practical-10

May 9, 2025

```python
[3]: import pandas as pd
```

```python
[4]: data = pd.read_csv("iris.csv")
```

```python
[5]: data.head()
```

```
[5]:    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
```

```python
[6]: data.sample(10)
```

```
[6]:      Id  SepalLengthCm  SepalWidthCm  PetalLengthCm  PetalWidthCm  \
    53    54            5.5           2.3            4.0           1.3
    137  138            6.4           3.1            5.5           1.8
    11    12            4.8           3.4            1.6           0.2
    15    16            5.7           4.4            1.5           0.4
    26    27            5.0           3.4            1.6           0.4
    109  110            7.2           3.6            6.1           2.5
    116  117            6.5           3.0            5.5           1.8
    111  112            6.4           2.7            5.3           1.9
    97    98            6.2           2.9            4.3           1.3
    37    38            4.9           3.1            1.5           0.1

             Species
    53   Iris-versicolor
    137   Iris-virginica
    11       Iris-setosa
    15       Iris-setosa
    26       Iris-setosa
    109   Iris-virginica
    116   Iris-virginica
    111   Iris-virginica
    97   Iris-versicolor
```

```
       37         Iris-setosa
```

[7]: `data.columns`

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

[8]: `data.shape`

[8]: (150, 6)

[9]: `print(data)`

```
        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
..     ...            ...           ...            ...           ...
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
0       Iris-setosa
1       Iris-setosa
2       Iris-setosa
3       Iris-setosa
4       Iris-setosa
..              ...
145  Iris-virginica
146  Iris-virginica
147  Iris-virginica
148  Iris-virginica
149  Iris-virginica

[150 rows x 6 columns]
```

[10]: `print(data[10:21])`

```
     Id  SepalLengthCm  SepalWidthCm  PetalLengthCm  PetalWidthCm      Species
10   11            5.4           3.7            1.5           0.2  Iris-setosa
11   12            4.8           3.4            1.6           0.2  Iris-setosa
12   13            4.8           3.0            1.4           0.1  Iris-setosa
```

```
13   14           4.3          3.0          1.1          0.1  Iris-setosa
14   15           5.8          4.0          1.2          0.2  Iris-setosa
15   16           5.7          4.4          1.5          0.4  Iris-setosa
16   17           5.4          3.9          1.3          0.4  Iris-setosa
17   18           5.1          3.5          1.4          0.3  Iris-setosa
18   19           5.7          3.8          1.7          0.3  Iris-setosa
19   20           5.1          3.8          1.5          0.3  Iris-setosa
20   21           5.4          3.4          1.7          0.2  Iris-setosa
```

[11]:
```
sliced_data=data[10:21]
print(sliced_data)
```

```
     Id  SepalLengthCm  SepalWidthCm  PetalLengthCm  PetalWidthCm      Species
10   11            5.4           3.7            1.5           0.2  Iris-setosa
11   12            4.8           3.4            1.6           0.2  Iris-setosa
12   13            4.8           3.0            1.4           0.1  Iris-setosa
13   14            4.3           3.0            1.1           0.1  Iris-setosa
14   15            5.8           4.0            1.2           0.2  Iris-setosa
15   16            5.7           4.4            1.5           0.4  Iris-setosa
16   17            5.4           3.9            1.3           0.4  Iris-setosa
17   18            5.1           3.5            1.4           0.3  Iris-setosa
18   19            5.7           3.8            1.7           0.3  Iris-setosa
19   20            5.1           3.8            1.5           0.3  Iris-setosa
20   21            5.4           3.4            1.7           0.2  Iris-setosa
```

[23]:
```
specific_data=data[["Species"]]
print(specific_data.head(10))
```

```
       Species
0  Iris-setosa
1  Iris-setosa
2  Iris-setosa
3  Iris-setosa
4  Iris-setosa
5  Iris-setosa
6  Iris-setosa
7  Iris-setosa
8  Iris-setosa
9  Iris-setosa
```

[24]:
```
data.iloc[5]
```

[24]:
```
Id                       6
SepalLengthCm          5.4
SepalWidthCm           3.9
PetalLengthCm          1.7
PetalWidthCm           0.4
Species        Iris-setosa
```

```
        total_values        17.4
        Name: 5, dtype: object
```

[26]: `data.loc[data["Species"] == "Iris-setosa"]`

[26]:
	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	
5	6	5.4	3.9	1.7	0.4	Iris-setosa	
6	7	4.6	3.4	1.4	0.3	Iris-setosa	
7	8	5.0	3.4	1.5	0.2	Iris-setosa	
8	9	4.4	2.9	1.4	0.2	Iris-setosa	
9	10	4.9	3.1	1.5	0.1	Iris-setosa	
10	11	5.4	3.7	1.5	0.2	Iris-setosa	
11	12	4.8	3.4	1.6	0.2	Iris-setosa	
12	13	4.8	3.0	1.4	0.1	Iris-setosa	
13	14	4.3	3.0	1.1	0.1	Iris-setosa	
14	15	5.8	4.0	1.2	0.2	Iris-setosa	
15	16	5.7	4.4	1.5	0.4	Iris-setosa	
16	17	5.4	3.9	1.3	0.4	Iris-setosa	
17	18	5.1	3.5	1.4	0.3	Iris-setosa	
18	19	5.7	3.8	1.7	0.3	Iris-setosa	
19	20	5.1	3.8	1.5	0.3	Iris-setosa	
20	21	5.4	3.4	1.7	0.2	Iris-setosa	
21	22	5.1	3.7	1.5	0.4	Iris-setosa	
22	23	4.6	3.6	1.0	0.2	Iris-setosa	
23	24	5.1	3.3	1.7	0.5	Iris-setosa	
24	25	4.8	3.4	1.9	0.2	Iris-setosa	
25	26	5.0	3.0	1.6	0.2	Iris-setosa	
26	27	5.0	3.4	1.6	0.4	Iris-setosa	
27	28	5.2	3.5	1.5	0.2	Iris-setosa	
28	29	5.2	3.4	1.4	0.2	Iris-setosa	
29	30	4.7	3.2	1.6	0.2	Iris-setosa	
30	31	4.8	3.1	1.6	0.2	Iris-setosa	
31	32	5.4	3.4	1.5	0.4	Iris-setosa	
32	33	5.2	4.1	1.5	0.1	Iris-setosa	
33	34	5.5	4.2	1.4	0.2	Iris-setosa	
34	35	4.9	3.1	1.5	0.1	Iris-setosa	
35	36	5.0	3.2	1.2	0.2	Iris-setosa	
36	37	5.5	3.5	1.3	0.2	Iris-setosa	
37	38	4.9	3.1	1.5	0.1	Iris-setosa	
38	39	4.4	3.0	1.3	0.2	Iris-setosa	
39	40	5.1	3.4	1.5	0.2	Iris-setosa	
40	41	5.0	3.5	1.3	0.3	Iris-setosa	

41	42	4.5	2.3	1.3	0.3	Iris-setosa
42	43	4.4	3.2	1.3	0.2	Iris-setosa
43	44	5.0	3.5	1.6	0.6	Iris-setosa
44	45	5.1	3.8	1.9	0.4	Iris-setosa
45	46	4.8	3.0	1.4	0.3	Iris-setosa
46	47	5.1	3.8	1.6	0.2	Iris-setosa
47	48	4.6	3.2	1.4	0.2	Iris-setosa
48	49	5.3	3.7	1.5	0.2	Iris-setosa
49	50	5.0	3.3	1.4	0.2	Iris-setosa

	total_values
0	11.2
1	11.5
2	12.4
3	13.4
4	15.2
5	17.4
6	16.7
7	18.1
8	17.9
9	19.6
10	21.8
11	22.0
12	22.3
13	22.5
14	26.2
15	28.0
16	28.0
17	28.3
18	30.5
19	30.7
20	31.7
21	32.7
22	32.4
23	34.6
24	35.3
25	35.8
26	37.4
27	38.4
28	39.2
29	39.7
30	40.7
31	42.7
32	43.9
33	45.3
34	44.6
35	45.6

```
36          47.5
37          47.6
38          47.9
39          50.2
40          51.1
41          50.4
42          52.1
43          54.7
44          56.2
45          55.5
46          57.7
47          57.4
48          59.7
49          59.9
```

[27]: `data["Species"].value_counts()`

[27]:
```
Species
Iris-setosa       50
Iris-versicolor   50
Iris-virginica    50
Name: count, dtype: int64
```

[28]:
```python
sum_data = data["SepalLengthCm"].sum()
mean_data = data["SepalLengthCm"].mean()
median_data = data["SepalLengthCm"].median()
print("Sum:",sum_data, "\nMean:", mean_data, "\nMedian:",median_data)
```

```
Sum: 876.5
Mean: 5.843333333333334
Median: 5.8
```

[29]:
```python
min_data=data["SepalLengthCm"].min()
max_data=data["SepalLengthCm"].max()
print("Minimum:",min_data, "\nMaximum:", max_data)
```

```
Minimum: 4.3
Maximum: 7.9
```

[30]:
```python
cols = data.columns
print(cols)
```

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

```
[31]: data1 = data[cols].select_dtypes(include='number')
      # Create the new 'total_values' column by summing the numeric values
      data["total_values"] = data1.sum(axis=1)
```

```
[32]: newcols={ "Id":"id",
      "SepalLengthCmh":"SepalLength",
      "SepalWidthCm":"SepalWidth"}
      data.rename(columns=newcols,inplace=True)
```

```
[33]: print(data.head())
```

```
   id  SepalLengthCm  SepalWidth  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

   total_values
0          22.4
1          23.0
2          24.8
3          26.8
4          30.4
```

```
[34]: data.style
```

```
[34]: <pandas.io.formats.style.Styler at 0x22abac1ef90>
```

```
[36]: # Highlight max values only in numeric columns
      data.style.highlight_max(subset=data.select_dtypes(include='number').columns)
```

```
[36]: <pandas.io.formats.style.Styler at 0x22abacb20d0>
```

```
[37]: data.isnull()
```

```
[37]:        id  SepalLengthCm  SepalWidth  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

```
      total_values
0             False
1             False
2             False
3             False
4             False
..              …
145           False
146           False
147           False
148           False
149           False

[150 rows x 7 columns]
```
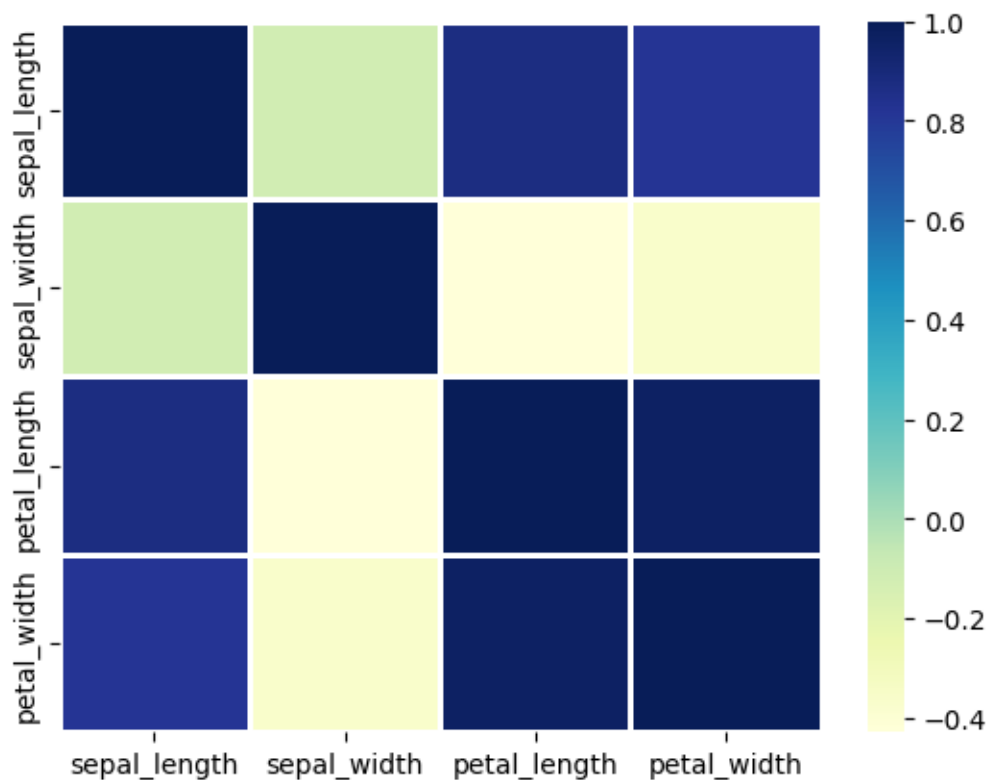
[38]:
```python
null_values = data.isnull().sum()
print(null_values)
```

```
id               0
SepalLengthCm    0
SepalWidth       0
PetalLengthCm    0
PetalWidthCm     0
Species          0
total_values     0
dtype: int64
```

[39]:
```python
import ssl
import seaborn as sns
ssl._create_default_https_context = ssl._create_unverified_context
iris = sns.load_dataset("iris")
sns.heatmap(iris.select_dtypes(include='number').corr(), cmap="YlGnBu",
linecolor='white', linewidths=1)
```

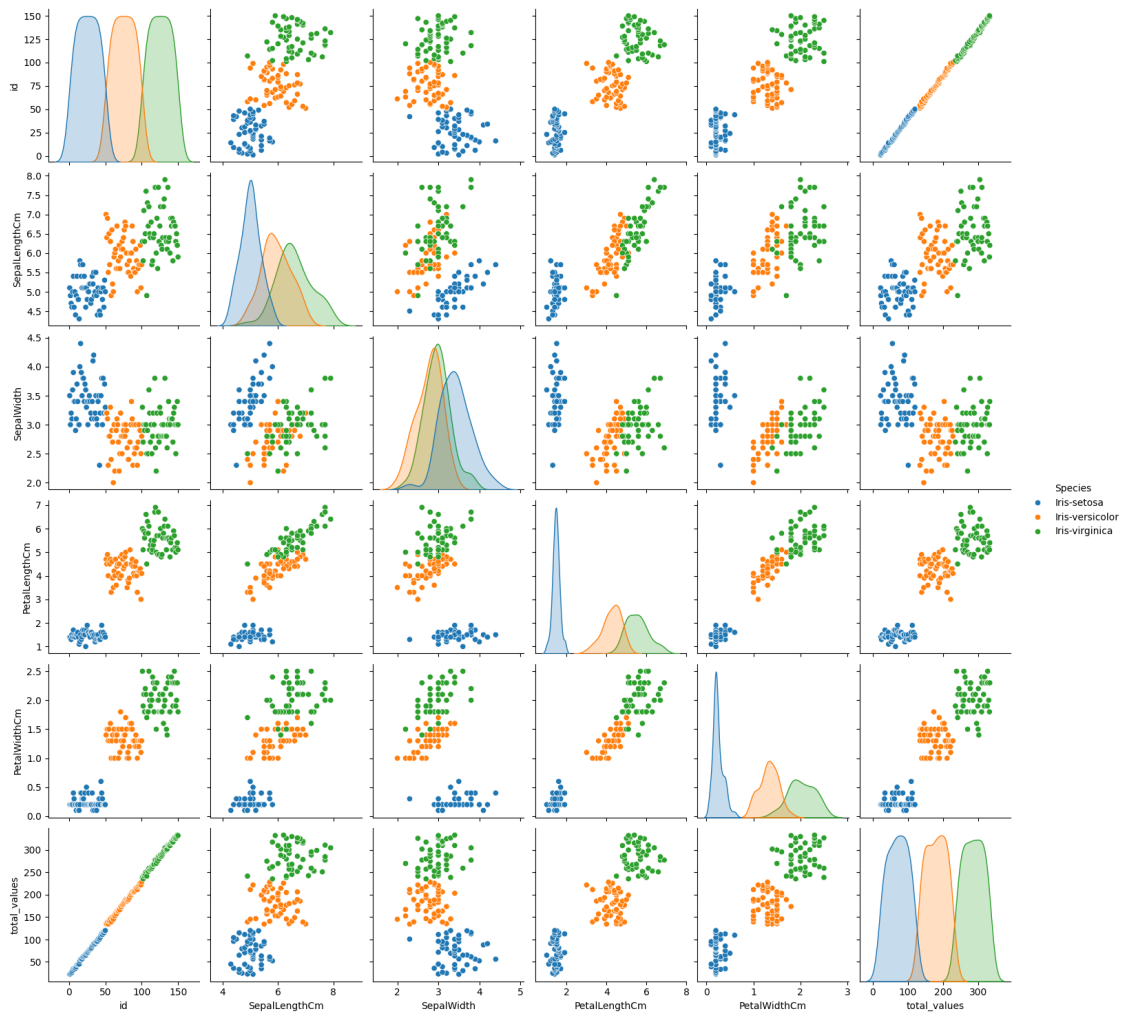[39]: <Axes: >

```
[40]: data.select_dtypes(include='number').corr(method='pearson')
```

```
[40]:                      id  SepalLengthCm  SepalWidth  PetalLengthCm  \
      id             1.000000       0.716676   -0.397729       0.882747
      SepalLengthCm  0.716676       1.000000   -0.109369       0.871754
      SepalWidth    -0.397729      -0.109369    1.000000      -0.420516
      PetalLengthCm  0.882747       0.871754   -0.420516       1.000000
      PetalWidthCm   0.899759       0.817954   -0.356544       0.962757
      total_values   0.999370       0.738612   -0.389048       0.897059

                     PetalWidthCm  total_values
      id                 0.899759      0.999370
      SepalLengthCm      0.817954      0.738612
      SepalWidth        -0.356544     -0.389048
      PetalLengthCm      0.962757      0.897059
      PetalWidthCm       1.000000      0.911904
      total_values       0.911904      1.000000
```

```
[41]: g=sns.pairplot(data,hue="Species")
```

[ ]: