LLESLIN DEVASSY PULLUPARABIL
PACHIPULUSU SAI PRADEEP
NIKHIT HEGDE

Case Study
Lead Scoring

X - E D U C A T I O N

STUDIO SHODWE

About X Education

An education company selling online courses to industry professionals.
The company markets its courses on several websites and search engines like Google.

When people fill in a form citing their interest.The captured data is classified as a lead.
X Education has a Lead conversion rate of 30%.

Problem Statement

01 X Education gets a lot of leads but its lead conversion rate is very poor.

02 To build a logistic regression model to help the organisation increase its lead conversion

03 Achive a conversion rate of 80%

Lead Conversion Process

Ads and Referrals

Potential Customer's fill in their details

Leads are followed up by the Tele Callers

END RESULT: 30% of the Leads are converted

Creating a Logistic Regression Model to increase the conversion rate to 80%

Proposed Solution

Hot Leads

Clustering the leads into categories based on their tendency / probability to convert.Thus getting a dedicated section to focus on

Focus on Hot Leads

As we have the potential hot leads , focusing on them with effective communication would be the key for a successful conversion

Better Conversion Rate

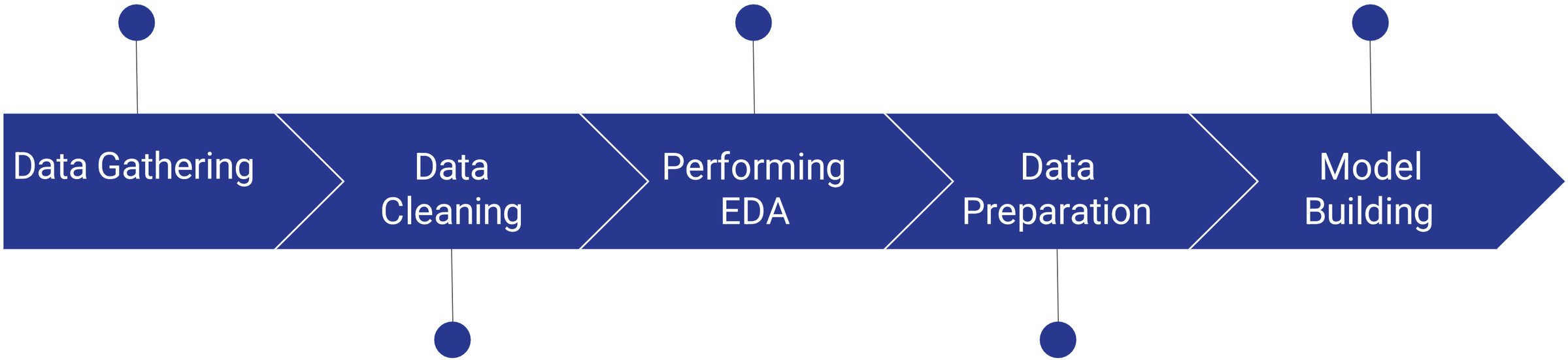
As we focus on the Hot leads, the conversion would be positive and would be a better use of Time , Energy and Money.

Implementation

Analyze the data

Univariate, Bivariate, and Heatmap for numerical and categorical columns

Performing pre-requisites for RFE and Logistic Regression

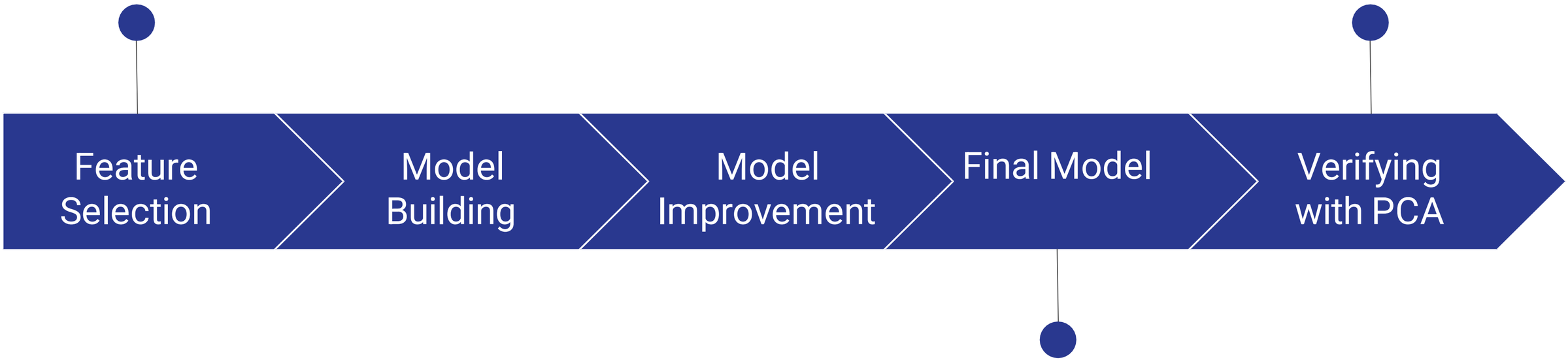


Duplicate removal, null value treatment, unnecessary column elimination, etc.

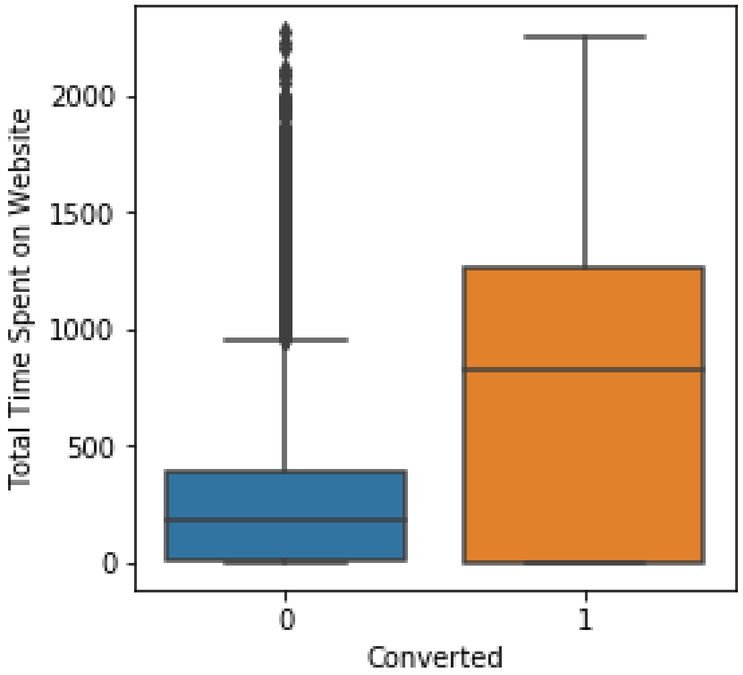
Outlier Treatment, Feature-Standardization

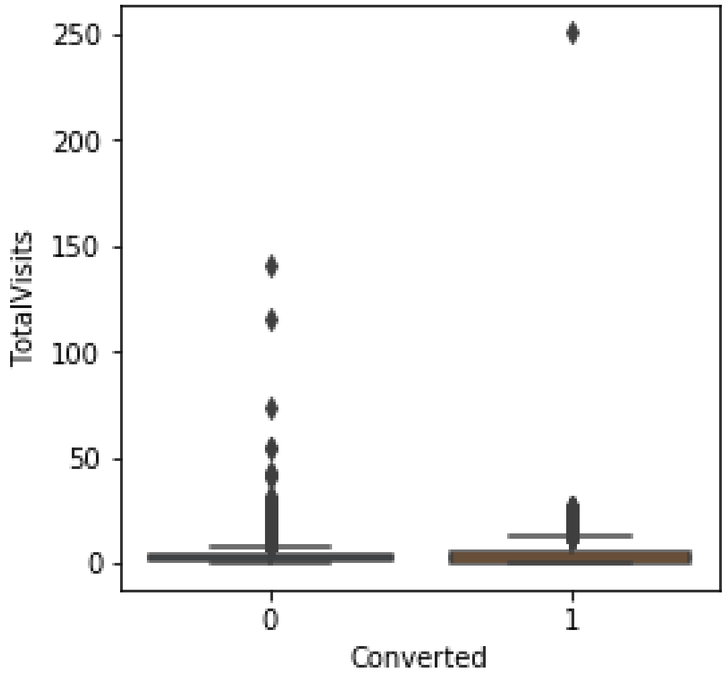
RFE 25 features

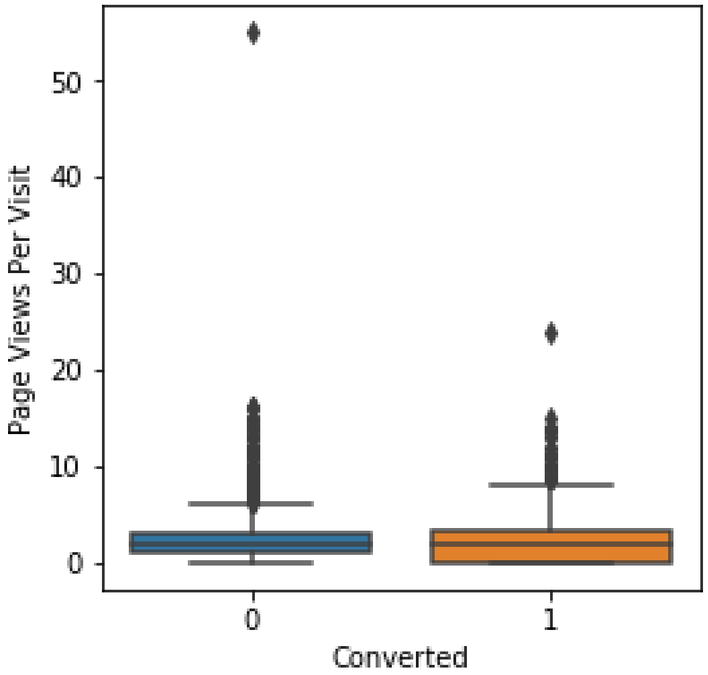
Verifying our Final Model Accuracy etc.



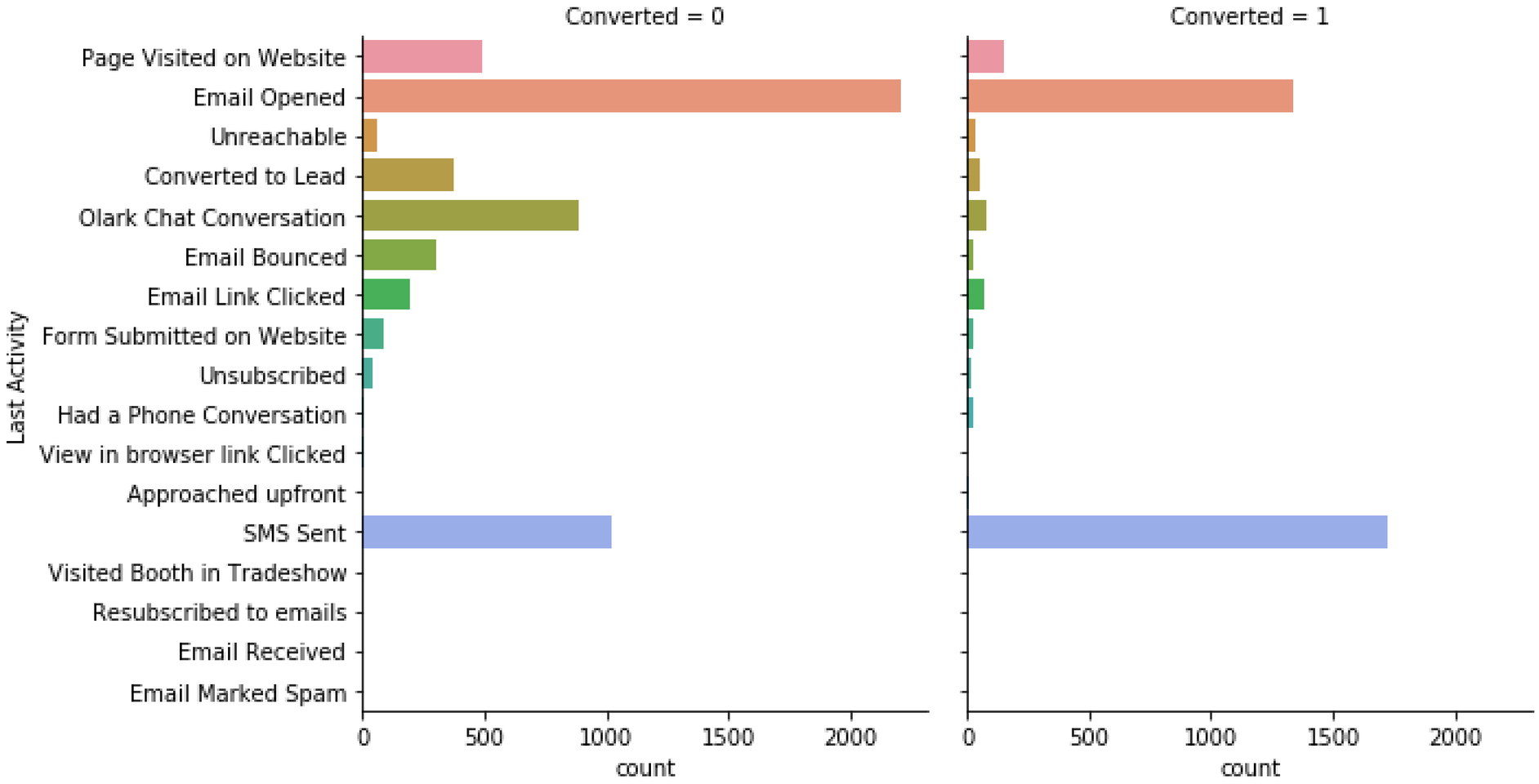
Model Analysis on Test Data



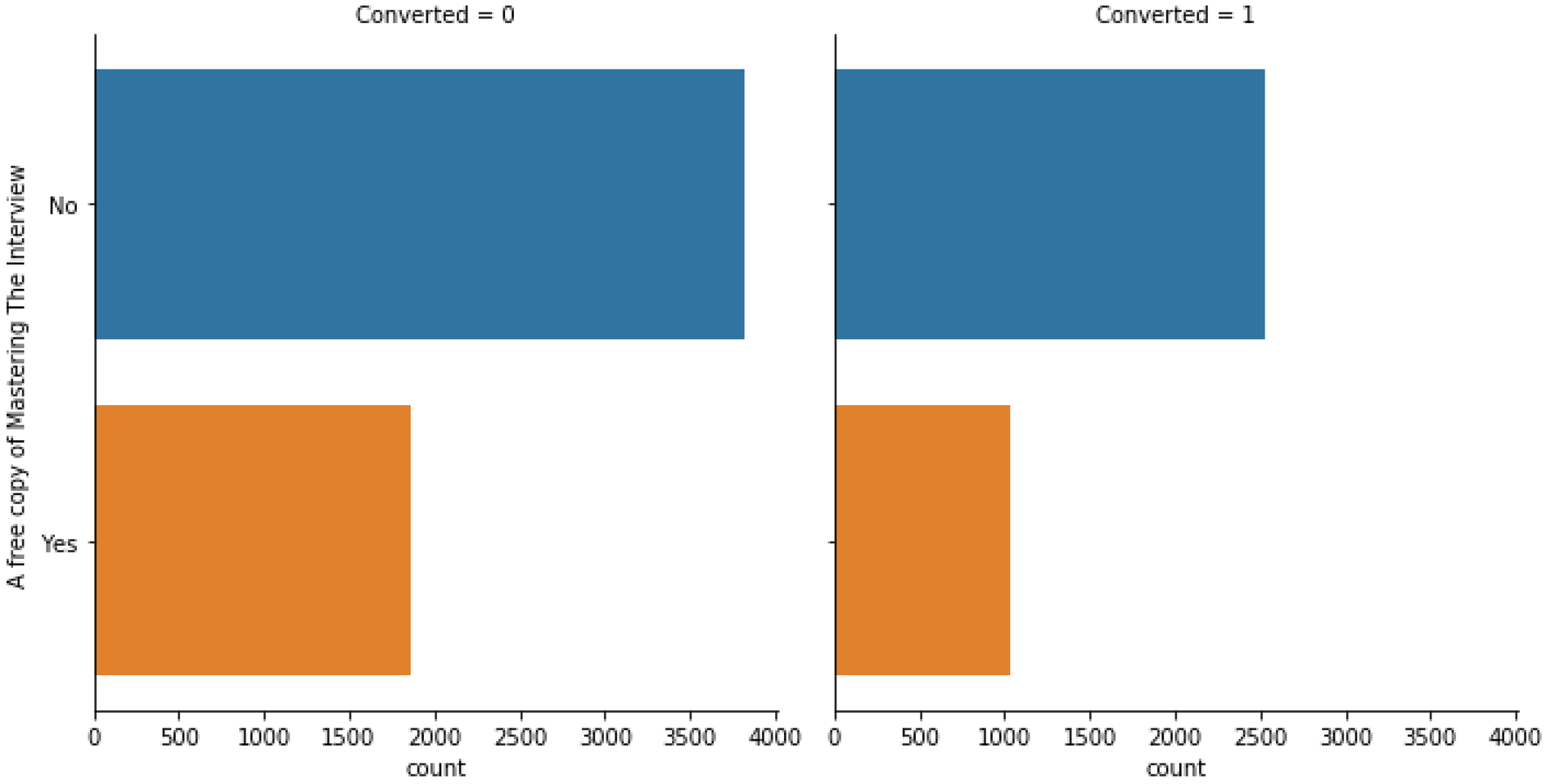




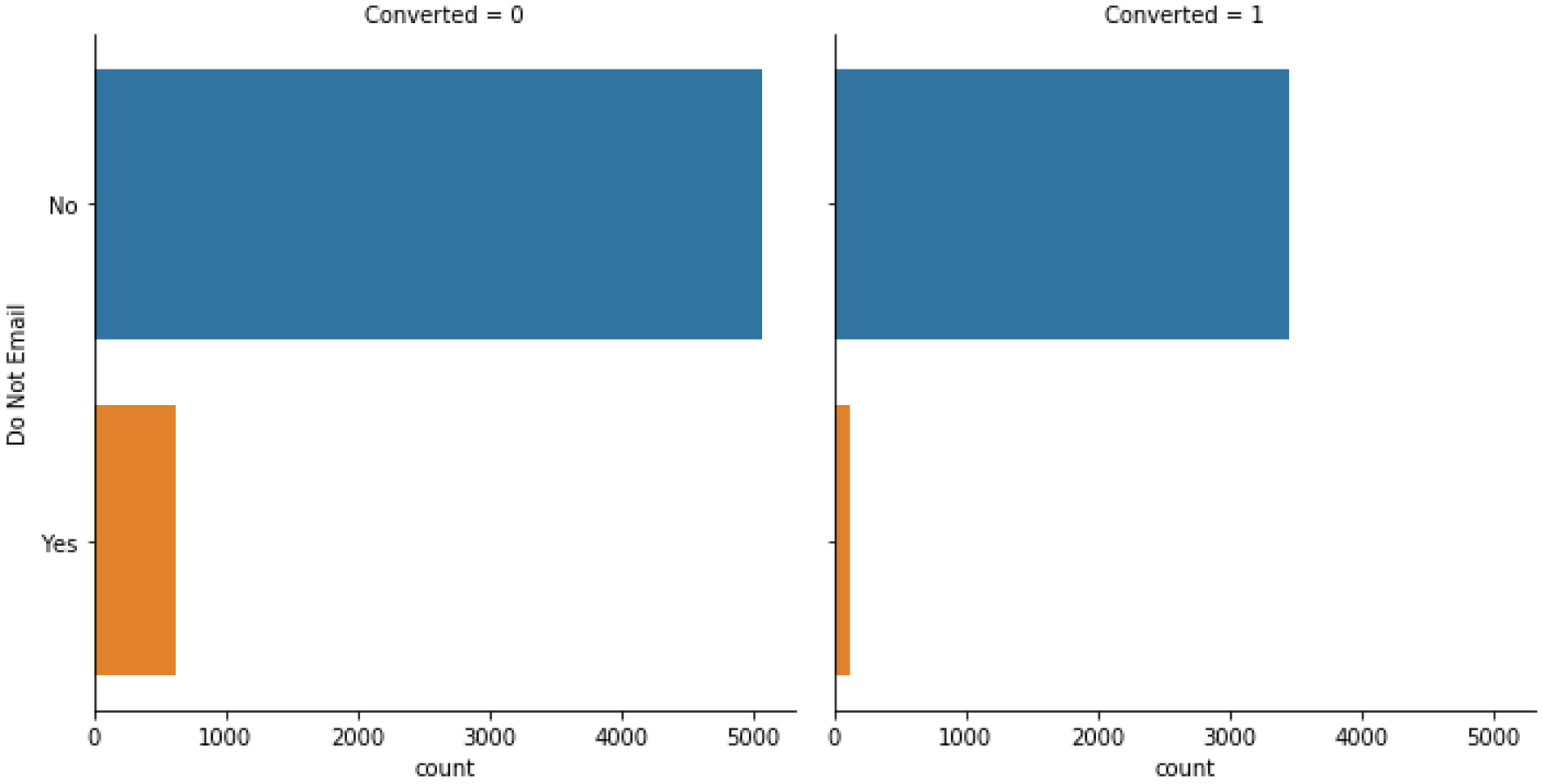
Converted Numerical Data



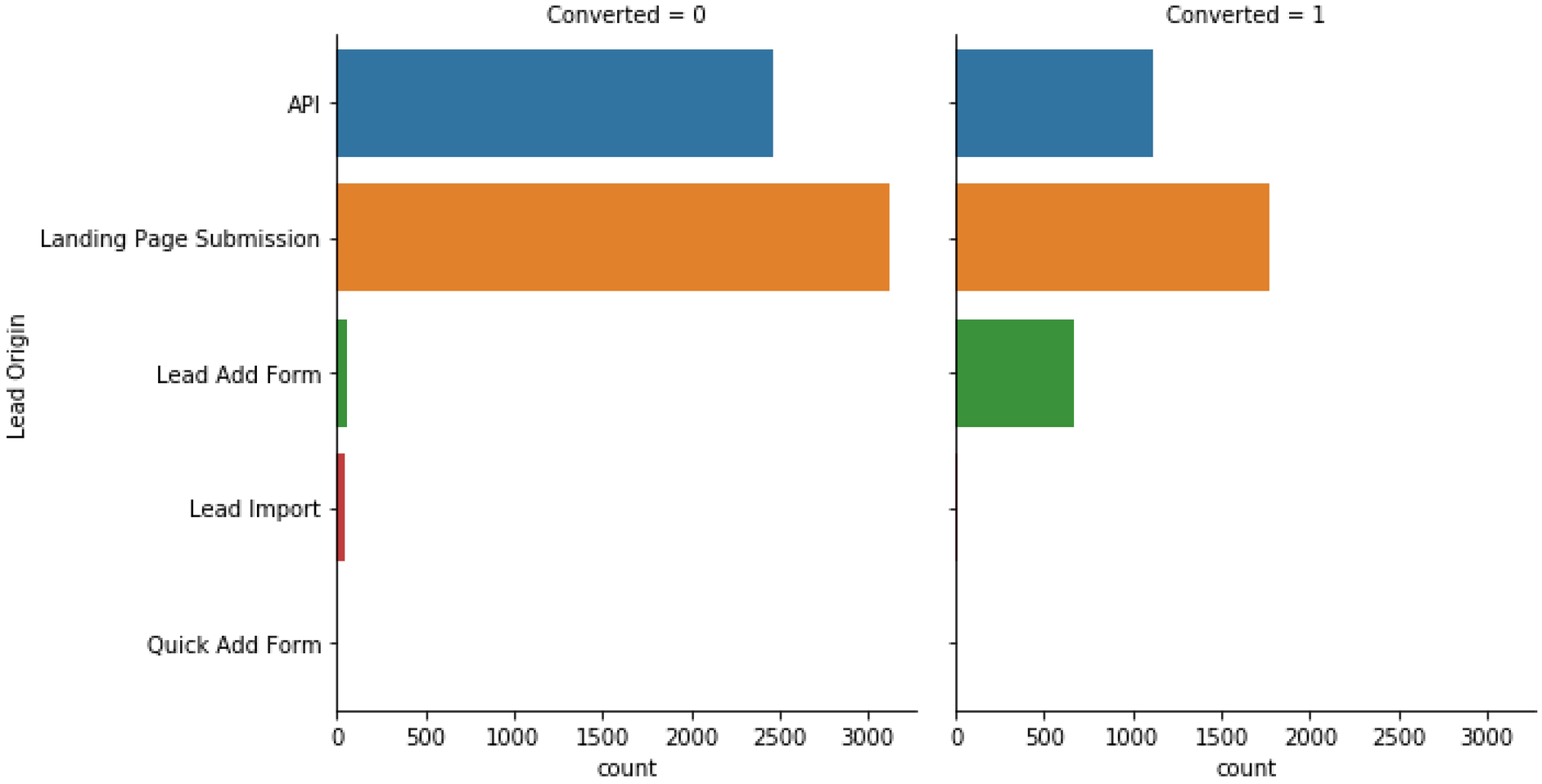
Variation in categorical column (Last Activity) for those who Converted and those who didn't.



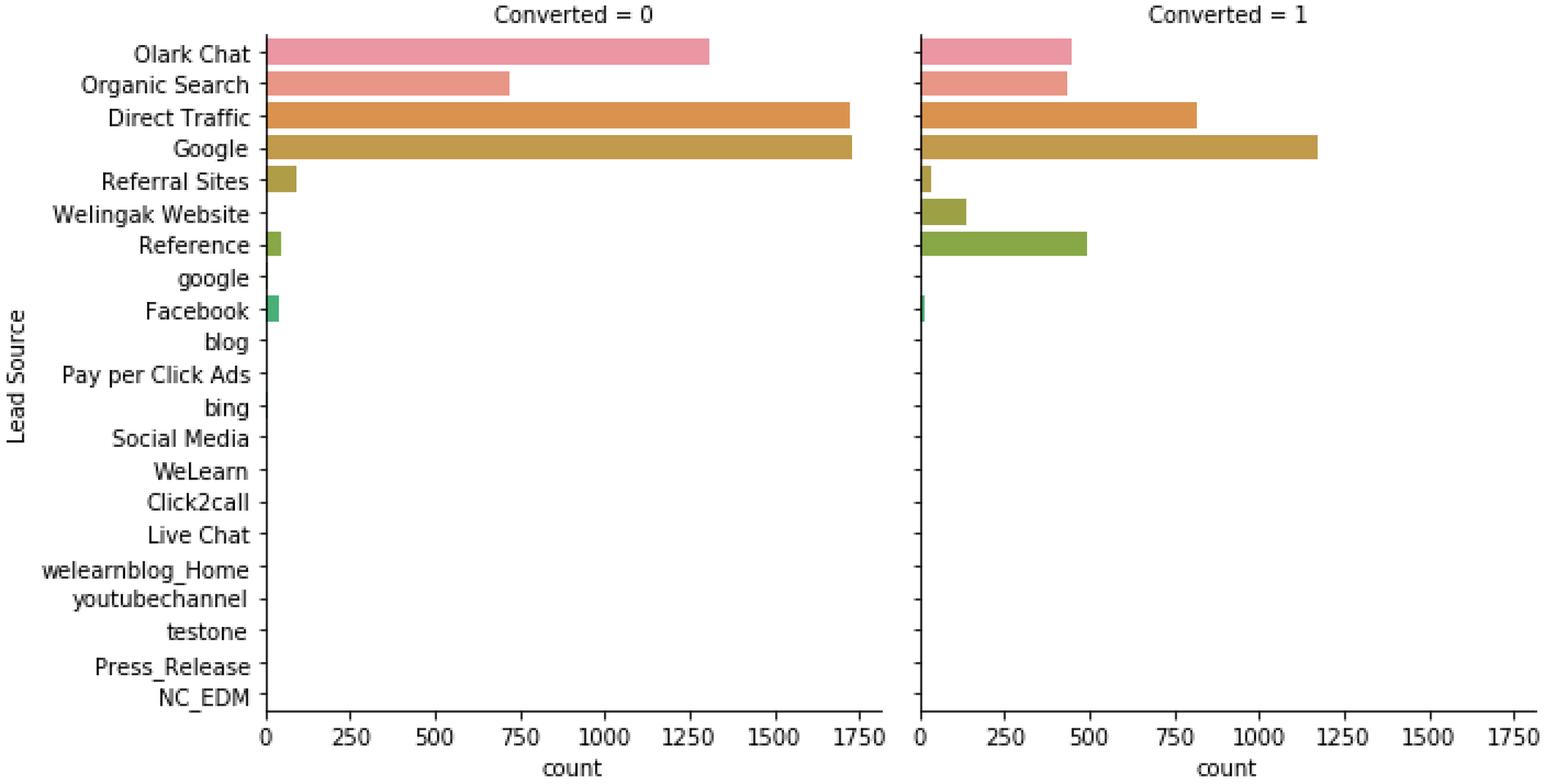
Variation in categorical column
(A free copy of Mastering The Interview)
for Convert and Non-Convert



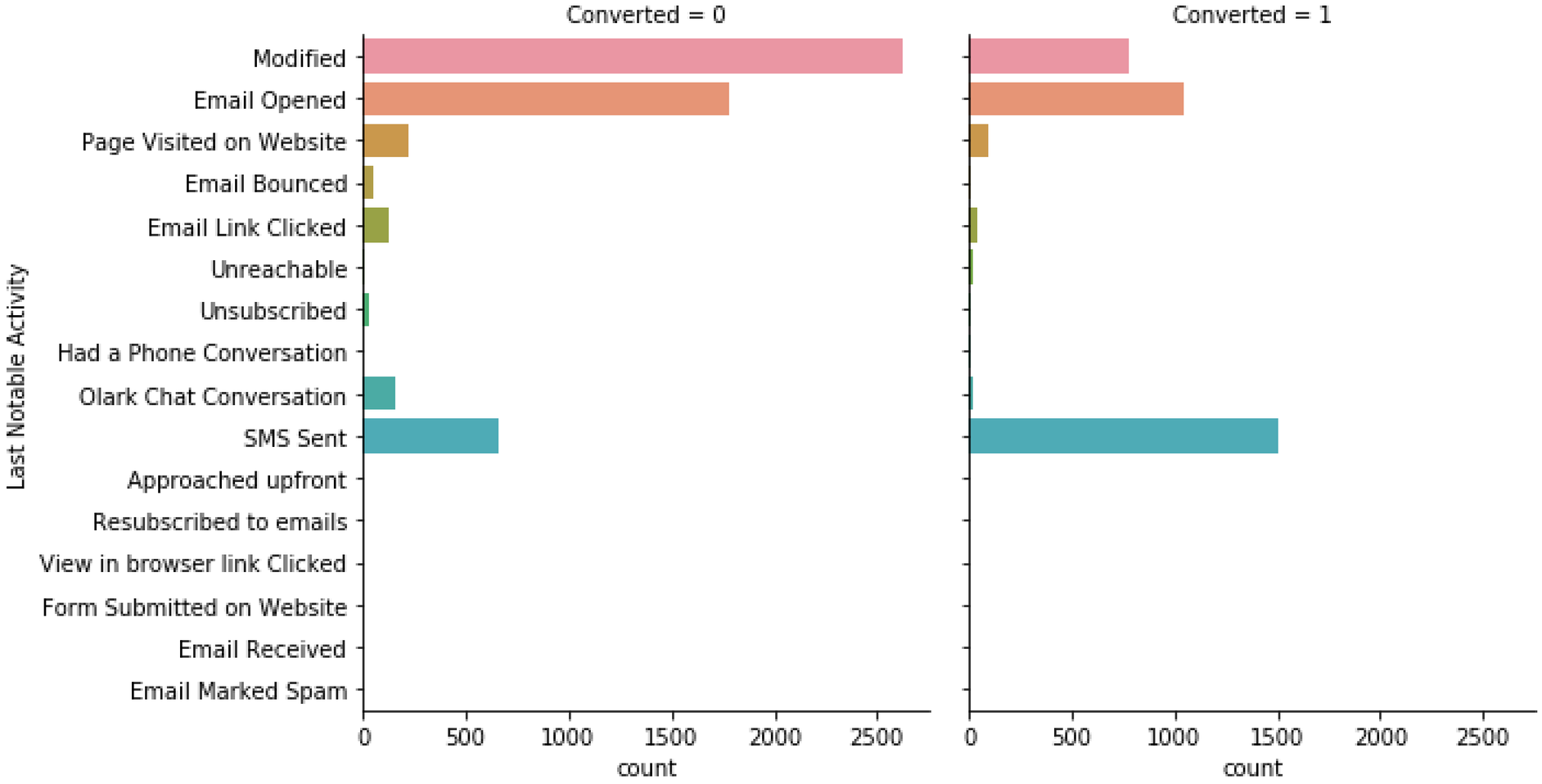
Variation in categorical column
(Do Not Email)
for Convert and Non-Convert



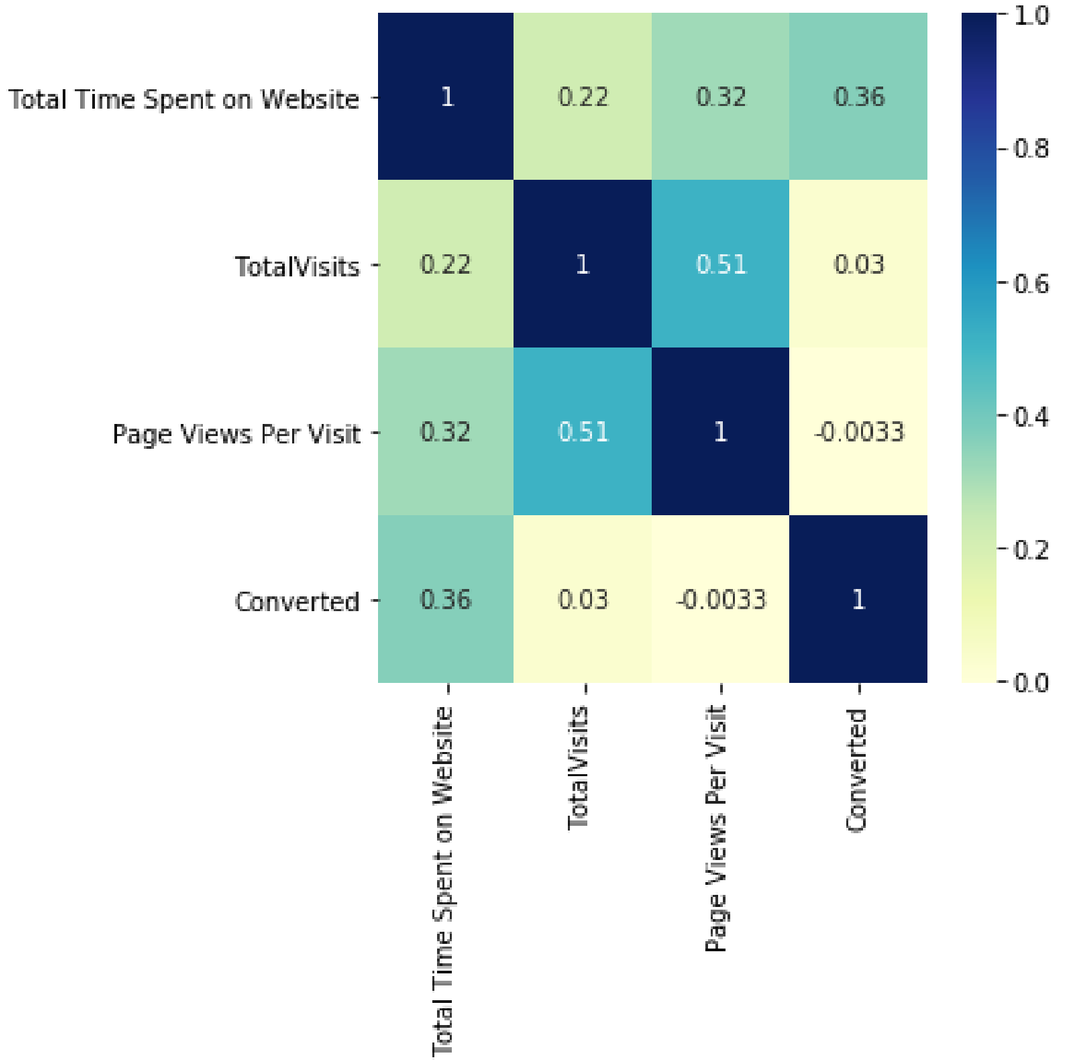
Variation in categorical column (Lead Origin) for Convert and Non-Convert

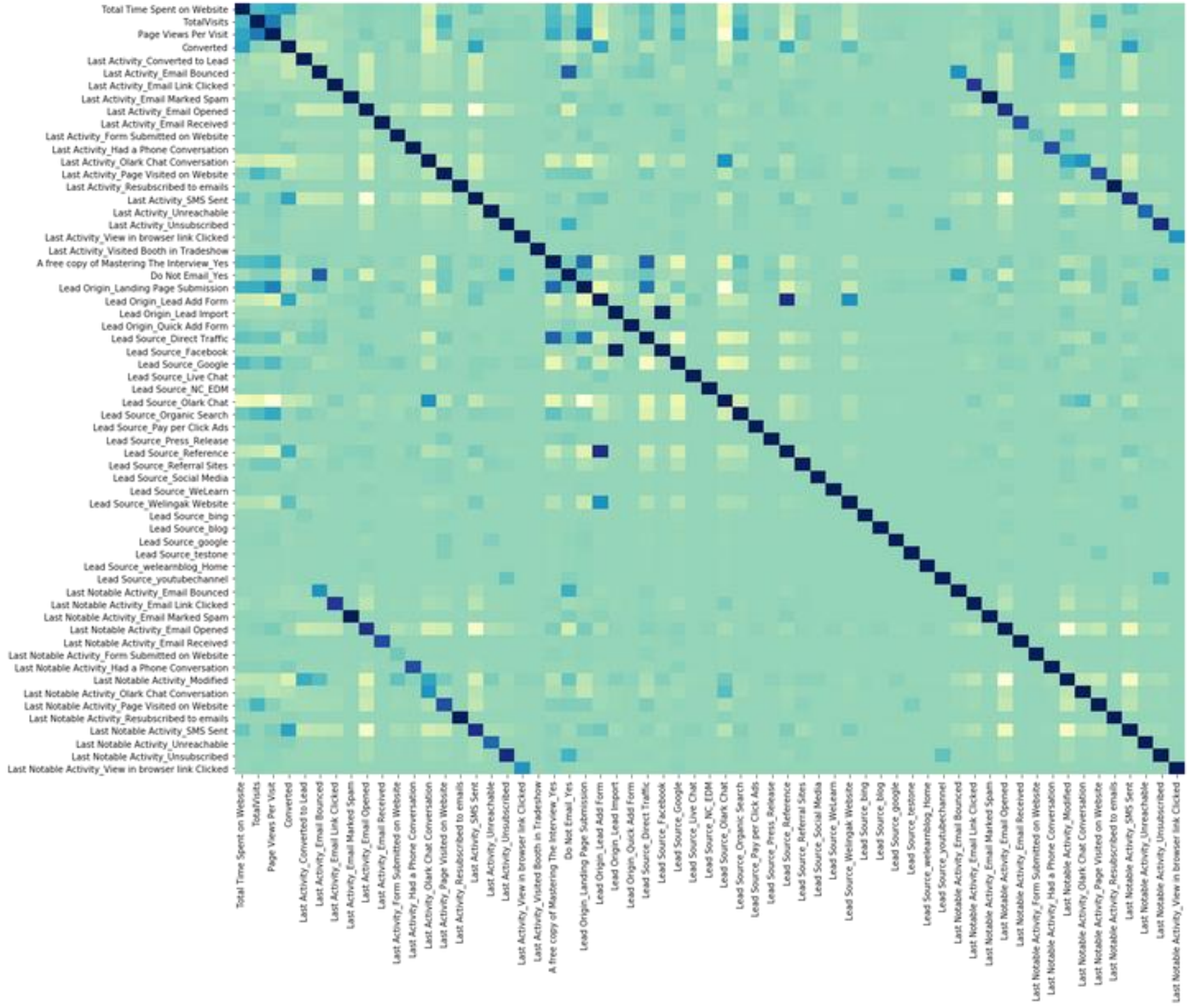


Variation in categorical column (Lead Source) for Convert and Non-Convert

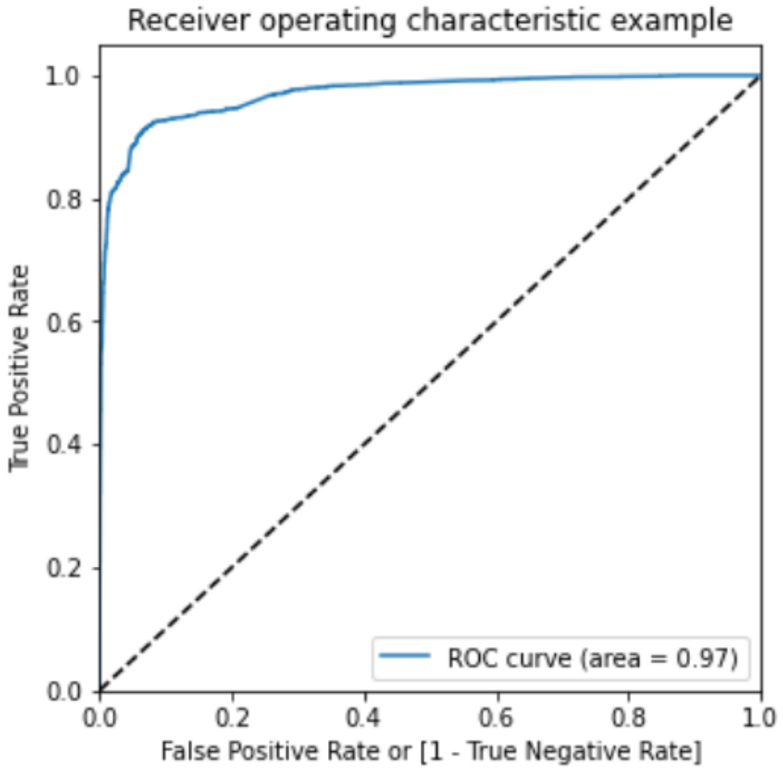


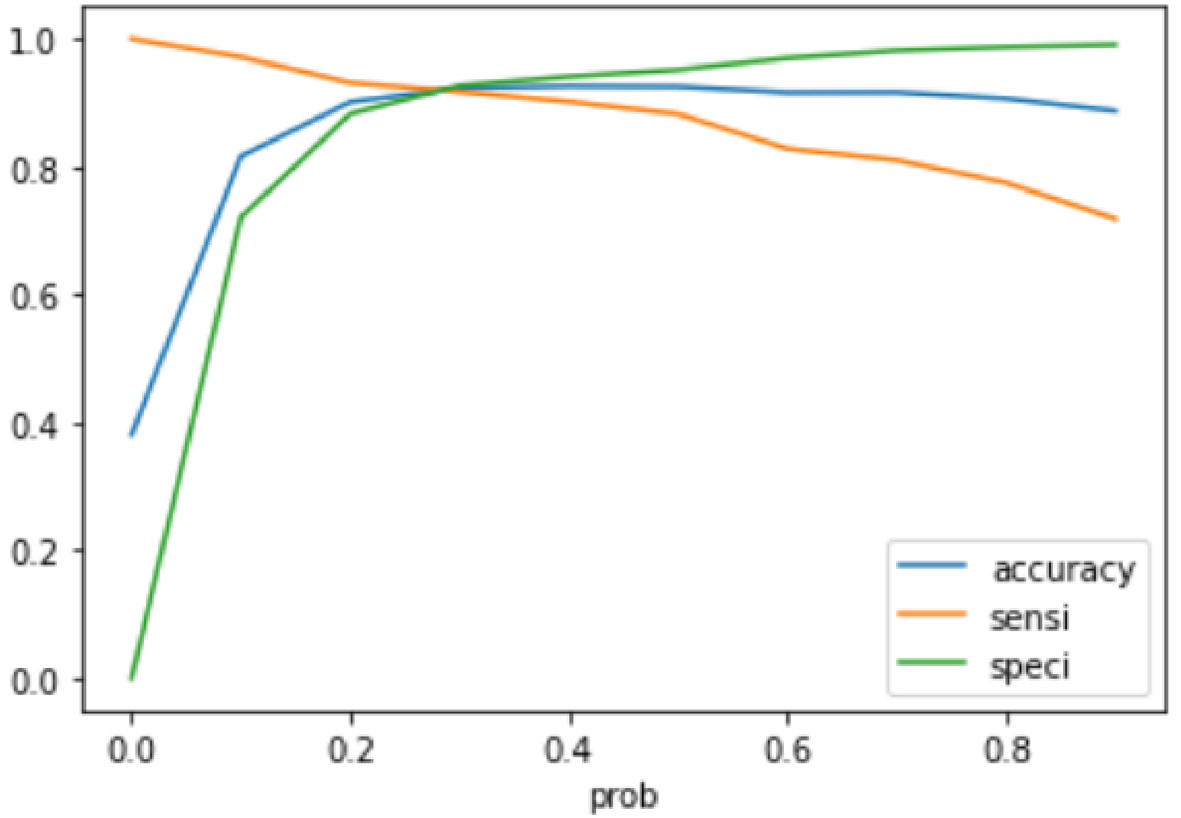
Variation in categorical column (Last Notable Activity) for Convert and Non-Convert





Heat Map Analysis



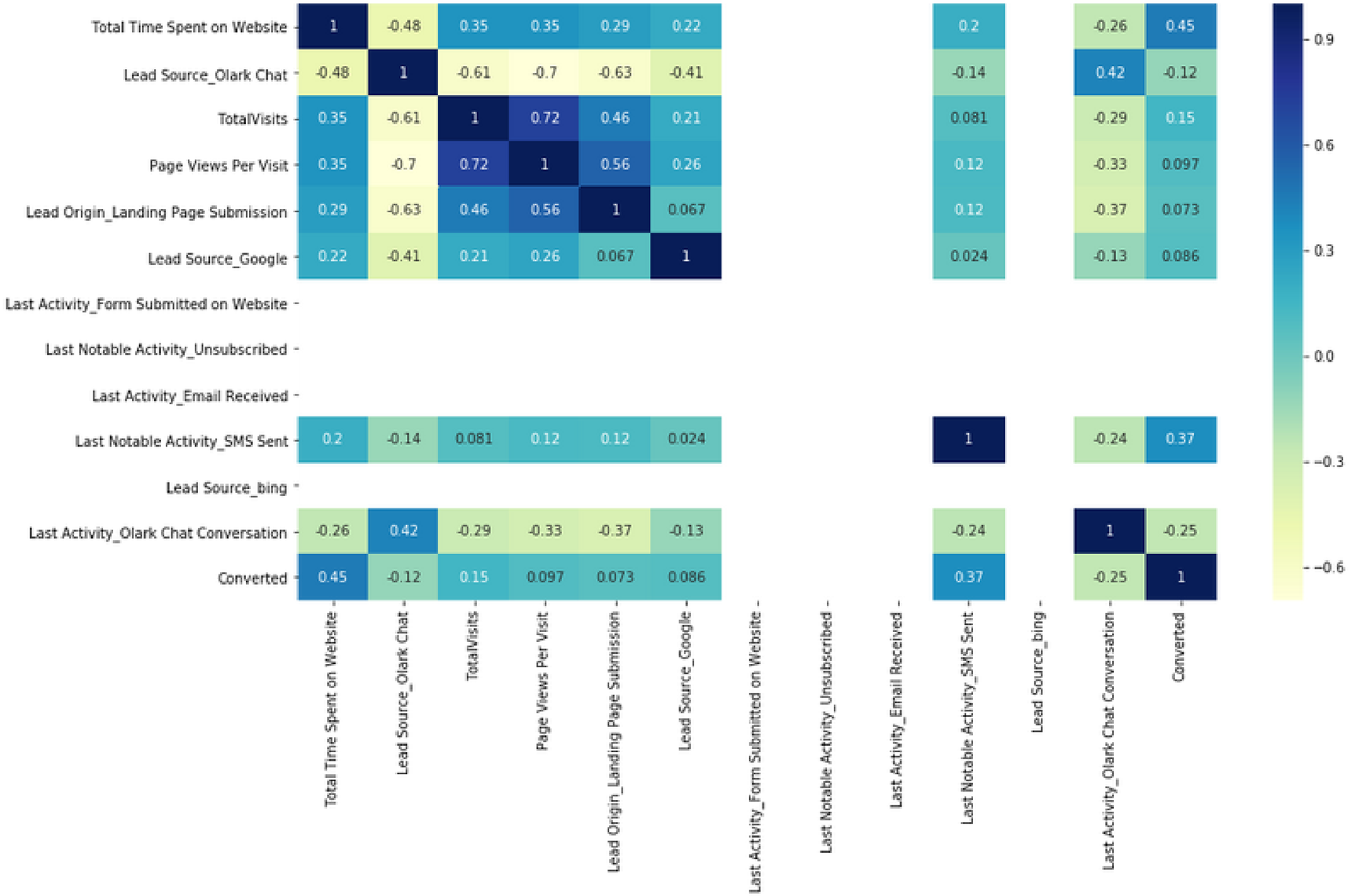


Linear Regression Final Model

Area under ROC =0.97

Intermediate cut-off =0.3

Final cut-off =0.3



Final Model Heat Map Analysis

INFERENCE

MODEL ANLYSIS

Model Performance

Train Data:

Accuracy : 92.29%

Sensitivity : 91.70%

Specificity : 92.66%

Test Data:

Accuracy : 92.78%

Sensitivity : 91.98%

Specificity : 93.26%

INSIGHTS

Business Insights Derived from our Model

Top 3 variables in model, that contribute towards lead conversion are:

- Total Time Spent on Website
- Last Notable Activity_SMS Sent
- TotalVisits

INSIGHTS

Business Insights Derived from our Model

Top 3 variables in my model, that should be focused are:

- Last Activity_SMS Sent (positively impacting)
- Last Activity_Olark Chat Conversation (negatively impacting)
- Lead Source_Olark Chat (negatively impacting)

Conclusion 1(LR Model)

Our Logistic Regression Model accurate enough,
92.78 % Accuracy on Test Set,
91.98 % Sensitivity &
93.26 % Specificity.
The Model seems to predict the Conversion Rate very well and we should be able to give the CEO confidence in making good calls based on this model
We can vary these parameters by varying the cut-off value and thus predict Hot leads based on scenarios like availability of extra resources and vice-versa.

Recommendation's

X Education Company needs to focus on following key aspects to improve the overall conversion rate:

- Increase user engagement on their website
- Increase on sending SMS notifications
- Get Total visits increased by advertising etc.
- Improve the Olark Chat service