covid-19

November 7, 2024

1. IMPORTING LIBRARIES

```python
[2]: import pandas as pd
     import numpy as np
     import seaborn as sns
     import matplotlib.pyplot as plt
```

2. IMPORTING THE DATASET

```python
[3]: data = pd.read_csv(r"H:\DA. Python\11. COVID 19 Micro Analysis\covid_19_data.
       ↪csv")
```

3. CHECKING THE DATASET.

```python
[4]: data
```

```
[4]:         Date      State         Region  Confirmed  Deaths  Recovered
     0    4/29/2020       NaN    Afghanistan       1939      60        252
     1    4/29/2020       NaN        Albania        766      30        455
     2    4/29/2020       NaN        Algeria       3848     444       1702
     3    4/29/2020       NaN        Andorra        743      42        423
     4    4/29/2020       NaN         Angola         27       2          7
     ..        ...       ...            ...        ...     ...        ...
     316  4/29/2020   Wyoming             US        545       7          0
     317  4/29/2020  Xinjiang  Mainland China         76       3         73
     318  4/29/2020     Yukon         Canada         11       0          0
     319  4/29/2020    Yunnan  Mainland China        185       2        181
     320  4/29/2020  Zhejiang  Mainland China       1268       1       1263

     [321 rows x 6 columns]
```
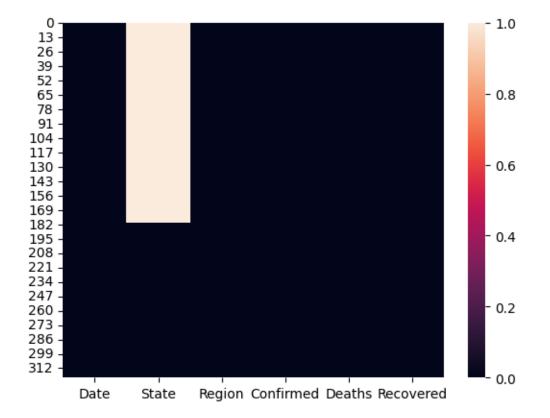
```python
[5]: data.shape
```

```
[5]: (321, 6)
```

```python
[6]: data.count()
```

```
[6]: Date       321
     State      140
     Region     321
```

```
Confirmed    321
Deaths       321
Recovered    321
dtype: int64
```

[7]: `data.isnull().sum()`

[7]:
```
Date          0
State       181
Region        0
Confirmed     0
Deaths        0
Recovered     0
dtype: int64
```

[8]:
```
sns.heatmap(data.isnull())
plt.show()
```


Show the number of Confirmed, Deaths and Recovered cases in each region.

```
[9]: data.head(2)
```

```
[9]:         Date State         Region  Confirmed  Deaths  Recovered
     0  4/29/2020   NaN   Afghanistan       1939      60        252
     1  4/29/2020   NaN       Albania        766      30        455
```

```
[10]: data.columns
```

```
[10]: Index(['Date', 'State', 'Region', 'Confirmed', 'Deaths', 'Recovered'],
      dtype='object')
```

```
[11]: data.groupby('Region')[['Confirmed', 'Deaths', 'Recovered']].sum()
```

```
[11]:                      Confirmed  Deaths  Recovered
     Region
     Afghanistan              1939      60        252
     Albania                   766      30        455
     Algeria                  3848     444       1702
     Andorra                   743      42        423
     Angola                     27       2          7
     ...                       ...     ...        ...
     West Bank and Gaza        344       2         71
     Western Sahara              6       0          5
     Yemen                       6       0          1
     Zambia                     97       3         54
     Zimbabwe                   32       4          5

     [187 rows x 3 columns]
```

Remove all the records where Confirmed cases is less than 10

```
[12]: data.head(2)
```

```
[12]:         Date State         Region  Confirmed  Deaths  Recovered
     0  4/29/2020   NaN   Afghanistan       1939      60        252
     1  4/29/2020   NaN       Albania        766      30        455
```

```
[15]: # This condition checks if the 'Confirmed' cases in the 'data' DataFrame are␣
      ↪less than 10.
      # It will return a Boolean series, where each element is True if the␣
      ↪'Confirmed' value is less than 10, and False otherwise.

      data.Confirmed < 10
```

```
[15]: 0      False
     1      False
     2      False
```

```
3      False
4      False
       …
316    False
317    False
318    False
319    False
320    False
Name: Confirmed, Length: 321, dtype: bool
```

```
[14]: # This line filters the 'data' DataFrame, returning only the rows where the
      ↪'Confirmed' cases are less than 10.
      # It uses a Boolean condition to select and display those rows that meet the
      ↪criteria.

      data[data.Confirmed < 10]
```

```
[14]:           Date                         State                    Region  \
      18   4/29/2020                           NaN                    Bhutan
      98   4/29/2020                           NaN                MS Zaandam
      105  4/29/2020                           NaN                Mauritania
      126  4/29/2020                           NaN          Papua New Guinea
      140  4/29/2020                           NaN   Sao Tome and Principe
      177  4/29/2020                           NaN            Western Sahara
      178  4/29/2020                           NaN                     Yemen
      184  4/29/2020                      Anguilla                        UK
      192  4/29/2020  Bonaire, Sint Eustatius and Saba         Netherlands
      194  4/29/2020         British Virgin Islands                      UK
      203  4/29/2020     Diamond Princess cruise ship                Canada
      272  4/29/2020          Northwest Territories                 Canada
      284  4/29/2020                     Recovered                 Canada
      285  4/29/2020                     Recovered                     US
      288  4/29/2020             Saint Barthelemy                 France
      289  4/29/2020     Saint Pierre and Miquelon                 France
      305  4/29/2020                         Tibet        Mainland China

           Confirmed  Deaths  Recovered
      18           7       0          5
      98           9       2          0
      105          8       1          6
      126          8       0          0
      140          8       0          4
      177          6       0          5
      178          6       0          1
      184          3       0          3
      192          5       0          0
      194          6       1          3
```

```
203          0          1          0
272          5          0          0
284          0          0      20327
285          0          0     120720
288          6          0          6
289          1          0          0
305          1          0          1
```

[16]: `data[~(data.Confirmed < 10)] #using tilde "~" to remove records`

[16]:
```
         Date     State        Region  Confirmed  Deaths  Recovered
0    4/29/2020       NaN   Afghanistan       1939      60        252
1    4/29/2020       NaN       Albania        766      30        455
2    4/29/2020       NaN       Algeria       3848     444       1702
3    4/29/2020       NaN       Andorra        743      42        423
4    4/29/2020       NaN        Angola         27       2          7
..         ...       ...           ...        ...     ...        ...
316  4/29/2020   Wyoming            US        545       7          0
317  4/29/2020  Xinjiang  Mainland China       76       3         73
318  4/29/2020     Yukon        Canada         11       0          0
319  4/29/2020    Yunnan  Mainland China      185       2        181
320  4/29/2020  Zhejiang  Mainland China     1268       1       1263

[304 rows x 6 columns]
```

[18]: `data = data[~(data.Confirmed < 10)] #Removing the records permanently`

[19]: `data`

[19]:
```
         Date     State        Region  Confirmed  Deaths  Recovered
0    4/29/2020       NaN   Afghanistan       1939      60        252
1    4/29/2020       NaN       Albania        766      30        455
2    4/29/2020       NaN       Algeria       3848     444       1702
3    4/29/2020       NaN       Andorra        743      42        423
4    4/29/2020       NaN        Angola         27       2          7
..         ...       ...           ...        ...     ...        ...
316  4/29/2020   Wyoming            US        545       7          0
317  4/29/2020  Xinjiang  Mainland China       76       3         73
318  4/29/2020     Yukon        Canada         11       0          0
319  4/29/2020    Yunnan  Mainland China      185       2        181
320  4/29/2020  Zhejiang  Mainland China     1268       1       1263

[304 rows x 6 columns]
```

[20]: `data.head(20)`

5

```
[20]:        Date State                    Region  Confirmed  Deaths  Recovered
     0   4/29/2020   NaN               Afghanistan       1939      60        252
     1   4/29/2020   NaN                   Albania        766      30        455
     2   4/29/2020   NaN                   Algeria       3848     444       1702
     3   4/29/2020   NaN                   Andorra        743      42        423
     4   4/29/2020   NaN                    Angola         27       2          7
     5   4/29/2020   NaN       Antigua and Barbuda         24       3         11
     6   4/29/2020   NaN                 Argentina       4285     214       1192
     7   4/29/2020   NaN                   Armenia       1932      30        900
     8   4/29/2020   NaN                   Austria      15402     580      12779
     9   4/29/2020   NaN                Azerbaijan       1766      23       1267
     10  4/29/2020   NaN                   Bahamas         80      11         23
     11  4/29/2020   NaN                   Bahrain       2921       8       1455
     12  4/29/2020   NaN                Bangladesh       7103     163        150
     13  4/29/2020   NaN                  Barbados         80       7         39
     14  4/29/2020   NaN                   Belarus      13181      84       2072
     15  4/29/2020   NaN                   Belgium      47859    7501      11283
     16  4/29/2020   NaN                    Belize         18       2          9
     17  4/29/2020   NaN                     Benin         64       1         33
     19  4/29/2020   NaN                   Bolivia       1110      59        117
     20  4/29/2020   NaN    Bosnia and Herzegovina       1677      65        710
```

In which Region, maximum number of Confirmed cases were recorded?

```
[21]: data.head(2)
```

```
[21]:        Date State       Region  Confirmed  Deaths  Recovered
     0   4/29/2020   NaN  Afghanistan       1939      60        252
     1   4/29/2020   NaN      Albania        766      30        455
```

```
[23]: data.groupby('Region')['Confirmed'].sum().sort_values(ascending=False).head(20)
```

```
[23]: Region
     US                1039909
     Spain              236899
     Italy              203591
     France             166536
     UK                 166432
     Germany            161539
     Turkey             117589
     Russia              99399
     Iran                93657
     Mainland China      82861
     Brazil              79685
     Canada              52860
     Belgium             47859
     Netherlands         38993
```

```
Peru              33931
India             33062
Switzerland       29407
Ecuador           24675
Portugal          24505
Saudi Arabia      21402
Name: Confirmed, dtype: int64
```

In which region, minimum number of Deaths cases were recorded?

[24]: `data.head(2)`

[24]:
```
       Date State      Region  Confirmed  Deaths  Recovered
0  4/29/2020   NaN  Afghanistan       1939      60        252
1  4/29/2020   NaN      Albania        766      30        455
```

[28]: `data.groupby("Region")["Deaths"].sum().sort_values(ascending=True).head(50)`

[28]:
```
Region
Cambodia                            0
Fiji                                0
Dominica                            0
Central African Republic            0
Eritrea                             0
Laos                                0
Holy See                            0
Nepal                               0
Madagascar                          0
Macau                               0
Mozambique                          0
Mongolia                            0
Namibia                             0
Grenada                             0
Saint Lucia                         0
Saint Vincent and the Grenadines    0
Timor-Leste                         0
Uganda                              0
Seychelles                          0
South Sudan                         0
Rwanda                              0
Saint Kitts and Nevis               0
Vietnam                             0
Liechtenstein                       1
Benin                               1
Gambia                              1
Eswatini                            1
Equatorial Guinea                   1
```

```
Guinea-Bissau              1
Maldives                   1
Cabo Verde                 1
Botswana                   1
Burundi                    1
Brunei                     1
Suriname                   1
Djibouti                   2
Chad                       2
West Bank and Gaza         2
Angola                     2
Libya                      2
Belize                     2
Malawi                     3
Ethiopia                   3
Syria                      3
Zambia                     3
Nicaragua                  3
Gabon                      3
Antigua and Barbuda        3
Monaco                     4
Zimbabwe                   4
Name: Deaths, dtype: int64
```

How many Confirmed, Deaths & Recovered cases where reported from India till 29 april 2020

[29]:
```
data.head(2)
```

[29]:
```
        Date State       Region  Confirmed  Deaths  Recovered
0  4/29/2020   NaN  Afghanistan       1939      60        252
1  4/29/2020   NaN      Albania        766      30        455
```

[32]:
```
data[data.Region == 'India']

# dataframe name - data
# column name - Region
# values - India
```

[32]:
```
        Date State Region  Confirmed  Deaths  Recovered
74  4/29/2020   NaN  India      33062    1079       8437
```

Sort the entire data wrt No. of Confirmed cases in ascending order

[37]:
```
data.head(2)
```

```
[37]:           Date State      Region  Confirmed  Deaths  Recovered
      0   4/29/2020    NaN  Afghanistan       1939      60        252
      1   4/29/2020    NaN      Albania        766      30        455
```

```
[34]: data.sort_values(by=['Confirmed'], ascending=True)
```

```
[34]:            Date      State      Region  Confirmed  Deaths  Recovered
      70   4/29/2020        NaN    Holy See         10       0          2
      59   4/29/2020        NaN      Gambia         10       1          8
      156  4/29/2020        NaN     Suriname        10       1          8
      27   4/29/2020        NaN      Burundi        11       1          4
      217  4/29/2020   Greenland     Denmark        11       0         11
      ..        ...        ...         ...        ...     ...        ...
      57   4/29/2020        NaN      France     165093   24087      48228
      168  4/29/2020        NaN          UK     165221   26097          0
      80   4/29/2020        NaN       Italy     203591   27682      71252
      153  4/29/2020        NaN       Spain     236899   24275     132929
      265  4/29/2020   New York          US     299691   23477          0

      [304 rows x 6 columns]
```

Sort the entire data wrt No. of Recovered cases in descending order

```
[35]: data.head(2)
```

```
[35]:           Date State      Region  Confirmed  Deaths  Recovered
      0   4/29/2020    NaN  Afghanistan       1939      60        252
      1   4/29/2020    NaN      Albania        766      30        455
```

```
[36]: data.sort_values(by=["Recovered"], ascending=False)
```

```
[36]:            Date                    State          Region  Confirmed  Deaths  \
      153  4/29/2020                     NaN            Spain     236899   24275
      61   4/29/2020                     NaN          Germany     161539    6467
      76   4/29/2020                     NaN             Iran      93657    5957
      80   4/29/2020                     NaN            Italy     203591   27682
      229  4/29/2020                   Hubei   Mainland China      68128    4512
      ..        ...                     ...              ...        ...     ...
      270  4/29/2020  Northern Mariana Islands            US         14       2
      299  4/29/2020            South Dakota              US       2373      13
      298  4/29/2020          South Carolina              US       5882     231
      302  4/29/2020               Tennessee              US      10366     195
      303  4/29/2020                   Texas              US      27257     754

           Recovered
      153     132929
      61      120400
```

9

```
76        73791
80        71252
229       63616
..          …
270           0
299           0
298           0
302           0
303           0

[304 rows x 6 columns]
```