```
In [1]:  import seaborn as sns

         # Load Iris dataset
         df = sns.load_dataset('iris')

         # Display first few rows
         df.head()
```

Out[1]:

	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

```
In [3]:  # Check for missing values
         df.isnull().sum()
```

```
Out[3]:  sepal_length    0
         sepal_width     0
         petal_length    0
         petal_width     0
         species         0
         dtype: int64
```
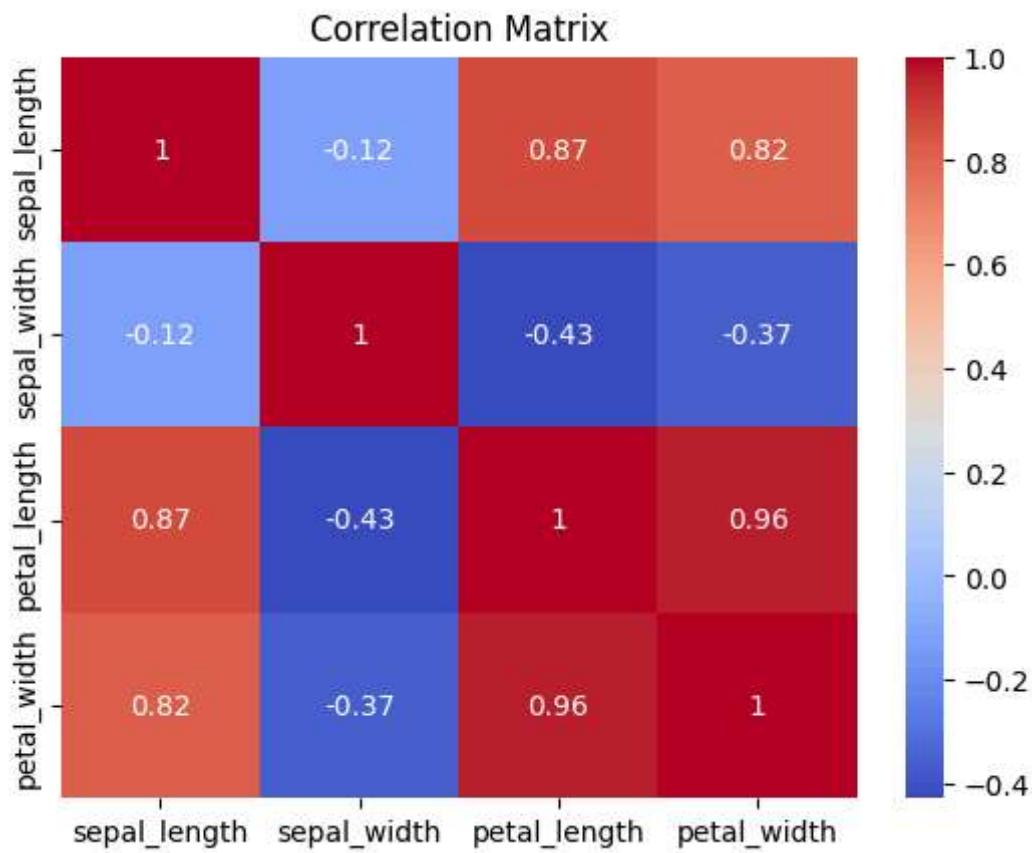
```
In [4]:  import matplotlib.pyplot as plt
         import seaborn as sns

         # Plotting distributions of all numerical columns
         sns.pairplot(df, hue='species')
         plt.show()
```
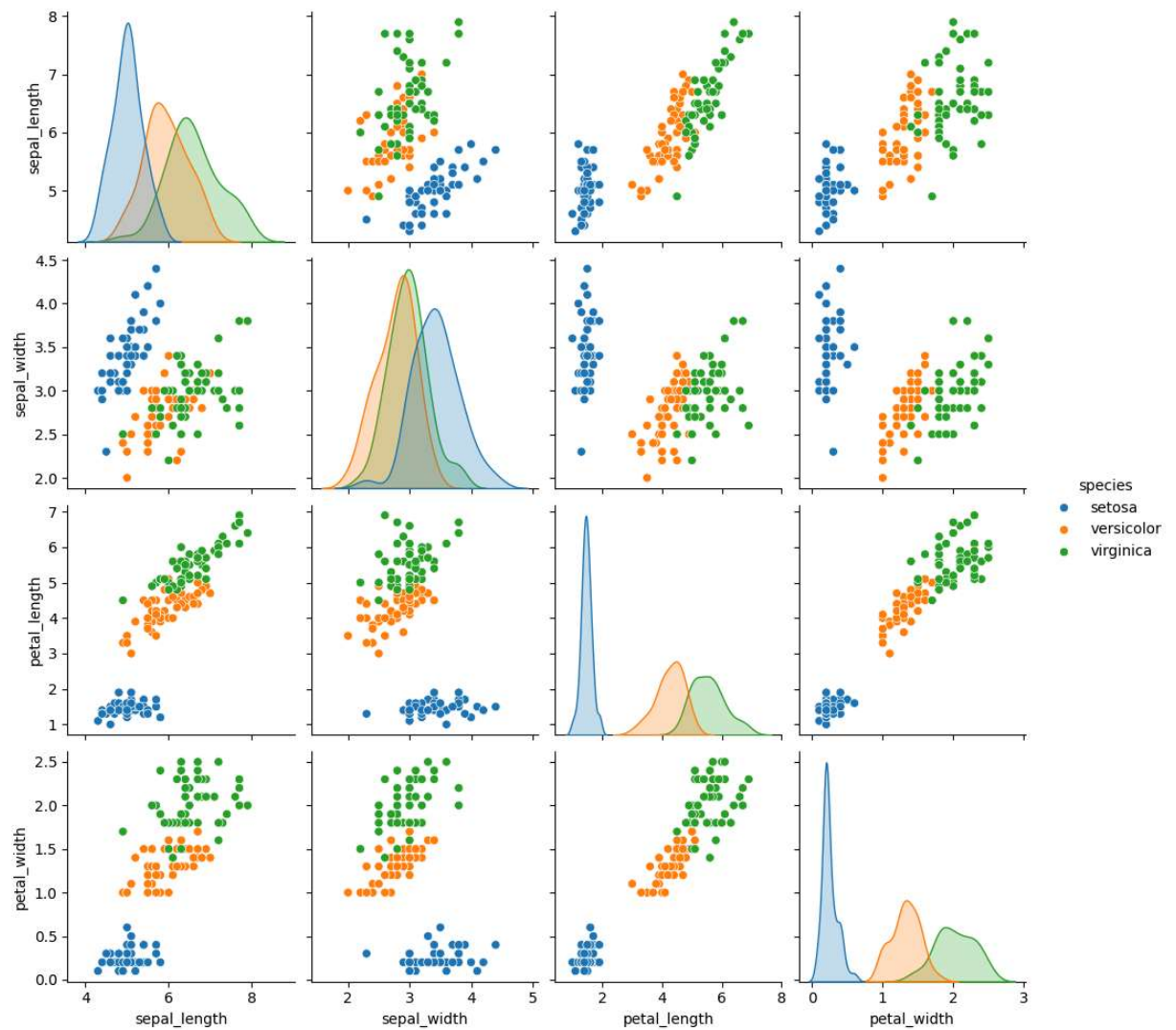
```
In [6]:   # Drop non-numeric column before correlation
          corr = df.drop("species", axis=1).corr()

          # Plot heatmap
          sns.heatmap(corr, annot=True, cmap="coolwarm")
          plt.title("Correlation Matrix")
          plt.show()
```

Correlation Matrix


```
In [7]:   # Boxplot for identifying outliers
          sns.boxplot(x='species', y='sepal_length', data=df)
          plt.title('Sepal Length Distribution by Species')
          plt.show()
```

Sepal Length Distribution by Species
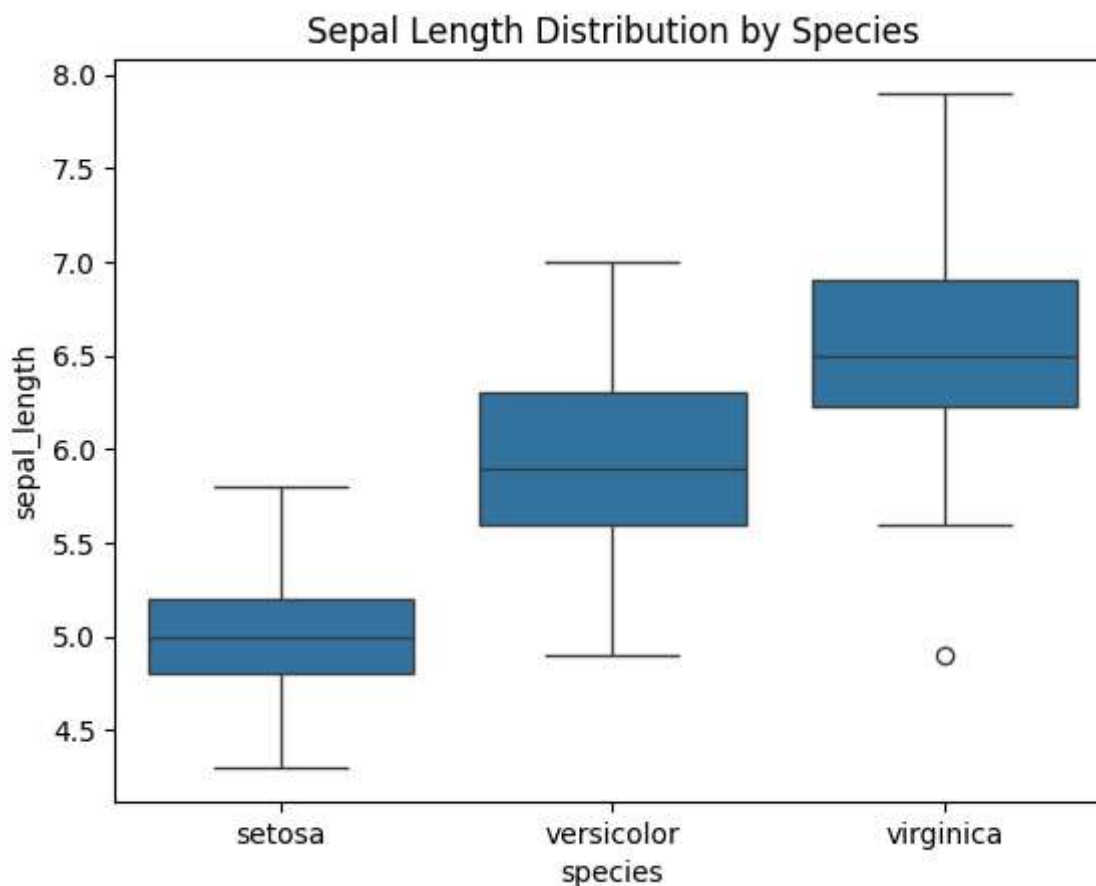

```
In [9]:  print("🔍 Generated Insights from EDA:\n")

         # 1. Correlation insight
         correlation = df.drop("species", axis=1).corr()
         petal_sepal_corr = correlation.loc['petal_length', 'sepal_length']
         print(f"1️⃣ The correlation between petal length and sepal length is {petal_sepal_c

         # 2. Average sepal length by species
         sepal_avg = df.groupby('species')['sepal_length'].mean()
         smallest_species = sepal_avg.idxmin()
         print(f"2️⃣ The species with the smallest average sepal length is **{smallest_speci

         # 3. Outliers detection using IQR for petal length in versicolor
         versicolor_petal = df[df['species'] == 'versicolor']['petal_length']
         Q1 = versicolor_petal.quantile(0.25)
         Q3 = versicolor_petal.quantile(0.75)
         IQR = Q3 - Q1
         outliers = versicolor_petal[(versicolor_petal < (Q1 - 1.5 * IQR)) | (versicolor_pet
         print(f"3️⃣ There are {outliers.shape[0]} outliers detected in petal length for the
```

🔍 Generated Insights from EDA:

1️⃣ The correlation between petal length and sepal length is 0.87 → High correlatio
n.
2️⃣ The species with the smallest average sepal length is **Setosa** (Avg: 5.01 cm).
3️⃣ There are 1 outliers detected in petal length for the species **Versicolor**.