Assignment 6 – Report Part 2

The Tweet clustering using K means algorithm has been implemented in java.

There are 2 java files namely

1. TweetsKMeans.java
2. Tweet.java

To compile the program:

> javac TweetsKMeans.java –cp "javax.json-1.0.jar"

To run the program (Windows):

> java –cp ".;./javax.json-1.0.jar" TweetsKMeans 25 InitialSeeds.txt Tweets.json <output_file_name>

We need to have the javax.json-1.0.jar, Tweets.json in the same folder as in KMeans_Algorithm.java

Output:

SSE for the given tweets is: 35.08911978506978