

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 coeficienților de variație pentru fiecare indicator. Se va salva pentru fiecare indicator numele indicatorului și coeficientul de variație. (**1 punct**)

Exemplu:

```
Indicator,CV
Air_quality_Carbon_Monoxide,1.383742005727142
...
```

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

2. Să se determine pentru fiecare continent indicatorul cu cel mai mare coeficient de variație. Datele vor fi salvate în fișierul *Cerinta2.csv*. Pentru fiecare continent se va salva numele continentului, indicatorul cu coeficientul de variație cel mai mare și valoarea coeficientului de variație. (**2 puncte**)

Exemplu:

```
Continent,Indicator,CV
Africa,Air_quality_Sulphur_dioxide,1.5996428123139814
...
```

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

B. Să se efectueze analiza în componente principale standardizată pe indicatorii privind calitatea aerului și să se furnizeze următoarele rezultate:

1. Variantele componentelor principale. Variantele vor fi afișate la consolă. (**1 punct**)

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

1. Scorurile asociate instanțelor. Scorurile vor fi salvate în fișierul *scoruri.csv*. (**2 puncte**)

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

3. Graficul scorurilor în primele două axe principale. (**2 puncte**)

Criteriul de acordare a punctajului: vizualizarea graficului

C. Într-un model de analiză discriminantă liniară aplicat pe 100 de instanțe descrise prin 10 variabile predictor, s-au obținut următoarele valori proprii: [0.9, 0.05, 0.8, 0.75, 0.5, 0.3, 0.2]. Să se afișeze la consolă numărul minim de variabile discriminante necesare pentru a acoperi peste 90% din puterea de discriminare a variabilelor discriminante. (**1 punct**)

Criteriul de acordare a punctajului: vizualizarea outputului și codului sursă