Iris DTC

January 13, 2022

0.1 Data visualisatie

[49]:
```
import pandas as pd
from sklearn.tree import DecisionTreeClassifier
from sklearn.model_selection import train_test_split
from sklearn.metrics import accuracy_score
import matplotlib.pyplot as plt
import numpy as np
import seaborn as sns
```

[51]:
```
iris_data = pd.read_csv("./Iris/Iris.csv")
```

[52]:
```
iris_data
```

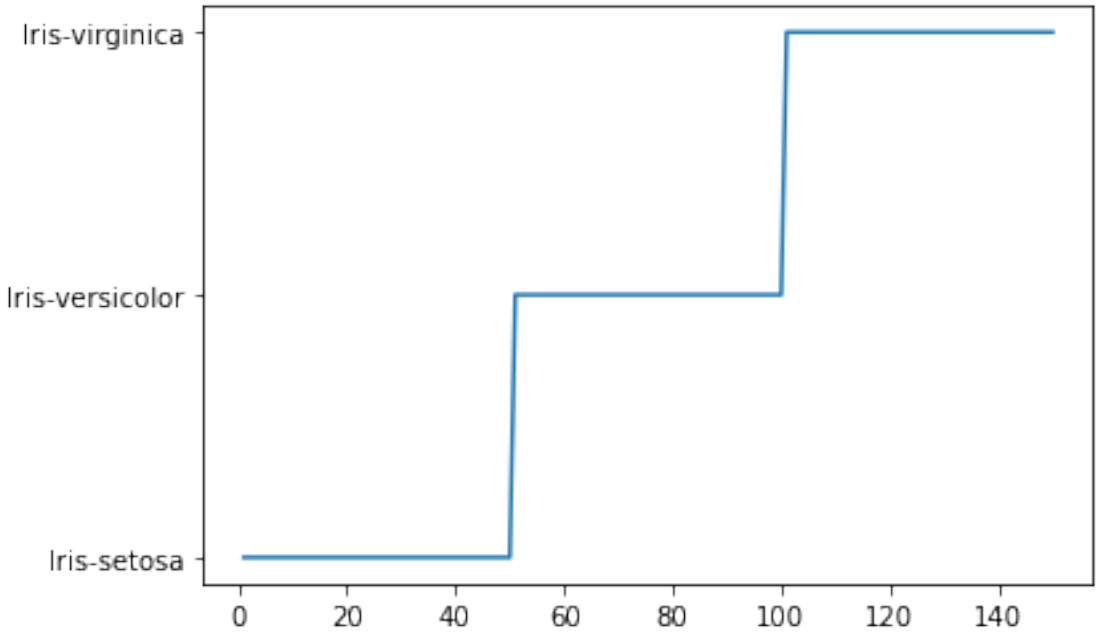
[52]:
```
      Id  SepalLengthCm  SepalWidthCm  PetalLengthCm  PetalWidthCm  \
0      1            5.1           3.5            1.4           0.2
1      2            4.9           3.0            1.4           0.2
2      3            4.7           3.2            1.3           0.2
3      4            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]
```

[73]:
```
plt.plot(iris_data.Id, iris_data["Species"])
plt.show()
```

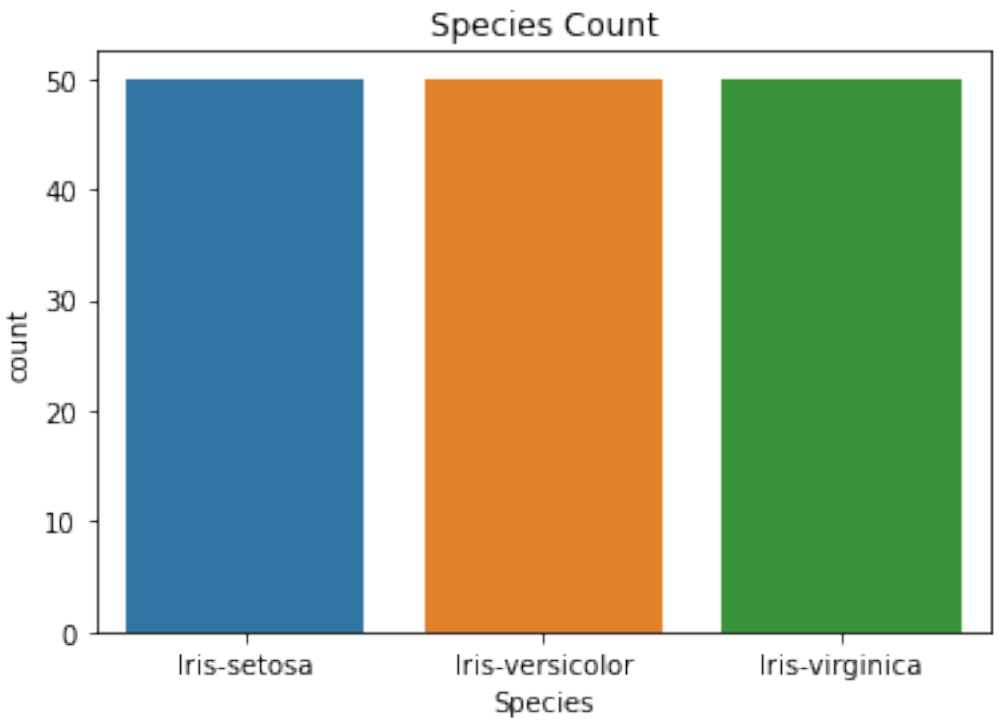


Use Seaborn

[58]:
```
sns.countplot(iris_data["Species"])
plt.title("Species Count")
```

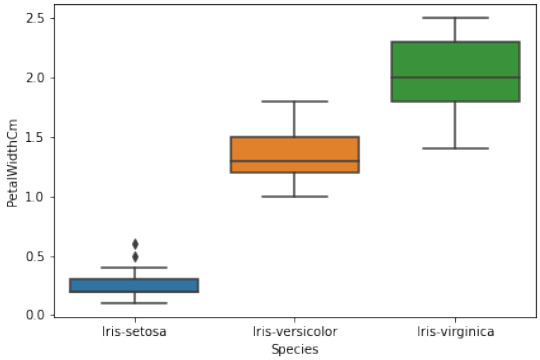
```
/opt/jupyterhub/anaconda/lib/python3.8/site-packages/seaborn/_decorators.py:36:
FutureWarning: Pass the following variable as a keyword arg: x. From version
0.12, the only valid positional argument will be `data`, and passing other
arguments without an explicit keyword will result in an error or
misinterpretation.
  warnings.warn(
```

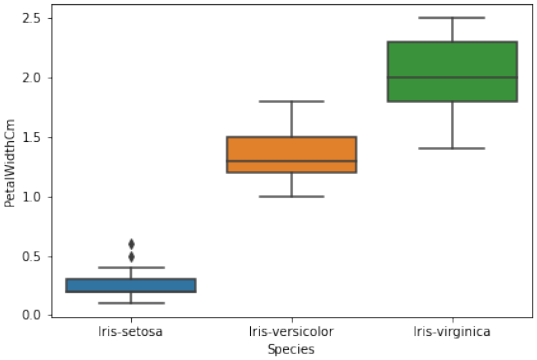
[58]:
```
Text(0.5, 1.0, 'Species Count')
```

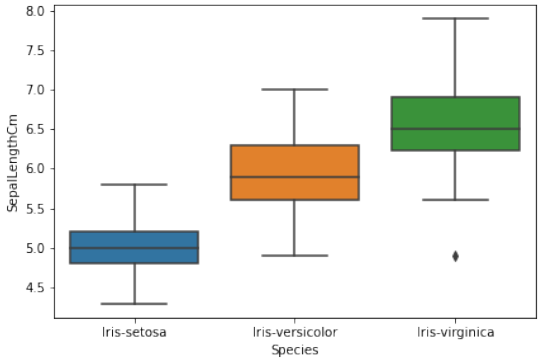


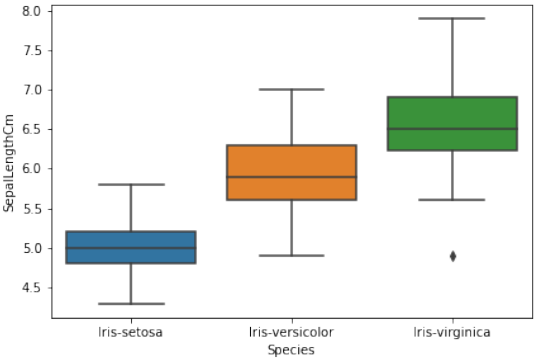
Boxplot

[65]:
```
fig, axes = plt.subplots(2, 2, figsize=(15,))
sns.boxplot( y="PetalWidthCm", x= "Species", data=iris_data, orient='v' ,␣
 ↪ax=axes[0, 0])
sns.boxplot( y="PetalWidthCm", x= "Species", data=iris_data, orient='v' ,␣
 ↪ax=axes[0, 1])
sns.boxplot( y="SepalLengthCm", x= "Species", data=iris_data, orient='v' ,␣
 ↪ax=axes[1, 0])
sns.boxplot( y="SepalLengthCm", x= "Species", data=iris_data, orient='v' ,␣
 ↪ax=axes[1, 1])
plt.show()
```


PetalWidthCm
2.5
2.0
1.5
1.0
0.5
0.0
Iris-setosa
Iris-versicolor
Iris-virginica
Species


PetalWidthCm
2.5
2.0
1.5
1.0
0.5
0.0
Iris-setosa
Iris-versicolor
Iris-virginica
Species


SepalLengthCm
8.0
7.5
7.0
6.5
6.0
5.5
5.0
4.5
Iris-setosa
Iris-versicolor
Iris-virginica
Species


SepalLengthCm
8.0
7.5
7.0
6.5
6.0
5.5
5.0
4.5
Iris-setosa
Iris-versicolor
Iris-virginica
Species