Executive Summary Report

The analysis done on the Lead data set and we got some useful insights. From this analysis X Education company get the most potential leads i.e. Hot Leads which lead to increase the lead conversion rate and get more people to join their courses.

The Steps we have done for analysis are :

1. **Cleaning The Data :**
- After importing the data set, cleaning on the data set done in this step, the data set variables contains a lot of null values which are no use of analysis so we dropped several variables which are having more than 3000 null values.
- Dropped those variables which are having one value majorly.
- Dropped null rows from the couple of variables.
- Renamed some of the column for better understanding.
- Checked the outliers present in the some important variables.

2. **Exploratory Data Analysis :**
- EDA techniques applied on the data set to check the condition of the data and we found that a lot of elements in the categorical variables were irrelevant.
- The heat map showed that there was no highly correlated variables present in the data set.

3. **Data Splitting :**
- The data was split into 70% of train data and 30% of test data.
4. Model Building :
- RFE method done to attain 15 relevant variables.
- The model was build by removing the variables manually depending on the VIF values and p-values. The condition for VIF and p-value are they should be <5 and <0.05 respectively.

5. **Model evaluation :**
- Here we made a confusion matrix.
- Based on the cutoff value (using ROC curve) was used to find the accuracy, sensitivity and specificity which are came to be around 80%.

6. **Model Predictions**
   (X-test) :
- The prediction was done on the test data set with an optimum cut-off value as 0.42 and we get :
   1. Accuracy : 76 %
   2. Sensitivity : 83 %
   3. Specificity : 70 %
   4. Precision : 78 %
   5. Recall : 78 %

(y-test) :

➕   We made the predictions on the test data set using cut-off value 0.42 and we get :

1. Accuracy : 76 %
2. Precision : 71%
3. Recall : 83 %

❖ **Major Contributors :**

The variables that mattered the most in potential buyers are as follows :

1.Total Visits
2.Time Spent On Website
3.Page Views Per Visit

These are the major contributed variables which contributed the most towards the probability of lead getting converted and sales increase.

❖ **Other Contributors :**

Along with above major variables we should look upon few other variables which are :

1. When the lead origin was :
   - Lead Add Form
2. When the lead source was :
   - Olark Chat
   - Reference
   - Welingak Website
3. When current occupation was :

   - Student
   - Working Professional
4. When the last activity was :
   - Phone Conversation
   - SMS Sent

By targeting the below variables X Education company can flourish as they have a very high chance to get almost all the potential buyers to change their mind and buy their courses. By focussing on below variables company can increase their lead conversion rate upto 80%.