DATA-FORENSIC DETERMINATION OF THE ACCURACY OF INTERNATIONAL COVID-19 REPORTING: USING ZIPF'S LAW FOR PANDEMIC INVESTIGATION

Aamo Iorliam [*1], Anthony T S Ho[2,3,4] , Santosh Tirunagari [5], and David Windridge[2,5]
[1]Department of Mathematics & Computer Science, BSU, Makurdi, Nigeria.
[2]Faculty of Engineering and Physical Sciences, University of Surrey, Guildford, Surrey, United Kingdom GU2 7XH.
[3]Tianjin University of Science and Technology, China.
[4]Wuhan University of Technology, China.
[5] Department of Computer Science, Middlesex University, London, UK.
Corresponding author: Aamo Iorliam: aiorliam@bsum.edu.ng

Abstract

Severe outbreaks of infectious disease occur throughout the world with some reaching the level of international pandemic: Coronavirus (COVID-19) is the most recent to do so. As such pandemics cause extensive loss of lives, hamper industrial operations, and cause economic losses in both developing and developed countries, it is critical to establish common standards of accuracy in the determination and reporting of cases. In particular, there are current concerns that countries are hiding or incorrectly reporting cases of COVID-19. In this paper, we set out a mechanism for using Zipf's law to establish the accuracy of international reporting of COVID-19 cases via a determination of whether an individual country's COVID-19 reporting follows a power-law for confirmed, recovered, and death cases of COVID-19.

We observe that the probability of Zipf's law (P-values) for COVID-19 confirmed cases show that Uzbekistan has the highest P-value of 0.940, followed by Belize (0.929), and Qatar (0.897). For COVID-19 recovered cases, Iraq had the highest P-value of 0.901, followed by New Zealand (0.888), and Austria (0.884). Furthermore, for COVID-19 death cases, Bosnia and Herzegovina had the highest P-value of 0.874, followed by Lithuania (0.843), and Morocco (0.825). China, where the COVID-19 pandemic began, is a significant outlier in recording P-values lower than 0.1 for the confirmed, recovered, and death cases. This raises important questions, not only for China, but also any country whose data exhibits P-values below this threshold. The main application of this work is to serve as an early warning for World Health Organization (WHO) and other health regulatory bodies to perform more investigations in countries where COVID-19 datasets deviate significantly from Zipf's law. To this end, we also provide a tool for illustrating Zipf's law P-values on a global map in order to convey the geographic distribution of reporting anomalies.

Keywords Zipf's law, COVID-19, Pandemics

1. INTRODUCTION

1.1 Historical Perspective on COVID-19 Pandemic Investigations

The first recorded pandemic was in 165 AD to 180 AD. This pandemic was referred to as the Antonine Plague (also known as the plague of Galen) and resulted in about 5 million deaths across the globe. Analysis of symptomology and infection pattern suggest that this was likely smallpox or measles [1].

In around the 735 AD – 737AD, the Japanese smallpox epidemic erupted (believed to be a variola virus), killing up to 1 million persons [2]. Later, around 541 AD – 542 AD, the Plague of Justina killed between 30 – 50 million persons, believed to be the world's first bubonic plague [3]. Procopius described the plague as that "by which the whole human race was near to being annihilated [4 - 5]."

The most devastating pandemic, in terms of its impact of the global population, occurred between 1347 AD – 1352 AD; this is the pandemic referred to as 'The Black Death', which claimed between 75 – 200 million lives. It is believed to have been caused by the bubonic plague [6]; Benedictow in [7] described this plague as "the greatest catastrophe ever"; Michael of Piazza, a Franciscan friar wrote contemporaneously that: ''the infection spread to everyone who had any intercourse with the disease'' [8].

It is recorded that around 1520 AD there was an outbreak of the New world smallpox, believed to be a Variola virus, resulting in 25 to 55 million deaths. The New world smallpox caused so much damage that Noble David Cook [9-10] estimated that "in the end, the regions least affected lost 80 percent of their populations; those most affected lost their full populations, and a typical society lost 90 percent of its population."

Around 1629-1631 AD, the Italian Plague erupted, believed to originate from Yersinia Pestis bacteria in rats/fleas. It claimed up to 1 million lives [12-13].

Around 1665 AD, the great plague of London claimed 75,000 to 100,000 lives, also believed to have its source from rats and fleas [11].

From 1817 to 1923 the Cholera Pandemic (caused by V. Cholera bacteria) killed more than 1 million people [14] in Europe. Around 1885, a third plague caused by Yersinia Pestis bacteria carried by rats and fleas resulted in around 12 million deaths in China and India [15]. Also in the late 1800s Yellow fever, its source is believed to be viruses/mosquitoes, resulted in more than 150,000 deaths. It targeted mostly South America and sub-Saharan Africa [16-17].

The Russian Influenza outbreak of around 1889 to 1890, transmitted via the H2N2 virus, claimed around 1 million lives [18-19]. The Spanish flu of 1918 to 1919, its source believed to be H1N1 virus, claimed about 40 to 50 million lives [18-19], by far the most deadly influenza pandemic. Asian flu, believed to be transmitted by the H2N2 virus, claimed around 2 million lives in 1968 – 1970 [19], while the 1968-1970 Hong Kong Flu, transmitted via the H3N2 virus claimed up to 1 million lives [20].

The HIV/AIDS epidemic, which commenced around 1981 has so far claimed 25 to 30 million lives [21, 22]. From 2002 to 2003, Severe Acute Respiratory Syndrome (SARS) resulting from Coronavirus in Bats and Civets killed up to 770 people [23]. Later, between 2009 and 2010, Swine flu, its source believed to be the H1N1 virus in pigs, killed about 200,000 people [19].

Around 2012 – 2016, an outbreak of the Ebola virus disease (EVD) killed up to 11,000 persons; it is believed to have arisen from Ebola virus in wild animal. EVD was itself recorded as far back as 1976 [24-25].

doi:10.20944/preprints202004.0531.v1

From 2015 till the present time, the Middle East Respiratory Syndrome (MERS) believed to be caused by MERS corona-virus (MERS-CoV) in wild animals has so far killed 850 persons [26]. Coronavirus Disease 2019 (COVID-19) [27], the subject of the current study, erupted in 2019; according to [28], COVID-19 has claimed 146,201 lives as of 17/ 04/2020.

As is clear from the above historical account of pandemic spread, the potential for negative global impact is very substantial indeed if unchecked. In the majority of the above cases, the reporting and compilation of pandemic statistics was substantially after the fact (sometimes by many centuries) given the limited contemporaneous statistical capabilities. In the absence of such statistics, compiled while the outbreak was still live, it would have been very difficult or impossible for authorities to make well-informed policy decisions in order to combat the pandemic spread.

It is therefore critical in the current COVID-19 pandemic that accurate compilation of international reporting is undertaken. However, given the potential for countries/individuals to falsify records, for political, offensive or financial purposes, it is necessary to have methods in place to distinguish authentic from forged records. In this paper, we propose using Zipf's law as a means for achieving this.

1.2 Motivation for the use of Zipf's law

Zipf's law was proposed in 1935 by the US linguist George K Zipf [29] and may be stated succinctly as follows: *given some corpus of natural language utterances, the frequency of any given word is inversely proportional to its rank in a frequency table*.

Newman [30] made this explicitly stochastic; when considering the probability of measuring a particular quantity (in our case, COVID-19 cases), and it is found that the quantity varies inversely as a power of that value, then the quantity may be said to follow Zipf's law [30]. Mathematically:

$$p(x) = cx^{-\alpha} \ldots\ldots. (1)$$

doi:10.20944/preprints202004.0531.v1

where $p(x)$is the distribution of the quantity x, α is the Zipf's law exponent and C is a constant [30].

Zipf's law is hence a Power law with small occurrences relatively common and large occurrences very uncommon. The null hypothesis in applying Zipf's law is hence that natural phenomena should follow a Power law and un-natural (or tampered-with) phenomena should deviate from this law [31-32]. Power laws have been applied to human language [30,32], the city populations [30,32], intensity of earthquakes [30,32], sizes of power outages [32], ranks of people watching a particular TV station at a given time [32, 33], stock market indices [33], gene expression [33], chess openings [33], the arts [33], paper citations [33], family names [33], personal donations [33], keystroke dynamics [32], the scales of Influenza A (H1N1) and Avian Influenza (H7N9) outbreaks [33, 34] amongst others.

In this paper, we propose the investigation of *reported* COVID-19 datasets using Zipf's law to establish veracity and accuracy, in particular because of presence of widespread allegations that countries may have hidden or systematically underreported the cases of COVID-19 [35].

We thus aim to establish probability values (P-values) in relation to Zipf's law calculation for each country affected by COVID-19. Furthermore, we represent the P-values of each country based on the Zipf's law calculation on a global map. This is hence an ongoing work as more data is compiled throughout the current COVID-19 outbreak.

2. EXPERIMENTS

Our primary goal is to investigate internationally reported cases of COVID-19 in order to determine consistency with Zipf's law. A secondary goal is to calculate the P-value for Zipf's law on each country's COVID-19 datasets. Lastly, we illustrate Zipf's law P-values on a global map to convey the geographic distribution of reporting anomalies.

We use the Power-law package developed by Clauset et. al. [36] to obtain P-values for reported cases of COVID-19 per country. As methodologically reported in [36] and evaluated in [32], we carry out experiments 1000 times on the COVID-19 datasets in order to obtain P-values in each case. The steps followed to test whether COVID-19 datasets follow a Power-law are set out in [32] and [36]. It should be noted that the P-values are generated using the Kolmogorov-Smirnov (KS) statistic goodness-of-fit test as specified in [32] and [36]. We used the standard COVID-19 datasets available at https://github.com/CSSEGISandData/COVID-19. The data consists in rows representing countries and columns representing the number of COVID-19 cases per each day from 1/22/2020 to 4/6/2020.

The evaluated hypotheses are thus:

1. Is Zipf's law applicable to COVID-19 case data?
2. How is Zipf's law behavior characterized if applicable?

2.2 Experimental Results

Tables 1 (Appendix I) show P-values for all the countries/regions using COVID-19 confirmed cases based on the Power law calculation. We further show the distribution of P-values across countries/regions for COVID-19 confirmed cases in Figure 1. These P-values range from 0 to 1, with any country having a value of less than 0.1 deemed to deviate from Zipf's law.

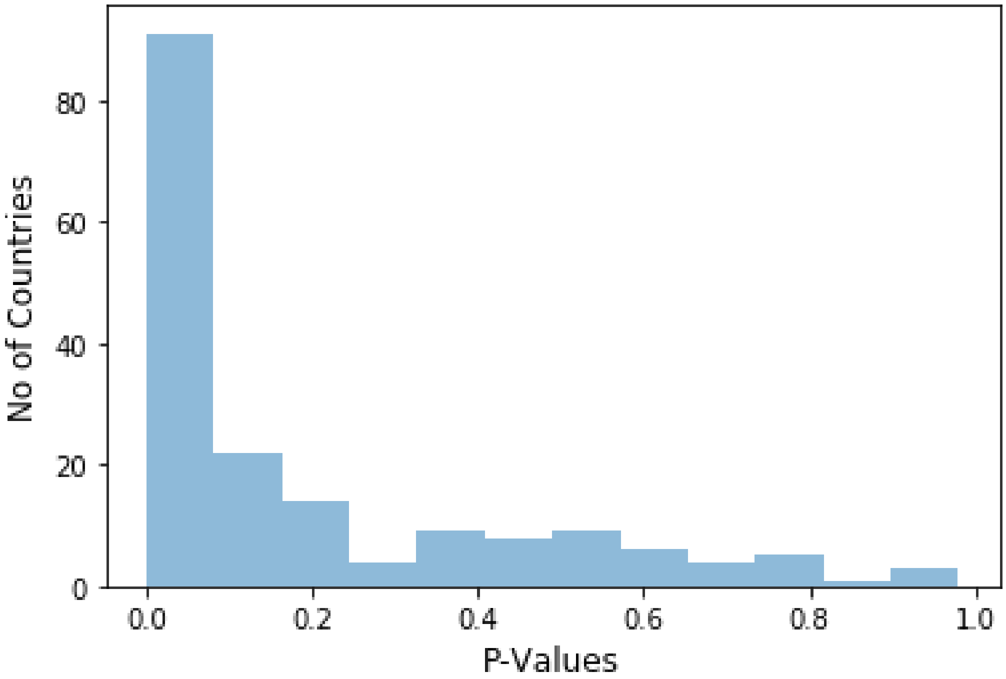


Figure 1: Distribution of P-values across Countries/Regions for Confirmed Cases

Table 2 extracts four countries'/regions' P-values from Table 1 to illustrate score dichotomy: Uzbekistan, Belize, Qatar have high P-values close to 1 while China has a P-value of 0 (to within the measured accuracy of the test). Countries/regions with the higher P-values indicate COVID-19 datasets in very close compliance with Zipf's law; China's reported statistics, along with those scoring similar P-values such as Australia, US, etc. are not possible to reconcile with Zipf's law.

Table 2: P-values of Four Countries/Regions for Confirmed Cases

Country	P-value
Uzbekistan	0.940
Belize	0.929
Qatar	0.897
China	0.000

Table 3 (Appendix II) further shows the P-values on all recovered cases of COVID-19. Due to the number of recovered cases being small in some countries, Zipf's law is not statistically robust in such cases; Table 3 therefore does not indicate these results.

We illustrate the distribution of P-values across countries/regions for COVID-19 recovered cases in Figure 2.

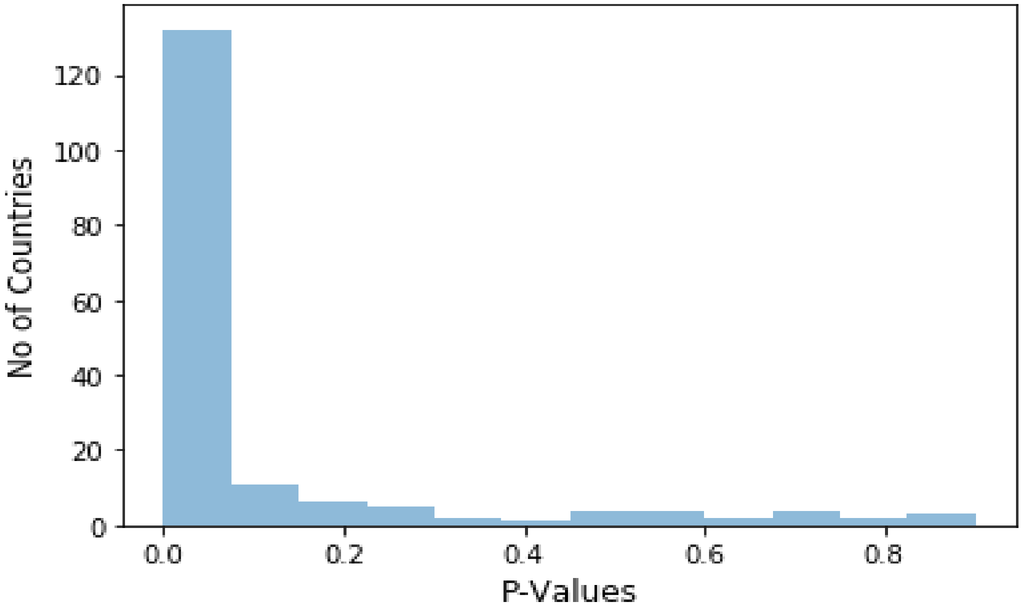


Figure 2: Distribution of P-values across Countries/Regions for Recovered Cases

Again, Table 4 extracts four extremal countries'/regions' P-values from Table 3: Iraq, New Zealand, Austria and China (with China having an extremely low P-value of 0.002). It should however be noted that other countries such as Ghana, US, etc. have P-values of 0 which are lesser than China when considering P-values for recovered cases.

Table 4: P-values of Four Countries/Regions for Recovered Cases

Country	P-value
Iraq	0.901
New Zealand	0.888
Austria	0.884
China	0.002

Table 5 (Appendix III) shows the P-values of all death cases of COVID-19 across each country. In countries in which there are only a few recorded cases of COVID-19 deaths power-law P-values are not presented.

We indicate the P-value distribution across countries/regions for COVID-19 death cases in Figure 3.

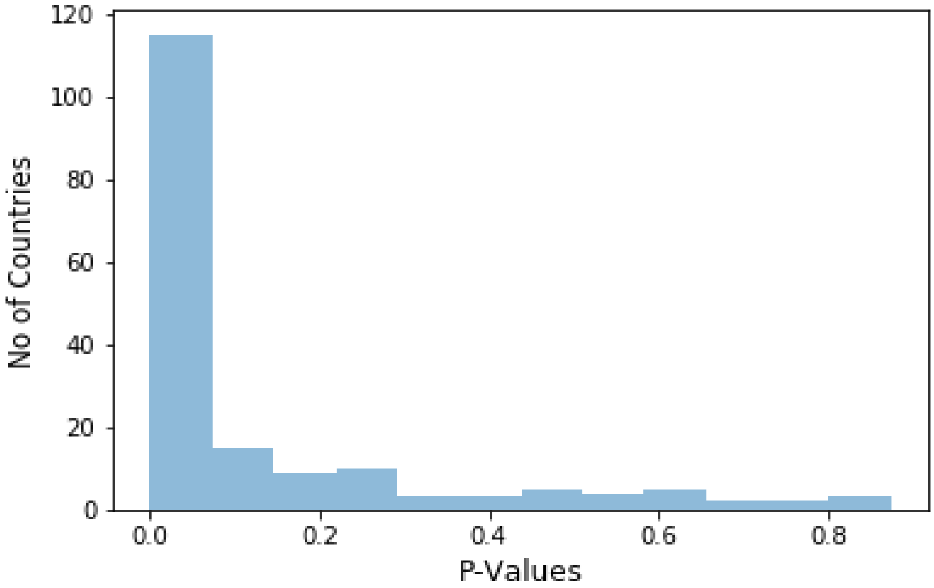


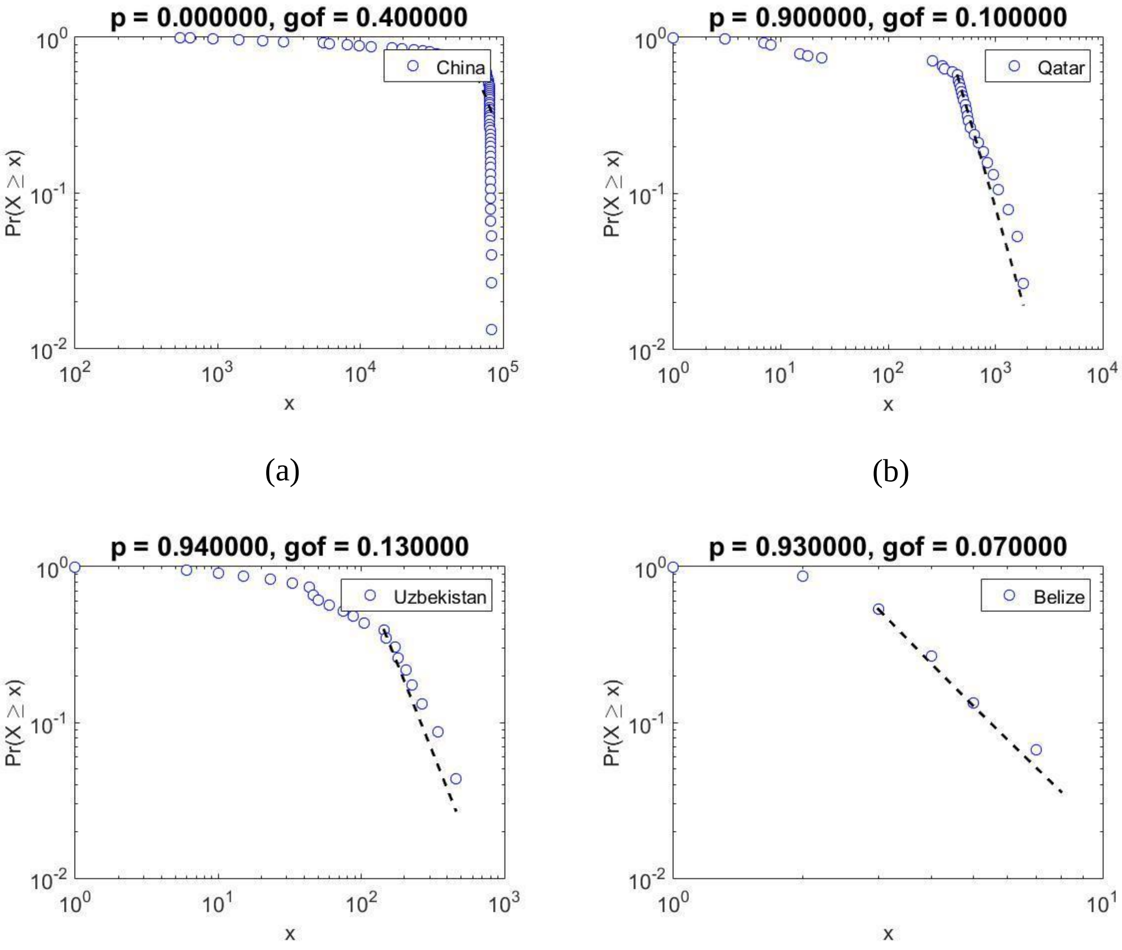
Figure 3: Distribution of P-values across Countries/Regions for Death Cases

Table 6 extracts four extremal P-values from Table 5; those for Bosnia and Herzegovina, Lithuania, Morocco and China; Bosnia and Herzegovina, Lithuania, Morocco have high P-values while China has a P-value of 0.000. Countries such as Costa Rica, Jordan, etc. have P-values of 0 similar to that of China.

Table 6: P-values of Four Countries/Regions for Death Cases

Country	P-value
Bosnia and Herzegovina	0.874
Lithuania	0.843
Morocco	0.825
China	0.000

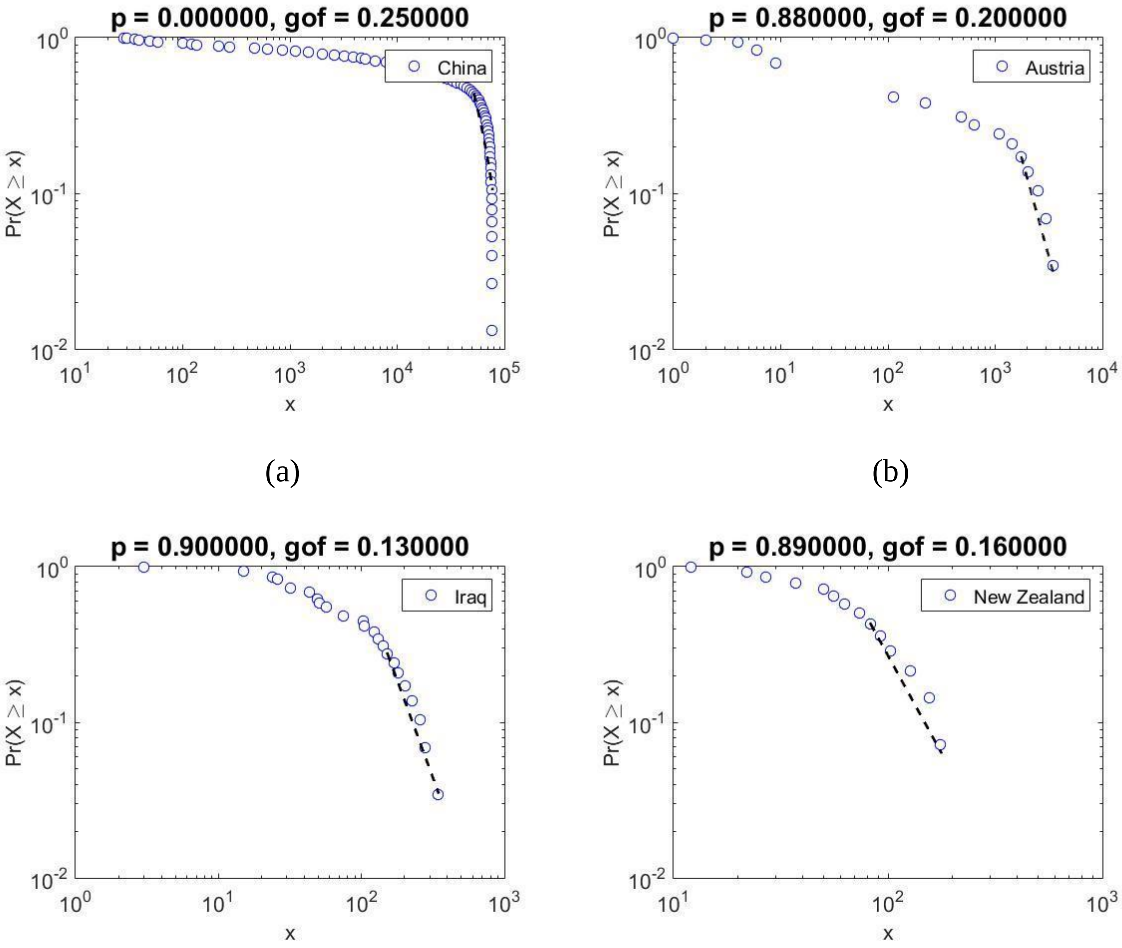
Zipf's law power-law graph fits for each of the four countries identified in Tables 2, 4, and 6 are shown in Figures 4, 5, and 6 for COVID-19 confirmed cases, recovered cases, and death cases, respectively.



(a) (b)

(c) (d)

Figure 4: Zipf's law power-law fits of COVID-19 Confirmed Cases for: (a) China, (b) Qatar, (c) Uzbekistan, (d) Belize

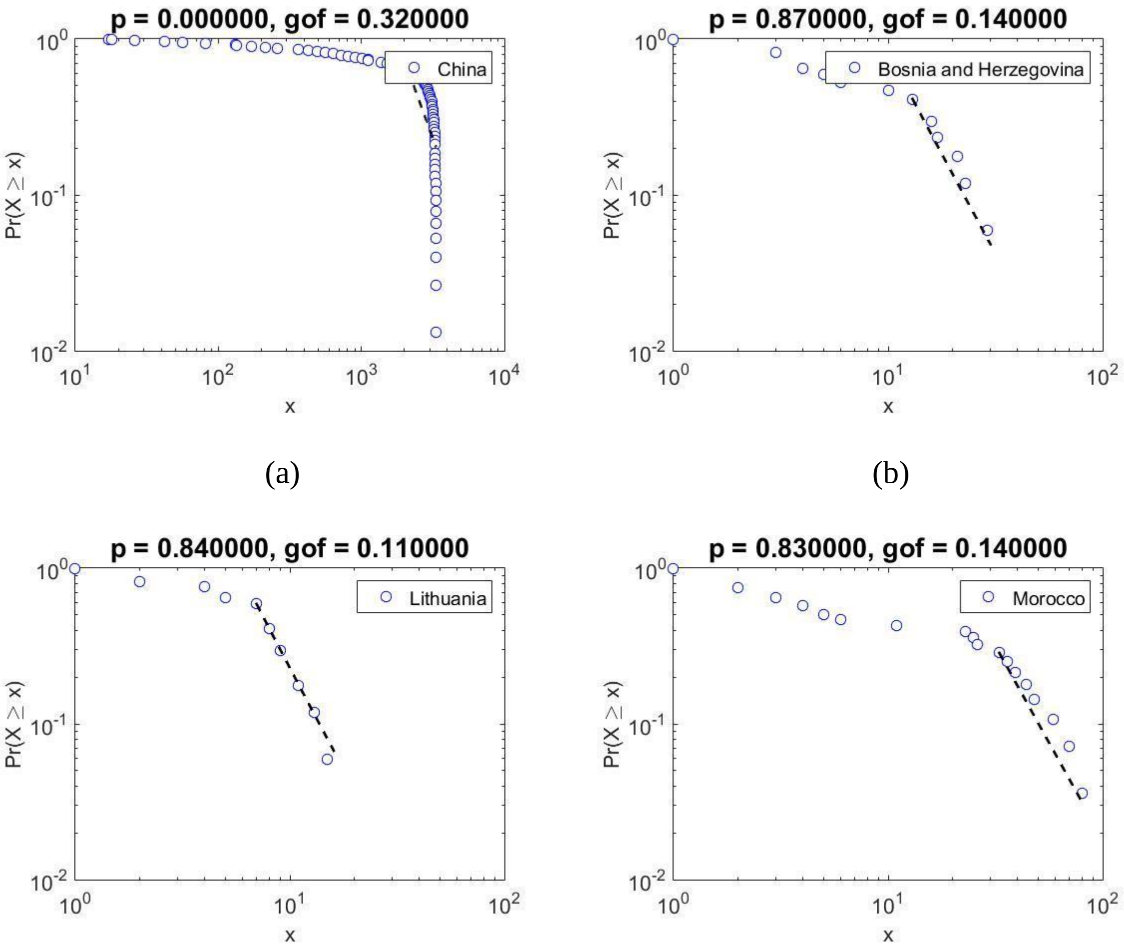


(a) (b)

(c) (d)

Figure 5: Zipf's law power-law fits of COVID-19 recovered Cases for: (a) China, (b) Austria, (c) Iraq, (d) New Zealand

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 30 April 2020 doi:10.20944/preprints202004.0531.v1



(a) (b)

(c) (d)

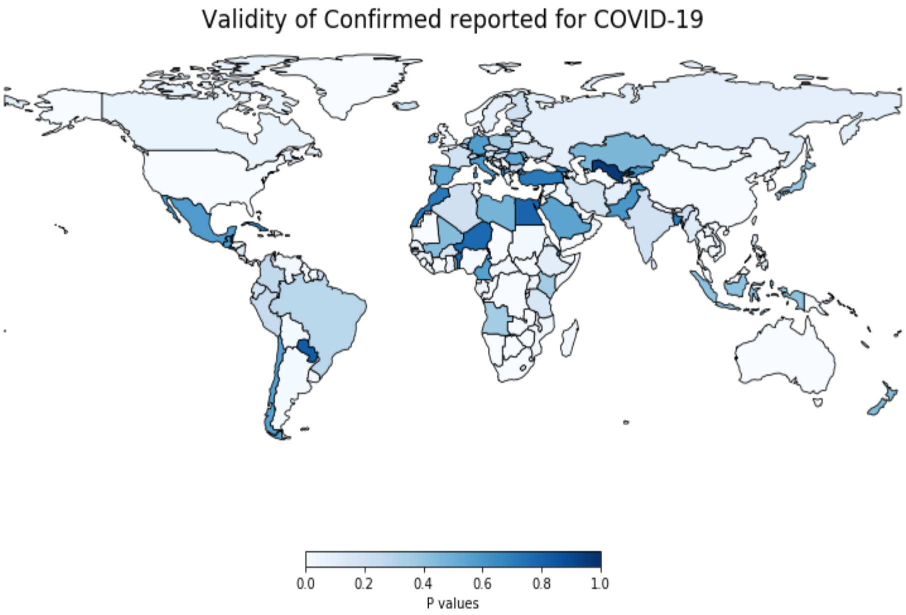
Figure 6: Zipf's law power-law fits of COVID-19 death Cases for (a) China, (b) Bosnia and Herzegovina, (c) Lithuania, (d) Morocco

Figures 7a, b, and c indicate, on a global map, P-values per country for the COVID-19 confirmed cases, recovered cases, and death cases respectively. These maps are interactive and can also be viewed online for the COVID-19 confirmed cases[1], recovered cases[2], and death cases[3] showing the P-values and country names when hovered on it.

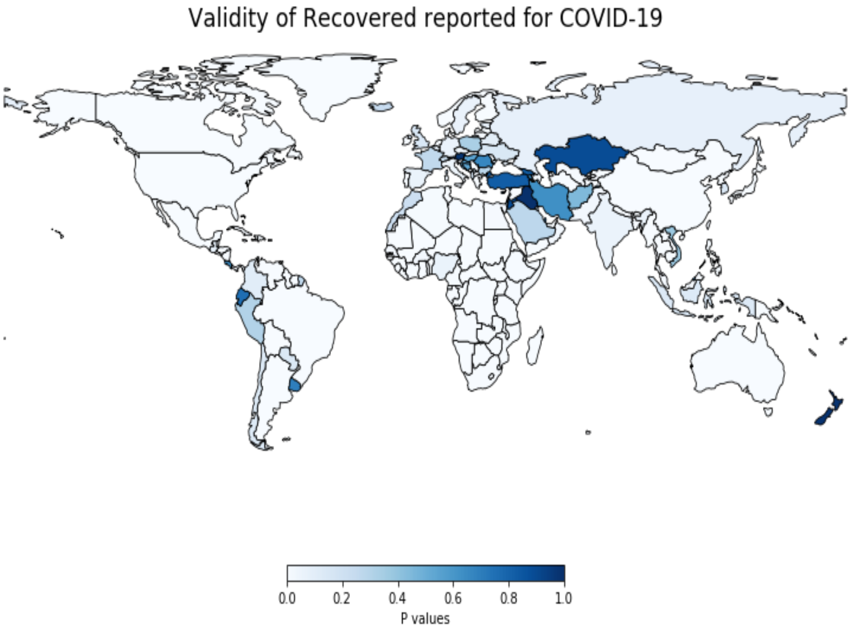
[1] Confirmed Cases: https://covido.volitionlabs.xyz/zipfs/confirmed

[2] Recovered Cases: https://covido.volitionlabs.xyz/zipfs/recovered

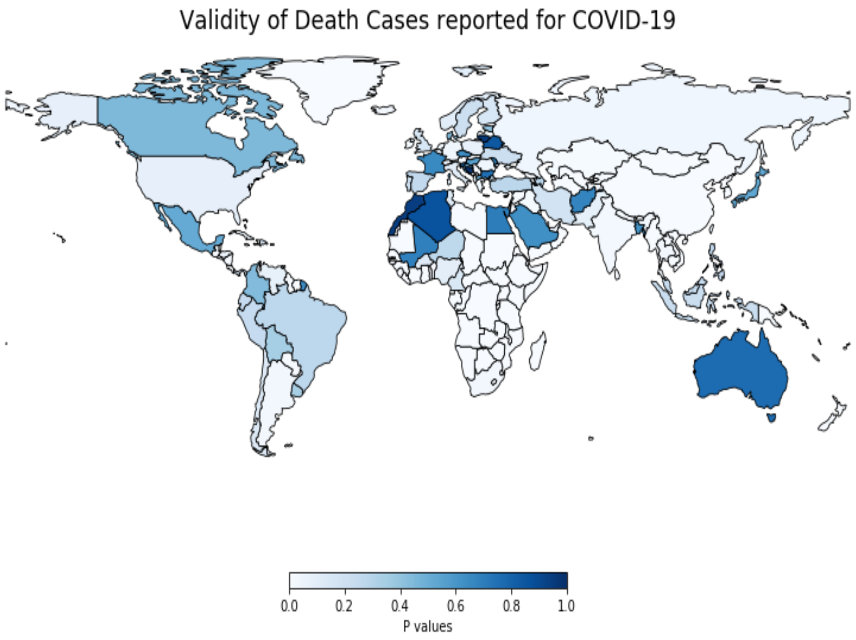
[3] Death Cases: https://covido.volitionlabs.xyz/zipfs/death


Validity of Confirmed reported for COVID-19
0.0
0.2
0.4
0.6
0.8
1.0
P values

(a)


Validity of Recovered reported for COVID-19
0.0
0.2
0.4
0.6
0.8
1.0
P values

(b)



(c)

Figure 7: Maps of: (a) COVID-19 confirmed cases, (b) COVID-19 recovered cases, (c) COVID-19 death cases

3.0 DISCUSSION

Experiments indicate that many countries' COVID-19 datasets follow a consistent power law for confirmed cases of COVID-19, recovered cases of COVID-19, and death cases of COVID-19. (Table 2 indicates that Uzbekistan had the highest P-value of 0.94, followed by Belize with a P-value of 0.929, and Qatar with a P-Value of 0.897; Table 4 indicates that Iraq's recovered cases data most closely follows Zipf's law with a P-value of 0.901, followed by New Zealand with a P-value of 0.888, and Austria with a P-value of 0.884; Table 6 indicates that Bosnia and Herzegovina had the highest P-value of 0.874, followed by Lithuania with a P-value of 0.843, and Morocco with a P-value of 0.825).

As can be seen in Figures 4, 5, 6, the CDF represented as Pr (x) is plotted as a function of frequency (x) for the COVID-19 dataset for confirmed, recovered, and death cases. The straight line (black line) shows the section of the plots where the Power law provided a good model fit [36] of the data considered.

From Figures 1 to 6 and Tables 1 to 6, however, we observe that several countries have P-values that are less than 0.1. Notably, China's confirmed cases, recovered cases, and death cases of COVID-19 all failed to follow Zipf's law despite an earlier commencement of data recording than those of other countries (being the pandemic origin).

This has raised some questions, not only for China, but also every other country whose power-law P-values are less than 0.1 (this threshold being the one selected to establish compliance with Zipf's law according to the reasoning in [32, 36]).

Based on the above discussion, we can conclude that:

1. Zipf's law can be applied to COVID-19 case data with reliability monotonically improving in relation to dataset size.
2. This analysis can potentially be used as an 'early warning system' for further investigation into COVID-19 datasets not consistent with Zipf's law.

4.0 CONCLUSIONS AND FUTURE WORK

In this paper, we have established that COVID-19 datasets for many countries can be shown to be consistent with Zipf's law. However, experiments also indicate that deviation of COVID-19 datasets from Zipf's law may be indicative of incorrect data reporting. The main application of this work is thus to serve as a potential early warning system for international health regulatory bodies such as the World Health Organization (WHO) in performing further investigations in countries where COVID-19 datasets have deviated from Zipf's law.

In future work, we plan to:

1. Re-perform this experiment with complete statistics once the COVID-19 pandemic has subsided.

2. Carry out experiments with other power laws variants (in particular Benford's law and Heap's law) both for COVID-19 and other pandemic datasets.

REFERENCES

[1] Duncan-Jones, R. P. (1996). The impact of the Antonine plague. *Journal of Roman Archaeology*, *9*, 108-136.

[2] Geddes, A. M. (2006). The history of smallpox. *Clinics in dermatology*, *24*(3), 152-157.

[3] Snowden, F. M. (2019). *Epidemics and Society: From the Black Death to the Present*. Yale University Press.

[4] Bray, R. S. (2004). *Armies of pestilence: the impact of disease on history*. James Clarke & Co..

[5] Findlay, R., & Lundahl, M. (2017). Demographic shocks and the factor proportions model: from the plague of Justinian to the Black Death. In *The Economics of the Frontier* (pp. 125-172). Palgrave Macmillan, London.

[6] Horrox, R. (2013). The black death.

[7] Benedictow, O. J. (2005). The Black Death: the greatest catastrophe ever. *History Today*, *55*(3), 42-49.

[8] Duncan, C. J., & Scott, S. (2005). What caused the black death?. *Postgraduate medical journal*, *81*(955), 315-320.

[9] Cook, N. D. (1998). *Born to die: disease and New World conquest, 1492-1650* (Vol. 1). Cambridge University Press.

[10] Nunn, N., & Qian, N. (2010). The Columbian exchange: A history of disease, food, and ideas. *Journal of Economic Perspectives*, *24*(2), 163-88.

[11] Roberts, R. S. (1966). The place of plague in English history. *Proceedings of the Royal Society of Medicine*, *59*(2), 101.

[12] Vicentini, C. B., Manfredini, S., Mares, D., Bonacci, T., Scapoli, C., Chicca, M., & Pezzi, M. (2020). Empirical "integrated disease management" in Ferrara during the Italian plague (1629–1631). *Parasitology International*, *75*, 102046.

[13] Alfani, G., & Murphy, T. E. (2017). Plague and lethal epidemics in the pre-industrial world. *the Journal of economic History*, *77*(1), 314-343.

[14] Barua, D. (1992). History of cholera. In *Cholera* (pp. 1-36). Springer, Boston, MA.

[15] Kausrud, K. L., Begon, M., Ari, T. B., Viljugrein, H., Esper, J., Büntgen, U., ... & Xu, L. (2010). Modeling the epidemiological history of plague in Central Asia: palaeoclimatic forcing on a disease system over the past millennium. *Bmc Biology*, *8*(1), 112.

[16] Delaporte, F. (1991). *History of yellow fever: an essay on the birth of tropical medicine*. MIT Press.

[17] Haddow, A. J. (1969). X.—The Natural History of Yellow Fever in Africa. *Proceedings of the Royal Society of Edinburgh, Section B: Biological Sciences*, *70*(3), 191-227.

[18] Valleron, A. J., Meurisse, S., & Boelle, P. Y. (2008). Historical Analysis of the 1889–1890 Pandemic in Europe. *International Journal of Infectious Diseases*, *12*, e95.

[19] Choi, A., & García-Sastre, A. (2020). Influenza forensics. In *Microbial Forensics* (pp. 89-104). Academic Press.

[20] Laver, W. G., & Webster, R. G. (1972). Studies on the origin of pandemic influenza: II. Peptide maps of the light and heavy polypeptide chains from the hemagglutinin subunits of A2 influenza viruses isolated before and after the appearance of Hong Kong influenza. *Virology*, *48*(2), 445-455.

[21] Ramogale, M. R., & Moodley, J. (2007). HIV-associated maternal mortality–primary causes of death at King Edward VIII Hospital, Durban. *South African Medical Journal*, *97*(5), 363-366.

[22] Grant, A. D., Djomand, G., & De, K. C. (1997). Natural history and spectrum of disease in adults with HIV/AIDS in Africa. *AIDS (London, England)*, *11*, S43-54.

[23] Hon, C. C., Lam, T. Y., Shi, Z. L., Drummond, A. J., Yip, C. W., Zeng, F., ... & Leung, F. C. C. (2008). Evidence of the recombinant origin of a bat severe acute respiratory syndrome (SARS)-like coronavirus and its implications on the direct ancestor of SARS coronavirus. *Journal of Virology*, *82*(4), 1819-1826.

[24] Weyer, J., Grobbelaar, A., & Blumberg, L. (2015). Ebola virus disease: history, epidemiology and outbreaks. *Current infectious disease reports*, *17*(5), 21.

[25] Pourrut, X., Kumulungui, B., Wittmann, T., Moussavou, G., Délicat, A., Yaba, P., ... & Leroy, E. M. (2005). The natural history of Ebola virus in Africa. *Microbes and infection*, *7*(7-8), 1005-1014.

[26] Ajlan, A. M., Ahyad, R. A., Jamjoom, L. G., Alharthy, A., & Madani, T. A. (2014). Middle East respiratory syndrome coronavirus (MERS-CoV) infection: chest CT findings. *American Journal of Roentgenology*, *203*(4), 782-787.

[27] Zu, Z. Y., Jiang, M. D., Xu, P. P., Chen, W., Ni, Q. Q., Lu, G. M., & Zhang, L. J. (2020). Coronavirus disease 2019 (COVID-19): a perspective from China. *Radiology*, 200490.

[28] Novel Coronavirus (COVID-19) Cases Data: https://data.humdata.org/dataset/novel-coronavirus-2019-ncov-cases, Retrieved: 17/4/2020.

[29] Manning, C. D., Manning, C. D., & Schütze, H. (1999). *Foundations of statistical natural language processing*. MIT press.

[30] Newman, M. E. (2005). Power laws, Pareto distributions and Zipf's law. *Contemporary physics*, *46*(5), 323-351.

[31] Kalankesh, L. R., Stevens, R., & Brass, A. (2012). The language of gene ontology: a zipf's law analysis. *BMC bioinformatics*, *13*(1), 127.

[32] Iorliam, A., Ho, A. T., Poh, N., Tirunagari, S., & Bours, P. (2015, March). Data forensic techniques using Benford's law and Zipf's law for keystroke dynamics. In *3rd International Workshop on Biometrics and Forensics (IWBF 2015)* (pp. 1-6). IEEE.

[33] Wang, L., Li, X., Zhang, Y. Q., Zhang, Y., & Zhang, K. (2011). Evolution of scaling emergence in large-scale spatial epidemic spreading. *PloS one*, *6*(7).

[34] Pan, Y. N., Lou, J. J., & Han, X. P. (2014). Outbreak patterns of the novel avian influenza (H7N9). *Physica A: Statistical Mechanics and its Applications*, *401*, 265-270.

[35] Lachmann, A. (2020). Correcting under-reported COVID-19 case numbers. *medRxiv*.

[36] Clauset, A., Shalizi, C. R., & Newman, M. E. (2009). Power-law distributions in empirical data. *SIAM review*, *51*(4), 661-703.

doi:10.20944/preprints202004.0531.v1

APPENDIX I: Table 1. P-Values for COVID-19 Confirmed Cases

Country	**P-values**	**Country**	**P-values**
Afghanistan	0.094	Cuba	0.652
Albania	0.724	Cyprus	0.161
Algeria	0.206	Czechia	0.190
Andorra	0.000	Denmark	0.096
Angola	0.352	Djibouti	0.543
Antigua and Barbuda	0.000	Dominican Republic	0.011
Argentina	0.011	Ecuador	0.175
Armenia	0.354	Egypt	0.778
Australia	0.000	El Salvador	0.804
Austria	0.160	Equatorial Guinea	0.000
Azerbaijan	0.037	Eritrea	0.108
Bahamas	0.036	Estonia	0.267
Bahrain	0.237	Eswatini	0.000
Bangladesh	0.686	Ethiopia	0.107
Barbados	0.020	Fiji	0.083
Belarus	0.044	Finland	0.198
Belgium	0.377	France	0.199
Benin	0.769	Gabon	0.047
Bhutan	0.002	Gambia	0.051
Bolivia	0.019	Georgia	0.648
Bosnia and Herzegovina	0.058	Germany	0.561
Brazil	0.285	Ghana	0.049
Brunei	0.000	Greece	0.009
Bulgaria	0.214	Guatemala	0.571
Burkina Faso	0.141	Guinea	0.076

Table 1. P-Values for COVID-19 Confirmed Cases

Country	P-values	Country	P-values
Cabo Verde	0.000	Guyana	0.009
Cambodia	0.000	Haiti	0.050
Cameroon	0.552	Holy See	0.000
Canada	0.060	Honduras	0.018
Central African Republic	0.028	Hungary	0.174
Chad	0.002	Iceland	0.187
Chile	0.554	India	0.202
China	0.000	Indonesia	0.404
Colombia	0.264	Iran	0.177
Congo (Brazzaville)	0.000	Iraq	0.017
Congo (Kinshasa)	0.139	Ireland	0.450
Costa Rica	0.094	Israel	0.041
Cote d'Ivoire	0.583	Italy	0.585
Croatia	0.203	Jamaica	0.189
Diamond Princess	0.000	Japan	0.366
Jordan	0.001	Portugal	0.240
Kazakhstan	0.449	Qatar	0.897
Kenya	0.355	Romania	0.536
Korea, South	0.000	Russia	0.092
Kuwait	0.980	Rwanda	0.096
Kyrgyzstan	0.579	Saint Lucia	0.003
Latvia	0.001	San Marino	0.000
Lebanon	0.000	Saudi Arabia	0.532
Liberia	0.038	Senegal	0.070
Liechtenstein	0.003	Serbia	0.575
Lithuania	0.281	Seychelles	0.026

doi:10.20944/preprints202004.0531.v1

Table 1. P-Values for COVID-19 Confirmed Cases

Country	P-values	Country	P-values
Luxembourg	0.128	Singapore	0.138
Madagascar	0.021	Slovakia	0.471
Malaysia	0.000	Slovenia	0.009
Maldives	0.000	Somalia	0.006
Malta	0.235	South Africa	0.001
Mauritania	0.000	Spain	0.522
Mauritius	0.535	Sri Lanka	0.153
Mexico	0.572	Sudan	0.021
Moldova	0.486	Suriname	0.008
Monaco	0.272	Sweden	0.081
Mongolia	0.009	Switzerland	0.149
Montenegro	0.465	Taiwan*	0.031
Morocco	0.707	Tanzania	0.157
Namibia	0.000	Thailand	0.000
Nepal	0.005	Togo	0.390
Netherlands	0.352	Trinidad and Tobago	0.145
New Zealand	0.430	Tunisia	0.008
Nicaragua	0.118	Turkey	0.702
Niger	0.769	Uganda	0.045
Nigeria	0.014	Ukraine	0.158
North Macedonia	0.495	United Arab Emirates	0.075
Norway	0.003	United Kingdom	0.026
Oman	0.040	Uruguay	0.023
Pakistan	0.587	US	0.000
Panama	0.024	Uzbekistan	0.940

Table 1. P-Values for COVID-19 Confirmed Cases

Country	P-values	Country	P-values
Peru	0.244	Vietnam	0.000
Philippines	0.100	Zambia	0.000
Poland	0.349	Zimbabwe	0.000
Dominica	0.001	West Bank and Gaza	0.171
Grenada	0.114	Guinea-Bissau	0.142
Mozambique	0.030	Mali	0.428
Syria	0.019	Saint Kitts and Nevis	0.065
Belize	0.929	Kosovo	0.140
Laos	0.056	Burma	0.112
Libya	0.467	MS Zaandam	0.000

APPENDIX II: Table 3. P-Values for COVID-19 Recovered Cases

Country	P-values	Country	P-values
Afghanistan	0.406	Chile	0.116
Albania	0.145	China	0.002
Algeria	0.010	Colombia	0.101
Andorra	0.002	Costa Rica	0.584
Argentina	0.001	Cote d'Ivoire	0.220
Armenia	0.581	Croatia	0.503
Australia	0.002	Diamond Princess	0.000
Austria	0.884	Cuba	0.007
Azerbaijan	0.010	Cyprus	0.554
Bahrain	0.001	Czechia	0.041
Bangladesh	0.005	Denmark	0.021
Belarus	0.095	Dominican Republic	0.026
Belgium	0.072	Ecuador	0.683
Bosnia and Herzegovina	0.486	Egypt	0.012
Brazil	0.000	Estonia	0.062
Brunei	0.498	Ethiopia	0.004
Bulgaria	0.254	Finland	0.000
Burkina Faso	0.019	France	0.241
Cambodia	0.000	Georgia	0.738
Cameroon	0.009	Germany	0.060
Canada	0.001	Ghana	0.000

doi:10.20944/preprints202004.0531.v1

Table 3. P-Values for COVID-19 Recovered Cases

Country	P-values	Country	P-values
Hungary	0.494	Malta	0.015
Iceland	0.215	Mexico	0.000
India	0.053	Moldova	0.016
Indonesia	0.117	Morocco	0.195
Iran	0.569	Netherlands	0.000
Iraq	0.901	New Zealand	0.888
Ireland	0.000	Nigeria	0.057
Israel	0.249	North Macedonia	0.012
Italy	0.037	Norway	0.001
Jamaica	0.077	Oman	0.021
Japan	0.001	Pakistan	0.028
Jordan	0.804	Panama	0.030
Kazakhstan	0.803	Paraguay	0.127
Kenya	0.005	Peru	0.283
Korea, South	0.124	Philippines	0.008
Kuwait	0.175	Poland	0.303
Lebanon	0.174	Portugal	0.000
Lithuania	0.000	Qatar	0.676
Luxembourg	0.000	Romania	0.601
Malaysia	0.012	Russia	0.077
Maldives	0.000	San Marino	0.157

Table 3. P-Values for COVID-19 Recovered Cases

Country	**P-values**	**Country**	**P-values**
Senegal	0.019	Turkey	0.745
Singapore	0.211	Ukraine	0.137
Slovakia	0.000	United Arab Emirates	0.001
Slovenia	0.000	United Kingdom	0.133
South Africa	0.000	Uruguay	0.631
Spain	0.057	US	0.000
Sri Lanka	0.000	Uzbekistan	0.014
Sweden	0.063	Venezuela	0.000
Switzerland	0.191	Vietnam	0.343
Taiwan*	0.017	West Bank and Gaza	0.006
Thailand	0.001	Kosovo	0.174
Togo	0.001		

APPENDIX III: Table 5. P-Values for COVID-19 Death Cases

Country	**P-values**	**Country**	**P-values**
Afghanistan	0.592	China	0.000
Albania	0.027	Colombia	0.389
Algeria	0.761	Congo (Kinshasa)	0.100
Andorra	0.376	Costa Rica	0.000
Argentina	0.024	Croatia	0.458
Armenia	0.005	Diamond Princess	0.029
Australia	0.671	Cuba	0.035
Austria	0.012	Cyprus	0.092
Azerbaijan	0.243	Czechia	0.561
Bahrain	0.000	Denmark	0.394
Bangladesh	0.544	Dominican Republic	0.153
Belarus	0.740	Ecuador	0.220
Belgium	0.078	Egypt	0.617
Bolivia	0.305	Estonia	0.356
Bosnia and Herzegovina	0.874	Finland	0.137
Brazil	0.243	France	0.571
Bulgaria	0.653	Germany	0.035
Burkina Faso	0.108	Ghana	0.004
Cameroon	0.132	Greece	0.032
Canada	0.394	Honduras	0.041
Chile	0.112	Hungary	0.489

Table 5. P-Values for COVID-19 Death Cases

Country	P-values	Country	P-values
Indonesia	0.200	North Macedonia	0.449
Iran	0.164	Norway	0.098
Iraq	0.017	Pakistan	0.077
Ireland	0.007	Panama	0.204
Israel	0.138	Paraguay	0.005
Italy	0.094	Peru	0.220
Japan	0.471	Philippines	0.231
Jordan	0.000	Poland	0.063
Kazakhstan	0.009	Portugal	0.192
Korea, South	0.002	Qatar	0.129
Lebanon	0.150	Romania	0.032
Lithuania	0.843	Russia	0.031
Luxembourg	0.221	San Marino	0.020
Malaysia	0.213	Saudi Arabia	0.550
Mauritius	0.036	Serbia	0.318
Mexico	0.454	Singapore	0.378
Moldova	0.611	Slovenia	0.671
Morocco	0.825	South Africa	0.024
Netherlands	0.012	Spain	0.228
Niger	0.239	Sri Lanka	0.006
Nigeria	0.090	Sweden	0.186

Table 5. P-Values for COVID-19 Death Cases

Country	**P-values**	**Country**	**P-values**
Thailand	0.028	United Kingdom	0.138
Trinidad and Tobago	0.240	Uruguay	0.301
Tunisia	0.104	US	0.075
Turkey	0.246	Venezuela	0.071
Ukraine	0.193	Mali	0.606
United Arab Emirates	0.042		