stage-2

March 13, 2023

[1]:
```
import pandas as pd
import numpy as np
import datetime
import plotly.express as px
from statistics import mode
```

[2]:
```
# Reading the confirmed cases data

confirmed = pd.read_csv("../../DATASETS/COVID DATASETS/covid_confirmed_usafacts.
 ↪csv")
confirmed
```

[2]:
```
      countyFIPS            County Name State  StateFIPS  2020-01-22  \
0              0  Statewide Unallocated    AL          1           0
1           1001         Autauga County    AL          1           0
2           1003         Baldwin County    AL          1           0
3           1005         Barbour County    AL          1           0
4           1007            Bibb County    AL          1           0
…            …                     …    …       …          …
3188       56037      Sweetwater County    WY         56           0
3189       56039          Teton County    WY         56           0
3190       56041          Uinta County    WY         56           0
3191       56043       Washakie County    WY         56           0
3192       56045         Weston County    WY         56           0

      2020-01-23  2020-01-24  2020-01-25  2020-01-26  2020-01-27  …  \
0              0           0           0           0           0  …
1              0           0           0           0           0  …
2              0           0           0           0           0  …
3              0           0           0           0           0  …
4              0           0           0           0           0  …
…            …         …         …         …         …       …  …
3188           0           0           0           0           0  …
3189           0           0           0           0           0  …
3190           0           0           0           0           0  …
3191           0           0           0           0           0  …
3192           0           0           0           0           0  …
```

```
      2023-01-07  2023-01-08  2023-01-09  2023-01-10  2023-01-11  2023-01-12  \
0              0           0           0           0           0           0
1          19205       19205       19205       19205       19205       19318
2          68182       68182       68182       68182       68182       68518
3           7120        7120        7120        7120        7120        7188
4           7808        7808        7808        7808        7808        7855
…          …         …         …         …         …         …
3188       12394       12394       12430       12430       12430       12437
3189       11997       11997       12035       12035       12035       12045
3190        6303        6303        6318        6318        6318        6333
3191        2717        2717        2727        2727        2727        2731
3192        1876        1876        1880        1880        1880        1881

      2023-01-13  2023-01-14  2023-01-15  2023-01-16
0              0           0           0           0
1          19318       19318       19318       19318
2          68518       68518       68518       68518
3           7188        7188        7188        7188
4           7855        7855        7855        7855
…          …         …         …         …
3188       12437       12437       12437       12437
3189       12045       12045       12045       12045
3190        6333        6333        6333        6333
3191        2731        2731        2731        2731
3192        1881        1881        1881        1881

[3193 rows x 1095 columns]
```

[3]:

```python
# filtering the data by removing 'statewide unallocated' data.

confirmed = confirmed[confirmed['County Name'] != 'Statewide Unallocated']
confirmed
```

[3]:

```
      countyFIPS         County Name State  StateFIPS  2020-01-22  2020-01-23  \
1           1001     Autauga County     AL          1           0           0
2           1003     Baldwin County     AL          1           0           0
3           1005     Barbour County     AL          1           0           0
4           1007        Bibb County     AL          1           0           0
5           1009      Blount County     AL          1           0           0
…          …                 …   …        …         …         …
3188       56037  Sweetwater County     WY         56           0           0
3189       56039       Teton County     WY         56           0           0
3190       56041       Uinta County     WY         56           0           0
3191       56043    Washakie County     WY         56           0           0
3192       56045      Weston County     WY         56           0           0

      2020-01-24  2020-01-25  2020-01-26  2020-01-27  …  2023-01-07  \
1              0          0          0          0  …       19205
2              0          0          0          0  …       68182
3              0          0          0          0  …        7120
4              0          0          0          0  …        7808
5              0          0          0          0  …       17952
…            …         …         …         …  …          …
3188           0          0          0          0  …       12394
3189           0          0          0          0  …       11997
3190           0          0          0          0  …        6303
3191           0          0          0          0  …        2717
3192           0          0          0          0  …        1876

      2023-01-08  2023-01-09  2023-01-10  2023-01-11  2023-01-12  2023-01-13  \
1          19205       19205       19205       19205       19318       19318
2          68182       68182       68182       68182       68518       68518
3           7120        7120        7120        7120        7188        7188
4           7808        7808        7808        7808        7855        7855
5          17952       17952       17952       17952       18057       18057
…            …         …         …         …         …          …
3188       12394       12430       12430       12430       12437       12437
3189       11997       12035       12035       12035       12045       12045
3190        6303        6318        6318        6318        6333        6333
3191        2717        2727        2727        2727        2731        2731
3192        1876        1880        1880        1880        1881        1881

      2023-01-14  2023-01-15  2023-01-16
1          19318       19318       19318
2          68518       68518       68518
3           7188        7188        7188
4           7855        7855        7855
5          18057       18057       18057
…            …         …         …
3188       12437       12437       12437
3189       12045       12045       12045
3190        6333        6333        6333
3191        2731        2731        2731
3192        1881        1881        1881

[3142 rows x 1095 columns]
```

[4]:

```
# Making a deep copy of the data and calculating the day wise confirmed cases.


confirmed_2 = confirmed.copy(deep=True)
for i in range(len(confirmed.columns)):
    if confirmed.columns[i] == "countyFIPS" or confirmed.columns[i]=="County␣
 ↪Name" or confirmed.columns[i]=="State" or confirmed.columns[i]=="StateFIPS":
        continue
    elif i>4:
        pres =  confirmed.columns[i]
        prev =  confirmed_2.columns[i-1]
        confirmed[pres] = confirmed[pres]-confirmed_2[prev]
```

```
C:\Users\venka\AppData\Local\Temp\ipykernel_18648\886717563.py:11:
SettingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame.
Try using .loc[row_indexer,col_indexer] = value instead

See the caveats in the documentation: https://pandas.pydata.org/pandas-
docs/stable/user_guide/indexing.html#returning-a-view-versus-a-copy
  confirmed[pres] = confirmed[pres]-confirmed_2[prev]
```

[5]:

```
# Day wise confirmed cases.

confirmed
```

[5]:

```
      countyFIPS         County Name State  StateFIPS  2020-01-22  2020-01-23  \
1           1001      Autauga County    AL          1           0           0
2           1003      Baldwin County    AL          1           0           0
3           1005      Barbour County    AL          1           0           0
4           1007         Bibb County    AL          1           0           0
5           1009       Blount County    AL          1           0           0
…          …                 …   …        …         …         …
3188       56037  Sweetwater County    WY         56           0           0
3189       56039       Teton County    WY         56           0           0
3190       56041       Uinta County    WY         56           0           0
3191       56043    Washakie County    WY         56           0           0
3192       56045      Weston County    WY         56           0           0

      2020-01-24  2020-01-25  2020-01-26  2020-01-27  …  2023-01-07  \
1              0           0           0           0  …           0
2              0           0           0           0  …           0
3              0           0           0           0  …           0
4              0           0           0           0  …           0
5              0           0           0           0  …           0
…          …         …         …         …  …         …
3188           0           0           0           0  …           0
3189           0           0           0           0  …           0
3190           0           0           0           0  …           0
3191           0           0           0           0  …           0
3192           0           0           0           0  …           0
```

```
      2023-01-08  2023-01-09  2023-01-10  2023-01-11  2023-01-12  2023-01-13  \
1              0           0           0           0         113           0
2              0           0           0           0         336           0
3              0           0           0           0          68           0
4              0           0           0           0          47           0
5              0           0           0           0         105           0
…            …          …          …          …          …          …
3188           0          36           0           0           7           0
3189           0          38           0           0          10           0
3190           0          15           0           0          15           0
3191           0          10           0           0           4           0
3192           0           4           0           0           1           0

      2023-01-14  2023-01-15  2023-01-16
1              0           0           0
2              0           0           0
3              0           0           0
4              0           0           0
5              0           0           0
…            …          …          …
3188           0           0           0
3189           0           0           0
3190           0           0           0
3191           0           0           0
3192           0           0           0

[3142 rows x 1095 columns]
```

[6]:

```python
# calculating the day of the starting and ending dates of the data, so we can add missing dates to get complete data for the week.

x=pd.to_datetime('2022-06-01')

y=pd.to_datetime('2022-12-31')



print(x.dayofweek,y.dayofweek)
```

```
2 5
```

[7]:

```python
# Removing the excess data and keeping only the columns we want.

requiredRange=confirmed.drop(confirmed.loc[:,'2020-01-22':'2022-05-29'],axis=1)
```

[8]:

```python
# Removing the excess data and keeping only the columns we want.

requiredRange.drop(requiredRange.loc[:,'2023-01-02':],axis=1,inplace=True)
requiredRange
```

```
[8]:       countyFIPS         County Name State  StateFIPS  2022-05-30  2022-05-31  \
     1           1001      Autauga County    AL          1           9          24
     2           1003      Baldwin County    AL          1          55         183
     3           1005      Barbour County    AL          1           1          12
     4           1007         Bibb County    AL          1           9           9
     5           1009       Blount County    AL          1           6          12
     …              …                 …   …      …       …         …
     3188       56037  Sweetwater County    WY         56           0          29
     3189       56039        Teton County    WY         56           0          62
     3190       56041        Uinta County    WY         56           0          23
     3191       56043     Washakie County    WY         56           0           3
     3192       56045       Weston County    WY         56           0           4

           2022-06-01  2022-06-02  2022-06-03  2022-06-04  …  2022-12-23  \
     1              6           9           0           0  …           0
     2             68          68           0           0  …           0
     3              3           4           0           0  …           0
     4              8           4           0           0  …           0
     5              4           7           0           0  …           0
     …              …         …         …           …  …              …
     3188           0           0           0           0  …           0
     3189           0           0           0           0  …           0
     3190           0           0           0           0  …           0
     3191           0           0           0           0  …           0
     3192           0           0           0           0  …           0

           2022-12-24  2022-12-25  2022-12-26  2022-12-27  2022-12-28  2022-12-29  \
     1              0           0           0           0           0           0
     2              0           0           0           0           0           0
     3              0           0           0           0           0           0
     4              0           0           0           0           0           0
     5              0           0           0           0           0           0
     …              …         …         …         …         …         …
     3188           0           0           0           0           0           0
     3189           0           0           0           0           0           0
     3190           0           0           0           0           0           0
     3191           0           0           0           0           0           0
     3192           0           0           0           0           0           0

           2022-12-30  2022-12-31  2023-01-01
     1              0           0           0
     2              0           0           0
     3              0           0           0
     4              0           0           0
     5              0           0           0
     …              …         …         …
     3188           0           0           0
```

```
3189            0           0           0
3190            0           0           0
3191            0           0           0
3192            0           0           0

[3142 rows x 221 columns]
```

[9]:

```
# Creating a new data frame for weekly data.

weekCases = requiredRange.loc[:,('countyFIPS','County␣
 ↪Name','State','StateFIPS')]
weekCases
```

[9]:

```
      countyFIPS          County Name State  StateFIPS
1           1001      Autauga County     AL          1
2           1003      Baldwin County     AL          1
3           1005      Barbour County     AL          1
4           1007         Bibb County     AL          1
5           1009       Blount County     AL          1
…              …                 …    …          …
3188       56037   Sweetwater County     WY         56
3189       56039        Teton County     WY         56
3190       56041        Uinta County     WY         56
3191       56043     Washakie County     WY         56
3192       56045       Weston County     WY         56

[3142 rows x 4 columns]
```

[10]:

```
# Calculating the weeekly sum for each week and represeting the sum at each␣
 ↪sunday of the week.

for i in range(10,len(requiredRange.columns),7):
    col_list=requiredRange.columns[i-6:i+1]

    col=requiredRange.columns[i]

    weekCases[col]= requiredRange[col_list].sum(axis=1)
```

[11]:

```
weekCases
```

[11]:

```
      countyFIPS          County Name State  StateFIPS  2022-06-05  2022-06-12  \
1           1001      Autauga County     AL          1          48         117
2           1003      Baldwin County     AL          1         374         518
3           1005      Barbour County     AL          1          20          34
4           1007         Bibb County     AL          1          30          35
```

```
5           1009      Blount County     AL          1          29          53
…          …                 …   …        …         …         …
3188       56037  Sweetwater County     WY         56          29          56
3189       56039       Teton County     WY         56          62         174
3190       56041       Uinta County     WY         56          23          21
3191       56043    Washakie County     WY         56           3           2
3192       56045      Weston County     WY         56           4          10

      2022-06-19  2022-06-26  2022-07-03  2022-07-10  …  2022-10-30  \
1            116         189         162         239  …          31
2            560         646         611         760  …          78
3             34          50          67          94  …           4
4             64          52          51          53  …          15
5             70          76         100         123  …          34
…          …         …         …         …  …         …
3188         211          50          69          70  …          18
3189         151         145         108         104  …          14
3190          51          31          31          41  …          10
3191         100          19          -8          16  …           3
3192          28           9           2          24  …           7

      2022-11-06  2022-11-13  2022-11-20  2022-11-27  2022-12-04  2022-12-11  \
1              0          60          21          20          68          72
2              0         240          55         212         250         221
3              0          20           9           9          12           9
4              0          29           8          15          10          16
5              0          66          27          47          40          59
…          …         …         …         …          …         …
3188          10          19          14          38          41          45
3189          20          17          16          21          23          36
3190          11          13          11          18          16          17
3191           1           8         -17          11           6          20
3192           7           5           4           5          11           2

      2022-12-18  2022-12-25  2023-01-01
1             95         114           0
2            270         275           0
3             18          20           0
4             15          24           0
5             89          83           0
…          …         …         …
3188          25          20           0
3189          21          29           0
3190          15          16           0
3191          16           0           0
3192           9           5           0
```

```
[3142 rows x 35 columns]
```

[12]:

```
# Using the melt function to covert wide dataframe into long dataframe so that␣
↪we can perform operations on the data easily.

weekNewCases=pd.melt(weekCases,id_vars=('countyFIPS','County␣
↪Name','State','StateFIPS'),var_name='Date',value_name='Number of new cases')
```

[13]:

```
weekNewCases
```

[13]:

```
       countyFIPS         County Name State  StateFIPS        Date   \
0            1001     Autauga County     AL          1  2022-06-05
1            1003     Baldwin County     AL          1  2022-06-05
2            1005     Barbour County     AL          1  2022-06-05
3            1007        Bibb County     AL          1  2022-06-05
4            1009      Blount County     AL          1  2022-06-05
…           …                  …   …        …        …
97397       56037  Sweetwater County     WY         56  2023-01-01
97398       56039       Teton County     WY         56  2023-01-01
97399       56041       Uinta County     WY         56  2023-01-01
97400       56043    Washakie County     WY         56  2023-01-01
97401       56045      Weston County     WY         56  2023-01-01

       Number of new cases
0                       48
1                      374
2                       20
3                       30
4                       29
…                     …
97397                    0
97398                    0
97399                    0
97400                    0
97401                    0

[97402 rows x 6 columns]
```

[14]:

```
# Converting the 'Date' column to datetime type.

weekNewCases['Date']=pd.to_datetime(weekNewCases['Date'])
```

[15]:

```
# Data types of the 'weeklyNewCases' dataframe.

weekNewCases.dtypes
```

[15]:

```
countyFIPS                         int64
County Name                       object
State                             object
StateFIPS                          int64
Date                      datetime64[ns]
Number of new cases                int64
dtype: object
```

[16]:

```
# Reading deaths data.

deaths = pd.read_csv("../../DATASETS/COVID DATASETS/covid_deaths_usafacts.csv")
deaths
```

[16]:

```
      countyFIPS            County Name State  StateFIPS  2020-01-22  \
0              0  Statewide Unallocated    AL          1           0
1           1001         Autauga County    AL          1           0
2           1003         Baldwin County    AL          1           0
3           1005         Barbour County    AL          1           0
4           1007            Bibb County    AL          1           0
…          …                   …   …        …         …
3188       56037      Sweetwater County    WY         56           0
3189       56039           Teton County    WY         56           0
3190       56041           Uinta County    WY         56           0
3191       56043        Washakie County    WY         56           0
3192       56045          Weston County    WY         56           0

      2020-01-23  2020-01-24  2020-01-25  2020-01-26  2020-01-27  …  \
0              0           0           0           0           0  …
1              0           0           0           0           0  …
2              0           0           0           0           0  …
3              0           0           0           0           0  …
4              0           0           0           0           0  …
…          …         …         …         …         …  …
3188           0           0           0           0           0  …
3189           0           0           0           0           0  …
3190           0           0           0           0           0  …
3191           0           0           0           0           0  …
3192           0           0           0           0           0  …

      2023-01-07  2023-01-08  2023-01-09  2023-01-10  2023-01-11  2023-01-12  \
0              0           0           0           0           0           0
1            230         230         230         230         230         230
2            719         719         719         719         719         721
3            103         103         103         103         103         103
4            108         108         108         108         108         108
…          …         …         …         …         …         …
3188         136         136         136         136         136         136
```

```
3189          16          16          16          16          16          16
3190          43          43          43          43          43          43
3191          47          47          47          47          47          47
3192          22          22          22          22          22          22

      2023-01-13  2023-01-14  2023-01-15  2023-01-16
0              0           0           0           0
1            230         230         230         230
2            721         721         721         721
3            103         103         103         103
4            108         108         108         108
…           …         …         …         …
3188         136         136         136         136
3189          16          16          16          16
3190          43          43          43          43
3191          47          47          47          47
3192          22          22          22          22

[3193 rows x 1095 columns]
```

[17]:
```
# Removing the excess data and keeping only the columns we want by removing␣
↪'Statewide Unallocated'

deaths = deaths[deaths['County Name'] != 'Statewide Unallocated']
deaths
```

[17]:
```
      countyFIPS         County Name State  StateFIPS  2020-01-22  2020-01-23  \
1           1001     Autauga County     AL          1           0           0
2           1003     Baldwin County     AL          1           0           0
3           1005     Barbour County     AL          1           0           0
4           1007        Bibb County     AL          1           0           0
5           1009      Blount County     AL          1           0           0
…           …                 …   …        …         …         …
3188       56037  Sweetwater County     AL         56           0           0
3189       56039       Teton County     WY         56           0           0
3190       56041       Uinta County     WY         56           0           0
3191       56043    Washakie County     WY         56           0           0
3192       56045      Weston County     WY         56           0           0

      2020-01-24  2020-01-25  2020-01-26  2020-01-27  …  2023-01-07  \
1              0           0           0           0  …         230
2              0           0           0           0  …         719
3              0           0           0           0  …         103
4              0           0           0           0  …         108
5              0           0           0           0  …         260
…           …         …         …         …  …         …
3188           0           0           0           0  …         136
```

```
3189           0           0           0           0  …         16
3190           0           0           0           0  …         43
3191           0           0           0           0  …         47
3192           0           0           0           0  …         22

      2023-01-08  2023-01-09  2023-01-10  2023-01-11  2023-01-12  2023-01-13  \
1            230         230         230         230         230         230
2            719         719         719         719         721         721
3            103         103         103         103         103         103
4            108         108         108         108         108         108
5            260         260         260         260         261         261
…             …          …          …          …          …          …
3188         136         136         136         136         136         136
3189          16          16          16          16          16          16
3190          43          43          43          43          43          43
3191          47          47          47          47          47          47
3192          22          22          22          22          22          22

      2023-01-14  2023-01-15  2023-01-16
1            230         230         230
2            721         721         721
3            103         103         103
4            108         108         108
5            261         261         261
…             …          …          …
3188         136         136         136
3189          16          16          16
3190          43          43          43
3191          47          47          47
3192          22          22          22

[3142 rows x 1095 columns]
```

[18]:

```python
# Making a deep copy of the data and calculating the day wise Deaths.

deaths_2 = deaths.copy(deep=True)
for i in range(len(deaths.columns)):
    if deaths.columns[i] == "countyFIPS" or deaths.columns[i]=="County Name" or␣
 ↪deaths.columns[i]=="State" or deaths.columns[i]=="StateFIPS":
        continue
    elif i>4:
        pres =  deaths.columns[i]
        prev =  deaths_2.columns[i-1]
        deaths[pres] = deaths[pres]-deaths_2[prev]
```

```
C:\Users\venka\AppData\Local\Temp\ipykernel_18648\3437054305.py:10:
SettingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame.
```

```
Try using .loc[row_indexer,col_indexer] = value instead

See the caveats in the documentation: https://pandas.pydata.org/pandas-
docs/stable/user_guide/indexing.html#returning-a-view-versus-a-copy
  deaths[pres] = deaths[pres]-deaths_2[prev]
```

[19]:

```
# Day wise Deaths.

deaths
```

[19]:

```
      countyFIPS         County Name State  StateFIPS  2020-01-22  2020-01-23  \
1           1001      Autauga County    AL          1           0           0
2           1003      Baldwin County    AL          1           0           0
3           1005      Barbour County    AL          1           0           0
4           1007         Bibb County    AL          1           0           0
5           1009       Blount County    AL          1           0           0
…           …                  …   …      …         …         …
3188       56037  Sweetwater County    WY         56           0           0
3189       56039       Teton County    WY         56           0           0
3190       56041       Uinta County    WY         56           0           0
3191       56043    Washakie County    WY         56           0           0
3192       56045      Weston County    WY         56           0           0

      2020-01-24  2020-01-25  2020-01-26  2020-01-27  …  2023-01-07  \
1              0           0           0           0  …           0
2              0           0           0           0  …           0
3              0           0           0           0  …           0
4              0           0           0           0  …           0
5              0           0           0           0  …           0
…           …         …         …         …   …         …
3188           0           0           0           0  …           0
3189           0           0           0           0  …           0
3190           0           0           0           0  …           0
3191           0           0           0           0  …           0
3192           0           0           0           0  …           0

      2023-01-08  2023-01-09  2023-01-10  2023-01-11  2023-01-12  2023-01-13  \
1              0           0           0           0           0           0
2              0           0           0           0           2           0
3              0           0           0           0           0           0
4              0           0           0           0           0           0
5              0           0           0           0           1           0
…           …         …         …         …         …         …
3188           0           0           0           0           0           0
3189           0           0           0           0           0           0
3190           0           0           0           0           0           0
3191           0           0           0           0           0           0
```

```
3192           0           0           0           0           0           0

      2023-01-14  2023-01-15  2023-01-16
1              0           0           0
2              0           0           0
3              0           0           0
4              0           0           0
5              0           0           0
…          …         …         …
3188           0           0           0
3189           0           0           0
3190           0           0           0
3191           0           0           0
3192           0           0           0

[3142 rows x 1095 columns]
```

[20]:

```
# Removing the excess data and keeping only the columns we want.

requiredDeathsRange=deaths.drop(deaths.loc[:,'2020-01-22':'2022-05-29'],axis=1)
```

[21]:

```
# Removing the excess data and keeping only the columns we want.

requiredDeathsRange.drop(requiredDeathsRange.loc[:,'2023-01-02':
 ↪],axis=1,inplace=True)
requiredDeathsRange
```

[21]:

```
      countyFIPS         County Name State  StateFIPS  2022-05-30  2022-05-31  \
1           1001      Autauga County    AL          1           0           0
2           1003      Baldwin County    AL          1           1           0
3           1005      Barbour County    AL          1           0           0
4           1007         Bibb County    AL          1           0           0
5           1009       Blount County    AL          1           0           1
…          …                 …   …     …         …         …
3188       56037   Sweetwater County    WY         56           0           0
3189       56039        Teton County    WY         56           0           0
3190       56041        Uinta County    WY         56           0           0
3191       56043     Washakie County    WY         56           0           0
3192       56045       Weston County    WY         56           0           0

      2022-06-01  2022-06-02  2022-06-03  2022-06-04  …  2022-12-23  \
1              0           0           0           0  …           0
2              0           0           0           0  …           0
3              0           0           0           0  …           0
4              0           0           0           0  …           0
5              0           0           0           0  …           0
…          …         …         …         …  …         …
```

```
3188           0           0           0           0  …           0
3189           0           0           0           0  …           0
3190           0           0           0           0  …           0
3191           0           0           0           0  …           0
3192           0           0           0           0  …           0

      2022-12-24  2022-12-25  2022-12-26  2022-12-27  2022-12-28  2022-12-29  \
1              0           0           0           0           0           0
2              0           0           0           0           0           0
3              0           0           0           0           0           0
4              0           0           0           0           0           0
5              0           0           0           0           0           0
…            …         …         …         …         …         …
3188           0           0           0           0           0           0
3189           0           0           0           0           0           0
3190           0           0           0           0           0           0
3191           0           0           0           0           0           0
3192           0           0           0           0           0           0

      2022-12-30  2022-12-31  2023-01-01
1              0           0           0
2              0           0           0
3              0           0           0
4              0           0           0
5              0           0           0
…            …         …         …
3188           0           0           0
3189           0           0           0
3190           0           0           0
3191           0           0           0
3192           0           0           0

[3142 rows x 221 columns]
```

[22]:

```
# Creating a new dataframe for weeklyDeaths.

weekDeaths = requiredDeathsRange.loc[:,('countyFIPS','County␣
 ↪Name','State','StateFIPS')]
weekDeaths
```

[22]:

```
      countyFIPS        County Name State  StateFIPS
1           1001     Autauga County    AL          1
2           1003     Baldwin County    AL          1
3           1005     Barbour County    AL          1
4           1007        Bibb County    AL          1
5           1009      Blount County    AL          1
…            …                 …   …         …
```

```
3188       56037  Sweetwater County     WY         56
3189       56039       Teton County     WY         56
3190       56041       Uinta County     WY         56
3191       56043    Washakie County     WY         56
3192       56045      Weston County     WY         56

[3142 rows x 4 columns]
```

[23]:

```
# Calculating the weeekly sum for each week and represeting the sum at each sunday of the week.

for i in range(10,len(requiredDeathsRange.columns),7):
    col_list=requiredDeathsRange.columns[i-6:i+1]

    col=requiredDeathsRange.columns[i]

    weekDeaths[col]= requiredDeathsRange[col_list].sum(axis=1)
```

[24]:

```
# weeklyDeaths dataframe

weekDeaths
```

[24]:

```
      countyFIPS        County Name State  StateFIPS  2022-06-05  2022-06-12  \
1           1001     Autauga County    AL          1           0           1
2           1003     Baldwin County    AL          1           1           0
3           1005     Barbour County    AL          1           0           0
4           1007        Bibb County    AL          1           0           0
5           1009      Blount County    AL          1           1           1
…            …                …   …        …          …           …
3188       56037  Sweetwater County    WY         56           0           0
3189       56039       Teton County    WY         56           0           0
3190       56041       Uinta County    WY         56           0           0
3191       56043    Washakie County    WY         56           0           0
3192       56045      Weston County    WY         56           0           1

      2022-06-19  2022-06-26  2022-07-03  2022-07-10  …  2022-10-30  \
1              0           0           0           0  …           0
2              0           0           0           0  …           0
3              0           0           0           0  …           0
4              0           0           0           0  …           0
5              0           0           1           0  …           0
…            …         …         …         …  …          …
3188           0           0           0           0  …           0
3189           0           0           0           0  …           0
3190           0           0           0           0  …           0
3191           0           0           0           0  …           0
3192           0           0           0           0  …           0
```

```
      2022-11-06  2022-11-13  2022-11-20  2022-11-27  2022-12-04  2022-12-11  \
1              0           1           0           1           0           0
2              0           0           0           0           0           1
3              0           0           0           0           0           0
4              0           0           0           0           0           0
5              0           0           0           0           1           1
…          …         …         …         …         …         …
3188           0           0           0           0           3           0
3189           0           0           0           0           0           0
3190           0           0           0           0           0           0
3191           0           0           0           0           0           0
3192           0           0           1           0           0           0

      2022-12-18  2022-12-25  2023-01-01
1              0           0           0
2              0           2           0
3              0           0           0
4              0           0           0
5              0           0           0
…          …         …         …
3188           1           1           0
3189           0           0           0
3190           0           0           0
3191           0           0           0
3192           0           0           0

[3142 rows x 35 columns]
```

[25]:
```
# Using the melt function to covert wide dataframe into long dataframe so that␣
 ↪we can perform operations on the data easily.

weekNewDeaths=pd.melt(weekDeaths,id_vars=('countyFIPS','County␣
 ↪Name','State','StateFIPS'),var_name='Date',value_name='Number of new Deaths')
```

[26]:
```
# Weekly deaths dataframe.

weekNewDeaths
```

[26]:
```
       countyFIPS         County Name State  StateFIPS        Date  \
0            1001      Autauga County    AL          1  2022-06-05
1            1003      Baldwin County    AL          1  2022-06-05
2            1005      Barbour County    AL          1  2022-06-05
3            1007         Bibb County    AL          1  2022-06-05
4            1009       Blount County    AL          1  2022-06-05
…            …                   …   …          …          …
97397       56037    Sweetwater County    WY         56  2023-01-01
```

```
97398       56039       Teton County     WY          56   2023-01-01
97399       56041       Uinta County     WY          56   2023-01-01
97400       56043    Washakie County     WY          56   2023-01-01
97401       56045      Weston County     WY          56   2023-01-01

       Number of new Deaths
0                         0
1                         1
2                         0
3                         0
4                         1
…                       …
97397                     0
97398                     0
97399                     0
97400                     0
97401                     0

[97402 rows x 6 columns]
```

[27]:
```
# Converting the 'date' column to datetime type.

weekNewDeaths['Date']=pd.to_datetime(weekNewDeaths['Date'])
```

[28]:
```
# 'weekNewDeaths' data types.

weekNewDeaths.dtypes
```

[28]:
```
countyFIPS                              int64
County Name                            object
State                                  object
StateFIPS                               int64
Date                           datetime64[ns]
Number of new Deaths                    int64
dtype: object
```

[29]:
```
# merging the cases and deaths dataframes.

weeklyCombined=pd.merge(weekNewCases,weekNewDeaths,on=('countyFIPS','County␣
 ↪Name','State','StateFIPS','Date'),how='inner')
```

[30]:
```
weeklyCombined
```

[30]:
```
       countyFIPS         County Name State  StateFIPS       Date  \
0            1001     Autauga County    AL          1 2022-06-05
1            1003     Baldwin County    AL          1 2022-06-05
2            1005     Barbour County    AL          1 2022-06-05
```

```
3             1007        Bibb County     AL          1  2022-06-05
4             1009      Blount County     AL          1  2022-06-05
…             …                 …   …       …         …
97397        56037  Sweetwater County     WY         56  2023-01-01
97398        56039       Teton County     WY         56  2023-01-01
97399        56041       Uinta County     WY         56  2023-01-01
97400        56043    Washakie County     WY         56  2023-01-01
97401        56045      Weston County     WY         56  2023-01-01

       Number of new cases  Number of new Deaths
0                       48                     0
1                      374                     1
2                       20                     0
3                       30                     0
4                       29                     1
…                      …                     …
97397                    0                     0
97398                    0                     0
97399                    0                     0
97400                    0                     0
97401                    0                     0

[97402 rows x 7 columns]
```

[31]:

```
# converting the daywise data from wide dataframe to long dataframe and merging␣
↪them.

dayWiseCases = pd.melt(requiredRange, id_vars=('countyFIPS','County␣
↪Name','State','StateFIPS'), var_name='Date', value_name='Number of new␣
↪cases')
dayWiseDeaths = pd.melt(requiredDeathsRange, id_vars=('countyFIPS','County␣
↪Name','State','StateFIPS'), var_name='Date', value_name='Number of new␣
↪Deaths')
dayWise = pd.merge(dayWiseCases,dayWiseDeaths, on=('countyFIPS','County␣
↪Name','State','StateFIPS','Date'), how = 'inner')
dayWise
```

[31]:

```
        countyFIPS        County Name State  StateFIPS        Date  \
0             1001     Autauga County    AL          1  2022-05-30
1             1003     Baldwin County    AL          1  2022-05-30
2             1005     Barbour County    AL          1  2022-05-30
3             1007        Bibb County    AL          1  2022-05-30
4             1009      Blount County    AL          1  2022-05-30
…             …                 …   …      …          …
681809       56037  Sweetwater County    WY         56  2023-01-01
681810       56039       Teton County    WY         56  2023-01-01
681811       56041       Uinta County    WY         56  2023-01-01
```

```
681812       56043    Washakie County     WY          56  2023-01-01
681813       56045      Weston County     WY          56  2023-01-01

        Number of new cases  Number of new Deaths
0                         9                     0
1                        55                     1
2                         1                     0
3                         9                     0
4                         6                     0
…                     …                   …
681809                    0                     0
681810                    0                     0
681811                    0                     0
681812                    0                     0
681813                    0                     0

[681814 rows x 7 columns]
```

[32]:

```
# converting the date column to datetime type.

dayWise['Date'] = pd.to_datetime(dayWise['Date'])
dayWise.dtypes
```

[32]:

```
countyFIPS                         int64
County Name                       object
State                             object
StateFIPS                          int64
Date                      datetime64[ns]
Number of new cases                int64
Number of new Deaths               int64
dtype: object
```

[33]:

```
# choosing a state from daywise data.

virginia = dayWise[dayWise["State"]=='VA']
virginia
```

[33]:

```
        countyFIPS             County Name State  StateFIPS       Date  \
2820         51001        Accomack County     VA          51 2022-05-30
2821         51003       Albemarle County     VA          51 2022-05-30
2822         51005       Alleghany County     VA          51 2022-05-30
2823         51007          Amelia County     VA          51 2022-05-30
2824         51009         Amherst County     VA          51 2022-05-30
…            …                     …   …        …        …
681620       51800         City of Suffolk    VA          51 2023-01-01
681621       51810  City of Virginia Beach    VA          51 2023-01-01
681622       51820      City of Waynesboro    VA          51 2023-01-01
```

```
681623       51830    City of Williamsburg    VA          51  2023-01-01
681624       51840      City of Winchester    VA          51  2023-01-01

        Number of new cases  Number of new Deaths
2820                      0                     0
2821                      0                     0
2822                      0                     0
2823                      0                     0
2824                      0                     0
…                       …                     …
681620                    0                     0
681621                    0                     0
681622                    0                     0
681623                    0                     0
681624                    0                     0

[28861 rows x 7 columns]
```

[34]:

```
# calculating mean, median and mode for the new cases in selected state.

casesStats = virginia.groupby([pd.Grouper(key='Date', freq ='w')])['Number of␣
 ↪new cases'].agg(['mean','median', lambda x : x.mode()[0]]).reset_index()
casesStats
```

[34]:

```
         Date       mean  median  <lambda_0>
0  2022-06-05  21.360902     2.0           0
1  2022-06-12  21.604726     5.0           0
2  2022-06-19  19.094522     4.0           0
3  2022-06-26  18.667025     4.0           0
4  2022-07-03  21.110634     5.0           0
5  2022-07-10  19.708915     4.0           0
6  2022-07-17  22.818475     5.0           0
7  2022-07-24  22.716434     6.0           0
8  2022-07-31  22.932331     6.0           0
9  2022-08-07  21.794844     6.0           0
10 2022-08-14  20.265306     6.0           0
11 2022-08-21  18.456498     6.0           0
12 2022-08-28  18.713212     6.0           0
13 2022-09-04  18.161117     6.0           0
14 2022-09-11  14.172932     3.0           0
15 2022-09-18  12.459721     3.0           0
16 2022-09-25  10.827068     3.0           0
17 2022-10-02   7.624060     1.0           0
18 2022-10-09   9.153598     2.0           0
19 2022-10-16   7.569280     2.0           0
20 2022-10-23   7.671321     2.0           0
21 2022-10-30   8.095596     2.0           0
```

```
22 2022-11-06   7.866810     2.0           0
23 2022-11-13   7.477981     2.0           0
24 2022-11-20   6.299678     0.0           0
25 2022-11-27   6.493018     0.0           0
26 2022-12-04  10.117078     3.0           0
27 2022-12-11  10.899033     3.0           0
28 2022-12-18  13.296455     3.0           0
29 2022-12-25  16.058002     4.0           0
30 2023-01-01   9.240602     0.0           0
```

[35]:
```
# renaming the column in the dataframe.


casesStats.rename(columns = {'<lambda_0>' : 'mode'},inplace = True)
```

[36]:
```
casesStats
```

[36]:
```
         Date       mean  median  mode
0  2022-06-05  21.360902     2.0     0
1  2022-06-12  21.604726     5.0     0
2  2022-06-19  19.094522     4.0     0
3  2022-06-26  18.667025     4.0     0
4  2022-07-03  21.110634     5.0     0
5  2022-07-10  19.708915     4.0     0
6  2022-07-17  22.818475     5.0     0
7  2022-07-24  22.716434     6.0     0
8  2022-07-31  22.932331     6.0     0
9  2022-08-07  21.794844     6.0     0
10 2022-08-14  20.265306     6.0     0
11 2022-08-21  18.456498     6.0     0
12 2022-08-28  18.713212     6.0     0
13 2022-09-04  18.161117     6.0     0
14 2022-09-11  14.172932     3.0     0
15 2022-09-18  12.459721     3.0     0
16 2022-09-25  10.827068     3.0     0
17 2022-10-02   7.624060     1.0     0
18 2022-10-09   9.153598     2.0     0
19 2022-10-16   7.569280     2.0     0
20 2022-10-23   7.671321     2.0     0
21 2022-10-30   8.095596     2.0     0
22 2022-11-06   7.866810     2.0     0
23 2022-11-13   7.477981     2.0     0
24 2022-11-20   6.299678     0.0     0
25 2022-11-27   6.493018     0.0     0
26 2022-12-04  10.117078     3.0     0
27 2022-12-11  10.899033     3.0     0
28 2022-12-18  13.296455     3.0     0
```

```
29 2022-12-25  16.058002     4.0     0
30 2023-01-01   9.240602     0.0     0
```

[37]:

```
# plotting weekly statistics for new cases for selected state.

fig_1 = px.line(casesStats, x='Date', y=['mean','median','mode'],title="Weekly␣
 ↪statistics for new cases ")
fig_1.show()
```

[38]:

```
# calculating weekly statistics for new deaths for selected.

deathsStats = virginia.groupby([pd.Grouper(key='Date', freq='w')])['Number of␣
 ↪new Deaths'].agg(['mean','median',lambda x : x.mode()[0]]).reset_index()
deathsStats
```

[38]:

```
         Date      mean  median  <lambda_0>
0  2022-06-05  0.051557     0.0           0
1  2022-06-12  0.018260     0.0           0
2  2022-06-19  0.012889     0.0           0
3  2022-06-26  0.070892     0.0           0
4  2022-07-03  0.119227     0.0           0
5  2022-07-10  0.118153     0.0           0
6  2022-07-17  0.065521     0.0           0
7  2022-07-24  0.070892     0.0           0
8  2022-07-31  0.054780     0.0           0
9  2022-08-07  0.146079     0.0           0
10 2022-08-14  0.097744     0.0           0
11 2022-08-21  0.111708     0.0           0
12 2022-08-28  0.114930     0.0           0
13 2022-09-04  0.114930     0.0           0
14 2022-09-11  0.133190     0.0           0
15 2022-09-18  0.098818     0.0           0
16 2022-09-25  0.096670     0.0           0
17 2022-10-02  0.085929     0.0           0
18 2022-10-09  0.116004     0.0           0
19 2022-10-16  0.080559     0.0           0
20 2022-10-23  0.081633     0.0           0
21 2022-10-30  0.087003     0.0           0
22 2022-11-06  0.035446     0.0           0
23 2022-11-13  0.010741     0.0           0
24 2022-11-20  0.056928     0.0           0
25 2022-11-27  0.087003     0.0           0
26 2022-12-04  0.097744     0.0           0
27 2022-12-11  0.087003     0.0           0
28 2022-12-18  0.058002     0.0           0
29 2022-12-25  0.061224     0.0           0
30 2023-01-01  0.022556     0.0           0
```

[39]:

```
# renaming the column in the dataframe.

deathsStats.rename(columns={'<lambda_0>':'mode'}, inplace=True)
deathsStats
```

[39]:

```
          Date      mean  median  mode
0   2022-06-05  0.051557     0.0     0
1   2022-06-12  0.018260     0.0     0
2   2022-06-19  0.012889     0.0     0
3   2022-06-26  0.070892     0.0     0
4   2022-07-03  0.119227     0.0     0
5   2022-07-10  0.118153     0.0     0
6   2022-07-17  0.065521     0.0     0
7   2022-07-24  0.070892     0.0     0
8   2022-07-31  0.054780     0.0     0
9   2022-08-07  0.146079     0.0     0
10  2022-08-14  0.097744     0.0     0
11  2022-08-21  0.111708     0.0     0
12  2022-08-28  0.114930     0.0     0
13  2022-09-04  0.114930     0.0     0
14  2022-09-11  0.133190     0.0     0
15  2022-09-18  0.098818     0.0     0
16  2022-09-25  0.096670     0.0     0
17  2022-10-02  0.085929     0.0     0
18  2022-10-09  0.116004     0.0     0
19  2022-10-16  0.080559     0.0     0
20  2022-10-23  0.081633     0.0     0
21  2022-10-30  0.087003     0.0     0
22  2022-11-06  0.035446     0.0     0
23  2022-11-13  0.010741     0.0     0
24  2022-11-20  0.056928     0.0     0
25  2022-11-27  0.087003     0.0     0
26  2022-12-04  0.097744     0.0     0
27  2022-12-11  0.087003     0.0     0
28  2022-12-18  0.058002     0.0     0
29  2022-12-25  0.061224     0.0     0
30  2023-01-01  0.022556     0.0     0
```

[40]:

```
# plotting the weekly statistics for new deaths for a selected state.

fig_2 = px.line(deathsStats, x='Date', y=['mean','median','mode'],␣
 ↪title='Weekly statistics for new deaths')
fig_2.show()
```

[41]:

```
# choosing 3 other states along with the above state.

compStates = dayWise[dayWise['State'].isin(['NC','AR','AZ','VA'])]
```

```
compStates
```

[41]:

```
        countyFIPS             County Name State  StateFIPS       Date  \
96            4001          Apache County     AZ          4 2022-05-30
97            4003         Cochise County     AZ          4 2022-05-30
98            4005        Coconino County     AZ          4 2022-05-30
99            4007            Gila County     AZ          4 2022-05-30
100           4009          Graham County     AZ          4 2022-05-30
…            …                     …   …        …        …
681620       51800         City of Suffolk    VA         51 2023-01-01
681621       51810  City of Virginia Beach    VA         51 2023-01-01
681622       51820      City of Waynesboro    VA         51 2023-01-01
681623       51830    City of Williamsburg    VA         51 2023-01-01
681624       51840      City of Winchester    VA         51 2023-01-01

        Number of new cases  Number of new Deaths
96                        0                     0
97                        0                     0
98                        0                     0
99                        0                     0
100                       0                     0
…                     …                   …
681620                    0                     0
681621                    0                     0
681622                    0                     0
681623                    0                     0
681624                    0                     0

[70091 rows x 7 columns]
```

[42]:

```
# calculating the mean values for new cases and new deaths for 4 states.

meanValues = compStates.groupby(['State', pd.
 ↪Grouper(key='Date',freq='w')])['Number of new cases','Number of new Deaths'].
 ↪agg({'Number of new cases':'mean', 'Number of new Deaths':'mean'}).
 ↪reset_index()
meanValues
```

```
C:\Users\venka\AppData\Local\Temp\ipykernel_18648\2671363035.py:3:
FutureWarning:

Indexing with multiple keys (implicitly converted to a tuple of keys) will be
deprecated, use a list instead.
```

[42]:

```
    State       Date  Number of new cases  Number of new Deaths
0      AR 2022-06-05             5.038095              0.051429
1      AR 2022-06-12             7.308571              0.026667
```

```
2      AR 2022-06-19            8.944762             0.040000
3      AR 2022-06-26           10.365714             0.059048
4      AR 2022-07-03           13.253333             0.040000
..    …         …                …                    …
119    VA 2022-12-04           10.117078             0.097744
120    VA 2022-12-11           10.899033             0.087003
121    VA 2022-12-18           13.296455             0.058002
122    VA 2022-12-25           16.058002             0.061224
123    VA 2023-01-01            9.240602             0.022556

[124 rows x 4 columns]
```

[43]:
```
# plotting the mean values for new cases and new deaths for the 4 states.

fig_3 = px.line(meanValues, x='Date',y=['Number of new cases','Number of new␣
 ↪Deaths'], color='State', title='Weekly mean values for 4 states')
fig_3.show()
```

[44]:
```
# calculating the median values for new cases and new deaths for the 4 states.


medianValues=compStates.groupby(['State',pd.Grouper(key='Date',␣
 ↪freq='w')])["Number of new cases",'Number of new Deaths'].agg({'Number of␣
 ↪new cases':'median','Number of new Deaths':'median'}).reset_index()
medianValues
```

```
C:\Users\venka\AppData\Local\Temp\ipykernel_18648\1151941052.py:4:
FutureWarning:

Indexing with multiple keys (implicitly converted to a tuple of keys) will be
deprecated, use a list instead.
```

[44]:
```
    State       Date  Number of new cases  Number of new Deaths
0      AR 2022-06-05                  1.0                   0.0
1      AR 2022-06-12                  2.0                   0.0
2      AR 2022-06-19                  3.0                   0.0
3      AR 2022-06-26                  3.0                   0.0
4      AR 2022-07-03                  4.0                   0.0
..    …         …                …                    …
119    VA 2022-12-04                  3.0                   0.0
120    VA 2022-12-11                  3.0                   0.0
121    VA 2022-12-18                  3.0                   0.0
122    VA 2022-12-25                  4.0                   0.0
123    VA 2023-01-01                  0.0                   0.0

[124 rows x 4 columns]
```

```
[45]: # plotting median values for the 4 states.

      fig_4 = px.line(medianValues, x='Date', y=['Number of new cases','Number of new␣
       ↪Deaths'], color='State', title='Weekly median values for 4 states')
      fig_4.show()
```

```
[46]: # calculating mode values for new cases and new deaths for the 4 states.

      modeValues = compStates.groupby(['State',pd.
       ↪Grouper(key='Date',freq='w')])['Number of new cases','Number of new Deaths'].
       ↪apply(lambda x : x.mode()).reset_index()
      modeValues
```

```
C:\Users\venka\AppData\Local\Temp\ipykernel_18648\3629254403.py:3:
FutureWarning:

Indexing with multiple keys (implicitly converted to a tuple of keys) will be
deprecated, use a list instead.
```

```
[46]:     State       Date  level_2  Number of new cases  Number of new Deaths
      0      AR 2022-06-05        0                    0                     0
      1      AR 2022-06-12        0                    0                     0
      2      AR 2022-06-19        0                    0                     0
      3      AR 2022-06-26        0                    0                     0
      4      AR 2022-07-03        0                    0                     0
      ..    …          …        …                  …                   …
      119    VA 2022-12-04        0                    0                     0
      120    VA 2022-12-11        0                    0                     0
      121    VA 2022-12-18        0                    0                     0
      122    VA 2022-12-25        0                    0                     0
      123    VA 2023-01-01        0                    0                     0

      [124 rows x 5 columns]
```

```
[47]: # plotting mode values for the 4 states.

      fig_5 = px.line(modeValues, x="Date", y=['Number of new cases','Number of new␣
       ↪Deaths'],color='State',title='Weekly mode values for 4 states')
      fig_5.show()
```

```
[48]: # Reading population data and filtering it.


      population = pd.read_csv("../../DATASETS/COVID DATASETS/
       ↪covid_county_population_usafacts.csv")
      population = population[population['County Name']!='Statewide Unallocated']
```

```python
population=population.drop(['County Name','State'],axis=1)
population
```

[48]:

```
      countyFIPS  population
1           1001       55869
2           1003      223234
3           1005       24686
4           1007       22394
5           1009       57826
…          …         …
3190       56037       42343
3191       56039       23464
3192       56041       20226
3193       56043        7805
3194       56045        6927

[3145 rows x 2 columns]
```

[49]:

```python
# merging population with new cases and deaths data.

weeklyCombinedNew = pd.merge(weeklyCombined,population, on ='countyFIPS')
```

[50]:

```python
weeklyCombinedNew
```

[50]:

```
       countyFIPS      County Name State  StateFIPS       Date  \
0            1001   Autauga County    AL          1 2022-06-05
1            1001   Autauga County    AL          1 2022-06-12
2            1001   Autauga County    AL          1 2022-06-19
3            1001   Autauga County    AL          1 2022-06-26
4            1001   Autauga County    AL          1 2022-07-03
…           …              …   …        …        …
97397       56045    Weston County    WY         56 2022-12-04
97398       56045    Weston County    WY         56 2022-12-11
97399       56045    Weston County    WY         56 2022-12-18
97400       56045    Weston County    WY         56 2022-12-25
97401       56045    Weston County    WY         56 2023-01-01

       Number of new cases  Number of new Deaths  population
0                       48                     0       55869
1                      117                     1       55869
2                      116                     0       55869
3                      189                     0       55869
4                      162                     0       55869
…                    …                   …         …
97397                   11                     0        6927
97398                    2                     0        6927
97399                    9                     0        6927
```

```
97400                     5                      0       6927
97401                     0                      0       6927

[97402 rows x 8 columns]
```

[51]:
```
# choosing states for comparing normalised weekly cases and deaths values.

normalisedCompare = weeklyCombinedNew[weeklyCombinedNew['State'].
 ↪isin(['NC','AR','AZ'])]
normalisedCompare
```

[51]:
```
       countyFIPS     County Name State  StateFIPS       Date  \
2976         4001  Apache County    AZ          4 2022-06-05
2977         4001  Apache County    AZ          4 2022-06-12
2978         4001  Apache County    AZ          4 2022-06-19
2979         4001  Apache County    AZ          4 2022-06-26
2980         4001  Apache County    AZ          4 2022-07-03
…               …             …   …          …          …
61685       37199  Yancey County    NC         37 2022-12-04
61686       37199  Yancey County    NC         37 2022-12-11
61687       37199  Yancey County    NC         37 2022-12-18
61688       37199  Yancey County    NC         37 2022-12-25
61689       37199  Yancey County    NC         37 2023-01-01

       Number of new cases  Number of new Deaths  population
2976                    80                     0       71887
2977                   174                     0       71887
2978                   225                     0       71887
2979                   258                     0       71887
2980                   294                     0       71887
…                      …                     …         …
61685                    5                     1       18069
61686                   13                     1       18069
61687                   14                     0       18069
61688                   12                     0       18069
61689                    0                     0       18069

[5890 rows x 8 columns]
```

[52]:
```
# grouping them according to the state and date.

normalisedCompare = normalisedCompare.groupby(['State','Date']).sum().
 ↪reset_index()
```

[53]:
```
normalisedCompare
```

```
[53]:    State        Date  countyFIPS  StateFIPS  Number of new cases  \
      0     AR 2022-06-05      380625        375                 2645
      1     AR 2022-06-12      380625        375                 3837
      2     AR 2022-06-19      380625        375                 4696
      3     AR 2022-06-26      380625        375                 5442
      4     AR 2022-07-03      380625        375                 6958
      ..   …          …           …        …                  …
      88    NC 2022-12-04     3710000       3700                 6842
      89    NC 2022-12-11     3710000       3700                11605
      90    NC 2022-12-18     3710000       3700                13190
      91    NC 2022-12-25     3710000       3700                16676
      92    NC 2023-01-01     3710000       3700                    0

          Number of new Deaths  population
      0                     27     3017804
      1                     14     3017804
      2                     21     3017804
      3                     31     3017804
      4                     21     3017804
      ..                   …           …
      88                    24    10488084
      89                   162    10488084
      90                    47    10488084
      91                    49    10488084
      92                     0    10488084

      [93 rows x 7 columns]
```

```
[54]: # calculating normalised cases and normalised deaths according to population.

      normalisedCompare['Normalised Cases'] = (normalisedCompare['Number of new␣
       ↪cases']/normalisedCompare['population'])*100000
      normalisedCompare['Normalised Deaths'] = (normalisedCompare['Number of new␣
       ↪Deaths']/normalisedCompare['population'])*100000
```

```
[55]: normalisedCompare
```

```
[55]:    State        Date  countyFIPS  StateFIPS  Number of new cases  \
      0     AR 2022-06-05      380625        375                 2645
      1     AR 2022-06-12      380625        375                 3837
      2     AR 2022-06-19      380625        375                 4696
      3     AR 2022-06-26      380625        375                 5442
      4     AR 2022-07-03      380625        375                 6958
      ..   …          …           …        …                  …
      88    NC 2022-12-04     3710000       3700                 6842
      89    NC 2022-12-11     3710000       3700                11605
      90    NC 2022-12-18     3710000       3700                13190
```

```
91     NC 2022-12-25      3710000       3700                   16676
92     NC 2023-01-01      3710000       3700                       0

    Number of new Deaths  population  Normalised Cases  Normalised Deaths
0                     27     3017804         87.646514           0.894690
1                     14     3017804        127.145434           0.463913
2                     21     3017804        155.609841           0.695870
3                     31     3017804        180.329803           1.027237
4                     21     3017804        230.565007           0.695870
..                   …           …                …                  …
88                    24    10488084         65.235938           0.228831
89                   162    10488084        110.649381           1.544610
90                    47    10488084        125.761769           0.448128
91                    49    10488084        158.999489           0.467197
92                     0    10488084          0.000000           0.000000

[93 rows x 9 columns]
```

[56]:
```
# plotting normalised cases and normalised deaths for the states.

fig_6 = px.line(normalisedCompare,x='Date',y=['Normalised Cases','Normalised␣
 ↪Deaths'],color='State',title='Log Normalised Graph for States')
fig_6.show()
```

[57]:
```
UsPattern = weeklyCombinedNew.groupby('Date')['Number of new cases','Number of␣
 ↪new Deaths','population'].sum().reset_index()
UsPattern
```

```
C:\Users\venka\AppData\Local\Temp\ipykernel_18648\1169135184.py:1:
FutureWarning:

Indexing with multiple keys (implicitly converted to a tuple of keys) will be
deprecated, use a list instead.
```

[57]:
```
          Date  Number of new cases  Number of new Deaths  population
0   2022-06-05               596616                  1226   328239523
1   2022-06-12               607086                  2062   328239523
2   2022-06-19               720738                  1771   328239523
3   2022-06-26               524811                  1540   328239523
4   2022-07-03               715190                  2478   328239523
5   2022-07-10               573480                  5225   328239523
6   2022-07-17               752889                  1220   328239523
7   2022-07-24               647246                  1831   328239523
8   2022-07-31               789033                  1901   328239523
9   2022-08-07               584134                  2346   328239523
10  2022-08-14               638027                  2125   328239523
```

```
11 2022-08-21               493742                  1962   328239523
12 2022-08-28               386887                  -948   328239523
13 2022-09-04               -26293                 -7218   328239523
14 2022-09-11               510896                  1958   328239523
15 2022-09-18               265845                  1947   328239523
16 2022-09-25               360031                  2107   328239523
17 2022-10-02               227630                  1569   328239523
18 2022-10-09               245389                  2171   328239523
19 2022-10-16               229679                  1872   328239523
20 2022-10-23               177433                  1551   328239523
21 2022-10-30               217652                  1803   328239523
22 2022-11-06               220576                  1484   328239523
23 2022-11-13               196964                   424   328239523
24 2022-11-20               222255                  1357   328239523
25 2022-11-27               210368                  1549   328239523
26 2022-12-04               310649                  2153   328239523
27 2022-12-11               320264                  1962   328239523
28 2022-12-18               326894                  1681   328239523
29 2022-12-25               299191                  1027   328239523
30 2023-01-01               241219                  2031   328239523
```

[58]:

```
UsPattern['Normalised Cases'] = (UsPattern['Number of new cases']/
 ↪UsPattern['population'])*100000
UsPattern['Normalised Deaths'] = (UsPattern['Number of new Deaths']/
 ↪UsPattern['population'])*100000
```

[59]:

```
fig_7 = px.line(UsPattern,x='Date',y=['Normalised Cases','Normalised␣
 ↪Deaths'],title='Log Normalised Graph for US')
fig_7.show()
```

Peaks of the graph: —> AR state Normalised cases peaked at july 17, 2022 with a value of 307.2433 Normalised deaths peaked at oct 09, 2022 with a value of 3.678171 —> AZ state Normalised cases peaked at july 24, 2022 with a value of 290.3671 Normalised deaths value is always 0 —> NC state Normalised cases peaked at july 31, 2022 with a value of 323.5291 Normalised deaths peaked at sep 4, 2022 with a value of 4.691038

The rates differ with each state because because of the population in state. Since NC is a popupalar growing state in terms of every field population got added in the recent years so the cases rate is higher compared to the other two states.Even though AR has less population it has high cases rate because of tourist spots are high populated in july. In the similar way deaths rates also differ with NC has highest death rate, followed by AR and then AZ.The cases are higher in summer as people roam on roads in summer. Deaths are high in winter due cold temperatures.

The pattern is almost similar with the US pattern like the cases are up to september 2022 and there is a decrease in september 2022 and after a slight increase, then they are almost constant with few fluctuations.

[60]:
```
# choosing a state to find the top 3 counties with highest cases and death␣
↪rates.

virginiaState = weeklyCombinedNew[weeklyCombinedNew['State']=='VA']
virginiaState
```

[60]:
```
       countyFIPS         County Name State  StateFIPS       Date  \
87420       51001     Accomack County    VA         51 2022-06-05
87421       51001     Accomack County    VA         51 2022-06-12
87422       51001     Accomack County    VA         51 2022-06-19
87423       51001     Accomack County    VA         51 2022-06-26
87424       51001     Accomack County    VA         51 2022-07-03
…             …                   …   …       …        …
91538       51840  City of Winchester    VA         51 2022-12-04
91539       51840  City of Winchester    VA         51 2022-12-11
91540       51840  City of Winchester    VA         51 2022-12-18
91541       51840  City of Winchester    VA         51 2022-12-25
91542       51840  City of Winchester    VA         51 2023-01-01

       Number of new cases  Number of new Deaths  population
87420                   39                     0       32316
87421                   63                     0       32316
87422                   55                     0       32316
87423                   83                     1       32316
87424                   58                     0       32316
…                    …                     …       …
91538                   74                     1       28078
91539                   53                     0       28078
91540                   46                     1       28078
91541                   41                     1       28078
91542                   36                     1       28078

[4123 rows x 8 columns]
```

[61]:
```
# sorting the values to find the top 3 counties.

virginiaState.groupby('countyFIPS').sum().sort_values(by=['Number of new␣
↪cases','Number of new Deaths'])
```

[61]:
```
            StateFIPS  Number of new cases  Number of new Deaths  population
countyFIPS
51678            1581                 -305                    -6      230826
51750            1581                 -206                    -4      565719
51091            1581                   82                    -1       67890
51017            1581                  128                     0      128557
51530            1581                  175                     3      200818
…                …                   …                     …       …
```

```
51810           1581             18078                     90    13949194
51041           1581             18244                     99    10936862
51107           1581             19152                     44    12819678
51153           1581             22203                     56    14580385
51059           1581             55369                    198    35573492

[133 rows x 4 columns]
```

[62]:

```
# since it is in the ascending order, so choosing the last three counties.

countyCompare = virginiaState[virginiaState['countyFIPS'].
 ↪isin([51107,51153,51059])]
countyCompare
```

[62]:

```
       countyFIPS            County Name State  StateFIPS        Date  \
88288       51059         Fairfax County    VA         51  2022-06-05
88289       51059         Fairfax County    VA         51  2022-06-12
88290       51059         Fairfax County    VA         51  2022-06-19
88291       51059         Fairfax County    VA         51  2022-06-26
88292       51059         Fairfax County    VA         51  2022-07-03
…           …                      …   …         …          …
89678       51153  Prince William County    VA         51  2022-12-04
89679       51153  Prince William County    VA         51  2022-12-11
89680       51153  Prince William County    VA         51  2022-12-18
89681       51153  Prince William County    VA         51  2022-12-25
89682       51153  Prince William County    VA         51  2023-01-01

       Number of new cases  Number of new Deaths  population
88288                 3265                     4     1147532
88289                 3207                     1     1147532
88290                 2718                     2     1147532
88291                 2899                     5     1147532
88292                 2881                     6     1147532
…                      …                     …         …
89678                  565                     2      470335
89679                  610                     1      470335
89680                  757                     1      470335
89681                  938                     1      470335
89682                  436                     1      470335

[93 rows x 8 columns]
```

[63]:

```
# plotting raw values of weekly cases and weekly deaths for the 3 counties.

fig_8 = px.line(countyCompare,x="Date",y=['Number of new cases','Number of new␣
 ↪Deaths'],color='County Name',title='Raw Graph for counties')
fig_8.show()
```

Fairfax County has high population during the last 6 months of 2022 so the new cases rate is higher than other two counties. Similarly the population difference between Loudoun County and Prince William County is very small compared to Fairfax County so the new case rates is almost simialr with Prince William County has higher rates then Loudoun County as the population is a little bit higher. Similarly deaths rates follow the same pattern. The cases are higher in summer as people roam on roads in summer. Deaths are high in winter due cold temperatures.

Peaks: —> Fairfax County New cases peaked at june 5, 2022 with a value of 3265. New deaths peaked at Sep 11,2022 with a value of 18 —> Loudoun County New cases peaked at June 12, 2022 with a value of 1247 New deaths peaked at Nov 27, 2022 with a value of 2 —> Prince William County New cases peaked at June 5, 2022 with a value of 1219 New deaths peaked at Aug 7 and Nov 27,2022 with a value of 4

The counties follow the state pattern of virginia, we can see the similarity from mean value graph of virginia state from fig_1.There is a dip cases from october to december 2022. The graph looks similar.

[64]:
```
# Calculating the log normalised cases and deaths for the 3 counties.

countyCompare['Log Normalised Cases'] = np.log((countyCompare['Number of new␣
 ↪cases']/countyCompare['population'])*100000)
countyCompare['Log Normalised Deaths'] = np.log((countyCompare['Number of new␣
 ↪Deaths']/countyCompare['population'])*100000)
```

```
C:\Users\venka\AppData\Local\Temp\ipykernel_18648\3630481622.py:3:
SettingWithCopyWarning:


A value is trying to be set on a copy of a slice from a DataFrame.
Try using .loc[row_indexer,col_indexer] = value instead

See the caveats in the documentation: https://pandas.pydata.org/pandas-
docs/stable/user_guide/indexing.html#returning-a-view-versus-a-copy

C:\Users\venka\anaconda3\lib\site-packages\pandas\core\arraylike.py:397:
RuntimeWarning:

divide by zero encountered in log

C:\Users\venka\anaconda3\lib\site-packages\pandas\core\arraylike.py:397:
RuntimeWarning:

invalid value encountered in log

C:\Users\venka\AppData\Local\Temp\ipykernel_18648\3630481622.py:4:
SettingWithCopyWarning:


A value is trying to be set on a copy of a slice from a DataFrame.
```

```
Try using .loc[row_indexer,col_indexer] = value instead

See the caveats in the documentation: https://pandas.pydata.org/pandas-docs/stable/user_guide/indexing.html#returning-a-view-versus-a-copy
```

[65]:

```
countyCompare
```

[65]:

```
       countyFIPS            County Name State  StateFIPS       Date  \
88288       51059         Fairfax County     VA         51 2022-06-05
88289       51059         Fairfax County     VA         51 2022-06-12
88290       51059         Fairfax County     VA         51 2022-06-19
88291       51059         Fairfax County     VA         51 2022-06-26
88292       51059         Fairfax County     VA         51 2022-07-03
…           …                      …   …         …         …
89678       51153  Prince William County     VA         51 2022-12-04
89679       51153  Prince William County     VA         51 2022-12-11
89680       51153  Prince William County     VA         51 2022-12-18
89681       51153  Prince William County     VA         51 2022-12-25
89682       51153  Prince William County     VA         51 2023-01-01

       Number of new cases  Number of new Deaths  population  \
88288                 3265                     4     1147532
88289                 3207                     1     1147532
88290                 2718                     2     1147532
88291                 2899                     5     1147532
88292                 2881                     6     1147532
…                        …                   …         …
89678                  565                     2      470335
89679                  610                     1      470335
89680                  757                     1      470335
89681                  938                     1      470335
89682                  436                     1      470335

       Log Normalised Cases  Log Normalised Deaths
88288              5.650816              -1.053904
88289              5.632893              -2.440199
88290              5.467453              -1.747051
88291              5.531922              -0.830761
88292              5.525694              -0.648439
…                         …                    …
89678              4.788551              -0.855128
89679              4.865184              -1.548275
89680              5.081088              -1.548275
89681              5.295475              -1.548275
89682              4.529367              -1.548275
```

```
[93 rows x 10 columns]
```

[66]:

```
# plotting the log normalised cases and deaths for the 3 counties.

fig_9 = px.line(countyCompare,x="Date",y=['Log Normalised Cases','Log␣
 ↪Normalised Deaths'],color='County Name',title='Log Normalised Graph for␣
 ↪counties')
fig_9.show()
```

References: 1) https://www.geeksforgeeks.org/line-chart-using-plotly-in-python/

2) https://pandas.pydata.org/pandas-docs/stable/reference/api/pandas.melt.html

3) https://pandas.pydata.org/docs/reference/api/pandas.Grouper.html