

În fișierul *Diversitate.csv* sunt prezențați indicii de diversitate economică la nivel de localitate pe ani. Variabile: *Siruta* - codul Siruta al localității, *Localitate* - denumirea localității
2008, 2009, ..., 2021 - anii pentru care sunt prezențați indicii de diversitate.
În fișierul *Coduri_Localitati.csv* se află codificarea localităților și împărțirea acestora pe județe. Variabile: *Siruta* - codul Siruta, *Localitate* - denumirea localității, *Judet* - indicativul de județ. *Siruta* este câmpul de legătură.

A. Să se implementeze următoarele cerințe:

1. Să se calculeze și să se salveze în fișierul *Cerinta1.csv*, valoarea medie a indicelui de diversitate (media pe ani) la nivel de localitate. Salvarea se va face în ordinea descrescătoare a indicelui de diversitate. Se va salva codul Siruta, denumirea localității și indicele mediu de diversitate. (**1 punct**)

Exemplu:

Siruta,Localitate,Diversitate
179132,Municipiul Bucuresti,0.7748850000000002
95060,Municipiul Iasi,0.7460085714285715
54975,Municipiul Cluj-Napoca,0.7419014285714285
...

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

2. Să se determine pentru fiecare județ și fiecare an, numărul de localități care au diversitatea 0. Rezultatul va fi salvat în fișierul *Cerinta2.csv*. Se va salva indicativul de județ și numărul de localități cu diversitate 0 pentru fiecare an. (**2 puncte**)

Exemplu:

Judet,2008,2009,2010,2011,2012,2013,2014,2015,2016,2017,2018,2019,2020,2021
ab,3,5,3,4,2,2,1,2,1,1,2,3,1,2
...

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

B. Să se efectueze analiza de clusteri prin metoda Ward, pentru indicii de diversitate 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. Graficul dendrogramă pentru partitia optimală. (**2 puncte**)

Criteriul de acordare a punctajului: vizualizarea graficului dendrogramă.

3. 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

C. Într-un model de analiză factorială cu 5 variabile observate standardizate și 2 factori comuni s-au obținut coeficienții factoriali *L* (*factor loadings*) din fișierul *g20.csv*. Să se afișeze la consolă care dintre factorii specifici are cea mai mare varianță. Se va afișa un număr de la 1 la 5. (**1 punct**)

Criteriul de acordare a punctajului: Urmărirea rezultatului și a codului care l-a generat.