Summary

The analysis is done for X Education and to find the most promising leads i.e the leads that are most likely to convert into potential buyers.

This dataset consists of various attributes such as Lead Source, Total Time Spent on Website, Total Visits, Last Activity, etc. which may or may not be useful in ultimately deciding whether a lead will be converted or not.

The target variable, in this case, is the column 'Converted' which tells whether a past lead was converted or not wherein 1 means it was converted and 0 means it wasn't converted.

The following are the steps followed :
1. Reading the dataset
2. Cleaning the dataset
    a. The null values in the data are first cleaned. Columns with more than 40% null values are removed.
    b. Remaining null values in the columns are imputed with the mean and median for numerical columns and mode for categorical columns.
    c. The columns which have "Select" are also replaced by null values.
3. EDA:
    a. A number of irrelevant categorical columns are found in the data on performing a basic EDA.
    b. Numerical columns are good with almost no outliers.
4. Creating Dummy variables:
    a. Dummy variables are created for categorical columns.
    b. For numerical columns Standardscaler is used.
5. Train - test split:
    a. The data was split into train and test data in 70:30 ratio.
6. Model building:
    a. Using RFE we selected the most relevant 15 variables.
    b. ELiminating features manually using the VIF and p-values. Features with vif <5 and p-value < 0.05 are kept.
7. Prediction:
    a. Prediction was done on the test data
8. Precision and Recall.

Insights:

- The Accuracy, Precision, Recall values of the test data is in acceptable range.
- This model has compatibility with the company's business requirements.
- The model is in a stable state.
- The important features that contribute to the probability of a lead to get converted are:
  - Lead Origin_Lead Add Form
  - Total Time Spent on Website
  - Last Notable Activity_SMS Sent
  - Tags_Will revert after reading the email
  - Last Notable Activity_Modified
  - Tags_Ringing
- X Education has to take these features into consideration and flourish as they have a very high chance to get almost all potential buyers to change their mind and buy their courses.