# SingleLink-CompleteLink

March 26, 2024

## 1 Import Libraries

[ ]:
```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
from scipy.cluster.hierarchy import dendrogram, linkage
from sklearn.datasets import load_iris
from sklearn.cluster import AgglomerativeClustering
import seaborn as sns
```

## 2 Load Dataset

[ ]:
```
# Load the iris dataset
iris = load_iris()
X = iris.data
```

[ ]:
```
df = pd.DataFrame(X)
df.corr()
```

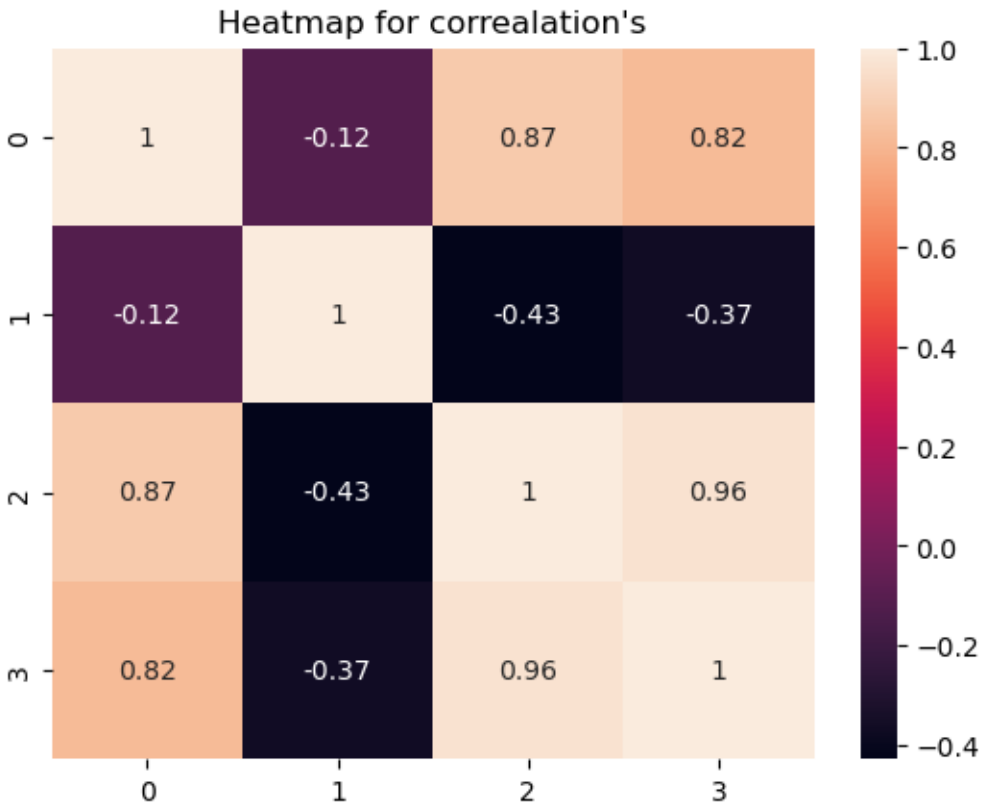
[ ]:

| | 0 | 1 | 2 | 3 |
|---|---|---|---|---|
| 0 | 1.000000 | -0.117570 | 0.871754 | 0.817941 |
| 1 | -0.117570 | 1.000000 | -0.428440 | -0.366126 |
| 2 | 0.871754 | -0.428440 | 1.000000 | 0.962865 |
| 3 | 0.817941 | -0.366126 | 0.962865 | 1.000000 |

## 3 Visualization's

- Heatmap

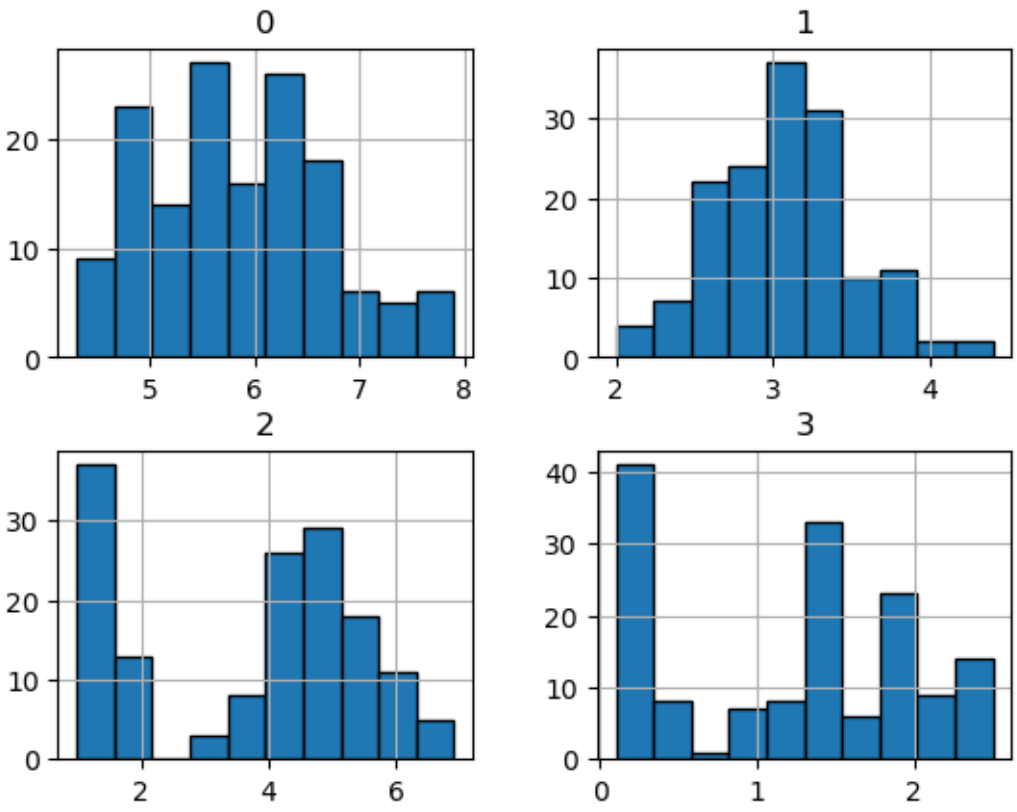
[ ]:
```
sns.heatmap(df.corr(),annot=True)
plt.title("Heatmap for correalation's")
plt.show()
```




- Histogram

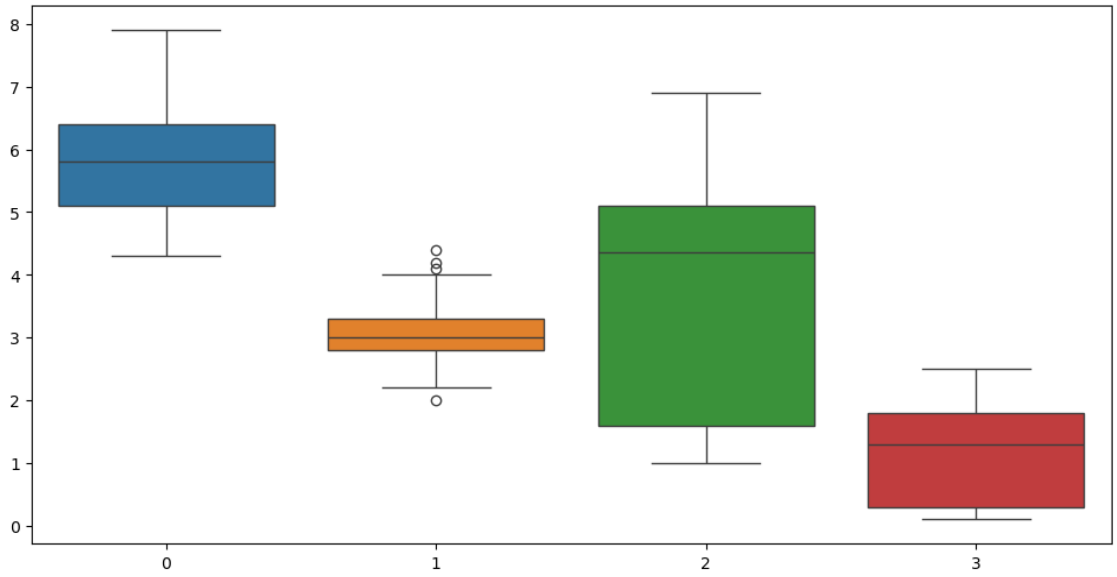
[ ]:
```
# Histogram
plt.figure(figsize=(12, 6))
df.hist(bins=10, edgecolor='black')
plt.show()
```

```
<Figure size 1200x600 with 0 Axes>
```



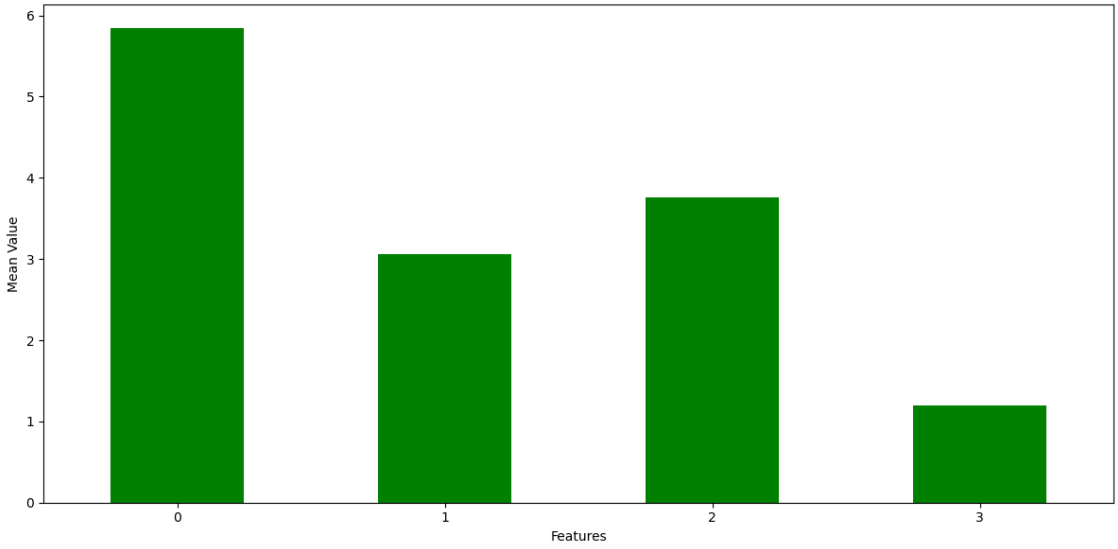

- BoxPlot

[ ]:
```
# Box Plot
plt.figure(figsize=(12, 6))
sns.boxplot(data=df)
plt.show()
```




- Bar Graph

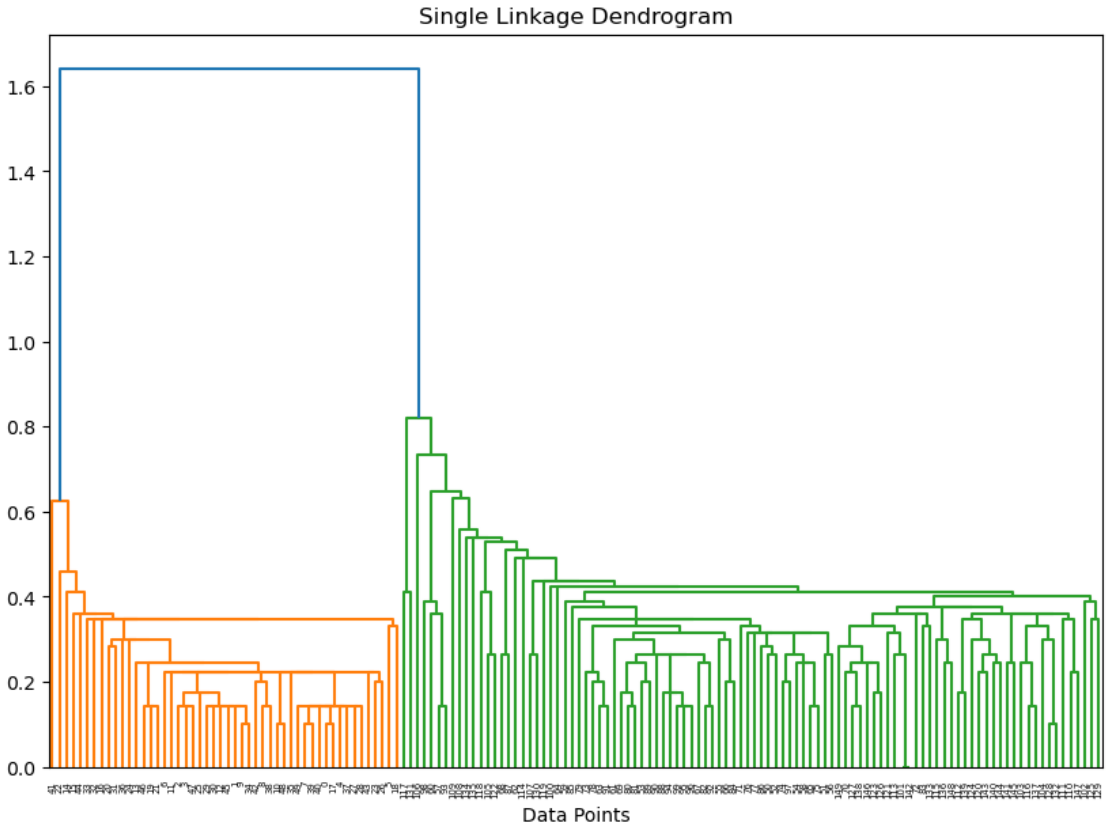
[ ]:
```
# Bar Graph
plt.figure(figsize=(12, 6))
df.mean().plot(kind='bar', rot=0, color='green')
plt.xlabel('Features')
plt.ylabel('Mean Value')
plt.tight_layout()
plt.show()
```



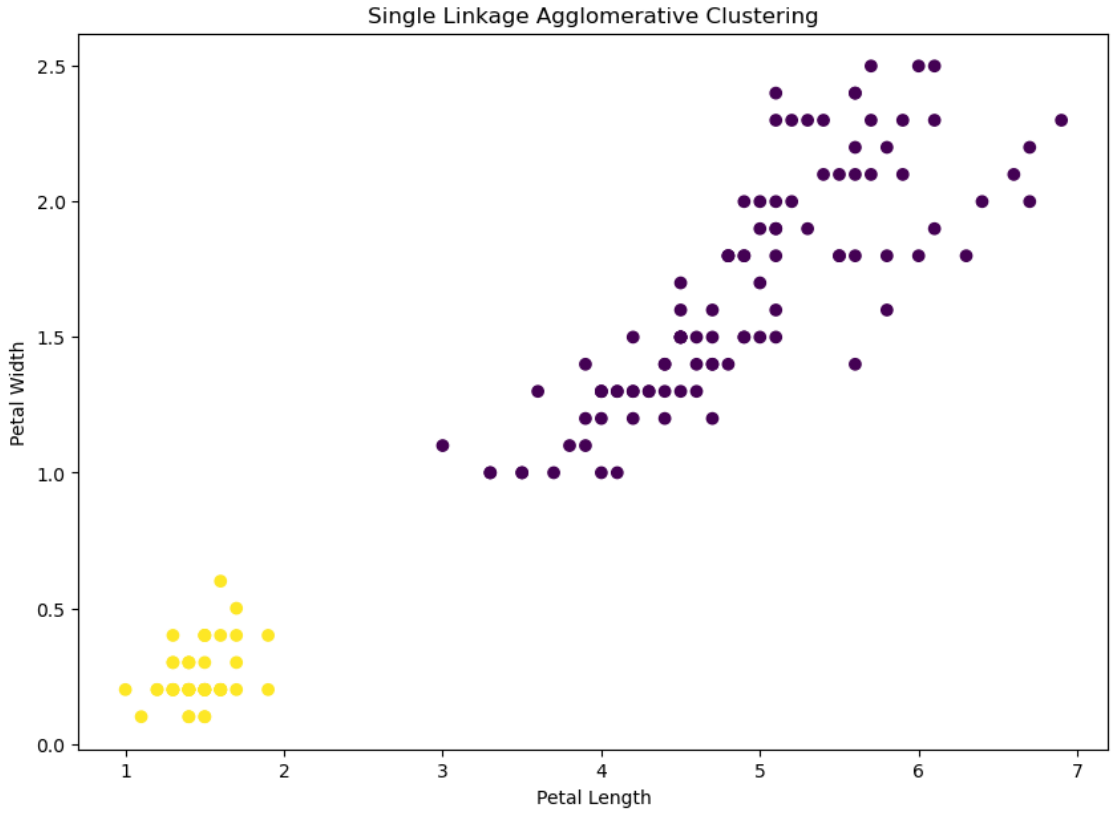
## 4 Clustering with Single Link

[ ]:
```
# Single Linkage Agglomerative Clustering
clustering_single = AgglomerativeClustering(linkage='single').fit(X)

# Plot the single linkage dendrogram
plt.figure(figsize=(10, 7))
dendrogram(linkage(X, method='single'))
plt.title('Single Linkage Dendrogram')
plt.xlabel('Data Points')
plt.show()
```



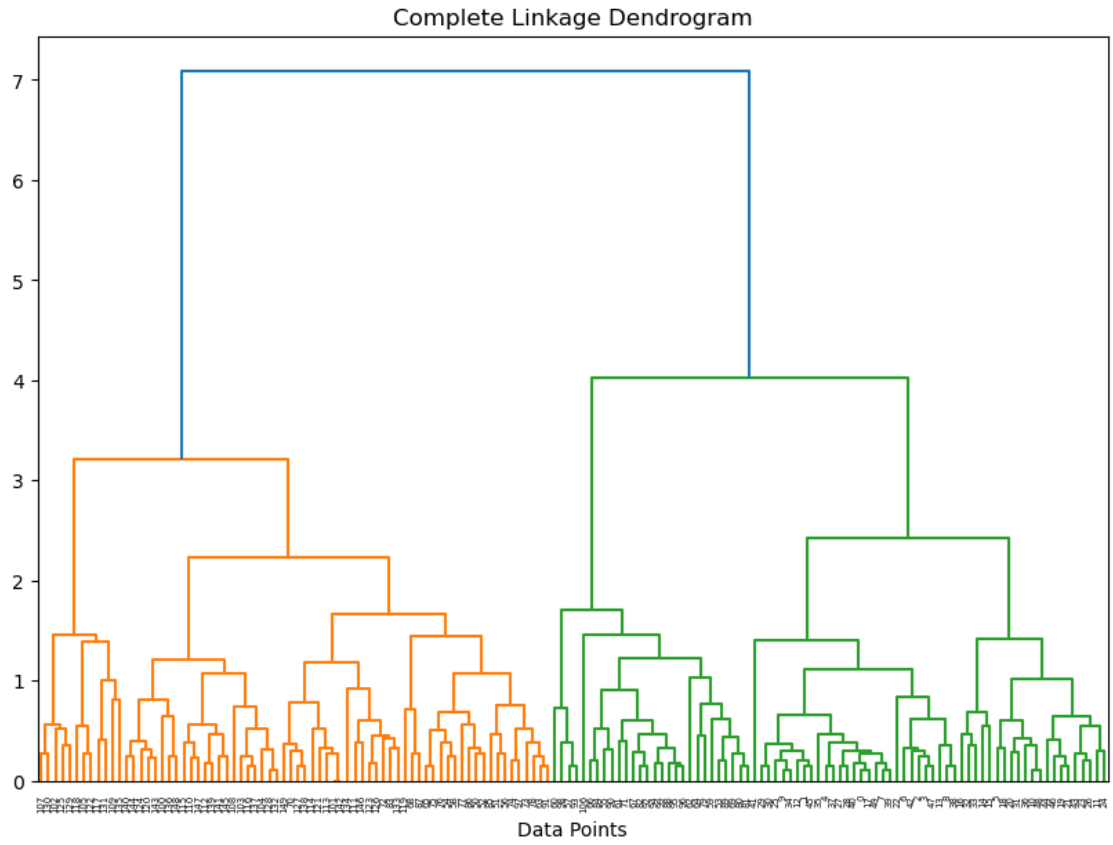

[ ]:
```
# Plot the single linkage clustering
plt.figure(figsize=(10, 7))
plt.scatter(X[:, 2], X[:, 3], c=clustering_single.labels_)
plt.title('Single Linkage Agglomerative Clustering')
plt.xlabel('Petal Length')
plt.ylabel('Petal Width')
plt.show()
```



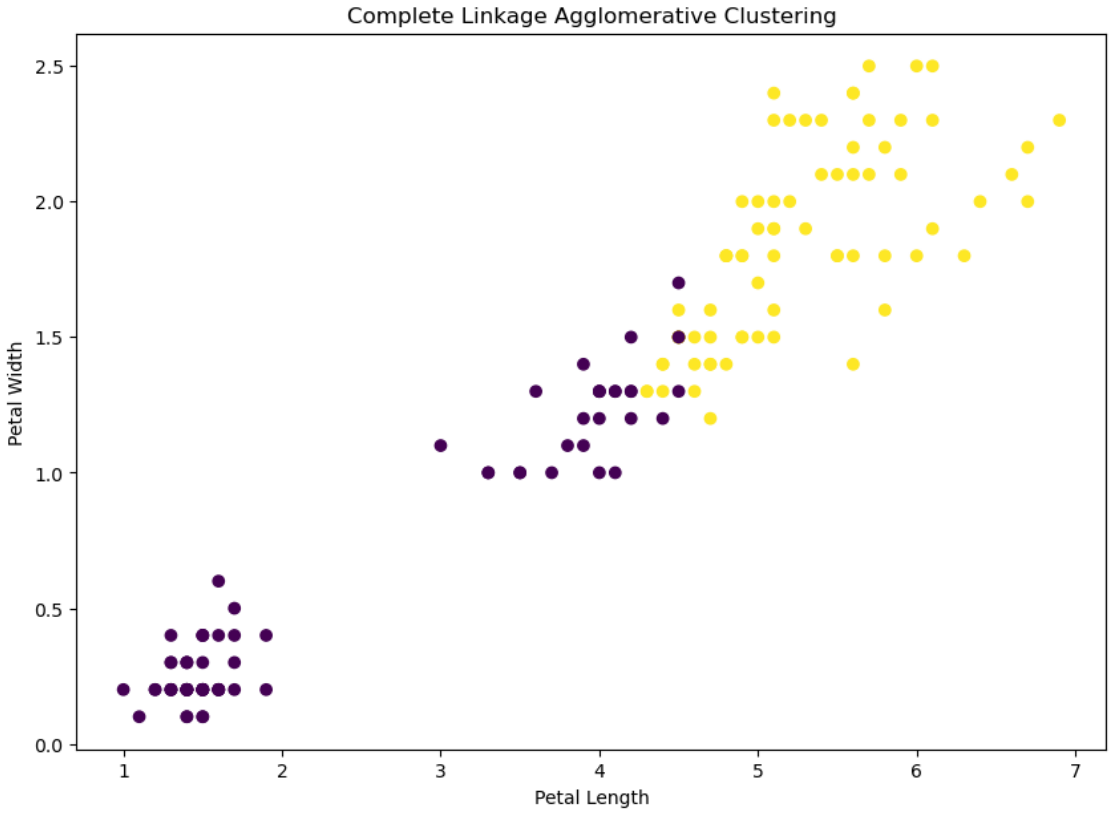

## 5 Clustering using Complete Link

[ ]:
```
# Complete Linkage Agglomerative Clustering
clustering_complete = AgglomerativeClustering(linkage='complete').fit(X)

# Plot the complete linkage dendrogram
plt.figure(figsize=(10, 7))
dendrogram(linkage(X, method='complete'))
plt.title('Complete Linkage Dendrogram')
plt.xlabel('Data Points')
plt.show()
```




[ ]:
```
# Plot the complete linkage clustering
plt.figure(figsize=(10, 7))
plt.scatter(X[:, 2], X[:, 3], c=clustering_complete.labels_)
plt.title('Complete Linkage Agglomerative Clustering')
plt.xlabel('Petal Length')
plt.ylabel('Petal Width')
plt.show()
```




**6 Conclusion: We can see that the data has been clustered differently with single link and complete link .And the cluster's are plotted using scatter plot and the cluster formation is done using dendogram , and there is clear seperation between cluster's showing the the two different specie's difference in Petal Length and Petal Width**