

Correlation between Covid-19 and weather/climate indicators: A Response

Ansari Saleh Ahmar

Department of Statistics, Universitas Negeri Makasar, Makassar, Indonesia

Abstract

This is a response to article by Dr. Ramadhan Tosepu and his colleagues entitled: “Correlation between the weather and the Covid-19 pandemic in Jakarta, Indonesia” in *Science of The Total Environment*, Vol. 725; and Dr. Muhammad Farhan Bashir and his colleagues entitled: “Correlation between the climate indicators and the Covid-19 pandemics in New York, USA.” in *Science of The Total Environment*, Vol. 728. We agree with the results of their research, but there are things that need to be clarified in relation to the selection of the Spearman-rank correlation as an analytical tool, because the selection of this analysis is without the normality test of the data. Spearman-rank correlation is useful for non-normally distributed continuous data; can be used for ordinal data; and is relatively robust for outliers.

© 2020 Author(s). All rights reserved.

Keywords: Weather, COVID-19, climate indicators.

Response

Response:

We read with great interest Article “Correlation between weather and Covid-19 pandemic in Jakarta, Indonesia” by Dr. Ramadhan Tosepu and his colleagues in *Science of The Total Environment*, Vol. 725; and Article “Correlation between climate indicators and Covid-19 pandemic in New York, USA” by Dr. Muhammad Farhan Bashir and his colleagues in *Science of The Total Environment*, Vol. 728. The article may give us insight into whether or not there is a correlation between Covid-19 and the weather/climate indicators. But there are a few things I would like to comment on in this article.

Spearman-rank correlation is a nonparametric test and (1) is useful for non-normally distributed continuous data (Akoglu, 2018; Gogtay and Thatte, 2017; Mukaka, 2012); (2) can be used for ordinal data; and (3) is relatively robust for outliers (Schober et al., 2018). To find out whether or not the data is normally distributed, a data normality test must be carried out. That is to say, in conducting a study that addresses the correlation needed to test for normality.

Based on the results of the research conducted by Ahmar, et.al. 2020, it is shown that the minimum temperature, maximum temperature and average temperature data are normally distributed for the City of Semarang (Indonesia) and this is in contrast to that provided by Dr. Tosepu, et al and Dr. Bashir, et al, who stated that the data does not normally be distributed without normal test results.

* Corresponding author.

E-mail address: ansari@unm.ac.id (Ansari Saleh Ahmar)

It is also necessary to consider the scale of the data when conducting a data analysis. Of course, the nominal value of money differs from the temperature; the level of air quality differs from the temperature; the rainfall differs from the wind speed; the number of Covid-19 patients differs from the wind speed because each data has a different maximum value. For example, humidity has a maximum value of 100%, while the maximum value of temperature, wind speed and number of patients may be above 100. The difference in the size of the data will certainly affect the analysis of the data. In order for this data scale to be the same, it is necessary to standardize the data (Heyer, 1999). This is in line with the opinion of (Gewers et al., 2018), who argued that standardization would make data of the same range of variations. Standardization is also used as a means of avoiding the bias of certain variables when the original variables have very different dispersions or scales. Standardization has little effect if the original data measured had a similar dispersion. In addition, (Kim, 2018) said that the correlation is obtained on the basis of a standardized scale resulting in a unit-less measure.

Based on this, I would like to suggest two things when conducting a correlation test: (1) conducting a data normality test to determine which correlation test is appropriate for use, whether it is a parametric correlation test or a nonparametric correlation test; (2) see the scale of the data whether it is the same size or different, if it is different, then standardize the data to minimize the dispersion of the data.

References

- A.S., 2020. Temperature and Covid-19 dataset in Semarang City: correlation between temperature and Covid-19. (Working Papers).
- Akoglu, H., 2018. User's guide to correlation coefficients. *Turkish J. Emerg. Med.* 18, 91–93. <https://doi.org/10.1016/j.tjem.2018.08.001>
- Gewers, F.L., Ferreira, G.R., de Arruda, H.F., Silva, F.N., Comin, C.H., Amancio, D.R., Costa, L. da F., 2018. *Principal Component Analysis: A Natural Approach to Data Exploration*.
- Gogtay, N.J., Thatta, U.M., 2017. Principles of correlation analysis. *J. Assoc. Physicians India* 65, 78–81.
- Heyer, L.J., 1999. Exploring Expression Data: Identification and Analysis of Coexpressed Genes. *Genome Res.* 9, 1106–1115. <https://doi.org/10.1101/gr.9.11.1106>
- Kim, H.-Y., 2018. Statistical notes for clinical researchers: covariance and correlation. *Restor. Dent. Endod.* 43. <https://doi.org/10.5395/rde.2018.43.e4>
- Mukaka, M.M., 2012. A guide to appropriate use of correlation coefficient in medical research. *Malawi Med. J.* 24, 69–71.
- Schober, P., Boer, C., Schwarte, L.A., 2018. Correlation Coefficients. *Anesth. Analg.* 126, 1763–1768. <https://doi.org/10.1213/ANE.0000000000002864>