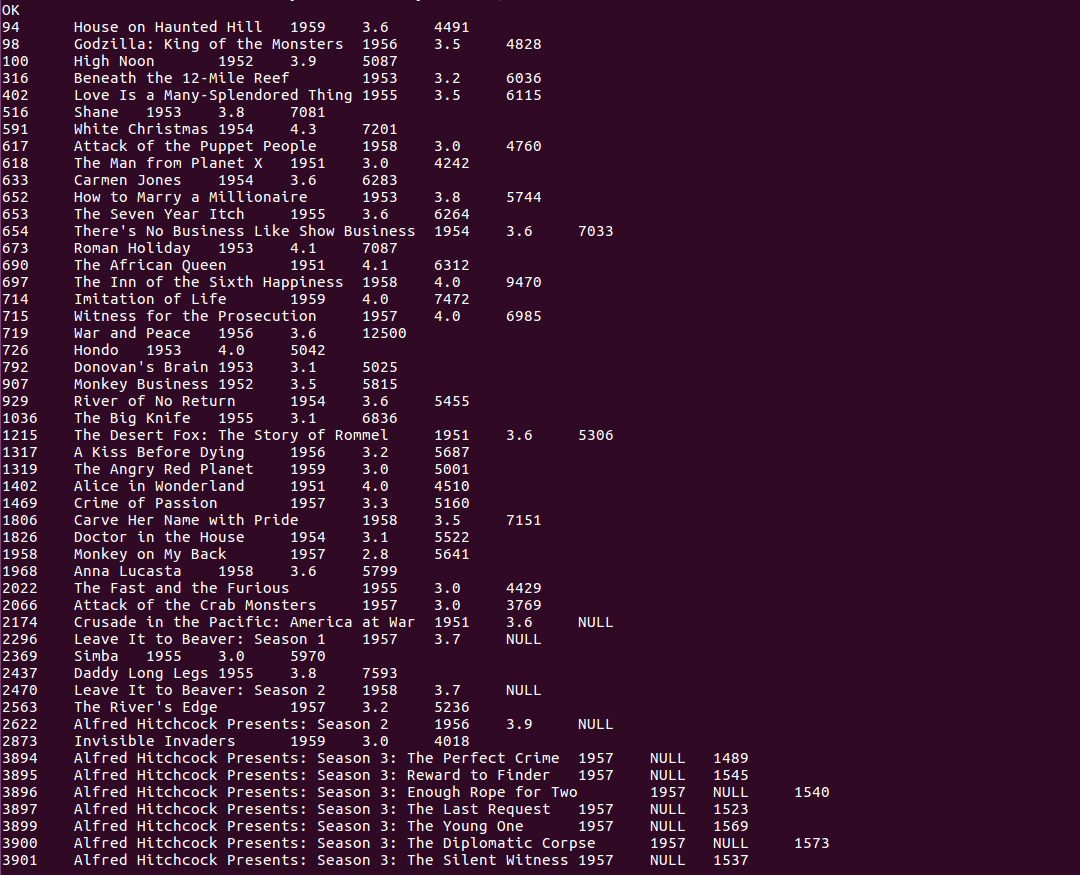
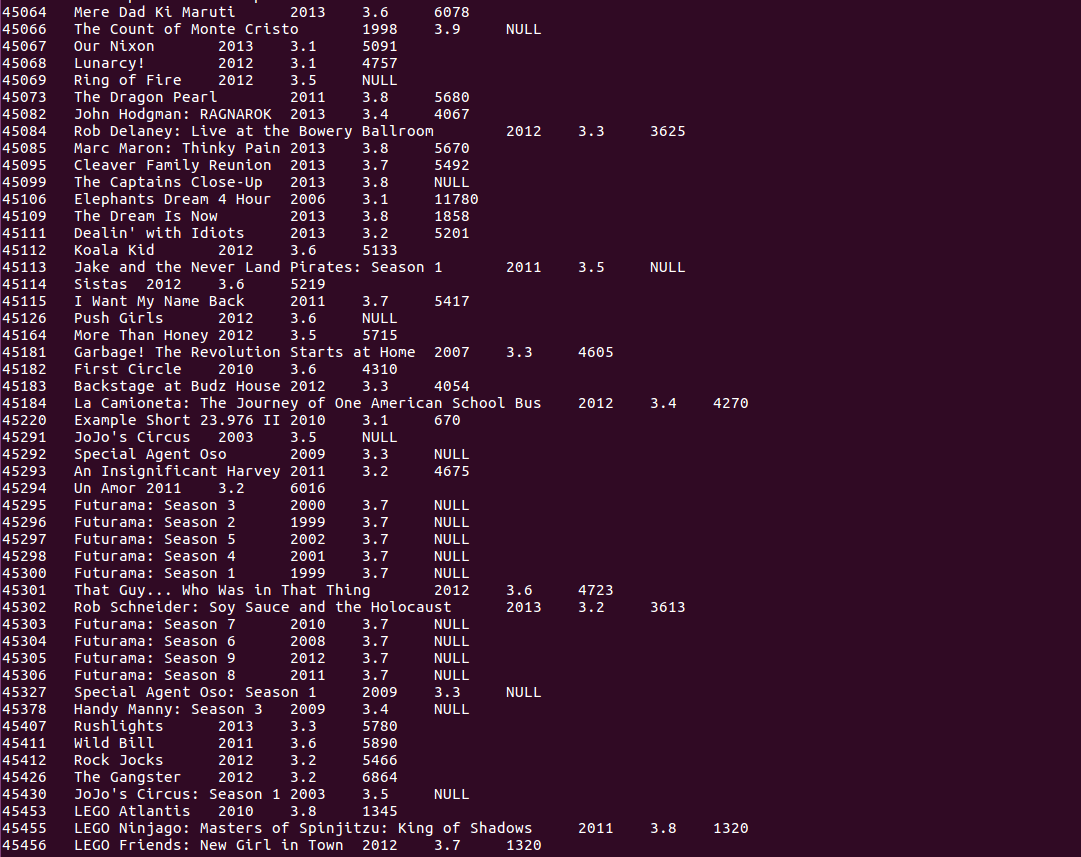
create table Movies(id int, name string, year int, rating double, duration int) row format delimited fields terminated by ',';  
  
load data local inpath '/home/iritika/Desktop/movies.txt' into table Movies;  
  
A)select \*from Movies where year>1950 and year<1960;

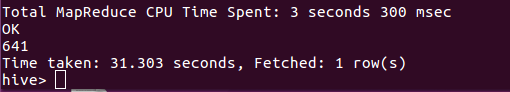


B)select count(id) from Movies where rating>4.0;

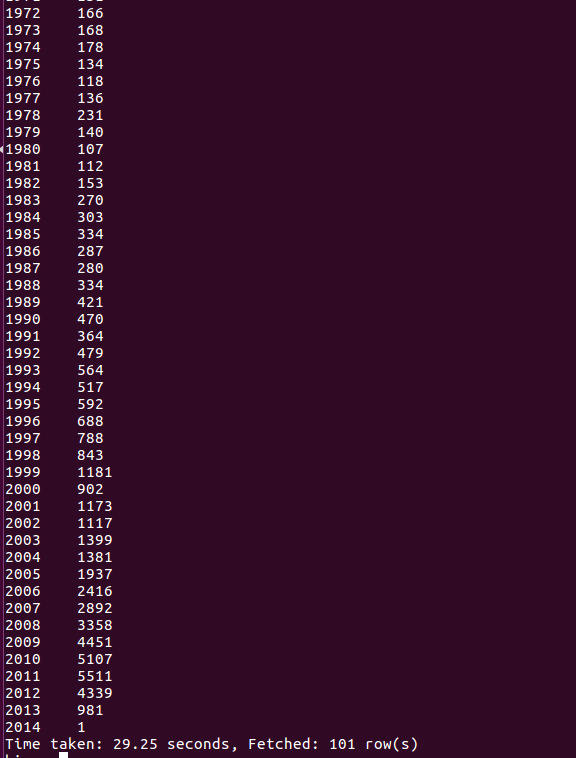


C)select \*from Movies where rating>3.0 and rating<4.0;

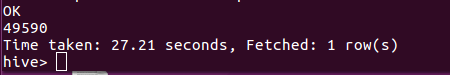
  
D)select count(id) from Movies where duration>7200;



E)select year, count(id) as c from Movies group by year;

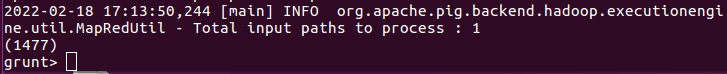


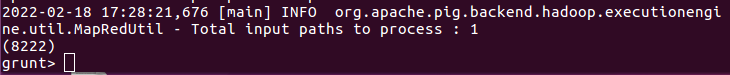
F)select count(id) from Movies;



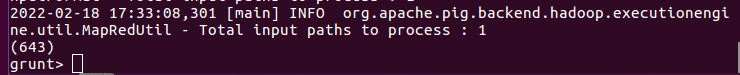
bin/hdfs dfs -put /home/iritika/Desktop/movies.txt /pig  
  
Movies = load '/pig/movies.txt' using PigStorage(',') as (id:int, name:chararray, year:int, rating:double, duration:int);  
  
A)A1 = filter Movies by year>=1950;  
A2 = filter A1 by year<=1960;  
A3 = group A2 all;  
A4 = foreach A3 generate COUNT(A2);  
dump A4;

  
  
B)B1 = filter Movies by rating>=4.0;  
B2 = group B1 all;  
B3 = foreach B2 generate COUNT(B1);  
dump B3;

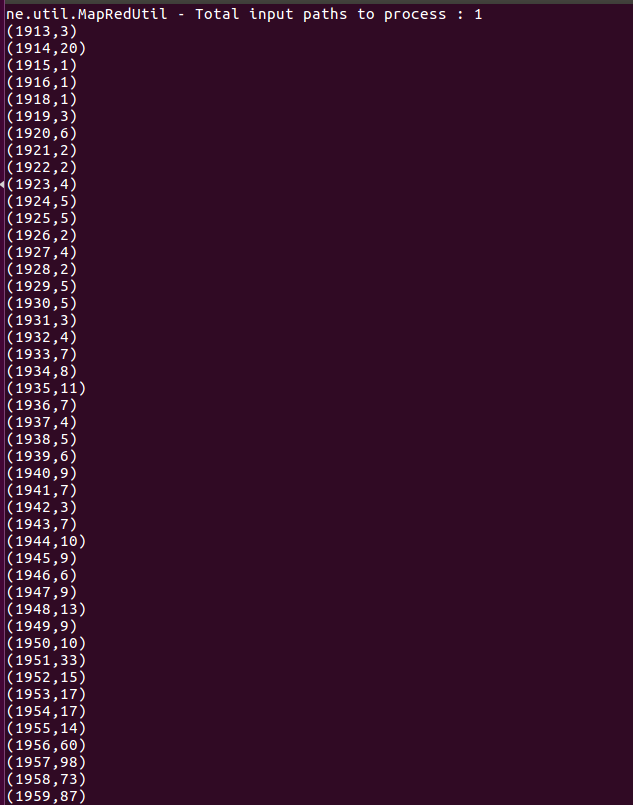
  
  
C)C1 = filter Movies by rating>=3.0;  
C2 = filter C1 by rating<=4.0;  
C3 = foreach C2 generate name;  
C4 = group C2 all;  
C5 = foreach C4 generate COUNT(C2);  
dump C5;

  
dump C3;

  
  
D)D1 = filter Movies by duration>=7200;  
D2 = group D1 all;  
D3 = foreach D2 generate COUNT(D1);  
dump D3;



E)E1 = foreach Movies generate year;  
E2 = group E1 by year;  
E3 = foreach E2 generate group, count(E1) as numOfMovies;  
dump E3;

  
  
F)F1 = group Movies all;  
F2 = foreach F1 generate COUNT(Movies);  
dump F4;

