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

```
df=pd.read_csv('/content/covid_toy - covid_toy (2).csv')
```

```
df.head(2)
```

	age	gender	fever	cough	city	has_covid
0	60	Male	103.0	Mild	Kolkata	No
1	27	Male	100.0	Mild	Delhi	Yes

Next steps: Generate code with df | New interactive sheet

```
df.isnull().sum()
```

	0
age	0
gender	0
fever	10
cough	0
city	0
has_covid	0

dtype: int64

```
df['fever']=df['fever'].fillna(df['fever'].mean())
```

```
df.isnull().sum()
```

	0
age	0
gender	0
fever	0
cough	0
city	0
has_covid	0

dtype: int64

```
df=df.drop(columns=['city'])
```

```
df.head(2)
```

	age	gender	fever	cough	has_covid
0	60	Male	103.0	Mild	No
1	27	Male	100.0	Mild	Yes

Next steps: Generate code with df | New interactive sheet

```
df['cough']=df['cough'].map({'yes':1,"no":0})
```

```
df['gender']=df['gender'].map({'male':1,'female':0})
```

```
df.head(2)
```

	age	gender	fever	cough	has_covid
0	60	NaN	103.0	NaN	No
1	27	NaN	100.0	NaN	Yes

Next steps: Generate code with df | New interactive sheet

```
x=df.drop(columns=['has_covid'])
y=df['has_covid']
```

```
from sklearn.model_selection import train_test_split
```

```
x_train,x_test,y_train,t_test = train_test_split(x,y,test_size=0.2,random_state=2)
```

```
np.round(x_train.describe(),2)
```

	age	gender	fever	cough
count	80.00	0.0	80.00	0.0
mean	43.15	NaN	100.80	NaN
std	24.66	NaN	1.90	NaN
min	5.00	NaN	98.00	NaN
25%	20.00	NaN	99.00	NaN
50%	43.00	NaN	100.84	NaN
75%	65.00	NaN	102.00	NaN
max	84.00	NaN	104.00	NaN

```
from sklearn.preprocessing import StandardScaler
```

```
ss=StandardScaler()
```

```
sc=ss.fit_transform(x_train)
```

```
/usr/local/lib/python3.12/dist-packages/sklearn/utils/extmath.py:1101: RuntimeWarning: invalid value encountered in divide
  updated_mean = (last_sum + new_sum) / updated_sample_count
/usr/local/lib/python3.12/dist-packages/sklearn/utils/extmath.py:1106: RuntimeWarning: invalid value encountered in divide
  T = new_sum / new_sample_count
/usr/local/lib/python3.12/dist-packages/sklearn/utils/extmath.py:1126: RuntimeWarning: invalid value encountered in divide
  new_unnormalized_variance -= correction**2 / new_sample_count
```

```
new=pd.DataFrame(x_train_sc,columns=x_train.columns)
```

```
np.round(new.describe(),2)
```

	age	gender	fever	cough
count	80.00	0.0	80.00	0.0
mean	0.00	NaN	-0.00	NaN
std	1.01	NaN	1.01	NaN
min	-1.56	NaN	-1.48	NaN
25%	-0.94	NaN	-0.95	NaN
50%	-0.01	NaN	0.03	NaN
75%	0.89	NaN	0.64	NaN
max	1.67	NaN	1.69	NaN