**Pokemon Data Analysis**

The Pokémon Fight League (PFL) management for the 2017 match has first of all decided a minimum criterion for the entry selection process that filters through the defense power for any Pokémon, which should ideally be greater than 55.

Hence, the eligible list will be randomly formed after filtering out the Pokémons with a defenseless than 55.

Furthermore,Our job is to give 2 list of names of those Pokémons who will be eligible for taking part in PFL this year from the list of all the participating 800 Pokémons.

Download Pokemon.csv and copy it to the local folder (/home/acadgild/Pokemon)

So First of all, we will load the dataset inside PIG.

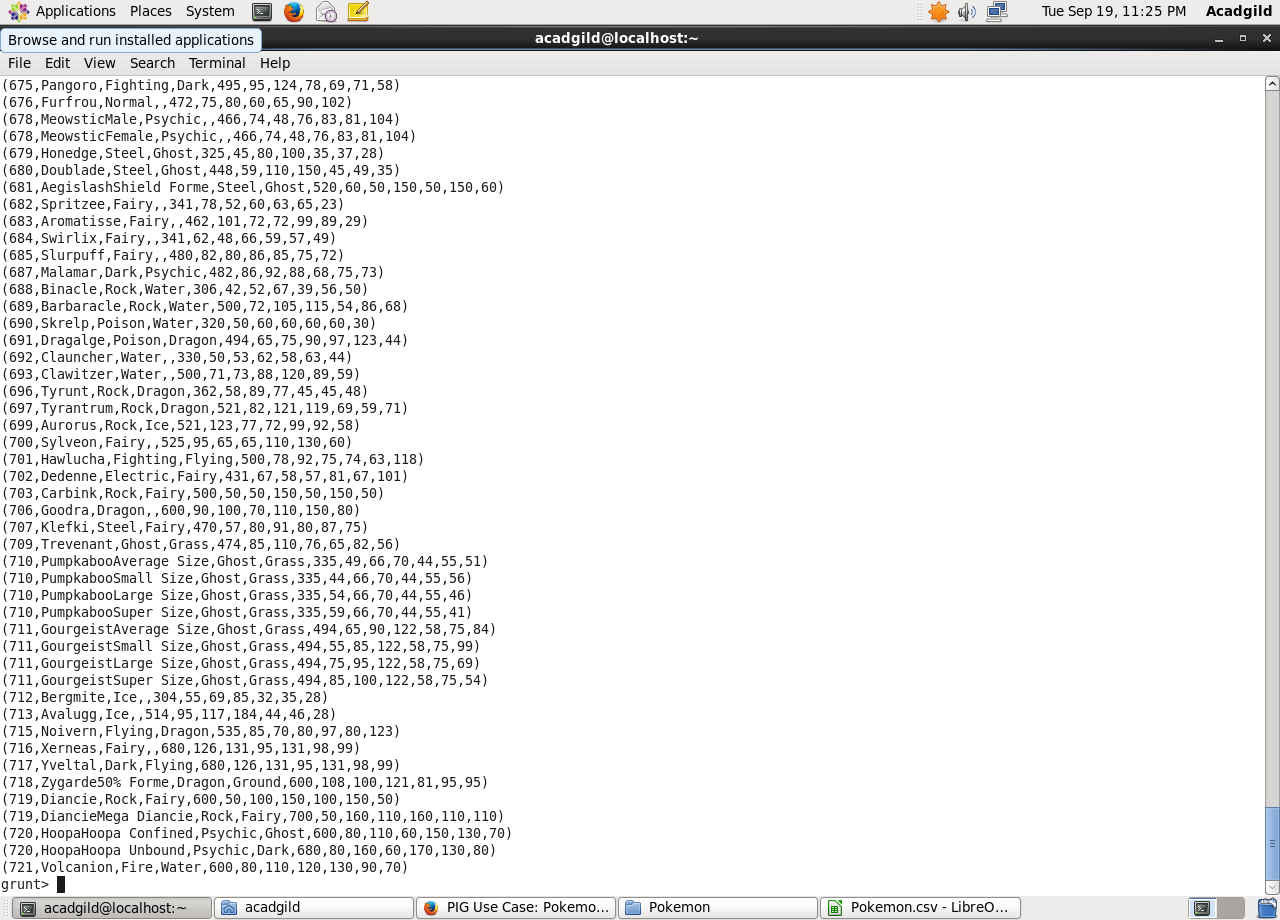
1. Load\_Data = LOAD ‘/home/acadgild/Pokemon/Pokémon.csv’ USING PigStorage(‘,’) AS(Sno:int,Name:chararray,Type1:chararray,Type2:chararray,Total:int,HP:int,Attack:int,Defense:int,SpAtk:int,SpDef:int,Speed:int);
2. Dump ;

**Ques 1: Find the list of players that have been selected in the qualifying round (DEFENCE>55)**

1. selected\_list = FILTER Load\_Data BY Defense>55;

2. Dump;

The dataset is filtered, and hence out of all the 800 Pokémons, only 544 are eligible to take part in the tournament. In order to get the count, refer the next problem statement.



**Ques 2: State the number of players taking part in the competition after getting selected in the qualifying round**.

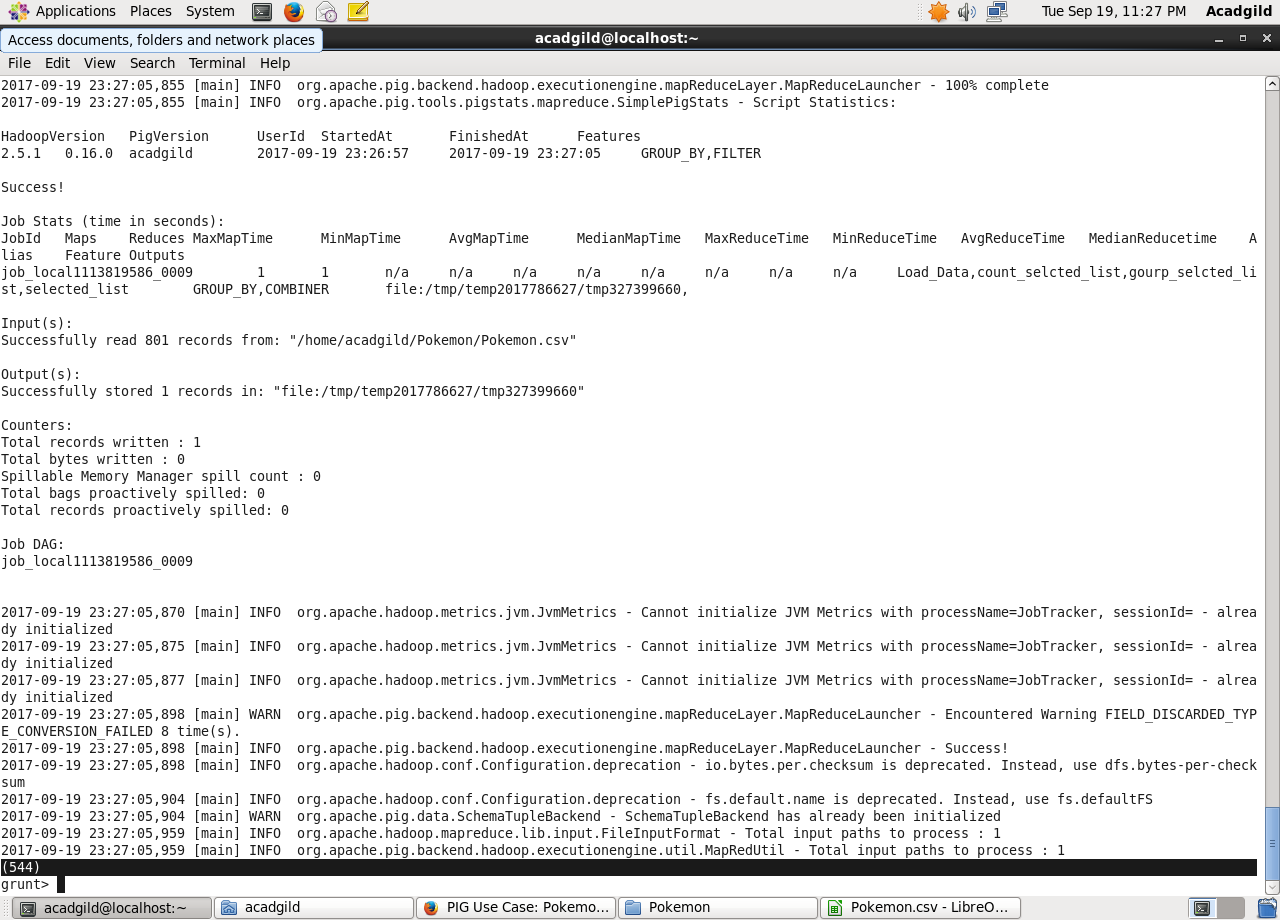
Command

1. gourp\_selcted\_list = Group selected\_list All;

2. count\_selcted\_list = foreach gourp\_selcted\_list GENERATE COUNT(selected\_list);

So, All the 544 players taking part will be alphabetically arranged and two teams of 5 Pokémons need to be extracted out randomly from the earlier list.

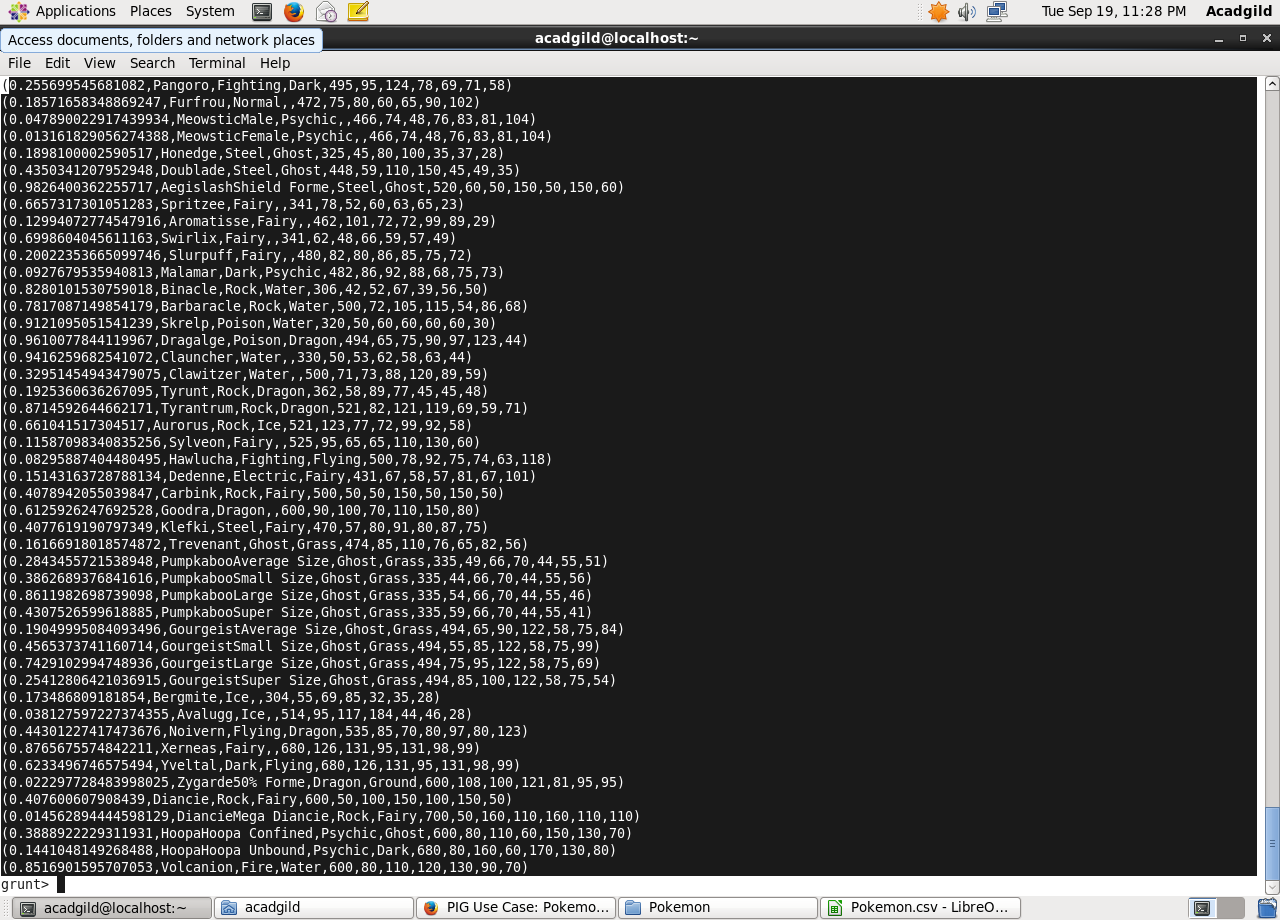
Seems like,this way we will have 2 lists containing 5 Pokémon each so to fight each other.



**Ques 3: Using random() generate random numbers for each Pokémon on the selected list.**

Command

1. random\_include1 = foreach selected\_list GENERATE RANDOM(),Name,Type1,Type2,Total,HP,Attack,Defense,SpAtk,SpDef,Speed;
2. dump;



**Ques 4: Arrange the new list in a descending order according to a column randomly.**

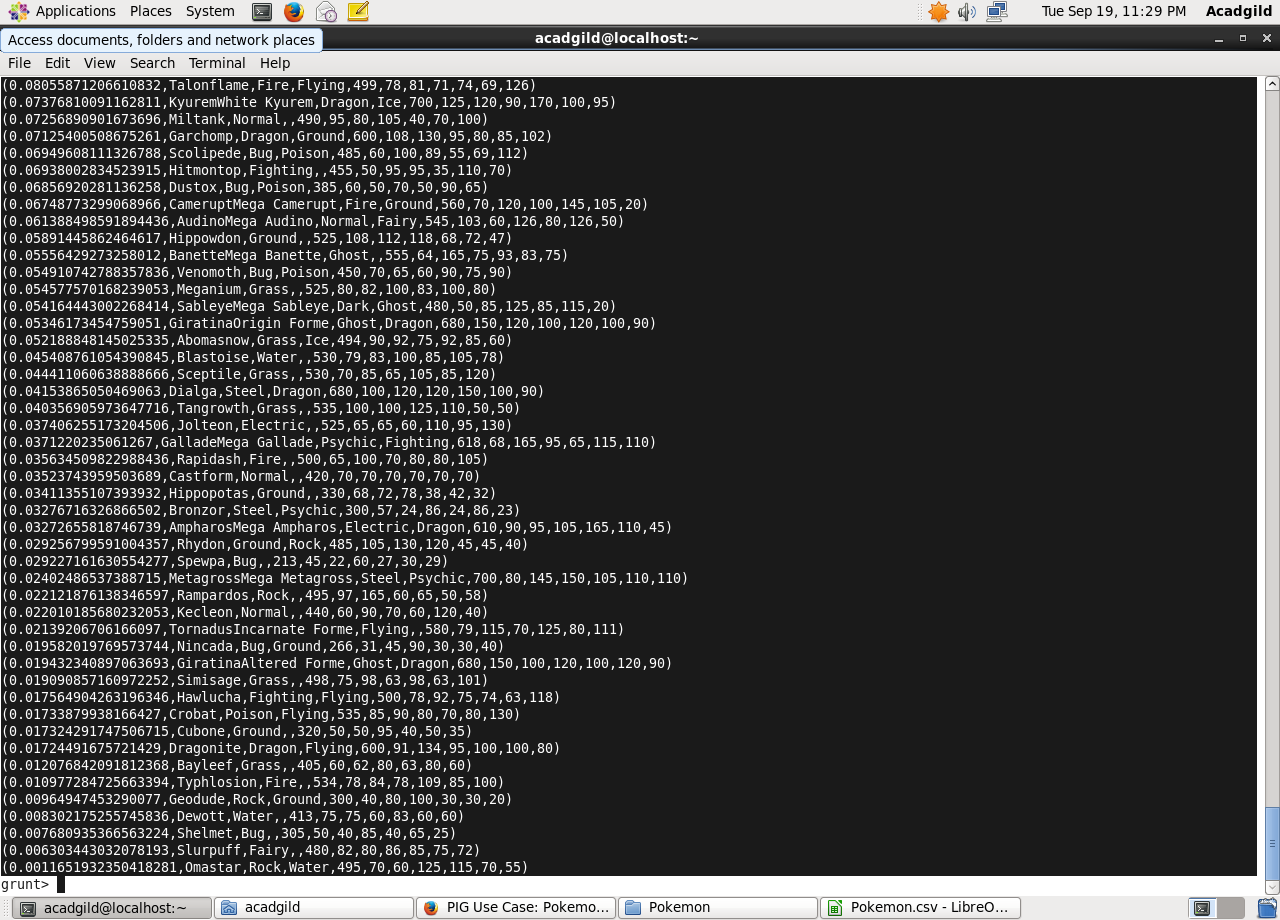
Explanation: This will give us consequently a layer arranged to pick the random list which 1st player will choose.

Command

1. random1\_desending = ORDER random\_include1 BY $0 DESC;

2. dump;

Yet we want 1 more list with random arrangements of Pokémons which will be therefore chosen by the 2nd player later on.

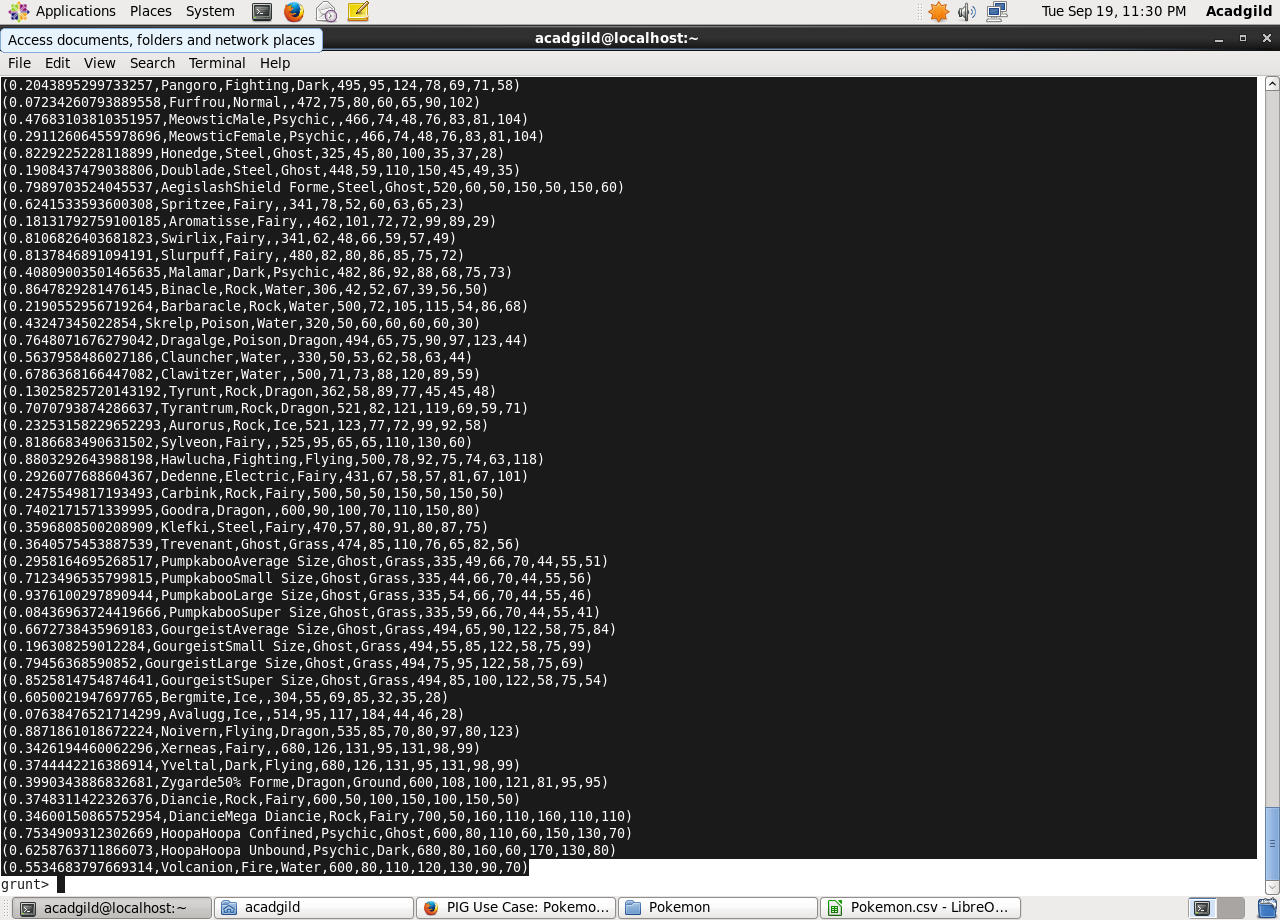


**Ques 5: Now on a new relation again associate random numbers for each Pokémon and arrange in descending order according to column random.**

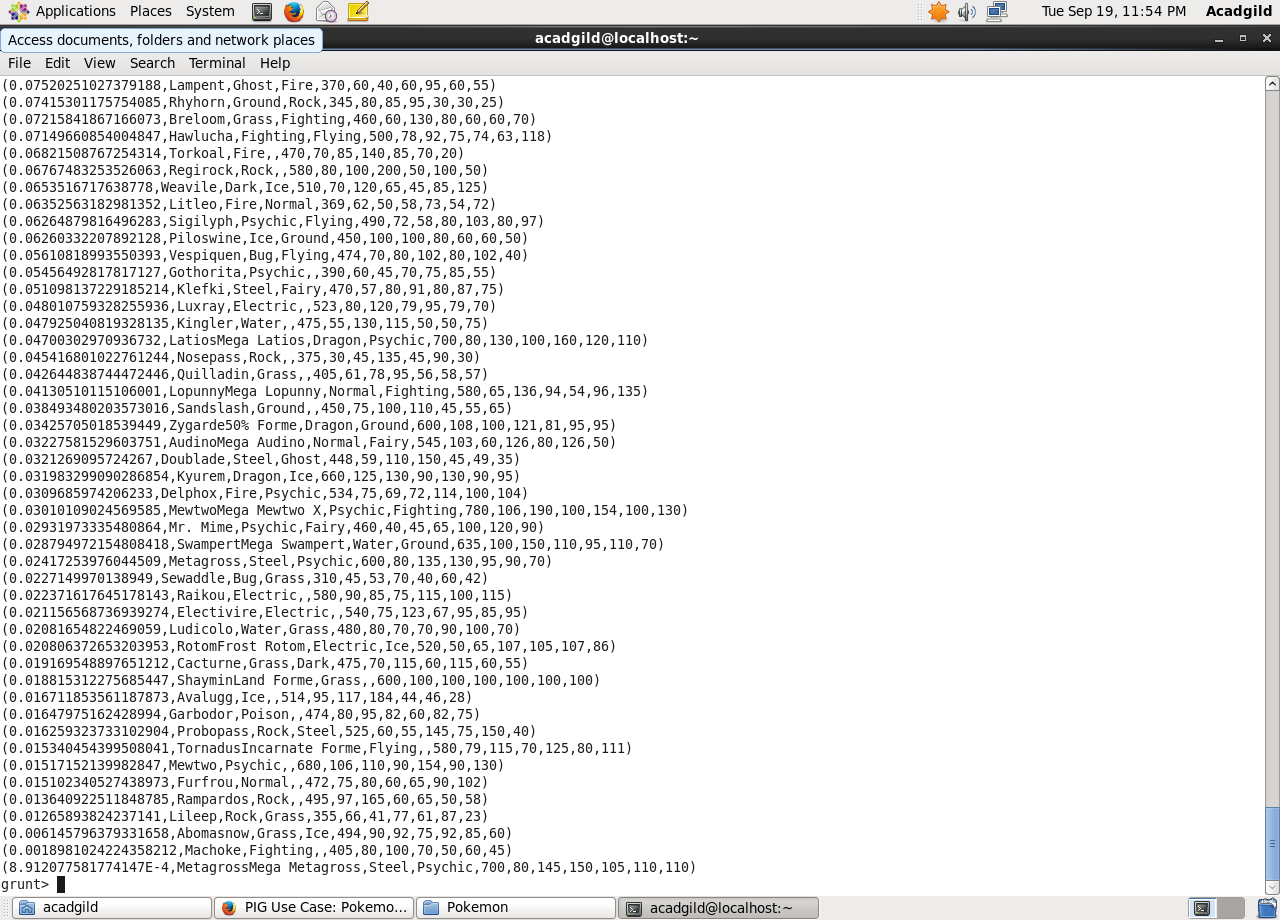
Explanation: We will be repeating above two steps again to form the 2nd list.

Command

1. random\_include2 = foreach selected\_list GENERATE RANDOM(),Name,Type1,Type2,Total,HP,Attack,Defense,SpAtk,SpDef,Speed;
2. Dump:



1. random2\_desending = ORDER random\_include2 BY $0 DESC;
2. dump;

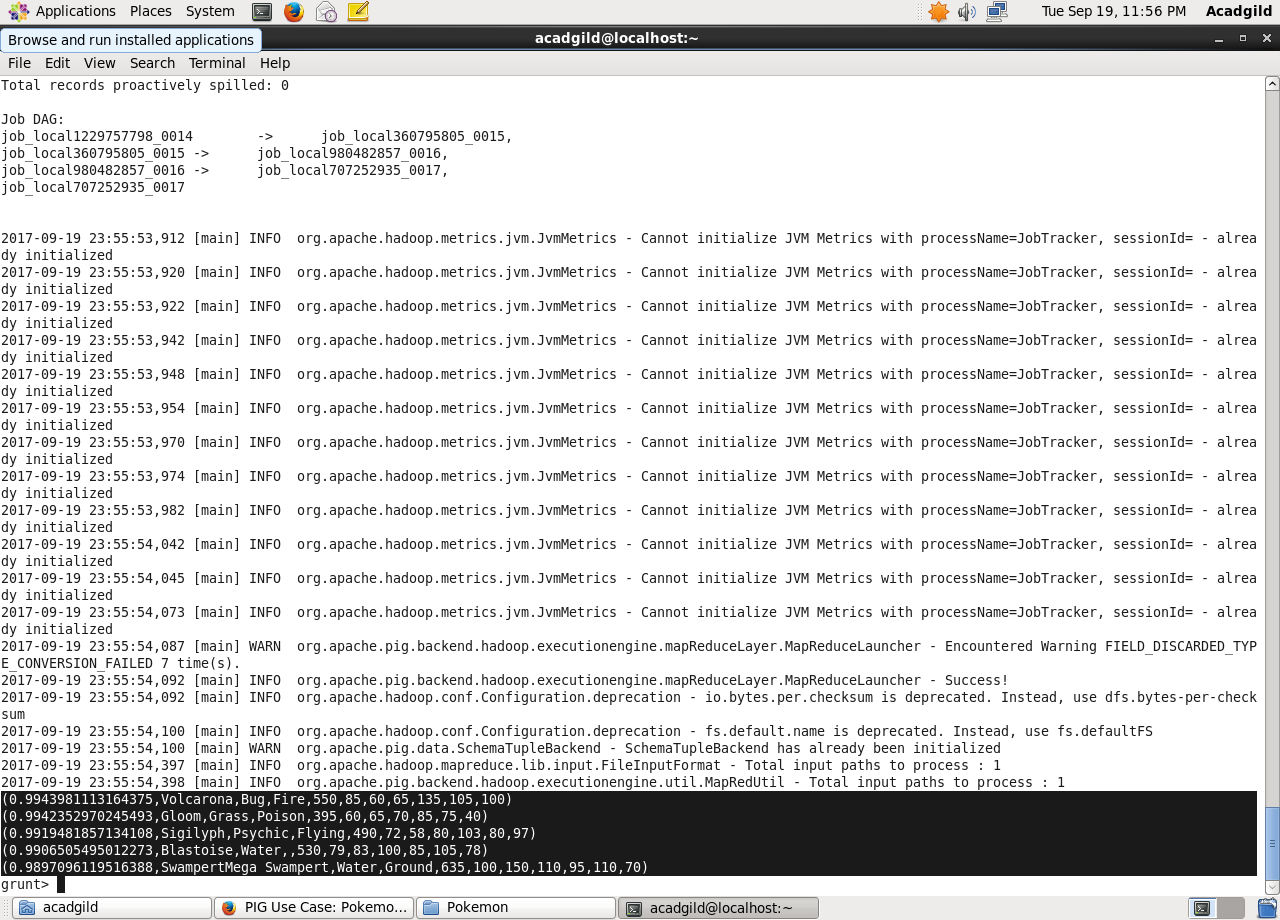


**Ques: From the two different descending lists of random Pokémons, select the top 5 Pokémons for 2 different players.**

Commands

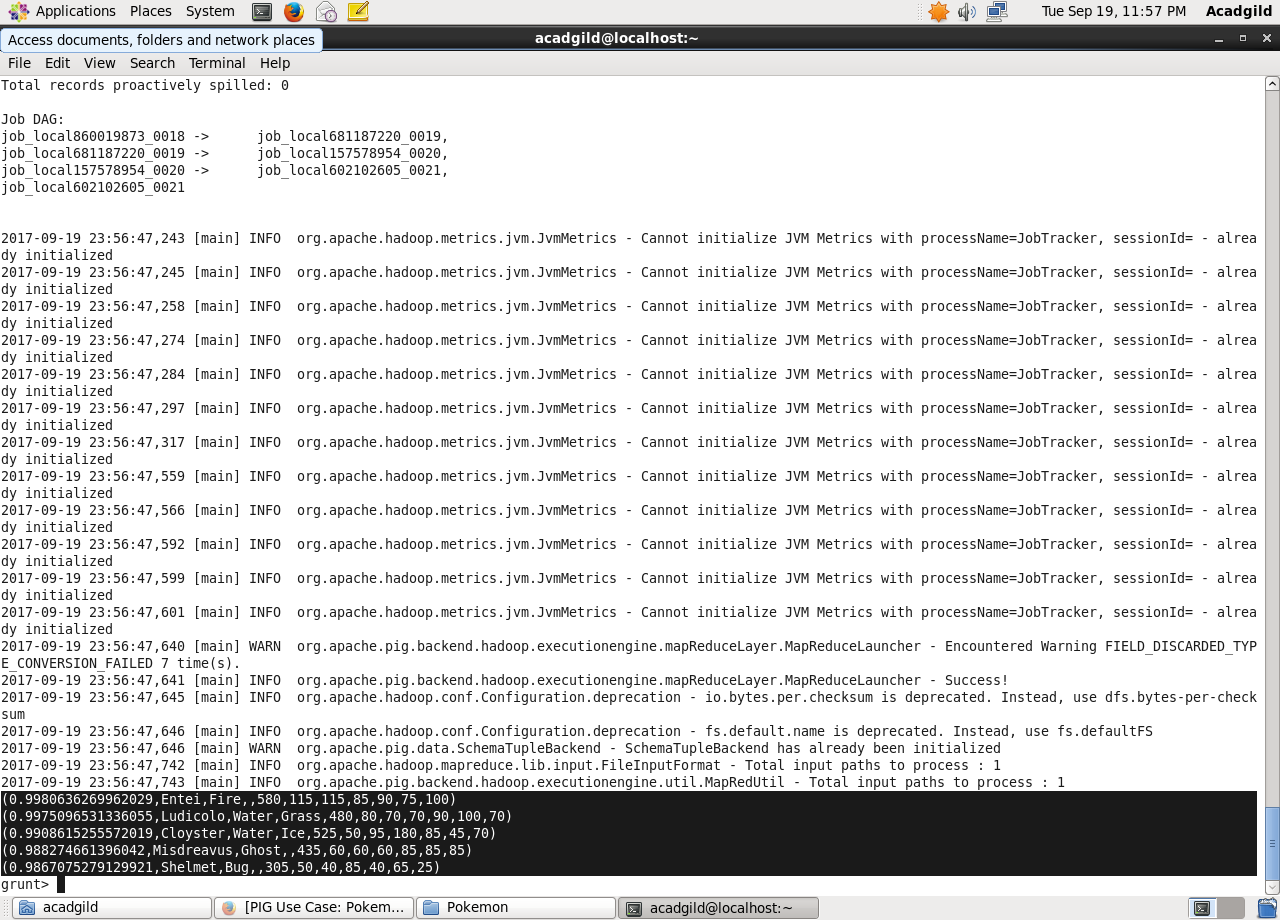
1. limit\_data\_random1\_desending = LIMIT random1\_desending 5 ;

Dump;



2. limit\_data\_random2\_desending = LIMIT random2\_desending 5 ;

Dump;

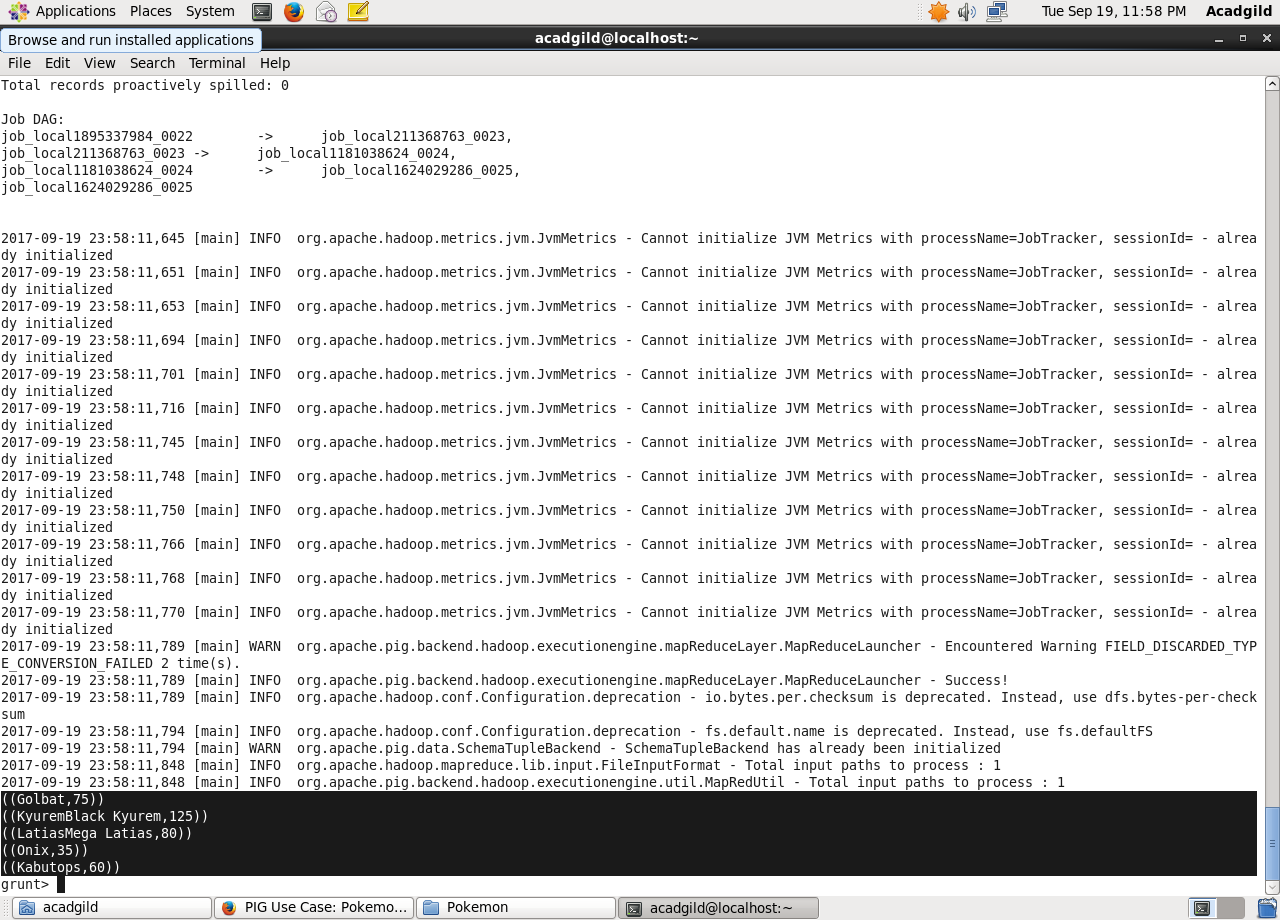


**Ques: Store the data on a local drive to announce for the final match. By the name player1 and player2 (only show the NAME and HP).**

Commands

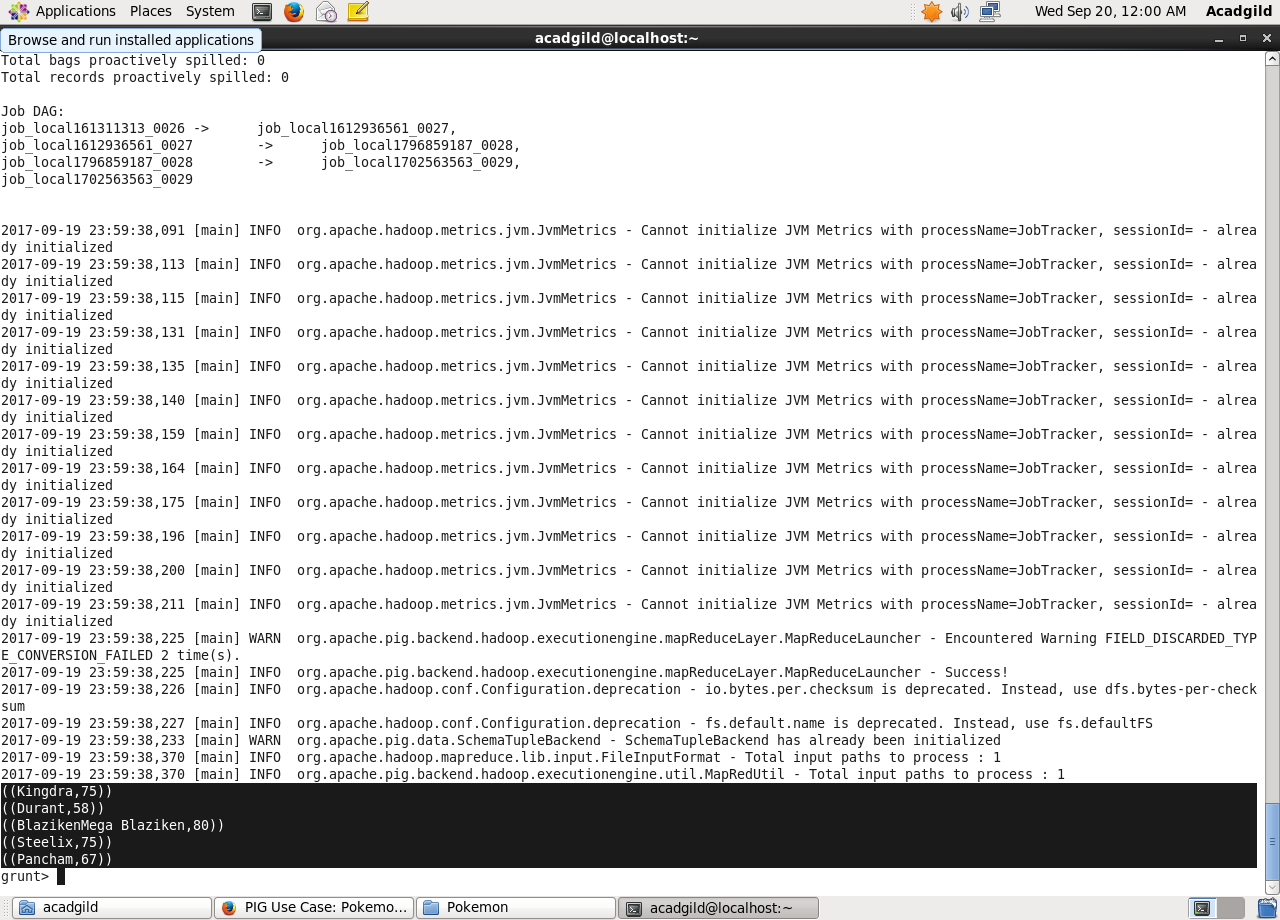
1. filter\_only\_name1 = foreach limit\_data\_random1\_desending Generate ($1,HP);

Dump;



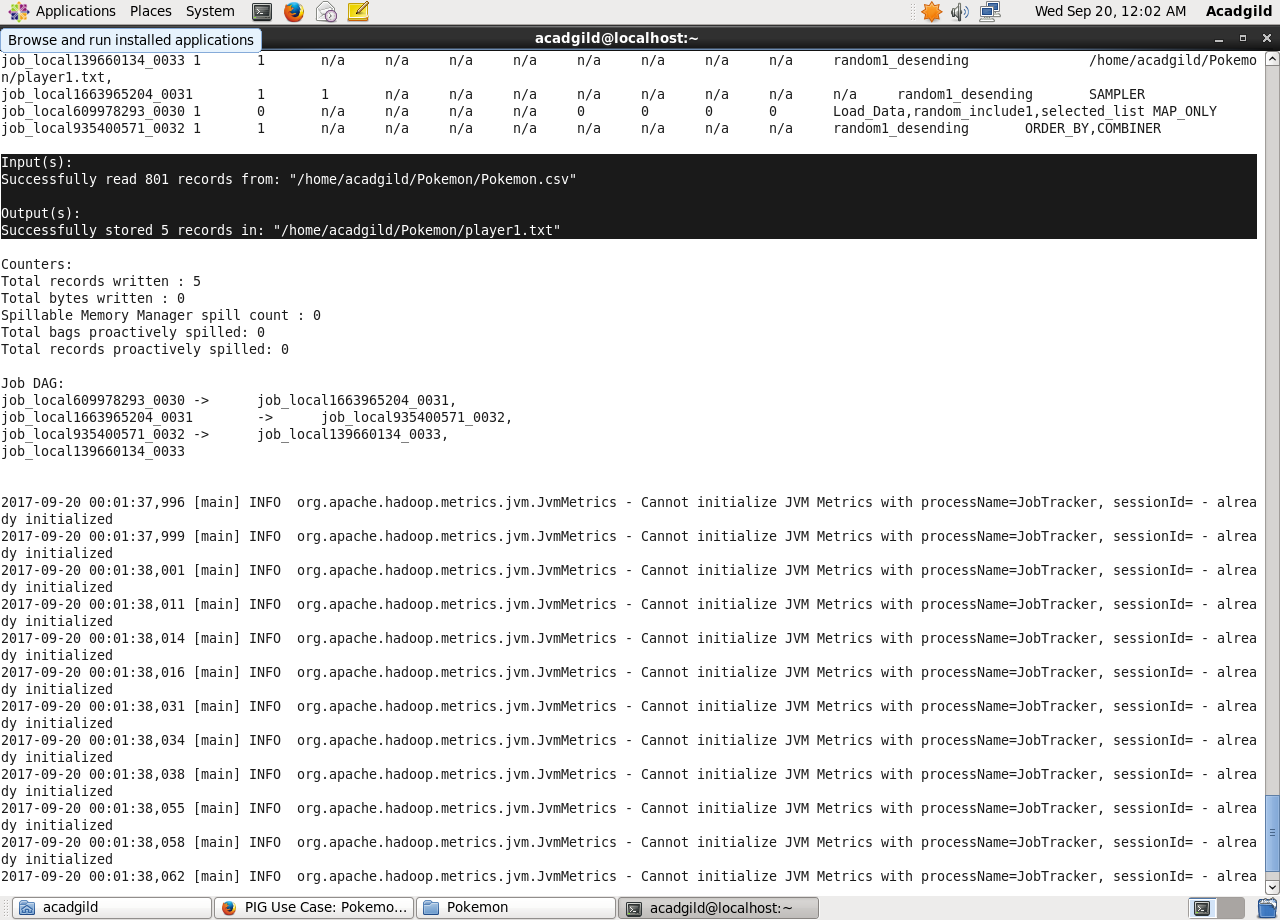
2. filter\_only\_name2 = foreach limit\_data\_random2\_desending Generate ($1,HP);

Dump;

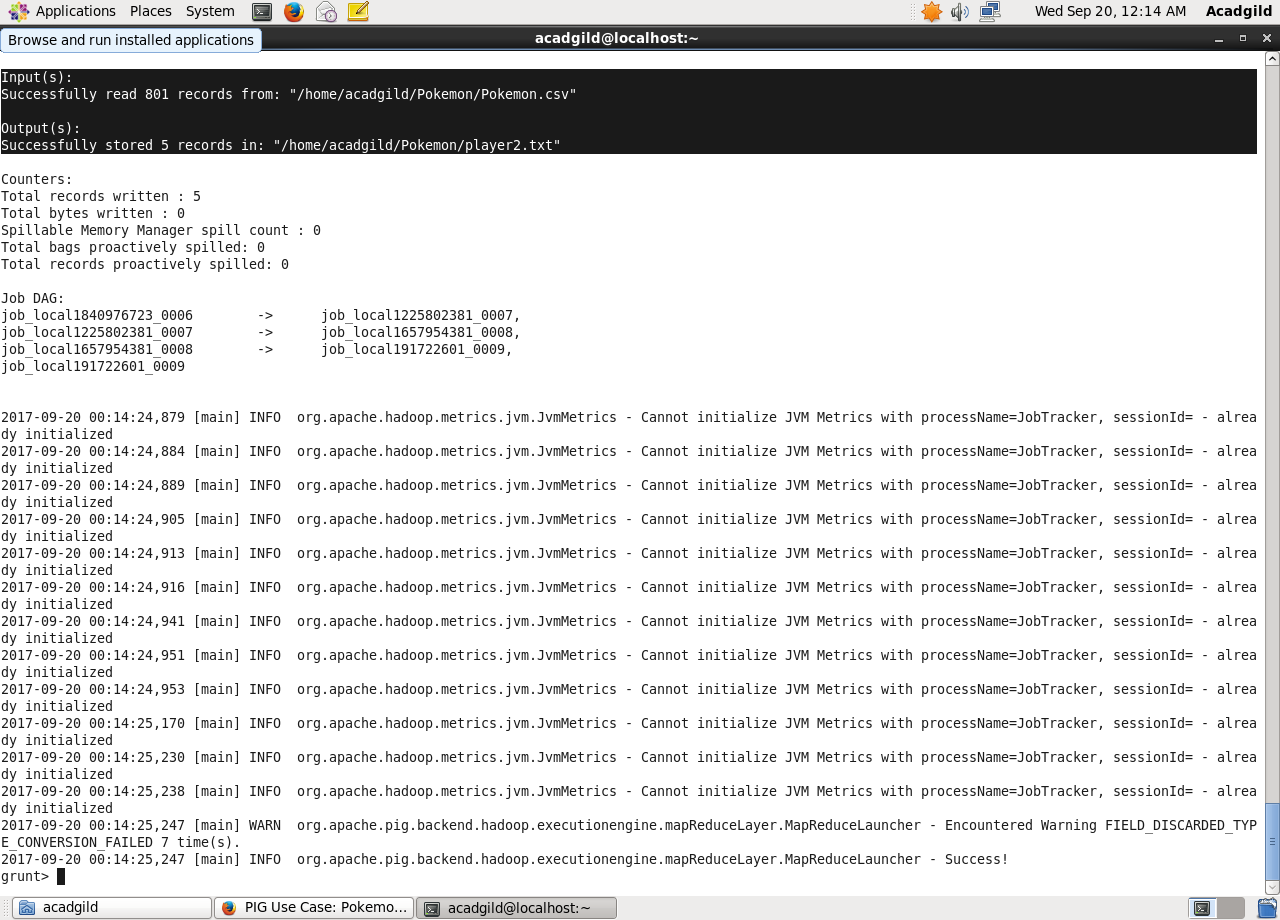


In conclusion, let’s store this result in our local system  using Store

1. STORE limit\_data\_random1\_desending INTO ‘/home/acadgild/Pokemon/player1.txt’;



1. STORE limit\_data\_random2\_desending INTO ‘/home/acadgild/Pokemon/player2.txt’;



As a result, the Pokémons for both players got selected . This player will be fighting consequently in the Finals with their respective Pokémons assigned.