

Experiment 22

Design a Python script to generate statistical reports (Minimum, Maximum, Count, Average, Sum etc) on public datasets.

```
In [1]: import pandas as pd
data=pd.read_csv("ebola_2014_2016_clean.csv")
```

```
In [2]: data.info()

<class 'pandas.core.frame.DataFrame'>
RangeIndex: 2485 entries, 0 to 2484
Data columns (total 4 columns):
#   Column                                                                 Non-Null Count  Dtype  
---  -
0   Country                                                                2485 non-null   object  
1   Date                                                                    2485 non-null   object  
2   Cumulative no. of confirmed, probable and suspected cases            2477 non-null   float64  
3   Cumulative no. of confirmed, probable and suspected deaths           2485 non-null   float64  
dtypes: float64(2), object(2)
memory usage: 77.8+ KB
```

```
In [3]: data.dtypes
```

```
Out[3]: Country                object
Date                          object
Cumulative no. of confirmed, probable and suspected cases  float64
Cumulative no. of confirmed, probable and suspected deaths  float64
dtype: object
```

```
In [4]: data.head()
```

```
Out[4]:
```

	Country	Date	Cumulative no. of confirmed, probable and suspected cases	Cumulative no. of confirmed, probable and suspected deaths
0	Guinea	2014-08-29	648.0	430.0
1	Nigeria	2014-08-29	19.0	7.0
2	Sierra Leone	2014-08-29	1026.0	422.0
3	Liberia	2014-08-29	1378.0	694.0
4	Sierra Leone	2014-09-05	1261.0	491.0

```
In [5]: data.tail()
```

```
Out[5]:
```

	Country	Date	Cumulative no. of confirmed, probable and suspected cases	Cumulative no. of confirmed, probable and suspected deaths
2480	Liberia	2016-03-23	10666.0	4806.0
2481	Italy	2016-03-23	1.0	0.0
2482	Liberia	2016-03-23	5.0	4.0
2483	Nigeria	2016-03-23	20.0	8.0
2484	United States of America	2016-03-23	4.0	1.0

```
In [6]: data.columns
```

```
Out[6]: Index(['Country', 'Date',
              'Cumulative no. of confirmed, probable and suspected cases',
              'Cumulative no. of confirmed, probable and suspected deaths'],
              dtype='object')
```

```
In [7]: data.items
```

```
Out[7]: <bound method DataFrame.items of
0          Guinea  2014-08-29
1          Nigeria  2014-08-29
2      Sierra Leone  2014-08-29
3          Liberia  2014-08-29
4      Sierra Leone  2014-09-05
...
2480          Liberia  2016-03-23
2481          Italy  2016-03-23
2482          Liberia  2016-03-23
2483          Nigeria  2016-03-23
2484  United States of America  2016-03-23

      Cumulative no. of confirmed, probable and suspected cases \
0          648.0
1          19.0
2         1026.0
3         1378.0
4         1261.0
...
2480        10666.0
2481          1.0
2482          5.0
2483         20.0
2484          4.0

      Cumulative no. of confirmed, probable and suspected deaths
0          430.0
1           7.0
2         422.0
3         694.0
4         491.0
...
2480        4806.0
```

```

2481
2482
2483
2484
0.0
4.0
8.0
1.0

```

```
[2485 rows x 4 columns]>
```

```
In [8]: data.describe()
```

```

Out[8]:
Cumulative no. of confirmed, probable and suspected cases
Cumulative no. of confirmed, probable and suspected deaths
count      2477.000000      2485.000000
mean      2553.678644      1028.347686
std       4427.118148      1656.064372
min         0.000000         0.000000
25%         1.000000         0.000000
50%         8.000000         6.000000
75%        3657.000000      2386.000000
max       14122.000000      4806.000000

```

```
In [9]: data.count(axis=0)
```

```

Out[9]: Country      2485
Date      2485
Cumulative no. of confirmed, probable and suspected cases      2477
Cumulative no. of confirmed, probable and suspected deaths      2485
dtype: int64

```

```
In [10]: data.count(axis=1)
```

```

Out[10]: 0      4
1      4
2      4
3      4
4      4
..
2480    4
2481    4
2482    4
2483    4
2484    4
Length: 2485, dtype: int64

```

```
In [11]: data.sum(axis=0)
```

```

Out[11]: Country      GuineaNigeriaSierra LeoneL
iberiaSierra LeoneNi...
Date      2014-08-292014-08-292014-0
8-292014-08-292014-0...
Cumulative no. of confirmed, probable and suspected cases
6.32546e+06

```

```
Cumulative no. of confirmed, probable and suspected deaths
2.55544e+06
dtype: object
```

```
In [12]: data['Cumulative no. of confirmed, probable and suspected cases'].sum()
```

```
Out[12]: 6325462.0
```

```
In [13]: data.median()
```

```
Out[13]: Cumulative no. of confirmed, probable and suspected cases    8.0
Cumulative no. of confirmed, probable and suspected deaths    6.0
dtype: float64
```

```
In [14]: data.mean()
```

```
Out[14]: Cumulative no. of confirmed, probable and suspected cases    2553.678644
Cumulative no. of confirmed, probable and suspected deaths    1028.347686
dtype: float64
```

```
In [15]: data['Cumulative no. of confirmed, probable and suspected deaths'].mean()
```

```
Out[15]: 1028.3476861167003
```

```
In [16]: data.sum()
```

```
Out[16]: Country                                GuineaNigeriaSierra LeoneL
iberiaSierra LeoneNi...
Date                                2014-08-292014-08-292014-0
8-292014-08-292014-0...
Cumulative no. of confirmed, probable and suspected cases
6.32546e+06
Cumulative no. of confirmed, probable and suspected deaths
2.55544e+06
dtype: object
```

```
In [17]: data['Cumulative no. of confirmed, probable and suspected deaths'].sum()
```

```
Out[17]: 2555444.0
```

```
In [18]: data.cummin()
```

```
Out[18]:
```

	Country	Date	Cumulative no. of confirmed, probable and suspected cases	Cumulative no. of confirmed, probable and suspected deaths
0	Guinea	2014-08-29	648.0	430.0
1	Guinea	2014-08-29	19.0	7.0
2	Guinea	2014-08-29	19.0	7.0

	Country	Date	Cumulative no. of confirmed, probable and suspected cases	Cumulative no. of confirmed, probable and suspected deaths
3	Guinea	2014-08-29	19.0	7.0
4	Guinea	2014-08-29	19.0	7.0
...
2480	Guinea	2014-08-29	0.0	0.0
2481	Guinea	2014-08-29	0.0	0.0
2482	Guinea	2014-08-29	0.0	0.0
2483	Guinea	2014-08-29	0.0	0.0
2484	Guinea	2014-08-29	0.0	0.0

2485 rows × 4 columns

```
In [19]: data['Cumulative no. of confirmed, probable and suspected cases'].cumprod()
```

```
Out[19]: 0      6.480000e+02
1      1.231200e+04
2      1.263211e+07
3      1.740705e+10
4      2.195029e+13
      ...
2480      NaN
2481      NaN
2482      NaN
2483      NaN
2484      NaN
Name: Cumulative no. of confirmed, probable and suspected cases, Length: 2485, dtype: float64
```

```
In [20]: data.cummin()
```

```
Out[20]:
```

	Country	Date	Cumulative no. of confirmed, probable and suspected cases	Cumulative no. of confirmed, probable and suspected deaths
0	Guinea	2014-08-29	648.0	430.0
1	Guinea	2014-08-29	19.0	7.0
2	Guinea	2014-08-29	19.0	7.0
3	Guinea	2014-08-29	19.0	7.0

	Country	Date	Cumulative no. of confirmed, probable and suspected cases	Cumulative no. of confirmed, probable and suspected deaths
4	Guinea	2014-08-29	19.0	7.0
...
2480	Guinea	2014-08-29	0.0	0.0
2481	Guinea	2014-08-29	0.0	0.0
2482	Guinea	2014-08-29	0.0	0.0
2483	Guinea	2014-08-29	0.0	0.0
2484	Guinea	2014-08-29	0.0	0.0

2485 rows × 4 columns

In [21]:

```
data.cummax()
```

Out[21]:

	Country	Date	Cumulative no. of confirmed, probable and suspected cases	Cumulative no. of confirmed, probable and suspected deaths
0	Guinea	2014-08-29	648.0	430.0
1	Nigeria	2014-08-29	648.0	430.0
2	Sierra Leone	2014-08-29	1026.0	430.0
3	Sierra Leone	2014-08-29	1378.0	694.0
4	Sierra Leone	2014-09-05	1378.0	694.0
...
2480	United States of America	2016-03-23	14122.0	4806.0
2481	United States of America	2016-03-23	14122.0	4806.0
2482	United States of America	2016-03-23	14122.0	4806.0
2483	United States of America	2016-03-23	14122.0	4806.0
2484	United States of America	2016-03-23	14122.0	4806.0

2485 rows × 4 columns

```
In [22]: data['Cumulative no. of confirmed, probable and suspected cases'].pct_change()
```

```
Out[22]: 0          NaN
1      -0.970679
2      53.000000
3       0.343080
4      -0.084906
...
2480    1184.111111
2481     -0.999906
2482      4.000000
2483      3.000000
2484     -0.800000
Name: Cumulative no. of confirmed, probable and suspected cases, Length: 2485, dtype: float64
```

By

Harsha Praneeth D

122010407003

ECE VLSI