Understanding the COVID-19 Pandemic as a Data Analytics Issue

For this task, you are required to answer queries regarding the COVID-19 pandemic. You are provided with data regarding the pandemic containing information regarding the daily new cases and cumulative cases for each country.

Dataset is attached below:

Coding:

You will load the given dataset and answer the following queries.

1) On a given day, find the top 20 countries with the most confirmed, most death, and most recovered cases.

2) Find the country(s) with the highest (new cases, new deaths) between two given dates.

3) Find the starting and ending days of the longest spread period for a given country. The spread period is defined as the period where daily new cases tend to increase. They may contain days where new confirmed cases were relatively lower or none at all.

For example, [5, 2, 9, 16, 11, 27, 14, 45, 11] has a longest spread period elapsing 7 days from day 2 (2) to day 8 (45) and elements are [2,9,16,27,45] or it may be elapsing 8 days from day 1(5) to day (8) and elements are [5,9,16,27,45] both have same number of elements i.e. 5.

4) Find the longest daily death toll decrease period for a given country. They may contain days where new confirmed deaths were relatively higher.

For example, [9, 5, 1, 16, 11, 23, 8, 3, 27, 14, 45, 11] has a longest daily death toll decrease period of 4 {16, 11, 8, 3}

5) Percentage of deaths.

6) Find top 10 most affected countries.

7) Find top 10 countries with the highest numbers of deaths.