

DSG

①

AKash Ranjan

1829036

CS&CE

Q1.- Country with highest recovery rate:-

Ans- Recovery Rate =  $\frac{\text{No. of Recovered}}{\text{No. of confirmed cases}} \times 100\%$

Recovery Rates:-

- i) Argentina = 42.55%
- ii) Belgium = 27.42%
- iii) Brazil = 68.64%
- iv) Ecuador = 43.56%
- v) Germany = 92.58%
- vi) India = 63.24%
- vii) Iran = 86.01%
- viii) Italy = 80.49%
- ix) USA = 30.74%
- x) Russia = 70.09%

∴ The country with highest recovery rate is Germany  
with 92.58% of recovery rate.

Akash

⑤

Q2.- Relationship bet<sup>n</sup> active cases & new cases

Ans- Active Cases =  $\frac{\text{Total Confirmed Cases} - \text{Recovered Cases}}{\text{Recovered Cases}}$

Country/ Region	Active Cases	New Cases
1. Argentina	63,848	4,236
2. Belgium	45,630	91
3. Brazil	6,16,650	39,924
4. Ecuador	39,688	759
5. Germany	14,890	434
6. India	3,56,089	32,676
7. Iran	37,000	23,590
8. Italy	47,490	162
9. USA	2,423,020	67,328
10. Russia	222,822	6,410

\* Active Cases & New Cases Vary linearly

akash

Q3.- From the above data what conclusion can you get about countries on how they handle covid 19 cases.

Ans- For controlling the covid 19 cases firstly Recovery Rate must be high for a country for this proper treatment facilities must be provided, Proper medicine, foods must be given. Secondly Case fatality Rate must be low means no. deaths must be low for this we need availability of beds in hospital, proper social distancing must be followed, etc.

Basic things countries must follow are:-

- i) Proper Lockdown Implementation
- ii) wearing of Mask must be kept compulsory
- iii) Proper Transportation facilities with proper Social Distancing
- iv) Awareness among people must be created
- v) Proper functioning for the poor ones
- vi) Digitization should be increased.
- vii) No. of tests must be increased
- viii) Proper functioning of Quarantine Centres
- ix) Gatherings must be prohibited.
- x) Proper hygiene maintenance (supply of mask, Sanitizer, Cleaning of hospitals, etc)

akash

4

Q4. - Can we get only satisfied with the recovery rates or we need another measurement about the specific countries if yes, then give which feature can be added and why?

Ans. Yes, Definitely we need another feature other than recovery rate because alone Recovery rate cannot determine everything. So, According to me Case fatality rate must be ~~introduced~~ introduced.

Case fatality Rate:-

It is the proportion of people who die from a specified disease among all individual diagnosed with the disease over a certain period of time.

Calculation of Case fatality Rate:-

$C(t)$ : No. of cases

$D(t)$ : No. of Deaths

Rate:-

$$D(t)/C(t)$$

According, to me it is also a very important parameter in order to know which country is performing better as lower the case fatality Rate more better the country is performing in this pandemic.

So, Two parameters

i) Recovery Rate:- More the Recovery Rate better the country in position

ii) Case fatality Rate:- Less the case fatality Rate better the country position.



Q5.- Insights about the health condition of countries from the data

(5)

Ans- Case fatality Rate of countries:-

1) Argentina = 1.84%

2) Belgium = 15.56%

3) Brazil = 3.83%

4) Ecuador = 7.33%

5) Germany = 4.51%

6) India = 2.57%

7) Iran = 5.06%

8) Italy = 14.37%

9) USA = 3.92%

10) Russia = 1.57%

\* As we can see Russia has the lowest Case fatality Rate.

Condition of Countries

i) Argentina:- Its recovery rate is not so good only 42.55% but its case fatality Rate is very less only 1.84%.

If I consider then Argentina health system is good only if its recovery rate is increased.

ii) Belgium:- Its Recovery rate is 27.42% which is really bad & its case fatality Rate is also very high 15.56%.

Hence, I can conclude Belgium health system is not good.

iii) Brazil:- Its Recovery rate is 68.64% which is really good & its case fatality Rate is 3.83% which is slightly high but its ok.

Hence, I can conclude Brazil health system is good.

⑥

iv) Ecuador: Its Recovery Rate is 43.56% which is not so good & its case fatality rate is 7.33% which is also pretty high.

Hence, I conclude its health condition is not in a good state.

v) Germany: Its recovery Rate is 92.58% which is extremely high & its case fatality rate is 4.51% which is slightly ~~high~~ high.

Hence, I conclude Germany has an excellent health condition.

vi) India: Its recovery Rate is 63.24% which is really good & its case fatality rate is

2.57% which is really good.

Hence, I conclude India has a ~~good~~ very good health condition.

vii) Iran: Its recovery rate is 86.01% which is really very very good & its case fatality Rate is 5.06% which is high.

Hence, I conclude Iran has a good health system.

viii) Italy: Its recovery rate is 80.49% which is really good & its case fatality rate is 14.37% which is extremely high.

Hence, I conclude Italy has not so good health system because of its high case fatality rate.

ix) USA: Its Recovery Rate is 30.74% which is extremely very low & its case fatality rate is 3.92% which is so good.

Hence, I conclude USA has not a good health system.

⑦

X) Russia: Its Recovery Rate is 70.09% which is good & its case fatality rate is 1.57% which is really very very good.  
Hence, I conclude It has a good health system.

Untitled3.ipynb - Colaboratory

colab.research.google.com/drive/1\_vY5aHqG9eQ8n40lZ\_JtHsNcicDcLcKe#scrollTo=aDSgaAL9tc8d

File Edit View Insert Runtime Tools Help [All changes saved](#)

Comment Share

Files

- ..
- sample\_data
- covid.csv

+ Code + Text

RAM Disk

Editing

## IMPORTING THE LIBRARIES

```
[1] import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
```

## ACCESSING THE DATASET

```
[7] dataset=pd.read_csv('covid.csv')
X=dataset.iloc[:, :-1].values
Y=dataset.iloc[:, -1].values
print(X)
print(Y)
```

```
[[ 63848]
 [ 45630]
 [ 616650]
 [ 39688]
 [ 14890]
 [ 356089]
 [ 37000]
```

Disk 76.83 GB available

Type here to search

01:22 21-07-2020



Untitled3.ipynb - Colaboratory

colab.research.google.com/drive/1\_vY5aHqG9eQ8n40lZ\_JtHsNcicDcLcKe#scrollTo=aDSgaAL9tc8d

File Edit View Insert Runtime Tools Help All changes saved

Comment Share

Files

- sample\_data
- covid.csv

+ Code + Text

```
print(Y)
```

```
[7] [[ 63848]
[ 45630]
[ 616650]
[ 39688]
[ 14890]
[ 356089]
[ 37000]
[ 47490]
[2423020]
[ 222822]]
[ 4236 91 39924 759 434 32676 23590 162 67328 6410]
```

RAM Disk

Editing

▼ SPLITTING THE DATASET INTO TEST SET AND TRAINING SET

```
[9] from sklearn.model_selection import train_test_split
X_train,X_test,Y_train,Y_test=train_test_split(X,Y,test_size=1/3,random_state=0)
print(X_train)
print(Y_train)
print(X_test)
print(Y_test)
```

Disk 76.83 GB available

Type here to search

01:23 21-07-2020

Untitled3.ipynb - Colaboratory

colab.research.google.com/drive/1\_vY5aHqG9eQ8n40lZ\_JtHsNcicDcLcKe#scrollTo=aDSgaAL9tc8d

Untitled3.ipynb

File Edit View Insert Runtime Tools Help All changes saved

Comment Share

RAM Disk

Editing

Files

..sample\_datacovid.csv

+

Code

+

Text

TRAINING THE SIMPLE REGRESSION MODEL ON THE TRAINING SET

```
[10] from sklearn.linear_model import LinearRegression
      regressor=LinearRegression()
      regressor.fit(X_train,Y_train)

      LinearRegression(copy_X=True, fit_intercept=True, n_jobs=None, normalize=False)
```

PREDICTING THE TEST SET RESULTS

```
[12] Y_pred=regressor.predict(X_test)
      print(Y_pred)

      [ 54043.50595611 206646.3064415  3206.60032765 20772.77199266]
```

VISUALIZING THE TRAINING SET RESULTS

```
[13] plt.scatter(X_train,Y_train,color="red")
      plt.plot(X_train,regressor.predict(X_train),color="blue")
```

Disk 76.83 GB available

Type here to search

01:23 21-07-2020

Untitled3.ipynb - Colaboratory

colab.research.google.com/drive/1\_vY5aHqG9eQ8n40lZ\_JtHsNcicDcLcKe#scrollTo=aDSgaAL9tc8d

Untitled3.ipynb

File Edit View Insert Runtime Tools Help All changes saved

Comment Share

RAM Disk

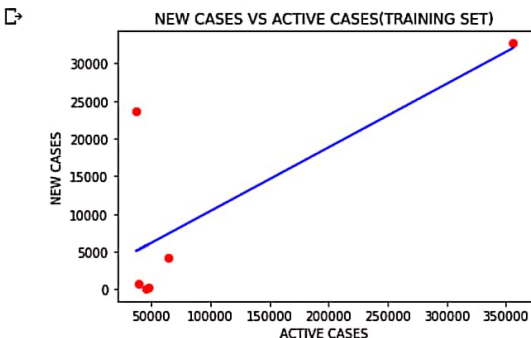
Editing

Files

- ..
- sample\_data
- covid.csv

### VISUALIZING THE TRAINING SET RESULTS

```
[13] plt.scatter(X_train,Y_train,color="red")
plt.plot(X_train,regressor.predict(X_train),color="blue")
plt.title("NEW CASES VS ACTIVE CASES(TRAINING SET)")
plt.xlabel("ACTIVE CASES")
plt.ylabel("NEW CASES")
plt.show()
```



Active Cases (X)	New Cases (Y)
50,000	24,000
75,000	4,000
100,000	1,000
125,000	1,000
350,000	32,000

Disk 76.83 GB available

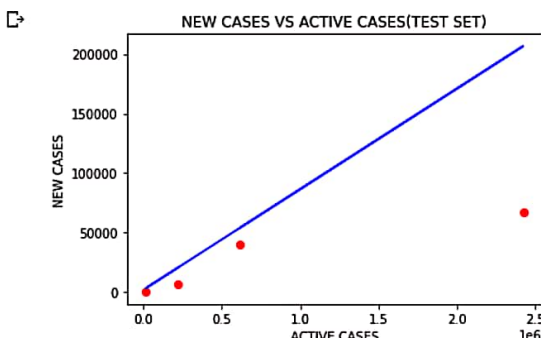
Type here to search

01:23 21-07-2020

Colaboratory interface showing a Jupyter Notebook titled "Untitled3.ipynb". The notebook is open to a code cell containing the following Python code:

```
[14] plt.scatter(X_test,Y_test,color="red")
plt.plot(X_test,Y_pred,color="blue")
plt.title("NEW CASES VS ACTIVE CASES(TEST SET)")
plt.xlabel("ACTIVE CASES")
plt.ylabel("NEW CASES")
plt.show()
```

The output of the code is a scatter plot titled "NEW CASES VS ACTIVE CASES(TEST SET)". The x-axis is labeled "ACTIVE CASES" and ranges from 0.0 to 2.5 (with a multiplier of  $10^6$ ). The y-axis is labeled "NEW CASES" and ranges from 0 to 200,000. The plot shows a positive linear correlation between active cases and new cases, with a blue line representing the predicted values and red dots representing the test set data points.



The interface also shows a file explorer on the left with a folder named "sample\_data" containing a file named "covid.csv". The bottom status bar indicates 76.83 GB available on the disk.



Menu



Home

Insert

Page Layout



Cut



Copy

Format  
Painter

Calibri

11

A<sup>+</sup>A<sup>-</sup>

≡

≡

≡

B

I

U



A



≡

≡

≡

B12

fx

A

B

C

D

E

	ACTIVE CASES	NEW CASES
1		
2	63848	4236
3	45630	91
4	616650	39924
5	39688	759
6	14890	434
7	356089	32676
8	37000	23590
9	47490	162
10	2423020	67328
11	222822	6410

--