Suppose we have datasets in local computer in download folder now importing it with pandas

In [1]:
```
import numpy as np
import pandas as pd
```

In [2]:
```
df=pd.read_csv(r"C:\Users\USER\Downloads\covid_toy.csv")
```

In [3]:
```
df.head(5)
```

Out[3]:

	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

In [4]:
```
df.sample(5)
```

Out[4]:

	age	gender	fever	cough	city	has_covid
93	27	Male	100.0	Mild	Kolkata	Yes
77	8	Female	101.0	Mild	Kolkata	No
55	81	Female	101.0	Mild	Mumbai	Yes
19	42	Female	NaN	Strong	Bangalore	Yes
9	64	Female	101.0	Mild	Delhi	No

Now, lets check null values, nominal and ordinal data so that if needed we can do column transformation using skitlearn

In [5]:
```
df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 100 entries, 0 to 99
Data columns (total 6 columns):
 #   Column     Non-Null Count  Dtype  
---  ------     --------------  -----  
 0   age        100 non-null    int64  
 1   gender     100 non-null    object 
 2   fever      90 non-null     float64
 3   cough      100 non-null    object 
 4   city       100 non-null    object 
 5   has_covid  100 non-null    object 
dtypes: float64(1), int64(1), object(4)
memory usage: 4.8+ KB
```

In [6]:
```
df.isnull().sum()
```

Out[6]:
```
age           0
gender        0
fever        10
cough         0
city          0
has_covid     0
dtype: int64
```

So, now before using column transfor or doing any encoder, we have to divide datasets into trainning and test.

In [7]:
```
X=df.iloc[:,0:5]
```

In [8]:
```
X.head()
```

Out[8]:

	age	gender	fever	cough	city
0	60	Male	103.0	Mild	Kolkata
1	27	Male	100.0	Mild	Delhi
2	42	Male	101.0	Mild	Delhi
3	31	Female	98.0	Mild	Kolkata
4	65	Female	101.0	Mild	Mumbai

In [9]:
```
y=df.iloc[:,-1]
y.head()
```

Out[9]:

```
0     No
1    Yes
2     No
3     No
4     No
Name: has_covid, dtype: object
```

In [10]:

```
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)
```

In [11]:

```
#lets see X_train : just to get sure
X_train
```

Out[11]:

	age	gender	fever	cough	city
43	22	Female	99.0	Mild	Bangalore
17	40	Female	98.0	Strong	Delhi
24	13	Female	100.0	Strong	Kolkata
91	38	Male	NaN	Mild	Delhi
32	34	Female	101.0	Strong	Delhi
...	...	...	...	...	...
58	23	Male	98.0	Strong	Mumbai
53	83	Male	98.0	Mild	Delhi
67	65	Male	99.0	Mild	Bangalore
56	71	Male	NaN	Strong	Kolkata
51	11	Female	100.0	Strong	Kolkata

80 rows × 5 columns

In [12]:

```
X_test
```

Out[12]:

	age	gender	fever	cough	city
36	38	Female	101.0	Mild	Bangalore
71	75	Female	104.0	Strong	Delhi
11	65	Female	98.0	Mild	Mumbai
21	73	Male	98.0	Mild	Bangalore
37	55	Male	100.0	Mild	Kolkata
75	5	Male	102.0	Mild	Kolkata
93	27	Male	100.0	Mild	Kolkata
90	59	Female	99.0	Strong	Delhi
57	49	Female	99.0	Strong	Bangalore
96	51	Female	101.0	Strong	Kolkata
13	64	Male	102.0	Mild	Bangalore
83	17	Female	104.0	Mild	Kolkata
76	80	Male	100.0	Mild	Bangalore
65	69	Female	102.0	Mild	Bangalore
55	81	Female	101.0	Mild	Mumbai
31	83	Male	103.0	Mild	Kolkata
85	16	Female	103.0	Mild	Bangalore
84	69	Female	98.0	Strong	Mumbai
64	42	Male	104.0	Mild	Mumbai
81	65	Male	99.0	Mild	Delhi

In [13]:

```
X_test.shape
```

Out[13]:

```
(20, 5)
```

Now, need to do column transfer: SimpleImputer for fever because there is 10 null values ; ordinalencoder for cough because of rank and order.

In [14]:

```
from sklearn.compose import ColumnTransformer
```

In [15]:

```
from sklearn.impute import SimpleImputer
from sklearn.preprocessing import OneHotEncoder
```

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

In [16]:

```python
transformer = ColumnTransformer(transformers=[
    ('tnf1',SimpleImputer(),['fever']),
    ('tnf2',OrdinalEncoder(categories=[['Mild','Strong']]),['cough']),
    ('tnf3',OneHotEncoder(sparse=False,drop='first'),['gender','city'])
],remainder='passthrough')
```

In [17]:

```python
X_train=transformer.fit_transform(X_train)
```

```
C:\Users\USER\anaconda3\Lib\site-packages\sklearn\preprocessing\_encoders.py:972: FutureWarning: `sparse` was r
enamed to `sparse_output` in version 1.2 and will be removed in 1.4. `sparse_output` is ignored unless you leav
e `sparse` to its default value.
  warnings.warn(
```

In [18]:

```python
X_train.shape
```

Out[18]:

```
(80, 7)
```

In [19]:

```python
X_test=transformer.transform(X_test)
```

In [20]:

```python
X_test.shape
```

Out[20]:

```
(20, 7)
```

In [21]:

```python
X_train
```

Out[21]:

```
array([[ 99.        ,   0.        ,   0.        ,   0.        ,
          0.        ,   0.        ,  22.        ],
       [ 98.        ,   1.        ,   0.        ,   1.        ,
          0.        ,   0.        ,  40.        ],
       [100.        ,   1.        ,   0.        ,   0.        ,
          1.        ,   0.        ,  13.        ],
       [100.82857143,   0.        ,   1.        ,   1.        ,
          0.        ,   0.        ,  38.        ],
       [101.        ,   1.        ,   0.        ,   1.        ,
          0.        ,   0.        ,  34.        ],
       [ 99.        ,   0.        ,   1.        ,   0.        ,
          0.        ,   0.        ,  72.        ],
       [ 98.        ,   0.        ,   0.        ,   0.        ,
          1.        ,   0.        ,  31.        ],
       [100.82857143,   1.        ,   0.        ,   0.        ,
          0.        ,   1.        ,  34.        ],
       [102.        ,   1.        ,   1.        ,   1.        ,
          0.        ,   0.        ,  20.        ],
       [101.        ,   0.        ,   0.        ,   1.        ,
          0.        ,   0.        ,  49.        ],
       [102.        ,   1.        ,   0.        ,   1.        ,
          0.        ,   0.        ,  33.        ],
       [104.        ,   0.        ,   0.        ,   0.        ,
          1.        ,   0.        ,   6.        ],
       [104.        ,   1.        ,   0.        ,   1.        ,
          0.        ,   0.        ,  34.        ],
       [100.        ,   0.        ,   1.        ,   0.        ,
          0.        ,   0.        ,  11.        ],
       [104.        ,   0.        ,   1.        ,   0.        ,
          0.        ,   0.        ,  51.        ],
       [100.82857143,   1.        ,   1.        ,   0.        ,
          1.        ,   0.        ,  79.        ],
       [100.82857143,   0.        ,   0.        ,   0.        ,
          0.        ,   0.        ,  84.        ],
       [100.        ,   0.        ,   0.        ,   0.        ,
          1.        ,   0.        ,  19.        ],
       [101.        ,   0.        ,   1.        ,   1.        ,
          0.        ,   0.        ,  15.        ],
       [102.        ,   1.        ,   0.        ,   0.        ,
          0.        ,   0.        ,  82.        ],
       [103.        ,   0.        ,   0.        ,   0.        ,
          1.        ,   0.        ,  48.        ],
       [100.82857143,   1.        ,   0.        ,   0.        ,
          0.        ,   1.        ,  20.        ],
       [ 98.        ,   1.        ,   0.        ,   0.        ,
          0.        ,   1.        ,   5.        ],
       [ 98.        ,   0.        ,   1.        ,   0.        ,
          1.        ,   0.        ,  24.        ],
       [100.        ,   0.        ,   0.        ,   0.        ,
          1.        ,   0.        ,   5.        ],
       [104.        ,   1.        ,   0.        ,   0.        ,
          0.        ,   0.        ,  56.        ],
       [102.        ,   1.        ,   0.        ,   0.        ,
          0.        ,   0.        ,  24.        ],
       [104.        ,   0.        ,   1.        ,   0.        ,
          0.        ,   1.        ,  44.        ],
       [101.        ,   0.        ,   0.        ,   0.        ,
          1.        ,   0.        ,   8.        ],
       [101.        ,   0.        ,   1.        ,   1.        ,
   0.        ,   0.        ,  42.        ],
 [104.        ,   0.        ,   1.        ,   0.        ,
   1.        ,   0.        ,  51.        ],
 [103.        ,   1.        ,   1.        ,   0.        ,
   1.        ,   0.        ,  70.        ],
 [100.        ,   0.        ,   1.        ,   1.        ,
   0.        ,   0.        ,  27.        ],
 [101.        ,   0.        ,   0.        ,   1.        ,
   0.        ,   0.        ,  64.        ],
 [ 98.        ,   0.        ,   0.        ,   0.        ,
   0.        ,   0.        ,  64.        ],
 [101.        ,   1.        ,   1.        ,   0.        ,
   0.        ,   0.        ,  14.        ],
 [ 98.        ,   0.        ,   0.        ,   1.        ,
   0.        ,   0.        ,  80.        ],
 [100.82857143,   0.        ,   1.        ,   0.        ,
   0.        ,   1.        ,  23.        ],
 [100.82857143,   1.        ,   0.        ,   0.        ,
   0.        ,   0.        ,  42.        ],
 [104.        ,   0.        ,   0.        ,   0.        ,
   0.        ,   0.        ,  18.        ],
 [ 98.        ,   1.        ,   1.        ,   0.        ,
   0.        ,   0.        ,  12.        ],
 [100.        ,   1.        ,   0.        ,   0.        ,
   0.        ,   0.        ,  47.        ],
 [102.        ,   0.        ,   1.        ,   0.        ,
   0.        ,   1.        ,  74.        ],
 [ 99.        ,   0.        ,   0.        ,   0.        ,
   0.        ,   1.        ,  14.        ],
 [100.        ,   0.        ,   1.        ,   0.        ,
   0.        ,   0.        ,  10.        ],
 [103.        ,   0.        ,   0.        ,   0.        ,
   1.        ,   0.        ,  69.        ],
 [ 98.        ,   0.        ,   0.        ,   0.        ,
   1.        ,   0.        ,  26.        ],
 [100.82857143,   0.        ,   1.        ,   0.        ,
   1.        ,   0.        ,  82.        ],
 [100.        ,   1.        ,   0.        ,   0.        ,
   0.        ,   0.        ,  19.        ],
 [101.        ,   0.        ,   0.        ,   0.        ,
   0.        ,   0.        ,  20.        ],
 [101.        ,   0.        ,   0.        ,   0.        ,
   0.        ,   1.        ,  19.        ],
 [100.        ,   0.        ,   1.        ,   1.        ,
   0.        ,   0.        ,  27.        ],
 [100.82857143,   0.        ,   0.        ,   1.        ,
   0.        ,   0.        ,  75.        ],
 [104.        ,   0.        ,   0.        ,   0.        ,
   0.        ,   0.        ,  12.        ],
 [103.        ,   0.        ,   0.        ,   1.        ,
   0.        ,   0.        ,  73.        ],
 [101.        ,   1.        ,   0.        ,   1.        ,
   0.        ,   0.        ,  68.        ],
 [104.        ,   0.        ,   1.        ,   0.        ,
   0.        ,   0.        ,  25.        ],
 [ 99.        ,   1.        ,   0.        ,   0.        ,
   1.        ,   0.        ,  25.        ],
 [ 98.        ,   1.        ,   0.        ,   0.        ,
   1.        ,   0.        ,  10.        ],
 [102.        ,   0.        ,   0.        ,   1.        ,
   0.        ,   0.        ,  49.        ],
 [ 98.        ,   1.        ,   0.        ,   0.        ,
   1.        ,   0.        ,  71.        ],
 [101.        ,   0.        ,   1.        ,   1.        ,
   0.        ,   0.        ,  19.        ],
 [102.        ,   1.        ,   0.        ,   0.        ,
   1.        ,   0.        ,  82.        ],
 [ 99.        ,   1.        ,   1.        ,   0.        ,
   0.        ,   0.        ,  66.        ],
 [ 99.        ,   0.        ,   0.        ,   0.        ,
   0.        ,   1.        ,  60.        ],
 [103.        ,   0.        ,   1.        ,   0.        ,
   1.        ,   0.        ,  60.        ],
 [ 98.        ,   1.        ,   1.        ,   0.        ,
   1.        ,   0.        ,  34.        ],
 [103.        ,   1.        ,   1.        ,   0.        ,
   0.        ,   0.        ,  46.        ],
 [101.        ,   0.        ,   0.        ,   0.        ,
   1.        ,   0.        ,  83.        ],
 [ 98.        ,   1.        ,   0.        ,   0.        ,
   0.        ,   1.        ,  81.        ],
 [101.        ,   1.        ,   1.        ,   0.        ,
   0.        ,   0.        ,  47.        ],
 [104.        ,   0.        ,   1.        ,   0.        ,
   1.        ,   0.        ,  16.        ],
 [104.        ,   1.        ,   0.        ,   0.        ,
   1.        ,   0.        ,  54.        ],
 [101.        ,   0.        ,   0.        ,   0.        ,
   0.        ,   1.        ,  65.        ],
       [103.        ,   0.        ,   0.        ,   0.        ,
          1.        ,   0.        ,  50.        ],
       [ 98.        ,   1.        ,   1.        ,   0.        ,
          0.        ,   1.        ,  23.        ],
       [ 98.        ,   0.        ,   1.        ,   1.        ,
          0.        ,   0.        ,  83.        ],
       [ 99.        ,   0.        ,   1.        ,   0.        ,
          0.        ,   0.        ,  65.        ],
       [100.82857143,   1.        ,   1.        ,   0.        ,
          1.        ,   0.        ,  71.        ],
       [100.        ,   1.        ,   0.        ,   0.        ,
          1.        ,   0.        ,  11.        ]])
```

In [22]:
```
X_test
```

Out[22]:
```
array([[101.,   0.,   0.,   0.,   0.,   0.,  38.],
       [104.,   1.,   0.,   1.,   0.,   0.,  75.],
       [ 98.,   0.,   0.,   0.,   0.,   1.,  65.],
       [ 98.,   0.,   1.,   0.,   0.,   0.,  73.],
       [100.,   0.,   1.,   0.,   1.,   0.,  55.],
       [102.,   0.,   1.,   0.,   1.,   0.,   5.],
       [100.,   0.,   1.,   0.,   1.,   0.,  27.],
       [ 99.,   1.,   0.,   1.,   0.,   0.,  59.],
       [ 99.,   1.,   0.,   0.,   0.,   0.,  49.],
       [101.,   1.,   0.,   0.,   1.,   0.,  51.],
       [102.,   0.,   1.,   0.,   0.,   0.,  64.],
       [104.,   0.,   0.,   0.,   1.,   0.,  17.],
       [100.,   0.,   1.,   0.,   0.,   0.,  80.],
       [102.,   0.,   0.,   0.,   0.,   0.,  69.],
       [101.,   0.,   0.,   0.,   0.,   1.,  81.],
       [103.,   0.,   1.,   0.,   1.,   0.,  83.],
       [103.,   0.,   0.,   0.,   0.,   0.,  16.],
       [ 98.,   1.,   0.,   0.,   0.,   1.,  69.],
       [104.,   0.,   1.,   0.,   0.,   1.,  42.],
       [ 99.,   0.,   1.,   1.,   0.,   0.,  65.]])
```

Loading [MathJax]/jax/output/CommonHTML/fonts/TeX/fontdata.js