Hierarchical Clustering N.C. = 4



Now, we can try to extract some basic features description for each cluster.

**Cluster 1** and **Cluster 2** are from **similar social class**, but **distinguished** by **gender**, with Cluster 1 mostly composed of women (87%) and Cluster 2 predominantly consisting of men (97.7%).

Their **education** level is **similar**, which the vast majority have only secondary education (84% and 86% respectively). Even though their **income has some discrepancy** (152k vs. 185k), this could be result of gender inequality by their employment situation.

And in the target level, the **Cluster 1 tends to overdue the payment (12,78%) slightly more than the Cluster 2 (3%).**

In the **Cluster 3**, the **gender** is more **homogeneously distributed**. Even though there is 62.9% of female (consequence of unbalanced gender distribution), the relative frequency is that 18.28% of female and 19.64% belongs to this cluster.

We believe that this cluster represents the grouping of individuals who belongs to a **higher social-economic** atmosphere, with higher income (210.535) and education (82.19%). This also can be observed by their job type, where core staff represents 22.7% of all the occupations. On the other hand, they have a more stable familiar situation (85% married, with higher nº child mean). In the target level, the percentage of overdue is 6.5%.

In the **Cluster 4**, there is a grouping of retired people, whose income is much lower (133k), their jobs status are exclusively pensioners. The average age of this cluster is 59 years old, and the percentage of widow is significantly higher than another cluster. In the target level, the percentage of overdue is 6.3%.

 