JOIV : Int. J. Inform. Visualization, 7(4) - December 2023 2263-2269

INTERNATIONAL JOURNAL
ON INFORMATICS VISUALIZATION

journal homepage : www.joiv.org/index.php/joiv

# Multi-Temporal Factors to Analyze Indonesian Government Policies regarding Restrictions on Community Activities during COVID-19 Pandemic

Syafrial Fachri Pane, Adiwijaya*, Mahmud Dwi Sulistiyo, Alfian Akbar Gozali

*Telkom University, Dayeuhkolot, Bandung, 40257, Indonesia*

*Corresponding author: adiwijaya@telkomuniversity.ac.id*

***Abstract*-Concerning the implementation of the government policy regarding the Restriction of Community Activities (PPKM) during the COVID-19 pandemic era, there are still discrepancies in the economic sector and population mobility. This issue emerges due to irrelevant data and information in one region of Indonesia. The data differences should be carefully solved when implementing the PPKM policy. Besides, the PPKM must also pay attention to some specific factors related to the real conditions of a region, such as the data on the epidemiology of COVID-19, economic situations, and population mobility. These three are called Multi Factors. Then, based on the data, COVID-19 has a specific spreading period that cannot be repeated and thus is called temporal. Therefore, using the Multi-Temporal Factors approach to identify their correlation with the PPKM policy by applying Machine Learning, such as the Multiple Linear Regression model and Dynamic Factors, is essential. This research aims to analyze the characteristics and correlations of the COVID-19 pandemic data and the effectiveness of the government's policy on community activities (PPKM) based on the data quality. The results show that the accuracy of the multiple linear regression models is 84%. The Dynamic Factor shows that the five most important factors are *idr_close*, *positive*, *retail_recreation*, *station*, and *healing*. Based on the ANOVA test, all independent variables significantly influence the dependent one. The linear multiple regression models do not display any symptoms of heteroscedasticity. Thus, based on the data quality, the implementation of PPKM by the government has a practical impact.**

***Keywords*—COVID-19; pandemic model; Indonesian government; community restriction policy; multi-temporal.**

*Manuscript received 12 Apr. 2022; revised 24 Dec. 2022; accepted 16 Feb. 2023. Date of publication 31 Dec. 2023.*


## I. INTRODUCTION

The spreading development of COVID-19 throughout the world is getting higher. According to information from the World Health Organization (WHO), there have been 242,348,657 people suffering from this virus, and 4,927,723 of them have been confirmed dead [1]. COVID-19 has become a serious global problem, significantly resulting in considerable losses in the world's health and economy [2]. The pandemic has forced all countries to innovate to reduce mortality and support economic recovery [3]. Now that the World Health Organization has declared the COVID-19 pandemic, [4]. There has been an increasing number of studies on this virus [3], such as the shift from pandemic to endemic status, which is becoming more difficult to forecast as death tolls rise globally [1]. Therefore, the Indonesian government needs to issue some policies or regulations to anticipate the risks caused by COVID-19. Many countries have imposed strict policies such as lockdown [6]-[7].

The number of positive COVID-19 cases until October 20, 2021, in Indonesia, is 4,236,287, with a recovery number of 4,076,541, and there have been 143,049 people who have passed away due to the virus [8]. Regarding mortality, Indonesia ranks fourth in Southeast Asia [9]. However, the death rate is predicted to increase if proper prevention programs are not carried out consistently. In response to the development and spread of COVID-19 in Indonesia, the government has implemented restrictions on community activities (PPKM) in the education, health, economics, and population mobility sectors [9]. The policy is intended to reduce the death rate.

The PPKM policy has had some positive effects in its implementation, especially in the health sector, where the mortality rate has dramatically decreased [10]. However, what about other fields like the economic sector and population mobility? Of course, everyone expects more stable conditions. There are still discrepancies or different health, economic, and population mobility results. The PPKM imposes considerable costs due to reduced economic activities [11]-[12]. The main reason is the unavailability of exact and relevant information on the real conditions of the spread of COVID-19 in many regions of Indonesia.

Moreover, the information authenticity still needs to be verified [13]. The data differences must be carefully watched when implementing the PPKM policy. Previous research explained that relevant information sources could be used to improve effective and efficient information transparency as the main key to performing better public services in informative policymaking [14]-[15].

Therefore, the PPKM policy must strongly consider what factors are relevant to the conditions of specific regions, such as precise data on the health, economy, and people's mobility sectors. Seeing the characteristics of the data, the spread of COVID-19 occurs during specific periods, and it cannot be repeated; So, it is called temporal. This study employs a Machine Learning technique, i.e., the Multiple Linear Regression model, to deal with the challenges posed by the data.

Regression refers to the additional research conducted to determine the link between the influence of COVID-19 and other parameters [16]-[17]. The influence of COVID-19 is quite complicated since it impacts several sectors, which are referred to as Multi-Temporal Factors. Based on the occurrence of complexity conditions, machine learning models [18] and statistics [19] are utilized to construct an effective computational model for intervening in COVID-19 pandemic policies [20]-[21].

Based on research problems, this research examines the data validity in some regions regarding implementing the PPKM policy. The data characteristics to be tested are taken from the health sector, including daily positive, daily healing, daily complete vaccine, and daily death. The data from the economic sector include the value of the Indonesian Rupiah and IHSG, while the population mobility sector covers retail recreation, grocery-pharm, parks, stations, workplaces, and residency. The data sources for this research originated from the Ministry of Health of the Republic of Indonesia and the COVID-19 Task Force, including the epidemiological, pandemic, economic, and population mobility data in a specific area during the implementation of the PPKM policy. The data are taken from August 9 to October 25, 2021. This research aims to analyze the characteristics and correlations of the data related to the spread of COVID-19 on the economic sector and population mobility to provide initial information on the effectiveness of government policies in handling the COVID-19 pandemic. The contributions of this research to knowledge development are as follows:

- Using multiple linear regressions to predict the mortality rate in DKI Jakarta, Indonesia, based on multi-temporal factors.
- Using Dynamic Factor to analyze significant Multi-Temporal Variables as the primary reason for mortality.
- Using the F-Test, T-Test, R-Square, and Assumption Test to Evaluate Multiple Linear Regression Models.

## II. Materials and Methods

The data set used to analyze the effects of COVID-19 on the specific sector, such as the economic conditions and population mobility, is taken from Yahoo Finance and Google Mobility. Meanwhile, epidemiological data comes from COVID-19 cases in the Special Region of Jakarta. The data set is a time series taken from August 9 to October 20, 2021. Next, this research uses the CRIPS-DM method, which includes Business Understanding, to analyze the characteristics and correlations of COVID-19 on the economic sector and population mobility.

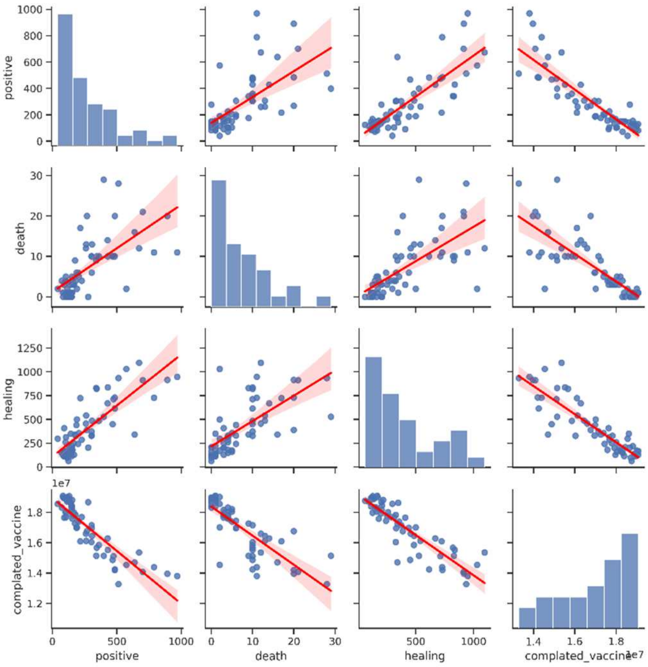


Fig. 1 Epidemiological Dataset Visualization

Figure 1 represents the epidemiological data distribution in DKI Jakarta, Indonesia.

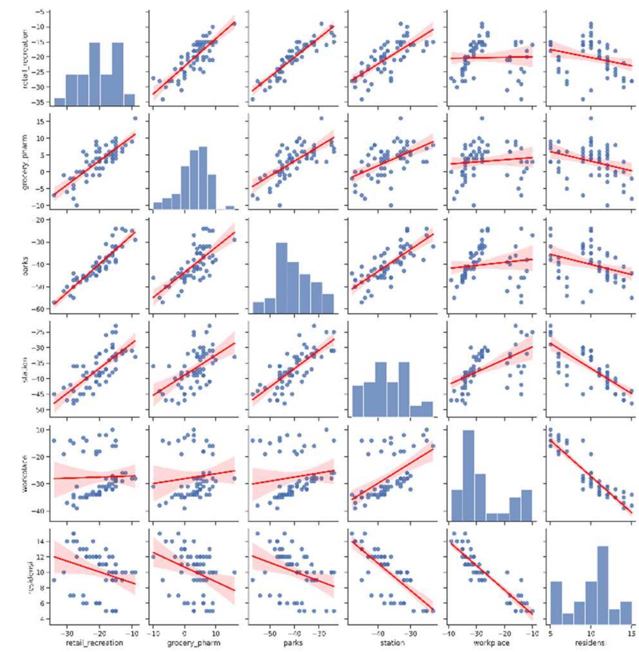

Fig. 2 Mobility Dataset Visualization

Figure 2 represents the mobility data distribution in DKI Jakarta, Indonesia.

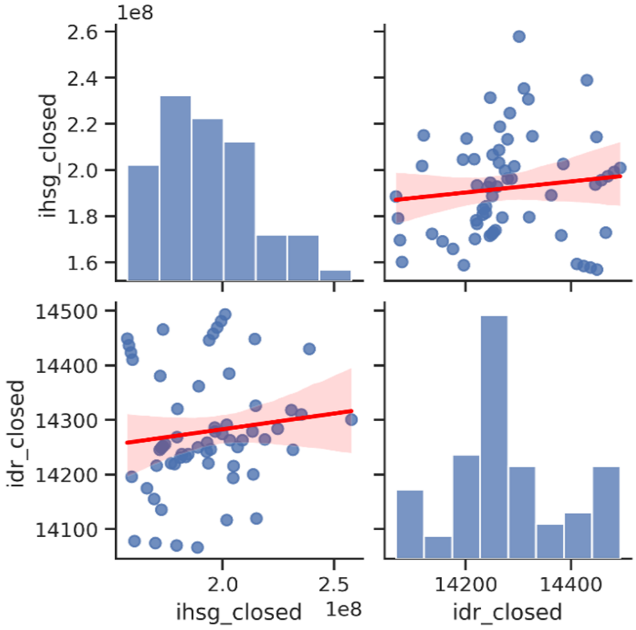


Fig. 3 Econo mic Dataset Visualization

Figure 3 represents the economic data distribution in DKI Jakarta, Indonesia. The set consists of epidemiological, economic, and population mobility data. The Data Preparation uses linear interpolation techniques. The Modeling uses Dynamic Factors and Multiple Linear Regression. The stage Evaluation identifies the quality and effectiveness in the modeling phase by F-Test (ANOVA), T-Test, R-Square, Assumption Test, Linearity Test, Normality Test, Auto Correlation Test, and Homoscedasticity Test. The stage Result means the findings on the research problems. Fig. 4 shows the methodology used in our research.

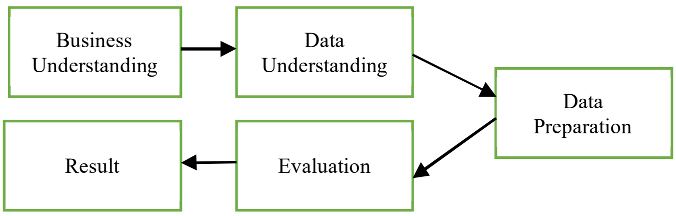


Fig. 4 Research Methodology

*A. Linear Interpolation*

The values are obtained between two data based on a linear equation. Linear interpolation is a method to specify a function value of a linear equation based on the law of proportionality [22]. Linear interpolation is used to input blank values in the data preparation process. The following is the linear interpolation formula:

$$l = \frac{r-t_{value}}{r-d.f} x(d.f - lowest.d.f) \quad (1)$$

Note:
$l$ = interpolation value
r-$t_{value}$ = range of $t_{value}$ from two closest $d.f$

Several studies have been added Linear interpolation was used to estimate the population in many research [23]. Investigate two paths to recovery where the immunization agenda objective (IA2030) is met by 2030 or less than 10% by linearly extrapolating coverage to that year [24]. During the COVID-19 epidemic, aggregate each daily time series and linearly interpolate each funeral time series to produce daily values [25].

*B. Dynamic Factor*

The Dynamic Factor (DF) is a parsimony model that can overcome issues like missing data and mixed frequency and can potentially estimate many variables [26]. Of course, due to the number of Multi-Temporal Factors variables, the DF model is used to generate futures that are the most influential on the impacts of COVID-19. DF is a flexible model for observing multivariate time series. It is a linear function of covariates and unobserved factors and has an autoregressive vector structure. Unobserved factors may also be a function of covariates. It can also find possible dependent variables that allow autocorrelation to occur.

Several studies have been added Dynamic factors are used to analyze dynamic factors that predict general macroeconomic trends [27]. The bootstrap approach improves the performance of limited samples from autoregression in dynamic factor models [28]. This study aims to use dynamic factor models to examine the correlation between school enrollment and unemployment rates at the national level [29].

*C. Multiple Linear Regression (MLR)*

MLR is a model used to investigate the relationship between several responses and predictor variables [30]. The goal of the MLR test is to make a prediction about the value of the dependent or responder variable ($Y$) based on the values of the independent or predictor variables ($X1$, $X2$, ..., $Xn$). Several studies have been add MLR is used to evaluate data in real-time to investigate new active cases of the COVID-19 pandemic [31]. The following is the MLR formula:

$$Y = \alpha + \beta 1X1 + \beta 2X2 + \beta nXn + e \quad (2)$$

Note:
Y = Independent or response variable.
X = Predictor or independent variable.
α = Constant.
β = Slope or estimate coefficient.

*D. F-Test, T-Test, and R-Square*

The F-test (ANOVA), also known as the analysis of variance, is a statistical test that may be used to characterize the connection between independent variables and predictors and determine the level of the variable that performed the best [32]. ANOVA is a decision-making tool to identify variations in process parameters [33] from predetermined predictors [32]. The F-test measures the magnitude of the difference in variance between two or more groups.

The T-test is a statistical method used to measure the effect of the independent variable on the primary variable by adjusting the existing hypothesis [34]. T-test measures the difference between two or more means among groups [34]. T-test aims to measure the difference between two or more means among groups. Meanwhile, the R-square indicates the determinant coefficient used to discover the significant relationship among several variables in a clearer sense. Several studies have been added to the F-Test feature selection in the ensemble stacking model for breast cancer prediction [35]. Several methods for performing ANOVA with binary dependent variables in a 2-way layout are compared to the parametric F-Test [36].

### E. Assumption Test

The assumption test is crucial to ensure that the estimates taken from the analysis are interpreted correctly. The most expedient method of the assumption test reminds back of an introductory statistics course, i.e., plotting the data. By visualizing the correlations, one can assess whether the intention or behavior data deviates from the normal distribution curve and whether there are asymmetries in the data distribution [37]. Classical Assumption Test is an analysis conducted to check for classical assumption problems in an Ordinary Least Square (OLS) linear regression model. The assumption tests used consist of Linearity, Normality, Autocorrelation, and Homoscedasticity tests[38]. Some studies have added assumption checks to estimate the standard error [39]. Putting the assumption test to the test may enhance the evidential value of research [40].

## III. Results and Discussion

### A. Dynamic Factor (DF)

DF is used to explain how much variability in a factor can be attributed to its connection with other related factors, and it is expressed as a number between 0.0 (failed to model the data properly at all) and 1.0 (perfect) [41]. Then, we can find out the important factors that greatly affect it. Fig. 5 shows the dependent and independent variables from the results of the coefficients. There are some of the most influential variables, i.e., *idr_close*, *positive*, *retail_recreation*, *station*, and *healing*.

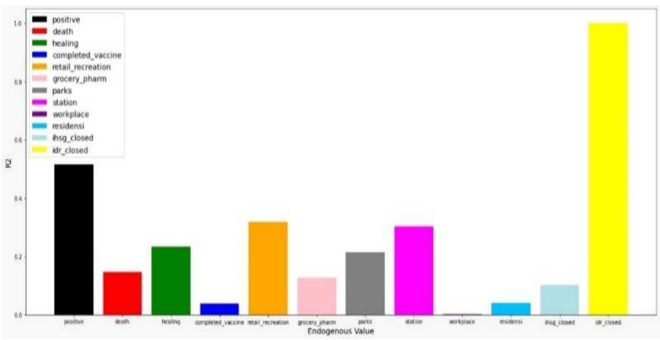


Fig. 5 Plot Coefficients of Determination

### B. Multiple Linear Regression

Of course, there are determined variables in the multiple linear regression models, namely the independent and dependent ones. The independent variables use parameters like *positive*, *healing*, *complete_vaccine*, *retail_recreation*, *grocery_pharm*, *parks*, *station*, *workplace*, *residency*, *ihsg_closed*, and *idr_closed*. Meanwhile, the dependent variables only use *death* as the main parameter.

#### 1) Model Validation:

The machine learning validation uses OLS models. Fig. 6 shows that the F-statistic value is 17.77 and the P-Value is 0.00. The fact that the F-Statistic is higher than the P-Value means the independent variables simultaneously become a significant explanation for the dependent ones, making the regression model feasible for the research. Based on the OLS model, the coefficient of determination (R-Square) is 0.84. It can be assumed that the mortality is influenced by a known factor of 0.84. The residual value of the coefficient of determination is 0.16, and other unidentified factors influence it.

```
                            OLS Regression Results
=======================================================================================
Dep. Variable:               kematian   R-squared (uncentered):                   0.841
Model:                            OLS   Adj. R-squared (uncentered):              0.794
Method:             Least Squares   F-statistic:                                17.77
Date:            Tue, 08 Feb 2022   Prob (F-statistic):                        1.46e-11
Time:                    20:11:24   Log-Likelihood:                           -137.24
No. Observations:              48   AIC:                                        296.5
Df Residuals:                  37   BIC:                                        317.1
Df Model:                      11
Covariance Type:        nonrobust
==============================================================================
                 coef    std err          t      P>|t|      [0.025      0.975]
------------------------------------------------------------------------------
x1            -0.0005      0.007     -0.077      0.939      -0.015       0.014
x2             0.0034      0.006      0.614      0.543      -0.008       0.015
x3          2.489e-06    1.5e-05      0.166      0.869   -2.79e-05    3.29e-05
x4            -1.4169      0.810     -1.750      0.088      -3.057       0.224
x5             0.7789      0.466      1.671      0.103      -0.166       1.723
x6             0.3497      0.355      0.986      0.331      -0.369       1.069
x7            -0.0119      0.641     -0.019      0.985      -1.311       1.287
x8            -0.0430      0.325     -0.132      0.895      -0.701       0.615
x9             0.1788      1.279      0.140      0.890      -2.414       2.771
x10           -0.0058      0.012     -0.493      0.625      -0.030       0.018
x11            0.0015      0.006      0.263      0.794      -0.010       0.013
==============================================================================
Omnibus:                        4.080   Durbin-Watson:                   1.739
Prob(Omnibus):                  0.130   Jarque-Bera (JB):                3.395
Skew:                           0.320   Prob(JB):                        0.183
Kurtosis:                       4.135   Cond. No.                     2.36e+05
==============================================================================

Notes:
[1] R² is computed without centering (uncentered) since the model does not contain a constant.
[2] Standard Errors assume that the covariance matrix of the errors is correctly specified.
[3] The condition number is large, 2.36e+05. This might indicate that there are
strong multicollinearity or other numerical problems.
```

Fig. 6 OLS Regression Result

#### 2) T-test:

From the result of the conducted T-Test, there are yielded hypotheses as follows:

$H0$ = no significant effect

$H1$ = significant effect

$\propto$ = 0.05 (significance level/threshold)

Based on the model validation using OLS, the t-test hypothesis can be described as follows:

- If the value of variable $\boldsymbol{X1}$ (***positive***) is higher than the significance level ⇒ hypothesis $H1$ is accepted.
- If the value of variable $\boldsymbol{X2}$ (***healing***) is higher than the significance level ⇒ hypothesis $H1$ is accepted.
- If the value of variable $\boldsymbol{X3}$ (***complete_vaccine***) is higher than the significance level ⇒ hypothesis $H1$ is accepted.
- If the value of $\boldsymbol{X4}$ (***retail_recreation***) is higher than the significance level ⇒ hypothesis $H1$ is accepted.
- If the value of the variable $\boldsymbol{X5}$ (***grocery_pharm***) is higher than the significance level ⇒ hypothesis $H1$ is accepted.
- If the value of the variable $\boldsymbol{X6}$ (***parks***) is higher the significance level ⇒ hypothesis $H1$ is accepted.
- If the value of the variable $\boldsymbol{X7}$ (***station***) is higher than the significance level ⇒ hypothesis $H1$ is accepted.
- If the value of the variable $\boldsymbol{X8}$ (***workplace***) is higher than the significance level ⇒ hypothesis $H1$ is accepted.
- If the value of the variable $\boldsymbol{X9}$ (***residency***) is higher than the significance level ⇒ hypothesis $H1$ is accepted.
- If the value of the variable $\boldsymbol{X10}$ (***ihsg_closed***) is higher than the significance level ⇒ hypothesis $H1$ is accepted.
- If the value of the variable $\boldsymbol{X11}$ (***idr_closed***) is higher than the significance level ⇒ hypothesis $H1$ is accepted.

It can be concluded that the independent variables that are lower than the level of significance ($X1$, $X2$, $X3$, $X4$, $X5$, $X6$, $X7$, $X8$, $X9$, $X10$, and $X11$) do not affect the dependent ones. Meanwhile, the independent variables that are higher than the level of significance ($X3$) can influence the dependent ones.

*3) Assumption Tests*

Residual analysis is conducted to validate the machine learning model. The following is a list of assumption tests that will be carried out to determine the model validity:

- Linearity Test

The curve regression method is used to conduct the linearity test. This method describes the linear connection between the $X$ and $Y$ variables. The following is the linearity assumption curve. The scattered plots in Fig. 7 show the residuals evenly distributed around the diagonal line, so it can be assumed that the independent and dependent variables are linked in a straight line.

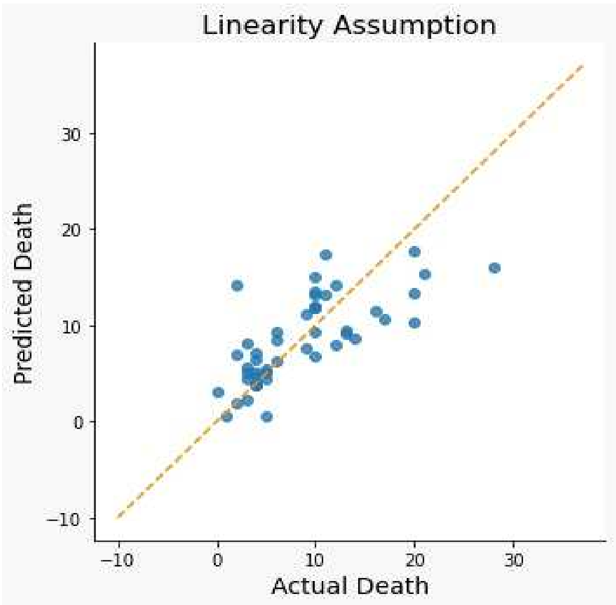


Fig. 7 Linearity Assumption

- Normality Test

If the dependent, independent, or both variables have a normal distribution, then the regression model passes the normality test. The regression model may be considered valid if the data follow a normal distribution or one very near to normal. Based on the assumptions above, the following hypotheses can be written:

$H0$ = normally distributed.

$H1$ = not normally distributed.

Fig. 8 shows that the P-value which is calculated using the Anderson-Darling method, is 0.032. This is lower than the specified threshold value of 0.05, which means that $H0$ is rejected and $H1$ is accepted.

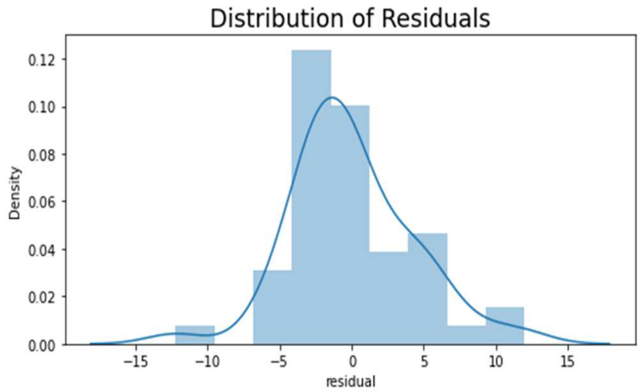


Fig. 8 Residual Distribution

- Autocorrelation

The calculation of the Durbin-Watson [42], results in a score of 1.73, meaning there is little or no autocorrelation, and the assumption is completed. Prior studies compensated for mistakes using neural network autocorrelation for time series [43], and identified a causal association in time series data sets[44].

- Homoscedasticity

Fig. 9 demonstrates that the residual points are distributed uniformly throughout the Y-axis, above and below 0 (zero). The conclusion holds that the regression model displays no signs of heteroscedasticity. The simultaneous evaluation of normality and homoscedasticity in linear fixed effects models is supported by prior research [45], [39].

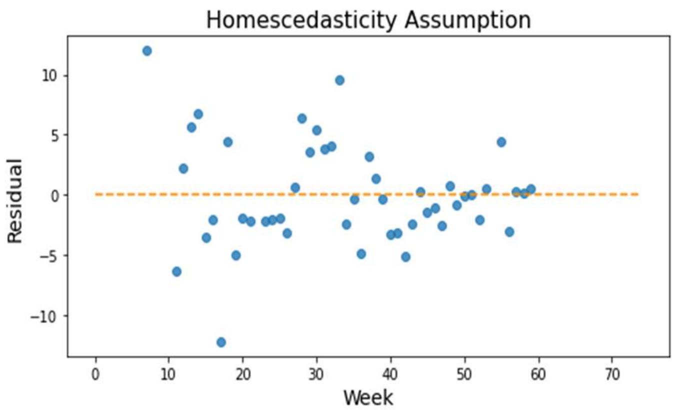


Fig. 9 Homoscedasticity Assumption

*C. Data Visualization*

The data visualization is also conducted to understand the data distribution pattern better. Three observed factors are health, economic sector, and people mobility.

*1) Health*

Fig. 10 Suggests that the health sector's adoption of the PPKM strategy has resulted in a significant drop in the death rate.

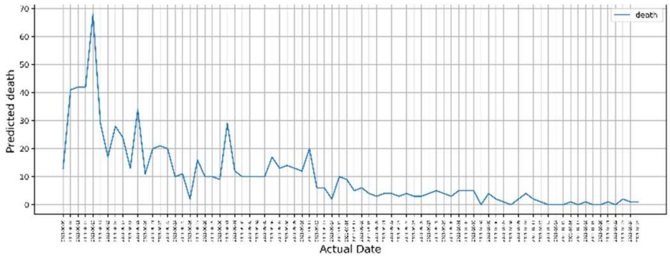


Fig. 10 Trend on the Death Rate

### 2) *Economic Sector*

From Fig. 11, it can be concluded that implementing the PPKM policy from the economic sector at the *idr_close* parameter has a practical impact, but the decreasing inflation is not so significant.

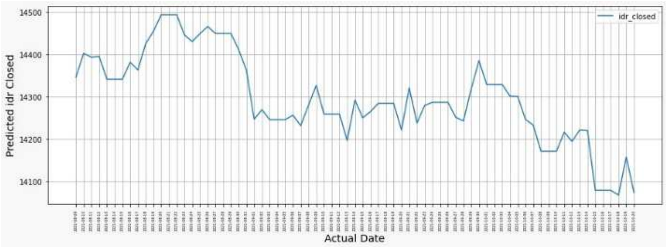


Fig. 11 Trend on *idr_close* Data

Fig. 12 shows that the implementation of the PPKM policy has an effective impact on the composite stock price index.

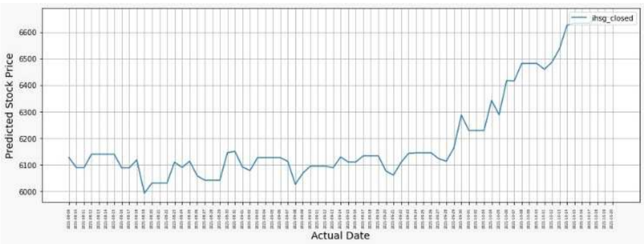


Fig. 12 Trend on the Stock Price Index Data

### 3) *People's Mobility*

Fig. 13 shows that people's mobility has been limited to stop the spread of COVID-19, but there is more mobility in retail, recreation, and grocery farms. This shows that the implementation of the PPKM policy poses a problem.


Fig. 13 Trend on People's Mobility Data

## IV. Conclusions

Based on the OLS model, the accuracy value is 84%. The residual value of the coefficient of determination is 0.16, meaning other unidentified factors influence 16%. This is a good accuracy value, so it can be stated that the multiple linear regression model performs well as expected. The multiple linear regression models have successfully passed all the tests in the model validation step. In the F-test, all independent variables significantly affect the dependent ones. Furthermore, based on the Assumption Test, there is a linear relationship between the independent and dependent variables. The normality assumption has been met, no autocorrelation found in the data, and no symptoms of heteroscedasticity in the model. Therefore, based on the quality of data on trends in health, economics, and people's mobility (the Multi-Temporal Factors), the implementation of PPKM policy by the government has effective impacts, although there are still some dilemmas found in the economic sector and people's mobility.

For future study, investigate first the most critical aspects that are believed to be the primary determinants of the COVID-19 pandemic model in Indonesia. They rank second, using numerous criteria besides death. Lastly, develop a new pandemic model by comparing current machine learning models. Fourth, a strategy to optimize the goal function of all mortality-influencing factors is needed. Increase the quantity of datasets available for future study. Sixth, location and time objects are added to compare the pandemic models and see whether they follow the same pattern. In addition to, R-Square and other variables such as RMSE, MAE, and MAPE is required to evaluate the model.

## Acknowledgment

We thank Telkom University for its financial support in this research.

## References

[1] Public Health Emergency of International Concern, "WHO Coronavirus (COVID-19) Dashboard," *WHO*, p. 1, 2021. [Online]. Available: https://covid19.who.int/

[2] E. W. Mansi, "Environmental Health Aspect of The Novel Corona Virus Disease and Its Global Impact," Journal of Applied Science, Engineering, Technology, and Education, vol. 3, no. 2, pp. 151–159, Nov. 2021, doi: 10.35877/454ri.asci131.

[3] P. Bose, S. Roy, and P. Ghosh, "A Comparative NLP-Based Study on the Current Trends and Future Directions in COVID-19 Research," IEEE Access, vol. 9, pp. 78341–78355, 2021, doi: 10.1109/access.2021.3082108.

[4] Public Health Emergency of International Concern, "WHO declared COVID-19 emergency a global pandemic," *WHO*, 2021. https://www.who.int/news-room/spotlight/a-year-without-precedent-who-s-covid-19-response (accessed October 20, 2021).

[5] L. J. Muhammad, E. A. Algehyne, S. S. Usman, A. Ahmad, C. Chakraborty, and I. A. Mohammed, "Supervised Machine Learning Models for Prediction of COVID-19 Infection using Epidemiology Dataset," SN Computer Science, vol. 2, no. 1, Nov. 2020, doi: 10.1007/s42979-020-00394-7.

[6] Q. Lin et al., "A conceptual model for the coronavirus disease 2019 (COVID-19) outbreak in Wuhan, China with individual reaction and governmental action," International Journal of Infectious Diseases, vol. 93, pp. 211–216, Apr. 2020, doi: 10.1016/j.ijid.2020.02.058.

[7] S. Lee, T.-Q. Peng, M. K. Lapinski, M. M. Turner, Y. Jang, and A. Schaaf, "Too stringent or too Lenient: Antecedents and consequences of perceived stringency of COVID-19 policies in the United States," Health Policy OPEN, vol. 2, p. 100047, Dec. 2021, doi: 10.1016/j.hpopen.2021.100047.

[8] SATGAS COVID 19 Indonesia, "Peta Sebaran COVID 19," *SATGAS COVID 19 Indonesia*, 2021. https://covid19.go.id/peta-sebaran-covid19

[9] Rohmi Aida Nur, "Daftar 10 Negara dengan Kasus Covid-19 Tertinggi di Asia, Indonesia Nomor 4," *Kompas*, 2020. https://www.kompas.com/tren/read/2020/12/11/160955265/daftar-10-negara-dengan-kasus-covid-19-tertinggi-di-asia-indonesia-nomor-

4?page=all
[10] K. D. N. R. Indonesia, "Pemberlakuan Pembatasan Kegiatan Masyarakat Darurat COVID 19 Di Wilayah Jawa dan BALI," 2021.
[11] L. Thunström, S. C. Newbold, D. Finnoff, M. Ashworth, and J. F. Shogren, "The Benefits and Costs of Using Social Distancing to Flatten the Curve for COVID-19," Journal of Benefit-Cost Analysis, vol. 11, no. 2, pp. 179–195, 2020, doi: 10.1017/bca.2020.12.
[12] A. Kaim, T. Gering, A. Moshaiov, and B. Adini, "Deciphering the COVID-19 Health Economic Dilemma (HED): A Scoping Review," International Journal of Environmental Research and Public Health, vol. 18, no. 18, p. 9555, Sep. 2021, doi: 10.3390/ijerph18189555.
[13] A. Akbar Gozali, "Pre-Development Analysis of Truther Framework: A Platform to Verify Information Authenticity," 2020 International Conference on Radar, Antenna, Microwave, Electronics, and Telecommunications (ICRAMET), Nov. 2020, doi: 10.1109/icramet51080.2020.9298635.
[14] R. Wang, "Measuring the Effect of Government Response on COVID-19 Pandemic: Empirical Evidence from Japan," COVID, vol. 1, no. 1, pp. 276–287, Aug. 2021, doi: 10.3390/covid1010022.
[15] L. Chen, D. Raitzer, R. Hasan, R. Lavado, and O. Velarde, "What Works to Control COVID-19? Econometric Analysis of a Cross-Country Panel," SSRN Electronic Journal, 2020, doi: 10.2139/ssrn.3785083.
[16] S. F. Pane, R. R. Fajri, A. G. Putrada, N. Alamsyah, M. N. Fauzan, and R. M. Awangga, "Non-Academic Factors Analysis Impacting Students' Online Learning During the COVID-19 Pandemic in Universitas Logistic and Bisnis Internasional," 2022 2nd International Conference on Intelligent Cybernetics Technology & Applications (ICICyTA), Dec. 2022, doi: 10.1109/icicyta57421.2022.10037986.
[17] S. F. Pane, Heriyanto, A. G. Putrada, N. Alamsyah, and M. N. Fauzan, "The Influence of The COVID-19 Pandemics in Indonesia On Predicting Economic Sectors," 2022 Seventh International Conference on Informatics and Computing (ICIC), Dec. 2022, doi: 10.1109/icic56845.2022.10006897.
[18] S. F. Pane, Adiwijaya, M. D. Sulistiyo, and A. A. Gozali, "LSTM and ARIMA for Forecasting COVID-19 Positive and Mortality Cases in DKI Jakarta and West Java," 2022 Seventh International Conference on Informatics and Computing (ICIC), Dec. 2022, doi: 10.1109/icic56845.2022.10006959.
[19] J. Rocklöv and A. A. Gayle, "Data Science and Machine Learning: Mathematical and Statistical Methods. D.P. Kroese, Z.I. Botev, T. Taimre, and R. Vaisman," International Journal of Epidemiology, vol. 49, no. 6, pp. 2094–2095, May 2020, doi: 10.1093/ije/dyaa072.
[20] L. Liu, Z. Zhang, H. Wang, S. Wang, S. Zhuang, and J. Duan, "Comparing modelling approaches for the estimation of government intervention effects in COVID-19: Impact of voluntary behavior changes," *PLoS One*, vol. 18, no. 2, p. e0276906, 2023.
[21] S. F. Pane and J. Ramdan, "Pemodelan Machine Learning: Analisis Sentimen Masyarakat Terhadap Kebijakan PPKM Menggunakan Data Twitter," *J. Sist. Cerdas*, vol. 5, no. 1, pp. 12–20, 2022.
[22] E. Puniach, W. Gruszczyński, T. Stoch, D. Mrocheń, P. P. E. Ćwi\kakała Paweland Sopata, and W. Matwij, "Determination of the coefficient of proportionality between horizontal displacement and tilt change using UAV photogrammetry," *Eng. Geol.*, vol. 312, p. 106939, 2023.
[23] N. Marquez, J. A. Ward, K. Parish, B. Saloner, and S. Dolovich, "COVID-19 incidence and mortality in federal and state prisons compared with the US population, April 5, 2020, to April 3, 2021," *JAMA*, vol. 326, no. 18, pp. 1865–1867, 2021.
[24] J. Toor et al., "COVID-19 impact on routine immunisations for vaccine-preventable diseases: Projecting the effect of different routes to recovery," Vaccine, vol. 40, no. 31, pp. 4142–4149, Jul. 2022, doi: 10.1016/j.vaccine.2022.05.074.
[25] A. Warsame *et al.*, "Excess mortality during the COVID-19 pandemic: a geospatial and statistical analysis in Mogadishu, Somalia," *Int. J. Infect. Dis.*, vol. 113, pp. 190–199, 2021.
[26] F. Bräuning and S. J. Koopman, "Forecasting macroeconomic variables using collapsed dynamic factor analysis," International Journal of Forecasting, vol. 30, no. 3, pp. 572–584, Jul. 2014, doi: 10.1016/j.ijforecast.2013.03.004.
[27] C. Doz and P. Fuleky, "Dynamic factor models," *Macroecon. Forecast. Era Big Data Theory Pract.*, pp. 27–64, 2020.
[28] M. Shintani and Z.-Y. Guo, "Improving the finite sample performance of autoregression estimators in dynamic factor models: A bootstrap approach," *Econom. Rev.*, vol. 37, no. 4, pp. 360–379, 2018.
[29] F. Blasques, M. H. Hoogerkamp, S. J. Koopman, and I. van de Werve, "Dynamic factor models with clustered loadings: Forecasting education flows using unemployment data," International Journal of Forecasting, vol. 37, no. 4, pp. 1426–1441, Oct. 2021, doi: 10.1016/j.ijforecast.2021.01.026.
[30] D. J. Eck, "Bootstrapping for multivariate linear regression models," Statistics & Probability Letters, vol. 134, pp. 141–149, Mar. 2018, doi: 10.1016/j.spl.2017.11.001.
[31] S. Rath, A. Tripathy, and A. R. Tripathy, "Prediction of new active cases of coronavirus disease (COVID-19) pandemic using multiple linear regression model," *Diabetes \& Metab. Syndr. Clin. Res. \& Rev.*, vol. 14, no. 5, pp. 1467–1474, 2020.
[32] D. V. Gradov et al., "Modelling of a continuous veneer drying unit of industrial scale and model-based ANOVA of the energy efficiency," Energy, vol. 244, p. 122673, Apr. 2022, doi: 10.1016/j.energy.2021.122673.
[33] A. Korgal, S. Upadhyaya, and A. T, "Grain refinement of aluminium 4032 alloy with the impact of vibration using Taguchi technique and analysis of variance (ANOVA)," Materials Today: Proceedings, vol. 54, pp. 507–512, 2022, doi: 10.1016/j.matpr.2021.11.504.
[34] Y. Feng, Y. Huang, and X. Ma, "The application of Student's t -test in internal quality control of clinical laboratory," Frontiers in Laboratory Medicine, vol. 1, no. 3, pp. 125–128, Sep. 2017, doi: 10.1016/j.flm.2017.09.002.
[35] R. Dhanya, I. R. Paul, S. S. Akula, M. Sivakumar, and J. J. Nair, "F-test feature selection in Stacking ensemble model for breast cancer prediction," Procedia Computer Science, vol. 171, pp. 1561–1570, 2020, doi: 10.1016/j.procs.2020.04.167.
[36] H. Luepsen, "ANOVA with binary variables: the F-test and some alternatives," *Commun. Stat. Comput.*, pp. 1–25, 2021.
[37] I. N. T. Sutaguna, M. Yusuf, R. Ardianto, and P. Wartono, "The Effect Of Competence, Work Experience, Work Environment, And Work Discipline On Employee Performance," *Asian J. Manag. Entrep. Soc. Sci.*, vol. 3, no. 01, pp. 367–381, 2023.
[38] P. Mugebe, M. Kizil, M. Yahyaei, and R. K. Y. Low, "Regression Analysis Applicability in Investigating Mining Technology and Macroeconomic Impact on Mining Companies' Share Price--Case Study: Iron Ore Suppliers," *Available SSRN 4354878*.
[39] A. F. Schmidt and C. Finan, "Linear regression and the normality assumption," Journal of Clinical Epidemiology, vol. 98, pp. 146–151, Jun. 2018, doi: 10.1016/j.jclinepi.2017.12.006.
[40] J. Uttley, "Power Analysis, Sample Size, and Assessment of Statistical Assumptions—Improving the Evidential Value of Lighting Research," LEUKOS, vol. 15, no. 2–3, pp. 143–162, Jan. 2019, doi: 10.1080/15502724.2018.1533851.
[41] M. Schwarz, J. Trippel, M. Engelhart, and M. Wagner, "Dynamic alpha factor prediction with operating data-a machine learning approach to model oxygen transfer dynamics in activated sludge," *Water Res.*, p. 119650, 2023.
[42] S. Adrianto, I. H. N. Balqis, C. Z. N. Soetanto, and M. Ohyver, "Cochrane orcutt method to overcome autocorrelation in modeling factors affecting the number of hotel visitors in Indonesia," Procedia Computer Science, vol. 216, pp. 630–638, 2023, doi: 10.1016/j.procs.2022.12.178.
[43] F.-K. Sun, C. Lang, and D. Boning, "Adjusting for autocorrelated errors in neural networks for time series," *Adv. Neural Inf. Process. Syst.*, vol. 34, pp. 29806–29819, 2021.
[44] J. Runge, "Discovering contemporaneous and lagged causal relations in autocorrelated nonlinear time series datasets," in *Conference on Uncertainty in Artificial Intelligence*, 2020, pp. 1388–1397.
[45] Y. Yang and T. Mathew, "The simultaneous assessment of normality and homoscedasticity in linear fixed effects models," *J. Stat. Theory Pract.*, vol. 12, pp. 66–81, 2018.