USING THE ARIMA FORECASTING METHOD

Haryadi Sarjono*
Bina Nusantara University, Jakarta, Indonesia

ABSTRACT

The purpose of this research was to forecast sales in the next period of a mini-market called Segar, which sells a variety of essential products to its local community in the residential area of Bogor, West Java. The focus product is bottled mineral water 600 ml of a particular brand. This research uses quantitative data, the data collection techniques being library research, field research, interviews and observations. Data processing analysis provides useful information in answering statistical problems. Analysis techniques include the method known as ARIMA (Autoregressive Integrated Moving Average), using SPSS and Minitab software. Versions of ARIMA used in this research are ARIMA 1.0.0, ARIMA 0.0.1, and ARIMA 1.0.1. Using these methods, the forecasting result for the month of November 2013 was 235.750 or 236 bottles of this mineral water. The best approach was found to be ARIMA 1.0.0 because the calculated result was obtained by MS (Means of Square), the smallest amounting to 328.89.

Keywords: forecasting, Arima, Minitab, mineral water

บทคัดย่อ

วัตถุประสงค์ของงานวิจัยนี้ เพื่อพยากรณ์ยอดขายในอนาคตของร้านค้าขนาดย่อมชื่อเซก้า ซึ่งขายสินค้าที่จำเป็นแก่ชุมชนท้องถิ่นในเขตโบกอร์ ชวาตะวันตก สินค้าที่ใช้ในการศึกษาครั้งนี้ คือ น้ำแร่ยี่ห้อหนึ่ง งานวิจัยนี้ใช้ข้อมูลเชิงปริมาณ โดยเก็บข้อมูลจากการค้นคว้าจากห้องสมุด การลงพื้นที่ การสัมภาษณ์ และการสังเกต การวิเคราะห์ประมวลผลข้อมูลให้ผลที่เป็นประโยชน์ในการตอบปัญหาทางสถิติ เทคนิคที่ใช้เพื่อวิเคราะห์ รวมถึง ARIMA, SPSS, และ ซอฟแวร์ Minitab โดยได้ผลพยากรณ์ยอดขายน้ำแร่จำนวน 236 ขวด สำหรับเดือนพฤศจิกายน 2556 ทั้งนี้เทคนิคที่ให้ผลดีที่สุด คือ ARIMA 1.0.0 โดยมีค่าเฉลี่ยความผิดพลาดที่น้อยที่สุด

* Mr. Sarjono ST, MM, ME is a lecturer in the Management Department, School of Business Management, and a book author. Email: haryadisarjono@yahoo.com

INTRODUCTION

Technological developments facilitate the obtaining of information through various media, needed by the corporate world for competitive advantage (Janvier, 2012). Consumer demands increase, not only for cheapness, product quality, and a variety of products, but for the ready availability of the product (Nyoman, 2005). This necessitates speedy response by producers, sales agents and retailers of products and services (Schroeder, 2007). The corollary is that companies must have adequate inventory with effective plans in order to meet consumer demands (Onawuni et al. (2011). To be able to have a well-planned stock or supplies, a company has to undertake skilled forecasting to predict consumer demand in the future or the next period (Rangkuti, 2005).

According to Gaspersz, (2002) forecasting can be defined as 'the art and science to predict events in the future, while the forecasting activity is a function of a business that tries to predict the sales of the business and the use of a product so that the product can be made in the right quantity'. Planned effective forecasting methods which are precise and accurate, are essential so that consumer needs can be met on time (Wisner and Stanley, 1994).

The Segar shop is a mini-shop located in a residential area in Bogor, West Java, which sells a variety of everyday products needed by the community, one product being bottled mineral water 600 ml of a particular brand. Mineral water is one of the best-selling products, and therefore forecasting is required at the Segar store to meet customer needs, especially this particular 600 ml bottle (Wacker and Lummus, 2002). The data used in this study is the sales data for January to October 2013.

The issues of this research are: 1). How big is the demand forecast for bottled mineral water 600ml in store for the month of November 2013, using the ARIMA program; 2). Which of the ARIMA methods produces the smallest value of MS (Means of Square) (Findlay and Martin, 2006).

BASIC THEORY

Understanding Operations Management

According to Heizer and Render (2011), operations management is a set of activities that produce value in the form of goods and services by transforming inputs into outputs. According to Subagyo (2000), operations management is the application of management science to organize production activities or operations to be performed efficiently. While according to Chase et al. (2006), operations and production management can be defined as a process of continuous and effective use of management functions to integrate vari-

ous resources efficiently in order to achieve the goal.

Understanding Forecasting

Understanding forecasting according to some experts, now follows.

- According to Ahmed et al. (2011) simple forecasting is not 'the truth', because it predict the future based only on estimations, while proper forecasting uses mathematical calculations for material consideration.
- According to Sarjono (2010) prediction is made in a simple conjecture about what will happen in the future, based on the information available at that moment.
- According to Supranto (1983), forecasting derived from the word 'prediction' which means that basically forecasts are assumptions or estimations regarding the occurrence of an event in the future.
- According to Nasution (2006), forecasting is the process of estimating how much needs are needed in the future that cover the needs in matters of quantity, quality, time and location that is needed in order to meet the demand for goods or services.
- According to Tampubolon (2004), forecasting is the use of data to outline upcoming events in determining the desired target.
- According to Heizer and Render (2011) forecasting is a process of a variable (events) in the future with variable data that have connection from the past events.
- Golden, Milewicz and Herbig (1995) define forecasting as follows: "*Forecasting is predicting, projecting, or Estimating some future event or condition roommates is outside an organization's control and provides a basis for managerial planning*"
- According to Hasibuan (2011), forecasting method is a way to predict or estimate quantitatively and qualitatively about what's going to happen in the future based on the relevant data from the past.

Understanding ARIMA

According to Heizer and Render (2009) ARIMA (*Autoregressive Integrated Moving Average*), is basically using the time series function, which requires a model approach to early identification and assessment of its parameters. An example is are forecasting foreign currency exchange rates.

Forecasting calculations using ARIMA with the Minitab program (known as *Box-Jenkins*) (Mulyono, 2000) is rarely used. Usually calculations with the ARIMA approach is often used in research by students majoring in Mathematics, Faculty of Mathematics and Natural Sciences, but seldom used by students in a Faculty of Economics (Accounting and Management)

From the results of previous ARIMA studies it can be concluded that:

- ARIMA is a non-theory model, because the variable used are the previous values with error values (Findlay and Martin, 2006).
- ARIMA has a level of forecasting accuracy that is quite high because after evaluating it with the level of measurement error in forecasting, MAE (Mean Absolute Error) value is close to zero (Abu-Shikah et al., 2011).
- Results have shown that the accuracy of forecasting insolvency firm with NNR method (*neural network regression*) is high at 91.67%; for the second order ARMA models the accuracy is 83.33%, and 66.37% for the Logit models, the worst result being for a model of OLS (*Ordinary Least Squares*) at only 58.33% accuracy (McElroy and Gagnon, 2008).

Table 1: Terms of ARIMA details

No	Explanation	ARIMA Terms
1	Short Term	Accurate
2	Long Term	Tends to be flat (data/constant) less accurate
3	Data	Dependent Variable
4	Observation	Time Series → Box-Jenkins
5	Classification Models	There are 5 Models: Auto Regressive (AR), Moving Average (MA) combined or mixed models of AR and MA (ARMA), ARIMA
6	Time Series Data	Stationary
7	Characteristic of Stationary	In the average value (mean) and the variance for each period.
8	Analysis	Probabilistic or stochastic from time series data
9	Minimum Data	72 Time Series Data (Firmansyah, 2000)

Source: Developed and processed by the author

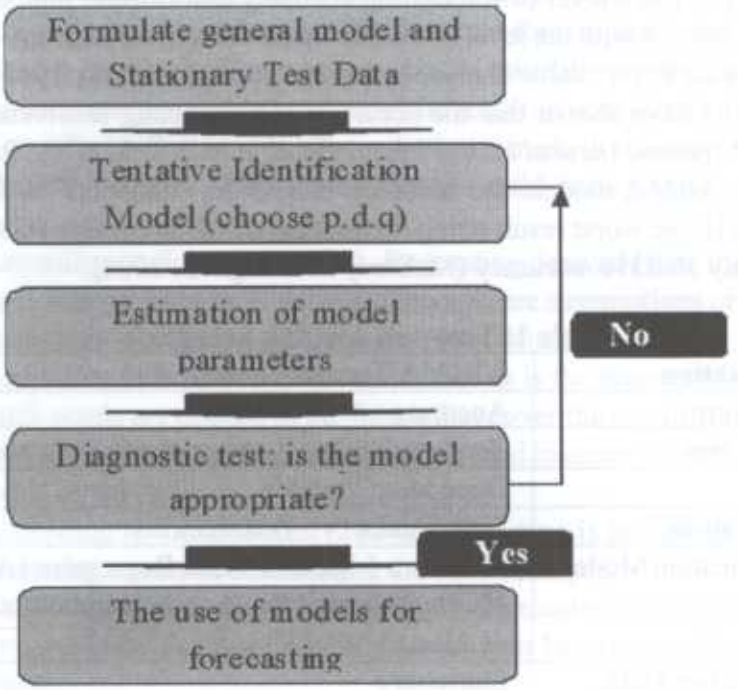
Calculation steps of the ARIMA Method

Important steps that must be done if using the ARIMA forecasting method:

1. Map the data over time, and examine the nature of trends, stability variance, and the presence of seasonal components, to determine the shape of the trans-formation linearity trend (if not linear), stability variance (if not stable), and elimination of the seasonal component (if any).
2. Calculate ACF and PACF and describe the korelogram of the original data and the data transformation results, to examine the order of differentiation and auto regression.
3. Create those models of ARIMA **(k, q, p)** that are likely to be suitable for the data held.
4. Perform estimation of parameters for each model built.
5. Perform a variance analysis or residual analysis to determine the forecast

model that will be used. The forecast model used is the significant model with the smallest standard error.

6. If necessary, try alternative models.



Source: Box and Jenkins (1976).

RESEARCH METHODOLOGY

Types and Sources of Data

Quantitative data is used for this analysis. The data obtained is the sales data of mineral water bottles 600 ml, with a particular brand of mini market, Segar store in Bogor, West Java, based on the number of monthly sales in the period from January to October 2013.

Data Collection Techniques

Data was collected by the researcher from library research, field research, interviews and observations.

Data Analysis Techniques

Analysis is the activity of processing data until it become useful information in answering statistical problems. The data analysis technique used is the ARIMA method, with the help of software SPSS and Minitab. However, previous calculations have to be done to examine the data in order to allow it to be counted. A process of differentiation and

testing determines the coefficient auto regression correlogram. To answer the existing problems and the test the hypothesis, analytical techniques are used, in the following steps:

Checking on Stationary of Data
Differencing Process
Determining the value of **p, d,** and **q** in ARIMA
ARIMA Model Parameter Estimation
Forecasting
Measurement errors in forecasting

RESULTS AND DISCUSSION

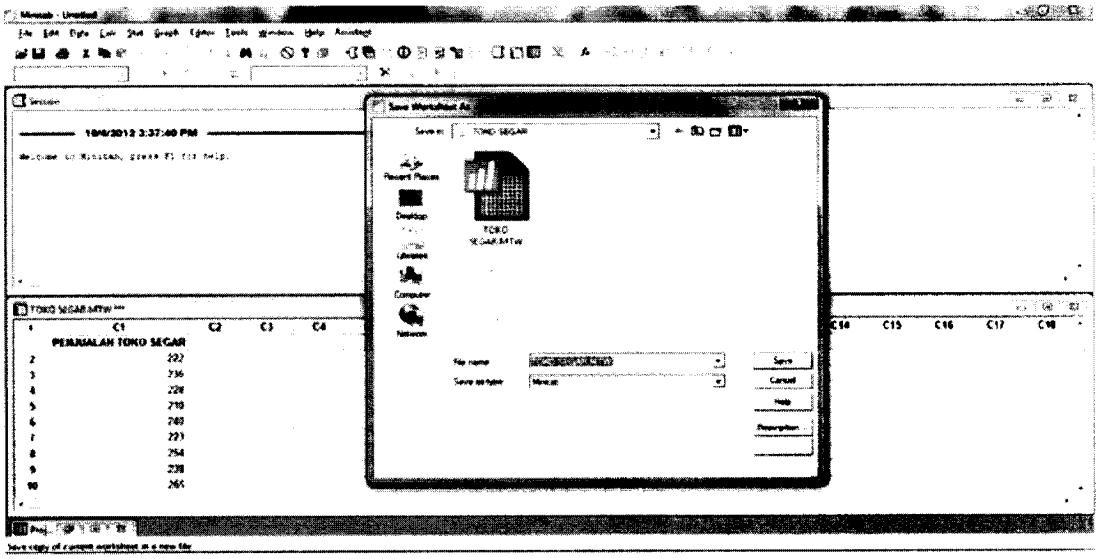
Data was obtained from the Segar Store for its sales of the bottled mineral water in the first 10 months of 2013, as presented in the following Table 2.

Table 2: Sales Data for the previous 10 months

Months	Sales
January	215
February	222
March	235
April	228
May	210
June	240
July	223
August	254
September	238
October	265
November	?

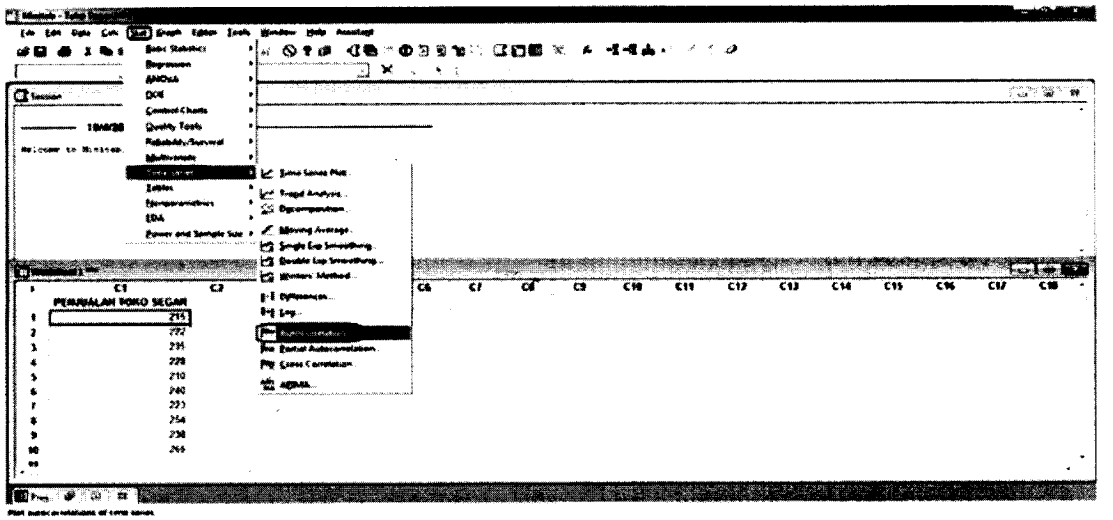
The process of step by step in using the Minitab Program, is now described.

1. Activate the Minitab Program, and then copy the data from the sales data table of Segar store in column C1. Name the column right between C1 and the first sales "Segar Store Sales" → Enter. After that, save the file in the desired folder with the format (File Name). MTW (Picture 1).



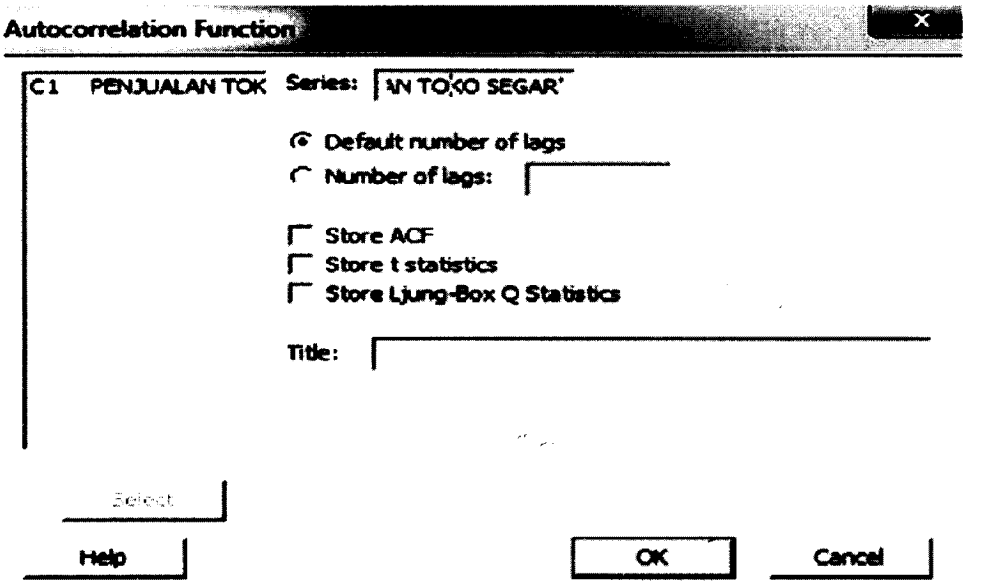
Picture 1. Save data

2. Perform identification of the data to see the pattern of the data whether it needs to do differencing by looking at the auto correlation. Steps are as follow (Picture 2).



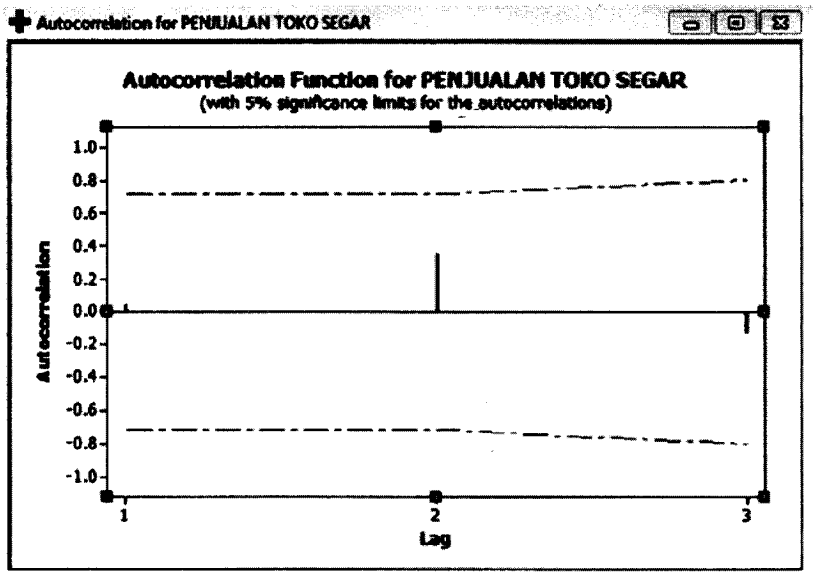
Picture 2. Autocorrelation test step

3. Input the variable "Segar Store Sales" to the series box by double clicking or select. Ignore other options beside default number of flags → OK (Picture 3).



Picture 3. Windows autocorrelation function

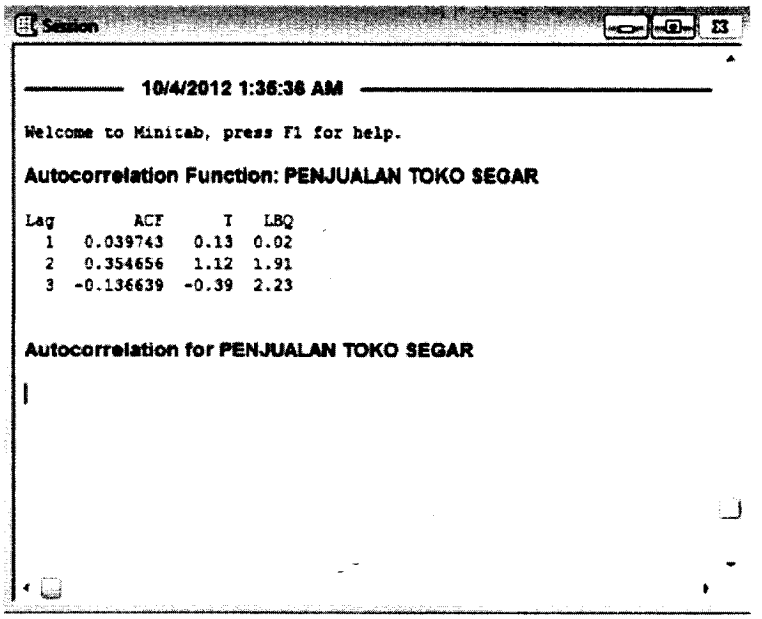
4. The resulting output shows that there is no need to do differencing. It can be seen from the graph of Windows autocorrelation Segar Store Sales, that lag or has no blue bar that run through the dotted line in red (Picture 4):



Picture 4. Chart autocorrelation

In addition, it can be seen from the Windows session ACF, that the numbers in the Table are no more than 0.5 and the numbers in the table T & LBQ tend to be small in value.

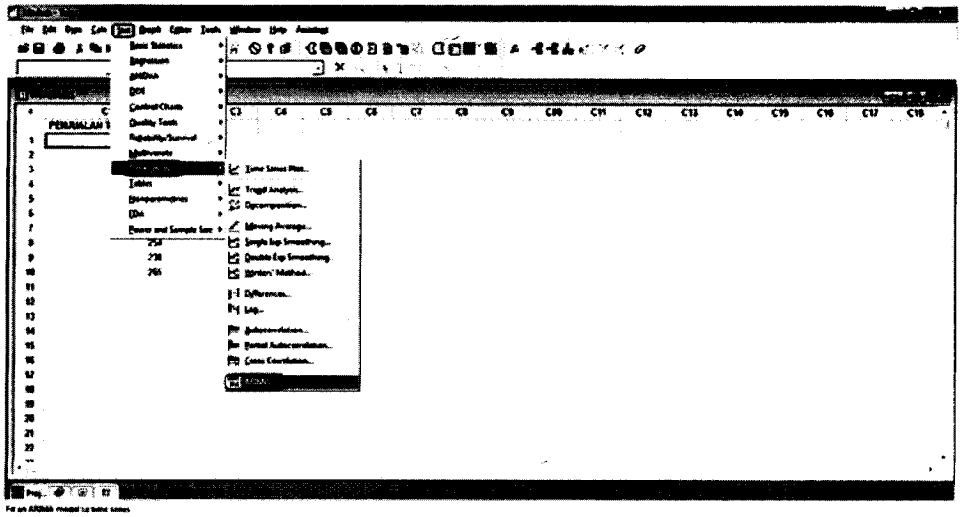
(Picture 5)



Picture 5. output autocorrelation

5. Because it is not necessary to do differencing, then the method of ARIMA (p, d, q) is given the value 0 (d = 0). Thus the models that can be used are:
   ARIMA (1,0,0)
   ARIMA (0,0,1)
   ARIMA (1,0,1)

The ARIMA model input to the above process is as follows (Picture 6):



Picture 6. ARIMA Steps

In inputting data to Windows ARIMA, note that there are three steps, where the third is different when the numbers are in the first section (p) such as ARIMA *(1,0,0)*; the numbers on the last (q) such as ARIMA *(0,0,1)*; and the numbers on the first and last (p & q) such as ARIMA (1,0,1)

The result is as follows (Picture 7):

ARIMA Model: PENJUALAN TOKO SEGAR

```
Estimates at each iteration

Iteration       SSE     Parameters
        0   2633.77   0.100   209.790
        1   2633.41   0.089   212.389
        2   2633.36   0.085   213.342
        3   2633.35   0.083   213.703
        4   2633.35   0.083   213.840
        5   2633.35   0.082   213.892
        6   2633.35   0.082   213.912

Relative change in each estimate less than 0.0010


Final Estimates of Parameters

Type          Coef  SE Coef       T       P
AR   1      0.0824   0.4669    0.18   0.864
Constant  213.912    5.958   35.90   0.000
Mean      233.123    6.493


Number of observations:  10
Residuals:     SS =  2631.14 (backforecasts excluded)
               MS =   328.89  DF = 8


Modified Box-Pierce (Ljung-Box) Chi-Square statistic

Lag             12   24   36   48
Chi-Square       *    *    *    *
DF               *    *    *    *
P-Value          *    *    *    *
Forecasts from period 10

                          95% Limits
Period  Forecast     Lower     Upper  Actual
    11   235.750   200.198   271.303
```

Later in the Windows session will appear like this.

Picture 7. ARIMA Output (1,0,0)

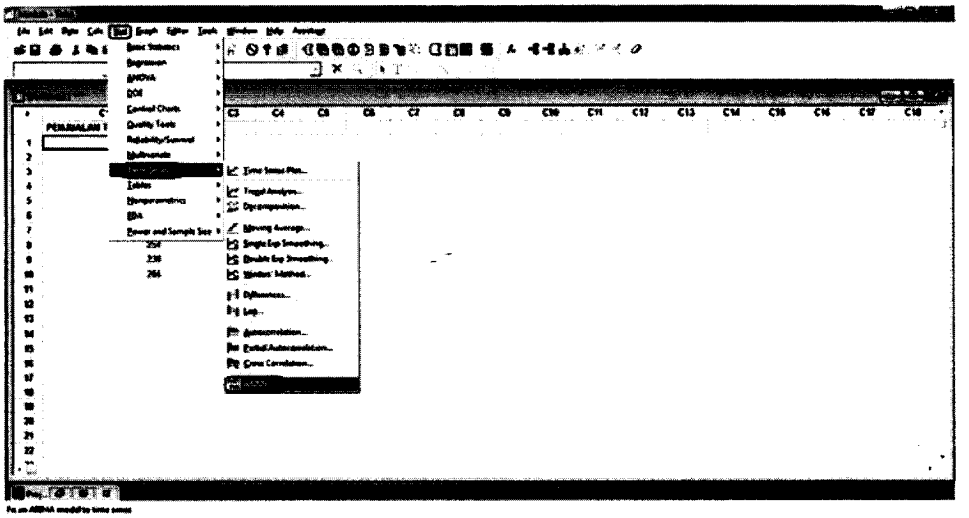
A comparison between all models is made from the amount of **MS** *(Means of Square)*, so that in the results obtained, ARIMA (1,0,0) has a MS of **328.89**, while the equation calculation obtained from the constant coefficients and coefficients of **AR** is:

$$Y_t = 213.912 + 0.0824\ Y_{(t-1)}$$

According to the ARIMA forecasting method (1,0,0), the next period which is period 11 (November) amounted to 235 750. To prove its accuracy, it can be done manually by entering the count $Y_{(t-1)}$ which is the data for October, so that it becomes:

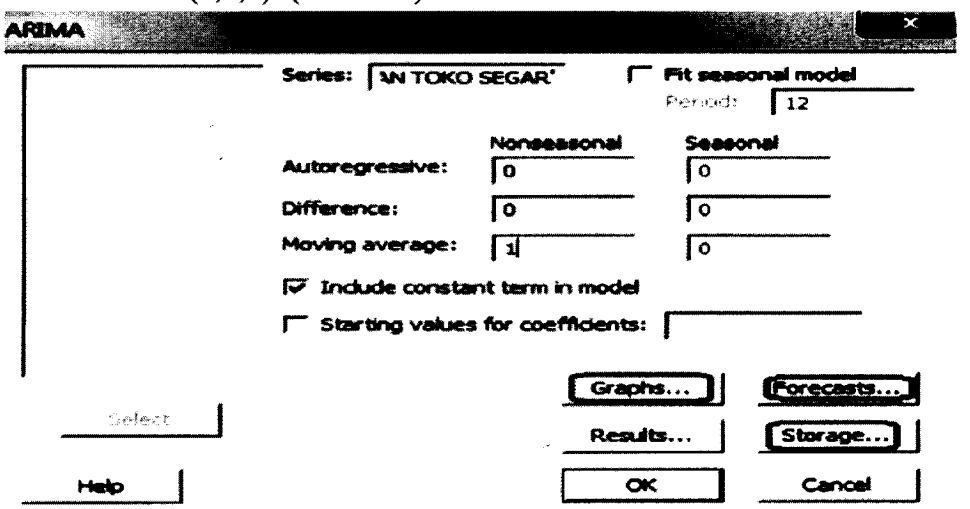
$$Y_t = 213.912 + 0.0824\ Y_{(t-1)}$$
$$Y_t = 213.912 + 0.0824\ (265)$$
$$\mathbf{Y_t = 235.748}$$

6. The next process is to process ARIMA (0,0,1) with the same steps (Picture 8).



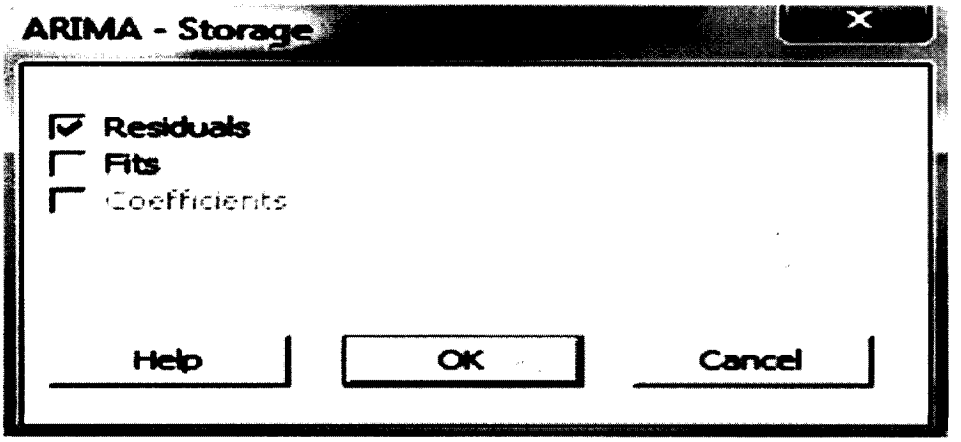
Picture 8. Steps ARIMA

But in this model, options were added to see the residual Storage options that serve to become Y_{t-1} specifically for the ARIMA model, where **q (p, d, q)** changed, or ARIMA (0,0,1). (Picture 9)



Picture 9. ARIMA (0,0,1)

7. Choose the Graphs and Forecast the same as determining ARIMA (1,0,0). Then select the option of storage and select residual → OK. After returning to the ARIMA box click OK to see the results of the processed data. (Picture 10)



Picture 10. Storage

8. Looking at the processed result of Minitab (Picture 11).

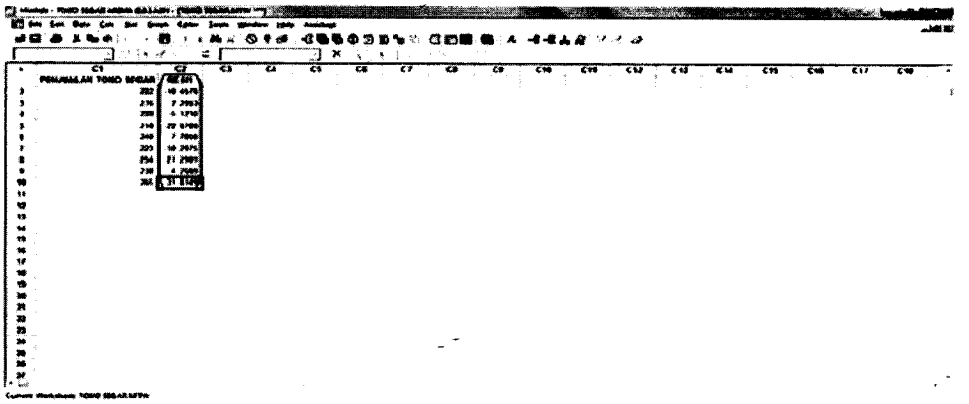
```
Relative change in each estimate less than 0.0010


Final Estimates of Parameters

Type          Coef  SE Coef      T      P
MA    1    -0.0326   0.4626  -0.07  0.945
Constant  233.046    6.109  38.15  0.000
Mean      233.046    6.109


Number of observations:  10
Residuals:     SS =  2638.21 (backforecasts excluded)
               MS =  329.78  DF = 8


Modified Box-Pierce (Ljung-Box) Chi-Square statistic

Lag            12   24   36   48
Chi-Square      *    *    *    *
DF              *    *    *    *
P-Value         *    *    *    *


Forecasts from period 10

                          95% Limits
Period  Forecast     Lower     Upper  Actual
    11   234.085  198.484   269.685
```

Picture 11. ARIMA Output (0,0,1)

The **MS** result of ARIMA (0,0,1) is greater than the **MS** of ARIMA (1,0,0) amounting to 329.78, because of negative **MA** which means that in the equation ormula do not use Y_{t-1} but use e_{t-1}, while the forecasting equation is:

$$Y_t = 233.046 - (-0.0326\ e\_(t-1))$$

e_{t-1} Residuals can be seen in the table that appears in the processing of the right table → Segar Store Sales. (Picture 12)

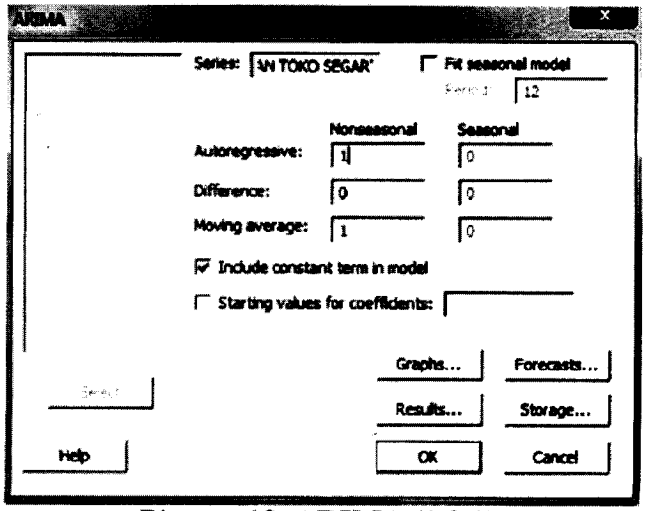


Picture 12. Residual Table

t-1 means that the data before the desired data if calculated manually, so it will look like:

$$Y_t = 233.046 - (-0.0326\ e_{t-1})$$
$$Y_t = 233.046 + (0.0326?31.8149)$$
$$Y_t = 234.082$$

9. The next step is to determine ARIMA (1,0,1). The steps are by STAT' TIME SERIES → ARIMA. Then input the method. (Picture 13)



Picture 13. ARIMA (1,0,1)

Use the same steps that have been performed before. The results are as follow (Picture 14):

```
Final Estimates of Parameters

Type          Coef  SE Coef      T      P
AR    1     0.8476   1.5180   0.56  0.594
MA    1     0.5151   1.6130   0.32  0.759
Constant    35.739    8.885   4.02  0.005
Mean        234.49    58.29


Number of observations:  10
Residuals:     SS =  2407.77 (backforecasts excluded)
               MS =  343.97  DF = 7


Modified Box-Pierce (Ljung-Box) Chi-Square statistic

Lag             12   24   36   48
Chi-Square       *    *    *    *
DF               *    *    *    *
P-Value          *    *    *    *


Forecasts from period 10

                          95% Limits
Period  Forecast    Lower    Upper  Actual
    11   246.183  209.825  282.541
```

Picture 14. ARIMA Output (1,0,1)

MS from ARIMA (1,0,1) and the result **343.97** are similar:

After all calculations have been completed, choose from the smallest value of MS

Method	MS	Forecast
ARIMA (1,0,0)	328,89	235,750
ARIMA (0,0,1)	329,78	234,085
ARIMA (1,0,1)	343,97	246,183

Choose the method of ARIMA (1,0,0) in determining the best for forecasting using an ARIMA method, because the result of the calculation for MS (Means of Square) is the smallest, and the forecast for Segar Store sales in November amounted to 235.750 or 236 units of bottled mineral water.

CONCLUSION

Based on the results and discussion above, there are a number of conclusions related to the research topic:

1. Based on calculations with the three ARIMA methods, the forecasting for November 2013 amounted to 236 bottles of this mineral water, when applying ARIMA (1.0.0) to the historical data from January to October 2013.
2. The third ARIMA method (1.0.0) is the best for forecasting because in the results of the calculations for MS (Means of Square) it had the smallest error, amounting to 328.89

REFERENCES

Abu-Shikhah, N., Fawwas, E., & Aloquili (2011). Medium-Term Load Forecasting using multivariable linear and non-linear regression. *Journal of Smart Grid and Renewable Energy,* 2, 126-135.

Ahmed, F., Shah, H., Rasa, I., & Saboor, A. (2011). Forecasting Milk Production in Pakistan, *Social Sciences Institute, National Agricultural Research Centre,* Pakistan. 24(1-4).

Box, G. & Jenkins. (1976). *Time series analysis,* forecasting and control. San Francisco, Holden-Day.

Chase, R.B., Jacobs, F.R., & Aquilano, N.J. (2006). *Operations Management for Competitive Advantage with Global Cases.* New York: McGraw-Hill.Robbins.

Findley, D.F. & Martin, D.E.K; (2006). Frequency Domain Analyses of SEATS and X-11/12-ARIMA Seasonal Adjustment Filters for Short and Moderate-Length Time Series. *Journal of Official Statistics,* 22(1) 1-34.

Gaspersz, V. (2005). *Production planning and inventory control,* Jakarta, Berdasakarn Pendekatan.

Golden, M. J. & Herbig, P. J. (1995). Forecasting Trials and Tribulations. *Management Decision,* 32(1) 33-36.

Hasibuan, M.S.P. (2011). *Manajemen Sumber Daya Manusia,* Jakarta, Aksara.

Heizer, J. & Render, B. (2009). *Manajemen Operasi,* (terjemahan), Buku 1, Edisi 9, Jakarta, Salemba Empat.

Heizer, J. & Render, B. (2011). *Operations Management.* New Jersey: Pearson.

Janvier, J.A.M (2012). A New Introduction to Supply Chains and Supply Chain Management: Definitions and Theories Perspective. *International Business Research Journal,* 5(1), 194-207

Mulyono, S. (2000). Peramalan Harga Saham dan Nilai Tukar: Teknik Box-Jenkins. *Jurnal Ekonomi dan Keuangan Indonesia,* XLVIII (2).

McElroy, T. & Gagnon, R. (2008). Finite Sample Revision Variances for ARIMA Model-Based Signal Extraction. *Journal of Official Statistics,* 24(3), 451-467.

Nasution, A.H. (2006). *Manajemen Industri,* Yogyakarta, Andi.

Nyoman, P. (2005). *Supply Chain Management,* Edisi Pertama, Guna Widya, Surabaya, Indonesia.

Onawuni, A.S., Oluleye, O.E., & Adebiyi, K.A. (2011). An Economic Order Quantity Model with Shortages, Price Break and Inflation. *International Journal Emerging Sciences,* 1(3): 465-477.

Rangkuti, F. (2005). *Manajemen Persediem,* Jakarta, PT Raja Grafindo Persada.

Render, B., Stair, R., & Hanna, M.E. (2012). *Quantitative Analysis for Management.* Essex, Pearson.

Sarjono, H. (2010). *Aplikasi Riset Operasi,* Salemba Empat, Jakarta, Indonesia.

Sarjono, H. (2011). *SPSS vs LISREL, sebuah pengantar aplikasi untuk riset,* Jakarta, Salemba Empat.

Schroeder, R.G. (2007). *Operations Management: Contemporary Concepts and Cases* New York: McGraw-Hill.

Subagyo, P. (2000). *Statistik Indusktir,* Yogyakarta, BPFE.

Supranto, J. (1983). *Managing the sales force,* Jakarta, Magister.

Tampubolon, M.P. (2004). *Perilaku Organisasi,* Yogyakarta, Ghalia Indonesia

Wacker, J.G. & Lummus, R.R. (2002). Sales forecasting for strategic resource planning, *International Journal of Operation & Production Management,* 22(9), 1014-1031.

Wisner, J. & Stanley, L. (1994). Forecasting Practices in Purchasing, *Journal of Supply Chain Management,* Winter, 30(1), 22.