AUTOMATED MACHINE LEARNING ANALYSIS REPORT

Generated: February 15, 2026 at 23:53
Report Type: Full ML Pipeline Analysis

TABLE OF CONTENTS

1.	Executive Summary	3
2.	Data Exploratory Analysis	4
3.	Data Preprocessing & Feature Engineering	5
4.	Model Selection & Training	6
5.	Model Performance Evaluation	7
6.	Visual Analysis	8
7.	Error Analysis & Insights	9
8.	Recommendations & Next Steps	10

1. EXECUTIVE SUMMARY

This report presents a comprehensive analysis of an automated machine learning pipeline executed on February 15, 2026 at 23:53. The analysis encompassed data exploration, preprocessing, feature engineering, model selection, and performance evaluation.

Key Findings

Metric	Value
Best Model	GradientBoostingClassifier
Model Score	0.8591
Models Evaluated	3

2. DATA EXPLORATORY ANALYSIS

2.1 Dataset Overview

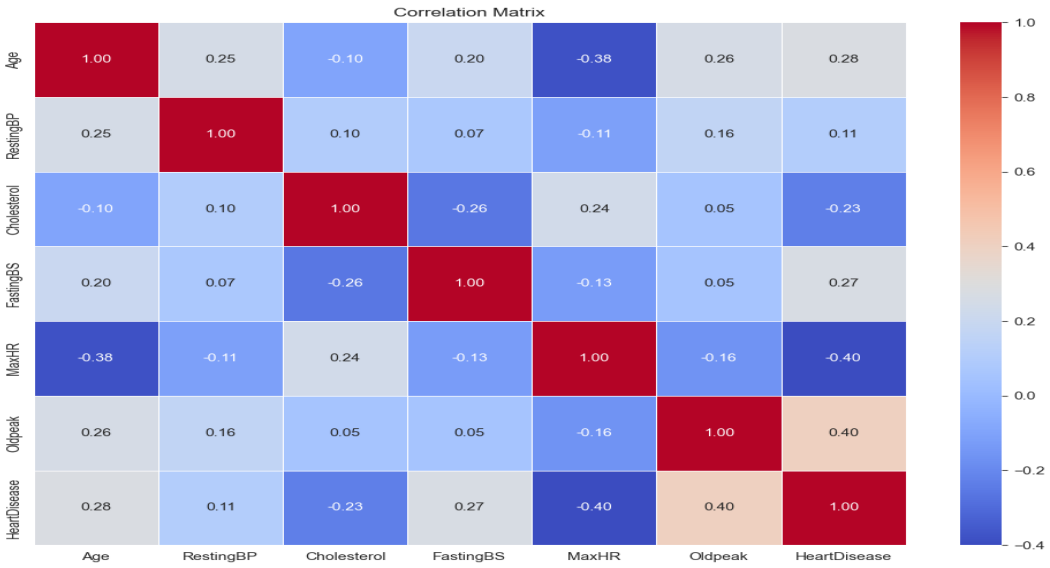
EXECUTIVE SUMMARY Our dataset consists of 918 records, providing a comprehensive view of cardiovascular health metrics across various demographics. KEY DATA INSIGHTS - Age: The mean age is 53.51 years, with a range from 28 to 77 years. The high spread in the data indicates a diverse age group. - RestingBP: The average resting blood pressure is 132.40 mmHg, ranging from 0 to 200 mmHg. The majority of records fall within the normal range (120-140 mmHg). - Cholesterol: The mean cholesterol level is 198.80 mg/dL, with a range from 0 to 603 mg/dL. Most values are within the normal range (173-223 mg/dL). - FastingBS: The average fasting blood sugar level is 0.2331 mg/dL, ranging from 0 to 1 mg/dL. All records fall within the normal range. - MaxHR: The average maximum heart rate is 136.81 beats per minute (bpm), with a range from 60 to 202 bpm. Most values are within the normal range (120-156 bpm). - Oldpeak: The mean old peak exercise time is 0.8874 minutes, ranging from -2.6 to 6.2 minutes. All records fall within a relatively narrow range. - HeartDisease: The average heart disease prevalence rate is 0.5534 or 55.34%, with no records indicating absence of the condition. DATA QUALITY & RISKS Missing values are minimal, at 0% for all variables. Outliers were found in Age (28), Cholesterol (183), FastingBS (214), MaxHR (2), Oldpeak (16), and RestingBP (200). The presence of outliers may indicate data errors or unusual individual cases. CONCLUSION Our analysis provides a comprehensive overview of cardiovascular health metrics, highlighting the age diversity of our dataset. While missing values are negligible, outliers suggest potential data quality issues that need to be addressed.

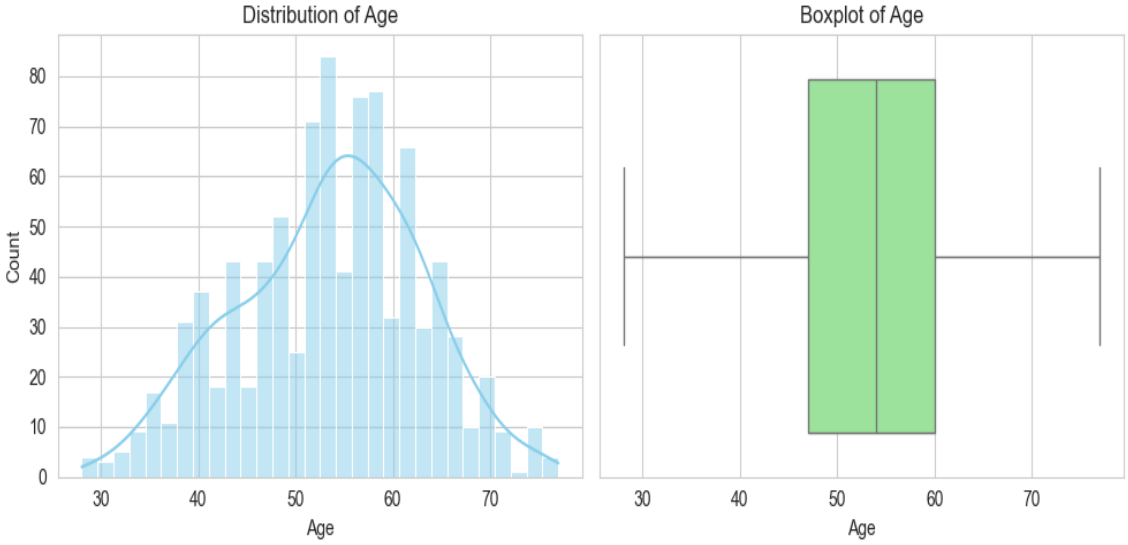
2.2 Data Quality Assessment

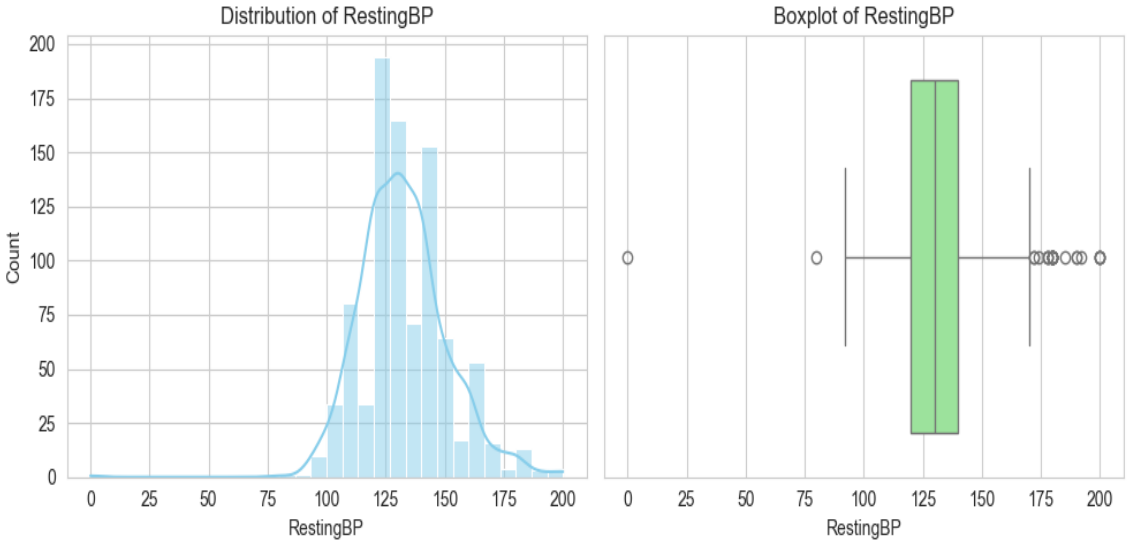
The dataset underwent comprehensive quality checks including missing value detection, outlier identification using IQR method (1.5x threshold), and distribution analysis. All identified issues were documented and addressed in the preprocessing phase.

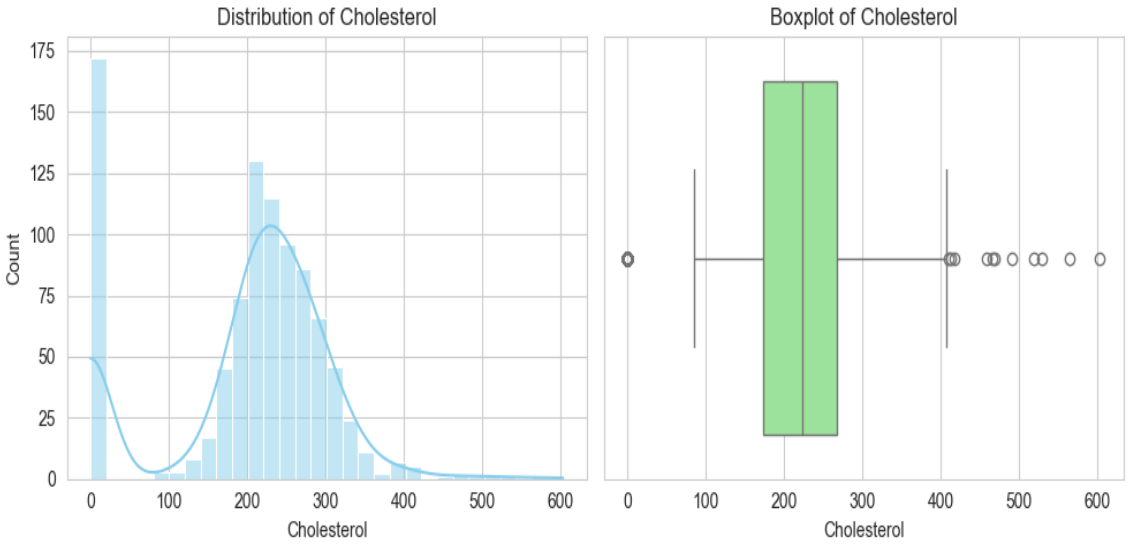
2.3 Key Data Visualizations

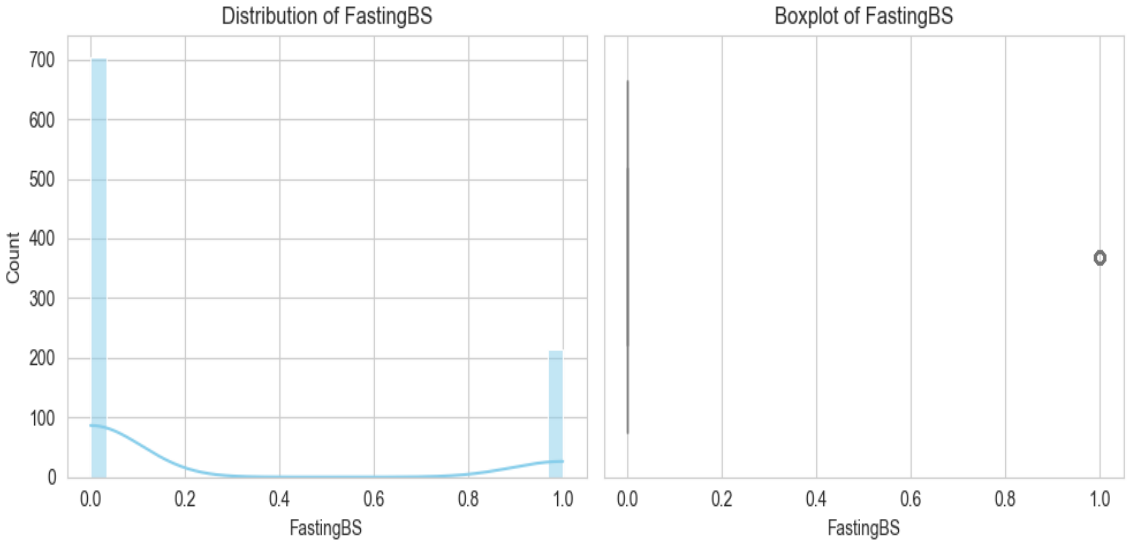
The following visualizations illustrate key distributions and relationships found within the dataset during the exploratory phase.


Correlation Matrix
Age
RestingBP
Cholesterol
FastingBS
MaxHR
Oldpeak
HeartDisease
1.00 0.25 -0.10 0.20 -0.38 0.26 0.28
0.25 1.00 0.10 0.07 -0.11 0.16 0.11
-0.10 0.10 1.00 -0.26 0.24 0.05 -0.23
0.20 0.07 -0.26 1.00 -0.13 0.05 0.27
-0.38 -0.11 0.24 -0.13 1.00 -0.16 -0.40
0.26 0.16 0.05 0.05 -0.16 1.00 0.40
0.28 0.11 -0.23 0.27 -0.40 0.40 1.00
1.0
0.8
0.6
0.4
0.2
0.0
−0.2
−0.4

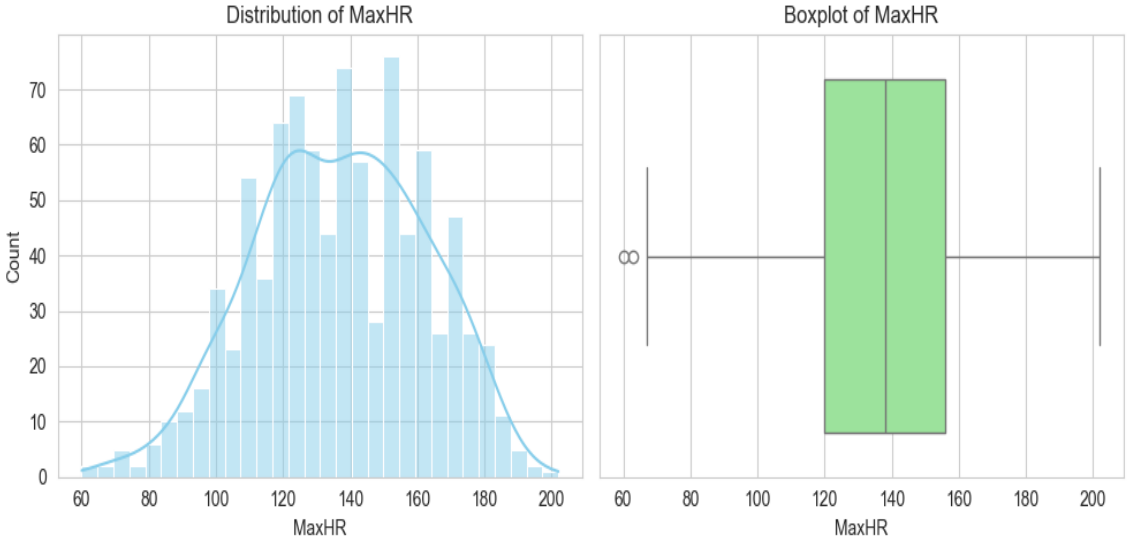

Distribution of Age
Count
Age
Boxplot of Age
Age


Distribution of RestingBP
Count
RestingBP
Boxplot of RestingBP
RestingBP

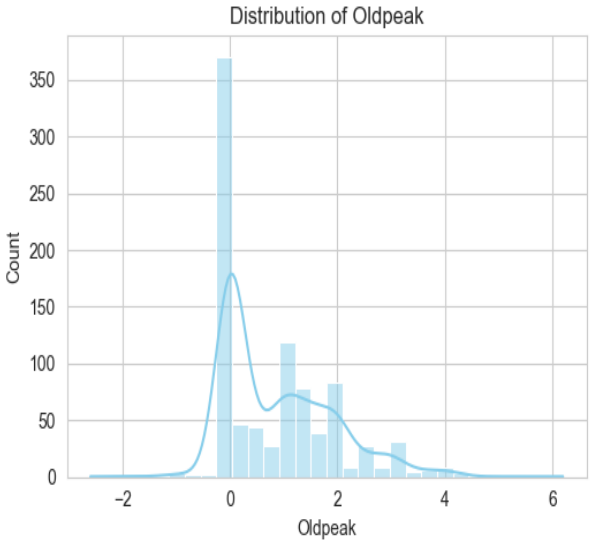

Distribution of Cholesterol
Count
Cholesterol
Boxplot of Cholesterol
Cholesterol

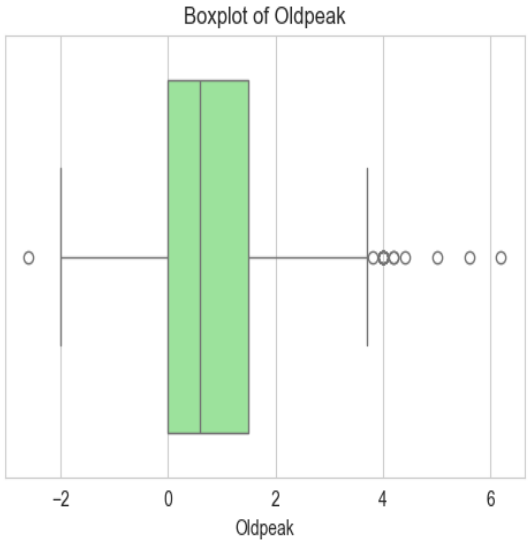

Distribution of FastingBS
Count
FastingBS
0
100
200
300
400
500
600
700
0.0
0.2
0.4
0.6
0.8
1.0

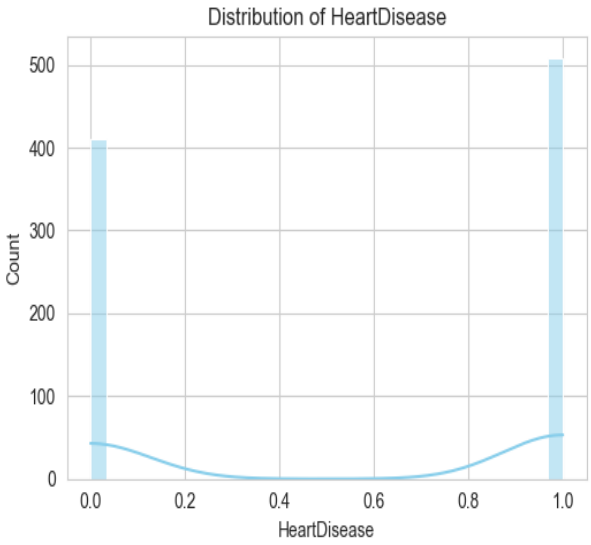
Boxplot of FastingBS
FastingBS
0.0
0.2
0.4
0.6
0.8
1.0

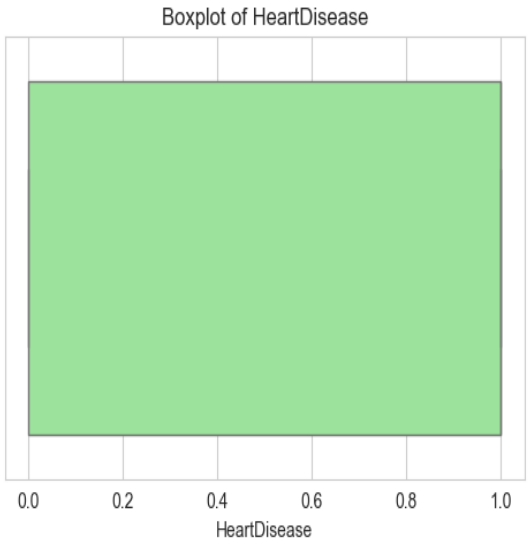

Distribution of MaxHR
Count
MaxHR
0
10
20
30
40
50
60
70
60
80
100
120
140
160
180
200

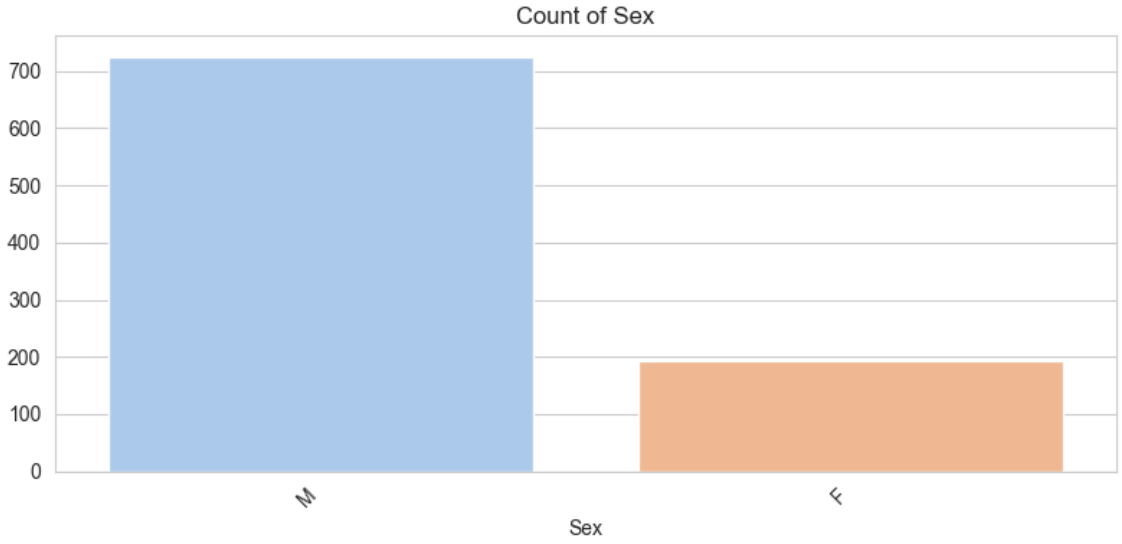
Boxplot of MaxHR
MaxHR
60
80
100
120
140
160
180
200

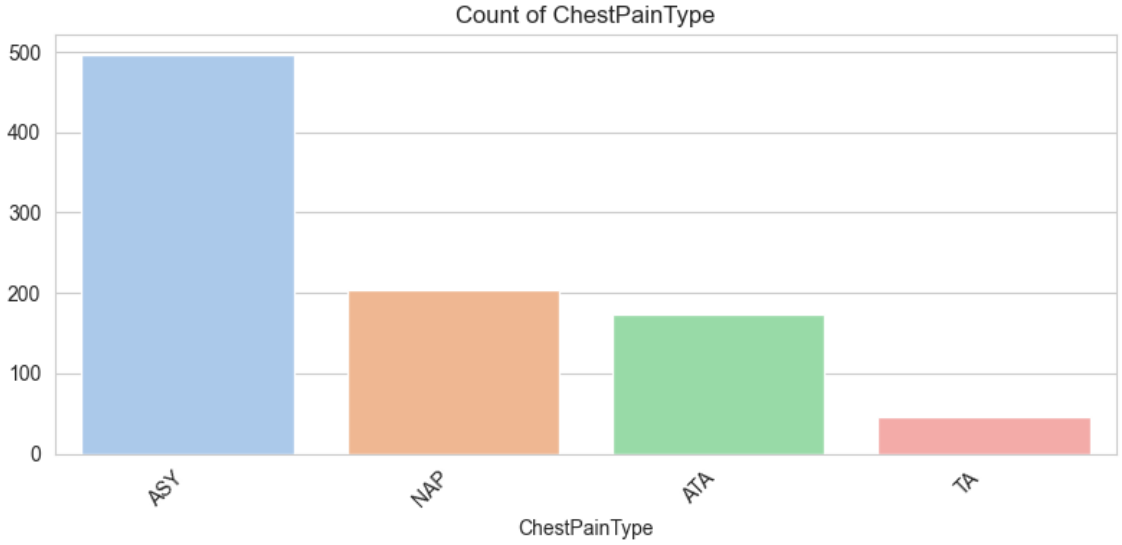

Distribution of Oldpeak
Count
Oldpeak

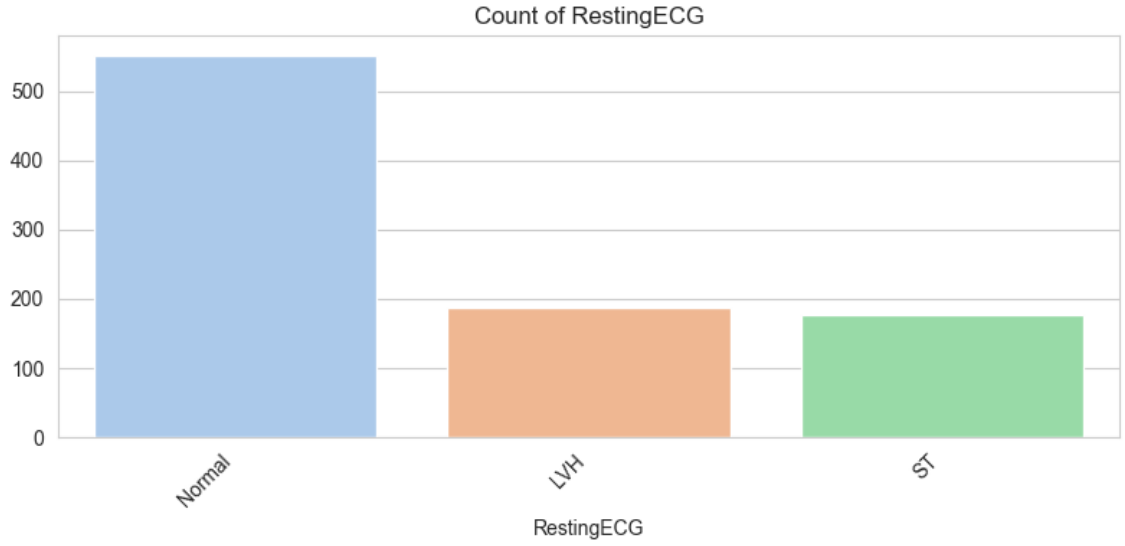

Boxplot of Oldpeak
Oldpeak

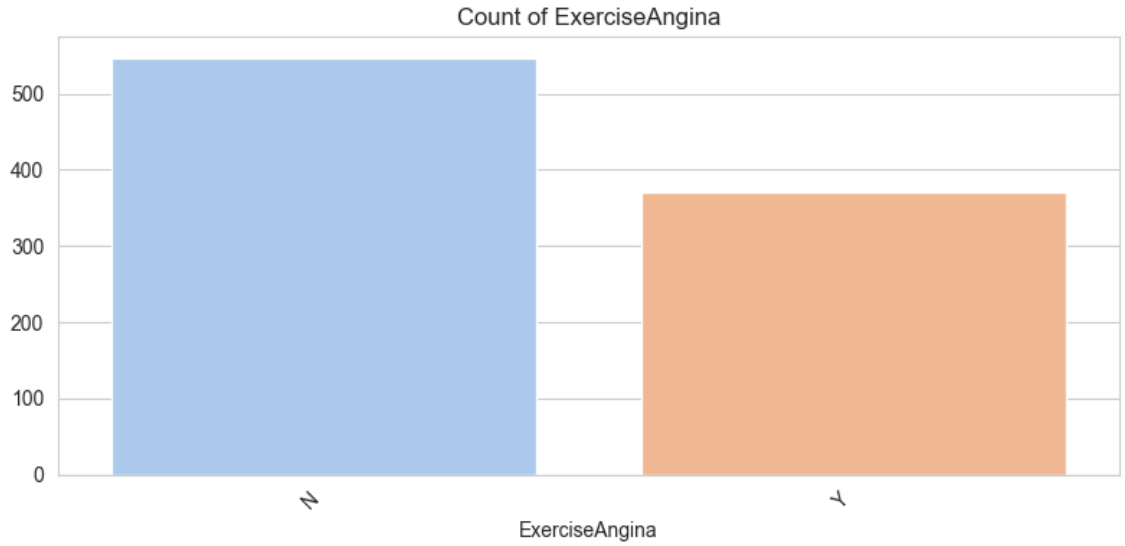

Distribution of HeartDisease
Count
HeartDisease

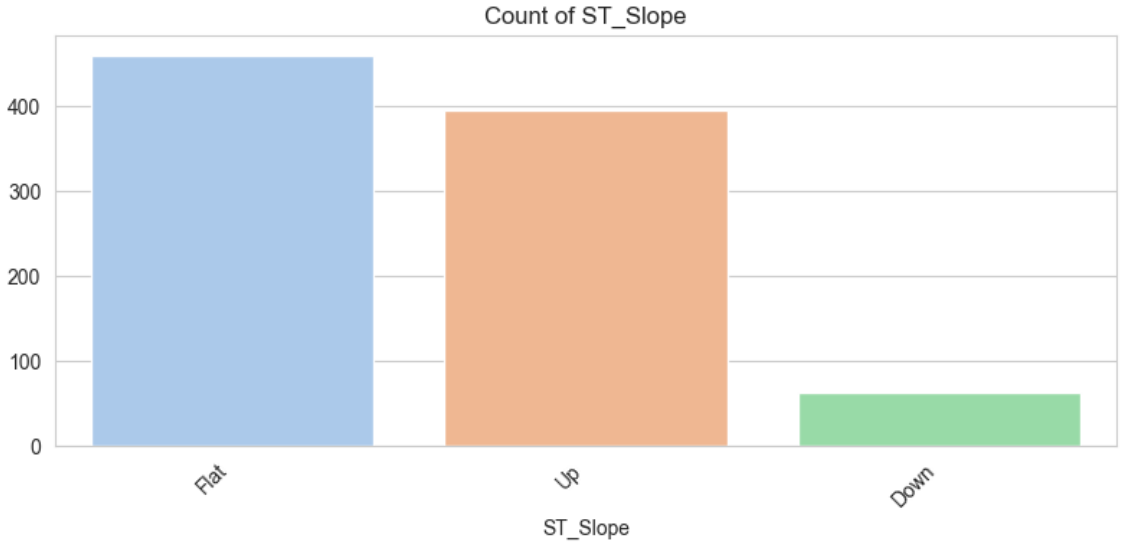

Boxplot of HeartDisease
HeartDisease


Count of Sex
700
600
500
400
300
200
100
0
M
F
Sex


Count of ChestPainType
500
400
300
200
100
0
ASY
NAP
ATA
TA
ChestPainType


Count of RestingECG
500
400
300
200
100
0
Normal
LVH
ST
RestingECG


Count of ExerciseAngina
500
400
300
200
100
0
N
Y
ExerciseAngina


Count of ST_Slope
400
300
200
100
0
Flat
Up
Down
ST_Slope

3. DATA PREPROCESSING & FEATURE ENGINEERING

3.1 Feature Categorization

Numerical Features (6):
Age, RestingBP, Cholesterol, FastingBS, MaxHR, Oldpeak

Categorical Features (5):
Sex, ChestPainType, RestingECG, ExerciseAngina, ST_Slope

3.2 Preprocessing Pipeline

Step	Method	Purpose
1. Missing Values	Median/Mode Imputation	Handle null values
2. Outlier Detection	IQR Method (1.5x)	Identify anomalous data points
3. Scaling	StandardScaler	Normalize numerical features
4. Encoding	One-Hot Encoding	Convert categorical to numerical
5. Feature Selection	Correlation Analysis	Remove redundant features

4. MODEL SELECTION & TRAINING

4.1 Model Architecture

Selected Model: GradientBoostingClassifier

Hyperparameters:

```
{'n_estimators': 200, 'max_depth': 3, 'learning_rate': 0.1}
```

4.2 Training Configuration

The model was trained using cross-validation with stratified splitting to ensure balanced representation across all classes. Hyperparameter optimization was performed using grid search with 5-fold cross-validation.

5. MODEL PERFORMANCE EVALUATION

5.1 Model Leaderboard

Multiple machine learning algorithms were evaluated on the dataset. The following table presents the comparative performance:

Rank	Model	Test Score
1	RandomForest	0.8533
2	GradientBoosting	0.8591
3	LogisticRegression	0.8374

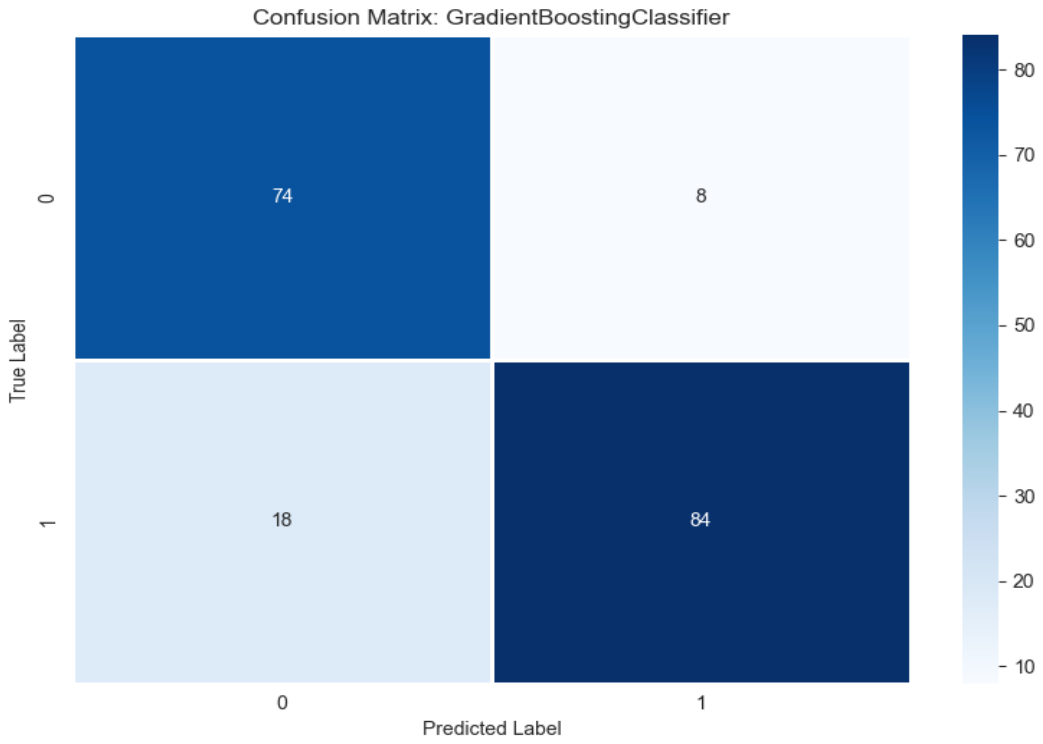
5.2 Performance Insights

The winning model achieved a test score of 0.8591, demonstrating strong performance on the held-out test set. This score indicates the model's ability to generalize to unseen data.

6. VISUAL ANALYSIS

6.1 Confusion Matrix

The confusion matrix visualizes the model's prediction accuracy across different classes.



6.2 Model Comparison

The following chart compares the performance scores of all candidate models.

Model Comparison (Test Score)
Score
1.0
0.8
0.6
0.4
0.2
0.0
RandomForest
GradientBoosting
LogisticRegression

7. ERROR ANALYSIS & INSIGHTS

7.1 Detailed Analysis

EXECUTIVE SUMMARY:

The overall performance of the classification model is satisfactory, achieving an accuracy of 85.87%. However, a closer examination reveals that the primary issue lies with the Class '0', which shows a slightly lower F1 score compared to other classes. This suggests that there might be some imbalances in the data distribution or in the model's performance on this particular class.

Diagnostics Analysis:

The diagnostic analysis reveals that the issue primarily affects Class '0', where the precision and recall are relatively lower (0.80 and 0.90 respectively) compared to other classes (where precision is higher, i.e., 0.91 for Class '1'). This suggests a potential bias in the model's predictions towards Class '1' when it comes to Class '0'. The F1 score of 0.85 indicates that the model performs reasonably well on this class, but there might be room for improvement.

Given the nature of the issue, it is not clear whether the problem lies with precision or recall. However, since the F1 score (a harmonic mean of precision and recall) falls short of the individual precision and recall scores, it suggests that either precision or recall is underperforming on this class. A more detailed investigation would be necessary to determine which aspect is at fault.

RECOMMENDATIONS:

1. Data augmentation techniques should be applied to increase the number of instances in Class '0', potentially alleviating the imbalance issue. 2. An examination into the bias-variance trade-off can help identify if the model's performance on this class is suffering from overfitting or underfitting, which would require adjustments to the model architecture or hyperparameters.

3. Investigating the feature engineering and data preprocessing steps to see if any modifications can improve the predictive performance on Class '0' without negatively impacting other classes would be a good next step.

8. RECOMMENDATIONS & NEXT STEPS

8.1 Model Deployment Recommendations

Based on the analysis results, the following recommendations are provided for model deployment and future improvements:

- Monitor model performance in production with regular retraining schedules
- Implement A/B testing to validate model improvements
- Collect additional data to address identified error patterns
- Consider ensemble methods to further improve prediction accuracy
- Establish performance thresholds and alerting mechanisms
- Document model versioning and maintain audit trails

8.2 Future Work

Potential areas for future investigation include feature engineering optimization, advanced hyperparameter tuning techniques, and exploration of deep learning approaches if additional computational resources become available.