CS 6375

ASSIGNMENT \_\_\_6\_\_\_\_\_\_\_\_\_

Names of students in your group:

Pranathi Peri(pxp162530)

Nitin VasiReddy(nxv160230)

Number of free late days used: \_\_0\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_   
Note: You are allowed a **total** of 4 free late days for the **entire semester**. You can use at most 2 for each assignment. After that, there will be a penalty of 10% for each late day.

Please list clearly all the sources/references that you have used in this assignment.

**Steps to compile and Run the code:**

The src folder contains the following files:

Kmeans.java

TweetData.java

clusterTweets.java

Steps to compile and run:

To compile the files:

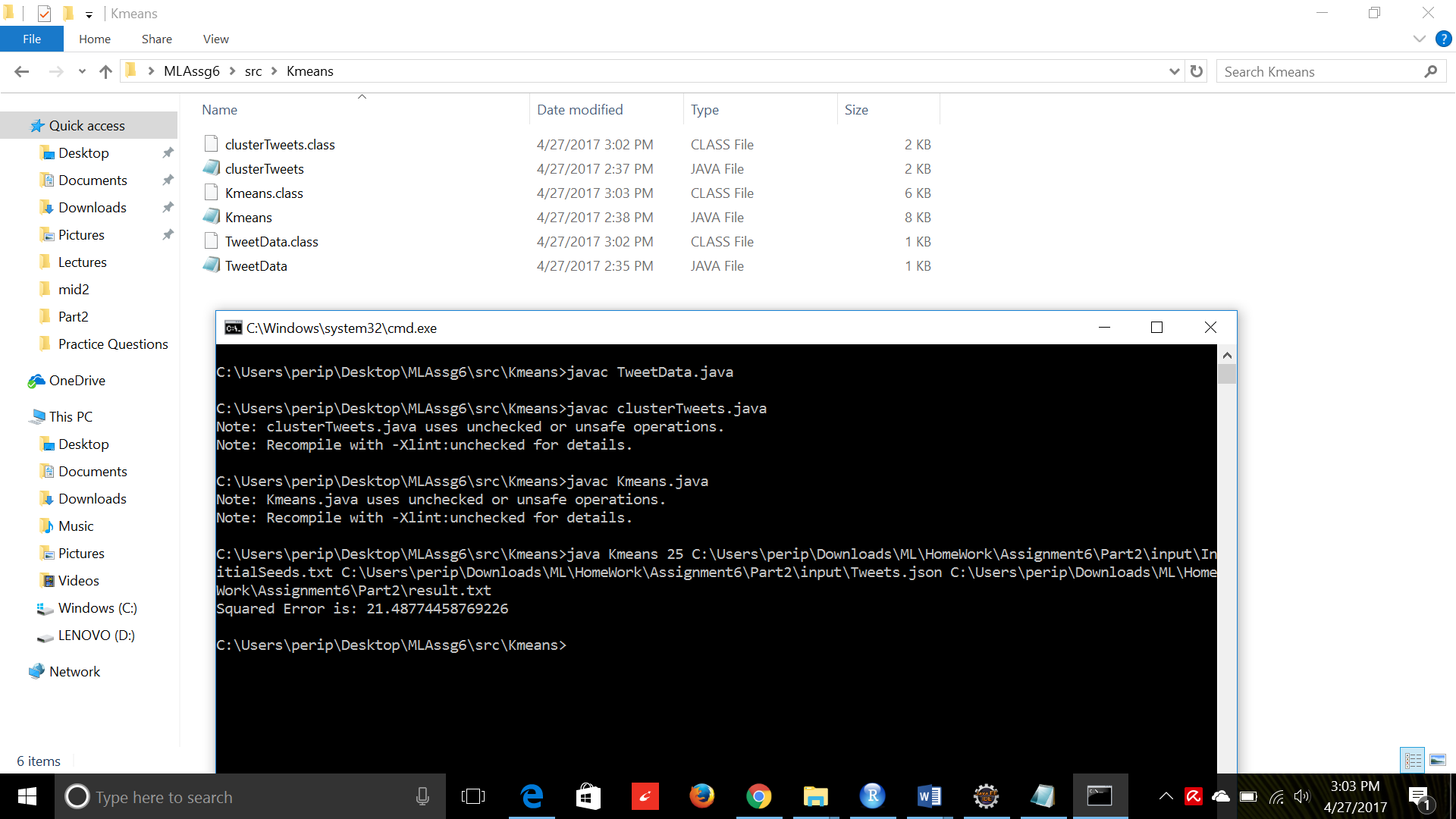
Traverse to the path where the source files are present. Shift+Right click and select Open command line window there.

javac <filename> (for all the three files)

The above command will create the .class files.

To run the file:

java Kmeans <numberOfClusters><initialSeedsFile><TweetsDataFile> <Outputfilename>



Sample Output:

ClusterNo TweetId

1 323906397609791488,323906397618196483,323906397735641088,323906397853073410,323906397962121216,323906398012461057,323906398230544385,323906398314438656,323906398352195585,323906398826164225,323906398993932289,323906399149109248,323906399295926273,323906399300100096,323906656318676993,323907087551836160,323907771256938496,323908455545049088,323908795254312962,323908795396943872,

2 323906483584655360,323906485249789952,323911610236293120,323915000567697409,323916051614138368,323920146454425600,323921510282702848,323923559799996417,323925264352546816,

The Squared Error is: 21.48774458769226