International Journal of Research (IJR)

p-ISSN: 2348-6848
e-ISSN: 2348-795X
Vol. 8 Issue 6
June 2021

The Effect of the Omission of Intercept term in a Linear Regression Model

by

Onu, Obineke Henry[1], Nelson, Maxwell[2] and Tonubari, Apere[3]

[1,2]Ignatius Ajuru University of Education, Rumuolumeni, Port Harcourt, Rivers State, Nigeria.

[3] Computer science Education Dept. Isaac Jasper Boro COE, Sagbama, Bayelsa State, Nigeria.

Corresponding email:onuobinekehenry@gmail.com

Abstract

The effect of the omission of intercept term in a linear regression model is considered for a simple linear regression with pooled explanatory variable, simple linear regression with single explanatory variable (sugar level) and multiple linear regression with two explanatory variables (age and sugar level). The study considered estimation of the intercept and slope for a simple linear model with intercept term and the estimation of slope for a model without intercept. SPSS software 23 version was applied in the analysis. Student t-test was conducted to ascertain the best model. Coefficient of determination and correlation were also applied. The model with intercept was observed to be significant in modelling the Blood pressure and the explanatory variables for the 100 persons. Age proved to be contributing more to the Blood pressure of human beings as compared to the sugar level. As the intercept is omitted from a model, the coefficient of determination, correlation and the sum of square error increases. The plots in Appendix G, H and I show that there exist unstable relationships between the response variable (blood pressures) and the explanatory variables (heights weights, ages, sex, sugar levels etc) for the 100 persons studied.

Keywords: intercept, slop, coefficient, correlation, coefficient of determination, t-test, sum of squares.

1. Introduction

A simple linear regression is a statistical approach (method) that is used in modeling a linear relationship that exists between the response variable otherwise called the dependent or output variable and one explanatory variable, known as the independent, predictor or input variable as seen in Iwundu and Onu (2017).

A multiple linear regression model is a linear model that measures the relationship between one response variable and more than one explanatory variable. See Sameera (2014). The aim of the study is to investigate the effect of the omission of intercept term in a linear regression model. The objectives of the study include:

1) Investigating the effect of the omission of intercept term in a simple linear regression model
2) Investigating the effect of the omission of the intercept term in a model with two explanatory variables, namely, sugar level and age.

p-ISSN: 2348-6848
e-ISSN: 2348-795X
Vol. 8 Issue 6
June 2021

Showing that for model without intercept, the slope obtained by using a least square method is equal to the slope obtained when the intercept term is substituted as zero in the already known least square method for intercept term estimation in a model with intercept. The knowledge of the contribution of intercept term in a linear model has been a serious problem over the years. Kutner et al (2005) stated that intercept term has no meaning on its own, but can be meaningful collectively with other variables in a model. As seen in Bremer (2012), Johnson (2003) and Sameera, (2014). The intercept is equal to $E(Y_i)$ when all the values of the explanatory variables amounts to zero, but if not, intercept says little on its own.

Recently Sameera (2014) has made comparison between models with and without intercept using full multiple linear regression model and a multiple linear regression model with additional point called leverage point was also built. The investigation did not consider the effect of omission of intercept term in

(1) A simple linear regression model with pooled x variables which included the Height, weight, age, sugar level, sex on the response y which is the blood pressure of 100 persons.

(2) A simple linear regression model with predictor variable x which is sugar level, in order to study the relationship between the blood pressure and sugar level for 100 persons.

(3) A linear regression model, with two predictor variables x_1 and x_2 which respectively represent sugar level and age of the twenty people on their blood pressure. Against this backdrop this research is conducted to investigate the effect of the omission of the intercept term in a linear model based on the above enumerated headings. Finally, the research work will show that for model without intercept the slope, obtained by using least square method on the model is equal to substituting the intercept to be zero in the already known least square estimate for intercept, in a model with intercept. In this study, a full and reduced linear regression model was built. The full linear regression model contains the intercept and also the slop, while the reduced linear regression model contains only slope, the intercept is omitted. Using the data of blood pressure and, height, weight, age, sugar level and sex used by Sameera in 2014. This study considers only simple linear model with one predictor variable x which is the combination of all the x values called pooled variables, it also considers a linear regression for a predictor variable which is sugar level and also considered a linear model with two explanatory variables, and they are sugar level and age on the blood pressure. The analysis considered the blood pressures for 100 persons and regressed it with the pooled variables, and with the sugar level and finally with the sugar level and ages of 100 persons. Investigation was also conducted to see whether or not the least square estimation of the slope in a model without intercept is equal to putting the intercept to be zero in the least square formula for estimating intercept for a model without intercept.

2. Full and Reduced Linear Model

Iwundu and Albert-Udochukwuka (2014) investigated the behaviour of D-optimal exact designs using first-order linear model, the study considered two types of linear polynomial model, which include the model with intercept term and that without intercept term and interaction being considered in the two model as seen.

$$y = m_0 + m_1x_1 + m_2x_2 + m_{12}x_1x_2 + e \qquad 2.1$$

p-ISSN: 2348-6848
e-ISSN: 2348-795X
Vol. 8 Issue 6
June 2021

The model in (2.1) contains intercept term β_0 and interaction term β_{12} as a result, it is a full-first-order linear model. While

$$y = m_1 x_1 + m x_2 + m_{12} x_1 x_2 + e \tag{2.2}$$

is the linear model without intercept term, hence, it is a reduced first-order linear model.

Sameera (2014) studied the comparison between models with intercept term and that without intercept term in a linear process and here leverage point was applied and it was observed that evaluating the leverage in the new points was equal to the evaluation when the linear regression model was forced through the origin, i.e $m_0 = 0$ in the full model. This was achieved by augmenting the data. For augmentation of design point, see Iwundu and Onu (2017). It was also discovered that intercept was significant in the full model, but it becomes insignificant when the leverage point was added, thereby forcing the model through the origin.

If $(A_1 Y_1), (A_2 Y_2), \ldots, (A_n Y_n)$ are r observation according to the model in Sameera (2014).

$Y_i = M_0 + M_1 A_i + ei$ where $e_i \sim IN(\,0,\ \delta^2)$ i.e independently and normally distributed random error with means 0 and variance δ^2. The least square estimate of M_0 and M_1 are given as:

$$\frac{\hat{M}_1 = \sum_{i=1}^{r} (ai - \bar{a})(y_i - \bar{y})}{\sum_{i=1}^{r} (ai - \bar{a})^2} \tag{2.3}$$

$$\hat{M}_0 = \bar{y} - \hat{M}_1 \bar{A} \tag{2.3b}$$

Where

$\bar{A} = \sum_{i=1}^{r} \frac{a_i}{r}$ and $\bar{Y} = \sum_{i=1}^{r} \frac{y_i}{r}$

Sameera (2014) considered the regression line through the origin, i.e $M_0 = 0$, then it was obtained by Sameera that the least square is;

$$\hat{m}_1 = \frac{\sum_{i=1}^{r} a_i y_i}{\sum_{i=1}^{r} a_i^2} \tag{2.4}$$

by the augmentation of the original data set with the observations

$(X_{r+1}, Y_r + 1) = (r^* \bar{A}, r^* \bar{Y})$

Where $r^* = \frac{r}{\sqrt{r+1} - 1}$

It was then observed that applying the full model to the augmented data set is equivalent to applying the reduced model when the model is from the origin. This is as shown in the verified identifies.

p-ISSN: 2348-6848
e-ISSN: 2348-795X
Vol. 8 Issue 6
June 2021

$$\sum_{i=1}^{r+1}(A_i - \bar{X}_{r+1})(Y_i - \bar{Y}_{r+1}) = \sum_{i=1}^{r+1} X_i Y_i \quad (2.5)$$

In this research work, for model without intercept term, it is assumed that the intercept term $M_0 = 0$, hence from (2.3b)

$\widehat{M}_0 = \bar{Y} - \widehat{M}_1 \bar{A}$ which can be written as:

$$\widehat{M}_0 = \sum_{i=1}^{r} \frac{y_i}{r} - \widehat{M}_1 \sum_{i=1}^{r} \frac{a_i}{r} \quad (2.6)$$

But if $M_0 = 0$

$$\sum_{i=1}^{r} \frac{y_i}{r} - \widehat{M}_1 \sum_{i=1}^{r} \frac{a_i}{r} = 0$$

Which can be written as

$$\sum_{i=1}^{r} yi = \widehat{M}_1 \sum_{i=1}^{r} ai$$

Solving for $\widehat{M}_1$ we have

$$\widehat{M}_1 = \frac{\sum_{i=1}^{r} y_i}{\sum_{i=1}^{r} a_i} \quad (2.7)$$

3. Material and Methods

Estimating the Intercept and Slope in A Linear Model

The estimation of the intercept is considered for a model with intercept and slope, while for a linear model without intercept, we estimate the slope.

The linear regression given as

$$y_i = M_0 + M_1 x_i + e_i \quad (3.1)$$

Is a simple linear regression model, where y_i is the response variable also known as the dependent variable whose values or outcome dependents on the input variable or the explanatory variable(s).

x_i is the input variable or the predictor variable also known as the independent variable whose values give rise to the estimated value of the response. It gives some level of interpretation on the response y_i.

M_0 and M_1 are the parameters of the regression model. Particularly M_0 is the y-intercept of the model and M_1 is the coefficient of the response variable x_i, it is the gradient of the regression model.

e_i is the measure of the stochastic or random error that may occur in the system under study.

The multiple linear regression model is as seen in Kutner et al (2005) and Nwagozie (2011) is given as

$$y_i = M_0 + M_1 x_{i1} + M_2 x_{i2} + e_i \tag{3.2}$$

and this is called first order model having two explanatory variables. One of the difference between the simple linear regression and the multiple linear regression is that, the linear regression given as

$$E(Y) = M_0 + M_1 x$$

is a real line while the multiple linear regression given as

$$E(Y) = M_0 + M_1 x_1 + M_2 x_2 \tag{3.3}$$

is a plane as seen in Kutner et al (2005) and Sameera (2014)

A first order model with more than one explanatory variable is generally given as

$$y_i = M_0 + M_1 x_{i1} + M_2 x_{i2} + \cdots + M_{n-1} x_{i,n-1} + e_i \tag{3.4}$$

Linear and noon linear model

As seen in Bremer (2012), Gujarati, (2004), Shalabh, (2017) and Kanpur (2001) it has been a clear knowledge that general regression model is not limited to only linear surfaces. The term Linear simply means that the model is linear in parameters but May or may not be linear in the explanatory variables. The models of the form are all linear models.

$$y = m_0 + m_1 x + e$$

$$y = m_0 + m_1 x + m_2 x^2 + e$$

$$y = m_0 + m_1 x + m_2 x^2 + m_3 x^3 + e$$

While the models

$$y = m_0 + m_1^2 x + e$$

$$y = m_0 + m_1 x^2 + m_2^2 x^3 + e$$

$$y = m_0 + m_1 m_2 x + m_3 x^2$$

Are non-linear model because they are not linear in parameters see also Kutner *et al* (2005).

Least square estimation of parameters for model with or without intercept term and proof

In the research done by Sameera (2014), for model without intercept term, to obtain the estimates of the slope parameters, the model was built and least square method was applied to obtain

$$\hat{\beta}_1 = \frac{\sum x_i y_i}{\sum x_i^2}$$

p-ISSN: 2348-6848
e-ISSN: 2348-795X
Vol. 8 Issue 6
June 2021

In this work, having looked in word, we saw the need to reduce the stress in computing $\sum x_i y_i$ and also$\sum x_i^2$. This will be equivalent to substituting β_0=0 in 2.3b to obtain the $\hat{\beta}_1$ which have been applied in this research work and the result obtained was equal to the result obtained when we used $\hat{\beta}_1 = \frac{\sum x_i y_i}{\sum x_i^2}$, there was no loss of information.

Lemma

Given a linear model without intercept term which is a reduced model, applying the least square method in obtaining the slope of the regression line, is equal or approximately equal to substituting the intercept term as zero in (2.3b) to obtain the slope.

Proof:

Let m_0 be the intercept term in a model with intercept and let m_1 be the slope, let x be the predictor or explanatory variable and let y be the response variable predicted by x.

Given: The linear model with intercept term

$$y = m_0 + m_1 x + e$$

And the model without intercept term as

$$y = m_1 x + e$$

We want to show that, is not a mere cancelation of the variables in the numerator

$\frac{\sum xy}{\sum x^2} = \frac{\sum y}{\sum x}$, the theorem is aimed at revealing that without going back to least square method to obtain slope, you can simply substitute the gradient as zero in the already known least square formula containing intercept and slope and hence solve for slope directly. i.e

From

$$y_i = m_1 x_i + e$$

applying least square estimate, to obtain the estimate of m_1 which is the slope, we have;

$$m_1 = \frac{\sum xiy}{\sum xi^2}$$

From

$y_i = m_0 + m_1 + e$ applying least square to obtain the intercept m_0, we have

$$m_0 = \frac{\sum y_i}{n} - m_1 \frac{\sum x_i}{n}$$

Letting $m_0 = 0$, we have

$$0 = \sum y_i - m_1 \sum xi$$

p-ISSN: 2348-6848
e-ISSN: 2348-795X
Vol. 8 Issue 6
June 2021

Then we have

$\sum yi = m_1 \sum xi$

$\therefore m_1 = \frac{\sum yi}{\sum xi}$

If you multiply the numerator and denominate by $\sum xi$, we have $m_1 = \frac{\sum xiyi}{\sum xi^2}$, hence yielding the same result with the least square method.

Coefficient of correlation and Determination

According to kutner *et al.* (2005) and Nwagaozie (2011), the coefficient of correlation is defined as the measure of the degree of linear relationships or association that exists between the response variable Y and the predictor variable X provided the two variables are random in nature. It is the square root of the coefficient of determination attached with plus or minus sign. It is given as

$$r = \pm\sqrt{R^2} \tag{3.5}$$

The value of $r = +1\ or - 1$ is attached in consideration of whether the slope of the model so far fitted is positive or negative. If the slope M_1 is positive then the value of $r = +1$, if the slope is negative, the value of $r = -1$. The coefficient of correlation has the range $-1 \leq r \leq 1$.

Coefficient of Determination R^2 is the proportionate decrease in the total variation associated with the explanatory variable X. The higher the value of R^2, the higher the total variation in Y variable is reduced. It is given as

$$R^2 = \frac{SSR}{SSTotal} \tag{3.6}$$

$$=1 - \frac{SSE}{SSTotal}$$

Coefficient of Determination has the range $0 \leq R^2 \leq 1$

Estimating the intercept and slope for a simple linear model in one variable

For a model with Intercept and slope, as given in (1.1) expressed in some well-organized literature like, Sameera (2014), Ijomah and Wali (2017) and Kutner *et al* (2005)

The model in (3.1) is given as:

$y_i = m_0 + m_1x_i + e_i$

In obtaining the parameters of the linear model in (3.1), we apply

(1) Matrix notation
(2) Least square estimation

p-ISSN: 2348-6848
e-ISSN: 2348-795X
Vol. 8 Issue 6
June 2021

Matrix notation

From the model in, (3.1) we obtain the matrix X using the column of the intercept m_0 as the column of 1's and the column of x_i's will be replaced by the values of x used in this research. The response values will be written in the column of the y_i, them, the model will be re-written as:

$\hat{y}_i = \hat{m}_0 + \hat{m}_1 x_i$

$y_i = (y_1, y_2 \dots, y_n)'$

$\hat{y}_i = (\hat{y}_1, \hat{y}_2, \dots, \hat{y}_n)'$

And

$x_i = (x_1, x_2, \dots x_n)'$

$m_0 = (1, 1, \dots 1)'$

This can be written in matrix form as;

$$\hat{y} = (y_1, y_2, \dots, y_n)' = \begin{pmatrix} 1 & x_1 \\ 1 & x_2 \\ \vdots & \vdots \\ . & . \\ 1 & x_n \end{pmatrix} \tag{3.7}$$

$$\therefore X_{n\times 2} = \begin{pmatrix} 1 & x_1 \\ 1 & x_2 \\ \vdots & \vdots \\ . & . \\ 1 & x_n \end{pmatrix}$$

We obtain the transpose of x, written as X', which can be given as:

$$X'_{n\times 2} = \begin{pmatrix} 1 & 1 & . & . & . & 1 \\ x_1 & x_2 & . & . & . & x_n \end{pmatrix}$$

Making the first column the first row and the second column the second row.

$X_{n\times 2}$ is change to $X_{2\times n}$

The hat matrix of the design is given as

$X'X=$

$$\begin{pmatrix} 1 & 1 & . & . & . & 1 \\ x_1 & x_2 & . & . & . & x_n \end{pmatrix} \begin{pmatrix} 1 & x_1 \\ 1 & x_2 \\ . & . \\ . & . \\ . & . \\ 1 & x_n \end{pmatrix} \quad (3.8)$$

$$= \begin{pmatrix} N & \sum_{i=1}^{n} x_i \\ \sum_{i=1}^{n} x_i & \sum_{i=1}^{n} x_i^2 \end{pmatrix}$$

We obtain $X'Y$ as shown

$X'Y$ =

$$\begin{pmatrix} 1 & 1 & . & . & . & 1 \\ x_1 & x_2 & . & . & . & x_n \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ . \\ . \\ . \\ y_n \end{pmatrix}$$

p-ISSN: 2348-6848
e-ISSN: 2348-795X
Vol. 8 Issue 6
June 2021

$$\begin{pmatrix} \sum_{i=1}^{n} x_i \\ \sum_{i=1}^{n} x_i y_i \end{pmatrix}$$

We use the least square equation as expressed in Iwundu and Onu (2017), Sameera, (2014), Kutner *et al.* (2005), and Huaglin & Welsch (1978).

$$\underline{\hat{m}} = (X'X)^{-1}(X'Y) \qquad (3.9)$$

We obtain the determinant of $(X'X)^{-1}$ as shown;

$$(X'X)^{-1} = \frac{AdJ(X'X)}{|X'X|} \qquad (3.10)$$

but Adj$(X'X)$= transpose of the matrix of co factors of $X'X$, but the $(X'X)^{-1}$ is a little bit easier for a 2×2 matrix i.e

$$\text{if } (X'X) = \begin{pmatrix} a_{11} & a_{12} \\ a_{12} & a_{22} \end{pmatrix}$$

Determinant of $(X'X)$ given as:

$|X'X| = a_{11}a_{12} - a_{12}a_{12}$ for symmetric matrix if not symmetric it will be

$|X'X| = a_{11}a_{22} - a_{12}a_{21}$, but here we are interest in symmetric matrix, in fact, multiplication of a matrix X' and it's corresponding matrix X makes it symmetric.

:. Adj$(X'X)$ for 2×2 is obtain by interchanging a_{11} and a_{12} and making a_{12} negative, i.e.

$$\text{Adj } (X'X) = \begin{pmatrix} a_{22} & -a_{12} \\ -a_{12} & a_{11} \end{pmatrix}$$

$$(X'X)^{-1} = \frac{1}{a_{11}a_{22}-a_{12}a_{12}} \begin{pmatrix} a_{22} & -a_{12} \\ -a_{12} & a_{11} \end{pmatrix} \qquad (3.11)$$

See Cohen et al (2003), Helsel and Hirsch (1991) and Iwundu and Onu (2017), for application of the alias matrix, and it is seen as;

$$\underline{\hat{m}} = \begin{pmatrix} \hat{m}_0 \\ \hat{m}_1 \\ \hat{m}_n \end{pmatrix} = \frac{1}{a_{11}a_{22}-a_{12}a_{12}} \begin{pmatrix} a_{22} & a_{12} \\ -a_{12} & a_{11} \end{pmatrix} \begin{pmatrix} \sum \mathrm{y_n} \\ \sum \mathrm{x_n y_n} \end{pmatrix}$$

p-ISSN: 2348-6848
e-ISSN: 2348-795X
Vol. 8 Issue 6
June 2021

$$\begin{pmatrix}\hat{m}_0\\ \hat{m}_1\end{pmatrix} = \begin{pmatrix}\frac{a_{22}\sum \mathrm{y_n} - a_{12}\sum \mathrm{x_n y_n}}{a_{11}\,a_{22} - a_{12}a_{12}}\\ \\ \frac{-a_{22}\sum \mathrm{y_n} + a_{11}\sum \mathrm{x_n y_n}}{a_{11}\,a_{22} - a_{12}a_{12}}\end{pmatrix}$$

$\hat{m}_0 = \frac{a_{22}\sum \mathrm{y_n} - a_{12}\sum \mathrm{x_n y_n}}{a_{11}\,a_{22} - a_{12}a_{12}}$ is the intercept term (3.12)

$$\hat{m}_1 = \frac{a_{11}\sum \mathrm{y_n} - a_{12}\sum \mathrm{y_n}}{a_{11}\,a_{22} - a_{12}a_{12}} \quad (3.13)$$

is the slope

The outlined approaches shall be employed in obtaining the estimate of the model parameters for a simple linear regression model, the approaches are applied in modeling the blood pressure of 100 persons as sampled for response variables over the explanatory variables x ranging from the height, weight, age, sugar level and sex for the same 100 persons. In order to maintain a simple linear regression analysis and for the purpose of studying the joint effect of all those variables on the response y, we added the values of each of them to obtain a single value and this is done for all the 100 persons. The summation of all these variables is referred to as pooled variables. The research also considers the contribution of sugar level as a predictor variable on the response y (Blood pressure). The study considers both model with or without intercept, in each case, the approaches outlined above are followed.

Least Square Estimation Method

The method of least square requires re-arrangement of the model that is to say, the model in (1.1) is re-arranged as;

$$\sum e_i^2 = (\sum y_i - m_0 - m_1 x_i)^2 \quad (3.14)$$

To obtain the estimate of the intercept term, m_0

to obtained the estimate of m_0 is as seen below:

$$\therefore \frac{\partial \sum e_i^2}{\partial m_0} = -2\sum(y_i - m_0 - m_1 x_i) = 0$$

Divide through by -2, we have

$$\sum(y_i - m_0 - m_1 x_i) = 0$$

Open the bracket in the L.H.S, we have:

$$\therefore \sum y_i - n m_0 - m_1 \sum x_i = 0$$

p-ISSN: 2348-6848
e-ISSN: 2348-795X
Vol. 8 Issue 6
June 2021

$nm_0 = \sum x_i - m_1 \sum x_i$

$\therefore m_0 = \frac{\sum y_i - m_1 \sum x_i}{n}$

$\widehat{m}_0 = \frac{\sum y_i}{n} = m_1 \frac{\sum x_i}{n}$

But $\sum y_i/_n = \bar{y}_i$ and $\frac{\sum xi}{n} = \bar{x}_i$

$$\therefore \widehat{m}_0 = \bar{y}_i - m_1 \bar{x} \quad (3.15)$$

Also, to obtain m_1 i.e the slope, we differentiate (3.14) with respect to m_1, as seen;

$\frac{\partial \sum e_i^2}{\partial m_1} = -2 \sum (y_i - m_0 - m_1 x_i)(x_i) = 0$

$\sum (y_i - m_0 - m_1 x_i)(x_i) = 0$

Open the bracket in the L.H.S, we have

$\sum x_i y_i - m_0 \sum x_i - m_1 \sum x_i{}^2 = 0$

$\sum x_i y_i - m_0 \sum x_i = m_1 \sum x_i{}^2$

$$m_1 = \frac{\sum x_i y_i - m_0 \sum x_i}{\sum x_i{}^2} \quad (3.16)$$

These are for linear model with intercept term. For model without intercept term the least square method is expressed in chapter 2 of this work.

Estimating the intercept and the slope of a model with two explanatory variables with or without intercept term

This approach follows from 3.1.1, but since the model have been added another predictor variable, the model becomes as seen:

$$y_i = m_0 + m_1 x_1 + m_2 x_2 + e_i \quad (3.17)$$

Re-arranging and squaring both side, we have,

$\sum e_i^2 = \sum (y_i - m_0 - m_1 x_1 - m_2 x_2)^2$

Differentiating w.r.t. m_0, we have

$\frac{\sum e_i^2}{\partial m_0} = -2 \sum (y_i - m_0 - m_1 x_1 - m_2 x_2) = 0$

$n\widehat{m}_0 = \sum y_i - m_1 \sum x_1 - m_2 \sum x_2$

$$\widehat{m}_0 = \frac{\sum y_i - m_1 \sum x_1 - m_2 \sum x_2}{n} \quad (3.18)$$

where m_1 and m_2 are the slopes

The x_1 and x_2 respectively.

Differentiating w.r.t m_1 we have;

$$\frac{\partial \sum e_i^2}{2m_1} = -2\sum(y_i - m_0 - m_1x_1 - m_2x_2) = 0$$

$$\sum x_i y_i - m_0 \sum x_i - m_1 \sum x_i{}^2 - m_2 \sum x_i \, x_2 = 0$$

$$m_1 \sum x_i{}^2 = \sum x_1 \, y_i - m_0 \sum \mathrm{x}_1 - m_2 \sum \mathrm{x}_1 \, \mathrm{x}_2$$

$$\widehat{m}_1 = \frac{\sum x_1 y_i - m_0 \sum x_1 - m_2 \sum x_1 x_1}{\sum x_1^2} \tag{3.19}$$

Also

$$\widehat{m}_2 = \frac{\sum x_1 y_i - m_0 \sum x_2 - m_1 \sum x_1 x_2}{\sum x_2^2} \tag{3.20}$$

Following from the above, the parameters for a multiple linear model can be obtained.

If for;

$$\widehat{m}_0 = \frac{\sum y_i - m_1 \sum x_1 - m_2 \sum x_2 \_\_ \ldots \_\_ m_n \sum x_n}{n} \tag{3.21}$$

$$\widehat{m}_1 = \frac{\sum_{i=1}^{n} x_i y_{i-} m_0 \sum_{i=1}^{n} x_i - m_2 - \sum_{i=1}^{n} x_i y_i - \ldots m_n \sum_{i=1}^{n} x_{n-1} x_n}{\sum_{i=1}^{n} x_n^2} \tag{3.22}$$

.

.

.

$$\widehat{m}_n = \frac{\sum_{i=1}^{n} x_i y_{i-} m_0 \sum_{i=1}^{n} x_i - m_2 - \sum_{i=1}^{n} x_i x_{i+1} - \ldots m_{n-1} \sum_{i=1}^{n} x_{n-1} x_n}{\sum_{i=1}^{n} x_n^2}$$

Application of analysis of variance (ANOVA).

Analysis of variance popularly known as ANOVA is applied to both the model with intercept and that without intercept. We obtain the sum of squares of the regression, between treatment, error sum of square and the sum of square total, see Keller and Warrack (2003). Sum of square treatment is the test statistic that is used to measure the similarities of the mean samples to each other. It is given as;

$$SS_{Treat} = \sum_{i=1} n_i \, (\bar{x}_i - \bar{\bar{x}})^2 \tag{3.23}$$

p-ISSN: 2348-6848
e-ISSN: 2348-795X
Vol. 8 Issue 6
June 2021

If a large difference is experienced in the between treatment means known as the sum of square treatment, it means that one and above sample means will considerably differ from the grand mean as see in Keller and Warrack (2003) in order to know whether or not to reject the null hypothesis, it is advisable to know how much variation that exist within treatments variation, and this in order word called sum of square error denoted as SSE. It is given as;

$$SSE = \sum_{j=1}^{n}\sum_{i=1}^{n}\left(x_{ij} - \bar{x}_j\right)^2 \quad (3.24)$$

Which can also be written by expansion as;

$SSE = (n_1 - 1)S_1^2 + (n_2 - 1)S_2^2 . \; . \; . + (n_K - 1)S_k^2$, this is as expressed in Keller and Warract (2003), Nwaogazie (2011) and Egbule (2008). We proceed to computing the mean squares, for which mean square for treatment is obtain as;

$MS_{treat} = \frac{SS_{treat}}{n-1}$, that is to say, the sum of square treatment is divided by the number of treatments in the sample minus 1.

While mean square error $= \frac{SSE}{\mathrm{N}-n}$ (3.25)

Where N is the total sample and n is the number of treatments.

We finally compute the F statistic given as

$$F = \frac{MS_{treat}}{MSE} \quad (3.26)$$

The hypothesis is built and conclusion drown as shown;

$H_0: m_0 = 0$

$H_1: m_0 \neq 0$ (3.27)

The Fstatistic tell us whether the value of SS_{treat} is large such that the null hypothesis can be rejected.

We reject null hypothesis if $F > F\alpha, n-1, \mu - n$ (3.28)

t test statistic was applied when interest shifted from observing whether their means μ_1 equal μ_2 or not to testing if μ_1 is greater than μ_2 and verse versa.

The t statistic used in this research is given as;

$$t = \frac{\hat{m}_0}{S(\hat{m}_0)} \quad (3.29)$$

Where $\hat{m}_0$ is the estimate of the intercept term in a model with intercept and $S(\hat{m}_0)$ is the standard deviation of the intercept term. For slope term $\hat{m}_1$, the T statistic is given as;

$t = \frac{\hat{m}_1}{S(\hat{m}_1)}$ see Kutner et al (2005)

But $S(\widehat{m}_0) = MSE\ \left[\frac{1}{n} + \frac{\bar{x}^2}{\sum(xi-\bar{x})^2}\right]$ (3.30)

According to Kutner et al (2005).

$SSE = \sum(y_i - \hat{y}_1)^2$

$SStotal = \sum(y_i - \bar{y})^2$

And

$SSR = \sum(\hat{y}_1 - \bar{y})^2$

And

$MSE_{Error} = \frac{\sum(y_i-\hat{y}_{i)}}{n-2} = \frac{SSE}{n-2}$

$MSE_{Reg} = \frac{\sum\widehat{(y_1}-\bar{y}_{i)}}{n-2} = \frac{SSE}{1} = SSR$

$MSE_{totla} = \frac{\sum(y_i-\bar{y})^2}{n-1} = \frac{SSE_{total}}{n-1}$

A typical example of a one way ANOVA is as shown in table 3.1

Table 3.1: A clear ANOVA table containing $E(MS)$ one way

Source of Variation	**Df**	**SS**	**MS**	**E-(MS)**
Regression (Treat)	1	$SSR = \sum(\hat{y}_i - \bar{y})^2$	$SS_R = \frac{SSR}{1}$	$\delta^2 + m_1^2 \sum(x_i - \bar{x})^2$
Error	$n-2$	$SSE = \sum(y_i - \hat{y}_i)^2$	$SSE = \frac{SSE}{n-2}$	δ^2
Total	$\boldsymbol{n-2}$	$\boldsymbol{SS_{Total} = \sum(y_i - \bar{y})^2}$		

The clearer and more simple table for ANOVA is presented in table 3.2 as seen below;

Table 3.2: One Way ANOVA

Source of Variation	**Df**	**SS**	**MS**	**Fcal**
Treatment (B/W)	(k-1)	SS_{treat}	MS_{treat}	$\frac{MS_{treat}}{MSE}$
Error (within)	$(\mu - k)$	SSE	MSE	
Total	$(\mu - k)$	$\boldsymbol{SST}$		

p-ISSN: 2348-6848
e-ISSN: 2348-795X
Vol. 8 Issue 6
June 2021

Results and Discussion

4. Investigating the omission of intercept term using pooled explanatory variables

The computations for pooled explanatory variables are presented in table 1

Investigating the omission of intercept term using single variable $x =$ sugar level without compressing the variables

The computations for sugar level are presented in table 2

Investigating the effect of the omission of the intercept term for two explanatory variables (age and sugar level).

The computations for age and sugar level are presented in table 3

testing of hypothesis by evaluating the standard deviation of the intercept term

The computations for the testing of hypothesis are presented in table 4

Application of Coefficient of Correlation and Determination

The computations for coefficient of correlation and determination are presented in table 5

Table 1:

(a) Result summary for model with intercept for blood pressure and pooled variable

Intercept(m_0)	Slope(m_1)	$\sum(y-\hat{y})$
163.88	-0.06	7.78

(b) Result summary for model without intercept for blood pressure and pooled variable

Slope(m_1)	$\sum(y-\hat{y})$
0.308	1.72

Table 2:

(a) Result summary for model with intercept for blood pressure and sugar level

Intercept(m_0)	Slope(m_1)	$\sum(y-\hat{y})$
81.40	0.324	16.84

(b) Result summary for model without intercept for blood pressure and sugar level

Slope(m_1)	$\sum(y-\hat{y})$
0.806	0.734

Table 3:

(a) Result summary for model with intercept for blood pressure and two predictor variables

Intercept(m_0)	Slope(m_1)	Slope(m_2)	$\sum(y-\hat{y})$
55.778	**0.235**	**0.202**	$\mathbf{6.5\times 10^2}$

(b) Result summary for model without intercept for blood pressure and two predictor variables

Slope(m_1)	Slope(m_2)	$\sum(y-\hat{y})$
0.49	**1.25**	**-25.14**

Table 4:

(a)

ONE WAY ANOVA FOR MODEL WITH INTERCEPT FOR COMPRESSED EXPLANATORY VARIABLES

```
ONEWAY y BY x
 /POLYNOMIAL=1
 /STATISTICS EFFECTS
 /PLOT MEANS
 /MISSING ANALYSIS.
```

Oneway

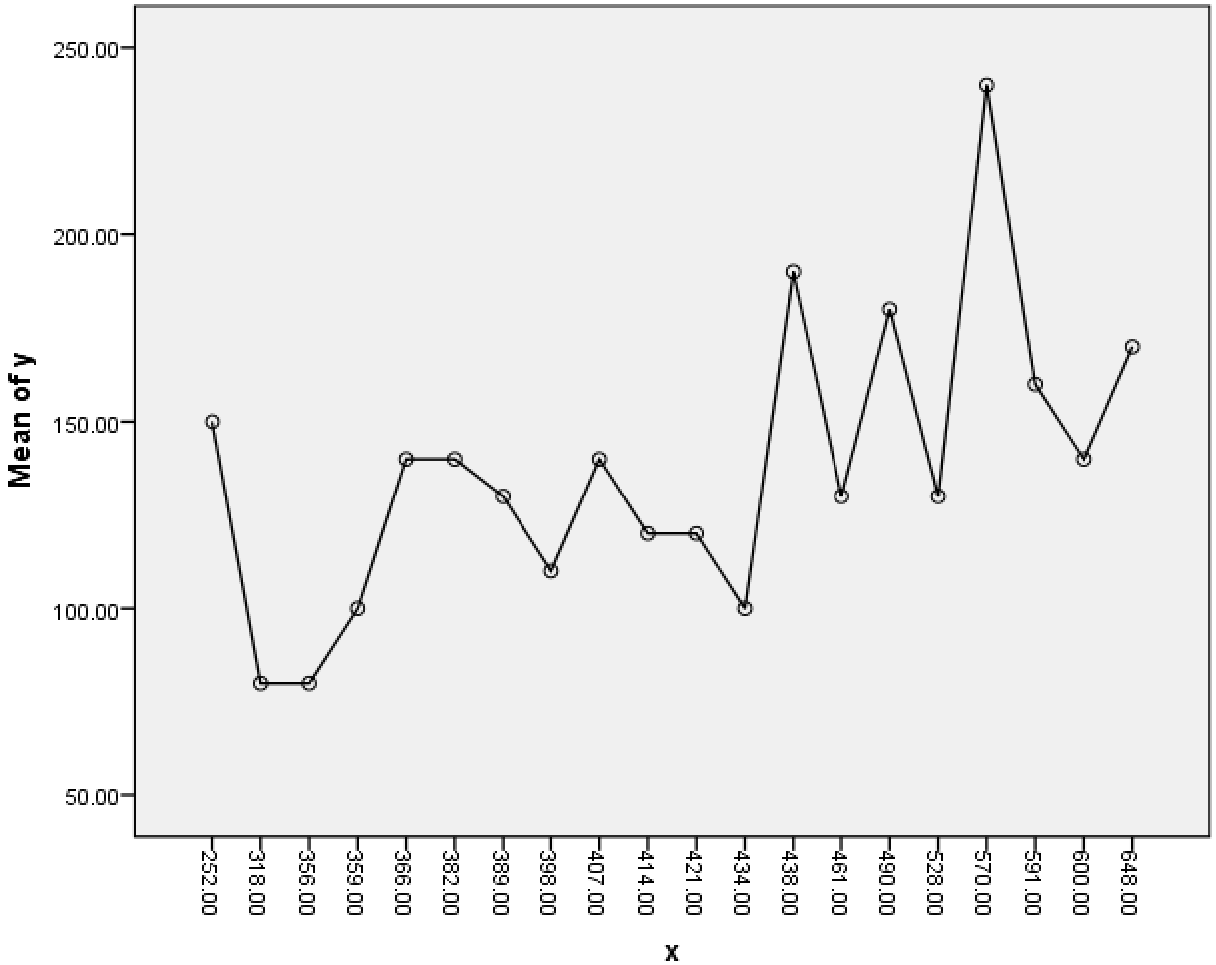
ANOVA

y

			Sum of Squares	Df	Mean Square	F	Sig.
Between Groups	(Combined)		27775.000	19	1461.842	1.33.	.
	Linear Term	Contrast	7964.350	1	7964.350	7.24.	
		Deviation	19810.650	18	1100.592	.	
Within Groups			.000	0	.		
Total			27775.000	19			

p-ISSN: 2348-6848
e-ISSN: 2348-795X
Vol. 8 Issue 6
June 2021

Means Plots



p-ISSN: 2348-6848
e-ISSN: 2348-795X
Vol. 8 Issue 6
June 2021

(b)
ONE WAY ANOVA FOR MODEL WITH INTERCEPT FOR UNCOMPRESSED EXPLANATORY VARIABLES

```
ONEWAY y BY x
  /POLYNOMIAL=1
  /STATISTICS EFFECTS
  /PLOT MEANS
  /MISSING ANALYSIS.
```

Oneway

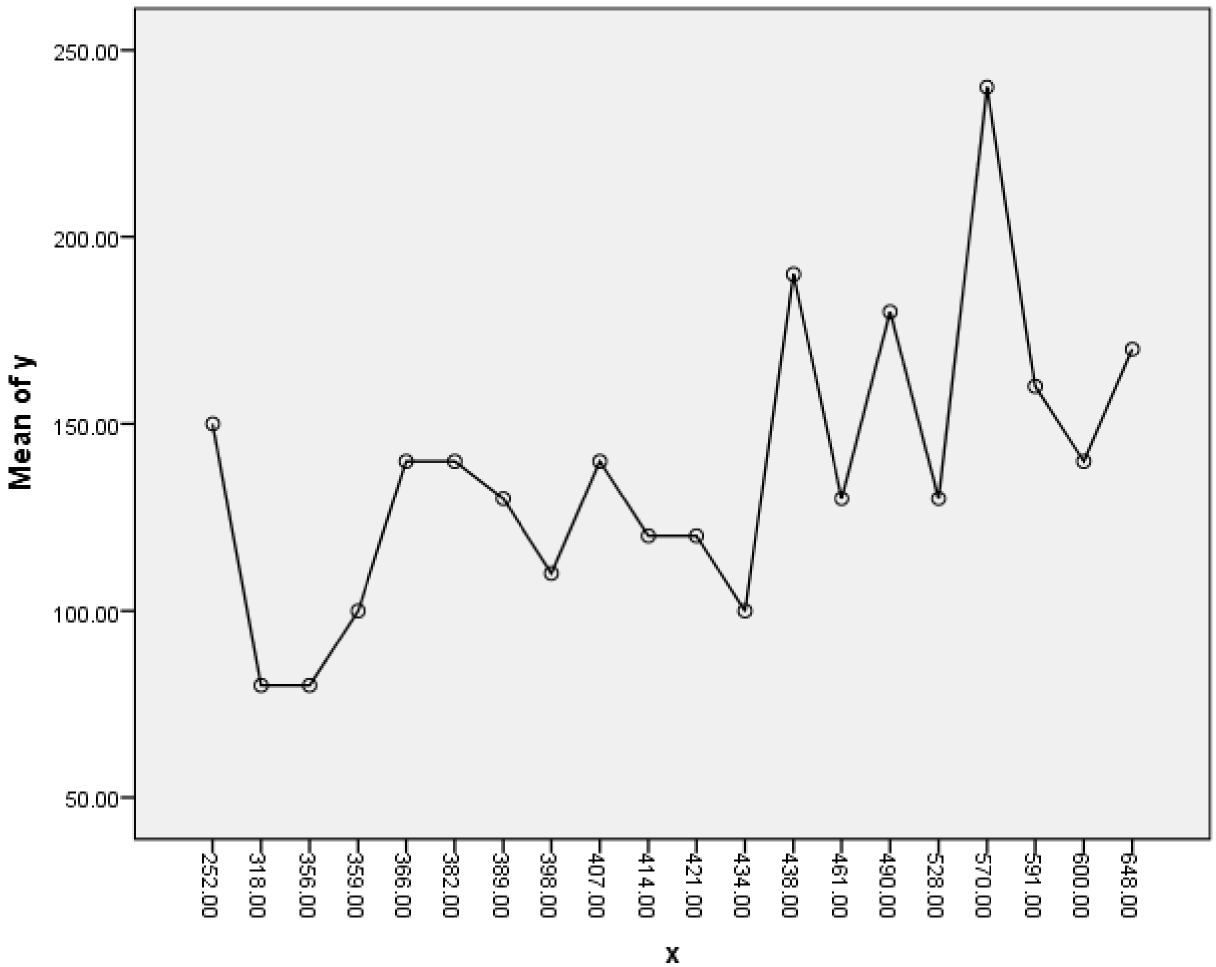
ANOVA

y

			Sum of Squares	df	Mean Square	F	Sig.
Between Groups	(Combined)		27775.000	19	1461.842	.	.
	Linear Term	Contrast	7964.350	1	7964.350	.	
		Deviation	19810.650	18	1100.592	.	
Within Groups			.000	0	.		
Total			27775.000	19			

Means Plots

p-ISSN: 2348-6848
e-ISSN: 2348-795X
Vol. 8 Issue 6
June 2021



International Journal of Research (IJR)
p-ISSN: 2348-6848
e-ISSN: 2348-795X
Vol. 8 Issue 6
June 2021

(c)

```
ONEWAY y BY x
  /POLYNOMIAL=1
  /STATISTICS EFFECTS
  /PLOT MEANS
  /MISSING ANALYSIS.
```

Oneway

[DataSet0]

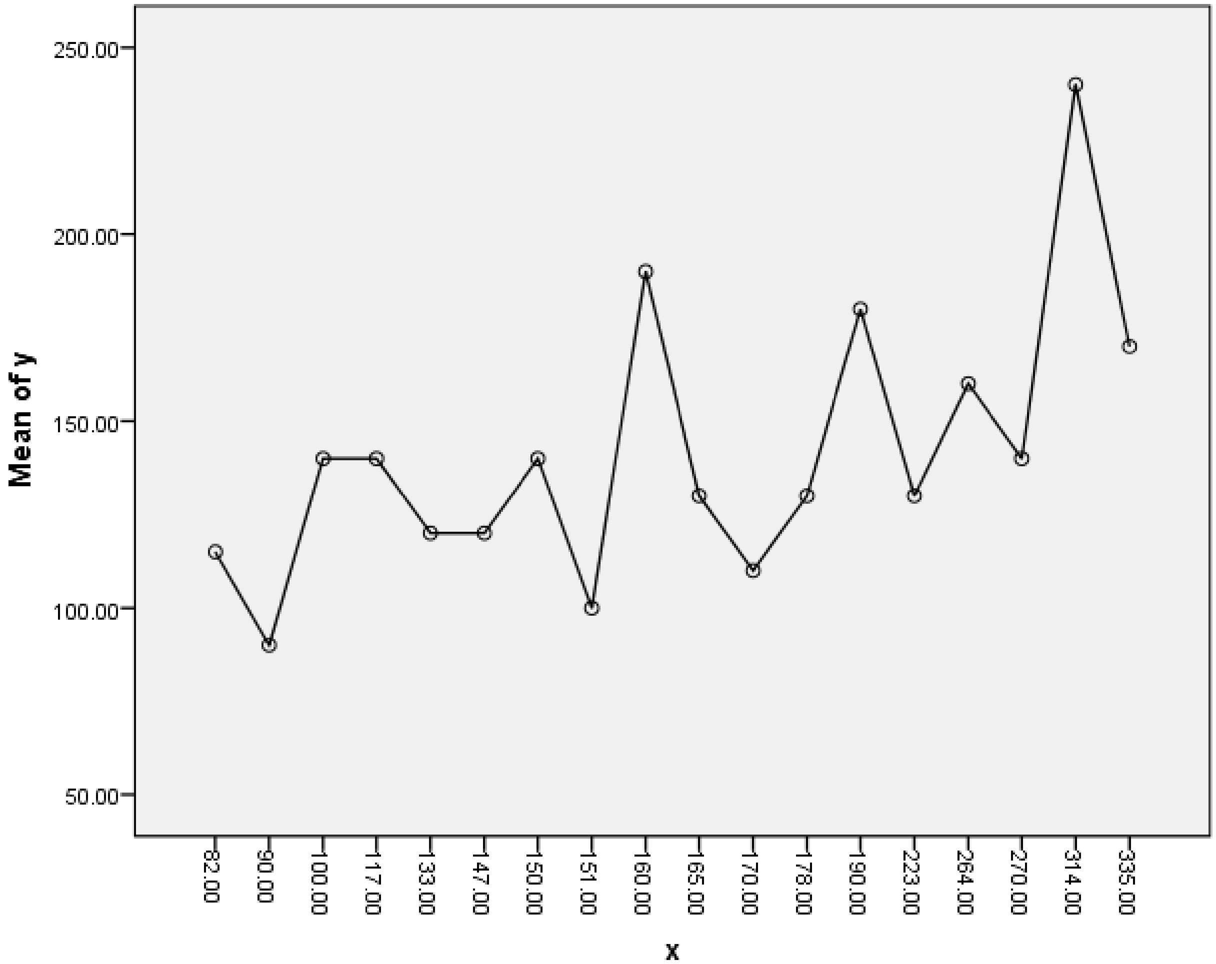
ANOVA

y

			Sum of Squares	df	Mean Square	F	Sig.
Between Groups	(Combined)		25125.000	17	1477.941	1.115	.574
	Linear Term	Weighted	11485.630	1	11485.630	8.668	.099
		Deviation	13639.370	16	852.461	.643	.758
Within Groups			2650.000	2	1325.000		
Total			27775.000	19			

Means Plots

p-ISSN: 2348-6848
e-ISSN: 2348-795X
Vol. 8 Issue 6
June 2021



p-ISSN: 2348-6848
e-ISSN: 2348-795X
Vol. 8 Issue 6
June 2021

(d)
ANALYSIS OF VARIANCE FOR INTERCEPT MODEL WITH TWO PREDICTOR VARIABLES

```
UNIANOVA y WITH X2 X1
  /METHOD=SSTYPE(1)
  /INTERCEPT=INCLUDE
  /EMMEANS=TABLES(OVERALL) WITH(X2=MEAN X1=MEAN)
  /CRITERIA=ALPHA(.05)
  /DESIGN=X2 X1.
```

Analysis of Variance for model with two predictor variables

Tests of Between-Subjects Effects

Dependent Variable: y

Source	Type I Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	11873.537[a]	2	5936.769	6.347	.009
Intercept	378125.000	1	378125.000	404.247	.000
X2	1535.620	1	1535.620	1.642	.217
X1	10337.917	1	10337.917	11.052	.004
Error	15901.463	17	935.380		
Total	405900.000	20			
Corrected Total	27775.000	19			

a. R Squared = .427 (Adjusted R Squared = .360)

p-ISSN: 2348-6848
e-ISSN: 2348-795X
Vol. 8 Issue 6
June 2021

(e)

ANALYSIS OF VARIANCE FOR MODEL WITHOUT INTERCEPT FOR TWO PREDICTOR VARIABLES

```
UNIANOVA y WITH X2 X1
  /METHOD=SSTYPE(1)
  /INTERCEPT=EXCLUDE
  /EMMEANS=TABLES(OVERALL) WITH(X2=MEAN X1=MEAN)
  /CRITERIA=ALPHA(.05)
  /DESIGN=X2 X1.
```

Tests of Between-Subjects Effects

Dependent Variable: y

Source	Type I Sum of Squares	df	Mean Square	F	Sig.
Model	381142.627[a]	2	190571.313	138.556	.000
X2	348350.626	1	348350.626	253.270	.000
X1	32792.001	1	32792.001	23.842	.000
Error	24757.373	18	1375.410		
Total	405900.000	20			

Squared = .939 (Adjusted R Squared = .932)

. R Squared = .939 (Adjusted R Squared = .932)

Table 5: summary of result for coefficient of correlation and determination

(a) Model with intercept for pooled variable

R	R^2
-0.536	**0.287**

(b) Model with intercept for sugar level

R	R^2
0.664	**0.441**

(c) model with intercept for two predictor variables (age and sugar level)

r_{x_1}	r_{x_2}	$R^2_{x_1}$	$R^2_{x_2}$	$R^2_{x_1x_2}$	$r_{x_1x_2}$
0.610	**0.235**	**0.372**	**0.055**	**0.427**	**0.654**

(d)) model without intercept for two predictor variables (age and sugar level)

p-ISSN: 2348-6848
e-ISSN: 2348-795X
Vol. 8 Issue 6
June 2021

r_{x_1}	r_{x_2}	$R^2_{x_1}$	$R^2_{x_2}$
0.926	**0.858**	**0.284**	**0.081**

4.6 Discussion of Result

For model with intercept with pooled x variables for a simple linear regression, it was discovered that the intercept term is positive and is equal to 163.88 which signifies a 50.12 reduction in the value of the intercept obtained by Sameera (2014) when a multiple linear regression of all the variables were considered. It was also discovered that the slope of the predictor variable (sugar level) is equal to -0.06, while that of the Sameera (2014) was estimated at 0.0839. the coefficient of correlation for -0.06 is negative and that of 0.0839 is positive. This shows that the combination of all the explanatory variables in a simple linear regression gives a contracting or opposite result to that obtained when the variables are fitted using a multiple linear regression. For model without intercept, using the pooled x variables for a simple linear regression model, it was observed that the slope of the data without intercept was 0.308 and that obtained by Sameera (2014) with a multiple linear regression for sugar level was 0.0850 both being positive. This shows that the estimation of the slope for a simple linear regression with pooled variables has better slope than that obtained by Sameera (2014) for multiple linear regression without intercept. It was observed that by omitting the intercept term, the value of the slope increases to 0.308 from -0.06 which shows a 24.4 percent increase in the slope.

For model with intercept for sugar level x for a simple linear regression, the intercept is estimated as 81.40 and the slope was 0.324, while for model without intercept for sugar level, the slope is 0.806, this also shows that by omitting the intercept term, the value of the slope increases to 0.806 from 0.324 showing a 48.2 percent increase in the slope. For model with intercept for sugar level and ages of the 20 persons, it was discovered that for model with two explanatory variables sugar level and age, the intercept is estimated at 55.778 and the slope for age is 0.235 and that of sugar level is estimated at 0.202, while for Sameera (2014), the intercept was estimated at 214 and the slope for age was 1.39 and slope for sugar level was 0.0839. this shows that the slope for age is greater than the slope for sugar level for both Sameera (2014) and this current work. This means that the age of a person contributes more to the person's blood pressure than the sugar level. But sugar level is seen contributing more than the age on the basis of coefficient of correlation and determination for model with intercept also, the opposite is the case for model without intercept, where the age out performs the sugar level on the basis of coefficient of correlation and determination. This generally reveals that, the model without intercept is opposite in performance to the model with intercept.

For model without intercept for in two explanatory variables (sugar level and age). The slope for age from Sameera (2014) is greater than that obtained in this research with about 20 percent increase while the slope for sugar level from this research is seen greater than the slope for sugar level from Sameera (2014) with about 40.95 percent increase. It is seen here that the age has higher slope than that of sugar level which have been interpreted above.

By testing hypothesis, it was observed that the model with intercept is significant, which means that, it is advisable to build a model with intercept term when modeling the blood pressure of

p-ISSN: 2348-6848
e-ISSN: 2348-795X
Vol. 8 Issue 6
June 2021

people with their sugar level and age. This is also true for the model with intercept term for blood pressure against the sugar level. Whereas for model without intercept, it is discovered that the intercept term is insignificant in modeling the blood pressure against the explanatory variables. That is to say, the exclusion of the intercept term forced the model through the origin. The analysis revealed that the data used for sugar level is better than the pooled data in fitting the blood pressure of human beings, this means that the addition of the explanatory variables to get a single explanatory variable amounts to loss of information. This was revealed by the value of mean square error of 1100.592 for pooled variable and 1325 for single variable which is sugar level. This clearly explains that it is better to conduct a simple linear regression for each of the explanatory variables under study against the response y, than to add all the explanatory variables to make then a single variable in order to use simple linear model, especially when multiple linear regression is not to be used for one reason or the other. When intercept is removed from a model, it was discovered that the mean square error increases than when intercept is not removed. The higher the mean square error, the smaller the F value. This research also proves both theoretically and numerically that obtaining least square estimate for the slope m_1 in a model without intercept yields the same result with that, when you substitute intercept $m_0 = 0$ in a least square estimate of the intercept in a model with intercept. Finally, as revealed in the plots for the blood pressure against the pooled explanatory variables in Appendix G, the blood pressure against the sugar level in Appendix H and the blood pressure against two explanatory variables (age, and sugar level) in Appendix I, the relationships between the response which is the blood pressure and the explanatory variables are not stable. This could be

Conclusion

The effect of the omission of intercept term in a linear model was considered using least square estimate and matrix approach in estimating intercept and slope for model with intercept and for estimating only slope for model without intercept. It was discovered that estimating the slope by method of least square yields the same result with substituting intercept as zero in the least square estimation of intercept in a model with intercept and the theorem was proposed to this respect and the prove was shown. It was revealed that the combination of all the predictor variables in a simple linear regression gives opposite result to what is obtainable when the combined predictor variables were used for a multiple linear regression. It was also observed that this research was in agreement to what was obtained by Sameera (2014) for both model with or without intercept, except for the pooled variables for linear regression and the multiple linear regression model built by Sameera (2014). sugar level is seen contributing more than the age on the basis of coefficient of correlation and determination for model with intercept also, the opposite is the case for model without intercept, where the age out performs the sugar level on the basis of coefficient of correlation and determination. This generally reveals that, the model without intercept is opposite in performance to the model with intercept.

The study reveals that there exists high level of unstable relationships between the response variable (blood pressures) of human beings and the explanatory variables which are, height, weight, sex, sugar level, age, it is evident in the plots seen in Appendix G, H AND I. The unstable relationship witnessed could be as a result of different body chemistry for different individuals.

p-ISSN: 2348-6848
e-ISSN: 2348-795X
Vol. 8 Issue 6
June 2021

References

Cohen et al (2003), Applied multiple regression/Correlation analysis for the Behavioural science (3rd edition): Mahwah, NJ: Lawrence Eribanm Associate-102-109.

Egbule J.F. (2008). Statistics for Researchers in the Behaviourial Sciences and Education Research Oriented edition: Ethiope Pubblishing cooperation, Benin City, Edo State Nigeria – 135 – 147.

Farnsworth D.L. (1990): The effect of a single point on correlation and slope; Internat J. Math & Math. Scio 13(4) – 789 – 800.

Gujarati (2004). Basic Econometrics (4th edition): The MacGraw-Hill Companies.

Helsel D.R. and Hirsch R.M. (2002). Statistical Methods in water resources techniques of water resources investigations: U.S. Geological Survey, Book 4-110-169.

Hoaglin D.C. and Welsch R.E. (1978). The hat matrix in regression and ANOVA: The American statistician 32(1) – 16 – 21.

Iwundu M.P. and Onu O.H. (2017). Preferences of equiradial designs with changing axial distances, design sizes and increase center points and their relationship to the N-point central composite design: International journal of advanced statistics and probability, 5(2)-77-82.

Iwundu M.P. and Udochukwuka (2014). An efficient algorithm for constructing D-optimal design: journal of applied science 14(24)-3547-3554

Johnson R.A. (2003) Probability and statistics for Engineers (6th edition): Pearson Education.

Keller G. and Warrack B. (2003). Statistics for management and economics (6th ed): Thomson Brooks /Cole – 603-641.

Kutner et al (2005). Applied linear statistical model: McGraw-Hill Irwin (5th edition).

Nwaogazie I.L. (2011). Probability and statistics for science and engineering practice. De-Adroit Innovation 13 Annang Str.Ogui N/L out Enugu-101-115.

Onu, O.H. (2017). Preferences of equiradial designs with changing axial distances, design sizes and increase center points and their relationship to the N-point central composite design: An M.Sc dissertation, Math/Statistics Uniport. Unpublished. .

Rawlings et al (2003). Multiple regression analysis using ANCOVA in University model: International Journal of Applied Physical and mathematics, 3(5) – 336 – 340.

Sameera (2014). Comparison between model with or without intercept: Gen Math notes, 21(1) - 118-127