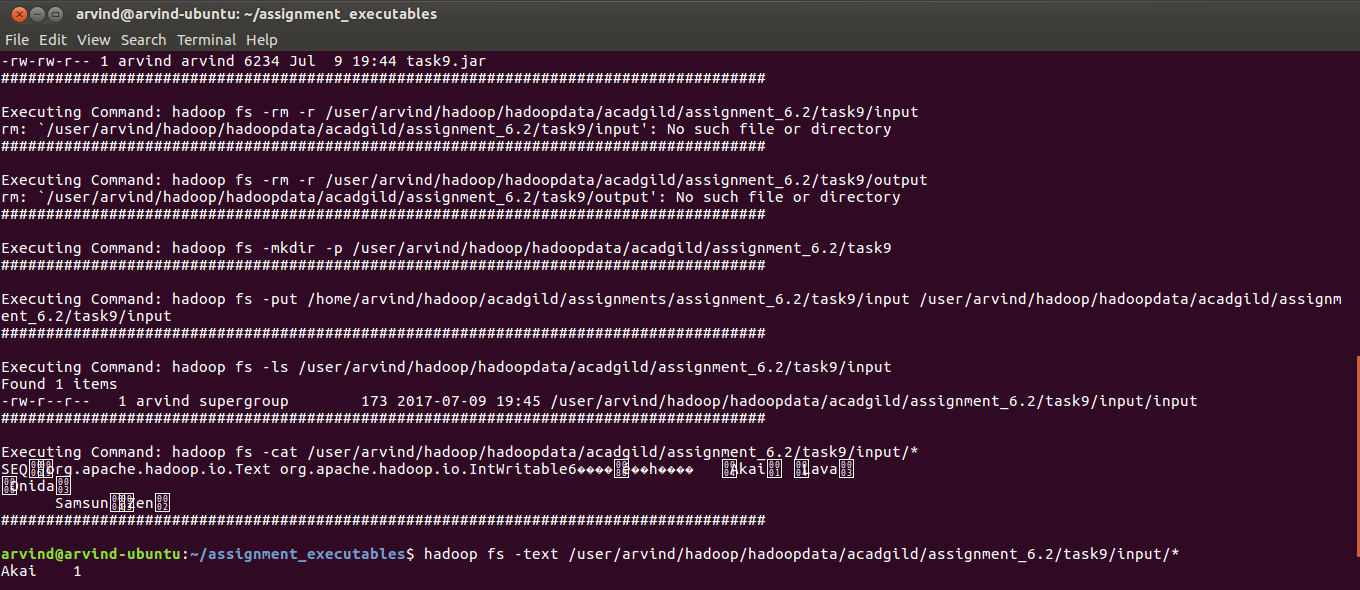
The map reduce program for this task takes as input the sequence file output generated by an earlier assignment. In this task, we are sorting the records in the input file based on the total units sold using multiple reducers

**make\_remote\_dirs\_and\_copy\_input.sh take 2 inputs, the assignment number and task number..**

1. It verifies if the input file and the jar for the assignment are kept in the appropriate location in the local file system using ls -l command.
2. Delete input and output directories in the hdfs for this assignment and task if already exists using hadoop fs -rm -r command(delete the folder and all its contents recursively)
3. Create a directory in HDFS for this assignment and task using hadoop fs -mkdir -p command
4. Copy the input from local file system to hdfs using hadoop fs -put command
5. Finally verify using hadoop fs -ls command if the input file has been copied into the HDFS

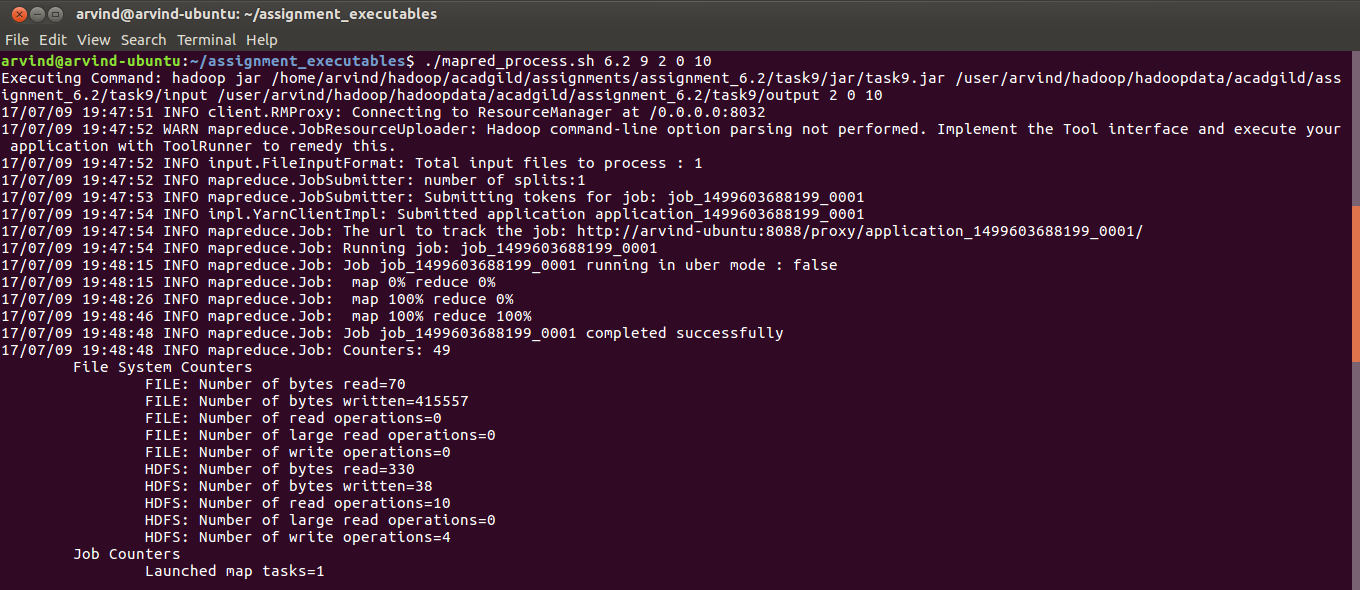
Note: A sequence file cannot be read by using the hadoop fs -cat command. Instead you need to use a



**mapred\_process also takes assignment number and task number as input. But for this task it takes three more inputs in this order – number of reducers, minimum number of total units sold and the maximum number of total units sold**

1. It executes the JAR from the local file system using hadoop jar command

Note: Here the number of reducers is 2, the minimum number of total units sold is 0 and the maximum number of total units sold is 10

****

**output.sh also takes assignment number and task number as input**

1. It lists all the files in the output directory for this assignment and task job in HDFS. (hadoop fs -ls)
2. It displays the content of all the output files (hadoop fs -cat)
3. Finally it copies the output files from HDFS to local file system

Note: The records in totality from all output files are sorted in ascending order of total units sold. Because of the custom partitioner logic, all keys less than 5 are sent to 1st reducer(pls refer screenshot below where the contents of part-r-00000 are displayed) and the keys greater than or equal to 5 are sent to the 2nd reducer (pls refer screenshot below where the contents of part-r-00001 are displayed)

