Dezvoltare Software Pentru Analiza Datelor

Proiect Unu – Analiza numarului de elevi inscrisi in ciclul primar si gimnazial in perioada 2018 – 2021

Student: Palcu Andrei

Grupa 1088, Seria D, Anul 3

Proiectul propus face referire la analiza numarului de elevi inscrisi in ciclul primar si gimnazial intre anii 2018 si 2021 in toate judetele din Romania. In fisierul ,,elevi\_inscrisi.csv” se regasesc datele aferente numarului de elevi, precum si judetele. In fisierul ,,judete.csv” se afla codurile judetelor. Scopul acestei analize este de a aplica testele statistice aferente, dupa cum urmeaza:

1. **Ponderea elevilor inscrisi in perioada 2018 - 2021**

Potrivit datelor din fisierul ,,PonderiElevi.cvs”, procentul de elevi inscrisi este aproape constant pentru fiecare an in parte. In anul 2018 ponderea pentru fiecare judet este cuprinsa intre 0.25 si 0.26, in 2019 intre 0.24 si 0.25, in 2020 valoarea este de 0.24 si 2021 intre 0.23 si 0.25. Numerele mentionate anterior sunt aproximate.

1. **Entropia**

Entropia masoara gradul de incertitdine al unui eveniment anume. Informaţia este un mesaj care aduce o precizare într-o problemă cu un anumit grad scazut de incredere. Incertitudinea scade o dată cu apariţia informaţiei. Comform datelor din fisierul ,,Entropie.csv”, valoarea entropiei pentru fiecare judet in parte este de aproximativ 1.99.

1. **Indicele Simpson**

Este utilizat pentru masurarea diversitatii unui esantion (in cazul curent: numarul de elevi inscrisi in cele 42 de judete). Cu cat valoarea indicelul este mai apropiata de 1, diversitatea este mica, iar daca este apropiata de 0 diversitatea este mai mare. Din datele ce reies din fisierul ,,Simpson.csv” rezulta ca valoarea indicelui este 0.74, ceea ce indica faptul ca esantionul nu este atat de diversificat.

1. **Standardizare**

Rezultatele dupa aplicarea functiei de standardizare se regasesc in fisierul ,,Standardizare.csv”.

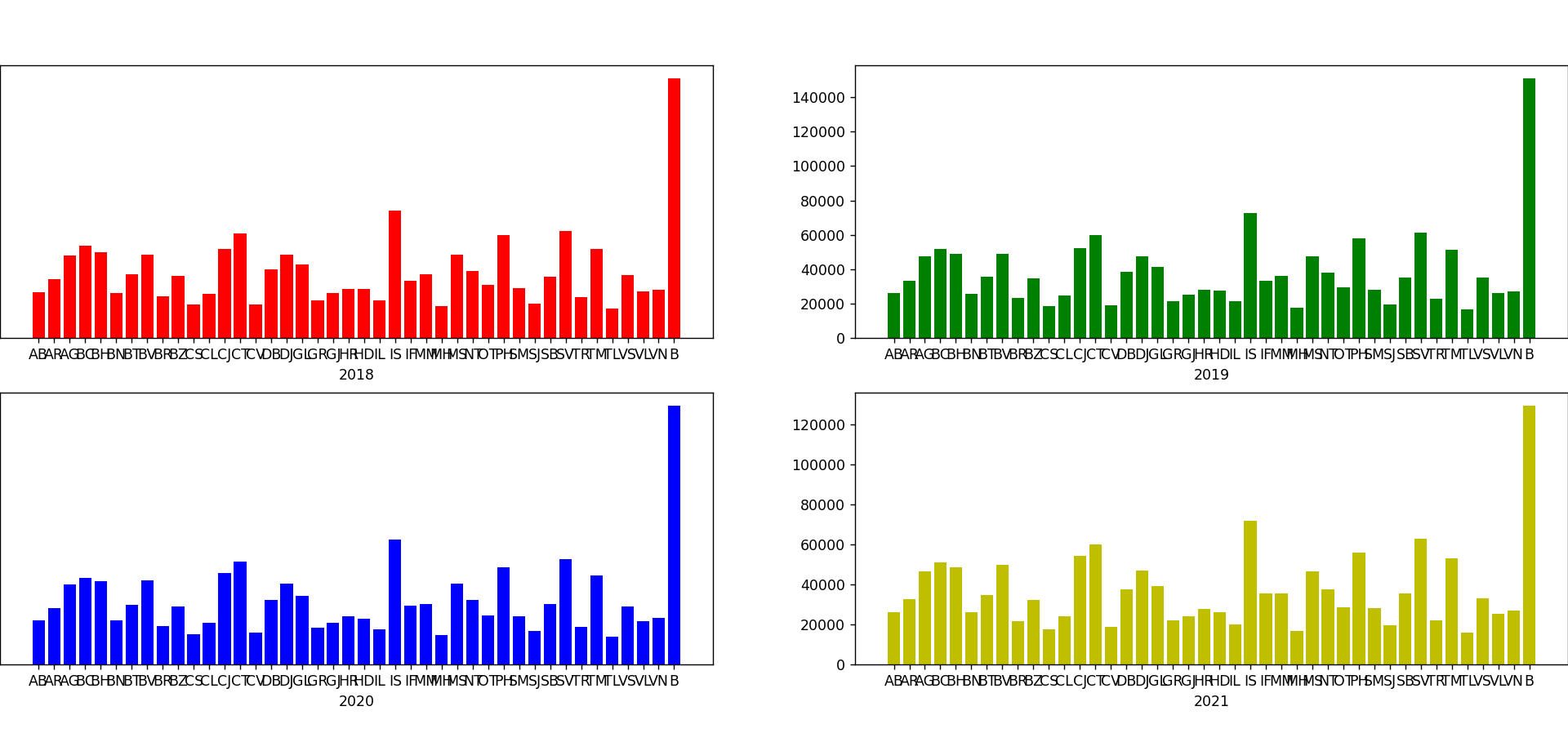
1. **Corelatia**

In cadrul fisierului ,,Corelatie.csv”, rezultatele sunt apropiate de 1, ceea ce reprezinta o legatura stransa intre acestea.

1. **Covarianta**

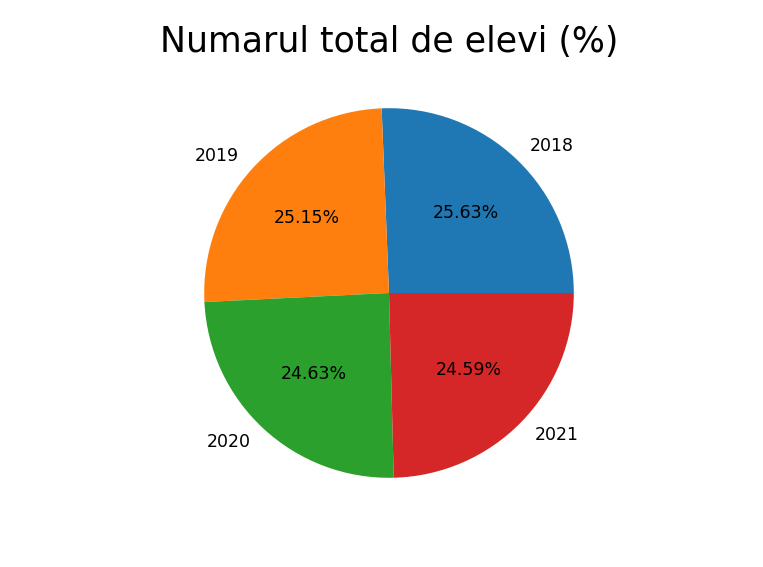
Rezultatele dupa aplicarea functiei de standardizare se regasesc in fisierul

,, Covarianta.csv”.

1. **Reprezentarea grafica a distributiei de elevi repartizati pe judete intre 2018 si 2021 (bar chart)**

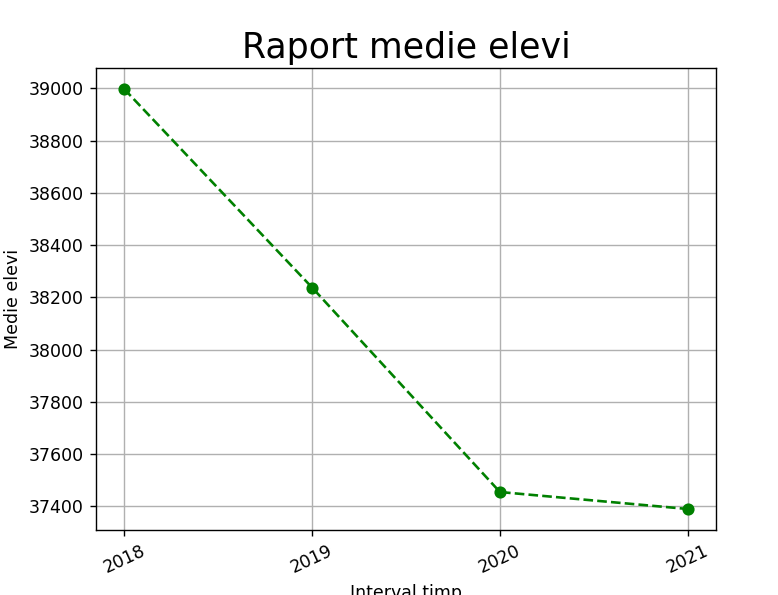
Comform graficului prezentat, repartitia elevilor pe parcursul celor 4 ani este aproximatv egala, insa cu mici diferente pentru anumite jutede. Municipiul Bucuresti este pe primul loc la numarul de inscrieri, urmat de Iasi, ambele orase mentinandu-si pozitiile timp de patru ani.

1. **Totalul elevilor inscrisi intre 2018 si 2021 (pie chart exprimat in procente)**



Din datele extrase din fisierul ,,total\_elevi.csv” se poate observa ca cei mai multi elevi inscrisi in ciclul primar si gimnazial provin din anul 2019. Anul 2018 a inregistrat cel mai mic rezultat, iar 2020 si 2021, chiar daca acesti ani au trecut printr-un context pandemic, au inregistrat 27.72%, respectiv 27.68% din totalul elevilor inscrisi in cei patru ani aflati in analiza.

1. **Numarul mediu de elevi inscrisi in perioada 2018-2021**

****

Din graficul propus rezulta ca in medie cei mai multi elevi au fost incrisi tot in anul 2019, o crestere semnificativa fata de anul precedent, urmand apoi o usoara scarede in anii urmatori. Motivul acestei cresteri poate fi justificat prin cresterea nivelului de trai si dezvoltarii infrastructurii, ceea ce a permis accesul multor copii la educatie.