Cluster Analysis on Algerian economic data

Description:

The project aims to perform a cluster analysis on Algerian economic data based on a set of economic indicators. The study should mainly focus on two aspects:

- data visualization
- and prediction/classification by clustering approach.

Student are asked to realize the following parts.

1- Data acquisition

The source is the World Bank website. Use the link: http://wdi.worldbank.org/tables

Students have to download files and construct two datasets:

- The first dataset contains the following economic indicators
    - From the **Data World tab**, download the **size of the economy** file. The 3 first columns (**population, surface area, population density**) are not used as economic indicators.
    - From the **economy tab**, add attributes that are included in: **growth of output, structure of output, structure of manufacturing and structure of merchandise exports** files.
    - From **people tab,** add **unemployment** data.
    - From **poverty and shared prosperity,** include **poverty data at international line** (2 files).

- The second dataset concerns the different resources (human, natural) and potentials. The indicators used in this dataset are:
    - **Population, surface area** and **population density** from **size of the economy** file.
    - From **environment tab**, add data from:
      rural environment and land use, **agricultural inputs**, **fresh water, energy production and use, contribution of natural resources to gross domestic product** files**.
    - From **people tab,** add **population dynamics, labor force structure, participation in education** and **education completion and outcomes** data**.

    Students may incorporate any additional data that they deem useful.

2- Data visualization

Students have to propose and produce data graphics of the two datasets.

3- Clustering

Students have to apply cluster analysis algorithms. Produce at least two clustering results using the datasets. Clustering results should also be presented with graphics if possible.

4- Reporting

Based on the previous results, students should prepare an analysis report illustrating a:
- Comparison between results of clustering.
- Prediction of economic classification based on resources and potentials clustering.