

Problem n.2

The file `streaming.txt` contains the data for 225 users of a music streaming service. The variable *minutes* contains the average minutes per day in 2021 for each user, while the variables *artists* the average number of single artists listened per day.

- a) Perform a cluster analysis of the users by using a hierarchical clustering method (Euclidean distance and single linkage). Report the plot of the dendrogram and the number of clusters you deem appropriate to represent the data.
- b) Report the centroids of the clusters, their size and the cophenetic coefficient.
- c) Provide Bonferroni intervals (global level 95%) for the mean of minutes and artists, within each of the clusters identified at point (a). Introduce and verify the appropriate assumptions.

Upload your results here:

<https://forms.office.com/Pages/ResponsePage.aspx?id=K3EXCvNtXUKAjjCd8ope6-9ASOGWf2lHjvGX24HiqFVUNUZCQkhIV1JPMEVDVUxMWUtYWUtOM004NS4u>