**ASSIGNMENT 7.6**

**PROBLEM STATEMENT :**

**1. Write a macro in pig to filter out the people with age >45 from the below given dataset.**

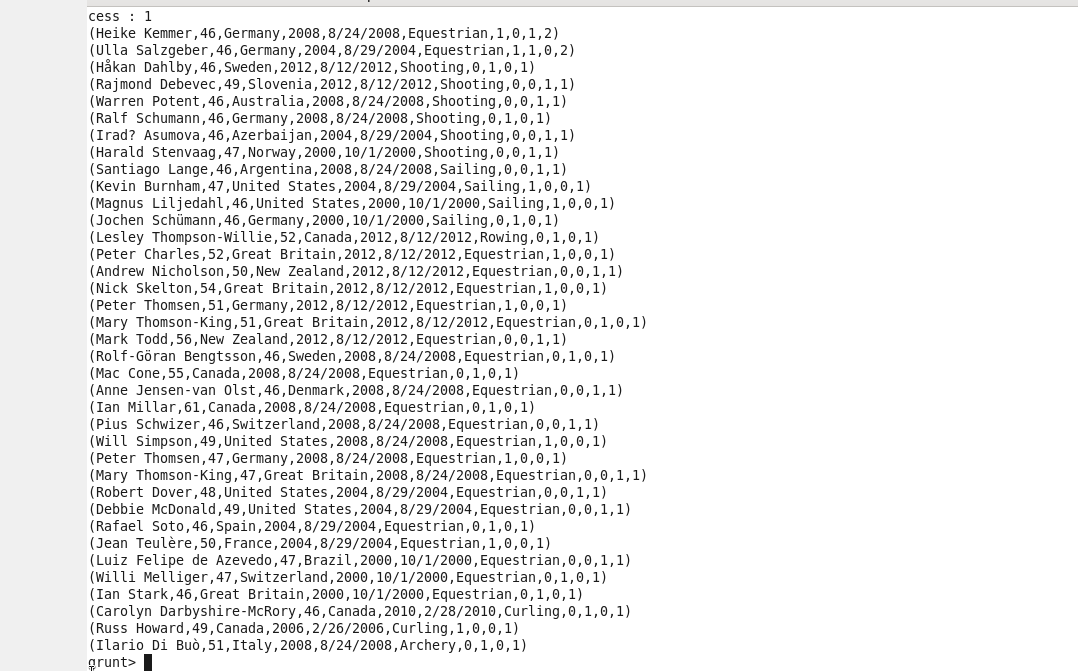
**STEP 1:**DEFINE myfilter(relvar,colvar) returns x{$x = filter $relvar by $colvar > 45;};

**STEP 2:**a = LOAD '/home/acadgild/olympix\_data.csv' USING PigStorage(',') as (name:chararray, age:int, country:chararray, year:chararray,c\_date:chararray, sport:chararray, gold:int, silver:int, bronze:int, total:int);

**STEP 3**: x1 = myfilter(a,age);

**STEP 4 :**dump x1;

**STEP 5:** OUTPUT REQUIRED



**2. Find out the number of medals each country won in archery.**

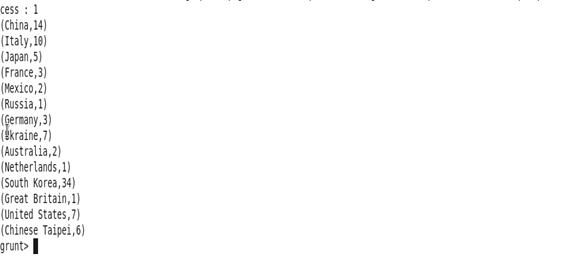
**STEP 1 :** a\_filter1 = FILTER a by sport =='Archery';

**STEP 2 :** athletes\_grp\_country = GROUP a\_filter1 by country;

**STEP 3 :** medal\_sum = FOREACH athletes\_grp\_country GENERATE group AS country, SUM(a\_filter1.total) as medal\_count;

**STEP 4 :** DUMP medal\_sum;

**STEP 5: OUTPUT REQUIRED**



**3. Find out the number of medals india won year wise.**

**STEP 1:** a\_filter = FILTER a by country == 'India';

**STEP 2 :** grp = GROUP a\_filter by year;

**STEP 3 :** total\_medals = FOREACH grp GENERATE group AS year, SUM(a\_filter.total) as total\_medal;



**STEP 4 :** dump total\_medals;

**STEP 5 : OUTPUT REQUIRED**

