Summary

Analysis for X Education: Enhancing Course Enrollment

This analysis seeks to identify the key factors that influence industry professionals' decisions to enroll in X Education's courses. By examining website interactions, engagement patterns, and demographic characteristics, we can formulate targeted strategies to enhance conversion rates.

Methodology

1. **Data Cleaning:** We addressed missing values and replaced irrelevant options with null values to improve data quality.
2. **Exploratory Data Analysis (EDA)**: Initial EDA indicated that categorical variables contained irrelevant elements, while numeric values were sound without outliers.
3. **Feature Engineering**: We created dummy variables, eliminated those with 'not provided' elements, and applied MinMaxScaler to numeric values.
4. **Train-Test Split:** The dataset was divided into 70% for training and 30% for testing.
5. **Model Building**: We used Recursive Feature Elimination (RFE) to identify the top 15 relevant variables, removing others based on Variance Inflation Factor (VIF) and p-values.
6. **Model Evaluation:** A confusion matrix was constructed, and the optimal cutoff value was determined using the ROC curve, achieving around 80% accuracy.

Key Findings

- **Total Time Spent on Website:** Longer engagement indicates higher interest.
- **Total Number of Visits**: Frequent visits suggest strong purchase intent.
- **Lead Source**: Google and direct traffic are most effective.
- **Last Activity**: Recent interactions reflect current interest.
- **Lead Origin**: Certain formats yield higher conversion rates.
- **Current Occupation:** Working professionals are more likely to invest in courses.

Focusing on these variables can enhance X Education's marketing strategies and drive business growth.