Nonparametric Approach to Multiple Regression

NWANKWO, CHIKE H. (Corresponding author).
Department of Statistics, Nnamdi Azikiwe University, PMB 5025, Awka. Anambra State, Nigeria.
E-mail of the corresponding author: chikeezoek@yahoo.com

OYEKA, I. C. A
Department of Statistics, Nnamdi Azikiwe University, PMB 5025, Awka. Anambra State, Nigeria.
E-mail: cyprianoyeka@yahoo.com

Abstract
Many real life problems require the use of multiple regression to characterize and solve. Sometimes however the independent variables come in forms that violate the assumptions of the parametric multiple regression. This paper developed a nonparametric approach which uses ranks of both the dependent and independent variables to achieve the objectives of multiple regression. This approach accommodates data as low as the ordinal scale and robust. A prediction procedure which is by interpolation, is also presented.

Keywords: Rank, Robust, Prediction, Interpolation, Rank Correlation

1.0 Introduction
If one has sample observations on dependant variable Y and a set of independent variables \(X_1, X_2 \ldots X_k\), and the dependent variable satisfies the assumptions necessary for the application of parametric regression methods, then one may use these methods to estimate the regression parameters and test appropriate null hypotheses. Thus if \(y_i\) is the ith observation independently made on the dependent variable Y and \(x_{i1}, x_{i2} \ldots x_{ik}\) are respectively the ith observation made on the independent variables \(X_1, X_2 \ldots X_k\) for i = 1, 2 ... n then we may fit the multiple regression model

\[ y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \ldots + \beta_k x_{ik} + e_i \]  

(1)

Where \(\beta_j\)’s are partial regression coefficients and \(e_i\)’s are independent error terms with \(E(e_i) = 0\) for \(i = 1, 2 \ldots n\). Equation 1 may alternatively be expressed in matrix form as

\[ y = X\beta + e \]

(2)

Where \(y\) is an n x 1 column vector of observations on the dependent variable Y. \(X\) is an n x r = n x (k + 1) design matrix of observations on the independent variables with full column rank r, the number of parameters in the model and \(k = r - 1\), the number of independent variables in the model; \(\beta\) is an r x 1 = (k + 1) x 1 column vector of partial regression coefficients, and \(e\) is an n x 1 column vector of error terms, with \(E(e) = 0\) [2].

Under these assumptions the usual least squares method may be used with Equation 1 or 2 to obtain unbiased estimates of the regression parameters \(\hat{\beta}\), from the normal equations [2].

\[ X'Y = X'X\hat{b} \]

(3)

in matrix form, which when solved yields

\[ \hat{b} = \hat{\beta} = (X'X)^{-1}X'y \]

(4)

Where \((X'X)^{-1}\) is the matrix inverse of XX.
If in addition Y can be assumed to be normally distributed with constant variance, then the usual parametric F test may be used to test the null hypothesis that Equation 1 or 2 fits [2][3], and other desired hypotheses.

However, if these assumptions cannot be satisfied then the parametric tests may not be validly performed. A non-parametric approach such as the following proposed method would then be indicated.

2.0 The Proposed Method
Let \(y_{ij}, x_{i1}, x_{i2} \ldots x_{ik}\) be as defined above for \(i = 1, 2 \ldots n\), where populations \(Y, X_1, X_2 \ldots X_k\) may now be measurements on as low as the ordinal scale and need not be continuous. Furthermore population Y need no longer be normally distributed for desired hypotheses to be tested.

Now instead of using the raw observations or scores in a regression model, these scores are now converted into ranks before being fitted in a regression model.

Thus let \(r_{yi}\) be the rank assigned to \(y_{ij}\), the ith observation for population Y, and \(r_{xij}\) be the rank assigned to \(x_{ij}\) the ith observation from population: \(X_{ij}\) for \(i = 1, 2 \ldots n; j = 1, 2 \ldots k\). All tied observations in each variable are, as usual, assigned their mean ranks. With these ranks we now set up a nonparametric multiple regression model of
the ranks assigned to observations from population Y regressing on the ranks assigned to observations from the
independent variables Xj’s obtaining
\[ r_{iy} = \beta_0 + \beta_1 r_{i1} + \beta_2 r_{i2} + \ldots + \beta_k r_{ik} + e_i \]
where the e_i’s are independent error terms with E(e_i) = 0. Or expressed in matrix notation
\[ r'_y = R\beta + e \]
where the e_i’s are independent error terms with \( E(e) = 0 \). Or expressed in matrix notation
\[ r'_y = (r_{iy}, r_{2y}, \ldots, r_{ny}) \] is an nx1 column vector of the ranks of Y
\[ R = \begin{pmatrix}
1 & r_{11} & r_{12} & \ldots & r_{1k} \\
1 & r_{21} & r_{22} & \ldots & r_{2k} \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
1 & r_{n1} & r_{n2} & \ldots & r_{nk}
\end{pmatrix} \]
is an n x r = n x (k + 1) design matrix of ranks; \( \beta' = (\beta_0, \beta_1, \beta_2, \ldots, \beta_k) \) is an r x 1 = (k + 1) x 1 column vector
of partial regression coefficients, not necessarily the same as those in equation 2; and e is an n x 1 column
vector of independent error terms with \( E(e) = 0 \). Notice that it is now no longer required that e is normally
distributed.
The method of least squares [2][1] may now be applied to either equation 5 or 6 to obtain an unbiased estimate \( \hat{b} \)
of the regression parameters \( \beta \) from the normal equations.
\[ R'R_{y} = R'\hat{\beta} \]
Solving equation 7 for \( \hat{b} \) we obtain the estimated partial regression coefficient as
\[ \hat{r}_y = Rb \]
Where \( (R'R)^{-1} \) is the matrix inverse of the matrix \( R'R \)
Using this estimate we have that the fitted or predicted nonparametric multiple regression equation is
\[ \hat{r}_y = Rb \]
Note that equation 7 may be expressed in terms of the ranks of the sample observations as
\[ \left( \begin{array}{c}
\sum_{i=1}^{n} r_{ij}r_{ij} \\
\sum_{i=1}^{n} r_{1j}r_{ij} \\
\sum_{i=1}^{n} r_{2j}r_{ij} \\
\sum_{i=1}^{n} r_{3j}r_{ij} \\
\vdots \\
\sum_{i=1}^{n} r_{kj}r_{ij}
\end{array} \right) = \left( \begin{array}{c}
n(n+1)/2 \\
n(n+1)/2 \\
n(n+1)/2 \\
n(n+1)/2 \\
\vdots \\
n(n+1)/2
\end{array} \right) \left( \begin{array}{c}
b_0 \\
b_1 \\
b_2 \\
b_3 \\
\vdots \\
b_k
\end{array} \right) \]
To evaluate equation 10, we note that \( r_{ij} \) and \( r_{ij} \)’s are each ranks and values of the first n positive integers
Hence
\[ \sum_{i=1}^{n} r_{ij} = \sum_{i=1}^{n} r_{ij} = \frac{n(n+1)}{2} \]
and
\[ \sum_{i=1}^{n} r_{ij}^2 = \sum_{i=1}^{n} r_{ij}^2 = \frac{n(n+1)(2n+1)}{6} \]
Also if \( \rho_{ij} \) is the Spearman’s rank correlation coefficient [4] between \( X_i \) and \( X_l \) whose \( i \)th observations have ranks \( r_{ij} \) and \( r_{lj} \) respectively, then
\[
\sum_{j=1}^{n} r_{ij} r_{lj} = \frac{n(n^2 - 1)}{12} \rho_{ij} + \frac{n(n+1)}{4} \left( \frac{(n-1)}{2} \rho_{ij} + \frac{(n+1)}{2} \right) - \frac{n^2}{2} - \frac{n}{2} \quad - - - (13)
\]
Using equations 11 – 13 in equation 10 yields, after simplification,
\[
\begin{pmatrix}
\frac{n+1}{2} \\
\frac{n+1}{2} \\
\frac{n+1}{2} \\
\frac{n+1}{2} \\
\frac{n+1}{2} \\
\rho_{1y} \\
\rho_{2y} \\
\rho_{hy} \\
\end{pmatrix} =
\begin{pmatrix}
1 & \frac{n+1}{2} & \frac{n+1}{2} & \ldots & \frac{n+1}{2} \\
0 & 1 & \rho_{12} & \rho_{13} & \ldots & \rho_{1k} \\
0 & \rho_{21} & 1 & \rho_{23} & \ldots & \rho_{2k} \\
\ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\
0 & \rho_{K1} & \rho_{K2} & \rho_{K3} & \ldots & 1 \\
\end{pmatrix}
\begin{pmatrix}
b_0 \\
b_1 \\
b_2 \\
\ldots \\
b_k \\
\end{pmatrix} - - - (14)
\]
Application of a few elementary operations on the above equation further reduces the normal equations to the simpler form
\[
\begin{pmatrix}
\frac{n+1}{2} \\
\rho_{1y} \\
\rho_{2y} \\
\rho_{hy} \\
\end{pmatrix} =
\begin{pmatrix}
1 & \frac{n+1}{2} & \frac{n+1}{2} & \ldots & \frac{n+1}{2} \\
0 & 1 & \rho_{12} & \rho_{13} & \ldots & \rho_{1k} \\
0 & \rho_{21} & 1 & \rho_{23} & \ldots & \rho_{2k} \\
\ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\
0 & \rho_{K1} & \rho_{K2} & \rho_{K3} & \ldots & 1 \\
\end{pmatrix}
\begin{pmatrix}
b_0 \\
b_1 \\
b_2 \\
\ldots \\
b_k \\
\end{pmatrix} - - - (15)
\]
Where \( \rho_{ij} = \rho_{ji} \); \( \rho_0 = 1 \); for \( i, j = 1, 2, 3, \ldots k \)

Equations 14 and 15 may be solved simultaneously or by the use of appropriate statistical package to obtain the results in equation 8, the required estimated partial regression coefficients \( b_i \). Although the regression coefficients are now easily obtained using these statistical packages, because nonparametric regression is a relatively novel concept, it may be instructive, for clearer understanding, to illustrate these calculations for cases in which \( k = 1, 2 \) and 3.

Thus for \( k = 1 \) we have from Equation 14 that
\[
\frac{n + 1}{2} \rho_{ij} = b_0 + \frac{(n+1)}{2} b_1 - - - - - (a)
\]
and
\[
\frac{\rho_{ij}}{6} + \frac{n + 1}{2} b_0 + \frac{(n+1)}{2} b_1 = b_0 + \frac{(n+1)}{2} b_1 + \frac{(n+1)}{3} b_1 \quad \text{Or}
\]
\[
\frac{(n+1)}{2} b_1 = b_0 + \frac{(n+1)}{2} b_1 - - - - - (b)
\]
Solving (a) and (b) simultaneously yields
\[
b_1 = \rho_{ij} \quad \text{and} \quad b_0 = \frac{(n+1)}{2} \left( 1 - \rho_{ij} \right) - - - - - (16)
\]
Using these values in equation 9 yields the fitted or predicted nonparametric simple linear regression line in terms of ranks as
\[
\hat{r}_y = \frac{(n+1)}{2} \left[ 1 - \rho_{ij} \right] + \rho_{ij} r_{ij} = \frac{(n+1)}{2} + \rho_{ij} \left( r_{ij} - \frac{(n+1)}{2} \right) - - - - - (17)
\]
For \( k = 2 \) the following normal equation from equation 14 may be used to obtain the required estimates of the regression coefficients;
\[
\frac{n+1}{2} = b_0 + \frac{(n+1)}{2} b_1 + \frac{(n+1)}{2} b_2 \\
\frac{n(n+1)}{12}((n-1)\rho_x + 3(n+1)) = \frac{n(n+1)}{2} b_0 + \frac{n(n+1)(2n+1)}{6} b_1 + \frac{n(n+1)}{12}((n-1)\rho_{12} + \frac{3(n+1)}{2}) b_2 \\
\frac{n(n+1)}{12}((n-1)\rho_{23} + 3(n+1)) = \frac{n(n+1)}{2} b_0 + \frac{n(n+1)(2n+1)}{6} b_1 + \frac{n(n+1)(2n+1)}{12}((n-1)\rho_{13} + \frac{3(n+1)}{2}) b_2 \\
\]

Solving these three equations simultaneously we obtain the estimates of the regression coefficients as

\[
b_1 = \frac{\rho_{1y} - \rho_{12} \rho_{2y}}{1 - \rho_{12}^2} ; \quad b_2 = \frac{\rho_{2y} - \rho_{12} \rho_{1y}}{1 - \rho_{12}^2} \\
\]

Substituting these values into (c) yields

\[
b_0 = \frac{(n+1)(1 - (\rho_{1y} - \rho_{2y}))}{2(1 + \rho_{12})} \\
\]

OR

\[
b_0 = \frac{(n+1)(1 - (\rho_{1y} - \rho_{2y}))}{2(1 + \rho_{12})} \\
\]

With these estimated regression coefficients, the fitted or estimated nonparametric bivariate regression model becomes

\[
\hat{y}_i = \frac{(n+1)(1 - (\rho_{1y} - \rho_{2y}))}{2(1 + \rho_{12})} + \frac{\rho_{1y} - \rho_{12} \rho_{2y}}{1 - \rho_{12}^2} r_{1i} + \frac{\rho_{2y} - \rho_{12} \rho_{1y}}{1 - \rho_{12}^2} r_{2i} \\
\]

For \(i = 1, 2, \ldots, n\)

As a last illustrative example, if there are \(k = 3\) independent variables, then the normal equations (equation 14) may be represented as

\[
\frac{n+1}{2} = b_0 + \frac{(n+1)}{2} b_1 + \frac{(n+1)}{2} b_2 + \frac{(n+1)}{2} b_3 \\
\frac{(n-1)}{6} \rho_{1y} + \frac{(n+1)}{2} = b_0 + \frac{(2n+1)}{3} b_1 + \frac{(n-1)}{6} \rho_{12} + \frac{(n+1)}{2} b_2 + \frac{(n-1)}{6} \rho_{13} + \frac{(n+1)}{2} b_3 \\
\frac{(n-1)}{6} \rho_{2y} + \frac{(n+1)}{2} = b_0 + \frac{(n-1)}{6} \rho_{12} + \frac{(n+1)}{2} b_1 + \frac{(2n+1)}{3} b_2 + \frac{(n-1)}{6} \rho_{23} + \frac{(n+1)}{2} b_3 \\
\frac{(n-1)}{6} \rho_{3y} + \frac{(n+1)}{2} = b_0 + \frac{(n-1)}{6} \rho_{13} + \frac{(n+1)}{2} b_1 + \frac{2n+1}{3} b_2 + \frac{(2n+1)}{3} b_3 \\
\]

Expectedly as \(k\), the number of independent variables in the model increases, manual computations become increasingly cumbersome. However, a few algebraic manipulations of the above set of four equations readily yield the estimated regression coefficients.

\[
b_2 = \frac{(\rho_{23} - \rho_{12} \rho_{13})(\rho_{3y} - \rho_{13} \rho_{1y}) - (1 - \rho_{13}^2)(\rho_{2y} - \rho_{12} \rho_{1y})}{(\rho_{23} - \rho_{12} \rho_{13})^2 - (1 - \rho_{12}^2)(1 - \rho_{13}^2)} \\
b_3 = \frac{(\rho_{23} - \rho_{12} \rho_{13})(\rho_{3y} - \rho_{12} \rho_{13}) - (1 - \rho_{12}^2)(\rho_{3y} - \rho_{13} \rho_{1y})}{(\rho_{23} - \rho_{12} \rho_{13})^2 - (1 - \rho_{12}^2)(1 - \rho_{13}^2)} \\
\]

The remaining coefficient \(b_0\) and \(b_1\) may similarly be obtained by substitution. Thus

\[
b_0 = \frac{(\rho_{13} - \rho_{12} \rho_{23})(\rho_{3y} - \rho_{23} \rho_{2y}) - (1 - \rho_{23}^2)(\rho_{1y} - \rho_{12} \rho_{2y})}{(\rho_{13} - \rho_{12} \rho_{23})^2 - (1 - \rho_{12}^2)(1 - \rho_{23}^2)} \\
b_1 = \frac{\rho_{(n+1)}(1 - (b_0 + b_2 + b_3))}{2} \\
\]

Using the estimated regression coefficients from Equations 4 – 24 we have the resulting fitted non parametric trivariate regression model as
\[ \hat{y}_i = b_0 + b_1 r_{i1} + b_2 r_{i2} + b_3 r_{i3} \]  
For \( i = 1, 2 \ldots n \)

In all cases, interest is usually in testing whether the hypothesized regression model is adequate. In other words, interest is in testing the null hypothesis that the population regression coefficients are all equal to zero with the hope of the null hypothesis being rejected. That is interest is in the null hypothesis.

\[ H_0 : \beta = Q \] versus \[ H_1 : \beta \neq Q \]

If \( H_0 \) is rejected then further interest may be in making pair-wise comparisons of the regression coefficients to determine whether they are statistically different. The adequacy of the regression model is tested using the Fisher F test which is fairly robust while the pair-wise comparisons are conducted using the student t test.

Note that in terms of ranks the total sum of squares SST for a multiple regression model is

\[ SST = n(n+1)(2n+1)/6 - n(n+1)^2/4 \]

That is

\[ SST = n(n^2 - 1)/12 \]

with \( n - 1 \) degrees of freedom

The regression sum of squares SSR is

\[ SSR = b' \left( R'R - \frac{n(n+1)^2}{4} \right) b \]

That is

\[ SSR = b' \left( \frac{R'R - n(n+1)^2/4}{b} \right) \]

With \( r - 1 = k \) degrees of freedom

The error sum of squares, SSE is the difference between the above two sums of squares, namely

\[ SSE = SST - SSR \]

These results are summarized in the following analysis of variance table (Table 1)

<table>
<thead>
<tr>
<th>Source of Variation</th>
<th>Sum of Squares (SS)</th>
<th>Degree of Freedom (df)</th>
<th>Mean Sum of Squares (MS)</th>
<th>F-Ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>Regression</td>
<td>SSR = b' \left( R'R - \frac{n(n+1)^2}{4} \right) b</td>
<td>( r - 1 = k )</td>
<td>MSR = \frac{SSR}{r - 1}</td>
<td>( F = \frac{MSR}{MSE} )</td>
</tr>
<tr>
<td>Error</td>
<td>SSE = \frac{n(n^2 - 1)}{12} - b' \left( R'R - \frac{n(n+1)^2}{4} \right) b</td>
<td>( n - r )</td>
<td>MSE = \frac{SSE}{n - r}</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>SST = \frac{n(n^2 - 1)}{12}</td>
<td>( n - 1 )</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The results of Table 1 may be used to test the null hypothesis of equation 26 that is to determine whether the hypothesized regression model is an adequate representation of the relationship between the dependent variable \( Y \) and the set of independent variables \( X_1, X_2, \ldots X_k \) represented by ranks.
As noted above, the F test is robust under fairly general conditions. Hence the null hypothesis \( H_0 \) of equation 26 is rejected at a specified significance level \( \alpha \) if the F-ratio in table 1 is greater than or equal to the tabulated F-value, that is if

\[
F \geq F_{1-\alpha; r-1, n-r}
\]

Otherwise \( H_0 \) is accepted.

If \( H_0 \) is rejected in which case not all the \( \beta_j \)'s are equal to zero, then one may proceed to use the usual student t test to further determine which pairs of these parameters may be statistically different.

Finally one may be interested in predicting values of a dependent variable on the basis of predicted or estimated ranks \( r_{iy} \) of equation 9. To do this, we may use the method of interpolation. Thus suppose \( Y_s \) and \( Y_z \) are the observed samples values of the dependent variable \( Y \) with assigned ranks \( r_{sy} \) and \( r_{zy} \) respectively. Furthermore suppose the rank predicted for a value to be predicted \( \hat{y} \) of the dependent variable \( Y \) is \( \hat{r}_y \) which is found to be closer in value to \( r_s \) than to \( r_z \), then the predicted or estimated value of \( Y \) at point or condition \( t \) namely \( \hat{y}_t \) is calculated as

\[
\hat{y}_t = \left( \frac{\hat{r}_y - r_{sy}}{r_{sy} - r_{zy}} \right) Y_s + \left( \frac{r_{zy} - \hat{r}_y}{r_{zy} - r_{sy}} \right) Y_z
\]

For \( t = 1, 2 \ldots \)

Provided the ranks specified for the observations from the independent variables for which the corresponding value of the dependent variable to be estimated or predicted are within or close to the range of the ranks assigned to the original observations for respective independent variables used in the model.

### 3.0 Illustrative Example

A researcher is interested in determining the effects age, body weight, blood pressure and packed cell volume (PCV) level have on the heartbeat of postpartum mothers, obtaining the data shown in Table 2 on a random sample of 38 postpartum mothers in a certain hospital maternity.
Table 2: Heartbeat, Age, Body weight, Blood Pressure and PCV Level of a Random Sample of Post Partion Mothers

<table>
<thead>
<tr>
<th>S/N</th>
<th>Pulse Rate of y_i (t_i)</th>
<th>Mothers Age x_i1</th>
<th>Rank of y_i (r_iy)</th>
<th>Mother weight x_i2</th>
<th>Rank of x_i2 (r_ix)</th>
<th>Systolic BP x_i3</th>
<th>Rank of x_i3 (r_i3)</th>
<th>Diastolic BP x_i4</th>
<th>Rank of x_i4 (r_i4)</th>
<th>PCV x_i5</th>
<th>Rank of x_i5 (r_i5)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>80 15.5 20 18 34 27 68 22</td>
<td>82 24 20 30 22.5 75 26</td>
<td>31 15 10 10.5 65 12.5</td>
<td>31 24 75 26.5 110 10.5</td>
<td>80 37 28</td>
<td>80 37 28 18 90 37</td>
<td>170 35.5 80</td>
<td>32 30 4.5</td>
<td>80 37 28</td>
<td>80 37 28 18 90 37</td>
<td>170 35.5 80</td>
</tr>
<tr>
<td>2.</td>
<td>86 32.5 30 30 22.5 55 10 103 3.5 63 7 31 10</td>
<td>82 24 24 10.5 60 5</td>
<td>102 1.5 68</td>
<td>20 38 33.5</td>
<td>82 24 24 10.5 60 5</td>
<td>102 1.5 68</td>
<td>20 38 33.5</td>
<td>82 24 24 10.5 60 5</td>
<td>102 1.5 68</td>
<td>20 38 33.5</td>
<td>82 24 24 10.5 60 5</td>
</tr>
<tr>
<td>3.</td>
<td>90 37 28 18 90 37 170 35.5 80 32 30 4.5</td>
<td>86 32.5 20 5 59 3.5</td>
<td>110 10.5</td>
<td>60 2.5 35 29.5</td>
<td>86 32.5 20 5 59 3.5</td>
<td>110 10.5</td>
<td>60 2.5 35 29.5</td>
<td>86 32.5 20 5 59 3.5</td>
<td>110 10.5</td>
<td>60 2.5 35 29.5</td>
<td>86 32.5 20 5 59 3.5</td>
</tr>
<tr>
<td>4.</td>
<td>82 24 24 10.5 60 5</td>
<td>102 1.5 68</td>
<td>20 38 33.5</td>
<td>82 24 24 10.5 60 5</td>
<td>102 1.5 68</td>
<td>20 38 33.5</td>
<td>82 24 24 10.5 60 5</td>
<td>102 1.5 68</td>
<td>20 38 33.5</td>
<td>82 24 24 10.5 60 5</td>
<td>102 1.5 68</td>
</tr>
</tbody>
</table>

We have analysed the data of Table 2 using the proposed nonparametric multiple regression method in this paper, this approach makes the analysis more general and avoids the often restrictive assumptions of parametric regression.

Thus we first rank the observations for each of the five variables from the smallest value ranked 1, to the largest value ranked 38. Tied values in each variable are assigned their mean ranks. The results of the ranking for each variable are shown beside the observations for that variable in Table 2. Pair-wise cross products of these ranks may now be taken to set up the normal equations (Equation 7) which may be solved simultaneously to obtain required estimates of the regression coefficients (equation 8) and hence the desired fitted nonparametric model (equation 9). However, manual computations are relatively difficult and cumbersome. Hence, using SPSS yields the model

\[
\hat{r}_i = 31.866 - 0.376 r_{i1} + 0.103 r_{i2} + 0.152 r_{i3} - 0.357 r_{i4} - 0.137 r_{i5}
\]

(32)

The corresponding analysis of variance table is
Table 3: Analysis of Variance (ANOVA) table for the data of Table 2.

<table>
<thead>
<tr>
<th>SV</th>
<th>SS</th>
<th>df</th>
<th>MS</th>
<th>F</th>
<th>P-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Regression</td>
<td>1205.743</td>
<td>5</td>
<td>241.149</td>
<td>2.207</td>
<td>0.078</td>
</tr>
<tr>
<td>Error</td>
<td>3497.126</td>
<td>32</td>
<td>109.285</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>4702.868</td>
<td>37</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The ANOVA table shows that the hypothesized regression model is adequate for a significance level of 0.1.

References
This academic article was published by The International Institute for Science, Technology and Education (IISTE). The IISTE is a pioneer in the Open Access Publishing service based in the U.S. and Europe. The aim of the institute is Accelerating Global Knowledge Sharing.

More information about the publisher can be found in the IISTE’s homepage:
http://www.iiste.org

CALL FOR JOURNAL PAPERS

The IISTE is currently hosting more than 30 peer-reviewed academic journals and collaborating with academic institutions around the world. There’s no deadline for submission. **Prospective authors of IISTE journals can find the submission instruction on the following page:** [http://www.iiste.org/journals/](http://www.iiste.org/journals/)  

The IISTE editorial team promises to the review and publish all the qualified submissions in a **fast** manner. All the journals articles are available online to the readers all over the world without financial, legal, or technical barriers other than those inseparable from gaining access to the internet itself. Printed version of the journals is also available upon request of readers and authors.

MORE RESOURCES


Recent conferences: [http://www.iiste.org/conference/](http://www.iiste.org/conference/)

**IISTE Knowledge Sharing Partners**

EBSCO, Index Copernicus, Ulrich's Periodicals Directory, JournalTOCS, PKP Open Archives Harvester, Bielefeld Academic Search Engine, Elektronische Zeitschriftenbibliothek EZB, Open J-Gate, OCLC WorldCat, Universe Digital Library, NewJour, Google Scholar