pratical-10

April 17, 2024

---

1 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 boxplot for each feature in the dataset. 4. Compare distributions and identify outliers.

[ ]:

```
# import seaborn as sns
# import matplotlib.pyplot as plt

# iris = sns.load_dataset('iris')
# iris

# iris.info

# iris.describe()

# type(iris.sepal_length)

# iris.sepal_length.dtype

# iris.sepal_width.dtype

# iris.petal_length.dtype

# iris.petal_width.dtype

# iris.species.dtype

# fig, axes = plt.subplots(2, 2, figsize = (10, 6))
# sns.histplot(iris['sepal_length'],ax = axes[0, 0])
# sns.histplot(iris['sepal_width'],ax = axes[0, 1])
# sns.histplot(iris['petal_length'],ax = axes[1, 0])
# sns.histplot(iris['petal_width'],ax = axes[1, 1])
# # For boxplot
# fig, axes = plt.subplots(2, 2, figsize = (16, 10))
# sns.boxplot(x= 'species', y= 'sepal_length', data= iris, ax= axes[0, 0])
# sns.boxplot(x= 'species', y= 'sepal_width', data= iris, ax= axes[0, 1])
# sns.boxplot(x= 'species', y= 'petal_length', data= iris, ax= axes[1, 0])
# sns.boxplot(x= 'species', y= 'petal_width', data= iris, ax= axes[1, 1])
```

[1]:

```
import seaborn as sns
import matplotlib.pyplot as plt
```

[2]:

```
iris = sns.load_dataset('iris')
iris
```

[2]:

```
     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 x 5 columns]
```

[3]:

```
iris.info
```

[3]:

```
<bound method DataFrame.info of      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 x 5 columns]>
```

[4]:

```
iris.describe()
```

[4]:

```
       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
```

[5]:

```
type(iris.sepal_length)
```

[5]:

```
pandas.core.series.Series
```

[6]:

```
iris.sepal_length.dtype
```

[6]:

```
dtype('float64')
```

[7]:

```
iris.sepal_width.dtype
```

[7]:

```
dtype('float64')
```

[8]:

```
iris.petal_length.dtype
```

[8]:

```
dtype('float64')
```

[9]:

```
iris.petal_width.dtype
```

[9]:

```
dtype('float64')
```

[10]:

```
iris.species.dtype
```

[10]:

```
dtype('O')
```

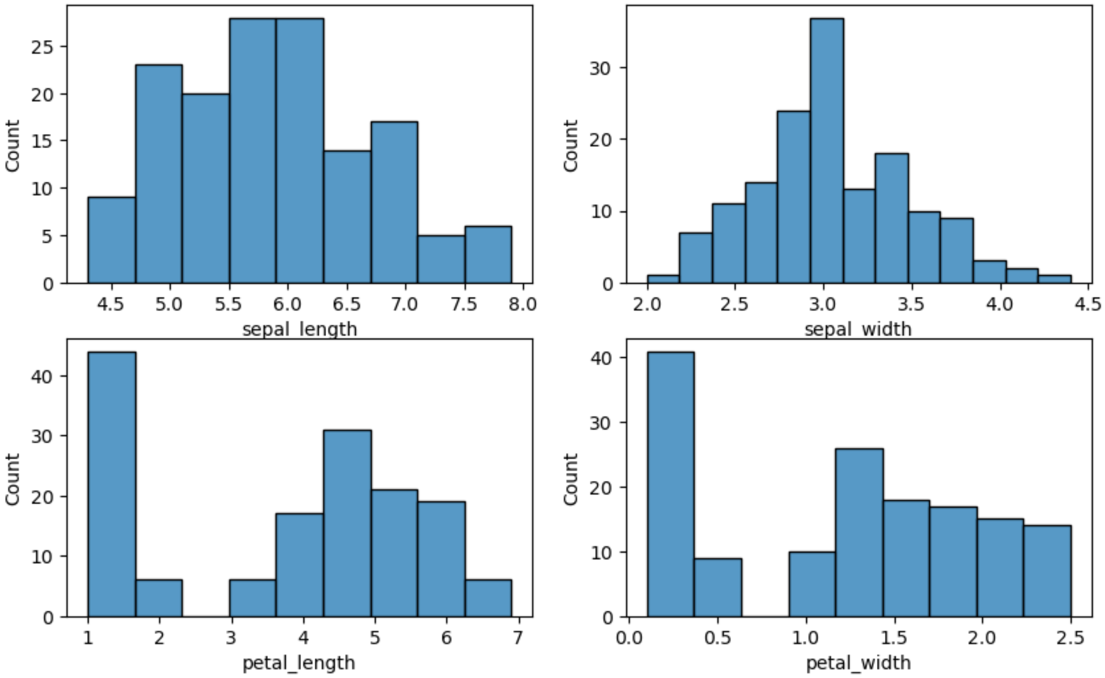
[11]:

```
fig, axes = plt.subplots(2, 2, figsize = (10, 6))
sns.histplot(iris['sepal_length'],ax = axes[0, 0])
sns.histplot(iris['sepal_width'],ax = axes[0, 1])
sns.histplot(iris['petal_length'],ax = axes[1, 0])
sns.histplot(iris['petal_width'],ax = axes[1, 1])
```

```
F:\Anaconda3\Lib\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):
F:\Anaconda3\Lib\site-packages\seaborn\_oldcore.py:1119: FutureWarning:
use_inf_as_na option is deprecated and will be removed in a future version.
```

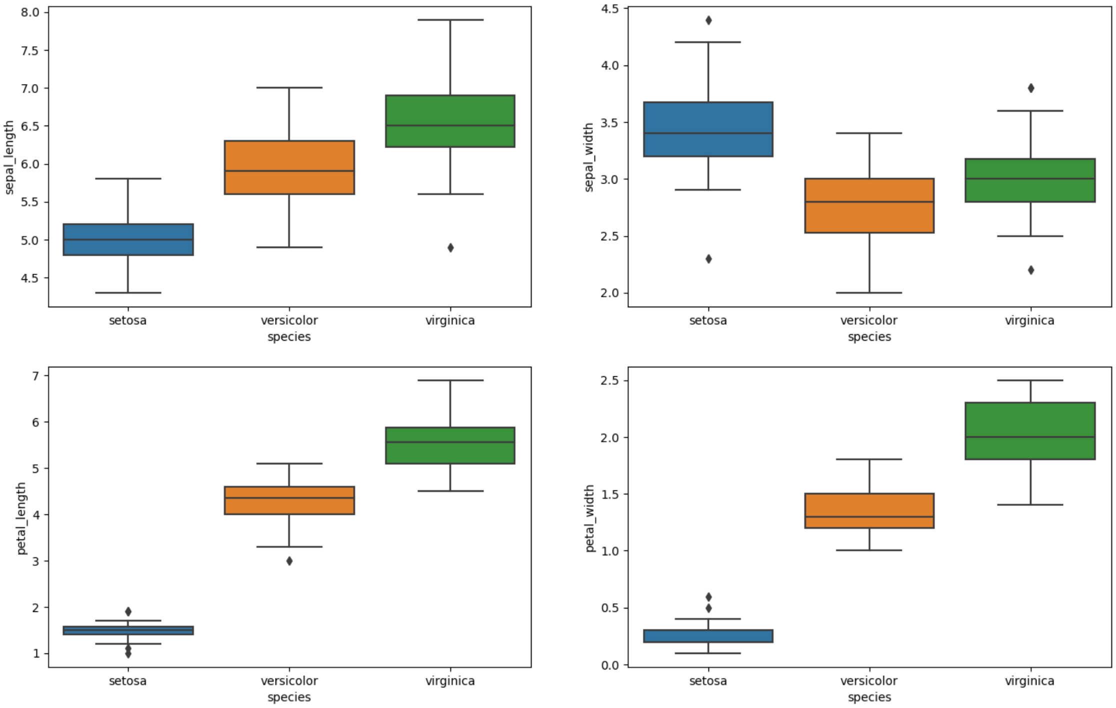
```
Convert inf values to NaN before operating instead.
  with pd.option_context('mode.use_inf_as_na', True):
F:\Anaconda3\Lib\site-packages\seaborn\_oldcore.py:1119: FutureWarning:
use_inf_as_na option is deprecated and will be removed in a future version.
Convert inf values to NaN before operating instead.
  with pd.option_context('mode.use_inf_as_na', True):
F:\Anaconda3\Lib\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):
```

```
[11]: <Axes: xlabel='petal_width', ylabel='Count'>
```



```
[12]: # For boxplot
fig, axes = plt.subplots(2, 2, figsize = (16, 10))
sns.boxplot(x= 'species', y= 'sepal_length', data= iris, ax= axes[0, 0])
sns.boxplot(x= 'species', y= 'sepal_width', data= iris, ax= axes[0, 1])
sns.boxplot(x= 'species', y= 'petal_length', data= iris, ax= axes[1, 0])
sns.boxplot(x= 'species', y= 'petal_width', data= iris, ax= axes[1, 1])
```

```
[12]: <Axes: xlabel='species', ylabel='petal_width'>
```


sepal_length
8.0
7.5
7.0
6.5
6.0
5.5
5.0
4.5
setosa
versicolor
virginica
species
sepal_width
4.5
4.0
3.5
3.0
2.5
2.0
setosa
versicolor
virginica
species
petal_length
7
6
5
4
3
2
1
setosa
versicolor
virginica
species
petal_width
2.5
2.0
1.5
1.0
0.5
0.0
setosa
versicolor
virginica
species

[ ]: