

Task 1

Data sets, "Data1.csv", "Data2.csv", "Data3.csv", "Data4.csv", "Data5.csv", "Data6.csv", "Data7.csv", "Data8.csv" contain the data points and their respective class information.

For each of the datasets follow the below steps

1. Use K-means and hierarchical clustering methods to generate clusters.
2. Evaluate the performance of the clustering algorithm using external validation metrics.
3. Plot (2D or 3D) the data points for each dataset and color them according to the original class.
4. Plot (2D or 3D) the data points for each dataset and color them according to the class allocated by the clustering algorithm.

Task 2

The world indicators dataset compares different countries based on selected attributes.

1. Use K-means and hierarchical clustering methods to group similar countries together.
2. Use Internal validation metrics to report the cluster quality.
3. Report the best clustering solution. Give a detailed list of all the groups and the countries included within the groups.
4. Generate three different scatter plots of your choice and color the data points according to the group.