Correlation and Regression Analysis Between Visitors and Buyers to The Selling Nominal Using Least Square

Nanum Sovia*, Ria Dhea Layla Nur Karisma
* Departement of Mathematics, UIN Maulana Malik Ibrahim Malang
*nanumsovia29@gmail.com, riadhea0863@gmail.com

ABSTRACT

Regression and correlation analysis is used to know the relationship between independent variable and dependent variable. This research uses data of 19 respondent data which include independent variable on visitors \(X_1\) and buyers \(X_2\), and dependent variable of selling nominal \(Y\). The model used in this research is Multiple Linear Regression. The result of the equation is

\[ Y = 161854.707 - 15627.262 X_1 + 46061.733 X_2 \]

The result of F test is calculated result of F value equal to 0.021 with alpha 5% meaning that variable of visitor and buyer influence to nominal variable of selling nominal. While the Individual test shows that value visitor \(<\) Significance value = 0.05 < 0.575, t value buyer \(<\) Significance value = 0.05 < 0.089 that the t value on the visitor less than the Significance value which means the decision for the test received \(H_0\). While the value of t arithmetic on the Buyer is also less than the significance value for the test receive \(H_0\). Then it can be concluded that there is no significant influence on the partial test. Correlation between visitors with buyers is significant. On the otherhand, there is a significant correlation between the buyer and the nominal purchase. Also, there is a significant correlation of the buyer with the nominal purchase there is a significant correlation. So it can be concluded that the influence between variables considered above have significant correlation value.

Keyword: Statistics, Multiple linear regression, Correlation, Visitors, Buyers, Selling Nominal

Copyright © 2017 Green Technology. All rights reserved.

Corresponding Author:
Ria Dhea Layla Nur Karisma
Departement of Mathematics, UIN Maulana Malik Ibrahim Malang, Jl. Gajayana No. 50 Malang, Jawa Timur, Indonesia 65144
Email: riadhea0863@gmail.com

1. INTRODUCTION

Buying and selling activities are common business activities between peoples, involving buyers and sellers. Buyers and sellers have their own significance in keeping the economic development in a country. One of the economic institutions, BPTP Jawa Timur has played an important role in the development of community welfare, mainly, people who work as an employee in public or private corporations. As to increase profits, BPTP Jawa Timur builds an “induced” corporation named Koperasi Pegawai Republik Indonesia (KPRI) “Agromart”. Actually, KPRI was established for the benefits of government employees. As to increase profits, it is useful to exploit a statistical method to make some forecast from the data in BPTP. The common method of forecasting, which use several correlated-variables in its process, is called regression method. Using the regression method, the forecast result of marketing variables such as marketing income and marketing lost are then obtained. Also, the correlation of several variables can be obtained, such as the correlation between the number of visitors (denoted by \(X_1\)) and the selling nominal (denoted by \(Y\)), the correlation between the number of buyers (denoted by \(X_2\)) and selling nominal (denoted by \(Y\)), and the correlation between the number of visitors (denoted by \(X_1\)) and the number of buyers (denoted by \(X_2\)). According to the above motivation, Researcher titled this article by
Regression and Correlation Analysis Between Visitors and Buyers to The Selling Nominal Using Least Square Method.

2. RESEARCH METHOD
1. Source of the Data and Date of Survey:
The data for this research is obtained from Agromart BPTP Jawa Timur. There are three kinds of data, that is the number visitors, the number of buyers and selling nominal. The main sample is the corresponding data along 19 days on July 2017. The survey held from Juny 2017 through August 2017 on Agromart BPTP Jawa Timur.
2. Steps:

3. RESULTS AND DISCUSSION

Data Description
Descriptive statistics can be seen from the size of centralization and dissemination of data that can be obtained from SPSS software, while the results are as follows.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Mean</th>
<th>Standard Deviation</th>
<th>Variance</th>
<th>Minimum</th>
<th>Maximum</th>
</tr>
</thead>
<tbody>
<tr>
<td>Visitors</td>
<td>17.2105</td>
<td>4.02187</td>
<td>16.175</td>
<td>12.00</td>
<td>27.00</td>
</tr>
<tr>
<td>Buyers</td>
<td>12.0000</td>
<td>4.32049</td>
<td>18.667</td>
<td>8.00</td>
<td>21.00</td>
</tr>
<tr>
<td>Selling Nominal</td>
<td>4.4564E5</td>
<td>2.33808E5</td>
<td>5.467E10</td>
<td>2.24E5</td>
<td>9.96E5</td>
</tr>
</tbody>
</table>

Based on Table 3.1 we know that mean of visitors is 17.2105, mean of buyers is 12.0000 and mean of selling nominal is 4.4564E5. The standard deviation of visitors is 4.02187, standard deviation of buyers is 4.32049, and then standard deviation of selling nominal is 2.33808E5.

3.2 Classic Assumptions Test
3.2.1 Multicollinearity Test
The multicollinearity test can be seen through the tolerance value and VIF (Variance Inflation Factor). If the value of VIF is around the number 1 and not greater than 10, and the tolerance value with tolerance $\frac{1}{VIF}$ is close to 1, then the data under study is free from multicollinearity problems. The hypothesis for this test is as follows:
$H_0$ : the residuals have a multicollinearity
$H_1$ : the residuals have not a multicollinearity
Statistics test: $VIF_k = \frac{1}{R_k^2}$
Critical Values: reject $H_0$ if $VIF < 5$
The result of multicollinearity test can be seen in this table:

Table 3.2 result of multicollinearity test

<table>
<thead>
<tr>
<th>Variable</th>
<th>VIF</th>
<th>Tolerance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Visitors</td>
<td>5.697</td>
<td>0.176</td>
</tr>
<tr>
<td>Buyers</td>
<td>5.697</td>
<td>0.176</td>
</tr>
</tbody>
</table>

Table 3.2 shows a Tolerance value greater than 0.1 or close to 1 and it appears that the VIF value is less than 10. This means that the data does not occur multicollinearity.

3.2.2 Heteroskedasticity Test

Heteroskedasticity test is used to find out whether in a regression model there is an inequality of variance error between one observation to another. The hypothesis for this test is as follows:

\[ H_0 : \sigma_1^2 = \sigma_2^2 = \cdots = \sigma_n^2 = \sigma^2 \]
\[ H_1 : \text{minimal one of them } \sigma_i^2 \neq \sigma^2 ; i = 1,2, \ldots, n \]

Statistics Test: \[ |t_{H01}| = \frac{\hat{\beta}_1}{\text{Var}(\hat{\beta}_1)} \]

Critical Values: reject \( H_0 \) if \( |t_{H01}| > t_{\alpha, n-2} \) or \( P_{value} < \alpha \)

The result of heteroskedasticity test from SPSS program can be seen in the following Figure 3.1:

**Figure 3.1 Scatterplot**

Based on Figure 3.1 the spots appear to spread between -2 to 2 and do not form a certain pattern so it is concluded that the regression model is homoscedasticity or not heteroscedasticity.

3.2.3 Normality Test

Tests on normality in this study using Kolmogorov-Smirnov test. This test can be used to determine whether or not a normal data being studied. The hypothesis for this test is as follows:

\[ H_0 : F_0(X) = F_e(X) \] (residual normally distributed)
\[ H_1 : F_0(X) \neq F_e(X) \] (residual is not normally distributed)

Statistics Test: \[ D = \max |F_e(x) - F_0(x)| \]

Critical Value: reject \( H_0 \) if \( D > D_{\alpha} \) or \( P_{value} < \alpha \)

Normality test results from SPSS program can be seen in the following picture:

**Table 3.3 Descriptive statistics**

<table>
<thead>
<tr>
<th>Skewness</th>
<th>Kurtosis</th>
</tr>
</thead>
<tbody>
<tr>
<td>Statistik</td>
<td>Std. Eror</td>
</tr>
<tr>
<td>0.478</td>
<td>0.524</td>
</tr>
</tbody>
</table>
Table 3.3 if the skewness and kurtosis ratio are between -2 to +2, then the data distribution is normal. Skewness ratio is the skewness value divided by the standard error skewness, while the Kurtois ratio is the kurtois value divided by the standard error kurtois.

It is seen that the ratio of Skewness is $0.478 / 0.534 = 0.9122137405$ and Kurtosis ratio is $0.265 / 1.014 = 0.2613412229$. Since the Skewness and Kurtosis ratios are between -2 to +2, it can be concluded that the data are normally distributed.

3.2.4 Autocorrelation Test

Autocorrelation is a correlation between data or observation time series or space series. Testing Autocorrelation problem using Durbin-Watson test with this hypothesis as follows:

$H_0: \rho = 0$ (there is autocorrelation in the residual)

$H_1: \rho \neq 0$ (no residual autocorrelation)

<table>
<thead>
<tr>
<th>Tabel 3.4 Autocorrelation Test</th>
</tr>
</thead>
<tbody>
<tr>
<td>Durbin Watson</td>
</tr>
<tr>
<td>2.547</td>
</tr>
</tbody>
</table>

Tabel 3.4 explains that the value of Durbin Watson is 2.547. Based on Durbin Watson table it can be seen that the value of $d_U$ is 1.5355 and $d_L$ is 1.0743, so the hose from $dw > d_L$ is 2.547 > 1.0743. Based on the calculation can be seen that the value of Durbin Watson then there is no positive autocorrelation.

3.3 Multivariate Regression Analysis

Using Multivariate Regression Analysis, the multivariate regression model corresponding to the data is $\hat{Y} = 161854.707 - 15627.262 X_1 + 46061.733 X_2$. From F-test, the resulting F is 0.021 with alpha value 5%. From the above survey, it can be concluded that the number of visitors and the number of buyers have a significant influence to the number of selling nominal. On the otherhand, partial-test concluded that resulting t of visitors < significance value $= 0.05 < 0.575$, the result t of buyers < significance value $= 0.05 < 0.089$ that is, the value of the result t of the number of visitors is least than the significance value. Then, we can concluded that hypothesis $H_0$ is accepted. On the otherhand the value of t hitung corresponding to the number of buyers is also least than the significance value. Then, the hypothesis $H_0$ is then accepted. From the above explanation, it can be concluded that, there is no significance from each variable in partial. The correlation between visitors and buyers, on the otherhand, there is a significance correlation between the number of visitors and the selling nominal, and between number of buyers and the selling nominal. Finally, it can be concluded the variables given in this survey are related, by some significance correlations, to each other.

3.4 Correlation Analysis

In correlation test in this study only used Pearson correlation coefficient. Then will be calculated correlation value with SPSS software, the calculation with SPSS:

<table>
<thead>
<tr>
<th>Tabel 3.5 Result of Correlation</th>
</tr>
</thead>
<tbody>
<tr>
<td>[24] Significance</td>
</tr>
<tr>
<td>[32] Selling Nominal</td>
</tr>
<tr>
<td>[37] Significance</td>
</tr>
</tbody>
</table>
Based on table 3.5, correlation between visitors with buyers is significant. On the other hand, there is a significant correlation between the buyer and the nominal purchase. Also, there is a significant correlation of the buyer with the nominal purchase. There is a significant correlation. So it can be concluded that the influence between variables considered above have significant correlation value.

CONCLUSION

1. The multivariate regression equation corresponding to the variables is $\hat{Y} = 161854.707 - 15627.262 X_1 + 46061.733 X_2$.

2. The correlation between the number of visitors and the number of buyers has a significance 0.000<0.05 which means that there is a significant correlation between them. In the other side, the correlation between the number of visitors and the selling nominal has a significance value 0.028<0.05 which means that there is a significant correlation between them. It then can finally be concluded that all the variables considered above have significant correlation to each other.

REFERENCES