

Covid

October 27, 2020

```
[1]: import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
```

```
[2]: dataframe = pd.read_csv("https://github.com/owid/covid-19-data/raw/master/
    ↳public/data/owid-covid-data.csv")
#importing the data
```

```
[4]: dataframe.info() #viewing the overall data
```

```
<class 'pandas.core.frame.DataFrame'>
```

```
RangeIndex: 52873 entries, 0 to 52872
```

```
Data columns (total 41 columns):
```

#	Column	Non-Null Count	Dtype
0	iso_code	52571 non-null	object
1	continent	52269 non-null	object
2	location	52873 non-null	object
3	date	52873 non-null	object
4	total_cases	49259 non-null	float64
5	new_cases	51979 non-null	float64
6	new_cases_smoothed	51185 non-null	float64
7	total_deaths	40323 non-null	float64
8	new_deaths	51979 non-null	float64
9	new_deaths_smoothed	51185 non-null	float64
10	total_cases_per_million	48993 non-null	float64
11	new_cases_per_million	51915 non-null	float64
12	new_cases_smoothed_per_million	51120 non-null	float64
13	total_deaths_per_million	40072 non-null	float64
14	new_deaths_per_million	51915 non-null	float64
15	new_deaths_smoothed_per_million	51120 non-null	float64
16	total_tests	20007 non-null	float64
17	new_tests	19844 non-null	float64
18	total_tests_per_thousand	20007 non-null	float64
19	new_tests_per_thousand	19844 non-null	float64
20	new_tests_smoothed	22220 non-null	float64
21	new_tests_smoothed_per_thousand	22220 non-null	float64
22	tests_per_case	20536 non-null	float64

```

23 positive_rate                20880 non-null float64
24 tests_units                  23217 non-null object
25 stringency_index             43997 non-null float64
26 population                    52571 non-null float64
27 population_density           50123 non-null float64
28 median_age                   47079 non-null float64
29 aged_65_older                46373 non-null float64
30 aged_70_older                46834 non-null float64
31 gdp_per_capita               46455 non-null float64
32 extreme_poverty              30975 non-null float64
33 cardiovasc_death_rate        47052 non-null float64
34 diabetes_prevalence          48759 non-null float64
35 female_smokers                36694 non-null float64
36 male_smokers                  36223 non-null float64
37 handwashing_facilities       22198 non-null float64
38 hospital_beds_per_thousand   42423 non-null float64
39 life_expectancy              51899 non-null float64
40 human_development_index      45419 non-null float64
dtypes: float64(36), object(5)
memory usage: 16.5+ MB

```

```

[5]: dataframe.handwashing_facilities.unique() # looking at the number of
      ↪handwashing stations

```

```

[5]: array([ nan, 37.746, 26.664, 94.043, 83.241,  6.144, 11.035, 11.877,
          34.808, 97.164, 90.083, 25.383, 88.469, 79.807, 16.603, 19.351,
           2.735,  4.472, 47.964, 65.386, 15.574, 83.841, 85.198, 55.182,
          83.741, 80.635, 89.827,  7.96 , 41.047, 17.45 ,  7.876,  6.403,
          24.64 , 76.665, 77.159, 84.169, 22.863, 64.204, 59.55 , 94.576,
          66.425, 98.999, 24.651, 89.22 , 66.229, 49.839,  1.188, 87.202,
           2.117, 86.979, 50.54 , 95.803, 87.847, 52.232, 79.287, 71.18 ,
          12.227, 15.95 ,  8.704, 44.6  ,  8.978, 41.949, 47.782, 97.4  ,
          59.607, 78.463, 79.602,  4.617, 23.437, 20.859, 35.89 , 19.275,
          90.65 ,  9.831, 97.719, 41.34 , 67.779, 24.097, 70.598,  5.818,
          10.475, 90.67 , 72.704, 28.178, 89.443, 78.687, 47.953, 21.222,
          85.847, 49.542, 43.993, 13.938, 36.791, 60.13 ])

```

```

[6]: dataframe.location.unique() #viewing the countries and world

```

```

[6]: array(['Aruba', 'Afghanistan', 'Angola', 'Anguilla', 'Albania', 'Andorra',
          'United Arab Emirates', 'Argentina', 'Armenia',
          'Antigua and Barbuda', 'Australia', 'Austria', 'Azerbaijan',
          'Burundi', 'Belgium', 'Benin', 'Bonaire Sint Eustatius and Saba',
          'Burkina Faso', 'Bangladesh', 'Bulgaria', 'Bahrain', 'Bahamas',
          'Bosnia and Herzegovina', 'Belarus', 'Belize', 'Bermuda',
          'Bolivia', 'Brazil', 'Barbados', 'Brunei', 'Bhutan', 'Botswana',
          'Central African Republic', 'Canada', 'Switzerland', 'Chile',

```

```
'China', 'Cote d'Ivoire', 'Cameroon',
'Democratic Republic of Congo', 'Congo', 'Colombia', 'Comoros',
'Cape Verde', 'Costa Rica', 'Cuba', 'Curacao', 'Cayman Islands',
'Cyprus', 'Czech Republic', 'Germany', 'Djibouti', 'Dominica',
'Denmark', 'Dominican Republic', 'Algeria', 'Ecuador', 'Egypt',
'Eritrea', 'Western Sahara', 'Spain', 'Estonia', 'Ethiopia',
'Finland', 'Fiji', 'Falkland Islands', 'France', 'Faeroe Islands',
'Gabon', 'United Kingdom', 'Georgia', 'Guernsey', 'Ghana',
'Gibraltar', 'Guinea', 'Gambia', 'Guinea-Bissau',
'Equatorial Guinea', 'Greece', 'Grenada', 'Greenland', 'Guatemala',
'Guam', 'Guyana', 'Hong Kong', 'Honduras', 'Croatia', 'Haiti',
'Hungary', 'Indonesia', 'Isle of Man', 'India', 'Ireland', 'Iran',
'Iraq', 'Iceland', 'Israel', 'Italy', 'Jamaica', 'Jersey',
'Jordan', 'Japan', 'Kazakhstan', 'Kenya', 'Kyrgyzstan', 'Cambodia',
'Saint Kitts and Nevis', 'South Korea', 'Kuwait', 'Laos',
'Lebanon', 'Liberia', 'Libya', 'Saint Lucia', 'Liechtenstein',
'Sri Lanka', 'Lesotho', 'Lithuania', 'Luxembourg', 'Latvia',
'Morocco', 'Monaco', 'Moldova', 'Madagascar', 'Maldives', 'Mexico',
'Macedonia', 'Mali', 'Malta', 'Myanmar', 'Montenegro', 'Mongolia',
'Northern Mariana Islands', 'Mozambique', 'Mauritania',
'Montserrat', 'Mauritius', 'Malawi', 'Malaysia', 'Namibia',
'New Caledonia', 'Niger', 'Nigeria', 'Nicaragua', 'Netherlands',
'Norway', 'Nepal', 'New Zealand', 'Oman', 'Pakistan', 'Panama',
'Peru', 'Philippines', 'Papua New Guinea', 'Poland', 'Puerto Rico',
'Portugal', 'Paraguay', 'Palestine', 'French Polynesia', 'Qatar',
'Romania', 'Russia', 'Rwanda', 'Saudi Arabia', 'Sudan', 'Senegal',
'Singapore', 'Solomon Islands', 'Sierra Leone', 'El Salvador',
'San Marino', 'Somalia', 'Serbia', 'South Sudan',
'Sao Tome and Principe', 'Suriname', 'Slovakia', 'Slovenia',
'Sweden', 'Swaziland', 'Sint Maarten (Dutch part)', 'Seychelles',
'Syria', 'Turks and Caicos Islands', 'Chad', 'Togo', 'Thailand',
'Tajikistan', 'Timor', 'Trinidad and Tobago', 'Tunisia', 'Turkey',
'Taiwan', 'Tanzania', 'Uganda', 'Ukraine', 'Uruguay',
'United States', 'Uzbekistan', 'Vatican',
'Saint Vincent and the Grenadines', 'Venezuela',
'British Virgin Islands', 'United States Virgin Islands',
'Vietnam', 'Wallis and Futuna', 'Kosovo', 'Yemen', 'South Africa',
'Zambia', 'Zimbabwe', 'World', 'International'], dtype=object)
```

```
[13]: df = dataframe[(dataframe.location != 'International')
                    &(dataframe.location != 'World')
                    ]
```

```
[14]: df.location.unique() #seeing the new data after it being cleaned up
```

```
[14]: array(['Aruba', 'Afghanistan', 'Angola', 'Anguilla', 'Albania', 'Andorra',
           'United Arab Emirates', 'Argentina', 'Armenia',
```

```
'Antigua and Barbuda', 'Australia', 'Austria', 'Azerbaijan',
'Burundi', 'Belgium', 'Benin', 'Bonaire Sint Eustatius and Saba',
'Burkina Faso', 'Bangladesh', 'Bulgaria', 'Bahrain', 'Bahamas',
'Bosnia and Herzegovina', 'Belarus', 'Belize', 'Bermuda',
'Bolivia', 'Brazil', 'Barbados', 'Brunei', 'Bhutan', 'Botswana',
'Central African Republic', 'Canada', 'Switzerland', 'Chile',
'China', 'Cote d'Ivoire', 'Cameroon',
'Democratic Republic of Congo', 'Congo', 'Colombia', 'Comoros',
'Cape Verde', 'Costa Rica', 'Cuba', 'Curacao', 'Cayman Islands',
'Cyprus', 'Czech Republic', 'Germany', 'Djibouti', 'Dominica',
'Denmark', 'Dominican Republic', 'Algeria', 'Ecuador', 'Egypt',
'Eritrea', 'Western Sahara', 'Spain', 'Estonia', 'Ethiopia',
'Finland', 'Fiji', 'Falkland Islands', 'France', 'Faeroe Islands',
'Gabon', 'United Kingdom', 'Georgia', 'Guernsey', 'Ghana',
'Gibraltar', 'Guinea', 'Gambia', 'Guinea-Bissau',
'Equatorial Guinea', 'Greece', 'Grenada', 'Greenland', 'Guatemala',
'Guam', 'Guyana', 'Hong Kong', 'Honduras', 'Croatia', 'Haiti',
'Hungary', 'Indonesia', 'Isle of Man', 'India', 'Ireland', 'Iran',
'Iraq', 'Iceland', 'Israel', 'Italy', 'Jamaica', 'Jersey',
'Jordan', 'Japan', 'Kazakhstan', 'Kenya', 'Kyrgyzstan', 'Cambodia',
'Saint Kitts and Nevis', 'South Korea', 'Kuwait', 'Laos',
'Lebanon', 'Liberia', 'Libya', 'Saint Lucia', 'Liechtenstein',
'Sri Lanka', 'Lesotho', 'Lithuania', 'Luxembourg', 'Latvia',
'Morocco', 'Monaco', 'Moldova', 'Madagascar', 'Maldives', 'Mexico',
'Macedonia', 'Mali', 'Malta', 'Myanmar', 'Montenegro', 'Mongolia',
'Northern Mariana Islands', 'Mozambique', 'Mauritania',
'Montserrat', 'Mauritius', 'Malawi', 'Malaysia', 'Namibia',
'New Caledonia', 'Niger', 'Nigeria', 'Nicaragua', 'Netherlands',
'Norway', 'Nepal', 'New Zealand', 'Oman', 'Pakistan', 'Panama',
'Peru', 'Philippines', 'Papua New Guinea', 'Poland', 'Puerto Rico',
'Portugal', 'Paraguay', 'Palestine', 'French Polynesia', 'Qatar',
'Romania', 'Russia', 'Rwanda', 'Saudi Arabia', 'Sudan', 'Senegal',
'Singapore', 'Solomon Islands', 'Sierra Leone', 'El Salvador',
'San Marino', 'Somalia', 'Serbia', 'South Sudan',
'Sao Tome and Principe', 'Suriname', 'Slovakia', 'Slovenia',
'Sweden', 'Swaziland', 'Sint Maarten (Dutch part)', 'Seychelles',
'Syria', 'Turks and Caicos Islands', 'Chad', 'Togo', 'Thailand',
'Tajikistan', 'Timor', 'Trinidad and Tobago', 'Tunisia', 'Turkey',
'Taiwan', 'Tanzania', 'Uganda', 'Ukraine', 'Uruguay',
'United States', 'Uzbekistan', 'Vatican',
'Saint Vincent and the Grenadines', 'Venezuela',
'British Virgin Islands', 'United States Virgin Islands',
'Vietnam', 'Wallis and Futuna', 'Kosovo', 'Yemen', 'South Africa',
'Zambia', 'Zimbabwe'], dtype=object)
```

```
[15]: df.handwashing_facilities.unique() #getting numbers from new data set
```

```
[15]: array([ nan, 37.746, 26.664, 94.043, 83.241,  6.144, 11.035, 11.877,
          34.808, 97.164, 90.083, 25.383, 88.469, 79.807, 16.603, 19.351,
           2.735,  4.472, 47.964, 65.386, 15.574, 83.841, 85.198, 55.182,
          83.741, 80.635, 89.827,  7.96 , 41.047, 17.45 ,  7.876,  6.403,
          24.64 , 76.665, 77.159, 84.169, 22.863, 64.204, 59.55 , 94.576,
          66.425, 98.999, 24.651, 89.22 , 66.229, 49.839,  1.188, 87.202,
           2.117, 86.979, 50.54 , 95.803, 87.847, 52.232, 79.287, 71.18 ,
          12.227, 15.95 ,  8.704, 44.6  ,  8.978, 41.949, 47.782, 97.4  ,
          59.607, 78.463, 79.602,  4.617, 23.437, 20.859, 35.89 , 19.275,
          90.65 ,  9.831, 97.719, 41.34 , 67.779, 24.097, 70.598,  5.818,
          10.475, 90.67 , 72.704, 28.178, 89.443, 78.687, 47.953, 21.222,
          85.847, 49.542, 43.993, 13.938, 36.791])
```

```
[16]: df['location'].head(10)
```

```
[16]: 0    Aruba
      1    Aruba
      2    Aruba
      3    Aruba
      4    Aruba
      5    Aruba
      6    Aruba
      7    Aruba
      8    Aruba
      9    Aruba
      Name: location, dtype: object
```

```
[ ]:
```

```
[20]: df.groupby('location').handwashing_facilities.min().
      ↪sort_values(ascending=False).iloc[1:10] #looking at min values
```

```
[20]: location
      Serbia                97.719
      Oman                  97.400
      Bosnia and Herzegovina  97.164
      Maldives               95.803
      Iraq                   94.576
      Armenia                94.043
      Thailand               90.670
      El Salvador            90.650
      Belize                 90.083
      Name: handwashing_facilities, dtype: float64
```

```
[ ]:
```

```
[19]: df.groupby('location').handwashing_facilities.max().
      ↪sort_values(ascending=False).iloc[1:10]
```

```
[19]: location
      Serbia                97.719
      Oman                  97.400
      Bosnia and Herzegovina  97.164
      Maldives               95.803
      Iraq                  94.576
      Armenia               94.043
      Thailand              90.670
      El Salvador           90.650
      Belize                90.083
      Name: handwashing_facilities, dtype: float64
```

```
[28]: total_max = df.groupby('location').total_cases.max().
      ↪sort_values(ascending=False).iloc[1:10]
      #looking at cases max
```

```
[35]: df.groupby('location').population_density.max().sort_values(ascending=False).
      ↪iloc[1:10]
      #viewing if population density has a correlation
      #result: no similar countries with most cases and density
```

```
[35]: location
      Singapore            7915.731
      Hong Kong            7039.714
      Gibraltar            3457.100
      Bahrain              1935.907
      Maldives             1454.433
      Malta                1454.037
      Bermuda              1308.820
      Bangladesh           1265.036
      Sint Maarten (Dutch part) 1209.088
      Name: population_density, dtype: float64
```

```
[27]: df.describe()
```

```
[27]:
```

	total_cases	new_cases	new_cases_smoothed	total_deaths	\
count	4.869100e+04	51613.000000	50824.000000	39781.000000	
mean	7.126040e+04	844.667894	831.155719	3166.584224	
std	4.319899e+05	4846.531305	4723.250632	14464.209161	
min	1.000000e+00	-8261.000000	-552.000000	1.000000	
25%	1.440000e+02	0.000000	0.714000	11.000000	
50%	1.797000e+03	12.000000	17.286000	74.000000	
75%	1.724150e+04	192.000000	200.857000	613.000000	
max	8.704524e+06	97894.000000	93198.571000	225735.000000	

	new_deaths	new_deaths_smoothed	total_cases_per_million \
count	51613.000000	50824.000000	48691.000000
mean	22.492337	22.512157	2602.304369
std	124.789333	115.002189	5133.853732
min	-1918.000000	-232.143000	0.001000
25%	0.000000	0.000000	76.188000
50%	0.000000	0.286000	499.162000
75%	3.000000	3.429000	2904.073500
max	4928.000000	2715.143000	55976.186000

	new_cases_per_million	new_cases_smoothed_per_million \
count	51613.000000	50824.000000
mean	31.478524	30.502448
std	112.281096	77.437669
min	-2212.545000	-269.978000
25%	0.000000	0.208000
50%	1.880000	3.471000
75%	22.218000	25.103750
max	8652.658000	2472.188000

	total_deaths_per_million ...	gdp_per_capita	extreme_poverty \
count	39781.000000 ...	46153.000000	30673.000000
mean	87.009752 ...	20772.812607	12.336648
std	171.660670 ...	20425.804680	19.444164
min	0.001000 ...	661.240000	0.100000
25%	3.473000 ...	5321.444000	0.500000
50%	18.026000 ...	13913.839000	1.800000
75%	73.826000 ...	31400.840000	18.100000
max	1237.551000 ...	116935.600000	77.600000

	cardiovasc_death_rate	diabetes_prevalence	female_smokers \
count	46750.000000	48457.000000	36392.000000
mean	252.220346	8.053628	10.802742
std	117.776840	4.170617	10.511085
min	79.370000	0.990000	0.100000
25%	155.898000	5.310000	1.900000
50%	240.208000	7.110000	6.300000
75%	318.991000	10.390000	19.600000
max	724.417000	23.360000	44.000000

	male_smokers	handwashing_facilities	hospital_beds_per_thousand \
count	35921.000000	21896.000000	42121.000000
mean	32.623496	52.129872	3.099639
std	13.497778	31.833114	2.527981
min	7.700000	1.188000	0.100000
25%	21.400000	20.859000	1.300000

50%	31.400000	52.232000	2.500000
75%	41.100000	83.741000	4.200000
max	78.100000	98.999000	13.800000

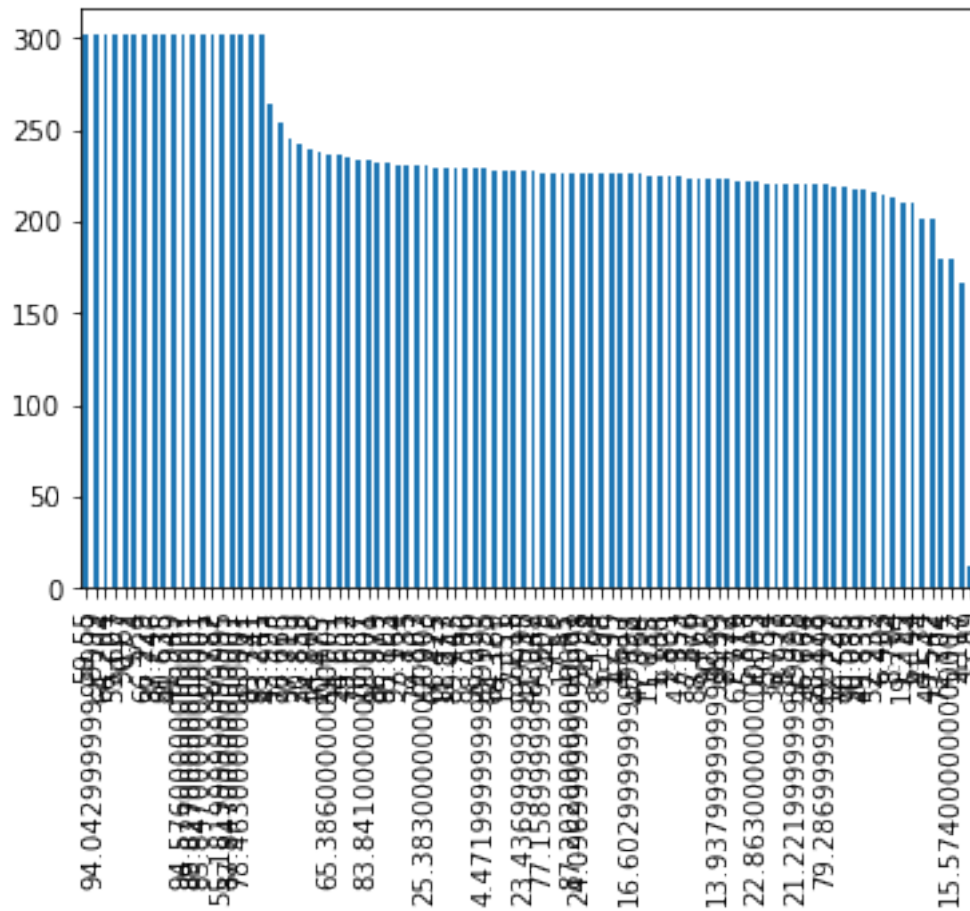
	life_expectancy	human_development_index
count	51597.000000	45419.000000
mean	73.979884	0.723251
std	7.410511	0.153115
min	53.280000	0.354000
25%	69.870000	0.601000
50%	75.400000	0.754000
75%	79.380000	0.847000
max	86.750000	0.953000

[8 rows x 36 columns]

[]:

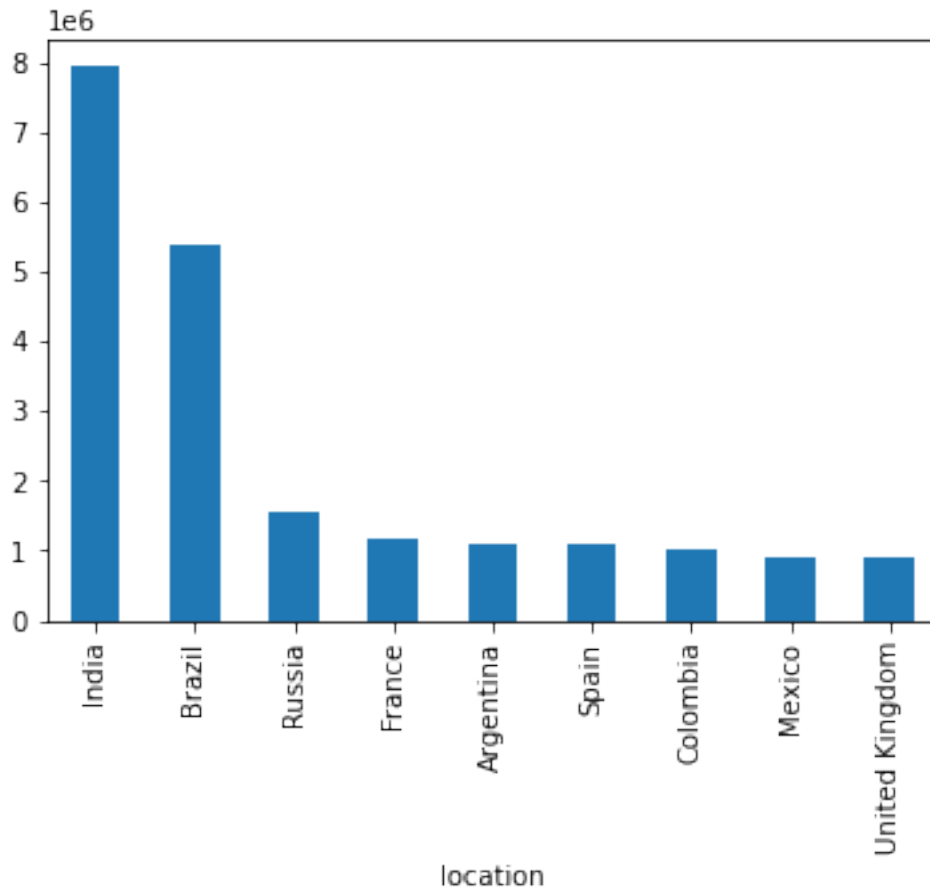
[32]: `pd.value_counts(df["handwashing_facilities"]).plot.bar()`

[32]: `<matplotlib.axes._subplots.AxesSubplot at 0x7ff8d08d4700>`



```
[48]: total_max.plot(x= 'location', y='total_cases', kind= 'bar')
      #messing around with a bar graph
```

```
[48]: <matplotlib.axes._subplots.AxesSubplot at 0x7ff8d03ebe50>
```



```
[47]: df.corr() #looking at the correlation of everything in teh data set
```

```
[47]:
```

	total_cases	new_cases	new_cases_smoothed	\
total_cases	1.000000	0.841949	0.867768	
new_cases	0.841949	1.000000	0.980825	
new_cases_smoothed	0.867768	0.980825	1.000000	
total_deaths	0.910993	0.764430	0.783014	
new_deaths	0.592009	0.730192	0.722780	
new_deaths_smoothed	0.657350	0.768107	0.786833	
total_cases_per_million	0.263493	0.195790	0.199255	
new_cases_per_million	0.099621	0.161193	0.137968	
new_cases_smoothed_per_million	0.144408	0.195391	0.194340	
total_deaths_per_million	0.273941	0.207921	0.210628	
new_deaths_per_million	0.084014	0.116865	0.112675	
new_deaths_smoothed_per_million	0.136195	0.171391	0.176385	
total_tests	0.935689	0.761913	0.773019	
new_tests	0.914756	0.905663	0.914507	
total_tests_per_thousand	0.111900	0.060844	0.059316	
new_tests_per_thousand	0.091810	0.077323	0.074209	

new_tests_smoothed	0.892340	0.864100	0.879060
new_tests_smoothed_per_thousand	0.089194	0.070857	0.069192
tests_per_case	-0.037098	-0.042904	-0.042721
positive_rate	0.038953	0.095841	0.095793
stringency_index	0.052751	0.095572	0.096065
population	0.308769	0.373536	0.382461
population_density	-0.019002	-0.019545	-0.019854
median_age	0.066413	0.058989	0.056818
aged_65_older	0.055864	0.050191	0.047378
aged_70_older	0.052667	0.046519	0.043502
gdp_per_capita	0.069860	0.046491	0.045280
extreme_poverty	-0.071829	-0.060010	-0.059338
cardiovasc_death_rate	-0.068436	-0.064372	-0.064462
diabetes_prevalence	0.035326	0.031783	0.033658
female_smokers	0.040547	0.033802	0.031024
male_smokers	-0.073605	-0.079544	-0.082188
handwashing_facilities	0.052219	0.047316	0.046337
hospital_beds_per_thousand	-0.027051	-0.034168	-0.037258
life_expectancy	0.048015	0.040126	0.038544
human_development_index	0.093660	0.080514	0.079272

	total_deaths	new_deaths \
total_cases	0.910993	0.592009
new_cases	0.764430	0.730192
new_cases_smoothed	0.783014	0.722780
total_deaths	1.000000	0.639454
new_deaths	0.639454	1.000000
new_deaths_smoothed	0.712564	0.915875
total_cases_per_million	0.260665	0.142126
new_cases_per_million	0.111746	0.105190
new_cases_smoothed_per_million	0.146368	0.143097
total_deaths_per_million	0.415523	0.213277
new_deaths_per_million	0.108466	0.353600
new_deaths_smoothed_per_million	0.180939	0.348282
total_tests	0.841018	0.479555
new_tests	0.828787	0.612653
total_tests_per_thousand	0.097982	-0.005617
new_tests_per_thousand	0.093418	0.015202
new_tests_smoothed	0.796592	0.572695
new_tests_smoothed_per_thousand	0.085682	0.003774
tests_per_case	-0.049380	-0.045187
positive_rate	0.067203	0.212362
stringency_index	0.040180	0.119794
population	0.260864	0.274401
population_density	-0.029762	-0.025882
median_age	0.124697	0.072055
aged_65_older	0.128378	0.067233

aged_70_older	0.129870	0.065403
gdp_per_capita	0.109822	0.053911
extreme_poverty	-0.116942	-0.086793
cardiovasc_death_rate	-0.144064	-0.098738
diabetes_prevalence	0.026215	0.026863
female_smokers	0.095537	0.045536
male_smokers	-0.104110	-0.096741
handwashing_facilities	0.116434	0.093445
hospital_beds_per_thousand	-0.028402	-0.039476
life_expectancy	0.108862	0.061579
human_development_index	0.158215	0.101955

	new_deaths_smoothed	total_cases_per_million \
total_cases	0.657350	0.263493
new_cases	0.768107	0.195790
new_cases_smoothed	0.786833	0.199255
total_deaths	0.712564	0.260665
new_deaths	0.915875	0.142126
new_deaths_smoothed	1.000000	0.158186
total_cases_per_million	0.158186	1.000000
new_cases_per_million	0.104784	0.435595
new_cases_smoothed_per_million	0.153708	0.627124
total_deaths_per_million	0.241931	0.590000
new_deaths_per_million	0.239184	0.187681
new_deaths_smoothed_per_million	0.380771	0.296306
total_tests	0.529080	0.228296
new_tests	0.663467	0.180727
total_tests_per_thousand	-0.005316	0.508445
new_tests_per_thousand	0.014708	0.402493
new_tests_smoothed	0.628527	0.185608
new_tests_smoothed_per_thousand	0.005063	0.441898
tests_per_case	-0.047661	-0.071458
positive_rate	0.218046	0.152023
stringency_index	0.128535	0.038760
population	0.299056	-0.045883
population_density	-0.028227	0.054461
median_age	0.077723	0.178015
aged_65_older	0.072755	0.070140
aged_70_older	0.070757	0.069289
gdp_per_capita	0.058627	0.379321
extreme_poverty	-0.093659	-0.242430
cardiovasc_death_rate	-0.108223	-0.200361
diabetes_prevalence	0.029513	0.128207
female_smokers	0.049265	0.098052
male_smokers	-0.104938	-0.044786
handwashing_facilities	0.101128	0.294176
hospital_beds_per_thousand	-0.043943	0.015554

life_expectancy	0.066623	0.271611
human_development_index	0.110552	0.273630

	new_cases_per_million \
total_cases	0.099621
new_cases	0.161193
new_cases_smoothed	0.137968
total_deaths	0.111746
new_deaths	0.105190
new_deaths_smoothed	0.104784
total_cases_per_million	0.435595
new_cases_per_million	1.000000
new_cases_smoothed_per_million	0.717471
total_deaths_per_million	0.262458
new_deaths_per_million	0.227299
new_deaths_smoothed_per_million	0.273032
total_tests	0.119425
new_tests	0.126278
total_tests_per_thousand	0.320563
new_tests_per_thousand	0.353896
new_tests_smoothed	0.128593
new_tests_smoothed_per_thousand	0.359180
tests_per_case	-0.083668
positive_rate	0.328647
stringency_index	0.120435
population	-0.027745
population_density	0.019137
median_age	0.143297
aged_65_older	0.086016
aged_70_older	0.082670
gdp_per_capita	0.198514
extreme_poverty	-0.181298
cardiovasc_death_rate	-0.122440
diabetes_prevalence	0.088617
female_smokers	0.115008
male_smokers	-0.007199
handwashing_facilities	0.249764
hospital_beds_per_thousand	0.033280
life_expectancy	0.142958
human_development_index	0.181439

	new_cases_smoothed_per_million \
total_cases	0.144408
new_cases	0.195391
new_cases_smoothed	0.194340
total_deaths	0.146368
new_deaths	0.143097

new_deaths_smoothed	0.153708
total_cases_per_million	0.627124
new_cases_per_million	0.717471
new_cases_smoothed_per_million	1.000000
total_deaths_per_million	0.333873
new_deaths_per_million	0.265621
new_deaths_smoothed_per_million	0.394760
total_tests	0.137097
new_tests	0.142409
total_tests_per_thousand	0.348536
new_tests_per_thousand	0.367167
new_tests_smoothed	0.146981
new_tests_smoothed_per_thousand	0.391379
tests_per_case	-0.093030
positive_rate	0.370019
stringency_index	0.152410
population	-0.038102
population_density	0.023557
median_age	0.159901
aged_65_older	0.089577
aged_70_older	0.085319
gdp_per_capita	0.227764
extreme_poverty	-0.208712
cardiovasc_death_rate	-0.153033
diabetes_prevalence	0.116950
female_smokers	0.132183
male_smokers	-0.013391
handwashing_facilities	0.294005
hospital_beds_per_thousand	0.029898
life_expectancy	0.196458
human_development_index	0.220967

	total_deaths_per_million	...	\
total_cases	0.273941	...	
new_cases	0.207921	...	
new_cases_smoothed	0.210628	...	
total_deaths	0.415523	...	
new_deaths	0.213277	...	
new_deaths_smoothed	0.241931	...	
total_cases_per_million	0.590000	...	
new_cases_per_million	0.262458	...	
new_cases_smoothed_per_million	0.333873	...	
total_deaths_per_million	1.000000	...	
new_deaths_per_million	0.216052	...	
new_deaths_smoothed_per_million	0.364244	...	
total_tests	0.274244	...	
new_tests	0.241499	...	

total_tests_per_thousand	0.208741	...
new_tests_per_thousand	0.195892	...
new_tests_smoothed	0.242344	...
new_tests_smoothed_per_thousand	0.191764	...
tests_per_case	-0.080626	...
positive_rate	0.082643	...
stringency_index	-0.051057	...
population	-0.028197	...
population_density	-0.024569	...
median_age	0.291655	...
aged_65_older	0.310024	...
aged_70_older	0.322438	...
gdp_per_capita	0.282312	...
extreme_poverty	-0.258657	...
cardiovasc_death_rate	-0.307682	...
diabetes_prevalence	-0.067840	...
female_smokers	0.305491	...
male_smokers	-0.104884	...
handwashing_facilities	0.263976	...
hospital_beds_per_thousand	0.045304	...
life_expectancy	0.340106	...
human_development_index	0.336861	...

	gdp_per_capita	extreme_poverty \
total_cases	0.069860	-0.071829
new_cases	0.046491	-0.060010
new_cases_smoothed	0.045280	-0.059338
total_deaths	0.109822	-0.116942
new_deaths	0.053911	-0.086793
new_deaths_smoothed	0.058627	-0.093659
total_cases_per_million	0.379321	-0.242430
new_cases_per_million	0.198514	-0.181298
new_cases_smoothed_per_million	0.227764	-0.208712
total_deaths_per_million	0.282312	-0.258657
new_deaths_per_million	0.066491	-0.121377
new_deaths_smoothed_per_million	0.100490	-0.176949
total_tests	0.096119	-0.057008
new_tests	0.090282	-0.048606
total_tests_per_thousand	0.477220	-0.209806
new_tests_per_thousand	0.503764	-0.220315
new_tests_smoothed	0.093949	-0.054533
new_tests_smoothed_per_thousand	0.538266	-0.236257
tests_per_case	0.030601	-0.068193
positive_rate	-0.158301	0.065564
stringency_index	-0.185554	0.015376
population	-0.071186	-0.031828
population_density	0.300043	-0.025178

median_age	0.630701	-0.693633
aged_65_older	0.483500	-0.568454
aged_70_older	0.475848	-0.551062
gdp_per_capita	1.000000	-0.497718
extreme_poverty	-0.497718	1.000000
cardiovasc_death_rate	-0.479492	0.191783
diabetes_prevalence	0.240242	-0.420077
female_smokers	0.344893	-0.398832
male_smokers	-0.110304	-0.192942
handwashing_facilities	0.654423	-0.767047
hospital_beds_per_thousand	0.278567	-0.425909
life_expectancy	0.684450	-0.750749
human_development_index	0.755392	-0.768519

	cardiovasc_death_rate	diabetes_prevalence \
total_cases	-0.068436	0.035326
new_cases	-0.064372	0.031783
new_cases_smoothed	-0.064462	0.033658
total_deaths	-0.144064	0.026215
new_deaths	-0.098738	0.026863
new_deaths_smoothed	-0.108223	0.029513
total_cases_per_million	-0.200361	0.128207
new_cases_per_million	-0.122440	0.088617
new_cases_smoothed_per_million	-0.153033	0.116950
total_deaths_per_million	-0.307682	-0.067840
new_deaths_per_million	-0.107247	-0.000562
new_deaths_smoothed_per_million	-0.159067	-0.002434
total_tests	0.014520	0.092432
new_tests	0.010850	0.100419
total_tests_per_thousand	-0.132791	0.147415
new_tests_per_thousand	-0.155710	0.124710
new_tests_smoothed	-0.001389	0.099458
new_tests_smoothed_per_thousand	-0.173915	0.151951
tests_per_case	-0.071292	-0.007406
positive_rate	0.050203	0.119026
stringency_index	0.072007	0.079080
population	0.017449	0.031001
population_density	-0.177535	0.011914
median_age	-0.338581	0.154064
aged_65_older	-0.338003	-0.070513
aged_70_older	-0.350589	-0.102834
gdp_per_capita	-0.479492	0.240242
extreme_poverty	0.191783	-0.420077
cardiovasc_death_rate	1.000000	0.042174
diabetes_prevalence	0.042174	1.000000
female_smokers	-0.189431	-0.098694
male_smokers	0.437215	0.162834

handwashing_facilities	-0.038977	0.562801
hospital_beds_per_thousand	0.002022	-0.106187
life_expectancy	-0.494998	0.271163
human_development_index	-0.440872	0.247534

	female_smokers	male_smokers \
total_cases	0.040547	-0.073605
new_cases	0.033802	-0.079544
new_cases_smoothed	0.031024	-0.082188
total_deaths	0.095537	-0.104110
new_deaths	0.045536	-0.096741
new_deaths_smoothed	0.049265	-0.104938
total_cases_per_million	0.098052	-0.044786
new_cases_per_million	0.115008	-0.007199
new_cases_smoothed_per_million	0.132183	-0.013391
total_deaths_per_million	0.305491	-0.104884
new_deaths_per_million	0.110731	-0.047989
new_deaths_smoothed_per_million	0.155210	-0.071195
total_tests	0.053011	-0.004740
new_tests	0.047120	-0.020875
total_tests_per_thousand	0.164963	0.026704
new_tests_per_thousand	0.192132	0.020184
new_tests_smoothed	0.063235	-0.022319
new_tests_smoothed_per_thousand	0.194557	0.017042
tests_per_case	-0.028191	0.005587
positive_rate	-0.193338	-0.114655
stringency_index	-0.189484	-0.040319
population	-0.133906	0.035413
population_density	-0.066630	-0.000568
median_age	0.680270	0.190647
aged_65_older	0.774867	0.087758
aged_70_older	0.778346	0.088238
gdp_per_capita	0.344893	-0.110304
extreme_poverty	-0.398832	-0.192942
cardiovasc_death_rate	-0.189431	0.437215
diabetes_prevalence	-0.098694	0.162834
female_smokers	1.000000	0.177382
male_smokers	0.177382	1.000000
handwashing_facilities	0.266021	0.335684
hospital_beds_per_thousand	0.487219	0.345966
life_expectancy	0.523000	0.054582
human_development_index	0.573898	0.083001

	handwashing_facilities \
total_cases	0.052219
new_cases	0.047316
new_cases_smoothed	0.046337

total_deaths	0.116434
new_deaths	0.093445
new_deaths_smoothed	0.101128
total_cases_per_million	0.294176
new_cases_per_million	0.249764
new_cases_smoothed_per_million	0.294005
total_deaths_per_million	0.263976
new_deaths_per_million	0.104773
new_deaths_smoothed_per_million	0.189763
total_tests	0.029016
new_tests	0.029783
total_tests_per_thousand	0.324369
new_tests_per_thousand	0.391435
new_tests_smoothed	0.031173
new_tests_smoothed_per_thousand	0.402904
tests_per_case	0.108685
positive_rate	0.044594
stringency_index	0.113095
population	0.040036
population_density	0.085360
median_age	0.783346
aged_65_older	0.632638
aged_70_older	0.630011
gdp_per_capita	0.654423
extreme_poverty	-0.767047
cardiovasc_death_rate	-0.038977
diabetes_prevalence	0.562801
female_smokers	0.266021
male_smokers	0.335684
handwashing_facilities	1.000000
hospital_beds_per_thousand	0.441046
life_expectancy	0.825678
human_development_index	0.842737

	hospital_beds_per_thousand	life_expectancy \
total_cases	-0.027051	0.048015
new_cases	-0.034168	0.040126
new_cases_smoothed	-0.037258	0.038544
total_deaths	-0.028402	0.108862
new_deaths	-0.039476	0.061579
new_deaths_smoothed	-0.043943	0.066623
total_cases_per_million	0.015554	0.271611
new_cases_per_million	0.033280	0.142958
new_cases_smoothed_per_million	0.029898	0.196458
total_deaths_per_million	0.045304	0.340106
new_deaths_per_million	0.006488	0.103934
new_deaths_smoothed_per_million	0.006908	0.162017

total_tests	0.010401	0.031479
new_tests	-0.008788	0.026869
total_tests_per_thousand	0.084997	0.296960
new_tests_per_thousand	0.097463	0.326203
new_tests_smoothed	-0.002329	0.038315
new_tests_smoothed_per_thousand	0.069440	0.345892
tests_per_case	-0.011145	0.089895
positive_rate	-0.264836	-0.200251
stringency_index	-0.183079	-0.140433
population	-0.043312	-0.035884
population_density	0.328169	0.207196
median_age	0.661741	0.849103
aged_65_older	0.651092	0.731059
aged_70_older	0.652528	0.718253
gdp_per_capita	0.278567	0.684450
extreme_poverty	-0.425909	-0.750749
cardiovasc_death_rate	0.002022	-0.494998
diabetes_prevalence	-0.106187	0.271163
female_smokers	0.487219	0.523000
male_smokers	0.345966	0.054582
handwashing_facilities	0.441046	0.825678
hospital_beds_per_thousand	1.000000	0.465983
life_expectancy	0.465983	1.000000
human_development_index	0.538400	0.913589

	human_development_index
total_cases	0.093660
new_cases	0.080514
new_cases_smoothed	0.079272
total_deaths	0.158215
new_deaths	0.101955
new_deaths_smoothed	0.110552
total_cases_per_million	0.273630
new_cases_per_million	0.181439
new_cases_smoothed_per_million	0.220967
total_deaths_per_million	0.336861
new_deaths_per_million	0.125723
new_deaths_smoothed_per_million	0.181338
total_tests	0.098284
new_tests	0.093286
total_tests_per_thousand	0.351194
new_tests_per_thousand	0.379162
new_tests_smoothed	0.101183
new_tests_smoothed_per_thousand	0.402099
tests_per_case	0.069522
positive_rate	-0.216824
stringency_index	-0.172363

population	-0.024289
population_density	0.149532
median_age	0.886772
aged_65_older	0.760137
aged_70_older	0.747496
gdp_per_capita	0.755392
extreme_poverty	-0.768519
cardiovasc_death_rate	-0.440872
diabetes_prevalence	0.247534
female_smokers	0.573898
male_smokers	0.083001
handwashing_facilities	0.842737
hospital_beds_per_thousand	0.538400
life_expectancy	0.913589
human_development_index	1.000000

[36 rows x 36 columns]

```
[46]: column1 = df["total_cases"]
      column2 = df["handwashing_facilities"]
      correlation = column1. corr(column2)
      print(correlation)

      #looking at the exact correlation between total cases and handwashing stations
```

0.052218559942789955

[]:

[]: