Available online at http://icsejournal.com/index.php/JCSE
Journal of Computer Science an Engineering (JCSE)
e-ISSN 2721-0251
Vol. 5, No. 1, February 2024, pp. 39-45

Forecasting Indonesia's Unemployment Rates Using Moving Average Methods

Mira Orisa[1*], Ahmad Faisol[2]

[12]Institut Teknologi Nasional Malang, Malang, Indonesia
[1]*mira.orisa@lecturer.itn.ac.id ; [2] mzfais@lecturer.itn.ac.id

ARTICLE INFO

Article History:
Received January 30, 2024
Revised February 24, 2024
Accepted February 27, 2024

Keywords:
Forecasting
Moving average
MAPE
Unemployment

Correspondence:
E-mail:
mira.orisa@lecturer.itn.ac.id

ABSTRACT

People who have completed more education generally have more access to better job prospects. Indonesia faces a significant social and economic problem with unemployment, and forecasting can help governments anticipate the country's yearly unemployment rate. The simple moving average is one forecasting strategy with an advantage over other methods when processing data with simpler variations. Using this technique, historical data from a given period is aligned to find underlying trends or patterns. The two most important steps in the moving average approach are choosing the time window and dividing the numbers in that window equally. Two time periods were found within the dataset, one lasting six months and the other twelve months, based on the results of the data decomposition. Predictions made from a 6-month timeframe are more accurate than those made from a 12-month period, as evidenced by the fact that the mean absolute percentage error (MAPE) for the 6-month period is smaller than that of the 12-month window. The volume of data (number of observations) and the precision of the simple moving average predictions are found to be positively correlated.

1. Introduction

As a developing nation, Indonesia has always prioritized sustainable development. It must attain economic growth by enhancing the quality of life for its citizens and elevating public opinion in order to accomplish this goal[1]. One metric that is frequently used to assess a nation's economic progress is its unemployment rate. Regardless of their degree of economic development, all nations face formidable obstacles to both economic growth and unemployment. Through economic policies, governments hope to lower unemployment and promote economic growth [2]. The impact of industrialization on the unemployment rate is positive but not statistically significant, suggesting that the service and agricultural sectors are more effective at reducing unemployment than industrialization [3]. Unemployment is a major problem for complex socioeconomic development [4]. Unemployment is caused by low-quality human resources and a work market that is not able to match the demand for labor [5]. The empirical results demonstrate that the majority of the reasons for the variation in the unemployment rate can be explained by the main regional characteristics, which include a young population, the industrial mix, the labor force participation rate, and educational attainment (or human capital). The mean regional variable indicates that more education greatly reduces regional unemployment, a finding that policymakers should take note of. Legislators also need to pass policies that encourage job expansion, especially in the industrial sector. In order to implement the promotion mechanism, attention must be paid to places with significant unemployment that are not underdeveloped, such as East Kalimantan, Jakarta, or West Java, rather than just underdeveloped regions[6].

The COVID-19 epidemic also caused a notable rise in joblessness. This study aims to anticipate the unemployment rate during the COVID-19 pandemic using historical data from the official labor force survey conducted by Badan Pusat Statistik (Statistics Indonesia) and Google Trends data query share for the keyword "phk" (job termination)[7]. From 1994 to 2001, Statistics Indonesia (BPS) changed

http://dx.doi.org/10.36596/jcse.v5i1.793

the definition of open unemployment twice in just seven years. These two changes resulted in a considerable increase in the stated unemployment rates. In 1994, the BPS first removed the need to be actively looking for work. Prior to 1994, an individual was considered actively seeking a job if they had made a sincere effort to find employment in the week preceding the poll. From 1994 onward, if a person has looked for work—even if it wasn't their most recent active search, and they are still waiting to hear back from potential employers, they are deemed actively seeking employment [8].

Forecasting techniques are becoming increasingly important in decision-making across a variety of industries as a result of technology breakthroughs and expanded data availability. A vast array of forecasting methods, from simple mathematical models to AI-supported methodologies, can be customized to specific needs and data complexity. The Single Moving Average (SMA) approach can be used to predict some widely used forecasting techniques, such as the Manurung 2020 study regarding Horden's sales in Umi Nala's shop company. The UMI Nala Store will use this technique to assist in deciding how many stock items to buy in the next period [9]. That same year, Harini used the Double Exponential Smoothing Method to forecast COVID-19 instances in Indonesia. The Double Exponential Smoothing method is one of the forecasting strategies that may be applied to maximize the ARIMA model's prediction with the completion parameter α [10]. Ning et al.'s 2021 study Model identification, parameter estimation, and model testing are carried out in order to create the ARIMA (p, d, q) model for forecasting the carbon emissions of the four regions, respectively. Ultimately, the model projects the data and analyzes their carbon emissions pattern over the next three years. The results can assist decision-makers in making the best choices and setting reasonable targets for cutting carbon emissions [11]. Regression analysis, the Box-Jenkins seasonal autoregressive integrated moving average model (SARIMA), and exploratory data analysis have all been used to examine and interpret Deretic's time series study from 2022. The analysis found that the time series demonstrates a seasonal character. The prognosis can be interpreted as reasonably accurate based on the model's 5.22% mean absolute percentage error (MAPE). Forecasting the frequency of traffic accidents may be utilized to achieve a range of purposes when combined with campaigns, strategies, and action plans for traffic safety to achieve the objectives stated in traffic safety strategies [12]. Moving averages can also be used to forecast daily intervals, such as five, ten, and twenty days, in analyzing stock price patterns [13].

In this study, we propose a moving average method for predicting unemployment in Indonesia according to finished education. The Indonesian Statistical Center provided data from 1986 to 2022 based on levels of education from primary to university, with accuracy testing performed using Mean Absolute Percentage Error.

2. Method

Four processes comprise the research process flow diagram, as depicted in Figure 1, gathering and choosing the dataset, cleaning the data, mining using the simple moving average approach, and utilizing MAPE (mean absolute percentage error) to assess the findings. This approach finds patterns and information that can be used to make decisions by assisting in understanding the context of the data being studied.

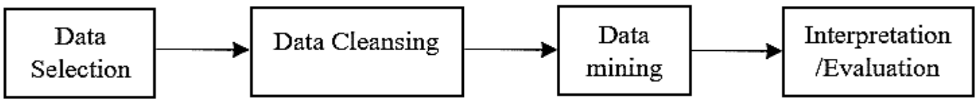

Figure 1. Flow diagram of the research process.

Dataset

According to Figure 1's flow diagram, the first step in the study process is gathering data from https://www.bps.go.id/id/statistics-table/1/OTcyIzE=/penganggur. The Indonesian Statistical Center collected open unemployment data based on higher education in Indonesia between 1986 and 2022,

http://dx.doi.org/10.36596/jcse.v5i1.793

and it is available in this dataset. It covers the total number of jobless people in Indonesia broken down by educational attainment, from primary school graduates to those with advanced degrees. The dataset is then subjected to data scaling and cleansing to eliminate missing values. After verification, only data of the object data type are still there, so the data type needs to be changed to numeric. From 1986 to 2022, the dataset produces 55 columns, one for each educational level. A basic moving average is used in the data mining to extract insights from the dataset. The simple moving average approach is assessed using the mean absolute percentage error (MAPE).

Table 1. Sample dataset unemployment data based on higher education in Indonesia between 1986 and 2022

No	Education Level	1986	1987	1988	1989	1990	1991	etc	2022
1	never attended school	52.166	47.944	41.695	21.424	36.387	30.354	etc	23.104
2	Not completed primary school	161.626	155.456	128.880	108.619	116.095	90.554	etc	92.356
3	Primary school (SD)	371.990	340.684	343.563	362.478	345.319	378.716	etc	421.822
4	Junior high school (SLTP)	316.344	269.332	333.506	331.783	313.481	359.134	etc	356.418
5	General/High School (SLTA Umum/SMU)	544.090	611.958	687.929	717.803	642.881	662.848	etc	748.491
6	Vocational High School (SLTA Kejuruan/SMK)	297.706	293.708	385.943	377.631	337.739	330.827	etc	349.988
7	Academy/Diploma	41.960	48.328	55.265	57.413	59.972	60.778	etc	77.183
8	University	31.790	52.097	63.937	61.007	59.926	78.904	etc	129.848
	Total	1.817.672	1.819.507	2.040.718	2.038.158	1.911.800	1.992.115	etc	16.828.084

Source: Survei Angkatan Kerja Nasional (Sakernas) Indonesia

Forecasting

Forecasting aims to provide future forecasts that are as accurate as possible given the information at hand, including previous data and awareness of any upcoming events that could affect the estimates. If you want your projections to align with your goals, you must plan by figuring out the best path of action. Given its potential significance across numerous business disciplines, forecasting ought to be essential to management's decision-making processes. Forecasts for the short, medium, and long terms are necessary for modern enterprises, depending on the particular use case. Labor, production, and transportation scheduling all require short-term estimates. Demand forecasts are frequently needed at every stage of the scheduling procedure. Before employing, making investments in machinery and equipment, or purchasing raw materials, medium-term projections of resource requirements must be prepared. Strategic planning makes use of long-term predictions. Market opportunities, outside factors, and internal resources must all be considered while making these judgments [14]. It should be noted that this definition of forecasting views forecasting as an organized projection of historical data rather than a prediction [15]. A time series known as a moving average is created by averaging multiple consecutive values from another time series [16].

Simple Moving Average Method

The simple moving average is one widely used forecasting method (SMA). It is easy to use and understand despite not having an adequate length selection mechanism or an underlying statistical model. The other study demonstrated how easy it is to find the optimal model length automatically and presented two statistical models that underlie the simple moving average. Next, we evaluate the proposed model on a real dataset and contrast its output with other popular elementary forecasting

http://dx.doi.org/10.36596/jcse.v5i1.793

methods. Regarding point predictions and prediction intervals, we find that the simple moving average performs better than cumulative and normal values [17]. This approach has certain distinctive features [18].

1. Historical data going back a specific amount of time is needed for forecasting.
2. The smoothing effect and resulting moving average will be more noticeable the longer the moving average. Although the reading of a shorter-term moving average is more erratic, it is closer to the original data.

The application user can choose how many numbers to enter in variable n. Average Terms Moving is utilized because, each time a new observation value emerges, a new average value can be computed by subtracting the previous observation value and inputting the most recent observation value or data. Formula moving average [19].

$$S_{t+1} = \frac{X_1 + X_{t-1} + \cdots + X_{t-n+1}}{n} \quad (1)$$

Description:
S_{t+1} : forecasting for period t+1
X_t : data on period t
n : moving average timeframe

Mean Absolute Percentage Error (MAPE)

The average of absolute percentage errors is known as MAPE (APE). Let At and Ft stand for the predicted and actual values, respectively, at data point t. Next, the definition of MAPE is [20]:

$$MAPE = \frac{1}{n} \sum_{t=1}^{n} |PE_t| \quad (2)$$

Details: t is a time period, and n is a large amount of data. The following equation yields the value of PE_t

$$PE_t = \left(\frac{X_t - F_t}{X_t}\right) x100\% \quad (3)$$

3. Results and Discussion

Results

Data visualization is done to examine the data distribution based on education level attributes after the dataset has been gathered and cleaned from empty data values. Graph Figure 2 displays the data visualization's outcomes. Figure 2 shows a graph of Indonesia's unemployment rate by educational attainment (1986–2022). From 1988 to 2022, the X axis represents the year, and the Y axis represents the number of unemployed people (millions) according to different educational levels (elementary, middle, high school, vocational, diploma, and university). Figure 2's data visualization illustrates how the number of unemployed people in Indonesia changed over the course of 1988 to 2022. The data pattern in the graph indicates that the number of jobless people rose between 2000 and 2008, then fell between 1997 and 1999 and between 2012 and 2016. With 3 million unemployed people in elementary, middle, and high school, 2006 saw the highest peak in the jobless rate. In 2020, enrollment in high schools and vocational schools reached 2.5 million. This was a result of the COVID-19 pandemic that year in Indonesia, which led to several layoffs in factories and other businesses that employed people with that level of education. With 500,000 individuals, the unemployment rate was at its lowest from 1988 to 1994. Let's examine it between 1994 and 2022. Those with only an elementary school education tend to be the most unemployed, followed by those with a junior high school education, who have the second highest unemployment rate after those with an elementary school education. Despite an uptick in 2012, junior high school education declined between 2008 and 2020. The graph in Figure 3 illustrates that the highest number of unemployed people are educated to an elementary school level, followed by junior high school, high school, and

http://dx.doi.org/10.36596/jcse.v5i1.793

vocational school, in that order. Compared to other education levels, diploma and university education levels are not very high. Figure 2's data distribution demonstrates that unemployment is still a major issue in Indonesia, particularly for those elementary and middle school graduates.

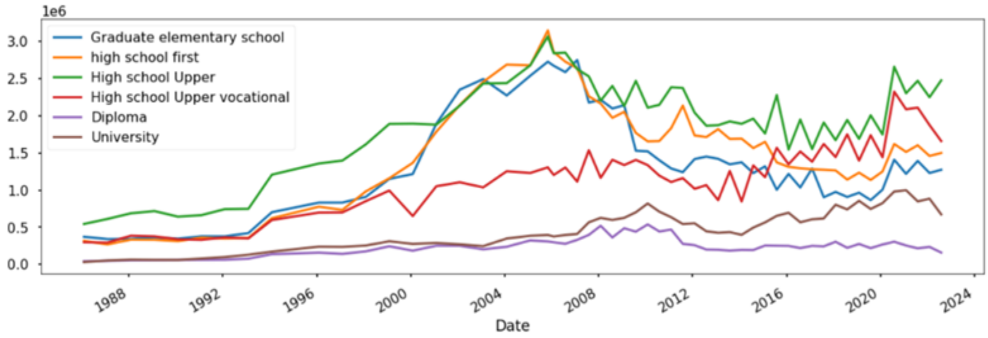

Figure 2. Graduation data based on the level of education completed

Data decomposition is performed on university graduate unemployment data by importing the stats models module, a Python package that offers a range of statistical and economic approaches for data analysis. A year's worth of decomp = seasonal_decompose visualization is displayed in Figure 3. It developed an ascending line pattern that has a trend and does not include seasonal data; it is stable over time and does not fluctuate regularly or repeatedly at a given period.

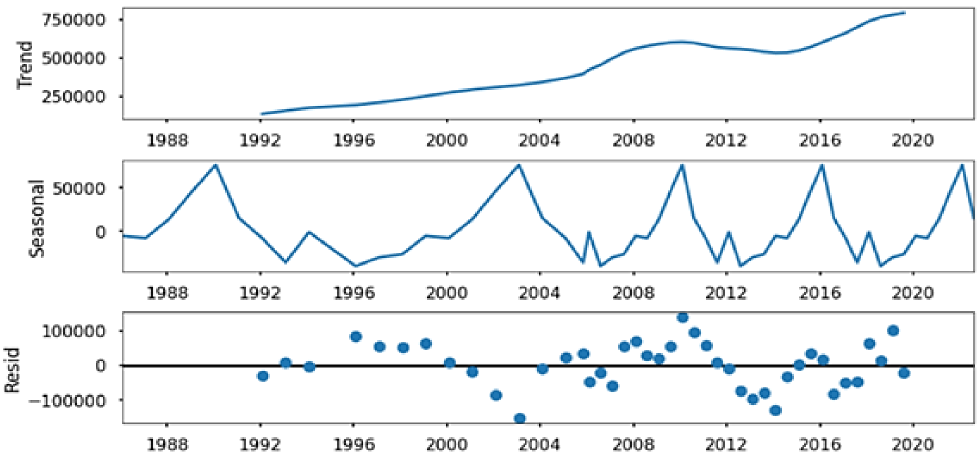

Figure 3. Trend and Resid on University Graduate Unemployment Data

Discussion

The data is first subjected to seasonal decomposition, and then the simple moving average (SMA) approach is applied to aid in the interpretation of patterns and trends in the data and to enable the development of more accurate predictions. Based on the gathered data, the value of the simple moving average was calculated using a rolling window with year and half-year intervals. For each half-year and one-year time window, Table 2 displays the unemployment estimates for primary school, high school first, high school, vocational high schools, diploma programs, or universities. The mean absolute percentage error, or MAPE, is computed to assess the method's accuracy. As indicated in Table 3, the Mean Absolute Percentage Error (MAPE) computation is based on the rolling window of six and twelve months.

http://dx.doi.org/10.36596/jcse.v5i1.793

Table 2. Result forecasting moving average

Graduate	Unemployment	
	Six-Month Period	*12-Month period*
Primary School	1255889.0	1120619.75
High School First	1492239.16	1357422.92
High School	2319826.83	2058306.08
Vocational High Schools	1918124.83	1737305.0
Diplomas	239695.67	246028.92
Universities	868761.5	868761.5

The MAPE results displayed in Table 2 indicate that there is a correlation between the way the Simple Moving Average (SMA) handles new and old data during the forecasting process. The six-month MAPE (mean absolute percentage error) is less than the twelve-month MAPE. The number of observations in the time window, which is used to get the average number of recent observations, affects how sensitive simple moving averages are to changes in the data. A simple moving average, which makes use of more previous data, typically yields predictions that are more consistent since changes in fresh data have less of an effect on the total projection. On the other hand, the forecast may become more volatile and have a tendency to closely follow data fluctuations if the number of observations is too small.

Table 3. The mean absolute percentage error (MAPE)

Graduate	The Mean Absolute Percentage Error (MAPE)	
	Six-Month Period	*12-Month period*
Primary School	16.42%	28.74%
High School First	15.02%	25.48%
High School	12.22%	18.18%
Vocational High Schools	13.47%	19.00%
Diplomas	19.29%	31.27%
Universities	17.35%	24.75%

4. Conclusion

The MAPE results displayed in Table 2 indicate that there is a correlation between the way the Simple Moving Average (SMA) handles new and old data during the forecasting process. The six-month MAPE (mean absolute percentage error) is less than the twelve-month MAPE. The number of observations in the time window, which is used to get the average number of recent observations, affects how sensitive simple moving averages are to changes in the data. Simple moving averages, which use a larger amount of historical data, typically produce predictions that are more stable since the overall prediction is less affected by changes in fresh data. On the other hand, the forecast may become more volatile and have a tendency to closely follow data fluctuations if the number of observations is too small.

References

[1] E. Kurniawan A., M. Awaluddin, F. Fitriadi, A. Busari, and D. C. Darma, "Contemporary Indonesian GDP: Context of Analysis at Unemployment, Labor Force and Poor People," International Journal of Economics and Financial Research, no. 74. Academic Research Publishing Group (Publications), pp. 143–154, Nov. 25, 2021. doi: 10.32861/ijefr.74.143.154

[2] Fatkhu Rokhim, "Factors Influencing Unemployment in Indonesia," Journal of Scientific Research, Education, and Technology (JSRET), vol. 2, no. 1. Kirana Publisher, pp. 122–131, Jan. 31, 2023. doi: 10.58526/jsret.v2i1.51

[3] E. A. R. Puspadjuita, "Factors that Influence the Rate of Unemployment in Indonesia," International Journal of Economics and Finance, vol. 10, no. 1. Canadian Center of Science and Education, p. 140, Dec. 10, 2017. doi: 10.5539/ijef.v10n1p140.

[4] P. Eko Prasetyo, “The Role of MSME on Unemployment in Indonesia”, TURCOMAT, vol. 12, no. 13, pp. 2519–2525, Jun. 2021. doi: 10.17762/turcomat.v12i13.8944

http://dx.doi.org/10.36596/jcse.v5i1.793

[5] A. Rahmawati and P. Putri, "The Effect of Education Variable on the Open Unemployment Rate", efficient, vol. 4, no. 2, pp. 1160-1173, Jun. 2021. doi: 10.15294/efficient.v4i2.46128

[6] E. K. Oktafianto, N. A. Achsani, and T. Irawan, "The Determinant of Regional Unemployment in Indonesia: The Spatial Durbin Models," Signifikan: Jurnal Ilmu Ekonomi, vol. 8, no. 2. LP2M Universitas Islam Negeri (UIN) Syarif Hidayatullah Jakarta, Jul. 06, 2019. doi: 10.15408/sjie.v8i2.10124.

[7] Fajar, M., Prasetyo, O. R., Nonalisa, S., & Wahyudi, W. Forecasting unemployment rate in the time of COVID-19 pandemic using Google trends data (case of Indonesia). 2020.

[8] Suryadarma, D., Suryahadi, A., & Sumarto, S. The measurement and trends of unemployment in Indonesia: The issue of discouraged workers. Available at SSRN 861464. 2005.

[9] Manurung, N. Application Of The Single Moving Average (SMA) Method For Forecasting Sales Of Horden In Umi Nala's Shop Business. In *International Conference on Social, Sciences and Information Technology* (Vol. 1, No. 1, pp. 297-306). 2020.

[10] Harini, S. . Identification COVID-19 cases in Indonesia with the double exponential smoothing method. *Jurnal Matematika MANTIK*, *6*(1), 66-75. 2020. doi: 10.15642/mantik.2020.6.1.66-75

[11] Ning, L., Pei, L., & Li, F. Forecast of China's carbon emissions based on Arima method. *Discrete Dynamics in Nature and Society*, 2021, 1-12. doi:10.1155/2021/1441942

[12] Deretić, N., Stanimirović, D., Awadh, M. A., Vujanović, N., & Djukić, A. SARIMA modelling approach for forecasting of traffic accidents. *Sustainability*, *14*(8), 4403. 2022. doi: 10.3390/su14084403

[13] Hari, Y., & Dewi, L. P. Forecasting system approach for stock trading with relative strength index and moving average indicator (Doctoral dissertation, Petra Christian University). 2018.

[14] Hyndman, R. J., & Athanasopoulos, G. *Forecasting: principles and practice*. OTexts. 2018.

[15] Chapman, S.N., 2006. The Fundamentals Of Production Planning And Control. [e-book] Prentice Hall : Upper Saddle River, New Jersey. Available : website < http://brharnetc.edu.in/br/wp-content/uploads/2018/11/22.pdf > [Accessed 03 Oktober 2023]

[16] R. J. Hyndman, "Moving Averages," International Encyclopedia of Statistical Science. Springer Berlin Heidelberg, pp. 866–869, 2011. doi: 10.1007/978-3-642-04898-2_380

[17] Svetunkov, Ivan, and Fotios Petropoulos. "Old dog, new tricks: a modelling view of simple moving averages." International Journal of Production Research 56.18. 6034-6047. 2018. doi: 10.1080/00207543.2017.1380326

[18] Nadhira, A. T., Gadisku, C. A., & Peranginangin, S. M. (2021). Demand Forecasting Comparison of Softex 1400-M using Single Moving Average Method and Single Exponential Smoothing Method. In *Proceedings of the International Conference on Industrial Engineering and Operations Management* (pp. 452-459).

[19] Yudaruddin.R(2019). Forecasting: for economic and business activities. RV.Pustaka Horizon. ISBN:978-602-5431-48-7.

[20] Liantoni, F., & Agusti, A. Forecasting bitcoin using double exponential smoothing method based on mean absolute percentage error. *JOIV: International Journal on Informatics Visualization*, *4*(2), 91-95. 2020. doi: 10.30630/joiv.4.2.335