



# Sistem de recomandare universitate

Călugărițoiu Ion-Victor  
Rizea Adelina-Maria  
Albuică Sergiu-Mihail

**IBD 2025**



# Funcționalități principale

## Jupyter Notebooks

1. Dataset cleaning
2. Analiza rezultatelor de la Bacalaureat folosind algoritmi de ML - Random Forest Classifier, K-Means Clustering
3. Vizualizarea dataset-urilor de la admiterile de la universități - histograme, boxplots, scatterplots, bar/pie charts
4. Recomandarea unei facultăți relevante în funcție de profilul unui candidat

# Web App - Microsoft Azure

## Funcționalități:

- Introducere profil unui candidat la Bacalaureat (rezultate, profil, specializare, etc)
- Recomandare profil facultate în funcție de datele introduse
- Comparatie între profilul introdus și cele din dataset-ul rezultatelor de la Bacalaureat
- Vizualizare date relevante din rezultatele dataset-urilor admiterilor

Aplicația web și Notebook-ul Jupyter rulează pe Azure.

The screenshot displays a web application interface with two main sections. The left section is a form for entering a student's profile, and the right section displays the results of a recommendation algorithm.

**Form Fields (Left):**

- City: Bucharest
- Profile: Real
- Specializare: Matematica-Informatica
- Subiect ea (e.g. Limba română (REAL)): Limba română (REAL)
- Subiect ec (e.g. Matematică MATE-INFO): Matematică MATE-INFO
- Subiect ed (e.g. Fizică TEO): Informatică M. C/C++
- NOTA\_FINALA\_EA: 7.01
- NOTA\_FINALA\_EC: 9.25
- NOTA\_FINALA\_ED: 9.5
- PUNCTAJ DIGITALE: 8.2
- ORAL\_PMO: 9.25
- Medie: 8.58

**Buttons:**

- Get Recommendation

**Detailed Scores (Right):**

Subject	Score
Computer Science	8.94
Mathematics	8.19
Journalism	8.06
Architecture	7.15
Political Science	5.22
Technical	4.64
Geography	4.49
Physical Education	3.57
Psychology	2.83
Medicine	2.58
Arts	2.35
Law	1.83

**Universities near Bucharest that might interest you:**












- Best Global Universities for Computer Science in Romania
- The Faculty of Mathematics and Computer Science
- List of Universities to study Computer Science And ...
- Department of Computer Science – University of Bucharest
- Best universities for a degree in Computer Science?

## Seturi de date

- Rezultate examen Bacalaureat

 bac2019.csv	5 hours ago
 probe_profil.csv	5 hours ago

- Rezultate admiteri

 ARHITECTURA_2024.pdf	5 hours ago
 DREPT_2024.pdf	5 hours ago
 FIZICA_2023.pdf	5 hours ago
 GEOGRAFIE_2024.pdf	5 hours ago
 PSIHOLOGIE_2024.pdf	5 hours ago
 SNSPA_2024.pdf	5 hours ago
 SOCIOLOGIE_2024.pdf	5 hours ago
 UMFCO_2019.pdf	5 hours ago
 UNIBUC_CTI_INFO_2024.pdf	5 hours ago
 UNIBUC_MATE_2024.pdf	5 hours ago
 UPB_ACS_2014.pdf	5 hours ago

# Data integration & cleaning

- Ștergerea coloanelor goale
- Eliminarea coloanelor nefolositoare
- Îmbinarea datelor
- Standardizarea datelor într-un format convenabil

ARHITECTURA_clean_2024.csv	5 hours ago
DREPT_clean_2024.csv	5 hours ago
GEOGRAFIE_2024.csv	5 hours ago
JURNALISM_2024.csv	5 hours ago
PSIHLOGIE_2024.csv	5 hours ago
SNSPA_2024.csv	5 hours ago
UMFCD_2019.csv	5 hours ago
UNIBUC_CTL_INFO_2024.csv	5 hours ago
UNIBUC_MATE_2024.csv	5 hours ago
UPB_ACS_2014.csv	5 hours ago

Printing ARHITECTURA csv...

[9]:	Nr.	Dosar	Medie BAC	Prezenta	Nota Admitere	Medie Generala	Status	Observatii	Loc repartizat
0	1	0084	9.63	Prezent	Olimpic	Olimpic	Admis Buget	INSCRIS	B-1
1	2	0859	9.31	Prezent	9.04	9.12	Admis Buget	INSCRIS	B-2
2	3	0347	9.73	Prezent	8.64	8.97	Admis Buget	INSCRIS	B-3
3	4	0168	9.51	Prezent	8.46	8.78	Admis Buget	INSCRIS	B-4
4	5	0098	9.86	Prezent	8.26	8.74	Admis Buget	INSCRIS	B-5

	Nr.	Medie BAC	Nota Română
0	1	9.48	9.85
1	2	9.26	8.55
2	3	9.06	8.35
3	4	9.60	9.40
4	5	9.36	9.65



# Bac Dataset - Algoritmi ML

**Random Forest Classifier** pentru a prezice dacă un elev promovează sau nu Bacalaureatul, folosind dataset-ul rezultatelor de la Bacalaureat:

- Modelul este antrenat pe dataset
- Coloanele care indică notele probelor, status-ul - PROMOVAT/NEPROMOVAT și media finală sunt ignorate pentru a evita influențarea modelului

## Output:

- Se prezice statusul promovării și se afișează raportul de clasificare.
- Se afișează matricea de confuzie și un grafic cu feature importance.

## Random Forest Classifier - rezultate

Classification Report (Random Forest):

	precision	recall	f1-score	support
Nepromovat	0.5833	0.6946	0.6341	11978
Promovat	0.8505	0.7778	0.8125	26748
accuracy			0.7521	38726

**Nepromovat: 0.5833** → Doar 58.33% dintre elevii prezis ca *Nepromovat* au fost corect clasificați.

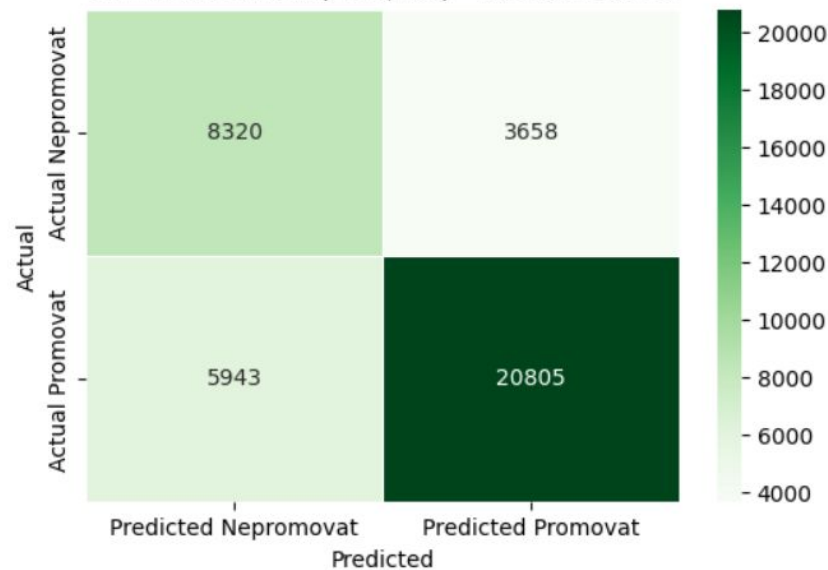
**Promovat: 0.8505** → 85.05% dintre elevii prezis ca *Promovat* au fost corect clasificați.

**Observație:** Modelul este mult mai precis în identificarea elevilor care promovează decât a celor care nu promovează.

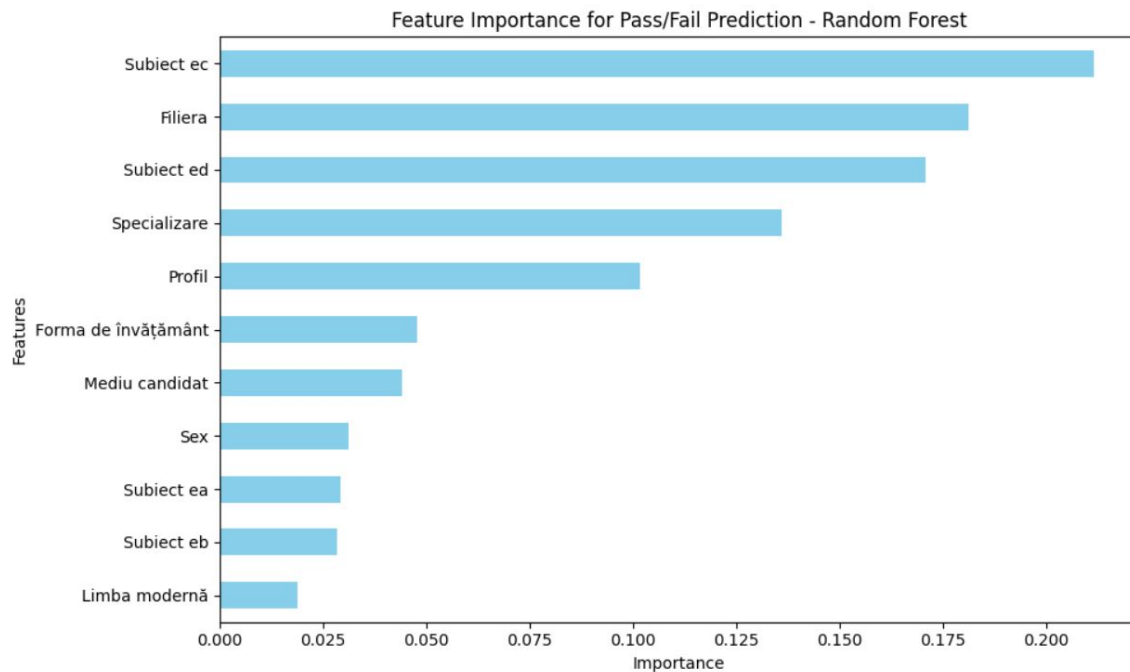
Confusion Matrix (Random Forest):

	Predicted Nepromovat	Predicted Promovat
Actual Nepromovat	8320	3658
Actual Promovat	5943	20805

Confusion Matrix (Pass/Fail) - Random Forest

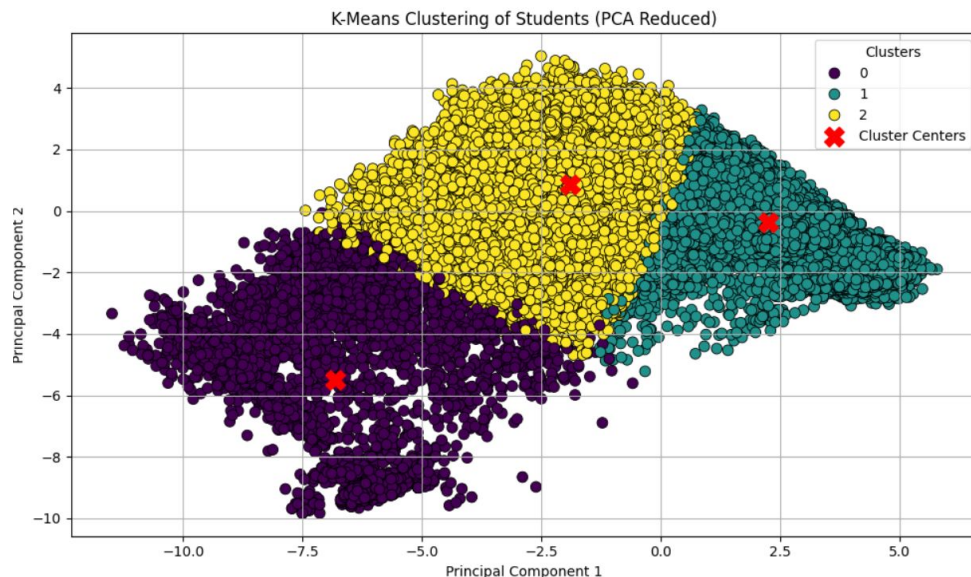


## Random Forest Classifier - Importanța coloanelor





## K-Means Clustering - grupare pe baza similarităților elevilor



### Cluster Summary:

- Cluster 0: 5169 students (3.80% of total)
- Cluster 1: 68526 students (50.35% of total)
- Cluster 2: 62396 students (45.85% of total)

### Top Specializations in Each Cluster:

#### Cluster 0:

- Tehnician mecanic pentru întreținere și reparații: 390 students
- Științe ale Naturii: 362 students
- Filologie: 288 students

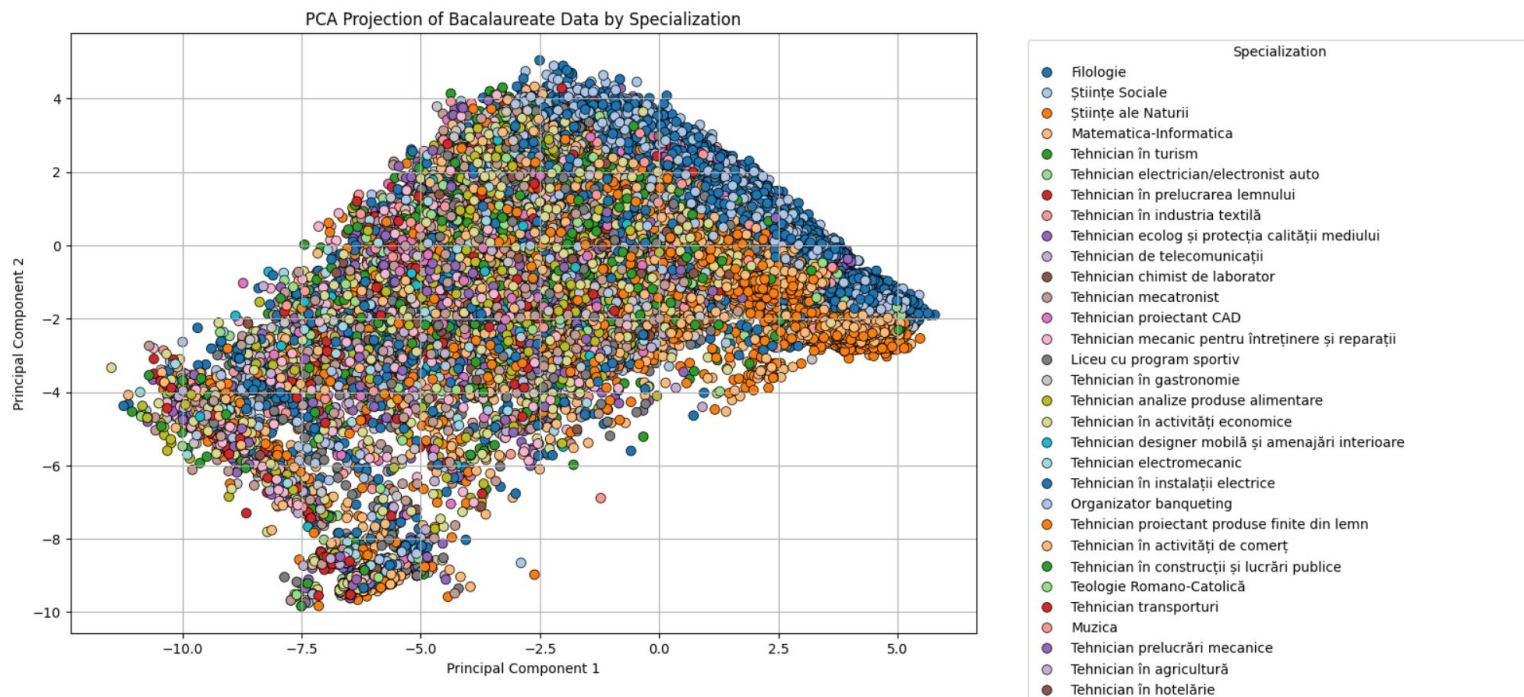
#### Cluster 1:

- Matematica-Informatica: 20245 students
- Filologie: 15064 students
- Științe ale Naturii: 12344 students

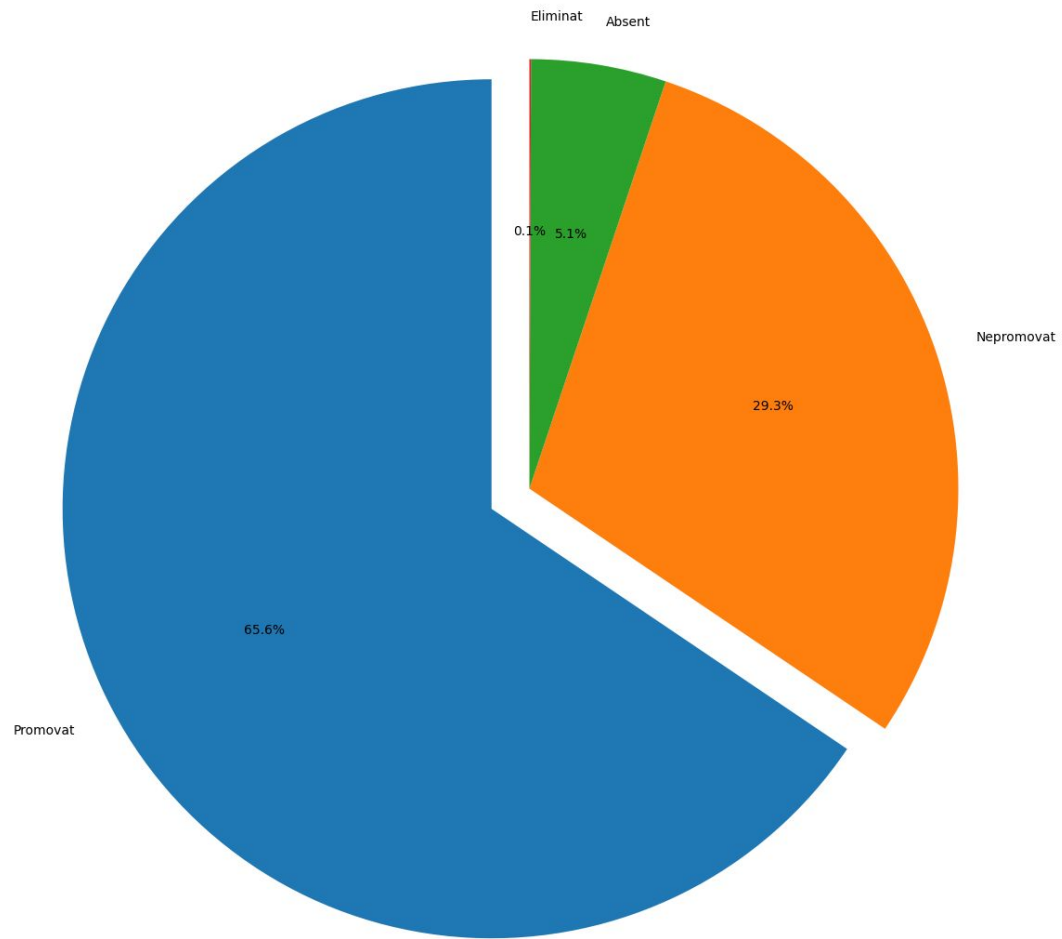
#### Cluster 2:

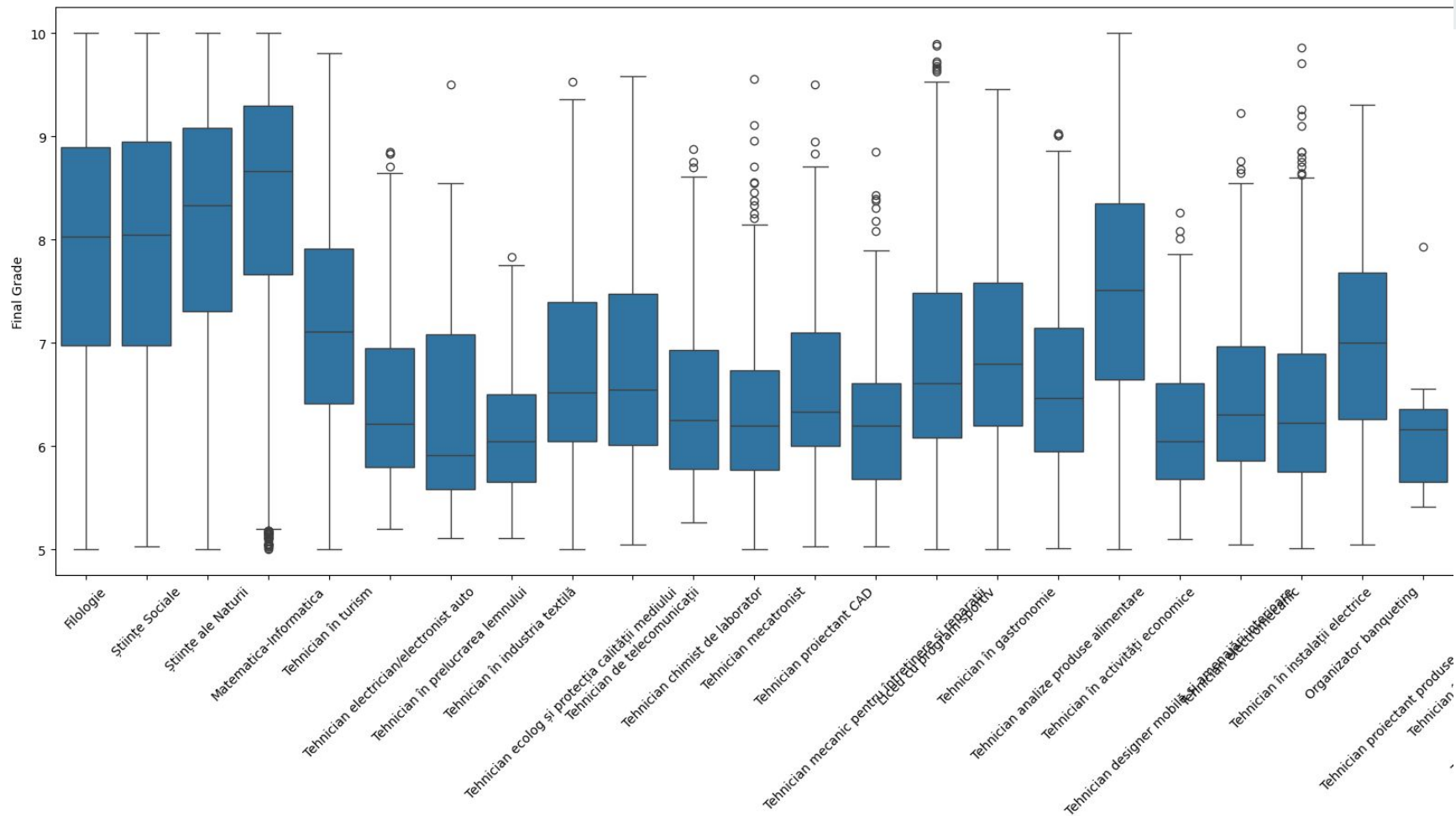
- Științe ale Naturii: 7200 students
- Tehnician în activități economice: 6328 students
- Filologie: 6149 students

## Principal Component Analysis (PCA) - grupare/specializare

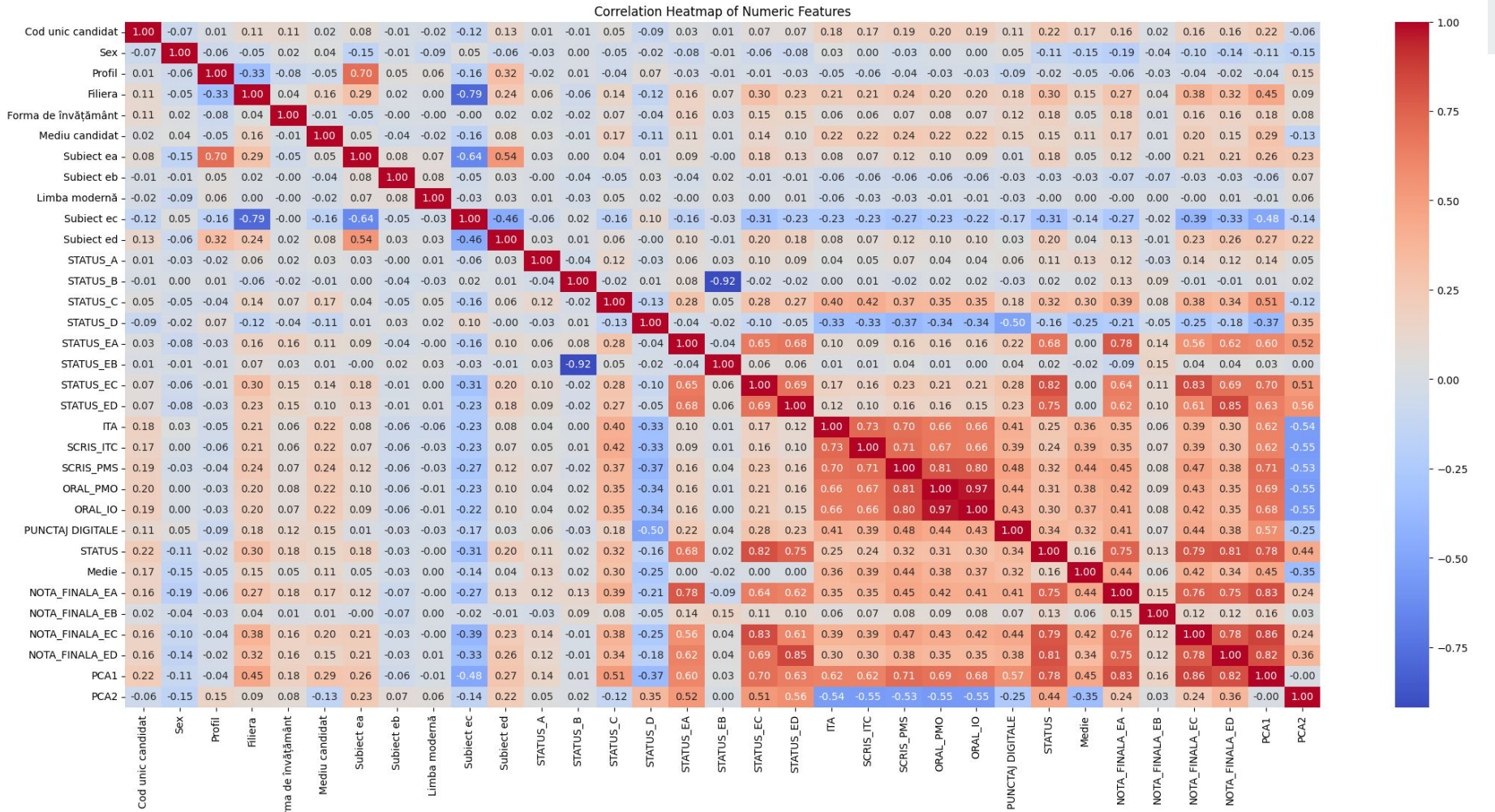


Pass/Fail Distribution









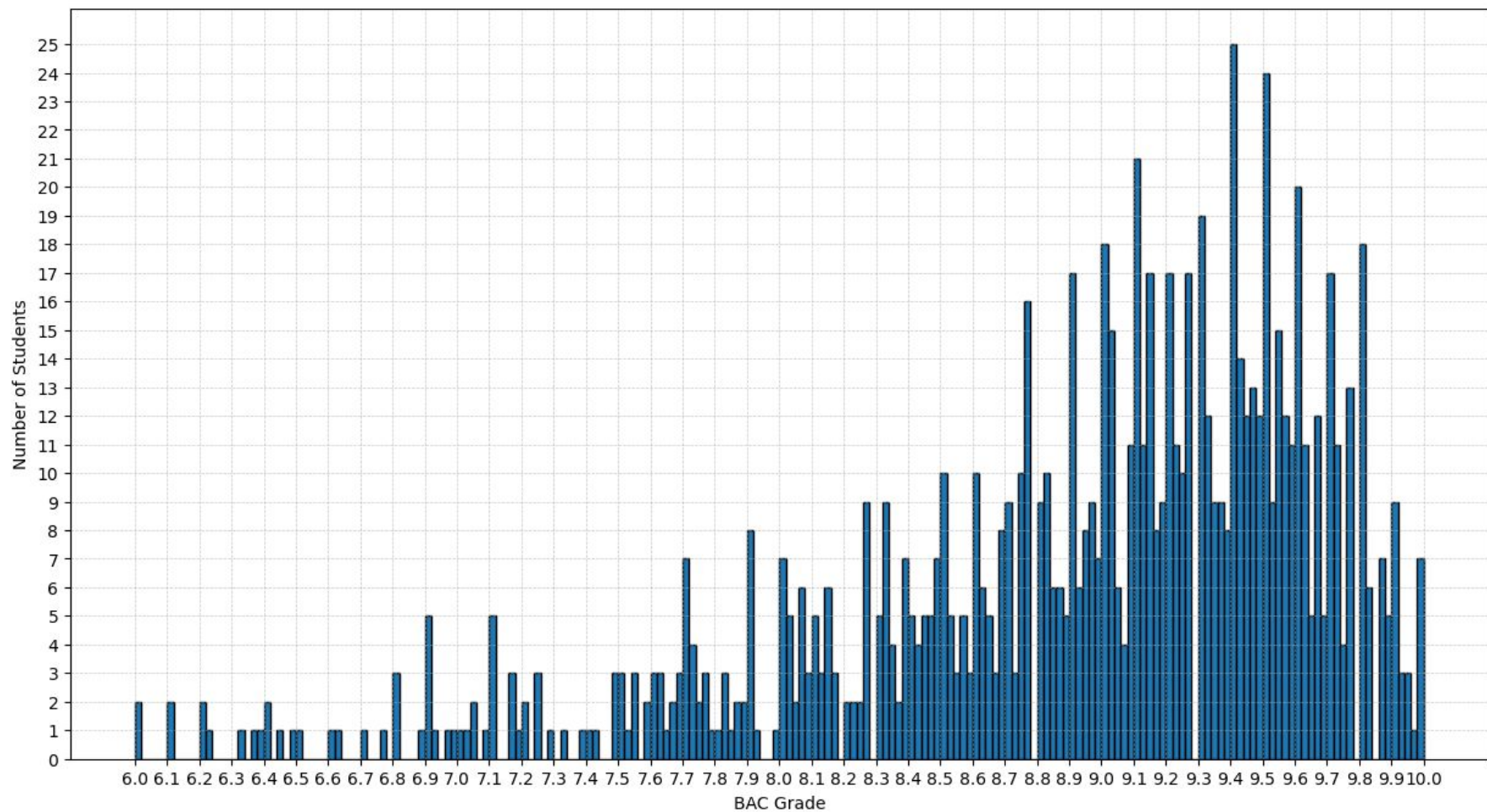


# Dataset Admiteri - Vizualizare

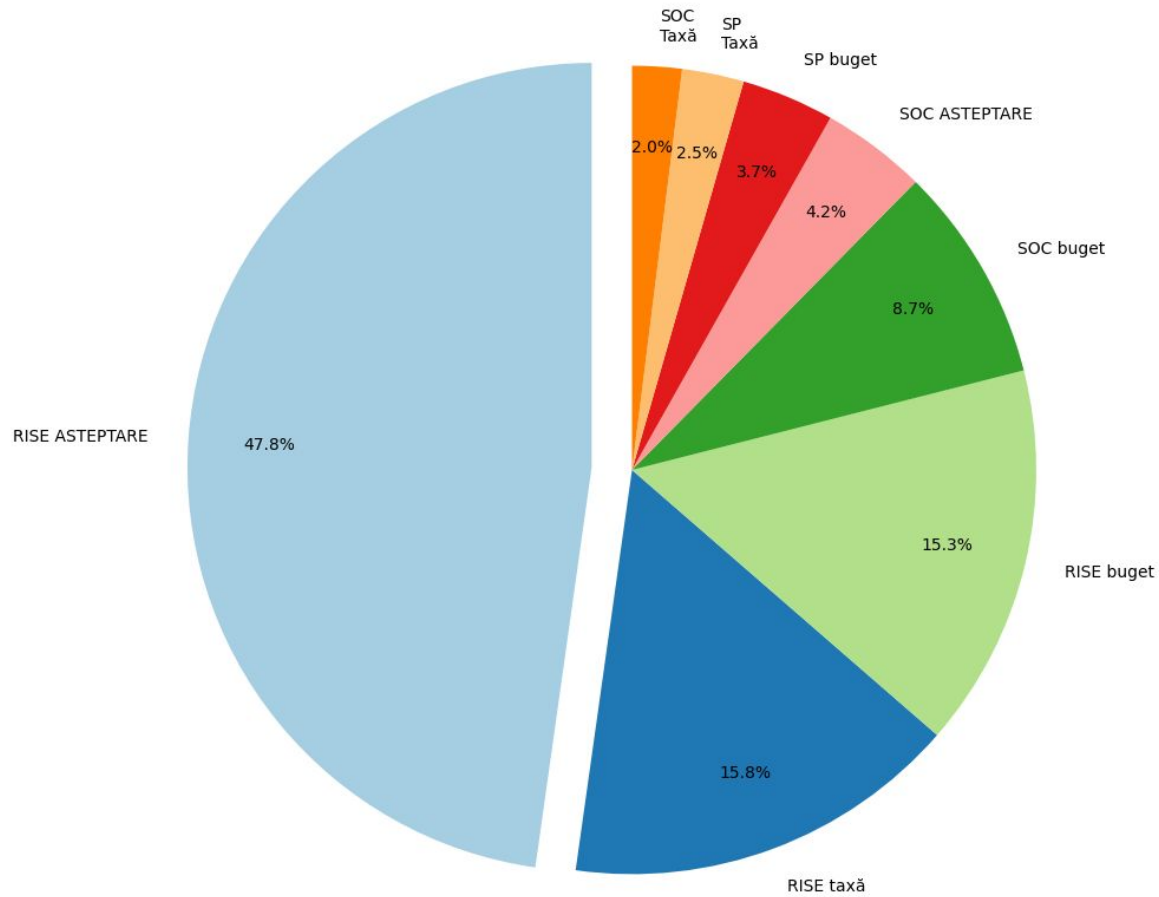
Această analiză își propune să evidențieze:

1. Distribuția notelor de admitere la diverse programe universitare din România, precum UNIBUC (CTI, Mate), UPB (ACS), UMFCD etc.
2. Corelația între performanțele la Bacalaureat și notele obținute la anumite examene specifice, precum examenul de Română (Drept).
3. Popularitatea specializărilor și statutul admiterii în funcție de datele disponibile pentru fiecare instituție.

Distribution of BAC Grades - UPB ACS

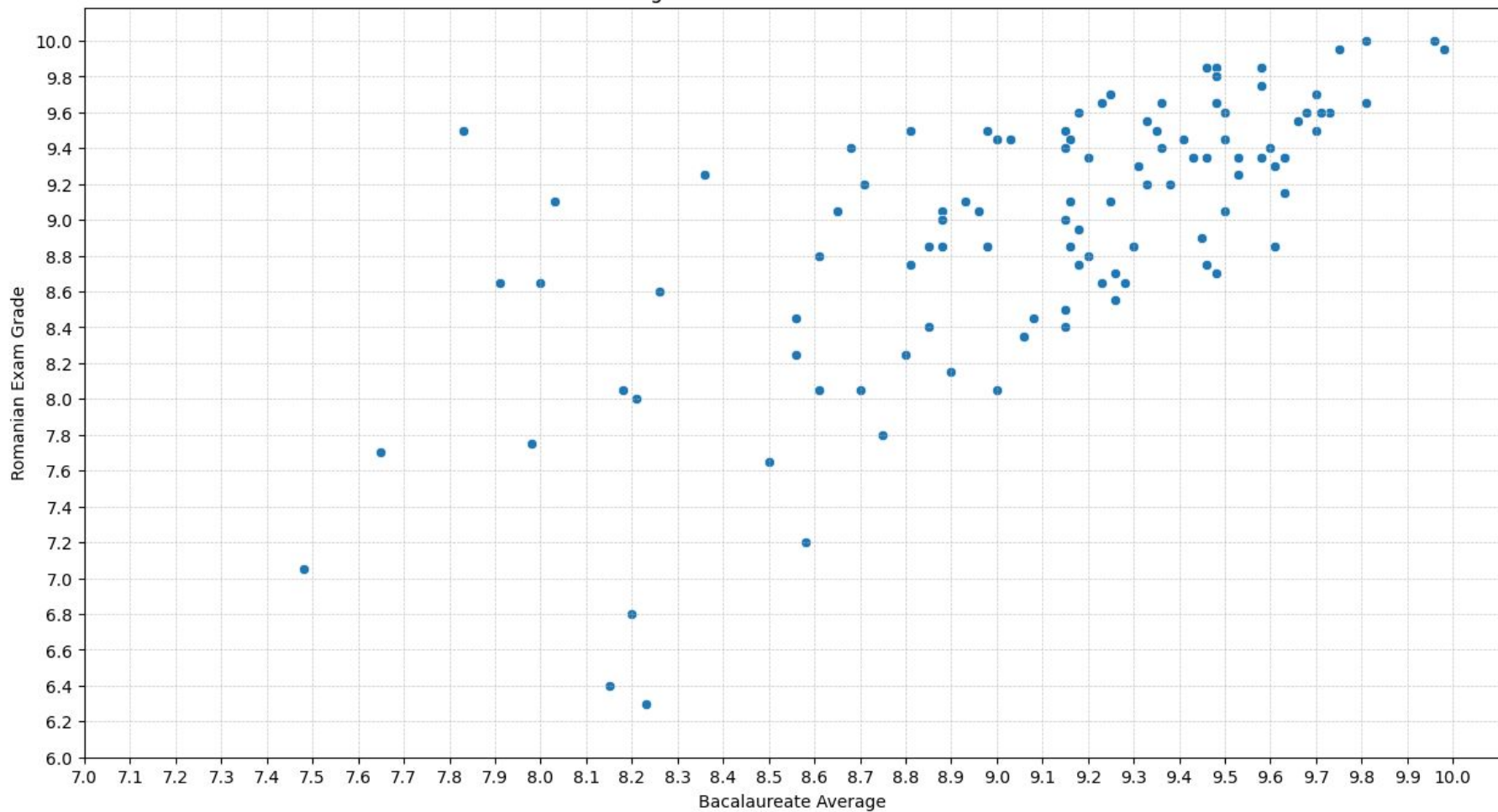


Distribution of Specializations - SNSPA

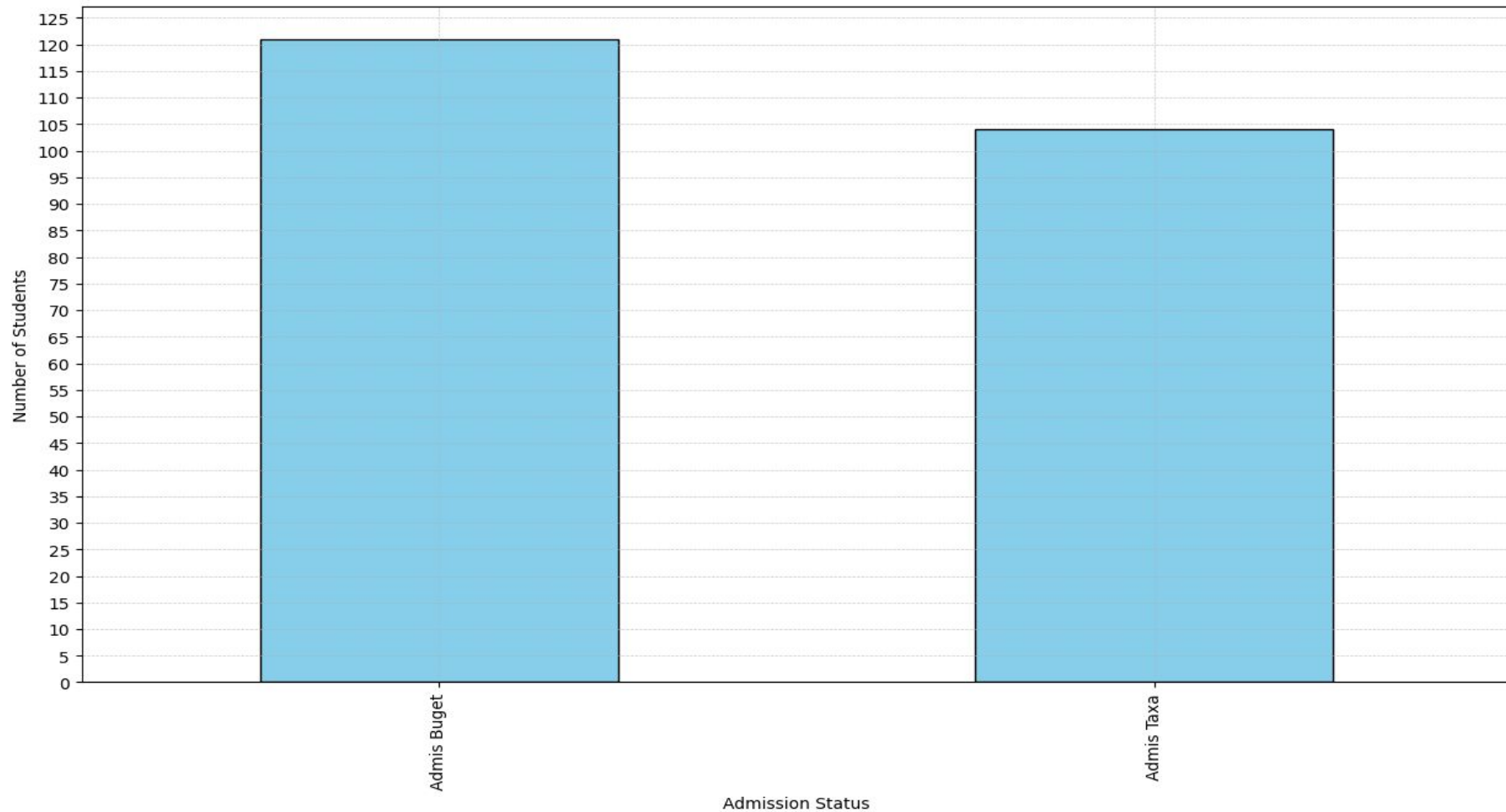




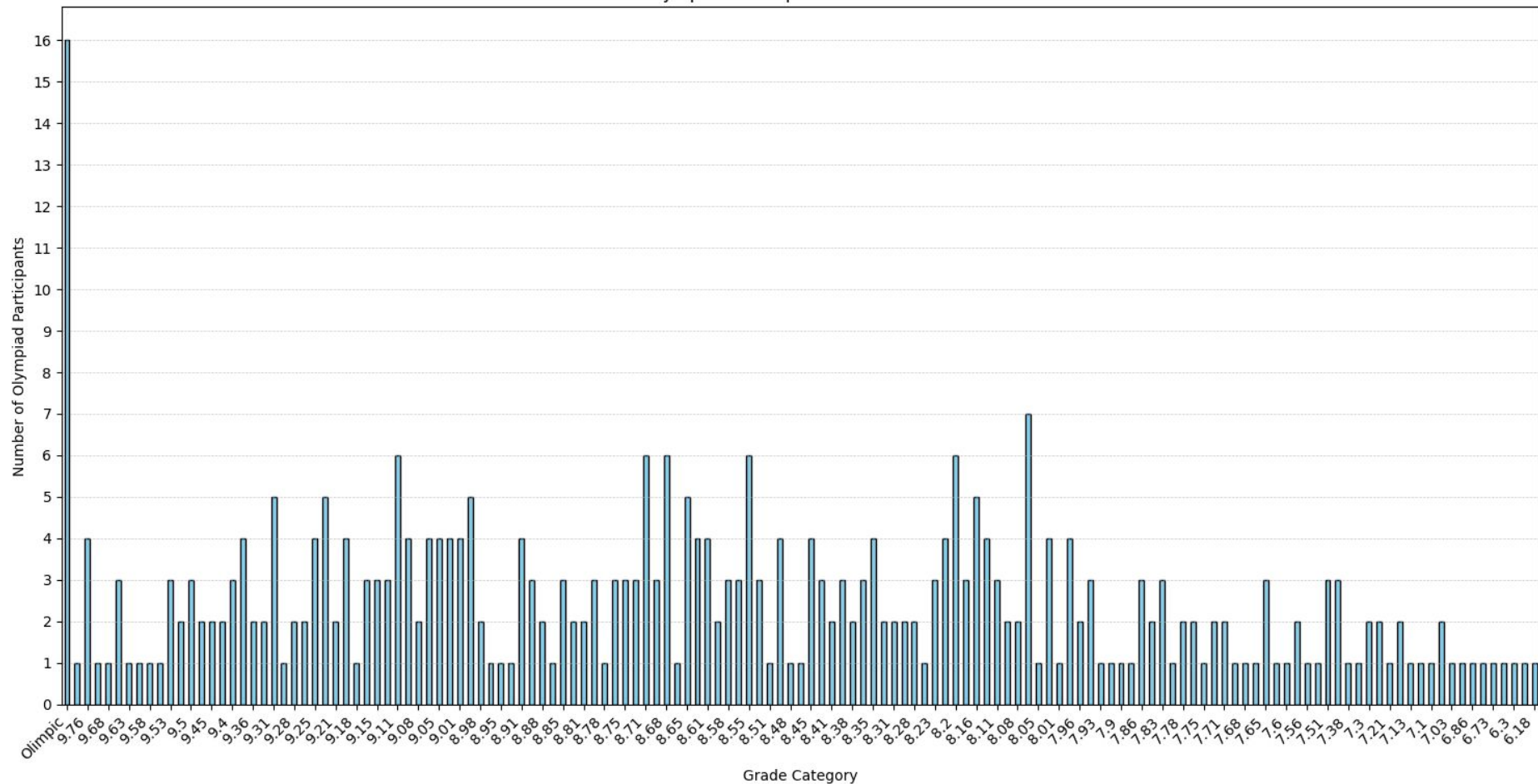
Bac Average vs. Romanian Exam Grade - DREPT



Admission Status Distribution - ARHITECTURA



# Olympiad Participation - GEOGRAFIE



# Demo Aplicație Web

Enter your BAC results

City:

Bucharest

Profil:

Real

Specializare:

Matematica-Informatica

Subiect ea (e.g. Limba română (REAL)):

Limba română (REAL)

Subiect ec (e.g. Matematică MATE-INFO):

Matematică MATE-INFO

Subiect ed (e.g. Fizică TEO):

Informatică MI C/C++

NOTA\_FINALA\_EA:

7.01

NOTA\_FINALA\_EC:

9.25

NOTA\_FINALA\_ED:

9.5

PUNCTAJ DIGITALE:

8.2

ORAL\_PMO:

9.25

Medie:

8.58

Get Recommendation

Recommended Field of Study: Computer Science

Detailed Scores:

Computer Science: 8.94

Mathematics: 8.19

Journalism: 8.06

Architecture: 7.15

Political Science: 5.22

Technical: 4.64

Geography: 4.49

Physical Education: 3.57

Psychology: 2.83

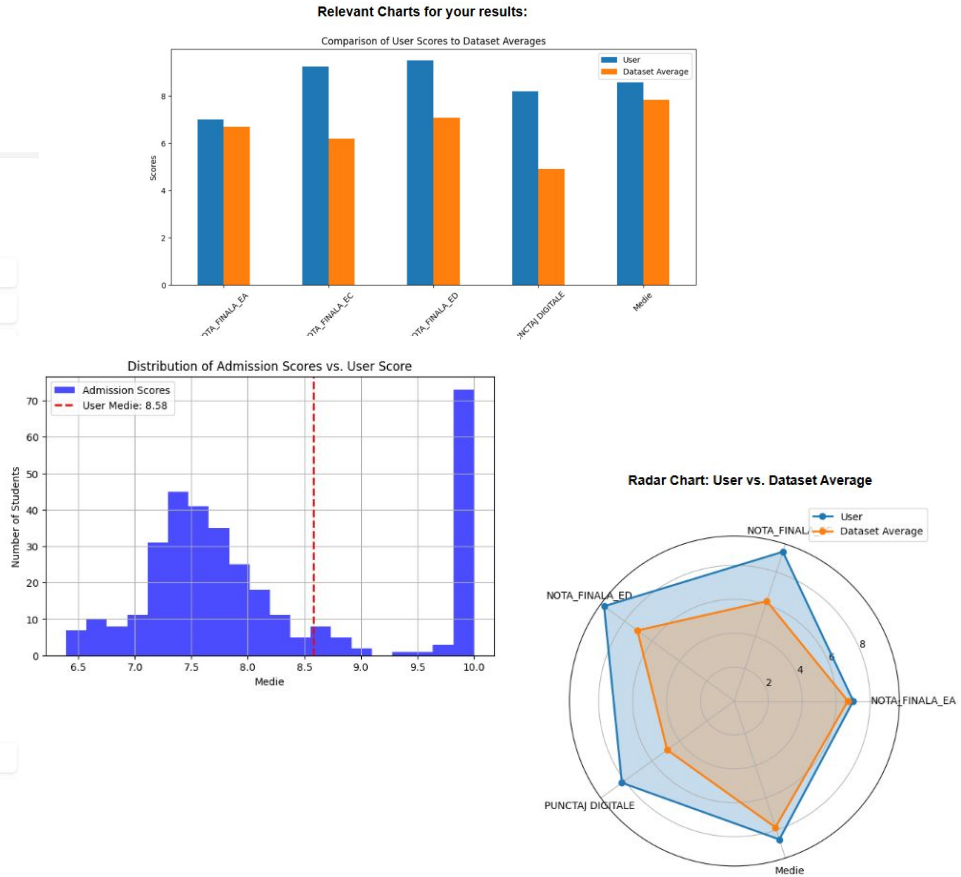
Medicine: 2.58

Arts: 2.35

Law: 1.83

Universities near Bucharest that might interest you:

Best Global Universities for Computer Science in Romania





## Referințe

- <https://docs.jupyter.org/en/latest/>
- <https://scikit-learn.org/stable/modules/generated/sklearn.ensemble.RandomForestClassifier.html>
- <https://realpython.com/k-means-clustering-python/>
- <https://flask.palletsprojects.com/en/stable/tutorial/>
- <https://serpapi.com/search-api>
- <http://ossasepia.com/bac-data/>