International Journal of Applied Information Systems (IJAIS) – ISSN : 2249-0868
Foundation of Computer Science FCS, New York, USA
Volume 12– No.32, August 2020 – www.ijais.org

Data Analytics and Visualization of Covid-19 Epidemic in Nigeria based on Recovered and Death Cases

Ajinaja Micheal
Department of Computer Science
Federal Polytechnic, Ile-Oluji
Ondo State, Nigeria

Olarinde Mobolaji
Department of Computer Science
Federal Polytechnic, Ile-Oluji
Ondo State, Nigeria

Olaniyan Olabisi
Department of Computer Science
Federal Polytechnic, Ile-Oluji
Ondo State, Nigeria

Adeola Ademola
Department of Computer Science
Federal Polytechnic, Ile-Oluji
Ondo State, Nigeria

ABSTRACT
The COVID-19 deadly virus was first reported in Wuhan, China and so far gone on to spread to more than 50 countries across the globe. WHO declared COVID-19 as a pandemic on 11th March 2020 and there has been rising cases of infection of the virus. This study present data analysis and visualization of the confirmed, recovered, active and death cases in Nigeria. Cases was also compared to global active, confirmed, recovered and death cases. Different line graph analysis, dot graph analysis and the use of SIR model to further provide diagrammatical view and analysis of the cases in Nigeria. The model and analysis will help to interpret patterns of the virus and assist in further research and provide relevant information to the public concerning the virus.

Keywords
COVID-19, data analytics, visualization, SIR

1. INTRODUCTION
The first reported case of the deadly covid-19 virus was in Wuhan, China on 31st December 2019 which as at that time was gaining grounds and growth steadily increasing. The World Health Organization (WHO) announced Covid-19 outbreak as a pandemic officially on 11th March, 2020 an since then the virus has continually infected more people around the globe. The first case outside of China was reported in Thailand on 13th January 2020 (Hui et al, 2020). No country has been spared as the virus continually ravages and cripples many nations economy around the world. The Covid-19 pandemic is a rapidly evolving and information about the virus is still emerging. The virus is primarily transmitted from person-to-person by coming into contact with an infected person's droplets (IAMAT, 2020). Different measures have been listed to halt the spread but the virus keeps on increasing in numbers.

As in the case in Nigeria, the federal ministry of health confirmed the first coronavirus disease case in Lagos state, Nigeria on 27th February 2020. The case was of an Italian citizen who works in Nigeria and returned from Milan, Italy to Lagos on 25th February 2020. The second case was detected on 9th March 2020 in Ewekoro, Ogun state a Nigerian citizen who had contact with the first case – the Italian citizen. Since that period, the country has withnessed an exponential increase in the number of confirmed Covid-19 cases and government finding it difficult to manage the infectious disease outbreak. An infectious disease outbreak is the occurrence of a disease that is not usually expected in a particular community, geographical region, or time period (Toppenberg et al, 2018). Typically, a rising infectious disease involves fast spreading, endangering the health of large numbers of people, and thus requires immediate action to prevent the disease at the community level (Lin et al, 2016).

In order to prevent the spread of this infection, World Health Organization (WHO) recommends frequent hand washing keeping unwashed hands away from the face, social distancing, and covering coughs and sneezes with a tissue or inner elbow. Some national health authorities recommend masks for the suspects and their caregivers. So far no vaccine or antiviral treatment is available for COVID-19 (Araujo & Naimi, 2020). This paper provides data analytics and modelling COVID-19 cases in Nigeria based on recovered and death cases currently available data as at time of this writing. Data analytics is provided first on the top ten total number of confirmed cases in the world, growth rate in Nigeria, total cases in Nigeria and the use of SIR model on the data. This study can give researchers guide on how this data can be used to prevent future spread of the disease in Nigeria.

2. RELATED WORKS
Numerous authors have worked on Covid-19 since the pandemic started. Researchers have started to extract information about the infected cases and analyzed the biomedical information and their medical histories to extract the main parameters that could cause coronavirus spreading. Researchers suggested that the spreading of coronavirus could be connected with sex, birth year, or the region they come from (AL-Rousan & AL-Najjar, 2020). SIR model is an epidemiological model that computes the theoretical number of people infected with a contagious illness in a closed population over time. Susceptible refers to individuals who can catch the infection and may become hosts if exposed, Exposed are individuals who are already infected but are asymptomatic, Infectious are individuals who are showing signs of infection and can transmit the virus, Removed or Recovered are individuals who are previously infected but are no longer infectious and already immune to the virus (Rachah & Torres, 2018).

Hamzaha et al 2020 worked on CoronaTracker: World-wide COVID-19 Outbreak Data Analysis and Prediction. Real-time data query was done and visualized and the queried data is used for Susceptible-Exposed-Infectious-Recovered (SEIR) predictive modelling. SEIR model was utilized to forecast COVID-19 outbreak within and outside of China based on daily observations (Hamzaha et al, 2020). Another similar work done was on data visualization and analyzation of COVID-19. The study focused on COVID-19 based on freely available datasets in which data analytics was provided on a number of aspects of COVID-19 including the symptoms of this disease, the difference of COVID-19 with other diseases caused by severe acute respiratory syndrome (SARS), Middle East respiratory syndrome (MERS), and swine flu (Khanam et al, 2020). Data visualization was provided on the comparison of infections in males/females which shows that males are more prone to this disease and the older people are more at risk. Based on the data, the pattern in the increase of confirmed cases is found to be an exponential curve in nature. Also, the relative number of confirmed, recovered and death cases in different countries were shown with data visualization in the study. S.C Gamoura, 2020 worked on Real-time data analytics and prediction of the covid-19 pandemic from February to March 26th, 2020 which visualizes current covid-19 cases in real time (Gamoura, 2020).

Mudr discussed the use of data analytics during the COVID-19 Pandemic in which he stated the number of challenges to the success posed by covid-19 to the success of clinical trials. (Mudr, 2020). Mohamed and Abderrahmane worked on predicting the covid-19 epidemic in Algeria using the SIR model. The study aimed in predicting the daily infected cases with Coronavirus (COVID-19) in Algeria. SIR model was applied on data from 25 February 2020 to 24 April 2020 for the prediction. Based on the simulation of the two models, the epidemic peak of COVID-19 was predicted to be attained on 24th July 2020 in a worst-case scenario, and the COVID-19 disease was expected to disappear in the period between September 2020 and November 2020 at the latest (Mohamed & Abderrahmane, 2020).

3. METHODOLOGY

SIR model is an epidemiological model that computes the theoretical number of people infected with a contagious illness in a closed population over time. For this study, the R-studio software was used in implementing the SIR model and also producing other data visualisation and analysis. Susceptible refers to individuals who can catch the infection and may become hosts if exposed, Exposed are individuals who are already infected but are asymptomatic, Infectious are individuals who are showing signs of infection and can transmit the virus, Removed or Recovered are individuals who are previously infected but are no longer infectious and already immune to the virus (Rachah & Torres, 2018). The model consists of three compartments:

S: The number of susceptible individuals. When a susceptible and an infectious individual come into "infectious contact", the susceptible individual contracts the disease and transitions to the infectious compartment.

I: The number of infectious individuals. These are individuals who have been infected and are capable of infecting susceptible individuals.

R: for the number of removed (and immune) or deceased individuals. These are individuals who have been infected and have either recovered from the disease and entered the removed compartment, or died. It is assumed that the number of deaths is negligible with respect to the total population.

The SIR model without demographics is given by the following non-linear system of ordinary differential equations (Harko et al, 2014);

$$\frac{dS(t)}{dt} = \beta S(t)I(t)$$

$$\frac{dI(t)}{dt} = \beta S(t)I(t) - \lambda I(t)$$

$$\frac{dR(t)}{d(t)} = \lambda I(t)$$

The function S(t) represents the number of susceptible cases at time t. The function I(t) represents the number of infected cases at time t, and the function R(t) is the compartment of removed cases from the disease (Kermack & McKendrick, 1927) where:

Removed cases = Death cases + Recovered cases.

The parameters β and λ are called the contact rate and the removed rate, respectively. The initial condition of this model S(0), I(0), R(0) must satisfy the condition S(0) + I(0) + R(0) = N, where N is the population size (Harko et al, 2014).

The model can be written also as follows:

$$\frac{dS(t)}{dt} = -\beta \frac{S(t)}{N} I(t),$$

$$\frac{dI(t)}{dt} = \beta S(t) \frac{I(t)}{N} - \lambda I(t)$$

$$\frac{dR(t)}{d(t)} = \lambda I(t)$$

With: $\frac{S(t) + I(t) + R(t)}{N} = 1$ and $S(0) > 0, I(0) > 0, R(0) = 0$ (Keeling et al 2008)

.....N

4. RESULTS AND DISCUSSION

After formulating the model necessary for simulating the data – the model was implemented using R-Studio Version 1.3.959 software. The Covid-19 analytics package which is among the many packages available in the R-Studio software was used in presenting the analysis and simulating the SIR model.

Figure 1 gives a plot analysis of Nigeria showing the relationship between the number of cases and date from which the first case was identified. It can be observed the steady rise of the number of cases started from the end of May. Nigeria locked down her two major states (Abuja and Lagos) for two weeks on 29th March 2020 and the virus spread was kept at minimal. The government further extended the lockdown to all states after two weeks elapsed to further halt the spread of the virus. This didn't go down well with the people as citizens continued to defy government orders thus the high infection rate as from month of May.

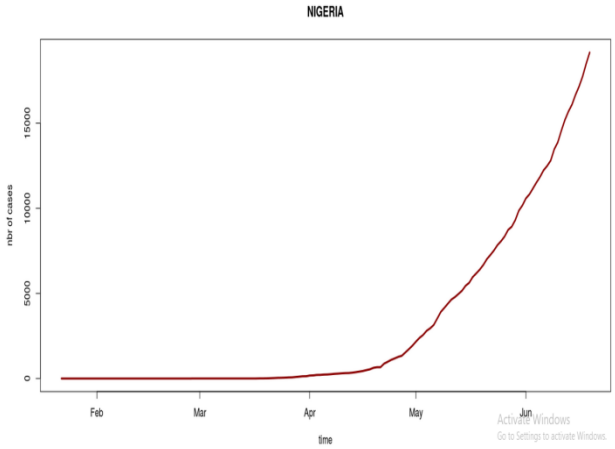

Figure 1: Plot diagram of number of cases and time

Figure 2 presents a plot diagram of the number of cases in logarithm and number of days. It is observed a sharp rise in the number of cases from day 40. Day 100 presented a steady rise in the number of cases as compared to day 40 increment.

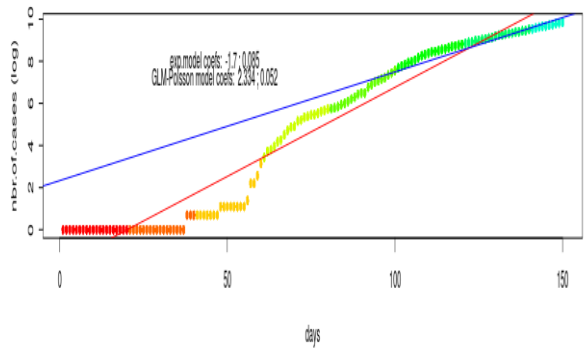

Figure 2: Plot diagram of number of cases (in log) and time

Figure 3 gives a bar diagram of steady increment of the virus in Nigeria. It is observed from the plot that there was increased covid-19 positive cases from 22nd April 2020. Cases reached more than 15,000 as from 13th June 2020 which showed peak value.

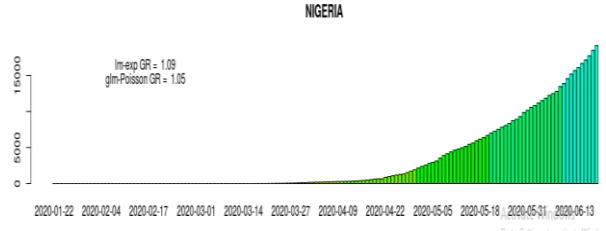

Figure 3: Plot diagram of number of cases (in log) and time

Figure 4 plot presented the number of changes in confirmed cases to the time of the month. It is observed from the plot, there was a steady difference in changes to number of confirmed cases from February to middle of March when the first case was detected. As cases gradually increased from middle of March, the difference remained almost same till middle of April when there was a sharp reduction in the difference in the number of cases. The cases picked up again towards the latter end of April and continually widened.

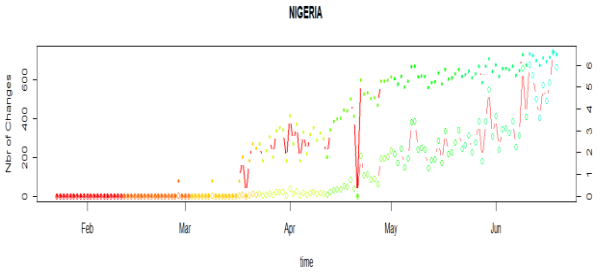

Figure 4: Plot diagram of number of changes (difference in the number of cases) and time

Figure 5 gives a bar chart of the growth rate of the virus in Nigeria. It is observed that growth rate peaked on 29th March 2020. 11th of April also witnessed peak but not as much as the previous date stated.

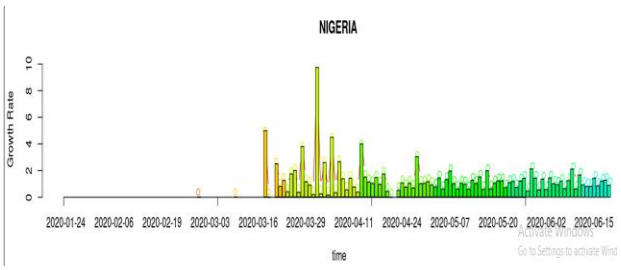

Figure 5: Plot diagram of growth rate and time

Figure 6 shows the linear graph of the number of cases across the globe to that of Nigeria. It is observed that compared to the number of cases in the world, the cases in the Nigeria are quite small and almost unnoticeable. The representation of each of the coloured (thick and thin) lines are specified in the graph. As at the time of this writing, the cases are far below 50,000 hence the line graph representation of Nigerian almost close to 0.

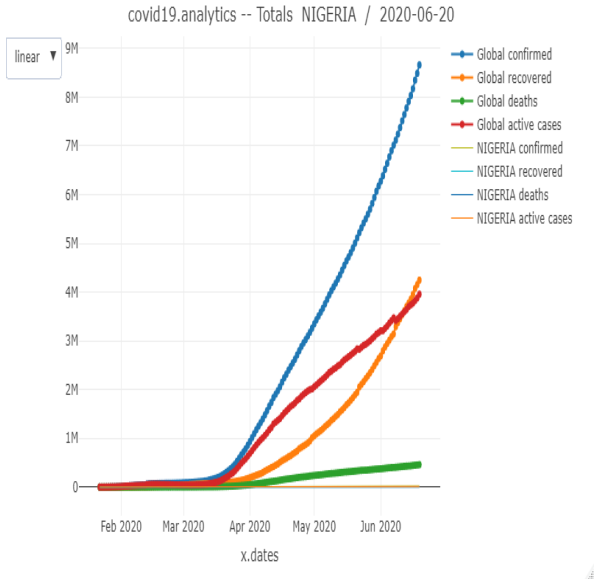

Figure 6: Line graph showing the relationship between the number of cases around the globe as compared to Nigeria and time

Figure 7 shows the log plot graph of the global confirmed cases, recovered and deaths as compared to Nigeria's confirmed, recovered, death and active cases. It is observed

that Nigeria cases is far below as compared to global cases. The representation of each of the line (thick and thin) in the graph is specified.

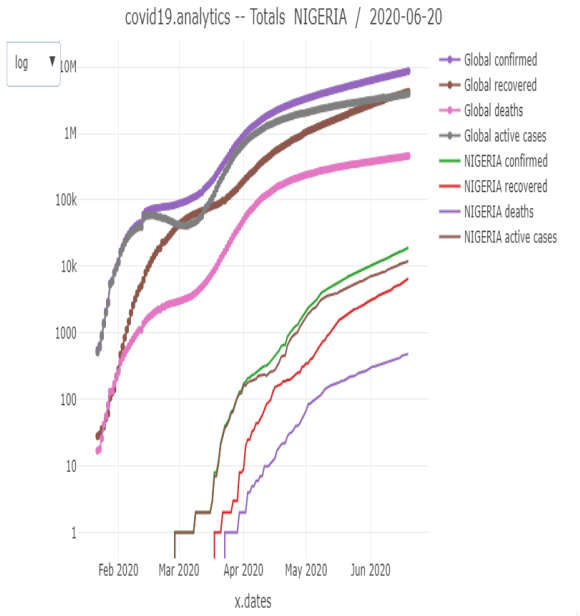

Figure 7: Log plot graph showing the relationship between global confirmed, recovered, deaths and Nigeria's confirmed, recovered and death cases.

Figure 8 present the dot line graphs of the total number of confirmed, recovered, deaths and active cases in Nigeria. It is observed from the graph the number of death cases is significantly low as compared to the confirmed and recovered cases although the number of death cases is gradually increasing but kept as the barest minimum. Confirmed cases continually increases rapidly as against the recovered, deaths and active cases.

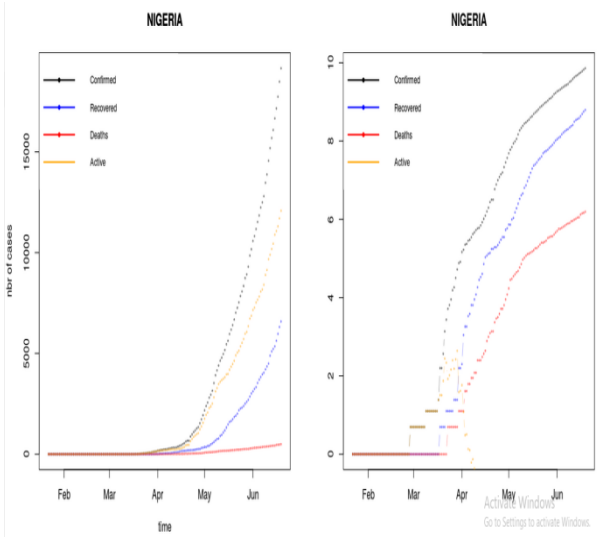

Figure 8: Dot line graph showing the relationship between the number of confirmed, recovered, death and active cases in Nigeria with respect to time

Figure 9 shows the result of the simulation of the SIR model on the covid-19 cases in Nigeria. The first two graph shows the model accuracy for the infected cases. The last graph show the summary of those who are susceptible to the virus, infected and recovered analysis.

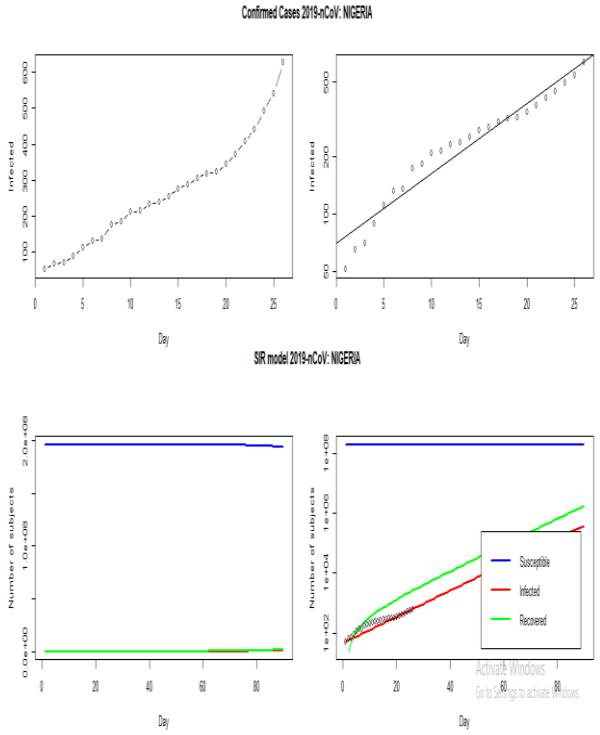

Figure 9: Line graph showing the result of the SIR model

5. CONCLUSION

The study provided data analysis and visualization summary of covid-19 cases in Nigeria based on the number of recovered, deaths, active and confirmed cases. The study presented different plots to show comparison among the attributes and global cases. SIR model was used in simulating the susceptible, infected and recovered cases available on R-studio software and the results of the model presented via the plot diagrams. It is believed that this analysis and visualizations will help government of Nigeria in knowing steps to take to curtail the spread of the virus.

6. REFERENCES

[1] A. Rachah and D. F. M. Torres, "Analysis, simulation and optimal control of a SEIR model for Ebola virus with demographic effects," Commun. Fac. Sac. Univ. Ank. Series A1, vol. 67, no. 1, pp. 179-197, 2018.

[2] Araujo M.B, Naimi B. Spread of SARSCoV-2 Coronavirus likely to be constrained by climate. Medrxiv; 2020

[3] COVID-19 Coronavirus Pandemic. Available: https://raw.githubusercontent.com/CSSEGISandData/COVID-19/master/csse_covid_19_data/csse_covid_19_daily_reports/06-18-2020.csv

[4] D. Hui et al, "The continuing 2019-nCoV epidemic threat of novel coronavirus to global health - The latest 2019 novel coronavirus outbreak in Wuhan, China," International Journal of Infectious Diseases, vol. 91, pp. 264-266, 2020

[5] D. Toppenberg-Pejcic, J. Noyes, T. Allen, N. Alexander, M. Vanderford and G. Gamhewage, "Emergency Risk

Communication: Lessons Learned from a Rapid Review of Recent Gray Literature on Ebola, Zika, and Yellow Fever," Health Communication, vol. 34, no. 4, pp. 437-455, 2018.

[6] Harko.T, Lobo. F. S. N, Mak. M. K.. Exact analytical solutions of the Susceptible-Infected- Recovered (SIR) epidemic model and of the SIR model with equal deaths and births. APPL MATH COMPUT. 2014; 236 (2014), 184–194

[7] Keeling, M. , Rohani, P. Modeling infectious diseases in humans and animals. Princeton University Press. 2008.

[8] Kermack, W.O.& McKendrick, A.G. A Contribution to the Mathematical Theory of Epidemics. Proc. Roy. Soc. Lond. 1927; A 115 700–721.

[9] L. Lin, R. McCloud, C. Bigman and K. Viswanath, "Tuning in and catching on? Examining the relationship between pandemic communication and awareness and knowledge of MERS in the USA," Journal of Public Health, p. fdw028, 2016.

[10] World Health Organization (WHO), "Coronavirus disease 2019 (COVID-19) SituationReport - 35," WHO, 2020.