Summary

In order to support X education and encourage more business professionals to enrol in their courses, this analysis is conducted. We learned a lot about the potential clients' visitation patterns, length of stay, mode of transportation, and other details from the basic data provided the rate of conversion.

The procedures employed are as follows:

- Cleaning Data:

With the exception of a few null values, the data was mostly clean. The option select has replaced as it didn't provide us with much information, with a null value.

- EDA:

We performed a brief EDA to assess the quality of our data. It was discovered that numerous components were unimportant in the category variables. The numbers appear to be in ranges with no anomalies were discovered.

- Dummy Variables:

After creating the dummy variables, the ones that had "not provided" elements were eliminated. The MinMaxScaler was utilised for numerical values.

- Train-Test Split:

For the train and test data, the split was performed at 70% and 30%, respectively.

- Model Building:

To begin with, REF was used to identify the top 15 pertinent variables. Afterwards, based on the p-values and VIF values, the remaining variables were manually eliminated.

- Model Evaluation:

A matrix of confusion was created. Afterwards, the ideal cut-off value was employed to determine the 80% accuracy, 80% sensitivity, and 80% specificity of each.

- Prediction:

An ideal cut off of 0.35 was used for the prediction, which was performed on the test data frame with 80% accuracy, sensitivity, and specificity.

- Precision – Recall:

Using this same procedure for a recheck, a cut off of 0.41 was discovered on the test data frame with a precision of about 73.24% and recall of around 76.61%.