

# LEAD SCORE CASE STUDY

**Presented by :-**

**Vaishali Agrawal**

**Sruthi Anith**

**Rohit Mishra**



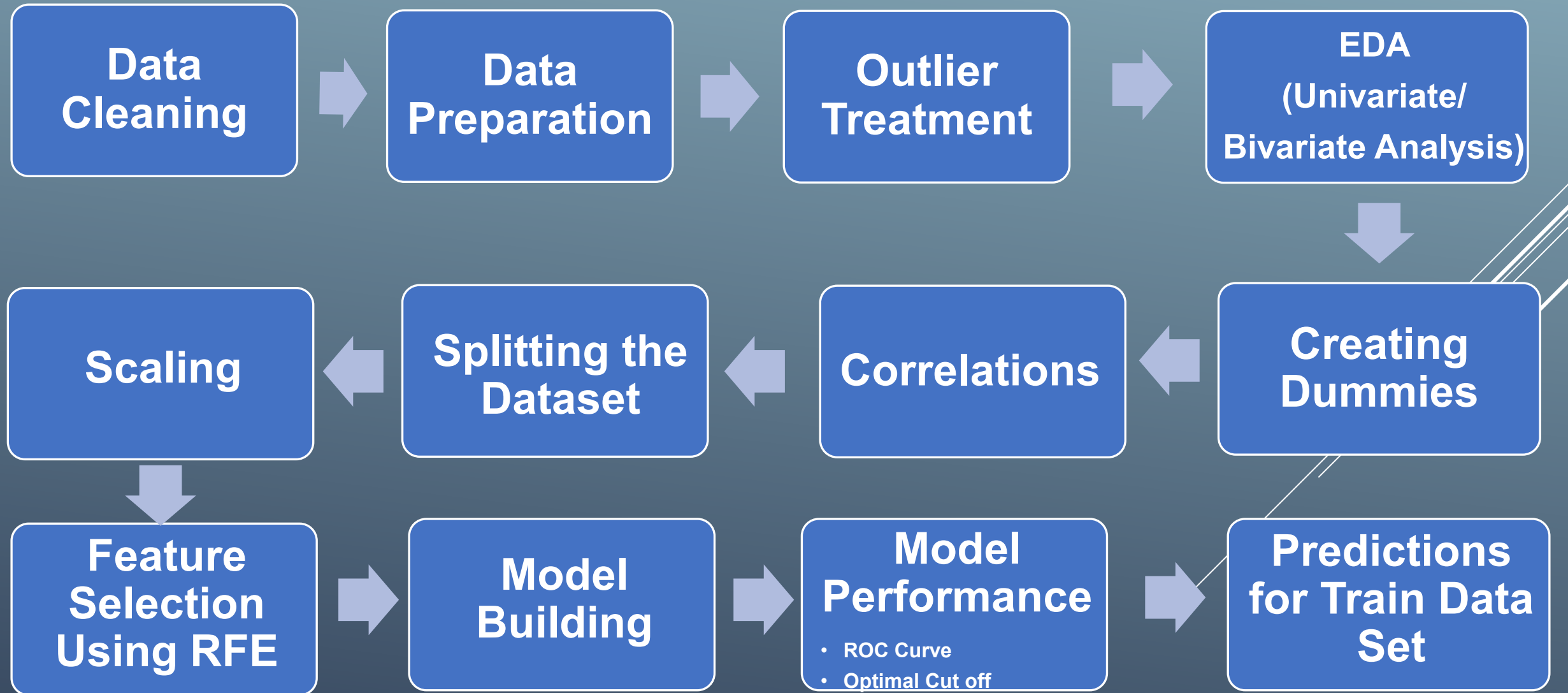
# **PROBLEM STATEMENT**

- An education company sells online courses to industry professional. The company markets its courses on several websites and search engines like Google.
- Once the potential users visits the website, they may browse the courses or fill up a form for the course.
- When these people fill up a form, they are classified to be a lead.
- The company as such receives a huge loads of leads
- Once these leads are acquired, 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 the company is around 30%.

# **BUSINESS OBJECTIVE**

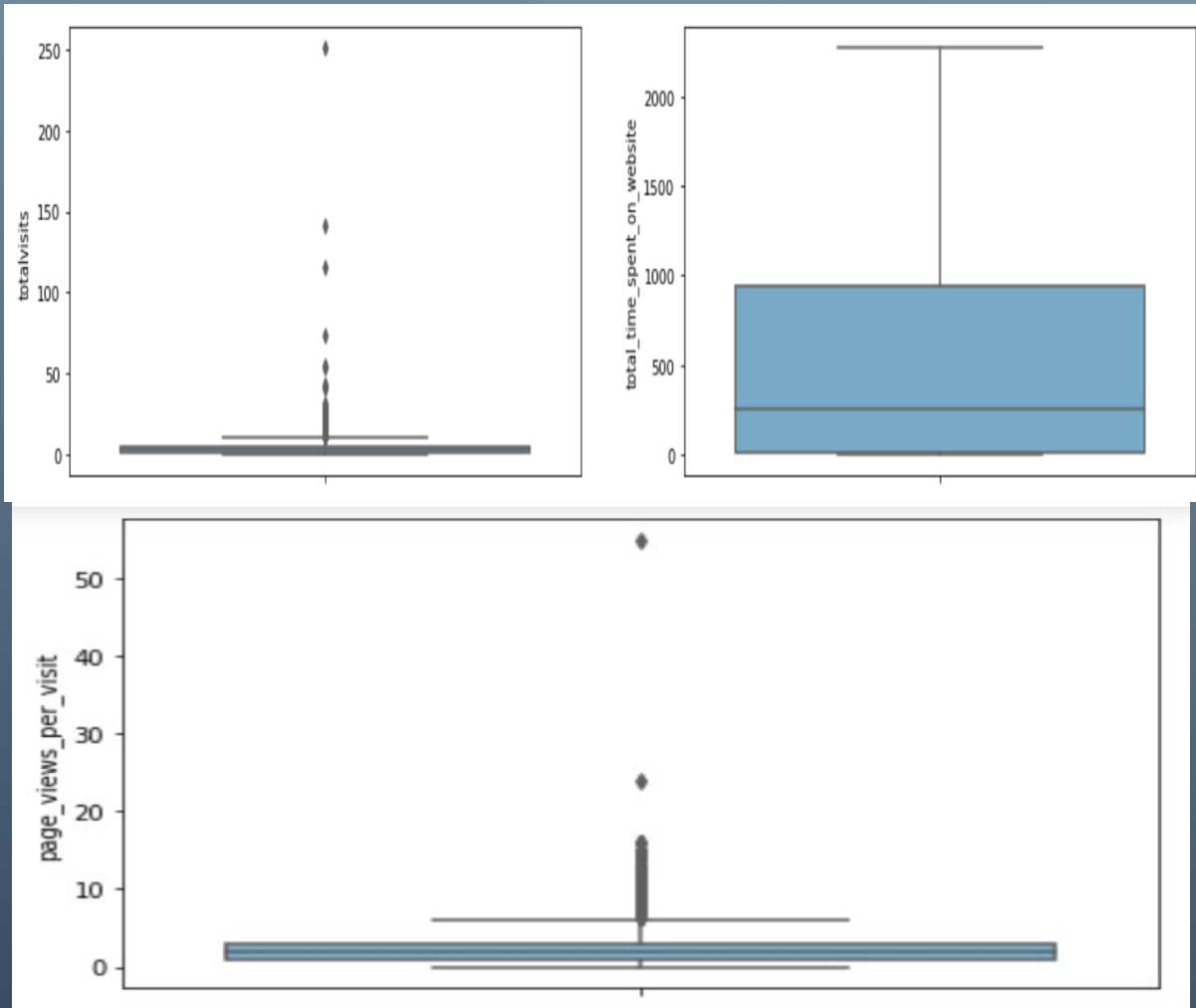
- The company needs help in building a model to identify the prospective leads thereby able to increase their conversion rate.
- The CEO has given a target lead conversion rate of 80% to achieve.
- In addition the company is looking forward to have a model which will help them in solving any future difficulties like deciding and implementing the possible approaches when the target is achieved.

# ANALYSIS APPROACH

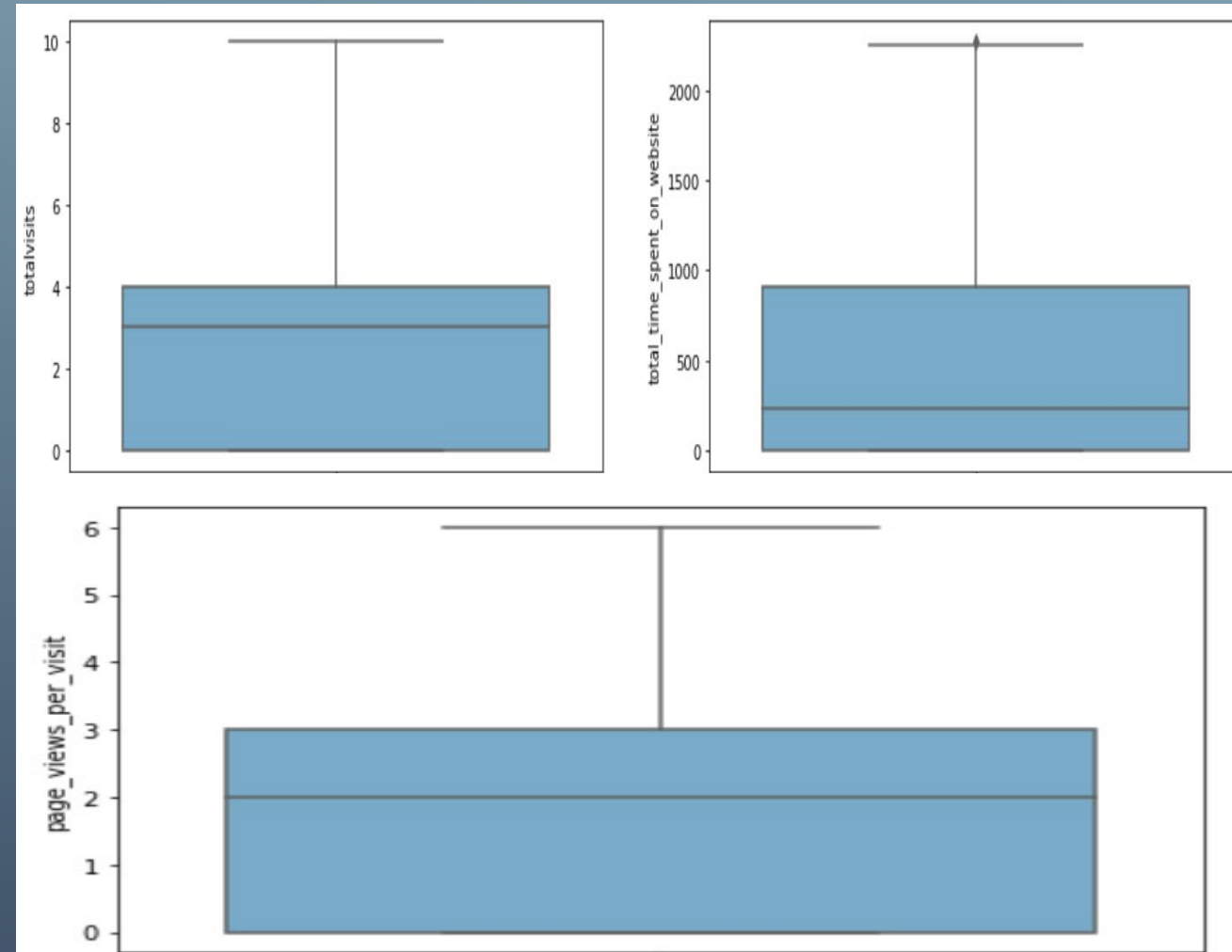


# OUTLIER TREATMENT

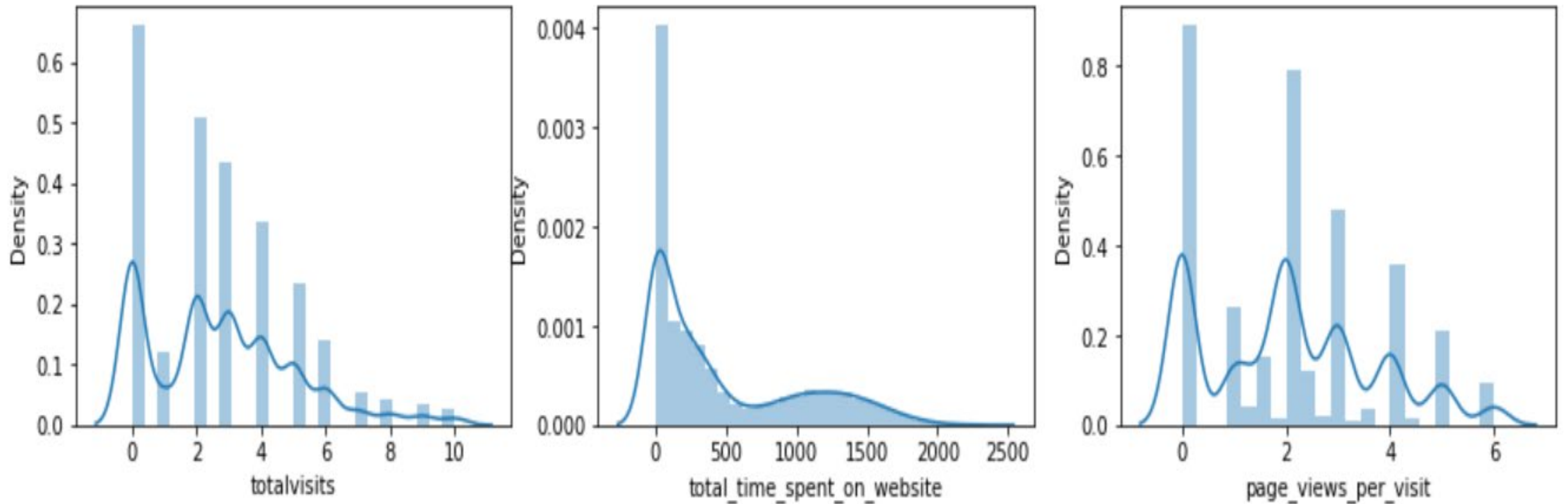
Before outlier



After Outlier

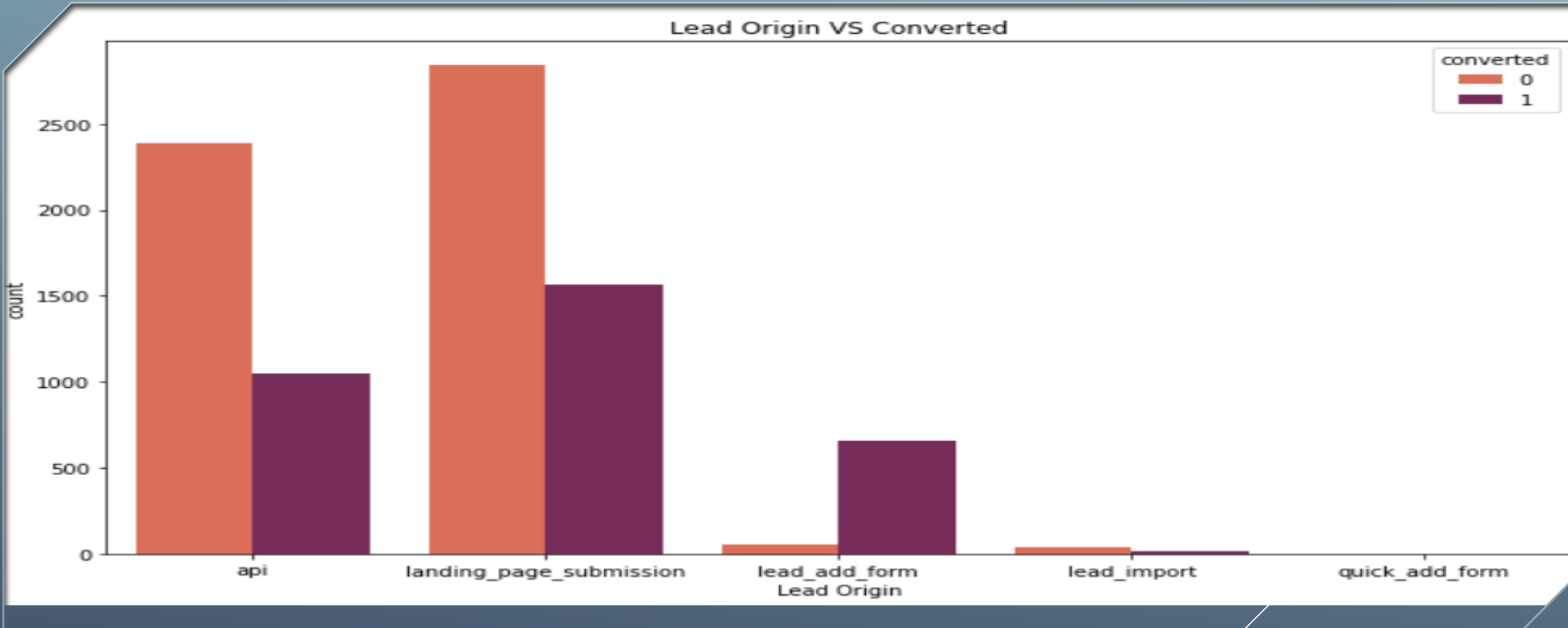


# EDA

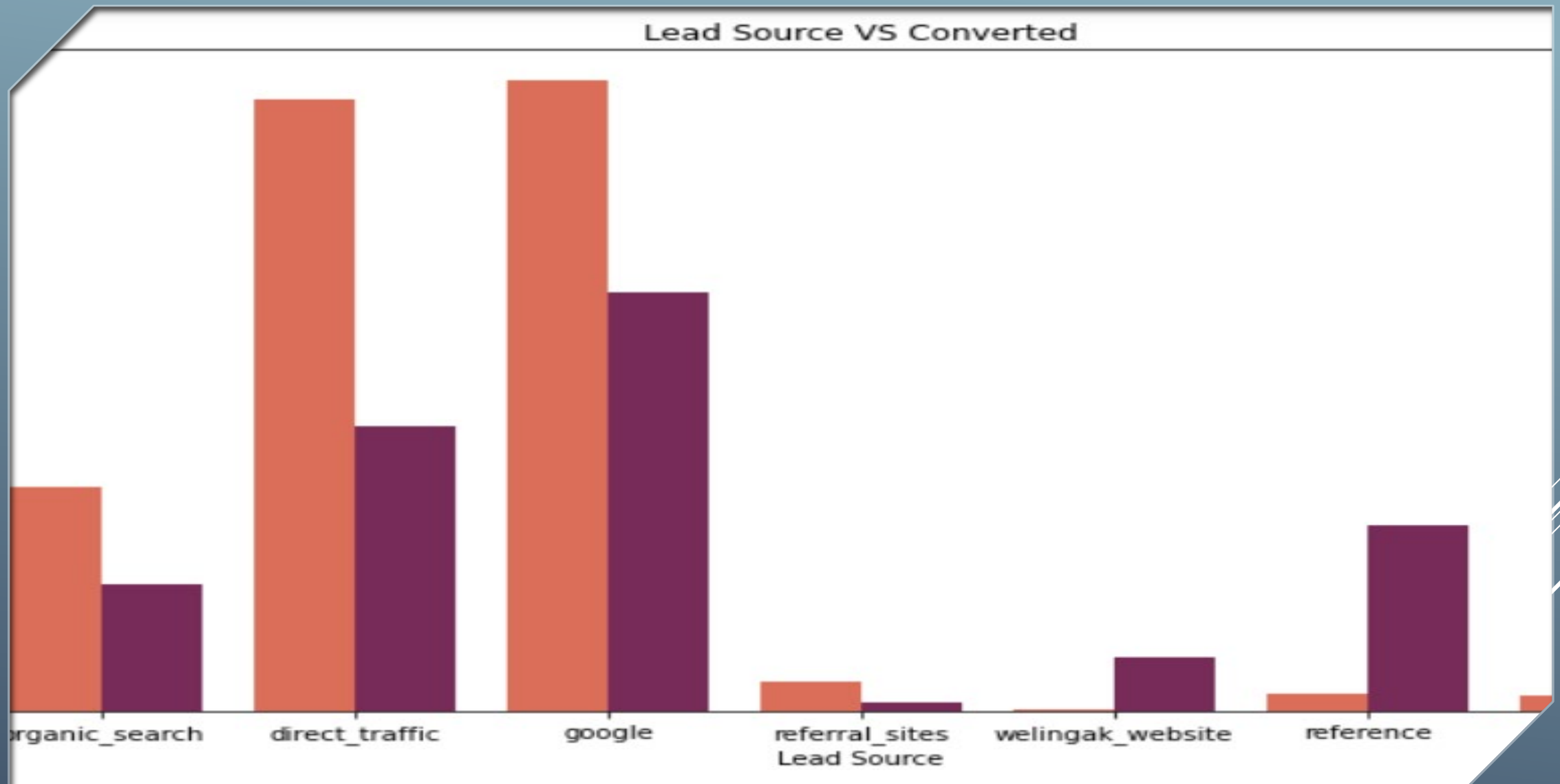


The graphical representation depicts the conversion rate of about 39%

# CATEGORICAL VARIABLES

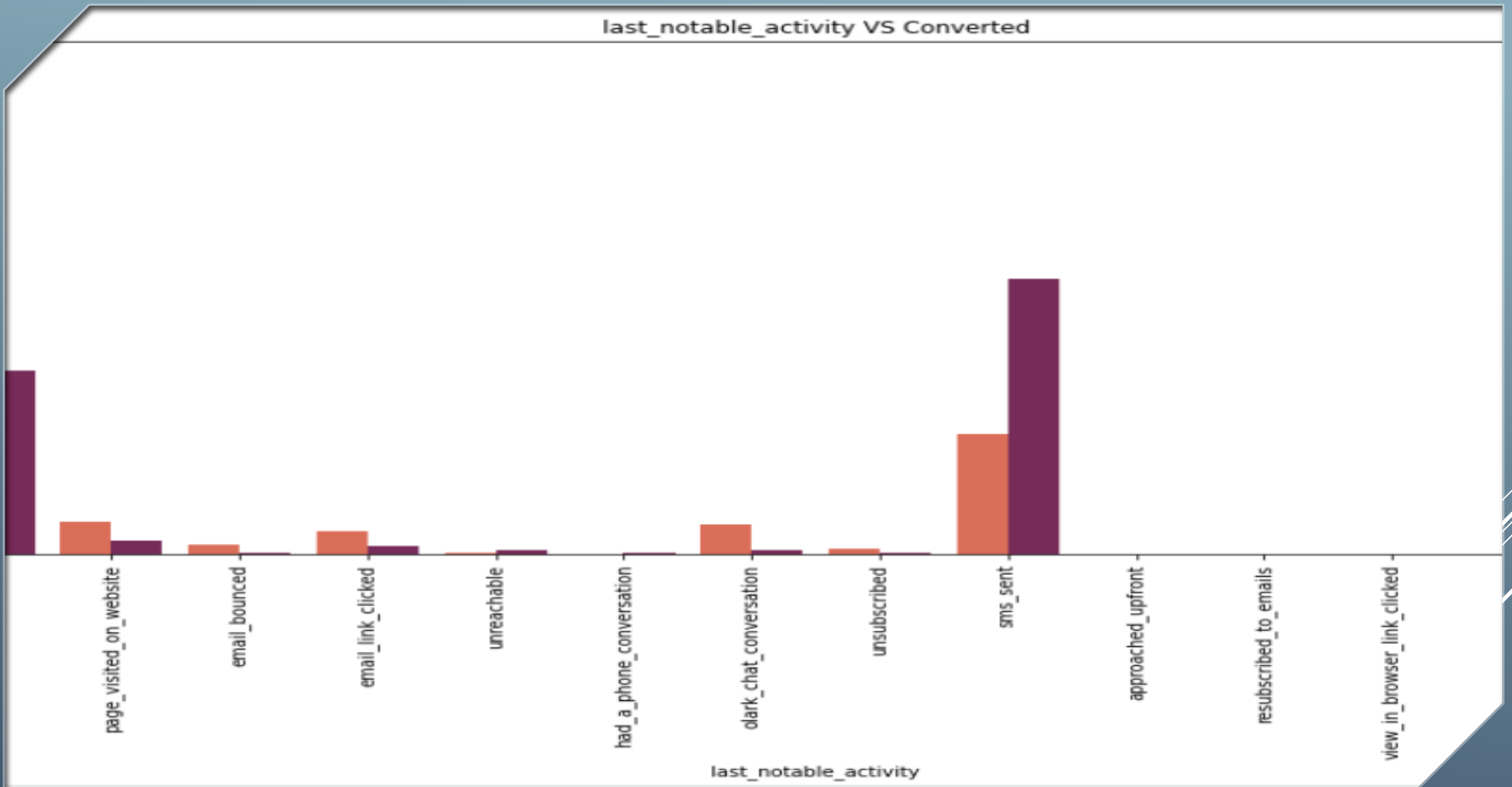


THE ABOVE GRAPH SHOWS THAT THE MAXIMUM CONVERSION TAKES PLACE FROM LANDING PAGE SUBMISSION.



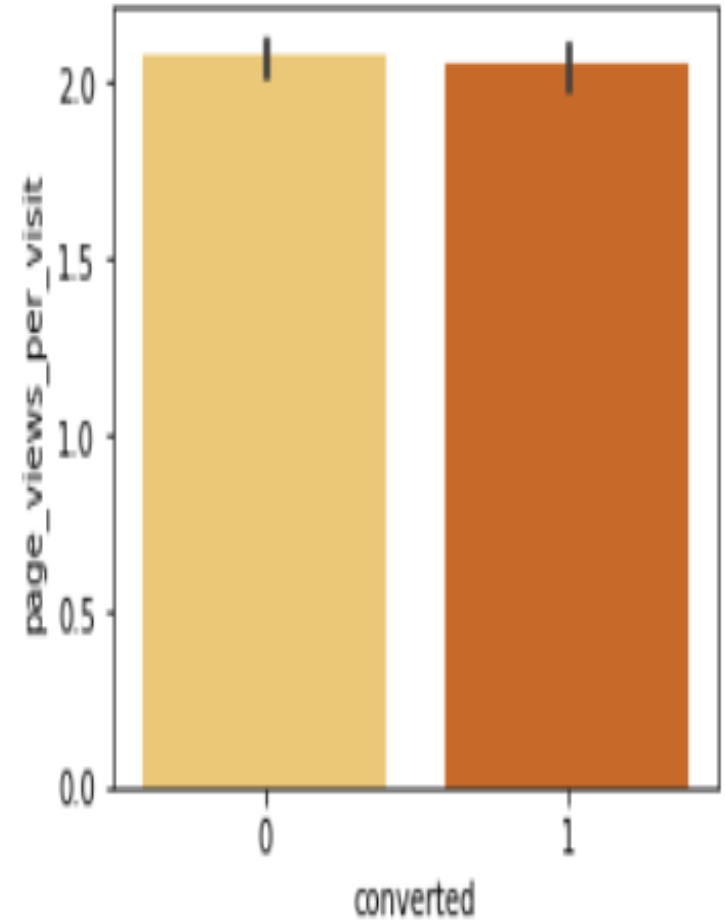
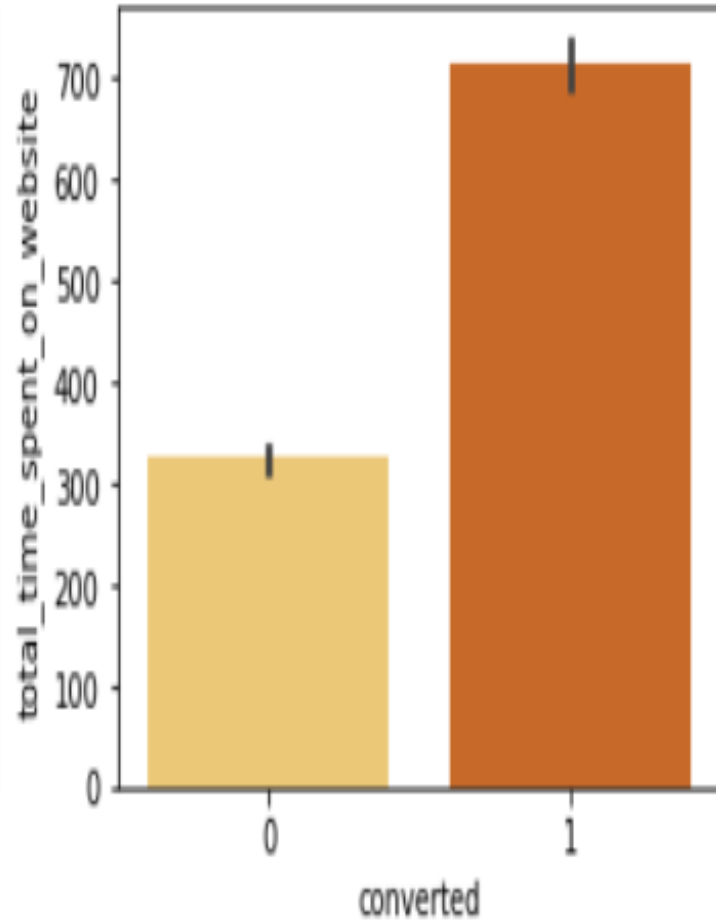
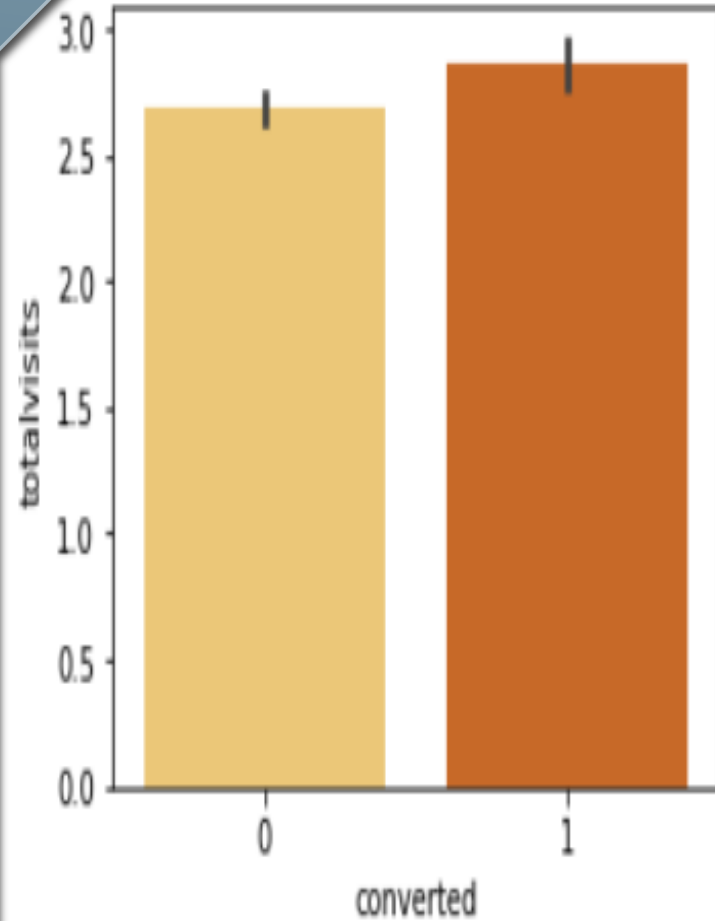
FROM THE GRAPH IT CAN BE INFERRED THAT MAXIMUM CONVERSION COMES FROM GOOGLE LEAD SOURCE

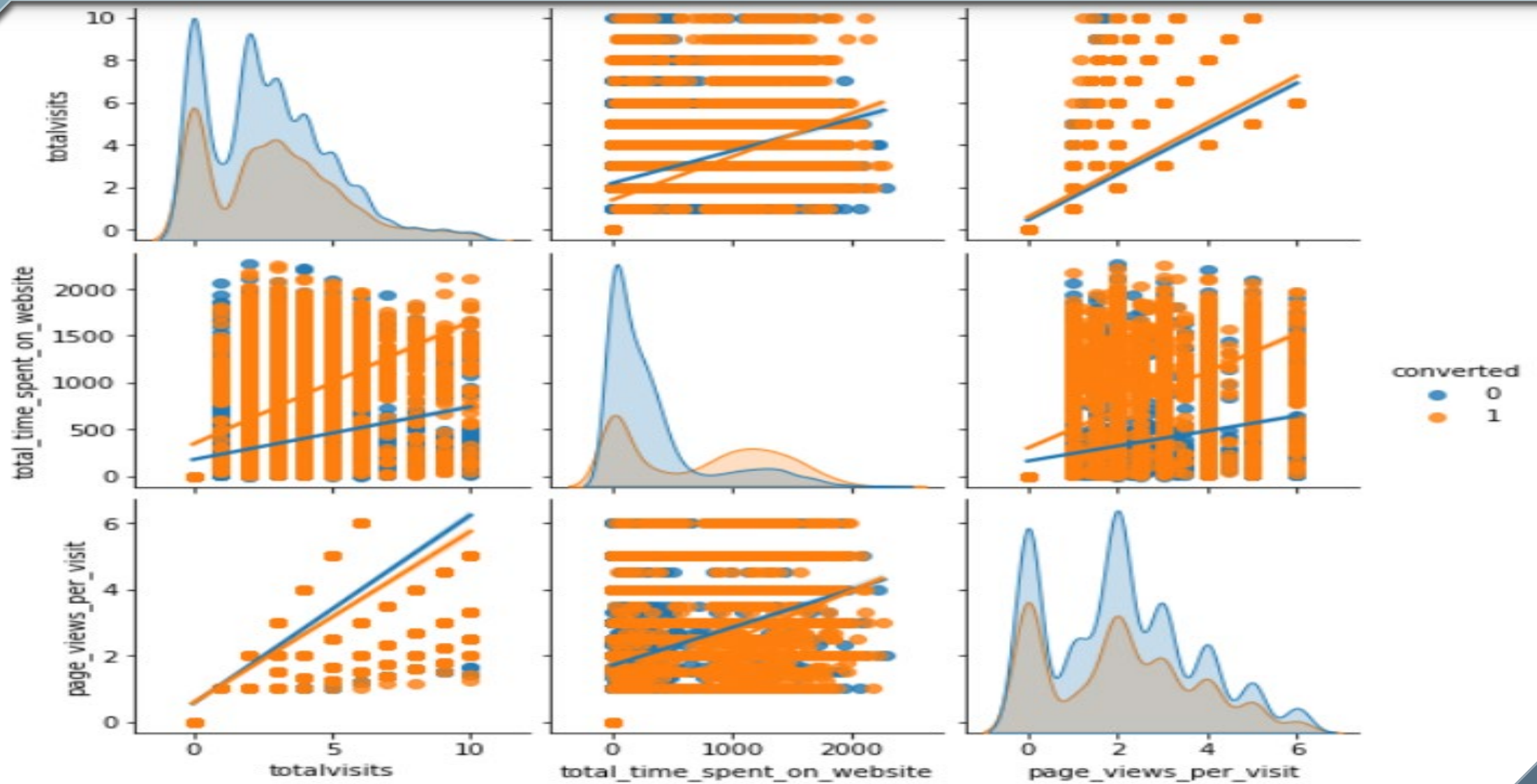




THE ABOVE GRAPH DEPICTS THAT SMS ACTIVITIES HAS GREATER CONVERSION FOLLOWED BY EMAIL OPENED

# NUMERICAL VARIABLES





TIME SPENT ON THE WEBSITE HAS THE HIGHEST IMPACT ON THE CONVERSION RATE

# MODEL BUILDING

## VIF Model 1

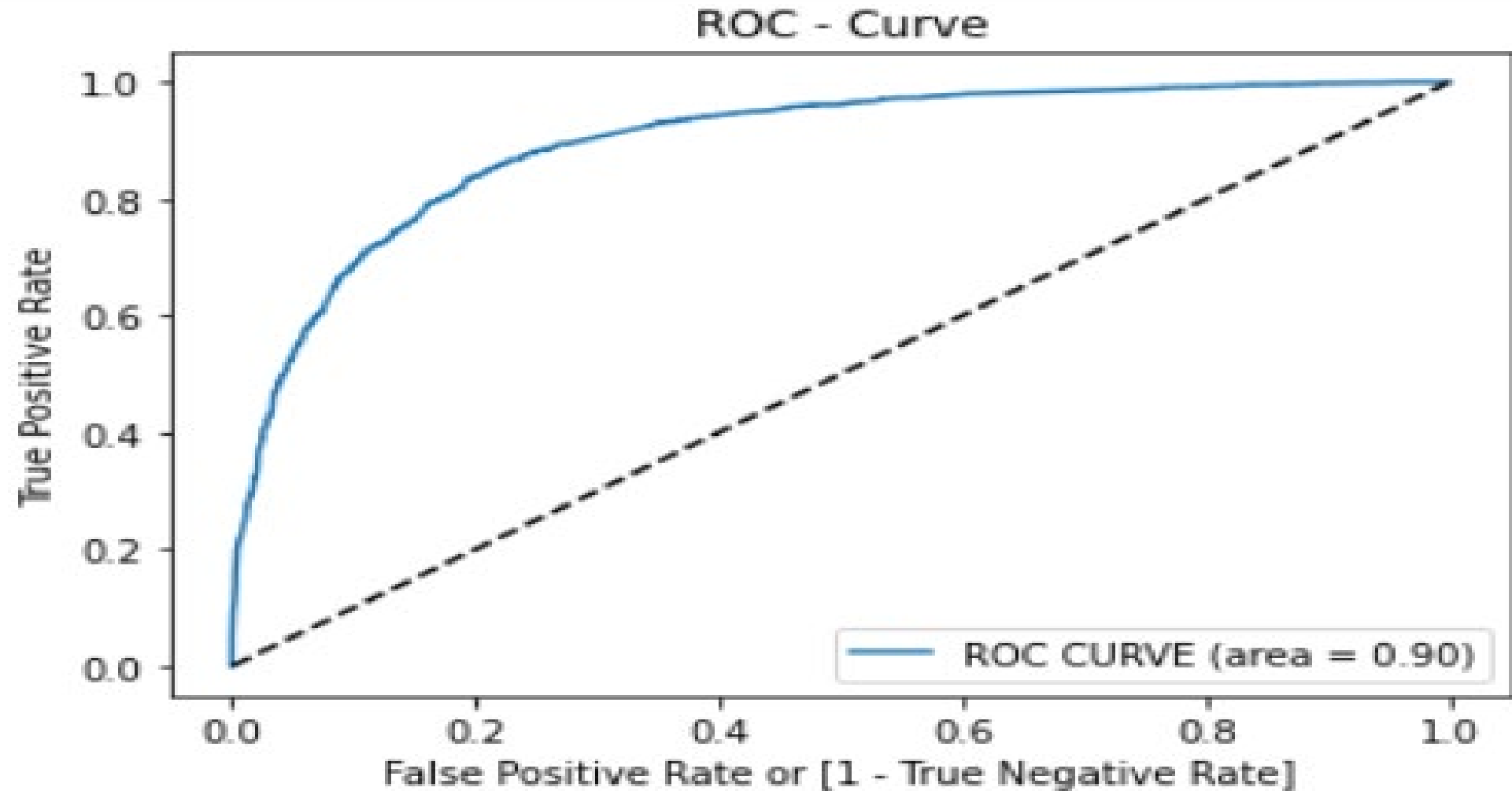
	features	VIF
2	what_is_your_current_occupation_unemployed	8.44
2	page_views_per_visit	8.01
0	totalvisits	6.35
3	lead_origin_landing_page_submission	5.80
11	what_is_your_current_occupation_not_provided	4.87
9	specialization_not_provided	4.09
8	country_not_provided	3.71
1	total_time_spent_on_website	2.31
4	lead_origin_lead_add_form	1.95
13	what_is_your_current_occupation_working_profes...	1.94
17	last_notable_activity_sms_sent	1.42
6	lead_source_welingak_website	1.28
7	do_not_email_yes	1.18
19	last_notable_activity_unsubscribed	1.08
15	last_notable_activity_olark_chat_conversation	1.06
5	lead_origin_lead_import	1.04
18	last_notable_activity_unreachable	1.01
10	what_is_your_current_occupation_housewife	1.01
14	last_notable_activity_had_a_phone_conversation	1.00
16	last_notable_activity_resubscribed_to_emails	1.00

## VIF Model 6

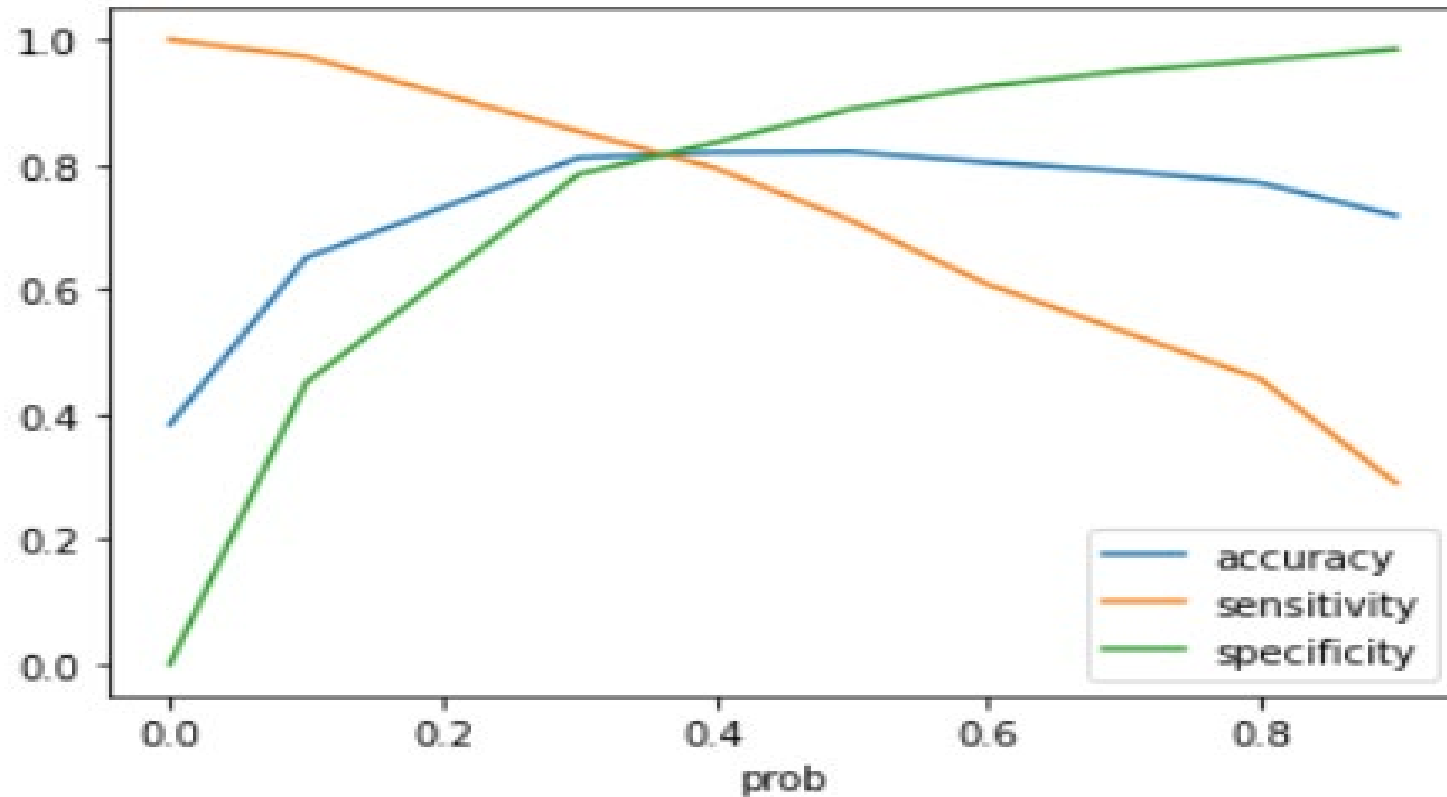
	features	VIF
0	totalvisits	3.59
2	lead_origin_landing_page_submission	3.34
6	country_not_provided	2.83
7	specialization_not_provided	2.81
1	total_time_spent_on_website	2.18
3	lead_origin_lead_add_form	1.87
8	what_is_your_current_occupation_not_provided	1.64
12	last_notable_activity_sms_sent	1.39
4	lead_source_welingak_website	1.27
9	what_is_your_current_occupation_working_profes...	1.21
5	do_not_email_yes	1.18
14	last_notable_activity_unsubscribed	1.08
11	last_notable_activity_olark_chat_conversation	1.06
13	last_notable_activity_unreachable	1.01
10	last_notable_activity_had_a_phone_conversation	1.00

Comparisons of before and after removal of multi-collinear features.

# ROC CURVE



# OPTIMAL CUT OFF



The optimal cut off point probability is around 0.37

# PRECISION FOR OPTIMAL CUT OFF

	Train Dataset Metrics	Test Dataset Metrics
Accuracy	0.8163265306122449	0.8053405572755418
Sensitivity	0.8054474708171206	0.778118609406953
Specificity	0.8231017770597738	0.821917808219178
Precision	0.7392857142857143	0.7268385864374403
Recall	0.8054474708171206	0.778118609406953

# CONCLUSION

- ▶ Features such as total time spent on website, lead source website have higher impact on the conversion rate.
- ▶ For higher probability of conversion, working professionals has to be focused on as they have better financial status to pay for services when compared to other categories.
- ▶ Conversion rates are more from google.
- ▶ Apart from phone conversation the 'SMS sent' or 'Email Opened' have a higher conversion rate.
- ▶ Focusing on tailored mail and calls may have a higher impact on conversion rate.