

Data Analysis Report

Dataset Characteristics

Number of Samples: 105

Number of Features: 5

Feature Density: 0.05

Class Balance: 0.84

Model Recommendations

* Small dataset detected. Consider using simpler models to avoid overfitting.

Data Summary

count: {'Id': 150.0, 'SepalLengthCm': 150.0, 'SepalWidthCm': 150.0, 'PetalLengthCm': 150.0, 'PetalWidthCm': 150.0, 'Species': 150}

unique: {'Id': nan, 'SepalLengthCm': nan, 'SepalWidthCm': nan, 'PetalLengthCm': nan, 'PetalWidthCm': nan, 'Species': 3}

top: {'Id': nan, 'SepalLengthCm': nan, 'SepalWidthCm': nan, 'PetalLengthCm': nan, 'PetalWidthCm': nan, 'Species': 'Iris-setosa'}

freq: {'Id': nan, 'SepalLengthCm': nan, 'SepalWidthCm': nan, 'PetalLengthCm': nan, 'PetalWidthCm': nan, 'Species': 50}

mean: {'Id': 75.5, 'SepalLengthCm': 5.843333333333334, 'SepalWidthCm': 3.0540000000000003, 'PetalLengthCm': 3.7586666666666666, 'PetalWidthCm': 1.1986666666666668, 'Species': nan}

std: {'Id': 43.445367992456916, 'SepalLengthCm': 0.828066127977863, 'SepalWidthCm': 0.4335943113621737, 'PetalLengthCm': 1.7644204199522626, 'PetalWidthCm': 0.7631607417008411, 'Species': nan}

min: {'Id': 1.0, 'SepalLengthCm': 4.3, 'SepalWidthCm': 2.0, 'PetalLengthCm': 1.0, 'PetalWidthCm': 0.1, 'Species': nan}

25%: {'Id': 38.25, 'SepalLengthCm': 5.1, 'SepalWidthCm': 2.8, 'PetalLengthCm': 1.6, 'PetalWidthCm':

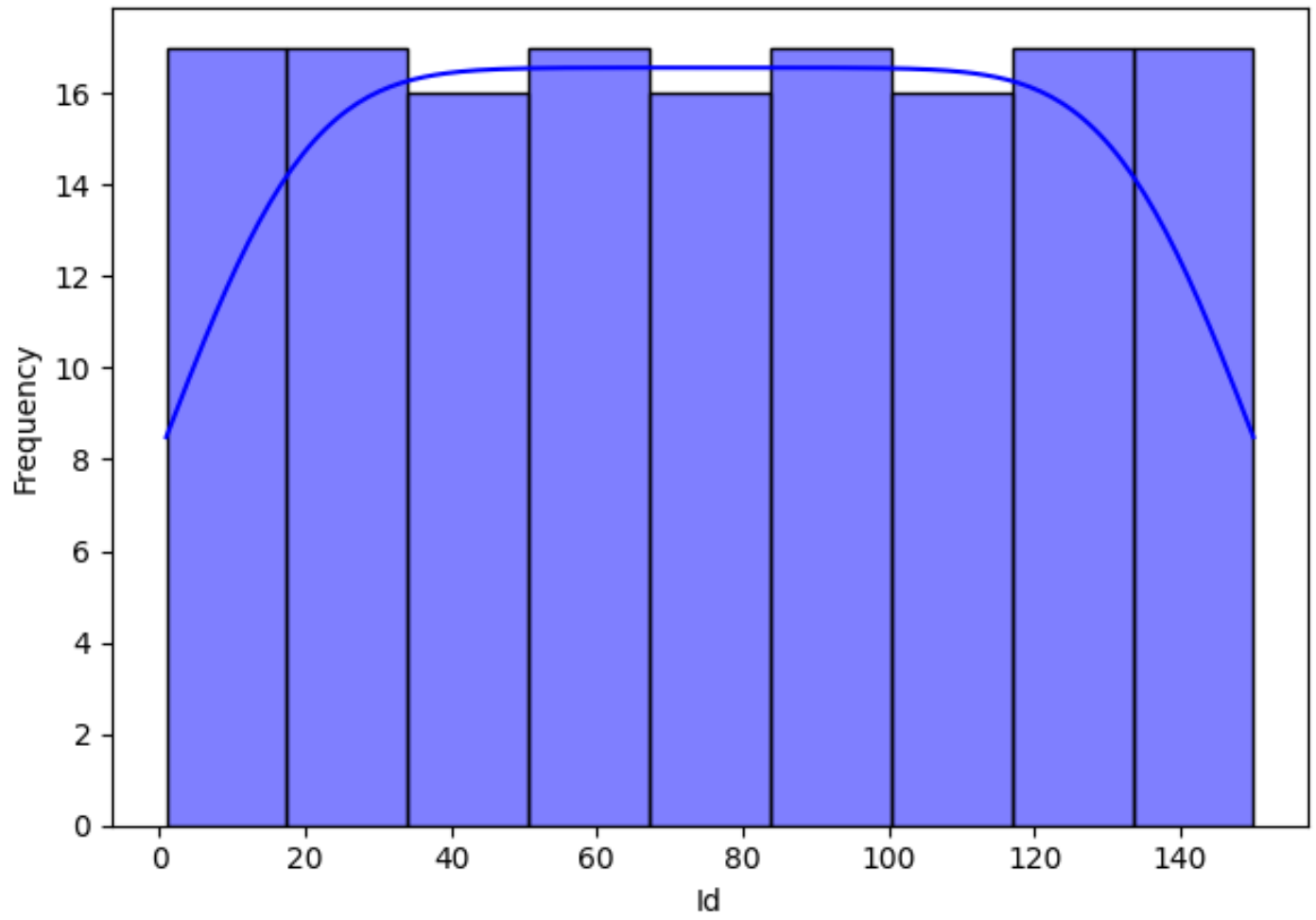
0.3, 'Species': nan}

50%: {'Id': 75.5, 'SepalLengthCm': 5.8, 'SepalWidthCm': 3.0, 'PetalLengthCm': 4.35, 'PetalWidthCm': 1.3, 'Species': nan}

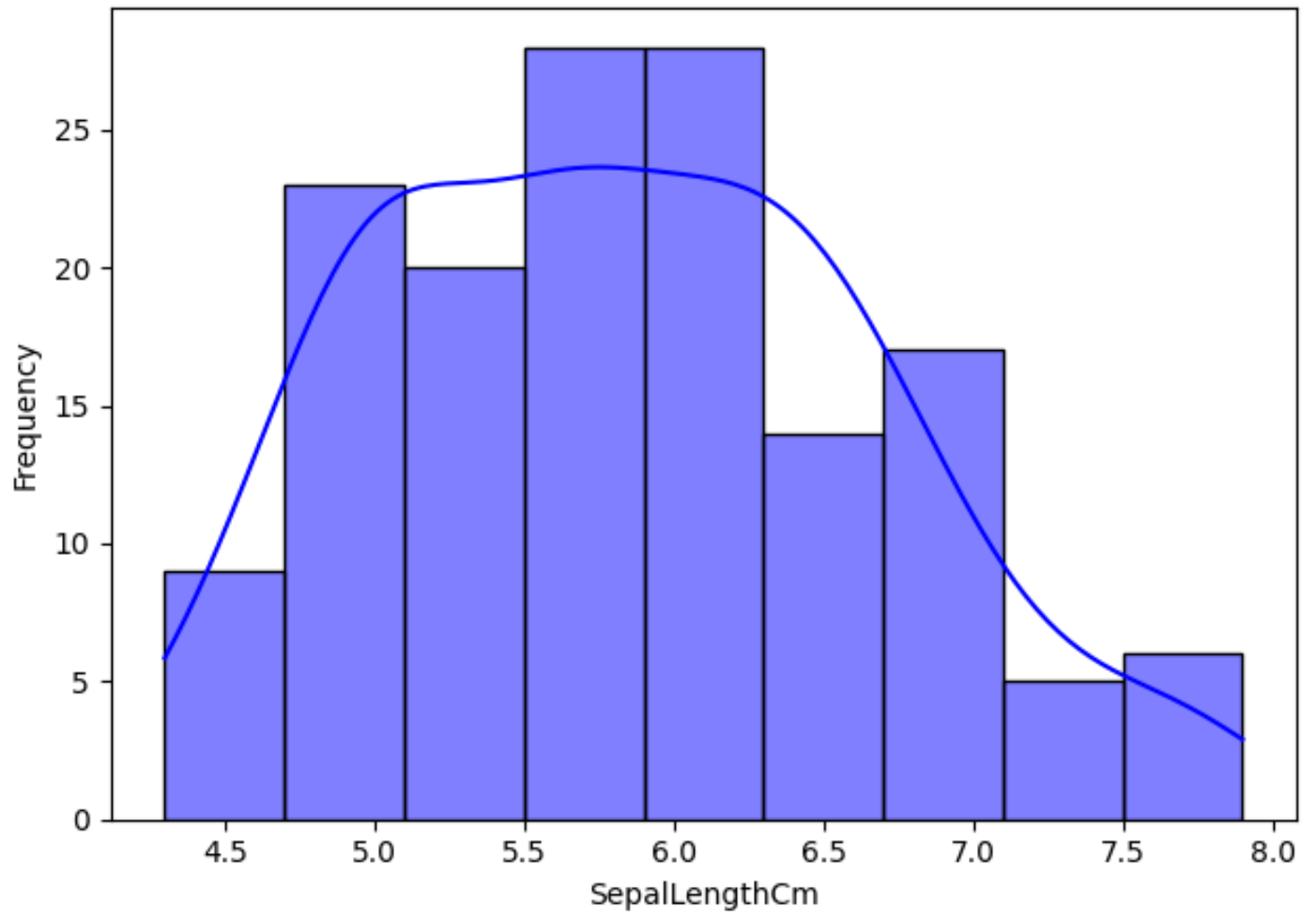
75%: {'Id': 112.75, 'SepalLengthCm': 6.4, 'SepalWidthCm': 3.3, 'PetalLengthCm': 5.1, 'PetalWidthCm': 1.8, 'Species': nan}

max: {'Id': 150.0, 'SepalLengthCm': 7.9, 'SepalWidthCm': 4.4, 'PetalLengthCm': 6.9, 'PetalWidthCm': 2.5, 'Species': nan}

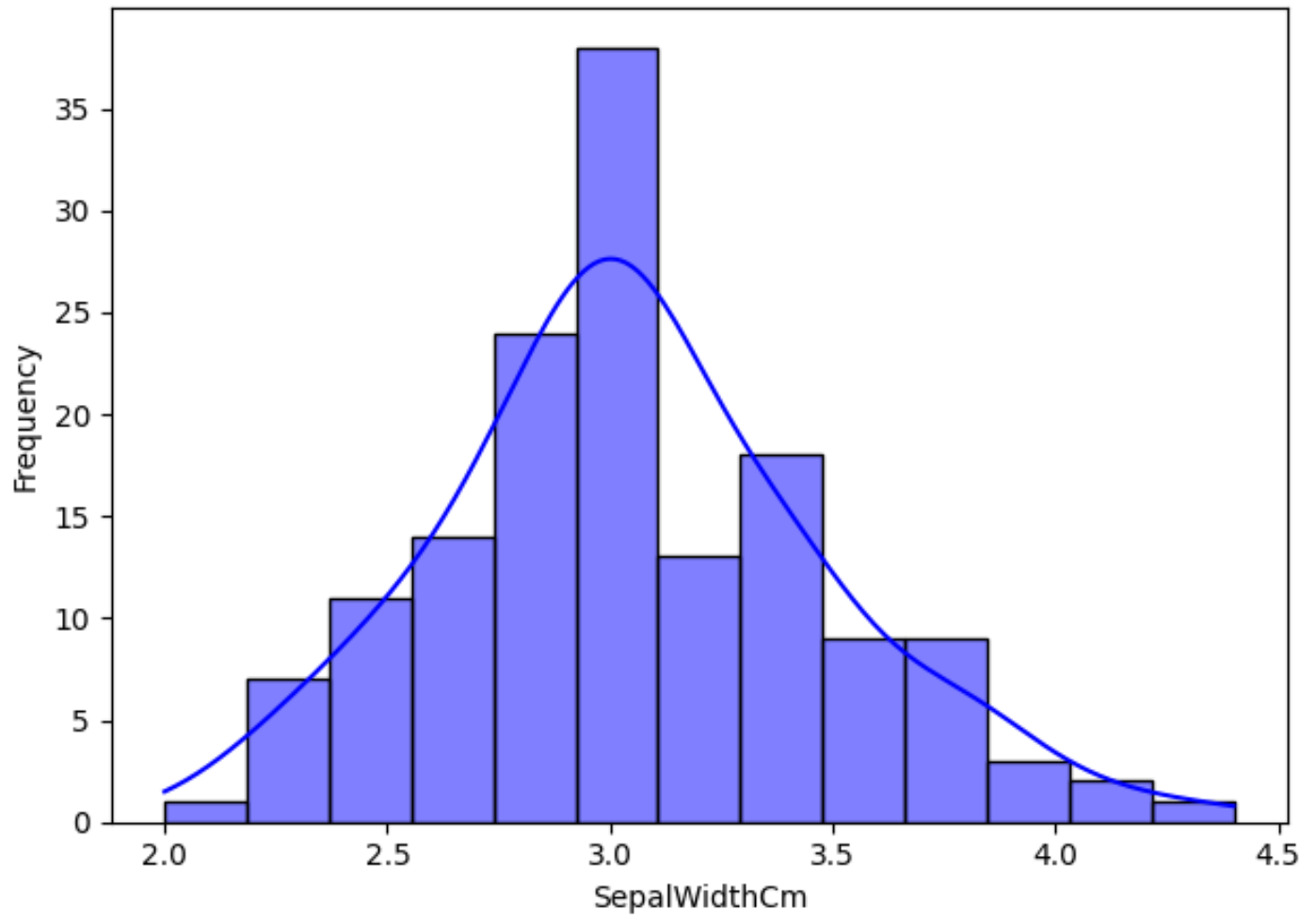
Histogram of Id



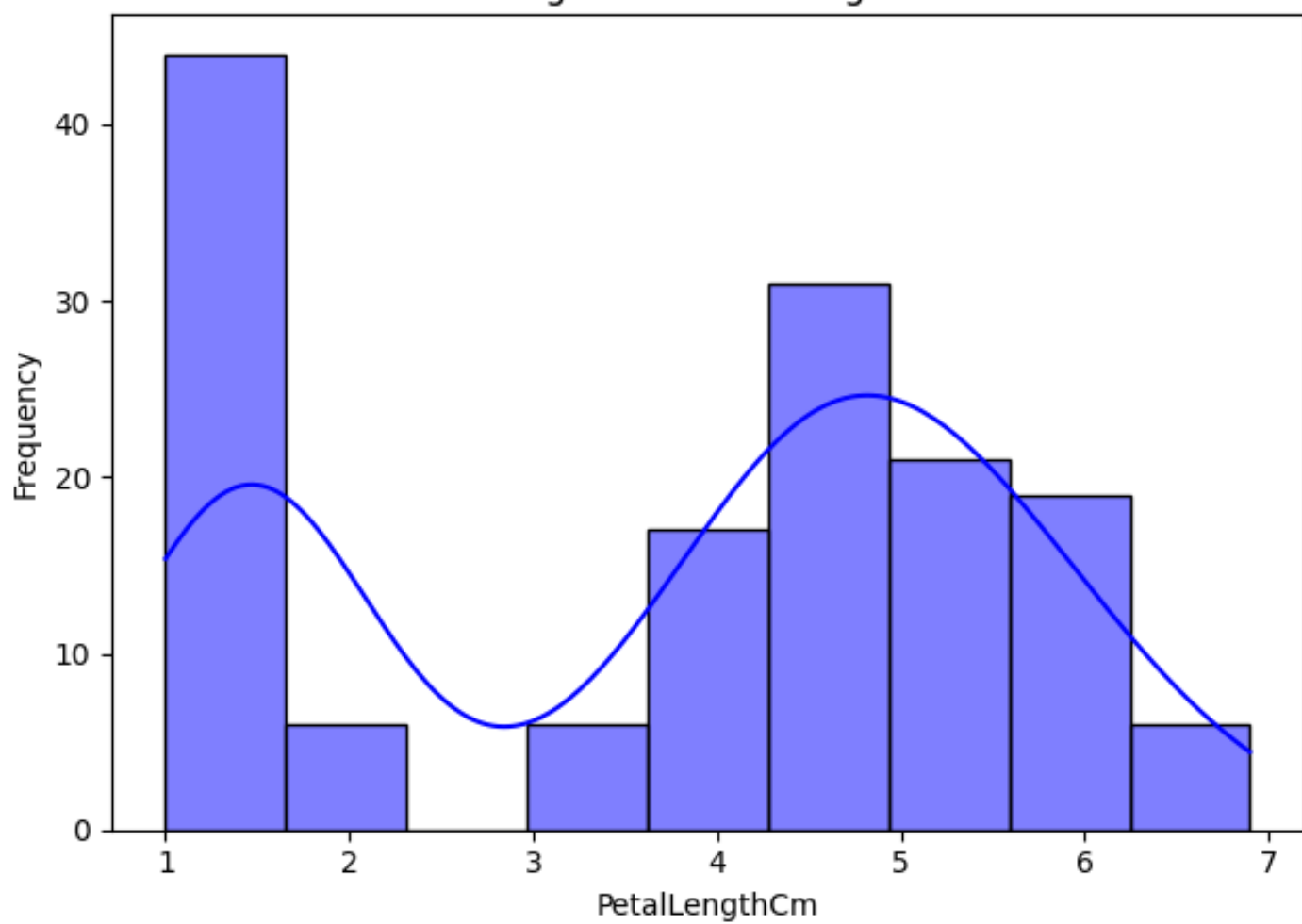
Histogram of SepalLengthCm



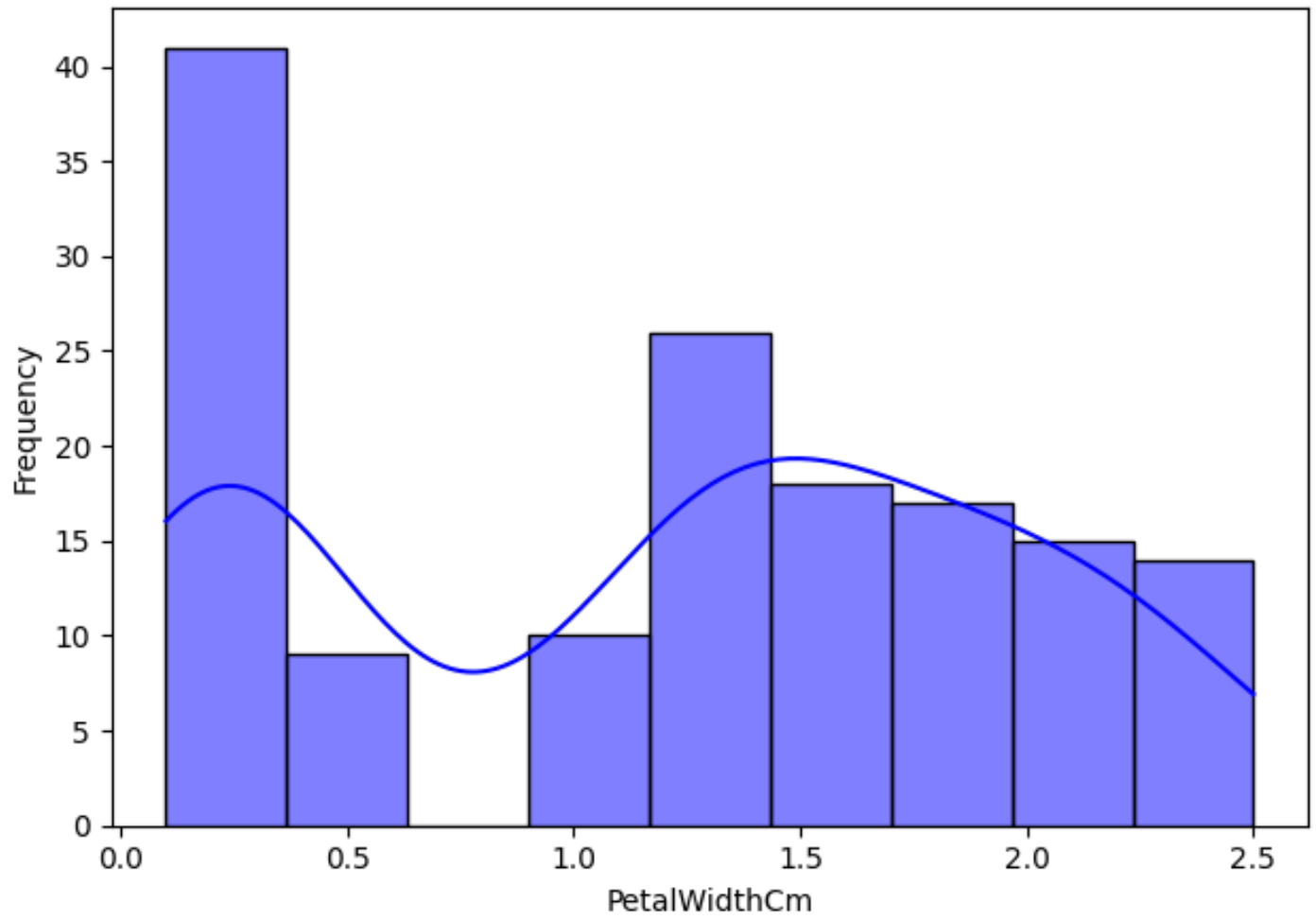
Histogram of SepalWidthCm



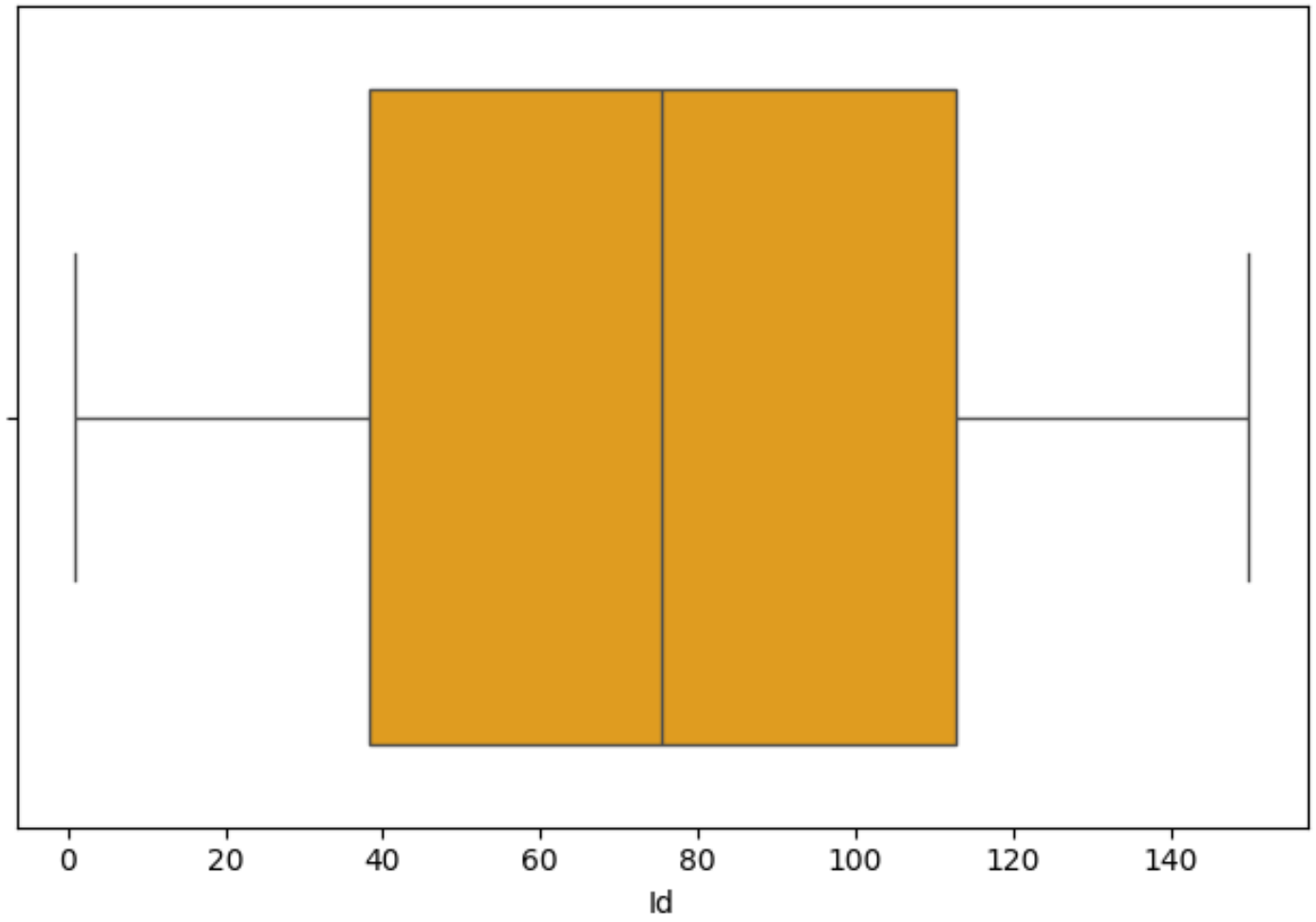
Histogram of PetalLengthCm



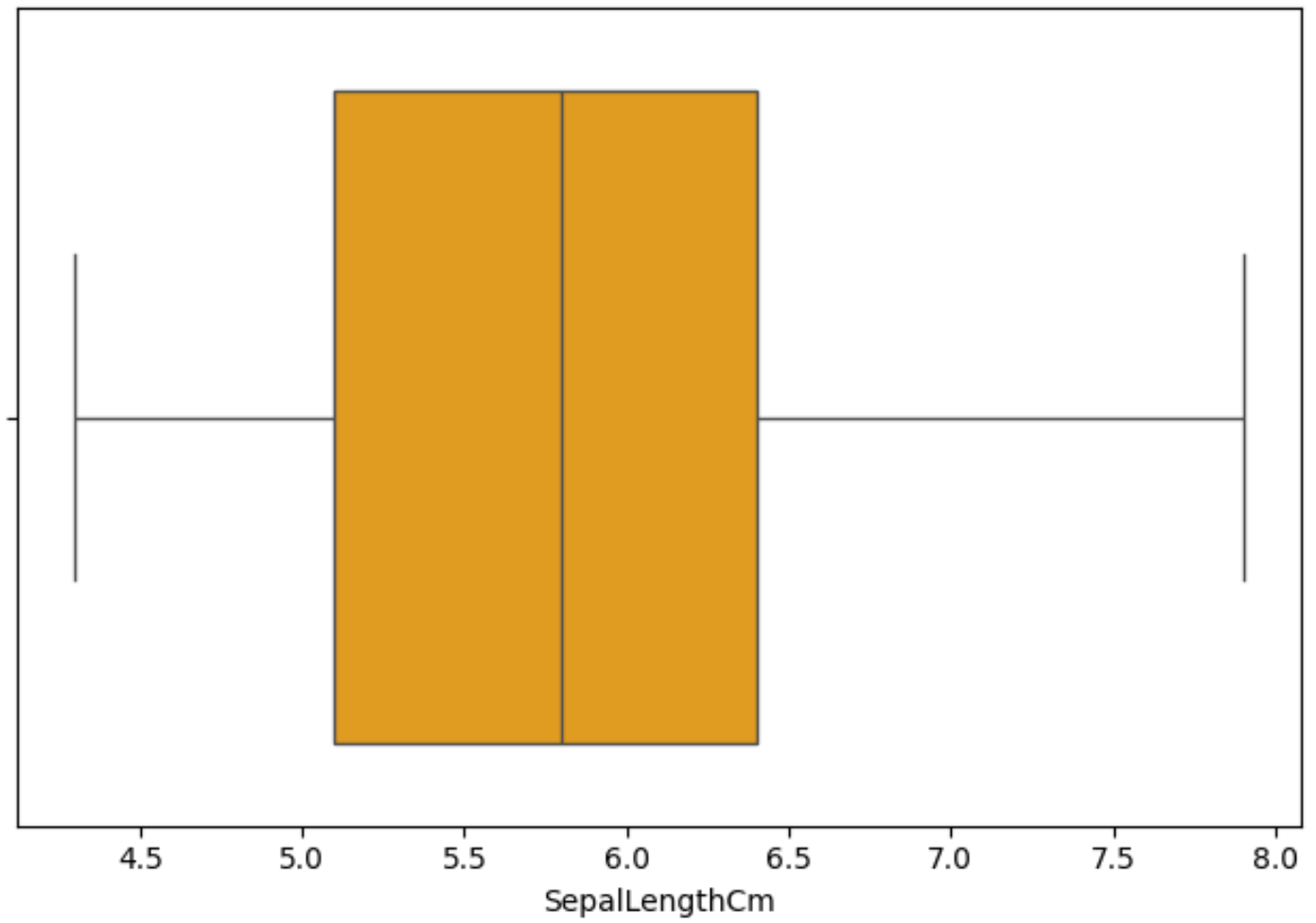
Histogram of PetalWidthCm



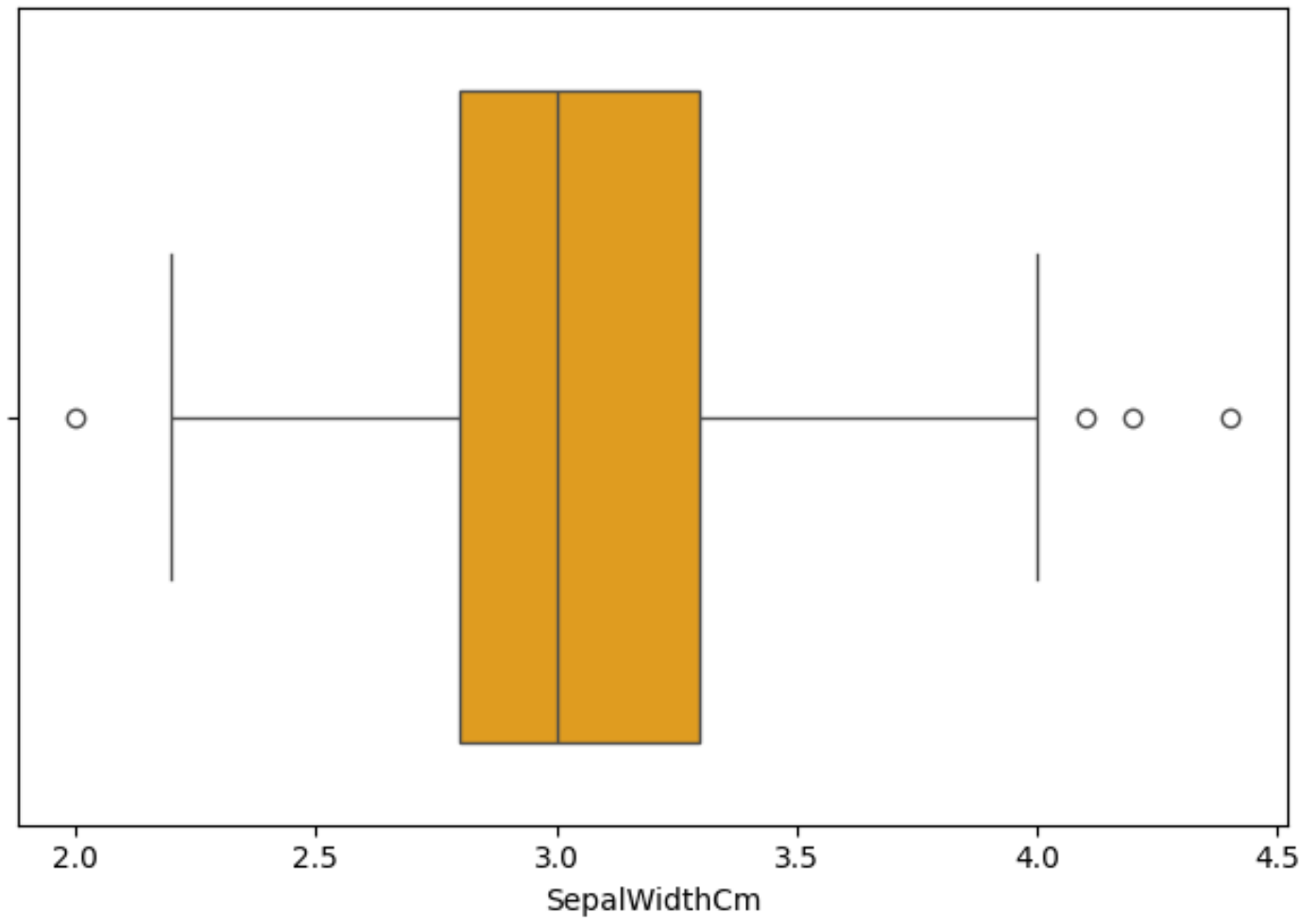
Boxplot of Id



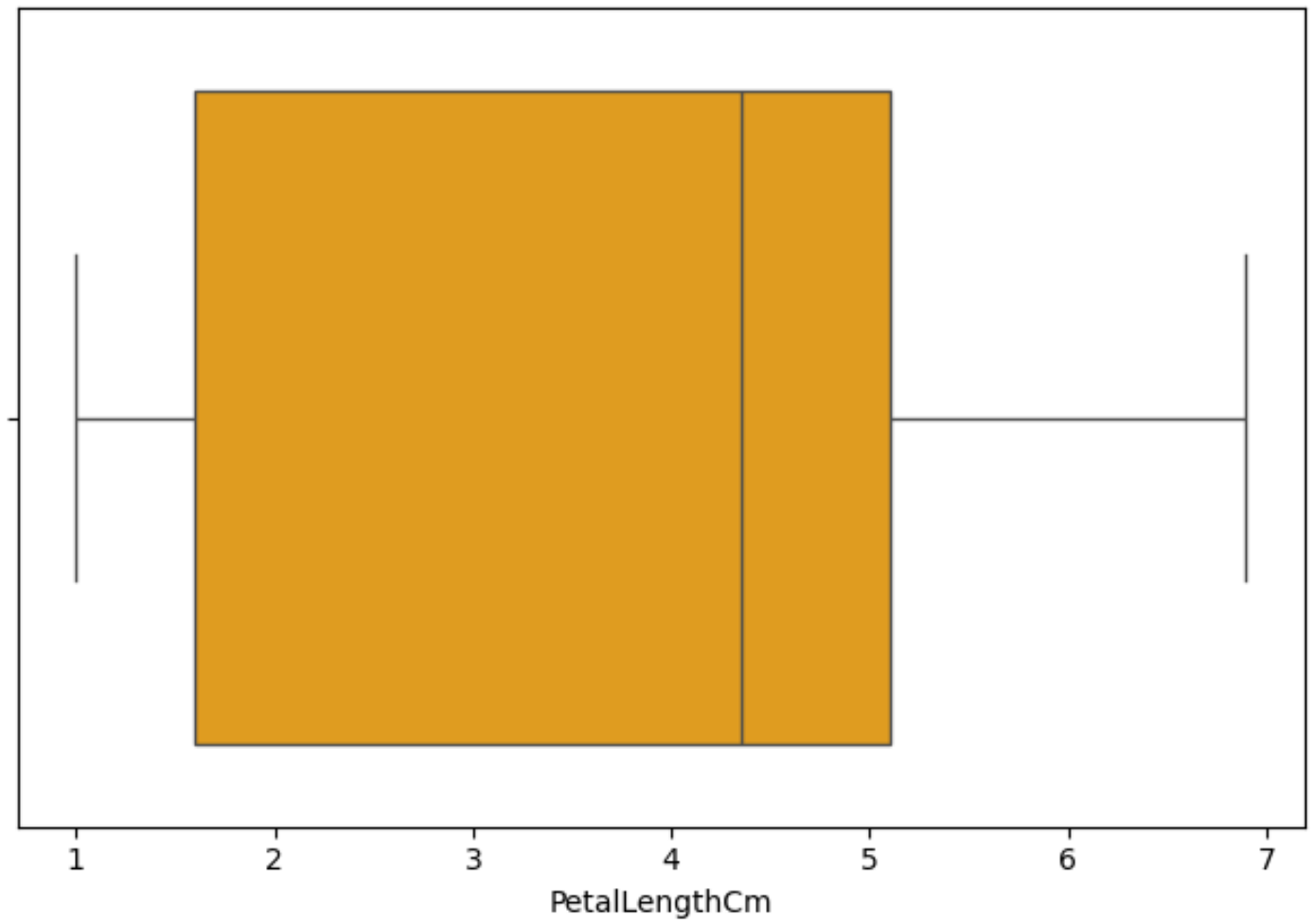
Boxplot of SepalLengthCm



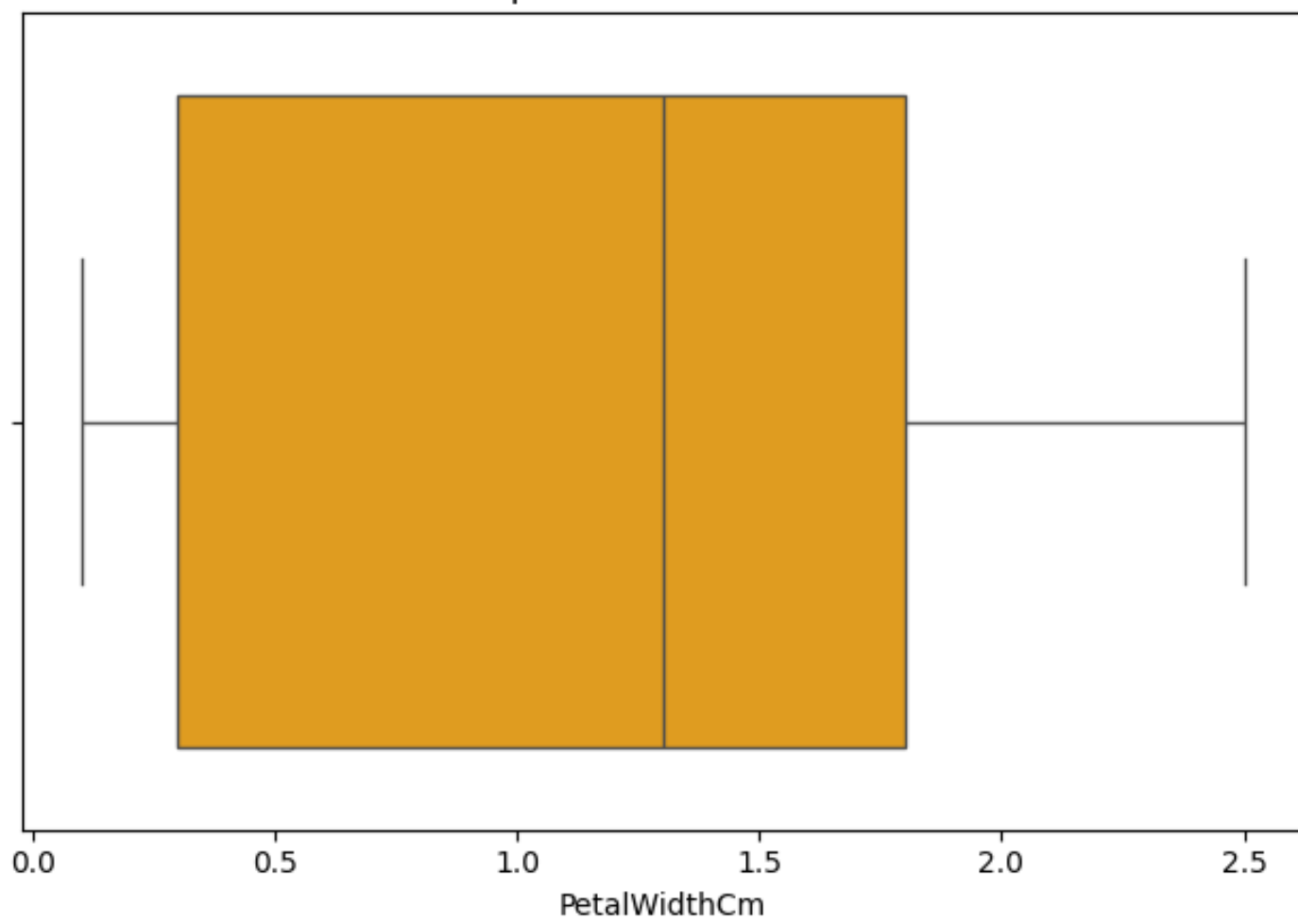
Boxplot of SepalWidthCm



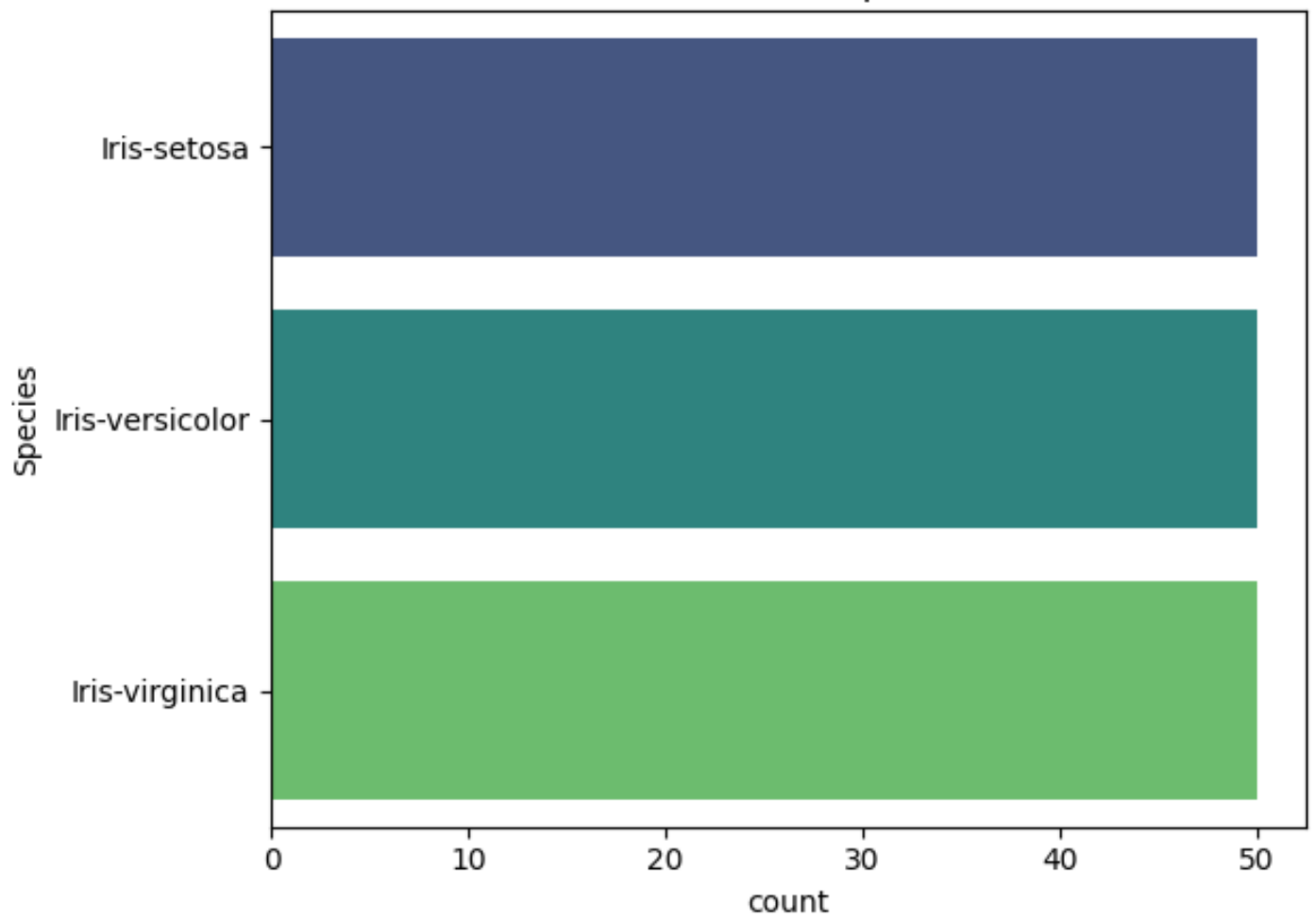
Boxplot of PetalLengthCm



Boxplot of PetalWidthCm



Count Plot of Species



Model Results

Model: Random Forest

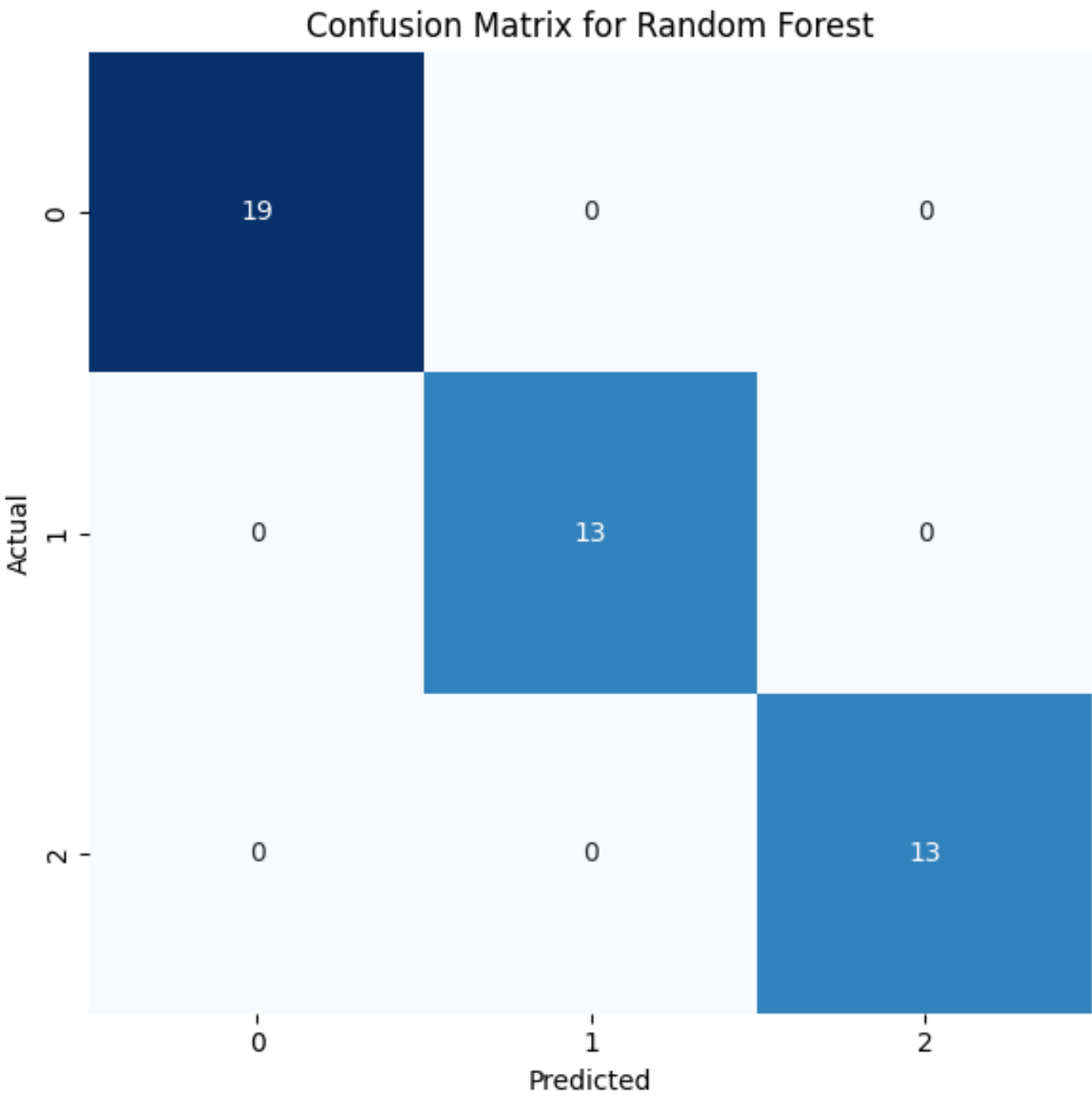
accuracy: 1.0

Model: Logistic Regression

accuracy: 1.0

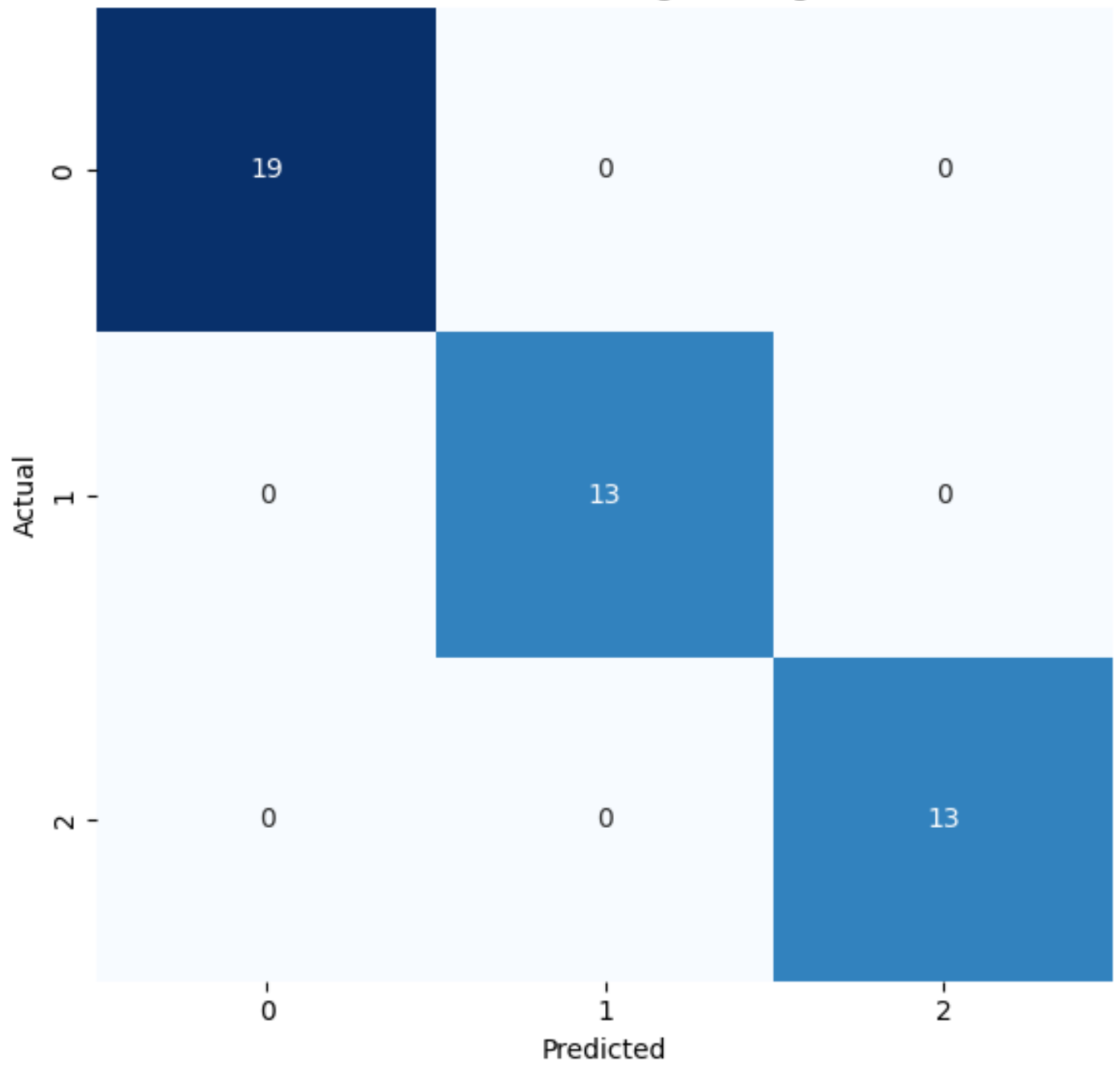
Model: K-Nearest Neighbors

accuracy: 1.0



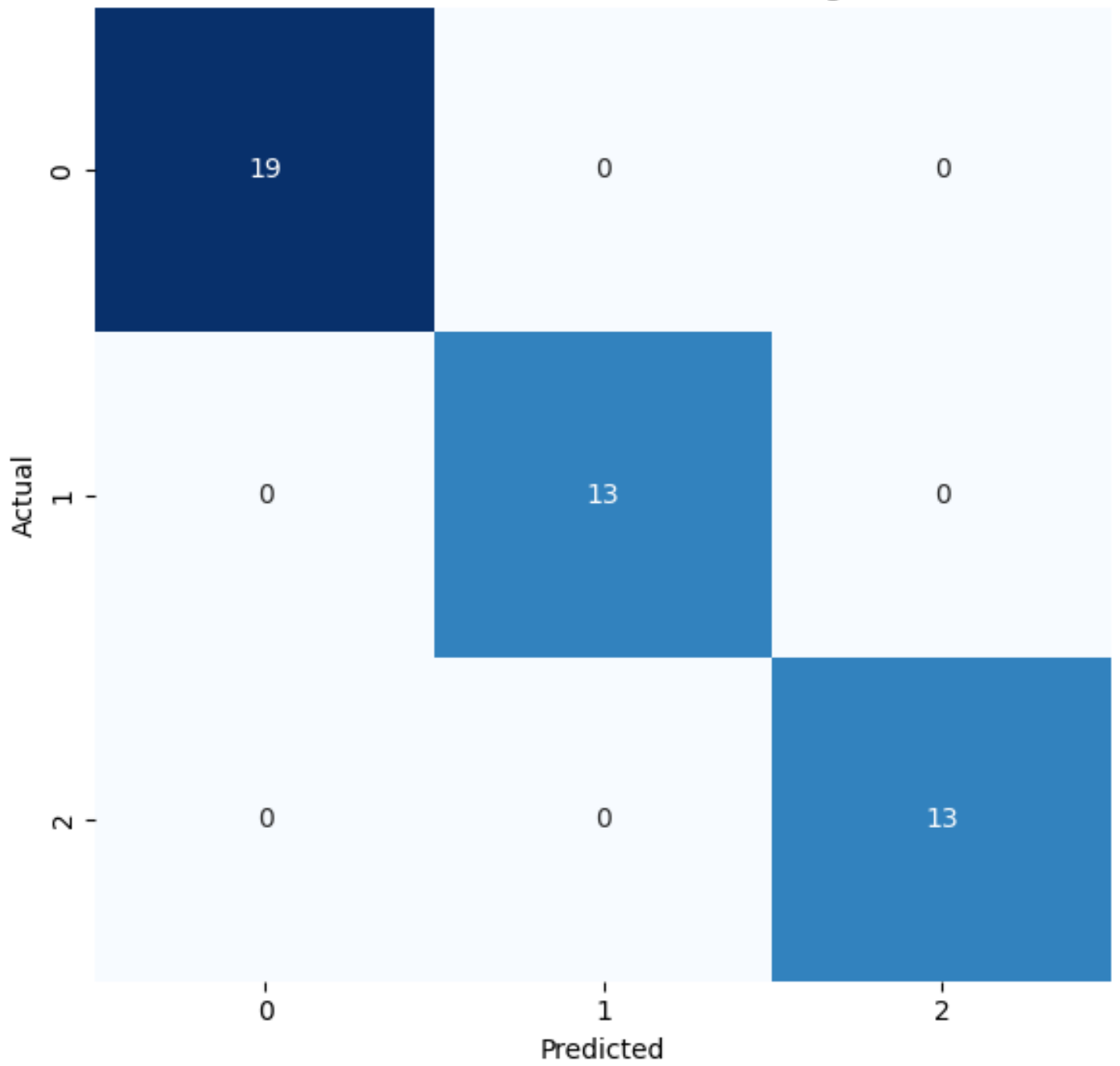
Confusion Matrix - Logistic Regression

Confusion Matrix for Logistic Regression



Confusion Matrix - K-Nearest Neighbors

Confusion Matrix for K-Nearest Neighbors



Model Comparison

