

```
import pandas as pd
import numpy as np
import seaborn as sns
import matplotlib.pyplot as plt
from sklearn.model_selection import train_test_split
from sklearn.linear_model import LinearRegression
from sklearn.metrics import ConfusionMatrixDisplay
from sklearn.metrics import classification_report
from sklearn.metrics import accuracy_score
```

```
df=pd.read_csv('/content/Covid Data.csv')
```

```
df
```

	USMER	MEDICAL_UNIT	SEX	PATIENT_TYPE	DATE_DIED	INTUBED	PNEUMONIA	AGE	PREGNANT	DIABETES	...	ASTHMA	INMSUPR	HIPERTEI
0	2		1	1	1 03/05/2020	97	1	65	2	2	...	2	2	
1	2		1	2	1 03/06/2020	97	1	72	97	2	...	2	2	
2	2		1	2	2 09/06/2020	1	2	55	97	1	...	2	2	
3	2		1	1	1 12/06/2020	97	2	53	2	2	...	2	2	
4	2		1	2	1 21/06/2020	97	2	68	97	1	...	2	2	
...	
57965	2		4	1	2 25/07/2020	2	1	84	2	2	...	2	2	
57966	2		4	2	2 25/07/2020	2	1	49	97	1	...	2	2	
57967	1		4	2	2 25/07/2020	1	1	71	97	2	...	2	2	
57968	2		4	2	2 25/07/2020	2	1	68	97	1	...	2	2	
57969	1		4	2	2 25/07/2020	1	1	51	97	1	...	2	2	

57970 rows × 21 columns



```
df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 57970 entries, 0 to 57969
Data columns (total 21 columns):
#   Column              Non-Null Count  Dtype
---  -
0   USMER                57970 non-null  int64
1   MEDICAL_UNIT        57970 non-null  int64
2   SEX                 57970 non-null  int64
3   PATIENT_TYPE        57970 non-null  int64
4   DATE_DIED           57970 non-null  object
5   INTUBED             57970 non-null  int64
6   PNEUMONIA           57970 non-null  int64
7   AGE                 57970 non-null  int64
8   PREGNANT            57970 non-null  int64
9   DIABETES            57970 non-null  int64
10  COPD                57970 non-null  int64
11  ASTHMA              57970 non-null  int64
12  INMSUPR             57970 non-null  int64
13  HIPERTENSION        57970 non-null  int64
14  OTHER_DISEASE       57970 non-null  int64
15  CARDIOVASCULAR      57970 non-null  int64
16  OBESITY             57970 non-null  int64
17  RENAL_CHRONIC       57969 non-null   float64
18  TOBACCO             57969 non-null   float64
19  CLASIFFICATION_FINAL 57969 non-null   float64
20  ICU                 57969 non-null   float64
dtypes: float64(4), int64(16), object(1)
memory usage: 9.3+ MB
```

```
df.describe()
```



	USMER	MEDICAL_UNIT	SEX	PATIENT_TYPE	INTUBED	PNEUMO
count	57970.000000	57970.000000	57970.000000	57970.000000	57970.000000	57970.000
mean	1.448560	3.655580	1.580128	1.686096	33.118561	2.813
std	0.497351	0.497251	0.493542	0.464082	44.900113	11.089
min	1.000000	1.000000	1.000000	1.000000	1.000000	1.000
25%	1.000000	3.000000	1.000000	1.000000	1.000000	1.000
50%	1.000000	4.000000	2.000000	2.000000	2.000000	2.000

df.shape

(57970, 21)

df.AGE.mean()

55.71619803346559

df[df==0]

	USMER	MEDICAL_UNIT	SEX	PATIENT_TYPE	DATE_DIED	INTUBED	PNEUMONIA	AGE	PREGNANT	DIABETES	...	ASTHMA	INMSUPR	HIPERT
0	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	...	NaN	NaN	
1	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	...	NaN	NaN	
2	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	...	NaN	NaN	
3	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	...	NaN	NaN	
4	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	...	NaN	NaN	
...	
57965	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	...	NaN	NaN	
57966	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	...	NaN	NaN	
57967	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	...	NaN	NaN	
57968	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	...	NaN	NaN	
57969	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	...	NaN	NaN	

57970 rows × 21 columns



df[df==0].sum()

USMER	0.0
MEDICAL_UNIT	0.0
SEX	0.0
PATIENT_TYPE	0.0
DATE_DIED	0
INTUBED	0.0
PNEUMONIA	0.0
AGE	0.0
PREGNANT	0.0
DIABETES	0.0
COPD	0.0
ASTHMA	0.0
INMSUPR	0.0
HIPERTENSION	0.0
OTHER_DISEASE	0.0
CARDIOVASCULAR	0.0
OBESITY	0.0
RENAL_CHRONIC	0.0
TOBACCO	0.0
CLASIFFICATION_FINAL	0.0
ICU	0.0
dtype:	object

df.dtypes

USMER	int64
MEDICAL_UNIT	int64
SEX	int64
PATIENT_TYPE	int64
DATE_DIED	object
INTUBED	int64
PNEUMONIA	int64
AGE	int64

3/5

4/5

	precision	recall	f1-score	support
0.0	0.99	0.88	0.93	450
1.0	0.45	0.90	0.60	50
accuracy			0.88	500
macro avg	0.72	0.89	0.76	500
weighted avg	0.93	0.88	0.90	500

```
accuracy_score(actual,predicted)

0.88
```