LEADS SCORING
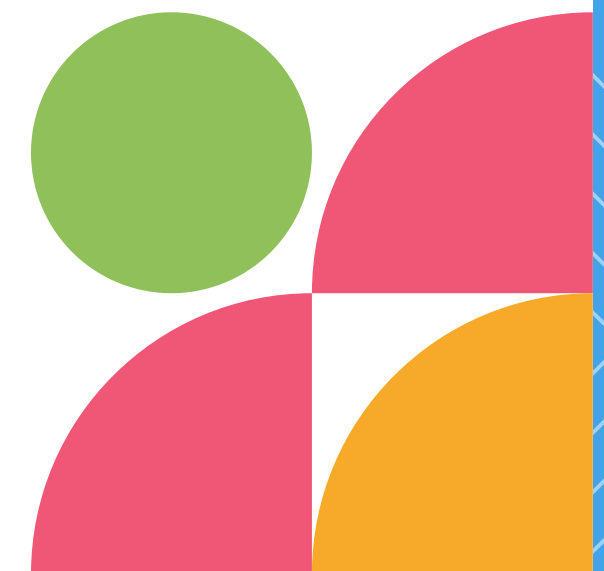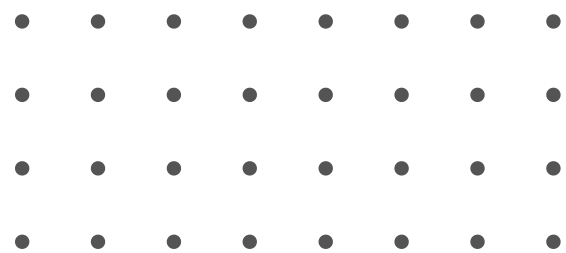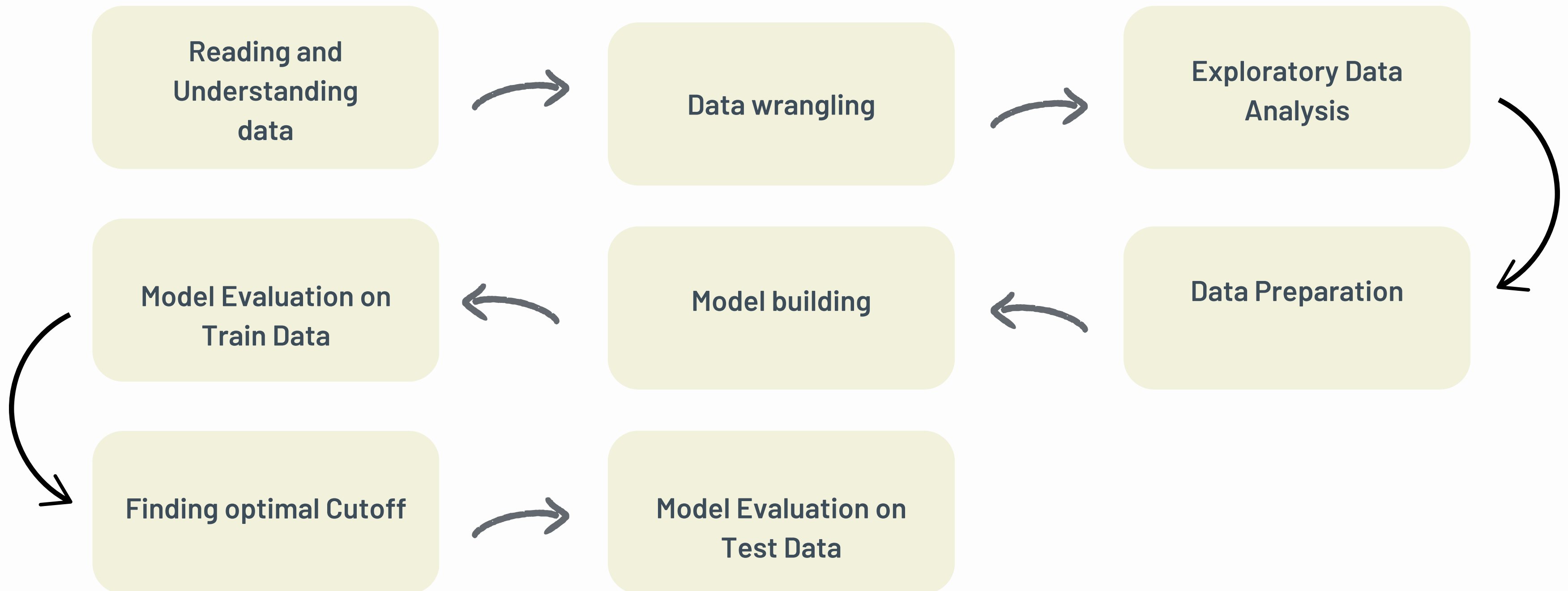
Case Study

INTRODUCTION

- X Education sells online courses to industry professionals.
- Acquires leads through its website, search engines, and referrals.
- High lead acquisition but poor lead conversion rate (~30%).
- Out of 100 leads, only 30 are converted into paying customers.
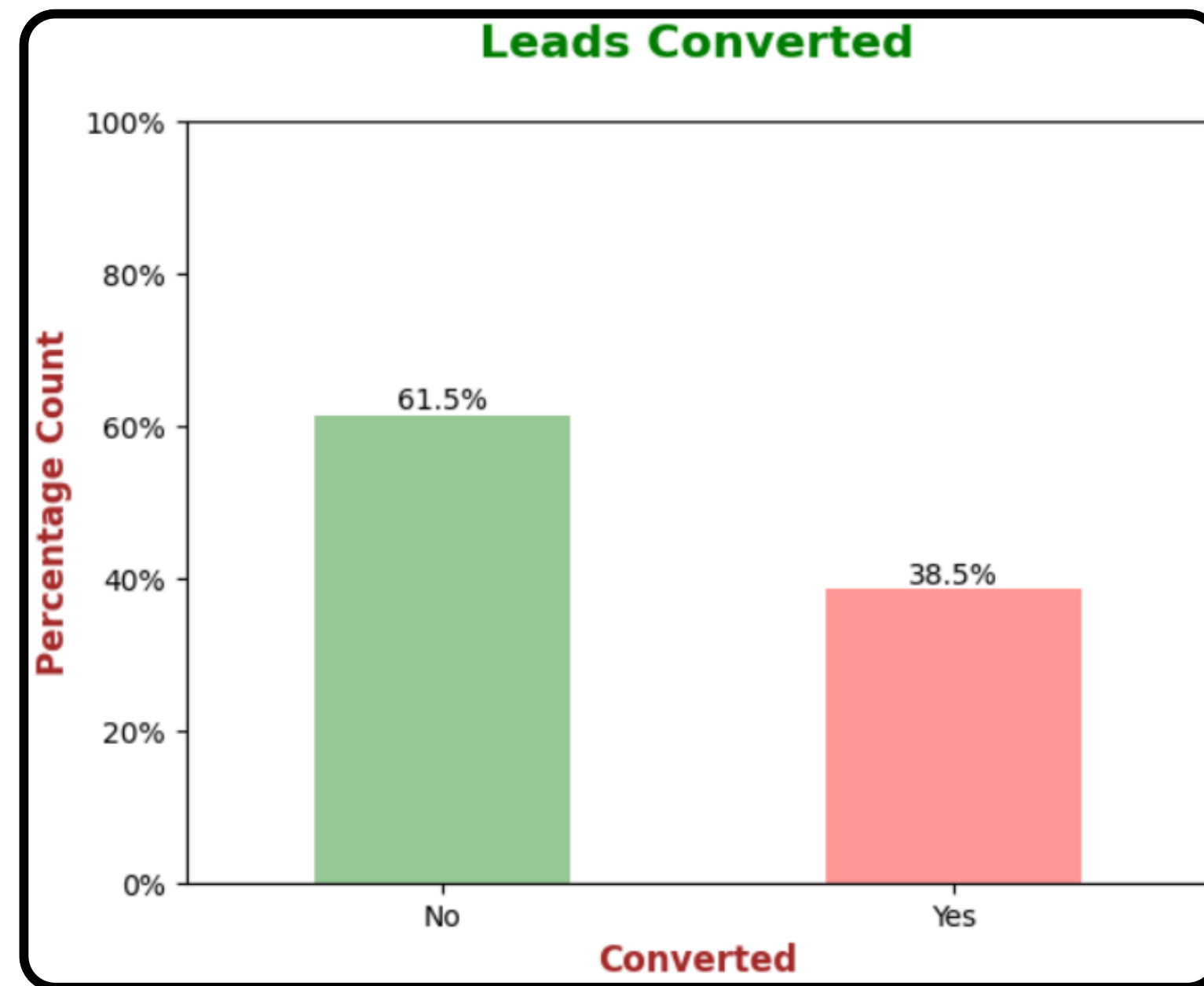- The company seeks to improve this conversion rate by identifying "Hot Leads."
-

- Historical data with 9000 leads.
- Includes features like Lead Source, Total Time Spent on the Website, and Last Activity.
- Handle levels in categorical variables like 'Select' which act as null values.
- By focusing on promising leads, the sales team can improve conversion efficiency.
- Potential increase in lead conversion rate as desired by the CEO.

OVERALL APPROACH

```
┌─────────────────────┐        ┌─────────────────────┐        ┌─────────────────────┐
│   Reading and       │        │                     │        │  Exploratory Data   │
│   Understanding     │  ──→   │   Data wrangling    │  ──→   │     Analysis        │
│      data           │        │                     │        │                     │
└─────────────────────┘        └─────────────────────┘        └─────────────────────┘
                                                                          │
                                                                          ↓
┌─────────────────────┐        ┌─────────────────────┐        ┌─────────────────────┐
│  Model Evaluation   │        │                     │        │                     │
│   on Train Data     │  ←──   │   Model building    │  ←──   │   Data Preparation  │
│                     │        │                     │        │                     │
└─────────────────────┘        └─────────────────────┘        └─────────────────────┘
         │
         ↓
┌─────────────────────┐        ┌─────────────────────┐
│  Finding optimal    │        │  Model Evaluation   │
│      Cutoff         │  ──→   │    on Test Data     │
│                     │        │                     │
└─────────────────────┘        └─────────────────────┘
```

LEADS CONVERTED

Leads Converted

- The conversion rate is 38.5%, meaning only 38.5% of the people have converted to leads. (Minority)
- While 61.5% of the people didn't convert to leads. (Majority)

UNI-VARIATE ANALYSIS

- Lead Origin: 53% of the customers came through "Landing Page Submission," while 39% were identified via "API."
- Current Occupation: 90% of the customers are categorized as Unemployed. Do Not Email: 92% of customers have opted out of receiving course-related emails.
- Lead Source: 58% of leads come from a combination of "Google" and "Direct Traffic."
- Last Activity: 68% of customer interactions involve "SMS Sent" and "Email Opened."
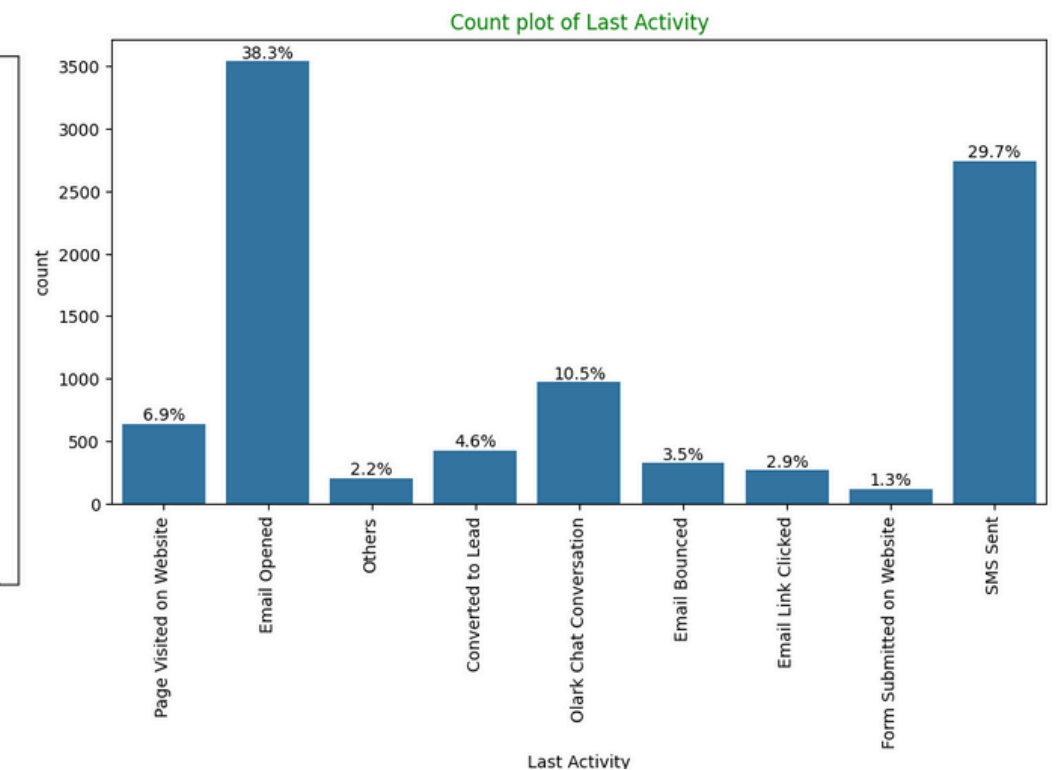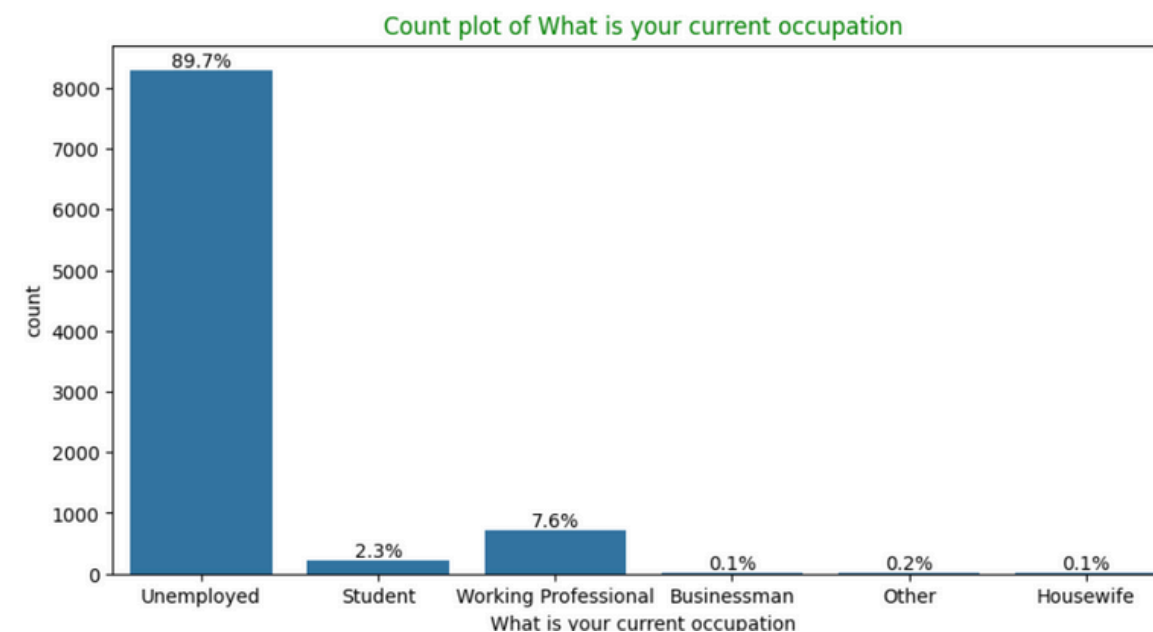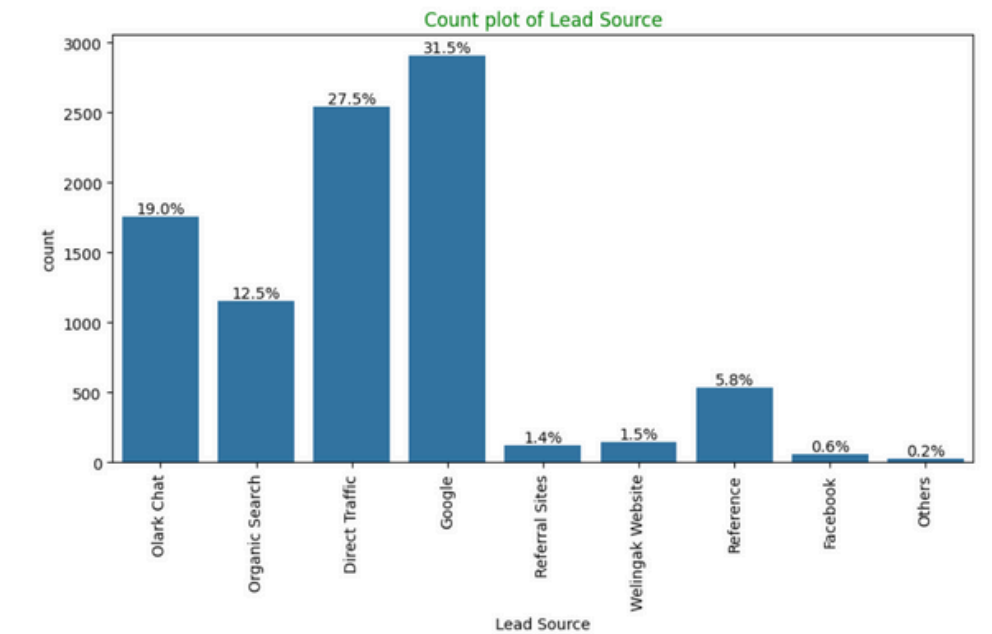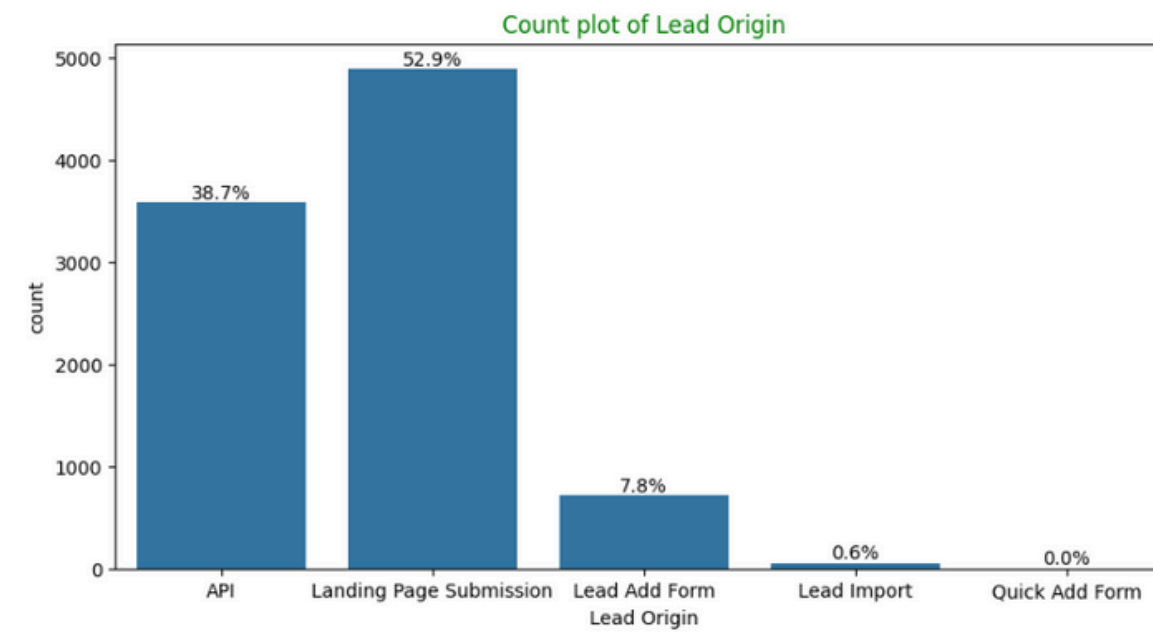
BI-VARIATE ANALYSIS

Lead Origin:
- Landing Page Submission accounts for 52% of leads with a 36% Lead Conversion Rate (LCR).
- API contributes 39% of leads with a 31% LCR.

Current Occupation:
- 90% of customers are Unemployed with a 34% LCR.
- Working Professionals comprise 7.6% of customers but have a high 92% LCR.

Do Not Email:
- 92% of people opted out of email communications.
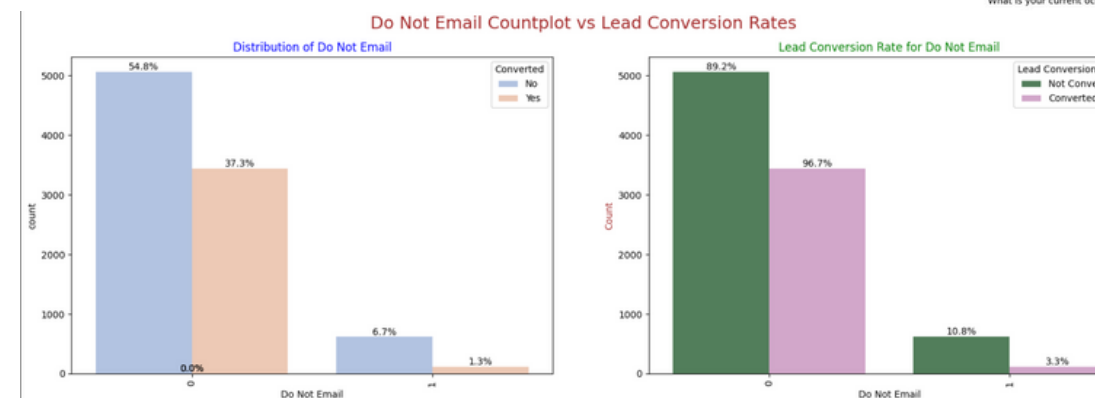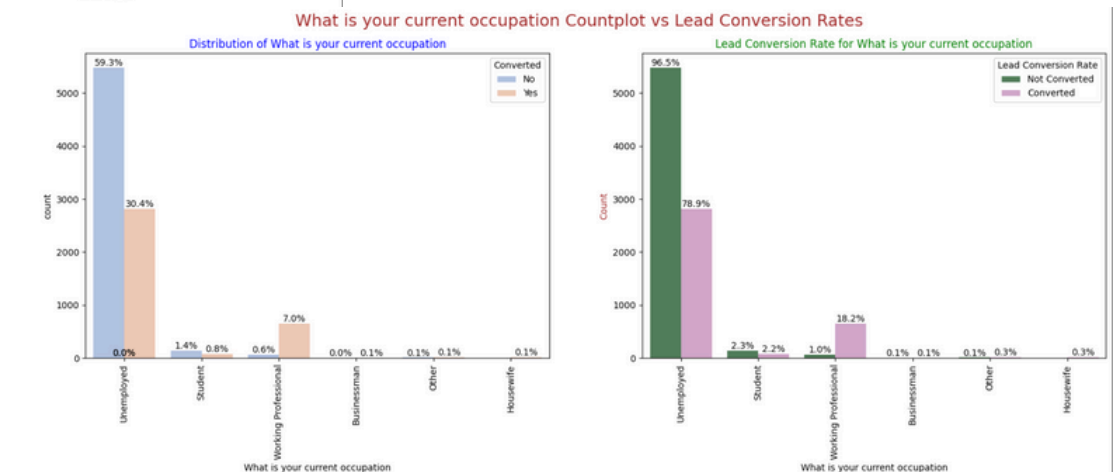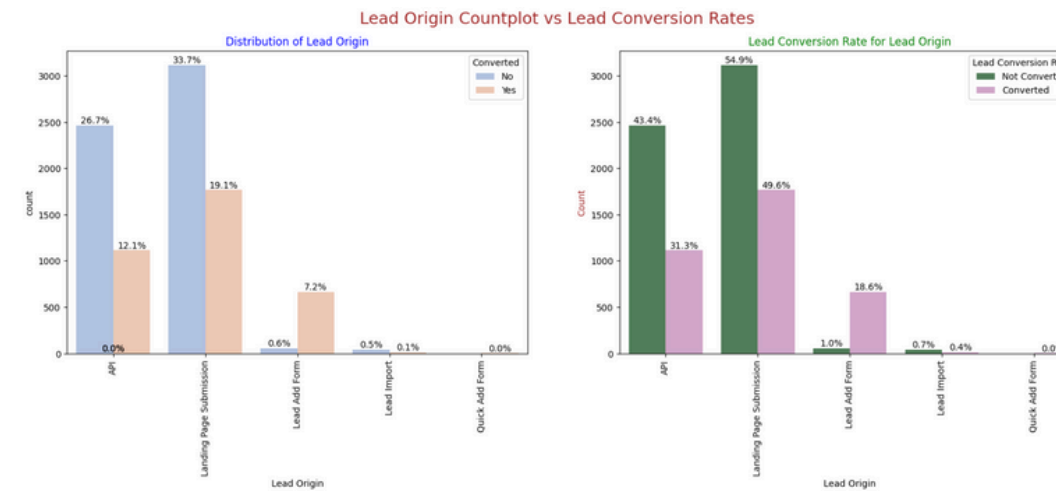
CONT.

Lead Source:
- Google has a 40% LCR, contributing 31% of leads.
- Direct Traffic: 32% LCR, 27% of leads.
- Organic Search: 37.8% LCR, 12.5% of leads.
- Reference: Highest LCR of 91%, but only 6% of leads come from this source.

Last Activity:
- SMS Sent has a 63% LCR, contributing 30% of last activities.
- Email Opened: 38% of last activities, with a 37% LCR.
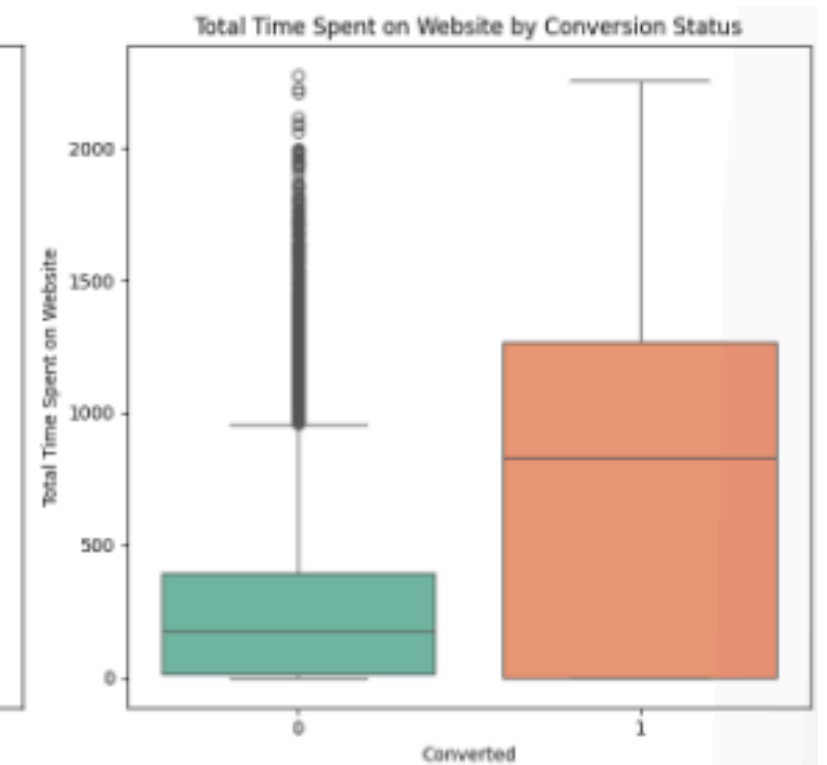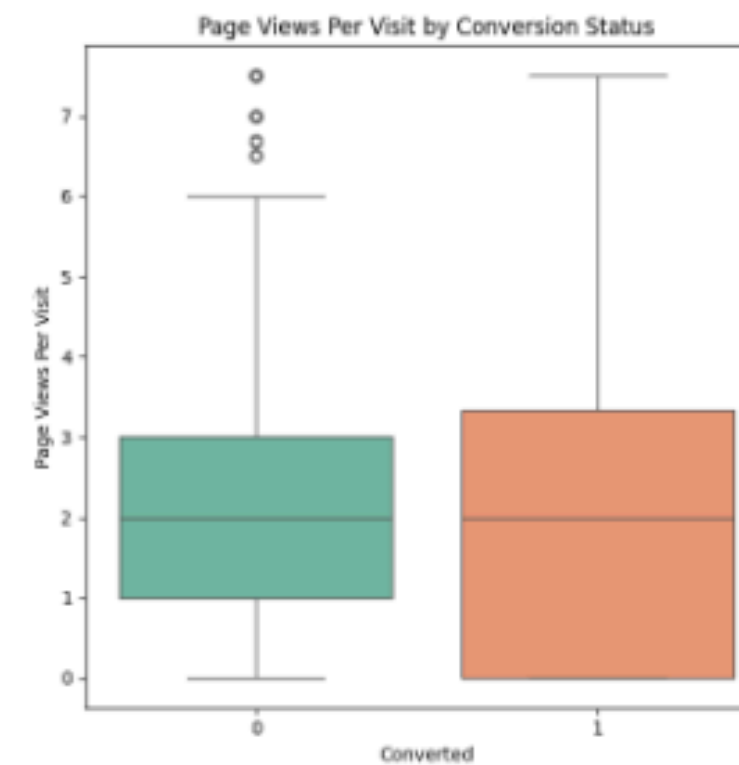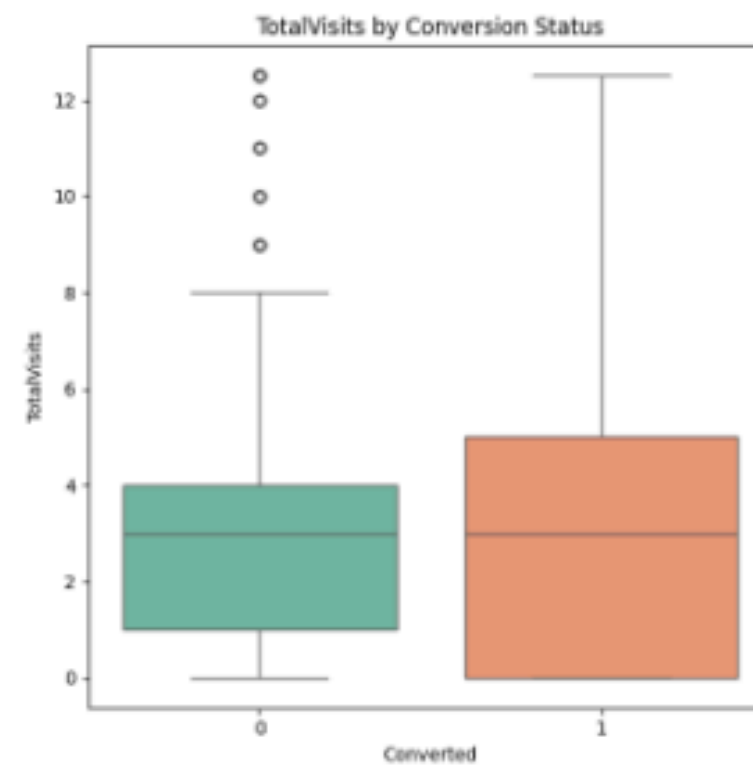
Specialization:
- Marketing, HR, and Finance Management show good contribution to lead conversions.
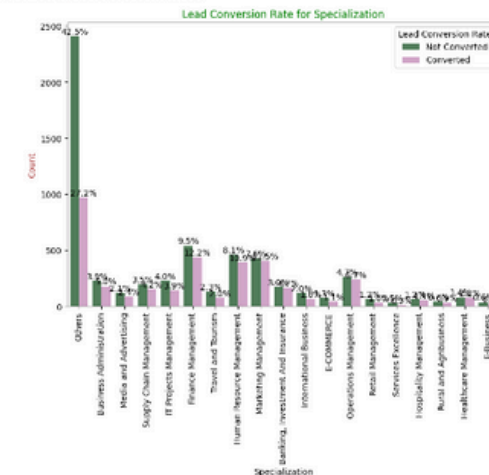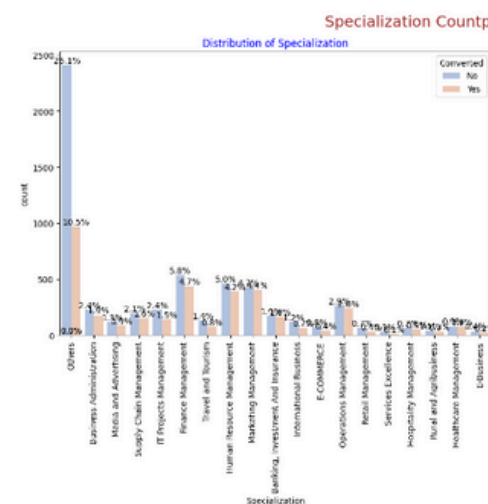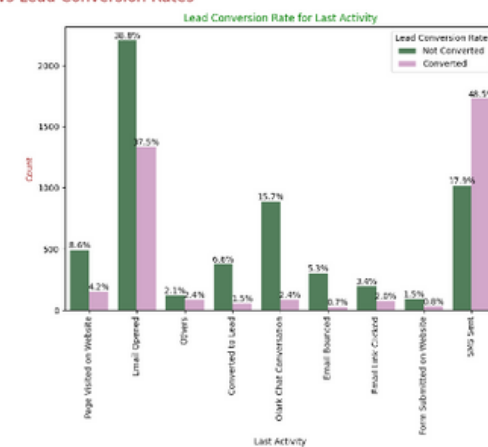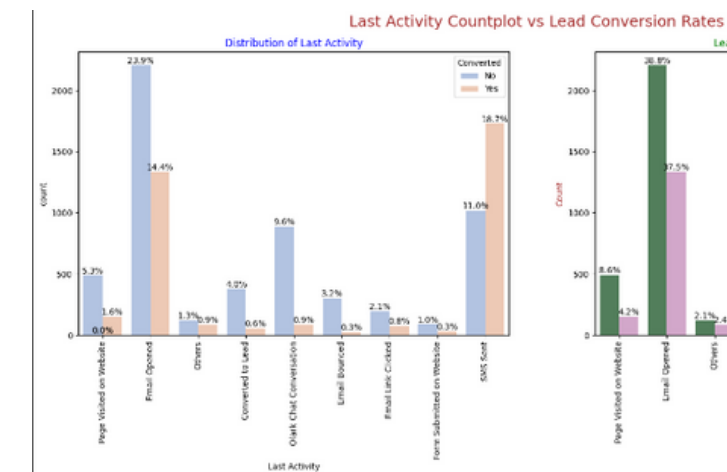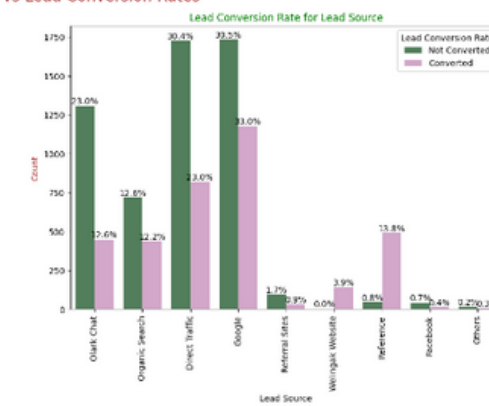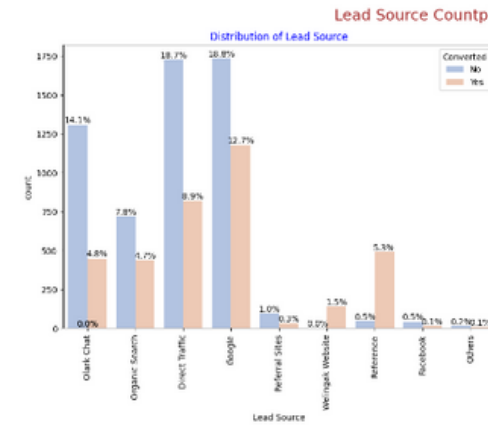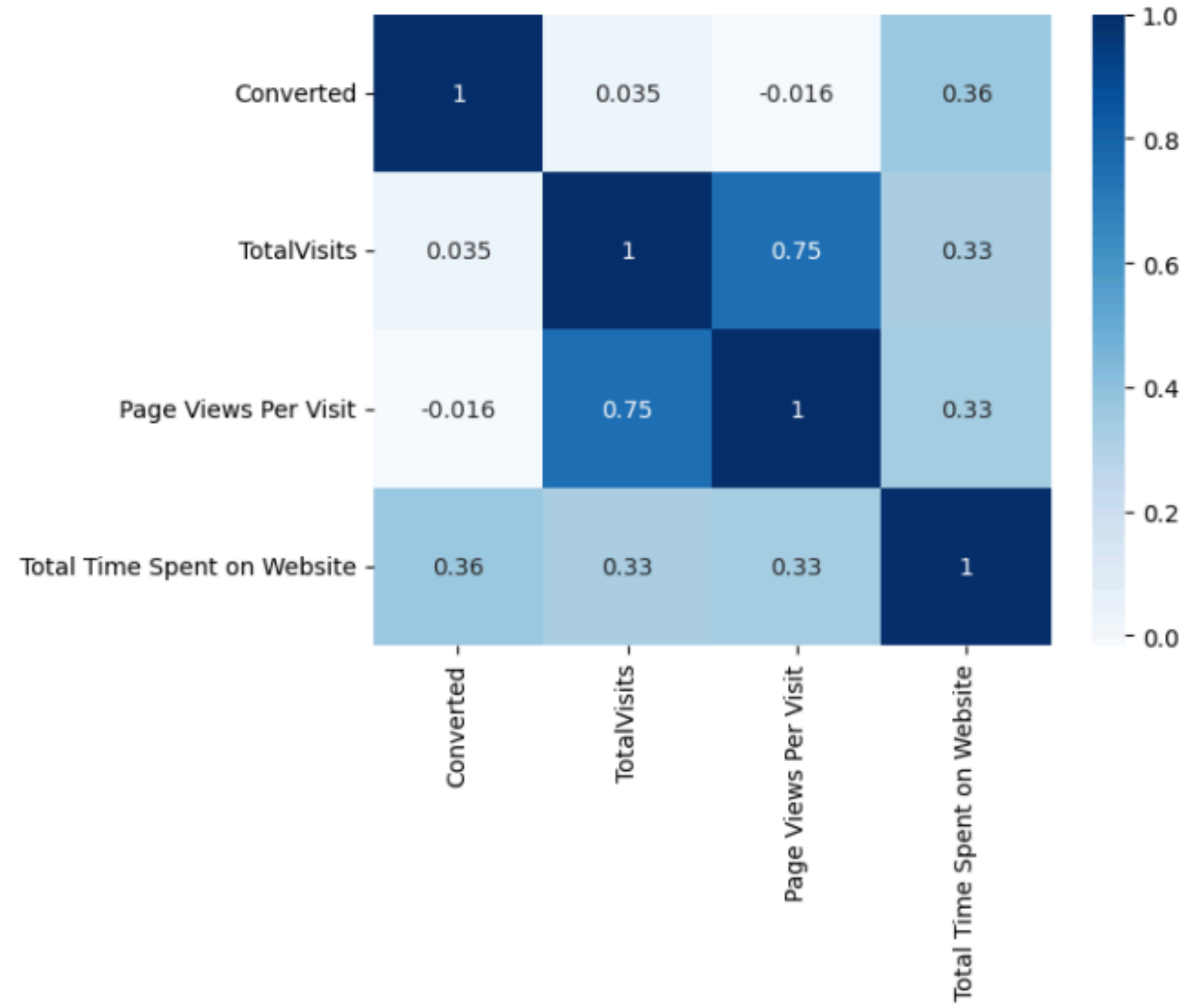
NUMERICAL VARIABLES

- The median value of total time spent on website for converted leads are considerably higher than the other group.
- Team should target those customers who are spending higher time on website.

HEATMAP

- There is a high correlation between Total Visits and page views per visit

MODEL EVALUATION

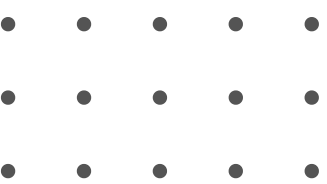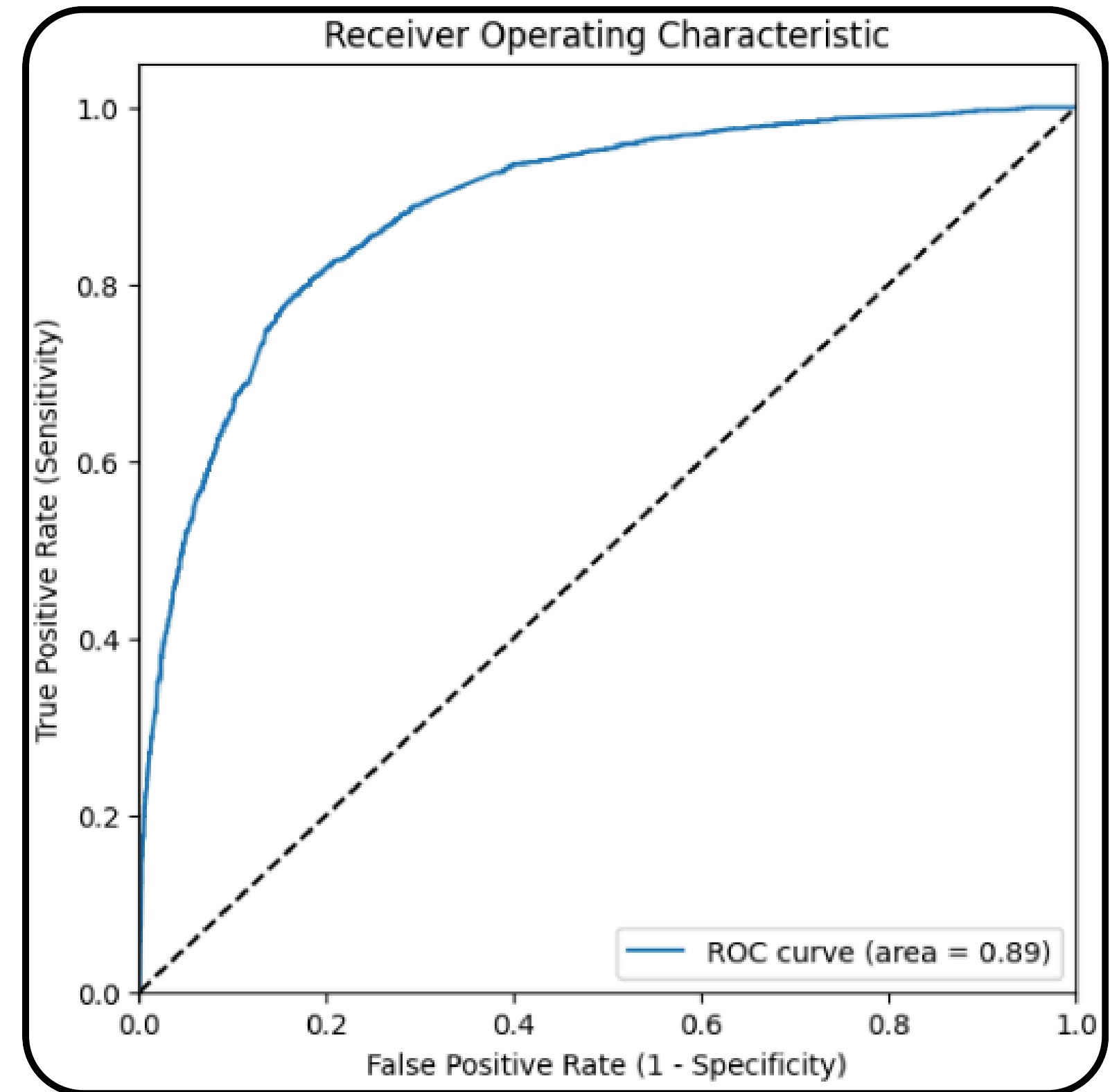**Train Data Set**

Accuracy : 81.11%

Sensitivity : 80.74%

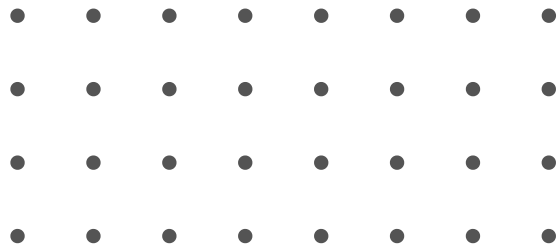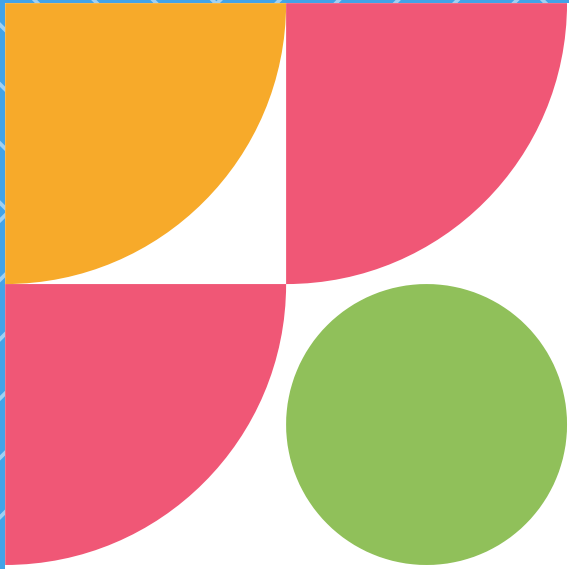Specifity : 81.33%

Test Data Set

Accuracy : 80.77%

Sensitivity : 80.27%

Specifity : 81.10%

Receiver Operating Characteristic

ROC curve (area = 0.89)

RECOMMENDATIONS

- Tailor messaging specifically for working professionals to boost engagement.
- Optimize communication methods based on how well leads respond to different channels.
- Increase budget for advertising on Welingak Website to enhance visibility.
- Offer incentives or discounts for referrals that convert, encouraging more people to refer others.
- Aggressively target working professionals, as they have a higher conversion rate and are more likely to afford premium fees.
- Focus marketing efforts on features with positive coefficients to better target potential customers.
- Create strategies to attract high-quality leads from the best-performing lead sources.
- Analyze specializations with negative coefficients to identify areas where offerings may need adjustment.
- Review the landing page submission process to find potential improvements for better lead conversions.

THANK YOU