About
This assignment is a part of screening process for candidates who applied for a role at our organization.

Data
You are being provided with the required information in dataset 'Dataset.csv'.

Background
The assignment focuses on checking the quality of clustering that has been performed. You are being provided with one of the required concept below, i.e, **Davies-Bouldin Index**. You can use the dataset 'Dataset.csv' to exploit your solution.

The davies-bouldin index has to be calculated for any value of n as follows:

$$DB = \frac{1}{n_c}\sum_{i=1}^{n_c} R_i, \text{ where}$$

$$R_i = \max_{j=1...nc, i\neq j}\left(R_{ij}\right), \; i = 1...n_c$$

$$R_{ij} = \frac{s_i + s_j}{d_{ij}}$$

$$d_{ij} = d\left(v_i, v_j\right), \; s_i = \frac{1}{\|c_i\|}\sum_{x\in c_i} d\left(x, v_i\right)$$

Where:

- d(x,y) is the Euclidean distance between x and y.
- ci is the cluster i.
- vi is the centroid of cluster ci
- $\|c_i\|$ refers to the norm of ci

Also, use **Silhouette Score** and **Calinski-Harabasz** index (CH) to check the Quality of Clustering.