Nigeria COVID-19 Data Analysis using Python

PRESENTED BY OLORUNTOBA IPINYEMI

Table of Contents

1. INTRODUCTION

The World Health Organization (WHO) was formerly notified of some cases of pneumonia in Wuhan City on 31 December 2019; by January, there were 59 confirmed cases. The cause of this acute respiratory syndrome was later known as COVID-19, which was a novel virus, and by 15 January 2019 there had been six deaths in Wuhan City, and 12 were in critical condition.

There were variations in the effect COVID-19 had on different countries but all in all, the normal way of doing things was drastically changed and affected and during these infectious waves, countries were on lockdown to be able to reduce the spread of the disease.

On February 27, 2020, the first confirmed case was reported in Nigeria, an Italian national in Lagos, and the second confirmed case was reported on 9 March 2020 in Ewekoro, Ogun State, a Nigerian citizen who interacted with the Italian National.

Nigeria is a fast-growing nation with the largest economy and population in Africa. COVID-19 had a devastating effect on Nigerians' livelihood because of a 5- week of lockdown, economic downturn, and hike in food prices and other essential commodities.

The Objective of this analysis is aimed at utilizing skills in python, data science, and analytics to collect data, explore data, perform analysis, create visualizations, and generate insights on the effect and impact of the COVID-19 pandemic in Nigeria.

2. OBJECTIVES

The objectives of this project are to understand the data collection process, in this case, web scraping, and importing from data sources

- Understand the data cleaning and manipulation process.
- Develop data wrangling skills & data intuition.
- Know how to ask the right questions & find ways to provide answers.
- Develop visualization skills through the use of open-source libraries in Python.
- Generate insights from analysis.
- Report writing

3. DATA COLLECTION AND METHODOLOGY

Data used for this analysis was sourced from different channels using various data collection skills like web scrapping and directly reading CSV files.

- The Nigeria Centre for Diseases Control (NCDC) monitors the country's COVID-19 situation and releases data on metrics across all 37 states in the country. Data could not be scrapped from NCDC COVID-19 official website, because at the time of this analysis, the web page was down. I downloaded it from the provided GitHub repository into my local drive and loaded the data from there.
    - COVIDNG = pd.read_csv("covidnig.csv")
- Data was collected from Johns Hopkins University Center for Systems Science and Engineering (JHU CSSE) publishes daily data on confirmed, death, and recovered cases across different countries. Nigeria's daily data was accessed from its repository , Nigeria's data was extracted from the global data and derived
    - related insights. (details in the submitted code notebook)
    - (DF_CONFIMED_DAILY = pd.read_csv("https://raw.githubusercontent.com/CSSEGISandData/COVID-19/master/csse_covid_19_data/csse_covid_19_time_series/time_series_covid19_confirmed_global.csv")

- DF_RECOVERED_DAILY =pd.read_csv ('https://raw.githubusercontent.com/CSSEGISandData/COVID-19/master/csse_covid_19_data/csse_covid_19_time_series/time_series_covid19_recovered_global.csv')

- DF_DAILY_DEATHS = pd.read_csv("https://raw.githubusercontent.com/CSSEGISandData/COVID-19/master/csse_covid_19_data/csse_covid_19_time_series/time_series_covid19_deaths_global.csv")

- 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 known 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. This data was got from the COVID-19 external data.

- Real Domestic Gross Product Data
  was used for this analysis and this data was on the Real Domestic Gross Product(GDP) data for Nigeria
- State Budget Data was used in this analysis as the data provides information on the reduced subsequent budget because of the impact of COVID-19 on the economy. The

percentage difference in the budget was computed and used to visualize the effects of COVID-19 on the budget of the affected states

4. ANALYSIS AND RESULT

Figures 1-3 below, show the top 10 confirmed, recovered, and death COVID-19 state cases in Nigeria. Lagos state shows the highest number of confirmed cases, discharged cases, and death cases.

Other states in the top 10 states in these categories of confirmed order are FCT, Kaduna, Plateau, Oyo, Rivers, Edo, Ogun, Kano, and Delta State.

Figure 1

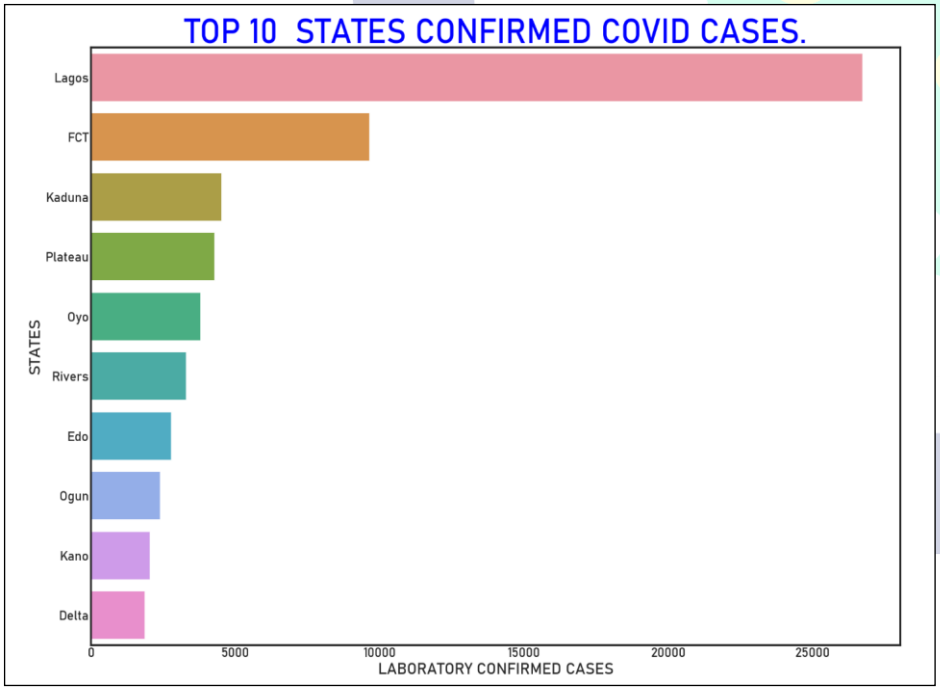
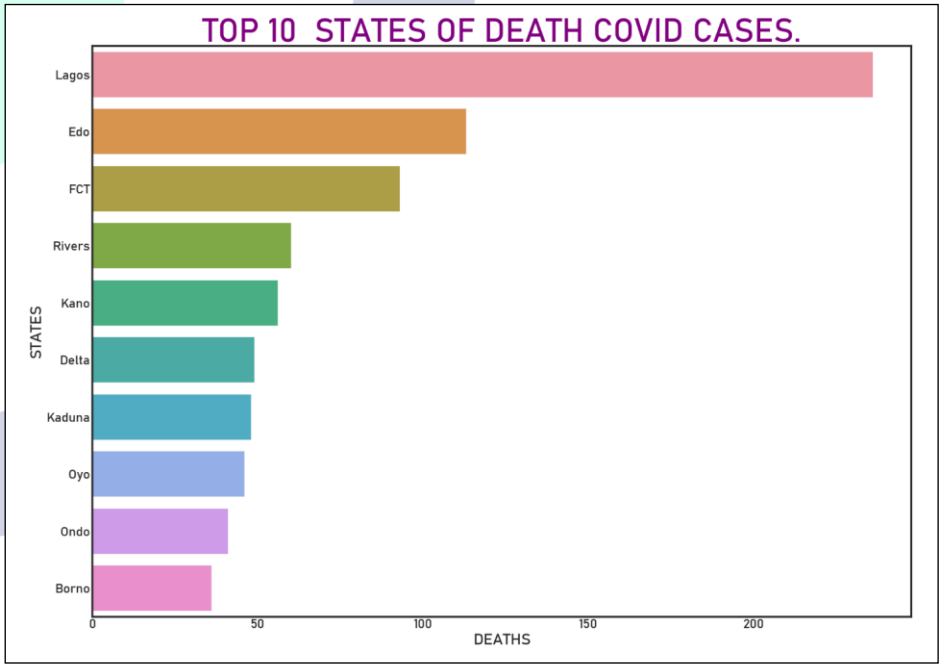
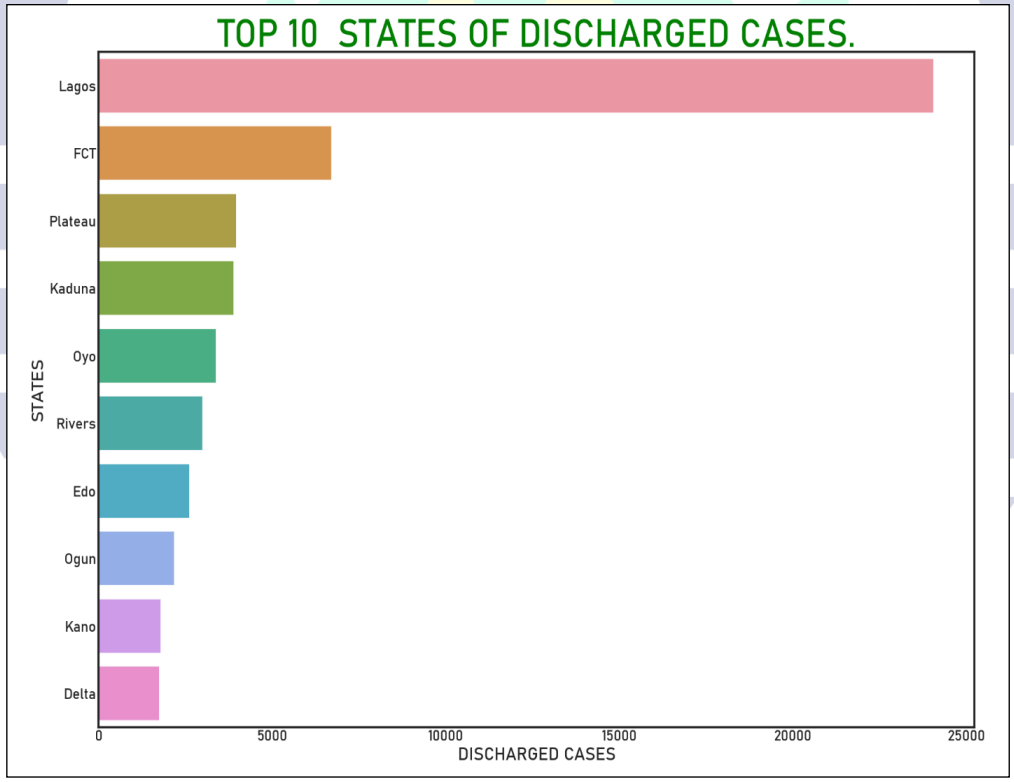


Figure 2



Edo state, even though they had the 7th largest confirmed cases had the second largest death cases. This might be due to some other factors like vulnerability. Lagos has the highest numbers of discharged cases , followed by the FCT.

Figure 3



There is steep rise in the daily recovery record up till september 2021 and the record becomes flat. Might be due to no records or a decline in confirmed cases. It was observed that there were no more data in recovered cases beyond September 2021, however, there were still reports of some confirmed cases.

Figure 4

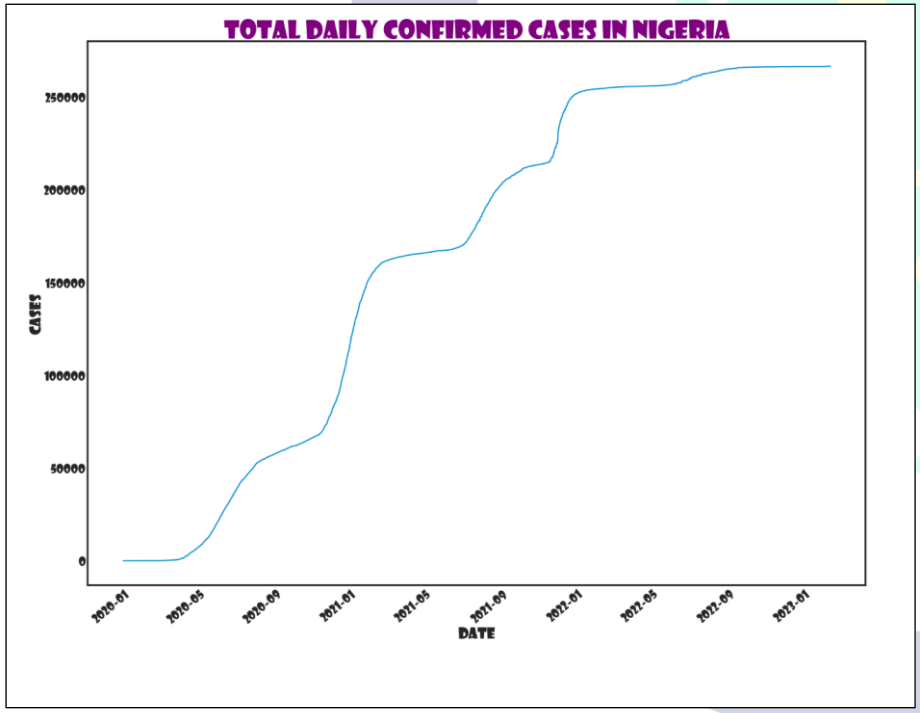


Figure 5

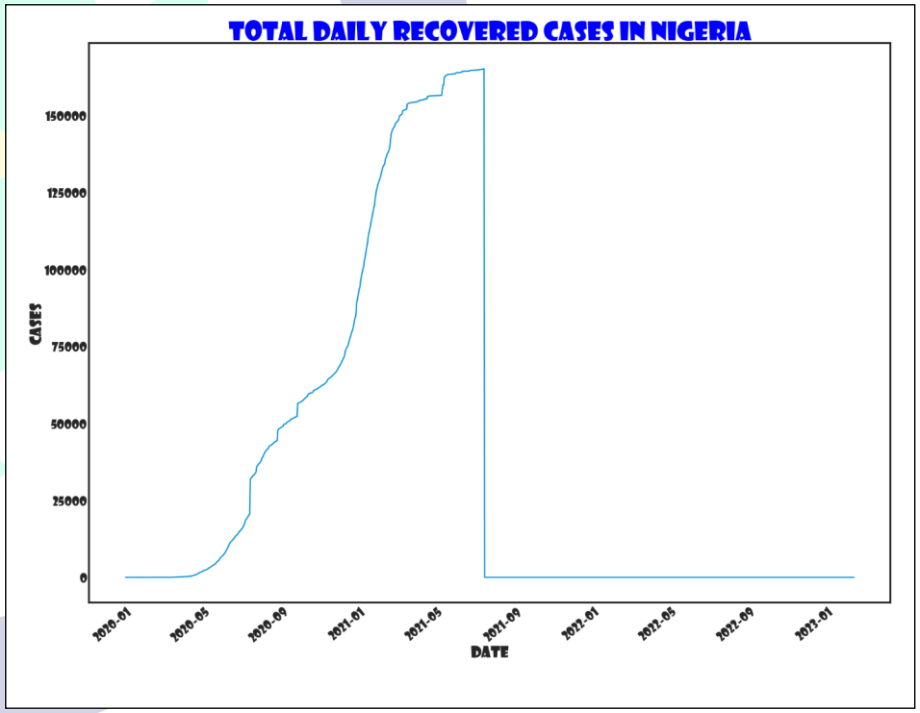
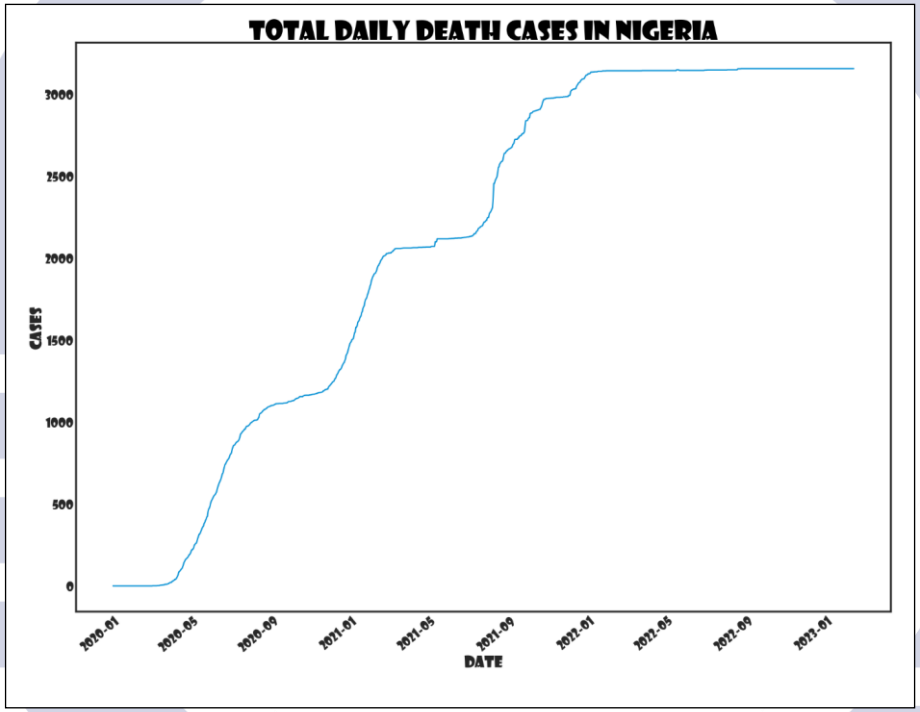


Figure 6


TOTAL DAILY DEATH CASES IN NIGERIA
3000
2500
2000
1500
1000
500
0
CASES
2020-01
2020-05
2020-09
2021-01
2021-05
2021-09
2022-01
2022-05
2022-09
2023-01
DATE

We have plots showing the rate of COVID-19 infection, recovery, and death on figure 7, and the rate of infection on a single plot in figure 8. Infection cases in Nigeria started around February 2020 which was the beginning of the first wave in the country, a decline in the rate of cases was noticed around September 2020. The second wave hit in January 2021 which was twice the cases recorded during the first wave. A third wave was noticed around August/September 2021. There was a short period of decline in cases and another peak in January 2022, afterward there was a large decline in reported cases.

Figure 7

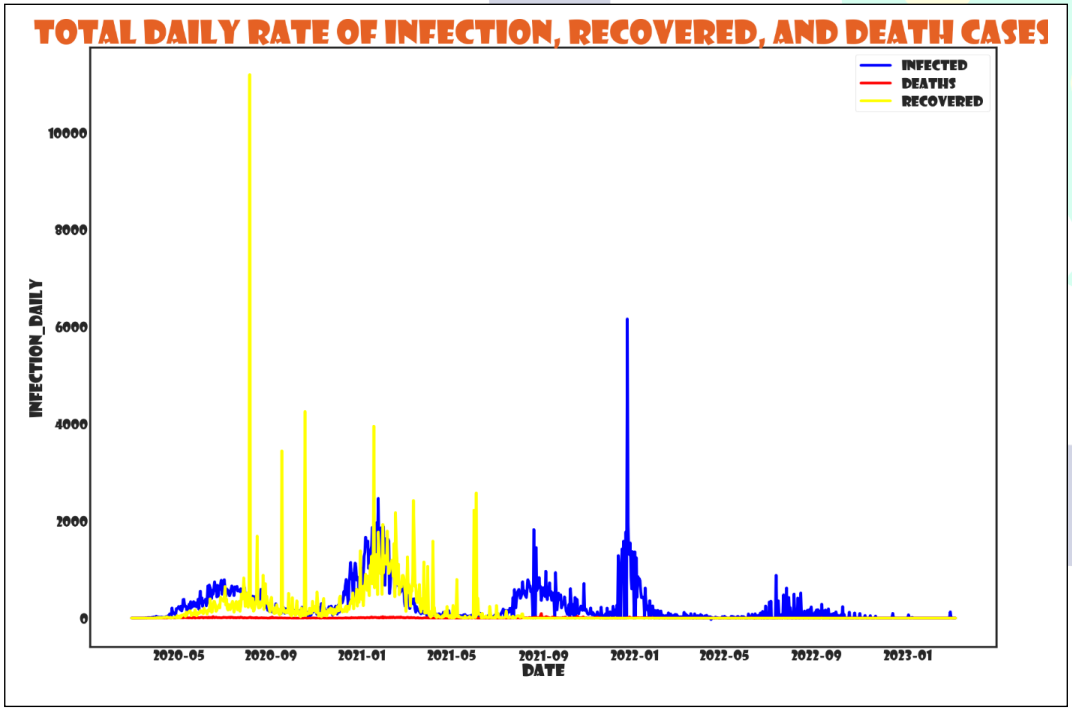
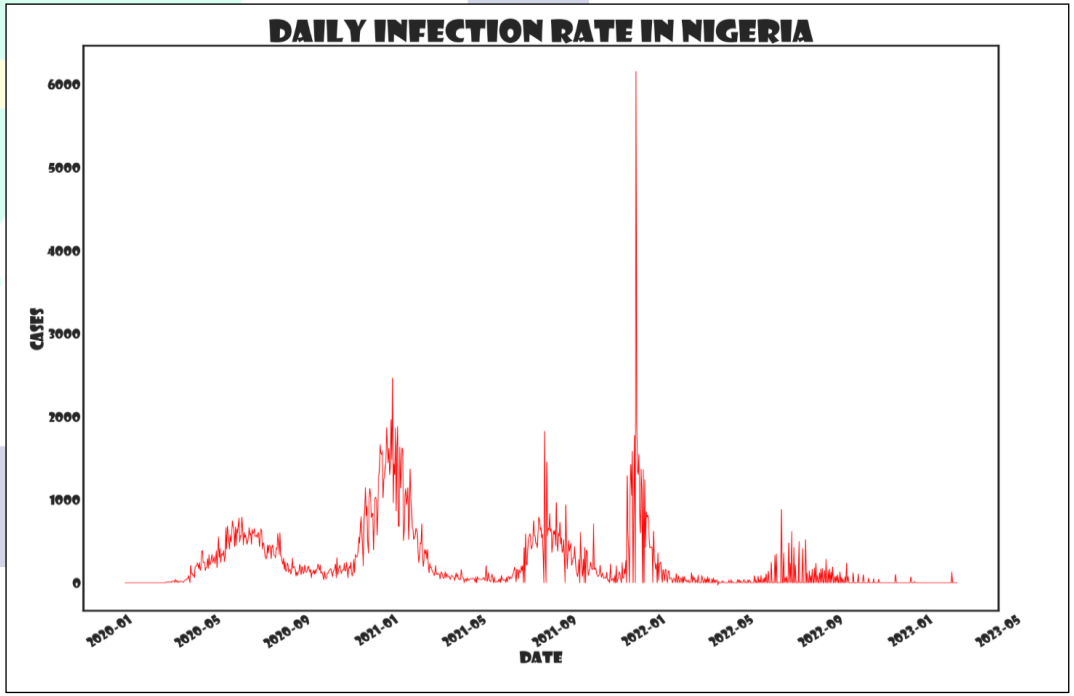
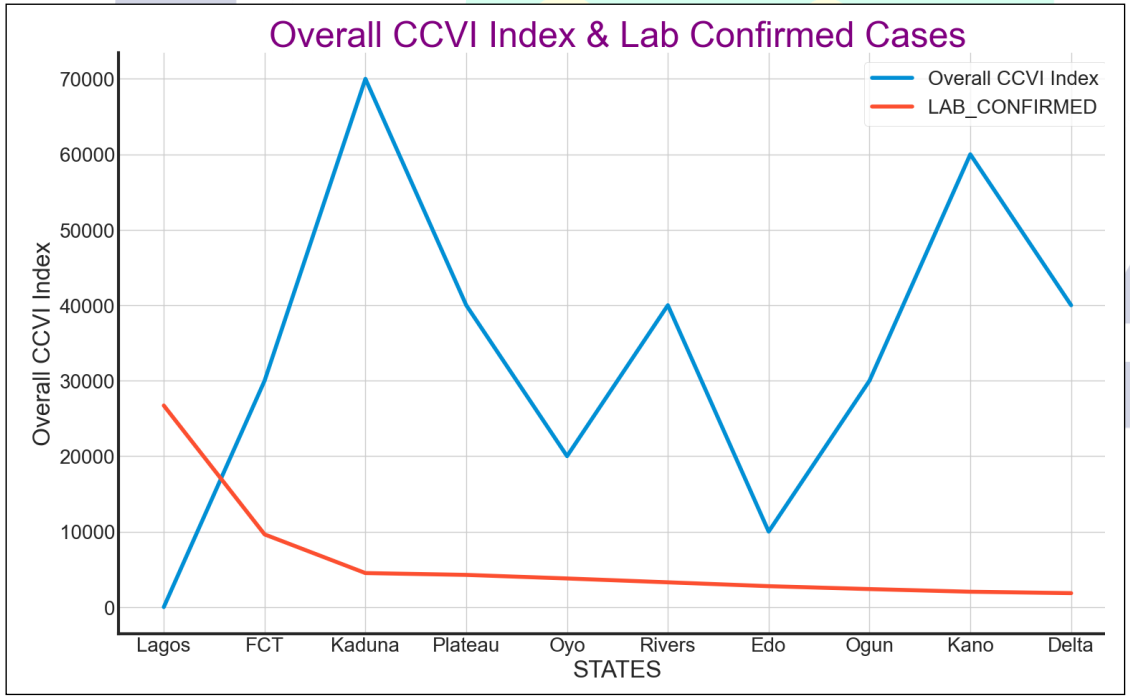


Figure 8



There is practically no direct correlation between the vulnerability index and confirmed cases. There are so many factors affecting the rate of infection beyond the vulnerability index. Lagos state has a vulnerability index of 0 yet has the highest number of lab-confirmed cases (26,708). Some other factors might be responsible for this high rate of infection in Lagos like Age, transportation availability, population, etc. Kaduna and Kano have a vulnerability index of 0.7 and 0.6 respectively, yet has lab-confirmed cases of 4504 and 2032 respectively. This shows further that other factors beyond the vulnerability index are more responsible for the spread of COVID-19.

Figure 9



The plot below shows that population density has more effect on the Lab-confirmed cases as there seems to be a positive correlation between these two factors. Low Age index score results to more confirmed cases and vice versa. Age index score has a negative corellation to confirmed cases.

Figure 10

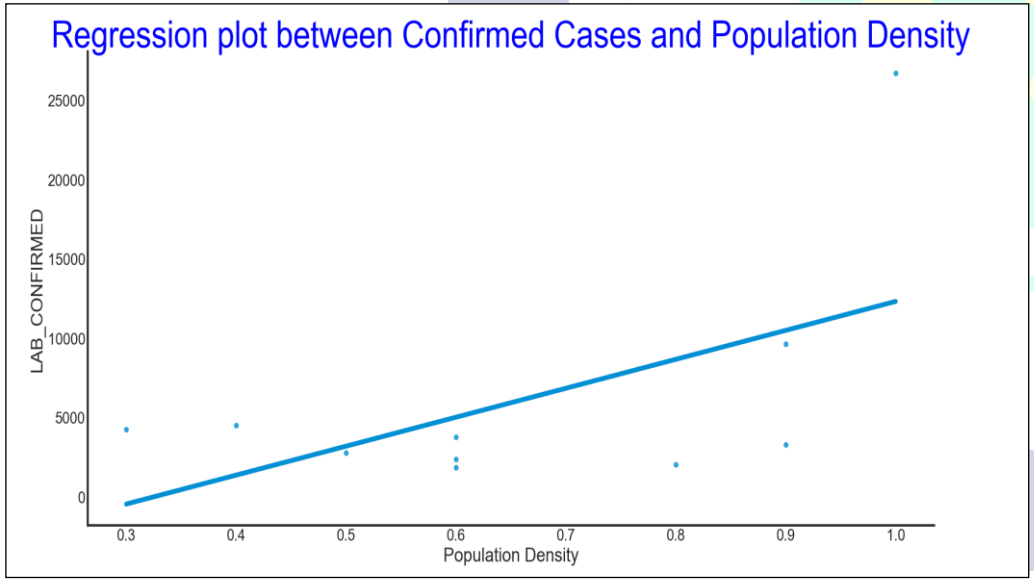
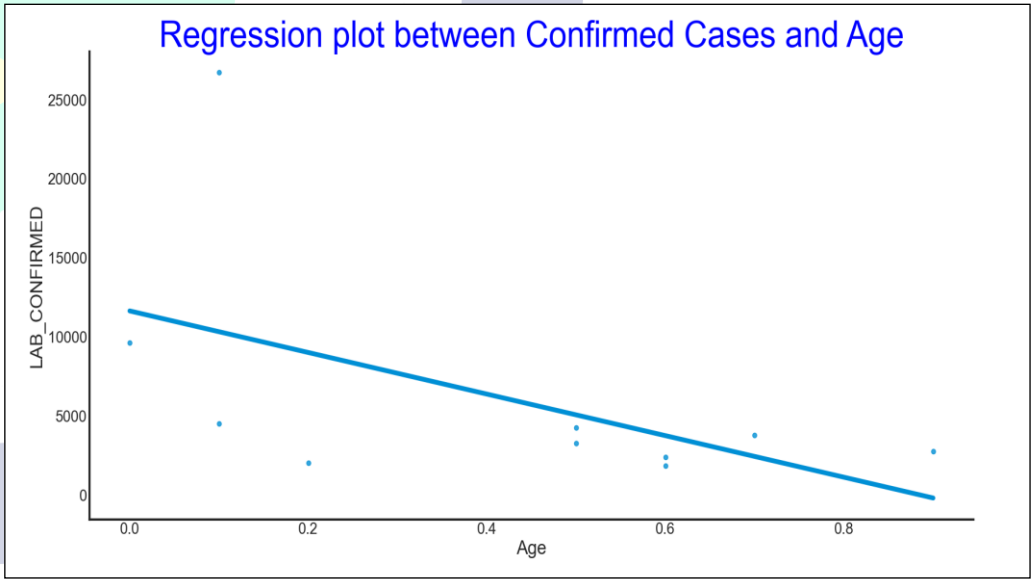


Figure 11



The plot below shows the real GDP between 2014 to 2020. This show pre-Covid to COVID-19 periodConcentrating on Q2,the value of real GDP was most reduced in 2020, although there was a gradual increase from 2014, the value dropped to the lowest in 2020.

Figure 12

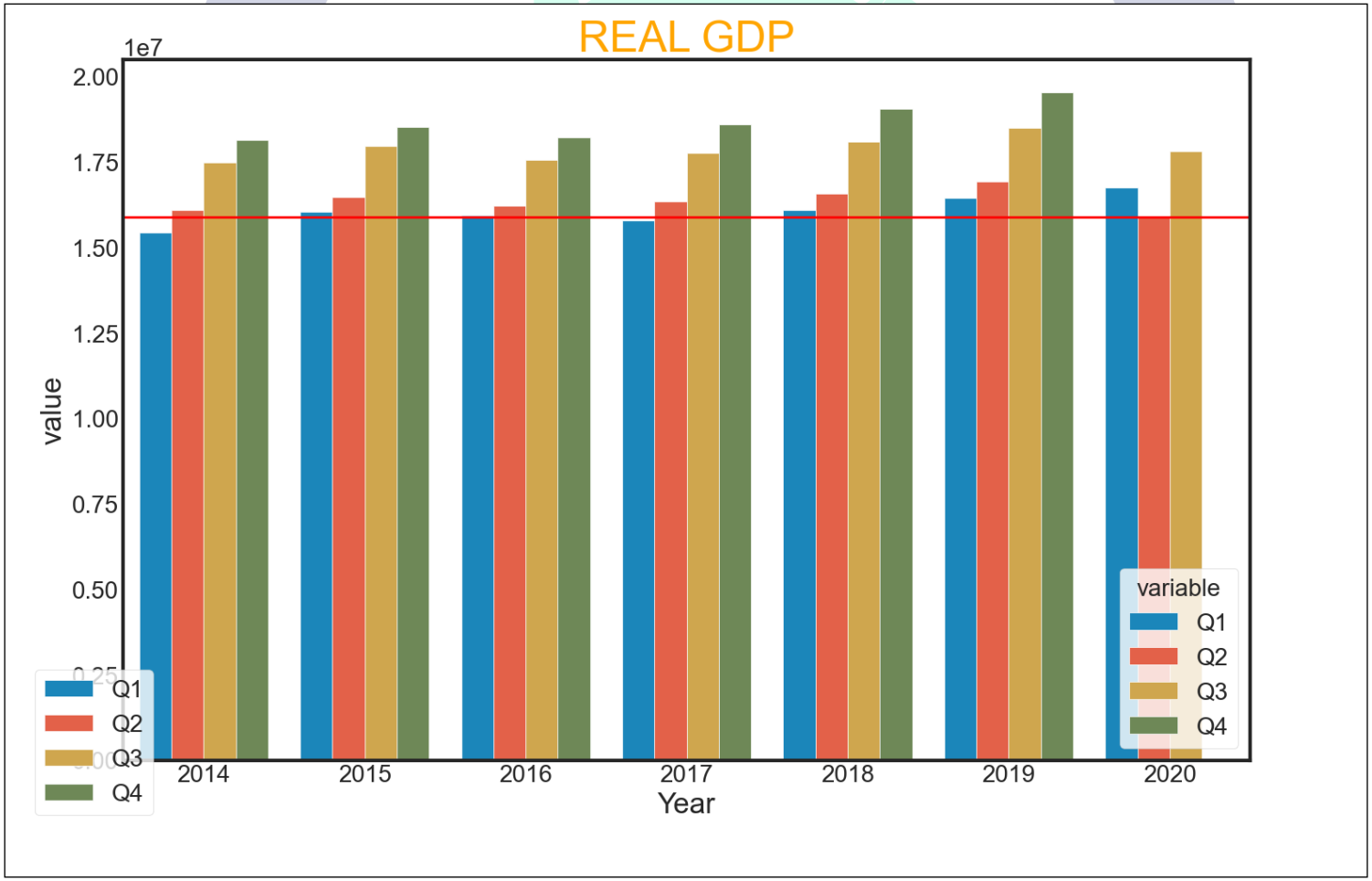
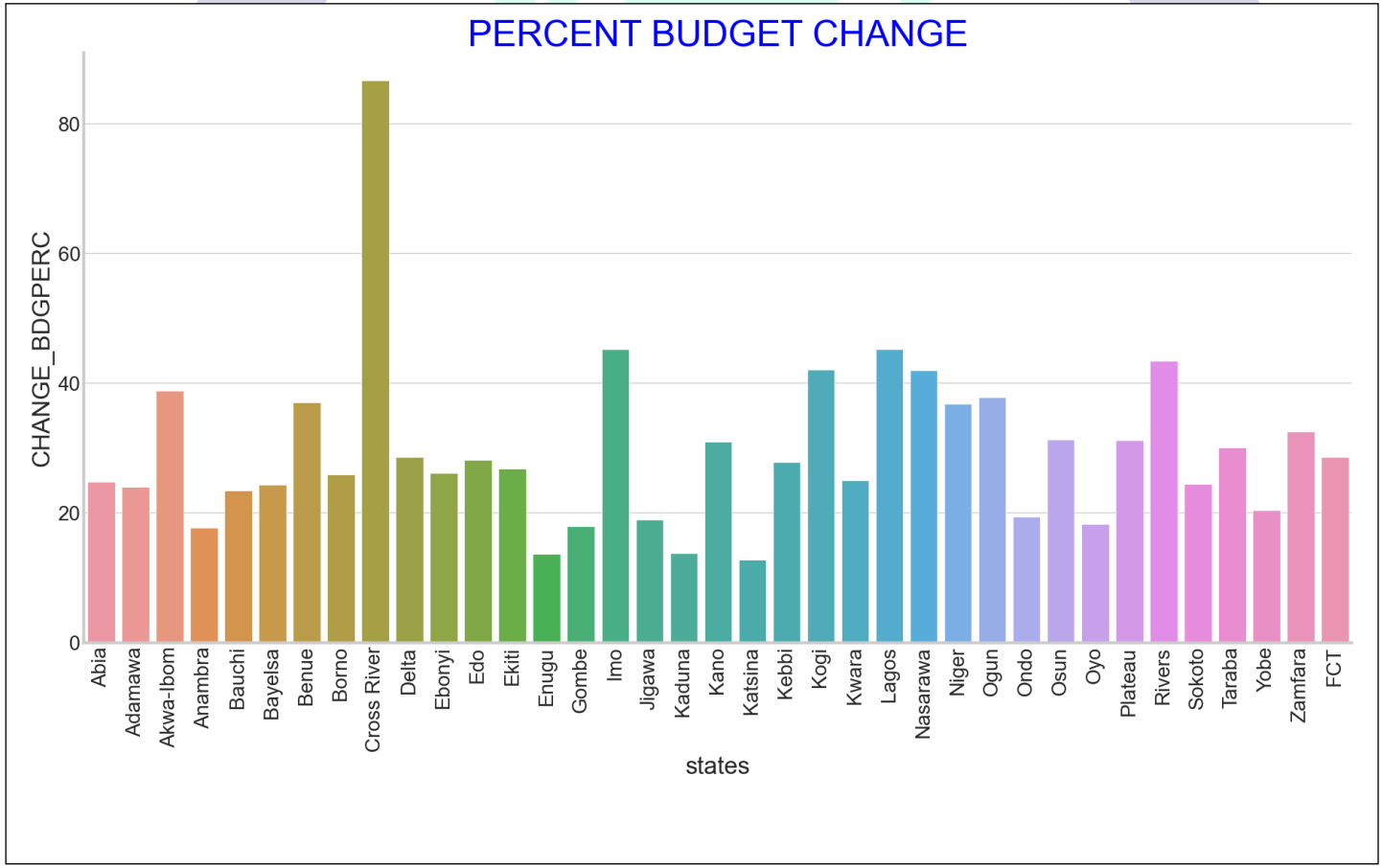


Figure 13 shows the percentage budget change across the 36 states in Nigeria. Cross Rivers state has the highest percentage budget change across all 36 states.

Figure 13



5. CONCLUSIONS/SUMMARY

1. Lagos state has the highest confirmed cases followed by the FCT and Kaduna, Plateau, Oyo, Rivers, Edo, Ogun, Kano, and Delta State.
2. Lagos state has the highest death cases followed by Edo, FCT, Rivers, Kano, Delta, Kaduna, Oyo, Ondo, Borno.
3. Lagos state has the maximum discharged cases, followed by FCT, Plateau, Kaduna, Oyo, Rivers, Edo, Ogun, Kano, Delta.
4. Daily maximum infection was on 2023-03-08 with 93 cases.
5. There is no direct correlation between the vulnerability index and confirmed cases. Factors like age index, and population density affected the rate of infection more compared to the vulnerability index.
6. There were waves of infection observed between February 2020 to March 2022. It is evident in the recovery rate marching up with the rate of infection and low death rate that the pandemic was properly managed and curtailed.
7. The economy was visibly affected as seen in the GDP, especially in the 2nd quarter of 2020.

FURTHER WORK TO BE DONE

More reliable data sources could be researched, and I personally will want to employ the knowledge of Machine learning and artificial intelligence to further make accurate guesses and proffer predictions on the data and the probability of using such predictions to combat similar futuristic occurrences.