

```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns
from sklearn.datasets import load_iris
```

```
iris = load_iris()
df = pd.DataFrame(data=np.c_[iris['data'], iris['target']],
                  columns=iris['feature_names'] + ['target'])
```

```
print(df.info())
print("\nFirst few rows of the dataset:")
print(df.head())
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 150 entries, 0 to 149
Data columns (total 5 columns):
 #   Column          Non-Null Count  Dtype
---  -
 0   sepal length (cm) 150 non-null    float64
 1   sepal width (cm)  150 non-null    float64
 2   petal length (cm) 150 non-null    float64
 3   petal width (cm)  150 non-null    float64
 4   target            150 non-null    float64
dtypes: float64(5)
memory usage: 6.0 KB
None
```

```
First few rows of the dataset:
   sepal length (cm)  sepal width (cm)  petal length (cm)  petal width (cm) \
0                5.1                3.5                1.4                0.2
1                4.9                3.0                1.4                0.2
2                4.7                3.2                1.3                0.2
3                4.6                3.1                1.5                0.2
4                5.0                3.6                1.4                0.2

   target
0      0.0
1      0.0
2      0.0
3      0.0
4      0.0
```

```
print("\nSummary statistics:")
print(df.describe())
```

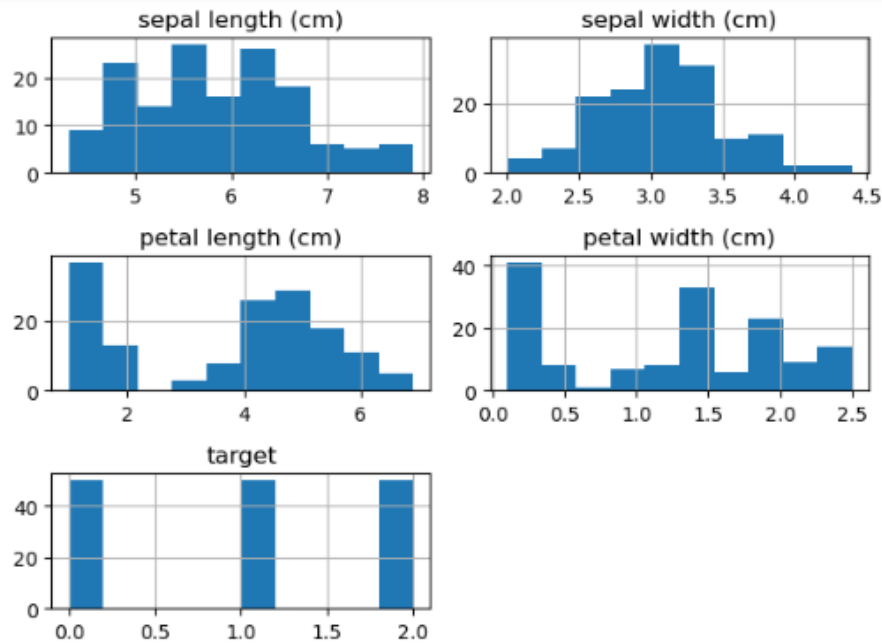
```
Summary statistics:
   sepal length (cm)  sepal width (cm)  petal length (cm) \
count      150.000000      150.000000      150.000000
mean         5.843333         3.057333         3.758000
std          0.828066         0.435866         1.765298
min          4.300000         2.000000         1.000000
25%          5.100000         2.800000         1.600000
50%          5.800000         3.000000         4.350000
75%          6.400000         3.300000         5.100000
max          7.900000         4.400000         6.900000

   petal width (cm)  target
count      150.000000  150.000000
mean         1.199333         1.000000
std          0.762238         0.819232
min          0.100000         0.000000
25%          0.300000         0.000000
50%          1.300000         1.000000
75%          1.800000         2.000000
max          2.500000         2.000000
```

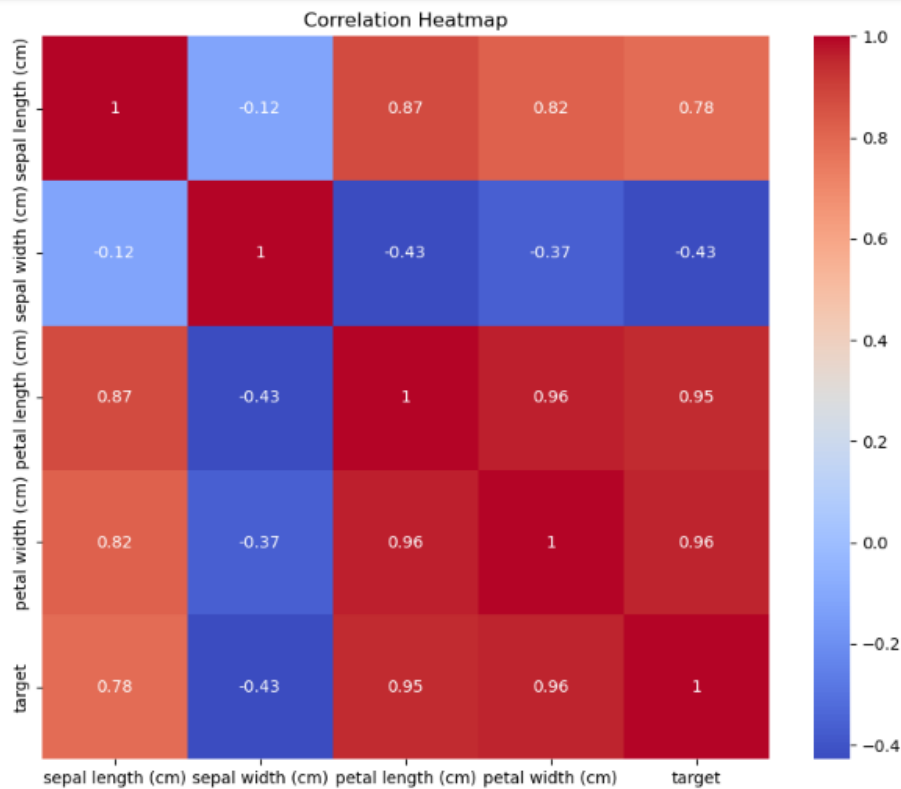
```
print("\nMissing values:")
print(df.isnull().sum())
```

```
Missing values:
sepal length (cm)  0
sepal width (cm)   0
petal length (cm)  0
petal width (cm)   0
target            0
dtype: int64
```

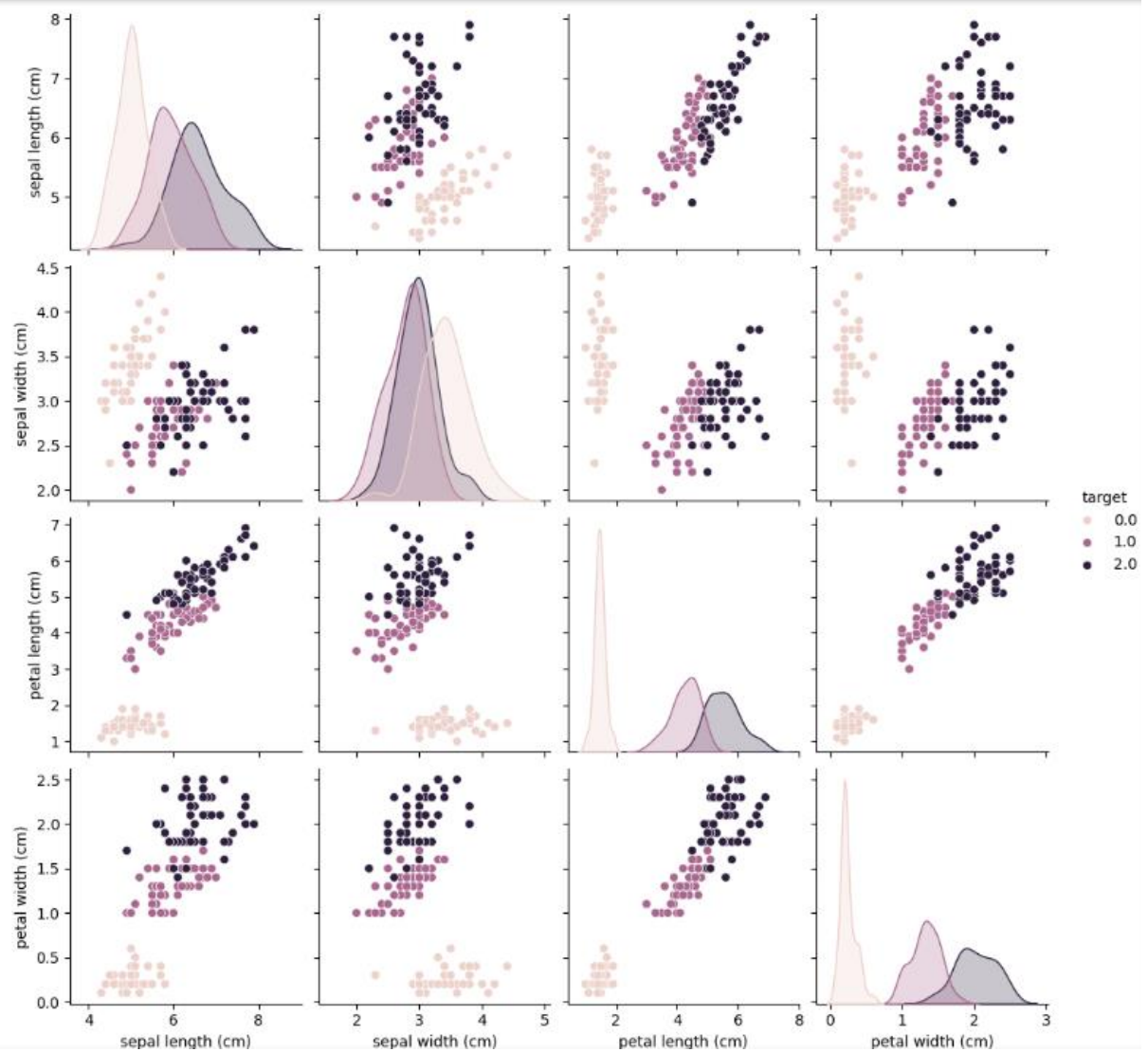
```
plt.figure(figsize=(12, 8))
df.hist()
plt.tight_layout()
plt.savefig('/tmp/histograms.png')
plt.close()
```



```
plt.figure(figsize=(10, 8))
sns.heatmap(df.corr(), annot=True, cmap='coolwarm')
plt.title('Correlation Heatmap')
plt.savefig('/tmp/correlation_heatmap.png')
plt.close()
```



```
sns.pairplot(df, hue='target')
plt.savefig('/tmp/pairplot.png')
plt.close()
```



```
plt.figure(figsize=(12, 8))
for i, feature in enumerate(iris['feature_names']):
    plt.subplot(2, 2, i+1)
    sns.boxplot(x='target', y=feature, data=df)
    plt.title(feature)
plt.tight_layout()
plt.savefig('/tmp/boxplots.png')
plt.close()
```

