Proceedings of the 2nd International Conference on Business and Policy Studies
DOI: 10.54254/2754-1169/11/20230561

Research on the Impact of COVID-19 in 225 Countries Based on SQL

Qianqian Yan[1,a,*]

[1]*Rutgers The State University Of NJ New Brunswick, NJ 08901, USA*
a, qy110@rutgers.edu
**corresponding author*

Abstract: The year 2022 marks the third year of humanity's fight against the novel coronavirus. With the deepening understanding of the novel coronavirus, the influencing factors of the number of infections and deaths have gradually been paid more attention by various countries. Based on the "COVID-19 Coronavirus Pandemic" for national populations; SQL was used to analyze the objective impact of COVID-19 on each country and the relationship between objective factors and mortality in different countries through the mortality rate. Studies have found that medical care, as well as population, are important factors in stopping the spread of the virus, The lack of medical care and supplies calls for better policies to prevent The spread of The virus; Countries with Large Populations should pay more attention to prevention methods.

Keywords: COVID-19, SQL

1. Introduction

It is necessary to set the question before the data calculation so that the calculation and analysis can be carried out with a target. The objective of this study is to obtain the main causes of virus transmission and suggestions and methods for preventing virus transmission based on data analysis. The number of deaths caused by a pandemic virus and the mortality rate is not negligible, because they are important factors to judge whether a pandemic virus has an impact on humans. More specifically, the number of deaths and the mortality rate are decisive factors, both in terms of the psychological impact on people and the policy decisions of governments. In The article The Emergence of COVID-19 as a cause of death in 2020 and its effect on mortality by disease of the respiratory System in Spain: Trends and Their Superiority Compared to 2019 "In - In 2020, COVID-19 produced a large increase (68.5%) in deaths by diseases of the respiratory system compared to the previous Year. " [1]. Clearly, in the case of Spain, the coronavirus has posed a great threat to human life. It's not hard to infer that COVID-19 deaths have occurred all over the world, not just in Spain. From this, we can set up The following three questions: The first question is The Number of Deaths per country and The five countries with The highest number of Deaths. The mortality rate and the total number of infections provided in the sample provide an estimate of the number of deaths due to COVID-19 in each country. The countries with the highest mortality rates provide a preliminary analysis of the factors contributing to the high mortality caused by COVID-19. By comparing the number of deaths in each country, we can analyze whether there is a link between the population of each country and the number of deaths caused by COVID-19, and provide advice on epidemic

DOI: 10.54254/2754-1169/11/20230561

prevention policies for countries affected by demographic factors. The second question is set to What is the infection rate in each country? The infection rate of COVID-19 in each country can be obtained based on the total number of infected persons and the total population of the country provided in the sample. The comparison of infection rates can provide analytical help for the impact of the development degree of the country on the transmission of COVID-19. The third question is set as the countries the death rate is higher than 5%. According to the sample, it is not difficult to see that only a few countries have a death rate higher than 5%. The purpose of this question is to screen out countries with a death rate higher than 5% and find out the similarities among these countries. To analyze the factors affecting the mortality of COVID-19.

The purpose of this study is to analyze the impact of COVID-19 on human health and life and to analyze ways to reduce the negative impact of COVID-19 based on the main factors of epidemic transmission, the population base, and the mortality and infection rates in each country. Advise countries with high rates of infection and death on policy and prevention. According to SQL calculation, the population base does have a certain impact on the number of infected people, but the decisive factors leading to the high mortality and infection rate are national policies and medical and economic conditions.

2. Data and Methods

2.1. Data Sample

The data of this study comes from Kaggle, a well-known data website. The data cover 225 countries and include variables such as population, total infections, total deaths, and mortality rates. The data is sorted by the first letter of a country's name, by summing multiple data types together in a single icon. Such data makes the differences between each country obvious, which greatly facilitates the study of data analysis. The data, obtained from Worldometer, aims to calculate the impact of the novel coronavirus on humans. By comparing the data from different countries, we can summarize better anti-virus methods and policies, find out the causes of the spread of the virus, and ways to reduce the spread of the virus. Therefore, this data can be regarded as a representative sample of global epidemic data and analyzed and summarized on the basis of the impact of the novel coronavirus on human beings reflected in the data.

Since the table contains duplicate data and data not required for this study, data screening is necessary. The original table included country; other names; ISO 3166-1 alpha-3 CODE; population; continent; Total cases; Total Deaths; Tot cases/1 M pop; Tot Deaths/ 1 M pop; Death percentage. After screening out duplicated data and unnecessary data in this study, the data range can be locked in the country. population; Total Cases; Total Deaths and Death Percentage. In this study, new data will be obtained by publicity on the basis of this data, and on this basis, data analysis will be conducted and the problem of setting in advance will be solved.

2.2. Research Methods

This study decided to use SQL for data analysis and calculation. Structured Query Language, as a data processing system, is good at processing large data. It can easily import large data and organize and filter it. Firstly, the samples used in this research are imported and the unnecessary data are screened out. The world of commercial computing has mostly failed to supply suitable work to justify the creation of commercial parallel systems. The advent of the relational Model, with its parallel language-SQL and the need to hold large amounts of data, has given this opportunity [8]. In addition, SQL uses a unified language, which makes it easier for users to manage databases. Import the data groups needed for this research into SQL and call them COVID19.

DOI: 10.54254/2754-1169/11/20230561

3. Research Results

By analyzing the different impacts of COVID-19 on individual countries, the study aims to determine the best way to prevent the spread of the virus. In order to solve the three problems raised, six models are compared in this study. The first three are the basic models, and three formulas are used to preliminarily sort out the whole data. On this basis, this study will carry out further data calculation and analysis and solve the problems raised.

3.1. Research Hypothesis

Perform 'From COVID19 Order by Population DESC Limit 5 'to get Table1. Sort the populations of 225 countries from largest to smallest and limit the number to 5 in the variable Population. Five of the most populous country of China, India, the USA, Indonesia and Pakistan.

Figure 1: Most populous countries in COVID-19

In Figure2, execute "From COVID19 Order by Total Cases DESC Limit5" and limit the number to 5 in the variable "Total Cases", The top five countries with the highest number of infections are the USA, India, Brazil, France, and Germany. Comparing Table1 and Table2, it is clear that India and USA appear in the top three countries with the largest population and the countries with the largest total number of infections.

The COVID-19 attack rate increases with city size and, in the absence of adequate controls, Larger Cities (and counties, as we assume) are expected to have more essential epidemics than smaller ones [2]. In the United States, for example, Los Angeles, Illinois, and Texas had the highest rates. And these are the three most densely populated cities in the United States. In other words, the more people are exposed to it, the more likely they are to become infected. From this, it can be tentatively assumed that countries with large populations have accelerated COVID-19 transmission. However, as the above two tables are the basis of data collation, no conclusion conducive to solving the problem can be analyzed, so further analysis and comparison should be carried out on this basis.

Figure 2: Top five countries with the highest total number of COVID-19 infections.

DOI: 10.54254/2754-1169/11/20230561

SELECT country ,round(sum(`Tot

	country	death
1	USA	1,008,222
2	Brazil	660,269
3	India	521,388
4	Russia	369,708
5	Mexico	323,212
6	Peru	212,328
7	UK	165,570
8	Italy	159,784
9	Indonesia	155,288
10	France	142,506

Figure 3: COVID-19 deaths by country.

For further research, it's possible to analyze the impact of COVID-19 on a country by counting the number of deaths. Run round (sum(Total cases * Death-percentage)/100,2) as death from COVID-19 group by country order by death DESC limit 10 ", the number of deaths in each country can be obtained by multiplying the variable Total cases and Death percentage and then sorting the number of deaths from the largest to the smallest to get the top ten countries with the largest number of deaths. It was found that among the five countries with the largest population, India, USA, and Indonesia all appeared in the top ten in the number of deaths. The two data are highly coincident. It can be compared between China and Pakistan and these countries to try to find out the optimal epidemic prevention policies of countries with a large population base. It can be tentatively assumed that countries with larger population bases will have higher death tolls.

3.2. Proof of Research Hypothesis

In Figure4, perform the "From COVID19 Order by Death Percentage DESC limit 5" with the variable "Death Percentage" limiting the number to 5 to obtain the five countries with the highest Death rates From COVID-19. Comparing table1 with table4, there is no duplication of data between the five countries with the highest populations in COVID-19 and the five countries with the highest death rates from COVID-19, suggesting that there is no direct impact on population and death rates. Therefore, it is necessary to re-analyze the reasons and give suggestions.

	country	Death percentage
1	Yemen	18.152
2	Western Sahara	10
3	Sudan	7.92
4	Peru	5.983
5	Mexico	5.705

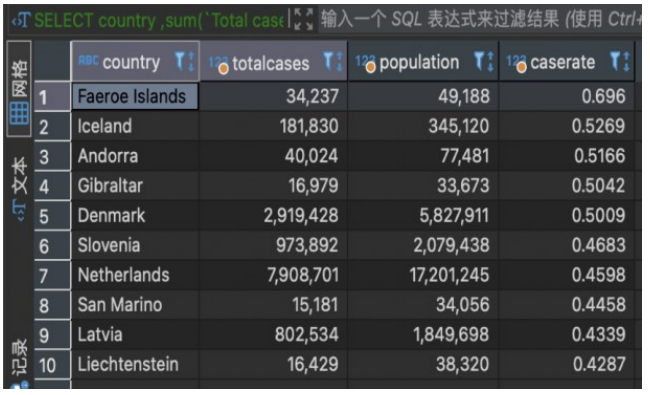
Figure 4: The five countries with the highest death rates from COVID-19.

Because of the influence of the total population of the country, only analyzing the data of the total number of infected people cannot control for a single variable, so it is one-sided and invalid. By

DOI: 10.54254/2754-1169/11/20230561

calculating the infection rate, you can really figure out which country is most affected by the novel coronavirus, and you can also analyze the impact of population size on the transmission of the virus.

Run the formula "sum(Total cases) as total cases, sum(population) as population, Sum (Total cases)/ SUM (population) as case rate from COVID-19 19 Group by Country Order by case rate DESC limit 10 ". The infection rate can be obtained by dividing the total number of infected people by the total population and then sorting the variable "case rate" to screen out the ten countries with the highest infection rate. It is clear that countries with high populations are not co-existing with countries with high infection rates. This is not consistent with the prediction and the conclusion of the hypothesis, and it is necessary to re-analyze the reasons and give suggestions.


SELECT country ,sum(`Total case 输入一个 SQL 表达式来过滤结果 (使用 Ctrl+

	country	totalcases	population	caserate
1	Faeroe Islands	34,237	49,188	0.696
2	Iceland	181,830	345,120	0.5269
3	Andorra	40,024	77,481	0.5166
4	Gibraltar	16,979	33,673	0.5042
5	Denmark	2,919,428	5,827,911	0.5009
6	Slovenia	973,892	2,079,438	0.4683
7	Netherlands	7,908,701	17,201,245	0.4598
8	San Marino	15,181	34,056	0.4458
9	Latvia	802,534	1,849,698	0.4339
10	Liechtenstein	16,429	38,320	0.4287

Figure 5: Infection rate

4. Discussion

This study found that there are many factors that influence the rate of infection and death from COVID-19, and these factors differ from country to country and even from person to person. As a result, COVID-19 has affected every country to varying degrees. According to the common infectious factors of an epidemic, population density is often an important factor affecting the infectious rate of an epidemic. With people living in denser conditions, Interactions between individuals and disease transmission tend to occur more easily [2]. Obviously, interactions between individuals and disease transmission tend to occur more easily. The impact of COVID-19 will be more severe in cities than in rural areas, where there will be more population density and more interaction between people. It follows that population density is closely related to infection rates.

In order to study the most important factors affecting the spread of COVID-19, the following two preliminary assumptions can be made based on the information provided by the data model: 1. Countries with larger population bases have higher death tolls. 2. Countries with large populations have accelerated COVID-19 transmission

Unexpectedly, the results of this study did not match expectations. Population size was not the most fundamental factor affecting mortality and infection rate. In order to verify hypothesis 1, need five countries with the most population and mortality rates are the highest, comparing the five countries in the most representative data to observe the relationship between the two groups of variables, therefore, calculating the population ranked table1 model and the calculation of the mortality of table3 model together, find two groups of the model is not repeated. More specifically, the top five countries by population do not have very high mortality rates, while the top five countries by highest mortality rates do not have very large populations. Therefore, hypothesis 1 can be overturned and it can be concluded that population size is not the fundamental factor causing high mortality. In order to verify hypothesis 2, it is necessary to compare the five countries with the largest population with several countries with high infection rates. In order to reduce the error and expand

DOI: 10.54254/2754-1169/11/20230561

the research scope, the limit of the infection rate model is expanded to 10. In other words, table5 will calculate the 10 countries with the highest infection rates, and divide the total number of infected people column by the population column to get a new series "infection rate". Using infection rates to make comparisons will result in more intuitive, more effective, and less error-prone research data. The ten countries with the highest infection rates all have small populations or even small populations. In other words, Hypothesis 2 is not true, and population size does not directly affect the COVID-19 infection rate.

Population, one of the most common infectious factors in an epidemic, has been shown by this data to have little to do with the rate of spread of COVID-19, so the data need to be reanalyzed to find out why. Looking closely at table4 and 5, the five countries with the highest death rates are Yemen, Western Sahara, Sudan, Peru, and Mexico, and the five countries with the highest infection rates are Faeroe Islands, Iceland, Andorra, Gibraltar, and Denmark. All ten countries have small populations, in some cases five figures, and none of them are developed. In other words, the real reason for the high rate of infection and death is economic strength and medical quality. In the case of low economic strength, masks, vaccines, and other epidemic prevention and medical supplies cannot be provided in time, and the number of medical equipment and medical staff cannot match the number of infected people. If such medical resources are not used, the infection rate will be greatly expanded. There is no medicine to treat the virus directly, but studies have shown that quarantine supplies such as masks can inhibit the spread of the virus. Numerical Evaluation of face Masks for Prevention of COVID-19 Airborne transmission points out that: "We find that wearing a face mask can effectively reduce the total mass and Sauter mean diameter of the residual Droplets after a single cough." [3] Clearly, countries lacking such medical supplies will increase the chances of COVID-19 transmission, as the virus will spread without any obstacles.

At the same time, the role of vaccines in curbing the spread of COVID-19 cannot be ignored, as countries with low COVID-19 vaccine coverage often have higher infection rates than countries with high coverage. Countries using mRNA-based vaccines have been the most effective in reducing the number of COVID-19 cases and deaths [4]. The COVID-19 vaccine has reduced the fatality rate and transmission rate of COVID-19 in a targeted manner, which has greatly enhanced people's confidence in fighting the virus. Countries that fail to provide COVID-19 vaccines in a timely manner will experience higher mortality and transmission rates. On the other hand, infected patients will die due to poor medical care, and lack of doctors and treatment resources. The mortality rate will increase greatly if the country has poor medical care.

Of course, one cannot, therefore, avoid the impact of population based on the epidemic. Going back to Tables 1 and 2, two of the top five countries by population are in the top five countries by total infections. Even if the data model for the total number of infections is not as good as the data model for the infection rate, the data still says something. First look at the countries that appear in both data models, namely USA and India. It is clear that these two countries have a policy role to play in promoting COVID-19. First, India's COVID-19 preparedness and management systems are inadequate. while the government was swift and decisive in its response, The pandemic exposed a weakness in policy coordination and inadequacies (C) in India's social protection System [5]. As a populous country, India has failed to effectively reduce people's interaction and maintain social safety distance and has not paid attention to the impact of population density on the infection rate [6]. Second, after the peak of the first wave of infections eased, the government let down its guard before the second wave hit [7,8]. This left India's health care system ill-prepared when the second wave of the virus hit, with insufficient facilities and beds to deal with infected patients [9]. Increased the risk of death threefold, regardless of stable disorders due to the surge in the number of infections, many people in India are not being provided with adequate medical resources and the death toll has risen dramatically. As a result, many believe the reported death toll in India is far below the true toll. At

DOI: 10.54254/2754-1169/11/20230561

the same time, the open-ended approach to pandemic control in the United States has had a similar effect. Under such policies, people's movements are not restricted, and some states do not even require masks. Clearly, such a policy is liberal, safeguards travel requirements, and sustain the economy. However, it is clear that such policies have drawbacks. Alarming numbers of infections and deaths have occurred in both countries, and it is clear that the population size of the US and India has played a role in accelerating the spread of COVID-19 under the influence of the policy. The policy of openness has led to continuous infecting of people, and the policy has led to continuous contact between those who are already infected and those who are not. Especially in countries with large populations, quarantine policies are necessary to protect those who are not infected and to protect them from repeated infection. As the pandemic continues, an appropriate duration of quarantine can help countries with low transmission rates to keep infections at a low level [6]. COVID-19 has an incubation period and onset period, and these two periods are very easy to transmit the virus. People should be isolated during infection, prevent interaction with others, maintain social distancing, and other important ways to prevent the spread of the virus.

On the other hand, China, as a populous country, has decisively chosen a series of mandatory measures, such as mandatory quarantines, mandatory masks, and travel restrictions, amid the rampant spread of COVID-19. People's travel has been severely affected, and tourism, transportation, and a number of other economies collapsed during this period. The sudden onslaught of COVID-19 has taken a heavy toll on China's macroeconomy and its market participants [10]. However, there is a positive side to such a policy. In the case of an epidemic, it is clear that such a policy has been effective, and China has not had a too noticeable infection and death rates. Overall, a quarantine policy is crucial for a populous country, The WHO and the Centers for Disease Control and Prevention (CDC) strongly recommended avoiding close and direct contact with people with acute respiratory infections, refraining from traveling to high-risk areas [7]. Each country can choose more appropriate epidemic prevention policies according to its own national conditions to reduce the number of infections and deaths.

5. Conclusion

This paper uses SQL to collate and analyze data from 225 countries, to study the impact of COVID-19 on each country, in order to find out the direct factors affecting the infection rate and mortality rate and to provide policy research advice for each country. The main conclusions are as follows: 1. Population size is not the most fundamental factor leading to high infection and mortality rates. 2. Health facilities and national policies are the most direct factors affecting infection and mortality rates. This study provides advice for countries with high mortality and infection rates to reduce the number of infections and deaths and slow the spread of COVID-19. However, because this study is based on only one data group, the conclusions obtained are often one-sided and limited, and the conclusions are obtained without considering other factors such as national conditions and citizens' will. I hope to draw experience from future research and do further and more comprehensive research.

References

[1] *Soriano, J. B. , et al. "The emergence of COVID-19 as a cause of death in 2020 and its effect on mortality by diseases of the respiratory system in Spain: Trends and their determinants compared to 2019." (2022).*

[2] *Lima, F. , N. Brown , and J. Duarte . "Understanding the Impact of Walkability, Population Density, and Population Size on COVID-19 Spread: A Pilot Study of the Early Contagion in the United States. " Entropy (Basel, Switzerland) 23.11(2021).*

[3] *Liu, Jiaxing , et al. "Numerical evaluation of face masks for prevention of COVID-19 airborne transmission." Environmental Science and Pollution Research 29.29(2022):44939-44953.*

[4] *Bukhari, Q. , Y. Jameel , and S. Khan . "Real-World Effectiveness of COVID-19 Vaccines: the Diverging Pattern of COVID-19 Cases and Deaths in Countries with High Vaccination Rates." Social Science Electronic Publishing.*

DOI: 10.54254/2754-1169/11/20230561

[5] *Taneja, P., and A. Bali. "India's Domestic and Foreign Policy Responses to Covid-19: Semantic Scholar." Undefined, 1 Jan. 1970, https://www.semanticscholar.org/paper/India%E2%80%99s-domestic-and-foreign-policy-responses-to-Taneja-Bali/41d953b599352d86449f45d50b758103a9be07ed*

[6] *Tu, H. , et al. "Effectiveness of 14 day quarantine strategy: Chinese experience of prevention and control." BMJ 375(2021):e12-20.*

[7] *M. Irani, "Review on the symptoms, transmission, therapeutics options and control the spread of the disease of COVID-19," Alborz University Medical Journal, vol. 9, no. 2, pp. 171–180, 2020.*
View at: Publisher Site | Google Scholar

[8] *Page, J. . "The benefits of database computers." Design and Application of Parallel Digital Processors, 1991. Second International Specialist Seminar on the IET, 2002.*

[9] *B, Adi Turjeman A , et al. "Assessing the impact of COVID-19 on mortality: A population-based matched case-control study." Clinical Microbiology and Infection (2022).*

[10] *Liu, X. , Y. Liu , and Y. Yan . "China macroeconomic report 2020: China's macroeconomy is on the rebound under the impact of COVID-19." Studies in Economics and Politics: English Edition 8.4(2020):41.*