

EXECUTIVE SUMMARY

ANALYSIS OF NIGERIA’S COVID-19 DATA



June 25, 2021

AUTHOR: OLUWASEYI SALAMI

# INTRODUCTION

Coronavirus disease 2019 [(COVID-19)](https://en.wikipedia.org/wiki/COVID-19), caused by the SARS-CoV-2 virus, was first identified in Wuhan, China in 2019 and subsequently spread all over the world leading to lockdowns and consequently economic disruption and uncertainty.

This is an exploration and analysis of publicly available COVID-19 data pertaining to Nigeria as part of Ustacky's Data Science Capstone Project.

# ANALYSIS STEPS

## Data collection

1. Scraping relevant data from the Nigeria Centre for Disease Control website and saving it to a csv file.
2. Importing daily global COVID-19 data from the John Hopkins University GitHub repository.
3. Saving other external data to Pandas dataframes.

## Viewing basic information about the collated data

* This was done by calling `head` and `info` methods on the dataframes to determine the need for data cleaning and preparation.

## Data cleaning and preparation

1. The scraped data was cleaned.
2. Pandas dataframes for Nigeria-specific cases were sliced and transposed.
3. Conversion to the appropriate data type.

## Analysis

* Specific info, plots and visualizations were generated highlighting various important aspects to the collated data.

# INSIGHTS

* A small proportion of blood samples tested returned positive for COVID-19. Most of the confirmed cases were discharged, with a tiny percentage of infections and cases resulting in death.
* Lagos state has the highest number of deaths, confirmed and recovered cases.
* A two-phased logarithmic growth pattern was seen on line plots of confirmed cases, recovered cases and deaths from COVID-19. The first increase in cases and deaths occurred around 5-6 months after the onset of the pandemic while the second increase occurred in early 2021.
* The most significant spike in the number of recoveries is seen around the 8th month of 2020 with the maximum no of daily recoveries, 11,188, seen on the 4th day of that month.
* Lagos has the highest number of cases and correspondingly the lowest overall CCVI (COVID-19 community vulnerability index).
* States with a low overall CCVI tend to have a higher number of COVID-19 deaths.
* Population density is a poor predictor of the number of cases and death.
* There is a slight decrease in Nigeria's second quarter 2020 GDP value compared to the pre-COVID 2019 value.
* All the years show an increase in GDP from Q1 to Q4 except 2020 in which there's a decrease in GDP from Q1 to Q2. This corresponds to the period of the first COVID-19 wave and national lockdown.
* The state with the largest relative decrease in budget is Cross River followed by Lagos state. However, Cross River is not among the top 10 states most affected by COVID-19.
* The highest number of monthly confirmed and recovered cases was seen in January followed by February.
* The highest number of monthly deaths is seen in February, 2021 following the spike in number of confirmed and discharged cases that occurred in January.
* However, the month within which the second highest number of deaths was recorded is June, 2020.
* A similar profile of the number of deaths per month during both waves may be an indication of a reduced mortality rate during the second wave.

# FUTURE PROPOSED ANALYSIS

* Data visualization using interactive plots created by other visualization libraries e.g., hvplot, holoviews and bokeh.
* Further analysis using weekly COVID-19 data.
* Comparison with other countries' data.

# VISUALISATION EXAMPLES

















