CARD International Journal of Science and Advanced Innovative Research (IJSAIR)
ISSN: 2536-7315 (Print) 2536-7323 (Online)
Volume 2, Number 2, June 2017
http://www.casirmediapublishing.com

# On use of Simple Linear Regression Models

**Wokoma Dagogo S.A. & Innocent Uchenna Amadi**

Department of Mathematics/Statistics
Port Harcourt Polytechnic, Rumuola, Rivers State
**Corresponding Author:** Wokoma Dagogo S.A.

## ABSTRACT

In this paper we focused on simple linear regression modeling techniques to Nigerian Oil Export between the periods of 2004-2013 obtained from statistical bulletin. Our result shows that there is a direct linear relationship between Oil Export(X) and Gross domestic product(Y).The coefficient of determination provided a good summary of the total variability explained by the chosen fitted model. We also observe that Oil export is statistically significant due to the fact that Oil export plays a key role in the economic growth of Nigeria. Hence the proposed model $\hat{y}=58123.115+0.723x$ is found to be adequate.

**Keywords**: Regression Analysis, forecasting, Oil Export, Modeling and Nigeria.

## INTRODUCTION

The importance of linear regression models cannot be over emphasized as its application are numerous in real life situation. However oil export to Nigerian economy can firstly be appreciated from the perspective of export and economic development. It includes the total of crude oil products sold to other foreign countries. The aim of this paper is to develop a model for possible forecasting that will aid in decision making.

However, a lot of researchers have shown renewed interest in the modeling using simple linear regression to model some physical and environmental problems. A few of these are: Mran Abbas *et al;* (2015), Nwankwo (2015), Samaila (2013), Emerole (2013), Oliver and Okpe (2015), Shehu (2012) and Amadi *et al;* (2016). Regression analysis are perhaps the most commonly used forms of statistical analysis and are invaluable when making a large number of business and economic decisions

Nwachukwu and Egbulonu (2000)., hopefully, the present work can provide further insight, as it widens the area of application.

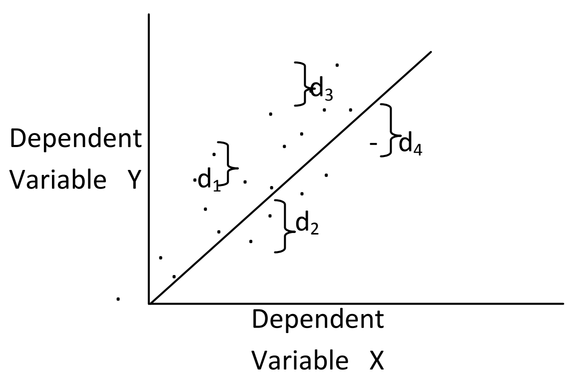
## MATERIALS AND METHOD

The data for this study is a secondary data, obtained from CBN Statistics Bulletin between the periods of 2004 – 2013.

### Simple Linear Regression Models

For simple linear regression we wish to determine the relationship between a single independent variable X and a dependent variable Y. We assume that through functional relationship there is a straight line and may be described by$Y = a + bx$.Where a is y- intercept and b is the slope of the line defined as the change in y for a unit change in Y or a unit change in X. The expected value of Y for each value of x is given $E(Y/X) = a + bx$, or $E(Y) = a + bx$. Suppose we have n pairs of observations $(Y_1, X_1)$, $(Y_2, X_2)$ … $(Y_n, X_n)$ and wish to fix the module $E(Y) = a+bx$ to the set of data points.

Figure 1: Scatter diagram of Simple Linear Regression




The convenient way to accomplish this and obtain estimated line with good properties is to minimize the sum of squares of the vertical deviations from the fitted lines.

Let $(\hat{Y}) = \hat{a} + \hat{b}x$ be the estimated line or an equations that describes the relationship between dependent and independent variable the predicted value at ith observation is $\hat{Y}_i = \hat{a} + \hat{b}x_i$. The deviation of the observe value of Y from ŷ is $d_i = y_i - \hat{y}_i$, for convenient we may call it the deviation error and express it as $e_i = y_i - \hat{y}_i$. our interest is to minimize the sum of squares of the deviations.

i.e Min $\sum_{i=1}^{n}(yi - \hat{y}i)^2$ (2.1)

Let $L = \sum_{i=1}^{n}(yi - \hat{y}i)^2$ (2.2)

$$\frac{\partial L}{\partial \hat{a}} = O \text{ and } \frac{\partial L}{\partial b} = 0 \quad (2.3)$$

$$L = \sum (yi - \hat{y}i)^2 \quad (2.4)$$

$$\text{Min } \sum (yi - \hat{y}i)^2 \quad (2.5)$$

$$\frac{\partial L}{\partial a} = \frac{\partial \Sigma(yi - \hat{y})2}{\partial a} \quad (2.6)$$

$$\frac{\partial \sum_{i=1}^{n}\left(yi - \hat{a} + \hat{b}x_i\right)^2}{\partial \hat{a}} \quad (2.7)$$

$$2\sum_{i=1}^{n}\left(yi - \left(\hat{a} - \hat{b}x_i\right)(-1) = 0\right. \quad (2.8)$$

$$-2\sum_{i=1}^{n}\left(yi - \hat{a} - \hat{b}x_i\right) = 0 \quad (2.9)$$

$$\sum_{i=1}^{n}\left(yi - \hat{a} - \hat{b}x_i\right) = 0 \quad (2.10)$$

$$\sum_{i=1}^{n}\left(y_i - \hat{a} - \hat{b}x_i\right) = 0 \quad (2.11)$$

$$\sum_{i=1}^{n} y_i - \sum_{i=1}^{n} \hat{a} - \hat{b}\sum_{i=1}^{n} x_i \quad (2.12)$$

$$\sum_{i=1}^{n} yi - n\hat{a} - \hat{b}\sum_{i=1}^{n} x_i \quad (2.13)$$

$$\frac{\partial L}{\partial \hat{b}} = \frac{\partial \Sigma\left(yi - \hat{a} - \hat{b}x_i\right)^2}{\partial \hat{b}} \quad (2.14)$$

$$2\Sigma\left(yi - \hat{a} - \hat{b}x_i\right)(-x_i) \quad (2.15)$$

$$-2\Sigma\left(yi - \hat{a} - \hat{b}x_i\right)(x_i) = 0 \quad (2.16)$$

$$\sum_{i=1}^{n}\left(yi - \hat{a} - \hat{b}x_i\right)x_i = 0 \quad (2.17)$$

$$\sum_{i=1}^{n}\left(x_i y_i - \hat{a}x_i - \hat{b}x_i^2\right) = 0 \quad (2.18)$$

$$\sum_{i=1}^{n} x_i y_i - \hat{a}\Sigma x_i - \Sigma x_1^2 = 0 \qquad (2.19)$$

To solve for $\hat{b}$ multiply equ (2.14) by sum of $x_i$

Thus we have $\Sigma x_i \, \Sigma y_i - n\hat{a}\Sigma x_i - \hat{b}(\Sigma x_i)^2 = 0$ (2.20)

$$n\Sigma x_i \, \Sigma y_i - n\hat{a}\Sigma x_i - n\hat{b}\Sigma x_i^{\ 2} = 0 \qquad (2.21)$$

Subtracting (2.22) – (2.21) gives

$$n\Sigma x_i \, \Sigma y_i - \Sigma x_i \Sigma y_i - n\hat{b}\Sigma x_i^{\ 2} + \hat{b}(\Sigma x_i)^2 = 0 \qquad (2.22)$$

$$n\Sigma x_i \, \Sigma y_i - \Sigma x_i \Sigma y_i - \hat{b}\left(n\Sigma x_i^{\ 2} - (\Sigma x_i)\right)^2 \qquad (2.23)$$

$$\hat{b} \quad = \quad \frac{n\Sigma x_i y_i - \Sigma x_i y_i}{n\Sigma x_i^2 - (\Sigma x_i)^2}$$

From (2.21)

$$\Sigma x \Sigma y - \hat{b}(\Sigma x_i)^2 = n\hat{a}\Sigma x_i \qquad (2.24)$$

Divide through by $n\Sigma x_i$

$$\frac{\Sigma x \Sigma y - \hat{b}(\Sigma x_i)^2}{n\Sigma x_i} = \frac{n\hat{a}\Sigma x_i}{n\Sigma x_i} \qquad (2.25)$$

$$\frac{\Sigma y}{n} - \frac{\hat{b}\Sigma x_i}{n} = \hat{a} \qquad (2.26)$$

$$\hat{y} - \hat{b}x = \hat{a} \qquad (2.27)$$

$$\hat{a} = \hat{y} - \hat{b}\bar{x}$$

Or

$$\hat{a} = \frac{\Sigma x^2 \Sigma y - \Sigma x \Sigma xy}{n\Sigma x^2 - (\Sigma x)^2}$$

## Inferences in Regression Analysis

$Y_i = \beta_0 + \beta_1 X_1 + \ldots + \beta_k X_{kj} + e_j$, Where j = 1,2, …, n

The residual estimate is given by $e_i = (y_i - \hat{y}_i)$.

The residual sum of squares (SS) is $\sum_{i=1}^{n} e_i^2 = \Sigma(y_i - \hat{y}_i)^2$

This can be shown to be equivalent to

$$\Sigma(y_i - \hat{y}_i)^2 = \Sigma(y_i - \bar{y}_i)^2 - \Sigma(y_i - \bar{y})\left[\beta_1(x_{ij} - \bar{x}_1) + \hat{\beta}_2(x_{2j} - \bar{x}_2)\right]$$

$$= \qquad \Sigma(y_i - \bar{y})^2 - \left(\hat{\beta}_1 \varsigma_{xiy} + \hat{\beta}_2 \varsigma_{x2y}\right)$$

Rearranging we have

$$\Sigma(y_i - \bar{y})^2 - (\beta_1 \varsigma_{x1y} + \hat{\beta}_2 \varsigma_{x2y} + \Sigma(y_i - \hat{y}_j)^2$$

Total SS = Regression SS + Residual SS
ESS is obtain by ESS = SYY - RSS

**Table 1: ANOVA Table for Simple Linear Regression**

| Source of Variation | Degrees of Freedom | Sum of Squares | Mean Squares | F – Ratio |
|---|---|---|---|---|
| Regression | K | SSR | MSR | F - $\frac{MSB}{MSE}$ |
| Error or Residual | n - p | SSE | MSE | |
| Total | n – 1 | SSY | | |

The coefficient of determination is given by $R^2 = 1 - \frac{SSE}{SST}$

An F – test is usually carried out to determine whether the overall regression of y on the independent variable is significant.

The Hypothesis is:
$H_0$: $\beta_1 = 0$
$H_1$: $\beta_1 \neq 0$
All analysis in this work was done using SPSS 20 statistical package.

## RESULTS AND DISCUSSIONS

The key results that we have obtained in this research are presented in this section using the above mentioned method of analysis.

**Table 2: Model summary of simple linear regression**

| Model | R | R Square | Adjusted R Square | Std. Error of the Estimate | Change Statistics | | | | | Durbin-Watson |
|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | R Square Change | F Change | Df1 | Df2 | Sig. F Change | |
| 1 | .723[a] | .523 | .463 | 186573.1860 | .523 | 8.757 | 1 | 8 | .018 | 2.698 |

a. predictors: (Constant), Oil exports
b. Dependent Variable: Gross Domestic Product

**Table 3: ANOVA table in testing for significance**

| Model | | Sum of squares | Df | Mean square | F | Sig. |
|---|---|---|---|---|---|---|
| 1 | Regression | 304833896914.164 | 1 | 304833896914.164 | 8.757 | .018b |
| | Residual | 278476429991.125 | 8 | 34809553748.891 | | |
| | Total | 583310326905.289 | 9 | | | |

a. Dependent Variable: Gross Domestic Product
b. Predictors: (Constant), Oil exports

**Table 4: Estimated Model Results**

| Model | Unstandardised coefficients | | Standardized coefficients | T | Sig. | 95.0% Confidence Interval for B | |
|---|---|---|---|---|---|---|---|
| | B | Std. Error | Beta | | | Lower Bound | Upper Bound |
| 1 Constant) | 58123.115 | 214474.787 | | .271 | .793 | -436456.630 | 552702.860 |
| Oil exports | .065 | .022 | .723 | 2.959 | 0.18 | .014 | .116 |

a. Dependent Variable: Gross Domestic Product
b. Predictors:( constant),Oil exports

**Table 5: Forecast of Estimated GDP**

| Year | GDP(Y) | Estimated GDP ($\hat{Y}$) |
|---|---|---|
| 2004 | 527576.0 | 439560.563 |
| 2005 | 56193.4 | 98750.9432 |
| 2006 | 595821.6 | 488902.1318 |
| 2007 | 634251.1 | 516686.6603 |
| 2008 | 672202.6 | 544125.5948 |
| 2009 | 718977.3 | 577943.7029 |
| 2010 | 775525.7 | 618828.1961 |
| 2011 | 842288.4 | 667097.6282 |
| 2012 | 909051.7 | 715367.4941 |
| 2013 | 951330.1 | 745934.7773 |

In table 2, R = 0.723 which indicates a high positive correlation, that is direct linear relationship between the GDP (y) and the oil exports (x). The coefficient of Determination ($r^2$) is 52% of the total variation in y is explained by the fitted model.

$H_o = \beta_1 = 0$ $V_s$ $H_1 = \beta_1 \neq 0$

Reject Ho if $F_{cal} > F_{tab}$, $V_1$, $V_2$ and accept Ho otherwise from tables: f0.05, 1, 9 = 5.12

Decision rule: Since $F_{cal} = 8.757$ is greater than $5.12 = f_{tab}$, we reject $H_o$ that $\beta 1 = 0$ at the 5% level of significance. Therefore, we conclude that there is underlying linear relation and hence oil exports (x) is useful as a predictor of GDP(y).

In table 3, we observed that $\beta = 58123.115$ which implies that if the independent variable (oil export) is held constant, the dependent variable (GDP) will increase by 58123.115 units. Hence the fitted model is given as $\hat{y} = 58123.115 + 0.723x$. The independent variable (oil export) is statistically significant; this is due to the fact that oil exports play a key role in the economic growth of Nigeria. This result agrees with work of Amadi *et al;* (2010).

However, we can be 95% confident that the slop of regression line is somehow between .014 and .116 (see table 4).
In table 5, you can agree with us that the variations between GDP(Y) and estimated GDP(Y) are not much. All this is to attest for adequacy of the fitted model equation.

## CONCLUSION

In particular, we have observed that Oil Export is statistically significant in terms of economic growth. This research have provided the basis for Nigerians not to compromise with any person or group of persons who may jeopardize the serious effort of federal Government in respect to Oil Export in Nigeria. It is an empirical approach for explaining for the growth of oil export. We shall be looking at the effectiveness of the fitted model in our next study.

## REFERENCES

Amadi, I.U., Anthony I. Wegbom and Bethel A. Anyamele (2016). An empirical Use of Multiple Regression Models On Oil Export and Non Oil Export in Nigeria. *International organization of*

*Scientific Research Journal of Mathematics ,Volume 12, Issue 3, pp 60*-62.

M.P. Iwundu (2016) *Regression Analysis lecture note, unpublished, Department of Mathematics,* RSUST, Nigeria.

Emerole G.E. Georgine (2013). Empirical Analysis of the Problem and Prospects of Export Marketing of Non-oil Agricultural Products in Nigeria.*Journal of Economics Volume 5, No. 23.pp132-147.*

Fatoki, O., U.A. Mbata, G.A. Olalude O. Abass (2010.) An application of Arima model to the Nigeria gross domestic product.*International journal of statistics and systems. Vol.5, NO1, PP.63-72.*

Nwachukwu, Vitals Obioma and Egbulonu, K.G (2000).*Elements of Statistical inference*. Owerri Imo State, Nigeria: Peace Enterprises.pp108-109.

Nwankwo , Odi, (2015).Diversification of Non-oil Export Product as a Precondition Accelerated Real Economic Growth in Nigeria. *Journal of Managerial Studies Research. Volume 3, Issue 7, pp 104 – 112.*

Oliver .I. Inyama and Okpe I. Ikechukwu (2015).Crude oil Production, Prices, Export and Foreign Exchange Rate. Do they Interact? Evidence from Nigeria.*International Journal of Development of Emerging Economic Volume 3 No. 2, pp 24 – 37.*

Omran Abbas Yusif Alla, Gadda Modhummed Awad, Adil Ali Ahmed Mohamed (2015). Some Economic Determinants of non-oil Exports in Sudan: An Empirical Investigation. *Journal of Business Studies Quarterly. Vol. 7, No1, pp 1-12.*

Samaila Idi Ningi (2013). An Analysis of Banks of Non-oil Exports in Nigeria.American *Journal of Contemporary Research Volume, 3, No. 1. pp 104-115.*

Shehu U.H. (2012).Corporate Governance Earning management and finance performance: A case study of Nigerian Manufacturing firms. *American international journal of contemporary Research Vol.2, No.7 pp 56-68.*