Use the simple Pig built-in functions to form two lists of 5 randomly selected Pokémons with some parameters set by the management.

Load the data:

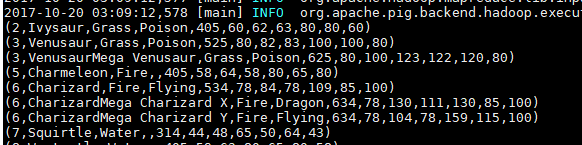
Load\_Data = LOAD 'Pokemon.csv' USING PigStorage(',') AS(Sno:int,Name:chararray,Type1:chararray,Type2:chararray,Total:int,HP:int,Attack:int,Defense:int,SpAtk:int,SpDef:int,Speed:int);

Dump Load\_Data;

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

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

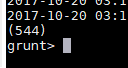
The dataset is filtered, and hence out of all the 800 Pokémons, only 544 are eligible to take part in the tournament. In order to get the count, refer the next problem statement.



1. State the number of players taking part in the competition after getting selected in the qualifying round.

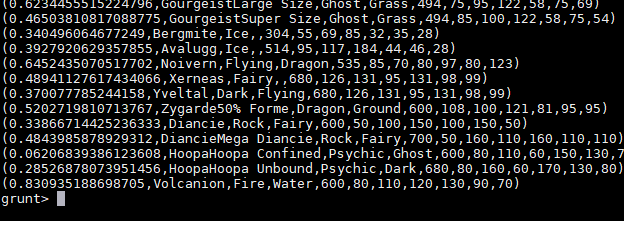
gourp\_selcted\_list = Group selected\_list All;

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



1. Using random() generate random numbers for each Pokémon on the selected list.

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



1. Arrange the new list in a descending order according to a column randomly.

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

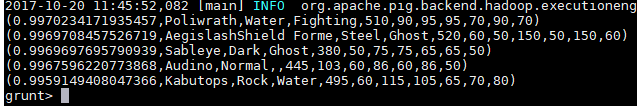
1. Now on a new relation again associate random numbers for each Pokémon and arrange in descending order according to column random.

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

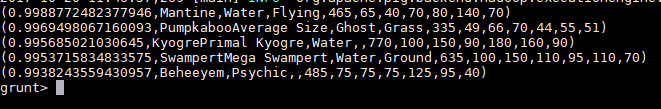
random2\_desending = ORDER random\_include2 BY $0 DESC;

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

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

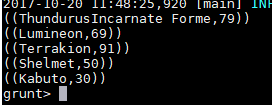


limit\_data\_random2\_desending = LIMIT random2\_desending 5 ;

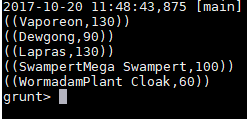


1. Store the data on a local drive to announce for the final match. By the name player1 and player2 (only show the NAME and HP).

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



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



STORE limit\_data\_random1\_desending INTO ‘player1.txt’;

STORE limit\_data\_random2\_desending INTO ‘player2.txt’;

