Data in Brief 33 (2020) 106424

Contents lists available at ScienceDirect

Data in Brief

journal homepage: www.elsevier.com/locate/dib

Data Article

An exploratory assessment of a multidimensional healthcare and economic data on COVID-19 in Nigeria

Ezekiel Ogundepo [1,*], Sakinat Folorunso [2,*], Olubayo Adekanmbi [1], Olalekan Akinsande [1], Oluwatobi Banjo [2], Emeka Ogbuju [3], Francisca Oladipo [3], Olawale Abimbola [1], Ehizokhale Oseghale [1], Oluwatobi Babajide [1]

[1] *Data Science Nigeria (DSN)*
[2] *The Department of Mathematical Sciences, Olabisi Onabanjo University, Ago Iwoye, Ogun State, Nigeria*
[3] *The Department of Computer Science, Federal University Lokoja, Nigeria*

ARTICLE INFO

Article history:
Received 4 October 2020
Revised 15 October 2020
Accepted 16 October 2020
Available online 21 October 2020

Keywords:
COVID-19 cases per state
Nigeria states budget
COVID-19 laboratories
Population
NCDC tweets
Healthcare facilities
Economy
Time series

ABSTRACT

The coronavirus disease of 2019 (COVID-19) is a pandemic that is ravaging Nigeria and the world at large. This data article provides a dataset of daily updates of COVID-19 as reported online by the Nigeria Centre for Disease Control (NCDC) from February 27, 2020 to September 29, 2020. The data were obtained through web scraping from different sources and it includes some economic variables such as the Nigeria budget for each state in 2020, population estimate, healthcare facilities, and the COVID-19 laboratories in Nigeria. The dataset has been processed using the standard of the FAIR data principle which encourages its *findability, accessibility, interoperability, and reusability* and will be relevant to researchers in different fields such as Data Science, Epidemiology, Earth Modelling, and Health Informatics.

* Corresponding authors.
E-mail addresses: ezekiel@datasciencenigeria.ai, ogundepoezekiel@gmail.com (E. Ogundepo), sakinat.folorunso@oouagoiwoye.edu.ng (S. Folorunso).
Social media: (E. Ogundepo)

https://doi.org/10.1016/j.dib.2020.106424

Specifications Table

Subject	Infectious diseases, Time series, Econometrics, Epidemiology, Decision science, Data science
Specific subject area	A data on the multidimensional study of the effect of COVID-19 on Government budgets, healthcare facilities, and COVID-19 testing laboratories in Nigeria.
Type of data	Tables, Charts, and Maps
How data were acquired	Data were collected from different sources such as NCDC, Geo-Referenced Infrastructure and Demographic Data for Development (GRID3), Geographic Population and Demographic Data (GeoPoDe), and National Bureau of Statistics (NBS). Some data were acquired directly from the host website using the Google Sheets function IMPORTHTML(). The daily updates of COVID-19 in Nigeria were also scrapped from NCDC timeline on Twitter (twitter.com/NCDCgov) using the rtweet package in R programming [15].
Data format	The raw datasets are in Microsoft Excel format and is available on the Mendeley data repository.
Parameters for data collection	Parameters for data collection comprises: 1. Geographical coordinates of all the 36 states and the Federal Capital Territory (FCT). 2. Population estimates of 0 to 100 years per state and by gender. 3. A cumulative report of COVID-19 cases in Nigeria per state. 4. Initial budget presented and the revised budget due to COVID-19 per state in Nigeria. 5. The state initial budget and the revised budget due to COVID-19. 6. Total healthcare facilities in Nigeria per state. 7. COVID-19 laboratories per state. 8. Daily updates of COVID-19 in Nigeria. 9. COVID-19 by the six (6) geopolitical zones in Nigeria. 10. NCDC daily tweets and retweets.
Description of data collection	The description of data collection is as follows: ■ The COVID-19 related dataset which included cases by State and laboratories were collected from the NCDC website [3] while daily COVID-19 cases in Nigeria were collected from NCDC Twitter updates [4]. ■ The initial and revised states' budget due to COVID-19 were collected from various news channels in Nigeria [5]. ■ The population estimate per state and by gender were collected from the GeoPoDe [8]. ■ The healthcare facilities per state were collected from GRID3 [6]. We used various tools and techniques to build and organise these datasets in a spreadsheet as advised by [16] and is available on the Mendeley data repository.
Data source location	Online at http://dx.doi.org/10.17632/8h5rtbbx7m.1
Data accessibility	Repository name: Mendeley Data Repository: http://dx.doi.org/10.17632/8h5rtbbx7m.1 Project URL: https://bit.ly/COVID-19data_project_repo

Value of the Data

- This data explains the effect of COVID-19 on the economy of Nigerian States in terms of budget reallocation and adjustment, and it will be useful for multidimensional studies while carrying out any post-COVID research on Nigeria by government agencies, international organizations and individual researcher.

- The data can help the government of Nigeria prepare the economy better for future pandemics (if any), and it serves as a basis for the appraisal of state government's performance at the end of the fiscal year.
- Data scientists can get hidden information and detect novel patterns from the data which can then be used to train different machine learning models to predict the future of COVID-19 in Nigeria.
- Public health services and disease control institutions can use the data for effective planning using its information on the facilities related to COVID-19 in the Nigerian States.
- It is also relevant for qualitative research works on expert opinion mining on the outbreak to determine both people and government's sentiments on different control measures.

1. Data Description

The World Health Organisation (WHO) announced Coronavirus (COVID-19) as a Public Health Emergency of International Concern on January 30, 2020 and a Pandemic on March 11, 2020 [1]. The COVID-19 pandemic pushed the global economy into a Great Lockdown, schools were shut down, travels banned, social distancing enforced, and many jobs were lost. The rate of the spread of the virus keeps increasing, yet no defined countermeasure, remedy or well-tested medications are handy for its eradication. Nigeria recorded her first case of COVID-19 on Thursday, February 27, 2020, and ever since there has been an exponential growth and spread of the virus all over the country. Due to COVID-19 pandemic, there have been some negative changes in the economy of Nigeria such as the total/partial lockdown, banning of international flights, decrease in the national revenue by over N320 billion, a downward revision of each state's budget, building of isolation centers, loss of jobs, and shutting down of schools [2]. This article presents datasets that are related to COVID-19 as reported by the NCDC, healthcare facilities, and laboratories for testing it, and population estimate by gender and budget for all the 36 states of Nigeria and the Federal Capital Territory (FCT). The datasets are in the Microsoft Excel Workbook with five (5) sheets as presented in Fig. 1.

Sheet 1 This sheet comprises the following features for each state in Nigeria: geocoordinates, population by gender, COVID-19 cases, healthcare facilities, COVID-19 Government laboratories, budget, and total available revenue in 2019.

Sheet 2 This sheet comprises daily cases of COVID-19 in Nigeria as reported by NCDC on https://twitter.com/NCDCgov.

Sheet 3 This sheet shows the aggregate of COVID-19 cases in the six (6) geopolitical zones in Nigeria which includes North-Central (Benue, Kogi, Kwara, Nasarawa, Niger, Plateau, and the FCT), North-East (Adamawa, Bauchi, Borno, Gombe, Taraba, Yobe), North-West (Jigawa, Kaduna, Kano, Katsina, Kebbi, Sokoto, Zamfara), South-East (Abia, Anambra, Ebonyi, Enugu,

Location		Geographical coordinates		Population estimates of 0 to 100 years from GEOPODE			COVID-19 as reported by NCDC on Tuesday, Sep			
States	Geopolitical zones	Latitude	Longitude	Population_Female	Population_Male	Population_Total	No. of Cases (Lab Confirmed)	No. of Cases (on admission)	No. Discharged	No. of Deaths
Abia	South East	5.532003041	7.486002487	2,117,742	1,965,828	4,083,570	894	14	872	8
Adamawa	North East	10.2703408	13.2700321	1,834,828	1,771,877	3,606,705	240	25	198	17
Akwa Ibom	South South	5.007996056	7.849998524	1,801,535	1,684,078	3,485,613	288	6	274	8
Anambra	South East	6.210433572	7.06999711	2,192,738	2,036,023	4,228,761	237	5	213	19
Bauchi	North East	11.68040977	10.19001339	3,140,988	3,038,042	6,179,030	699	17	668	14
Bayelsa	South South	4.7719	6.0699	560,865	539,087	1,099,952	399	6	372	21
Benue	North Central	7.190399596	8.129984089	2,286,249	2,161,538	4,447,787	481	58	413	10
Borno	North East	10.62042279	12.18999467	2,883,631	2,692,682	5,576,313	745	4	705	36
Cross River	South South	4.960406513	8.330023558	1,554,583	1,414,947	2,969,530	87	4	74	9
Delta	South South	5.890427265	5.680004434	1,880,125	1,785,026	3,665,151	1,802	101	1,652	49
Ebonyi	South East	6.178	7.9593	1,438,935	1,332,597	2,771,532	1,040	3	1,007	30
Edo	South South	6.340477314	5.620008096	2,057,366	1,963,003	4,020,369	2,626	24	2,495	107
Ekiti	South West	7.630372741	5.219980834	1,049,887	1,023,826	2,073,713	321	12	303	6
Enugu	South East	6.867034321	7.383362995	1,482,946	1,371,914	2,854,860	1,289	102	1,166	21
Gombe	North East	10.29044293	11.16995357	1,473,771	1,387,304	2,861,075	883	111	747	25
Imo	South East	5.492997053	7.026003588	3,670,761	3,480,781	7,151,542	568	285	271	12

Data | Daily updateof COVID-19 in Nig. | COVID-19 by Geopolitical zone | COVID-19 Lab | NCDC timelines

Fig. 1. A glimpse of the data.

Table 1
Description of variables.

Variable Name	Description	Source
State	List of all 36 states in Nigeria and the FCT	
Geopolitical zones	List of the six (6) geopolitical zones where each state and FCT falls	
Longitude, Latitude	A geographic coordinate of each state in Nigeria	
No. of Cases (Lab Confirmed)	Cumulative of COVID-19 lab-confirmed cases per state	[3]
No. of Cases (On Admission)	Cumulative of COVID-19 cases on admission per state	[3]
No. Discharged	Cumulative of COVID-19 discharged cases per state	[3]
No. of Deaths	Cumulative of COVID-19 death cases per state	[3]
Discharge rate	Proportion of COVID-19 discharged cases and lab-confirmed cases per state expressed in percentage	[3]
Fatality rate	Proportion of COVID-19 deaths cases and lab-confirmed cases expressed in percentage	[3]
COVID-19 government laboratories	Total government COVID-19 testing laboratories for each state in Nigeria	[3]
Healthcare facilities	Total primary, secondary and tertiary healthcare facilities available for each state in Nigeria	[6]
2020 initial budget (bn) presented	Initial budget presented for the year 2020 by the state government before COVID-19	[5]
2020 revised budget (bn) due to COVID-19	Revised budget for the year 2020 by the state government due to COVID-19	[5]
Percentage budget reduction	The percentage of budget reduction for each state in Nigeria	[5]
Population estimate (0-100) years	Population estimate by gender for each state in Nigeria	[8]
Date	Datetime of NCDC tweet	[4]
Tweet	Tweets posted or retweet by NCDC	[4]
Tweet_type	Whether the tweet is emanating from NCDC directly or retweet	[4]
Hashtags	Hashtags included in the tweet	[4]
Media_url	Images included in the tweet	[4]
Mentions_twitter_handle	People mentioned/tagged in the tweet	[4]

The first 16 variables in Table 1 are categorized into six (6) geopolitical zones in Nigeria while the features such as discharge rate, fatality rate, and percentage budget reduction were derived by the authors.

Imo), South-South (Akwa Ibom, Bayelsa, Cross River, Rivers, Delta, Edo) and South-West (Ekiti, Lagos, Ogun, Ondo, Osun, Oyo).

Sheet 4 This sheet shows the general COVID-19 testing laboratories in Nigeria.

Sheet 5 This sheet contains NCDC daily tweets from December 1, 2019 to September 29, 2020.

The datasets were collected through web scraping of different sources with features referenced and described in Table 1. They contain the daily updates of confirmed, recovered and death cases of COVID-19 in Nigeria [4], cumulative laboratory confirmed cases, patients on admission, number of patients discharged, number of death cases, discharge rate and fatality rate of COVID-19 per state in Nigeria [3], COVID-19 Government laboratories [3], total healthcare facilities per state in Nigeria [6], initial, revised, and percentage budget reduction due to COVID-19 [5] total revenue available per state in 2019 [7], and the population estimate per state from 2016 to 2017 [8].

1.1. Nigeria population estimate and healthcare facilities

The Nigeria population estimate per state and the FCT in Fig. 2 was generated from the data that was produced by the WorldPop Research Group at the University of Southampton [8] and it represented bottom-up gridded population estimates (∼100 m grid cells) from 2016 to 2017.

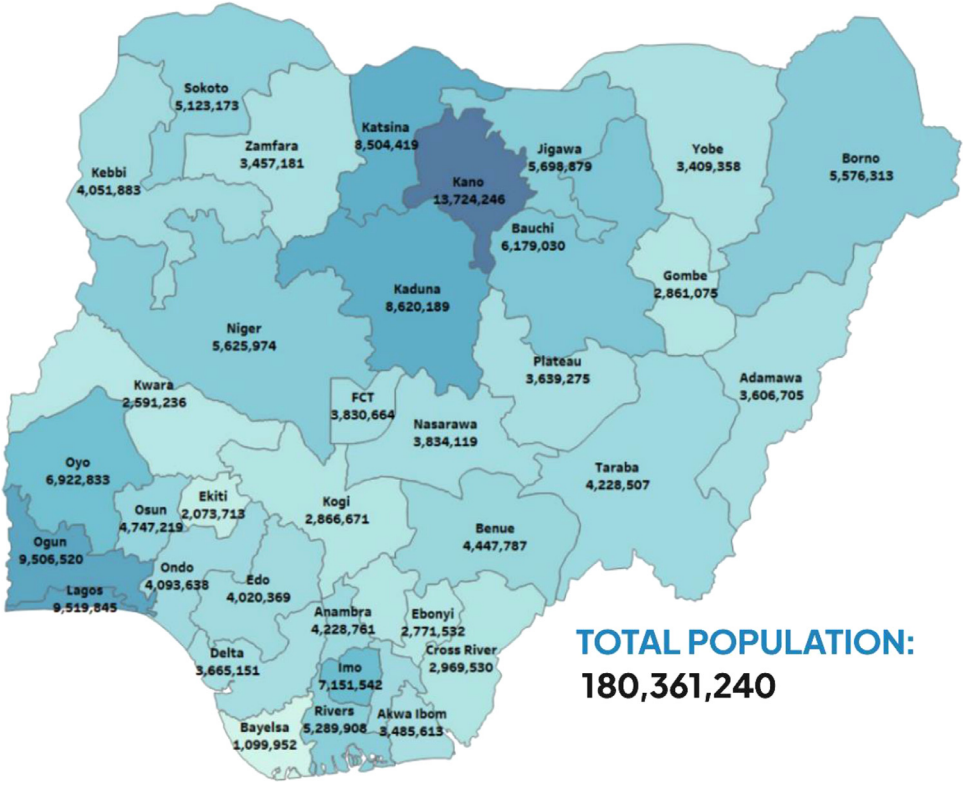

Fig. 2. Nigerian states population estimate from 2016 to 2017.

The data included the estimates of the number of people belonging to individual age-sex groups but only total population estimates for all gender was shown in Fig. 2.

The Fig. 3 shows the total primary, secondary, and tertiary healthcare facilities available in Nigeria.

1.2. COVID-19 data analytics

1.2.1. COVID-19 cases by geopolitical zone

Fig. 4 shows the distribution of COVID-19 for each geopolitical zone. It can be seen that COVID-19 affected the South-West region more when compared to other zones in Nigeria.

1.2.2. COVID-19 laboratories in Nigeria

To contain the spread of COVID-19 in Nigeria, the government of each state has taken the decision in line with the NCDC directives to equip laboratories and provide testing kits for COVID-19. The government laboratories were spread across the six (6) geopolitical zones and they are as follows: North-East (7, 10.4%), South-East (8, 11.9%), North-Central (15, 22.4%), South-West (15, 22.4%), North-West (12, 17.9%) and South-South (10, 14.9%) as shown in Fig. 5.

Figs. 6 and 7 show different kinds of COVID-19 testing laboratories at which anyone can test for COVID-19 in Nigeria and they include government laboratories, fee paying private laboratories, and corporate laboratories. Corporate laboratories are owned by the corporate bodies such as Shell Petroleum Development Company of Nigeria. As reported by NCDC, a laboratory can

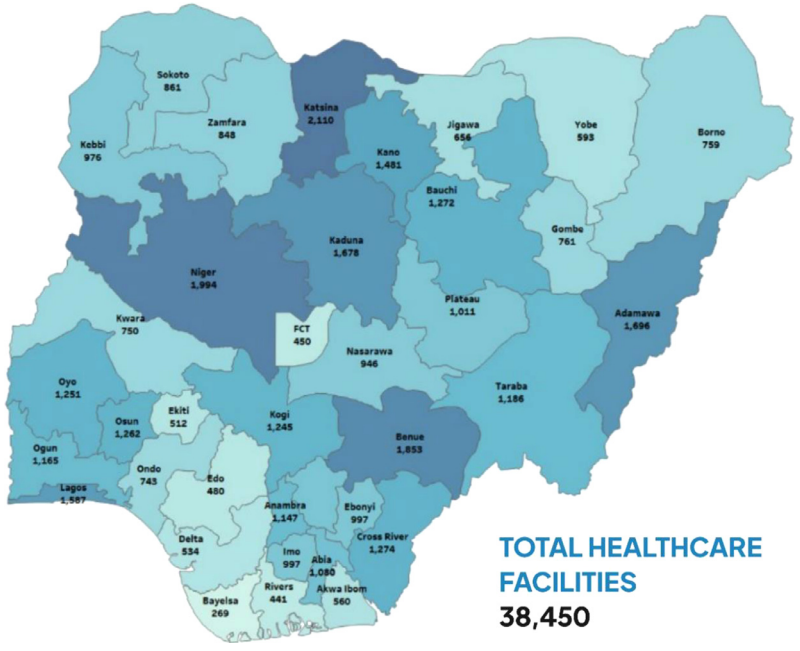

Fig. 3. Healthcare facilities available for each state in Nigeria.

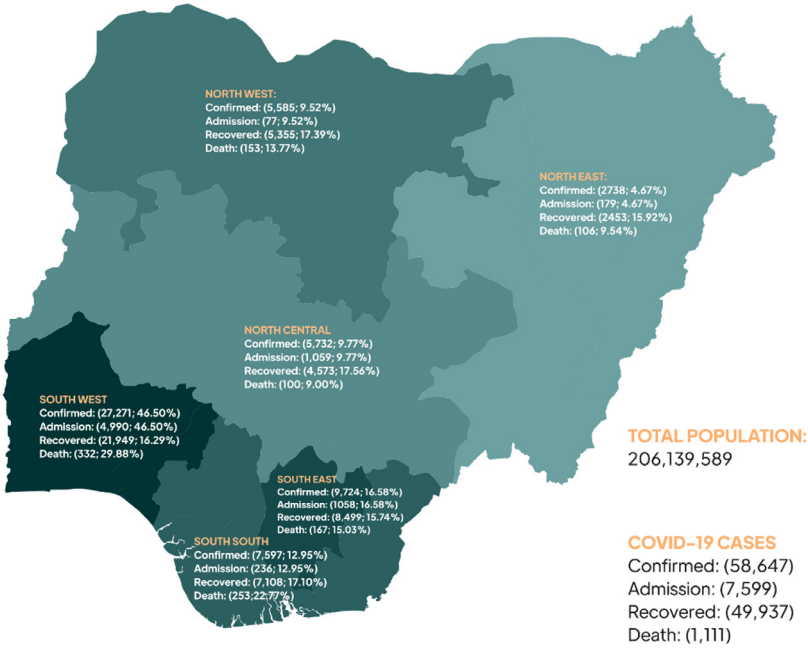

Fig. 4. COVID-19 cases as of September 29, 2020 by geopolitical zones and Nigeria 2020 population estimate at mid-year by the UN.

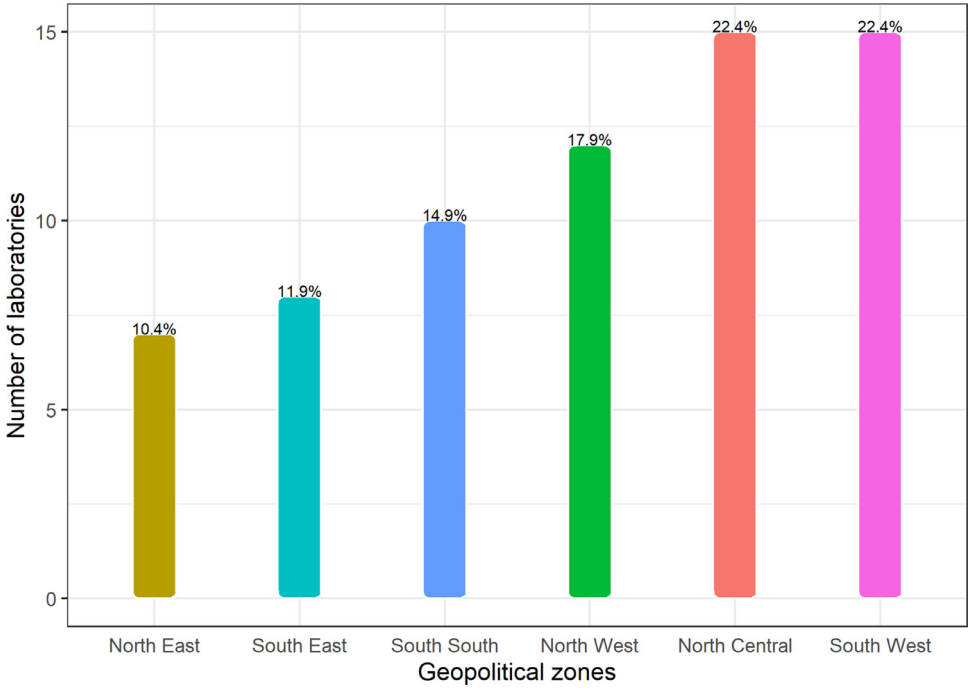

Fig. 5. Government Laboratories by Geopolitical Zones as of September 29, 2020.

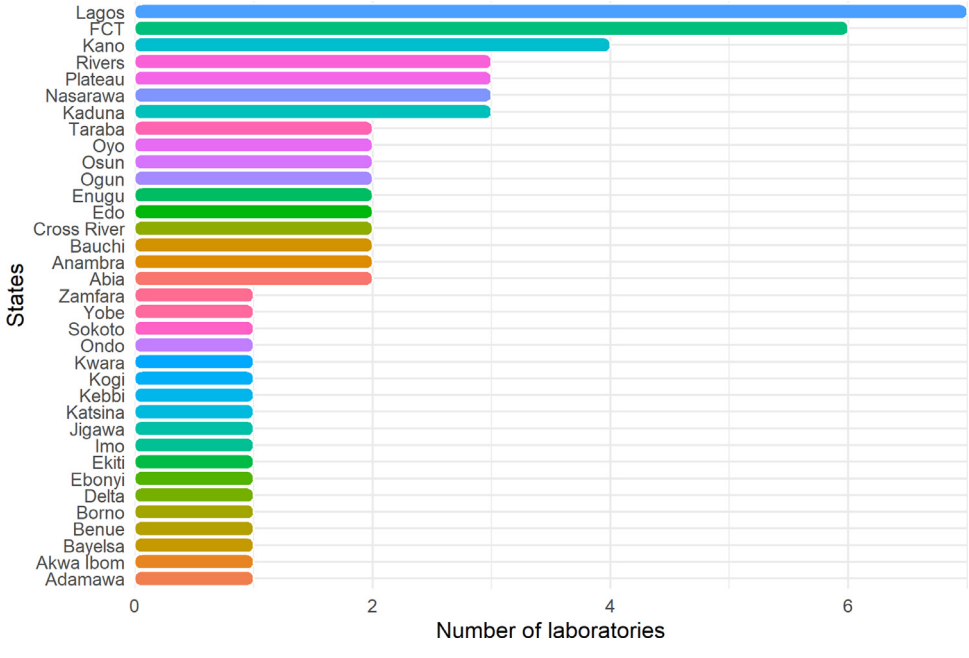

Fig. 6. Government laboratories for testing COVID-19 as of September 29, 2020.

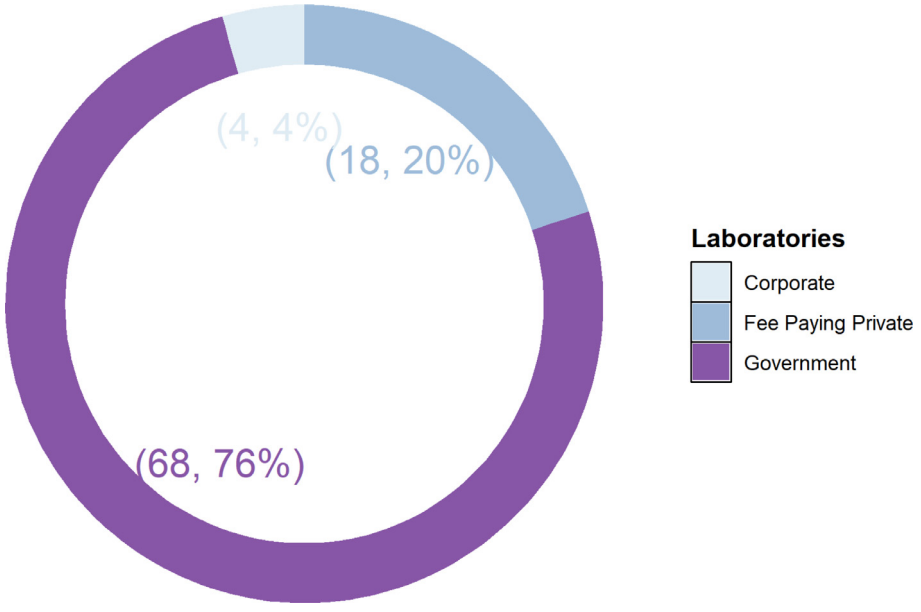

Fig. 7. General COVID-19 testing laboratories as of September 29, 2020.

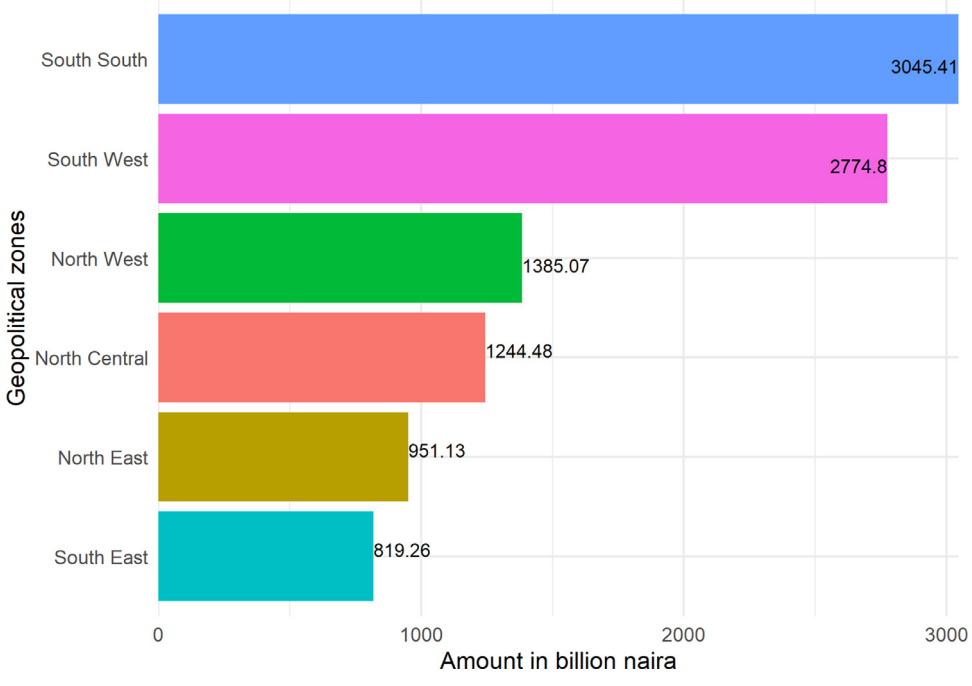

Fig. 8. Initial budget by geopolitical zones in Nigeria.

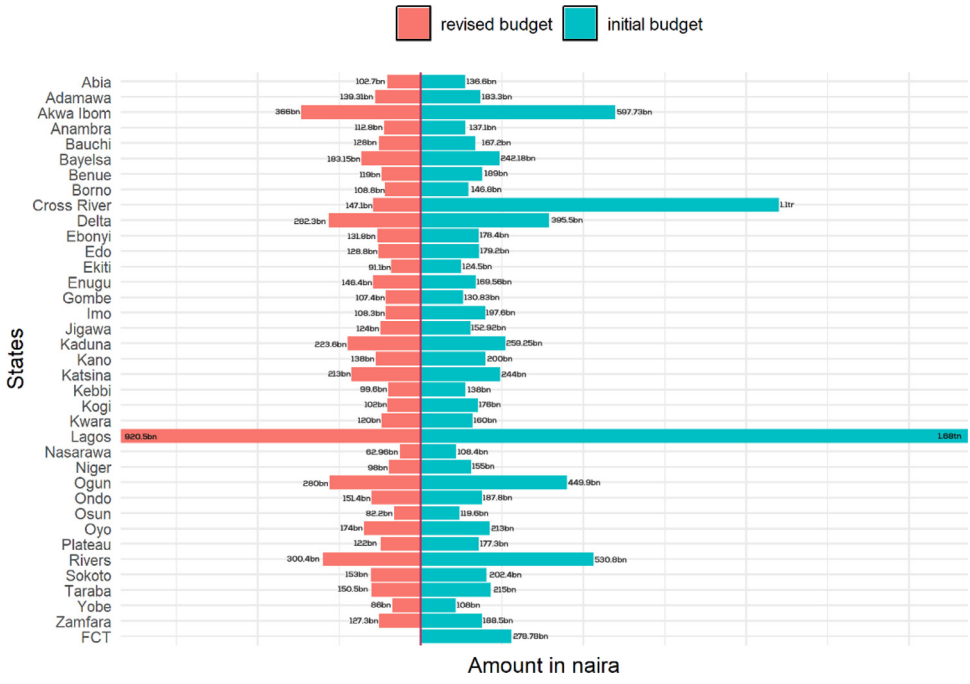

Fig. 9. 36 States of Nigeria reduced 2020 budget due to COVID-19 Pandemic.

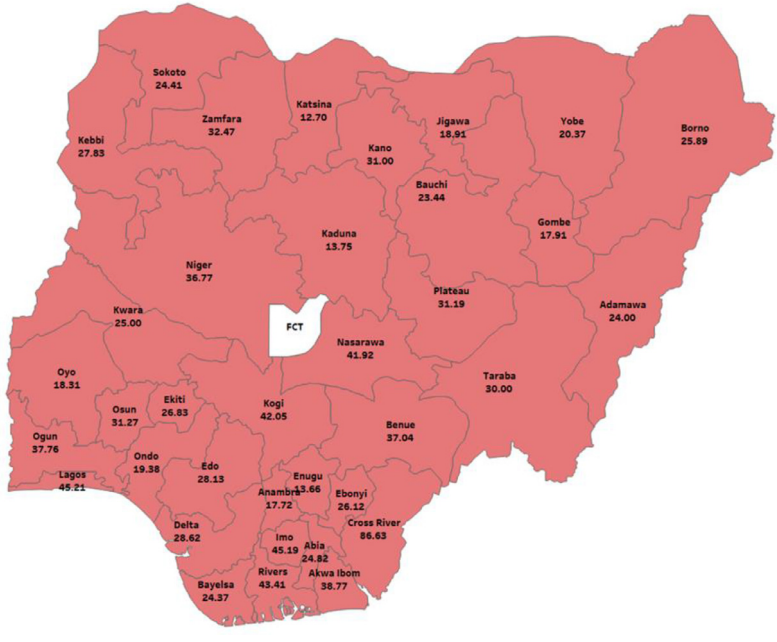

Fig. 10. States reduced budget for 2020 due to COVID-19 (Values in percentage).

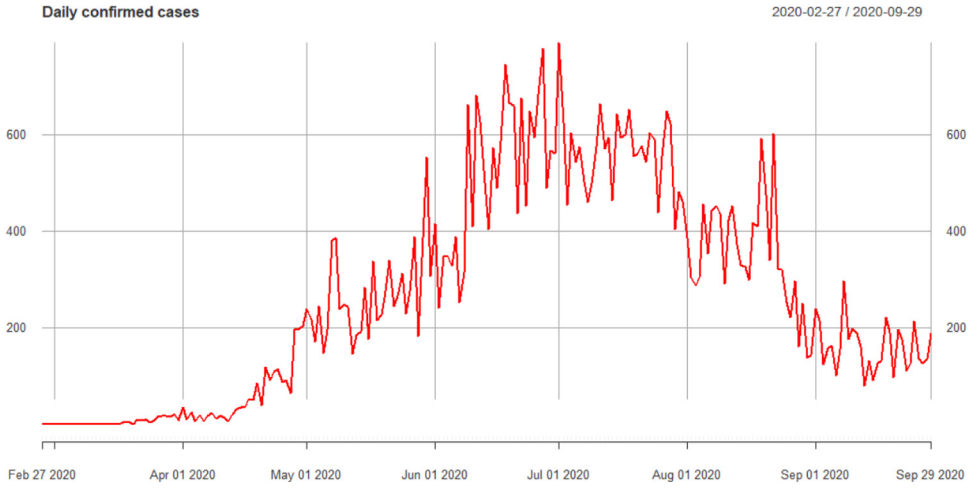

Fig. 11. Daily confirmed cases of COVID-19 in Nigeria.

serve more than one state e.g. African Centre of Excellence for Genomics of Infectious Diseases serves Osun and Ondo States [3].

1.3. Budget analysis

Before the incidence of COVID-19 in Nigeria, all the 36 states including the FCT had presented 2020 budgets to their respective House of Assemblies while some had even signed the appropriation bill into law. According to the data that we collated for the initial 2020 states budget, the 36 states and FCT had budgeted a total sum of N10.220 trillion for 2020 compared to the total sum of N9.140 trillion budgeted in 2019. The South-South region with the following states; Cross River, Akwa Ibom, Rivers, Delta, Bayelsa, Edo and in the order of budget ranking has the highest budget of N3.045 trillion combined while the South-West region with the following states; Lagos, Ogun, Oyo, Ekiti, Ondo, Osun and in the order of budget ranking has a total budget of N2.774 trillion as shown in Fig. 8.

All the 36 states excluding the FCT reduced their cumulative budget of N9.941 trillion to N6.131 trillion, making 38.32 % budget reduction due to COVID-19 as shown in Fig. 9.

As shown in Fig. 9, the top five states that have the highest initial budget are Lagos (N1.68 trillion), Cross River (N1.1 trillion), Akwa Ibom (N597.73 billion), Rivers (N530.8 billion) and Ogun (N449.9 billion) while the top 5 states with the highest revised budget due to COVID-19 are Lagos (N920.5 billion), Akwa Ibom (N366 billion), Rivers (N300.4 billion), Delta (N282.3 billion) and Ogun (N280 billion). This suggests that COVID-19 really affected the Cross River state economy forcing them to reduce their budget by 86.63% as shown in Fig. 10 thereby making their initial budget of N1.1 trillion unrealistic.

Also, as shown in Table 2, the South-South region reduced their budget of N3.045 trillion by 53.77% while the South-West region reduced their budget of N2.77 trillion by 38.76%.

2. Experimental Design, Materials and Methods

We organized our various datasets using tidy principles; each variable is a column, each observation is a row, and each type of observational unit is a table [9]. Tidy datasets are always easy to manipulate and visualize. Majority of the data wrangling was done in R programming

Table 2
Initial and revised 2020 budget by geopolitical zones in Nigeria (36 states + FCT).

Geopolitical zones	Cumulative Initial budget	Cumulative Revised budget	% reduction
North-Central	1,244.48	623.96	49.86%
North-East	951.13	720.01	24.30%
North-West	1,385.07	1,078.5	22.13%
South-East	819.26	602	26.52%
South-South	3,045.41	1,407.75	53.77%
South-West	2,774.8	1,699.2	38.76%
Grand total	10,220.15	6,131.42	40.01%

Budget in billion naira.

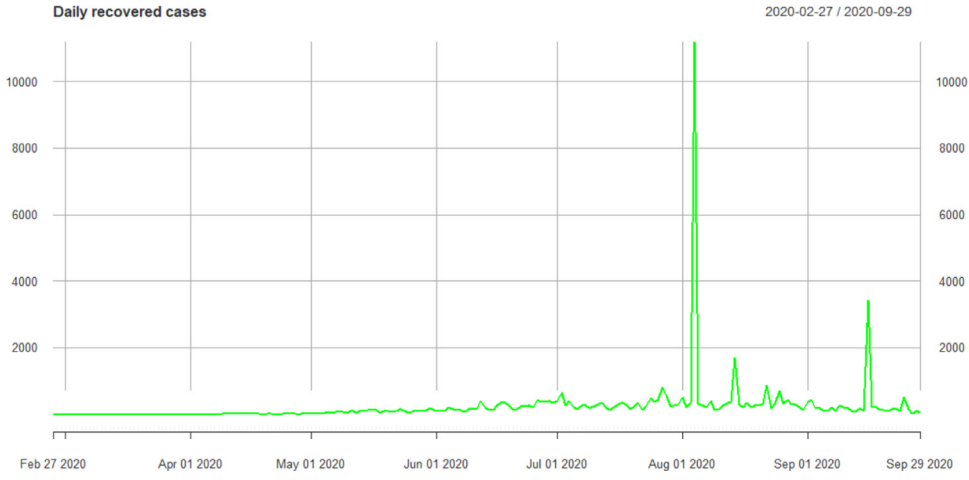

Fig. 12. Daily recovered cases of COVID-19 in Nigeria.

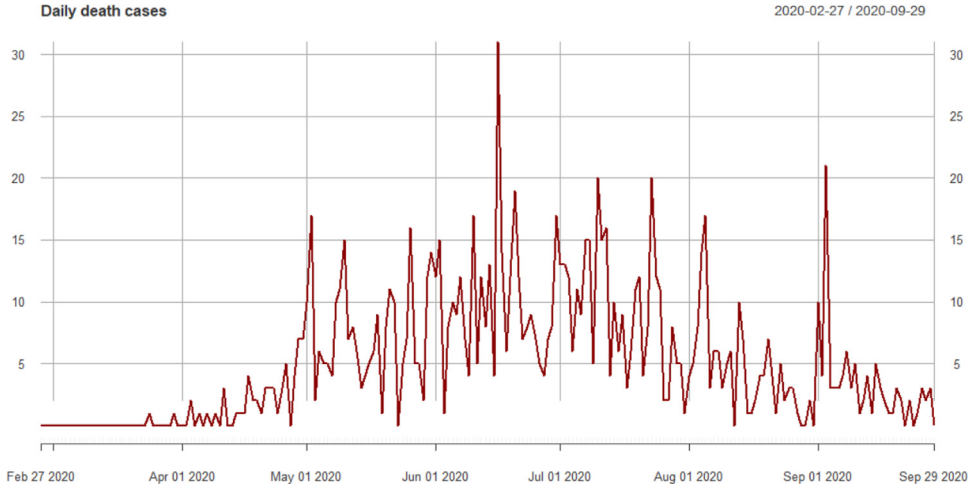

Fig. 13. Daily death cases of COVID-19 in Nigeria.

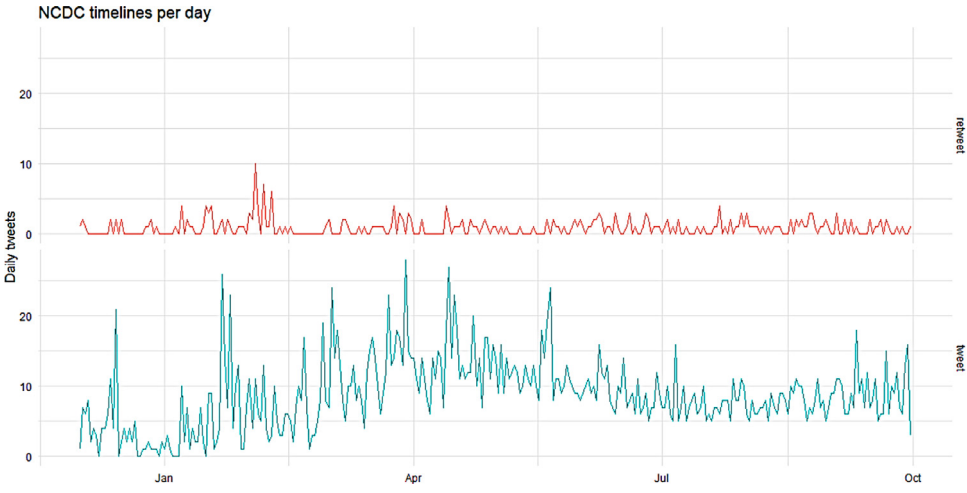

Fig. 14. The NCDC daily tweets and retweets from December 1, 2019 to September 29, 2020.

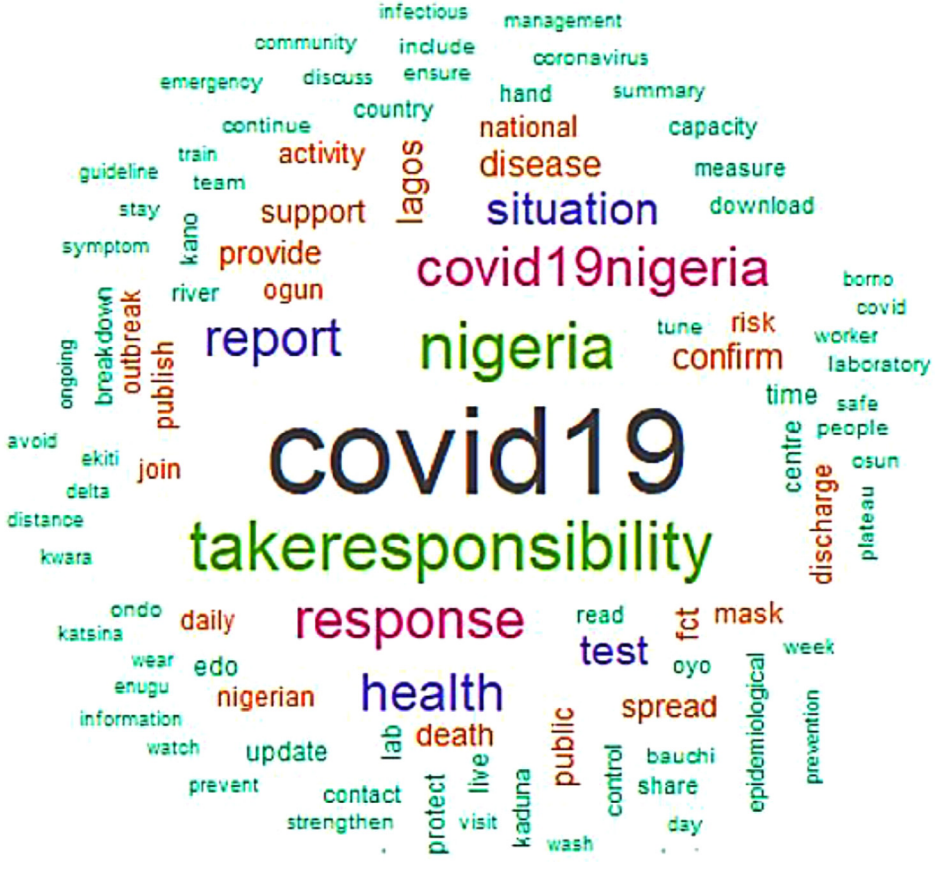

Fig. 15. Word cloud of NCDC tweets from December 1, 2019 to September 29, 2020.

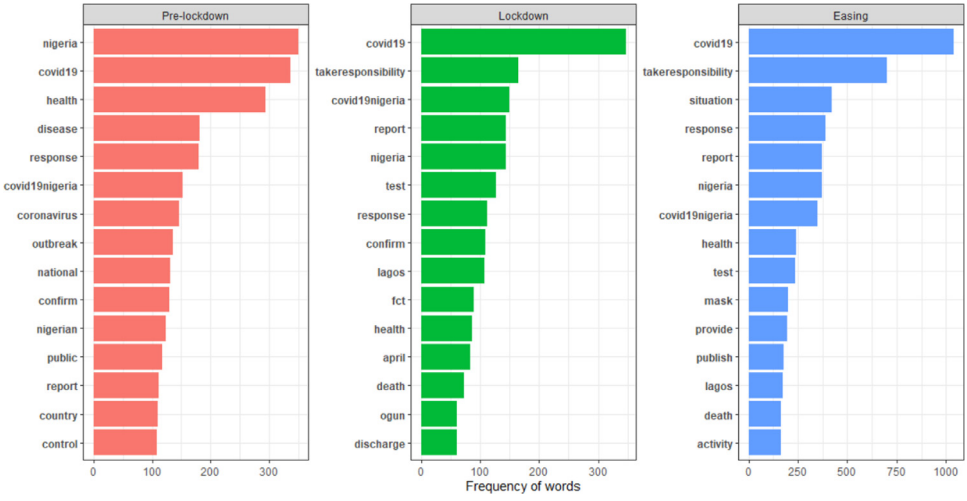

Fig. 16. Top 15 words found in NCDC tweets at different COVID-19 phases in Nigeria.

version 4.0.2 while the dataset resides in the Microsoft Excel 2016. Packages used in R included dplyr and ggplot2 for data analysis and visualization [10,11]. We also used the readxl package [12] to import the Excel file to R programming and the janitor package [13] to clean the data. We used xts package [14] to create an extensible time-series object which can be ordered by time index to plot the daily updates of COVID-19 confirmed, recovered and death cases in Nigeria as shown in Figs. 11 to 13. Dataon NCDC timelines on Twitter was scrapped using rtweet [15] and the frequency of daily tweets and retweets from December 1, 2019 to September 29, 2020 is shown in Fig. 14 while most of the common words found in the tweets are shown in Fig. 15.

Fig. 16 shows the frequency of most common words found in NCDC tweets at various COVID-19 phases in Nigeria which were partitioned into pre-lockdown (December 1, 2020 to March 29, 2020), lockdown (March 30, 2020 to May 4, 2020), and lockdown easing (May 5, 2020 to September 29, 2020) to provide a deep understanding and variation in the importance of the mentioned words in Fig. 15. For example, it could be seen that the word 'takeresponsibility' is more frequently used during the lockdown and easing phases.

We carried out most of the data visualization in RStudio and also used Tableau and CorelDRAW to bring out those visualizations into perfect shape. RStudio project script can be found via https://bit.ly/COVID-19data_project_repo.

Ethics Statement

The process of data collection does not violate any social media privacy, involve experiment, or human sample.

Declaration of Competing Interest

This research is supported by Data Science Nigeria (DSN), 174b, Muritala Mohammed way, Yaba, Lagos, Nigeria, https://www.datasciencenigeria.org.

Acknowledgments

Our gratitude goes to the Virus Outbreak and Disease Network (VODAN) Africa for the training and workshops in creating the science of FAIR data in Africa. We also sincerely appreciate the Nigerian team of VODAN-Africa, Data Science Nigeria (DSN), and Olabisi Onabanjo University for providing data stewardship, and Oniyitan Oluwaseyi for visualizing Fig. 4 using CorelDRAW. We also thank Wole Ademola Adewole, the Implementation Coordinator of GRID3 Nigeria for his time and support in providing all relevant information about GRID3 data that were used in this data article.

Supplementary Materials

Supplementary material associated with this article can be found in the online version at doi:10.1016/j.dib.2020.106424.

References

[1] S.P. Adhikari, S. Meng, Y.J. Wu, Y.P. Mao, R.X. Ye, Q.Z. Wang, H. Zhou, Epidemiology, causes, clinical manifestation and diagnosis, prevention and control of coronavirus disease (COVID-19) during the early outbreak period: a scoping review, Infect. Dis. Poverty 9 (1) (2020) 1–12.
[2] M. Nicola, Z. Alsafi, C. Sohrabi, A. Kerwan, A. Al-Jabir, C. Iosifidis, M. Agha, R. Agha, The socio-economic implications of the coronavirus pandemic (COVID-19): a review, Int. J. Surg. (London, England) 78 (2020) 185–193 https://doi.org/10.1016/j.ijsu.2020.04.018.
[3] NCDC COVID-19 Microsite. https://covid19.ncdc.gov.ng/.
[4] NCDC COVID-19 daily updates on Twitter. https://twitter.com/NCDCgov.
[5] Nigeria States Budget (2020). Retrieved from https://docs.google.com/spreadsheets/d/1GtMaF-JDx3F9prxNVbefnY_bNScfV_vFCRHuhZqSQgw/edit?usp=sharing
[6] eHealth Africa, Proxy LogicsGeo-Referenced Infrastructure and Demographic Data for Development (GRID3), 2019 Accessed July 2, 2020.
[7] National Bureau of Statistics. https://www.nigerianstat.gov.ng/.
[8] WorldPopBottom-Up Gridded Population Estimates for Nigeria,, version 1.2., WorldPop, University of Southampton, 2019, doi:10.5258/SOTON/WP00655.
[9] H. Wickham, Tidy data, J. Stat. Soft. 59 (10) (2014) 1–23.
[10] H. Wickham, F. Romain, H. Lionel, M. Kirill, dplyr: A Grammar of Data Manipulation, 2020 R package version 1.0.0 https://CRAN.R-project.org/package=dplyr .
[11] H. Wickham, ggplot2: Elegant Graphics for Data Analysis, Springer-Verlag, New York, 2016.
[12] H. Wickham, J. Bryan, readxl: Read Excel Files., version 1.3.1, R package, 2019 https://CRAN.R-project.org/package=readxl.
[13] S. Firke, Janitor: Simple Tools for Examining and Cleaning Dirty Data., version 2.0.1, R package, 2020 https://CRAN.R-project.org/package=janitor.
[14] A.R. Jeffrey, M.U. Joshua, xts: eXtensible Time Series., version 0.12-0, R package, 2020 https://CRAN.R-project.org/package=xts.
[15] M.W. Kearney, rtweet: Collecting and analyzing twitter data, J. Open Source Soft. 4 (42) (2019) 1829 (R package version 0.7.0), doi:10.21105/joss.01829.
[16] K.W. Broman, K.H. Woo, Data organization in spreadsheets, The Am Stat. 72 (1) (2018) 2–10.