Practical No. 10

by Kartik Deshpande

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). Scan the dataset and give the inference as:

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

In [6]:
```python
import seaborn as sb
ds = sb.load_dataset('iris')
ds.head()
```

Out[6]:

	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

In [13]:
```python
import matplotlib.pyplot as plt
fig, axes = plt.subplots(2,2,figsize = (16,9))
sb.histplot(ds['sepal_length'], ax = axes[0,0])
sb.histplot(ds['sepal_width'], ax = axes[0,1])
sb.histplot(ds['petal_length'], ax = axes[1,0])
sb.histplot(ds['petal_width'], ax = axes[1,1])
```

/home/kartik/anaconda3/lib/python3.11/site-packages/seaborn/_oldcore.py:1119: FutureWarning: use_inf_as_na option is deprecated and will be removed in a future version. Convert inf values to NaN before operating instead.
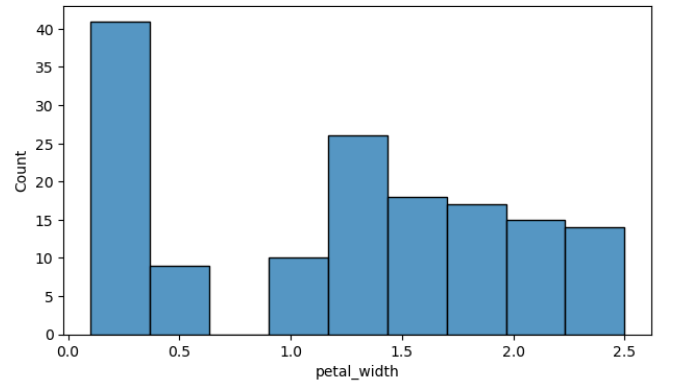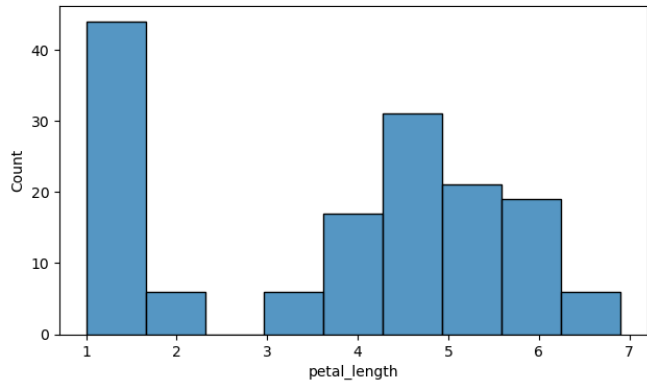 with pd.option_context('mode.use_inf_as_na', True):
/home/kartik/anaconda3/lib/python3.11/site-packages/seaborn/_oldcore.py:1119: FutureWarning: use_inf_as_na option is deprecated and will be removed in a future version. Convert inf values to NaN before operating instead.
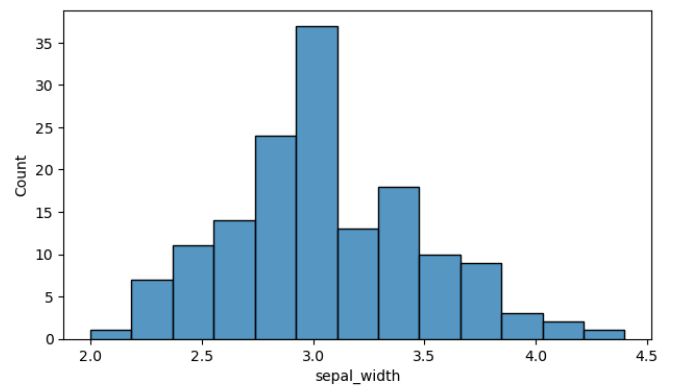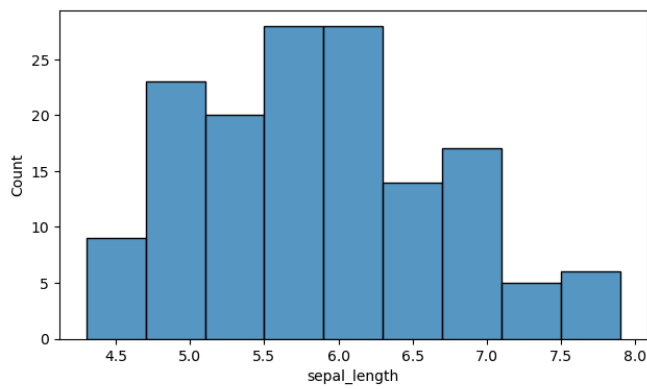 with pd.option_context('mode.use_inf_as_na', True):
/home/kartik/anaconda3/lib/python3.11/site-packages/seaborn/_oldcore.py:1119: FutureWarning: use_inf_as_na option is deprecated and will be removed in a future version. Convert inf values to NaN before operating instead.
  with pd.option_context('mode.use_inf_as_na', True):
/home/kartik/anaconda3/lib/python3.11/site-packages/seaborn/_oldcore.py:1119: FutureWarning: use_inf_as_na option is deprecated and will be removed in a future version. Convert inf values to NaN before operating instead.
  with pd.option_context('mode.use_inf_as_na', True):

Out[13]: <Axes: xlabel='petal_width', ylabel='Count'>

```
In [17]: import matplotlib.pyplot as plt
         fig, axes = plt.subplots(2,2,figsize=(16,9))
         sb.boxplot(x='species',y='petal_length',data = ds, ax=axes[0,0])
         sb.boxplot(x='species',y='petal_width',data = ds, ax=axes[0,1])
         sb.boxplot(x='species',y='sepal_length',data = ds, ax=axes[1,0])
         sb.boxplot(x='species',y='sepal_width',data = ds, ax=axes[1,1])
```

Out[17]: <Axes: xlabel='species', ylabel='sepal_width'>