



USTACKY MICRO DEGREE COURSE

DATA SCIENTIST MICRODEGREE CAPSTONE PROJECT

Executive Summary - Nigeria-COVID-19-Data-Analysis-Using-Python

NAME: ADEBANJO ESTHER CHIZOBA

(Student)

13th, July, 2023

Contents

Contents	2
Introduction	3
Description	3
Data collection	3
Data cleaning and preparation	3
Data analysis	4
Insight	4

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 lockdowns, curfews, and reviving its economy.

Description

In this project, data science & analytics skills (web scraping) were employed to collect data, explore the data, perform analysis, create visualizations, and generate insights.

The data source is divided into different parts, and some data were combined to create another dataset used to perform analysis and provide insights.

Data collection

After importing all required libraries, datasets were collected from The Johns Hopkins University Center for Systems Science and Engineering (JHU CSSE) repository using the `Urllib.request` since the NCDC COVID-19 official website could not be accessed. These datasets include:

- Global daily confirmed cases
- Global daily recovered cases
- Global daily death cases
- Covid_external data
- Budget
- Real GDP

Data cleaning and preparation

- Data were cleaned by dropping columns with a null value and zero value, removing commas from numerical values
- Data were also converted to appropriate datatypes
- Nigeria data was extracted from the world COVID-19 record for analysis
- The dataframe was transposed to give it a better presentation
- Pandas DataFrame for Daily Confirmed, Recovered and deaths Cases were gotten

Data analysis

The following analysis was carried out:

- A bar plot that shows the Top 10 states in terms of Confirmed Covid (Laboratory test) cases, discharged, and death cases were generated.
- A line plot for the total daily confirmed, recovered, and death cases in Nigeria were each generated.
- The daily infection rate was determined to show the derivate of the total cases on a bar plot
- The daily confirmed, recovered, and death cases datasets were combined and viewed on a bar chart.
- The daily infection rate was determined and plotted on a line graph.
- The maximum infection rate for a day was calculated signifying the number of new cases and the date found was identified.
- The relationship between the external dataset and the NCDC COVID-19 dataset was determined based on the top 10 confirmed cases versus the overall community vulnerability index as well as the least 10 cases by merging the two datasets on a common column (states) and plotting their result on the same axis of a line plot.
- A regression plot between two variables (Confirmed Cases and Population Density) was generated to visualize their linear relationships.
- A pie chart was generated to visualize the proportion/percentage of confirmed cases per region within the country.
- The percentage difference between the pre-covid_19 and the covid_19 Q2 was determined.

- The state with the highest and lowest initial and revised budget was determined, also the percentage difference between the initial budget and the revised budget was determined. All these were added on to the original dataset and plotted on a bar plot.

Insight

- The global covid_19 dataset consists of 289 rows and 1144 valid columns while Nigeria's data consist of 1 row and 506 valid columns.
- Lagos has the most confirmed cases in Nigeria with 26708 cases while Delta state has the least cases with 1843 cases as shown in Figure 1.
- Lagos has the most discharged cases in Nigeria with 24037 cases while Delta state has the least cases with 1737 cases as shown in Figure 2.
- Lagos has the most death cases in Nigeria (236 cases) while Delta state has the least cases (36 cases) as shown in Figure 3.
- It was observed that there were very few death cases in Nigeria.
- Maximum daily infection was observed to be 6158.0 and the date it occurred was observed to be 2021-12-22 as shown in Figure 8.
- The top 10 and least 10 confirmed cases versus the overall community vulnerability index merge show that the Overall CCVI Index in these states is very low, which means that the impact of the virus on these communities after the virus arrives is low as shown in Figures 9 and 10 respectively.
- Analysis of the real GDP data shows that the GDP dropped from 413801074.72 Pre-Covid_19 to 15890000.0 during Q2 2020, a %change of 96%. A visualization was also presented as shown in Figure 13.
- Lagos state has the highest initial and revised budget, Yobe state has the lowest initial budget and Nasarawa state has the lowest revised budget as shown in Figure 14.

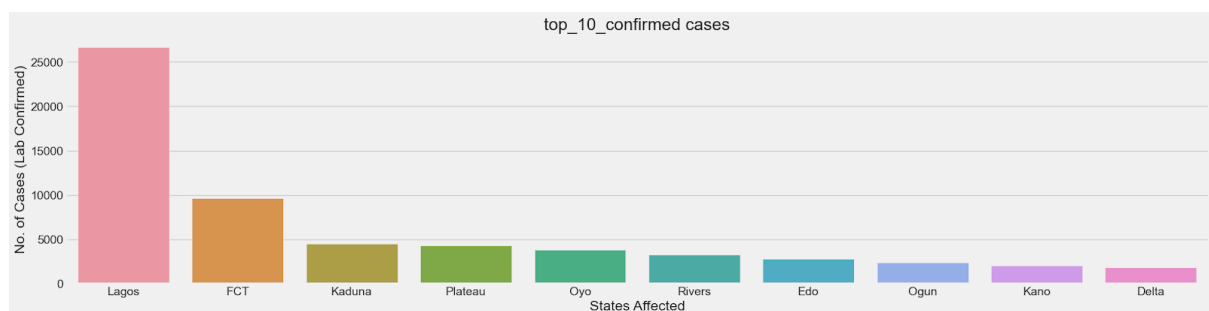


Figure 1: Top 10 Confirmed Covid cases

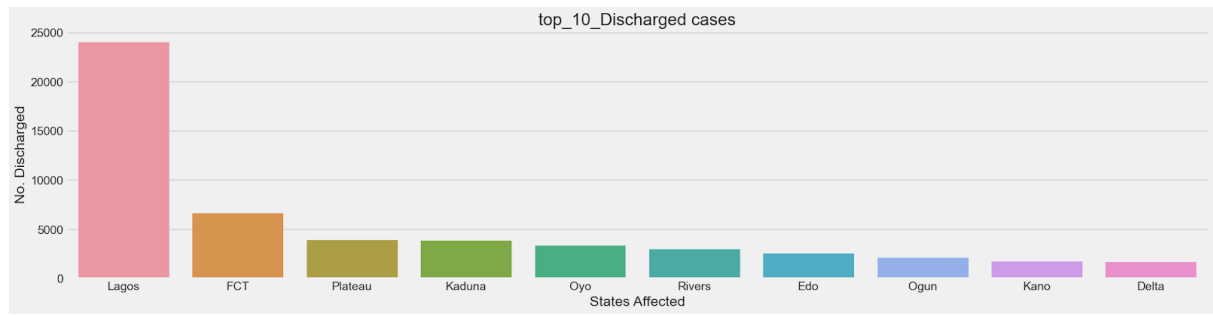


Figure 2: Top 10 Discharged cases

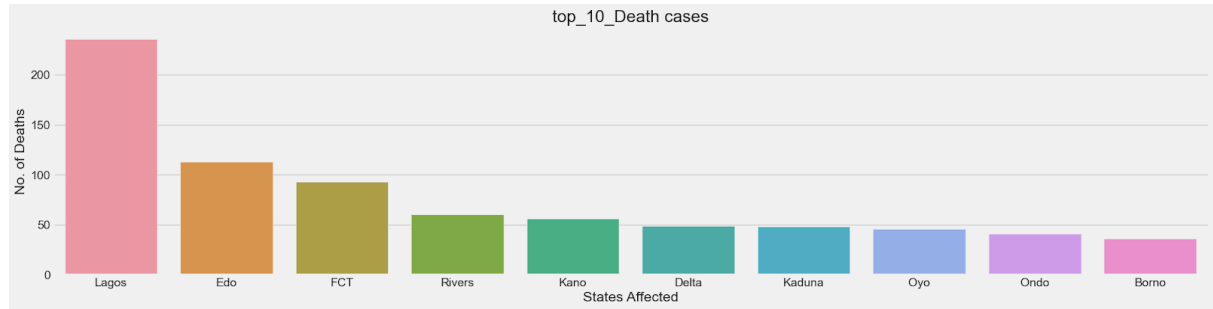


Figure 3: The top 10 Death cases

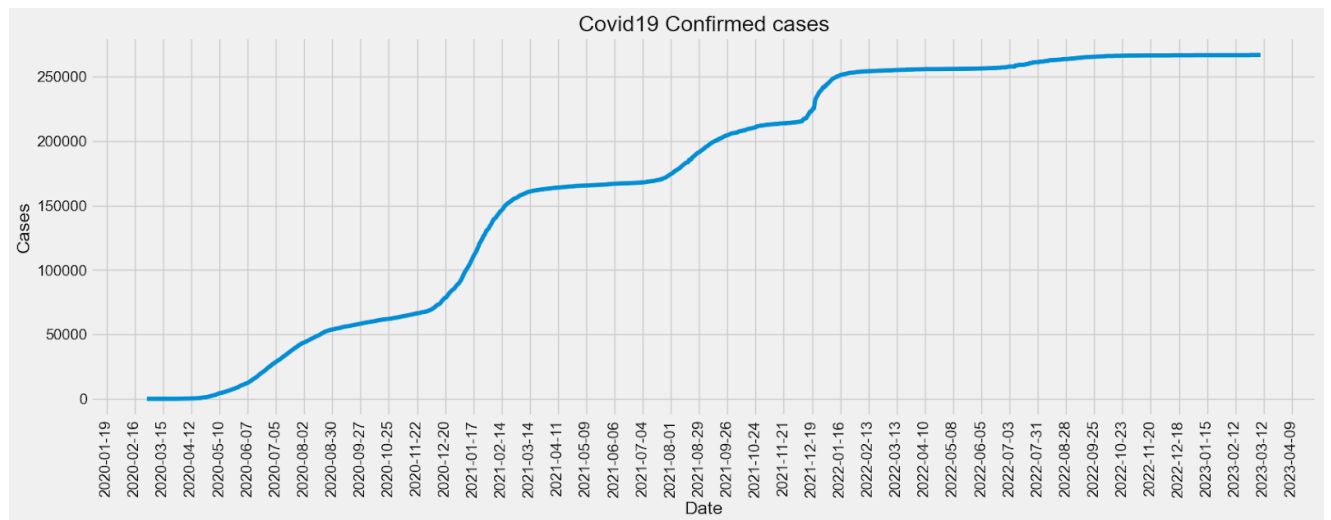


Figure 4: Total daily confirmed cases

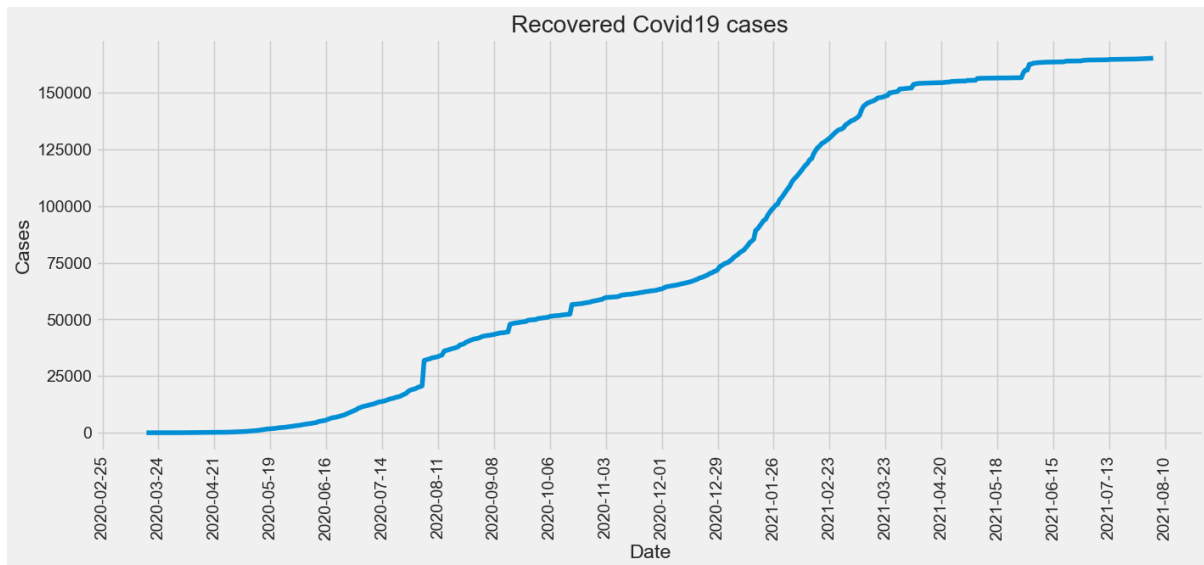


Figure 5: Total daily recovered cases

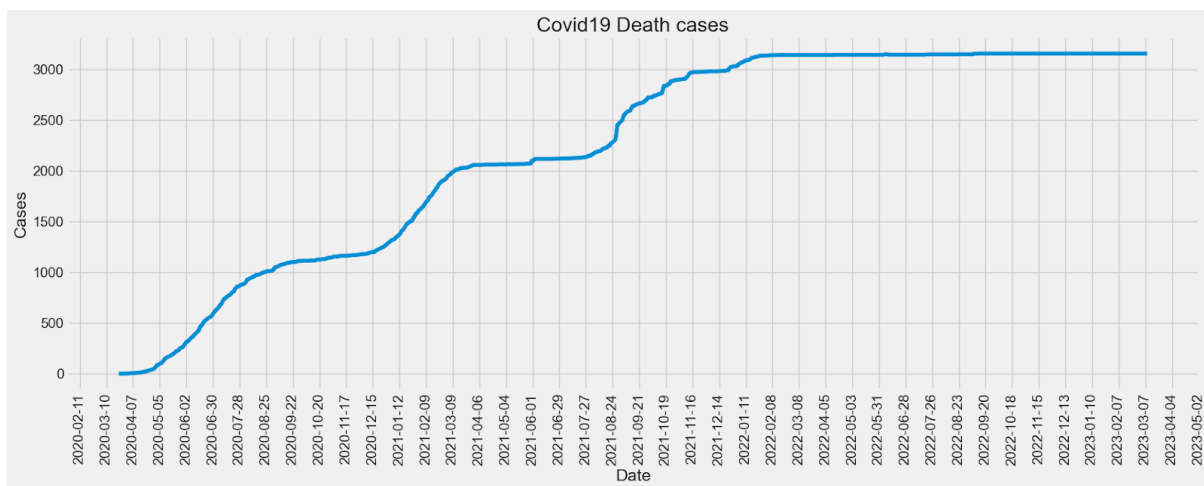


Figure 6: Total daily death cases



Figure 7: Combine datasets for daily confirmed, recovered, and death cases

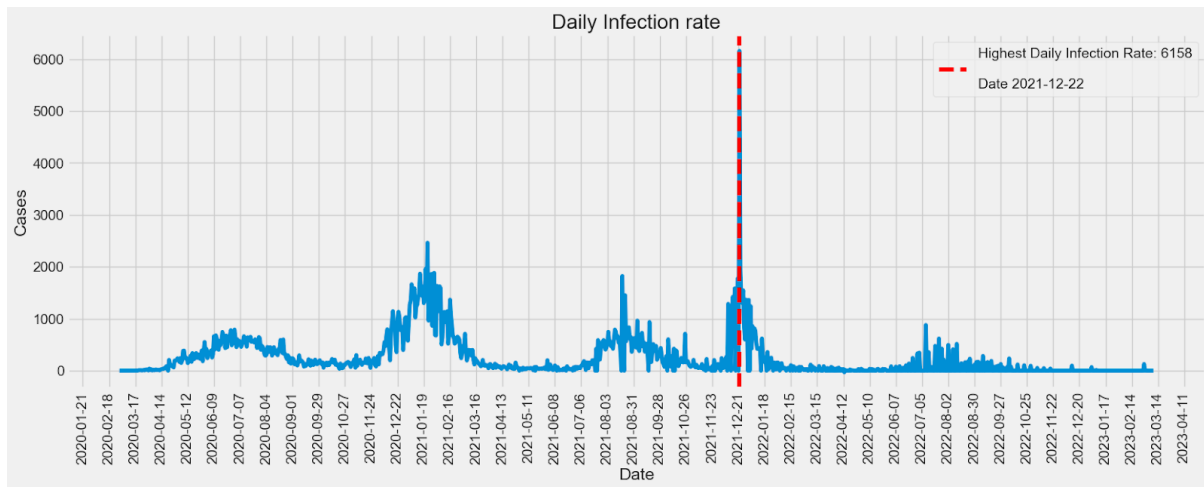


Figure 8: Daily infection rate

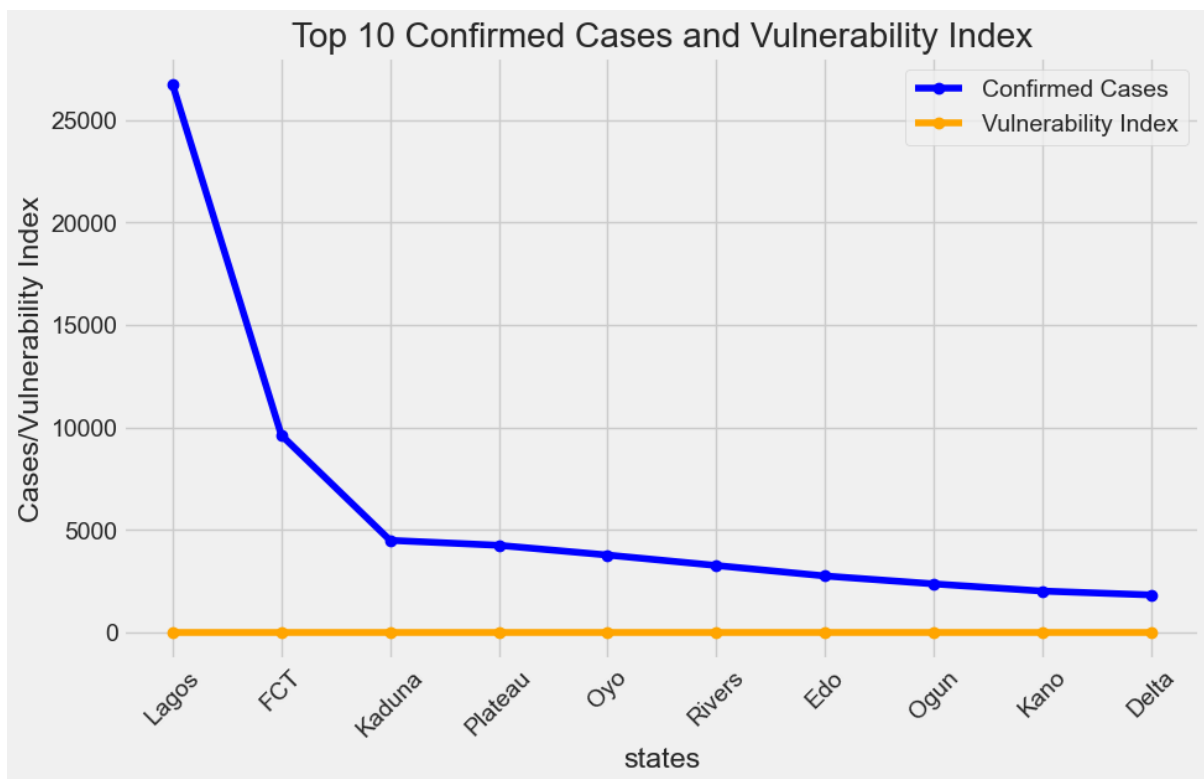


Figure 9: Top 10 confirmed cases versus the overall community vulnerability index

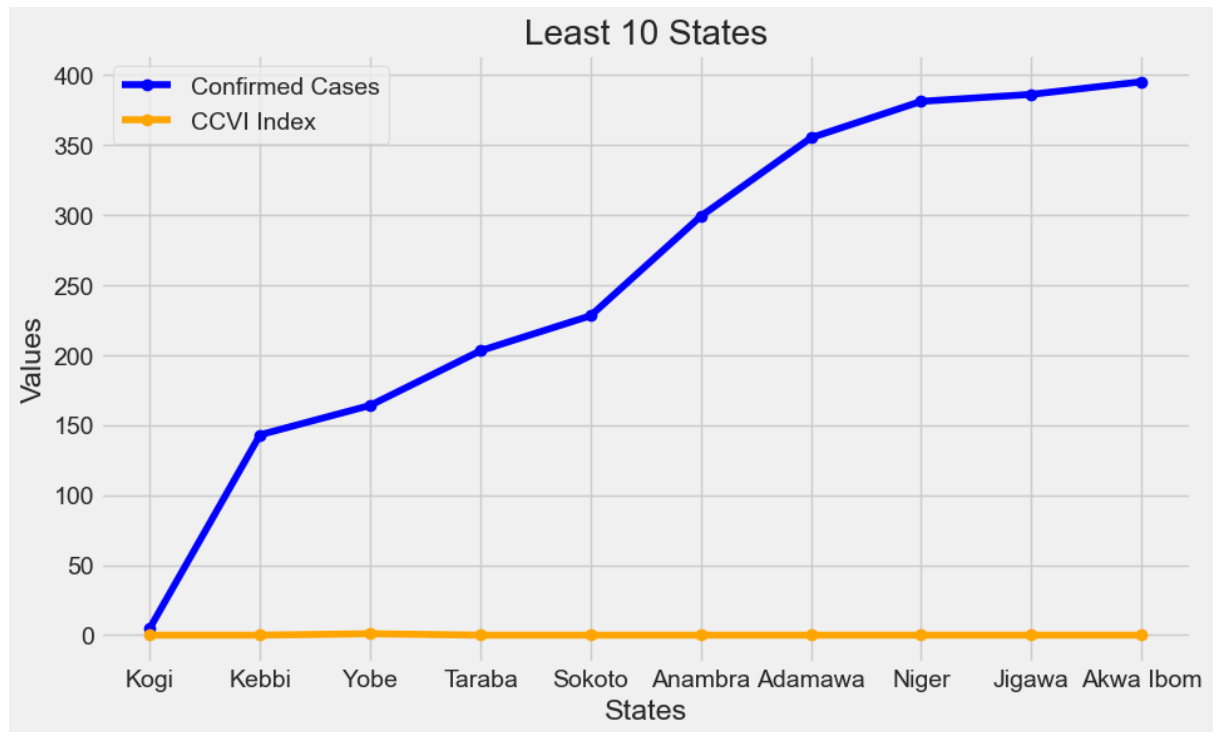


Figure 10: Least 10 confirmed cases versus the overall community vulnerability index

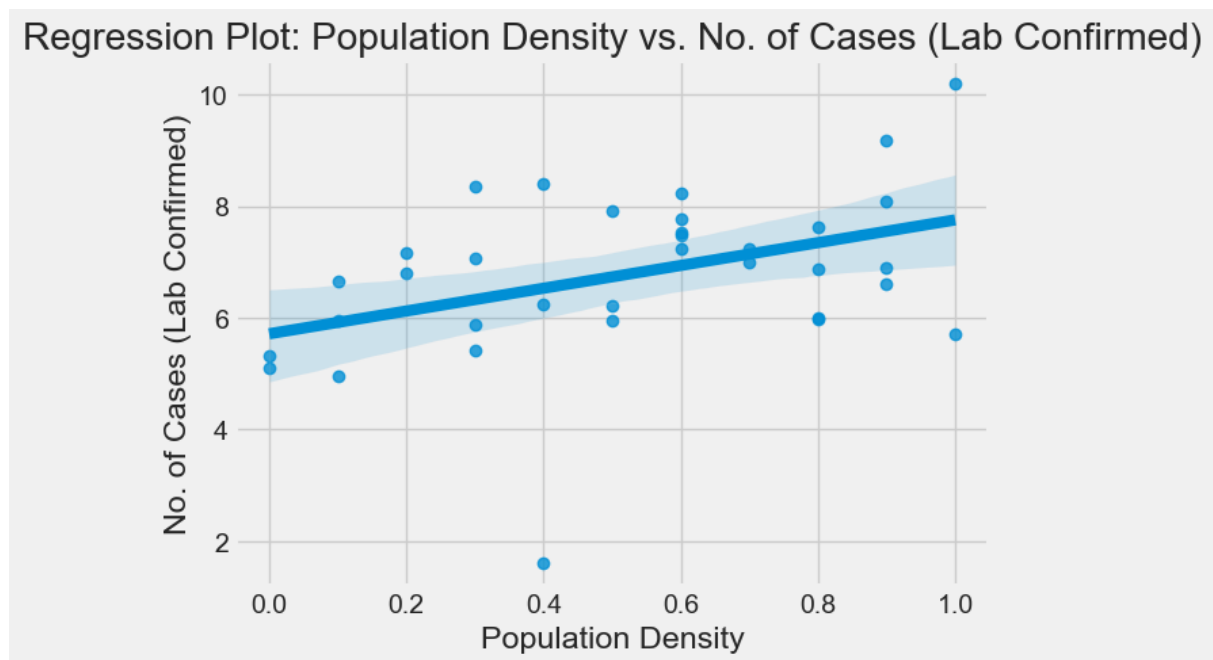


Figure 11: Regression plot of confirmed Cases and Population Density

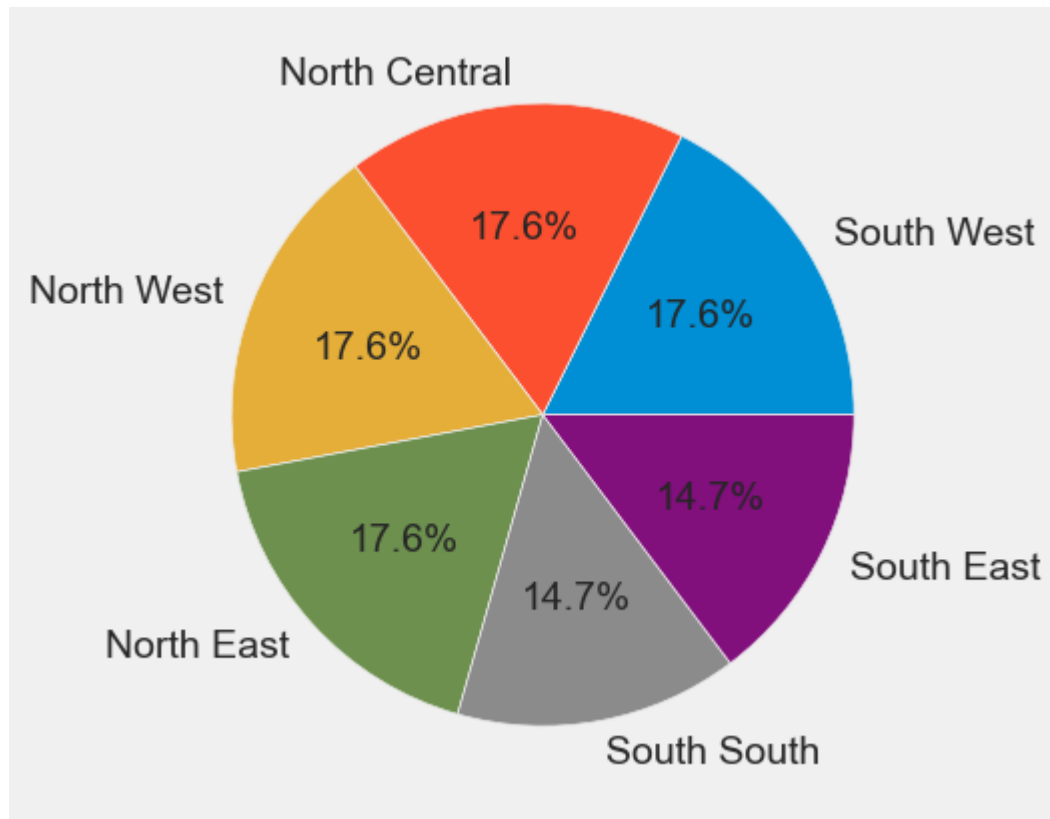


Figure 12: A pie chart visualizing the proportion/percentage of confirmed cases per region.

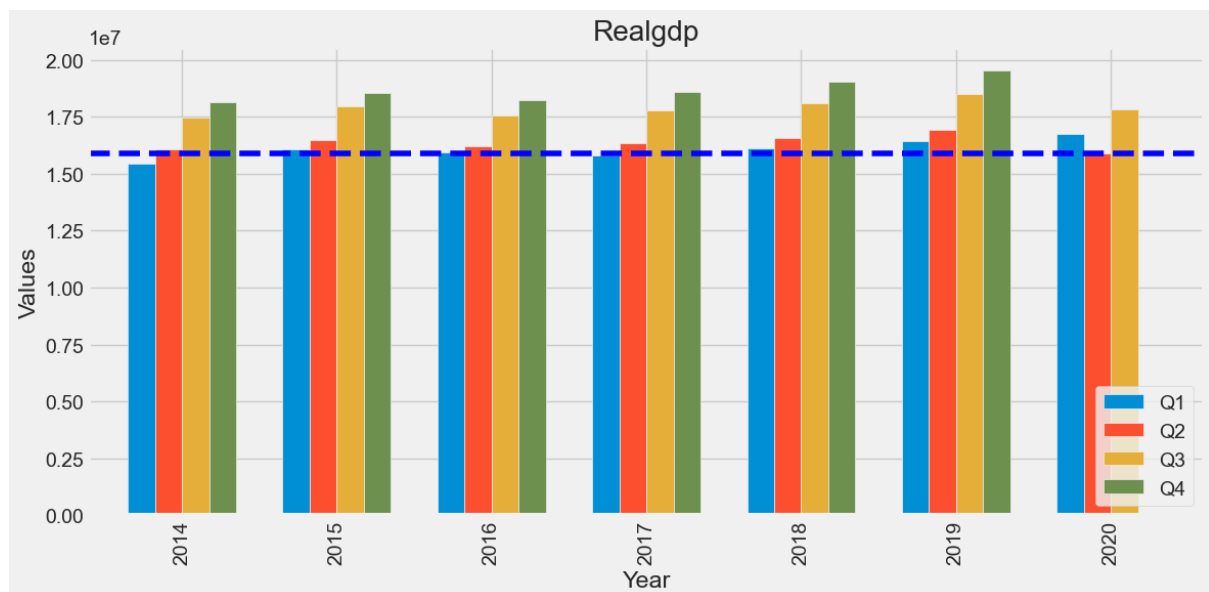


Figure 13: Real GDP analysis

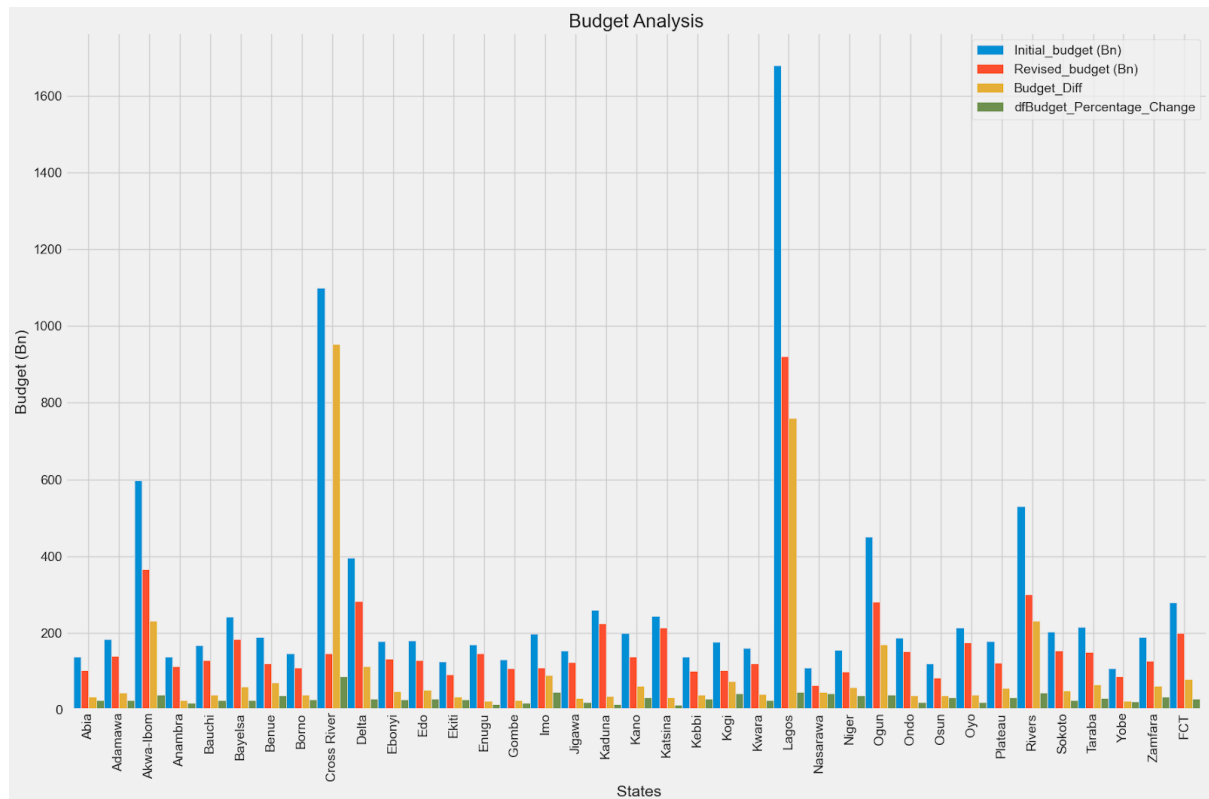


Figure 14: Budget analysis

Observations

- It is observed that the top 10 states with confirmed Covid_19 cases are: Lagos, FCT, Kaduna, Plateau, Oyo, Rivers, Edo, Ogun, Kano, and Delta states in descending order as represented in Figure 1.
- It is observed that the top 10 states with discharged Covid_19 cases are: Lagos, FCT, Plateau, Kaduna, Oyo, Rivers, Edo, Ogun, Kano, and Delta states in descending order as represented in Figure 2.
- It is observed that the top 10 states with Covid_19 death cases are: Lagos, Edo, FCT, Rivers, Kano, Delta, Kaduna, Oyo, Ondo, and Borno states in descending order as represented in Figure 3.
- Figure 4 shows that the first case was confirmed in February 2020, however, confirmed cases came to its peak between September 2022 and October 2022 and since became steady till around February 2023
- Figure 5 shows that recovery started around February 2020, however, peak recovery occurred between July 2021 and August 2021.

- Figure 6 shows that the first death case occurred in March 2020, however, death cases came to its peak between January 2022 and February 2022 and since became steady till around March 2023.
- Figure 7 shows the variation between the Confirmed, Recovered, and Death cases
- Figure 8 represents the daily infection rate and shows that the highest daily infection is 6158 which occurred on the 22nd of December 2022.
- Figure 9 shows the top 10 confirmed cases with respect to their vulnerability index which happens to (very low) zero.
- Figure 10 shows the least 10 confirmed cases with respect to their vulnerability index which happens to (very low) zero. With Kogi state being the least Covid_19 confirmed case.
- Figure 11 shows a linear regression of the confirmed cases distribution versus population density.
- Figure 12 shows a pie representation of the various regions within the country.
- Figure 13 shows a drop in the GDP during the second quarter of the year 2020 that is almost equivalent to the GDP recorded in 2014 Q2.
- Figure 14 shows that Lagos state has the highest initial and revised budget, Yobe state has the lowest initial budget and Nasarawa state has the lowest revised budget, I further show the difference and percentage change in the initial and revised budget.