Sign in | Inbox (100) - revathyaramam@gm | WhatsApp | Home | Clustering_algorithm

localhost:8888/notebooks/Clustering_algorithm.ipynb

jupyter Clustering_algorithm Last Checkpoint: 1 minute ago

File Edit View Run Kernel Settings Help

Trusted

Code

JupyterLab

Python 3 (ipykernel)

[8]:
```
import pandas as pd

df = pd.read_excel("iris.xlsx")
df.head()
```

[8]:

	Id	SepalLengthCm	SepalWidthCm	PetalLengthCm	PetalWidthCm	Species
0	1	5.1	3.5	1.4	0.2	Iris-setosa
1	2	4.9	3.0	1.4	0.2	Iris-setosa
2	3	4.7	3.2	1.3	0.2	Iris-setosa
3	4	4.6	3.1	1.5	0.2	Iris-setosa
4	5	5.0	3.6	1.4	0.2	Iris-setosa

[9]:
```
X = df.drop(columns=['Species'])
```

[ ]:
```
###KMeans Clustering
KMeans is a centroid-based clustering algorithm that partitions data into K distinct clusters by minimizing
the variance within each cluster. It starts with random centroids, assigns points to the nearest one, and
updates centroids until convergence.
```

[10]:
```


from sklearn.cluster import KMeans
import seaborn as sns
import matplotlib.pyplot as plt

# KMeans clustering
kmeans = KMeans(n_clusters=3, random_state=42)
df['KMeans_Cluster'] = kmeans.fit_predict(X)

# Visualization
sns.pairplot(df, hue='KMeans_Cluster', palette='Set1')
```

28°C Mostly cloudy

Search

ENG IN

20:10 27-03-2025

Sign in | Inbox (100) - revathyaramam@gm | WhatsApp | Home | Clustering_algorithm

localhost:8888/notebooks/Clustering_algorithm.ipynb

jupyter Clustering_algorithm Last Checkpoint: 2 minutes ago

File Edit View Run Kernel Settings Help

Trusted

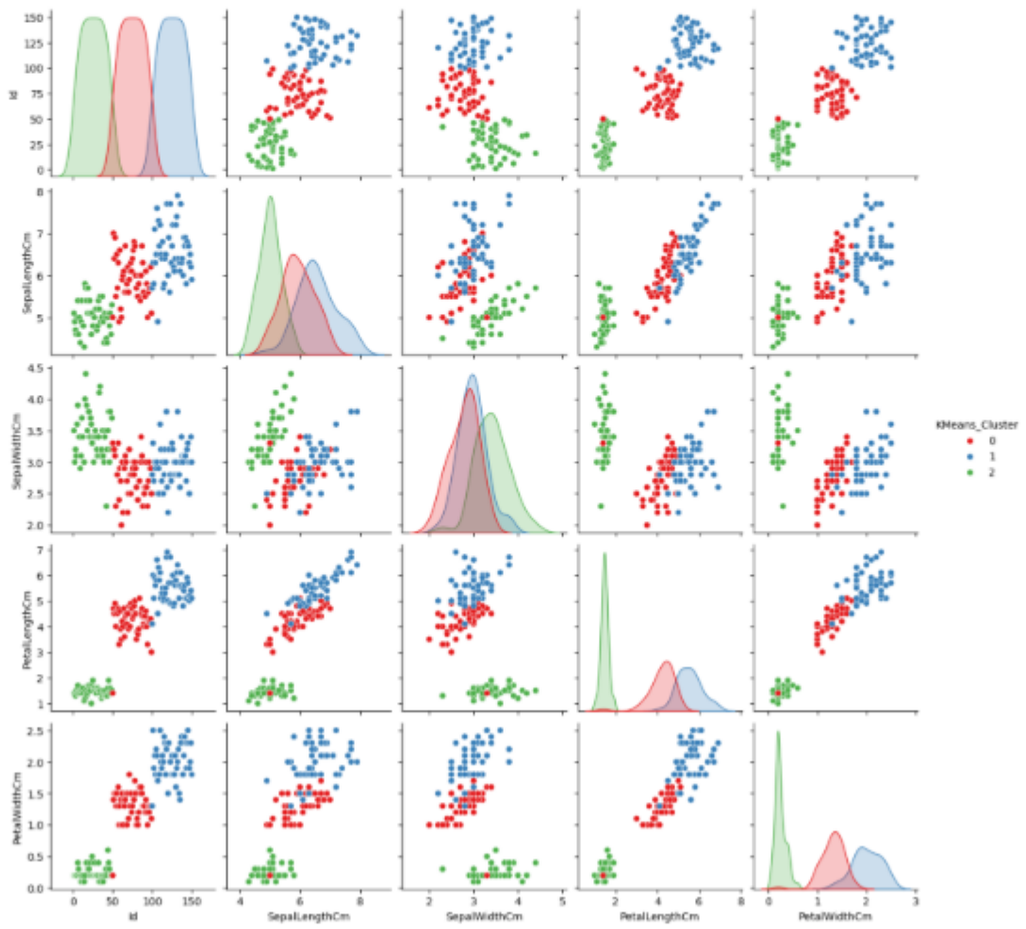
Code

JupyterLab | Python 3 (ipykernel)

```
# Visualization
sns.pairplot(df, hue='KMeans_Cluster', palette='Set1')
plt.suptitle("KMeans Clustering on Iris Dataset", y=1.02)
plt.show()
```

```
C:\Users\amrut\anaconda3\Lib\site-packages\sklearn\cluster\_kmeans.py:1429: UserWarning: KMeans is known to have a memory leak on Windows with MKL, when there are less chunks than available threads. You can avoid it by setting the environment variable OMP_NUM_THREADS=1.
  warnings.warn(
```

KMeans Clustering on Iris Dataset



28°C Mostly cloudy

Search

ENG IN

20:11 27-03-2025

jupyter Clustering_algorithm Last Checkpoint: 3 minutes ago

File Edit View Run Kernel Settings Help

Trusted

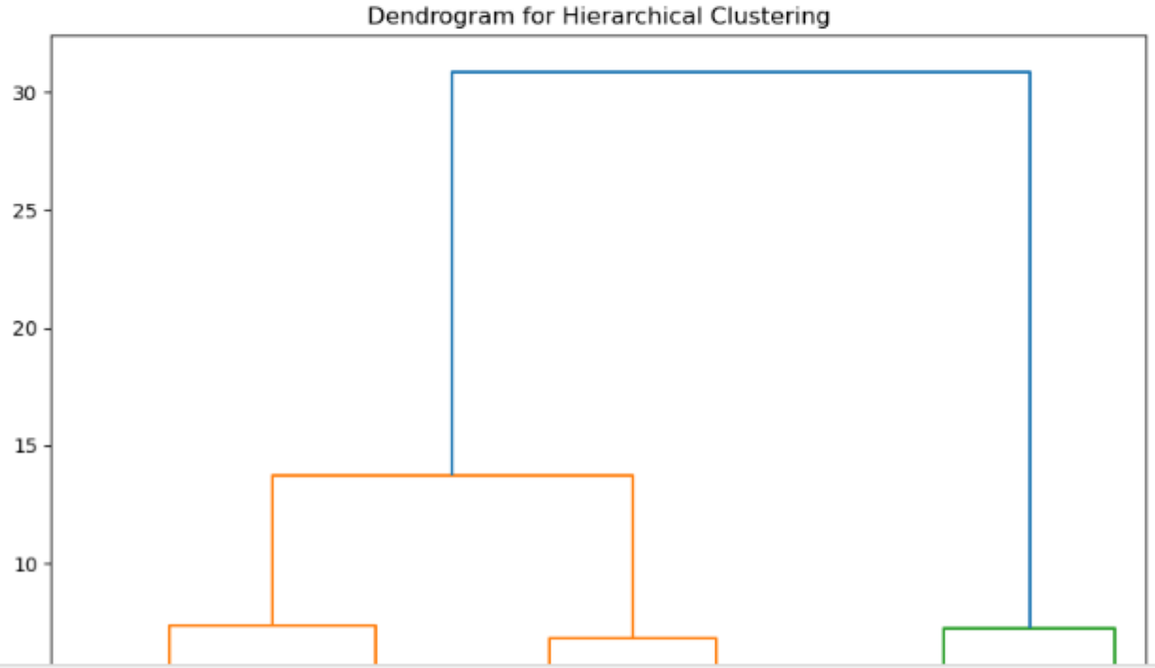
Code

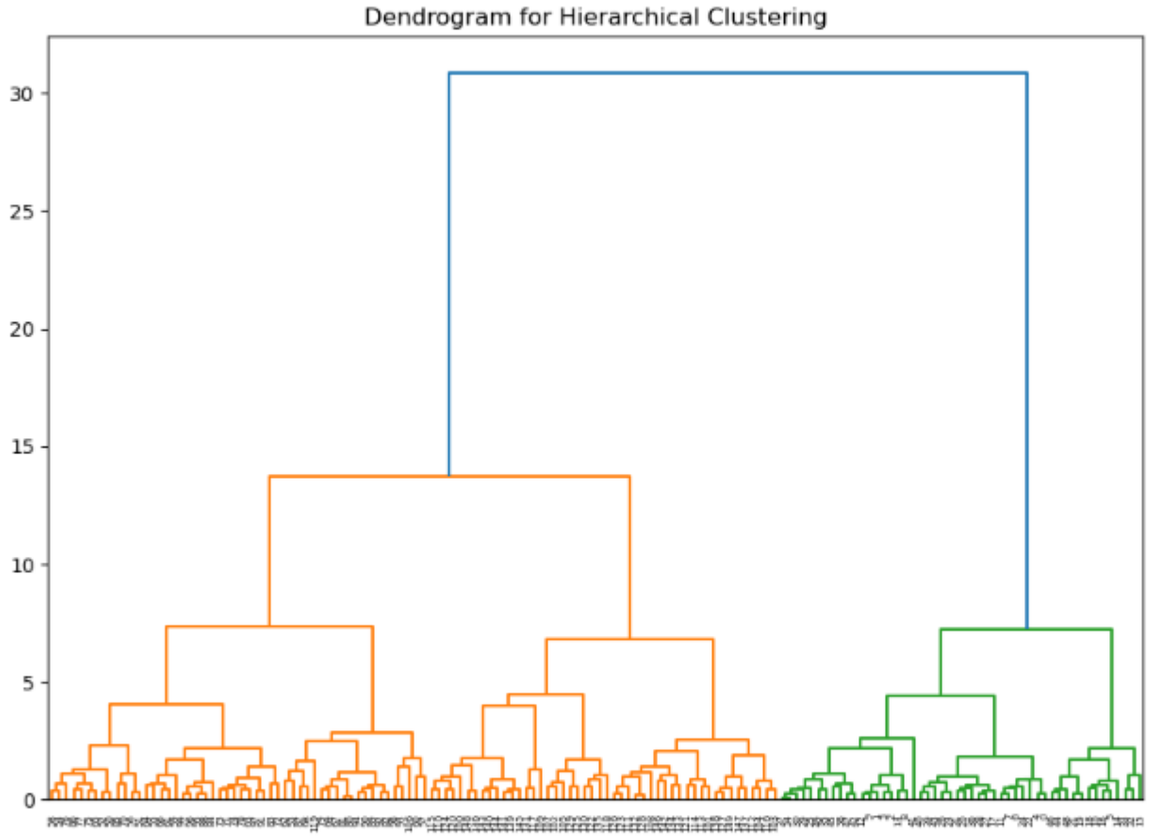
JupyterLab Python 3 (ipykernel)

[11]:
```
from sklearn.preprocessing import StandardScaler
from scipy.cluster.hierarchy import dendrogram, linkage
from sklearn.cluster import AgglomerativeClustering

# Standardize features
scaler = StandardScaler()
X_scaled = scaler.fit_transform(X)

# Dendrogram
linked = linkage(X_scaled, method='ward')
plt.figure(figsize=(10, 7))
dendrogram(linked, orientation='top', distance_sort='descending', show_leaf_counts=True)
plt.title("Dendrogram for Hierarchical Clustering")
plt.show()
```





[ ]:
```
###Hierarchial Clustering
Hierarchical clustering creates a hierarchy of clusters using a bottom-up approach (agglomerative).
It starts by treating each point as a single cluster and merges the closest pairs based on linkage criteria.
```

[12]:
```
hc = AgglomerativeClustering(n_clusters=3)
```

Sign in | Inbox (100) - revathyaramam@gn | WhatsApp | Home | Clustering_algorithm

localhost:8888/notebooks/Clustering_algorithm.ipynb

jupyter Clustering_algorithm Last Checkpoint: 4 minutes ago

File Edit View Run Kernel Settings Help

Trusted

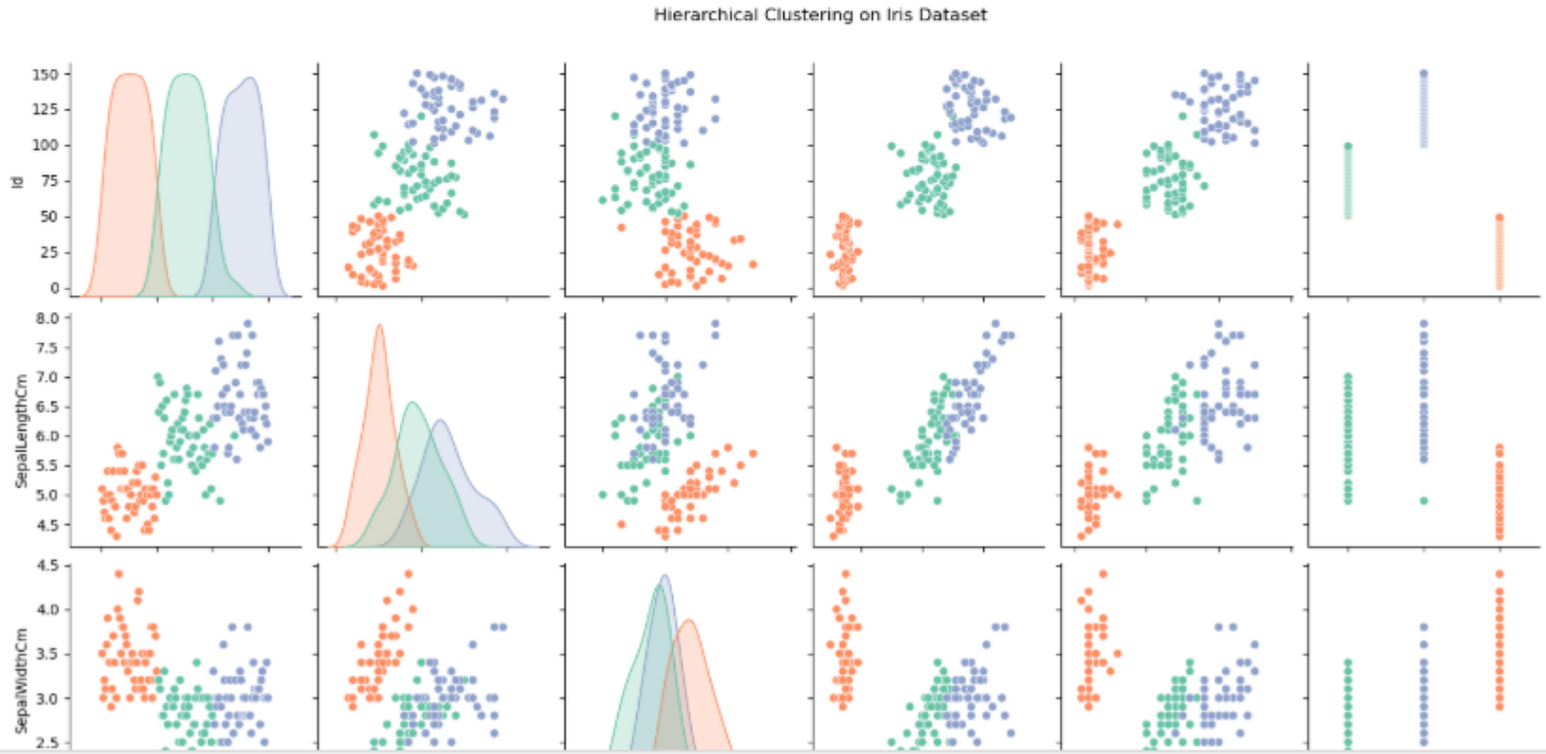
Code

JupyterLab

Python 3 (ipykernel)

[ ]:
```
###Hierarchial Clustering
Hierarchical clustering creates a hierarchy of clusters using a bottom-up approach (agglomerative).
It starts by treating each point as a single cluster and merges the closest pairs based on linkage criteria.
```

[12]:
```
hc = AgglomerativeClustering(n_clusters=3)
df['HC_Cluster'] = hc.fit_predict(X_scaled)
sns.pairplot(df, hue='HC_Cluster', palette='Set2')
plt.suptitle("Hierarchical Clustering on Iris Dataset", y=1.02)
plt.show()
```



28°C Mostly cloudy

Search

ENG IN

20:13 27-03-2025

Sign in | Inbox (100) - revathyaramam@gm | WhatsApp | Home | Clustering_algorithm

localhost:8888/notebooks/Clustering_algorithm.ipynb

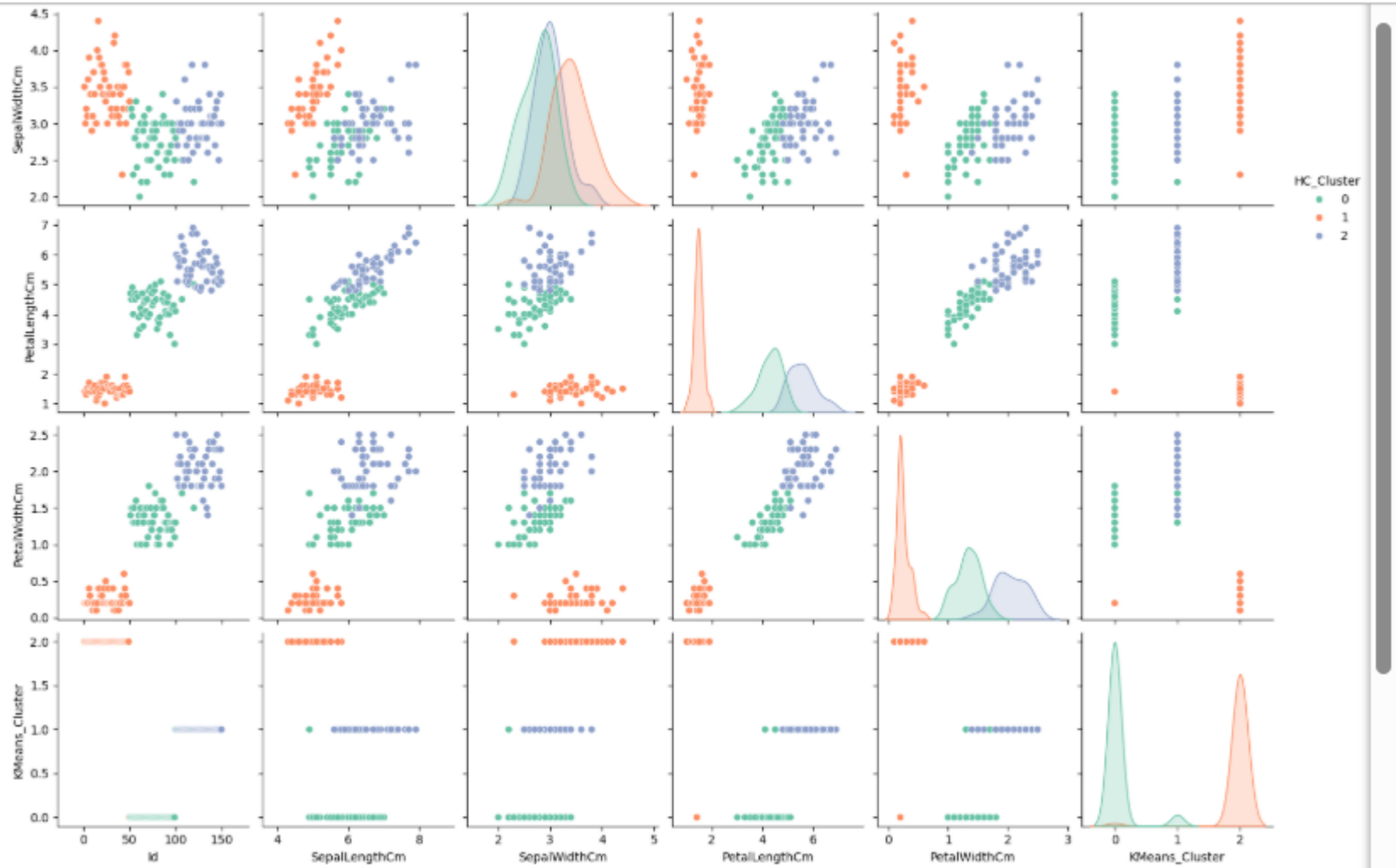
jupyter Clustering_algorithm Last Checkpoint: 7 minutes ago

File Edit View Run Kernel Settings Help

Trusted

Code

JupyterLab | Python 3 (ipykernel)



28°C Mostly cloudy

Search

ENG IN

20:16 27-03-2025