Modelling academic performance of students in Bolgatanga East District using Panel Regression

Samson Akanlewine Ayamga[1] , Issah Fatawu[1], Jakperik Dioggban[2]*

[1]Department of Statistics and Actuarial Science, School of Mathematical Sciences, C. K. Tedam University of Technology and Applied Sciences, Box 24, Navrongo.

[2]*Department of Biometry, School of Mathematical Sciences, C. K. Tedam University of Technology and Applied Sciences, Box 24, Navrongo.

*Corresponding Author: jdioggban@cktutas.edu.gh

DOI:https://doi.org/10.30210/JMSO.202402.013

Submitted: Oct. 02, 2024 Accepted: Nov. 25, 2024

ABSTRACT

The findings of a study that modeled the academic performance of the pupils in the Bolgatanga East District using panel regression analysis are presented in this abstract. The study concentrated on how student performance was impacted by attendance, residence, age, class size, parental occupation, and sex. To identify patterns and trends in the academic performance of pupils, secondary data was gathered from 15 junior high schools in the Bolgatanga East District. The various elements that affect pupils' academic success were identified using panel regression and nested design models. The study's findings demonstrated the significance of attendance, age, and gender as predictors of academic achievement. Furthermore, it was discovered that student age played a considerable role, with younger pupils outperforming older ones in terms of marks. The performance of male and female students differed slightly, with girls outperforming males in some courses. Exam scores were higher for students who frequently attended classes, indicating that attendance had a beneficial effect on performance. Overall, the study highlights the significance of attendance, age, and sex in predicting students' outcomes and offers insights into the elements impacting academic achievement in the Bolgatanga East District. Interventions and policies targeted at raising district students' achievement can be informed by the data. It is crucial to educate stakeholders on the importance of these elements in relation to district students' academic success.

Keywords: Performance, Attendance, Residence, Age, and Sex

1. Introduction

Students' performance on an exam is referred to as their academic performance. Student achievement is influenced by a number of things. [4]. Final exam results and average marks from ongoing evaluations are used to evaluate it. Academic performance serves as a gauge of the efficacy of the educational system by assessing students' comprehension of the material and abilities, their ability

to complete assignments, and their overall accomplishments. Academic achievement also has a big impact on people's well-being. [15]. It has an impact on both the academic achievement of present pupils and their prospects for future employment. [25]. Poor academic performance is associated with higher dropout rates, higher rates of substance misuse, and detrimental effects on social, mental, and physical health. [15].

The ultimate purpose of educational research is frequently thought to be determining what factors influence school success. [23]. For many years, scholars and decision-makers have been examining the primary factors that contribute to successful educational outcomes. How well a school achieves its objectives in terms of student learning, development, and general well-being is what is known as school effectiveness [14]. Similarly, an investigation by Ismail et al. [9] highlighted how important a positive school climate with strong relationships between teachers and students is to raising student achievement.

For both teachers and students, increasing class sizes is without a doubt the best way to handle large enrollments. The disadvantages of double-shifting are so severe that they can affect classes of up to 100 students. Research has indicated that teacher effectiveness is linked to the presence of established classroom routines and procedures, as well as the establishment of robust classroom systems, which can result in better student results [18]. At the same time, greater teacher effectiveness enables teachers to provide more effective pedagogical support, which benefits students' learning outcomes.

To achieve universal access to primary education by 2015, Ghana has made tremendous progress toward achieving the two Millennium Development Goals. In 2006/2007, 90.80% of students were enrolled in elementary school, according to the Institute of Statistical, Social, and Economic Research [8]. In Ghanaian elementary schools, the government has also implemented a number of measures to help with student registration, daily attendance, and retention. One example of such innovative initiatives is the Ghanaian government's provision of educational resources, including school uniforms, school meals, and instructional materials.

The results of the Basic Education and Certificate Examination (BECE) from 2017–18 to 2020–21 indicate that gender In the past, guys did better in mathematics than girls, although this difference has narrowed over time. Exam results in Ghanaian and English are comparable for both boys and girls. In mathematics, boys used to do better than girls, but by 2020–21, there were no discernible differences between the two groups' exam results. In 2017–18, for instance, boys scored an average of 2 percentage points better than girls (46 vs 44); nevertheless, by 2020–21, the difference was negligible (49.4 versus 49.2). Boys and girls have continuously performed similarly over the years in Ghanaian and English. Ghana's commitment to attaining gender equality in educational attainment is demonstrated by these favorable outcomes.

The educational sector still faces numerous obstacles in spite of all of its achievements. Many students are failing to meet the minimal requirements for learning as well as to acquire the necessary abilities and fundamental skills, according to ISSER [8]. According to results from the 2006 Basic Education Certificate Examination (BECE), 190,921 pupils out of 208,379 who took the test were able

to receive aggregate scores between 6 and 30, or 62.00%. The 2008 BECE results showed that more than 282,202 pupils received a pass grade, or 62.20%. Between 2019 and 2021, Bolgatanga East District's BECE performance decreased from 46.60% in 2019 to 33.80% in 2020, then marginally improved to 38.10% in 2021.

A student's age, educational institution, involvement in class, socioeconomic status of parents or guardians, residential area (rural or urban), educational trends, and the medium of instruction used in educational institutions are just a few of the many factors that determine a learner's success. [19].

Learning outcomes are thought to be essential to the effectiveness of the overall educational system. It is believed that the success of students dictates whether a learning institution succeeds or fails [15]. Education and academic effect on the nation's socioeconomic activities are the foundation of any nation's development [21]. This is based on the academic success of the students.

According to Farooq et al. [5], teachers' first concern should be pupils' success. Even exam results and continuous student assessments are used to gauge learning objectives, they claim. Others, such as learning about factors that influence students' success, have attracted a lot of attention from scholars and educators [2, 5, 21]. These studies demonstrate that enhancing students' academic performance is influenced by a variety of factors.

Any country's ability to succeed and develop is greatly influenced by the quality of education it offers. All individuals should have access to high-quality, inclusive education, according to Sustainable Development Goal 4. The significance of the foundational level of education is linked to the growth and advancement of contemporary civilization. At the primary school level, pupils work to improve their reading, writing, and numeracy abilities.

The Junior High School, formerly the Junior Secondary School, was founded to provide students with foundational technical, career, and scientific knowledge. The Senior High School facility serves to reinforce knowledge and skills learned during basic education and educational attainment, while also providing opportunities for students to develop a variety of skills and abilities. This study therefore attempts to predict factors that influence junior high school students' performance in the Bolgatanga East district.

2. Methodology

2.1 Source of Data

Secondary data from the Bolgatanga East District Education Directorate's student academic records was used. The schools provided the students' examination results for the three years (2020–2022) based on terms.

This study's target population consists of all third-year students from public junior high schools in the Bolgatanga East District. Purposive sampling was used to evaluate fifteen (15) public junior high schools. Three main subjects were taken into consideration in the study: integrated science, mathematics,

and English language. Each academic year's termly student scores for each of the chosen schools were taken into account.

2.2 Panel Regression

Often called longitudinal or cross-sectional time series data, panel data represents information gathered from multiple entities over a range of periods. One or more variables are seen throughout several periods in time series data sets, while numerous variables are observed at a single moment in time in cross-sectional data.

Unobserved heterogeneity does not cause bias when using panel regression. Data from multiple individuals is incorporated into a pooled model without taking into consideration possible individual variances that could produce different coefficients.

The model for Panel Regression (PR) is expressed by;

$$y_{it} = \beta_0 + \beta_1 X_{it,2} + \ldots + \beta_k X_{it} + \varepsilon_{it}, \quad (1)$$

$$(i = 1, \ldots, N;\ t = 1, \ldots, T, an\ k = 1, \ldots, K).$$

where;

i is the unit of observation,

t is the period,

k indicates the k^{th} explanatory variable,

β_0 is the intercept,

β_k is the coefficient of each explanatory variable,

ε_{it} is the error term

[24].

The error term, ε_{it}, can be broken down into two components: a unit-specific error related to the cross-section, a_i, and an individual-specific error u_{it}.. The error term, ε_{it} can be decomposed into two sections such as a cross-sectional unit-specific error, a_i and an idiosyncratic error u_{it}.

$$\varepsilon_{it} = a_i + u_{it}. \quad (2)$$

The cross-sectional unit-specific error, denoted as a_i, remains constant over time, while the idiosyncratic error, u_{it} changes across different cross-sectional units and over time. The rationale and advantages of breaking down the error terms into components are aimed at reducing certain portions by utilizing panel data, which is beneficial in lessening concerns related to omitted variable bias that arises from unobserved unit-specific factors. By integrating equation (2) into equation (1), the following equation can be derived.

$$y_{it} = \beta_0 + \beta_1 X_{it,2} + \ldots + \beta_k X_{it} + \varepsilon_{it} + a_i + u_{it}. \quad (3)$$

[24]

2.2.1 Assumptions of panel regression

1. Linearity: There must be a linear relationship between the dependent and independent variables.

2. Homoskedasticity and Non-autocorrelation: The variance of the residuals must be constant across the predicted values.
3. Independent variables are not stochastic.
4. No multicollinearity: No high correlation between the independent variables.

2.3 Nested Design

Nested design refers to a research design where one-factor level (such as Factor B) is hierarchically subsumed under another level of a factor (such as Factor A). The statistical model for the nested design is

$$y_{ijk} = \mu + \alpha_j + \beta_{j(i)} + \mathcal{E}_{(ij)k,} \quad (4)$$

$(i = 1,\ldots,n;\ j = 1,\ldots,p, and\ k = 1,\ldots,q).$

Where;

y_{ijk} presents the score of the i^{th} observation within the j^{th} level of Factor A and the k^{th} level of Factor B,

μ indicates the grand mean, serving as a constant for the population of observations,

α_j signifies the effect of the j^{th} treatment condition of Factor A; mathematically, it represents the deviation of the population means (μ_j) from the grand mean (μ). All observations share the same dependent score in the j^{th} condition, subject to the requirement that the sum of all α_j equals zero across all treatment conditions.

$\beta_{k(j)}$ denotes the effect of the k^{th} treatment condition of Factor B, which is nested within the j^{th} level of Factor A; mathematically, it reflects the deviation of the population mean (μ_{jk}) in the combined k^{th} and j^{th} level from the grand mean (μ). This effect remains constant for all dependent scores of observations in the k^{th} condition, nested within the j^{th} condition of Factor A. It is assumed that the effect follows a normal distribution in its underlying population.

$\mathcal{E}_{i(jk)}$ represents the random error effect associated with the i^{th} observation in the j^{th} condition of Factor A and the k^{th} condition of Factor B. This term is a random variable, distributed normally in the underlying population and is independent of $\beta_{k(j)}$.

[13]

2.4. Assumptions of Nested Design

1. Linearity: There must be a linear association between the dependent and independent variables.
2. Normality: The error residuals are assumed to be distributed normally.
3. Homoscedasticity: The residuals are assumed to have a constant variance.
4. Independence of observations (No autocorrelation of the model's residuals).

3. Results

The means for attendance, age, and class size of the students were 57.39, 18.69, and 33.69, respectively as contained in Table 1 below. The standard deviations for the same variables were obtained as 16.52, 1.58, and 11.29, respectively. The minimum and maximum values for attendance recorded are 14.50 and 99.50, age is 14.00 and 26.00, and class size is 18.00 and 58.00, respectively.

Table 1. Descriptive Statistics

Variables	Mean	Std. Deviation	Minimum	Maximum
Attendance	57.39	16.52	14.50	99.50
Age	18.69	1.58	14.00	26.00
Class-size	33.69	11.29	18.00	58.00

3.1 Correlation Coefficients

The Pearson product-moment correlation coefficient (r) assesses how strongly two quantitative variables are linearly associated with a sample. Each individual or case needs to have measurements for both quantitative variables. The significance test for r determines if there is a linear connection between the two variables in the overall population. Pearson Correlation analysis was conducted to evaluate the relationship between the variables. The findings are shown in Table 2.

Table 2. Pearson Correlation Coefficients for the Variables

Variables	Attendance	Class-size	Age	Scores
Attendance	1.0000	0.0120	0.0600	-0.0200
Class-size		1.0000	-0.0310	0.0000
Age			1.0000	0.0000
Scores				1.0000

In general, there exists a weak relationship between the variables, which satisfies the criteria for conducting panel regression analysis.

3.2 Boxplot

The boxplot shows the skewness of the distribution of the dataset collected for the study.

Figure 1 presents a boxplot for the gender of students where boxplot boys were negatively skewed (skewed left) with an outlier. Boxplot of girls also shows the distribution of the data was approximately symmetric with outliers.

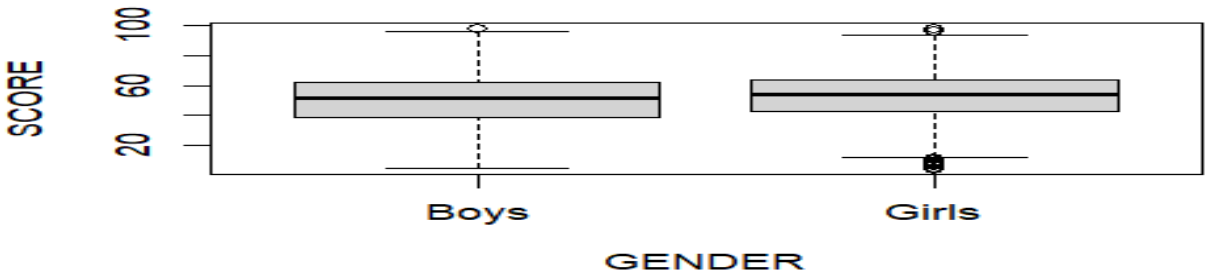

Figure 1. Boxplot for Gender of Students

Figure 2 also displays the place of residence for students in the Bolgatanga East District. Both boxplots show the data distribution was negatively skewed with outliers for the rural residents.

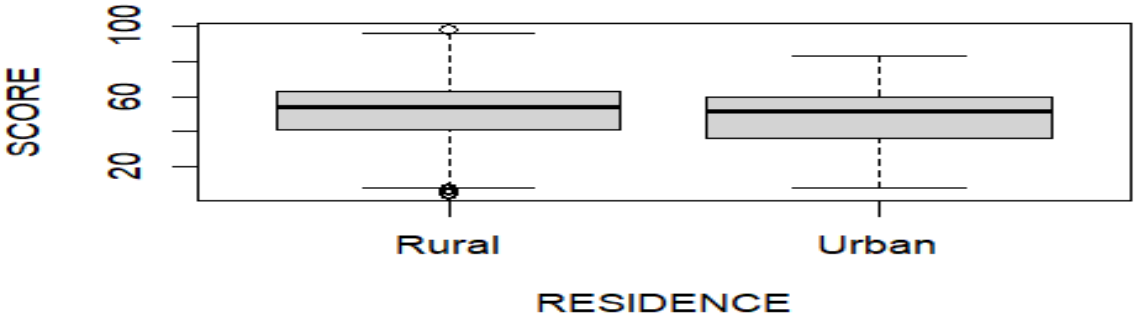

Figure 2. Boxplot for Place of Residence of Students

Figure 3 also presents the occupational status of the parents of students who were considered in this study. The study indicates that the data for both categories were negatively skewed with outliers for the informal sector of the study.

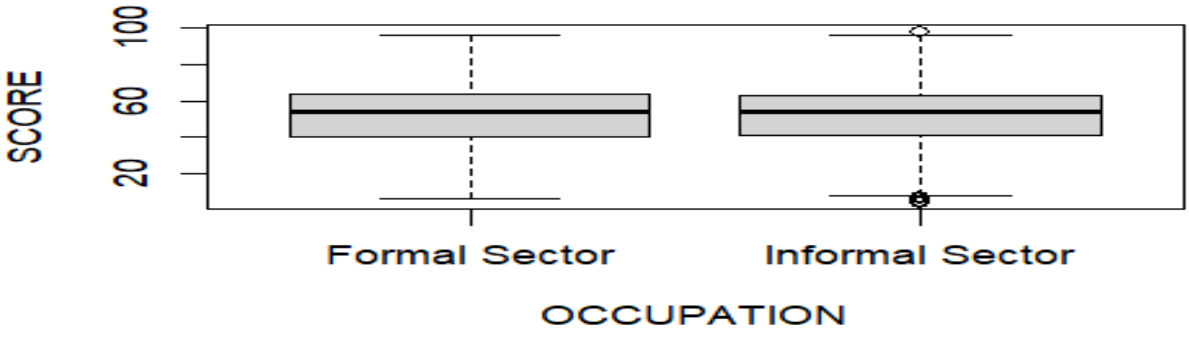

Figure 3 Boxplot for Parental Occupation of Students

Figure 4 presents a boxplot for the age of students who were considered in the study. The boxplots show that the data for all boxplots were negatively skewed with outliers in the majority of them except the last boxplot which was approximately symmetric.

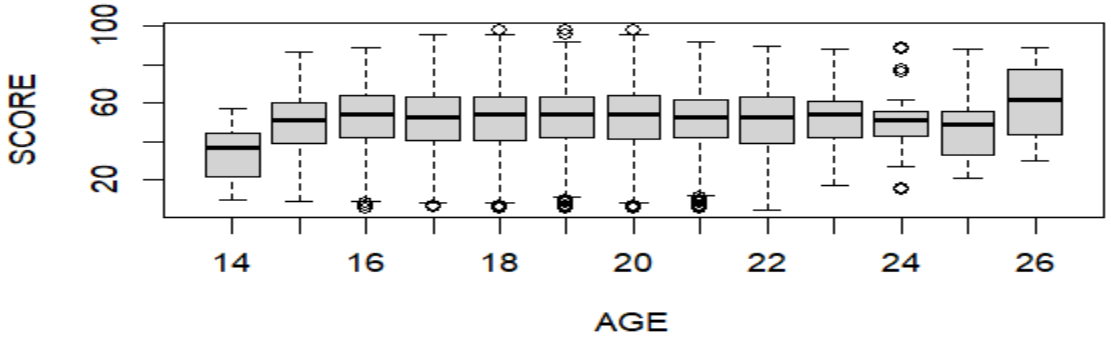

Figure 4. Boxplot for Ages of Students

3.3 Panel Regression Model

Table 3 shows that the coefficients of occupation, site of residence, and attendance are all negative and significantly correlated with students' academic performance. Gender, on the other hand, has a high and positive correlation with academic success. Furthermore, all but age and class size had positive coefficients that are substantially correlated with academic achievement.

From Table 3, it can be observed that attendance and gender are significant with p-values less than 0.05. The coefficient of attendance is negative While that of gender is positive.

Table 3. Pooled OLS Regression Estimated Coefficients

Variable	**Coefficient**	**Std. Error**	**t-value**	**Pr(>|t|)**
Intercept	48.5130	2.6066	18.6116	0.0000
Attendance	-0.0229	0.0090	-2.5436	0.0110
Age	0.0113	0.0939	0.1204	0.9042
Class Size	0.0082	0.0131	0.6218	0.5341
Residence				
Urban	2.5767	1.7771	1.4500	0.1471
Gender				
Female	1.9604	0.3175	6.1747	0.0000
Occupation				
Informal Sector	-0.1177	0.5616	-0.2095	0.8341

3.4 Nested Design Analysis

3.4.1 Academic term nested within a year

The nested design was used to predict the academic achievement of learners in the Bolgatanga East District. The findings from the nested design model are shown below.

Table 4. Estimated Coefficients of Academic Term Nested Within a Year

Variable	Coefficient	Std. Error	t-value	Pr(>|t|)
Attendance	-0.0474	0.0122	-3.8900	0.0001
Age	-0.3779	0.1091	-3.4630	0.0005
Class Size	0.0070	0.0131	0.5310	0.5954
Term				
Term 2	1124.0000	730.4000	1.5390	0.1239
Year				
2021	-48130000.0000	1456000000.0000	-0.0330	0.9736
2022	-96260000.0000	2912000000.0000	-0.0330	0.9736
Gender				
Female	1.9640	0.3165	6.2040	0.0000
Residence				
Urban	2.5320	1.7720	1.4290	0.1531
Occupation				
Informal Sector	-0.1002	0.5598	-0.1790	0.8579
Year: Term	-0.5572	0.3614	-1.5420	0.1232

The model has an adjusted $R^2 = 0.8985$ indicating that about 90% of the variation in academic achievement of students is explained by the explanatory variables. The model is significant with a P-value of 0.00 and a residual standard error of 17.23 with 13622 degrees of freedom with an F- statistic of 12060.00.

It was discovered that student gender, age, and attendance in class were statistically significant. The academic term and year, however, showed a negligible effect on the academic performance of children in the Bolgatanga East District.

3.4.2 Subjects nested within a term

The variables in the model that have topics nested within significant terms include the student's age, gender, and class attendance, as well as subjects like Integrated Science and mathematics. The remaining factors were deemed inconsequential at the 5% level of significance, along with the academic period and subjects nested inside a term.

Table 5. Estimated Coefficients of Subjects Nested Within a Term

Variabless	Coefficient	Std. Error	t-value	Pr(>|t|)
Attendance	-0.0477	0.0117	-4.0750	0.0000
Age	-0.3777	0.1091	-3.4620	0.0005
Class Size	0.0069	0.0131	0.5290	0.5971
Term				
Term 2	-0.6776	0.7807	-0.8680	0.3854
Subject				

Core Mathematics	3.4275	0.7192	4.7650	0.0000
Core Int. Science	4.7046	1.1643	4.0410	0.0001
Gender				
Female	1.9637	0.3165	6.2050	0.0000
Residence				
Urban	2.5321	1.7722	1.3680	0.1531
Occupation				
Informal Sector	-0.1004	0.5598	-0.1790	0.8576
Term: Subject	-0.5573	0.3614	-1.5420	0.1231

The model has an adjusted $R^2 = 0.8986$ showing that about 90% of the variation in the performance of students is explained by the explanatory variables. The model is significant with a p-value of 0.0000.

It is observed that the points in the normal Q-Q plots follow quite a linear relationship, especially for the majority of the points. It is observed that the points predominantly align with the straight diagonal line, exhibiting slight variations at both ends. According to this graph, the data is normally distributed.

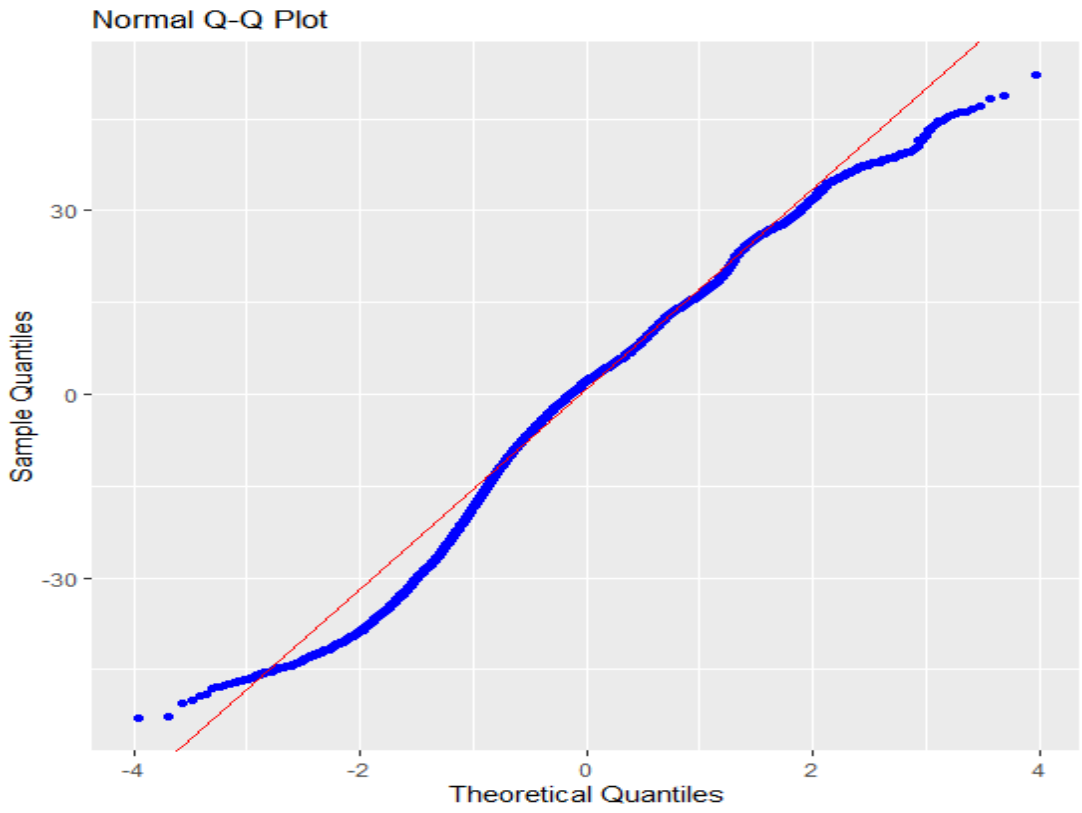

Figure 5. Nested Design Fitted

3.5 Discussion of Results

Attendance in class has been shown to be a strong indicator of academic success. According to the study, kids' academic performance in the Bolgatanga East District decreases when their attendance at class increases. The results of other studies are in conflict with this [12, 1, 3]. This may be primarily because the majority of students in the Bolgatanga East District work in farming and gold mining, also known as "galamsey." As a result, they perform poorly academically because they are unable to focus in class and do not find time to revise after school. The Upper East Region is one of the poorest regions in Ghana, and thus the students could have been pushed to “Galamsey” by the effects of poverty [25]. Generally, females are seldom involved in the “Galamsey” due to its manual nature and it is not surprising that the female students are performing better than the male students.

The gender of students also plays an important role in the academic performance of students in the study area. The study revealed that male and female students perform differently academically. Compared to male students, female students achieve nearly twice as well academically. Empirical evidence suggests that gender has an impact on students' academic performance. This is consistent with the findings of [16] and [18], whose researchers discovered that factors like students' gender are linked to academic achievement. This could be because of the gender roles they play at their respective destinations.

Students' academic performance was impacted by both gender and age, which were statistically significant. According to a study by [11] and [10], this is consistent. Some students may have enrolled late, which could explain the findings. Older pupils, on the other hand, might be busy with home chores, which would limit their time for review and result in poor results.

The Bolgatanga East District's academic performance is not significantly influenced by class size. A study by [6] found no statistically significant improvement in students' performance in either smaller or bigger courses, which contradicts the study's findings.

Additionally, the study found that adolescents' academic performance and well-being were not significantly impacted by their parents' careers or where they lived. This supported the findings of other researchers that domicile has little bearing on academic achievement [22].

4. Conclusions

According to the study, the academic performance of kids in the Bolgatanga East District is significantly influenced by their age, gender, and attendance in class. The district's students' academic performance was also significantly influenced by gender and class attendance, according to both models. According to the layered design, one factor affecting pupils' academic performance is their age. Stakeholders should increase stakeholder education, particularly in rural areas, to deter student absences and raise the academic achievement of students from underprivileged areas.

Acknowledgements

To my supervisor, Prof. Dioggban Jakperik, I truly appreciate your dedication. May the good Lord bless you and your family for your excellent guidance, insightful feedback, constructive pressure, and

criticisms. May God grant you abundant blessings. I am particularly grateful to my family for their unwavering support in every aspect of my life, especially my sister, Mrs. Magdalene Kannae, whose insights have been pivotal to my achievement. I genuinely value the significant impact she has had on my life. I express my heartfelt gratitude to the Examination Officer and head teachers of the Bolgatanga East District who willingly contributed their part in the data compilation.

References

[1] AbuRuz, M.E. Is there a correlation between frequent class absences and diminished academic performance? European Scientific Journal, 2015, 11(7).

[2] Ali, A., Ambler, G., Strydom, A., Rai, D., Cooper, C.A., McManus, S. and Hassiotis. The link between happiness and IQ: the role of socio-economic and clinical elements. Psychological Medicine, 2013, 43(6), 1303-1312.

[3] Alghamdi, A., Yamani, A., Khalil, A., Albarkati, B., Alrehili, O. and Salih, M. The frequency, reasons, and effects of absenteeism among medical students at UQU. Education, 2016, 6(1), 9-12.

[4] Eze, V.A., Ezenwafor, J.I. and Obidile, I.J. The effort put forth by students contributes to their success in exams. IOSR Journal of Humanities and Social Science, 2016, 32(3), 93-107.

[5] Farooq, M.S., Chaudhry, A.H., Shafiq, M. and Berhanu, G. Factors affecting students' quality of academic performance: a case of secondary school level. Journal of Quality and Technology Management, 2011, 7(2), 1-14.

[6] Hoxby, C.M. The impact of class size on student performance: fresh insights from population differences. The Quarterly Journal of Economics, 2000, 115(4), 1239-1285.

[7] Hsiao, C., Lahiri, K., Lee, L. and others. Examination of Panel Data and Models with Limited Dependent Variables. Cambridge: Cambridge University Press, 1999.

[8] Institute of Statistical, Social and Economic Research (ISSER). The condition of the Ghanaian economy in 2007. Accra: Sundel Services, 2008.

[9] Ismail, M., KhatiBi, A. and Azam, S.F. The influence of school culture on the effectiveness of government schools in the Maldives. Participatory Educational Research, 2022, 9(2), 261-279.

[10] Joseph, F.A.B.U.N.M.I. The relationship between adolescents' gender and age about career selection among secondary school students in Southwest Nigeria. Advances in Social Sciences Research Journal, 2017, 4(4).

[11] John, G. and Johnes, J. International Handbook on the Economics of Education. Cheltenham, U.K. and Northampton, Massachusetts: Elgar, 2004, pages 860.

[12] Kirby, N. and McElroy, B. The impact of attendance on grades for first-year economics students at the University College Cork. Econ Soc Rev, 2003, 34, 311–326.

[13] Kurihara, Y., Tripoli, G.J. and Bender, M.A. Development of a flexible nested-mesh primitive equation model. Monthly Weather Review, 1979, 107(3), 239-249.

[14] Leithwood, K., Seashore, K., Anderson, S. and Wahlstrom, K. A summary of studies: The impact of leadership on student learning. 2004.

[15] Nankana and Asuogyaman Districts in Ghana. Academic Arena, 2009, 1(4), 73-96.

[16] Nasri, H. and Ahmed, E. Elements influencing student achievement. Journal of Business Education, 2006, 82(5), 282-290.

[17] Nguyen, N.T., Allen, L. and Fraccastoro, K. Personality can predict academic achievement: exploring the influence of gender as a moderating variable. Journal of Higher Education Policy and Management, 2005, 27(1), 105–116.

[18] Perera, H.N., Maghsoudlou, A., Miller, C.J., McIlveen, P., Barber, D., Part, R. and Reyes, A.L. The connections between self-efficacy in science teaching and teaching methods, student performance and support, along teacher job satisfaction. Contemporary Educational Psychology, 2022, 69, 102041.

[19] Schwerdt, G. and Wuppermann, A.C. Investigating how teacher evaluations relate to student assessment outcomes in Washoe County. Peabody Journal of Education, 2008, 79(4), 54–78.

[20] Sheard, M. The factors of hardiness, commitment, gender, and age distinguish university academic success. British Journal of Educational Psychology, 2009, 79(1), 189-204.

[21] Shoukat, A., Ilyas, M., Azam, R. and Ch, A.H. The influence of parental education on the academic success of children. Secondary Education Journal, 2013, 2(1), 53-59.

[22] Stephens, N.M., Fryberg, S.A., Markus, H.R., Johnson, C.S. and Covarrubias, R. Unseen disadvantage: how American universities' focus on independence undermines the academic performance of first-generation college students. Journal of Personality and Social Psychology, 2012, 102(6), 1178.

[23] Teig, N. and Steinmann, I. Utilizing extensive assessments to promote effective and fair educational practices: an examination of the Nordic nations. Large-Scale Assessments in Education, 2023, 11(1), 21.

[24] Zulfikar, R. and STp, M.M. A review of the estimation model and selection techniques for panel data regression: an overview of common effects, fixed effects, and random effects models. JEMA: Jurnal Ilmiah Bidang Akuntansi, 2018, 1-10.

[25] Jakperik, D., Otieno, R.O. and Orwa, G.O. Inference on poverty indicators for Ghana. Journal of Statistical and Econometric Methods, 2019, 8(1), 83-96.