Logistic Regression Summary Report (Mode-Standardize)

Confusion Matrices

Training Confusion Matrix

2708	37
233	111

Testing Confusion Matrix

900	23
90	17

Evaluation Metrics Summary

	Train	Test
Accuracy	0.8994	0.8971
Precision	0.615	0.5796
Recall	0.218	0.2144
Class 0 F1	0.9457	0.9443
Class 1 F1	0.3219	0.3088
F1 Score	0.6338	0.6266
Weighted F1 Score	0.8774	0.8749

Interpretation of Results

1. General Performance:

- Logistic Regression shows strong overall accuracy with 91.25% on training and 89.02% on test.

2. Precision and Recall:

- Precision drops to 0.425 in test, suggesting more false positives.
- Recall is low (0.1588), indicating difficulty in detecting actual positives.

3. Class-wise F1:

- Class 0 F1 is high (>0.9), but Class 1 F1 drops from 0.45 to 0.23 - this shows underperformance on minority class.

4. Weighted Metrics:

- Weighted F1 is fairly stable due to majority class dominance.

5. Confusion Matrix:

- Many false negatives for Class 1 (train: 233, test: 90), need improvement in recall.

Recommendations and Model Tuning Suggestions

1. Class Imbalance Handling:

- Try oversampling (e.g. SMOTE) or set class_weight='balanced' to improve Class 1 recall.

2. Threshold Tuning:

- Adjust decision threshold to improve sensitivity (recall).

3. Hyperparameter Optimization:

- Tune regularization strength (C) and penalty using GridSearchCV.

4. Feature Engineering:

- Enhance feature set with domain knowledge or interactions.

5. Consider Alternative Models:

- Evaluate Random Forest, SVM, or boosting models for better recall and precision.