**Assignment7.3**

**Employee Details:**

101,Amitabh,20000,1

102,Shahrukh,10000,2

103,Akshay,11000,3

104,Anubhav,5000,4

105,Pawan,2500,5

106,Aamir,25000,1

107,Salman,17500,2

108,Ranbir,14000,3

109,Katrina,1000,4

110,Priyanka,2000,5

111,Tushar,500,1

112,Ajay,5000,2

113,Jubeen,1000,1

114,Madhuri,2000,2

**Employee Expenses:**

101 200

102 100

110 400

114 200

119 200

105 100

101 100

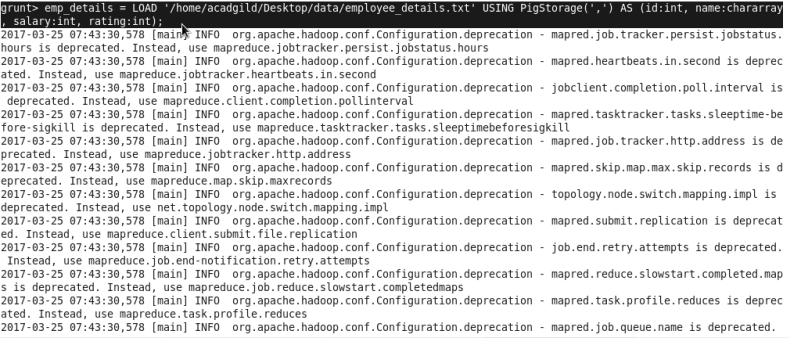
104 300

102 400

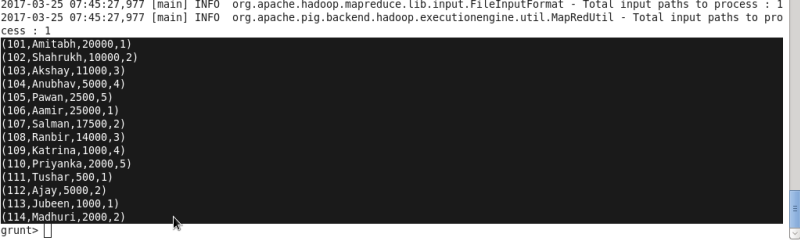
Load emp details:

emp\_details = LOAD '/home/acadgild/Desktop/data/employee\_details.txt' USING PigStorage(',') AS (id:int, name:chararray, salary:int, rating:int);

Load terminal:



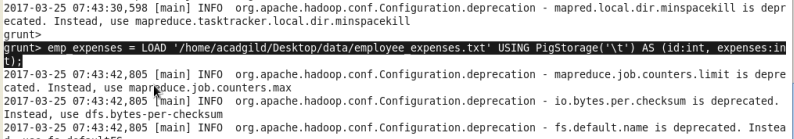
Load output:



Load employee exp details:

emp\_expenses = LOAD '/home/acadgild/Desktop/data/employee\_expenses.txt' USING PigStorage('\t') AS (id:int, expenses:int);

Load employee exp terminal:



Load employee exp output:



(a) Top 5 employees (employee id and employee name) with highest rating. (In case two employees have same rating, employee with name coming first in dictionary should get preference)

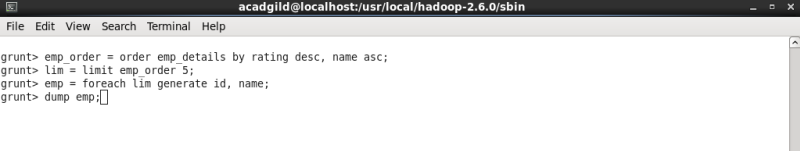
Query:

emp\_order = order emp\_details by rating desc, name asc;

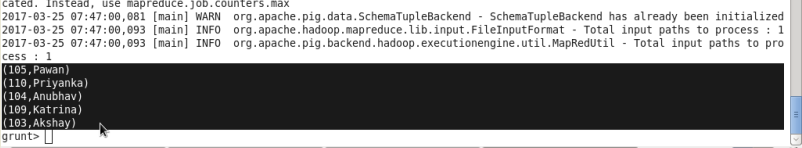
lim = limit emp\_order 5;

emp = foreach lim generate id, name;

dump emp;



output:



(b) Top 3 employees (employee id and employee name) with highest salary, whose employee id is an odd number. (In case two employees have same salary, employee with name coming first in dictionary should get preference)

Query:

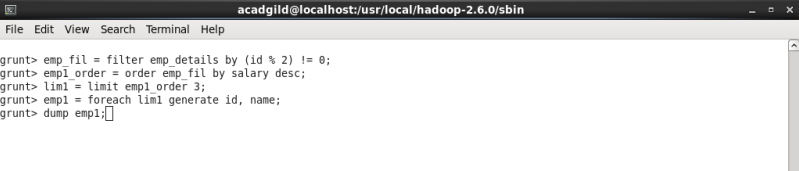
emp\_fil = filter emp\_details by (id % 2) != 0;

emp1\_order = order emp\_fil by salary desc;

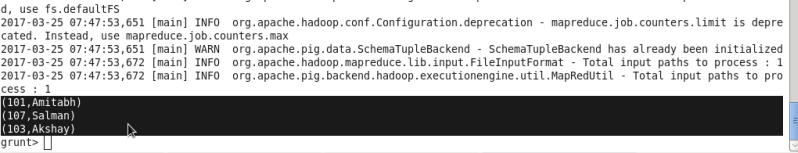
lim1 = limit emp1\_order 3;

emp1 = foreach lim1 generate id, name;

dump emp1;



output:



(c) Employee (employee id and employee name) with maximum expense (In case two employees have same expense, employee with name coming first in dictionary should get preference)

Query:

emp\_join = join emp\_details by id, emp\_expenses by id;

emp2\_order = order emp\_join by expenses desc, name asc;

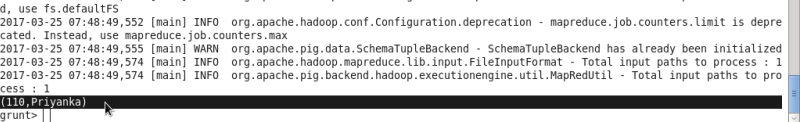
lim2 = limit emp2\_order 1;

emp2 = foreach lim2 generate $0,$1;

dump emp2;



Output:



(d) List of employees (employee id and employee name) having entries in employee\_expenses file.

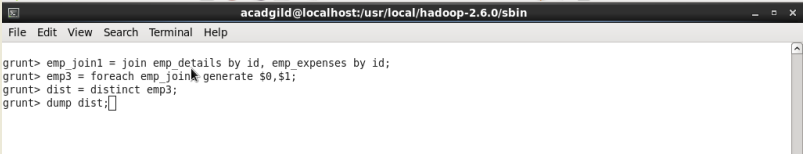
Query:

emp\_join1 = join emp\_details by id, emp\_expenses by id;

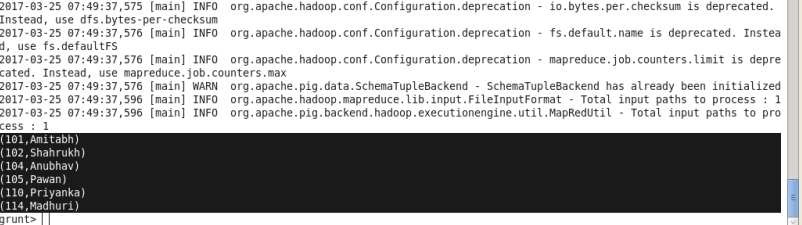
emp3 = foreach emp\_join1 generate $0,$1;

dist = distinct emp3;

dump dist;



Output:



(e) List of employees (employee id and employee name) having no entry in employee\_expenses file.

Query:

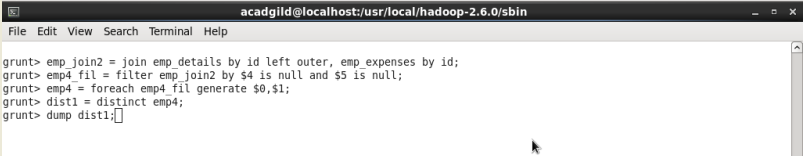
emp\_join2 = join emp\_details by id left outer, emp\_expenses by id;

emp4\_fil = filter emp\_join2 by $4 is null and $5 is null;

emp4 = foreach emp4\_fil generate $0,$1;

dist1 = distinct emp4;

dump dist1;



output:

