Lead Score Assignment

29/12/2022

Done By:

Apoorva Devarmani

Prashanth Kumar

Shabbir Abbas

Problem Statement

- X Education, which is a company that sells online courses, want to increase their lead conversion rate.

- They want to identify the most potential leads which is also known as the hot leads.

- This is so that they can concentrate on these leads and create more sales conversions.

Solution Steps

- First checked and handled duplicate data. Also handled null and "NA" values.

- Dropped the unnecessary columns and imputed values where necessary.

- Handled outliers in the data.

- Then we performed EDA and did validation of the model.

- Then we arrived with conclusions and recommendations.

EDA
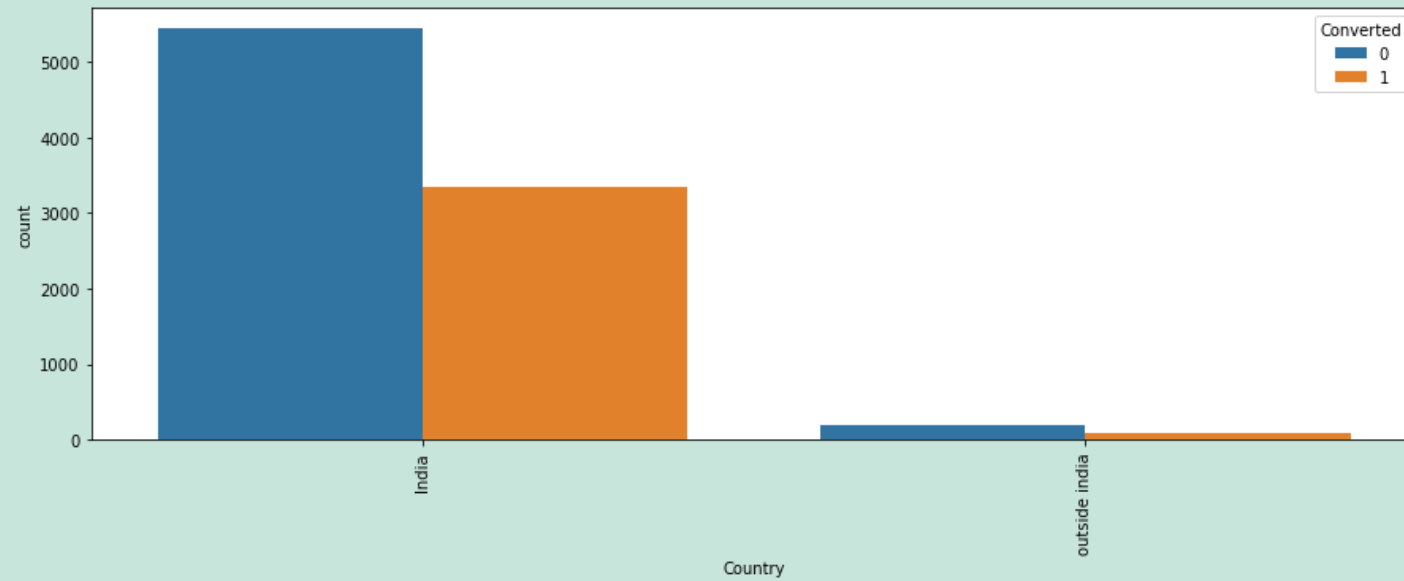
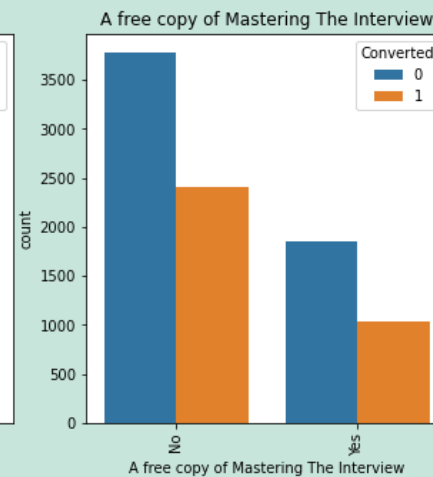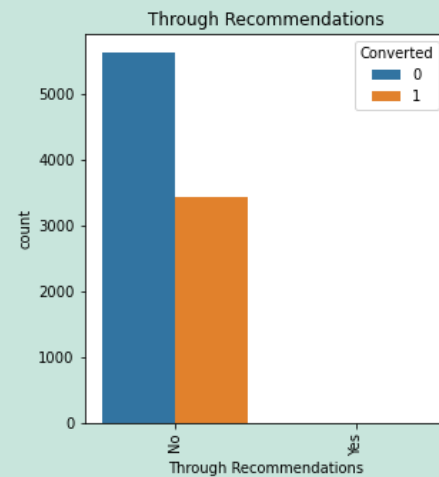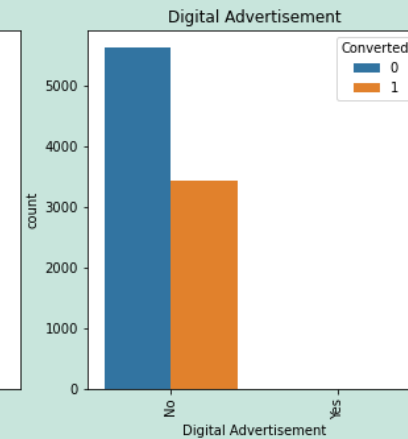◈ We can see that most of the data points are from India

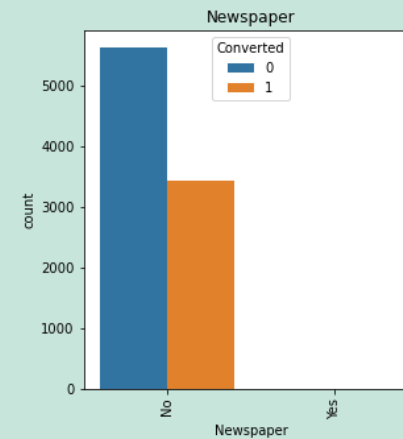# EDA
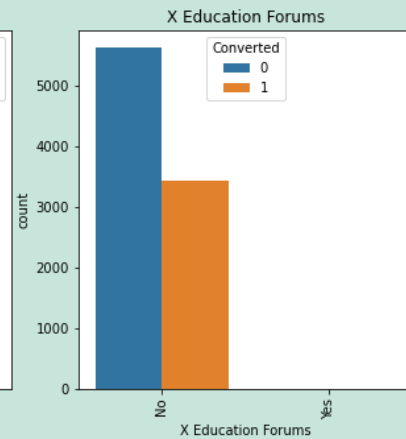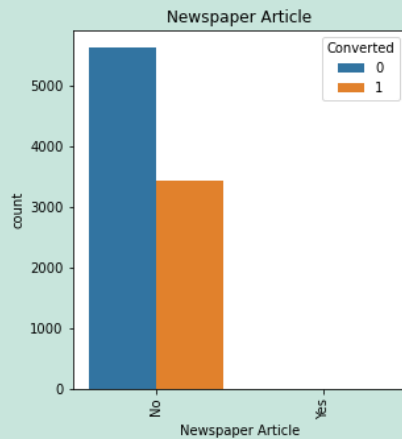

What matters most to you in choosing a course / Search
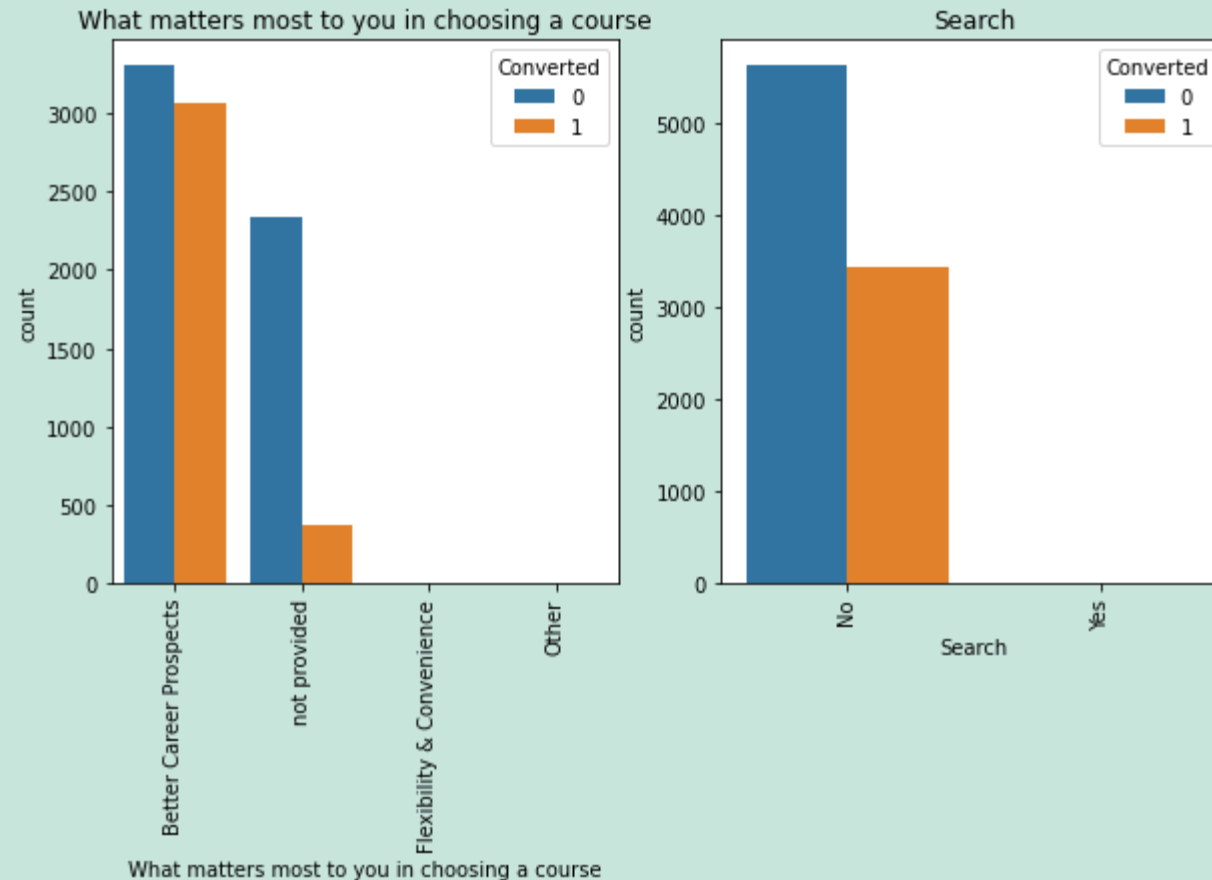
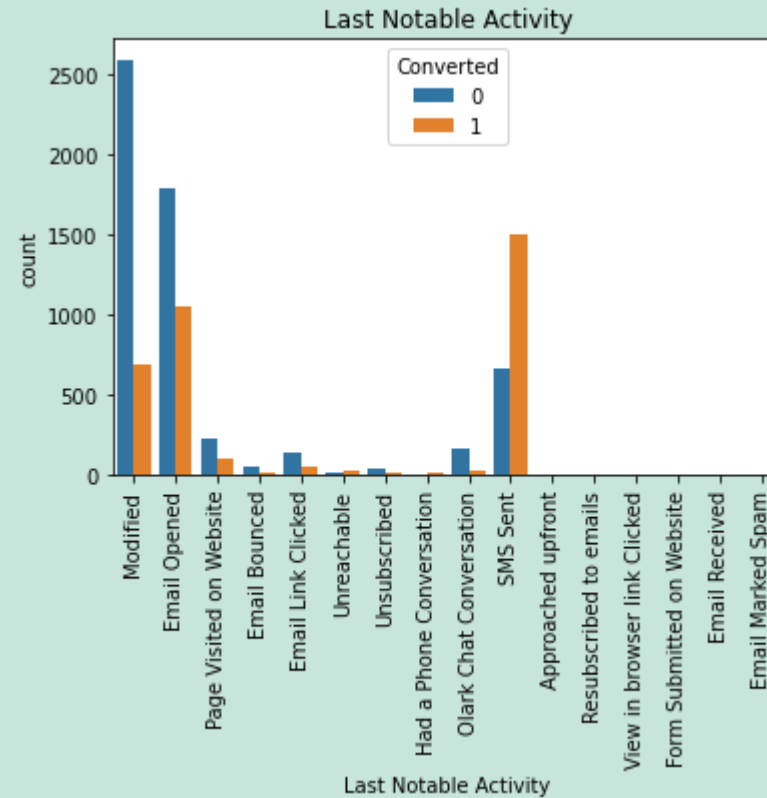◈ We can see that those with the interest of "Better Career Prospects" would most likely convert to a sale.

EDA

◈ Following is the different channels and conversions for the same.

EDA

◈ These were the last notable activities that users did.
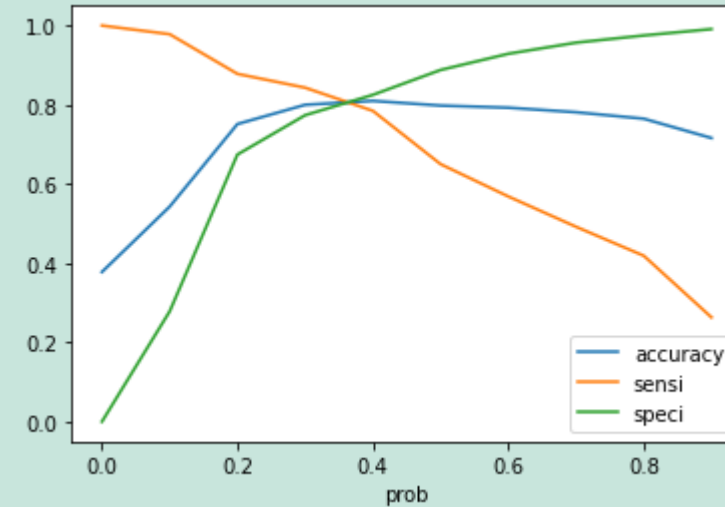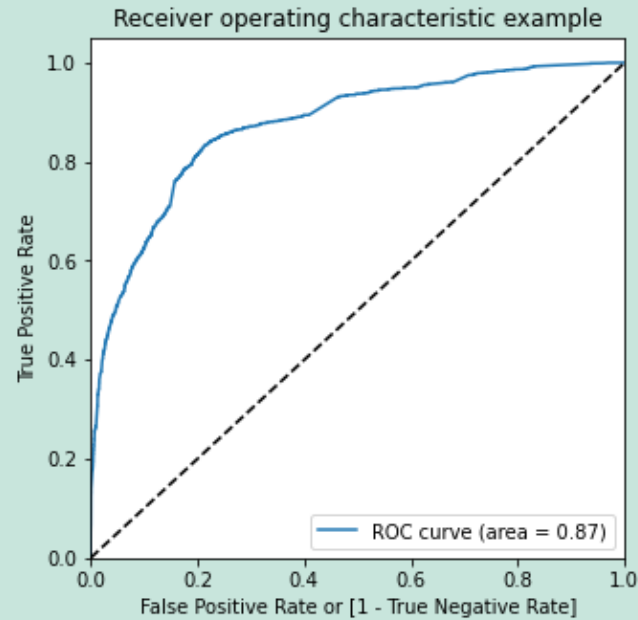

Last Notable Activity

◈ We can see that those marketed by SMS are most likely to convert.

Model Building

◈ Splitting the data into training and test sets using the ratio 70:30

◈ Use RFE for feature selection with 15 variables.

◈ Use RFE for feature selection with 15 variables.

◈ Remove the variables whose P value is greater than 0.05 and VIF value is greater than 5.

ROC Curve

◈ We are trying to find an optimal cut off probability to get a balanced sensitivity and specificity. The optimal cut off is around 0.35.

Conclusion

◈ Variables that mattered the most:

    1) Total time spent on website

    2) Total number of visits

    3) When lead source was the following: Google, Direct Traffic, Organic Search, Welingak Website

    4) When the last activity was the following: SMS, Olark chat conversation

    5) When the current occupation is working professional

When X Education concentrates on the above variables, it will be ale to flourish with more leads.