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

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

```
df.head(6)
```

	age	gender	fever	cough	city	has_covid
0	60	Male	103.0	Mild	Kolkata	No
1	27	Male	100.0	Mild	Delhi	Yes
2	42	Male	101.0	Mild	Delhi	No
3	31	Female	98.0	Mild	Kolkata	No
4	65	Female	101.0	Mild	Mumbai	No
5	84	Female	NaN	Mild	Bangalore	Yes

Next steps: Generate code with df New interactive sheet

```
df=df.drop(columns=['fever','age'])
```

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

	gender	cough	city	has_covid
0	Male	Mild	Kolkata	No
1	Male	Mild	Delhi	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,y_test=train_test_split(x,y,test_size=0.2,random_state=2)
```

```
df['cough'].value_counts()
```

cough	count
Mild	62
Strong	38

dtype: int64

```
df['city'].value_counts()
```

city	count
Kolkata	32
Bangalore	30
Delhi	22
Mumbai	16

dtype: int64

```
df['gender'].value_counts()
```

gender	count
Female	59
Male	41

dtype: int64

```
df['has_covid'].value_counts()
```

has_covid	count
No	55
Yes	45

dtype: int64

```
from sklearn.preprocessing import OrdinalEncoder
```

```
oe=OrdinalEncoder(categories=[['Male','Female'],
                  ['Mild','Strong'],
                  ['Kolkata','Bangalore','Delhi','Mumbai']])
```

```
encoded = oe.fit_transform(x_train)
```

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

```
new.head(2)
```

	gender	cough	city
0	1.0	1.0	1.0
1	1.0	0.0	3.0

Next steps: Generate code with new | New interactive sheet