# Load libraries

import json

import pandas

from IPython.display import display, HTML

# load data

raw = pandas.read\_csv("raw.csv", encoding="big5", index\_col=[0,1])

raw = raw.reset\_index("區域")

X = raw.reset\_index().iloc[:,7:10]

'''X = pandas.concat([X.iloc[:,[0,4]].rename(columns={'7-11 (%)':'persentage'}),

X.iloc[:,[1,4]].rename(columns={'OK (%)':'persentage'}),

X.iloc[:,[2,4]].rename(columns={'全家 (%)':'persentage'}),

X.iloc[:,[3,4]].rename(columns={'萊爾富 (%)':'persentage'})]).reset\_index().iloc[:,1:3]'''

display(X)

'''y = pandas.concat([raw.iloc[:,0], raw.iloc[:,0], raw.iloc[:,0], raw.iloc[:,0]]).reset\_index().iloc[:,1:2]

display(y)

region = {"北":0, "中":1, "南":2, "東":3, "外":4}

y = y["區域"].map(region)

display(y)'''

region = {"北":0, "中":1, "南":2, "東":3, "外":4}

y = raw["區域"].map(region)

display(y)

from mpl\_toolkits.mplot3d import Axes3D

import matplotlib.pyplot as plt

plt.rcParams['font.sans-serif'] = ['Microsoft JhengHei']

#接下來匯入KMeans函式庫

from sklearn.cluster import KMeans

#請KMeans分成三類

clf = KMeans(n\_clusters=3)

#開始訓練！

clf.fit(X)

#這樣就可以取得預測結果了！

print(clf.labels\_)

#最後畫出來看看

#真的分成三類！太神奇了………無意義的資料也能分～

fig = plt.figure(figsize=(10,10))

ax = fig.add\_subplot(111, projection='3d')

ax.scatter(X.iloc[:,2], X.iloc[:,1], X.iloc[:,0], c=clf.labels\_)

ax.set\_zlabel('7-11%')

ax.set\_ylabel('全家 (%)')

ax.set\_xlabel('萊爾富 (%)')

plt.show()