**Goal:**

The goal of this project is to try to mimic the COVID-19 Dashboard showed on the following website: https://coronavirus.jhu.edu/map.html

**Aim:**

The project aim is to demonstrate the ability to perform data pre-processing using a combination of python and Excel. And able to use Power BI to create high quality of data visuals. Most importantly try to gain insights from data visualization.

**Data source**: <https://github.com/CSSEGISandData/COVID-19>

Mainly I have used (not ordered):

1. <https://github.com/CSSEGISandData/COVID-19/tree/master/archived_data/archived_daily_case_updates>
2. <https://github.com/CSSEGISandData/COVID-19/tree/master/csse_covid_19_data/csse_covid_19_daily_reports>
3. <https://github.com/CSSEGISandData/COVID-19/blob/master/csse_covid_19_data/UID_ISO_FIPS_LookUp_Table.csv>

**Things Considered**:

1. Use Python

* Concatenated over 200 CSV files into one data frame.
* Detected, and updated inconsistent column names and row names.
* Removed and duplicated rows and unnecessary columns.
* Changed date format and filled empty cells with zeros or blank spaces.
* More.

1. Use Excel

* Used VLOOKUP() function to detect and import latitude and longitude information from UID\_ISO\_FIPS\_LookUp\_Table.
* Created new columns/attributes: Daily\_new\_cases, Daily\_deaths, Daily\_avtives, Incidence\_Rate and Case-Fatality rate from existing attributes.
* Included Smoothing.

1. Use Power BI

* Included heat map, line chart, area chat, cards, etc.
* Enabled animation.
* Considered order hierarchy of column names.
* Implemented joins with two different tables via foreign key.