Input files :

employee\_detail.txt

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.txt

101 200

102 100

110 400

114 200

119 200

105 100

101 100

104 300

102 400

(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)

script:

emp\_details = LOAD 'employee\_details.txt' USING PigStorage(',') AS

(emp\_id:int, emp\_name:chararray,emp\_salary:int,emp\_rating:int);

ordered = order emp\_details by emp\_rating DESC, emp\_name ASC;

top = limit ordered 5;

dump top;

output:

(105,Pawan,2500,5)

(110,Priyanka,2000,5)

(104,Anubhav,5000,4)

(109,Katrina,1000,4)

(103,Akshay,11000,3)

~

(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)

script:

emp\_details = LOAD 'employee\_details.txt' USING PigStorage(',') AS

(emp\_id:int, emp\_name:chararray,emp\_salary:int,emp\_rating:int);

ordered = order emp\_details by emp\_salary DESC, emp\_name ASC;

odd\_empid\_fields = filter ordered by emp\_id%2 != 0;

top = limit odd\_empid\_fields 3;

dump top;

output:

(101,Amitabh,20000,1)

(107,Salman,17500,2)

(103,Akshay,11000,3)

(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)

script:

emp\_details = LOAD 'employee\_details1.txt' USING PigStorage(',') AS

(emp\_id:int, emp\_name:chararray,emp\_salary:int,emp\_rating:int);

employee\_expenses = LOAD 'employee\_expenses.txt' USING

PigStorage('\t') AS (emp\_id:int,emp\_expenses:int);

join\_employee\_details = JOIN emp\_details by emp\_id,

employee\_expenses by emp\_id;

ordered = order join\_employee\_details by emp\_expenses DESC, emp\_name

ASC;

dump ordered;

output:

(110,Priyanka,2000,5,110,400)

(102,Shahrukh,10000,2,102,400)

(104,Anubhav,5000,4,104,300)

(101,Amitabh,20000,1,101,200)

(114,Madhuri,2000,2,114,200)

(101,Amitabh,20000,1,101,100)

(105,Pawan,2500,5,105,100)

(102,Shahrukh,10000,2,102,100)

(d) List of employees (employee id and employee name) having entries in

employee\_expenses

file.

script:

emp\_details = LOAD 'employee\_details1.txt' USING PigStorage(',') AS

(emp\_id:int, emp\_name:chararray,emp\_salary:int,emp\_rating:int);

employee\_expenses = LOAD 'employee\_expenses.txt' USING

PigStorage('\t') AS (emp\_id:int,emp\_expenses:int);

join\_employee\_details = JOIN emp\_details by emp\_id,

employee\_expenses by emp\_id;

dump join\_employee\_details;

output:

(101,Amitabh,20000,1,101,100)

(101,Amitabh,20000,1,101,200)

(102,Shahrukh,10000,2,102,400)

(102,Shahrukh,10000,2,102,100)

(104,Anubhav,5000,4,104,300)

(105,Pawan,2500,5,105,100)

(110,Priyanka,2000,5,110,400)

(114,Madhuri,2000,2,114,200)

(e) List of employees (employee id and employee name) having no entry in

employee\_expenses file.

script :

emp\_details = LOAD 'employee\_details1.txt' USING PigStorage(',') AS

(emp\_id:int, emp\_name:chararray,emp\_salary:int,emp\_rating:int);

employee\_expenses = LOAD 'employee\_expenses.txt' USING

PigStorage('\t') AS (emp\_id:int,emp\_expenses:int);

join\_employee\_details = JOIN emp\_details by emp\_id LEFT OUTER,

employee\_expenses by emp\_id;

dump join\_employee\_details;

output:

(101,Amitabh,20000,1,101,100)

(101,Amitabh,20000,1,101,200)

(102,Shahrukh,10000,2,102,400)

(102,Shahrukh,10000,2,102,100)

(103,Akshay,11000,3,,)

(104,Anubhav,5000,4,104,300)

(105,Pawan,2500,5,105,100)

(106,Aamir,25000,1,,)

(107,Salman,17500,2,,)

(108,Ranbir,14000,3,,)

(109,Katrina,1000,4,,)

(110,Priyanka,2000,5,110,400)

(111,Tushar,500,1,,)

(112,Ajay,5000,2,,)

(113,Jubeen,1000,1,,)

(114,Madhuri,2000,2,114,200)