# EXECUTIVE SUMMARY

**APRIL 2021**

## NIGERIA COVID19 DATA ANALYSIS USING PYTHON

***Authored by***:

**Sadiya Ovaino Idris**

# INTRODUCTION

Coronavirus disease (COVID-19) is an infectious disease caused by a newly discovered coronavirus, and it has affected major parts of the world. Nigeria, a West-African country, has also been affected by the COVID-19 pandemic after recording its first case on 27th February 2020.

Nigeria is a country with 37 states - Federal Capital Territory included and a fast-growing economic environment with about 200 million citizens. COVID-19 has affected several country activities as the country steadily progressed from its first case to shutting down major airports, state-wide lockdown, curfews, and reviving its economy.

In this project, data science and analytics skills are employed to collect data, explore the data, perform analysis, create visualizations and generate valuable insights.

# PROCESSES

## Data Overview

The data source is divided into different parts, and the data was combined to perform analysis and provide insights.

1. **The Nigeria Centre for Diseases Control (NCDC**) monitors the country's COVID-19 situation, and releases data on metrics across all the 37 states in the country. From NCDC COVID-19 official website, data was obtained by performing a web scraping.

   **2. The Johns Hopkins University Center for Systems Science and Engineering (JHU CSSE)** publishes daily data on confirmed, death and recovered cases across different countries. Daily data for Nigeria from their repository was derived with related insights.

   **3. Nigeria Community Vulnerability Index data** shows how the vulnerability index was computed by considering several factors such as socio-economic status, age, population density, housing type, transportation, epidemiological and health system.

   **4. Real Domestic Gross Product Data** contains data on the Real Domestic Gross Product(GDP) for Nigeria. This will help you determine the impact of COVID-19 on the economy.

   **5. State Budget Data** contains states across the country with their initial and revised budgets during the pandemic.

**And additional datasets from:**

**6. WHO COVID19 table data** which provides information on the countries, date reported, WHO Region, Cumulative confirmed cases and transmission classification

**7. Simplemaps.com**, a website that contains Nigerian states with their longitude and latitude coordinates in csv format.

**8. Esri**, a website that provides updated data on coronavirus for countries with their cumulative total confirmed, recovered, active and death cases

## Methods

### A. Data Wrangling

- Libraries were imported
- The data sets were loaded
- Data was cleaned to remove unneccessary columns and rows, chage datatypes a nd rename columns for simplicity

### B. Data Visualization

The findings were communicated and presented using chart visualizations from Python visualization libraries (Folium, Plotly, Seaborn and Matplotlib). The charts were genera ted as:

- Bar plots
- Scatter plots
- Geographical map
- Line plots
- Pie charts
- Regression plots
- Heatmap

## Data Analysis

**Descriptive analysis:** This was done on the NCDC covid19 data using the describe() method:

***Output:***

```
ncdc_covid19_data.describe()
```

| | Lab Confirmed Cases | Admitted Cases | Discharged Cases | Death Cases |
|---|---|---|---|---|
| count | 37.000000 | 37.000000 | 37.000000 | 37.000000 |
| mean | 4452.864865 | 208.945946 | 4188.189189 | 55.729730 |
| std | 9806.325629 | 420.197371 | 9623.486738 | 77.399271 |
| min | 5.000000 | 0.000000 | 3.000000 | 2.000000 |
| 25% | 930.000000 | 6.000000 | 747.000000 | 17.000000 |
| 50% | 1909.000000 | 25.000000 | 1744.000000 | 32.000000 |
| 75% | 3242.000000 | 99.000000 | 2814.000000 | 57.000000 |
| max | 58317.000000 | 1996.000000 | 56990.000000 | 439.000000 |

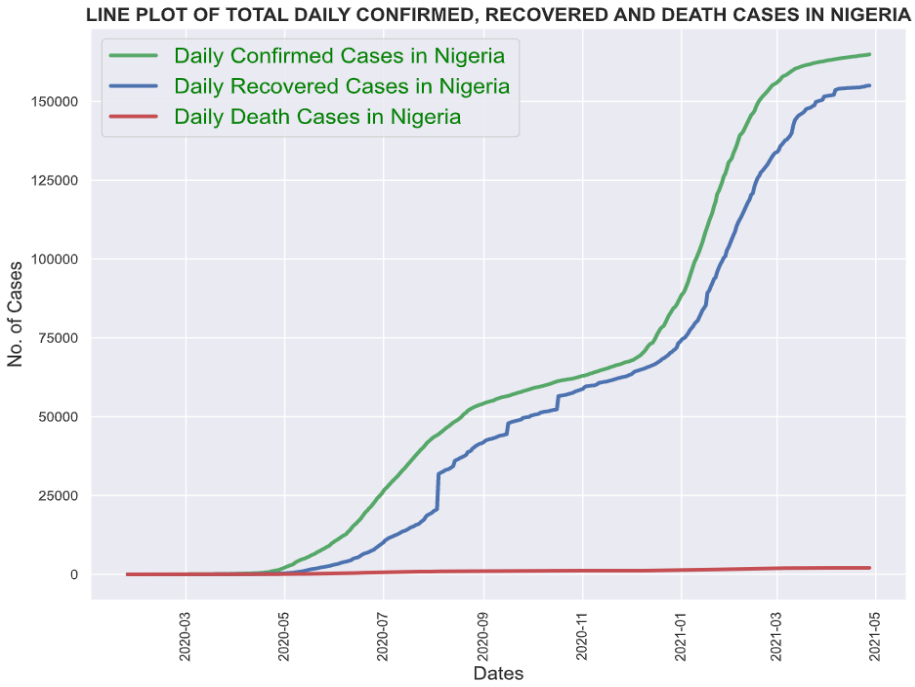
***Insights:***

*From this analysis, the following were determined:*

- *Average number of COVID19 deaths across all states in Nigeria is approximately 56.*
- *Highest number of deaths recorded in a state is 439 and minimum is 2.*
- *75% of cases result into approximately 57 deaths.*
- *Highest number of discharged cases recorded in a state is 56990 and minimum is 3.*
- *Average number of COVID19 admitted cases across all states in Nigeria is approximately 4188.*

**Exploratory Analysis** was done to determine:

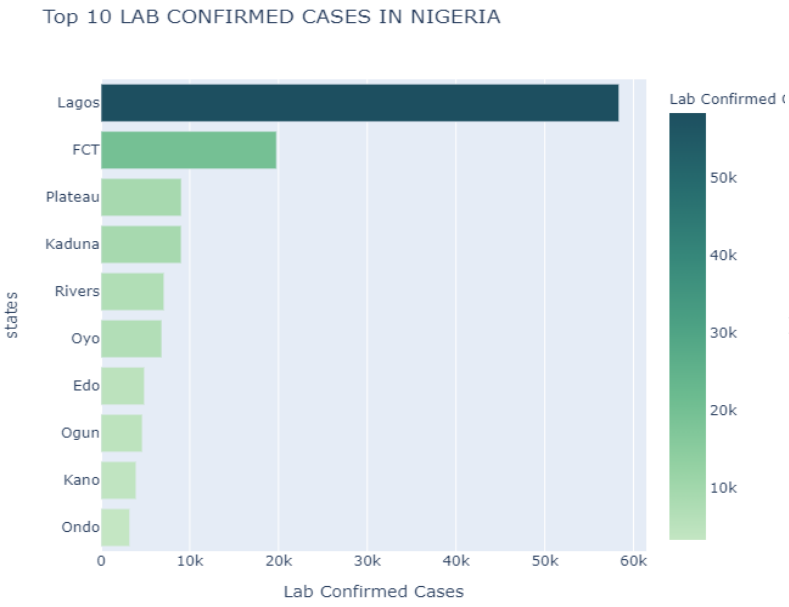
1. Total daily confirmed, recovered and death cases in Nigeria

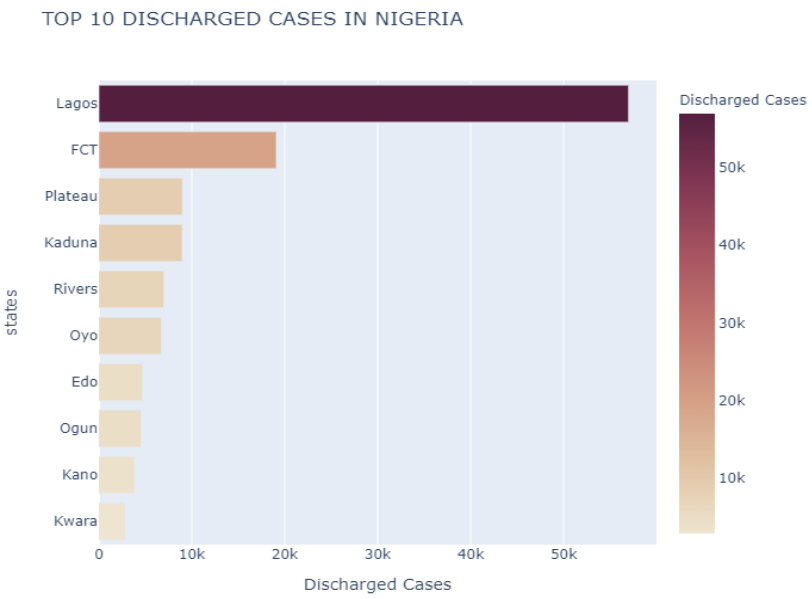


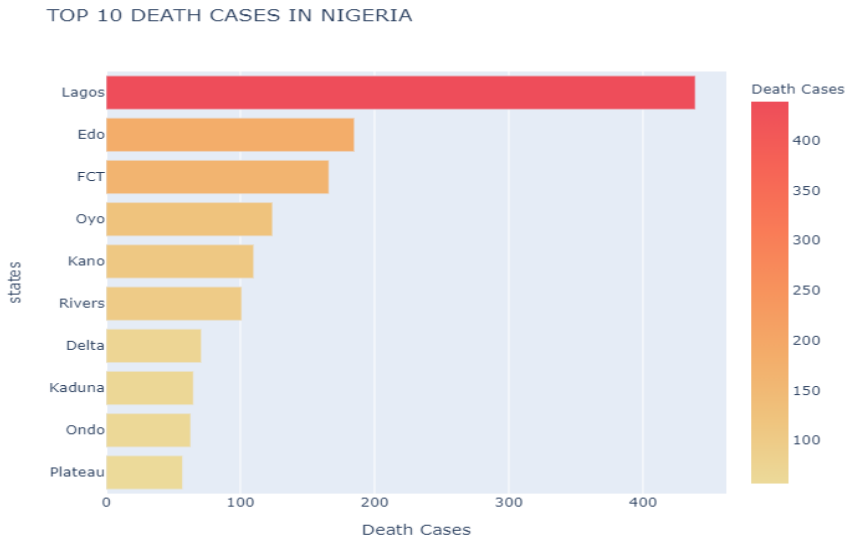

***Insight:***

*From this line plot, it is evident that the rate of daily confirmed cases moves in almost the same rate as daily recovered cases. Beyond that, COVID19 cases began to rise around May 2020 there is an odd jump in January 2021, then the rate of new cases slows down for a while, then speeds up again in March. We need to dig deeper to see what is happening.*

## 2. Top 10 Confirmed, Recovered and Death cases



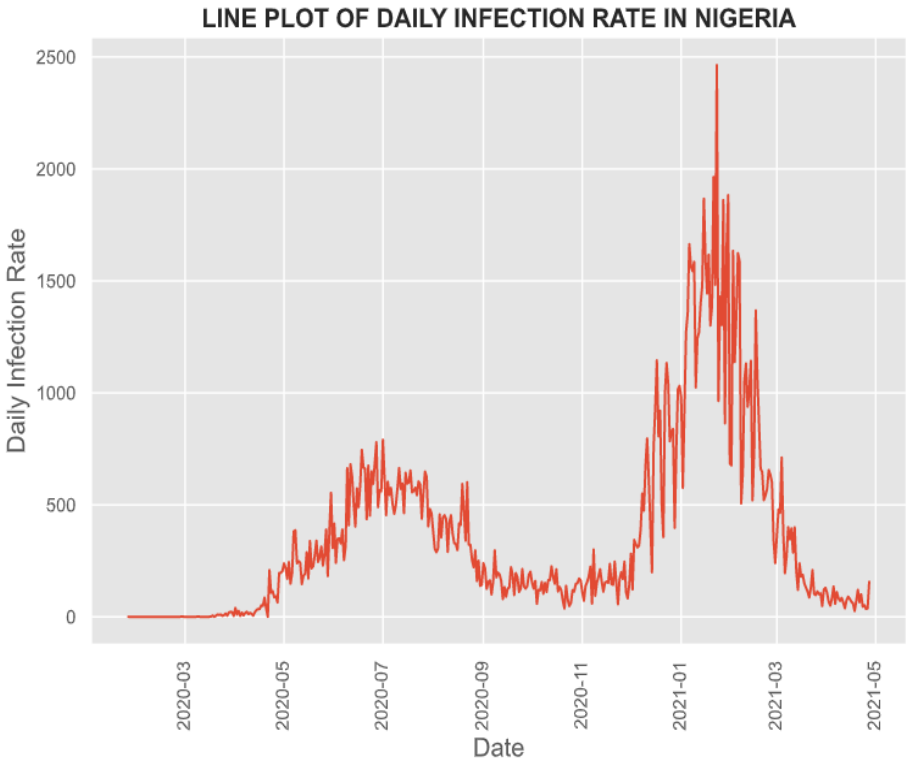




***Insights:***

*From this chart, it is evident that the top 10 states with the most confirmed cases are: Lagos, FCT, Plateau, Kaduna, Rivers, Oyo, Edo, Ogun, Kano and Ondo, while the top 10 states with the most discharged cases are: Lagos, FCT, Plateau, Kaduna, Rivers, Oyo, Edo, Ogun, Kano and Kwara, also, the top 10 states with death cases are: Lagos, Edo, FCT, Oyo, Kano, Rivers, Delta, Kaduna, Ondo and Plateau.*

*It should be noted that most of the states with the most lab confirmed cases also have the most discharged cases and death cases.*

3. Daily infection rate in Nigeria

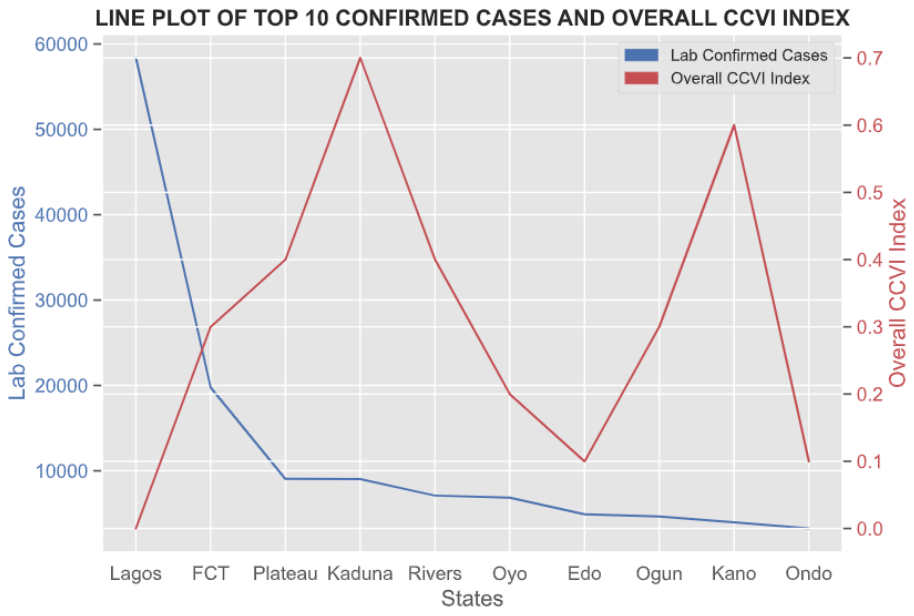



***Insight:***

*From this line plot, it is evident that the rate of daily infection began to rise around April 2020, then slows down for a while, then again, speeds up in December 2020 till it got to its highest peak around late January 2021, after which, the rate of daily infection keeps slowing down up to this time of analysis.*

***The maximum infection rate** for a day in Nigeria is **2464.0***
***The date of maximum infection rate** is **23rd January, 2021***

4. Relationship between Lab Confirmed Cases and Overall CCVI Index




***Insight:***

*This line plot shows the top 10 states with Confirmed COVID19 Cases and their Overall Community Vunerability Index (CCVI). From the plot, the following observations are deduced:*

1. *Lagos state, which has the highest number of COVID19 cases, actually has the lowest Overall CCVI Index of 0.0, in other words, it is the least vunerable to COVID19 with all factors (like age, fragility, etc.,) being considered.*
2. *Edo and Ondo states both have a low Overall CCVI Index of 0.1.*
3. *The rest of the states have a higher Overall CCVI Index compared to Lagos, Edo and Ondo, with Kaduna and Kano having the highest CCVI Index of 0.7 and 0.6 respectively, meaning they are highly vunerable to COVID19, suffice to say, this two states have earned their spots at the top 10 states with Confirmed COVID19 Cases.*

*Although the same cannot be said about Lagos, Edo and Ondo, which begs the question:*

***Why would states with very low Overall CCVI Index be among the top 10 states with Confirmed COVID19 Cases??? Does this mean Overall CCVI Index is not really a strong factor in measuring confirmed cases???***

5. Relationship between Lab Confirmed Cases and Population Density

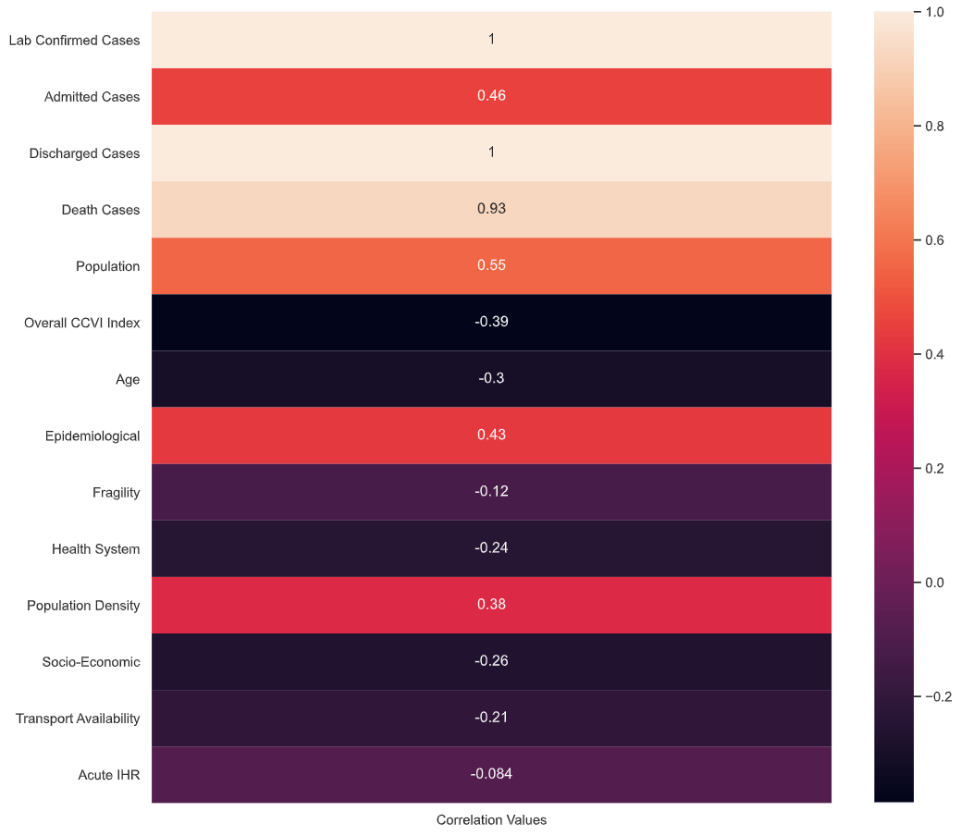

**Insight:**

*This regression plot shows the linear relationship between all Confirmed Cases in the states and their Population Density. From the plot, it is evident that there is little correlation between the two variables.*

6. Degree of Correlation between Lab Confirmed Cases and various risk factors

*A **degree of correlation** indicates how a change in one variable affects another variable and can be positive or negative. This correlation calculation is performed on the states in Nigeria.*

*Values will range from 1 (**very strong positive correlation,** i.e., as one goes up, the other tends to also) to -1 (**very strong negative correlation,** i.e., as one goes up, the other will tends to push down, or vice-versa) and 0 (**no linear trend**).*




***Insights:***

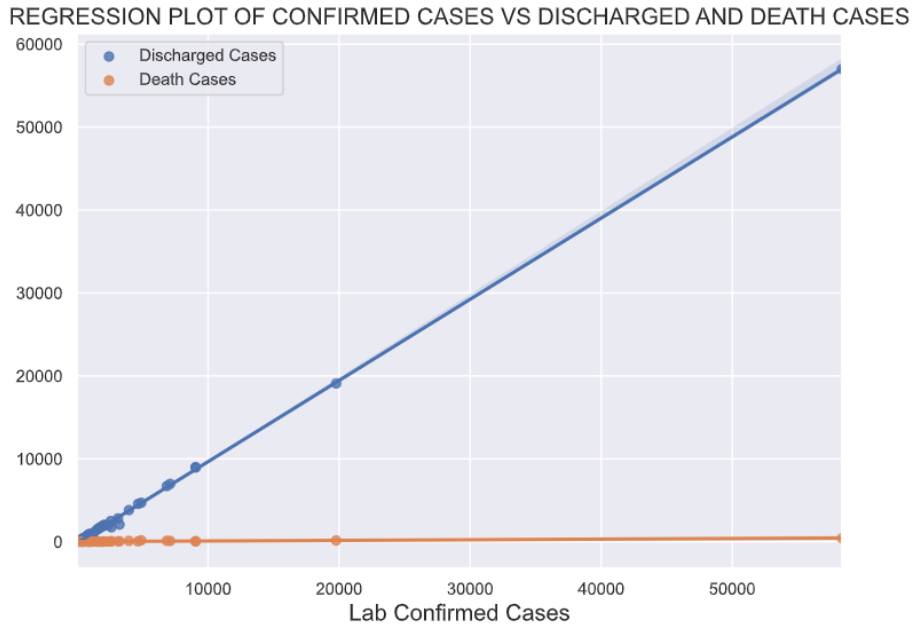
*Checking correlations between variables:*

- *Number of discharged cases and death cases are highly positively correlated with number of Lab confirmed cases.*
- *Population is positively correlated with number of Lab confirmed cases with a value of 0.55.*
- *Number of admitted cases and epidemiological factor have an average positive correlation with Number of lab cases.*
- *Overall CCVI Index, Age, Fragility, Health System, Socio-Economic, Transport Availability and Acute IHR factors all have negative correlations with number of Lab Confirmed Cases*

*Also note that the total lab confirmed cases obviously has a direct correlation value of 1 with itself. However, most of the correlation values are not high correlation values which are usually greater than 0.6 or lesser than -0.6. Perhaps, this goes to show how devastatingly widespread the effects of this contagious disease is.*

*Let's use regression plot to see how the number of discharged cases and death cases are highly positively correlated with number of Lab confirmed cases:*

A. Lab Confirmed Cases Vs Discharged Cases and Death Cases



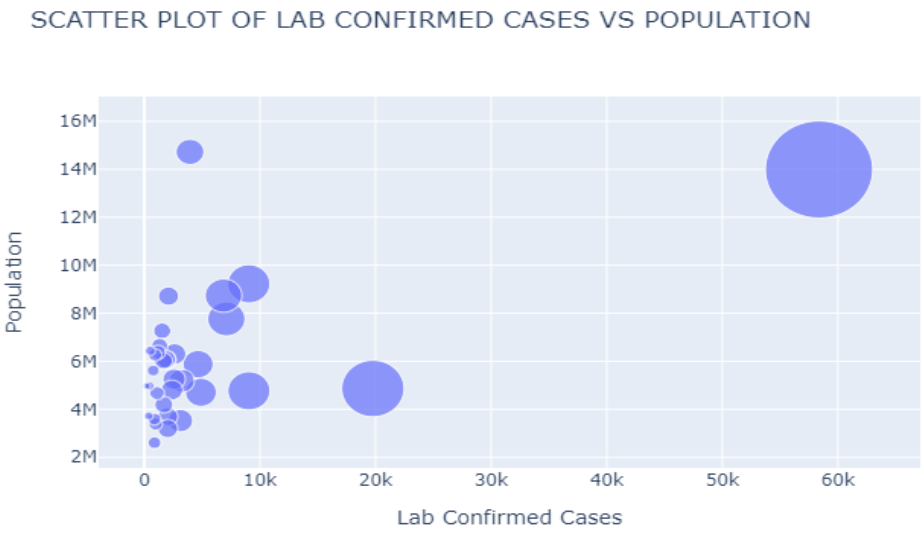

***Insights***

*This shows that there is a linear relationship between:*

- *Lab confirmed cases and Discharged cases,*
- *Lab confirmed cases and Death cases,*

*Since it is obvious from the regression plot that an increase in one leads to a corresponding increase in the other.*

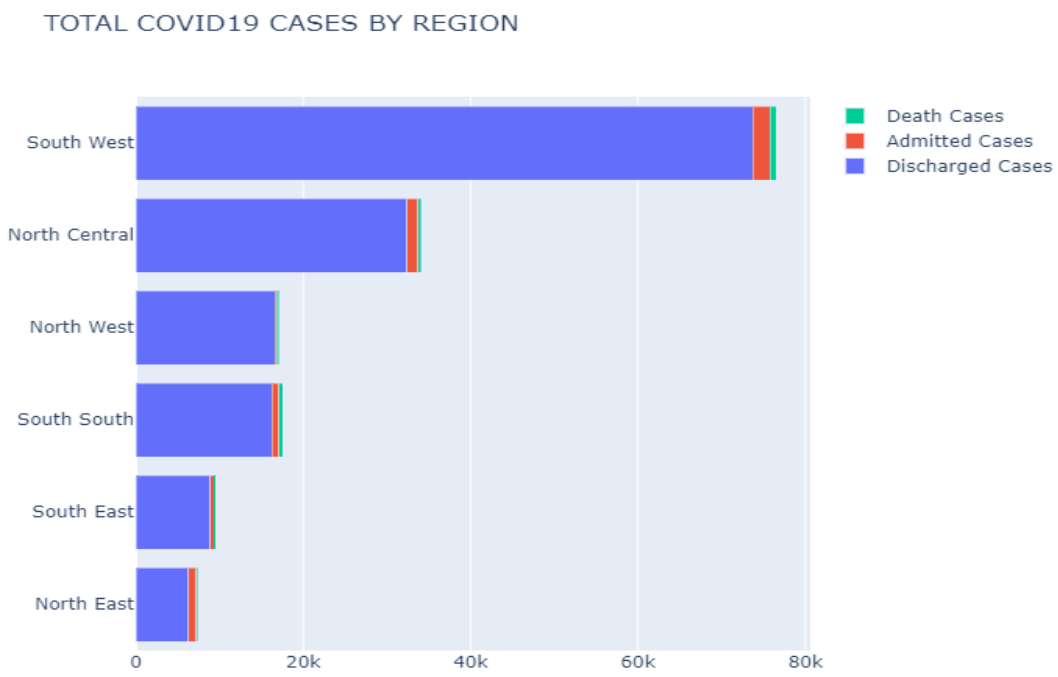
B. Lab Confirmed Cases and Population



***Insights***

*This shows that densely populated states have a higher number of lab confirmed cases while sparsely populated states have lower number of lab confirmed cases.*

7. Nigeria Region-wise case distribution




***Insights:***

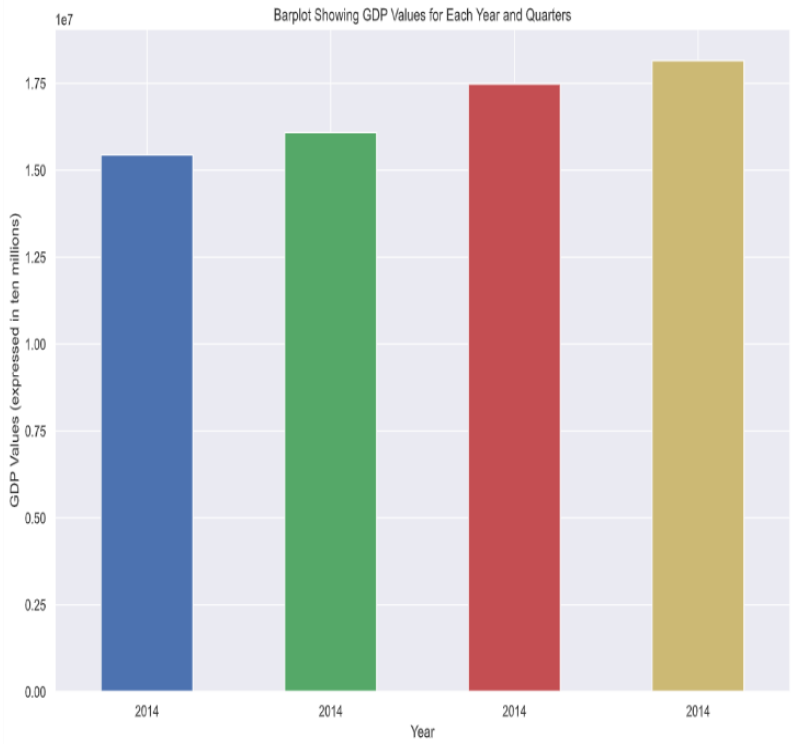
*This bar chart represents the total confirmed COVID19 Cases base on Regions.*

*From this chart, the following deductions were made:*

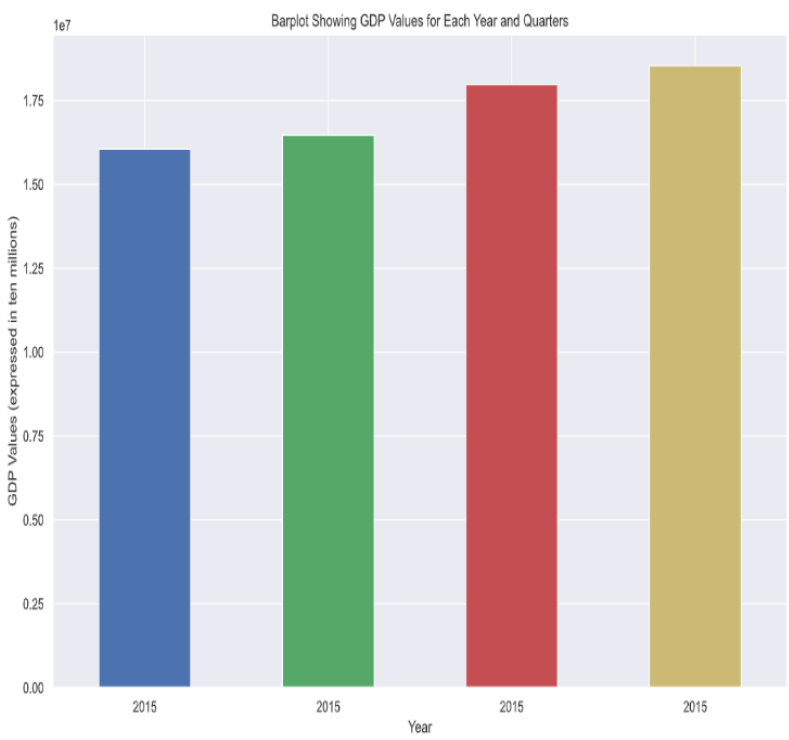
- *The **South West** region has the highest number of Confirmed COVID19 Cases which totals to 76,023 cases.*
- *While the **North East** region has the lowest number of Confirmed COVID19 Cases which totals to 7,235 cases.*

## 8. Effect of the pandemic on Nigeria's GDP

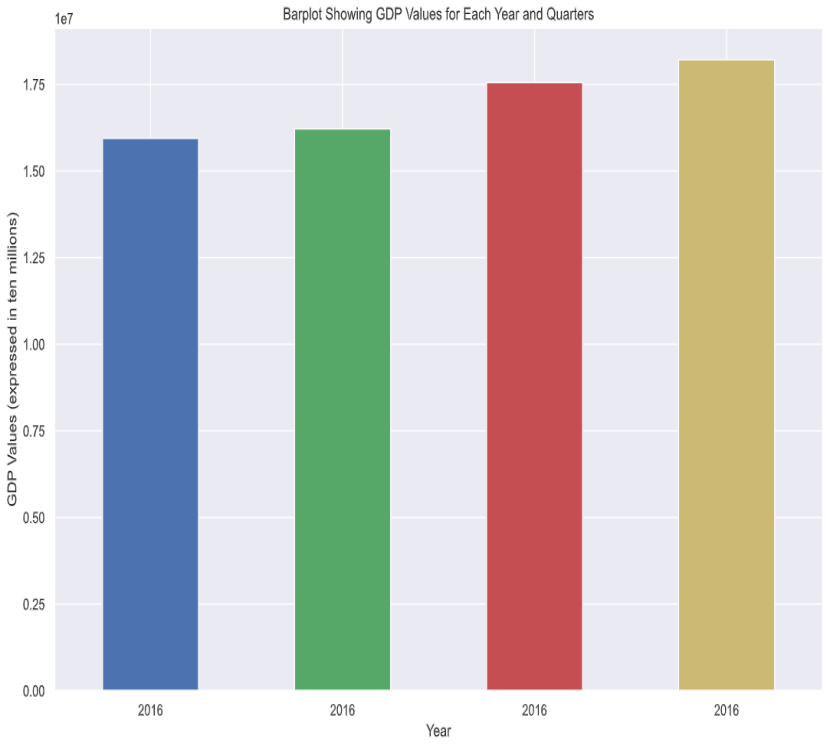
**2014**



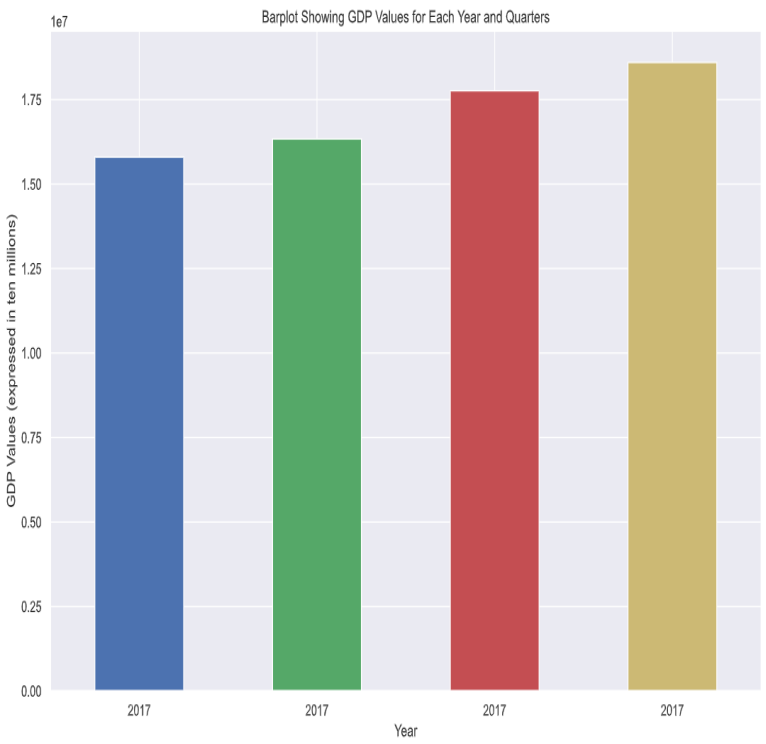

**2015**



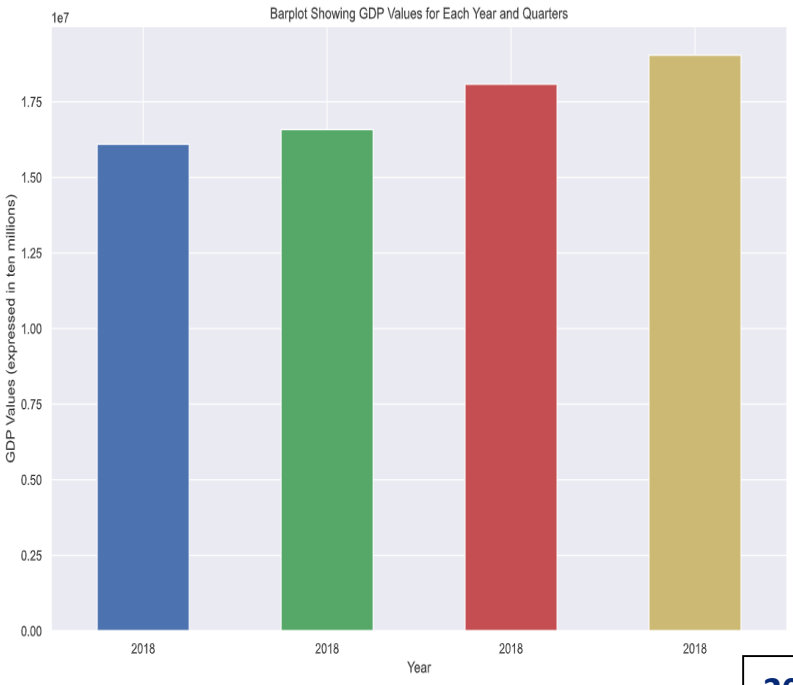

**2016**



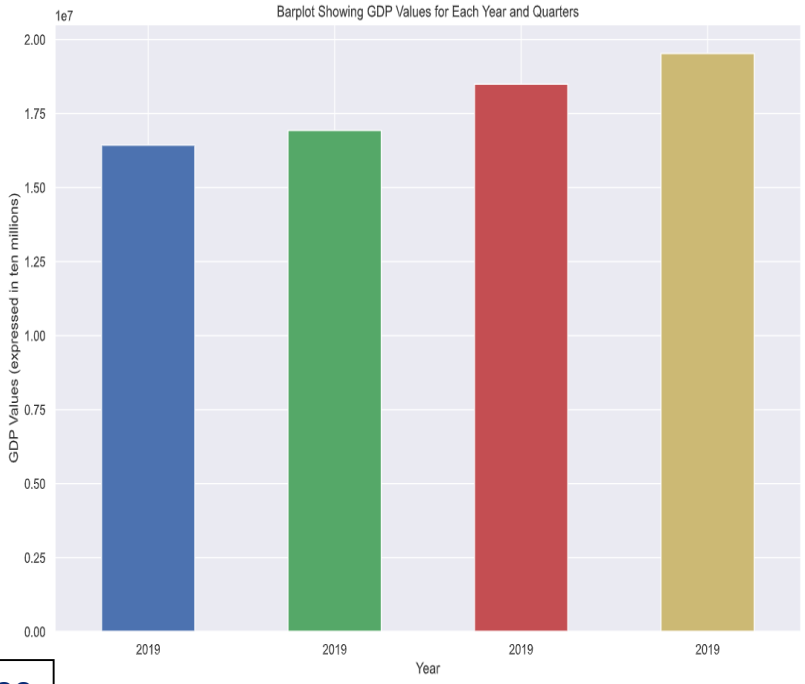

**2017**



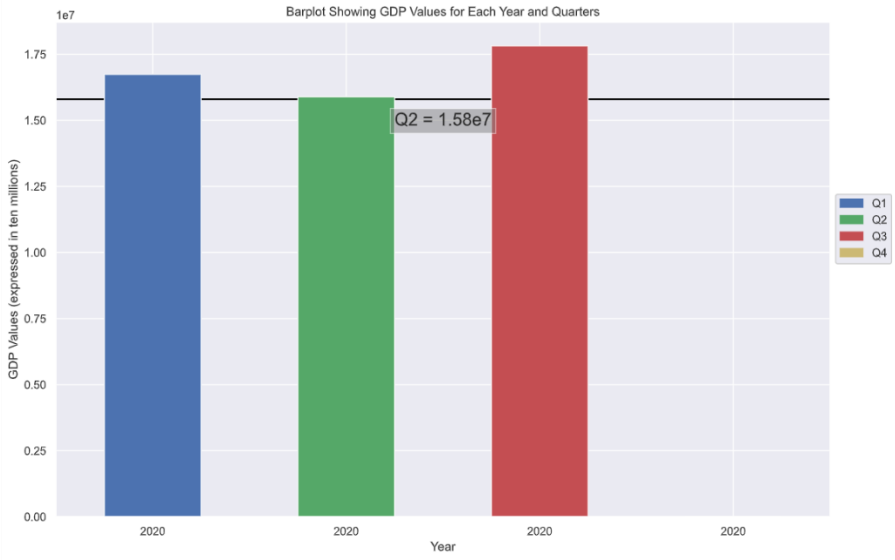
**2018**




**2019**




**2020**

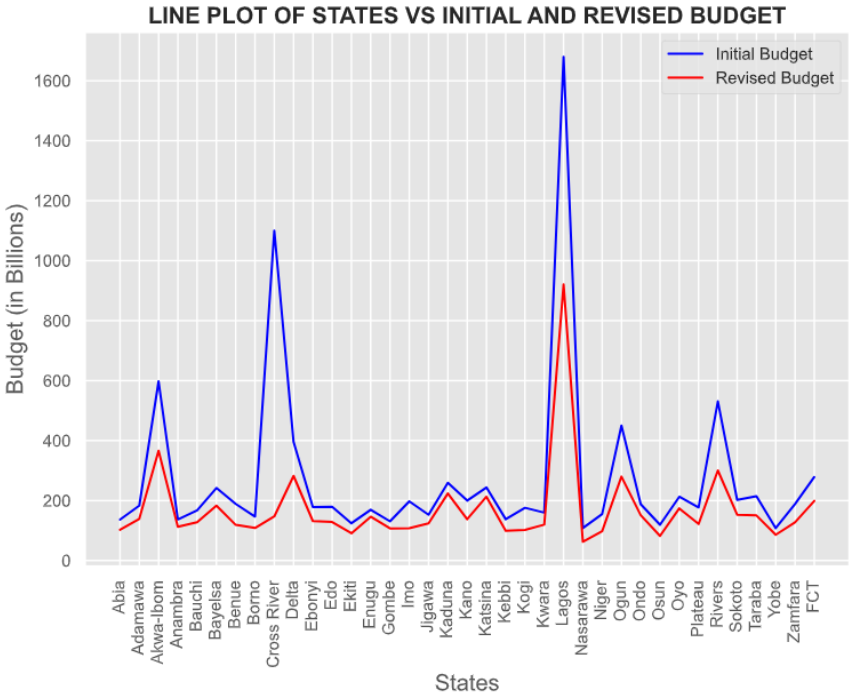



***Insights:***

*From the bar plots, the following deductions were made:*

*Year 2020 recorded its lowest Real GDP as of the time of this analysis, in its second quarter(Q2) with a value of 15.89 million which is the lowest Real GDP for Q2 recorded from the Year 2014 till 2020 and this could be associated with the effect of the pandemic on the country as at the time, since the Real GDP of Nigeria Pre-COVID19, especially during the second quarters (Q2) is higher than the present Q2 for 2020.*

## 9. Effect of the pandemic on Nigeria's Budget



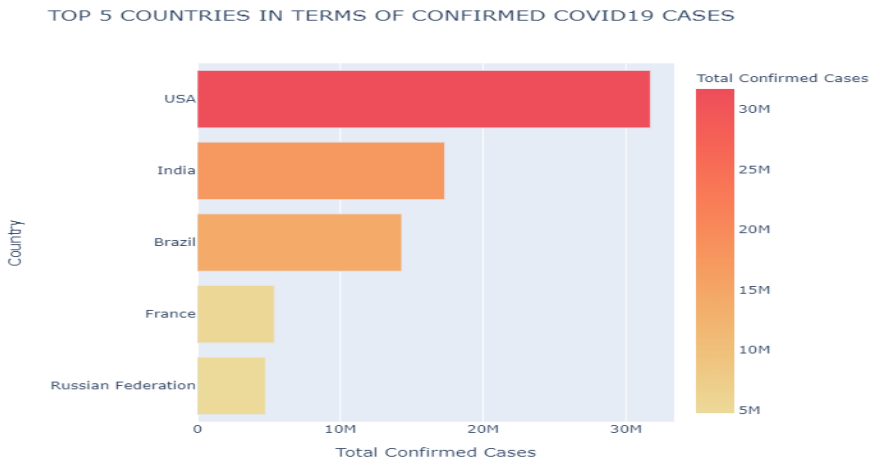

***Insights:***

*From this line chart, it is evident that all states in Nigeria had to revise their initial budget to a lower value. This is due to the effect that that COVID19 has on the economy.*

## 10. Comparing Nigeria

A. How does Nigeria's COVID19 case compare to top 5 affected countries?

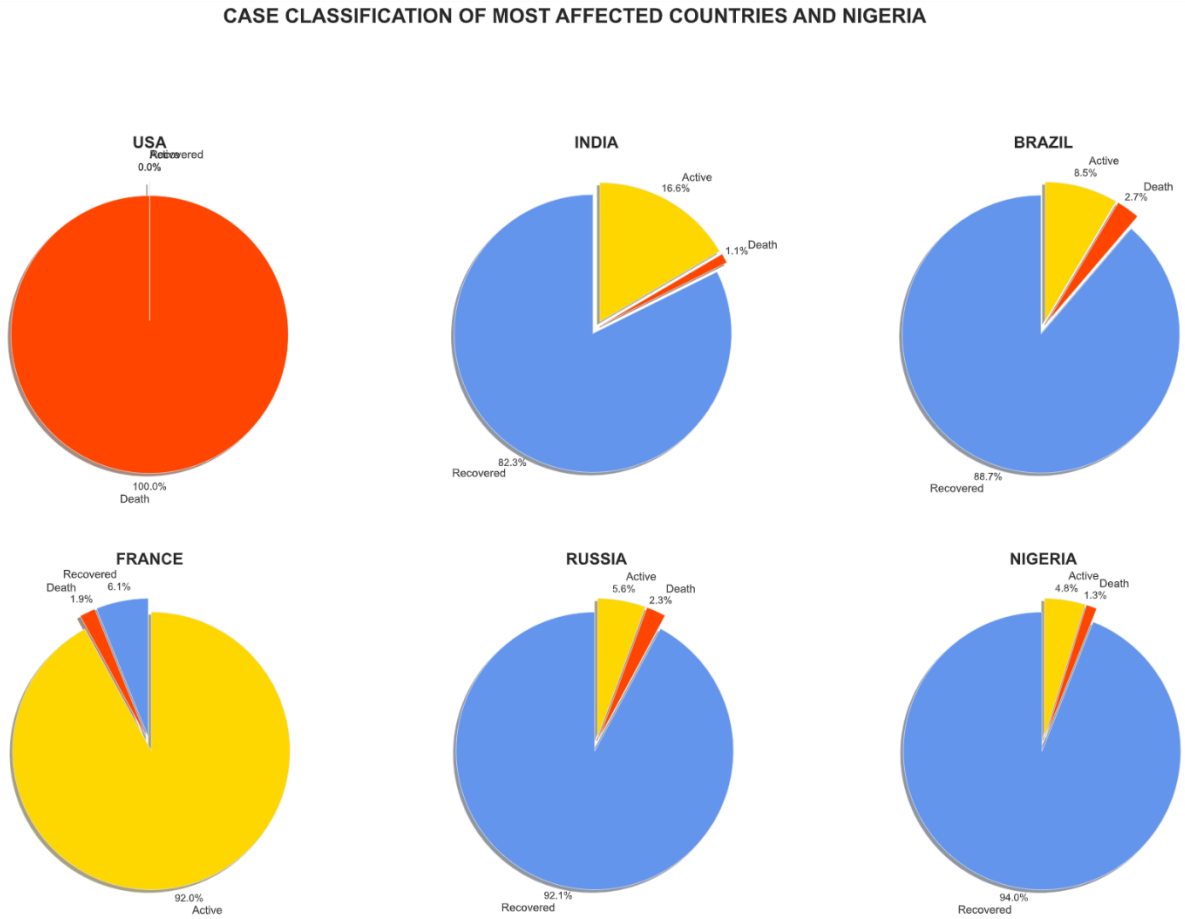
i. Top 5 COVID19 affected countries in the world




***Insights***

*The top 5 countries with the highest confirmed cases of COVID19 Cases are:* ***USA****,* ***India****,* ***Brazil****,* ***France*** *and* ***Russia***

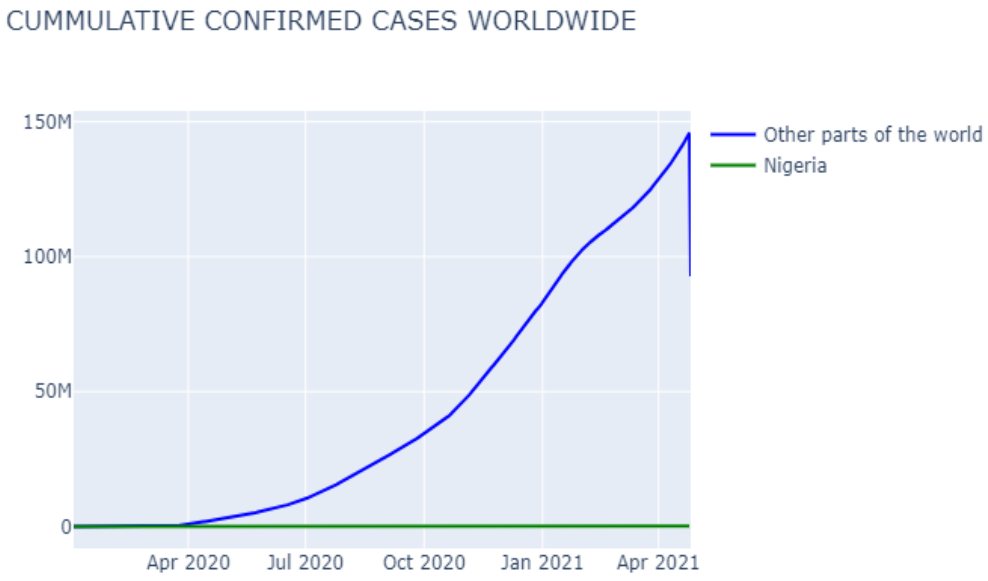
ii. Case Distribution of Top 5 Affected Countries and Nigeria using Pie Chart



***Insights***

*In this group classification, it can be seen that Nigeria has the highest recovery rate of 94.1% and lowest rate of active cases with the value of 4.7% when compared to the other top 5 affected countries. Although the death rate in Nigeria is low with the value of 1.3% but this is not the lowest when compared with the other 5 countries. Surprisingly, India actually has the lowest death rate of 1.1%.*

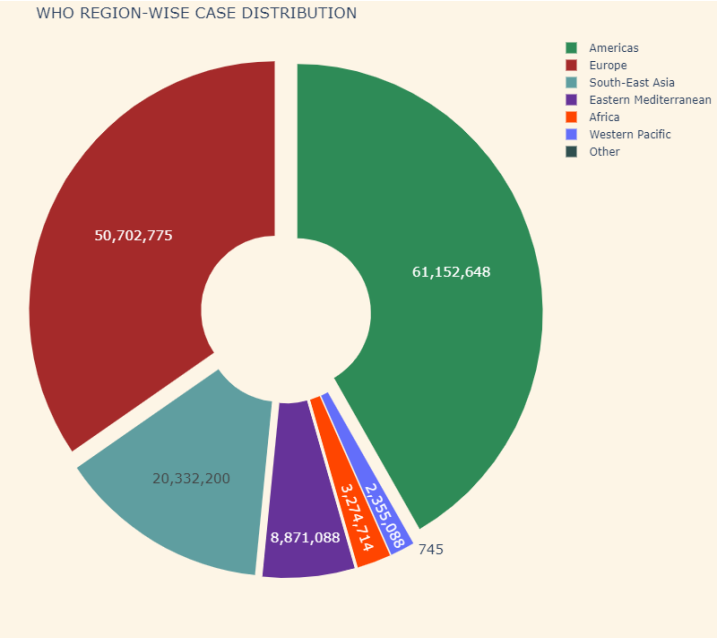
B. How does Nigeria's COVID19 case compare to the rest of the world?




***Insights***

*This line plot clearly shows that when compared to the rest of the world, Nigeria's confirmed cases do not match even one third of the world's cases. Suffice to say that most countries actually have it worse.*

## 11. WHO Region-wise case distribution

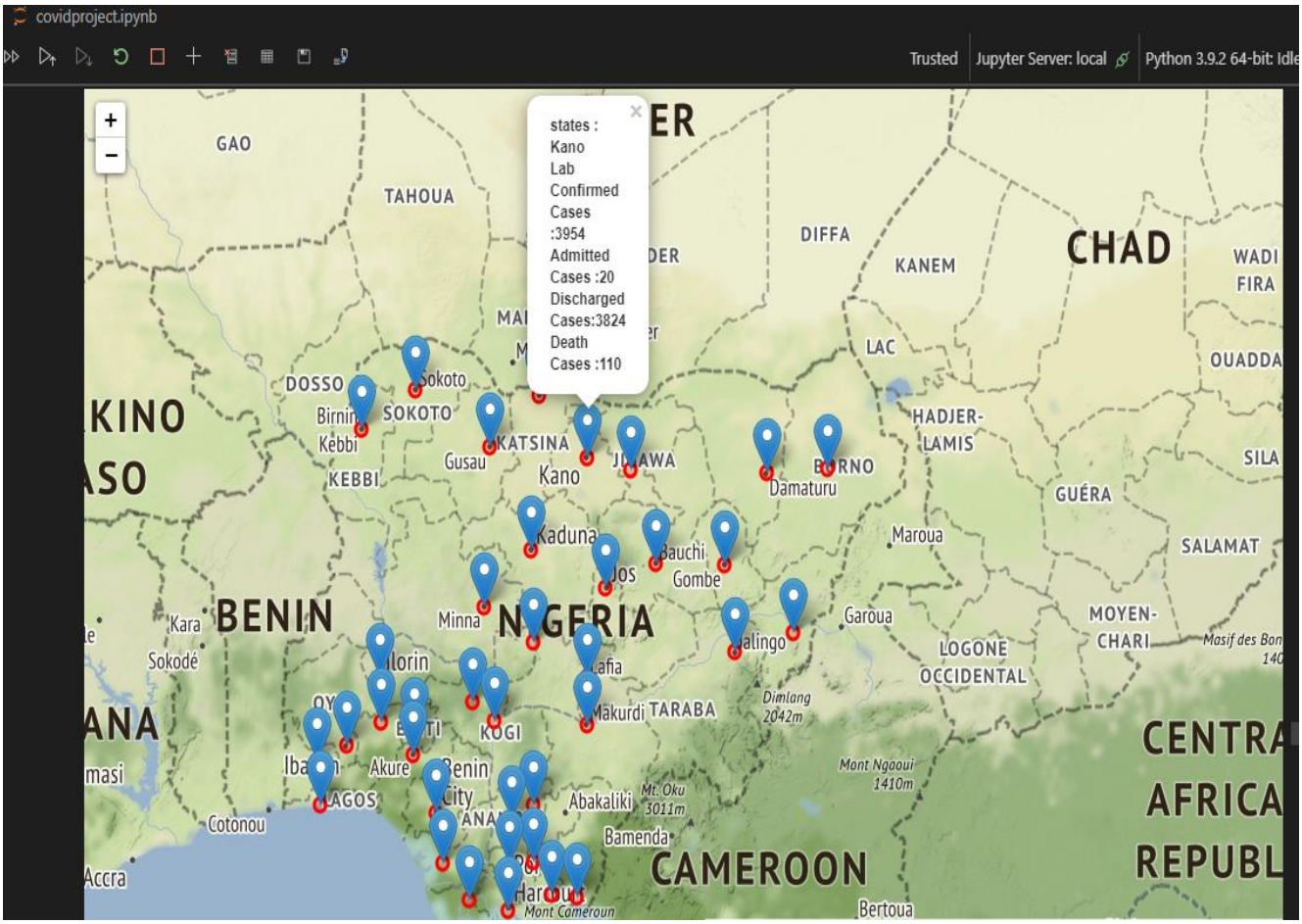



***Insights:***

*From this chart, we can see that based on the total confirmed COVID19 cases, the hardest hit regions are **Americas**, **Europe** and **South-East Asia** while the least hit regions are **Eastern Mediterrean**, **Africa**, **Western Pacific** and **Other**. Well Nigeria is in the African region which is one of the least hit regions.*

12. Map overview of Nigerian states and COVID19 Cases

This is a map visualization of Nigerian States through a map showing the various states and how they have been affected by the virus using their total confirmed, recovered, active and death cases.



# CONCLUSION

COVID-19 disrupts the globe: a typical case of unintended consequences of globalization. The flow of people aids the flow of infectious diseases (such as the Ebola virus disease and coronavirus). From a few imported cases, most nations are now battling with thousands of cases and deaths.

Although, Nigeria is recording a decline in infection rate but still there is an immense need for Nigeria to really equip adequate facilities, uplift our fragile health system and also implement immediate emergency response team to COVID19 Cases.

Also, there is a need for people to still practice COVID19 safety measures by regular handwashing, use of nose mask and avoiding crowded space.