JENER Journal of Empirical and Non-Empirical Research
Volume 2, Issue 1, Page 343-352 | Article No. 40
DOI: 10.4898/jener.v2i1.a40
ISSN: 7895-9216
Journal homepage: https://jenerjournal.com

JENER Journal
JENER Journal of Empirical and Non-Empirical Research

Binary Logistic Regression Analysis of the Factors Affecting the Academic Performance of Pupils in Lgea Primary School, Ichwa, Makurdi, Benue State

M.K. Adamu[1]*, P. Onuche[2] and U.S. Usu[3]

[1]*Department of of Statistics, Joseph Sarwuan Tarka University, Makurdi*
[2]*Department of of Statistics, Joseph Sarwuan Tarka University, Makurdi*
[3]*Department of of Statistics, Joseph Sarwuan Tarka University, Makurdi*
*Corresponding Author: M. K. Adamu, Email: adamu.michael@uam.edu.ng

APA Citation and Referencing: Adamu, M.K., Onuche, P., & Usu, U.S. (2026). Binary Logistic Regression Analysis of the Factors Affecting the Academic Performance of Pupils in Lgea Primary School, Ichwa, Makurdi, Benue State. *JENER Journal of Empirical and Non-Empirical Research*, 2(1), 343-352

ARTICLE INFORMATION

Article history:
Published on 27th Jan 2026

Keywords:
Binary Logistic regression
Maximum Likelihood Estimator
BMI
Academic performance

ABSTRACT

In this study, we considered the application of binary logistic regression model to predict and determine the factors affecting academic performance of pupils. The choice of this model becomes imperious as a result of dichotomous relationship existing in the model (either pass or fail). 242 from 650 pupils of the LGEA primary school Ichwa were selected for the purpose of the study. The R Statistical package was used for the analysis. The results show that only the occupation of parents was a significant predictor while gender, age, weight, height, BMI, were all insignificant predictors. The large difference between Null deviance and residual deviance tells us that the model is a good fit since the greater the difference better the model. Further investigation is needed to test the performance of and validate the model. The fitted model indicated that fitted binary logistic regression model could be used to predict the future performance of pupils.

1. Introduction

Over the years, logistic regression model has become a useful tool in business, medicine, epidemiology sociology and marketing and so on as it can predict the ailment disease, the granting of a credit to a specific individual, the success or failure of business, among many others. On the flip side, the same model could be used for predicting whether a particular student will pass or fail when the number of hours studied is provided as a feature and the variable. For the response has two values pass and fail. The learning process is one of the most important activities in the life of a child. The learning process is connected with academic achievement and the achievement influences the life satisfaction of children as well as all human (Kubiatko *et al.*, 2018). Student’s academic performances are affected due to social, psychological, economic, environmental and personal variables. The variables might be the type and location of school, quality of teaching, study habit, economic and educational background of parents, peer influence, unhealthy lifestyle etc. Academic achievement affects lives of students to the full extent (Koc *et al.*, 2018). For this reason, educational institutions all around the world consider it an important target to improve the student’s academic success.

Also, number of teachers, instructors, coaches and mentors lack required specialized education in the subjects as well as technical skills for comprehending the necessities of particular topics, as a result, this produces negative impacts on students learning and achievement (Kayalar, 2017). Each student has different way of processing information and different spaces of working. They have different backgrounds and bring different knowledge and experiences to the class room environment. They show various approaches to completing task such interacting and communicating in school environment and processing information (Rao and Meo, 2016).

1.1 Statement of the problem

This study addressed the factors affecting performance of primary school pupil which includes; Teaching method, unhealthy lifestyle, an uncomfortable learning environment, learning infrastructure, difficulty in understanding teacher to pupil ratio, information overload, performance pressure, family background etc. Overweight and obesity have significant impacts on both physical and psychological health with snore and long-term adverse effects. People who are obese and overweight are more likely to have risk factor for heart disease, hypotension and more. The interest of the study was therefore to investigate factors that significantly influence pupil academic performance in primary education. Adopting a policy for Body Mass Index (BMI) reporting program labeling of all food beverages to promote pupil’s awareness of calories as well as taxes on sugar sweetened beverages will provide means of reducing the prevalence of obesity and overweight.

1.2 Aim and Objectives of the study

The aim of the study is to fit a binary logistics regression model to determine the factors affecting the academic performance of pupils in primary school. The objectives of this study are:

i. To assess and identify the factors which influence pupil academic performance using binary logistic regression model.

ii. To determine whether there is a relationship between academic performance of primary school pupil and their body mass index.

iii. To predict academic performance of primary school pupil based on the significant factors affecting primary education.

2. Literature Review

Frenette et al. (2015) conducted research on Academic Outcomes of Public and Private High School Students: What Lies Behind the Differences? Private high school students score significantly higher than public high school students on reading, mathematics, and science assessments at age 15, and have higher levels of educational attainment by age 23. Students who attended private high schools were more likely to have socio-economic characteristics positively associated with academic success and to have school peers with university-educated parents. School resources and practices accounted for little of the differences in academic outcomes. Maria et al. (2016) used logistic regression model to investigate the academic performance of first-year medical students in the biomedical area. The study showed that student self-perception of his or her own ability to perform the required tasks of a course is a factor that predicts academic performance. Therefore, teachers must consider the importance of reassuring the student self-esteem when learning new concepts. Olufemi et al. (2018) researched on the factors affecting students' academic performance in colleges of education in southwest, Nigeria. The result revealed that students' factors, parental background, school factors, and teachers' factors have serious influence on students' academic performance. Ayoola F.J. et al. (2018) conducted research on Binary Logistic Analysis on the Assessment of Students' Academic Performance on School-Based Factors. The study employed the binary logistic regression to estimate the academic performance. The result shows that students' performance will be better if the number of students per teacher follows international best practice (IBP) and library and laboratory are well equipped. Jarrod and Ramlee (2019) conducted research on binary logistic regression analysis of instructional leadership factors affecting English language literacy in primary schools. The findings revealed that supervising and evaluating instruction emerged as strong predictors of literacy. Schools were more than 17 times more likely to achieve 100% literacy rate when the instruction processes were supervised and evaluated by the headmasters. Okoedion et al. (2019) examined possible factors affecting students' academic performance in Nigerian universities. Student related factors, lecturer related factors, institutional related factors, parental levels of education and income, textbooks availability and accessibility, libraries, practical laboratory, meals provision and teachers have tremendous effects on the academic performance of students at school. The study recommends that these factors should be regularly monitored and adjusted to meet students' needs and aspirations. Noora Shrestha (2019) Applied Binary Logistics Regression Model to Assess the Likelihood of Overweight. The binary regression analysis was executed to determine the influences of gender, physical activity index, and physical measurements on the likelihood that the subjects fall in overweight category. The prevalence of overweight is higher (27.8%) in female than in male (14.7%) subjects. The odds ratio for gender reveals that the likelihood of subjects falling to overweight category is 2.6 times higher in female compared to male subjects. Brew et al. (2021) carried out research on a literature review of academic performance. The paper elucidated how these factors negatively affect academic performance. The study found out that truancy affects academic performance drastically and sometimes even leads to school dropout. The multiple linear regressions result showed that school distance, availability of school facilities and level of awareness about gender issues were the school factors that influence academic performance of female students. Ketema et al. (2022) conducted a study titled Factors that Influence an academic Performance of female students. The result shows various factors, such as personal factors, socioeconomic factors and home related factors. The findings indicated that the major factors affecting students' academic performance are student, lecturer and institutional related factors, with a significant joint contribution on poor academic performance of students.

3. Methodology

3.1 Research design

The design of this study was based on survey research in which data was collected for the objectives of the study. The research was based on the study of factors affecting the academic performances of LGEA Ichwa primary school pupil.

3.2 Sample size and sampling procedure

A population size of 650 pupils will be used to determine sample size from a given population.

3.3 Sampling technique

The Simple random sampling was used to select a sample of 242 pupils out of 650 pupils, which was used so that every member of the population has an equal chance of being selected in the sample.

3.4 Method of Data Collection

For the purpose of this study, interview method of data collection was used.

3.5 Data Collection Procedure

The researcher conducted direct personal interviews on the negative effect of obesity and other factors that affects academic performance of primary school pupils.

3.6 Data Description

The study used variables such as Gender, Height, Weight, BMI status, Age, Grade average for a term, and occupation of parent. Grade average code as 0-pass and 1-fail was the dependent variable, while Age, BMI, occupation of parent and Gender are the independent variables.

3.7 Binary Logistic Regression Model

Binary logistic regression determines the impact of multiple independent variables presented simultaneously to predict membership of one or other of the two dependent variable categories.

3.8 Logistic Sigmoid Function

Due to the dichotomous nature of the outcome variable in a logistic regression, unlike the linear regression (with an output consisting of continues number of values), the logistic regression transforms it output using logistic sigmoid function. The logistic sigmoid function fits a set of data with independent variable(s) taking any real value, and the dependent variable being eith er o or 1. Therefore, the logistic sigmoid function maps predicted values to the probabilities as it maps an y real value into another value between 0 and 1, this result to a perfect output representation of the probabilities as the function always lies in the range of 0 to 1. For the parameter, say z, the logistic sigmoid function is given as:

$$h\,(Z) = 1 = \frac{1}{1-e^{-z}} \qquad 3.1$$

Where:

h (z) = output between 0 and 1 (probability estimate)

z = input to the function (your algorithm's prediction e.g. mx + b)

e = base of natural log

The Binomial Distribution as an Error Distribution in logistic Regression

The binomial distribution is appropriate to use as an error distribution in logistic regression because: the outcome of inter est is dichotomous (a success or a failure); and number of independent trials were considered in obta ining the required information for the analysis-. For the purpose of this research work,

let Yi= 1 if the ith secondary school student will perform well and

Yi = 0 if the ith secondary school student will not perform well

where yi is the category of student performance of secondary school i. Here, yi is considered as a realization of a random variable Yi that can take the values one and zero with probabilities pi and 1−pi respectively. The distribution of Yi is called a Bernoulli distribution parameter pi and can be written as

$$pr(Y_i{=}y_i) = p_i^{y_i}]\,(1{-}pi)^{1-y_i} \qquad 3.\ 2$$

For y_i = 0,1. If y_i=1 we obtain p_i, and if y_i = 0 we obtain 1− p_i

The Expected value and variance of Y_i are $E\,(Y_i\ ) = \mu_i = \mathrm{p_i}$ and var $(Y_i) = \sigma_i^2 = p_i(1{-}p_i)$

The mean and variance depend on the underlying probability $\mathrm{p_i}$ and any factor that affects the probability will alter not just the mean but also the variance of the observations. This suggests a linear model that allows for k independent pupils $y_1,\ldots, k$, (k=1…,). The ith pupil can be treated as a realization of a random variable Yi. We assume that Yi has a binomial distribution $Y_i \sim Bin\ (n_i,\, p_i)$ with binomial denominator n_i and probability p_i. For individual data, n_i=1 for all i. The outcome variable (pupil's performance) used in the analysis is dichotomous and the number of independent trials were considered in the analysis.

3.9 Logistic Regression Model

Noora Shrestha (2019) applied the binary logistic regression analysis to predict the likelihood that a subject falls into any one of the two groups of a dichotomous dependent variable (overweight or no overweight) based on the independent variables that were continuous, and categorical. The reasons for selecting this model were i) it is particularly flexible and ii) gives momentous interpretation in health studies. Let us assume that a sample of n independent observations of the pair $(x_i\ y_i)$, i = 1, 2…. n. The probability distribution of the outcome variable is Binomial i.e. $y_i \sim$ Bin $(n_i,\, \pi(x_i))$ where, y_i denote the value of a dichotomous response variable, and x_i denotes the value of the independent variable for the ith subject.

Now,

$$\mathrm{y}_i = \begin{cases} \text{1,if the subject is overweight} \\ 0, otherwise \end{cases}$$

$$\text{Logit } \pi(\mathrm{x}) = \ln\left[\frac{\pi(\mathrm{x})}{1-\pi(\mathrm{x})}\right] = \beta_0 + \beta_1\mathrm{x} \qquad 3.3$$

Here, π (x) is the predicted probability of the event which is coded with "1" overweight rather than "0" no overweight. 1 - π(x) is the predicted probability of the other decision and x is the independent variable. The logit value may be continuous and ran ge; Let us consider the conditional mean, as π (x), the expected

Value of y given the value of x in logistic regression:

$$\pi\,(\mathrm{x}) \quad = \frac{e^{\beta_0 + \beta_1 \mathrm{x}}}{1+\mathrm{e}^{\,\beta_0 + \beta_1 \mathrm{x}}} \qquad 3.4$$

To fit the logistic regression model in equation (3.3) to a set of data, the unknown parameters β_0 and β_1 have to be estimated and for dichotomous data, 0 ≤ π (x) ≤1. The binary regression model intends to predict the logit, which is the natural log of the odds of subjects to be overweight or no overweight with predictors such as gender, physical measurement, and physical activity index. The term logit transformation in this model is the transformation of p(x) which is defined as:

from $-\infty$ to $+\infty$.

The expression for π (x) in equation (3.4) provides the conditional probability P (Y=1|x) and the term 1- π (x) gives the conditional probability $P\ (Y = 0|x)$ for an arbitrary parameter $\beta = \beta_0$ and β_1, the vector parameters.

For those pairs (x_i, y_i), if y_i = 1, the contribution to the likelihood function is $\pi(x_i)$, and if yi = 0, it is 1- π (x_i); where $\pi(x_i)$ is the value of π(x) computed at xi. It can be expressed as:

$$\pi(x_i)^{\,y_i}\ [1-\ \pi(x_i)]^{\ 1-y_i} \qquad 3.5$$

For logistic regression, the observations are assumed to be independent, so the likelihood function is obtained as follows:

$$L(\beta) = \prod_{i=1}^{n} \pi(x_i)^{\,y_i}\ [1\pi(x_i)]^{\ 1-y_i} \qquad 3.6$$

For ease of mathematical calculations, log of equation (iv), log likelihood, can be written as:

$$\text{L}\ (\beta) = \ln\ [\text{L}\ (\beta)] = \sum_{i=1}^{n}\{ y_i\ \ \ln[\ \pi(x_i)] + (1-y_i)\ \ln[1-\pi(x_i\ ]\} \qquad 3.7$$

Differentiating equation (5) with respect to β_0 and β_1 for maximizing likelihood function, L (β) and solving for β we get the two likelihood equations as:

$$\sum[y_i\ -\ \pi(x_i)] = 0 \qquad 3.8$$

$$\sum x_i[y_i\ -\ \pi(x_i)]=0 \qquad 3.9$$

For binary logistic regression the terms in equations 3.8 and 3.9 are nonlinear in β_0 and β_1, and hence require iterative methods for solution which is obtained by using an iterative weighted least square technique. Then, the value of maximum likelihood estimate β will be obtained. In the present study, there are few major independent variables such as gender, physical measurement, physical activity index for which the expression of binary logistic regression model (i= 1, 2…, n subjects) is given by

$$\ln\left[\frac{\pi(x_i)}{1-\pi(x_i)}\right] = \beta_0\,, + \beta 1 x_{i1} + \beta 2\ x_{i2} + \ldots \beta n x_{in} \qquad 3.10$$

Here, x_{i1}, x_{i2}…., x_{in} are categorical or continuous independent variables. From equation 3, the equation for the prediction of the probability can be derived and solved the logit equation for $\pi\ (x_i)$ to obtain

$$\pi(x_i) = \frac{e^{\beta_0+\ \beta_1 x_{i1}\ + \beta_2\, x_{i2},\ldots,\beta_n x_{in}}}{1+\ e^{\beta_0+\ \beta_1 x_{i1} + \beta_2\, x_{i2},\ldots,\beta_n x_{in}}} \qquad 3.11$$

3.10 Model assumptions

I. Logistic regression model does not require a linear relationship between explanatory variable and response variable.

ii. The errors terms or residuals are not normally distributed.

iii. Homoscedasticity is not required.

iv. Logistics regression model assumes independent variable are linearly related to the log odds

v. There is no Multicollinearity among the explanatory variables.

3.11 Method for Estimating the Parameters of the Model

Maximum Likelihood Estimation

Suppose a binary random variable y where $= (y_{1\,,} y_{2,}\, \ldots, y_{N\,,})^{\mathrm{T}}$, each $y_{i,}\ \ i = 1, \ldots, N$ follows a Bernoulli distribution that y_i take either the value 1 or the value 0 with probability $\pi_i\,(x), 1-\pi_i\ (x)$ respectively where $\pi_i\,(x)$ is the mean of the binary variable representing the conditional probability $p(y_i\ =\ 1|x)$ and $x = \left(x_{i0,} x_{i1,} x_{i2\ ,}\ \ldots, x_{ik,}\right) \in \mathbb{R}^{K+1}$ is a vector of $K+1$ explanatory variables.

The described form of the LRM with unknown parameters $\beta = (\beta_0, \beta_1, \ldots, \beta_k\ )^T$ with $x_{i0} = 1$ given by

$$\pi_i(x) = \frac{e^{\sum_{k=0}^{K}\beta_k x_{ik}}}{1+e^{\sum_{k=0}^{K}\beta_k x_{ik}}} \qquad 3.12$$

The specific transformation of $\pi_i\ (x)$ is called the logit transformation; it is given by

$$\text{logit}\ (\ \pi_i\,(x)) = \ln\left(\frac{\pi_i\,(x)}{1-\ \pi_i\,(x)}\right)$$

Equation (3.12) can be rewritten as logit ($\pi_i\,(x)) = X^T\ \beta$

If a sample of $N-$independent observation $\{(y_i, x_i)\}_{i=1,\ldots,N}$ where $\in [\{\ 0{,}1\}\times\ \mathbb{R}^{K+1}]^N$ where y_i represents a binary outcome value, and x_i is the predictor value for the i^{th} subject (Rashid, 2008).

To find the MLE for β, the likelihood function $L(\beta)$ is defined as

$$L(\beta)\ = \prod_{i=1}^{N} \pi_i\ (x)^{y_{i(1-\pi_i\,(x)\,)^{n_i-y_i}}} \qquad 3.13$$

Substituting 3.12 in 3.13, obtain:

$$L(\beta)\ = \prod_{i=1}^{N}\left(\frac{e^{\sum_{k=0}^{K}\beta_k x_{ik}}}{1+e^{\sum_{k=0}^{K}\beta_k x_{ik}}}\right)^{y_i}\left(1-\frac{e^{\sum_{k=0}^{K}\beta_k x_{ik}}}{1+e^{\sum_{k=0}^{K}\beta_k x_{ik}}}\right)^{n_i-y_i}$$

$$=\prod_{i=1}^{N}\left(e^{\,y_i\sum_{k=0}^{K}\beta_k x_{ik}}\right)\left(1+e^{\sum_{k=0}^{K}\beta_k x_{ik}}\right)^{n_i} \qquad 3.14$$

Taking the natural log of (3.14) yields the log likelihood function:

$$l(\beta)\ = log\ \ L(\beta)\ = \sum_{i=1}^{N}\left[y_i\,(\textstyle\sum_{k=0}^{K}\beta_k\,x_{ik}) - n_i\,log\left(1+e^{\sum_{k=0}^{K}\beta_k x_{ik}}\right)\right] \qquad 3.15$$

To find the critical points of $l(\beta)$

$$\frac{\partial l(\beta)}{\partial\beta_k} = \sum_{i=1}^{N}[y_i x_{ik} - n_i\pi_i(x) x_{ik}] = 0 \qquad 3.16$$

Then, equation (3.16) can be expressed as a matrix multiplication as follows:

$$G(\beta) =\ l'(\beta) = X^T\ \beta \qquad 3.17$$

Where $l'(\beta)$ is a column vector of length $K+1$ whose elements are $\frac{\partial l(\beta)}{\partial \beta_k}, k = 0,1 \dots, K,$ likewise μ is a column vector of length N with elements $\mu_i = n_i \pi_i$, i.e. $\mu = (n_1\pi_1, n_2\pi_2, \dots, n_N \pi_N,)^T$, and X^T is a $(K+1) \times N$ matrix.
The ML estimates for β can be found by setting each of the K+1 equations in (3.16) and (3.17) equal to zero and solving for each β_k. In order to maximize the estimates each of $K+1$ equations are differentiated with respect to β_k, denoted by $\beta_k{}'$. The general form of matrix of second partial derivatives is

$$\frac{\partial^2 l(\beta)}{\partial \beta_k \partial \beta_{k'}} = \frac{\partial l(\beta)}{\partial \beta_{k'}} \sum_{i=1}^{N} [y_i x_{ik} - n_i \pi_i(x) x_{ik}]$$

$$= \; -\sum_{i=1}^{N} n_i x_{ik} \pi_i(x)(1 - \pi_i(x)) \, x_{ik'} \qquad 3.18$$

Where $k = 0,1, \dots, K$ and $k' = 0,1, \dots, K$.
Equation (3.18) can be expressed in term of matrix multiplication

$$H(\beta) = -X^T DX$$

$H(\beta)$ is called Hessian matrix, where D is $N \times N$ diagonal matrix.
Setting the equations in (3.16) and (3.17) equal to zero results in a system of $K+1$ nonlinear equations each with $K+1$ unknown variables. The solution of the system is a vector with elements, β_k. After verifying that the matrix of second partial derivatives is negative definite, and that the solution is the global maximum rather than a local maximum, then we can conclude that this vector contains the parameter estimates for which the observed data would have the highest probability of occurrence (Van Den Berg, et, al.1984).

4. Findings

Table 1 Summary Statistics

	Age	Sex	Height (m)	Weight (kg)	BMI	Academic average
N	242	242	242	242	242	242
Mean	10.49		1.3528	28.053	15.067	50.587
Median	10.00		1.3600	28.600	14.900	48.550
Mode	10		1.36	30.3	11.1	36.7
Std.deviation	2.558		0.12961	8.2060	2.7693	13.2507
Range	12		0.62	44.4	17.6	69.1

4.1 Exploratory Data Analysis

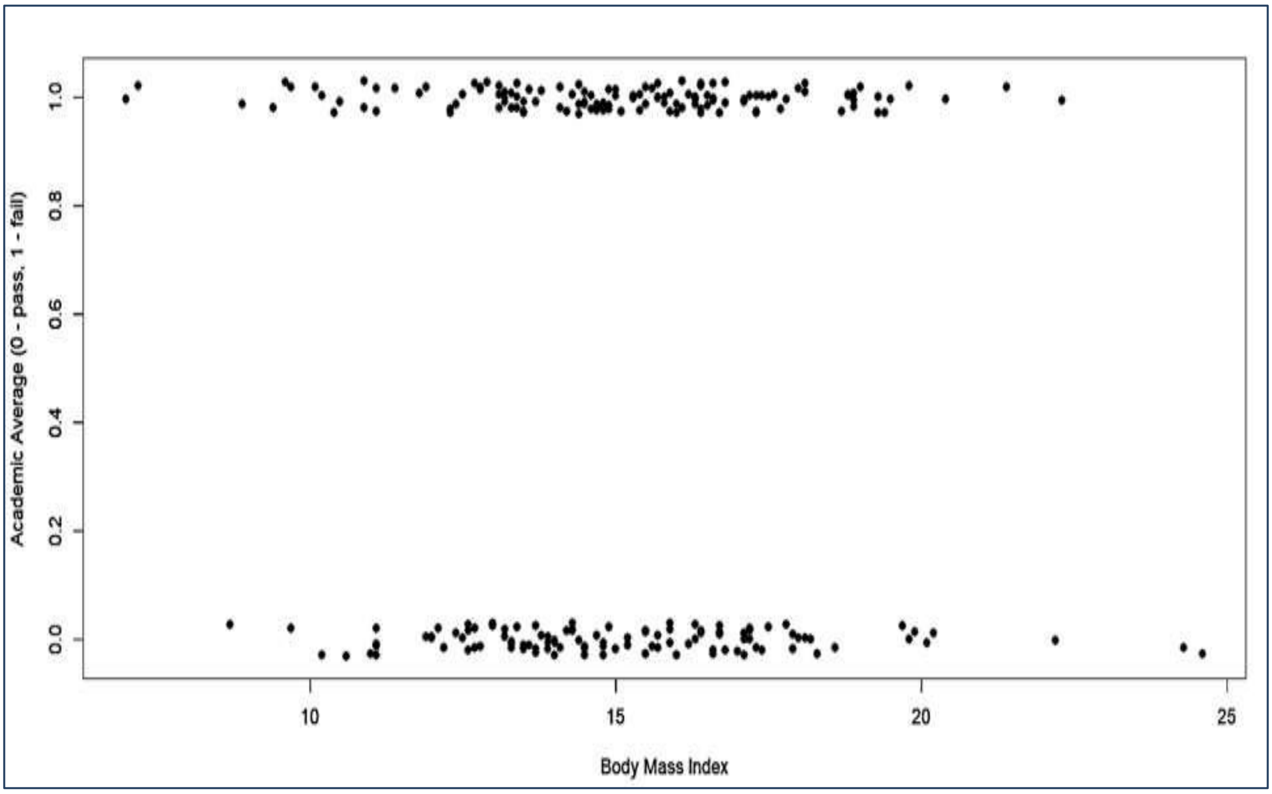

Figure 1: Binary Logistic Graph

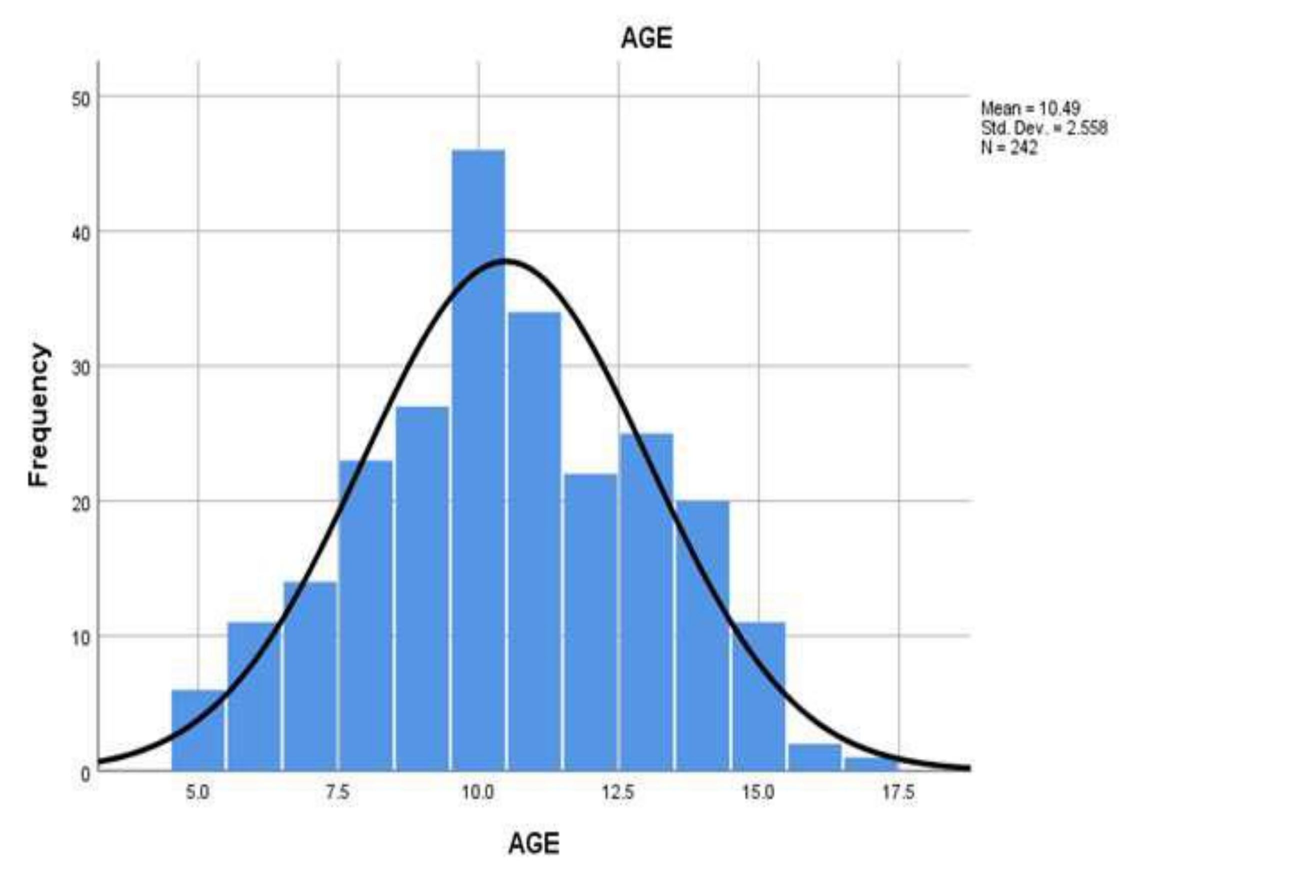

Figure 2: A graph of Age distribution of respondents

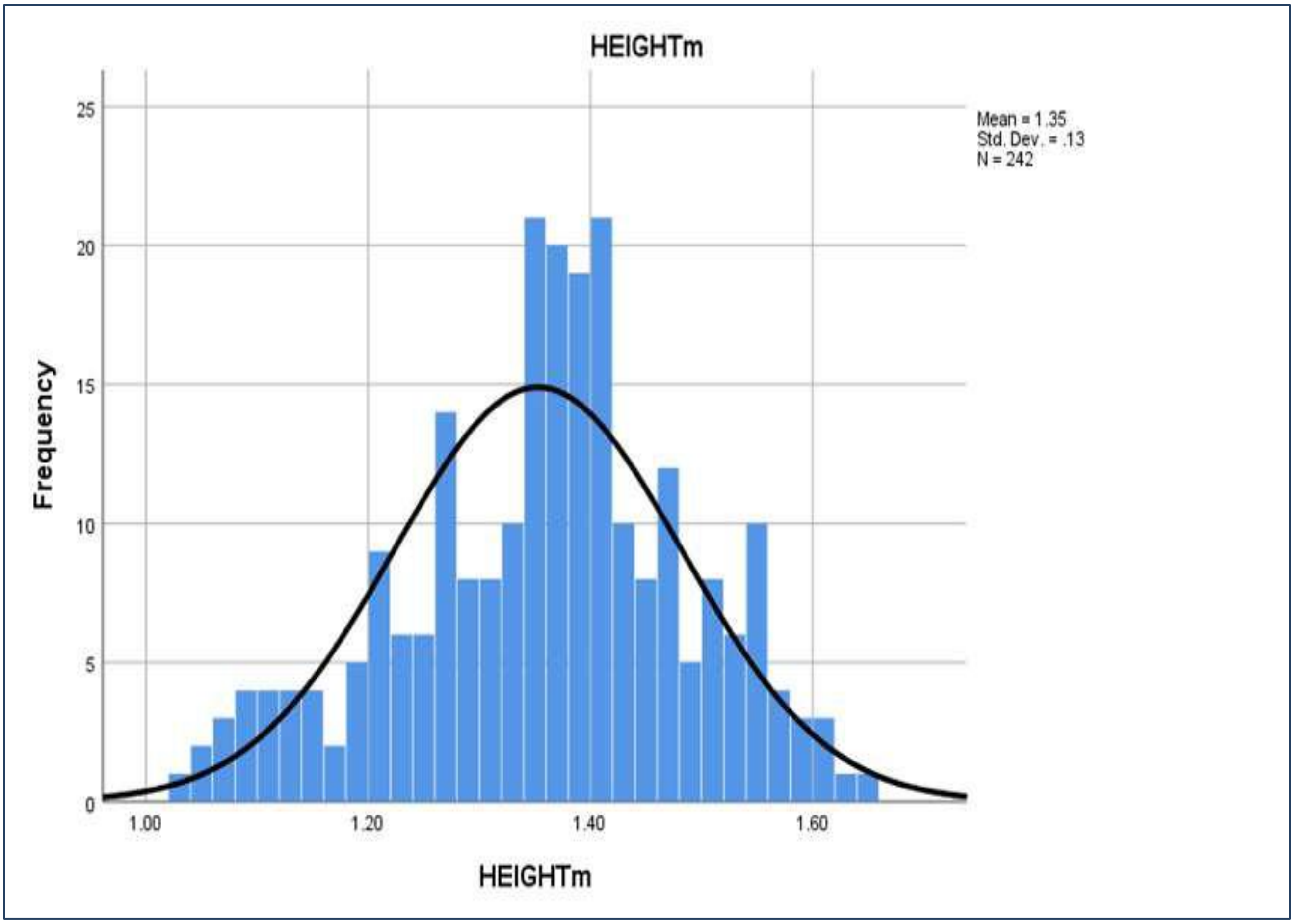

Figure 3: a graph of height distribution of respondents.

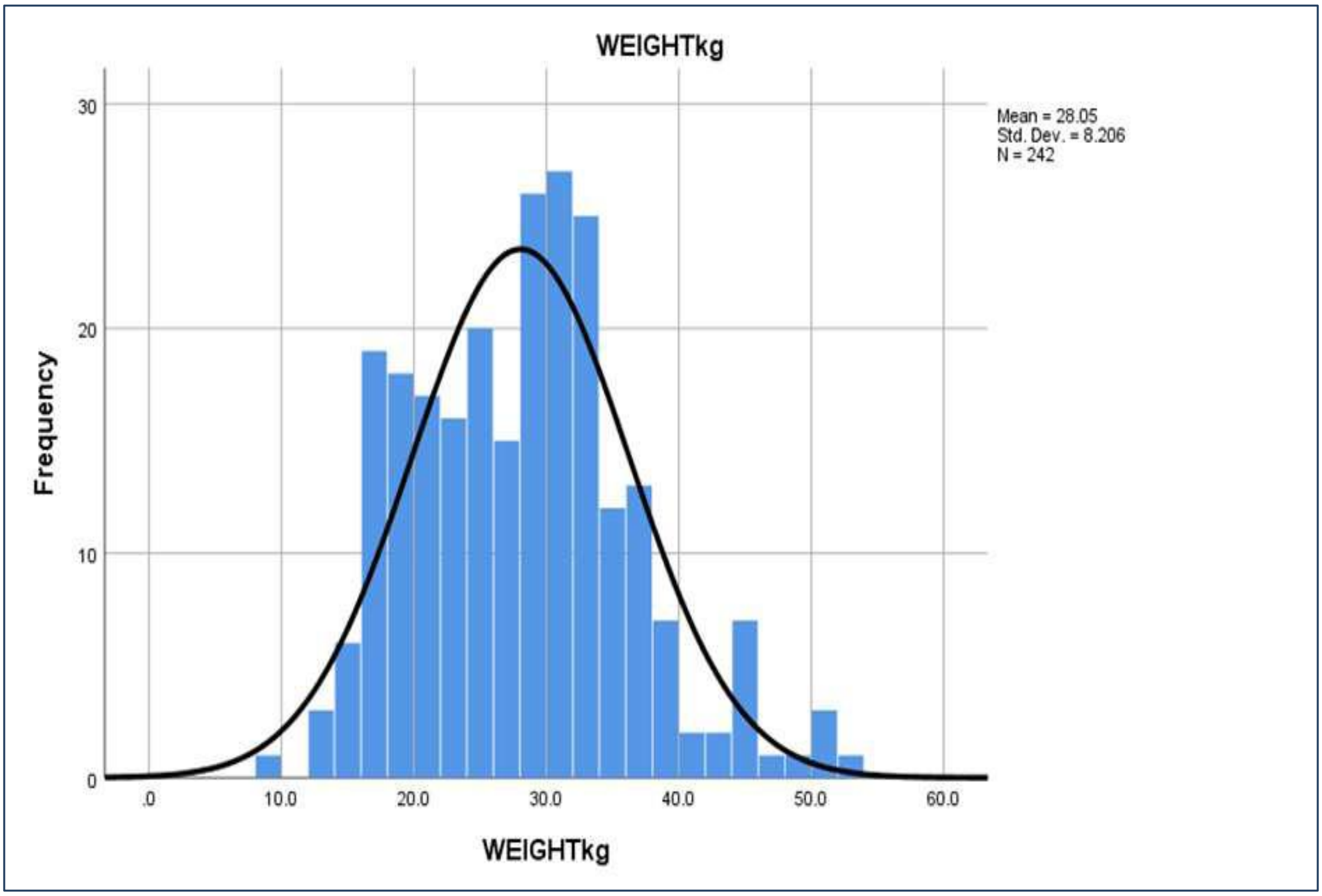

Figure 4: A graph of weight distribution of the respondent.

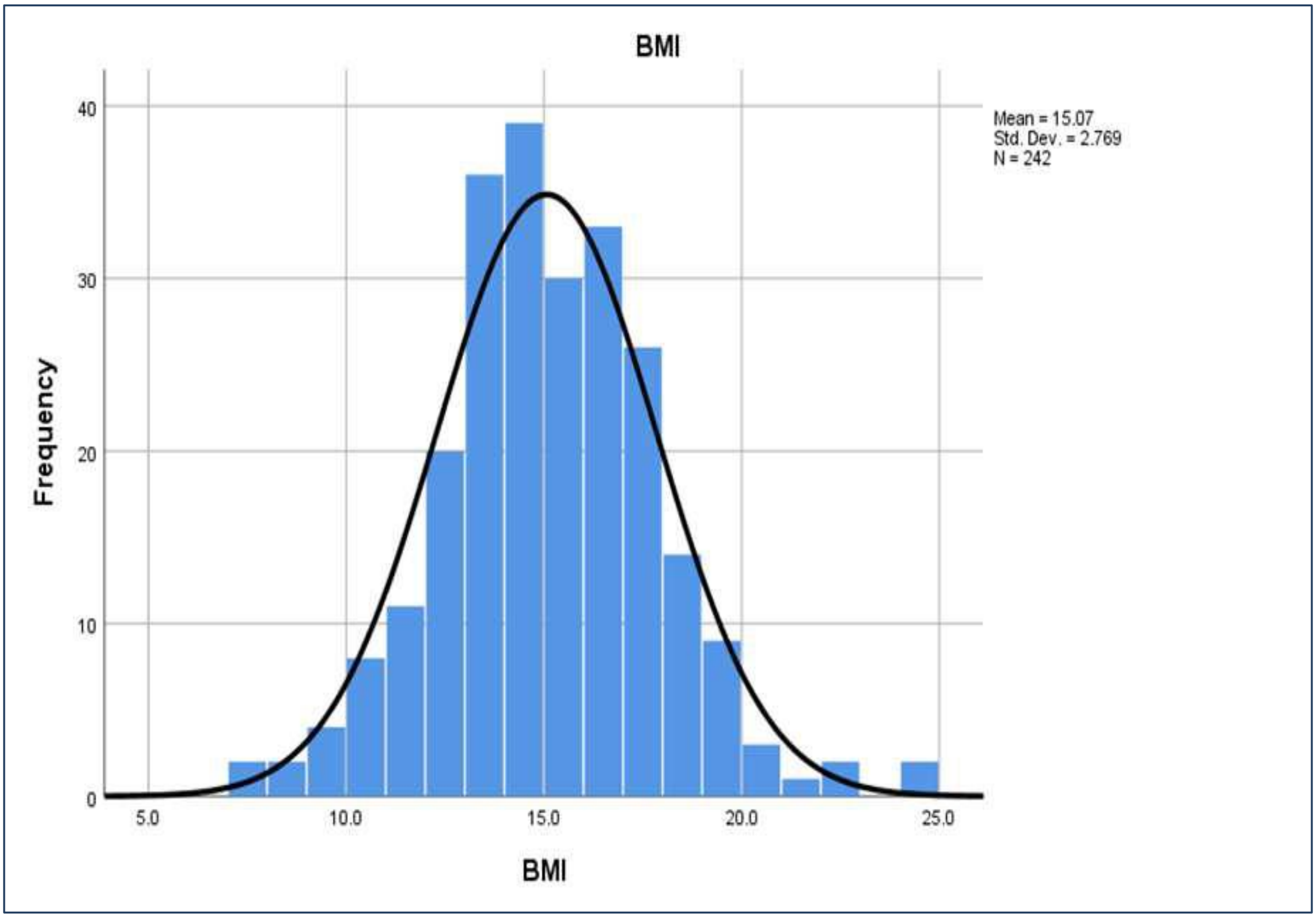

Figure 5: A graph of BMI distribution of the respondent.

Table 2: Parameter Estimate of the Binary Regression Model.

Coefficients	Estimates	Std error	Z value	P-value
Intercept	-0.4996	0.9375	-0.533	0.5941
Age	-0.0678	0.0550	-1.232	0.2181
Sex	0.1402	0.2683	0.523	0.6013
BMI	0.0010	0.0503	0.020	0.9837
Occupation of parents	0.7552	0.2642	2.858	0.0043 **

4.2 The Logistic Regression Model

$$\text{logit}(\text{Academic average}) = \; -0.4996 - 0.0678\,(\text{Age}) + 0.1402\,(\text{Sex}) + 0.0010(\text{BMI}) + 0.75552(\text{occupation of parent})$$

4.3 Odd Ratio

Table 3: Parameter Estimates and Their Odd Ratio

		OR	2.5%	97.5%
Intercept	-0.4996	0.6068	0.0952	3.8169
Age	-0.0678	0.9345	0.8379	1.0404
Sex	0.1402	1.1505	0.6807	1.9522
BMI	0.0010	1.0010	0.9067	1.1053
Occupation	0.7552	2.1280	1.2721	3.5892

4.4 Predicted Probabilities

Table 4: Predicted Probabilities

S/no	Academic Average	Age	Sex	BMI	Occupation of parent	Predicted probabilities
1	0	6	1	9.7	1	0.50
2	1	6	1	9.7	2	0.68
3	1	10	2	12.5	2	0.65
4	1	5	1	7.0	1	0.52
5	0	10	2	12.1	2	0.65
6	1	7	2	10.4	2	0.70
7	1	8	2	10.2	2	0.68
8	1	7	2	10.1	1	0.52
9	0	5	1	11.1	1	0.52
10	1	5	1	9.4	2	0.69
11	0	9	2	11.1	1	0.48
12	0	11	1	15.9	1	0.42
13	1	9	1	10.9	2	0.63
14	1	10	1	14.8	1	0.43
15	1	9	2	15.3	2	0.67
16	1	9	2	12.4	2	0.67
17	1	7	1	10.5	2	0.67
18	0	8	1	17.2	1	0.47
19	0	11	2	17.8	1	0.45
20	1	10	1	13.5	2	0.65

Table 5: Classification table

Academic Average	False	True
0	60	51
1	45	86

4.5 Sensitivity and Specificity

The sensitivity of events correctly predicted is at 65.6% and specificity of non-events correctly predicted is at 54.1%.

4.6 Discussion

The mean from table 1 above indicated that the mean value of the respondent for age was 10.49, that of height was 1.3528, weight was 28.053, BMI was 15.067. The median from the table 1 above indicates the median value of respondent for age was 10.00, height was 1.3600, weight was 28.600, BMI was 14.900. From figure 1, plot, all of our data points take either the value 0 (pass) and 1 (fail) to depict the binary characteristics of the academic average against the body mass index. From figure 2, the highest percentage of the respondents were between the ages of 9-10 years old with a mean age of 10 and a median age of 10. From figure

3, the data was normally distributed with the mean of 1.35 and standard deviation of 0.13. From figure 4, the data was normally distributed with the mean of 28.06 and standard deviation of 8.206. From figure 5, the data was normally distributed with the mean of 15.07 and standard deviation of 2.769, and from table 2, since p-value is < 0.05 for Occupation of parents, this means that the Occupation of Parents is significant while other variables show no significant contribution to the model. From the binary logistic model in equation (4.0) the Intercept (-0.4996) is the value of the logit (Academic average) when other variables are zero. The coefficient of age is -0.0678, meaning that for every unit increase in logit (Academic average) there is a decrease of -0.0678 in age of the respondent. The coefficient of sex is 0.1402, meaning that for every unit increase in logit (academic average) there is an increase of 0.1402 in sex of the respondent. The coefficient of BMI is 0.0010, meaning that for every unit increase in logit (academic average) there is an increase of 0.0010 in BMI of the respondent. The coefficient of occupation of parent is 0.7552, meaning that for every unit increase in logit (academic average) there is an increase of 0.7552 in occupation of parent of the respondent. The results shows that only one out of the five predictor variables are significant, at $p < 0.05$ level of significance. The large difference between Null deviance and residual deviance tells us that the model is a good fit since the greater the difference better the model. Null deviance is the value when you only have intercept in your equation with no variables and residual deviance is the value when all the variables are taken into account. It makes sense to consider the model good if that difference is big enough. From table 3, the odd ratio for occupation of parents is approximately 2.13 which mean that for one unit change in occupation of parents, the odds of obtaining a pass in academic average is 2.13, From table 4, the last column contains the probabilities generated using the fitted model, from table 5, the accuracy rate is 60 plus 86, divided by 242. Hence the accuracy is 60%., also, from table 5, True indicates 'predicted pass' and false indicates 'predicted fail'. There were 60 correctly predicted fail and 86 correctly predicted pass. There are 51 wrongly predicted as fail and 45 wrongly predicted as pass.

5. Conclusion and Recommendations

5.1 Conclusion

Binary logistic regression analysis was used to investigate the factors that affect the academic performance of primary school pupils. The model in this study identified variables that predicted academic performance among primary school pupils. A logistic model was fitted to data on some variables provided by interview administered on 242 pupils out of 650 pupils in LGEA primary school Ichwa Northbank Makurdi. The Parameter Estimate of the binary regression model indicates that the p-value is < 0.05 for Occupation of parents; this independent variable is significant while other variables show no significant contribution to the model. The results shows that only one out of the five predictor variables are significant, at $p < 0.05$ level of significance. The large difference between Null deviance and residual deviance tells us that the model is a good fit since the greater the difference better the model. The findings of the study revealed that occupational status of parents played an important role in academic performance of primary school pupils.

5.2 Recommendations

Based on the result from this study, we recommended that further research should be carried out to incorporate more factors affecting the performance of pupil in primary schools, as sufficient evidence was not available in this study to justify that only occupation of parents affects the academic performance of a pupil.

6. References

[1] Ayoola F.J., and Balogun K.O, "Binary Logistic Analysis on the Assessment of Students' Academic Performance on School-Based Factors in Some Selected Public Secondary Schools in Akinyele Local Government, Oyo State, Nigeria." American *Journal of Applied Mathematics and Statistics*, vol. 6, no. 2 (2018): 61-66. doi: 10.12691/ajams-6-2-4.

[2] Berg, C., Christensen, J. P. R., & Ressel, P. (1984). Harmonic Analysis on Semigroups: Theory of Positive Definite and Related Functions (Graduate Texts in Mathematics, Vol. 100). New York, NY: Springer-Verlag.

[3] Brew, E.A., Nketiah, B. and Koranteng, R. (2021) A Literature Review of Academic Performance, an Insight into Factors and their Influences on Academic Outcomes of Students at Senior High Schools. *Open Access Library Journal*, 8, 1-14. doi: 10.4236/oalib.1107423.

[4] Frenette, M. & Chan, P.C. (2015). Academic outcomes of public and private high school students: What lies behind the differences? *Statistics Canada Social Analysis and Modeling*, 367.

[5] Jarrod, K. and Ramlee, R. (2019): Schooling and Human capital, oxford university press, New York NY, USA

[6] Kayalar, F., and Kayalar, F. (2017). The Effects of Auditory Learning Strategy on Learning Skills of Language Learners (students' views). *IOSR Journal of Humanities and Social Science*, 22(10), 4–10.

[7] Ketema, S., Shukri, A. and Shimelis, B. (2022) Factors That Influence an Academic Performance of Female Students in Kabridahar District, Somali Regional State, Ethiopia. *Open Journal of Social Sciences*, 10, 360-375. doi: 10.4236/jss.2022.104027.

[8] Kubiatko, M., Hsieh, M. Y., Ersozlu, Z. N., and Usak, M. (2018). The Motivation Toward Learning among Czech High School Students and Influence of Selected Variables on Motivation. *Revista de Cercetare si Interventie Sociala*, 60, 79–93.

[9] Maria, R.A., Northrup, K. L., & Cottrell, L.A (2016). Children's Aerobic Fitness and Academic Achievement: a longitudinal examination of students during their fifth and seventh grade years. *American journal of public health*, 102(12), 2303-2307

[10] Noora Shrestha, Application of Binary Logistic Regression Model to Assess the Likelihood of Overweight (2019). American *Journal of Theoretical and Applied Statistics*. Volume 8, Issue 1, pp. 18-25. doi: 10.11648/j.ajtas.20190801.13

[11] Okoedion, E. G., Okolie, U. C., & Udom, I. D. (2019). Perceived factors affecting students' academic performance in Nigerian Universities. *Studi Sulla Formazione/Open Journal of Education*, 22(2), 409-422.

[12] Olufemi, O. T., Adediran, A. A., and Oyediran, W. O. (2018). Factors Affecting Student's Academic Performance in Colleges of Education in South-West, Nigeria. *British Journal of Education*, 6(10), 43 – 46.
[13] Rao, K., and Meo, G. (2016). Using Universal Design for Learning to Design Standards-Based Lessons. *SAGE Open*, 6(4), 1–12. https://doi.org/10.1177/2158244016680688
[14] Shi Y, Yu H, Di S and Ma C (2022) Body Mass Index and Academic Achievement Among Chinese Secondary School Students: The Mediating Effect of Inhibitory Control and the Moderating Effect of Social Support. *Front. Psychol.* 13:835171. doi: 10.3389/fpsyg.2022.835171