

În fișierul *CalitateaAeruluiTari.csv* sunt prezentate informații privind calitatea aerului la nivel de țară. Indicatorii care măsoară calitatea aerului sunt: *Air_quality_Carbon_Monoxide*, *Air_quality_Ozone*, *Air_quality_Nitrogen_dioxide*, *Air_quality_Sulphur_dioxide*, *Air_quality_PM2.5*, *Air_quality_PM10*. Aceștia cuantifică prezența diverselor particule în aer (ozon, sulf etc.). Celelalte variabile sunt: *CountryId* - Codul de țară, *Country* - numele țării.

În fișierul *CoduriTari.csv* se află codificări ale țărilor și împărțirea acestora pe continente. *CountryId* este câmpul de legătură.

A. Să se implementeze următoarele cerințe:

1. Salvarea în fișierul *Cerinta1.csv* a țării în care se înregistrează cea mai mare valoare pentru fiecare indicator. Se va salva pentru fiecare indicator numele indicatorului și țara în care s-a înregistrat cea mai mare valoare. (**1 punct**)

Exemplu:

```
Indicator,Country
Air_quality_Carbon_Monoxide,Indonesia
...
...
```

Criteriul de acordare a punctajului: vizualizarea fișierului output și a codului sursă

2. Salvarea în fișierul *Cerinta2.csv* a țărilor în care s-au înregistrat cele mai mari valori la nivel de continent, pentru fiecare indicator. Se va salva numele continentului și țările în care s-au obținut cele mai mari valori pentru fiecare continent. (**2 puncte**)

Exemplu:

```
Continent,Air_quality_Carbon_Monoxide,Air_quality_Ozone,Air_quality_Nitrogen_dioxide,Air_quality_Sulphur_dioxide,Air_quality_PM2.5,Air_quality_PM10
Africa,Ethiopia,Morocco,Algeria,South Africa,Burkina Faso,Mauritania
...
...
```

Criteriul de acordare a punctajului: vizualizarea fișierului output și a codului sursă

B. Să se efectueze analiza de clusteri prin metoda Ward, pentru indicatorii calității aerului și să se furnizeze următoarele rezultate:

1. Matricea ierarhie cu informații privind joncțiunile făcute. Pentru fiecare joncțiune se va specifica clusterii intrați în joncțiune, distanța dintre cei doi clusteri și numărul de instanțe în clusterul nou format. Matricea va fi afișată la consolă. (**1 punct**)

Criteriul de acordare a punctajului: vizualizarea outputului și a codului sursă

2. Componența partiției optimale. Pentru fiecare instanță se determină clusterul de care aparține. Partiția se va salva în fișierul *popt.csv*. (**2 puncte**)

Criteriul de acordare a punctajului: vizualizarea fișierului output și a codului sursă

3. Graficele histogramă pentru una din variabile (la alegere - se va inițializa în cod). (**2 puncte**)

Criteriul de acordare a punctajului: vizualizarea graficelor și a codului sursă

C. Într-un model de analiză factorială cu 5 variabile observate, X₁, ..., X₅ s-au obținut coeficienții de corelație din fișierul *g21_1.csv* și coeficienții de corelație parțială din fișierul *g21_2.csv*. Să se calculeze și să se afișeze la consolă indexul KMO global. (1 punct**)**

Criteriul de acordare a punctajului: vizualizarea rezultatului în consolă și urmărirea codului sursă.