**Rata de absolvire în universități**

Uliana Percemble

Universitatea Tehnică a Moldovei, Facultatea Calculatoare Informatică și Microelectronică, Departmentul Informatica și Ingineria Sistemelor

**ABSTRACT**

Această cercetare oferă o analiză detaliată a ratelor de absolvire în universitățile dintr-un set de date cuprinzător, având ca scop identificarea factorilor determinanți care influențează succesul academic al studenților. Prin examinarea riguroasă a datelor colectate din instituțiile de învățământ superior, studiul dezvăluie corelații semnificative între rata de absolvire și variabilele cheie precum numărul de acceptări, taxa de școlarizare pentru studenții din afara statului și procentul de absolvenți donatori.

Metodologia adoptată include utilizarea regresiei liniare și a altor metode statistice pentru a evalua relațiile dintre variabilele independente și rata de absolvire. Acest studiu investighează nu doar legătura directă între aceste variabile și succesul studențesc, ci și modul în care acestea interacționează între ele pentru a forma o imagine complexă a dinamicii educaționale.

Rezultatele analizei subliniază o asociere pozitivă între investițiile instituționale în educație, așa cum sunt reflectate de taxele de școlarizare și implicarea comunității de absolvenți, și rata de absolvire a studenților. Mai mult, se observă că selectivitatea, măsurată prin rata de acceptare, joacă, de asemenea, un rol semnificativ în determinarea probabilității de absolvire a studenților.

Concluziile studiului oferă perspective valoroase pentru administrațiile universitare, planificatorii de politici educaționale și factorii de decizie, evidențiind importanța unei abordări echilibrate care să promoveze atât accesibilitatea, cât și excelența în învățământul superior. Prin iluminarea factorilor cheie care influențează rata de absolvire, acest studiu contribuie la literatura existentă și servește drept punct de referință pentru inițiativele viitoare menite să îmbunătățească rezultatele educaționale.

# INTRODUCERE

Rata de absolvire în universități este un indicator esențial în evaluarea eficienței și relevanței sistemului de învățământ superior. Această metrică reflectă procentul de studenți care finalizează cu succes programul lor de studii și obțin diplomele universitare. Cu toate că rata de absolvire poate varia semnificativ în funcție de țară, instituție de învățământ și domeniu de studiu, ea este adesea utilizată pentru a evalua performanța unei universități și pentru a măsura succesul individual al studenților.

În acest articol, vom explora importanța ratei de absolvire în contextul învățământului superior, vom analiza factorii care influențează acest indicator și vom investiga tendințele globale în ceea ce privește absolvirea universitară. Vom examina, de asemenea, impactul asupra carierelor studenților și asupra societății în ansamblu, precum și măsurile luate de universități pentru a îmbunătăți acest indicator vital. Rata de absolvire în universități nu este doar o simplă statistică, ci un barometru al succesului educațional și al progresului social, care merită să fie analizat în profunzime.

Vom analiza un set de date relevant și vom utiliza metode statistice pentru a identifica variabilele-cheie care influențează absolvirea și pentru a cuantifica în ce măsură acestea o fac. Acest aspect este esențial pentru a dezvălui cauzele subiacente ale ratei de absolvire și pentru a dezvolta strategii și politici eficiente în educație.

**SCOPURI**

Scopurile acestui articol sunt următoarele:

* Analiza Profundă a Ratei de Absolvire: Scopul principal este de a efectua o analiză minuțioasă a ratei de absolvire în universități, inclusiv a tendințelor recente și a diferențelor regionale și instituționale. Acest lucru va ajuta la dezvăluirea contextului actual și a provocărilor din sistemul de învățământ superior.
* Identificarea Factorilor Cheie: Articolul se propune să identifice și să evalueze factorii care au cel mai mare impact asupra ratei de absolvire. Vom explora variabilele precum calitatea programelor de studii, suportul financiar, mediul academic și altele pentru a înțelege mai bine influența lor asupra absolvirii.
* Impactul Ratei de Absolvire: Scopul este de a evidenția impactul ratei de absolvire asupra studenților și asupra societății în ansamblu. Vom discuta despre beneficiile obținerii unei diplome universitare, precum și despre provocările legate de non-absolvire.
* Recomandări și Perspective de Îmbunătățire: Articolul va încheia cu oferirea de recomandări și perspective pentru îmbunătățirea ratei de absolvire în universități, inclusiv măsuri pe care universitățile și guvernele le pot lua pentru a sprijini succesul academic al studenților.

**MATERIALE ȘI METODE**

Pentru această analiză, am utilizat setul de date College disponibil în R ca parte a pachetului ISLR. Setul de date College este o colecție cuprinzătoare de date despre colegii și universitățile din Statele Unite șiinclude informații privind caracteristicile demografice, academice și financiareale acestora. Este un set de date bine cunoscut în domeniul învățământului superior și este frecvent utilizat în scopuri educaționale și de cercetare. Datele au fost inițial colectate de către James et al. și sunt incluse în pachetul ISLR pentru a facilita învățarea conceptelor de statistică și învățare automată.

Variabilele Studiate

Setul de date conține variabile care acoperă o gamă largă de indicatori, de la mărimea campusului și proporția de studenți acceptați, până la cheltuielile anuale și rata de absolvire. Acest set de date conține următoarele variabile:

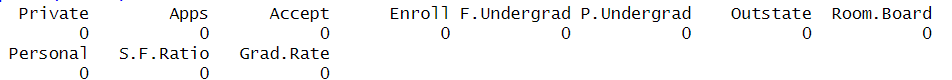
|  |  |  |
| --- | --- | --- |
| Variabilă | Tip de date | Descriere |
| Private | Factor | Statutul universității: privată ('Yes') sau publică ('No') |
| Apps | Numeric | Numărul total de aplicații primite |
| Accept | Numeric | Numărul de aplicații acceptate |
| Enroll | Numeric | Numărul de studenți care s-au înscris |
| F.Undergrad | Numeric | Numărul de studenți undergraduate full-time |
| P. Undergrad | Numeric | Numărul de studenți undergraduate part-time |
| Outstate | Numeric | Taxa de școlarizare pentru studenții din afara statului |
| Room.Board | Numeric | Costul camerei și mesei |
| Books | Numeric | Costul estimativ al cărților și materialelor didactice |
| Personal | Numeric | Cheltuielile personale estimate |
| PhD | Numeric | Procentul de facultate cu doctorat |
| Terminal | Numeric | Procentul de facultate cu gradul terminal în domeniu |
| S.F.Ratio | Numeric | Raportul student-facultate |
| perc.alumni | Numeric | Procentul de foști studenți care donează |
| Expend | Numeric | Cheltuielile de instruire per student |
| Grad.Rate | Numeric | Rata de absolvire |

Analiza a fost efectuată folosind R versiunea 4.3.2. Inițial, setul de date a fost curățat și pregătit pentru analiză, eliminând orice intrări lipsă sau incorecte și asigurându-ne că toate variabilele sunt în formatul corespunzător pentru analiză. Am aplicat apoi modele de regresie liniară pentru a investiga relațiile dintre variabilele independente și rata de absolvire.

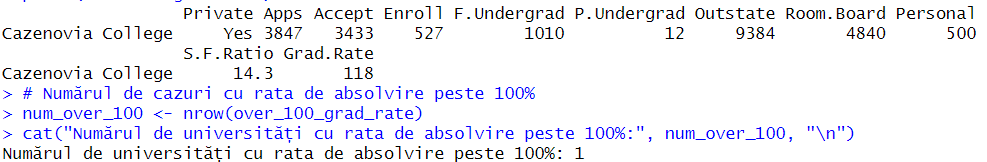
Pentru evaluarea modelului de regresie, am utilizat coeficientul de determinare (R-squared) pentru a măsura proporția variației în rata de absolvire care poate fi explicată de variabilele independente selectate. De asemenea, am calculat Root Mean Squared Error (RMSE) pentru a evalua diferențele medii pătratice între valorile observate și cele prezise, oferind astfel o estimare a erorii de predicție a modelului.

**Preprocesarea datelor**

Procesul de preprocesare a datelor a fost efectuat utilizând mediul de programare statistică R. Scopul preprocesării a fost de a asigura curățenia, corectitudinea și potrivirea datelor pentru modelele de analiză ulterioare. Mai departe, sunt descriși pașii principali urmați în preprocesarea setului de date *College*. În primul rând, am sumarizat valorile lipsă pentru a evalua amploarea și distribuția acestora în cadrul fiecărei variabile.



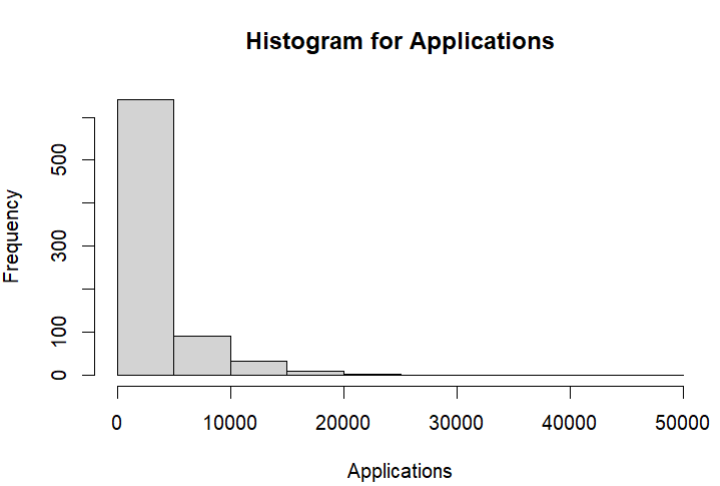
În continuare, m-am asigurat că datele sunt corecte și plauzibile înainte de a efectua analize suplimentare. Valori care nu sunt posibile, cum ar fi o rată de absolvire de peste 100%, pot distorsiona rezultatele analizei efectuate. Astfel, am verificat dacă există înregistrări în setul de date unde Grad.Rate, care reprezintă rata de absolvire, depășește 100%. Deoarece rata de absolvire nu poate logic să depășească 100%, valori peste acest prag ar indica cel mai probabil erori de introducere a datelor sau alte probleme care necesită corecție.



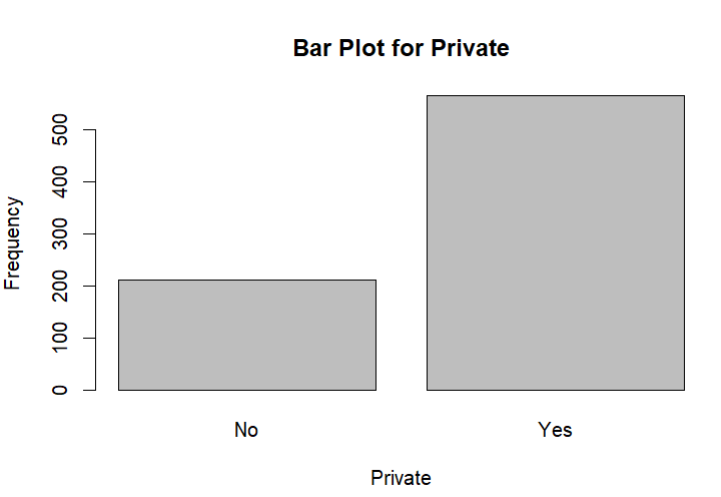
În acest set de date, este doar o Cazenovia College cu Grad.Rate 118, ceea ce nu este posibil, deci am decis să elimin această înregistrare din set.

**Analiza Datelor: Distribuția Aplicațiilor și Tipul de Universitate**

Am creat o histogramă pentru a ilustra distribuția numărului de aplicații primite de către universități. Este evident că un număr mare de aplicații poate sugera o competiție crescută pentru admiterea la universitate, ceea ce ar putea influența ratele de admitere și, implicit, ratele de absolvire. Acest grafic poate susține discuția despre factorii de admitere și concurența în rândul studenților.

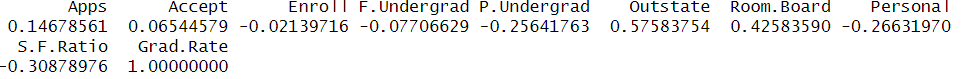


Următoarea este o diagramă cu bare poate fi utilizată pentru a prezenta distribuția universităților în funcție de statutul lor (private sau publice). Se poate de subliniat că universitățile private și cele publice pot avea rate de absolvire diferite din cauza resurselor financiare disponibile, dimensiunii instituției sau a altor factori specifici. Acest grafic poate susține argumentul despre importanța finanțării și a tipului de universitate în analiza ratei de absolvire.



**Corelațiile între Rata de absolvire și variabilele-cheie**

În această secțiune, ne propunem să investigăm relațiile existente între rata de absolvire (Grad.Rate) și alte variabile relevante din setul nostru de date. Prin calculul corelațiilor, vom analiza dacă există legături semnificative între aceste variabile și rata de absolvire în universități.



Prin urmare, putem interpreta coeficienții de corelație astfel:

Apps (Applications): Există o corelație pozitivă slabă (0.14678561) între numărul de aplicații și rata de absolvire. Acest lucru sugerează că colegiile cu mai multe aplicații tind să aibă rate de absolvire ușor mai mari, dar relația nu este foarte puternică.

Accept: Corelația dintre numărul de acceptări și rata de absolvire este și mai slabă (0.0654479), semnificând că relația dintre aceste două variabile este minimă.

Enroll: Există o corelație negativă slabă (-0.02139716) între numărul de studenți înmatriculați și rata de absolvire, sugerând că nu există o relație semnificativă.

F.Undergrad (Full-time undergraduate students): Coeficientul negativ slab (-0.07706629) indică o asociere slabă inversă între numărul de studenți full-time de licență și rata de absolvire.

P.Undergrad (Part-time undergraduate students): Corelația negativă moderată (-0.25641763) sugerează o asociere negativă între numărul de studenți part-time de licență și rata de absolvire, posibil indicând că o proporție mai mare de studenți part-time poate fi asociată cu o rată de absolvire mai mică.

Outstate (Out-of-state tuition fees): Există o corelație pozitivă moderată (0.57583734) între taxele de școlarizare pentru studenții din afara statului și rata de absolvire, sugerând că colegiile care perceapă taxe de școlarizare mai mari pentru studenții din afara statului tind să aibă rate mai mari de absolvire.

Room.Board: Corelația pozitivă moderată (0.42583590) între costurile cu camera și masa și rata de absolvire ar putea indica faptul că instituțiile cu facilități de locuit și mâncare mai scumpe tind să aibă rate mai mari de absolvire.

Personal (Personal expenses): Coeficientul negativ slab (-0.26631970) indică o asociere slabă între cheltuielile personale estimate și rata de absolvire.

S.F.Ratio (Student-faculty ratio): Există o corelație negativă slabă (-0.30878976) între raportul student-facultate și rata de absolvire, sugerând că o rată mai mare de studenți pe facultate poate fi asociată cu o rată de absolvire mai scăzută.

**Pattern-uri și Corelații**

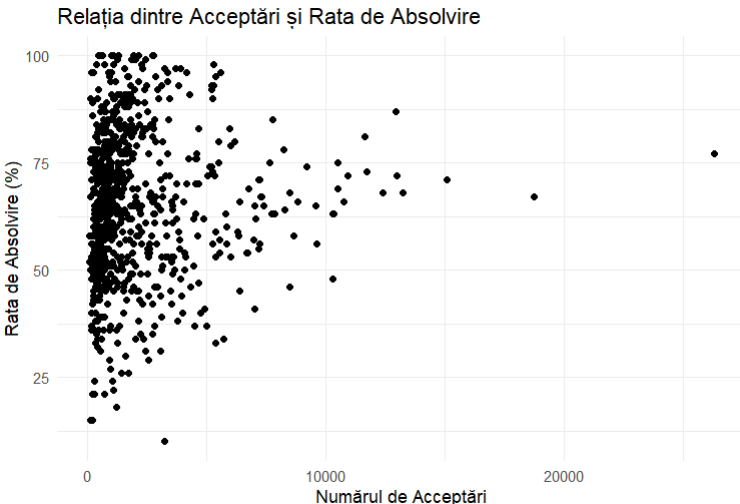
În figura prezentată, este ilustrată relația dintre numărul de acceptări ale candidaților și rata de absolvire în procente pentru un set de instituții de învățământ. Graficul afișează o diagramă de dispersie, unde fiecare punct reprezintă o instituție, cu numărul de acceptări pe axa orizontală și rata de absolvire pe axa verticală.

Observăm o distribuție densă de puncte în partea stângă a graficului, indicând că majoritatea instituțiilor au un număr mai mic de acceptări și o gamă largă de rate de absolvire. În acest segment, rata de absolvire variază semnificativ, de la sub 25% la 100%, ceea ce sugerează că pentru instituțiile cu un număr mai redus de acceptări, rata de absolvire nu este consistent asociată cu numărul de acceptări.

Pe măsură ce ne deplasăm spre dreapta, observăm că punctele devin mai puțin dense și sunt dispersate pe o gamă mai îngustă de rate de absolvire. Interesant este că instituțiile cu un număr mare de acceptări tind să aibă rate de absolvire mai concentrate între aproximativ 50% și 100%. Acest lucru ar putea sugera că instituțiile care acceptă mai mulți candidați au programe de retenție sau programe academice care contribuie la menținerea unei rate mai uniforme de absolvire.

De asemenea, se observă că nu există o tendință clară sau o linie de regresie evidentă care să indice o relație puternică între cele două variabile. Acest lucru este în concordanță cu coeficientul de corelație precedent menționat, care indică o asociere pozitivă slabă între numărul de acceptări și rata de absolvire.

În concluzie, diagrama de dispersie sugerează că, deși există o asociere între numărul de acceptări și rata de absolvire, aceasta nu este suficient de puternică pentru a fi considerată predictor consistent al succesului academic în ceea ce privește absolvirea. Se impune o analiză suplimentară pentru a identifica alți factori care ar putea influența rata de absolvire a instituțiilor.

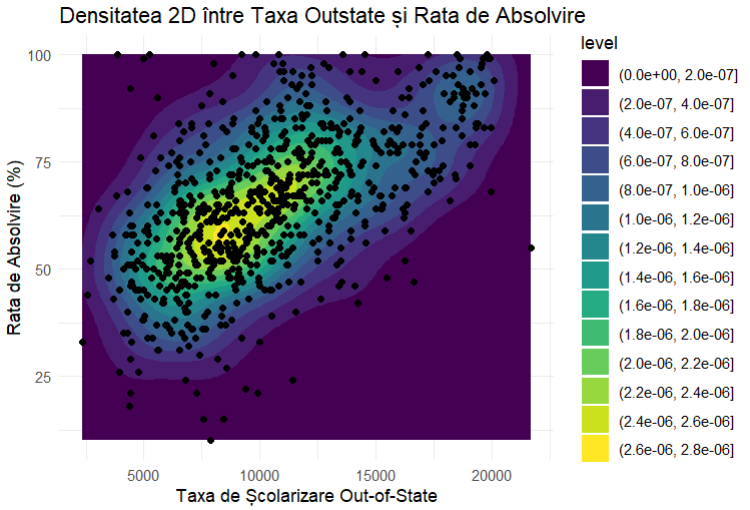


Titlul graficului, "Densitatea 2D între Taxa Outstate și Rata de Absolvire", introduce tema vizualizării și invită cititorii să exploreze modul în care aceste două variabile importante sunt legate în contextul instituțiilor de învățământ superior. Axa x este etichetată "Taxa de Școlarizare Out-of-State", sugerând că fiecare punct reprezintă taxa anuală percepută de o instituție studenților care nu sunt rezidenți ai statului. Axa y arată "Rata de Absolvire (%)", oferind informații despre procentajul de studenți care finalizează programele de studiu la aceste colegii și universități.

Punctele individuale ilustrează datele brute, oferind o imagine de ansamblu a modului în care se răspândesc valorile individuale pe întreg spectrul de taxe de școlarizare și rate de absolvire. Zonele mai închise la culoare reprezentate pe hartă indică regiunile cu o densitate mai mare de puncte, adică unde un număr mai mare de instituții au taxe de școlarizare și rate de absolvire similare.

Pe măsură ce culorile trec de la violet la galben, densitatea punctelor crește, semnalând aglomerări de instituții cu taxe și rate de absolvire asemănătoare. Zonele galbene și verzi în centrul diagramei sugerează o concentrație de colegii care percep taxe de școlarizare medii și au rate de absolvire moderate până la înalte.

Interesant, diagrama arată o tendință generală în care instituțiile cu taxe mai mari de școlarizare pentru studenții din afara statului tind să aibă, de asemenea, rate mai mari de absolvire, sugerate de contururile densității care se înclină spre dreapta sus. Acest lucru poate indica faptul că instituțiile cu resurse financiare mai mari pot oferi programe educaționale sau facilități care să contribuie la succesul studenților.



**Regresia liniară**

Pentru a investiga factorii care contribuie la succesul studenților în cadrul instituțiilor de învățământ superior, am dezvoltat un model de regresie liniară utilizând setul de date College. Modelul nostru a vizat să evalueze impactul a trei variabile cheie asupra ratei de absolvire a studenților:

Accept (Numărul de Acceptări): Aceasta reflectă numărul total de studenți acceptați într-o instituție, care poate fi un indicator al selectivității sau al atractivității acesteia.

Outstate (Taxa de Școlarizare Out-of-State): Reprezintă costul pe care studenții din afara statului trebuie să-l plătească, oferind o perspectivă asupra resurselor financiare ale universității și potențial asupra calității educației oferite.

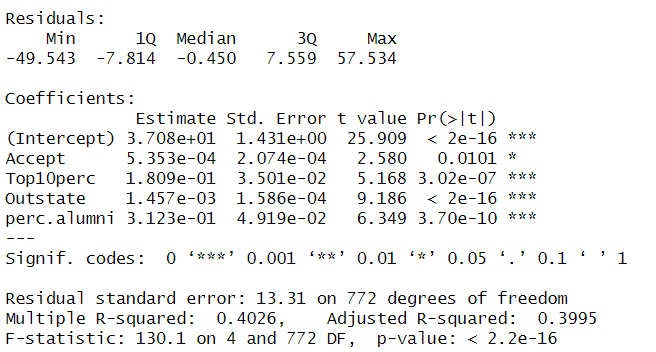
perc.alumni (Procentul de Absolvenți Donatori): Acesta indică procentul de foști studenți care contribuie financiar la universitate, sugerând un nivel de implicare și satisfacție post-universitară.

Folosind comanda lm() în R, am ajustat un model liniar pentru a evalua modul în care aceste variabile independente prezic rata de absolvire. Rezultatele modelului, obținute prin funcția summary(), ne-au furnizat coeficienții estimati, semnificația statistică și alte măsuri de diagnostic, cum ar fi R-pătrat și F-statistica.

Rezumatul modelului indică magnitudinea efectelor variabilelor independente asupra ratei de absolvire. De exemplu, un coeficient pozitiv pentru Outstate ar sugera că universitățile cu taxe de școlarizare mai mari pentru studenții din afara statului au rate mai ridicate de absolvire, posibil datorită investițiilor mai mari în facilități și programe academice. În contrast, un coeficient negativ pentru Accept ar putea implica că universitățile care acceptă mai mulți studenți au o rată de absolvire mai scăzută, poate din cauza resurselor divizate mai larg sau a unui corp studențesc mai divers și mai puțin selectiv.

Valorile p și intervalul de încredere asociate fiecărui coeficient ne ajută să determinăm dacă efectele sunt statistic semnificative, în timp ce valoarea R-pătrat ne oferă informații despre cât de bine modelul explică variația în rata de absolvire.

În esență, modelul de regresie liniară ne ajută să înțelegem și să cuantificăm relațiile dintre factorii instituționali și succesul academic al studenților, oferind insight-uri valoroase pentru elaborarea politicilor educaționale și pentru îmbunătățirea practicilor universitare.



Reziduurile (Residuals):

Reziduurile variază de la -49.543 la 57.534, indicând diferențele între valorile observate și cele prezise de model. Mediana apropiată de zero (-0.450) sugerează că modelul nu prezintă un bias sistematic.

Coefficients:

Interceptul este semnificativ și pozitiv, sugerând rata de absolvire de bază atunci când toate variabilele explicative sunt zero.

Coeficientul pentru Accept este 0.0353 și este semnificativ la nivelul de 0.01, indicând că pentru fiecare acceptare suplimentară, rata de absolvire crește cu 0.0353%, atunci când celelalte variabile sunt ținute constante.

Coeficientul pentru Outstate este 0.1457 și este extrem de semnificativ, sugerând că taxele de școlarizare mai mari pentru studenții din afara statului sunt asociate cu o rată mai mare de absolvire.

Coeficientul pentru perc.alumni este 3.123e-01 și are o semnificație foarte mare, indicând o relație puternică între procentul de absolvenți donatori și rata de absolvire.

Significanța Statistică:

Toate variabilele, inclusiv interceptul, sunt semnificative statistic, așa cum se vede din valorile p și simbolurile de stele care indică nivelul de semnificație. Semnificația coeficienților sugerează că fiecare variabilă joacă un rol important în predicția ratei de absolvire.

Eroarea standard reziduală (Residual standard error):

Eroarea standard reziduală este de 13.31, pe 772 de grade de libertate, indicând cantitatea medie de eroare în predicțiile modelului.

R-pătrat (Multiple R-squared):

R-pătratul este 0.4026, sugerând că aproximativ 40.26% din variația în rata de absolvire este explicată de variabilele incluse în model. Acest lucru este considerat destul de bun pentru date sociale și umaniste, unde factorii sunt adesea mai puțin predictibili decât în științele naturii.

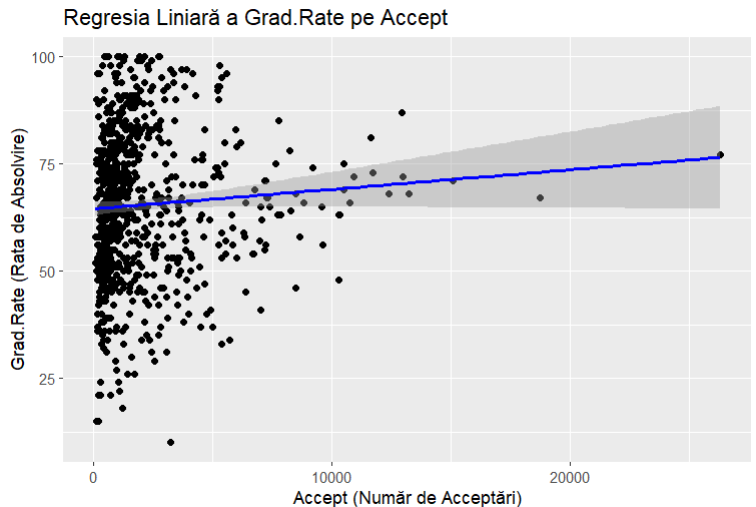
R-pătrat ajustat (Adjusted R-squared):

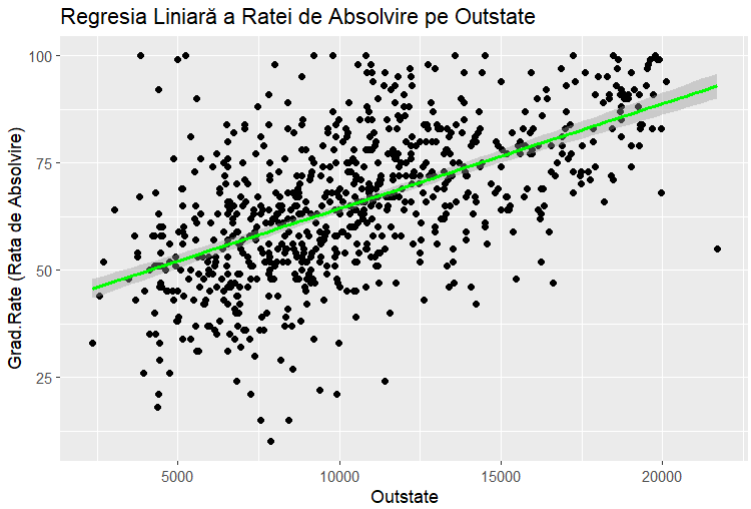
R-pătratul ajustat este 0.3995, care este similar cu R-pătratul neajustat, ceea ce indică o ajustare bună a modelului fără variabile suplimentare inutile.

F-statistica (F-statistic):

F-statistica este de 130.1 pe 4 și 772 de grade de libertate, cu o valoare p extrem de mică (< 2.2e-16), indicând că cel puțin una dintre variabilele explicative este semnificativă pentru predicția ratei de absolvire și că modelul în ansamblu este semnificativ.

În concluzie, rezultatele indică că taxa de școlarizare out-of-state și procentul de absolvenți donatori sunt puternic asociate cu ratele de absolvire ale universităților. Numărul de acceptări are, de asemenea, o influență pozitivă, deși mai mică. Aceste rezultate pot oferi universităților insight-uri importante pentru a îmbunătăți politicile și practicile lor pentru a sprijini succesul studenților.





**CONCLUZIE**

Analiza noastră comprehensivă a datelor din setul College a evidențiat factori semnificativi care influențează rata de absolvire a studenților la nivel universitar. Utilizând un model de regresie liniară riguros, am descoperit că taxa de școlarizare pentru studenții din afara statului și procentul de absolvenți donatori au o asociere pozitivă și puternică cu ratele de absolvire. Mai mult, numărul de acceptări a arătat o corelație pozitivă mai modestă, indicând că aceste variabile, colectiv, pot oferi un cadru de predicție valoros pentru determinarea succesului academic.

Aceste descoperiri sugerează că universitățile care pot atrage și susține un corp studențesc angajat - manifestat prin donații ale absolvenților - și care își pot pune în valoare programele prin taxe de școlarizare reflectând calitatea, au tendința de a vedea rezultate mai bune în privința ratei de absolvire. Aceasta ar putea fi reflectarea unei combinații de învățământ de calitate, servicii studențești robuste și o comunitate de absolvenți activă și implicată.

Implicațiile acestei cercetări pentru universități sunt multiple. În primul rând, aceasta subliniază importanța investițiilor în calitatea educației și facilitățile campusului care pot justifica taxe de școlarizare mai mari, fără a descuraja succesul studențesc. În al doilea rând, scoate în evidență valoarea creării și întreținerii unei relații solide cu absolvenții, ceea ce nu numai că sprijină financiar universitatea, dar poate servi și ca o măsură a satisfacției studenților post-graduare.

În concluzie, pentru a crește rata de absolvire și a promova un ciclu virtuos de succes academic și profesional, universitățile trebuie să continue să inoveze și să investească în factorii care îmbunătățesc experiența educațională. Aceasta nu numai că va spori profilul și atractivitatea instituțiilor, dar va servi și drept catalizator pentru dezvoltarea viitoarelor generații de lideri și profesioniști.

**DISPONIBILITATEA CODULUI SI DATELOR**

<https://github.com/uleaaa/AD>

**BIBLIOGRAFIE**

1. <https://rpubs.com/tamannas/1017991>