

COVID-19: About the inequality of the territorial distribution of the number of deaths in the world and the indicator of the quality of epidemic management

Antoni Wilinski (✉ awilin@o2.pl)

WSB University in Gdansk, Poland

Arti MK

NSUT East Campus (AIACTR), New Delhi, INDIA

Lukasz Kupracz

Koszalin University of Technology, Poland

Research Article

Keywords: Keywords: COVID-19, Gini coefficient, management of pandemics

Posted Date: July 7th, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-688333/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

COVID-19: About the inequality of the territorial distribution of the number of deaths in the world and the indicator of the quality of epidemic management

Antoni Wilinski¹, Arti M.K.², Lukasz Kupracz³

1 – WSB University in Gdansk, Poland

2 - NSUT East Campus (AIACR), New Delhi, INDIA

3 - Koszalin University of Technology, Poland

Highlights

1. The number of deaths in COVID-19 is a more stable indicator of the severity of the pandemic
2. Good pandemic management is about distributing resources fairly among the region
3. A good indicator of the even distribution of deaths is the Gini index.
4. There are evident differences among the leader states affected by the pandemic

Abstract

The aim of the article is to present some kind of injustice of fate in the aspect of uneven territorial distribution of post-Covid-19 deaths. On a global scale, this unevenness is evident because it is not controlled by anything. In particular countries, it should be the result of efficient and honest management of the spread of the epidemic. The authors assume in the proposed algorithm that the number of deaths in individual organizational units of the state (regions, states, provinces, etc.) is known in the existing administrative division. By considering the population numbers in these units, the Lorentz curve is prepared, and the Gini coefficient is calculated - for the entire world and for individual countries such as the USA, India, Brazil, Poland and the Balkan and Eastern European countries with the highest number of deaths per million inhabitants in the world. Moreover, an attempt was made to present the universal mortality rate in a given country in the form of a bicriterion combining the Gini index and the number of deaths per million inhabitants achieved.

Keywords: COVID-19, Gini coefficient, management of pandemics

Introduction

Even in a first view at the world map with marked on the confirmed cases of infections or deaths intensity allows the first observation of a clear uneven geographic distribution of the pandemic victims. The main source of information on the distribution of infections and deaths worldwide will be the website of the CSSE institute of Johns Hopkins University in Baltimore (CSSE JHU) [8]. This is the primary source of knowledge cooperating with the WHO and meticulously registering cases from almost all countries in the world since January 20, 2020. The home page of this site also displays a graphical world map with countries and regions highlighted. The number of infections is visualized

using circles of the appropriate diameter. The first glance at the board confirms the uneven distribution of the number of deaths. The numerical raw data is located in the GitHub repository. The second important data source used in this paper is the worldometers page [10], which provides a lot of additional information, including indicators derived from basic data, e.g. the number of infections and deaths per million inhabitants of a given country. Therefore, from this source we will obtain information about the number of inhabitants of a given country.

Unfortunately, based on data [10,11] it is not possible to assess the distribution of deaths among regions of a given country. This requires additional research in a space of ever-changing data. An important, almost philosophical finding in the world of numbers is the authors' decision to focus solely on deaths. These are definitely more reliable data than the confirmed cases of contagions.

There will not be analyzed the issues of the accuracy of recording these deaths by local authorities, their relationship with comorbidities, political and social aspects of data disclosure, etc. In this article we deal only with the statistics disclosed on the CSSE JHU websites [8]. The intention of the authors can be summarized briefly - we are dealing with the statistical effects of a pandemic, leaving the research of the cause to other researchers. It is worth mentioning, however, that among the countries considered here, there may be large differences in formal, local law-related methods of pandemic management resulting, for example, from the autonomy of regions and economic differentiation. These differences that result in the distribution of deaths can be measured. The so-called Gini index is used here. It enables a numerical assessment of the degree of non-uniformity of the distribution of the considered feature after plotting the Lorentz curve. The more even the distribution among regions / provinces / counties or states, the higher the society of a given country should assess pandemic crisis management efficiency. Of course, a natural dilemma will arise, what is more important? Or in other words, what's worse? Whether the high average level of cumulative deaths in the entire country or a high disproportion among the regions of a given country. There is an attached table in this paper that allows to compare these two assessments. An attempt will also be made to develop a bi-criteria taking into account both the Gini coefficient and the cumulative number of deaths in the selected country.

Many authors deal with the problem of deaths in a pandemic. Due to its relatively short history, during the analysis of related the publications, it is also worth noting a number of preprints. Most of the scientific literature related to pandemic modelling is devoted to the spread of infection and to a lesser extent the deaths. The last four publications on this issue were selected [1-4]. Neither of these works deals with the degree of uniformity of infections and deaths, but rather focuses on the risk of infection and death under certain specific conditions. However, we found three publications [5-7] in which similar statistical treatments are used, which differentiate the regions of the country or entire countries.

The research is performed for data available on June 17, 2021.

Model of uneven territorial distribution of the number of deaths.

Let the target territory (e.g. country, a group of countries, or even the whole world) T be given, which can be conventionally divided into N disjoint smaller territories T_i consisting of the target territory T :

$$\forall T_i \in T, \sum_{i=1}^N T_i = T \quad (1)$$

Each T_i territory has two features: P_i and D_i , where

P_i – population in millions measured in 2021 year

D_i – cumulated deaths caused by SARS-CoV-2 till the end of May 2021

Let's introduce a variable – deaths' number for each million of habitants of i -th territory of a given country (T_i) as:

$$L_i = \frac{D_i}{P_i}; i = 1, 2, \dots, N \quad (2)$$

Where P_i should be given in millions of habitants (there are sources calculating these statistics in 100 thousands of habitants).

In further considerations we use the integer variable M_i of the one-million population granules in a given region of T_i , which can be determined after rounding:

$$M_i = \text{round}\left(\frac{P_i}{L_i}\right); i = 1, 2, \dots, N \quad (3)$$

The idea of determining the uneven distribution of infections and deaths in a given country will be targeted to the Lorentz curve model known from economics as a function of income distribution in a given society. It allows you to determine what percentage of the society earns income above or below a certain level.

For the purposes of this model, let us imagine that the entire population of a given country is divided into one-million granules that differ from each other by a second feature - the number of deaths per million. Accordingly, in every T_i territory there will be granules identical to D_i . If we sort these granules from those with the lowest D_i value starting with those with the highest D_i value, they will form a stepped curve. The lengths of the horizontal sections of such a curve will correspond to the number of people in particular regions of T_i , and step changes on the ordinate axis will correspond to changes in the D_i feature - the number of deaths per million. The integral of such a step curve is called the Lorentz curve. It will be a polyline curve with the features of the income curve already mentioned in economics.

Formally, the Lorentz's curve can be explained with below pseudocode:

```
k = 0; %one-million granules counter
for i = 1 : N %regions from 1 to N
    for j = 1 : Mi % Mi - number of granules in i-th region
        k = k+1;
        zk = Li
    end for j
end for i
```

Execution of the above loops effects in creation of a vector Z_k , $k=1, 2, \dots, k$ granules of a different size, approximately equal to the number of millions of habitants in a given country (sum of P_i). The Z_k vector length is marked as K :

$$K = \text{sum}(M_i) \quad (4)$$

Let's sort these one-million granules accordingly to their L_i value ascending

$$Z_S = \text{sort}(Z_k), \text{ascend} \quad (5)$$

Then we calculate the integral of the above curve according to the formula:

$$Z_C(k) = \text{sum}(1:k) Z_S \text{ for } k = 1, 2, \dots, K \quad (6)$$

In order to normalize this curve lets calculate a sum of all Z_S as $Z = \text{sum}(1:K) Z_C$

$$Z_n(k) = \frac{Z_C(k)}{Z} \quad (7)$$

We also put a normalized variable $k_n = \frac{k}{K}$ on the abscissa axis. In this way, we move along the demographic resources of a given country to express them in the form of a number belonging to the range [0,1].

Lorentz's curves [5] and the resulting Gini coefficients [6] we calculate for the different countries and the whole world.

Death distribution studies for various countries and for the whole world

In order to check the uniformity of the distribution of the number of deaths in the scale of the whole world, the data on the www.worldometers.info [10] pages was used.

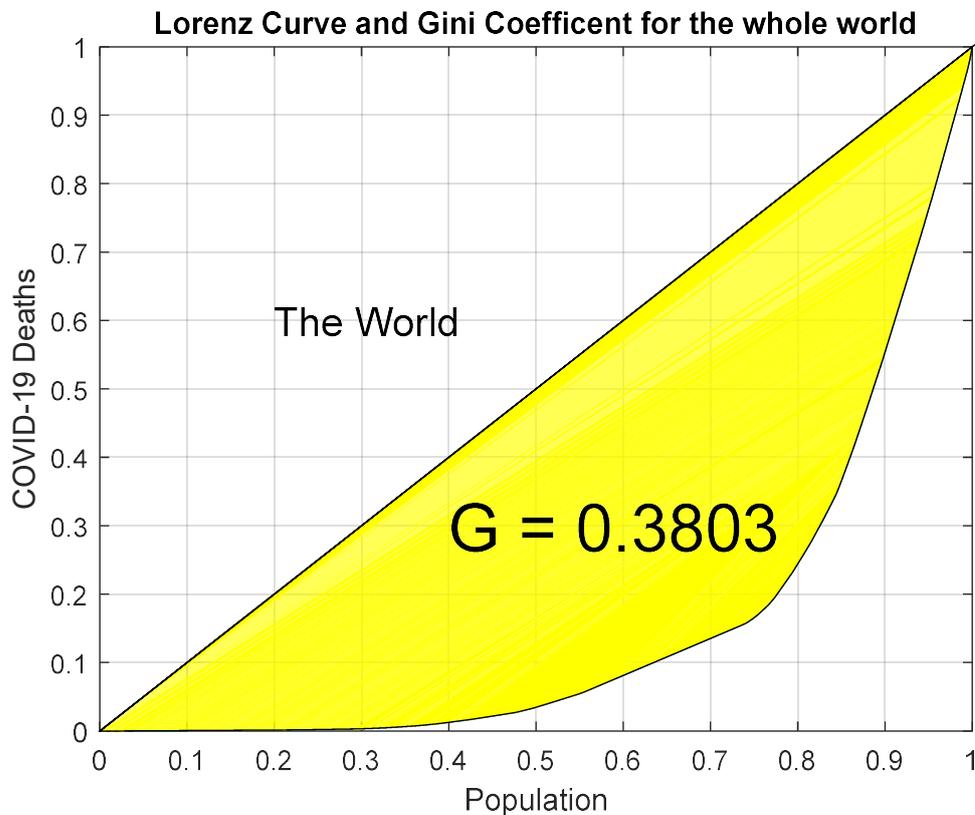


Fig. 1. The Gini coefficient indicates a definitely uneven distribution of the number of deaths on a global scale among individual countries.

Taking into account the data of 220 countries listed in [10], the world average number of deaths per million inhabitants is 580.9 and the standard deviation is 917.8 (end of May 2021). After performing the operation in accordance with the described pseudocode, it is possible to determine the Lorentz curve [5] for the whole world, which in Fig. 1 is shown by an arched broken curve that delimits the yellow field from below. If the area under the Lorentz curve is denoted as A, then the Gini coefficient [9] is the area G expressed as a part (percentage) of the lower triangle of the normalized plot.

$$G = 0.5 - A \tag{8}$$

In the formula above, 0.5 denotes the area of the bottom triangle in the normalized 1 x 1 square. The diagonal of this square is the Lorentz curve for the idealized completely uniform distribution of the feature under study.

The graph - Fig. 1 shows the characteristic Pareto Law [18] - approximately 80% of the world's population lives in regions that experienced only 20% of all covid deaths.

Similar studies were carried out for a few selected countries. The USA, India, Brazil, Poland and a group of Balkan and Eastern European countries were selected. The latter choice was dictated by the high global concentration of very large numbers of deaths. At the end of May, the top 10 countries in Eastern Europe and the Balkans were among the 20 countries with the highest number of deaths per million inhabitants in the world.

To conduct the research of a selected country (or a group of countries), two vectors P_i (number of inhabitants in particular regions of a given country) and D_i (number of cumulative deaths in these regions) defined in (1) are necessary.

Results are depicted in Fig. 2.

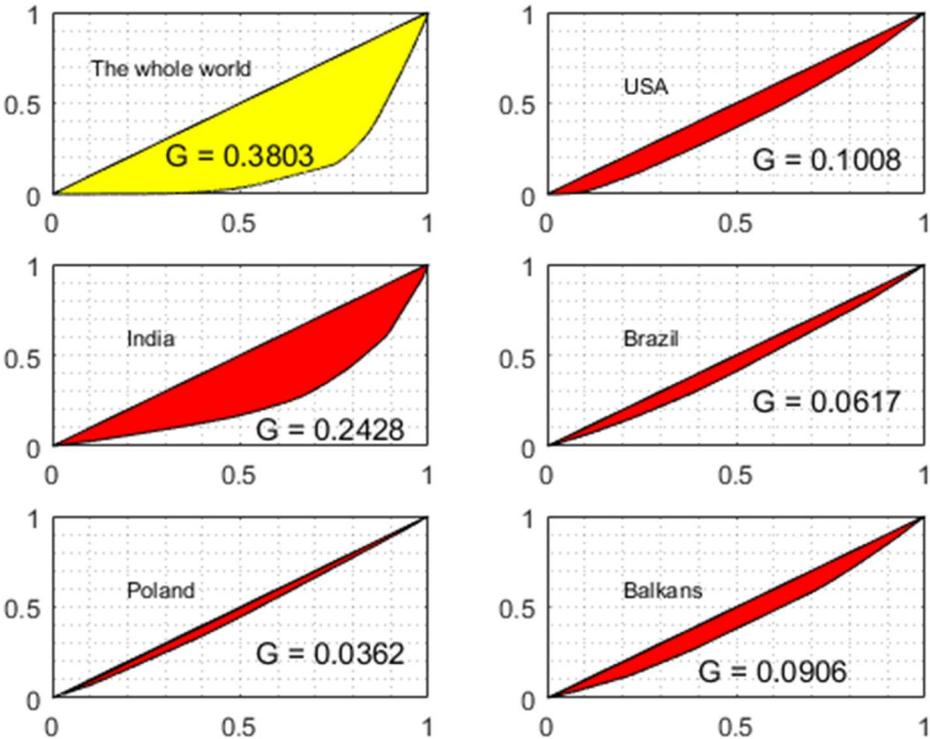


Fig. 2. Gini indexes for selected countries

The results presented in Fig. 2 show a fairly even distribution of deaths between provinces in Poland and quite equal, but worse among 51 US states. In turn, the Gini coefficient in India is very high, confirming the assumption about the huge ethnic, cultural and demographic diversity of this country. The Balkans and Eastern Europe are an artificially created set of states and the high Gini index should not be a surprise here.

Detailed data on the administrative divisions and statistical features of the distribution of deaths are presented in Tab.1.

Universal indicator of the quality of epidemic management

Taking into account both considered factors - the Gini index G and the number of deaths per million inhabitants D , one could try to fuse the influences of these factors by creating a bi-criterion. There is a problem of the weighting of both these factors in such a common criterion. Indicators of this type are created mainly for their comparative or classification functions. So how to set any reference level in this case, how to determine the importance (weight) of the first and second factors?

Perhaps a convincing object for comparisons can be found? Let it be, for example, the whole world.

For the whole world, the Gini index was $G_W = 0.3803$ and the number of deaths per million inhabitants $D_W = 580.9$. Let us assume that we treat both of these factors as equivalent and that we will try to define the global mortality level as a benchmark to which we will refer the respective levels in different countries.

Let the contractual level of management of the pandemic M_W in terms of global mortality be defined as

$$M_W = G_W \times k_G + D_W \times k_D \quad (9)$$

and let's assign this variable the value 1:

$$M_W = 1 - \text{for the entire world,} \quad (10)$$

If the both factors should be equally important, then

$$G_W \times k_G = 0,5; \text{ and} \quad (11)$$

$$D_W \times k_D = 0,5 \quad (12)$$

From this, the value of the multipliers k follows:

$$k_G = \frac{0,5}{G_W} \quad (13)$$

$$k_D = \frac{0,5}{D_W} \quad (14)$$

After calculation – $k_G = \frac{0,5}{0,3803} = \mathbf{1,31}$ and $k_D = \frac{0,5}{580,9} = \mathbf{0,00086}$

Obviously, the k_G and k_D multipliers would be different if we made a different assumption, instead of equivalence, to change the significance of each of the factors G and D , e.g. $(\frac{0,6}{0,4}$ or $\frac{0,8}{0,2}$ or $\frac{0,7}{0,3}$ etc.).

Calculated for the entire globe multipliers k_G and k_D we are going to treat as constant allowing to compare M_g indicators representing particular countries with indicator M_W representing the entire world.

$$M_{USA} = k_G \times G_{USA} + k_D \times D_{USA} \quad (15)$$

Then:

$$M_{USA} = 1,31 \times 0,0780 + 0,00086 \times 1619,1 = 1,495$$

Indicators for USA were calculated based on the data from [16], [17], [19], [21] and [24] indicators for India based on [12] and [13], for Brazil based on [7], [19], [21] and [22], for Poland based on [14], [15], [21], [22], and [23].

The M_{USA} index is greater than 1, which means that the American way of managing the pandemic gave a worse result than the chaotic, uncontrollable behavior of the entire globe. This is no surprise when you consider that it is the leader of the global rankings of infections. This position is favored by all circumstances, from the right to widely understand freedom of American society to the completely irresponsible behavior of the previous president. The result has been and is the infamous leadership of the world in terms of the scale of infections and deaths.

Tab. 1. Data and statistical results for several countries, a group of countries and the whole world on epidemic indicators and the quality of management in a given country

Countries	Population [millions]	Number of regions/states	Gini G indicator	Average number of deaths	Death std. dev.	Pandemic management indicator
World	7821.7	220	0.3803	580.9	917.8	1.0
USA	332.8	51	0.1008	1537.7	715.2	1.454
India	1392.5	36	0.2428	397.2	399.9	0.659
Brazil	211.7	27	0.0617	2321.8	568.2	2.077
Poland	37.8	16	0.0362	1946.5	231.3	1.721
Balkans	134.5	16	0.0880	2044.5	708.8	1.873

Another pole is the runner-up in this classification - India. It is a country demographically four times larger than the USA, however, after appropriate conversion, the index $M_{India} = 0,659$. We will comment on this in the Discussion. Other countries or a group of countries obtained the following M-index values:

$$M_{Brazil} = 2.077$$

$$M_{Poland} = 1.721;$$

$$M_{Balkans} = 1.873.$$

Discussion

The region of Eastern Europe with the Balkan states was selected because it is a rather hard-pressed place on Earth. Just look at the table [1]. In the course of this work, revolutionary changes were made in Peru. As a result of the verification of the number of deaths and the methods of their registration, the government published new data more than doubling the total number of deaths in the name of transparency and responsibility for the country [20]. In this way, Peru took the infamous first place in the world statistics with the number of 5,680 deaths per million inhabitants. Peru is followed by a

whole group of European countries considered here, starting with Hungary.

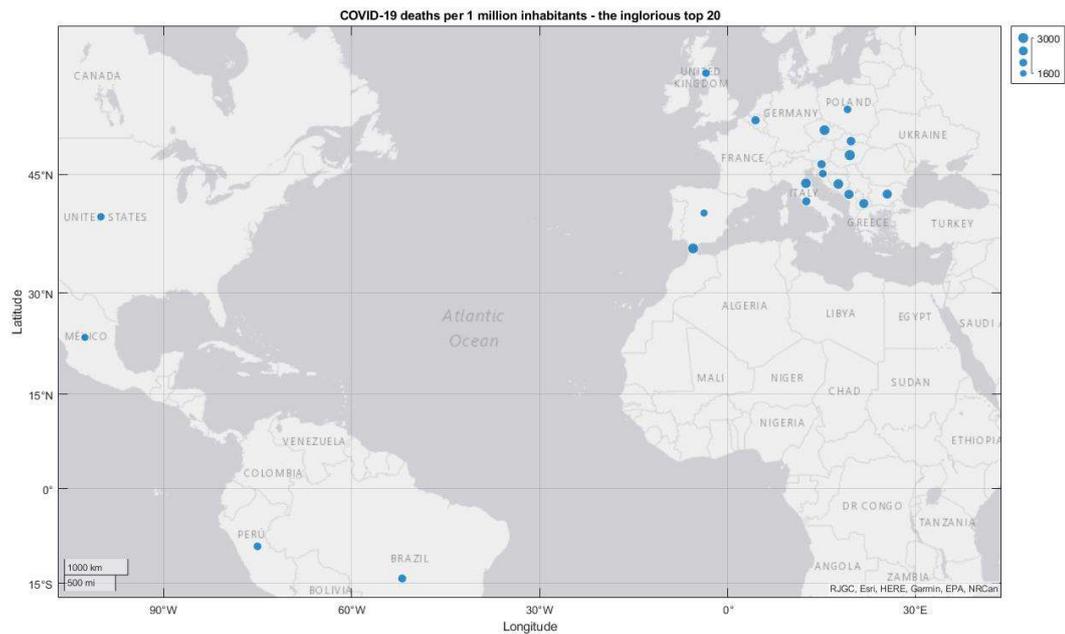


Fig. 3. An example of the concentration of countries with high deaths per million inhabitants among the Balkan and Eastern European countries.

This situation is presented in Figure 3, which shows the 20 countries from around the world who are the leaders in the number of cumulative deaths per million inhabitants. There are no Asian and African countries on this map. This map fully explains the choice of this particular territory for the calculation of the indicators.

The choice of the remaining countries is also obvious - they are the largest countries in terms of the number of infections according to which the world pandemic rankings are most recently drawn up, and Poland - an average country belonging to the endangered region of the world.

In India, during first wave of COVID-19 the cumulative death toll was around 158,396 till 1st March 2021. The second wave of COVID-19 was more severe and during this wave, death toll increased to 388,164 till 20th June 2021. It is not a big number as compared to the death toll of other countries of the world because of the large population of India [10]. The major factor which controls the spread of COVID-19 and deaths in India was one of the longest lockdowns, that was imposed from 24th March 2020 to 31st May 2020. During second wave of COVID-19 also, most of the states of India announced curfews, lockdown, several other restrictions, and encouraged vaccination. These facts are reflected in Table 1, as the average death in India is 397.2 and pandemic management indicator is 0.659, which is lowest among all considered countries in Table 1.

The situation is completely different in Brazil, which is rapidly approaching the position of a leader in the world, both in terms of the number of confirmed cases of infection and the number of deaths we are interested in. While in the most important category, the number of confirmed infections, in which the Johns Hopkins University presents the global COVID-19 statistics ranking, Brazil is in third place behind the US and India, it is second in the death category and is approaching the American leader. Therefore, the M index for Brazil, the highest in Table 1, is not accidental. The authors reserved the scope of this study to the statistical effects of the pandemic, not to its causes. In any country, however, any free citizen may try to become aware of what constitutes the bad or good

effects of a pandemic. The territorially uneven distribution of public health efforts is certainly one of the factors that can be assessed in terms of the fair and honest management of the state.

In the GitHub data repository there is a matlab script containing data for Poland and making it possible to calculate the Gini index.

Conclusion

According to the authors, the pandemic management quality index M reflects the important conditions determining pandemic management. This indicator can be adjusted to different than the authors' assessments of the importance of the two factors included in it (by changing their weights) - the factor resulting from the number of deaths in a given country and from the distribution of these deaths among regions. Of course, the high values of each of these factors are a bad sign of the authorities of a given country. However, there are a number of objectively existing causes (economic, cultural, regional autonomy, etc.) that we do not consider here. Comparing the values of M with tab. In the three main affected countries, some rather unexpected conclusions can be drawn. If in line with the comments already made in this paper, we consider that the deaths are an absolute and cruel consequence of a pandemic, then Brazil will be the country most affected worldwide. Against this background, many opinions expressed in the media are definitely positively influenced by India. They look better than the rest of the world on average. The United States, despite many months of leadership in both the number of confirmed cases of the disease and the number of deaths, is not as dramatically bad as it was a few months earlier. In the above-mentioned three world dominators, they are in the middle. By changing the weights for the factors under consideration, it is possible to reduce or increase the differences in this infamous classification, but the distances between these countries are, however, significant. Of course, it should be remembered that the pandemic is not over. Everything is subject to change.

CRedit author statement:

Antoni Wilinski: Conceptualization, Methodology, Writing, Software;

Arti MK.: Validation, Data Curation, Writing;

Lukasz Kupracz: Formal analysis, Software, Data Curation, Review&Editing

Acknowledgment:

The authors thank professor Rubens Junqueira Magalhães Afonso for valuable tips on the data on the distribution of deaths in Brazil and student Ryszard Karol Lisiecki for the preparation of Fig. 3.

References:

1. Bertsimas, D., Lukin, G., Mingardi, L., Nohadani, O., Orfanoudaki, A., Stellato, B., ... & Hellenic COVID-19 Study Group. (2020). COVID-19 mortality risk assessment: An international multi-center study. *PloS one*, 15(12), e0243262.
2. Barda, N., Riesel, D., Akriv, A., Levy, J., Finkel, U., Yona, G., ... & Dagan, N. (2020). Developing a COVID-19 mortality risk prediction model when individual-level data are not available. *Nature communications*, 11(1), 1-9.
3. Chowdhury, M. E., Rahman, T., Khandakar, A., Al-Madeed, S., Zughair, S. M., Doi, S. A., ... & Islam, M. T. (2021). An early warning tool for predicting mortality risk of COVID-19 patients using machine learning. *Cognitive Computation*, 1-16.
4. Gao, Y., Cai, G. Y., Fang, W., Li, H. Y., Wang, S. Y., Chen, L., ... & Gao, Q. L. (2020). Machine learning based early warning system enables accurate mortality risk prediction for COVID-19. *Nature communications*, 11(1), 1-10.
5. Oronce, C. I. A., Scannell, C. A., Kawachi, I., & Tsugawa, Y. (2020). Association between state-level income inequality and COVID-19 cases and mortality in the USA. *Journal of general internal medicine*, 35(9), 2791-2793.
6. Lindström, M. (2020). A commentary on "The trouble with trust: Time-series analysis of social capital, income inequality, and COVID-19 deaths in 84 countries". *Social Science & Medicine*, 263, 113386.
7. Raymundo, C. E., Oliveira, M. C., Eleuterio, T. D. A., André, S. R., da Silva, M. G., Queiroz, E. R. D. S., & Medronho, R. D. A. (2021). Spatial analysis of COVID-19 incidence and the sociodemographic context in Brazil. *Plos one*, 16(3), e0247794.
8. <https://www.economicshelp.org/blog/glossary/lorenz-curve/>
9. <https://worldpopulationreview.com/country-rankings/gini-coefficient-by-country>
10. www.worldometers.info
11. <https://gisanddata.maps.arcgis.com/apps/dashboards/>
12. <https://statisticstimes.com/demographics/india/indian-states-population.php>
13. <https://www.mygov.in/corona-data/covid19-statewise-status/>
14. https://pl.wikipedia.org/wiki/Podzia%C5%82_administracyjny_Polski
15. <https://konkret24.tvn24.pl/polska,108/zgony-z-covid-19-na-milion-mieszkanow-polska-piata-w-unii-europejskiej,1062030.html>
16. <https://usafacts.org/visualizations/coronavirus-covid-19-spread-map/>
17. <https://americatracker.com/u-s-population-by-states/>
18. <https://www.thebalancecareers.com/pareto-s-principle-the-80-20-rule-2275148>
19. <https://www.statista.com/statistics/1104709/coronavirus-deaths-worldwide-per-million-inhabitants/>
20. <https://cnnphilippines.com/world/2021/6/2/Peru-COVID-19-death-toll.html>
21. <https://www.statista.com/statistics/1107109/brazil-coronavirus-deaths-state/>
22. <https://gitlab.com/rubensjma/sigmoid-covid-19>
23. [Koronawirus w Polsce - aktualne dane, wykresy i mapa \(koronawirus.w.pl\)](https://koronawirus.w.pl)
24. [List of States By Population, Based on 2020 US Census \(1keydata.com\)](https://www.1keydata.com)