PYTHON

ICP – 6

Authored By

Nikhitha Kolluri

Lesson Overview:

In this lesson, we will focus on clustering techniques

Clustering using KMeansand Dimension reduction using PCA

Use Case Description:

1. KMeans Clusteringand PCA

Programming elements:

KMeans Clustering andData Analysis

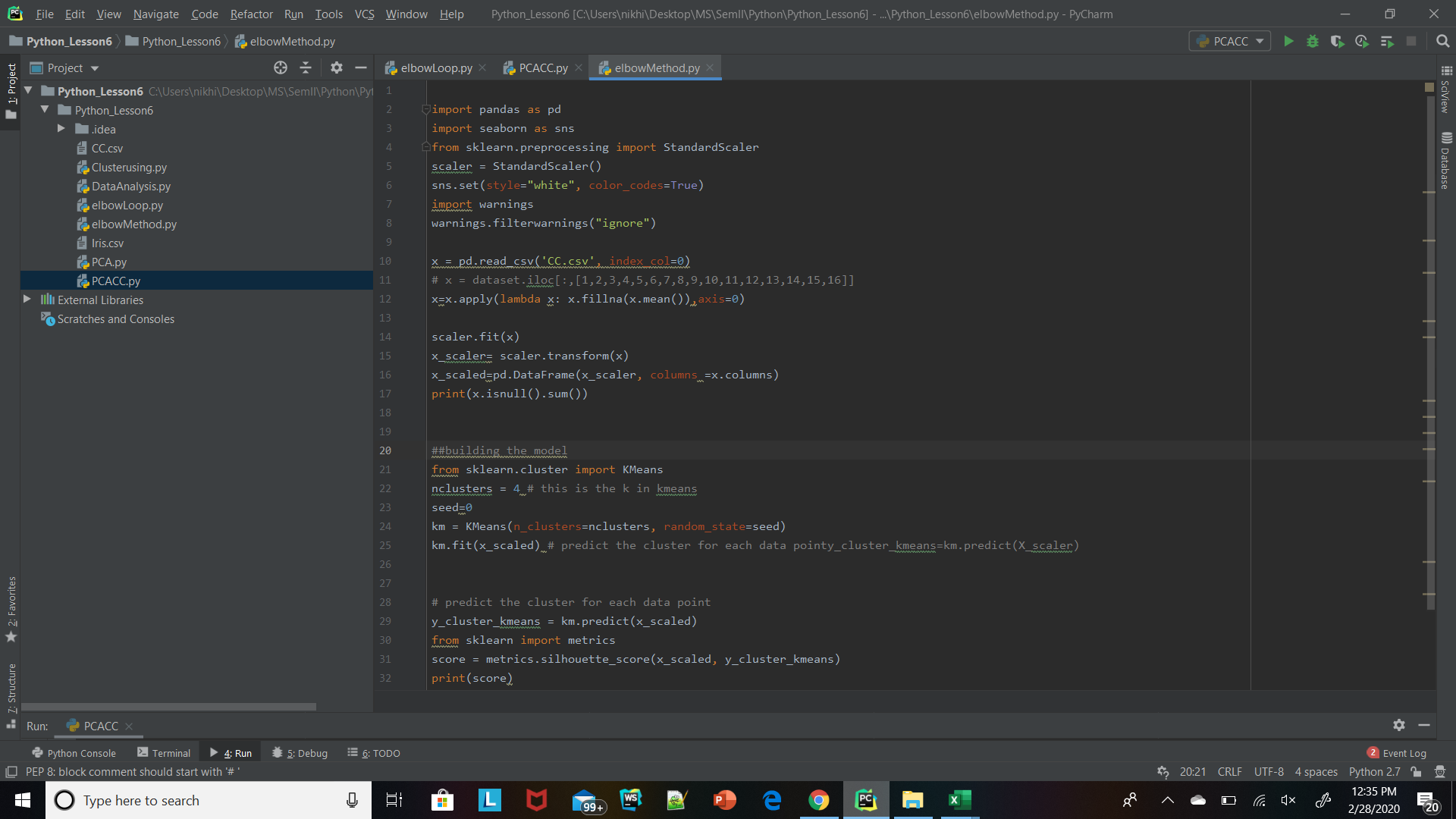
Source Code:

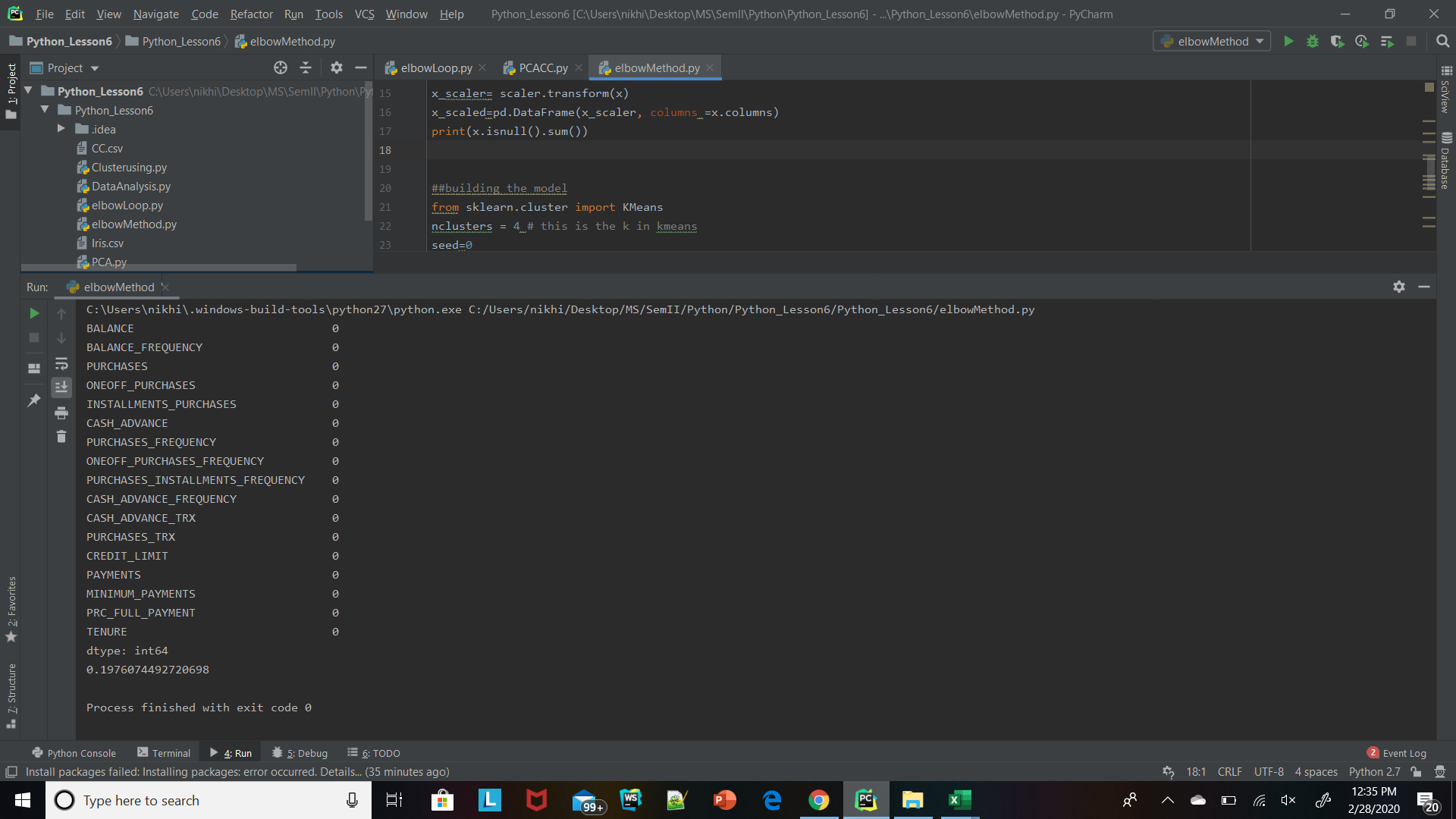
<https://umkc.box.com/s/pgqpapgr12wyuj63sgs3162wztpe3q77>

In class programming:

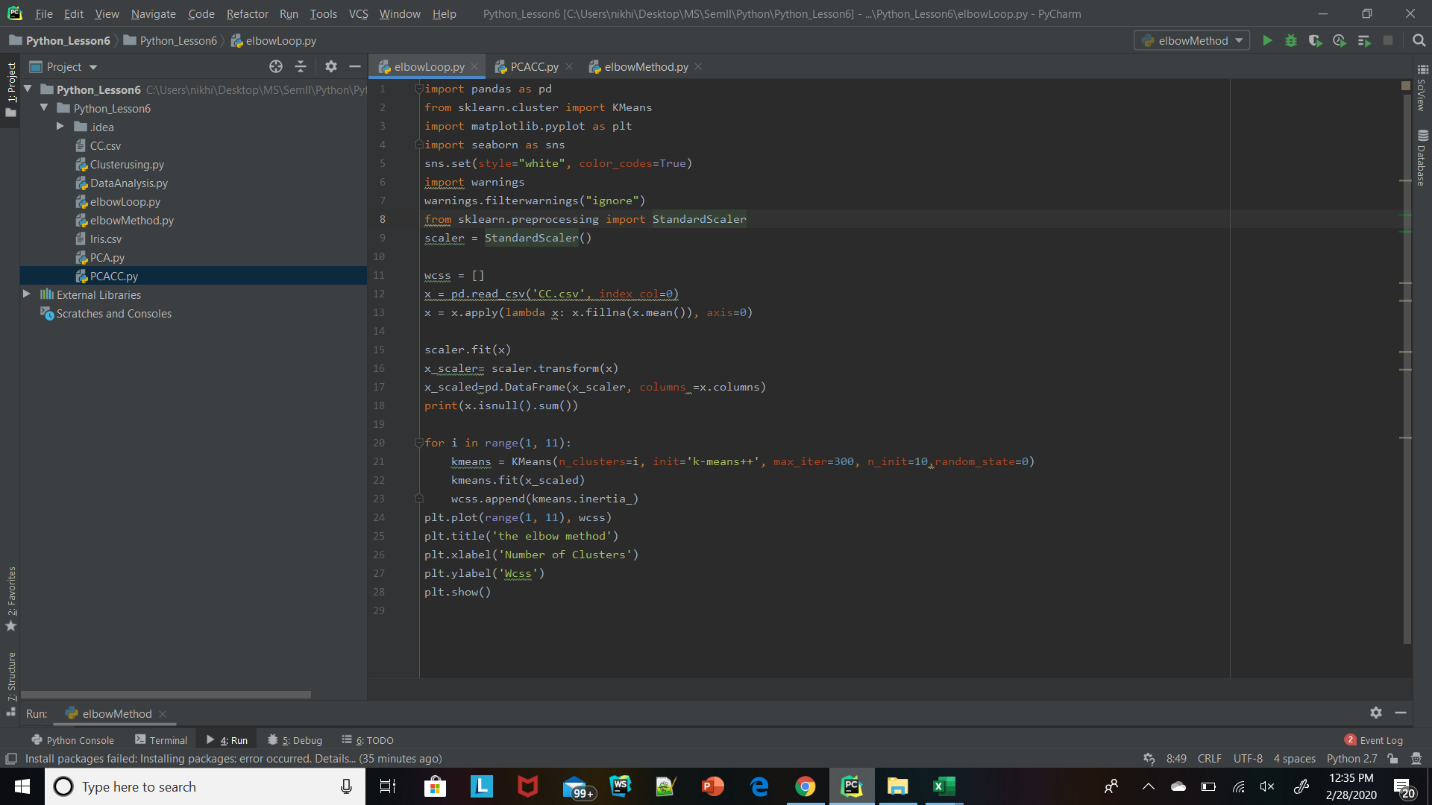
1.Apply K means clustering in this data set provided below:https://umkc.box.com/s/s15r7m0gnxu7b1s2kaobvc5w7da2nc1c

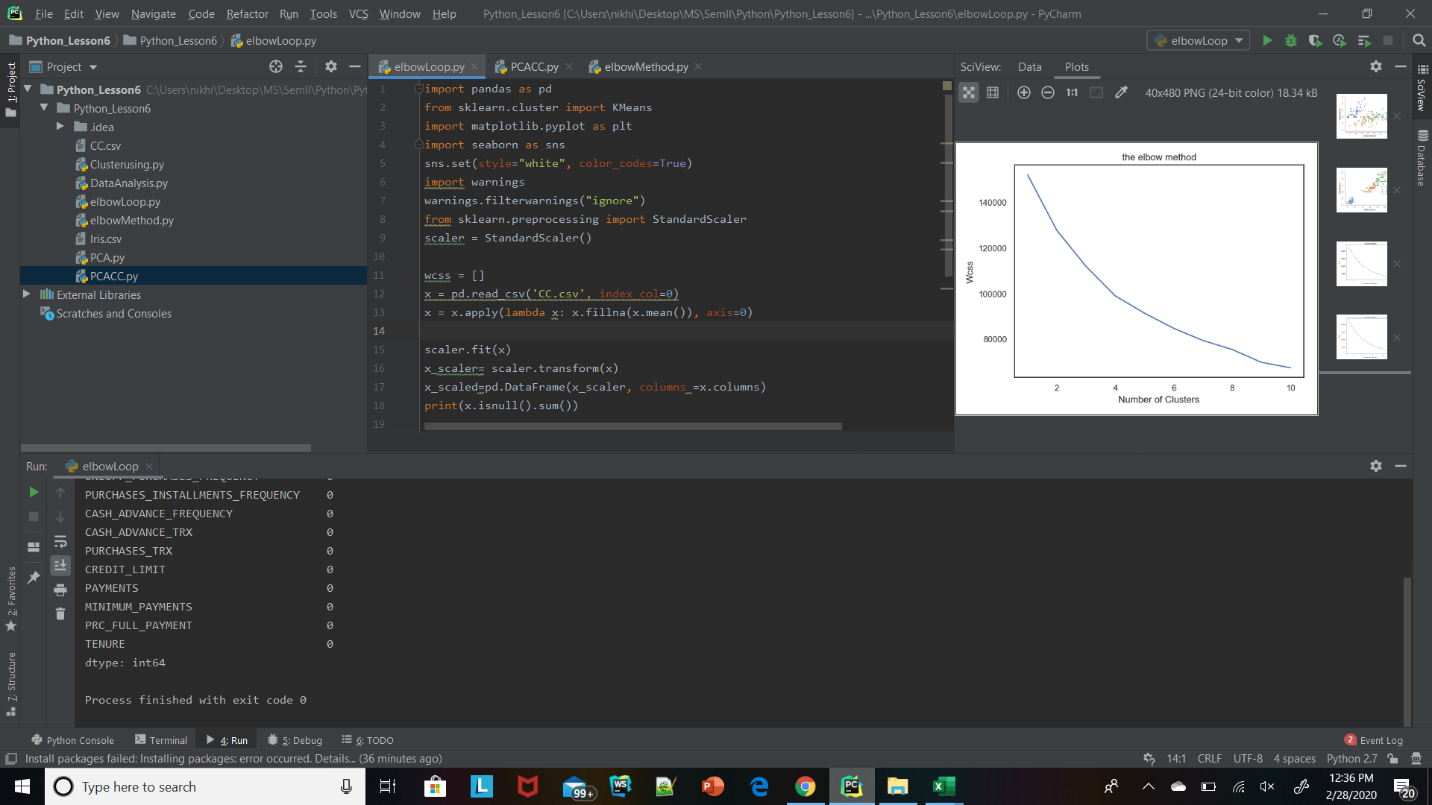
•Remove any null values by the mean.



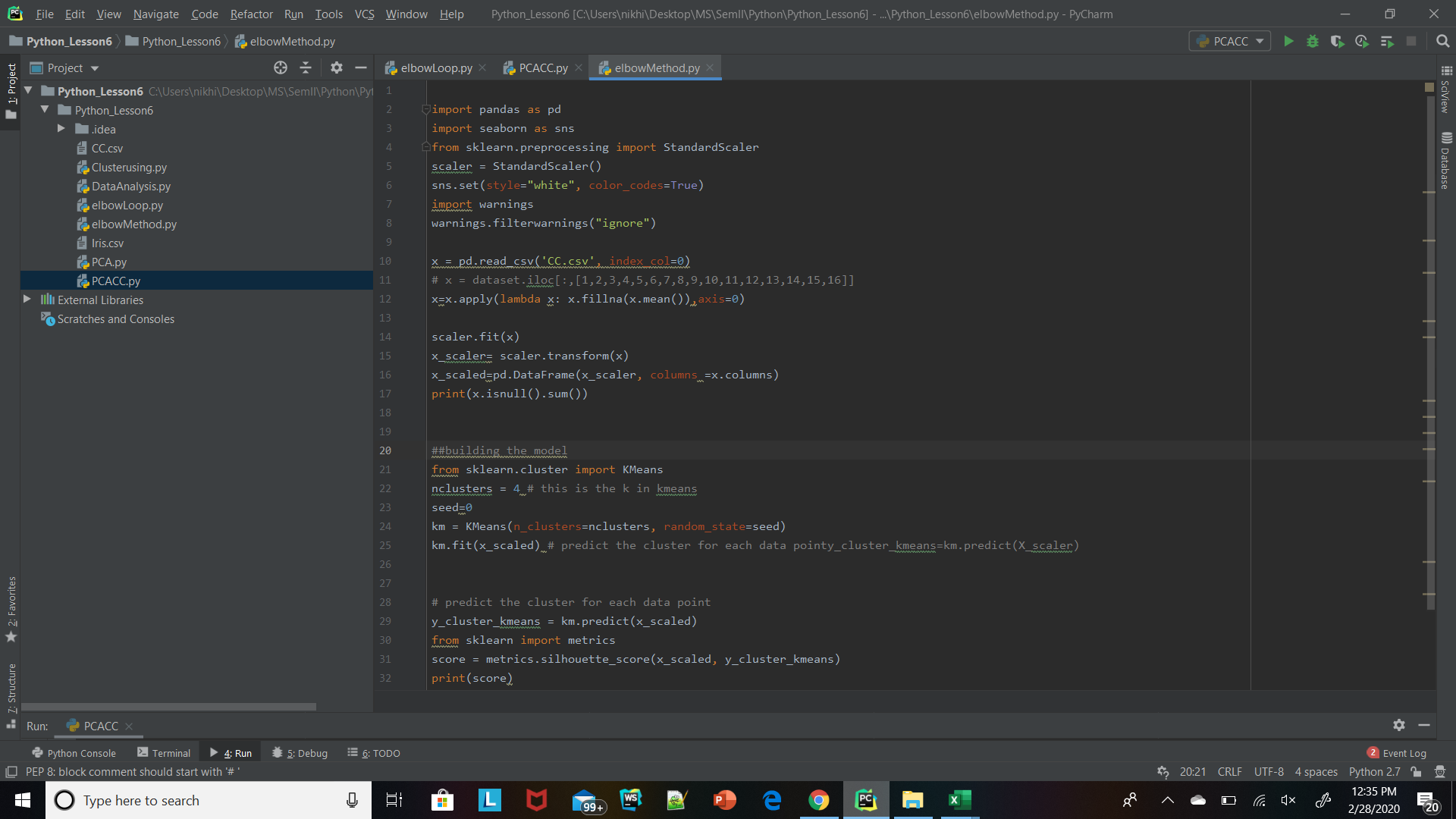


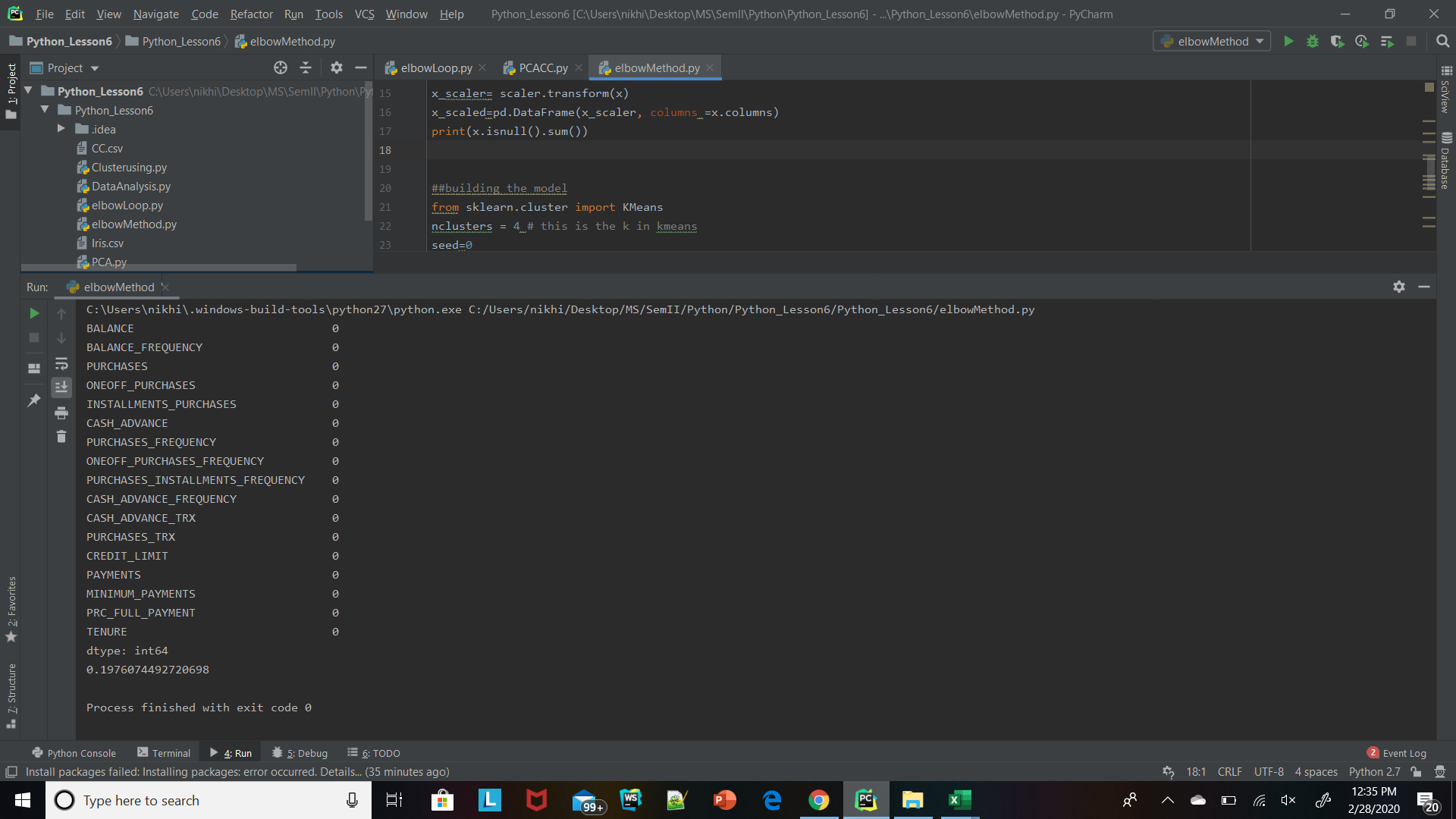
* Use the elbow method to find a good number of clusters with the KMeans algorithm





1. Calculate the silhouette score for the above clustering
2. Try feature scaling to see if it willl improve the Silhouette score





1. Apply PCA on the same dataset.

