# 2. DATA ACQUISITION AND CLEANING

## 2.1 DATA COLLECTION

The datasets to be used for the analysis are collected across different platforms which is depended on the data requirement. The data on the status of the nouvelle coronavirus cases in Nigeria are collected from the National Centre for Disease Control through [this link](https://covid19.ncdc.gov.ng/) which highlights the confirmed cases, active cases and death toll across the nation. The population forecast of Nigeria for 2019 is obtained from the National Bureau of Statistics through [this link](https://nigeria.opendataforafrica.org/ifpbxbd/state-population-2006). Other datasets such as the Gross Domestic Products (GDP), the literacy level and the geographical locations are obtained from [here](https://en.wikipedia.org/wiki/List_of_Nigerian_states_by_GDP), [here](https://en.wikipedia.org/wiki/List_of_Nigerian_states_by_literacy_rate) and [here](https://www.researchgate.net/figure/The-36-states-of-Nigeria-and-their-coordinates_tbl1_273009753) respectively.

These data are not collected for exactly the same period of time but they are all obtained within the last five years, thence, it could be assumed that there have not been much changes in the demographics of the country.

The data on the location of hospitals at the central cities of each state would be collected using the Foursquare Application Programming Interface (API). The collection would be based on a radius of 5000m from the central location and a limit of 500 responses would be imposed.