**Assignment 9 2**

Loading data:

**studentDS**= LOAD 'studentDS.txt' USING PigStorage(',') AS (name:chararray,col:chararray,dob:chararray,course\_stream:chararray,grade:float,state:chararray,city:chararray);

**Task1:**

grade= FILTER studentDS BY grade<5;

grouped= GROUP grade all;

count= FOREACH grouped GENERATE COUNT(grade.grade);

dump count;

**Output:**

(188)

**Task2:**

filtered= FILTER studentDS BY state=='alaska';

result= FOREACH filtered GENERATE name;

DUMP result;

**Output:**

(BELL)

(DAVIS)

(LOGAN)

(YOUNG)

(COWEN)

(DISHONG)

(GIBSON)

(IMMILE)

(KEYSER)

(POLK)

(ROBTS&CMRN)

(STEPHENS)

(WEIRS)

(JACOB)

(BARBARA)

(GEORGE)

(SARAH)

(WILLIAM)

(JOHN)

(JOHN)

(STILLWEL)

(Albright)

(Bowman)

(Brunner)

(Carson)

(Diller)

(Denyes)

(Ellmaker)

(Fisher)

(Gumpf)

(Hartman)

(Halbach)

(Huber)

(Houser)

(Jordan)

(Keubler)

(Kendig)

(Long)

(McGinnis)

(Miller)

(Nourse)

(Quinn)

(Rotharmel)

(Sherer)

(Sherwood)

(Voight)

(Waters)

(Wiley)

(Zahm)

(Matthew)

(Peter)

(Christian)

(George)

(Alexander)

(Frederick)

(Michael)

(Richard)

(Jacob)

(Joseph)

(John W)

(Michael O)

(Daniel)

(William)

(William)

(Joseph)

(Henry)

(Samuel)

(John)

**Task3:**

filtered= FILTER studentDS BY ((state=='alabama') AND (col=='goverenment'));

grouped= GROUP filtered all;

counted\_result= FOREACH grouped GENERATE COUNT(filtered.col);

DUMP counted\_result;

**Output:**

(2198)

**Task4:**

filtered= FILTER studentDS BY ((state=='oregon') AND (course\_stream=='BE'));

names= FOREACH filtered GENERATE name;

DUMP names;

**Output:**

(BLOOM)

(PATTERSON)

(WAGGONER)

(IMMILE)

(SHOCK)

(WHITE)

(JOHN)

(RICHARD)

(JACOB)

(Albright)

(Buckius)

(Bowman)

(Cordes)

(Danner)

(Donas)

(Ehrisman)

(Frick)

(Grimler)

(Heinitsh)

(Himmelsbach)

(Hubley)

(Kitch)

(Lambert)

(McNair)

(Naumann)

(Royal)

(Shindle)

(Wayne)

(Reformed)

(P.K.)

(Charles H)

(Jacob)

(Daniel)

(George)

(John)

(Peter)

(George B)

(Philip)

(Dennis)

(George)

(Edwin)

(Frederick)