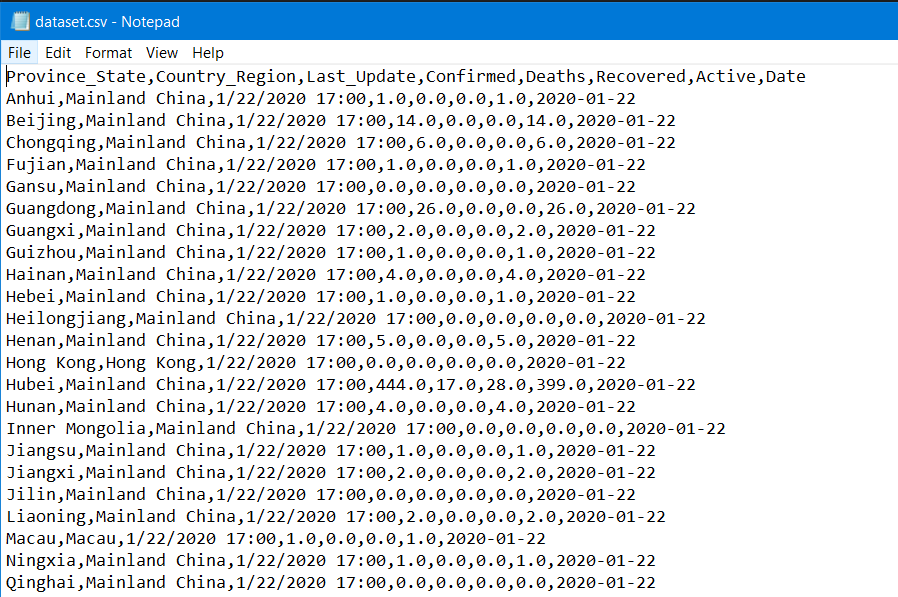
**Running:**

1. **The datasets collected from the “**<https://github.com/CSSEGISandData/COVID-19>**” were placed in the “datasets” directory**
2. **“parseDataset.py” will read all the .csv files containing day to day stats of the COVID-19 and convert the data into Dataframe which will be stored inside the “input/dataset.csv”**
3. **“preProcessing.py” will read the .csv file and preprocess the data by removing the columns and also conversion of data forms**
4. **“plotGraphs.py” consists of functions which take dataframe as argument and constructs the graphs from it. Those graphs will be stored in “output/\*GRAPHNAME\*.png”**
5. **“analysis.py” will read dataset.csv from input Directory and performs the analysis**

**Data:**

**The csv file generated by the “parseDataset.py” will be as follows**

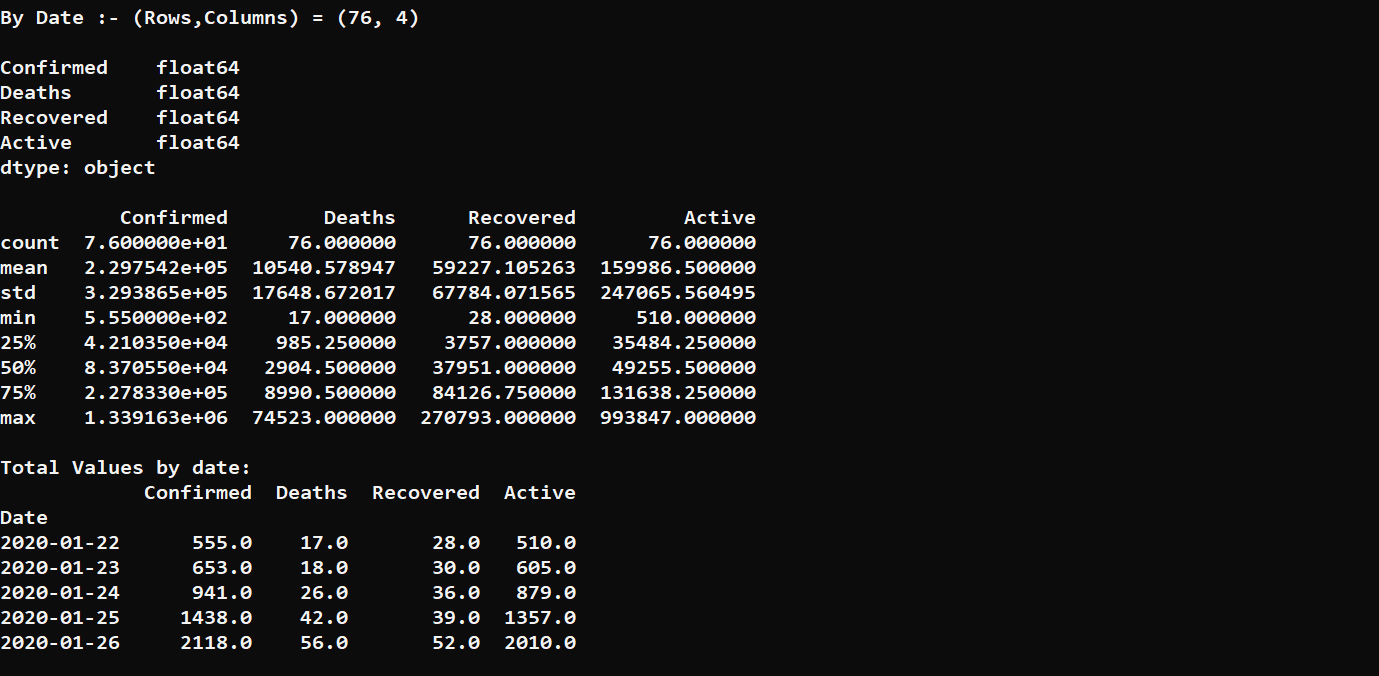


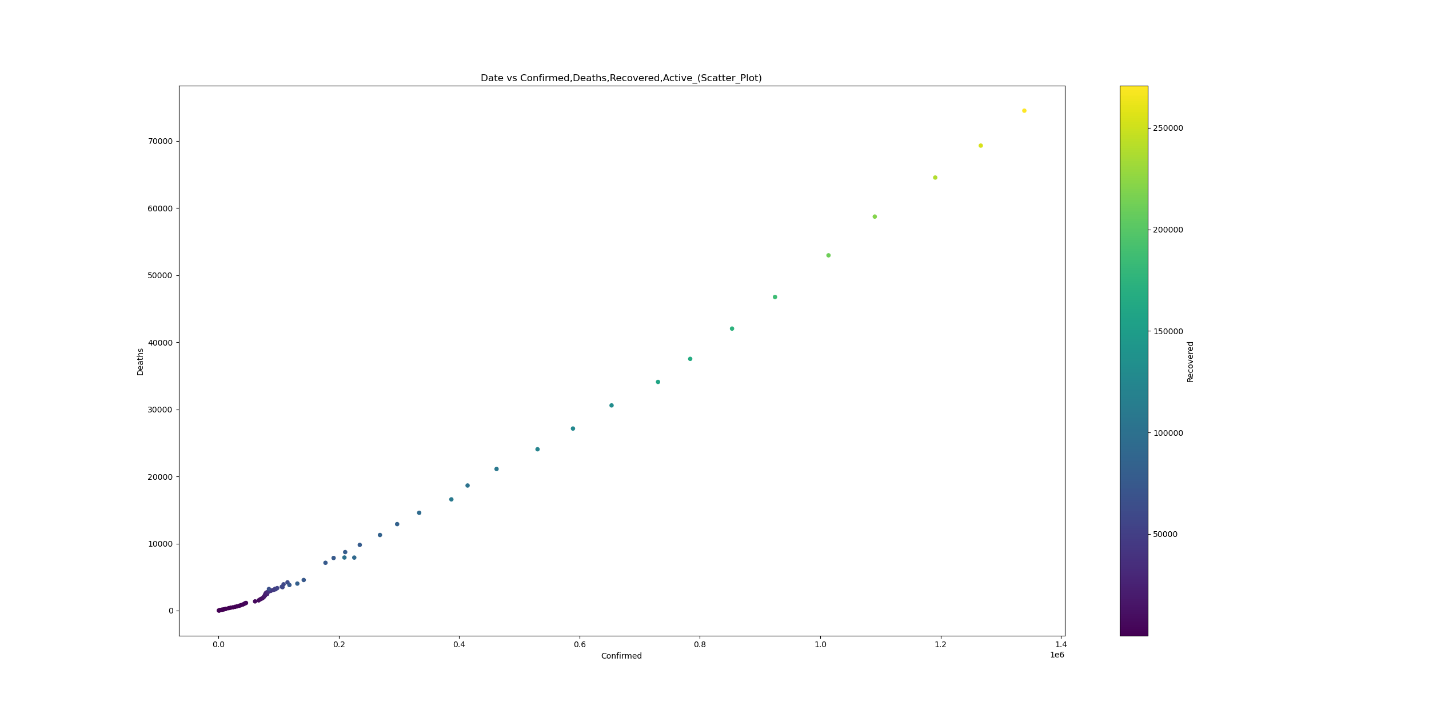
**Fields:**

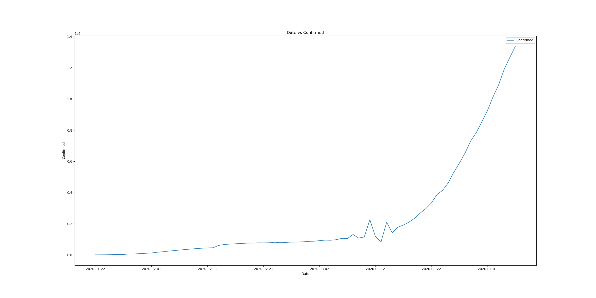
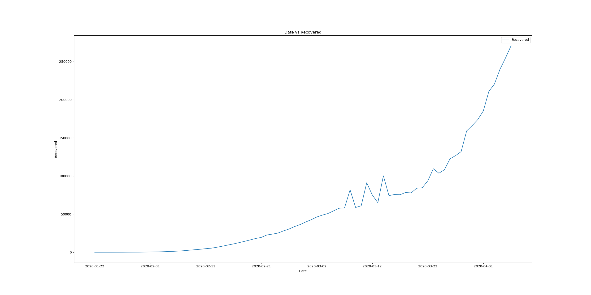
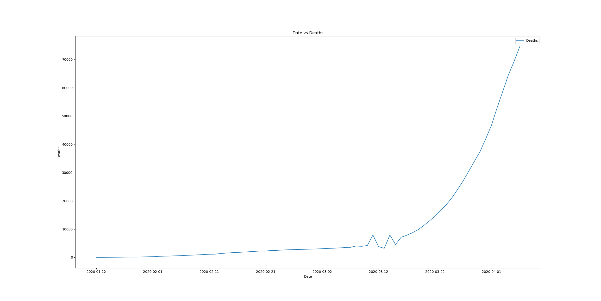
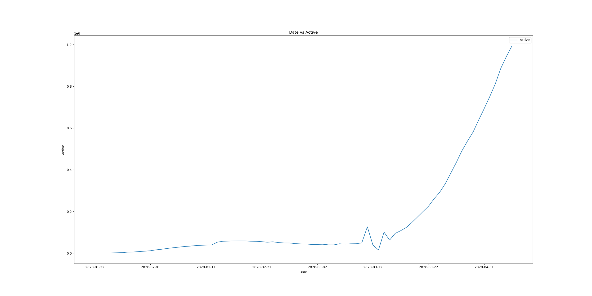
* **Province\_State** is used to indicate the “STATE” in the country and the data type is string
* **Country\_Region** is used to indicate the “COUNTRY” OR “PART OF THE COUNTRY” and the data type is string
* **Last\_Update** is a timestamp indicating the time when the stat has been recorded
* **Confirmed** is a integer value indicating number of COVID-19 confirmed victims of that time in the country or state
* **Deaths** is a integer value indicating number of COVID-19 victims who are passed away in the country or state
* **Recovered** is a integer value indicating number of COVID-19 victims who got recovered from virus in the country or state
* **Active** is a integer value indicating number of COVID-19 victims who are still infected neither dead nor recovered from virus in the country or state
* **Date** is a string indicating to which that date that particular row is related to

**Analysis:**

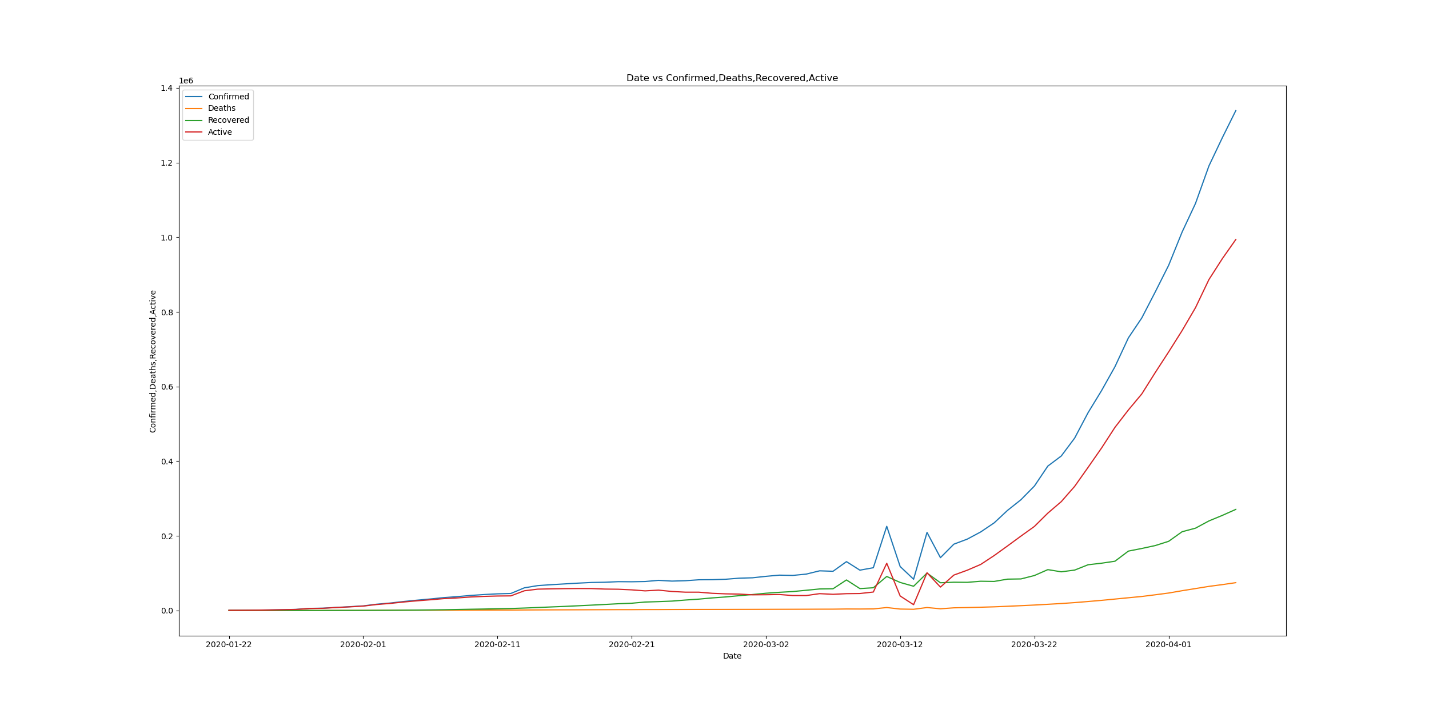
* **By Date:**



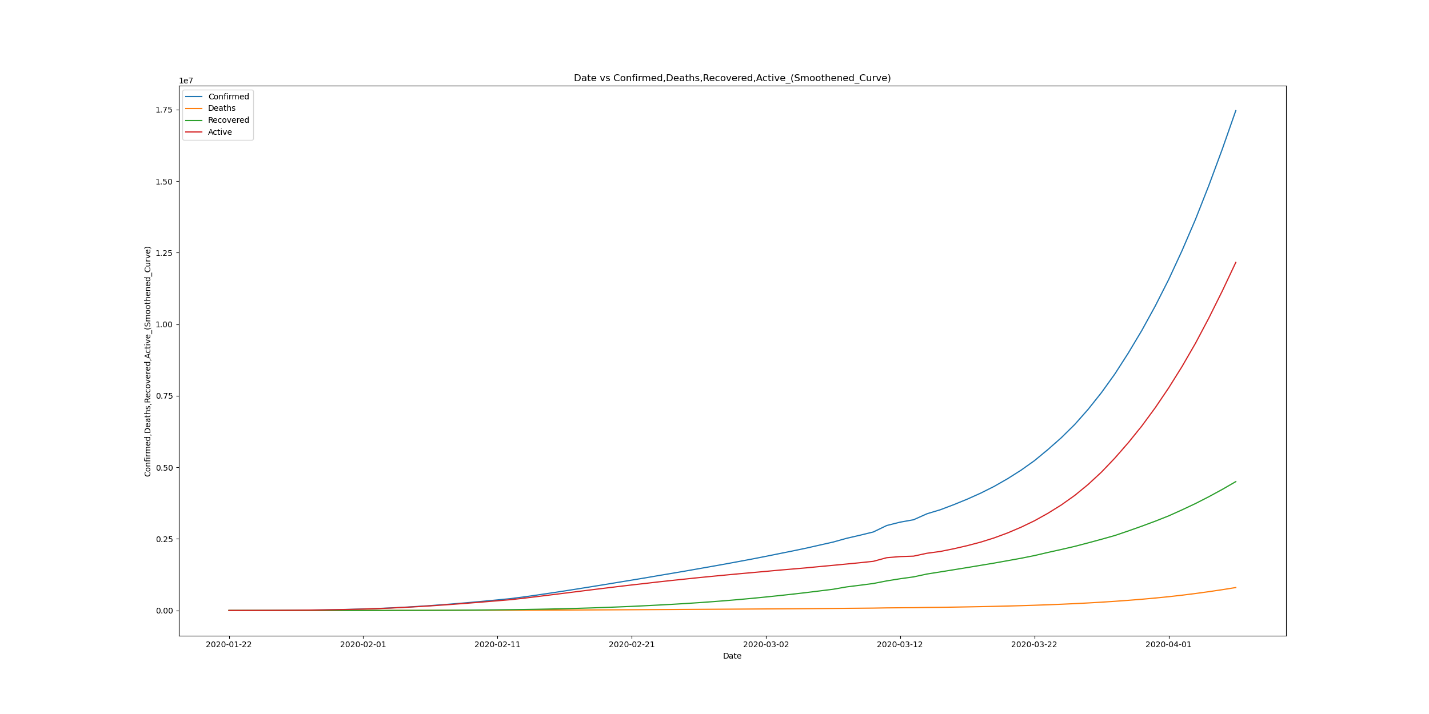
* + **Scatter Plot :** This plot shows how COVID-19 Confirmed vs Deaths varies via Recovered. 
  + **Line Plot :** These plots are plotted in multiple ways to show how the they are varied from day to day.

****  **** 

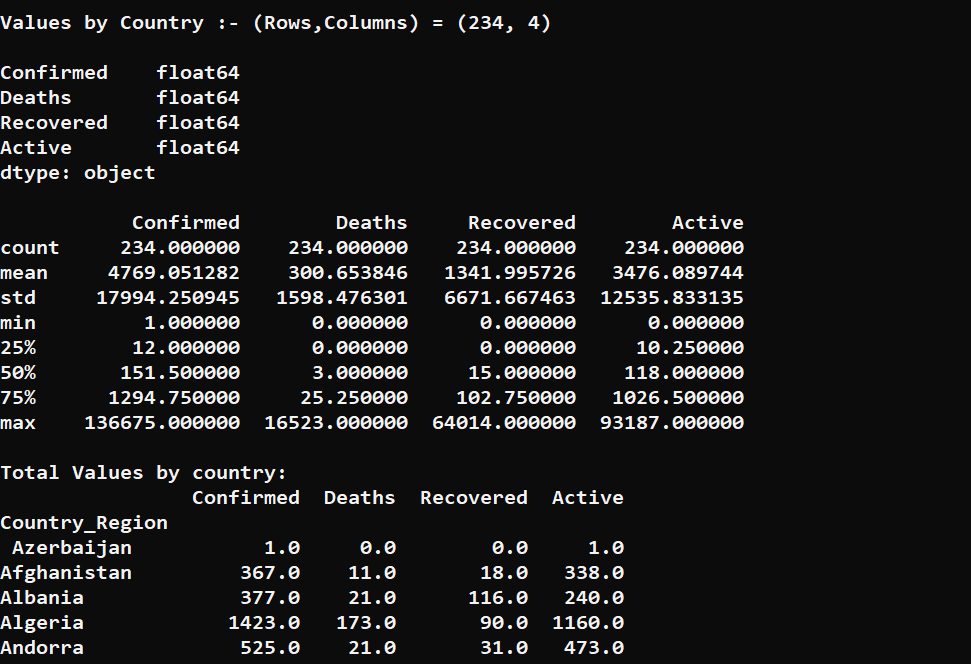
The comparison between Date vs Confirmed, Deaths, Active, Recovered is as follows

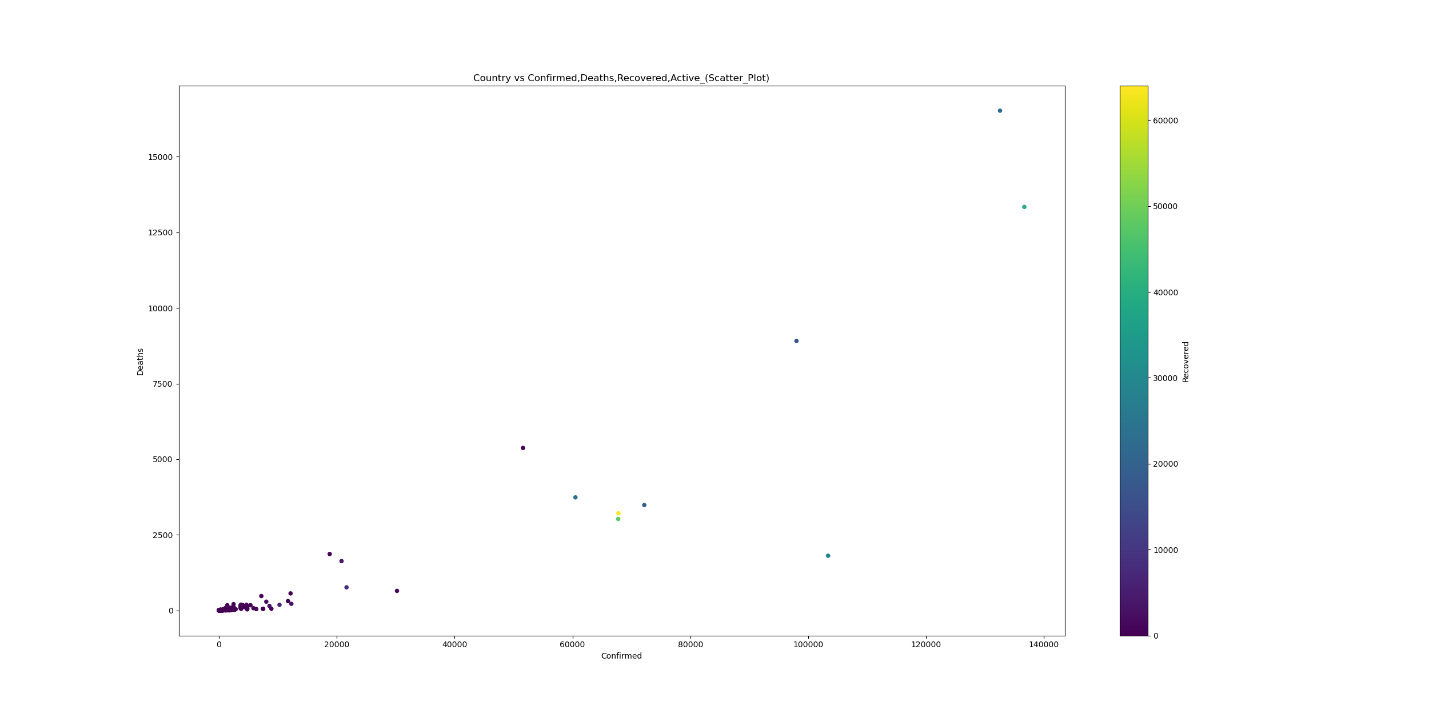


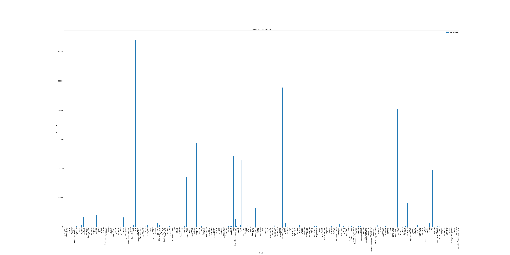
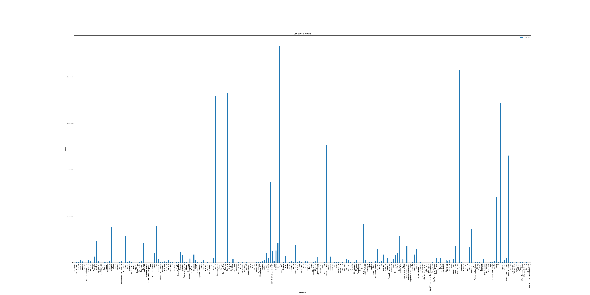
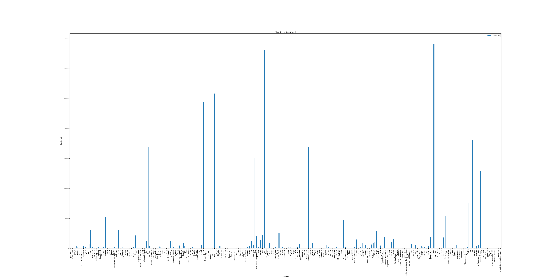
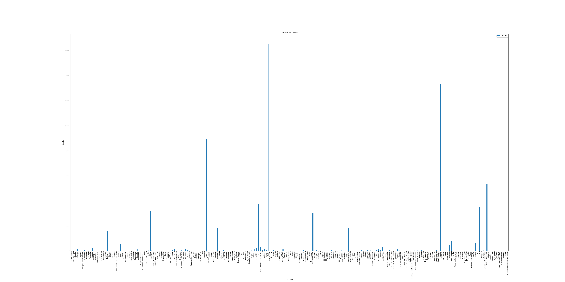
The normalize curve of the above line plot is



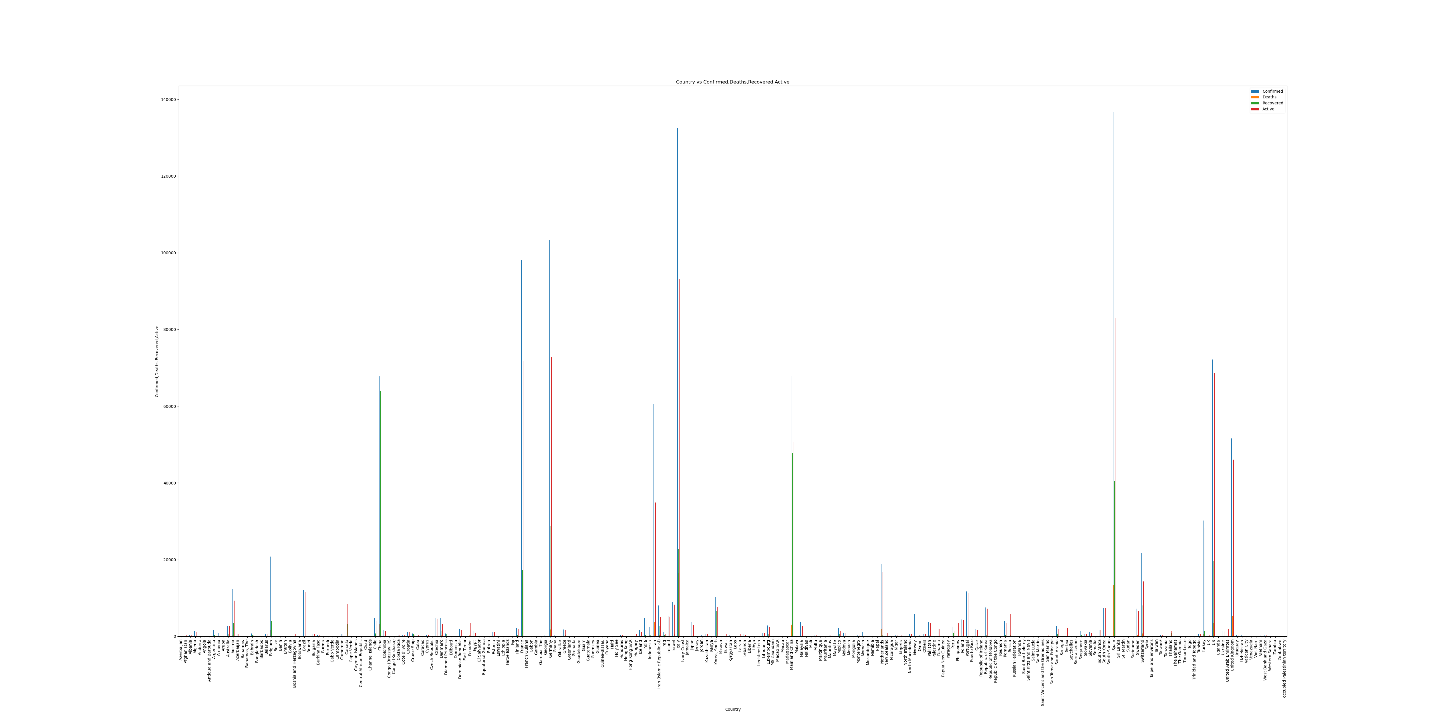
* **By Country:**



* + **Scatter Plot :** This plot shows how COVID-19 Confirmed vs Deaths varies via Recovered. 
  + **Bar Plot :** These plots are plotted in multiple ways to show how the they are varied from country to country.



The comparison between Country vs Confirmed, Deaths, Active, Recovered is as follows



By the Visualizations we can observe that the COVID-19 has been populating so fast among the countries. Italy followed by Spain has been major victims of this virus.