

```
import numpy as np
import pandas as pd

X=[
    (0.1, 0.6),
    (0.15, 0.71),
    (0.08,0.9),
    (0.16, 0.85),
    (0.2,0.3),
    (0.25,0.5),
    (0.24,0.1),
    (0.3,0.2)
]

class KMeans:
    def __init__(self,k):
        self.k=k

    def distances(self,x,y):
        return ((x[0]-y[0])**2 + (x[1]-y[1])**2)**0.5

    def fit(self,X,centroids):
        clusters = [set() for i in range(self.k)]
        print(clusters)
        for x in X:
            idx=0
            for i in range(1,self.k):
                if self.distances(centroids[idx],x) > self.distances(centroids[i],x):
                    idx=i
            clusters[idx].add(x)
        print(clusters)

        for i in range(self.k):
            centroids[i] = np.mean(list(clusters[i]), axis = 0)
        print(centroids)
        return clusters,centroids

c1 = KMeans(2)

clusters,centroids = c1.fit(X,[(0.1,0.6),(0.3,0.2)])

[set(), set()]
[ {(0.08, 0.9), (0.16, 0.85), (0.1, 0.6), (0.15, 0.71), (0.25, 0.5)}, {(0.2, 0.3), (0.24, 0.1), (0.3, 0.2)} ]
[array([0.148, 0.712]), array([0.24666667, 0.2      ])]

df = pd.DataFrame()

x = []
y = []
category = []

for i in range(len(clusters)):
    for pt in clusters[i]:
        x.append(pt[0])
        y.append(pt[1])
        category.append(f'{i}')

for c in centroids:
    x.append(c[0])
    y.append(c[1])
    category.append('category')

df['x']=x
df['y']=y
df['category']=category
df


```

	x	y	category
0	0.080000	0.900	0
1	0.160000	0.850	0
2	0.100000	0.600	0
3	0.150000	0.710	0
4	0.250000	0.500	0
5	0.200000	0.300	1
6	0.240000	0.100	1
7	0.300000	0.200	1
8	0.148000	0.712	category
9	0.246667	0.200	category

```
import matplotlib.pyplot as plt

import seaborn as sns
sns.set(rc={'figure.figsize':(7, 7)})
sns.scatterplot(data=df,x='x',y='y',hue='category')
```

<AxesSubplot:xlabel='x', ylabel='y'>

