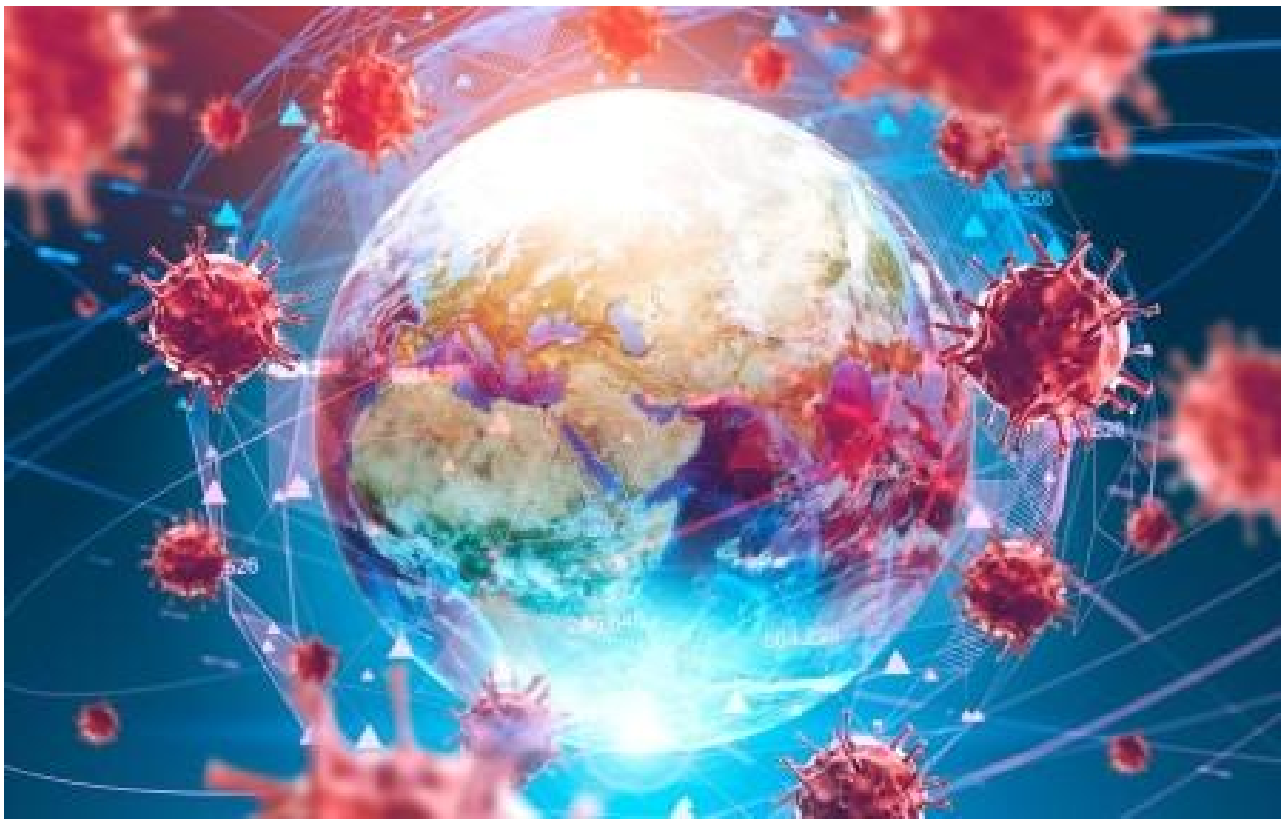
Nigeria COVID-19 Data Analysis Report Using Python



By

Daniel Akobundu

Ustacky Capstone Project

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.

This project is aimed at employing data science & analytics skills to collect data, explore the data, perform analysis, create visualizations, and generate insights on the effect of the virus on the country.
This analysis shows:

- The South-West geo-political region of the country is most affected by the virus,
- The negative effect of the virus on the Gross Domestic Product (GDP) and budget change of the country,
- The relationship between the Vulnerability index (CCVI Index) and the number of confirmed cases,
- The top states with the most confirmed cases of the virus and casualty,
- The relationship between the confirmed cases in each state and the population,
- The daily infection, recovery and casualty rate of cases in the country.

DATA OVERVIEW AND METHOD

The data for this analysis was sourced from different location and combined to perform analysis and provide insights. The data sources include:

1. The Nigeria Center for Diseases Control (NCDC) monitors the country's COVID-19 situation, and releases data on metrics across all the 37 states in the country. Data from NCDC COVID-29 official website (www.covid19.ncdc.gov.ng) was obtained by performing a web extraction/ web scraping. This was done by querying the website API and passing the data into a pandas DataFrame using pd.read_html().

2. Nigeria Community Vulnerability Index Data:

   The vulnerability index was computed by considering several factors such as socio-economic status, population density, housing type, transportation, epidemiological, health system etc., these factors are knows as themes. Each theme was broken into subthemes, and data was gathered from them to compute the overall vulnerability index score by weighing equally each theme.

- The Overall Vulnerability Index (CCVI Index) refers to the measure of the impact of the virus on a community after the virus arrives. It ranks from Very Low (0) to Very High (+1)

3. The Johns Hopkins University Data:

   The Johns Hopkins University Center for Systems Science and Engineering (JHU CSSE) publishes daily data on confirmed, death and recovered cases across different countries. The data for Nigeria was accessed from the data repository, analyzed and insight was created from the data.

4. Real Gross Domestic Product Data:

   The data on Gross Domestic Product (GDP) for Nigeria was provided with the project details. This data was used to determine the impact of COVID-19 on the economy by comparing the Real GDP (Pre COVID-19) with Real GDP (during COVID-19). The analysis was done by plotting the GDP of each year (by Quarter) to show the deference and determine the impact of the negative virus on the economy

   Resource Link - https://www.pwc.com/ng/en/assets/pdf/economic-alert-october-2020.pdf

5. State Budget Data:

   States across Nigeria reduced their initial budget due to the impact of COVID-19 on the economy. The budget data was provided with the project details. This data was analyzed to determine the negative impact of the virus on the economy. This analysis was done by calculating the difference between the initial and new budget of each state and calculating the percentage difference.

ANALYSIS AND RESULTS

1. Fig 1-3 - Maps of Top 10 States Confirmed, Discharged/Recovered cases COVID-19 cases in Nigeria.

Fig. 1 shows Lagos State has the highest cases of COVID-19 confirmed, discharged and death cases, recording almost 58,393 confirmed cases, 66,990 discharged and 439 death cases as at 1st May, 2021. Other states have a significantly lower confirmed case with FCT being the second highest recording up to 19,776 confirmed cases, 19,095 discharged and 515 death cases, and Ondo been the 10th highest recording up to 3,242 confirmed cases, 2,074 discharged and 63 death cases of COVID-19.

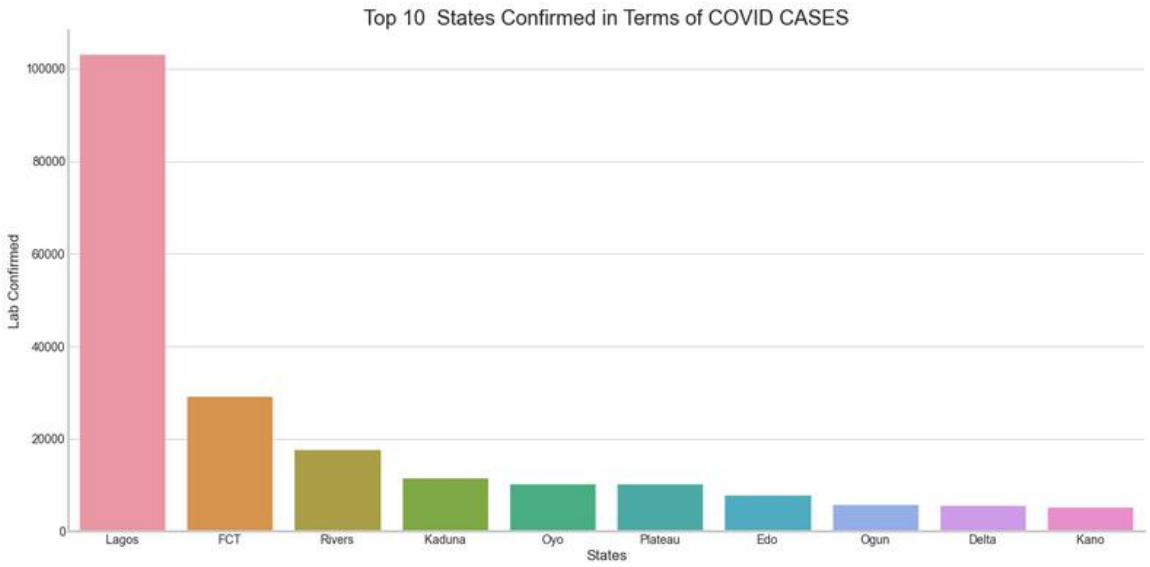


Fig. 1

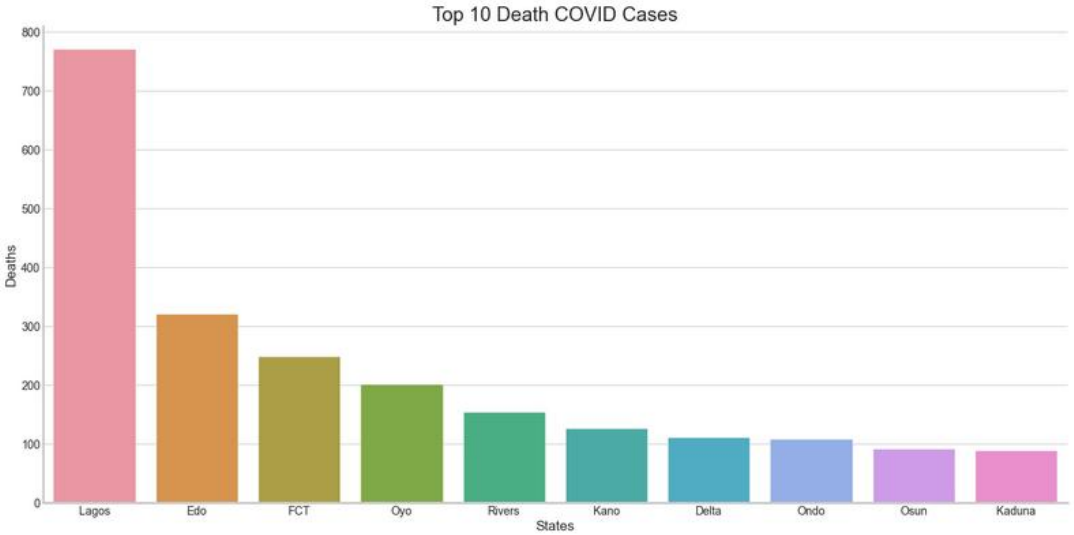


Fig. 2

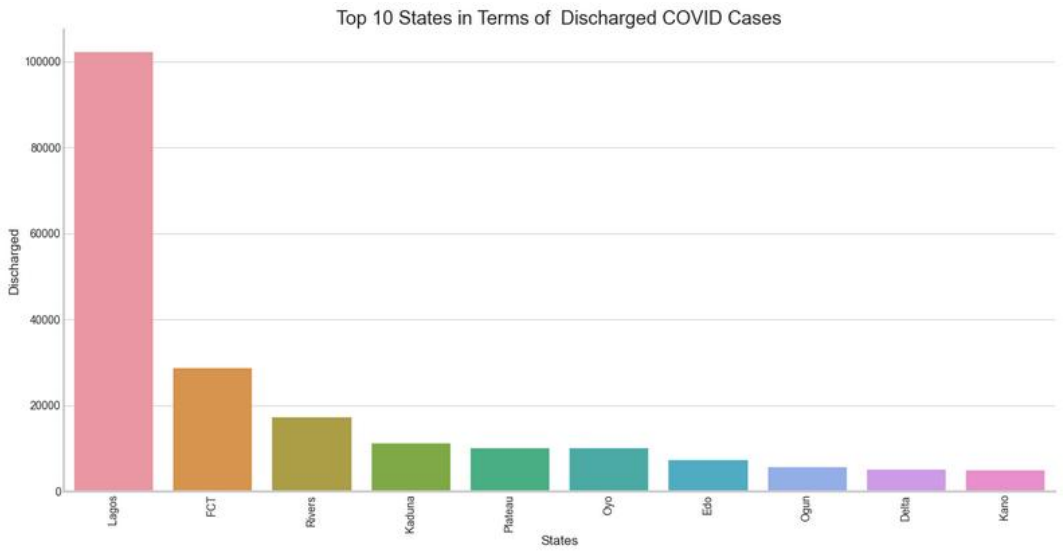


Fig. 3

2. Fig. 4 shows the daily infection rate of the virus from 28th February, 2020 till date. There are three (3) peak periods, early 2021, towards summer of 2021 and the highest rate in Christmas period of 2021

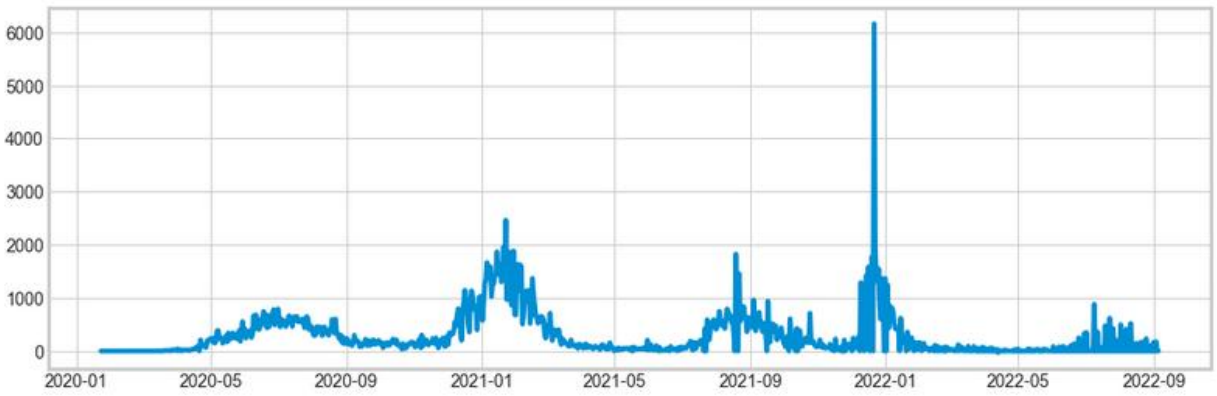


Fig. 4

3. Fig. 5 Shows the relationship between the overall CCVI Index (Vulnerability Index) and the number of confirmed cases of COVID-19 in the top 10 infected states in Nigeria. The vulnerability index is a measure of how susceptible a state is to the virus. The maps show no direct relationship between CCVI Index and the number of confirmed cases in each state. Lagos States, which has the highest number of confirmed cases (58,393) of the virus has a CCVI Index of 0.0. Kaduna and Kano state with a CCVI Index of 0.7 and 0.6 respectively, which should be more susceptible to the virus and hence, have more cases, both have less than 10,000 cases each.

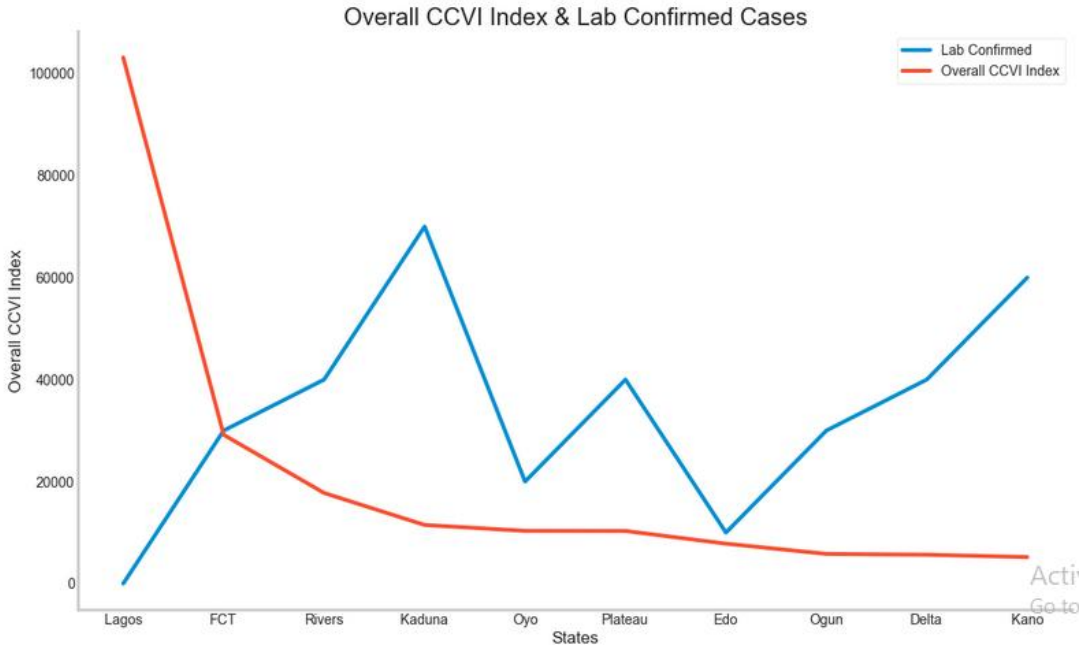


Fig. 5

4. Fig. 6 Shows the regression plot between the number of confirmed cases and population density of each states.

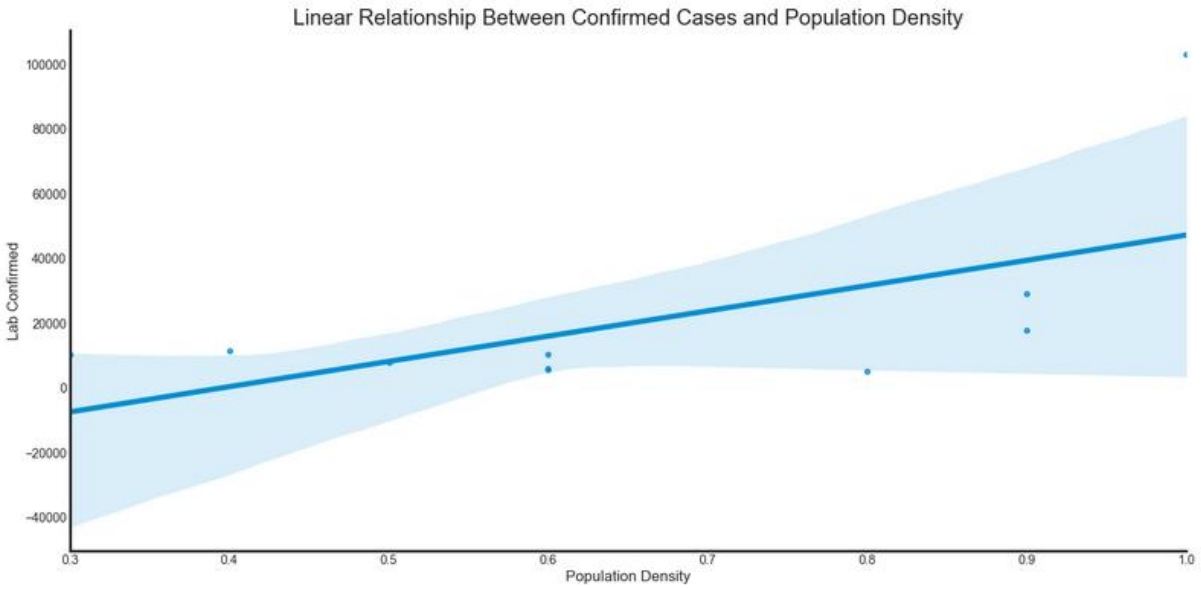


Fig. 6

5. Fig. 7 is a map of the confirmed, recovered and death cases of COVID-19 in Nigeria. The map shows a continued growth in the confirmed cases and a sharp drop in the recovered cases of the virus around August 2021. The death rate indicates an effective approach to curtailing the virus.

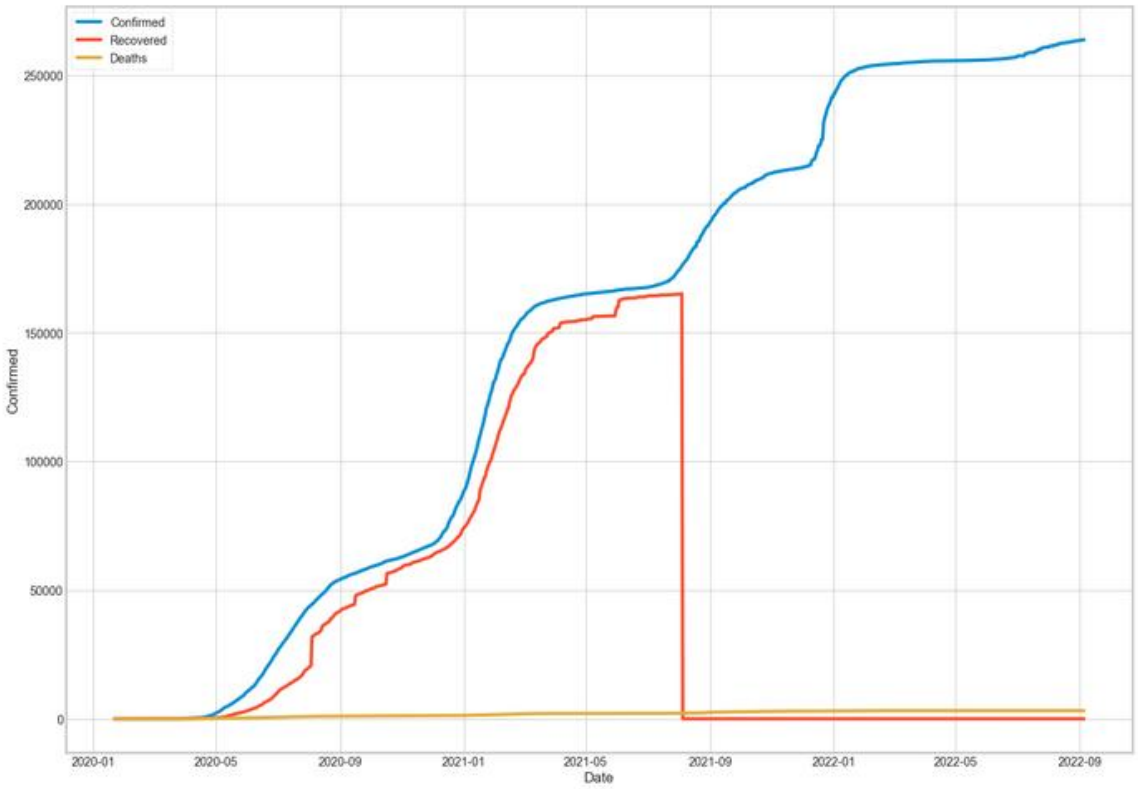


Fig. 7

6. Fig. 8 Shows the Gross Domestic Product of the country between 2014 – 2020. This shows the second and third quarter of 2020 (the COVID period) has the lowest GDP since 2014.

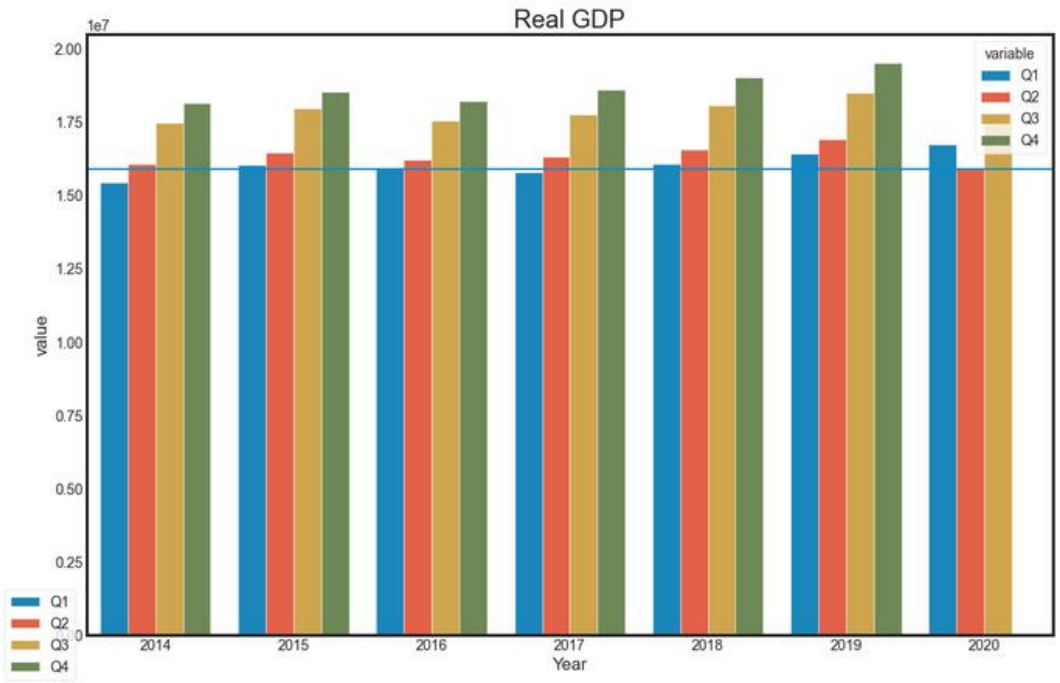


Fig. 8

CONCLUSIONS

The purpose of the project is to analyze COVID-19 data and provide insight on the data. Analysis of the data used for this project shows different results which has led to conclusions that:

1. Lagos state and FCT are the top two states with the confirmed, recovered and death cases of COVID-19 virus in the country,
2. The South-West region of Nigeria has the highest number of confirmed cases in the Country.