Dataset explanation

There are three datasets used in the project. Confirmed cases, deaths, recovered.

These datasets contain information on province,country,latitude,longitude, dates which shows the number of cases on that date

Why we used numpy arrays

LINEAR REGRESSION

1. We use Sklearn library for performing linear regression
2. First we see the confirmed cases
3. Here we convert the required confirmed cases into a numpy array as each element of the array the sum of each column..
4. We split the model into train and test using respective function in the sklearn library
5. Then we fit the best line using the training data and predict the output using test data.
6. We plot the graph with the matplotlib library.
7. Then we find the mean absolute error, mean squared error, root mean squared error.
8. We repeat the similar method for the datasets of death and recovery.
9. Show plots.