Homework 3

January 18, 2018

- Find the top 100 words of the file quixote.txt with map-reduce.

- Find the number of words that start with each letter of the English alphabet in the file quixote.txt.

- Write a pySpark program that implements the social network friendship recommendation algorithm, which was discussed in class, using the data soc-data.txt. You may need to use –driver-memory 8G to set the runtime memory to 8GB. We run