Politecnico di Milano
Scuola di Ingegneria Industriale e dell'Informazione
Applied Statistics
July 12th, 2022

Problem n.2

The Demogorgons are attacking students in the city of Hawkins. Lucas and Mike collected 97 positions of sighting (latitude and longitude) in the file `demogorgons.txt` and are interested in characterizing the most dangerous areas of the city.

a) Identify possible clusters within the data using a hierarchical clustering algorithm (Euclidean distance, average linkage). Provide the plot of the dendrogram and qualitatively identify the optimal number of clusters.

b) Formulate a MANOVA model for the latitude and longitude of the sightings as a function of the clustering membership. Report the formulation of the model, the estimates of the parameters and verify the assumptions. Is there statistical evidence to state that the membership to a cluster has an effect on the mean positions?

c) Report the centers of each cluster. For each cluster, provide a confidence region at level 95% for the mean of the longitude and latitude and plot them on a scatterplot with the data.

Upload your results here:
`https://forms.office.com/Pages/ResponsePage.aspx?id=K3EXCvNtXUKAjjCd8ope6-9AS0GWf2lHjvGX24HiqFVUODhIWTdWV1NBOVJYMU9CWlMxTTJVOU1JNi4u`