```
import pandas as pd
import numpy as np

#测试用数据集
def loadExData():
    return np.mat([[0, 0, 0, 0, 0, 4, 0, 0, 0, 0, 5],
           [0, 0, 0, 3, 0, 4, 0, 0, 0, 0, 3],
           [0, 0, 0, 0, 4, 0, 0, 1, 0, 4, 0],
           [3, 3, 4, 0, 0, 0, 0, 2, 2, 0, 0],
           [5, 4, 5, 0, 0, 0, 0, 5, 5, 0, 0],
           [0, 0, 0, 0, 5, 0, 1, 0, 0, 5, 0],
           [4, 3, 4, 0, 0, 0, 0, 5, 5, 0, 1],
           [0, 0, 0, 4, 0, 4, 0, 0, 0, 0, 4],
           [0, 0, 0, 2, 0, 2, 5, 0, 0, 1, 2],
           [0, 0, 0, 0, 5, 0, 0, 0, 0, 4, 0],
           [1, 0, 0, 0, 0, 0, 0, 1, 2, 0, 0]])

def loadData():
    attribute = pd.read_csv('./code/attribute.data', header=None, names=['A', 'B',
 'C', 'D', 'E']).reset_index(drop=True)
    xKey = attribute['B'].tolist()
    #xKey.sort()
    data = pd.read_csv('./code/anonymous-msweb.data',header=None,names=['A','B','C
']).reset_index(drop=True)
    caseLineList = data[data['A']=='C'].index.tolist()
    n = max(xKey)-min(xKey)
    m = len(caseLineList)
    dataMatrix = np.zeros((m,n),dtype=int)
    for i in range(1, m):
        #caesNum = caseLineList[i]
        cStartPoint = caseLineList[i-1]
        cStopPoint = caseLineList[i]
        vPointList = data.iloc[cStartPoint+1 : cStopPoint]['B'].tolist()
        matrixRow = data.iloc[cStartPoint]['B'] - 10001
        for j in vPointList:
            dataMatrix[matrixRow][j - 1000] = 1
    return dataMatrix

def compuDiagMatrix(dataMatrix):
    tempMatrix = np.dot(dataMatrix.T, dataMatrix)
    eigenvalues, featureVector = np.linalg.eig(tempMatrix)
    totalSum = sum( eigenvalues )
    partSum = 0
    i = 0
    diagList = []
    while(partSum < totalSum*0.9):
        diagList.append(np.sqrt(eigenvalues[i]))
        partSum += eigenvalues[i]
        i = i + 1
    return np.diag(diagList), i



if __name__ == '__main__':
    dataMatrix = loadData()
    #dataMatrix = loadExData()
    resMatrix, num = compuDiagMatrix(dataMatrix)
    print("奇异值分解之后的对角矩阵")
    print(resMatrix)
    print("原矩阵维数：",dataMatrix.shape[1])
    print("分解之后矩阵维数：",resMatrix.shape[1])
    #print('Well done!')
```

```
奇异值分解之后的对角矩阵
[[143.77711852   0.           0.         ...   0.           0.
    0.        ]
 [  0.          94.55432015   0.         ...   0.           0.
    0.        ]
 [  0.           0.          80.29654019 ...   0.           0.
    0.        ]
 ...
 [  0.           0.           0.         ...  14.5063684    0.
    0.        ]
 [  0.           0.           0.         ...   0.          13.92835253
    0.        ]
 [  0.           0.           0.         ...   0.           0.
   14.09230649]]
原矩阵维数： 297
分解之后矩阵维数： 66
```