

Abstract geometric lines in the top-left corner of the page, consisting of several thin, black, overlapping lines that form a complex, non-representational shape.

LEAD SCORE CASE STUDY

Girish, Saurabh and Monika

EXPLORATORY DATA ANALYSIS

1. Data Cleaning
2. Univariate Analysis
3. Bivariate Analysis

FEATURE ENGINEERING

1. Converting categorical columns to dummies
2. Identifying and dropping highly correlated columns (>0.7)

MODEL BUILDING

1. Feature Selection through RFE
2. Rebuilding Model based on VIF and P value analysis
3. ROC and Determining Optimal Cut off Point

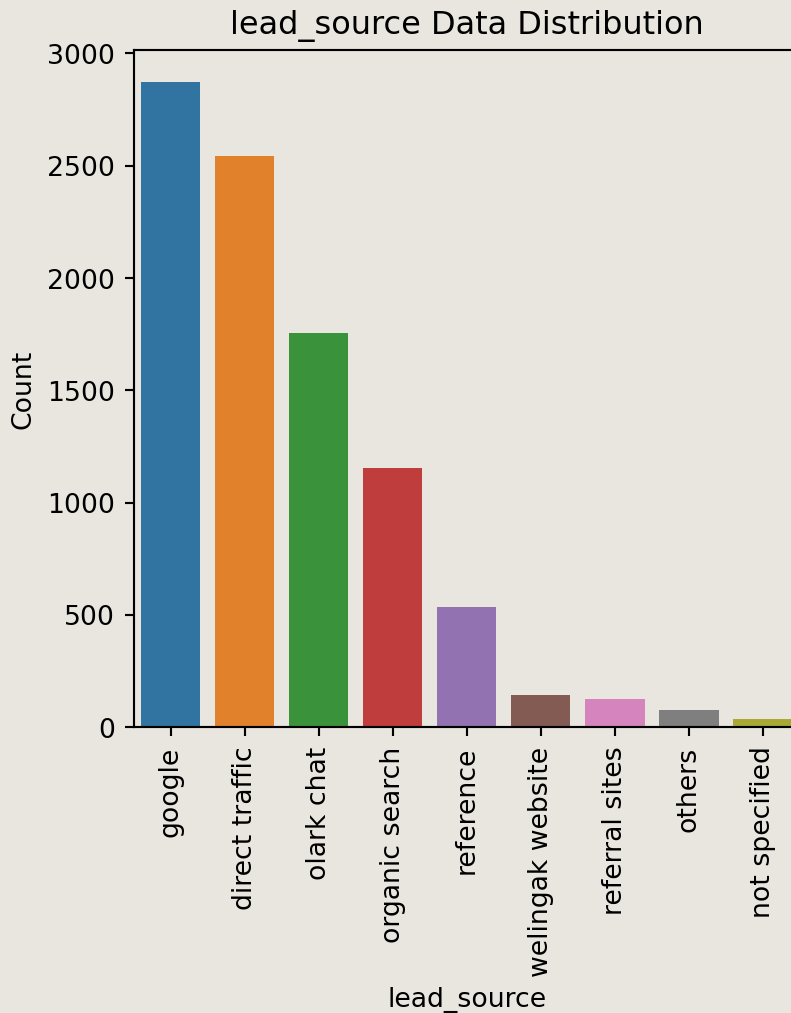
MODEL EVALUATION

1. Validating Cut off point Using Precision & Recall Curve
2. Predicting test datasets
3. Comparing Performance of Train v/s Test data set

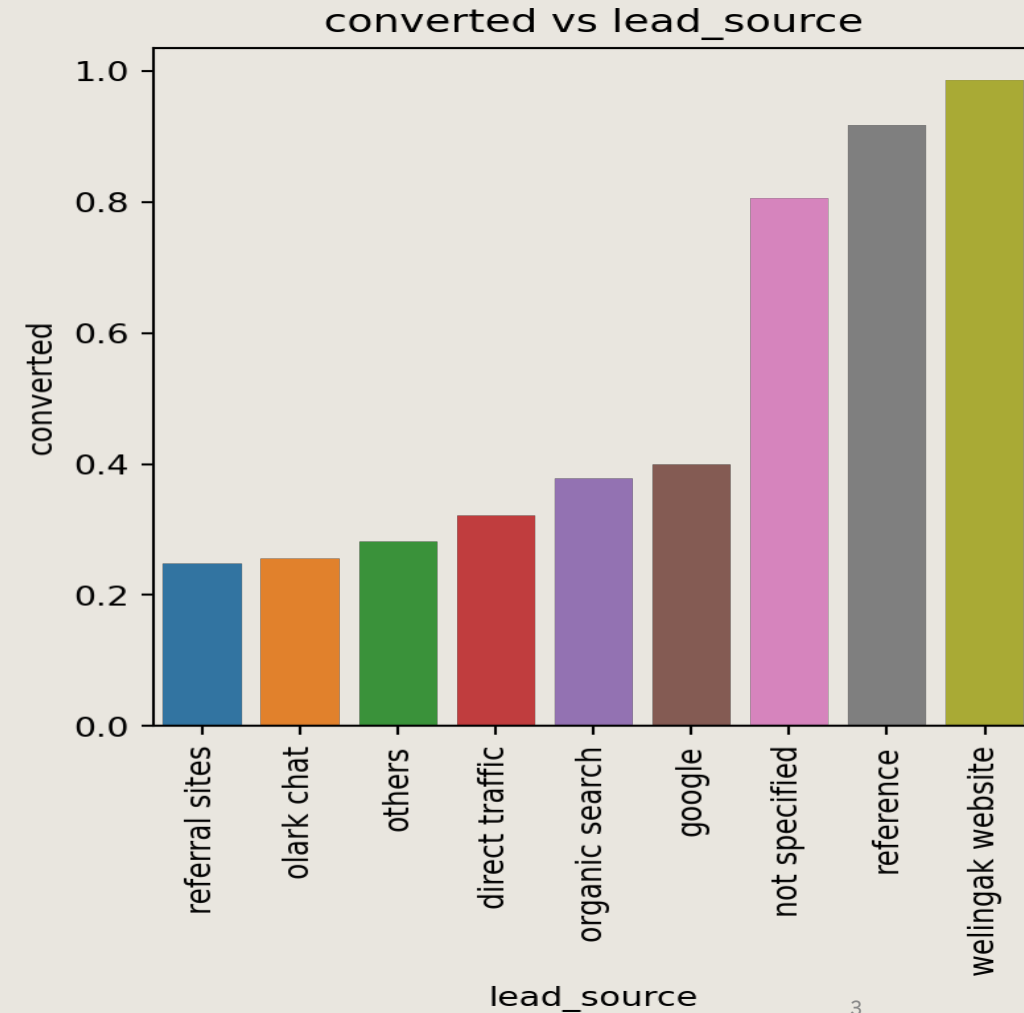
STEPS

INSIGHTS FROM EDA

Lead Source

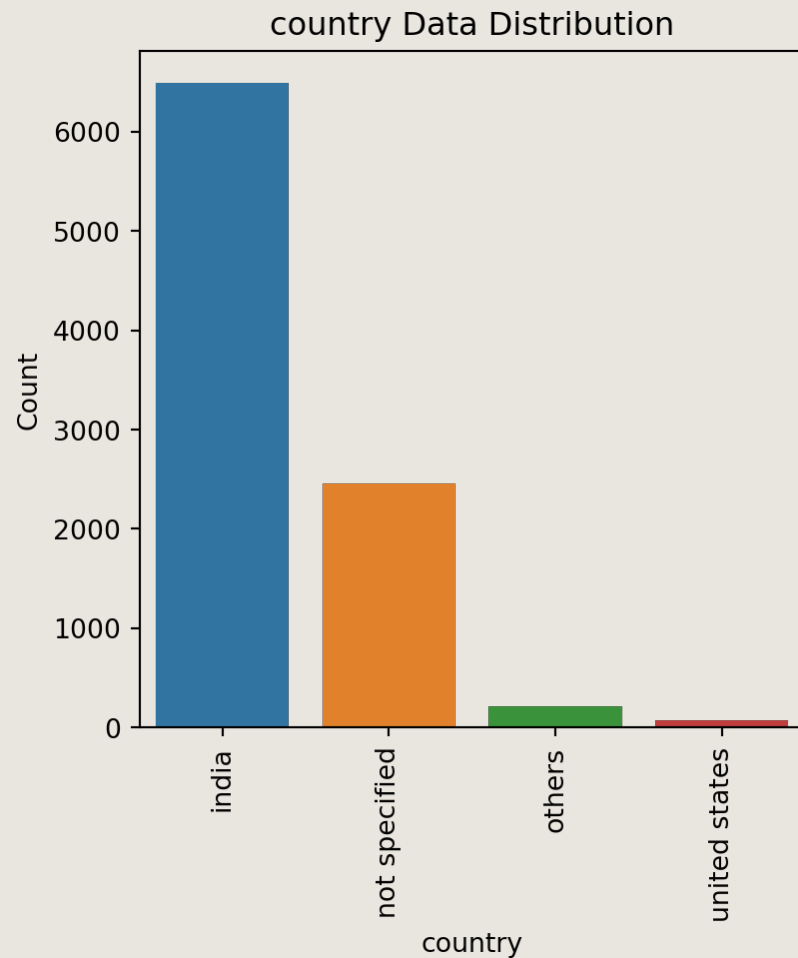


- Though Leads from is High, Conversion is low
- Conversely, Leads from Reference is low, but Conversion is High

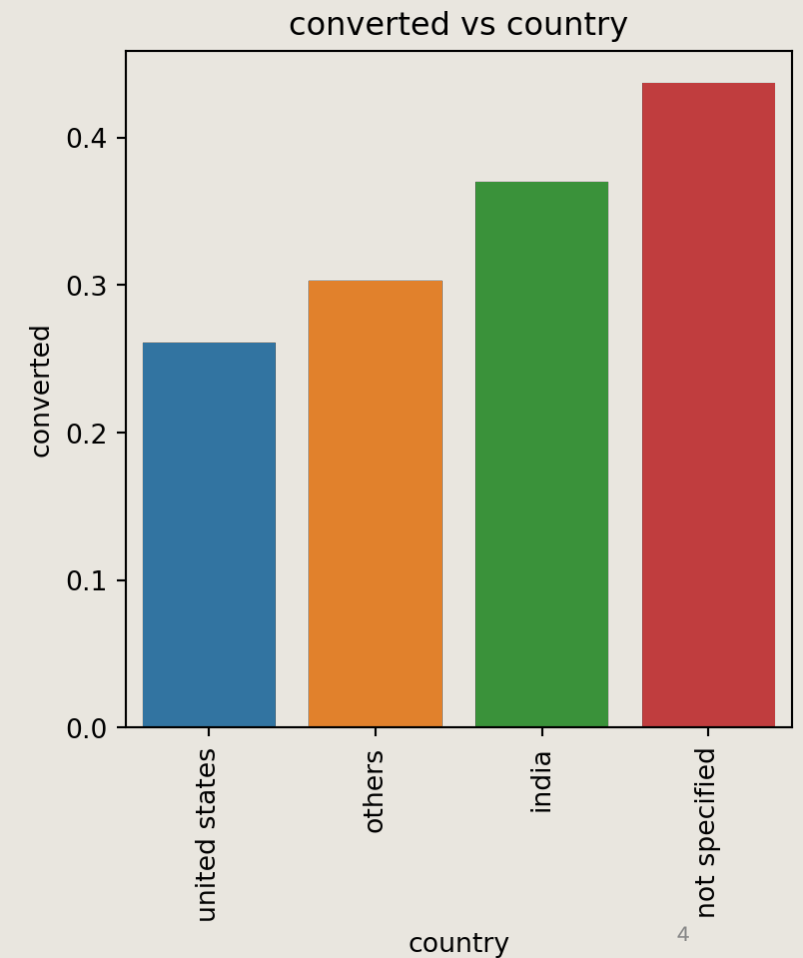


INSIGHTS FROM EDA

Country



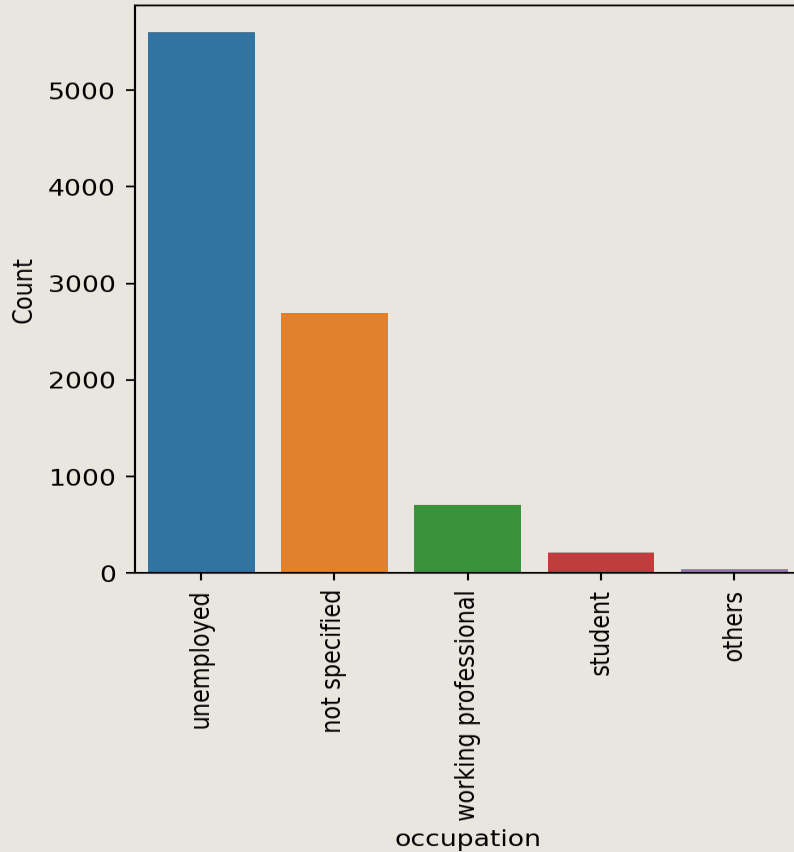
- Leads from India is High
- Conversion rate of india is also High
- When user doesn't specify country, Its Safe to assume that they belong to india as almost 75% leads are from india



INSIGHTS FROM EDA

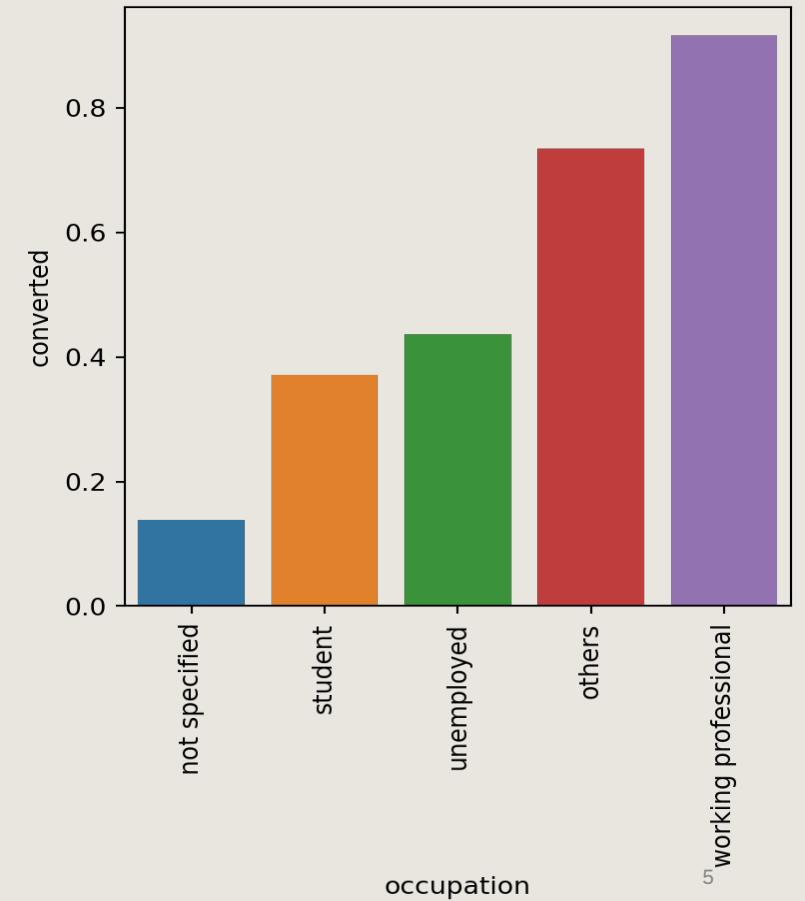
Occupation

occupation Data Distribution



- Working Professionals are more likely to join the course
- Leads from Unemployed Customers have moderate chance of conversion
- Students are less likely to join the course

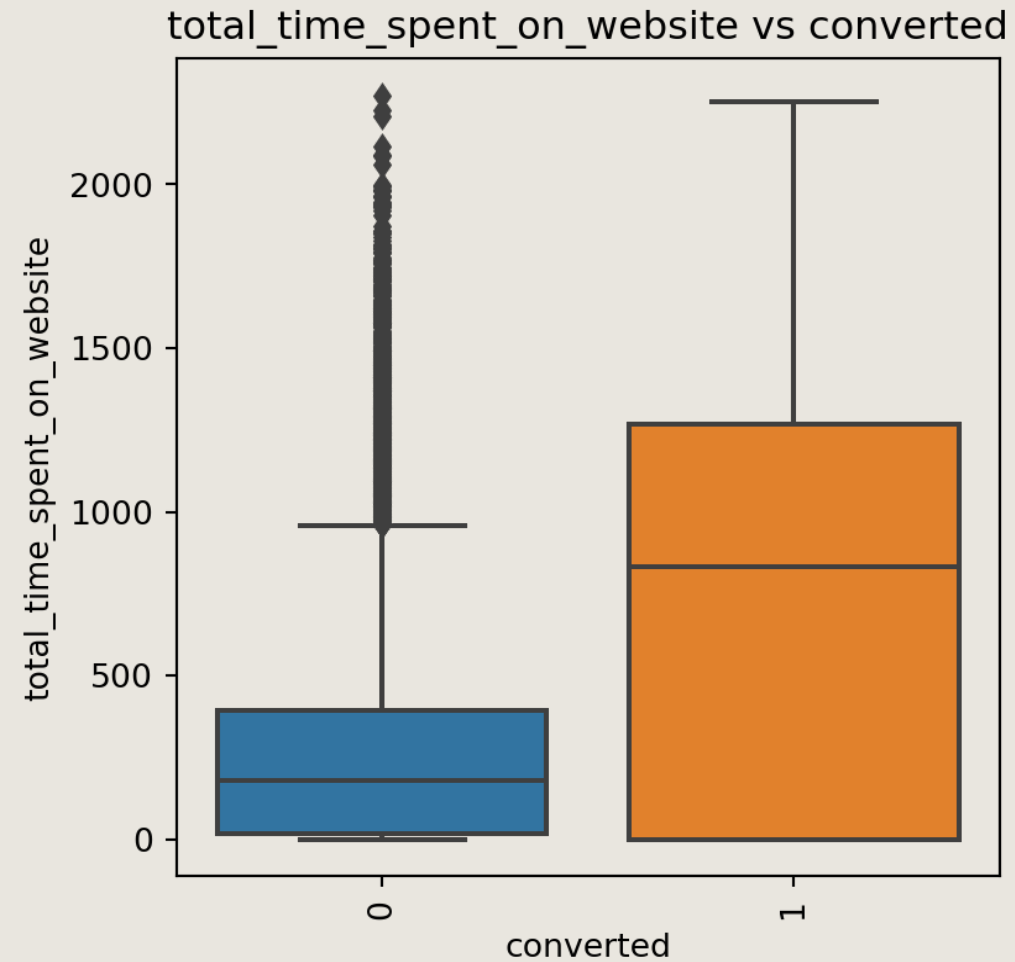
converted vs occupation



INSIGHTS FROM EDA

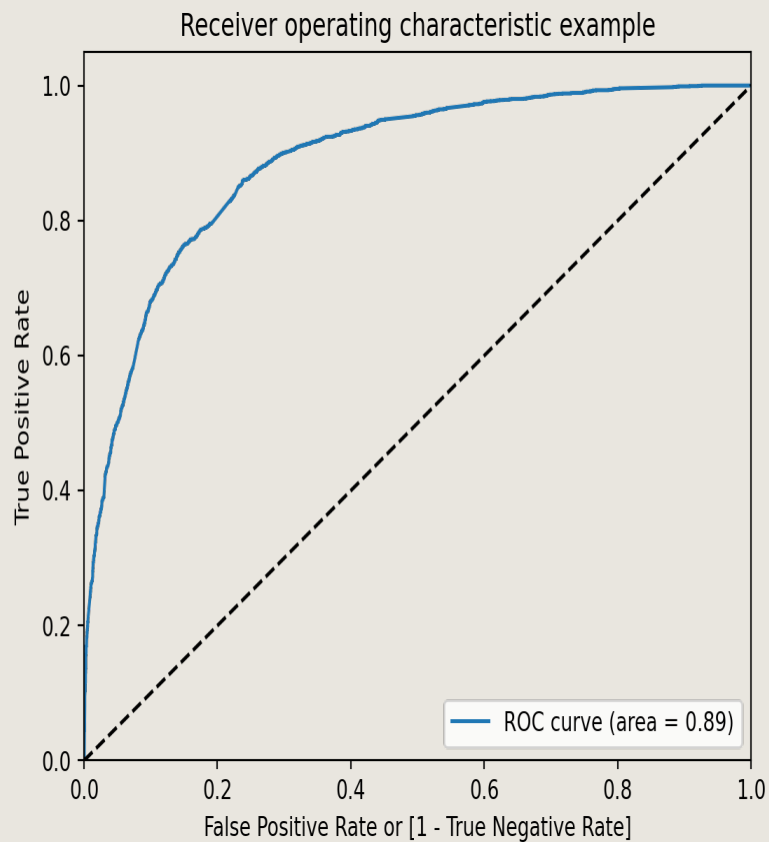
Total Time Spent On Website

- People Who spend more time on website are more likely to convert

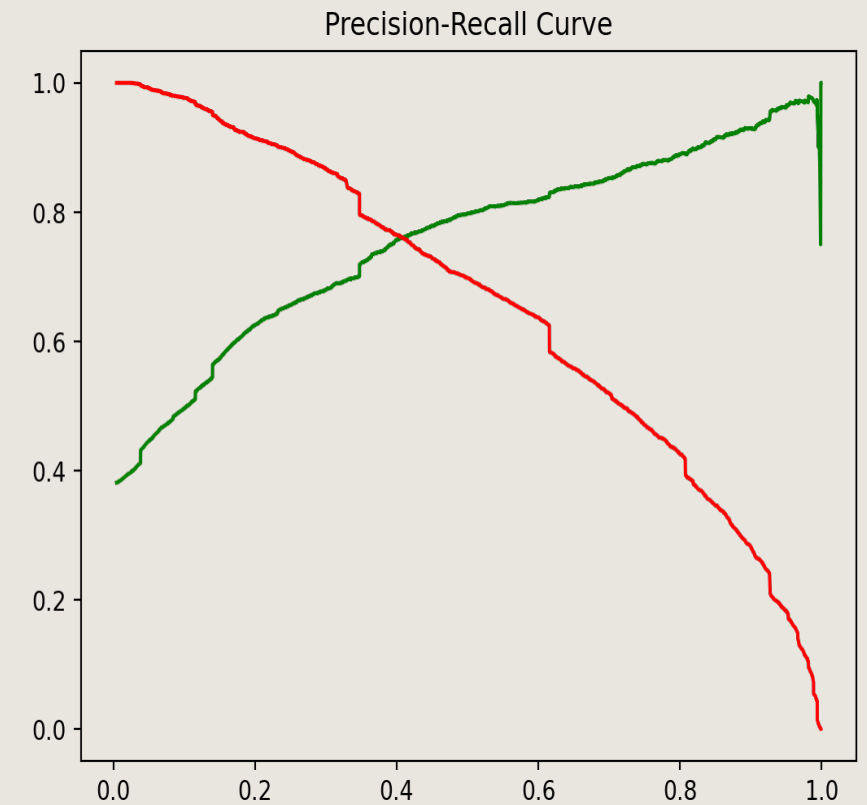


MODEL EVALUATION

ROC



- ROC is 0.89, which indicates that our model is statistically good
- Optimal Cut off is Around 0.37 – 0.4, we can Consider 0.37



MODEL EVALUATION

Accuracy, Specificity, Sensitivity/Recall, Precision and FPR

KEY METRICS		
	Train	Test
Accuracy	0.81	0.80
Specificity	0.78	0.77
Sensitivity/Recall	0.82	0.83
Precision	0.74	0.75
FPR	0.17	0.16

TOP 10 FEATURES

KEY METRICS	
	Co-efficient
lead_source_welingak website	3.983940
occupation_working professional	3.717404
occupation_others	2.224231
lead_source_reference	2.062279
last_activity_others	1.699773
lead_source_google	-1.144451
do_not_email_1	-1.239934
lead_source_referral sites	-1.244498
lead_source_organic search	-1.318488
lead_source_direct traffic	-1.560190



GROWTH STRATEGY

- Support team can contact if a lead belongs to any of these categories:
 1. Customer is Working Professional
 2. Lead was obtained from Welingak Website
 3. Lead was obtained through referral
- Support team can give lower priority if lead belongs to any of these categories:
 1. Lead was obtained from Google, direct visit or referral sites
 2. Customer did not opt for email communication

A series of white, thin, overlapping geometric lines on a black background, forming a complex, abstract pattern on the left side of the slide.

THANK YOU

Girish, Saurabh and Monika