

# **LEAD SCORING CASE STUDY**

By:-

Iqbal Rais

Abhas Ramteke

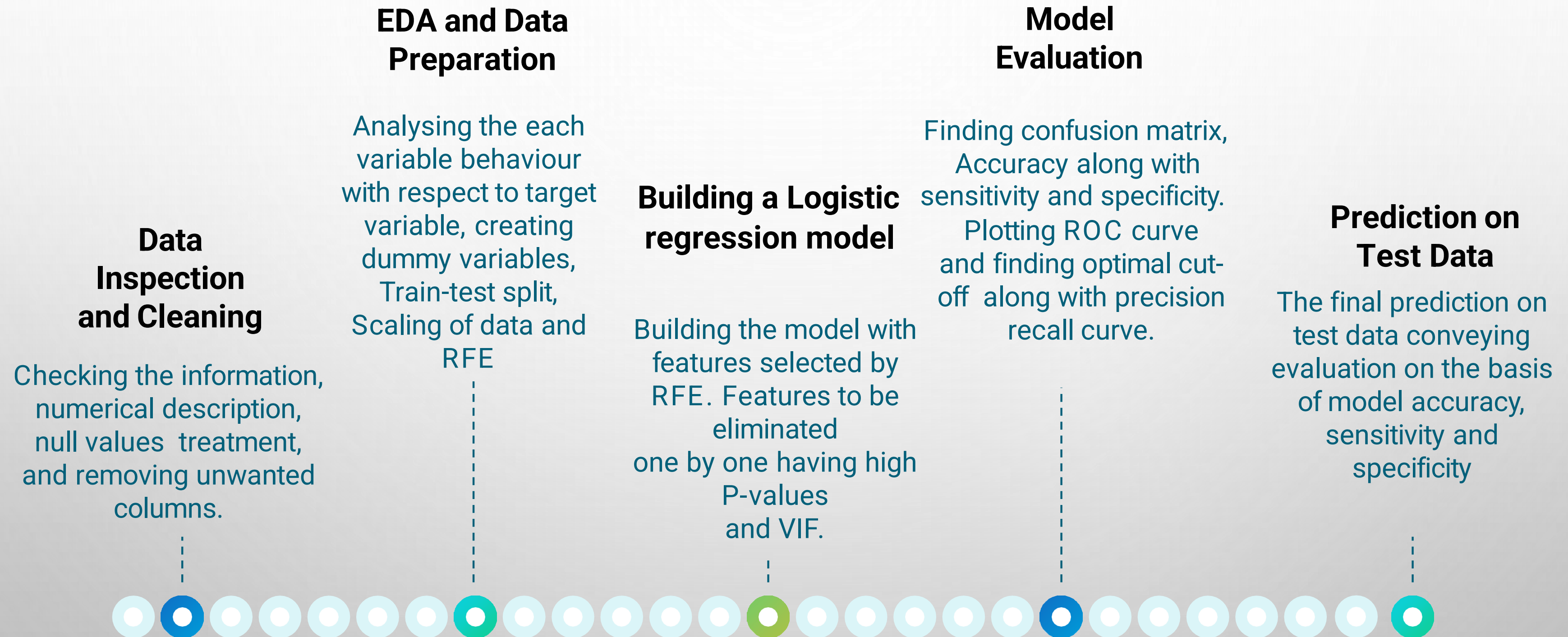
# PROBLEM STATEMENT

- X Education company sells online courses to industry professionals.
- X Education company gets a lots of leads & conversion rate is low as 30%.
- To make this more efficient, company wishes to identify most potential leads, also known as “hot leads” to get segregated as per the sequence of 0-100.
- 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.

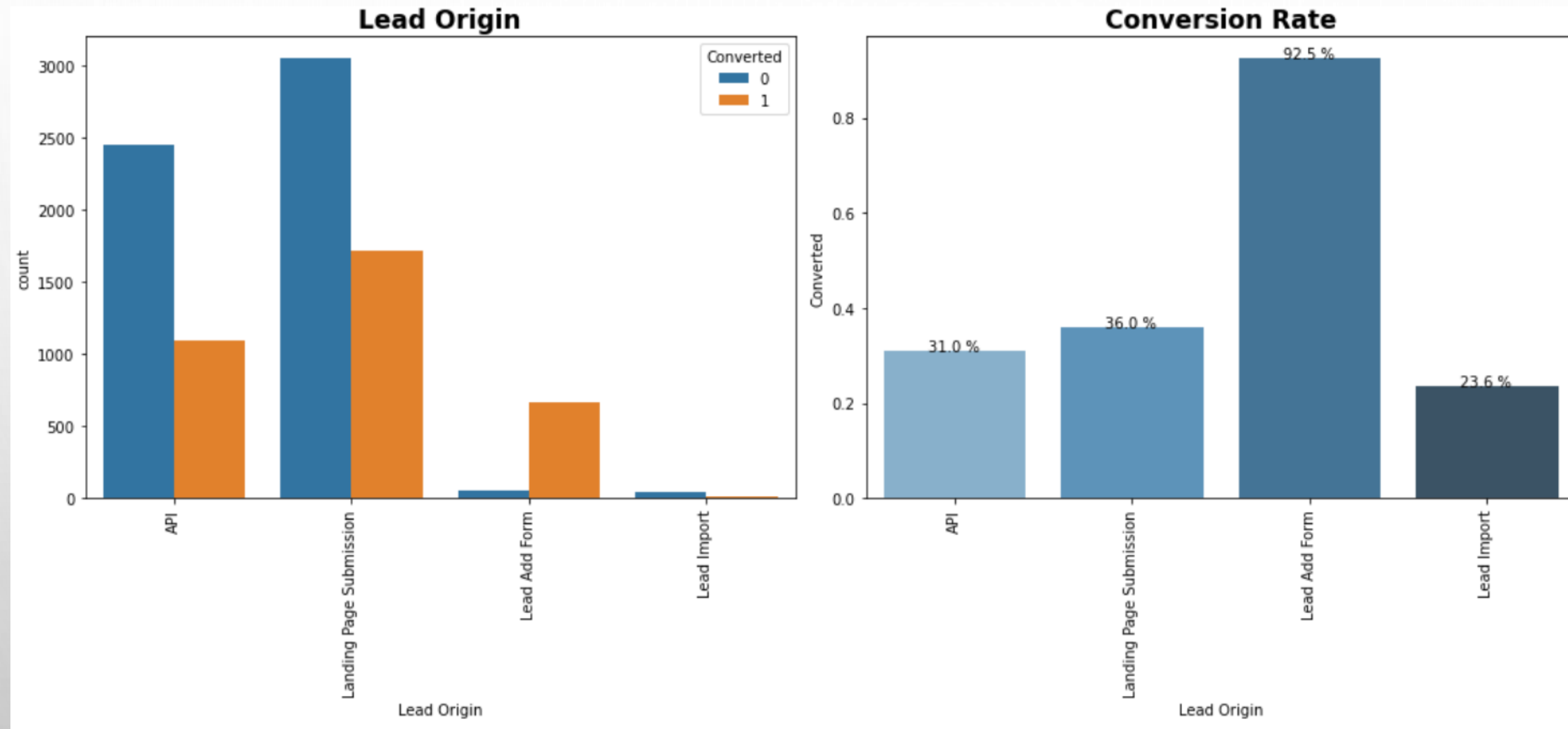
# BUSINESS OBJECTIVE

- To help X-education for selection of Most Promising Leads( Hot Leads).
- To create a logistic regression model and award each lead a score between 0 and 100 that may be utilised by the business to target potential leads and grow sales.

# ANALYSIS APPROACH



# ANALYSIS - LEAD ORIGIN

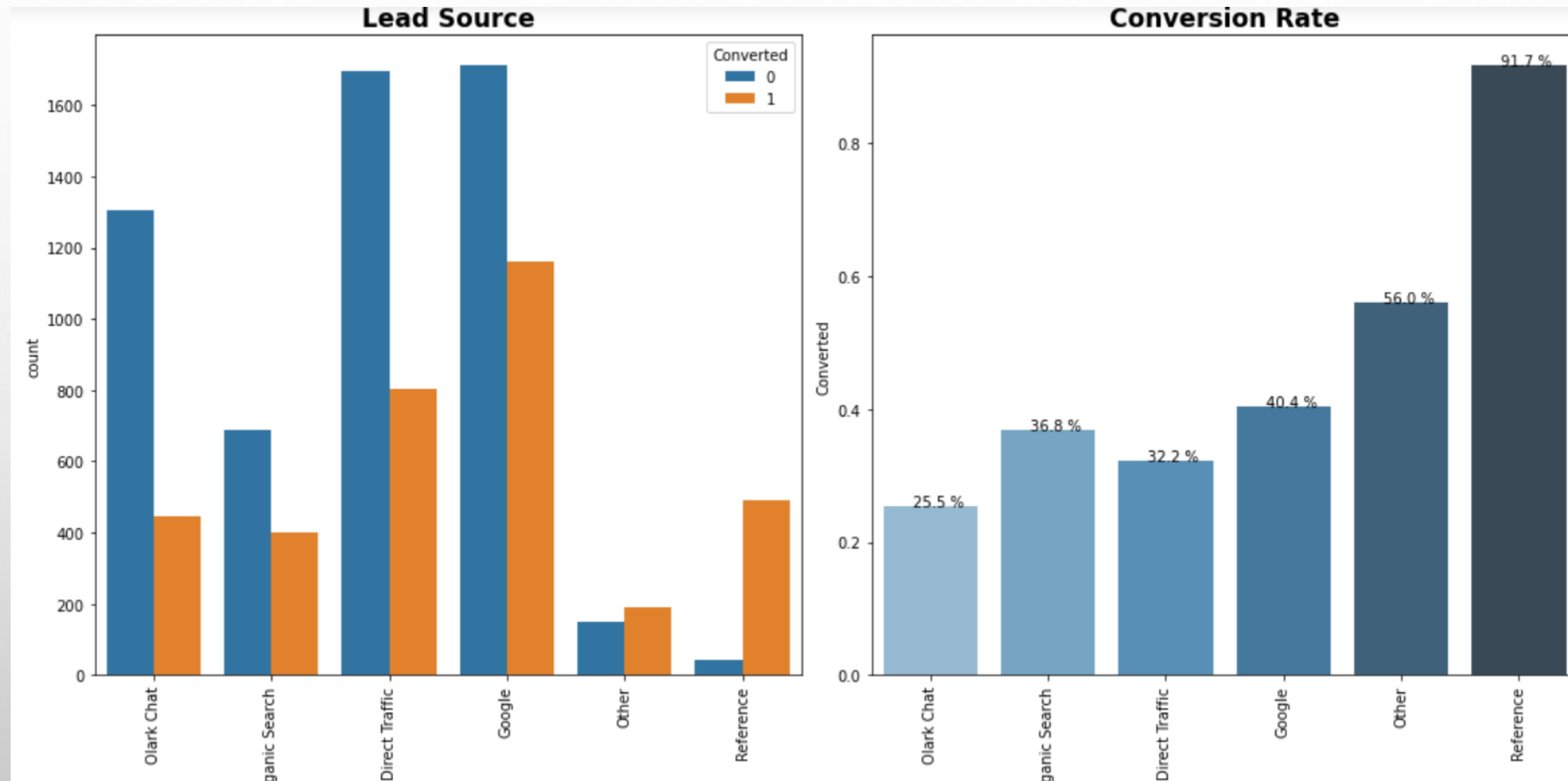


## Observation:

Most of the leads of leads coming from Landing Page Submission from where about 30% of leads are converted. 90% of the 718 leads in this category came via Lead Add Form, which also has the greatest conversion rate. Lead imports are rare and have a poor conversion rate. We must focus on enhancing lead conversion from API and landing page submission origins and increasing lead generation from lead add forms if we want to boost total lead conversion rates.



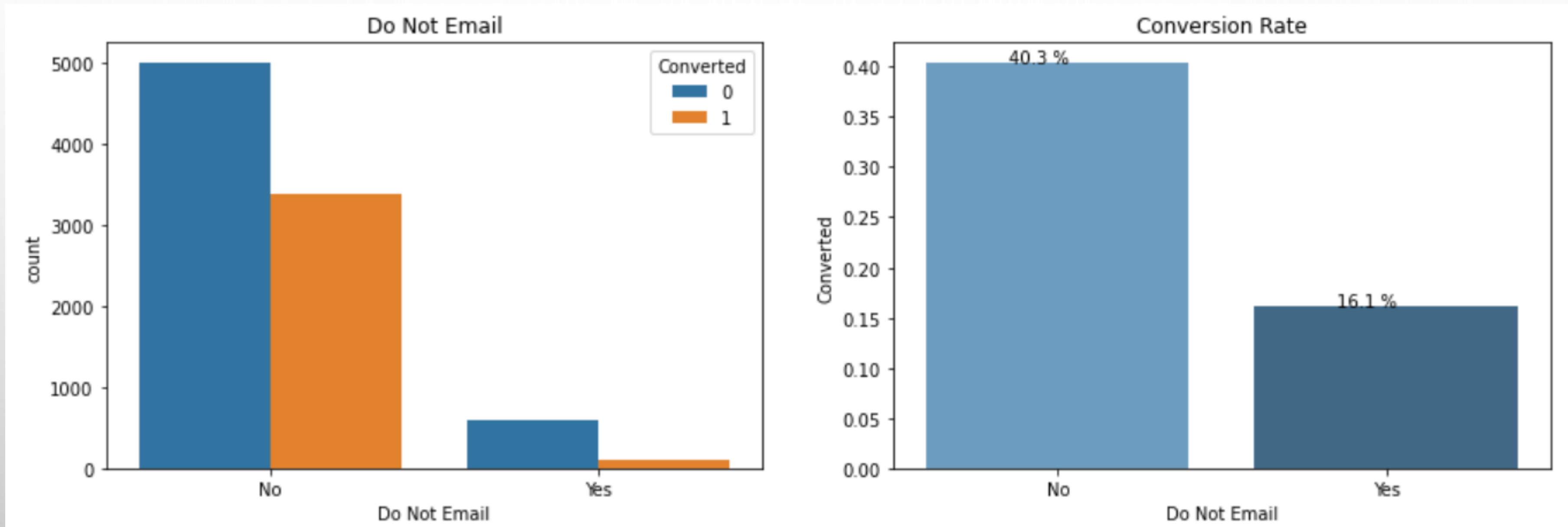
# ANALYSIS - LEAD SOURCE



## Observation:

Google generated the most leads, accounting for 40% of all conversions, followed by Direct Traffic, Organic Search, and Olark Chat. From a total of 534 leads, a lead generated by a referral has a conversion rate of more than 90%. Welingak's website has a nearly perfect lead conversion rate. This option should be investigated further in order to improve lead conversion. To increase lead count, initiatives should be implemented so that existing members increase their referrals.

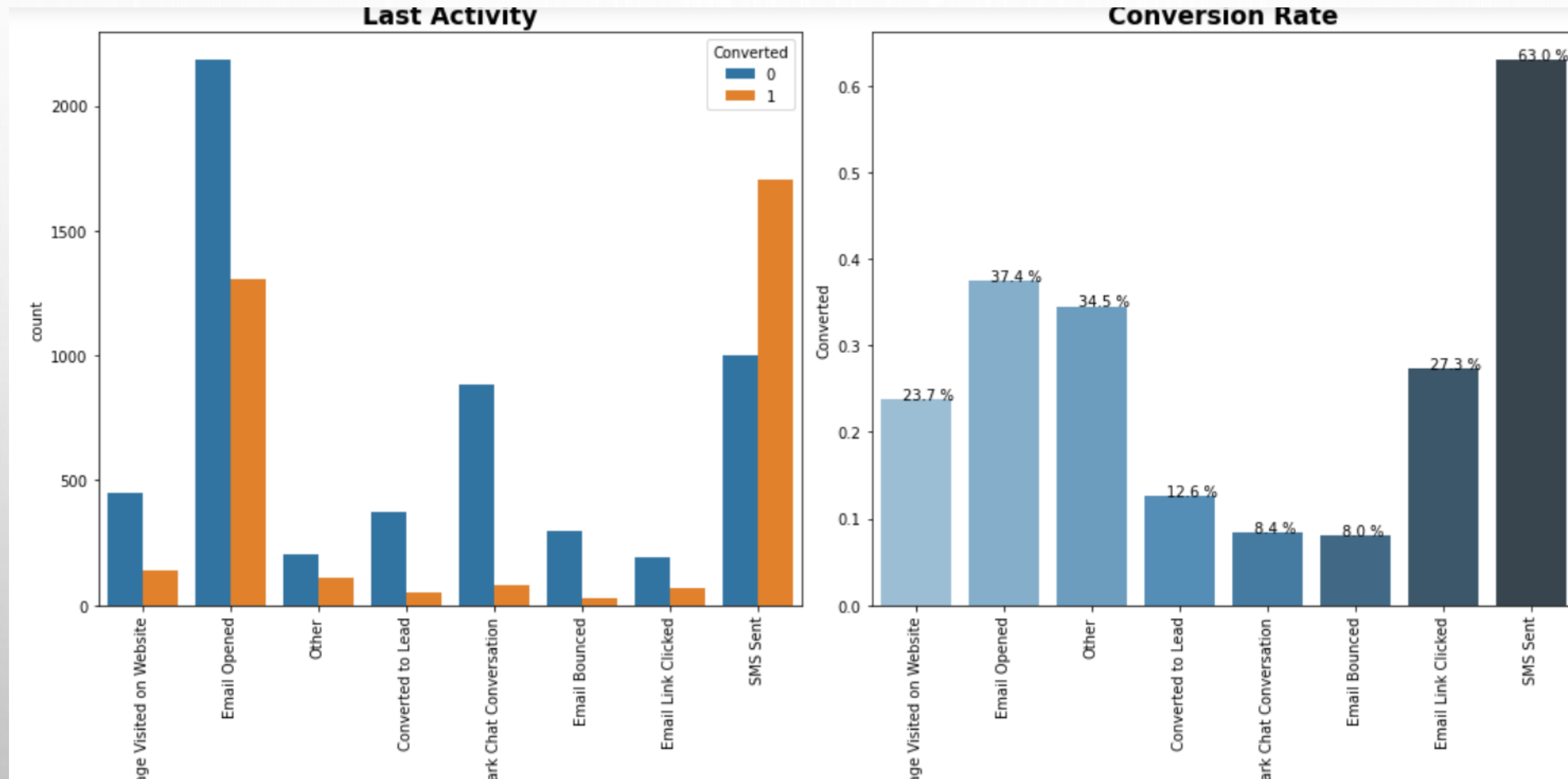
# ANALYSIS - DO NOT EMAIL



## Observation:

- Majority of the people want Email (~92%)
- People who have opted to receive Email has higher rate of conversion (40%)

# ANALYSIS - LAST ACTIVITY

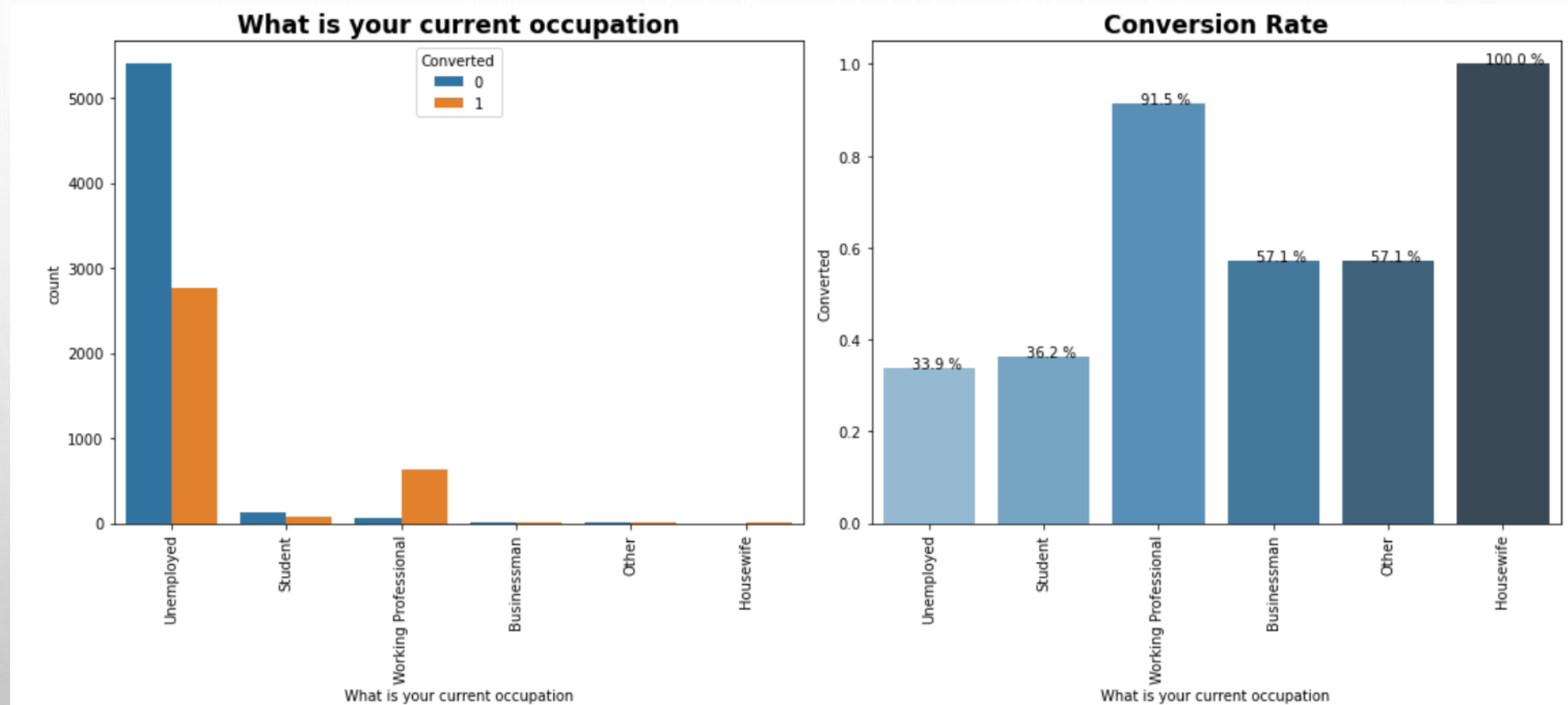


## Observation:

- As their most recent activity, the majority of the leads have opened their email. Lead conversion is very high (70 percent) after combining smaller Last Activity types as Other Activities. The conversion rate for leads with the most recent activity as SMS Sent is nearly 60%.



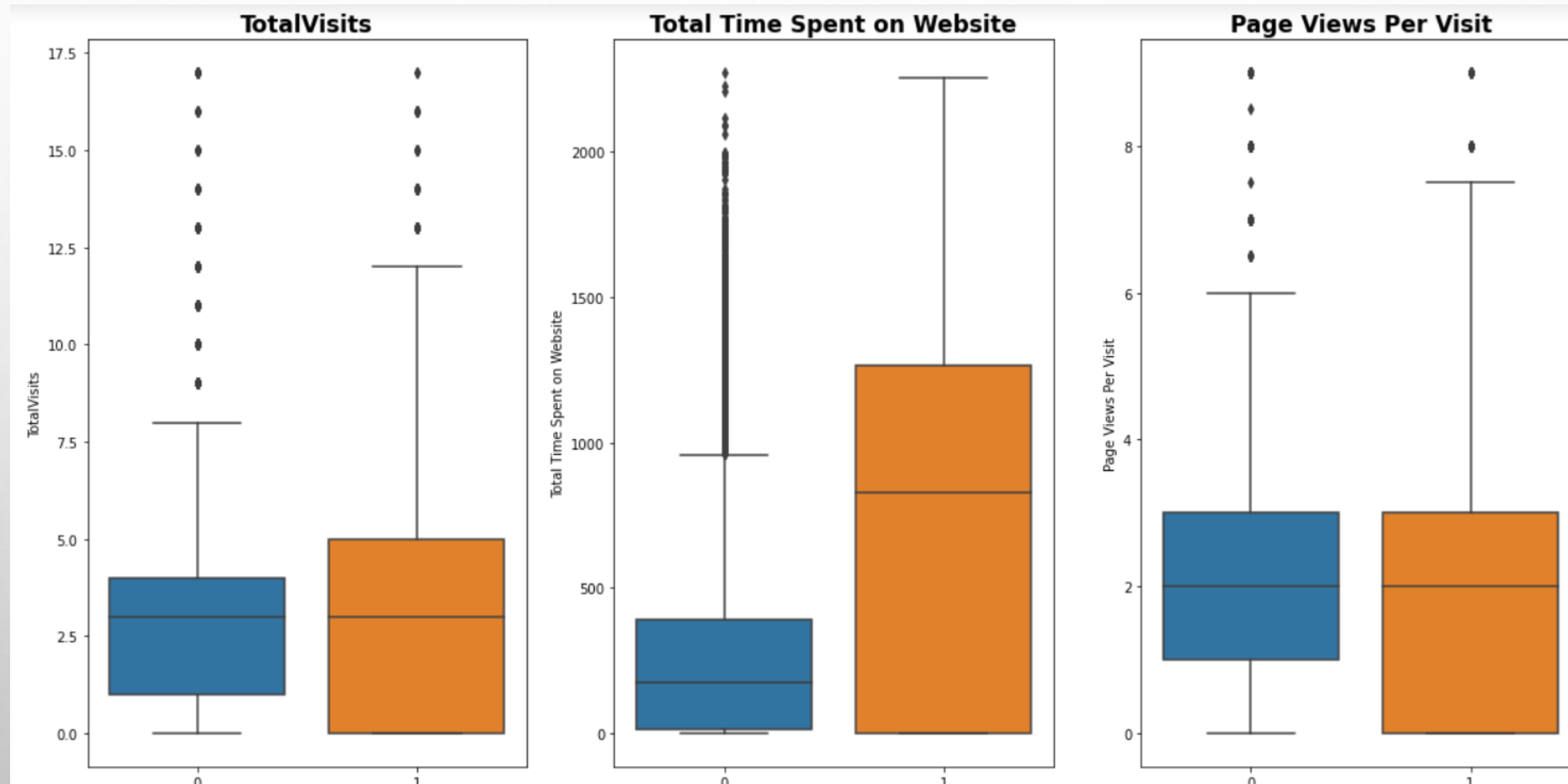
# ANALYSIS - CURRENT OCCUPATION



## Observation:

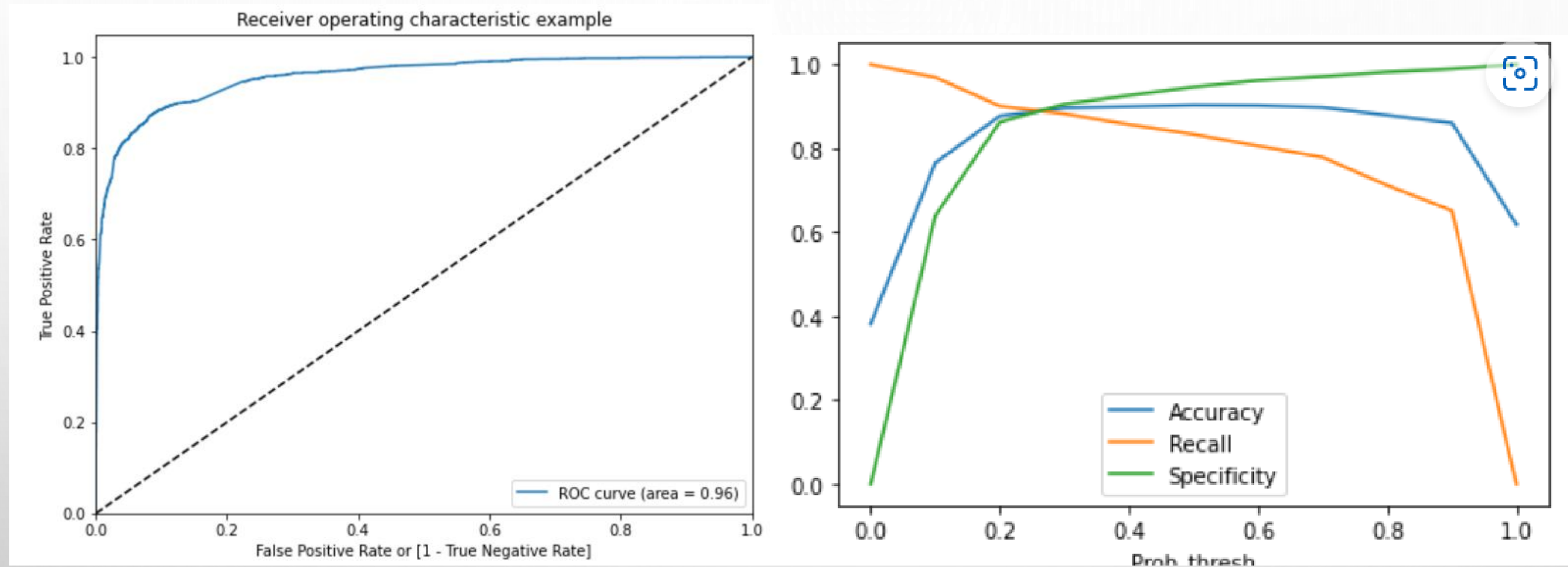
- Housewives are less in numbers, but have 100% conversion rate
- Working professionals, Businessmen and Other have high conversion rate
- Leads with Unemployed occupation is highest in number, but the conversion rate is low (~40%)

# ANALYSIS – TOTAL VISITS/TOTAL TIME SPENT ON WEBSITE & PAGE VIEWS PER VISIT



As the customers spend more time on website the chance of conversion is more and mean and meadian of Total Visits and Page views per visit is almost same.

# ROC CURVE & PROBABILITY THRESHOLD

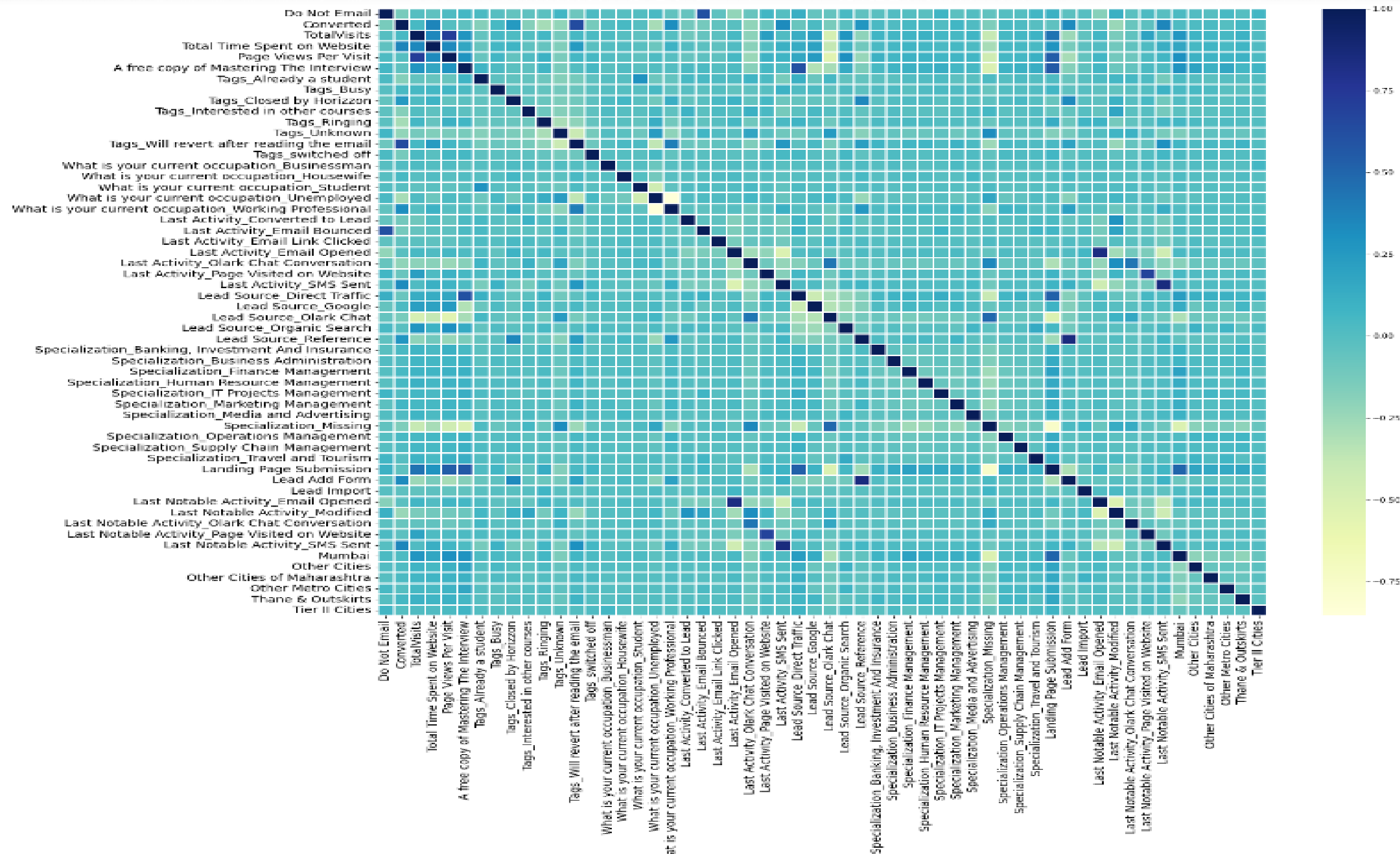


ROC CURVE

Probability Threshold

\* As you can see that around 0.25, you get the optimal values of the three metrics. So we have taken 0.25 threshold in consideration.

# ANALYSIS- MULTIVARIATE



# MODEL EVALUATION – TRAIN DATA

0.25 Probability Threshold

Accuracy  
89.08

Sensitivity  
89.17

Specificity  
89.02

AUC  
0.96

F1 Score  
86.17

Recall  
89.17

Precision  
83.37

Positive predictive value  
79.94

Negative predictive value  
86.92



# MODEL EVALUATION – TEST DATA

0.25

Probability Threshold

Accuracy

89.08

Sensitivity

89.88

Specificity

89.87

AUC

0.96

F1 Score

87.42

Recall

89.88

Precision

85.09

Positive predictive  
value

79.08

Negative predictive  
value

86.13

# FINAL PREDICTION

## Top 5 Records

	Converted	Convert_prob	Lead Score	predicted
6906	1.0	0.999069	99.91	1
1873	0.0	0.011742	1.17	0
771	0.0	0.000870	0.09	0
4495	0.0	0.002596	0.26	0
9061	1.0	0.715664	71.57	1

**Lead Score = 100 \* Conversion Probability**

- Higher the lead score, higher is the probability of a lead getting converted and vice versa,
- Since, we had used 0.25 as our final Probability threshold for deciding if a lead will convert or not, any lead with a lead score of 25 or above will have a value of '1' in the Final predicted column.

# FEATURE IMPORTANCE

1. If the value of the variable "Tags Closed by Horizzon" is True or 1, the log-odds increase by 6.11.
2. "Tags Will revert after reading the email": The log odds rise by 4.13 if the current status or tag is "Will revert after reading the email."
3. "Tags switched off": The log odds drop by 4.55 if the current status or tag is "turned off."
4. 'Tags Ringing': The log odds drop by 3.44 if the current status or tag is 'Ringing'.

# RECOMMENDATION

- 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 conversion rate, can be improved by focusing more on improving lead conversion of API and Landing Page Submission origin and generate more leads from Lead Add Form.
- Though Google is the highest source to get leads, the lead conversion through Google is low comparatively.
- Focus on Working Professional which has high conversion.
- Website should be made more engaging to make leads spend more time.
- Improve the Olark Chat service since this is affecting the conversion negatively.