AN ASSESSMENT OF HIV/AIDS ON AGE AND DIFFERENTIALS IN GENDER USING REGRESSION METHOD

ADAMU, A[1*], A. U. SHELLENG[1] AND A. U. KINAFA[1]

[1]Department of Mathematics, Gombe State University, Gombe
Corresponding Author: absadstat22@gmail.com; Tel: +2348022361062

ABSTRACT

HIV/AIDS is a contagious disease as a result of unprotected sex or through blood transfusion or through sharing of sharp objects. When first infected. A person can develop an acute infection which can range in severity from a very mild illness with few or no symptoms to a serious condition requiring hospitalization. Acute HIV/AIDS can manifest within the first 6 months of contacting the disease. The disease become one of the factors for the death of many people in our society for over 3 decades. The method used for this study is multiple regression method and t-test. The data was secondary collected from the record department of Gombe specialist hospital for a period of ten years from 2008 – 2017. The result of this study shows that, there is significance difference in both gender and age. Such we conclude that gender and age contributed immensely in the spread of HIV/AIDS. Majority of the victims are female, also almost 88% of the victims are youth of the age 0 – 49 years while elders constitute the remaining percent. Hence, from the result, government and non-governmental organization need to embark on mass awareness and sensitization campaign to all nook and crannies of the state with a target of reducing the risk of contracting the disease.

Keywords: Regression; Dependent; Explanatory; Human Immune Virus; Acquired Immunodeficiency Virus.

INTRODUCTION

Human Immune Virus and Acquired Immune Deficiency Syndrome (HIV/AIDS) in our state become a cause of alarm as a result of many innocent promising lives are being waste on daily basis in our community through what could be aptly described as avoidable incidence. The present situation bring about the creation of independent bodies like Gombe State Action Committee on Aids (GOMSACA) and the National Action Committee on Aids (NACA), a body statutorily mandated to ensure the safety of lives and properties in the state (Health, 2003). The Doctors says that statistically in some previous years, the list of patients living with HIV/AIDS cannot be quantified but statistics shows that the rate of infection and death due to human immune virus (HIV) disease is on the increase. It is also pathetic to note that the degree of agony and loss of human lives brought about by the spread of the diseases in our dear state has become a challenge to humanity. Many factors have been implicated as responsible for the incidence in the state. These include deplorable condition of our community, absence of prevention or careless over the prevention, some of these human factors could be treated to the reliance of the ignorance and careless. Shortly after, Doctors and scientist worldwide realized they were up against new and little understood viral foe (HIV) with a difficult ability to defeat the most powerful fighters, the human immune virus (AIDS, 2008). In turn public fears mounted as new reports detailed lack of medical power/weapons with which to assault this new frightening disease and its potential ability to spread to those previously not thought of to be at risk. Acquired Immune Deficiency Syndrome (AIDS) is cause by a virus called Human Immune System and progressively destroys the body ability to fight infection and certain cancers. HIV is most commonly spread through sexual intercourse with an infected partner. There are also other means of transmission like through blood transfusion, sharing of sharp objects like razor blade with infected person and so forth. Research has shown that 80% of HIV transmission account for the rest and the unintended consequences of rapid modernization and urbanization (including a thriving commercial sex trade, child labour, overcrowding and force of voluntary migration) (WHO, 2004). This is perhaps not surprising as many in the state and indeed many Nigerian do not belief that the disease Human Immune Virus and Acquired Immune Deficiency Syndrome (HIV/AIDS) is real.

It is arguable that the disease is not real or the disease is due to poverty, but all these will be secondary, if the human mind is constantly turned to the fact that deficiencies are always present and there is a need to make the best advantage of a worst circumstance. It is so worthy to note that fundamental to good usage. Also, culture is non – existing in the present generation, especially in most of our places like village and without this, it might be on uphill task to minimally reduce the number of infections. It needs to be mentioned that because of the frequent cases of the incidence and staggering number of deaths that occur in such diseases. The Gombe State was adjusted as the second most dangerous state to the infection in the country. This informed the establishment of the Gombe State Action Committee on Aids (GOMSACA) and some other non – governmental organization. As part of the preparatory work for the setting up of safety outfit of various state law scheme were studied and found from state to state throughout the federation. This informed the basis for standardization in the practice of how to prevent. Therefore, we are going to use regression and students t – test in order to ascertain the growth and spread of the disease in the state.

Multiple regression analysis studies the simultaneous emotions that two or more independent variables may have over one dependent variable (Lefter, 2004), Regression methods continue to be an area of active research. In recent decades, new methods have been developed for robust regression, regression involving correlated responses such as time series and growth curves, regression in which the predictor or response variables are curves, images, graphs, or other complex data objects, regression methods accommodating various types of missing data, nonparametric regression, Bayesian methods for regression,

regression in which the predictor variables are measured with error, regression with more predictor variables than observations, and causal inference with regression. The term "regression" was coined by (Galton, 1894) in the nineteenth century to describe a biological phenomenon. The phenomenon was that the heights of descendants of tall ancestors tend to regress down towards a normal average (a phenomenon also known as regression toward the mean). For Galton, regression had only this biological meaning, but his work was later extended by Yule Udny and Pearson Karl to a more general statistical context. In the work of (Yule & Kendall, 1937) and (Pearson, 1896), the joint distribution of the response and explanatory variables is assumed to be Gaussian. This assumption was weakened by (Fisher, 1935) in his works. Fisher assumed that the conditional distribution of the response variable is Gaussian, but the joint distribution need not be. In this respect, Fisher's assumption is closer to Gauss's formulation of 1821. T – test is used to test two variables in order to clarify whether there is statistically significant between the two variables or not.

MATERIALS AND METHOD

The Research Design: Regression Analysis and Students t- Test

A regression analysis is a statistical technique for investigating and modelling the relationship between variables. It is relationship between the regressor and the response, for example the relationship may be suggested by theoretical considerations. Regression analysis can aid confirming a cause effect relationship but it cannot be the sole basis of such a claim.

$$\hat{Y} = \alpha + \beta X \qquad (1)$$

where

represent the predicted value of Y

α represent the intercept of the best fitted line

β represent the slope of the line

X represent the independent variable

It is important to note that regression analysis is part of a broader data analytic approach to problem solving, that is regression equation itself may not be the primary objective of the study. It is usually more important to gain insight and understanding concerning generating the data.

Least Square Estimation

The parameters α and β are unknown and must be estimated using simple data in case of one independent variable. The method of ordinary least square is used to estimate α and β so that the sum of the squares of the differences between the observations Y_i and the straight line as shown in equation 2 below;

$$Y_i = \alpha + \beta X_i + \varepsilon_i \qquad (2)$$

$$\sum_{i=1}^{n} \varepsilon_i = \sum \left(Y_i - \alpha - \beta X_i \right)^2 \qquad (3)$$

We take partial derivative of $\sum_{i=1}^{n} \varepsilon_i^2$ with respect to α and β and equate each to zero

$$\frac{\partial \sum \varepsilon_i^2}{\partial \alpha_i} = -2 \sum \left(Y_i - \alpha - \beta X_i \right) \qquad (4)$$

$$\frac{\partial\sum\varepsilon_i^2}{\partial\beta_i} = -2\sum(Y_i - \alpha - \beta X_i)X_i \quad (5)$$

$$\sum(Y_i - \alpha - \beta X_i) = 0 \quad (6)$$

$$\sum(Y_i - \alpha - \beta X_i)X_i = 0 \quad (7)$$

From equation 5 we have

$$\sum Y_i - n\alpha - \beta\sum_{i=1}^{n} X_i = 0 \quad (8)$$

$$\hat{\alpha} = \frac{\sum Y_i}{n} - \hat{\beta}\frac{\sum_{i=1}^{n} X_i}{n} \quad (9)$$

$$\hat{\alpha} = \bar{Y} - \hat{\beta}\bar{X} \quad (10)$$

Also from equation 6 we have

$$\sum Y_i X_i - \alpha\sum X_i - \beta\sum_{i=1}^{n} X_i^2 = 0 \quad (11)$$

$$\hat{\beta}\sum_{i=1}^{n} X_i^2 = \sum Y_i X_i - \hat{\alpha}\sum X_i \quad (12)$$

$$\hat{\beta}\sum_{i=1}^{n} X_i^2 = \sum Y_i X_i - (\bar{Y} - \hat{\beta}\bar{X})\sum X_i \quad (13)$$

$$\hat{\beta}\sum_{i=1}^{n} X_i^2 = n\sum Y_i X_i - \left(\sum Y_i - \hat{\beta}\sum X_i\right)\sum X_i \quad (14)$$

$$\hat{\beta}\sum_{i=1}^{n} X_i^2 = n\sum Y_i X_i - \sum Y_i\sum X_i - \hat{\beta}\left(\sum X_i\right)^2 \quad (15)$$

$$\hat{\beta}\left(\sum_{i=1}^{n} X_i^2 - \left(\sum_{i=1}^{n} X_i\right)^2\right) = n\sum Y_i X_i - \sum Y_i\sum X_i \quad (16)$$

$$\hat{\beta} = \frac{n\sum Y_i X_i - \sum Y_i\sum X_i}{\left(\sum_{i=1}^{n} X_i^2 - \left(\sum_{i=1}^{n} X_i\right)^2\right)} \quad (17)$$

Multiple Regressions Model

The multiple linear regression model was used to analyze the variables and the model is expressed as

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_p X_p + \epsilon_i \quad (18)$$

Where Y is the dependent variable, $X_1, X_2, \cdots, X_p$ are the independent variables, otherwise, called the predictors, $\beta_0, \beta_{1,} \cdots \beta_p$ are the parameters of the model indicating the number of units in Y caused by one unit increase in the independent variables and p is the number of independent variables in the model. Equation 18 is called the intercept model with β_0 as the intercept. When the intercept is not significant, it is often better to remove it from the model and specify the model as,

$$Y = \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_p X_p + \epsilon_i \quad (19)$$

The following notations were used to represent the variables in the data: $X_1 = Number\ of\ Male\ Patient$, $X_2 = Number\ of\ Female\ Patient$, $X_3 = Number\ of\ Patients\ Age\ 0 - 49\ Years$, $X_4 = Number\ of\ Patients\ Age\ 50\ and\ above$, X_5 = Number of Death Patients and Y= Total number of HIV patients.

Thus, this regression model involves more than one regressor's variable and it can be written in matrix form as

$$Y = X\beta + Ei \quad (20)$$

Where $Y = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}$, $X = \begin{bmatrix} 1 & x_{11} & x_{21} & \dots & x_{k1} \\ 1 & x_{12} & x_{22} & \dots & x_{k2} \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & x_{1n} & x_{2n} & \dots & x_{kn} \end{bmatrix}$

$\beta = \begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_k \end{bmatrix}$ And $U = \begin{bmatrix} E_1 \\ E_2 \\ \vdots \\ E_k \end{bmatrix}$

The above equation can also be written as

$$Y = \alpha + \beta_1 X_1 + \beta_2 X_2 + \varepsilon \quad (21)$$

In general, Y is an n x 1 vector of the observations; x is an $n \times k$ matrix of the level of regressors variables, β is a k x 1 vector of the regression coefficients and ε is an n x 1 vector of random errors.

The expression to obtain the parameter $\hat{\beta}$ is

$$S(\beta) = \sum \varepsilon_i^2 = \varepsilon^2 \varepsilon = (y - x\beta)'(y - x\beta)$$

$$= y'y - \beta'x'y - y'x\beta + \beta x'x\beta$$

$$= y'y - 2\beta'x'y - \beta'x'x\beta \quad (22)$$

Since $\beta'x'y$ is a 1x1 matrix and its transpose $(\beta'x'y) = y'\beta$ is the sale. The least square estimates must satisfy

$$\frac{2s}{2\beta} = -2x'y + 2x'x\hat{\beta}.$$

$$x'x\hat{\beta} = x'y \quad (23)$$

Multiplying both sides by the inverse of $x'x$ to get the least square estimator of β we have

$$\hat{\beta} = (x'x)^{-1}x'y \quad (24)$$

Provided $(X'X)^{-1}$ matrix exists, and it always exists if the regressors are linearly independent that is, if no column of the X matrix is a linear combination of the columns.

$$\begin{bmatrix} n & \sum x_{i1} & \dots & \sum x_{ik} \\ \sum x_{1i} & \sum x_{1i}^2 & \dots & \sum x_{1i}x_{ki} \\ \vdots & \vdots & \ddots & \vdots \\ \sum x_{ki} & \sum x_{ki}x_{1i} & \dots & \sum x_{1i}^2 \end{bmatrix} \quad 25$$

Student t - Test

The student t distribution requires the assumption that the two variables possessed independent natural distribution with equal standard deviation. If these assumptions are measurably satisfied, then one may treat the variables.

Test Statistics

$$t = (X_1 - X_2 - (\theta_1 - \theta_2)) / \hat{\sigma}\sqrt{\frac{1}{n_1} + \frac{1}{n_2}} \quad 26$$

where $\hat{\sigma} = \sqrt{\frac{[n_1-1)S_1^2+(n_2-1)}{n_1+n_2-2}}$

Hypothesis H$_o$: $\theta_1 = \theta_2$ vs H$_o$: $\theta_1 \neq \theta_2$

Decision: Reject H$_0$ if $t_{cal} \geq t_{n1-n2-2}$ (Nwaosu, 2008)

Test for the Significance of Regression

Test for significance of regression is a test to determine if there is a linear relationship between the response y and any of the regressor variables $x_1, x_2, \dots, x_k$. The approximation hypothesis is

H$_o$: $\beta_1 = \beta_2 =, \dots, \beta_k = 0$

H$_1$: $\beta_1 \neq 0$ for atleast one of j

Rejection of the null hypothesis implies that at least one of the regressors $x_1, x_2, \dots, x_k$ contribute significantly to the model., the test procedure is a generalization of the Analysis of variance used in simple linear regression. The sum of squares total SSt is partitioned into a sum of square due to regression, SS_r and sum of error is SSe

Where;

$SSr = SSt - SSe$ 27

$SSe = \sum_{i=1}^{n}(y_i - \bar{y})^2$ 28

$SSr = \sum_{i=1}^{n}(\bar{y} - \hat{y})^2$ 29

$SSt = \sum_{i=1}^{n}(y_i - \hat{y})^2$ 30

If the null hypothesis is true, then $\frac{SSr}{\sigma^2}$ follows χ_k^2 distribution which has the same number of degree of freedom as number of regression variable in the model. $\frac{SSe}{\sigma^2} \sim \chi_{n-k-1}^2$

$F = \frac{SSr/k}{SSe/(n-k-1)} = \frac{MSr}{MSe}$, Which follow $F_{k,n-k-1}$ distribution.

Table 1: General ANOVA table

Source of variation	Sum of squares	Degree of freedom	Mean of square	F_0
Regression	SSR	K – 1	MSR	$\frac{MSR}{MSE}$
Residual	SSE	N – k	MSE	
Total	SST	N - 1		

The expected mean square indicates that if the observed value of F_0 is large, then it is likely that at least one $\beta \neq 0$, then F_0 follows a non-central F distribution with k and n-k-1 degree of freedom. Therefore, to test the hypothesis$H_0: \beta_1 = \beta_2 = \cdots = \beta_k = 0$, conjugate the test statistics F_0 and reject H_0 if $F_0 > F_{\alpha, (K-1, N-k)}$.

RESULTS

Data Analysis

This analysis is carried out to verify whether the disease is increasing or decreasing as time goes on in Gombe State and the result is shown below. The data were analyzed using STATA version 12.0 and the result of the analysis is as follows.

Table 2: Regression ANOVA for general model

Source	SS	df	MS		
				Number of obs	= 30
				F(5, 24)	= 385.28
Model	262674.475	5	52534.895	Prob > F	= 0.0000
Residual	3272.49142	24	136.353809	R-squared	= 0.9877
				Adj R-squared	= 0.9851
Total	265946.967	29	9170.58506	Root MSE	= 11.677

Table 2 shows the ANOVA for general model which indicates that there is significant difference between the factors (treatment) at 5% level of significant. Therefore, we reject H_o and accept H_a and conclude that the rate of spread of HIV decreases according to time, based on the data obtained. R-squared also explained that 99% of the variability is accounted for which Root MSE as 11.677.

Table 3 shows the regression model as $Y = 1.914971 + 0.5378001X_1 + 1.236468X_2 - 0.0216408X_3 - 0.0481885X_4 + 1.137794X_5$ which indicates positive relationship between gender and death, while negative relationship between age, t-test indicates that varibles MALE, FEMALE and Death are significant while AGE 0 -49 years and 50 > years are not significant. Hence, we reject H_o in Male, Female and Death and conclude that there is significant difference among them. In other word, we accept H_o in Age 0 – 49 years and 50 > years and conclude that there is no significant difference between them.

Table 3: Regression coefficient

TOTAL	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
MALE	.5378001	.0764142	7.04	0.000	.3800891	.6955112
FEMALE	1.236468	.0616295	20.06	0.000	1.109271	1.363666
AGE 0 – 49 YEARS	-.0216408	.0313252	-0.69	0.496	-.0862929	.0430113
AGE 50 >	-.0481885	.0468431	-1.03	0.314	-.1448679	.048491
DEATH	1.137794	.2210704	5.15	0.000	.6815271	1.594061
_cons	1.914971	8.323586	0.23	0.820	-15.26407	19.09401

Table 4 shows the ANOVA for Male and Female which indicates that there is significant difference between the factors (treatment) at 5% level of significant. Therefore, we reject H_o and accept H_a and conclude that the rate of spread of HIV which respect to gender decreases as time goes on, based on the data obtained. R-squared also explained that 97% of the variability is accounted for which Root MSE as 17.

Table 4: Regression ANOVA for gender

Source	SS	df	MS			
Model	258143.923	2	129071.961	Number of obs	=	30
Residual	7803.0438	27	289.001622	F(2, 27)	=	446.61
Total	265946.967	29	9170.58506	Prob > F	=	0.0000
				R-squared	=	0.9707
				Adj R-squared	=	0.9685
				Root MSE	=	17

Table 5 shows the regression model as $Y = 16.08256 + 0.578561X_1 + 1.244448X_2$ which indicates positive relationship between gender, t-test indicates that variables MALE, FEMALE is significant. Hence, we reject H_o in Male and Female and conclude that there is significant difference between them.

Table 5: Regression coefficients for gender

TOTAL	Coef.	Std. Err.	t	P>|t|	[95% Conf. Interval]	
MALE	.578561	.1095728	5.28	0.000	.3537362	.8033859
FEMALE	1.244448	.068492	18.17	0.000	1.103914	1.384981
_cons	16.08256	11.2754	1.43	0.165	-7.052641	39.21777

Table 6 shows the ANOVA for Age which indicates that there is significant difference between the factors (treatment) at 5% level of significant. Therefore, we reject H_o and accept H_a and conclude that the rate of spread of HIV which respect to Age decreases as time goes on, based on the data obtained. R-squared also explained that 47% of the variability is accounted for which Root MSE as 72.

Table 7 shows the regression model as $Y = 159.0498 + 0.620592X_1 + 0.5720893X_2$ which indicates positive relationship between age, t-test indicates that varibles 0 – 49years and 50years > are significant. Hence, we reject H_o in age 0 – 49years and 50years > and conclude that there is significant difference between them.

Table 6: ANOVA for age

Source	SS	df	MS		
				Number of obs	= 30
				F(2, 27)	= 12.00
Model	125173.993	2	62586.9965	Prob > F	= 0.0002
Residual	140772.974	27	5213.81384	R-squared	= 0.4707
				Adj R-squared	= 0.4315
Total	265946.967	29	9170.58506	Root MSE	= 72.207

Table 7: Regression coefficients for age

TOTAL	Coef.	Std. Err.	t	P>|t|	[95% Conf. Interval]	
AGE 0 – 49years	.620592	.12803	4.85	0.000	.357896	.8832879
AGE 50 >	.5720893	.2509704	2.28	0.031	.0571406	1.087038
CONSTANT	159.0498	39.71991	4.00	0.000	77.55128	240.5483

Table 8 shows the ANOVA for Death patients which indicates that there is significant difference between the factors (treatment) at 5% level of significant. Therefore, we reject H_o and accept H_a and conclude that HIV is a disease that killed if care is not taken, based on the data obtained. R-squared also explained that 15% of the variability is accounted for which Root MSE as 89.92.

Table 8: ANOVA for death patients

Source	SS	df	MS		
				Number of obs	= 30
				F(1, 28)	= 4.89
Model	39567.393	1	39567.393	Prob > F	= 0.0353
Residual	226379.574	28	8084.98477	R-squared	= 0.1488
				Adj R-squared	= 0.1184
Total	265946.967	29	9170.58506	Root MSE	= 89.917

Table 9 shows the regression model as $Y = 261.0718 + 3.184926X_1$ which indicates positive relationship, t-test indicates that variables Death is significant. Hence, we reject H_o in Death patients and conclude that there is significant different.

Table 9: Regression coefficient

TOTAL	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
DEATH	3.184926	1.439694	2.21	0.035	.2358463	6.134005
_cons	261.0718	39.12331	6.67	0.000	180.9314	341.2123

Figure 1 shows the number of HIV/AIDS patients for a period of ten years from 2008 to 2017. Based on the chart female are the highest victim across the period, 2009 has the highest number of female patients, followed by 2013 while male patients have the highest number in 2008, followed by 2009 in 2007 and 2016 recorded the least number of patients for both male and female. Therefore, from the above chart it indicates that female has the highest risk of contracting HIV/AIDS than their male counterpart because of their number.

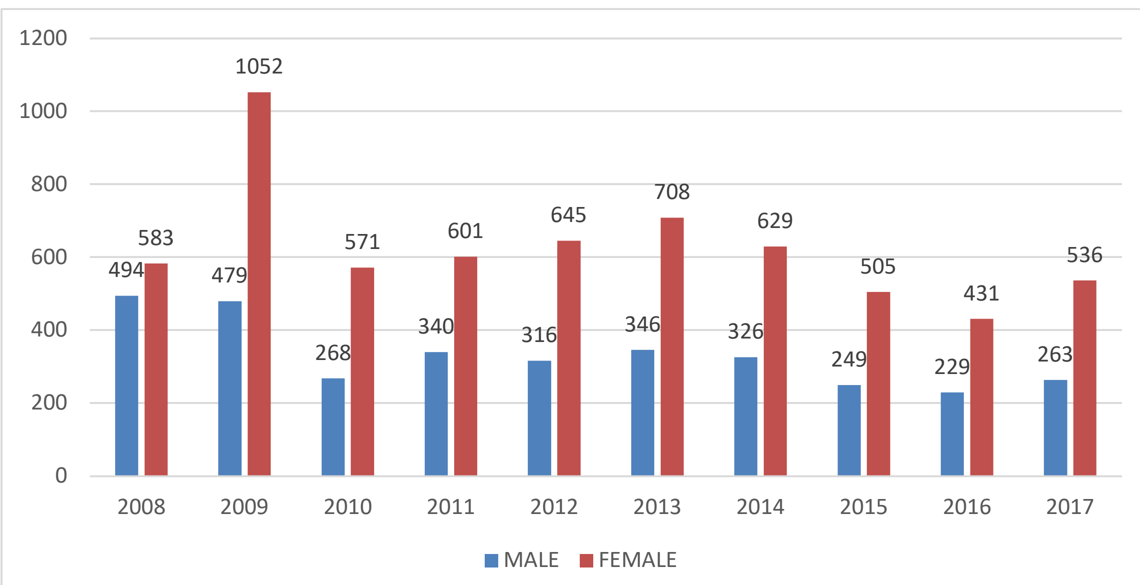

Figure 1: Cluster bar graph of HIV/AIDS transmission in gender

Figure 2 indicates that 2017 has the highest number of death HIV/AIDS patients from the data obtained at specialist hospital Gombe, followed by 2008 while 2015 recorded the least number of death. This gives us information that the number of death patients as a result of HIV/AIDS is fluctuating yearly, it irregular in nature.

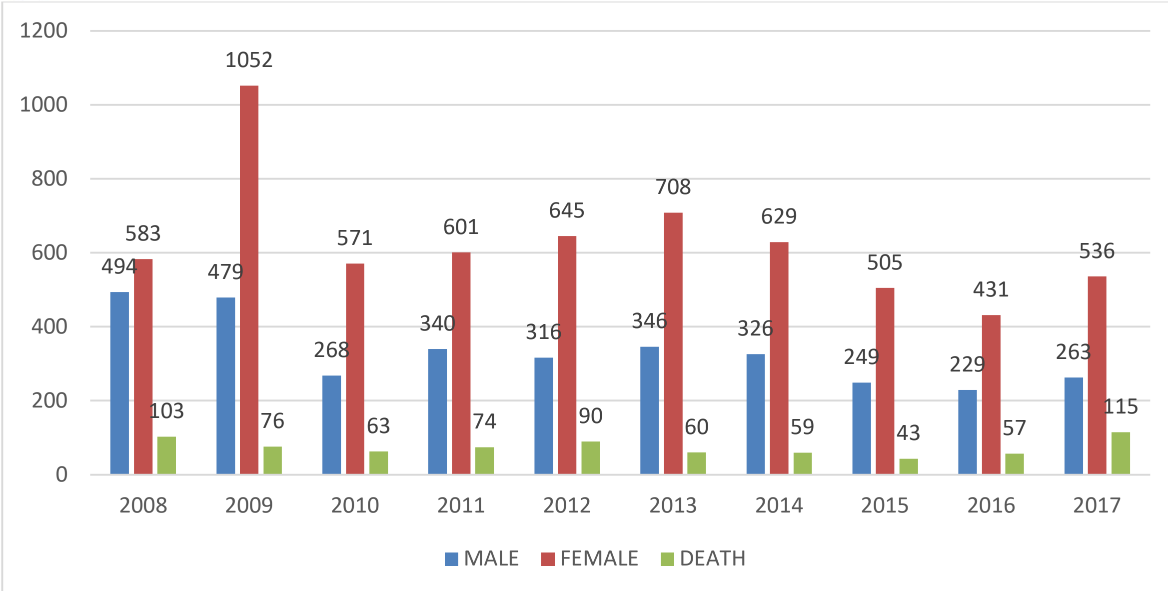

Figure 2: Cluster Bar Chart for HIV/AIDS Transmission of Gender and Death

Figure 3 shows the number of HIV/AIDS patients with respect to age, based on the chart 2009 has the highest number of HIV/AIDS for the people of age 0 – 49 years, followed by 2012 while 2017 has the highest number of HIV/AIDS patients for people of age 50 and above years followed by 2013 and 2017 has the least number of HIV/AIDS patients for the age 0 – 49 years and 2010 for age 50 >.

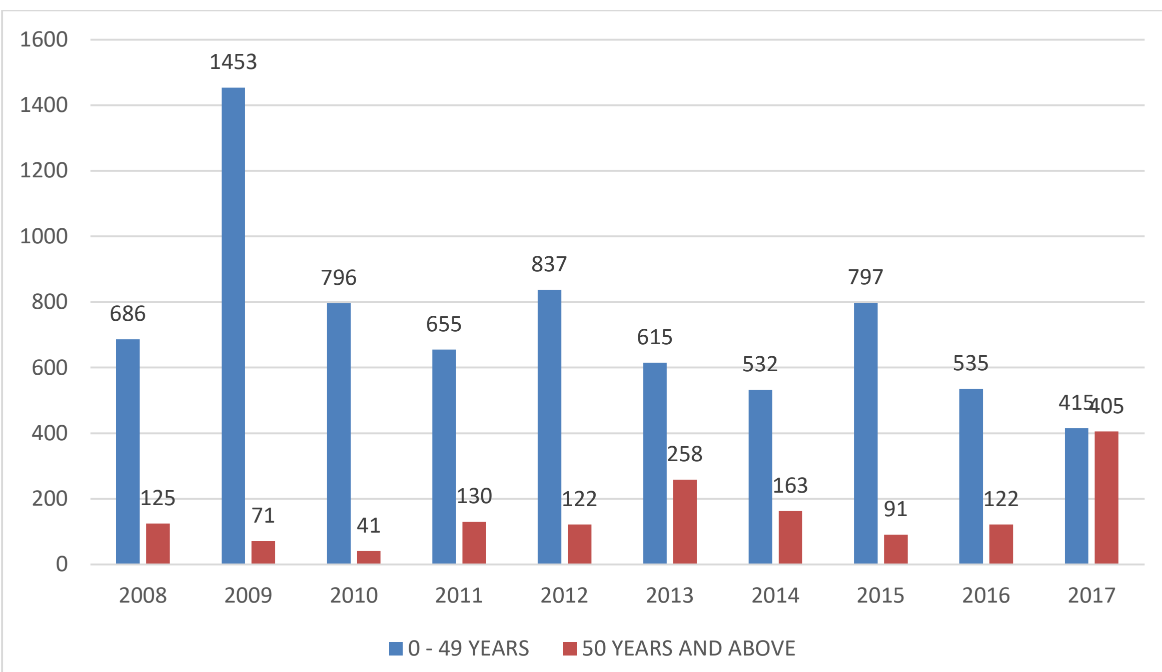

Figure 3: Cluster Bar Chart for HIV/AIDS Transmission of Age (0 – 49y and 50>)

DISCUSSION

Based on the above tables and figure we conclude that all the ANOVA tables have significant difference at 5% level of significant while regression models for the above tables have positive relationship except age barrier in general regression model which have negative relationship. result for t-test indicates that all the coefficient has significant differences for all the above table except coefficient for age differences. R – square values for the first two tables gives good values for the variability while the last two gives weak result (values).

Result from the above chart indicates that female HIV/AIDS are more vulnerable to the disease than their male counterpart because, they have the highest risk of contracting the disease either through sex, abortion, delivery e.t.c. The highest number of female HIV patients recorded is in 2009 while for male is in 2008. Figure 2 shows that people of age 0 – 49 years has the highest number of HIV/AIDS patients than the people of age 50 years and above. Thus, youth are more vulnerable to HIV/AIDS than their old age counterparts. Also, number of death people as a result of HIV/AIDS for the period of the study is irregular or random because of it is fluctuations as time goes on.

REFERENCES

AIDS, G. S. A. C. O., (2008). *Daily Increase and Fear of HIV Victims ,* Gombe: Ministry of Health.

Fisher, R. A., (1935). *The Design of Experiment.* Scotland: Oliver and Boyd.

Galton, F., (1894). *Natural Inheritance.* 5th edition ed. New York: Macmillan and Company.

Health, F. M. O., (2003). *National HIV/AIDS and Reproductive Health Survey,* Abuja: Ministry of Health.

Lefter, C., (2004). *Marketing Researches.* Brasov: Infomarket.

Nwaosu, S. C., (2008). *Notes on Regression Analysis,* Gombe: s.n.

Organization, W. H., (2004). *Fifty Seven World Health Assembly,* Geneva: World Organization.

Pearson, K., (1896). *Mathematical Contributions to the Theory of Evolution. III. Regression, Heredity and Panmixia..* London: Transactions of the Royal Society of London.

Yule, G. U. & Kendall, M. G., (1937). *An Introduction to the Theory of Statistics.* 2nd ed. london: Griffin.

Singh, R. K., Dhama, K., Malik, Y. S., Ramakrishnan, M. A., Karthik, K., Tiwari, R., Joshi, S. K. (2016). Zika virus–emergence, evolution, pathology, diagnosis, and control: current global scenario and future perspectives–a comprehensive review. *Veterinary Quarterly, 36*(3), 150-175.