Xinyi Gu
xinyigu2

Number of P	0N	1N	2N	3N	4N	0c	1c	2c	3c	4c
Dataset IÊ	4.54247	0.38345	0.17556	0.14178	0.16084	4.54311903	0.38461353	0.17781528	0.14444051	0.16083836
Dataset IIÊ	4.54247	0.64109	0.71563	0.90839	1.11566	4.54953899	0.64864211	0.75062113	0.94197282	1.11565786
Dataset IIIÊ	4.54247	1.29037	1.96724	2.65084	3.65328	4.55747296	1.32346215	2.11974805	3.02737992	3.65327973
Dataset IVÊ	4.54247	0.79994	0.82808	0.98495	1.194	4.56619867	0.84061416	1.2070898	1.27119197	1.194
Dataset V	4.54247	1.91777	3.33172	4.54826	5.13927	4.919928	2.83567943	4.6514345	4.97124727	5.13926667

```
import numpy as np
import pandas as pd
import math
import random
import csv
from sklearn.decomposition import PCA
from sklearn.metrics import mean_squared_error

data1 = pd.read_csv("/Users/Sunny/Desktop/CS498/HW3/hw3-data/dataI.csv").as_matrix()
data2 = pd.read_csv("/Users/Sunny/Desktop/CS498/HW3/hw3-data/dataII.csv").as_matrix()
data3 = pd.read_csv("/Users/Sunny/Desktop/CS498/HW3/hw3-data/dataIII.csv").as_matrix()
data4 = pd.read_csv("/Users/Sunny/Desktop/CS498/HW3/hw3-data/dataIV.csv").as_matrix()
data5 = pd.read_csv("/Users/Sunny/Desktop/CS498/HW3/hw3-data/dataV.csv").as_matrix()
iris  = pd.read_csv("/Users/Sunny/Desktop/CS498/HW3/hw3-data/iris.csv").as_matrix()
def pca_rep(data,fit, pc):
    mean = np.mean(data, axis=0)
    pca = PCA(n_components = pc)
    pca = pca.fit(fit)
    Xhat = np.dot(pca.transform(data)[:,:pc], pca.components_[:pc,:])
    Xhat += mean
    return Xhat

def pca_rep_nl(data,fit, pc):
    mean = np.mean(fit, axis=0)
    pca = PCA()
    pca = pca.fit(fit)
    Xhat = np.dot(pca.transform(data)[:,:pc], pca.components_[:pc,:])
    Xhat += mean
    return Xhat

def get_mse(data,pc):
    datare = pca_rep(data,iris,pc)
    mse = mean_squared_error(iris,datare)*4
    return mse

with open('xinyigu2-recon.csv','wt') as csvfile:
    writer = csv.writer(csvfile)
    writer.writerow(("X1","X2","X3","X4"))
    array =  pca_rep(data2,data2,2)
    for i in range(len(array)):
        writer.writerow((array[i]))
```