Summary Report

Introduction:

X Education, a provider of online courses for industry professionals, faces the challenge of improving a low lead conversion rate of approximately 30%. The objective of this project was to leverage predictive modelling to better identify and prioritize 'hot leads', aiming to increase conversion rates significantly.

Methodology:

The analysis began with a comprehensive review of the dataset, comprising approximately 9,000 data points with variables critical to lead generation such as Lead Source, Total Time Spent on Website, and Total Visits. Initial data preparation involved cleaning activities such as handling missing values and removing irrelevant identifiers like Prospect ID and Lead Number.

Exploratory Data Analysis (EDA):

EDA revealed critical insights into variables significantly impacting lead conversion. Visualizations highlighted the distribution of 'Converted' leads and underscored the effectiveness of specific lead sources and origins in driving conversions.

Model Development:

A logistic regression model was chosen for its appropriateness in predicting binary outcomes like lead conversion. The model identified key features influencing lead conversion:

1. ***Lead Origin - Lead Add Form:*** This feature's high positive coefficient (+2.063) suggested its strong impact in increasing the likelihood of lead conversion.
2. ***Lead Profile - Student of Some School:*** This feature's significant negative coefficient (-1.844) indicated a lower likelihood of conversion for this segment.
3. ***Lead Source - Welingak Website:*** A positive coefficient (+1.517) highlighted the effectiveness of this source in enhancing lead conversion chances.

These findings informed the development of a strategic lead scoring system.

Lead Scoring System:

The model's probability outputs were translated into a lead scoring mechanism, assigning scores from 0 to 100 based on conversion likelihood. This system enabled the sales team to prioritize leads effectively, with a focus on those showing the highest potential for conversion.

Results and Business Impact:

The implementation of the lead scoring system has been instrumental in optimizing lead management processes, contributing to a more focused approach in targeting high-potential leads. This targeted approach is expected to significantly boost overall conversion rates.

Learnings:

The project emphasized the value of data-driven strategies in sales and marketing. It highlighted the importance of clean data, the effectiveness of logistic regression in binary outcome prediction, and the benefits of implementing a systematic lead scoring system. Continuous adaptation based on data-driven insights was identified as crucial for ongoing improvement.

Conclusion:

This analytics project has set a new standard for lead management at X Education, integrating advanced data analysis with traditional sales strategies to enhance efficiency and effectiveness in meeting sales targets.