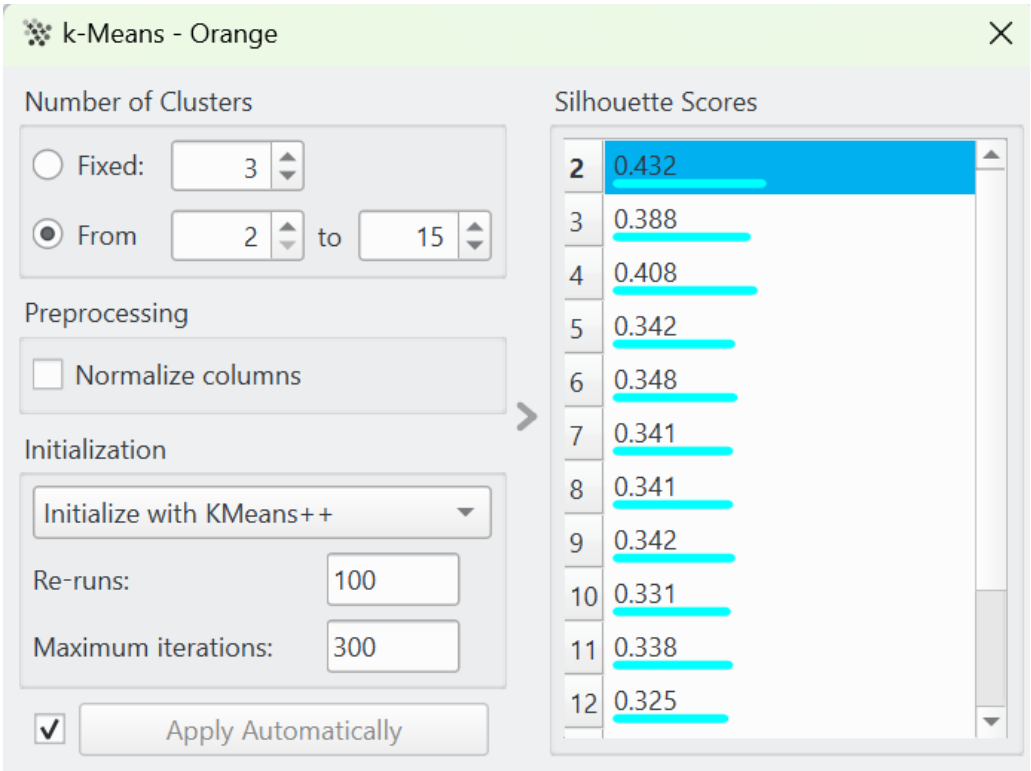
Analysis of Out-of-Hospital Cardiac Arrest Mortality in Alicante Using Clustering Algorithms

Aim

To determine the mortality rate of out-of-hospital cardiac arrest in Alicante (Spain) and identify the factors independently associated with death before hospital arrival.

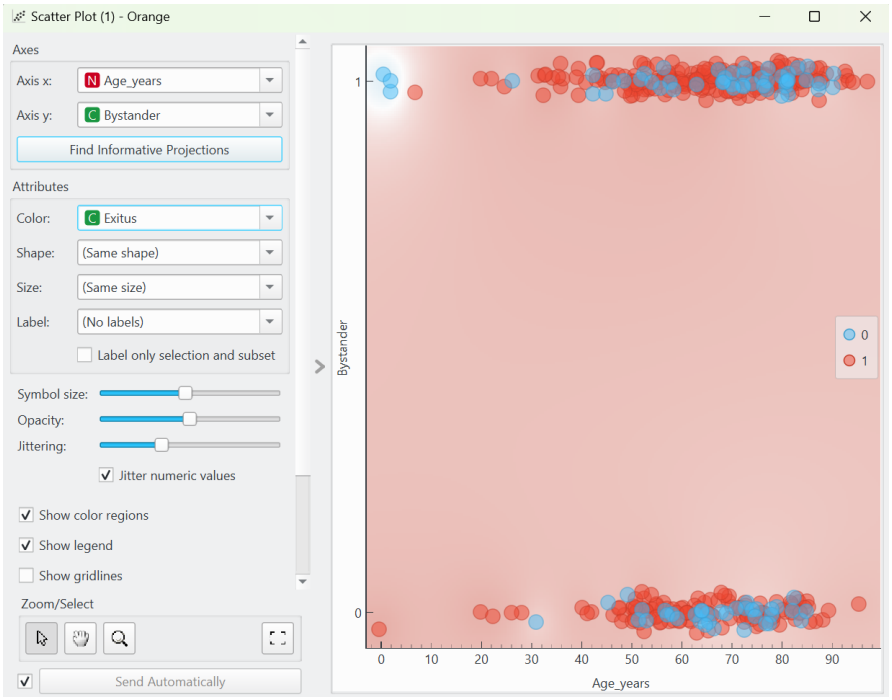
1. K-means Clustering:

The best silhouette score (0.432) was achieved with 2 clusters. This indicates a moderately good clustering structure, suggesting the data naturally separates into two distinguishable groups.



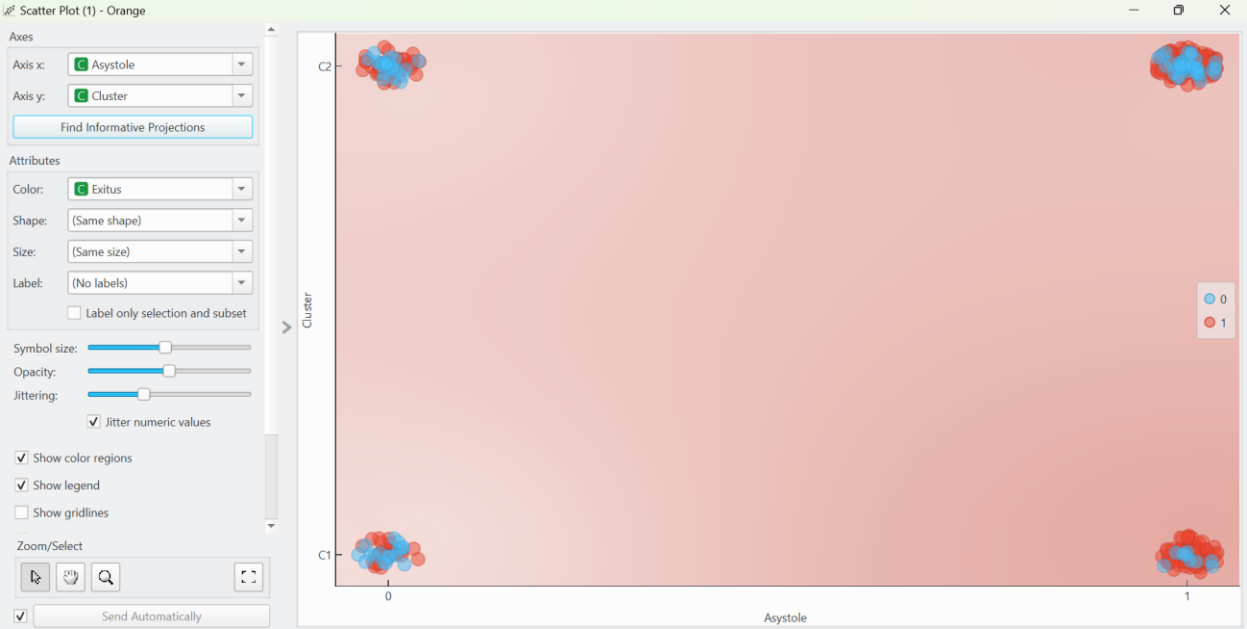
Scatter Plot: Age vs Bystander CPR (Colored by Exitus):

Absence of bystander CPR is associated with a higher likelihood of death. However, even with bystander CPR, mortality remains high, indicating that other factors (e.g., EMS delay, functional status, rhythm) are also critical. Bystander CPR may improve chances of survival, but not sufficient alone to ensure survival, especially in older patients or with delayed EMS.

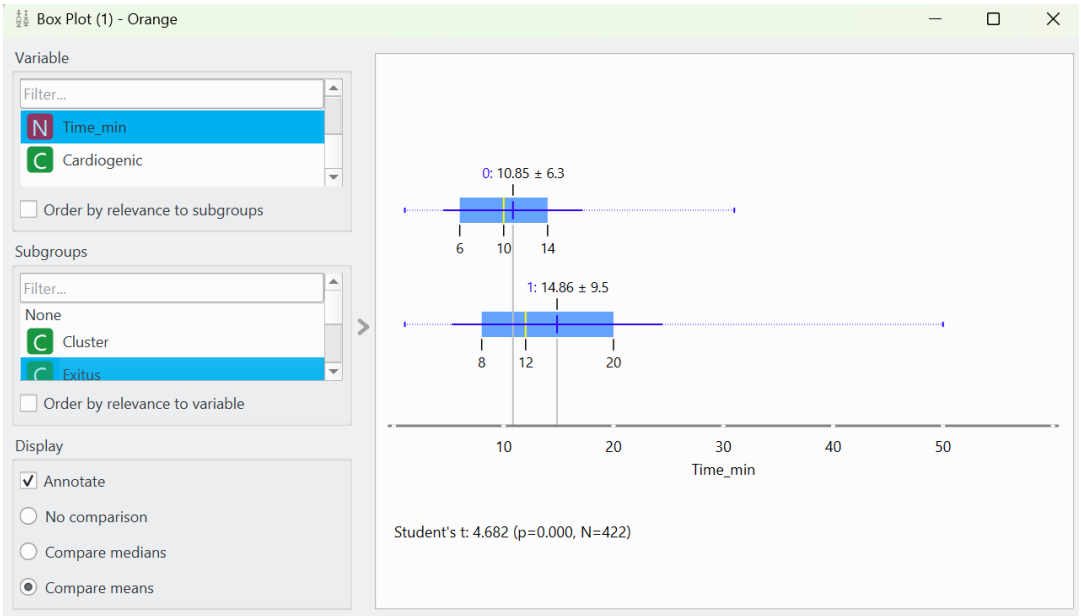


Scatter Plot: Asystole vs Cluster (Colored by Exitus):

Asystole is more common in Cluster C2, suggesting this cluster groups patients with more severe cardiac arrest presentations. Mortality is higher among patients with asystole. The K-Means algorithm partially captured clinical severity by grouping non-asystole (C1) and mixed but asystole-dominant (C2) cases. However, the clusters are not cleanly separated by survival outcome, indicating limited clustering utility for predicting mortality alone.



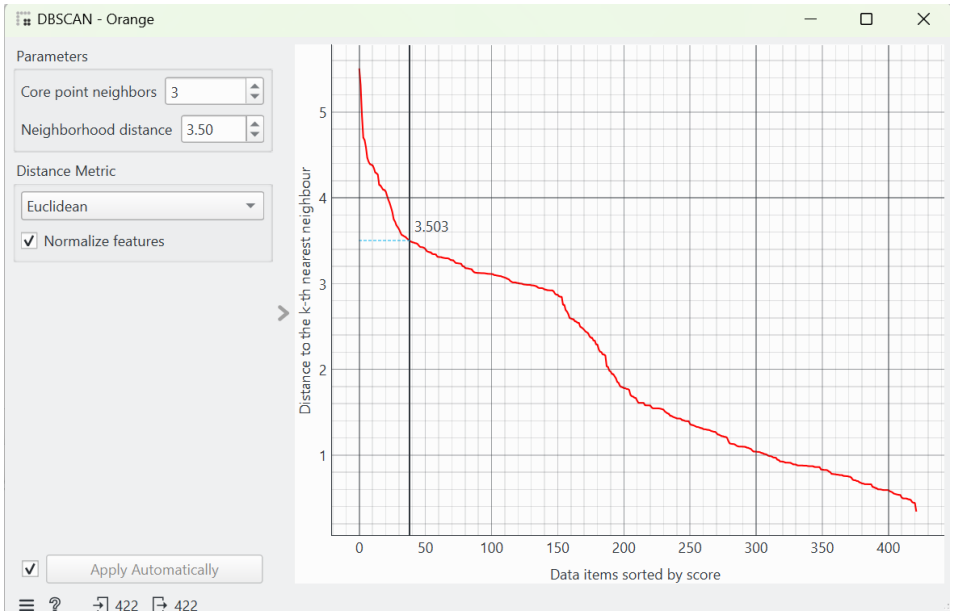
Box Plot: EMS Arrival Time (Time_min) vs Exitus:



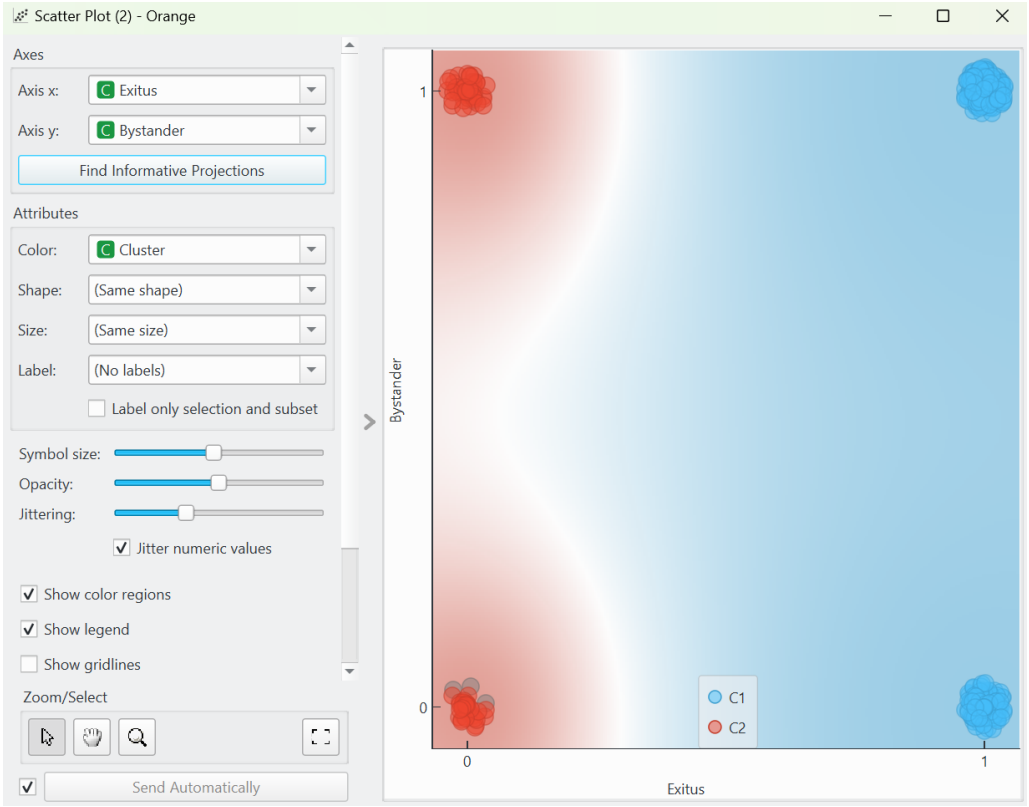
Patients who died before reaching the hospital experienced significantly longer EMS response times compared to those who survived. This difference is statistically significant, indicating that delay in EMS arrival is strongly associated with increased mortality. This supports the need for rapid emergency response as a critical factor in improving out-of-hospital cardiac arrest survival.

2. DBSCAN:

DBSCAN identified 2 dense clusters based on patients' survival patterns. The elbow of the curve appears around ε = 3.5, confirming the selected neighborhood distance is appropriate. This value separates dense clusters from noise/outliers effectively.



Scatter Plot: Exitus vs Bystander (Color by Cluster):



Cluster	Exitus Pattern	Bystander Presence	Interpretation
C1	Mostly survived (Exitus = 0)	Mostly with bystander present (Bystander = 1)	Cluster C1 indicates better survival outcomes when bystander is present.
C2	Mostly not survived (Exitus = 1)	Mostly without bystander (Bystander = 0)	Cluster C2 shows poorer outcomes in absence of bystander intervention.

Cluster	Age Distribution	Exitus Outcome	Interpretation
C1	Mixed age (mostly younger)	More 0 (survived)	Survival higher in **younger patients** (Cluster C1).
C2	Older age group	More 1 (not survived)	Higher mortality in **older patients** (Cluster C2).

Bystander presence is strongly associated with survival.

MDS Plot:

The MDS plot effectively shows the spatial separation of two DBSCAN clusters.

1. Cluster C1 (blue) is more concentrated with Exitus = 1 (non-survivors).
2. Cluster C2 (red) has a higher density of Exitus = 0 (survivors).

The Kruskal Stress = 0.011 indicates that the 2D projection accurately preserves the structure of high-dimensional data.

FreeViz Plot:

Feature	Interpretation
Functional_status	Strongly points away from the center, plays an important role in separation.
Time_min	Strong directional vector: more time is associated with higher mortality.
Endotracheal_intub	Moderate separation impact, may relate to survival or severity.
Cardiac_arrest_at_home	Contributes mildly to separation, direction indicates higher mortality.
Asystole	Points toward red area, related to negative outcomes.
Bystander	Direction shows influence, but less prominent.

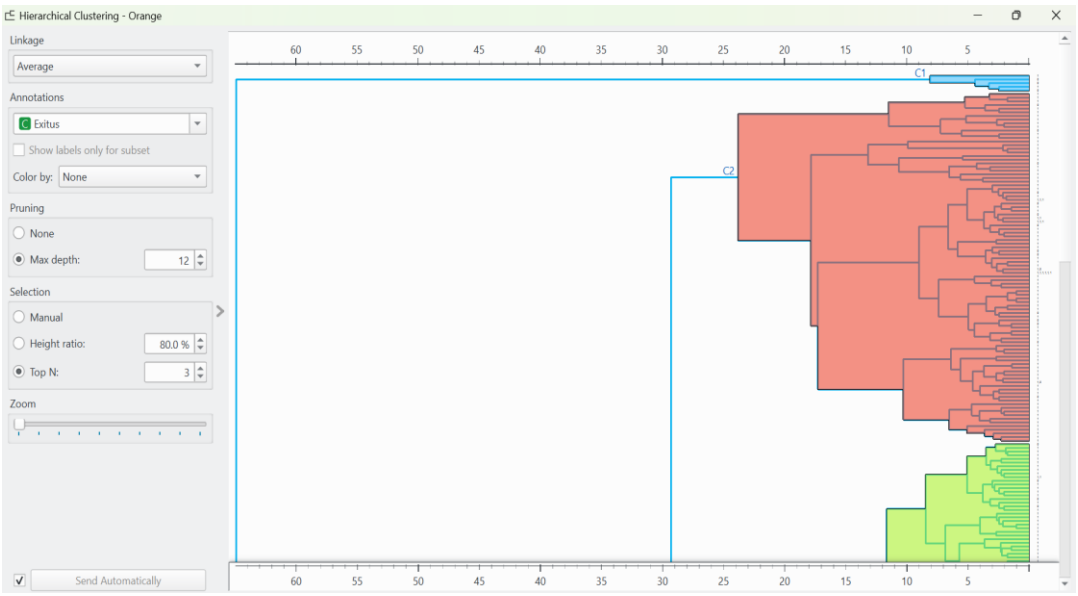
Age_years & Cardiogenic	Minor contributions as they are shorter and less distinct in direction.

Conclusion from FreeViz

- **Functional_status** and **Time_min** are the strongest discriminators of outcome (Exitus).
- Patients with **asystole**, **longer response times**, and **poor functional status** tend to have worse outcomes.
- The red and blue points overlap significantly, features alone do not perfectly distinguish survival, but certain variables are clearly more influential.

3. Hierarchical Clustering:

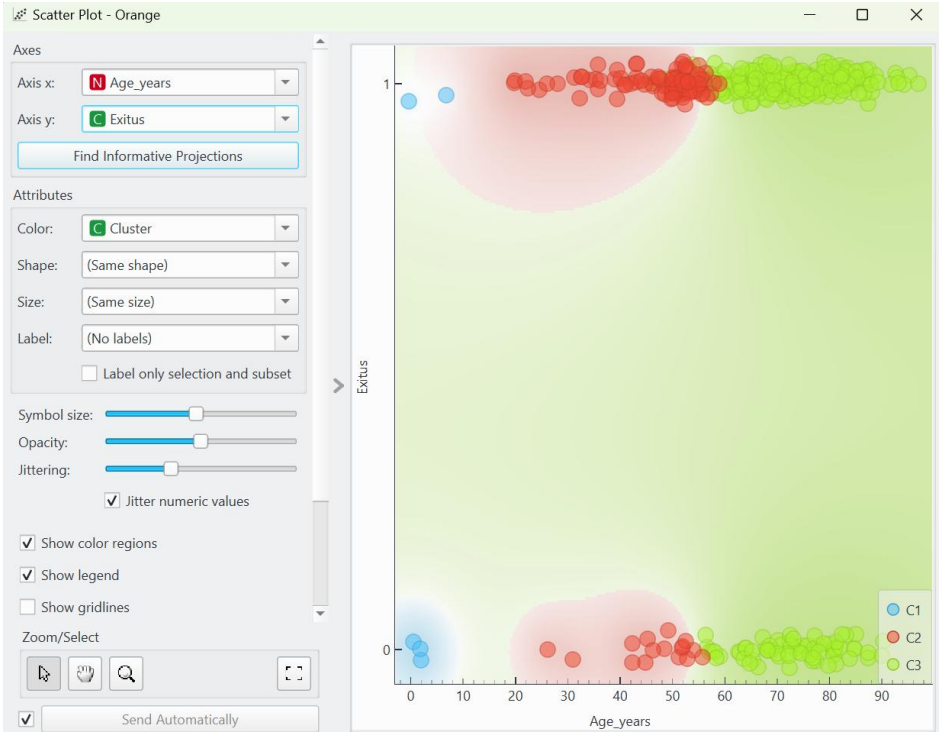
The dendrogram shows that the data has been successfully grouped into three main clusters (C1, C2, C3). The height of the linkages represents the dissimilarity between clusters: larger vertical lines mean more distinct subgroups. The cut at Top N = 3 indicates that this was a natural division based on the similarity of cases (patients) across all features. The depth and branching suggest how patient profiles diverge, revealing underlying structure in the cardiac arrest data.



Hierarchical clustering successfully divided the cardiac arrest patients into three distinct clusters:

- **C1** captured younger survivors.
- **C2** identified a high-risk group with poor outcomes.
- **C3** clustered older individuals with better-than-expected survival.

Scatter Plot:



Cluster	Age Pattern	Exitus Pattern	Interpretation
C1 (Blue)	Mostly **young** patients	Majority are **alive (Exitus = 0)**	This group reflects younger patients with better survival outcomes.
C2 (Red)	Middle-aged group	Predominantly **dead (Exitus = 1)**	This may represent high-risk patients with unfavorable outcomes.
C3 (Green)	Mostly **older** patients	Many are **alive (Exitus = 0)**	Older groups possibly benefited from timely care or lower-risk factors.