



Prediction of Affected and Deceased Population Trend of COVID-19 in India using Statistical Analysis

A. Poompavai¹, N.Manjula Devi², Manimannan G^{3*}, A. Poongothai⁴

^{1,4}Assistant Professor of Statistics, Apollo College of Arts and Science, Chennai

²Senior Post Graduate Students, Department of Bio Statistics, ICMR-NIE, Chennai

³Assistant Professor of Statistics, TMG College of Arts and Science, Chennai

***Corresponding Author:** MANIMANNAN G, Assistant Professor, Department of Mathematics, TMG College of Arts and Science, Chennai

Abstract: This research paper attempts to identify the 7 day moving average, regression analysis and predict the data from February 2 to April 3, 2020. The secondary source of data were collected from Wikipedia. The data consists of the following parameters- New cases, Total cases, new death and total death of Indian States and Union Territories. The COVID-19 panic the entire world. India went on lockdown from March 24 to April 13, 2020. The aim was to find the trends in total affected and deaths population caused by Coronavirus Disease 2019(COVID-19) in Indian States and Union Territories. The main objective of this research paper were identified by using 7-day moving average, trend line, and regression analysis. In addition, the Data was extrapolated and predicted for the next 3 days after 3rd April 2020.

Keywords: COVID-19, Moving Average, Trend Line, Regression Analysis, ANOVA and Prediction.

1. INTRODUCTION

In recent days a big issue, an outbreak of the corona virus (COVID-19) in China has caused global concern. This virus came from a seafood and meat market in Wuhan, China, in December 2019. It has spread to other countries, including all the super power countries. Despite Wuhan and other Chinese cities being quarantined, the COVID-19 has spread to almost more than 160 countries internationally.

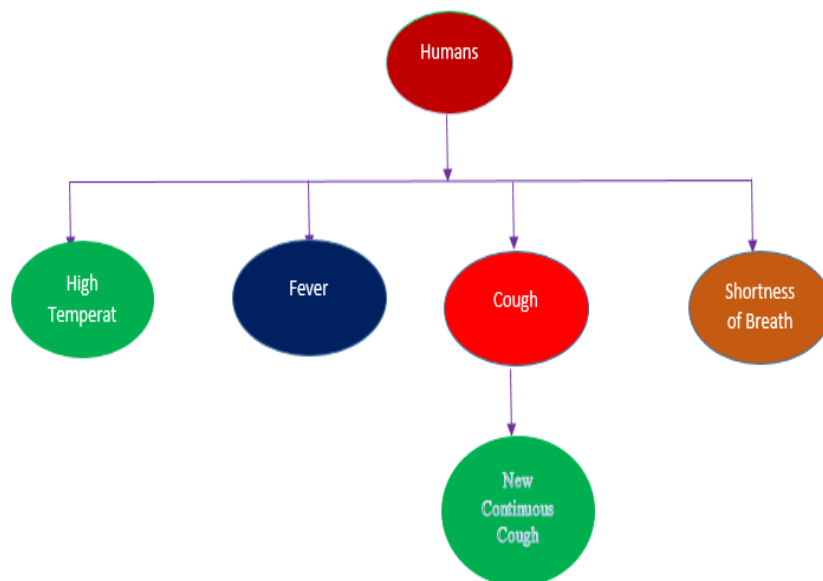
This one initiated by a member of the corona virus family that has never been encountered before. Comparable with other corona viruses, it takes transferred to humans from animals. The World Health Organisation (WHO) has declared it a epidemic. Corona virus is a virus that is found in animals and, hardly, can be transmitted from animals to humans and then spread person to person. In addition to COVID-19, other human. Corona viruses have included:

The MERS virus or Middle East respiratory syndrome.

The Severe Acute Respiratory Syndrome (SARS) virus, which first occurred in the Guangdong province in southern China.

Rendering to the WHO, the most common symptoms of Covid-19 are fever, tiredness and a dry cough. Certain patients may also have a runny nose, sore throat, nasal mobbing and aches and pains or diarrhoea. Around 80 percent of people who get COVID-19 knowledge a minor case, around as serious as a regular cold and improve without demanding any special treatment [1].

Approximately one in six people, the WHO says, develop seriously ill. The ageing and people with medical problems like high blood pressure, heart problems or diabetes, or chronic respiratory conditions, are at a greater risk of serious illness from COVID-19. Now in UK, the National Health Service (NHS) has identified specific symptoms to look for as experiencing either: The following structures are explained the symptoms of COVID-19.[2]



1. High temperature - you feel hot to touch on your chest or back
2. New continuous cough - this means you've started coughing repeatedly.
3. Fever
4. Cough
5. Shortness of breath

A new Corona virus or COVID-19 case in India is nearing the dreaded 500 mark. With the latest data from ICMR, 471 persons have been ill by the new virus as of March 23 of which 349 are Indians and 41 foreign nationals. Thirty five persons have been cured and discharged; one has been transferred. The States and Central governments in the meantime have decided to totally lock down 80 districts across the country where more number of corona virus cases have been reported. States such as Uttar Pradesh, Maharashtra, Punjab, Karnataka, Tamil Nadu and Kerala announced lockdown in many districts. The capital city of India, Delhi has put on locked down from 6 am on March 23 till midnight on March 31. Interstate passenger bus and trains services have been suspended till March 31 as well. India has also suspended visas till April 15. According to official figures, over 15 lakh people have been screened at various airports in India so far. The first case occurred by 2019–20 corona virus pandemic in India was reported on 30 January 2020 from Kerala, originating from China. As of 24 March 2020, the Indian Council of Medical Research and Ministry of Health and Family Welfare have confirmed a total of 494 cases, 36 recoveries, 1 migration and 9 deaths in the country. The virus infection proportion of COVID-19 in India is reported to be 1.7, which is remarkably lower than in the worst affected countries.

The corona virus extensive has been declared an epidemic in more than a dozen states and union territories, where provisions of the Epidemic Diseases Act, 1897 have been invoked, and most of the educational institutions and many commercial establishments have been shut down. India proceeds postponed all tourist visas, as a majority of the confirmed cases were linked to other countries. The government has also issued a complete lockdown in 23 States and 7 Union Territories covering 548 districts.

2. REVIEW OF LITERATURE

The Corona Virus Disease -19 (COVID-19) is a respiratory illness that can spread from person to person. The virus that causes COVID-19 is an innovative corona virus that was first identified during an examination into an outbreak in Wuhan, China. [3]. Many research study showed the significance of SARS details in the following research papers. The epidemic of Severe Acute Respiratory Syndrome (SARS) in 2002–2003 represented a serious public health threat to the international community. Its rapid spread to regions far away from the initial outbreak created great concern for the potential ability of the virus to affect a large number of countries and required a coordinated effort aimed at its containment [4]. Most importantly, it clearly pointed out that people's mobility and travelling along

commercial airline routes is the major channel for emerging disease propagation at the global scale. Spatiotemporal structures of human movements thus need to be considered for a global analysis of epidemic outbreaks [5], as for example in [6], which incorporates the airline network structure of the largest 500 airports of the world.

In the recent research using mathematical modelling based On 31 December 2019, the World Health Organization (WHO) China Country Office was informed about cases of pneumonia from unknown etiology (unknown cause) detected in Wuhan City, Hubei Province of China, and WHO reported that a novel Corona virus (2019-nCoV), which was named as severe acute respiratory syndrome corona virus 2 (SARS-CoV-2) by International Committee on Taxonomy of Viruses on 11 February, 2020, was identified as the causative virus by Chinese authorities on 7 January [7]. It is reported that the virus might be bat origin [6], and the transmission of the virus might related to a seafood market. The genetic features and some clinical findings of the infection have been reported recently. Potentials for international spread via commercial air travel had been assessed [8]. Public health concerns are being paid globally on how many people are infected and suspected.

However, the bat origin and the transmission route from the seafood market to people were not considered in the published models. In India COVID-19 spread from overseas Indian community, it is slowly spread over all the Indian States except very few states and union Territories. Indian population is very high followed by China in the world community. Indian government takes necessary steps day by day, but still it is in increasing trends. Indian population isolated from every states and within the districts. This study is based on the data source from Indian Health Organization up to March 24, 2020. The main objective of this research is to analyse and cluster the Indian states using k-means with silhouette distance measure of data mining.

3. DATABASE

The database were collected from Wikipedia during the period of February 2, 2020 to April 3, 2020. The data consists of various indian states and union territories with parameter of new cases, total cases, new deaths and total deaths that has occurred during this period. It includes 27 states and union territories.[9].

4. METHODOLOGY

Corona virus disease 2019 (COVID-19) was a chain originating in Wuhan, Hubei, China. Corona virus is a threat for all countries and threatens the lives of many people along with the economies of the world. It has spread further having already infected more than 1, 009, 885 people, death rate 59.197and recovered people 229,975 across the globe till date. In India, the affected population is 2902, death rate 68and recovered is 184. In this research paper,7 day moving averages was used to find the casualties caused by the new virus and new death along with the trend line for the expected casualties of the virus.

4.1. Method 1 Moving Average Algorithm

Step 1: Select Data Analysis Tool Package and Select Moving average method.

Step 2: Select the input range from (\$C\$2:\$C\$37) and choose interval as 7 days.

Step 3: Select the output range from empty cells (\$D\$2:\$D\$37)

Step 4: Select chart output and display the moving average chart.

Step 5: Repeat the steps 1 to 4 for every parameters.

4.2. Method 2. Regression Analysis

Step 1: Select input Y range and X range for infected new case and infected total case.

Step 2: Select data labels and confidence interval (95%)

Step 3: Output options, choose new Workbook.

Step 4: Choose Residuals: Residuals, Residual plots, Standardized Residuals and Line Fit Plots

Step 5: Select normal probability plots.

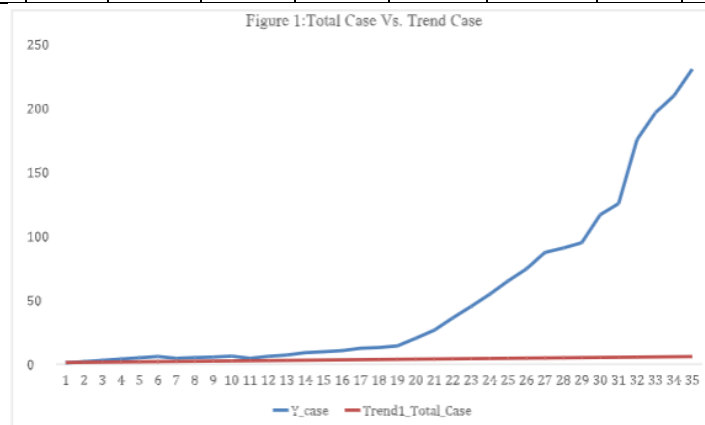
Step 6: Repeat the steps 1 to 5 for all the parameters [8].

5. RESULT AND DISCUSSION

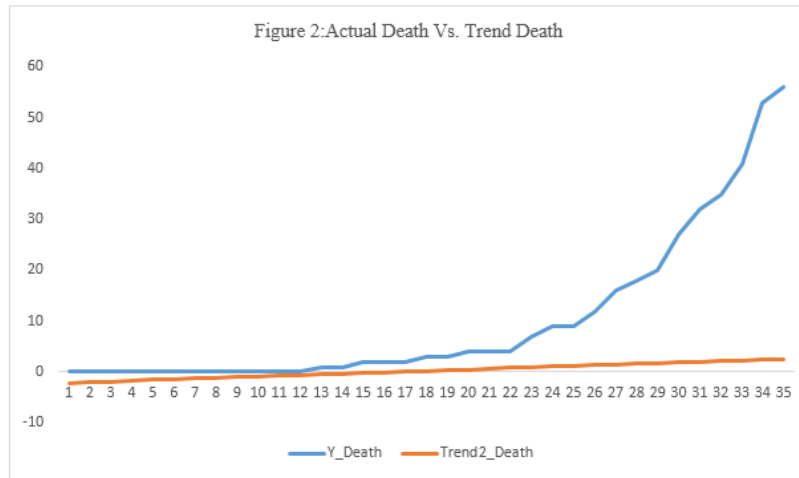
The following table shows that 7 days moving average and trend line of total affected population and total death in India (Table 1). Figures 1 and 2 explains the new cases vs. total infected cases of trend line and new death vs, total death trend line up to April 3, 2020.

Table1. Data for Moving Average

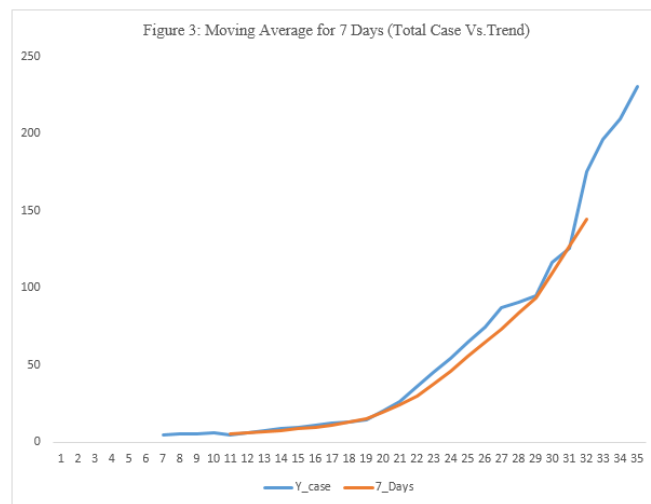
Date	Day_No	N_Case	Y_case	7_Days	N_Death	Y_Death	7_Days	x1=d-17.5	T_Case	T_Death	Predi.1	Predil
Feb-02	1	1	2		0	0		-16.5	1.32386	-2.12175	67.41694	3.836455
Feb-03	2	1	3		0	0		-15.5	1.45968	-1.98295	67.41694	3.972271
Mar-02	3	2	5		0	0		-14.5	1.59549	-1.84415	73.44626	4.243904
Mar-03	4	1	6	12.33333	0	0	0	-13.5	1.73131	-1.70536	67.41694	4.37972
Mar-04	5	22	28	15	0	0	0	-12.5	1.86712	-1.56656	194.0326	7.367675
Mar-05	6	2	30	19.57143	0	0	0	-11.5	2.00294	-1.42776	73.44626	7.639307
Mar-06	7	1	31	24.71429	0	0	0	-10.5	2.13876	-1.28896	67.41694	7.775123
Mar-07	8	3	34	30.28571	0	0	0.142857	-9.5	2.27457	-1.15017	79.47557	8.182572
Mar-08	9	5	39	36.57143	0	0	0.285714	-8.5	2.41039	-1.01137	91.53421	8.861652
Mar-09	10	5	44	41.14286	0	0	0.571429	-7.5	2.54620	-0.87257	91.53421	9.540733
Mar-10	11	6	50	47.28571	0	0	0.857143	-6.5	2.68202	-0.73378	97.56353	10.35563
Mar-11	12	10	60	54.42857	0	0	1.142857	-5.5	2.81784	-0.59498	121.6808	11.71379
Mar-12	13	13	73	63.42857	1	1	1.571429	-4.5	2.95365	-0.45618	139.7688	13.4794
Mar-13	14	8	81	73.14286	0	1	2.000000	-3.5	3.08947	-0.31739	109.6222	14.56593
Mar-14	15	16	97	83.71429	1	2	2.428571	-2.5	3.22528	-0.17859	157.8567	16.73899
Mar-15	16	10	107	96.14286	0	2	2.857143	-1.5	3.3611	-0.03979	121.6808	18.09715
Mar-16	17	11	118	109.1429	0	2	3.142857	-0.5	3.49692	0.09901	127.7101	19.59113
Mar-17	18	19	137	123.4286	1	3	3.857143	0.5	3.63273	0.2378	175.9447	22.17163
Mar-18	19	14	151	143.7143	0	3	4.857143	1.5	3.76855	0.3766	145.7981	24.07306
Mar-19	20	22	173	170.2857	1	4	5.714286	2.5	3.90436	0.5154	194.0326	27.06101
Mar-20	21	50	223	206.4286	0	4	7.000000	3.5	4.04018	0.65419	362.8536	33.85182
Mar-21	22	60	283	251.5714	0	4	8.714286	4.5	4.17600	0.79299	423.1467	42.00079
Mar-22	23	77	360	306.1429	3	7	10.71429	5.5	4.31181	0.93179	525.6452	52.45863
Mar-23	24	74	434	371.1429	2	9	13.00000	6.5	4.44763	1.07058	507.5572	62.50902
Mar-24	25	85	519	445.5714	0	9	15.85714	7.5	4.58344	1.20938	573.8797	74.0534
Mar-25	26	87	606	532.8571	3	12	19.14286	8.5	4.71926	1.34818	585.9384	85.8694
Mar-26	27	88	694	623.5714	4	16	22.85714	9.5	4.85508	1.48698	591.9677	97.82122
Mar-27	28	140	834	718.4286	2	18	27.00000	10.5	4.99089	1.62577	905.4923	116.8355
Mar-28	29	84	918	835.1429	2	20	32.28571	11.5	5.12671	1.76457	567.8504	128.244
Mar-29	30	106	1024	960.5714	7	27	37.71429	12.5	5.26252	1.90337	700.4954	142.6405
Mar-30	31	227	1251	1136	5	32	40.66667	13.5	5.39834	2.04216	1430.043	173.4708
Mar-31	32	146	1397	1332.429	3	35	43.40000	14.5	5.53416	2.18096	941.6682	193.3
Apr-01	33	437	1834		6	41		15.5	5.66997	2.31976	2696.2	252.6516
Apr-02	34	235	2069		12	53		16.5	5.80579	2.45855	1478.278	284.5684
Apr-03	35	232	2301		3	56		17.5	5.9416	2.59735	1460.19	316.0777



According to the above figure, the total case fatalities are expected to reach very low level in new case and total case of population in Table 1. By 6th April in the long term it shows a constant decrease although in reality it might be decrease due to the manner in which the disease spreads and external factors like lockdown and social distancing in India.



The above figure shows that the total case fatalities are expected to reach low level new death cases and total death cases in Table 1. By 6th April in the long term it shows a constant decrease although in reality it might be decrease due to the manner in which the disease spreads and external factors like lockdown and social distancing in India.



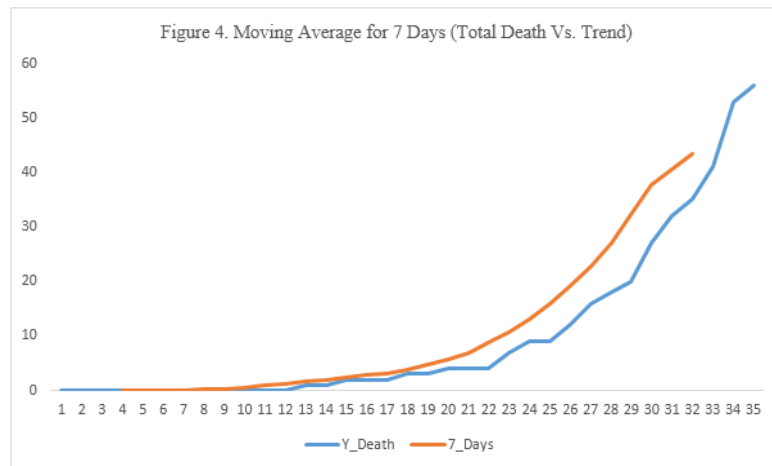
The above figure indicates that averages of 7 days of new active and total active case caused by COVID-19 in India from 2nd February to April 3rd, 2020. In this graph it clearly shows that, the trend is rapidly increasing every day and has not declined and thus we do not expect any decline in this trend. The rate of deaths increased after 18th day.

The above figure indicates that averages of 7 days of new deaths and total active death cases caused by COVID-19 in India from 2nd February to April 3rd, 2020. In this the graph shows that the trend is rapidly increasing every day and has not declined and thus we do not expect any decline in this trend. The rate of deaths has increased after the 18th day.

Table 2. Regression Analysis for Total Cases

Regression Statistics	
Multiple R	0.904919
R Square	0.818879
Adjusted R Square	0.81339
Standard Error	40.71937
Observations	35

ANOVA								
	df	SS	MS	F	Significance F			
Regression	1	247380.9	247380.9	149.1984	8.64E-14			
Residual	33	54716.2	1658.067					
Total	34	302097.1						
Regression Coefficients								
	Coefficients	Standard Error	t Stat	P-value	Lower 95%	Upper 95%	Lower 95.0%	Upper 95.0%
Intercept	3.564823	8.559326	0.416484	0.679752	-13.8493	20.9789	-13.8493	20.9789
X Variable 1	0.135816	0.011119	12.21468	8.64E-14	0.113194	0.158438	0.113194	0.158438

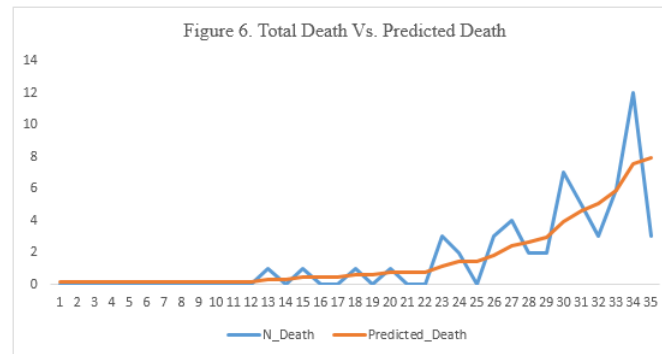
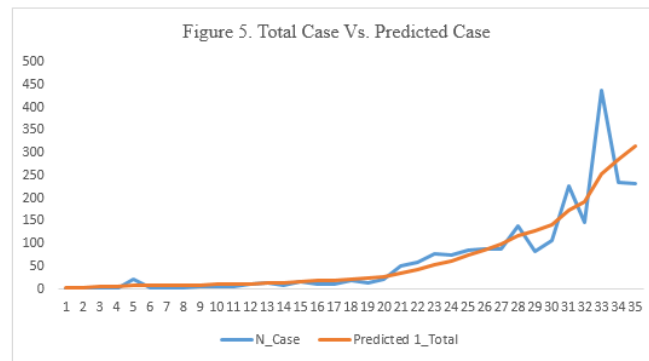


From the above table it is clear that R-square value is 0.818, ANOVA significant value is 0.000 and Regression coefficients for the best fitted model of new cases and total cases of affected population, the ANOVA significantly differs from new affected case and Total affected case of COVID-19. The regression coefficient is used for Trend line of Total new affected case and total case of population from COVID-19.

Table3. Regression Analysis for Death Cases

Regression Statistics								
Multiple R	0.829251							
R Square	0.687656							
Adjusted R Square	0.678191							
Standard Error	1.483124							
Observations	35							
ANOVA								
	df	SS	MS	F	Significance F			
Regression	1	159.8114	159.8114	72.65289	7.53E-10			
Residual	33	72.58864	2.199656					
Total	34	232.4						
Regression Coefficient								
	Coefficients	Standard Error	t Stat	P-value	Lower 95%	Upper 95%	Lower 95.0%	Upper 95.0%
Intercept	0.168404	0.301755	0.55808	0.580557	-0.44552	0.78233	-0.44552	0.78233
X Variable 0	0.138797	0.016284	8.523667	7.53E-10	0.105668	0.171927	0.105668	0.171927

The above table shows that R-square value is 0.6876, ANOVA significant value is 0.000 and Regression coefficients for the best model fitted, the ANOVA significantly differs from new death cases and Total death cases of COVID-19. The regression coefficient is used to identify the Trend line of Total new death cases and total death of population from COVID-19. The following figures 5 and 6 shows the originally affected and death cases of population from COVID-19.



6. CONCLUSION

This research paper concludes with the research findings by using the moving average, regression analysis and the prediction of total affected case and total death case of population from COVID-19 in various Indian states and union territories from 2nd February to 3rd April, 2020. It shows the rate at which active and deaths have been increasing and has triggered a panic due to the current trend of Indian scenario. The Central and State government takes utmost care of Indian citizens to the maximum effort and lockdown the entire country from March 23 to April 13, 2020. The overall performance is steadily increasing with the affected and death population in India.

REFERENCES

- [1] Huang C, Wang Y, Li X, Ren L, Zhao J, Hu Y, et al. Clinical features of patients infected with 2019 novel corona virus in Wuhan, China. *Lancet*. 2020.
- [2] World Health Organization. Corona virus. World Health Organization, cited January 19, 2020. Available
- [3] Dye C, Gay N: Modelling the SARS epidemic. *Science* 2003, 300:1884-1885.
- [4] Hufnagel L, Brockmann D, Geisel T: Forecast and control of epidemics in a globalized world. *Proc Natl Acad Sci USA* 2004.101:15124-15129
- [5] Li Q, Guan X, Wu P, Wang X, Zhou L, Tong Y, et al. Early transmission dynamics in Wuhan, China, of novel corona virus-infected pneumonia. *N Engl J Med*. 2020.
- [6] Zhu N, Zhang D, Wang W, Li X, Yang B, Song J, et al. A novel corona virus from patients with pneumonia in China, 2019. *N Engl J Med*. 2020..
- [7] Bogoch II, Watts A, Thomas-Bachli A, Huber C, Kraemer MUG, Khan K. Pneumonia of unknown etiology in Wuhan, China: Potential for International Spread Via Commercial Air Travel. *J Travel Med*. 2020. .
- [8] Wu JT, Leung K, Leung GM. Nowcasting and forecasting the potential domestic and international spread of the 2019-nCoV outbreak originating in Wuhan, China: a modelling study. *Lancet*. 2020.
- [9] A.B.M. Shawkat Ali and Saleh A, Wasimi, 2009, Data mining: Methods and Techniques, Cengage Learning India Private Ltd. Delhi, India.

Citation: MANIMANNAN, et.al., Prediction of Affected and Deceased Population Trend of COVID-19 in India using Statistical Analysis, *International Journal of Scientific and Innovative Mathematical Research (IJSIMR)*, vol. 8, no. 2, pp. 37-43, 2020. Available : DOI: <http://dx.doi.org/10.20431/2347-3142.0705005>

Copyright: © 2020 Authors. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.