

## Summary

This analysis was conducted for X Education to determine methods for attracting more industry professionals to their courses. The provided data gave us valuable insights into the behavior of potential customers, including their website visit frequency and duration, as well as how they found the site and the conversion rate.

The steps taken to analyze this information are:

1. Cleaning data: Removing null values and replacing them with appropriate information. Creating categories for better analysis.
2. Exploratory Data Analysis (EDA): Checking the quality and structure of the data.
3. Creating Dummy Variables: Encoding categorical variables for analysis.
4. Train-Test Split: Splitting the data into training and testing sets for model building and evaluation.
5. Model Building: Selecting relevant variables using Recursive Feature Elimination (RFE) and removing others based on VIF values and p-values.
6. Model Evaluation: Using a confusion matrix to determine accuracy, sensitivity and specificity. Optimal cut off found to be 0.35.
7. Prediction: Applying the model on the test data with a precision and recall of 73% and 75% respectively with cut off 0.41.
8. Precision-Recall: This method was also employed to verify the results, and a cutoff of 0.41 was determined, yielding a precision of 73% and a recall of 75% on the test data set.

It was found that the following variables matter the most to potential buyers, in descending order:

- Total time spent on the website
- Total number of visits
- Lead source (Google, direct traffic, organic search, Welingak website)
- Last activity (SMS, Olark chat conversation)
- Lead origin (Lead add format)
- Current occupation as a working professional

By considering these factors, X Education has the potential to attract a high number of industry professionals to enroll in their courses.