k-means

# example code

|  |
| --- |
| from sklearn import datasets  import pandas as pd  import numpy as np  import matplotlib.pyplot as plt  iris = datasets.load\_iris()  df = pd.DataFrame(iris.data)  df.columns=['Sepal length','Sepal width','Petal length','Petal width']  df = df[['Sepal length','Sepal width']].copy()  df.head()  x1,y1 = 5, 3.5  x2,y2 = 6, 3  x3,y3 = 7.5, 4  # before  plt.figure(figsize=(7,5))  plt.title("Before", fontsize=15)  plt.plot(df["Sepal length"], df["Sepal width"], "o", label="Data")  plt.plot([x1,x2,x3], [y1,y2,y3], "rD", markersize=12, label='init\_Centroid')  plt.xlabel("Sepal length", fontsize=12)  plt.ylabel("Sepal width", fontsize=12)  plt.legend()  plt.grid()  plt.show()  # after  from sklearn.cluster import KMeans  kmeans = KMeans(n\_clusters=3, init=np.array([(x1,y1),(x2,y2),(x3,y3)])).fit(df)  df['cluster'] = kmeans.labels\_  final\_centroid = kmeans.cluster\_centers\_  plt.figure(figsize=(7,5))  plt.title("After", fontsize=15)  plt.scatter(df['Sepal length'],df['Sepal width'],c=df['cluster'])  plt.plot(final\_centroid[:,0], final\_centroid[:,1], "rD", markersize=12, label='final\_Centroid')  plt.xlabel("Sepal length", fontsize=12)  plt.ylabel("Sepal width", fontsize=12)  plt.legend()  plt.grid()  plt.show() |

# testing result

|  |
| --- |
|  |