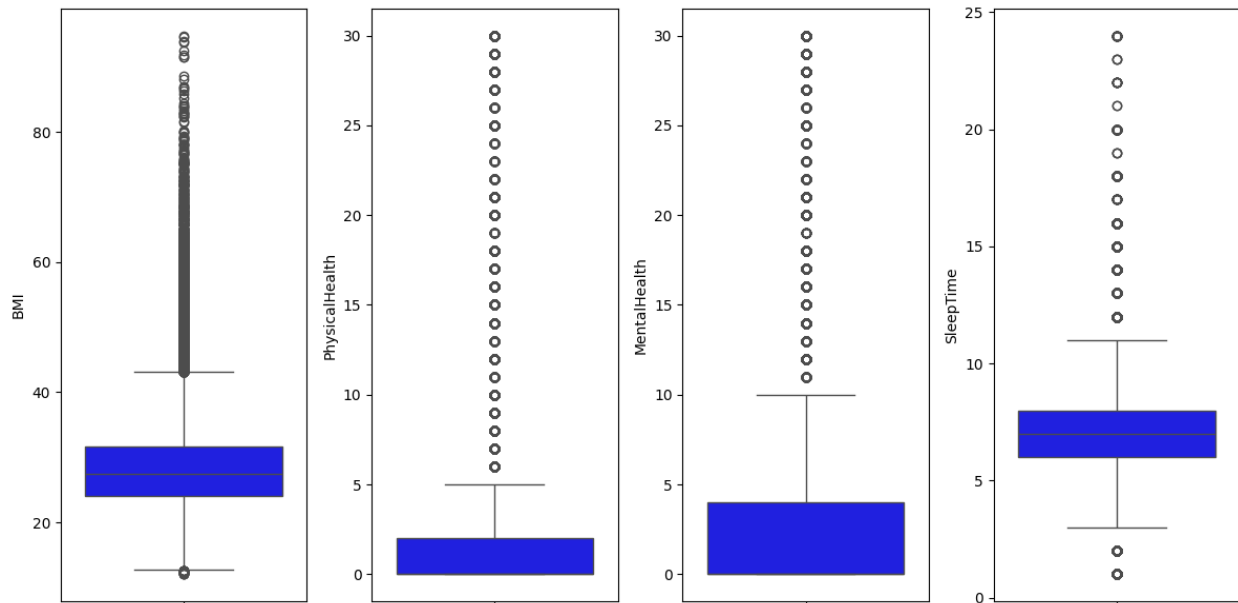# Project Output Report

Heart Disease Prediction Using Machine Learning
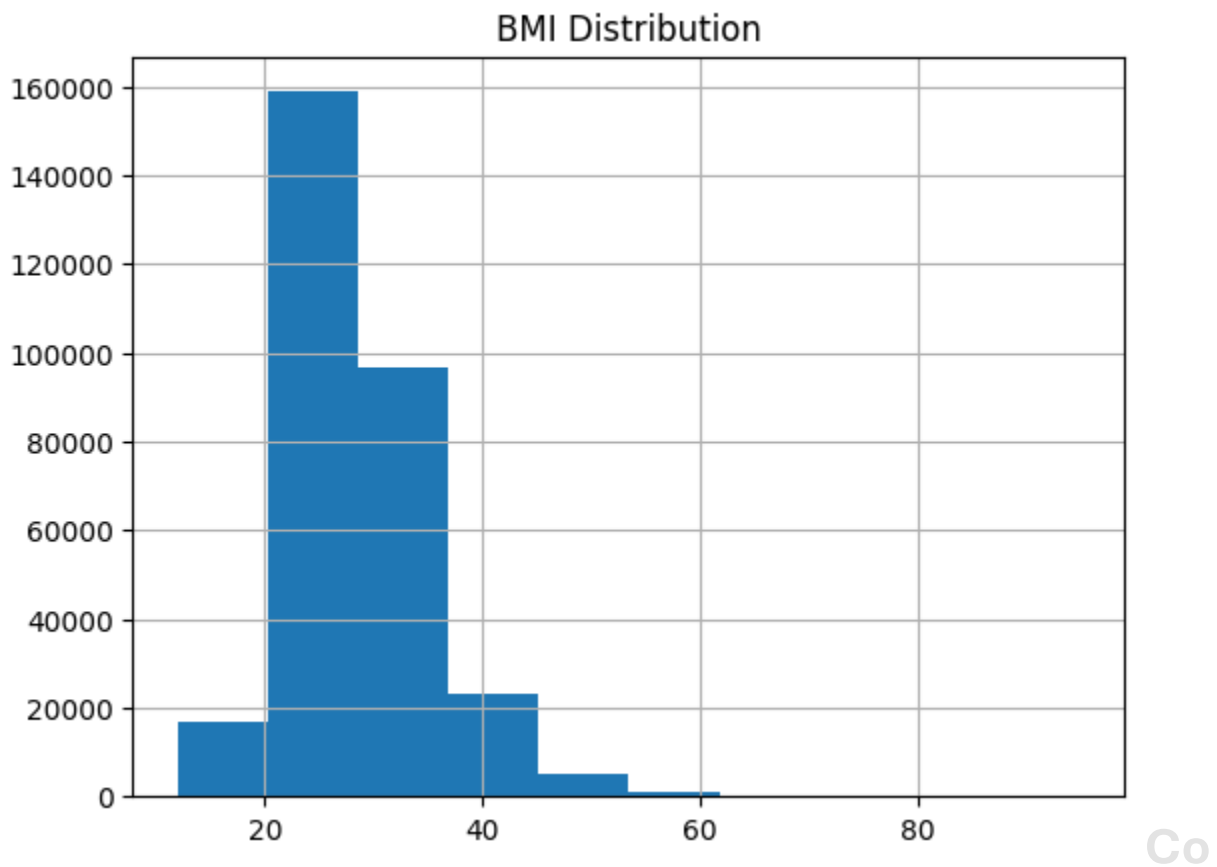
Team Members

- Aisha Samir
- Rola Hany
- Menna Mohsen
- Menna Akram
- Nada Etman

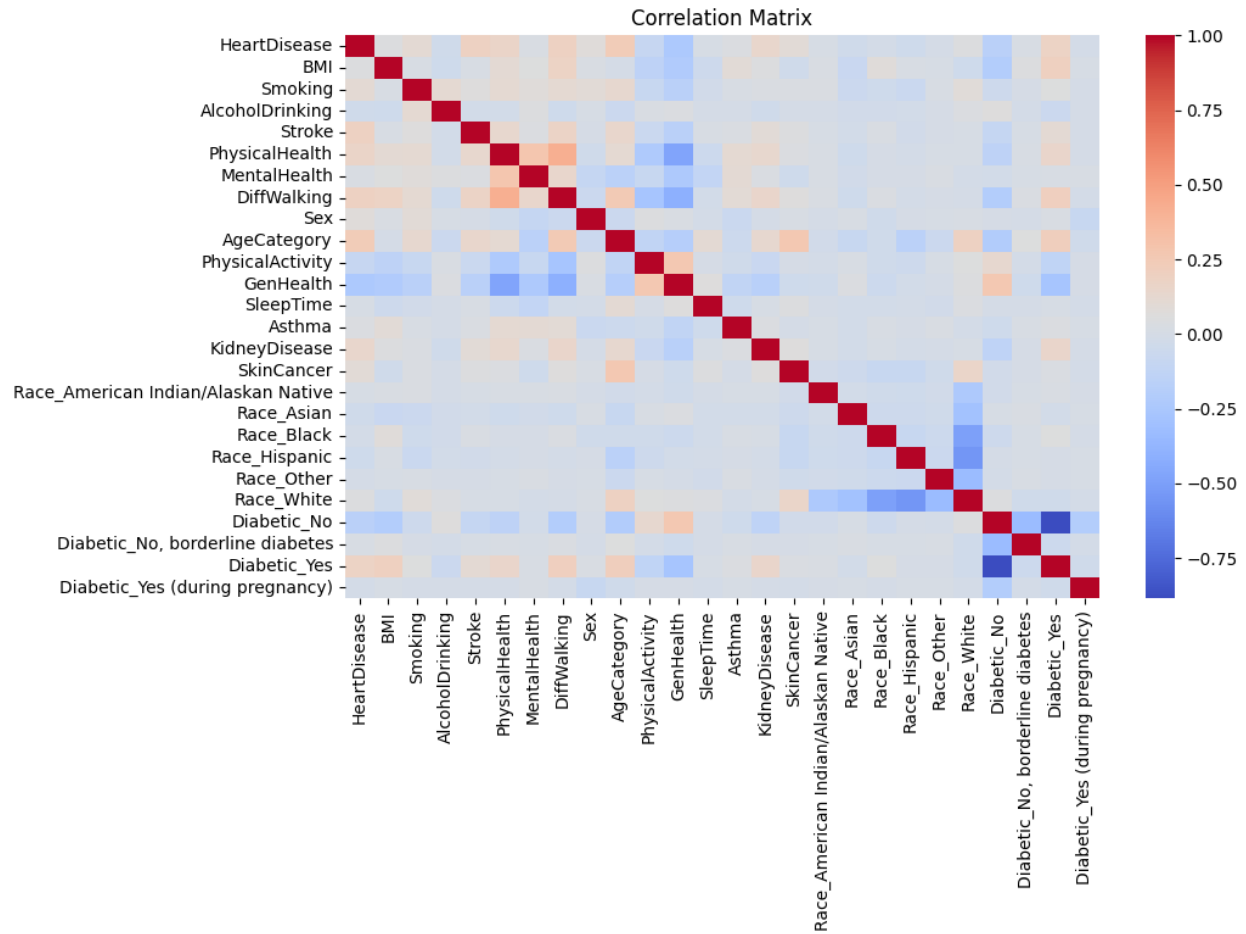## Exploratory Data Analysis (EDA) and Preprocessing Output:
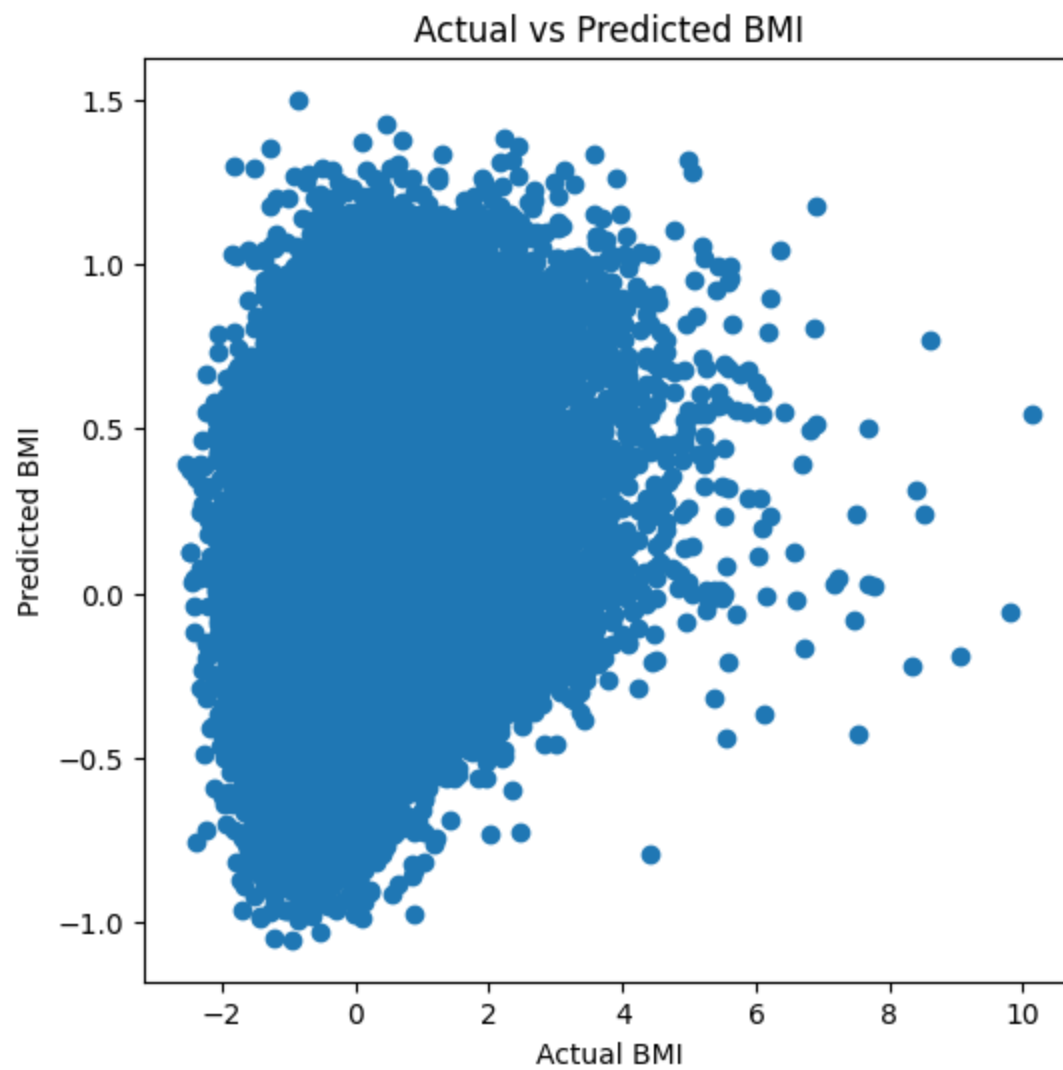
Outliers

Distribution

BMI Distribution

Co

Correlation

Correlation Matrix

Regression Metrics

Regression Metrics: MSE: 0.891205761918424

RMSE: 0.944036949445531

R^2 Score: 0.10667930509316825

Actual vs Predicted BMI

Classification Task Output

Samples: 301717

Features: 25

Target classes: [0 1]

Make predictions

[[0.99540401 0.00459599]

 [0.99679749 0.00320251]

 [0.98444963 0.01555037]

 [0.93295839 0.06704161]

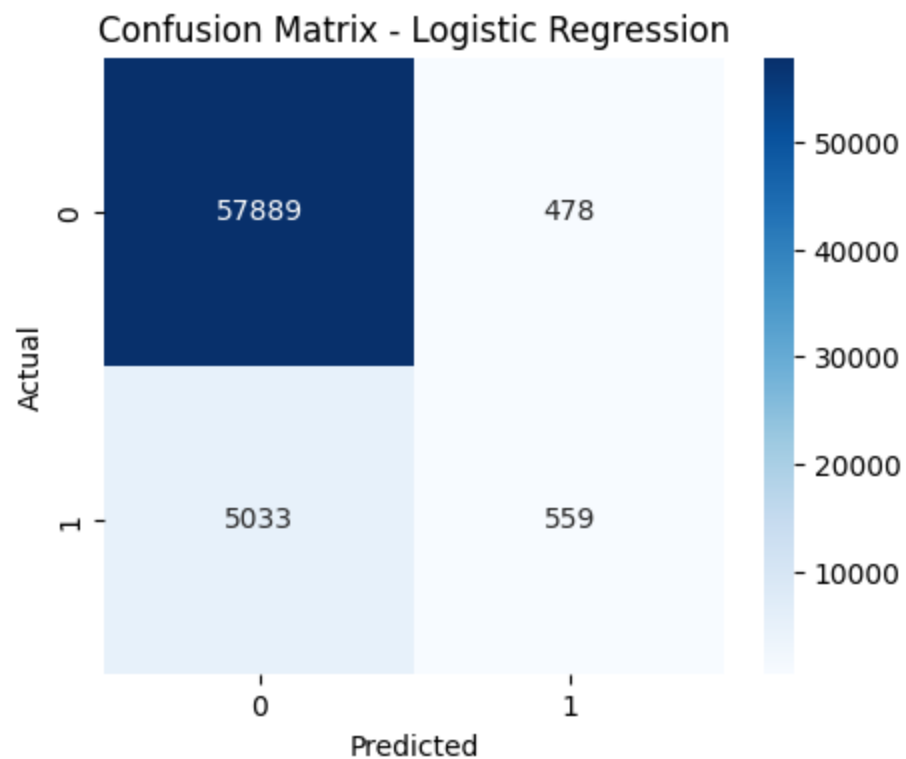 [0.77948135 0.22051865]]

Evaluate logistic regression model

Accuracy: 0.9138

 Confusion Matrix:

 [[57889 478] [5033 559]]

 Precision: 0.5390549662487946

Recall: 0.09996423462088698

Confusion Matrix - Logistic Regression

Clustering Task Output

Classification Results:

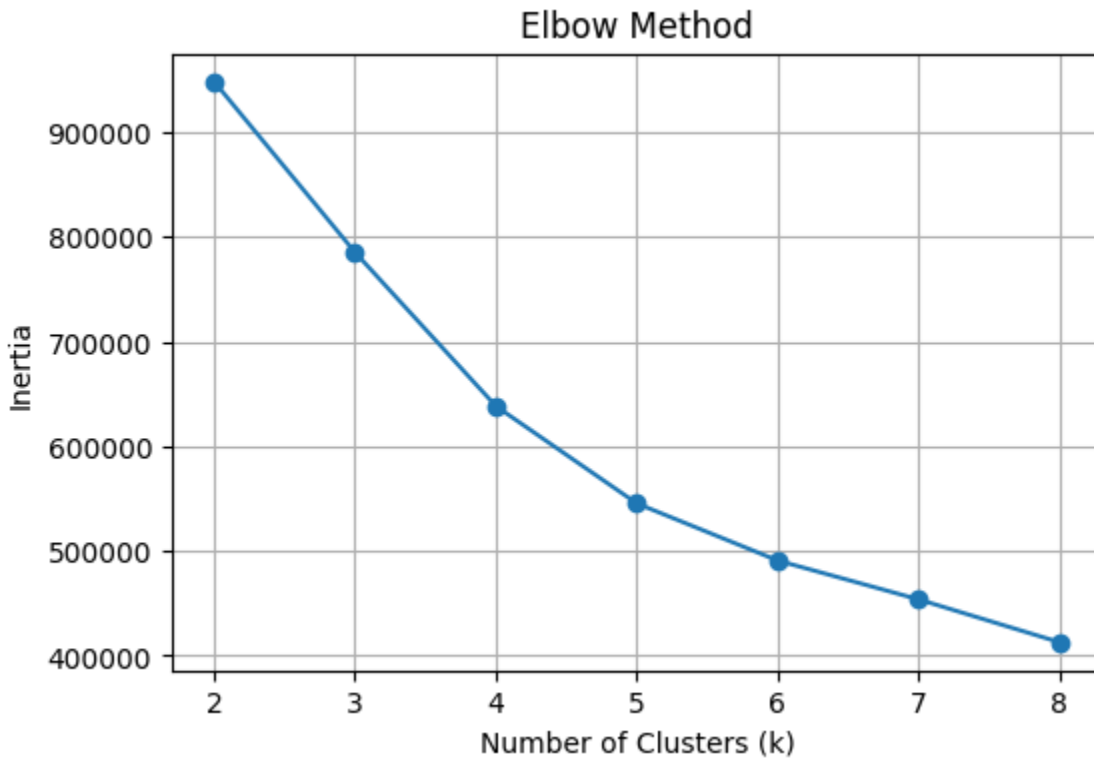**Accuracy: 0.9137**

Precision: 0.5339

Recall: 0.1041

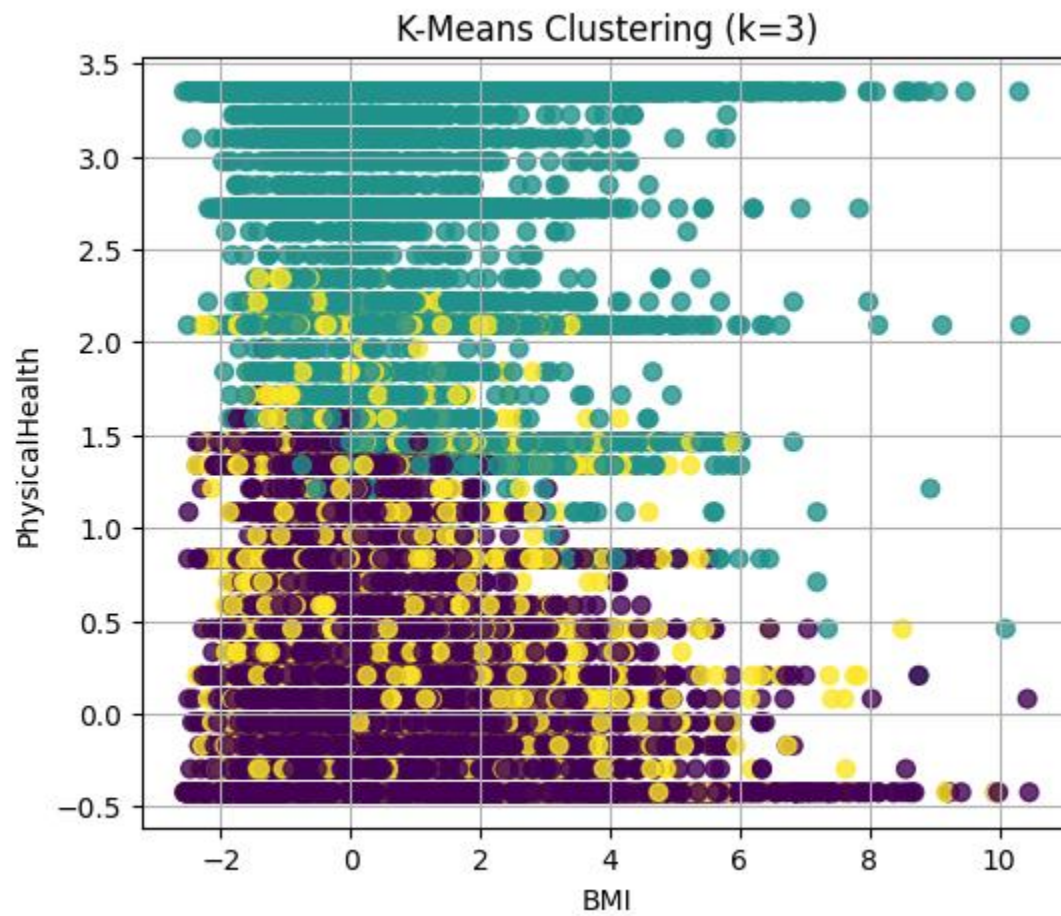## Regression Results:
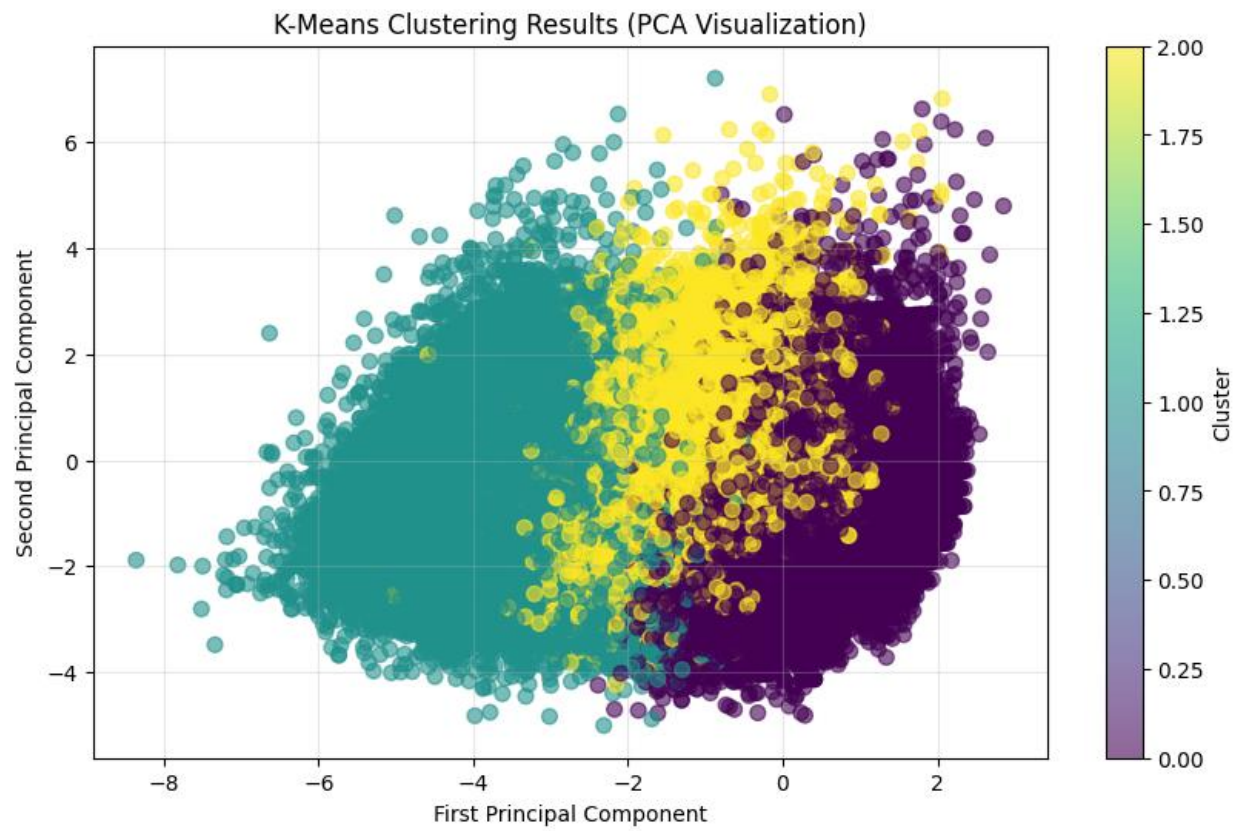
R^2 Score: 0.1097

MSE: 0.8821

RMSE: 0.9392

Elbow Methed

Elbow Method

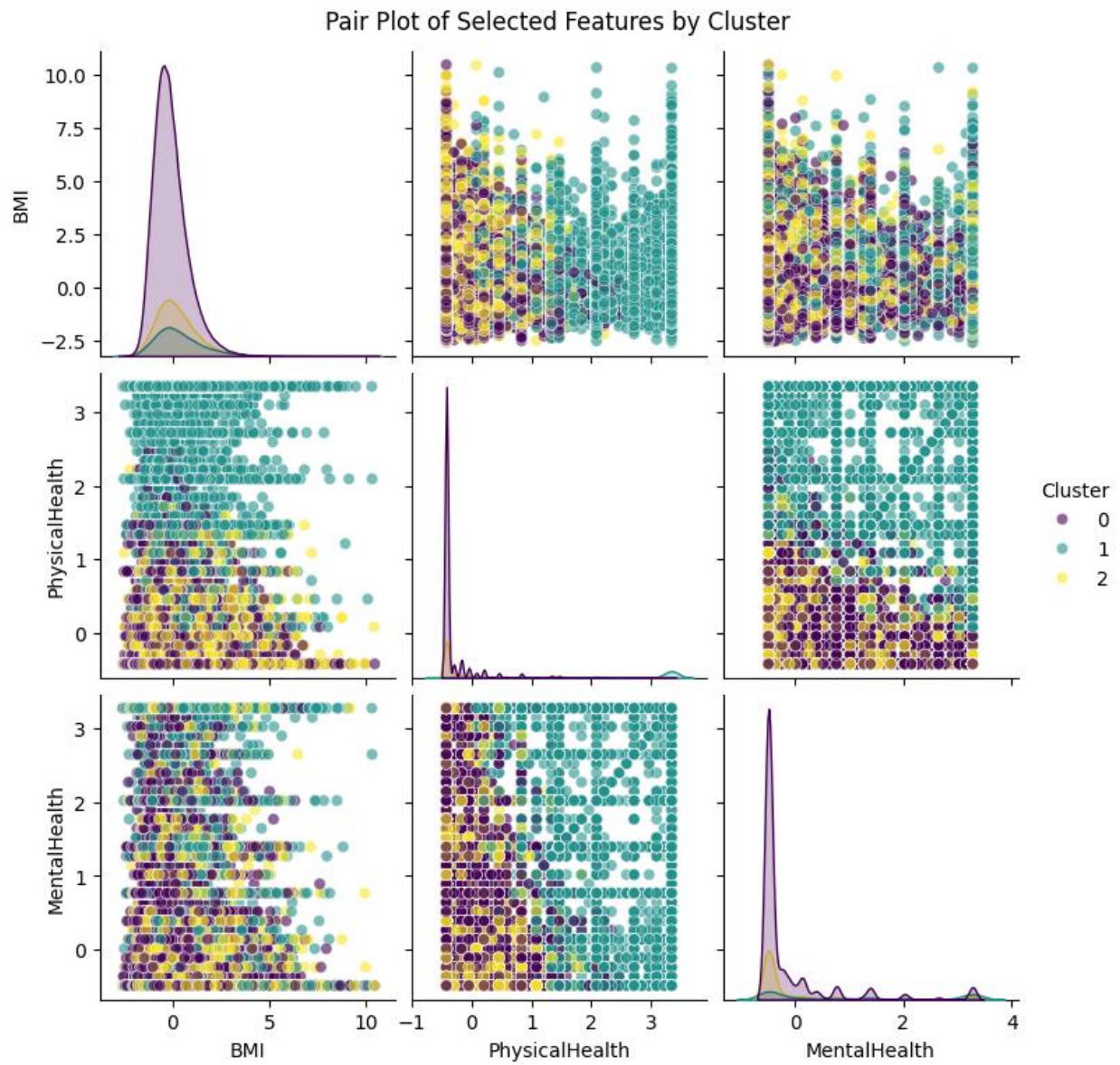# Silhouette Score for k=3: 0.465

K-Means Clustering (k=3)

K-Means

K-Means Clustering Results (PCA Visualization)

<Figure size 1000x800 with 0 Axes>

Pair Plot of Selected Features by Cluster

boxplot_clusters