

Lead score case study

Submitted by:
Mohini fulpagar
Rahul singh
Rishi solanki

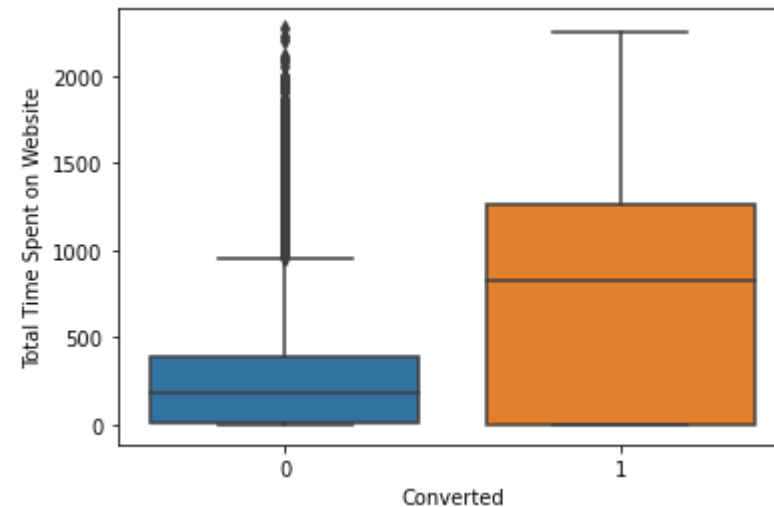
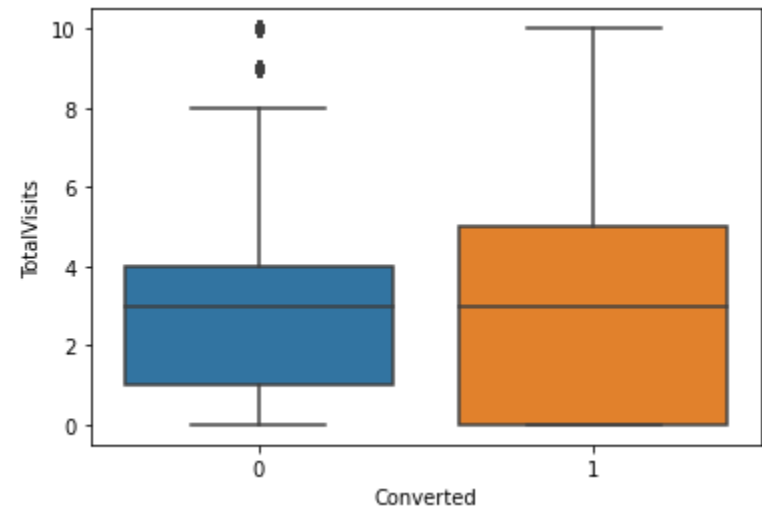
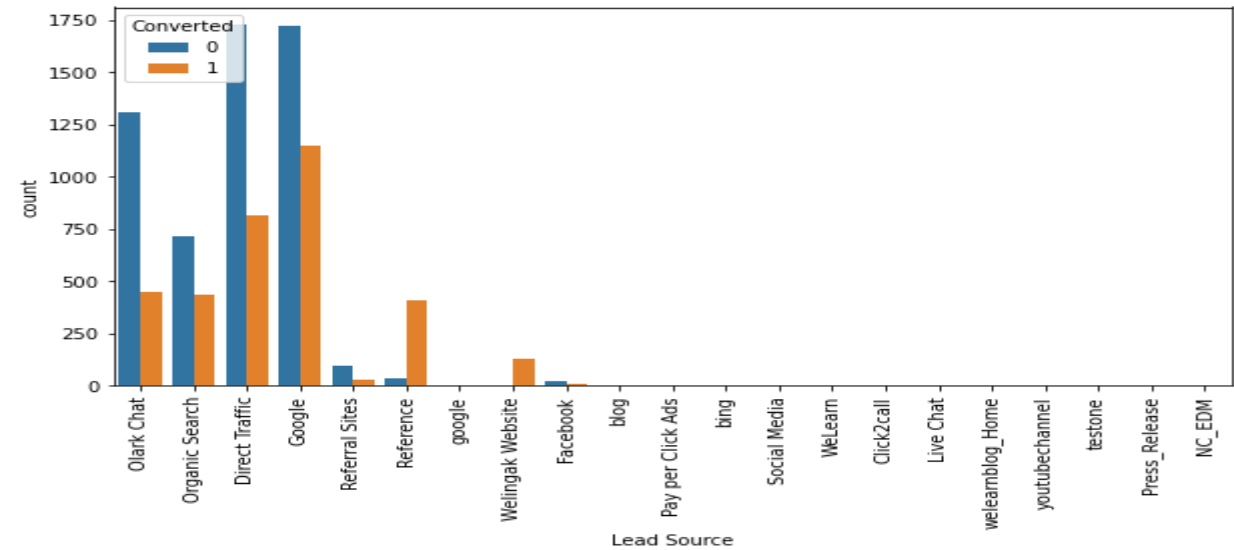
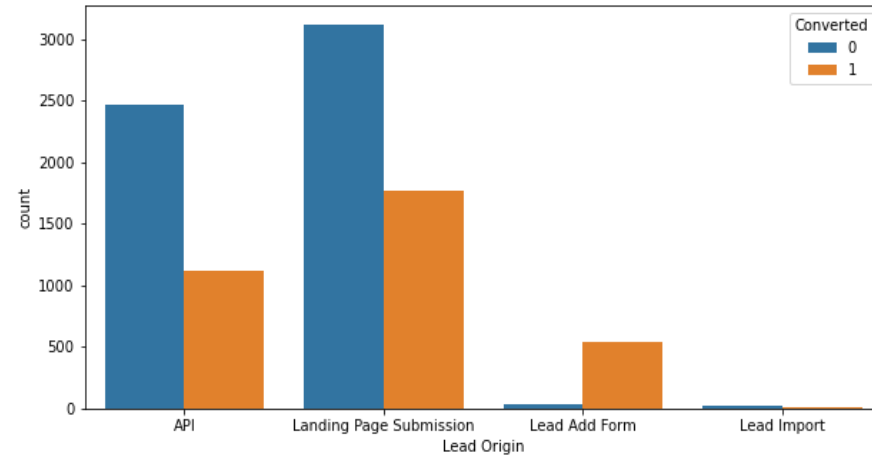
PURPOSE :-

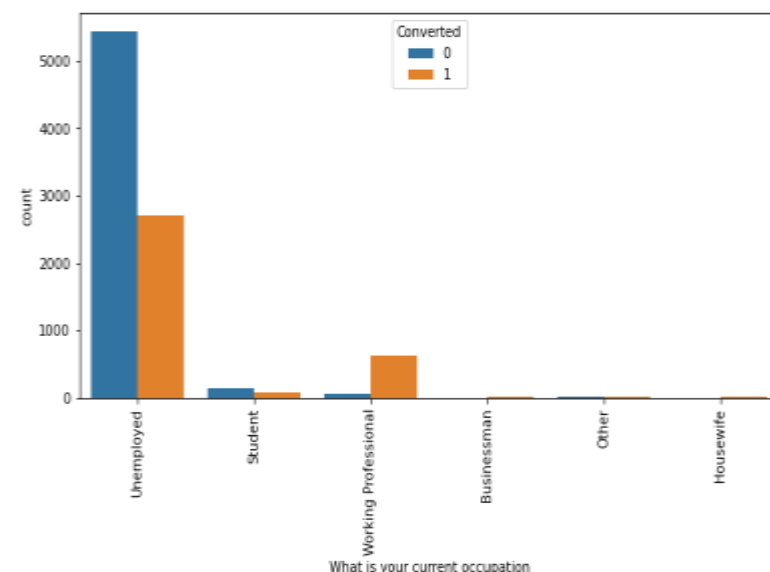
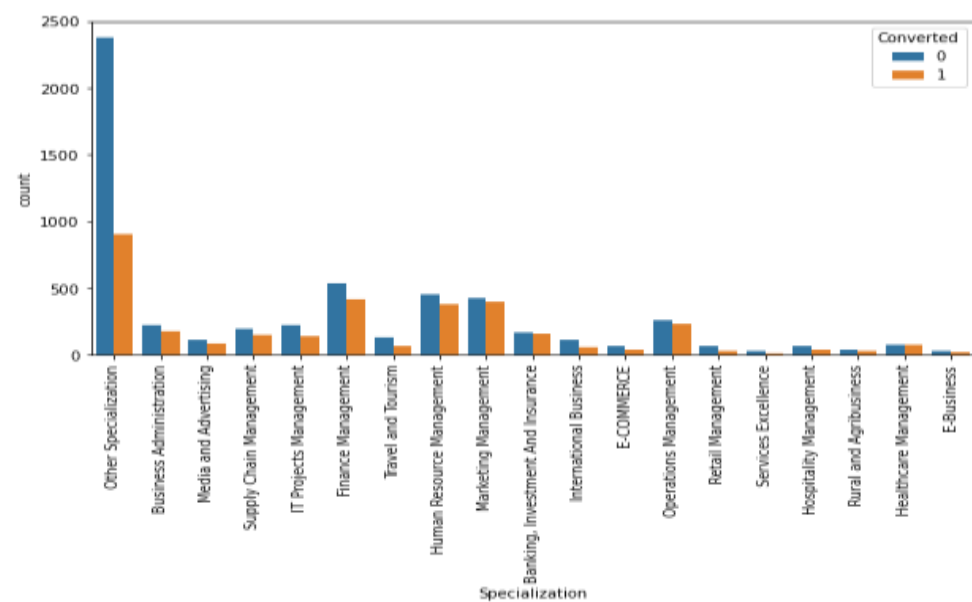
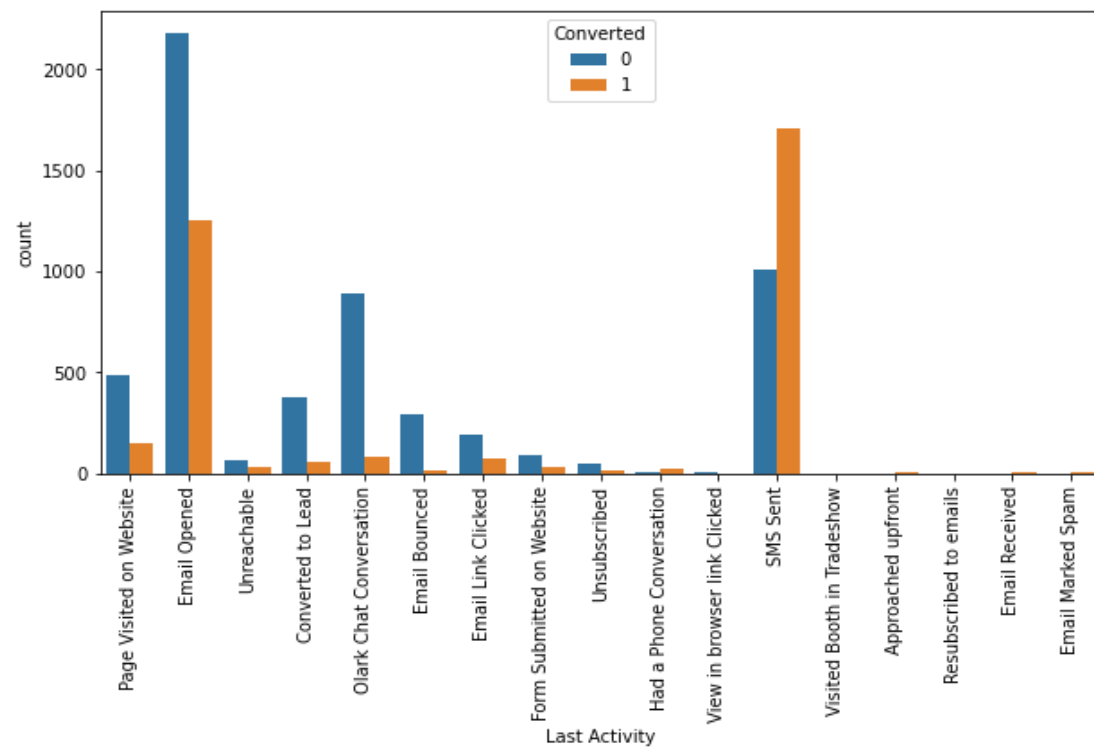
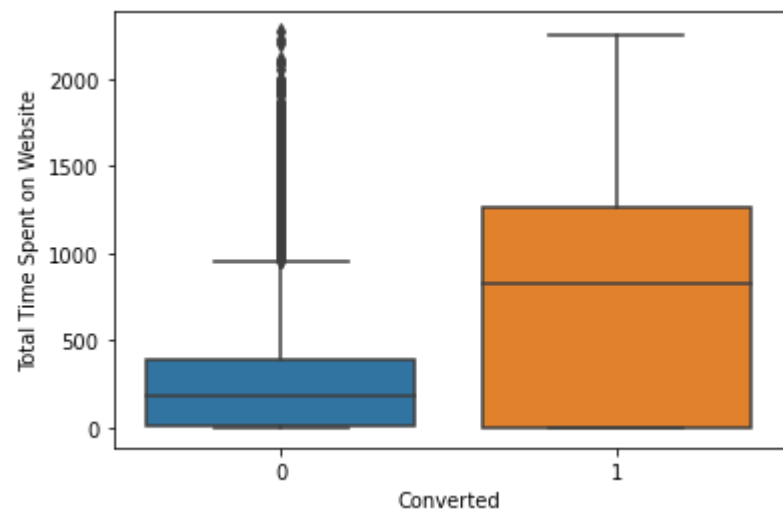
- This analysis is done for X Education and to find ways to get more industry professionals to join their courses. The basic data provided gave us a lot of information about how the potential customers visit the site, the time they spend there, how they reached the site and the conversion rate.
- To build a model to describe the which factor have higher conversion rate or which sector are likely to convert with some effort.

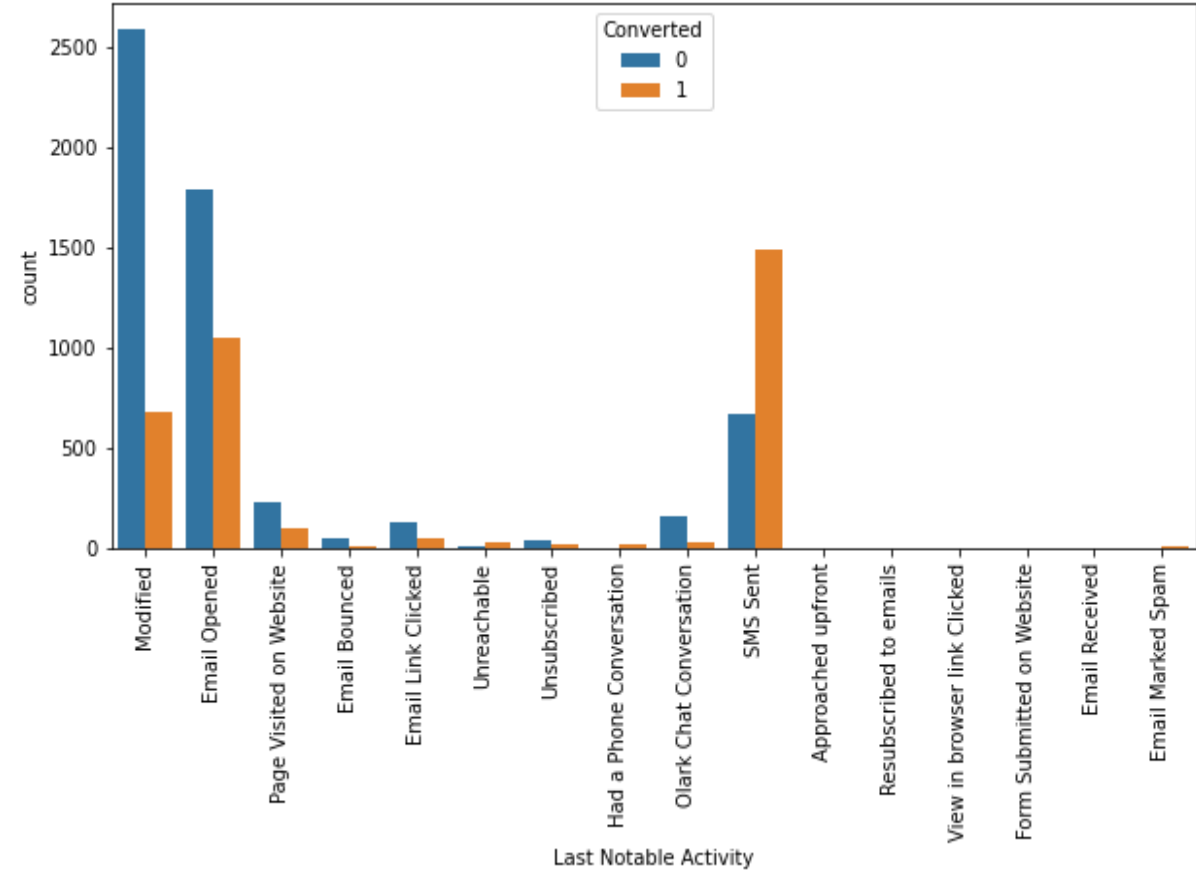
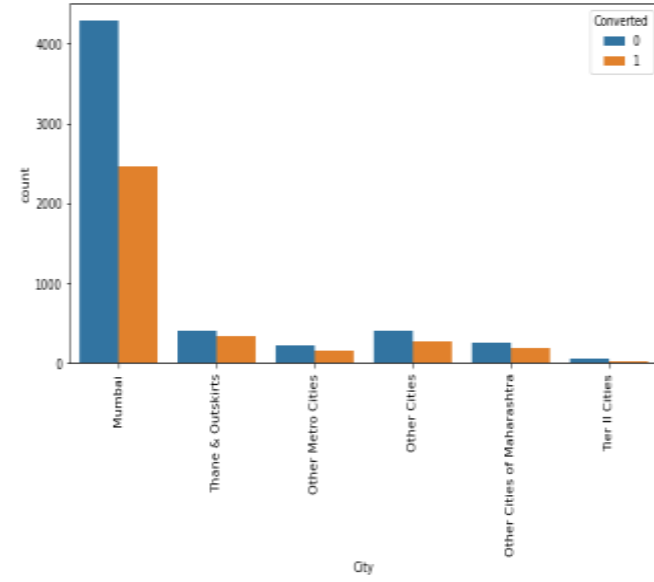
Steps include:-

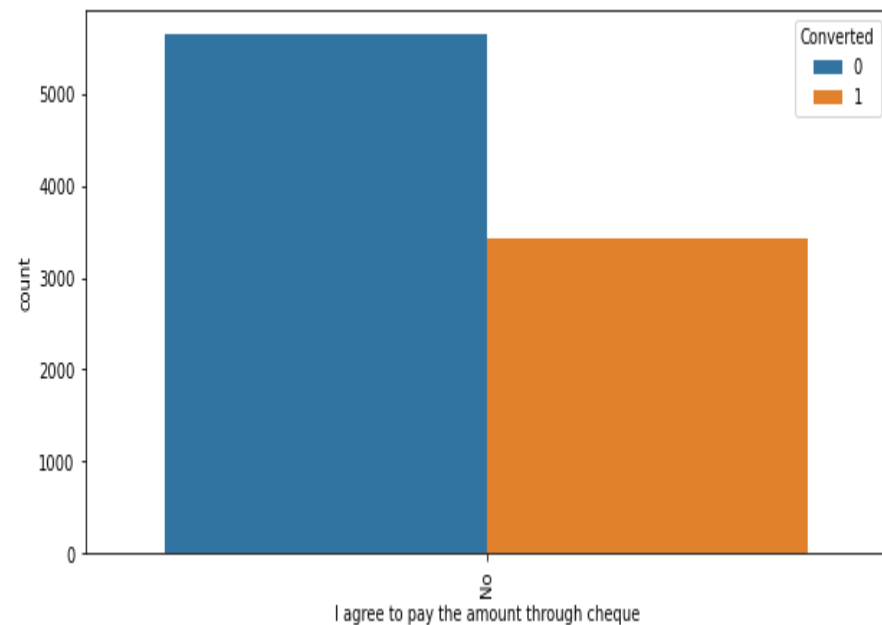
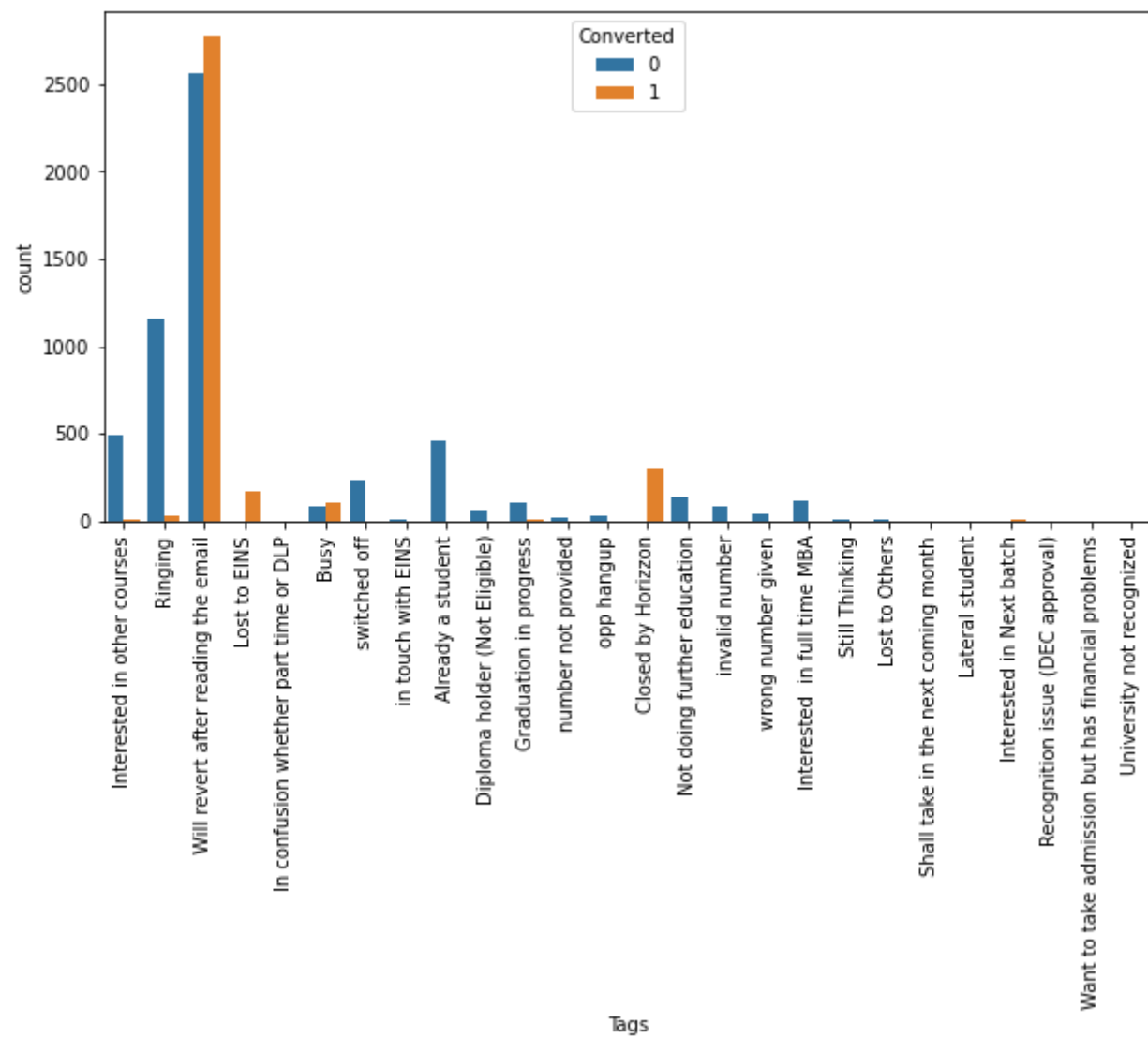
- Understanding the data
- Data cleaning and manipulation
- EDA
- Dummy variable
- Test train splitting
- Model building
- Model evaluation
- Prediction

EDA wrt to target variable









Note: following points to be consider after EDA

- landing page submission and API have conversion rate
- lead Add form have higher conversion rate
- google and direct traffic have high rate for conversion
- people are more engaging on website they have more chance of converted
- lead conversion for SMS sent have high rate as last activity
- those who have Email opened as last activity have conversion rate
- most cases from india
- Other specialization have chance to convert
- working professional have high RATE of joining
- unemployed people have more chance of joining the course

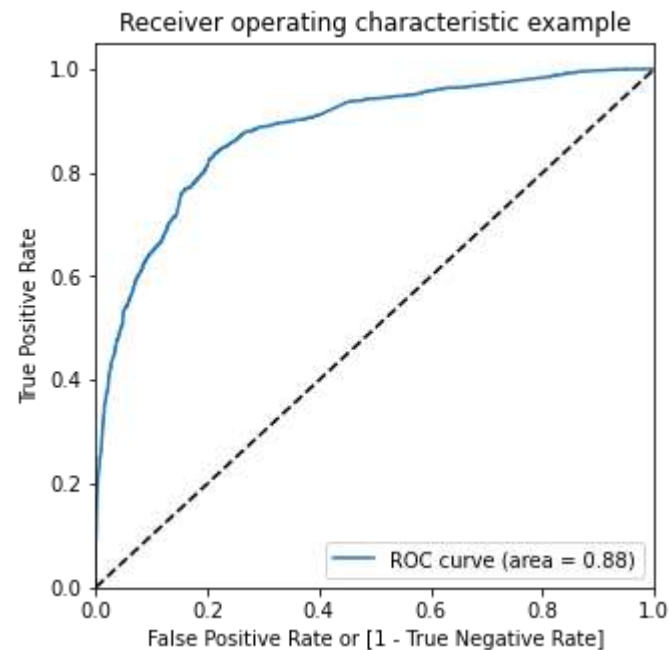
Model building

After building parameter seen as :

| Generalized Linear Model Regression Results | | | |
|---|------------------|-------------------|----------|
| Dep. Variable: | Converted | No. Observations: | 6351 |
| Model: | GLM | Df Residuals: | 6335 |
| Model Family: | Binomial | Df Model: | 15 |
| Link Function: | Logit | Scale: | 1.0000 |
| Method: | IRLS | Log-Likelihood: | -2650.3 |
| Date: | Sun, 19 Mar 2023 | Deviance: | 5300.5 |
| Time: | 21:55:42 | Pearson chi2: | 6.31e+03 |

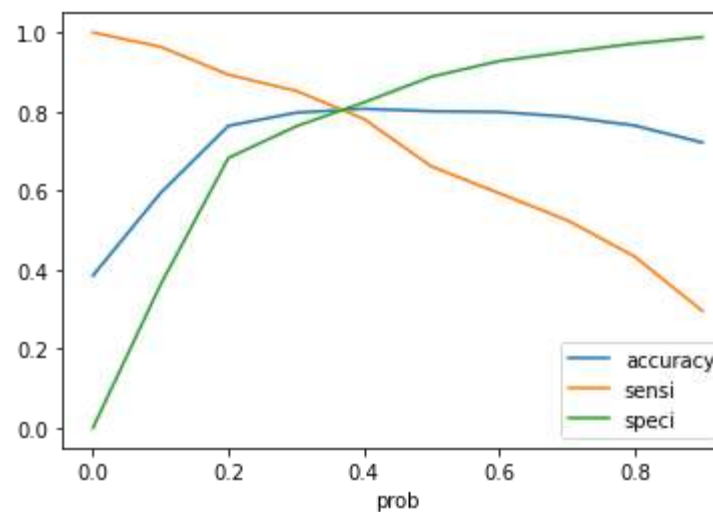
| | coef | std err | z | P> z | [0.025 | 0.975] |
|---|---------|---------|---------|-------|--------|--------|
| const | 1.7476 | 0.130 | 13.459 | 0.000 | 1.493 | 2.002 |
| Do Not Email | -1.7849 | 0.180 | -9.915 | 0.000 | -2.138 | -1.432 |
| Total Time Spent on Website | 1.0050 | 0.037 | 27.023 | 0.000 | 0.932 | 1.078 |
| Lead Origin_Landing Page Submission | -1.2643 | 0.126 | -10.038 | 0.000 | -1.511 | -1.017 |
| Lead Origin_Lead Add Form | 2.8703 | 0.228 | 12.577 | 0.000 | 2.423 | 3.318 |
| Lead Source_Direct Traffic | -0.3914 | 0.091 | -4.295 | 0.000 | -0.570 | -0.213 |
| Lead Source_Organic Search | -0.3846 | 0.108 | -3.562 | 0.000 | -0.596 | -0.173 |
| Lead Source_Referral Sites | -0.6804 | 0.327 | -2.081 | 0.037 | -1.321 | -0.040 |
| Lead Source_Welink Website | 2.6659 | 0.765 | 3.486 | 0.000 | 1.167 | 4.165 |
| What is your current occupation_ Working Professional | 2.5838 | 0.193 | 13.382 | 0.000 | 2.205 | 2.962 |
| Last Notable Activity_Email Link Clicked | -1.7685 | 0.259 | -6.825 | 0.000 | -2.276 | -1.261 |
| Last Notable Activity_Email Opened | -1.4433 | 0.088 | -16.328 | 0.000 | -1.617 | -1.270 |
| Last Notable Activity_Modified | -2.1188 | 0.093 | -22.823 | 0.000 | -2.301 | -1.937 |
| Last Notable Activity_Olark Chat Conversation | -2.5678 | 0.327 | -7.850 | 0.000 | -3.209 | -1.927 |
| Last Notable Activity_Page Visited on Website | -1.7320 | 0.200 | -8.674 | 0.000 | -2.123 | -1.341 |
| Specialization_Other Specialization | -1.1047 | 0.122 | -9.031 | 0.000 | -1.344 | -0.865 |

Roc curve



Finding the cutoff point from curve:

- Approx 0.37 is the cutoff point



Accuracy, sensitivity and specificity

| For train data set | | |
|--------------------|-------------|-------------|
| Accuracy | Sensitivity | Specificity |
| 80% | 80% | 80% |
| For test data set | | |
| Accuracy | Sensitivity | Specificity |
| 80% | 78% | 81% |

Lead score

- Total 364 lead are hot lead having score more than 85

conclusion

| | |
|--|-----------|
| Lead Origin_Lead Add Form | 2.870277 |
| Lead Source_Welingak Website | 2.665893 |
| What is your current occupation_Working Professional | 2.583803 |
| const | 1.747645 |
| Total Time Spent on Website | 1.004997 |
| Lead Source_Organic Search | -0.384635 |
| Lead Source_Direct Traffic | -0.391358 |
| Lead Source_Referral Sites | -0.680416 |
| Specialization_Other Specialization | -1.104708 |
| Lead Origin_Landing Page Submission | -1.264281 |
| Last Notable Activity_Email Opened | -1.443274 |
| Last Notable Activity_Page Visited on Website | -1.731995 |
| Last Notable Activity_Email Link Clicked | -1.768517 |
| Do Not Email | -1.784873 |
| Last Notable Activity_Modified | -2.118759 |
| Last Notable Activity_Olark Chat Conversation | -2.567752 |

- Lead origin add form, lead source welingak website, current occupation working professional have higher chances of conversion.
- those who spent more on website convert as lead.
- Company should not call to last notable activity olark chat conversion as they have would not convert as lead.