**01\_data\_preprocessing\_results**

Dataset Shape: (303, 14)

Missing Values per Column:

age 0

sex 0

cp 0

trestbps 0

chol 0

fbs 0

restecg 0

thalach 0

exang 0

oldpeak 0

slope 0

ca 4

thal 2

num 0

dtype: int64

Missing Values After Cleaning:

age 0

sex 0

cp 0

trestbps 0

chol 0

fbs 0

restecg 0

thalach 0

exang 0

oldpeak 0

slope 0

ca 0

thal 0

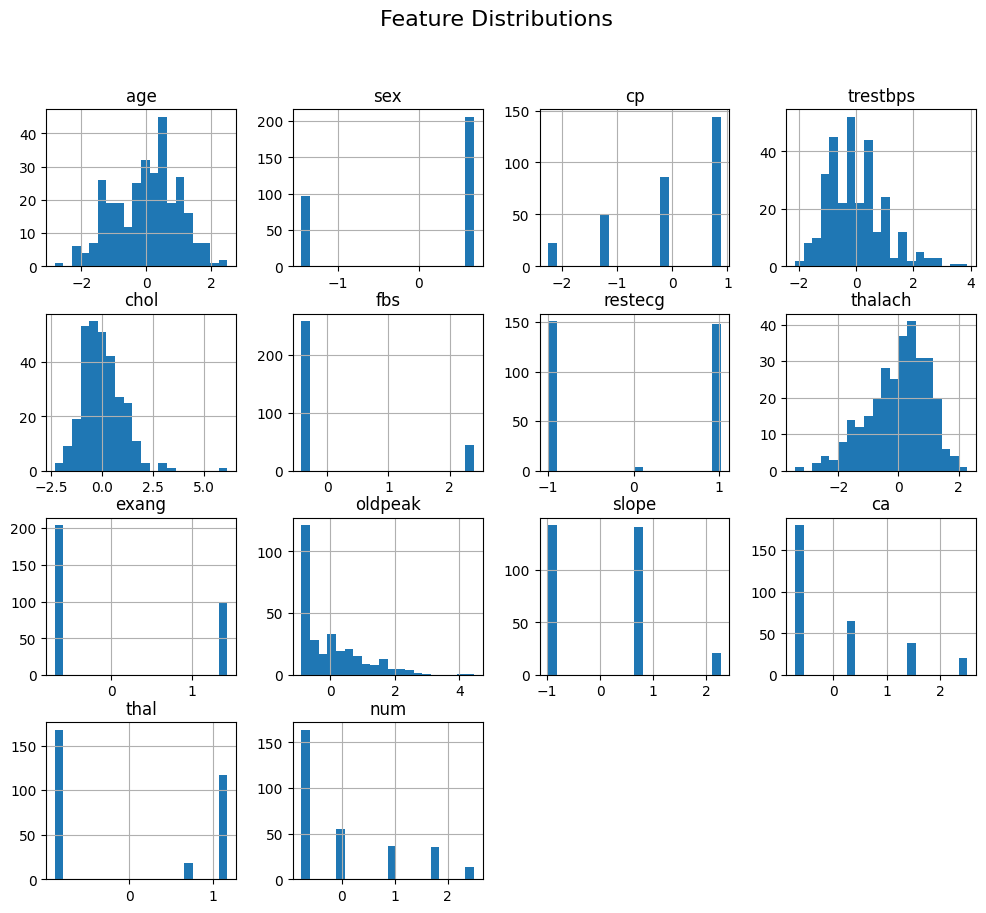
num 0

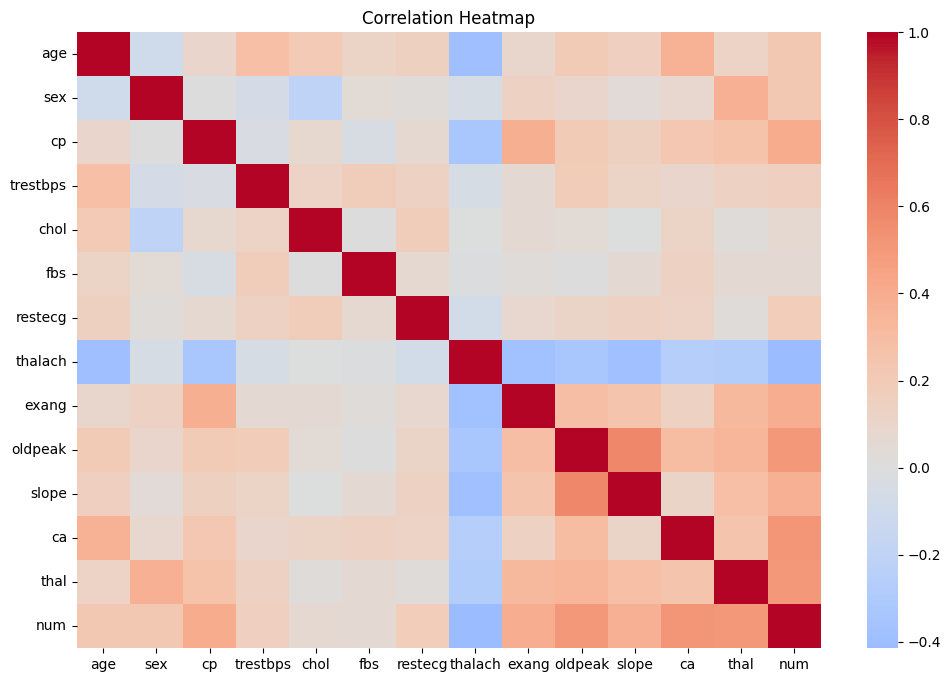
dtype: int64

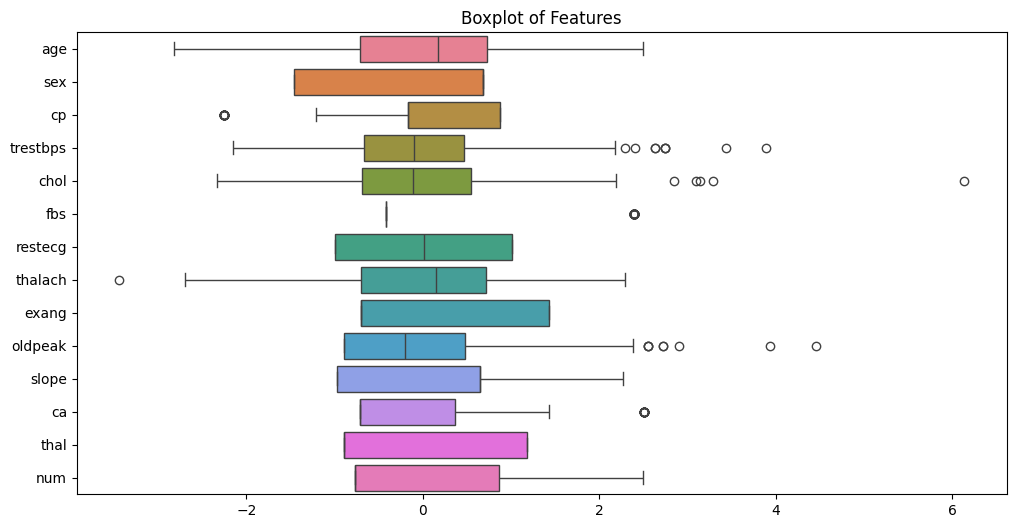
Index(['age', 'sex', 'cp', 'trestbps', 'chol', 'fbs', 'restecg', 'thalach',

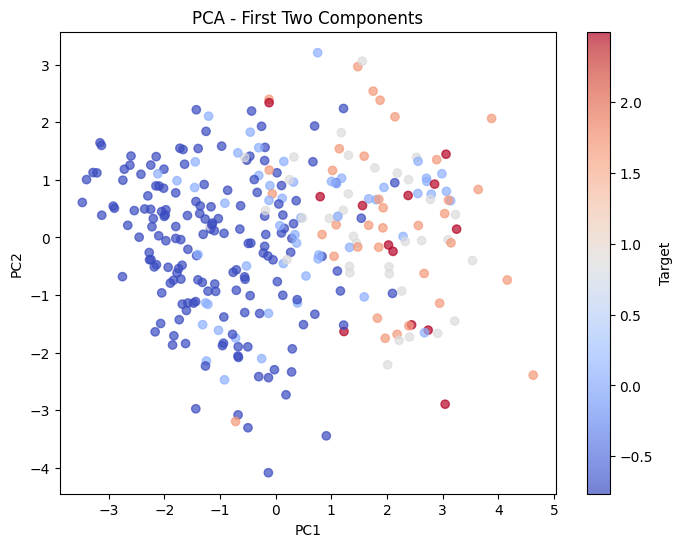
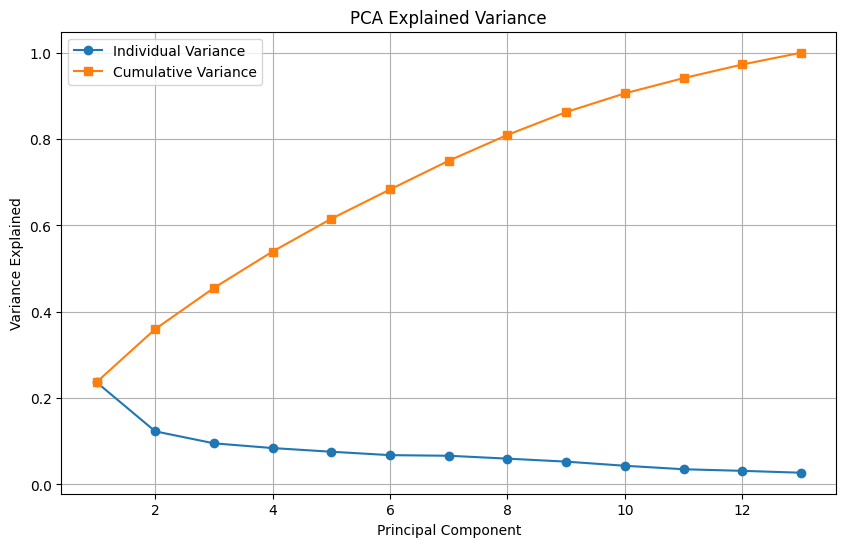
'exang', 'oldpeak', 'slope', 'ca', 'thal', 'num'],

dtype='object')

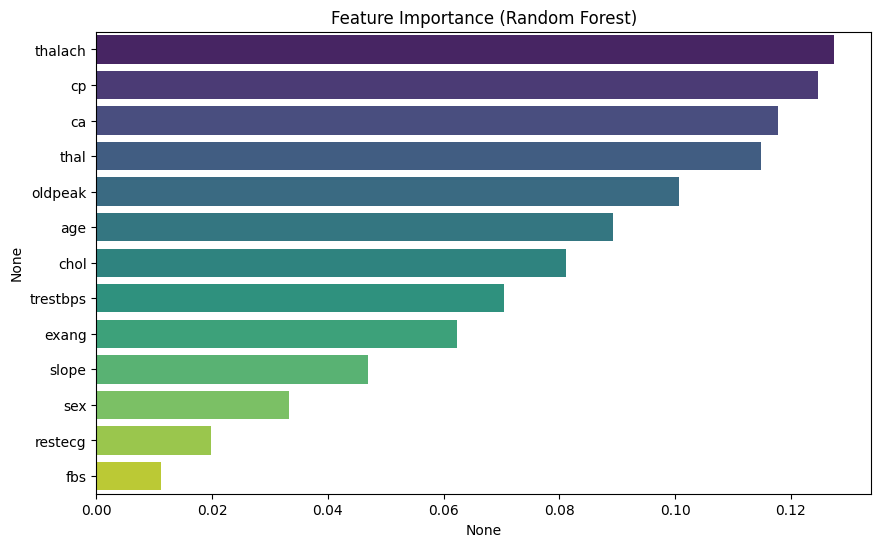


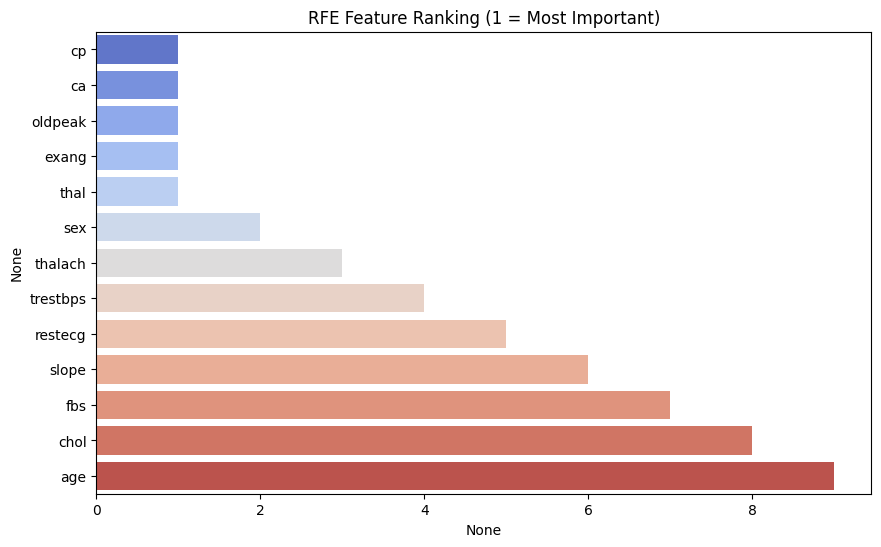




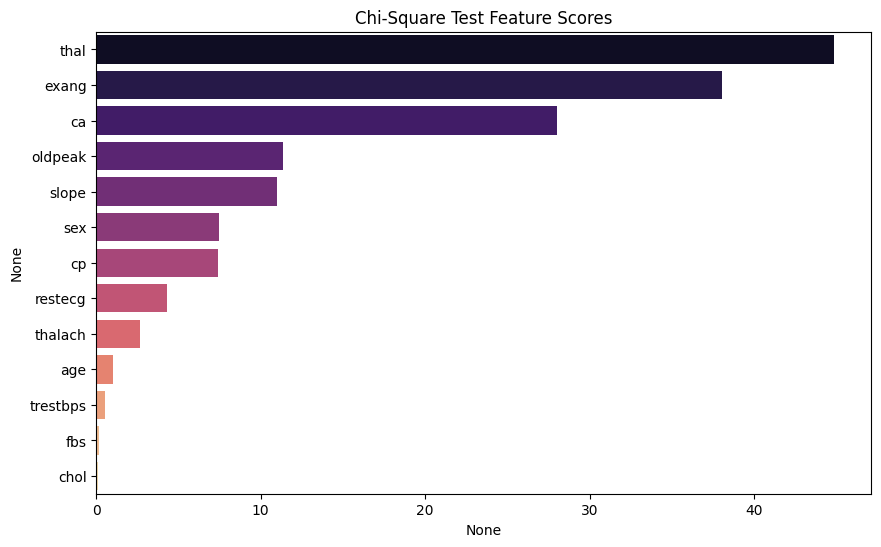
**02\_pca\_analysis\_results**

**03\_feature\_selection\_results**

****

****

Top 5 Features from RFE: ['cp', 'exang', 'oldpeak', 'ca', 'thal']



top Features from Chi-Square Test: ['age', 'sex', 'cp', 'trestbps', 'chol', 'fbs', 'restecg', 'thalach', 'exang', 'oldpeak', 'slope', 'ca', 'thal']

Final Selected Features for Modeling: ['thal', 'exang', 'thalach', 'slope', 'oldpeak', 'ca', 'cp']

**04\_supervised\_learning\_results**

Dataset shape: (303, 8)

Columns: ['thalach', 'oldpeak', 'slope', 'thal', 'ca', 'exang', 'cp', 'num']

Train/Test sizes: (242, 7) (61, 7)

=== LogisticRegression ===

Accuracy: 0.8689 Precision: 0.8125 Recall: 0.9286 F1: 0.8667 ROC\_AUC: 0.9404761904761905

Classification report:

precision recall f1-score support

0 0.93 0.82 0.87 33

1 0.81 0.93 0.87 28

accuracy 0.87 61

macro avg 0.87 0.87 0.87 61

weighted avg 0.88 0.87 0.87 61

=== DecisionTree ===

Accuracy: 0.7705 Precision: 0.6944 Recall: 0.8929 F1: 0.7812 ROC\_AUC: 0.7797619047619049

Classification report:

precision recall f1-score support

0 0.88 0.67 0.76 33

1 0.69 0.89 0.78 28

accuracy 0.77 61

macro avg 0.79 0.78 0.77 61

weighted avg 0.79 0.77 0.77 61

=== RandomForest ===

Accuracy: 0.8525 Precision: 0.8276 Recall: 0.8571 F1: 0.8421 ROC\_AUC: 0.9393939393939394

Classification report:

precision recall f1-score support

0 0.88 0.85 0.86 33

1 0.83 0.86 0.84 28

accuracy 0.85 61

macro avg 0.85 0.85 0.85 61

weighted avg 0.85 0.85 0.85 61

=== SVC ===

Accuracy: 0.8689 Precision: 0.8333 Recall: 0.8929 F1: 0.8621 ROC\_AUC: 0.935064935064935

Classification report:

precision recall f1-score support

0 0.90 0.85 0.88 33

1 0.83 0.89 0.86 28

accuracy 0.87 61

macro avg 0.87 0.87 0.87 61

weighted avg 0.87 0.87 0.87 61

RandomForest best params: {'clf\_\_max\_depth': 6, 'clf\_\_n\_estimators': 50} best score(cv): 0.8732151503890634

SVC best params: {'clf\_\_C': 1, 'clf\_\_gamma': 'scale', 'clf\_\_kernel': 'rbf'} best score(cv): 0.8964173507651768

=== RandomForest\_Tuned ===

Accuracy: 0.9016 Precision: 0.9231 Recall: 0.8571 F1: 0.8889 ROC\_AUC: 0.9523809523809524

Classification report:

precision recall f1-score support

0 0.89 0.94 0.91 33

1 0.92 0.86 0.89 28

accuracy 0.90 61

macro avg 0.90 0.90 0.90 61

weighted avg 0.90 0.90 0.90 61

=== SVC\_Tuned ===

Accuracy: 0.8689 Precision: 0.8333 Recall: 0.8929 F1: 0.8621 ROC\_AUC: 0.935064935064935

Classification report:

precision recall f1-score support

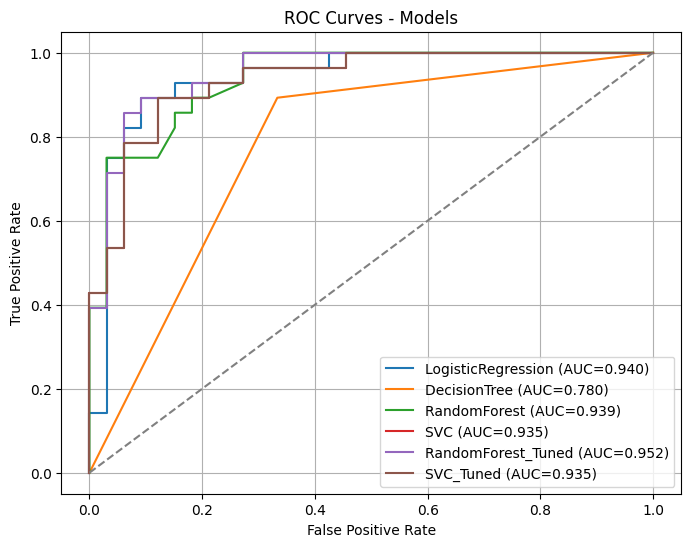
0 0.90 0.85 0.88 33

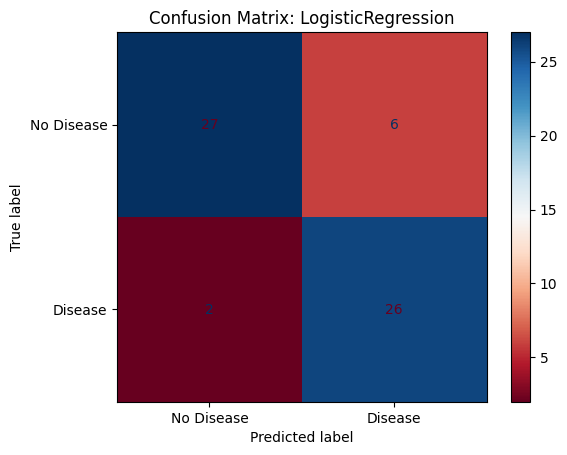
1 0.83 0.89 0.86 28

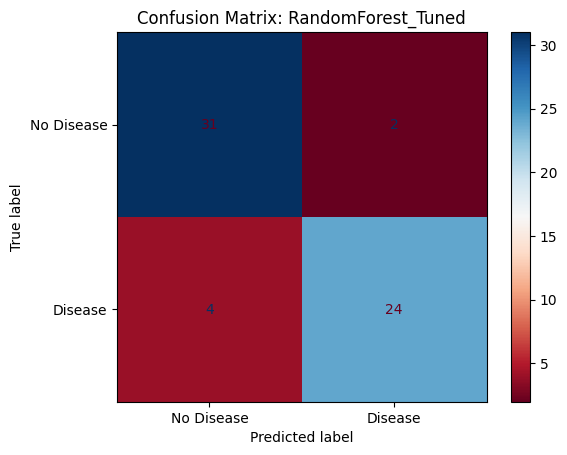
accuracy 0.87 61

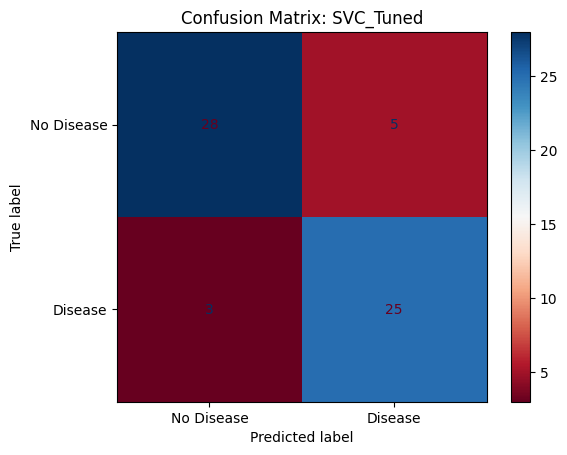
macro avg 0.87 0.87 0.87 61

weighted avg 0.87 0.87 0.87 61

****

****

****

****

**Best model by ROC\_AUC: RandomForest\_Tuned AUC: 0.9523809523809524**

**Saved best model to model\_RandomForest\_Tuned.pkl**

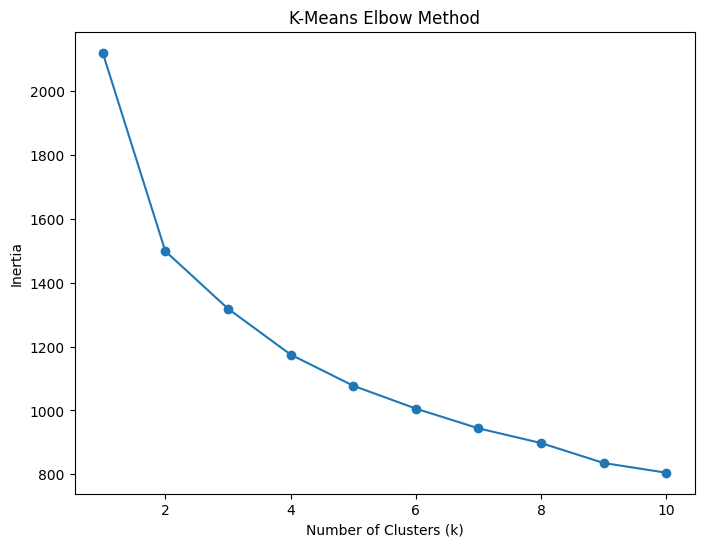
**Saved evaluation summary to results/evaluation\_metrics.txt**

**Cross-val ROC AUC scores (5-fold) for selected classifiers:**

**LogisticRegression: mean AUC=0.896 std=0.028**

**RandomForest\_Tuned: mean AUC=0.882 std=0.033**

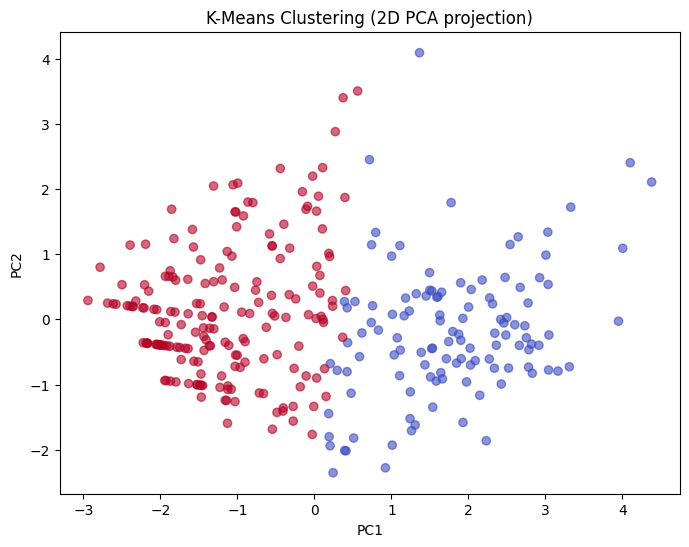
**05\_unsupervised\_learning\_results**

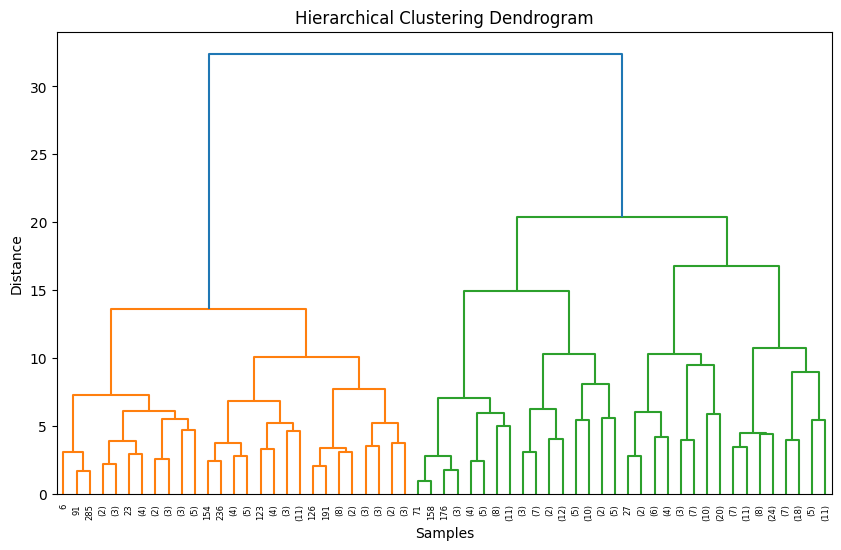


K-Means Results:

ARI (Adjusted Rand Index): 0.455860651499769

Silhouette Score: 0.2898160696666745

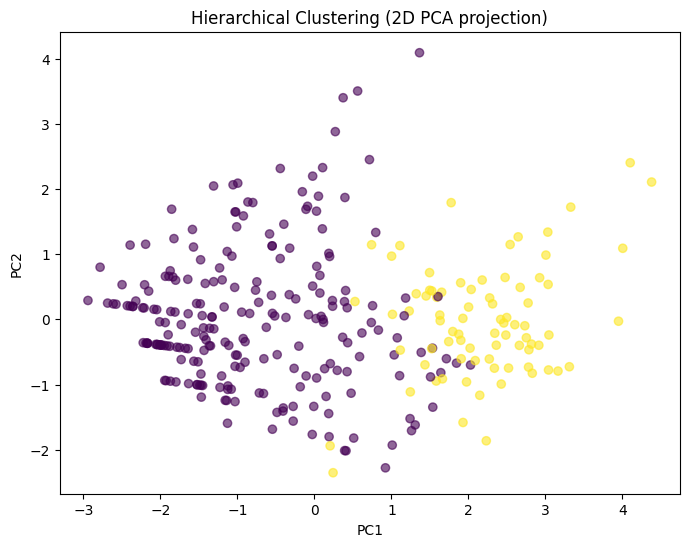




Hierarchical Clustering Results:

ARI (Adjusted Rand Index): 0.28647269720254115

Silhouette Score: 0.2686359215592109



**06\_hyperparameter\_tuning\_results**

**Train/Test sizes: (242, 7) (61, 7)**

**Best Logistic Regression Params: {'clf\_\_C': 10, 'clf\_\_penalty': 'l2', 'clf\_\_solver': 'lbfgs'}**

**=== Logistic Regression (Tuned) ===**

**Accuracy: 0.8688524590163934**

**Precision: 0.8125**

**Recall: 0.9285714285714286**

**F1: 0.8666666666666667**

**ROC-AUC: 0.9383116883116882**

**Best Decision Tree Params: {'clf\_\_max\_depth': 3, 'clf\_\_min\_samples\_leaf': 2, 'clf\_\_min\_samples\_split': 2}**

**=== Decision Tree (Tuned) ===**

**Accuracy: 0.8688524590163934**

**Precision: 0.8571428571428571**

**Recall: 0.8571428571428571**

**F1: 0.8571428571428571**

**ROC-AUC: 0.8701298701298701**

**Best Random Forest Params: {'clf\_\_n\_estimators': 100, 'clf\_\_min\_samples\_split': 5, 'clf\_\_min\_samples\_leaf': 2, 'clf\_\_max\_depth': 10}**

**=== Random Forest (Tuned) ===**

**Accuracy: 0.9016393442622951**

**Precision: 0.8928571428571429**

**Recall: 0.8928571428571429**

**F1: 0.8928571428571429**

**ROC-AUC: 0.9556277056277056**

**Best SVC Params: {'clf\_\_C': 1, 'clf\_\_gamma': 'scale', 'clf\_\_kernel': 'rbf'}**

**=== SVC (Tuned) ===**

**Accuracy: 0.8688524590163934**

**Precision: 0.8333333333333334**

**Recall: 0.8928571428571429**

**F1: 0.8620689655172413**

**ROC-AUC: 0.935064935064935**