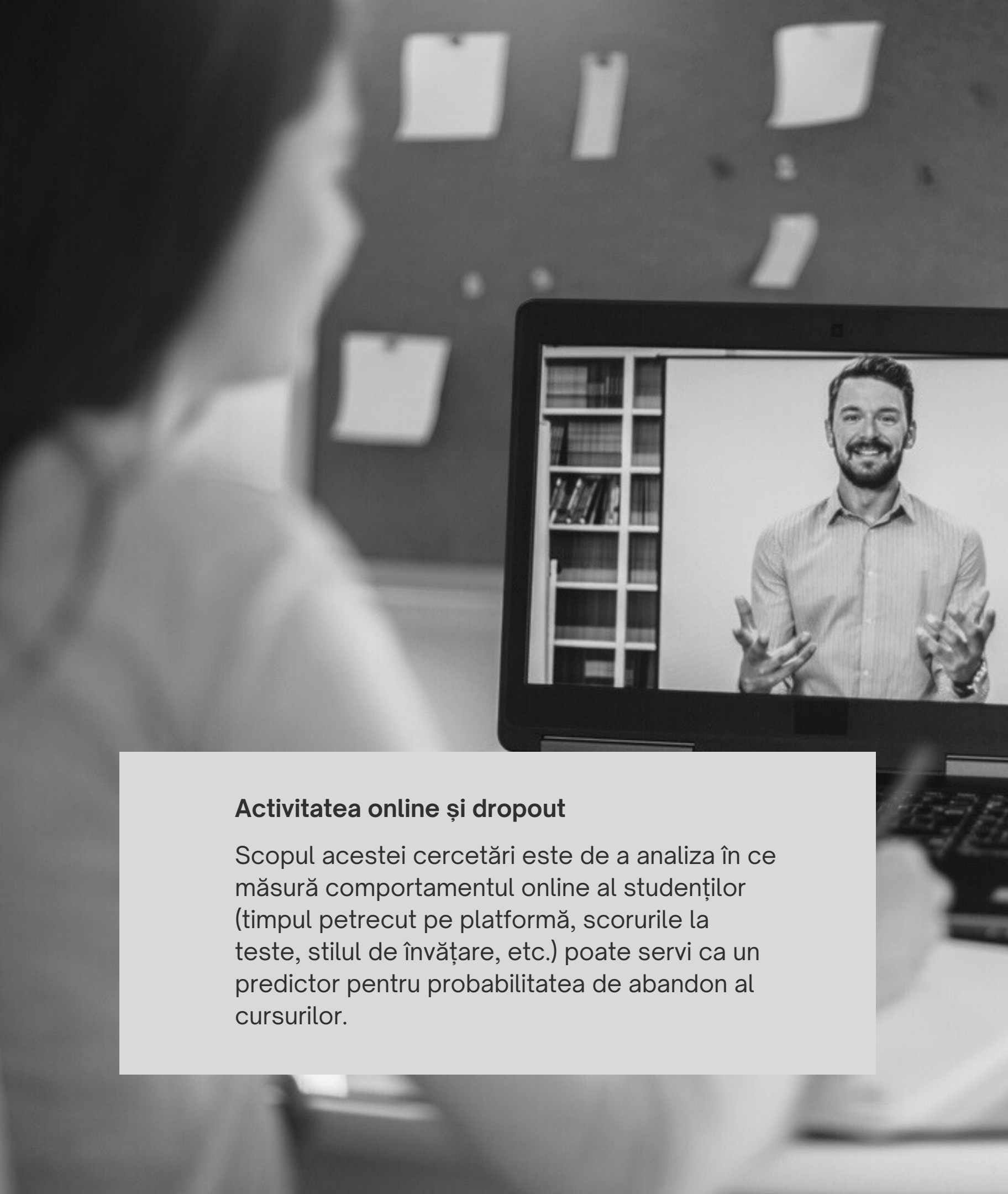


Învățarea adaptivă și predicția abandonului în educația online

19 Iunie, 2025

Laura Ciobotaru & Ignat Călin





Activitatea online și dropout

Scopul acestei cercetări este de a analiza în ce măsură comportamentul online al studenților (timpul petrecut pe platformă, scorurile la teste, stilul de învățare, etc.) poate servi ca un predictor pentru probabilitatea de abandon al cursurilor.



Metode avansate de cercetare
în științele comunicării



Contextul
cercetării
noastre

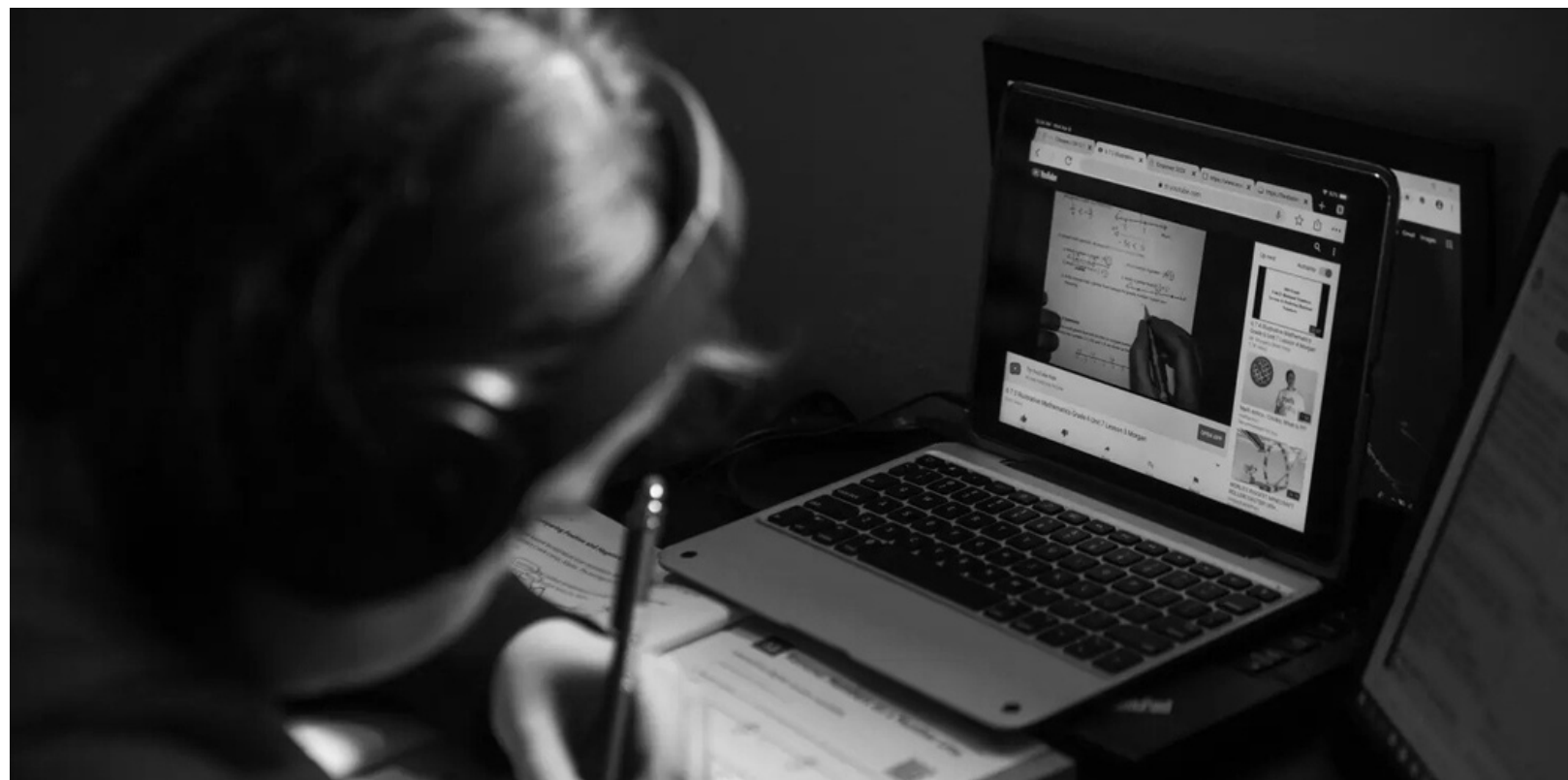


Digitalizarea și școala

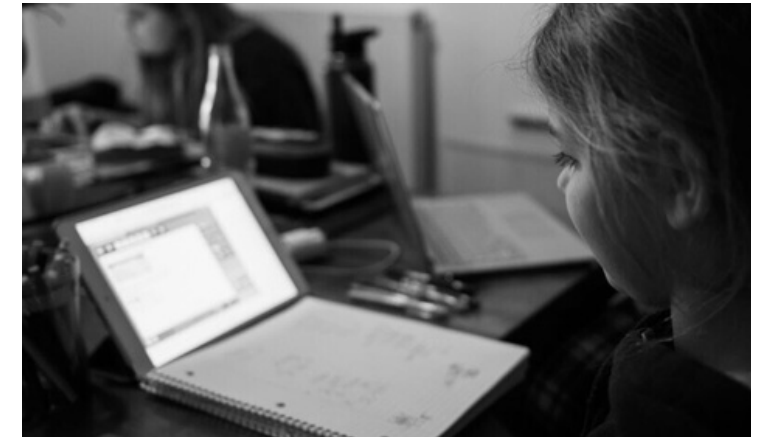
Digitalizarea educației a cunoscut o expansiune accelerată, determinată atât de progresele tehnologice, cât și de nevoia tot mai mare de acces flexibil la resursele educaționale. Astfel, învățarea adaptivă a devenit o soluție promițătoare pentru creșterea eficienței educației online.

Introducere

Context și întrebarea de cercetare



Definirea conceptelor
cheie



Dropout

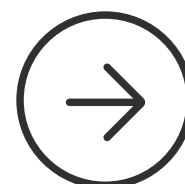
Dropout se referă la întreruperea cursului de către un participant înainte de finalizare. Acest fenomen reflectă lipsa de motivație, dificultăți de adaptare sau o nepotrivire între așteptările cursantului și oferta educațională.

Engagementul educațional

Engagementul educațional descrie nivelul de implicare activă al cursantului cu platforma, materialele și comunitatea de învățare (Time_Spent_on_Videos, Forum_Participation și Assignment_Completion_Rate).

Stilul de învățare

Stilul de învățare reflectă preferințele cognitive individuale în procesul de învățare, cum ar fi vizual, auditiv, kinestezic sau citit/scriș (Learning_Style).





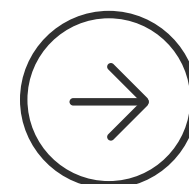
Ce cunoaștem și ce nu?

Cât de gravă este problema abandonului?

Rata de abandon în cadrul Massive Open Online Courses depășește frecvent 85%. Această rată ridicată a abandonului poate însemna și pierderea unor oportunități reale de dezvoltare profesională.

Ce comportamente pot prezice abandonul?

Timpul petrecut pe platformă, participarea activă în forumuri sau scorurile obținute la evaluările intermediare sunt frecvent asociate cu reușita sau eșecul cursantului. Scăderea bruscă a timpului de logare și numărul redus de interacțiuni sunt predictorii ai riscului de dropout.



Rolul stilurilor de învățare în adaptativitate și abandon

Stilurile de învățare pot influența motivația și nivelul de retenție. Cursanții ale căror stiluri nu sunt „adresate” de designul platformei pot manifesta un nivel mai scăzut de satisfacție și implicare, crescând astfel riscul de abandon.

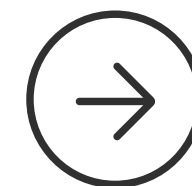


De ce este necesară cercetarea aceasta?

Există încă o lipsă de modele integrate care să combine variabilele cognitive cu cele comportamentale. Acest studiu propune o analiză ce combină ambele dimensiuni pentru a identifica mai clar tiparele asociate cu riscul de abandon.

Perspective și direcții

Modelele viitoare trebuie să integreze indicatori cantitativi de performanță, cât și informații calitative privind motivația sau stilurile de învățare. Această abordare ar putea duce la dezvoltarea unor sisteme educaționale adaptive.



Ce cunoaștem și ce nu?

⚙ Metode avansate de cercetare în științele comunicării

Întrebarea noastră de cercetare reprezintă punctul de plecare al studiului nostru, ghidând procesul de investigație și definind clara ce anume ne propunem să descoperim. În acest context, întrebările formulate sunt clare și precise, astfel încât să poată fi abordate eficient.

Care este capacitatea predictivă a comportamentelor de învățare online precum timpul petrecut pe platformă, participarea în forumuri, scorurile la teste și stilul de învățare, în estimarea riscului de abandon al cursurilor online?



Întrebarea de cercetare



Ipotezele cercetării noastre

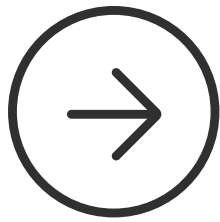
Ipotezele ce stau la baza cercetării sunt faptul că acei cursanți cu nivel scăzut de engagement au o probabilitate semnificativ mai mare de dropout, cursanții cu stil vizual sau kinestezic obțin scoruri semnificativ mai mari la examenul final, participarea frecvență la forumuri este asociată cu o scădere a probabilității de abandon,

iar combinarea activității video cu numărul de teme completate oferă o predicție mai bună a scorului final decât oricare dintre ele separat. Astfel, această cercetare urmărește să înțeleagă factorii care influențează succesul sau abandonul în mediul educațional online, oferind perspective valoroase pentru optimizarea platformelor de învățare.



Răspunsuri existente & ipoteze

Sursa și descrierea datasetului



Personalized Learning & Adaptive Education Dataset

Datasetul prezintă comportamentele cursanților într-un mediu online, oferind o bază solidă pentru cercetarea relațiilor dintre factori cognitivi, comportamentali și performanțele acedemice. Datasetul include peste 10.000 de înregistrări, fiecare corespunzând unui cursant anonim și este structurat pe mai multe categorii de variabile:

- Profilul cursantului, vârsta, genul și nivelul educațional
- Comportamente de învățare online, ca timpul petrecut urmărind materiale video, numărul de tentative la teste, nivelul de participare în forumuri, rata de completare a temelor și nivelul general de implicare
- Performanțele academice care sunt evaluate prin scorul mediu la teste, scorul obținut la examenul final și feedback-ul oferit cursului
- Stilul cognitiv și probabilitatea de abandon

	A	B	C	D	E	F	G	H	I	J	K	L	M	N
1	Student_ID	Age	Gender	Education_Level	Course_Name	Time_Spent	Quiz_Attempts	Quiz_Scores	Forum_Participation	Assignments	Engagement_Level	Final_Exam_Score	Learning_Style	Feedback
2	S00001	15	Female	High School	Machine Learning	171		4	67	2	89	Medium	51	Visual
3	S00002	49	Male	Undergraduate	Python Basics	156		4	64	0	94	Medium	92	Reading/Writing
4	S00003	20	Female	Undergraduate	Python Basics	217		2	55	2	67	Medium	45	Reading/Writing
5	S00004	37	Female	Undergraduate	Data Science	489		1	65	43	60	High	59	Visual
6	S00005	34	Female	Postgraduate	Python Basics	496		3	59	34	88	Medium	93	Visual
7	S00006	34	Male	Undergraduate	Web Development	184		1	87	34	70	Medium	43	Visual
8	S00007	45	Male	High School	Cybersecurity	454		3	69	46	83	Low	37	Kinesthetic
9	S00008	47	Male	High School	Cybersecurity	425		2	62	23	52	High	35	Reading/Writing
10	S00009	48	Male	Undergraduate	Cybersecurity	359		1	59	10	88	Medium	49	Reading/Writing
11	S00010	45	Female	Undergraduate	Data Science	263		4	63	30	99	Low	61	Auditory
12	S00011	29	Male	Postgraduate	Data Science	247		1	80	42	85	High	46	Visual
13	S00012	48	Male	Undergraduate	Python Basics	25		1	77	36	58	High	83	Reading/Writing
14	S00013	35	Male	High School	Python Basics	190		2	97	4	50	Medium	55	Auditory
15	S00014	29	Female	Undergraduate	Machine Learning	240		2	97	31	99	Medium	61	Kinesthetic
16	S00015	16	Male	High School	Data Science	157		1	30	46	73	Medium	49	Auditory
17	S00016	44	Male	Postgraduate	Python Basics	312		4	53	21	90	High	78	Kinesthetic
18	S00017	37	Male	High School	Machine Learning	239		1	83	44	65	High	53	Auditory
19	S00018	24	Female	Undergraduate	Machine Learning	150		3	98	20	83	Medium	62	Kinesthetic
20	S00019	38	Female	Undergraduate	Web Development	297		4	30	42	42	Low	86	Visual
21	S00020	40	Male	High School	Cybersecurity	409		4	45	8	50	Low	37	Visual
22	S00021	48	Male	Undergraduate	Machine Learning	222		4	65	28	95	Medium	44	Visual
23	S00022	28	Male	Postgraduate	Cybersecurity	96		3	51	38	83	High	69	Auditory
24	S00023	24	Female	Postgraduate	Cybersecurity	338		2	32	15	54	Medium	93	Visual
25	S00024	28	Female	High School	Data Science	452		2	85	49	45	Medium	68	Visual
26	S00025	40	Male	Undergraduate	Web Development	295		2	84	42	95	Medium	36	Auditory
27	S00026	23	Male	High School	Data Science	371		1	67	46	79	Low	32	Auditory
28	S00027	30	Male	High School	Python Basics	309		4	54	39	59	Medium	90	Visual
29	S00028	26	Male	Postgraduate	Cybersecurity	484		2	59	15	89	Low	87	Visual
30	S00029	30	Female	Undergraduate	Cybersecurity	67		4	99	27	89	Low	50	Reading/Writing
31	S00030	44	Male	Undergraduate	Web Development	408		4	44	44	90	Medium	65	Kinesthetic
32	S00031	17	Female	Postgraduate	Machine Learning	276		2	94	31	61	Medium	62	Kinesthetic
33	S00032	42	Female	High School	Cybersecurity	74		1	89	31	86	High	55	Kinesthetic
34	S00033	30	Female	Undergraduate	Data Science	350		2	54	49	52	Low	79	Kinesthetic
35	S00034	25	Male	Postgraduate	Data Science	321		4	45	28	46	High	61	Visual
36	S00035	17	Other	Undergraduate	Web Development	83		4	46	11	44	High	88	Auditory
37	S00036	21	Male	High School	Python Basics	77		2	35	29	64	High	67	Reading/Writing
38	S00037	16	Female	High School	Python Basics	435		2	46	7	67	Low	95	Auditory
39	S00038	37	Female	High School	Python Basics	258		1	35	15	90	High	31	Kinesthetic
40	S00039	46	Female	Undergraduate	Web Development	227		3	52	3	88	Medium	52	Visual



Justificare & preprocesare

Am ales acest dataset deoarece conține variabile cantitative, reflectă în mod realist comportamentele uzuale din învățarea online și este compatibil cu o gamă largă de metode statistice.

Pentru a asigura calitatea cercetării, am eliminat valorile lipsă, am identificat și tratat valorile extreme pentru variabilele comportamentale și am codificat valorile categorice pentru a putea fi utilizate în modelele statistice.



Datasetul

De ce este important?

- 49% femei (4.886 persoane)
- 47% bărbați (4.699 persoane)
- 4% altă identitate de gen („Other” – 415 persoane)

Aceasta înseamnă că genurile sunt distribuite aproape egal, fără ca un grup să fie dominant numeric.

Echilibru între grupuri

→ Permite comparații corecte între bărbați, femei și alte identități, fără ca rezultatele să fie influențate de un dezechilibru numeric.

Incluziune

→ Prezența categoriei „Other” arată că setul de date ia în considerare diversitatea de gen, ceea ce este important pentru studiile moderne în educație.

Engagement și abandon

→ Faptul că nu sunt mari diferențe de gen sugerează că rata de abandon sau gradul de implicare (engagement) nu sunt cauzate de gen, ci mai degrabă de alți factori, cum ar fi:

- stilul de învățare (vizual, auditiv etc.)
- timpul investit în platformă
- motivația personală

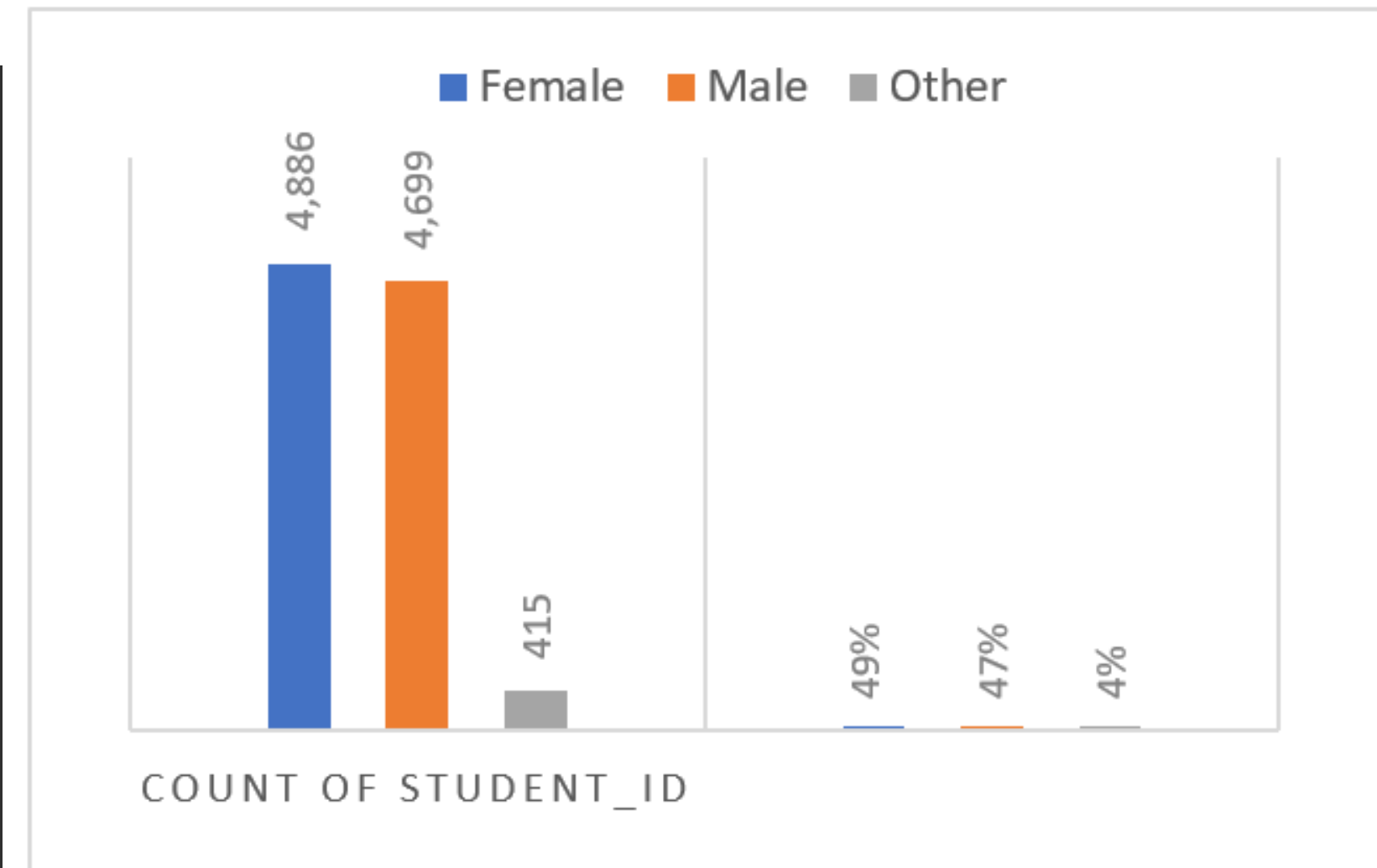


Figura 1 Distribuția cursanților după gen și identitate de gen

Demografie: gen și identitate

De ce este important?

- **Undergraduate (51%)** - majoritatea cursanților sunt studenți în formare
- **High School (29%)** - utilizatori preuniversitari, interesați de pregătire timpurie
- **Postgraduate (20%)** - reflectă interesul pentru educația continuă.

Această distribuție arată că platforma răspunde nevoilor unui public divers, de la liceeni la absolvenți de studii superioare. Faptul că este frecventată de utilizatori cu niveluri educaționale diferite sugerează:

- ✓ Accesibilitate ridicată
- ✓ Interes crescut pentru învățare continuă
- ✓ Potențial de personalizare pentru nevoi educaționale variate.

➔ O bază solidă pentru dezvoltarea unor programe flexibile și incluzive.

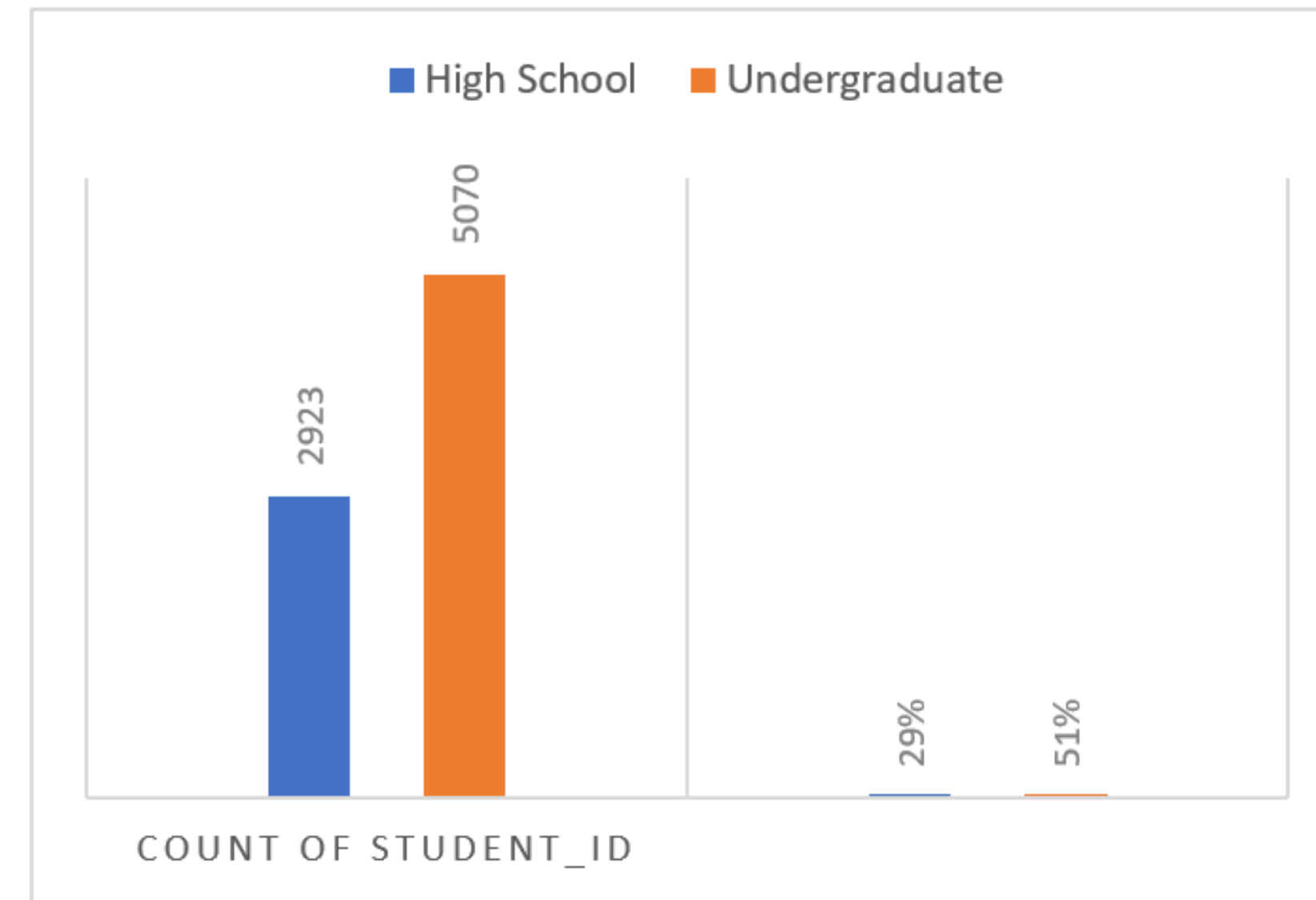


Figura 2 Repartitia cursanților în funcție de nivelul de educație

Nivel de educație

Scorul mediu general: 64.70

Undergraduate are media cea mai ridicată (64.95), urmat de High School (64.72), apoi Postgraduate (64.01)

Distribuția este relativ echilibrată, cu diferențe moderate

Cursanții „Other” la nivel Postgraduate au scor semnificativ mai mic (60.97) → posibil efect de eșantion redus sau factori de excluziune/demotivare

Implicație:

Formatul platformei pare mai adaptat pentru începători. Se recomandă explorarea nevoilor grupurilor subreprezentate pentru a crește incluziunea și performanța.

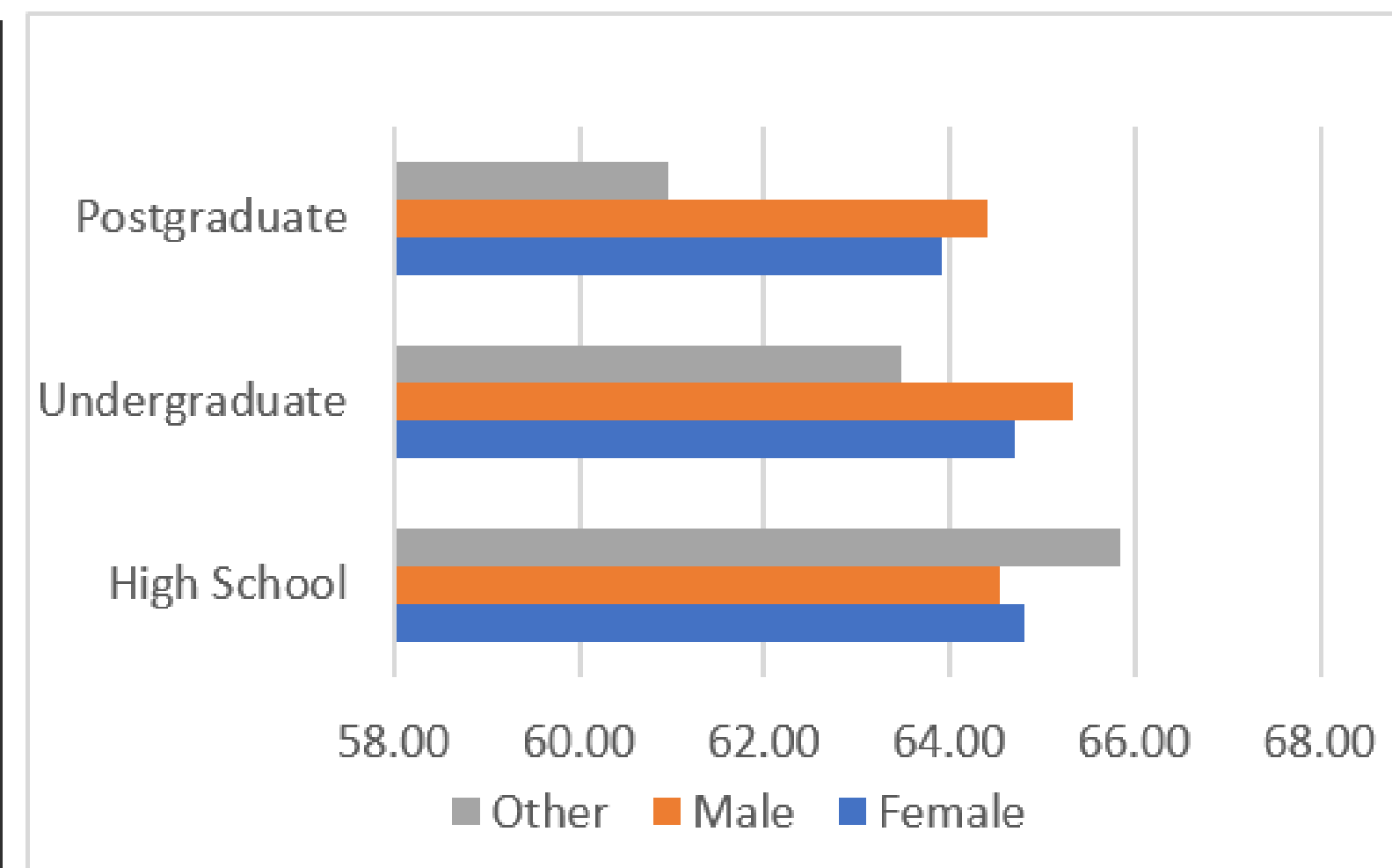


Figura 3 Scoruri medii finale la examen în funcție de gen și nivelul de educație

Performanță academică

Repartiție echilibrată între stiluri, cu un ușor avantaj pentru Reading/Writing (2554 cursanți)

Nivelul de engagement „High” este relativ uniform (729–764 cursanți per stil)

Auditory și Kinesthetic au o ușoară tendință spre engagement „Low”

Implicație:

Aceste stiluri ar putea beneficia de conținut mai interactiv sau aplicativ, adaptat preferințelor senzoriale ale cursanților.

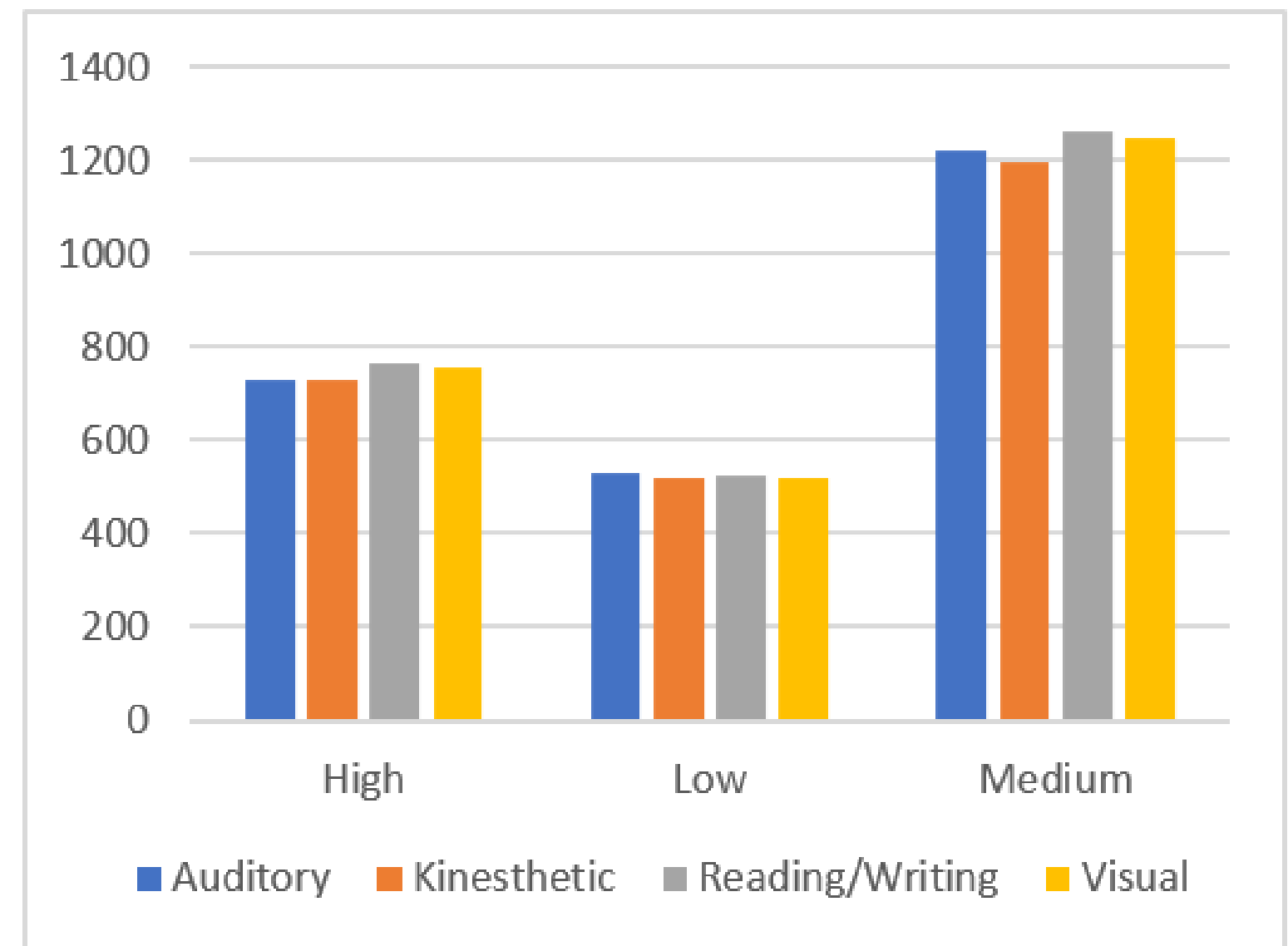


Figura 4 Distribuția stilurilor de învățare și nivelul de engagement asociat

Stiluri de învățare și engagement

Repartiție echilibrată între stiluri, cu un ușor avantaj pentru Reading/Writing (2554 cursanți)

Nivelul de engagement „High” este relativ uniform (729–764 cursanți per stil)

Auditory și Kinesthetic au o ușoară tendință spre engagement „Low”

Implicație:

Aceste stiluri ar putea beneficia de conținut mai interactiv sau aplicativ, adaptat preferințelor senzoriale ale cursanților.

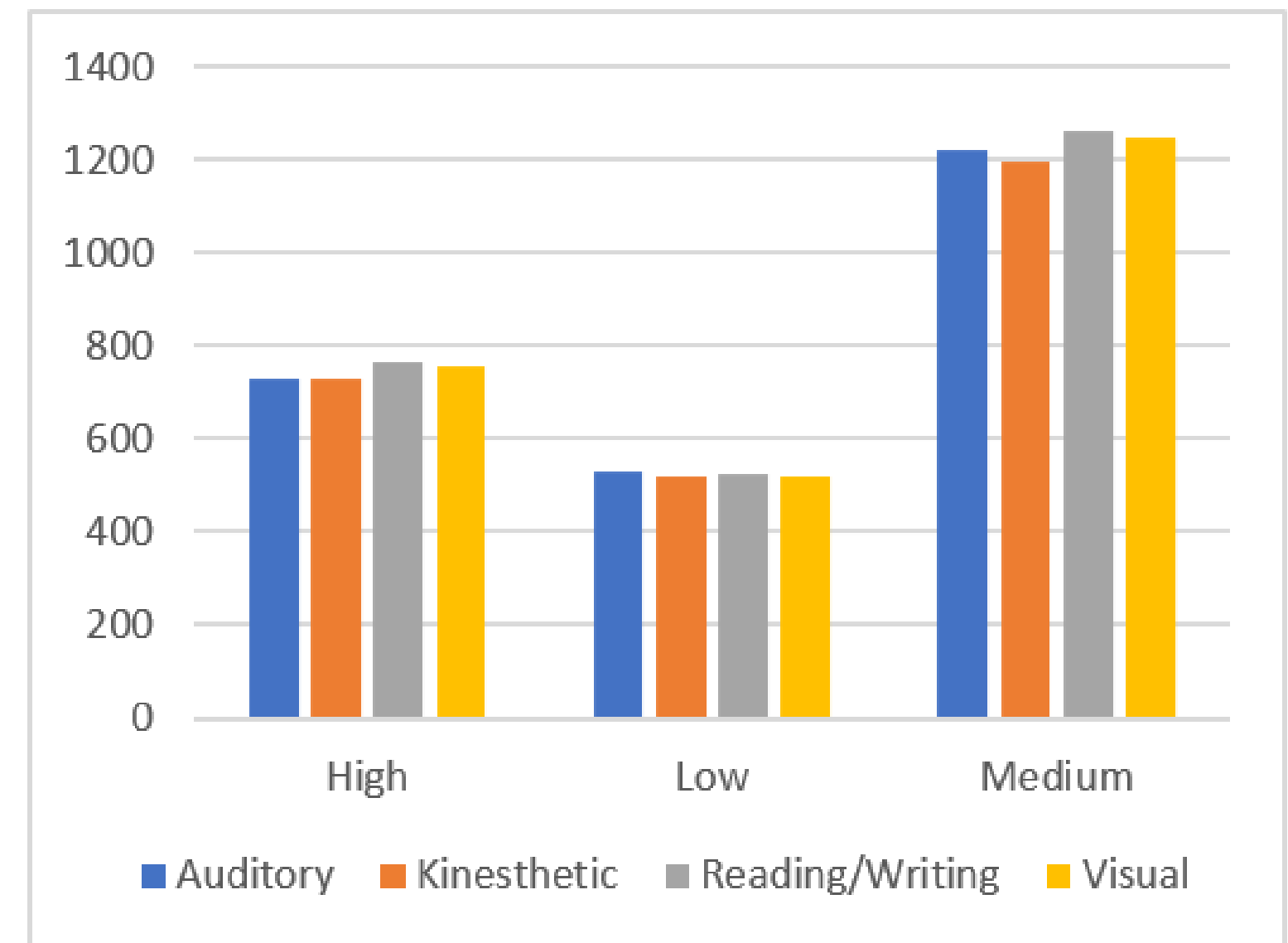


Figura 4 Distribuția stilurilor de învățare și nivelul de engagement asociat

Stiluri de învățare și engagement

- Variații minore între stiluri (< 1 punct procentual)
- Media generală a abandonului: ~19.6%
- Stilul de învățare nu este un predictor puternic al abandonului

Probabil că alți factori, precum nivelul de engagement sau participarea activă, influențează mai mult decizia de a finaliza cursul

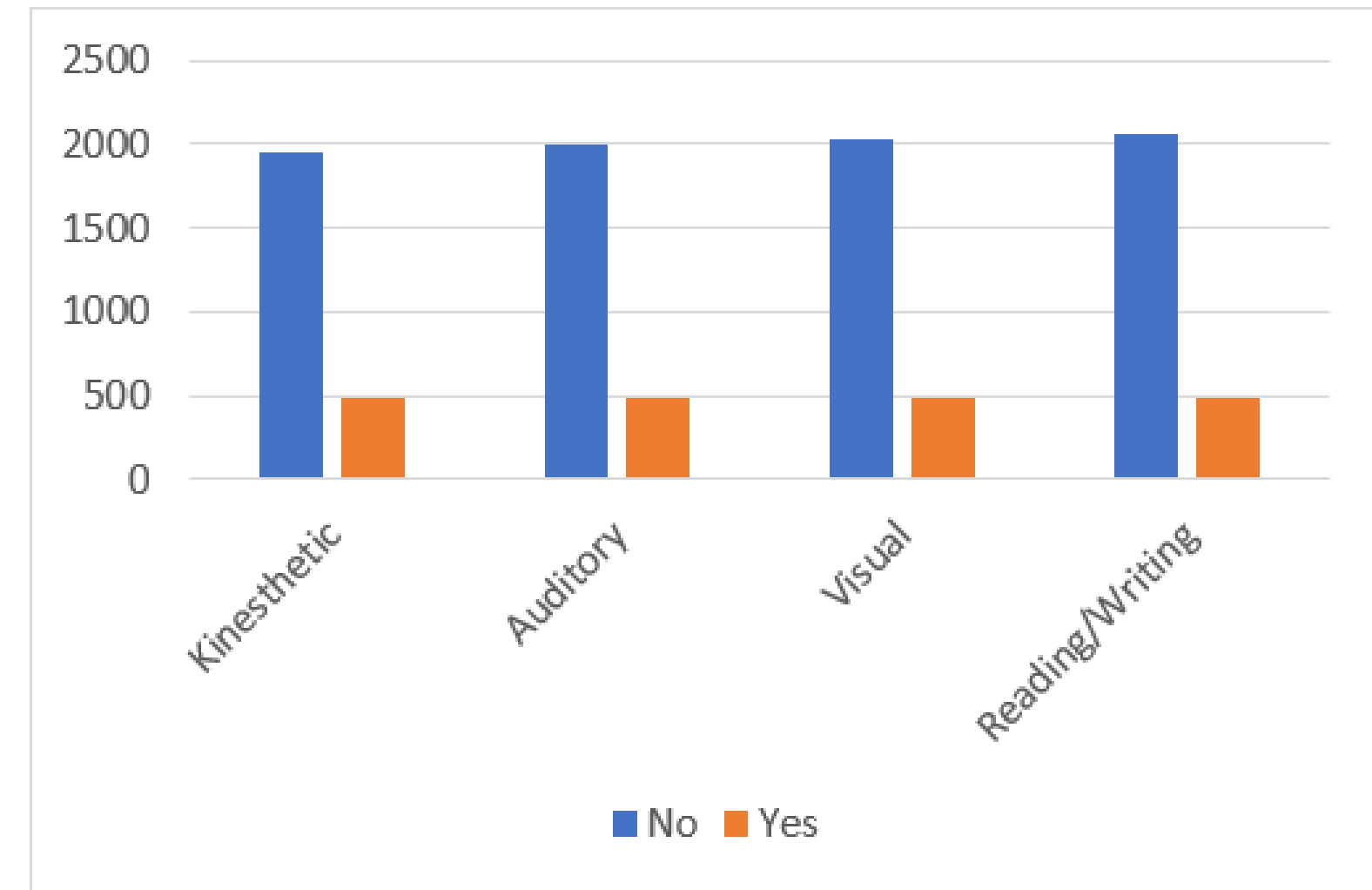


Figura 5 Rata de abandon în funcție de stilul de învățare

Dropout și stil cognitiv

Compararea performanțelor între quizuri și examen final după stilul de învățare

- Auditory: cea mai mare creștere a scorurilor (+1.13)
- Kinesthetic: scoruri constante, evidențiind consistență ridicată
- Reading/Writing: ușoară scădere la examenul final
- Posibilă nepotrivire între stilul de învățare și formatul evaluării finale

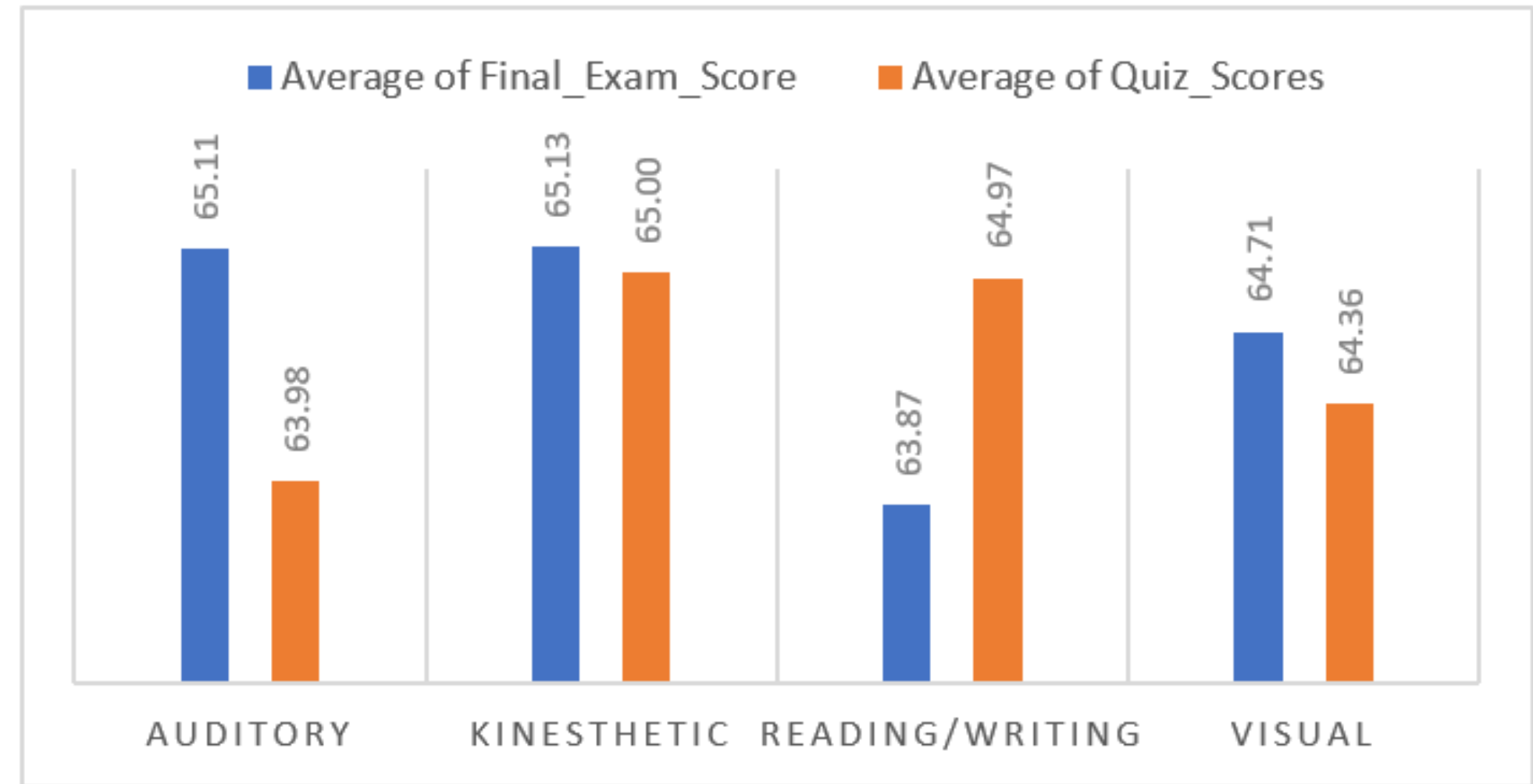


Figura 6 Evolutia scorurilor medii între quizuri și examenul final pe stiluri de învățare

Causes of the Industrial Revolution

Diferențe de sub 1 postare în medie

Kinesthetic are ușor cea mai mare medie → ar putea fi legat de preferința pentru „learning by doing”, inclusiv discutând în forumuri.

Reading/Writing și Visual sunt apropiate, confirmând ideea că aceste stiluri sunt confortabile în interacțiunea scrisă.

Nivelul educațional nu influențează puternic comportamentul de a relua quizurile.

Cu toate acestea, studenții postuniversitari au un ușor avantaj - poate datorită experienței academice sau auto-disciplinei.

Participare activă (forumuri & quizuri)

Learning_Style	Total
Auditory	24.44
Kinesthetic	24.92
Reading/Writing	24.57
Visual	24.54

Figura 7 Participarea medie în forumuri în funcție de stilul de învățare (ilustrează variația medie a postărilor pe forum pentru fiecare stil cognitiv: Auditory, Kinesthetic, Reading/Writing, Visual)

Average of Quiz_Attempts

Education_Level	Total
High School	2.48
Postgraduate	2.53
Undergraduate	2.50
Grand Total	2.50

Figura 8 Numărul mediu de tentative la quizuri în funcție de nivelul educațional (prezintă comparația între High School, Undergraduate și Postgraduate în ceea ce privește reluarea evaluărilor intermediare)

Teme versus dropout

Diferența de ~0.44 puncte procentuale este foarte mică. Asta sugerează că procentul temelor completate nu este un predictor puternic al abandonului, cel puțin nu singur.

Diferența de ~0.44 puncte procentuale" înseamnă că între cursanții care au abandonat cursul și cei care nu au abandonat, există o diferență foarte mică în procentul temelor finalizate (sub jumătate de procent).

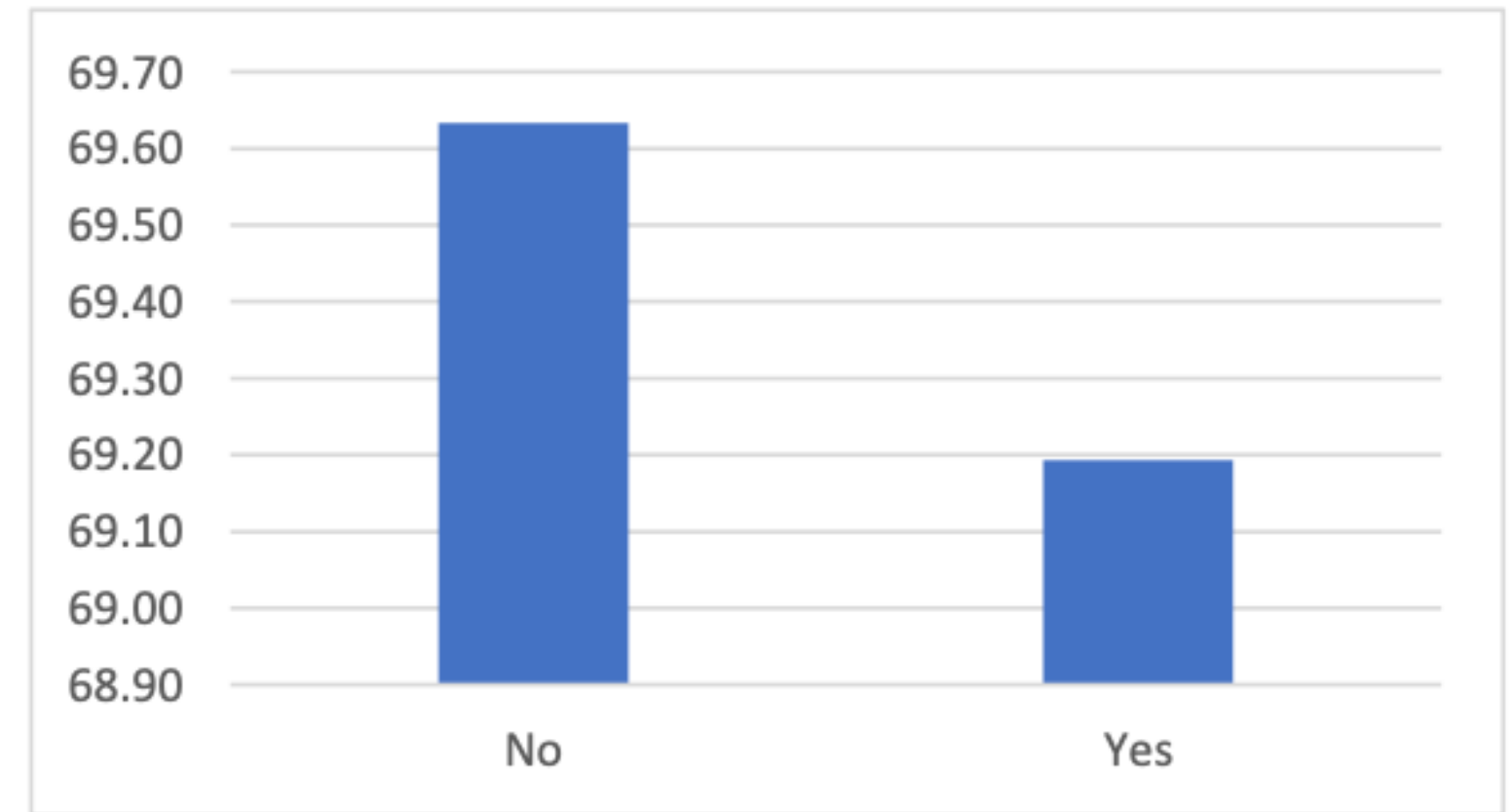


Figure 9 Participarea medie în forumuri și numărul mediu de tentative la quizuri după stiluri de învățare și nivel educațional

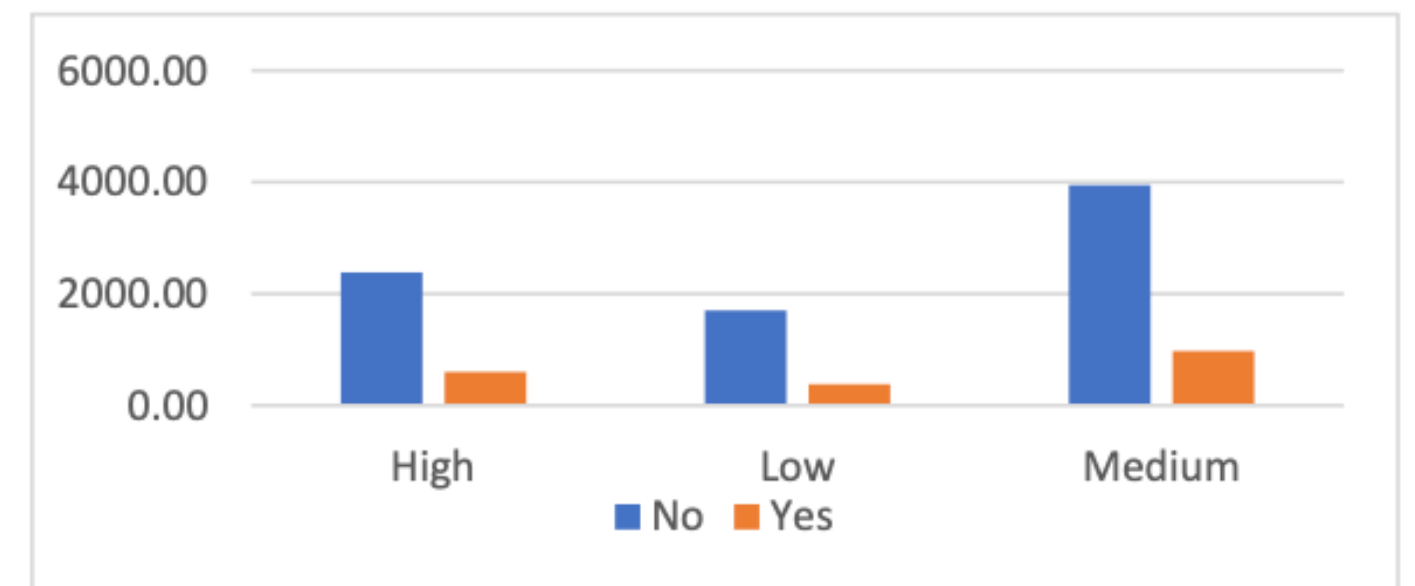


Figura 10 Rata de abandon în funcție de nivelul de engagement (Low, Medium, High)

Comportamente + rezultate (engagement + completare)

Abandon și felul în care învață

Diferențele între stiluri sunt foarte mici între 0.19 și 0.20). Acest lucru sugerează că stilul de învățare, izolat, nu este un predictor puternic al abandonului.

Poți integra această observație în analiza ta, întărind ideea că este nevoie de modele predictive mai complexe, care să integreze factori multipli (ex. engagement, stiluri + performanță).

Analiza descriptivă asupra datelor din eșantionul de 10.000 de cursanți arată că rata medie de abandon diferă foarte puțin în funcție de stilul de învățare declarat. Stilurile „Auditory” și „Reading” prezintă o rată de 0.19, în timp ce „Kinesthetic” și categoria „Other” ajung la 0.20.

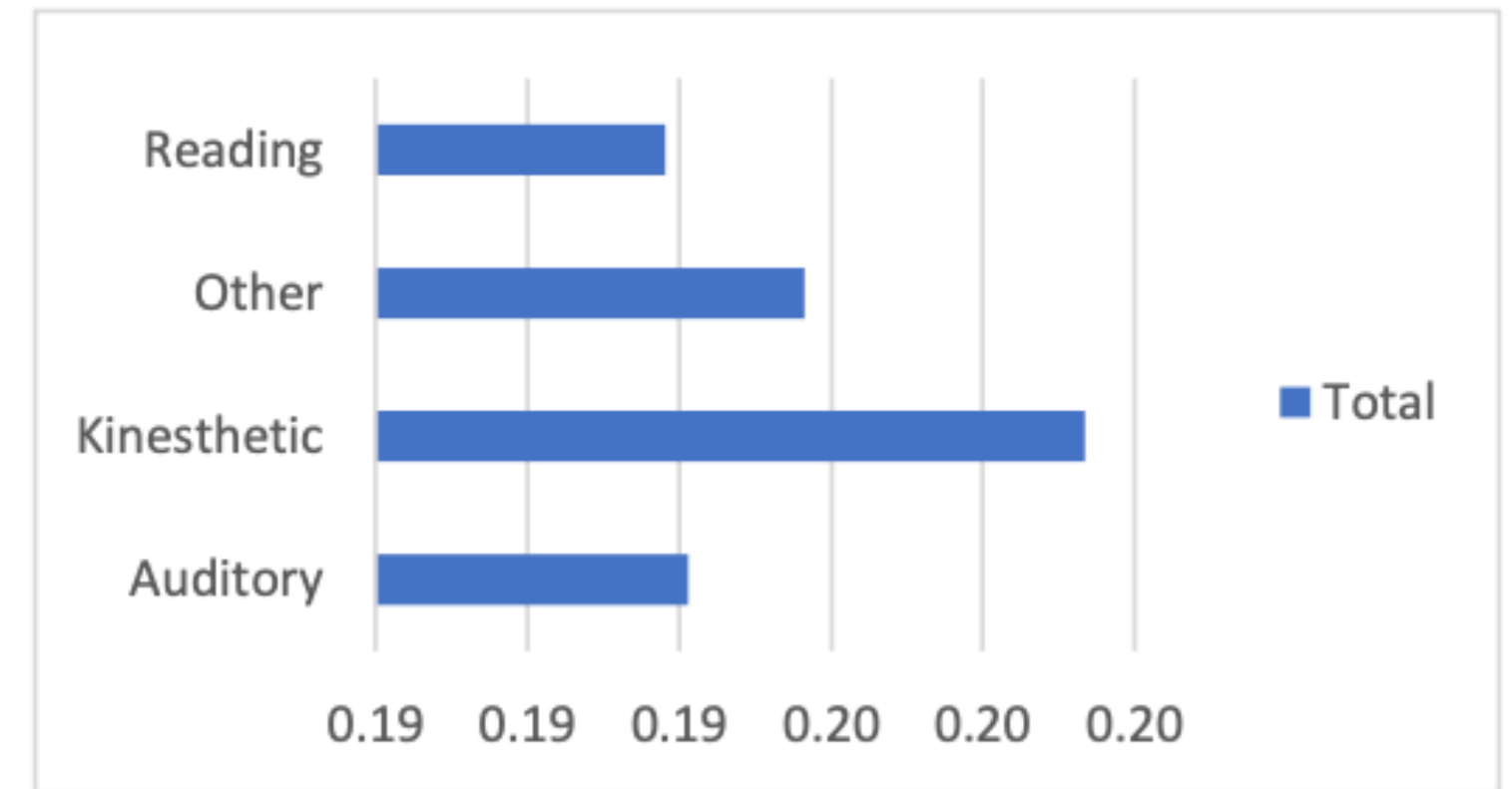


Figura 11 Rata medie de abandon pe categorii de stiluri de învățare

Predictia abandonului în funcție de stilul de învățare

Cursanții și cursurile

Distribuția cursanților pe cele cinci cursuri din platformă este relativ echilibrată, cu variații minore (între ~1.950 și ~2.050 cursanți per curs). Această omogenitate ajută la evitarea efectelor de eșantionare disproporționată și conferă robustețe rezultatelor comparative între cursuri.

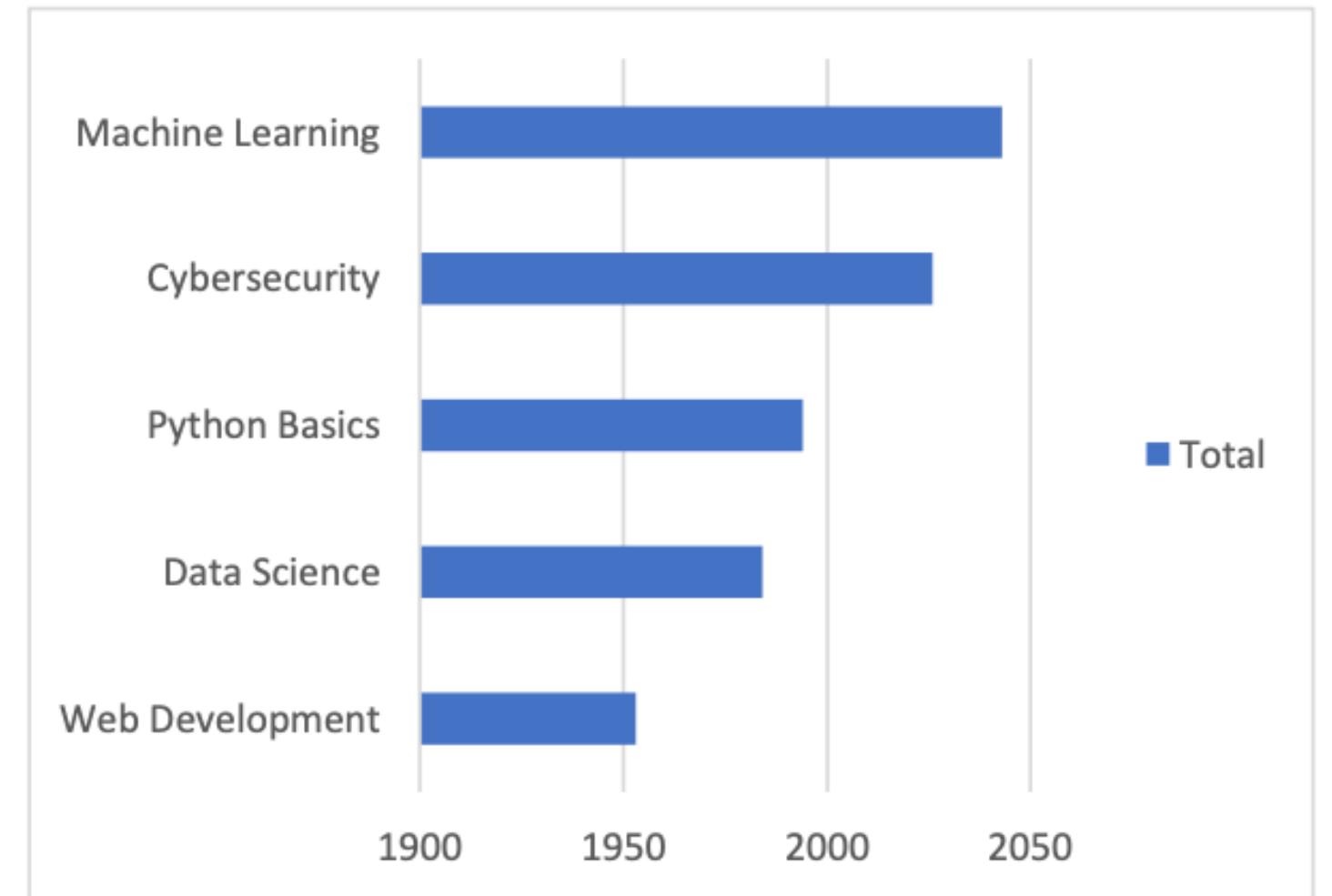


Figura 12 Numărul cursanților înregistrați pe fiecare curs

Repartiția cursanților pe cursuri (volum)

Cursurile și abandonul

Rata estimată de abandon este relativ constantă între cursuri, situându-se între 17% și 21%. Cursul „Machine Learning” înregistrează cel mai mare număr de cursanți cu probabilitate de abandon (Yes = 429), urmat de „Cybersecurity” (398). În schimb, „Web Development” are cea mai scăzută valoare (357).

Aceste diferențe moderate sugerează că tipul cursului poate influența ușor riscul de abandon, dar nu în mod decisiv. Este posibil ca factori precum dificultatea percepută, profilul cursanților sau structura conținutului să joace un rol în variațiile observate.

Rata de abandon în funcție de cursul urmat

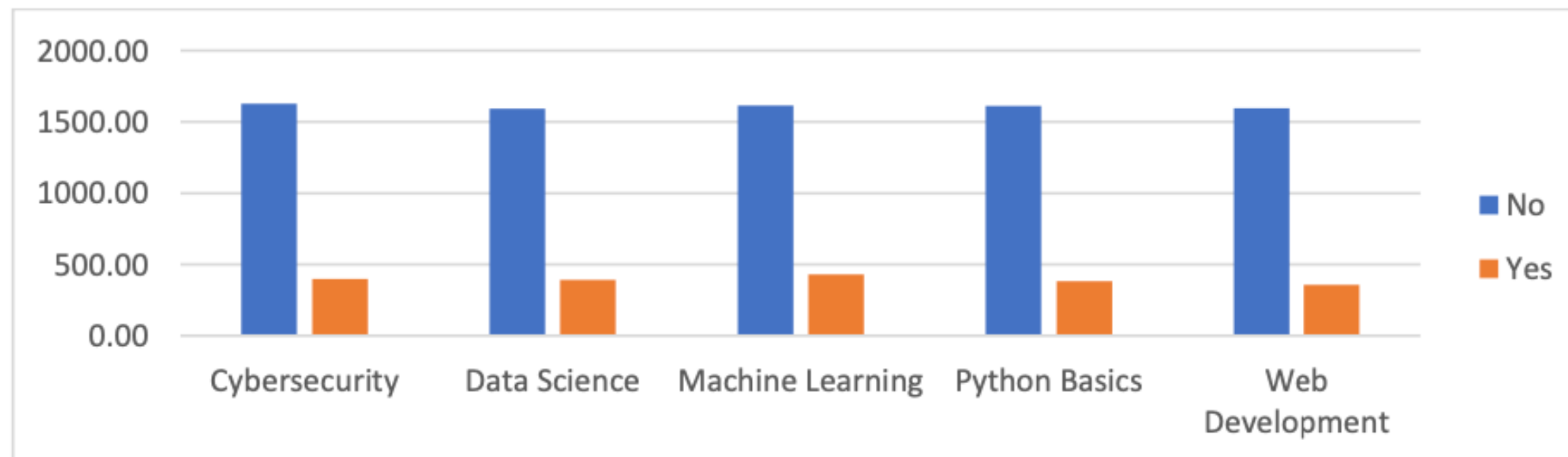


Figura 13 Rata de abandon estimată pe cursuri

Scoruri finale și cursuri

Datele arată că performanța medie la examenele finale și la quizuri este relativ uniformă între cursuri, variind foarte puțin în jurul valorii de 64–65 de puncte. Cybersecurity și Data Science se evidențiază printr-un echilibru perfect între scorul la quizuri și la examenul final (65–65).

Machine Learning are un scor ușor mai mic la examenul final (64), posibil sugerând o dificultate mai mare la evaluarea finală. Python Basics și Web Development înregistrează scoruri puțin mai scăzute la quizuri (64), ceea ce poate indica o dificultate crescută în sarcinile intermediare sau o lipsă de familiarizare timpurie cu materialul.

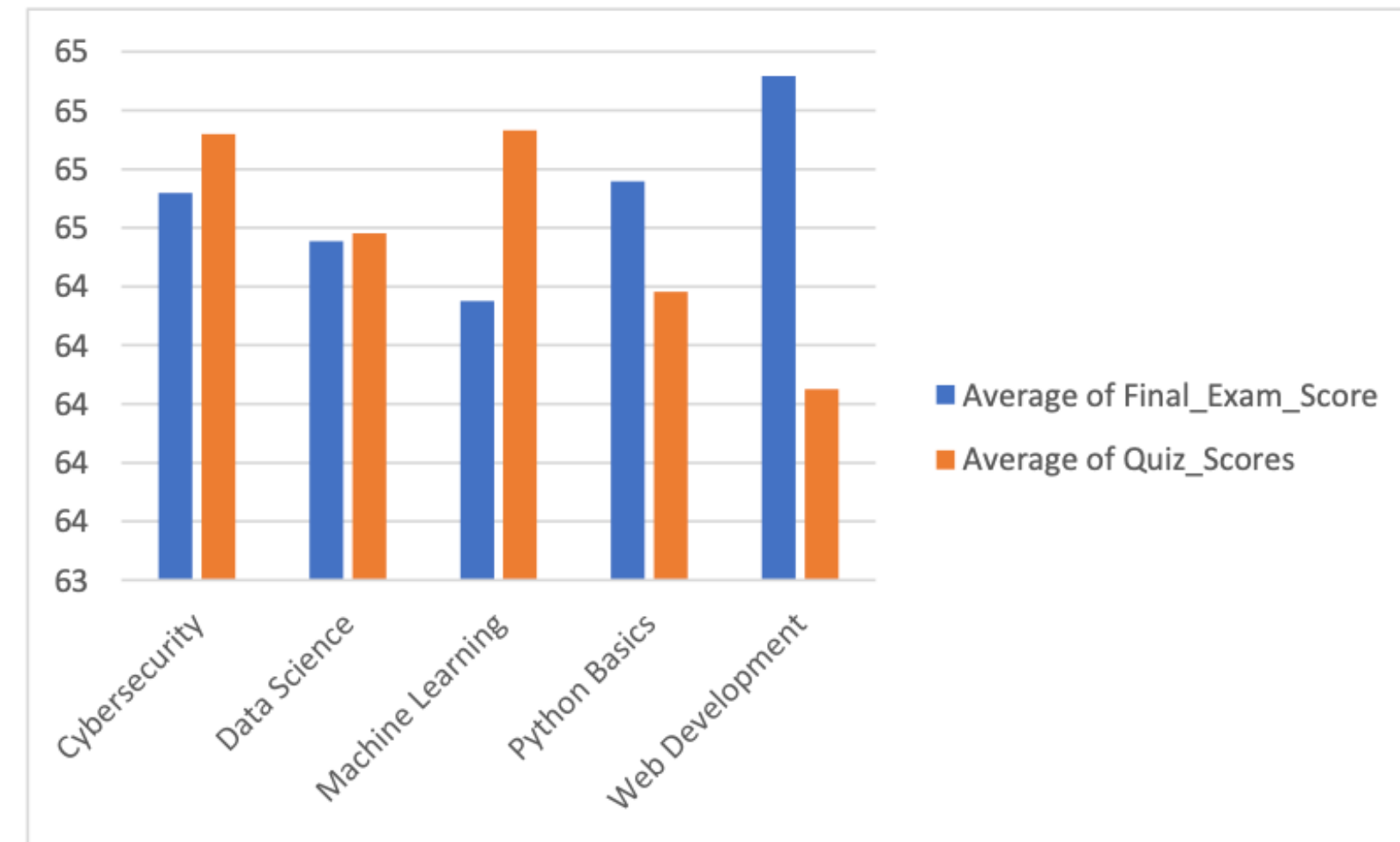


Figura 14 Scoruri medii la examenul final și quizuri pe cursuri

Scoruri finale medii per curs

Key Factors Driving Industrialization

Valorile pentru fiecare curs sunt foarte apropiate, toate în jurul valorii de 3, ceea ce indică o uniformitate a performanței sau a unei metrici relevante (poate engagement, satisfacție, dificultate medie etc.) în toate cursurile.

Python Basics are valoarea cea mai mare (3.045), iar Web Development cea mai mică (2.97), dar diferența este mică.

Media generală pentru toate cursurile este 3.01, confirmând un nivel similar al acestei măsurători în cadrul celor 5 cursuri. Valorile medii ale scorului (sau metricii respective) pentru fiecare curs indică o relativă omogenitate între acestea, cu valori în jurul lui 3, variind ușor între cursuri.

Evaluarea cursurilor de către participanți (Feedback Score)

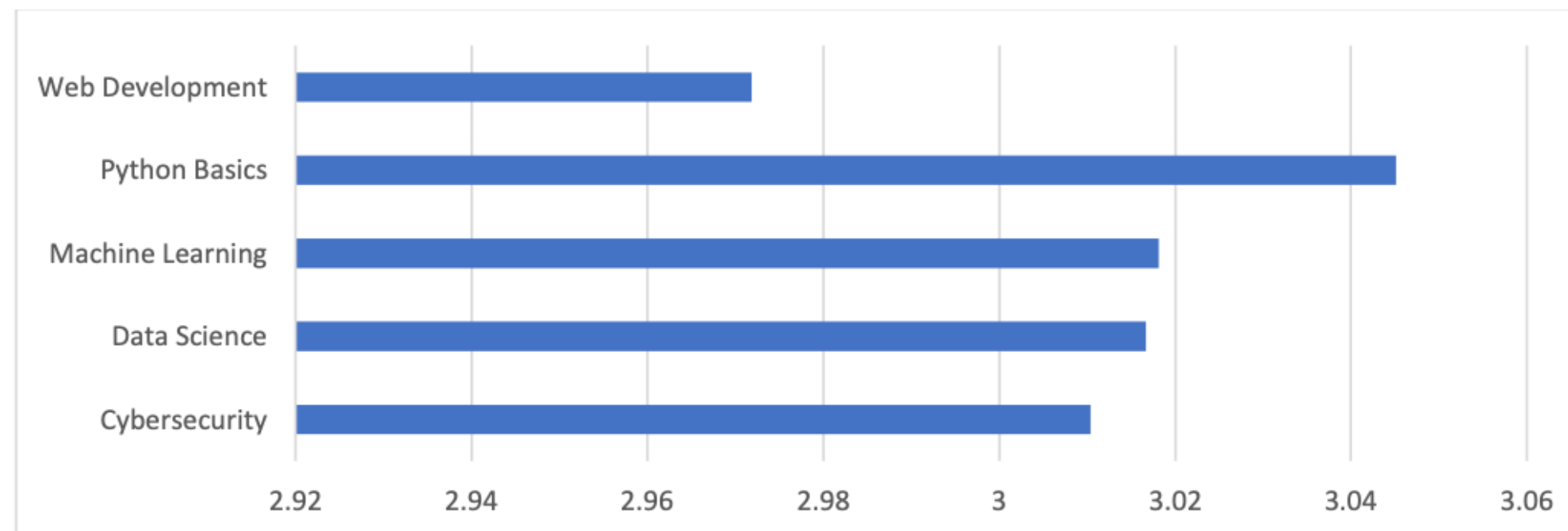


Figura 15 Scorul mediu de feedback pentru cursuri

Discuții, implicații & concluzii



Principalele concluzii

Rata de abandon și performanțele academice nu diferă semnificativ în funcție de stilul de învățare sau engagement. Participarea și comportamentele de învățare sunt echilibrate între grupuri, indicând o platformă educațională echitabilă și incluzivă.

Limitări

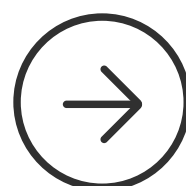
Limitările studiului includ: utilizarea unei variabile agregate pentru engagement, lipsa analizelor statistice inferențiale, folosirea unui set de date sintetice și absența datelor longitudinale, ceea ce restrânge validitatea și aplicabilitatea concluziilor în contexte reale.

Oportunități

Viitorul educației digitale trebuie să se bazeze pe date reale de comportament, cu sisteme adaptive și indicatori dinamici pentru intervenții timpurii, combinând stiluri de învățare cu comportamente observate pentru o personalizare eficientă.

Implicațiile rezultatelor pentru societate

Rezultatele arată că platformele educaționale online pot fi medii incluzive și echitabile, oferind șanse egale cursanților din diverse medii. Lipsa diferențelor semnificative între genuri, niveluri educaționale și stiluri cognitive indică că digitalizarea educației poate reduce inegalitățile.



Recomandări



În primul rând, este esențială dezvoltarea unor sisteme adaptive mai sofisticate, care să utilizeze metode de învățare automată pentru a detecta tipare comportamentale relevante și pentru a personaliza intervențiile educaționale în timp real. În al doilea rând, se impune colectarea și analiza de date reale și longitudinale, din platforme existente precum Coursera, Moodle sau edX, pentru a valida și extinde rezultatele obținute în acest studiu. În final, se recomandă integrarea unor mecanisme de feedback și intervenție personalizată, care să reacționeze rapid la semnalele comportamentale ale cursanților, astfel încât sprijinul pedagogic să fie oferit proactiv și relevant.



Echipa

Thank You!



Laura Bianca Ciobotaru



Călin Ignat