

# Factors related to academic performance among engineering students: a descriptive correlational research study

Jaime A. Gutiérrez-Monsalve <sup>a</sup>, Juan Garzón <sup>b</sup>, Difariney Gonzalez-Gómez <sup>c</sup> & Angela M. Segura-Cardona <sup>d</sup>

<sup>a</sup> Grupo de Investigación en Educación Superior (GIES). Escuela de Graduados, Doctorado en Epidemiología y Bioestadística, Universidad CES, Medellín, Colombia. jagutierrez33@gmail.com

<sup>b</sup> Facultad de Ingenierías, Grupo de Investigación en Ingenierías Multidisciplinar (GIMU), Universidad Católica de Oriente UCO, Rionegro, Antioquia, Colombia, fgarzon@uco.edu.co

<sup>c</sup> Grupo de Investigación en Demografía y Salud. Universidad de Antioquia, Medellín, Colombia, difariney.gonzalez@udea.edu.co

<sup>d</sup> Grupo de investigación en Epidemiología y Bioestadística, Universidad CES. Medellín, Colombia, asegura@ces.edu.co

Received: February 6<sup>th</sup>, 2023. Received in revised form: May 25<sup>th</sup>, 2023. Accepted: June 23<sup>th</sup>, 2023.

## Abstract

Academic performance (AP) is a proper pedagogical strategy to determine acquisition of learning outcomes. Monitoring is essential for decision-making regarding accompanying plans and improving educational quality. Saber-Pro is a standardized test applied by the Colombian Government to establish the ability in quantitative, English, critical reading, citizen comprehension, and written communication. The main objective is to determine which sociodemographic, pedagogical, and institutional factors will arise in AP. A canonical discriminant analysis was used to classify 100% of the student, and the risk of low AP was estimated. The model supports variables that explain potentiating risk factors associated with AP in engineering students according to Saber-Pro.

**Keywords:** academic performance; dropout; engineering education; logistic regression.

# Factores asociados al rendimiento académico en ingeniería: un estudio correlacional descriptivo

## Resumen

El Rendimiento Académico Universitario (RAU) es un indicador adecuado para determinar el grado de adquisición de los resultados de aprendizaje. Su seguimiento y gestión son fundamentales para la toma de decisiones sobre estrategias de acompañamiento y mejora de la calidad educativa. Saber-Pro es una prueba estandarizada aplicada por el Gobierno colombiano para establecer el RAU de los estudiantes universitarios en razonamiento cuantitativo, inglés, lectura crítica, comprensión ciudadana y comunicación escrita. Se pretende determinar los factores sociodemográficos, pedagógicos e institucionales relacionados con el bajo RAU. Se utilizó un análisis discriminante canónico que clasificó al 100% de los estudiantes con bajo RAU. La regresión logística permitió establecer las variables que explican los factores de riesgo potenciales del bajo RAU en estudiantes de ingeniería. Los resultados permiten a los gestores universitarios conocer las variables a mejorar para mejorar las competencias evaluadas en Saber-Pro.

**Palabras clave:** rendimiento académico; deserción; enseñanza en ingeniería; regresión logística.

## 1 Introduction

Academic performance (AP) could be defined as the value attributed to the learning outcomes of students in each subject area compared to the level of knowledge expected from their peers. AP is an indicator of university assurance and considers pedagogical, institutional, sociodemographic -

past AP, and psychosocial variables [1,2]. Dropout caused by AP is a common indicator to estimate student achievement. According to national regulations, each university determines the evaluative criteria to give an assessment of AP considering the weighted average in each assignment, the number of credits enrolled, the number of subjects passed, or another effectiveness indicator in a specific period [3,4].

**How to cite:** Gutiérrez-Monsalve, J.A., Garzón, J., Gonzalez-Gómez, D. and Segura-Cardona A.M., Factors related to academic performance among engineering students: a descriptive correlational research study. DYNA, 90(227), pp. 35-44, July - September, 2023.

Unlike graduation and dropout rates, AP can be measured in the early stages of the educational process and has different ways of quantifying (average grades per subject, rate of courses earned per semester). Other advantages of AP could be measured continuously, monitored at any time of university trajectory, and is a good predictor of dropout [5-7]. In general, many studies are interested in determining the factors that affect AP to generate explicative and predictive models that allow early intervention before students drop out [5-7].

In higher education context, it is usual to identify AP as a learning result. In fact, AP can be classified into two categories: immediate, determined by the achievement of students in their academic trajectory until they obtain their professional qualifications. Mediate refers to the learning results or achievement obtained through applying the competencies received in the student trajectory and determines the effectiveness of labor incorporation [8]. This type of mediate AP is frequently measured by government standardized tests in most countries. In Colombia this kind of test is called Saber-Pro [9].

Saber-Pro is presented by all Colombian university students. This test is proposed by the Colombian Institute for Education Evaluation (ICFES) to measure the mediate AP acquired throughout their trajectory. According to Colombian Law, Saber-Pro was instituted as compulsory and included evaluation of verbal, quantitative, and English proficiency, critical reading, and written communication, as well as evaluating the subject area of knowledge for several fields. Like Saber-Pro, Graduate Record Examination (GRE) evaluates and predicts mediate AP in verbal, quantitative, and analytics abilities for admit students in many United States higher education institutions [10,11]. Scholastic Assessment Test (SAT) and American College Test (ACT) are two other examples of standardized tests widely used by student students, and they are mandatory to access postgrads and labor organizations in many countries worldwide [12].

University quality manifests a broad relationship with academic results of standardized tests, this kind of examination could indicate accomplished objectives related to skills development and learning results at the end of the academic program, establishing a comparison between institutions and programs over time [13,14]. Therefore, Saber-Pro constitutes an objective mediate AP indicator that help to know how each program or university is positioned concerning its reference group [15].

Learning outcomes acquired from the development of academic activities are relevant indicators when evaluating students and the academic programs in which they are enrolled. To this, some Latin American higher education institutions have implemented strategies measuring variables associated with AP according to standardized tests. Nevertheless, these studies have represented a significant advance in establishing criteria for improving the quality of higher education and academic performance, impacting the productive and social environment [16-18].

Satisfactory AP is achieved by the interaction of different variables, including the knowledge acquired by students according to the learning process, pedagogy, didactics, and course organization. Pedagogical, sociodemographic, and

institutional factors are essential to classify students with differential mediate AP [19-21].

From the pedagogical factor, mediate AP is understood from the effectiveness of all the teaching strategies used by professors, evaluation, and didactic material, as well as the procedures that professors followed to cover all the learning objectives and contents [22]. From institutional factors, university structures are included, as well as their influence on the particularities of students. The program of choice, number of students per course, maximum number of credits enrolling, availability of scholarships or subsidies, the policies, structure, and functionality of each educational program, number of students per course, number of times a subject is repeated, academic level, and the program of choice would be relevant [1,23].

Sociodemographic factor concentrates on variables that allow students to be classified by differential traits, define the social and cultural guidelines where the student develops, and can affect and mediate AP. These include socioeconomic income, sex, proceeding region, high school graduation, parents' educational level, employment status, access to information and consultation resources, previous AP [20,24-26].

Different studies have concluded that nations' quality of life and future sustainability depend on technological progress and the development of engineering knowledge and techniques [27,28]. For this reason, there is a sustained effort by higher education institution to develop strategies for recruiting and retaining students in science, technology, engineering, and mathematics. Much of this effort is based on predictive and explicative models of AP [29,30]. Engineering, specially in areas like computer, industrial, environmental and electronic play an important role in deploying sustainable, innovative and nature-based solution to deliver Sustainable Development Goals (SDG's) according to 2030 agenda.

In this research, results of Saber-Pro 2019 from environmental, industrial, computer, and electronics engineering were integrated with academic institutional databases of a private university of Colombia. This research aims to identify and quantify the risk associated with low AP according to Saber-Pro test. We proposed a model to support predicting and estimating potentiate risk factors of AP in engineering students. Until this research, these variables have not yet been explored to explain and predict student AP using an integration of academic institutional databases and mandatory standardized tests.

## 2 Method

To determine factors associated with Academic Performance (AP), a retrospective case-control study according to a quantitative approach was carried out. Cases were defined as students with low AP in Saber Pro. According to sociodemographic factors, variables such as gender (female and male), nature of the high school (public or private), region of origin, program in which they are enrolled, and age were considered. Within the institutional factor, variables like having a subsidy or scholarship, number of subjects taken per semester, Grade Point Average (GPA)

in the midterm, follow-up, and final grades were considered. At the same time, the number of times a student repeats an assignment, credits approval, and the semester in which they presented Saber-Pro were considered. In pedagogical factors, type of contract, age, salary, level of study, and qualification scores of professors in pedagogy, management, institutional identity, and didactics were considered. According to students' previous AP, the GPA of students in subjects related to citizen comprehension (CC), writing communication (WC), critical reading (CR), and quantitative reasoning (QR) during their academic trajectory also were included.

The sample consisted of 77 students from environmental, industrial, computer, and electronics engineering in a private university located in the east of Colombia who presented Saber-Pro tests in 2019. Students with scores below 159.3 in Saber-Pro were classified with low AP; this score corresponds to the average established for high-quality accredited universities in Colombia in 2019 [31]. The scope of this study was correlational and explanatory. Correlational in order to determine the relation of variables considered in all factors with AP according to Saber Pro. Explanatory in order to explain and predict risk measures of all AP-related variables.

A bivariate statistical analysis was initially carried out. In the case of qualitative variables, Chi-squared independence was used, and in quantitative variables, ANOVA or Kruskal-Wallis were performed if data were distributed parametrically or not. For the definitive logistic model, those variables with *p-values* lower than .25 were chosen to be entered according to Hosmer-Lemeshow criteria. All analysis was performed with 95% of confidence.

Canonical discriminant analysis was proposed to validate if a correct classification was achieved between students with low and average AP. The classification rules were assigned to identify 100% of students with low AP. All statistical component was developed in R software using the package *candisc* [32].

### 2.1 Logistic regression approach

A logistic regression model was performed to identify the association between factors and AP. This model allowed calculating Odds Ratio (OR) for each variable according to low AP.

The model's predictive quality was evaluated using the goodness-of-fit criteria and area under the curve (AUC) based on sensitivity and specificity. The model was set with 2/3 of the data and the remaining 1/3 for validation.

In the logistic model, the dependent variable  $Y_i$  (defined as the event that a student has low AP) is calculated according to expression (1):

$$Y_i = \left( \frac{\Pi_i}{1 - \Pi_i} \right) \quad (1)$$

Where:

$\Pi_i$ : is the probability that student  $i$  will experience low academic performance, and it is the value of the dependent variable in case  $i$  for the regression coefficients  $\alpha$  and  $\beta$ .

$$Y_i = \alpha + \beta X_i \quad (2)$$

Expressed as regression, the logistic model can be represented as:

$$\Pi_i = \left( \frac{1}{1 + e^{-(\alpha + \beta X_i)}} \right) \quad (3)$$

The risk ratio OR indicate the probability that a student has low AP concerning one who has normal AP according to the value assumed by the independent variable  $X_i$ . This value constitutes a level of the relative effect of each variable in AP. According to this, the OR can be estimated from the logistic regression model as follows:

$$\begin{aligned} OR_i &= \log \left( \frac{\Pi_i}{1 - \Pi_i} \right) = \left( \frac{1}{\frac{e^{-(\alpha + \beta X_i)}}{1 + e^{-(\alpha + \beta X_i)}}} \right) \\ &= \log \left( \frac{1}{e^{-(\alpha + \beta X_i)}} \right) = \alpha + \beta X_i \end{aligned} \quad (4)$$

For logistic regression, all hypotheses were tested with a significance of 0.05.

We hypothesized that variables associated with sociodemographic, pedagogical, institutional, and past academic performance were included as risk variables and are significantly correlated with AP according to the Saber-Pro test.

## 3 Results

Most of the students included had small incomes (about 85%). More than 78% of them came from public schools, and nearly 7% live far from the campus in rural areas. Finally, 10% came from other Colombian regions but live in local university residencies. As of 2019-1, the engineering faculty has four programs with 536 students, where 15.41% study environmental engineering, 46.64% in computer engineering, 29.85% in industrial engineering, and 7.65% in electronics engineering.

### 3.1 Classification of university students

Considering the mean score of high-quality accredited universities in Colombia, 46 (59.74%) were in low AP, and the remaining 31 (40.26%) obtained normal AP.

A descriptive analysis is reported from pedagogical (Tables 1 and 2), institutional (Table 3), and sociodemographic factors (Table 4). Statistical differences were tested with low AP according to the Hosmer-Lemeshow methodology, where variables with a *p-value* < .25 or less are considered significant [33].

Pedagogical variables associated with low AP are if chair professors directed courses ( $p < .001$ ), older ( $p = .045$ ), paid less for their work ( $p = .002$ ), without a master's degree ( $p = .23$ ), and with low scores in didactics ( $p = .17$ ). Students with low scores in subjects of quantitative reasoning (QR), English (EN), critical reading (CR), citizen comprehension (CC) and written communication (WC) ( $p < .001$ ) were associated with low AP (Table 1 and 2). These variables were selected to be included in the multivariate logistic regression model.

Table 1. Bivariate analysis of AP according to variables in pedagogical factor

Pedagogical factor				
Variable	Professor Contract		Test	
	Low AP	Normal AP	Sig.	Test
Full time	56.16%	43.84%		
Variable	Professor Degree		Test	
	Low AP	Normal AP	Sig.	Test
Professional	60.40%	39.60%	.23	Chi-2
Specialization	59.96%	40.03%		
Master's degree	56.39%	43.61%		
Ph.D.	63.12%	36.87%		
Variable	Professor Qualification		Test	
	Low AP	Normal AP	Sig.	Test
High in pedagogy	58.28%	41.71%	.41	Chi-2
Low in pedagogy	61.68%	38.31%		
High in course management	58.49%	41.51%	.97	Chi-2
Low in course management	58.69%	41.30%		
High in institutional identity	58.51%	41.48%	.93	Chi-2
Low in institutional identity	57.97%	42.03%		
High in didactics	58.06%	41.93%	.17	Chi-2
Low in didactics	62.83%	37.17%		
Variable	Types of subjects		Test	
	Low AP	Average AP	Sig.	Test
Lost in citizen comprehension (CC)	58%	41.71 %	<.001**	Chi-2
Lost in writing communication (WC)	53.07 %	46.93 %		
Lost in critical Reading (CR)	60.63 %	39.36 %		
Lost in quantitative reasoning (QR)	57.24 %	42.76 %		

Source: own

Table 2. Quantitative analysis of AP according to variables in pedagogical factor

Variable	Quantitative variables		Test	
	Low AP	Average AP	Sig.	Test
Professor age (Me ± MAD)	38 ± 6	38 ± 5	.049*	Kruskal-Wallis
Professor salary US\$ (Me ± DAM)	635 ± 400.28	741.94 ± 414	.002**	Kruskal-Wallis
Overall semester professor qualification (Me ± DAM)	86.8 ± 4.02	87.6 ± 4.195	<.001**	Kruskal-Wallis

\*\*\* and \*\* indicates statistical differences with .05, .02 y >.01 of significance.

Source: own

According to institutional factors, variables associated with low AP were student has not had any subsidy ( $p < .01$ ), has enrolled in few concurrent subjects/semester ( $p < .01$ ), has

obtained low partial ( $p < .01$ ), and final grades ( $p < .01$ ). Student who has high repetition rate ( $p < .01$ ), has been in an advanced semester at the time of presenting Saber-Pro ( $p = .18$ ) and has had a lower number of subject approval ( $p < .01$ ) has been associated with low AP (Table 3).

In the sociodemographic factor, the female sex ( $p = .02$ ) comes from rural and foreign regions ( $p < .01$ ), enrolled in industrial engineering ( $p = .02$ ), and being older ( $p < .01$ ) are variables associated with low AP. Critical remarks suggest high student competencies in CC, WC, CR, and QR acquired during their student trajectory increment AP. According to Saber-Pro, students with poor subject results are likelier to have low AP (Table 4).

Table 3. Bivariate analysis of AP according to variables in institutional factor.

Variable	Institutional factors		Test	
	Low AP	Normal AP	Sig.	Test
Have a subsidie	53.41%	46.59%	>0.01**	Chi-2
Not have a subsidy	63.60%	36.40%		
Concurrent subject/semester (Me ± DAM)	5 ± 1	5 ± 1	0.061	Kruskal-Wallis
Cancelation subject/semester (Me ± DAM)	1 ± 1	1 ± 0	>0.01**	Kruskal-Wallis
Midterm exam scores (Me ± DAM)	3.7 ± 0.7	4 ± 0.6	>0.01**	Kruskal-Wallis
Follow-up score (Me ± DAM)	4.2 ± 0.4	4.2 ± 0.3	>0.01**	Kruskal-Wallis
Final exam score (Me ± DAM)	3.8 ± 0.8	4 ± 0.7	>0.01**	Kruskal-Wallis
Repetition rate (Me ± DAM)	1 ± 0	1 ± 0	0.029**	Kruskal-Wallis
Level or semester	9 ± 1	9 ± 0	0.18	Kruskal-Wallis
Credits per semester (Me ± DAM)	10 ± 3	10 ± 3	0.27	Kruskal-Wallis
Approval credits per semester (Me ± DAM)	9.5 ± 3.5	9 ± 3	0.41	Kruskal-Wallis
Average semester grades ( $\bar{x} \pm Sd$ )	4 ± 0.52	4.232 ± 0.47	0.048**	ANOVA
Accumulated average grades (Me ± DAM)	3.8 ± 0.2	3.9 ± 0.1	0.027**	Kruskal-Wallis
Approved subject per semester (Me ± DAM)	4 ± 1	5 ± 1	>0.01**	Kruskal-Wallis

\*\*\* and \*\* indicates statistical differences with .05, .02 y >.01 of significance.

Source: own

Table 4. Bivariate analysis of AP according to variables in sociodemographic factor.

Variable	Sociodemographics factors		Test	
	Low AP	Normal AP	Sig.	Test
Female	68.50%	31.50%	0.02	Chi-2
Male	57.80%	42.20%		
Public School	59.90%	40.10%	0.47	Chi-2
Private School	63.00%	37.00%		
From metropolitan Region	65.70%	34.30%	>0.01	Chi-2

From rural nearest areas	69.30%	30.70%		
From colombian foreign regions	100%	0%		
From nearest big metropoli - Medellin	31.10%	68.90%		
Study environmental engineering	59.20%	40.80%		
Study computer engineering	45.20%	54.80%	0.02	Chi-2
Study industrial engineering Electronics engineering	78.40%	21.60%		
Age (Me ± DAM)	22 ± 4.45	21 ± 2.97	>0.01	Kruskall-Wallis

\*\*, \*\* and \*\*\* indicates statistical differences with .05, .02 y >.01 of significance.  
Source: own

**3.2 Validation of the cut-off points and discriminant variables between low and normal academic performance**

Figs. 1 - 3 show a discriminant analysis for pedagogical, institutional, and sociodemographic factors. This analysis allows us to determine which variables significantly classify between low and average AP. These models also explain 100% of the variance in their first dimension.

Canonical discriminant analysis in pedagogical factors identifies 56.71% of students with low or average (Normal) AP (Fig. 1). Variables that best discriminate against low AP were if the professor had no full-time contract (Prof.Vinculate). This result could be associated with their limited time to accompany students during extracurricular hours. Another variable that favors low AP is the professor's age (Prof.Age): older had more students with low AP. These results suggest that teaching methodologies must be updated to new disciplinary trends that older teachers did not incorporate. Another consideration is that older teachers must be more empathic with their younger students.

Nevertheless, variables that favor average AP in pedagogical factors (Fig.1) are high professor salary, high overall professor qualification, and high grades in pedagogy and didactics. According to students, good grades in a subject associated with writing communication (WC) and quantitative reasoning (QR) favored average AP. These results suggest that competencies of WC and QR acquired during student trajectory could be decisive in improving AP. A high professor's salary could be associated with motivation and responsibility, and students recognize the effort and competence of an excellent professor to give them good qualifications [34].

Canonical discriminant in sociodemographic factors was able to identify 65.34% of students with low AP. According to Fig. 3, students from rural regions, being females, studying industrial engineering and older, are associated with low AP in engineering. These results suggest reviewing industrial engineering curricula, competencies, and teaching support is a strategy to improve AP in this specific program. Finally, older working students, have families, and have less time for their academic activities could be associated with low AP [35].

The canonical discriminant model of institutional factor allows for identifying the 57.48% of students with low and normal (Normal) AP (Fig. 2). In accordance, institutional variables like have not a subsidy, repetition rate, and low accumulated, midterm, final, and follow-up grades are associated with low AP. The number of concurrent subjects and credits enrolled per semester (Credits and Concurrent. Subj) and if they had in a lower semester when students present Saber Pro were associated with low AP. This result suggests that providing economic or subsidy support could be an effective strategy to improve in a meaningful manner AP. Maintaining and promoting subsidies implies they make extraordinary efforts to have a good AP in their student trajectory. In the same way, low AP is also associated with high grades in university. According to Saber Pro, these results suggest that better performance in student trajectory is associated with good AP.

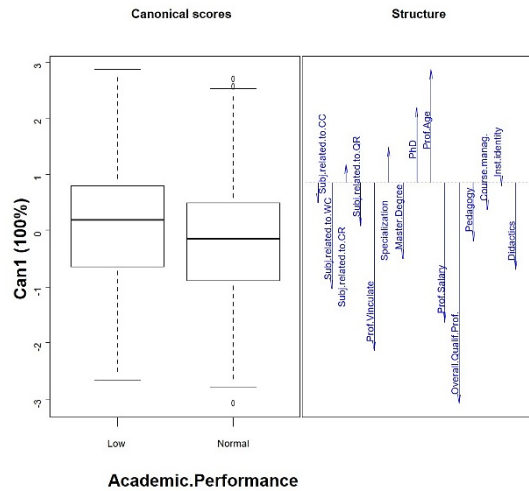


Figure 1. canonical discriminant analysis of variables influencing AP according to pedagogical factor. This model allows identifying by cross-validation the 56.71 % of all student students.  
Source: own

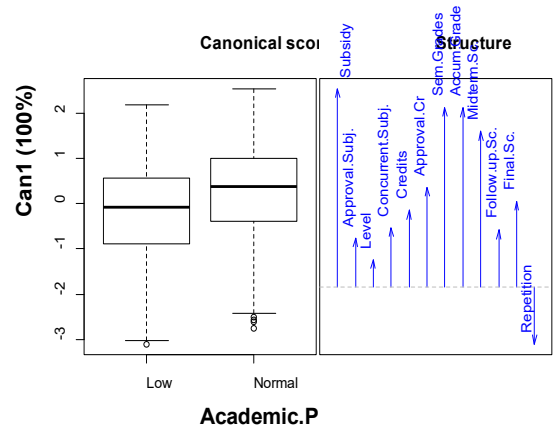


Figure 2. canonical discriminant analysis of variables influencing AP according to institutional factor. This model allows identifying by cross-validation the 57.48% of all students.  
Source: own

### 3.3 Logistic regression to estimate risk ratios

Based on logistic regression performed as multivariate analysis and Hosmer-Lemeshow criteria using a stepwise methodology, it was possible to select those variables that best explain low AP and risk association Odds Ratio (OR) (Table 5). OR establishes a proportional relationship of the variable with AP; that is, if for females, the adjusted OR (adj.OR) obtained is 1.3, this implies that women present 1.3 times more feasibility of having low AP than men. In consequence, the female will be a risk variable. On the other hand, if the adj.OR is less than 1, the variable could be considered protective.

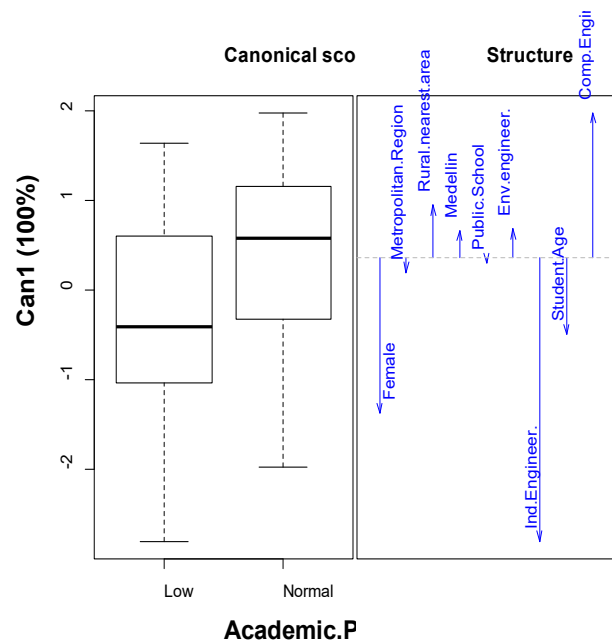


Figure 3: canonical discriminant analysis of variables influencing AP according to sociodemographic factor. This model allows identifying by cross-validation the 65.34% of all students. Source: own

The logistic regression in Table 5 allows correctly classifying 73.56% of students. In addition, this model identifies 82.49% with low and 61.20% of normal AP by cross-validation. According to this, the model performs well in explaining variables affecting AP according to Saber Pro for a cut-off probability of 0.5.

Regarding pedagogical factor, the variable that most affects AP is professor degree: professors with master's and specialization have students with better AP in Saber-Pro (adj.OR 0.59 [ 0.38,0.92]) and (adj.OR 0.63 [0.4,0.99]) respectively, compared with professional (adj. OR 0.56 [0.33,0.95]) and Ph.D. which was the reference category. These results suggest that the investment made by professors and universities in postgraduate training contributes significantly to improve AP.

Logistic regression also presented risk variables considering the institutional factor (Table 4). It was observed that having a subsidy (adj.OR 0.53 [0.43,0.65]), a high ratio of approved credits per semester (adj.OR 0.43 [0.2,0.94]),

good grades in subjects related to critical reading at university trajectory (adj.OR 0.43 [0.2,0.94]), good grades in midterm (adj.OR 0.86 [0.77,0.96]) and final exams (adj.OR 0.99 [0.81,0.99]) becomes protective factor. On the other hand, a student who was at a high level or semester when presenting Saber Pro was considered at risk of low AP (OR 1.51 [0.95, 2.42]).

According to Saber-Pro, this result suggests that good performance in the university training trajectory guarantees average AP. In the sociodemographic factor, males (adj.OR 0.75 [0.59,0.95]) who had been enrolled in computer engineering (adj.OR 0.51 [0.4,0.67]) were considered protective. Older age at the moment of presenting Saber Pro (adj.OR 1.03 [1.00,1.06]), comes from a rural region (adj.OR 3.13 [1.88,5.21]), had graduated from a public high school (adj.OR 1.42 [1.09,1.84]), and had been enrolled in industrial engineering program (adj.OR 2.92 [2.35,3.65]) were risk factors of low AP according to logistic regression.

Table 5. Crude and adjusted Odds Ratio (OR) of variables related to AP in Saber-Pro.

Factor	Variable	Adj. OR 95% CI in parenthesis	p	Variable description
Pedagogical	Professor professional – PhD	0.56 (0.33, 0.95)	.032	Protector
	Professor specialization - PhD	0.63 (0.4, 0.99)	.044	Protector
	Professor master's degree – PhD	0.59 (0.38, 0.92)	.02	Protector
	Professor salary	1 (1, 1)	<	-----
	Have subsidy	0.53 (0.43, 0.65)	<	Protector
Institutional	Good grades in Critical Reading (CR)	0.43 (0.2, 0.94)	.033	Protector
	Good grades in Midterm Exams	0.86 (0.77, 0.96)	.007	Protector
	Good grades in the Final Exam	0.9 (0.81, 1)	.041	Protector
	Be in High level or semester	1.51 (0.94, 2.42)	.088	Risk
	Male – Women	0.75 (0.59, 0.95)	.016	Protector
Sociodemographic	Student age	1.03 (1, 1.06)	.081	Risk
	Metropolitan region – Medellin	1.73 (1.29, 2.33)	<	Risk
	Rural nearest area – Medellin	3.13 (1.88, 5.21)	<	Risk
	Public – private school	1.42 (1.09, 1.84)	.009	Risk
	Computer eng. – Environment eng.	0.51 (0.4, 0.67)	<	Protector
	Industrial eng. - Environment eng.	2.92 (2.35, 3.65)	<	Risk

Source: own



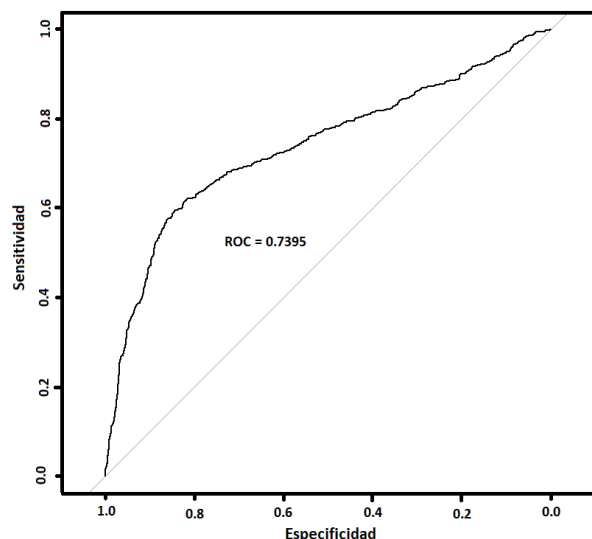


Figure 4. ROC curve establishes that the logistic model is explanatory.  
Source: own

Finally, the Area Under Curve (AUC) of the Receiver Operating Characteristic (ROC) (Fig. 4) was 0.739, which establishes that this model is only for explanation, and it is not possible to make accurate predictions. It is necessary to make modifications and adjust the model with variables that provide a complement to obtain an adequate ROC. Cognitive and non-cognitive variables in psychosocial factors could be considered.

#### 4 Discussion

The classification criteria based on the average general result of 159.3 in high-quality accredited universities presented consistency when approaching the research since the discriminant analyzes were able to explain 100% of the variance after validating the classification. According to Saber Pro, pedagogical variables such as writing communication (WC) and quantitative reasoning (QR) acquired during the student trajectory could be decisive in having a normal AP. This result is similar to the results by Gardner and Milne (2011) and Peiss et al. (2013). They demonstrated that high levels of communication skills (listening, speaking, and writing communication) in the early stages of university trajectory positively correlate with academic performance and achievement in students [37,38].

According to QR as a predictor of average AP and success in student trajectory, Mckee and Caldarella (2016) conclude that grades in subjects associated with math at high school were all strong predictors of student AP [39]. In the same way, Mutawah et al. (2018) demonstrated that student motivated with math and science in high school shows good AP when they are admitted to a student program [40]. Following these findings, universities could concentrate their effort on reinforcing competencies in subjects related to WC and QR in the curricula in all student trajectories to improve AP in Saber Pro.

Another variable that explains differences between

students with low AP is the type of professor contracted full time, high professor salary, high qualification (master's degrees and PhD.), and better grades in professor's performance are related to normal AP. According to this, much research argues that good AP in students is related to the quality of the professor. Master's degrees and Ph.D. provide students with valued knowledge and technical support that is important to labor success and student motivation [41]. The results presented here are coherent with Rodríguez et al. (2019). They declared that the more significant investment in higher education of university professors is positively correlated with the more excellent knowledge acquired in all areas, which impacts better learning results for students [42].

Paying professors more for their work and the type of professor contract could not affect academic performance in big university scenarios. However, regional or low-income scenarios could be associated with labor motivation. Finally, students recognize that the effort and competence of a good professor give them good grades in the professor's performance [43-45].

Interesting results were obtained when we showed that industrial engineering, female sex, older students, and region are associated with low AP in students, otherwise, proceeding high school (public or private) was not. This result agrees with Iregui et al. (2007) and Gutierrez-Monsalve et al. (2021); they studied the efficiency of student education in Colombia and compared AP programs of many students with several variables, including if students come from public or private high schools. The main results of their research conclude that private schools could benefit from the more favorable income condition. However, when public schools are assumed to be in environments equivalent to private ones, there are no statistical differences between the AP from both types of high schools [17,46]. In the case of gender, Lin et al. (2014) and Monroy-Varela et al. (2022) reported significant associations between gender, AP, and retention indicators in engineering students in favor of women but not men. Our research is one of the few studies that report low AP in favor of women [47,48]. In the case of the region, Adeoiti et al. (2020) concluded that the proceeding region of students can affect academic performance in a university context. They concluded that rural students could have low AP compared to those from urban areas [49]. Other reports argue that most students of lower age assimilated and performed better in class than older ones. In fact, this may be due to older students almost work, have families and have less time for their academic activities, but if they work together with younger, their AP could be similar [50-52]. To our knowledge, there are no reports that publish the association between the AP considering student programs such as industrial, computer, and environmental engineering together.

Data from the risk ratio (OR) in logistic regression showed that only 16 of 56 variables had a high association with AP. According to the model, only nine have a protective effect, and the remaining seven have a risk effect. A subsidy is one of the most relevant variables that act as a protective variable. Providing economic support, especially for a student with low income, could be improved in a meaningful

manner AP. Subsidy or scholarships implies that students making more significant efforts to have a good AP in their trajectory, so students with subsidy have better AP than others without [53–56].

Finally, these results present a broad relationship with the findings of Gutiérrez-Monsalve et al. (2021), where significant associations of low AP with the student's grade trajectory have subsidy and repetition rate.

## 5 Conclusions

According to factors related to AP in Saber Pro, professors with Ph.D. or master's degrees and well-paid were variables related to good AP. According to Saber Pro, professors with high qualifications in didactics and pedagogy had fewer students with low AP.

One interesting result is associated with student trajectory AP measure as a grade in assignments related to Citizen Comprehension (CC), Writing Communication (WR), Critical Reading (CR), and Quantitative Reasoning (QR). Students with good performance in their academic trajectory show better AP in Saber Pro. This result implies that curricula development through the career could be an essential predictor of AP.

According to sociodemographic and institutional variables, having a subsidy, a high ratio of approved credits (associated with past AP), and good grades in their university trajectory were related to good AP in Saber-Pro. Have subsidy, be male, enrolled in computer engineering, younger, comes from the urban region and had graduated from private high school, improved AP of Saber-Pro.

Based on this entire study, this model is explanatory. Therefore, prediction criteria cannot be established based on the results of the ROC curve. More variables that are interrelated with AP must be included to build a predictive model, especially in the psychosocial factor.

The study's main limitation is associated with the absence of cognitive and noncognitive variables that demonstrate a good prediction of Academic performance in engineering students [3].

## Acknowledgements

The authors would like to recognize the effort of Universidad Católica de Oriente and Universidad CES to provide data, valuable technical support, and statistical analysis to complete the results.

The authors declare no conflict of interest.

## References

- [1] Montero-Rojas, E. y Villalobos-Palma, J., Factores institucionales, pedagógicos, psicosociales y sociodemográficos asociados al rendimiento académico en la Universidad de Costa Rica: un análisis multinivel, RELIEVE. Revista Electrónica de Investigación y evaluación educativa, 13(2), pp. 215-234, 2007.
- [2] Namoun, A. and Alshantit, A., Predicting student performance using data mining and learning analytics techniques: a systematic literature review, Appl. Sci., 11(1), art. 237, 2020. DOI: <https://doi.org/10.3390/app11010237>
- [3] Ridgell, S.D., and Lounsbury, J.W., Predicting academic success: general intelligence, 'Bigfive' personality traits, and work drive., Coll. Stud. J., 38(4), pp. 607-618, 2004, [Online]. Available at: <http://search.ebscohost.com/login.aspx?direct=true&db=ehh&AN=15493261&lang=es&site=ehost-live&scope=site>
- [4] Miguéis, V.L., Freitas, P.J., Garcia, V., and Silva, A., Early segmentation of students according to their academic performance: a predictive modelling approach, Decis. Support Syst., 115, pp. 36-51, 2018. DOI: <https://doi.org/10.1016/j.dss.2018.09.001>
- [5] O'Connor, M.C., and Paunonen, S.V., Big Five personality predictors of post-secondary academic performance, Pers. Individ. Dif., 43(5), pp. 971-990, 2007. DOI: <https://doi.org/10.1016/j.paid.2007.03.017>
- [6] Zuniga-Jara, S., Pizarro-Díaz, R., Escudero-López, M.E. y Honores-Marín, G., Rendimiento Académico universitario y conductas alimentarias, Form. Univ., 11(2), pp. 57-64, 2018. DOI: <https://doi.org/10.4067/S0718-50062018000200057>
- [7] Kahn, J.H., and Nauta, M.M., Social-cognitive predictors of first-year college persistence: the importance of proximal assessment., Res. High. Educ., 42(6), pp. 633-652, 2001. DOI: <https://doi.org/10.1023/A:1012225510213>
- [8] Hellas, A., Ihtola, P., Petersen, A., Ajanovski, V.V., Gutica, M., Hynninen, T., and Liao, S.N., Predicting academic performance: a systematic literature review, in Proceedings companion of the 23<sup>rd</sup> annual ACM conference on innovation and technology in computer science education, 2018, pp. 175-199. DOI: <https://doi.org/10.1145/3293881.3295783>
- [9] Ayán, M.N.R. y Díaz, M.Á.R., Indicadores de rendimiento de estudiantes universitarios: calificaciones versus créditos acumulados, Rev. Educ., 355, pp. 467-462, 2011. DOI: 10-4438/1988-592X-RE-2011-355-033
- [10] Razmi, M.H., Khabir, M., and Tilwani, S.A., Meta-Analysis on the predictive validity of Graduate Record Examination (GRE) general test., Int. J. Lang. Test., 11(2), pp. 51-63, 2021.
- [11] Kuncel, N.R., Hezlett, S.A., and Ones, D.S., A comprehensive meta-analysis of the predictive validity of the graduate record examinations: implications for graduate student selection and performance., Psychol. Bull., 127(1), art. 162, 2001. DOI: <https://doi.org/10.1037/0033-2909.127.1.162>
- [12] Wai, J., Brown, M.I., and Chabris, C.F., Using standardized test scores to include general cognitive ability in education research and policy, J. Intell., 6(3), art. 37, 2018. DOI: <https://doi.org/10.3390/jintelligence6030037>
- [13] Urbano, M.C.T., Aporte de los exámenes de estado, saber pro, al desarrollo curricular y la calidad en dos programas de psicología, MSc. Tesis, Universidad Católica de Colombia, Bogotá, Colombia, 2017. Disponible en: <https://repository.ucatolica.edu.co/bitstream/10983/15028/1/APORTE%20DE%20LOS%20EX%20C3%81MENES%20DE%20ESTADO1.pdf>.
- [14] Mishra, S., Social networks, social capital, social support and academic success in higher education: a systematic review with a special focus on 'underrepresented' students, Educ. Res. Rev., 29, art. 100307, 2020. DOI: <https://doi.org/10.1016/j.edurev.2019.100307>
- [15] Celis, M.T., Jimenez, O. y Jaramillo, J.F., ¿Cuál es la brecha de la calidad educativa en Colombia en la educación básica y en la superior?, Saber Investig., pp. 67-98, 2015.
- [16] Bahamón, J. y Reyes L., Caracterización de la capacidad intelectual, factores sociodemográficos y académicos de estudiantes con alto y bajo desempeño en los exámenes Saber Pro - año 2012, Avances en Psicología Latinoamericana, 32(3), pp. 459-476, 2014. DOI: <https://doi.org/10.12804/apl32.03.2014.01>
- [17] Gutiérrez-Monsalve, J.A., Garzón, J. y Segura-Cardona, A.M., Factores asociados al rendimiento académico en estudiantes universitarios, Form. Univ., 14(1), pp. 13-24, 2021. DOI: <https://doi.org/10.4067/S0718-50062021000100013>
- [18] Castellanos, W.A.O., Perdomo, A.D.M. y Flórez, T.E.R., Factores asociados al rendimiento en las Pruebas Saber Pro en estudiantes de ingeniería civil en universidades colombianas, Innovaciencia Fac. Ciencias Exactas, Físicas y Nat., 2(1), pp. 22-29, 2014. DOI: <https://doi.org/10.15649/2346075x.234>
- [19] Castrillón, J.J.C., Ceron, A.C., Collazos, A.K., Molina, A.M., Osorio, J., Ospina, A. y Zambrano, O.E., Factores que inciden en la motivación académica en un programa de medicina, Manizales, Colombia, 2010,



- Arch. Med., 12(1), pp. 46-61, 2012. DOI: <https://doi.org/10.15649/2346075X.234>
- [20] Castaño, E., Gallón, S., Gómez, K. y Vásquez, J., Análisis de los factores asociados a la deserción y graduación estudiantil universitaria, *Lect. Econ.*, (65), 2006.
- [21] Trussel, J.M., and Burke-Smalley, L., Demography and student success: early warning tools to drive intervention, *J. Educ. Bus.*, 93(8), pp. 363-372, 2018. DOI: <https://doi.org/10.1080/08832323.2018.1496893>
- [22] Gissel, S.T., and Buch, B., A systematic review of research on how students and teachers use didactic learning materials in L1, *Learn. Tech.*, (7), pp. 90-129, 2020. DOI: <https://doi.org/10.7146/lt.v5i7.117281>
- [23] Caballero, B.F., Higher education: factors and strategies for student retention., *HETS Online J.*, 10(2), pp. 1D-1D, 2020.
- [24] Helal, S., Li, J., Liu, L., Ebrahimie, E., Dawson, S., Murray, D.J., and Long, Q., Predicting academic performance by considering student heterogeneity, *Knowledge-Based Syst.*, 161, pp. 134-146, 2018. DOI: <https://doi.org/10.1016/j.knosys.2018.07.042>
- [25] Selvitopu, A., and Kaya, M., A meta-analytic review of the effect of socioeconomic status on academic performance, *J. Educ.*, 0(0), art. 1978, 2021. DOI: <https://doi.org/10.1177/00220574211031978>
- [26] Salazar Uribe, J.C., Lopera-Gómez, C., and Jaramillo-Elorza, M.C., Identification of factors that affect the loss of student status using a logit survival model for discrete time data, *Dyna*, 79(171), pp. 16-22, 2012.
- [27] Al-Sheeb, B.A., Hamouda, A.M., and Abdella, G.M., Modeling of student academic achievement in engineering education using cognitive and non-cognitive factors, *J. Appl. Res. High. Educ.*, 11(2), pp. 178-198, 2019. DOI: <https://doi.org/10.1108/JARHE-10-2017-0120>
- [28] Ridner, S.L., Newton, K.S., Staten, R.R., Crawford, T.N. and Hall, L.A., Predictors of well-being among college students, *J. Am. Coll. Heal.*, 64(2), pp. 116-124, 2016, DOI: <https://doi.org/10.1080/07448481.2015.1085057>
- [29] Jørgensen, U., and Valderrama, A., The politics of engineering professionalism and education, in *Engineering Professionalism*, Brill Sense, 2016, pp. 281-309.
- [30] Veenstra, C.P., Dey, E.L., and Herrin, G.D., A model for Freshman engineering retention., *Adv. Eng. Educ.*, 1(3), p.33, 2009.
- [31] ICFES, Guía de orientación Saber Pro 2019. Disponible en: <https://www.icfes.gov.co/gu%C3%ADas-de-orientaci%C3%B3n-examen-saber-pro> . 2019.
- [32] Friendly, M., Fox, J., and Friendly, M.M., Visualizing generalized canonical discriminant and canonical correlation analysis, R Packag. 'candisc', version 0.6-5, 2013.
- [33] Hosmer Jr, D.W., Lemeshow, S., and Sturdivant, R.X, *Applied Logistic Regression*. Wiley, 2013. DOI: <https://doi.org/10.1002/9781118548387>.
- [34] McCallen, L.S., and Johnson, H.L. The role of institutional agents in promoting higher education success among first-generation college students at a public urban university., *J. Divers. High. Educ.*, 13(4), art. 320, 2020. DOI: <https://doi.org/10.1037/dhe0000143>
- [35] Sothan, S., The determinants of academic performance: evidence from a Cambodian University, *Stud. High. Educ.*, 44(11), pp. 2096-2111, 2019. DOI: <https://doi.org/10.1080/03075079.2018.1496408>
- [36] Millea, M., Wills, R., Elder, A., and Molina, D., What matters in college student success? Determinants of college retention and graduation rates, *Education*, 138(4), pp. 309-322, 2018, [Online]. Available at: <http://search.ebscohost.com/login.aspx?direct=true&db=ejh&AN=130165011&lang=es&site=ehost-live&scope=site>
- [37] Gardner, C.T., Milne, M.J., Stringer, C.P., and Whiting, R.H., Oral and written communication apprehension in accounting students: curriculum impacts and impacts on academic performance, *Account. Educ.*, 14(3), pp. 313-336, 2005. DOI: <https://doi.org/10.1080/06939280500077269>
- [38] Preiss, D.D., Castillo, J.C., Grigorenko, E.L., and Manzi, J., Argumentative writing and academic achievement: A longitudinal study, *Learn. Individ. Differ.*, 28, pp. 204-211, 2013. DOI: <https://doi.org/10.1016/j.lindif.2012.12.013>
- [39] McKee, M.T., and Caldarella, P., Middle school predictors of high school performance: a case study of dropout risk indicators, *Education*, 136(4), pp. 515-529, 2016.
- [40] Al-Mutawah, M.A., and Fateel, M.J., Students' achievement in math and science: how grit and attitudes influence? *Int. Educ. Stud.*, 11(2), pp. 97-105, 2018.
- [41] Singh, S., and Ryhal, P.C., The influence of teachers' emotional intelligence on academic performance with mediating effect of job satisfaction, *J. Educ.*, 0(0), art. 2314, 2021. DOI: <https://doi.org/10.1177/00220574211032314>
- [42] Jimenez, O.R.R., Peña, L.M.C., Toloza, J.A.M., Miranda, B.A.P. y Fúquene, S.F.S., Factores asociados al desempeño en la Prueba Saber Pro-2017: un estudio de los programas de psicología, *Jan. 2019*.
- [43] Goldhaber, D., and Anthony, E., Teacher quality and student achievement. *Urban Diversity Series.*, 2003.
- [44] Johnson, S.M., Kraft, M.A., and Papay, J.P., How context matters in high-need schools: the effects of teachers' working conditions on their professional satisfaction and their students' achievement, *Teach. Coll. Rec.*, 114(10), pp. 1-39, 2012. DOI: <https://doi.org/10.1177/016146811211401004>
- [45] Hoffmann, F., and Oreopoulos, P., Professor qualities and student achievement, *Rev. Econ. Stat.*, 91(1), pp. 83-92, 2009. DOI: <https://doi.org/10.1162/rest.91.1.83>
- [46] Iregui, A.M., Melo, L. y Ramos, J., Análisis de eficiencia de la educación en Colombia, *Rev. Econ. del Rosario*, 10(1), pp. 21-4, 2007.
- [47] Lin, J.J., and Imbrie, P.K., Modeling retention and graduation of engineering students of difference sexes, in: *6<sup>th</sup> First Year Engineering Experience Conference*, TX, August 7, 8, 2014.
- [48] Monroy-Varela, S.E., Gallego-Vega, L.E., Amórtégui-Gil, F.J., Vega-Herrera, J.M., and Díaz-Morales, H., Impact of the COVID 19 pandemic on the student's academic performance at the School of Engineering-Universidad Nacional de Colombia, Bogotá Campus, *Dyna*, 89(222), pp. 38-47, 2022. DOI: <https://doi.org/10.15446/dyna.v89n222.101308>
- [49] Yusuf, F.A., Okanlawon, A.E., and Oladayo, T.R., Investigation into factors affecting students' academic performance in tertiary institutions as expressed by students, *J. Educ. Black Sea Reg.*, 5(2), pp. 62-75, 2020. DOI: <https://doi.org/10.31578/jeds.v5i2.200>
- [50] Adeyemi, A.M., and Adeyemi, S.B., Personal factors as predictors of students' academic achievement in colleges of education in South Western Nigeria, 2014. DOI: <https://doi.org/10.5897/ERR2014.1708>
- [51] Hakimi, S., Hejazi, E., and Lavasani, M.G., The relationships between personality traits and students' academic achievement, *Procedia-Social Behav. Sci.*, 29, pp. 836-845, 2011. DOI: <https://doi.org/10.1016/j.sbspro.2011.11.312>
- [52] Heirany, F., Mahmood, M., Jaber, Z., Nassim, S., and Ehsan, K., Impact of personal characteristics on the academic achievement of the accounting students of Islamic Azad University of Yazd, *Int. J. Acad. Res. Intro Gressive Educ. Dev.*, 2(4), pp. 190-199, 2013.
- [53] Williams, J.M., Bryan, J., Morrison, S., and Scott, T.R., Protective factors and processes contributing to the academic success of students living in poverty: implications for counselors, *J. Multicult. Couns. Devel.*, 45(3), pp. 183-200, 2017. DOI: <https://doi.org/10.1002/jmcd.12073>
- [54] Moreira, G.D.O., Passeri, S., Velho, P.E.N.F., Ferraresi, F., Appenzeller, S., and Amaral, E., The academic performance of scholarship students during medical school, *Rev. Bras. Educ. Med.*, 43, pp. 163-169, 2019. DOI: <https://doi.org/10.1590/1981-52712015v43n3RB20180180>
- [55] Rubin, L.M., and Rosser, V.J., Comparing division IA scholarship and non-scholarship student-athletes: a discriminant analysis., *J. Issues Intercoll. Athl.*, 7, art. 2752, 2014. <http://dx.doi.org/10.34917/4332752>
- [56] Richburg-Hayes, L., Brock, T., LeBlanc, A., Paxson, C.H., Rouse, C.E., and Barrow, L., Rewarding persistence: effects of a performance-based scholarship program for low-income parents, Available SSRN 1353360, 2009. DOI: <https://dx.doi.org/10.2139/ssrn.1353360>

**J.A. Gutiérrez-Monsalve**, is MSc., PhD(c) in epidemiology and biostatistics from Universidad CES. MSc. in Engineering and Process Engineering from the EAFIT University, Medellín, Colombia. Researcher in applied statistics, data science, and machine learning. He is working on

statistical models to improve crop protection and building models to predict and accompany the low academic performance and high-risk dropout students. MSc., from the Escuela de Graduados, Doctorado en Epidemiología y Bioestadística, Universidad CES, Medellín, Colombia, G.I. en Educación Superior  
ORCID: 0000-0003-3118-6824

**J. Garzón**, is PhD. Specialist in Telecommunications, MSc. in Information and Communication Technologies all of them from the Pontifical Bolivarian University, Medellín, Colombia. PhD. in Computer Engineering from Complutense University of Madrid, Spain. Full-Time professor in Universidad Católica de Oriente, Rionegro, Antioquia, Colombia. Skills and expertise related to: teaching and learning, electronics and communication, engineering teaching, e-learning, academic writing, pedagogy and education engineering, curriculum development, higher education. Facultad de Ingenierías, Grupo de Investigación en Ingenierías Multidisciplinar (GIMU), Universidad Católica de Oriente UCO, Colombia,  
ORCID: 0000-0002-0374-8570

**D. Gonzalez-Gómez**, is PhD. Mathematics professional from the Universidad Nacional University of Colombia. Master in Statistical Sciences. Nacional de Colombia, campus Medellín. Dr. in Education, from the University of Antioquia, Colombia. Grupo de investigación en Demografía y Salud. Universidad de Antioquia, Colombia.  
ORCID: 0000-0002-0175-2613

**A.M. Segura-Cardona**, is PhD. Disciplines: public health, epidemiology, infectious diseases, statistics and biostatistics. Skills and expertise: infectious disease epidemiology, tropical diseases, mortality, quality of life, desing sampling, heath risk management, elderly and epidemiology. Grupo de investigación en Epidemiología y Bioestadística, Universidad CES, Colombia,  
ORCID: 0000-0002-0010-1413