

COVID-19 Data Visualization Using Python

The project dives into the realm of COVID-19 data visualization using Python 3, exploring various data-driven visualizations that offer practical guidance on crafting compelling representations of COVID-19 data. Beginning with the importation of worldwide COVID-19 datasets and progressing to the creation of essential data for impactful visualizations, our goal is to uncover significant patterns and trends to contribute to the ongoing battle against the COVID-19 pandemic.

Importing modules

```
import pandas as pd
import matplotlib.pyplot as plt
import numpy as np
import datetime as dt
import warnings
warnings.filterwarnings('ignore')
plt.rcParams['figure.dpi'] = 100

dataset_url = "https://raw.githubusercontent.com/CSSEGISandData/COVID-19/master/data/countries-agggregated.csv"
df = pd.read_csv(dataset_url)
```

Let's check the dataframe of this dataset

```
df.head()
```

```
   Date      Country  Confirmed  Recovered  Deaths
0  2020-02-22  Afghanistan      0         0         0
1  2020-02-22  Afghanistan      0         0         0
2  2020-02-24  Afghanistan      0         0         0
3  2020-02-25  Afghanistan      0         0         0
4  2020-02-26  Afghanistan      0         0         0
```

```
df.tail()
```

```
   Date      Country  Confirmed  Recovered  Deaths
16596  2020-04-12  Zimbabwe      247984      0      5480
16596  2020-04-13  Zimbabwe      247980      0      5480
16596  2020-04-14  Zimbabwe      247930      0      5482
16596  2020-04-15  Zimbabwe      247937      0      5482
16597  2020-04-16  Zimbabwe      247937      0      5482
```

```
df.info()
df.describe()
```

let's do some preprocessing

```
df = df[df['Confirmed'] > 0]
```

```
df.head()
```

```
   Date      Country  Confirmed  Recovered  Deaths
33  2020-02-22  Afghanistan      5         0         0
34  2020-02-25  Afghanistan      5         0         0
35  2020-02-26  Afghanistan      5         0         0
36  2020-02-27  Afghanistan      5         0         0
37  2020-02-28  Afghanistan      5         0         0
```

let's see data related to a country for example Italy

```
df[df['Country'] == 'Italy']
```

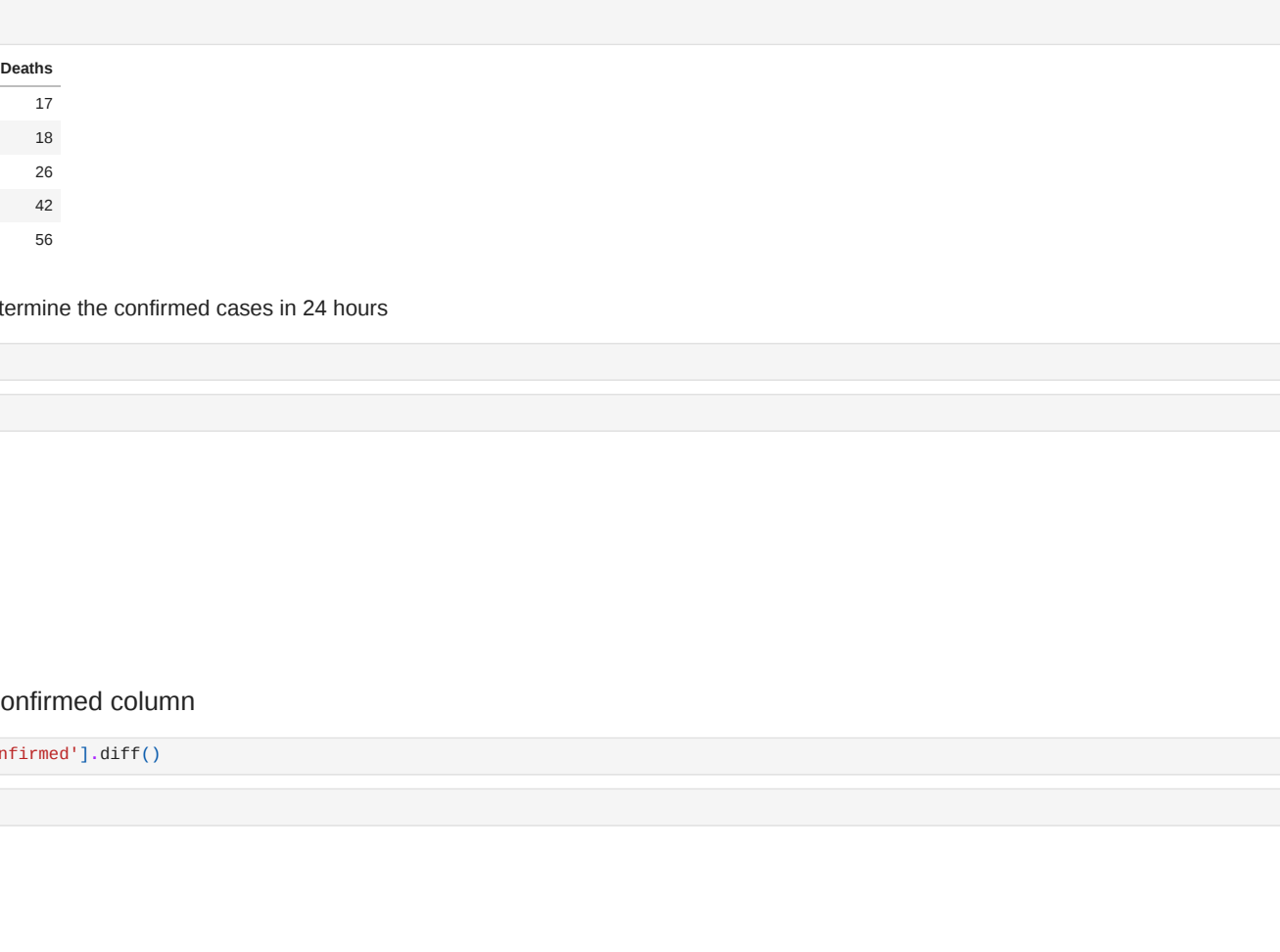
```
df.info()
df.describe()
```

```
   Date      Country  Confirmed  Recovered  Deaths
78395  2020-03-01      Italy      2         0         0
78396  2020-03-01      Italy      2         0         0
78397  2020-03-02      Italy      2         0         0
78398  2020-03-03      Italy      2         0         0
78399  2020-03-04      Italy      2         0         0
...
78987  2020-04-12      Italy     1540480      0     10352
78987  2020-04-13      Italy     15407390      0     103187
78989  2020-04-14      Italy     15502021      0     103236
78990  2020-04-15      Italy     15502021      0     103420
78991  2020-04-16      Italy     15505825      0     103502
```

807 rows x 5 columns

Let's see Global spread of Covid19

```
fig, ax = plt.subplots(df['locations'].unique(), figsize=(15, 10))
ax = ax.flatten()
```



Let's visualize how intensive the Covid-19 Transmission has been in each of the country

```
df[df['Country'] == 'China']
df.info()
df.describe()
```

```
   Date      Country  Confirmed  Recovered  Deaths
36392  2020-02-02      China      640      26      17
36393  2020-02-03      China      640      26      18
36394  2020-02-04      China      820      36      26
36395  2020-02-05      China     1438      39      42
36396  2020-02-06      China     2075      49      56
```

Let's select the columns that we need to determine the confirmed cases in 24 hours

```
df_china = df[df['date'] == '2020-01-22']
```

```
df_china.head()
```

```
   Date  Confirmed
36392  2020-01-22      540
36393  2020-01-22      640
36394  2020-01-24      820
36395  2020-01-25     1405
36396  2020-01-26     2075
```

let's calculate the first derivation of confirmed column

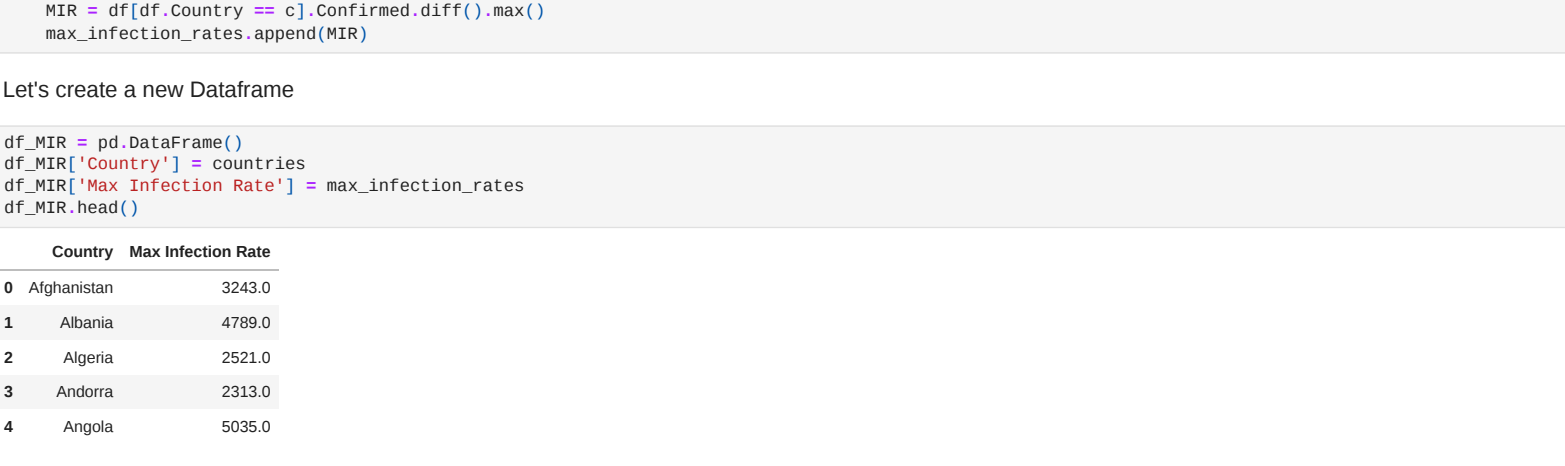
```
df_china['infection_rate'] = df_china['confirmed'].diff()
```

```
df.info()
df_china.head()
```

```
   Date  Confirmed  Infection Rate
36392  2020-01-22      540      NaN
36393  2020-01-22      640      80.0
36394  2020-01-24      820     277.0
36395  2020-01-25     1405     480.0
36396  2020-01-26     2075     650.0
```

From the table above, the infection rate on 2020-01-22 is NaN because that's the date pandemic occurred

```
fig, line = plt.subplots(x = 'date', y = ['Confirmed', 'Infection Rate'])
```



```
df_china['infection_rate'].max()
```

```
df.info()
```

Let's calculate Maximum infection rate for all of the countries

```
df.head()
```

```
   Date      Country  Confirmed  Recovered  Deaths
33  2020-02-24  Afghanistan      5         0         0
34  2020-02-25  Afghanistan      5         0         0
35  2020-02-26  Afghanistan      5         0         0
36  2020-02-27  Afghanistan      5         0         0
37  2020-02-28  Afghanistan      5         0         0
```

```
countries = list(df['Country'].unique())
max_infection_rate = []
for c in countries:
    MCR = df[df['Country'] == c].Confirmed.diff().max()
    max_infection_rate.append(MCR)
```

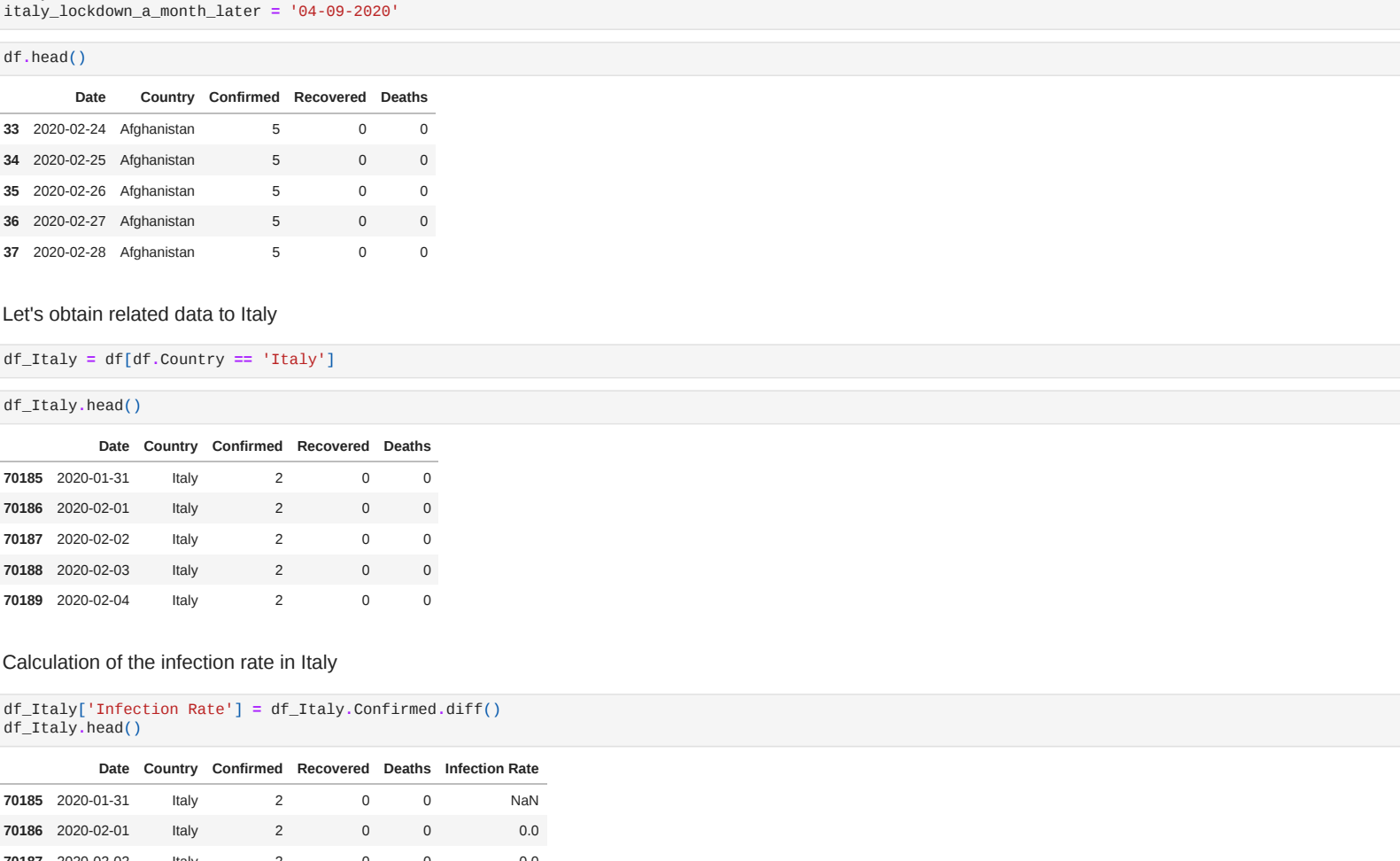
Let's create a new DataFrame

```
df_MCR = pd.DataFrame()
df_MCR['Country'] = countries
df_MCR['Max Infection Rate'] = max_infection_rate
df_MCR.head()
```

```
   Country  Max Infection Rate
0  Afghanistan      254.0
1  Albania      4786.0
2  Algeria      262.0
3  Angola      253.0
4  Angola      658.0
```

Let's plot the bar chart: maximum infection rate of each country

```
fig, bar = plt.subplots(x = 'Country', y = 'Max Infection Rate', color = 'Country', title = 'Global Maximum Infection Rate', log_y = True)
```



Impact of National Lockdowns in Italy During the Pandemic

COVID-19 Pandemic Lockdown in Italy

On March 19, 2020, the government of Italy under Prime Minister Giuseppe Conte imposed a national quarantine, restricting the movement of the population except for necessity, work, and health circumstances, in response to the growing pandemic of the COVID-19 in the country. [Source](#)

```
Italy_lockdown_start_date = "20-03-2020"
Italy_lockdown_a_month_later = "04-04-2020"
```

```
df.head()
```

```
   Date      Country  Confirmed  Recovered  Deaths
33  2020-02-24  Afghanistan      5         0         0
34  2020-02-25  Afghanistan      5         0         0
35  2020-02-26  Afghanistan      5         0         0
36  2020-02-27  Afghanistan      5         0         0
37  2020-02-28  Afghanistan      5         0         0
```

Let's obtain related data to Italy

```
df_Italy = df[df['Country'] == 'Italy']
```

```
df_Italy.head()
```

```
   Date      Country  Confirmed  Recovered  Deaths
78395  2020-03-01      Italy      2         0         0
78396  2020-03-01      Italy      2         0         0
78397  2020-03-02      Italy      2         0         0
78398  2020-03-03      Italy      2         0         0
78399  2020-03-04      Italy      2         0         0
```

Calculation of the infection rate in Italy

```
df_Italy['infection_rate'] = df_Italy['Confirmed'].diff()
df_Italy.head()
```

```
   Date      Country  Confirmed  Recovered  Deaths  Infection Rate
78395  2020-03-01      Italy      2         0         0      NaN
78396  2020-03-01      Italy      2         0         0      0.0
78397  2020-03-02      Italy      2         0         0      0.0
78398  2020-03-03      Italy      2         0         0      0.0
78399  2020-03-04      Italy      2         0         0      0.0
```

```
fig = plt.figure(df_Italy, x = 'date', y = ['Infection Rate', 'Before and After Lockdown in Italy'], log_y = True)
fig.add_subplot(2, 1, 1)
```

```
df_Italy.info()
df_Italy.describe()
```

```
   Date      Country  Confirmed  Recovered  Deaths  Infection Rate
78395  2020-03-01      Italy      2         0         0      NaN
78396  2020-03-01      Italy      2         0         0      0.0
78397  2020-03-02      Italy      2         0         0      0.0
78398  2020-03-03      Italy      2         0         0      0.0
78399  2020-03-04      Italy      2         0         0      0.0
```

```
fig.add_subplot(2, 1, 2)
```

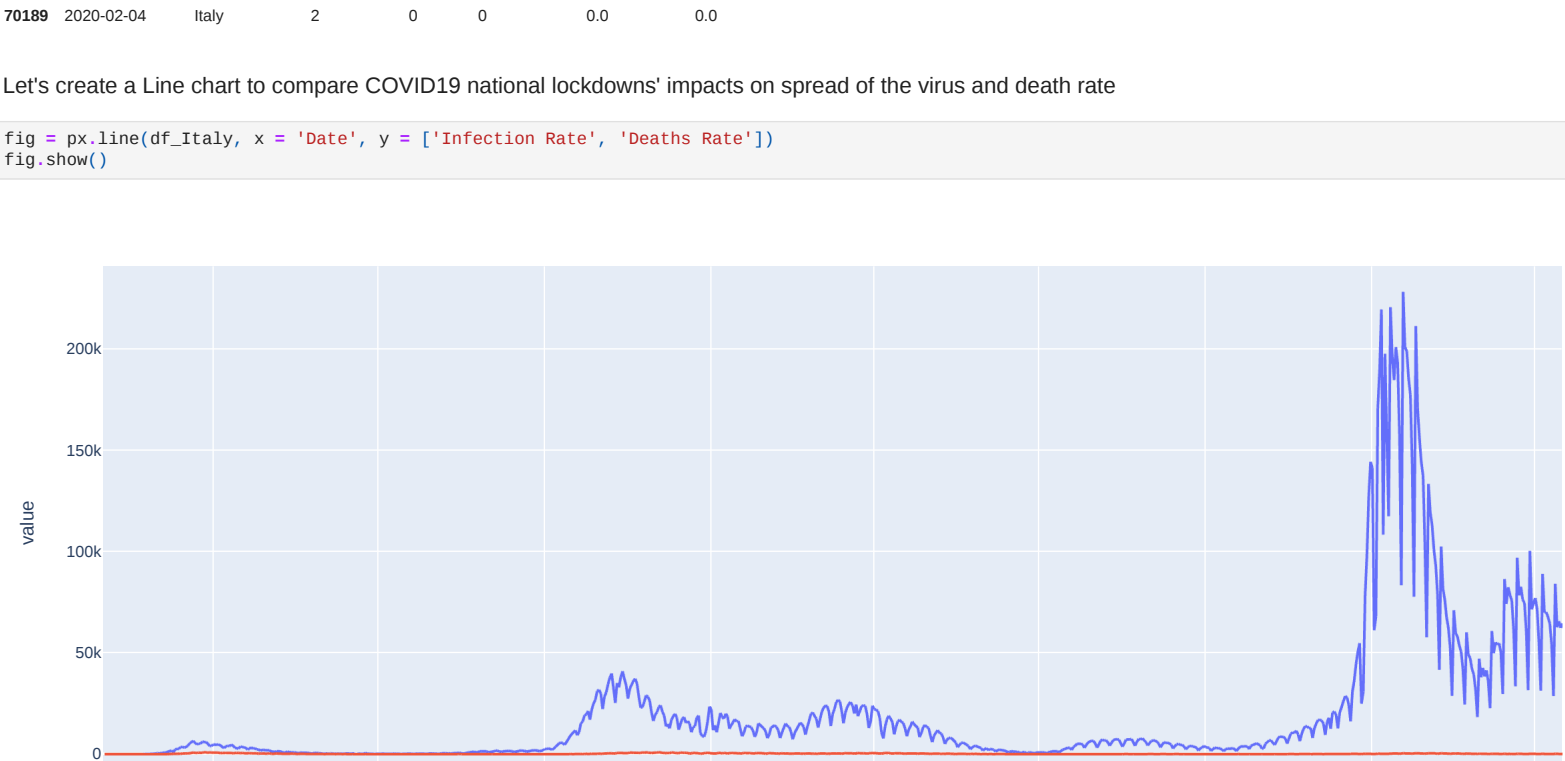
```
   Date      Country  Confirmed  Recovered  Deaths  Infection Rate
78395  2020-03-01      Italy      2         0         0      NaN
78396  2020-03-01      Italy      2         0         0      0.0
78397  2020-03-02      Italy      2         0         0      0.0
78398  2020-03-03      Italy      2         0         0      0.0
78399  2020-03-04      Italy      2         0         0      0.0
```

```
fig.add_subplot(2, 1, 3)
```

```
   Date      Country  Confirmed  Recovered  Deaths  Infection Rate
78395  2020-03-01      Italy      2         0         0      NaN
78396  2020-03-01      Italy      2         0         0      0.0
78397  2020-03-02      Italy      2         0         0      0.0
78398  2020-03-03      Italy      2         0         0      0.0
78399  2020-03-04      Italy      2         0         0      0.0
```

Let's create a Line chart to compare COVID19 national lockdowns' impacts on spread of the virus and death rate

```
fig = plt.figure(df_Italy, x = 'Date', y = ['Infection Rate', 'Deaths Rate'])
fig.subplots()
```



```
df_Italy['infection_rate'] = df_Italy['infection_rate'].diff()
df_Italy['death_rate'] = df_Italy['Deaths Rate'].diff()
```

Let's create a line chart to display the Infection and Deaths Rate of Italy

```
df_Italy.info()
df_Italy.describe()
```

```
   Date      Country  Confirmed  Recovered  Deaths  Infection Rate
78395  2020-03-01      Italy      2         0         0      NaN
78396  2020-03-01      Italy      2         0         0      0.0
78397  2020-03-02      Italy      2         0         0      0.0
78398  2020-03-03      Italy      2         0         0      0.0
78399  2020-03-04      Italy      2         0         0      0.0
```

```
fig.add_subplot(2, 1, 1)
```

```
   Date      Country  Confirmed  Recovered  Deaths  Infection Rate
78395  2020-03-01      Italy      2         0         0      NaN
78396  2020-03-01      Italy      2         0         0      0.0
78397  2020-03-02      Italy      2         0         0      0.0
78398  2020-03-03      Italy      2         0         0      0.0
78399  2020-03-04      Italy      2         0         0      0.0
```

```
fig.add_subplot(2, 1, 2)
```

```
   Date      Country  Confirmed  Recovered  Deaths  Infection Rate
78395  2020-03-01      Italy      2         0         0      NaN
78396  2020-03-01      Italy      2         0         0      0.0
78397  2020-03-02      Italy      2         0         0      0.0
78398  2020-03-03      Italy      2         0         0      0.0
78399  2020-03-04      Italy      2         0         0      0.0
```

```
fig.add_subplot(2, 1, 3)
```

```
   Date      Country  Confirmed  Recovered  Deaths  Infection Rate
78395  2020-03-01      Italy      2         0         0      NaN
78396  2020-03-01      Italy      2         0         0      0.0
78397  2020-03-02      Italy      2         0         0      0.0
78398  2020-03-03      Italy      2         0         0      0.0
78399  2020-03-04      Italy      2         0         0      0.0
```

```
fig.add_subplot(2, 1, 4)
```

```
   Date      Country  Confirmed  Recovered  Deaths  Infection Rate
78395  2020-03-01      Italy      2         0         0      NaN
78396  2020-03-01      Italy      2         0         0      0.0
78397  2020-03-02      Italy      2         0         0      0.0
78398  2020-03-03      Italy      2         0         0      0.0
78399  2020-03-04      Italy      2         0         0      0.0
```

```
fig.add_subplot(2, 1, 5)
```

```
   Date      Country  Confirmed  Recovered  Deaths  Infection Rate
78395  2020-03-01      Italy      2         0         0      NaN
78396  2020-03-01      Italy      2         0         0      0.0
78397  2020-03-02      Italy      2         0         0      0.0
78398  2020-03-03      Italy      2         0         0      0.0
78399  2020-03-04      Italy      2         0         0      0.0
```

```
fig.add_subplot(2, 1, 6)
```

```
   Date      Country  Confirmed  Recovered  Deaths  Infection Rate
78395  2020-03-01      Italy      2         0         0      NaN
78396  2020-03-01      Italy      2         0         0      0.0
78397  2020-03-02      Italy      2         0         0      0.0
78398  2020-03-03      Italy      2         0         0      0.0
78399  2020-03-04      Italy      2         0         0      0.0
```

```
fig.add_subplot(2, 1, 7)
```

```
   Date      Country  Confirmed  Recovered  Deaths  Infection Rate
78395  2020-03-01      Italy      2         0         0      NaN
78396  2020-03-01      Italy      2         0         0      0.0
78397  2020-03-02      Italy      2         0         0      0.0
78398  2020-03-03      Italy      2         0         0      0.0
78399  2020-03-04      Italy      2         0         0      0.0
```

```
fig.add_subplot(2, 1, 8)
```

```
   Date      Country  Confirmed  Recovered  Deaths  Infection Rate
78395  2020-03-01      Italy      2         0         0      NaN
78396  2020-03-01      Italy      2         0         0      0.0
78397  2020-03-02      Italy      2         0         0      0.0
78398  2020-03-03      Italy      2         0         0      0.0
78399  2020-03-04      Italy      2         0         0      0.0
```

```
fig.add_subplot(2, 1, 9)
```

```
   Date      Country  Confirmed  Recovered  Deaths  Infection Rate
78395  2020-03-01      Italy      2         0         0      NaN
78396  2020-03-01      Italy      2         0         0      0.0
78397  2020-03-02      Italy      2         0         0      0.0
78398  2020-03-03      Italy      2         0         0      0.0
78399  2020-03-04      Italy      2         0         0      0.0
```

```
fig.add_subplot(2, 1, 10)
```

```
   Date      Country  Confirmed  Recovered  Deaths  Infection Rate
78395  2020-03-01      Italy      2         0         0      NaN
78396  2020-03-01      Italy      2         0         0      0.0
78397  2020-03-02      Italy      2         0         0      0.0
78398  2020-03-03      Italy      2         0         0      0.0
78399  2020-03-04      Italy      2         0         0      0.0
```

```
fig.add_subplot(2, 1, 11)
```

```
   Date      Country  Confirmed  Recovered  Deaths  Infection Rate
78395  2020-03-01      Italy      2         0         0      NaN
78396  2020-03-01      Italy      2         0         0      0.0
78397  2020-03-02      Italy      2         0         0      0.0
78398  2020-03-03      Italy      2         0         0      0.0
78399  2020-03-04      Italy      2         0         0      0.0
```

```
fig.add_subplot(2, 1, 12)
```

```
   Date      Country  Confirmed  Recovered  Deaths  Infection Rate
78395  2020-03-01      Italy      2         0         0      NaN
78396  2020-03-01      Italy      2         0         0      0.0
78397  2020-03-02      Italy      2         0         0      0.0
78398  2020-03-03      Italy      2         0         0      0.0
78399  2020-03-04      Italy      2         0         0      0.0
```

```
fig.add_subplot(2, 1, 13)
```

```
   Date      Country  Confirmed  Recovered  Deaths  Infection Rate
78395  2020-03-01      Italy      2         0         0      NaN
78396  2020-03-01      Italy      2         0         0      0.0
78397  2020-03-02      Italy      2         0         0      0.0
78398  2020-03-03      Italy      2         0         0      0.0
78399  2020-03-04      Italy      2         0         0      0.0
```

```
fig.add_subplot(2, 1, 14)
```

```
   Date      Country  Confirmed  Recovered  Deaths  Infection Rate
78395  2020-03-01      Italy      2         0         0      NaN
78396  2020-03-01      Italy      2         0         0      0.0
78397  2020-03-02      Italy      2         0         0      0.0
78398  2020-03-03      Italy      2         0         0      0.0
78399  2020-03-04      Italy      2         0         0      0.0
```

```
fig.add_subplot(2, 1, 15)
```

```
   Date      Country  Confirmed  Recovered  Deaths  Infection Rate
78395  2020-03-01      Italy      2         0         0      NaN
78396  2020-03-01      Italy      2         0         0      0.0
78397  2020-03-02      Italy      2         0         0      0.0
78398  2020-03-03      Italy      2         0         0      0.0
78399  2020-03-04      Italy      2         0         0      0.0
```

```
fig.add_subplot(2, 1, 16)
```

```
   Date      Country  Confirmed  Recovered  Deaths  Infection Rate
78395  2020-03-01      Italy      2         0         0      NaN
78396  2020-03-01      Italy      2         0         0      0.0
78397  2020-03-02      Italy      2         0         0      0.0
78398  2020-03-03      Italy      2         0         0      0.0
78399  2020-03-04      Italy      2         0         0      0.0
```

```
fig.add_subplot(2, 1, 17)
```

```
   Date      Country  Confirmed  Recovered  Deaths  Infection Rate
78395  2020-03-01      Italy      2         0         0      NaN
78396  2020-03-01      Italy      2         0         0      0.0
78397  2020-03-02      Italy      2         0         0      0.0
78398  2020-03-03      Italy      2         0         0      0.0
78399  2020-03-04      Italy      2         0         0      0.0
```

```
fig.add_subplot(2, 1, 18)
```

```
   Date      Country  Confirmed  Recovered  Deaths  Infection Rate
78395  2020-03-01      Italy      2         0         0      NaN
78396  2020-03-01      Italy      2         0         0      0.0
78397  2020-03-02      Italy      2         0         0      0.0
78398  2020-03-03      Italy      2         0         0      0.0
78399  2020-03-04      Italy      2         0         0      0.0
```

```
fig.add_subplot(2, 1, 19)
```

```
   Date      Country  Confirmed  Recovered  Deaths  Infection Rate
78395  2020-03-01      Italy      2         0         0      NaN
78396  2020-03-01      Italy      2         0         0      0.0
78397  2020-03-02      Italy      2         0         0      0.0
78398  2020-03-03      Italy      2         0         0      0.0
78399  2020-03-04      Italy      2         0         0      0.0
```

```
fig.add_subplot(2, 1, 20)
```

```
   Date      Country  Confirmed  Recovered  Deaths  Infection Rate
78395  2020-03-01      Italy      2         0         0      NaN
78396  2020-03-01      Italy      2         0         0      0.0
78397  2020-03-02      Italy      2         0         0      0.0
78398  2020-03-03      Italy      2         0         0      0.0
78399  2020-03-04      Italy      2         0         0      0.0
```

```
fig.add_subplot(2, 1, 21)
```

```
   Date      Country  Confirmed  Recovered  Deaths  Infection Rate
78395  2020-03-01      Italy      2         0         0      NaN
78396  2020-03-01      Italy      2         0         0      0.0
78397  2020-03-02      Italy      2         0         0      0.0
78398  2020-03-03      Italy      2         0         0      0.0
78399  2020-03-04      Italy      2         0         0      0.0
```

```
fig.add_subplot(2, 1, 22)
```

```
   Date      Country  Confirmed  Recovered  Deaths  Infection Rate
78395  2020-03-01      Italy      2         0         0      NaN
78396  2020-03-01      Italy      2         0         0      0.0
78397  2020-03-02      Italy      2         0         0      0.0
78398  2020-03-03      Italy      2         0         0      0.0
78399  2020-03-04      Italy      2         0         0      0.0
```

```
fig.add_subplot(2, 1, 23)
```

```
   Date      Country  Confirmed  Recovered  Deaths  Infection Rate
78395  2020-03-01      Italy      2         0         0      NaN
78396  2020-03-01      Italy      2         0         0      0.0
78397  2020-03-02      Italy      2         0         0      0.0
78398  2020-03-03      Italy      2         0         0      0.0
78399  2020-03-04      Italy      2         0         0      0.0
```

```
fig.add_subplot(2, 1, 24)
```

```
   Date      Country  Conf
```