

X EDUCATION – LEAD SCORE CASE STUDY

Identification of Hot Leads to focus more on them and thus enhancing the conversion ratio for X Education

HARSHA INDURKAR

BACKGROUND

- ❖ X Education , An education company named 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 some videos
- ❖ When these people fill up a form providing their email address or phone number, they are classified to be 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 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 leads, 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 to 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 has given a ballpark of the target lead conversion rate to be 80%.

LEAD CONVERSION PROCESS

Lead to
Conversion
Process

Lead Generation:

1. Ads on websites like Google
2. Referrals

Visit to X
Education
website by these
potential
customers
(professionals)

Visitors either
provide Email id &
Contact Details
Or
View videos etc

Tele calling and
Emailing activity
to all the leads

~30% leads get
converted

Proposed Solution:
A model to filter leads
so that leads to
conversion ratio is
80%+

PROPOSED SOLUTION

Selection of Hot Leads

Leads Clustering

We cluster the leads into certain categories based on their tendency or probability to convert, thus, getting a smaller section of hot leads to focus more on.

Communicating with Hot Leads

Focus Communication

Since we would have a smaller set of leads to have communication with, we might make more impact with effective communication.

Conversion of Hot Leads

Increase conversion

Since we focussed on hot leads, which were more probable to convert, we would have a better conversion rate, and hence we can achieve the 80% target.

IMPLEMENTATION



**Selection of top 25
features using RFE**

**Reduction of columns
and Model re-building**

**Verifying our Final
Model Accuracy etc.
with model built with
PCA**

**Feature
Selection**

**Model
Building**

**Model
Improvement**

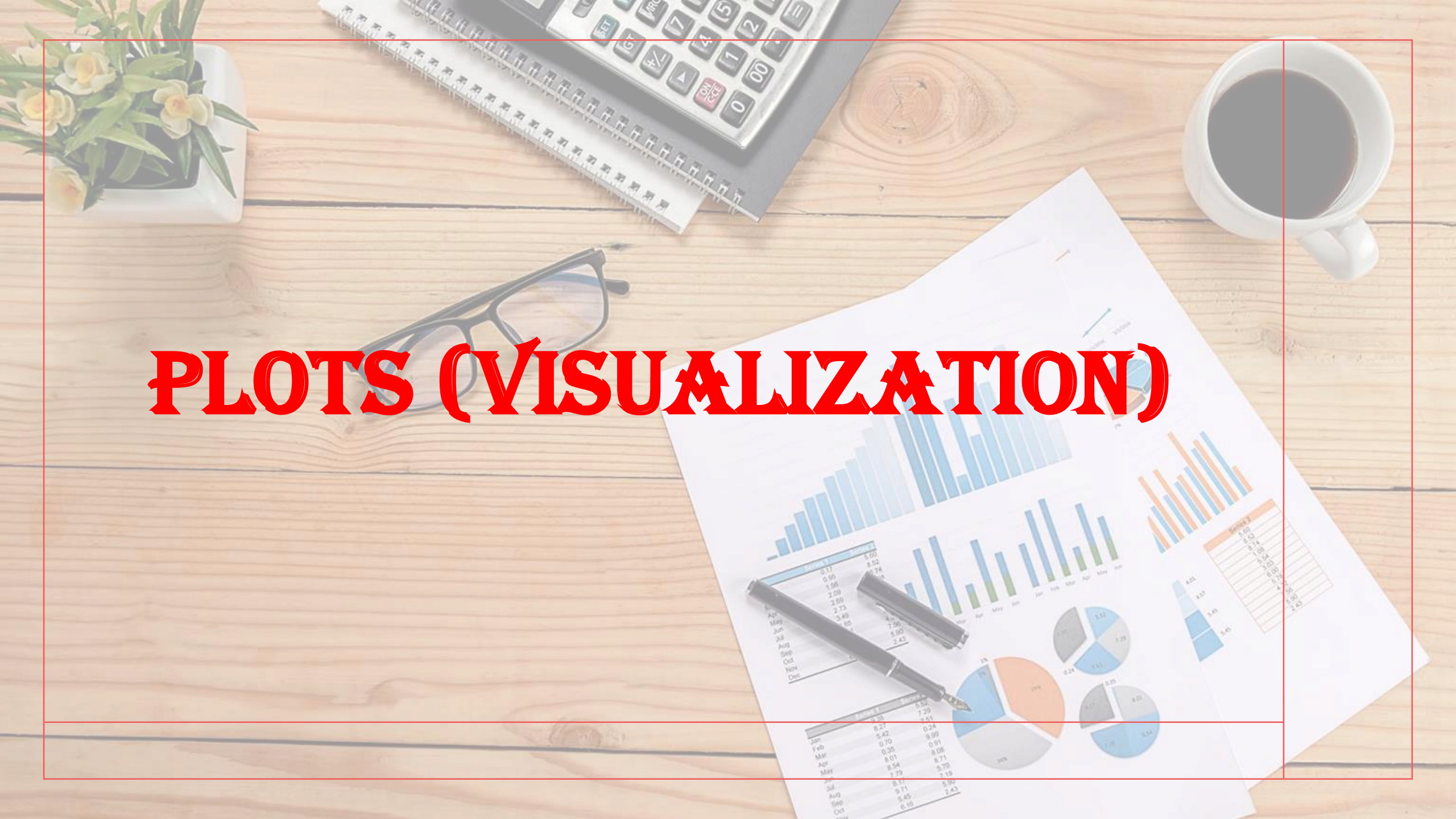
Final Model

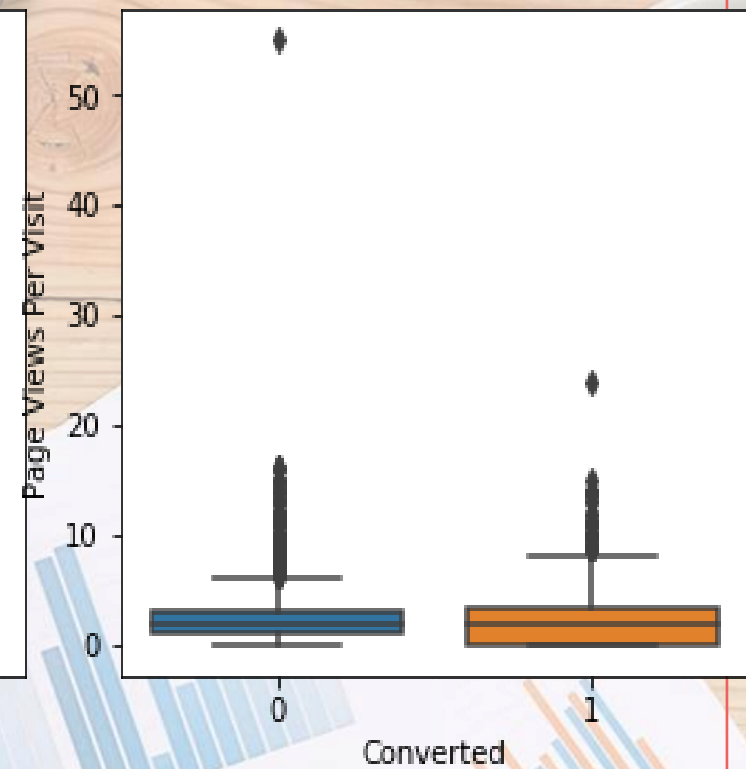
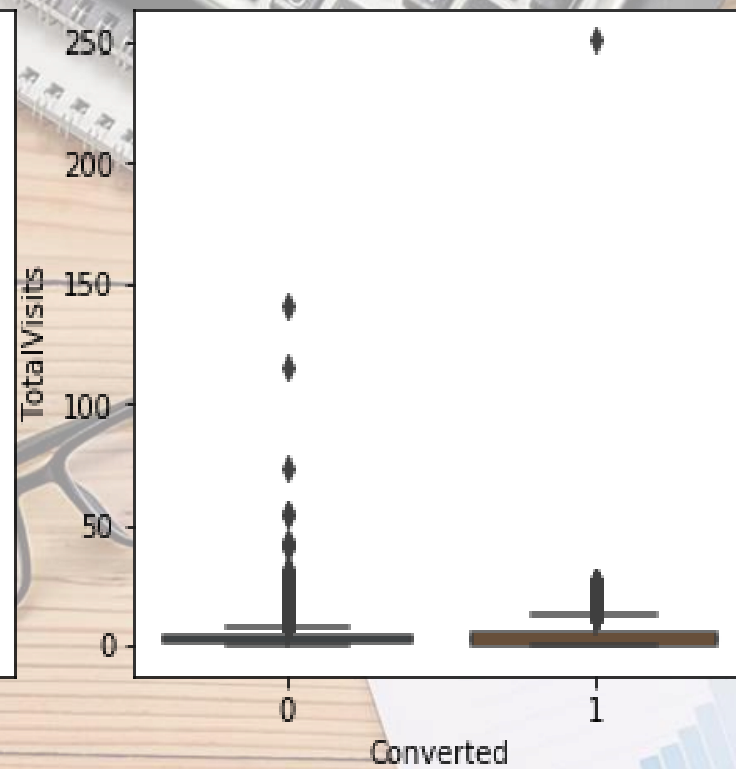
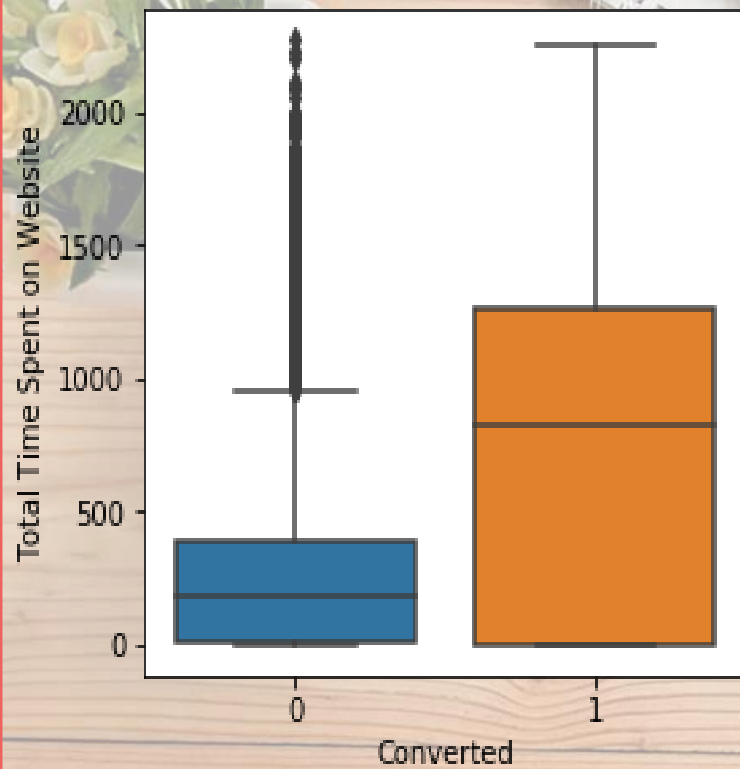
**Verifying
with PCA**

**Model building using RFE for
selected columns**

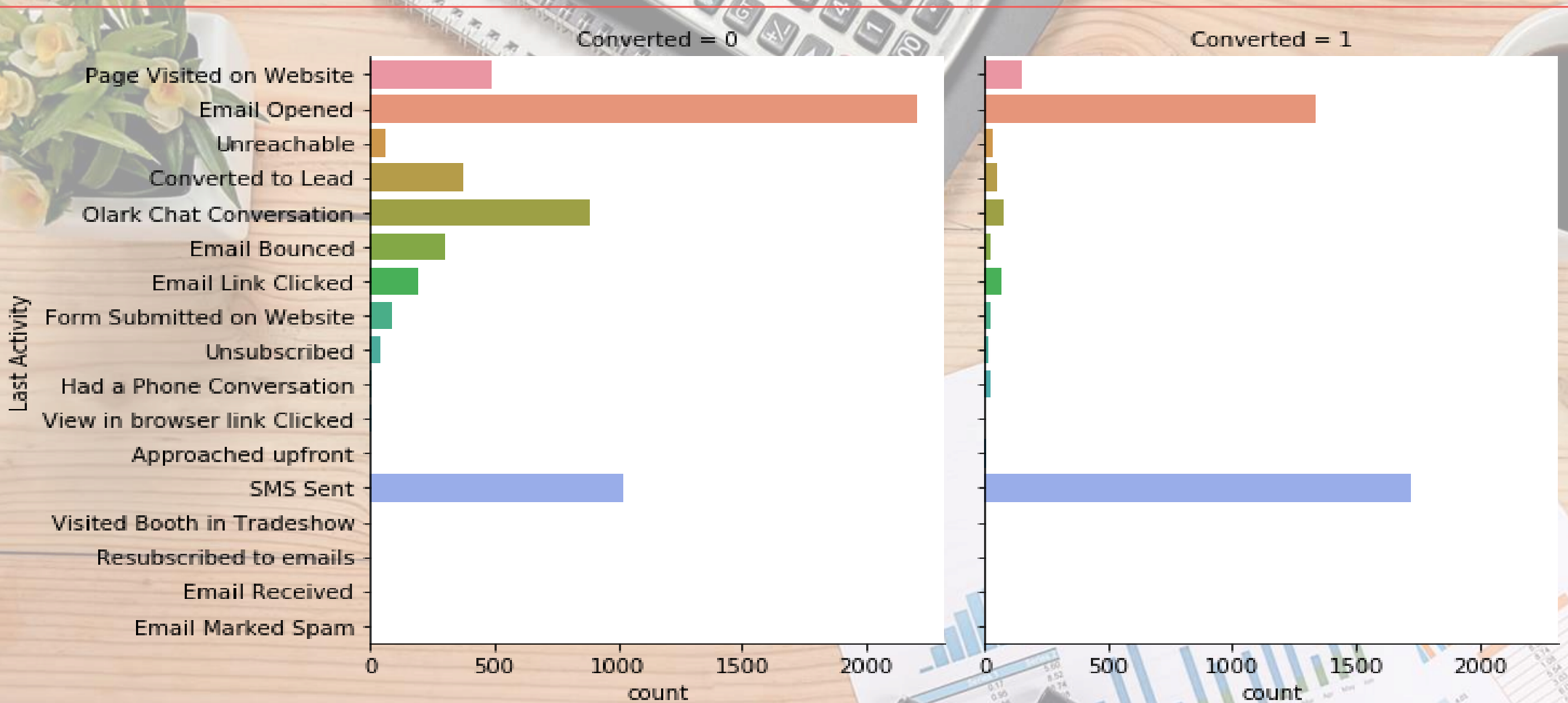
**Final Model Analysis
and performance on
Test Data**

PLOTS (VISUALIZATION)





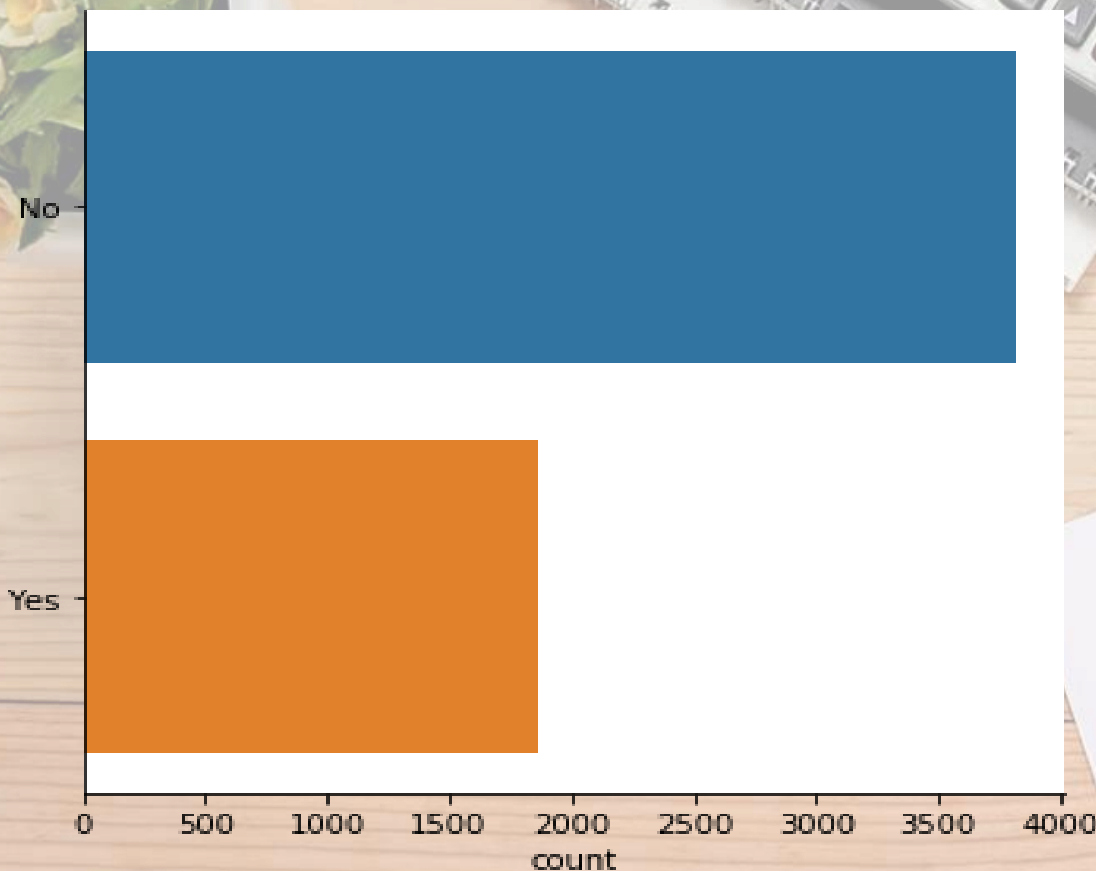
EDA plots depicting variation in numerical columns for those who Converted and those who didn't.



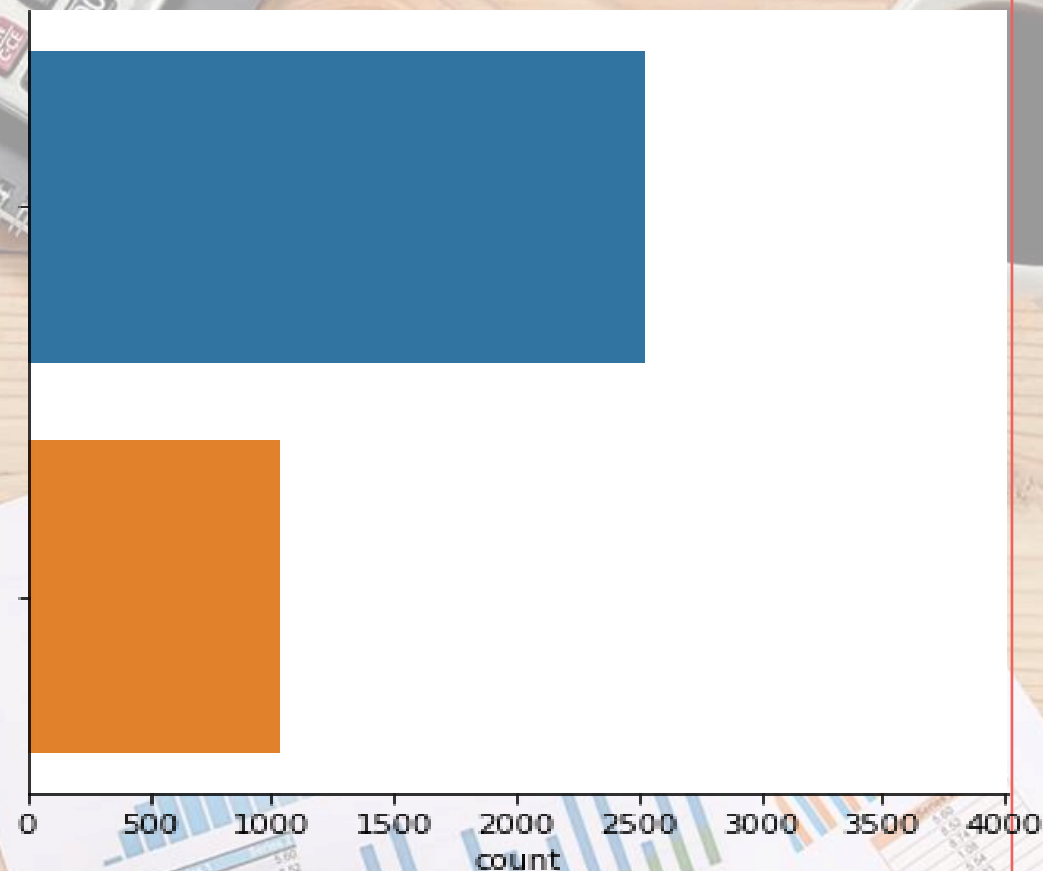
EDA plots depicting variation in categorical column (Last Activity) for those who Converted and those who didn't.

Converted = 0

A free copy of Mastering The Interview

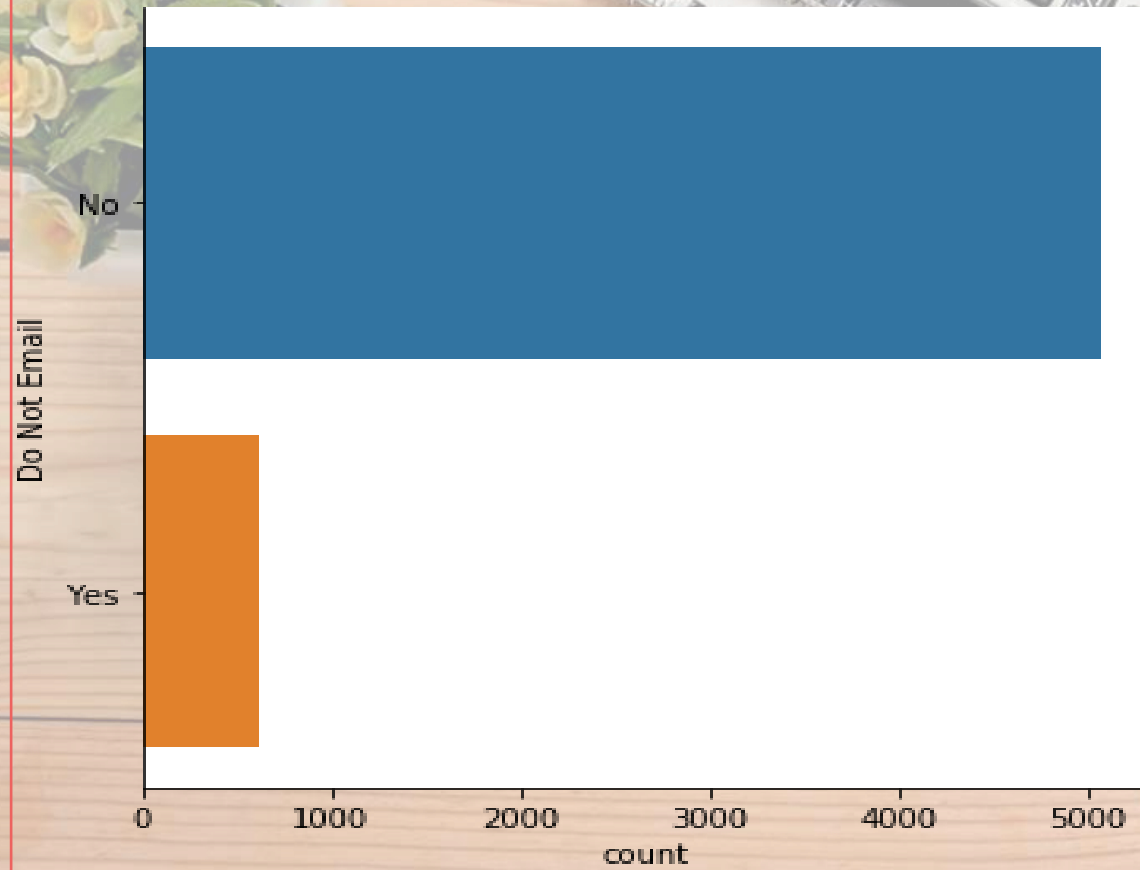


Converted = 1

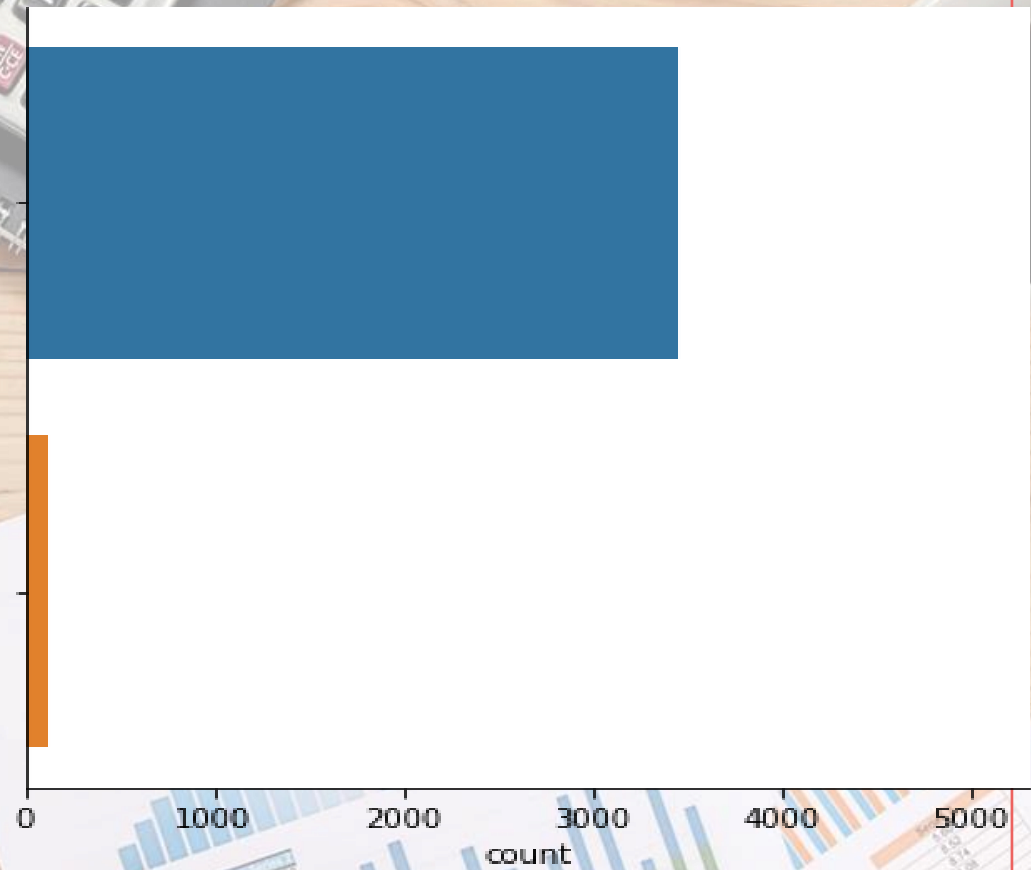


EDA plots depicting variation in categorical column (A free copy of Mastering The Interview) for those who Converted and those who didn't.

Converted = 0



Converted = 1

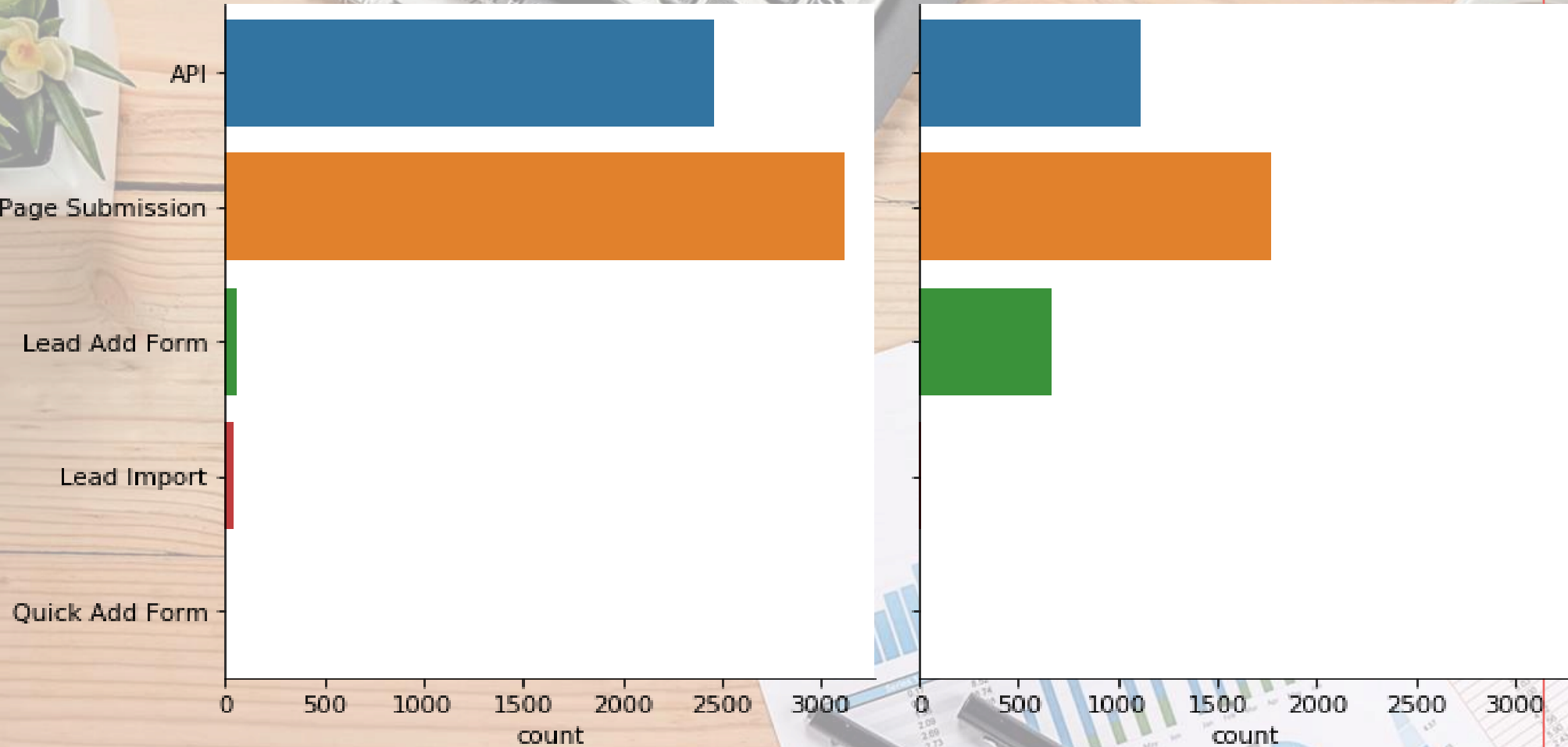


EDA plots depicting variation in categorical column (Do Not Email) for those who Converted and those who didn't.

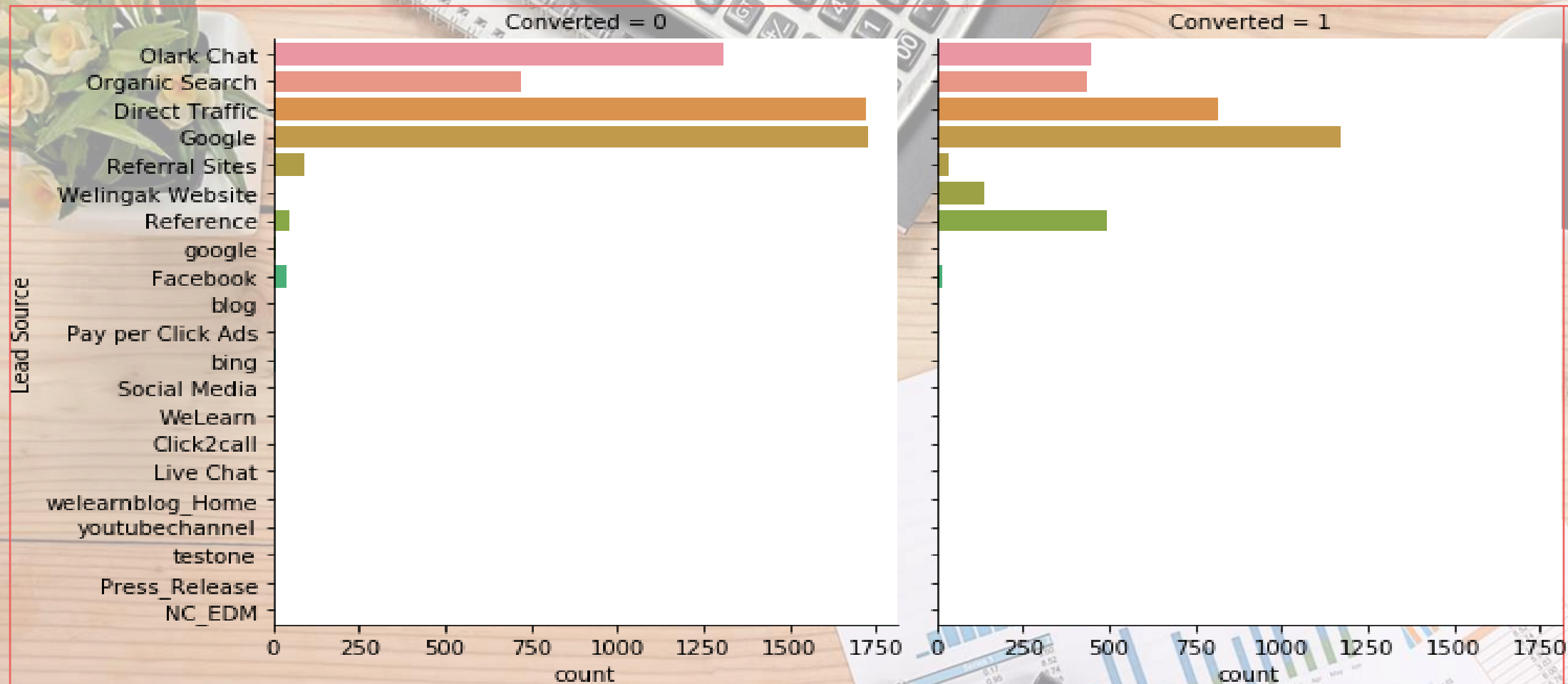
Lead Origin

Converted = 0

Converted = 1



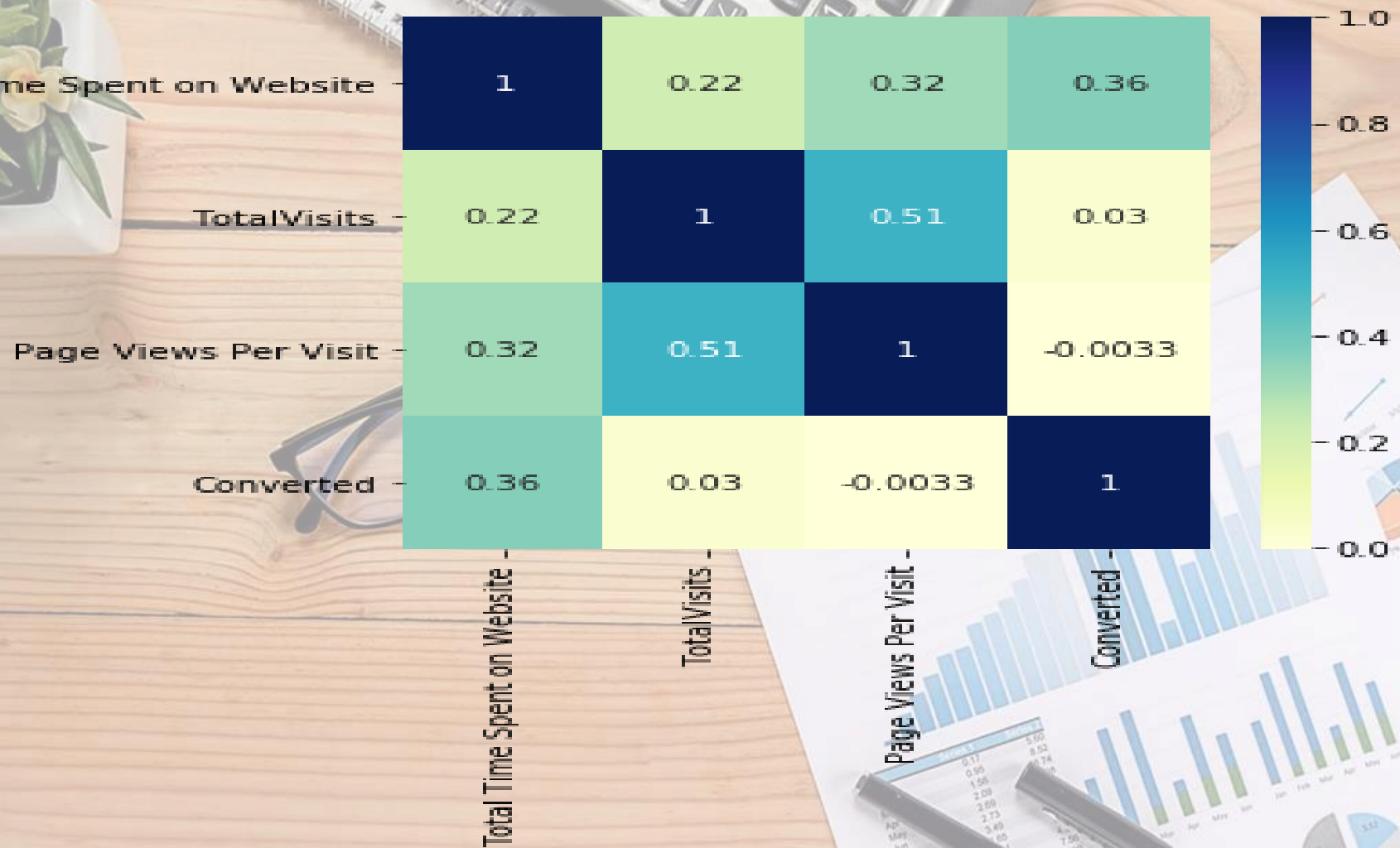
EDA plots depicting variation in categorical column (Lead Origin) for those who Converted and those who didn't.



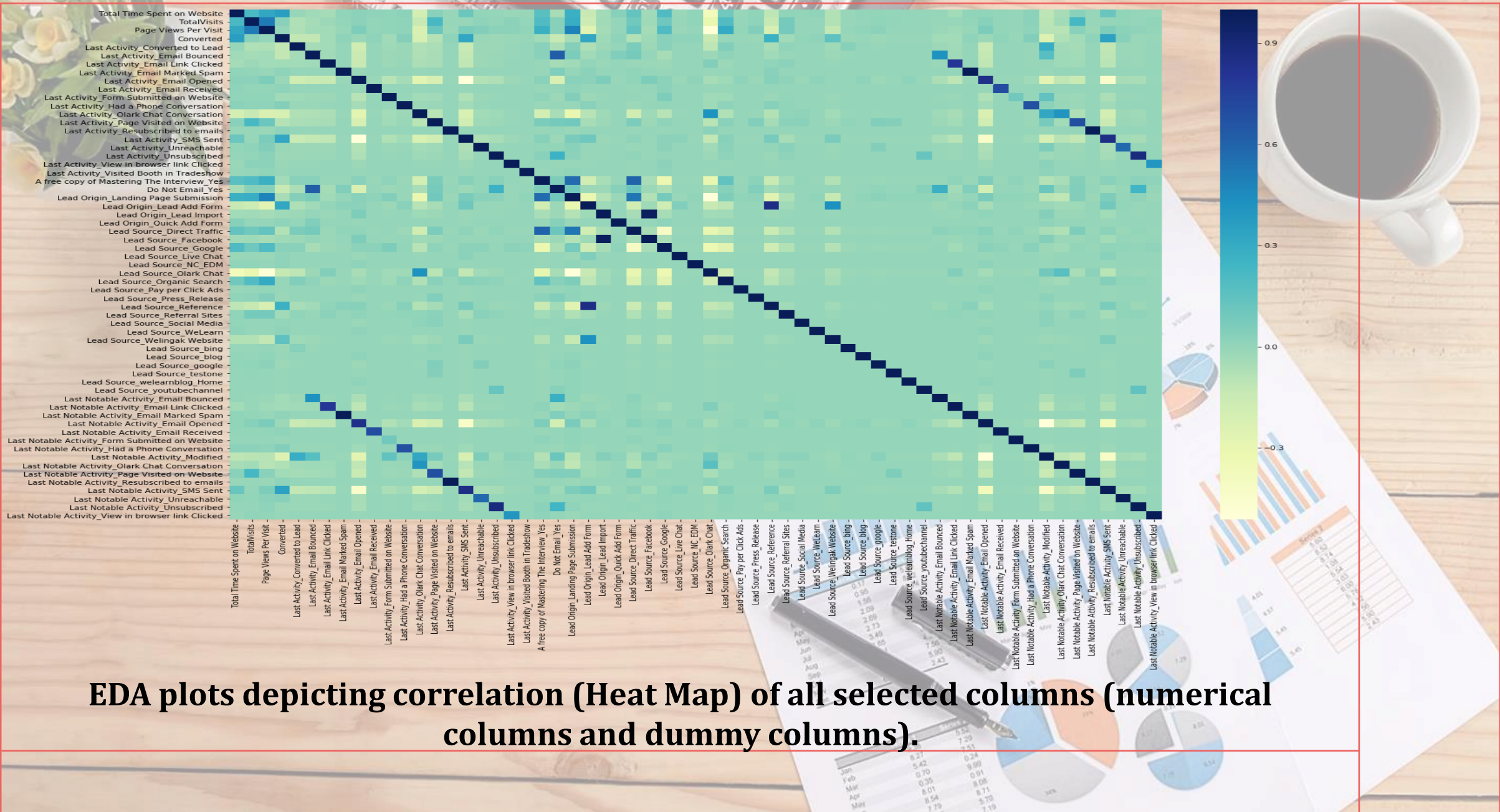
EDA plots depicting variation in categorical column (Lead Source) for those who Converted and those who didn't.



EDA plots depicting variation in categorical column (Last Notable Activity) for those who Converted and those who didn't.

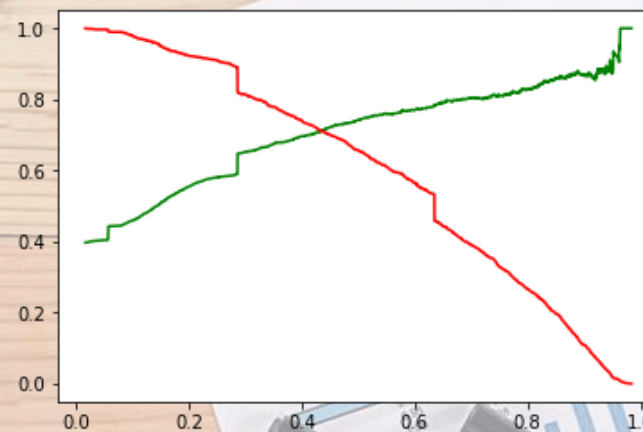
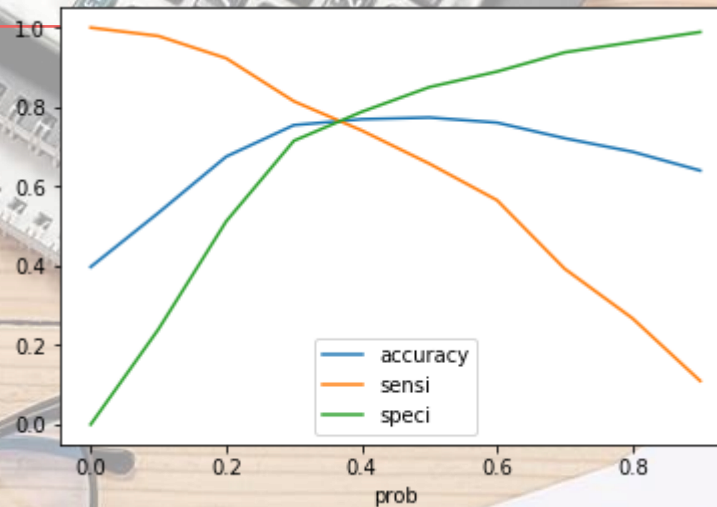
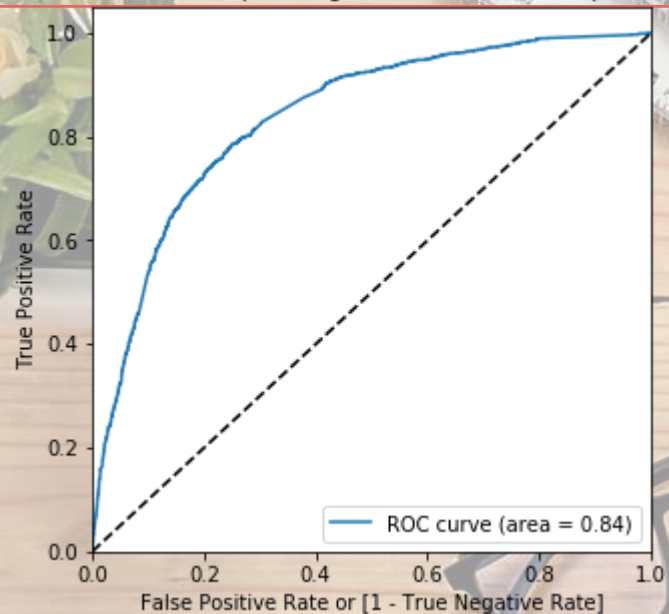


EDA plots depicting correlation (Heat Map) of all selected numerical columns.



EDA plots depicting correlation (Heat Map) of all selected columns (numerical columns and dummy columns).

Receiver operating characteristic example

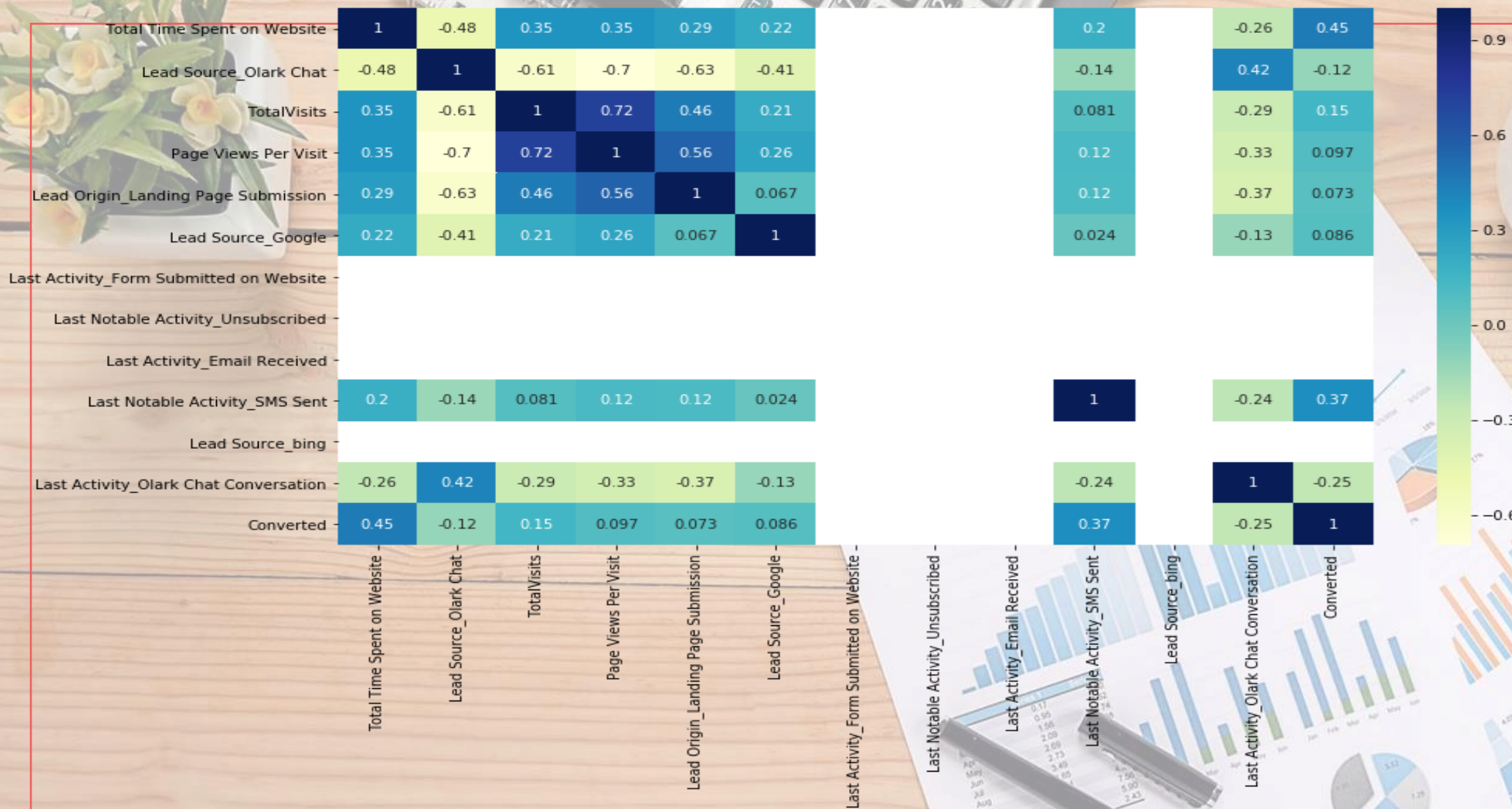


Linear Regression Final Model Parameters

Area under ROC = 0.84

Intermediate cut-off = 0.35

Final cut-off = 0.42



EDA plots depicting correlation (Heat Map) of all selected columns (numerical columns and dummy columns) in our final Model.

CONCLUSION





MODEL ANALYSIS

PERFORMANCE OF THE MODEL

- ❖ Overall accuracy on Test set: 0.786
- ❖ Sensitivity of our logistic regression model: 0.733
- ❖ Specificity of our logistic regression model: 0.823

INFERENCES FROM THE MODEL

BUSINESS INSIGHTS DERIVED FROM THE MODEL

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

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

Top 3 variables in model, that should be focused on:

- ❖ **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 is decent and accurate enough, when compared to the model derived using PCA, with 78.6 % Accuracy on Test Set, 73.3 % Sensitivity and 82.3 % Specificity. 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.

CONCLUSION 2 (RECOMMENDATION)

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

- Increase user engagement on their website since this helps in higher conversion
- Increase on sending SMS notifications since this helps in higher conversion
- Get Total visits increased by advertising etc. since this helps in higher conversion
- Improve the Olark Chat service since this is affecting the conversion negatively