

day3of1000

February 16, 2026

```
[3]: import pandas as pd
import seaborn as sns
import matplotlib.pyplot as plt
from scipy import stats

# Load Iris dataset
iris = sns.load_dataset("iris")

# 1. Descriptive Statistics

print("Mean, Median, and Std Deviation by Feature:\n")
for col in iris.columns[:-1]:
    mean = iris[col].mean()
    median = iris[col].median()
    std = iris[col].std()
    print(f"{col}: mean={mean:.2f}, median={median:.2f}, std={std:.2f}")

# 2. Grouped Statistics by Species

print("\nMean and Std by Species:\n")
grouped = iris.groupby('species').agg(['mean', 'std'])
print(grouped)

# 3. ANOVA Test (Are species significantly different?)

anova_result = stats.f_oneway(
    iris[iris['species']=='setosa']['petal_length'],
    iris[iris['species']=='versicolor']['petal_length'],
    iris[iris['species']=='virginica']['petal_length']
)
print("\nANOVA result for petal_length across species:", anova_result)

# 4. Visualizations
```

```

# Boxplots by species
for col in iris.columns[:-1]:
    plt.figure(figsize=(6,4))
    sns.boxplot(data=iris, x='species', y=col)
    plt.title(f"{col} by Species")
    plt.show()

# Histogram & KDE
for col in iris.columns[:-1]:
    plt.figure(figsize=(6,4))
    sns.histplot(data=iris, x=col, hue='species', kde=True, element='step')
    plt.title(f"Distribution of {col} by Species")
    plt.show()

```

Mean, Median, and Std Deviation by Feature:

```

sepal_length: mean=5.84, median=5.80, std=0.83
sepal_width: mean=3.06, median=3.00, std=0.44
petal_length: mean=3.76, median=4.35, std=1.77
petal_width: mean=1.20, median=1.30, std=0.76

```

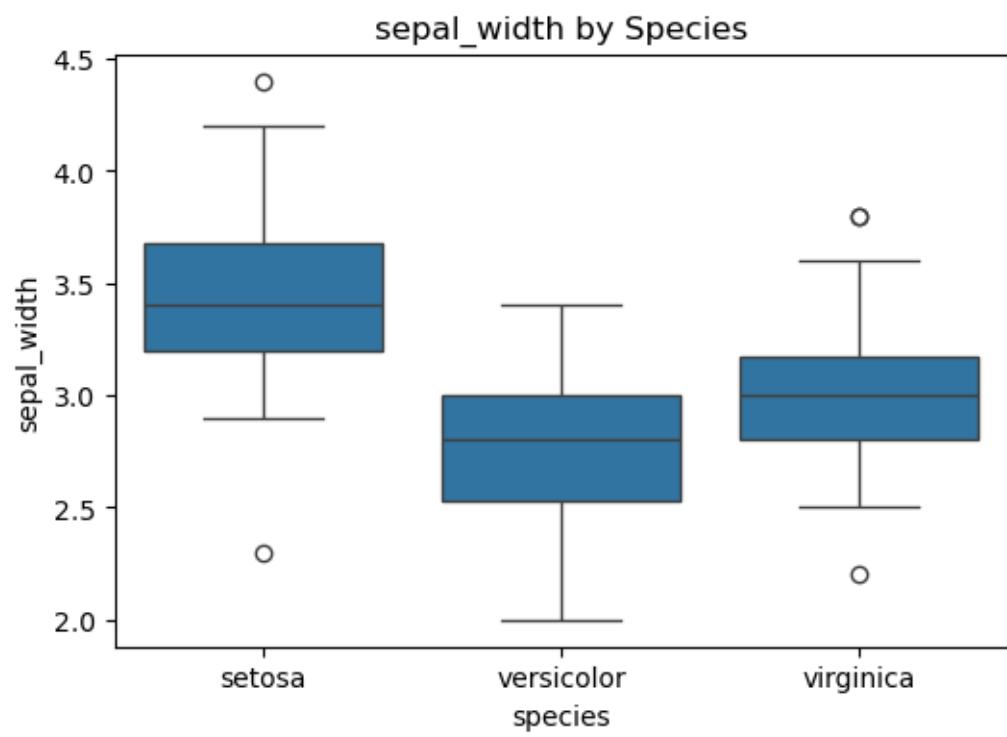
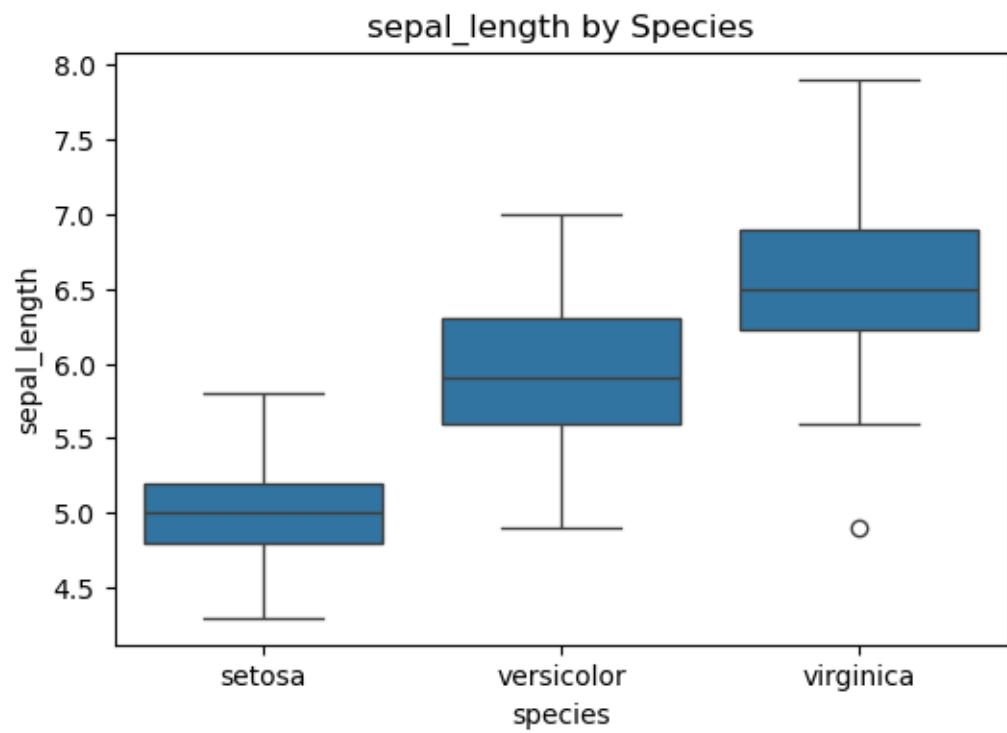
Mean and Std by Species:

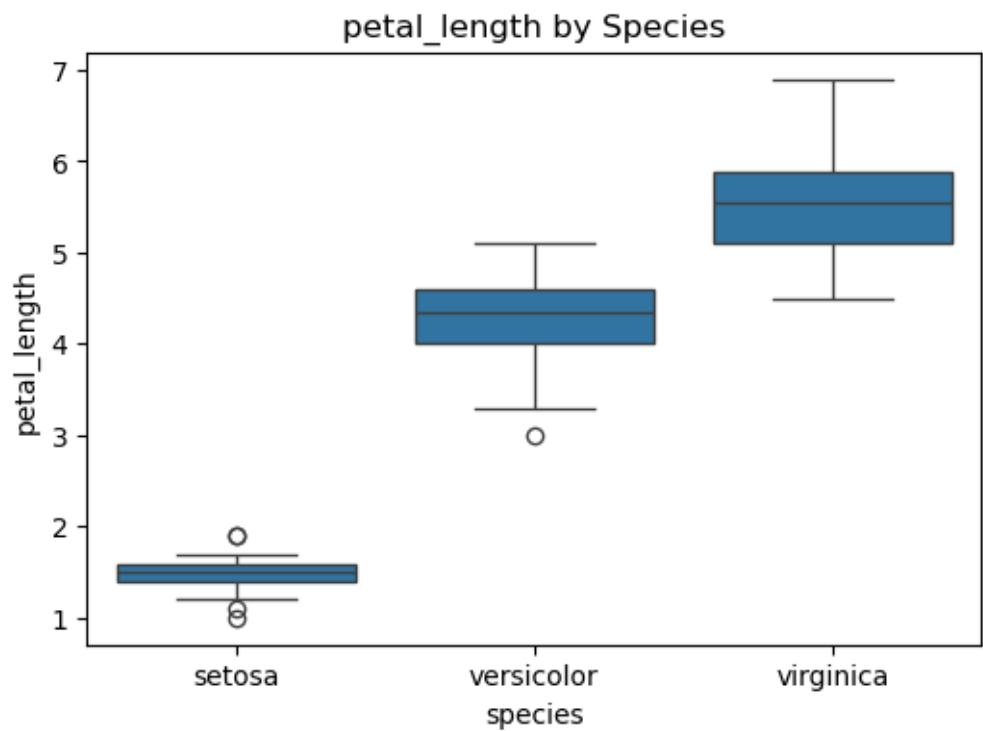
species	sepal_length		sepal_width		petal_length \	
	mean	std	mean	std	mean	std
setosa	5.006	0.352490	3.428	0.379064	1.462	
versicolor	5.936	0.516171	2.770	0.313798	4.260	
virginica	6.588	0.635880	2.974	0.322497	5.552	
	petal_width					
	std	mean	std			

species	petal_width		
	std	mean	std
setosa	0.173664	0.246	0.105386
versicolor	0.469911	1.326	0.197753
virginica	0.551895	2.026	0.274650

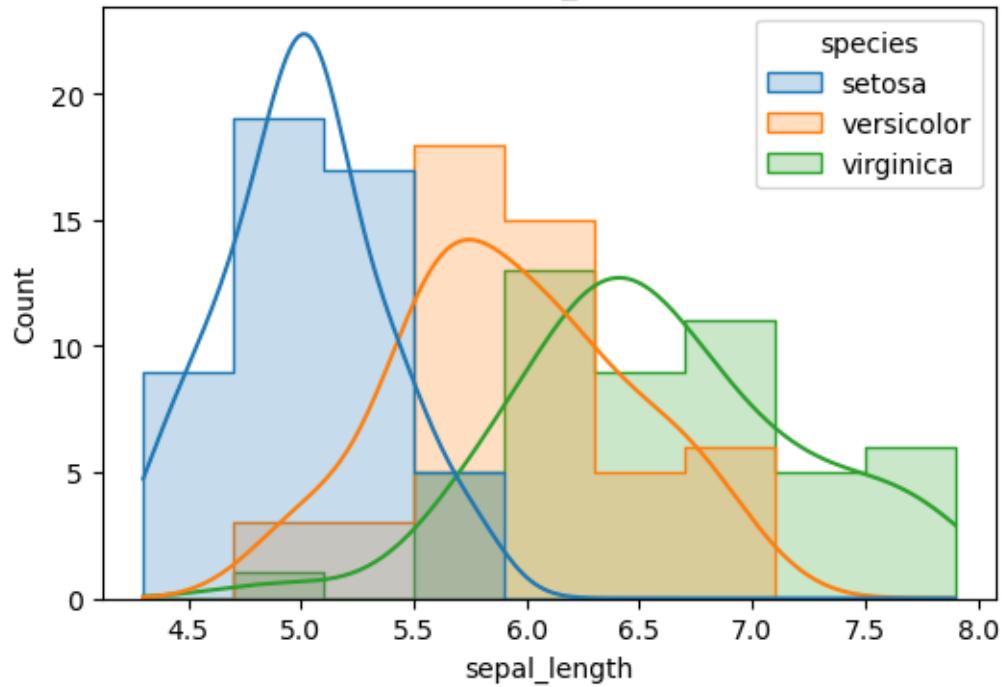
ANOVA result for petal_length across species:

```
F_onewayResult(statistic=1180.161182252981, pvalue=2.8567766109615584e-91)
```

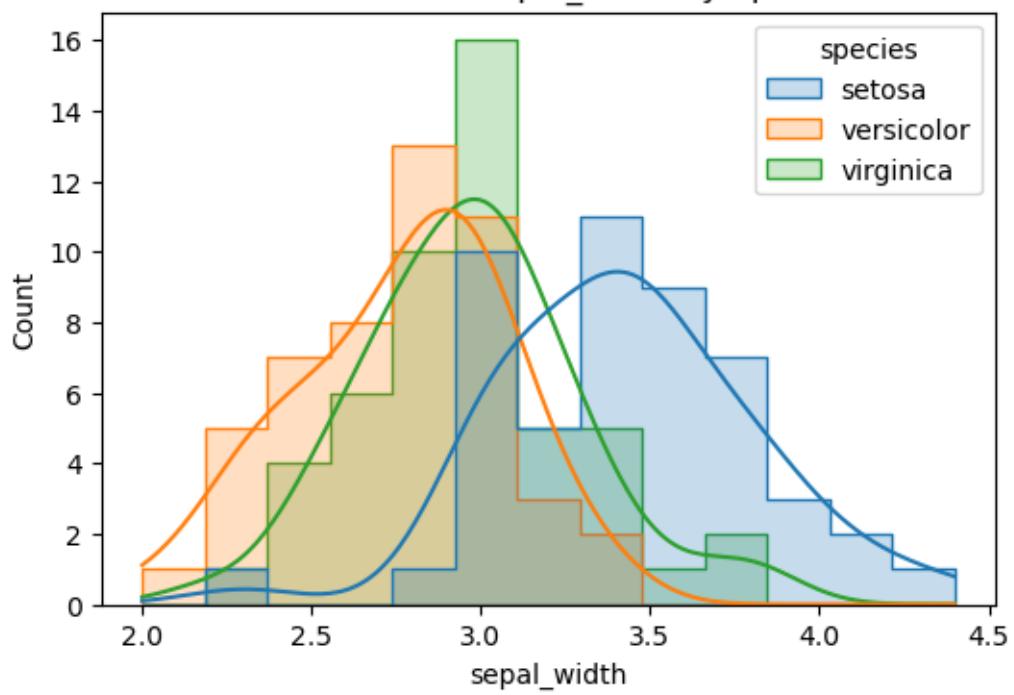




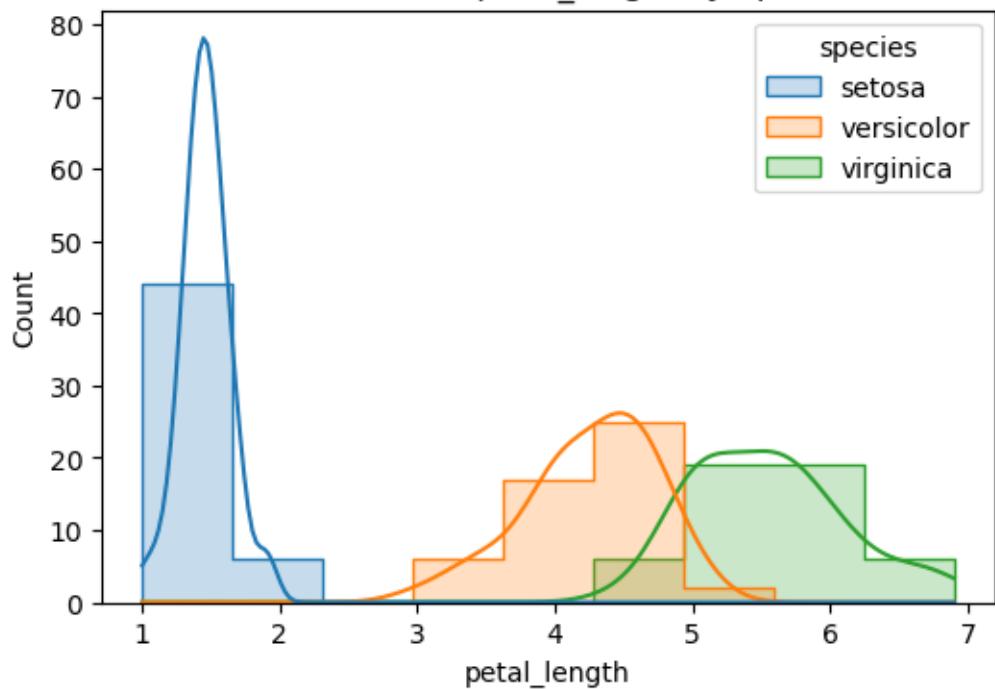
Distribution of sepal_length by Species



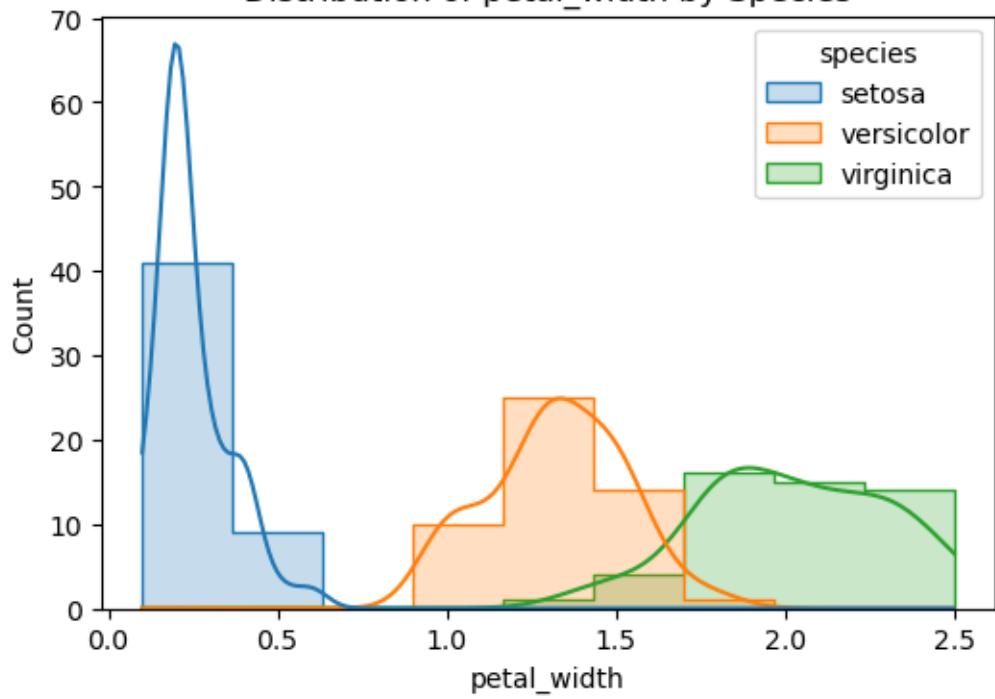
Distribution of sepal_width by Species



Distribution of petal_length by Species



Distribution of petal_width by Species



[]: