Session 5

Assignment 3

|  |  |
| --- | --- |
| **Prepared For:** | AcadGild |
|  |  |
| **Document Approval:** | **AcadGild** |
|  |  |
|  |  |
|  |  |
|  |  |
| **Project Title:** | Session 5 – Assignment 3Pokemon |
|  |  |
| **Prepared By:** | Duncan Burgess |
|  |  |
|  | dburgess@duncb.com |
|  |  |
| **Primary Engineer:** | Duncan Burgess |
|  |  |
| **Document Reference:** | **Session 5 – Assignment 3** |
|  |  |
| **Start Date:** | 15/09/2017 |
|  |  |
|  |  |



# 

# Contents

[Contents 2](#_Toc493471435)

[Change History 3](#_Toc493471436)

[1. Problem Statement 4](#_Toc493471437)

[2. Setting up the PIG environment 5](#_Toc493471438)

[3. Solutions 6](#_Toc493471439)

[3.1. Filter Pokemon with defence greater than 55 6](#_Toc493471440)

[3.2. Create Randon numbers for each Pokemon List1. 7](#_Toc493471441)

[3.3. Create Randon numbers for each Pokemon List2. 7](#_Toc493471442)

[3.4. Select 5 descending random Pokemon for 2 Players 7](#_Toc493471443)

[3.5. Player 1 and Player 2 selected Pokemon 8](#_Toc493471444)

[3.6. Storing results to Local Store 8](#_Toc493471445)

# Change History

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Document Revision** | **Date** | **Authored By** | **Authorised By** | **Sections Affected** | **Reason for Change** |
| Rev 01 | 18/09/2017 | Duncan Burgess |  | All | Initial release. |
|  |  |  |  |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |

# Problem Statement

Implement the use case present in below blog link and share the complete steps along with

screenshot(s) from your end.

NOTE: You must submit a word file containing steps and screenshots.

https://acadgild.com/blog/pig-use-case-pokemon-data-analysis/

# Setting up the PIG environment

**One dataset created to implement use case.**

**Pokemon.csv**

#,Name,Type1,Type2,Total,HP,Attack,Defense,Sp.Atk,Sp.Def,Speed  
001,Bulbasaur,Grass,Poison,318,45,49,49,65,65,45  
002,Ivysaur,Grass,Poison,405,60,62,63,80,80,60  
003,Venusaur,Grass,Poison,525,80,82,83,100,100,80  
003,VenusaurMegaVenusaur,Grass,Poison,625,80,100,123,122,120,80  
004,Charmander,Fire,,309,39,52,43,60,50,65  
005,Charmeleon,Fire,,405,58,64,58,80,65,80  
006,Charizard,Fire,Flying,534,78,84,78,109,85,100  
006,CharizardMegaCharizardX,Fire,Dragon,634,78,130,111,130,85,100  
006,CharizardMegaCharizardY,Fire,Flying,634,78,104,78,159,115,100  
007,Squirtle,Water,,314,44,48,65,50,64,43  
008,Wartortle,Water,,405,59,63,80,65,80,58

**Copy files to hdfs**

$ hadoop fs -copyFromLocal pokemon.csv

**Start Pig**

pig -x local - *(local mode)*

2017-09-14 14:32:48,658 INFO [main] pig.ExecTypeProvider: Trying ExecType : LOCAL

2017-09-14 14:32:48,658 INFO [main] pig.ExecTypeProvider: Picked LOCAL as the ExecType

you will be presents with the pig shell: grunt>

We need to load the files now: ***Note PigStorage is cap sensitive.***

# Solutions

The following demonstrations show the use of Pig commands in Pokemon use case.

grunt> *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);*

grunt> *describe load\_Data;*

**Result of dump**

grunt> dump load\_Data;

(,Name,Type 1,Type 2,,,,,,,)

(1,Bulbasaur,Grass,Poison,318,45,49,49,65,65,45)

(2,Ivysaur,Grass,Poison,405,60,62,63,80,80,60)

(3,Venusaur,Grass,Poison,525,80,82,83,100,100,80)

(3,VenusaurMega Venusaur,Grass,Poison,625,80,100,123,122,120,80)

(4,Charmander,Fire,,309,39,52,43,60,50,65)

(5,Charmeleon,Fire,,405,58,64,58,80,65,80)

(6,Charizard,Fire,Flying,534,78,84,78,109,85,100)

(6,CharizardMega Charizard X,Fire,Dragon,634,78,130,111,130,85,100)

(6,CharizardMega Charizard Y,Fire,Flying,634,78,104,78,159,115,100)

(7,Squirtle,Water,,314,44,48,65,50,64,43)

(8,Wartortle,Water,,405,59,63,80,65,80,58)

## Filter Pokemon with defence greater than 55

grunt> *selected\_list = FILTER load\_Data BY Defense>55;*

grunt> *dump selected\_list;*

(2,Ivysaur,Grass,Poison,405,60,62,63,80,80,60)

(3,Venusaur,Grass,Poison,525,80,82,83,100,100,80)

(3,VenusaurMega Venusaur,Grass,Poison,625,80,100,123,122,120,80)

(5,Charmeleon,Fire,,405,58,64,58,80,65,80)

(6,Charizard,Fire,Flying,534,78,84,78,109,85,100)

(6,CharizardMega Charizard X,Fire,Dragon,634,78,130,111,130,85,100)

....

(719,Diancie,Rock,Fairy,600,50,100,150,100,150,50)

(719,DiancieMega Diancie,Rock,Fairy,700,50,160,110,160,110,110)

(720,HoopaHoopa Confined,Psychic,Ghost,600,80,110,60,150,130,70)

(720,HoopaHoopa Unbound,Psychic,Dark,680,80,160,60,170,130,80)

(721,Volcanion,Fire,Water,600,80,110,120,130,90,70)

Find the eligible players and get a count.

*grunt> group\_selcted\_list = Group selected\_list All;*

*grunt> count\_selcted\_list = foreach group\_selcted\_list GENERATE COUNT(selected\_list);*

*Grunt> dump count\_selcted\_list*

**Result**

(544)

## Create Randon numbers for each Pokemon List1.

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

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

*grunt> dump random1\_desending*

**Result**

(0.9979652685294028,Geodude,Rock,Ground,300,40,80,100,30,30,20)

(0.9978704118468474,Terrakion,Rock,Fighting,580,91,129,90,72,90,108)

(0.9971863401310382,PumpkabooAverage Size,Ghost,Grass,335,49,66,70,44,55,51)

(0.9969268911741044,Turtwig,Grass,,318,55,68,64,45,55,31)

(0.996481639321043,Primeape,Fighting,,455,65,105,60,60,70,95)

(0.992081576633216,Bergmite,Ice,,304,55,69,85,32,35,28)

(0.9915178659195364,Mamoswine,Ice,Ground,530,110,130,80,70,60,80)

(0.9905079897759242,Seaking,Water,,450,80,92,65,65,80,68)

(0.985219089518211,Ho-oh,Fire,Flying,680,106,130,90,110,154,90)

## Create Randon numbers for each Pokemon List2.

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

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

*grunt> dump random2\_descending;*

(0.998572827650494,Banette,Ghost,,455,64,115,65,83,63,65)

(0.9959782052060961,Tauros,Normal,,490,75,100,95,40,70,110)

(0.9943638921933604,KeldeoResolute Forme,Water,Fighting,580,91,72,90,129,90,108)

(0.9926680451482541,Machoke,Fighting,,405,80,100,70,50,60,45)

(0.9911086663617373,Octillery,Water,,480,75,105,75,105,75,45)

(0.9858048895453777,Shellder,Water,,305,30,65,100,45,25,40)

## Select 5 descending random Pokemon for 2 Players

*grunt> limit\_data\_random1\_desending = LIMIT random1\_desending 5;*

*grunt> dump limit\_data\_random1\_desending*

**Results**

(0.9998321751032035,Clamperl,Water,,345,35,64,85,74,55,32)

(0.9980352273848314,Scrafty,Dark,Fighting,488,65,90,115,45,115,58)

(0.9966692074145164,Gengar,Ghost,Poison,500,60,65,60,130,75,110)

(0.9965362460858387,Bisharp,Dark,Steel,490,65,125,100,60,70,70)

(0.9933730260940183,AbomasnowMega Abomasnow,Grass,Ice,594,90,132,105,132,105,30)

*grunt> limit\_data\_random2\_desending = LIMIT random2\_desending 5;*

*grunt> dump limit\_data\_random2\_desending*

**Results**

(0.998343993042376,Emolga,Electric,Flying,428,55,75,60,75,60,103)

(0.998083010717732,Golbat,Poison,Flying,455,75,80,70,65,75,90)

(0.9973035243331635,Rapidash,Fire,,500,65,100,70,80,80,105)

(0.9968065998257111,Hariyama,Fighting,,474,144,120,60,40,60,50)

(0.9964452813504475,Bagon,Dragon,,300,45,75,60,40,30,50)

## Player 1 and Player 2 selected Pokemon

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

*grunt> dump filter\_only\_name1*

**Results**

((Bellossom,75))

((MeowsticMale,74))

((Blaziken,80))

((Gurdurr,85))

((GiratinaOrigin Forme,150))

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

*grunt> dump filter\_only\_name2*

**Results**

((Omanyte,35))

((Mandibuzz,110))

((Sandshrew,50))

((Scolipede,60))

((Uxie,75))

## Storing results to Local Store

*grunt> STORE limit\_data\_random1\_desending INTO '/home/acadgild/Documents/Player\_01.txt';*

**Results**

Input(s):

Successfully read 801 records from: "file:///home/acadgild/Documents/Pokemon.csv"

Output(s):

Successfully stored 5 records in: "/home/acadgild/Documents/Player\_01.txt"

Counters:

Total records written : 5

*grunt> STORE limit\_data\_random2\_desending INTO '/home/acadgild/Documents/Player\_02.txt';*

**Results**

Input(s):

Successfully read 801 records from: "file:///home/acadgild/Documents/Pokemon.csv"

Output(s):

Successfully stored 5 records in: "/home/acadgild/Documents/Player\_02.txt"

Counters:

Total records written : 5