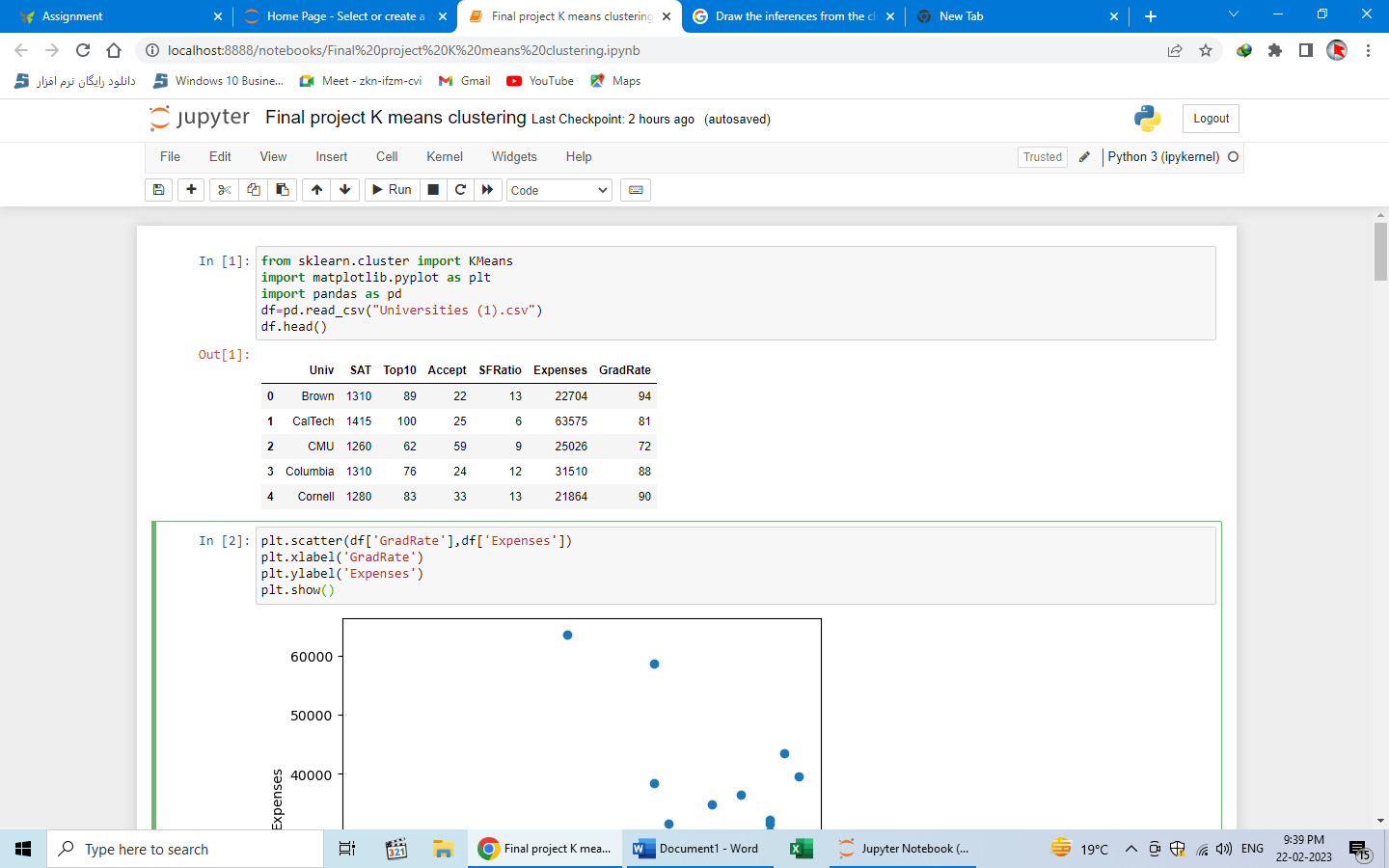
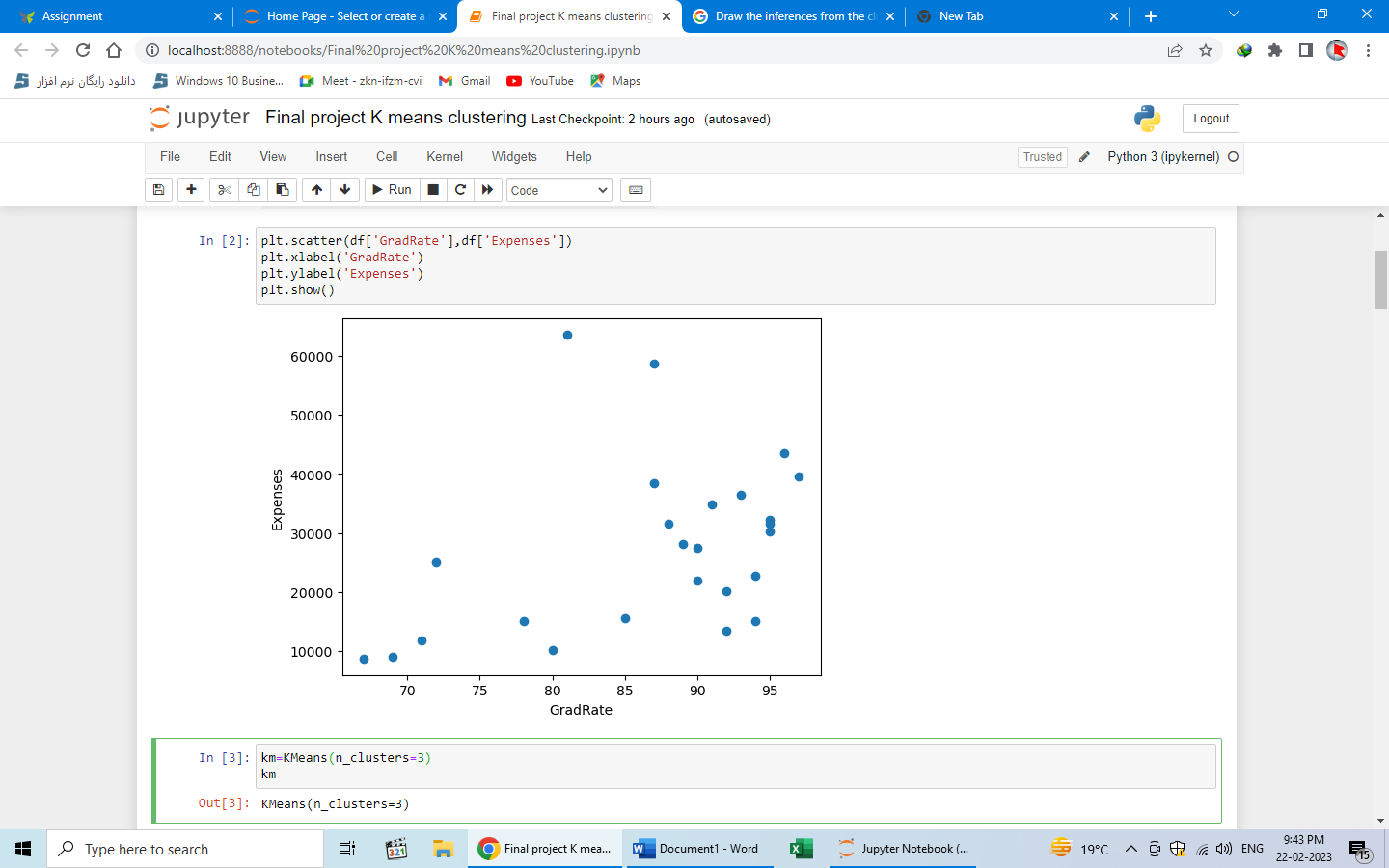
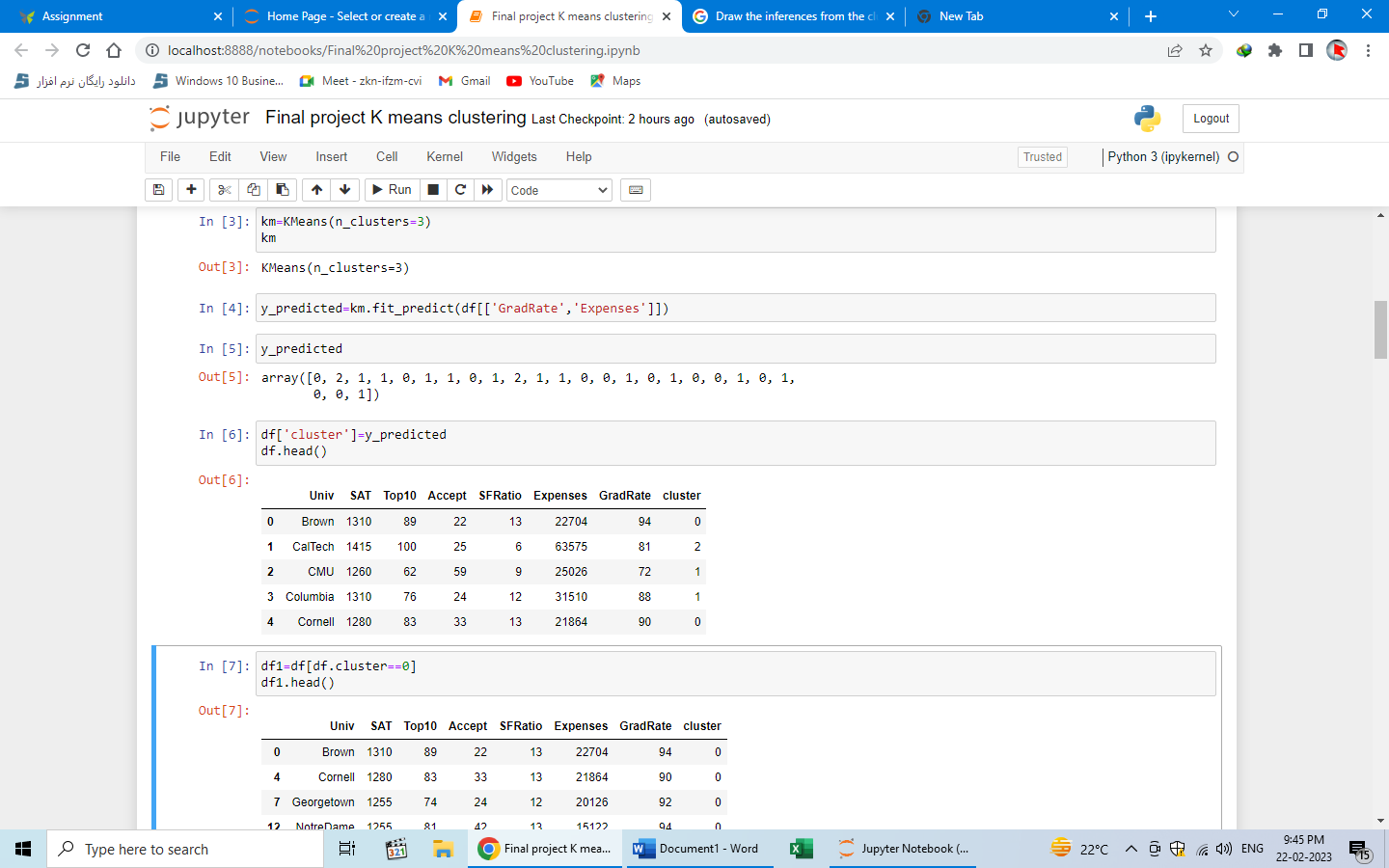
Final project

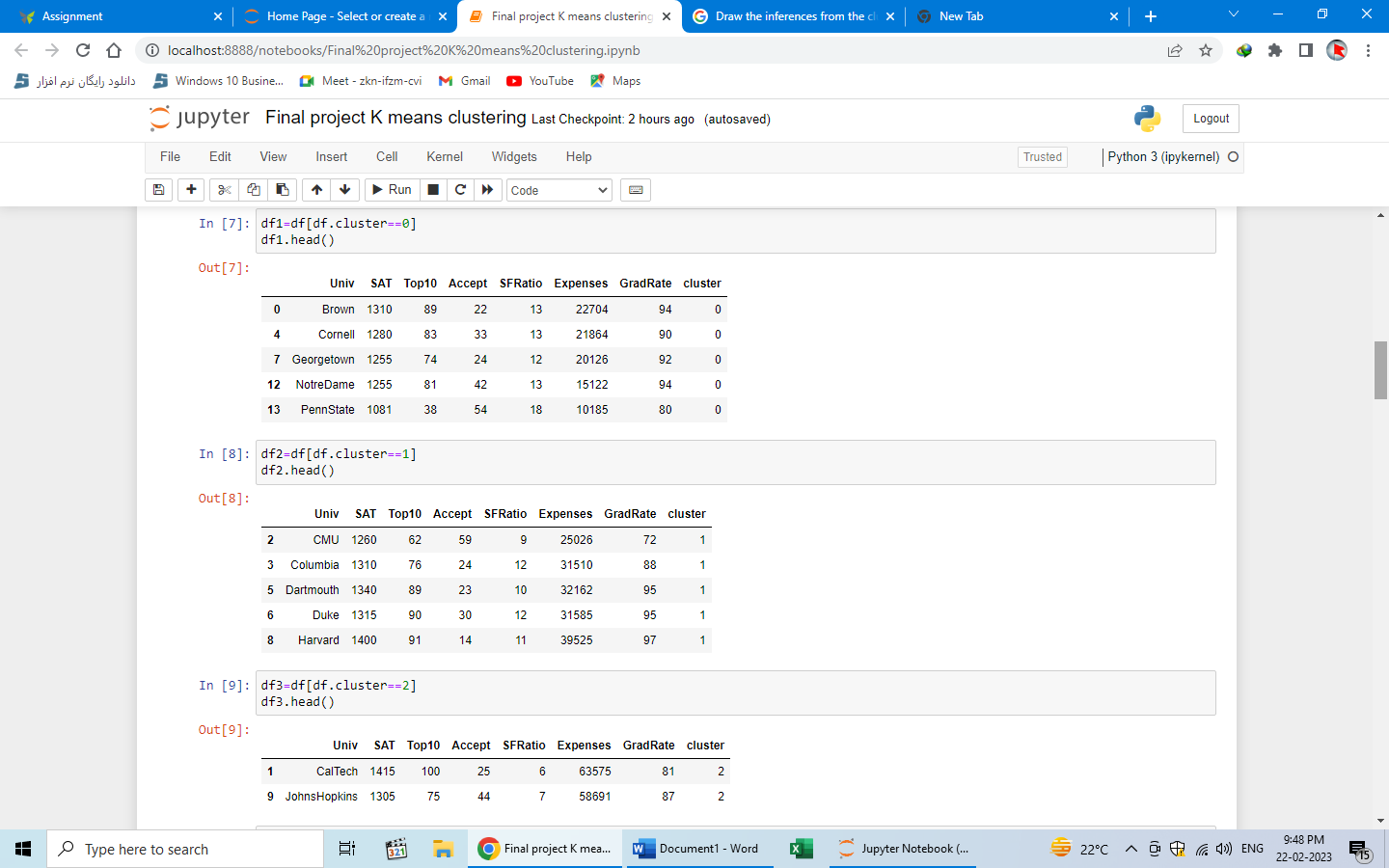
Assignment of unsupervised learning

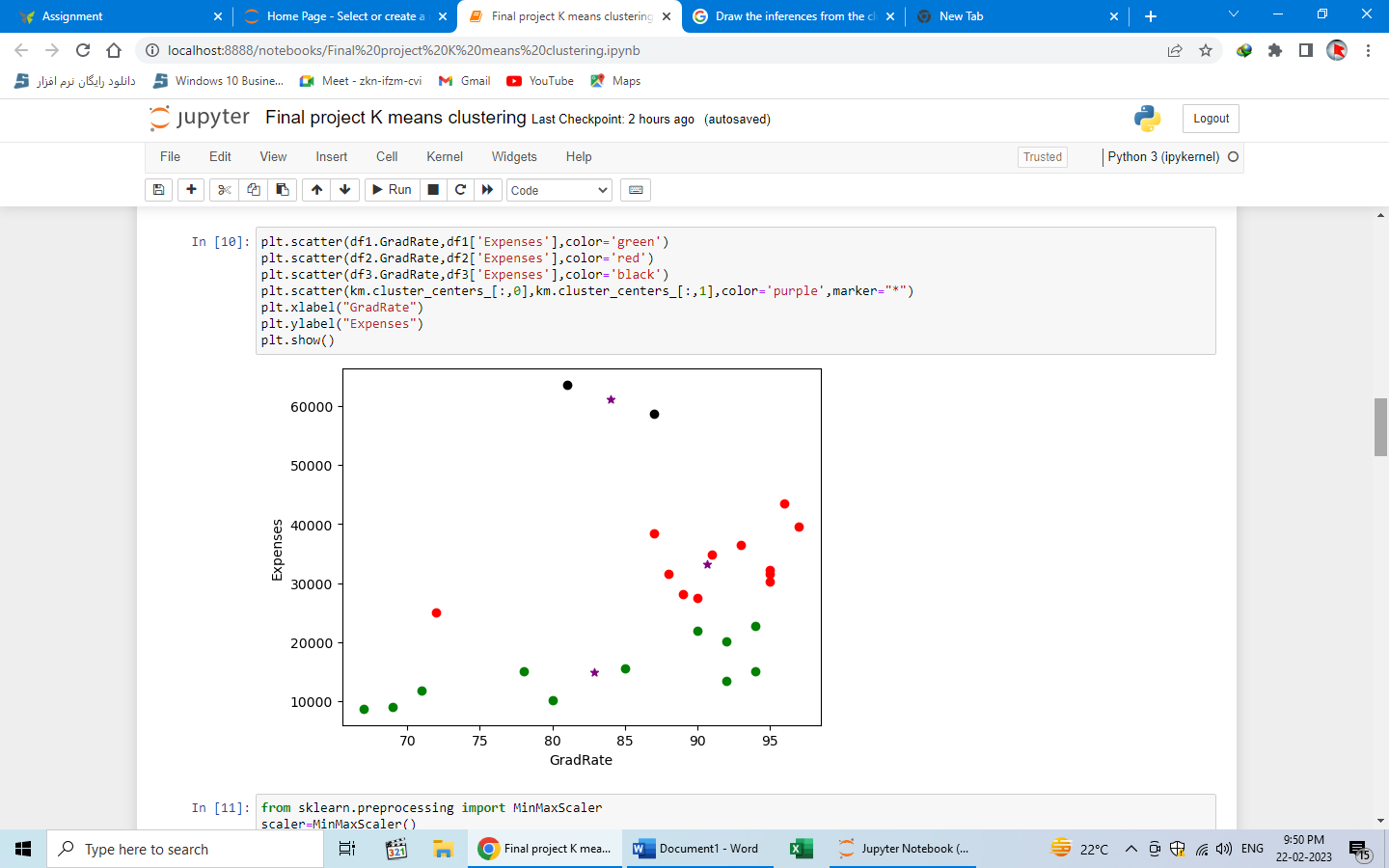
Q. Perform clustering (K means clustering) for the given data to obtain optimum number of clusters. Draw the inferences from the clusters obtained. (Universities (1).csv)

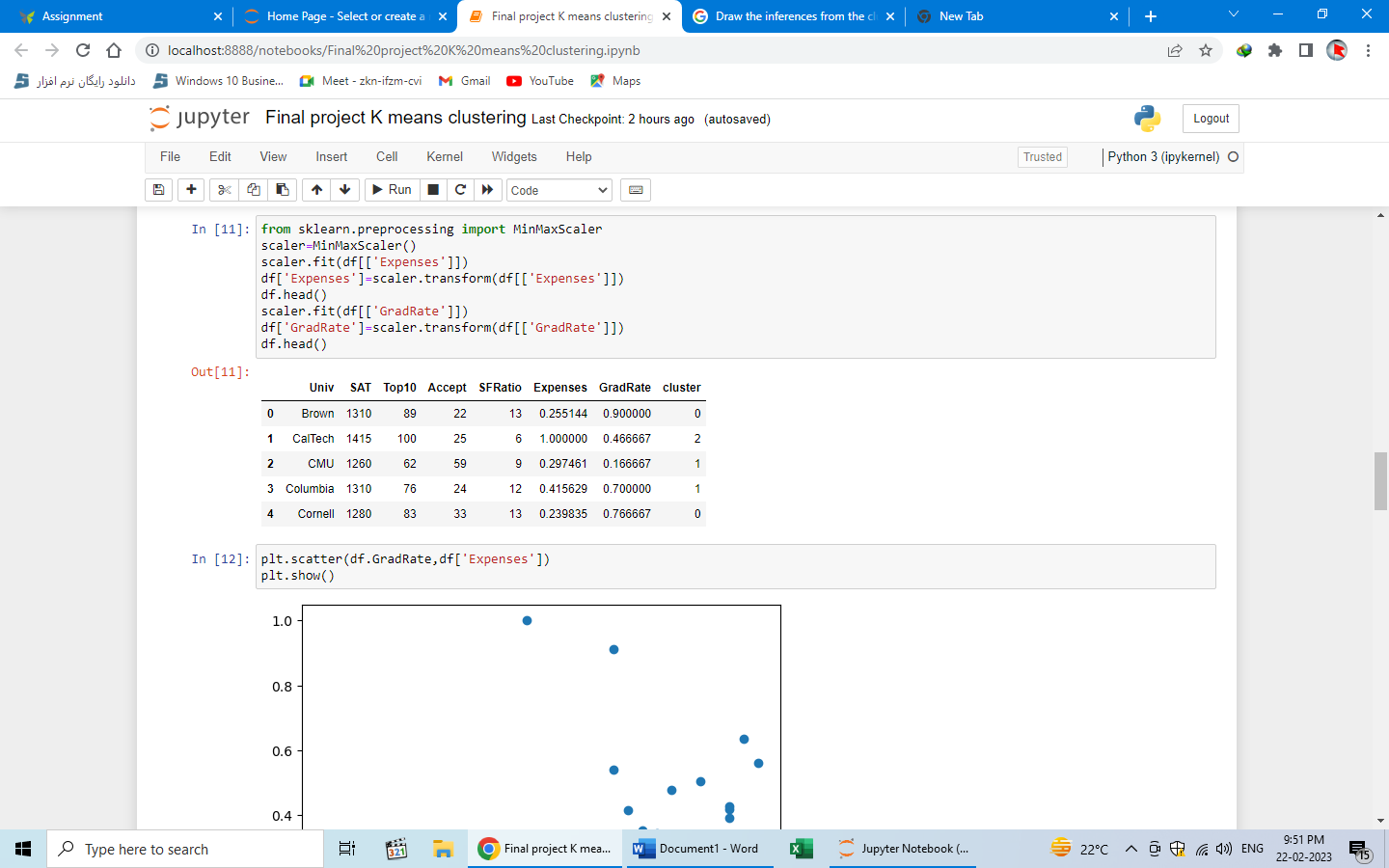


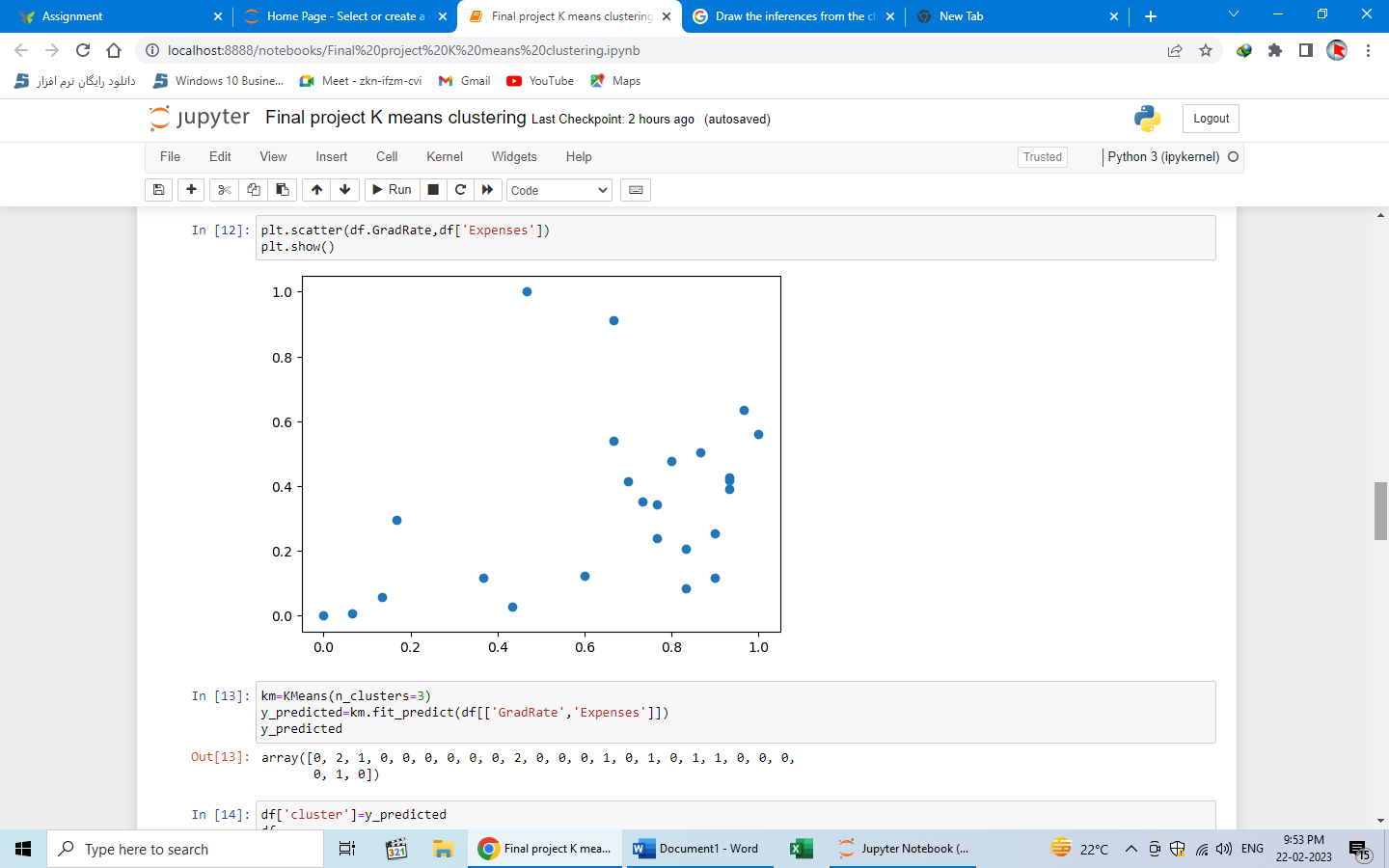


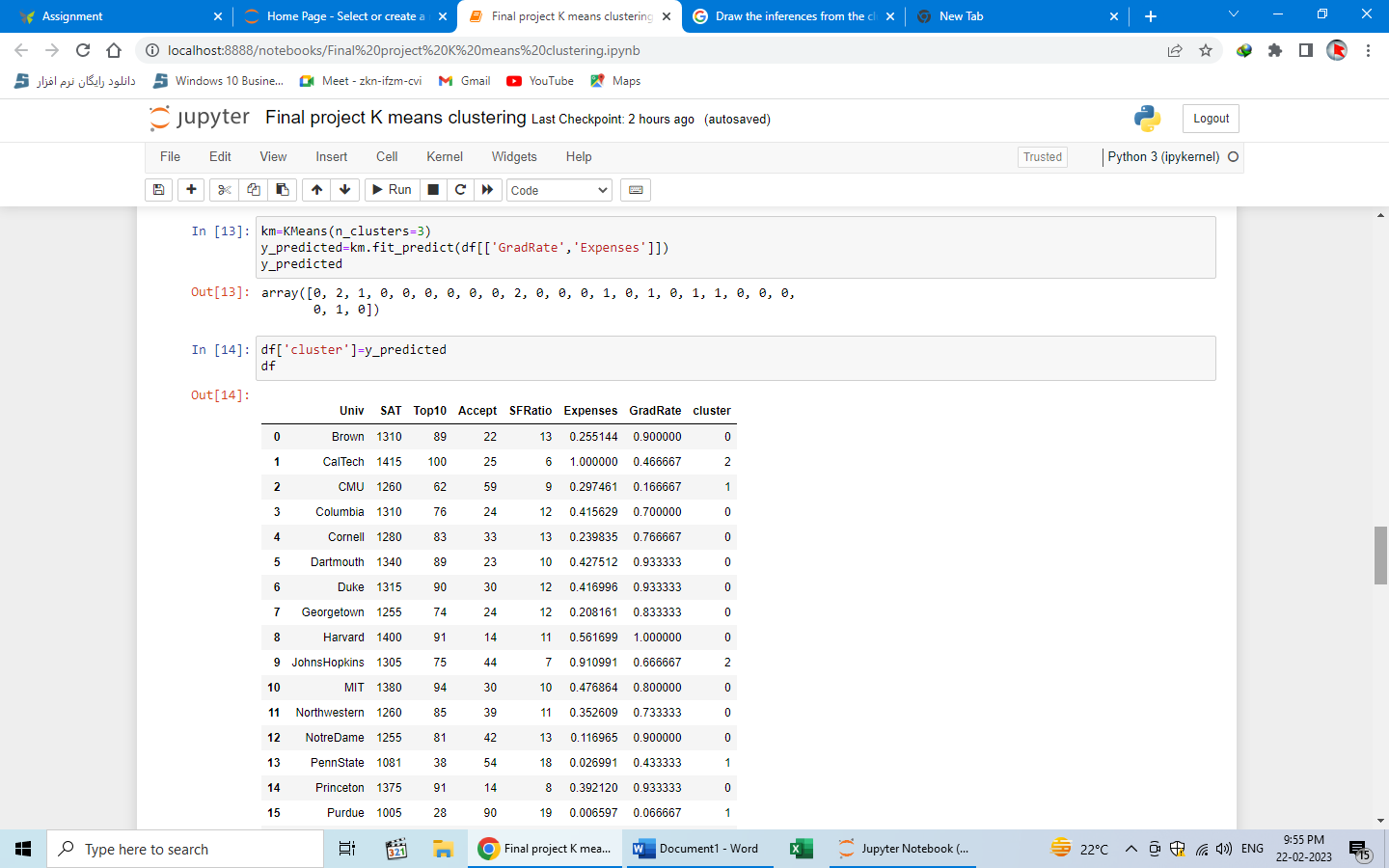


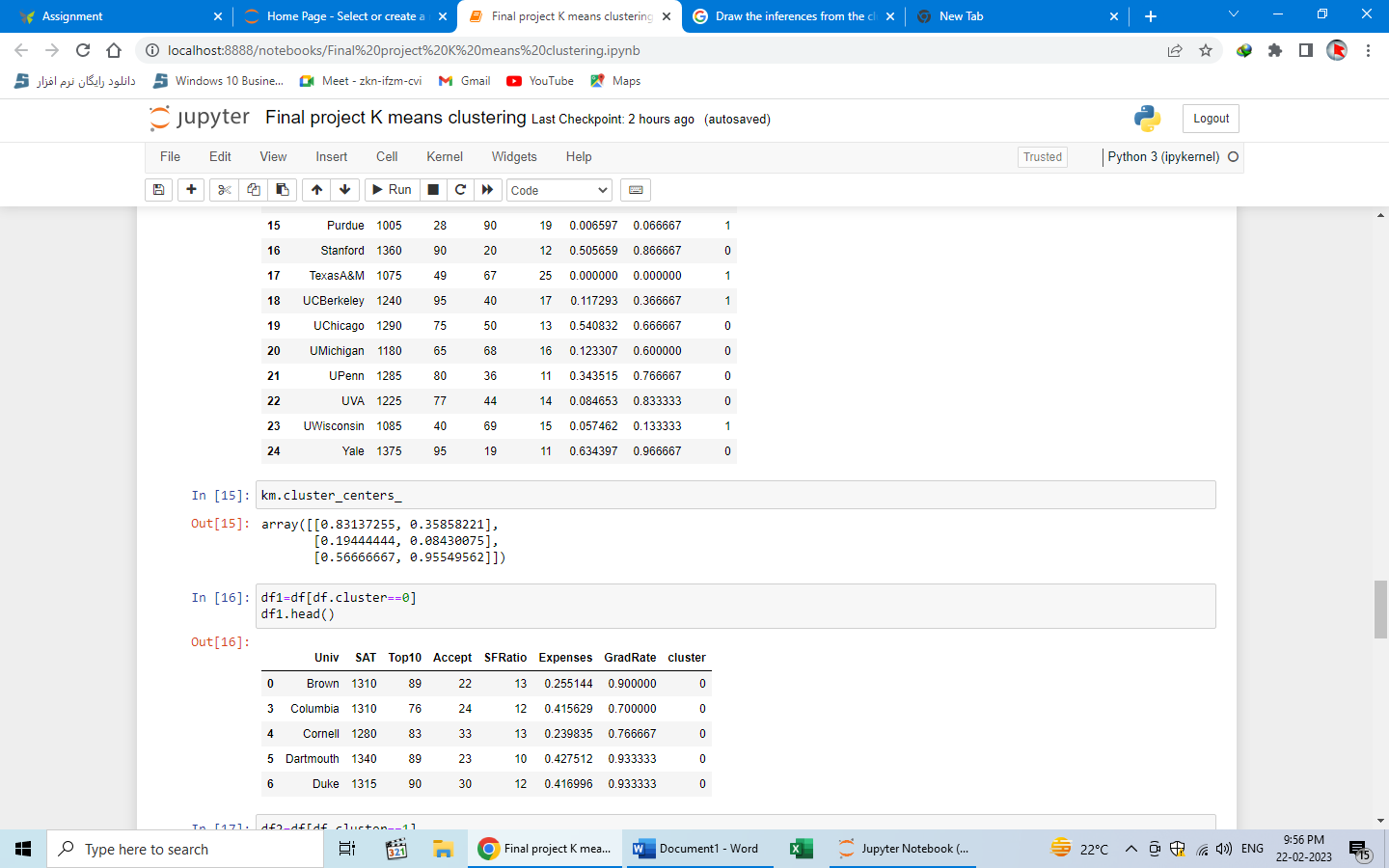


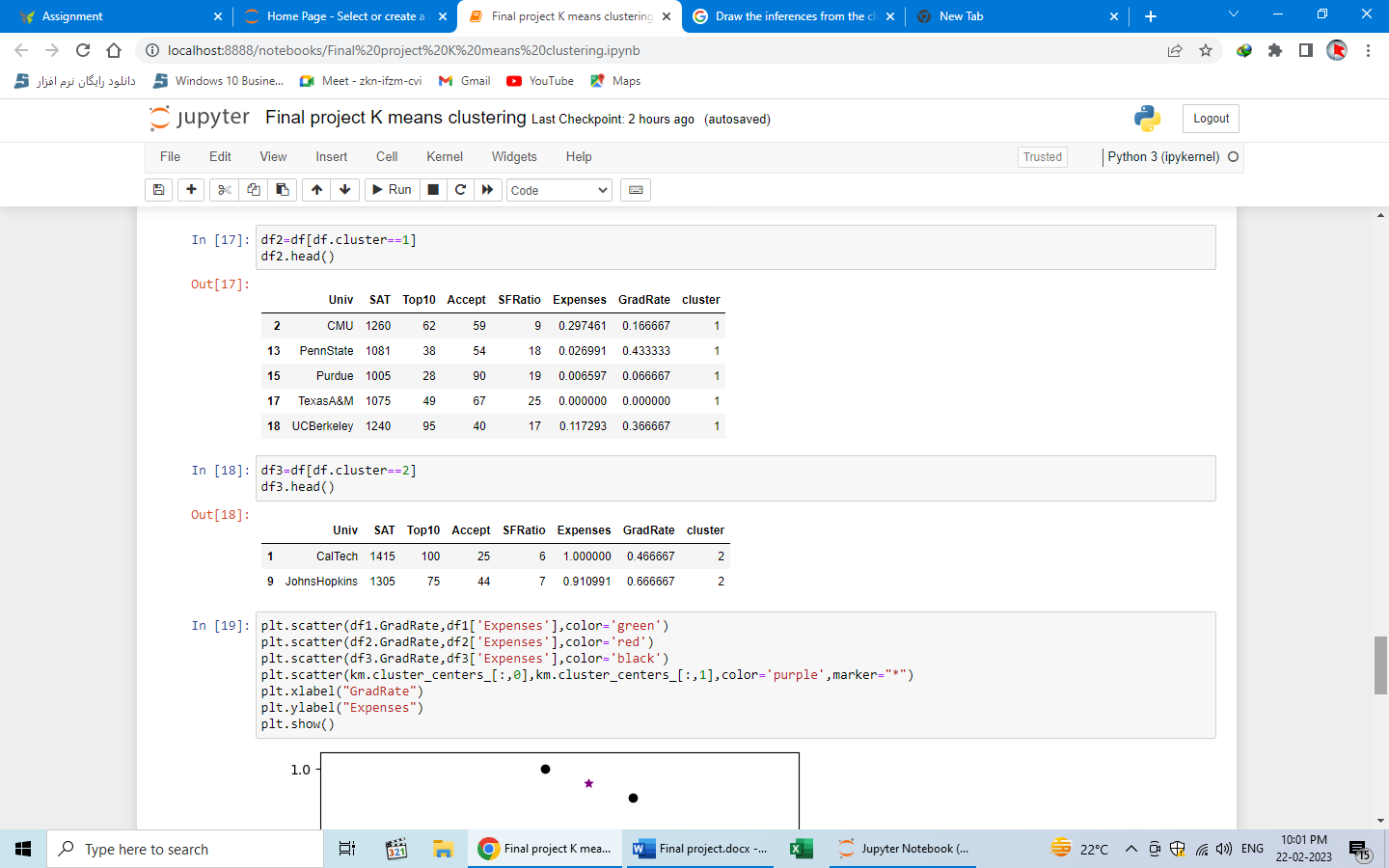


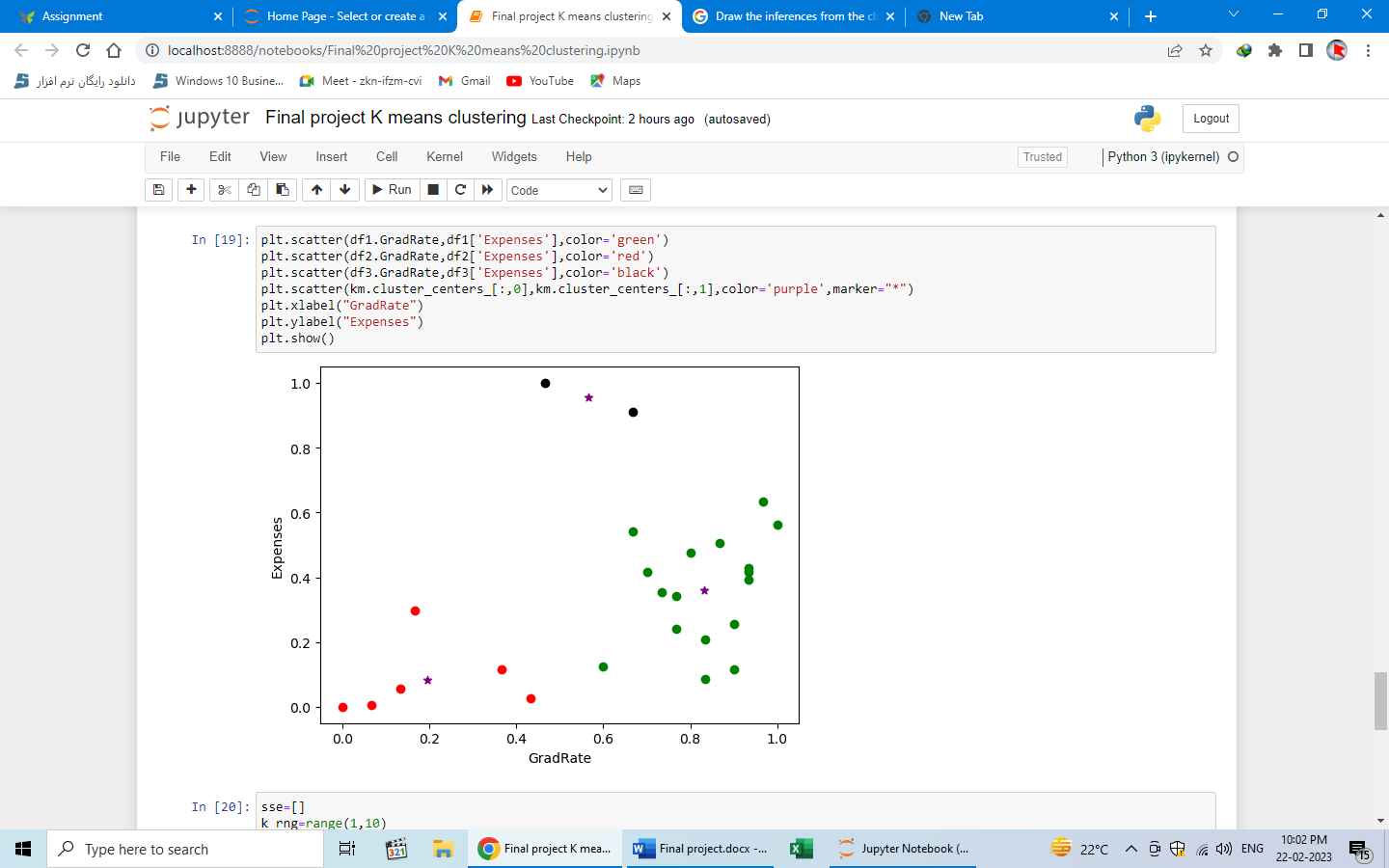


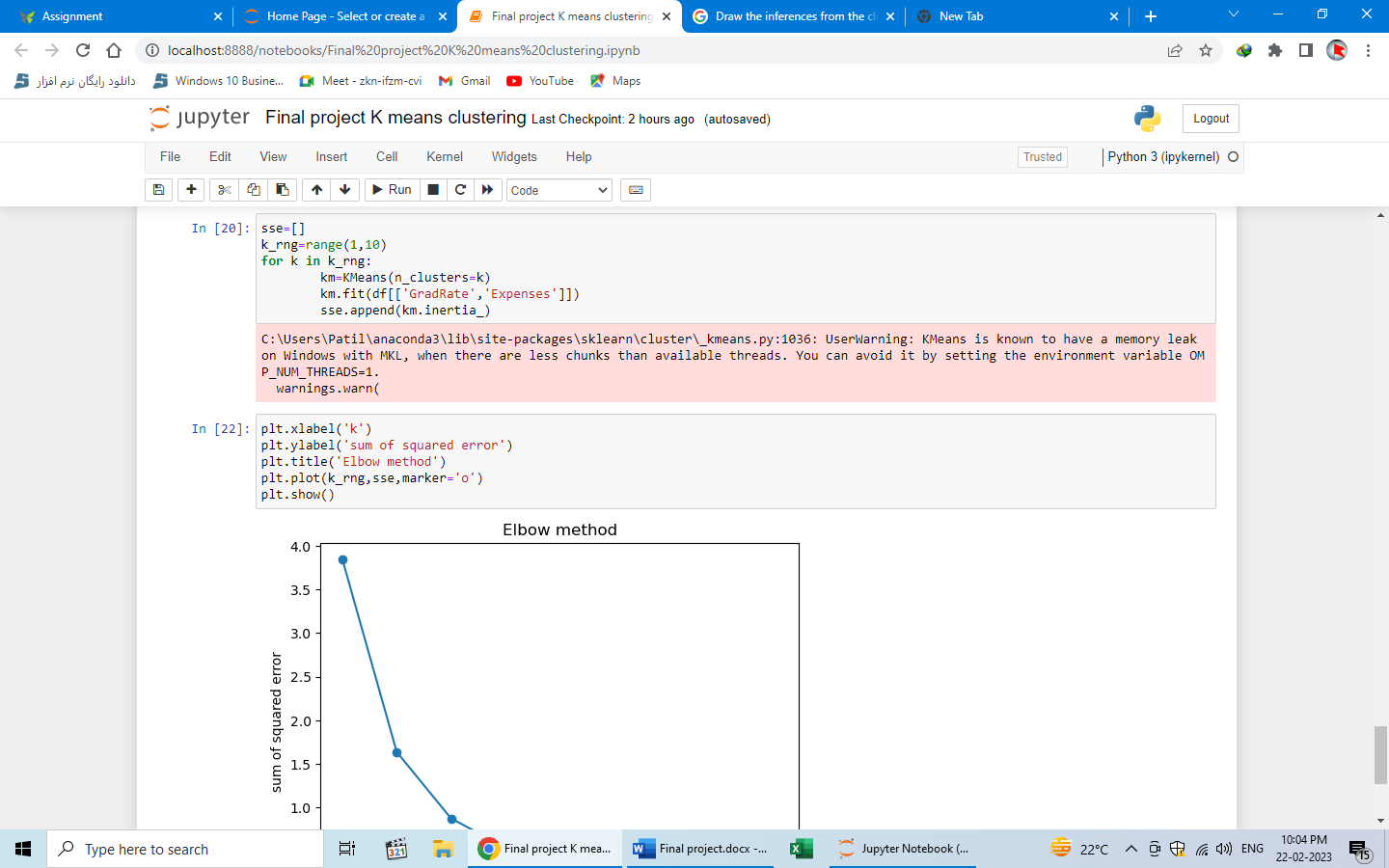


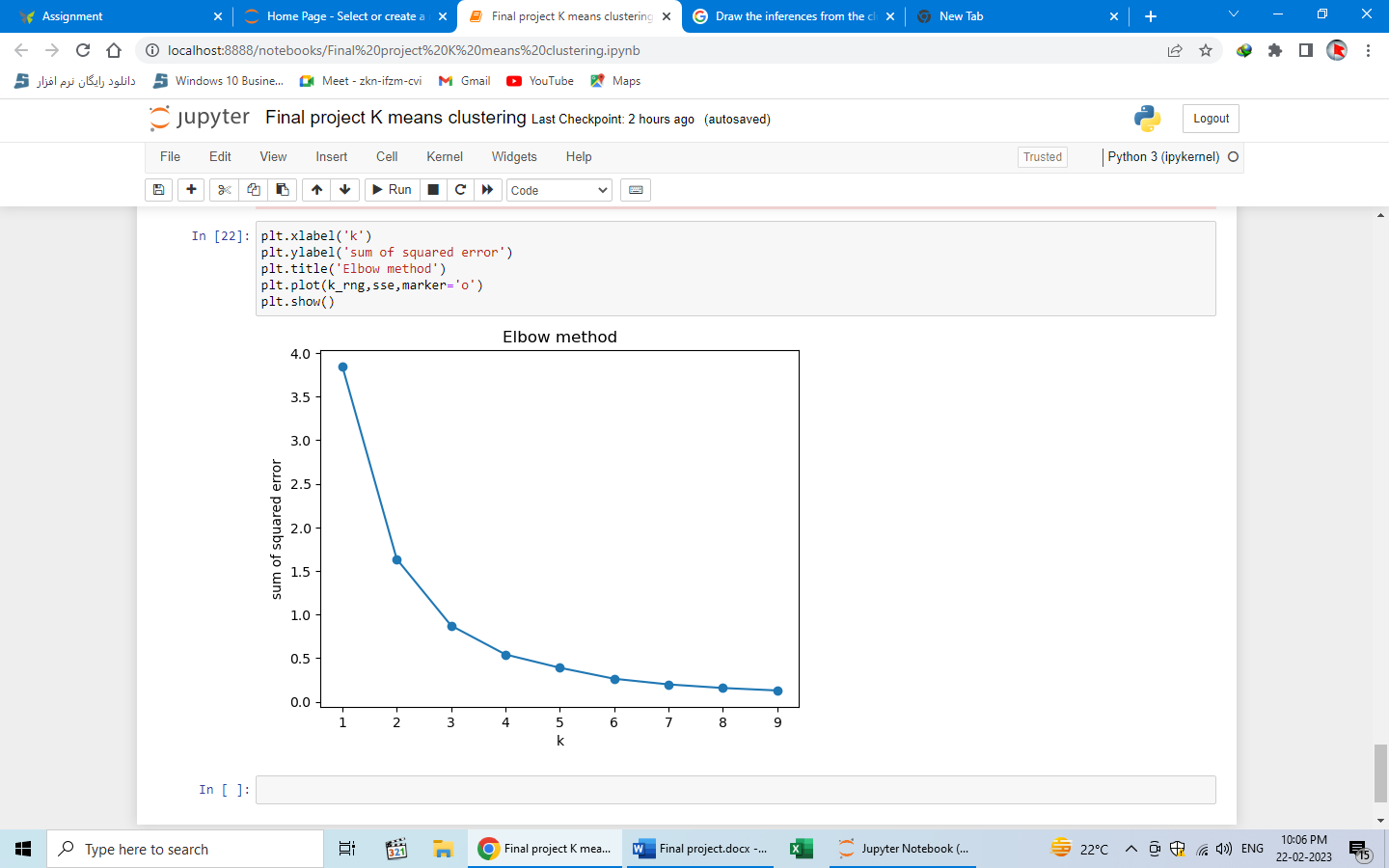












Conclusion:

K-means clustering is the unsupervised machine learning algorithm that is part of a much deep pool of data techniques and operations in the realm of Data Science. It is the fastest and most efficient algorithm to categorize data points into groups even when very little information is available about data. similar to other unsupervised learning, it is necessary to understand the data before adopting which technique fits well on a given dataset to solve problems. Considering the correct algorithm, in return, can save time and efforts and assist in obtaining more accurate results.

In this universities data set I am calculate that k means clustering The first step is to define the K number of clusters in which we will group the data. Let’s select K=3.then Centroid is the center of a cluster but initially, the exact center of data points will be unknown so, we select random data points and define them as centroids for each cluster. We will initialize 3 centroids in the dataset. Now thatcentroids are initialized, the next step is to assign data points *X*n to their closest cluster centroid *C*k and then In this step, we will first calculate the distance between data point X and centroid C using Euclidean Distance metric And then choose the cluster for data points where the distance between the data point and the centroid is minimum. centroids and the assignments of data points to correct clusters

K means clustering point is k3