Lead Score Case Study.

Presented by Akshata Gurayya Annigeri.

Problem Statement

X company sells online courses to industry professionals. The company markets it's courses on several search engines example Google.

Leads come through various ways like advertisements, past referrals, google searches etc. Although many leads come in the rate of conversion of leads to paid customers is less. Hence , X company wants to start concentrating more on the leads that are highly likable to get converted as paid customers.

X company has the Lead conversion rate of 30%. The implementation process of lead generating attributes are not helpful in helping conversions.

Business Goal

X company needs help targeting the most promising leads that are likely to be converted to paying customers.

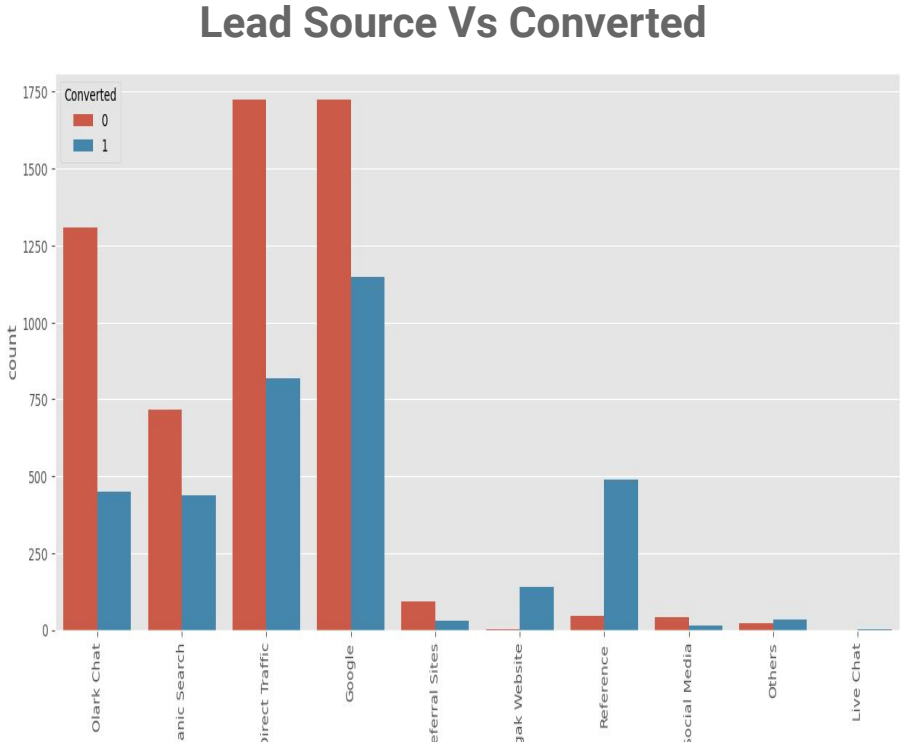
The company is in need of a model where a lead score is assigned to each of the leads such that the customer with higher lead value have a higher conversion chance.

The CEO, in particular has given a ballpark of the target lead conversion rate to be around 80%.

Strategy

- **Import Data**
- **Clean and prepare imported data for further analysis.**
- **Exploratory Data Analysis.**
- **Scaling**
- **Preparing the Data for model building.**
- **Build the Logistic Regression model.**
- **Evaluate the model using different metrics like Specificity, Sensitivity, Precision and Recall.**
- **Applying the best model in Test data.**

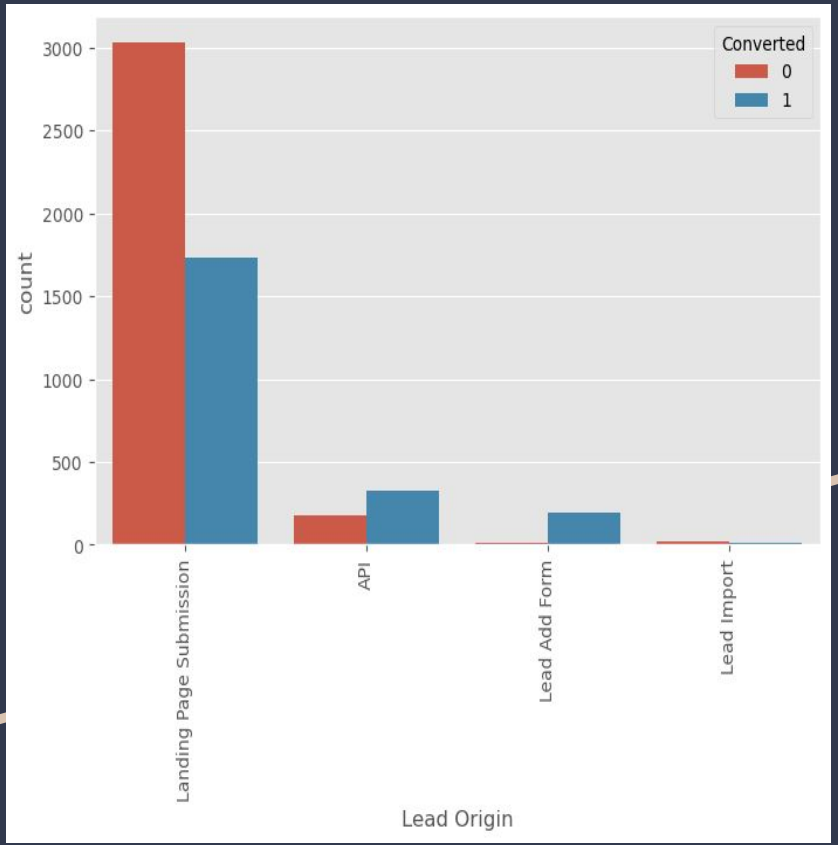
Exploratory Data Analysis.



Inferences:

- **Maximum number of leads are generated by Google and Direct traffic.**
- **Conversion Rate of reference leads and leads through welingak website is high.**
- **To improve overall lead conversion rate, focus should be on improving lead conversion of olark chat, organic search, direct traffic, and google leads and generate more leads from reference and welingak website.**

Lead Origin Vs Converted



Inference:

1)API and Landing Page Submission bring higher number of leads as well as conversion.

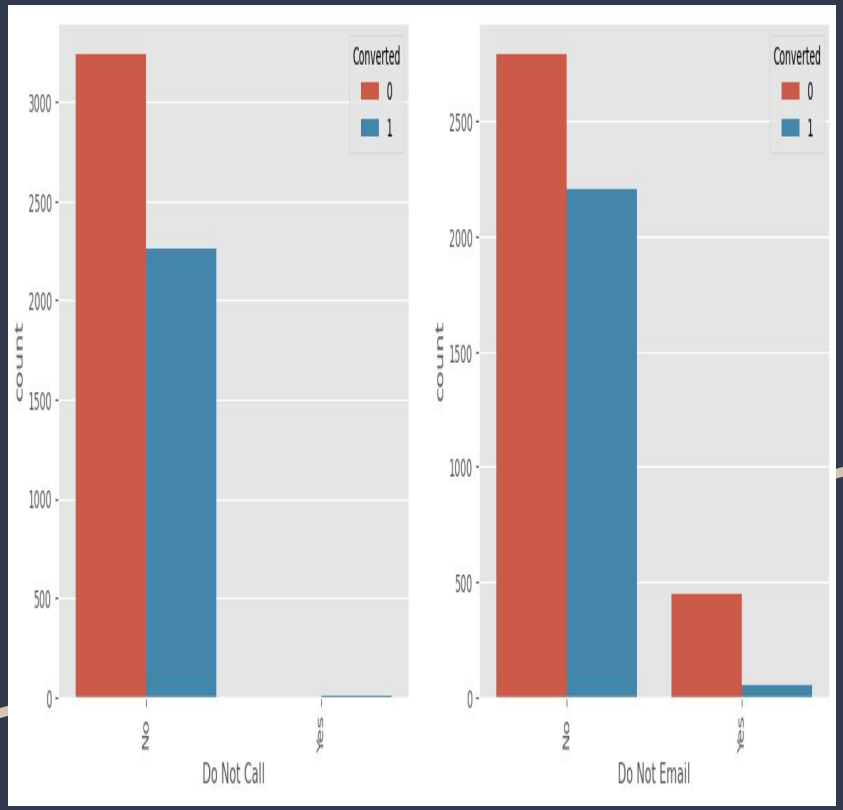
2)Lead Add Form has a very high conversion rate but count of leads are not very high.

3)Lead Import and Quick Add Form get very few leads.

4)In order to improve overall lead conversion rate, we have to improve lead conversion of API and Landing Page Submission origin and generate more leads from Lead Add Form.

Do Not Call
Vs
Converted

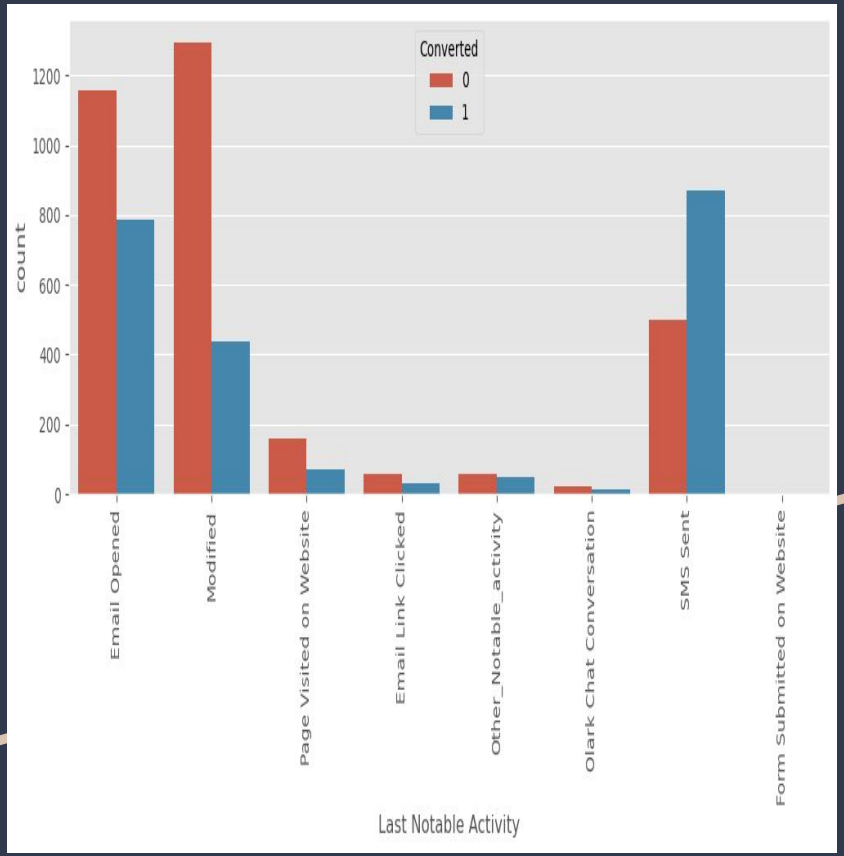
Do Not Email
Vs
Converted



Inferences:

- **Most Leads prefer not to be informed through phone.**
- **Google search has high conversion compared to other modes, whilst references has high conversion rate.**

Last Notable Activity Vs Converted



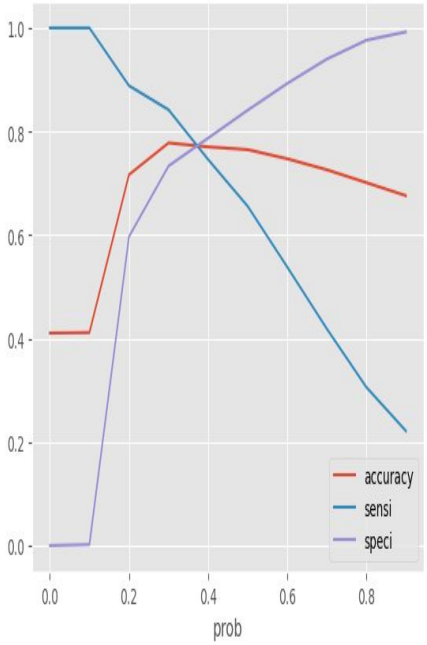
Inferences:

- **SMS is proven to be best method for positive conversions.**
- **Emails also has high conversion.**

Model Building Steps:

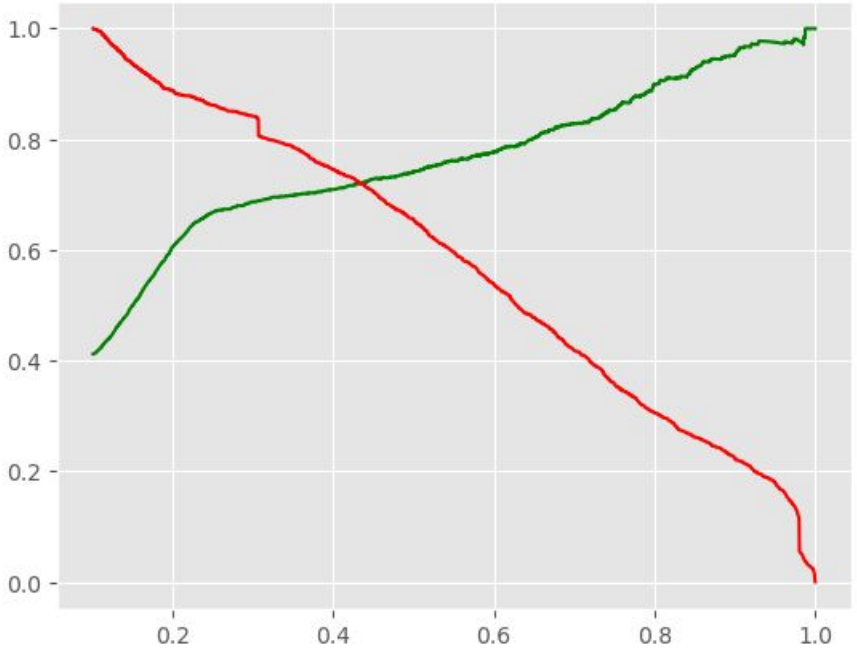
- Splitting the Dataframe into train and test set.
- Scale variables in train set.
- Build the first model.
- Use RFE to eliminate less relevant variables.
- Build next model.
- Eliminate variables based on high p-values.
- Check VIF value for all the existing columns.
- Predict using train set.
- Evaluate the various merics.
- Predict using test set.
- Precision and recall analysis on test prediction.

Model Evaluation (Train)



- Accuracy=>0.7778
- Sensitivity=>0.8419
- Specificity=>0.7332
- Precision=>0.6872
- Recall=>0.8419

Model Evaluation (Test)



- **Accuracy=> 0.7709**
- **Sensitivity=>0.8373**
- **Specificity=>0.7248**
- **Precision=>0.6788**
- **Recall=>0.8373**