sda

November 17, 2022

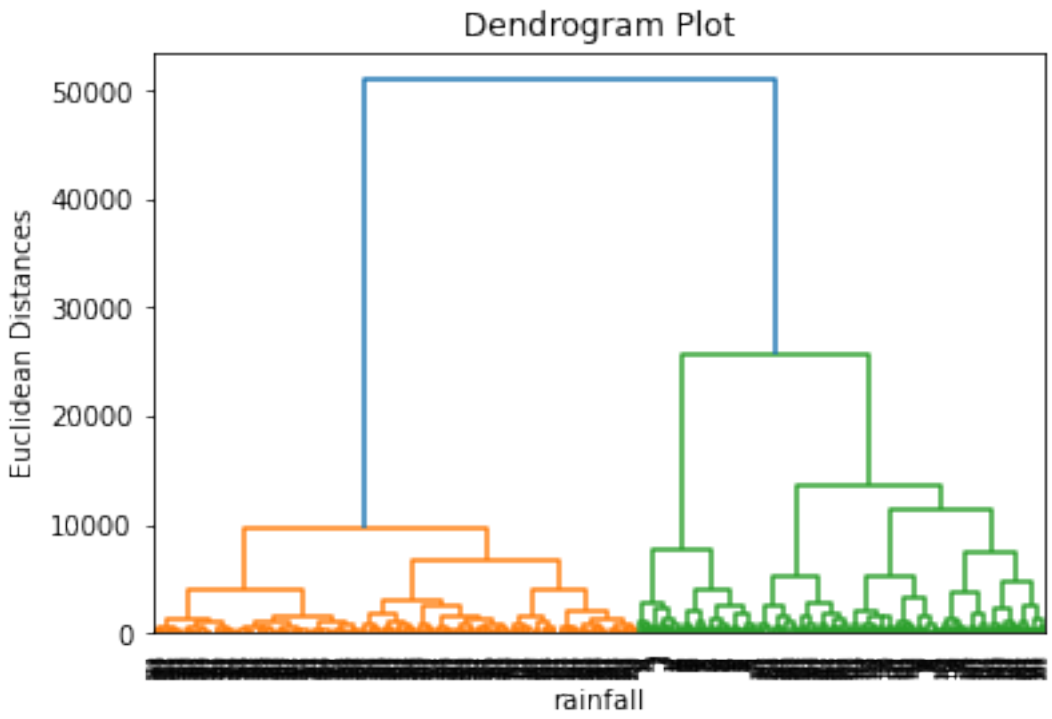
[1]:
```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
%matplotlib inline
```

[2]:
```
#dfpre=pd.read_csv("/new_pre.csv") #containig precepetation file at 525 grid
#dftemp=pd.read_csv("/new_temp.csv")#containing temperature data at 525 grid
df=pd.read_csv("/content/Mean_skew_rain.csv")# containing mean and skew␣
 ↪features of 525 grid for rain
df1=pd.read_csv("/content/Mean_skew_temp.csv")# containing mean and skew␣
 ↪features of 525 grid for temperature
df2=pd.read_csv("/content/Mean_kurtosis_rain.csv") # containing mean and␣
 ↪kurtosis features of 525 grid for rain
df3=pd.read_csv("/content/Mean_kurtosis_temp.csv")# containing mean and␣
 ↪kurtosis features of 525 grid for temperature
```

[3]:
```
x=df.values
x1=df1.values
x2=df2.values
x3=df3.values
```

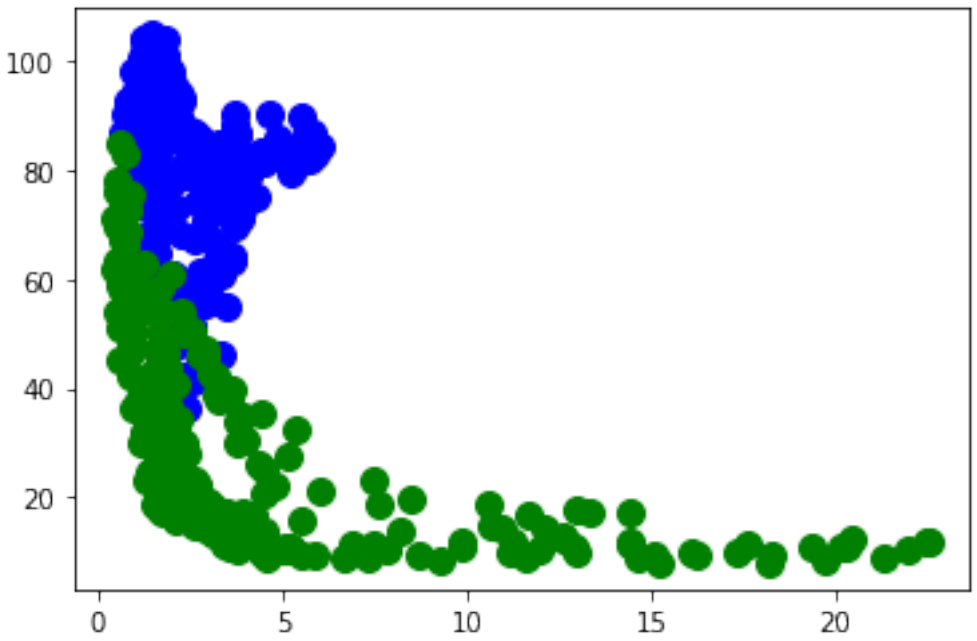
Mean and skewness based hirerichal clustering

[4]:
```
import scipy.cluster.hierarchy as shc
dendro = shc.dendrogram(shc.linkage(x, method="ward"))
plt.title("Dendrogram Plot")
plt.ylabel("Euclidean Distances")
plt.xlabel("rainfall")
plt.show()
```



[5]:
```
from sklearn.cluster import AgglomerativeClustering
hc= AgglomerativeClustering(n_clusters=2, affinity='euclidean', linkage='ward')
y_pred= hc.fit_predict(x)
```

[6]:
```
plt.scatter(x[y_pred == 0, 0], x[y_pred == 0, 1], s = 100, c = 'blue', label =␣
↪'Cluster 1')
plt.scatter(x[y_pred == 1, 0], x[y_pred == 1, 1], s = 100, c = 'green', label =␣
↪'Cluster 2')
plt.show()
```



[8]:
```
dflatlong=pd.read_csv("/content/lat_long.csv") #containing latitude and
↪longitude
```

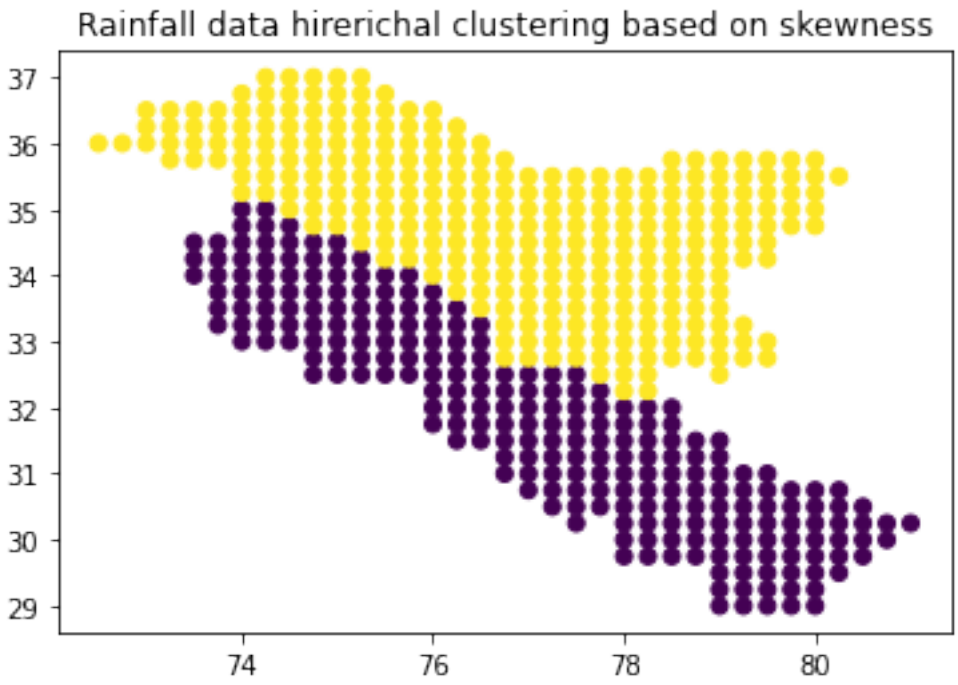
[9]:
```
from sklearn.metrics import silhouette_score
silhouette_avg = silhouette_score(x, y_pred)
```

[10]:
```
print("rainfall silhourtte test bsed on skewness",silhouette_avg)
```

```
rainfall silhourtte test bsed on skewness 0.5457798238080422
```

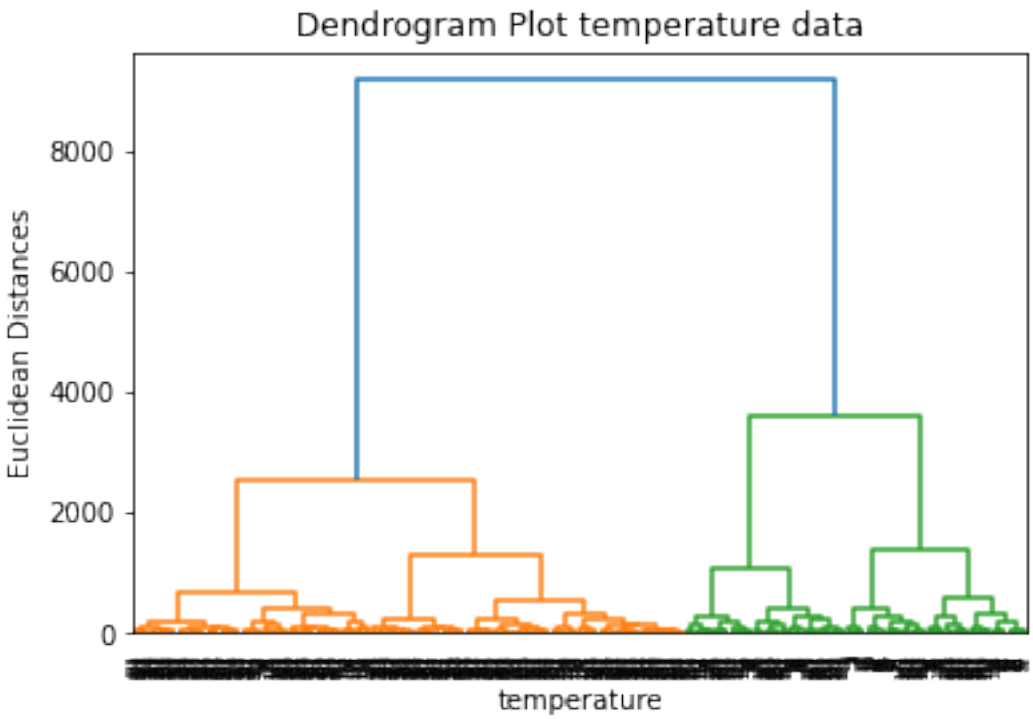
[11]:
```
plt.scatter(dflatlong.x,dflatlong.y,c=y_pred)
plt.title("Rainfall data hirerichal clustering based on skewness")
```

[11]:
```
Text(0.5, 1.0, 'Rainfall data hirerichal clustering based on skewness')
```



Mean and skewness based Hirerichal clustering on temperature data

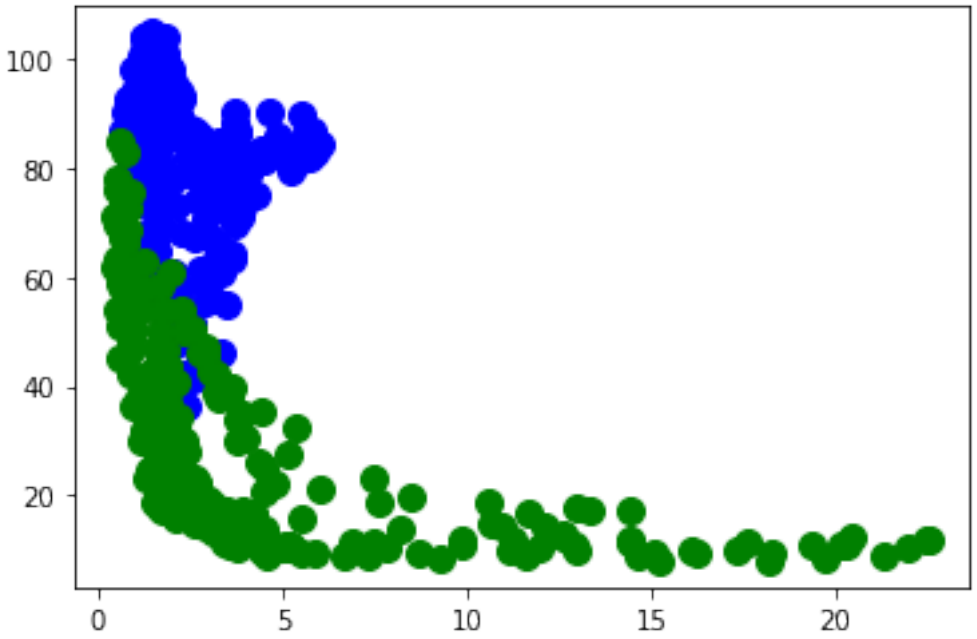
[12]:
```
import scipy.cluster.hierarchy as shc
dendro = shc.dendrogram(shc.linkage(x1, method="ward"))
plt.title("Dendrogram Plot temperature data")
plt.ylabel("Euclidean Distances")
plt.xlabel("temperature")
plt.show()
```



[13]:
```
from sklearn.cluster import AgglomerativeClustering
hc= AgglomerativeClustering(n_clusters=2, affinity='euclidean', linkage='ward')
y_pred1= hc.fit_predict(x1)
```

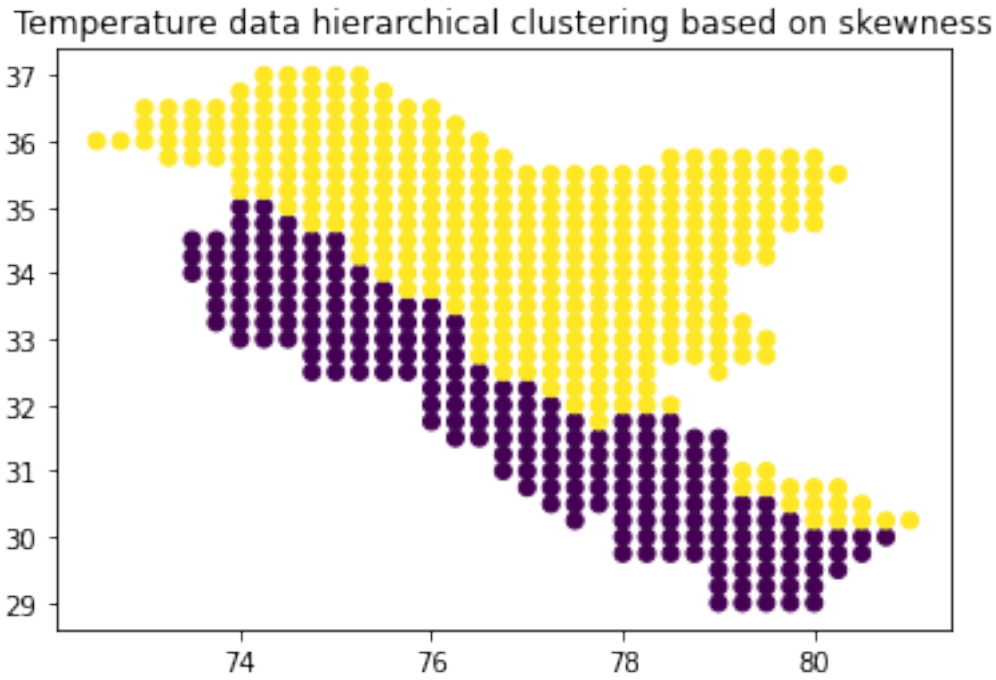
[14]:
```
plt.scatter(x[y_pred == 0, 0], x[y_pred == 0, 1], s = 100, c = 'blue', label =␣
↪'Cluster 1')
plt.scatter(x[y_pred == 1, 0], x[y_pred == 1, 1], s = 100, c = 'green', label =␣
↪'Cluster 2')
```

[14]:
```
<matplotlib.collections.PathCollection at 0x7f552a616fd0>
```



[15]:
```
plt.scatter(dflatlong.x,dflatlong.y,c=y_pred1)
plt.title("Temperature data hierarchical clustering based on skewness")
```

[15]: Text(0.5, 1.0, 'Temperature data hierarchical clustering based on skewness')



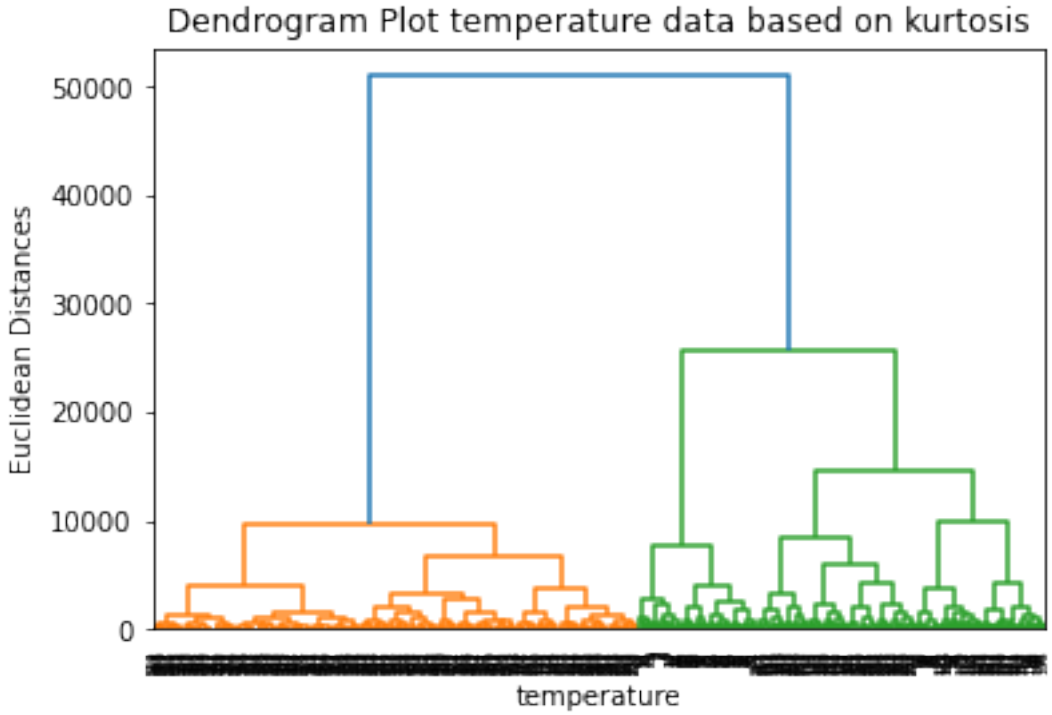
[16]:
```
silhouette_avg = silhouette_score(x1, y_pred1)
```

[17]:
```
print("Silhouette_test for temperature data",silhouette_avg)
```

```
Silhouette_test for temperature data 0.6494728851859553
```

Mean and kurtosis based hirerichal clustering on temperature data

[18]:
```
import scipy.cluster.hierarchy as shc
dendro = shc.dendrogram(shc.linkage(x2, method="ward"))
plt.title("Dendrogram Plot temperature data based on kurtosis")
plt.ylabel("Euclidean Distances")
plt.xlabel("temperature")
plt.show()
```



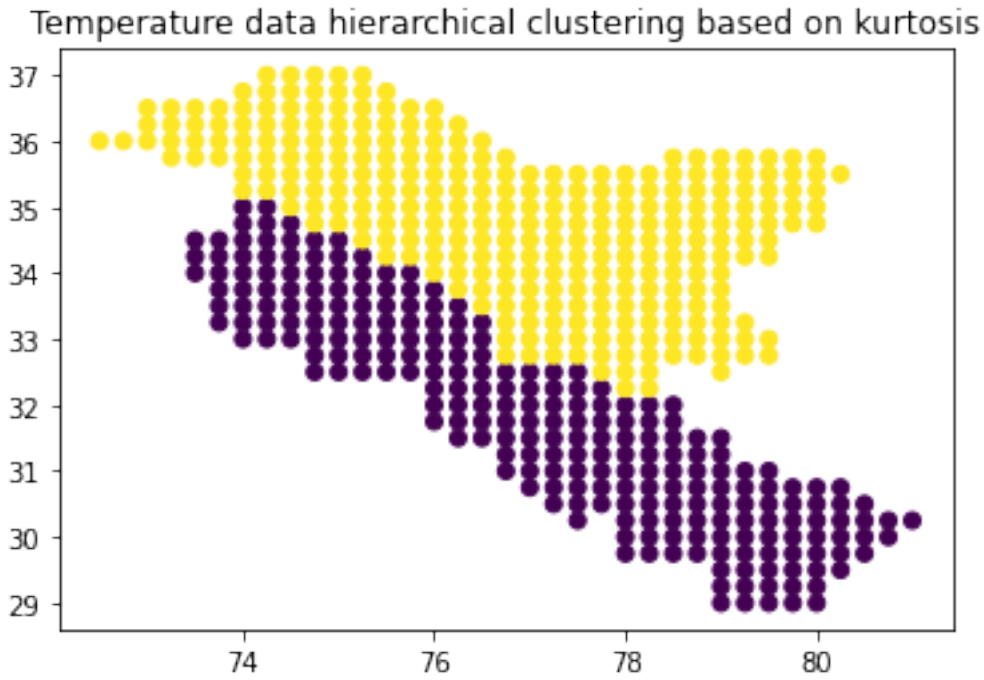
[19]:
```
from sklearn.cluster import AgglomerativeClustering
hc= AgglomerativeClustering(n_clusters=2, affinity='euclidean', linkage='ward')
y_pred2= hc.fit_predict(x2)
```

[20]:
```
silhouette_avg = silhouette_score(x2, y_pred2)
print("Silhouette_test for temperature data based on kurtosis",silhouette_avg)
```

```
Silhouette_test for temperature data based on kurtosis 0.5457833873261578
```

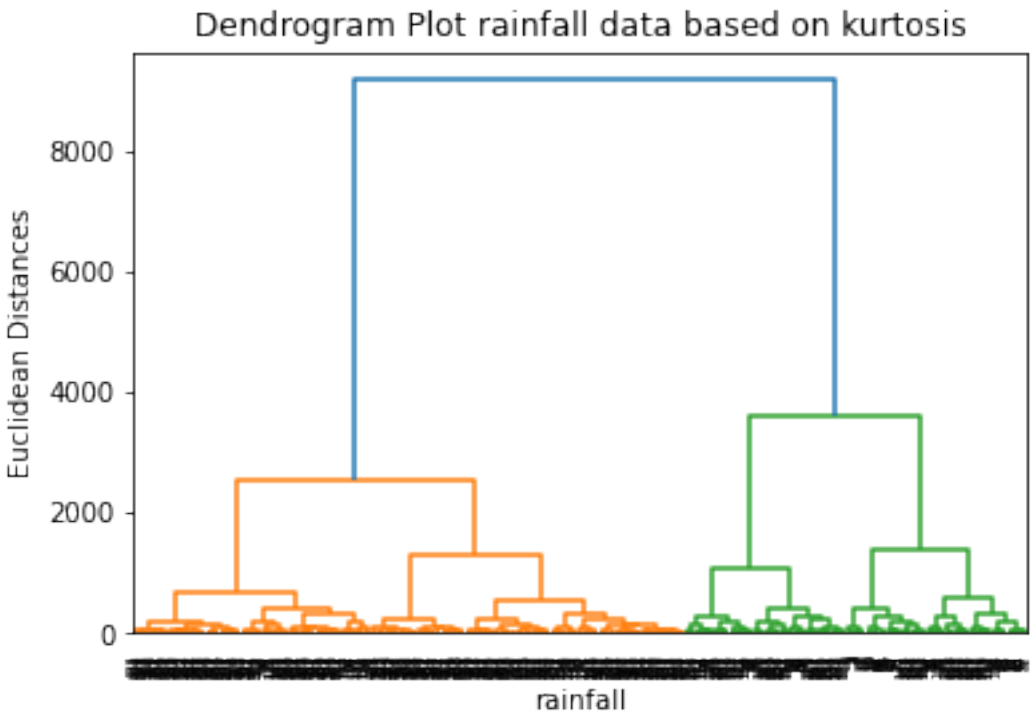
[21]:
```
plt.scatter(dflatlong.x,dflatlong.y,c=y_pred2)
plt.title("Temperature data hierarchical clustering based on kurtosis")
```

[21]:
```
Text(0.5, 1.0, 'Temperature data hierarchical clustering based on kurtosis')
```



Mean and kurtosis based hirerichal clustering on rainfall data

[22]:
```
import scipy.cluster.hierarchy as shc
dendro = shc.dendrogram(shc.linkage(x3, method="ward"))
plt.title("Dendrogram Plot rainfall data based on kurtosis")
plt.ylabel("Euclidean Distances")
plt.xlabel("rainfall")
plt.show()
```



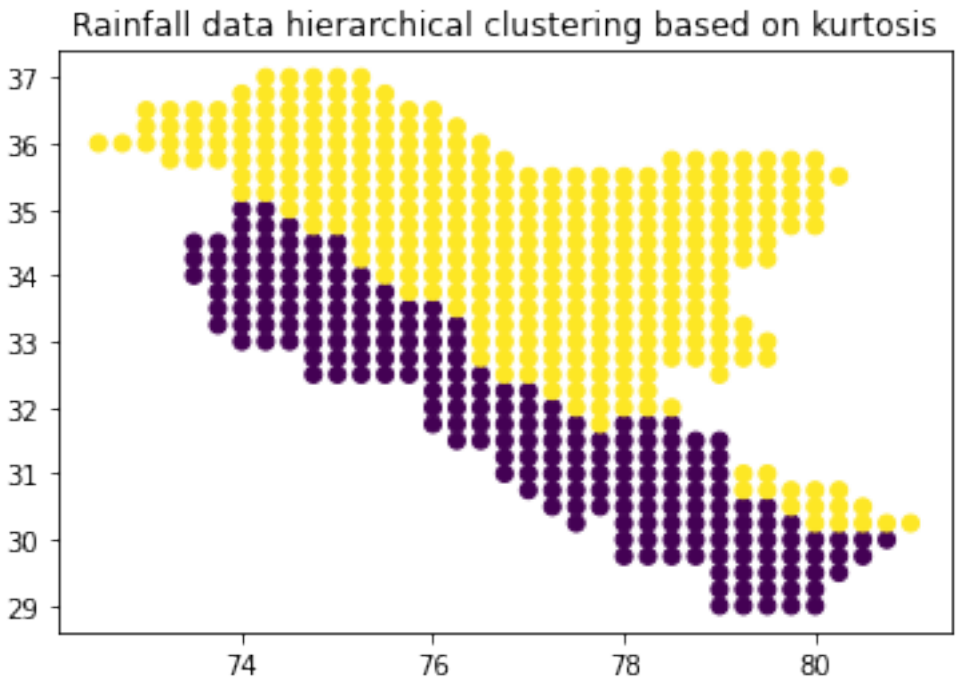
[23]:
```
hc= AgglomerativeClustering(n_clusters=2, affinity='euclidean', linkage='ward')
y_pred3= hc.fit_predict(x3)
```

[24]:
```
silhouette_avg = silhouette_score(x3, y_pred3)
print("Silhouette_test for rainfall data based on kurtosis",silhouette_avg)
```

```
Silhouette_test for rainfall data based on kurtosis 0.6494728744329883
```

[25]:
```
plt.scatter(dflatlong.x,dflatlong.y,c=y_pred3)
plt.title("Rainfall data hierarchical clustering based on kurtosis")
```

[25]:
```
Text(0.5, 1.0, 'Rainfall data hierarchical clustering based on kurtosis')
```


Rainfall data hierarchical clustering based on kurtosis
37
36
35
34
33
32
31
30
29
74
76
78
80