High School Scores to College Performance: Exploring the Predictive Validity of Grade 12 Standardized Test Scores for University GPA

Abdurrahman Ahmad Wahab [1], and Hemin Ibrahim [2]

[1] Department of English Language Teaching, Faculty of Education, Tishk International University, Erbil, Iraq.
[2] Department of Information Technology, Faculty of Science, Tishk International University, Erbil, Iraq.
Correspondence: Abdurrahman Ahmad Wahab, Department of English Language Teaching, Faculty of Education, Tishk International University, Erbil, Iraq.
Email: abdurrahman.wahab@tiu.edu.iq

Doi: 10.23918/ijsses.v11i4p120

Abstract: This study examines the correlation between High School Scores (HSS) and academic achievement in the ELT Department at Tishk International University, Erbil, Iraq, aiming to assess the validity of HSS as a predictor of university GPA. The study employs traditional statistical methods and advanced AI-based algorithms to categorize and analyze a dataset comprising 225 students admitted between 2014 through 2019. As AI-based algorithms, Linear Regression model yielded a mean square error (MSE) of 0.2228 and R^2 of 0.1774, while Neural Network model produced an MSE of 0.2383 and R^2 of 0.1204. These results indicate a weak to moderate correlation between HSS and GPA, indicating that other factors significantly influence GPA. The study concludes that while HSS can serve as an initial indicator of potential academic success, it is not a definitive predictor, suggesting the need for a more comprehensive approach to evaluating student performance in the ELT program.

Keywords: High School Score, College Admission, University GPA, Machine Learning, Linear Regression, Neural Network

1. Introduction: High School Standardized Testing in the Kurdistan Region of Iraq (KRI)

The high school standardized testing in the Kurdistan Region of Iraq (KRI), known as the Baccalaureate Exam or Ministerial Exam, serves as the sole criterion for secondary education graduation and university admission (Mohammed et al., 2021; Vernez et al., 2016; Wahab, 2019). Over the past three decades, the examination has undergone significant changes. However, it has consistently focused on the contents of the grade 12 school textbooks for each academic subject. Since 2014, each subject's exam consists of 50 multiple-choice questions, each is worth 2 points, for a total score of 100. While a passing score in each subject is 50% or above (MoE-KRG, 2009), the emphasis is on the total score, which determines university placement and program admission in the KRI.

All public and private schools adhering to the Ministry of Education's curriculum culminate in the standardized Ministerial Exam. However, international private schools follow a different path, often administering Cambridge IGCSE A-Level exams or the SAT.

Received: 08.09.2024
Accepted: 23.09.2024
Wahab, A. A., & Ibrahim, H. (2024). High School Scores to College Performance: Exploring the Predictive Validity of Grade 12 Standardized Test Scores for University GPA. *International Journal of Social Sciences & Educational Studies, 11*(4), 120-138.

The Ministry of Education then equalizes these scores to align with the Ministerial Exam scores. This standardization ensures that all students in the KRI are admitted to higher education institutions based on a uniform criterion.

1.1 Test Scores and Academic Hierarchy

In the Kurdistan Region's higher education system, a clear hierarchy of colleges and academic programs is primarily determined by high school test scores (hereafter referred to as HSS). Historically, medical colleges and programs have occupied the top tier, accepting only the highest HSS. They are followed by engineering, material sciences, and humanities and social sciences programs. At the bottom of this hierarchy are two-year technical and vocational programs that award a high diploma.

There is also a hierarchy between public and private universities. This hierarchy is shaped both by government policy design and by favorability. Public universities are generally more favored due to the tuition-free education offered to most students. However, in recent years, the Ministry of Higher Education and Scientific Research-KRG has introduced a parallel admission system. This system allows some students to enter their desired programs by paying tuition, granted that their HSS still falls within the accepted range. Moreover, the tuition fees in the parallel system are lower than those in private universities. Public universities also benefit from the full political and financial support of the KRG, which has historically made the public and employers recognize their diplomas without much hesitation. Additionally, the Ministry of Higher Education and Scientific Research-KRG has set lower minimum HSS requirements for programs at private universities compared to their public counterparts, suggesting a more rigorous admissions criterion in public institutions.

1.2 A Centralized University Admission System

Students applying to public universities in the KRI use Zankoline, which is an online electronic application system administered jointly by the KRG Ministries of Education and Higher Education. By entering their unique numbers, students can view the universities and programs available to them based on their total HSS. They then select their preferred universities and colleges from the options presented. The system allows students to choose multiple universities and programs if their HSS qualifies them for those choices. The availability of seats in the desired universities and programs, which is set by the Ministry of Higher Education, determine the students' final placement. Students with higher scores naturally have a better chance of being admitted to their preferred programs.

1.3 HSS as the Primary Admission Criterion

HSS serves as the central criterion for students applying to university programs. University admission is based on a combination of HSS and the availability of programs for which their scores qualify them. Public universities do not directly admit students; instead, they inform the Ministry of Higher Education about the number of students they can accommodate in specific programs. The availability of seats varies each year, influenced by the number of high school graduates and government planning for overall student admissions. Typically, a minimum score is established for each academic program, and the electronic system then allocates students across universities and their programs accordingly.

1.4 Admission Process in Private Universities

The admission process for private universities follows a slightly different path. Like public universities, the Ministry of Higher Education determines the maximum number of students each program can admit and the minimum HSS required for admission in all private universities. This ensures uniform admission criteria for similar programs across private institutions in the KRI. Students wishing to apply to private universities obtain application forms from each university and indicate their desired programs. These forms are then submitted to the universities' admissions offices, where the information is entered into an electronic system similar to Zankoline. However, access to this system is restricted to admissions personnel and the Ministry. The Ministry announces the results of the admission process through this system, usually months after the public university admissions process has concluded. Consequently, students often choose to study at private universities only after assessing their chances of admission to public universities. Those satisfied with their public university placements typically remain in those programs, while those who did not secure their desired spots may turn to private universities as an alternative.

1.5 HSS in the ELT Department

The Ministry of Higher Education-KRG sets lower minimum HSS for programs at private universities compared to their public counterparts. Medical and engineering programs at private institutions are at the top of the academic hierarchy, with their minimum HSS fluctuating only slightly over the years. Humanities and social sciences programs rank lower in this hierarchy, resulting in even lower HSS requirements for admission. For instance, to be admitted to the English Language Teaching (ELT) Department at TIU, a minimum HSS of 57 is required. The Ministry also caps the number of students the ELT Department can admit at 120 per academic year. However, the highest number of students admitted since the department's establishment has been 84.

1.6 HSS and Academic Assumptions

Relying on HSS as the sole criterion for high school graduation and university admission in the KRI underscores several assumptions about the significance of standardized test scores in academic and financial decision-making. Given that HSS is the only measure of college readiness and a predictor of academic success in higher education programs in the KRI, it is assumed that higher HSS correlates with greater potential for success in college. University programs at the top of the academic hierarchy admit only students with the highest HSS, assuming these scores reflect academic rigor.

Additionally, a higher HSS provides students with more options in terms of college admissions, granting them access to a wider range of academic programs and institutions. It also has implications for financial aid. In public universities, students admitted through the parallel system pay tuition fees because their HSS is slightly lower than those of their peers. Similarly, many private universities offer financial aid, including scholarships and discounts, based on students' HSS. This system assumes that higher HSS is an indicator of future college performance. For example, the standard financial aid offered by the ELT Department at TIU has been 50% of tuition fees for students with an HSS of 80 or above, aimed at attracting high-scoring

students to the program. There have even been instances where students with an HSS of 90 and above have received a 100% scholarship on tuition fees.

Ultimately, while these assumptions grant significant weight to the HSS, shaping major policy and decision-making in the KRI educational system, they are not supported by specific research or empirical evidence tailored to the context of the Region.

2. Literature Review: Standardized Test Scores are at Best Controversial

Standardized testing as a criterion for assessing school and student performance, college admission, and other academic decision-making has been a subject for debate and controversy. Some argue for their complete elimination in admission processes, while others call for a more diverse approach where a combination of criteria is used to present a better predictive measure of academic and professional success (Wightman, 2000; Seyfarth et al., 1994).

Proponents of standardized testing prioritize these tests mainly because they represent a uniform, reliable, and objective tool for decision-making and student selection mechanisms. Test scores collected from large-scale testing provide a simple, easy-to-use criterion for sorting out applicants who have diverse backgrounds and experiences. Lee et al. (2023) argue that many colleges have prioritized using standardized testing as a tool to manage large pools of college applicants. In contrast to high school GPA, standardized tests have been viewed as "methodologically rigorous", that they provide "a more uniform and valid yardstick for assessing student ability and achievement" (Geiser & Santelices, 2007, p. 4). These tests are also presented as being more precise, reliable, and consistent in measuring students' overall ability and academic achievement (p. 4). Some researchers also see the value of the unique information that standardized testing adds to any admission process, particularly in helping decision-makers "compare applicants who graduate from schools that differ, sometimes substantially, in grading rigor and academic standards" (Wightman, 2000, p. 95). Others further highlight the reliability and objectivity of standardized test scores "because all students are judged based on the same tasks under the same conditions" (Allensworth & Clark, 2020, p. 198), hereby eliminating bias and misperception in the process.

However, critics of standardized testing offer counter arguments that undermine or at least explicate the underlying assumptions of standardized testing. From the outset, standardized testing is criticized for its scope by focusing on "a very narrow set of skills and abilities" instead of stressing other traits and competencies that are required for success in academia and beyond (Powers & Kaufman, 2002). Allensworth and Clark (2020) challenge a key assumption of using standardized tests in college admission processes, which is that such exams are strong and consistent indicators of college readiness. They argue that research shows that other measures, such as high school GPA, are better and more consistent indicators of college readiness and success. Wightman (2000) contends the limited meaning and scope of test scores as indicators of student success and other factors in college admission, suggesting instead to move from basing admission on single test scores to a "battery or profile of strengths and weaknesses among applicants" (p. 90). Geiser (2008) argues that other indicators, such as high school GPA that is "based on repeated sampling of student performance over several years", are much stronger and more reliable predictors of college success than scores obtained through single-sitting tests (p. 2). Along the same lines, Seyfarth at al. (1994) argue that admission processes can rely on other, more superior assessment methods

that measure student and school performance throughout a longer period of time and in different areas of relevance.

Furthermore, Allensworth and Clark (2020) argue that standardized tests do not reflect student success in college. Rather, test scores "represent factors associated with the student's school rather than the student" (p. 209). In fact, as Zwick (2012; 2013) also argues, there is a direct connection between socioeconomic status (SES) of students and schools with test scores. Hence, standardized test scores are often indicators of the socioeconomic status of the students, the school, and the students' communities rather than their academic skills. In other words, test scores do not indicate academic capabilities as much as they reflect SES, mainly because wealthier families and communities can access better test preparation and coaching services that eventually translate into higher scores (Zwick, 2012).

Researchers also refer to the negative impact of relying heavily on standardized test scores on the meaning and value of education and school activities that are not generally reflected in such tests. Proponents of standardized testing often confuse measurements with values, assuming that using specific measurements in education will improve educational values (Portelli & Vilbert, 2002). Furthermore, Casas and Meaghan (2001) argue that using large scale testing for sorting and ranking students for college placement does not help improve the quality of education. Rather, such tests have been problematic and have negatively impacted especially low income and minority children (p. 147). Moreover, when high schools determine their success in educating and preparing students for college based on test scores only without tracking their students' success in college, schools "may be misestimating the effects of their practices on students' college readiness" (Allensworth & Clark, 2020, p. 210). Schools might prioritize practices, such as teaching to the test and test coaching, that could lead to raising test scores without fully understanding if these scores translate into success in college. Also, since standardized testing heavily depends on multiple-choice question formats, these tests "may inadvertently lead to the creation of false knowledge" (Roediger III & Marsh, 2005, p. 1155). Moreover, even though advocates of standardized tests highlight objectivity and reliability of these tests, critics argue that standardized tests have not resolved gender, racial, and socioeconomic biases. They have criticized the "non-holistic nature" of these tests in admission processes (Lee et al., 2023).

The debate over standardized testing as a criterion for academic decision-making highlights significant concerns and conflicting perspectives. While proponents emphasize the objectivity, reliability, and uniformity of these tests, critics challenge their narrow focus, questioning their effectiveness in predicting long-term success and highlighting the socioeconomic biases they may perpetuate. As a result, there is a growing consensus on the need for a more comprehensive and nuanced approach to student assessment and college admissions, one that integrates multiple indicators to better capture university applicants' diverse talents and potential.

2.1 Research Questions

Considering the controversial nature of standardized testing and the significant assumptions underpinning the heavy reliance on HSS as the sole indicator of successful secondary education completion, college readiness, and academic success in higher education in the KRI, this study seeks to address the following questions:

1. What is the correlation between HSS and academic achievement in the ELT Department?
2. How reliable is HSS in predicting university GPA?

3. Methodology

To answer the research questions, this study employs statistical methods and AI-based algorithms to analyze data from students enrolled in the English Language Teaching (ELT) Department at Tishk International University (TIU) in Erbil since 2014. AI-based algorithms are computational processes that use machine learning models to analyze data, which, unlike traditional statistical analysis, focus on predictive accuracy rather than interpretability of the data. AI-based algorithms recognize patterns and optimize relationships to provide highly accurate predictions. Utilizing both methods enables a comprehensive evaluation of the high school standardized test score (HSS) as a reliable predictor of university achievement.

3.1 Data Collection

The study examines two major variables: students' HSS and their college GPA. Although the ELT Department has admitted students since 2008, the study focuses on electronically available data, starting from 2014. Data from students who are currently enrolled (those admitted in 2020, 2021, 2022, 2023, and those from previous years who are yet to graduate) has also been excluded as these students do not yet have a complete GPA. Additionally, GPA scores below 2.0 have been excluded, as they are not considered indicative of success or graduation from the ELT Department.

The study also considers the students' high school tracks—literary or scientific. Without conducting an in-depth analysis, the study seeks to observe how students' high school tracks may have influenced their achievement in the ELT Department, a social science department more aligned with the high school literary track.

The data for this study was provided by the Office of Student Affairs at TIU. While some data was available in MS Excel format, making it easier to extract, much of the data was in PDF format, requiring additional effort to extract and clean. Ensuring the accuracy of the data involved double-checking each student's full name (including Kurdish and Arabic names), student number, HSS, and GPA. There were challenges to the data cleaning process, such as mismatches between spelling names in students' registration records and their transcripts (due to varying transliterations of Kurdish and Arabic names), duplicate names, as well as difficulties in extracting data from PDF forms. For instance, some records mistakenly listed students' second names as their first names and their first names as their last names, complicating the process of matching student names and their corresponding data.

After thorough data cleaning and confirmation, the study achieved a solid dataset comprising 225 students from the academic years 2014-2015 through 2019-2020. These students have all been admitted to and successfully completed their academic program in the ELT Department at TIU in Erbil.

3.2 Traditional Statistical Method

This study therefore classifies the data into quartiles and ranges to investigate the relationship between students' HSS and their GPA using the traditional statistical methods of mean, median, and standard deviation. The dataset is arranged into three separate categories: Quartile Category, HSS Range Category, and GPA Range Category. By calculating the mean and standard deviation for each category, the study assesses the central tendency and variability of academic performance within each of them. The median provides insights into the central value of the distributions, allowing for a detailed analysis of how HSS relates to GPA across different categories.

The Quartile Category shows a division of the dataset based on student HSS, from lowest to highest. This approach enables us to focus on the trends in GPA and HSS through different quartiles, thereby allowing us to understand the relationship between HSS and GPA. Comparing these quartiles, the study identifies patterns and potential discrepancies between students' high and low HSS vis-à-vis their academic performance.

Additionally, the HSS Range and GPA Range Categories offer complementary perspectives. The HSS Range Category groups students based on predefined intervals of HSS, starting from the lowest HSS in the dataset and increasing each category by 10 (such as 58-68, 68-78, 78-88, and 88-98). This categorization explains how GPA distributions vary within each HSS range, helping to explain the relationship between high school performance and university achievement.

The GPA Range Category, on the other hand, groups students based on their GPA ranges, starting from lowest GPA in the dataset (which is 2.09) and moving up by half a point (e.g., 2.09-2.5, 2.51-3.0, 3.01-3.5, and 3.5-4.0). This categorization provides insights into how HSS ranges connect with different levels of GPA, allowing us to observe whether higher HSS scores consistently correspond with higher GPAs.

Together, these categorizations offer a nuanced view of the dataset, allowing for a comprehensive analysis of how HSS influences university achievement and identifying trends that can inform educational policy and decision making.

3.3 Machine Learning Methods

In addition to traditional statistics, the study has adopted advanced machine learning algorithms to further explain the correlation of HSS and university GPA. Machine learning methods can offer a robust approach to detect and show complex patterns that may not be marked from simple statistical methods. In this study we have applied two machine learning algorithms: Linear Regression and Neural Networks.

3.3.1 Linear Regression

Linear regression is a simple machine learning model of supervised learning for predicting numerical or continuous outcomes (James et al., 2023). It is an approach for predicting a quantitative response Y on a single predictor variable X. It presumes that there is nearly a linear relationship between the variable X and the continuous output value of Y. The general formula of linear regression is:

$$y = \beta_{0+}\beta_1 x$$

Where y is the predicted outcome, x is the independent variable or predictor. $\beta 1$ is the slope, indicating how much y changes with a unit increase in x. $\beta 0$ is the y-intercept, indicating the starting value of y when $x = 0$.

In this study linear regression has been adopted to investigate the relationship between HSS and university GPA and how much it will affects getting higher GPA at university. Linear regression as a fundamental statistical approach is used to this relationship. The main goal of using linear regression in this work is to identify the strength of the association between high school scores and university GPA.

Additionally, we split the dataset to train set and test set by 80% and 20% respectively. Also, only the HSS feature is considered for predicting GPA. This ensures a focused model of how HSS influences university academic success. The train set is used to train the linear regression model, while the test set is used for evaluating.

3.3.2 Neural Network Model

A neural network is a computational machine learning model inspired by the same mechanisms of biological neural networks found in the human brain (De Wilde, 2013). These networks are capable of uncovering complex patterns and solve various problems through flexible learning algorithms.

In this study, both neural network and linear regression models are used to analyze the correlation between HSS and university GPA. The neural network model contains three main layers: an input layer, a hidden layer, and an output layer. The input layer receives the relevant feature from the students' HSS. Next, the hidden layer enables the network to learn and discover complex patterns and relationships between the variables within the dataset. The output layer produces the result that makes predictions for continuous values. Additionally, it predicts the students' university GPA by using the learned patterns. Compiling the Keras model (Chollet, 2015), the Adam optimizer (Kingma & Ba, 2014) has been utilized, and the model was trained in 50 epochs, and each time one batch size of 10 was used. This means, the algorithm does 50 complete iterations (epochs) with the whole dataset. In each iteration, the whole dataset is divided into smaller sets called batches; each batch contains 10 samples. The model processes these 10 samples each time and updates its parameters after each batch. Once all the batches in the dataset have been processed, it is said that the algorithm has finished one epoch. Then comes the next epoch and repeats the iterations a total of 50 times.

To understand the strength of the relationship between HSS and university GPA, both linear regression and neural network models use Mean Squared Error (MSE) and the coefficient of determination (R^2), which are effective for measuring accuracy and identifying the relationship between the variables.

Mean Squared Error (MSE) measures the average squared differences between HSS and university GPA as explained in the following equation:

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2$$

The variable y_i represents each student's actual university GPA, while $\hat{y}_i$ represents the GPA predicted by the model based on the HSS. To measure the accuracy of these predictions, we calculate the difference between the actual and predicted GPAs for each student (from 1 to n), square those differences, and then take the average of all these squared differences across all students. This average gives us the Mean Squared Error (MSE).

A lower MSE suggests that the model is a better fit since it indicates smaller discrepancies between the predicted and actual values. This metric is useful in evaluating how closely the model's predicted university GPA aligns with the actual outcomes based on HSS.

The coefficient of determination (R^2) is another important statistical measure that shows how well the model's predictions align with the actual data. It also indicates the percentage of the variation in university GPA that can be explained by the HSS. The following equation shows the R^2:

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y}_i)^2}$$

Where y_i represents the actual university GPA for each student, $\hat{y}_i$ denotes the predicted GPA based on the HSS. The $\sum_{i=1}^{n}(y_i - \hat{y}_i)^2$ computes the total squared error between the actual GPA and the predicted GPA. Additionally, $\sum_{i=1}^{n}(y_i - \bar{y}_i)^2$ represents how much all the actual GPA differ from the average GPA, $\bar{y}$.

R^2 ranges between 0 to 1, with higher values indicating a better fit of the model to the data, and it means indicating a strong relationship between HSS and GPA. Equally, lower R^2 indicates a weak relationship between HSS and GPA.

By combining both traditional statistical method and machine learning algorithms to analyze the dataset comprising 225 students and evaluate the correlation of their HSS and GPA, this study offers an analytical model of a robust understanding and future prediction of such correlation.

3.4 Study Limitations

While this study examines key variables, namely HSS and GPA, to understand students' academic progression from high school to university, it does not account for variations within the HSS (such as individual scores in the English language test) that could influence and clarify the overall findings. Additionally, the study does not explore the numerous factors within the ELT program at TIU that could affect student GPA. Other significant variables, such as gender, socioeconomic status, age, and students' urban or rural backgrounds, are also not considered, despite their potential impact on students' academic journey and progression over the four years of study in the ELT Department.

To provide a more comprehensive understanding of the relationship between secondary and tertiary education, future studies could consider a narrative approach that tracks how student grades evolve throughout their undergraduate studies. These studies could also compare students' grades in high school coursework, not just their HSS obtained from standardized exams.

While this study focuses on data from the ELT Department at TIU, the same model could be expanded to include data from other programs at TIU, as well as programs from other private and public universities in the Kurdistan Region of Iraq.

4. Results and Discussion

This study examines the data of 225 students who have been admitted to and successfully completed the program in the ELT Department at TIU in Erbil during the academic years 2014-2015 through 2019-2020. The analysis focuses only on the data available on admissions and graduations from 2014 through 2019, excluding students admitted in 2020 and later, as they have not yet completed their program.

Table 1. GPA and High School Score (HSS) across the dataset, including scientific and literary categories

Category	Metric	Overall	Scientific (144 students)	Literary (81 students)
University GPA	Mean	2.97	3.04	2.85
	Median	2.91	3.03	2.72
	Standard Deviation	0.53	0.53	0.49
	Minimum / Maximum	2.09 / 3.98	2.15/3.98	2.09/3.96
High School Score	Mean	67.36	66.78	68.38
	Median	65.00	64.59	66.14
	Standard Deviation	7.50	7.40	7.62
	Minimum / Maximum	58.00 / 94.06	58.00/94.06	58.43/93.70

As Table highlights, the minimum and maximum HSS scores in the dataset are 58.00 and 94.06, and the minimum and maximum GPA achieved are 2.09 and 3.98. It also shows the mean, median and standard deviation for each of these achievements.

To analyze the correlation between students' HSS and their university GPA, the study has employed two distinct methods: traditional statistical analysis and machine learning methods. In the traditional statistical method, the data is categorized in three distinct ways: Quartile Category, HSS Range Category, and GPA Range Category. Each categorization offers a unique perspective on the connection between the two major variables, providing a more thorough and nuanced analysis. This approach enables the study to generate more comprehensive and actionable insights.

4.1 Quartile Category

The students are divided into four quartiles based on their HSS, from the lowest (Q1) to the highest (Q4) (Figure). Each quartile represents approximately 25% of the total student population included in the study, with Q1 containing 25.33% and Q2, Q3, and Q4 each containing 24.89%.

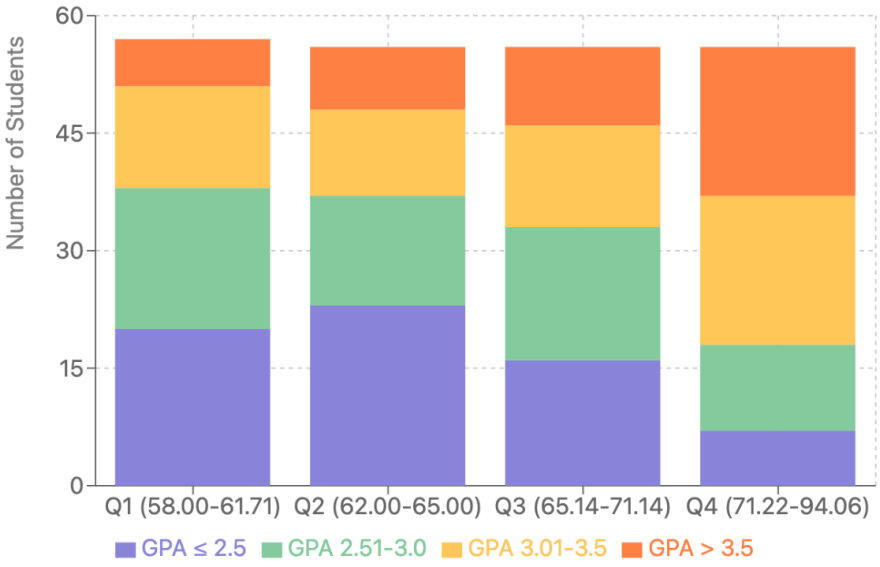

Figure 1: GPA distribution across quartiles (Q1-Q4) for the whole 225 students in the dataset

4.1.1 Q1 Category

In the Q1 category, the students have the lowest HSS, with a mean of 60.02 with a standard deviation (SD) ± 1.00. Despite this, their university GPA mean is 2.84 with SD ± 0.47, with a median GPA of 2.69, indicating a fairly wide range of academic performance (2.15 - 3.92).

4.1.2 Q2 Category

The Q2 category shows a slight increase in the mean HSS to 63.69 with SD ± 0.90. However, this increase in HSS by 3.67 points compared to Q1 does not correspond to any change in the mean GPA, which stays the same at 2.84 with SD ± 0.55. This suggests that for the first half of the student population (Q1 and Q2), a higher HSS does not necessarily translate into a higher GPA.

4.1.3 Q3 Category

In the Q3 category, the mean HSS increases further to 67.81 with SD ± 1.91, resulting in a modest increase in the mean GPA to 2.96 with SD ± 0.53. This slight improvement (0.12) in GPA, despite a 4.12-point increase in HSS, indicates that the connection between HSS and GPA is still weak for this group.

4.1.4 Q4 Category

Finally, in the Q4 category, which includes students with the highest HSS (mean of 78.04 with SD ± 6.16), the mean GPA increases more significantly to 3.25 with SD ± 0.45. This reflects a 0.29-point increase in GPA, corresponding to a 10.23-point increase in mean HSS compared to Q3. This trend suggests that the relationship between HSS and GPA becomes stronger in the higher quartiles, with students in Q4 showing the highest academic performance at the university level.

According to the quartile categorization (Figure), the data suggests a weak to moderate connection between HSS and University GPA. In the Q1 and Q2 categories (lower quartiles), despite a noticeable increase in HSS (from 60.02 to 63.69), there is no corresponding increase in GPA (both have a mean GPA of 2.84). This suggests a very weak or negligible connection between HSS and GPA in the lower quartiles. As HSS in the Q3 category (medium-high quartile) increases further (mean of 67.81), there is a slight improvement in GPA (mean of 2.96). This indicates a weak positive connection between HSS and GPA in this quartile. The strongest relationship is seen in the Q4 category (high quartile). A substantial increase in HSS (mean of 78.04) is associated with a more noticeable increase in GPA (mean of 3.25). This shows a moderate positive connection in the highest quartile.

Overall, the connection between HSS and GPA appears to be weak in the lower quartiles (Q1 and Q2) but strengthens in the higher quartiles (Q3 and Q4). This suggests that higher HSS is somewhat predictive of higher GPA, particularly for students in the top quartile. However, the correlation is not consistent across all quartiles, implying that other factors may significantly influence GPA, especially for students with lower HSS.

4.2 HSS Range Category

The students were categorized based on their High School Scores (HSS) into four distinct groups (Figure): Low Score Category (LSC), Medium-Low Score Category (MLSC), Medium-High Score Category (MHSC), and High Score Category (HSC). These categories span HSS ranges from 58 to 98, with each group representing a different percentile. As we move from LSC to HSC, the size of the groups decreases significantly.

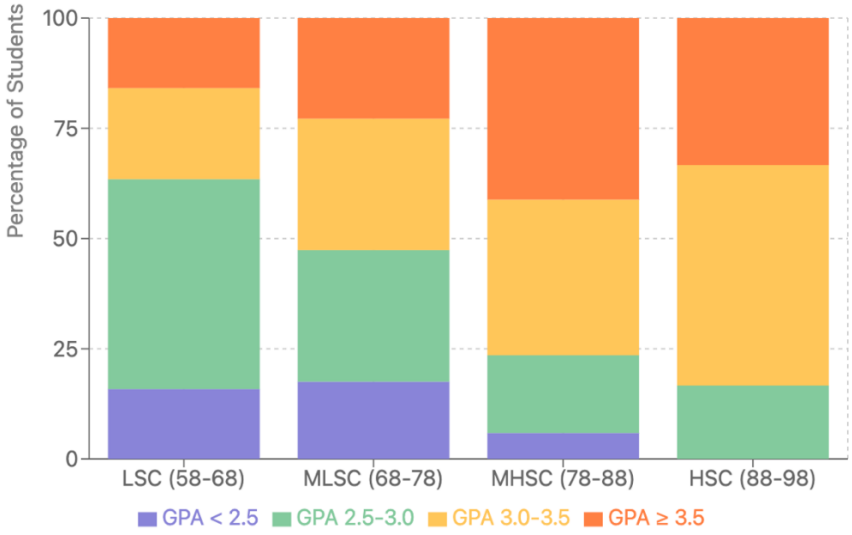

Figure 2: GPA distribution across High School Score (HSS) ranges representing different percentiles

4.2.1 Low Score Category (LSC): HSS 58 – 68

The HSS range in this category is 58 to 68, which includes 145 students (64.44% of the total). The university GPA range in this category is 2.15 to 3.96. Despite being the largest group with the lowest HSS, the LSC exhibits a wide range of university GPAs, reflecting a significant variability in academic performance. The standard deviation of GPA in this category is 0.51, the highest among all groups. Remarkably, 23 students in this category achieved a GPA of 3.5 or higher, despite their lower HSS.

4.2.2 Medium-Low Score Category (MLSC): HSS 68 – 78

The HSS range in this category is 68 to 78, which includes 57 students (25.33% of the total). The university GPA range in this category is 2.09 to 3.97. Although it has a smaller number of students, this category still shows substantial GPA variation, with the GPA range covering both the lowest (2.09) and the second highest (3.97) scores in the dataset. 13 students in this group have achieved a GPA of 3.5 or above.

4.2.3 Medium-High Score Category (MHSC): HSS 78 – 88

The HSS range in this category is 78 to 88, which includes 17 students (7.56% of the total). The university GPA range in this category is 2.51 to 3.93. With a further increase in HSS, the MHSC group has a smaller size and a more concentrated GPA range. Although only 7 students in this category have achieved a GPA of 3.5 or higher, the group's overall performance suggests a moderate connection between HSS and university GPA.

4.2.4 High Score Category (HSC): HSS 88 – 98

The HSS range in this category is 88 to 98, which includes only 6 students (2.67% of the total). The university GPA range in this category is 3.03 to 3.98. The HSC, representing students with the highest HSS, is the smallest group but includes the highest university GPA in the dataset (they don't have lower

than 3.03). The GPA range is narrower, indicating less variability, and includes the highest GPA recorded in the ELT department. However, only 2 students in this category achieved a GPA of 3.5 or higher.

Overall, the data shows (Figure) that while HSS has some impact on GPA, the relationship is not straightforward. The LSC group has the most variability in GPA, suggesting that factors other than HSS may play a significant role in determining university performance. In contrast, as HSS increases, particularly in the HSC, the GPA range narrows, indicating a stronger, though not absolute, connection between higher HSS and better academic performance.

4.3 GPA Range Category

The dataset is categorized based on students' university GPA ranges, organized into four groups from low to high (Figure): Low Grade Category (LGC), Medium-Low Grade Category (MLGC), Medium-High Grade Category (MHGC), and High Grade Category (HGC).

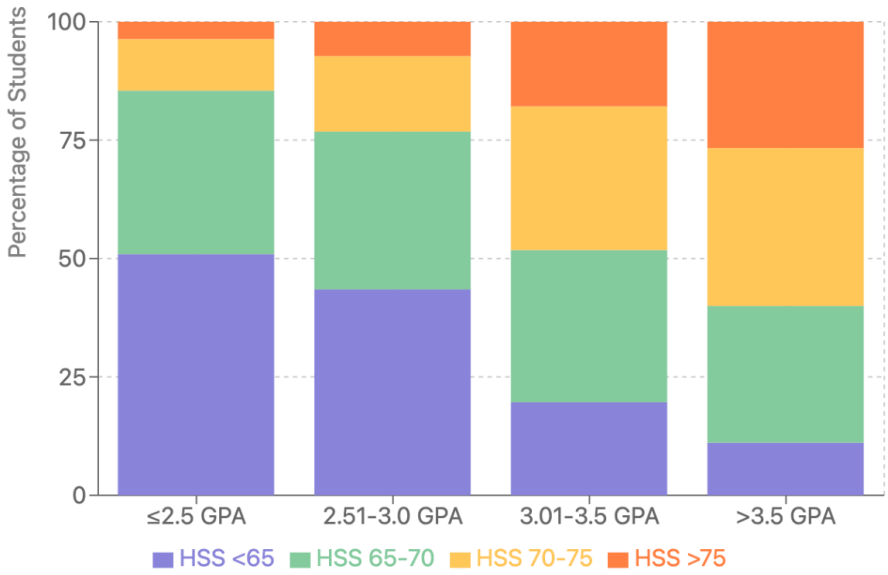

Figure 3. High School Score (HSS) distribution across university GPA ranges from low to high

4.3.1 Low Grade Category (LGC): GPA 2.09 – 2.50

The LGC includes 55 students who comprise %24.44 of the total number of students. The HSS of students in this category range between 58.86 and 75.57. The students in the LGC category generally have lower GPAs, and their HSS is also on the lower end of the spectrum. Interestingly, 18 students in this category have HSS of 65 or above, 65 being the median of HSS in the dataset as shown in Table . However, despite their relatively higher HSS, these students' GPA still remain in the lower category, suggesting that HSS alone is not a definitive predictor of university success.

4.3.2 Medium-Low Grade Category (MLGC): GPA 2.51 – 3.0

The MLGC includes 69 students who comprise 30.67% of the total number of students. This category shows a broader range of HSS, from as low as 58 (the lowest in the dataset as shown in Table) to as high as 82.91, indicating that students with varying HSS can still achieve similar GPA outcomes. Notably, 30 students in this group have HSS of 65 or above, indicating that while these students performed better in the high school standardized exam, it did not necessarily translate to a higher GPA in university compared to those in the higher GPA categories.

4.3.3 Medium-High Grade Category (MHGC): GPA 3.01 – 3.50

The MHGC includes 56 students who comprise 24.89% of the total number of students. This category includes students with a wider range of HSS, with some as low as 58.43 and others as high as 93.70. A substantial proportion (37 students) have HSS of 65 or above, indicating that higher HSS generally correlates with a better GPA, but there are still students with lower HSS achieving stronger university results.

4.3.4 High Grade Category (HGC): GPA > 3.50

There are 45 students in this category who comprise 20.00% of the total number. This category represents the highest achievers in terms of GPA, with a notable range of HSS from 58.86 to 94.06. Despite high GPAs, only 25 students have HSS of 65 or above, while 20 students achieved these high GPAs despite having HSS below 65, showing that some students significantly outperformed their high school performance in university.

It is noted (Figure) that in the two highest GPA categories (MHGC and HGC), 39 students have had HSS below 65, demonstrating that while HSS is a factor in predicting GPA, it is not the sole determinant. Some students are able to achieve high GPAs despite lower HSS. Furthermore, what is consistent across categories is that while higher HSS generally correlates with higher GPA, there is considerable overlap across categories, with students achieving a wide range of GPAs regardless of their HSS.

This study has also adopted both linear regression and neural network models to investigate the correlation between HSS and university GPA. The performance of both models was evaluated using two main metrics: Mean Squared Error (MSE) and the coefficient of determination (R^2).

4.4 Linear Regression Model

Using the linear regression model, we analyzed the whole dataset to obtain a slope to represent the rate at which the GPA changes with respect to the HSS. Figure illustrates the slope as 0.0232. This means that for every 1 score in HSS, the university GPA is predicted to increase on average by 0.0232 points. To put it differently, any 10 score increase in HSS would predict a 0.232 increase in university GPA.

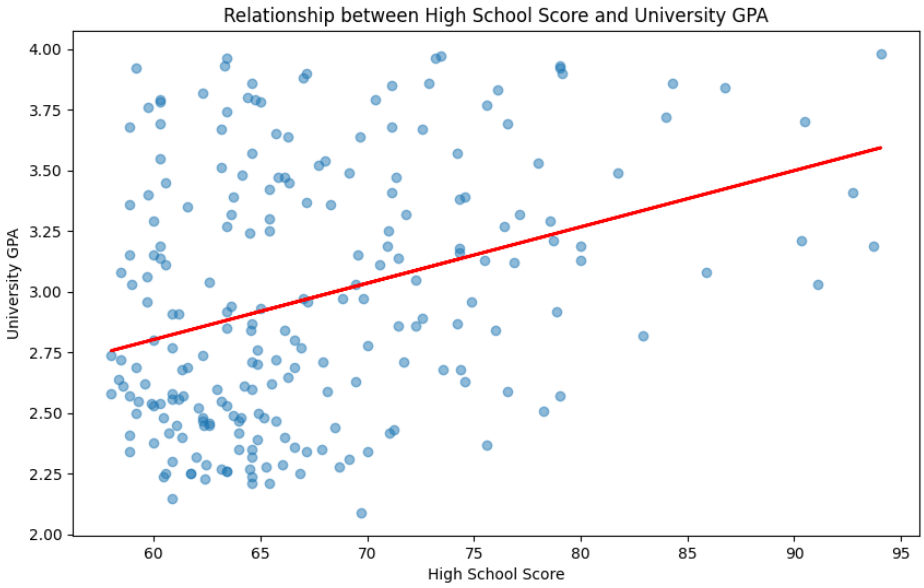

Figure 4: Correlation between High School Score and university GPA

Furthermore, when trained by dividing the dataset into 80% train set and 20% test set, the linear regression model yielded an MSE of 0.2228, which indicates that there is a relationship between HSS and GPA. However, this relationship is not very strong, for there is an apparent gap between predictions based on HSS and the GPA. Additionally, the 0.2228 suggests that a major feature of the change in university GPA is not explained well by HSS alone.

The linear regression model also yielded an R^2 value of 0.1774, which offers a significant insight into the relationship between HSS and GPA. This value means that approximately 17.74% of the variance in GPA can be explained by HSS, which indicates a weak to moderate relationship between them. While there is a positive correlation between HSS and GPA, 82.26% of university GPA is not explained by HSS.

4.5 Neural Network Model

With an MSE of 0.2383 and an R^2 of 0.1204, the neural network model largely corroborates the findings from the linear regression model, supporting our understanding of the relationship between HSS and GPA. The neural network model shows a similar, although slightly weaker, predictive performance compared to the linear regression model (MSE: 0.2228, R^2: 0.1774).

Employing a complex neural network model in this study provides strong support for the findings from both the traditional statistical methods and the linear regression model. Additionally, the similar results from both linear regression and neural network models provide stronger evidence for the nature of the HSS-GPA relationship.

5. Conclusion

This study successfully addresses the two research questions regarding the relationship between HSS and academic performance in the ELT Department, as well as the reliability of the HSS as a predictor of university GPA. By combining traditional statistical methods with advanced AI-based algorithms, such as linear regression and neural network models, this study shows a weak to moderate correlation between HSS and GPA. Although higher HSS corresponds generally to a higher GPA, especially for students in the top HSS range, this trend is not consistent across all ranges. The results of the analysis suggest that while HSS is a variable in the prediction of university performance, it is not the sole factor. Other factors are likely to play a significant role, particularly for students with lower HSS, as is seen in the significant overlap in GPA across the different categories of HSS. Consequently, although HSS demonstrates a degree of predictive capability, it fails to explain a significant portion of the variance observed in university GPA, stressing the intricate nature of academic success within the ELT Department.

The study further highlights a significant observation for policy and the decision-making: while there is some correlation between high HSS and high GPA, the number of students whose HSS is very high is rather low. This is particularly relevant to university admissions and scholarship strategies that target students with high HSS-as illustrated, for instance, by offering tuition discounts to applicants with HSS of 80 and above to attract highly qualified students. The analysis indicates that such policies would affect only a limited number of students, which does not significantly impact the overall pool of students. Hence, reliance on HSS as a primary factor for such decisions may not be sufficiently indicative of the potential for academic success among a broader range of students. Therefore, alternative or additional admission criteria must be taken into consideration.

Although HSS can still serve as an initial indicator to identify students with academic potential, the study findings indicate that university scholarship policies can be revised to include lower thresholds than 80 for HSS to attract a broader pool of students who might perform strongly in the ELT Department. The data shows that students with HSS lower than 80 are still likely to achieve high GPA, indicating that a rigid HSS cut-off may exclude students who have the potential for success but do not meet the higher threshold. Therefore, it would be advisable for scholarship and other incentive programs to not only consider initial HSS but also incorporate academic progression within the college programs. This could be done by offering incentives or scholarships to the students who achieve top rankings during their undergraduate study. This approach would ensure that scholarships are aligned with actual performance and development within the ELT program and provide a more holistic and dynamic evaluation of student success.

This study does not expect that it will present a definitive solution to the debate over standardized testing as a reliable predictive measure to college performance. However, the findings in this study add weight to the literature that complicates the topic, deeming the need for further and continued research in the field.

References

Allensworth, E. M., & Clark, K. (2020). High school GPAs and ACT scores as predictors of college completion: Examining assumptions about consistency across high schools. *Educational Researcher, 49*(3), 198-211. https://doi.org/10.3102/0013189X20902110

Casas, F. R., & Meaghan, D. E. (2001). Renewing the debate over the use of standardized testing in the evaluation of learning and teaching. *Interchange, 32*, 147–181. https://doi.org/10.1023/A:1011946403091

Chollet, F. (2015). *Keras: The Python deep learning library*. https://keras.io

De Wilde, P. (2013). *Neural network models: Theory and projects*. Springer Science & Business Media.

Geiser, S., & Santelices, M. V. (2007). *Validity of high-school grades in predicting student success beyond the freshman year: High-school record vs. standardized tests as indicators of four-year college outcomes*. University of California, Berkeley, Center for Studies in Higher Education. https://escholarship.org/content/qt7306z0zf/qt7306z0zf.pdf

Geiser, S. (2009). Back to the basics: In defense of achievement (and achievement tests) in college admissions. *Change: The Magazine of Higher Learning, 41*(1), 16–23. https://doi.org/10.3200/CHNG.41.1.16-23

James, G., Witten, D., Hastie, T., Tibshirani, R., & Taylor, J. (2023). Linear regression. In *An introduction to statistical learning: With applications in Python* (Springer Texts in Statistics, pp. 69-134). Springer, Cham. https://doi.org/10.1007/978-3-031-38747-0_3

Kingma, D. P., & Ba, J. (2014). Adam: A method for stochastic optimization. *arXiv* preprint *arXiv:1412.6980*.

Lee, H., Kizilcec, R. F., & Joachims, T. (2023). Evaluating a learned admission-prediction model as a replacement for standardized tests in college admissions. In *Proceedings of the Tenth ACM Conference on Learning@ Scale, Copenhagen, Denmark* (pp. 195–203). https://doi.org/10.1145/3573051.3593382

Ministry of Education-KRG (MoE-KRG). (2009). Secondary education school system. https://gov.krd/moe-en/system-and-education/basic-and-secondary-systems/

Mohammed, P. J., Casas, J., Saleh, S. T., & Kovács, K. E. (2021). An international comparison of educational systems: The Columbian, Iraqi, and Kurdish cases. *Central European Journal of Educational Research, 3*(2), 110-120. https://doi.org/10.37441/cejer/2021/3/2/9803

Portelli, J. P., & Vilbert, A. B. (2002). Standards, equity, and the curriculum of life. *Analytic Teaching, 22*(1).

Powers, D. E., & Kaufman, J. C. (2002). Do standardized multiple-choice tests penalize deep-thinking or creative students? *ETS Research Report Series, 2002*(2), i-11. https://doi.org/10.1002/j.2333-8504.2002.tb01882.x

Roediger, H. L. III, & Marsh, E. J. (2005). The positive and negative consequences of multiple-choice testing. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 31*(5), 1155–1159. https://doi.org/10.1037/0278-7393.31.5.1155

Seyfarth, J. T., Simon, D. J., & Schlesinger, J. (1994, February). *Assessing student performance: Are our assumptions valid?* Paper presented at the Annual Meeting of the American Association of Colleges for Teacher Education, Chicago, IL, United States. https://files.eric.ed.gov/fulltext/ED367635.pdf

Vernez, G., Culbertson, S., Constant, L., & Karam, R. (2016). *Initiatives to improve quality of education in the Kurdistan Region-Iraq*. RAND Education. https://www.rand.org/content/dam/rand/pubs/research_reports/RR900/RR960/RAND_RR960.pdf

Wahab, A. (2019). The futility of market-based reform in education: The case of the Iraqi Kurdistan Region. In S. Chitpin & J. P. Portelli (Eds.), *Confronting Educational Policy in Neoliberal Times: International Perspectives* (1st ed., pp. 132-144). Routledge. https://doi.org/10.4324/9781315149875

Wightman, L. F. (2000). The role of standardized admission tests in the debate about merit, academic standards, and affirmative action. *Psychology, Public Policy, and Law, 6*(1), 90–100. https://doi.org/10.1037/1076-8971.6.1.90

Zwick, R. (2012). The role of admissions test scores, socioeconomic status, and high school grades in predicting college achievement. *Pensamiento Educativo, 49*(2). https://onomazein.letras.uc.cl/index.php/pel/article/view/25803

Zwick, R. (2013). Disentangling the role of high school grades, SAT® scores, and SES in predicting college achievement. *ETS Research Report Series, 2013*(1), 1-20. https://doi.org/10.1002/j.2333-8504.2013.tb02316.x