### **1. Handling Data Quality Issues**

**Problem:** Missing values, duplicate records, and inconsistent data types affected the dataset. **Solution:**

* Imputed missing values using mean/median for numerical variables and mode for categorical variables.
* Removed duplicate entries to avoid redundancy.
* Standardized data formats to maintain consistency.

### **2. Feature Selection and Engineering**

**Problem:** Some features added noise and did not contribute significantly to the prediction. **Solution:**

* Performed correlation analysis to identify highly correlated variables.
* Used feature importance scores to retain only the most relevant predictors.
* Created new derived features where necessary to enhance model performance.

### **3. Model Performance Improvement**

**Problem:** The initial logistic regression model had suboptimal performance with misclassifications. **Solution:**

* Optimized model parameters using hyperparameter tuning.
* Implemented regularization techniques (L1 and L2) to prevent overfitting.
* Balanced the dataset if necessary to reduce bias towards a particular class.

### **4. Model Evaluation and Business Interpretation**

**Problem:** Evaluating model accuracy alone was insufficient for business decisions. **Solution:**

* Multiple evaluation metrics such as precision, recall, F1-score, and ROC AUC were used.
* Analyzed confusion matrix to understand false positive and false negative rates.
* Interpreted the model results in business terms to aid decision-making.

### **5. Scalability and Future Adaptability**

**Problem:** Business requirements may evolve, needing changes in the model. **Solution:**

* Ensured the model can be retrained easily with new data.
* Designed the pipeline to allow integration with advanced models if needed.
* Provided recommendations for continuous model monitoring and updates.

### **6. Business Strategy Recommendations**

**Problem:** The company needed actionable insights from the model results. **Solution:**

* Prioritize follow-up on leads with high scores to optimize resource allocation.
* Regularly evaluate the model's predictions against actual outcomes.
* Incorporate additional features, such as customer engagement data, for better accuracy.

By implementing these solutions, the logistic regression model provides a robust framework for lead scoring and business decision-making.