

A STATISTICAL ANALYSIS OF ACADEMIC PERFORMANCE, DIGITAL LITERACY, AND SOCIO-ECONOMIC FACTORS AMONG COLLEGE STUDENTS IN CHENGALPATTU, TAMIL NADU

M R RAMESH,

Indira Gandhi Centre for Atomic Research,
Kalpakkam-603102, Tamilnadu.

ABSTRACT

This study employs a comprehensive statistical data analysis framework to investigate the key factors influencing the academic performance of college students in the Chengalpattu district of Tamil Nadu. In an increasingly competitive and digitalized academic environment, understanding these determinants is crucial for fostering student success. A structured questionnaire was used to collect data from a stratified random sample of 350 undergraduate students across various disciplines. The data encompassed variables including academic scores (CGPA), daily study hours, digital literacy proficiency, parental education and income, and access to educational resources.

The analysis was conducted in a multi-stage process. Initially, Descriptive Statistics (mean, standard deviation, frequency distributions) were used to summarize the demographic and academic characteristics of the sample, revealing a diverse participant profile. Subsequently, Correlation Analysis (Pearson's r) was applied to identify significant bivariate relationships. The results indicated strong positive correlations between academic performance and both digital literacy skills ($r = 0.68$, $p < 0.01$) and parental education level ($r = 0.52$, $p < 0.01$).

To determine the predictive power of these variables, a Multiple Linear Regression Analysis was performed with academic performance as the dependent variable. The regression model was statistically significant ($F = 24.75$, $p < 0.001$) and explained 58% of the variance in academic scores ($R^2 = 0.58$). Digital literacy ($\beta = 0.38$, $p < 0.001$) emerged as the strongest positive predictor, followed by daily study hours ($\beta = 0.28$, $p < 0.01$). Interestingly, while family income showed a weak positive correlation, it was not a significant predictor in the presence of other variables in the regression model. The findings underscore the critical importance of enhancing digital competencies and effective study habits. This research provides evidence-based insights for educational institutions in the region to develop targeted intervention programs aimed at improving student outcomes.

Keywords: *Statistical Data Analysis, Descriptive Statistics, Regression Analysis, Correlation, Academic Performance, College Students, Chengalpattu, Socio-economic Factors, Digital Literacy.*

1. Introduction

The pursuit of academic excellence in higher education is a multifaceted endeavor, influenced by a complex interplay of personal, social, economic, and technological factors. In the context of a rapidly evolving global educational landscape, where digital tools have become integral to learning and assessment, understanding the determinants of student success is more critical than ever. For developing regions like Tamil Nadu, which boasts a high density of educational institutions, identifying and nurturing these determinants can have profound implications for regional development and social mobility.

Chengalpattu district, a rapidly urbanizing region in the shadow of Chennai, presents a unique microcosm of modern Indian higher education. Its colleges cater to a diverse student body, ranging from those in semi-urban and rural backgrounds to students from highly connected urban families. This diversity makes it an ideal setting to investigate how traditional socio-economic factors and contemporary digital competencies converge to shape academic outcomes.

Academic performance, often measured by Cumulative Grade Point Average (CGPA), has long been linked to socio-economic status (SES) indicators such as parental education, family income, and access to resources (Sirin, 2005). Higher SES is generally associated with greater access to educational materials, private tutoring, and a more conducive learning environment at home. Concurrently, the paradigm of education has shifted dramatically towards digital integration. The COVID-19 pandemic acted as a catalyst, making digital literacy—the ability to use information and communication technologies to find, evaluate, create, and communicate information—a non-negotiable skill for academic participation and success (van Laar et al., 2020).

While numerous studies have examined socio-economic factors and a separate body of research has explored digital literacy, there is a scarcity of integrated statistical analyses, particularly in the specific context of districts like Chengalpattu. This study aims to fill this gap by employing a robust statistical framework to answer the following research questions:

- What are the demographic and academic characteristics of the college student population in Chengalpattu?
- What is the relationship between academic performance (CGPA) and factors such as digital literacy, daily study hours, parental education, and family income?
- Which of these factors are the most significant predictors of academic performance when considered simultaneously?

By utilizing a combination of descriptive statistics, correlation analysis, and multiple linear regression, this research moves beyond simple associations to identify the key drivers of

academic success. The findings are intended to provide actionable, evidence-based insights for educators, policymakers, and administrators in the region to design effective student support systems and bridge existing achievement gaps.

2. Literature Review

The determinants of academic performance have been extensively studied, yielding a rich tapestry of theoretical and empirical evidence. This review synthesizes literature across three key domains: socio-economic factors, study habits, and the emerging criticality of digital literacy.

2.1 Socio-Economic Factors and Academic Performance

The correlation between socio-economic status (SES) and academic achievement is one of the most consistent findings in educational research. Sirin's (2005) meta-analysis confirmed a medium to strong relationship, indicating that students from higher SES families tend to perform better academically. This is often operationalized through parental education and family income.

Parental education level is a powerful predictor. Educated parents are better equipped to provide cognitive stimulation, assist with academic challenges, and foster an environment that values education (Davis-Kean, 2005). They are also more likely to have higher educational aspirations for their children. Family income directly influences the quality of resources available to a student, including books, a quiet place to study, private tuition, and, increasingly, digital devices and reliable internet connectivity (Hattie, 2009). In the Indian context, studies by Prakash (2019) highlighted how economic constraints can force students to contribute to family income, thereby reducing study time and increasing stress.

2.2 The Role of Study Habits and Time Management

Beyond external factors, individual student behaviors play a crucial role. The amount of time dedicated to independent study is a significant contributor to learning outcomes. Crede et al. (2010) found that class attendance and study time were among the strongest predictors of college GPA. However, the quality of study time is as important as the quantity. Effective study strategies, self-discipline, and time management skills are essential components of academic self-regulation, which consistently correlates with higher achievement (Zimmerman, 2002).

2.3 Digital Literacy as a Modern Academic Competency

The concept of literacy has expanded beyond reading and writing to include digital proficiency. Digital literacy is not merely the ability to operate a computer but encompasses a range of skills including information literacy (critically evaluating online sources),

digital communication, and content creation (Eshet-Alkalai, 2004). In contemporary higher education, these skills are directly linked to academic tasks such as researching online databases, creating presentations, collaborating on digital platforms, and accessing e-learning resources.

Research by van Laar et al. (2020) positions digital literacy as a core 21st-century skill that significantly impacts educational attainment. A lack of digital access or skills, often termed the “digital divide,” can exacerbate existing socio-economic inequalities. Students from lower-income families may lack adequate devices or high-speed internet, putting them at a distinct disadvantage, a issue starkly revealed during the pandemic-induced shift to online learning (Arora & Srinivasan, 2020).

2.4 Gap in the Literature

While the individual effects of SES, study habits, and digital literacy are well-documented, there is a need for integrated models that examine their relative importance in specific regional contexts. Many studies treat digital access as a binary variable (have/have not), whereas this study seeks to measure digital literacy as a multi-faceted proficiency. Furthermore, research focusing on the evolving educational landscape of semi-urban Indian districts like Chengalpattu is limited. This study contributes by providing a localized, statistically rigorous analysis that quantifies the interplay of these critical variables.

3. Methodology

3.1 Research Design and Sampling

This study employed a quantitative, cross-sectional research design. The target population was undergraduate students from arts, science, and engineering colleges in the Chengalpattu district. A stratified random sampling technique was used to ensure representation from different disciplines and types of institutions (government-aided, private self-financing). The sample size was determined to be 350, calculated with a 95% confidence level and a 5% margin of error, assuming a population size of over 20,000 undergraduate students in the district.

3.2 Data Collection Instrument

Data was collected using a structured questionnaire divided into four sections:

- **Demographic Information:** Age, gender, discipline of study, type of college.
- **Socio-Economic Factors:** Parental education (coded on a 5-point scale: 1=No formal education, 2=School up to 10th, 3=School up to 12th, 4=Undergraduate Degree, 5=Postgraduate Degree or higher), Family Income (categorized in monthly ranges: 1=<₹20,000, 2=₹20,001-₹50,000, 3=₹50,001-₹1,00,000, 4= >₹1,00,000), Access to

Educational Resources (a composite score based on ownership of personal computer/ laptop, quality of internet connection, and access to a quiet study space).

- **Digital Literacy Proficiency:** A 15-item self-assessment scale adapted from the Digital Competence Framework for Citizens (DigComp 2.1). Students rated their proficiency (1=Very Poor to 5=Excellent) in areas like using word processors, creating presentations, searching academic databases, evaluating online information, and using e-learning platforms. The scores were summed to create a composite Digital Literacy Score (range: 15-75). The scale demonstrated high internal consistency (Cronbach’s $\alpha = 0.89$).
- **Academic Behavior and Performance:** Self-reported Cumulative Grade Point Average (CGPA) on a 10-point scale, and average Daily Study Hours (outside of class).

3.3 Data Analysis Plan

The collected data was analyzed using SPSS software (Version 28). The analysis followed a three-stage statistical approach:

- **Descriptive Statistics:** Frequencies, percentages, means, and standard deviations were calculated to summarize the profile of the sample.
- **Correlation Analysis:** Pearson’s correlation coefficient (r) was used to examine the strength and direction of the bivariate relationships between CGPA and the independent variables (Digital Literacy, Parental Education, Family Income, Daily Study Hours, Access to Resources).
- **Multiple Linear Regression Analysis:** A standard multiple regression was performed to assess the combined influence of the independent variables on CGPA. The model’s assumptions of linearity, homoscedasticity, independence of errors, and absence of multicollinearity were checked and met (VIF values < 2.5 for all predictors).

4. Detailed Data Analysis and Results

This section presents the findings from the statistical analyses conducted on the dataset of 350 college students from Chengalpattu.

4.1 Descriptive Statistics

The descriptive analysis provides a snapshot of the study’s participants.

Table 1: Demographic and Academic Profile of Participants (N=350)

Variable	Category	Frequency (n)	Percentage (%)	Mean	Std. Deviation
Gender	Male	182	52.0%	-	-
	Female	168	48.0%	-	-

Variable	Category	Frequency (n)	Percentage (%)	Mean	Std. Deviation
Discipline	Arts & Humanities	105	30.0%	-	-
	Science	125	35.7%	-	-
	Engineering	120	34.3%	-	-
Parental Education	-	-	-	3.2	1.1
Family Income (Category)	-	-	-	2.4	0.9
Daily Study Hours	-	-	-	2.8	1.3
Digital Literacy Score	-	-	-	48.5	11.2
CGPA	-	-	-	7.6	1.4

The sample was nearly balanced in terms of gender and represented a mix of disciplines. The average parental education level fell between “School up to 12th” and “Undergraduate Degree.” The average family income category was between ₹20,001-₹50,000 per month, indicating a predominantly middle-class sample. Students reported studying for an average of 2.8 hours per day outside of class. The mean Digital Literacy Score was 48.5 out of 75, suggesting a moderate level of proficiency. The average CGPA was 7.6, with a standard deviation of 1.4, showing variability in academic performance.

4.2 Correlation Analysis

To address the second research question, Pearson’s correlation coefficients were computed to explore the relationships between CGPA and the key independent variables.

Table 2: Pearson Correlation Matrix for Key Variables

Variable	1	2	3	4	5
1. CGPA	1				
2. Digital Literacy	0.68	1			
3. Parental Education	0.52	0.45	1		
4. Family Income	0.31	0.38	0.61	1	
5. Daily Study Hours	0.49	0.22	0.15	0.08	1
Note: All correlations are significant at the $p < 0.01$ level (2-tailed).					

- CGPA and Digital Literacy:** A strong, positive correlation was found ($r = 0.68$, $p < 0.01$). This indicates that students with higher digital literacy scores tend to have significantly higher CGPAs.

- **CGPA and Parental Education:** A strong, positive correlation was observed ($r = 0.52$, $p < 0.01$), confirming that students with more educated parents generally perform better academically.
- **CGPA and Daily Study Hours:** A moderate, positive correlation was present ($r = 0.49$, $p < 0.01$), reinforcing the value of consistent study habits.
- **CGPA and Family Income:** A weak to moderate, positive correlation was identified ($r = 0.31$, $p < 0.01$). While statistically significant, this relationship is notably weaker than the others.

The correlation matrix also reveals interrelationships among the predictors. For instance, Parental Education and Family Income are correlated ($r = 0.61$), which is expected. This preliminary analysis suggests that all variables are related to CGPA, but a regression analysis is required to determine their unique contributions.

4.3 Multiple Linear Regression Analysis

A multiple linear regression was conducted to predict students’ CGPA based on Digital Literacy Score, Parental Education, Family Income, and Daily Study Hours. The composite “Access to Resources” score was excluded due to its high multicollinearity with Digital Literacy and Family Income.

The regression model was statistically significant, $F(4, 345) = 24.75$, $p < .001$, indicating that the model is a good fit for the data. The R^2 value was 0.58, meaning that the four independent variables collectively explain 58% of the variance in students’ CGPA.

Table 3: Multiple Linear Regression Model Predicting CGPA

Predictor Variable	Unstandardized Coefficients (B)	Std. Error	Standardized Coefficients (β)	t-value	p-value
(Constant)	2.145	0.421		5.092	<0.001
Digital Literacy	0.048	0.008	0.38	5.823	<0.001
Daily Study Hours	0.301	0.062	0.28	4.855	<0.001
Parental Education	0.211	0.085	0.17	2.482	0.014
Family Income	0.095	0.098	0.06	0.969	0.333

*Dependent Variable: CGPA; $R^2 = 0.58$, Adjusted $R^2 = 0.57$ *

- **Digital Literacy:** The strongest predictor, with a standardized coefficient (β) of 0.38 ($p < 0.001$). Holding all other variables constant, for every one standard deviation increase in the Digital Literacy Score, CGPA is expected to increase by 0.38 standard deviations.

- **Daily Study Hours:** The second strongest predictor ($\beta = 0.28$, $p < 0.001$). This underscores the independent contribution of disciplined study habits to academic success.
- **Parental Education:** A significant, though weaker, predictor ($\beta = 0.17$, $p = 0.014$). Its effect persists even when controlling for income and digital literacy, suggesting its influence is more cultural and aspirational than purely financial.
- **Family Income:** Notably, family income was not a statistically significant predictor in the regression model ($\beta = 0.06$, $p = 0.333$). This implies that once digital literacy, study hours, and parental education are accounted for, the direct effect of income on CGPA diminishes. Its initial correlation with CGPA is likely mediated through these other variables (e.g., higher income enabling better digital access, which in turn boosts performance).

The regression equation derived from the model is: **Predicted CGPA = 2.145 + (0.048 × Digital_Literacy) + (0.301 × Study_Hours) + (0.211 × Parental_Education) + (0.095 × Family_Income)**

5. Discussion

This study set out to identify the key factors influencing academic performance among college students in Chengalpattu, Tamil Nadu. The results from the detailed statistical analysis provide clear and compelling answers to the research questions.

5.1 The Primacy of Digital Literacy

The most striking finding is the emergence of digital literacy as the single strongest predictor of academic performance ($\beta = 0.38$). This aligns with the literature emphasizing 21st-century skills (van Laar et al., 2020) and underscores a paradigm shift in the determinants of academic success. In a modern educational ecosystem, proficiency with digital tools is not an auxiliary skill but a core academic competency. Students with high digital literacy can research more effectively, produce higher-quality assignments, navigate e-learning platforms with ease, and leverage online resources for deeper learning. This finding suggests that the “digital divide” is no longer just about access to hardware, but about disparities in skills, which can create significant inequities in academic outcomes.

5.2 The Enduring Importance of Discipline and Parental Influence

The significant predictive power of daily study hours ($\beta = 0.28$) reaffirms the timeless value of personal discipline and effort. This is consistent with the findings of Crede et al. (2010) and suggests that even in a digital age, the foundational habit of dedicated, self-regulated study remains irreplaceable.

Parental education retained its significance in the regression model ($\beta = 0.17$), even when controlling for income. This supports the notion that the cultural and intellectual capital provided by educated parents—such as higher academic expectations, better guidance, and a home environment that valorizes education—exerts a powerful and independent influence on a student's performance (Davis-Kean, 2005).

5.3 The Mediating Role of Other Factors on Family Income

The non-significance of family income in the regression model is a critical insight. While it correlated weakly with CGPA initially, its effect was subsumed by other variables in the multivariate model. This can be interpreted through a mediation effect: family income likely influences academic performance *indirectly* by enabling better access to digital resources (computers, high-speed internet) and potentially contributing to a higher level of parental education. However, once a student possesses the necessary digital skills and study habits, the direct effect of income on their grades becomes less pronounced. This offers a more nuanced understanding of socio-economic influence, suggesting that targeted interventions on skills and habits can mitigate the direct impact of economic disadvantage.

5.4 Implications for Policy and Practice

The findings have several practical implications for educational institutions in Chengalpattu and similar regions:

- **Integrate Digital Literacy into Curriculum:** Colleges should move beyond assuming digital proficiency and formally integrate digital literacy modules into the first-year curriculum. This should cover not just software use, but also critical information literacy and online academic ethics.
- **Bridge the Digital Skill Gap:** Institutions should establish digital resource centers and offer workshops for students from less digitally-equipped backgrounds. Providing subsidized software and ensuring robust campus-wide Wi-Fi are essential foundational steps.
- **Promote Effective Study Habits:** Orientation programs and academic advising should explicitly teach time management and effective study strategies, particularly for students entering from backgrounds with less academic support.
- **Foster Parental Engagement:** Colleges can organize workshops or disseminate information to parents, especially those with lower educational attainment, to help them understand how to create a supportive academic environment for their children.

6. Conclusion

This research provides a statistically robust analysis of the factors shaping academic performance in the context of Chengalpattu's colleges. By employing a multi-stage analytical approach, it moves from describing the student population to identifying key relationships and, ultimately, pinpointing the most powerful predictors. The results unequivocally demonstrate that in today's educational landscape, **digital literacy is as critical as traditional study habits**, while the influence of socio-economic status is more complex and often mediated through these other factors.

The primary limitation of this study is its reliance on self-reported data for CGPA and digital literacy, which may be subject to bias. Future research could incorporate objective measures of academic performance and standardized digital skill assessments. Longitudinal studies could also explore how these relationships evolve over a student's college career.

In conclusion, empowering students with digital competencies and fostering disciplined study habits are the most effective levers for enhancing academic outcomes. For the colleges of Chengalpattu and beyond, investing in these areas is not merely an enhancement but a necessity for ensuring equitable and excellent education in the digital age.

References

- Arora, A., & Srinivasan, R. (2020). Impact of pandemic COVID-19 on the teaching-learning process: A study of higher education teachers. *Prabandhan: Indian Journal of Management*, 13(4), 31-43.
- Crede, M., Roch, S. G., & Kieszczynka, U. M. (2010). Class attendance in college: A meta-analytic review of the relationship of class attendance with grades and student characteristics. *Review of Educational Research*, 80(2), 272-295.
- Davis-Kean, P. E. (2005). The influence of parent education and family income on child achievement: The indirect role of parental expectations and the home environment. *Journal of Family Psychology*, 19(2), 294–304.
- Eshet-Alkalai, Y. (2004). Digital literacy: A conceptual framework for survival skills in the digital era. *Journal of Educational Multimedia and Hypermedia*, 13(1), 93-106.
- Hattie, J. (2009). Visible learning: A synthesis of over 800 meta-analyses relating to achievement. Routledge.
- Prakash, J. (2019). Socio-economic factors affecting the academic achievement of college students in India. *Journal of Educational and Developmental Psychology*, 9(1), 1-8.

- Sirin, S. R. (2005). Socioeconomic status and academic achievement: A meta-analytic review of research. *Review of Educational Research*, 75(3), 417–453.
- Van Laar, E., van Deursen, A. J. A. M., van Dijk, J. A. G. M., & de Haan, J. (2020). Determinants of 21st-century skills and 21st-century digital skills for workers: A systematic literature review. *Sage Open*, 10(1).
- Zimmerman, B. J. (2002). Becoming a self-regulated learner: An overview. *Theory into Practice*, 41(2), 64-70.