Unsupervised Analysis to Examine the Relationship Between PHQ-9 scores and Heart Failure Indicators

Introduction

Heart failure is a serious and life-threatening condition that not only affects physical health but is also closely linked to mental health issues such as depression. Studies have shown that patients with heart failure often experience symptoms of depression, which can negatively impact treatment outcomes, quality of life, and even survival rates. Understanding this connection is important for better diagnosis, treatment, and overall patient care.

This project focuses on exploring the relationship between depression severity and heart failure condition using unsupervised machine learning techniques. The goal is to examine whether patients with higher depression scores also tend to have worse heart health indicators, such as lower ejection fraction, elevated BNP levels, higher hospitalization rates, and increased mortality. The project uses clustering algorithms to group patients based on similarities in clinical and psychological features, without relying on predefined labels.

The main aims of this project are:

A. To apply multiple clustering methods (K-Means, Hierarchical Clustering, and DBSCAN) to identify patient subgroups.

B. Comparing depression scores (PHQ-9) across clusters and evaluate their relationship with heart health outcomes.

C. To determine whether high depression severity is linked with worse heart failure progression.

The analysis is based on a publicly available dataset titled “Depression and Heart Failure”, originally published as part of a scientific study on patient outcomes. The dataset includes 425 patient records.

By combining medical variables with depression scores, the dataset provides a rich foundation for analyzing the interaction between mental and physical health. Through clustering and data visualization in Orange, this project aims to provide insights into how depression severity may help predict heart failure outcomes and highlight the importance of mental health screening in cardiac care.

Process Methodology

This project was carried out using the Orange Data Mining software, which provides a visual workflow environment for data analysis and machine learning. The methodology followed a structured process to prepare the data, apply clustering algorithms, and interpret results using visual tools. The steps are outlined below:

1. Data Import and Preprocessing

The dataset, containing 425 patient records and several clinical variables, was first imported into Orange using the **File** widget. The **Preprocess** widget was used to handle missing values, normalize numeric features, and ensure that all variables were in the correct format. This step helped improve the performance and accuracy of clustering algorithms.

2. Feature Selection

Relevant features were selected using the **Select Columns** widget. Key variables included:

- PHQ-9 (depression score)
- Ejection fraction (%)
- BNP/NT-proBNP levels
- Death and hospitalization status
- Kidney function markers (e.g., blood urea nitrogen, eGFR)
- Blood pressure and age

These features were chosen based on their clinical relevance to both depression and heart failure.

3. Clustering Algorithms

Three unsupervised clustering algorithms were applied to the preprocessed dataset to uncover natural patterns among patients:

1. **K-Means Clustering**

   The optimal number of clusters was selected based on silhouette scores. Patients were grouped by similarity in clinical and psychological features. The results were further analyzed using scatter plots, MDS visualizations, and summary tables.

2. **Hierarchical Clustering**

   This method was used to build a tree-like structure based on average linkage. Clusters were selected using a height ratio of around 89%, and results were interpreted using box plots to compare depression scores and health outcomes across groups.

3. **DBSCAN (Density-Based Spatial Clustering)**

   DBSCAN was used to detect clusters based on density rather than distance. The optimal epsilon value was determined using the k-distance plot. This method helped identify core clusters as well as outlier patients.

4. Data Visualization and Analysis

To interpret and compare the results of each clustering method, multiple visual tools were used:

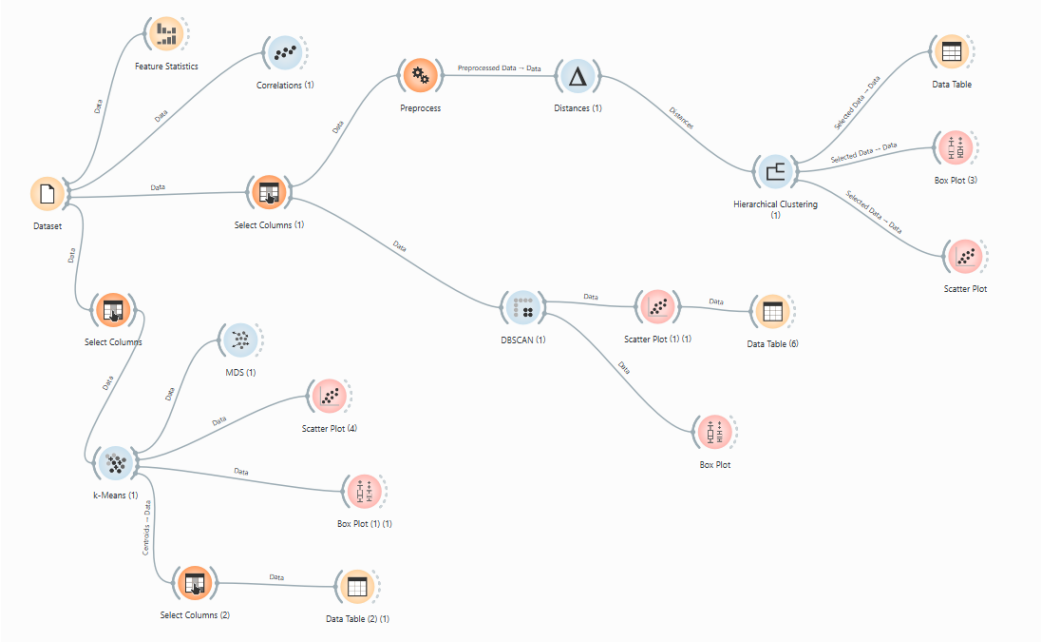
- **Box Plots** to compare PHQ-9 scores, ejection fraction, and other clinical measures across clusters.
- **Scatter Plots** to explore the relationships between depression and heart-related variables (e.g., BNP levels, mortality, kidney function).
- **MDS and Distance Maps** to visualize the separation and structure of clusters in a reduced dimensional space.

5. Hypothesis Evaluation

Finally, all clustering results were analyzed to evaluate the hypothesis:

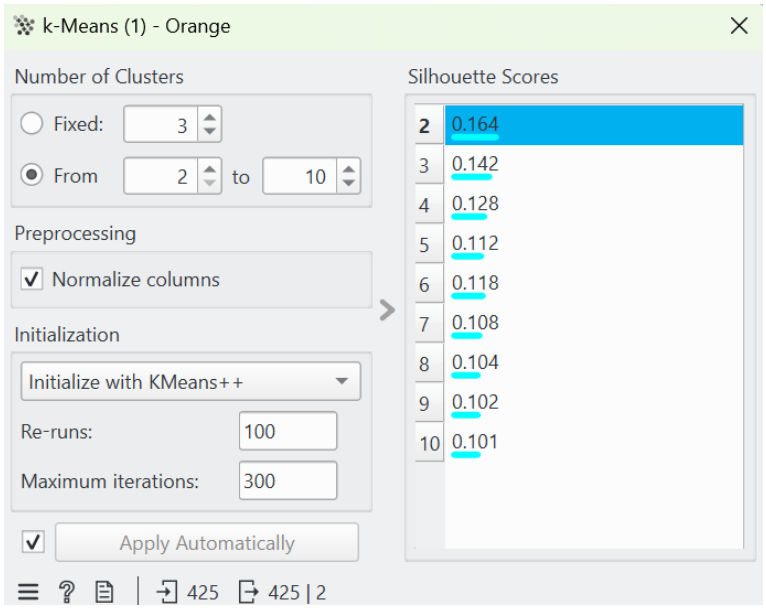
"Patients with higher PHQ-9 scores (more depression) tend to have worse heart failure conditions."

This was done by comparing cluster characteristics across different algorithms and identifying consistent patterns linking depression severity with poor health indicators.



1. K-Means Clustering

K-Means clustering was applied to identify patterns between depression severity (measured by PHQ-9 scores) and indicators of heart health. Based on silhouette scores, two clusters provided the best grouping of data. Cluster assignments were used to analyze differences in health outcomes across patient groups.



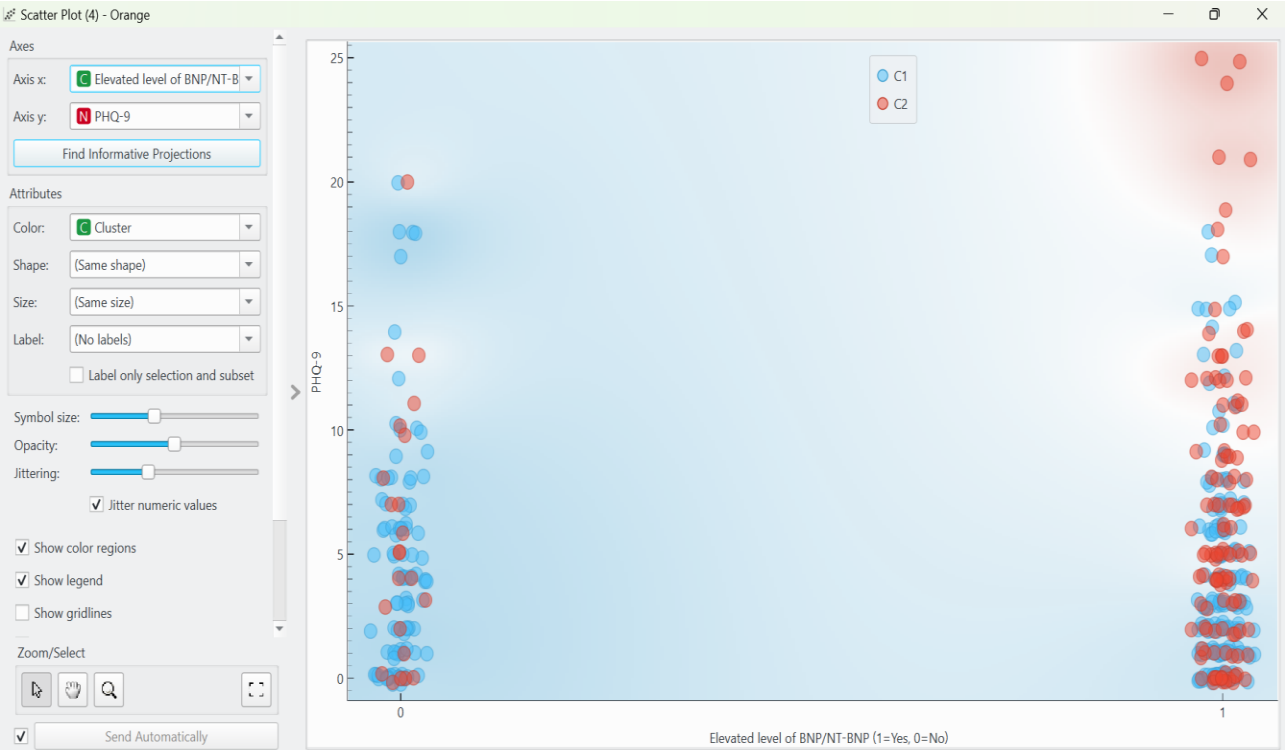
Found that people in Cluster C2 had higher depression scores (PHQ-9). These same people also showed signs of worse heart health, they had lower heart pumping strength (measured by ejection fraction), higher death rates, and more often had high BNP levels, which is a clear sign of heart failure.

Data Table (2) (1) - Orange

	Cluster	Silhouette	PHQ-9	jection fraction (%	ogy HF(1=Yes, 0=N	r diabetes mellitu	of BNP/NT-BNP (1	ath (1=Yes, 0=No)	talized (1=Yes, 0=I	ale (1=Yes, 0=No)
1	C1	0.563301	-0.162	0.0157465	0.376307	0.655052	0.344948	0.965157	0.390244	0.560976
2	C2	0.526902	0.337	-0.0327482	0.528986	0.536232	0.166667	0.355072	0.101449	0.608696

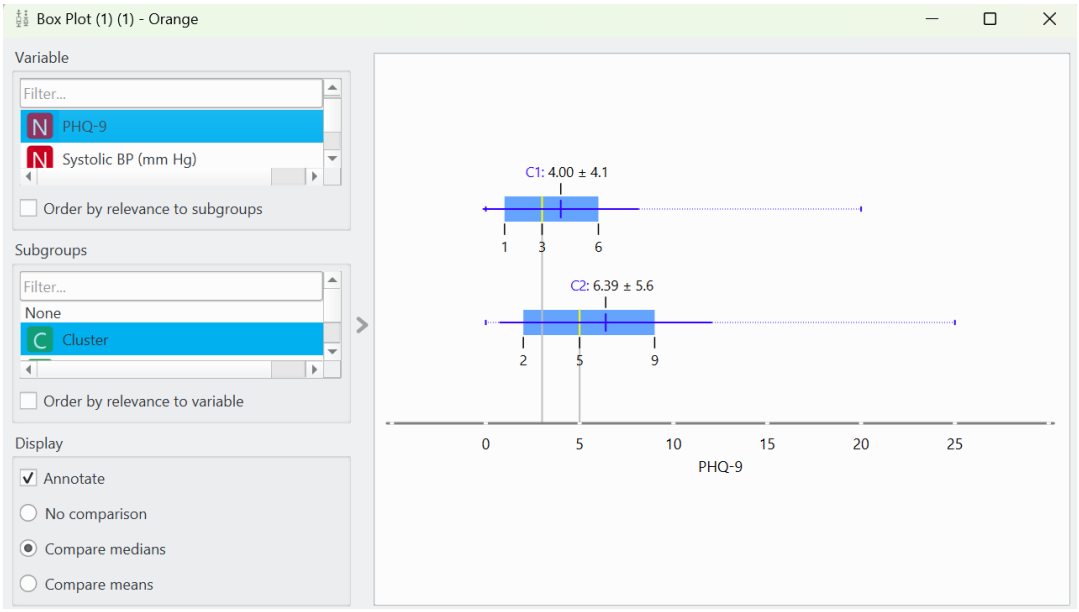
Alongside depression, several heart-related indicators also differed between the clusters. Patients in Cluster C2 had lower average ejection fraction, meaning their heart was pumping less effectively. This cluster also had a much higher death rate (96.5%), more frequent hospitalizations (39.0%), and a higher number of cases with elevated BNP/NT-BNP levels (34.4%), which are biomarkers of heart failure. In contrast, Cluster C1 showed better outcomes on all these measures.

Scatter plot of PHQ-9 scores versus BNP levels showed that most patients with both high depression and elevated BNP belonged to Cluster C2.

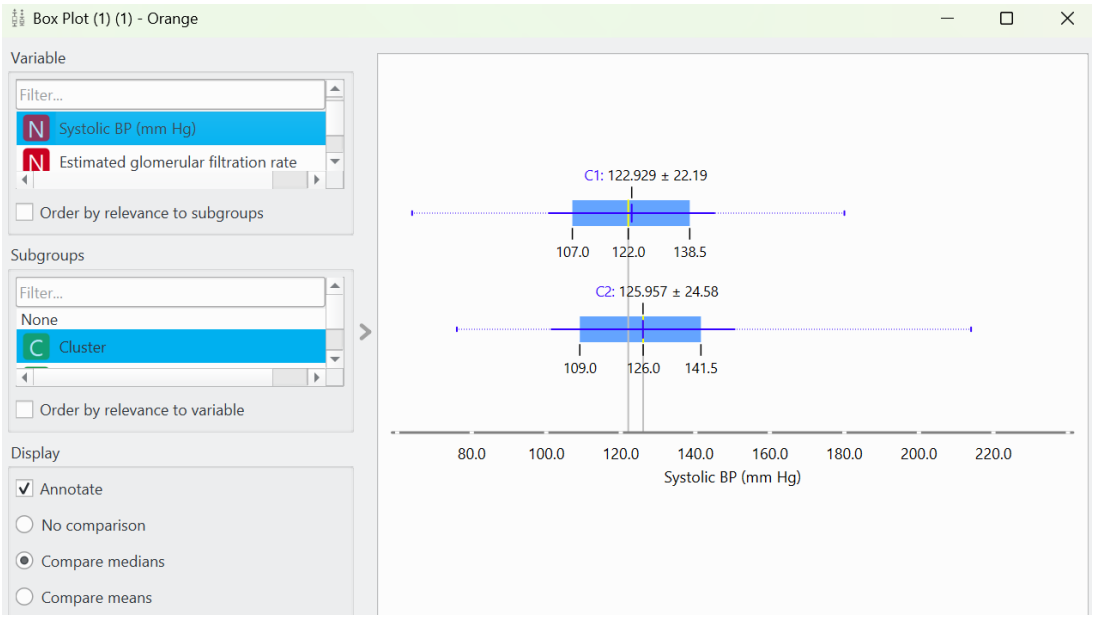


PHQ-9 scores and systolic BP:

Box plots were used to compare the PHQ-9 depression scores and systolic blood pressure values across the two clusters formed by K-Means. This helped to visually understand how depression levels and heart-related measures varied among patient groups.

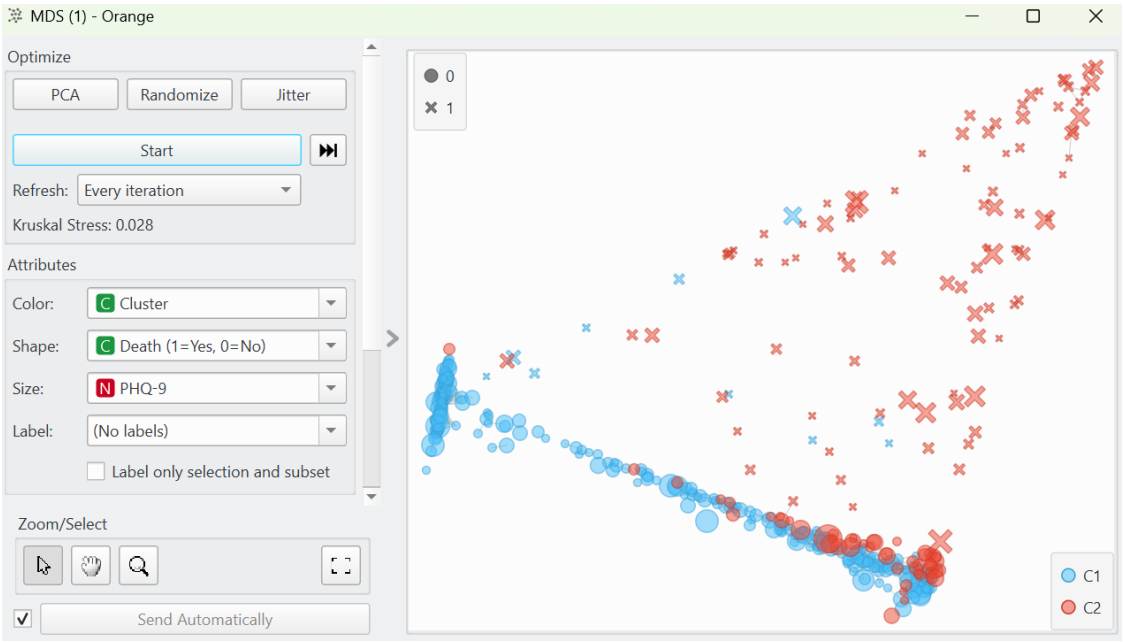


The first box plot showed the distribution of PHQ-9 scores, which reflects the severity of depression. Patients in Cluster C2 had a higher average PHQ-9 score of 6.39 ± 5.6, while those in Cluster C1 had a lower average score of 4.00 ± 4.1. The median score in Cluster C2 was also clearly higher than in Cluster C1. This difference confirms that patients at Cluster C2 were more depressed. Since clustering was based on health-related variables, the higher depression in C2 already suggests a connection with more severe heart conditions.



The second box plot compared systolic blood pressure between the clusters. Patients in Cluster C2 had a higher average systolic BP (125.96 ± 24.58 mm Hg) compared to Cluster C1 (122.92 ± 22.19 mm Hg). Although this difference is not large, it shows a trend of slightly higher blood pressure in the more depressed cluster. Elevated blood pressure increases the workload on the heart and is commonly associated with risk of heart failure.

Box plots are especially useful because they show the spread and central values of the data, allowing for a clear comparison. In both plots, Cluster C2 consistently shows higher values for PHQ-9 and systolic BP, which are important indicators of mental and physical stress.



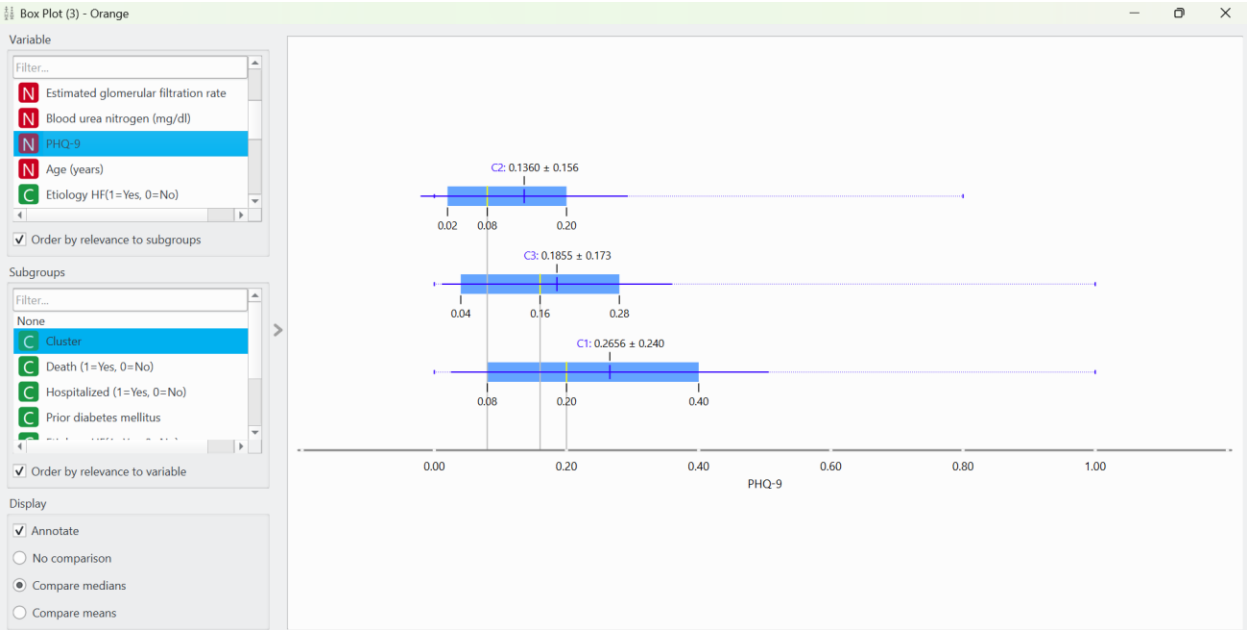
The MDS plot visually confirmed the separation of clusters. Patients in Cluster C2 (marked in red) had both higher PHQ-9 scores (represented by larger circle sizes) and a higher incidence of death (cross shapes), reinforcing the numeric summary. The MDS plot confirmed that the two clusters were well-separated, and most patients who had died were from the high-depression cluster.

2. Hierarchical Clustering

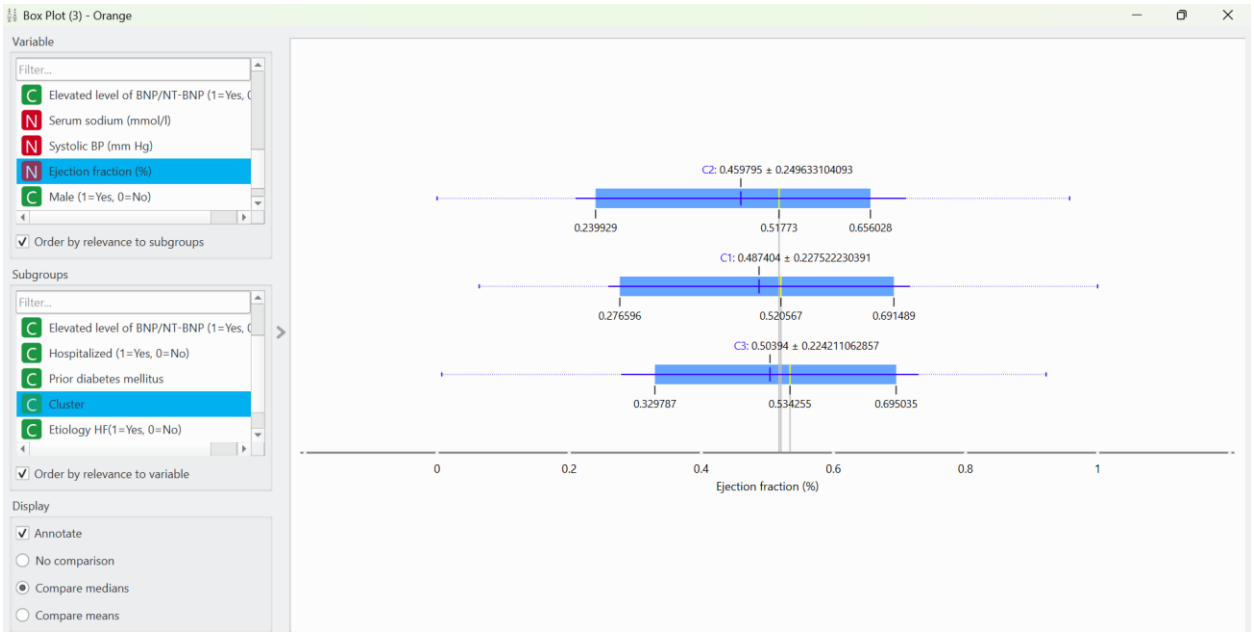
Hierarchical clustering was used to explore the relationship between depression severity and heart health. The linkage method was set to "Average", and three main clusters were selected using a height ratio of 89.4%. This allowed the grouping of patients based on similarities in their health and depression-related features.

PHQ-9 & Ejection Fraction:

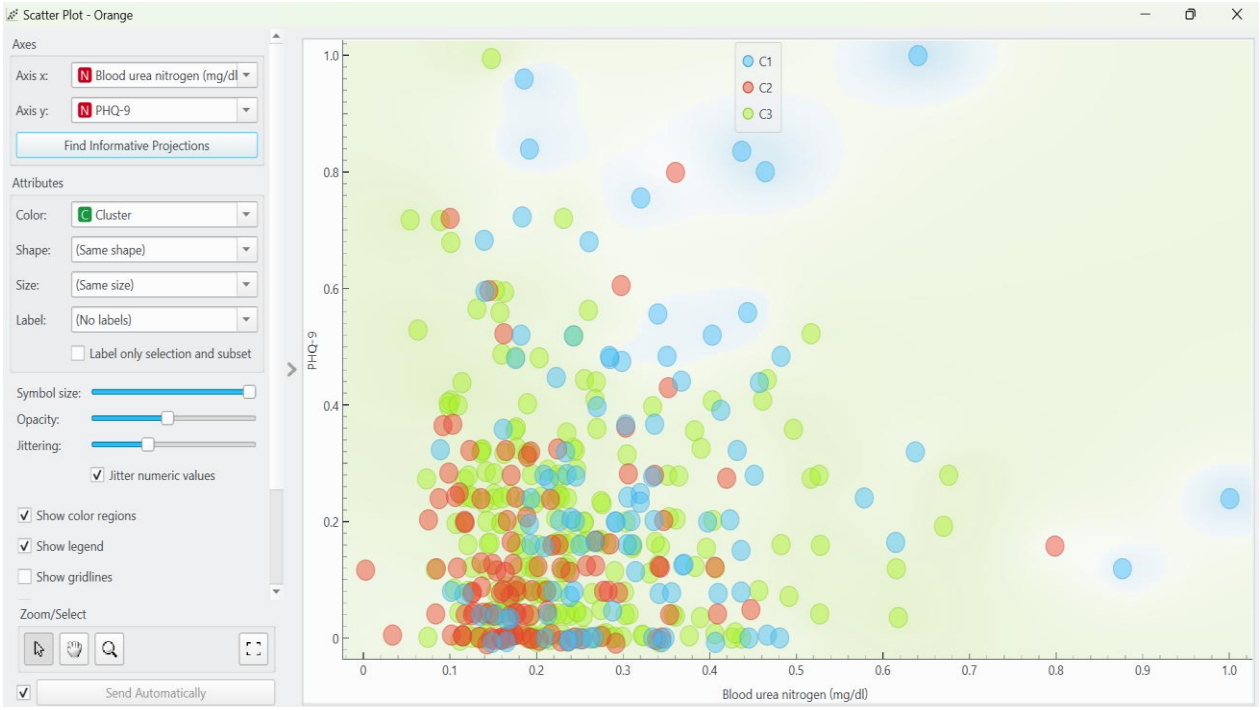
The clusters were then analyzed through various box plots and scatter plots. The box plot of PHQ-9 scores showed that Cluster C1 had the highest average depression score (0.2656 ± 0.240), followed by Cluster C3 (0.1855), while Cluster C2 had the lowest score (0.1360). This suggests that patients in Cluster C1 were experiencing more severe depression symptoms.



To understand the heart condition of these groups, a box plot of ejection fraction, a key measure of heart pumping strength, was used. The results showed that Cluster C1 had the lowest average ejection fraction (0.487) among the three clusters, which means their heart function was weaker. Cluster C2 and C3 had higher average ejection fractions, suggesting better cardiac function.



Relationship b/w PHQ-9 & BUN:

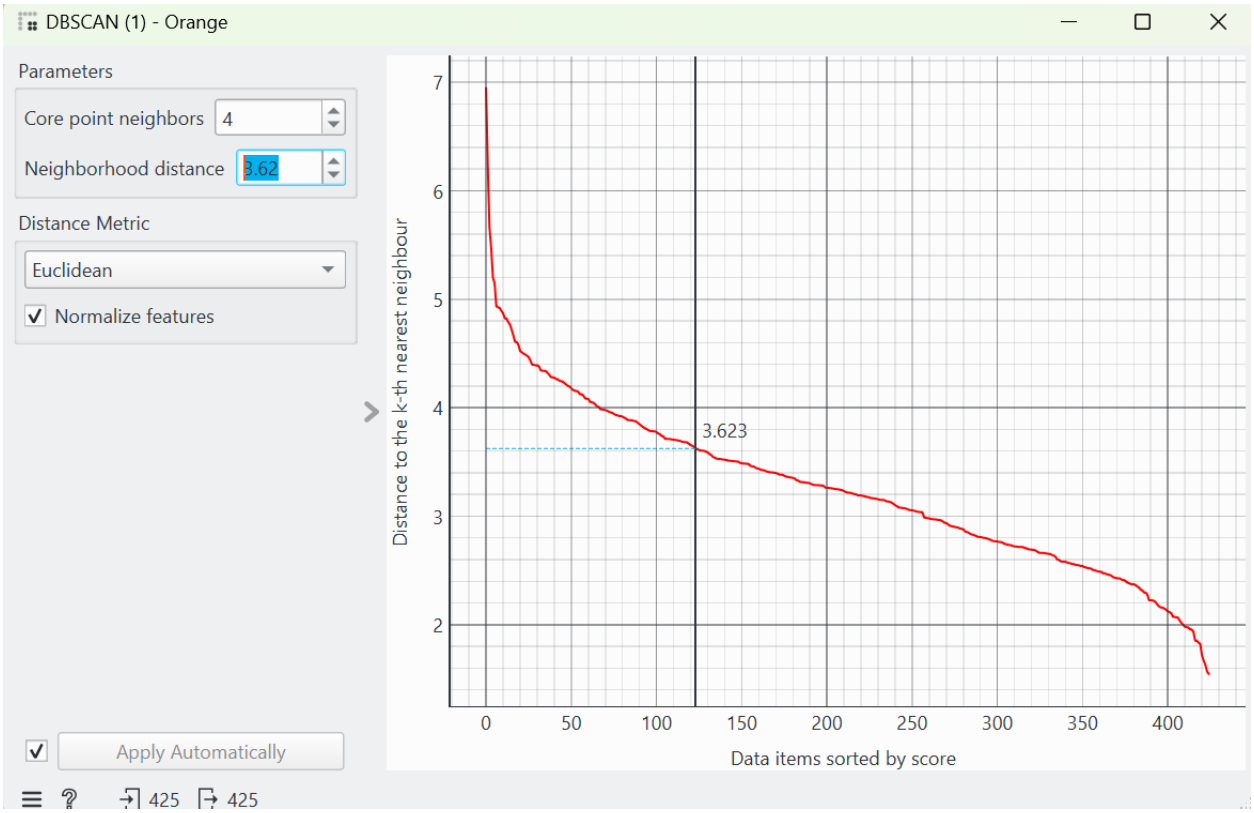


The plot shows that Cluster C1, which earlier showed higher depression scores and lower heart function, also includes many patients with elevated BUN levels. This pattern adds further support to the idea that patients with more severe depression (higher PHQ-9) tend to fall into the group with worse physical health markers.

Although values are spread across all clusters, the upper-right section of the plot representing high BUN and high PHQ-9, is mostly occupied by patients from Cluster C1, reinforcing the earlier box plot results. This overlap between psychological and physical stress markers strengthens the connection between depression and poor heart health.

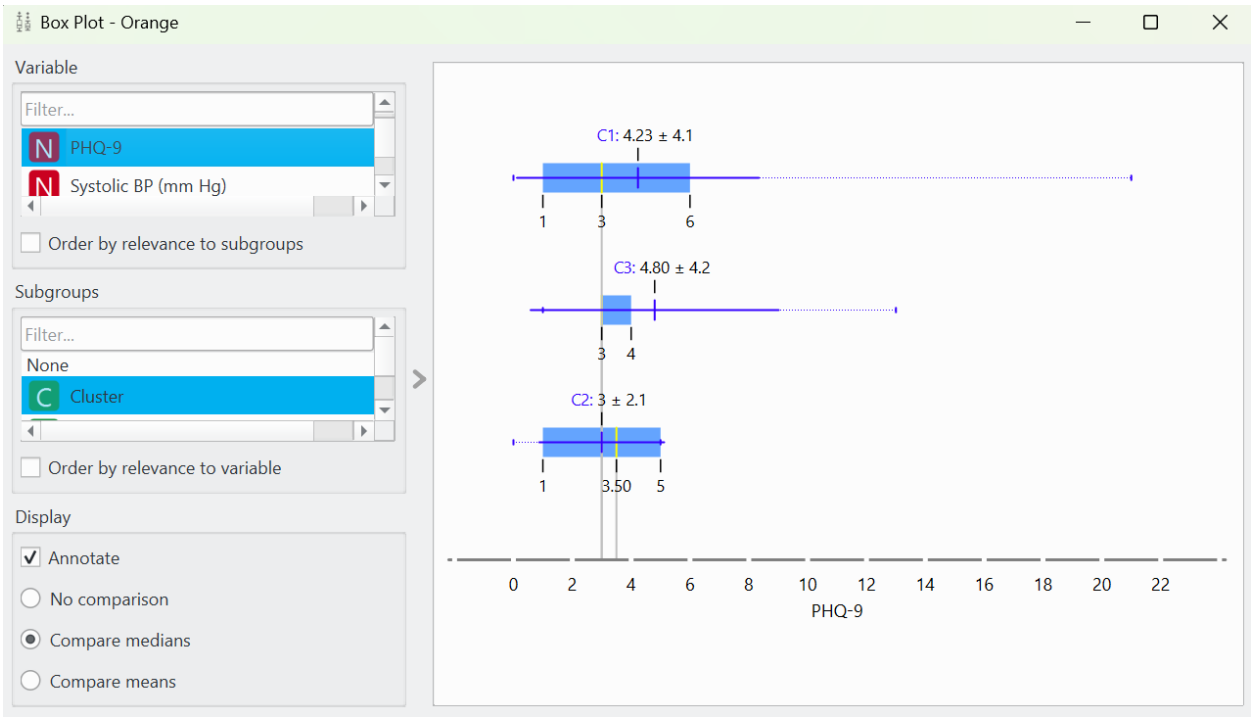
3. DBSCAN Clustering

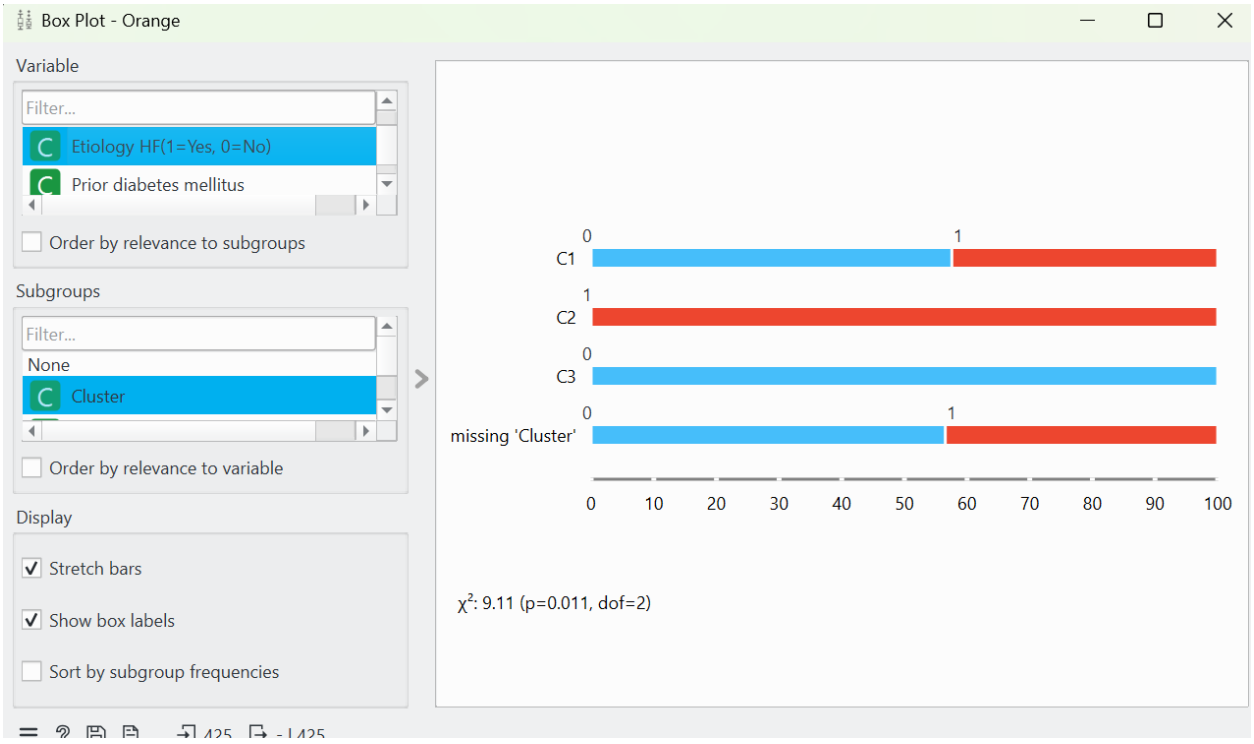
DBSCAN was used to identify natural groupings in the data without pre-setting the number of clusters. Based on the elbow point in the k-distance graph, a neighborhood distance of 3.62 was selected. This helped separate core clusters from noise effectively.



PHQ-9 and Etiology HF:

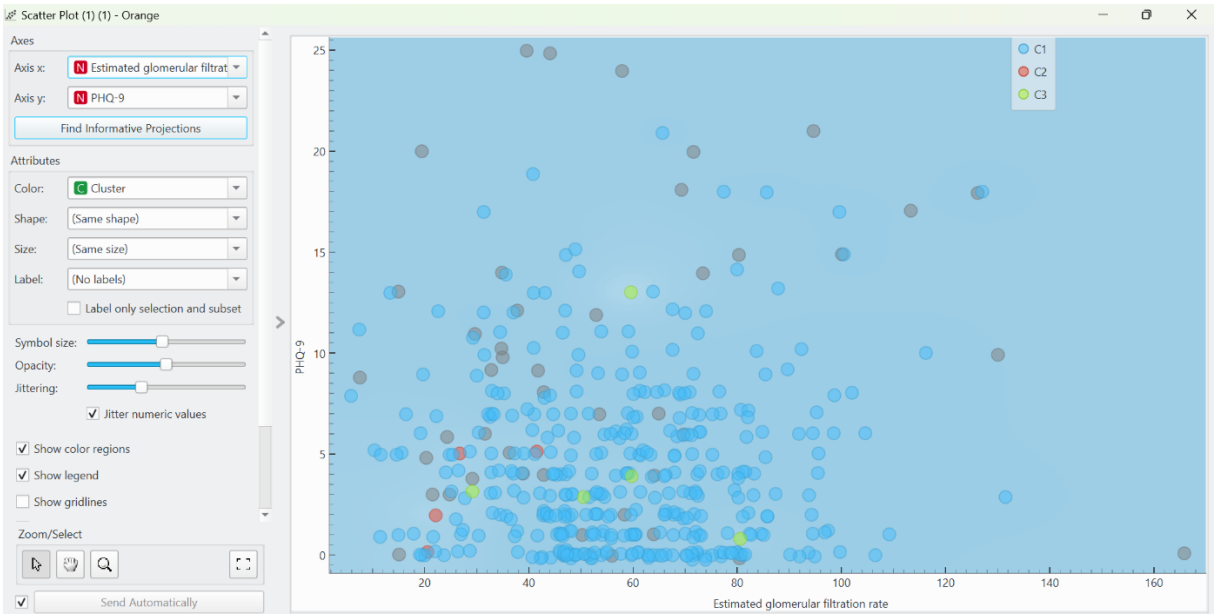
The box plot of PHQ-9 scores showed that Cluster C3 had the highest average depression score (4.80 ± 4.2), followed by C1 (4.23 ± 4.1), while C2 had the lowest (3.0 ± 2.1). This indicates that patients in Clusters C1 and C3 were more depressed than those in C2.





In another boxplot, it is showing that heart failure causes (Etiology HF) across clusters, Cluster C2 had a higher share of patients with known heart failure etiology. However, depression scores in this group were lower, while those with higher PHQ-9 scores (especially C3) had more variable clinical conditions.

The scatter plot, “Estimated glomerular filtration rate” against “PHQ-9” showed that most patients with low kidney function and high depression scores were grouped in Cluster C1. This pattern suggests a link between depression and worsening organ function, which is commonly seen in advanced heart failure.



Taken together, the box plots and scatter plots support the hypothesis: clusters with higher PHQ-9 depression scores also showed signs of worse heart conditions. DBSCAN results confirm that depression and heart health severity are related, consistent with earlier findings from K-Means and Hierarchical clustering.

Conclusion

This project aimed to explore the relationship between depression severity and heart failure outcomes using unsupervised clustering techniques in Orange Data Mining. Through the use of K-Means, Hierarchical Clustering, and DBSCAN, patients were grouped based on similarities in clinical and psychological features, allowing for meaningful analysis without prior labels.

Across all three clustering methods, a consistent pattern was observed: patients with higher PHQ-9 scores tended to fall into clusters that showed worse heart-related conditions. These conditions included lower ejection fraction, higher levels of BNP (a marker of heart failure), increased hospitalization rates, and higher mortality. Additional factors such as kidney function markers and blood pressure also supported these findings. The box plots and scatter plots provided visual confirmation, showing clear differences in health outcomes among patient groups with varying levels of depression.

These results strongly support the initial hypothesis that higher depression severity is linked with poorer heart failure progression. The study highlights the importance of considering mental health alongside physical health when assessing patients with chronic cardiac conditions. Identifying patients with elevated depression scores could help in early intervention, leading to better management of both psychological and physiological aspects of heart failure.