

```

1 import matplotlib.pyplot as plt
2 import seaborn as sns
3 import pandas as pd
4 from sklearn.datasets import load_iris
5
6 # Load a sample dataset (Iris dataset)
7 iris = load_iris(as_frame=True)
8 iris_data = iris.frame
9
10 # 1. Scatter Plot (Sepal Length vs Sepal Width)
11 plt.figure(figsize=(6, 4))
12 plt.scatter(iris_data['sepal length (cm)'], iris_data['sepal width (cm)'], c=iris_data['target'], cmap='viridis')
13 plt.title('Sepal Length vs Sepal Width')
14 plt.xlabel('Sepal Length (cm)')
15 plt.ylabel('Sepal Width (cm)')
16 plt.colorbar(label='Species')
17 plt.show()
18
19 # 2. Histogram (Distribution of Sepal Length)
20 plt.figure(figsize=(6, 4))
21 plt.hist(iris_data['sepal length (cm)'], bins=10, color='skyblue', edgecolor='black')
22 plt.title('Distribution of Sepal Length')
23 plt.xlabel('Sepal Length (cm)')
24 plt.ylabel('Frequency')
25 plt.show()
26 # Conclusion: Sepal lengths are distributed with most values concentrated between 5 and 6.5 cm.
27
28 # 3. Box Plot (Sepal Width by Species)
29 plt.figure(figsize=(6, 4))
30 species = ['Setosa', 'Versicolor', 'Virginica']
31 data_by_species = [iris_data[iris_data['target'] == i]['sepal width (cm)'] for i in range(3)]
32 plt.boxplot(data_by_species, labels=species)
33 plt.title('Sepal Width Distribution by Species')
34 plt.ylabel('Sepal Width (cm)')
35 plt.show()
36 # Conclusion: Setosa has the widest Sepal width range, while Virginica's is more compact.
37
38 # 4. Line Plot (Cumulative Sepal Length)
39 plt.figure(figsize=(6, 4))
40 plt.plot(iris_data['sepal length (cm)'].cumsum(), label='Cumulative Sepal Length', color='red')
41 plt.title('Cumulative Sepal Length')
42 plt.xlabel('Index')
43 plt.ylabel('Cumulative Sepal Length (cm)')
44 plt.legend()
45 plt.show()
46 # Conclusion: The cumulative sepal length shows a smooth upward trend as data is added.
47
48 # 5. Pie Chart (Species Distribution)
49 plt.figure(figsize=(6, 6))
50 species_counts = iris_data['target'].value_counts()
51 plt.pie(species_counts, labels=species, autopct='%1.1f%%', startangle=140, colors=['gold', 'lightgreen', 'skyblue'])
52 plt.title('Species Distribution')
53 plt.show()
54 # Conclusion: The dataset is evenly distributed among the three species.
55
56 # ----- Seaborn Plots -----
57 # 1. Pairplot (Relationships among features)
58 sns.pairplot(iris_data, hue='target', palette='deep')
59 plt.suptitle('Pairplot of Iris Features', y=1.02)
60 plt.show()
61 # Analysis: Pairwise relationships show clear separations for Setosa, while Versicolor and Virginica overlap.
62
63 # 2. Heatmap (Correlation Matrix)
64 correlation_matrix = iris_data.corr()
65 sns.heatmap(correlation_matrix, annot=True, cmap='coolwarm', fmt='.2f')
66 plt.title('Heatmap of Feature Correlations')
67 plt.show()
68 # Analysis: Petal length and width are strongly correlated, while Sepal features have weaker correlations.
69
70 # 3. Violin Plot (Petal Length by Species)
71 sns.violinplot(x=iris_data['target'], y=iris_data['petal length (cm)'], palette='muted')
72 plt.title('Petal Length Distribution by Species')
73 plt.xlabel('Species')
74 plt.ylabel('Petal Length (cm)')
75 plt.xticks([0, 1, 2], species)
76 plt.show()

```

```

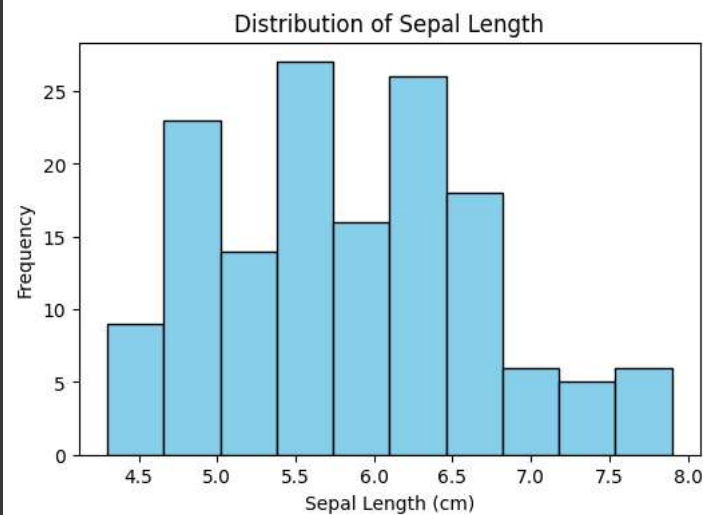
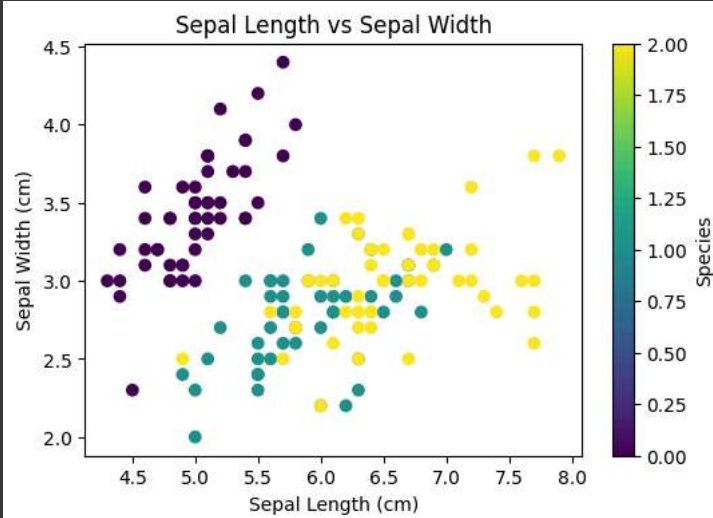
77 # Analysis: Setosa's petal length is distinctly smaller compared to the other species.
78
79 # 4. Swarm Plot (Sepal Length by Species)
80 sns.swarmplot(x='target', y='sepal length (cm)', data=iris_data, palette='pastel')
81 plt.title('Sepal Length by Species')
82 plt.xlabel('Species')
83 plt.ylabel('Sepal Length (cm)')
84 plt.xticks([0, 1, 2], species)
85 plt.show()
86 # Analysis: Virginica shows the longest Sepal lengths, while Setosa shows the shortest.
87
88 # 5. Count Plot (Species Count)
89 sns.countplot(x='target', data=iris_data, palette='bright')
90 plt.title('Count of Each Species')
91 plt.xlabel('Species')
92 plt.ylabel('Count')
93 plt.xticks([0, 1, 2], species)
94 plt.show()
95 # Analysis: All species are equally represented in the dataset (50 samples each).
96 import numpy as np
97 import matplotlib.pyplot as plt
98 from mpl_toolkits.mplot3d import Axes3D
99
100 # Load data (Iris dataset)
101 from sklearn.datasets import load_iris
102 iris = load_iris(as_frame=True)
103 iris_data = iris.frame
104
105 # 1. **3D Scatter Plot**: Sepal Length, Sepal Width, and Petal Length
106 fig = plt.figure(figsize=(8, 6))
107 ax = fig.add_subplot(111, projection='3d')
108
109 scatter = ax.scatter(
110     iris_data['sepal length (cm)'],
111     iris_data['sepal width (cm)'],
112     iris_data['petal length (cm)'],
113     c=iris_data['target'], cmap='viridis', s=50
114 )
115 ax.set_xlabel('Sepal Length (cm)')
116 ax.set_ylabel('Sepal Width (cm)')
117 ax.set_zlabel('Petal Length (cm)')
118 ax.set_title('3D Scatter Plot of Iris Features')
119 plt.colorbar(scatter, ax=ax, label='Species')
120 plt.show()
121 # Conclusion: Setosa is distinctly separate in 3D space, while Versicolor and Virginica show slight overlap.
122
123 # 2. **Subplots with Shared Axes**: Comparing Petal and Sepal Distributions
124 fig, axs = plt.subplots(2, 2, figsize=(10, 8), sharex=True, sharey=False)
125 fig.suptitle('Feature Comparisons', fontsize=16)
126
127 # Petal Length
128 axs[0, 0].hist(iris_data['petal length (cm)'], bins=15, color='skyblue', edgecolor='black')
129 axs[0, 0].set_title('Petal Length')
130 axs[0, 0].set_ylabel('Frequency')
131
132 # Petal Width
133 axs[0, 1].hist(iris_data['petal width (cm)'], bins=15, color='lightcoral', edgecolor='black')
134 axs[0, 1].set_title('Petal Width')
135
136 # Sepal Length
137 axs[1, 0].hist(iris_data['sepal length (cm)'], bins=15, color='lightgreen', edgecolor='black')
138 axs[1, 0].set_title('Sepal Length')
139 axs[1, 0].set_ylabel('Frequency')
140 axs[1, 0].set_xlabel('Measurement (cm)')
141
142 # Sepal Width
143 axs[1, 1].hist(iris_data['sepal width (cm)'], bins=15, color='gold', edgecolor='black')
144 axs[1, 1].set_title('Sepal Width')
145 axs[1, 1].set_xlabel('Measurement (cm)')
146
147 plt.tight_layout(rect=[0, 0, 1, 0.96])
148 plt.show()
149 # Conclusion: Petal features have more variation compared to Sepal features.
150
151 # 3. **Advanced Line Plot with Annotations**: Cumulative Sum of Petal Length
152 cumsum = iris_data['petal length (cm)'].cumsum()
153 fig, ax = plt.subplots(figsize=(8, 4))

```

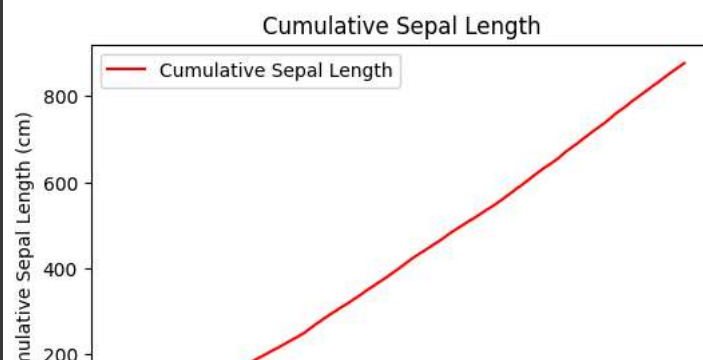
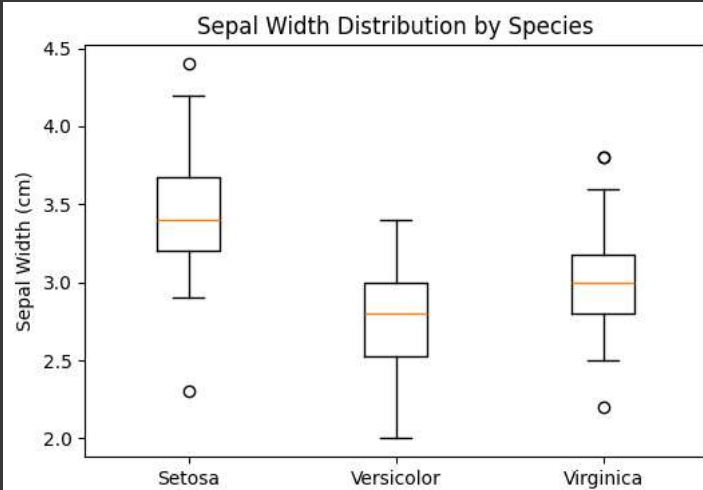
```

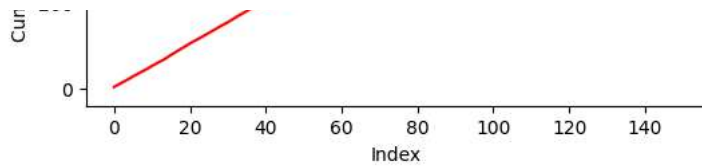
154 ax.plot(cumsum, label='Cumulative Petal Length', color='purple', lw=2)
155 ax.set_title('Cumulative Sum of Petal Length', fontsize=14)
156 ax.set_xlabel('Index')
157 ax.set_ylabel('Cumulative Petal Length (cm)')
158
159 # Annotating specific points
160 for idx in [30, 75, 120]:
161     ax.annotate(f'{cumsum[idx]:.1f}', xy=(idx, cumsum[idx]), xytext=(idx + 5, cumsum[idx] + 5),
162               arrowprops=dict(facecolor='black', arrowstyle='->'))
163 plt.legend()
164 plt.grid()
165 plt.show()
166 # Conclusion: The cumulative increase is linear due to evenly distributed values across species.
167
168 # 4. **Filled Area Plot**: Sepal Length by Species
169 species_colors = ['gold', 'lightgreen', 'lightcoral']
170 fig, ax = plt.subplots(figsize=(10, 5))
171 for i in range(3):
172     data = iris_data[iris_data['target'] == i]['sepal length (cm)']
173     ax.fill_between(range(len(data)), data, alpha=0.5, label=iris.target_names[i], color=species_colors[i])
174
175 ax.set_title('Sepal Length Distribution by Species')
176 ax.set_xlabel('Sample Index')
177 ax.set_ylabel('Sepal Length (cm)')
178 ax.legend()
179 plt.show()
180 # Conclusion: Virginica generally has the longest sepal lengths, while Setosa's values are consistently smaller.
181
182 # 5. **Polar Plot**: Proportions of Sepal vs Petal Dimensions
183 averages = [
184     iris_data['sepal length (cm)'].mean(),
185     iris_data['sepal width (cm)'].mean(),
186     iris_data['petal length (cm)'].mean(),
187     iris_data['petal width (cm)'].mean()
188 ]
189
190 angles = np.linspace(0, 2 * np.pi, len(averages), endpoint=False).tolist()
191 angles += angles[:1] # Close the loop
192
193 fig, ax = plt.subplots(figsize=(6, 6), subplot_kw=dict(polar=True))
194 averages += averages[:1] # Close the loop
195 ax.plot(angles, averages, color='blue', linewidth=2, label='Average Values')
196 ax.fill(angles, averages, color='blue', alpha=0.25)
197
198 ax.set_yticks([1, 2, 3, 4])
199 ax.set_xticks(angles[:-1])
200 ax.set_xticklabels(['Sepal Length', 'Sepal Width', 'Petal Length', 'Petal Width'])
201 ax.set_title('Radar Chart of Average Dimensions')
202 ax.legend(loc='upper right')
203 plt.show()
204 # Conclusion: Petal dimensions show higher variability, while Sepal dimensions are relatively balanced.
205

```

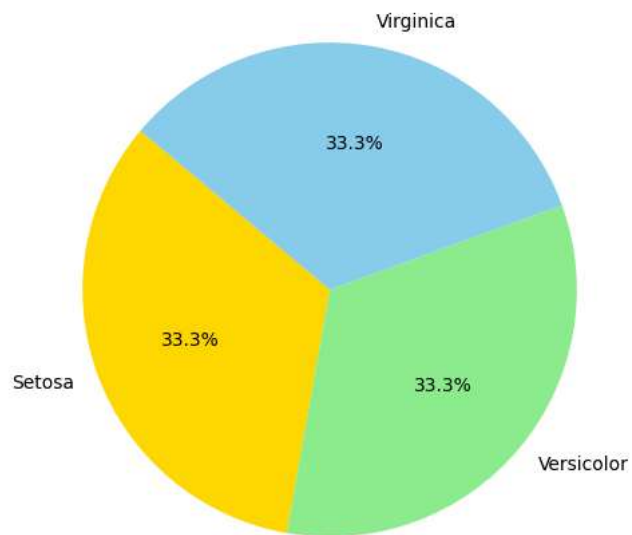


```
<ipython-input-3-74c1d8543e06>:32: MatplotlibDeprecationWarning: The 'labels' parameter of boxplot() has been renamed 'tick_labels' s
plt.boxplot(data_by_species, labels=species)
```

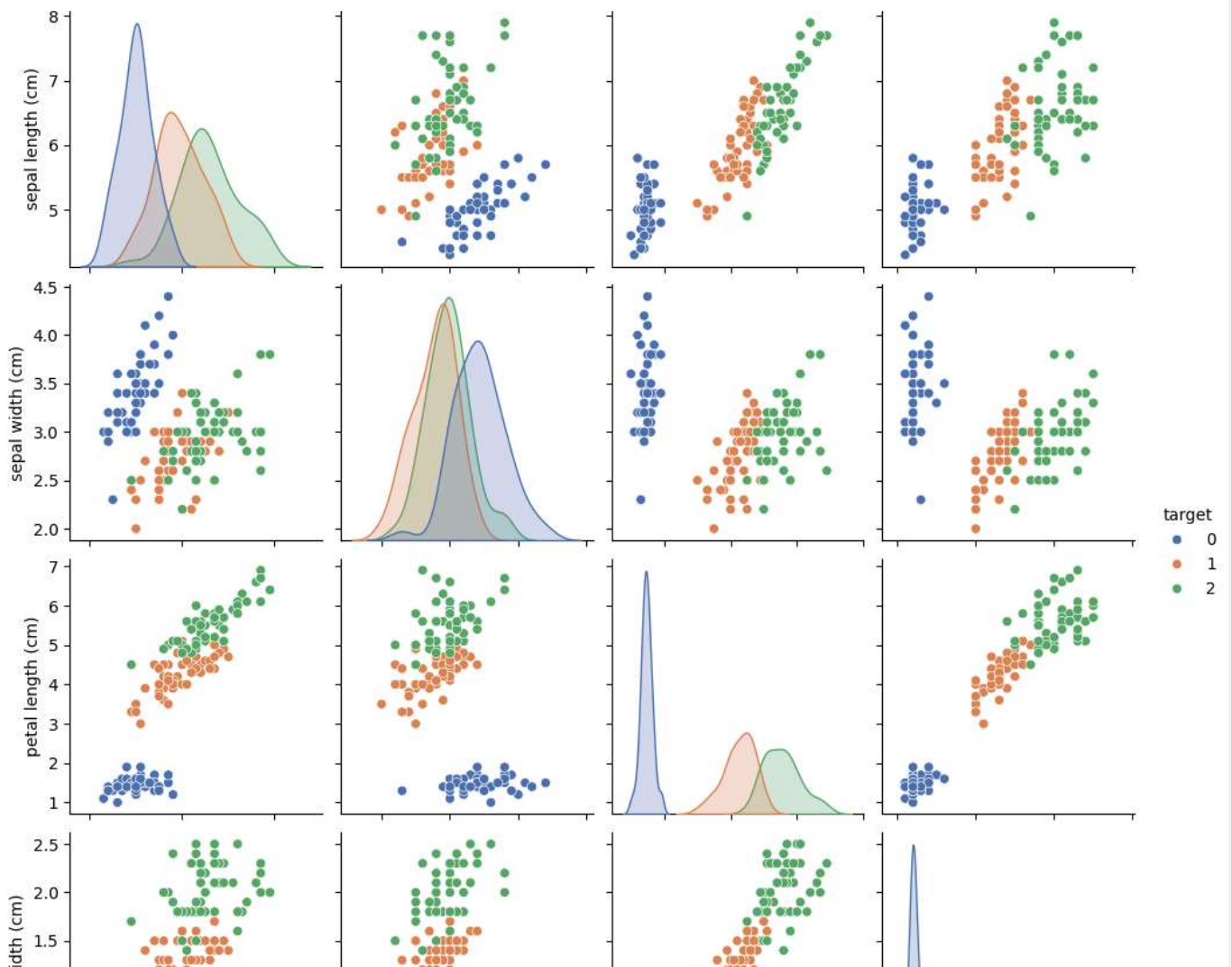


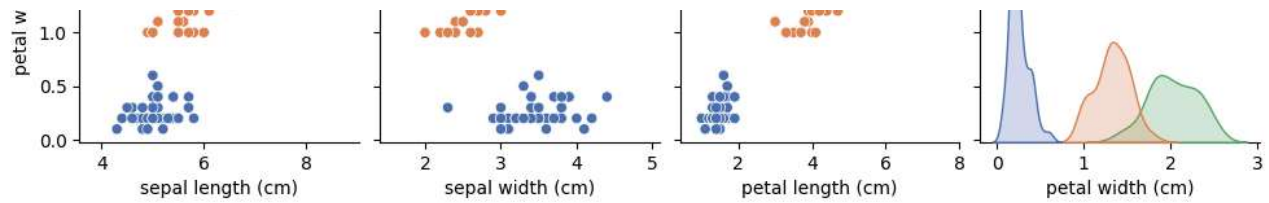


Species Distribution

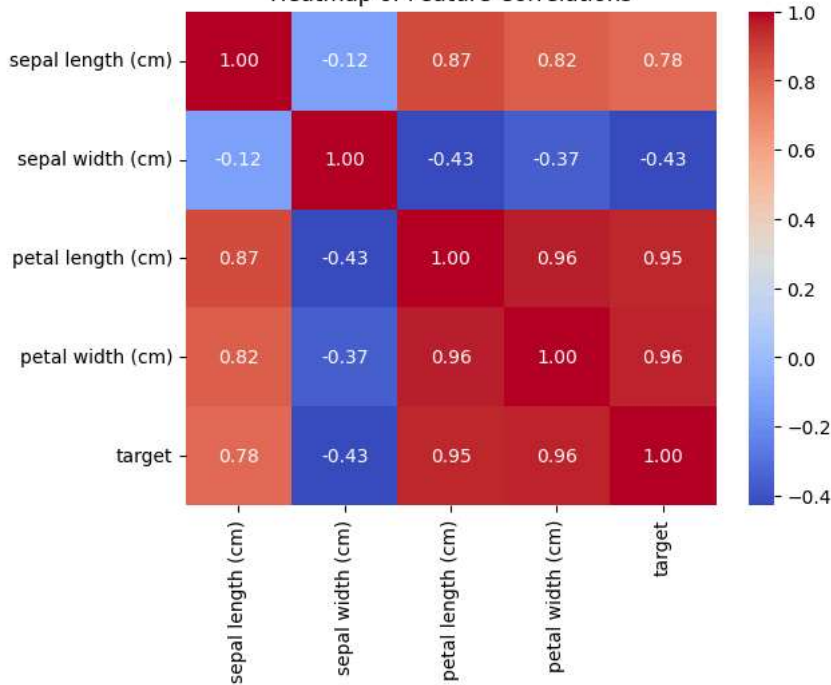


Pairplot of Iris Features





Heatmap of Feature Correlations

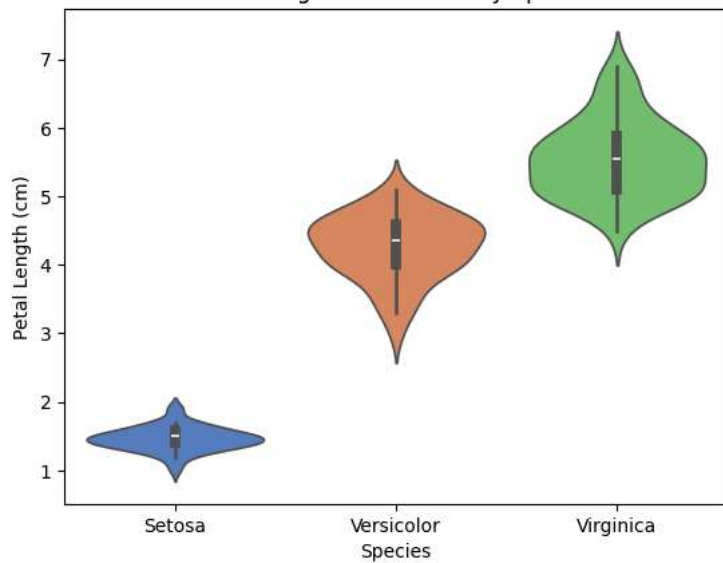


```
<ipython-input-3-74c1d8543e06>:71: FutureWarning:
```

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')
```

Petal Length Distribution by Species



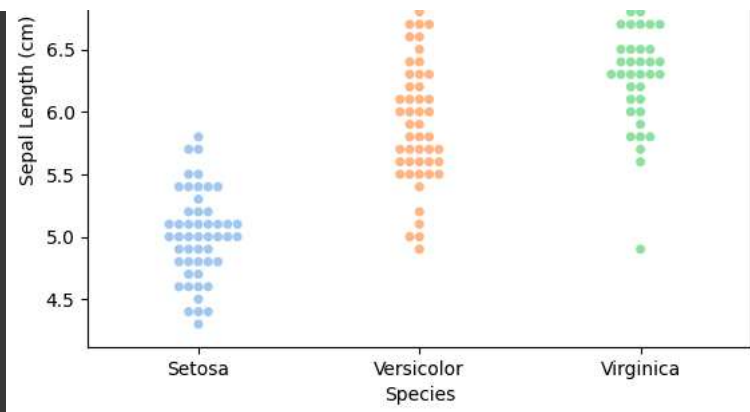
```
<ipython-input-3-74c1d8543e06>:80: FutureWarning:
```

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')
```

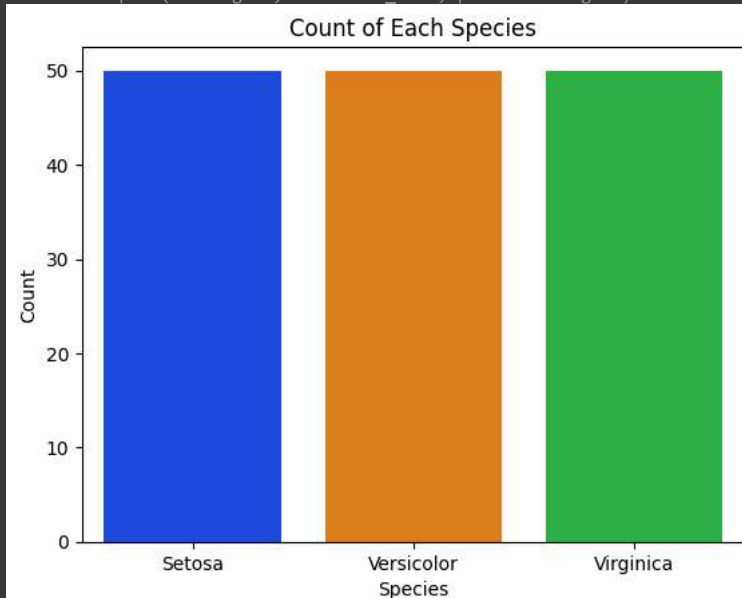
Sepal Length by Species



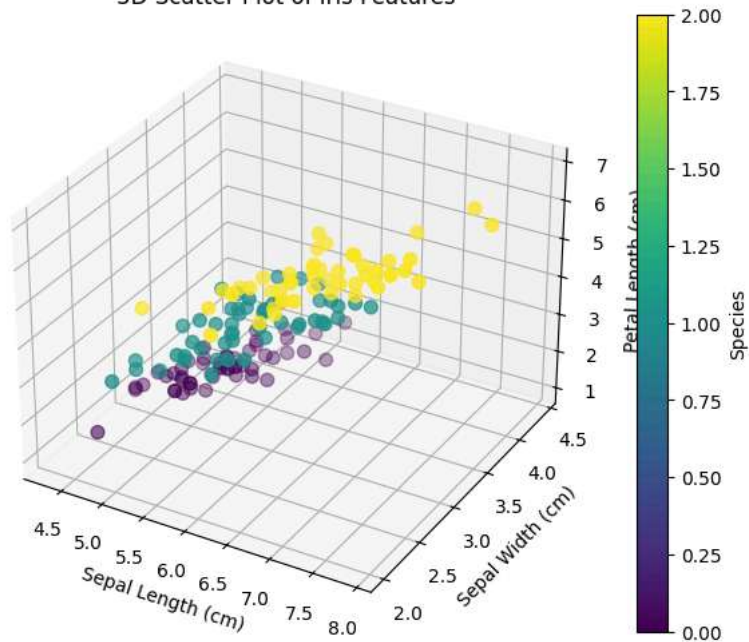


```
<ipython-input-3-74c1d8543e06>:89: FutureWarning:
```

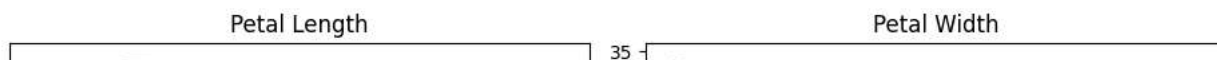
```
Passing `palette` without assigning `hue` is deprecated and will be removed in v0.14.0. Assign the `x` variable to `hue` and set `leg  
sns.countplot(x='target', data=iris_data, palette='bright')
```

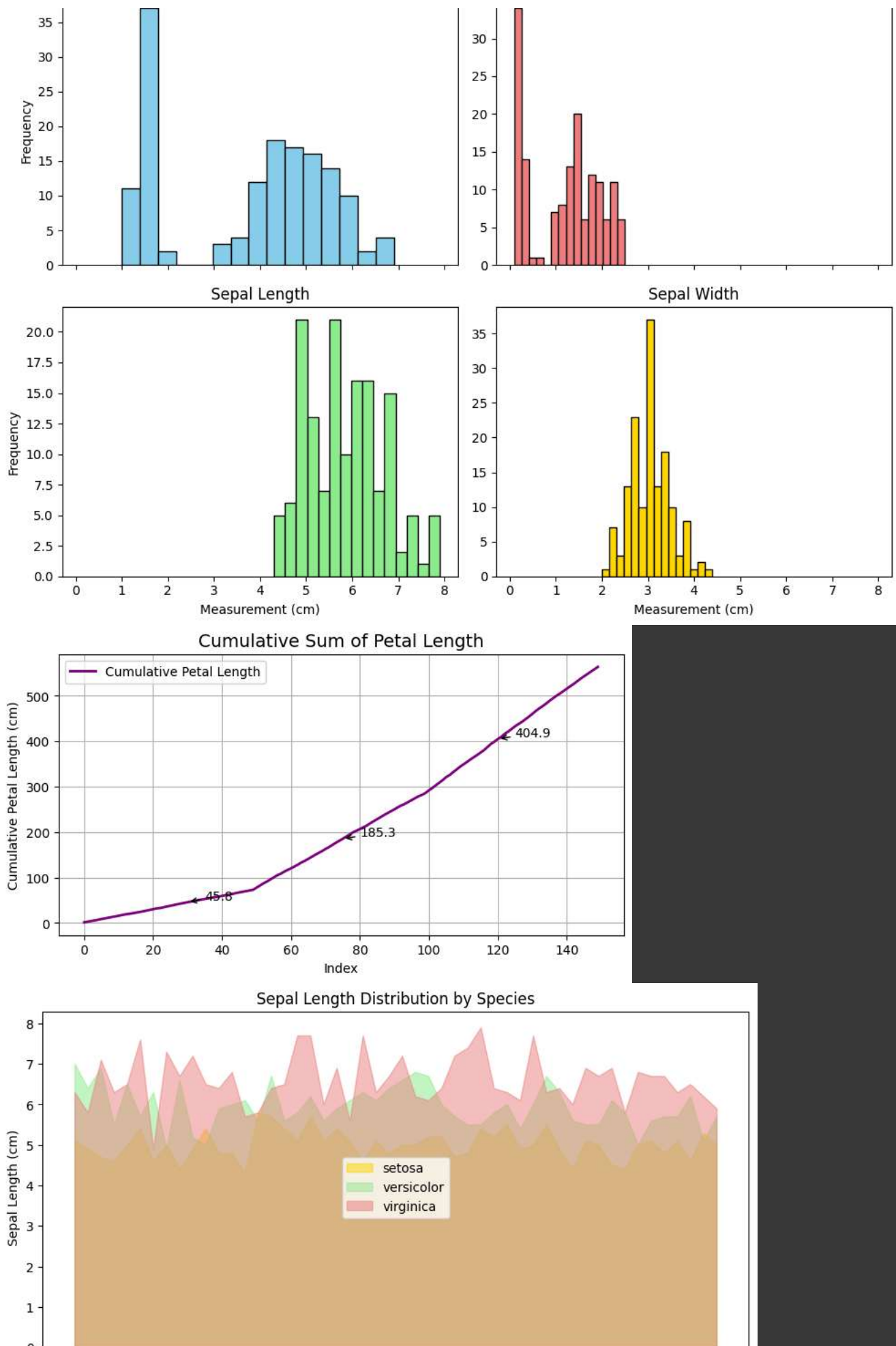


3D Scatter Plot of Iris Features

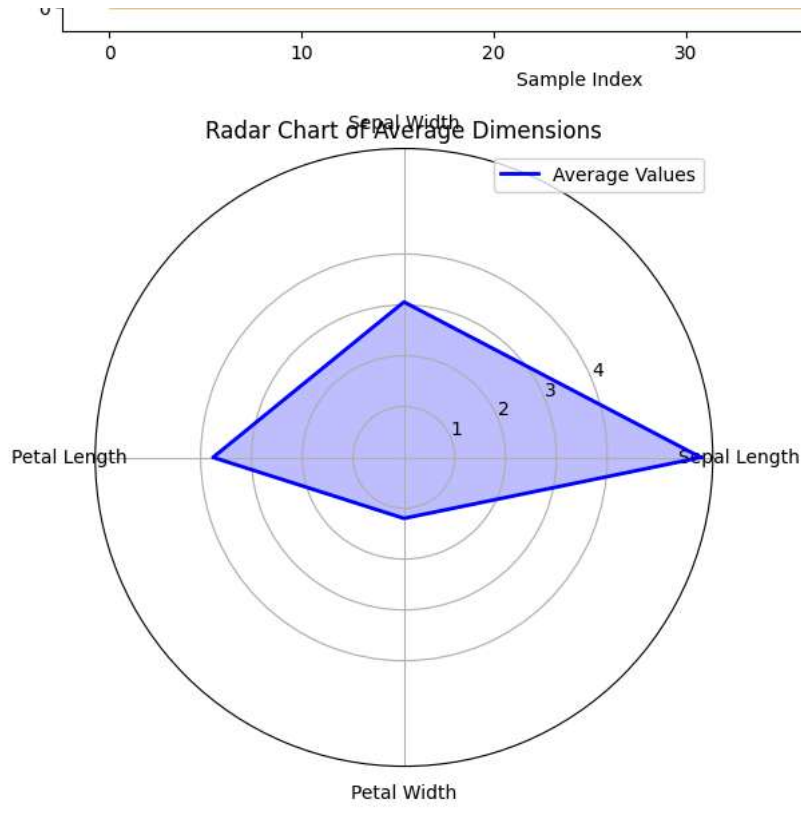


Feature Comparisons









```

1 # -*- coding: utf-8 -*-
2 """Untitled17.ipynb
3
4 Automatically generated by Colab.
5
6 Original file is located at
7 https://colab.research.google.com/drive/1c4w-nRN7Kr-C1VBDaimwgvG2Wo2GnMKw
8 """
9
10 import matplotlib.pyplot as plt
11 import seaborn as sns
12 import pandas as pd
13 from sklearn.datasets import load_iris
14
15 # Load a sample dataset (Iris dataset)
16 iris = load_iris(as_frame=True)
17 iris_data = iris.frame
18
19 # ----- Matplotlib Plots -----
20 # 1. Scatter Plot (Sepal Length vs Sepal Width)
21 plt.figure(figsize=(6, 4))
22 plt.scatter(iris_data['sepal length (cm)'], iris_data['sepal width (cm)'], c=iris_data['target'], cmap='viridis')
23 plt.title('Sepal Length vs Sepal Width')
24 plt.xlabel('Sepal Length (cm)')
25 plt.ylabel('Sepal Width (cm)')
26 plt.colorbar(label='Species')
27 plt.show()
28 # Conclusion: Different species have slight overlaps in Sepal dimensions but show clusters.
29
30 # 2. Histogram (Distribution of Sepal Length)
31 plt.figure(figsize=(6, 4))
32 plt.hist(iris_data['sepal length (cm)'], bins=10, color='skyblue', edgecolor='black')
33 plt.title('Distribution of Sepal Length')
34 plt.xlabel('Sepal Length (cm)')
35 plt.ylabel('Frequency')
36 plt.show()
37 # Conclusion: Sepal lengths are distributed with most values concentrated between 5 and 6.5 cm.
38
39 # 3. Box Plot (Sepal Width by Species)
40 plt.figure(figsize=(6, 4))
41 species = ['Setosa', 'Versicolor', 'Virginica']
42 data_by_species = [iris_data[iris_data['target'] == i]['sepal width (cm)'] for i in range(3)]
43 plt.boxplot(data_by_species, labels=species)
44 plt.title('Sepal Width Distribution by Species')
45 plt.ylabel('Sepal Width (cm)')
46 plt.show()
47 # Conclusion: Setosa has the widest Sepal width range, while Virginica's is more compact.
48
49 # 4. Line Plot (Cumulative Sepal Length)
50 plt.figure(figsize=(6, 4))
51 plt.plot(iris_data['sepal length (cm)'].cumsum(), label='Cumulative Sepal Length', color='red')
52 plt.title('Cumulative Sepal Length')
53 plt.xlabel('Index')
54 plt.ylabel('Cumulative Sepal Length (cm)')
55 plt.legend()
56 plt.show()
57 # Conclusion: The cumulative sepal length shows a smooth upward trend as data is added.
58
59 # 5. Pie Chart (Species Distribution)
60 plt.figure(figsize=(6, 6))
61 species_counts = iris_data['target'].value_counts()
62 plt.pie(species_counts, labels=species, autopct='%1.1f%%', startangle=140, colors=['gold', 'lightgreen', 'skyblue'])
63 plt.title('Species Distribution')
64 plt.show()
65 # Conclusion: The dataset is evenly distributed among the three species.
66
67 # ----- Seaborn Plots -----
68 # 1. Pairplot (Relationships among features)
69 sns.pairplot(iris_data, hue='target', palette='deep')
70 plt.suptitle('Pairplot of Iris Features', y=1.02)
71 plt.show()
72 # Analysis: Pairwise relationships show clear separations for Setosa, while Versicolor and Virginica overlap.
73
74 # 2. Heatmap (Correlation Matrix)
75 correlation_matrix = iris_data.corr()
76 sns.heatmap(correlation_matrix, annot=True, cmap='coolwarm', fmt='.2f')
77 plt.title('Heatmap of Feature Correlations')

```

```

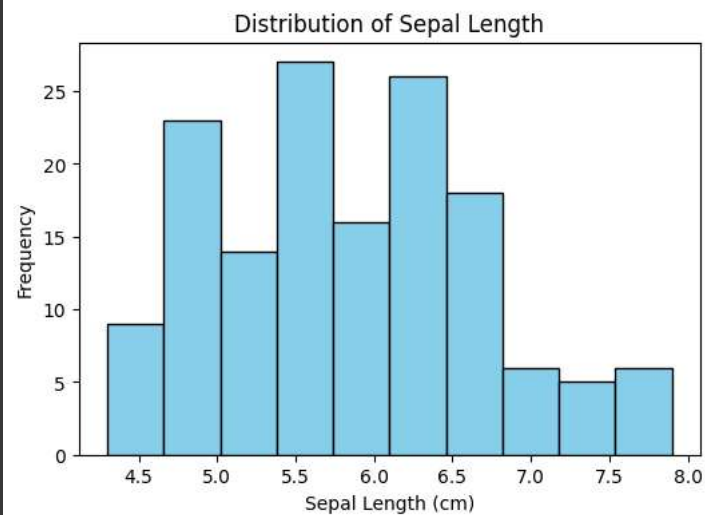
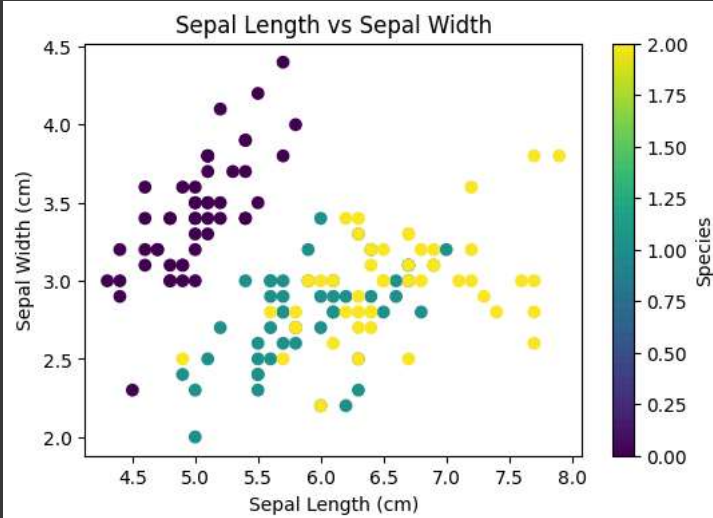
78 plt.show()
79 # Analysis: Petal length and width are strongly correlated, while Sepal features have weaker correlations.
80
81 # 3. Violin Plot (Petal Length by Species)
82 sns.violinplot(x=iris_data['target'], y=iris_data['petal length (cm)'], palette='muted')
83 plt.title('Petal Length Distribution by Species')
84 plt.xlabel('Species')
85 plt.ylabel('Petal Length (cm)')
86 plt.xticks([0, 1, 2], species)
87 plt.show()
88 # Analysis: Setosa's petal length is distinctly smaller compared to the other species.
89
90 # 4. Swarm Plot (Sepal Length by Species)
91 sns.swarmplot(x='target', y='sepal length (cm)', data=iris_data, palette='pastel')
92 plt.title('Sepal Length by Species')
93 plt.xlabel('Species')
94 plt.ylabel('Sepal Length (cm)')
95 plt.xticks([0, 1, 2], species)
96 plt.show()
97 # Analysis: Virginica shows the longest Sepal lengths, while Setosa shows the shortest.
98
99 # 5. Count Plot (Species Count)
100 sns.countplot(x='target', data=iris_data, palette='bright')
101 plt.title('Count of Each Species')
102 plt.xlabel('Species')
103 plt.ylabel('Count')
104 plt.xticks([0, 1, 2], species)
105 plt.show()
106 # Analysis: All species are equally represented in the dataset (50 samples each).
107
108 import numpy as np
109 import matplotlib.pyplot as plt
110 from mpl_toolkits.mplot3d import Axes3D
111
112 # Load data (Iris dataset)
113 from sklearn.datasets import load_iris
114 iris = load_iris(as_frame=True)
115 iris_data = iris.frame
116
117 # 1. **3D Scatter Plot**: Sepal Length, Sepal Width, and Petal Length
118 fig = plt.figure(figsize=(8, 6))
119 ax = fig.add_subplot(111, projection='3d')
120
121 scatter = ax.scatter(
122     iris_data['sepal length (cm)'],
123     iris_data['sepal width (cm)'],
124     iris_data['petal length (cm)'],
125     c=iris_data['target'], cmap='viridis', s=50
126 )
127 ax.set_xlabel('Sepal Length (cm)')
128 ax.set_ylabel('Sepal Width (cm)')
129 ax.set_zlabel('Petal Length (cm)')
130 ax.set_title('3D Scatter Plot of Iris Features')
131 plt.colorbar(scatter, ax=ax, label='Species')
132 plt.show()
133 # Conclusion: Setosa is distinctly separate in 3D space, while Versicolor and Virginica show slight overlap.
134
135 # 2. **Subplots with Shared Axes**: Comparing Petal and Sepal Distributions
136 fig, axs = plt.subplots(2, 2, figsize=(10, 8), sharex=True, sharey=False)
137 fig.suptitle('Feature Comparisons', fontsize=16)
138
139 # Petal Length
140 axs[0, 0].hist(iris_data['petal length (cm)'], bins=15, color='skyblue', edgecolor='black')
141 axs[0, 0].set_title('Petal Length')
142 axs[0, 0].set_ylabel('Frequency')
143
144 # Petal Width
145 axs[0, 1].hist(iris_data['petal width (cm)'], bins=15, color='lightcoral', edgecolor='black')
146 axs[0, 1].set_title('Petal Width')
147
148 # Sepal Length
149 axs[1, 0].hist(iris_data['sepal length (cm)'], bins=15, color='lightgreen', edgecolor='black')
150 axs[1, 0].set_title('Sepal Length')
151 axs[1, 0].set_ylabel('Frequency')
152 axs[1, 0].set_xlabel('Measurement (cm)')
153
154 # Sepal Width

```

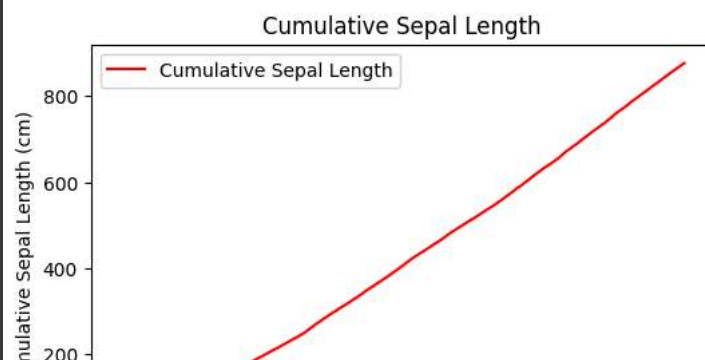
```

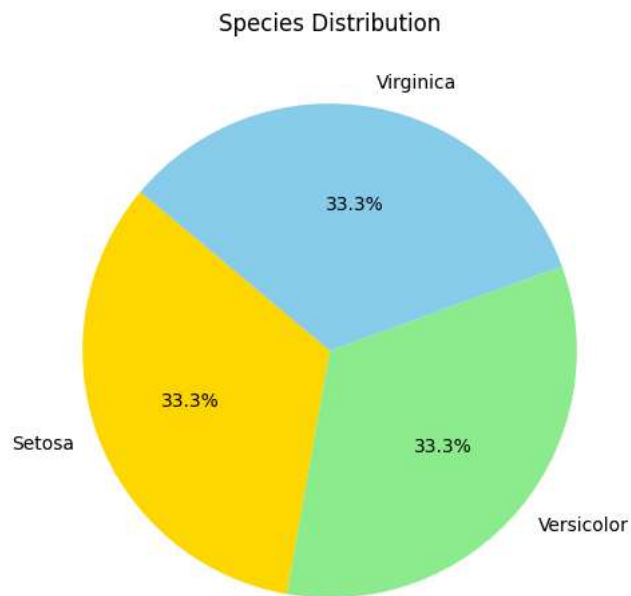
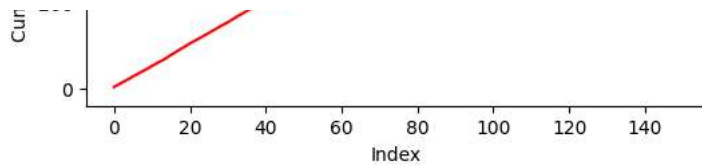
155 axs[1, 1].hist(iris_data['sepal width (cm)'], bins=15, color='gold', edgecolor='black')
156 axs[1, 1].set_title('Sepal Width')
157 axs[1, 1].set_xlabel('Measurement (cm)')
158
159 plt.tight_layout(rect=[0, 0, 1, 0.96])
160 plt.show()
161 # Conclusion: Petal features have more variation compared to Sepal features.
162
163 # 3. **Advanced Line Plot with Annotations**: Cumulative Sum of Petal Length
164 cumsum = iris_data['petal length (cm)'].cumsum()
165 fig, ax = plt.subplots(figsize=(8, 4))
166 ax.plot(cumsum, label='Cumulative Petal Length', color='purple', lw=2)
167 ax.set_title('Cumulative Sum of Petal Length', fontsize=14)
168 ax.set_xlabel('Index')
169 ax.set_ylabel('Cumulative Petal Length (cm)')
170
171 # Annotating specific points
172 for idx in [30, 75, 120]:
173     ax.annotate(f'{cumsum[idx]:.1f}', xy=(idx, cumsum[idx]), xytext=(idx + 5, cumsum[idx] + 5),
174               arrowprops=dict(facecolor='black', arrowstyle='->'))
175 plt.legend()
176 plt.grid()
177 plt.show()
178 # Conclusion: The cumulative increase is linear due to evenly distributed values across species.
179
180 # 4. **Filled Area Plot**: Sepal Length by Species
181 species_colors = ['gold', 'lightgreen', 'lightcoral']
182 fig, ax = plt.subplots(figsize=(10, 5))
183 for i in range(3):
184     data = iris_data[iris_data['target'] == i]['sepal length (cm)']
185     ax.fill_between(range(len(data)), data, alpha=0.5, label=iris.target_names[i], color=species_colors[i])
186
187 ax.set_title('Sepal Length Distribution by Species')
188 ax.set_xlabel('Sample Index')
189 ax.set_ylabel('Sepal Length (cm)')
190 ax.legend()
191 plt.show()
192 # Conclusion: Virginica generally has the longest sepal lengths, while Setosa's values are consistently smaller.
193
194 # 5. **Polar Plot**: Proportions of Sepal vs Petal Dimensions
195 averages = [
196     iris_data['sepal length (cm)'].mean(),
197     iris_data['sepal width (cm)'].mean(),
198     iris_data['petal length (cm)'].mean(),
199     iris_data['petal width (cm)'].mean()
200 ]
201
202 angles = np.linspace(0, 2 * np.pi, len(averages), endpoint=False).tolist()
203 angles += angles[:1] # Close the loop
204
205 fig, ax = plt.subplots(figsize=(6, 6), subplot_kw=dict(polar=True))
206 averages += averages[:1] # Close the loop
207 ax.plot(angles, averages, color='blue', linewidth=2, label='Average Values')
208 ax.fill(angles, averages, color='blue', alpha=0.25)
209
210 ax.set_yticks([1, 2, 3, 4])
211 ax.set_xticks(angles[:-1])
212 ax.set_xticklabels(['Sepal Length', 'Sepal Width', 'Petal Length', 'Petal Width'])
213 ax.set_title('Radar Chart of Average Dimensions')
214 ax.legend(loc='upper right')
215 plt.show()
216 # Conclusion: Petal dimensions show higher variability, while Sepal dimensions are relatively balanced.

```

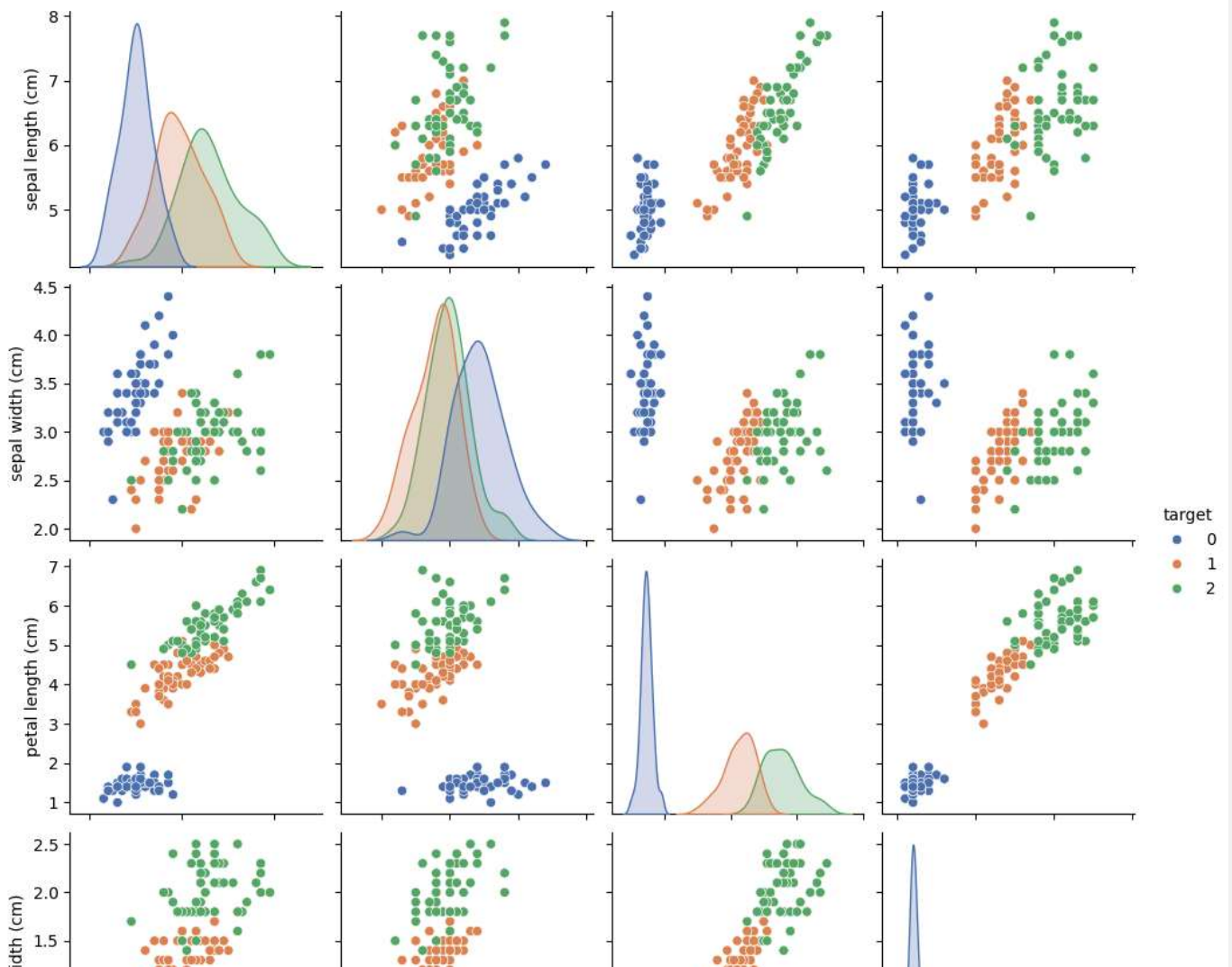


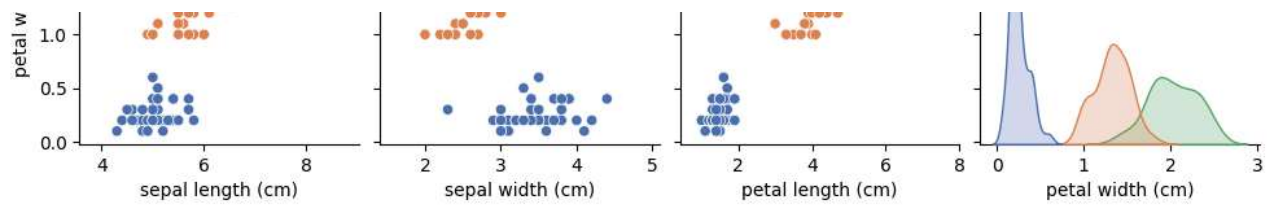
```
<ipython-input-4-27152afe7d4b>:43: MatplotlibDeprecationWarning: The 'labels' parameter of boxplot() has been renamed 'tick_labels' s
plt.boxplot(data_by_species, labels=species)
```



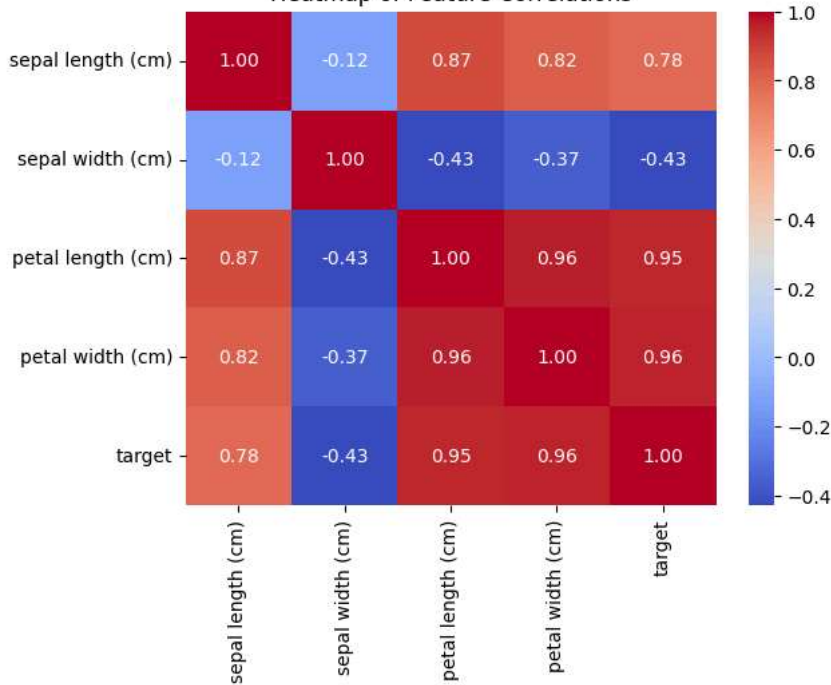


Pairplot of Iris Features





Heatmap of Feature Correlations

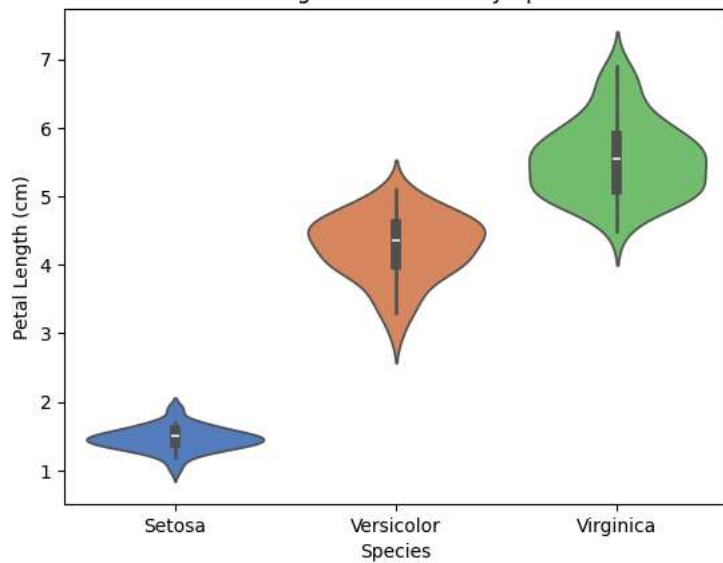


```
<ipython-input-4-27152afe7d4b>:82: FutureWarning:
```

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')
```

Petal Length Distribution by Species

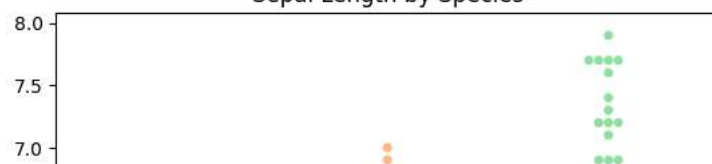


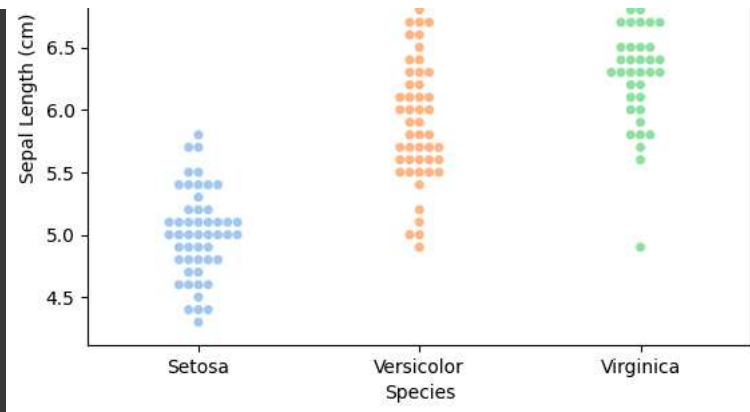
```
<ipython-input-4-27152afe7d4b>:91: FutureWarning:
```

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')
```

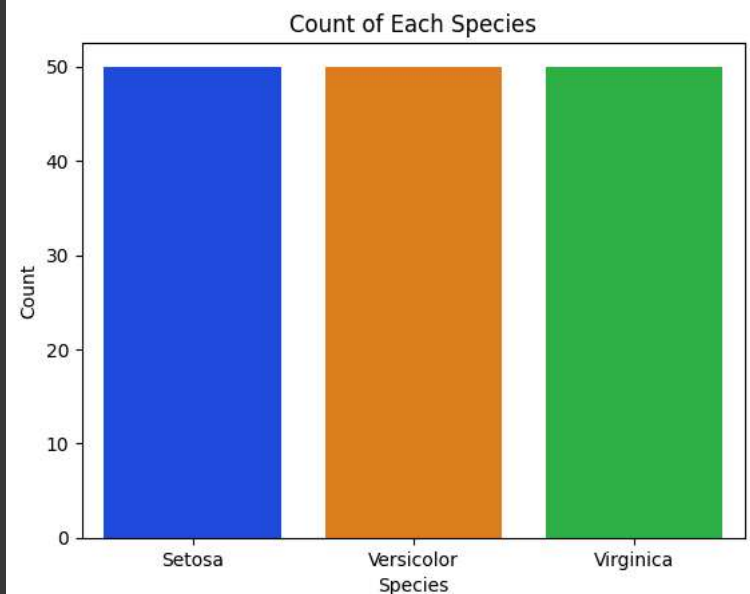
Sepal Length by Species



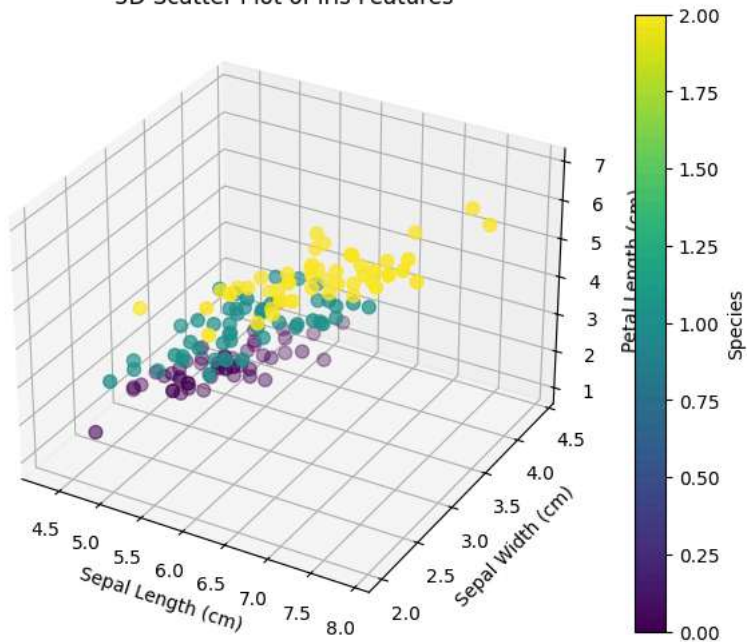


<ipython-input-4-27152afe7d4b>:100: FutureWarning:

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')



3D Scatter Plot of Iris Features



Feature Comparisons

