

Visualization and Prediction The Number of Patients Visit in Curug Health Center

Milda Aprilia¹, Wella^{2✉}

^{1,2} Information System Study Program, Faculty of Engineering and Informatics, Universitas Multimedia Nusantara, Indonesia

Article Information

Article History

Received : April 15, 2024

Revised : April 29, 2024

Accepted : May 15, 2024

Keywords:

Accuracy; Autoregressive Integrated Moving Average (ARIMA); Linear Regression; Prediction; Single Exponential Smoothing.

Kata Kunci:

Akurasi; Autoregressive Integrated Moving Average (ARIMA); Linear Regression; Prediksi; Single Exponential Smoothing.

ABSTRACT

The problem with this research is the services provided by Puskesmas Curug seem to have excessively long waiting times. In light of this problem, it is imperative to establish a forecast in order to improve the readiness of healthcare facilities in delivering patient care. The forecast for the number of patient visits at Puskesmas Curug was made using data from the years 2015-2018, including visits from both the general public and those covered by the BPJS. Three techniques, specifically linear regression, single exponential smoothing, and autoregressive integrated moving average, are utilised to predict and ascertain the most precise forecasting accuracy. Furthermore, this study has created a representation that can be used by the Curug health centre. The research findings suggest that the ARIMA technique demonstrated the highest level of accuracy, with a precision rate of 73%. The visualisation result will be transmitted to the Puskesmas and then evaluated.

ABSTRAK

Sebagai hasil dari penelitian ini, layanan Puskesmas Curug terkesan terlalu lama. Akibatnya, prediksi diperlukan untuk membuat puskesmas lebih siap melayani pasien. Data dari tahun 2015–2018 digunakan untuk memprediksi jumlah kunjungan pasien UMUM dan BPJS di Puskesmas Curug. Tiga pendekatan digunakan untuk menghasilkan prediksi yang paling akurat: regresi linier, pemulusan eksponensial tunggal, dan metode rata-rata bergerak autoregresif terintegrasi. Studi ini juga membuat visualisasi yang akan digunakan oleh Puskesmas Curug. Hasil penelitian menunjukkan bahwa metode ARIMA memiliki akurasi tertinggi sebesar 73%, dan hasil visualisasi akan dikirim ke Puskesmas untuk dievaluasi. Visualisasi dibangun dengan menggunakan data yang tersedia, dengan menggunakan Microsoft Excel sebagai alat bantu yang dipilih karena sudah familiar dan kompatibel dengan Puskesmas, sehingga memungkinkan terciptanya visualisasi yang baik.

Corresponding Author:

Wella

Information System Study Program, Faculty of Engineering and Informatics, Universitas Multimedia Nusantara, Indonesia

Jl. Scientia Boulevard, Gading Serpong, Tangerang

E-mail: wella@umn.ac.id

INTRODUCTION

Health Centres, also known as Puskesmas, are operational entities that implement comprehensive, integrated, fair, acceptable, and economical healthcare initiatives for the

community (Valentina et al., 2023). By actively involving the community and utilising the outcomes of suitable scientific and technological advancements, while considering costs that can be supported by both the government and the community (M. Saputra et al., 2015). The Curug Health Centre, situated in Curug, Tangerang Regency, was selected as one of the primary research sites for this study. Based on the acquired data, the Curug Health Centre recorded a total of 16,427 outpatient visits in the year 2015, averaging at 1,369 visits per month.

In 2016, the Curug Health Centre recorded a total of 29,598 outpatient visits, averaging 2,466 patient visits per month. The pattern has exhibited a steady upward trend until the year 2018. The Curug Health Centre is currently open 24 hours a day, but outpatients can only register until 12:00 pm from Monday to Thursday and until 11:30 am on Fridays. Every day, there were hundreds of patients who came for outpatient treatment. The Puskesmas Curug faced complaints from visitors due to a lengthy service time, caused by a combination of limited registration time and a high volume of outpatients. In order to anticipate the number of outpatients attending Puskesmas Curug in the coming year, it is necessary to make a prognosis based on the typical number of visits and the associated issues. Prediction is essential to ensure that Puskesmas can adequately anticipate the needs of outpatients who will seek medical services. In order to generate forecasts and graphical representations, it is necessary to employ a technique for making predictions and visualisations (Hullman et al., 2018; Wiratman & Wella, 2024). The Curug Health Centre predicts patient visits using three different methods: linear regression (Rahmawati & Laras, 2023), single exponential smoothing (Aprilia, 2017; Maulidiyah et al., 2021) and autoregressive integrated moving average (ARIMA) (Wardani, 2017). The visualisation strategy is congruent with the data being utilised. Data visualisation is the process of presenting information to individuals in a clear and concise manner using various types of charts, such as tables and graphs (D. F. Saputra, 2018; Wiratman & Wella, 2024).

RESEARCH METHODS

Research Methodology

This research focused on the prediction and visualization. The main steps could be seen on figure below.

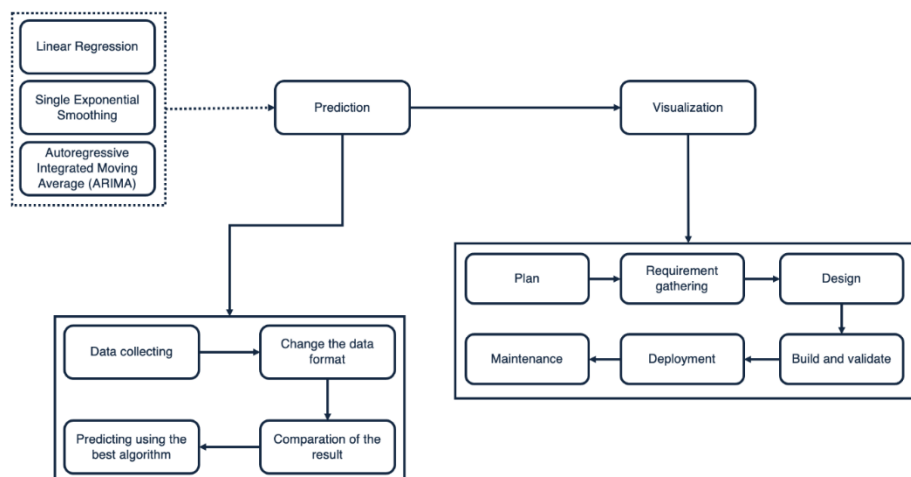


Figure 1. Whole flow of the research methodology

In Figure 1 it can be seen that there will be two processes carried out: prediction and visualisation. In order to generate a prediction, it is vital to possess a methodology that can serve as a dependable frame of reference. When formulating this forecast, three methodologies are employed: Linear Regression, Single Exponential Smoothing, and Autoregressive Integrated Moving Average (ARIMA). The process of producing predictions comprises four components. The initial phase comprises data retrieval, which involves acquiring consent from the Tangerang

District Health Office and requesting the requisite data from the Curug Health Centre. The second stage entails transforming the data format from PDF to Excel. This conversion is carried out to facilitate the utilisation of the data in the computational process. In the third step, exploration outcomes are compared using three data processing methods: linear regression, single exponential smoothing, and autoregressive integrated moving average (ARIMA). Perform predictions and examine the data to ascertain the precision of each forecast. The last phase is predicting the number of patient visits anticipated in the year 2020. After attaining a notable degree of accuracy in all inquiries, the method was chosen to calculate projections for the year 2020. Moreover, after the prediction phase, the next step is to generate a visual depiction of the data obtained from the previously described forecasts. There are a total of six steps involved in the process of visualising data. The project begins with the planning phase, which is followed by the requirement gathering stage, the design stage, the build and validation stage, the deployment stage, and ultimately the maintenance stage.

Linear Regression Method

Linear regression is a statistical method employed to approximate, anticipate, or project the values of a variable by considering the values of other variables. Additionally, it aids in establishing the correlation between two variables (Riadi & Setiawan, 2019). The Y variable denotes the values that are to be forecasted. The term "criterion" is commonly used to describe it. The values of variable X are fundamental for making predictions. According to the research conducted by (Marhenta et al., 2018), variable X is designated as the predictor.

Single Exponential Smoothing

The exponential smoothing approach is an iterative process that involves conducting calculations using the most recent observable data. The Single Exponential Smoothing method is a technique that applies decreasing weights to previous observation values according to an exponential function. In essence, a more recent value is assigned greater importance as compared to an older observation value. This method utilises an exponential weighting system to calculate the moving averages of all previous observed values (Hartono et al., 2012).

Autoregressive Integrated Moving Average (ARIMA)

ARIMA is an independent variable-agnostic model used for forecasting. ARIMA utilises past and present values of the dependant variable to make accurate forecasts for the immediate future. ARIMA is suitable when there is a statistical correlation between observations in a time series, showing interdependence (Mubin et al., 2012).

Visualization

Visualisation aids users in the process of analysing and drawing logical conclusions from data and evidence. Visualisation enhances the accessibility, comprehensibility, and utility of intricate data. Users have the ability to perform various analytical tasks, such as drawing comparisons or comprehending quality, by utilising graphs that demonstrate comparisons or quality as a result of their work. Tables are typically employed to display specific values of a variable, whilst graphs of diverse sorts are utilised to visualise patterns or correlations in data involving one or more variables (Syaripul & Bachtiar, 2016).

RESULTS AND DISCUSSION

Table 1. Results of Data Exploration

Data used	Methods Used	Level of Accuracy
2015 – 2018	Linear Regression	66%
	Single Exponential Smoothing	65%
	Autoregressive Integrated Moving Average (ARIMA)	73%
2016 – 2018	Linear Regression	61%
	Single Exponential Smoothing	65%
	Autoregressive Integrated Moving Average (ARIMA)	60%

Data used	Methods Used	Level of Accuracy
2017 – 2018	Linear Regression	52%
	Single Exponential Smoothing	65%
	Autoregressive Integrated Moving Average (ARIMA)	56%
2018	Linear Regression	39%
	Single Exponential Smoothing	66%
	Autoregressive Integrated Moving Average (ARIMA)	60%

Three approaches were employed to produce forecasts, and the ARIMA technique exhibited the highest level of accuracy when analysing the data from 2015 to 2018. The table shown illustrates the outcomes of data analysis performed on datasets covering the years 2015 to 2018, specifically the time periods 2015-2018, 2016-2018, 2017-2018, and 2018. Three different approaches were employed for this analysis: Linear Regression, Single Exponential Smoothing, and Autoregressive Integrated Moving Average (ARIMA). Upon careful examination of the projections for outpatient visits in 2019, the precision of each analysis has been determined. The accuracy rate of 73% is considered to be high. The rate was determined by a research that analysed data from 2015 to 2018, using the Autoregressive Integrated Moving Average (ARIMA) methodology. The method exhibiting the utmost precision will be selected as the technique employed for future projections, specifically in forecasting the quantity of patient visits in 2020.

	Jan	Feb	Mar	Apr	May	Jun
2020	158.5037	423.6509	915.2702	1426.3011	1944.4496	2393.8373
2021	3250.4430	3179.2533	3103.5115	3032.1179	2970.8587	2922.7329
2022	2880.8460	2891.6556	2901.5109	2909.7224	2915.9745	2920.2424
2023	2918.0751	2916.7282	2915.6384	2914.8359	2914.3129	2914.0362

	Jul	Aug	Sep	Oct	Nov	Dec
2020	2761.7822	3030.9962	3207.5467	3300.7460	3327.4147	3304.8383
2021	2888.5267	2867.3982	2857.4635	2856.2984	2861.3459	2870.2101
2022	2922.7066	2923.6717	2923.4986	2922.5515	2921.1605	2919.5993
2023	2913.9583	2914.0263	2914.1889	2914.4013	2914.6270	2914.8395

Figure 2. Prediction using R for BPJS patient visits

Figure 2 displays the projected outcomes for the overall count of outpatient visits in 2020, utilising the ARIMA methodology. Figure 3 illustrates the representation of the anticipated outcomes. It can be seen that the prediction data using ARIMA almost follows the pattern in the previous data. At the beginning of the year, the patient experience was quite low, then as the months went by, there was a significant increase. Followed by a significant decline until the end of the year.

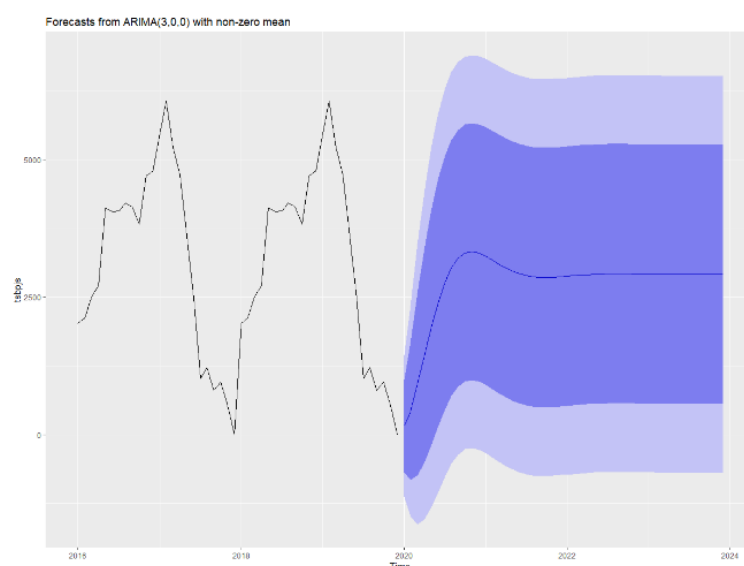


Figure 3. R Graph for future BPJS patient visits in 2020

	Jan	Feb	Mar	Apr	May
2020	3263.20222	3479.71410	1161.64054	3592.57720	2490.12173
2021	3234.49796	3457.28375	1144.74318	3579.49174	2480.19707
2022	3233.32656	3456.38669	1144.05706	3578.96646	2479.79522
2023	3233.27929	3456.35052	1144.02937	3578.94527	2479.77901
	Jun	Jul	Aug	Sep	Oct
2020	1357.78253	1868.72126	1731.91458	1363.41320	1783.38111
2021	1350.13608	1862.89941	1727.44220	1360.00053	1780.76382
2022	1349.82849	1862.66406	1727.26207	1359.86270	1780.65833
2023	1349.81608	1862.65457	1727.25480	1359.85714	1780.65408
	Nov	Dec			
2020	974.43323	-38.55965			
2021	972.43360	-40.09179			
2022	972.35288	-40.15357			
2023	972.34962	-40.15606			

Figure 4. Prediction using R for UMUM patient visits

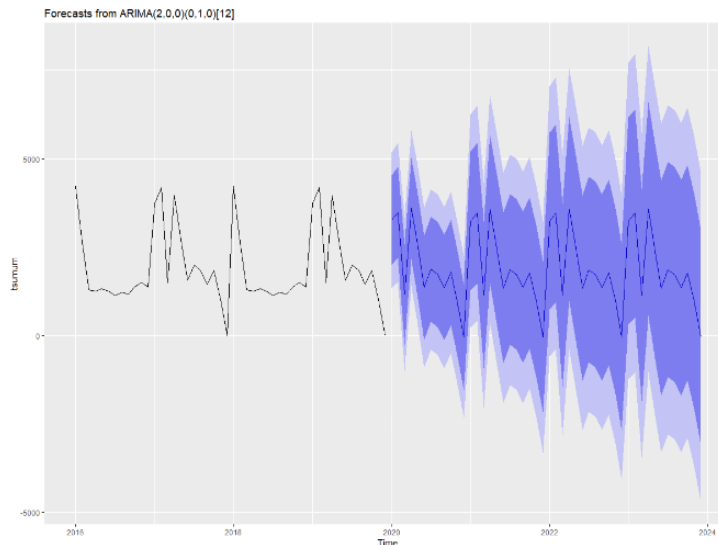


Figure 5. R Graph for UMUM patient visits in 2020

Figure 6 presents a graphical representation of the outpatient visit statistics at the Curug Community Health Centre spanning from 2015 to 2019. This dataset displays the monthly count of BPJS outpatients, spanning from January to December. In March 2018, there were a total of 3157 patient visits, marking the highest number of visits by BPJS patients. The aim is for this visualization to assist researchers in making informed judgments. The Curug Health Center is capable of ensuring an ample supply of skilled personnel, medical equipment, and medications. In order to ensure optimal care for all patients.

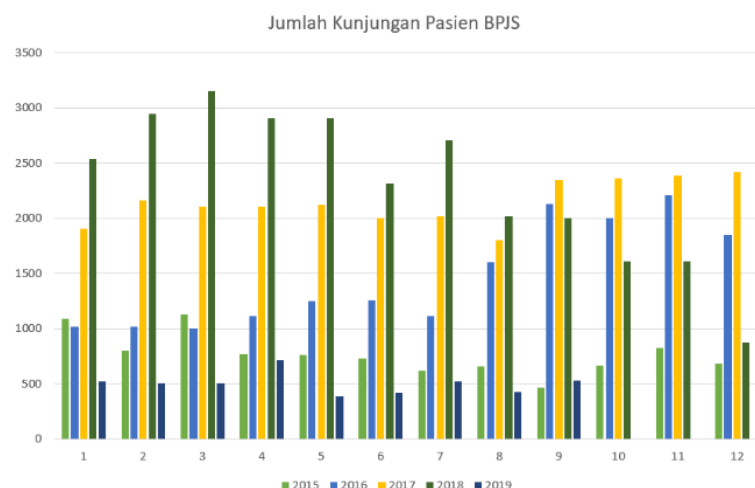


Figure 6. Visualization BPJS patient

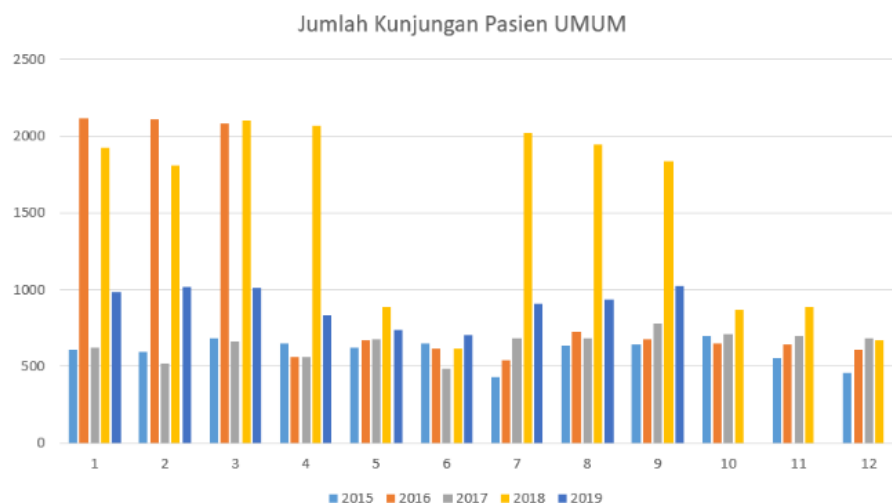


Figure 7. Visualization UMUM patient

Figure 7 presents a visualisation depicting the outpatient visit data at the Curug Community Health Centre from 2015 to 2019. This dataset displays the monthly count of outpatients categorised as GENERAL, spanning from January to December. In January 2016, there were a total of 2114 patient visits, which was the highest number of visits for general patients.

CONCLUSION

The study concluded that the autoregressive integrated moving average (ARIMA) technique had the highest level of accuracy when applied to a dataset covering a period of four years. The ARIMA model is used to explore the data, using a dataset spanning four years, in order to get the highest level of accuracy. The forecast outcomes for GENERAL patient visits produced a mean value of 1357, with the highest value observed at 3592 and the lowest value at 974. Additionally, the patient visits for BPJS had an average value of 2293, with the highest value recorded as 3327 and the lowest value recorded as 158. The visualisation is built using the available data, employing Microsoft Excel as the selected tool due to its familiarity and compatibility for the Puskesmas, allowing for the creation of good visualisation.

ACKNOWLEDGEMENTS

The author would like to thank Universitas Multimedia Nusantara, especially the Information Systems Study Program, which has provided both material and non-material support in the preparation of this article.

REFERENCES

- Aprilia, D. (2017). Penerapan Metode Forecast Exponential Smoothing pada Jumlah Pasien Puskesmas. *Jurnal Biometrika Dan Kependudukan*, 5(2), 146. <https://doi.org/10.20473/jbk.v5i2.2016.146-156>
- Hartono, A., Dwijana, D., & Handiwidjojo, W. (2012). Perbandingan Metode Single Exponential Smoothing Dan Metode Exponential Smoothing Adjusted For Trend (Holt's Method) Untuk Meramalkan Penjualan. Studi Kasus: Toko Onderdil Mobil "PRODI, PURWODADI." *Jurnal EKSIS*, 5(1).
- Hullman, J., Kay, M., Kim, Y.-S., & Shrestha, S. (2018). Imagining Replications: Graphical Prediction & Discrete Visualizations Improve Recall & Estimation of Effect

- Uncertainty. *IEEE Transactions on Visualization and Computer Graphics*, 24(1), 446–456. <https://doi.org/10.1109/TVCG.2017.2743898>
- Marhenta, Y. B., Satibi, S., & Wiedyaningsih, C. (2018). The Effect of BPJS Service Quality Level and Patient Characteristics to Patient Satisfaction in Primary Health Facilities. *JURNAL MANAJEMEN DAN PELAYANAN FARMASI (Journal of Management and Pharmacy Practice)*, 8(1), 18. <https://doi.org/10.22146/jmpf.34438>
- Maulidiyah, R., Puspitasari, S. T., & Redjeki, E. S. (2021). Hubungan Faktor Kapabilitas, Motivasi, dan Pendayagunaan SDM terhadap Kinerja pada Pegawai Puskesmas. *Sport Science and Health*, 3(8), 586–594. <https://doi.org/10.17977/um062v3i82021p586-594>
- Mubin, L. F., Anggraeni, W., & Vinarti, A. (2012). Prediksi Jumlah Kunjungan Pasien Rawat Jalan Menggunakan Metode Genetic Fuzzy Systems Studi Kasus : Rumah Sakit Usada Sidoarjo. *Jurnal Teknik ITS*, 1(1), 482–487.
- Rahmawati, E. N., & Laras, M. A. (2023). PREDIKSI KUNJUNGAN PASIEN RAWAT JALAN DI RSAU Dr. SISWANTO LANUD ADI SOEMARMO KARANGANYAR TAHUN 2022-2026. *Infokes: Jurnal Ilmiah Rekam Medis Dan Informatika Kesehatan*, 13(1), 52–63. <https://doi.org/10.47701/infokes.v13i1.2561>
- Riadi, B., & Setiawan, J. (2019). Visualisasi Data dan Prediksi Perkembangan Impor Barang di Indonesia Tahun 2010 – 2014. *Ultima InfoSys*, 9(2), 92–97. <https://doi.org/10.31937/si.v9i2.998>
- Saputra, D. F. (2018). Visualisasi Data Di Sistem Manajemen Perpustakaan. *Jurnal Perpustakaan Pertanian*, 26(2), 82. <https://doi.org/10.21082/jpp.v26n2.2017.p82-86>
- Saputra, M., Marlinae, L., Rahman, F., & Rosadi, D. (2015). Program Jaminan Kesehatan Nasional Dari Aspek Sumber Daya Manusia Pelaksana Pelayanan Kesehatan. *Jurnal Kesehatan Masyarakat*, 11(1), 32. <https://doi.org/10.15294/kemas.v11i1.3462>
- Syaripul, N. A., & Bachtiar, A. M. (2016). Visualisasi Data Interaktif Data Terbuka Pemerintah Provinsi Dki Jakarta: Topik Ekonomi Dan Keuangan Daerah. *Jurnal Sistem Informasi*, 12(2), 82. <https://doi.org/10.21609/jsi.v12i2.481>
- Valentina, F., Andayani, L. S., & Lubis, R. (2023). Impact Factors of Use of Integrated Non-communicable Diseases Development Post at Community Health Centers. *Randwick International of Social Science Journal*, 4(1), 28–37. <https://doi.org/10.47175/rissj.v4i1.603>
- Wardani, R. (2017). Trend Analisis Peningkatan Jumlah Kunjungan Pasien Ditinjau Dari Marketing MIX. *IKESMA*, 13(1). <https://doi.org/10.19184/ikesma.v13i1.7025>
- Wiratman, A. B., & Wella, W. (2024). Personalized Learning Models Using Decision Tree and Random Forest Algorithms in Telecommunication Company. *JOIV : International Journal on Informatics Visualization*, 8(1), 318. <https://doi.org/10.62527/joiv.8.1.1905>