

GILDART HAASE SCHOOL OF COMPUTER SCIENCES AND ENGINEERING

University College: Arts – Sciences – Professional Studies

Fairleigh Dickinson University

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

*Student’s Name: Professor:*

Desham,RajeshReddy (1682826) *Alex Rudniy,Ph.D* **LastName, First Name(Student I.D)**

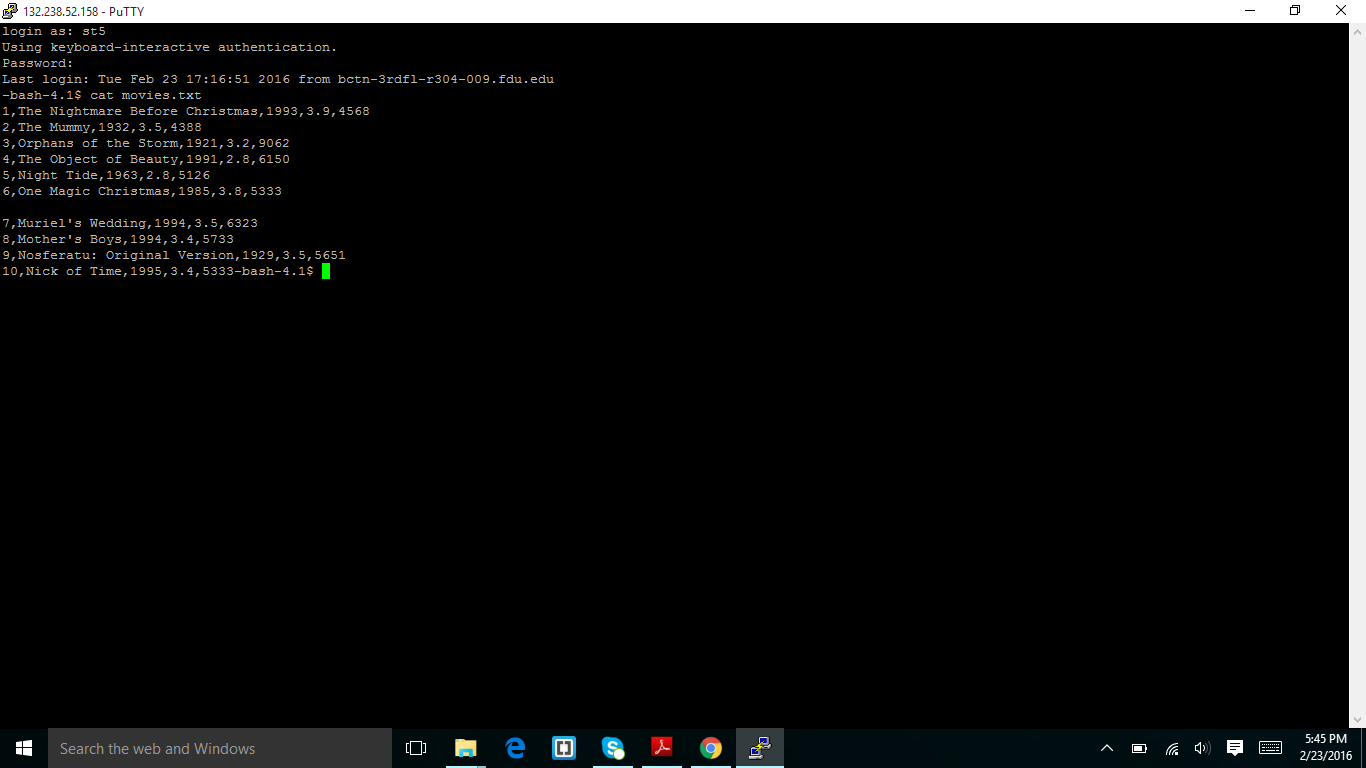
**CSCI 6885\_81**

**Big Data Analytics with Hadoop and R**

**SPRING 2016**

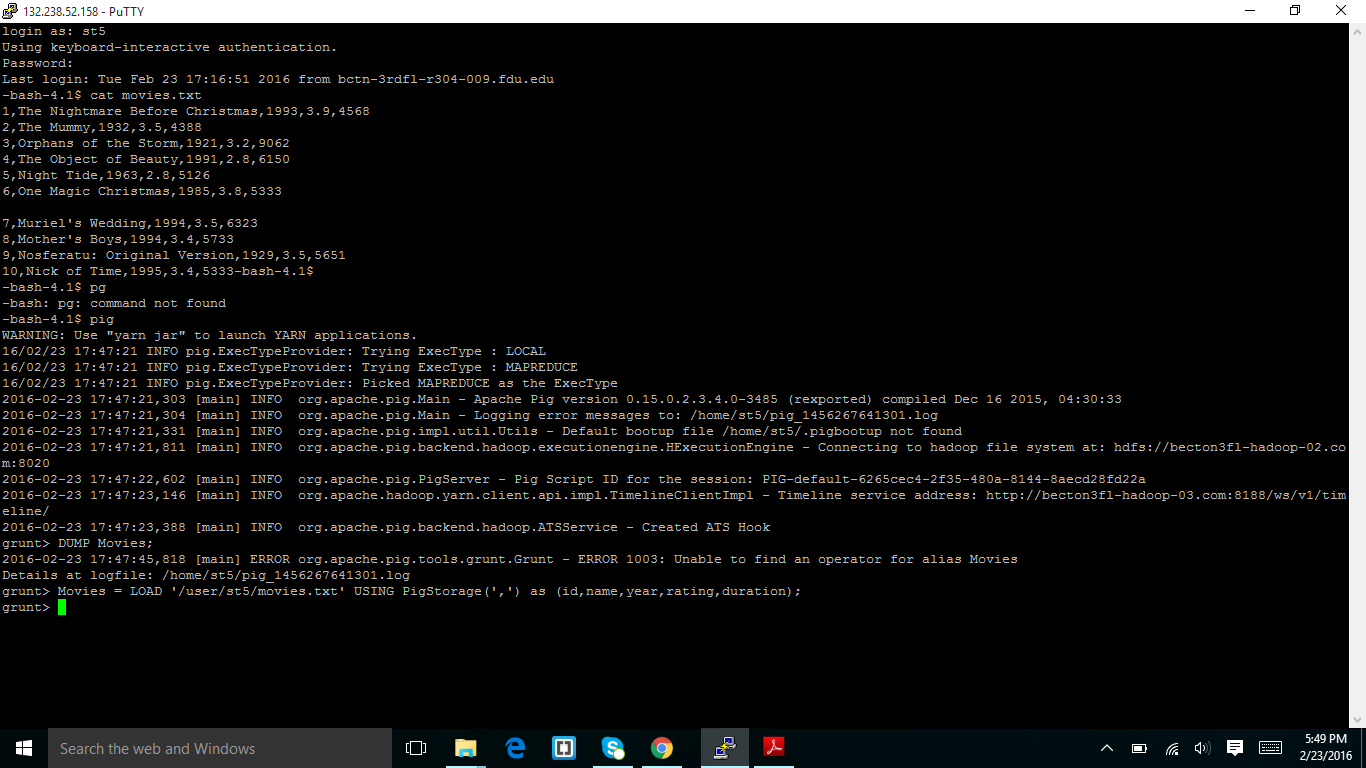
Assignment-4

**STEP1:**



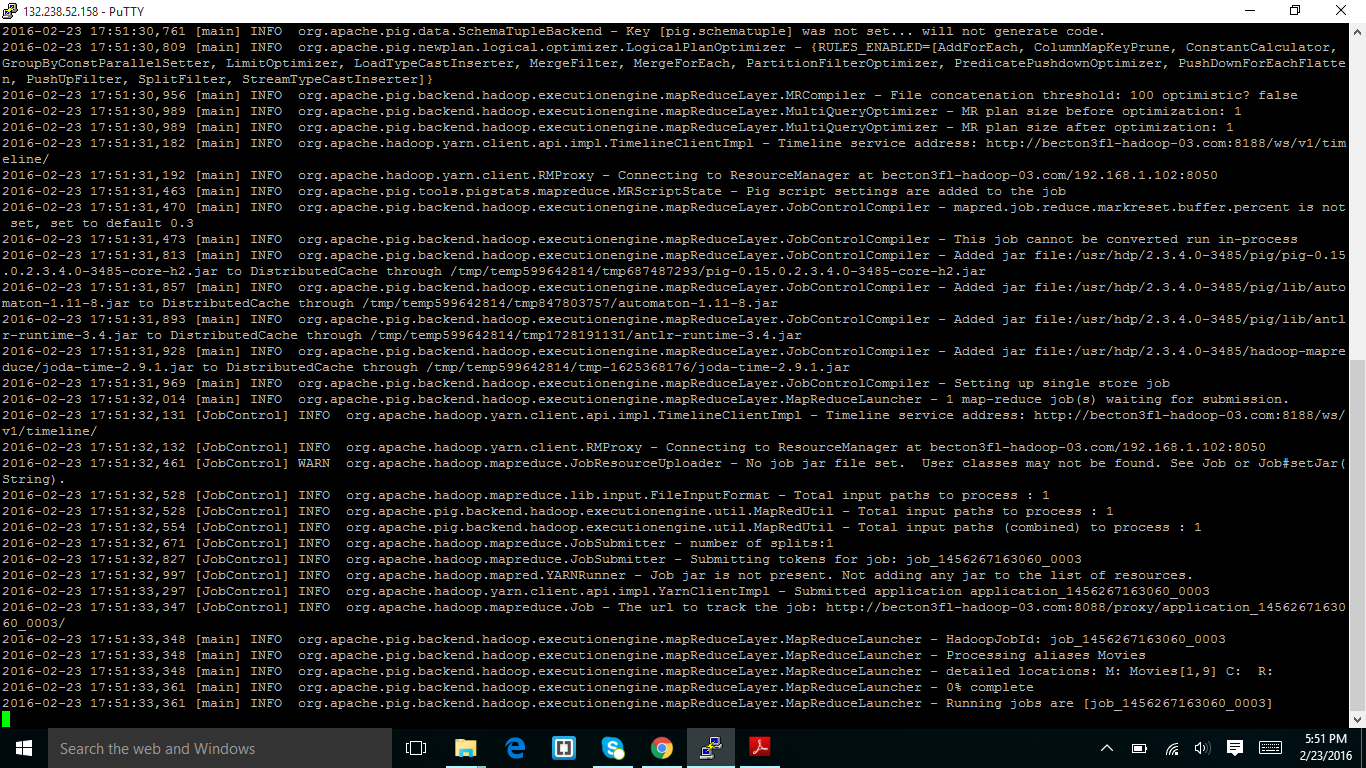
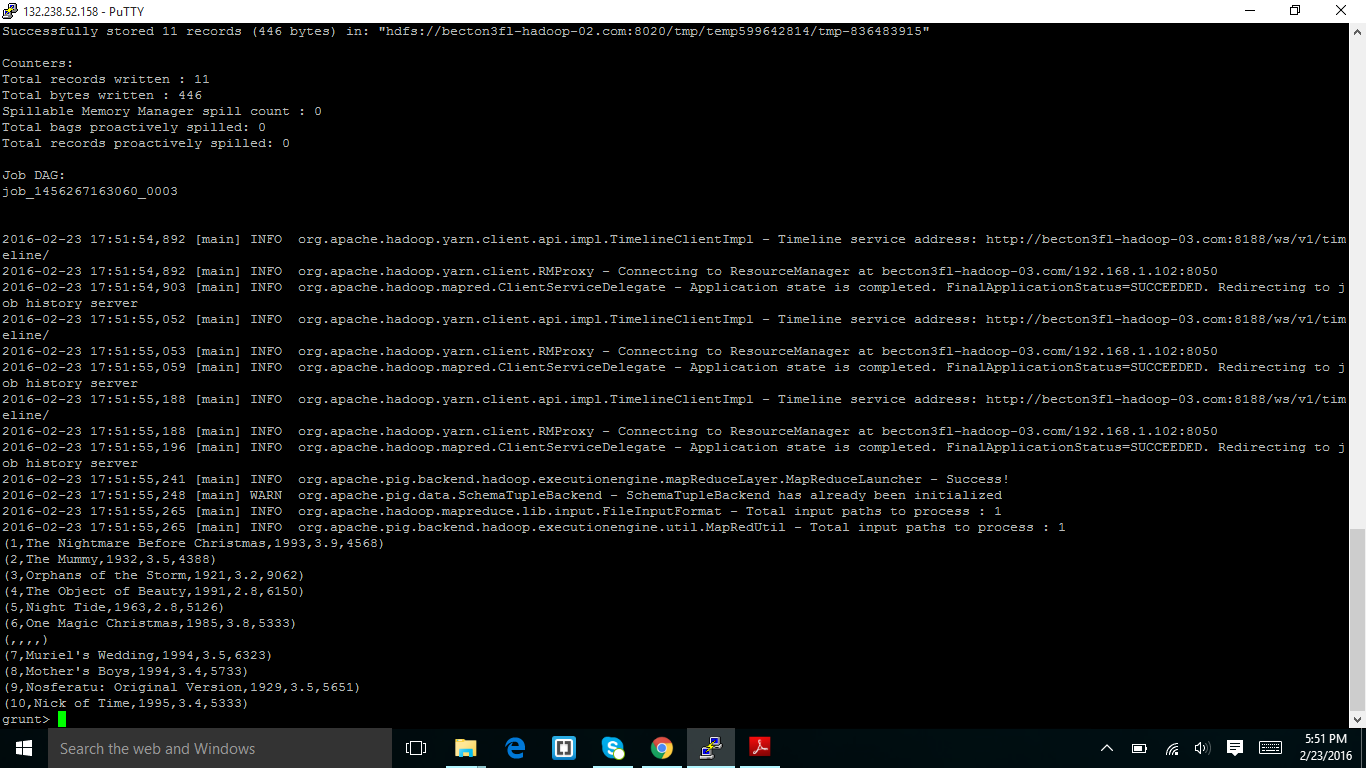
I had login in to the putty and I created a movies.txt file in the linux machine and I displayed it by using the command (cat movies.txt).

**STEP2:**



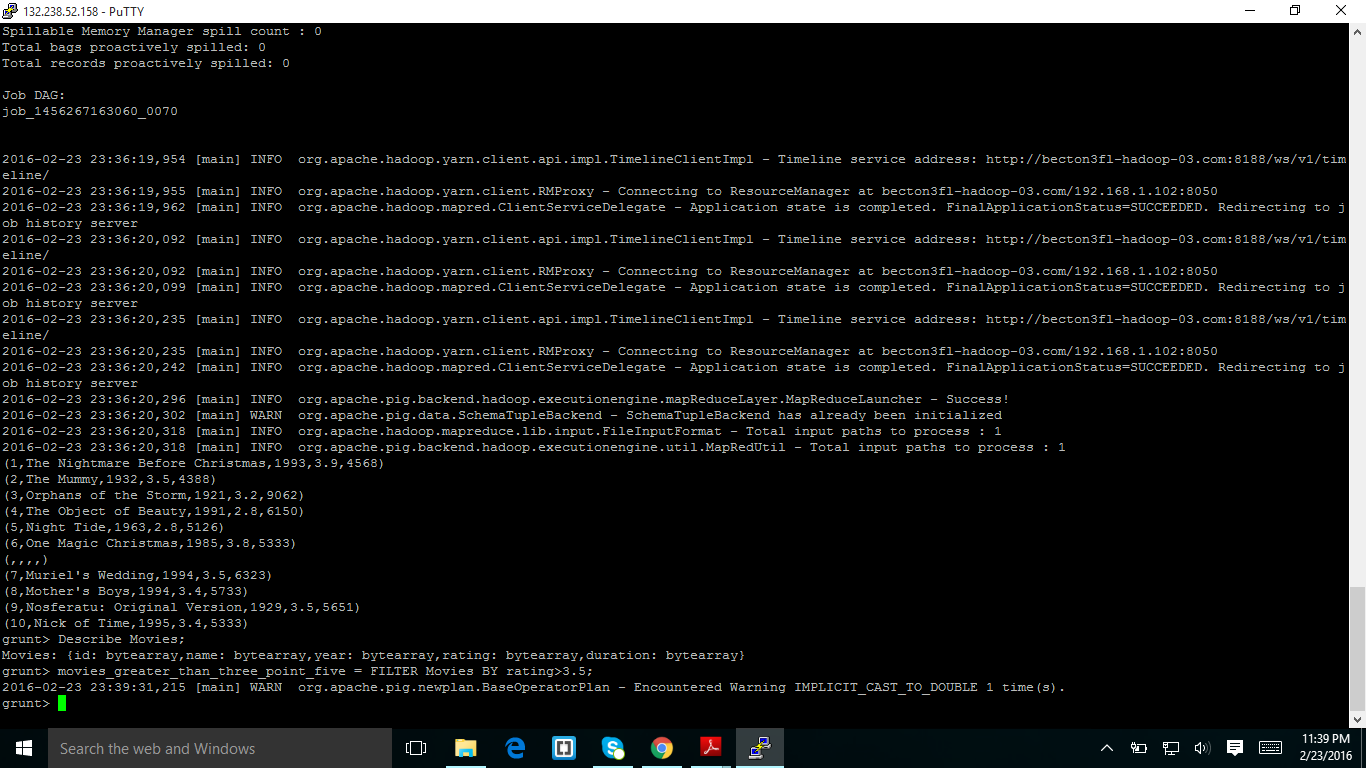
Before the screen shot I load the movies.txt into the HDFS directory and in above the screen shot I started pig and I load the movies.txt using pig storage with id ,name,year,rating,duration.

**STEP3:**



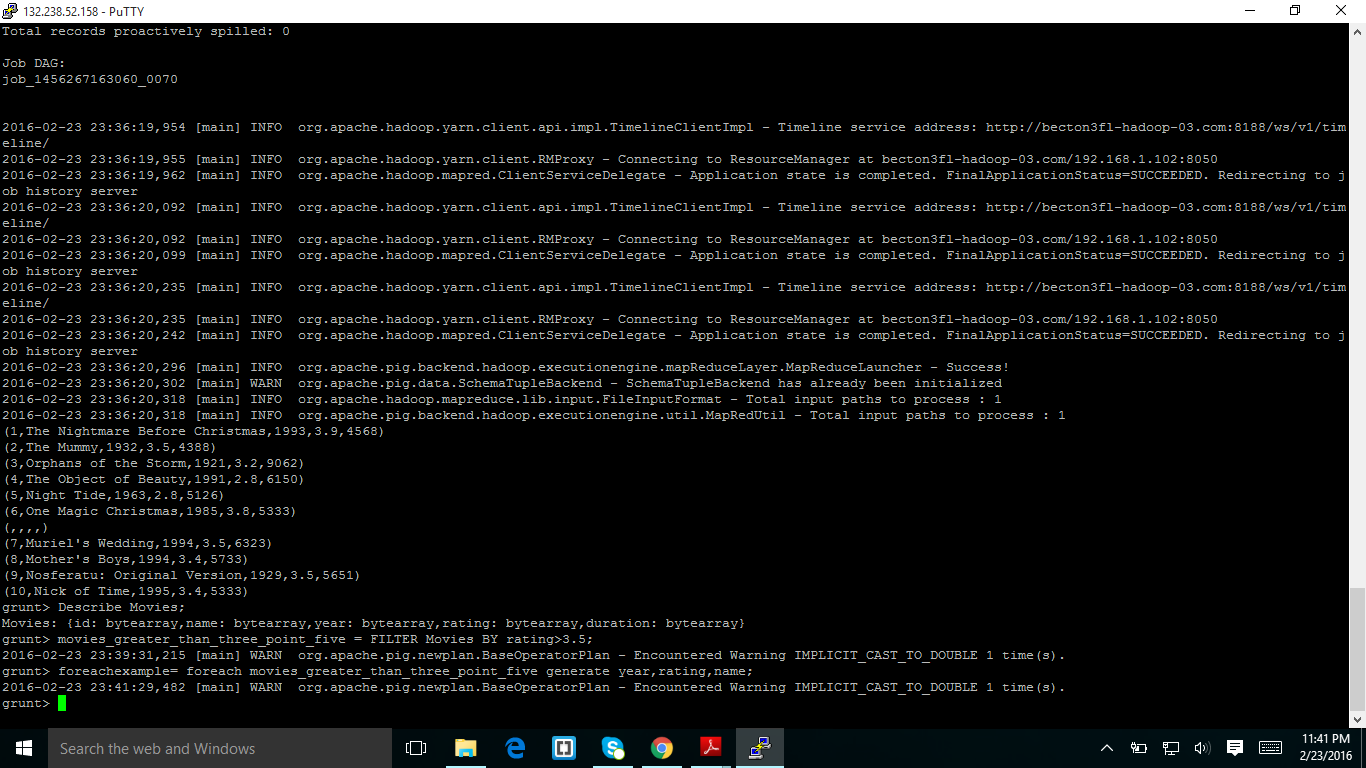
In above screen I dumped the movies using the command(DUMP movies.txt) and it displayed the 11 results.

**STEP4:**



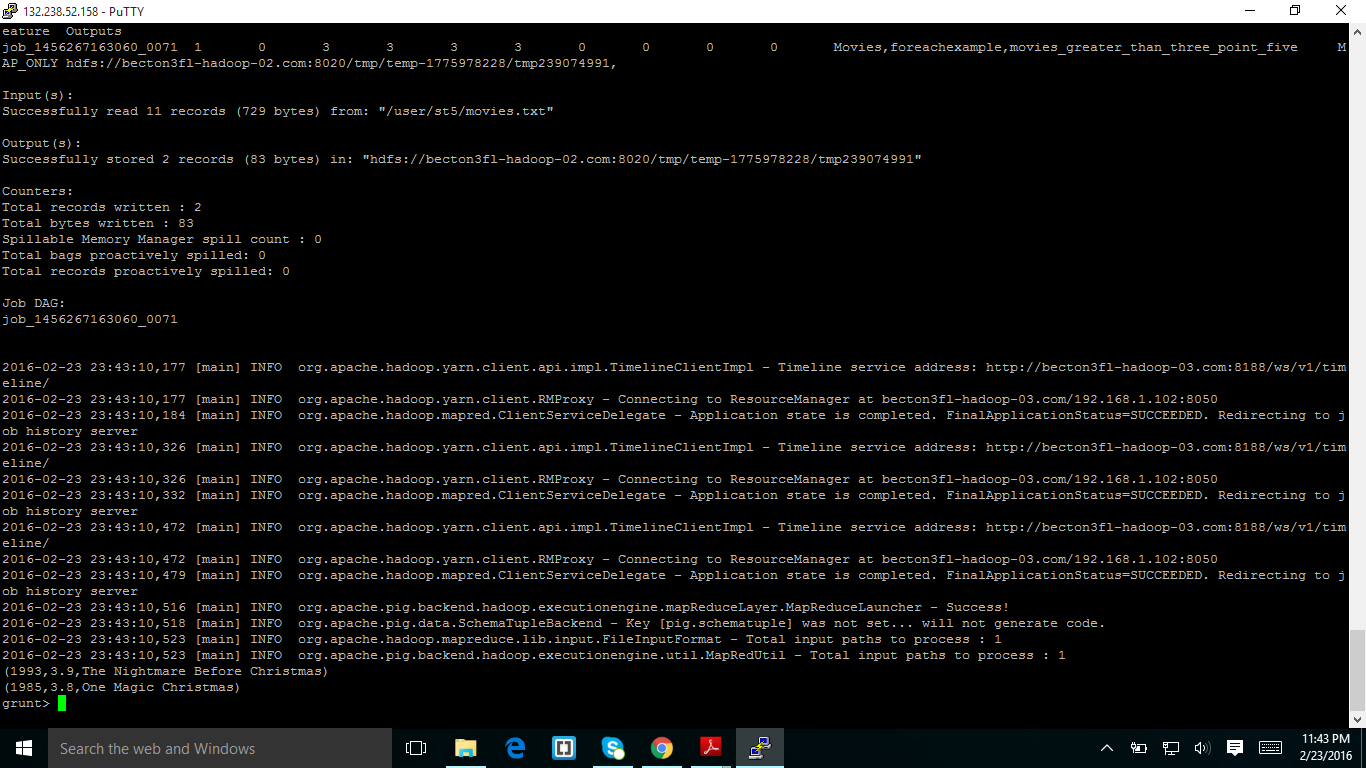
In above screen shot I describe the movies variable with the command describe and it has displayed the attribute types.

**STEP5:**



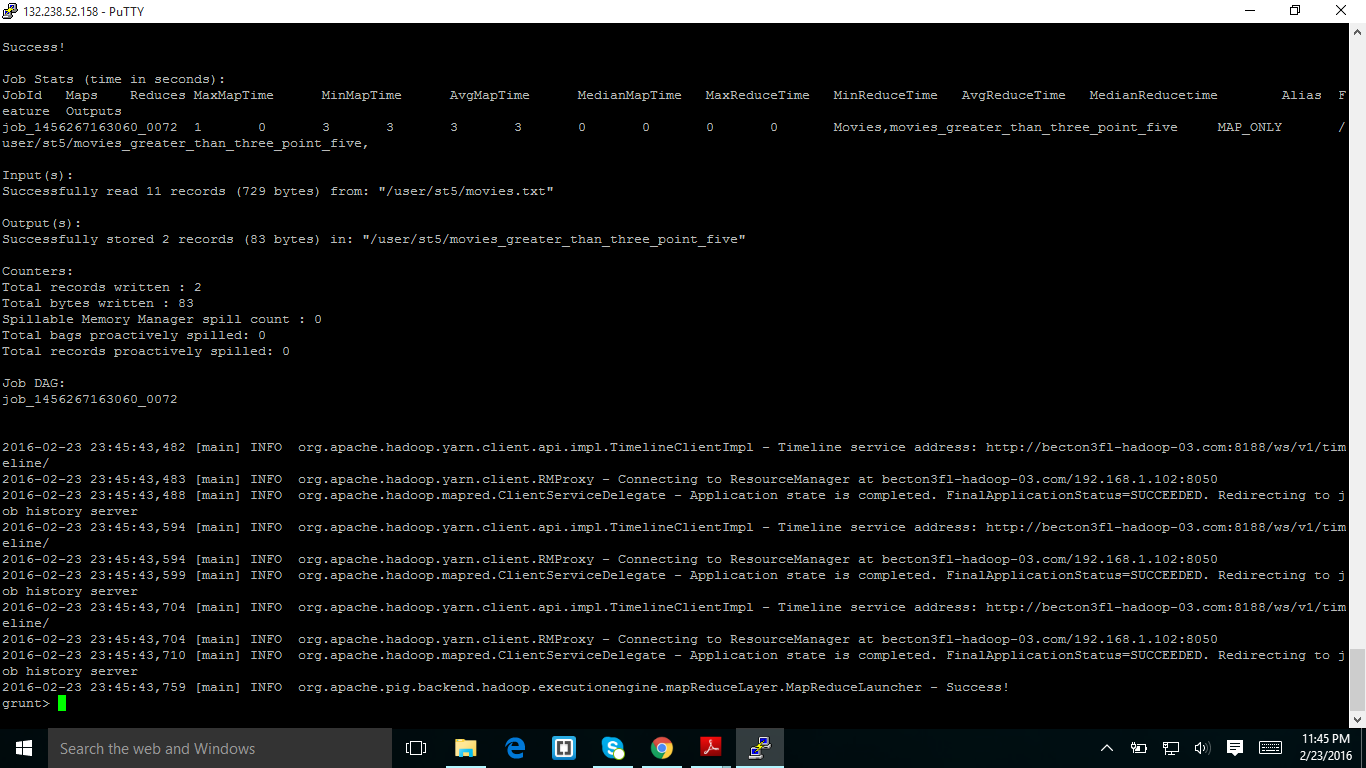
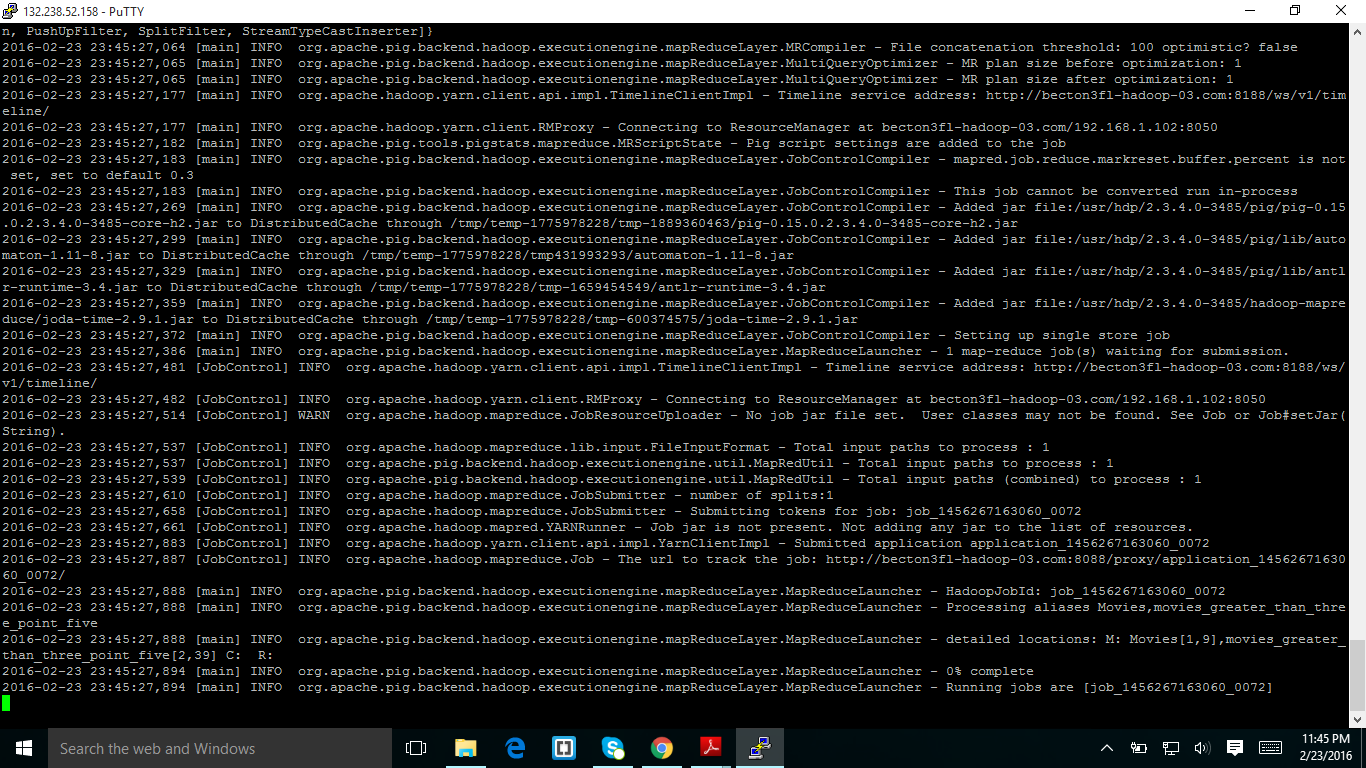
In above screen shot I filter the data for movies which are greater than 3.5 and from the variable From the variable movies\_greater\_than\_three\_point\_five I extract the values of year,rating,name and I stored in the other variable name foreachexample.

**STEP6**:



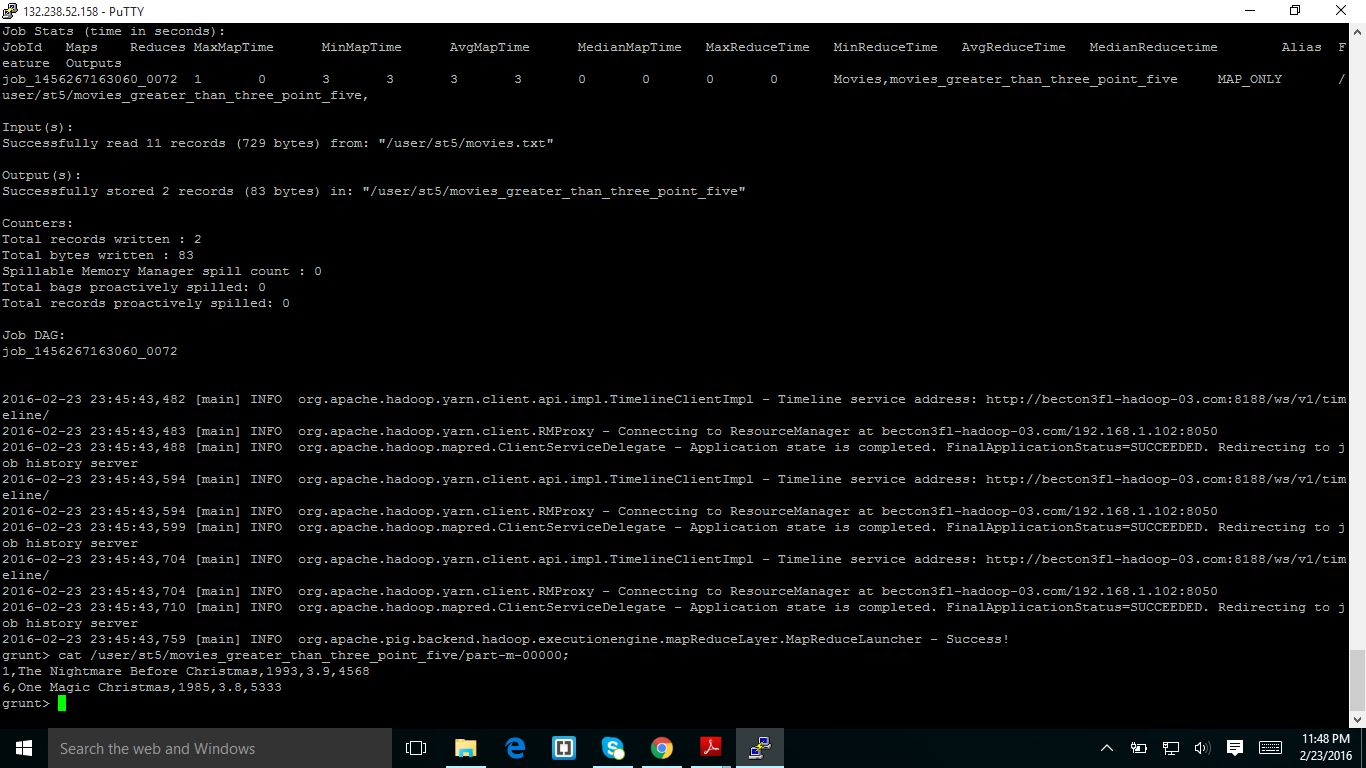
In the above screen shot I displayed the values of variable for each example using dump command.

**STEP7**:



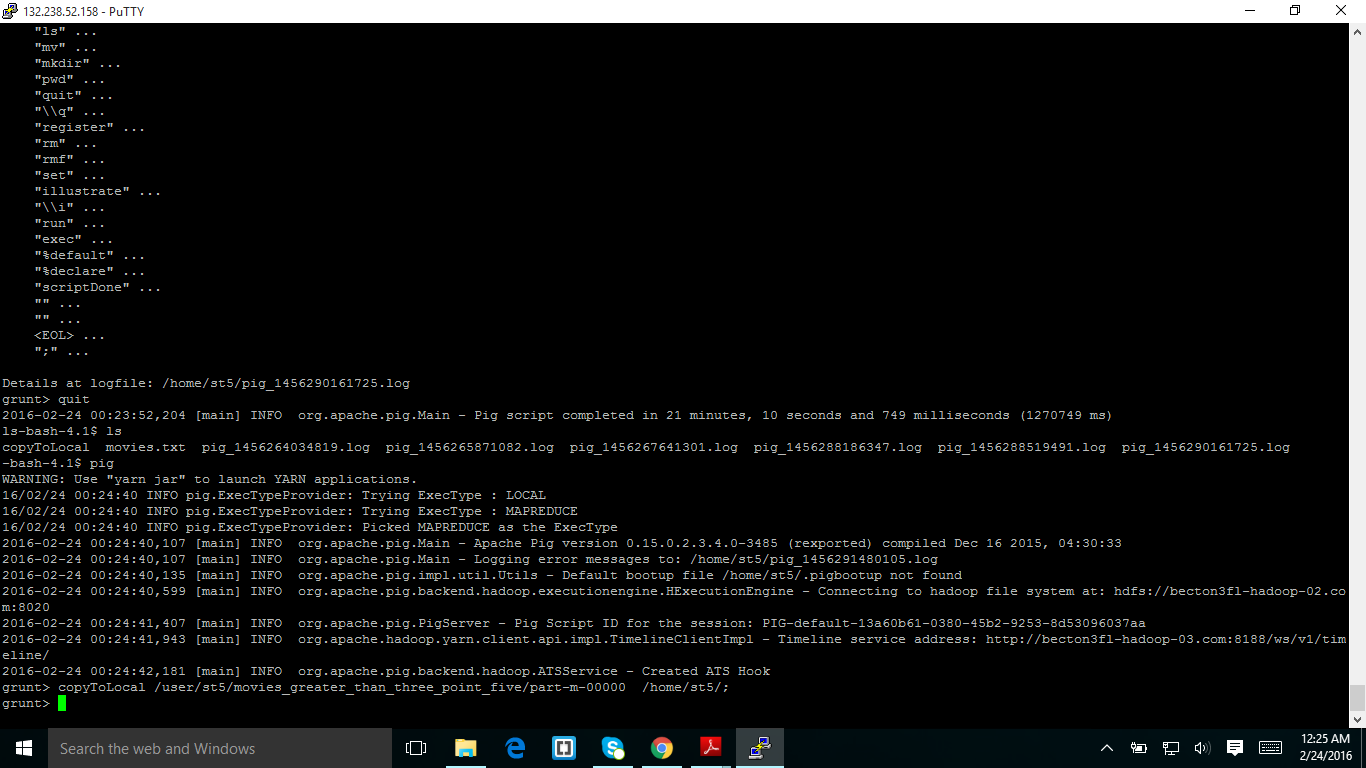
In the above screen shots I stored the variable movies\_greater\_than\_three\_point\_five in HDFS file system.and the above screen shot result save in the file name of part-m-00000 placed in the sub folder of movies\_greater\_than\_three\_point\_five.

**STEP8**:



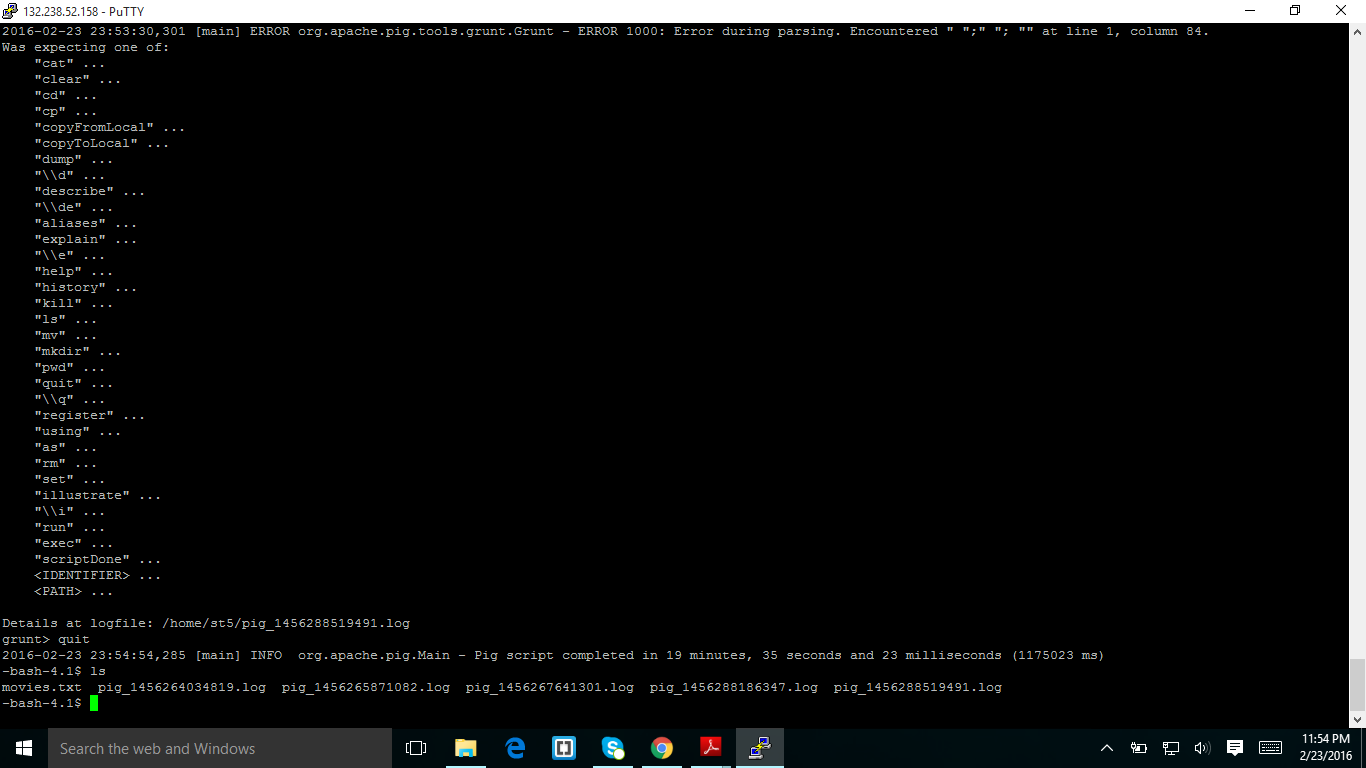
Using the above command I opened the processed file using cat command.

**STEP9**:



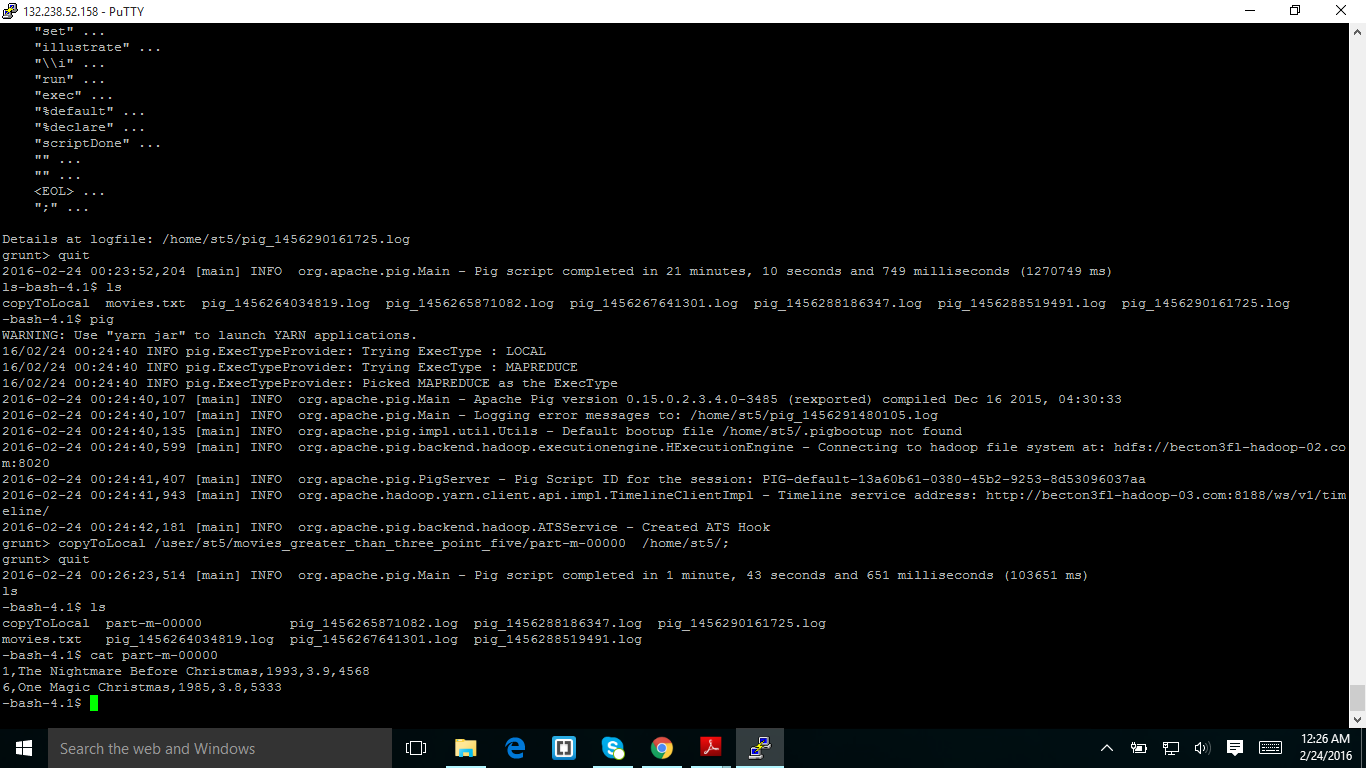
I copied the file from HDFS to local file system using the command copyto local.

**STEP10:**



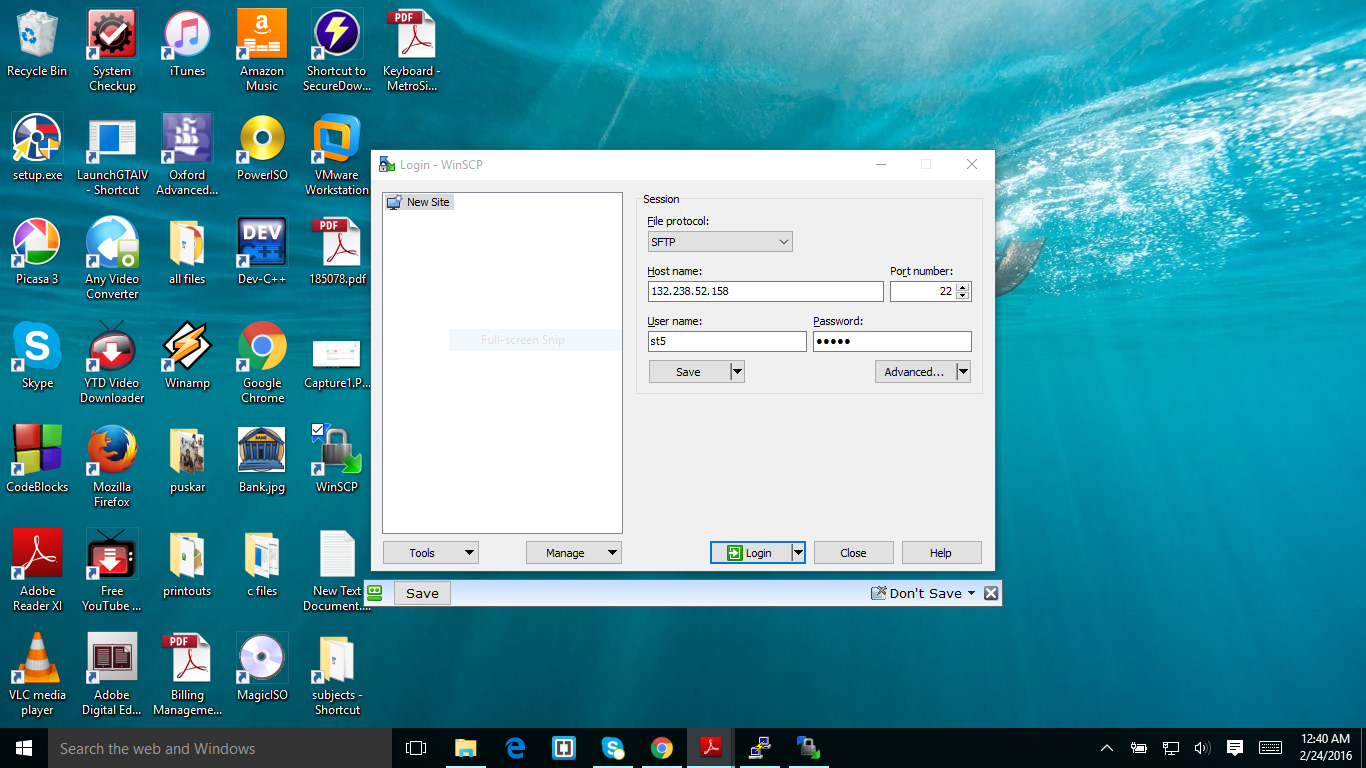
In above screen shot I quit from pig by using quit command and I displayed my files in my directory using ls command using linux machine.

**STEP11**:

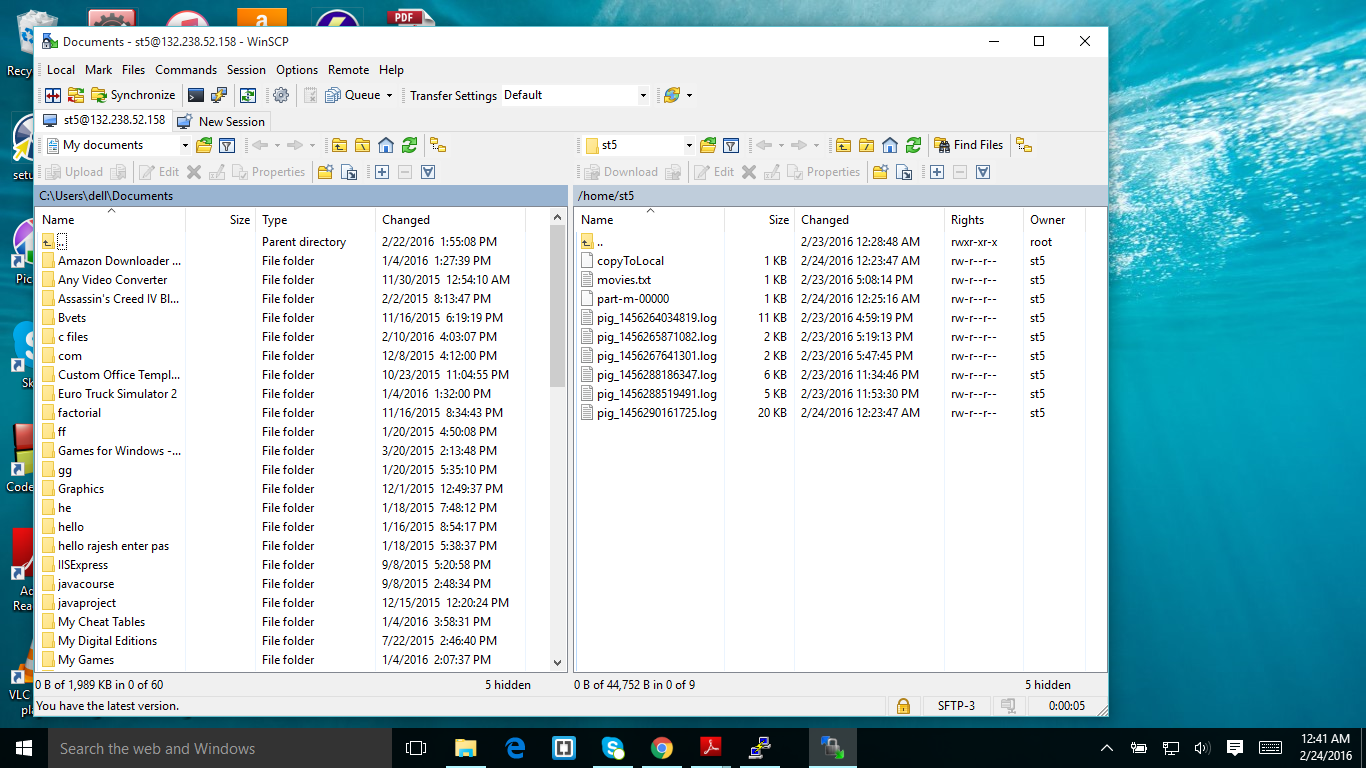


Using cat command I opened the file which was copied from hdfs to linux machine(part-m-00000) and it displayed the results.

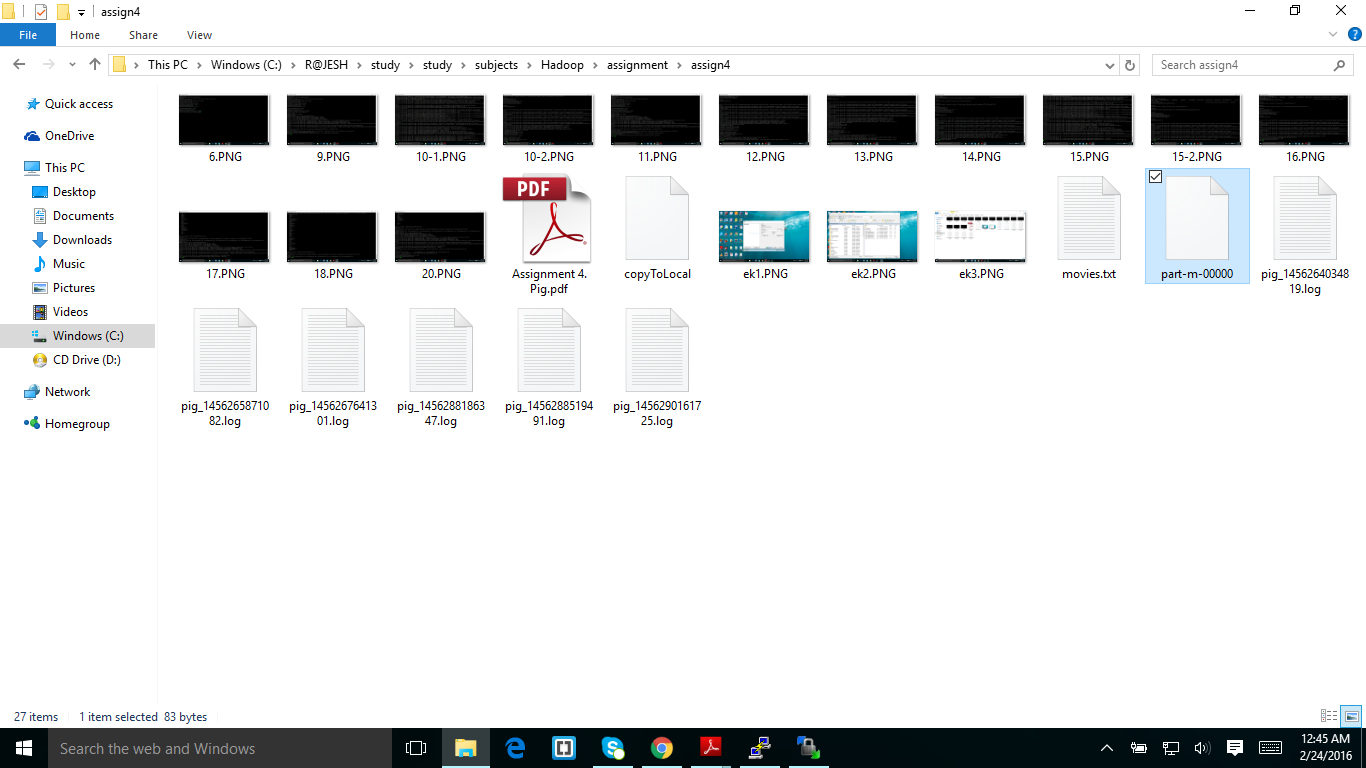
**EXTRA CREDIT ASSIGNMENT**:



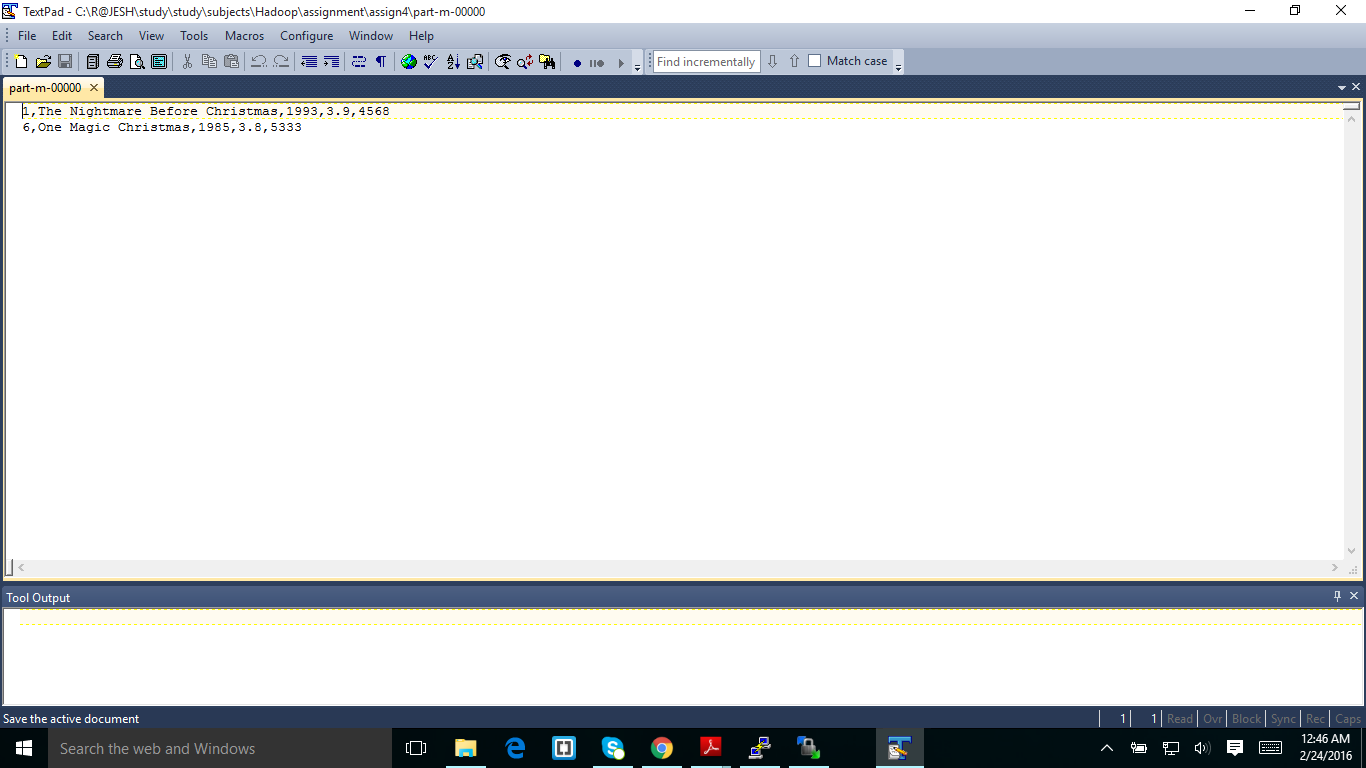
In above screen shot I was login to linux machine using winscp



The screen shot was taken after login and it displayed my folder in linux machine.



I click the download button of selecting all and I downloaded in to my folder.and I copied all log files also.



In the above screen shot I displayed the result of part-m-00000.