In [31]:
```
import pandas as pd
import numpy as np
```

In [32]:
```
df=pd.read_csv('D:\Project\Covid 19\StatewiseTestingDetails.csv')
```

In [33]:
```
df.shape
```

Out[33]:
```
(1923, 5)
```

In [34]:
```
df.describe()
```

Out[34]:

	TotalSamples	Negative	Positive
count	1923.000000	1485.000000	1915.000000
mean	70284.702028	69458.684175	2913.050653
std	108860.912783	109976.196257	8259.917376
min	58.000000	0.000000	0.000000
25%	4732.000000	4505.000000	37.000000
50%	20877.000000	19316.000000	365.000000
75%	89046.500000	82356.000000	2102.000000
max	638846.000000	601363.000000	90787.000000

In [35]:
```
df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 1923 entries, 0 to 1922
Data columns (total 5 columns):
 #   Column        Non-Null Count  Dtype  
---  ------        --------------  -----  
 0   Date          1923 non-null   object 
 1   State         1923 non-null   object 
 2   TotalSamples  1923 non-null   float64
 3   Negative      1485 non-null   float64
 4   Positive      1915 non-null   float64
dtypes: float64(3), object(2)
memory usage: 60.2+ KB
```

In [36]:
```
df.isnull().sum()
```

Out[36]:
```
Date             0
State            0
TotalSamples     0
Negative       438
Positive         8
dtype: int64
```

In [37]:
```
df.fillna(0,inplace=True)
```

In [38]:
```
df.isnull().sum()
```

Out[38]:
```
Date           0
State          0
TotalSamples   0
Negative       0
Positive       0
dtype: int64
```

In [39]:
```
df.head(5)
```

Out[39]:

	Date	State	TotalSamples	Negative	Positive
0	2020-04-17	Andaman and Nicobar Islands	1403.0	1210.0	12.0
1	2020-04-24	Andaman and Nicobar Islands	2679.0	0.0	27.0
2	2020-04-27	Andaman and Nicobar Islands	2848.0	0.0	33.0
3	2020-05-01	Andaman and Nicobar Islands	3754.0	0.0	33.0
4	2020-05-16	Andaman and Nicobar Islands	6677.0	0.0	33.0

In [41]:
```
path=('D:\Project\Covid 19\StatewiseTestingDetails.csv')
df.to_csv(path)
```