

În fișierul *ElectricityProduction.csv* se află date privind structura producției de energie electrică la nivelul țărilor din Europa, pe tipuri de centrale. Variabile: *CountryCode* - codul de țară, *Country* - numele țării, de la variabila *coal* până la variabila *other* sunt procentele de energie electrică produsă prin centrale pe bază de cărbune, țiței, gaze naturale, combustibil bio, șamd până la alte tipuri de resurse. În fișierul *Emissions.csv* se află informații privind emisiile de particule la nivelul țărilor din Europa. Variabile: *CountryCode* - codul de țară, *Country* - numele țării, *AirEmiss* - emisii de carbon, *Sulphur* - emisii de sulf, *Nitrogen* - emisii de nitrogen, *Ammonia* - emisii de amoniac, *NonMeth* - compuși organici volatili nemetanici, *Partic* - amestec de particule solide, *GreenGE* și *GreenGIE* - emisii de gaze cu efect de seră. Emisiile sunt măsurate în tone, cu excepția *GreenGE* și *GreenGIE* care sunt măsurate în mii tone. În fișierul *PopulatieEuropa.csv* se află împărțirea pe regiuni a țărilor din Europa și populația fiecărei țări. Regiunile sunt codificate: *EasternE*, *WesternE*,

A. Să se implementeze următoarele cerințe:

1. Salvarea în fișierul *Cerinta1.csv* a emisiilor totale de particule, la nivel de țară, **exprimate în tone**. Se va salva codul de țară, numele țării și emisiile totale. (**1 punct**)

Exemplu

CountryCode,Country,Emisii_total_tone

AUT,Austria,2453643.49

BEL,Belgium,2419287.04

...

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

2. Salvarea în fișierul *Cerinta2.csv* a emisiilor de particule la 100000 locuitori, la nivel de regiune, pe fiecare tip de particule (tone/100000loc). Pentru fiecare regiune se salvează codul regiunii și emisiile de particule la 100000 de locuitori. (**2 puncte**)

Exemplu

Region,AirEmiss,Sulphur,Nitrogen,Ammonia,NonMeth,Partic,GreenGE_tone,GreenGIE_tone

British,2416.49,4.018,6.07,0.024,1.926,0.872,7291.122,2685.739

...

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

B. Să se studieze legătura dintre structura producției de energie electrică din fișierul *ElectricityProduction.csv* și emisiile de particule din fișierul *Emissions.csv*, utilizând analiza canonică. Vor fi folosite variabilele de la *coal* la *other* pentru primul set și variabilele de la *AirEmiss* la *GreenGIE* pentru setul 2.

Cerințe:

1. Calculul scorurilor canonice pentru cele două seturi și salvarea lor în fișierele *z.csv* și *u.csv* (**2 puncte**)

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

2. Calculul corelațiilor canonice și salvarea lor în fișierul *r.csv*. (**1 punct**)

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

3. Aplicarea testului Bartlett pentru determinarea semnificației statistice a perechilor canonice. Se vor afișa la consolă rezultatele testului, valorile pvalues și numărul de rădăcini cu un prag de semnificație mai mic decât 0.01 (**2 puncte**)

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

C. Într-un model de analiză în componente principale aplicat pe un set de date cu trei variabile observate s-a obținut următoarea matrice a vectorilor proprii așezați pe coloane:

$$a = \begin{bmatrix} 0.4 & -0.25 & 0.8 \\ 0.5 & 0.75 & 0.6 \\ 0.1 & 0.75 & 0.5 \end{bmatrix}$$

Determinați și afișați valoarea cosinus a instanței $x = [2, 4, 2]$ în raport cu axa celei de a treia componente.