File   Edit   View   Run   Kernel   Settings   Help

Trusted

Code ∨

JupyterLab   Python 3 (ipykernel)

```python
[84]: #importing the libraries required.
      import pandas as pd
      import seaborn as sns
      import matplotlib.pyplot as plt
      import numpy as np
```

```python
[85]: #Opening the file or dataset as a CVS file.
      dataset = pd.read_csv(r"C:\Users\Elias\Downloads\Compressed\archive_3\Iris.csv")
```

```python
[86]: dataset.head(10)#Getting the first 10 rows of our data set
```

[86]:
	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

```python
[87]: dataset.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 150 entries, 0 to 149
Data columns (total 6 columns):
 #   Column         Non-Null Count  Dtype
---  ------         --------------  -----
 0   Id             150 non-null    int64
 1   SepalLengthCm  150 non-null    float64
 2   SepalWidthCm   150 non-null    float64
 3   PetalLengthCm  150 non-null    float64
 4   PetalWidthCm   150 non-null    float64
 5   Species        150 non-null    object
dtypes: float64(4), int64(1), object(1)
memory usage: 7.2+ KB
```

```python
[88]: dataset
```

[88]:
	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
...	...	...	...	...	...	...
145	146	6.7	3.0	5.2	2.3	Iris-virginica
146	147	6.3	2.5	5.0	1.9	Iris-virginica
147	148	6.5	3.0	5.2	2.0	Iris-virginica
148	149	6.2	3.4	5.4	2.3	Iris-virginica
149	150	5.9	3.0	5.1	1.8	Iris-virginica

150 rows × 6 columns

```python
[89]: dataset.shape
```

[89]: (150, 6)

```python
[90]: dataset.describe()
```

[90]:
	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

[91]:
```python
#Checking for Total Missing Values in our file or dataset
dataset.isnull().sum()
```

[91]:
```
Id               0
SepalLengthCm    0
SepalWidthCm     0
PetalLengthCm    0
PetalWidthCm     0
Species          0
dtype: int64
```
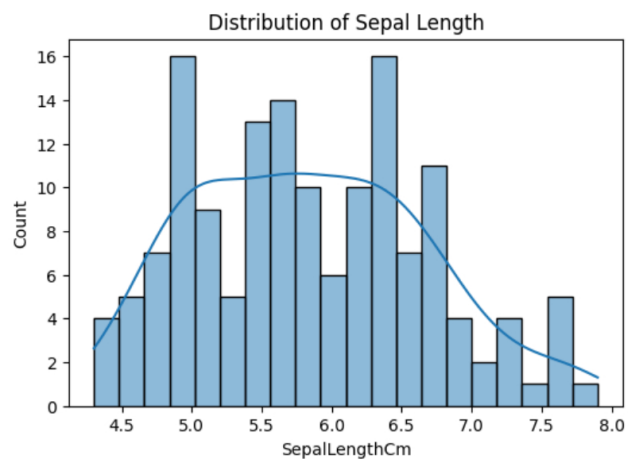
[92]:
```python
#Checking for the last 20 rows in our dataset
dataset.tail(20)
```
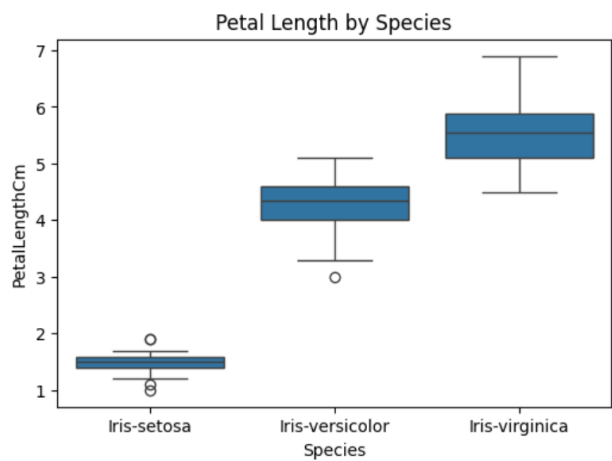
[92]:

	Id	SepalLengthCm	SepalWidthCm	PetalLengthCm	PetalWidthCm	Species
130	131	7.4	2.8	6.1	1.9	Iris-virginica
131	132	7.9	3.8	6.4	2.0	Iris-virginica
132	133	6.4	2.8	5.6	2.2	Iris-virginica
133	134	6.3	2.8	5.1	1.5	Iris-virginica
134	135	6.1	2.6	5.6	1.4	Iris-virginica
135	136	7.7	3.0	6.1	2.3	Iris-virginica
136	137	6.3	3.4	5.6	2.4	Iris-virginica
137	138	6.4	3.1	5.5	1.8	Iris-virginica
138	139	6.0	3.0	4.8	1.8	Iris-virginica
139	140	6.9	3.1	5.4	2.1	Iris-virginica
140	141	6.7	3.1	5.6	2.4	Iris-virginica
141	142	6.9	3.1	5.1	2.3	Iris-virginica
142	143	5.8	2.7	5.1	1.9	Iris-virginica
143	144	6.8	3.2	5.9	2.3	Iris-virginica
144	145	6.7	3.3	5.7	2.5	Iris-virginica
145	146	6.7	3.0	5.2	2.3	Iris-virginica
146	147	6.3	2.5	5.0	1.9	Iris-virginica
147	148	6.5	3.0	5.2	2.0	Iris-virginica
148	149	6.2	3.4	5.4	2.3	Iris-virginica
149	150	5.9	3.0	5.1	1.8	Iris-virginica

[93]:
```python
# Histogram of sepal length
plt.figure(figsize=(6,4))
sns.histplot(dataset["SepalLengthCm"], kde=True, bins=20)
plt.title("Distribution of Sepal Length")
plt.show()

# Boxplot for petal length by species
plt.figure(figsize=(6,4))
sns.boxplot(x="Species", y="PetalLengthCm", data=dataset)
plt.title("Petal Length by Species")
plt.show()
```
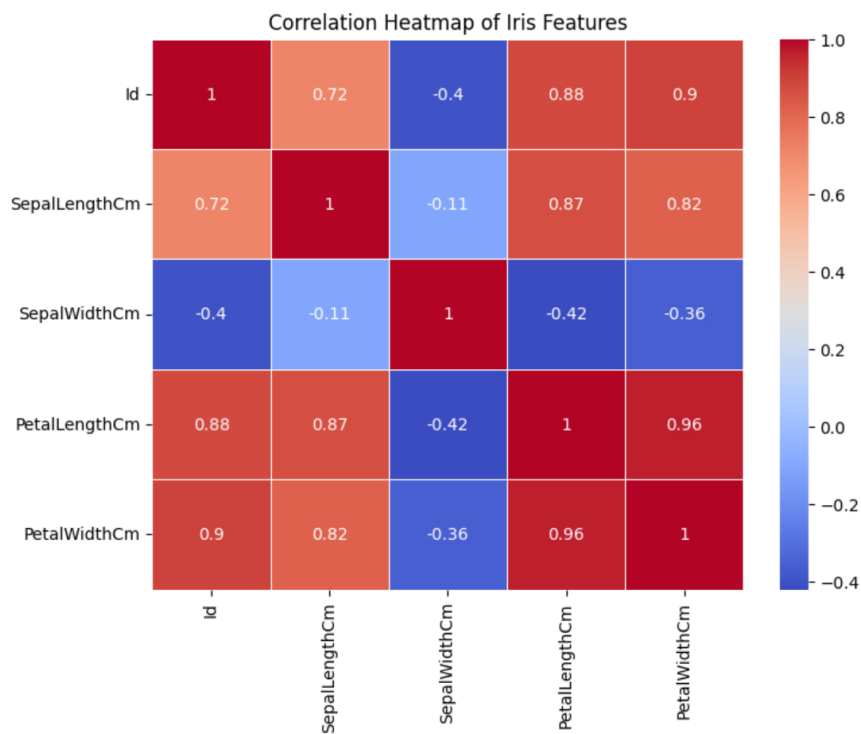
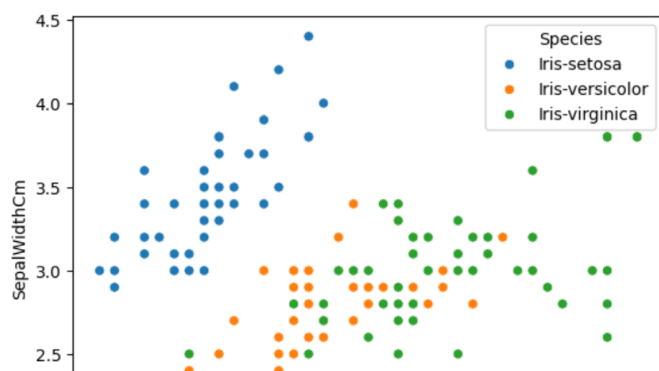Petal Length by Species

[96]:
```
# Compute correlation matrix
corr = dataset.corr(numeric_only=True)

# Plot heatmap
plt.figure(figsize=(8,6))
sns.heatmap(corr, annot=True, cmap="coolwarm", linewidths=0.5)
plt.title("Correlation Heatmap of Iris Features")
plt.show()
```
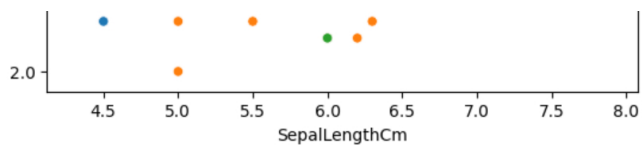

Correlation Heatmap of Iris Features

[100]:
```
# Scatter plot by species
sns.scatterplot(x='SepalLengthCm', y='SepalWidthCm', hue='Species', data=dataset)
plt.show()
```

2.0

4.5   5.0   5.5   6.0   6.5   7.0   7.5   8.0
SepalLengthCm

[98]:

[ ]: