

Zadanie 1

```
import pandas as pd
import numpy as np

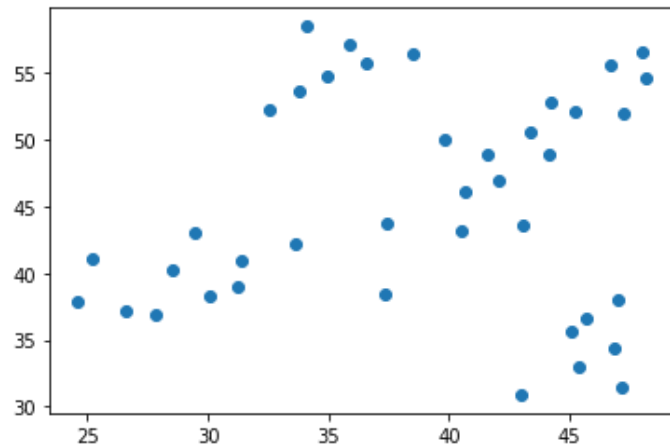
dataEx1 = pd.read_csv('k_means_data.csv')
dataEx1
```

	X	Y
0	25.23	41.09
1	45.10	35.69
2	26.59	37.21
3	27.80	36.93
4	28.56	40.21
5	29.49	43.05
6	30.04	38.33
7	31.25	39.03
8	31.35	40.98
9	42.95	30.91
10	37.30	38.42
11	37.39	43.69
12	32.53	52.18
13	33.60	42.20
14	24.60	37.88
15	33.76	53.60
16	34.97	54.72
17	35.84	57.04
18	47.16	31.52
19	36.58	55.76
20	46.82	34.33
21	46.98	38.03
22	45.34	33.02



```
import matplotlib.pyplot as plt

plt.scatter(dataEx1.iloc[:,0],dataEx1.iloc[:,1])
plt.show()
```



```
from sklearn.cluster import AgglomerativeClustering
```

```
k = 3
model = AgglomerativeClustering(linkage='single',n_clusters=k,affinity='euclidean', distance_threshold=None)
model.fit(dataEx1)
```

```
AgglomerativeClustering(linkage='single', n_clusters=3)
```

```
model.labels_
```

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

```
from sklearn.cluster import KMeans
```

```
k = 3
model = KMeans(n_clusters=3).fit(dataEx1)
model.labels_
```

```
array([2, 1, 2, 2, 2, 2, 2, 2, 2, 1, 2, 2, 0, 2, 2, 0, 0, 0, 1, 0, 1, 1,
       1, 1, 0, 0, 0, 0, 0, 1, 0, 0, 0, 1, 0, 0, 0, 0, 0, 0], dtype=int32)
```

```
K=model.cluster_centers_  
K
```

```
array([[40.91210526, 52.77789474],  
       [44.84         , 36.31444444],  
       [30.26666667, 39.91833333]])
```

Zadanie 2

```
data = pd.read_csv('k_means_data.csv')  
print(data)  
data = np.array(data)
```

	X	Y
0	25.23	41.09
1	45.10	35.69
2	26.59	37.21
3	27.80	36.93
4	28.56	40.21
5	29.49	43.05
6	30.04	38.33
7	31.25	39.03
8	31.35	40.98
9	42.95	30.91
10	37.30	38.42
11	37.39	43.69
12	32.53	52.18
13	33.60	42.20
14	24.60	37.88
15	33.76	53.60
16	34.97	54.72
17	35.84	57.04
18	47.16	31.52
19	36.58	55.76

20	46.82	34.33
21	46.98	38.03
22	45.34	33.02
23	45.70	36.63
24	48.12	54.53
25	47.25	51.99
26	38.48	56.41
27	34.10	58.45
28	39.81	49.98
29	40.47	43.18
30	40.66	46.02
31	41.59	48.86
32	42.05	46.89
33	43.04	43.52
34	43.34	50.49
35	44.17	48.94
36	44.23	52.81
37	45.19	52.05
38	46.68	55.59
39	47.98	56.47

```
Label = np.zeros((len(data),1),dtype=np.float64)
```

```
Label
```

```
array([[0.],  
       [0.],  
       [0.],  
       [0.],  
       [0.],  
       [0.],  
       [0.],  
       [0.],  
       [0.],  
       [0.],  
       [0.],  
       [0.],  
       [0.],  
       [0.],  
       [0.],  
       [0.],  
       [0.],  
       [0.],  
       [0.]])
```

```
NewData = np.concatenate([data,Label],axis=1)
NewData
```

<https://colab.research.google.com/drive/1gbDLqgzYsrLh-RMgVZNC66ghyUGPt8r3#scrollTo=-XkIHqGLUUL&printMode=true>

```
[35.84, 57.04, 0. ],
[47.16, 31.52, 0. ],
[36.58, 55.76, 0. ],
[46.82, 34.33, 0. ],
[46.98, 38.03, 0. ],
[45.34, 33.02, 0. ],
[45.7 , 36.63, 0. ],
[48.12, 54.53, 0. ],
[47.25, 51.99, 0. ],
[38.48, 56.41, 0. ],
[34.1 , 58.45, 0. ],
[39.81, 49.98, 0. ],
[40.47, 43.18, 0. ],
[40.66, 46.02, 0. ],
[41.59, 48.86, 0. ],
[42.05, 46.89, 0. ],
[43.04, 43.52, 0. ],
[43.34, 50.49, 0. ],
[44.17, 48.94, 0. ],
[44.23, 52.81, 0. ],
[45.19, 52.05, 0. ],
[46.68, 55.59, 0. ],
[47.98, 56.47, 0. ]])
```

```
NewData[:,2]
```

```
array([0., 0., 0., 0., 0., 0., 0., 0., 0., 0., 0., 0., 0., 0., 0., 0., 0.,
       0., 0., 0., 0., 0., 0., 0., 0., 0., 0., 0., 0., 0., 0., 0., 0.,
       0., 0., 0., 0., 0.])
```

```
minX = int(np.min(data[:,0]))
maxX = int(np.max(data[:,0]))
minY = int(np.min(data[:,1]))
maxY = int(np.max(data[:,1]))
print(minX," ",maxX," ",minY," ",maxY)
```

```
24    48    30    58
```

```
import random
CentersX = []
CentersY = []
```

```

for _ in range(0,3):
    CentersX.append(random.randrange(minX,maxX))
    CentersY.append(random.randrange(minY,maxY))

CentersX = np.array(CentersX).reshape(3,1)
CentersY = np.array(CentersY).reshape(3,1)

Centers = np.concatenate([CentersX,CentersY],axis=1).astype(np.float64)

```

Centers

```

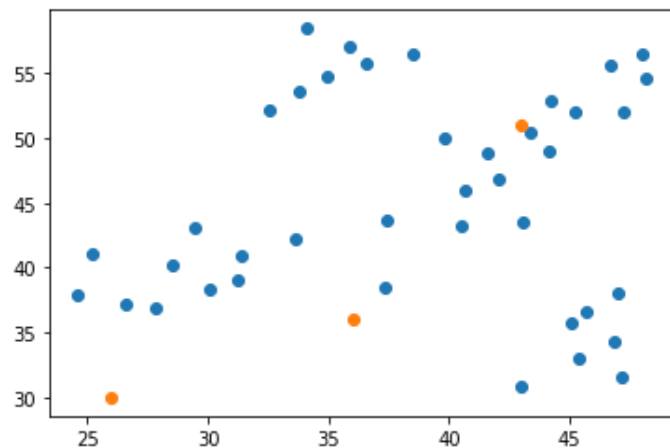
array([[36., 36.],
       [43., 51.],
       [26., 30.]])

```

```

import matplotlib.pyplot as plt
plt.scatter(data[:,0],data[:,1])
plt.scatter(Centers[:,0],Centers[:,1])
plt.show()

```



```
import math
```

```

for point in NewData:
    print(point[0],point[1])
    i = -1

```



```

j= -1
print(" ")
for cent in Centers:
    print("Cent",cent[0],cent[1])
    i= i+1
    euclidesian=math.sqrt(((cent[1]-point[1])*(cent[1]-point[1]))+((cent[0]-point[0])*(cent[0]-point[0])))
    if(i==0):
        min=euclidesian
        j=i
    if(min>euclidesian):
        min=euclidesian
        j=i

print(euclidesian)
print("Max"," i ", min,i,j)
point[2]=j

Cent 26.0 30.0
26.842423512045258
Max i 0.6129437168288784 2 1
44.17 48.94

Cent 36.0 36.0
15.303349306606053
Max i 15.303349306606053 0 0
Cent 43.0 51.0
2.3690715480964295
Max i 2.3690715480964295 1 1
Cent 26.0 30.0
26.24638070287025
Max i 2.3690715480964295 2 1
44.23 52.81

Cent 36.0 36.0
18.716543484308207
Max i 18.716543484308207 0 0
Cent 43.0 51.0
2.1883783950679097
Max i 2.1883783950679097 1 1
Cent 26.0 30.0
29.199811643228113
Max i 2.1883783950679097 2 1
45.19 52.05

```

```

Cent 36.0 36.0
18.494826303590955
Max i 18.494826303590955 0 0
Cent 43.0 51.0
2.428703357761088
Max i 2.428703357761088 1 1
Cent 26.0 30.0
29.231123823760175
Max i 2.428703357761088 2 1
46.68 55.59

```

```

Cent 36.0 36.0
22.312115542906284
Max i 22.312115542906284 0 0
Cent 43.0 51.0
5.883068927014202
Max i 5.883068927014202 1 1
Cent 26.0 30.0
32.90152732017163
Max i 5.883068927014202 2 1
47.98 56.47

```

```

Cent 36.0 36.0
23.717953115730705
Max i 23.717953115730705 0 0
Cent 43.0 51.0
7.397384672977332
Max i 7.397384672977332 1 1
Cent 26.0 30.0
34.406123001582145
Max i 7.397384672977332 2 1

```

NewData

```

array([[25.23, 41.09, 2. ],
       [45.1 , 35.69, 0. ],
       [26.59, 37.21, 2. ],
       [27.8 , 36.93, 2. ],
       [28.56, 40.21, 0. ],
       [29.49, 43.05, 0. ],
       [30.04, 38.33, 0. ],
       [31.25, 39.03, 0. ],

```

```
[ 31.35, 40.98, 0. ],
[ 42.95, 30.91, 0. ],
[ 37.3 , 38.42, 0. ],
[ 37.39, 43.69, 0. ],
[ 32.53, 52.18, 1. ],
[ 33.6 , 42.2 , 0. ],
[ 24.6 , 37.88, 2. ],
[ 33.76, 53.6 , 1. ],
[ 34.97, 54.72, 1. ],
[ 35.84, 57.04, 1. ],
[ 47.16, 31.52, 0. ],
[ 36.58, 55.76, 1. ],
[ 46.82, 34.33, 0. ],
[ 46.98, 38.03, 0. ],
[ 45.34, 33.02, 0. ],
[ 45.7 , 36.63, 0. ],
[ 48.12, 54.53, 1. ],
[ 47.25, 51.99, 1. ],
[ 38.48, 56.41, 1. ],
[ 34.1 , 58.45, 1. ],
[ 39.81, 49.98, 1. ],
[ 40.47, 43.18, 1. ],
[ 40.66, 46.02, 1. ],
[ 41.59, 48.86, 1. ],
[ 42.05, 46.89, 1. ],
[ 43.04, 43.52, 1. ],
[ 43.34, 50.49, 1. ],
[ 44.17, 48.94, 1. ],
[ 44.23, 52.81, 1. ],
[ 45.19, 52.05, 1. ],
[ 46.68, 55.59, 1. ],
[ 47.98, 56.47, 1. ]])
```

```
NewData[:,2]
```

```
array([2., 0., 2., 2., 0., 0., 0., 0., 0., 0., 0., 0., 1., 0., 2., 1., 1.,
       1., 0., 1., 0., 0., 0., 0., 1., 1., 1., 1., 1., 1., 1., 1., 1.,
       1., 1., 1., 1., 1., 1.])
```

```
OldCenters = Centers
```

```
while True:
```

```
    x0=0
```

```
    y0=0
```

```

i0=0
x1=0
y1=0
i1=0
x2=0
y2=0
i2=0

Cluster1X=[]
Cluster1Y=[]
Cluster2X=[]
Cluster2Y=[]
Cluster3X=[]
Cluster3Y=[]
for point in NewData:
    print(point[0],point[1],point[2])
    if(point[2]==0):
        x0=x0+point[0]
        Cluster1X.append(point[0])
        y0=y0+point[1]
        Cluster1Y.append(point[1])
        i0=i0+1
    if(point[2]==1):
        x1=x1+point[0]
        Cluster2X.append(point[0])
        y1=y1+point[1]
        Cluster2Y.append(point[1])
        i1=i1+1
    if(point[2]==2):
        x2=x2+point[0]
        Cluster3X.append(point[0])
        y2=y2+point[1]
        Cluster3Y.append(point[1])
        i2=i2+1

NewCentersX=[]
NewCentersY=[]

#NewCentersX.append(x0/i0)
#NewCentersY.append(y0/i0)
#NewCentersX.append(x1/i1)
#NewCentersY.append(y1/i1)

```

```
#NewCentersX.append(x1/i1)
#NewCentersY.append(y1/i1)
#NewCentersX.append(x2/i2)
#NewCentersY.append(y2/i2)
NewCentersX.append(np.mean(Cluster1X))
NewCentersX.append(np.mean(Cluster2X))
NewCentersX.append(np.mean(Cluster3X))
NewCentersY.append(np.mean(Cluster1Y))
NewCentersY.append(np.mean(Cluster2Y))
NewCentersY.append(np.mean(Cluster3Y))
NewCentersX = np.array(NewCentersX).reshape(3,1)
NewCentersY = np.array(NewCentersY).reshape(3,1)

print(np.mean(Cluster1X))
print(np.mean(Cluster1Y))
print(np.mean(Cluster2X))
print(np.mean(Cluster2Y))
print(np.mean(Cluster3X))
print(np.mean(Cluster3Y))

NewCenters = np.concatenate([NewCentersX,NewCentersY],axis=1).astype(np.float64)

print(x0,y0,i0)
print(x1,y1,i1)
print(x2,y2,i2)
plt.scatter(NewData[:,0],NewData[:,1])
plt.scatter(NewCenters[:,0],NewCenters[:,1])
plt.show()

print(OldCenters)
print(NewCenters)
print("Space")

if np.array_equal(OldCenters,NewCenters)==True:
    break
OldCenters = NewCenters
#NewCenters
```

```
for point in NewData:
    #print(point[0],point[1])
```

```
i = -1
j= -1
#print(" ")
for cent in NewCenters:
    print("Cent",cent[0],cent[1])
    i= i+1
    #euclidesian=math.sqrt(((cent[1]-point[1])*(cent[1]-point[1]))+((cent[0]-point[0])*(cent[0]-point[0])))
    euclidesian=math.sqrt(((cent[0]-point[0])*(cent[0]-point[0]))+((cent[1]-point[1])*(cent[1]-point[1])))
    if(i==0):
        min=euclidesian
        j=i
    if(min>euclidesian):
        min=euclidesian
        j=i

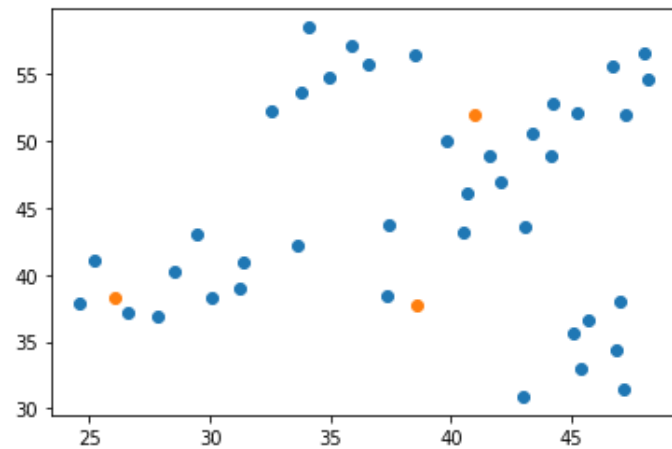
#print(euclidesian)
#print("Max"," i ", min,i,j)
point[2]=j
```

```
25.23 41.09 2.0
45.1 35.69 0.0
26.59 37.21 2.0
27.8 36.93 2.0
28.56 40.21 0.0
29.49 43.05 0.0
30.04 38.33 0.0
31.25 39.03 0.0
31.35 40.98 0.0
42.95 30.91 0.0
37.3 38.42 0.0
37.39 43.69 0.0
32.53 52.18 1.0
33.6 42.2 0.0
24.6 37.88 2.0
33.76 53.6 1.0
34.97 54.72 1.0
35.84 57.04 1.0
47.16 31.52 0.0
36.58 55.76 1.0
46.82 34.33 0.0
46.98 38.03 0.0
45.34 33.02 0.0
45.7 36.63 0.0
48.12 54.53 1.0
47.25 51.99 1.0
38.48 56.41 1.0
34.1 58.45 1.0
39.81 49.98 1.0
40.47 43.18 1.0
40.66 46.02 1.0
41.59 48.86 1.0
42.05 46.89 1.0
43.04 43.52 1.0
43.34 50.49 1.0
44.17 48.94 1.0
44.23 52.81 1.0
45.19 52.05 1.0
46.68 55.59 1.0
47.98 56.47 1.0
38.602
37.736
40.99238095238094
51.879999999999999
```

```

26.055
38.2775
579.03000000000001 566.04 15
860.84 1089.4799999999998 21
104.22 153.11 4

```



```

[[36. 36.]
 [43. 51.]
 [26. 30.]]
[[38.602      37.736      ]
 [40.99238095 51.88      ]
 [26.055      38.2775     ]]

```

Space

```

Cent 38.602 37.736
Cent 40.99238095238094 51.879999999999999
Cent 26.055 38.2775
Cent 38.602 37.736
Cent 40.99238095238094 51.879999999999999
Cent 26.055 38.2775
Cent 38.602 37.736
Cent 40.99238095238094 51.879999999999999
Cent 26.055 38.2775
Cent 38.602 37.736
Cent 40.99238095238094 51.879999999999999
Cent 26.055 38.2775
Cent 38.602 37.736
Cent 40.99238095238094 51.879999999999999
Cent 26.055 38.2775
Cent 38.602 37.736
Cent 40.99238095238094 51.879999999999999
Cent 26.055 38.2775

```


<https://colab.research.google.com/drive/1gbDLqgzYsrLh-RMgVZNC66ghyUGPt8r3#scrollTo=-XkiHYqGLUUL&printMode=true>

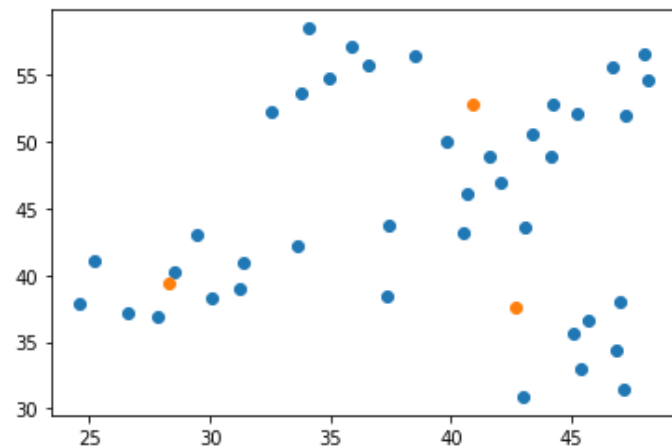
[illegible]

```
Cent 26.055 38.2775
Cent 38.602 37.736
Cent 40.99238095238094 51.879999999999999
Cent 26.055 38.2775
Cent 38.602 37.736
Cent 40.99238095238094 51.879999999999999
Cent 26.055 38.2775
Cent 38.602 37.736
Cent 40.99238095238094 51.879999999999999
Cent 26.055 38.2775
Cent 38.602 37.736
Cent 40.99238095238094 51.879999999999999
Cent 26.055 38.2775
25.23 41.09 2.0
45.1 35.69 0.0
26.59 37.21 2.0
27.8 36.93 2.0
28.56 40.21 2.0
29.49 43.05 2.0
30.04 38.33 2.0
31.25 39.03 2.0
31.35 40.98 2.0
42.95 30.91 0.0
37.3 38.42 0.0
37.39 43.69 0.0
32.53 52.18 1.0
33.6 42.2 0.0
24.6 37.88 2.0
33.76 53.6 1.0
34.97 54.72 1.0
35.84 57.04 1.0
47.16 31.52 0.0
36.58 55.76 1.0
46.82 34.33 0.0
46.98 38.03 0.0
45.34 33.02 0.0
45.7 36.63 0.0
48.12 54.53 1.0
47.25 51.99 1.0
38.48 56.41 1.0
34.1 58.45 1.0
39.81 49.98 1.0
40.47 43.18 0.0
40.66 46.02 1.0
41.59 48.86 1.0
```

```

42.05 46.89 1.0
43.04 43.52 0.0
43.34 50.49 1.0
44.17 48.94 1.0
44.23 52.81 1.0
45.19 52.05 1.0
46.68 55.59 1.0
47.98 56.47 1.0
42.654166666666666
37.594999999999999
40.9121052631579
52.77789473684211
28.323333333333334
39.41222222222222
511.84999999999997 451.13999999999993 12
777.33 1002.7799999999999 19
254.91 354.71000000000004 9

```



```

[[ 38.602      37.736      ]
 [ 40.99238095 51.88       ]
 [ 26.055      38.2775      ]]
[[ 42.65416667 37.595      ]
 [ 40.91210526 52.77789474]
 [ 28.32333333 39.41222222]]

```

Space

```

Cent 42.654166666666666 37.594999999999999
Cent 40.9121052631579 52.77789473684211
Cent 28.323333333333334 39.41222222222222
Cent 42.654166666666666 37.594999999999999
Cent 40.9121052631579 52.77789473684211
Cent 28.323333333333334 39.41222222222222

```

— — — — —

<https://colab.research.google.com/drive/1gbDLqgzYsrLh-RMgVZNC66ghyUGPt8r3#scrollTo=-XkiHYqGLUUL&printMode=true>

```
Cent 28.323333333333334 39.41222222222222
Cent 42.654166666666666 37.59499999999999
Cent 40.9121052631579 52.77789473684211
Cent 28.323333333333334 39.41222222222222
Cent 42.654166666666666 37.59499999999999
Cent 40.9121052631579 52.77789473684211
Cent 28.323333333333334 39.41222222222222
Cent 42.654166666666666 37.59499999999999
Cent 40.9121052631579 52.77789473684211
Cent 28.323333333333334 39.41222222222222
Cent 42.654166666666666 37.59499999999999
Cent 40.9121052631579 52.77789473684211
Cent 28.323333333333334 39.41222222222222
Cent 42.654166666666666 37.59499999999999
Cent 40.9121052631579 52.77789473684211
Cent 28.323333333333334 39.41222222222222
Cent 42.654166666666666 37.59499999999999
Cent 40.9121052631579 52.77789473684211
Cent 28.323333333333334 39.41222222222222
Cent 42.654166666666666 37.59499999999999
Cent 40.9121052631579 52.77789473684211
Cent 28.323333333333334 39.41222222222222
Cent 42.654166666666666 37.59499999999999
Cent 40.9121052631579 52.77789473684211
Cent 28.323333333333334 39.41222222222222
Cent 42.654166666666666 37.59499999999999
Cent 40.9121052631579 52.77789473684211
Cent 28.323333333333334 39.41222222222222
25.23 41.09 2.0
45.1 35.69 0.0
26.59 37.21 2.0
27.8 36.93 2.0
28.56 40.21 2.0
29.49 43.05 2.0
30.04 38.33 2.0
31.25 39.03 2.0
31.35 40.98 2.0
42.95 30.91 0.0
37.3 38.42 0.0
37.39 43.69 0.0
32.53 52.18 1.0
33.6 42.2 2.0
24.6 37.88 2.0
33.76 53.6 1.0
34.97 54.72 1.0
35.84 57.04 1.0
47.16 31.52 0.0
-- -- -- -- --
```