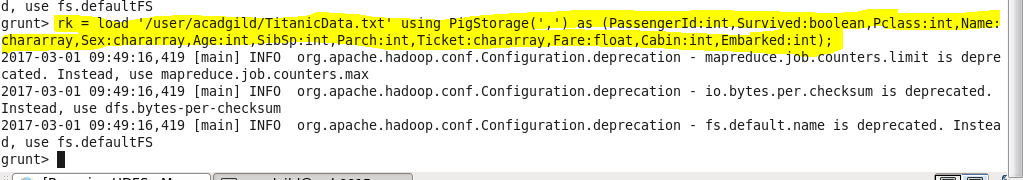
ASSIGNMENT6.4

1)A pig script to find the average of male and female persons died in the titanic tragedy.

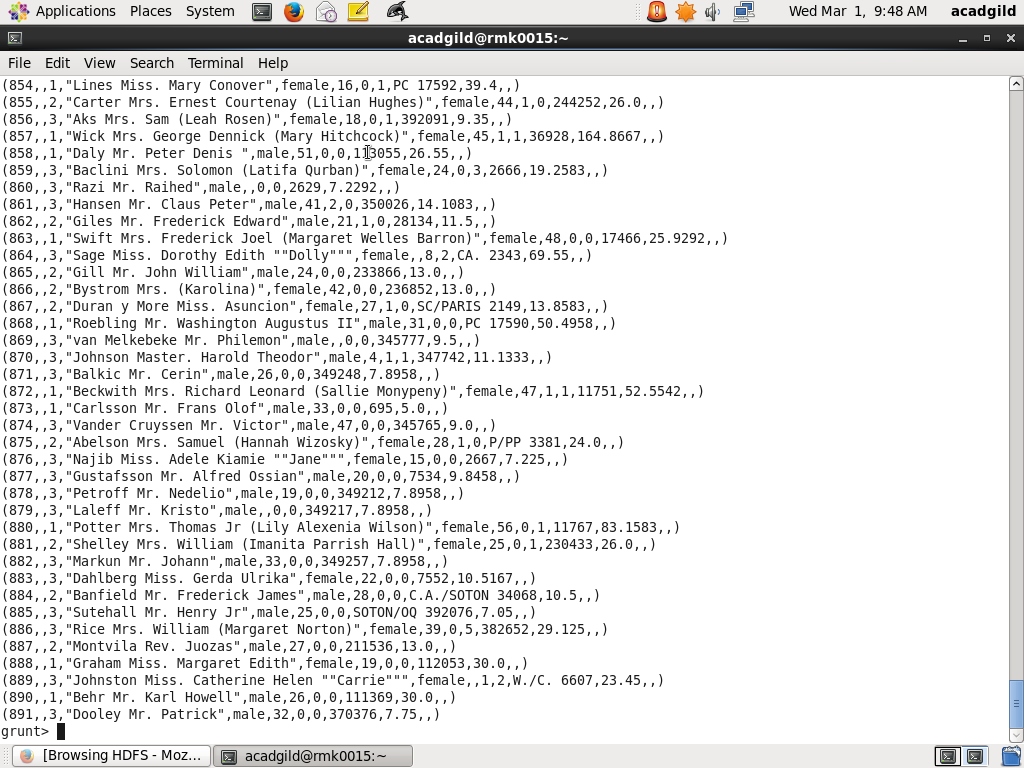
A) rk = load '/user/acadgild/TitanicData.txt' using PigStorage(',') as

(Passengerid:int,Survived:boolean,Pclass:int,Name:chararray,Sex:chararray,

Age:int,SibSp:int,Parch:int,Ticket:chararray,Fare:float,Cabin:chararray,Embarked:)



dump rk;

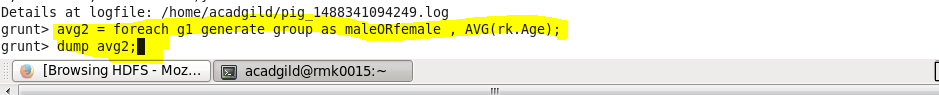


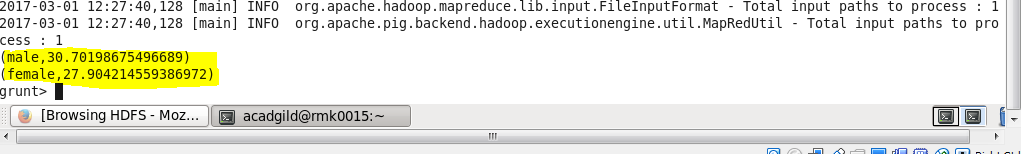
g2 = group rk by Sex;

dump g2;

avg2 = foreach g1 generate group as maleORfemale,AVG(rk.Age);

dump avg2;





THE AVERAGE OF MALES IN THE TITANIC TRAGEDY =30.70

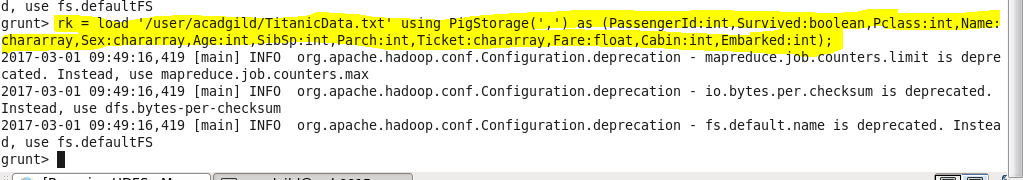
THE AVERAGE OF ALL FEMALES IN TITANIC TRAGEDY =27.90 shown in screenshot

2)A pig script to find the number of people died and survived in the each class.

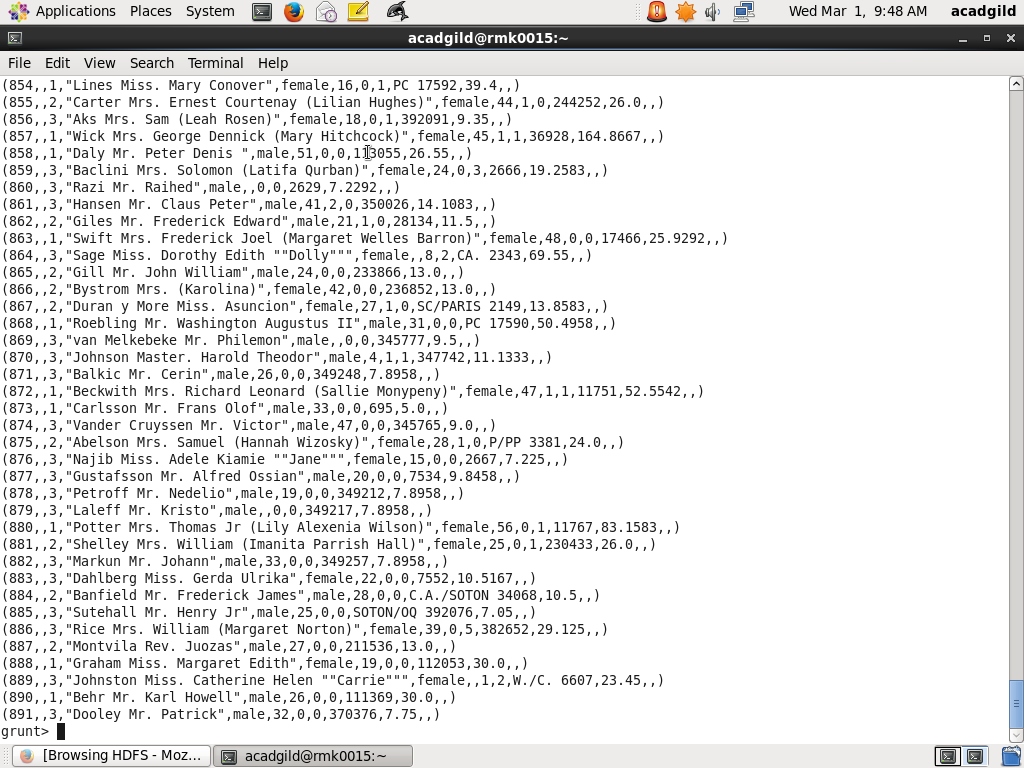
A) rk = load '/user/acadgild/TitanicData.txt' using PigStorage(',') as

(Passengerid:int,Survived:int,Pclass:int,Name:chararray,Sex:chararray,

Age:int,SibSp:int,Parch:int,Ticket:chararray,Fare:float,Cabin:chararray,Embarked:);



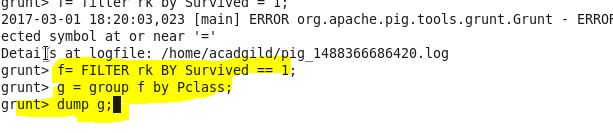
dump rk;



f = filter rk by Survived==1;

g = group f by Pclass;

dump g;



C1 = foreach g generate group ,COUNT(f.PassengerId);

dump c1;

