In [6]:

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

In [7]:

```
df=pd.read_csv("IRIS.csv")
```

In [8]:

```
df
```

Out[8]:

	sepal_length	sepal_width	petal_length	petal_width	species
0	5.1	3.5	1.4	0.2	Iris-setosa
1	4.9	3.0	1.4	0.2	Iris-setosa
2	4.7	3.2	1.3	0.2	Iris-setosa
3	4.6	3.1	1.5	0.2	Iris-setosa
4	5.0	3.6	1.4	0.2	Iris-setosa
...	...	...	...	...	...
145	6.7	3.0	5.2	2.3	Iris-virginica
146	6.3	2.5	5.0	1.9	Iris-virginica
147	6.5	3.0	5.2	2.0	Iris-virginica
148	6.2	3.4	5.4	2.3	Iris-virginica
149	5.9	3.0	5.1	1.8	Iris-virginica

150 rows × 5 columns

In [9]:

```
df.head()
```

Out[9]:

	sepal_length	sepal_width	petal_length	petal_width	species
0	5.1	3.5	1.4	0.2	Iris-setosa
1	4.9	3.0	1.4	0.2	Iris-setosa
2	4.7	3.2	1.3	0.2	Iris-setosa
3	4.6	3.1	1.5	0.2	Iris-setosa
4	5.0	3.6	1.4	0.2	Iris-setosa

In [10]:

```
df.tail()
```

Out[10]:

	sepal_length	sepal_width	petal_length	petal_width	species
145	6.7	3.0	5.2	2.3	Iris-virginica
146	6.3	2.5	5.0	1.9	Iris-virginica
147	6.5	3.0	5.2	2.0	Iris-virginica
148	6.2	3.4	5.4	2.3	Iris-virginica
149	5.9	3.0	5.1	1.8	Iris-virginica

In [11]:

```
df.index
```

Out[11]:

```
RangeIndex(start=0, stop=150, step=1)
```

In [12]:

```
df.columns
```

Out[12]:

```
Index(['sepal_length', 'sepal_width', 'petal_length', 'petal_width',
       'species'],
      dtype='object')
```

In [13]:

```
df.shape
```

Out[13]:

```
(150, 5)
```

In [14]:

```
df.dtypes
```

Out[14]:

```
sepal_length    float64
sepal_width     float64
petal_length    float64
petal_width     float64
species          object
dtype: object
```

In [15]:

```
df.columns.values
```

Out[15]:

```
array(['sepal_length', 'sepal_width', 'petal_length', 'petal_width',
       'species'], dtype=object)
```

In [16]:

```
df.columns.values.tolist()
```

Out[16]:

```
['sepal_length', 'sepal_width', 'petal_length', 'petal_width', 'species']
```

In [17]:

```
df['sepal_length']
```

Out[17]:

```
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: sepal_length, Length: 150, dtype: float64
```

In [18]:

```
df['petal_length']
```

Out[18]:

```
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: petal_length, Length: 150, dtype: float64
```

In [19]:

```
df.to_numpy()
```

Out[19]:

```
array([[5.1, 3.5, 1.4, 0.2, 'Iris-setosa'],
       [4.9, 3.0, 1.4, 0.2, 'Iris-setosa'],
       [4.7, 3.2, 1.3, 0.2, 'Iris-setosa'],
       [4.6, 3.1, 1.5, 0.2, 'Iris-setosa'],
       [5.0, 3.6, 1.4, 0.2, 'Iris-setosa'],
       [5.4, 3.9, 1.7, 0.4, 'Iris-setosa'],
       [4.6, 3.4, 1.4, 0.3, 'Iris-setosa'],
       [5.0, 3.4, 1.5, 0.2, 'Iris-setosa'],
       [4.4, 2.9, 1.4, 0.2, 'Iris-setosa'],
       [4.9, 3.1, 1.5, 0.1, 'Iris-setosa'],
       [5.4, 3.7, 1.5, 0.2, 'Iris-setosa'],
       [4.8, 3.4, 1.6, 0.2, 'Iris-setosa'],
       [4.8, 3.0, 1.4, 0.1, 'Iris-setosa'],
       [4.3, 3.0, 1.1, 0.1, 'Iris-setosa'],
       [5.8, 4.0, 1.2, 0.2, 'Iris-setosa'],
       [5.7, 4.4, 1.5, 0.4, 'Iris-setosa'],
       [5.4, 3.9, 1.3, 0.4, 'Iris-setosa'],
       [5.1, 3.5, 1.4, 0.3, 'Iris-setosa'],
```

In [20]:

```
df.T
```

Out[20]:

	0	1	2	3	4	5	6	7	8	9	...
sepal_length	5.1	4.9	4.7	4.6	5.0	5.4	4.6	5.0	4.4	4.9	...
sepal_width	3.5	3.0	3.2	3.1	3.6	3.9	3.4	3.4	2.9	3.1	...
petal_length	1.4	1.4	1.3	1.5	1.4	1.7	1.4	1.5	1.4	1.5	...
petal_width	0.2	0.2	0.2	0.2	0.2	0.4	0.3	0.2	0.2	0.1	...
species	Iris-setosa	Iris-setosa	Iris-setosa	Iris-setosa	Iris-setosa	Iris-setosa	Iris-setosa	Iris-setosa	Iris-setosa	Iris-setosa	...

5 rows × 150 columns

In [21]:

```
df.sort_index(axis=1, ascending=False)
```

Out[21]:

	species	sepal_width	sepal_length	petal_width	petal_length
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

150 rows × 5 columns

In [22]:

```
df.sort_values(by="sepal_length")
```

Out[22]:

	sepal_length	sepal_width	petal_length	petal_width	species
13	4.3	3.0	1.1	0.1	Iris-setosa
42	4.4	3.2	1.3	0.2	Iris-setosa
38	4.4	3.0	1.3	0.2	Iris-setosa
8	4.4	2.9	1.4	0.2	Iris-setosa
41	4.5	2.3	1.3	0.3	Iris-setosa
...	...	...	...	...	...
122	7.7	2.8	6.7	2.0	Iris-virginica
118	7.7	2.6	6.9	2.3	Iris-virginica
117	7.7	3.8	6.7	2.2	Iris-virginica
135	7.7	3.0	6.1	2.3	Iris-virginica
131	7.9	3.8	6.4	2.0	Iris-virginica

150 rows × 5 columns

In [23]:

```
df.isnull()
```

Out[23]:

	sepal_length	sepal_width	petal_length	petal_width	species
0	False	False	False	False	False
1	False	False	False	False	False
2	False	False	False	False	False
3	False	False	False	False	False
4	False	False	False	False	False
...	...	...	...	...	...
145	False	False	False	False	False
146	False	False	False	False	False
147	False	False	False	False	False
148	False	False	False	False	False
149	False	False	False	False	False

150 rows × 5 columns

In [24]:

```
df.notnull()
```

Out[24]:

	sepal_length	sepal_width	petal_length	petal_width	species
0	True	True	True	True	True
1	True	True	True	True	True
2	True	True	True	True	True
3	True	True	True	True	True
4	True	True	True	True	True
...	...	...	...	...	...
145	True	True	True	True	True
146	True	True	True	True	True
147	True	True	True	True	True
148	True	True	True	True	True
149	True	True	True	True	True

150 rows × 5 columns

In [25]:

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

Out[25]:

```
sepal_length    0
sepal_width     0
petal_length    0
petal_width     0
species         0
dtype: int64
```

In [26]:

```
df.iloc[5]
```

Out[26]:

```
sepal_length            5.4
sepal_width             3.9
petal_length            1.7
petal_width             0.4
species         Iris-setosa
Name: 5, dtype: object
```

In [27]:

```
df[0:5]
```

Out[27]:

	sepal_length	sepal_width	petal_length	petal_width	species
0	5.1	3.5	1.4	0.2	Iris-setosa
1	4.9	3.0	1.4	0.2	Iris-setosa
2	4.7	3.2	1.3	0.2	Iris-setosa
3	4.6	3.1	1.5	0.2	Iris-setosa
4	5.0	3.6	1.4	0.2	Iris-setosa

In [28]:

```
df.loc[:,["petal_length","petal_width"]]
```

Out[28]:

	petal_length	petal_width
0	1.4	0.2
1	1.4	0.2
2	1.3	0.2
3	1.5	0.2
4	1.4	0.2
...	...	...
145	5.2	2.3
146	5.0	1.9
147	5.2	2.0
148	5.4	2.3
149	5.1	1.8

150 rows × 2 columns

In [29]:

```
df.iloc[:199,:]
```

Out[29]:

	sepal_length	sepal_width	petal_length	petal_width	species
0	5.1	3.5	1.4	0.2	Iris-setosa
1	4.9	3.0	1.4	0.2	Iris-setosa
2	4.7	3.2	1.3	0.2	Iris-setosa
3	4.6	3.1	1.5	0.2	Iris-setosa
4	5.0	3.6	1.4	0.2	Iris-setosa
...	...	...	...	...	...
145	6.7	3.0	5.2	2.3	Iris-virginica
146	6.3	2.5	5.0	1.9	Iris-virginica
147	6.5	3.0	5.2	2.0	Iris-virginica
148	6.2	3.4	5.4	2.3	Iris-virginica
149	5.9	3.0	5.1	1.8	Iris-virginica

150 rows × 5 columns

In [30]:

```
df.iloc[3]
```

Out[30]:

```
sepal_length            4.6
sepal_width             3.1
petal_length            1.5
petal_width             0.2
species         Iris-setosa
Name: 3, dtype: object
```

In [31]:

```
df.iloc[3:5,0:2]
```

Out[31]:

	sepal_length	sepal_width
3	4.6	3.1
4	5.0	3.6

In [32]:

```
df.iloc[:,1:3]
```

Out[32]:

	sepal_width	petal_length
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 × 2 columns

In [33]:

```
df.iloc[1,2]
```

Out[33]:

```
1.4
```

In [34]:

```
df[["sepal_width","petal_width"]]
```

Out[34]:

	sepal_width	petal_width
0	3.5	0.2
1	3.0	0.2
2	3.2	0.2
3	3.1	0.2
4	3.6	0.2
...	...	...
145	3.0	2.3
146	2.5	1.9
147	3.0	2.0
148	3.4	2.3
149	3.0	1.8

150 rows × 2 columns

In [35]:

```
cols_2_4=df.columns[2:4]
cols_2_4=df[cols_2_4]
cols_2_4.iloc[5:10]
```

Out[35]:

	petal_length	petal_width
5	1.7	0.4
6	1.4	0.3
7	1.5	0.2
8	1.4	0.2
9	1.5	0.1

In [36]:

```
df[df.columns[2:4]].iloc[5:10]
```

Out[36]:

	petal_length	petal_width
5	1.7	0.4
6	1.4	0.3
7	1.5	0.2
8	1.4	0.2
9	1.5	0.1

In [37]:

```
df.describe(percentiles=None, include=None, exclude=None)
```

Out[37]:

	sepal_length	sepal_width	petal_length	petal_width
count	150.000000	150.000000	150.000000	150.000000
mean	5.843333	3.054000	3.758667	1.198667
std	0.828066	0.433594	1.764420	0.763161
min	4.300000	2.000000	1.000000	0.100000
25%	5.100000	2.800000	1.600000	0.300000
50%	5.800000	3.000000	4.350000	1.300000
75%	6.400000	3.300000	5.100000	1.800000
max	7.900000	4.400000	6.900000	2.500000

In [38]:

```
df
```

Out[38]:

	sepal_length	sepal_width	petal_length	petal_width	species
0	5.1	3.5	1.4	0.2	Iris-setosa
1	4.9	3.0	1.4	0.2	Iris-setosa
2	4.7	3.2	1.3	0.2	Iris-setosa
3	4.6	3.1	1.5	0.2	Iris-setosa
4	5.0	3.6	1.4	0.2	Iris-setosa
...	...	...	...	...	...
145	6.7	3.0	5.2	2.3	Iris-virginica
146	6.3	2.5	5.0	1.9	Iris-virginica
147	6.5	3.0	5.2	2.0	Iris-virginica
148	6.2	3.4	5.4	2.3	Iris-virginica
149	5.9	3.0	5.1	1.8	Iris-virginica

150 rows × 5 columns

In [39]:

```
df.fillna(0)
```

Out[39]:

	sepal_length	sepal_width	petal_length	petal_width	species
0	5.1	3.5	1.4	0.2	Iris-setosa
1	4.9	3.0	1.4	0.2	Iris-setosa
2	4.7	3.2	1.3	0.2	Iris-setosa
3	4.6	3.1	1.5	0.2	Iris-setosa
4	5.0	3.6	1.4	0.2	Iris-setosa
...	...	...	...	...	...
145	6.7	3.0	5.2	2.3	Iris-virginica
146	6.3	2.5	5.0	1.9	Iris-virginica
147	6.5	3.0	5.2	2.0	Iris-virginica
148	6.2	3.4	5.4	2.3	Iris-virginica
149	5.9	3.0	5.1	1.8	Iris-virginica

150 rows × 5 columns

In [40]:

```
df.mean()
```

```
C:\Users\91755\AppData\Local\Temp\ipykernel_22608\3698961737.py:1: FutureW
arning: The default value of numeric_only in DataFrame.mean is deprecated.
In a future version, it will default to False. In addition, specifying 'nu
meric_only=None' is deprecated. Select only valid columns or specify the v
alue of numeric_only to silence this warning.
  df.mean()
```

Out[40]:

```
sepal_length    5.843333
sepal_width     3.054000
petal_length    3.758667
petal_width     1.198667
dtype: float64
```

df.median()

df

In [41]:

```
df.median()
```

```
C:\Users\91755\AppData\Local\Temp\ipykernel_22608\530051474.py:1: FutureWa
rning: The default value of numeric_only in DataFrame.median is deprecate
d. In a future version, it will default to False. In addition, specifying
'numeric_only=None' is deprecated. Select only valid columns or specify th
e value of numeric_only to silence this warning.
  df.median()
```

Out[41]:

```
sepal_length    5.80
sepal_width     3.00
petal_length    4.35
petal_width     1.30
dtype: float64
```

In [42]:

```
df.mode()
```

Out[42]:

	sepal_length	sepal_width	petal_length	petal_width	species
0	5.0	3.0	1.5	0.2	Iris-setosa
1	NaN	NaN	NaN	NaN	Iris-versicolor
2	NaN	NaN	NaN	NaN	Iris-virginica

In [43]:

```
df.mode(axis=0)
```

Out[43]:

	sepal_length	sepal_width	petal_length	petal_width	species
0	5.0	3.0	1.5	0.2	Iris-setosa
1	NaN	NaN	NaN	NaN	Iris-versicolor
2	NaN	NaN	NaN	NaN	Iris-virginica

In [44]:

```
df['sepal_width']=df['sepal_width'].astype('int')
```

In [45]:

```
df
```

Out[45]:

	sepal_length	sepal_width	petal_length	petal_width	species
0	5.1	3	1.4	0.2	Iris-setosa
1	4.9	3	1.4	0.2	Iris-setosa
2	4.7	3	1.3	0.2	Iris-setosa
3	4.6	3	1.5	0.2	Iris-setosa
4	5.0	3	1.4	0.2	Iris-setosa
...	...	...	...	...	...
145	6.7	3	5.2	2.3	Iris-virginica
146	6.3	2	5.0	1.9	Iris-virginica
147	6.5	3	5.2	2.0	Iris-virginica
148	6.2	3	5.4	2.3	Iris-virginica
149	5.9	3	5.1	1.8	Iris-virginica

150 rows × 5 columns

In [46]:

```
df.drop(columns='sepal_width',inplace=True)
```

In [47]:

```
df
```

Out[47]:

	sepal_length	petal_length	petal_width	species
0	5.1	1.4	0.2	Iris-setosa
1	4.9	1.4	0.2	Iris-setosa
2	4.7	1.3	0.2	Iris-setosa
3	4.6	1.5	0.2	Iris-setosa
4	5.0	1.4	0.2	Iris-setosa
...	...	...	...	...
145	6.7	5.2	2.3	Iris-virginica
146	6.3	5.0	1.9	Iris-virginica
147	6.5	5.2	2.0	Iris-virginica
148	6.2	5.4	2.3	Iris-virginica
149	5.9	5.1	1.8	Iris-virginica

150 rows × 4 columns

In [48]:

```
df.head()
```

Out[48]:

	sepal_length	petal_length	petal_width	species
0	5.1	1.4	0.2	Iris-setosa
1	4.9	1.4	0.2	Iris-setosa
2	4.7	1.3	0.2	Iris-setosa
3	4.6	1.5	0.2	Iris-setosa
4	5.0	1.4	0.2	Iris-setosa

In [49]:

```
dummies = pd.get_dummies(df.species)
dummies.tail()
```

Out[49]:

	Iris-setosa	Iris-versicolor	Iris-virginica
145	0	0	1
146	0	0	1
147	0	0	1
148	0	0	1
149	0	0	1

In [50]:

```
merged_data = pd.concat([df,dummies], axis=1)
```

In [51]:

```
merged_data
```

Out[51]:

	sepal_length	petal_length	petal_width	species	Iris-setosa	Iris-versicolor	Iris-virginica
0	5.1	1.4	0.2	Iris-setosa	1	0	0
1	4.9	1.4	0.2	Iris-setosa	1	0	0
2	4.7	1.3	0.2	Iris-setosa	1	0	0
3	4.6	1.5	0.2	Iris-setosa	1	0	0
4	5.0	1.4	0.2	Iris-setosa	1	0	0
...	...	...	...	...	...	...	...
145	6.7	5.2	2.3	Iris-virginica	0	0	1
146	6.3	5.0	1.9	Iris-virginica	0	0	1
147	6.5	5.2	2.0	Iris-virginica	0	0	1
148	6.2	5.4	2.3	Iris-virginica	0	0	1
149	5.9	5.1	1.8	Iris-virginica	0	0	1

150 rows × 7 columns

In [52]:

```
final_data = merged_data.drop(columns='species')
```

In [53]:

```
final_data
```

Out[53]:

	sepal_length	petal_length	petal_width	Iris-setosa	Iris-versicolor	Iris-virginica
0	5.1	1.4	0.2	1	0	0
1	4.9	1.4	0.2	1	0	0
2	4.7	1.3	0.2	1	0	0
3	4.6	1.5	0.2	1	0	0
4	5.0	1.4	0.2	1	0	0
...	...	...	...	...	...	...
145	6.7	5.2	2.3	0	0	1
146	6.3	5.0	1.9	0	0	1
147	6.5	5.2	2.0	0	0	1
148	6.2	5.4	2.3	0	0	1
149	5.9	5.1	1.8	0	0	1

150 rows × 6 columns

In [54]:

```
final_data.head()
```

Out[54]:

	sepal_length	petal_length	petal_width	Iris-setosa	Iris-versicolor	Iris-virginica
0	5.1	1.4	0.2	1	0	0
1	4.9	1.4	0.2	1	0	0
2	4.7	1.3	0.2	1	0	0
3	4.6	1.5	0.2	1	0	0
4	5.0	1.4	0.2	1	0	0

In [ ]:

```

```

In [ ]:

```

```