

## Proiect Analiza Datelor

Realizați un proiect la disciplina *Analiza Datelor* care are ca scop recapitularea cunoștințelor referitoare la analiză exploratorie a datelor utilizând funcții speciale pentru reprezentări grafice și teste statistice, analiza variației și analiza în componente principale. Proiectul va fi trimis la adresa [tinelle@yahoo.com](mailto:tinelle@yahoo.com) până la data de **10 ianuarie 2020**. Proiectul va fi salvat *nume\_prenume.pdf*, iar la mesajul mail-ului se va preciza numele și prenumele studentului, specializare, grupa.

Proiectul va avea următoarea structură:

- A. *Formularea problemei*
- B. *Metode și tehnici de soluționare a problemei* – noțiuni teoretice necesare abordării problemei enunțate (scurtă descriere).
- C. *Experimente efectuate* – prezentarea etapelor pentru atingerea obiectivelor (rezolvarea cerințelor)
  - 1. Încărcarea fișierului sursă – prezentare structura fișier (x variabile numerice, y variabile categoriale etc.).
  - 2. Analiza exploratorie a datelor - statistici descriptive ale variabilelor considerate. Discuție.
  - 3. Realizarea de grafice corespunzătoare variabilelor (minim 3 grafice) și interpretarea acestora.
  - 4. Verificarea corelației între variabilele considerate. Exemplu, interpretare rezultate.
  - 5. Teste statistice pentru compararea mediilor – exemplu, interpretare rezultate.
  - 6. Analiza variației (ANOVA) – exemplu, interpretare rezultate.
  - 7. Analiza în componente principale – exemplu, interpretare rezultate.
- D. *Concluzii generale*
- E. *Referințe bibliografice/webografie.*

Note:

- Fiecare student va alege un fișier sursă dintr-o bază de date publică precum UCI Machine Learning Repository.
- Proiectul va avea minim 5 pagini și va conține, pe lângă partea teoretică (descrișă succint, doar ce este necesar), experimentele realizate însoțite de explicații și capturi de ecran.