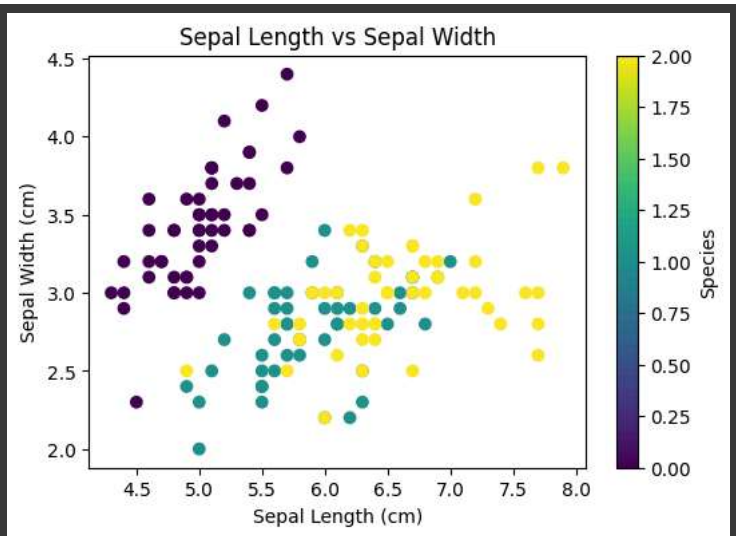
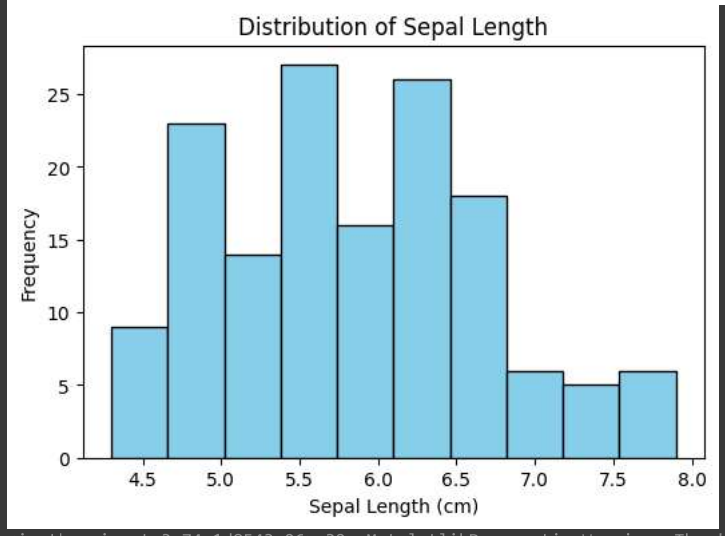
```python
import matplotlib.pyplot as plt
import seaborn as sns
import pandas as pd
from sklearn.datasets import load_iris

# Load a sample dataset (Iris dataset)
iris = load_iris(as_frame=True)
iris_data = iris.frame

# 1. Scatter Plot (Sepal Length vs Sepal Width)
plt.figure(figsize=(6, 4))
plt.scatter(iris_data['sepal length (cm)'], iris_data['sepal width (cm)'], c=iris_data['target'], cmap='viridis')
plt.title('Sepal Length vs Sepal Width')
plt.xlabel('Sepal Length (cm)')
plt.ylabel('Sepal Width (cm)')
plt.colorbar(label='Species')
plt.show()

# 2. Histogram (Distribution of Sepal Length)
plt.figure(figsize=(6, 4))
plt.hist(iris_data['sepal length (cm)'], bins=10, color='skyblue', edgecolor='black')
plt.title('Distribution of Sepal Length')
plt.xlabel('Sepal Length (cm)')
plt.ylabel('Frequency')
plt.show()
# Conclusion: Sepal lengths are distributed with most values concentrated between 5 and 6.5 cm.

# 3. Box Plot (Sepal Width by Species)
plt.figure(figsize=(6, 4))
species = ['Setosa', 'Versicolor', 'Virginica']
data_by_species = [iris_data[iris_data['target'] == i]['sepal width (cm)'] for i in range(3)]
plt.boxplot(data_by_species, labels=species)
plt.title('Sepal Width Distribution by Species')
plt.ylabel('Sepal Width (cm)')
plt.show()
# Conclusion: Setosa has the widest Sepal width range, while Virginica's is more compact.

# 4. Line Plot (Cumulative Sepal Length)
plt.figure(figsize=(6, 4))
plt.plot(iris_data['sepal length (cm)'].cumsum(), label='Cumulative Sepal Length', color='red')
plt.title('Cumulative Sepal Length')
plt.xlabel('Index')
plt.ylabel('Cumulative Sepal Length (cm)')
plt.legend()
plt.show()
# Conclusion: The cumulative sepal length shows a smooth upward trend as data is added.

# 5. Pie Chart (Species Distribution)
plt.figure(figsize=(6, 6))
species_counts = iris_data['target'].value_counts()
plt.pie(species_counts, labels=species, autopct='%1.1f%%', startangle=140, colors=['gold', 'lightgreen', 'skyblue'])
plt.title('Species Distribution')
plt.show()
# Conclusion: The dataset is evenly distributed among the three species.

# ----------------- Seaborn Plots -----------------
# 1. Pairplot (Relationships among features)
sns.pairplot(iris_data, hue='target', palette='deep')
plt.suptitle('Pairplot of Iris Features', y=1.02)
plt.show()
# Analysis: Pairwise relationships show clear separations for Setosa, while Versicolor and Virginica overlap.

# 2. Heatmap (Correlation Matrix)
correlation_matrix = iris_data.corr()
sns.heatmap(correlation_matrix, annot=True, cmap='coolwarm', fmt='.2f')
plt.title('Heatmap of Feature Correlations')
plt.show()
# Analysis: Petal length and width are strongly correlated, while Sepal features have weaker correlations.

# 3. Violin Plot (Petal Length by Species)
sns.violinplot(x=iris_data['target'], y=iris_data['petal length (cm)'], palette='muted')
plt.title('Petal Length Distribution by Species')
plt.xlabel('Species')
plt.ylabel('Petal Length (cm)')
plt.xticks([0, 1, 2], species)
plt.show()
```

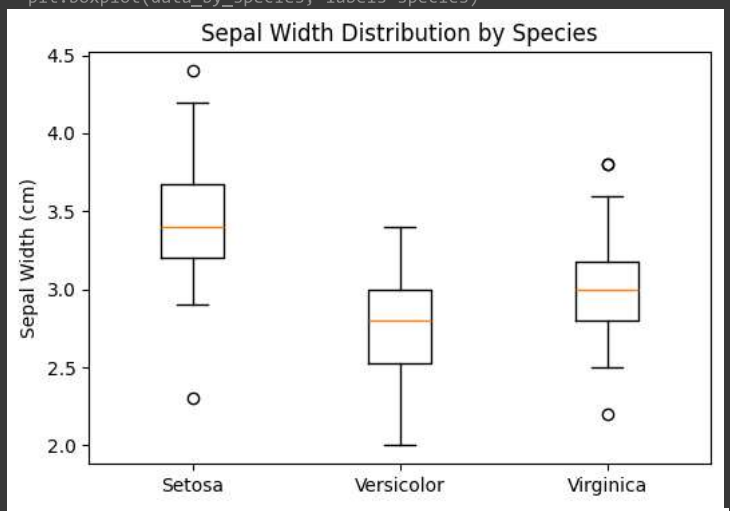
```
# Analysis: Setosa's petal length is distinctly smaller compared to the other species.

# 4. Swarm Plot (Sepal Length by Species)
sns.swarmplot(x='target', y='sepal length (cm)', data=iris_data, palette='pastel')
plt.title('Sepal Length by Species')
plt.xlabel('Species')
plt.ylabel('Sepal Length (cm)')
plt.xticks([0, 1, 2], species)
plt.show()
# Analysis: Virginica shows the longest Sepal lengths, while Setosa shows the shortest.

# 5. Count Plot (Species Count)
sns.countplot(x='target', data=iris_data, palette='bright')
plt.title('Count of Each Species')
plt.xlabel('Species')
plt.ylabel('Count')
plt.xticks([0, 1, 2], species)
plt.show()
# Analysis: All species are equally represented in the dataset (50 samples each).
import numpy as np
import matplotlib.pyplot as plt
from mpl_toolkits.mplot3d import Axes3D

# Load data (Iris dataset)
from sklearn.datasets import load_iris
iris = load_iris(as_frame=True)
iris_data = iris.frame

# 1. **3D Scatter Plot**: Sepal Length, Sepal Width, and Petal Length
fig = plt.figure(figsize=(8, 6))
ax = fig.add_subplot(111, projection='3d')

scatter = ax.scatter(
    iris_data['sepal length (cm)'],
    iris_data['sepal width (cm)'],
    iris_data['petal length (cm)'],
    c=iris_data['target'], cmap='viridis', s=50
)
ax.set_xlabel('Sepal Length (cm)')
ax.set_ylabel('Sepal Width (cm)')
ax.set_zlabel('Petal Length (cm)')
ax.set_title('3D Scatter Plot of Iris Features')
plt.colorbar(scatter, ax=ax, label='Species')
plt.show()
# Conclusion: Setosa is distinctly separate in 3D space, while Versicolor and Virginica show slight overlap.

# 2. **Subplots with Shared Axes**: Comparing Petal and Sepal Distributions
fig, axs = plt.subplots(2, 2, figsize=(10, 8), sharex=True, sharey=False)
fig.suptitle('Feature Comparisons', fontsize=16)

# Petal Length
axs[0, 0].hist(iris_data['petal length (cm)'], bins=15, color='skyblue', edgecolor='black')
axs[0, 0].set_title('Petal Length')
axs[0, 0].set_ylabel('Frequency')

# Petal Width
axs[0, 1].hist(iris_data['petal width (cm)'], bins=15, color='lightcoral', edgecolor='black')
axs[0, 1].set_title('Petal Width')

# Sepal Length
axs[1, 0].hist(iris_data['sepal length (cm)'], bins=15, color='lightgreen', edgecolor='black')
axs[1, 0].set_title('Sepal Length')
axs[1, 0].set_ylabel('Frequency')
axs[1, 0].set_xlabel('Measurement (cm)')

# Sepal Width
axs[1, 1].hist(iris_data['sepal width (cm)'], bins=15, color='gold', edgecolor='black')
axs[1, 1].set_title('Sepal Width')
axs[1, 1].set_xlabel('Measurement (cm)')

plt.tight_layout(rect=[0, 0, 1, 0.96])
plt.show()
# Conclusion: Petal features have more variation compared to Sepal features.

# 3. **Advanced Line Plot with Annotations**: Cumulative Sum of Petal Length
cumsum = iris_data['petal length (cm)'].cumsum()
fig, ax = plt.subplots(figsize=(8, 4))
```

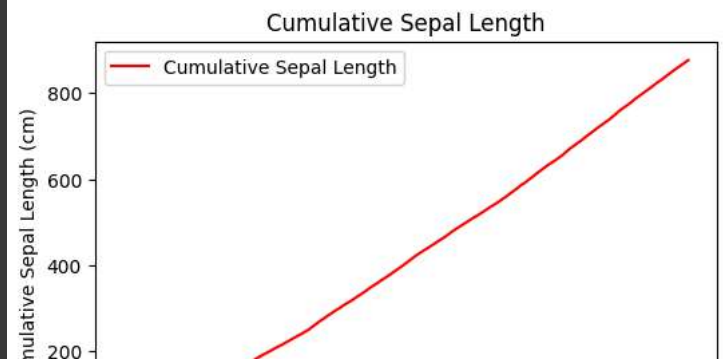
```
ax.plot(cumsum, label='Cumulative Petal Length', color='purple', lw=2)
ax.set_title('Cumulative Sum of Petal Length', fontsize=14)
ax.set_xlabel('Index')
ax.set_ylabel('Cumulative Petal Length (cm)')

# Annotating specific points
for idx in [30, 75, 120]:
    ax.annotate(f'{cumsum[idx]:.1f}', xy=(idx, cumsum[idx]), xytext=(idx + 5, cumsum[idx] + 5),
                arrowprops=dict(facecolor='black', arrowstyle='->'))
plt.legend()
plt.grid()
plt.show()
# Conclusion: The cumulative increase is linear due to evenly distributed values across species.

# 4. **Filled Area Plot**: Sepal Length by Species
species_colors = ['gold', 'lightgreen', 'lightcoral']
fig, ax = plt.subplots(figsize=(10, 5))
for i in range(3):
    data = iris_data[iris_data['target'] == i]['sepal length (cm)']
    ax.fill_between(range(len(data)), data, alpha=0.5, label=iris.target_names[i], color=species_colors[i])

ax.set_title('Sepal Length Distribution by Species')
ax.set_xlabel('Sample Index')
ax.set_ylabel('Sepal Length (cm)')
ax.legend()
plt.show()
# Conclusion: Virginica generally has the longest sepal lengths, while Setosa's values are consistently smaller.

# 5. **Polar Plot**: Proportions of Sepal vs Petal Dimensions
averages = [
    iris_data['sepal length (cm)'].mean(),
    iris_data['sepal width (cm)'].mean(),
    iris_data['petal length (cm)'].mean(),
    iris_data['petal width (cm)'].mean()
]

angles = np.linspace(0, 2 * np.pi, len(averages), endpoint=False).tolist()
angles += angles[:1]  # Close the loop

fig, ax = plt.subplots(figsize=(6, 6), subplot_kw=dict(polar=True))
averages += averages[:1]  # Close the loop
ax.plot(angles, averages, color='blue', linewidth=2, label='Average Values')
ax.fill(angles, averages, color='blue', alpha=0.25)

ax.set_yticks([1, 2, 3, 4])
ax.set_xticks(angles[:-1])
ax.set_xticklabels(['Sepal Length', 'Sepal Width', 'Petal Length', 'Petal Width'])
ax.set_title('Radar Chart of Average Dimensions')
ax.legend(loc='upper right')
plt.show()
# Conclusion: Petal dimensions show higher variability, while Sepal dimensions are relatively balanced.

```

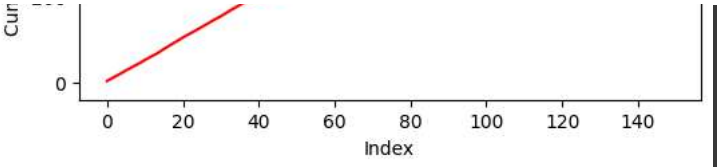


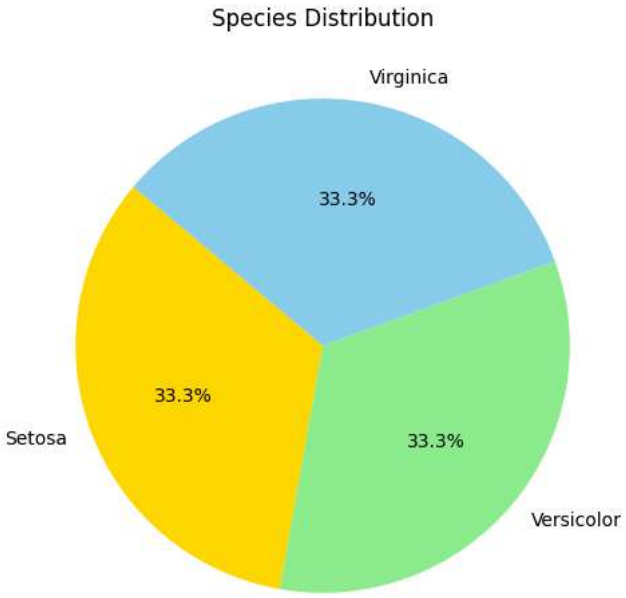


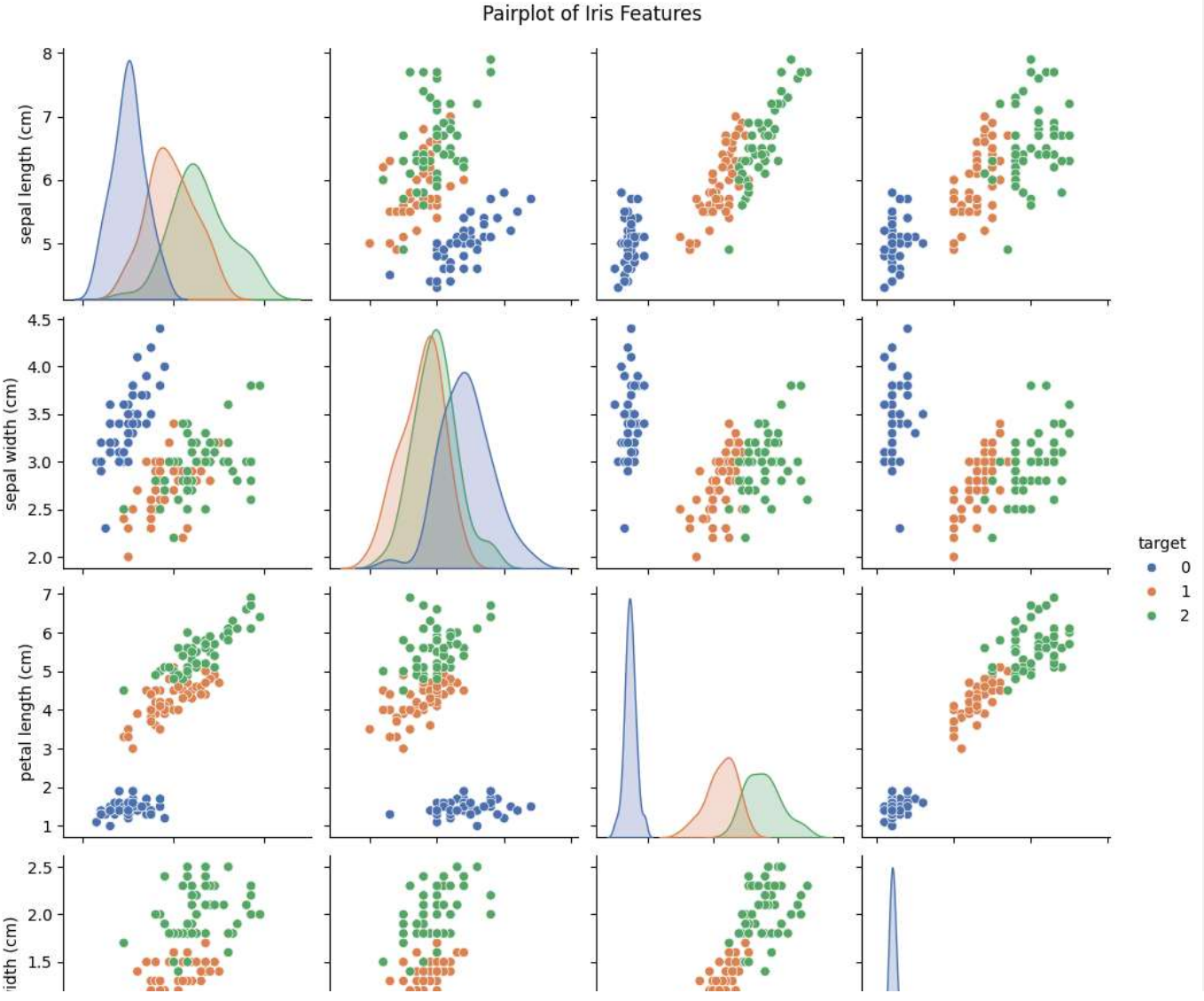

```
<ipython-input-3-74c1d8543e06>:32: MatplotlibDeprecationWarning: The 'labels' parameter of boxplot() has been renamed 'tick_labels' s
  plt.boxplot(data_by_species, labels=species)
```

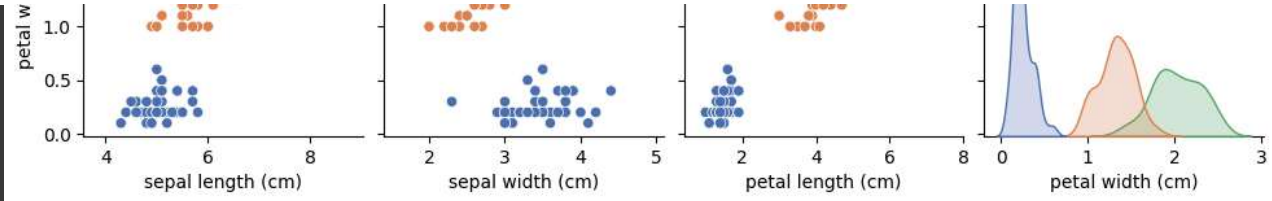


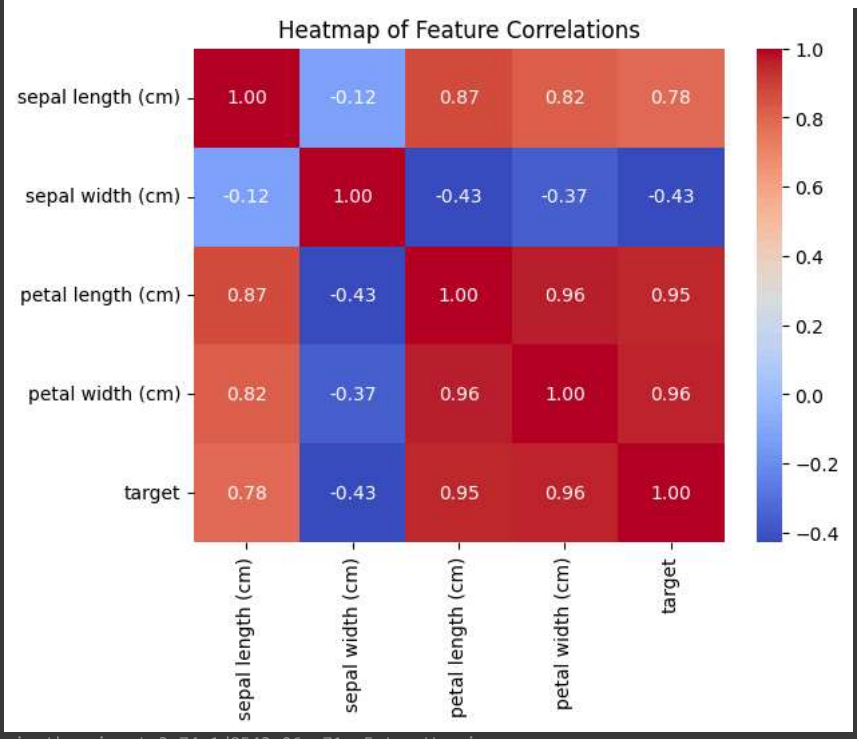



0
0
20
40
60
80
100
120
140
Index

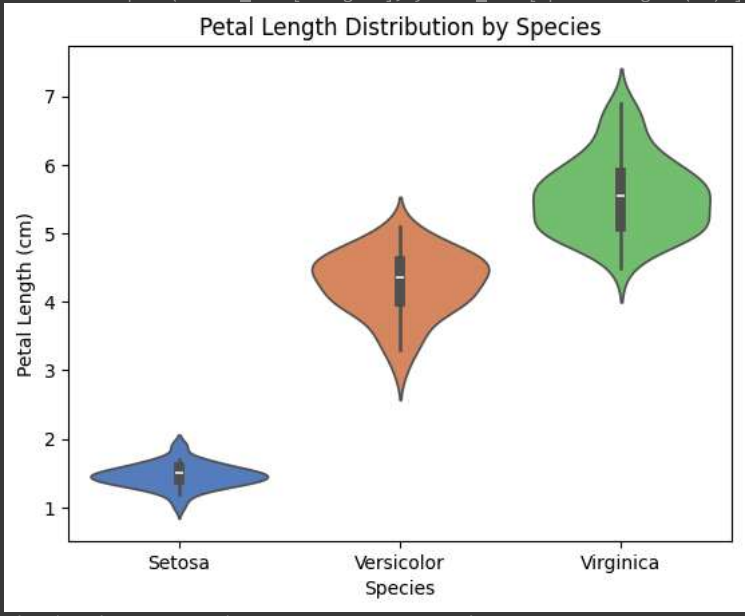

Species Distribution
Virginica
33.3%
33.3%
Setosa
33.3%
Versicolor


Pairplot of Iris Features
sepal length (cm)
sepal width (cm)
petal length (cm)
target
0
1
2

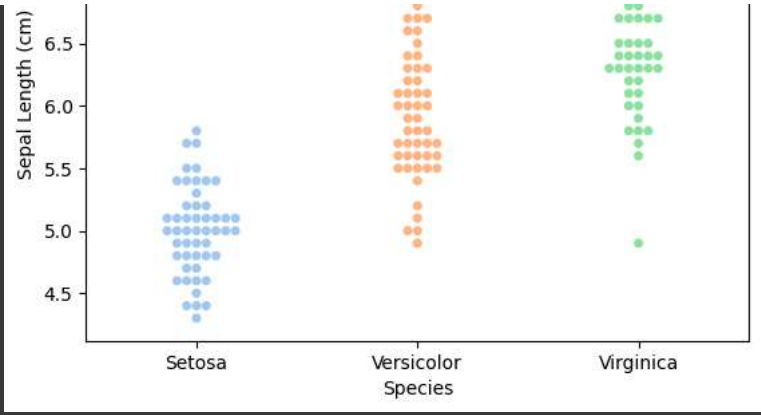




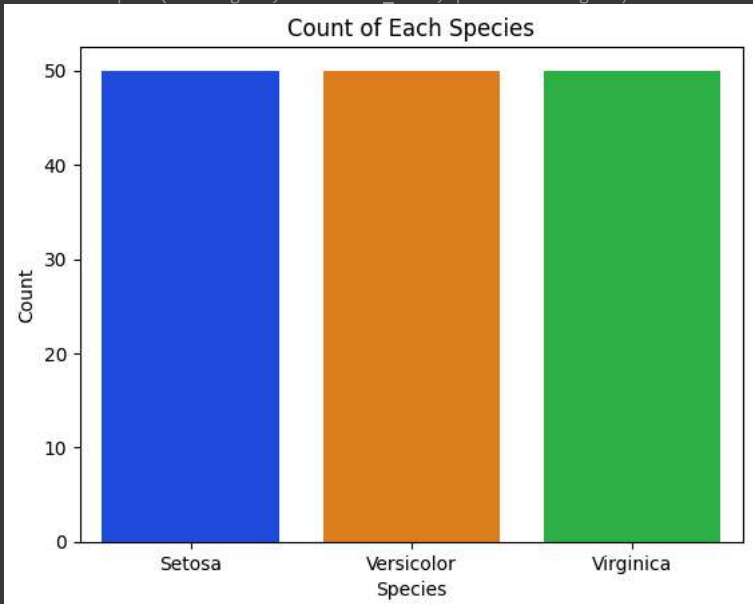

```
<ipython-input-3-74c1d8543e06>:71: FutureWarning:

Passing `palette` without assigning `hue` is deprecated and will be removed in v0.14.0. Assign the `x` variable to `hue` and set `leg

  sns.violinplot(x=iris_data['target'], y=iris_data['petal length (cm)'], palette='muted')
```

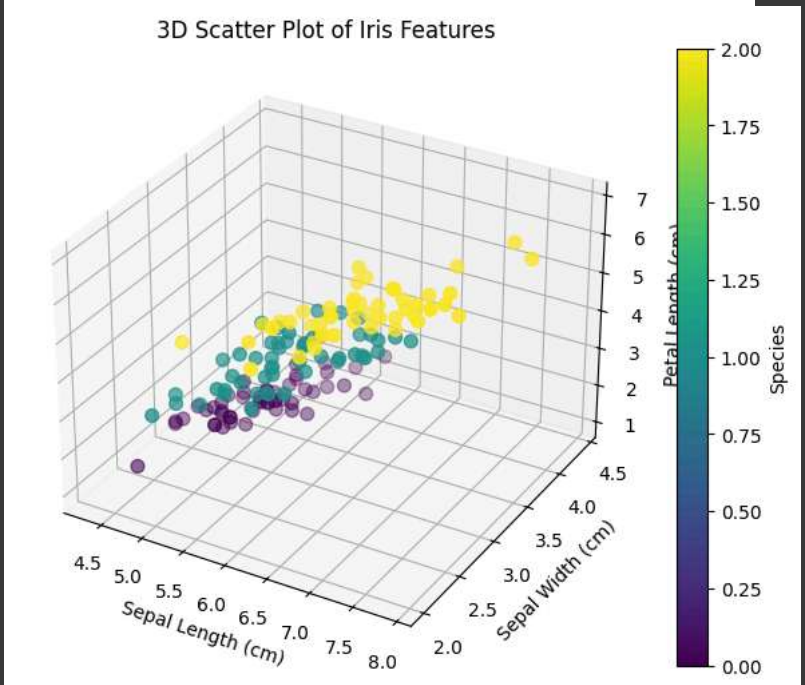


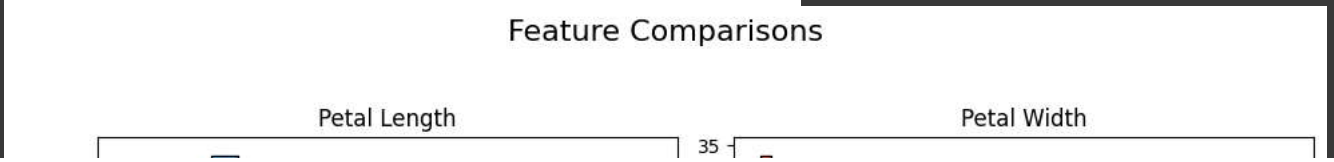

```
<ipython-input-3-74c1d8543e06>:80: FutureWarning:

Passing `palette` without assigning `hue` is deprecated and will be removed in v0.14.0. Assign the `x` variable to `hue` and set `leg

  sns.swarmplot(x='target', y='sepal length (cm)', data=iris_data, palette='pastel')
```

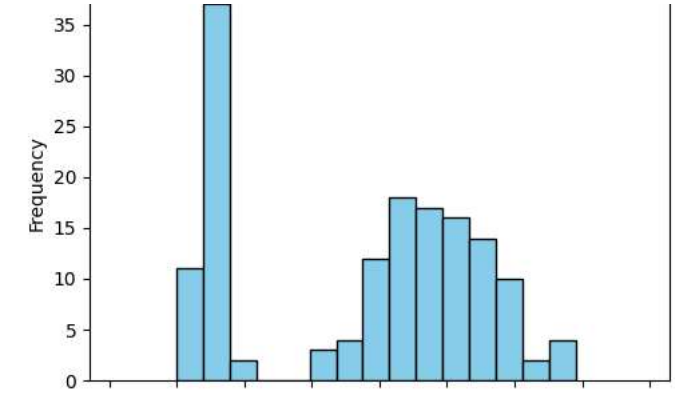


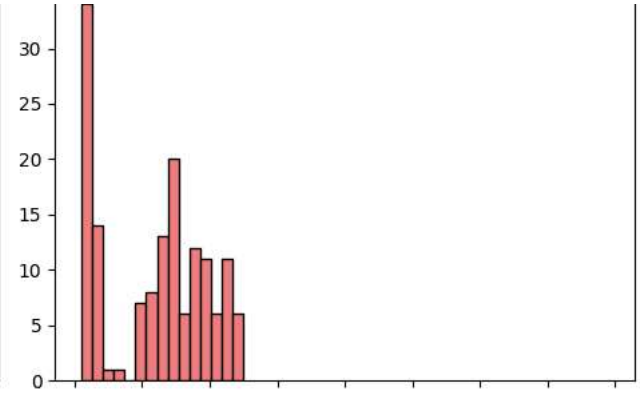

```
<ipython-input-3-74c1d8543e06>:89: FutureWarning:

Passing `palette` without assigning `hue` is deprecated and will be removed in v0.14.0. Assign the `x` variable to `hue` and set `leg

  sns.countplot(x='target', data=iris_data, palette='bright')
```

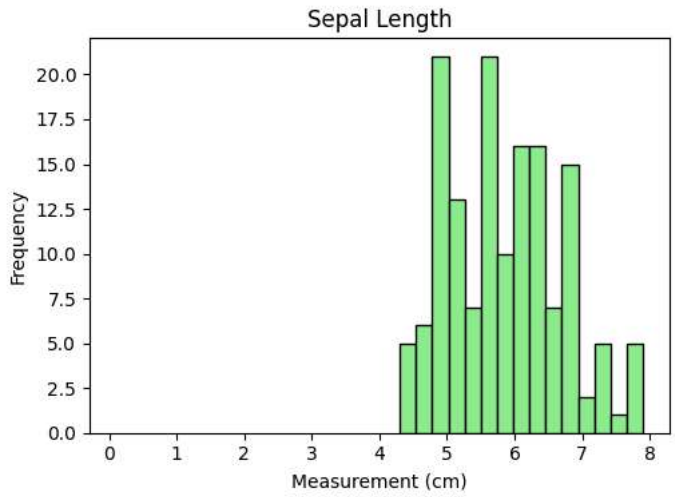


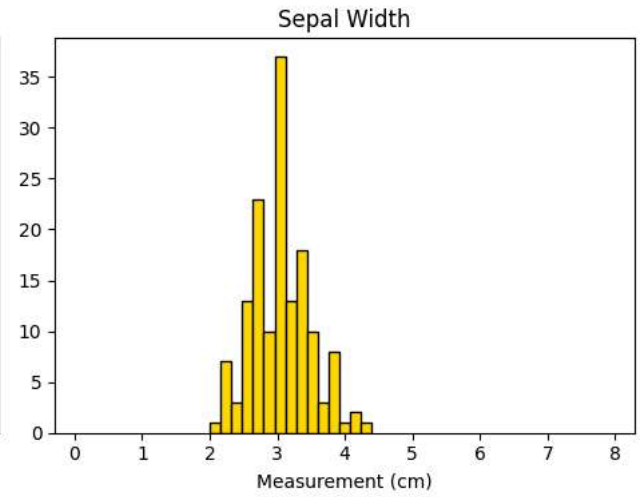


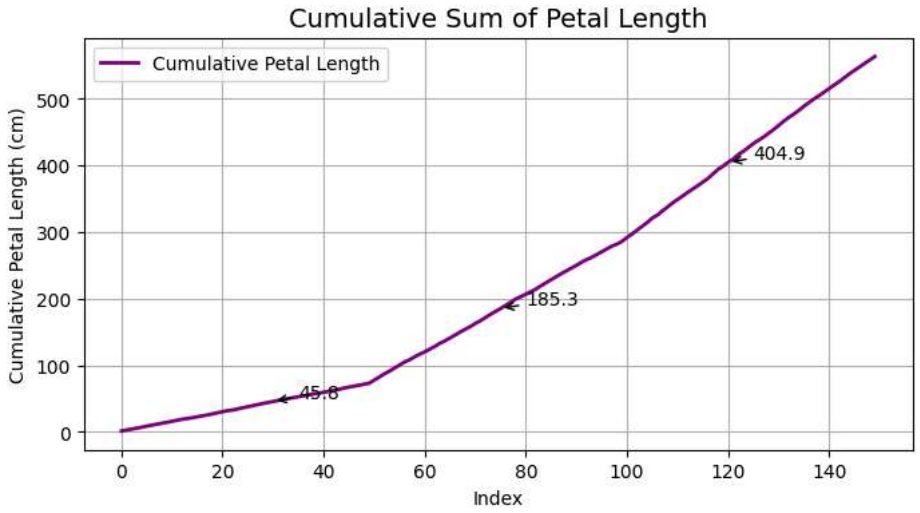


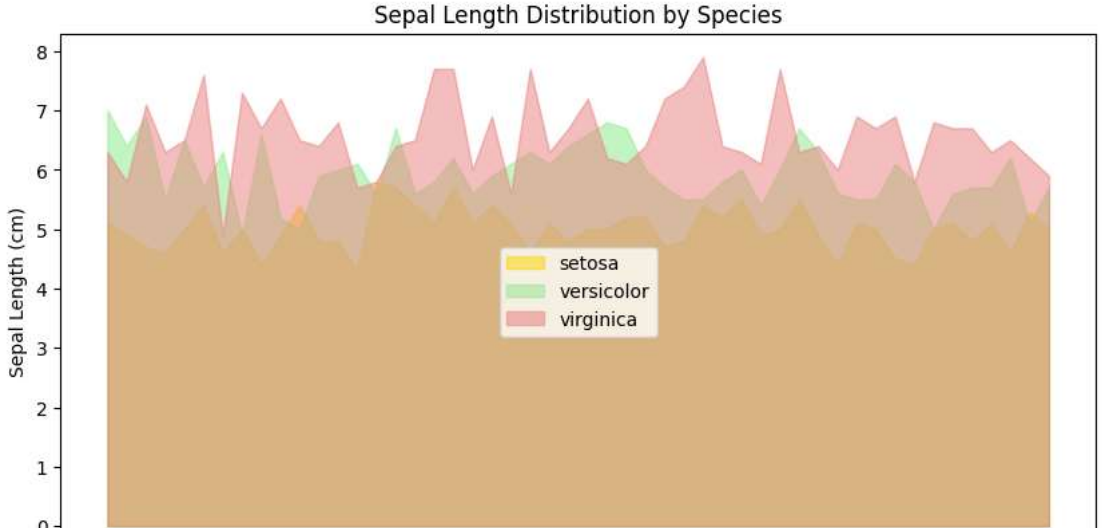

Frequency

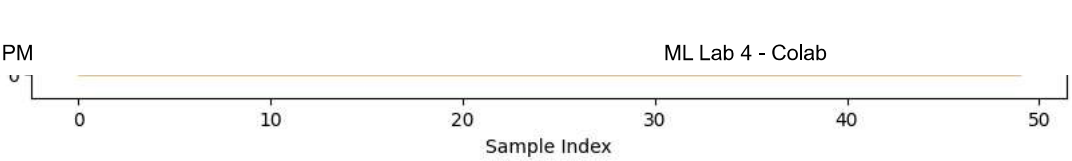


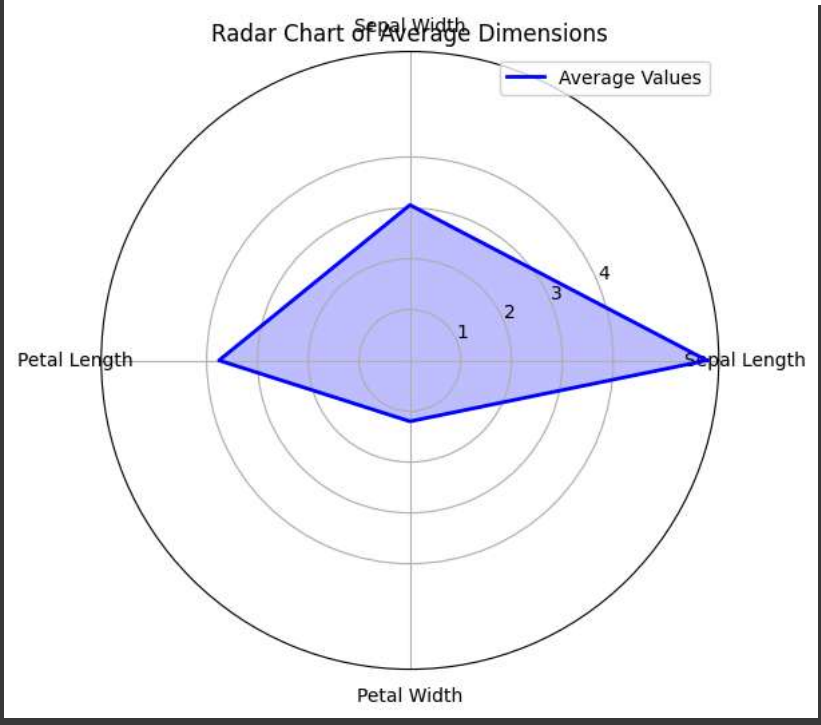

Sepal Length
Frequency
Measurement (cm)


Sepal Width
Measurement (cm)


Cumulative Sum of Petal Length
Cumulative Petal Length
Cumulative Petal Length (cm)
404.9
185.3
45.8
Index


Sepal Length Distribution by Species
Sepal Length (cm)
setosa
versicolor
virginica

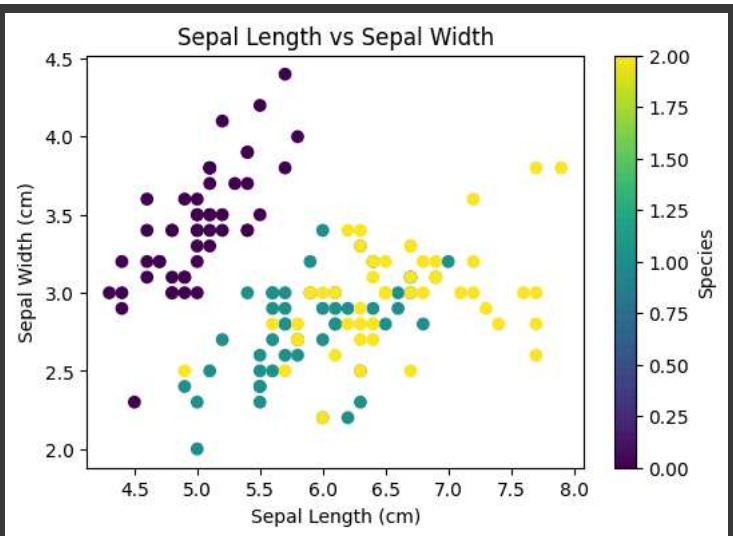

0
0
10
20
30
40
50
Sample Index

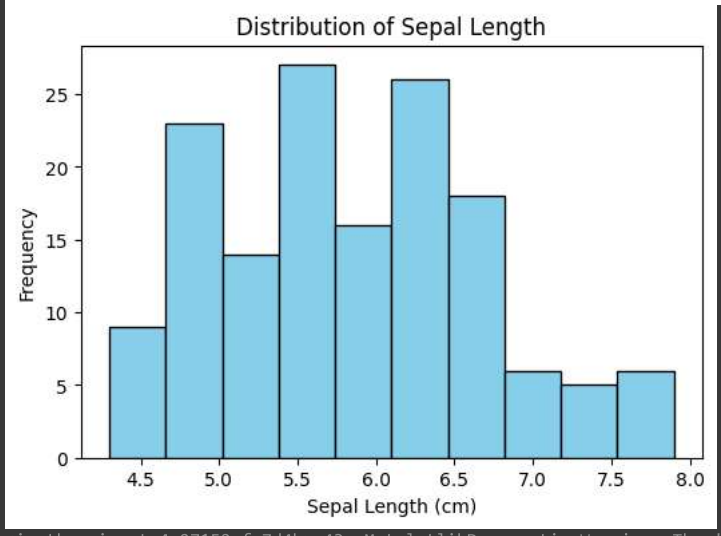

Radar Chart of Average Dimensions
Average Values
Sepal Width
Petal Length
Sepal Length
Petal Width
1
2
3
4

```
# -*- coding: utf-8 -*-
"""Untitled17.ipynb

Automatically generated by Colab.

Original file is located at
    https://colab.research.google.com/drive/1c4w-nRN7Kr-C1VBDaimwgvG2Wo2GnMKw
"""

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

# Load a sample dataset (Iris dataset)
iris = load_iris(as_frame=True)
iris_data = iris.frame

# ---------------- Matplotlib Plots ----------------
# 1. Scatter Plot (Sepal Length vs Sepal Width)
plt.figure(figsize=(6, 4))
plt.scatter(iris_data['sepal length (cm)'], iris_data['sepal width (cm)'], c=iris_data['target'], cmap='viridis')
plt.title('Sepal Length vs Sepal Width')
plt.xlabel('Sepal Length (cm)')
plt.ylabel('Sepal Width (cm)')
plt.colorbar(label='Species')
plt.show()
# Conclusion: Different species have slight overlaps in Sepal dimensions but show clusters.

# 2. Histogram (Distribution of Sepal Length)
plt.figure(figsize=(6, 4))
plt.hist(iris_data['sepal length (cm)'], bins=10, color='skyblue', edgecolor='black')
plt.title('Distribution of Sepal Length')
plt.xlabel('Sepal Length (cm)')
plt.ylabel('Frequency')
plt.show()
# Conclusion: Sepal lengths are distributed with most values concentrated between 5 and 6.5 cm.

# 3. Box Plot (Sepal Width by Species)
plt.figure(figsize=(6, 4))
species = ['Setosa', 'Versicolor', 'Virginica']
data_by_species = [iris_data[iris_data['target'] == i]['sepal width (cm)'] for i in range(3)]
plt.boxplot(data_by_species, labels=species)
plt.title('Sepal Width Distribution by Species')
plt.ylabel('Sepal Width (cm)')
plt.show()
# Conclusion: Setosa has the widest Sepal width range, while Virginica's is more compact.

# 4. Line Plot (Cumulative Sepal Length)
plt.figure(figsize=(6, 4))
plt.plot(iris_data['sepal length (cm)'].cumsum(), label='Cumulative Sepal Length', color='red')
plt.title('Cumulative Sepal Length')
plt.xlabel('Index')
plt.ylabel('Cumulative Sepal Length (cm)')
plt.legend()
plt.show()
# Conclusion: The cumulative sepal length shows a smooth upward trend as data is added.

# 5. Pie Chart (Species Distribution)
plt.figure(figsize=(6, 6))
species_counts = iris_data['target'].value_counts()
plt.pie(species_counts, labels=species, autopct='%1.1f%%', startangle=140, colors=['gold', 'lightgreen', 'skyblue'])
plt.title('Species Distribution')
plt.show()
# Conclusion: The dataset is evenly distributed among the three species.

# ---------------- Seaborn Plots ----------------
# 1. Pairplot (Relationships among features)
sns.pairplot(iris_data, hue='target', palette='deep')
plt.suptitle('Pairplot of Iris Features', y=1.02)
plt.show()
# Analysis: Pairwise relationships show clear separations for Setosa, while Versicolor and Virginica overlap.

# 2. Heatmap (Correlation Matrix)
correlation_matrix = iris_data.corr()
sns.heatmap(correlation_matrix, annot=True, cmap='coolwarm', fmt='.2f')
plt.title('Heatmap of Feature Correlations')
```

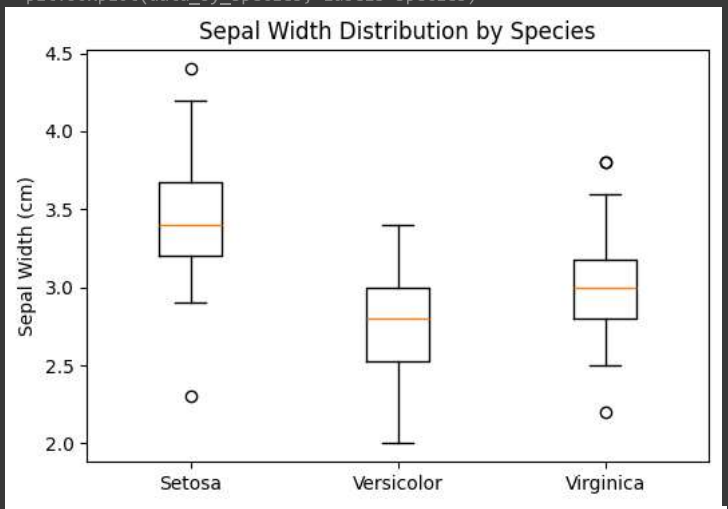
```
plt.show()
# Analysis: Petal length and width are strongly correlated, while Sepal features have weaker correlations.

# 3. Violin Plot (Petal Length by Species)
sns.violinplot(x=iris_data['target'], y=iris_data['petal length (cm)'], palette='muted')
plt.title('Petal Length Distribution by Species')
plt.xlabel('Species')
plt.ylabel('Petal Length (cm)')
plt.xticks([0, 1, 2], species)
plt.show()
# Analysis: Setosa's petal length is distinctly smaller compared to the other species.

# 4. Swarm Plot (Sepal Length by Species)
sns.swarmplot(x='target', y='sepal length (cm)', data=iris_data, palette='pastel')
plt.title('Sepal Length by Species')
plt.xlabel('Species')
plt.ylabel('Sepal Length (cm)')
plt.xticks([0, 1, 2], species)
plt.show()
# Analysis: Virginica shows the longest Sepal lengths, while Setosa shows the shortest.

# 5. Count Plot (Species Count)
sns.countplot(x='target', data=iris_data, palette='bright')
plt.title('Count of Each Species')
plt.xlabel('Species')
plt.ylabel('Count')
plt.xticks([0, 1, 2], species)
plt.show()
# Analysis: All species are equally represented in the dataset (50 samples each).

import numpy as np
import matplotlib.pyplot as plt
from mpl_toolkits.mplot3d import Axes3D

# Load data (Iris dataset)
from sklearn.datasets import load_iris
iris = load_iris(as_frame=True)
iris_data = iris.frame

# 1. **3D Scatter Plot**: Sepal Length, Sepal Width, and Petal Length
fig = plt.figure(figsize=(8, 6))
ax = fig.add_subplot(111, projection='3d')

scatter = ax.scatter(
    iris_data['sepal length (cm)'],
    iris_data['sepal width (cm)'],
    iris_data['petal length (cm)'],
    c=iris_data['target'], cmap='viridis', s=50
)
ax.set_xlabel('Sepal Length (cm)')
ax.set_ylabel('Sepal Width (cm)')
ax.set_zlabel('Petal Length (cm)')
ax.set_title('3D Scatter Plot of Iris Features')
plt.colorbar(scatter, ax=ax, label='Species')
plt.show()
# Conclusion: Setosa is distinctly separate in 3D space, while Versicolor and Virginica show slight overlap.

# 2. **Subplots with Shared Axes**: Comparing Petal and Sepal Distributions
fig, axs = plt.subplots(2, 2, figsize=(10, 8), sharex=True, sharey=False)
fig.suptitle('Feature Comparisons', fontsize=16)

# Petal Length
axs[0, 0].hist(iris_data['petal length (cm)'], bins=15, color='skyblue', edgecolor='black')
axs[0, 0].set_title('Petal Length')
axs[0, 0].set_ylabel('Frequency')

# Petal Width
axs[0, 1].hist(iris_data['petal width (cm)'], bins=15, color='lightcoral', edgecolor='black')
axs[0, 1].set_title('Petal Width')

# Sepal Length
axs[1, 0].hist(iris_data['sepal length (cm)'], bins=15, color='lightgreen', edgecolor='black')
axs[1, 0].set_title('Sepal Length')
axs[1, 0].set_ylabel('Frequency')
axs[1, 0].set_xlabel('Measurement (cm)')

# Sepal Width
```

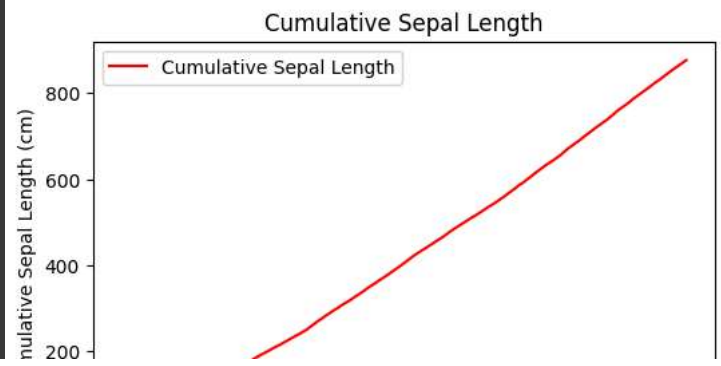
```
axs[1, 1].hist(iris_data['sepal width (cm)'], bins=15, color='gold', edgecolor='black')
axs[1, 1].set_title('Sepal Width')
axs[1, 1].set_xlabel('Measurement (cm)')

plt.tight_layout(rect=[0, 0, 1, 0.96])
plt.show()
# Conclusion: Petal features have more variation compared to Sepal features.

# 3. **Advanced Line Plot with Annotations**: Cumulative Sum of Petal Length
cumsum = iris_data['petal length (cm)'].cumsum()
fig, ax = plt.subplots(figsize=(8, 4))
ax.plot(cumsum, label='Cumulative Petal Length', color='purple', lw=2)
ax.set_title('Cumulative Sum of Petal Length', fontsize=14)
ax.set_xlabel('Index')
ax.set_ylabel('Cumulative Petal Length (cm)')

# Annotating specific points
for idx in [30, 75, 120]:
    ax.annotate(f'{cumsum[idx]:.1f}', xy=(idx, cumsum[idx]), xytext=(idx + 5, cumsum[idx] + 5),
                arrowprops=dict(facecolor='black', arrowstyle='->'))
plt.legend()
plt.grid()
plt.show()
# Conclusion: The cumulative increase is linear due to evenly distributed values across species.

# 4. **Filled Area Plot**: Sepal Length by Species
species_colors = ['gold', 'lightgreen', 'lightcoral']
fig, ax = plt.subplots(figsize=(10, 5))
for i in range(3):
    data = iris_data[iris_data['target'] == i]['sepal length (cm)']
    ax.fill_between(range(len(data)), data, alpha=0.5, label=iris.target_names[i], color=species_colors[i])

ax.set_title('Sepal Length Distribution by Species')
ax.set_xlabel('Sample Index')
ax.set_ylabel('Sepal Length (cm)')
ax.legend()
plt.show()
# Conclusion: Virginica generally has the longest sepal lengths, while Setosa's values are consistently smaller.

# 5. **Polar Plot**: Proportions of Sepal vs Petal Dimensions
averages = [
    iris_data['sepal length (cm)'].mean(),
    iris_data['sepal width (cm)'].mean(),
    iris_data['petal length (cm)'].mean(),
    iris_data['petal width (cm)'].mean()
]

angles = np.linspace(0, 2 * np.pi, len(averages), endpoint=False).tolist()
angles += angles[:1]  # Close the loop

fig, ax = plt.subplots(figsize=(6, 6), subplot_kw=dict(polar=True))
averages += averages[:1]  # Close the loop
ax.plot(angles, averages, color='blue', linewidth=2, label='Average Values')
ax.fill(angles, averages, color='blue', alpha=0.25)

ax.set_yticks([1, 2, 3, 4])
ax.set_xticks(angles[:-1])
ax.set_xticklabels(['Sepal Length', 'Sepal Width', 'Petal Length', 'Petal Width'])
ax.set_title('Radar Chart of Average Dimensions')
ax.legend(loc='upper right')
plt.show()
# Conclusion: Petal dimensions show higher variability, while Sepal dimensions are relatively balanced.
```

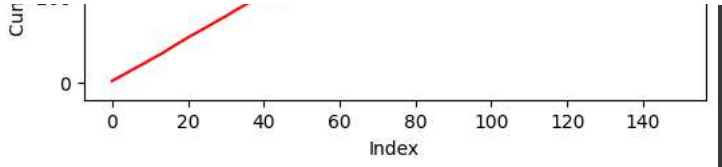


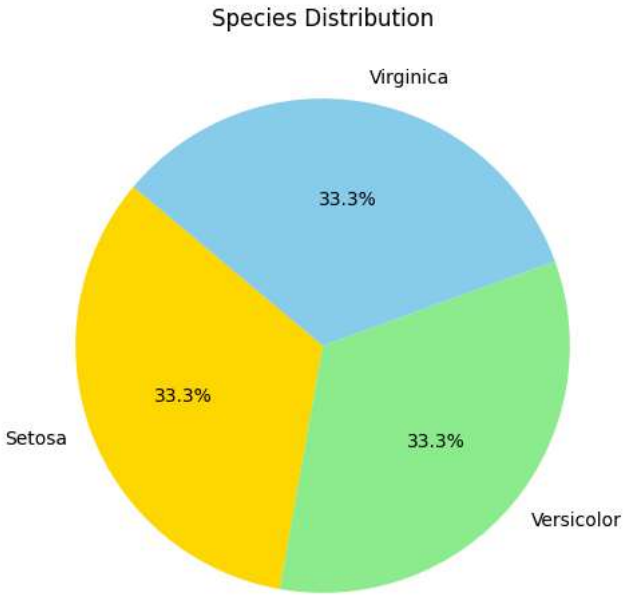


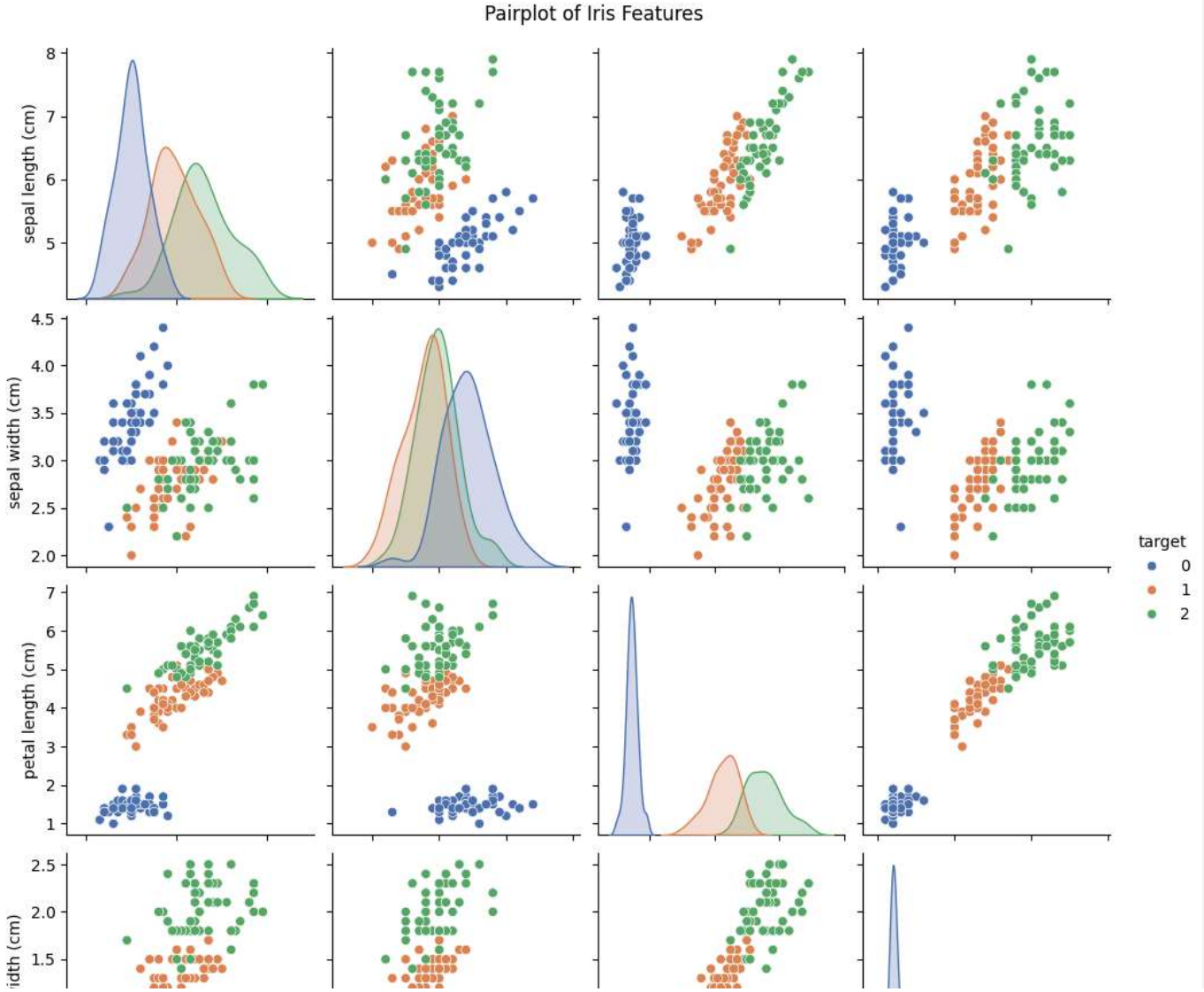

```
<ipython-input-4-27152afe7d4b>:43: MatplotlibDeprecationWarning: The 'labels' parameter of boxplot() has been renamed 'tick_labels' s
  plt.boxplot(data_by_species, labels=species)
```

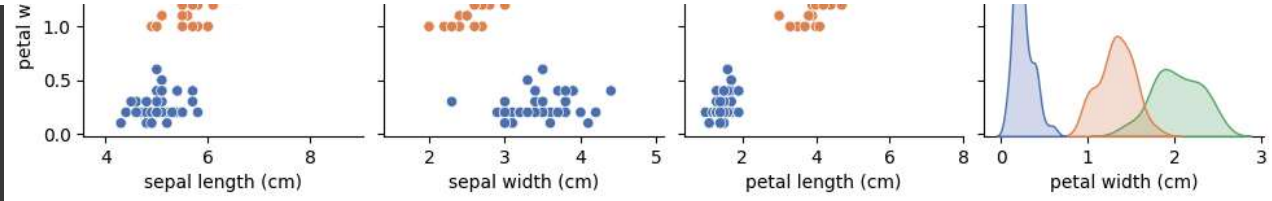


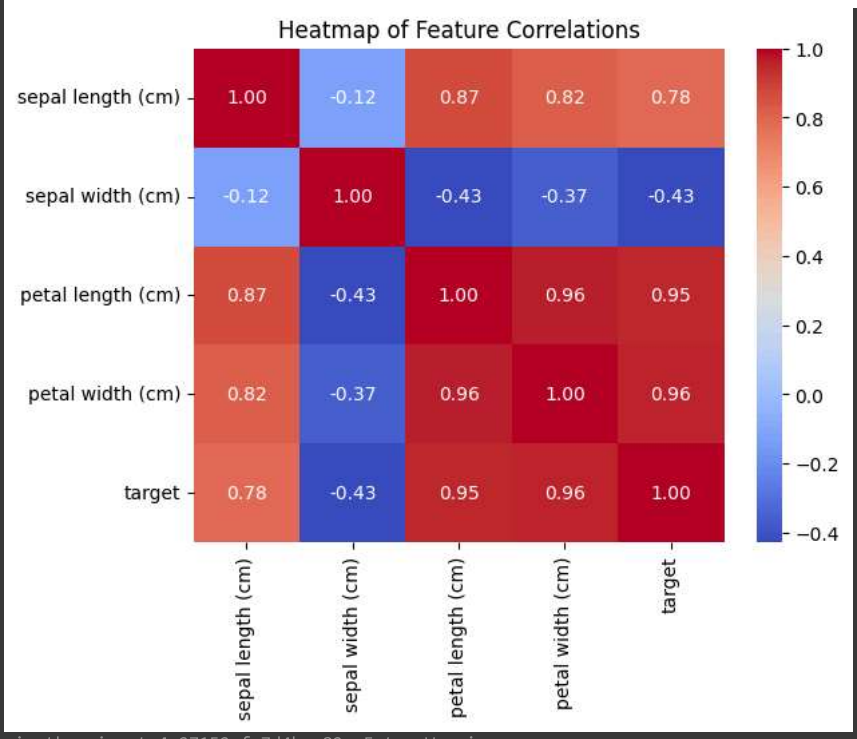



0
0
20
40
60
80
100
120
140
Index

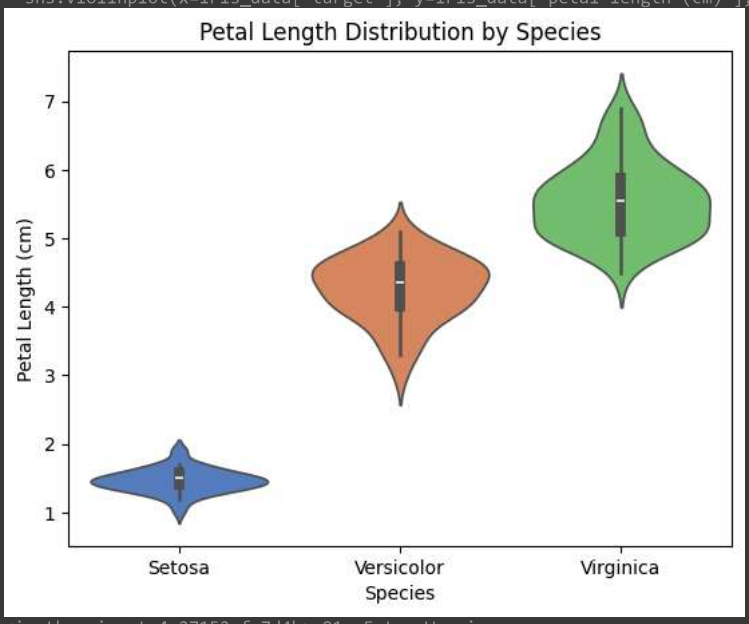

Species Distribution
Virginica
33.3%
33.3%
Setosa
33.3%
Versicolor


Pairplot of Iris Features
sepal length (cm)
sepal width (cm)
petal length (cm)
target
0
1
2

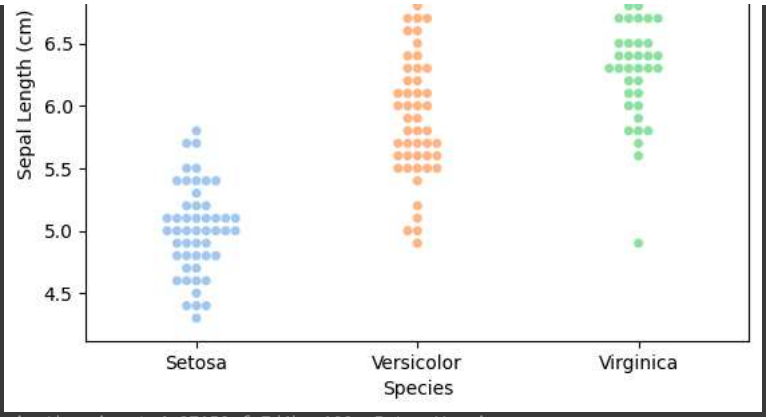





```
<ipython-input-4-27152afe7d4b>:82: FutureWarning:

Passing `palette` without assigning `hue` is deprecated and will be removed in v0.14.0. Assign the `x` variable to `hue` and set `leg

  sns.violinplot(x=iris_data['target'], y=iris_data['petal length (cm)'], palette='muted')
```




```
<ipython-input-4-27152afe7d4b>:91: FutureWarning:

Passing `palette` without assigning `hue` is deprecated and will be removed in v0.14.0. Assign the `x` variable to `hue` and set `leg

  sns.swarmplot(x='target', y='sepal length (cm)', data=iris_data, palette='pastel')
```




```
<ipython-input-4-27152afe7d4b>:100: FutureWarning:

Passing `palette` without assigning `hue` is deprecated and will be removed in v0.14.0. Assign the `x` variable to `hue` and set `leg

  sns.countplot(x='target', data=iris_data, palette='bright')
```