Analysis of academic dropout data to estimate relevant variables using computer tools at a Higher Education Institution

Análisis de datos de abandono académico para estimar variables relevantes utilizando herramientas informáticas en una Institución de Educación Superior

http://doi.org/10.17981/culteduso c.16.1.2025.5777

Received: 09-04-2024 Accepted: 10-03-2025 Published: 21-03-2025

Monica-Lizeth Sánchez-Arévalo
Institución Universitaria Politécnico Grancolombiano. Bogotá (Colombia)
mlizsanchez@poligran.edu.co

Roberto Ferro-Escobar
Universidad Distrital Francisco José de Caldas. Bogotá (Colombia)
rferro@udistrital.edu.co

Luisa-Fernanda Chaparro-Sierra
Tecnológico de Monterrey. Monterrey (México)
lchaparr@tec.mx

How to cite this article:
Sánchez-Arévalo, M. L., Ferro-Escobar, R. & Chaparro-Sierra, L. F. (2025). Analysis of academic dropout data to estimate relevant variables using computer tools at a Higher Education Institution. *Cultura Educación y Sociedad, 16*(1), e5777. http://doi.org/10.17981/culteduso c.16.1.2025.5777

Abstract

Student dropout, which refers to students abandoning their curricular programs due to various influences, is a critical concern in education, impacting individuals, institutions, families, and society. This research, centered at a Higher education Institution, examines the Tinto Dropout Model and explores the impact of Problem-Based Learning on student retention. It also analyzes undergraduate dropout rate trends from 2014 to 2020, identifying patterns and disparities among student groups through ANOVA. This comprehensive approach is designed to contribute to informed decision-making and the continuous improvement of higher education in Mexico. The research strives to understand why students discontinue their studies and to develop effective strategies to mitigate dropout rates, ultimately benefiting both the students and the institution.

Keywords: Student Dropout; Data Hub in Higher education; education innovation; ANOVA.

Resumen

La deserción estudiantil, se refiere a los estudiantes que abandonan sus programas curriculares debido a diversas influencias, es una preocupación crítica en la educación, que afecta a los individuos, las instituciones, las familias y la sociedad. Esta investigación examina el Modelo de Abandono definido por Tinto y explora el impacto del Aprendizaje Basado en Problemas en la retención estudiantil. También analiza las tendencias de la tasa de deserción estudiantil de licenciatura de 2014 a 2020, identificando patrones y disparidades entre grupos de estudiantes a través de ANOVA. Este enfoque integral está diseñado para contribuir a la toma de decisiones informadas y a la mejora continua de la educación superior in México. La investigación busca entender por qué los estudiantes interrumpen sus estudios y desarrollar estrategias efectivas para mitigar las tasas de deserción, beneficiando en última instancia tanto a los estudiantes como a la institución.

Palabras clave: Abandono estudiantil; Data Hub en educación superior; innovación educativa; ANOVA.

INTRODUCTION

This research delves into a critical issue in the Mexican educational landscape, as well as in many parts of the world: dropout rates in higher education. Over recent years, the technological institution under study has witnessed students who, having taken the significant step of enrolling and attending at least one semester, unfortunately, end up discontinuing their studies. This phenomenon raises profound questions and demands an in-depth understanding of the dynamics behind these students' decisions.

In a context where education is fundamental to a nation's economic and social development, university dropout represents a challenge that cannot be ignored. Dropout not only impacts individual students but also has ramifications for educational institutions, families, and society. However, the resources invested in the education of these students go to waste, and the potential contributions they could have made to their communities and the country remain unrealized. On the one hand, renowned psychologist Vincent Tinto has proposed an influential model that highlights the pivotal role of educational institutions in student retention. His model, based on interactive variables, experiences, goals, and university commitments, provides valuable insight into understanding why some students choose to abandon their studies while others continue.

Nevertheless, the issue of student dropout is multifaceted and goes beyond existing theories and models. It is influenced by personal, familial, economic, and social factors that intertwine in complex ways. In this regard, the System Dynamics methodology presents itself as a promising tool to address this complexity. By considering the interaction of multiple variables within a dynamic system, it can gain a more comprehensive understanding of the causes and consequences of student dropout.

On the other hand, this research not only aims to identify and analyze the factors contributing to student dropout but also to propose effective solutions. By gaining a better understanding of the dynamics behind students' decisions, it can develop more informed strategies and policies that help to prevent dropouts and promote student retention. Moreover, by focusing on a specific time between 2014 and 2020, this research may shed light on trends and changes in dropout rates over these years, providing relevant insights for future decision-making. This study seeks to contribute to the development of more effective education policies and the ongoing improvement of the quality of higher education in Mexico. Through a rigorous focus on System Dynamics and a comprehensive analysis of the factors influencing student dropout, we hope to provide a solid foundation for informed decision-making at the technological institution under study and, ultimately, within the broader Mexican education system.

LITERATURE REVIEW

General context of dropout

Education is one of the social challenges in a nation's economic and political development. However, there is complexity in providing equal access to education, both in quantity and quality. Recent studies confirm that there is an inverse relationship

between poverty and academic level; individuals with lower levels of education are significantly more likely to face precarious employment and remain in poverty (World Bank, 2024).

Research shows a growing interest in the early detection, based on the individual's cognitive profile, of the vulnerability to academic failures that subsequently trigger dropout. This approach is favorable for all educational institutions that wish to design and implement policies to reduce abandonment rates (Schneider & Preckel, 2017). Student dropout has become a broad field of research, and several authors have dealt with the problem from different areas and aspects. Table 1 shows the authors and the relevant characteristics of their studies on academic desertion, achieving contributions for recognizing and treating variables, allowing universities to create and implement strategies in their action plans to curb the increase in the desertion rate that is disproportionate in the entry-level university programs. Current trends and challenges in higher education imply the need to rethink the mission and its role in society, identifying new priorities for the near future (UNESCO, 2022).

The contextualization of the problem focuses on identifying the set of external entities that can, directly and indirectly, affect the educational process of young people in the world. To enter higher education, a student must have economic capital to a greater extent than what was invested in secondary schooling. Thus, the economic factor is one of many generators of desertion on the university path, (UNESCO, 2016).

Table 1. *Compilation of authors and characteristics of their studies on academic dropout*

Area	Authors (Year)	Characteristics
Biological	Fishbean & Ajzen (1975) Attinasi (1986) Ethington (1990) Bean & Eaton (2001)	Student attributes and personality traits without regard to institutional conditions
Sociological	Spady (1970)	Identification of external factors associated with the institution and the family
Organizational	Kammens (1971) Braxton (2000) Cabrera (1990) Tillman (2002) Himmel (2002)	Dropout based on the characteristics ofthe institution
Economical	Stampen & Hansen (1990) Cabrera Nora & Castañeda (1992) St Jhon, Paulsen & Starkey (1996) Porto et al (2001)	Identification of cost/benefit variables and subsidy targeting
Interactionist model	Vincent Tinto Representative author (1975)	Integration and adaptation of the student are priorities in the development of their process; such interaction enables their professional and academic development.

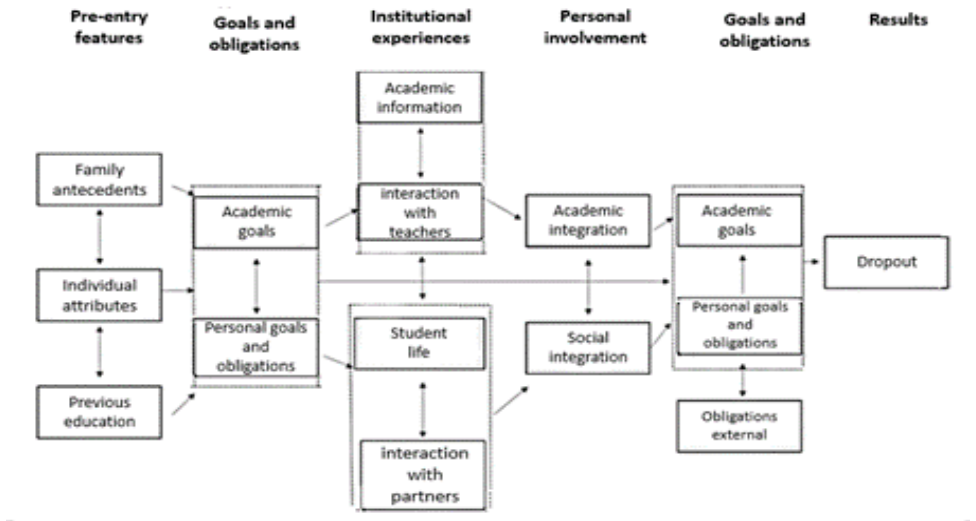
Source: The Authors (2024)

For the Secretaría de Educación Pública de México (SEP), dropout can have causes or origins related to personal, family, socioeconomic, and educational aspects. Dropout is one of the issues that government and higher education institutions face today

(SEP, 2024). It affects the academic, social, economic, and emotional areas where the student, the educational institution, the family, and the government intervene. The psychologist Vincent Tinto proposed the most influential model in desertion issues. This model highlighting the role of educational institutions in the construction of results (Tinto, 1975) was postulated in 1975 as a system of interacting variables, experiences, goals, and university commitments, presenting academic desertion because of varying situations. This contribution is based on various research and considers the individual's duty within the Higher education institution (HEI) of enrollment and its academic objectives as transcendental in the student's permanence or abandonment.

Commitments can be affected by personal aspects, the student's environment, and the experiences of entering university. The factors' correlation results in the student's decision: remain in the institution, change to another, or drop out of the educational system for good. Vincent Tinto proposed a model that is based on the causes of desertion. The model integrates university students' factors, family and individual attributes, previous higher education, teaching methods, institutional support, economic situation, family events, health status, and academic and social integration. Those factors directly influence personal and institutional objectives, which are decisive in deciding to withdraw or remain in the curricular program of the university that the student attends (Tinto, 1975). Figure 1 shows the model.

Figure 1. *Tinto's Model (1975).*



Tinto's model also presents variables that determine the outcome and the experience with the Higher education institution (HEI):

The first set of variables consists of the entry characteristics, where family background, individual aspects, and previous academic education are presented.

The second set concerns commitments on the part of students concerning academic and personal interest's vis-a-vis the institution.

The third refers to the students' experiences within the institution, which are classified into two moments: academic experiences, where the student's performance and interaction with teachers are considered, and social practices with peers and

extracurricular activities, which could alter the student's initial commitments in terms of educational goals and those of the institution.

The fourth set refers to the social and academic integration of the student based on experiences in the institution. The last group considers the objectives, goals, and commitments with the university; these are reflected in the desired result (graduation) or not desired (dropout or lag) (Saldaña Villa & Barriga, 2010). The model shows that the academic and social system interaction is conclusive for student dropout or retention.

The research done by the Universidad Austral de Argentina for instance, concluded the need for early detection of student dropout. Researchers from this university-based their conclusions on their cognitive profile evaluated through educational aptitudes to find those potentially vulnerable to academic failure and having a high probability of dropping out. It became an instrument for counseling policies and specialized training from admission onwards (Corengia, Pita, Mesurado & Centeno, 2013). This proposal is justified by the need to conduct a study with a holistic analysis to model the behavior of dropouts in the technological institution under study to describe the influence of individual factors on the student's decision-making process. This problem is characterized by the difficulty in controlling and intervening in many social and individual variables (Vargas Tamayo, Contreras Bravo & Tristancho Ortiz, 2014). The elaboration involved relevant social, academic, and economic factors, achieving tools for design strategies and institutional policies that benefit students by significantly reducing academic desertion. Identifying and locating students in vulnerable situations can allow universities to provide them with early assistance and avoid the truncation of their academic training. The results are expected to be helpful to benefit the university community and society, serving as a reference source for Higher Education Institutions or researchers as a guide for future projects. Furthermore, the study may generate researchers' interest in investigating impactful academic and social phenomena (Sánchez, 2015).

Recognizing the evolving landscape of higher education, the technological institution under study continues to adapt its educational model to address contemporary challenges and opportunities. Recent initiatives focus on integrating technology-enhanced learning, promoting student well-being, and fostering interdisciplinary collaboration to prepare graduates for success in a rapidly changing world (Narváez et al., 2023). These efforts reflect a commitment to innovation and a dedication to providing students with a transformative educational experience.

Context Dropout in Mexico

In recent reports prepared by the Universidad Autónoma de México and non-governmental organizations, Mexico is the fourth highest country in school dropout. This situation worsened with the effects of isolation and measures to contain the COVID-19 pandemic (UNAM, 2021) (Instituto Nacional de Estadística y Geografía (INEGI), 2020). According to UNICEF Mexico, for 2017, the official figures of the Secretary of Public Education (SEP) indicated that more than 4 million children and adolescents in Mexico did not go to school, and around 600 thousand were at risk of abandoning it (UNICEF). According to a study by the OECD (Organization for Economic Cooperation and Development), the country has the least access to higher education. Only 17% of Mexican

adults over 25 could finish a university degree before the pandemic (Organization for Economic Co-operation and Development-OECD, 2017). According to the INEGI (2022) , the dropout rate for the 2020/2021 school year was like those at the beginning of the century, without any considerable improvement in those 20 years. One of the leading causes of the deterioration of higher education in Mexico is the low public investment of the Mexican State, which invests 7,889 dollars per year for each person enrolled in higher education. According to the Organization for Economic Development Cooperation (OECD), it is a meager sum. The INEGI Institute indicates that the two leading causes of desertion at the university level are dislike or little interest in the study generated by various factors and economic issues (INEGI, 2020).

Regarding the first cause of desertion, the lack of interest is mainly due to unemployment and low pay for university graduates; in Mexico, the hiring rate of professionals with a degree is 79%, and in other OECD member countries, it is 83%. The percentage of employment of people without a university degree is 62%, which becomes a key figure since there is little difference between being unemployed or employed for people with a university degree vs. those without one. On the other hand, the high tuition fees in private universities are also one of the elements that help explain university dropout due to socioeconomic issues. According to INEGI, tuition at a private university range from 14,500 to 111,000 Mexican pesos per semester (INEGI, 2020).

Definition of the technological institution educational model

The educational model of the technological institution under study is a model that has matured and has been continuously updated over time. Here is a set of didactic techniques and methodologies for the teaching-learning process of course topics taught in class. The model is centered on the students. It makes them responsible for their learning, where they delve into the topics and find adequate knowledge to put into practice; here, the professors become guides and facilitators. They make brief presentations, structure concepts, and clarify questions (Tecnológico de Monterrey, 2019). The didactic technique selection depends on the area of knowledge and the students' training level. Collaborative work is an essential soft competency incorporated into all the techniques (Tecnológico de Monterrey, 2018).

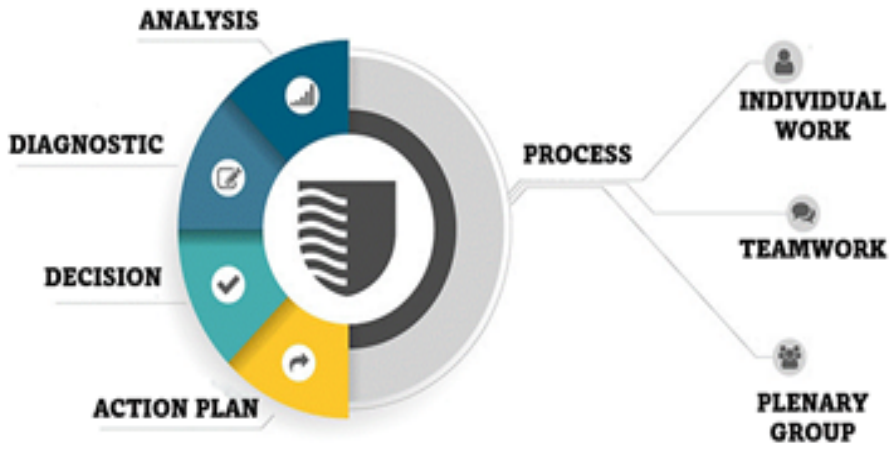
Didactic techniques in the technological institution educational model

a. Problem-based learning (PBL): Adopted by different universities worldwide, such as the University of Maastricht in the Netherlands, it is applied in careers such as medicine, law, economics, administration, and cultural sciences. PBL is oriented to learning where students address real problems in small groups and under the supervision of a tutor. The curriculum comprises thematic modules or blocks. Teams of teachers create these units according to the following, (a) Clarification of the terms and concepts of the problem, (b) Problem definition, (c) Problem analysis (brainstorming), (d) Organization of ideas, (e) Formulation of learning objectives, (f) Obtaining new information and (g) Reporting of results.

Figure 2. *Problem-based learning*

Adopted by different universities worldwide, such as the University of Maastricht in the Netherlands, it is applied in careers such as medicine, law, economics, administration, and cultural sciences. PBL is oriented to learning where students address real problems in small groups and under the supervision of a tutor. The curriculum comprises thematic modules or blocks. Teams of teachers create these units according to the following, (a) Clarification of the terms and concepts of the problem, (b) Problem definition, (c) Problem analysis (brainstorming), (d) Organization of ideas, (e) Formulation of learning objectives, (f) Obtaining new information and (g) Reporting of results.

Figure 3. *Application of the method*



b. Project Oriented Learning (POL): The objective is that the student learns by doing, acquiring an adequate methodology to face the problems that will occur in his future professional practice. It is sought that the student "learns to learn". To apply this strategy, it is necessary that the teacher first defines the type of model to be used: (a) Model 1- The project is carried out independently within each course, (b) Model 2- The project is carried out at the end of the semester, considering content from different courses and (c) Model 3- The project is carried out during the semester, considering content from different courses. The models are applicable in graduate studies, at the end of the degree, and in disciplines where knowledge can be integrated with a specific project.

c. Collaborative Learning (CL): It is not only a technique. It is also a personal way of working and can be used at all levels and subjects. In all situations, people meet in groups with a shared authority based on the consensus built from the cooperation of the group members. An essential aspect is the conformation of small groups within the large group constituted by the students and the teacher. By groups: (a) Plenary Group: Formed by all the students, including the teacher, (b) Base Group: To help each other, duration between 2 or 3 weeks until the end of the course, (c) Formal Group: It can last a part of the class or several sessions to carry out specific projects and (d) Informal Group: The groups are randomly created, no matter the roles, and consume a few minutes of the class.

Components of the educational model

a. Relevant and deep knowledge: The educational model proposes to manage the theories, concepts, and procedures contained in the programs to create a teaching system connected to reality, where students become knowledgeable of history and, thus, understand the world around them to apply this knowledge to problem-solving and constructive activities.

b. Challenge-based learning: It is a pedagogical approach that closely links the student to a problematic situation. The linking occurs with the environment through academic spaces such as Week I and Semester I, spaces oriented to address situations that require defining a challenge and implementing a solution.

c. Flexibility: Flexibility offers the student decision options around their training in which they can personalize and maximize their learning experience. The elements of flexibility are embodied in two parts: flexibility in the curriculum and flexibility in the learning experience

d. Formation of leading professionals with human sensitivity and entrepreneurial spirit: This component offers spaces that guarantee opportunities for voluntary or configured participation in programs that allow learning and development of competencies for life and the professional performance of the participants. As a fundamental pillar of internationalization, this component makes integrating a global, diverse, and multicultural learning community possible.

e. New Teaching and Learning Process: The educational model change is based on teachers' development, oriented toward accompanying students in their learning process and developing competencies through the experience of challenges. The teacher appropriates the role of advisor, evaluator, mentor, challenge designer, and professor. They comply with fundamental characteristics such as staying up-to-date, inspiring, connected, immersed in the teaching processes, innovating, and facile with information technologies. Thus, educators comply with three dimensions of their academic work: teaching, intellectual vitality, and leadership services.

f. Self-directed learning: The preparation of students requires continuous self-regulated education, where they participate in experiences in which they autonomously achieve the construction of their knowledge. That happens when they research, analyze, and contrast information on their own, propose ways of working, and organize the times

to carry them out, and when they present the solutions, they defend their ideas with solid arguments, and communicate them adequately to others for their understanding.

g. Continuous Learning Improvement: The students improve their learning through continuous evaluation, allowing them to reflect on the activities learned and develop the ability of continuous improvement, which is, in short, the ability to learn. Likewise, teachers should obtain continuous information ahead of their learners' evolution.

h. Collaborative Learning: Collaborative learning is an experience of socialization oriented to achieve in the student a solidary way of life where lies the educational essence of developing the mental capacity of the human being. With their peer group and under the direction and support of the teacher, students learn to expound their points of view, accept their differences, help each other, be supportive, work on joint projects, create group rules, and fulfill the commitments collectively adopted.

Enablers of the educational Model

a. Academic Communities: These spaces allow sharing and enhancing the academic experience and successful practices of those who collaborate at this educational level, enriching teaching practice, strengthening intellectual vitality, demonstrating service and leadership skills, etc. The interaction and collaboration in these groups contribute not only to the improvement of academic quality and the generation of educational innovations but also to the construction of the new "Tec21" educational Model and the consolidation of the institution as "one Tec" so that the same institutional spirit and values guide the academic communities.

b. Educational Innovation: Innovation is one of the institutional values of the technological institution under study. It consists of generating and making ideas a reality, breaking paradigms, taking risks, and learning from mistakes. It also implies creating opportunities and acting as generators of change. educational innovation in the "Tec21 Educational Model" is manifested in the formative programs and teaching practices.

c. Educational Spaces: Educational spaces are catalysts for cognitive processes. At the same time, they are the scenario where teachers implement their pedagogical designs. Facilitating interactions among multitasking and highly connected students gives an organic perspective to the teaching-learning process where flexible learning models are fostered with furniture and technology that dynamically propitiate collaborative work and self-management of learning.

d. Linkage with the environment: It is how disciplinary and transversal competencies are expressed and create value in organizations. They are an appropriate vehicle to strengthen skills necessary to incorporate students into working life and(or) business activity.

METHODOLOGY

His research employs a mixed-methods approach to comprehensively investigate student dropout, combining quantitative and qualitative techniques. The study utilizes

a sequential explanatory design, beginning with a quantitative analysis of student dropout data followed by a qualitative exploration of underlying reasons and contributing factors. The System Dynamics methodology is integrated to address the complexity inherent in this phenomenon.

The research population for the quantitative analysis consists of all undergraduate students enrolled at the technological institution under study between 2014 and 2020. The sample for the quantitative phase includes data from students who dropped out of their programs during this period to analyze dropout rate trends. A representative subsample of these students will be selected for in-depth qualitative interviews. The units of analysis for the quantitative phase are individual student records, including demographic data, academic performance, financial aid information, and enrollment history. For the qualitative phase, the units of analysis are individual students who withdrew from the technological institution under study between 2014 and 2020.

Inclusion criteria for the quantitative phase are undergraduate students enrolled in the technological institution under study between 2014 and 2020 with complete student records. For the qualitative phase, the inclusion criteria are students who voluntarily withdrew from their undergraduate programs at the technological institution under study between 2014 and 2020 and are willing to participate in an interview. Exclusion criteria for the quantitative phase are students with incomplete records or those who transferred to another institution without formally withdrawing. Exclusion criteria for the qualitative phase are students who transferred to another institution without formally withdrawing, students who are unwilling or unable to provide informed consent, or those who cannot be reached for an interview.

Data collection techniques include data mining and statistical analysis for the quantitative phase and semi-structured interviews for the qualitative phase. Student records from the technological institution under the study database were extracted and analyzed. The database includes demographic information, academic performance, financial aid data, and enrollment history, with Statistical Analysis made through ANOVA. Data cleaning and validation procedures were implemented to ensure the accuracy and completeness of the quantitative data. Statistical analysis was conducted to assess the reliability of the extracted data.

The quantitative data will be analyzed using descriptive statistics, correlation analysis, and regression analysis to identify factors associated with student dropout. Machine Learning models will be developed to predict student dropouts and identify at-risk students. The interview data will be analyzed using thematic analysis to identify recurring themes, patterns, and insights related to student dropout.

RESULTS

Dropout context in the technological institution under study

The technological institution under study is a university with 25 campuses throughout Mexico. At the level of undergraduate programs, it has 45 careers (disciplinary programs) in its educational offerings (Alvarado Uribe et al., 2022). As of 2022, Tec had 60,169

undergraduate students, of which 54% benefited from some scholarship programs or financial support (Instituto Tecnológico y de Estudios Superiores de Monterrey, 2023).

During 2014-2020, which corresponded to seven cohorts of new admissions, each ending as new students entered the August-December period, the dropout levels behaved as follows (see Table 2) (Instituto Tecnológico y de Estudios Superiores de Monterrey, 2021).

Table 2. *Variation of dropouts at the technological institution at the undergraduate level*

Comparative periods	Dropout (%)
2014-2015	↓ -0,89
2015-2016	↓ -0,38
2016-2017	↓ -0,11
2017-2018	↓ -0,94
2018-2019	↑ +1,96
2019-2020	↓ -1,48

Source: The Authors (2024)

However, this behavior does not correspond to the level of students who enrolled in the programs. The number of students enrolled increased year after year. The absolute number of dropouts as a percentage of the total number of students on Tec campuses can be seen in Table 3 (Instituto Tecnológico y de Estudios Superiores de Monterrey, 2021).

Table 3. *Absolute percentage of dropouts at the technological institution at the undergraduate level. This value is measured over the total number of students enrolled in the Period.*

Cohort	Dropout (%)
2014	10,09
2015	9,20
2016	8,82
2017	8,71
2018	7,77
2019	9,43
2020	7,95

Source: The Authors (2024)

It is important to remember that since 2019, the new Tec21 educational model has been implemented. It changed its form of evaluation from the traditional teaching methodology to a competency-based model, and the teaching times changed from semester courses to 5-week courses.

Exploratory Data Analytics

For 2020 and 2021, the technological institution under study presented 121.584 general dropouts, 63,76% corresponding to dropouts from an undergraduate program (See Table 4).

Table 4. *Filter data*

Characteristic	Description	Values
Global	Global Dropout	121.584
Undergraduate	Level of studies	77.518
Dropout	Students who do not continue Studying at the technological institution	6.813
Tec20	Students who do not continue at the technological institution, Model TEC20	4.685
Tec21	Students who do not continue at the technological institution, Model TEC21	2.128

Source: The Authors (2024)

In the database treated, information is presented on the level of undergraduate education of people who do not continue studying in the technological institution under study. For the 6,813 cases, an invariant and covariate analysis was performed, differentiating the two types of "Educational Models," Tec20 and Tec21. The following 16 variables were evaluated for the presence of seasonality, the correlations, and the function that describe such behavior (Table 5).

Table 5. Variables

N	Variable	Data Type	Description	Values
(1)	Generation	String	A unique indicator that denotes the Generation (class year) to which the student belongs	7
(2)	Gender	String	Student's gender	2
(3)	Age	Integer	Student's age	32
(4)	Father.education. complete	String	Description of the last level of studies completed by the father	9
(5)	Mother.education. complete	String	Description of the last level of studies completed by the mother	9
(6)	Parents.exatec	String	The indicator that denotes if either parent is exatec (a student of the technological institution)	3
(7)	TEC.No.TEC	String	Indicator that denotes if the student comes from a school that belongs to the technological institution	2
(8)	Foreign	String	Indicator to identify if the student is a foreigner (Yes: Foreigner), if the student's birthplace is different from the location of the school campus (Yes: National) or belongs to the same location (Local)	3
(9)	Zone.type	String	Description of the type of zone to which the student's address belongs	4
(10)	First.generation	String	Indicates that the student is the first person in the family to study for a professional career	3
(11)	School	String	Acronyms of the school to which the student's academic program belongs	7
(12)	Program	String	Acronyms of the academic program to which the student belongs	76
(13)	Region	String	Code of the region for the campus where the student is enrolled	5
(14)	Scholarship.perc	Float	Scholarship percentage	26
(15)	Scholarship.type	String	Scholarship type	9
(16)	Socioeconomic.level	String	Socioeconomic level	8

Source: The Authors (2024)

A multivariate test was performed on each variable to determine the homogeneity of the data. For this purpose, contingency tables were made for each of the variables. Tests of uniformity of variance analysis (ANOVA) were performed. The decomposition of the total variation of the data for the OVERALL MEAN allows us to determine if there are significant statistical differences between samples.

For the analysis of variance (ANOVA), data was exported into Microsoft Excel. We first ensured data was grouped according to categorical variables. Then, we used the "Data Analysis" Toolpak, which required enabling the add-in within Excel options. After selecting the ANOVA test, the corresponding data range was selected for each category, setting the significance level at 0.05. Excel then automatically calculated F-statistics, p-values, and degrees of freedom, which were used to determine the statistical significance of differences between groups, to confirm homogeneity.

On the contrary, based on the statistical analysis performed, it can be presumed that the means of the compared groups do not differ significantly concerning the variable being analyzed. To formally test this presumption, two hypotheses were proposed: a null hypothesis (asserting no difference between the means) and an alternative hypothesis (asserting a difference). These hypotheses were evaluated using a pre-determined significance level of 0.05. This significance level indicates that there is a 5% risk of rejecting the null hypothesis when it is true (a Type I error).

Null hypothesis (H0)
H0 = All stockings are equal Alternative hypothesis (Ha)
Ha = At least of the averages differ

The ANOVA test performs an analysis of the variance in which the dispersion of the values of each sample from its meaning is measured. An analysis of the variance of the dispersion of the sample means compared to the overall mean is performed, i.e., it allows the effect of one or more factors on the mean of a variable to be analyzed (Juarez, 2015) (Table 6).

Table 6. *Calculations for analysis of variance*

Sources of variation	Sum of squares	Degrees of freedom	Mean square
Between groups	SCE	gl SCE	SCE/gl SCE
Within groups	SCD	Gl	SCD/gl SCD
SCE	SCT	SCD	Fc
SCT-SCD	$\sum X^2 - \frac{(\sum X)^2}{n}$ All groups	$\sum X^2 - \frac{(\sum X)^2}{n}$ Each group	$\frac{SCE/glSCE}{SCD/glSCD}$ $F_c = F_{a,k-1,N-k}$

Source: The Authors (2024)

To decide that a statistically significant difference exists, the numerical value of F must be greater than or equal to the value of Fc, or it can also indicate that the following decision-making is in place:

if F < Fc, (H0) is approved,
if F > Fc, (H0) is rejected.

To analyze academic desertion data with computer tools, we used the Excel statistical add-in "Data Analysis" in the option "Analysis of variance of one factor", selecting 0,05 as the alpha value for the ANOVA function: this being the minimum acceptable for the social sciences. A total of 169 "One-factor analyses of variance" were constructed for the Tec20 Model and 255 for the Tec21 Model; the latter had two variables included in the run time of that educational model. Figure 4 shows an example of the table obtained in the test for the variable "Generation" (AD19-AD20), classifying the data into "Gender" (Female- Male) for the Tec21 model.

Figure 4. *Example ANOVA - Analysis of variance of a factor*

ANALYSIS OF VARIANCE OF A FACTOR

Origin of variations	*Sum of squares*	*Degrees of freedom*	*Average squares*	*Probability*	*F*	*Fc*
Between groups	7396	1	7396	0,2955551	**1,9702**	**18,51**
Within groups	7508	2	3754			
Total	14904	3				

SUMMARY

Groups	*Account*	*Sum*	*Average*	*Variance*
AD19	2	1150	575	6050
AD20	2	978	489	1458

Source: The Authors (2024). Note: Prepared for EXCEL

An ANOVA was applied to the Model-Tec20 database, resulting in the relationship shown in Figure 5, where A indicates that the Ho is approved and R that the Ho is rejected. No information was available for variables 9 (Zone.type) and 16 (Socioeconomic. level).

Figure 5. *ANOVA applied to the Model-Tec20*

Variables	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
1		A	A	A	A	A	A	A		A	A	A	A	A	A	
2	R		A	A	A	A	A	A		A	A	A	A	A	A	
3	R	R		A	A	R	R	A		A	A	R	R	A	A	
4	R	R	A		A	A	R	A		A	R	R	R	A	A	
5	R	R	A	A		A	R	A		A	R	R	R	A	A	
6	A	R	A	A	A		A	A		A	A	R	R	A	A	
7	R	A	A	A	A	A		A		A	A	R	A	A	A	
8	R	R	A	A	A	A	A			A	R	R	R	A	A	
9																
10	A	R	A	A	A	A	R	A			R	R	R	A	A	
11	R	A	A	A	A	A	R	A		A		A	R	A	A	
12	R	R	A	A	A	R	R	A		A	A		R	A	A	
13	R	R	A	A	A	A	A	A		A	A	R		A	A	
14	R	R	A	A	A	R	R	R		R	R	R	R		A	
15	R	R	A	A	A	R	R	A		A	R	R	R	R		
16																

Source: The Authors (2024)

The data analysis led to the following conclusions for the Tec20 Model:

1. The variable "Generation" does not present significant statistical differences in the means for the tests applied to the 13 variables. • The variable "Gender" does

not present significant differences in the means for the analysis of 12 variables, only for the application of the test with “Generation”, which indicates that there was a significant difference between “Female” and “Male” in each of the generations AD14, AD15, AD16, AD17, and AD18.

2. The variable “Age” presents significant differences in the means for the variables “Generation”, “Gender”, “Parents.exatec”, “TEC.No.TEC”, “Program” and “Region”, “Gender”, “Parents.exatec”, “TEC.No.TEC”, “Program” and “Region”, which indicates that there is no homogeneity for these variables. They are classified by age, being a variable to be studied.

3. The variables “Father.education.complete” and “Mother.education.complete” present similarities in their behavior, which is why they present differences in the means for the factors “Generation”, “Gender”, “TEC.No.TEC”, “School”, “Program” and “Region”. “Gender”, “TEC.No.TEC”, “School”, “Program” and “Region”, for which the means of each group should be observed, and the one(s) with the highest value should be identified. For the variable “Parents.exatec”, there were 3 variables, “Gender”, “Program”, and “Region” with significant differences in the means of each of the internal classifications.

4. Regarding the variable “TEC.No.TEC”, the variables “Generation” and “Program” were identified as having significant differences in the means of their internal classification.

5. The variable “Foreign”, when compared with the 13 variables, shows a significant difference with the means of the variables “Generation”, “Gender”, “School”, “Program” and “Region”, which makes it necessary for future studies to carry out additional tests on these factors. A similar situation is presented for “Zone.type”, in which there is no homogeneity for the variables, “Gender”, “TEC.No.TEC”, “School”, “Program” and “Region”.

6. The variables “Generation”, “TEC.No.TEC”, and “Region” are not homogeneous since they present significant differences in the means when grouped according to the variable “School.

7. For the classification in “Program”, there are five variables where there is no homogeneity “Generation”, “Gender”, “Parents.exatec”, “TEC.No.TEC” and “Region”. These also obtained the same rejection decision in the variable “Scholarship.type” adding “School”, “Program” and “Scholarship.perc”.

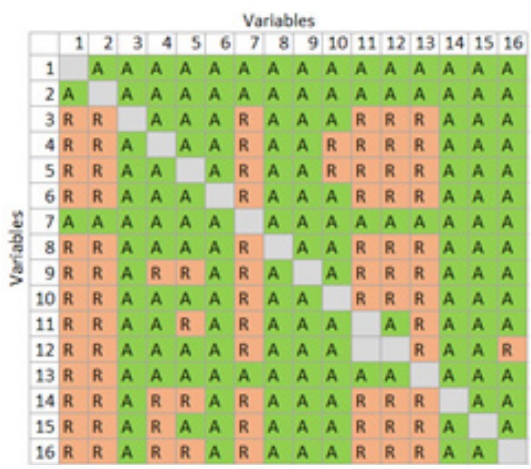
8. The variable “Region” presents homogeneity in most of the variables, except for “Generation”, “Gender”, and “Program”.

9. The variable “Scholarship.perc” approves the null hypothesis for “Age”, “Father. education”, “Mother.education.complete” and “Scholarship.type”.

An ANOVA was applied to the Model-Tec21 database, resulting in the relationship shown in Figure 6, where A indicates that the Ho is approved and R that the Ho is rejected. In statistical terms, ANOVA (Analysis of Variance) is a method used to test significant differences between the means of two or more groups. The null hypothesis (Ho) assumes that there are no significant differences between the group means. Therefore, in Figure 6, ‘A’ signifies that the p-value associated with the ANOVA test was greater

than the chosen significance level (alpha, typically 0.05), leading to the acceptance of the null hypothesis, meaning we did not find statistically significant evidence to reject the assumption that the group means are equal. Conversely, 'R' indicates that the p-value was less than or equal to the significance level, leading to the rejection of the null hypothesis and suggesting that there is a statistically significant difference between at least two of the group means.

Figure 6. *ANOVA applied to the Model-Tec21*



Source: The Authors (2024)

The data analysis led to the following conclusions in the Tec21 Model:

1. The variables "Generation", "Gender" and "TEC.No.TEC" do not present significant differences in the means for the tests applied in the 15 variables defined for each one.
2. The variables "Age", "Parents.exatec", "Foreign" and "First.generation", have the same behavior in terms of non-homogeneity in the means of the variables "Generation", "Age", "TEC.No.TEC", "School", "Program" and "Region".
3. In the case of the variables "Father.education.complete" and "Mother.education.complete", they have the same behavior, whereas the null hypothesis is rejected for the variables "Generation", "Age", "TEC.No.TEC", "First.generation", "School", "Program" and "Region".
4. The variables "Zone.type", "Scholarship.perc" and "Socioeconomic.level" share the behavior of the tests performed on the variables where the acceptance of the null hypothesis is concluded for "Age", "Parents.exatec", "Foreign", "First.generation", "Scholarship.perc", "Scholarship.type", and "Socioeconomic.level".
5. In the case of the variable "School", there are five rejections of the null hypothesis for "Generation", "Gender", "Mother.education.complete", "TEC.No.TEC", and "Region".
6. For the classification in "Program", there are five variables where there is no homogeneity "Generation", "Gender", "TEC.No.TEC", "Region", and "Socioeconomic.level".
7. The variable "Region" shows a significant difference in the means for "Generation" and "Gender".

8. The variable "Scholarship.type", presents rejection in the null hypothesis for "Generation", "Gender", "Father.education.complete", "TEC.No.TEC", "School", "Program" and "Region".

An ANOVA was applied to the Model-Tec21 database, resulting in the relationship shown in Figure 6, where A indicates that the Ho is approved, and R indicates that the Ho is rejected. The classification of each variable in Table 6 is shown in Table 7, allowing us to understand the factor analysis presented in Figures 5 and 6. As a continuation of this analysis, these data from Table 7 facilitate the interpretation of Figure 6, in which «A» means that the p-value associated with the ANOVA test was greater than the chosen significance level (usually 0.05), leading to acceptance of the null hypothesis. However, «R» indicates that the p-value was less than or equal to the significance level, leading to rejection of the null hypothesis, suggesting a statistically significant difference between the means of at least two groups.

Analysis of results - Model Tec-21

To understand social phenomena, one must perform and interpret the inferential statistical analyses and their links with the theory, which is why we decided to apply the "one-factor analysis of variance (ANOVA) to analyze whether more than two groups differed significantly in their means and variances and to identify whether an independent variable influenced the dependent variable. This procedure makes it possible to determine whether, when a particular treatment is applied to a population, it significantly impacts the meaning.

The analysis of the 16 representative variables applied to the educational model of the technological institution under study using ANOVA analysis, where the variables did not show a significant difference in the means between AD19 and AD20, found that the EIC represents 39,75% of the dropouts under the Tec21 Model, which was the one in force at the university. The data indicated the following characteristics of the school´s 846 dropouts: (a) There were 32,51% female and 67,49% male dropouts, (b) 91,84% were between 17 and 19 years of age, (c) About 50% of the dropouts had as a characteristic that their parents "Received undergraduate degree" as their educational level, (d) 83,92% were not "Parents.exatec", which could significantly impact not knowing about the technological institution directly. Also, 61,23% of the dropouts did not come from a high school in the technological institution educational system, (e) 76,8% were from the region of their campus and urban housing, (f) 79,6% were not the first generation, i.e., their parents, siblings, or family already had a university education, (g) 65% did not have any scholarships, and (h) 55,32% did not indicate the socioeconomic level to which they belong, so the absence of this information may generate errors in the type of strategy selected to reduce dropout.

For the case of the EIC school, future work is proposed to implement a causal model, enabling the verification of variable correlations and dependencies. Additionally, expanding and detailing the available variables is suggested, and even a new study using MANOVA could provide valuable insights. This analysis could incorporate a wider range of dependent variables, such as student performance, satisfaction, and

persistence, offering a more nuanced understanding of the factors influencing student success in the EIC school. As well as providing a greater understanding of the data obtained.

DISCUSSION

This research aimed to contribute to the understanding of student dropout in the technological institution under study by examining the Tinto Dropout Model, exploring the impact of Problem-Based Learning (PBL), analyzing dropout trends, and investigating specific causes using data analysis and simulation. The findings offer several points for discussion, grounded in the theoretical framework and contrasted with existing literature.

a. Tinto's Model and its Contemporary Relevance:

The study reaffirmed the enduring relevance of Tinto's model (1975) as a valuable lens through which to examine student departure in higher education. Tinto's emphasis on the interplay between individual characteristics, institutional commitments, and academic/social integration remains a cornerstone of dropout research (Spady, 1970). However, the contemporary context of higher education demands a nuanced understanding of these factors. While Tinto's model provides a robust framework, the specific manifestations of these variables have evolved (Braxton et al., 2004). Our research suggests that while Tinto's model provides valuable insight into developing a strong database where it is possible to integrate information on each student from the four axes defined by Tinto, the results show a significant lack of relevant information within each program. Future research is required, and it would be suitable to perform a deep dive into the database, where each program could extract more information, to strengthen their findings. These factors of desertion will allow for a more detailed study of how to prevent desertion in each of the schools and programs.

The literature emphasizes the importance of considering students' sense of belonging and the role of institutional support systems in fostering retention. Furthermore, the influence of external factors, such as economic pressures and family responsibilities, cannot be overlooked. In the Mexican context, where socioeconomic disparities are pronounced, these external factors may play a particularly significant role.

b. The Tec21 Model and its Impact on Student Retention:

The Tec21 Educational Model at the technological institution, with its focus on student-centered learning and problem-based approaches, represents an innovative attempt to enhance student engagement and promote competence development. However, our findings also highlight the potential challenges associated with adapting to this non-traditional educational environment. Modifying traditional education can be considered a factor for some students as positively challenging and for others as difficult to adapt. The latter is considered one of the reasons for high dropout rates in university education. They need to continuously update, change, and seek strategies to motivate students to develop disciplinary and transversal competencies. Likewise, it is necessary to conduct

statistical treatment and analysis of information regarding students' characteristics and perceptions that facilitate strategy designs for their permanence in education.

Research suggests that PBL can positively impact student motivation and learning outcomes, but its effectiveness depends on factors such as the design of the learning tasks, the quality of instructor facilitation, and students' prior knowledge and skills. Furthermore, the Tec21 model's emphasis on flexibility and autonomy may not be suitable for all students, particularly those who require more structured guidance and support.

c. Trends in Dropout Rates and Disparities among Student Groups:

The analysis of undergraduate dropout rate trends from 2014 to 2020 revealed distinct patterns and disparities among student groups. The statistical treatment showed that characteristics are specific to each of the six schools of the technological institution and, in turn to each program. It will be necessary to characterize each population and define specifically the factors of desertion. This variability highlights the importance of disaggregating data and tailoring interventions to the specific needs of different student populations. These results are in line with previous studies in the region that show high dropout rates on specific programs due to economic situations.

Research suggests that first-generation students, students from low-income backgrounds, and students from underrepresented minority groups are often at higher risk of dropping out. These students may face a range of challenges, including financial constraints, academic preparation gaps, and a lack of social support. Addressing these disparities requires a multi-faceted approach that includes targeted financial aid, academic support services, and culturally responsive interventions.

d. Leveraging Data Analysis and Simulation to Understand Dropout Causes:

This research demonstrates the potential of leveraging data analysis, and simulation to inform strategies to mitigate attrition rates. Data mining techniques can identify at-risk students and predict dropout with reasonable accuracy. System Dynamics modeling allows for the creation of simulations to model different scenarios and their complex behavior for decision-making. These can provide insights into the potential impact of interventions and inform policy decisions.

However, it is important to acknowledge the limitations of relying solely on quantitative data. Qualitative insights, such as those gained through student interviews, are essential for understanding the lived experiences and motivations behind student decisions. Integrating both quantitative and qualitative data provides a richer and more nuanced understanding of student dropout.

e. Critical Reflection and Future Directions:

This research has provided valuable insights into the dynamics of student dropout. The application of a mixed-methods approach, combined with a robust theoretical framework, has yielded a comprehensive understanding of the factors contributing to student attrition. However, the limitations of this study must also be acknowledged. The focus on a single institution limits the generalizability of the findings to other contexts.

Furthermore, the reliance on historical data may not fully capture the evolving nature of student dropout in the context of rapid technological and societal change.

Future research should explore the following avenues:

- **Longitudinal Studies:** Tracking students over time to examine the long-term impact of interventions and policies.
- **Comparative Studies:** Comparing dropout patterns and interventions across different institutions and contexts.
- **Qualitative Research:** Conducting in-depth interviews with students who have successfully persisted in their studies to identify factors that promote resilience and success.
- **Experimental Studies:** Evaluating the effectiveness of specific interventions using randomized controlled trials.

It is crucial for universities and educational systems to continuously update, change, and seek strategies to motivate students to develop disciplines and transversal competencies. Likewise, it is necessary to conduct statistical treatment and analysis of information regarding students' characteristics and perceptions that facilitate strategy designs for their permanence in education.

CONCLUSIONS

This research addressed critical aspects of student dropout in higher education. Several important conclusions can be drawn from the analysis:

Tinto's Model and Problem-Based Learning: The study reaffirmed the relevance of Tinto›s model as a valuable framework for understanding student departure, emphasizing the pivotal role of educational institutions in fostering student retention. While the Tec21 Educational Model at the technological institution aims to foster student competencies through a student-centered, flexible approach, the transition from traditional education can pose challenges for some students. Further research is needed to explore how problem-based learning (PBL) within Tec21 impacts student retention, particularly its influence on the academic and social integration aspects highlighted by Tinto's model.

Trends in Dropout Rates (2014-2020): The analysis of undergraduate dropout rate trends from 2014 to 2020 revealed distinct patterns and disparities among student groups. The statistical treatment showed that characteristics are specific to each of the six schools of the technological institution and, in turn to each program. It will be necessary to characterize each population and define specifically the factors of desertion. Though the Tec21 Educational Model aims to foster student competencies, further information will be required for a wider scope of the findings of each school and program of the institution.

Specific Causes of Dropout: To effectively address dropout, institutions must guarantee a robust database where it is possible to integrate information on each student from the four axes defined by Tinto. The study underscores the importance

of a comprehensive approach to identifying the specific causes of student dropout. This research demonstrates the potential of leveraging data analytics, and simulation to inform strategies to mitigate attrition rates. To minimize dropout, educators and educational stakeholders must collect and analyze ongoing data and combine various strategies to address academic, social, family, institutional, and other identified factors.

Therefore, for future lines of investigation, and with the found results, it will be suitable to analyses the data and results from a closer perspective to each program. In conclusion, combining a strong database system that facilitates obtaining student characteristics will be highly relevant to guarantee the information of the four axes proposed by Tinto. These factors of desertion will allow for a more detailed study of how to prevent desertion in each of the schools and programs.

FINANCING

No external funding has been received for the development of this research.

CONFLICTS OF INTEREST

The authors declare that the submitted study does not represent any conflict of interest between them, with the magazine, the publisher or the financing entities.

ETHICAL CONSIDERATIONS

This study was carried out in abidance of the scientific, technical and administrative standards for health research set out in Resolution No. 8430 of the Ministry of Health and Protection of Colombia (1993).

ACKNOWLEDGEMENTS

The authors would like to acknowledge the Living Lab & Data Hub of the Institute for the Future of Education, Tecnologico de Monterrey, Mexico, for the data published through the Call “Bringing New Solutions to the Challenges of Predicting and Countering Student Dropout in Higher Education” used in the production of this work.

AUTHOR’S CONTRIBUTION

Mónica-Lizeth Sánchez-Arévalo: Conceptualization, methodology, data curation, formal analysis, research, visualization and writing.

Roberto Ferro-Escobar: Conceptualization, research and visualization.

Luisa-Fernanda Chaparro-Sierra: Research and validation.

REFERENCES

Alvarado Uribe, J., Mejía Almada, P., Masetto Herrera, A., Molontay R, H., Hegde, V., Montemayor Gallegos, J,. Ceballos, H. (2022). Student Dataset from Tecnologico de Monterrey in Mexico to Predict Dropout in Higher Education. https://www.mdpi.com/2306-5729/7/9/119#cite

Corengia, A., Pita, M., Mesurado, & Centeno, A. (2013). Predicting academic performance and attrition in undergraduate students. *Liberabit*, 101-112. http://www.scielo.org.pe/scielo.php?pid=S1729-48272013000100010&script=sci_abstract&tlng=en

Donoso Díaz, S., & Schiefelbein, E. (2007). Analysis of explicative models of student's retention at the university: A vision from the social inequality. *Estudios Pedagógicos*, 7-27. https://www.researchgate.net/publication/279910482_Analysis_of_explicative_models_of_student's_retention_at_the_university_A_vision_from_the_social_inequality

Flores López, W., Gutiérrez Restrepo, E., León Corredor, O., Sarraipa, J., Pantoja Lima, C., Merino, C.,. Boticario, J. (2016). Centros de Apoyo y Desarrollo Educativo Profesional para la observación y disminución de la deserción universitaria. *Ciencia e Interculturalidad*, 48-62. https://doi.org/10.5377/rci.v18i1.3049

Instituto Nacional de Estadística y Geografía (INEGI). (2020). Encuesta para la Medición del Impacto COVID-19 en la Educación (ECOVID-ED) . https://www.inegi.org.mx/investigacion/ecovided/2020/

Instituto Nacional de Estadística y Geografía (INEGI). (2022). Tasa de abandono escolar por entidad federativa según nivel educativo, ciclos escolares seleccionados de 2000/2001 a 2020/2021. https://www.inegi.org.mx/app/tabulados/interactivos/?pxq=9171df60-8e9e-4417-932e-9b80593216ee

Instituto Tecnológico y de Estudios Superiores de Monterrey. (2021). Los campus del Tec de Monterrey y otros datos interesantes. México. https://conecta.tec.mx/es/noticias/nacional/institucion/campus-tec-de-monterrey

Instituto Tecnológico y de Estudios Superiores de Monterrey. (2023). Datos y Cifras. México: Tecnológico de Monterrey. https://tec.mx/es/datos-y-cifras

Juarez Lugo, C. (2015). Análisis de varianza de un factor ANOVA I. Centro Universitario UAEM Ecatepec. http://ri.uaemex.mx/bitstream/handle/20.500.11799/33860/secme-16960.pdf?sequence=1

Narváez, J. L., y Gómez Bustamante, E. (2023). Factores asociados con la deserción escolar en una institución educativa pública en Cartagena (Colombia). *Panorama Económico, 31*(3), 267–290. https://doi.org/10.32997/pe-2023-470

Organization for Economic Co-operation and Development-OECD. (2017). Nota Pais: Panorama de la Educación 2017. https://www.oecd.org/education/skills-beyond-school/EAG2017CN-Mexico-Spanish.pdf

Ruiz Ramírez, R., García Cué, J., & Pérez Olvera, M. (2014). Causas y consecuencias de la deserción escolar en el bachillerato: Caso Universidad Autónoma de Sinaloa. Ra Ximhai. https://www.redalyc.org/articulo.oa?id=46132134004

Saldaña Villa, M., & Barriga, O. (2010). Adaptación del modelo de deserción universitaria de Tinto a la Universidad Católica de la Santísima Concepción,Chile. *Revista de Ciencias Sociales*, 616 - 628. http://ve.scielo.org/scielo.php?script=sci_arttext&pid=S1315-95182010000400005&nrm=iso

Sánchez, M. L. (2015). Modelo representativo de deserción estudiantil voluntaria en carreras de pregrado de la Facultad de Ingeniería de la Universidad Distrital Francisco José de Caldas-Bajo la dirección de PhD. Roberto Ferro Escobar. Universidad Distrital Francisco José de Caldas. http://repository.udistrital.edu.co/bitstream/11349/2939/1/SanchezArevalo MonicaLizeth2015.pdf

Schneider, B. S., & Preckel, F. (2017). Variables associated with achievement in higher education: A systematic review of meta-analyses. *Educational Psychology Review, 29*(4), 757-786.

SEP. (2024). Análisis de matrícula, abandono escolar, eficiencia terminal y cobertura en Educación Básica de 2019 a 2023. Secretaría de Educación Pública. https://educacionbasica.sep.gob.mx/analisis-de-matricula-abandono-escolar-eficiencia-terminal-y-cobertura-en-educacion-basica-de-2019-a-2023/

Tecnológico de Monterrey. (2018). Modelo educativo Tec21. https://tec.mx/es/modelo-tec

Tecnológico de Monterrey. (2019). Reinventamos la educación con el Modelo educativo Tec21:Insignia del Plan 2020. https://tec.mx/sites/default/files/inline-files/Informe_Anual_Tec_2019.pdf

Tinto, V. (1975). Dropout from higher education: A theoretical synthesis of recent research. *Review of Educational Research, 45*(1), 89-125.

UNAM. (2021). Aumento de abandono escolar y trabajo infantil, consecuencia del coronavirus. https://ciencia.unam.mx/leer/1120/aumento-de-abandono-escuela-y-trabajo-infantil-consecuencia-del-coronavirus

UNESCO. (1996). Policy Paper for Change and Development in Higher Education by UNESCO. *Higher Education*, 529-533. http://www.jstor.org/stable/3447824

UNESCO. (2016). Data availability for the calculation of SDG 4-Education 2030 indicators. Analysis covering Latin America and the Caribbean. Quebec: tech. rep. https://unesdoc.unesco.org/ark:/48223/pf0000246263

UNESCO. (2022). Reimagining Our Futures Together: A New Social Contract for Education. UNESCO.

UNICEF. (s.f.). Asistencia a la Escuela. https://www.unicef.org/mexico/asistencia-la-escuela

Vargas Tamayo, L. F., Contreras Bravo, L. E., & Tristancho Ortiz, J. A. (2014). Desarrollo de herramientas computacionales para la prevención de bajo rendimiento

académico y deserción en la educación superior. *Revista Q*. http://hdl.handle.net/20.500.11912/6822

World Bank. (2024). Education Overview: Development news, research, data. https://www.worldbank.org/en/topic/education/overview

Monica-Lizeth Sánchez-Arévalo: Master's degree in industrial engineering from Universidad Distrital Francisco José de Caldas - Colombia. Industrial Engineer from Universidad Distrital Francisco José de Caldas - Colombia. Professor Research at Institución Universitaria Politécnica Grancolombiano (Bogotá, Colombia). Her main research topics include university dropout and social and business issues through mathematical models. ORCID: https://orcid.org/0000-0001-8871-1912

Roberto Ferro-Escobar: Ph.D. in Informatics Engineering from Universidad Pontificia de Salamanca- España. Master's degree intel informatics from Universidad Distrital Francisco José de Caldas - Colombia. Electronic Engineer from Universidad Distrital Francisco José de Caldas - Colombia. Her main research topics include Smart cities. Professor at Universidad Distrital Francisco José de Caldas (Bogotá, Colombia). ORCID: https://orcid.org/0000-0002-8978-538X

Luisa- Fernanda Chaparro- Sierra: Ph.D. in High Energy Physics from Universidad de Los Andes. Postdoctoral research in the Mathematics Department at Universidad Nacional de Colombia in Data Science applications to quantify the crime in Bogotá (Colombia). Research Professor at Tecnológico de Monterrey, Campus Monterrey (Monterrey, México). Her main research topics include computational modeling and phenomenology analysis of High Energy Physics and its applications. ORCID: https://orcid.org/0000-0002-6570-9075

ANNEX A

Table 7. *Calculations for analysis of variance*

N	Variable	Number of unique values	Values
(1)	Generation	7	1) AD14 (Tec20) 2) AD15 (Tec20) 3) AD16 (Tec20) 4) AD17 (Tec20) 5) AD18 (Tec20) 6) AD19 (Tec21) 7) AD20 (Tec21)
(2)	Gender	2	1) Female 2) Male
(3)	Age	32 (4 ranges)	1) 16-17 2) 17-21 3) 21-26 4) 26+
(4)	Father.education.complete	9	1) Attended university but did not graduate 2) Graduated from elementary or middle school 3) Graduated from high school 4) No information 5) No educational degree 6) Received master's degree 7) Received Ph.D. 8) Received technical or commercial degree 9) Received undergraduate degree
(5)	Mother.education.complete	9	1) Attended university but did not graduate 2) Graduated from elementary or middle school 3) Graduated from high school 4) No information 5) No educational degree 6) Received master's degree 7) Received Ph.D. 8) Received technical or commercial degree 9) Received undergraduate degree
(6)	Parents.exatec	3	1) Yes 2) No 3) No information
(7)	TEC.No.TEC	2	1) Tec 2) No Tec
(8)	Foreign	3	1) Local 2) Yes: Foreigner 3) Yes: National
(9)	Zone.type	4	1) No information 2) Rural 3) Semiurban 4) Urban

N	Variable	Number of unique values	Values
(10)	First.generation	3	1) Yes 2) No 3) No information
(11)	School	7	1) EHE: Escuela de Humanidades y Educación 2) EAAD: Escuela de Arquitectura, arte y Diseño 3) ECSG: Escuela de Ciencias Sociales y Gobierno 4) EIC: Escuela de Ingenieria y Ciencias 5) EMCS: Escuela de Medicina y Ciencias de la Salud 6) EN: Escuela de Negocios 7) No information
(13)	Region	5	1) DR: Región Desarrollo Regional 2) RCM: Región Ciudad de México 3) RCS: Región Centro-Sur 4) RM: Región Monterrey 5) RO: Región Occidente
(14)	Scholarship.perc	26 (6 ranges)	1) 0% 2) 5%-30% 3) 30%-40% 4) 40%-50% 5) 50%-60% 6) 60%-90%
(15)	Scholarship.type	9	1) No scholarship 2) Academic Talent 3) Army/Navy scholarship 4) Child of Professor/Employee/Director 5) Cultural Talent 6) Leadership Talent 7) Leaders of Tomorrow Scholarship 8) Sport Talent 9) Traditional
(16)	Socioeconomic.level	8	1) Level 1 2) Level 2 3) Level 3 4) Level 4 5) Level 5 6) Level 6 7) Level 7 8) No information

End Table