# Untitled

September 16, 2022

### 0.0.1 The Data is downloaded from the website our world in data which is a tabular representation of early covid-19 cases in India.

Here's what we can tell by looking at the dataframe: • The file provides four day-wise counts for COVID-19 in India • The metrics reported are new cases, deaths, and patients cured • Data is provided for 959 days: from Jan 30, 2020, to Sept 15, 2022 Keep in mind that these are officially reported numbers. The actual number of cases & deaths may be higher, as not all cases are diagnosed.

[1]:
```
from urllib.request import urlretrieve
```

[61]:
```
india_covid_url = 'https://gist.githubusercontent.com/shekharankur4/
↪67fc5eb4db9a5a2bda98c02ebd792b71/raw/
↪01109aaf827f2362fcecf39baab237ced159f076/covid-india-daywise.csv'

urlretrieve(india_covid_url, 'india-covid-daywise.csv')
```

[61]:
```
('india-covid-daywise.csv', <http.client.HTTPMessage at 0x7fb16cbe8700>)
```

[62]:
```
!pip install pandas --upgrade --quiet
```

[63]:
```
import pandas as pd
```

[64]:
```
covid_df = pd.read_csv('india-covid-daywise.csv')
```

[65]:
```
type(covid_df)
```

[65]:
```
pandas.core.frame.DataFrame
```

[66]:
```
covid_df
```

[66]:

| | iso_code | continent | location | date | total_cases | new_cases | \ |
|---|---|---|---|---|---|---|---|
| 0 | IND | Asia | India | 2020-01-30 | 1.0 | 1.0 | |
| 1 | IND | Asia | India | 2020-01-31 | 1.0 | 0.0 | |
| 2 | IND | Asia | India | 2020-02-01 | 1.0 | 0.0 | |
| 3 | IND | Asia | India | 2020-02-02 | 2.0 | 1.0 | |
| 4 | IND | Asia | India | 2020-02-03 | 3.0 | 1.0 | |
| .. | … | … | … | … | … | … | |

```
955      IND      Asia    India  2022-09-11   44500580.0     5221.0
956      IND      Asia    India  2022-09-12   44502363.0     1783.0
957      IND      Asia    India  2022-09-13   44510057.0     7694.0
958      IND      Asia    India  2022-09-14   44516479.0     6422.0
959      IND      Asia    India  2022-09-15   44522777.0     6298.0

     new_cases_smoothed  total_deaths  new_deaths  new_deaths_smoothed  …  \
0                   NaN           NaN         NaN                  NaN  …
1                   NaN           NaN         NaN                  NaN  …
2                   NaN           NaN         NaN                  NaN  …
3                   NaN           NaN         NaN                  NaN  …
4                   NaN           NaN         NaN                  NaN  …
..                  …             …           …                    …  …
955            5447.857      528165.0        15.0               22.571  …
956            5071.571      528185.0        20.0               22.143  …
957            5770.857      528216.0        31.0               22.714  …
958            5406.143      528250.0        34.0               22.857  …
959            5435.429      528273.0        23.0               21.714  …

     female_smokers  male_smokers  handwashing_facilities  \
0               1.9          20.6                   59.55
1               1.9          20.6                   59.55
2               1.9          20.6                   59.55
3               1.9          20.6                   59.55
4               1.9          20.6                   59.55
..              …            …                      …
955             1.9          20.6                   59.55
956             1.9          20.6                   59.55
957             1.9          20.6                   59.55
958             1.9          20.6                   59.55
959             1.9          20.6                   59.55

     hospital_beds_per_thousand  life_expectancy  human_development_index  \
0                          0.53            69.66                    0.645
1                          0.53            69.66                    0.645
2                          0.53            69.66                    0.645
3                          0.53            69.66                    0.645
4                          0.53            69.66                    0.645
..                          …              …                        …
955                        0.53            69.66                    0.645
956                        0.53            69.66                    0.645
957                        0.53            69.66                    0.645
958                        0.53            69.66                    0.645
959                        0.53            69.66                    0.645

     excess_mortality_cumulative_absolute  excess_mortality_cumulative  \
0                                     NaN                          NaN
```

```
1                                    NaN                                  NaN
2                                    NaN                                  NaN
3                                    NaN                                  NaN
4                                    NaN                                  NaN
..                                   …                                    …
955                                  NaN                                  NaN
956                                  NaN                                  NaN
957                                  NaN                                  NaN
958                                  NaN                                  NaN
959                                  NaN                                  NaN

     excess_mortality  excess_mortality_cumulative_per_million
0                 NaN                                     NaN
1                 NaN                                     NaN
2                 NaN                                     NaN
3                 NaN                                     NaN
4                 NaN                                     NaN
..                …                                       …
955               NaN                                     NaN
956               NaN                                     NaN
957               NaN                                     NaN
958               NaN                                     NaN
959               NaN                                     NaN

[960 rows x 67 columns]
```

[67]:
```
covid_df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 960 entries, 0 to 959
Data columns (total 67 columns):
 #   Column                                Non-Null Count  Dtype
---  ------                                --------------  -----
 0   iso_code                              960 non-null    object
 1   continent                             960 non-null    object
 2   location                              960 non-null    object
 3   date                                  960 non-null    object
 4   total_cases                           960 non-null    float64
 5   new_cases                             960 non-null    float64
 6   new_cases_smoothed                    955 non-null    float64
 7   total_deaths                          919 non-null    float64
 8   new_deaths                            917 non-null    float64
 9   new_deaths_smoothed                   912 non-null    float64
 10  total_cases_per_million               960 non-null    float64
 11  new_cases_per_million                 960 non-null    float64
 12  new_cases_smoothed_per_million        955 non-null    float64
 13  total_deaths_per_million              919 non-null    float64
```

```
14  new_deaths_per_million                          917 non-null    float64
15  new_deaths_smoothed_per_million                 912 non-null    float64
16  reproduction_rate                               912 non-null    float64
17  icu_patients                                    0 non-null      float64
18  icu_patients_per_million                        0 non-null      float64
19  hosp_patients                                   0 non-null      float64
20  hosp_patients_per_million                       0 non-null      float64
21  weekly_icu_admissions                           0 non-null      float64
22  weekly_icu_admissions_per_million               0 non-null      float64
23  weekly_hosp_admissions                          0 non-null      float64
24  weekly_hosp_admissions_per_million              0 non-null      float64
25  total_tests                                     810 non-null    float64
26  new_tests                                       795 non-null    float64
27  total_tests_per_thousand                        810 non-null    float64
28  new_tests_per_thousand                          795 non-null    float64
29  new_tests_smoothed                              824 non-null    float64
30  new_tests_smoothed_per_thousand                 824 non-null    float64
31  positive_rate                                   816 non-null    float64
32  tests_per_case                                  816 non-null    float64
33  tests_units                                     831 non-null    object
34  total_vaccinations                              589 non-null    float64
35  people_vaccinated                               589 non-null    float64
36  people_fully_vaccinated                         560 non-null    float64
37  total_boosters                                  236 non-null    float64
38  new_vaccinations                                572 non-null    float64
39  new_vaccinations_smoothed                       608 non-null    float64
40  total_vaccinations_per_hundred                  589 non-null    float64
41  people_vaccinated_per_hundred                   589 non-null    float64
42  people_fully_vaccinated_per_hundred             560 non-null    float64
43  total_boosters_per_hundred                      236 non-null    float64
44  new_vaccinations_smoothed_per_million           608 non-null    float64
45  new_people_vaccinated_smoothed                  608 non-null    float64
46  new_people_vaccinated_smoothed_per_hundred      608 non-null    float64
47  stringency_index                                958 non-null    float64
48  population                                      960 non-null    float64
49  population_density                              960 non-null    float64
50  median_age                                      960 non-null    float64
51  aged_65_older                                   960 non-null    float64
52  aged_70_older                                   960 non-null    float64
53  gdp_per_capita                                  960 non-null    float64
54  extreme_poverty                                 960 non-null    float64
55  cardiovasc_death_rate                           960 non-null    float64
56  diabetes_prevalence                             960 non-null    float64
57  female_smokers                                  960 non-null    float64
58  male_smokers                                    960 non-null    float64
59  handwashing_facilities                          960 non-null    float64
60  hospital_beds_per_thousand                      960 non-null    float64
61  life_expectancy                                 960 non-null    float64
```

```
 62  human_development_index                      960 non-null    float64
 63  excess_mortality_cumulative_absolute         0 non-null      float64
 64  excess_mortality_cumulative                  0 non-null      float64
 65  excess_mortality                             0 non-null      float64
 66  excess_mortality_cumulative_per_million      0 non-null      float64
dtypes: float64(62), object(5)
memory usage: 502.6+ KB
```

[68]:
```
covid_df.describe()
```

[68]:

| | total_cases | new_cases | new_cases_smoothed | total_deaths | \ |
|---|---|---|---|---|---|
| count | 9.600000e+02 | 960.000000 | 955.000000 | 919.000000 | |
| mean | 2.220383e+07 | 46377.892708 | 46602.526549 | 296126.988030 | |
| std | 1.689469e+07 | 76017.973083 | 75503.596419 | 200964.140602 | |
| min | 1.000000e+00 | 0.000000 | 0.000000 | 1.000000 | |
| 25% | 5.970382e+06 | 6818.750000 | 7245.357000 | 119756.000000 | |
| 50% | 2.685066e+07 | 18537.000000 | 18315.000000 | 374305.000000 | |
| 75% | 3.798062e+07 | 46344.000000 | 46230.428500 | 493644.500000 | |
| max | 4.452278e+07 | 414188.000000 | 391232.000000 | 528273.000000 | |

| | new_deaths | new_deaths_smoothed | total_cases_per_million | \ |
|---|---|---|---|---|
| count | 917.000000 | 912.000000 | 960.000000 | |
| mean | 568.047983 | 568.726985 | 15774.653869 | |
| std | 850.264969 | 811.462150 | 12002.790863 | |
| min | 0.000000 | 0.286000 | 0.001000 | |
| 25% | 77.000000 | 93.821250 | 4241.641750 | |
| 50% | 306.000000 | 323.928500 | 19075.980500 | |
| 75% | 614.000000 | 638.107000 | 26983.233500 | |
| max | 4529.000000 | 4190.000000 | 31631.089000 | |

| | new_cases_per_million | new_cases_smoothed_per_million | \ |
|---|---|---|---|
| count | 960.000000 | 955.000000 | |
| mean | 32.949055 | 33.108658 | |
| std | 54.006787 | 53.641346 | |
| min | 0.000000 | 0.000000 | |
| 25% | 4.844500 | 5.147500 | |
| 50% | 13.169500 | 13.012000 | |
| 75% | 32.924750 | 32.844000 | |
| max | 294.259000 | 277.950000 | |

| | total_deaths_per_million | … | female_smokers | male_smokers | \ |
|---|---|---|---|---|---|
| count | 919.000000 | … | 9.600000e+02 | 9.600000e+02 | |
| mean | 210.382641 | … | 1.900000e+00 | 2.060000e+01 | |
| std | 142.774451 | … | 3.110245e-14 | 3.910022e-14 | |
| min | 0.001000 | … | 1.900000e+00 | 2.060000e+01 | |
| 25% | 85.080500 | … | 1.900000e+00 | 2.060000e+01 | |
| 50% | 265.924000 | … | 1.900000e+00 | 2.060000e+01 | |

```
75%                  350.708500  …    1.900000e+00  2.060000e+01
max                  375.310000  …    1.900000e+00  2.060000e+01

       handwashing_facilities  hospital_beds_per_thousand  life_expectancy  \
count            9.600000e+02                9.600000e+02     9.600000e+02
mean             5.955000e+01                5.300000e-01     6.966000e+01
std              9.526236e-13                1.299638e-14     1.720410e-12
min              5.955000e+01                5.300000e-01     6.966000e+01
25%              5.955000e+01                5.300000e-01     6.966000e+01
50%              5.955000e+01                5.300000e-01     6.966000e+01
75%              5.955000e+01                5.300000e-01     6.966000e+01
max              5.955000e+01                5.300000e-01     6.966000e+01

       human_development_index  excess_mortality_cumulative_absolute  \
count             9.600000e+02                                   0.0
mean              6.450000e-01                                   NaN
std               8.886414e-15                                   NaN
min               6.450000e-01                                   NaN
25%               6.450000e-01                                   NaN
50%               6.450000e-01                                   NaN
75%               6.450000e-01                                   NaN
max               6.450000e-01                                   NaN

       excess_mortality_cumulative  excess_mortality  \
count                          0.0               0.0
mean                           NaN               NaN
std                            NaN               NaN
min                            NaN               NaN
25%                            NaN               NaN
50%                            NaN               NaN
75%                            NaN               NaN
max                            NaN               NaN

       excess_mortality_cumulative_per_million
count                                      0.0
mean                                       NaN
std                                        NaN
min                                        NaN
25%                                        NaN
50%                                        NaN
75%                                        NaN
max                                        NaN

[8 rows x 62 columns]
```

[69]:
```
covid_df.columns
```

[69]:
```
Index(['iso_code', 'continent', 'location', 'date', 'total_cases', 'new_cases',
       'new_cases_smoothed', 'total_deaths', 'new_deaths',
       'new_deaths_smoothed', 'total_cases_per_million',
       'new_cases_per_million', 'new_cases_smoothed_per_million',
       'total_deaths_per_million', 'new_deaths_per_million',
       'new_deaths_smoothed_per_million', 'reproduction_rate', 'icu_patients',
       'icu_patients_per_million', 'hosp_patients',
       'hosp_patients_per_million', 'weekly_icu_admissions',
       'weekly_icu_admissions_per_million', 'weekly_hosp_admissions',
       'weekly_hosp_admissions_per_million', 'total_tests', 'new_tests',
       'total_tests_per_thousand', 'new_tests_per_thousand',
       'new_tests_smoothed', 'new_tests_smoothed_per_thousand',
       'positive_rate', 'tests_per_case', 'tests_units', 'total_vaccinations',
       'people_vaccinated', 'people_fully_vaccinated', 'total_boosters',
       'new_vaccinations', 'new_vaccinations_smoothed',
       'total_vaccinations_per_hundred', 'people_vaccinated_per_hundred',
       'people_fully_vaccinated_per_hundred', 'total_boosters_per_hundred',
       'new_vaccinations_smoothed_per_million',
       'new_people_vaccinated_smoothed',
       'new_people_vaccinated_smoothed_per_hundred', 'stringency_index',
       'population', 'population_density', 'median_age', 'aged_65_older',
       'aged_70_older', 'gdp_per_capita', 'extreme_poverty',
       'cardiovasc_death_rate', 'diabetes_prevalence', 'female_smokers',
       'male_smokers', 'handwashing_facilities', 'hospital_beds_per_thousand',
       'life_expectancy', 'human_development_index',
       'excess_mortality_cumulative_absolute', 'excess_mortality_cumulative',
       'excess_mortality', 'excess_mortality_cumulative_per_million'],
      dtype='object')
```

[70]:
```
covid_df.shape
```

[70]:
```
(960, 67)
```

[71]:
```
 !pip install jovian --upgrade --quiet
```

[72]:
```
import jovian
```

[73]:
```
jovian.commit(project='python-pandas-covid-india-data-analysis-')
```

```
<IPython.core.display.Javascript object>
```

```
[jovian] Updating notebook "shekharankur4/python-pandas-covid-india-data-
analysis" on https://jovian.ai
[jovian] Committed successfully! https://jovian.ai/shekharankur4/python-pandas-
covid-india-data-analysis
```

[73]:
```
'https://jovian.ai/shekharankur4/python-pandas-covid-india-data-analysis'
```

[74]:
```
covid_df['new_cases']
```

[74]:
```
0         1.0
1         0.0
2         0.0
3         1.0
4         1.0
        ...
955    5221.0
956    1783.0
957    7694.0
958    6422.0
959    6298.0
Name: new_cases, Length: 960, dtype: float64
```

[75]:
```
type(covid_df['new_cases'])
```

[75]:
```
pandas.core.series.Series
```

[76]:
```
covid_df['new_cases'][246]
```

[76]:
```
79476.0
```

[77]:
```
covid_df.at[246, 'new_cases']
```

[77]:
```
79476.0
```

[78]:
```
covid_df.at[240, 'new_deaths']
```

[78]:
```
1124.0
```

[79]:
```
covid_df.new_cases
```

[79]:
```
0         1.0
1         0.0
2         0.0
3         1.0
4         1.0
        ...
955    5221.0
956    1783.0
957    7694.0
958    6422.0
959    6298.0
Name: new_cases, Length: 960, dtype: float64
```

[80]:
```
cases_df = covid_df[['date', 'new_cases']]
cases_df
```

[80]:
```
           date  new_cases
0    2020-01-30        1.0
1    2020-01-31        0.0
2    2020-02-01        0.0
3    2020-02-02        1.0
4    2020-02-03        1.0
..          …         …
955  2022-09-11     5221.0
956  2022-09-12     1783.0
957  2022-09-13     7694.0
958  2022-09-14     6422.0
959  2022-09-15     6298.0

[960 rows x 2 columns]
```

[81]:
```
covid_df_copy = covid_df.copy()
```

[82]:
```
covid_df.loc[243]
```

[82]:
```
iso_code                                        IND
continent                                      Asia
location                                      India
date                                     2020-09-29
total_cases                               6225763.0
                                                …
human_development_index                       0.645
excess_mortality_cumulative_absolute            NaN
excess_mortality_cumulative                     NaN
excess_mortality                                NaN
excess_mortality_cumulative_per_million         NaN
Name: 243, Length: 67, dtype: object
```

[83]:
```
type(covid_df.loc[243])
```

[83]:
```
pandas.core.series.Series
```

[84]:
```
covid_df.head(5)
```

[84]:
```
  iso_code continent location        date  total_cases  new_cases  \
0      IND      Asia    India  2020-01-30          1.0        1.0
1      IND      Asia    India  2020-01-31          1.0        0.0
2      IND      Asia    India  2020-02-01          1.0        0.0
3      IND      Asia    India  2020-02-02          2.0        1.0
4      IND      Asia    India  2020-02-03          3.0        1.0
```

```
   new_cases_smoothed  total_deaths  new_deaths  new_deaths_smoothed  …  \
0                 NaN           NaN         NaN                  NaN  …
1                 NaN           NaN         NaN                  NaN  …
2                 NaN           NaN         NaN                  NaN  …
3                 NaN           NaN         NaN                  NaN  …
4                 NaN           NaN         NaN                  NaN  …

   female_smokers  male_smokers  handwashing_facilities  \
0             1.9          20.6                   59.55
1             1.9          20.6                   59.55
2             1.9          20.6                   59.55
3             1.9          20.6                   59.55
4             1.9          20.6                   59.55

   hospital_beds_per_thousand  life_expectancy  human_development_index  \
0                        0.53            69.66                    0.645
1                        0.53            69.66                    0.645
2                        0.53            69.66                    0.645
3                        0.53            69.66                    0.645
4                        0.53            69.66                    0.645

   excess_mortality_cumulative_absolute  excess_mortality_cumulative  \
0                                   NaN                          NaN
1                                   NaN                          NaN
2                                   NaN                          NaN
3                                   NaN                          NaN
4                                   NaN                          NaN

   excess_mortality  excess_mortality_cumulative_per_million
0               NaN                                     NaN
1               NaN                                     NaN
2               NaN                                     NaN
3               NaN                                     NaN
4               NaN                                     NaN

[5 rows x 67 columns]
```

[85]:
```
covid_df.tail(4)
```

[85]:
```
    iso_code continent location        date  total_cases  new_cases  \
956      IND      Asia    India  2022-09-12   44502363.0     1783.0
957      IND      Asia    India  2022-09-13   44510057.0     7694.0
958      IND      Asia    India  2022-09-14   44516479.0     6422.0
959      IND      Asia    India  2022-09-15   44522777.0     6298.0

     new_cases_smoothed  total_deaths  new_deaths  new_deaths_smoothed  …  \
```

```
956           5071.571      528185.0        20.0               22.143  …
957           5770.857      528216.0        31.0               22.714  …
958           5406.143      528250.0        34.0               22.857  …
959           5435.429      528273.0        23.0               21.714  …

     female_smokers  male_smokers  handwashing_facilities  \
956             1.9          20.6                   59.55
957             1.9          20.6                   59.55
958             1.9          20.6                   59.55
959             1.9          20.6                   59.55

     hospital_beds_per_thousand  life_expectancy  human_development_index  \
956                        0.53            69.66                    0.645
957                        0.53            69.66                    0.645
958                        0.53            69.66                    0.645
959                        0.53            69.66                    0.645

     excess_mortality_cumulative_absolute  excess_mortality_cumulative  \
956                                   NaN                          NaN
957                                   NaN                          NaN
958                                   NaN                          NaN
959                                   NaN                          NaN

     excess_mortality  excess_mortality_cumulative_per_million
956               NaN                                     NaN
957               NaN                                     NaN
958               NaN                                     NaN
959               NaN                                     NaN

[4 rows x 67 columns]
```

[86]:
```
covid_df.at[0, 'new_deaths']
```

[86]:
```
nan
```

[87]:
```
type(covid_df.at[0, 'new_deaths'])
```

[87]:
```
numpy.float64
```

[88]:
```
covid_df.new_cases_smoothed.first_valid_index()
```

[88]:
```
5
```

[89]:
```
covid_df.loc[108:113]
```

[89]:
```
    iso_code continent location        date  total_cases  new_cases  \
108      IND      Asia    India  2020-05-17      95698.0     5050.0
```

```
109      IND      Asia    India  2020-05-18     100328.0     4630.0
110      IND      Asia    India  2020-05-19     106475.0     6147.0
111      IND      Asia    India  2020-05-20     112028.0     5553.0
112      IND      Asia    India  2020-05-21     118226.0     6198.0
113      IND      Asia    India  2020-05-22     124794.0     6568.0

     new_cases_smoothed  total_deaths  new_deaths  new_deaths_smoothed  …  \
108            4076.714        3025.0       154.0              116.143  …
109            4222.857        3156.0       131.0              123.143  …
110            4597.571        3302.0       146.0              126.714  …
111            4853.286        3434.0       132.0              126.143  …
112            5175.571        3584.0       150.0              133.571  …
113            5572.857        3726.0       142.0              139.000  …

     female_smokers  male_smokers  handwashing_facilities  \
108             1.9          20.6                   59.55
109             1.9          20.6                   59.55
110             1.9          20.6                   59.55
111             1.9          20.6                   59.55
112             1.9          20.6                   59.55
113             1.9          20.6                   59.55

     hospital_beds_per_thousand  life_expectancy  human_development_index  \
108                        0.53            69.66                    0.645
109                        0.53            69.66                    0.645
110                        0.53            69.66                    0.645
111                        0.53            69.66                    0.645
112                        0.53            69.66                    0.645
113                        0.53            69.66                    0.645

     excess_mortality_cumulative_absolute  excess_mortality_cumulative  \
108                                   NaN                          NaN
109                                   NaN                          NaN
110                                   NaN                          NaN
111                                   NaN                          NaN
112                                   NaN                          NaN
113                                   NaN                          NaN

     excess_mortality  excess_mortality_cumulative_per_million
108               NaN                                     NaN
109               NaN                                     NaN
110               NaN                                     NaN
111               NaN                                     NaN
112               NaN                                     NaN
113               NaN                                     NaN

[6 rows x 67 columns]
```

[90]:
```
covid_df.sample(10)
```

[90]:

| | iso_code | continent | location | date | total_cases | new_cases | \ |
|---|---|---|---|---|---|---|---|
| 175 | IND | Asia | India | 2020-07-23 | 1288108.0 | 49310.0 | |
| 541 | IND | Asia | India | 2021-07-24 | 31371901.0 | 39742.0 | |
| 649 | IND | Asia | India | 2021-11-09 | 34388579.0 | 11466.0 | |
| 434 | IND | Asia | India | 2021-04-08 | 13060542.0 | 131968.0 | |
| 453 | IND | Asia | India | 2021-04-27 | 17997113.0 | 360927.0 | |
| 284 | IND | Asia | India | 2020-11-09 | 8591730.0 | 38073.0 | |
| 777 | IND | Asia | India | 2022-03-17 | 43004005.0 | 2528.0 | |
| 869 | IND | Asia | India | 2022-06-17 | 43283793.0 | 13216.0 | |
| 354 | IND | Asia | India | 2021-01-18 | 10581823.0 | 10050.0 | |
| 194 | IND | Asia | India | 2020-08-11 | 2329638.0 | 60963.0 | |

| | new_cases_smoothed | total_deaths | new_deaths | new_deaths_smoothed | … | \ |
|---|---|---|---|---|---|---|
| 175 | 40610.857 | 30601.0 | 740.0 | 714.143 | … | |
| 541 | 37976.571 | 420551.0 | 535.0 | 991.714 | … | |
| 649 | 11491.286 | 461849.0 | 460.0 | 379.714 | … | |
| 434 | 108201.571 | 167642.0 | 780.0 | 606.571 | … | |
| 453 | 340140.429 | 201187.0 | 3293.0 | 2662.000 | … | |
| 284 | 46301.000 | 127059.0 | 448.0 | 566.000 | … | |
| 777 | 2820.571 | 516281.0 | 149.0 | 81.000 | … | |
| 869 | 10051.143 | 524840.0 | 23.0 | 11.857 | … | |
| 354 | 14663.429 | 152556.0 | 137.0 | 175.571 | … | |
| 194 | 60197.714 | 46091.0 | 834.0 | 899.429 | … | |

| | female_smokers | male_smokers | handwashing_facilities | \ |
|---|---|---|---|---|
| 175 | 1.9 | 20.6 | 59.55 | |
| 541 | 1.9 | 20.6 | 59.55 | |
| 649 | 1.9 | 20.6 | 59.55 | |
| 434 | 1.9 | 20.6 | 59.55 | |
| 453 | 1.9 | 20.6 | 59.55 | |
| 284 | 1.9 | 20.6 | 59.55 | |
| 777 | 1.9 | 20.6 | 59.55 | |
| 869 | 1.9 | 20.6 | 59.55 | |
| 354 | 1.9 | 20.6 | 59.55 | |
| 194 | 1.9 | 20.6 | 59.55 | |

| | hospital_beds_per_thousand | life_expectancy | human_development_index | \ |
|---|---|---|---|---|
| 175 | 0.53 | 69.66 | 0.645 | |
| 541 | 0.53 | 69.66 | 0.645 | |
| 649 | 0.53 | 69.66 | 0.645 | |
| 434 | 0.53 | 69.66 | 0.645 | |
| 453 | 0.53 | 69.66 | 0.645 | |
| 284 | 0.53 | 69.66 | 0.645 | |
| 777 | 0.53 | 69.66 | 0.645 | |
| 869 | 0.53 | 69.66 | 0.645 | |

```
354                        0.53            69.66                     0.645
194                        0.53            69.66                     0.645

     excess_mortality_cumulative_absolute  excess_mortality_cumulative  \
175                                   NaN                          NaN
541                                   NaN                          NaN
649                                   NaN                          NaN
434                                   NaN                          NaN
453                                   NaN                          NaN
284                                   NaN                          NaN
777                                   NaN                          NaN
869                                   NaN                          NaN
354                                   NaN                          NaN
194                                   NaN                          NaN

     excess_mortality  excess_mortality_cumulative_per_million
175               NaN                                     NaN
541               NaN                                     NaN
649               NaN                                     NaN
434               NaN                                     NaN
453               NaN                                     NaN
284               NaN                                     NaN
777               NaN                                     NaN
869               NaN                                     NaN
354               NaN                                     NaN
194               NaN                                     NaN

[10 rows x 67 columns]
```

[91]:
```
import jovian
```

[92]:
```
jovian.commit()
```

```
<IPython.core.display.Javascript object>
```

```
[jovian] Updating notebook "shekharankur4/python-pandas-covid-india-data-
analysis" on https://jovian.ai
[jovian] Committed successfully! https://jovian.ai/shekharankur4/python-pandas-
covid-india-data-analysis
```

[92]:
```
'https://jovian.ai/shekharankur4/python-pandas-covid-india-data-analysis'
```

### 0.1 Analyzing data from data frames

Q: What are the total number of reported cases and deaths related to Covid-19 in India?

[93]:
```
total_cases = covid_df.new_cases.sum()
total_deaths = covid_df.new_deaths.sum()
```

[94]:
```
print('The number of reported cases is {} and the number of reported deaths is␣
↪{}.'.format(int(total_cases), int(total_deaths)))
```

```
The number of reported cases is 44522777 and the number of reported deaths is
520900.
```

Q: What is the overall death rate (ratio of reported deaths to reported cases)?

[95]:
```
death_rate = covid_df.new_deaths.sum() / covid_df.new_cases.sum()
```

[96]:
```
print("The overall reported death rate in Italy is {:.2f} %.".
↪format(death_rate*100))
```

```
The overall reported death rate in Italy is 1.17 %.
```

[97]:
```
import jovian
```

[98]:
```
jovian.commit()
```

```
<IPython.core.display.Javascript object>
```

```
[jovian] Updating notebook "shekharankur4/python-pandas-covid-india-data-
analysis" on https://jovian.ai
[jovian] Committed successfully! https://jovian.ai/shekharankur4/python-pandas-
covid-india-data-analysis
```

[98]:
```
'https://jovian.ai/shekharankur4/python-pandas-covid-india-data-analysis'
```

## 0.2 Querying and sorting rows

[99]:
```
high_new_cases = covid_df.new_cases > 1000
```

[100]:
```
high_new_cases
```

[100]:
```
0      False
1      False
2      False
3      False
4      False
       …
955     True
956     True
957     True
958     True
959     True
Name: new_cases, Length: 960, dtype: bool
```

[101]:
```
covid_df[high_new_cases]
```

[101]:

| | iso_code | continent | location | date | total_cases | new_cases | \ |
|---|---|---|---|---|---|---|---|
| 67 | IND | Asia | India | 2020-04-06 | 4778.0 | 1190.0 | |
| 74 | IND | Asia | India | 2020-04-13 | 10453.0 | 1248.0 | |
| 75 | IND | Asia | India | 2020-04-14 | 11487.0 | 1034.0 | |
| 77 | IND | Asia | India | 2020-04-16 | 13430.0 | 1108.0 | |
| 79 | IND | Asia | India | 2020-04-18 | 15722.0 | 1370.0 | |
| .. | … | … | … | … | … | … | |
| 955 | IND | Asia | India | 2022-09-11 | 44500580.0 | 5221.0 | |
| 956 | IND | Asia | India | 2022-09-12 | 44502363.0 | 1783.0 | |
| 957 | IND | Asia | India | 2022-09-13 | 44510057.0 | 7694.0 | |
| 958 | IND | Asia | India | 2022-09-14 | 44516479.0 | 6422.0 | |
| 959 | IND | Asia | India | 2022-09-15 | 44522777.0 | 6298.0 | |

| | new_cases_smoothed | total_deaths | new_deaths | new_deaths_smoothed | … | \ |
|---|---|---|---|---|---|---|
| 67 | 503.857 | 136.0 | 37.0 | 14.857 | … | |
| 74 | 810.714 | 358.0 | 27.0 | 31.714 | … | |
| 75 | 882.286 | 393.0 | 35.0 | 34.714 | … | |
| 77 | 957.857 | 448.0 | 43.0 | 31.714 | … | |
| 79 | 1039.429 | 521.0 | 35.0 | 33.286 | … | |
| .. | … | … | … | … | … | |
| 955 | 5447.857 | 528165.0 | 15.0 | 22.571 | … | |
| 956 | 5071.571 | 528185.0 | 20.0 | 22.143 | … | |
| 957 | 5770.857 | 528216.0 | 31.0 | 22.714 | … | |
| 958 | 5406.143 | 528250.0 | 34.0 | 22.857 | … | |
| 959 | 5435.429 | 528273.0 | 23.0 | 21.714 | … | |

| | female_smokers | male_smokers | handwashing_facilities | \ |
|---|---|---|---|---|
| 67 | 1.9 | 20.6 | 59.55 | |
| 74 | 1.9 | 20.6 | 59.55 | |
| 75 | 1.9 | 20.6 | 59.55 | |
| 77 | 1.9 | 20.6 | 59.55 | |
| 79 | 1.9 | 20.6 | 59.55 | |
| .. | … | … | … | |
| 955 | 1.9 | 20.6 | 59.55 | |
| 956 | 1.9 | 20.6 | 59.55 | |
| 957 | 1.9 | 20.6 | 59.55 | |
| 958 | 1.9 | 20.6 | 59.55 | |
| 959 | 1.9 | 20.6 | 59.55 | |

| | hospital_beds_per_thousand | life_expectancy | human_development_index | \ |
|---|---|---|---|---|
| 67 | 0.53 | 69.66 | 0.645 | |
| 74 | 0.53 | 69.66 | 0.645 | |
| 75 | 0.53 | 69.66 | 0.645 | |
| 77 | 0.53 | 69.66 | 0.645 | |
| 79 | 0.53 | 69.66 | 0.645 | |
| .. | … | … | … | |
| 955 | 0.53 | 69.66 | 0.645 | |

```
956                         0.53            69.66                     0.645
957                         0.53            69.66                     0.645
958                         0.53            69.66                     0.645
959                         0.53            69.66                     0.645

     excess_mortality_cumulative_absolute  excess_mortality_cumulative  \
67                                    NaN                          NaN
74                                    NaN                          NaN
75                                    NaN                          NaN
77                                    NaN                          NaN
79                                    NaN                          NaN
..                                    …                            …
955                                   NaN                          NaN
956                                   NaN                          NaN
957                                   NaN                          NaN
958                                   NaN                          NaN
959                                   NaN                          NaN

     excess_mortality  excess_mortality_cumulative_per_million
67                NaN                                     NaN
74                NaN                                     NaN
75                NaN                                     NaN
77                NaN                                     NaN
79                NaN                                     NaN
..                …                                       …
955               NaN                                     NaN
956               NaN                                     NaN
957               NaN                                     NaN
958               NaN                                     NaN
959               NaN                                     NaN

[871 rows x 67 columns]
```

[102]:
```
high_cases_df = covid_df[covid_df.new_cases > 1000]
```

[103]:
```
high_cases_df
```

[103]:
```
    iso_code continent location        date  total_cases  new_cases  \
67       IND      Asia    India  2020-04-06       4778.0     1190.0
74       IND      Asia    India  2020-04-13      10453.0     1248.0
75       IND      Asia    India  2020-04-14      11487.0     1034.0
77       IND      Asia    India  2020-04-16      13430.0     1108.0
79       IND      Asia    India  2020-04-18      15722.0     1370.0
..       …        …       …          …           …           …
955      IND      Asia    India  2022-09-11   44500580.0     5221.0
956      IND      Asia    India  2022-09-12   44502363.0     1783.0
957      IND      Asia    India  2022-09-13   44510057.0     7694.0
```

```
958      IND      Asia    India  2022-09-14   44516479.0     6422.0
959      IND      Asia    India  2022-09-15   44522777.0     6298.0

     new_cases_smoothed  total_deaths  new_deaths  new_deaths_smoothed  …  \
67               503.857         136.0        37.0               14.857  …
74               810.714         358.0        27.0               31.714  …
75               882.286         393.0        35.0               34.714  …
77               957.857         448.0        43.0               31.714  …
79              1039.429         521.0        35.0               33.286  …
..                   …             …           …                    …  …
955             5447.857      528165.0        15.0               22.571  …
956             5071.571      528185.0        20.0               22.143  …
957             5770.857      528216.0        31.0               22.714  …
958             5406.143      528250.0        34.0               22.857  …
959             5435.429      528273.0        23.0               21.714  …

     female_smokers  male_smokers  handwashing_facilities  \
67              1.9          20.6                   59.55
74              1.9          20.6                   59.55
75              1.9          20.6                   59.55
77              1.9          20.6                   59.55
79              1.9          20.6                   59.55
..              …            …                      …
955             1.9          20.6                   59.55
956             1.9          20.6                   59.55
957             1.9          20.6                   59.55
958             1.9          20.6                   59.55
959             1.9          20.6                   59.55

     hospital_beds_per_thousand  life_expectancy  human_development_index  \
67                         0.53            69.66                    0.645
74                         0.53            69.66                    0.645
75                         0.53            69.66                    0.645
77                         0.53            69.66                    0.645
79                         0.53            69.66                    0.645
..                          …              …                        …
955                        0.53            69.66                    0.645
956                        0.53            69.66                    0.645
957                        0.53            69.66                    0.645
958                        0.53            69.66                    0.645
959                        0.53            69.66                    0.645

     excess_mortality_cumulative_absolute  excess_mortality_cumulative  \
67                                    NaN                          NaN
74                                    NaN                          NaN
75                                    NaN                          NaN
77                                    NaN                          NaN
```

```
79                                    NaN                                  NaN
..                                    …                                    …
955                                   NaN                                  NaN
956                                   NaN                                  NaN
957                                   NaN                                  NaN
958                                   NaN                                  NaN
959                                   NaN                                  NaN

     excess_mortality  excess_mortality_cumulative_per_million
67                NaN                                      NaN
74                NaN                                      NaN
75                NaN                                      NaN
77                NaN                                      NaN
79                NaN                                      NaN
..                …                                        …
955               NaN                                      NaN
956               NaN                                      NaN
957               NaN                                      NaN
958               NaN                                      NaN
959               NaN                                      NaN

[871 rows x 67 columns]
```

[104]:
```
from IPython.display import display
with pd.option_context('display.max_rows', 100):
    display(covid_df[covid_df.new_cases > 1000])
```

```
    iso_code continent location        date  total_cases  new_cases  \
67       IND      Asia    India  2020-04-06       4778.0     1190.0
74       IND      Asia    India  2020-04-13      10453.0     1248.0
75       IND      Asia    India  2020-04-14      11487.0     1034.0
77       IND      Asia    India  2020-04-16      13430.0     1108.0
79       IND      Asia    India  2020-04-18      15722.0     1370.0
..       …        …       …          …           …          …
955      IND      Asia    India  2022-09-11   44500580.0     5221.0
956      IND      Asia    India  2022-09-12   44502363.0     1783.0
957      IND      Asia    India  2022-09-13   44510057.0     7694.0
958      IND      Asia    India  2022-09-14   44516479.0     6422.0
959      IND      Asia    India  2022-09-15   44522777.0     6298.0

     new_cases_smoothed  total_deaths  new_deaths  new_deaths_smoothed  …  \
67              503.857         136.0        37.0               14.857  …
74              810.714         358.0        27.0               31.714  …
75              882.286         393.0        35.0               34.714  …
77              957.857         448.0        43.0               31.714  …
79             1039.429         521.0        35.0               33.286  …
..                  …           …          …                  …  …
```

```
955           5447.857      528165.0        15.0              22.571  …
956           5071.571      528185.0        20.0              22.143  …
957           5770.857      528216.0        31.0              22.714  …
958           5406.143      528250.0        34.0              22.857  …
959           5435.429      528273.0        23.0              21.714  …

     female_smokers  male_smokers  handwashing_facilities  \
67              1.9          20.6                   59.55
74              1.9          20.6                   59.55
75              1.9          20.6                   59.55
77              1.9          20.6                   59.55
79              1.9          20.6                   59.55
..              …            …                      …
955             1.9          20.6                   59.55
956             1.9          20.6                   59.55
957             1.9          20.6                   59.55
958             1.9          20.6                   59.55
959             1.9          20.6                   59.55

     hospital_beds_per_thousand  life_expectancy  human_development_index  \
67                         0.53            69.66                    0.645
74                         0.53            69.66                    0.645
75                         0.53            69.66                    0.645
77                         0.53            69.66                    0.645
79                         0.53            69.66                    0.645
..                          …              …                        …
955                        0.53            69.66                    0.645
956                        0.53            69.66                    0.645
957                        0.53            69.66                    0.645
958                        0.53            69.66                    0.645
959                        0.53            69.66                    0.645

     excess_mortality_cumulative_absolute  excess_mortality_cumulative  \
67                                    NaN                          NaN
74                                    NaN                          NaN
75                                    NaN                          NaN
77                                    NaN                          NaN
79                                    NaN                          NaN
..                                    …                            …
955                                   NaN                          NaN
956                                   NaN                          NaN
957                                   NaN                          NaN
958                                   NaN                          NaN
959                                   NaN                          NaN

     excess_mortality  excess_mortality_cumulative_per_million
67                NaN                                     NaN
74                NaN                                     NaN
```

```
75                NaN                                    NaN
77                NaN                                    NaN
79                NaN                                    NaN
..                …                                    …
955               NaN                                    NaN
956               NaN                                    NaN
957               NaN                                    NaN
958               NaN                                    NaN
959               NaN                                    NaN

[871 rows x 67 columns]
```

### 0.2.1 Sorting rows using column values

[105]:
```
covid_df.sort_values('new_cases', ascending=False).head(10)
```

[105]:

| | iso_code | continent | location | date | total_cases | new_cases | \ |
|---|---|---|---|---|---|---|---|
| 462 | IND | Asia | India | 2021-05-06 | 21491598.0 | 414188.0 | |
| 461 | IND | Asia | India | 2021-05-05 | 21077410.0 | 412431.0 | |
| 464 | IND | Asia | India | 2021-05-08 | 22296081.0 | 403405.0 | |
| 456 | IND | Asia | India | 2021-04-30 | 19164969.0 | 401993.0 | |
| 463 | IND | Asia | India | 2021-05-07 | 21892676.0 | 401078.0 | |
| 457 | IND | Asia | India | 2021-05-01 | 19557457.0 | 392488.0 | |
| 455 | IND | Asia | India | 2021-04-29 | 18762976.0 | 386555.0 | |
| 460 | IND | Asia | India | 2021-05-04 | 20664979.0 | 382146.0 | |
| 454 | IND | Asia | India | 2021-04-28 | 18376421.0 | 379308.0 | |
| 458 | IND | Asia | India | 2021-05-02 | 19925517.0 | 368060.0 | |

| | new_cases_smoothed | total_deaths | new_deaths | new_deaths_smoothed | … | \ |
|---|---|---|---|---|---|---|
| 462 | 389803.143 | 234083.0 | 3915.0 | 3679.000 | … | |
| 461 | 385855.571 | 230168.0 | 3980.0 | 3619.429 | … | |
| 464 | 391232.000 | 242347.0 | 4077.0 | 3829.286 | … | |
| 456 | 364926.857 | 211853.0 | 3523.0 | 3187.000 | … | |
| 463 | 389672.429 | 238270.0 | 4187.0 | 3773.857 | … | |
| 457 | 371040.714 | 215542.0 | 3689.0 | 3318.714 | … | |
| 455 | 357040.143 | 208330.0 | 3498.0 | 3058.571 | … | |
| 460 | 381123.714 | 226188.0 | 3780.0 | 3571.571 | … | |
| 454 | 349378.143 | 204832.0 | 3645.0 | 2882.143 | … | |
| 458 | 373193.429 | 218959.0 | 3417.0 | 3405.143 | … | |

| | female_smokers | male_smokers | handwashing_facilities | \ |
|---|---|---|---|---|
| 462 | 1.9 | 20.6 | 59.55 | |
| 461 | 1.9 | 20.6 | 59.55 | |
| 464 | 1.9 | 20.6 | 59.55 | |
| 456 | 1.9 | 20.6 | 59.55 | |
| 463 | 1.9 | 20.6 | 59.55 | |
| 457 | 1.9 | 20.6 | 59.55 | |

```
455             1.9          20.6                   59.55
460             1.9          20.6                   59.55
454             1.9          20.6                   59.55
458             1.9          20.6                   59.55

     hospital_beds_per_thousand  life_expectancy  human_development_index  \
462                        0.53            69.66                    0.645
461                        0.53            69.66                    0.645
464                        0.53            69.66                    0.645
456                        0.53            69.66                    0.645
463                        0.53            69.66                    0.645
457                        0.53            69.66                    0.645
455                        0.53            69.66                    0.645
460                        0.53            69.66                    0.645
454                        0.53            69.66                    0.645
458                        0.53            69.66                    0.645

     excess_mortality_cumulative_absolute  excess_mortality_cumulative  \
462                                   NaN                          NaN
461                                   NaN                          NaN
464                                   NaN                          NaN
456                                   NaN                          NaN
463                                   NaN                          NaN
457                                   NaN                          NaN
455                                   NaN                          NaN
460                                   NaN                          NaN
454                                   NaN                          NaN
458                                   NaN                          NaN

     excess_mortality  excess_mortality_cumulative_per_million
462               NaN                                     NaN
461               NaN                                     NaN
464               NaN                                     NaN
456               NaN                                     NaN
463               NaN                                     NaN
457               NaN                                     NaN
455               NaN                                     NaN
460               NaN                                     NaN
454               NaN                                     NaN
458               NaN                                     NaN

[10 rows x 67 columns]
```

[106]:
```
covid_df.sort_values('new_deaths', ascending=False).head(10)
```

[106]:
```
    iso_code continent location        date  total_cases  new_cases  \
474      IND      Asia    India  2021-05-18   25496330.0   267334.0
```

```
479      IND      Asia    India  2021-05-23   26752447.0   222315.0
473      IND      Asia    India  2021-05-17   25228996.0   263533.0
476      IND      Asia    India  2021-05-20   26031991.0   259551.0
467      IND      Asia    India  2021-05-11   23340938.0   348421.0
477      IND      Asia    India  2021-05-21   26289290.0   257299.0
463      IND      Asia    India  2021-05-07   21892676.0   401078.0
481      IND      Asia    India  2021-05-25   27157795.0   208921.0
468      IND      Asia    India  2021-05-12   23703665.0   362727.0
472      IND      Asia    India  2021-05-16   24965463.0   281386.0

     new_cases_smoothed  total_deaths  new_deaths  new_deaths_smoothed  …  \
474           307913.143      283248.0      4529.0             4150.143  …
479           255283.429      303720.0      4454.0             4190.000  …
473           319497.000      278719.0      4329.0             4103.857  …
476           283597.429      291331.0      4209.0             4144.857  …
467           382279.857      254197.0      4205.0             4001.286  …
477           273769.000      295525.0      4194.0             4188.286  …
463           389672.429      238270.0      4187.0             3773.857  …
481           237352.143      311388.0      4157.0             4020.000  …
468           375179.286      258317.0      4120.0             4021.286  …
472           328984.000      274390.0      4106.0             4039.143  …

     female_smokers  male_smokers  handwashing_facilities  \
474             1.9          20.6                   59.55
479             1.9          20.6                   59.55
473             1.9          20.6                   59.55
476             1.9          20.6                   59.55
467             1.9          20.6                   59.55
477             1.9          20.6                   59.55
463             1.9          20.6                   59.55
481             1.9          20.6                   59.55
468             1.9          20.6                   59.55
472             1.9          20.6                   59.55

     hospital_beds_per_thousand  life_expectancy  human_development_index  \
474                        0.53            69.66                    0.645
479                        0.53            69.66                    0.645
473                        0.53            69.66                    0.645
476                        0.53            69.66                    0.645
467                        0.53            69.66                    0.645
477                        0.53            69.66                    0.645
463                        0.53            69.66                    0.645
481                        0.53            69.66                    0.645
468                        0.53            69.66                    0.645
472                        0.53            69.66                    0.645

     excess_mortality_cumulative_absolute  excess_mortality_cumulative  \
```

```
474                                  NaN                                  NaN
479                                  NaN                                  NaN
473                                  NaN                                  NaN
476                                  NaN                                  NaN
467                                  NaN                                  NaN
477                                  NaN                                  NaN
463                                  NaN                                  NaN
481                                  NaN                                  NaN
468                                  NaN                                  NaN
472                                  NaN                                  NaN

     excess_mortality  excess_mortality_cumulative_per_million
474               NaN                                     NaN
479               NaN                                     NaN
473               NaN                                     NaN
476               NaN                                     NaN
467               NaN                                     NaN
477               NaN                                     NaN
463               NaN                                     NaN
481               NaN                                     NaN
468               NaN                                     NaN
472               NaN                                     NaN

[10 rows x 67 columns]
```

[107]:
```
covid_df.sort_values('new_cases').head(10)
```

[107]:
```
    iso_code continent location        date  total_cases  new_cases  \
344      IND      Asia    India  2021-01-08   10413417.0        0.0
20       IND      Asia    India  2020-02-19          3.0        0.0
21       IND      Asia    India  2020-02-20          3.0        0.0
22       IND      Asia    India  2020-02-21          3.0        0.0
23       IND      Asia    India  2020-02-22          3.0        0.0
24       IND      Asia    India  2020-02-23          3.0        0.0
25       IND      Asia    India  2020-02-24          3.0        0.0
19       IND      Asia    India  2020-02-18          3.0        0.0
26       IND      Asia    India  2020-02-25          3.0        0.0
28       IND      Asia    India  2020-02-27          3.0        0.0

     new_cases_smoothed  total_deaths  new_deaths  new_deaths_smoothed  …  \
344           15375.571      150570.0         0.0              193.143  …
20                0.000           NaN         NaN                  NaN  …
21                0.000           NaN         NaN                  NaN  …
22                0.000           NaN         NaN                  NaN  …
23                0.000           NaN         NaN                  NaN  …
24                0.000           NaN         NaN                  NaN  …
25                0.000           NaN         NaN                  NaN  …
```

```
19                0.000           NaN         NaN                   NaN  …
26                0.000           NaN         NaN                   NaN  …
28                0.000           NaN         NaN                   NaN  …

     female_smokers  male_smokers  handwashing_facilities  \
344             1.9          20.6                   59.55
20              1.9          20.6                   59.55
21              1.9          20.6                   59.55
22              1.9          20.6                   59.55
23              1.9          20.6                   59.55
24              1.9          20.6                   59.55
25              1.9          20.6                   59.55
19              1.9          20.6                   59.55
26              1.9          20.6                   59.55
28              1.9          20.6                   59.55

     hospital_beds_per_thousand  life_expectancy  human_development_index  \
344                        0.53            69.66                    0.645
20                         0.53            69.66                    0.645
21                         0.53            69.66                    0.645
22                         0.53            69.66                    0.645
23                         0.53            69.66                    0.645
24                         0.53            69.66                    0.645
25                         0.53            69.66                    0.645
19                         0.53            69.66                    0.645
26                         0.53            69.66                    0.645
28                         0.53            69.66                    0.645

     excess_mortality_cumulative_absolute  excess_mortality_cumulative  \
344                                   NaN                          NaN
20                                    NaN                          NaN
21                                    NaN                          NaN
22                                    NaN                          NaN
23                                    NaN                          NaN
24                                    NaN                          NaN
25                                    NaN                          NaN
19                                    NaN                          NaN
26                                    NaN                          NaN
28                                    NaN                          NaN

     excess_mortality  excess_mortality_cumulative_per_million
344               NaN                                     NaN
20                NaN                                     NaN
21                NaN                                     NaN
22                NaN                                     NaN
23                NaN                                     NaN
24                NaN                                     NaN
```

```
25                NaN                                        NaN
19                NaN                                        NaN
26                NaN                                        NaN
28                NaN                                        NaN

[10 rows x 67 columns]
```

[108]:
```
covid_df.loc[169:175]
```

[108]:
```
    iso_code continent location        date  total_cases  new_cases  \
169      IND      Asia    India  2020-07-17    1039084.0    35252.0
170      IND      Asia    India  2020-07-18    1077781.0    38697.0
171      IND      Asia    India  2020-07-19    1118206.0    40425.0
172      IND      Asia    India  2020-07-20    1155338.0    37132.0
173      IND      Asia    India  2020-07-21    1193078.0    37740.0
174      IND      Asia    India  2020-07-22    1238798.0    45720.0
175      IND      Asia    India  2020-07-23    1288108.0    49310.0

     new_cases_smoothed  total_deaths  new_deaths  new_deaths_smoothed  …  \
169           31166.857       26273.0       671.0              592.857  …
170           32608.429       26816.0       543.0              591.857  …
171           34278.857       27497.0       681.0              617.571  …
172           35512.286       28082.0       585.0              622.143  …
173           36699.571       28732.0       650.0              631.857  …
174           38563.000       29861.0      1129.0              706.714  …
175           40610.857       30601.0       740.0              714.143  …

     female_smokers  male_smokers  handwashing_facilities  \
169             1.9          20.6                   59.55
170             1.9          20.6                   59.55
171             1.9          20.6                   59.55
172             1.9          20.6                   59.55
173             1.9          20.6                   59.55
174             1.9          20.6                   59.55
175             1.9          20.6                   59.55

     hospital_beds_per_thousand  life_expectancy  human_development_index  \
169                        0.53            69.66                    0.645
170                        0.53            69.66                    0.645
171                        0.53            69.66                    0.645
172                        0.53            69.66                    0.645
173                        0.53            69.66                    0.645
174                        0.53            69.66                    0.645
175                        0.53            69.66                    0.645

     excess_mortality_cumulative_absolute  excess_mortality_cumulative  \
169                                   NaN                          NaN
```

```
170                                    NaN                                    NaN
171                                    NaN                                    NaN
172                                    NaN                                    NaN
173                                    NaN                                    NaN
174                                    NaN                                    NaN
175                                    NaN                                    NaN

     excess_mortality  excess_mortality_cumulative_per_million
169               NaN                                      NaN
170               NaN                                      NaN
171               NaN                                      NaN
172               NaN                                      NaN
173               NaN                                      NaN
174               NaN                                      NaN
175               NaN                                      NaN

[7 rows x 67 columns]
```

[109]:
```
covid_df.at[172, 'new_cases'] = (covid_df.at[171, 'new_cases'] + covid_df.
↪at[173, 'new_cases'])/2
```

[110]:
```
import jovian
```

[111]:
```
jovian.commit()
```

```
<IPython.core.display.Javascript object>

[jovian] Updating notebook "shekharankur4/python-pandas-covid-india-data-
analysis" on https://jovian.ai
[jovian] Committed successfully! https://jovian.ai/shekharankur4/python-pandas-
covid-india-data-analysis
```

[111]:
```
'https://jovian.ai/shekharankur4/python-pandas-covid-india-data-analysis'
```

### 0.3 Working with dates

While we've looked at overall numbers for the cases, tests, positive rate, etc., it would also be useful to study these numbers on a month-by-month basis.

[112]:
```
covid_df.date
```

[112]:
```
0      2020-01-30
1      2020-01-31
2      2020-02-01
3      2020-02-02
4      2020-02-03
          ...
955    2022-09-11
```

```
956    2022-09-12
957    2022-09-13
958    2022-09-14
959    2022-09-15
Name: date, Length: 960, dtype: object
```

[113]:
```
covid_df['date'] = pd.to_datetime(covid_df.date)
```

[114]:
```
covid_df['date']
```

[114]:
```
0     2020-01-30
1     2020-01-31
2     2020-02-01
3     2020-02-02
4     2020-02-03
         …
955   2022-09-11
956   2022-09-12
957   2022-09-13
958   2022-09-14
959   2022-09-15
Name: date, Length: 960, dtype: datetime64[ns]
```

[116]:
```
covid_df['year'] = pd.DatetimeIndex(covid_df.date).year
covid_df['month'] = pd.DatetimeIndex(covid_df.date).month
covid_df['day'] = pd.DatetimeIndex(covid_df.date).day
covid_df['weekday'] = pd.DatetimeIndex(covid_df.date).weekday
```

[117]:
```
covid_df
```

[117]:
```
    iso_code continent location       date  total_cases  new_cases  \
0        IND      Asia    India 2020-01-30          1.0        1.0
1        IND      Asia    India 2020-01-31          1.0        0.0
2        IND      Asia    India 2020-02-01          1.0        0.0
3        IND      Asia    India 2020-02-02          2.0        1.0
4        IND      Asia    India 2020-02-03          3.0        1.0
..       …        …       …         …           …          …
955      IND      Asia    India 2022-09-11   44500580.0     5221.0
956      IND      Asia    India 2022-09-12   44502363.0     1783.0
957      IND      Asia    India 2022-09-13   44510057.0     7694.0
958      IND      Asia    India 2022-09-14   44516479.0     6422.0
959      IND      Asia    India 2022-09-15   44522777.0     6298.0

     new_cases_smoothed  total_deaths  new_deaths  new_deaths_smoothed  …  \
0                   NaN           NaN         NaN                  NaN  …
1                   NaN           NaN         NaN                  NaN  …
2                   NaN           NaN         NaN                  NaN  …
```

```
3                   NaN           NaN         NaN                   NaN  …
4                   NaN           NaN         NaN                   NaN  …
..                  …             …           …                     …    …
955            5447.857      528165.0        15.0                22.571  …
956            5071.571      528185.0        20.0                22.143  …
957            5770.857      528216.0        31.0                22.714  …
958            5406.143      528250.0        34.0                22.857  …
959            5435.429      528273.0        23.0                21.714  …

     life_expectancy  human_development_index  \
0              69.66                    0.645
1              69.66                    0.645
2              69.66                    0.645
3              69.66                    0.645
4              69.66                    0.645
..               …                      …
955            69.66                    0.645
956            69.66                    0.645
957            69.66                    0.645
958            69.66                    0.645
959            69.66                    0.645

     excess_mortality_cumulative_absolute  excess_mortality_cumulative  \
0                                     NaN                          NaN
1                                     NaN                          NaN
2                                     NaN                          NaN
3                                     NaN                          NaN
4                                     NaN                          NaN
..                                    …                            …
955                                   NaN                          NaN
956                                   NaN                          NaN
957                                   NaN                          NaN
958                                   NaN                          NaN
959                                   NaN                          NaN

     excess_mortality  excess_mortality_cumulative_per_million  year  month  \
0                 NaN                                      NaN  2020      1
1                 NaN                                      NaN  2020      1
2                 NaN                                      NaN  2020      2
3                 NaN                                      NaN  2020      2
4                 NaN                                      NaN  2020      2
..                …                                        …    …         …
955               NaN                                      NaN  2022      9
956               NaN                                      NaN  2022      9
957               NaN                                      NaN  2022      9
958               NaN                                      NaN  2022      9
959               NaN                                      NaN  2022      9
```

```
     day  weekday
0     30        3
1     31        4
2      1        5
3      2        6
4      3        0
..   …       …
955   11        6
956   12        0
957   13        1
958   14        2
959   15        3

[960 rows x 71 columns]
```

[119]:
```
# Query the rows for May
covid_df_may = covid_df[covid_df.month == 5]
```

[120]:
```
# Extract the subset of columns to be aggregated
covid_df_may_metrics = covid_df_may[['new_cases', 'new_deaths',␣
↪'new_cases_smoothed']]
```

[121]:
```
# Get the column-wise sum
covid_may_totals = covid_df_may_metrics.sum()
```

[122]:
```
covid_may_totals
```

[122]:
```
new_cases             9247465.000
new_deaths             125089.000
new_cases_smoothed    9894896.997
dtype: float64
```

[124]:
```
type(covid_may_totals)
```

[124]:
```
pandas.core.series.Series
```

[125]:
```
covid_df[covid_df.month == 5][['new_cases', 'new_deaths',␣
↪'new_cases_smoothed']].sum()
```

[125]:
```
new_cases             9247465.000
new_deaths             125089.000
new_cases_smoothed    9894896.997
dtype: float64
```

[126]:
```
# Overall average
covid_df.new_cases.mean()
```

[126]: 46379.924479166664

[127]:
```
# Average for Sundays
covid_df[covid_df.weekday == 6].new_cases.mean()
```

[127]: 45530.99270072993

[128]:
```
import jovian
```

[129]:
```
jovian.commit()
```

```
<IPython.core.display.Javascript object>
```

```
[jovian] Updating notebook "shekharankur4/python-pandas-covid-india-data-
analysis" on https://jovian.ai
[jovian] Committed successfully! https://jovian.ai/shekharankur4/python-pandas-
covid-india-data-analysis
```

[129]: 'https://jovian.ai/shekharankur4/python-pandas-covid-india-data-analysis'

### 0.4 Grouping and aggregation

As a next step, we might want to summarize the day-wise data and create a new dataframe with month-wise data

[130]:
```
covid_month_df = covid_df.groupby('month')[['new_cases', 'new_deaths',␣
↪'new_cases_smoothed']].sum()
```

[131]:
```
covid_month_df
```

[131]:

| month | new_cases | new_deaths | new_cases_smoothed |
|---|---|---|---|
| 1 | 7078822.0 | 20154.0 | 6504004.996 |
| 2 | 1816179.0 | 20546.0 | 2419874.716 |
| 3 | 1205548.0 | 12964.0 | 1089793.861 |
| 4 | 7030183.0 | 52707.0 | 6072267.717 |
| 5 | 9247465.0 | 125089.0 | 9894896.997 |
| 6 | 2939864.0 | 72685.0 | 3201126.432 |
| 7 | 2923688.5 | 45260.0 | 2819285.711 |
| 8 | 3550263.0 | 44961.0 | 3519713.576 |
| 9 | 3663718.0 | 43071.0 | 3706096.140 |
| 10 | 2390605.0 | 33531.0 | 2523189.570 |
| 11 | 1589689.0 | 26320.0 | 1632411.570 |
| 12 | 1088703.0 | 23612.0 | 1122751.568 |

[132]:
```
covid_month_mean_df = covid_df.groupby('month')[['new_cases', 'new_deaths',␣
↪'new_cases_smoothed']].mean()
```

[133]:
```
covid_month_mean_df
```

[133]:

| month | new_cases | new_deaths | new_cases_smoothed |
|---|---|---|---|
| 1 | 110606.593750 | 325.064516 | 104903.306387 |
| 2 | 21366.811765 | 366.892857 | 29510.667268 |
| 3 | 12962.881720 | 158.097561 | 11718.213559 |
| 4 | 78113.144444 | 585.633333 | 67469.641300 |
| 5 | 99435.107527 | 1345.043011 | 106396.741903 |
| 6 | 32665.155556 | 816.685393 | 35568.071467 |
| 7 | 31437.510753 | 486.666667 | 30314.900118 |
| 8 | 38174.870968 | 483.451613 | 37846.382538 |
| 9 | 48849.573333 | 574.280000 | 49414.615200 |
| 10 | 38558.145161 | 540.822581 | 40696.605968 |
| 11 | 26494.816667 | 438.666667 | 27206.859500 |
| 12 | 17559.725806 | 380.838710 | 18108.896258 |

[134]:
```
covid_df['total_cases'] = covid_df.new_cases.cumsum()
```

[135]:
```
covid_df['total_deaths'] = covid_df.new_deaths.cumsum()
```

### 0.5 Merging data from multiple sources

To determine other metrics like test per million, cases per million, etc., we require some more information about the country, viz. its population. Let's download another file `locations.csv` that contains health-related information for many countries, including Italy.

[146]:
```
urlretrieve('https://gist.githubusercontent.com/aakashns/
↪8684589ef4f266116cdce023377fc9c8/raw/
↪99ce3826b2a9d1e6d0bde7e9e559fc8b6e9ac88b/locations.csv',
            'locations.csv')
```

[146]:
```
('locations.csv', <http.client.HTTPMessage at 0x7fb16c3ccee0>)
```

[147]:
```
locations_df = pd.read_csv('locations.csv')
```

[148]:
```
locations_df
```

[148]:

| | location | continent | population | life_expectancy | \ |
|---|---|---|---|---|---|
| 0 | Afghanistan | Asia | 3.892834e+07 | 64.83 | |
| 1 | Albania | Europe | 2.877800e+06 | 78.57 | |
| 2 | Algeria | Africa | 4.385104e+07 | 76.88 | |
| 3 | Andorra | Europe | 7.726500e+04 | 83.73 | |
| 4 | Angola | Africa | 3.286627e+07 | 61.15 | |
| .. | … | … | … | … | |
| 207 | Yemen | Asia | 2.982597e+07 | 66.12 | |
| 208 | Zambia | Africa | 1.838396e+07 | 63.89 | |

```
209       Zimbabwe    Africa  1.486293e+07            61.49
210          World       NaN  7.794799e+09            72.58
211  International       NaN           NaN              NaN

     hospital_beds_per_thousand  gdp_per_capita
0                         0.500        1803.987
1                         2.890       11803.431
2                         1.900       13913.839
3                           NaN             NaN
4                           NaN        5819.495
..                          …              …
207                       0.700        1479.147
208                       2.000        3689.251
209                       1.700        1899.775
210                       2.705       15469.207
211                         NaN             NaN

[212 rows x 6 columns]
```

[149]:
```
locations_df[locations_df.location == "India"]
```

[149]:
```
    location continent    population  life_expectancy  \
90     India      Asia  1.380004e+09            69.66

    hospital_beds_per_thousand  gdp_per_capita
90                        0.53        6426.674
```

[150]:
```
covid_df['location'] = "India"
```

[151]:
```
covid_df
```

[151]:
```
     iso_code continent location       date  total_cases  new_cases  \
0         IND      Asia    India 2020-01-30          1.0        1.0
1         IND      Asia    India 2020-01-31          1.0        0.0
2         IND      Asia    India 2020-02-01          1.0        0.0
3         IND      Asia    India 2020-02-02          2.0        1.0
4         IND      Asia    India 2020-02-03          3.0        1.0
..        …        …       …         …            …          …
955       IND      Asia    India 2022-09-11   44502530.5     5221.0
956       IND      Asia    India 2022-09-12   44504313.5     1783.0
957       IND      Asia    India 2022-09-13   44512007.5     7694.0
958       IND      Asia    India 2022-09-14   44518429.5     6422.0
959       IND      Asia    India 2022-09-15   44524727.5     6298.0

     new_cases_smoothed  total_deaths  new_deaths  new_deaths_smoothed  …  \
0                   NaN           NaN         NaN                  NaN  …
1                   NaN           NaN         NaN                  NaN  …
```

```
2                  NaN           NaN         NaN                  NaN  …
3                  NaN           NaN         NaN                  NaN  …
4                  NaN           NaN         NaN                  NaN  …
..                 …             …           …                    …    …
955            5447.857      520792.0        15.0                22.571  …
956            5071.571      520812.0        20.0                22.143  …
957            5770.857      520843.0        31.0                22.714  …
958            5406.143      520877.0        34.0                22.857  …
959            5435.429      520900.0        23.0                21.714  …

     life_expectancy  human_development_index  \
0              69.66                    0.645
1              69.66                    0.645
2              69.66                    0.645
3              69.66                    0.645
4              69.66                    0.645
..               …                      …
955            69.66                    0.645
956            69.66                    0.645
957            69.66                    0.645
958            69.66                    0.645
959            69.66                    0.645

     excess_mortality_cumulative_absolute  excess_mortality_cumulative  \
0                                     NaN                          NaN
1                                     NaN                          NaN
2                                     NaN                          NaN
3                                     NaN                          NaN
4                                     NaN                          NaN
..                                    …                            …
955                                   NaN                          NaN
956                                   NaN                          NaN
957                                   NaN                          NaN
958                                   NaN                          NaN
959                                   NaN                          NaN

     excess_mortality  excess_mortality_cumulative_per_million  year  month  \
0                 NaN                                      NaN  2020      1
1                 NaN                                      NaN  2020      1
2                 NaN                                      NaN  2020      2
3                 NaN                                      NaN  2020      2
4                 NaN                                      NaN  2020      2
..                …                                        …    …         …
955               NaN                                      NaN  2022      9
956               NaN                                      NaN  2022      9
957               NaN                                      NaN  2022      9
958               NaN                                      NaN  2022      9
```

```
959              NaN                                    NaN  2022      9

     day  weekday
0     30        3
1     31        4
2      1        5
3      2        6
4      3        0
..   …       …
955   11        6
956   12        0
957   13        1
958   14        2
959   15        3

[960 rows x 71 columns]
```

[152]:
```
merged_df = covid_df.merge(locations_df, on="location")
```

[153]:
```
merged_df
```

[153]:
```
    iso_code continent_x location       date  total_cases  new_cases  \
0        IND        Asia    India 2020-01-30          1.0        1.0
1        IND        Asia    India 2020-01-31          1.0        0.0
2        IND        Asia    India 2020-02-01          1.0        0.0
3        IND        Asia    India 2020-02-02          2.0        1.0
4        IND        Asia    India 2020-02-03          3.0        1.0
..       …          …       …          …           …          …
955      IND        Asia    India 2022-09-11   44502530.5     5221.0
956      IND        Asia    India 2022-09-12   44504313.5     1783.0
957      IND        Asia    India 2022-09-13   44512007.5     7694.0
958      IND        Asia    India 2022-09-14   44518429.5     6422.0
959      IND        Asia    India 2022-09-15   44524727.5     6298.0

     new_cases_smoothed  total_deaths  new_deaths  new_deaths_smoothed  …  \
0                   NaN           NaN         NaN                  NaN  …
1                   NaN           NaN         NaN                  NaN  …
2                   NaN           NaN         NaN                  NaN  …
3                   NaN           NaN         NaN                  NaN  …
4                   NaN           NaN         NaN                  NaN  …
..                  …          …          …                  …  …
955            5447.857      520792.0        15.0               22.571  …
956            5071.571      520812.0        20.0               22.143  …
957            5770.857      520843.0        31.0               22.714  …
958            5406.143      520877.0        34.0               22.857  …
959            5435.429      520900.0        23.0               21.714  …
```

```
     excess_mortality_cumulative_per_million  year  month  day  weekday  \
0                                        NaN  2020      1   30        3
1                                        NaN  2020      1   31        4
2                                        NaN  2020      2    1        5
3                                        NaN  2020      2    2        6
4                                        NaN  2020      2    3        0
..                                       …    …    …  …       …
955                                      NaN  2022      9   11        6
956                                      NaN  2022      9   12        0
957                                      NaN  2022      9   13        1
958                                      NaN  2022      9   14        2
959                                      NaN  2022      9   15        3

     continent_y  population_y  life_expectancy_y  \
0           Asia  1.380004e+09              69.66
1           Asia  1.380004e+09              69.66
2           Asia  1.380004e+09              69.66
3           Asia  1.380004e+09              69.66
4           Asia  1.380004e+09              69.66
..           …           …                 …
955         Asia  1.380004e+09              69.66
956         Asia  1.380004e+09              69.66
957         Asia  1.380004e+09              69.66
958         Asia  1.380004e+09              69.66
959         Asia  1.380004e+09              69.66

     hospital_beds_per_thousand_y  gdp_per_capita_y
0                            0.53          6426.674
1                            0.53          6426.674
2                            0.53          6426.674
3                            0.53          6426.674
4                            0.53          6426.674
..                            …               …
955                          0.53          6426.674
956                          0.53          6426.674
957                          0.53          6426.674
958                          0.53          6426.674
959                          0.53          6426.674

[960 rows x 76 columns]
```

[156]:
```
merged_df['cases_per_million'] = merged_df.total_cases * 1e6 / merged_df.
↪population_y
```

[157]:
```
merged_df['deaths_per_million'] = merged_df.total_deaths * 1e6 / merged_df.
↪population_y
```

[158]:
```
merged_df['tests_per_million'] = merged_df.total_tests * 1e6 / merged_df.
↪population_y
```

[159]:
```
merged_df
```

[159]:
```
    iso_code continent_x location       date  total_cases  new_cases  \
0        IND        Asia    India 2020-01-30          1.0        1.0
1        IND        Asia    India 2020-01-31          1.0        0.0
2        IND        Asia    India 2020-02-01          1.0        0.0
3        IND        Asia    India 2020-02-02          2.0        1.0
4        IND        Asia    India 2020-02-03          3.0        1.0
..       …          …       …        …           …          …
955      IND        Asia    India 2022-09-11   44502530.5     5221.0
956      IND        Asia    India 2022-09-12   44504313.5     1783.0
957      IND        Asia    India 2022-09-13   44512007.5     7694.0
958      IND        Asia    India 2022-09-14   44518429.5     6422.0
959      IND        Asia    India 2022-09-15   44524727.5     6298.0

     new_cases_smoothed  total_deaths  new_deaths  new_deaths_smoothed  …  \
0                   NaN           NaN         NaN                  NaN  …
1                   NaN           NaN         NaN                  NaN  …
2                   NaN           NaN         NaN                  NaN  …
3                   NaN           NaN         NaN                  NaN  …
4                   NaN           NaN         NaN                  NaN  …
..                  …             …           …                    …  …
955            5447.857      520792.0        15.0               22.571  …
956            5071.571      520812.0        20.0               22.143  …
957            5770.857      520843.0        31.0               22.714  …
958            5406.143      520877.0        34.0               22.857  …
959            5435.429      520900.0        23.0               21.714  …

     day  weekday  continent_y  population_y  life_expectancy_y  \
0     30        3         Asia  1.380004e+09              69.66
1     31        4         Asia  1.380004e+09              69.66
2      1        5         Asia  1.380004e+09              69.66
3      2        6         Asia  1.380004e+09              69.66
4      3        0         Asia  1.380004e+09              69.66
..   …      …           …           …                  …
955   11        6         Asia  1.380004e+09              69.66
956   12        0         Asia  1.380004e+09              69.66
957   13        1         Asia  1.380004e+09              69.66
958   14        2         Asia  1.380004e+09              69.66
959   15        3         Asia  1.380004e+09              69.66

     hospital_beds_per_thousand_y  gdp_per_capita_y  cases_per_million  \
0                            0.53          6426.674           0.000725
1                            0.53          6426.674           0.000725
```

```
2                          0.53          6426.674          0.000725
3                          0.53          6426.674          0.001449
4                          0.53          6426.674          0.002174
..                          …                …                 …
955                        0.53          6426.674      32248.108038
956                        0.53          6426.674      32249.400063
957                        0.53          6426.674      32254.975407
958                        0.53          6426.674      32259.629016
959                        0.53          6426.674      32264.192769

     deaths_per_million  tests_per_million
0                   NaN                NaN
1                   NaN                NaN
2                   NaN                NaN
3                   NaN                NaN
4                   NaN                NaN
..                  …                  …
955          377.384308                NaN
956          377.398801                NaN
957          377.421264                NaN
958          377.445902                NaN
959          377.462569                NaN

[960 rows x 79 columns]
```

[160]:
```
import jovian
```

[161]:
```
jovian.commit()
```

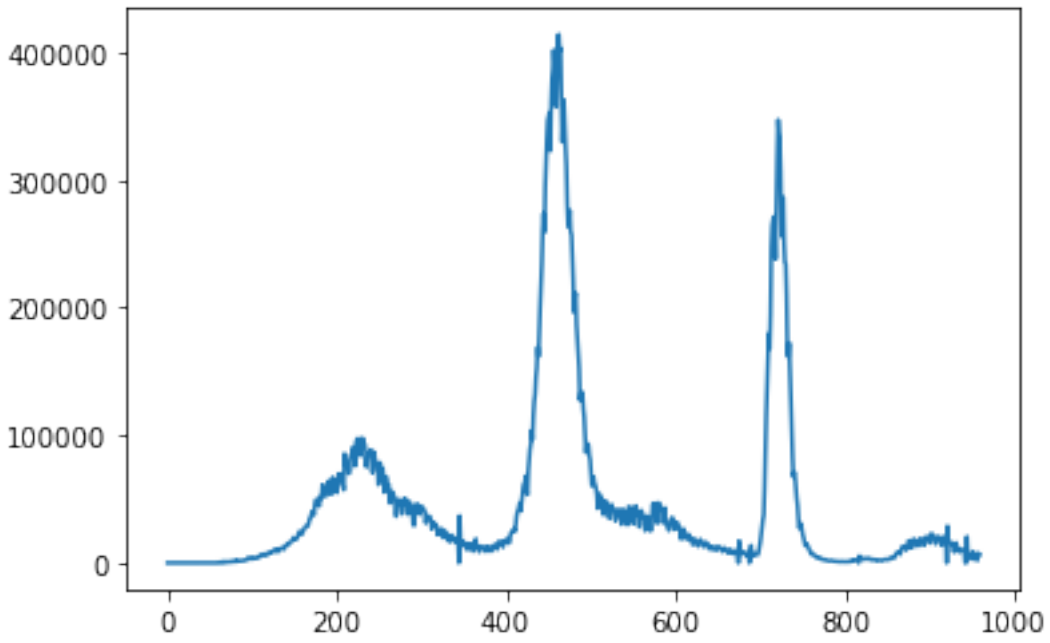
```
<IPython.core.display.Javascript object>
```

```
[jovian] Updating notebook "shekharankur4/python-pandas-covid-india-data-
analysis" on https://jovian.ai
[jovian] Committed successfully! https://jovian.ai/shekharankur4/python-pandas-
covid-india-data-analysis
```

[161]:
```
'https://jovian.ai/shekharankur4/python-pandas-covid-india-data-analysis'
```

### 0.6 Plotting with Pandas

[162]:
```
covid_df.new_cases.plot();
```




[163]:
```
covid_df.set_index('date', inplace=True)
```

[164]:
```
covid_df
```

[164]:

| date | iso_code | continent | location | total_cases | new_cases |
|---|---|---|---|---|---|
| 2020-01-30 | IND | Asia | India | 1.0 | 1.0 |
| 2020-01-31 | IND | Asia | India | 1.0 | 0.0 |
| 2020-02-01 | IND | Asia | India | 1.0 | 0.0 |
| 2020-02-02 | IND | Asia | India | 2.0 | 1.0 |
| 2020-02-03 | IND | Asia | India | 3.0 | 1.0 |
| … | … | … | … | … | … |
| 2022-09-11 | IND | Asia | India | 44502530.5 | 5221.0 |
| 2022-09-12 | IND | Asia | India | 44504313.5 | 1783.0 |
| 2022-09-13 | IND | Asia | India | 44512007.5 | 7694.0 |
| 2022-09-14 | IND | Asia | India | 44518429.5 | 6422.0 |
| 2022-09-15 | IND | Asia | India | 44524727.5 | 6298.0 |

| date | new_cases_smoothed | total_deaths | new_deaths | new_deaths_smoothed |
|---|---|---|---|---|
| 2020-01-30 | NaN | NaN | NaN | NaN |
| 2020-01-31 | NaN | NaN | NaN | NaN |
| 2020-02-01 | NaN | NaN | NaN | NaN |
| 2020-02-02 | NaN | NaN | NaN | NaN |
| 2020-02-03 | NaN | NaN | NaN | NaN |
| … | … | … | … | … |

```
2022-09-11            5447.857      520792.0        15.0               22.571
2022-09-12            5071.571      520812.0        20.0               22.143
2022-09-13            5770.857      520843.0        31.0               22.714
2022-09-14            5406.143      520877.0        34.0               22.857
2022-09-15            5435.429      520900.0        23.0               21.714

            total_cases_per_million  …  life_expectancy  \
date                                 …
2020-01-30                    0.001  …            69.66
2020-01-31                    0.001  …            69.66
2020-02-01                    0.001  …            69.66
2020-02-02                    0.001  …            69.66
2020-02-03                    0.002  …            69.66
…                                  …  …              …
2022-09-11                31615.319  …            69.66
2022-09-12                31616.586  …            69.66
2022-09-13                31622.052  …            69.66
2022-09-14                31626.614  …            69.66
2022-09-15                31631.089  …            69.66

            human_development_index  excess_mortality_cumulative_absolute  \
date
2020-01-30                    0.645                                   NaN
2020-01-31                    0.645                                   NaN
2020-02-01                    0.645                                   NaN
2020-02-02                    0.645                                   NaN
2020-02-03                    0.645                                   NaN
…                                 …                                   …
2022-09-11                    0.645                                   NaN
2022-09-12                    0.645                                   NaN
2022-09-13                    0.645                                   NaN
2022-09-14                    0.645                                   NaN
2022-09-15                    0.645                                   NaN

            excess_mortality_cumulative  excess_mortality  \
date
2020-01-30                          NaN               NaN
2020-01-31                          NaN               NaN
2020-02-01                          NaN               NaN
2020-02-02                          NaN               NaN
2020-02-03                          NaN               NaN
…                                   …                 …
2022-09-11                          NaN               NaN
2022-09-12                          NaN               NaN
2022-09-13                          NaN               NaN
2022-09-14                          NaN               NaN
2022-09-15                          NaN               NaN
```

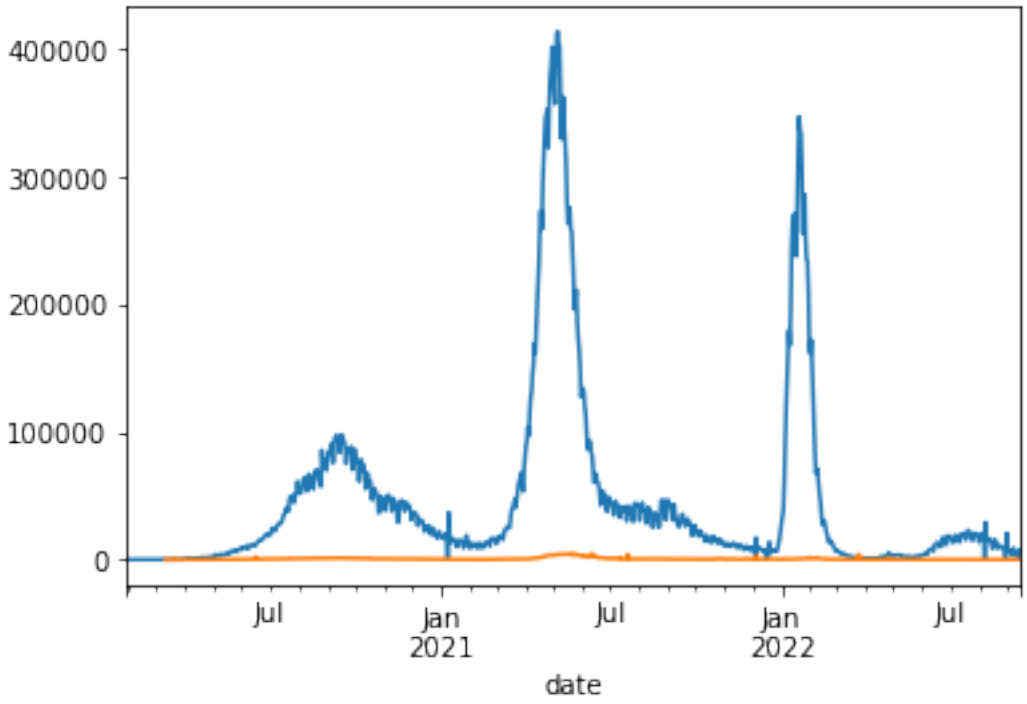
```
            excess_mortality_cumulative_per_million  year  month  day  weekday
date
2020-01-30                                      NaN  2020      1   30        3
2020-01-31                                      NaN  2020      1   31        4
2020-02-01                                      NaN  2020      2    1        5
2020-02-02                                      NaN  2020      2    2        6
2020-02-03                                      NaN  2020      2    3        0
…                                               …    …    …  …       …
2022-09-11                                      NaN  2022      9   11        6
2022-09-12                                      NaN  2022      9   12        0
2022-09-13                                      NaN  2022      9   13        1
2022-09-14                                      NaN  2022      9   14        2
2022-09-15                                      NaN  2022      9   15        3

[960 rows x 70 columns]
```

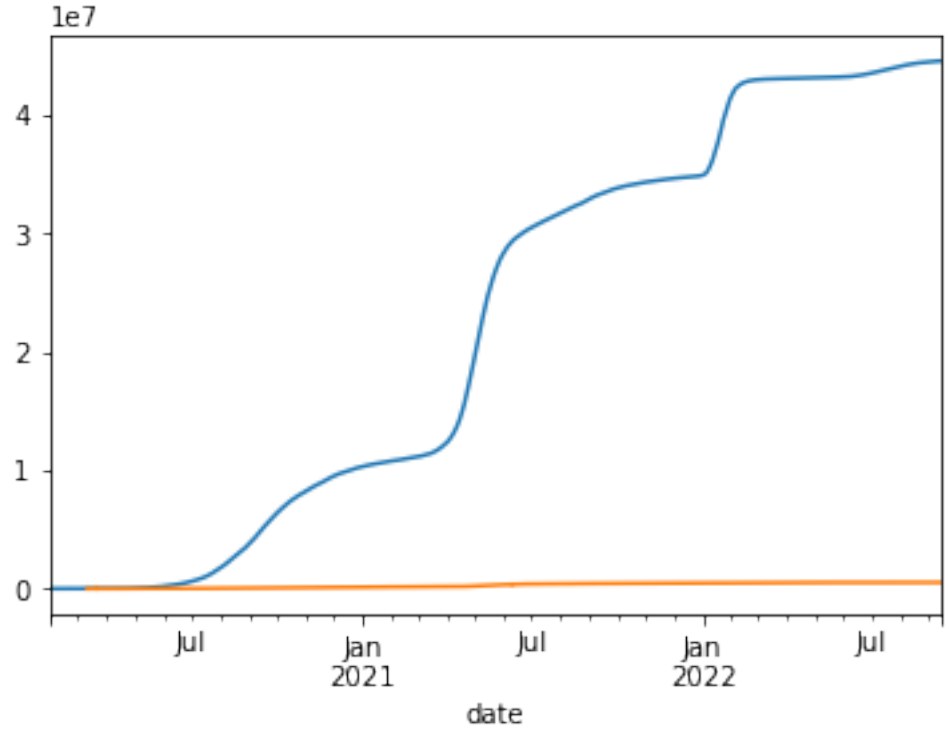
[165]:
```
covid_df.loc['2020-09-01']
```

[165]:
```
iso_code                                    IND
continent                                  Asia
location                                  India
total_cases                          3771473.5
new_cases                              78357.0
                                           …
excess_mortality_cumulative_per_million    NaN
year                                      2020
month                                        9
day                                          1
weekday                                      1
Name: 2020-09-01 00:00:00, Length: 70, dtype: object
```

[166]:
```
covid_df.new_cases.plot()
covid_df.new_deaths.plot();
```

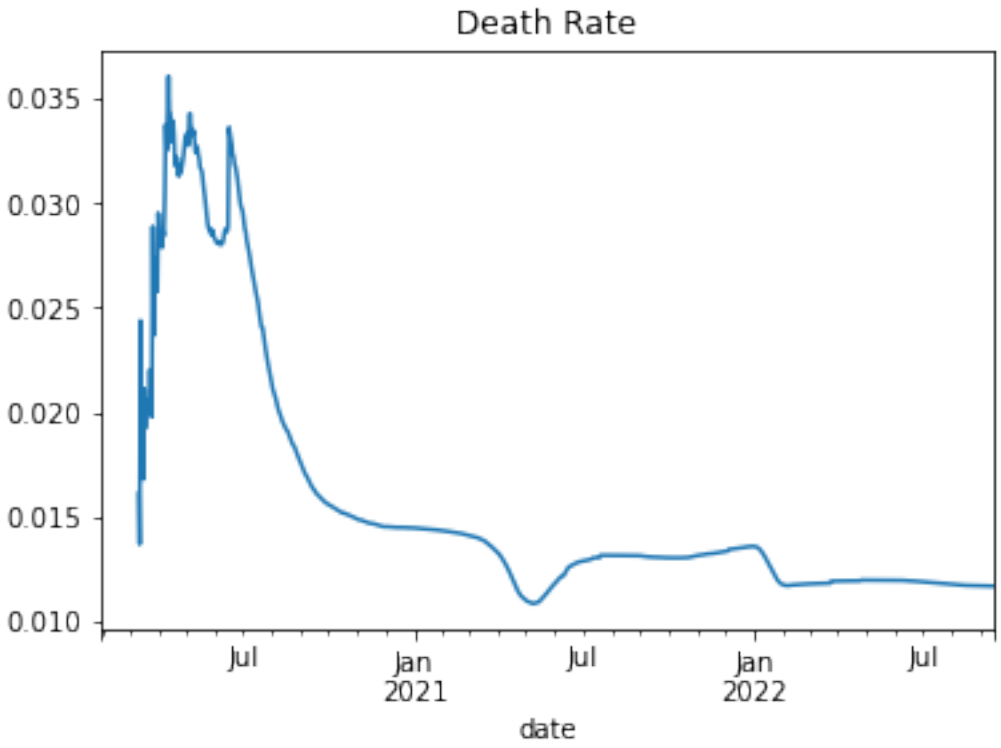



[167]:
```
covid_df.total_cases.plot()
covid_df.total_deaths.plot();
```

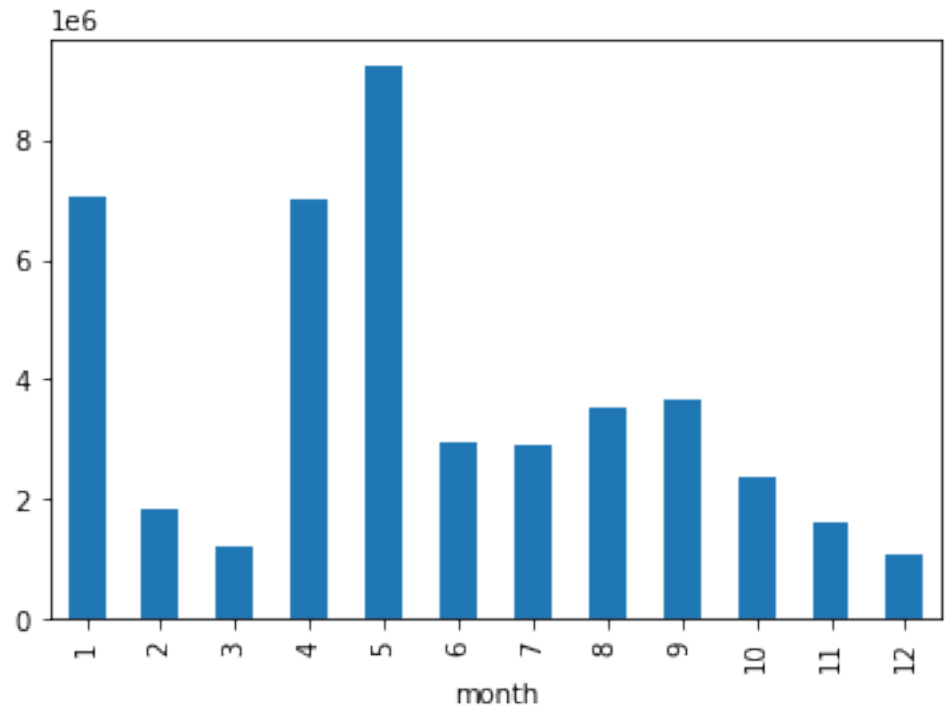


[168]:
```
death_rate = covid_df.total_deaths / covid_df.total_cases
```

[169]:
```
death_rate.plot(title='Death Rate');
```

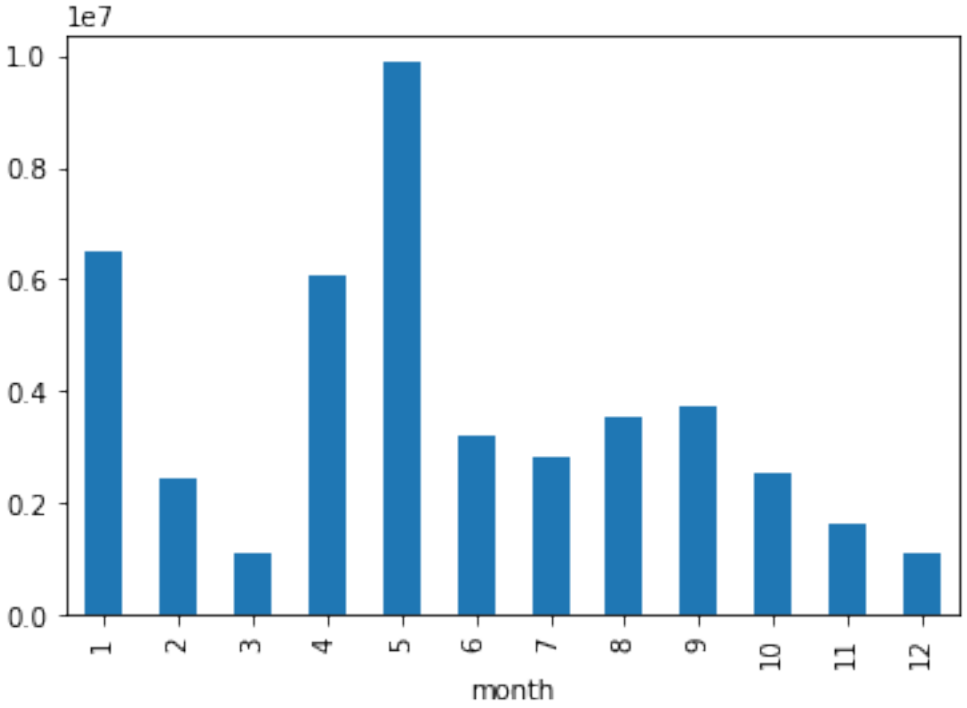



[170]:
```
covid_month_df.new_cases.plot(kind='bar');
```




[172]:
```
covid_month_df.new_cases_smoothed.plot(kind='bar')
```

[172]: <AxesSubplot:xlabel='month'>



[173]:
```
import jovian
```

[174]:
```
jovian.commit()
```

```
<IPython.core.display.Javascript object>
```

```
[jovian] Updating notebook "shekharankur4/python-pandas-covid-india-data-
analysis" on https://jovian.ai
[jovian] Committed successfully! https://jovian.ai/shekharankur4/python-pandas-
covid-india-data-analysis
```

[174]:
```
'https://jovian.ai/shekharankur4/python-pandas-covid-india-data-analysis'
```

[ ]:
```

```