EXECUTIVE SUMMARY

Coronavirus (COVID-19) is an infectious disease caused by a newly discovered virus, 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, I have employed data science and analytics skills to collect data, explore the data, perform analysis, create visualizations, and generate insights.

The data used for this project include:

* Data from NCDC Covid-19 official website. I was able to get this data by scraping it from NCDC’s website using Python’s *Beautiful Soup*. Features in this data include: Number of cases confirmed in the laboratory and on admission, Number of discharged patients and Number of patients who died.
* Data from The Johns Hopkins University Center for Systems Science and Engineering repository, they publish daily data on confirmed death and recovered cases across different countries in their repository.
* The Nigeria Community Vulnerability Index data, which 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.
* Real Domestic Gross Product Data of Nigeria, and the State Budget Data.

**After performing exploratory analysis on the data, I was able to derive the following insights:**

* The top ten states in terms of confirmed covid-19 cases by Laboratory test are: Lagos, FCT, Kaduna, Plateau, Rivers, Oyo, Edo, Ogun, Kano and Ondo.
* The top ten states with the highest number of discharged covid-19 patients are: Lagos, FCT, Kaduna, Plateau, Rivers, Oyo, Edo, Ogun, Kano and Ondo.
* The top ten states with the highest number of deaths are: Lagos, Edo, FCT, Oyo, Kano, Rivers, Delta, Kaduna, Ondo and Plateau.
* After merging NCDC data with Nigeria Community Vulnerability Index data, I was able to detect a negative relationship between number of cases confirmed in Lab and overall CCVI Index that is the higher the number of confirmed cases, the lower the overall community vulnerability index in the states except in FCT.
* There is a direct relationship between population density and number of cases confirmed in the lab, the higher the population density the higher the number of cases confirmed in the laboratory.
* There is also a negative correlation between Age Index Score and number of cases confirmed in Lab with Age Index Score and Number of Deaths; that is, when we have a low Age index score, there are more confirmed cases and more number of deaths and vice versa.

Written By:

Ojini Winifred for UStacky Capstone Project