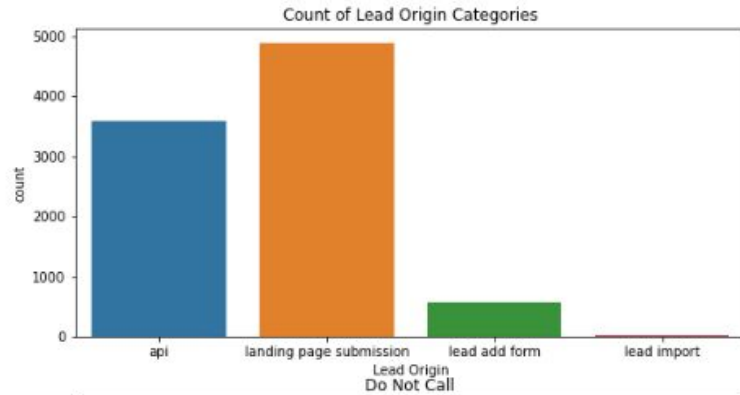


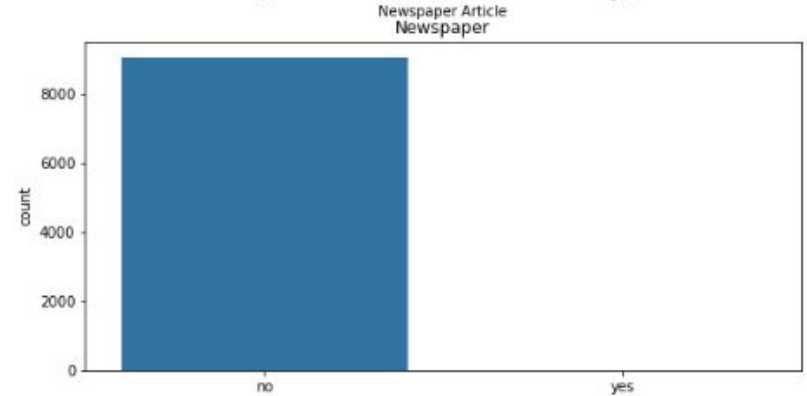
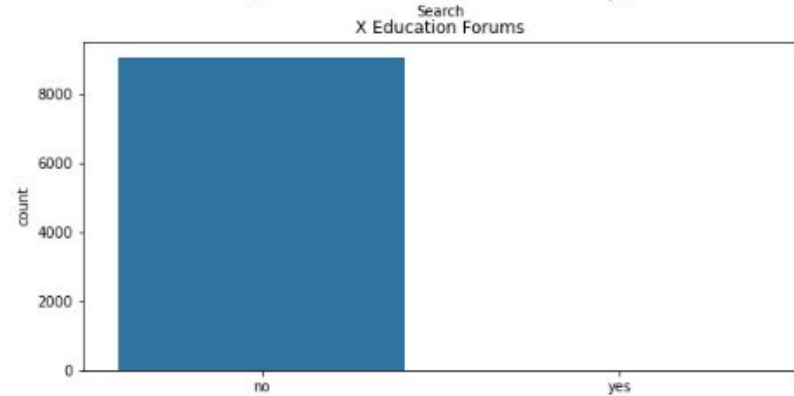
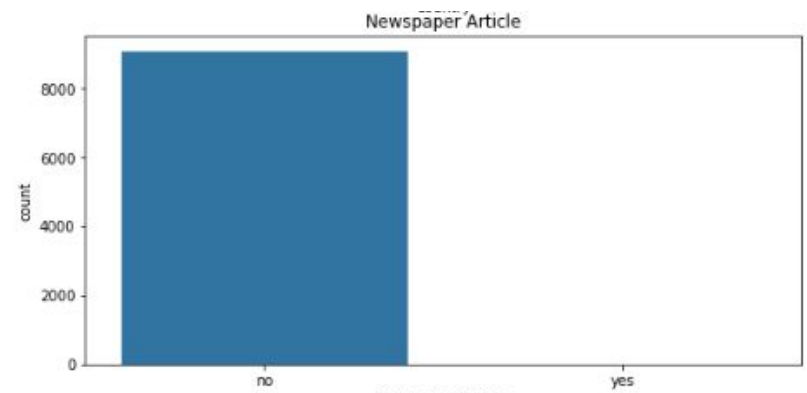
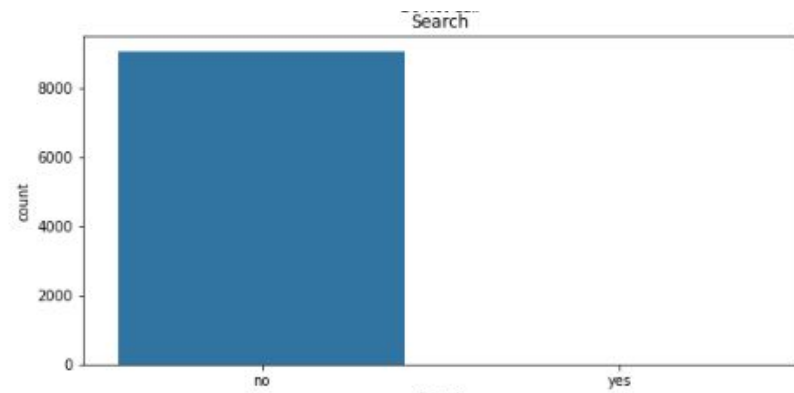
Lead Scoring Case Study

Bhargav D

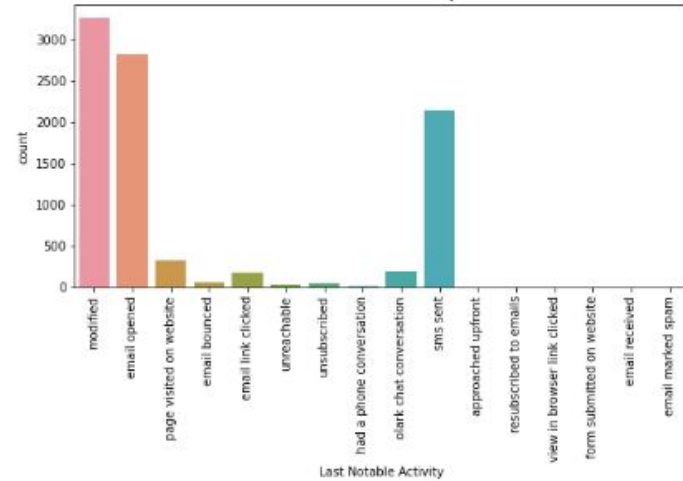
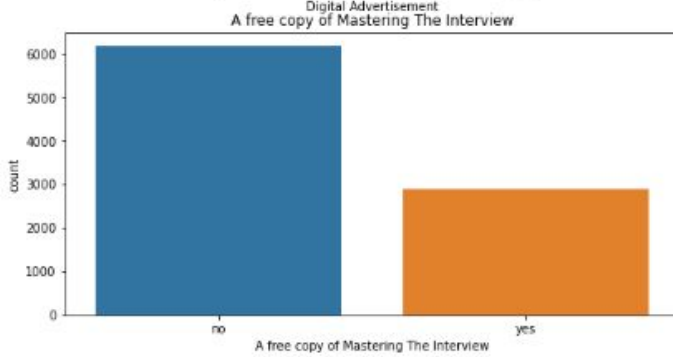
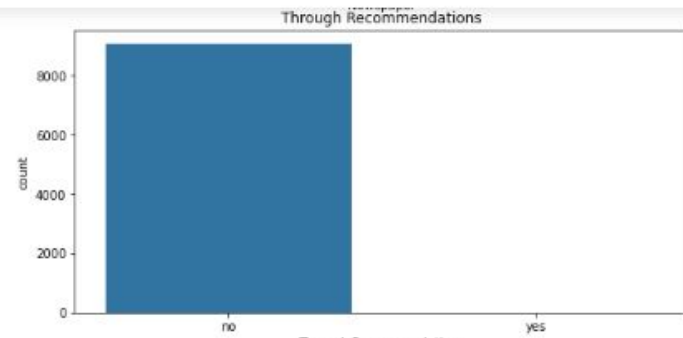
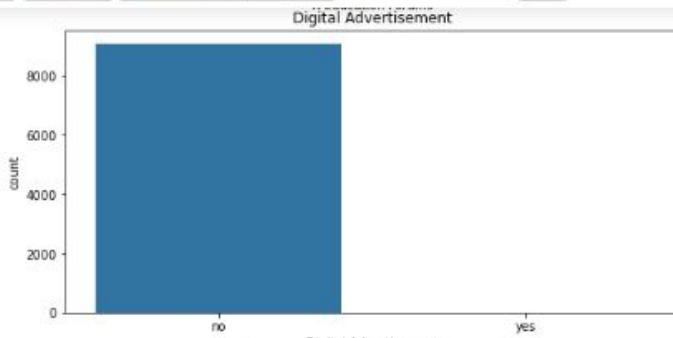
CATEGORICAL UNIVARIATE ANALYSIS



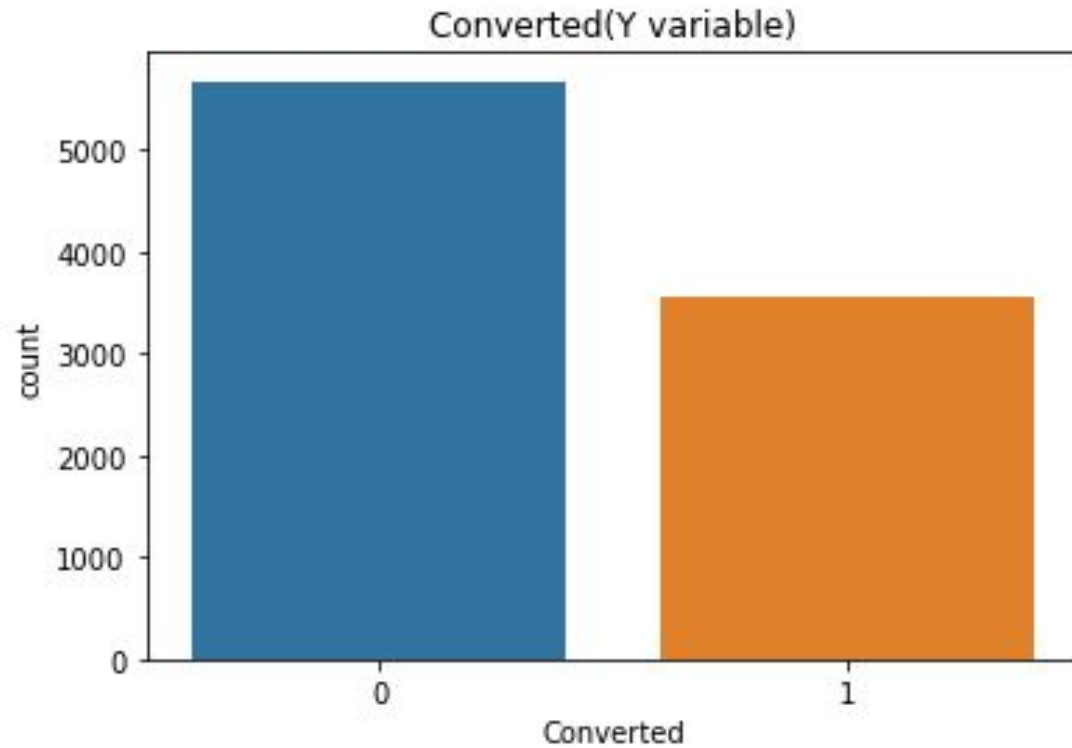
- Visualizing the count distribution of Lead Origin, Do Not Email, Do Not Call, Country.



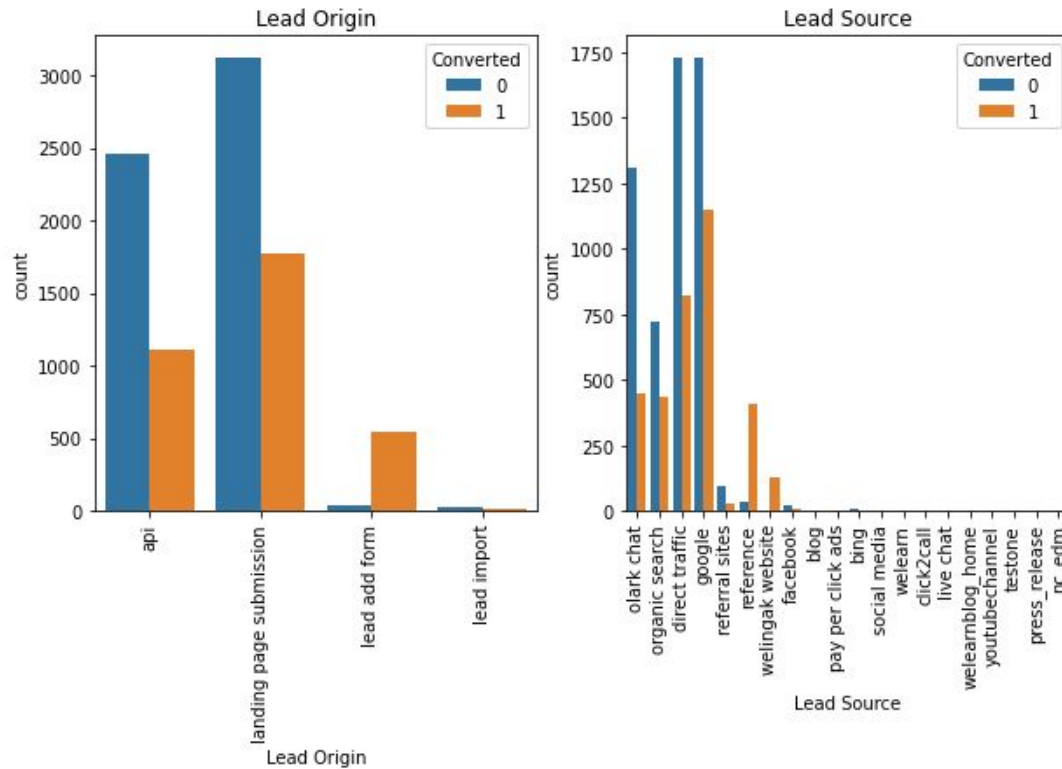
- Visualizing the count distribution of Search, Newspaper Article, X Education Forums, Newspaper.



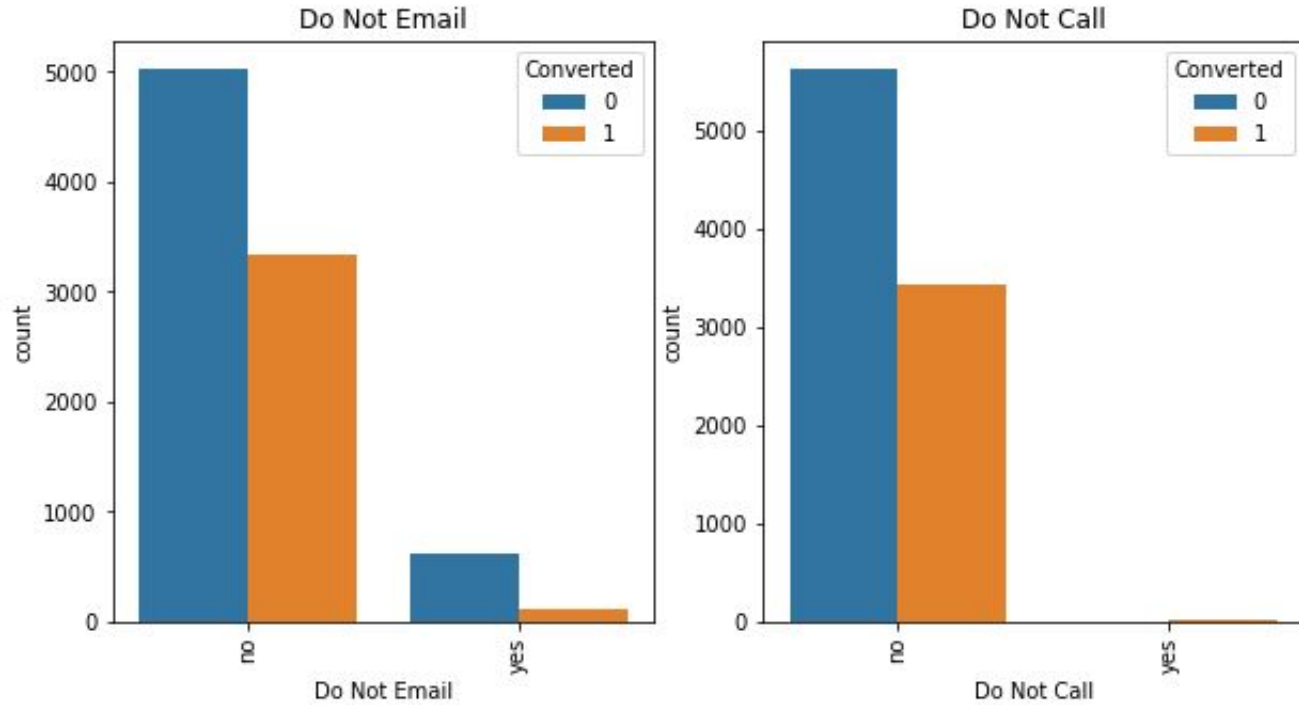
- Visualizing the count distribution of Digital Advertisement, Through Recommendations, A Free Copy of Mastering The Interview, Last Notable Activity.



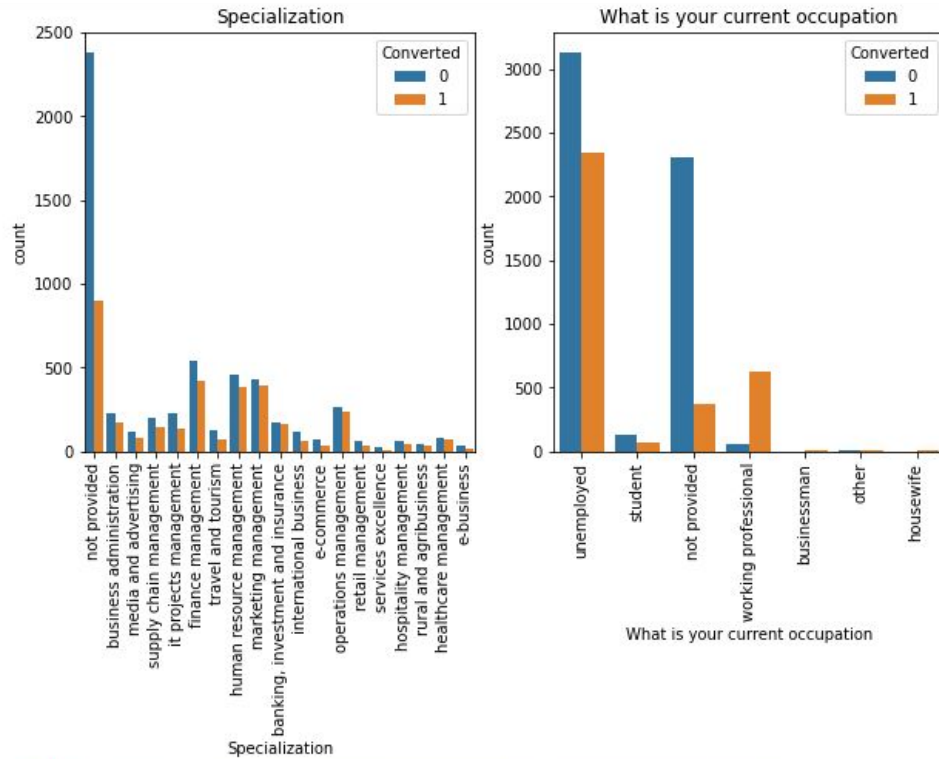
- Countplot of converted Y variable with the data df



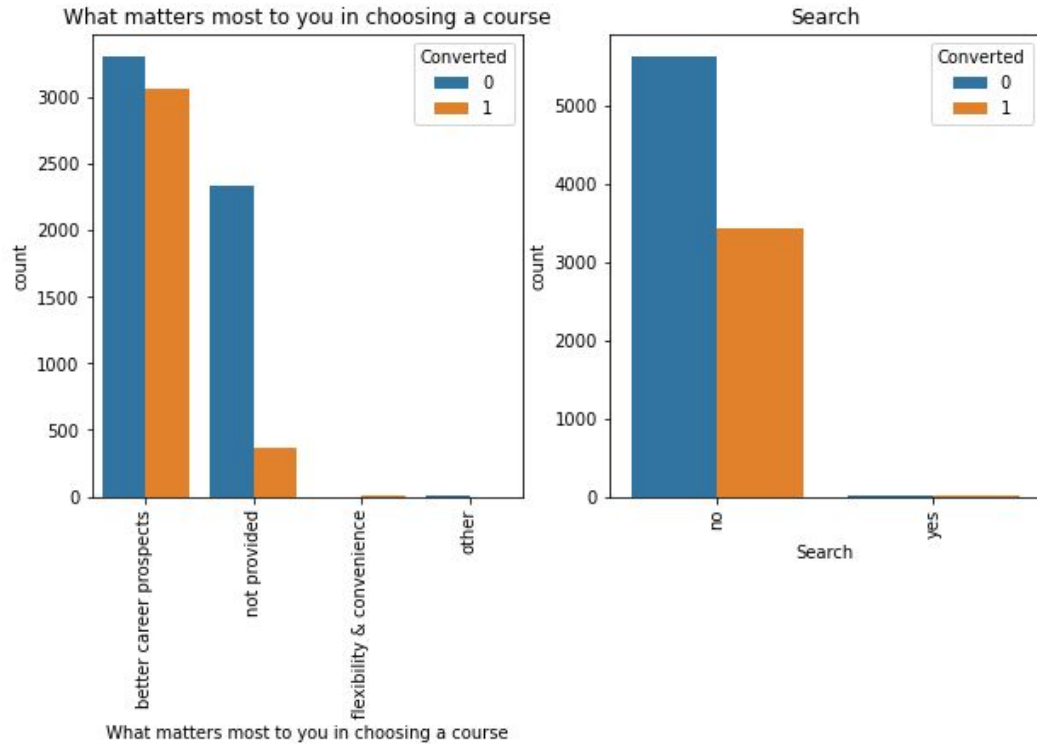
- Count Plot of Relating variables Lead Origin and Lead Source to converted variable



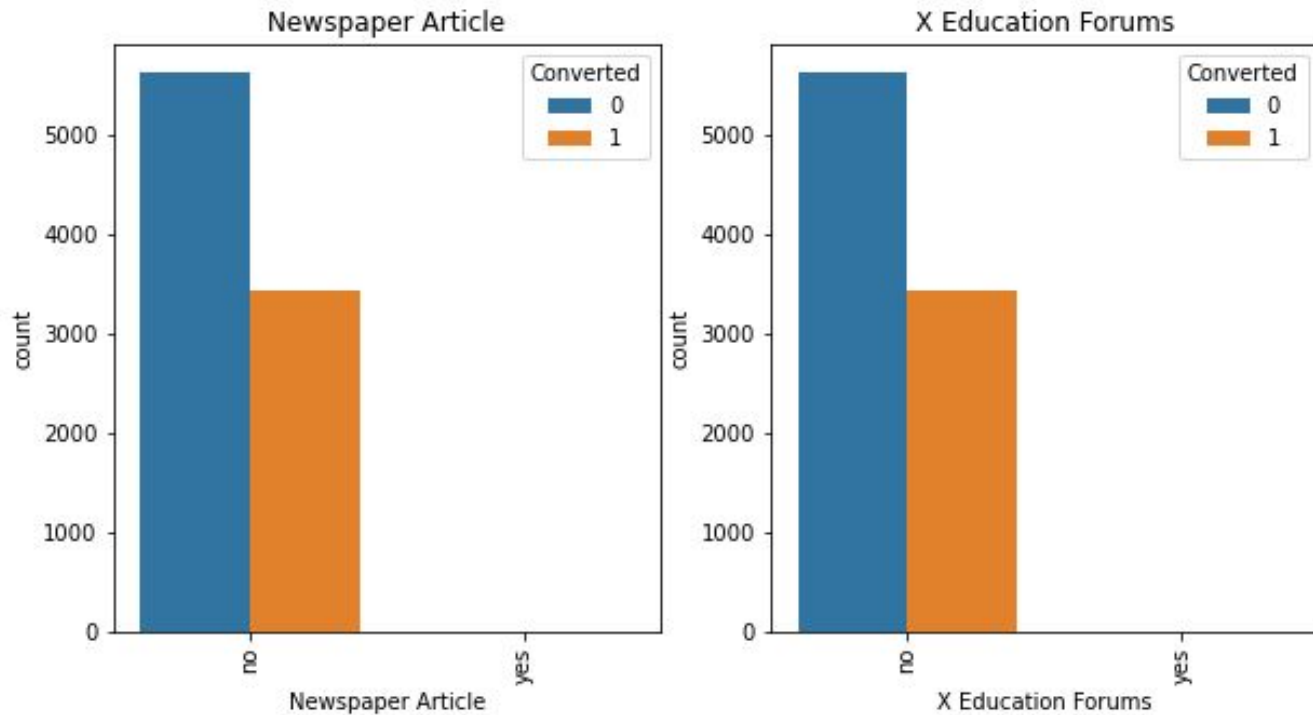
- Count Plot of Relating variables Do Not Email and Do Not Call to converted variable



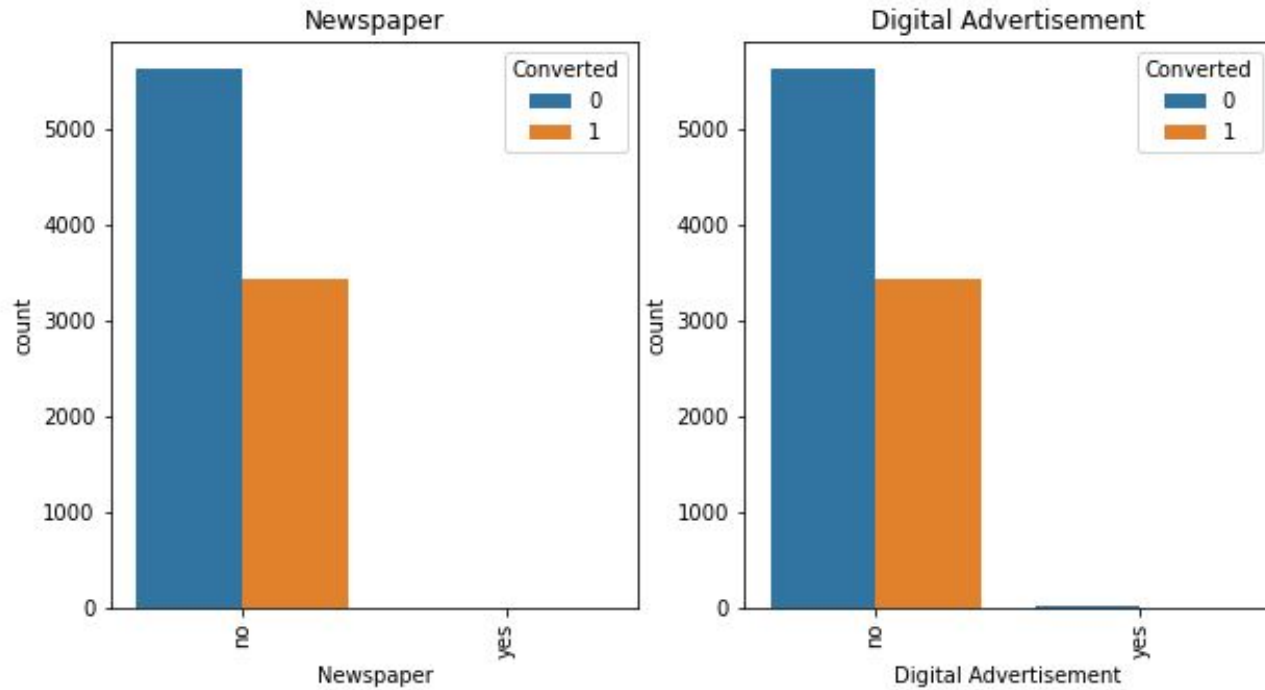
- Count Plot of Relating variables Specialization and What is your current occupation to converted variable



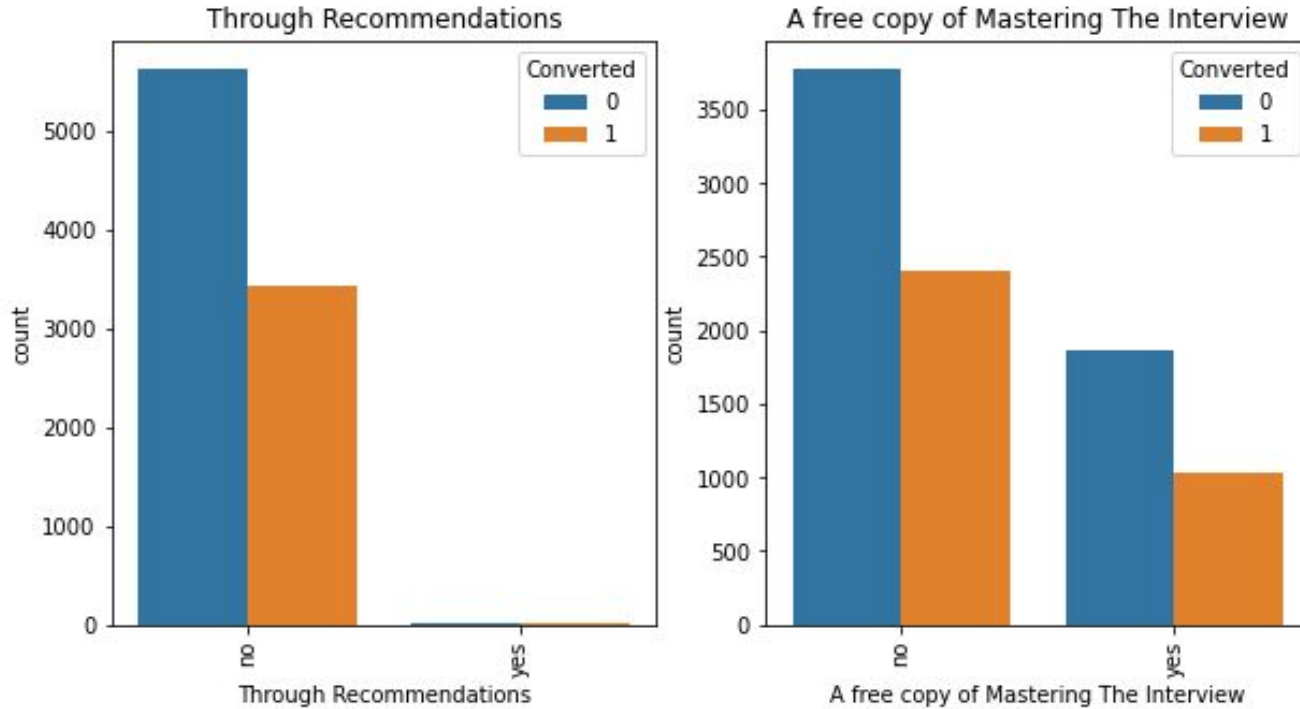
- Count Plot of Relating variables What matters most to you in choosing a course and search to converted variable



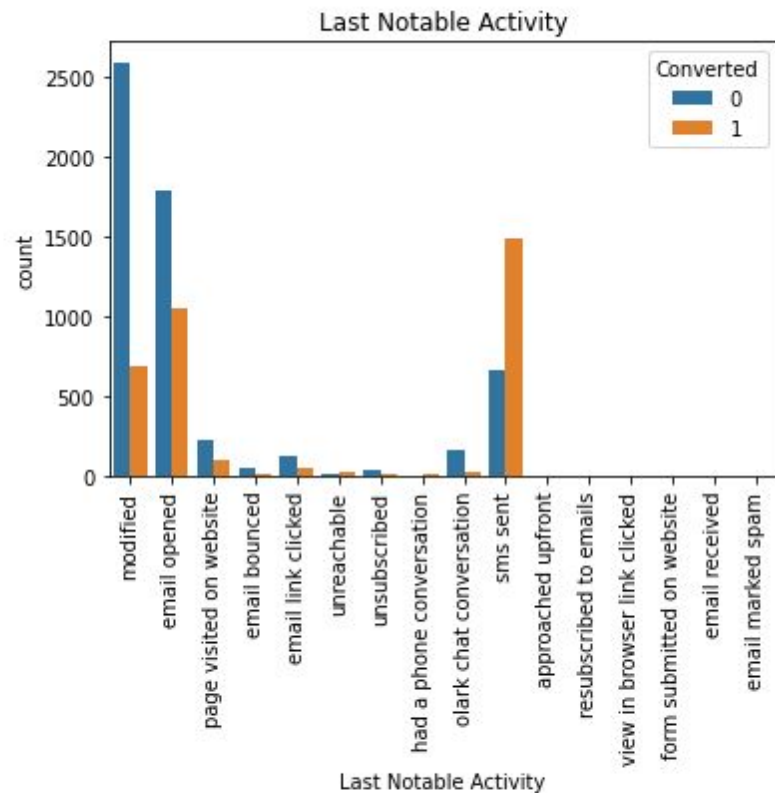
- **Count Plot of Relating variables Newspaper Article and X Education Forums to converted variable**



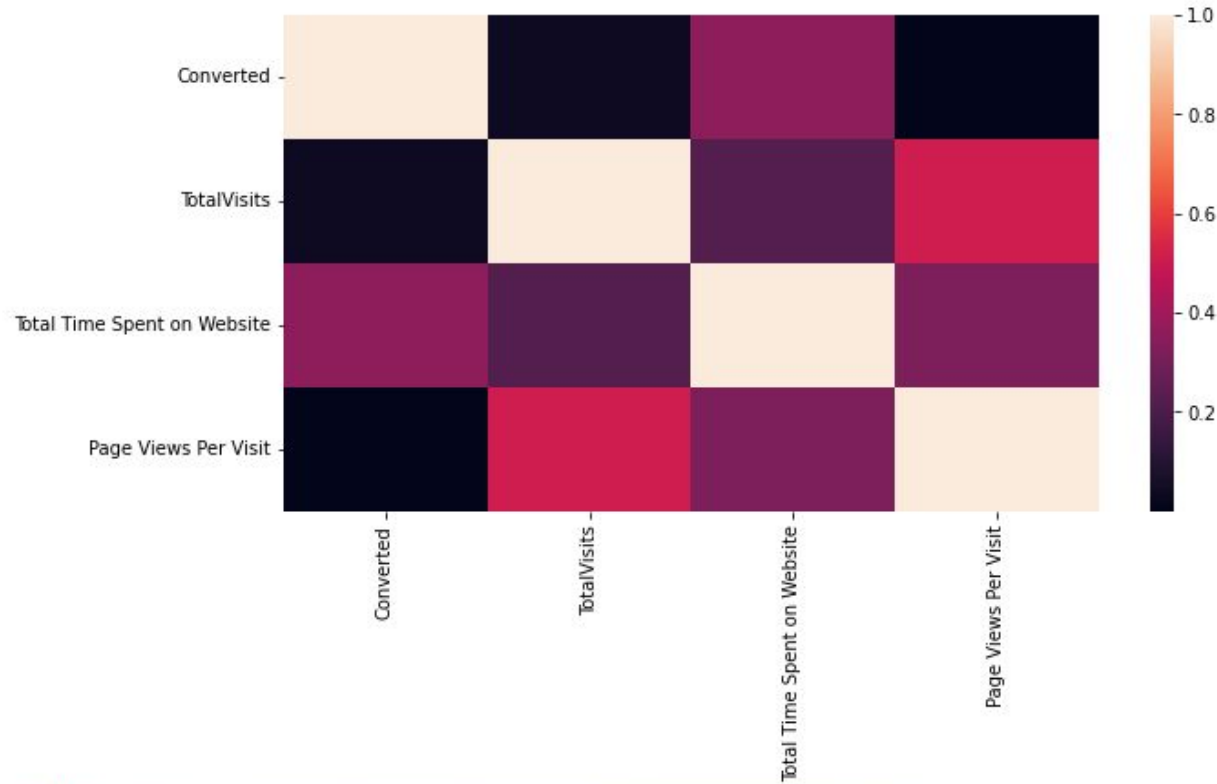
- **Count Plot of Relating variables Newspaper and Digital Advertisement to converted variable**



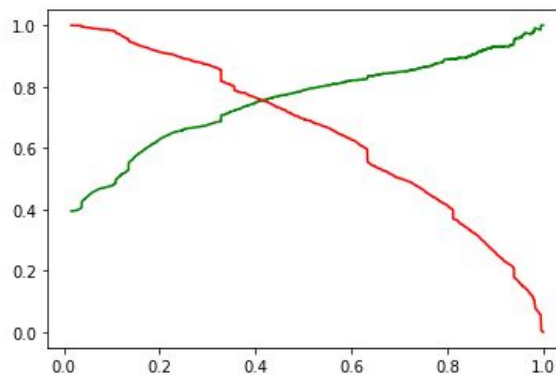
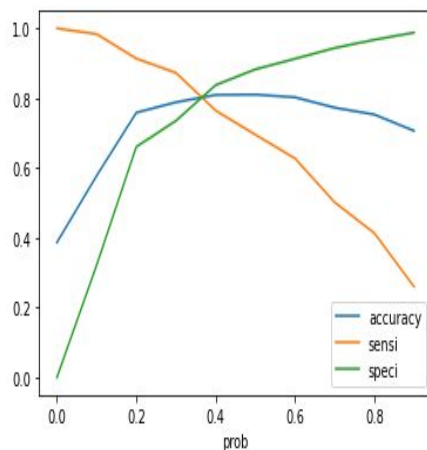
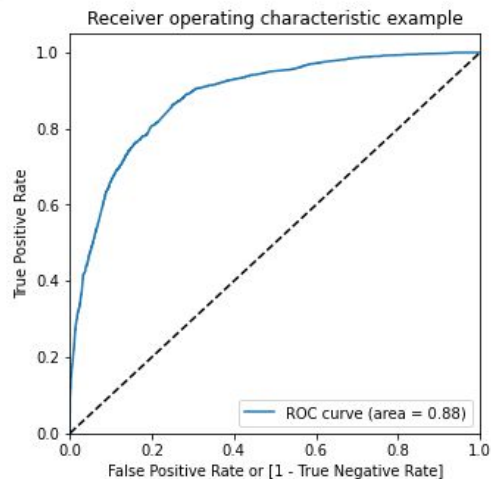
- Count Plot of Relating variables Through Recommendations and A free copy of mastering the interview to converted variable.



- Count Plot of Relating Last Notable Activity to converted variable



- The correlation of all selected numerical columns
- It is understandable from the above EDA that there are many elements that have very little data and so will be of less relevance to our analysis.



- **Linear Regression final model parameters under**
 - **ROC = 0.84**
 - **Intermediate cut-off = 0.35**
 - **final cut-off = 0.42**

Conclusion :

It was found that the variables that mattered the most in the potential buyers are (In descending order) :

1. The total time spent on the Website.
 2. Total number of visits.
 3. When the lead source was:
 - a. Google
 - b. Direct traffic
 - c. Organic search
 - d. Welingak website
 4. When the last activity was:
 - a. SMS
 - b. Olark chat conversation
 5. When the lead origin is Lead add format.
 6. When their current occupation is as a working professional.
- 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.