

## Task 2: Exploratory Data Analysis (Iris Dataset)

```
In [1]: import seaborn as sns
import matplotlib.pyplot as plt
from sklearn.datasets import load_iris
import pandas as pd

# Load Iris
iris = load_iris(as_frame=True)
df = iris.frame

# Summary statistics
print(df.describe())
```

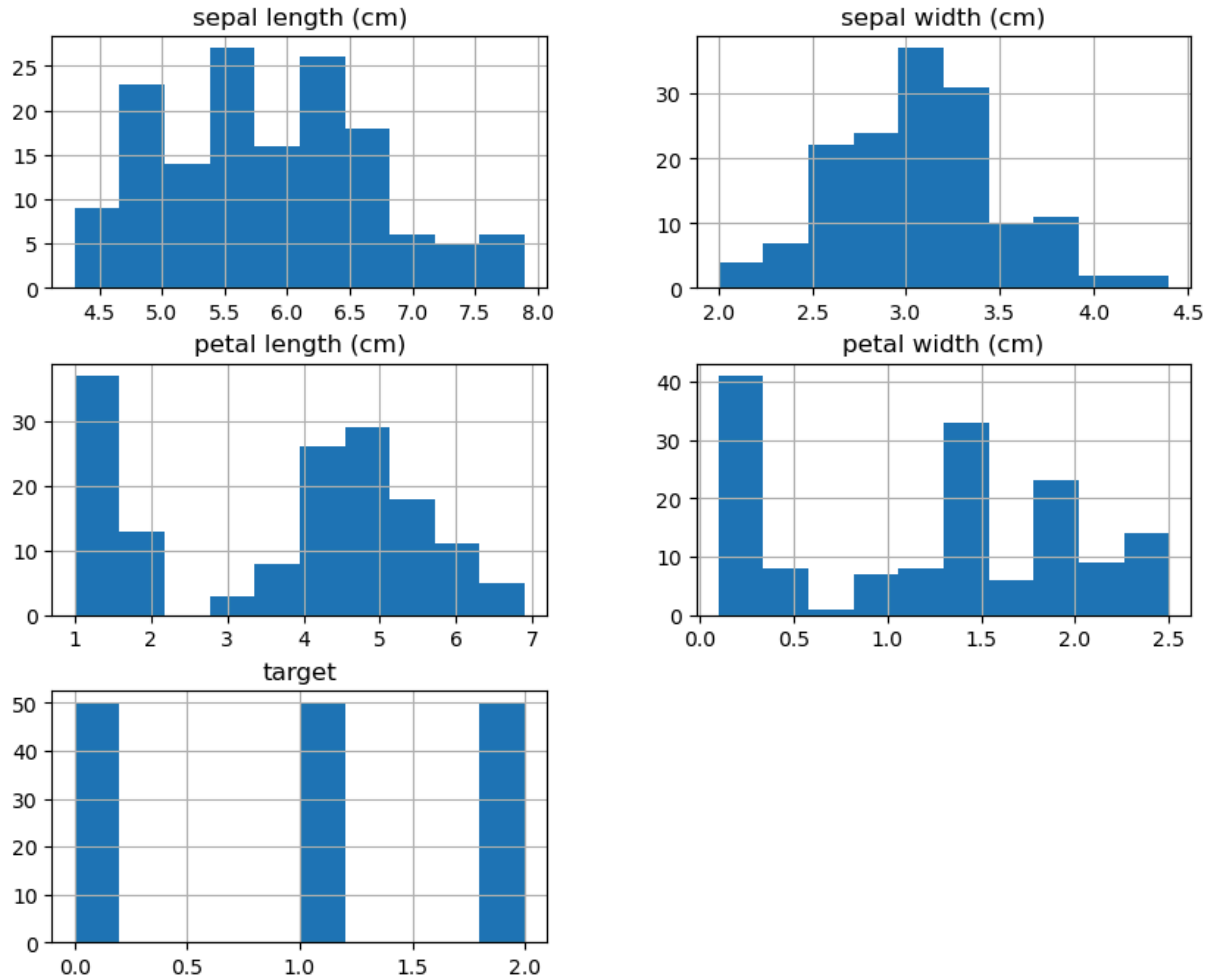
	sepal length (cm)	sepal width (cm)	petal length (cm)	\
count	150.000000	150.000000	150.000000	
mean	5.843333	3.057333	3.758000	
std	0.828066	0.435866	1.765298	
min	4.300000	2.000000	1.000000	
25%	5.100000	2.800000	1.600000	
50%	5.800000	3.000000	4.350000	
75%	6.400000	3.300000	5.100000	
max	7.900000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

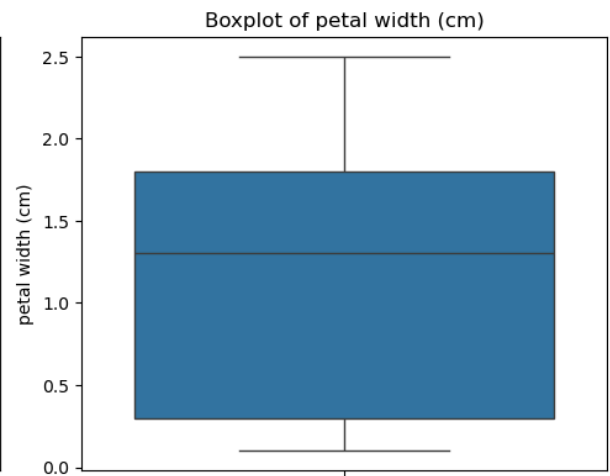
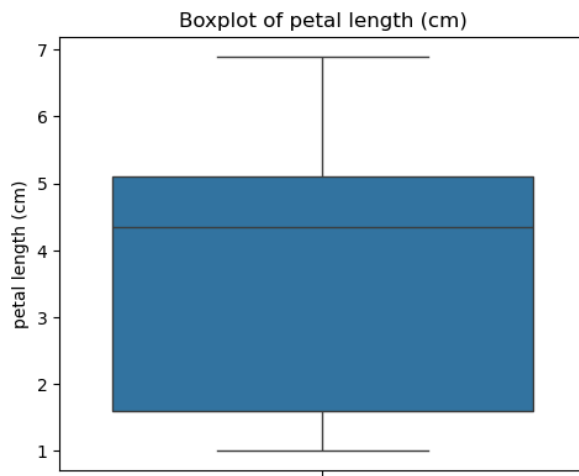
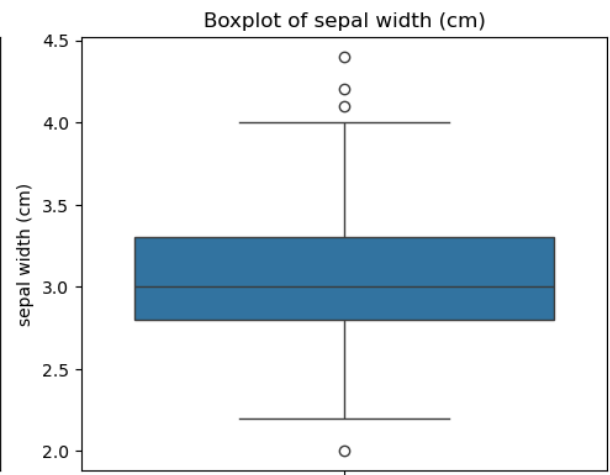
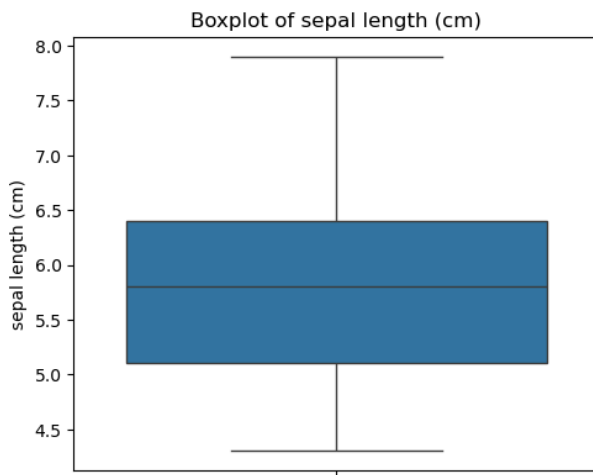
	petal width (cm)	target
count	150.000000	150.000000
mean	1.199333	1.000000
std	0.762238	0.819232
min	0.100000	0.000000
25%	0.300000	0.000000
50%	1.300000	1.000000
75%	1.800000	2.000000
max	2.500000	2.000000

```
In [2]: #Visualizations
# Histograms
df.hist(figsize=(10,8))
plt.suptitle("Histograms of Iris Features")
plt.show()
```

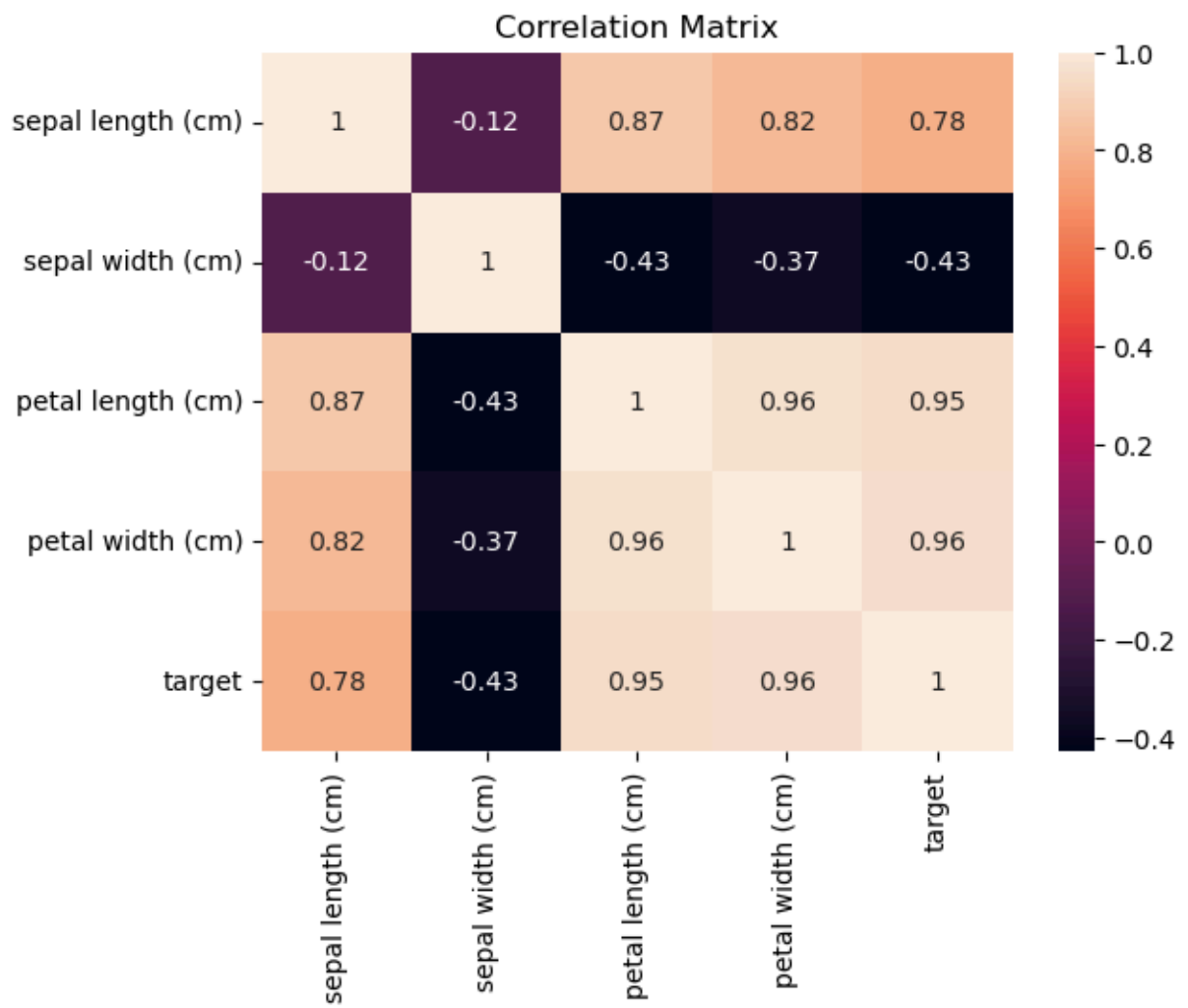
## Histograms of Iris Features



```
In [3]: # Boxplots
plt.figure(figsize=(10,8))
for i, col in enumerate(df.columns[:-1], 1):
    plt.subplot(2,2,i)
    sns.boxplot(y=df[col])
    plt.title(f"Boxplot of {col}")
plt.tight_layout()
plt.show()
```



```
In [5]: # Correlation matrix
sns.heatmap(df.corr(), annot=True)
plt.title("Correlation Matrix")
plt.show()
```



In [ ]: