Journal of the Nigerian Association of Mathematical Physics
Volume **25 (November, 2013), pp 311 – 320**

Criterion for Choosing Among Alternative Competitive Models for Assessing the Fit of Regression Models

[1]Osemeke R. F., [2]Efeizomor R.O and [3]Azagbaekwue A.

[1, 3]Department of Mathematics,
College of Education, Agbor, Delta State.and
[2]College of Education, Agbor, Delta State

Abstract

Several statistical measures such as Mallows C_p statistic, coefficient determination r^2, adjusted r^2, standard error of estimates and forward stepwise regression are used as a criterion for the selection of best subsets regression models in a multiple regression analysis. The best subset fitted models are selected among competitive models based on Cp statistic $\leq (P + 1)$ which means a small biased, the highest value of adjusted r^2, highest value of r^2, lowest value of standard error of estimates, low bivariate correlation among the predictors. The predictors X_3 (PARKING) and X_5 (INCOME) was removed from the model due to non significant effects .The selected best fitted model through studentized residuals (STR) against the predicted value $\widehat{(y)}$ are used to evaluate the aptness of the fitted model .The model X_4 (SHOPCNTR) demonstrate some anomalous features and was improve upon by log transformation .The final fitted model was

$Yi = 37.82 – 0.0021X_1 – 0.531X_2 + 0l0gX_4$.With iteration method of outlier detention, row 5, row 7 and row 18 of Table 8 was removed from the model because each of the value for standardized residuals is outside the range of 2 x standard deviation or -2 x standard deviation. At each evaluation process, there was a greater improvement in the regression coefficient. The standardized residuals, leverage points, and studentized residuals of Table 8 were used to detect outliers as influential. For studentized residuals, any value that exceed +2(up) and -2(down) are regarded as an outlier. The average leverage value is $\frac{P}{N}$, where p is the number of predictors (the number of coefficients plus one for the constant) and n is the sample size. Leverage point greater than $\frac{2k+2}{n}$ should be carefully examined.

***Keywords*:** Best Subset Analysis; Stepwise Estimation; Correlation Matrix; Optimal Model; Residual Plots; Outliers

1.0 Introduction

Regression analysis is an important statistical technique, whereby one(X_1) or more independent variables(X_1, X_2,…..,X_k) are used to predict a single dependent variable (y). In this case, the optimal model is not ascertained. In a complex multiple regression situations, when there is large number of independent variables (X_1, X_2, X_3,…….,Xn) which may or may not be relevant for making predictions about the dependent variable(y), it is important to be able to reduce the models to contain only the variables which provide statistical significant information about the independent variables. To understand this regression effectively, the researchers must be aware of and uses of the diagnostics measures and plots that have been be developed for assessing the best fitted model.

There are several methods available in literature for selection of best independent variables among several predictors. The selection of the best optimal model is based on the best subset analysis and forward stepwise method using collinearity diagnoses. The forward stepwise procedures help to ascertained the significant and low significant models to the dependent variable(y) and the non significant model is removed and the model re estimated again for further transformation.

Transformations on the predictors are done by trial by error means. Various transformations are tried until one get a satisfactory model. The log Transformation used validates the regression assumption of linearity, normality and constant variance of the error term and improves regression coefficients. But deciding which variables to be included in the regression model is not always trivial, due to tedious nature of data computation; statistical software was used to evaluate the data for easy results.

Corresponding author: ***Osemeke R. F.,*** E-mail: reubitosmek@yahoo.com, Tel.: +2348065917356

1.1. Theoretical Analysis and its Applications

The Coefficient of Determination ($r^{2)}$ and Adjusted R^2 Criterion

Coefficient of Determination ($r^{2)}$ is a statistic that gives some information about the goodness of fit of a model. In a regression, r^2 is a statistical measure of how well the regression line approximates the real data points. Everitt **[1]** defines r^2 as the Coefficient of determination. R^2 is the square of the correlation between two variables. Nagelkere [2] Indicates that r^2 in the range of -1 to 1 perfectly fits the regression line. From Table 2, $r^2\ of\ 0.700$ is 70% at which the variation in the response variable can be explained by the independent variables? The remaining 30% can be attributed to unknown

The model with the highest value of r^2 provides the closet fit. However, the major drawback of r^2 is that as the model increased, r^2 goes high whether the extra variables provide any important information about the dependent variable(y) or not. Therefore, it makes no sense to define the best model as the model with the largest r^2 value. A common way to avoid this problem is to used the adjusted version of r^2 instead of r^2 itself. The adjusted r^2 statistic for a model with k explanatory variables is given by

$$r^2_{adj} = [1 - r^2_{y.26.....p})\ \frac{n-1}{n-p-1}] \qquad (1)$$

Where

P denotes the number of explanatory variables in the regression equation.

$r^2_{y.26.....p}$ denotes the coefficient of determination for full regression model

The r^2 adjusted does not necessarily increase when the numbers of predictor variables increase. It increased when the data has significant effects on the model. According to coefficient of determination ($r^{2)}$ and adjusted r^2 criterion, one should choose the model that has the largest adjusted r^2 and r^2

Some researchers suggest that the adjusted r^2 be computed to reflect both the number of explanatory variables in the model and the sample size. Adjusted r^2 for full model was computed with the above formula of equation (1)

$1 - \left[(1 - 0.700)\frac{20-1}{20-5-1}\right] = 1 - \left[(0.3)\frac{19}{14}\right].$

Adjusted $r^2 = 0.593$

Mallows Cp Statistic

Gilmour[3] and Mallows[4] sees Cp statistic as a measure for assessing the fit of a regression model that has been estimated using ordinary least squares. It is applied in the context of model selection, where a number of predictor variables are available for predicting some outcome and the goal is to find the best model involving a subset of these predictors.

Mallows[4] have suggested using Cp as the best criterion for choosing a model among alternative competitive models .The model are unbiased when Cp $\leq P + 1$. For other illustration and comments on interpretation sees Mallows[4],Goldman and Toman[5] or Daniel and Wood[6] . One disadvantage of Cp is that it seems to be necessary to evaluate Cp for all or most of the possible subsets to allow interpretation. The C_p statistic as defined by Mallows [4] is denoted by

$$Cp = \frac{(1-R_P^2)(n-T)}{1-R_T^2)} - [n - 2(p + 1)] \qquad (2)$$

Where

P denotes number of independent variables included in the regression model

T denotes Total number of parameters (including the intercept) to be estimated in the full regression model

R_P^2 denotes coefficient of multiple determinations for a regression model that has P independent variables

R_T^2 denotes coefficient of multiple determinations for a full regression model that contains all T estimated parameters.

The Cp for full model was computed with the above formula of equation (2)

N = 20 P = 5, T = 5+1 = 6, $R_p^2 = 0.700, R_T^2 = 0.700$

$C_p = \frac{(1-0.700)(20-6)}{1-0.700} - [20 - 2(5 + 1)],$

$C_p = 6$

Standard Error of the Estimate (SEE)

The standard error of estimate is another measure of the accuracy of our predictors and a means of model inclusion. It is the square root of the sum of the squared errors divided by the degrees of freedom. It represents a measure of variation around the regression line. It is also used in estimating the size of the confidence interval for the predictions. For more details see Neter,Wassermanann and Kunter [7]

The standard of estimates is defined as

$$\sqrt{\frac{sum\ of\ squares\ errors}{n-2}} = \sqrt{\frac{\sum r^2}{n-2}}. \qquad (3)$$

The standard error for full model was computed with the above formula of equation (1). The error sum of squares is denoted as 225.110 and n is 20.

Therefore $\sqrt{\frac{225.110}{20-2}}$ is 4.0099

2.0 Data Analysis with Applications of Theoretical Formula of Equation (1), (2), and (3)

Table 1: Collection of Twenty independent pharmacies in an attempt to predict prescriptions volume (sales per month). The data in Table 1 will be used to detect all the values of the best subset analysis

FLOOR_SP(X1)	PRESC_RX(X2)	PARKING(X3)	SHOPCNTR(X4)	INCOME(X5)
4900	9	40	1	18
5800	10	50	1	20
5000	11	55	1	17
4400	12	30	0	19
3850	**13**	**42**	**0**	**10**
5300	15	20	1	22
4100	**20**	**25**	**0**	**8**
4700	22	60	1	15
5600	24	45	1	16
4900	27	82	1	14
3700	28	56	0	12
3800	31	38	0	8
2400	36	35	0	6
1800	37	28	0	4
3100	40	43	0	6
2300	41	20	0	5
4400	42	46	1	7
3300	**42**	**15**	**0**	**4**
2900	45	30	1	9
2400	46	16	0	3

Source of data: Hilderland and Lyman[8]

Table 2:The output of Best Subsets Analysis for Drugstore Volume Obtained from Table 1 Analysis

NUMBER IN MODEL	R^2	C_P	ADJ R^2	STD ERROR	P + 1	VARIABLES IN MODEL	CONSIDER THE MODEL
2 *	0.647	2.474	0.606	3.947	3	X_2X_4	YES
2 *	0.666	1.606	0.626	3.842	3	X_1X_2	YES
3 *	0.663	3.75	0.599	3.98	4	$X_2X_3X_4$	YES
3 *	0.666	3.571	0.604	3.96	4	$X_1X_2X_5$	YES
3 *	0.679	2.963	0.619	3.877	4	$X_1X_2X_3$	YES
3 *	0.690	2.436	0.633	3.81	4	$X_1X_2X_4$	YES
4 *	0.681	4.909	0.595	3.998	5	$X_1X_2X_3X_5$	YES
4 *	0.693	4.318	0.611	3.918	5	$X_1X_2X_4X_5$	YES
4 *	0.698	4.062	0.618	3.88	5	$X_1X_2X_3X_4$	YES
5 **	**0.700**	6	0.593	4.009	6	$X_1X_2X_3X_4X_5$	YES

Each values of r^2 in Table 2 are high, but the full model (******) has the highest r^2 value of 0.700. Based on r^2 criterion, the full model is selected for prediction. Contrarily, the full model violates the criterion for adjusted r^2 and standard error of estimates. The full model has a lower adjusted r^2 value of 0.593 and a high standard error of 4.009. Based on adjusted r^2 value and standard error criterion, it is not eligible for model inclusion. The adjusted r^2 value of 0.63 and lowest standard error of estimates of 3.81 in Table 2 shows that model X_1, X_2 and X_4 is included for prediction. The entire model with (*) satisfies the condition $for\ Cp \leq P+1$ and are good for model inclusion according to Cp statistic. This is an indication that the Mallows Cp statistic is confusing, biased and requires some further estimation in other to get the best optimal model. This uncertainties of best subset analysis resulted to the use of stepwise estimation

3.0 Stepwise Estimation

This is another strong measure for model inclusion in line with the best subset analysis. The stepwise estimation is a more advanced, reliable and tedious ways of choosing the best fitted model for further analysis. Many researchers prefer the used of stepwise estimation as the best form of model inclusion. In building optimal model with stepwise, the steps below has to be thoroughly diagnosed to ensure easy analysis

Steps 1: The first variable included is the one that has the highest r^2 value for predicting y; assume that this variable is called x_1.or the variable that has the highest negative or positive value correlation with y is selected.

Steps 2: The second variable is the model when combined x_1 , yields the highest r^2 value; call x_2.If there is any degree of collinearity among the x's, x_2 may not have the largest r^2_{yx} value

Steps 3: The third variable included by forward selection yields the highest r^2 value when combined with x_1 and x_2. The process continues in this same manner

Step 4: The process stops when there is no additional increased of r^2 when others models are added

Step 5: The best fitted model through studentized residuals against the predictors of interest are used to test the regression assumptions of linearity, constant variance, normality and independence of the error terms

Step6: Violations of any of the regression assumptions of step 5, requires a transformation to improve on the model and to some extend validate the assumption

Step 7: Check if an outliers is detected. Any value of the standardized residuals outside the range of 2 x standard to -2 x standard deviation should be removed from the model. Alternatively, the standardized residuals, the studentized residuals and leverage points are another means of outlier detention.

Step 8: The final model is re estimated

Model Building with Stepwise Estimation:

Table 3: Correlation Matrix of Independent Pharmacies Data: Selecting the First variable

	Correlations					
Predictors	(Y)	(X1)	(X2)	(X3)	(X4)	(X5)
VOLUME(Y)	1.000	.183	-.663	-.069	-.203	.385
FLOOR_SP(X1)	.183	1.000	-.751	.504	.710	.863
PRESC_RX(X2)	-.663	-.751	1.000	-.328	-.341	-.845
PARKING(X3)	-.069	.504	-.328	1.000	.482	.393
SHOPCNTR(X4)	-.203	.710	-.341	.482	1.000	.645
INCOME(X5)	.385	.863	-.845	.393	.645	1.000

Table 3 display's the correlation among the five predictors and their correlations with the dependent variable(Y).Examination of the correlation matrix of Table 3 indicate that the predictor(X_2) is most closely correlated with Y, having a high negative correlation of -0.663. Although X_2 has a very high correlation with X_5 and X_1 causing serious multicollinearity problem, but has a very low correlation with X_3 and X_4, upon that, the first step is to build a regression equation with model X_2

Table 4: Inclusion of Model X_2 (PRESC_RX)

Multiple R	**Multiple R^2**	**Adjusted R^2**	**Std Error of Estimate**	**F Statistics**	**P Value**
0.663	0.439	0.408	4.835	14.105	0.001

Table 5: Inclusion of Model X_4 (SHOPCNTR)

Multiple R	**Multiple R^2**	**Adjusted r^2**	**Std Error of Estimate**	**T test**
0.804	0.647	0.606	3.947	15.584

The output of correlation matrix of Table 3 shows that model X_5 has the next higher correlation with Y of 0.385, but the problem with X_5 in Table 3 is that it has a collinearity problem due to high correlation with X_1(0.863), X_4(0.710), X_3(0.504) and X_2(-0.751). On that note, X_5 was not chosen for inclusion. Predictor X_4 was included into the model with X_2 because it has a low collinearity problem with X_2(-0.341), and X_3(0.482). The multiple R and r^2 values have both increased with the addition of X_4. The r^2 has increased by the 19.8%

Table 6: Inclusion of Model X_1 (FLOOR_SP)

Multiple R	**Multiple R^2**	**Adjusted R^2**	**Std Error of Estimate**	**F Statistics**	**Durbin-Watson Statistic**
0.831	0.691	0.633	3.809	11.911	2.376

With model X_1 into the regression equation, the value of r^2 in Table 6 increased by 0.27%. No additional values will be gained by adding the model X_3 and X_5

4.0 Evaluating the Final Model to Assessed the Assumptions of Regression Analysis

The final fitted model

$$Y = b_0 + b_1X_1 + b_2X_2 + b_4X_4 + ei$$

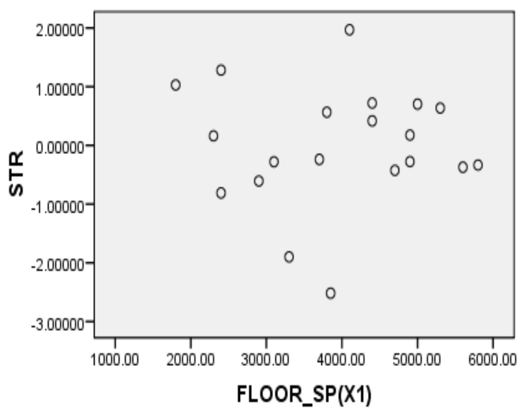
Will be used to address two issues

(i) Meeting the regression assumptions

(ii) Identifying the influential data points called outliers. Outliers are data points that are far away from the mean when analyses. Outliers are removed from a regression model when the ordered standardized residuals (Zres) is outside the range of 2 x standard deviation and -2 x standard deviation and the model is re estimated again. This iteration continues until no further outliers are detected.

The principal measure used in evaluating the regression assumption is the residuals. For comparison purpose, Mosteller and Tuckey[9] prefer the used of studentized residuals versus the predicted(y^) and the predictors($X_1,X_2,\ldots,X_k$) to display assumption validation.

Standardized Studentized Residuals (STR) Plots of the fitted model



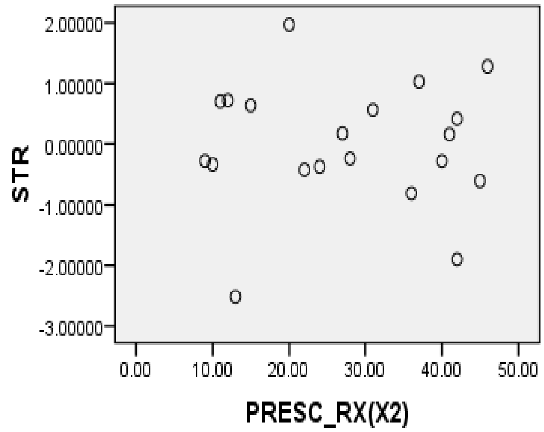
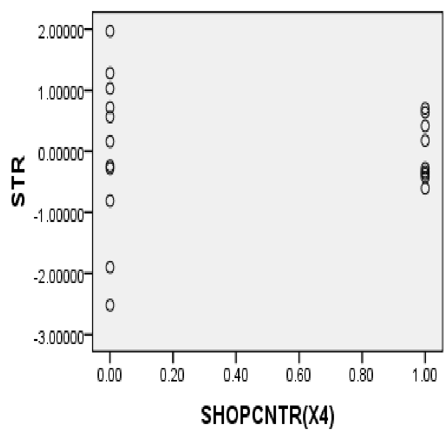


Fig 1: STR Floor_Sp Residual Plots

Fig 2: STR Presc_Rx Residual Plot



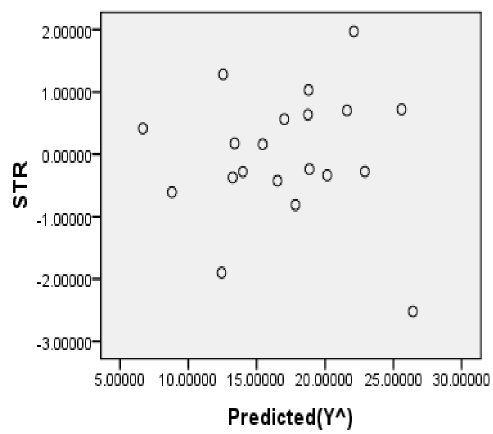


Fig 3: STR Shopcntr Residual Plot

Fig 4:STR Predicted(y^) Residual Plot

The scatter plots of Fig 1, Fig 2 and Fig 4 are quite define; thus they have strong and significant effects in the regression equation.Fig4 is less well defined , due to heteroscedasticity tendency.

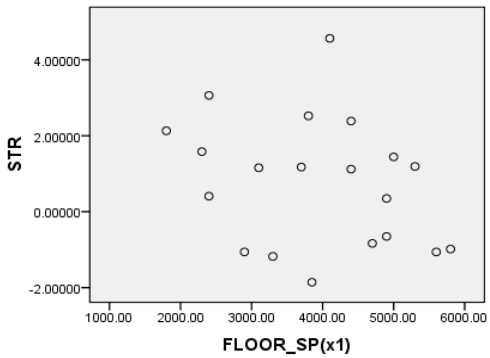
Table 7: Applying Remedies for Assumption Violations of Fig 3:The Output

Multiple R	R^2	Adj R^2	F change	
0.964	0.929	0.906	39.514	

The log transformation of predictor X_4 and re fitting the model
$Y = b_0 + b_1X_1 + b_2X_2 + b_4\log X_4 + ei$ yields the output of Table 7
In Table 2, the r^2 value is 0.700 and adj r^2 value is 0.593. Table 7 shows a greater improvement of 0.929 for r^2 and 0.906 for adj r^2. The difference in r^2 is 22.9% and adjusted r^2 difference is 31.3%. This shows that there was a greater improvement in the logarithm transformation of predictor x4. The fitted model was

$Y\hat{} = 37.826 - 0.002X_1 - 0.531X_2 + 0\text{Log}X_4$. The model X_3 and X_5 was removed from the model because it has no significant statistical contribution. The improvement are shown in the graph below

Improved Standardized Studentized Residuals(STR) Plots



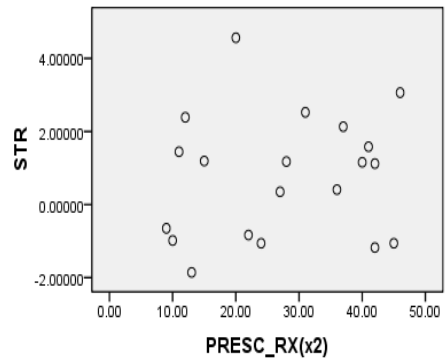


Fig 5: Improved STR Floor_Sp Residual Plot. Fig 6 Improved STR Presc_Rx residual Plot

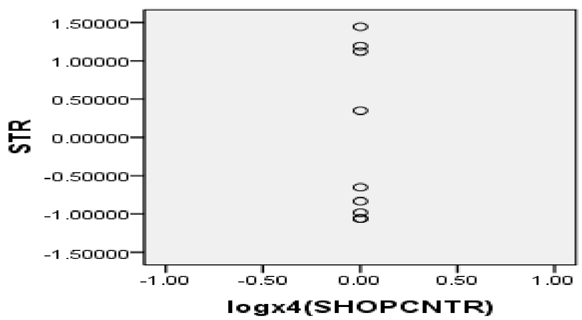


Fig 7: Improve STR LogX_4 (Shopcntr) Residual Plot

5.0 Identification of Outliers Using Numerical Test

Table 8: Detecting Outliers as Influential Observations

OBS	**(Y)**	**(X1)**	**(X2)**	**(X4)**	**Standardized residuals(ZRES)**	**Studentized Residuals(STR)**	**Leverages**	**Deleted Row**
1	22	4900	9	1	-0.238	-0.276	**0.207**	
2	19	5800	10	1	-0.302	-0.335	0.136	
3	24	5000	11	1	0.6300	0.703	0.1474	
4	28	4400	12	0	0.6318	0.72	0.181	
5	18	3850	13	0	**-2.212**	**-2.518**	0.179	**Row 5**
6	21	5300	15	1	0.5930	0.637	0.083	
7	29	4100	20	0	1.8115	**1.970**	0.105	**Row 7**
8	15	4700	22	1	-0.399	-0.42	0.067	
9	12	5600	24	1	-0.324	-0.372	0.190	
10	14	4900	27	1	0.163	0.175	0.08	
11	18	3700	28	0	-0.224	-0.238	0.065	
12	19	3800	31	0	0.520	0.57	0.104	
13	15	2400	36	0	-0.7434	-0.81	0.1064	
14	22	1800	37	0	0.8434	1.030	**0.280**	
15	13	3100	40	0	-0.2590	-0.279	0.090	
16	16	2300	41	0	0.14872	0.1615	0.1025	
17	8	4400	42	1	0.3484	0.417	**0.252**	
18	6	3300	42	0	-1.690	-1.899	0.16	**Row 18**
19	7	2900	45	1	-0.471	-0.607	**0.347**	
20	**17**	2400	46	0	**1.171**	**1.282**	**0.116**	

Table 8 represents the final fitted model using stepwise estimation. Finally row 5, row 7 and row 18 was removed from the model. Table 9, 10, 11 and 12 gives an insight on how outliers was removed from the model

Table 9: Deleting Row 5

Standard Deviation for Standardized residuals	**Deleted row**
1.836	5

Table 8 with 20 observations, has standardized residuals as -2.21171. Row 5 of standardized residuals is outside the range of 1.836 and -1.836. Row 5 was deleted from the model and the model was left with 19 observations .Re estimated with model with 19 observations gives the output of Table 10

Table 10: Regression Output with the Res Estimation of 19 Observations

Multiple R	R^2	Adj r^2	Std Error of Estimates	P value	Durbin Watson Statistic	Standard Deviation for Standardized Residuals	Deleted Row
0.902	0.813	0.776	3.0564	0.000	2.306	1.826	18

Row 18 of Table 8 was deleted from the model after res estimation. The Standardized residuals are outside the range of 1.826 and -1.826 and are regarded as an outlier. The model is left with 18 observations.

Table 11 Regression Output with the Re Estimation of 18 Observations

Multiple R	**R^2**	**Adj r^2**	**Std Error of Estimates**	**F value**	**P value**	**Durbin Watson Statistic**	**Standard Deviation for Standardized Residuals**	**Deleted row**
0.924	0.854	0.823	2.54147	**27.310**	0.000	2.262	1.814	7

The Standard deviation for standardized residuals is 1.814. *Row 7 was deleted from the mode because the value is outside the range of 1.814 and -1.814. You are left with 17 obs*ervations.

Table 12: Regression Output with the Re Estimation of 17 Observations

Multiple r	**R^2**	**Adj r^2**	**Std Error of Estimates**	**F value**	**P value**	**Durbin Watson Statistic**	**Standard Deviation for Standardized Residuals**	**Deleted row**
0.932	0.868	0.838	2.21751	28.540	0.000	2.594	1.802	18

The fitted model is $\hat{y} = 44.195 - 0.002X_1 - 0.541X_2 - 4.187X_4$. Standard deviation for standardized residuals is 1.802. Hence no further outlier was detected because no value for standardized residuals is outside the range of 1.802 to -1.802. Finally, row 5, row7 and row 18 of Table 8 was deleted from the model.

Final Residual Plots with the Non Inclusion of Row5, Row7 and Row 18 from The Model

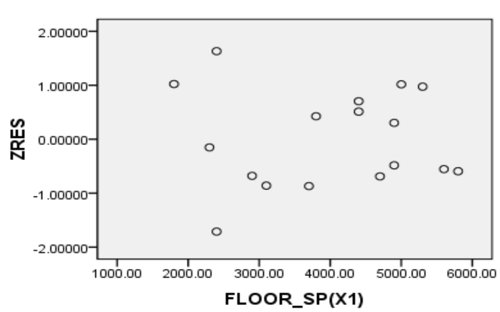


Fig 8: ZRES VERSUS FLOOR_SP

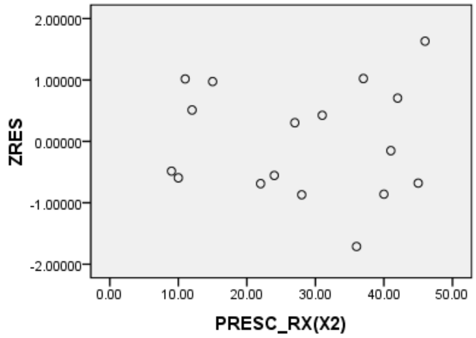


Fig 9: ZRES VERSUS PRESC_RX

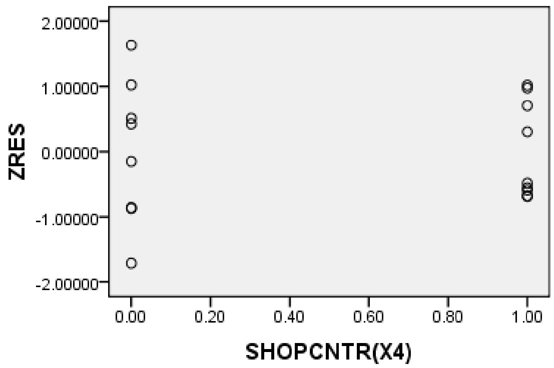


Fig 10: ZRES VERSUS FITTED

1.5.1 Alternative Method of Outlier Detention

The standardized residuals, the studentized residuals and leverage points are another means of

outliers detention.

The most widely used value is 1.96 thus, identifying statistically significant residuals as those with residuals that is greater than 1.96. For studentized residuals values, row 5 of -2.518 and row 7 of 1.97 in Table 8 is outside the range of 1.96. For standardized residuals, row 5 of Table 8 is an outlier and should be removed and the model re estimated with 19 observations.

Also for studentised residuals, the most widely used is 95 percent confidence ($\propto = 0.05$).The corresponding t value is 1.96, thus identifying statistically significant residuals as those residuals greater than 1.96 and less than -1.96. Attention should be on studentized residuals that exceed +2 or -2 and even more concerned about residuals that exceed 3.

The average value for leverage point is $\frac{P}{N}$, where p is the number of predictors (the number of coefficients plus one for the constant). The rule of thumb for situations where p is greater than 10 and the sample size exceeds 50 is to select observations with a leverage value greater than twice the average ($\frac{2P}{N}$).When the number of predictors or the sample size is less, use of three times the average ($\frac{3P}{N}$) is suggested.

Now let's look at the leverage's to identify observations that will have potential great influence on regression coefficient estimates. Generally, a point with leverage greater than (2k+2)/*n* should be carefully examined, where *k* is the number of predictors and *n* is the number of observations.

For leverage point, row1, of 0.207, row 14 of 0.280, row 17 of 0.252 and row 19 of 0.347 in Table 8 is a leverage point. Using ($\frac{2P}{N}$), no outlier was detected

6.0 Conclusion

The goal of this paper was to raise awareness of the importance of selecting the best optimal model from the set of predictors and presents a simplified introduction to the rationale and fundamental concepts underlying multiple regression analysis. It emphasizes that multiple regression analysis can describes and predict the relationship between two or more independent variables. Also, multiple regression analysis, which can be used to examine the incremental and total explanatory power of many variables, is a great improvement over the sequential analysis approach necessary with univariate techniques. Both stepwise estimation and best subset analysis were used to estimate the best optimal regression equation and uses the model through studentized residuals to assessed the validation and non-validation of the regression assumptions and detect outliers as influential observations. However, we have seen that the uses of best subset analysis always give rises to different models as the best optimal models making statistical analysis difficult, biased, and confusing. Using coefficient of determination, gave the full models as optimal model. Using Cp statistic $(Cp \leq P + 1)$ gave many models for inclusion. Using adjusted r^2 and standard error of the estimates, the model X_1, X_2 and X_4 was chosen for inclusion but when using stepwise regression the model X_1, X_2 and X_4 was chosen for inclusion and model X_3 and X_5 was removed from the model.

Finally, adjusted r^2, standard error of estimates and stepwise estimation method gave us the same result, but the stepwise estimation gave us a more precise, understanding, and more reliable than any of the tools used. Familiarity with the concepts presented in this paper will help you better understand the more complex and detailed technical presentation in other textbooks.

We therefore regard the stepwise regression as an indispensable tool for choosing best subsets regression model for validation of regression assumption. It helps to choose the best optimal model and also facilitates the job of the analysis

References

[1] Everitt, B.S. (2002). The Cambridge Dictionary of Statistics (2nd Ed.).Cambridge University Press Publisher, New York

[2] Nagelkere, N.J. (1991). A Note on a General Definition of the Coefficient of determination. Biometrices; 78(3):pp691-692

[3] Gilmour, and Steven (1996). The interpretation of Mallows Cp- statistics. Journal of the Royal Statistical Society, Series 45(1): pp45-56

[4] Mallows, C. L. (1973). *Some Comments on Cp*. Technometrics (American society for Quality) 15. Pp661-675

[5] Gorman, J.W. and Toman, R.J. (1966) Selection of variables for fitting equations to data (2nd ed.). Wliey & Sons Inc;New York.

[6] Daniel C. and F.S. Wood. (1980). *Fitting Equations to Data*. 2d ed.: New York: Wiley and Sons Inc New York,

[7] Neter, J.W, Wassermann, and Kunter .M. H.. Applied Linear Regression Models.McGraw hill International, New York

[8] David, K. Hilderland and Lyman. O. (1991). Statistical Thinking For managers (3rd ed.), Pws – Kent's publishers

[9] Mosteller, F., and Tukey,J.W(1977). Data Analysis and Regression. Addison – Wesley Publisher, New York