To perform a hierarchical analytical process, we decided to select the same dataset employed for the DBScan application process.

First of all, we selected the DBScan’s subset of numeric attributes by classifying all the records included in that dataset. We represented the hierarchy graphically employing dendrograms: showing all the instances on the abscissa axis, and the distance between the clusters on the ordinates. We eventually grouped the informations again to provide an easier understanding.

Depending on the methods applied, we can see how hierarchical clustering can yield better results than the density based approach.

As expected, the “single link” method resulted to be the one who provides the worst output. As a matter of fact, its clusters tend to agglomerate very quickly with a very close range between each other. This graphical attitude, just like we saw in DBScan in the previous analysis, results to be due to the continuity of the records and their proximity.

We can conclude by saying that both metrics, euclidean and cosine, provide an unsatisfactory amount of informations.pasted-image.tiffpasted-image.tiff

(sx single link euclidean - DX SINGLE LINK COSINE)

Going deeper with the inspection, we found out that, unlike the criteria applied in the previous step, the Ward, Complete and Average methods are able to show better and appreciable results, capable of providing more informations.

Starting from the Ward’s method -which tends to minimise the variance within the clusters (SSW)-, we can say that it looks like showing results close to the K-means’s ones: grouping datas in three clusters of different sizes. Subsequently we analysed the Complete and Average methods, coming out with a very similar output: both of them partitioned the dataset and grouped the values into six clusters, showing very similar shape features.

To be precise, we can specify that employing the Complete method, it is possible to define six clusters that choosing the maximum distance criteria causes the clusters to be partitions.pasted-image.tiffpasted-image.tiff

pasted-image.tiff