1.1-exploratory-analysis-mortality

October 21, 2020

1 Imports

```
[1]: import sys
     import os

     import matplotlib as mpl

     project_path = os.path.abspath(os.path.join('..'))

     if project_path not in sys.path:
         sys.path.append(f'{project_path}/src/visualizations/')

     from covid_data_viz import CovidDataViz
```

2 Setup

```
[2]: mpl.rcParams['figure.figsize'] = (9, 5)
```

3 Goal

My goal is to visualize various aspect of the `COVID-19` pandemic. In this notebook we view country
level data.

4 Data sources

In this notebook I use data from the following sources: -
https://github.com/CSSEGISandData/COVID-19 - JHU CSSE COVID-19 Data. -
https://datahub.io/JohnSnowLabs/country-and-continent-codes-list - country codes and con-
tinents.

5 Data loading

```
[3]: cdv = CovidDataViz()
```

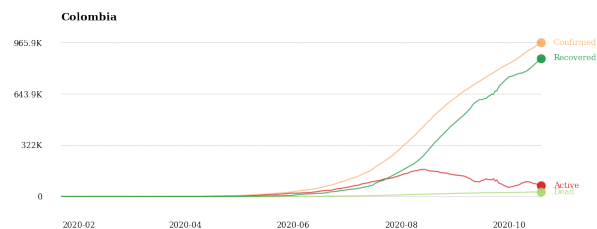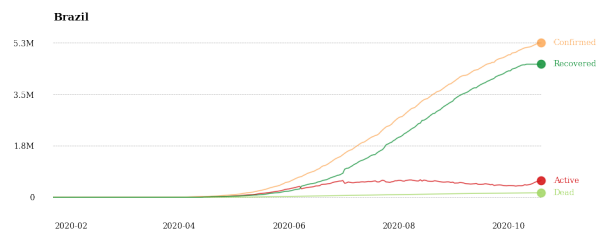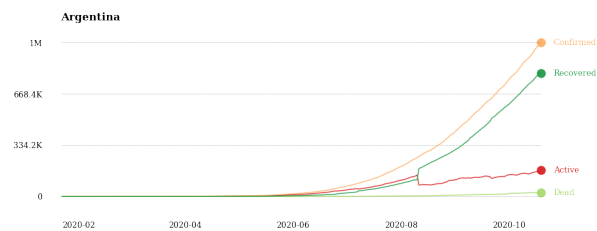# 6 Countries with most confirmed cases

Now we shall "zoom in" on the data and view it from a country level. We examine the countries listed in the previous notebook. Note that the sharp jumps in the time series are probably stem from the fact that the data collection process varies from country to country. However these plots still allow us to gain some insight.
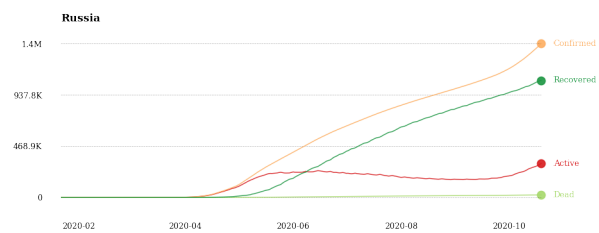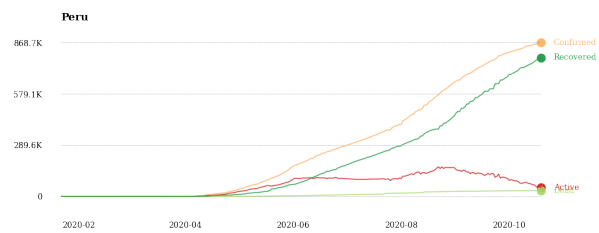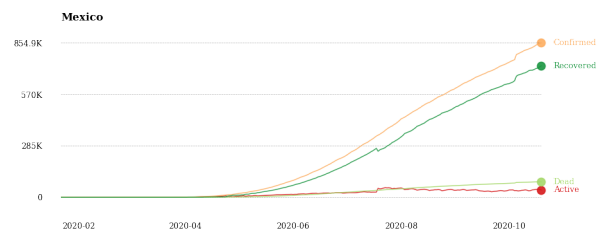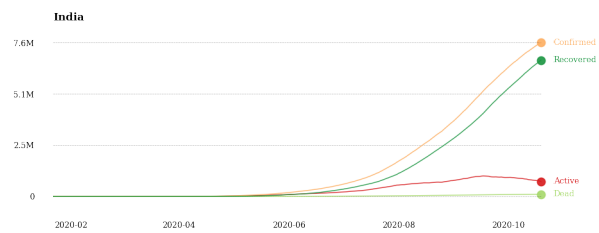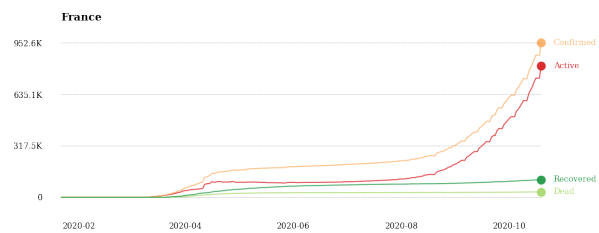
Note that the reason for `Asia`'s exponential growth is in fact `India`. Furthermore observe that `confirmed` cases are stil growing in `Spain` and `Argentina`.
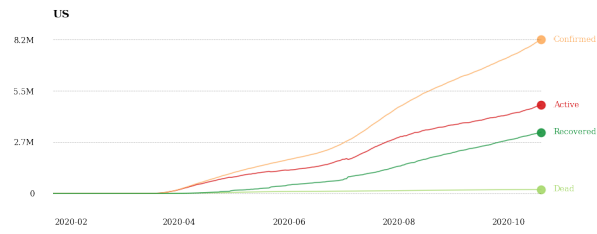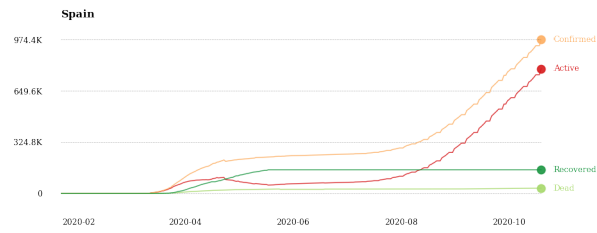
6.1 Cases over time

```
[4]: top_confirmed = sorted(cdv.get_most_cases('Confirmed')['Country'])
```

```
[5]: for country in top_confirmed:
         cdv.plot_country_cases(country=country)
```
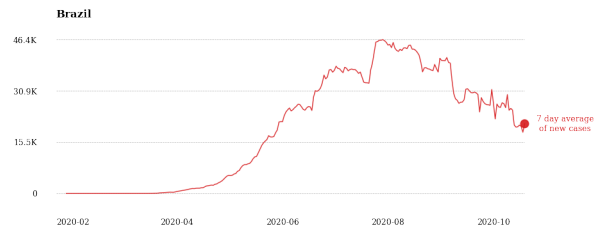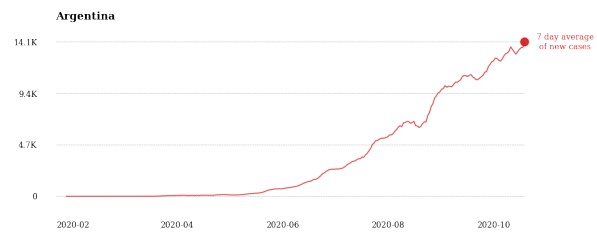
France

- Confirmed
- Active
- Recovered
- Dead

952.6K
635.1K
317.5K
0

2020-02 2020-04 2020-06 2020-08 2020-10

India

- Confirmed
- Recovered
- Active
- Dead

7.6M
5.1M
2.5M
0

2020-02 2020-04 2020-06 2020-08 2020-10

Mexico

- Confirmed
- Recovered
- Dead
- Active

854.9K
570K
285K
0

2020-02 2020-04 2020-06 2020-08 2020-10

Peru

- Confirmed
- Recovered
- Active

868.7K
579.1K
289.6K
0

2020-02 2020-04 2020-06 2020-08 2020-10

Russia

- Confirmed
- Recovered
- Active
- Dead

1.4M
937.8K
468.9K
0

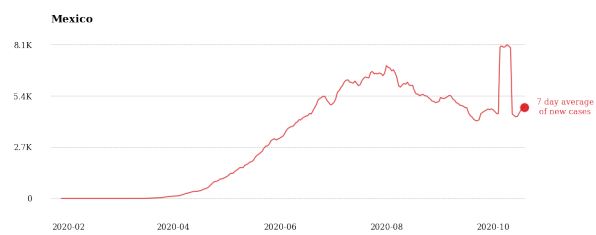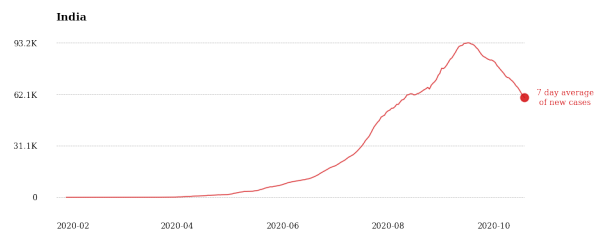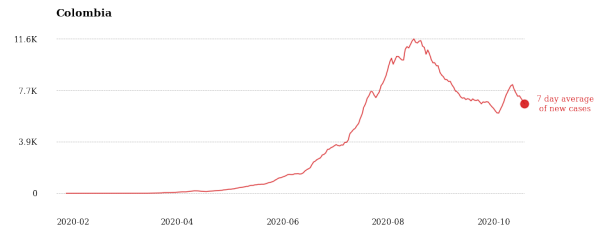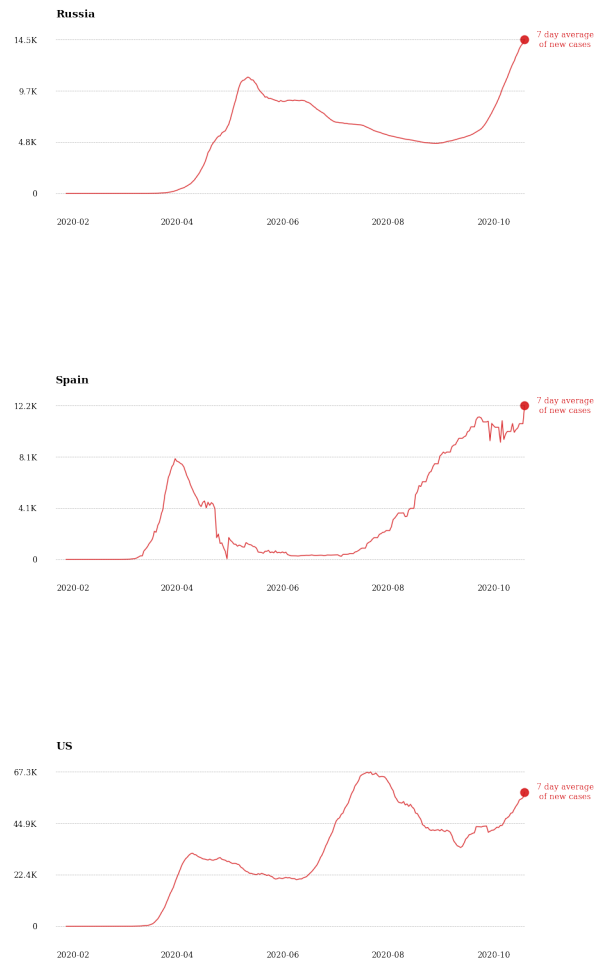2020-02 2020-04 2020-06 2020-08 2020-10

Spain

US

6.2 Change in cases over time

```
[6]: for country in top_confirmed:
         cdv.plot_country_cases_chg(country=country)
```


Argentina

Brazil

Colombia

11.6K

7.7K

3.9K

0

2020-02   2020-04   2020-06   2020-08   2020-10

7 day average
of new cases

France

25.2K

16.8K

8.4K

0

2020-02   2020-04   2020-06   2020-08   2020-10

7 day average
of new cases

India

93.2K

62.1K

31.1K

0

2020-02   2020-04   2020-06   2020-08   2020-10

7 day average
of new cases

Mexico

8.1K

5.4K

2.7K

0

2020-02   2020-04   2020-06   2020-08   2020-10

7 day average
of new cases

Peru

8.6K

5.7K

2.9K

0

2020-02   2020-04   2020-06   2020-08   2020-10

7 day average
of new cases

# 7 Countries with highest mortality rate

Note the huge surge of cases in `Italy`. This was the first country in `Europe` to be seriously hit during the pandemic and it made headlines around the world.
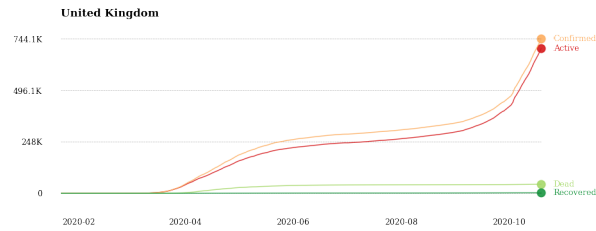
7.1 Cases over time

```
[7]: top_mortality = sorted(cdv.get_most_cases('Mortality')['Country'])
```

```
[8]: for country in top_mortality:
         cdv.plot_country_cases(country)
```
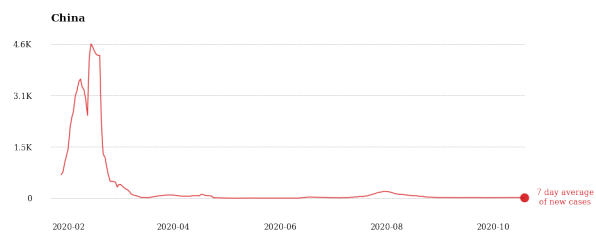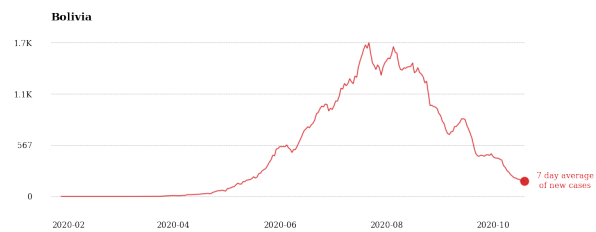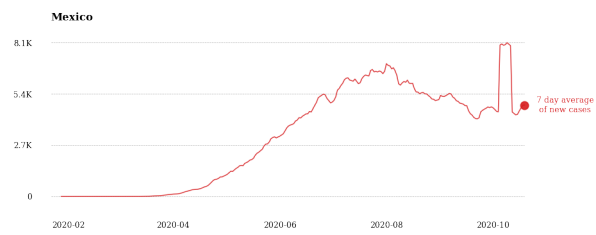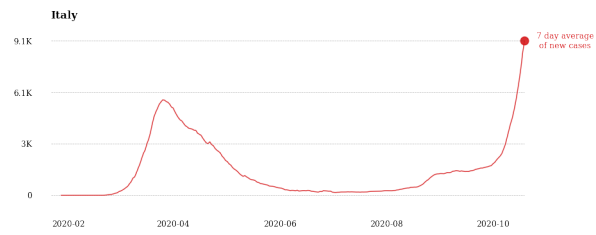
Bolivia

China

Ecuador

Egypt

Iran

Italy

Mexico

Sudan

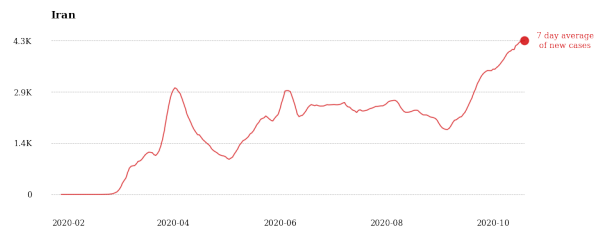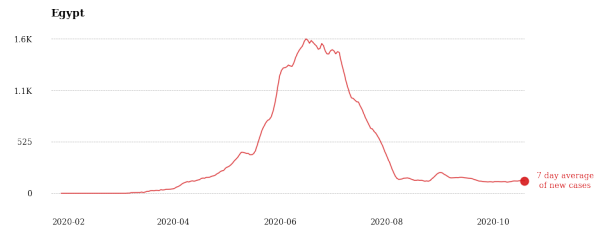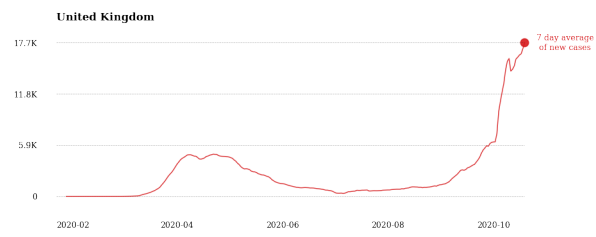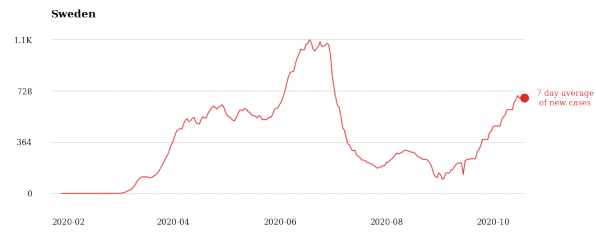**Sweden**
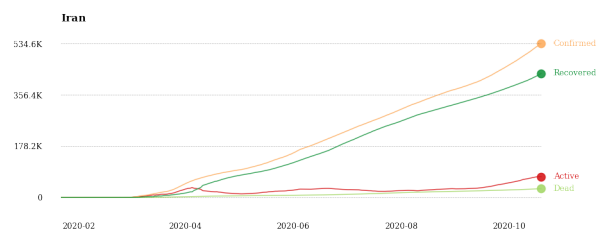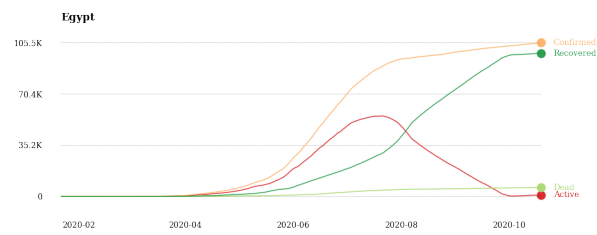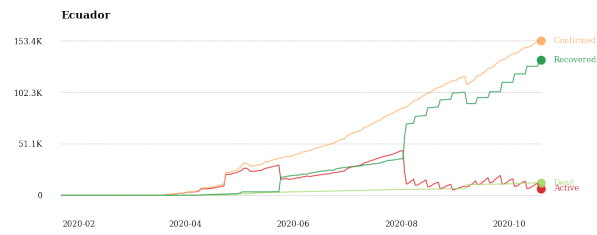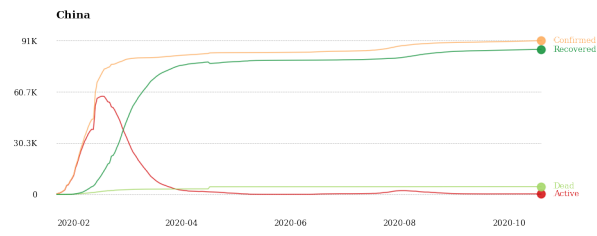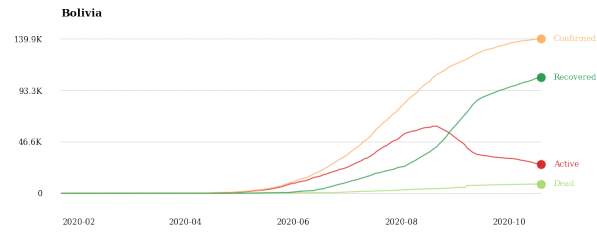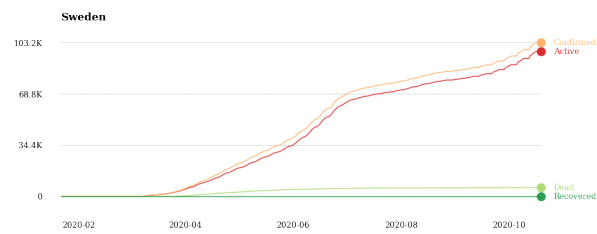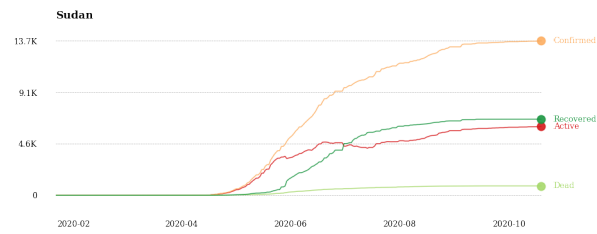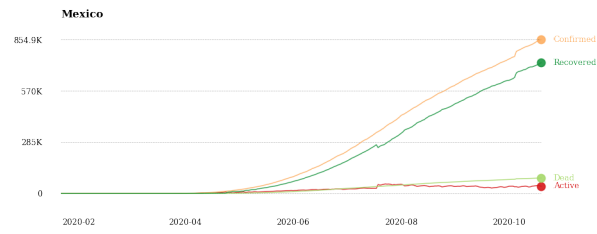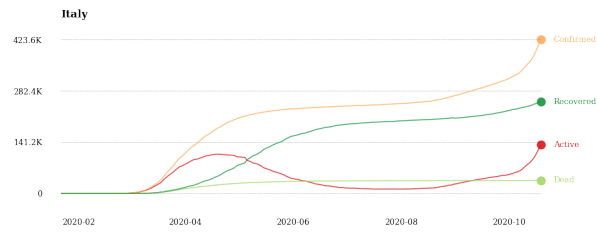
United Kingdom

7.2 Change in cases over time

Observe how the timeseries of changes in confirmed cases form different patterns. Some of these plots suggest an existence of a second wave. For example see the plot for `United Kingdom`.

```
[9]:  for country in top_mortality:
          cdv.plot_country_cases_chg(country=country)
```


Bolivia

China

Ecuador

Egypt

Iran

Italy

Mexico

Sudan

Sweden

7 day average
of new cases

United Kingdom

7 day average
of new cases

See the next notebook for more.