

# A Panel Regression Analysis for the COVID-19 Epidemic in the United States: From the Quarantine Policy and Environmental Perspective

**Yinpei Guo**

Jilin University School of Public Health

**Bo Li**

Jilin University School of Public Health

**Tonghua Duan**

Jilin University School of Mathematics

**Nan Yao**

Jilin University School of Public Health

**Han Wang**

Jilin University School of Public Health

**Yixue Yang**

Jilin University School of Public Health

**Shoumeng Yan**

Jilin University School of Public Health

**Mengzi Sun**

Jilin University School of Public Health

**Ling Wang**

Jilin University School of Public Health

**Yan Yao**

Jilin University School of Public Health

**Yuchen Sun**

JHU Whiting School of Engineering: Johns Hopkins University Whiting School of Engineering

**Jiwei Jia**

Jilin University School of Mathematics, Jilin National Applied Mathematical Center

**Siyu Liu** (✉ [liusiyu@jlu.edu.cn](mailto:liusiyu@jlu.edu.cn))

Jilin University <https://orcid.org/0000-0002-0966-4589>

---

**Research Article**

**Keywords:** COVID-19, panel data, quarantine policy, environmental factors

**Posted Date:** June 8th, 2021

**DOI:** <https://doi.org/10.21203/rs.3.rs-578957/v1>

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

[Read Full License](#)

---

# **A panel regression analysis for the COVID-19 epidemic in the United States: from the quarantine policy and environmental perspective**

Yinpei Guo<sup>1</sup>, Bo Li<sup>1</sup>, Tonghua Duan<sup>2</sup>, Nan Yao<sup>1</sup>, Han Wang<sup>1</sup>, Yixue Yang<sup>1</sup>, Shoumeng Yan<sup>1</sup>, Mengzi Sun<sup>1</sup>, Ling Wang<sup>1</sup>, Yan Yao<sup>1</sup>, Yuchen Sun<sup>3</sup>, Jiwei Jia<sup>2, 4\*</sup>, Siyu Liu<sup>1\*</sup>

<sup>1</sup> Department of Epidemiology and Biostatistics, School of Public Health, Jilin University

<sup>2</sup> Department of Computational Mathematics, School of Mathematics, Jilin University

<sup>3</sup> Department of Applied Mathematics & Statistics, Whiting School of Engineering, Johns Hopkins University

<sup>4</sup> Jilin National Applied Mathematical Center, Jilin University

## **\*Corresponding authors:**

**Siyu Liu:** Department of Epidemiology and Biostatistics, School of Public Health, Jilin University, 1163 Xinmin Avenue, Changchun, 130021, P. R. China.

**Jiwei Jia:** Department of Computational Mathematics, School of Mathematics; Jilin National Applied Mathematical Center, Jilin University, Qianjin street 2699, Changchun, 130012, P. R. China

## **Authors' Email Addresses:**

**Siyu Liu:** Tel: +86 15143080289      E-mail: liusiyu@jlu.edu.cn

**Jiwei Jia:** Tel: +86 13596197575      E-mail: jjajiwei@jlu.edu.cn

## **Abstract**

**Background.** The COVID-19 epidemic has had an extreme impact on society. This study aimed to discuss this epidemic in the U.S. and explore the association between COVID-19 daily incidence rate and influencing factors including people's implementation of states' quarantine policy and environmental factors including temperature, humidity and so on.

**Methods.** Data of 50 states in U.S. were used as the research subjects. A panel data model was established based on the daily incidence rate and influencing factors from 15 March to 30 September, 2020. The period was analyzed both unsegmented and segmented. The k-means clustering method was used to cluster the states, and panel linear regression method was used for correlation analysis.

**Results.** The characteristics of the daily incidence rate and factors of the three categories were different after clustering. The daily residents at home, proportion of travel people, humidity and incidence rate were negatively correlated, while the daily temperature and incidence rate were positively correlated after unsegmented multivariate analysis. While after segmented analysis, the air pressure and the temperature showed a trend that was negatively correlated with the daily incidence rate respectively in the first and the fifth segment, other indicators showed the analogous results. At the same time, this study also completed the regression analysis after classification of the three groups. Compared with results without classification, there was a decrease of the number of significant independent variables.

**Conclusions.** The spread of COVID-19 in 50 states in U.S. was related to quarantine

measures, temperature and humidity. The progress of the epidemic would be relatively slow if people chose to stay at home. Besides, the increase in temperature ( $<84.2^{\circ}\text{F}$ ) could be conducive to the spread of the epidemic, while the increase in relative humidity (40~70%) might inhibit the spread of the virus to a certain degree.

## **Keywords**

COVID-19, panel data, quarantine policy, environmental factors

## Background

The rapid spread of COVID-19 has seriously affected people's health and daily life and imposed a great burden on almost every country and society <sup>[1-2]</sup>.

On 20 January, 2020, the Centers for Disease Control and Prevention of United States confirmed the first male case of infection <sup>[3]</sup>. On 31 January, the first case of human-to-human transmission was diagnosed in the U.S. <sup>[4]</sup>. On 5 March, 100 cases were diagnosed. On 18 March, the number of cases surged to 10,000. The outbreak was inevitable. Until 12 April, 50 states had issued a "Disaster Declaration". New York was the worst-hit area in the U.S. As a result of the limited implementation of domestic policies from the government and the intense discussions agued by American on "wearing masks" and "freedom and human rights" <sup>[5]</sup>, as well as a series of protests triggered by major hot events, the epidemic didn't reach a plateau after a surge like other countries <sup>[6]</sup>, inversely, it had been in a stage of rapid growth. Until 30 September, the cumulative number of confirmed cases had reached 7,400,589.

The COVID-19 is caused by the spread of the new coronavirus SARS-CoV-2, which can exist in the air or droplets and spread through breathing and talking when people contacting with others face to face <sup>[7-9]</sup>. In addition, this type of virus also has the characteristics of aerosol transmission <sup>[10-11]</sup>. Compared with SARS, the fatality rate of SARS-CoV-2 is a bit lower, but its community communication power shows a higher level <sup>[12-13]</sup>.

The spread of SARS-CoV-2 depends on droplets or aerosols, thus the influence of environmental factors such as temperature, humidity, rainfall, air pressure should also

be considered <sup>[19-21]</sup>. Previous studies showed that the survival and spread of coronavirus are significantly affected by weather factors. The daily incidence rate reaches its peak at around 16-28°C <sup>[22-23]</sup>. Some experimental researches showed that high temperature and high humidity will accelerate the inactivation of the virus <sup>[24-25]</sup>.

The global economy showed a downward trend under the influence of COVID-19, and major industrial chains have suffered heavy blows <sup>[14-15]</sup>. With reference to the SARS event in 2003, tertiary industries such as optional consumer industries (education and entertainment) were severely affected <sup>[16-18]</sup>. The unemployment of people working in the catering and tourism industries, or the sudden decline in the income of service personnel may lead to group social behaviors, such as demonstrations and large gatherings. This is a social problem that makes COVID-19 spread through gatherings.

This study aims to discuss the development of the COVID-19 epidemic in the U.S. and analyze the association between COVID-19 daily incidence rate (DIR) and people's implementation of states' quarantine policy and environmental factors.

## **Methods**

### **Data sources**

In this study, data of 50 states in the U.S. were used as the research subjects. According to the data released by Johns Hopkins University <sup>[6]</sup>, the DIR data of each state from 15 March to 30 September, 2020 were included as the dependent variable. Independent variable data included the proportion of daily residents at home (AHR),

the proportion of daily travel (TR) was obtained from the website of the Bureau of Transportation Statistics <sup>[26]</sup>. TR equals to the number of travel times of American divided by the daily exposure population. Missing values were completed by linear interpolation.

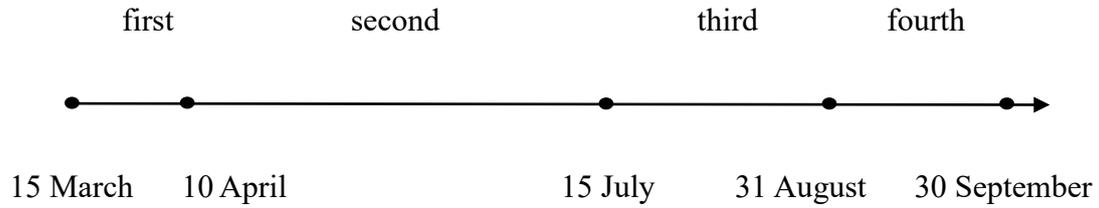
Environmental factors including daily temperature (T), humidity (H), air pressure (AP), wind speed (WS) and precipitation (PPTN) of state were collected from Weather Underground website <sup>[27]</sup>. Monthly unemployment rate (UR) and labor force participation rate (LFPR), overseas importing situation (OIS) data equals to importing frequency divided by the monthly exposed population were obtained from official website of the U.S. Department of Labor <sup>[28]</sup> and the Bureau of Foreign Import Statistics <sup>[29]</sup>.

## **Preprocessing**

The onset of COVID-19 has a certain incubation period. Previous studies found that the effect time of exposure to coronavirus is 5-7 days <sup>[30-31]</sup>. Although there are some reports suggesting that it may be much longer <sup>[32]</sup>, this study finally used 7-days incubation period determined by previous studies as the standard to preprocess the data, which means the respective variables (in days) would be correspond to the DIR after 7 days.

In addition to the panel regression analysis of the DIR and the independent variables of 50 states from March to September, 2020, this study also took the time when the quarantine policy was issued and major events happened as the time-splitting points, divided March to September into five segments. The main influencing

factors of DIR at each segment were explored. The segmentation method is as follows: The isolation and class suspension policies were promulgated in most states around April 10. Starting from July 15th, U.S. states gradually resumed work and production until the coming of the election month in September. Therefore, based on the above time nodes, this study classified 15 March to 10 April as the first segment, 11 April to 15 July as the second segment, 16 July to 31 August as the third segment and September as the fourth segment. Separately analysis between DIR and influencing factors was done after the stratification above.



### Panel data model

Panel data is a set of two-dimensional cross-sectional data that contains both time and space. It can be understood as a set of data formed by intercepting certain characteristic values of  $i$  objects at  $t$  different time nodes<sup>[33]</sup>. Therefore, the panel data can be represented by double subscript variable  $y_{it}$ .

$$y_{it} = \alpha_i + \sum_{k=1}^k x_{itk} \beta_k + u_{it} \quad (1)$$

$\alpha_i$  : intercept

$i = 1, 2, 3 \dots N$  (Number of subjects)

$t = 1, 2, 3 \dots T$  (Point of observation of each individual)

$k =$  Number of explanatory variables

$u_{it}$  : random error

We can also regard multi-indicator panel data as a composite matrix. For the same moment, the observation values of different indicators of all samples can form a time matrix. For the same indicator, several samples can be selected to observe it at each moment and these data can form an indicator matrix.

This study used a panel data model to fit the daily incidence rates of 50 states in the U.S., and considered the development of COVID-19 both in the vertical - time dimension, and the horizontal – states dimension. Through the cluster and multiple linear regression model analysis of the panel data, the characteristic of both space and time dimensions of the epidemic can be separately explored.

### **Statistical analysis**

Firstly, the cluster analysis of the panel data model was conducted based on the traditional classic K-Means algorithm. The data in this study can be expressed as an  $n \times d$  matrix  $X$ , while  $n$  is the number of samples ( $n=8450$  in our study),  $d$  is the dimension of the samples ( $d=9$  in our study).  $k$  cluster centers are expressed as  $k \times d$  matrix  $C$ , while  $k=3$ , and each row of  $C$  represents a cluster center. The distance from the sample to the  $k$  centers is expressed as an  $n \times k$  matrix  $D$ .

According to the optimization problem (1) to assign each sample point to the new nearest class center (2) to form  $k$  classes and update the sample mean of this class as the class center. Then, update the class center iteratively until the class center keep stable.

$$\min_{c_1 \wedge c_k} \sum_{j=1}^k \sum_{i=1}^n \|x_i - c_j\|_2^2 \quad (2)$$

$$c_j^{(t)} = \frac{1}{n} \sum_{i=1}^n x_i \quad (3)$$

Group visualization is completed according to the maximum number of days that each research object belongs to a certain category in the research time. For example, according to our study, Alaska has most days in the third category, thus, we classified it into the third category (Figure 1).

Except for 50 states regression, regression analysis of every one of the three classification of was also completed in order to test the impact of clustering results on the regression model.

The hausman test was used to select random effects model or fixed effects model for panel regression analysis. And the fitting of the linear regression of the panel data model mostly performed by the ordinary least square method (OLS). After pre-analysis, the regression model of 50 states had a low degree of fit (R-square<0.2). While based on the characteristics of panel data: the disturbance items between different individuals are independent of each other, but there is often autocorrelation among the disturbance items of the same individual in different periods, so this study used the robust command to perform regression analysis under the clustering robust standard error to reduce the overestimation of the influence of the independent variable on the dependent variable to obtain a more accurate linear trend.

The python-based software code was built for cluster analysis and statistical description, stata16.0 software was used for panel linear regression analysis.  $X \pm SD$  and M (P<sub>25</sub>, P<sub>75</sub>) were respectively used to describe normal and non-normal data.  $\alpha = 0.05$ .

## Results

### Cluster and basic situation

The above independent variables were used to cluster the 50 states, and the frequency distribution cluster graph is shown in Figure 1. The first category contains the states with characteristics of high DIR, low daily AHR and TR, high T, H, AP, PPTN, and low WS represented by Louisiana and Mississippi, contains 5 states. The second category are the states contains 7 states represented by Colorado and New Mexico, which has low DIR, high daily AHR and TR, low T, H, AP, PPTN and high WS. The third category is represented by Alaska and Maine, containing the rest 38 states with low DIR, high daily AHR and TR, low T, high H, AP, WS and PPTN.

The weekly incidence rate of the representative states in above categories were used to conduct a line chart drawn to observe the development trend (Figure 2), showing overall trends of these 6 states are continuously rising, with no platform period of the COVID-19 epidemic coming out. The basic situation of influencing factors of these 6 states is shown in Table 1.

### Multivariate analysis

After completing the hausman test, the fixed-effects model was selected for multivariate regression analysis. According to the unsegmented results, the daily AHR, daily TR, daily H and DIR are negatively correlated, and the daily T and DIR are positively correlated (Table 2). The linear regression equation can be written as:

$$DIR_{ud} = -5.572e-04 \times AHR - 8.78e-05 \times TR + 1.77e-06 \times T - 3.17e-07 \times H + 5.20e-04$$

The regression results of the three categories after clustering were also listed in table 2. The results showed some differences after being divided into three clustering groups. In the unsegmented model, the number of significant variables of the three categories were all less than that given by the regression model of the 50 states. In R-square's best regression model for the second clustering group, only AHR, TR and T were significant, and the significance of H only appeared in the model of the first clustering group.

The monthly labor force participation rate (LFPR) is negatively correlated with the monthly incidence rate (MIR) of each state. The linear regression equation is:

$$MIR_{ud} = -4.60e-04 \times LFPR + 3.08e-02$$

In the first segment, the regression model for 50 states showed one more significant negative correlation variable AP than the unsegmented model, while AHR and T have no significant significance. The first and third categories also had such a significant AP variable. In connection with the previous clustering results, the AP level of these two categories were high, which is what they have in common. Besides, it is worth noting that in the regression model of the third category, daily H and DIR actually showed a negative correlation.

In the second segment, the daily AHR, TR and H showed a negative correlation with DIR, the daily T and DIR are positively correlated. Similar to the unsegmented results, the regression results for 50 states showed more significant variables than that from three categories regression models.

In the third segment, the regression model of 50 states and three categories all

showed that the daily AHR and DIR were negatively correlated. Except for the significant positive correlation between T and DIR given by the regression model of the second categories, the impact of other environmental factors on DIR was not significant.

In the fourth segment, the daily TR was positively correlated with DTR, and the daily T was negatively correlated with DIR. And in the second category regression model, there additionally existed the negative correlation between AHR and DIR.

The linear regression equations can be written as below:

$$DIR_{s1} = - 4.38e-05 \times TR - 6.11e-05 \times AP + 1.95e-03$$

$$DIR_{s2} = - 4.92e-04 \times AHR - 8.11e-05 \times TR + 9.81e-07 \times T - 2.64e-07 \times H + 4.54e-04$$

$$DIR_{s3} = - 8.33e-04 \times AHR - 4.38e-04$$

$$DIR_{s4} = 4.96e-05 \times TR - 5.43e-07 \times T$$

## Discussion

Our research found that the impact of social factors like quarantine behavior on the COVID-19 epidemic is of significant. In addition, environmental factors such as temperature and humidity may also be the important factors for the development of COVID-19 epidemic.

According to the cluster results, except Arkansas, which is a non-port city located north of Louisiana and Mississippi, the first category represented by Louisiana and Mississippi has the highest DIR among the 3 categories clustering from 50 states in the U.S. The remaining 5 states are all coastal cities located in the southern part of the

American continent. Even if we have not obtained any positive association between OIS and DIR (Table 2), we can slightly speculate that the rising number of imported cases may bring a certain impact on the DIR in such states. If appropriate and timely measures are taken, the epidemic can be relatively controlled <sup>[34]</sup>. The DIR levels of the two other categories are not very high. The second category which includes 7 states located in the western area of American continent except 3 coast states, only has a slightly more severe epidemic, obviously lower humidity, and slightly lower air pressure compared to the third category, playing a certain role in prompting the subsequent segmented / unsegmented linear regression analysis results.

Besides, the first category has the highest temperature, with a median above 70 °F, and the second, third category have a median between 60-70 °F and 40-50 °F, corresponding to their DIR levels decreasing successively, which suggests that temperature is of great significance to the development of the epidemic (Table 1). The quarantine measures were not carried out successfully by American, as well as such tight medical conditions in the U.S. at that period, patients cannot be completely quarantined from susceptible people even if they have reduced the frequency of going out. Therefore, the unsegmented results may show a certain significance for the COVID-19 development. In other words, the development of COVID-19 epidemic in the U.S. is relatively very close to the natural epidemic model.

Echoing the above clustering results, daily T and DIR are positively correlated based on March to August multivariate analysis results, while in September, the temperature and daily incidence are inversely correlated. Although most conclusion

obtained in laboratories showed that lower temperature can lead a higher survival ability of virus <sup>[35-37]</sup>, not only the resistance of the virus itself, but also the impact of environmental factors on transmission should be considered in the actual epidemic process. According to the research on the COVID-19 epidemic of various cities in China, the rise in temperature showed a positive effect on the incidence <sup>[22]</sup>. Previous studies <sup>[38]</sup> have also shown that the vitality of the virus is closely related to the increase in temperature. Referring to the spread of SARS in 2003 and MERS in 2012, the spread of coronavirus reached its peak at 82.4°F <sup>[39]</sup>. Our results just fit this characteristic for the U.S. is located in the northern hemisphere with lower temperature than 82.4°F for most dates during March to August, while the temperature in September is relatively higher – which homologizes winter to summer - corresponding to the positive correlation from unsegmented regression model and the negative correlation in September regression model (Table 2). Therefore, the temperature rise from winter to summer (<82.4°F) may indeed lead to the acceleration of the virus spread in the U.S., while the higher temperature in September did inhibit the spread of the virus to a certain degree. Thus, we can speculate that the increasing vitality of SARS-2 virus occupies the main effect even if human immunity is also gradually rising.

Since the third category contained 38 of the 50 states, its regression model result may have a significant representative meaning for the overall 50 states model. As mentioned in the results section above, the third category regression model showed the daily T and DIR actually were negative correlated, which may correspond to the

study analyzed the epidemic situation from 20 January to 29 February, 2020<sup>[40]</sup>. In the cold winter, the relative increase in temperature may inhibit the increase in daily incidence. This may be resulted from that a small increase in temperature in winter shows benefit for human immunity increasing and air ventilation with little change in virus activity at the same time.

The average humidity of most states was basically between 40%-85% during the study period- the air was relatively moist, only several states had the lowest humidity around 20% (Table 1). According to the research of Noti<sup>[41]</sup>, the most suitable humidity for the spread of coronavirus is around 40%, which is basically consistent with the negative correlation result between daily humidity and DIR obtained in this study (Table 2). However, except for the second segment, the correlation between daily H and DIR in the models of other segments was not significant. For the results of the daily AP and DIR in the first segment regression model of the 50 states, the first and third categories were negatively correlated, suggesting that the increase in air pressure might relatively inhibit the spread of the virus in the high-pressure weather in winter. Therefore, more research should be carried out in the future.

For the quarantine policy implemented by people - the daily transportation situation, the first category has the lowest daily AHR, suggesting that the effect of quarantine on DIR is of certain significance. The multivariate results of unsegmented analysis show that the higher the daily AHR, the lower DIR, which confirms the crowd gathering characteristics of COVID-19 epidemic. The most effective self-protection measure for susceptible people is to get away from the source of infection

and wear a mask when facing a respiratory disease transmitted by droplets. Therefore, the inadequacy of protective measures caused the epidemic approaching to the natural development trend, as well as the reason why the epidemic in the U.S. did not reach any plateau for a long time. However, the factor daily TR has a different performance – the daily TR of the first category is of a lower percentage, while that of the second and third categories are significantly higher. Multivariate regression results show that daily TR is negatively correlated with DIR, which seems to be inconsistent with the actual situation. There may exist two explanations: Firstly, the real proportion of daily TR is more than 100%, nearly reaching 200% (Table 1), which means every one of residents in these states has an average trip number around twice per day. We can infer that as long as people did go out (whether going downstairs to throw trash, going to the supermarket, et al.), any time was counted as a trip. Even a relatively low number of trips can lead to a certain degree of virus transmission yet, for community communication of COVID-19 is the most common way and the root cause for the whole world pandemic. Secondly, different developments of the epidemic situation in each state could have different impact on residents' travel - most people may choose not to travel when they find the situation is too serious. Therefore, further research will be carried out in the next step in order to judge whether the daily low TR due to the severe epidemic, or the daily high TR aggravated the epidemic and requires further research to discuss who is the real dependent variable. The reason why labor force participation rate carried out in monthly data was negatively correlated with DIR may be also similar to the daily TR, and needs to be further explored.

## **Limitation**

Our research also has certain limitations. We only considered the situation of COVID-19 in the U.S. for inferring the factors affecting the global pandemic even if the development of it was relatively close to the natural model. In addition, there are a wide range of factors affecting the spread of COVID-19. We only included independent variables from two perspectives - which was subjective to a certain extent. Therefore, the R-square of most regression models was not high, and the results and conclusions could only serve as a partial reference. Even so, it was not easy to show the degree of correlation between significant variables and the DIR of COVID-19 in such models. Further research needs to be completed in the future.

## **Conclusion**

The spread of COVID-19 in 50 states in U.S. was related to quarantine measures, temperature and humidity. The progress of the epidemic would be relatively slow if people chose to stay at home. Besides, the increase in temperature ( $<84.2^{\circ}\text{F}$ ) could be conducive to the spread of the epidemic, while the increase in relative humidity (40~70%) might inhibit the spread of the virus to a certain degree.

## **List of abbreviations**

COVID-19	Coronavirus disease 2019
U.S.	the United States
DIR	the daily incidence rate

AHR	residents at home
TR	proportion of travel people
T	temperature
H	humidity
AP	air pressure
WS	wind speed
PPTN	precipitation
UR	unemployment rate
LFPR	labor force participation rate
OIS	overseas importing situation

## **Declarations**

### ***Ethical Approval and Consent to participate***

Not applicable.

### ***Consent for publication***

Not applicable.

### ***Availability of data and materials***

The data are available from the public websites - <https://www.bts.gov/>.

<https://www.wunderground.com/>. <http://dol.gov/>. <https://www.wunderground.com/>.

An anonymous dataset is available to researchers upon request to the corresponding author after reviewing by the study team.

### ***Competing interests***

The authors declare that they have no competing interests.

### ***Funding***

This work was partially supported by National Natural Science Foundation of China (Grant Nos.41704116, 11901234), Jilin Provincial Excellent Youth Talents Foundation (Grant No.20180520093JH), Scientific Research Project of Education Department of Jilin Province (Grant No. JJKH20200933KJ).

### ***Authors' contributions***

BL, SL and JJ formulated or developed overall research goals and objectives; YG, TD, NY, HW, YY, LW collected data, conducted a research; YG and TD created original draft preparation; SL, JJ, NY, HW, YY, SY, MS, LW and YY reviewed and edited the draft. All authors approved the final version of the manuscript.

### ***Acknowledgments***

Not applicable.

### ***Authors' information***

Not applicable.

Figure 1. The frequency distribution cluster graph

Figure 2. The average weekly incidence rate trends of 6 representative states

### **References**

- [1]. Chinazzi M, Davis J T, Ajelli M, et al. The effect of travel restrictions on the spread of the 2019 novel coronavirus (COVID-19) outbreak[J]. *ence*, 368.
- [2]. Grein J, Ohmagari N, Shin D, et al. Compassionate Use of Remdesivir for

- Patients with Severe Covid-19[J]. *New England Journal of Medicine*, 2020, 382(24).
- [3]. Holshue ML, DeBolt C, Lindquist S, et al. First Case of 2019 Novel Coronavirus in the United States. *N Engl J Med*. 2020;382(10):929-936. doi:10.1056/NEJMoa2001191.
- [4]. Ghinai I, McPherson TD, Hunter JC, et al. First known person-to-person transmission of severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) in the USA. *Lancet*. 2020;395(10230):1137-1144. doi:10.1016/S0140-6736(20)30607-3.
- [5]. Doogan C, Buntine W, Linger H, et al. Public Perceptions and Attitudes Toward COVID-19 Nonpharmaceutical Interventions Across Six Countries: A Topic Modeling Analysis of Twitter Data. *J Med Internet Res*. 2020;22(9):e21419. Published 2020 Sep 3. doi:10.2196/21419.
- [6]. <https://coronavirus.jhu.edu/data/cumulative-cases>
- [7]. Cai J, Sun W, Huang J, Gamber M, Wu J, He G. Indirect Virus Transmission in Cluster of COVID-19 Cases, Wenzhou, China, 2020. *Emerg Infect Dis*. 2020;26(6):1343-1345.
- [8]. Lake MA. What we know so far: COVID-19 current clinical knowledge and research. *Clin Med (Lond)*. 2020;20(2):124-127.
- [9]. Li LQ, Huang T, Wang YQ, et al. COVID-19 patients' clinical characteristics, discharge rate, and fatality rate of meta-analysis. *J Med Virol*. 2020;92(6):577-583.
- [10]. Yanko R, Klitinich V, Haviv Y, et al. Inhalation Sedation During the COVID-19 Outbreak: An Expert Opinion. *Isr Med Assoc J*. 2020 Oct;22(10):533-535.
- [11]. da Silva PG, Nascimento MSJ, Soares RRG, et al. Airborne spread of infectious SARS-CoV-2: Moving forward using lessons from SARS-CoV and MERS-CoV [published online ahead of print, 2020 Oct 8]. *Sci Total Environ*. 2020;142802. doi:10.1016/j.scitotenv.2020.142802.
- [12]. Hu T, Liu Y, Zhao M, et al. A comparison of COVID-19, SARS and MERS. *PeerJ*. 2020 Aug 19;8:e9725. doi: 10.7717/peerj.9725. PMID: 32879801;

PMCID: PMC7443081.

- [13]. Cascella M, Rajnik M, Cuomo A, et al. Features, Evaluation, and Treatment of Coronavirus. 2020 Oct 4. In: StatPearls [Internet]. Treasure Island (FL): StatPearls Publishing; 2020 Jan–. PMID: 32150360.
- [14]. Harmooshi NN, Shirbandi K, Rahim F. Environmental concern regarding the effect of humidity and temperature on 2019-nCoV survival: fact or fiction. *Environ Sci Pollut Res Int.* 2020 Oct;27(29):36027-36036.
- [15]. Bi P, Wang J, Hiller JE. Weather: driving force behind the transmission of severe acute respiratory syndrome in China?. *Intern Med J.* 2007;37(8):550-554.
- [16]. Chan KH, Peiris JS, Lam SY, Poon LL, Yuen KY, Seto WH. The Effects of Temperature and Relative Humidity on the Viability of the SARS Coronavirus. *Adv Virol.* 2011; 2011:734690. Sarkodie SA, Owusu PA. Global assessment of environment, health and economic impact of the novel coronavirus (COVID-19). *Environ Dev Sustain.* 2020 Jun 5:1-11.
- [17]. Xie J, Zhu Y. Association between ambient temperature and COVID-19 infection in 122 cities from China. *Sci Total Environ.* 2020;724:138201.
- [18]. Van Doremalen N., Bushmaker T., Munster V. Stability of Middle East respiratory syndrome coronavirus (MERS-CoV) under different environmental conditions. *Eurosurveillance.* 2013;18.
- [19]. Biryukov J, Boydston JA, Dunning RA, et al. Increasing Temperature and Relative Humidity Accelerates Inactivation of SARS-CoV-2 on Surfaces. *mSphere.* 2020 Jul 1;5(4):e00441-20.
- [20]. Pani SK, Lin N-H, Babu SR. Association of COVID-19 pandemic with meteorological parameters over Singapore. *Sci. Total Environ.* 2020;740:2.
- [21]. Schneemann C, Weber E, Wolter MI, et al. Which Industries Are Economically Relevant to the System? *Wirtschaftsdienst,* 2020;100(9):687-693.
- [22]. Nakat Z, Bou-Mitri C. COVID-19 and the food industry: Readiness assessment. *Food Control.* 2021 Mar;121:107661.
- [23]. Sarkodie SA, Owusu PA. Global assessment of environment, health and economic impact of the novel coronavirus (COVID-19). *Environ Dev Sustain.*

2020 Jun 5:1-11.

- [24]. Tran BL, Chen CC, Tseng WC, et al. Tourism under the Early Phase of COVID-19 in Four APEC Economies: An Estimation with Special Focus on SARS Experiences. *Int J Environ Res Public Health*. 2020 Oct 16;17(20):E7543.
- [25]. Institute of Medicine (US) Forum on Microbial Threats. Learning from SARS: Preparing for the Next Disease Outbreak: Workshop Summary. Knobler S, Mahmoud A, Lemon S, Mack A, Sivitz L, Oberholtzer K, editors. Washington (DC): National Academies Press (US); 2004. PMID: 22553895.
- [26]. <https://www.bts.gov/>
- [27]. <https://www.wunderground.com/>
- [28]. <http://dol.gov/>
- [29]. <https://www.wunderground.com/>
- [30]. Lauer S.A, Grantz K.H, Bi Q, et al. The incubation period of coronavirus disease 2019 (COVID-19) from publicly reported confirmed cases: estimation and application. *Annals of internal medicine*. 2020;172(9):577–582.
- [31]. Wassie GT, Azene AG, Bantie GM, Dessie G, Aragaw AM. Incubation period of SARS-CoV-2: A systematic review and meta-analysis. *Curr Ther Res Clin Exp*. 2020 Oct 11:100607.
- [32]. Bai Y, Yao L, Wei T, Tian F, Jin DY, Chen L, Wang M. Presumed Asymptomatic Carrier Transmission of COVID-19. *JAMA*. 2020 Apr 14;323(14):1406-1407.
- [33]. Guo X, Ren D, Shi J. Carbon emissions, logistics volume and GDP in China: empirical analysis based on panel data model. *Environ Sci Pollut Res Int*. 2016;23(24):24758-24767.
- [34]. Bai Y, Yao L, Wei T, Tian F, Jin DY, Chen L, Wang M. Presumed Asymptomatic Carrier Transmission of COVID-19. *JAMA*. 2020 Apr 14;323(14):1406-1407
- [35]. Wu B, Lei ZY, Wu KL, He JR, Cao HJ, Fu J, Chen F, Chen Y, Chen B, Zhou XL, Huang T, Wu T, Du YG, Chen SX, Xiao FR, Gao ZL, He J, Lin F, Lin BL. Compare the epidemiological and clinical features of imported and local COVID-

- 19 cases in Hainan, China. *Infect Dis Poverty*. 2020 Oct 19;9(1):143.
- [36]. Morris DH, Yinda KC, Gamble A, Rossine FW, Huang Q, Bushmaker T, Fischer RJ, Matson MJ, van Doremalen N, Vikesland PJ, Marr LC, Munster VJ, Lloyd-Smith JO. The effect of temperature and humidity on the stability of SARS-CoV-2 and other enveloped viruses. *bioRxiv [Preprint]*. 2020 Oct 16:2020.10.16.341883.
- [37]. Casanova LM, Rutala JS, Weber WA, Sobsey MD. Effects of air temperature and relative humidity on coronavirus survival on surfaces. *Appl. Environ. Microbiol.* 2020;76:2712–2717.
- [38]. Bashir MF, Ma B, Bilal KB, Bashir MA, Tan D, Bashir M. Correlation between climate indicators and COVID-19 pandemic in New York, USA. *Sci. Total Environ.* 2020;728:138835.
- [39]. Tan J, Mu L, Huang J, et al. An initial investigation of the association between the SARS outbreak and weather: with the view of the environmental temperature and its variation. *J. Epidemiol Community Health.* Mar. 2005;59(3):186–192.
- [40]. Shi P, Dong Y, Yan H, Zhao C, Li X, Liu W, He M, Tang S, Xi S. Impact of temperature on the dynamics of the COVID-19 outbreak in China. *Sci Total Environ.* 2020 Aug 1;728:138890.
- [41]. Noti JD, Blachere FM, McMillen CM, et al. High humidity leads to loss of infectious influenza virus from simulated coughs. *PLoS One.* 2013; 8(2):e57485.

# Figures

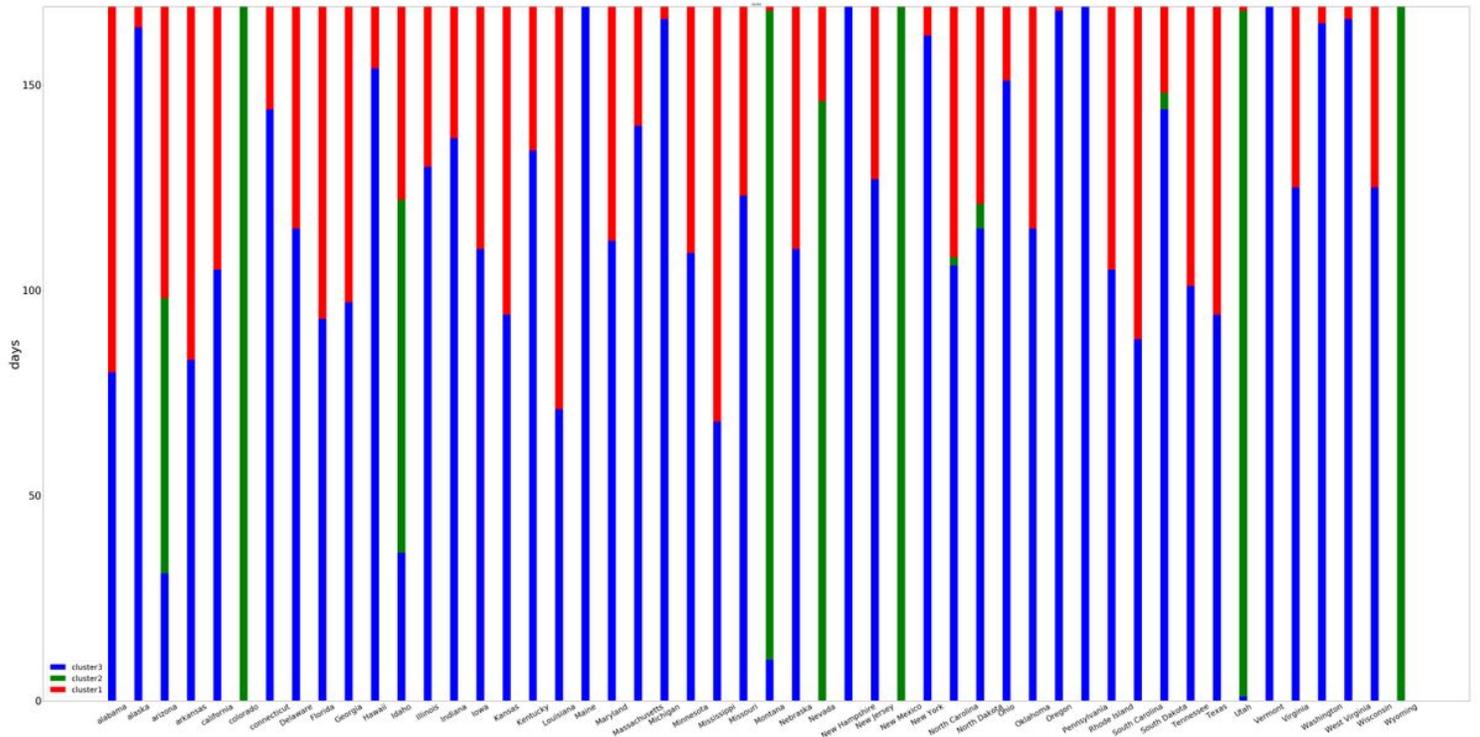
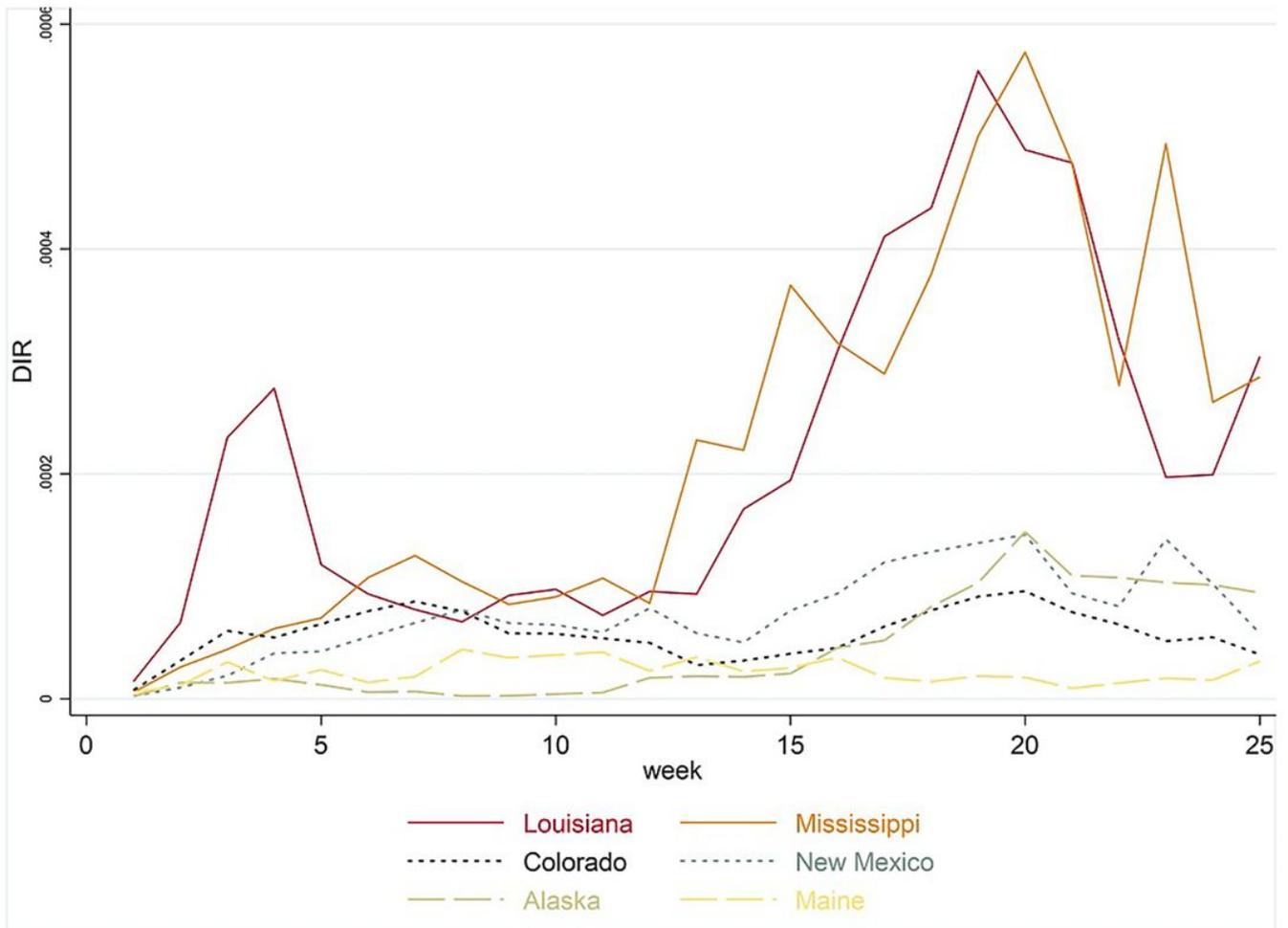


Figure 1

The frequency distribution cluster graph



**Figure 2**

The average weekly incidence rate trends of 6 representative states

## Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [table2.docx](#)