



*Correspondence:
Tapalina Bhattasali, St.
Xavier's College (Autonomous), Kolkata, India,
tapalina@sxccal.edu

Pandemic Analytics to Assess Risk of Covid-19 Outbreak

Tapalina Bhattasali

St. Xavier's College (Autonomous), Kolkata, India, tapalina@sxccal.edu

Abstract

COVID-19 pandemic has spread all over the world within a short period. It has changed every aspect of our daily lives significantly. The number of infected cases and the number of deaths are increasing day by day in many countries. Consequently, the situation becomes out of control. Due to recent advances in computational technologies, this paper focuses on the analytics part to assess various risks associated with the COVID-19 outbreak, which can be used to combat the severe effects of pandemics. Pandemic analytics is used to understand the spread pattern of pandemics by using the concept of artificial intelligence, machine learning, blockchain, and big data analytics. It is also required to evaluate policy for disease control. Based on the nature of the pandemic, a theoretical mathematical model is designed to predict the risks associated with the population all over the world. The analysis part is capable of forecasting the status and of answering various questions that arise from various parts of the world, such as the dependency of COVID-19 infection with sex, age, location, temperature, etc. Pandemic analytics is also used to visualize the official data before making any significant decisions.

Keyword: Pandemic; Analytics; Outbreak; Coronavirus; COVID-19

1. Introduction

The future status of any disasters needs to be predicted to take proactive actions. Prediction (World Health Organization, 2019). is based on the analysis of the data, which characterize the observed phenomena. The formulation of the theoretical model is very much needed to assess the disaster. If the predictions differ with the experimental observations because of its dynamic nature, the model can be customized accordingly. The procedure is continued until the discrepancies are minimized. The year 2020 has witnessed the emergence of coronavirus outbreak, which is caused by a novel virus SARS-CoV-2. The rapid spread of COVID-19 all over the world shows the limitations of existing infrastructure handle pandemic like disasters. Pandemic analytics, which is based on innovative technologies such as Artificial Intelligence (AI), Machine Learning, Big Data Analytics, and Blockchain, is considered a promising

solution to combat the COVID-19 outbreak. Early detection of an outbreak is possible as well as privacy protected while tracking the outbreak. COVID-19 can be considered as the data analysis application area. It is an infectious disease triggered by novel coronavirus. World Health Organization (WHO) (World Health Organization, 2019) has declared COVID-19 as pandemic as it quickly spreads all over the world and affects many people at the same time. A sudden increase in the number of infected cases gives rise to the outbreak of the disease.

Negative impact of the disease can be reduced if the right innovation can be utilized towards the right direction. Analytics bridges the gap between pandemic data and decision making. The data needs to be aggregated, refined, and analyzed in order to take accurate decisions. Analytics translates the technical data into useable information to facilitate decision making. It is the conglomeration and examination of information from numerous sources to urge the vision. Pandemic analytics is the way to assess the status of the outbreak, which in turn can be used to control the speedy spread of irresistible diseases like COVID-19. It is a valuable technique to combat this sort of worldwide crisis. As widespread advances, real-time modeling can enhance our knowledge towards the patterns of the outbreak. This type of modeling can be used as a simulator to assess the COVID-19 outbreak. However, real-time modeling of continuous widespread is not exceptionally simple. Pandemic analytics is used here for the statistical analysis of surveillance data. The model can provide an increasingly accurate assessment, and predictions of the pandemic evolution as data are available over time. Mathematical modelling (Giordano, G., Blanchini, F., Bruno, R., et al., 2020). is a crucial component to design strategies for pandemic mitigation. Pandemic analytics is increasingly used to study the spread of the virus with advances in data availability. This type of data analytics is considered for making better predictions using past data and analysis techniques, including statistical and learning methods. Discovery of patterns in input data and anticipating what is likely to happen is the main objective of the analytics. Statistical analysis and predictive modeling are the main pillars of pandemic analytics. The selection of an appropriate model decides how efficiently we can make better insights and useful decisions. For example, it can be used to predict the number of COVID-19 infected patients in the next few days. Pandemic analytics demand the design of a system model to analyze the official data, which can be used to address various predictive questions in the context of COVID-19. The classification model can be used for decision problems where only Yes or No is the answer.

The model classifies the data into multiple categories using previous data. The clustering model arranges data in several logical classes based on well-known attributes. Forecast models can be used to predict new values learned from previous data. The outlier model can be used to identify inconsistent entries within the dataset. The time series model can be used for short-term prediction using data collected from various time-periods. With the advent of advanced technology, sometimes the theoretical model may not require for the scenarios, where predictions can be made

by processing the existing data directly. However, prediction without understanding the details may be dangerous. Sometimes spurious correlation can emerge, where it is challenging to differentiate between exact cause, and it is hard to discriminate between a pure coincidence and a real cause. Prediction is challenging as the phenomenon is going on, and the nature of the virus is changing during mutation.

As large amount of official data about COVID-19 are readily available now; many researchers are developing predictive models to analyze data. However, many attempts have not produced fully satisfying results to predict the future of the COVID-19 pandemic. It may be due to the complexity in spreading procedure, various mutated forms of the virus, or limitation in official data. With the advent of pandemic analytics, it becomes easier to access huge amount of heterogeneous data related to the virus and the disease, collected from various sources. As a result, the facts can be visualized in more straightforward and practical ways.

Prediction of the status of the pandemic is possible with the application of technological know-how. Pandemic analytics is used here to assess COVID-19 status for a future time and a given location. There are various strategies to forecast the status of the virus. Warnings related to the pandemic outbreak are essential as the forecast is used here to protect the existence and the environment. Considering risk prediction is an essential step towards analytics of disease to generate risk scores.

The remaining parts of the paper are presented as follows. Section 2 gives a brief overview of the background of the COVID-19 outbreak. Section 3 presents a few research works related to the COVID19 pandemic outbreak. Section 4 gives an idea of the methodology of pandemic analytics to assess the status of COVID-19. Detailed analysis of data and relevant interpretation is described in section 5. Finally, section 6 concludes the paper.

2. Background

Coronaviruses are an own big family of various viruses. Some of them cause the common cold in human beings. Others infect animals, including bats, camels, and cattle. Experts say SARS-CoV-2 is originated in bats. That is also how the coronaviruses in the back of Middle East breathing syndrome (MERS) and severe acute breathing syndrome (SARS) were given started. In overdue December 2019, the novel Coronavirus was recognized in Wuhan, the capital of China, causing extreme respiratory disease. Since then, the virus has been spread. It is biologically called extreme acute breathing syndrome coronavirus 2 (SARS-CoV-2).

The sickness caused as a result of infection is known as coronavirus disease (COVID-19). COVID-19 has been labeled as an airborne High Consequence Infections Disease. SARS-CoV-2 is spreading between humans globally. Health Government in China aforementioned to the World Health Organization (WHO) about the cluster of viral infection cases of unknown reason in Wuhan, Hubei on 31st December 2019. In early January, an investigation was launched for this. On 30th January, the outbreak

is declared as Public Health Emergency of International Concern (PHEIC) by WHO (World Health Organization, 2019).

WHO diagnosed the spread of COVID-19 as a virulent disease on 11th March 2020. As SARS-CoV-2 spread both inside and outside of China, it infected humans who have had no direct contact with animals. That meant the virus is transmitted from one human to another. It is now spreading all around the world, which means that humans are unwittingly catching and passing on the coronavirus. This growing worldwide transmission has to turn out to be a plague. As a newly diagnosed virus, no human immunity is built towards this virus. At the identical time, no vaccine is to be had to save you contamination. Antibiotics are not an effective treatment as viral contamination.

Globally, the contamination rate has grown at such an alarming rate that a number of nations have already declared a state of emergency (World Health Organization, 2019). They have incorporated complete lockdown to stop further spread. Recent facts show that spreading rate becomes double than that for seasonal flu and fatality rate is ten times than that for seasonal flu. The increasing global contamination figures have shown that COVID-19 cases are rising sharply all over the world without any awareness. In some places, it becomes uncontrollable. The complicated cases become easier to spread and difficult to detect because of asymptomatic carriers, shortage of check kits, new pathogen of COVID-19. The increased infection instances will inevitably crush the health system as doctors and nurses are overworked to preserve up with new cases and therefore the number of clinic facilities, especially in-depth care units (ICU). Scientists have started works to find out remedies and vaccines to control the outbreak. However, drug discovery may be a lengthy process, which will probably take minimum of one year. There are wide ranges of challenges posed by coronavirus. It starts from early detection, containment, and isolation, to mitigation and segmentation, and in the end to the stop of the disease.

3. Related Works

Because of the fast spread of COVID-19, in a short time, several studies on the prediction of Trend and the effects thereof. This section discusses all these recent studies that are Linked to theoretical forecasts. Giulia Giordano et al. suggest an epidemic Predictive model comparing infected density with symptom frequency. Creators have proposed a SIDARTHE Model, which assists with rethinking multiplication number and reproduction results additionally show that the proposed model gives exact results in the wake of contrasting the discoveries and genuine information on the COVID-19 pandemic in Italy. Melanie Banister et al. (2020) introduced an intriguing examination to build up the connection of temperature and proof of COVID-19 in Europe. Authors guarantee that the regular variety, basically in the temperature, incredibly sway the spread of COVID-19. The study expresses that higher average temperature is the potential possibility to restrain the spread of COVID-19.

Lucia Russo et al. (2020) introduced an instrument to locate the principal day of

contaminations and forecasts of COVID-19 in Italy. Contingent on proposed work, creators can appraise that the genuine tally of uncovered instances of COVID-19. Vitaly Volpert et al. (2020) introduced the impact of the isolated model on the spread of coronavirus disease utilizing information investigation. The principle point of this work is to introduce the evaluation of set isolate components utilizing numerical demonstrating. Albertine Weber et al. (2020) introduced the pattern examination of the COVID-19 pandemic in China, utilizing a universally acknowledged SIR model in this investigation. The dataset utilized in this examination is taken from the Johns Hopkins University site for investigation, and it is found that scourge was contained in China. The fundamental point of the investigation introduced by Feng Zhang et al. (2020) is to give control measures to be considered universally for worldwide control of this pandemic. The period of the dataset is from 3-10 February 2020, and creators introduced a period arrangement model to foresee the number of tainted cases and the defining moment where the spread is at top. Possibility investigation of controlling COVID-19 spread by segregating tainted cases and isolate is introduced by Joel Hellewell et al. (2020) The proposed probabilistic model introduced in this examination considered differed situations like beginning diseases, fundamental reoccur number; also, the likelihood of contacts followed, and the pace of clinical contaminations. The outcomes appear that, in scourge circumstance, the seclusion of tainted individuals and contact following is not adequate to limit the pace of spread. Demonstrating of COVID-19 plague in China for an aggregate number of tainted cases utilizing information accessible in early stage in light of differential condition is introduced by Z. Liu et al. (2020) Simulation results show that on the off chance that the limitations would have been applied multi-week prior, at that point, there would have been a critical decrease in the number of tainted cases. From the literature study, it can be said that simple and robust approaches become increasingly important to assess the risks associated with COVID-19 like pandemics.

4. Methodology

Pandemic analytics is used to depict irresistible illness in an efficient way to interpret suspicions and information concerning illness transmission into the quantitative estimation of how a pandemic advances over time and space. Crichton's hazard triangle (Park, S. W., Bolker, B. M., Champredon, D., et al., 2020) is considered here to address pandemic risks. It can be seen from three viewpoints- hazards, exposure,

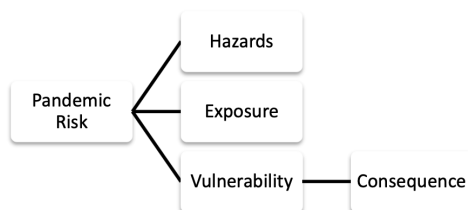


Fig. 1 Pandemic Risk Factors

and vulnerability.

- Hazard – the degree, seriousness, and likelihood of coronavirus’s danger, the potential danger, and its ability to cause a damage
- Exposure – the degree to which people at hazard may come into contact with the infected one; to measure how human is exposed to the potential damage
- Vulnerability – the susceptibility to harm the people at risk due to the infection at a specific escalated (co-morbidity); attitude to be affected when exposed

Hazard is related to the probability that a virus can spread. Initially, the spread is assumed as homogeneous. How the virus has been transmitted among people depends on a set of spatial and socio-economic factors. If a hazard represents the level of infection diffusion over an area, exposure is related to the size of the population in that area to represent the number of people affected by the infectious disease. Vulnerability represents the ability of an infected person to become sick or die – that is how he or she responds. It is a relationship between the infection and the health damage. Vulnerability is the degree of damage when the person is exposed to hazard. He or she may be recovered or die. This is the consequence. Vulnerability is related to the age and initial health conditions before the infection. The number of people who become affected due to virus can be easily measured if the vulnerability is combined with exposure. Consequently, it can be analyzed whether the person’s death is only due to COVID-19 or due to co-morbidity.

Due to the inter-relationships between these measurements, they have been conveniently conceptualized as a risk triangle. It can be surveyed in terms of various external factors like contamination, individual mobility, temperature, population density, average age, gender, and locality.

Three risk factors are considered to calculate the pandemic risk index (PRI). It can be used for exploring the correlations between this risk ranking and the observed effects of the COVID-19 outbreak

$$PR = aPE * (bPH * cPV).$$

Here PR represents a pandemic risk; PE represents exposure to the pandemic, PH represents hazards due to pandemic, and PV represents a vulnerability for the pandemic. a, b, c are the external factors which are responsible for consequences (Pc). The value of PC can be either 0.0 or 0.5 or 1.

Pandemic Risk Index for any location, $PRI, l = PR * P_c$, where $l=1,2,\dots$

0.0 implies death because of only COVID-19;

0.5 implies death due to co-morbidity (infected with COVID-19 along with other diseases);

1.0 implies recovery.

Pandemic modeling is considered here to describe the infectious disease in a systematic way to translate assumptions and data related to the transmission of disease into the quantitative estimation of pandemic progress through space and time. Pandemics of infectious diseases among humans are driven by transmission of the

virus among humans either directly or through various agents (like droplets) in the environment. The transmission of the disease depends on environmental, behavioral, and social factors. Pandemic dynamics can be expressed in terms of mathematical equations to analyze the data quantitatively.

During pandemic modeling, it is considered here that individuals go through eight health states with specific rates - Unwell, Endangered, Symptomatic, Asymptomatic, Confirmed, Quarantined, Recovered, Deceased. Unwell is the state of the individual when he or she is not feeling well due to cough, fever, cold, and difficulty breathing. These symptoms may arise due to general flue. Therefore, it can be said that individuals may feel unwell, which may not always ensure Corona Virus infection. If an individual is exposed to another COVID-19 patient, or he or she may suffer from lung disease, heart disease, diabetes, hypertension, or he or she has travel history. The endangered state has a high probability of being infected. Symptomatic is when an individual's symptoms are easily identified, while he or she is carrying the virus, and proper actions can be taken to reduce contamination. Asymptomatic is the state where the individual's symptoms are not visible while he or she is carrying the virus. Therefore, this type of individual's high probability is to spread the virus. Confirmed is the state when an individual is truly affected by the virus. Quarantined is the state when an individual is diagnosed and quarantined. Recovered is the state when an individual gets rid of the virus and recovers from the disease. Deceased is the state when an individual dies due to the coronavirus. This may be due to the only coronavirus or may be due to co-morbidity (where coronavirus is not only the reason for death; individual is already

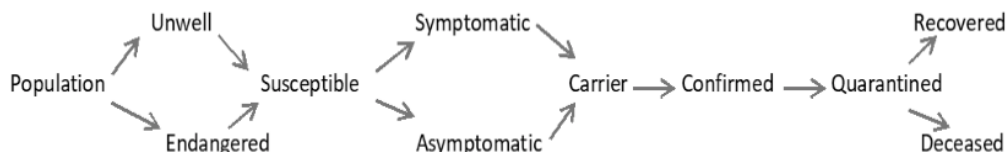


Fig. 2 Pandemic Modeling

affected by various diseases at the same time.

The state of the individual becomes confirmed when unwell or endangered individuals are infected by the virus. Confirmed individuals can transmit the virus. Confirmed individuals are in a recovered state after the infectious period. The individual may become immune to the disease after recovery. Deceased can also spread the virus.

Here a population of size N is partitioned into eight classes of individuals- Unwell (U), Endangered (E), Symptomatic(S), Asymptomatic (A), Confirmed (C), Quarantined (Q), Recovered (R), Deceased (D). Unwell (U) and Endangered (E) both are under Susceptible (S*) class. Symptomatic (S) and Asymptomatic (A) are both under Carrier (C*) class.

Therefore, $N = S^* + C^* + C + Q + R + D$

$$= U+E+S+A+C+Q+R+D, \text{ where } S^*=U+E \text{ and } C^*=S+A.$$

UESACQRD model is considered to analyze the status of a pandemic outbreak all over the world. Initially, a population is fully under susceptible (S) class of UESACQRD model of the COVID19 pandemic outbreak. It is just like the global population in December 2019, when the COVID-19 outbreak initiates. The susceptible population may become symptomatic carrier q and asymptomatic carrier $1-q$ with a rate of b . The symptomatic carrier may become confirmed at a rate of f , and asymptomatic carrier may become confirmed at a rate of g . Symptomatic and Asymptomatic carriers may be recovered after undergoing treatment with a rate of m and s , respectively. Individual who loses their temporary immunity becomes susceptible at a rate of y . New infection can be due to active contact with either the symptomatic carrier or asymptomatic carriers.

p_0 represents the average number of cases that result from the introduction of a single confirmed case in a totally susceptible population during the crisis period. The average number of unexposed individuals infected by a single confirmed individual is known as reproduction number and denoted by p_0 in UESACQRD model. The major estimation parameters include reproduction number (p_0), case lethality, or fatality rate (f), case recovery rate (h), and case confirmation rate (i). Reproduction rate (p_0) can predict whether the infectious disease will spread into a population or die out. If the value of p_0 is less than one, then the infection will die out quickly as each infectious person passes on the disease, on average, to less than one other individual. The outbreak cannot sustain its spread. If p_0 is larger than one, then the outbreak will grow exponentially.

The reproduction number wraps up all the troubles in the data about disease transmission and makes it transparent. From the manner by which the disease develops in the body to the method of transmission – and even the structure of the societies inside which it spreads – it catches all the key features of the outbreak and permits us to respond likewise.

The following parameters are considered here.

- ◆ proportion of the population that are unwell or endangered
- ◆ proportion of susceptible become symptomatic or asymptomatic carriers
- ◆ rate at which susceptible becomes a symptomatic or asymptomatic carrier
- ◆ rate at which symptomatic becomes confirmed
- ◆ rate at which asymptomatic becomes confirmed
- ◆ rate at which symptomatic becomes recovered
- ◆ rate at which asymptomatic becomes recovered
- ◆ rate at which confirmed becomes recovered
- ◆ rate at which confirmed becomes deceased

P_0 is divided into three components- the size of the population, the rate at which susceptible becomes confirmed, and the rate of recovery or death from the disease. Increasing the first two of these factors increases p_0 , whereas increasing the recovery

rate reduces it. The bigger the population and the quicker the disease spreads between people, the larger the outbreak is probably going to be. If individual's recovery becomes faster, individual has less scope to spread the disease to other people. Consequently, it will be simpler to manage the outbreak.

Then the reproduction number becomes effective. This is the average number of infections caused by the confirmed individual at a given point of time during the progression of the outbreak. If the effective reproduction number can be brought below 1 by intervention, then the disease will die out.

Isolation and confinement can productively lessen the transmission rate. Consequently, the effective reproduction number is also reduced. Confining infective patients decreases the pace of spread while isolating sound people lessens the effective susceptible population. The two activities add to diminishing the effective reproduction number. That is why social distancing and self-isolation are the basic methodologies for handling COVID-19.

Let us assume that the number of unwell, endangered, confirmed, recovered and deceased persons at time t in the population of size N is represented by $U(t)$, $E(t)$, $S(t)$, $A(t)$, $C(t)$, $Q(t)$, $R(t)$, $D(t)$. The following equations describe pandemic dynamics:

$$\begin{aligned} U(t) &= U(t-1) - (x/N)U(t-1)C(t-1), \\ E(t) &= E(t-1) - (x/N)E(t-1)C(t-1), \\ S(t) &= S(t-1) + (x/N)U(t-1)S(t-1) + (x/N)E(t-1)S(t-1) - fS(t-1) - hS(t-1), \\ A(t) &= A(t-1) + (x/N)U(t-1)A(t-1) + (x/N)E(t-1)A(t-1) - fA(t-1) - hA(t-1), \\ C(t) &= C(t-1) + (x/N)U(t-1)C(t-1) + (x/N)E(t-1)C(t-1) - fC(t-1) - hC(t-1), \\ Q(t) &= Q(t-1) + (x/N)U(t-1)Q(t-1) + (x/N)E(t-1)Q(t-1) - fQ(t-1) - hQ(t-1), \\ R(t) &= R(t-1) + fC(t-1), \\ D(t) &= D(t-1) + hC(t-1). \end{aligned}$$

The above equations are defined in discrete time points $t = 1, 2, \dots$, with the initial condition of the pandemic $U(0) = N - 1$, $E(0) = N - 1$, $S(0) = 1$, $A(0) = 1$, $C(0) = 1$, $Q(0) = 1$, $R(0) = D(0) = 0$. The parameters of the UESACQRD model are calibrated to fit the reported data.

Initially, the population N is considered as susceptible (S). p_0 can be estimated by the parameters of UESACQRD model as follows $p_0 = x/(f+h)$.

Here, $\Delta U(t) = U(t) - U(t - 1)$,

$$\begin{aligned} \Delta E(t) &= E(t) - E(t - 1), \\ \Delta S(t) &= S(t) - S(t - 1), \\ \Delta A(t) &= A(t) - A(t - 1), \\ \Delta C(t) &= C(t) - C(t - 1), \\ \Delta Q(t) &= Q(t) - Q(t - 1), \\ \Delta R(t) &= R(t) - R(t - 1), \\ \Delta D(t) &= D(t) - D(t - 1). \end{aligned}$$

The above equations represent reported new cases of unwell, endangered, confirmed, recovered, and dead at time t . $L\Delta U(t)$, $L\Delta E(t)$, $L\Delta S(t)$, $L\Delta A(t)$, $L\Delta C(t)$, $L\Delta Q(t)$, $L\Delta R(t)$, $L\Delta D(t)$ are the cumulative numbers of unwell cases, endangered

cases, symptomatic cases, asymptomatic cases, confirmed cases, quarantined cases, recovered cases and deceased cases at time t .

$$L\Delta Z(t) = \text{Sum}(Z_i(t)), \text{ where } i=1 \text{ to } t \text{ and } Z=\{U,E,S, A, C, Q R,D\}.$$

The resulting model includes flexible control of model parameters. It applies pandemic analytics to optimize resources, improve situational awareness, and ensure stability in demand-planning.

Algorithm:

A node which may not be reachable from the initially infected set of nodes will have zero expectation of infection. If every edge had a 100% chance of transmitting disease, then every node is reachable from the initially infected nodes. Therefore probability of getting infected is high here. This is often not the case. If the set of nodes are reachable from the initially infected nodes, it needs to be ensured that expected outbreak length is calculated according to the algorithm presented below. Computation becomes worst case linear within the dimension of the graph as computation of topological (or pre-order) sort of the vertices within the directed acyclic graph takes time. In practice, it is best to compute it for the vertices reachable from vertices within the outbreak seeding set. Use of easy breadth-first search saves significant time for several seeding sets. This computation is often done at the equal time because the computation of expectation values occurs free. If it is pre-computed, it takes time just like the expectation computation traversal.

Here, the objective is to calculate the probability that each node in a time-expanded graph is infected in pandemic at a given start node.

Input: A time-expanded graph $G = (V, E)$,

$N_s \leftarrow$ node where the pandemic will be seeded

$P_t \leftarrow$ a mapping of edges in G to probabilities that those edges will transmit infection if the source of edge is infected

Output: A mapping from nodes in the time-expanded graph to probabilities that each is infected in an

pandemic seeded at N_s

$T \leftarrow$ DAG of all nodes of G reachable from N_s

$O \leftarrow$ a topological ordering of nodes in T

$D \leftarrow$ an empty mapping

for (u, t) in O do

$P_t \leftarrow$ parents of (u, t) in T

$pNI \leftarrow 1$

$pNI \rightarrow$ prob. of non-infected

for (v, t') in P_t

do

$pl \leftarrow P_t[(v, t') \rightarrow (u, t)]$

$pl \rightarrow$ prob of infected

$pNI \leftarrow pNI \times (1 - pl(D[(v, t')]))$

end for

$D[(u, t)] \leftarrow 1 - \text{probNotInfected}$

end for return D

5. Analysis

This section includes a detailed analysis of data to assess the impact of COVID-19 all over the world. SPSS encapsulates mathematical and statistical knowledge to extract the analysis part and forecast reports, which are adaptive in nature to improve outcomes. It is expected to predict the outcome of the disease before they occur by using the benefits offered by SPSS. Publicly available official data (Roser, M., Ritchie, H., Ortiz-Ospina, E., & Hasell, J., 2020) are analyzed. Few points may differ due to several constraints imposed on the privacy of the data. Similarly, all the data may not be accurately available all the time.

Descriptive Statistics

	Mean	Std. Deviation	N
new_cases	16214.04	29339.028	184
new_deaths	1142.39	2297.360	184

Correlations

		new_cases	new_deaths
new_cases	Pearson Correlation	1	.963**
	Sig. (2-tailed)		.000
	N	184	184
new_deaths	Pearson Correlation	.963**	1
	Sig. (2-tailed)	.000	
	N	184	184

** . Correlation is significant at the 0.01 level (2-tailed).

It is seen that number of deceases are increasing with the number of new cases.

Correlations

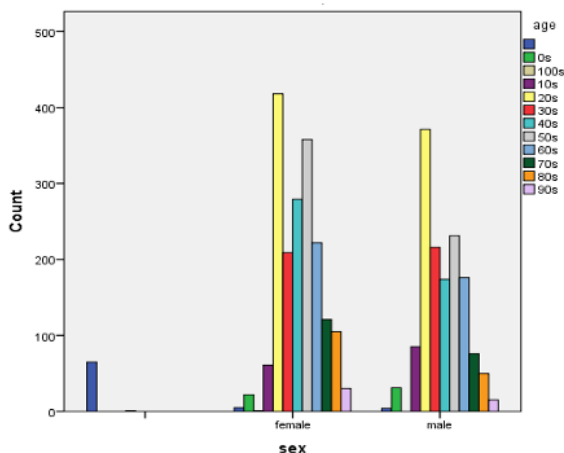
		Test	confirmed
test	Pearson Correlation	1	.966**
	Sig. (2-tailed)		.000
	Sum of Squares and Cross-products	3701543084915.304	79692278251.565
	Covariance	40676297636.432	875739321.446
	N	92	92
confirmed	Pearson Correlation	.966**	1
	Sig. (2-tailed)	.000	
	Sum of Squares and Cross-products	79692278251.565	1839423832.478
	Covariance	875739321.446	20213448.709
	N	92	92

** . Correlation is significant at the 0.01 level (2-tailed).

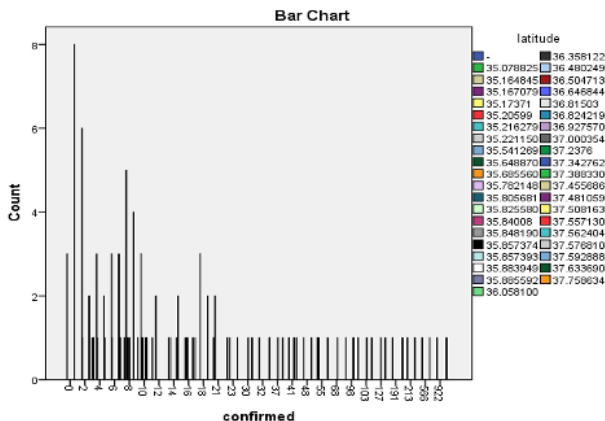
It can be said that there is a positive correlation between the number of tests and the number of confirmed cases.

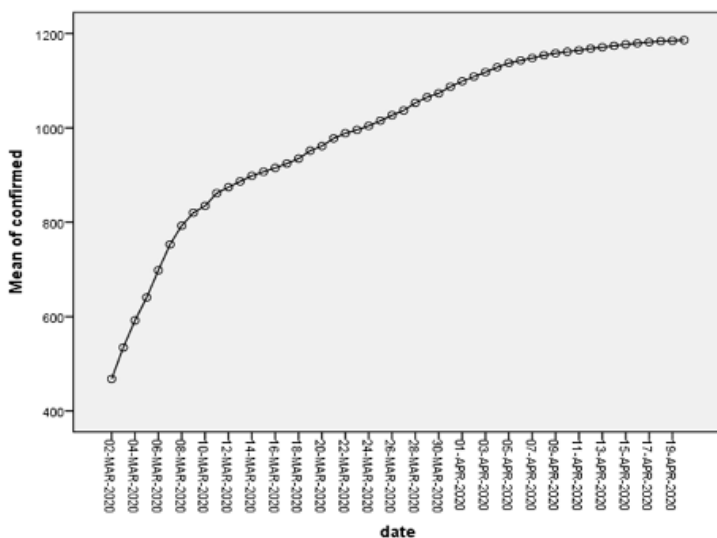
*sex * age Crosstabulation*

Count		age												Total
		0s	100s	10s	20s	30s	40s	50s	60s	70s	80s	90s		
sex		65	0	0	0	1	0	0	0	0	0	0	0	66
	female	5	22	1	61	418	209	279	358	222	121	105	30	1831
	male	4	31	0	85	371	216	174	231	176	76	50	15	1429
Total		74	53	1	146	790	425	453	589	398	197	155	45	3326



It can be said that there is no positive correlation between sex and confirmed cases.





6. Conclusions

Pandemic analytics modeling will stay an essential technique for infectious disease control in the future. Unused data on novel coronavirus and the development of the episode ended up on hand at an unusual pace within the automatic and globalized world. Still, important questions stay unanswered. Unique solutions for looking forward to the drift of the outbreak simply cannot be made in this way. We earnestly underline the vulnerability of official data, especially related to the real pattern of contaminated cases.

Pandemic analytics can dramatically speed up the information evaluation required to reap observations, responses, and suggestions for coping with the disaster generated from the COVID19 pandemic with more than predicted precision. It will take a leap forward in supplying a deeper perception into the developments and conduct of people, allowing correct predictions of the disease. With responsible recommendations in place, and when it becomes the incident of national emergency, analytics can do things quicker and more, as it should contain, mitigate, and end this global problem. This paper considers a few assumptions when building the model. We ignore the impact of factors such as population birth rate and natural mortality when we design a dynamic discrete version for a specific period for COVID19. This paper is based totally on the collected information for a particular period to fit and estimate the regeneration number, infection rate, and recovery rate of COVID19. The analysis part is based on government published statistics. Some records can be missing, and different factors may appear. More massive scale research wished to verify these findings.

Due to the pandemic of COVID-19, all nations are looking towards mitigation plan to govern the spread with the assist a few modeling techniques. This work empowers

the predictions targeting all the stakeholders involved in the COVID-19 pandemic. This paper provides research communities new insights on pandemic analytics to improve COVID-19 situation and motivates them for further studies in combating the outbreak of pandemics like COVID-19.

References

Bannister-Tyrrell, M., Meyer, A., Faverjon, C., & Cameron, A. (2020). Preliminary evidence that higher temperatures are associated with lower incidence of COVID-19, for cases reported globally up to 29th February 2020. *medRxiv*.

Giordano, G., Blanchini, F., Bruno, R., et al. (2020). A SIDARTHE model of COVID-19 epidemic in Italy. *arXiv preprint arXiv:2003.09861*.

Hellewell, J., Abbott, S., Gimma, A., Bosse, N. I., et al. (2020). Feasibility of controlling COVID-19 outbreaks by isolation of cases and contacts. *The Lancet Global Health*.

Liu, Z., Magal, P., Seydi, O., & Webb, G. (2020). Predicting the cumulative number of cases for the COVID-19 epidemic in China from early data. *arXiv preprint arXiv:2002.12298*.

Park, S. W., Bolker, B. M., Champredon, D., et al. (2020). Reconciling early-outbreak estimates of the basic reproductive number and its uncertainty: framework and applications to the novel coronavirus (SARS-CoV-2) outbreak. *MedRxiv*.

Roser, M., Ritchie, H., Ortiz-Ospina, E., & Hasell, J. (2020). Coronavirus disease (COVID-19)–Statistics and research. *Our World in data*.

Russo, L., Anastassopoulou, C., Tsakris, A., et al. (2020). Tracing DAY-ZERO and Forecasting the Fade out of the COVID-19 Outbreak in Lombardy, Italy: A Compartmental Modelling and Numerical Optimization Approach. *medRxiv*.

Volpert, V., Banerjee, M., & Petrovskii, S. (2020). On a quarantine model of coronavirus infection and data analysis. *Mathematical Modelling of Natural Phenomena*, 15, 24.

Weber, A., Ianelli, F., & Goncalves, S. (2020). Trend analysis of the COVID-19 pandemic in China and the rest of the world. *arXiv preprint arXiv:2003.09032*.

World Health Organization. (2019). Retrieved from: <https://www.who.int/emergencies/diseases/novel-coronavirus-2019>

Zhang, F., Zhang, J., Cao, M., & Hui, C. (2020). A simple ecological model captures the transmission pattern of the coronavirus COVID-19 outbreak in China. *medRxiv*.

Submitted 07.03.2020

Accepted 19.05.2020