Assignment 5.3

Problem Statement 1:-

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

Solution:-

Pokemon_Data = LOAD '/home/acadgild/hadoop/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);

DUMP Pokemon_Data;

```
grunt> Pokemon_Data = LOAD '/home/acadgild/hadoop/Pokemon.csv' USING PigStorage(',') AS(Sno:int,Name:chararray,Type1:chararray,Type2:chararray,Total:int,H
P:int,Attack:int,Defense:int,SpAtk:int,SpDef:int,Speed:int);
2017-11-06 22:54:16,126 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-ch
ecksum
2017-11-06 22:54:16,126 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated. Instead, use fs.defaultFS
grunt> dump Pokemon_Data;
```

list_Defense_more_55 = FILTER Pokemon_Data BY Defense>55;

dump list_Defense_more_55;

```
grunt> list_Defense_more_55 = FILTER Pokemon_Data BY Defense>55;
grunt> dump list_Defense_more_55;
```

Problem Statement 2:-

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

Solution:-

group_Defense = Group list_Defense_more_55 All;

count_Defense = foreach group_Defense GENERATE COUNT(list_Defense_more_55);

dump count_Defense;

```
grunt> group_Defense = Group list_Defense_more_55 All;
grunt> count_Defense = foreach group_Defense GENERATE COUNT(list_Defense_more_55);
grunt> dump count_Defense;
```

```
2017-11-06 23:08:01,70
(544)
grunt>
```

Problem Statement 3:-

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

Solution:-

random_list = foreach list_Defense_more_55 GENERATE RANDOM(),Name,Type1,Type2,Total,HP,Attack,Defense,SpAtk,SpDef,Speed;

dump random_list;

```
(544)
grunt> random_list = foreach list_Defense_more_55 GENERATE RANDOM(),Name,Type1,Type2,Total,HP,Attack,Defense,SpAtk,SpDef,Speed;
grunt> dump random_list;
```

Problem Statement 4:-

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

Solution:-

random_list_desending = ORDER random_list BY $0 DESC;

dump random_list_desending;

```
grunt>
grunt> random_list_desending = ORDER random_list BY $0 DESC;
grunt> dump random_list_desending;
```

Problem Statement 5:-

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

Solution:-

random_list2 = foreach list_Defense_more_55 GENERATE RANDOM(),Name,Type1,Type2,Total,HP,Attack,Defense,SpAtk,SpDef,Speed;

random_list2_desending = ORDER random_list2 BY $0 DESC;

dump random_list2_desending;

```
grunt>
grunt> random_list2 = foreach list_Defense_more_55 GENERATE RANDOM(),Name,Type1,Type2,Total,HP,Attack,Defense,SpAtk,SpDef,Speed;
grunt> random_list2_desending = ORDER random_list2 BY $0 DESC;
grunt> dump random_list2_desending;
```

Problem Statement 6:-

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

Solution:-

limit_data_random_list_desending = LIMIT random_list_desending 5 ;

dump limit_data_random_list_desending;

```
grunt>
grunt> limit_data_random_list_desending = LIMIT random_list_desending 5 ;
grunt> dump limit_data_random_list_desending;
```

```
(0.9998532122056683,Flareon,Fire,,525,65,130,60,95,110,65)
(0.9991946252459442,Chespin,Grass,,313,56,61,65,48,45,38)
(0.9984968166164956,Yveltal,Dark,Flying,680,126,131,95,131,98,99)
(0.9959792930079532,Seviper,Poison,,458,73,100,60,100,60,65)
(0.9958855390708033,Fearow,Normal,Flying,442,65,90,65,61,61,100)
grunt>
```

limit_data_random_list2_desending = LIMIT random_list2_desending 5 ;

dump limit_data_random_list2_desending;

```
grunt>
grunt> limit_data_random_list2_desending = LIMIT random_list2_desending 5 ;
grunt> dump limit_data_random_list2_desending;
```

```
2017-11-06 23:26:07,859 [main] INFO  org.apache.pig.backend.hadoop.e
(0.9995404469572298,Ninetales,Fire,,505,73,76,75,81,100,100)
(0.9940956585427244,Nidoking,Poison,Ground,505,81,102,77,85,75,85)
(0.992922483610844,Aerodactyl,Rock,Flying,515,80,105,65,60,75,130)
(0.9902804960937007,Wartortle,Water,,405,59,63,80,65,80,58)
(0.9902171306332544,Swalot,Poison,,467,100,73,83,73,83,55)
grunt>
```

Problem Statement 6:-

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).

Solution:-

filter_only_name1 = foreach limit_data_random_list_desending Generate ($1,HP);

dump filter_only_name1;

```
grunt>
grunt> filter_only_name1 = foreach limit_data_random_list_desending Generate ($1,HP);
grunt> dump filter_only_name1;
```

```
2017-11-06 23:34:07,94
2017-11-06 23:34:07,94
((Mr. Mime,40))
((Gigalith,85))
((Kingler,55))
((Sceptile,70))
((Wartortle,59))
grunt>
```

filter_only_name2 = foreach limit_data_random_list2_desending Generate ($1,HP);

dump filter_only_name2;

```
grunt>
grunt> filter_only_name2 = foreach limit_data_random_list2_desending Generate ($1,HP);
grunt> dump filter_only_name2;
```

```
2017-11-06 23:35:14,576 [m
((Persian,65))
((Ninetales,73))
((Slakoth,60))
((Roserade,60))
((Tyranitar,100))
grunt>
```

STORE limit_data_random_list_desending INTO
‘/home/acadgild/hadoop/Pokemon_player1.txt’;

```
grunt>
grunt> STORE limit_data_random_list_desending INTO '/home/acadgild/hadoop/Pokemon_player1.txt';
2017-11-06 23:36:38,191 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation - io.bytes
ecksum
2017-11-06 23:36:38,191 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation - fs.defau
```

```
-rw-rw-r--. 1 acadgild acadgild 247963212 Nov  5 22:34 DelayedFlights.csv
-rw-rw-r--. 1 acadgild acadgild     37792 Nov  6 22:39 Pokemon.csv
drwxrwxr-x. 2 acadgild acadgild      4096 Nov  6 23:36 Pokemon_player1.txt
[acadgild@localhost hadoop]$ cat Pokemon_player1.txt/
part-r-00000       .part-r-00000.crc  _SUCCESS           ._SUCCESS.crc
[acadgild@localhost hadoop]$ ls -lrt Pokemon_player1.txt/
total 4
-rw-r--r--. 1 acadgild acadgild 300 Nov  6 23:36 part-r-00000
-rw-r--r--. 1 acadgild acadgild   0 Nov  6 23:36 _SUCCESS
[acadgild@localhost hadoop]$
```

```
[acadgild@localhost hadoop]$ cat Pokemon_player1.txt/part-r-00000
0.9995808638163947      Pupitar Rock    Ground  410     70      84      70      65      70      51
0.9994868094492321      Shellder        Water           305     30      65      100     45      25      40
0.9954408925356402      Nincada Bug     Ground  266     31      45      90      30      30      40
0.9952287803304086      Crobat  Poison  Flying  535     85      90      80      70      80      130
0.9940342541125476      Vigoroth        Normal          440     80      80      80      55      55      90
[acadgild@localhost hadoop]$
```

STORE limit_data_random_list2_desending INTO
‘/home/acadgild/hadoop/Pokemon_player2.txt’;

```
grunt>
grunt> STORE limit_data_random_list2_desending INTO '/home/acadgild/hadoop/Pokemon_player2.txt';
```

```
[acadgild@localhost hadoop]$ ls -lrt
total 245640
drwxr-xr-x. 3 acadgild acadgild      4096 Aug 12  2016 namenode
drwx------. 3 acadgild acadgild      4096 Aug 12  2016 datanode
-rw-rw-r--. 1 acadgild acadgild     21007 Oct  9 01:16 sample_temperature_dataset.csv
-rw-rw-r--. 1 acadgild acadgild      5445 Oct  9 23:45 mapreduce-0.0.1-SNAPSHOT.jar
-rw-rw-r--. 1 acadgild acadgild       210 Oct 11 22:24 max-temp.txt
-rw-rw-r--. 1 acadgild acadgild   3194099 Oct 14 13:21 NYSE_daily
drwxrwxr-x. 2 acadgild acadgild      4096 Oct 15 01:15 maxout
-rw-rw-r--. 1 acadgild acadgild      5338 Oct 15 01:22 pig_1508008101045.log
-rw-rw-r--. 1 acadgild acadgild        48 Oct 21 10:59 dept_data.csv
-rw-rw-r--. 1 acadgild acadgild       375 Nov  1 00:44 student_details.txt
drwxrwxr-x. 2 acadgild acadgild      4096 Nov  1 02:09 student_gpa_sumout
-rw-rw-r--. 1 acadgild acadgild       273 Nov  5 11:10 employee_details.txt
-rw-rw-r--. 1 acadgild acadgild        79 Nov  5 11:10 employee_expenses.txt
-rw-rw-r--. 1 acadgild acadgild      1367 Nov  5 11:43 pig_1509861709152.log
-rw-rw-r--. 1 acadgild acadgild    244438 Nov  5 22:33 airports.csv
-rw-rw-r--. 1 acadgild acadgild 247963212 Nov  5 22:34 DelayedFlights.csv
-rw-rw-r--. 1 acadgild acadgild     37792 Nov  6 22:39 Pokemon.csv
drwxrwxr-x. 2 acadgild acadgild      4096 Nov  6 23:36 Pokemon_player1.txt
drwxrwxr-x. 2 acadgild acadgild      4096 Nov  6 23:40 Pokemon_player2.txt
[acadgild@localhost hadoop]$ ls -lrt Pokemon_player2.txt/
total 4
-rw-r--r--. 1 acadgild acadgild 293 Nov  6 23:40 part-r-00000
-rw-r--r--. 1 acadgild acadgild   0 Nov  6 23:40 _SUCCESS
[acadgild@localhost hadoop]$ cat Pokemon_player2.txt/part-r-00000
0.9989392616764946      Garbodor        Poison          474     80      95      82      60      82      75
0.9987629973565375      Venipede        Bug     Poison  260     30      45      59      30      39      57
0.9984482719139836      Regice  Ice             580     80      50      100     100     200     50
0.9983744685897152      Seviper Poison          458     73      100     60      100     60      65
0.9969259197902556      Cinccino        Normal          470     75      95      60      65      60      115
[acadgild@localhost hadoop]$
```

Submitted By:-

Hardik Kaushik