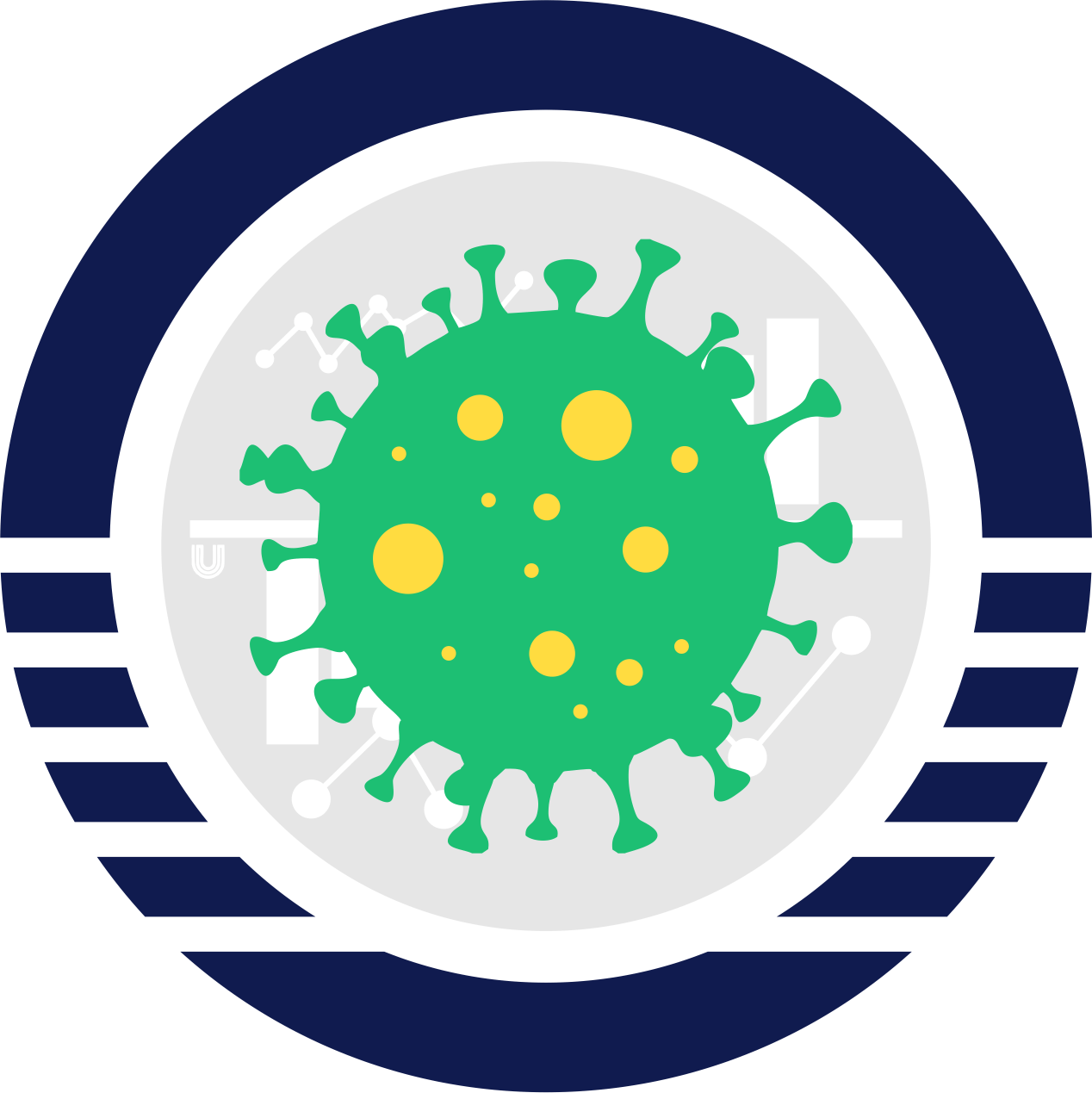
**NIGERIA COVID-19 DATA ANALYSIS USING PYTHON**

****

**By**

**ADEPOJU Yusuf Abiodun**

**USTACKY DATA SCIENCE CAPSTONE PROJECT**

**Jan 2023**

[**INTRODUCTION**](https://www.who.int/emergencies/diseases/novel-coronavirus-2019/question-and-answers-hub/q-a-detail/coronavirus-disease-covid-19)

The COVID-19 pandemic in Nigeria is part of the [worldwide pandemic](https://en.wikipedia.org/wiki/COVID-19_pandemic) of coronavirus disease 2019 ([COVID-19](https://en.wikipedia.org/wiki/COVID-19)) caused by [severe acute respiratory syndrome coronavirus 2](https://en.wikipedia.org/wiki/Severe_acute_respiratory_syndrome_coronavirus_2) (SARS-CoV-2). The first confirmed case in [Nigeria](https://en.wikipedia.org/wiki/Nigeria) was announced on 27 February 2020, when an [Italian](https://en.wikipedia.org/wiki/Italian_people) national in [Lagos](https://en.wikipedia.org/wiki/Lagos) tested positive for the virus. On 9 March 2020, a second case of the virus was reported in [Ewekoro](https://en.wikipedia.org/wiki/Ewekoro), [Ogun State](https://en.wikipedia.org/wiki/Ogun_State).

The effect of the virus in Nigeria has become notable worldwide for being extremely understate, as

there have been just under 266,381 confirmed cases in a country of over 200 million; Deaths however have been minimal (3,155). This has been credited to a warmer [climate](https://en.wikipedia.org/wiki/Climate_change), far younger populations

(Fewer people in care homes), faster government responses, and, crucially, experience in dealing with recent epidemics, such as the [Ebola](https://en.wikipedia.org/wiki/Ebola_in_Nigeria) virus, that most Western countries lacked.

Despite previous analysis and reports, there are only limited information and insights about the effect of covid-19 in various states. In this project, Data science & analytical skills are used to collect data about Covid-19 cases in Nigeria, explore the data, perform analysis, create visualizations, and generate insights.

The dataset used in this analysis consists of data from Nigeria Centre for Diseases Control (NCDC) that monitors the country’s COVID-19 situation official website web extraction or web scraping. Also Johns Hopkins University Center for Systems Science and Engineering (JHU CSSE) daily data on confirmed, death and recovered cases across different countries and Nigeria Community Vulnerability Index data which consists data of socio-economic status, population density, housing type, transportation, epidemiological, health system etc.

At the end of this report, top 10 states affected in terms of confirmed, discharged and death cases are focused on, relationship between the vulnerability index and states are revealed.

**DATA OVERVIEW**

**Nigeria Community Vulnerability Index Data**

Nigeria Community Vulnerability Index data consist of socio-economic status, population density, housing type, transportation, epidemiological, health system etc. Socioeconomic status gives information about the position of an individual or group on the socioeconomic scale, which is determined by a combination of social and economic factors such as income, amount and kind of education, type and prestige of occupation, place of residence, and in some societies or parts of society ethnic origin or religious background.

**Real Domestic Gross Product Data**

This data provided information on the Real Domestic Gross Product (GDP) data for Nigeria at different quarters starting from some years back. This will help to determine and compare the impact of COVID-19 on the economy at pre and post covid-19 era.

**State Budget Data**

This data consist of initial and revised budget of states as States across the country reduced their initial budget due to the impact of COVID-19 on the economy. This gives information to determine the impact of COVID-19 on the economy.

**Global Covid-19 data**

The data in this category are daily recorded recovered, death and confirmed cases from John Hopkins University, it consists Covid-19data of different countries.

**METHODOLOGY**

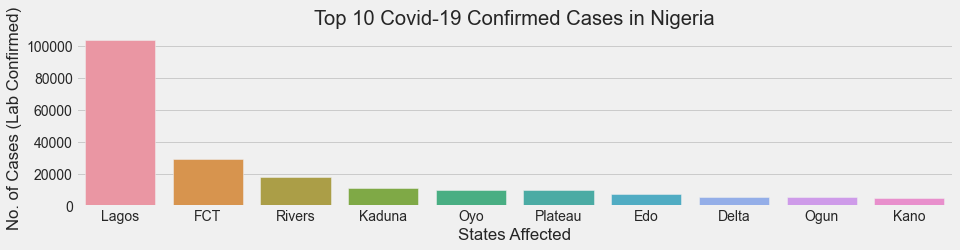
The analysis was carried out with the following method

* Data Gathering: The dataset used was gathered by performing web scrapping to obtain up to date information about covid-19 in Nigeria and cloning data from official repositories of JHU universities. The scraping was done by using beautiful soup, a module for web extraction and a request module to query for information.

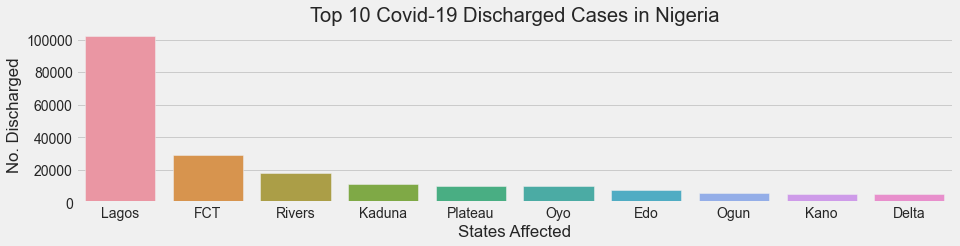
With the help of the request module, a request was sent to the official NCDC website to get the text data. The text data was used in creating a web soup, parsing the soup with html parser. Information’s about the table columns and their observations was extracted.

* Data Cleaning and wrangling: Data cleaning is the process of removing incorrect or otherwise erroneous data from the dataset. These errors include incorrectly formatted data, redundant entries, mislabeled data, and null values. Data cleaning improves the quality of our data as well as any insights or decisions that you draw based on the data.
* Exploratory Data Analysis: Exploratory data analysis (EDA) is used to analyze and investigate data sets and summarize their main characteristics, also used in data visualization methods. It helps determine how best to manipulate data to get the answers to business problems, making it easier to discover patterns, spot anomalies, test a hypothesis, or check assumptions.

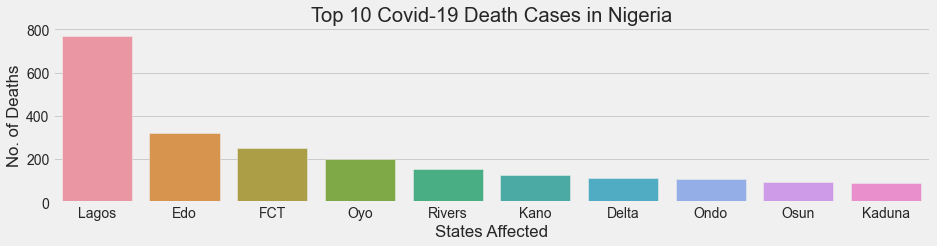
The analysis include finding top 10 states in terms of Confirmed Covid-19 cases by Laboratory test



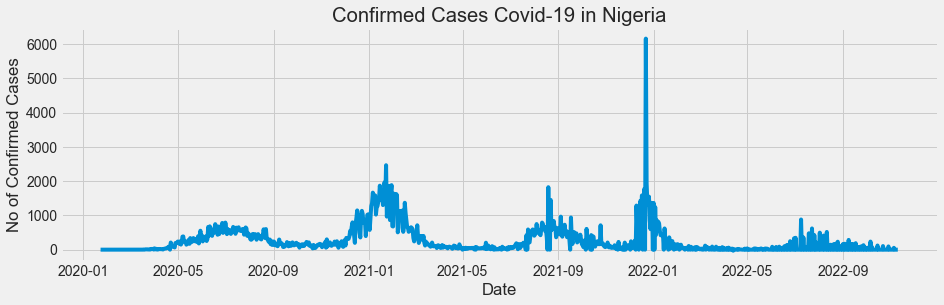
Top 10 states in terms of Discharged Covid-19 cases



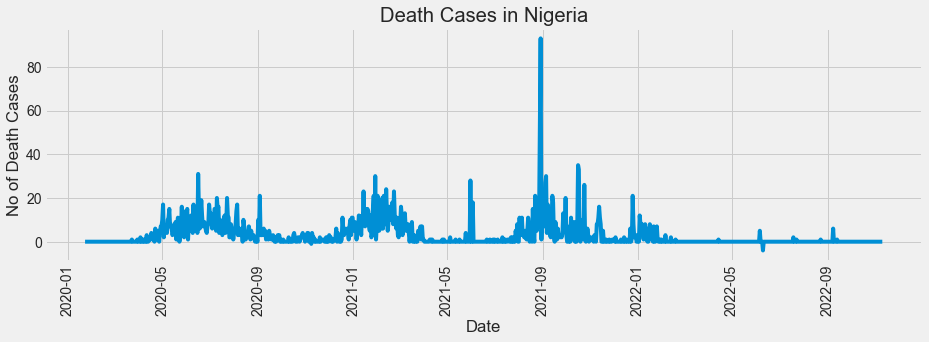
Top 10 Death cases



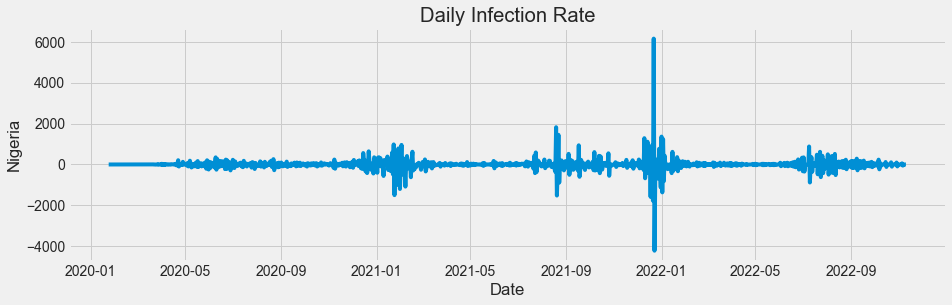
Confirmed Cases Covid-19 in Nigeria



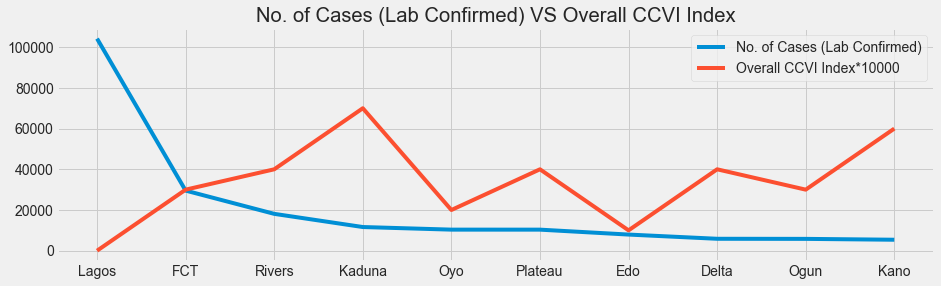
Death Cases in Nigeria



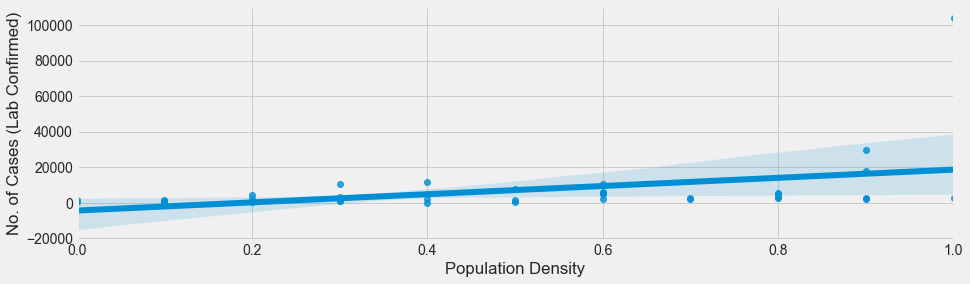
Daily Infection Rate in Nigeria



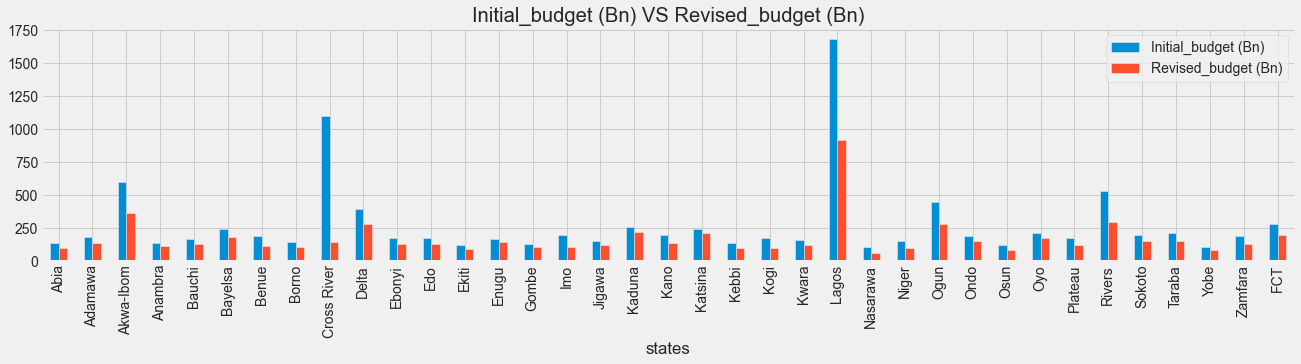
Plot showing No of Cases (Lab Confirmed) and overall CCVI index



Regression plot between Confirmed Cases and Population Density



Initial budget (Bn) VS Revised budget (Bn)



GDP values for each year & quarters.



* Lagos state has the highest number of covid-19 cases, leading the second highest by FCT by 71% while Ogun and Kano are the 9th and 10th respectively.
* Lagos state has the highest number of discharged cases, leading the second highest FCT by 71% while Kano and Delta are the 9th and 10th state with discharged cases respectively.
* Lagos state has the highest number of death cases of covid-19 patients, leading the second highest Edo state by 58% while Osun and Kaduna state are the 9th and 10th state with death cases of covid-19 patients.

- On 2022-11-10 highest number of active confirmed cases in Nigeria was recorded with 6158 numbers of active patients, on 2022-11-10, 93 death was also recorded as the highest.

**Conclusions**

The analysis is carried out to investigate the cases of covid-19 in the country, emphasis are laid on top 10 states are and the following conclusions were drawn.

This analysis consists of explanatory and exploratory data analysis, the basic information was provided and overview of covid-19 was explored to identify missing gaps and areas where further analysis might be needed especially in this era of subsequent waves of covid-19 to make provision to ratify its effect.

Lagos state has the highest number of covid-19 cases, leading the second highest by FCT by 71% while Ogun and Kano are the 9th and 10th respectively.

Lagos state has the highest number of discharged cases, leading the second highest FCT by 71% while Kano and Delta are the 9th and 10th state with discharged cases respectively.

Lagos state has the highest number of death cases of covid-19 patients, leading the second highest Edo state by 58% while Osun and Kaduna state are the 9th and 10th state with death cases of covid-19 patients.

On 2022-11-10 highest number of active confirmed cases in Nigeria was recorded with 6158 numbers of active patients, 2022-11-10 number of death also recorded.

On 2022-11-10 highest number of active confirmed cases in Nigeria was recorded with 6158 numbers of active patients, on 2022-11-10, 93 death was also recorded as the highest.

Gombe and Yobe has the overall highest CCVI index of 1, followed by Sokoto, Jigawa, Zamfara, Borno with CCVI of 0.9, Akwa Ibom, Cross River, Lagos, Rivers are the states with highest difference in revised and initial budgets.

There is a fluctuation in GDP at first quarters over the years, rise across the second, third and fourth quarters with a sharp fall in 2020.