

DataAnalysis forCOVID-19

*E-mailIdofcorrespondingauthor-*alwinece2022@gmail.com

***Abstract***

**Corona Virus Disease- 19 (COVID-19) was first time reported in Wuhan, China. This disease has covered more than 200countries till May 2020. World Health Organisation (WHO) has declared COVID-19 as Public Health Emergency ofInternational Concern (PHEIC) on 30 January 2020. COVID- 19 causes severe acute respiratory syndrome coronavirus 2(SARS-CoV-2) which was progressive earlier in China but now in maximum countries. Therefore, the different onlineplatform are used which provides the latest update of confirmed corona cases throughout the globe for the analysis ofdata. The aim of data analysis for CIVID-19 is to aware of the community against the infectious disease and forecast theCOVID-19 confirmed cases, deaths, and recoveries through the data analysis methods. Different models are also used tostudy the behavior of the disease. The models help to forecast the patterns of public sentiments on health information withboththepoliticalandeconomicalinfluenceofthespreadofthevirus.DataanalysismethodswhichareusedareExploratoryData Analysis (EDA) in which the numberof confirmed cases, death,and recovered data are recorded,model like Susceptible-Exposed-Infectious-Recovered (SEIR) model is used to predict the time and the rate taken for thespreading up of disease throughout the globe. A statistical model can also be used to compare the data among differentcountriestomakehumans awareof theinfection.**

***Keywords:***2019-nCoV,SARS-CoV-2,Coronavirus,COVID-19,dataanalysis,visualization.

1. INTRODUCTION

The outbreak of the new disease in Wuhan, Chinawas caused by novel Coronavirus (2019-nCoV) [1]. This diseaseis a form of pneumonia. Coronavirus belongs to the *Orthocoronavirinae* subfamily. The first case was observed atthe Chinese Center for Disease Control and Prevention (CDC) on 12 December 2019 and was considered as a non-SARS novel coronavirus [2]. The family to which Coronavirus belongs is *Coronaviridae* which consists of a large,single RNA strand of plus sign [3]. Viruses of these family show the symptoms of common cold, diarrhea in humanbeings. In the year 2003, it was seen the outbreak of coronavirus i.e. severe acute respiratory syndrome coronavirus(SARS-CoV) [4]. In December 2019 at Wuhan, China's symptoms closely resembled the same as pneumonia [5].Several cases of approximately 1974 were confirmed in China according to thecouncil information office inBeijing, China's capital on 26th January,2020. Virus started spreading inmany other countries likethevery firstcase after China was reported in Thailand, Japan and two cases were also seen in Korea on 16 January 2020. Recentresearches have shown some evidence of the origin of the virus from the bat and it was also seen that transmission ofthevirusistakingplacefromhumantohuman.Thesituationstarted gettingworstfrom19January2020 daybyday,

sototakesomeseriousactionforthecontrolandpreventionfromthedisease.WorldHealthOrganisation(WHO)on 30 January 2020 declared that Coronavirus Disease was an outbreak emergency of international concern after theattack of H1N1 in 2009, the emergence of Ebola virus in 2014, polio in 2014 and Zika virus in 2016[6] [29].Finally, on 11 February 2020, World Health Organization (WHO) gave the name of the novel disease which wascaused by the corona virus as Corona Virus Disease- 19 (COVID-19) [7] [32]. Record maintenance on 24 February2020 showed that more than 78, 000 patients were suffering from COVID-19 throughout many countries. Themaximum patients were from China according to the World Health Organization (WHO) which were approximately77,000 and 2500 death [8]. According to the World Health Organization (WHO) the rest of the countries reported2000 confirmed cases and 300 deaths as on 7 March 2020. In Wuhan, China lockdown orders of all the trains, fightsand public transport were passed on 23 January 2020. The exact origin of COVID-19 was not reported but throughdifferent researches, it was seen that coronavirus possibly has originated from the bat. According to the Centers forDisease Control and Prevention (CDC), the novel disease COVID-19 was transmitted from person to person throughdroplets,and thesymptomsseen werefever,shortnessofbreath,and cough which wasseenafter 14 days[9].

The International Committee on Virus Taxonomy replaced the name of 2019-nCoV as SARS-CoV-2 (severe acuterespiratory coronavirus-2 syndrome) [10]. The outbreak of novel SARS-CoV-2 was increasing at an alarming rate inChina as global intimidating as pandemic throughout the World. Different methods were used to analyze dataregarding epidemiology which were exploratory data analysis (EDA) methods and visualization model. These twomethods showed the awareness among the communities and were noticed according to the data analysis that thegovernment, health workers and the public have to cooperate throughout the World to prevent the spreading of theCOVID-19[11].

Data was collected from different sources from different countries regarding COVID-19 [12]. The maximum datarelated to COVID-19 was available at Google, WHO, CDC, ECDC, NHC of the PRC, JHU CSSE, DXY, QQwebsites [13]. With the help of these data from different sources helped to analyze the people getting affected byCOVID-19 and the rate of recovery and deaths were also analyzed daily. The dataset was recorded from 22 January2020. Dataset was analyzed to record the death, survival, recovery, and people who were affected by COVID-19.The very first data suggested that males of age above years were at higher risk of infection with COVID-19.According to two other well-known diseases which are caused by coronavirus i.e SARS and MERS (Middle Eastrespiratory syndrome) the reproduction rate of the virus was 2.5 to 3.0 days in the early stage of the outbreak. Afterthereproductionrate,itwasreported thatthe doublingtimeoftheviruswas7.5days[14].

1. METHODSofDATAANALYSIS

Thedatasetwasretrievedfromdifferentsourceswhichwerefurtherusedforanalysisandvisualizationmethods.The methods which were used for analysis were able to track the spreading up of COVID-19 throughout the World.The data included the number of confirmed cases, recovered rate, and death rate in different countries. The datasetcanbe seeninFig.1asupdatedon 17thMay2020[15].

5000000

4500000

4000000

3500000

3000000

2500000

2000000

1500000

1000000

500000

0

Totalconfirmed

TotalrecoveredTotaldeaths

Feb,2020 March,2020April, 2020 May,2020

Fig.1:Outbreaktrendsovertime

1. *Data Source*

Data related to COVID-19 was retrieved from verified sources like Google, WHO DingXiangYuan, a website that isauthorized by the Chinese government [16]. These sites provide information about the confirmed COVID-19 cases,the number of people recovered from the disease and the number of deaths that took place by infection of the virus[17].

1. *DataVisualization*

The retrieved data from different sites can be used to track the status of the corona [16]. The data collected from adifferent source can be seen in the table 1 [15]. The updates from the different countries can be seen through thedifferentcountries' COVID-19portalor WHO[18].

1. *ExploratoryDataAnalysis(EDA)*

Exploratory Data Analysis is used to analyze data and visualize the dataset provided by different sources regardingthe emergence of the disease. The exploratory data analysis was used to record the dataset of the outbreak ofCOVID-19 throughout the World [19]. The first dataset was visualized and analyzed between 22 January 2020 to 10March 2020. It was seen that the rate of people affected by COVID-19 was more from China than the rest of theWorld, the few affected countries were neighbors of China. After 10 March 2020 more than 30 countries and 32states in China were affected by COVID-19 [20]. Outside China not many deaths were reported, only ten deathreports were noticed until 11 March 2020. It was noticed that the rate of recovery was more than the rate of deathsand by 15 May 2020, more than 200 countries were affected by the corona. Table I shows the number of confirmedcases,thenumberofdeathsandthenumberofsurvivaltill17thMay2020[15].

TABLEI

Reportedcasesasof17thMay2020[15]

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Sr.No.** | **Country** | **Totalconfirmed** | **TotalRecovered** | **TotalDeaths** |
| 1 | USA | 1,510,286 | 339,578 | 90,178 |
| 2 | Russia | 281,752 | 67,373 | 2,631 |
| 3 | Spain | 277,719 | 195,945 | 27,650 |
| 4 | UK | 240,161 | 135 | 34,466 |
| 5 | Brazil | 233,511 | 89,672 | 15,662 |
| 6 | Italy | 224,760 | 122,810 | 31,763 |
| 7 | France | 179,365 | 61,066 | 27,625 |
| 8 | Germany | 176,450 | 153,400 | 8,027 |
| 9 | Turkey | 148,067 | 108,137 | 4,096 |
| 10 | Iran | 120,198 | 94,464 | 6,988 |
| 11 | India | 92,239 | 35,603 | 2,911 |
| 12 | Peru | 88,541 | 28,272 | 2,523 |
| 13 | China | 82,947 | 78,227 | 4,633 |
| 14 | Canada | 75,864 | 37,819 | 5,679 |
| 15 | Belgium | 55,280 | 14,630 | 9,052 |
| 16 | SaudiArabia | 54,752 | 25,722 | 312 |
| 17 | Mexico | 47,144 | 31,848 | 5,045 |
| 18 | Netherlands | 43,995 | 3 | 5,680 |
| 19 | Chile | 41,428 | 18,014 | 421 |
| 20 | Pakistan | 40,151 | 11,341 | 873 |
| 21 | Ecuador | 32,763 | 3,433 | 2,688 |
| 22 | Qatar | 32,604 | 4,370 | 15 |
| 23 | Switzerland | 30,587 | 27,400 | 1,881 |
| 24 | Sweden | 30,143 | 4,971 | 3,679 |
| 25 | Belarus | 29,650 | 9,932 | 165 |
| 26 | Portugal | 29,036 | 4,636 | 1,218 |
| 27 | Singapore | 28,038 | 8,342 | 22 |
| 28 | Ireland | 24,048 | 19,470 | 1,533 |
| 29 | UAE | 23,358 | 8,512 | 220 |
| 30 | Bangladesh | 22,268 | 4,373 | 328 |
| 31 | Poland | 18,394 | 7,451 | 919 |
| 32 | Ukraine | 18,291 | 5,116 | 514 |
| 33 | Indonesia | 17,514 | 4,129 | 1,148 |
| 34 | Romania | 16,871 | 9,890 | 1,104 |
| 35 | Israel | 16,607 | 12,884 | 271 |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 36 | Austria | 16,242 | 14,563 | 629 |
| 37 | Japan | 16,237 | 10,338 | 725 |
| 38 | Colombia | 14,939 | 3,587 | 562 |
| 39 | Kuwait | 14,850 | 4,093 | 112 |
| 40 | SouthAfrica | 14,355 | 6,478 | 261 |
| 41 | Philippines | 12,513 | 2,635 | 824 |
| 42 | DominicanRepublic | 12,110 | 3,726 | 428 |
| 43 | Egypt | 11,719 | 2,950 | 612 |
| 44 | S.Korea | 11,050 | 9,888 | 262 |
| 45 | Denmark | 10,927 | 9,227 | 547 |
| 46 | Serbia | 10,610 | 4,713 | 230 |
| 47 | Panama | 9,449 | 6,080 | 269 |
| 48 | Czechia | 8,457 | 5,422 | 297 |
| 49 | Norway | 8,244 | 32 | 232 |
| 50 | Argentina | 7,805 | 2,569 | 366 |
| 51 | Australia | 7,045 | 6,367 | 98 |
| 52 | Bahrain | 6,930 | 2,774 | 12 |
| 53 | Malaysia | 6,894 | 5,571 | 113 |
| 54 | Algeria | 6,821 | 3,409 | 542 |
| 55 | Morocco | 6,798 | 3,645 | 192 |
| 56 | Afghanistan | 6,664 | 778 | 169 |
| 57 | Finland | 6,347 | 5,000 | 298 |
| 58 | Kazakhstan | 6,157 | 3,090 | 34 |
| 59 | Moldova | 5,934 | 2,344 | 207 |
| 60 | Ghana | 5,735 | 1,754 | 29 |
| 61 | Nigeria | 5,621 | 1,472 | 176 |
| 62 | Oman | 5,186 | 1,465 | 22 |
| 63 | Armenia | 4,472 | 1,925 | 60 |
| 64 | Luxembourg | 3,930 | 3,699 | 104 |
| 65 | Bolivia | 3,826 | 473 | 165 |
| 66 | Hungary | 3,509 | 1,396 | 451 |
| 67 | Azerbaijan | 3,274 | 2,015 | 39 |
| 68 | Iraq | 3,260 | 2,126 | 121 |
| 69 | Cameroon | 3,105 | 1,567 | 140 |
| 70 | Thailand | 3,028 | 2,856 | 56 |
| 71 | Greece | 2,819 | 1,374 | 162 |
| 72 | Uzbekistan | 2,741 | 2,213 | 11 |
| 73 | Guinea | 2,658 | 1,133 | 16 |
| 74 | Honduras | 2,565 | 278 | 138 |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 75 | Senegal | 2,480 | 973 | 25 |
| 76 | BosniaandHerzegovina | 2,290 | 1,436 | 133 |
| 77 | Sudan | 2,289 | 222 | 97 |
| 78 | Croatia | 2,226 | 1,936 | 95 |
| 79 | Bulgaria | 2,211 | 598 | 108 |
| 80 | IvoryCoast | 2,061 | 987 | 25 |
| 81 | Cuba | 1,872 | 1,495 | 79 |
| 82 | Iceland | 1,802 | 1,786 | 10 |
| 83 | NorthMacedonia | 1,792 | 1,293 | 101 |
| 84 | Estonia | 1,774 | 938 | 63 |
| 85 | Guatemala | 1,763 | 138 | 33 |
| 86 | Lithuania | 1,541 | 997 | 56 |
| 87 | NewZealand | 1,499 | 1,433 | 21 |
| 88 | Slovakia | 1,494 | 1,163 | 28 |
| 89 | Slovenia | 1,466 | 273 | 104 |
| 90 | DRC | 1,455 | 270 | 61 |
| 91 | Somalia | 1,357 | 148 | 55 |
| 92 | El Salvador | 1,338 | 462 | 27 |
| 93 | Djibouti | 1,331 | 950 | 4 |
| 94 | Tajikistan | 1,322 | 0 | 36 |
| 95 | Gabon | 1,320 | 244 | 11 |
| 96 | Mayotte | 1,312 | 627 | 18 |
| 97 | Kyrgyzstan | 1,138 | 804 | 14 |
| 98 | Maldives | 1,078 | 58 | 4 |
| 99 | HongKong | 1,056 | 1,024 | 4 |
| 100 | Tunisia | 1,037 | 807 | 45 |
| 101 | Latvia | 1,008 | 662 | 19 |
| 102 | Guinea-Bissau | 969 | 26 | 4 |
| 103 | SriLanka | 964 | 538 | 9 |
| 104 | Albania | 946 | 715 | 31 |
| 105 | Cyprus | 914 | 515 | 17 |
| 106 | Lebanon | 911 | 247 | 26 |
| 107 | Niger | 889 | 689 | 51 |
| 108 | Kenya | 887 | 301 | 50 |
| 109 | Mali | 860 | 494 | 52 |
| 110 | CostaRica | 853 | 551 | 10 |
| 111 | BurkinaFaso | 782 | 604 | 51 |
| 112 | Paraguay | 778 | 198 | 11 |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 113 | Andorra | 761 | 615 | 51 |
| 114 | Zambia | 753 | 188 | 7 |
| 115 | Uruguay | 733 | 558 | 19 |
| 116 | DiamondPrincess\* | 712 | 651 | 13 |
| 117 | Georgia | 695 | 425 | 12 |
| 118 | SanMarino | 654 | 201 | 41 |
| 119 | Jordan | 607 | 404 | 9 |
| 120 | EquatorialGuinea | 594 | 22 | 7 |
| 121 | Malta | 553 | 454 | 6 |
| 122 | Jamaica | 517 | 121 | 9 |
| 123 | Tanzania | 509 | 183 | 21 |
| 124 | SierraLeone | 505 | 141 | 32 |
| 125 | Venezuela | 504 | 241 | 10 |
| 126 | Chad | 474 | 111 | 50 |
| 127 | Réunion | 443 | 354 | 0 |
| 128 | Taiwan | 440 | 395 | 7 |
| 129 | Palestine | 381 | 335 | 2 |
| 130 | Haiti | 358 | 29 | 20 |
| 131 | Benin | 339 | 83 | 2 |
| 132 | IsleofMan | 335 | 285 | 24 |
| 133 | Mauritius | 332 | 322 | 10 |
| 134 | CaboVerde | 328 | 84 | 3 |
| 135 | CAR | 327 | 13 | 0 |
| 136 | Montenegro | 324 | 311 | 9 |
| 137 | Vietnam | 320 | 260 | 0 |
| 138 | Ethiopia | 317 | 113 | 5 |
| 139 | Madagascar | 304 | 114 | 1 |
| 140 | Togo | 298 | 99 | 11 |
| 141 | Nepal | 295 | 36 | 2 |
| 142 | Rwanda | 289 | 178 | 0 |
| 143 | SaoTomeandPrincipe | 235 | 4 | 7 |
| 144 | Uganda | 227 | 63 | 0 |
| 145 | Liberia | 226 | 120 | 21 |
| 146 | Eswatini | 202 | 72 | 2 |
| 147 | FrenchGuiana | 197 | 125 | 1 |
| 148 | Martinique | 192 | 91 | 14 |
| 149 | FaeroeIslands | 187 | 187 | 0 |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 150 | Myanmar | 182 | 96 | 6 |
| 151 | Guadeloupe | 155 | 109 | 13 |
| 152 | Gibraltar | 147 | 145 | 0 |
| 153 | Brunei | 141 | 136 | 1 |
| 154 | Mozambique | 137 | 43 | 0 |
| 155 | Mongolia | 136 | 21 | 0 |
| 156 | Bermuda | 123 | 73 | 9 |
| 157 | Cambodia | 122 | 122 | 0 |
| 158 | Yemen | 122 | 1 | 18 |
| 159 | Guyana | 117 | 43 | 10 |
| 160 | TrinidadandTobago | 116 | 107 | 8 |
| 161 | Aruba | 101 | 93 | 3 |
| 162 | Monaco | 96 | 87 | 4 |
| 163 | Bahamas | 96 | 42 | 11 |
| 164 | CaymanIslands | 94 | 55 | 1 |
| 165 | Barbados | 86 | 67 | 7 |
| 166 | Liechtenstein | 82 | 55 | 1 |
| 167 | Libya | 65 | 28 | 3 |
| 168 | Malawi | 65 | 24 | 3 |
| 169 | FrenchPolynesia | 60 | 59 | 0 |
| 170 | Syria | 51 | 36 | 3 |
| 171 | Angola | 48 | 17 | 2 |
| 172 | Macao | 45 | 43 | 0 |
| 173 | Zimbabwe | 44 | 17 | 4 |
| 174 | Mauritania | 40 | 7 | 4 |
| 175 | Eritrea | 39 | 39 | 0 |
| 176 | SaintMartin | 39 | 30 | 3 |
| 177 | PuertoRico | 39 | 1 | 2 |
| 178 | Guam | 32 | 0 | 1 |
| 179 | Antiguaand Barbuda | 25 | 19 | 3 |
| 180 | Nicaragua | 25 | 7 | 8 |
| 181 | Timor-Leste | 24 | 24 | 0 |
| 182 | Botswana | 24 | 17 | 1 |
| 183 | Gambia | 23 | 12 | 1 |
| 184 | Grenada | 22 | 14 | 0 |
| 185 | Bhutan | 21 | 5 | 0 |
| 186 | Laos | 19 | 14 | 0 |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 187 | NewCaledonia | 18 | 18 | 0 |
| 188 | SaintLucia | 18 | 18 | 0 |
| 189 | Belize | 18 | 16 | 2 |
| 190 | Fiji | 18 | 15 | 0 |
| 191 | St.VincentGrenadines | 17 | 14 | 0 |
| 192 | U.S.VirginIslands | 17 | 0 | 0 |
| 193 | Dominica | 16 | 15 | 0 |
| 194 | Curaçao | 16 | 14 | 1 |
| 195 | Namibia | 16 | 13 | 0 |
| 196 | SaintKittsandNevis | 15 | 14 | 0 |
| 197 | Burundi | 15 | 7 | 1 |
| 198 | Turksand Caicos | 12 | 10 | 1 |
| 199 | VaticanCity | 12 | 2 | 0 |
| 200 | Greenland | 11 | 11 | 0 |
| 201 | Montserrat | 11 | 10 | 1 |
| 202 | Seychelles | 11 | 10 | 0 |
| 203 | Comoros | 11 | 3 | 1 |
| 204 | Suriname | 10 | 9 | 1 |
| 205 | PapuaNewGuinea | 8 | 8 | 0 |
| 206 | BritishVirginIslands | 8 | 6 | 1 |
| 207 | WesternSahara | 6 | 6 | 0 |
| 208 | St.Barth | 6 | 6 | 0 |
| 209 | Anguilla | 3 | 3 | 0 |
| 210 | Kosovo | 2 | 0 | 0 |
| 211 | Lesotho | 1 | 0 | 0 |

1. *VisualExploratoryDataAnalysis(EDA)*

Visual Exploratory Data Analysis (EDA) is a method used to analyze the rate at which COVID-19 was spreadingthroughout the globe. In this method, the data was analyzed through a map that helps an individual to understand theepidemiological nature of COVID-19 as shown in Fig. 2. According to the data, it was noticed that China reportedthe highest rate of cases confirmed with COVID-19 and the highest death rate by the virus (Till 17 March 2020)followed by Italy [21]. EDA provides a piece of good knowledge about the time taken by the virus to spreadthroughout the globe. The data analysis through EDA is also useful in analyzing the behavior of the disease. EDAhelpsinunderstandingthesituationoftheCOVID-19.ThedataforCOVID-19isavailableatURL

<http://samratdey.me/visualization.html>and by 15 May 2020, it was seen that highest number of corona cases werereportedbyUSAandRussia followedbySpain[15].

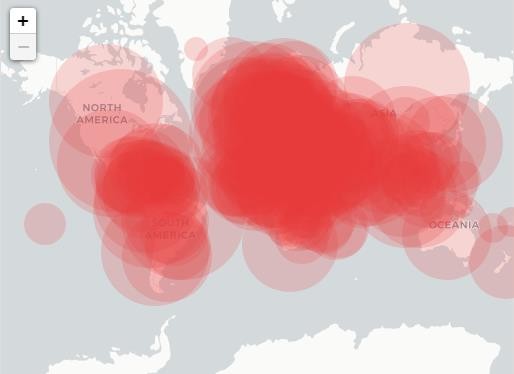


Fig.2**:**Worldmapofaffectedregion,wherethedarkerredregionsinthemappredictnumberofinfectedcases[15]

1. *PredictiveModeling –SEIRModel*

Susceptible-Exposed-Infectious-Recovered (SEIR) model is used to predict the time and the rate taken for thespreading up of disease throughout the globe. In this modeling method, real-time data is collected and visualized toforecast the rate of increasing cases for COVID-19 [22]. SEIR model predicts according to the previous dataprovided to forecast the number of cases that may take place in the future, it also predicts the death rate that mayoccur in the future because of COVID-19. SEIR model is designed to analyze and classify the news into positive andnegative sentiments [31]. The result of news on the behavior of peoples both economically and politically. Theproperties of Susceptible-Exposed-Infected- Removed (SEIR) system is used to study the outbreak of COVID-19throughout the World [23]. SEIR is considered to be the model for simulation studies for the disease spreading,where parameters are Susceptible (S), Exposed (E), Infections (I) and Recovered (R). In Susceptible (S), people mayor may not have infection were considered, in Exposed (E), people who were incubated after encountering of thevirus, in Infections (I)people afterincubationshowing symptomswerekept andinRecovered(R)parameteritreferstothestatewherenooneisinfectedwiththe diseaseor disease-freepeople[30].

TheparametersinSEIRmodelsareasfollows:-

1. Betaisrepresentedfortherateofspread,whichistherateatwhichdiseaseistransmittedbetweenasusceptibleand infectiousperson.
2. The incubation rate is given by sigma, in which incubation to the individual is given and the rate is being recordedin whichanindividualwillgetinfected.The durationofincubationgivenis5days.
3. Recoveryrateisgivenbygamma, inthis,theaveragedurationoftherecoveryfromtheinfectionisrecorded.
4. Aftertherecoveryphase,recoveredcandidatesarekeptundertheremovedphase[24]asshowninFig.3.

Susceptible(S)

Beta

Exposed(E)

sigma

Infections(I)

gamma

Recovered(R)

**Fig.3**:SEIRModel

1. *SentimentAnalysis*

Sentiment analysis is done to keep a record of data which is neither too long nor too short and it is the result of theSEIR predictive model [25]. Sentiment analysis consists of a summary containing a description of more than eightwordsofthe trainedmodel[26][27].

1. *Statisticalchallengesofanalyzing COVID-19data*

After the outbreak of COVID-19 in Wuhan, China, the statistical model plays a major role in comparing the numberofconfirmedCOVID-19cases,thenumberofrecoveriesandthenumberofdeathratethatistakingplacethroughout the globe as shown in Fig. 4, Fig. 5 and Fig. 6 respectively. The statistical model compares the data fromorigin i.e China to the data of different countries with respect to time in the form of a bar graph. The data fromdifferent countries of the confirmed cases are recorded from the very start of the outbreak of the disease. Theseparate data is maintained in a statistical model for thecases which are recovered from the infection and thenumber of deaths caused by COVID-19. The two protocols are maintained under in which closed COVID-19 casesarerecordedwhichare asfollows:-

* 1. InternationalSevereAcuteRespiratoryandEmergingInfectionConsortium(ISARIC)(isaric.tghn.org)
  2. LeanEuropeanOpenSurveyonSARS-CoV-2InfectedPatients(LEOSS)(leoss.net).

For COVID-19 patients, the most important clinical endpoints are the record of intensive care, invasive ventilation,andsurvival.Theless relevant endpoint is supportiveoxygen.According tothese twoendpoints datacan beanalyzed on a statistical model that will be dependent on time. The data is further collected from ISARIC andLEOSStoanalyze data inthe standardprotocol[28].



**NumberofConfirmedCases**

16,00,000

14,00,000

12,00,000

10,00,000

8,00,000

6,00,000

4,00,000

2,00,000

0

NumberofConfirmed

Cases

Fig.4:MostaffectedcountriesshowingnumberofconfirmedCOVID-19cases.



**NumberofRecoverdcases**

4,00,000

3,50,000

3,00,000

2,50,000

2,00,000

1,50,000

1,00,000

50,000

0

NumberofRecoverd

cases

Fig.5:MostaffectedcountriesshowingnumberofrecoveredCOVID-19cases.



**NumberofDeaths**

1,00,000

90,000

80,000

70,000

60,000

50,000

40,000

30,000

20,000

10,000

0

NumberofDeaths

Fig.6:MostaffectedcountriesshowingnumberofrecoveredCOVID-19cases.

1. CONCLUSION

COVID-19 outbreakwhich took place in Chinawas recorded and visualized throughdifferent online platforms.Data analysis was done through several methods. Exploratory Data Analysis was used to analyze data and visualizethe dataset provided by different sources regarding the emergence of the disease. Visual Exploratory Data Analysis(EDA) was used as a method to analyze the rate at which COVID-19 was spreading throughout the globe. In thismethod, the data was analyzed through a map that helps an individual to understand the epidemiological nature ofCOVID-19. Susceptible-Exposed-Infectious-Recovered (SEIR)modelwas used to predict the time and the ratetaken for the spreading up of disease throughout the globe. In this modeling method, real-time datawas collectedand visualized to forecast the rate of increasing cases for COVID-19 and was also used to forecast the analysis ofinfection. The results from the SEIR model were further used to analyze data for sentiment analysis among thecommunity regarding the outbreak of COVID-19. The COVID-19 outbreak spreads not only through the country’spolicy but also through the social responsibility of each individual.The different online platform, updates theviewers with the situation of the disease including the number of confirmed cases, number of recoveries and numberof deaths taking place throughout the world. Data analysis for COVID-19 is done to make aware humans against theinfectioncausedbyCorona.

ACKNOWLEDGEMENT

I wouldliketothanksDr.AbhimanyuKumarJha forhiscontinuoussupportthroughouttheresearch.

**REFERENCES**

1. ZhuN,ZhangD,WangW,etal,“AnovelcoronavirusfrompatientswithpneumoniainChina”,*NEnglJMed*,https://doi.org/10.1056/NEJMoa2001017,2019.
2. CentersforDiseaseControlandPrevention,“NovelCoronavirus(2019-nCoV),Wuhan,China”,https://[www.cdc.gov/coronavirus/2019-](http://www.cdc.gov/coronavirus/2019-)nCoV/summary.html, 2019.
3. ChenY,LiuQ,GuoD,“Emergingcoronaviruses:genomestructure,replication,andpathogenesis,”*JMedVirol*.https://doi.org/10.1002/jmv.25681,2020.