# Covid 19 Data analysis

## Load the covid 19 data into a spark data frame (country\_wise\_latest.csv) with the correct schema definition.

## 

## The are some column names which are long, contains special characters, spaces etc. Rename all such column names accordingly. Example Country/Region - > country New cases - > New\_cases etc

### Schema

A screenshot of a computer program

Description automatically generated

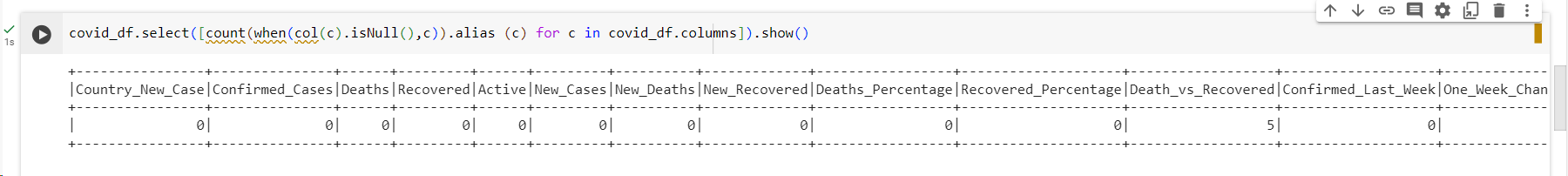
A screenshot of a computer program

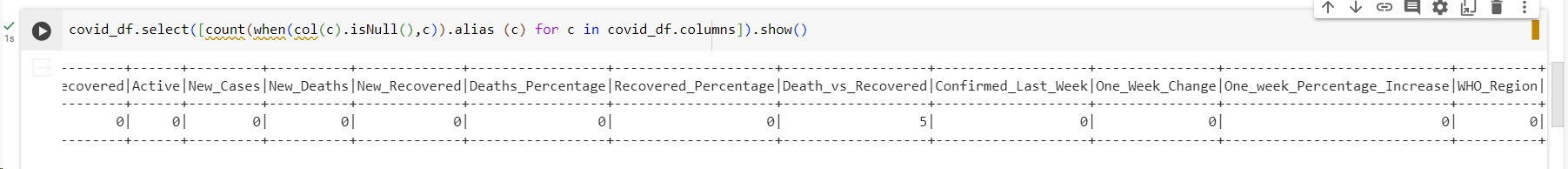
Description automatically generated

A screenshot of a computer

Description automatically generated

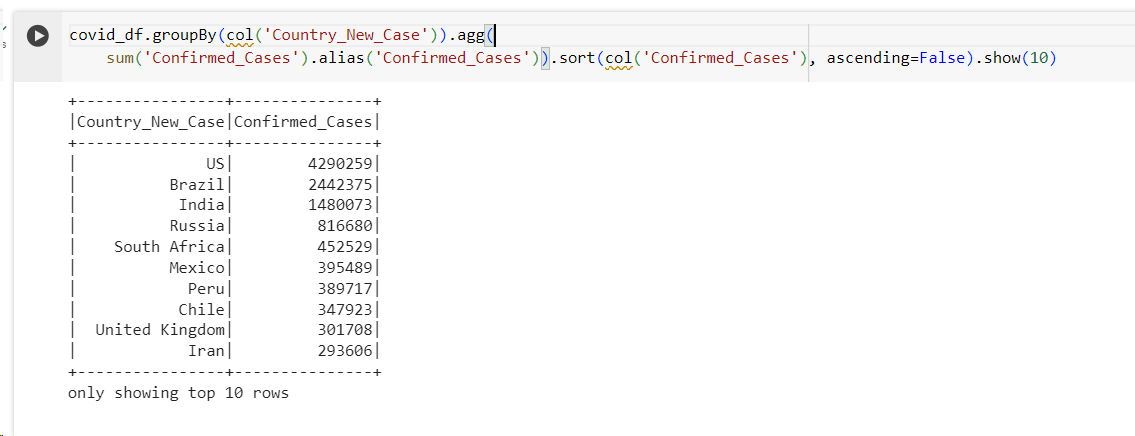
## Count and check if there any null values in any of the columns





## A screenshot of a computer Description automatically generated

## What are the top 10 countries under the WHO region with covid 19 Confirmed cases



## What are the bottom 10 countries under the WHO region with covid 19 Confirmed cases

A screenshot of a computer

Description automatically generated

## What are the total number of countries/ total no. of WHO regions and also list the various WHO regions

### Total no of countries



### Total no. of WHO regions

A close-up of a computer screen

Description automatically generated

### List the various WHO regions

A screenshot of a computer

Description automatically generated