



# **LEAD SCORING CASE STUDY**

**LANCHANA  
LOKESH**

**KANISHKA  
BALWANI**

**MALHAAR  
SINHA**




# PROBLEM STATEMENT

- An education company named X Education sells online courses to industry professionals. On any given day, many professionals who are interested in the courses land on their website and browse for courses.
- The company markets its courses on several websites and search engines like Google. Once these people land on the website, they might browse the courses or fill up a form for the course or watch some videos. When these people fill up a form providing their email address or phone number, they are classified to be a lead. Moreover, the company also gets leads through past referrals. Once these leads are acquired, employees from 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 X education is around 30%.
- There are a lot of leads generated in the initial stage (top) but only a few of them come out as paying customers from the bottom. In the middle stage, you need to nurture the potential leads well (i.e. educating the leads about the product, constantly communicating etc. ) in order to get a higher lead conversion.
- X Education has appointed you to help them select the most promising leads, i.e. the leads that are most likely to convert into paying customers. The company requires you to build a model wherein you need to assign a lead score to each of the leads such that the customers with a higher lead score have a higher conversion chance and the customers with a lower lead score have a lower conversion chance. The CEO, in particular, has given a ballpark of the target lead conversion rate to be around 80%.

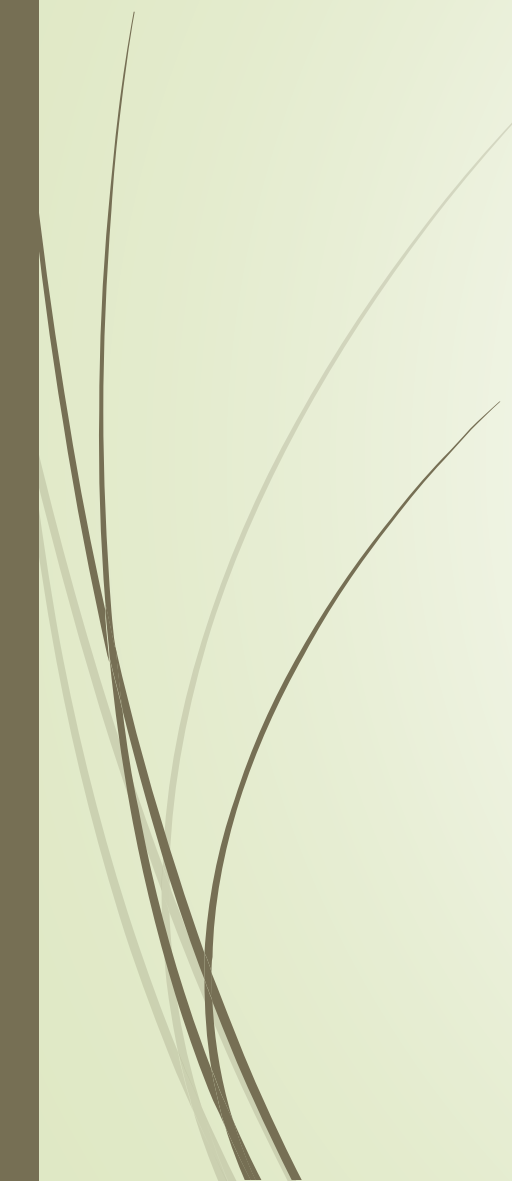


# BUSINESS OBJECTIVE

- Lead X wants us to build a model to give every lead a lead score between 10-100. So that they can identify the Hot Leads and increase their conversion rate as well.
  - The CEO want to achieve a lead conversion rate of 80%.
  - They want the model to be able to handle future constraints as well like Peak time actions required, how to utilize full man power and after achieving target what should be the approaches.
- 

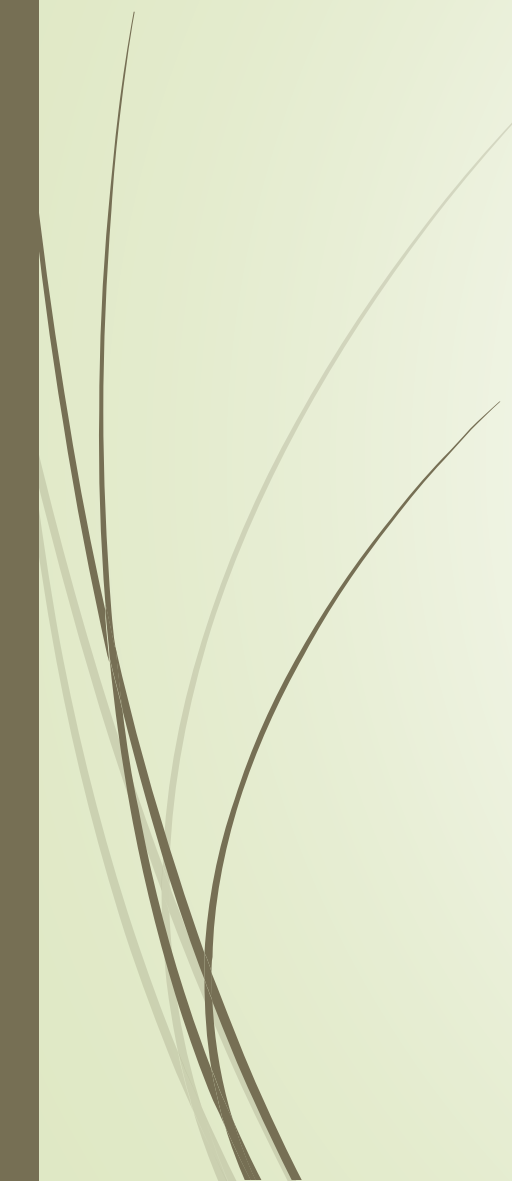


# APPROACH

- ❖ Importing the data and inspecting the data frame
  - ❖ Data preparation
  - ❖ Dummy variable creation
  - ❖ Test-Train split
  - ❖ Feature Scaling
  - ❖ Correlations
  - ❖ Model Building (RFE Rsquared VIF and pvalues)
  - ❖ Model Evaluation
  - ❖ Making predictions on test set.
- 




# DATA SOURCING & CLEANING

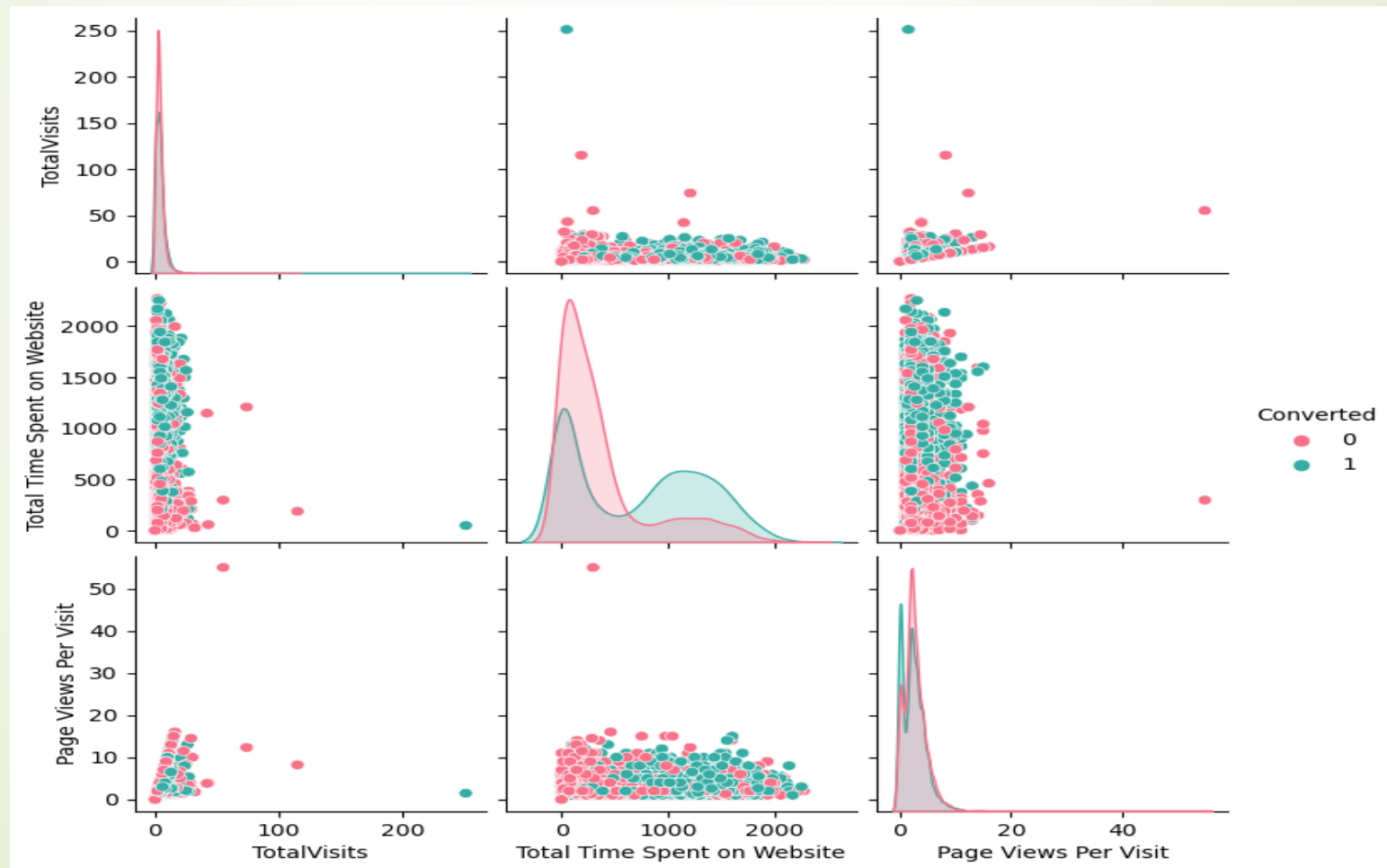
- ▶ Read the data from CSV File
  - ▶ Data cleaning -Handling Null Values & removing higher Null values data
  - ▶ Removing Redundant columns in the data
  - ▶ Imputing Null Values
- 



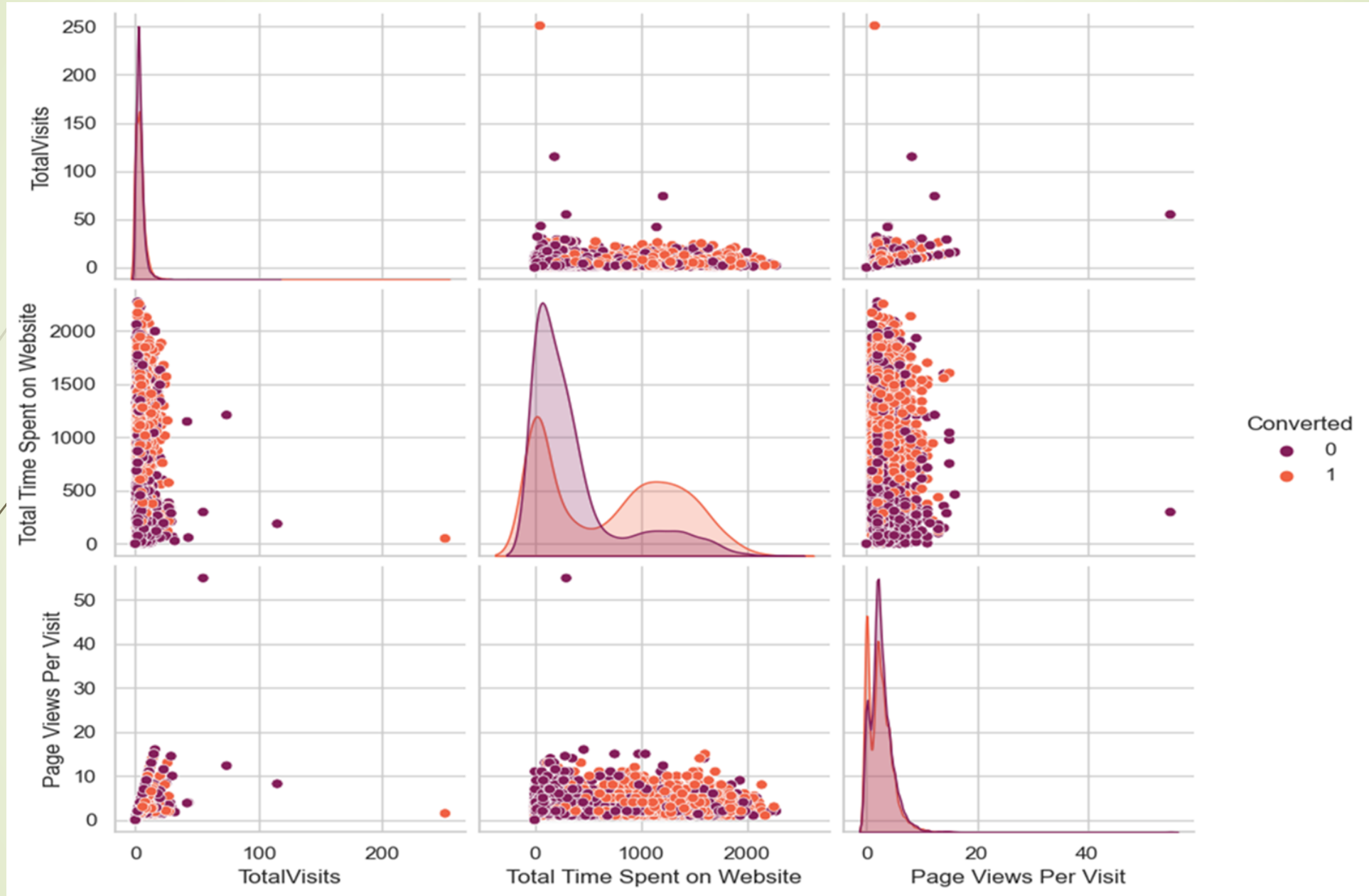
# MODEL BUILDING

- Splitting into train and test set
  - Scale variables in train set
  - Build the first model
  - Use RFE to eliminate less relevant variables
  - Build the next model
  - Eliminate variables based on high p-values
  - Check VIF value for all the existing columns
  - Predict using train set
  - Evaluate accuracy and other metric
  - Predict using test set
- 

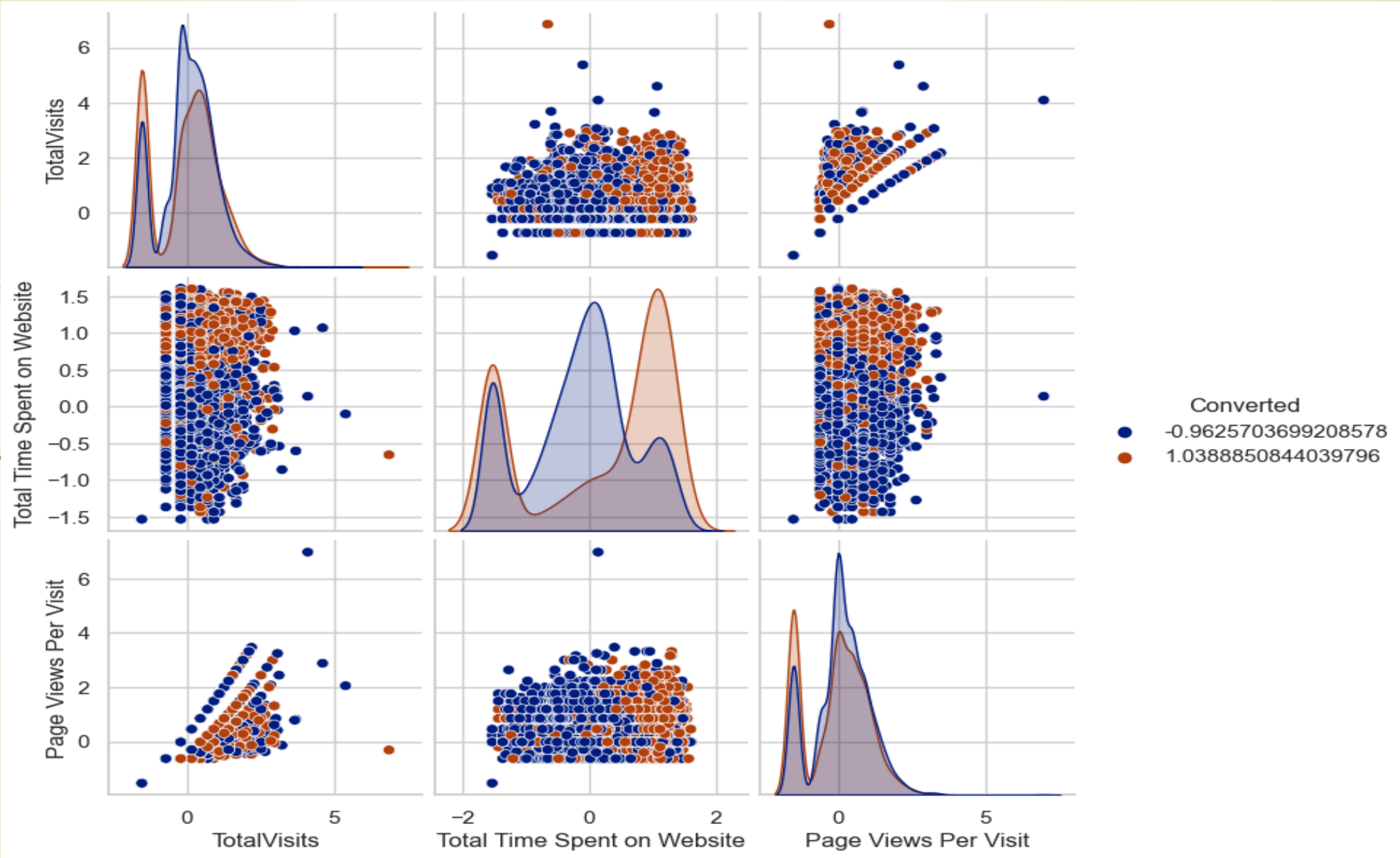
# PLOTS & CORRELATIONS

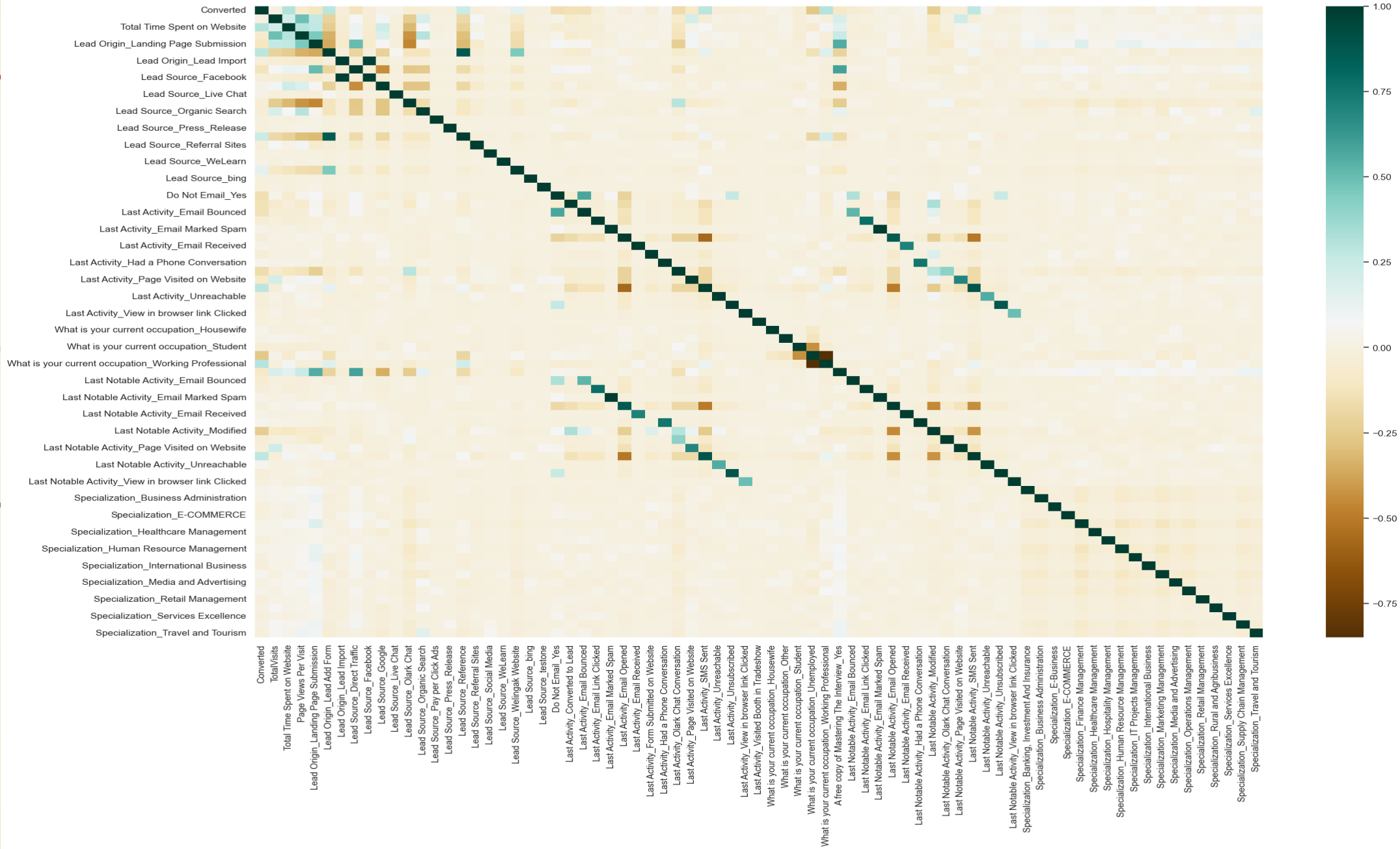




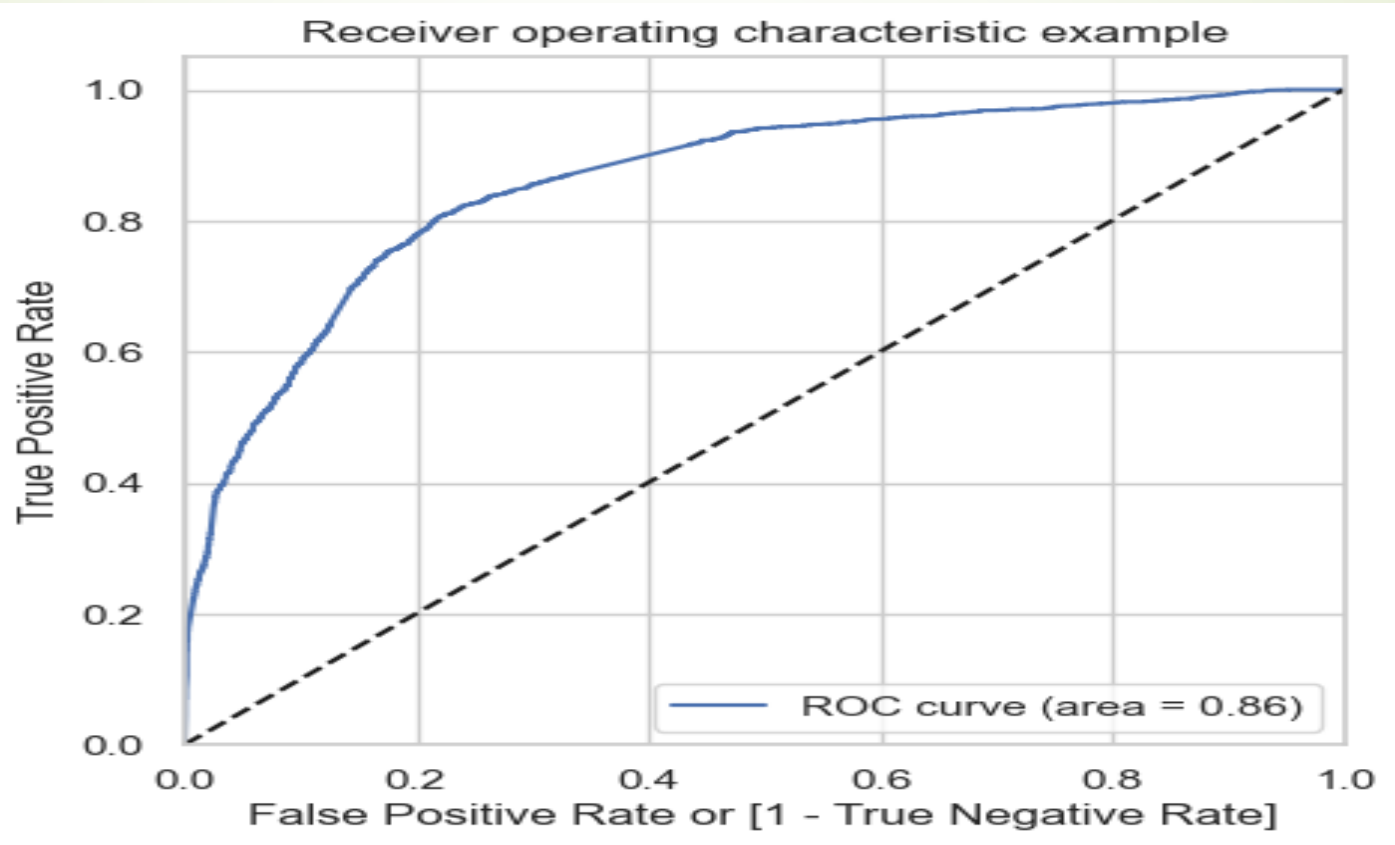






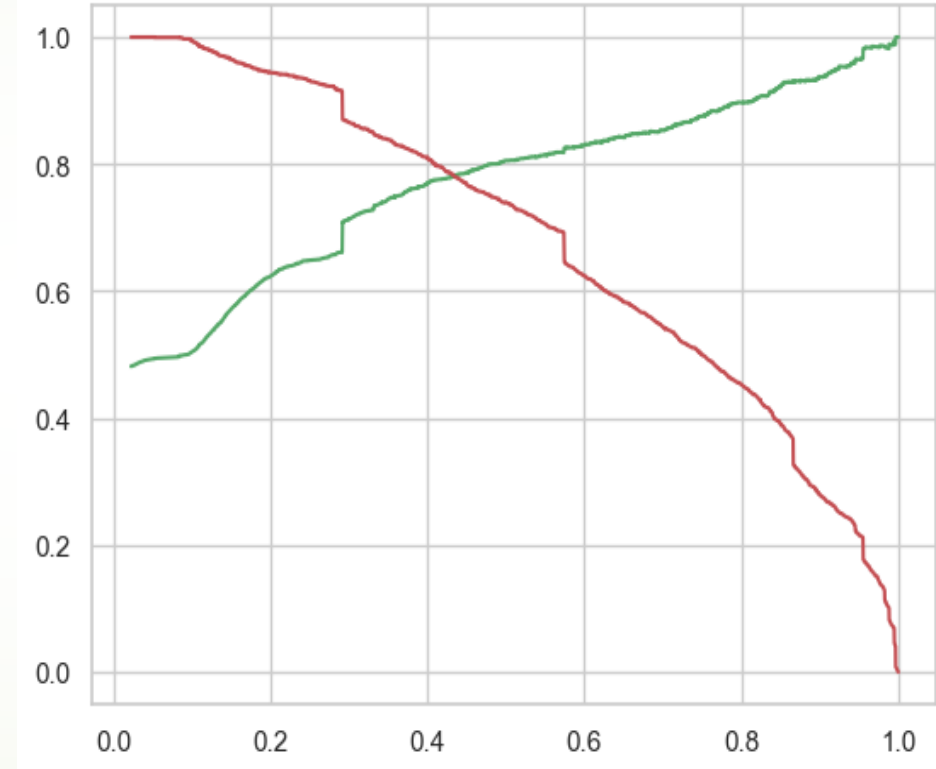
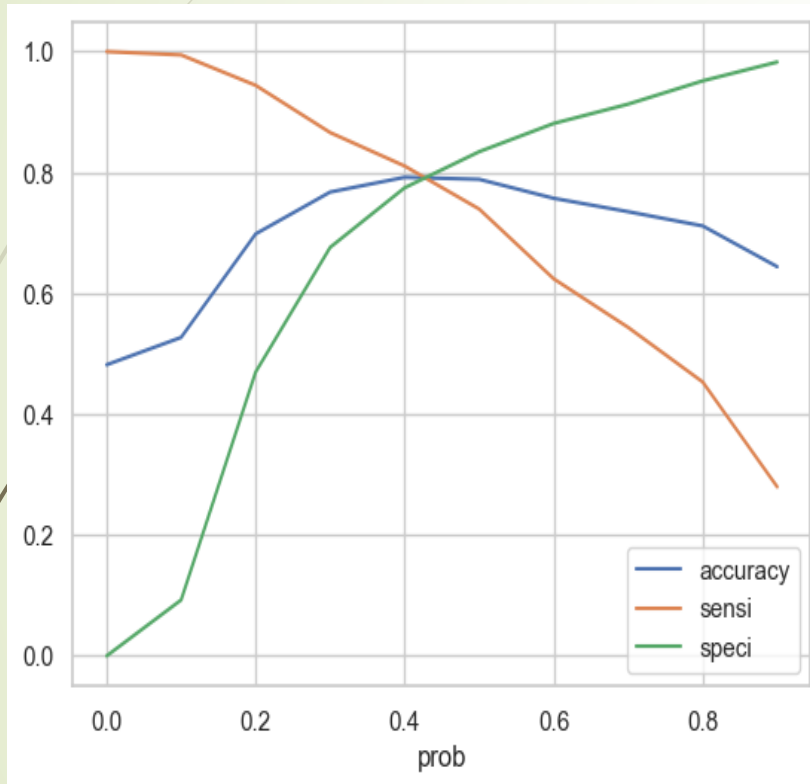


# ROC CURVE



The area under the curve of the ROC is 0.86 which is quite good.

# MODEL EVALUATION PLOTS



0.42 is the tradeoff between Precision and Recall -  
Thus we can safely choose to consider any Prospect Lead with  
Conversion Probability higher than  
42 % to be a hot Lead.



# MODEL EVALUATION

Train Set	Test Set
Accuracy = 78.95%	Accuracy = 78.61%
Sensitivity = 79.34%	Sensitivity = 79.70%
Specificity = 78.85%	Specificity = 77.31%
Precision = 78.40%	Precision = 77.54%
Recall = 77.71%	Recall = 77.84%



# CONCLUSION

- ❖ Engaging with individuals who spend more time than average could be a strategic move for increasing conversions.
- ❖ SMS campaigns could significantly influence lead conversion rates. Submissions on landing pages are an effective method to generate additional leads.
- ❖ Specializations in marketing management and human resources management exhibit high conversion rates, indicating that individuals from these fields may be valuable leads.
- ❖ Implementing referral incentives could enhance conversion rates.
- ❖ Alerts and informational messages have been observed to effectively convert leads.