In [7]:
```
import pandas as pd
import numpy as np
import seaborn as sb
```

In [8]:
```
sb.get_dataset_names()
```

Out[8]:
```
['anagrams',
 'anscombe',
 'attention',
 'brain_networks',
 'car_crashes',
 'diamonds',
 'dots',
 'dowjones',
 'exercise',
 'flights',
 'fmri',
 'geyser',
 'glue',
 'healthexp',
 'iris',
 'mpg',
 'penguins',
 'planets',
 'seaice',
```

In [10]:
```
df = sb.load_dataset("iris")
```

In [11]:
```
df
```

Out[11]:

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

150 rows × 5 columns

In [19]:
```
1 df.describe()
```

Out[19]:

	sepal_length	sepal_width	petal_length	petal_width
count	150.000000	150.000000	150.000000	150.000000
mean	5.843333	3.057333	3.758000	1.199333
std	0.828066	0.435866	1.765298	0.762238
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 [21]:
```
1 df.loc[:,'sepal_length'].mean()
```

Out[21]: 5.843333333333334

In [22]:
```
1 df.loc[:,'sepal_width'].mean()
```

Out[22]: 3.0573333333333337

In [23]:
```
1 df.loc[:,'petal_length'].mean()
```

Out[23]: 3.7580000000000005

In [24]:
```
1 df.loc[:,'petal_width'].mean()
```

Out[24]: 1.1993333333333336

In [25]:
```
1 df.loc[:,'sepal_length'].median()
```

Out[25]: 5.8

In [26]:
```
1 df.loc[:,'sepal_width'].median()
```

Out[26]: 3.0

In [27]:
```
1 df.loc[:,'petal_length'].median()
```

Out[27]: 4.35

In [28]:
```
1 df.loc[:,'petal_width'].median()
```

Out[28]: 1.3

In [29]:
```
df.loc[:,'sepal_length'].mode()
```

Out[29]:
```
0    5.0
Name: sepal_length, dtype: float64
```

In [30]:
```
df.loc[:,'sepal_width'].mode()
```

Out[30]:
```
0    3.0
Name: sepal_width, dtype: float64
```

In [31]:
```
df.loc[:,'petal_length'].mode()
```

Out[31]:
```
0    1.4
1    1.5
Name: petal_length, dtype: float64
```

In [36]:
```
df.loc[:,'petal_width'].mode()
```

Out[36]:
```
0    0.2
Name: petal_width, dtype: float64
```

In [33]:
```
df.loc[:,'sepal_length'].std()
```

Out[33]:
```
0.8280661279778629
```

In [34]:
```
df.loc[:,'sepal_width'].std()
```

Out[34]:
```
0.435866284936698
```

In [35]:
```
df.loc[:,'petal_length'].std()
```

Out[35]:
```
1.7652982332594667
```

In [37]:
```
df.loc[:,'petal_width'].std()
```

Out[37]:
```
0.7622376689603465
```

In [38]:
```
df.groupby(['species'])['sepal_length'].mean()
```

Out[38]:
```
species
setosa        5.006
versicolor    5.936
virginica     6.588
Name: sepal_length, dtype: float64
```

In [39]:
```
df.groupby(['species'])['sepal_width'].mean()
```

Out[39]:
```
species
setosa        3.428
versicolor    2.770
virginica     2.974
Name: sepal_width, dtype: float64
```

In [40]:
```
1 df.groupby(['species'])['petal_length'].mean()
```

Out[40]:
```
species
setosa        1.462
versicolor    4.260
virginica     5.552
Name: petal_length, dtype: float64
```

In [41]:
```
1 df.groupby(['species'])['petal_width'].mean()
```

Out[41]:
```
species
setosa        0.246
versicolor    1.326
virginica     2.026
Name: petal_width, dtype: float64
```

In [42]:
```
1 df.groupby(['species'])['sepal_length'].median()
```

Out[42]:
```
species
setosa        5.0
versicolor    5.9
virginica     6.5
Name: sepal_length, dtype: float64
```

In [43]:
```
1 df.groupby(['species'])['sepal_width'].median()
```

Out[43]:
```
species
setosa        3.4
versicolor    2.8
virginica     3.0
Name: sepal_width, dtype: float64
```

In [44]:
```
1 df.groupby(['species'])['petal_length'].median()
```

Out[44]:
```
species
setosa        1.50
versicolor    4.35
virginica     5.55
Name: petal_length, dtype: float64
```

In [45]:
```
1 df.groupby(['species'])['petal_width'].median()
```

Out[45]:
```
species
setosa        0.2
versicolor    1.3
virginica     2.0
Name: petal_width, dtype: float64
```

In [66]:
```
from sklearn import preprocessing
enc = preprocessing.OneHotEncoder()
enc_df = pd.DataFrame(enc.fit_transform(df[['species']]).toarray())
enc_df
```

Out[66]:

	0	1	2
0	1.0	0.0	0.0
1	1.0	0.0	0.0
2	1.0	0.0	0.0
3	1.0	0.0	0.0
4	1.0	0.0	0.0
...	...	...	...
145	0.0	0.0	1.0
146	0.0	0.0	1.0
147	0.0	0.0	1.0
148	0.0	0.0	1.0
149	0.0	0.0	1.0

150 rows × 3 columns

In [49]:
```
df101 = (df['species'] == 'setosa')
print(df101)
```

```
0       True
1       True
2       True
3       True
4       True
       ...
145    False
146    False
147    False
148    False
149    False
Name: species, Length: 150, dtype: bool
```

In [67]:
```
print("setosa")
print(df[df101].describe())
```

```
setosa
       sepal_length  sepal_width  petal_length  petal_width
count      50.00000    50.000000     50.000000    50.000000
mean        5.00600     3.428000      1.462000     0.246000
std         0.35249     0.379064      0.173664     0.105386
min         4.30000     2.300000      1.000000     0.100000
25%         4.80000     3.200000      1.400000     0.200000
50%         5.00000     3.400000      1.500000     0.200000
75%         5.20000     3.675000      1.575000     0.300000
max         5.80000     4.400000      1.900000     0.600000
```

In [68]:
```
df102 = (df['species'] == 'versicolor')
print("versicolor")
print(df[df102].describe())
```

```
versicolor
       sepal_length  sepal_width  petal_length  petal_width
count     50.000000    50.000000     50.000000    50.000000
mean       5.936000     2.770000      4.260000     1.326000
std        0.516171     0.313798      0.469911     0.197753
min        4.900000     2.000000      3.000000     1.000000
25%        5.600000     2.525000      4.000000     1.200000
50%        5.900000     2.800000      4.350000     1.300000
75%        6.300000     3.000000      4.600000     1.500000
max        7.000000     3.400000      5.100000     1.800000
```

In [69]:
```
df102 = (df['species'] == 'virginica')
print("virginica")
print(df[df102].describe())
```

```
virginica
       sepal_length  sepal_width  petal_length  petal_width
count      50.00000    50.000000     50.000000     50.00000
mean        6.58800     2.974000      5.552000      2.02600
std         0.63588     0.322497      0.551895      0.27465
min         4.90000     2.200000      4.500000      1.40000
25%         6.22500     2.800000      5.100000      1.80000
50%         6.50000     3.000000      5.550000      2.00000
75%         6.90000     3.175000      5.875000      2.30000
max         7.90000     3.800000      6.900000      2.50000
```

```
# Swayambhu Bhapkar TE_13121
```