Forecasting Hotel Expenses using The Arima Method

T Syahromi*, I D Sumitra
Program of Study Magister of System Information, Universitas Komputer Indonesia, Bandung, Indonesia.

Email : tiassyahromi@gmail.com

Abstract. The purpose of the research is for the purpose of the paper to forecast expenses in hotels forecasting costs incurred in hotels using the ARIMA method that is the data used is the cost of expenditure data from January 2014 to January 2019. In addition to the data obtained by using a descriptive method that is documentation with collect data and in the literature. This method was chosen because of the suitability of the amount and pattern of existing data. Based on the results of research using the ARIMA method is suitable for forecasting data that has trends. That means using the ARIMA method will bring good data to broadcasts that can make good analysis for the company. In addition, the results of the study indicate that using the ARIMA method (1,1,1) an estimated result is obtained from 788450868.46 - 758272912.67 in January 2019 and has a MAPE of 6.616718.

1. Introduction
Revenue Management is a process of applying past data and current data to predict future demand. by understanding customer expectations and habits, managing revenue successfully can determine market segmentation [1]. Often during this time, budget renewal is only to see from the previous year, so that the costs incurred need to be tightened, this causes the implementation of building or equipment maintenance is often delayed or failed.

The prediction of the amount expenses of each month is very much needed because with these predictions, the hotel can plan how to realize and estimate the expenses that must be spent by the hotel in the following year. Along with the increasing competition among hotel properties, the determination of the right income and expenditure is needed, in order to facilitate operations within the hotel [2-3]. In recent years, hoteliers have highlighted the need for short and high term to keep competing in highly competitive markets [4].in running a hotel if the business actor can calculate the company's burden and the number of occupancy that can be sold, and then the entrepreneur can maintain a good ratio between expenses and income. But with a small sales area it will be difficult to determine the forecast [5]. Arima is included in the time series model. Time series models look at time patterns (trends, seasonal) which are used in the Arima [6].

The purpose of this study was to determine the accuracy of the Arima method in forecasting hotel expense. This method was chosen because the data used has the characteristics needed by the ARIMA method

2. Method
Time series methods are methods that pay attention to data based on certain intervals such as daily, weekly, monthly or yearly. George Box and Gwilym Jenkins first developed ARIMA for time series analysis modeling. ARIMA is often also called Box-Jenkins models. ARIMA represents three models, namely from the Auto Regressive model (AR), Moving Average (MA), and Auto Regressive and Moving Average models (ARMA). The stages of implementation in the search model are:

1) Identify the temporary model by using past data to get a model from ARIMA. The identification stage is carried out by observing the estimation pattern of ACF (Autocorrelation Function) and the Partial Autocorrelation Function PACF obtained from the data which is then used to get the guesswork of the model that matches the data pattern.
2) Interpretation or parameter estimation of the ARIMA model using past data.
3) Diagnostic testing to test the feasibility of the model. If the model is not feasible, then do the steps of identification, estimation, diagnostic testing to get a proper model.
4) Application, which is forecasting the value of the periodic data series that will come using the method that has been tested.

The Box-Jenkins Model (ARIMA) is divided into 3 groups, namely: The Auto Regressive (AR) model, the Moving Average (MA), and the ARIMA (Autoregressive Moving Average) mixed model that has the characteristics of the first two models.

1) Auto Regressive Model (AR). The general form of the Auto Regressive model with order p (AR (p)) or the ARIMA model (p, 0,0) is stated as follows:

\[ X_t = \mu + \phi_1 X_{t-1} + \phi_2 X_{t-2} + \ldots + \phi_p X_{t-p} + \epsilon_t \]  

Where:
\[ \mu \] = a constant
\[ \phi_p \] = Autoregressive parameter -p 
\[ \epsilon_t \] = error value at the time t

2) Moving Average Model (MA). The general form of the order moving average model q (MA (q)) or ARIMA (0,0, q) is stated as follows:

\[ X_t = \mu + \epsilon_t - \theta_1 \epsilon_{t-1} - \theta_2 \epsilon_{t-2} - \ldots - \theta_q \epsilon_{t-q} \]  

Where:
\[ \mu \] = a constant
\[ \theta_1 \ldots \theta_q \] is a Moving Average parameter
\[ \epsilon_{t-k} \] = error value at the time t - k

3) Mixed model
a) ARMA process
The general model for a mixture of pure AR (1) and MA (1) processes, for example ARIMA (1,0,1) is stated as follows:

\[ X_t = \mu + \phi_1 X_{t-1} + \epsilon_t - \theta_1 \epsilon_{t-1} \]  

or
\[ (1 - \phi_1 B) X_t = \mu + (1 - \theta_1 B) \epsilon_t \]  

Where:
\[ (1 - \phi_1 B) \] = AR (1)
\[ (1 - \theta_1 B) \] = MA (1)

3. Results and Discussion
The data used for this study is hotel expenditure data, from January 2014 to January 2019. If we see from Table 1, we can see a trend that goes up every year. Data from Table 1 is 61 data with monthly calculations. The data will be tested for stationarity, graph data from Table 1 and then tested stationary data using Box – Cox.
Table 1. 2014-2016 expenditure data

| Month      | 2014            | 2015            | 2016            | 2017            | 2018            | 2019            |
|------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|
| January    | 618,042,087     | 589,768,258     | 714,075,691     | 699,070,063     | 811,201,173     | 630,870,000     |
| February   | 673,395,170     | 633,886,778     | 712,096,920     | 696,635,552     | 762,939,788     |                 |
| March      | 672,399,050     | 633,678,917     | 711,755,279     | 698,357,489     | 735,272,010     |                 |
| April      | 725,506,097     | 627,997,884     | 714,634,268     | 693,217,588     | 755,570,883     |                 |
| May        | 782,216,816     | 673,745,182     | 803,395,430     | 723,382,644     | 741,512,473     |                 |
| June       | 787,597,750     | 725,627,552     | 743,556,504     | 850,570,884     | 905,687,349     |                 |
| July       | 839,868,971     | 893,056,920     | 831,070,063     | 720,170,445     | 752,680,500     |                 |
| August     | 727,248,665     | 733,708,206     | 796,686,703     | 705,382,395     | 776,600,703     |                 |
| September  | 723,519,360     | 790,463,485     | 772,228,421     | 731,170,709     | 740,200,360     |                 |
| October    | 719,596,632     | 794,515,009     | 796,686,703     | 756,122,760     | 820,480,618     |                 |
| November   | 756,117,088     | 803,421,618     | 823,959,772     | 781,926,328     | 825,686,703     |                 |
| December   | 806,352,872     | 844,728,860     | 854,552,760     | 810,937,533     | 870,937,533     |                 |

From Figure 1 can be seen that the data has a seasonal pattern where the pattern repeats every 1 year in terms of 12 months of seasonal patterns. In Figure 2 shows the results of transformation using Box-cox. there is shown the data is stationary because the value is 1.

![Time Series Plot of Expenses](image-url)

**Figure 1.** Time series plot
**Figure 2.** Box-cox plot

**Figure 3.** Autocorelation Function (ACF)
In Table 2 we can see the results of forecasting data, comparing existing data with forecasting data.

**Table 2. Results of comparison of actual and forecasting data**

| MONTH           | FORCAST     | ACTUAL     |
|-----------------|-------------|------------|
| JANUARY 2018    | 788450868.46| 811201173.90|
| FEBRUARY 2018   | 775558583.98| 762939788.81|
| MARCH 2018      | 768167047.77| 735272010.03|
| APRIL 2018      | 763929256.74| 755570883.10|
| MAY 2018        | 761499603.10| 741512473.85|
| JUNE 2018       | 760106609.26| 905687349.13|
| JULY 2018       | 759307963.81| 752680500.00|
| AUGUST 2018     | 758850076.24| 776600703.64|
| SEPTEMBER 2018  | 758587555.46| 740200360.69|
| OCTOBER 2018    | 758437044.37| 820480618.92|
| NOVEMBER 2018   | 758350751.81| 825686703.64|
| DECEMBER 2018   | 758301277.68| 870937533.52|
| JANUARY 2019    | 758272912.67| 630870000.00|

After trying to use the ARIMA method, we get the smallest MAPE, in Table 3 we can calculate the results of MAPE 6.616718. The value generated using this method is fairly good. According to previous literature, in Table 4 based on existing forecasting data and MAPE value, the existing forecasting value is said to be excellent [10].
Table 3. The signification of MAPE [10]

| MAPE  | Signification                   |
|-------|---------------------------------|
| < 10% | Excellent forecasting ability   |
| 10-20%| Good forecasting ability        |
| 20-50%| Reasonable forecasting ability  |
| >50%  | Bad forecasting ability         |

4. Conclusion
In this paper obtained, the MAPE 6.616718% value can be said that this value can be considered excellent. Hence within predicting data with the shape and nature of data like this the ARIMA method can be a good method to use.

Acknowledgment
The researcher thanked “UNIKOM” for its support for the making of the paper

References
[1] Weatherford, L. R., and Kimes, S. E. (2003). A comparison of forecasting methods for hotel revenue management. *International Journal of Forecasting, 19*(3), pp. 401-415.
[2] Weatherford, L. R., Kimes, S. E., and Scott, D. A. (2001). Forecasting for hotel revenue management: Testing aggregation against disaggregation. *Cornell Hotel and Restaurant Administration Quarterly, 42*(4), pp. 53-64.
[3] Chen, I., and Claire, P. (2010). Hotel revenue management: Investigating the interaction of information technology and judgmental forecasting. *UNLV Theses, Dissertations, Professional Papers, and Capstones, 68*(5), pp. 647.
[4] Soegoto, E. S., and Akbar, R. (2018, August). Effect of the internet in improving business transactions with online market methods. In *IOP Conference Series: Materials Science and Engineering, 407*(1), pp. 012051).
[5] Pan, B., and Yang, Y. (2017). Forecasting destination weekly hotel occupancy with big data. *Journal of Travel Research, 56*(7), pp. 957-970.
[6] Condratov, I., and Stanciu, P. (2013). The Use of ARIMA Models for Forecasting the Supply and Demand Indicators from Tourism Sector. *The USV Annals of Economics and Public Administration, 12*(2 (16)), pp. 234-244.
[7] Song, H., Li, G., Witt, S. F., and Athanasopoulos, G. (2011). Forecasting tourist arrivals using time-varying parameter structural time series models. *International Journal of Forecasting, 27*(3), pp. 855-869.
[8] Wardhani, A. R. (2010). Studi Analisis Peramalan Dengan Metode Deret Berkala. *Widya Teknika, 18*(2) 225-234.
[9] Küsters, U., McCullough, B. D., and Bell, M. (2006). Forecasting software: Past, present and future. *International Journal of Forecasting, 22*(3), 599-615.
[10] Kurniawan, B. (2018, August). Integrated Information System for Radio Frequency Identification Based Administration and Academic Activities on Higher Education. In *IOP Conference Series: Materials Science and Engineering, 407*(1), pp. 012097).