```
import numpy as np
import pandas as pd
import seaborn as sns
import matplotlib.pyplot as plt
```

```
from google.colab import files
uploaded = files.upload()
```

Choose Files IRIS.csv

- **IRIS.csv**(text/csv) - 4617 bytes, last modified: 3/16/2023 - 100% done

```
Saving IRIS.csv to IRIS.csv
```

```
df = pd.read_csv('IRIS.csv')
```

```
df
```

	sepal_length	sepal_width	petal_length	petal_width	species
0	5.1	3.5	1.4	0.2	Iris-setosa
1	4.9	3.0	1.4	0.2	Iris-setosa
2	4.7	3.2	1.3	0.2	Iris-setosa
3	4.6	3.1	1.5	0.2	Iris-setosa
4	5.0	3.6	1.4	0.2	Iris-setosa
...	...	...	...	...	...
145	6.7	3.0	5.2	2.3	Iris-virginica
146	6.3	2.5	5.0	1.9	Iris-virginica
147	6.5	3.0	5.2	2.0	Iris-virginica
148	6.2	3.4	5.4	2.3	Iris-virginica
149	5.9	3.0	5.1	1.8	Iris-virginica

150 rows × 5 columns

```
df.describe()
```

	sepal_length	sepal_width	petal_length	petal_width
count	150.000000	150.000000	150.000000	150.000000
mean	5.843333	3.054000	3.758667	1.198667
std	0.828066	0.433594	1.764420	0.763161
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

```
df.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
```

```
df.dtypes
```

```
sepal_length    float64
sepal_width     float64
petal_length    float64
petal_width     float64
```

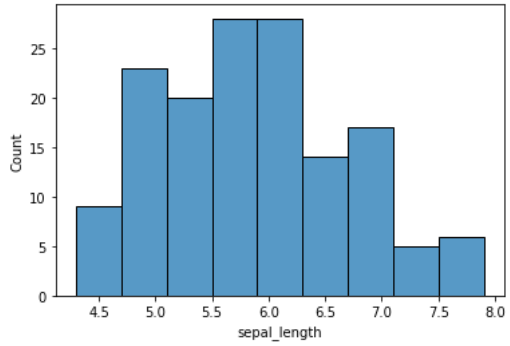
```
species          object
dtype: object
```

```
np.unique(df['species'])
```

```
array(['Iris-setosa', 'Iris-versicolor', 'Iris-virginica'], dtype=object)
```

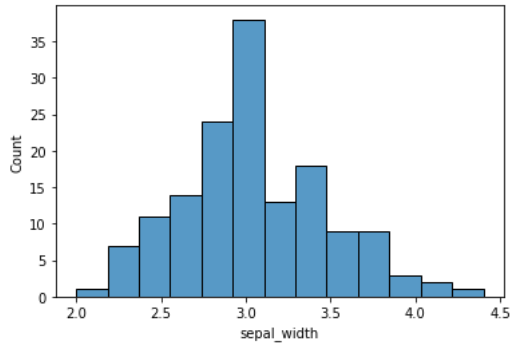
```
sns.histplot(df['sepal_length'])
```

```
<AxesSubplot:xlabel='sepal_length', ylabel='Count'>
```



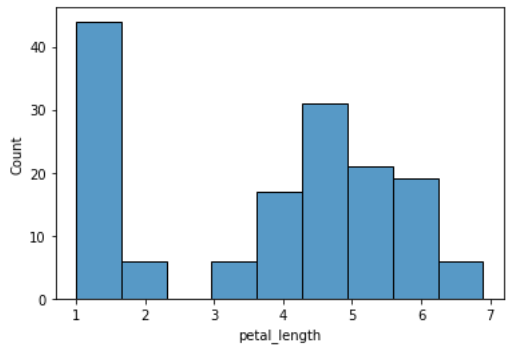

```
sns.histplot(df['sepal_width'])
```

```
<AxesSubplot:xlabel='sepal_width', ylabel='Count'>
```



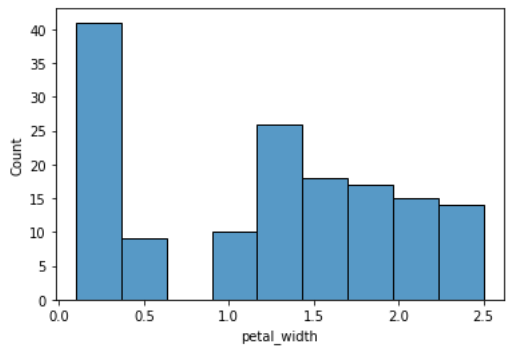

```
sns.histplot(df['petal_length'])
```

```
<AxesSubplot:xlabel='petal_length', ylabel='Count'>
```



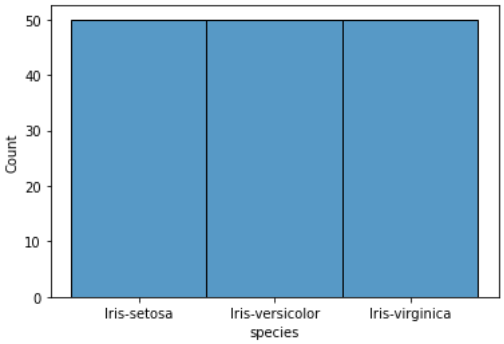

```
sns.histplot(df['petal_width'])
```

```
<AxesSubplot:xlabel='petal_width', ylabel='Count'>
```



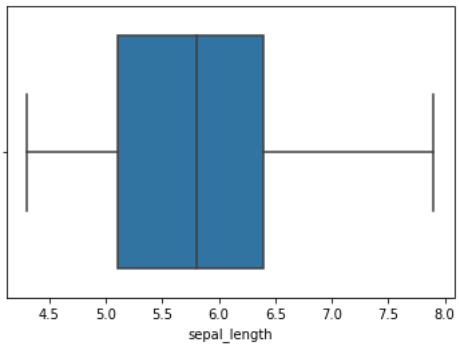
```
sns.histplot(df['species'])
```

```
<AxesSubplot:xlabel='species', ylabel='Count'>
```



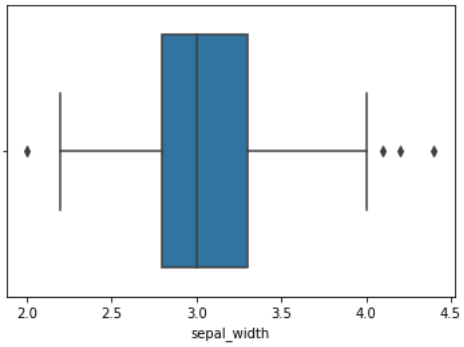

```
sns.boxplot(df['sepal_length'])
```

```
/usr/local/lib/python3.9/dist-packages/seaborn/_decorators.py:36: FutureWarning: Pass the following variable as a keyword a
  warnings.warn(
<AxesSubplot:xlabel='sepal_length'>
```



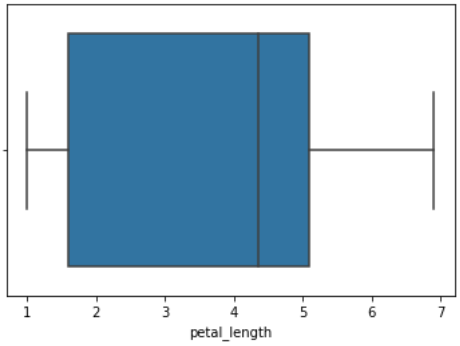

```
sns.boxplot(df['sepal_width'])
```

```
/usr/local/lib/python3.9/dist-packages/seaborn/_decorators.py:36: FutureWarning: Pass the following variable as a keyword a
  warnings.warn(
<AxesSubplot:xlabel='sepal_width'>
```



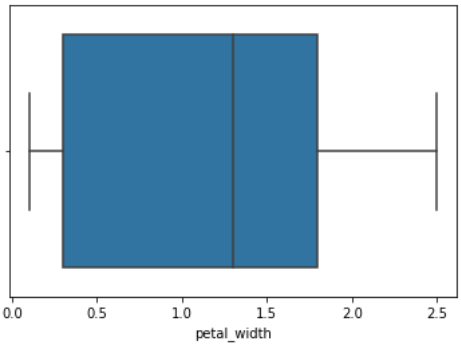

```
sns.boxplot(df['petal_length'])
```

```
/usr/local/lib/python3.9/dist-packages/seaborn/_decorators.py:36: FutureWarning: Pass the following variable as a keyword a
  warnings.warn(
<AxesSubplot:xlabel='petal_length'>
```



```
sns.boxplot(df['petal_width'])
```

```
/usr/local/lib/python3.9/dist-packages/seaborn/_decorators.py:36: FutureWarning: Pass
  warnings.warn(
<AxesSubplot:xlabel='petal_width'>
```




```
data_to_plot = [df['sepal_length'],df['sepal_width'],df['petal_length'],df['petal_width']]
# Comparing distributions and identifying outliers
```

```
fig = plt.figure(1, figsize=(12,8))
 # 1 is unique identifier
 # 12 and 8 is width and height
```

```
<Figure size 864x576 with 0 Axes>
```

```
ax = fig.add_subplot(111)
# Creating an axes instance
```

```
bp = ax.boxplot(data_to_plot)
# Creating the boxplot
```

```
bp
```

```
{'whiskers': [<matplotlib.lines.Line2D at 0x7f254b3b9f70>,
  <matplotlib.lines.Line2D at 0x7f254b6c1490>,
  <matplotlib.lines.Line2D at 0x7f254b39dee0>,
  <matplotlib.lines.Line2D at 0x7f254b22b1f0>,
  <matplotlib.lines.Line2D at 0x7f254b237310>,
  <matplotlib.lines.Line2D at 0x7f254b2375e0>,
  <matplotlib.lines.Line2D at 0x7f254b235700>,
  <matplotlib.lines.Line2D at 0x7f254b2359d0>],
 'caps': [<matplotlib.lines.Line2D at 0x7f254b39d0d0>,
  <matplotlib.lines.Line2D at 0x7f254b39d3a0>,
  <matplotlib.lines.Line2D at 0x7f254b22b4c0>,
  <matplotlib.lines.Line2D at 0x7f254b22b790>,
  <matplotlib.lines.Line2D at 0x7f254b2378b0>,
  <matplotlib.lines.Line2D at 0x7f254b237b80>,
  <matplotlib.lines.Line2D at 0x7f254b235ca0>,
  <matplotlib.lines.Line2D at 0x7f254b235f70>],
 'boxes': [<matplotlib.lines.Line2D at 0x7f254b3b98e0>,
  <matplotlib.lines.Line2D at 0x7f254b39dc10>,
  <matplotlib.lines.Line2D at 0x7f254b237040>,
  <matplotlib.lines.Line2D at 0x7f254b235430>],
 'medians': [<matplotlib.lines.Line2D at 0x7f254b39d670>,
  <matplotlib.lines.Line2D at 0x7f254b22ba60>,
  <matplotlib.lines.Line2D at 0x7f254b237e50>,
  <matplotlib.lines.Line2D at 0x7f254ac08280>],
 'fliers': [<matplotlib.lines.Line2D at 0x7f254b39d940>,
  <matplotlib.lines.Line2D at 0x7f254b22bd30>,
  <matplotlib.lines.Line2D at 0x7f254b235160>,
  <matplotlib.lines.Line2D at 0x7f254ac08550>],
 'means': []}
```

✓ 0s completed at 1:58 AM