K-Mean cluster

Date: 29/05/2020 Name: D.Saravanan

Program:

```
#!/usr/bin/env python3
import numpy as np
import pandas as pd

# read csv file and create dataframe
df = pd.read_csv("objects.csv")

# randomly select rows of size = 2
dfs = df.sample(n=2, axis=0, replace=False)
df1 = dfs.iloc[0]; df2 = dfs.iloc[1]

df1x = df1.x; df1y = df1.y; df1z = df1.z
df2x = df2.x; df2y = df2.y; df2z = df2.z

print("random 1 values: x:{} y:{} z:{}".format(df1x,df1y,df1z))
print("random 2 values: x:{} y:{} z:{}".format(df2x,df2y,df2z))

m = 0
while m < 4:

    D1=[sum([abs(df1x-df.x[n]),abs(df1y-df.y[n]),abs(df1z-df.z[n])]) for n in range(len(df))]
    D2=[sum([abs(df2x-df.x[n]),abs(df2y-df.y[n]),abs(df2z-df.z[n])]) for n in range(len(df))]

    # adding columns with values of D1 and D2 to dataframe
    df['Dist. from C1({:.1f},{:.1f},{:.1f})'.format(df1x,df1y,df1z)] = D1
    df['Dist. from C2({:.1f},{:.1f},{:.1f})'.format(df2x,df2y,df2z)] = D2

    print("\nk-mean cluster: \n{}".format(df))
    df = df[["Objects", "x", "y", "z"]]

    cluster_1 = []
    cluster_2 = []
    for n in range(len(df)):
        if D1[n] < D2[n]: cluster_1.append(df.iloc[n])
        else: cluster_2.append(df.iloc[n])

    c1 = pd.DataFrame(cluster_1)
    c2 = pd.DataFrame(cluster_2)

    print("\nCluster 1: \n{}".format(c1))
    df1x = np.mean(c1.x); df1y = np.mean(c1.y); df1z = np.mean(c1.z)
    print("\nCluster 1: x_mean = {:.1f}, y_mean = {:.1f}, z_mean = {:.1f}".format(df1x, df1y,
    df1z))

    print("\nCluster 2: \n{}".format(c2))
    df2x = np.mean(c2.x); df2y = np.mean(c2.y); df2z = np.mean(c2.z)
    print("\nCluster 2: x_mean = {:.1f}, y_mean = {:.1f}, z_mean = {:.1f}".format(df2x, df2y,
    df2z))

    print("
    ----------------------------------------------------------------------------------------")

    m = m + 1
```

Output:

```
random 1 values: x:1 y:1 z:1
random 2 values: x:2 y:1 z:1

k-mean cluster:
  Objects  x  y  z  Dist. from C1(1.0,1.0,1.0)  Dist. from C2(2.0,1.0,1.0)
0    OB-1  1  4  1                           3                           4
1    OB-2  1  2  2                           2                           3
2    OB-3  1  4  2                           4                           5
3    OB-4  2  1  2                           2                           1
4    OB-5  1  1  1                           0                           1
5    OB-6  2  4  2                           5                           4
6    OB-7  1  1  2                           1                           2
7    OB-8  2  1  1                           1                           0

Cluster 1:
  Objects  x  y  z
0    OB-1  1  4  1
1    OB-2  1  2  2
2    OB-3  1  4  2
4    OB-5  1  1  1
6    OB-7  1  1  2

Cluster 1: x_mean = 1.0, y_mean = 2.4, z_mean = 1.6

Cluster 2:
  Objects  x  y  z
3    OB-4  2  1  2
5    OB-6  2  4  2
7    OB-8  2  1  1

Cluster 2: x_mean = 2.0, y_mean = 2.0, z_mean = 1.7
--------------------------------------------------------------------------------

k-mean cluster:
  Objects  x  y  z  Dist. from C1(1.0,2.4,1.6)  Dist. from C2(2.0,2.0,1.7)
0    OB-1  1  4  1                         2.2                    3.666667
1    OB-2  1  2  2                         0.8                    1.333333
2    OB-3  1  4  2                         2.0                    3.333333
3    OB-4  2  1  2                         2.8                    1.333333
4    OB-5  1  1  1                         2.0                    2.666667
5    OB-6  2  4  2                         3.0                    2.333333
6    OB-7  1  1  2                         1.8                    2.333333
7    OB-8  2  1  1                         3.0                    1.666667

Cluster 1:
  Objects  x  y  z
0    OB-1  1  4  1
1    OB-2  1  2  2
2    OB-3  1  4  2
4    OB-5  1  1  1
6    OB-7  1  1  2

Cluster 1: x_mean = 1.0, y_mean = 2.4, z_mean = 1.6

Cluster 2:
  Objects  x  y  z
3    OB-4  2  1  2
5    OB-6  2  4  2
7    OB-8  2  1  1

Cluster 2: x_mean = 2.0, y_mean = 2.0, z_mean = 1.7
```

```
--------------------------------------------------------------------------------

k-mean cluster:
  Objects  x  y  z  Dist. from C1(1.0,2.4,1.6)  Dist. from C2(2.0,2.0,1.7)
0    OB-1  1  4  1                         2.2                   3.666667
1    OB-2  1  2  2                         0.8                   1.333333
2    OB-3  1  4  2                         2.0                   3.333333
3    OB-4  2  1  2                         2.8                   1.333333
4    OB-5  1  1  1                         2.0                   2.666667
5    OB-6  2  4  2                         3.0                   2.333333
6    OB-7  1  1  2                         1.8                   2.333333
7    OB-8  2  1  1                         3.0                   1.666667

Cluster 1:
  Objects  x  y  z
0    OB-1  1  4  1
1    OB-2  1  2  2
2    OB-3  1  4  2
4    OB-5  1  1  1
6    OB-7  1  1  2

Cluster 1: x_mean = 1.0, y_mean = 2.4, z_mean = 1.6

Cluster 2:
  Objects  x  y  z
3    OB-4  2  1  2
5    OB-6  2  4  2
7    OB-8  2  1  1

Cluster 2: x_mean = 2.0, y_mean = 2.0, z_mean = 1.7
--------------------------------------------------------------------------------

k-mean cluster:
  Objects  x  y  z  Dist. from C1(1.0,2.4,1.6)  Dist. from C2(2.0,2.0,1.7)
0    OB-1  1  4  1                         2.2                   3.666667
1    OB-2  1  2  2                         0.8                   1.333333
2    OB-3  1  4  2                         2.0                   3.333333
3    OB-4  2  1  2                         2.8                   1.333333
4    OB-5  1  1  1                         2.0                   2.666667
5    OB-6  2  4  2                         3.0                   2.333333
6    OB-7  1  1  2                         1.8                   2.333333
7    OB-8  2  1  1                         3.0                   1.666667

Cluster 1:
  Objects  x  y  z
0    OB-1  1  4  1
1    OB-2  1  2  2
2    OB-3  1  4  2
4    OB-5  1  1  1
6    OB-7  1  1  2

Cluster 1: x_mean = 1.0, y_mean = 2.4, z_mean = 1.6

Cluster 2:
  Objects  x  y  z
3    OB-4  2  1  2
5    OB-6  2  4  2
7    OB-8  2  1  1

Cluster 2: x_mean = 2.0, y_mean = 2.0, z_mean = 1.7
--------------------------------------------------------------------------------
```

Assignment Problem:
Program:

```
#!/usr/bin/env python3
import numpy as np
import pandas as pd

# read csv file and create dataframe
df = pd.read_csv("data.csv")

# randomly select rows of size = 2
dfs = df.sample(n=2, axis=0, replace=False)
df1 = dfs.iloc[0]; df2 = dfs.iloc[1]

df1x = df1.x; df1y = df1.y; df1z = df1.z
df2x = df2.x; df2y = df2.y; df2z = df2.z

print("random 1 values: x:{} y:{} z:{}".format(df1x,df1y,df1z))
print("random 2 values: x:{} y:{} z:{}".format(df2x,df2y,df2z))

m = 0
while m < 4:

    D1=[sum([abs(df1x-df.x[n]),abs(df1y-df.y[n]),abs(df1z-df.z[n])]) for n in range(len(df))]
    D2=[sum([abs(df2x-df.x[n]),abs(df2y-df.y[n]),abs(df2z-df.z[n])]) for n in range(len(df))]

    # adding columns with values of D1 and D2 to dataframe
    df['Dist. from C1({:.1f},{:.1f},{:.1f})'.format(df1x,df1y,df1z)] = D1
    df['Dist. from C2({:.1f},{:.1f},{:.1f})'.format(df2x,df2y,df2z)] = D2

    print("\nk-mean cluster: \n{}".format(df))
    df = df[["Objects", "x", "y", "z"]]

    cluster_1 = []
    cluster_2 = []
    for n in range(len(df)):
        if D1[n] < D2[n]: cluster_1.append(df.iloc[n])
        else: cluster_2.append(df.iloc[n])

    c1 = pd.DataFrame(cluster_1)
    c2 = pd.DataFrame(cluster_2)

    print("\nCluster 1: \n{}".format(c1))
    df1x = np.mean(c1.x); df1y = np.mean(c1.y); df1z = np.mean(c1.z)
    print("\nCluster 1: x_mean = {:.1f}, y_mean = {:.1f}, z_mean = {:.1f}".format(df1x, df1y,
    df1z))

    print("\nCluster 2: \n{}".format(c2))
    df2x = np.mean(c2.x); df2y = np.mean(c2.y); df2z = np.mean(c2.z)
    print("\nCluster 2: x_mean = {:.1f}, y_mean = {:.1f}, z_mean = {:.1f}".format(df2x, df2y,
    df2z))

    print("
    ----------------------------------------------------------------------------------------")

    m = m + 1
```

Output:

```
random 1 values: x:2 y:4 z:1
random 2 values: x:2 y:2 z:2

k-mean cluster:
  Objects  x  y  z  Dist. from C1(2.0,4.0,1.0)  Dist. from C2(2.0,2.0,2.0)
0    OB-1  2  4  1                           0                           3
1    OB-2  2  2  2                           3                           0
2    OB-3  1  2  1                           3                           2
3    OB-4  2  2  1                           2                           1

Cluster 1:
  Objects  x  y  z
0    OB-1  2  4  1

Cluster 1: x_mean = 2.0, y_mean = 4.0, z_mean = 1.0

Cluster 2:
  Objects  x  y  z
1    OB-2  2  2  2
2    OB-3  1  2  1
3    OB-4  2  2  1

Cluster 2: x_mean = 1.7, y_mean = 2.0, z_mean = 1.3
--------------------------------------------------------------------------------

k-mean cluster:
  Objects  x  y  z  Dist. from C1(2.0,4.0,1.0)  Dist. from C2(1.7,2.0,1.3)
0    OB-1  2  4  1                         0.0                    2.666667
1    OB-2  2  2  2                         3.0                    1.000000
2    OB-3  1  2  1                         3.0                    1.000000
3    OB-4  2  2  1                         2.0                    0.666667

Cluster 1:
  Objects  x  y  z
0    OB-1  2  4  1

Cluster 1: x_mean = 2.0, y_mean = 4.0, z_mean = 1.0

Cluster 2:
  Objects  x  y  z
1    OB-2  2  2  2
2    OB-3  1  2  1
3    OB-4  2  2  1

Cluster 2: x_mean = 1.7, y_mean = 2.0, z_mean = 1.3
--------------------------------------------------------------------------------

k-mean cluster:
  Objects  x  y  z  Dist. from C1(2.0,4.0,1.0)  Dist. from C2(1.7,2.0,1.3)
0    OB-1  2  4  1                         0.0                    2.666667
1    OB-2  2  2  2                         3.0                    1.000000
2    OB-3  1  2  1                         3.0                    1.000000
3    OB-4  2  2  1                         2.0                    0.666667

Cluster 1:
  Objects  x  y  z
0    OB-1  2  4  1

Cluster 1: x_mean = 2.0, y_mean = 4.0, z_mean = 1.0

Cluster 2:
```

```
  Objects  x  y  z
1    OB-2  2  2  2
2    OB-3  1  2  1
3    OB-4  2  2  1

Cluster 2: x_mean = 1.7, y_mean = 2.0, z_mean = 1.3
--------------------------------------------------------------------------------

k-mean cluster:
  Objects  x  y  z  Dist. from C1(2.0,4.0,1.0)  Dist. from C2(1.7,2.0,1.3)
0    OB-1  2  4  1                         0.0                   2.666667
1    OB-2  2  2  2                         3.0                   1.000000
2    OB-3  1  2  1                         3.0                   1.000000
3    OB-4  2  2  1                         2.0                   0.666667

Cluster 1:
  Objects  x  y  z
0    OB-1  2  4  1

Cluster 1: x_mean = 2.0, y_mean = 4.0, z_mean = 1.0

Cluster 2:
  Objects  x  y  z
1    OB-2  2  2  2
2    OB-3  1  2  1
3    OB-4  2  2  1

Cluster 2: x_mean = 1.7, y_mean = 2.0, z_mean = 1.3
--------------------------------------------------------------------------------
```