PROJECT OVERVIEW.

INTRODUCTION

The project focus is on the effect of the SARS COV – 19 virus on Nigeria, there are varying data used containing GDP , Budgets , overall community vulnerability index , NCDC report etc. The report intends to shed light on findings, problem encountered, and report from the analysis.

From the Data, the most affected state in terms of infection and death is Lagos, due to its population and small land mass. The virus being a communicable disease must have thrived in a state like Lagos which is overpopulated, followed closely is FCT, which is the federal capital territory. States with expatriates, foreigners and people with means of leaving the country regularly is also a very important factor in deducing the cause of high infection rate in certain states. Another factor to take close notice is the availability of testing centers and disease control units, States without proper health systems show high number of cases leading to death due to poor management of victims/patients

One major problem this dataset must have faced is people awareness, a factor that greatly affects the credibility of the data and its insights, as occupants of remote and rural areas might get infected and be asymptomatic and even if they are symptomatic, might not be able to access testing centers and will easily get misdiagnosed. These number of people will easily go under the radar and affect the statistics of the data.

This I believe is the reason for the least number of confirmed cases being situated in the North. Another problem will be mishandling of data, but without evidence pointing towards any such cases its effects are diminutive and can be ignored.

As stated the state with largest confirmed cases is Lagos, while the state with the least number of confirmed cases is Kogi.

DATA

NCDC data, budget data, GDP data, infect rate data, recovery rate data, death rate data, overall community vulnerability index data, these are the data used, and they contained integers, floats, objects and datetime columns

Methods

WEB SCRAPING

First input the link into a variable, using the request.get() function to bring the link to your working environment. To scrape the data from its source employ Beautiful soup, with little knowledge of HTML and the help of the inspect option on the web page, identify the table we intend to extract. Then extract the header and then the body and input into a data frame.

DATA ANALYSIS

An inquisitive approach comes in handy when dealing with data, in this case guidelines have already been provided which makes the process easier. There is still a process of inspecting the data, cleaning the data and preprocessing which consist of Feature engineering as to put the data in a required form for analysis. Lastly information extraction or data analysis, this is to get insights from the data, either by following a guide or asking the right questions

RESULTS

Lagos has the highest number of confirmed cases this could be due to the state being overpopulated, Lagos also has the highest number of discharged patients and recorded deaths. The state with least recorded infection rate is Kogi.

Albeit Lagos having the lowest Community Vulnerability Index it had the highest confirmed cases and in the top 10, Kaduna and Kano have the higher Community Vulnerability index but their number of confirmed cases is very low compared to other states with lower Community Vulnerability index

There is a correlation between the level of each regions health system and how many patients could be taken care of and discharged as the regions with better health care systems discharged more patients.

The Pandemic greatly affected budgets and GDP, causing Budget review and drop in GDP