Summary of Model Building and Prediction

1. **Exploratory Data Analysis (EDA):**
    - Checked for missing values and removed columns with over 45% missing data.
    - Replaced missing values in important columns with "not provided" and imputed "India" for a significant portion of missing data.
    - Dropped the "India" column due to its dominance (97% of data).
    - Addressed outliers and worked on dummy variables for categorical features.
2. **Train-Test Split & Scaling:**
    - Split data into training (70%) and testing (30%) sets.
    - Applied Min-Max scaling to the following variables: TotalVisits, Page Views Per Visit, and Total Time Spent on Website.
3. **Model Building:**
    - Used Recursive Feature Elimination (RFE) for feature selection, keeping the top 15 relevant variables.
    - Removed irrelevant variables manually based on Variance Inflation Factor (VIF) and p-values.
    - Accuracy achieved on the model: 80.91% with a confusion matrix.
4. **Model Evaluation:**
    - **Sensitivity-Specificity Evaluation:**
        - Training Data:
            - Optimal cutoff value = 0.35.
            - Accuracy: 80.91%, Sensitivity: 79.94%, Specificity: 81.50%.
        - Test Data:
            - Accuracy: 80.02%, Sensitivity: 79.23%, Specificity: 80.50%.
    - **Precision-Recall Evaluation:**
        - Training Data:
            - With cutoff 0.35: Precision: 77.82%, Recall: 70.31%.
            - Adjusted cutoff (0.44): Accuracy: 80.60%, Precision: 74.46%, Recall: 75.85%.
        - Test Data:

        - Accuracy: 80.86%, Precision: 72.53%, Recall: 75.28%.

5. **Optimal Cutoff Values:**
    - Sensitivity-Specificity: Optimal cutoff = 0.35.
    - Precision-Recall: Optimal cutoff = 0.44.
6. **Conclusion:**
    - Key factors influencing conversion rates:

1. Total time spent on the website.
2. Total number of visits.
3. Lead sources (Google, Direct Traffic, Organic Search, Welingak website).
4. Last activity (SMS, Olark chat).
5. Lead origin (Lead ad format).
6. Current occupation (Working professionals).
    - The model predicts the conversion rate well, offering X Education the opportunity to effectively target potential buyers and improve conversion rates, leading to potential growth.