Visualising and analysing data on Covid-19 pandemic outbreak in Nigeria

Franklin Ovuolelolo Okorodudu* and Onyeacholem Ifeanyi Joshua

Department of Computer Science, Delta State University Abraka Nigeria.
*Corresponding author. E-mail: okoroblackx4@yahoo.co.uk.

Accepted 3rd November, 2021

The novel corona-virus (COVID-19) that was first reported at the closing stages of 2019 has dwindled almost every aspect of life beyond reasonable doubt. This paper focuses on the incidence of the virus in the 37 states in Nigeria. The first confirmed cases were Lagos State, which was established on the 27th day of February 2020, and was the earliest case to be reported in Nigeria since the start of the outburst in China in January 2020. Using chart method, the authors will analyse the daily and increasing prevalence of COVID-19 in the 37 states. This Paper displayed and analysed the data on covid-19 pandemic outbreak in Nigeria, by using jupyter notebook which is a programming tool for visualizing, analysing and comparing the rate of affected persons in various states in Nigeria. Presently a lot of visualization and analysis has been carried out by various researchers in order to ascertain the actual number of the affected persons. For the cause of this Paper, the authors used two data analytical tools such as R and Python (computer programming) to analyse this data and make prediction using data reports from Nigeria centre for disease and control (NCDC). The result displayed an intriguing report of the virus trajectory in Nigerian state, which also gives the three tiers of government measures in tackling the virus to avoid further spread.

Key words: Visualization, big data, analytics, pandemic, COVID-19, outbreak, Jupyter notebook.

INTRODUCTION

During the COVID-19 pandemic of 2020, epidemic curves have become a commonplace in scientific and mainstream media most curves exhibit the everyday number of cases over time or the escalating figure of cases over time. These plague curves are helpful for communicating the size (increasing number of cases) and total growth rate (new cases on a daily basis) of the outbreak. Conversely, no single graphical technique can successfully convey all significant aspects of a pandemic. Classic epidemic curves (showed in natural units) contain the following shortcomings (Perneger et al., 2020). This paper made use of classic epidemic graph - counts of daily events or swelling events over time in various states in Nigeria -emphasize temporal changes in the increase or size of epidemic outbreaks.

LITERATURE REVIEW

The COVID-19 pandemic has grown to be a terror across the world. Measuring the level or impact of fear in diverse populations may help spot populations and areas in need of public health and education campaigns. The authors were concerned in the diagnostic tests developed to assess or diagnose COVID-19-related fear or phobia (Ashley Elizabeth Muller et al., 2020). For the purpose of this paper, it uses Bar plot and scatter plot to visualize the data in the various 37 states in Nigeria and to ascertain the fluctuation in the increase and decrease of number of affected persons with respect to the number of deaths recorded.

Near the beginning stage of the outbreak and figure estimated for crucial measures of the infectiousness of the disease, as well as the basic reproduction number, growth rate, and doubling

time. Estimates of the vital reproduction numeral were discovered to be bigger than 1 in both countries, with values being between 2 and 3 for Italy, and 2.5 and 4 for Spain. The analytical capability of the log-linear regression model was established to give a better fit, and simple estimate of the daily prevalence for both countries were computed.

According to Jeffrey Chu (2021), estimates were also computed, for the more active helpful reproduction number, which showed that since the first cases were confirmed in the respective countries, the harshness has by and large been decreasing. The rising fear of the on-going COVID-19 epidemic in South Africa has necessitated the application of modelling strategies to forecast the COVID-19 cases and deaths.

Short and long-term forecasts of COVID-19 cases and deaths, at the national, state and local level, are a key aspect of the strategy to handle the COVID-19 epidemic in the country (Tarylee et al., 2021). Successful and complimentary solution for data visualization application in a dispersed environment will be of great reward.

WHO (2020), affirmed that, globally new corona-viruses emerge once in a while in diverse areas, including SARS in 2002 and MERS in 2012? Several known corona-viruses are circulating in animals that have not yet infected humans.

As observation improves more corona-viruses are liable to be known. WHO (On February 19, 2021), H5 Reference the State Laboratory of Research Centre for Virology and Biotechnology (VECTOR) of the Russian Federation informed the WHO regarding cases of human illness with avian influenza A(H5N8) virus? These are the first laboratory established human cases of A (H5N8) virus infectivity to be reported in all parts of the world. The genome sequences of the viruses from a human case and from poultry are shared in GISAID, (2020) (EPI_ISL_1038924).

De Standard, 2020 number of new admissions to hospitals as a result of an established coronavirus infection increased by 17% in a week, from 3 to 9 August, to a mean of 44.7 new admissions per day. Traces of the coronavirus have been found in sewers. Aquafin is confident that they won't make them sick. 'But it gives a good picture of the spread of the virus. To administer these datasets is troublesome with the traditional method.

Furthermore, data storage, transition, visualization, penetrating, analysis, protection, data privacy violations and sharing propose diverse uphill challenges that the “Big Data Analytics” reinforces.

METHODS SPECIFICATION

Imported libraries from python into jupyter notebook: jupyter notebook is a web base integrated development environment (IDE) use in writing code for data analysis in decision making process.

```
import pandas as pd
import numpy as np
import pandas as pd
import seaborn as sns
import matplotlib.pyplot as plt
from matplotlib import style
%matplotlib inline
import re
from sklearn.linear_model import LinearRegression
from statsmodels.tsa.stattools import adfuller
```

DISCUSSION AND RESULTS

Table 1 show the list of data states with total cases, number of admitted person, and number of discharged and total number of death recorded during covid-19 outbreak in the various states in Nigeria. (Data source URL: www.ncdc.gov.ng).

Data was visualized using jupyter notebook in python programming

Using R to visualize the same data

Below are the plots showing the pool of charts displayed in scatter plot with the following sub headings: state, confirmed cases, number of recovery persons, number of deaths, number of active cases and number of tested persons.

Result interpretations on data visualization

From Figures 1 and 2, it is observed that Kogi

Table 1. List of data states with total cases, number of admitted person, and number of discharged and total number of death recorded during covid-19 outbreak in the various states in Nigeria.

	State	Confirmed cases	Numbers recoveries	Numbers of death	Active cases	Numers testing
1	Abia	1,683	1,645	22	16	21,939
2	Adamawa	1,063	274	32	757	18,586
3	Akwa Ibom	1,843	1,739	14	90	18,612
4	Anambra	1,909	1,826	19	64	24,675
5	Bauchi	1,540	1,515	17	8	24,921
6	Bayelsa	885	834	26	25	17,582
7	Benue	1,188	591	22	575	16,733
8	Borno	1,337	1,200	38	99	20,403
9	Cross River	394	367	18	9	6,968
10	Delta	2,617	1,744	71	802	36,024
11	Ebonyi	2,030	1,965	32	33	15,471
12	Edo	4,897	4,707	185	5	42,897
13	Ekiti	869	847	11	11	16,539
14	Enugu	2,292	2,013	29	250	24,757
15	FCT	19,747	19084	165	498	258,940
16	Gombe	2,034	1,986	44	4	43,480
17	Imo	1,655	1,592	37	26	35,664
18	Jigawa	527	485	16	42	17,221
19	Kaduna	9,014	8,915	65	34	86,535
20	Kano	3,944	3,811	110	23	95,071
21	Katsina	2,097	2,049	34	23	39,397
22	Kebbi	450	392	16	42	17,221
23	Kogi	5	3	2	0	17,740
24	Kwara	3,120	2,814	55	251	24,377
25	Lagos	58,119	56,990	439	690	461,747
26	Nassarawa	2,380	373	13	1,994	22,951
27	Niger	930	913	17	0	17,740
28	Ogun	4,639	4,571	49	19	71,150
29	Ondo	3,226	2,080	63	1,083	24,527
30	Osun	2,574	2,485	52	37	18,676
31	Oyo	6,840	6,506	123	211	58,709
32	Plateau	9,046	8,982	57	7	67,712
33	Rivers	7,049	6,896	101	52	175,002
34	Sokoto	775	746	28	1	18,801
35	Taraba	910	864	22	24	12,637
36	Yobe	365	307	9	49	12,765
37	Zamfara	240	221	8	11	7,848

and Zamfara state has the least number of affected persons and active cases while Lagos have the peak number of affected persons happened. From Figure 3, the following states are having the peak number of actives, Nasarawa, Lagos, Ondo, Delta, Adamawa, respectively.

From Figure 3, it is observed that Lagos, Edo, Abuja (FCT), Kano and Oyo respectively are having high numbers of death cases and that of Figure 4 showing the numbers of recovered persons. Figures 5 and 6 shows the data visualization using R.

In python programming with respect to

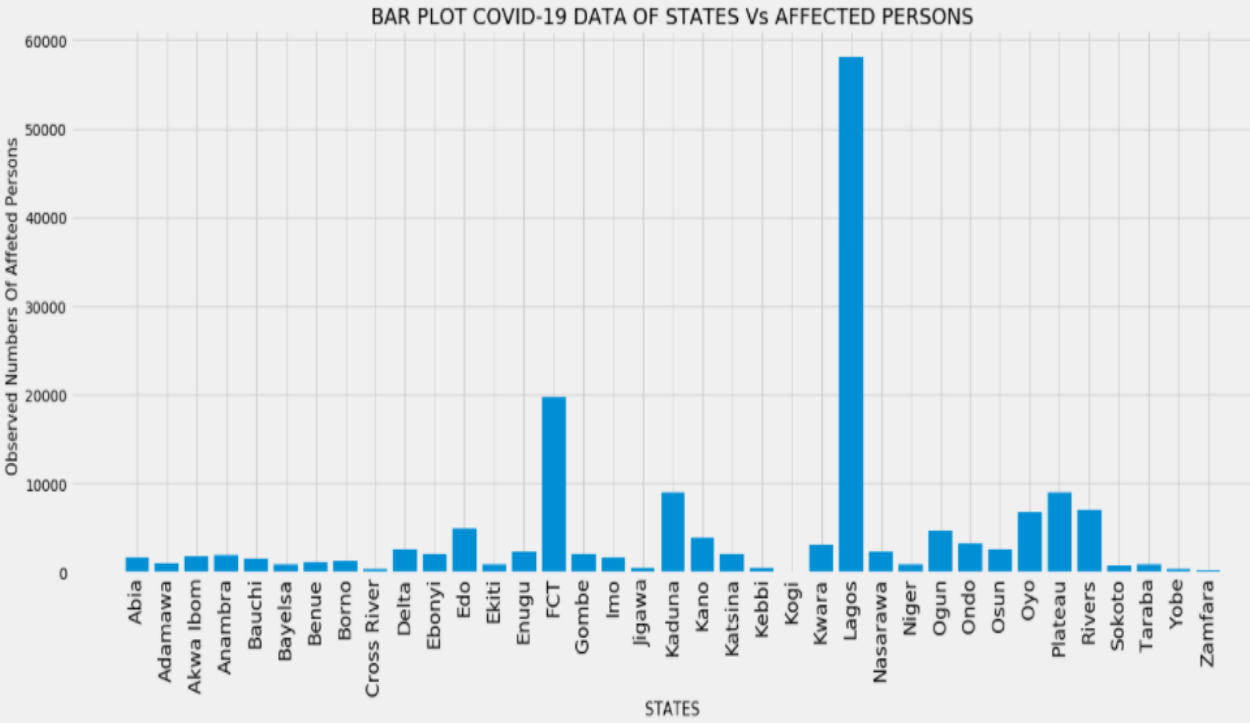

Figure 1. Number of affected persons displayed on bar chart.

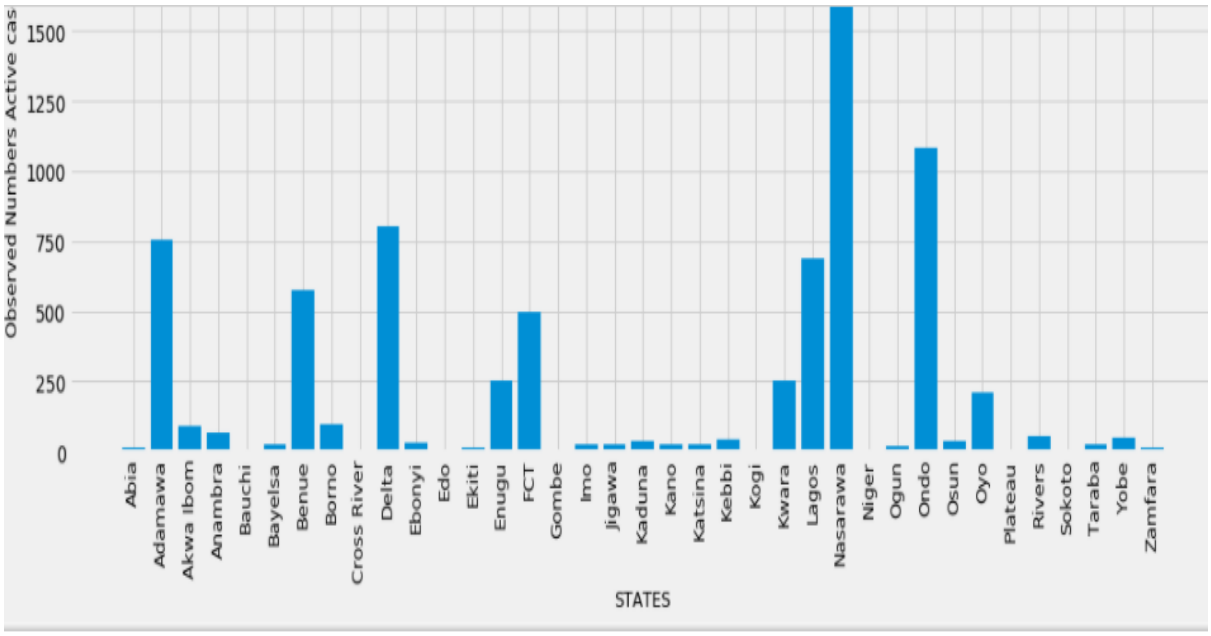

Figure 2. Bar chart showing number of active cases.

application to machine learning (ML), equation and model are not same like mathematical equations, below are the equation and linear regression model used in forecasting the figures in the table. Table 2 shows the current data figures.

Train and test data model

Train= data drop (['state','numbers_recoveries','active_cases','numbers_testing'], axis=1)

Test = data ['numbers_ deaths']

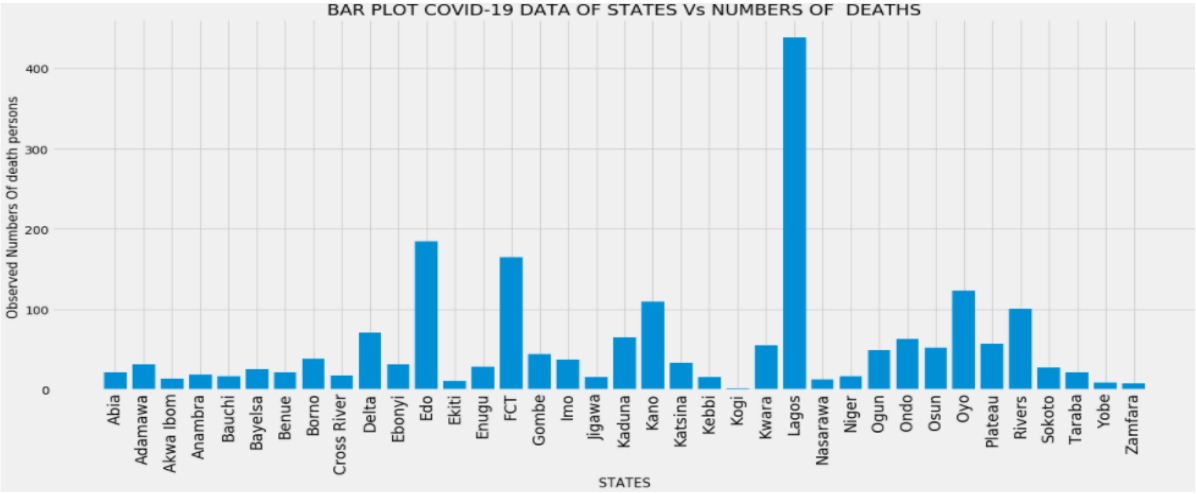

Figure 3. Covid-19 data of State vs number of death.

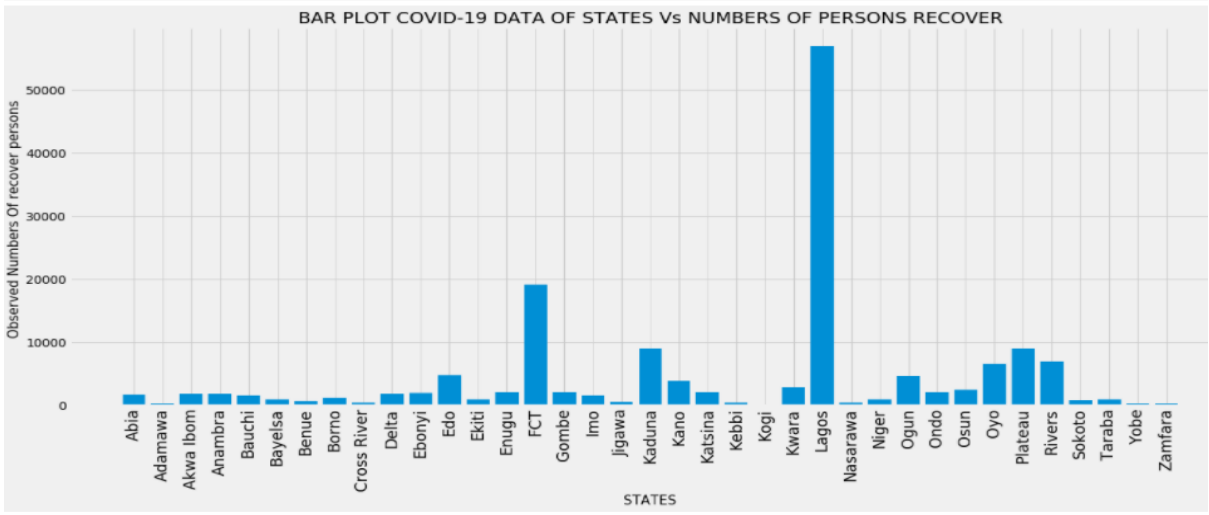

Figure 4. Covid-19 data of State vs Number of persons recovered.

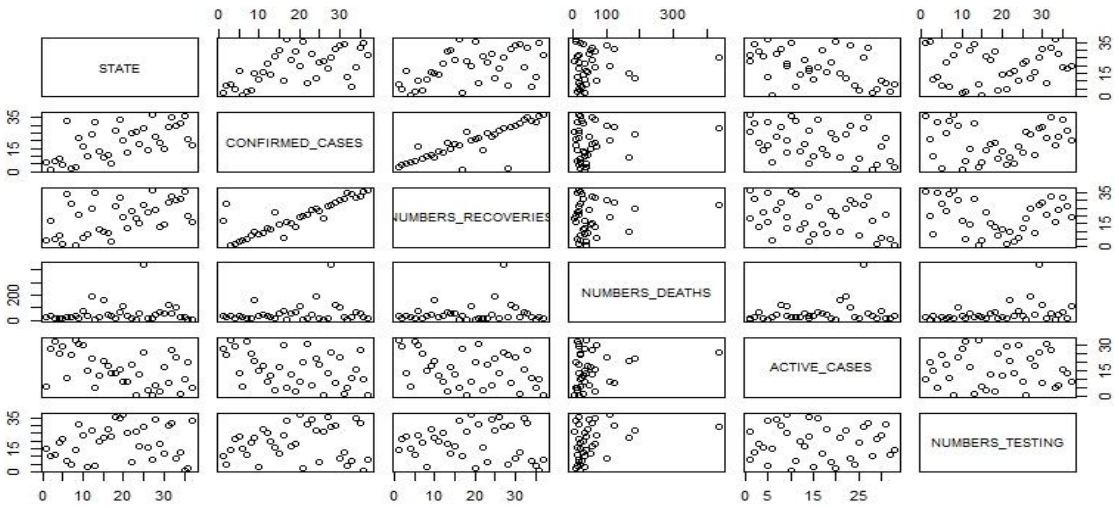

Figure 6. Charts displayed in scatter plot with the following sub headings: state, confirmed cases, number of recovery persons, number of deaths, number of active cases and number of tested persons.

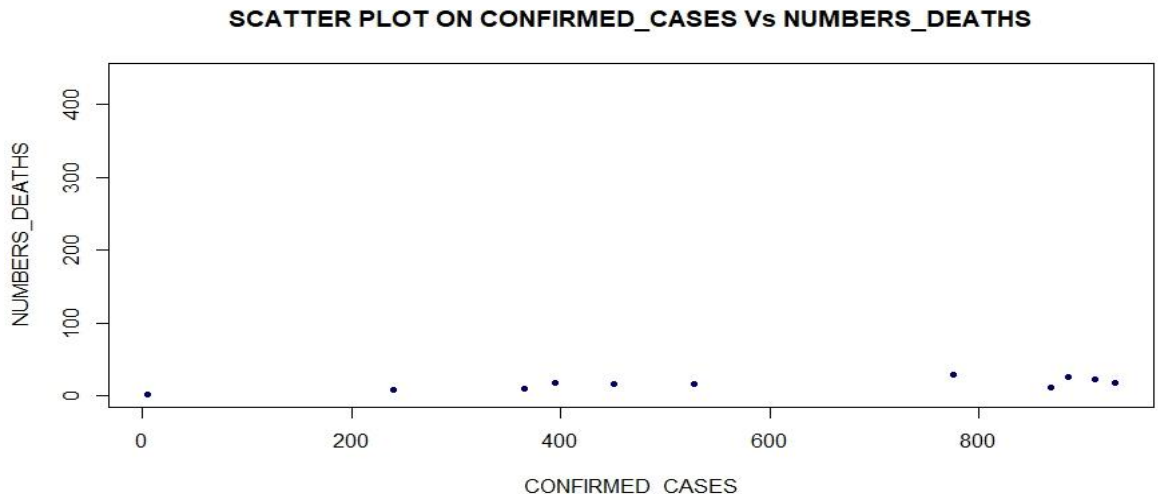

Figure 7. Showing scattered confirmed cases.

Table 2. Showing current data figures.

Current numbers of death cases	2061
Current number of affected persons	7849
Current numbers of recovered persons	154332
Current number of confirmed cases	164233

Model:

x_train, X_test, y_train, y_test = train_test_split(train, test, test_size=0.3, random_state=2)
regr = LinearRegression() => linear regression built-in function in python programming.

Interpretation for the forecasted years

Interpretation for the forecasted years from Table 3, it is observed that in the year 2022 the number of death record will Increase by 750204, Followed by 1502469 in 2023, in 2024 it will increase by 2254734, while in 2025 the number will increase by 3006999 drastically and in 2026 the number will increase by 3759264 respectively.

Table 3. Showing current forecasted figures for each year.

Year	Figure
2022	750204
2023	1502469
2024	2254734
2025	3006999
2026	3759264

Conclusion

From the analysis and visualization, it is observed that the proportion of affected number of persons will continue to increase drastically if proper preventive measures are not in place, and it advised that people residing in these respective states with high numbers of affected persons and high numbers of death should take preventive measures strictly seriously to avoid wide spread of the virus.

From the forecast in the next five years, it is observed that numbers of death cases will continue to increase from the current number: 2061 to 6766263 cumulatively, hence drastic measures should be put in place to curb the spread of the deadly virus.

RECOMMENDATION

1) From those state having high numbers of affected person, should be placed on total lockdown to avoid the wide spread of the virus.
2) The government should made available enough relief materials to those states that are placed on total lockdown.
3) lastly, preventive equipment's should go round to all the nooks and crannies of the states and it is expected of everyone to put on nose marks, and wash up their hands every time and observe social distance.

CONFLICTS OF INTERESTS

The authors have not declared any conflicts of interests.

REFERENCES

Ashley, Elizabeth Muller , Elisabet Vivianne Hafstad , Jan Peter William, Himmels Geir Smedslund, Signe Flottorp Synne, Øien Stensland Stijn, Stroobants Stijn Vande, and Velde Gunn Elisabeth Vist (2020). "The mental health impact of the covid19 pandemic on health care workers, and interventions to help them: A rapid systematic review"Psychiatry Research Volume 293**,** November 2020, 113441

Chu J (2021). A statistical analysis of the novel coronavirus (COVID-19) in Italy and Spain. PLoS ONE 16(3): e0249037. Retrieved May 16, 2021 https://doi.org/10.1371/journal.pone.0249037

De -standard (2020). "The number of new admissions to hospitals as a result of an established coronavirus infection increased by 17 percent" Retrieved February 6, 2021 from https://www.standaard.be/

De Standard; (2020b). Virus in wastewater: not a pathogen, but a good indicator Available Retrieved March 10, 2021 from: https://www.standaard.be/cnt/DMF20200410_04919728.

European Centre for Disease Prevention and Control (ECDC) (2021). Primerscan.ecdc.europa.eu. Stockholm: ECDC;. Available from: https://primerscan.ecdc.europa.eu/?assay=Overview GISAID 2020 [9 June 2021]. Available from: https://www.gisaid.org/.

Jian Xiao, Min Fang, Qiong Chen and Bixiu He. (2020). "SARS, MERS and COVID-19 among healthcare workers: A narrative review" Journal of Infection and Public Health: Volume 13, Issue 6, June 2020, Pages 843-848 Retrieved April 6, 2021 https://doi.org/10.1016/j.jiph.2020.05.019,

Nigeria Centre for Disease Control (2021): Retrieved February 26, 2021 from: https://covid19.ncdc.gov.ng/

Perneger, T., Kevorkian, A., Grenet, T., and Gallée, H. Gayet-Ageron, A. (2020). Alternative graphical displays for the monitoring of epidemic outbreaks, with application to COVID-19 mortality. *BMC Med Res Methodol* 20, 248 (2020). Retrieved May 16, 2021 https://doi.org/10.1186/s12874-020-01122-8

Souvik, Dubey; Payel Biswas, Ritwik Ghosh , Subhankar Chatterjee; Mahua Jana Dubey , Subham Chatterjee; Durjoy Lahiri and Carl J. Lavie, (2020). *Diabetes & Metabolic Syndrome: Clinical Research & Reviews* Volume 14, Issue 5, September–October 2020, Pages 779-788, Retrieved April 23, 2021 from: https://doi.org/10.1016/j.dsx.2020.05.035

Tarylee Reddy , Ziv Shkedy , Charl Janse van Rensburg1 , Henry Mwambi3 , Pravesh Debba4 , Khangelani Zuma5 and Samuel Manda (2021). Short-term real-time prediction of total number of reported COVID-19 cases and deaths in South Africa: a data driven approach, BMC Medical Research Methodology (2021) 21:15 Retrieved February 6, 2021. From: https://doi.org/10.1186/s12874-02001165-x

World Health Organization (WHO) (2020). WHO Statement regarding cluster of pneumonia cases in Wuhan, China 2020 [14 Januray 2021]. Available from: https://www.who.int/china/news/detail/09-01-2020-who-statement-regarding-cluster-of-pneumonia-cases-in-wuhan-china.

World Health Organization (WHO) (2020). Novel Coronavirus – China 2020 [14 January 2020]. Available from: https://www.who.int/csr/don/12-january-2020-novel-coronavirus- china/en/.

Yin Y, Wunderink R.G,(2018). MERS, SARS and other coronaviruses as causes of pneumonia. Respirology (Carlton, Vic). 2018 Feb;23(2):130-7.