In [45]:

```
import pandas as pd
import numpy as np
```

In [ ]:

```
pip install statsmodels
```

In [ ]:

```
from statsmodels.stats import weightstats as stests
```

In [ ]:

```
from scipy.stats import f_oneway
from scipy import stats
```

In [46]:

```
pd.set_option('display.max_columns',80)
df1 = pd.read_csv('data covid-19 - Sheet1.csv')
```

In [47]:

```
df1
```

Out[47]:

	#	Country,	Total	New	Total.1	New.1	Total.2	Active	Serious,	Tot Cases/	Deaths/	
0	NaN	Other	Cases	Cases	Deaths	Deaths	Recovered	Cases	Critical	1M pop	1M pop	
1	World	64,998,913	168,709	1,502,357	3,979	45,092,256	18,404,300	106,693	8,339	192.7	NaN	
2	1	USA	14,314,265	NaN	279,867	NaN	8,462,434	5,571,964	26,187	43,138	843	199,
3	2	India	9,534,964	1,493	138,657	NaN	8,973,373	422,934	8,944	6,881	100	143,
4	3	Brazil	6,436,650	NaN	174,531	NaN	5,698,353	563,766	8,318	30,191	819	25,
...	...	...	...	...	...	...	...	...	...	...	...	...
218	217	Marshall Islands	4	NaN	NaN	NaN	4	0	NaN	67	NaN	
219	218	Wallis and Futuna	3	NaN	NaN	NaN	1	2	NaN	269	NaN	
220	219	Samoa	2	NaN	NaN	NaN	NaN	2	NaN	10	NaN	
221	220	Vanuatu	1	NaN	NaN	NaN	1	0	NaN	3	NaN	
222	NaN	Total:	64,998,913	168,709	1,502,357	3,979	45,092,256	18,404,300	106,693	8,338.80	192.7	

223 rows × 14 columns

In [48]:

```
df2=df1[1:]
df2
```

Out[48]:

	#	Country,	Total	New	Total.1	New.1	Total.2	Active	Serious,	Tot Cases/	Deaths/	
1	World	64,998,913	168,709	1,502,357	3,979	45,092,256	18,404,300	106,693	8,339	192.7	NaN	
2	1	USA	14,314,265	NaN	279,867	NaN	8,462,434	5,571,964	26,187	43,138	843	199,5
3	2	India	9,534,964	1,493	138,657	NaN	8,973,373	422,934	8,944	6,881	100	143,5

	#	Country,	Total	New	Total.1	New.1	Total.2	Active	Serious,	Tot Cases/	Deaths/	
4	3	Brazil	6,436,650	NaN	174,531	NaN	5,698,353	563,766	8,318	30,191	819	25,7
5	4	Russia	2,375,546	28,145	41,607	554	1,859,851	474,088	2,300	16,275	285	77,6
...	...	...	...	...	...	...	...	...	...	...	...	...
218	217	Marshall Islands	4	NaN	NaN	NaN	4	0	NaN	67	NaN	
219	218	Wallis and Futuna	3	NaN	NaN	NaN	1	2	NaN	269	NaN	
220	219	Samoa	2	NaN	NaN	NaN	NaN	2	NaN	10	NaN	
221	220	Vanuatu	1	NaN	NaN	NaN	1	0	NaN	3	NaN	
222	NaN	Total:	64,998,913	168,709	1,502,357	3,979	45,092,256	18,404,300	106,693	8,338.80	192.7	

222 rows × 14 columns

In [51]:

```
df3=df2.drop(['#'],axis=1)
df3
```

Out[51]:

	Country,	Total	New	Total.1	New.1	Total.2	Active	Serious,	Tot Cases/	Deaths/	Total.3
1	64,998,913	168,709	1,502,357	3,979	45,092,256	18,404,300	106,693	8,339	192.7	NaN	NaN
2	USA	14,314,265	NaN	279,867	NaN	8,462,434	5,571,964	26,187	43,138	843	199,906,513
3	India	9,534,964	1,493	138,657	NaN	8,973,373	422,934	8,944	6,881	100	143,557,647
4	Brazil	6,436,650	NaN	174,531	NaN	5,698,353	563,766	8,318	30,191	819	25,700,000
5	Russia	2,375,546	28,145	41,607	554	1,859,851	474,088	2,300	16,275	285	77,693,654
...	...	...	...	...	...	...	...	...	...	...	...
218	Marshall Islands	4	NaN	NaN	NaN	4	0	NaN	67	NaN	NaN
219	Wallis and Futuna	3	NaN	NaN	NaN	1	2	NaN	269	NaN	1,149
220	Samoa	2	NaN	NaN	NaN	NaN	2	NaN	10	NaN	NaN
221	Vanuatu	1	NaN	NaN	NaN	1	0	NaN	3	NaN	NaN
222	Total:	64,998,913	168,709	1,502,357	3,979	45,092,256	18,404,300	106,693	8,338.80	192.7	NaN

222 rows × 13 columns

In [53]:

```
for i in df3.columns:
    if (i!='Country,'):
        df3[i]=df3[i].str.replace(',','')
```

In [54]:

```
for i in df3.columns:
    if (i!='Country,'):
        df3[i] = df3[i].astype(float)
```

In [55]:

```
df3
```

Out[55]:

	Country,	Total	New	Total.1	New.1	Total.2	Active	Serious,	Tot Cases/	Deaths/	Tot
1	64,998,913	168709.0	1502357.0	3979.0	45092256.0	18404300.0	106693.0	8339.0	192.7	NaN	

	Country,	Total	New	Total.1	New.1	Total.2	Active	Serious,	Tot Cases/	Deaths/	To
2	USA	14314265.0	NaN	279867.0	NaN	8409464.0	5571961.0	96187.0	43138.0	842.0	1999065
3	India	9534964.0	1493.0	138657.0	NaN	8973373.0	422934.0	8944.0	6881.0	100.0	1435576
4	Brazil	6436650.0	NaN	174531.0	NaN	5698353.0	563766.0	8318.0	30191.0	819.0	257000
5	Russia	2375546.0	28145.0	41607.0	554.0	1859851.0	474088.0	2300.0	16275.0	285.0	776936
...	...	...	...	...	...	...	...	...	...	...	
218	Marshall Islands	4.0	NaN	NaN	NaN	4.0	0.0	NaN	67.0	NaN	I
219	Wallis and Futuna	3.0	NaN	NaN	NaN	1.0	2.0	NaN	269.0	NaN	11
220	Samoa	2.0	NaN	NaN	NaN	NaN	2.0	NaN	10.0	NaN	I
221	Vanuatu	1.0	NaN	NaN	NaN	1.0	0.0	NaN	3.0	NaN	I
222	Total:	64998913.0	168709.0	1502357.0	3979.0	45092256.0	18404300.0	106693.0	8338.8	192.7	I

222 rows × 13 columns

In [56]:

```
df4=df3.fillna(0)
```

In [57]:

```
ztest ,pval1 = stests.ztest(df4['Total'], x2=df4['Total.1'], value=0,alternative='two-sid
ed')
print(float(pval1))
if pval1<0.05:
    print("reject null hypothesis")
else:
    print("accept null hypothesis")
```

```
0.059221746469648465
accept null hypothesis
```

In [58]:

```
ztest ,pval1 = stests.ztest(df4['Total'], x2=df4['Total.2'], value=0,alternative='two-sid
ed')
print(float(pval1))
if pval1<0.05:
    print("reject null hypothesis")
else:
    print("accept null hypothesis")
```

```
0.7930067348834116
accept null hypothesis
```

In [59]:

```
ztest ,pval1 = stests.ztest(df4['Total.1'], x2=df4['Total.2'], value=0,alternative='two-s
ided')
print(float(pval1))
if pval1<0.05:
    print("reject null hypothesis")
else:
    print("accept null hypothesis")
```

```
0.036732288546715 09
reject null hypothesis
```

In [ ]:

```

```