MS4S10_Week4_Jupyter_lab

December 13, 2019

[1]:
```
import pandas as pd
```

[2]:
```
geyser = pd.read_csv("C:\\Users\\masif\\Downloads\\faithful.csv")
```

[3]:
```
geyser.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 272 entries, 0 to 271
Data columns (total 3 columns):
Unnamed: 0    272 non-null int64
eruptions     272 non-null float64
waiting       272 non-null int64
dtypes: float64(1), int64(2)
memory usage: 6.5 KB
```

[4]:
```
geyser.head()
```

[4]:

	Unnamed: 0	eruptions	waiting
0	1	3.600	79
1	2	1.800	54
2	3	3.333	74
3	4	2.283	62
4	5	4.533	85

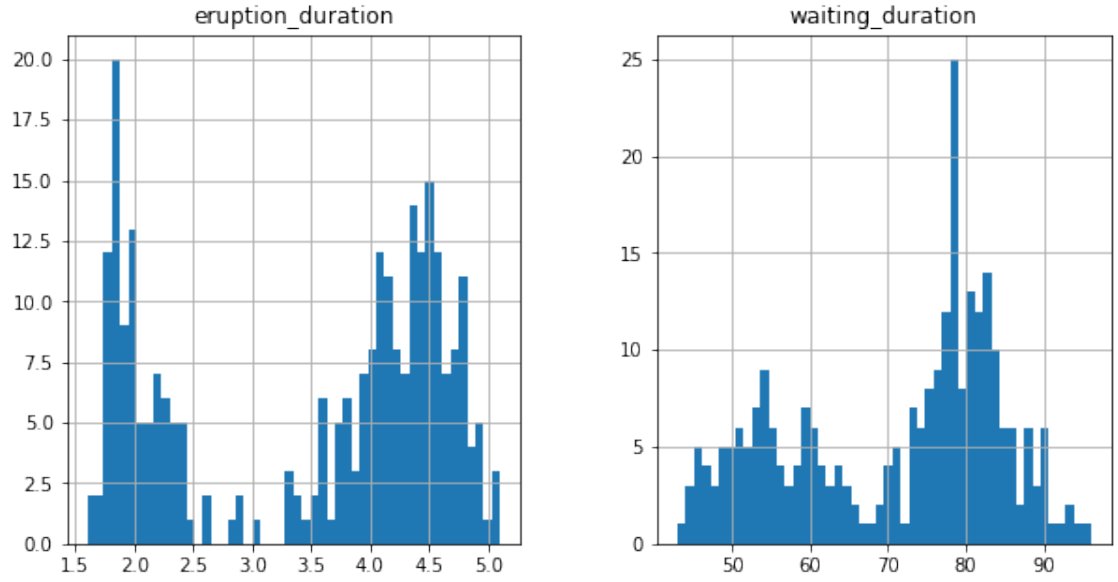
[5]:
```
geyser = geyser.drop(geyser.columns[0],axis = 1)
geyser. columns = ['eruption_duration', 'waiting_duration']
geyser.head()
```

[5]:

	eruption_duration	waiting_duration
0	3.600	79
1	1.800	54
2	3.333	74
3	2.283	62
4	4.533	85

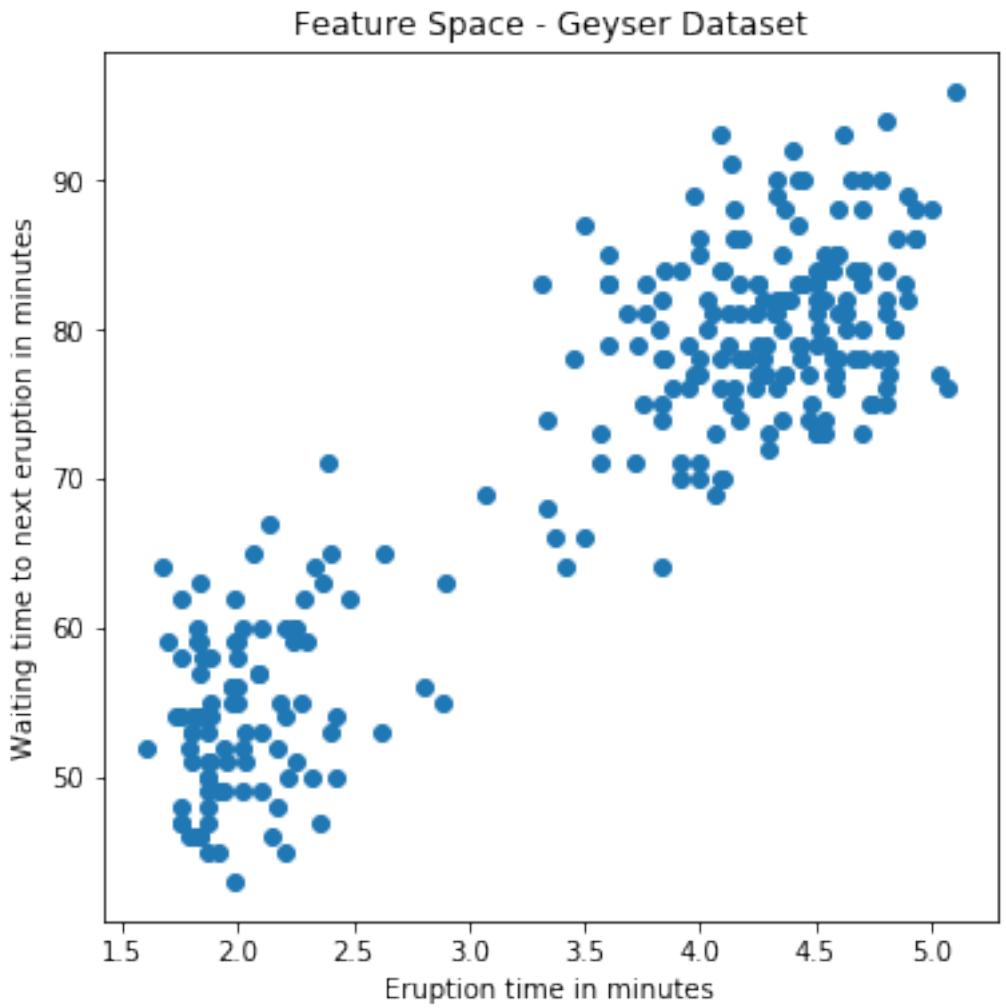
[6]:
```
import matplotlib.pyplot as plt
%matplotlib inline

geyser.hist(bins=50, figsize=(10,5))
```

[6]:
```
array([[<matplotlib.axes._subplots.AxesSubplot object at 0x0000020EE189F808>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x0000020EE524BA48>]],
      dtype=object)
```



[7]:
```
plt.figure(figsize = (6,6))
plt.scatter(geyser.iloc[:, 0],
            geyser.iloc[:,1])

plt.xlabel('Eruption time in minutes')
plt.ylabel('Waiting time to next eruption in minutes')
plt.title('Feature Space - Geyser Dataset')
```

[7]:
```
Text(0.5, 1.0, 'Feature Space - Geyser Dataset')
```



[9]:
```
from sklearn.preprocessing import StandardScaler as ss
geyser_scaled = ss().fit_transform(geyser)
geyser_scaled = pd.DataFrame(geyser_scaled, columns = ['eruption_duration',␣
↪'waiting_duration'])
geyser_scaled.head()
```

[9]:

	eruption_duration	waiting_duration
0	0.098499	0.597123
1	-1.481459	-1.245181
2	-0.135861	0.228663
3	-1.057503	-0.655644
4	0.917443	1.039277

[10]:
```
from sklearn.cluster import KMeans
```

[11]:
```
geyser_initial = geyser.iloc[0:259, :]
geyser_scaled_initial = geyser_scaled.iloc[0:259, :]
```

```
geyser_later = geyser.iloc[260:273, :]
geyser_scaled_later = geyser_scaled.iloc[260:273, :]
```

[15]:
```
km_1 = KMeans(n_clusters = 2, max_iter = 20, verbose = 1).fit(geyser_initial)
```

```
Initialization complete
start iteration
done sorting
end inner loop
Iteration 0, inertia 13371.137613107134
start iteration
done sorting
end inner loop
Iteration 1, inertia 8896.849594428875
start iteration
done sorting
end inner loop
Iteration 2, inertia 8393.843993207567
start iteration
done sorting
end inner loop
Iteration 3, inertia 8376.791060111938
start iteration
done sorting
end inner loop
Iteration 4, inertia 8376.791060111938
center shift 0.000000e+00 within tolerance 9.109693e-03
Initialization complete
start iteration
done sorting
end inner loop
Iteration 0, inertia 9487.947996633855
start iteration
done sorting
end inner loop
Iteration 1, inertia 8393.843993207567
start iteration
done sorting
end inner loop
Iteration 2, inertia 8376.791060111938
start iteration
done sorting
end inner loop
Iteration 3, inertia 8376.791060111938
center shift 0.000000e+00 within tolerance 9.109693e-03
Initialization complete
start iteration
```

```
done sorting
end inner loop
Iteration 0, inertia 9487.947996633855
start iteration
done sorting
end inner loop
Iteration 1, inertia 8393.843993207567
start iteration
done sorting
end inner loop
Iteration 2, inertia 8376.791060111938
start iteration
done sorting
end inner loop
Iteration 3, inertia 8376.791060111938
center shift 0.000000e+00 within tolerance 9.109693e-03
Initialization complete
start iteration
done sorting
end inner loop
Iteration 0, inertia 8410.259562077626
start iteration
done sorting
end inner loop
Iteration 1, inertia 8376.791060111938
start iteration
done sorting
end inner loop
Iteration 2, inertia 8376.791060111938
center shift 0.000000e+00 within tolerance 9.109693e-03
Initialization complete
start iteration
done sorting
end inner loop
Iteration 0, inertia 8551.648257333312
start iteration
done sorting
end inner loop
Iteration 1, inertia 8376.791060111938
start iteration
done sorting
end inner loop
Iteration 2, inertia 8376.791060111938
center shift 0.000000e+00 within tolerance 9.109693e-03
Initialization complete
start iteration
done sorting
end inner loop
```

```
Iteration 0, inertia 8376.791060111938
start iteration
done sorting
end inner loop
Iteration 1, inertia 8376.791060111938
center shift 0.000000e+00 within tolerance 9.109693e-03
Initialization complete
start iteration
done sorting
end inner loop
Iteration 0, inertia 8514.391768418809
start iteration
done sorting
end inner loop
Iteration 1, inertia 8376.791060111938
start iteration
done sorting
end inner loop
Iteration 2, inertia 8376.791060111938
center shift 0.000000e+00 within tolerance 9.109693e-03
Initialization complete
start iteration
done sorting
end inner loop
Iteration 0, inertia 8376.791060111938
start iteration
done sorting
end inner loop
Iteration 1, inertia 8376.791060111938
center shift 0.000000e+00 within tolerance 9.109693e-03
Initialization complete
start iteration
done sorting
end inner loop
Iteration 0, inertia 9183.686962357926
start iteration
done sorting
end inner loop
Iteration 1, inertia 8410.259562077626
start iteration
done sorting
end inner loop
Iteration 2, inertia 8376.791060111938
start iteration
done sorting
end inner loop
Iteration 3, inertia 8376.791060111938
center shift 0.000000e+00 within tolerance 9.109693e-03
```

```
Initialization complete
start iteration
done sorting
end inner loop
Iteration 0, inertia 8376.791060111938
start iteration
done sorting
end inner loop
Iteration 1, inertia 8376.791060111938
center shift 0.000000e+00 within tolerance 9.109693e-03
```

[16]:
```
km_1.cluster_centers_
```

[16]:
```
array([[ 2.09876842, 54.96842105],
       [ 4.29060976, 80.27439024]])
```

[21]:
```
km_1.labels_
```

[21]:
```
array([1, 0, 1, 0, 1, 0, 1, 1, 0, 1, 0, 1, 1, 0, 1, 0, 0, 1, 0, 1, 0, 0,
       1, 1, 1, 1, 0, 1, 1, 1, 1, 1, 0, 1, 1, 0, 0, 1, 0, 1, 1, 0, 1, 0,
       1, 1, 0, 0, 1, 0, 1, 1, 0, 1, 0, 1, 1, 0, 1, 1, 0, 1, 0, 1, 0, 1,
       1, 1, 0, 1, 1, 0, 1, 1, 0, 1, 0, 1, 1, 1, 1, 1, 1, 0, 1, 1, 1, 1,
       0, 1, 0, 1, 0, 1, 0, 1, 1, 1, 0, 1, 0, 1, 0, 1, 1, 0, 1, 0, 1, 1,
       1, 0, 1, 1, 0, 1, 0, 1, 0, 1, 0, 1, 1, 0, 1, 1, 0, 1, 0, 1, 0, 1,
       0, 1, 0, 1, 0, 1, 0, 1, 1, 0, 1, 1, 1, 0, 1, 0, 1, 0, 1, 1, 0, 1,
       1, 1, 1, 1, 0, 1, 0, 1, 0, 1, 0, 1, 0, 1, 0, 1, 0, 0, 1, 1, 1, 1,
       1, 0, 1, 1, 0, 1, 1, 1, 0, 1, 1, 0, 1, 0, 1, 0, 1, 1, 1, 1, 1, 1,
       0, 1, 0, 1, 1, 0, 1, 0, 1, 1, 0, 1, 1, 1, 0, 1, 0, 1, 0, 1, 0, 1,
       0, 1, 0, 1, 1, 1, 1, 1, 1, 1, 1, 0, 1, 0, 1, 0, 0, 1, 1, 0, 1, 0,
       1, 0, 1, 1, 0, 1, 0, 1, 0, 1, 1, 1, 1, 1, 1, 1, 0])
```

[17]:
```
km_1.inertia_
```

[17]:
```
8376.791060111938
```

[18]:
```
km_1_ss = KMeans(n_clusters = 2, max_iter = 2, verbose = 1).
↪fit(geyser_scaled_initial)
```

```
Initialization complete
start iteration
done sorting
end inner loop
Iteration 0, inertia 76.26914949135022
start iteration
done sorting
end inner loop
Iteration 1, inertia 76.20433662586007
Initialization complete
start iteration
done sorting
```

```
end inner loop
Iteration 0, inertia 82.25398666653072
start iteration
done sorting
end inner loop
Iteration 1, inertia 76.26914949135022
Initialization complete
start iteration
done sorting
end inner loop
Iteration 0, inertia 77.49917079101512
start iteration
done sorting
end inner loop
Iteration 1, inertia 76.24234090685127
Initialization complete
start iteration
done sorting
end inner loop
Iteration 0, inertia 77.49917079101512
start iteration
done sorting
end inner loop
Iteration 1, inertia 76.24234090685127
Initialization complete
start iteration
done sorting
end inner loop
Iteration 0, inertia 77.9206690263899
start iteration
done sorting
end inner loop
Iteration 1, inertia 76.26914949135022
Initialization complete
start iteration
done sorting
end inner loop
Iteration 0, inertia 78.53314617814105
start iteration
done sorting
end inner loop
Iteration 1, inertia 76.42936722079398
Initialization complete
start iteration
done sorting
end inner loop
Iteration 0, inertia 76.26914949135022
start iteration
```

```
done sorting
end inner loop
Iteration 1, inertia 76.20433662586007
Initialization complete
start iteration
done sorting
end inner loop
Iteration 0, inertia 82.25398666653072
start iteration
done sorting
end inner loop
Iteration 1, inertia 76.26914949135022
Initialization complete
start iteration
done sorting
end inner loop
Iteration 0, inertia 76.24234090685127
start iteration
done sorting
end inner loop
Iteration 1, inertia 76.20433662586007
Initialization complete
start iteration
done sorting
end inner loop
Iteration 0, inertia 78.53314617814105
start iteration
done sorting
end inner loop
Iteration 1, inertia 76.42936722079398
```

[19]:
```
km_1_ss.cluster_centers_
```

[19]:
```
array([[ 0.70328731,  0.67525491],
       [-1.25809373, -1.185752  ]])
```

[20]:
```
km_1_ss.inertia_
```

[20]:
```
76.20433662586007
```

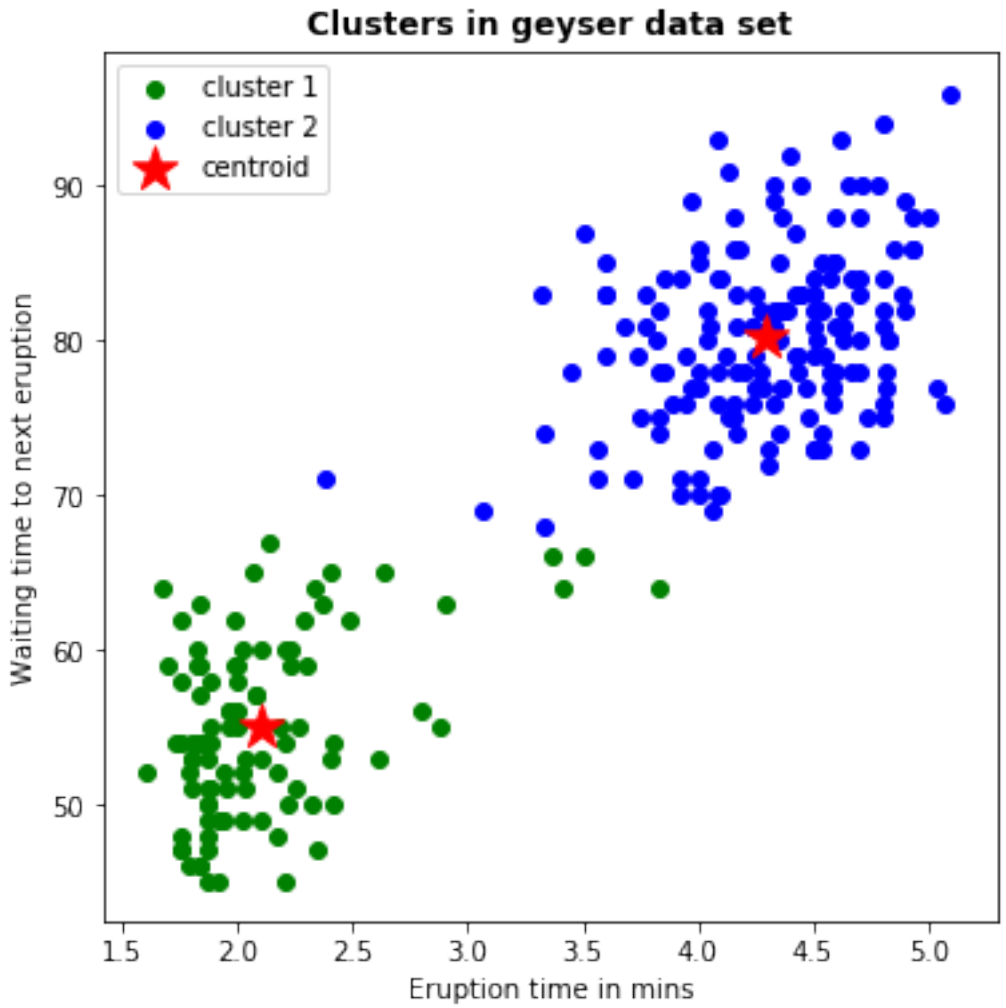
[22]:
```
km_1_ss.labels_
```

[22]:
```
array([0, 1, 0, 1, 0, 1, 0, 0, 1, 0, 1, 0, 0, 1, 0, 1, 1, 0, 1, 0, 1, 1,
       0, 0, 0, 0, 1, 0, 0, 0, 0, 0, 0, 0, 0, 1, 1, 0, 1, 0, 0, 1, 0, 1,
       0, 0, 0, 1, 0, 1, 0, 0, 1, 0, 1, 0, 0, 1, 0, 0, 1, 0, 1, 0, 1, 0,
       0, 0, 1, 0, 0, 1, 0, 0, 1, 0, 1, 0, 0, 0, 0, 0, 0, 1, 0, 0, 0, 0,
       1, 0, 1, 0, 1, 0, 1, 0, 0, 0, 1, 0, 1, 0, 1, 0, 0, 1, 0, 1, 0, 0,
       0, 1, 0, 0, 1, 0, 1, 0, 1, 0, 1, 0, 0, 1, 0, 0, 1, 0, 1, 0, 1, 0,
       1, 0, 1, 0, 1, 0, 1, 0, 0, 1, 0, 0, 0, 1, 0, 1, 0, 1, 0, 0, 1, 0,
       0, 0, 0, 0, 1, 0, 1, 0, 1, 0, 0, 0, 1, 0, 1, 0, 1, 1, 0, 0, 0, 0,
```

```
       0, 1, 0, 0, 1, 0, 0, 0, 1, 0, 0, 1, 0, 1, 0, 1, 0, 0, 0, 0, 0, 0,
       1, 0, 1, 0, 0, 1, 0, 1, 0, 0, 1, 0, 1, 0, 1, 0, 1, 0, 1, 0, 1, 0,
       1, 0, 1, 0, 0, 0, 0, 0, 0, 0, 0, 1, 0, 1, 0, 1, 1, 0, 0, 1, 0, 1,
       0, 1, 0, 0, 1, 0, 1, 0, 1, 0, 0, 0, 0, 0, 0, 0, 1])
```

[28]:
```
fig, ax = plt.subplots(figsize = (6,6)) #initialise the plots

#first cluster
plt.scatter(x = geyser_initial.iloc[geyser_initial.index[km_1.labels_ == 0].
 ↪tolist() , [0]],
            y = geyser_initial.iloc[geyser_initial.index[km_1.labels_ == 0].
 ↪tolist() , [1]],
            c = 'green',
            label = 'cluster 1')

#second cluster
plt.scatter(x = geyser_initial.iloc[geyser_initial.index[km_1.labels_ == 1].
 ↪tolist() , [0]],
            y = geyser_initial.iloc[geyser_initial.index[km_1.labels_ == 1].
 ↪tolist() , [1]],
            c = 'blue',
            label = 'cluster 2')

#centroid
plt.scatter(x = km_1.cluster_centers_[:, [0]],
             y = km_1.cluster_centers_[:, [1]],
             marker = '*', s = 300, label = 'centroid', c = 'r')

plt.legend()
plt.xlabel('Eruption time in mins')
plt.ylabel('Waiting time to next eruption')
plt.title('Clusters in geyser data set', fontweight = 'bold')
```

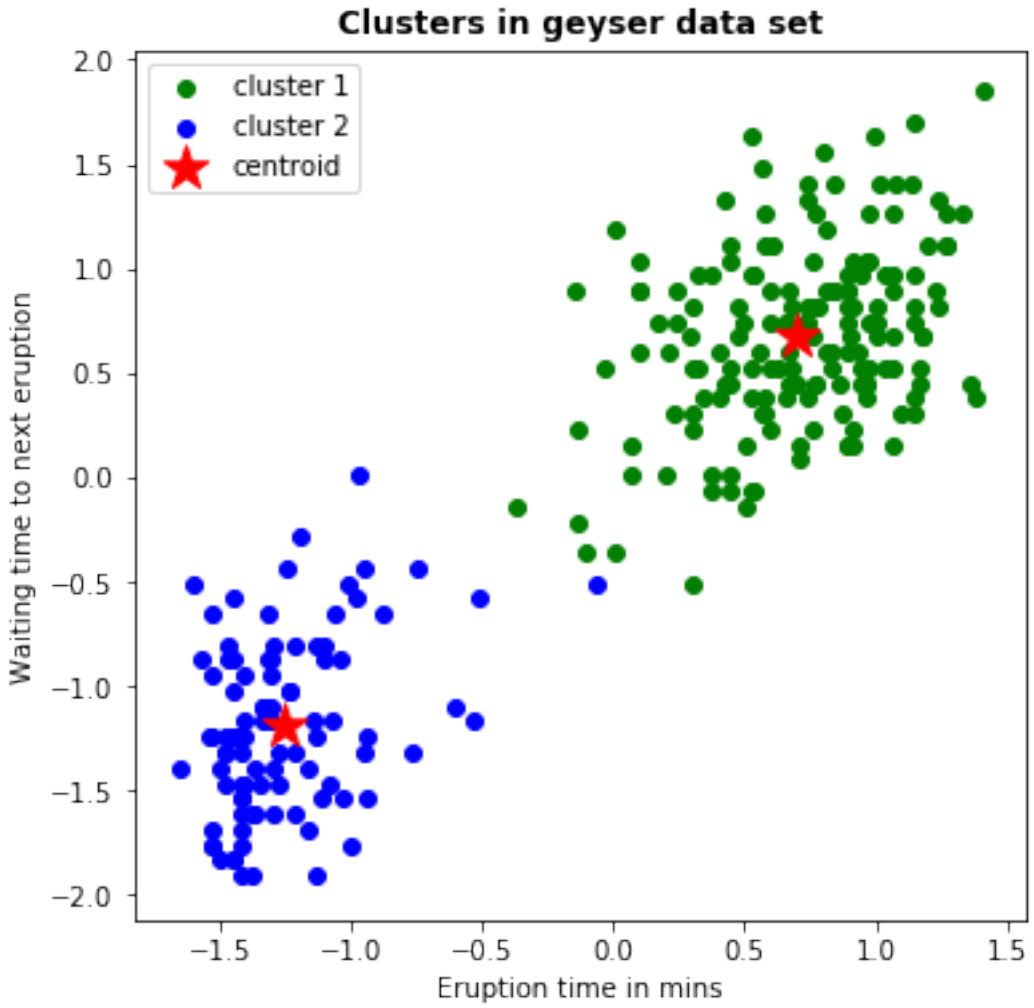
[28]:
```
Text(0.5, 1.0, 'Clusters in geyser data set')
```



[30]:
```
fig, ax = plt.subplots(figsize = (6,6)) #initialise the plots

#first cluster
plt.scatter(x = geyser_scaled_initial.iloc[geyser_scaled_initial.index[km_1_ss.
 ↪labels_ == 0].tolist() , [0]],
            y = geyser_scaled_initial.iloc[geyser_scaled_initial.index[km_1_ss.
 ↪labels_ == 0].tolist() , [1]],
            c = 'green',
            label = 'cluster 1')

#second cluster
plt.scatter(x = geyser_scaled_initial.iloc[geyser_scaled_initial.index[km_1_ss.
 ↪labels_ == 1].tolist() , [0]],
            y = geyser_scaled_initial.iloc[geyser_scaled_initial.index[km_1_ss.
 ↪labels_ == 1].tolist() , [1]],
            c = 'blue',
```

```
            label = 'cluster 2')

#centroid
plt.scatter(x = km_1_ss.cluster_centers_[:, [0]],
            y = km_1_ss.cluster_centers_[:, [1]],
            marker = '*', s = 300, label = 'centroid', c = 'r')

plt.legend()
plt.xlabel('Eruption time in mins')
plt.ylabel('Waiting time to next eruption')
plt.title('Clusters in geyser data set', fontweight = 'bold')
```

[30]:
```
Text(0.5, 1.0, 'Clusters in geyser data set')
```



[32]:
```
km_2_ss = KMeans(n_clusters = 3, max_iter = 20, verbose = 1).
↪fit(geyser_scaled_initial)
```

```
Initialization complete
start iteration
done sorting
end inner loop
Iteration 0, inertia 54.133160121207325
start iteration
done sorting
end inner loop
Iteration 1, inertia 53.927501057633286
start iteration
done sorting
end inner loop
Iteration 2, inertia 53.8046635643386
start iteration
done sorting
end inner loop
Iteration 3, inertia 53.72328202973945
start iteration
done sorting
end inner loop
Iteration 4, inertia 53.58904057565307
start iteration
done sorting
end inner loop
Iteration 5, inertia 53.244737533279995
start iteration
done sorting
end inner loop
Iteration 6, inertia 53.13740995050338
start iteration
done sorting
end inner loop
Iteration 7, inertia 53.13740995050338
center shift 0.000000e+00 within tolerance 9.883149e-05
Initialization complete
start iteration
done sorting
end inner loop
Iteration 0, inertia 63.31550836980897
start iteration
done sorting
end inner loop
Iteration 1, inertia 62.35687492741287
start iteration
done sorting
end inner loop
Iteration 2, inertia 61.96109922499922
start iteration
```

```
done sorting
end inner loop
Iteration 3, inertia 61.92589173896566
start iteration
done sorting
end inner loop
Iteration 4, inertia 61.92589173896566
center shift 0.000000e+00 within tolerance 9.883149e-05
Initialization complete
start iteration
done sorting
end inner loop
Iteration 0, inertia 56.91129224492146
start iteration
done sorting
end inner loop
Iteration 1, inertia 54.24760854103371
start iteration
done sorting
end inner loop
Iteration 2, inertia 53.56562457142724
start iteration
done sorting
end inner loop
Iteration 3, inertia 53.16702859579839
start iteration
done sorting
end inner loop
Iteration 4, inertia 53.11526887354445
start iteration
done sorting
end inner loop
Iteration 5, inertia 53.092293761447195
start iteration
done sorting
end inner loop
Iteration 6, inertia 53.092293761447195
center shift 0.000000e+00 within tolerance 9.883149e-05
Initialization complete
start iteration
done sorting
end inner loop
Iteration 0, inertia 56.61228907618525
start iteration
done sorting
end inner loop
Iteration 1, inertia 55.19363871408681
start iteration
```

```
done sorting
end inner loop
Iteration 2, inertia 54.23575019641109
start iteration
done sorting
end inner loop
Iteration 3, inertia 53.76530064665067
start iteration
done sorting
end inner loop
Iteration 4, inertia 53.49549235402023
start iteration
done sorting
end inner loop
Iteration 5, inertia 53.32046592282783
start iteration
done sorting
end inner loop
Iteration 6, inertia 53.13740995050338
start iteration
done sorting
end inner loop
Iteration 7, inertia 53.13740995050338
center shift 0.000000e+00 within tolerance 9.883149e-05
Initialization complete
start iteration
done sorting
end inner loop
Iteration 0, inertia 64.32120613952196
start iteration
done sorting
end inner loop
Iteration 1, inertia 58.28826912259212
start iteration
done sorting
end inner loop
Iteration 2, inertia 54.52754432694722
start iteration
done sorting
end inner loop
Iteration 3, inertia 53.34224648872054
start iteration
done sorting
end inner loop
Iteration 4, inertia 53.14614252529202
start iteration
done sorting
end inner loop
```

```
Iteration 5, inertia 53.10378851613723
start iteration
done sorting
end inner loop
Iteration 6, inertia 53.10378851613723
center shift 0.000000e+00 within tolerance 9.883149e-05
Initialization complete
start iteration
done sorting
end inner loop
Iteration 0, inertia 56.38679016373704
start iteration
done sorting
end inner loop
Iteration 1, inertia 55.113753775545696
start iteration
done sorting
end inner loop
Iteration 2, inertia 54.23575019641109
start iteration
done sorting
end inner loop
Iteration 3, inertia 53.76530064665067
start iteration
done sorting
end inner loop
Iteration 4, inertia 53.49549235402023
start iteration
done sorting
end inner loop
Iteration 5, inertia 53.32046592282783
start iteration
done sorting
end inner loop
Iteration 6, inertia 53.13740995050338
start iteration
done sorting
end inner loop
Iteration 7, inertia 53.13740995050338
center shift 0.000000e+00 within tolerance 9.883149e-05
Initialization complete
start iteration
done sorting
end inner loop
Iteration 0, inertia 53.52286174745079
start iteration
done sorting
end inner loop
```

```
Iteration 1, inertia 53.09082746696128
start iteration
done sorting
end inner loop
Iteration 2, inertia 53.09082746696128
center shift 0.000000e+00 within tolerance 9.883149e-05
Initialization complete
start iteration
done sorting
end inner loop
Iteration 0, inertia 65.01299927783818
start iteration
done sorting
end inner loop
Iteration 1, inertia 56.92501496550979
start iteration
done sorting
end inner loop
Iteration 2, inertia 54.275568113706385
start iteration
done sorting
end inner loop
Iteration 3, inertia 53.91820543833627
start iteration
done sorting
end inner loop
Iteration 4, inertia 53.8046635643386
start iteration
done sorting
end inner loop
Iteration 5, inertia 53.72328202973945
start iteration
done sorting
end inner loop
Iteration 6, inertia 53.58904057565307
start iteration
done sorting
end inner loop
Iteration 7, inertia 53.244737533279995
start iteration
done sorting
end inner loop
Iteration 8, inertia 53.13740995050338
start iteration
done sorting
end inner loop
Iteration 9, inertia 53.13740995050338
center shift 0.000000e+00 within tolerance 9.883149e-05
```

```
Initialization complete
start iteration
done sorting
end inner loop
Iteration 0, inertia 54.927959437708004
start iteration
done sorting
end inner loop
Iteration 1, inertia 53.927501057633286
start iteration
done sorting
end inner loop
Iteration 2, inertia 53.8046635643386
start iteration
done sorting
end inner loop
Iteration 3, inertia 53.72328202973945
start iteration
done sorting
end inner loop
Iteration 4, inertia 53.58904057565307
start iteration
done sorting
end inner loop
Iteration 5, inertia 53.244737533279995
start iteration
done sorting
end inner loop
Iteration 6, inertia 53.13740995050338
start iteration
done sorting
end inner loop
Iteration 7, inertia 53.13740995050338
center shift 0.000000e+00 within tolerance 9.883149e-05
Initialization complete
start iteration
done sorting
end inner loop
Iteration 0, inertia 61.81673119143812
start iteration
done sorting
end inner loop
Iteration 1, inertia 61.09155355813671
start iteration
done sorting
end inner loop
Iteration 2, inertia 61.09155355813671
center shift 0.000000e+00 within tolerance 9.883149e-05
```

[45]:
```
km_3_ss = KMeans(n_clusters = 4, max_iter = 20, verbose = 1).
↪fit(geyser_scaled_initial)
```

```
Initialization complete
start iteration
done sorting
end inner loop
Iteration 0, inertia 47.04572186639287
start iteration
done sorting
end inner loop
Iteration 1, inertia 45.91263226909665
start iteration
done sorting
end inner loop
Iteration 2, inertia 45.65206274766136
start iteration
done sorting
end inner loop
Iteration 3, inertia 45.5138288705064
start iteration
done sorting
end inner loop
Iteration 4, inertia 45.39458473922034
start iteration
done sorting
end inner loop
Iteration 5, inertia 45.38477634499685
start iteration
done sorting
end inner loop
Iteration 6, inertia 45.38477634499685
center shift 0.000000e+00 within tolerance 9.883149e-05
Initialization complete
start iteration
done sorting
end inner loop
Iteration 0, inertia 50.47912975152537
start iteration
done sorting
end inner loop
Iteration 1, inertia 48.12753420461566
start iteration
done sorting
end inner loop
Iteration 2, inertia 47.31883270202541
start iteration
```

```
done sorting
end inner loop
Iteration 3, inertia 46.986124961839835
start iteration
done sorting
end inner loop
Iteration 4, inertia 46.81091955069067
start iteration
done sorting
end inner loop
Iteration 5, inertia 46.70648351504147
start iteration
done sorting
end inner loop
Iteration 6, inertia 46.366592079411795
start iteration
done sorting
end inner loop
Iteration 7, inertia 45.78230489421425
start iteration
done sorting
end inner loop
Iteration 8, inertia 45.70150763215784
start iteration
done sorting
end inner loop
Iteration 9, inertia 45.56974340487012
start iteration
done sorting
end inner loop
Iteration 10, inertia 45.31868935272981
start iteration
done sorting
end inner loop
Iteration 11, inertia 44.865550836927966
start iteration
done sorting
end inner loop
Iteration 12, inertia 44.64519661848327
start iteration
done sorting
end inner loop
Iteration 13, inertia 44.48969466773778
start iteration
done sorting
end inner loop
Iteration 14, inertia 44.4086346381217
start iteration
```

```
done sorting
end inner loop
Iteration 15, inertia 44.351039100120346
start iteration
done sorting
end inner loop
Iteration 16, inertia 44.28205692466044
start iteration
done sorting
end inner loop
Iteration 17, inertia 44.26932200376645
start iteration
done sorting
end inner loop
Iteration 18, inertia 44.26227241338869
start iteration
done sorting
end inner loop
Iteration 19, inertia 44.25320396950309
Initialization complete
start iteration
done sorting
end inner loop
Iteration 0, inertia 43.928658135479154
start iteration
done sorting
end inner loop
Iteration 1, inertia 42.57473653938675
start iteration
done sorting
end inner loop
Iteration 2, inertia 41.98895383854935
start iteration
done sorting
end inner loop
Iteration 3, inertia 41.84596444818291
start iteration
done sorting
end inner loop
Iteration 4, inertia 41.82176579417316
start iteration
done sorting
end inner loop
Iteration 5, inertia 41.82176579417316
center shift 0.000000e+00 within tolerance 9.883149e-05
Initialization complete
start iteration
done sorting
```

```
end inner loop
Iteration 0, inertia 46.184209472642074
start iteration
done sorting
end inner loop
Iteration 1, inertia 44.50258341972906
start iteration
done sorting
end inner loop
Iteration 2, inertia 44.40696612606821
start iteration
done sorting
end inner loop
Iteration 3, inertia 44.39180243246967
start iteration
done sorting
end inner loop
Iteration 4, inertia 44.39180243246967
center shift 0.000000e+00 within tolerance 9.883149e-05
Initialization complete
start iteration
done sorting
end inner loop
Iteration 0, inertia 49.69915459221494
start iteration
done sorting
end inner loop
Iteration 1, inertia 44.925568437094306
start iteration
done sorting
end inner loop
Iteration 2, inertia 43.78264742636159
start iteration
done sorting
end inner loop
Iteration 3, inertia 43.24422701541962
start iteration
done sorting
end inner loop
Iteration 4, inertia 43.04851938131404
start iteration
done sorting
end inner loop
Iteration 5, inertia 43.02544696496648
start iteration
done sorting
end inner loop
Iteration 6, inertia 43.000809805108844
```

```
start iteration
done sorting
end inner loop
Iteration 7, inertia 42.99136812295667
center shift 9.705341e-03 within tolerance 9.883149e-05
Initialization complete
start iteration
done sorting
end inner loop
Iteration 0, inertia 45.87082293427251
start iteration
done sorting
end inner loop
Iteration 1, inertia 45.2449161064165
start iteration
done sorting
end inner loop
Iteration 2, inertia 44.94492507073117
start iteration
done sorting
end inner loop
Iteration 3, inertia 44.61431867179136
start iteration
done sorting
end inner loop
Iteration 4, inertia 44.61431867179136
center shift 0.000000e+00 within tolerance 9.883149e-05
Initialization complete
start iteration
done sorting
end inner loop
Iteration 0, inertia 44.710850356454536
start iteration
done sorting
end inner loop
Iteration 1, inertia 43.19821683899244
start iteration
done sorting
end inner loop
Iteration 2, inertia 42.32522414377977
start iteration
done sorting
end inner loop
Iteration 3, inertia 41.95014852288341
start iteration
done sorting
end inner loop
Iteration 4, inertia 41.85223853759469
```

```
start iteration
done sorting
end inner loop
Iteration 5, inertia 41.82803988358493
start iteration
done sorting
end inner loop
Iteration 6, inertia 41.82803988358493
center shift 0.000000e+00 within tolerance 9.883149e-05
Initialization complete
start iteration
done sorting
end inner loop
Iteration 0, inertia 43.72099302268987
start iteration
done sorting
end inner loop
Iteration 1, inertia 42.4495609758474
start iteration
done sorting
end inner loop
Iteration 2, inertia 41.88118409401799
start iteration
done sorting
end inner loop
Iteration 3, inertia 41.816338030367945
start iteration
done sorting
end inner loop
Iteration 4, inertia 41.79994425752656
start iteration
done sorting
end inner loop
Iteration 5, inertia 41.79247565841723
start iteration
done sorting
end inner loop
Iteration 6, inertia 41.79247565841723
center shift 0.000000e+00 within tolerance 9.883149e-05
Initialization complete
start iteration
done sorting
end inner loop
Iteration 0, inertia 48.06690801227036
start iteration
done sorting
end inner loop
Iteration 1, inertia 46.51566654075059
```

```
start iteration
done sorting
end inner loop
Iteration 2, inertia 45.99739240855007
start iteration
done sorting
end inner loop
Iteration 3, inertia 45.88888204944219
start iteration
done sorting
end inner loop
Iteration 4, inertia 45.87595404679715
start iteration
done sorting
end inner loop
Iteration 5, inertia 45.87595404679715
center shift 0.000000e+00 within tolerance 9.883149e-05
Initialization complete
start iteration
done sorting
end inner loop
Iteration 0, inertia 45.32197272381871
start iteration
done sorting
end inner loop
Iteration 1, inertia 43.378158142886946
start iteration
done sorting
end inner loop
Iteration 2, inertia 43.146377005154434
start iteration
done sorting
end inner loop
Iteration 3, inertia 43.05791264497849
start iteration
done sorting
end inner loop
Iteration 4, inertia 43.041580793748516
start iteration
done sorting
end inner loop
Iteration 5, inertia 42.9626708285376
start iteration
done sorting
end inner loop
Iteration 6, inertia 42.94008276295207
center shift 9.118850e-03 within tolerance 9.883149e-05
```

[46]:
```
km_4_ss  = KMeans(n_clusters = 5, max_iter = 20, verbose = 1).
 ↪fit(geyser_scaled_initial)
```

```
Initialization complete
start iteration
done sorting
end inner loop
Iteration 0, inertia 37.399491420663296
start iteration
done sorting
end inner loop
Iteration 1, inertia 34.57644958656211
start iteration
done sorting
end inner loop
Iteration 2, inertia 34.15383136147889
start iteration
done sorting
end inner loop
Iteration 3, inertia 33.79792055521861
start iteration
done sorting
end inner loop
Iteration 4, inertia 33.34652646703114
start iteration
done sorting
end inner loop
Iteration 5, inertia 33.28365589149314
start iteration
done sorting
end inner loop
Iteration 6, inertia 33.22647608296299
start iteration
done sorting
end inner loop
Iteration 7, inertia 33.16071118809525
start iteration
done sorting
end inner loop
Iteration 8, inertia 33.0486231803025
start iteration
done sorting
end inner loop
Iteration 9, inertia 32.9659758880172
start iteration
done sorting
end inner loop
```

```
Iteration 10, inertia 32.9659758880172
center shift 0.000000e+00 within tolerance 9.883149e-05
Initialization complete
start iteration
done sorting
end inner loop
Iteration 0, inertia 46.10421469568994
start iteration
done sorting
end inner loop
Iteration 1, inertia 38.024833895628106
start iteration
done sorting
end inner loop
Iteration 2, inertia 35.4760761414839
start iteration
done sorting
end inner loop
Iteration 3, inertia 34.274141269859044
start iteration
done sorting
end inner loop
Iteration 4, inertia 33.967122207698466
start iteration
done sorting
end inner loop
Iteration 5, inertia 33.724128869927675
start iteration
done sorting
end inner loop
Iteration 6, inertia 33.502779634021834
start iteration
done sorting
end inner loop
Iteration 7, inertia 33.37251170827247
start iteration
done sorting
end inner loop
Iteration 8, inertia 33.20663779693536
start iteration
done sorting
end inner loop
Iteration 9, inertia 33.14607689329255
start iteration
done sorting
end inner loop
Iteration 10, inertia 33.09439881250289
start iteration
```

```
done sorting
end inner loop
Iteration 11, inertia 33.06497183511591
start iteration
done sorting
end inner loop
Iteration 12, inertia 33.053220302365276
start iteration
done sorting
end inner loop
Iteration 13, inertia 33.04031591837729
start iteration
done sorting
end inner loop
Iteration 14, inertia 33.022383767362555
start iteration
done sorting
end inner loop
Iteration 15, inertia 33.012084063049116
start iteration
done sorting
end inner loop
Iteration 16, inertia 33.006162714893115
start iteration
done sorting
end inner loop
Iteration 17, inertia 32.99798177314992
center shift 9.826366e-03 within tolerance 9.883149e-05
Initialization complete
start iteration
done sorting
end inner loop
Iteration 0, inertia 39.403680661012075
start iteration
done sorting
end inner loop
Iteration 1, inertia 37.10514002475424
start iteration
done sorting
end inner loop
Iteration 2, inertia 36.230118107672396
start iteration
done sorting
end inner loop
Iteration 3, inertia 35.76205235259458
start iteration
done sorting
end inner loop
```

```
Iteration 4, inertia 35.31480898427516
start iteration
done sorting
end inner loop
Iteration 5, inertia 34.91698030966555
start iteration
done sorting
end inner loop
Iteration 6, inertia 34.78566860830841
start iteration
done sorting
end inner loop
Iteration 7, inertia 34.78566860830841
center shift 0.000000e+00 within tolerance 9.883149e-05
Initialization complete
start iteration
done sorting
end inner loop
Iteration 0, inertia 36.52981440002759
start iteration
done sorting
end inner loop
Iteration 1, inertia 34.48891281265959
start iteration
done sorting
end inner loop
Iteration 2, inertia 33.53242100419933
start iteration
done sorting
end inner loop
Iteration 3, inertia 32.98178260468936
start iteration
done sorting
end inner loop
Iteration 4, inertia 32.93718885190608
start iteration
done sorting
end inner loop
Iteration 5, inertia 32.93718885190608
center shift 0.000000e+00 within tolerance 9.883149e-05
Initialization complete
start iteration
done sorting
end inner loop
Iteration 0, inertia 42.5473002854821
start iteration
done sorting
end inner loop
```

```
Iteration 1, inertia 35.797294791205616
start iteration
done sorting
end inner loop
Iteration 2, inertia 34.14890123684396
start iteration
done sorting
end inner loop
Iteration 3, inertia 33.35427114338516
start iteration
done sorting
end inner loop
Iteration 4, inertia 33.10280232445153
start iteration
done sorting
end inner loop
Iteration 5, inertia 32.93845214277911
start iteration
done sorting
end inner loop
Iteration 6, inertia 32.92731757607033
start iteration
done sorting
end inner loop
Iteration 7, inertia 32.92731757607033
center shift 0.000000e+00 within tolerance 9.883149e-05
Initialization complete
start iteration
done sorting
end inner loop
Iteration 0, inertia 37.06326856126147
start iteration
done sorting
end inner loop
Iteration 1, inertia 35.08508552563065
start iteration
done sorting
end inner loop
Iteration 2, inertia 34.815407333365705
start iteration
done sorting
end inner loop
Iteration 3, inertia 34.798300151101316
start iteration
done sorting
end inner loop
Iteration 4, inertia 34.78751957492365
start iteration
```

```
done sorting
end inner loop
Iteration 5, inertia 34.78751957492365
center shift 0.000000e+00 within tolerance 9.883149e-05
Initialization complete
start iteration
done sorting
end inner loop
Iteration 0, inertia 42.63743678380038
start iteration
done sorting
end inner loop
Iteration 1, inertia 36.7975131091675
start iteration
done sorting
end inner loop
Iteration 2, inertia 35.50384711356361
start iteration
done sorting
end inner loop
Iteration 3, inertia 35.11077156368909
start iteration
done sorting
end inner loop
Iteration 4, inertia 34.85499676025428
start iteration
done sorting
end inner loop
Iteration 5, inertia 34.75888126314138
start iteration
done sorting
end inner loop
Iteration 6, inertia 34.75888126314138
center shift 0.000000e+00 within tolerance 9.883149e-05
Initialization complete
start iteration
done sorting
end inner loop
Iteration 0, inertia 40.549934053084144
start iteration
done sorting
end inner loop
Iteration 1, inertia 37.35579756235956
start iteration
done sorting
end inner loop
Iteration 2, inertia 36.37526033532895
start iteration
```

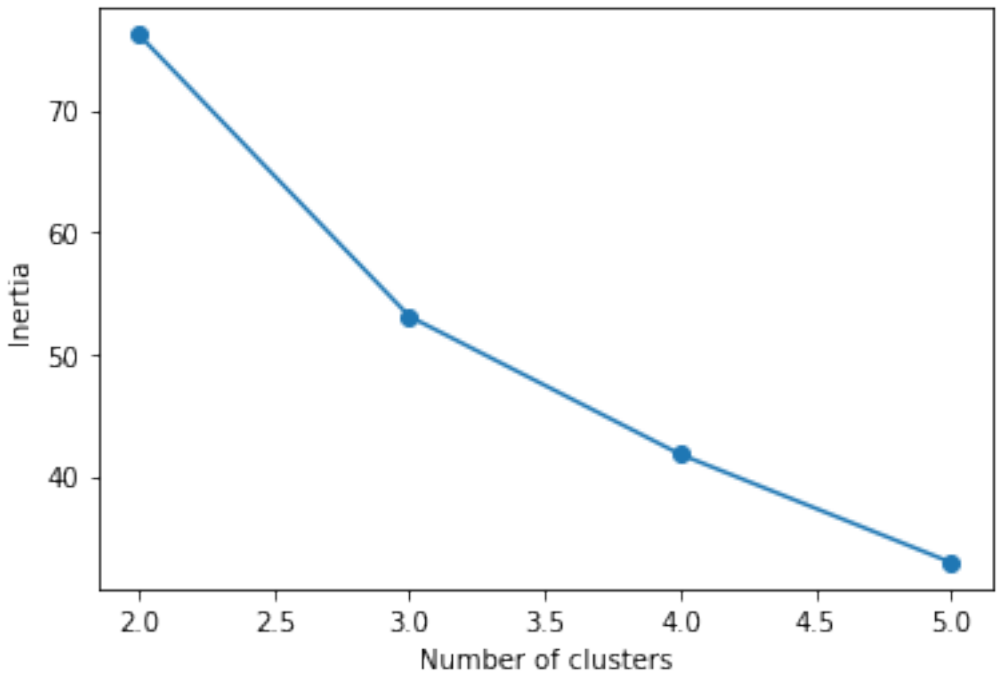
```
done sorting
end inner loop
Iteration 3, inertia 36.23643947478161
start iteration
done sorting
end inner loop
Iteration 4, inertia 36.212932457159205
start iteration
done sorting
end inner loop
Iteration 5, inertia 36.20365509756911
start iteration
done sorting
end inner loop
Iteration 6, inertia 36.20365509756911
center shift 0.000000e+00 within tolerance 9.883149e-05
Initialization complete
start iteration
done sorting
end inner loop
Iteration 0, inertia 39.28753016369022
start iteration
done sorting
end inner loop
Iteration 1, inertia 38.2458932796004
start iteration
done sorting
end inner loop
Iteration 2, inertia 37.60087277804972
start iteration
done sorting
end inner loop
Iteration 3, inertia 37.098624988158775
start iteration
done sorting
end inner loop
Iteration 4, inertia 36.91964179528218
start iteration
done sorting
end inner loop
Iteration 5, inertia 36.88407771745328
start iteration
done sorting
end inner loop
Iteration 6, inertia 36.81329816874053
start iteration
done sorting
end inner loop
```

```
Iteration 7, inertia 36.6324445594442
start iteration
done sorting
end inner loop
Iteration 8, inertia 36.311312569354314
start iteration
done sorting
end inner loop
Iteration 9, inertia 35.909995474763605
start iteration
done sorting
end inner loop
Iteration 10, inertia 35.49884455091038
start iteration
done sorting
end inner loop
Iteration 11, inertia 35.117429284446786
start iteration
done sorting
end inner loop
Iteration 12, inertia 34.5497800915533
start iteration
done sorting
end inner loop
Iteration 13, inertia 34.19985688147621
start iteration
done sorting
end inner loop
Iteration 14, inertia 34.11213659084957
start iteration
done sorting
end inner loop
Iteration 15, inertia 34.11213659084957
center shift 0.000000e+00 within tolerance 9.883149e-05
Initialization complete
start iteration
done sorting
end inner loop
Iteration 0, inertia 40.704142132536816
start iteration
done sorting
end inner loop
Iteration 1, inertia 35.37392025808401
start iteration
done sorting
end inner loop
Iteration 2, inertia 33.46363578456407
start iteration
```

```
done sorting
end inner loop
Iteration 3, inertia 33.160948205580254
start iteration
done sorting
end inner loop
Iteration 4, inertia 33.09275272353049
start iteration
done sorting
end inner loop
Iteration 5, inertia 33.05251395889687
start iteration
done sorting
end inner loop
Iteration 6, inertia 33.022976979223245
start iteration
done sorting
end inner loop
Iteration 7, inertia 33.00765914093242
start iteration
done sorting
end inner loop
Iteration 8, inertia 32.98335856982964
start iteration
done sorting
end inner loop
Iteration 9, inertia 32.97429012594404
start iteration
done sorting
end inner loop
Iteration 10, inertia 32.97429012594404
center shift 0.000000e+00 within tolerance 9.883149e-05
```

[47]:
```
inertias = [km_1_ss.inertia_, km_2_ss.inertia_,km_3_ss.inertia_,km_4_ss.
 ↪inertia_]
clusters = [2,3,4,5]
```

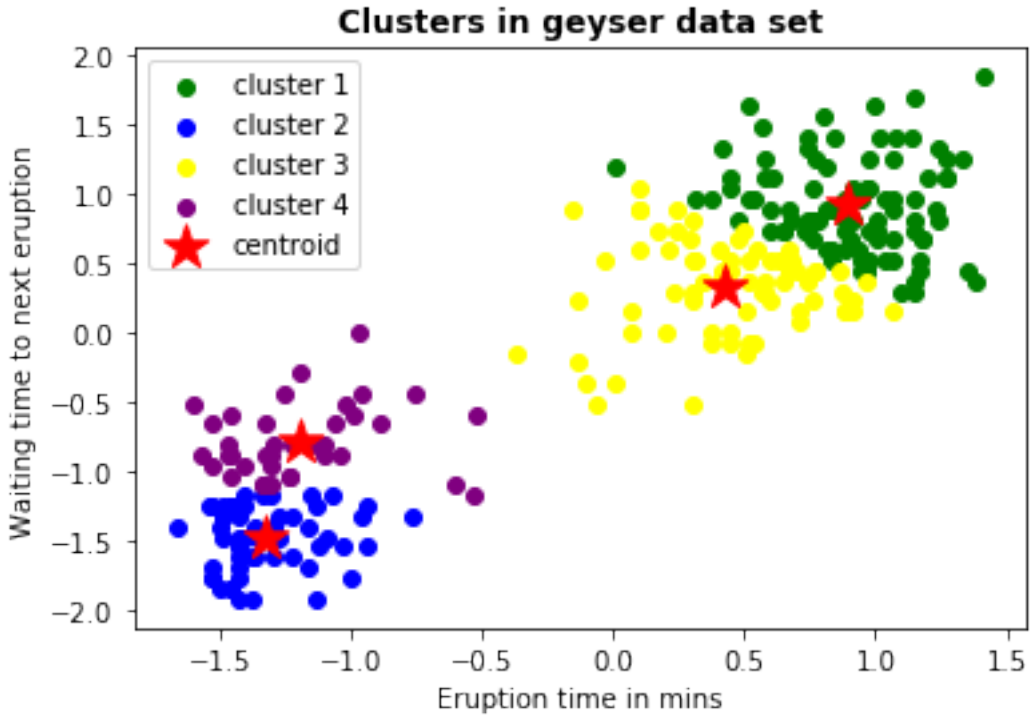
[48]:
```
plt.plot(clusters,inertias,
        marker = 'o')
plt.xlabel('Number of clusters')
plt.ylabel('Inertia')
plt.show()
```



[49]:
```
#first cluster
plt.scatter(x = geyser_scaled_initial.iloc[geyser_scaled_initial.index[km_3_ss.
 ↪labels_ == 0].tolist() , [0]],
            y = geyser_scaled_initial.iloc[geyser_scaled_initial.index[km_3_ss.
 ↪labels_ == 0].tolist() , [1]],
            c = 'green',
            label = 'cluster 1')

#second cluster
plt.scatter(x = geyser_scaled_initial.iloc[geyser_scaled_initial.index[km_3_ss.
 ↪labels_ == 1].tolist() , [0]],
            y = geyser_scaled_initial.iloc[geyser_scaled_initial.index[km_3_ss.
 ↪labels_ == 1].tolist() , [1]],
            c = 'blue',
            label = 'cluster 2')
#third cluster
plt.scatter(x = geyser_scaled_initial.iloc[geyser_scaled_initial.index[km_3_ss.
 ↪labels_ == 2].tolist() , [0]],
            y = geyser_scaled_initial.iloc[geyser_scaled_initial.index[km_3_ss.
 ↪labels_ == 2].tolist() , [1]],
            c = 'yellow',
            label = 'cluster 3')

#fourth cluster
```

```
plt.scatter(x = geyser_scaled_initial.iloc[geyser_scaled_initial.index[km_3_ss.
 ↪labels_ == 3].tolist() , [0]],
            y = geyser_scaled_initial.iloc[geyser_scaled_initial.index[km_3_ss.
 ↪labels_ == 3].tolist() , [1]],
            c = 'purple',
            label = 'cluster 4')



#centroid
plt.scatter(x = km_3_ss.cluster_centers_[:, [0]],
             y = km_3_ss.cluster_centers_[:, [1]],
             marker = '*', s = 300, label = 'centroid', c = 'r')

plt.legend()
plt.xlabel('Eruption time in mins')
plt.ylabel('Waiting time to next eruption')
plt.title('Clusters in geyser data set', fontweight = 'bold')
```

[49]:
```
Text(0.5, 1.0, 'Clusters in geyser data set')
```



[50]:
```
km_3_ss.cluster_centers_
```

[50]:
```
array([[ 0.89758426,  0.91952675],
       [-1.32263294, -1.47563674],
```

```
       [ 0.42979541,  0.32830265],
       [-1.19448041, -0.77315286]])
```

[52]:
```
from matplotlib import cm
from sklearn.metrics import silhouette_samples
import numpy as np
```

[55]:
```
km_1_ss_sl = KMeans(n_clusters = 3, n_init = 10, max_iter = 20, random_state =␣
 ↪0)
pred_km_1_ss_sl = km_1_ss_sl.fit_predict(geyser_scaled_initial)
```

[56]:
```
pred_km_1_ss_sl
```

[56]:
```
array([2, 1, 2, 1, 0, 1, 0, 2, 1, 0, 1, 0, 2, 1, 0, 1, 1, 0, 1, 2, 1, 1,
       2, 2, 2, 2, 1, 2, 2, 0, 2, 2, 2, 2, 2, 1, 1, 0, 1, 0, 0, 1, 0, 1,
       2, 2, 2, 1, 0, 1, 0, 0, 1, 0, 1, 0, 2, 1, 0, 0, 1, 0, 1, 0, 1, 0,
       2, 0, 1, 2, 0, 1, 0, 2, 1, 0, 1, 0, 2, 2, 2, 0, 2, 1, 2, 0, 2, 0,
       1, 0, 1, 0, 1, 0, 1, 2, 0, 2, 1, 0, 1, 0, 1, 0, 2, 1, 0, 1, 0, 2,
       0, 1, 0, 0, 1, 0, 1, 0, 1, 0, 1, 2, 2, 1, 0, 2, 1, 0, 1, 0, 1, 0,
       1, 0, 1, 0, 1, 0, 1, 2, 0, 1, 0, 0, 2, 1, 0, 1, 0, 1, 0, 2, 1, 0,
       2, 2, 0, 0, 1, 0, 1, 0, 1, 2, 2, 2, 1, 0, 1, 0, 1, 1, 0, 2, 0, 0,
       2, 1, 0, 2, 1, 0, 0, 2, 1, 0, 0, 1, 0, 1, 0, 1, 0, 0, 2, 0, 0, 2,
       1, 0, 1, 0, 0, 1, 0, 1, 2, 0, 1, 0, 1, 0, 1, 2, 2, 2, 1, 0, 1, 2,
       1, 0, 1, 2, 2, 2, 2, 2, 2, 0, 2, 1, 0, 1, 0, 1, 1, 2, 2, 1, 2, 1,
       0, 1, 0, 2, 1, 0, 1, 2, 1, 0, 2, 2, 0, 2, 2, 0, 1])
```

[64]:
```
cluster_labels = np.unique(pred_km_1_ss_sl)
n_clusters = cluster_labels.shape[0]
silhouette_vals = silhouette_samples(geyser_scaled_initial,
                                     pred_km_1_ss_sl,
                                    metric = 'euclidean')
```

[65]:
```
silhouette_vals
```

[65]:
```
array([ 0.37646427,  0.81828654,  0.44830146,  0.61947171,  0.56414232,
        0.5225464 ,  0.54792125,  0.02194472,  0.81662274,  0.50330386,
        0.82136003,  0.12425949,  0.22056058,  0.76318746,  0.55105653,
        0.80115174,  0.68061863,  0.55200394,  0.78794479,  0.02208882,
        0.81113237,  0.76318746,  0.39615537,  0.29297049,  0.19266684,
        0.15022271,  0.8186915 ,  0.44996734,  0.41923049,  0.20692772,
        0.40618465, -0.00671657,  0.35871371,  0.15592371,  0.53239677,
        0.81942025,  0.78286595,  0.43263641,  0.75978379,  0.5133258 ,
        0.23481369,  0.78078367,  0.56420757,  0.77123731,  0.23817794,
        0.19432631,  0.41208431,  0.81352474,  0.52404024,  0.75890087,
        0.08391952,  0.51673021,  0.82136003,  0.43263641,  0.80979146,
        0.52067287,  0.51984906,  0.61842454,  0.10343727,  0.30713518,
        0.72195447,  0.5516564 ,  0.77638464,  0.51455591,  0.736316  ,
        0.4509364 ,  0.25426277,  0.28766752,  0.5552814 ,  0.10292206,
        0.06992473,  0.8099941 ,  0.26939518,  0.51551294,  0.67674644,
        0.21892313,  0.73310805,  0.21551613,  0.49892504,  0.15022271,
```

```
 0.46132994,  0.4005692 ,  0.4827436 ,  0.31805592,  0.5093002 ,
 0.52064595,  0.49090103,  0.37514953,  0.76277747,  0.29168029,
 0.70476865,  0.46958491,  0.80729067,  0.31439928,  0.65175613,
 0.41419541,  0.56789122,  0.51596736,  0.81584046,  0.49466512,
 0.5418585 ,  0.50470829,  0.78515732,  0.53398082,  0.00968625,
 0.76812372,  0.56584302,  0.81494464,  0.54848533,  0.25933586,
 0.04140711,  0.70345536,  0.51503203,  0.18934139,  0.74777726,
 0.48079672,  0.7575734 ,  0.57184742,  0.75888307,  0.52808193,
 0.66821005,  0.47462084,  0.26801285,  0.8099941 ,  0.54674474,
 0.23412017,  0.73410361,  0.49752677,  0.77467283,  0.51670778,
 0.73478916,  0.30679664,  0.5493941 ,  0.48317362,  0.75181665,
 0.43780479,  0.81643695,  0.53171762,  0.8206504 ,  0.37594313,
 0.21259636,  0.69684321,  0.50789495,  0.24362433,  0.26223117,
 0.75991671,  0.42112204,  0.79291246,  0.39311533,  0.81946884,
 0.26819977,  0.43690284,  0.4502504 ,  0.475099  ,  0.50169785,
 0.50350106,  0.44122247,  0.3534682 ,  0.81946884,  0.31800693,
 0.71081866,  0.39971537,  0.77900212,  0.42150971,  0.43515585,
-0.02644238,  0.55275445,  0.50698422,  0.82264961,  0.4570531 ,
 0.7968634 ,  0.78791832,  0.11874043,  0.43339905,  0.1265723 ,
 0.32290318,  0.26501128,  0.73138798,  0.2564097 ,  0.46523536,
 0.81825468,  0.11874043,  0.38426652,  0.0686762 ,  0.81226219,
 0.07576154,  0.28357029,  0.75181665,  0.50082262,  0.79344259,
 0.48021222,  0.7937924 ,  0.16172063,  0.299259  ,  0.4477103 ,
 0.21259636,  0.01313202,  0.13208061,  0.78017245,  0.27480954,
 0.72336418,  0.41425155,  0.39018553,  0.82410362,  0.23894348,
 0.74992745,  0.1307685 ,  0.06195255,  0.79668642,  0.53398082,
 0.13743129,  0.43231153,  0.79609187,  0.52162488,  0.28070766,
 0.35569667,  0.74622338,  0.44046987,  0.81717117,  0.41496785,
 0.80729067,  0.34018588,  0.81221844,  0.17721992,  0.37348671,
 0.19149503,  0.32426966,  0.14042581,  0.51193054,  0.30520785,
 0.48722943,  0.73750328,  0.42084294,  0.77872826,  0.49646233,
 0.82376926,  0.82243878,  0.23281055,  0.31778298,  0.52384812,
 0.45296029,  0.71537412,  0.53171762,  0.22481922,  0.57072584,
 0.11155985,  0.78791832,  0.42688417,  0.45293268,  0.34930727,
 0.79549387,  0.51649711,  0.5044546 ,  0.26501128,  0.41328064,
 0.29304708,  0.52427948,  0.51649711,  0.80848012])
```

[74]:
```
y_ax_lower = 0
y_ax_upper = 0
yticks = []
```

[75]:
```
for i, c in enumerate(cluster_labels):
    c_silhouette_vals = silhouette_vals[pred_km_1_ss_sl == c]
    c_silhouette_vals.sort()
    y_ax_upper = y_ax_upper + len(c_silhouette_vals)
    color = cm.jet(i/n_clusters)
    plt.barh(range(y_ax_upper, y_ax_lower),
             c_silhouette_vals,
```

```
            height = 1.0,
            edgecolor = 'none',
            color = color)
    yticks.append((y_ax_lower + y_ax_upper)/2)
    y_ax_lower = y_ax_lower + len(c_silhouette_vals)


silhouette_avg = np.mean(silhouette_vals)
plt.axvline(silhouette_avg,
           color = 'red',
           linestyle = '--')
plt.yticks(yticks, cluster_labels + 1)

plt.ylabel('Cluster')
plt.xlabel('Silhouette coefficent')
plt.show()
```

```
      ␣
↪---------------------------------------------------------------------------

       ValueError                                Traceback (most recent call␣
↪last)

       <ipython-input-75-bfa4aec22308> in <module>
         8             height = 1.0,
         9             edgecolor = 'none',
   ---> 10             color = color)
        11     yticks.append((y_ax_lower + y_ax_upper)/2)
        12     y_ax_lower = y_ax_lower + len(c_silhouette_vals)


       C:\ProgramData\Anaconda\lib\site-packages\matplotlib\pyplot.py in␣
↪barh(y, width, height, left, align, **kwargs)
      2452 def barh(y, width, height=0.8, left=None, *, align='center',␣
↪**kwargs):
      2453     return gca().barh(
   -> 2454         y, width, height=height, left=left, align=align, **kwargs)
      2455
      2456


       C:\ProgramData\Anaconda\lib\site-packages\matplotlib\axes\_axes.py in␣
↪barh(self, y, width, height, left, align, **kwargs)
      2595         kwargs.setdefault('orientation', 'horizontal')
      2596         patches = self.bar(x=left, height=height, width=width,␣
↪bottom=y,
```

```
   -> 2597                             align=align, **kwargs)
      2598         return patches
      2599
```

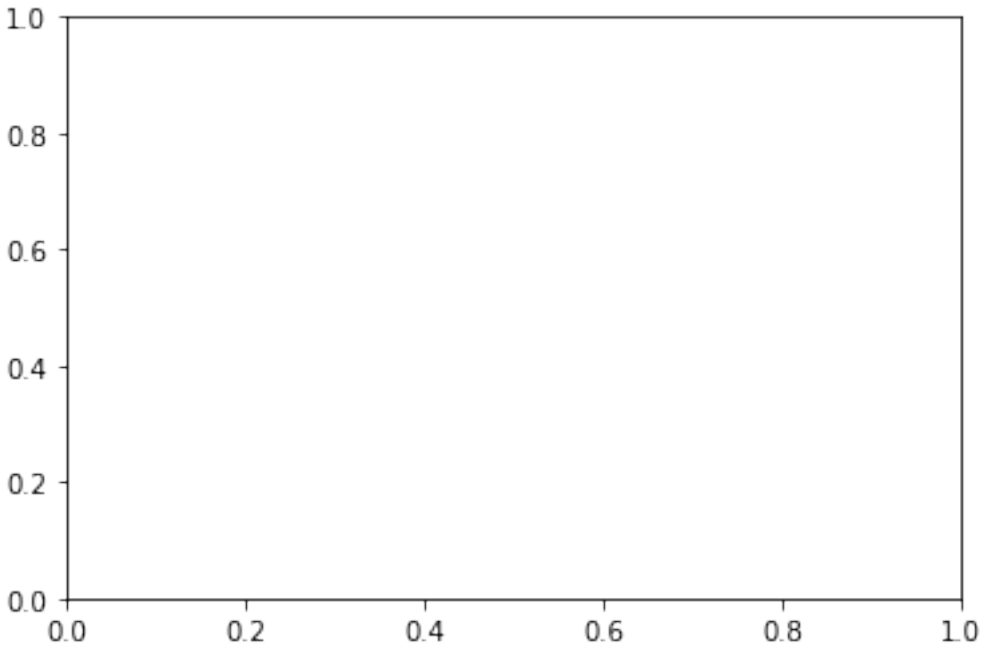
```
      C:\ProgramData\Anaconda\lib\site-packages\matplotlib\__init__.py in␣
↪inner(ax, data, *args, **kwargs)
      1599     def inner(ax, *args, data=None, **kwargs):
      1600         if data is None:
   -> 1601             return func(ax, *map(sanitize_sequence, args), **kwargs)
      1602
      1603         bound = new_sig.bind(ax, *args, **kwargs)
```

```
      C:\ProgramData\Anaconda\lib\site-packages\matplotlib\axes\_axes.py in␣
↪bar(self, x, height, width, bottom, align, **kwargs)
      2373         x, height, width, y, linewidth = np.broadcast_arrays(
      2374             # Make args iterable too.
   -> 2375             np.atleast_1d(x), height, width, y, linewidth)
      2376
      2377         # Now that units have been converted, set the tick locations.
```

```
      C:\ProgramData\Anaconda\lib\site-packages\numpy\lib\stride_tricks.py in␣
↪broadcast_arrays(*args, **kwargs)
       257     args = [np.array(_m, copy=False, subok=subok) for _m in args]
       258
   --> 259     shape = _broadcast_shape(*args)
       260
       261     if all(array.shape == shape for array in args):
```

```
      C:\ProgramData\Anaconda\lib\site-packages\numpy\lib\stride_tricks.py in␣
↪_broadcast_shape(*args)
       191     # use the old-iterator because np.nditer does not handle size 0␣
↪arrays
       192     # consistently
   --> 193     b = np.broadcast(*args[:32])
       194     # unfortunately, it cannot handle 32 or more arguments directly
       195     for pos in range(32, len(args), 31):
```

```
      ValueError: shape mismatch: objects cannot be broadcast to a single shape
```



[73]:
```
c_silhouette_vals
```

[73]:
```
array([0.01313202, 0.04140711, 0.06195255, 0.06992473, 0.07576154,
       0.08391952, 0.10343727, 0.11874043, 0.11874043, 0.12425949,
       0.1265723 , 0.16172063, 0.18934139, 0.20692772, 0.21259636,
       0.21259636, 0.21551613, 0.21892313, 0.23481369, 0.23894348,
       0.24362433, 0.2564097 , 0.26819977, 0.26939518, 0.27480954,
       0.28357029, 0.28766752, 0.29168029, 0.299259  , 0.30520785,
       0.30679664, 0.30713518, 0.31439928, 0.31800693, 0.32290318,
       0.34018588, 0.3534682 , 0.37514953, 0.38426652, 0.39018553,
       0.39311533, 0.39971537, 0.4005692 , 0.41328064, 0.41425155,
       0.42084294, 0.42112204, 0.42688417, 0.43231153, 0.43263641,
       0.43263641, 0.43780479, 0.44046987, 0.44122247, 0.4509364 ,
       0.4570531 , 0.46958491, 0.475099  , 0.48021222, 0.48079672,
       0.48317362, 0.49466512, 0.49646233, 0.49752677, 0.50082262,
       0.50330386, 0.50470829, 0.50698422, 0.50789495, 0.5133258 ,
       0.51455591, 0.51503203, 0.51649711, 0.51649711, 0.51670778,
       0.51673021, 0.52064595, 0.52067287, 0.52404024, 0.52808193,
       0.53171762, 0.53171762, 0.53398082, 0.53398082, 0.54674474,
       0.54792125, 0.54848533, 0.55105653, 0.5516564 , 0.55200394,
       0.56414232, 0.56420757, 0.56584302, 0.56789122, 0.57072584,
       0.57184742])
```

[82]:
```
!pip install mglearn
```

```
Collecting mglearn
  Downloading https://files.pythonhosted.org/packages/fb/01/8d3630ecc767c9de96a9
```

```
c46e055f2a3a5f9e14a47d3d0348a36a5005fe67/mglearn-0.1.7.tar.gz (540kB)
Requirement already satisfied: numpy in c:\programdata\anaconda\lib\site-
packages (from mglearn) (1.16.5)
Requirement already satisfied: matplotlib in c:\programdata\anaconda\lib\site-
packages (from mglearn) (3.1.1)
Requirement already satisfied: scikit-learn in c:\programdata\anaconda\lib\site-
packages (from mglearn) (0.21.2)
Requirement already satisfied: pandas in c:\programdata\anaconda\lib\site-
packages (from mglearn) (0.25.1)
Requirement already satisfied: pillow in c:\programdata\anaconda\lib\site-
packages (from mglearn) (6.1.0)
Requirement already satisfied: cycler in c:\programdata\anaconda\lib\site-
packages (from mglearn) (0.10.0)
Requirement already satisfied: imageio in c:\programdata\anaconda\lib\site-
packages (from mglearn) (2.5.0)
Requirement already satisfied: kiwisolver>=1.0.1 in
c:\programdata\anaconda\lib\site-packages (from matplotlib->mglearn) (1.1.0)
Requirement already satisfied: pyparsing!=2.0.4,!=2.1.2,!=2.1.6,>=2.0.1 in
c:\programdata\anaconda\lib\site-packages (from matplotlib->mglearn) (2.4.2)
Requirement already satisfied: python-dateutil>=2.1 in
c:\programdata\anaconda\lib\site-packages (from matplotlib->mglearn) (2.8.0)
Requirement already satisfied: scipy>=0.17.0 in
c:\programdata\anaconda\lib\site-packages (from scikit-learn->mglearn) (1.3.1)
Requirement already satisfied: joblib>=0.11 in c:\programdata\anaconda\lib\site-
packages (from scikit-learn->mglearn) (0.13.2)
Requirement already satisfied: pytz>=2017.2 in c:\programdata\anaconda\lib\site-
packages (from pandas->mglearn) (2019.2)
Requirement already satisfied: six in c:\programdata\anaconda\lib\site-packages
(from cycler->mglearn) (1.12.0)
Requirement already satisfied: setuptools in c:\programdata\anaconda\lib\site-
packages (from kiwisolver>=1.0.1->matplotlib->mglearn) (41.0.1)
Building wheels for collected packages: mglearn
  Building wheel for mglearn (setup.py): started
  Building wheel for mglearn (setup.py): finished with status 'done'
  Created wheel for mglearn: filename=mglearn-0.1.7-py2.py3-none-any.whl
size=582710
sha256=156c29477908ff96e1e0d8baf0297d2ae9fc63af88f34be1f0367f2bd4b56678
  Stored in directory: C:\Users\masif\AppData\Local\pip\Cache\wheels\74\cf\8d\04
f4932d15854a36726c6210763c7127e62de28f5c8ddfcf3b
Successfully built mglearn
Installing collected packages: mglearn
Successfully installed mglearn-0.1.7
```

[84]:
```
from sklearn.datasets import make_moons
import mglearn
```

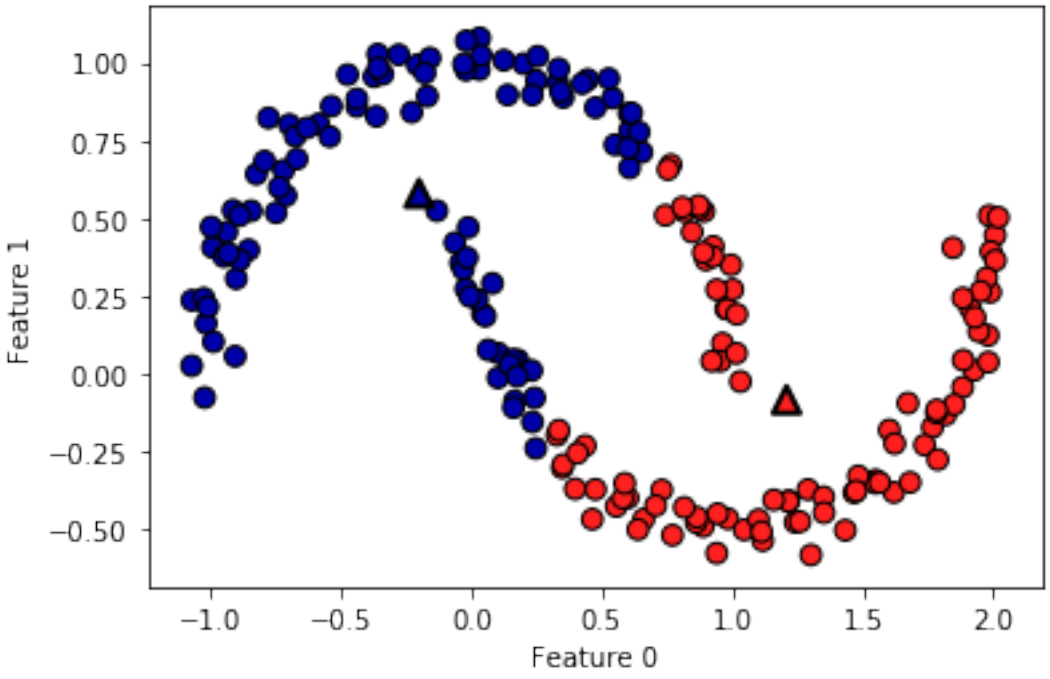
```
C:\ProgramData\Anaconda\lib\site-packages\sklearn\externals\six.py:31:
DeprecationWarning: The module is deprecated in version 0.21 and will be removed
```

```
in version 0.23 since we've dropped support for Python 2.7. Please rely on the
official version of six (https://pypi.org/project/six/).
  "(https://pypi.org/project/six/).", DeprecationWarning)
C:\ProgramData\Anaconda\lib\site-
packages\sklearn\externals\joblib\__init__.py:15: DeprecationWarning:
sklearn.externals.joblib is deprecated in 0.21 and will be removed in 0.23.
Please import this functionality directly from joblib, which can be installed
with: pip install joblib. If this warning is raised when loading pickled models,
you may need to re-serialize those models with scikit-learn 0.21+.
  warnings.warn(msg, category=DeprecationWarning)
```

[93]:
```
x,y = make_moons(n_samples = 200, noise = 0.05, random_state = 0)
```

[87]:
```
km_m = KMeans(n_clusters = 2)
km_m.fit(x)
y_pred = km_m.predict(x)

plt.scatter(x [:, 0], x[:,1], c = y_pred,
            cmap = mglearn.cm2,
            s=60,
            edgecolor = 'k')
plt.scatter(km_m.cluster_centers_[:,0],
            km_m.cluster_centers_[:,1],
            marker = '^',
            c = [mglearn.cm2(0), mglearn.cm2(1)], s = 100, linewidth = 2,
                edgecolor = 'k')

plt.xlabel('Feature 0')
plt.ylabel('Feature 1')
```

[87]:
```
Text(0, 0.5, 'Feature 1')
```



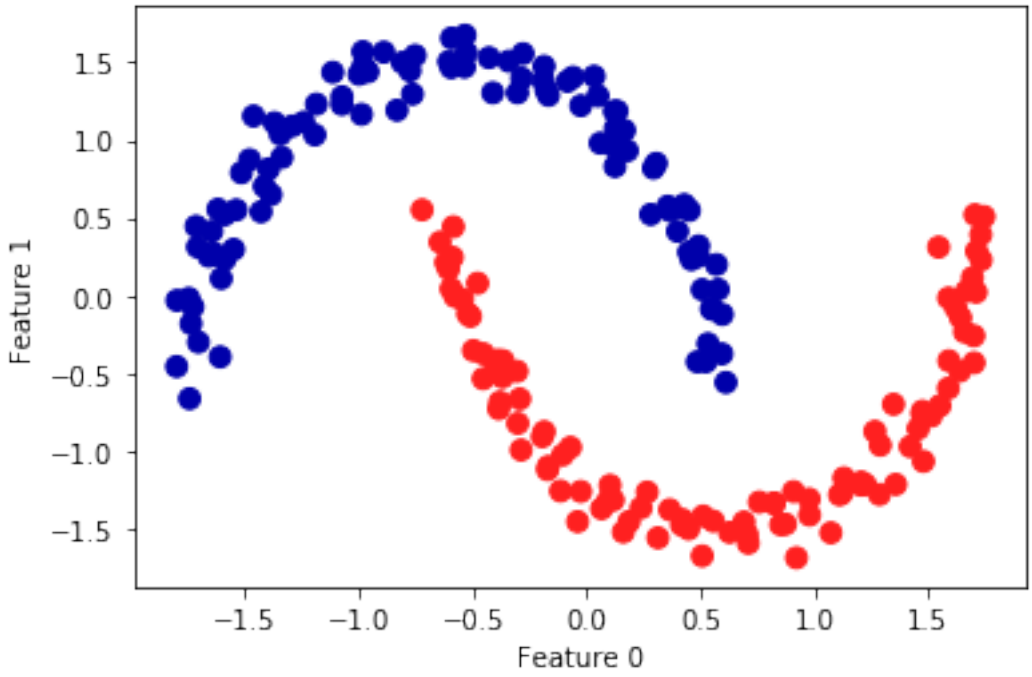
[88]:
```
from sklearn.cluster import DBSCAN
```

[97]:
```
dbscan = DBSCAN()

scaler = ss()
scaler.fit(x)
x_scaled = scaler.transform(x)
```

[101]:
```
cluster = dbscan.fit_predict(x_scaled)
```

[102]:
```
plt.scatter(x_scaled[:,0], x_scaled[:,1],
            c = cluster,
            cmap = mglearn.cm2,
            s=60)

plt.xlabel('Feature 0')
plt.ylabel('Feature 1')
```

[102]:
```
Text(0, 0.5, 'Feature 1')
```



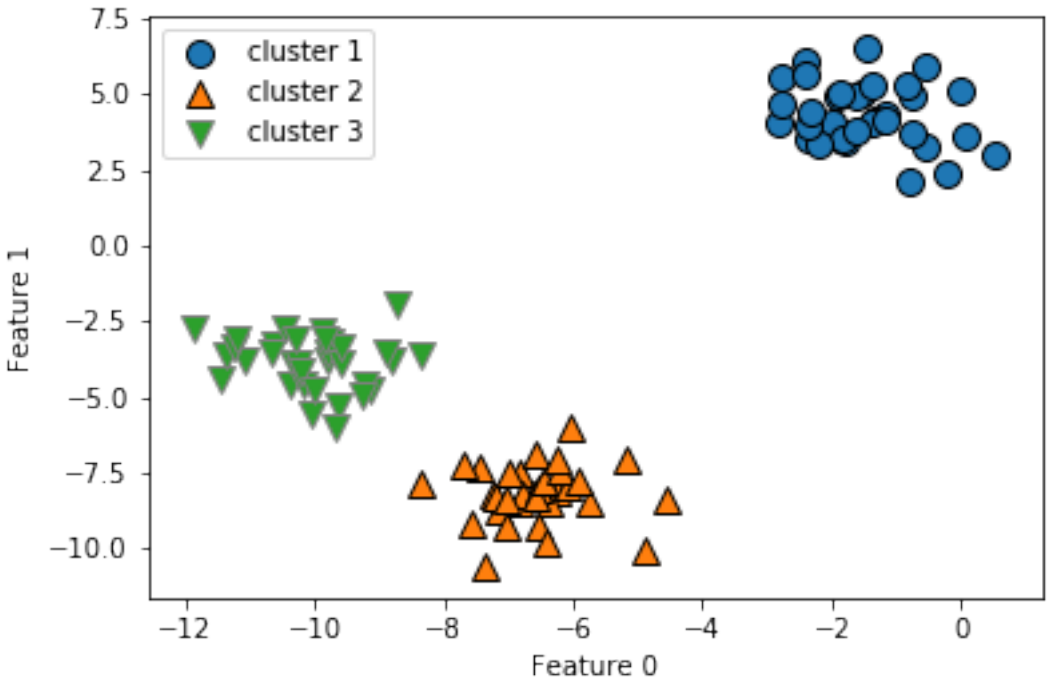
[107]:
```
from sklearn.cluster import AgglomerativeClustering
from sklearn.datasets import make_blobs
```

[108]:
```
x,y = make_blobs(random_state = 1)
```

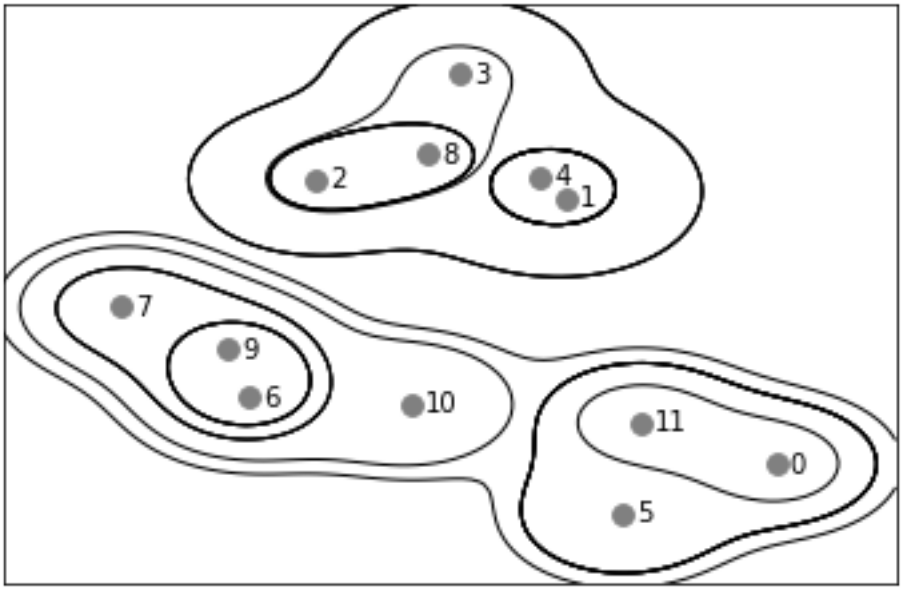
[109]:
```
agg = AgglomerativeClustering(n_clusters = 3)
assignment = agg.fit_predict(x)
```

[113]:
```
mglearn.discrete_scatter(x[:, 0], x[:,1], assignment)
plt.legend(['cluster 1', 'cluster 2', 'cluster 3'], loc = 'best')
plt.xlabel('Feature 0')
plt.ylabel('Feature 1')
```

[113]:
```
Text(0, 0.5, 'Feature 1')
```



[114]:
```
mglearn.plots.plot_agglomerative()
```



[119]:
```
from scipy.cluster.hierarchy import dendrogram, ward
```

[120]:
```
linkage_array = ward(x)
```

[124]:
```
dendrogram(linkage_array)
ax = plt.gca()
bounds = ax.get_xbound()
```