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

```
df =
pd.read_csv('https://raw.githubusercontent.com/YBI-Foundation/Dataset/
main/Fruits.csv')
```

```
df.head()
```

```
   Fruit Category Fruit Name  Fruit Weight  Fruit Width  Fruit Length
\
0               1      Apple           192          8.4           7.3

1               1      Apple           180          8.0           6.8

2               1      Apple           176          7.4           7.2

3               1      Apple           178          7.1           7.8

4               1      Apple           172          7.4           7.0


   Fruit Colour Score
0                0.55
1                0.59
2                0.60
3                0.92
4                0.89
```

```
df.info
```

```
<bound method DataFrame.info of     Fruit Category Fruit Name  Fruit
Weight  Fruit Width  Fruit Length  \
0                1      Apple           192          8.4           7.3

1                1      Apple           180          8.0           6.8

2                1      Apple           176          7.4           7.2

3                1      Apple           178          7.1           7.8

4                1      Apple           172          7.4           7.0

5                1      Apple           166          6.9           7.3

6                1      Apple           172          7.1           7.6

7                1      Apple           154          7.0           7.1

8                1      Apple           164          7.3           7.7
```

9	1	Apple	152	7.6	7.3
10	1	Apple	156	7.7	7.1
11	1	Apple	156	7.6	7.5
12	1	Apple	168	7.5	7.6
13	1	Apple	162	7.5	7.1
14	1	Apple	162	7.4	7.2
15	1	Apple	160	7.5	7.5
16	1	Apple	156	7.4	7.4
17	1	Apple	140	7.3	7.1
18	1	Apple	170	7.6	7.9
19	2	Orange	86	6.2	4.7
20	2	Orange	84	6.0	4.6
21	2	Orange	80	5.8	4.3
22	2	Orange	80	5.9	4.3
23	2	Orange	76	5.8	4.0
24	2	Orange	342	9.0	9.4
25	2	Orange	356	9.2	9.2
26	2	Orange	362	9.6	9.2
27	2	Orange	204	7.5	9.2
28	2	Orange	140	6.7	7.1
29	2	Orange	160	7.0	7.4
30	2	Orange	158	7.1	7.5
31	2	Orange	210	7.8	8.0
32	2	Orange	164	7.2	7.0

33	2	Orange	190	7.5	8.1
34	2	Orange	142	7.6	7.8
35	2	Orange	150	7.1	7.9
36	2	Orange	160	7.1	7.6
37	2	Orange	154	7.3	7.3
38	2	Orange	158	7.2	7.8
39	2	Orange	144	6.8	7.4
40	2	Orange	154	7.1	7.5
41	2	Orange	180	7.6	8.2
42	2	Orange	154	7.2	7.2
43	3	Lemon	97	7.2	10.3
44	3	Lemon	70	7.3	10.5
45	3	Lemon	93	7.2	9.2
46	3	Lemon	80	7.3	10.2
47	3	Lemon	98	7.3	9.7
48	3	Lemon	87	7.3	10.1
49	3	Lemon	66	5.8	8.7
50	3	Lemon	65	6.0	8.2
51	3	Lemon	58	6.0	7.5
52	3	Lemon	59	5.9	8.0
53	3	Lemon	60	6.0	8.4
54	3	Lemon	58	6.1	8.5
55	3	Lemon	58	6.3	7.7
56	3	Lemon	58	5.9	8.1

57	3	Lemon	76	6.5	8.5
58	3	Lemon	59	6.1	8.1

```
    Fruit Colour Score
0                 0.55
1                 0.59
2                 0.60
3                 0.92
4                 0.89
5                 0.93
6                 0.92
7                 0.88
8                 0.70
9                 0.69
10                0.69
11                0.67
12                0.73
13                0.83
14                0.85
15                0.86
16                0.84
17                0.87
18                0.88
19                0.80
20                0.79
21                0.77
22                0.81
23                0.81
24                0.75
25                0.75
26                0.74
27                0.77
28                0.72
29                0.81
30                0.79
31                0.82
32                0.80
33                0.74
34                0.75
35                0.75
36                0.76
37                0.79
38                0.77
39                0.75
40                0.78
41                0.79
42                0.82
```

```
43                 0.70
44                 0.72
45                 0.72
46                 0.71
47                 0.72
48                 0.72
49                 0.73
50                 0.71
51                 0.72
52                 0.72
53                 0.74
54                 0.71
55                 0.72
56                 0.73
57                 0.72
58                 0.70  >
```

```
df.describe
```

```
<bound method NDFrame.describe of     Fruit Category Fruit Name  Fruit
Weight  Fruit Width  Fruit Length  \
0                1      Apple          192           8.4           7.3

1                1      Apple          180           8.0           6.8

2                1      Apple          176           7.4           7.2

3                1      Apple          178           7.1           7.8

4                1      Apple          172           7.4           7.0

5                1      Apple          166           6.9           7.3

6                1      Apple          172           7.1           7.6

7                1      Apple          154           7.0           7.1

8                1      Apple          164           7.3           7.7

9                1      Apple          152           7.6           7.3

10               1      Apple          156           7.7           7.1

11               1      Apple          156           7.6           7.5

12               1      Apple          168           7.5           7.6

13               1      Apple          162           7.5           7.1
```

14	1	Apple	162	7.4	7.2
15	1	Apple	160	7.5	7.5
16	1	Apple	156	7.4	7.4
17	1	Apple	140	7.3	7.1
18	1	Apple	170	7.6	7.9
19	2	Orange	86	6.2	4.7
20	2	Orange	84	6.0	4.6
21	2	Orange	80	5.8	4.3
22	2	Orange	80	5.9	4.3
23	2	Orange	76	5.8	4.0
24	2	Orange	342	9.0	9.4
25	2	Orange	356	9.2	9.2
26	2	Orange	362	9.6	9.2
27	2	Orange	204	7.5	9.2
28	2	Orange	140	6.7	7.1
29	2	Orange	160	7.0	7.4
30	2	Orange	158	7.1	7.5
31	2	Orange	210	7.8	8.0
32	2	Orange	164	7.2	7.0
33	2	Orange	190	7.5	8.1
34	2	Orange	142	7.6	7.8
35	2	Orange	150	7.1	7.9
36	2	Orange	160	7.1	7.6
37	2	Orange	154	7.3	7.3

```
38               2     Orange           158          7.2           7.8

39               2     Orange           144          6.8           7.4

40               2     Orange           154          7.1           7.5

41               2     Orange           180          7.6           8.2

42               2     Orange           154          7.2           7.2

43               3      Lemon            97          7.2          10.3

44               3      Lemon            70          7.3          10.5

45               3      Lemon            93          7.2           9.2

46               3      Lemon            80          7.3          10.2

47               3      Lemon            98          7.3           9.7

48               3      Lemon            87          7.3          10.1

49               3      Lemon            66          5.8           8.7

50               3      Lemon            65          6.0           8.2

51               3      Lemon            58          6.0           7.5

52               3      Lemon            59          5.9           8.0

53               3      Lemon            60          6.0           8.4

54               3      Lemon            58          6.1           8.5

55               3      Lemon            58          6.3           7.7

56               3      Lemon            58          5.9           8.1

57               3      Lemon            76          6.5           8.5

58               3      Lemon            59          6.1           8.1
```

```
    Fruit Colour Score
0                 0.55
1                 0.59
2                 0.60
3                 0.92
```

4	0.89
5	0.93
6	0.92
7	0.88
8	0.70
9	0.69
10	0.69
11	0.67
12	0.73
13	0.83
14	0.85
15	0.86
16	0.84
17	0.87
18	0.88
19	0.80
20	0.79
21	0.77
22	0.81
23	0.81
24	0.75
25	0.75
26	0.74
27	0.77
28	0.72
29	0.81
30	0.79
31	0.82
32	0.80
33	0.74
34	0.75
35	0.75
36	0.76
37	0.79
38	0.77
39	0.75
40	0.78
41	0.79
42	0.82
43	0.70
44	0.72
45	0.72
46	0.71
47	0.72
48	0.72
49	0.73
50	0.71
51	0.72
52	0.72
53	0.74

```
54                0.71
55                0.72
56                0.73
57                0.72
58                0.70  >
```

```
df.shape
```

```
(59, 6)
```

```
df.columns
```

```
Index(['Fruit Category', 'Fruit Name', 'Fruit Weight', 'Fruit Width',
       'Fruit Length', 'Fruit Colour Score'],
      dtype='object')
```

```
df['Fruit Category'].value_counts()
```

```
2    24
1    19
3    16
Name: Fruit Category, dtype: int64
```

```
df.groupby('Fruit Category').mean()
```

```
                Fruit Weight  Fruit Width  Fruit Length  Fruit Colour 
Score
Fruit Category

1                 165.052632     7.457895      7.342105 
0.783684
2                 170.333333     7.220833      7.195833 
0.776250
3                  71.375000     6.512500      8.856250 
0.718125
```

```
y = df['Fruit Category']
```

```
y.shape
```

```
(59,)
```

```
y
```

```
0     1
1     1
2     1
3     1
4     1
5     1
6     1
7     1
8     1
```

9	1
10	1
11	1
12	1
13	1
14	1
15	1
16	1
17	1
18	1
19	2
20	2
21	2
22	2
23	2
24	2
25	2
26	2
27	2
28	2
29	2
30	2
31	2
32	2
33	2
34	2
35	2
36	2
37	2
38	2
39	2
40	2
41	2
42	2
43	3
44	3
45	3
46	3
47	3
48	3
49	3
50	3
51	3
52	3
53	3
54	3
55	3
56	3
57	3

```
58    3
Name: Fruit Category, dtype: int64
```

```
x = df[['Fruit Weight', 'Fruit Width', 'Fruit Length', 'Fruit Colour
Score']]
```

```
x.shape
```

```
(59, 4)
```

```
x
```

```
    Fruit Weight  Fruit Width  Fruit Length  Fruit Colour Score
0            192          8.4           7.3                0.55
1            180          8.0           6.8                0.59
2            176          7.4           7.2                0.60
3            178          7.1           7.8                0.92
4            172          7.4           7.0                0.89
5            166          6.9           7.3                0.93
6            172          7.1           7.6                0.92
7            154          7.0           7.1                0.88
8            164          7.3           7.7                0.70
9            152          7.6           7.3                0.69
10           156          7.7           7.1                0.69
11           156          7.6           7.5                0.67
12           168          7.5           7.6                0.73
13           162          7.5           7.1                0.83
14           162          7.4           7.2                0.85
15           160          7.5           7.5                0.86
16           156          7.4           7.4                0.84
17           140          7.3           7.1                0.87
18           170          7.6           7.9                0.88
19            86          6.2           4.7                0.80
20            84          6.0           4.6                0.79
21            80          5.8           4.3                0.77
22            80          5.9           4.3                0.81
23            76          5.8           4.0                0.81
24           342          9.0           9.4                0.75
25           356          9.2           9.2                0.75
26           362          9.6           9.2                0.74
27           204          7.5           9.2                0.77
28           140          6.7           7.1                0.72
29           160          7.0           7.4                0.81
30           158          7.1           7.5                0.79
31           210          7.8           8.0                0.82
32           164          7.2           7.0                0.80
33           190          7.5           8.1                0.74
34           142          7.6           7.8                0.75
35           150          7.1           7.9                0.75
36           160          7.1           7.6                0.76
37           154          7.3           7.3                0.79
```

```
38           158          7.2           7.8                0.77
39           144          6.8           7.4                0.75
40           154          7.1           7.5                0.78
41           180          7.6           8.2                0.79
42           154          7.2           7.2                0.82
43            97          7.2          10.3                0.70
44            70          7.3          10.5                0.72
45            93          7.2           9.2                0.72
46            80          7.3          10.2                0.71
47            98          7.3           9.7                0.72
48            87          7.3          10.1                0.72
49            66          5.8           8.7                0.73
50            65          6.0           8.2                0.71
51            58          6.0           7.5                0.72
52            59          5.9           8.0                0.72
53            60          6.0           8.4                0.74
54            58          6.1           8.5                0.71
55            58          6.3           7.7                0.72
56            58          5.9           8.1                0.73
57            76          6.5           8.5                0.72
58            59          6.1           8.1                0.70
```

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

```
x_train, x_test, y_train, y_test = train_test_split(x,y,
test_size=0.3, random_state=2529)
```

```
x_train.shape, x_test.shape, y_train.shape, y_test.shape
```

```
((41, 4), (18, 4), (41,), (18,))
```

```
from sklearn.linear_model import LogisticRegression
```

```
model = LogisticRegression(max_iter=500)
```

```
model.fit(x_train, y_train)
```

```
LogisticRegression(max_iter=500)
```

```
y_pred = model.predict(x_test)
```

```
y_pred.shape
```

```
(18,)
```

```
y_pred
```

```
array([3, 3, 2, 2, 2, 2, 2, 3, 2, 1, 3, 1, 2, 3, 1, 2, 1, 3])
```

```
model.predict_proba(x_test)
```

```
array([[2.41632966e-03, 3.19679220e-03, 9.94386878e-01],
       [2.13813513e-03, 2.41144478e-03, 9.95450420e-01],
       [4.00734862e-01, 5.98780620e-01, 4.84518233e-04],
       [4.38505956e-01, 5.61492022e-01, 2.02148976e-06],
       [4.08832600e-01, 5.91142577e-01, 2.48232453e-05],
       [4.53050894e-01, 5.46408107e-01, 5.40998978e-04],
       [2.52105201e-01, 7.01822439e-01, 4.60723596e-02],
       [8.66358590e-03, 1.85227743e-03, 9.89484137e-01],
       [4.95660140e-01, 5.04153971e-01, 1.85888493e-04],
       [6.04355858e-01, 3.95230058e-01, 4.14083784e-04],
       [1.42234794e-02, 1.02050293e-02, 9.75571491e-01],
       [5.75881620e-01, 4.23402795e-01, 7.15584880e-04],
       [4.17695783e-01, 5.82304217e-01, 1.39382161e-13],
       [7.46818111e-04, 1.09501728e-04, 9.99143680e-01],
       [6.72343729e-01, 3.27648588e-01, 7.68383938e-06],
       [3.63803252e-01, 6.36098813e-01, 9.79356349e-05],
       [5.76094513e-01, 4.23889814e-01, 1.56732761e-05],
       [5.00639637e-03, 3.81527341e-03, 9.91178330e-01]])
```

```
from sklearn.metrics import confusion_matrix, classification_report
```

```
print(confusion_matrix(y_test, y_pred))
```

```
[[4 2 0]
 [0 6 0]
 [0 0 6]]
```

```
print(classification_report(y_test, y_pred))
```

```
              precision    recall  f1-score   support

           1       1.00      0.67      0.80         6
           2       0.75      1.00      0.86         6
           3       1.00      1.00      1.00         6

    accuracy                           0.89        18
   macro avg       0.92      0.89      0.89        18
weighted avg       0.92      0.89      0.89        18
```

```
df_new = df.sample(1)
```

```
df_new
```

```
    Fruit Category Fruit Name  Fruit Weight  Fruit Width  Fruit Length
\
30               2     Orange           158          7.1           7.5



    Fruit Colour Score
30                0.79
```

```
x_new = df_new[['Fruit Weight', 'Fruit Width','Fruit Length', 'Fruit
Colour Score']]
```

```
x_new.shape
```

```
(1, 4)
```

```
y_pred_new = model.predict(x_new)
```

```
y_pred_new
```

```
array([2])
```

```
model.predict_proba(x_new)
```

```
array([[4.00734862e-01, 5.98780620e-01, 4.84518233e-04]])
```