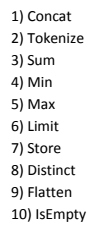
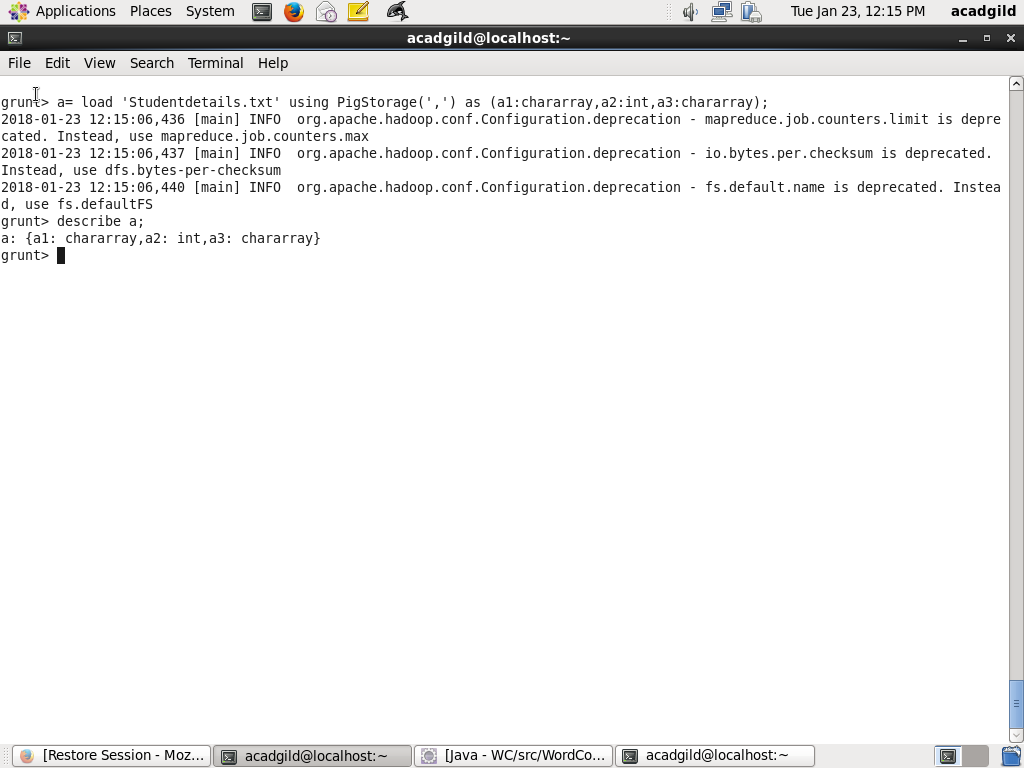
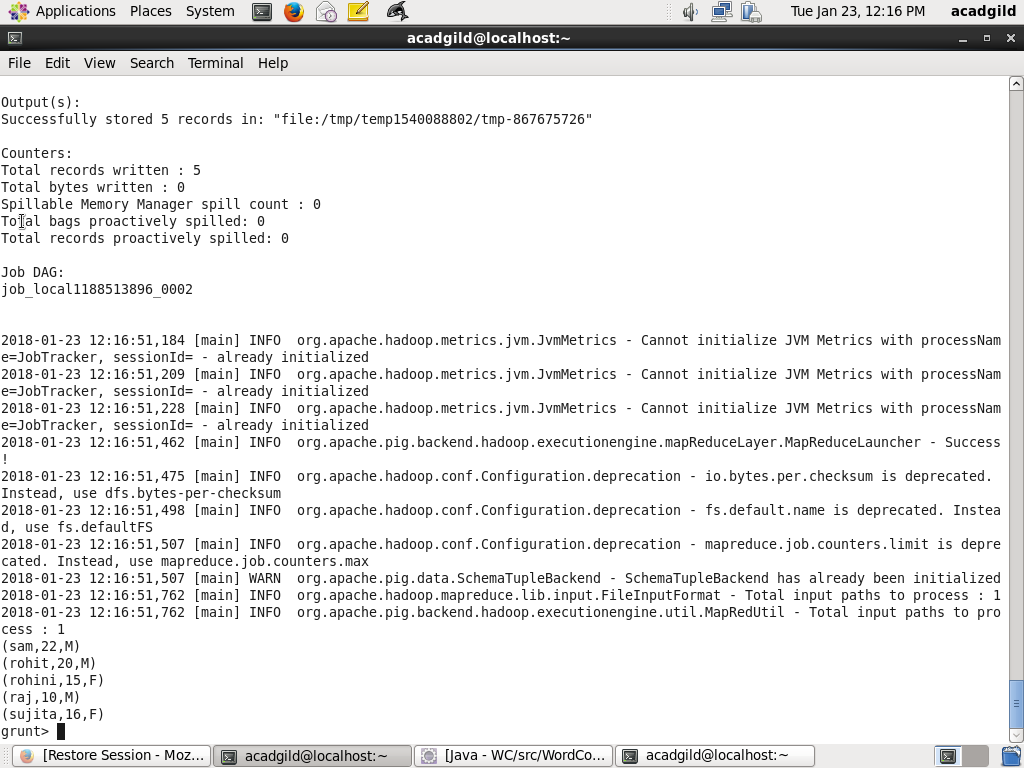
**Assignment 4.2**

Applying following Pig commands on the Student dataset as show in snapshots

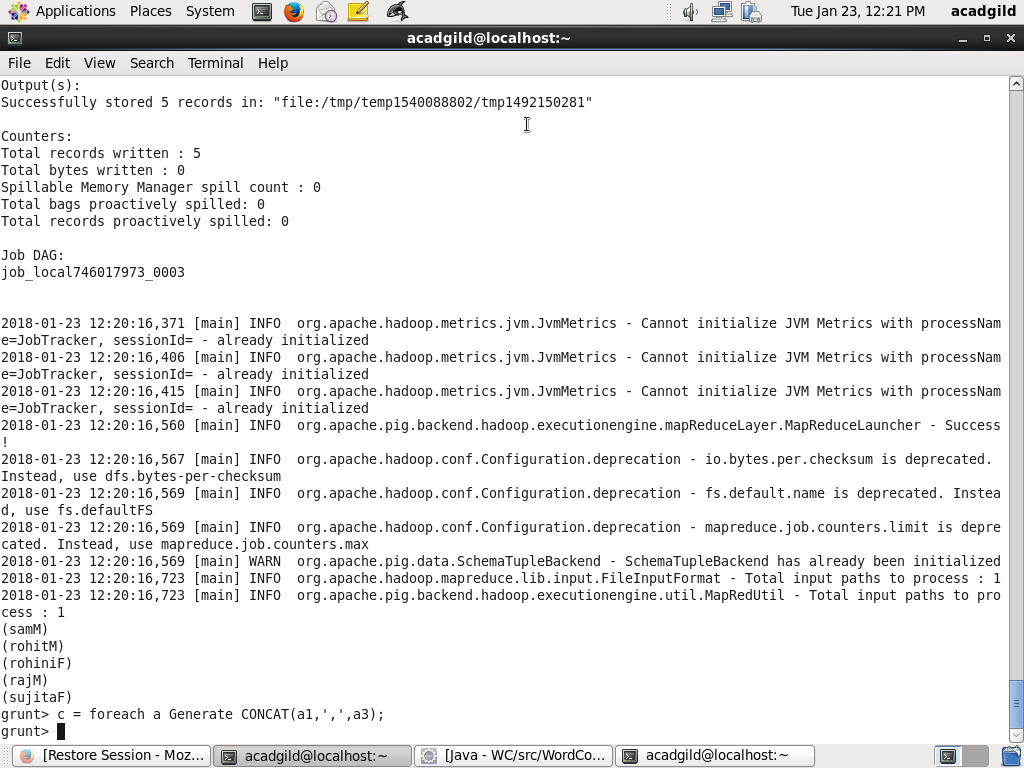


**1.CONCAT:**  For this I have created a table name Studentdetails.txt which contains studentname, age and gender . I have loaded this table into a as shown in grunt shell.

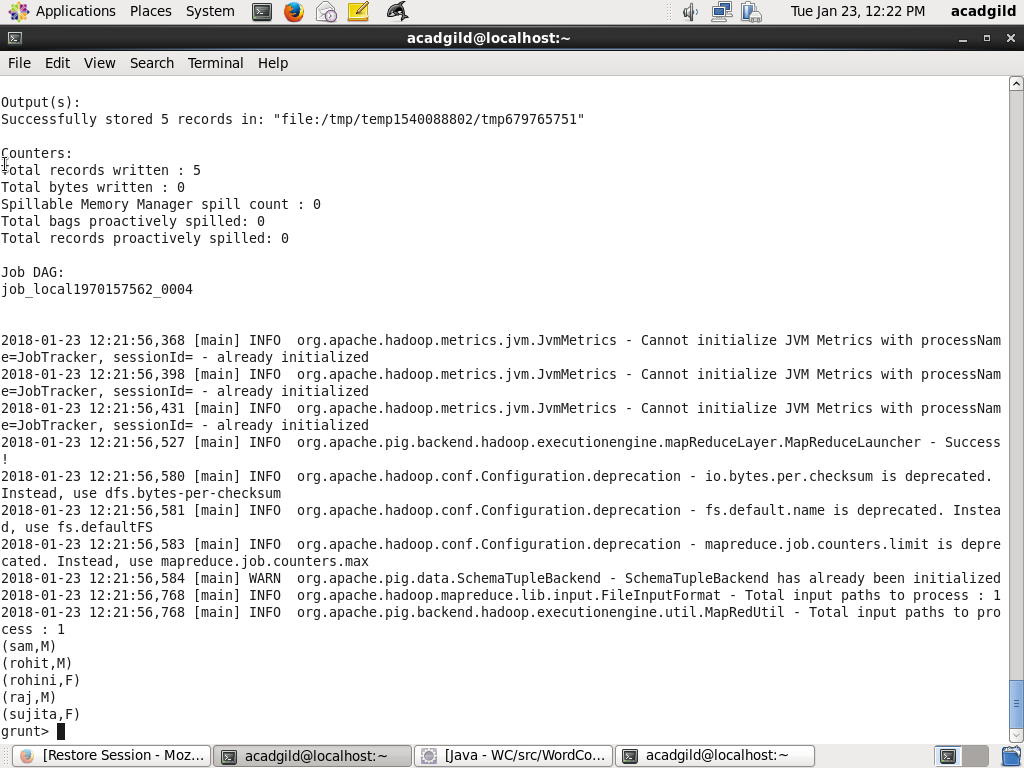


These are the table entries as shown below. 

Now im applying CONCAT and merging the name and gender of students as shown below

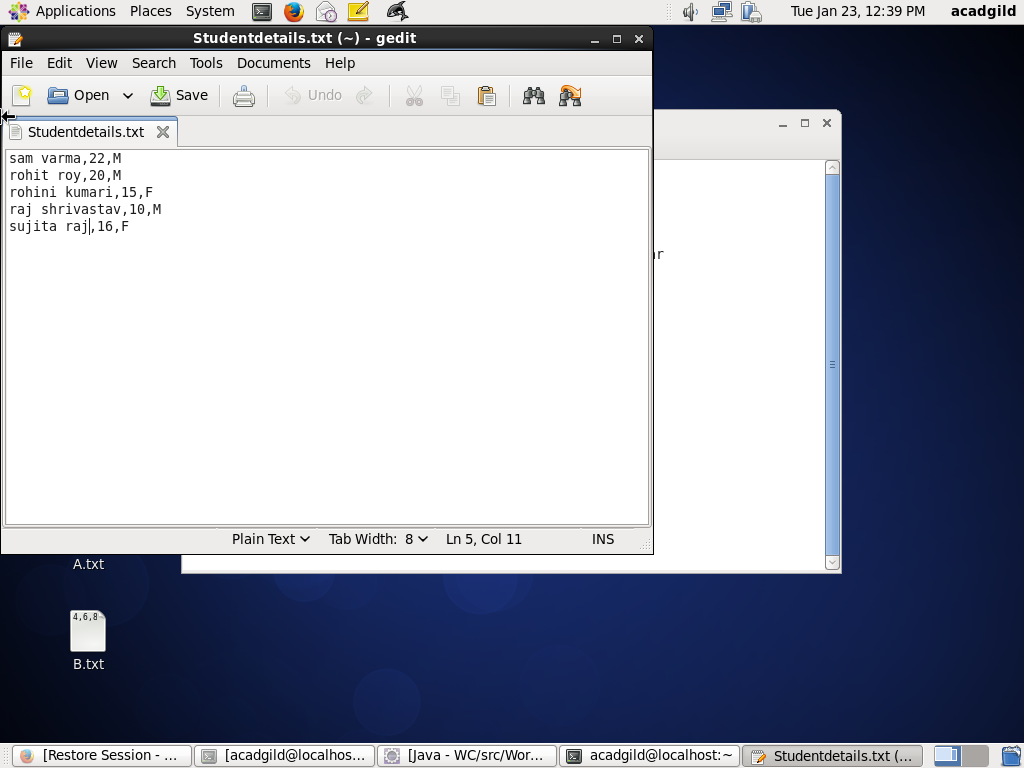


Here is the result



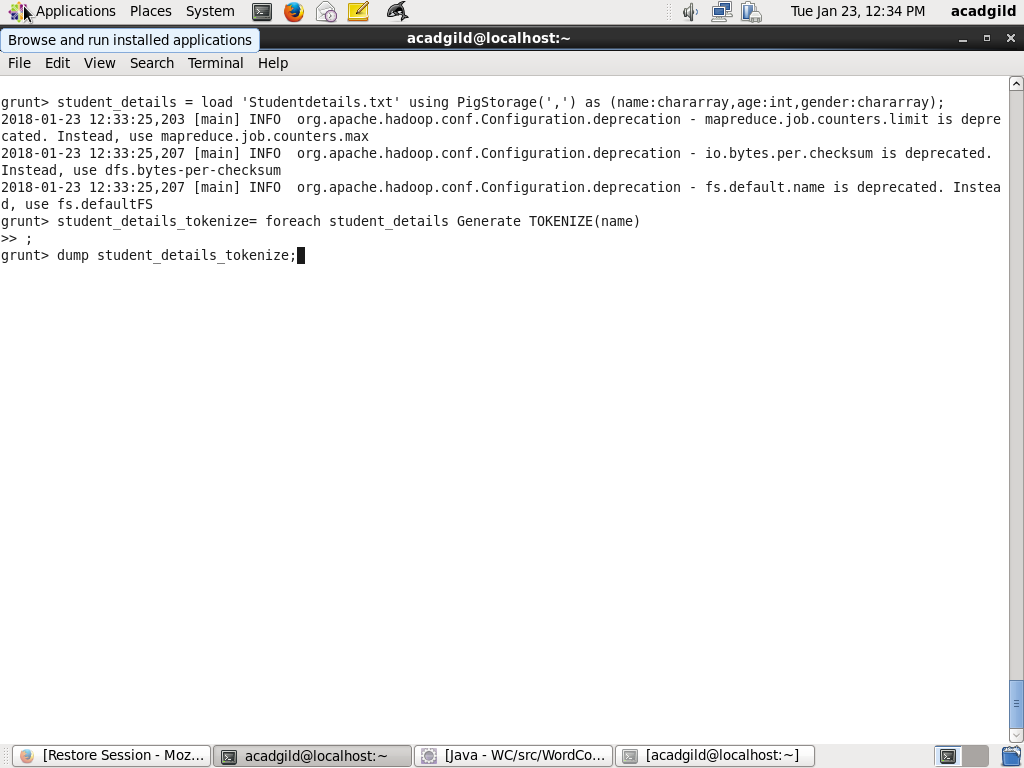
**2.TOKENIZE**

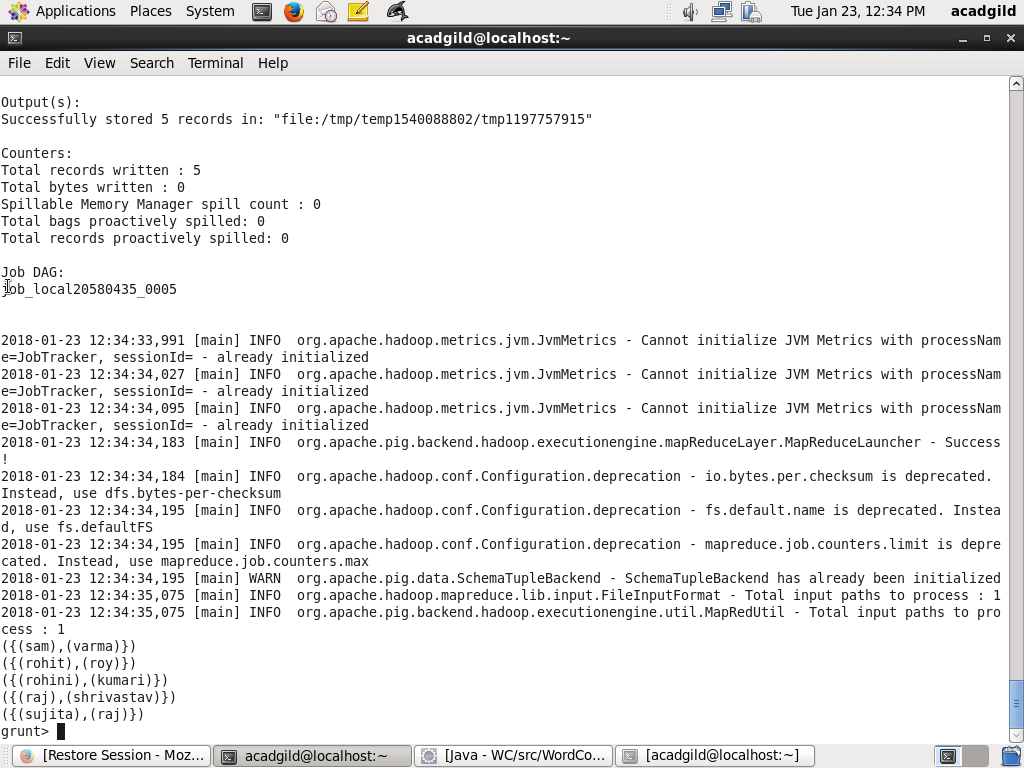
Here is the Studentdetails.txt table with

.

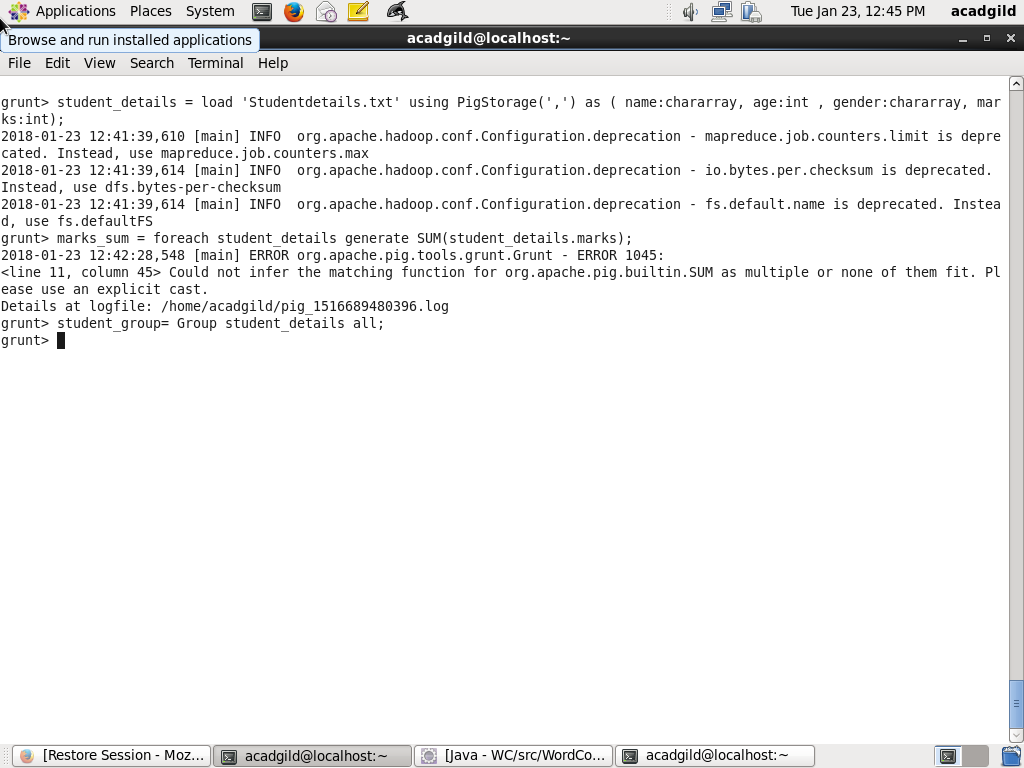
Here im loading Studentdetails.txt into student\_details inside the grunt shell.

Then im applying TOKENIZE to the names

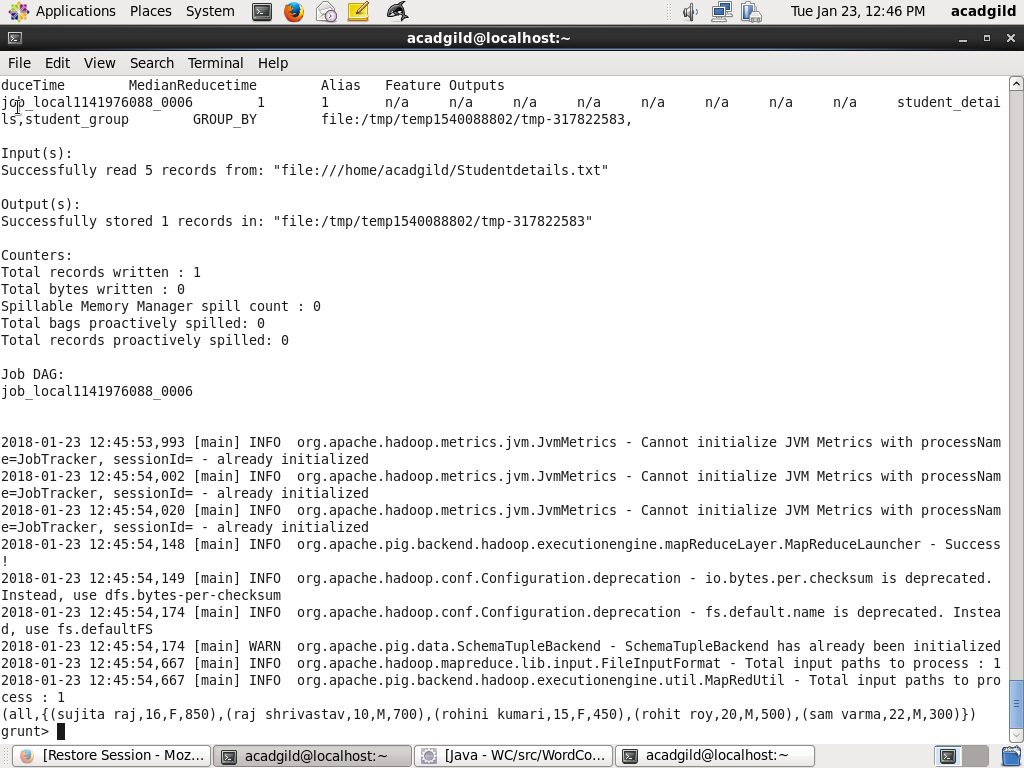




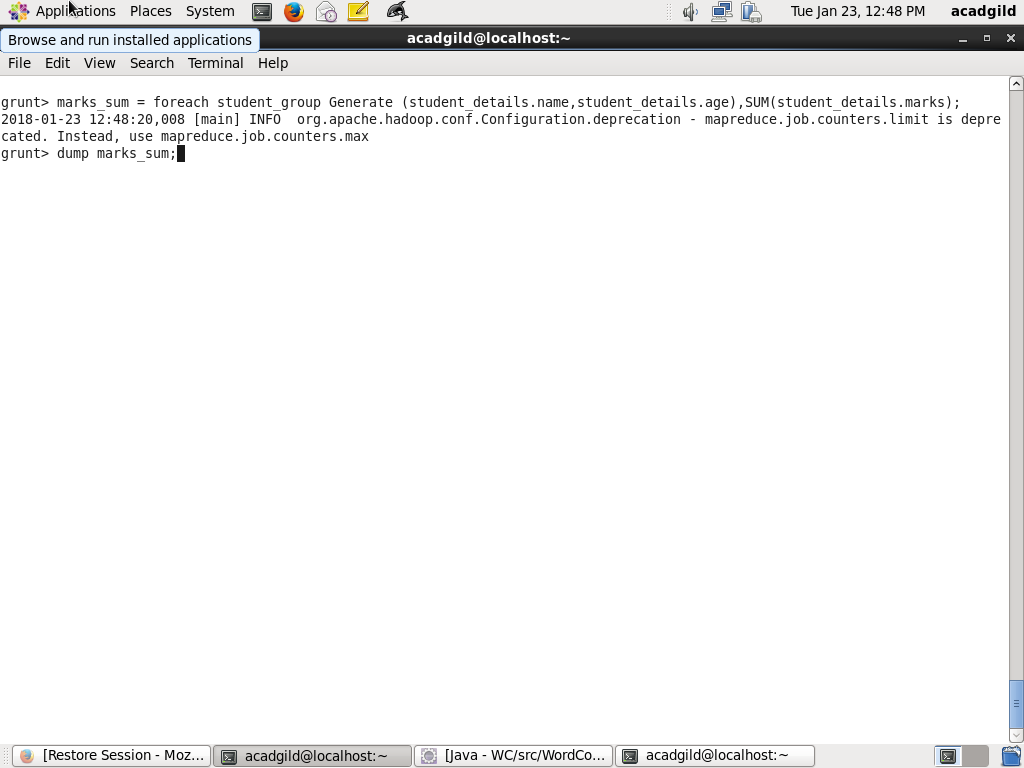
**3.SUM**

For this first applying group to the student\_details table

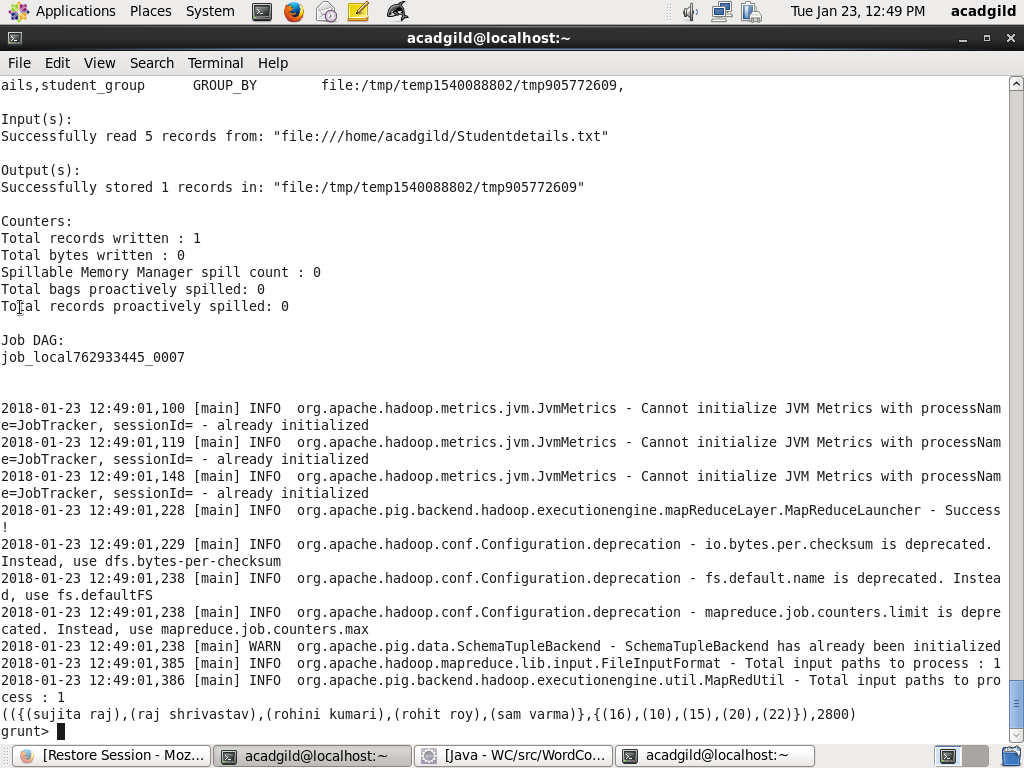
Dump student\_group table result



Then performing SUM function on marks of the student\_group as shown below

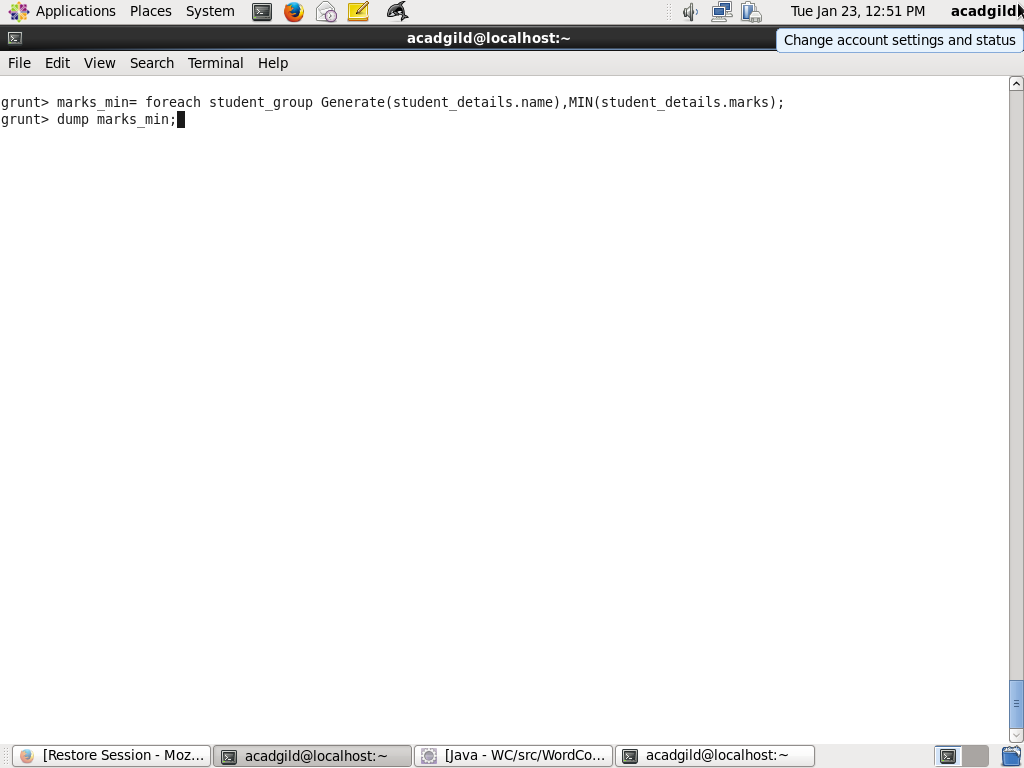


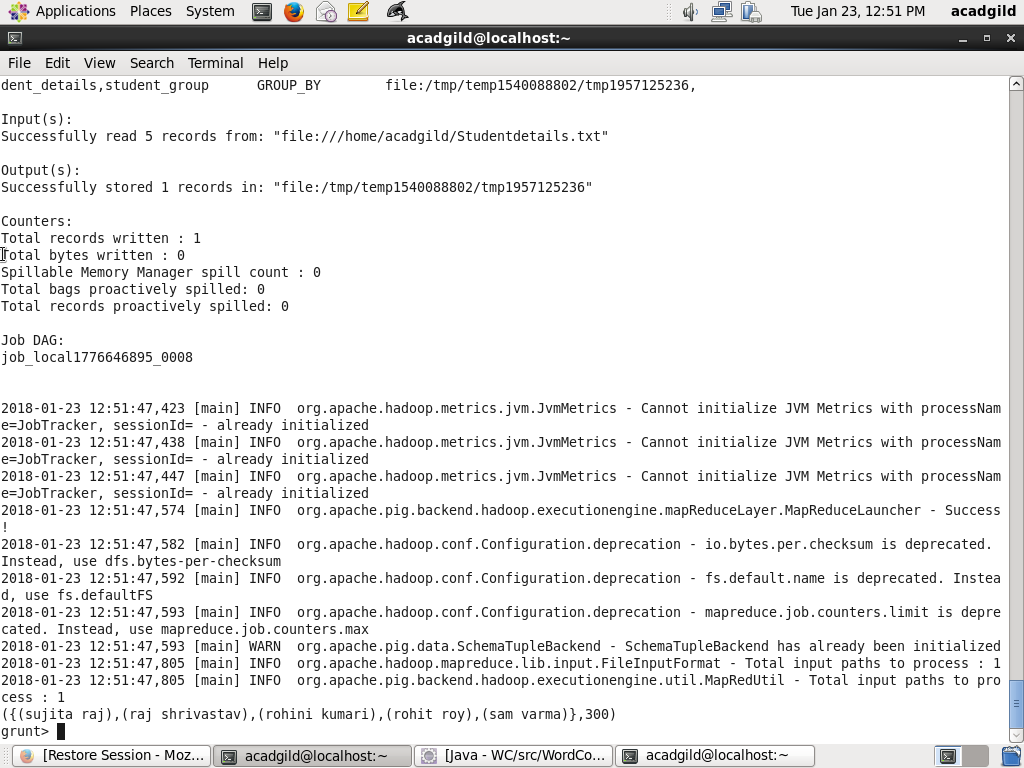
Dump mark\_sum in the grunt shell. Following is the result.



**4.MIN**

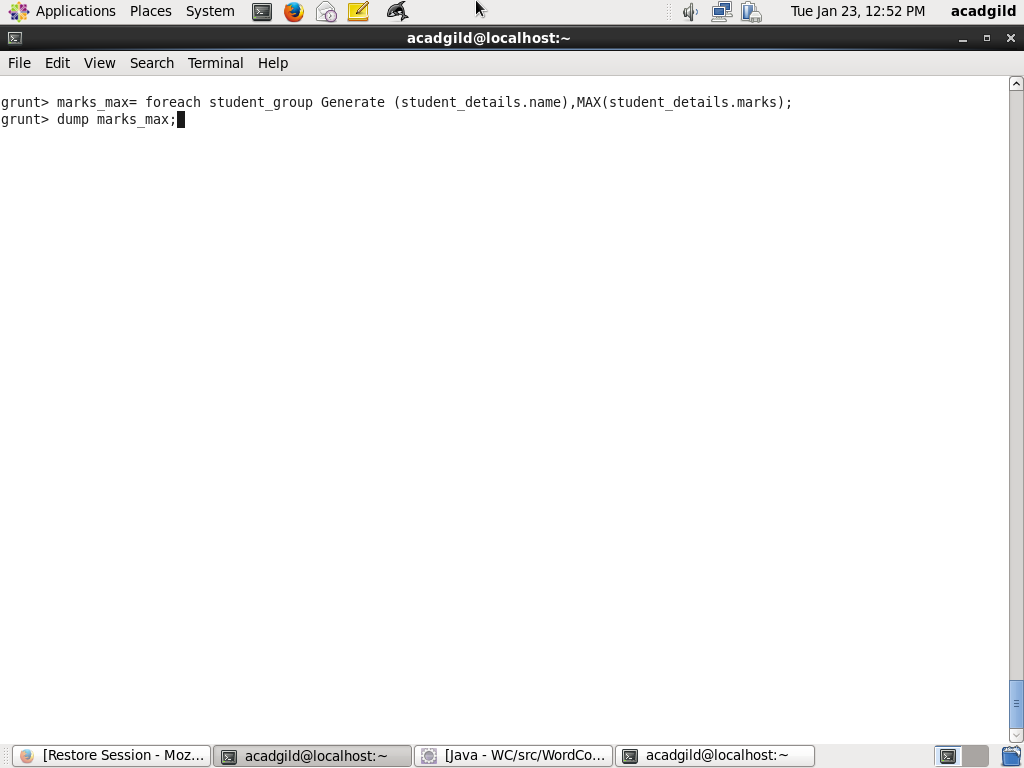
Now applying MIN to the student\_group to get the minimum marks from that as shown below

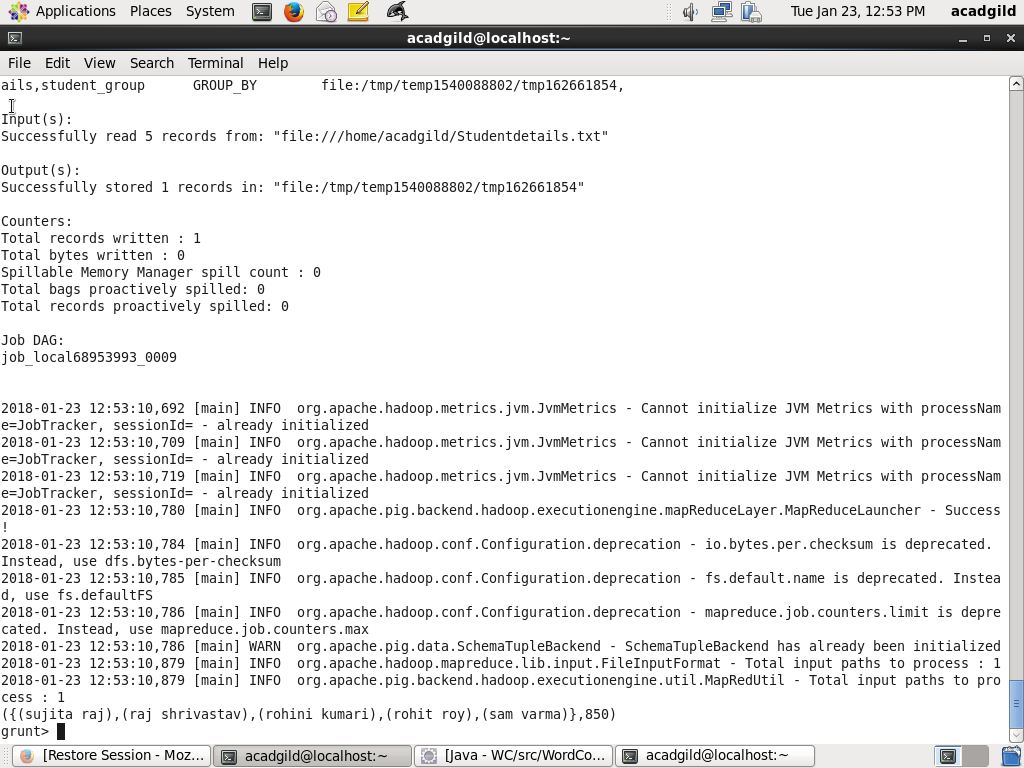




**5.MAX**

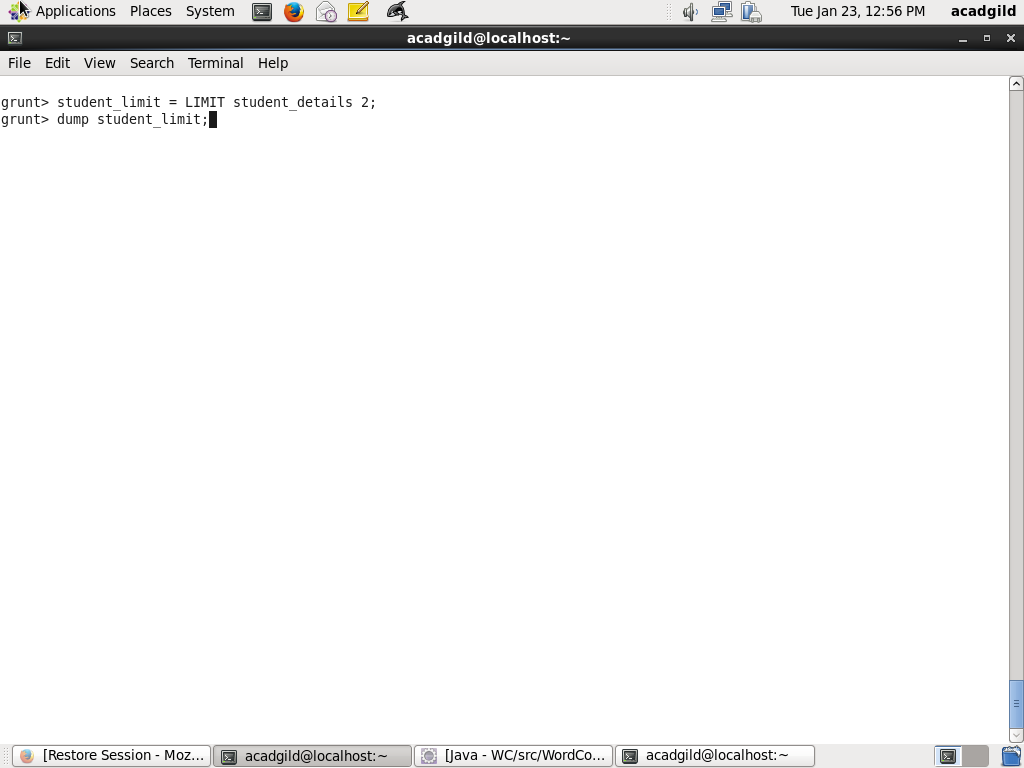
MAX is applied on student\_group to get the maximum marks .

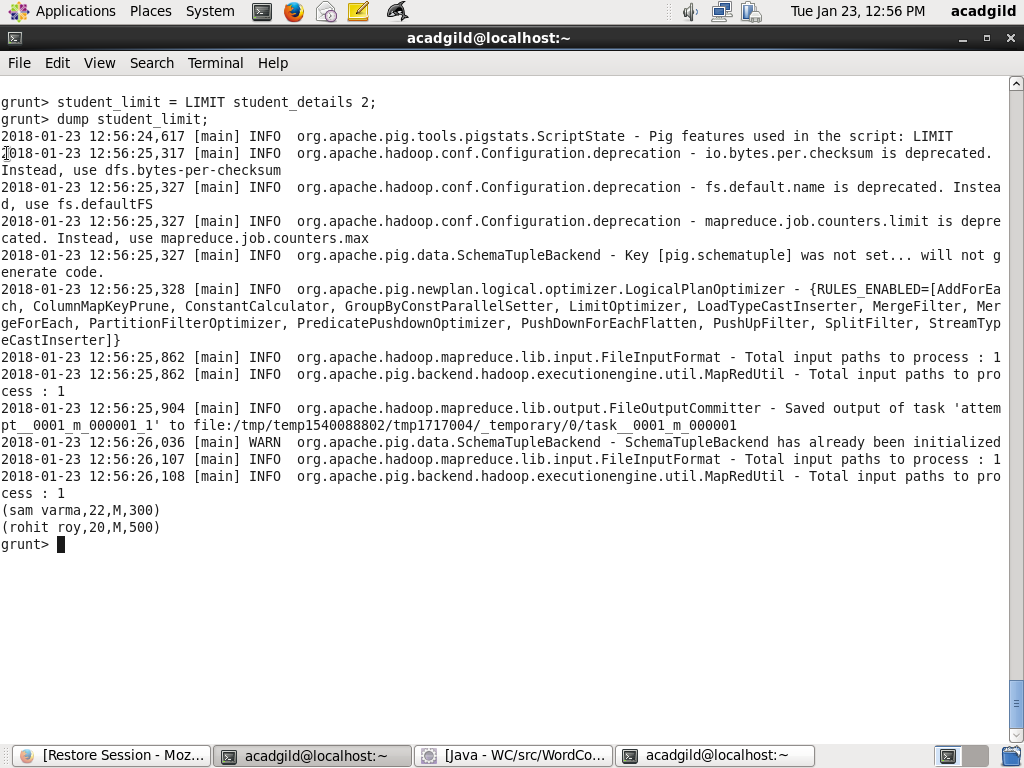




**6.LIMIT**

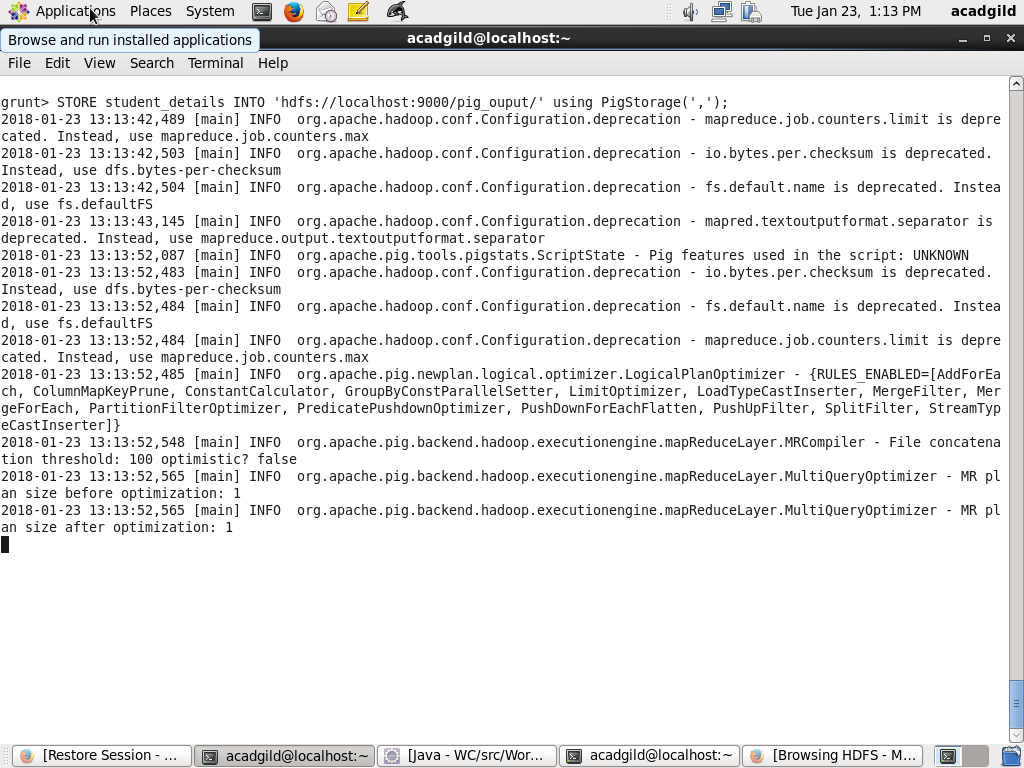
Applying LIMIT to the desired number of records from the table as shown below

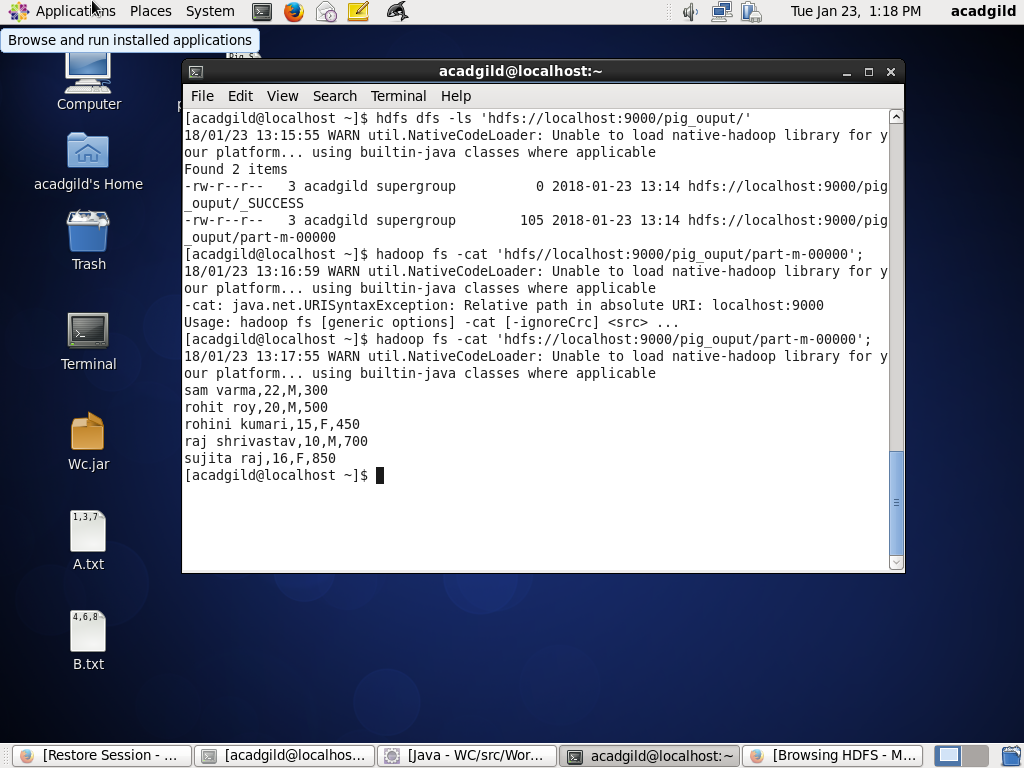




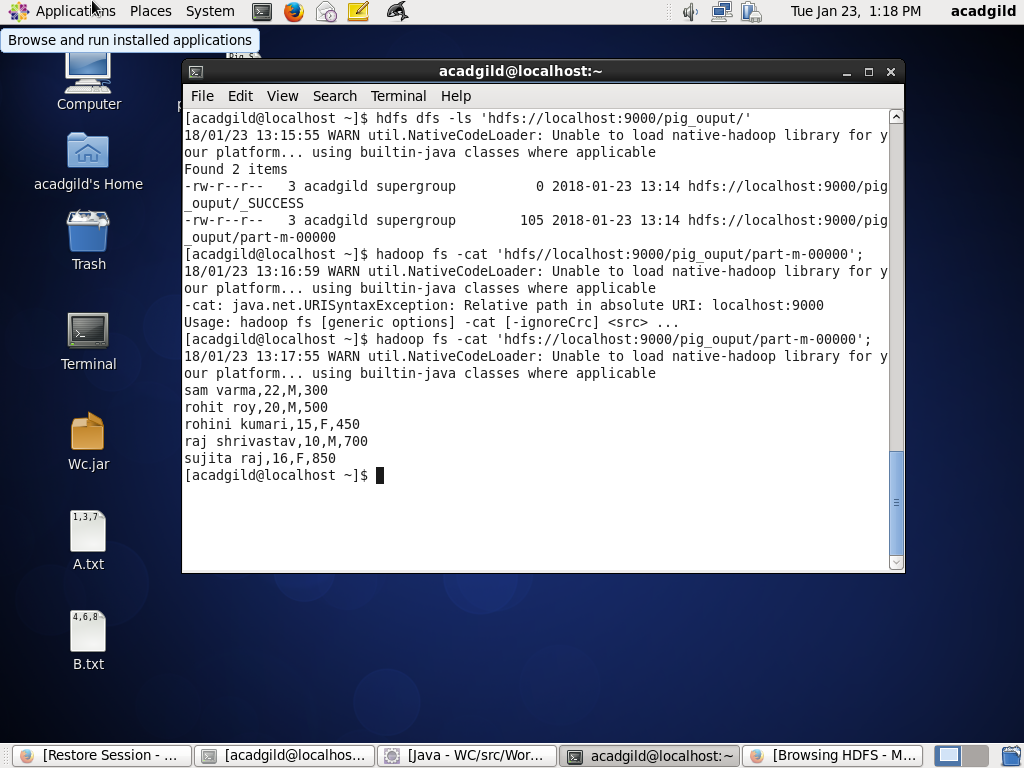
**7.STORE**

Im uing STORE command to store student details inside the hdfs directly as shown below



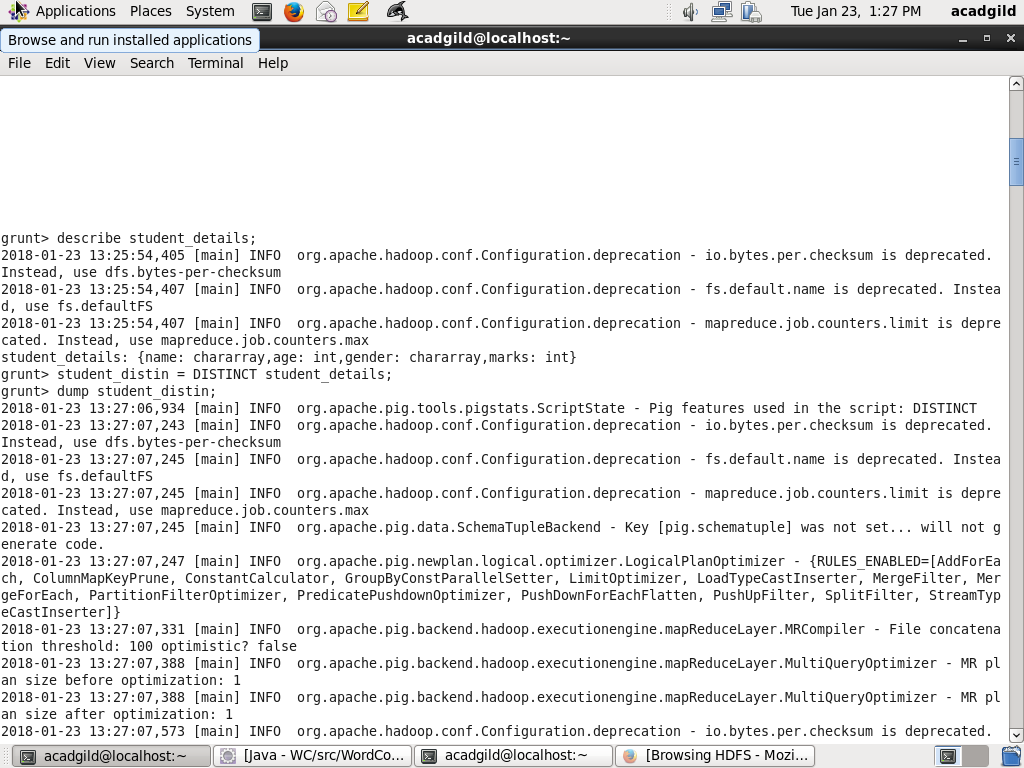


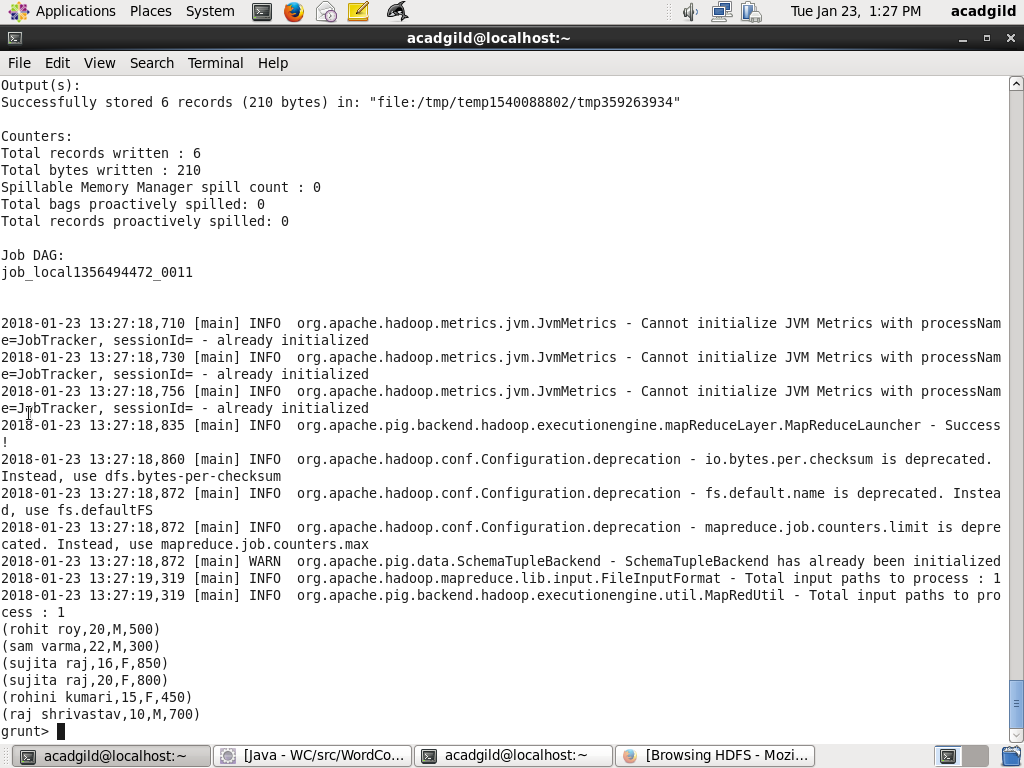
Now displaying record from hdfs using cat command



**8.DISTINCT**

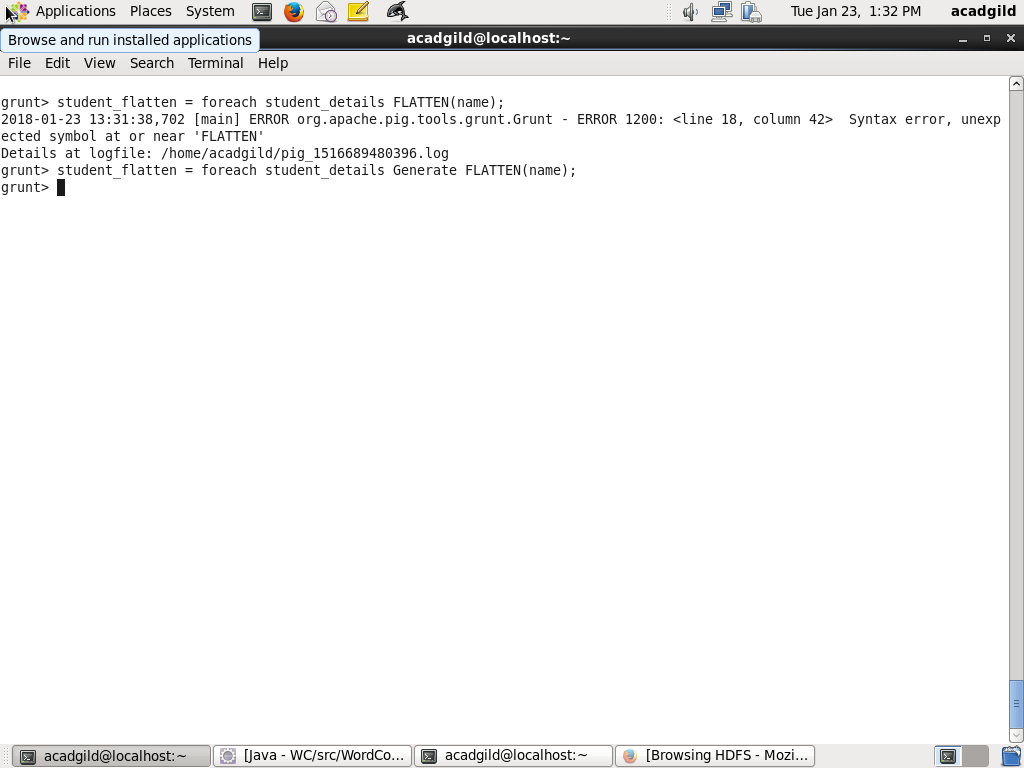
Now getting the unique records from student\_details as shown below by applying DISTINCT command

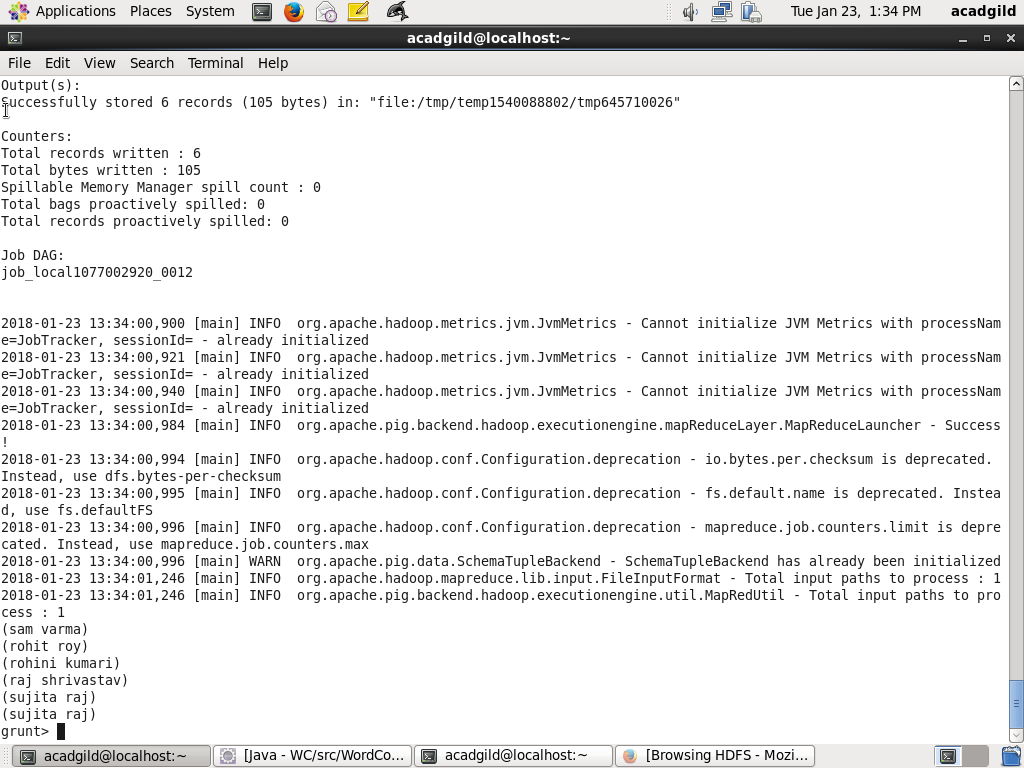




**9.FLATTEN**

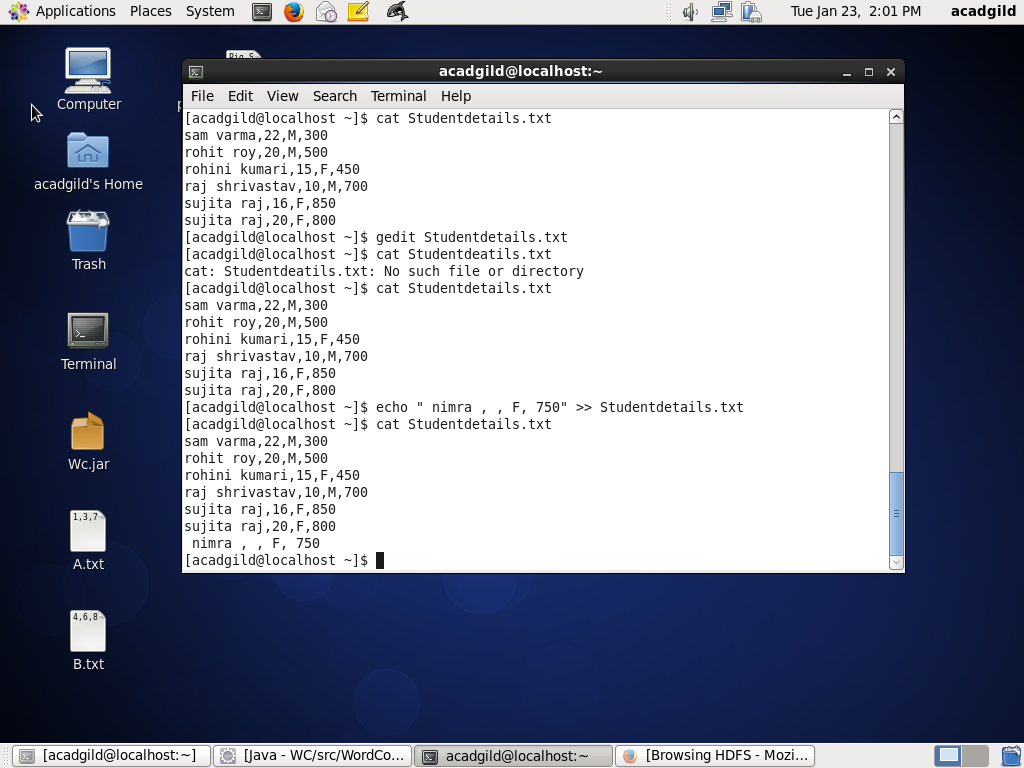
Applying FLLATEN command on name of student\_details.



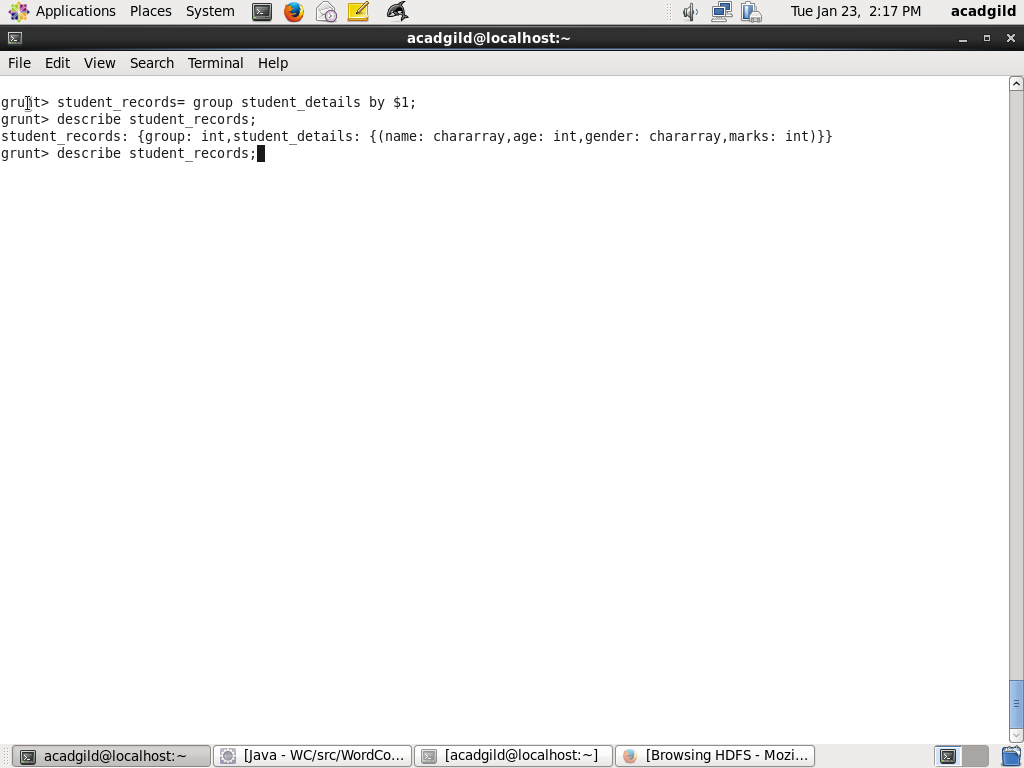
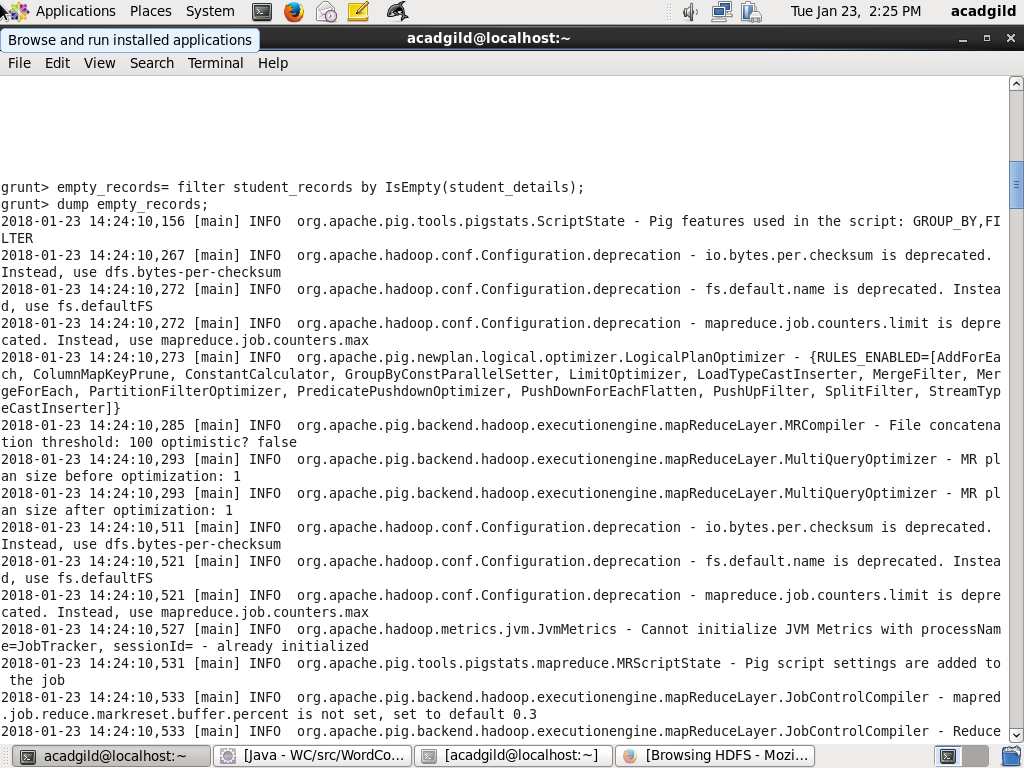


**10.IsEmpty**

Firstly inserting empty record in the studentdetails table as shown below



Applying group by $1 which is age as follows then applying IsEmpty command as follows

  
**{(nimra, , F, 750)}**