

X Education - Lead Scoring Case Study

Increasing lead conversion rate for X education

Group Members:

- **Prasoon**
- **Mayukh Das**
- **Cyril Philip**

Background

X Education Company

- ❖ X Education , An education company sells online courses to industry professionals.
- ❖ Many interested professionals land on their website.
- ❖ The company markets its courses on several websites like Google. Once these people land on the website, they might browse the courses or fill up a form for the course or watch relevant videos.
- ❖ When these people fill up a form providing their email address or phone number, they are classified as a lead.
- ❖ Once these leads are acquired, employees from the sales team start making calls, writing emails, etc. Through this process, some of the leads get converted while most do not.
- ❖ The typical lead conversion rate at X education is around 30%.

Problem Statement

X Education Company's Problem

- ❖ X Education gets a lot of leads but its lead conversion rate is very poor.
- ❖ To make this process more efficient, the company wishes to identify the most potential lead, 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.
- ❖ We will help them select the most promising leads, i.e. the leads that are most likely to convert into paying customers.
- ❖ We are required to build a model wherein we need to assign a lead score to each of the leads such that the customers with higher lead score have a higher conversion chance.
- ❖ The CEO, in particular, has given a ballpark of the target lead conversion rate to be 80%.

Solution

Finding Hot Leads accurately

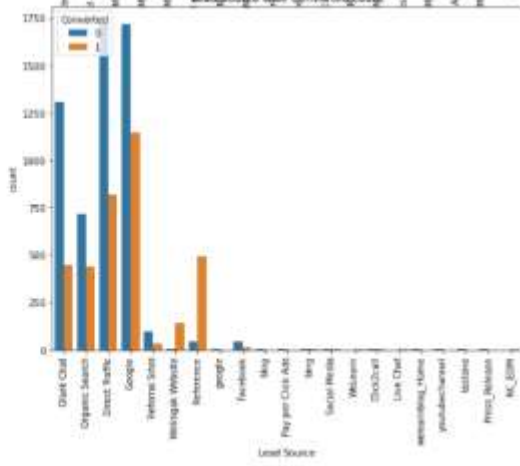
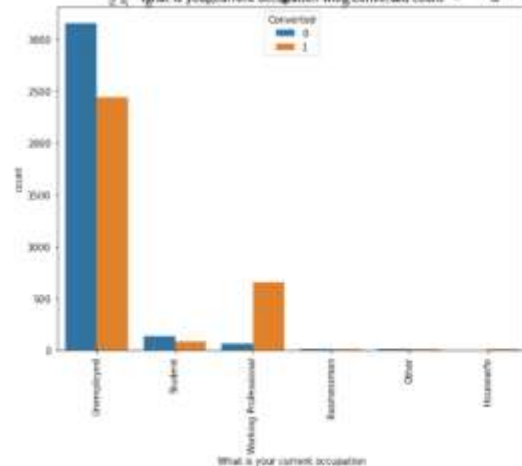
For the Solution to the mentioned problem, the crucial part is to accurately identify hot leads.

Higher Conversions seem directly proportional to the accuracy in obtaining the hot leads

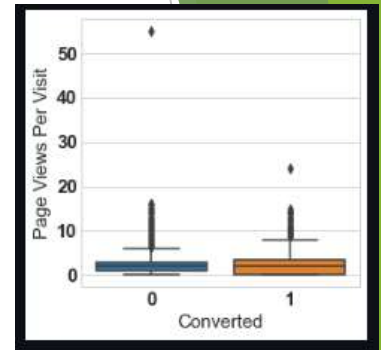
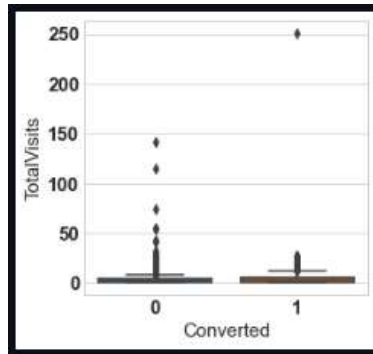
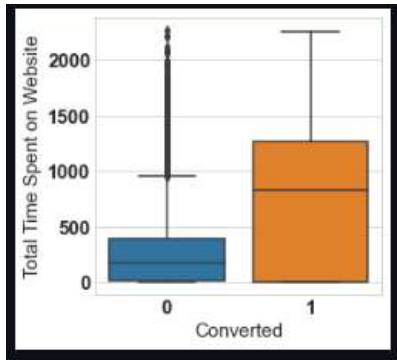
Hence, high number of hot leads are required in order to fulfill the criteria of 80% conversion rate.

EDA- Visualization

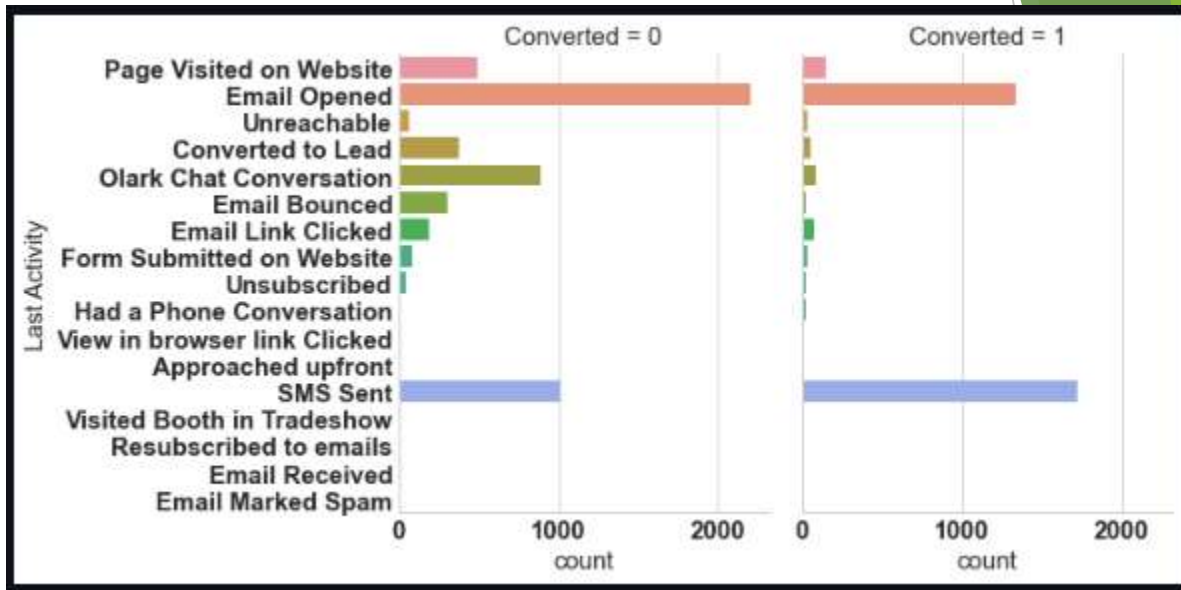
The background of the slide is black. On the right side, there is a decorative graphic consisting of several overlapping, semi-transparent green triangles and polygons of various shades, ranging from dark forest green to bright lime green. A thin white line also runs diagonally across the lower right portion of the slide, intersecting the green shapes.



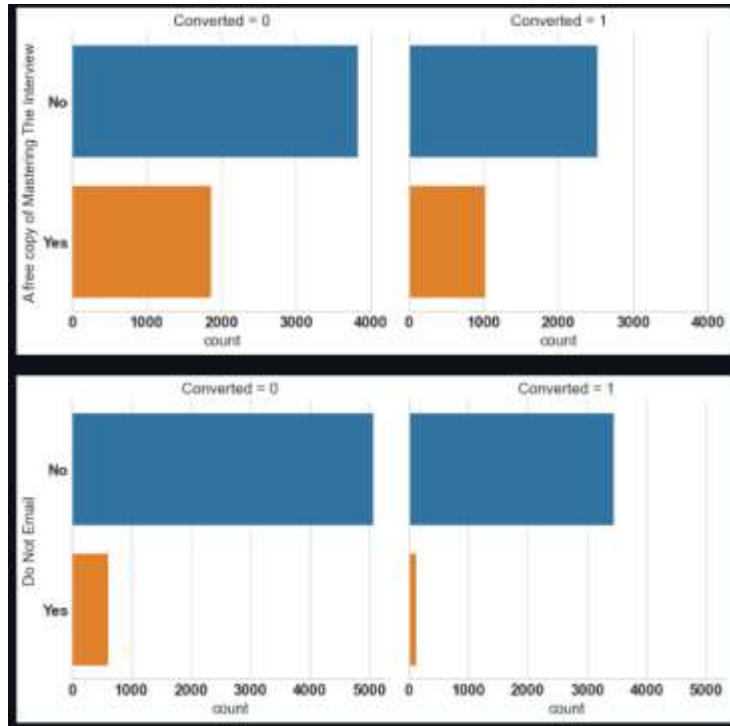
1



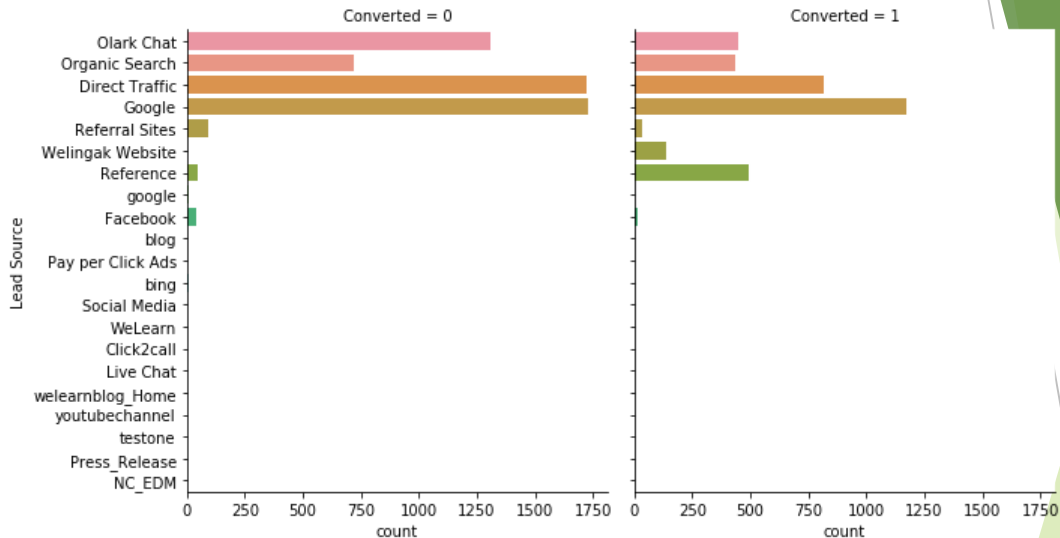
EDA plots depicting variation in 'Time spent', 'Total visits' and 'Views per visit' columns for Converted and Not-Converted.



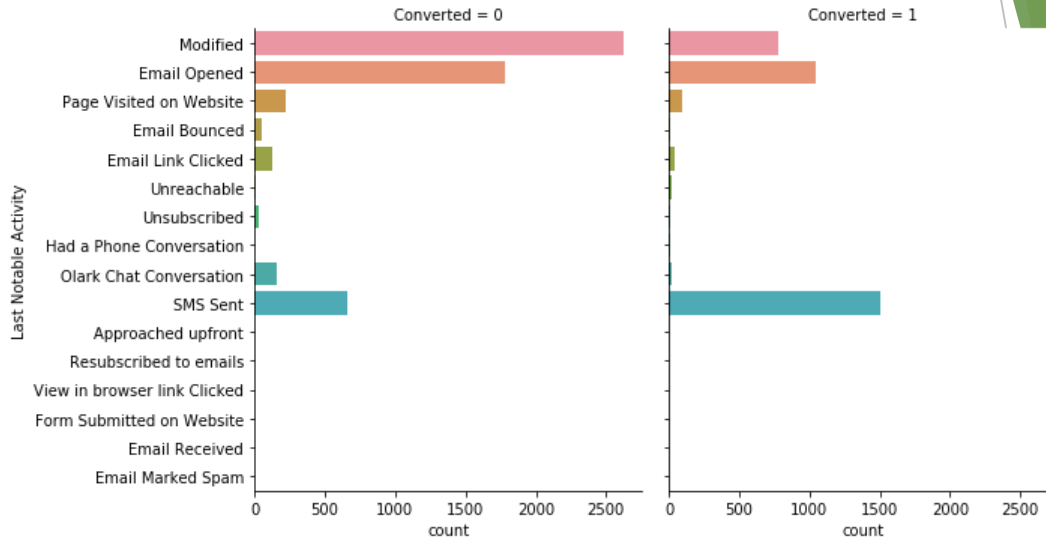
EDA plots depicting variation 'Last Activity' column for Converted and Not-Converted.



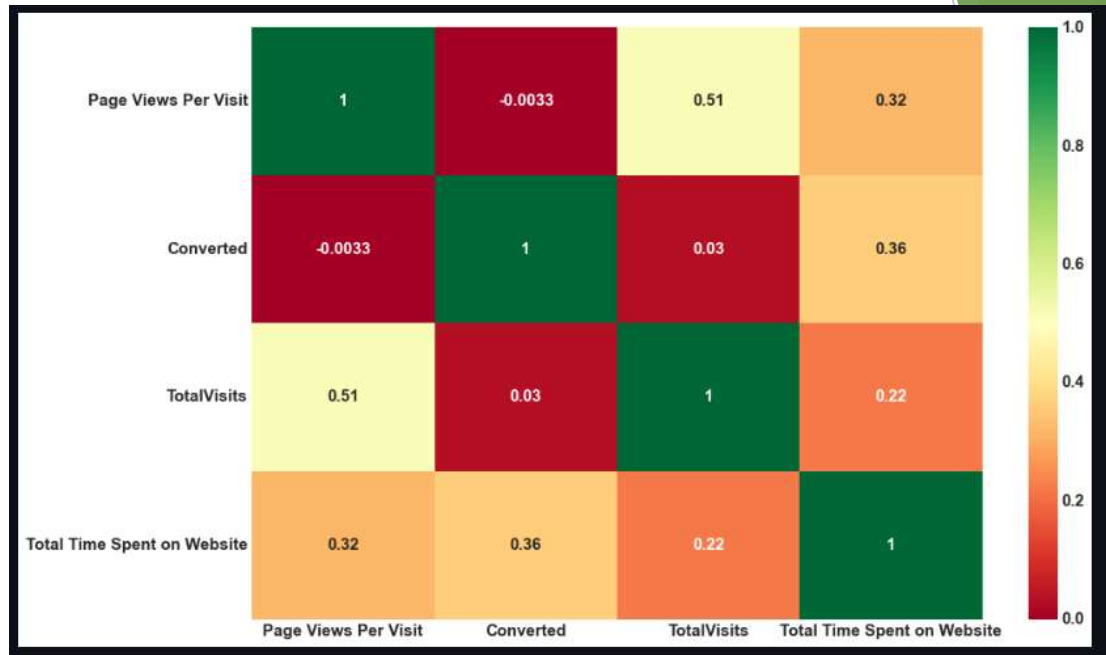
Variation in columns 'A free copy of Mastering The Interview' and 'Do not email' for Converted and Not-Converted.



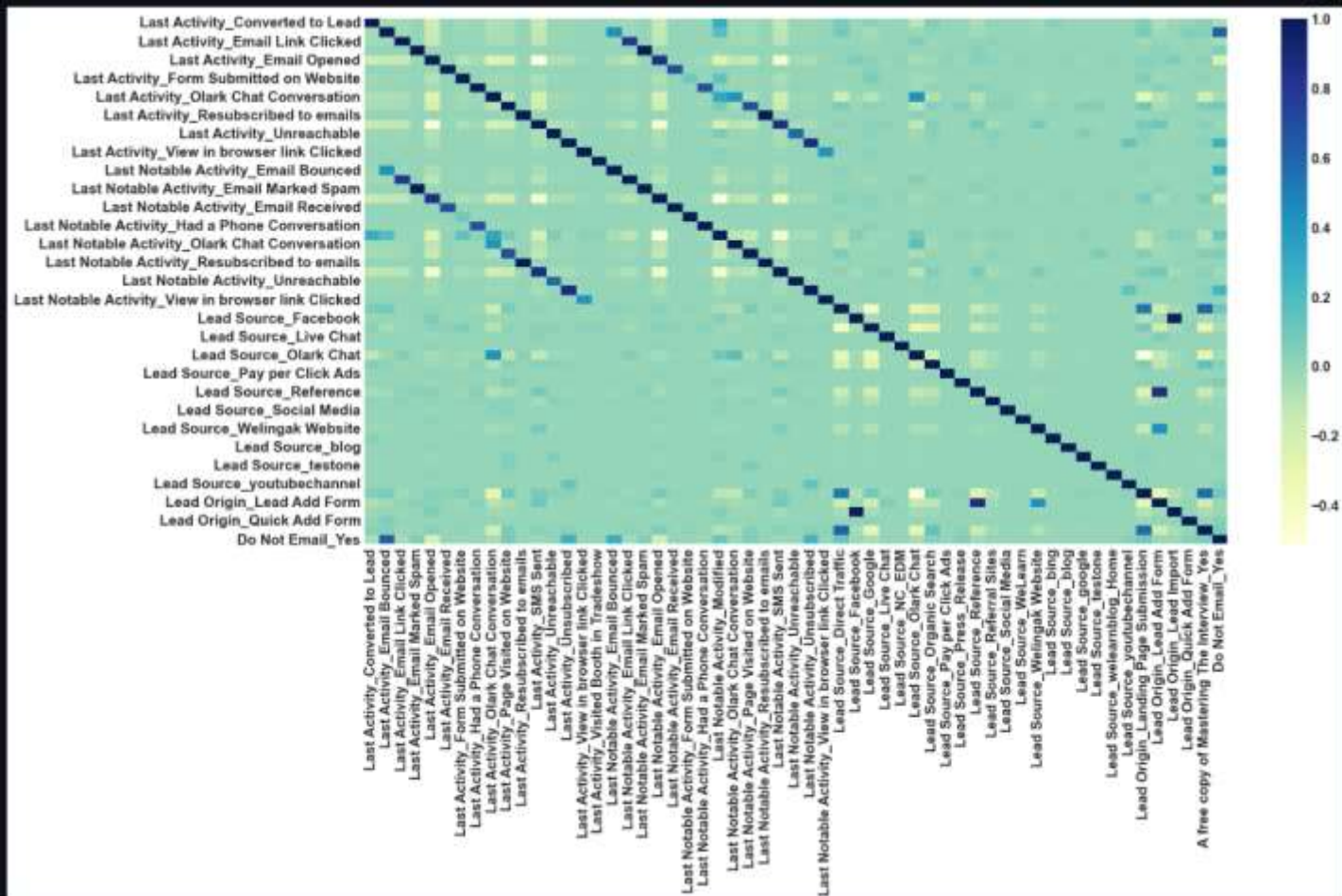
Visualization plots showing the variation in column 'Lead Source' for Converted and Not-Converted.



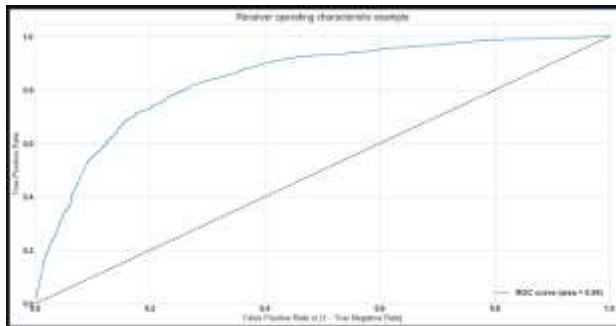
Visualization plots showing variation in column 'Last Notable Activity' for those Converted and Not-Converted.



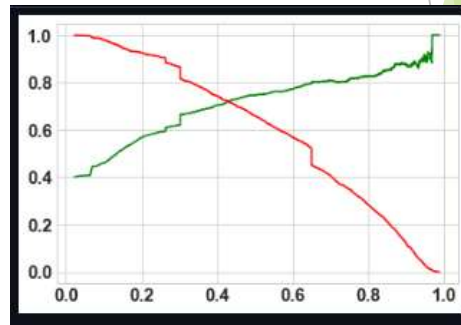
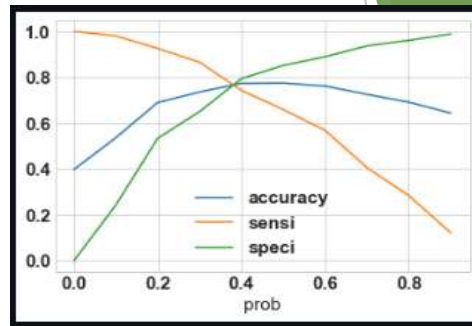
'Heat Map' for all finally selected numerical columns after EDA.



‘Heat Map’ for all selected columns (numerical as well as dummy columns).



Linear Regression Final Model Parameters
Area under ROC = 0.84
Intermediate cut-off = 0.35
Final cut-off = 0.42



Inference / Conclusion

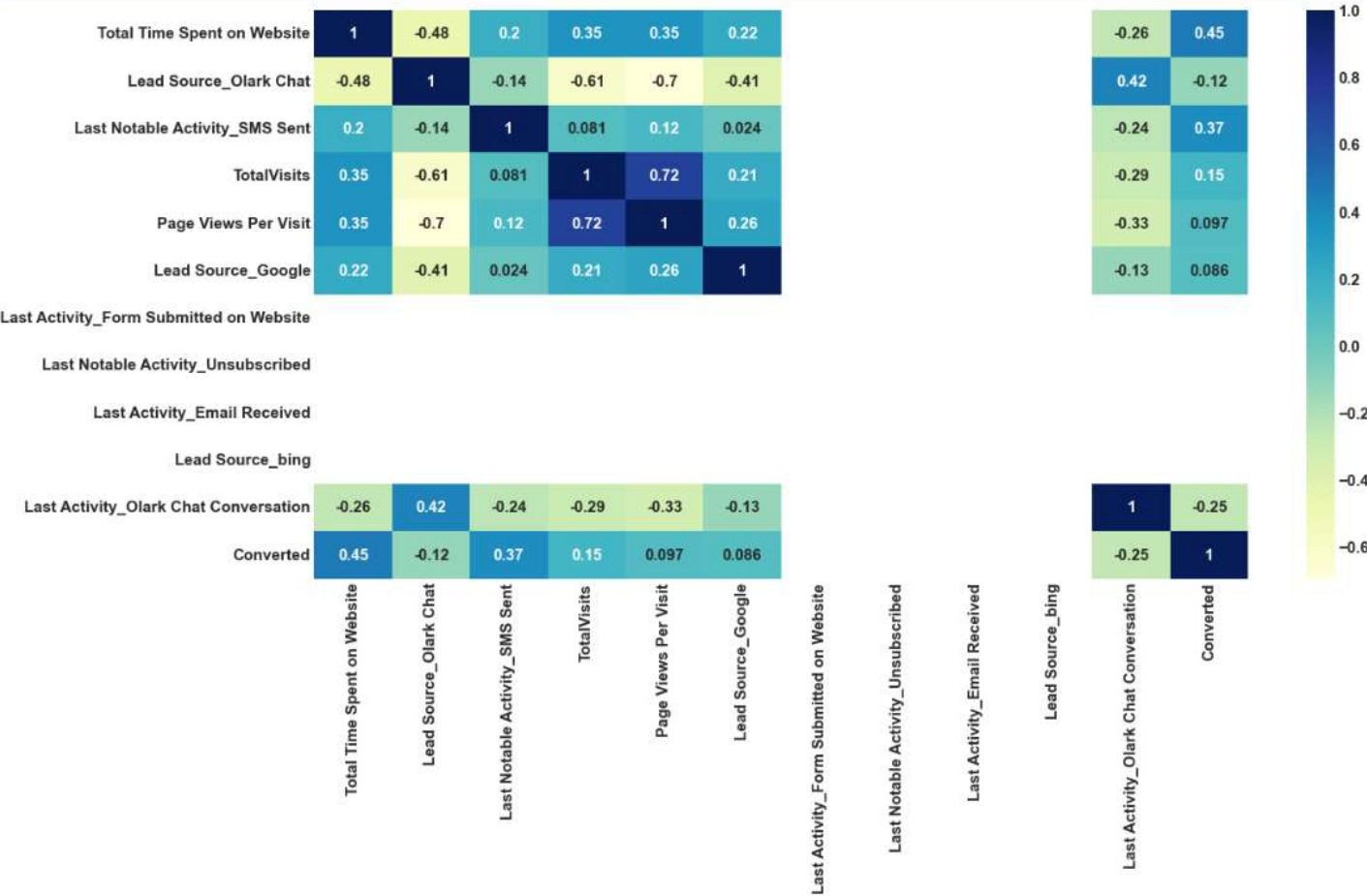
Model Analysis

Performance of Final Model

Overall accuracy on Test set: 0.76

Sensitivity for the logistic regression model: 0.73

Specificity for the logistic regression model: 0.81



Correlation Between Model selected indicators

Generalized Linear Model Regression Results

```

=====
Dep. Variable:          Converted    No. Observations:          4181
Model:                  GLM         Df Residuals:              4173
Model Family:          Binomial     Df Model:                  7
Link Function:         Logit        Scale:                     1.0000
Method:                IRLS         Log-Likelihood:           -2027.5
Date:                  Wed, 12 Oct 2022    Deviance:                 4054.9
Time:                  16:15:08          Pearson chi2:             4.27e+03
No. Iterations:        5              Pseudo R-squ. (CS):       0.3124
Covariance Type:       nonrobust
=====

```

```

=====
                                coef    std err          z      P>|z|      [0.025    0.975]
-----
const                        -1.2557     0.075    -16.770     0.000    -1.402    -1.109
Total Time Spent on Website     1.1661     0.047     24.647     0.000     1.073     1.259
TotalVisits                     0.3169     0.057      5.538     0.000     0.205     0.429
Lead Source_Olark Chat         1.5555     0.153     10.146     0.000     1.255     1.856
Last Notable Activity_SMS Sent  1.5186     0.086     17.720     0.000     1.351     1.687
Page Views Per Visit           -0.2094     0.064     -3.290     0.001    -0.334    -0.085
Lead Source_Google              0.3233     0.089      3.640     0.000     0.149     0.497
Last Activity_Form Submitted on Website      0         0         nan         nan         0         0
Last Notable Activity_Unsubscribed           0         0         nan         nan         0         0
Last Activity_Email Received                0         0         nan         nan         0         0
Lead Source_bing                       0         0         nan         nan         0         0
Last Activity_Olark Chat Conversation    -1.7857     0.185     -9.630     0.000    -2.149    -1.422
=====

```

Relevant stats

Conclusion (Recommendation)

X Education Company needs to work on below mentioned key aspects in order to improve the overall conversions:

- ⌘ Look for ways to increase user engagement on the website in order to have higher conversions.
- ⌘ Get Total visits increased by advertising etc.
- ⌘ SMS notifications seems an effective way to increase conversions.
- ⌘ Olark Chat service has to be improved as this seems helpful.

Thank-you

The background features abstract, overlapping green geometric shapes, primarily triangles and polygons, in various shades of green. These shapes are concentrated on the right side and bottom of the slide, creating a modern, layered effect. The rest of the slide is a plain, light gray background.