Impact of Demographic and Co-Curricular Factors on the Academic Success of Students from Low-Income Families: A Scholarship Program Study

Andres Gonzalez-Nucamendi
School of Engineering and Science
Tecnologico de Monterrey
Mexico City, Mexico
anucamen@tec.mx

Luis Neri
School of Engineering and Science
Tecnologico de Monterrey
Mexico City, Mexico
neri@tec.mx

Rosa M.G. García-Castelán
School of Engineering and Science
Tecnologico de Monterrey
Mexico City, Mexico
rmggarci@tec.mx

Víctor Robledo-Rella
School of Engineering and Science
Tecnologico de Monterrey
Mexico City, Mexico
vrobledo@tec.mx

Jorge Valverde-Rebaza
School of Engineering and Science
Tecnologico de Monterrey
Mexico City, Mexico
jvalverr@tec.mx

Julieta Noguez
School of Engineering and Science
Tecnologico de Monterrey
Mexico City, Mexico
jnoguez@tec.mx

***Abstract*—This research paper describes the factors that contribute to the academic success of university students who receive scholarships due to their outstanding performance in high school, along with a track record of social impact and leadership activities, despite their disadvantaged economic status. The data analysis encompasses information from 2014 to 2023 for a sample of 1,796 students at Tecnologico de Monterrey, with over 60% enrolled in engineering programs. This study examines the factors that positively impact the academic performance of students enrolled in the "Leaders of Tomorrow" scholarship program. The most relevant factors are participation in student group activities, international experiences, community activities, and projects related to technological development or research. Additionally, certain demographic factors, such as not being the first family member to attend university, younger age, coming from urban areas, proximity to the university campus, and female gender, also had a positive impact on academic results. It was found that previous work experiences, sports or cultural activities improve academic performance when properly balanced with the time dedicated to the academic study. Furthermore, the scope of student projects in high school, participation in leadership activities or in academic contests were not relevant to the university academic performance. Understanding the diversity of factors that contribute to academic success will enable more effective support, and the implementation of timely interventions aimed at maximizing the educational opportunities of these students.**

Keywords—academic success factors, co-curricular activities, high-performance students, economically disadvantaged students, scholarships, demographic factors, higher education, educational innovation

I. INTRODUCTION

The 17 Sustainable Development Goals (SDGs) represent humanity's most visionary blueprint for crafting a more equitable and feasible world for present and future generations, with the ambition of achieving tangible progress by the year 2030 [1]. The fourth sustainable development goal is to increase access to quality education. Along this line, Tecnologico de Monterrey has strengthened its distinguished scholarship initiative known as the "*Leaders of Tomorrow*" program (hereinafter referred to as LTP) [2], designed for talented youth in need of full financial support to embark on their undergraduate academic journey at this institution. Cultivating visionary leaders for a societal transformation is the main goal of this endeavor. Applicants are required to actively engage in initiatives aimed at promoting community well-being. They must demonstrate leadership attributes and social empathy. A minimum average grade of 90/100 in high school is required o be eligible to apply for this type of scholarship. Moreover, substantiating financial need is imperative. The scholarship grants full tuition, comprehensive Health Care Insurance and support for the acquisition of books during the whole 4-year program. Formative guidance, mentorship networks, and advisory support for project development are integral parts of the program. Leadership advancement and sound psychological capital will advance over time. The bachelor's degree in computer engineering, followed by biotechnology engineering and mechatronics, are the most enrolled careers. Upon receiving the scholarship, it is required to maintain an accumulated grade average of 85/100 throughout all the semesters. Furthermore, it is mandatory to pass all courses and fulfill the assigned scholarship service duties. The student record, which contains information on academic achievements, extracurricular activities, and mentoring assignments, must be kept updated [2].

To ensure the long-term sustainability of the LTP, it is essential to identify and assess the factors that most significantly impact the academic performance of participating students. Hence, the central research question of this study is:

What factors determine a good performance in low-income and high-performance students enrolled in the *Leaders of Tomorrow Scholarship Program*?

The objective that emerges from the research question is:

To determine the factors that determine academic excellence among low-income students with high academic performance enrolled in the Leaders of Tomorrow Scholarship Program.

The paper's organization is as follows. In Section II the Related Work is outlined. The Theoretical Framework is included in Section III. Section IV discusses the Methodology followed in this work. The Results and Analysis of this research are presented in Section V. Finally, the main conclusions and future work that emerge from the study are highlighted in Section VI.

II. Related Work

In Mexico, several institutions provide scholarships to assist low-income students. In addition to Tecnologico de Monterrey, entities such as banks [3-5], Telmex (a huge Mexican telephone company; [6]), Carlos Slim Foundation [7], IMJUVE (Mexican Institute for Youth; [8]), the Organization of American States (OAS; [9]), the National Institute for Indigenous Groups (INPI; [10]), among others, allocate funds for scholarships to support less-privileged sectors. Carlos Slim Foundation has gone further, creating the *APRENDE* platform, which provides free online education through open-source courses when the Telmex internet connection is used. *APRENDE* has benefited more than 4 million users, mainly in the areas of mathematics, engineering, chemistry, biology, health, humanities, and business [7]. The OEA-CONAHCYT-AMEXCID scholarship consortium has also focused on providing scholarships for science and engineering students who want to pursue graduate studies [9]. As of 2023, the Mexican Government had allocated 142 billion pesos (approximately 8.4 billion dollars) to several educational initiatives aimed at reducing educational disparities [11]. However, no data analysis studies have been detected to identify success factors and to follow up on the trajectory of students in these scholarship programs from all the previously mentioned institutions.

Several studies address the challenges and solutions related to academic success among disadvantaged students. Tomliston & Javis [12] explored how teachers and schools support high-potential minority students from low-income backgrounds. Rivera's 2021 dissertation [13], highlighted the difficulties principals face in under-resourced schools, emphasizing the need for community engagement to bridge the achievement gap caused by economic disparity. Ribar & Rubenstein [14] examined scholarship dynamics in Georgia, finding that students of color, economically disadvantaged students, and male students, face higher risks of losing scholarships, thereby exacerbating existing inequities. Despite these insights, these studies do not account for students' individual characteristics in their academic success.

III. Theoretical Framework

This section presents definitions and key concepts to facilitate the understanding of the logic behind this work and its results. It begins with contextualizing the concept of dropout-prone or vulnerable students. Then, the statistical tools and techniques used in this research are summarized.

A. Vulnerable Students

To define vulnerable students, two approaches are considered: the first is based on a minimum average requirement of 85/100, which provides an objective placement of the student irrespective of context. The second approach pertains to low outliers, specifically students who fall significantly below the range of the majority group, with their positioning being contextual and contingent upon the distribution of the group.

B. Visualization Tools and Statistical Tests for Mean Comparison

The basic visualization tools are histograms and boxplots. Histograms display the frequency distribution, facilitating the understanding of the shape and trends of the data. A Kernel Density Estimation (KDE) curve can be added to the histogram to provide a more precise visualization of the distribution of probabilities [15]. Boxplots represent the dispersion and median of the data, efficiently identifying outliers. Additionally, the Welch's *t*-test is employed, as it is particularly useful for comparing the means of independent groups with different variances in scientific research [16].

IV. Methodology

This section delineates the methodology employed for analyzing the success factors among high-achieving students from economically disadvantaged backgrounds. It comprises three pivotal stages: (a) Data definition, (b) Variable selection and data preprocessing, and (c) Selection of statistical techniques and acquisition of results.

A. Data definition

Tecnologico de Monterrey has provided the authors with access to a database (DB) containing the academic trajectories of $N = 1796$ students from 2014 to 2023. This comprehensive dataset comprises 22,718 meticulously organized entries, delineating the developmental journey of each student throughout their educational tenure. The database, originally comprising 45 distinct columns, delineates 16 numerical variables encapsulating facets of academic performance prior to and during enrollment, and granular insights into the students' personalities. The remaining 29 columns are devoted to demographic data, including gender, age, nationality, and scholar antecedents, such as the prior academic institution and detailed accounts of admission examinations and evaluations. Moreover, it also catalogs the students' engagement in social initiatives, detailing levels of involvement in athletic, cultural,

and leadership pursuits, as well as ongoing academic programs and extracurricular engagements. In sum, the database provides academic and demographic information about LTP students. Given that most students in the sample were enrolled in computing or engineering careers, (about 63%), the results of this study mainly reflect the characteristics and profiles of these students.

B. *Variable Selection and Data Preprocessing*

1) Target Variable

Upon data analysis, the variable "*term_gpa_program*" was chosen as the indicator of student academic performance. This variable represents the cumulative overall average grade of students up to a specific period and provides a more comprehensive and stable representation of academic performance compared to solely utilizing the average of a specific period.

2) Selection between Traditional Learning model and Competency-based model

The database contains information from two different educational models at our institution: the past "Traditional Learning Model" (TLM hereafter), which was used until Spring 2019, and the current "Competency-Based Model" (CBM hereafter). The sample size, number of records in the DB and periods for each model are given in Table I. Although a comparison of some results for these models is outlined below, the analysis of the present work is mainly focused on the CBM.

TABLE I. SAMPLE SIZES AND PERIODS FOR THE TWO MODELS

	Number of records	Number of students	Period
Traditional Learning Model (TLM)	13,444	984	Fall 2014 to Spring 2019
Competence-based model (CBM)	9274	826	Fall 2019 to Fall 2023

3) Selection of predictor variables

A careful inspection of the DB revealed 15 categorical variables worthy of detailed analysis. To examine the academic performance of students based on the alternatives presented in each of the categorical variables, the average of "*term_gpa_program*" was calculated for several subgroups within every categorical variable. The selected variables from the DB used in the following analysis and their explanation are given in Table II below. A further study of the remaining variables in the DB will be presented in a forthcoming paper.

Additionally, the proportions of students considered "at risk," defined as those with an average grade below 85/100 (the threshold required to retain their scholarships), were scrutinized. Boxplots were employed to identify outliers, defined as data points falling below the lower quartile minus 1.5 times the interquartile range or above the upper quartile plus 1.5 times the interquartile range. Furthermore, their occurrence percentages in both models were compared, and the interquartile ranges in each scenario were assessed.

Regarding the CBM alone, the categorical variables with the greatest impact on academic performance were identified to calculate and compare the *term_gpa_program* averages within each variable. Moreover, hypothesis tests were conducted to compare two means using the *t*-Welch statistic, carefully selecting two subgroups of interest and considering the possibility of unequal variances. These tests were performed using Python libraries, such as *pandas*, *numpy*, *matplotlib*, *seaborn*, and *scipy.stats*.

V. RESULTS AND ANALYSIS

A. *Comparison of Educational Models*

The selected dependent variable was *term_gpa_program* grade, which varies from 0 – 100. The range 70 – 100 was selected to improve data visualization, since there were very few students with *term_gpa_program* below 70. Only 12 students from the TLM were removed from this study (0.09%), and 19 students from the CBM (0.20%).

In Figure 1, shows the histogram of the number of students and the corresponding KDE curves as function of *term_gpa_program* for TLM and CBM, for *term_gpa_program* > 70. As it is evident, the highest TLM frequency box centers around *term_gpa_program* = 92.5, whereas for CBM, it is centered around 95. Clearly, the TLM distribution surpasses the CBM distribution for values between 80 and 94, while the opposite is observed for values above 94. Therefore, based on the *term_gpa_program* variable, CBM students perform better than the TLM students.

B. *Selection of Statistical Techniques*

To assess the difference in academic performance between LPT students under the TLM and CBM educational models, a variety of analytical methods were implemented. Histograms were employed to observe the frequency distributions of the *term_gpa_program* variable for students within both TLM and CBM models.

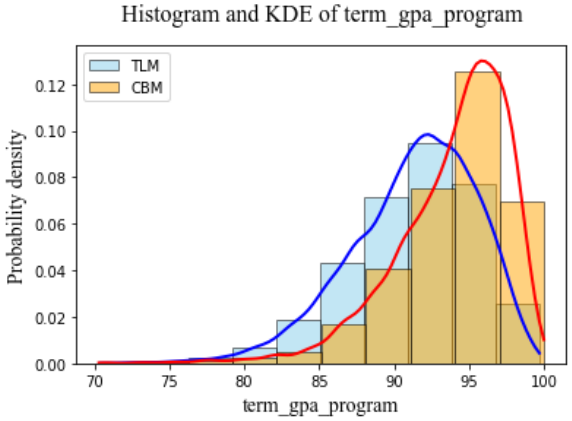

Fig. 1. Histograms and corresponding KDE curves for TLM and CBM.

TABLE II. VARIABLES NAMES AND SUBGROUPS DESCRIPTION

Variable name	Description of Subgroups
Project Scope	**Student completed project scope.** Carried out 1 = with close friends or family 2 = with neighbors or schoolmates 3 = at state level 4 = at national level Does not apply = Assessment was not considered
Project type	**Project classification into one of three subgroups.** Technology development or research Leadership Social impact project Does not apply = The assessment was not considered
Sports	**Previous level of participation in a sporting activity.** The student 1 = has not participated in sports activities 2 = has participated in any sporting activities 3 = has been part of a representative team, participated in sports competitions individually, or selected at the municipal, state or regional level 4 = has been selected at the national or international level, or has obtained recognition as a high-performance athlete in a team or individually Does not apply = The assessment was not considered
Cultural	**Previous level of participation in a cultural activity.** The student 1 = has not participated in cultural activities 2 = has participated in any cultural activities 3 = has participated in any cultural activities, occupying leading positions, or the student has received recognition at the municipal, state, or regional level 4 = The student has participated in any cultural activities and has received national or international recognition Does not apply = The assessment was not considered
Leadership	**Previous level of participation in a leadership activity.** The student 1 = has not participated in leadership activities 2 = has participated in a leadership or entrepreneurship project inside or outside of their school 3 = has coordinated a leadership or entrepreneurship project inside or outside of their school 4 = has developed and led a project inside or outside of their school Does not apply = The assessment was not considered
Student activities	**Previous level of participation in student activities.** The student 1 = has not participated in student activities 2 = has collaborated in any student group activities, without being a member of the board of directors, or the student has participated in some school activities 3 = has been a member of the board of directors of a student group or the students has represented their school in some school activities 4 = has created a student group at their school or the student has won national or international competitions in some school activities Does not apply = The assessment was not considered
Academic achievement	**Previous level of participation in academic achievements.** The student 1 = has no academic achievements 2 = has participated in any interscholastic, local, regional, or state competitions, obtaining recognition 3 = has participated in any national or international competitions, obtaining recognition 4 = has been among the top five places of their generation or the student has participated in a higher-level Advanced Placement or International Baccalaureate course Does not apply = The assessment was not considered
Work experience	**Previous level of work experience.** The student 1 = has not had work experience 2 = has participated in a work project without receiving remuneration or recognition for it 3 = has participated in a work project and has received remuneration or recognition for it 4 = has led a work team in a work project with or without remuneration and recognition for it Does not apply = The assessment was not considered
International	**Previous level of participation in international experience.** The student 1 = has not had international experience 2 = went on a study trip abroad with her school or attended a competition abroad (academic, sports, or cultural) 3 = carried out studies abroad during up to three months 4 = carried out studies abroad for at least one semester Does not apply = The assessment was not considered
Community	**Previous level of participation in a community activity.** The student 1 = has not participated in community support activities 2 = has collaborated in a group or institutions, such as churches, schools, clubs, or on their own 3 = has been a board member of a group or community that carries out philanthropy or community service activities 4 = has created a group or community that carries out philanthropy or community service activities Does not apply = The assessment was not considered
First generation	**Whether students are the first generation in their family to study a professional career.** Yes = First generation No = Not first generation No information = Missing information
Foreign	**Whether the student is local or foreigner.** 0 = Local (student residence address is in the same city as the Campus where they are enrolled,) 1 = Foreigner (otherwise) No information
Address zone	**Type of zone classification of the student's address.** Rural = Population < 2,500 Semiurban = Population between 2,500 and 15,000 Urban = Population > 15,000 No information
Age	**Student's age.** ≤ 18; 19 to 21; ≥ 22
Gender	**Student's gender.** Female; Male

In Figure 2, the average accumulated grade (*term_gpa_program*) boxplots for students in the TLM and the CBM are presented. The quartiles and students falling outside them ("outliers") are clearly shown. In alignment with the findings from Figure 1, the data presented in Figure 2 corroborate that, on average, students enrolled in CBM demonstrated a better academic performance than those in TLM.

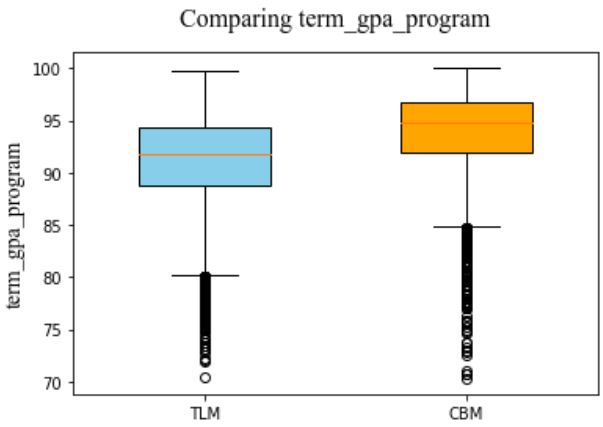

Fig. 2. Boxplots for TLM and CBM average accumulated grades.

In a similar way, in Table III the percentages of students at risk of dropping out the LTP are shown, for both TLM and CBM models, based on their academic performance (defined as *term_gpa_program* < 85). The data clearly indicate that TLM students exhibit a higher probability of dropping out of the LTP.

TABLE III. STUDENTS AT RISK IN THE TWO EDUCATIONAL MODELS BASED ON ACADEMIC PERFORMANCE

Model Number of students	Percentage of Outliers (Median ± 1.5 IQR)	Percentage of Students Below 85
TLM $N = 984$	1.26 %	8.09 %
CBM $N = 826$	2.43 %	2.57 %

Despite the overall high standards of academic performance exhibited by all LTP students, it remains crucial to identify the segment of academically vulnerable students.

C. *Analysis of categorical variables in academic performance for CBM students*

This section studies the relation between *term_gpa_program* and the categorical variables defined in Table II, to discriminate between students having (or not) a solid academic performance. It is important to note that this section corresponds exclusively to the CBM data (N = 826). The objective is to map out the primary determinants of these strengths or vulnerabilities. In this task, the search was restricted only to those variables with five or fewer subgroups, as statistical or machine learning techniques for data dimensionality reduction are not being considered (refer to Table II).

Table IV presents the program accumulated average grades (*term_gpa_program*) for the various subgroups within each categorical variable, as described in Table II for CBM students. The sample size for each subgroup ordered by decreasing percentage is also included.

The average and sample size percentages for students with missing information in the variables are also provided in the table (No information). However, it is important to note that the results for these students without information are not considered in the subsequent analyses.

Hypothesis tests were conducted to compare the means in academic performance (*term_gpa_program*) among the various subgroups presented in the categorical variables. Given the focus of this study on describing the scholarship program and its outcomes we opted to use Welch's *t*-test to contrast the maximum and minimum means, assuming unequal variances. While ANOVA or Factorial Analysis could provide a more rigorous statistical examination, they were deemed beyond the scope of this paper, which prioritized a simple statistical approach. This approach was chosen to maintain the paper's emphasis on context description rather than shifting towards an exhaustive statistical analysis, which is planned for a forthcoming paper using the same dataset. All test results yielded significant findings, with *p*-values less than 0.001 in all evaluated variables. Consequently, the null hypothesis was rejected in favor of the alternative hypothesis in all cases, indicating significant differences in academic performance among the different alternatives in the analyzed categorical variables.

From Table IV it can be seen that:

(i) The level of students' project scope (locally, within a community, statewide, or nationally) has not an important impact on students' grades; the difference among the different subgroups is only 0.5 points out of 100. The highest percentage population (33.89 %) corresponds to subgroup 3 (state level).

(ii) Students who pursued projects related to "Technology development or research" received slightly higher grades (94.4) compared to those focusing on "Social impact" (93.8) or "Leadership" (92.8). However, most student projects were centered around "social impact" (87.1%), with the other two subgroups comprising approximately 6 to 7% each.

(iii) Participating in sports prior to starting undergrad studies has a modestly positive influence on student' academic performance throughout their program. Subgroups 2 (engaging in any sports) and 3 (involvement in representative teams or individual competitions at municipal, state, or regional levels) exhibit only slight differences of 0.1 and 0.9 points, respectively, compared to subgroup 1 (no sports). Interestingly, students who compete at national or international levels or achieve recognition as high-performance athletes (subgroup 4) not only attain lower average grades than those engaged locally or individually (subgroups 2 and 3), but also when compared to

students in subgroup 1. This suggests that the commitment and time invested in national or international sports may somewhat divert students' attention from their academic pursuits. The largest segment of the student population falls into subgroup 2 (23.6%), participating in sports without involvement in competitions or representative teams. It is important to mention that although these findings were based on students' high school sports activities, it is probable that students maintain similar levels of sports involvement during their university studies.

TABLE IV. ACADEMIC PERFORMANCE BY SUBGROUPS AND SUBGROUPS RELATIVE SIZE FOR CBM STUDENTS

Variable	Subgroup	Means of academic performance by subgroup (scale: 0 to 100)	Size of each subgroup (%)
proj_scope	1	93.9	17.39
	2	94.1	19.05
	3	93.7	33.89
	4	93.6	29.43
	No information	83.6	0.24
proj_type	Leadership	92.8	6.26
	Social impact project	93.8	87.11
	Technology development or research	94.4	6.62
sport	1	93.9	21.89
	2	94	23.59
	3	94.8	13.85
	4	91.4	1.43
	No information	93.3	39.24
cultural	1	94.3	14.43
	2	93.9	28.33
	3	94.6	15.65
	4	92.1	2.36
	No information	93.3	39.24
leadership	1	94.6	10.9
	2	93.3	8.58
	3	94.7	8.82
	4	93.9	32.46
	No information	93.3	39.24
student_activt	1	93.7	18.29
	2	94.1	16.88
	3	94.3	20.38
	4	94.5	5.22
	No information	93.3	39.24
academic_achievements	1	94.2	11.3
	2	94.1	21.06
	3	93.9	15.41
	4	94.2	12.99
	No information	93.3	93.3
work_experience	1	94	23.03
	2	94.5	11.24
	3	94.1	20.13
	4	93.8	6.16
	No information	93.3	39.44
international	1	94	53.94
	2	94.2	5.49
	3	94.5	0.61
	4	94.3	0.72
	No information	93.3	39.24
community	1	93.5	4.69
	2	94	26.41
	3	94.1	13.58
	4	94.2	16.09
	No information	93.3	39.24
isFirstGeneration	No	94	49.58
	Yes	93.5	49.19
	No information	91.2	1.23
isForeign	0	94.1	49.45
	1	93.4	46.72
	No information	93.7	3.83
address_zone	Rural	92.9	7.44
	Semiurban	93.3	15.13
	Urban	94	70.84
	No information	93	6.59
age	18 and below	93.8	30.49
	19 to 21	93.9	63.61
	22 and above	91.8	5.9
gender	Female	94	54.76
	Male	93.5	45.24

(iv) Similarly to participation in sports activities, involvement in cultural activities prior to university enrollment appears to have minimal impact on students' average grades. Additionally, students classified in the highest level of involvement (subgroup 4; those receiving recognition at national or international levels) achieved the lowest average grades among the student sample. This also suggest that the dedication and time allocated to international cultural activities might distract students from their academic endeavors. The largest proportion of the student population (28.3%) comprises individuals engaged in cultural activities locally or regionally (subgroups 2 and 3), without national or international involvement.

(v) The extent of involvement in leadership activities does not significantly affect student academic performance. Students who did not participate in these activities (subgroup 1) achieved higher average grades (94.6) compared to students in subgroups 2 and 4 (93.3 and 93.9, respectively), and only slightly lower than students in subgroup 3 (94.7). Notably, 90.0% of the student sample participated to some extent in leadership activities, with the largest percentage belonging to subgroup 4

(32.5%), where students initiated and managed projects either within or outside their high schools.

(vi) The level of participation in student group activities before starting university positively correlates with the average grades of LTP students. Students who did not participate in such activities (subgroup 1), attained lower average grades (93.7) compared to those who did engage in such activities. Furthermore, grades increased with greater involvement or breadth of activities, ranging from 94.1 for subgroup 2 to 94.5 for subgroup 4.

(vii) The extent of participation in academic achievements prior to university, including involvement in local to international academic competitions, attainment of top positions within their high school cohorts, and enrollment in higher-level Advanced Placement or International Baccalaureate courses, did not significantly impact the average grades of LTP students. Grades of students with no participation (subgroup 1; 94.2) were comparable to those of students with varying levels of participation (subgroups 2, 3, and 4; 94.1, 93.9, and 94.2, respectively), despite that most students were engaged in such activities during high school.

(viii) The extent of involvement in work experience before commencing university positively influences the outcomes of LTP students, particularly when such involvement does not necessitate extensive time commitment as leaders of work teams. Students in subgroups 2 and 3 achieve higher average grades (94.5 and 94.1, respectively) compared to those in subgroup 1 (94.0; no participation); whereas students in subgroup 4 attain lower grades (93.8). Work experience is valuable for receiving compensation or recognition, but it can also distract students from their academic pursuits. It is noteworthy that most LTP students have had prior work experience before entering university.

(ix) The degree of involvement in international experiences prior to university enrollment positively influences LTP students' average grades. Students engaged to some extent in international activities (subgroups 2, 3, and 4) achieved higher grades (94.2, 94.5, and 94.3, respectively) compared to those with no participation (subgroup 1; 94.0). However, over 50% of LTP students were unable to participate in international activities probably due to their disadvantaged economic circumstances during high school.

(x) The extent of involvement in community activities prior to university enrollment is another factor that positively impacts the academic performance of LTP students. Students who did not participate obtained a lower grade (93.5) compared to those engaged in community activities to varying degrees (subgroups 2, 3, and 4), with average grades of 94.0, 94.1, and 94.2, respectively. It is noteworthy that a significant portion (more than 90%) of LTP students were engaged in community participation during their high school years.

(xi) Students who were not the first in their families to attend university, achieved slightly higher grades (0.5 points) compared to those who are the first generation to pursue higher education. The proportion of these two groups within the student sample is nearly identical, with each comprising approximately 50% of the sample.

(xii) Local students (subgroup 0) achieved slightly higher grades (94.1) compared to students originating from areas outside the city where the campus is located (subgroup 1, 93.4). The student sample sizes of both subgroups are very similar.

(xiii) The type of zone of the student's address in high school has noticeable impact on LTP students' learning outcomes at the university. Students coming from rural zones (subgroup 1) obtained smaller average grades (92.9) than those coming from semirural (subgroup 2; 93.3) or urban environments (subgroup 3; 94.0). In our student sample more than 70% of the students came from urban areas.

(xiv) Younger students demonstrated superior academic performance compared to their older counterparts. Among students aged 18 and under or between 19 and 21, the program's average grade stood at 93.8 and 93.9, respectively. However, for students aged 22 years or older, this average slightly decreased to 92.8. The largest portion of the student demographic fell within the age range of 19 to 21 years (63.6%), followed by students aged 18 years or younger at 30.5%, while only 5.90% comprised students aged 22 years or older.

(xv) Lastly, concerning the gender of LTP students, females obtained slightly higher grades than males (94.0 vs. 93.5).

Based on the previous analyses, the factors or activities undertaken during high school that correlate with LTP undergrad students' academic achievements can be categorized into three groups:

(a) Factors that positively influenced undergrad LTP students' academic performance to varying degrees include: (i) participation in student group activities, (ii) engagement in international experiences, (iii) involvement in community activities. Additionally, certain demographic factors showed positive impacts on LTP students' academic outcomes, such as: (i) not being the first in their families to attend university, (ii)

being of a younger age, (iii) originating from urban residential areas rather than semirural or rural environments, (iv) residing in the same city as the university campus, and (v) being female.

(b) The factors that positively influenced student academic performance, provided they did not consume excessive time that could detract from their studies included: (i) prior work experience, (ii) participation in sports activities, and (iii) engagement in cultural activities.

(c) Factors that demonstrated little effect on student academic performance include: (i) the scope of students' projects, (ii) the extent of involvement in leadership activities, and (iii) participation in local, national, or international academic competitions.

VI. Conclusions and Future Work

The Tecnologico de Monterrey, a premier private university in Mexico, launched its *Leaders of Tomorrow Scholarship Program* (LTP) in 2013. This initiative aims to provide comprehensive support to young outstanding students, who also exhibit leadership qualities within their communities, by granting them a 100% scholarship to pursue their university degree. This program represents a significant initiative that should serve as model for other mayor universities globally, as it effectively nurtures and empowers valuable talent spread in far, low-income communities, especially in developing countries such as Mexico.

This study highlights the notable positive correlation between student participation in cocurricular activities and improved academic outcomes among students enrolled in the LTP. This suggests the promotion of cocurricular activities by academic institutions at suitable levels of intensity and extent to not distract student time from their academic activities. Furthermore, this research highlights the importance of designing policies to support LTP students from rural areas and those attending campuses situated outside their communities.

Nevertheless, it is important to acknowledge certain limitations within this study, such as the complexity arising from the analysis of multiple variables, which may hinder some results' interpretation. For future research endeavors, the adoption of dimensionality reduction techniques, including Principal Component Analysis, Clustering, or Classification, is recommended. These methods could offer more intuitive visual outcomes, aiding more effectively in the identification of underlying patterns within the student's data. Moreover, considering personality traits as variables in future investigations could shed light on those conducive to academic success.

In future work, we aim to extend this analysis using multivariate techniques, including ANOVA and Factorial Analysis, to provide a more comprehensive statistical evaluation of the trends revealed by the database.

Acknowledgments

The authors acknowledge financial support from the Writing Lab, and the Institute for the Future of Education, from the Tecnologico de Monterrey, México, in the production and presentation of this work. We also thank the Cyber Learning and Data Science Laboratory, and the Science Department and Computer Department of the School of Engineering and Science of the Tecnologico de Monterrey, Mexico City Campus.

References

[1] United Nations (2024) [Online]. Available: https://unric.org/en/united-nations-sustainable-development-goals/

[2] Tecnológico de Monterrey (2024). Líderes del Mañana. [Online]. Available: https://lideresdelmanana.itesm.mx/?_gl=1*1ky3iq5*_ga*OTU1MzEzMTQ1LjE3MTM5Mjk4Nzk.*_ga_D9LSDN87GD*MTcxNTAxMDY5NC4xMi4xLjE3MTUwMTE4MTcuMzQuMC4w

[3] Fundación BBVA (2024). [Online]. Available: https://www.fundacionbbva.mx/convocatoria-becas2024/

[4] Banco Santander (2024). [Online]. Available: https://www.bancosantander.es/universidades/educacion/becas

[5] Citibanamex (2024). [Online]. Available: https://www.banamex.com/compromiso-social/programa/?id=30

[6] Fundación Telmex (2024). [Online]. Available: https://www.fundaciontelmextelcel.org/becas

[7] Fundación Carlos Slim (2024). [Online]. Available: https://fundacioncarlosslim.org/educacion-2/

[8] Gobierno de México IMJUVE (2024). [Online]. Available: https://www.gob.mx/imjuve/acciones-y-programas/programa-de-becas-academicas-333477

[9] OAS (2024). [Online]. Available: http://www.oas.org/es/becas/

[10] Gobierno de México INPI (2024). [Online]. Available: https://www.gob.mx/inpi

[11] Gobierno de México boletín 18 (2024). [Online]. Available: https://www.gob.mx/sep/es/articulos/boletin-18-destina-gobierno-de-mexico-mas-de-142-mil-mdp-en-becas-y-planteles-educativos-sep.

[12] C.A. Tomlinson & J.M. Jarvis, "Case Studies of Success: Supporting Academic Success for Students with High Potential From Ethnic Minority and Economically Disadvantaged Backgrounds", Journal for the Education of the Gifted. Vol. 7(3), 2014, pp. 191-219. Available: https://doi.org/10.1177/0162353214540826

[13] G.A. Rivera. "Leading Schools in Poverty: Principal Perceptions, Key Factors, and Strategies to Academic Success". Ph.D. Thesis, West Texas A&M University, USA, 2021, Available: https://wtamu-ir.tdl.org/items/40847bde-db9c-443d-a965-3ce83751409d.

[14] D. C. Ribar & R. Rubenstein, "Gaining, Losing, and Regaining Merit-based Scholarships", Education Finance and Policy, Vol. 18, Num. 4, 2023, Pp. 597–622, Available: https://doi.org/10.1162/edfp_a_00385

[15] M.P. Wand & M.C. Jones, (1994). *Kernel Smoothing* 1st. ed, Chapman and Hall/CRC, 1994. Available: https://doi.org/10.1201/b14876.

[16] B. L. Welch, (1938). "The Significance of the Difference Between Two Means when the Population Variances are Unequal". Biometrika, Vol. *29,* Num. 3/4, 1938, Pp. 350–362, Available: https://doi.org/10.2307/2332010.