practical-1

May 9, 2025

[61]:
```
import pandas as pd
```

[62]:
```
iris=pd.read_csv("iris.csv")
```

[63]:
```
iris.describe()
```

[63]:

	Id	SepalLengthCm	SepalWidthCm	PetalLengthCm	PetalWidthCm
count	150.000000	150.000000	150.000000	150.000000	150.000000
mean	75.500000	5.843333	3.054000	3.758667	1.198667
std	43.445368	0.828066	0.433594	1.764420	0.763161
min	1.000000	4.300000	2.000000	1.000000	0.100000
25%	38.250000	5.100000	2.800000	1.600000	0.300000
50%	75.500000	5.800000	3.000000	4.350000	1.300000
75%	112.750000	6.400000	3.300000	5.100000	1.800000
max	150.000000	7.900000	4.400000	6.900000	2.500000

[64]:
```
iris.head()
```

[64]:

	Id	SepalLengthCm	SepalWidthCm	PetalLengthCm	PetalWidthCm	Species
0	1	5.1	3.5	1.4	0.2	Iris-setosa
1	2	4.9	3.0	1.4	0.2	Iris-setosa
2	3	4.7	3.2	1.3	0.2	Iris-setosa
3	4	4.6	3.1	1.5	0.2	Iris-setosa
4	5	5.0	3.6	1.4	0.2	Iris-setosa

[65]:
```
iris.tail()
```

[65]:

	Id	SepalLengthCm	SepalWidthCm	PetalLengthCm	PetalWidthCm	\
145	146	6.7	3.0	5.2	2.3	
146	147	6.3	2.5	5.0	1.9	
147	148	6.5	3.0	5.2	2.0	
148	149	6.2	3.4	5.4	2.3	
149	150	5.9	3.0	5.1	1.8	

	Species
145	Iris-virginica
146	Iris-virginica

```
147  Iris-virginica
148  Iris-virginica
149  Iris-virginica
```

[66]:
```
iris.index
```

[66]:
```
RangeIndex(start=0, stop=150, step=1)
```

[67]:
```
iris.columns
```

[67]:
```
Index(['Id', 'SepalLengthCm', 'SepalWidthCm', 'PetalLengthCm', 'PetalWidthCm',
       'Species'],
      dtype='object')
```

[68]:
```
iris.shape
```

[68]:
```
(150, 6)
```

[69]:
```
iris.dtypes
```

[69]:
```
Id                int64
SepalLengthCm    float64
SepalWidthCm     float64
PetalLengthCm    float64
PetalWidthCm     float64
Species           object
dtype: object
```

[70]:
```
iris.columns.values
```

[70]:
```
array(['Id', 'SepalLengthCm', 'SepalWidthCm', 'PetalLengthCm',
       'PetalWidthCm', 'Species'], dtype=object)
```

[71]:
```
iris.describe(include="all")
```

[71]:

	Id	SepalLengthCm	SepalWidthCm	PetalLengthCm	PetalWidthCm	\
count	150.000000	150.000000	150.000000	150.000000	150.000000	
unique	NaN	NaN	NaN	NaN	NaN	
top	NaN	NaN	NaN	NaN	NaN	
freq	NaN	NaN	NaN	NaN	NaN	
mean	75.500000	5.843333	3.054000	3.758667	1.198667	
std	43.445368	0.828066	0.433594	1.764420	0.763161	
min	1.000000	4.300000	2.000000	1.000000	0.100000	
25%	38.250000	5.100000	2.800000	1.600000	0.300000	
50%	75.500000	5.800000	3.000000	4.350000	1.300000	
75%	112.750000	6.400000	3.300000	5.100000	1.800000	
max	150.000000	7.900000	4.400000	6.900000	2.500000	

```
             Species
count            150
unique             3
top     Iris-setosa
freq              50
mean             NaN
std              NaN
min              NaN
25%              NaN
50%              NaN
75%              NaN
max              NaN
```

[72]:
```
iris['SepalLengthCm']
```

[72]:
```
0      5.1
1      4.9
2      4.7
3      4.6
4      5.0
      …
145    6.7
146    6.3
147    6.5
148    6.2
149    5.9
Name: SepalLengthCm, Length: 150, dtype: float64
```

[73]:
```
iris['PetalLengthCm']
```

[73]:
```
0      1.4
1      1.4
2      1.3
3      1.5
4      1.4
      …
145    5.2
146    5.0
147    5.2
148    5.4
149    5.1
Name: PetalLengthCm, Length: 150, dtype: float64
```

[74]:
```
iris.sort_index(axis=1,ascending=False)
```

[74]:
```
            Species  SepalWidthCm  SepalLengthCm  PetalWidthCm  PetalLengthCm  \
0       Iris-setosa           3.5            5.1           0.2            1.4
1       Iris-setosa           3.0            4.9           0.2            1.4
2       Iris-setosa           3.2            4.7           0.2            1.3
3       Iris-setosa           3.1            4.6           0.2            1.5
4       Iris-setosa           3.6            5.0           0.2            1.4
..              …             …              …             …              …
145  Iris-virginica           3.0            6.7           2.3            5.2
146  Iris-virginica           2.5            6.3           1.9            5.0
147  Iris-virginica           3.0            6.5           2.0            5.2
148  Iris-virginica           3.4            6.2           2.3            5.4
149  Iris-virginica           3.0            5.9           1.8            5.1

      Id
0      1
1      2
2      3
3      4
4      5
..   …
145  146
146  147
147  148
148  149
149  150

[150 rows x 6 columns]
```

[75]:
```
iris.sort_values(by="SepalWidthCm")
```

[75]:
```
      Id  SepalLengthCm  SepalWidthCm  PetalLengthCm  PetalWidthCm  \
60    61            5.0           2.0            3.5           1.0
62    63            6.0           2.2            4.0           1.0
119  120            6.0           2.2            5.0           1.5
68    69            6.2           2.2            4.5           1.5
87    88            6.3           2.3            4.4           1.3
..   …              …             …              …             …
5      6            5.4           3.9            1.7           0.4
14    15            5.8           4.0            1.2           0.2
32    33            5.2           4.1            1.5           0.1
33    34            5.5           4.2            1.4           0.2
15    16            5.7           4.4            1.5           0.4

             Species
60   Iris-versicolor
62   Iris-versicolor
119   Iris-virginica
```

```
68   Iris-versicolor
87   Iris-versicolor
..               …
5        Iris-setosa
14       Iris-setosa
32       Iris-setosa
33       Iris-setosa
15       Iris-setosa

[150 rows x 6 columns]
```

[76]:
```
iris.iloc[5]
```

[76]:
```
Id                      6
SepalLengthCm         5.4
SepalWidthCm          3.9
PetalLengthCm         1.7
PetalWidthCm          0.4
Species       Iris-setosa
Name: 5, dtype: object
```

[77]:
```
iris[0:3]
```

[77]:

	Id	SepalLengthCm	SepalWidthCm	PetalLengthCm	PetalWidthCm	Species
0	1	5.1	3.5	1.4	0.2	Iris-setosa
1	2	4.9	3.0	1.4	0.2	Iris-setosa
2	3	4.7	3.2	1.3	0.2	Iris-setosa

[79]:
```
iris.loc[:, ["SepalLengthCm","SepalWidthCm"]]
```

[79]:
```
     SepalLengthCm  SepalWidthCm
0              5.1           3.5
1              4.9           3.0
2              4.7           3.2
3              4.6           3.1
4              5.0           3.6
..             …             …
145            6.7           3.0
146            6.3           2.5
147            6.5           3.0
148            6.2           3.4
149            5.9           3.0

[150 rows x 2 columns]
```

[80]:
```
iris.iloc[:5, :]
```

[80]:

	Id	SepalLengthCm	SepalWidthCm	PetalLengthCm	PetalWidthCm	Species
0	1	5.1	3.5	1.4	0.2	Iris-setosa
1	2	4.9	3.0	1.4	0.2	Iris-setosa
2	3	4.7	3.2	1.3	0.2	Iris-setosa
3	4	4.6	3.1	1.5	0.2	Iris-setosa
4	5	5.0	3.6	1.4	0.2	Iris-setosa

[81]:
```
iris.iloc[:, :4]
```

[81]:

	Id	SepalLengthCm	SepalWidthCm	PetalLengthCm
0	1	5.1	3.5	1.4
1	2	4.9	3.0	1.4
2	3	4.7	3.2	1.3
3	4	4.6	3.1	1.5
4	5	5.0	3.6	1.4
..	…	…	…	…
145	146	6.7	3.0	5.2
146	147	6.3	2.5	5.0
147	148	6.5	3.0	5.2
148	149	6.2	3.4	5.4
149	150	5.9	3.0	5.1

[150 rows x 4 columns]

[82]:
```
iris.iloc[:7, :5]
```

[82]:

	Id	SepalLengthCm	SepalWidthCm	PetalLengthCm	PetalWidthCm
0	1	5.1	3.5	1.4	0.2
1	2	4.9	3.0	1.4	0.2
2	3	4.7	3.2	1.3	0.2
3	4	4.6	3.1	1.5	0.2
4	5	5.0	3.6	1.4	0.2
5	6	5.4	3.9	1.7	0.4
6	7	4.6	3.4	1.4	0.3

[83]:
```
iris.iloc[3:5, 0:2]
```

[83]:

	Id	SepalLengthCm
3	4	4.6
4	5	5.0

[84]:
```
iris.iloc[[1, 2, 4],[0,2]]
```

[84]:

	Id	SepalWidthCm
1	2	3.0
2	3	3.2
4	5	3.6

[85]:
```
iris.iloc[1:3, :]
```

[85]:

	Id	SepalLengthCm	SepalWidthCm	PetalLengthCm	PetalWidthCm	Species
1	2	4.9	3.0	1.4	0.2	Iris-setosa
2	3	4.7	3.2	1.3	0.2	Iris-setosa

[86]:
```
iris.iloc[:, 1:3]
```

[86]:

	SepalLengthCm	SepalWidthCm
0	5.1	3.5
1	4.9	3.0
2	4.7	3.2
3	4.6	3.1
4	5.0	3.6
..	…	…
145	6.7	3.0
146	6.3	2.5
147	6.5	3.0
148	6.2	3.4
149	5.9	3.0

[150 rows x 2 columns]

[87]:
```
iris.iloc[1,1]
```

[87]:
```
np.float64(4.9)
```

[90]:
```
iris['PetalLengthCm'].iloc[5]
```

[90]:
```
np.float64(1.7)
```

[91]:
```
cols_2_4=iris.columns[2:4]
iris[cols_2_4]
```

[91]:

	SepalWidthCm	PetalLengthCm
0	3.5	1.4
1	3.0	1.4
2	3.2	1.3
3	3.1	1.5
4	3.6	1.4
..	…	…
145	3.0	5.2
146	2.5	5.0
147	3.0	5.2
148	3.4	5.4
149	3.0	5.1

```
[150 rows x 2 columns]
```

[92]:
```
iris[iris.columns[2:4]].iloc[5:10]
```

[92]:
```
   SepalWidthCm  PetalLengthCm
5           3.9            1.7
6           3.4            1.4
7           3.4            1.5
8           2.9            1.4
9           3.1            1.5
```

[93]:
```
iris.isnull()
```

[93]:
```
        Id  SepalLengthCm  SepalWidthCm  PetalLengthCm  PetalWidthCm  Species
0    False          False         False          False         False    False
1    False          False         False          False         False    False
2    False          False         False          False         False    False
3    False          False         False          False         False    False
4    False          False         False          False         False    False
..     …              …             …              …             …        …
145  False          False         False          False         False    False
146  False          False         False          False         False    False
147  False          False         False          False         False    False
148  False          False         False          False         False    False
149  False          False         False          False         False    False
```

```
[150 rows x 6 columns]
```

[94]:
```
data={'Name':['Jai','Princi','Gaurav','Anuj','Ravi','Natasha','Riya'],
'Age':[17,17,18,17,18,17,17],
'Gender':['M','F','M','M','M','F','F'],
'Marks':[90,76,'NaN',74,65,'NaN',71]}
```

[95]:
```
df=pd.DataFrame(data)
```

[107]:
```
df
```

[107]:
```
      Name  Age Gender  Marks
0      Jai   17      M   90.0
1   Princi   17      F   76.0
2   Gaurav   18      M   75.2
3     Anuj   17      M   74.0
4     Ravi   18      M   65.0
5  Natasha   17      F   75.2
6     Riya   17      F   71.0
```

[109]:
```
c = avg = 0
for ele in df['Marks']:
    if str(ele).isnumeric():  # Ensure 'ele' is numeric
        c += 1
        avg += float(ele)  # Convert to float to avoid type issues

if c != 0:
    avg /= c
else:
    avg = 0  # To avoid ZeroDivisionError
```

[110]:
```
df=df.replace(to_replace="NaN", value=avg)
```

[111]:
```
df
```

[111]:

	Name	Age	Gender	Marks
0	Jai	17	M	90.0
1	Princi	17	F	76.0
2	Gaurav	18	M	75.2
3	Anuj	17	M	74.0
4	Ravi	18	M	65.0
5	Natasha	17	F	75.2
6	Riya	17	F	71.0

[112]:
```
df['Gender'] = df['Gender'].map({'M': 0,'F': 1, }).astype(float)
```

[113]:
```
df
```

[113]:

	Name	Age	Gender	Marks
0	Jai	17	0.0	90.0
1	Princi	17	1.0	76.0
2	Gaurav	18	0.0	75.2
3	Anuj	17	0.0	74.0
4	Ravi	18	0.0	65.0
5	Natasha	17	1.0	75.2
6	Riya	17	1.0	71.0

[114]:
```
df = df[df['Marks'] >= 75]
```

[115]:
```
df = df.drop(['Age'], axis=1)
```

[116]:
```
df
```

[116]:

	Name	Gender	Marks
0	Jai	0.0	90.0
1	Princi	1.0	76.0
2	Gaurav	0.0	75.2

```
5  Natasha     1.0   75.2
```

[ ]: