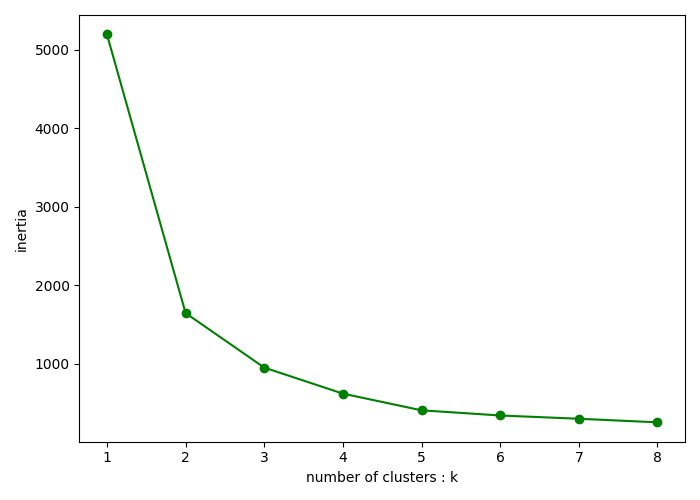
**Assignment 12 kmean**

## **1. take k = 3 and use k-means sklearn library routing for kmeans (random initialization and use the defaults). Take k = 1, 2, . . . 7, 8 and compute the distortion vs. k. Use the ”knee” method to find out the best k.**



According to the “knee” method, the best k is 3

## **2. for this optimal k, examine your clusters and for each cluster compute the percentage of ”green” and ”red” weeks in that cluster.**



0: green

1: red

## **3. does your k-means clustering find any ”pure” clusters (percent of red or green weeks in a cluster is more than, say,90% of all weeks in that cluster)?**

No, there is no cluster higher than 90%.