1. X Education aimed to build a logistic regression model to assign lead scores to potential leads, optimizing customer acquisition.

2. The dataset underwent preprocessing, including handling missing values by removing features with over 30% missing data and treating features with constant values.

3. During exploratory data analysis (EDA), significant insights were gathered:

- API and Landing Page submissions generated the most leads, while Lead Add Form had a higher conversion rate.
- Lead Import and Quick Add Form contributed few leads, with Quick Add Form showing zero conversion.
- Direct Traffic and Olark Chat brought in many leads but with low conversion, while Google had good lead inputs and decent conversion.
- References showed the highest conversion rate.
- Activities like having a Phone Conversation and sending SMS appeared to generate leads with higher conversion rates.
- Although unemployed individuals comprised most leads, their conversion rate was low, about half of the leads.
- Businessmen and Working Professionals had higher conversion rates.
- While Housewives generated fewer leads, they tended to convert well.

4. Data formatting included standard scaling of numerical features and encoding categorical variables.

5. The dataset was split into 70% training and 30% testing data.

6. Model building involved feature selection based on significance (p-values), removing features with p-values above 0.05, and checking for multicollinearity using the Variance Inflation Factor (VIF).

7. Predictions were made on both train and test datasets, and performance metrics such as Accuracy, Sensitivity, Specificity were calculated, achieving an accuracy close to 80%.

8. Finally, lead scores were computed for each customer, aiding in targeting potential leads more effectively.