

## Covariance Matrix

```
In [17]: import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns
import warnings
warnings.filterwarnings('ignore')
```

```
In [18]: iris = sns.load_dataset('iris')
iris
```

```
Out[18]:
```

	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

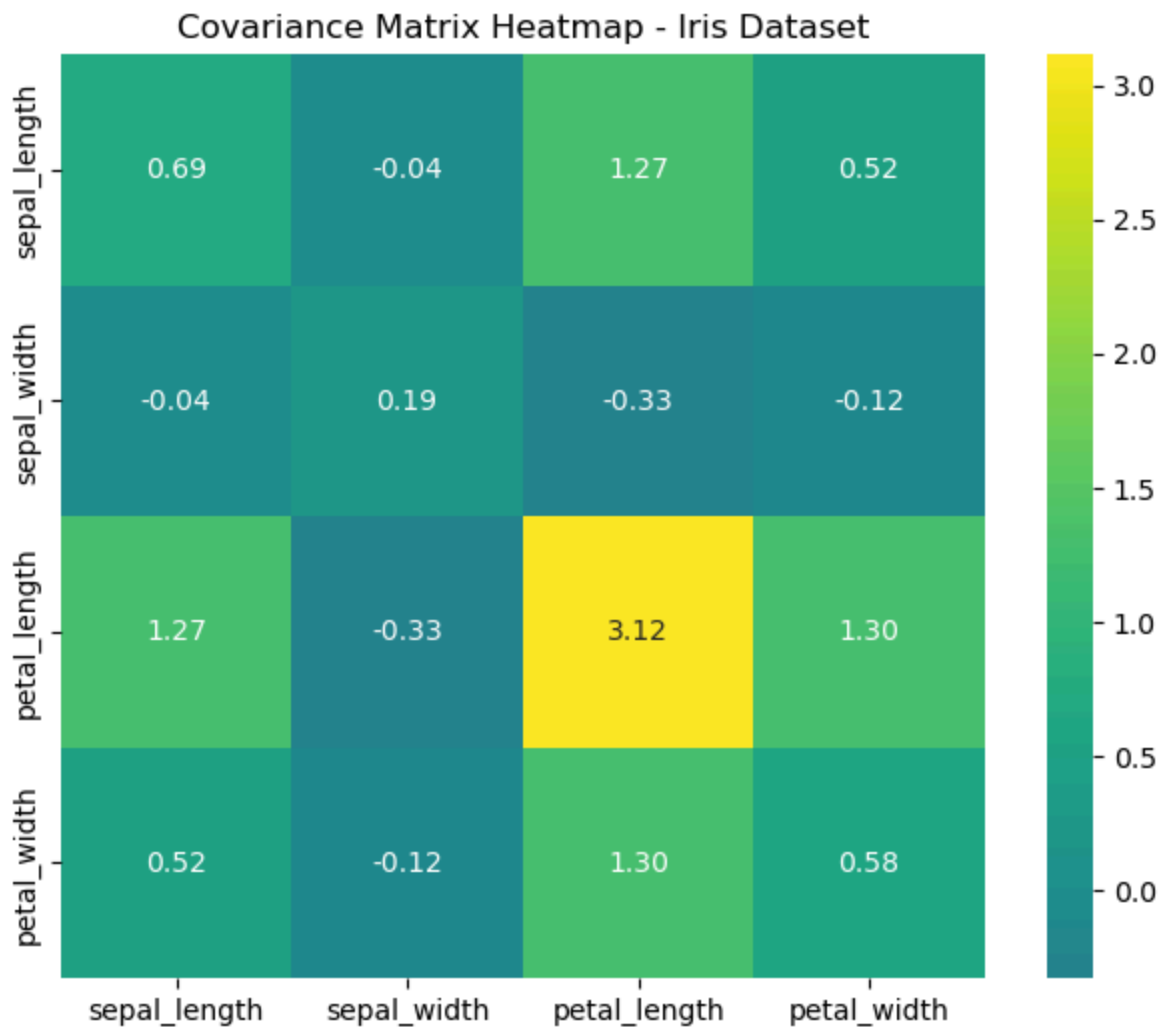
150 rows × 5 columns

```
In [20]: # Select only numeric columns for covariance matrix
iris_numeric = iris.select_dtypes(include='number')
# Compute covariance matrix
cov_matrix = iris_numeric.cov()
cov_matrix
```

```
Out[20]:
```

	sepal_length	sepal_width	petal_length	petal_width
<b>sepal_length</b>	0.685694	-0.042434	1.274315	0.516271
<b>sepal_width</b>	-0.042434	0.189979	-0.329656	-0.121639
<b>petal_length</b>	1.274315	-0.329656	3.116278	1.295609
<b>petal_width</b>	0.516271	-0.121639	1.295609	0.581006

```
In [21]: # Plot covariance matrix heatmap
plt.figure(figsize=(8, 6))
sns.heatmap(cov_matrix, annot=True, fmt=".2f", cmap="viridis", center=0, square=True)
plt.title('Covariance Matrix Heatmap - Iris Dataset')
plt.show()
```



**Interpretation:** The heatmap shows that petal length has the largest variance (3.1163), meaning it varies most among the features. The covariance between petal length & petal width (1.2956) represents that longer petals tend to come with wider petals. On the other hand, sepal width & petal length (-0.3297) has a negative covariance indicating that, wider sepals tend to have slightly shorter petals.