For both problem, we create RDD use local file. Please put the dataset file in the current directory when executing the code.

For problem1, paste the code to scala shell and then execute it.

For problem2, please use spark-submit run the .jar file or pack the .scala file by sbt.

The class name is Problem2 and the jar file name is spark-scala-wordcount-network-assembly-1.0

So please input ./spark-submit --name problem2 --class Problem2 spark-scala-wordcount-network-assembly-1.0.jar from terminal to run it.