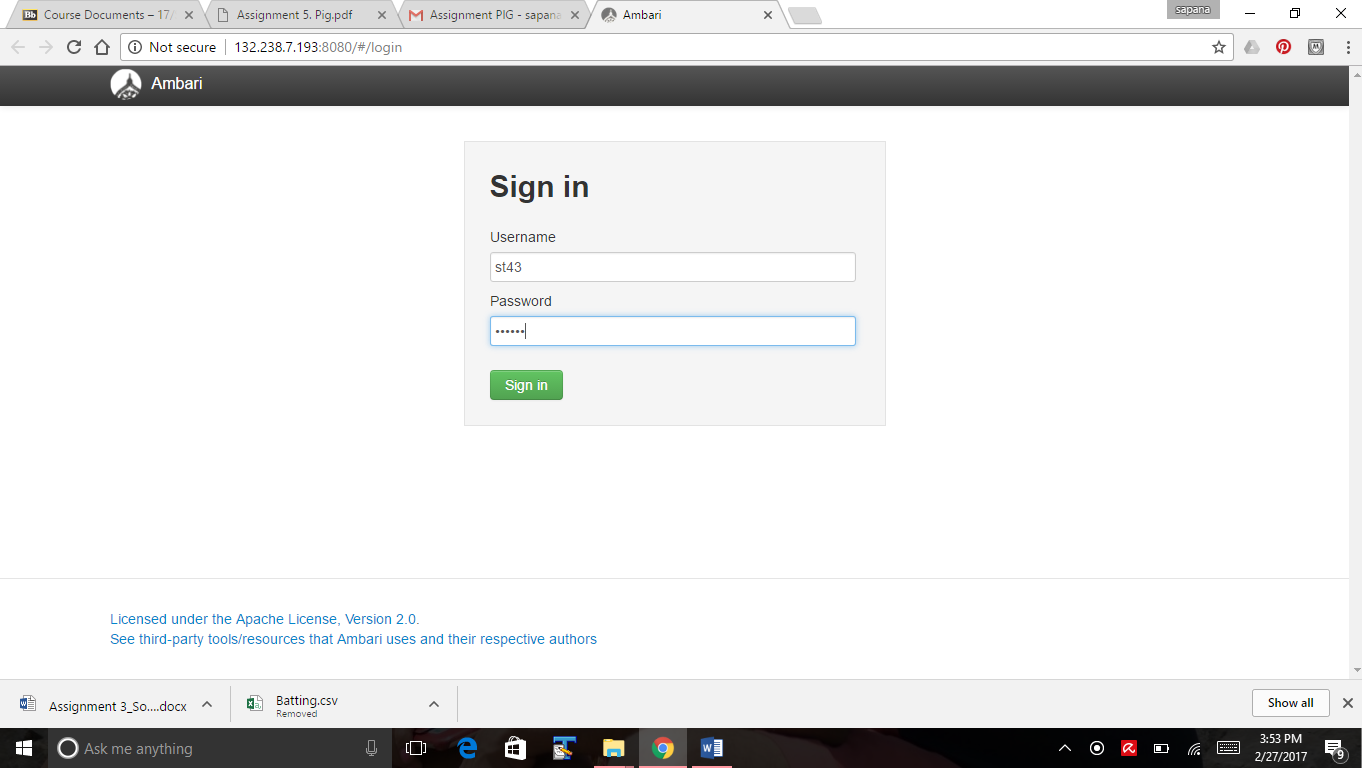
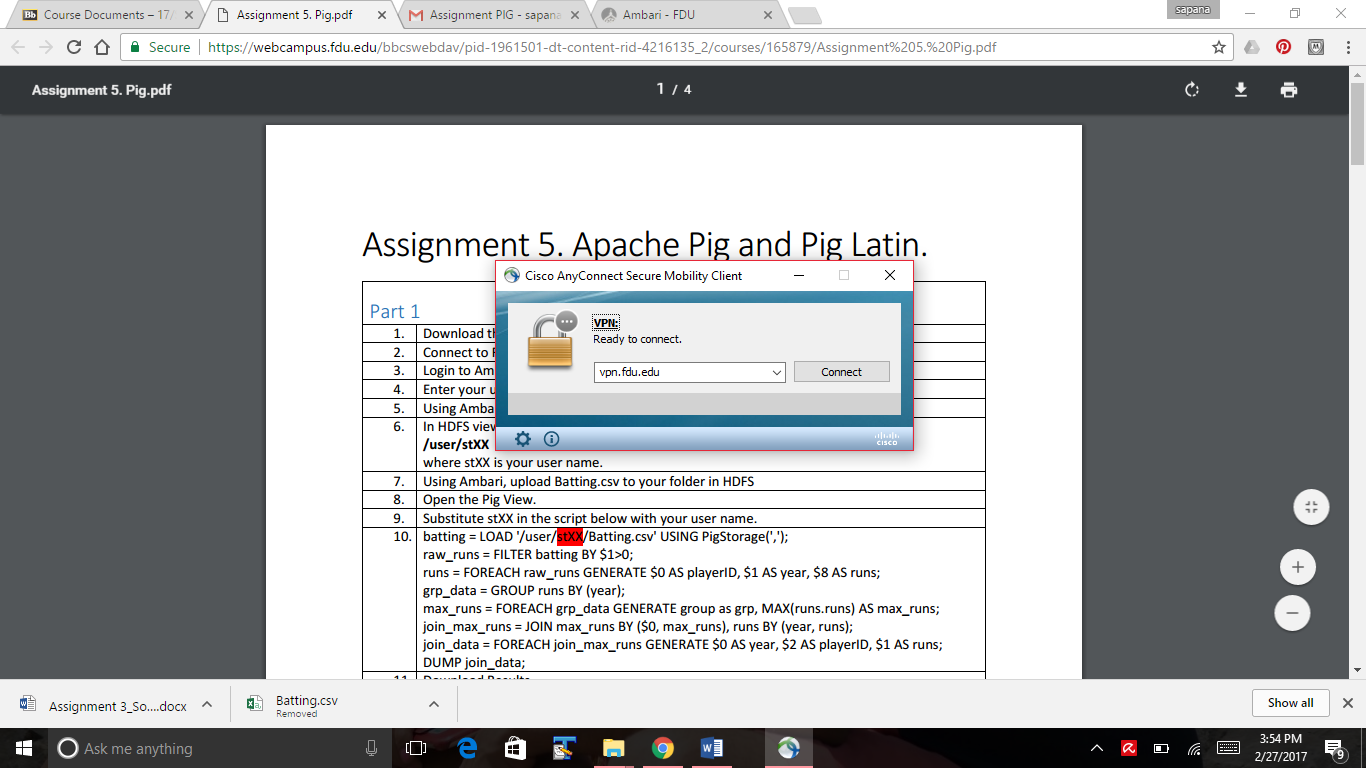
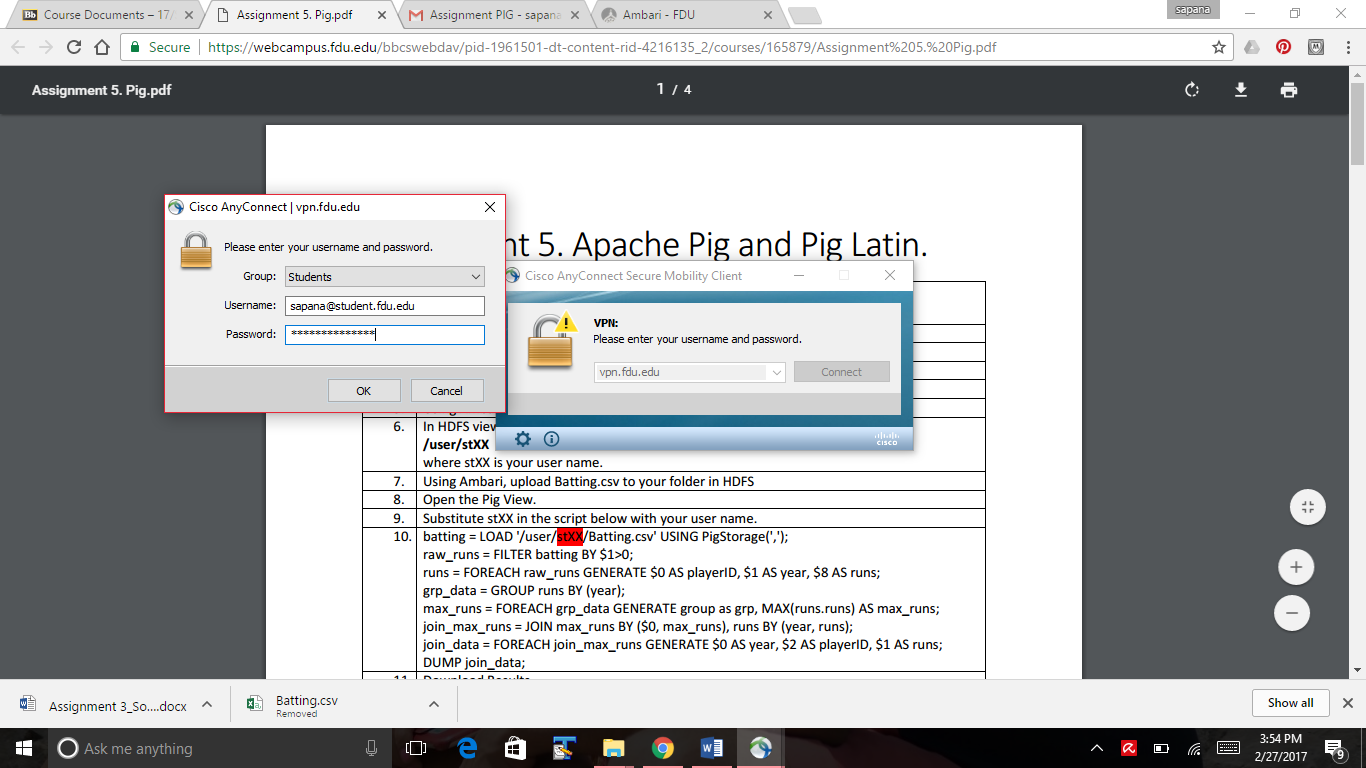
**­PART-1**

Sign in to ambari

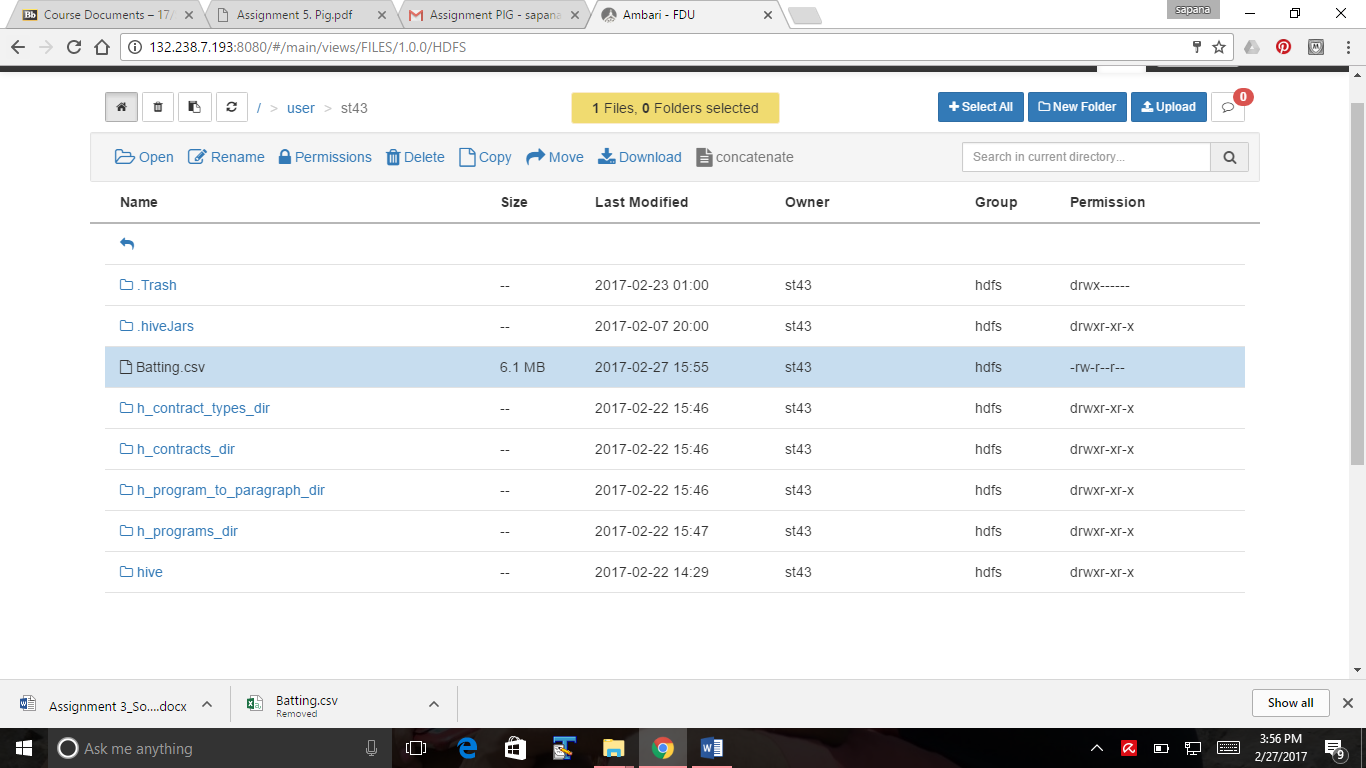


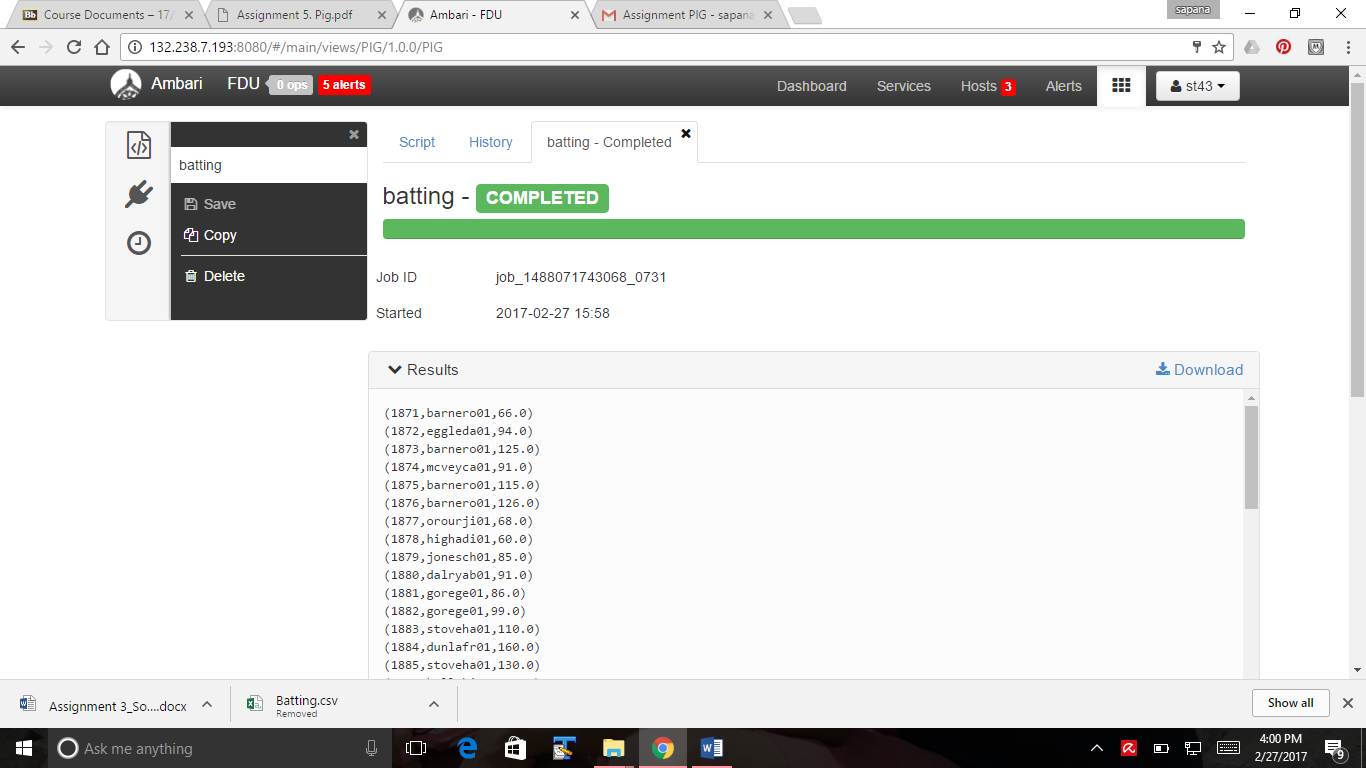
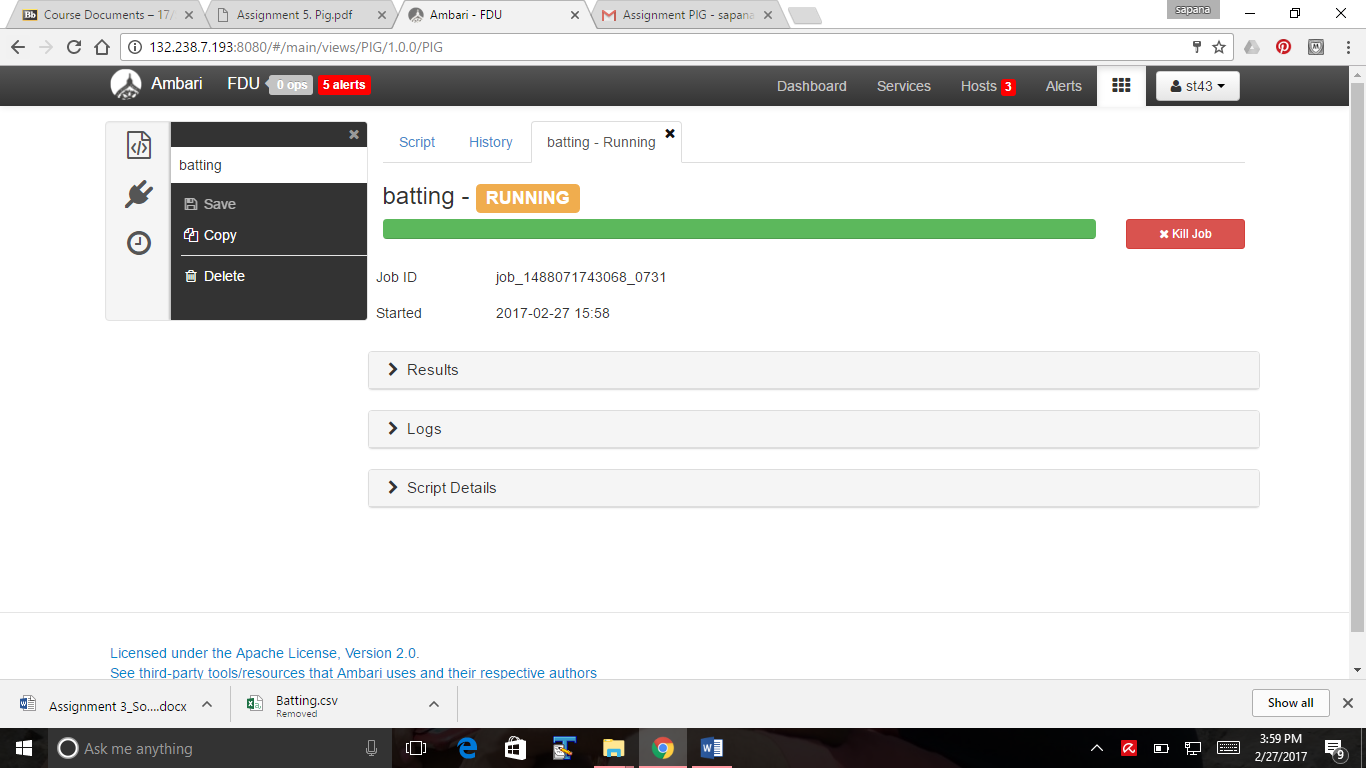
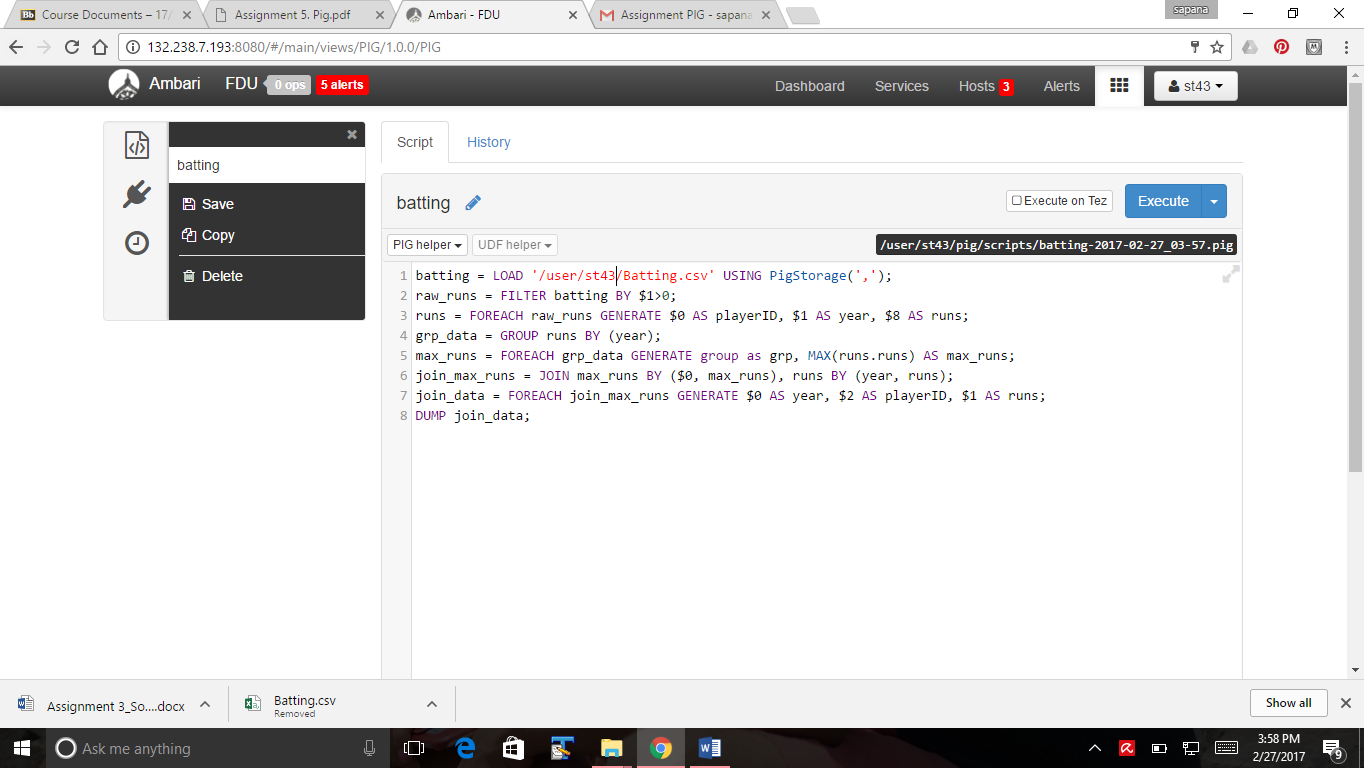
Start VPN

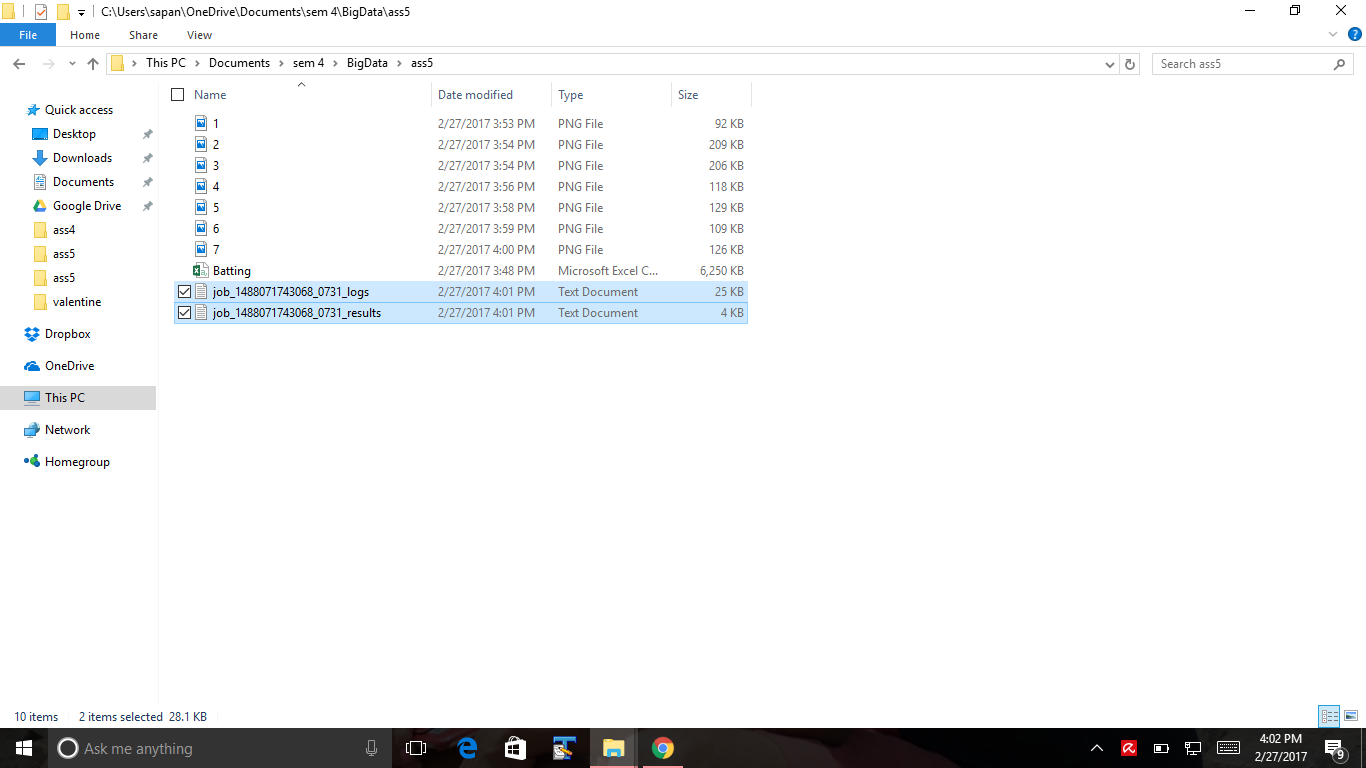




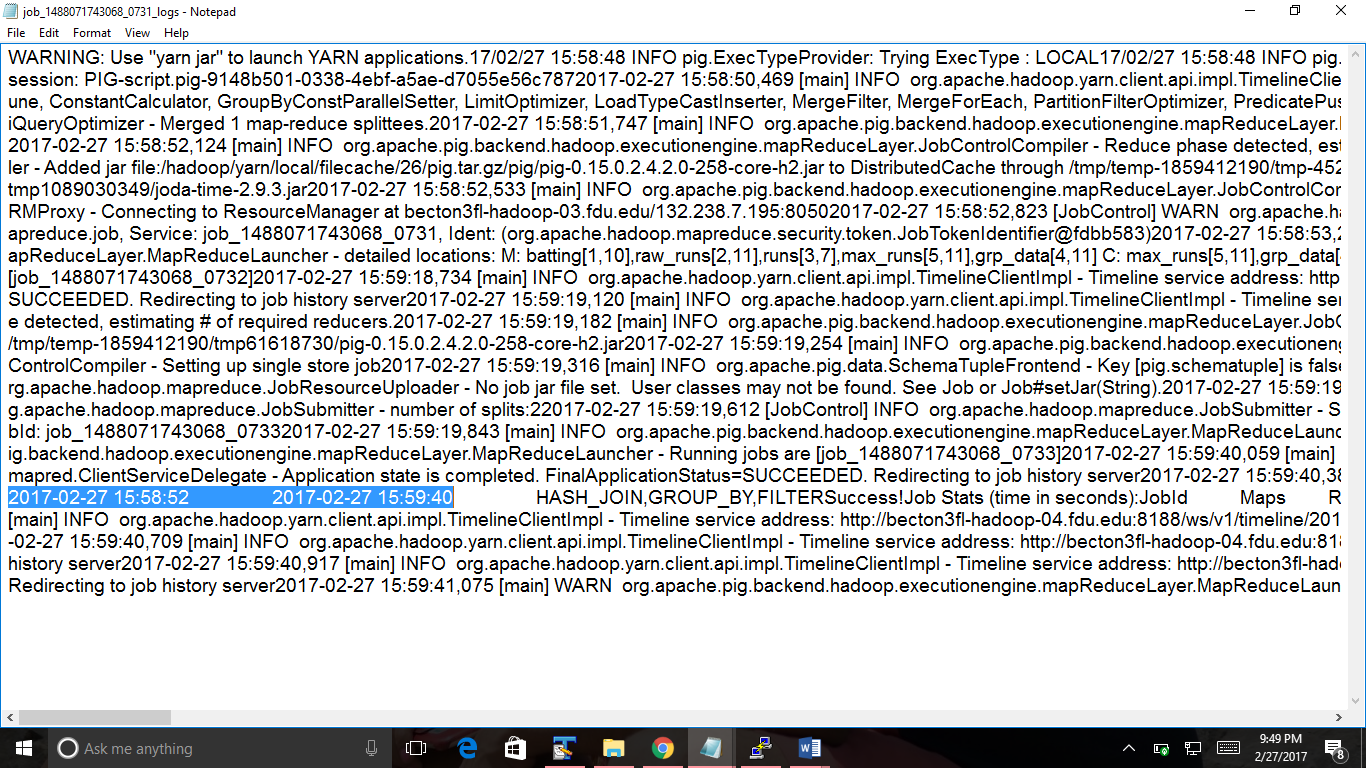
Upload Batting.csv file to my folder st43





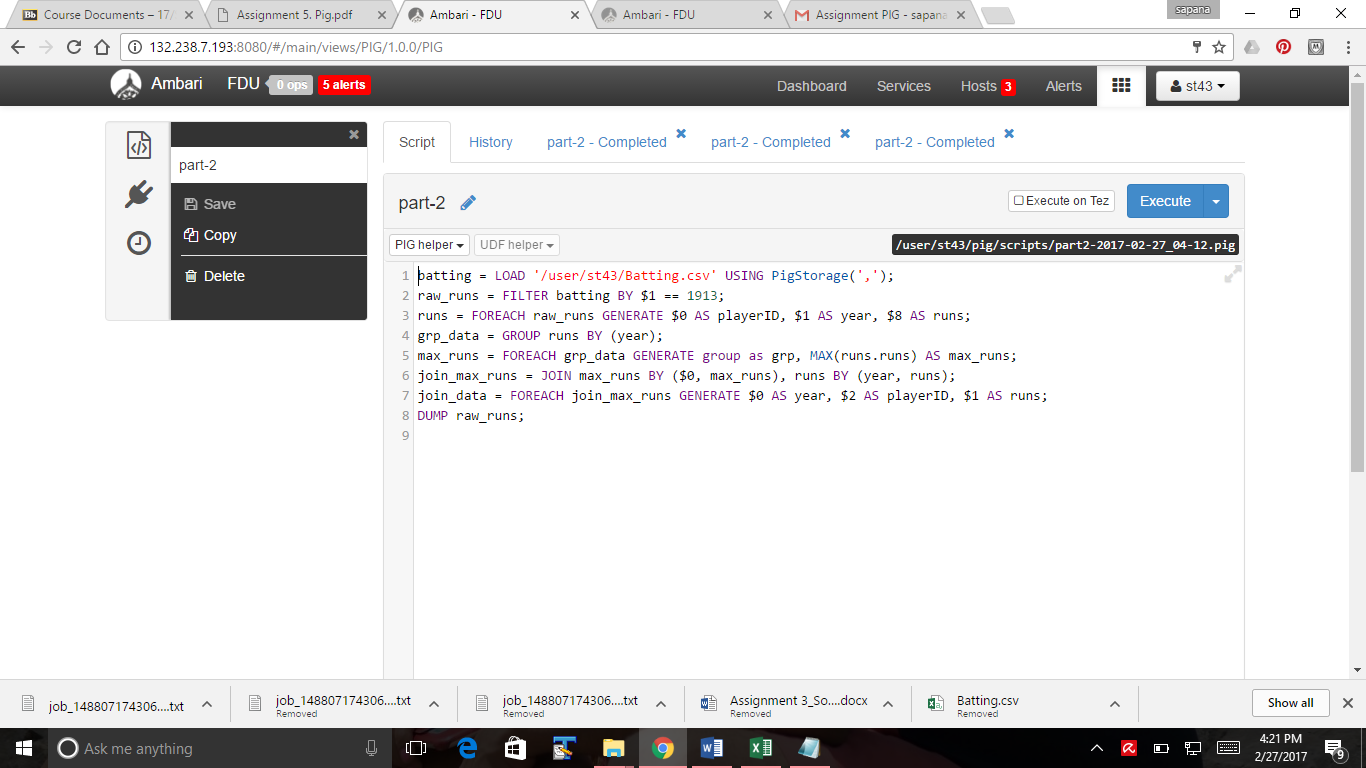


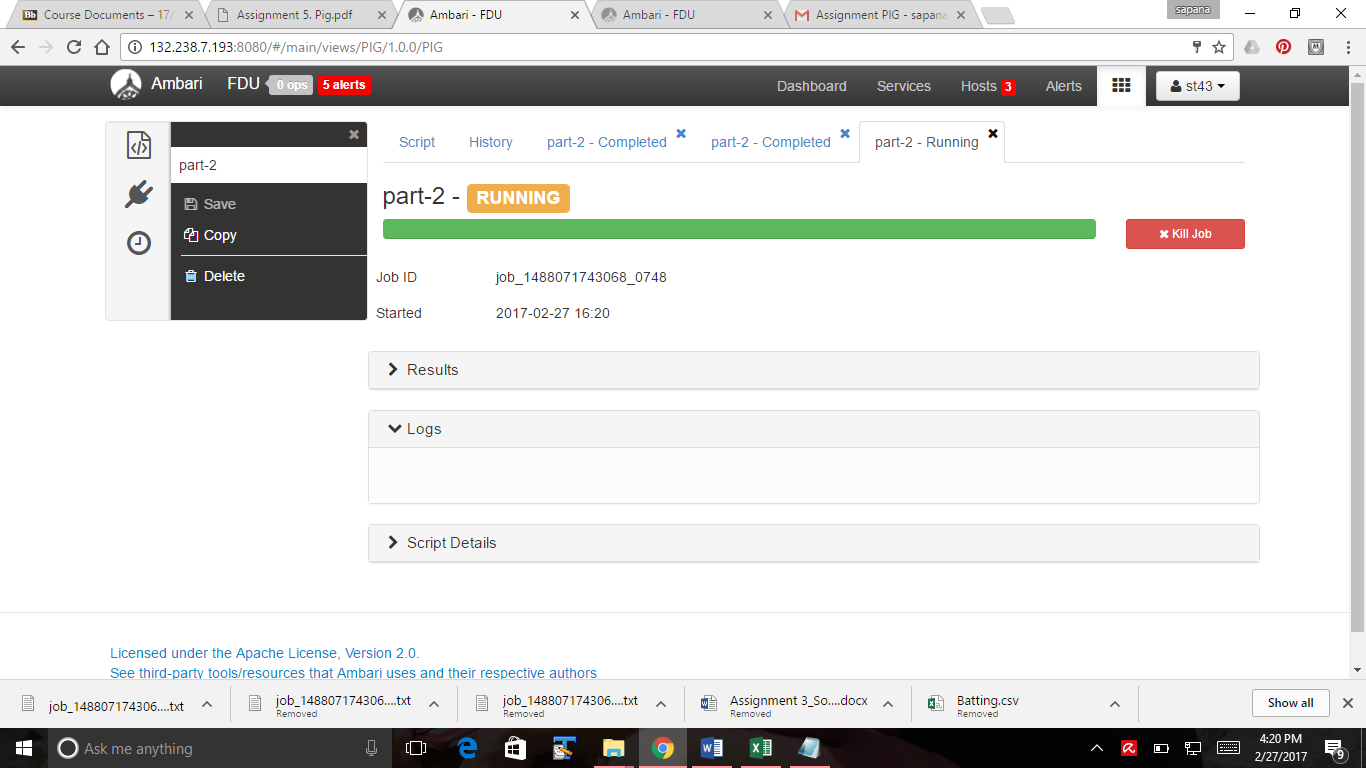
date and time when the script started and ended execution

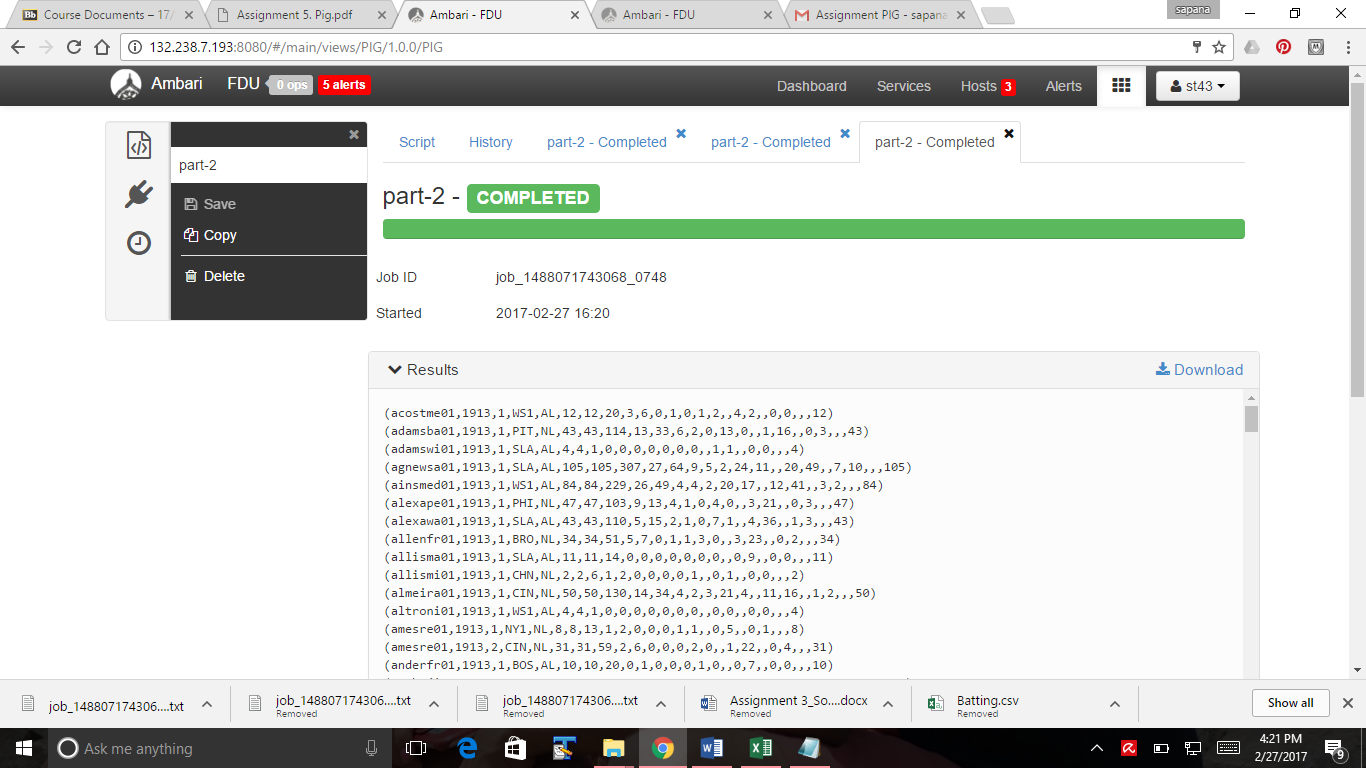


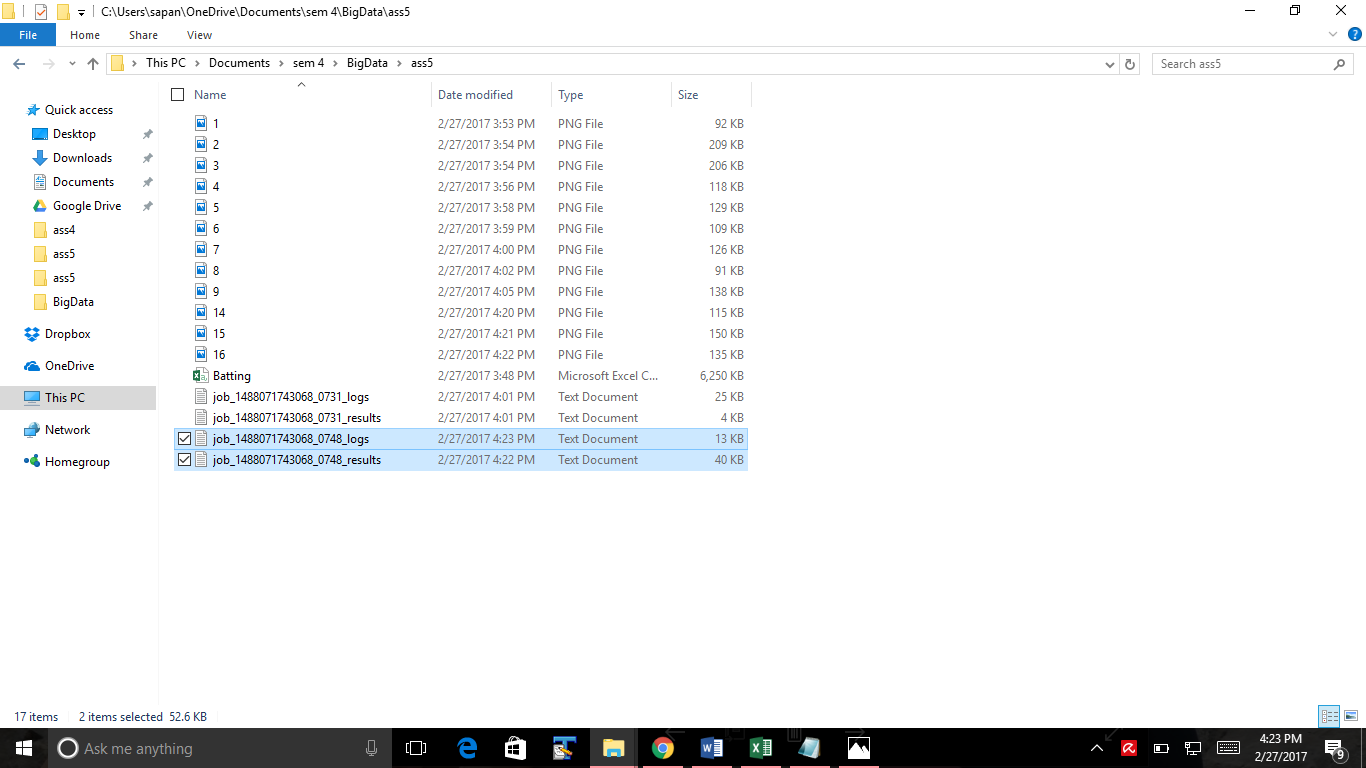
As per the Log file the script took 52 seconds

**PART-2**

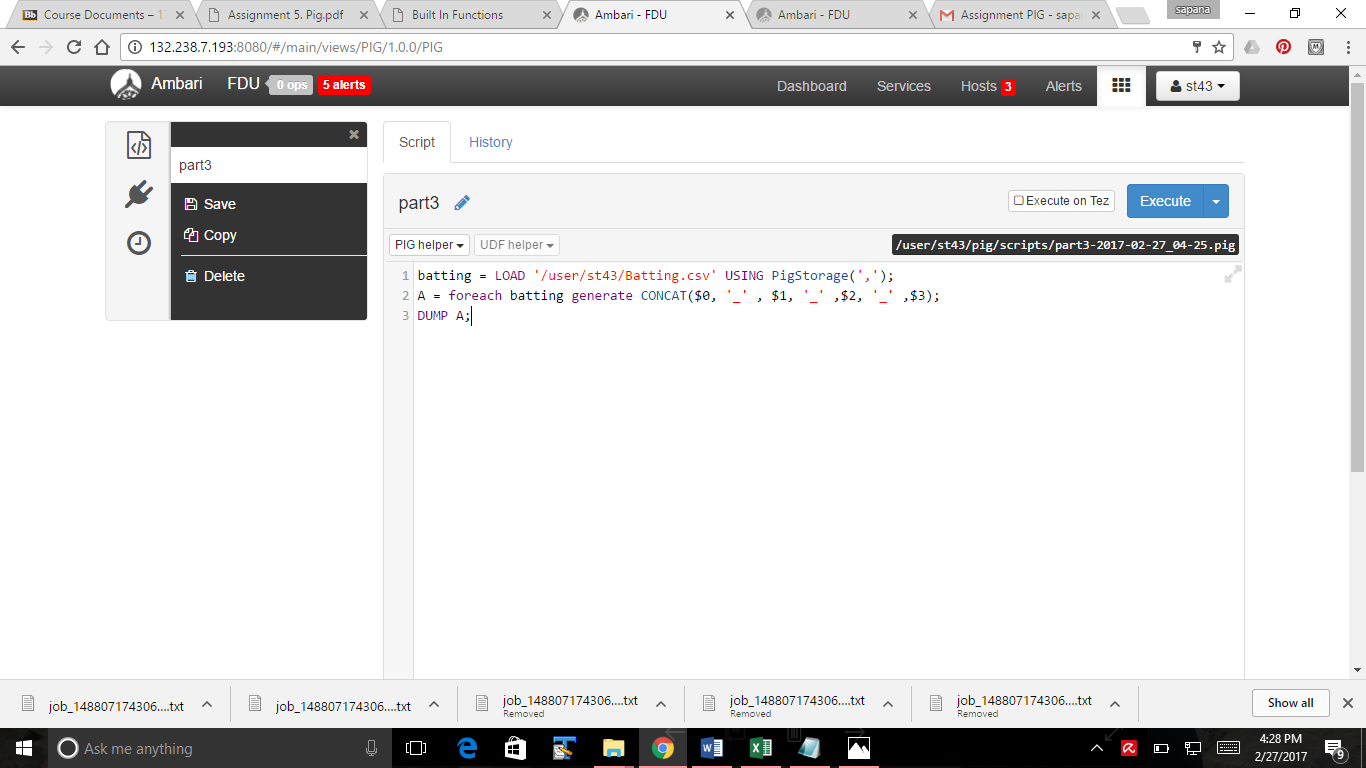


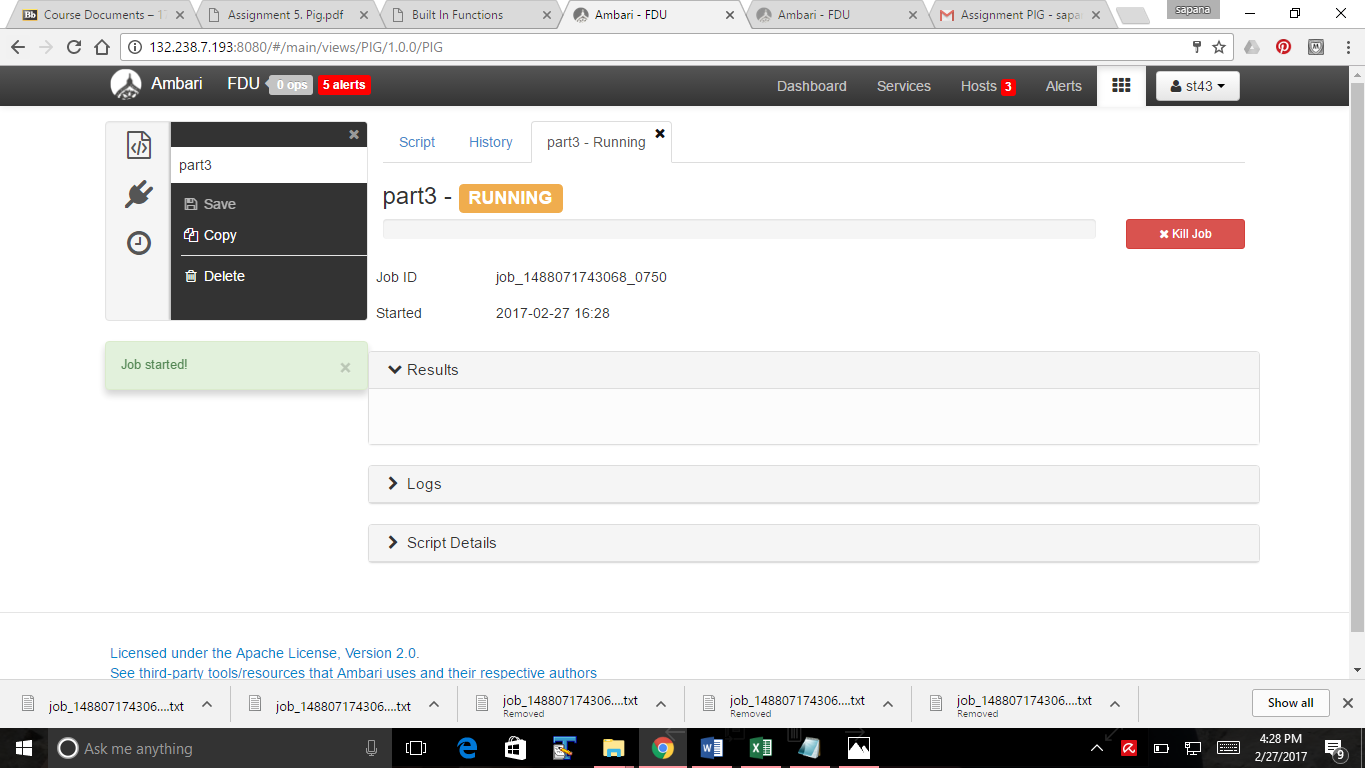


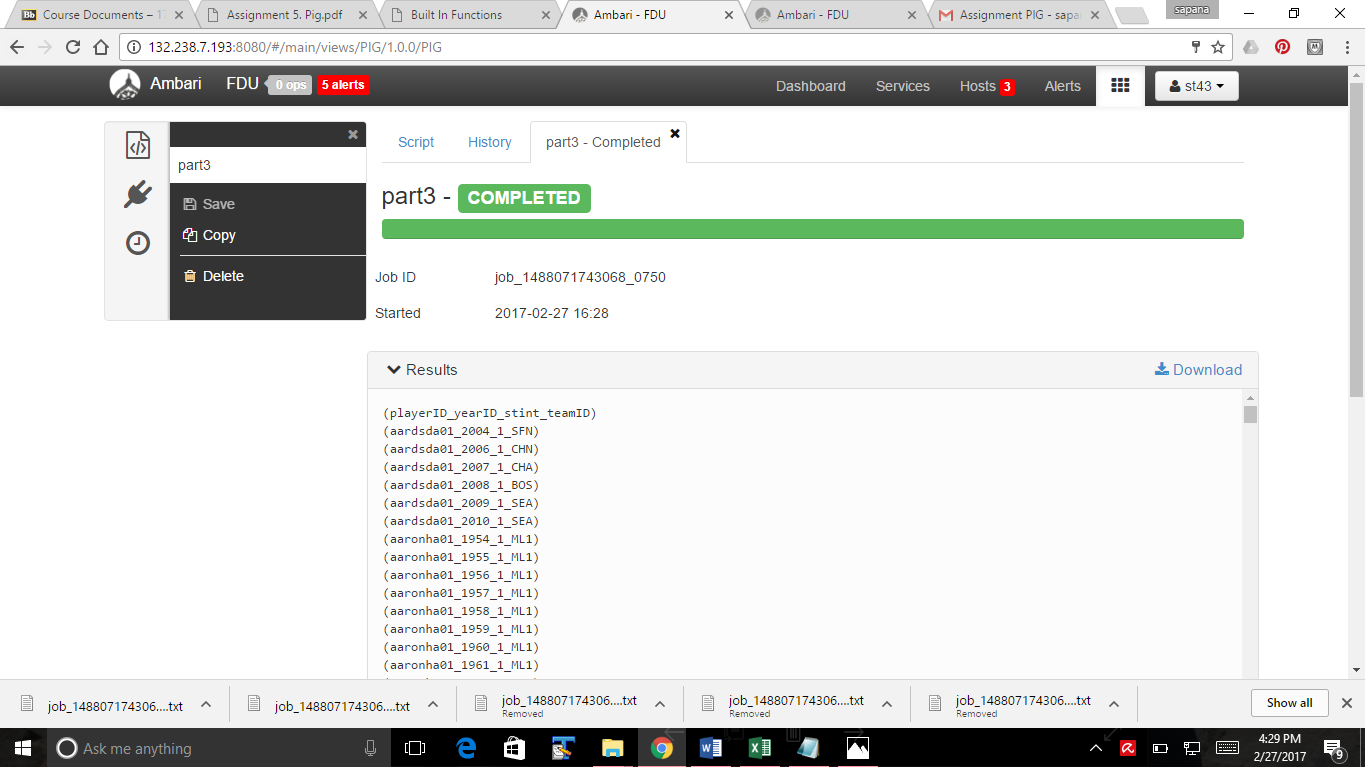




**PART-3**

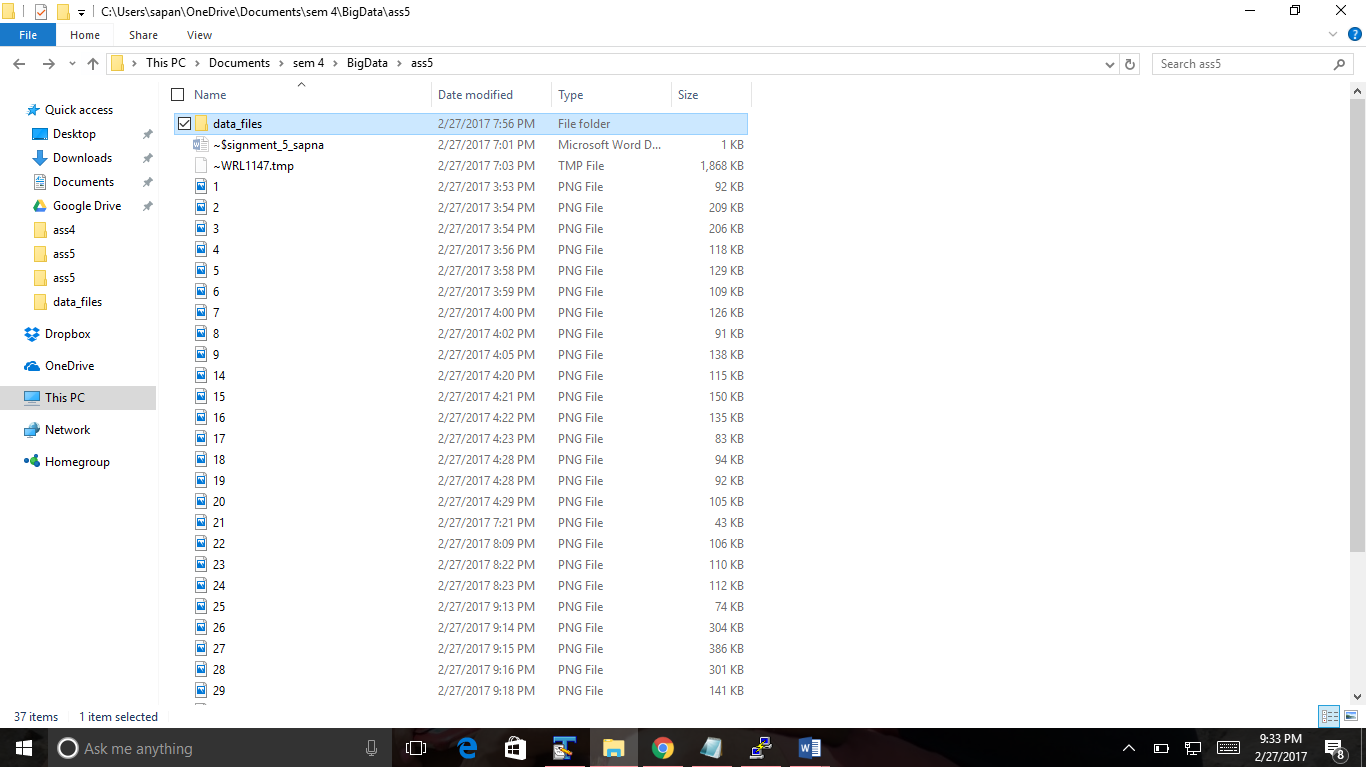




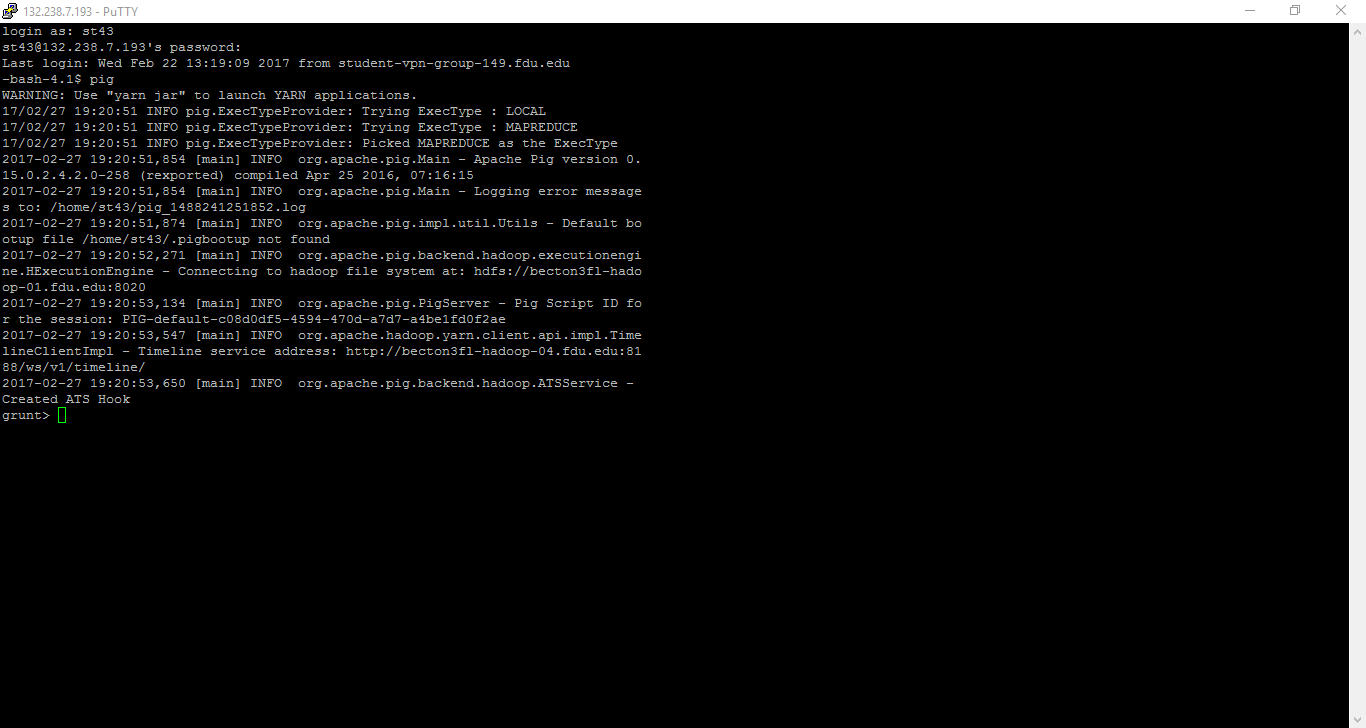


**PART-4**

(1) Download data\_files



(2)Start Grunt (Pig shell) by typing pig command



**(3) & (4)Load the datafiles by following command:**

1.categories = LOAD '/user/st43/project/Categories.txt' USING PigStorage(',') AS (cat\_Number:int, cat\_Name:chararray);

2.constants = LOAD '/user/st43/project/Constants.txt' USING PigStorage(',') AS (seqno:int, taxrate:float, latefee:float);

3.creditcards = LOAD '/user/st43/project/CreditCards.txt' USING PigStorage(',') AS (card\_Name:chararray);

4.customers = LOAD '/user/st43/project/Customers.txt' USING PigStorage(',') AS (c\_id:int,c\_lname:chararray, c\_fname:chararray, c\_phone:chararray, c\_street:chararray, c\_city:chararray,c\_state:chararray, c\_zip:chararray, c\_email:chararray, c\_sex:chararray, c\_DOB:chararray,c\_licence:chararray, c\_creditcard:chararray, c\_ccnumber:chararray, c\_ccexpiry:chararray,c\_memberdate:chararray);

5.language = LOAD '/user/st43/project/Language.txt' USING PigStorage(',') AS (l\_No:int,l\_Name:chararray);

6.ratings= LOAD '/user/st43/project/Ratings.txt' USING PigStorage(',') AS

(r\_No:int, r\_Rating:chararray);

7.formats = LOAD 'user/st43/project/Formats.txt' USING PigStorage(',') AS(f\_No:int, f\_Formats:chararray);

8.latecharges = LOAD 'user/st43/project/LateCharges.txt' USING PigStorage(',') AS(lc\_num:int, lc\_date:chararray, lc\_cust:int, lc\_movie:int, lc\_amount:float, lc\_paid:int);

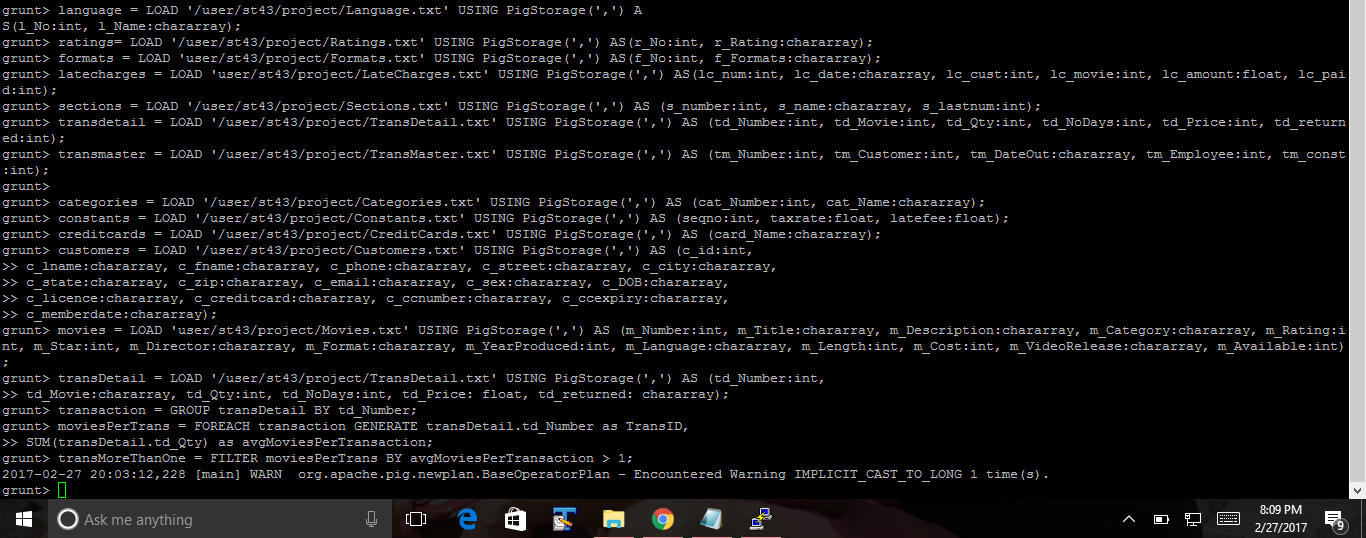
9.sections = LOAD '/user/st43/project/Sections.txt' USING PigStorage(',') AS (s\_number:int, s\_name:chararray, s\_lastnum:int);

10.transdetail = LOAD '/user/st43/project/TransDetail.txt' USING PigStorage(',') AS (td\_Number:int, td\_Movie:int, td\_Qty:int, td\_NoDays:int, td\_Price:int, td\_returned:int);

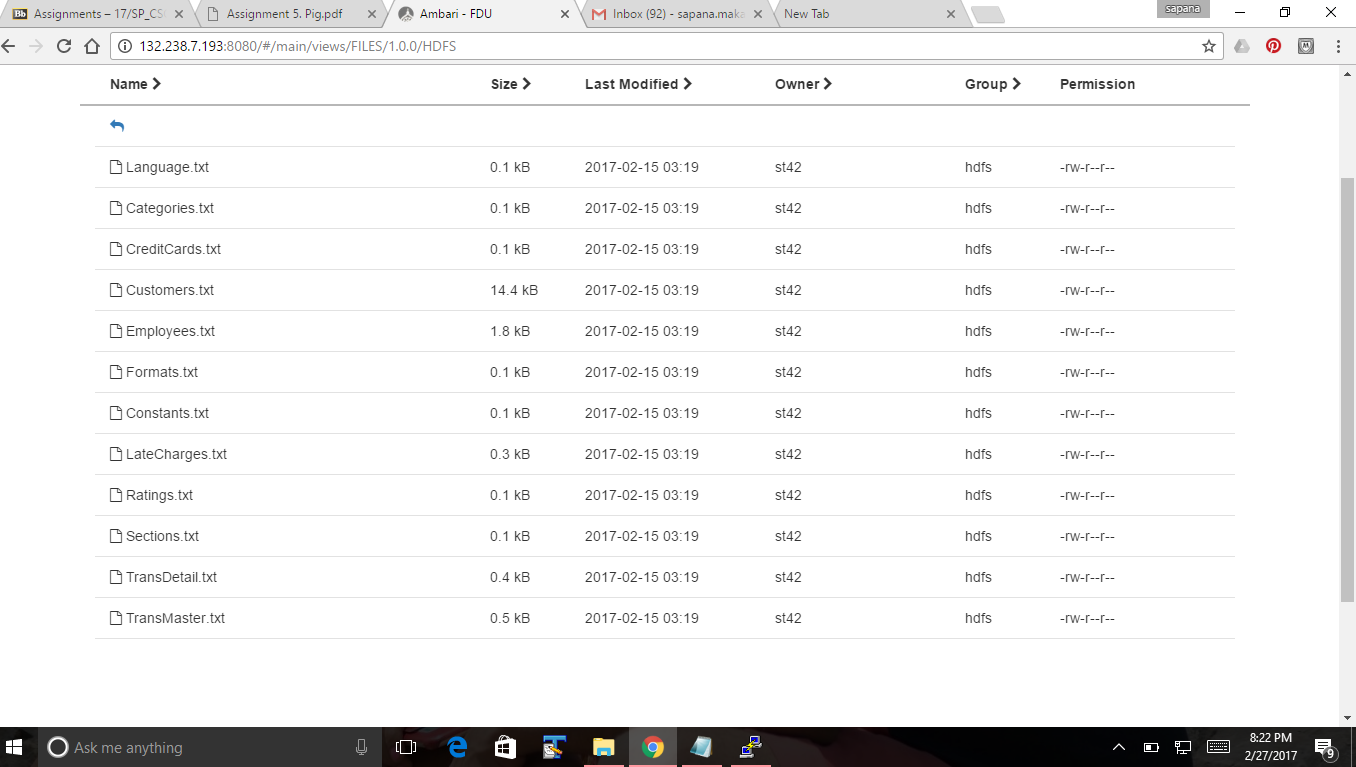
11.transmaster = LOAD '/user/st43/project/TransMaster.txt' USING PigStorage(',') AS (tm\_Number:int, tm\_Customer:int, tm\_DateOut:chararray, tm\_Employee:int, tm\_const:int);

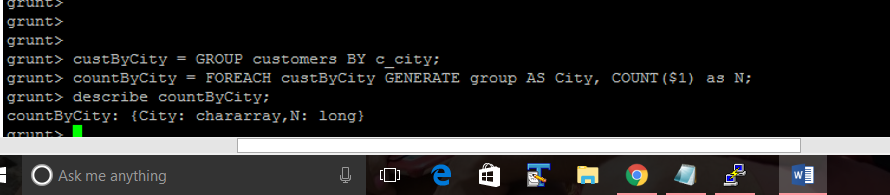
12.movies = LOAD 'user/st43/project/Movies.txt' USING PigStorage(',') AS (m\_Number:int, m\_Title:chararray, m\_Description:chararray, m\_Category:chararray, m\_Rating:int, m\_Star:int, m\_Director:chararray, m\_Format:chararray, m\_YearProduced:int, m\_Language:chararray, m\_Length:int, m\_Cost:float, m\_VideoRelease:chararray, m\_Available:int);

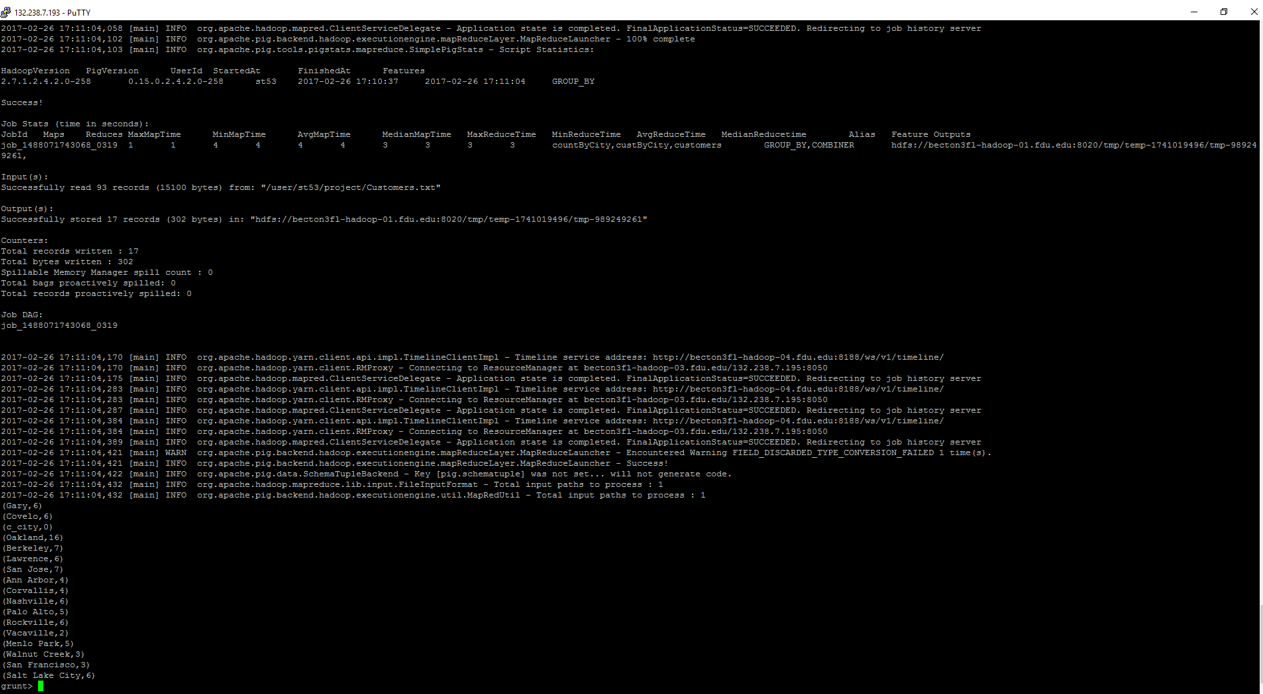
13.Employees= LOAD '/user/st43/project/ Employees.txt' USING PigStorage(',') AS (emp\_fname:chararray,emp\_id:int,emp\_minit:chararray,emp\_lname:chararray,emp\_job\_id:int,emp\_hire\_date:chararray);



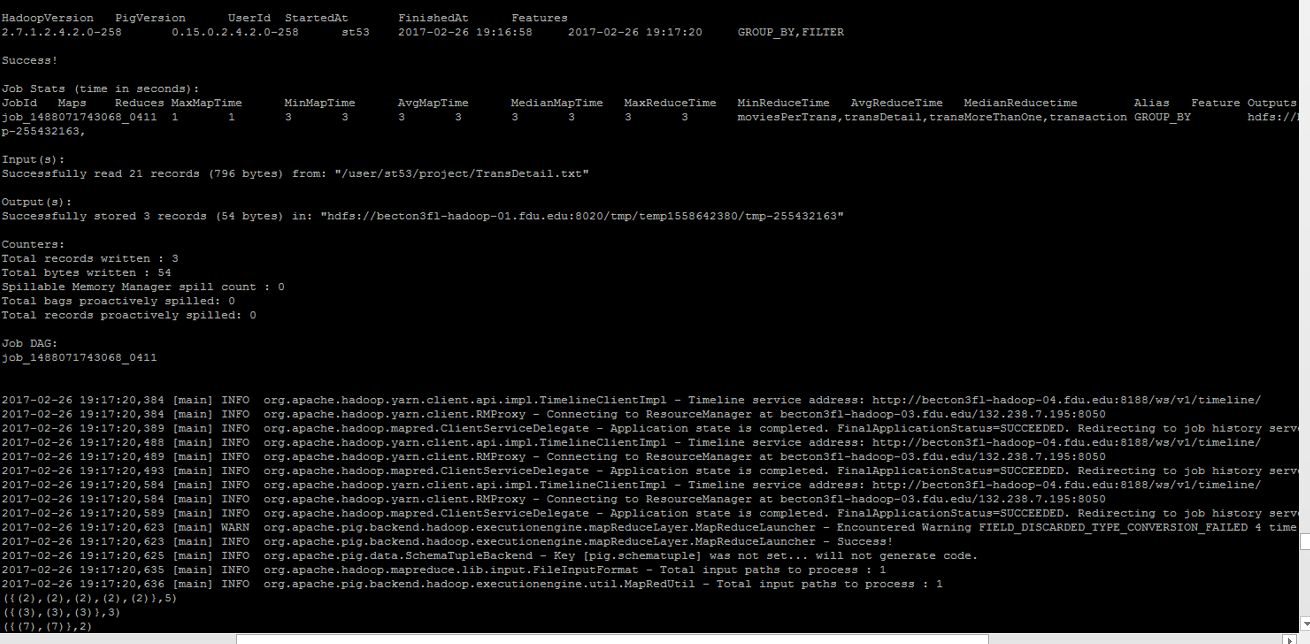




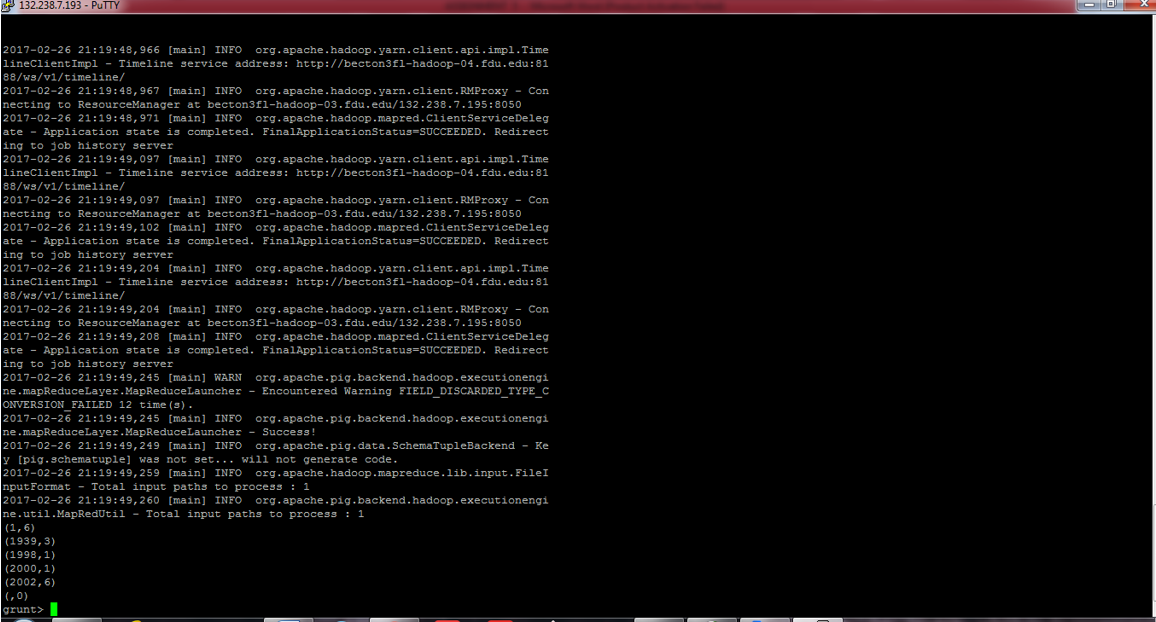
(5)



Display number of movies rented for each transaction



display transaction containing more than one movie



Executing the combined query on Apache HUE

