Name : Suryal D . Khirade
Roll NO: T190424399
Assignment No :10

##Data Visualization III Download the Iris flower dataset or any other dataset into a DataFrame. (e.g., https://archive.ics.uci.edu/ml/datasets/Iris (https://archive.ics.uci.edu/ml/datasets/Iris)). Scan the dataset and give the inference as:

List down the features and their types (e.g., numeric, nominal) available in the dataset. Create a histogram for each feature in the dataset to illustrate the feature distributions. Create a boxplot for each feature in the dataset. Compare distributions and identify outliers.

In [4]:
```python
import pandas as pd
import numpy as np
import seaborn as sns
import matplotlib.pyplot as plt
```

In [3]:
```python
iris = sns.load_dataset("iris")
```

In [5]:
```python
iris
```

Out[5]:

	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

1.List down the features and their types (e.g., numeric, nominal) available in the dataset

In [8]: `iris.info()`

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 150 entries, 0 to 149
Data columns (total 5 columns):
 #   Column        Non-Null Count  Dtype
---  ------        --------------  -----
 0   sepal_length  150 non-null    float64
 1   sepal_width   150 non-null    float64
 2   petal_length  150 non-null    float64
 3   petal_width   150 non-null    float64
 4   species       150 non-null    object
dtypes: float64(4), object(1)
memory usage: 6.0+ KB
```

In [7]: `iris.describe()`

Out[7]:

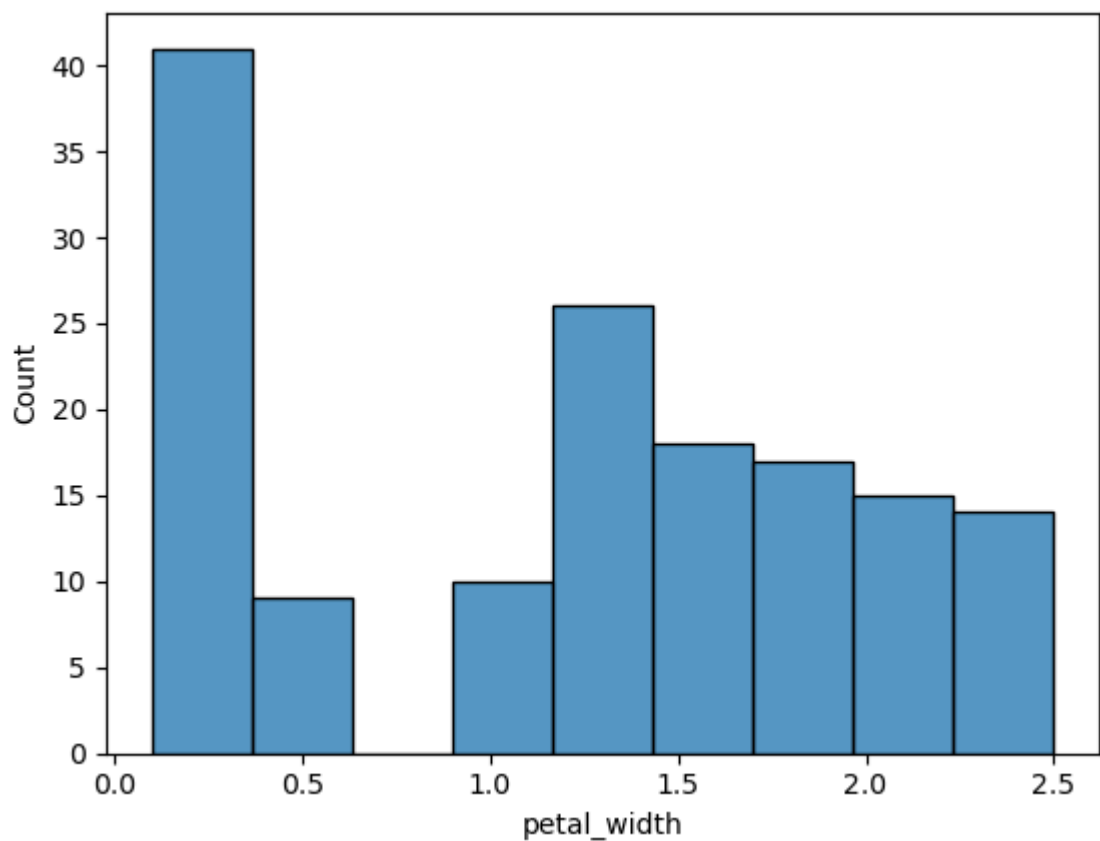|	sepal_length	sepal_width	petal_length	petal_width
count	150.000000	150.000000	150.000000	150.000000
mean	5.843333	3.057333	3.758000	1.199333
std	0.828066	0.435866	1.765298	0.762238
min	4.300000	2.000000	1.000000	0.100000
25%	5.100000	2.800000	1.600000	0.300000
50%	5.800000	3.000000	4.350000	1.300000
75%	6.400000	3.300000	5.100000	1.800000
max	7.900000	4.400000	6.900000	2.500000

In [10]: `iris.shape`

Out[10]: `(150, 5)`

In [11]: `iris.species.unique()`

Out[11]: `array(['setosa', 'versicolor', 'virginica'], dtype=object)`

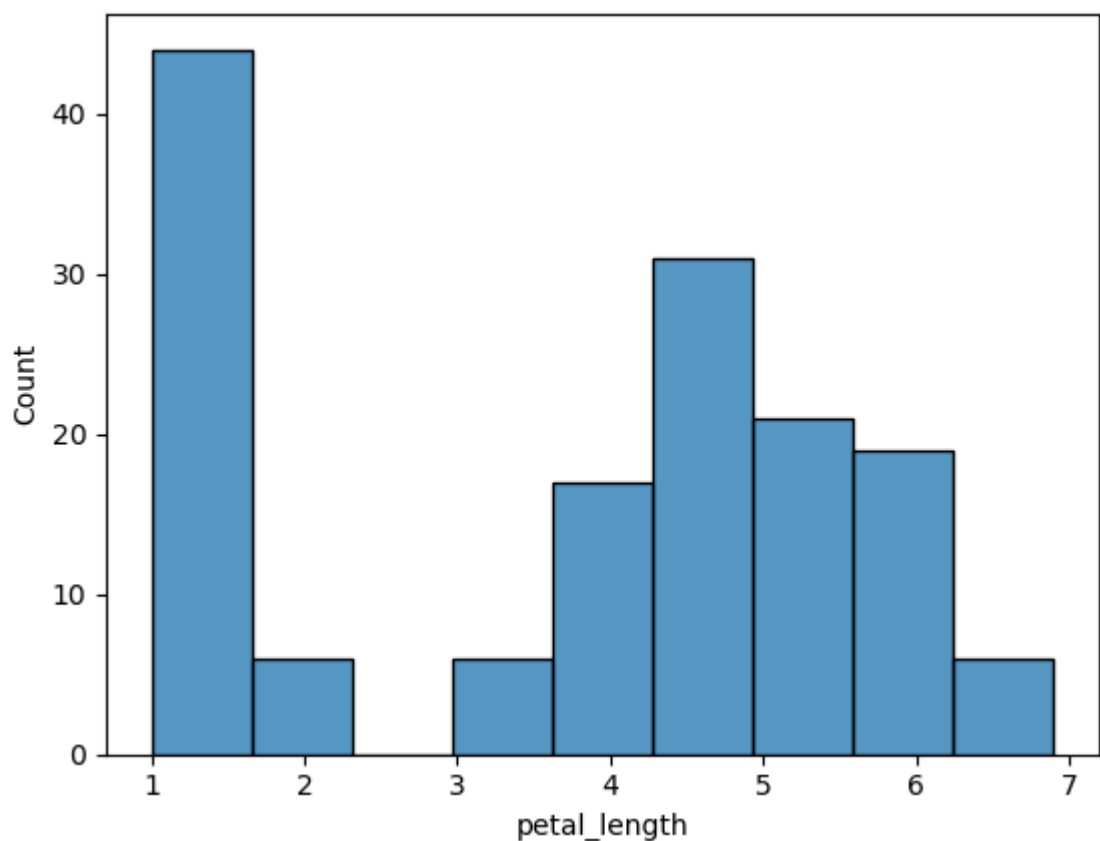2.Create a histogram for each feature in the dataset to illustrate the feature distributions.

In [12]: `sns.histplot(iris.petal_width)`
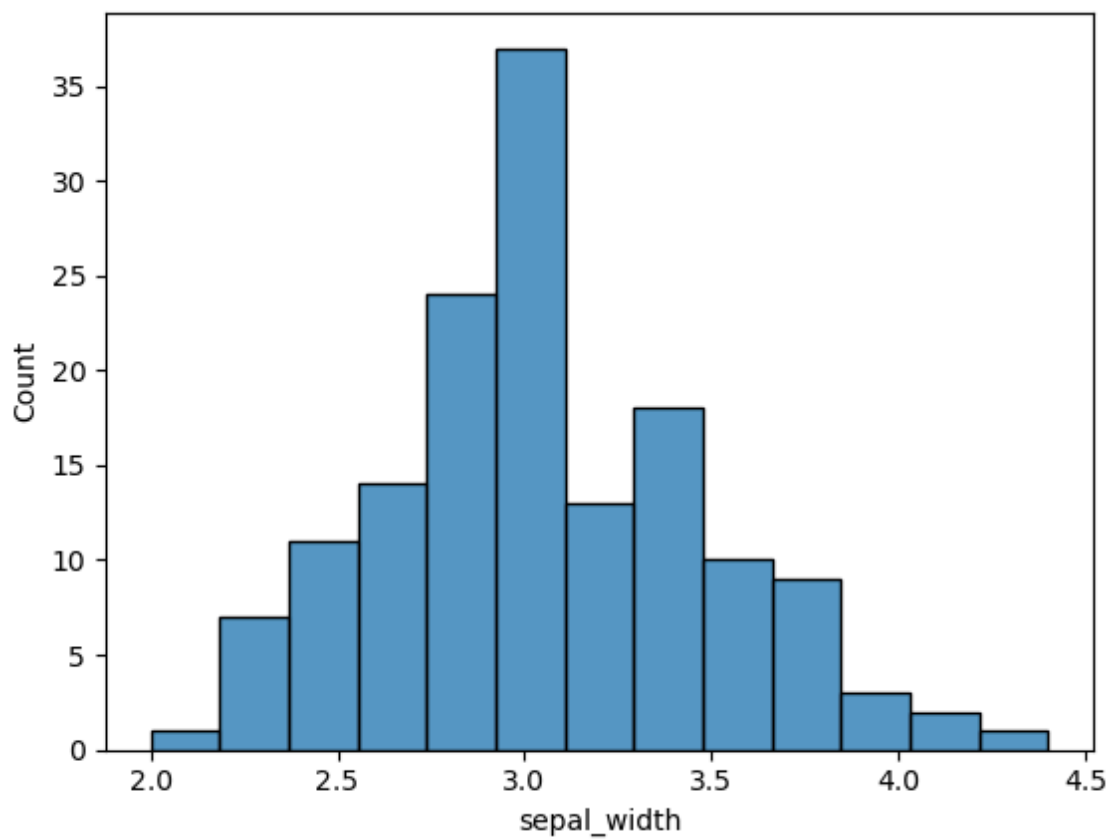
Out[12]: `<Axes: xlabel='petal_width', ylabel='Count'>`

In [13]: `sns.histplot(iris.petal_width)`
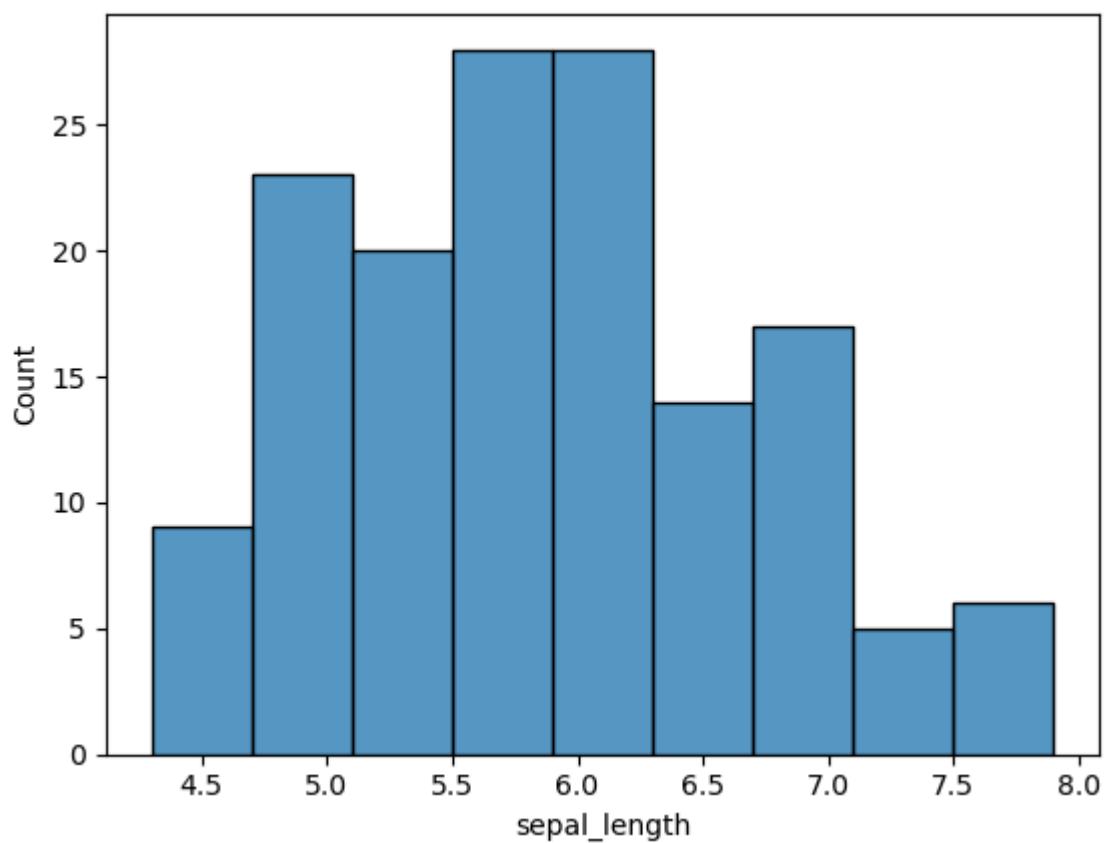
Out[13]: `<Axes: xlabel='petal_length', ylabel='Count'>`

In [14]: `sns.histplot(iris.sepal_width)`

Out[14]: `<Axes: xlabel='sepal_width', ylabel='Count'>`
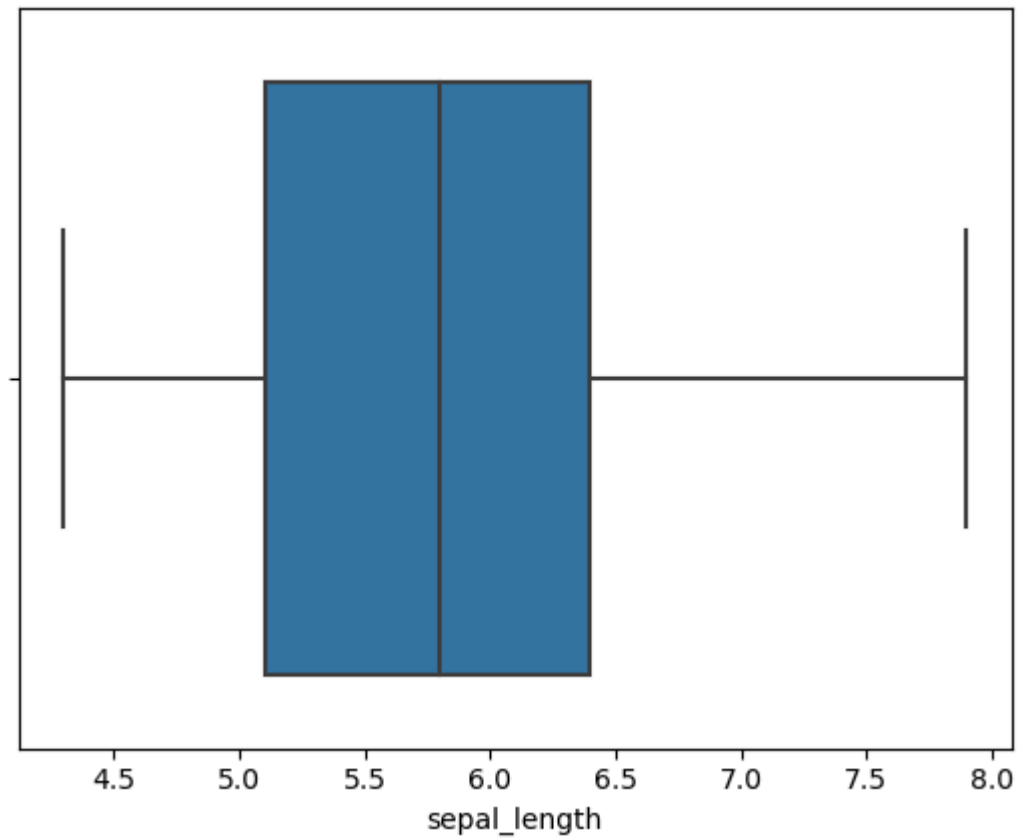

In [15]: `sns.histplot(iris.sepal_length)`

Out[15]: `<Axes: xlabel='sepal_length', ylabel='Count'>`
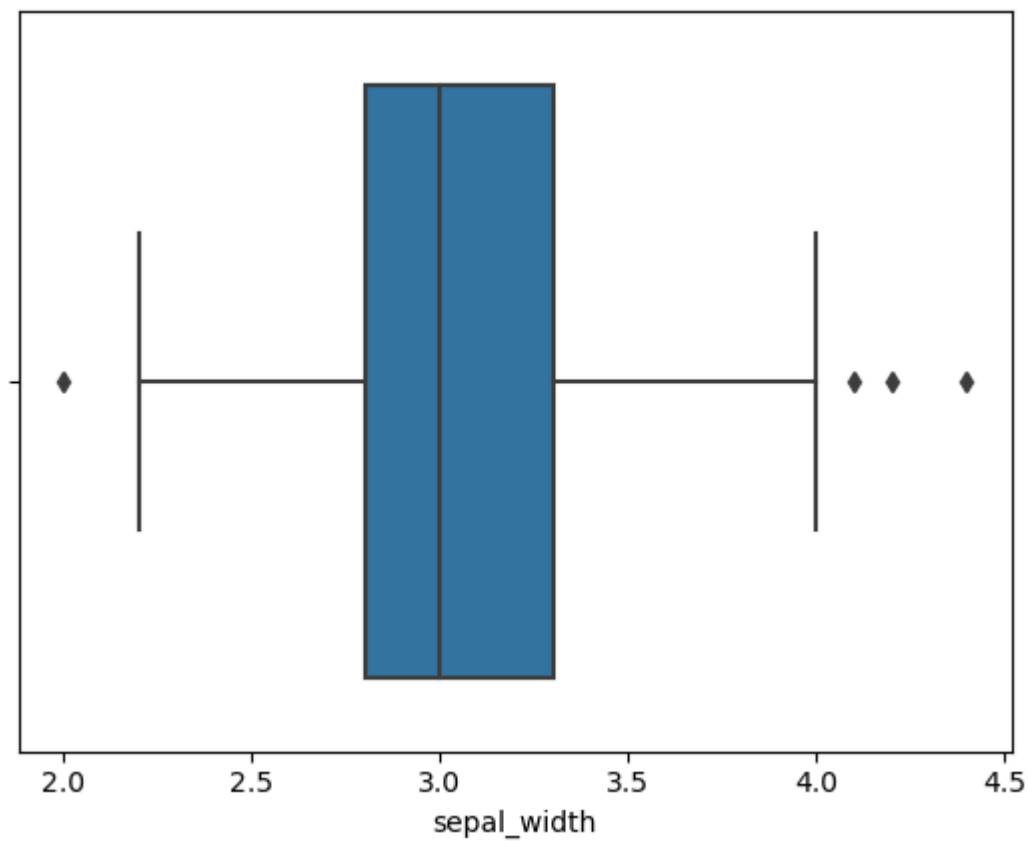
3.Create a boxplot for each feature in the dataset

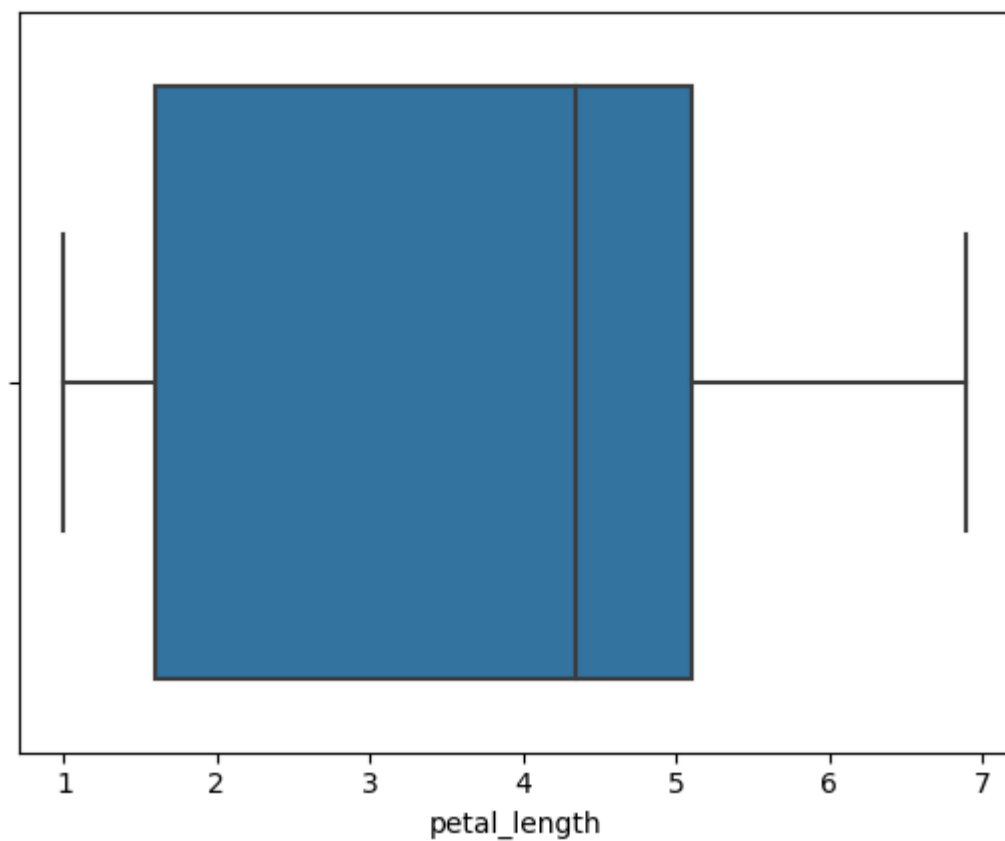In [19]: `sns.boxplot(data = iris , x= "sepal_length" )`

Out[19]: `<Axes: xlabel='sepal_length'>`

In [20]: `sns.boxplot(data = iris , x= "sepal_width" )`
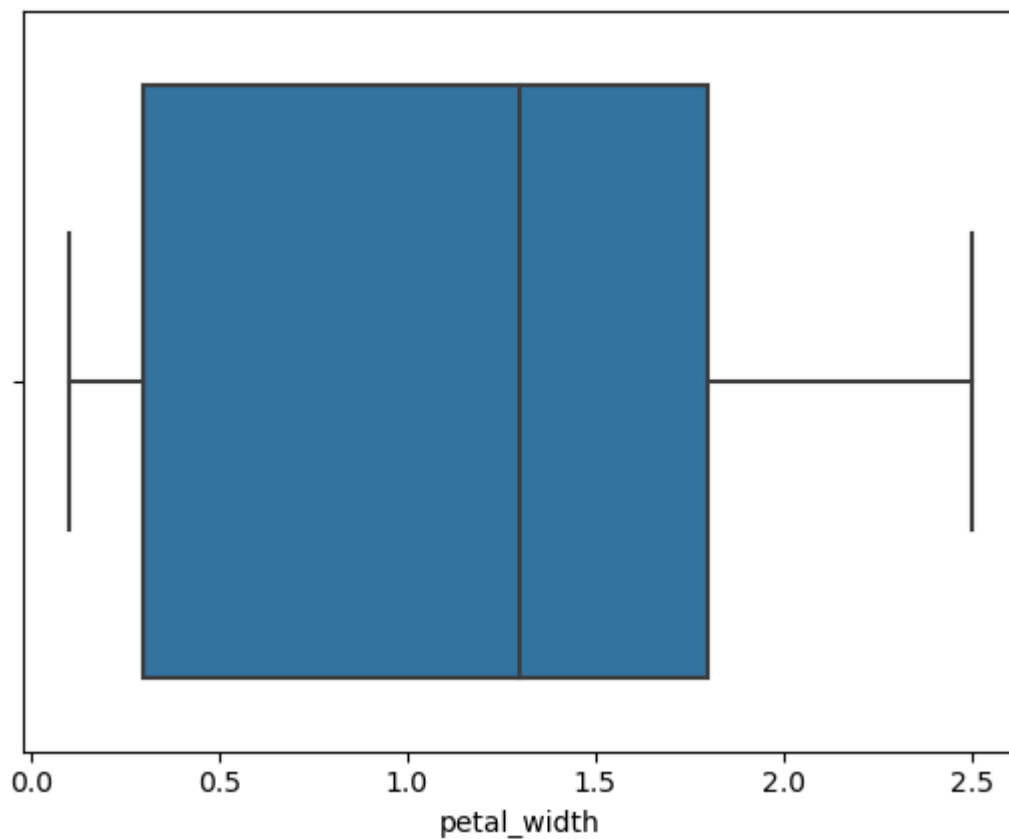
Out[20]: `<Axes: xlabel='sepal_width'>`

In [21]: `sns.boxplot(data = iris , x= "petal_length" )`

Out[21]: `<Axes: xlabel='petal_length'>`
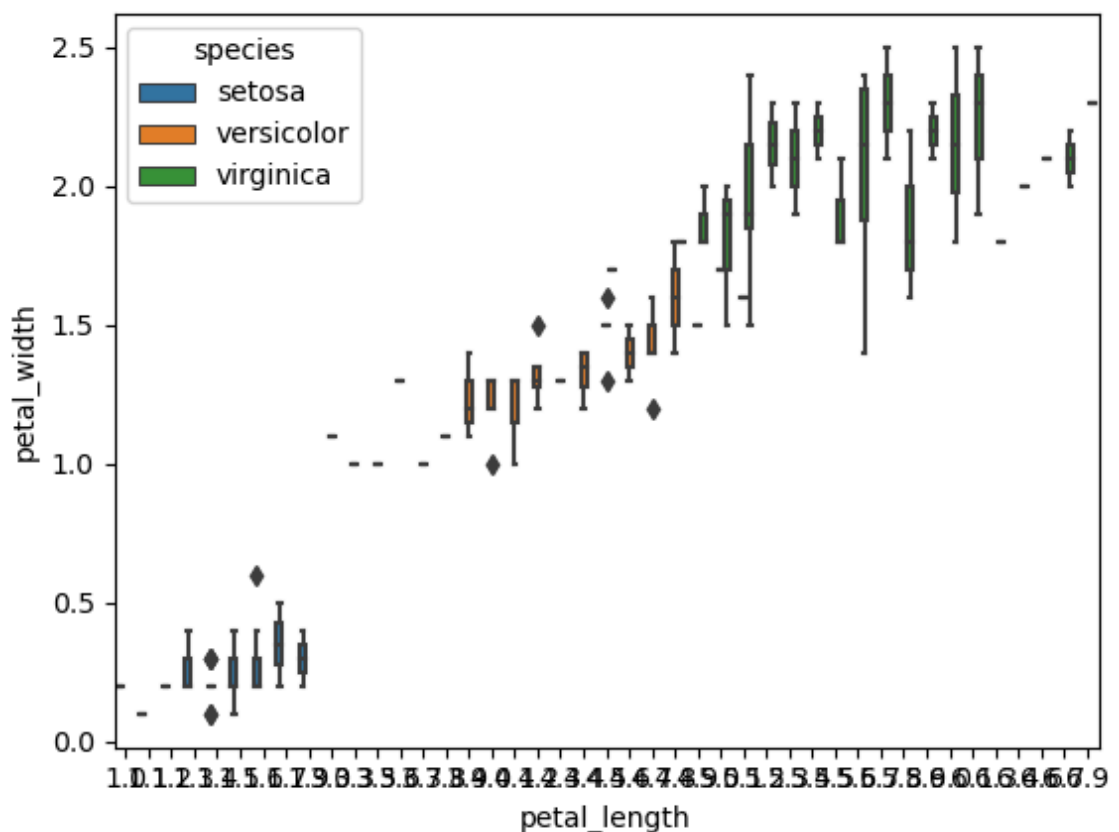
In [22]: `sns.boxplot(data = iris , x= "petal_width" )`

Out[22]: `<Axes: xlabel='petal_width'>`

In [27]: `sns.boxplot(data=iris ,x="petal_length", y = "petal_width" ,hue ="species")`

Out[27]: `<Axes: xlabel='petal_length', ylabel='petal_width'>`

4.Compare distributions and identify outliers.

In [ ]: