

Ejercicio 1

Los valores de los atributos parecen tener distribuciones aproximadamente uniformes, sin outliers.

Los resultados del k-means con k=4 y max runs=10 fueron los siguientes (sin normalizar):

Number of Clusters: 4

Distance Measure: Mixed Euclidean Distance

Average Cluster Distance: 148.949

Davies-Bouldin Index: 0.512

Cluster 0

118

Average Distance: 146.534

Colesterol is on average **22.61%** larger, **Peso** is on average **17.54%** larger, **Sexo** is on average **14.22%** smaller

Cluster 1

154

Average Distance: 162.484

Peso is on average **83.89%** larger, **Colesterol** is on average **70.85%** larger, **Sexo** is on average **15.03%** larger

Cluster 2

140

Average Distance: 140.096

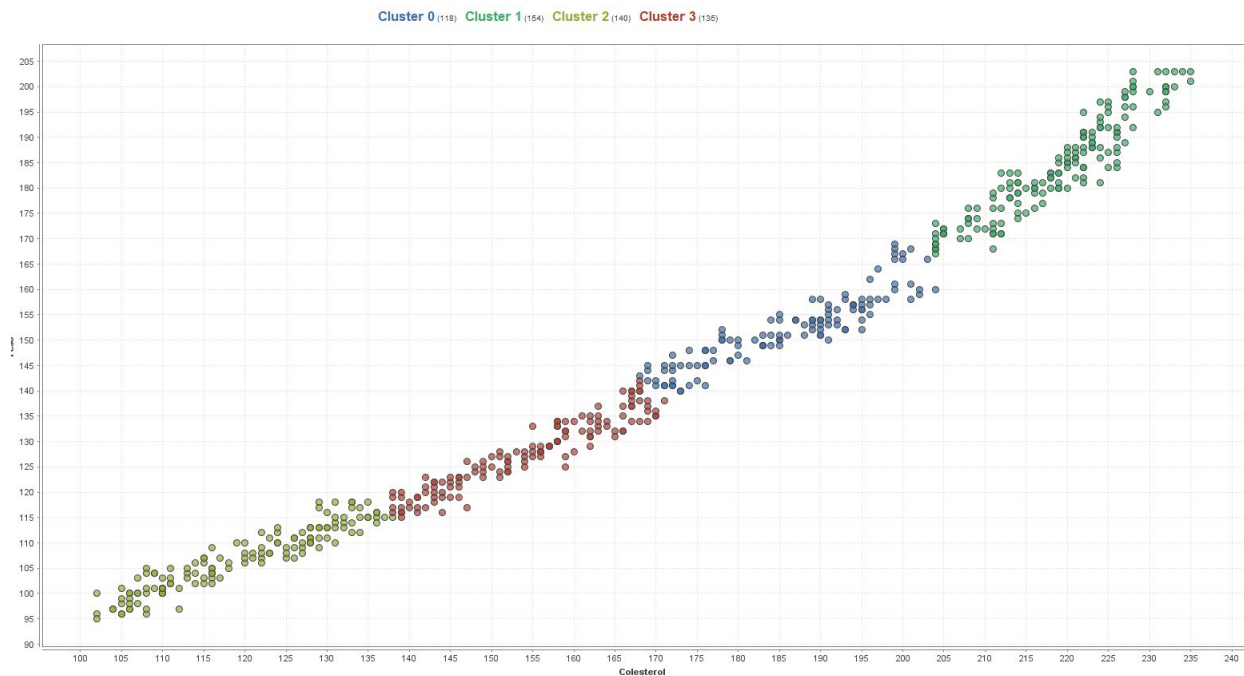
Peso is on average **75.60%** smaller, **Colesterol** is on average **74.38%** smaller, **Sexo** is on average **5.67%** larger

Cluster 3

135

Average Distance: 144.803

Peso is on average **32.62%** smaller, **Colesterol** is on average **23.45%** smaller, **Sexo** is on average **10.60%** smaller



Ejercicio 2

3 experimentos con diferentes parametros para el DBSCAN:

Parámetros por defecto:

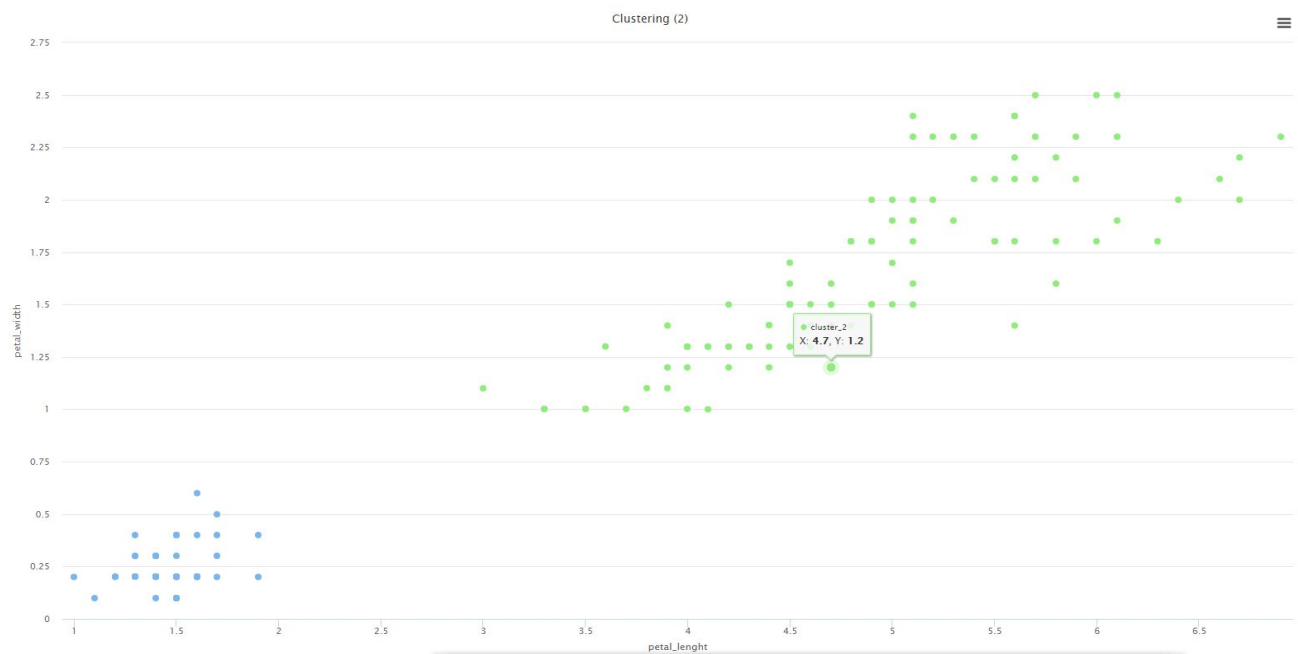
PerformanceVector:

Avg. within cluster distance: -76.595

Avg. within cluster distance for cluster 0: unknown

Avg. within cluster distance for cluster 1: -11.979

Avg. within cluster distance for cluster 2: -108.902



Resultado con epsilon=0.1:

PerformanceVector:

Avg. within cluster distance: -145.440

Avg. within cluster distance for cluster 0: -201.741

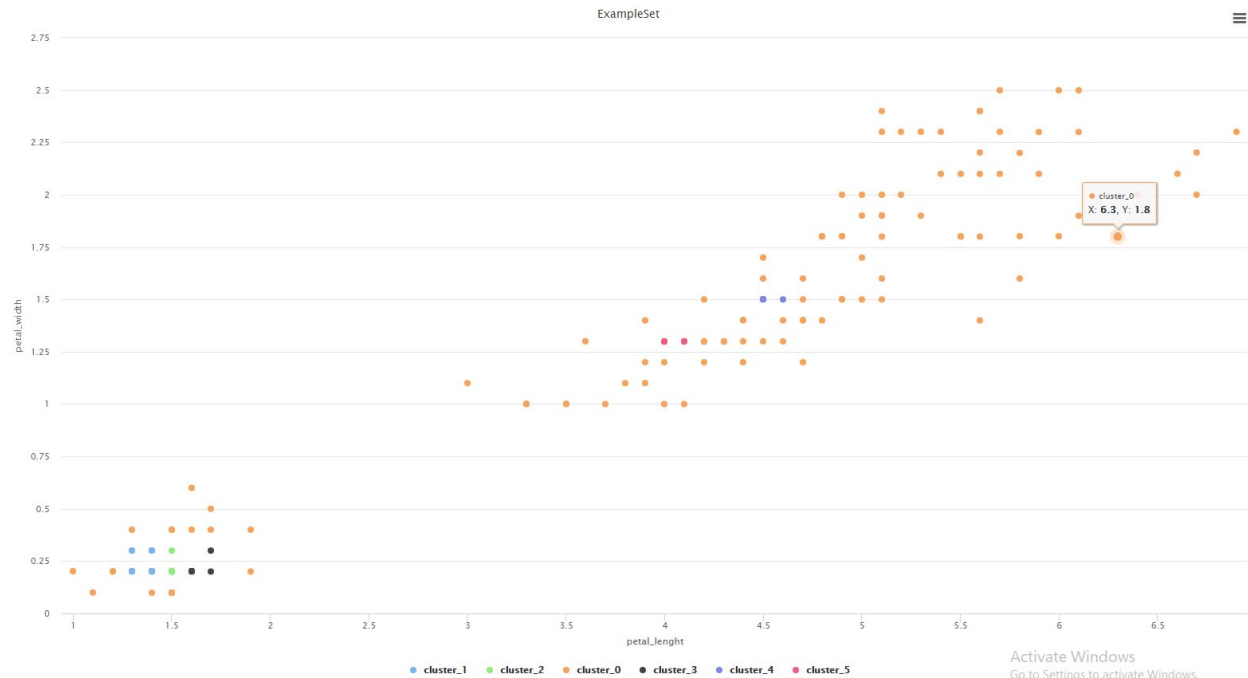
Avg. within cluster distance for cluster 1: -1.289

Avg. within cluster distance for cluster 2: -0.171

Avg. within cluster distance for cluster 3: -0.373

Avg. within cluster distance for cluster 4: -0.167

Avg. within cluster distance for cluster 5: -0.240



El epsilon pequeño hizo que los clusters sean muy pequeños y solo incluyan las observaciones que se encuentran extremadamente cerca entre ellas, dejando afuera como ruido a la gran mayoría de las observaciones del dataset.

Epsilon=0.2 y MinPoints=3:

PerformanceVector:

Avg. within cluster distance: -21.597

Avg. within cluster distance for cluster 0: -25.962

Avg. within cluster distance for cluster 1: -11.979

Avg. within cluster distance for cluster 2: -35.113

Avg. within cluster distance for cluster 3: -6.505

Avg. within cluster distance for cluster 4: -0.133

Avg. within cluster distance for cluster 5: -0.322

