anurag-dsbdal-pr3

February 21, 2024

[1]:
```
import pandas as pd
import numpy as np
import seaborn as sb
```

[3]:
```
datanames=sb.get_dataset_names()
print(datanames)
```

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

[4]:
```
df = sb.load_dataset("iris")
df
```

[4]:

	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 x 5 columns]

[5]:
```
df.describe()
```

[5]:

	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
```

[6]:
```
df.loc[:,'sepal_length'].mean()
```

[6]:
```
5.843333333333334
```

[7]:
```
df.loc[:,"sepal_width"].mean()
```

[7]:
```
3.0573333333333337
```

[8]:
```
df.loc[:,'petal_length'].mean()
```

[8]:
```
3.7580000000000005
```

[9]:
```
df.loc[:,'petal_width'].mean()
```

[9]:
```
1.1993333333333336
```

[10]:
```
df.loc[:,'sepal_length'].mode()
```

[10]:
```
0    5.0
Name: sepal_length, dtype: float64
```

[11]:
```
df.loc[:,'sepal_length'].median()
```

[11]:
```
5.8
```

[12]:
```
df.loc[:,"sepal_width"].mode()
```

[12]:
```
0    3.0
Name: sepal_width, dtype: float64
```

[13]:
```
df.loc[:,"sepal_width"].median()
```

[13]:
```
3.0
```

[14]:
```
df.loc[:,'petal_length'].mode()
```

[14]:
```
0    1.4
1    1.5
Name: petal_length, dtype: float64
```

[15]:
```
df.loc[:,'petal_length'].median()
```

[15]:
```
4.35
```

[16]:
```
df.loc[:,'petal_width'].mode()
```

[16]:
```
0    0.2
Name: petal_width, dtype: float64
```

[17]:
```
df.loc[:,'petal_width'].median()
```

[17]:
```
1.3
```

[18]:
```
df.loc[:,'sepal_length'].std()
```

[18]:
```
0.8280661279778629
```

[19]:
```
df.loc[:,'sepal_width'].std()
```

[19]:
```
0.435866284936698
```

[20]:
```
df.loc[:,'petal_length'].std()
```

[20]:
```
1.7652982332594667
```

[21]:
```
df.loc[:,'petal_width'].std()
```

[21]:
```
0.7622376689603465
```

[22]:
```
df.groupby(['species'])['sepal_length'].mean()
```

[22]:
```
species
setosa        5.006
versicolor    5.936
virginica     6.588
Name: sepal_length, dtype: float64
```

[23]:
```
df.groupby(['species'])['sepal_width'].mean()
```

[23]:
```
species
setosa        3.428
versicolor    2.770
virginica     2.974
Name: sepal_width, dtype: float64
```

[24]:
```
df.groupby(['species'])['petal_length'].mean()
```

[24]:
```
species
setosa        1.462
versicolor    4.260
virginica     5.552
```

```
Name: petal_length, dtype: float64
```

[25]:
```
df.groupby(['species'])['sepal_width'].mean()
```

[25]:
```
species
setosa        3.428
versicolor    2.770
virginica     2.974
Name: sepal_width, dtype: float64
```

[26]:
```
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
```

[27]:
```
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
```

[28]:
```
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
```

[29]:
```
df103 = (df['species'] == 'virginica')
print("virginica")
print(df[df103].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
```

[ ]:
```
Name-Anurag Jadhav
Roll No-13171
Div-A
```