POKEMON ANALYSIS USING APACHE PIG

===============================================================

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.

----------------------------------------------------------------------------------------------------------------

**Description of Input Dataset Pokemon.csv**

#,Name,Type 1,Type 2,Total,HP,Attack,Defense,Sp. Atk,Sp. Def,Speed

----------------------------------------------------------------------------------------------------------------

**Problem Statement 1:**

**Find the list of players that have been selected in the qualifying round.**

Input Pig Commands 1:

-------------------------------

pig -x local

REGISTER '/home/cloudera/Downloads/piggybank-0.17.0.jar';

p1 = LOAD '/home/cloudera/Downloads/Pokemon.csv' USING PigStorage(',')

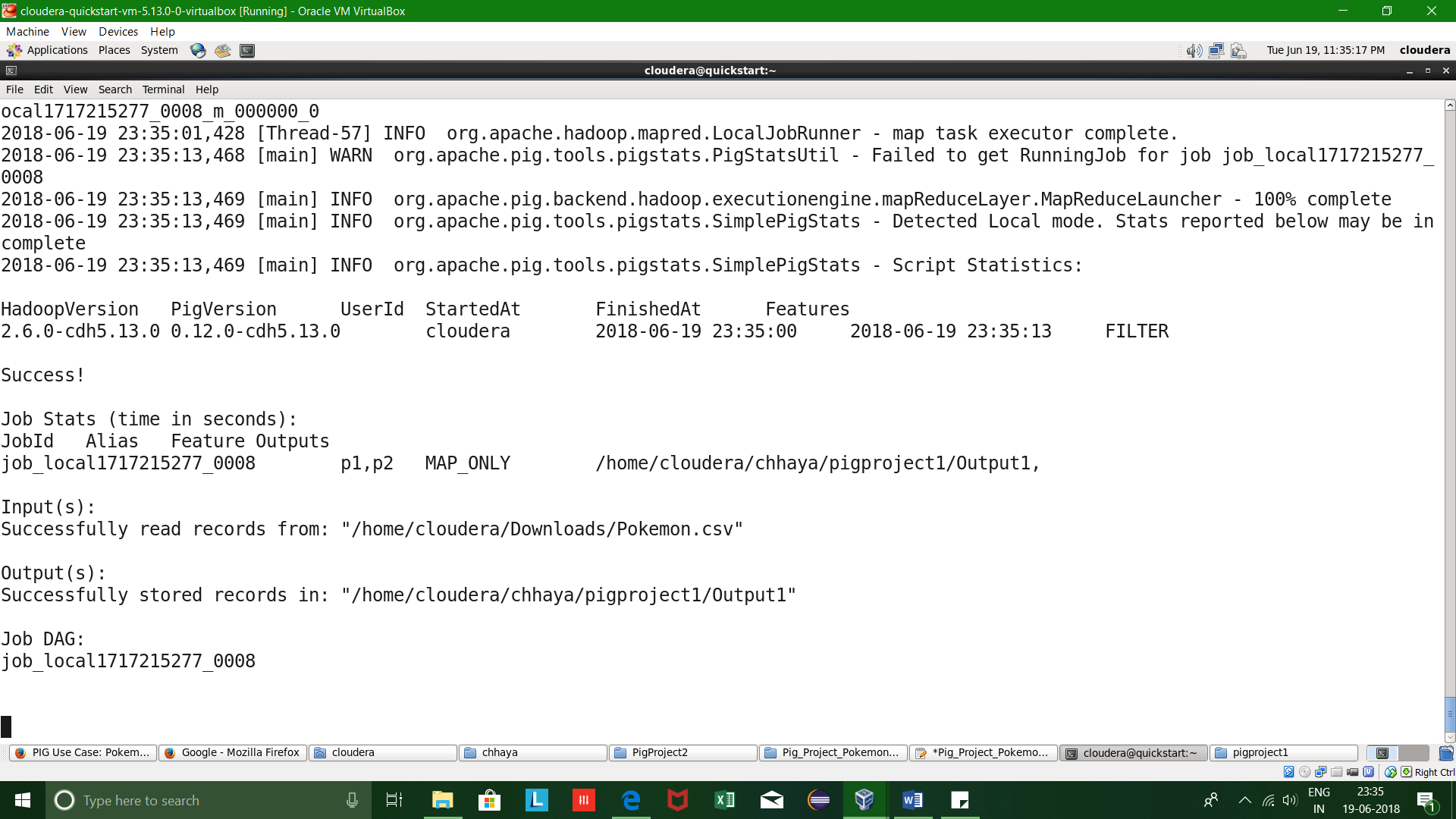
AS (sr:int , pokemonname:chararray , type\_1:chararray , type\_2:chararray, total:int , hp:int, attack:int , defense:int , specialattack:int , specialdefense:int , speed:int);

p2 = FILTER p1 BY ( defense > 55);

DUMP p2;

STORE p2 INTO '/home/cloudera/chhaya/pigproject1/Output1' USING PigStorage('\t');

Output Screenshots 1:



================================================================

Problem Statement 2:

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

Input Pig Commands 2:

----------------------------

REGISTER '/home/cloudera/Downloads/piggybank-0.17.0.jar';

p1 = LOAD '/home/cloudera/Downloads/Pokemon.csv' USING PigStorage(',')

AS (sr:int , pokemonname:chararray , type\_1:chararray , type\_2:chararray, total:int , hp:int, attack:int , defense:int , specialattack:int , specialdefense:int , speed:int);

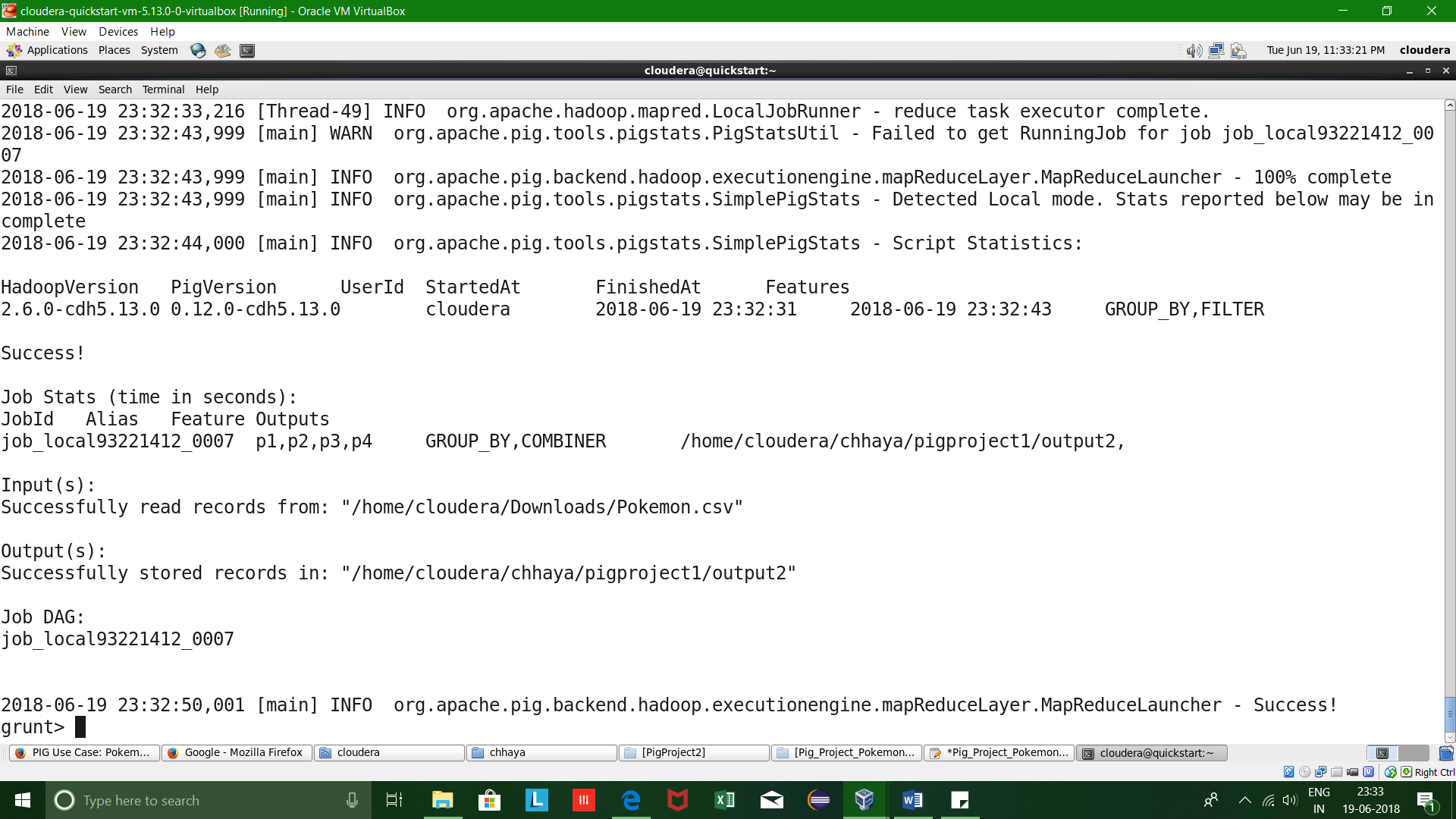
p2 = FILTER p1 BY ( defense > 55);

p3 = GROUP p2 ALL;

p4 = FOREACH p3 generate group, COUNT(p2);

STORE p4 INTO '/home/cloudera/chhaya/pigproject1/output2' USING PigStorage('\t');

Output Screenshots 2:



================================================================

Problem Statement 3:

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

Input Pig Commands 3:

----------------------------

REGISTER '/home/cloudera/Downloads/piggybank-0.17.0.jar';

p1 = LOAD '/home/cloudera/Downloads/Pokemon.csv' USING PigStorage(',')

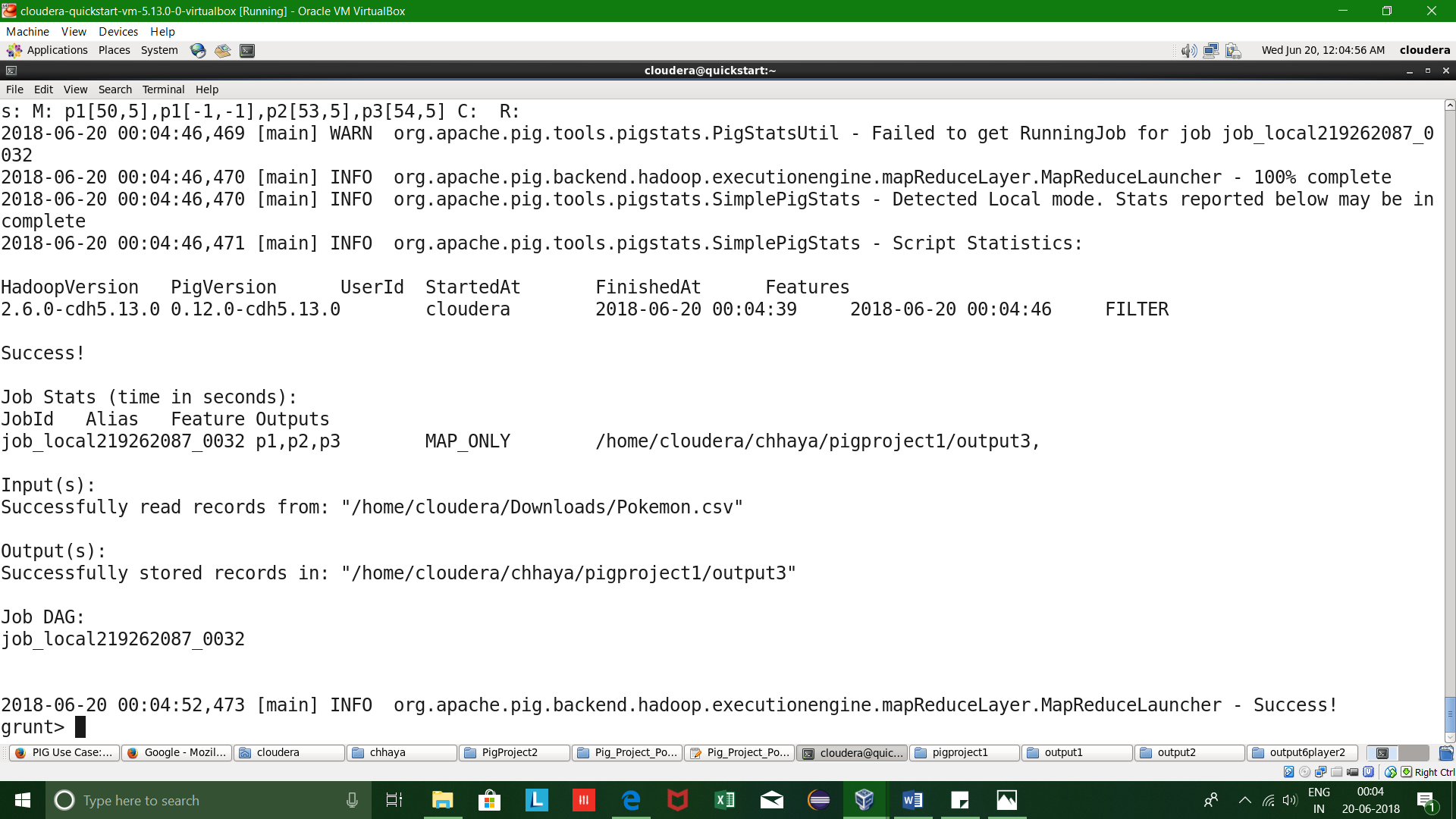
AS (sr:int , pokemonname:chararray , type\_1:chararray , type\_2:chararray, total:int , hp:int, attack:int , defense:int , specialattack:int , specialdefense:int , speed:int);

p2 = FILTER p1 BY ( defense > 55);

p3 = FOREACH p2 GENERATE RANDOM(), pokemonname, type\_1, type\_2 , total, hp, attack, defense, specialattack, specialdefense, speed;

STORE p3 INTO '/home/cloudera/chhaya/pigproject1/output3' USING PigStorage('\t');

Output Screenshots 3:



================================================================

Problem Statement 4:

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

Input Pig Commands 4:

----------------------------

REGISTER '/home/cloudera/Downloads/piggybank-0.17.0.jar';

p1 = LOAD '/home/cloudera/Downloads/Pokemon.csv' USING PigStorage(',')

AS (sr:int , pokemonname:chararray , type\_1:chararray , type\_2:chararray, total:int , hp:int, attack:int , defense:int , specialattack:int , specialdefense:int , speed:int);

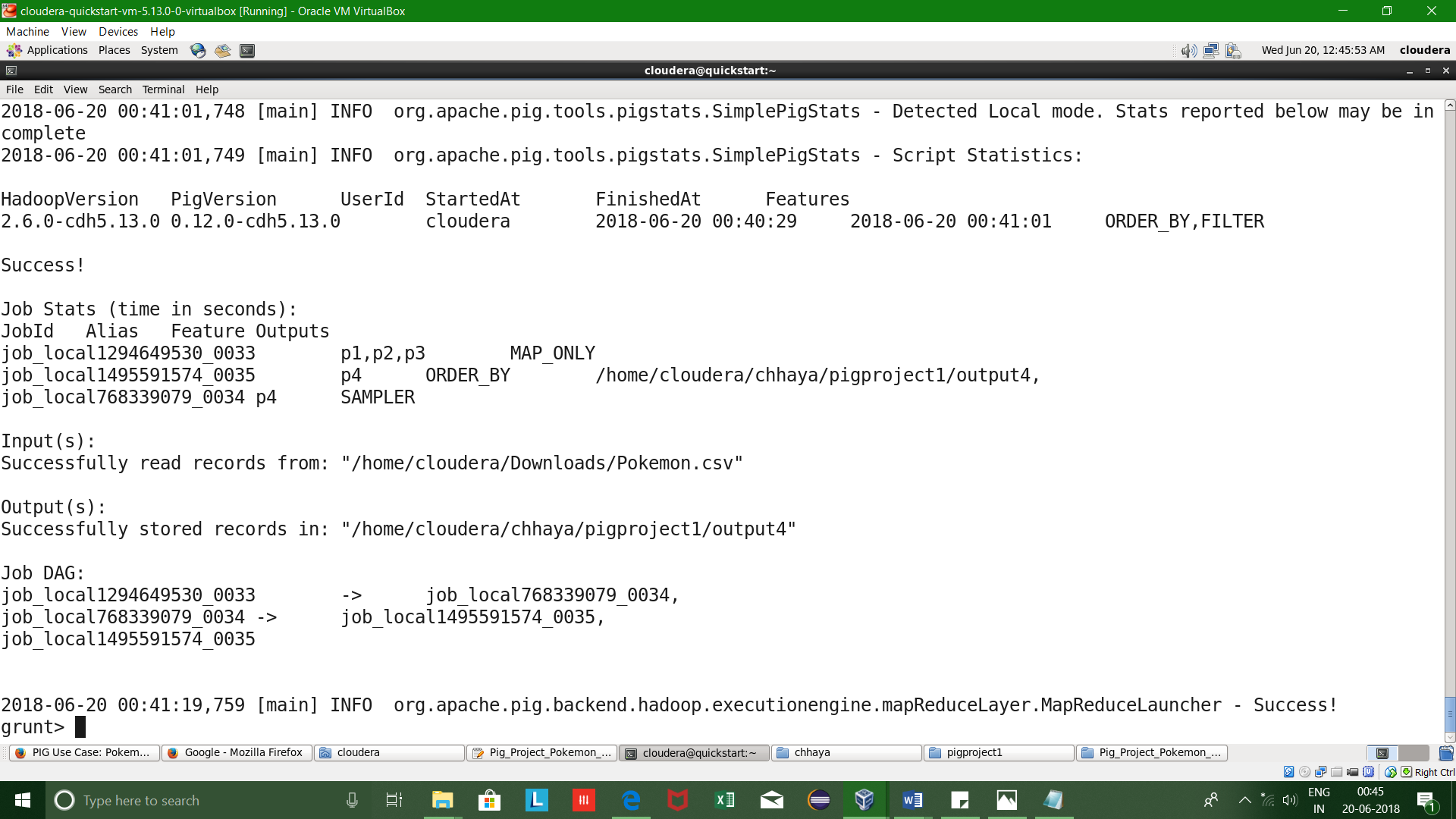
p2 = FILTER p1 BY ( defense > 55);

p3 = FOREACH p2 GENERATE RANDOM() as random, pokemonname, type\_1, type\_2 , total, hp, attack, defense, specialattack, specialdefense, speed;

p4 = ORDER p3 BY random desc;

STORE p4 INTO '/home/cloudera/chhaya/pigproject1/output4' USING PigStorage('\t');

Output Screenshots 4:



================================================================

Problem Statement 5:

**Now on a new relation again associate random numbers for each Pokémon and arrange in descending order according to column random.From the two different descending lists of random Pokémons, select the top 5 Pokémons for 2 different players.**

Input Pig Commands 5:

------------------

REGISTER '/home/cloudera/Downloads/piggybank-0.17.0.jar';

p1 = LOAD '/home/cloudera/Downloads/Pokemon.csv' USING PigStorage(',')

AS (sr:int , pokemonname:chararray , type\_1:chararray , type\_2:chararray, total:int , hp:int, attack:int , defense:int , specialattack:int , specialdefense:int , speed:int);

p2 = FILTER p1 BY ( defense > 55);

random1 = FOREACH p2 GENERATE RANDOM() as random, pokemonname, type\_1, type\_2 , total, hp, attack, defense, specialattack, specialdefense, speed;

random1\_desc = ORDER random1 by random desc;

random2 = FOREACH p2 GENERATE RANDOM() as random, pokemonname, type\_1, type\_2 , total, hp, attack, defense, specialattack, specialdefense, speed;

random2\_desc = ORDER random2 by random desc;

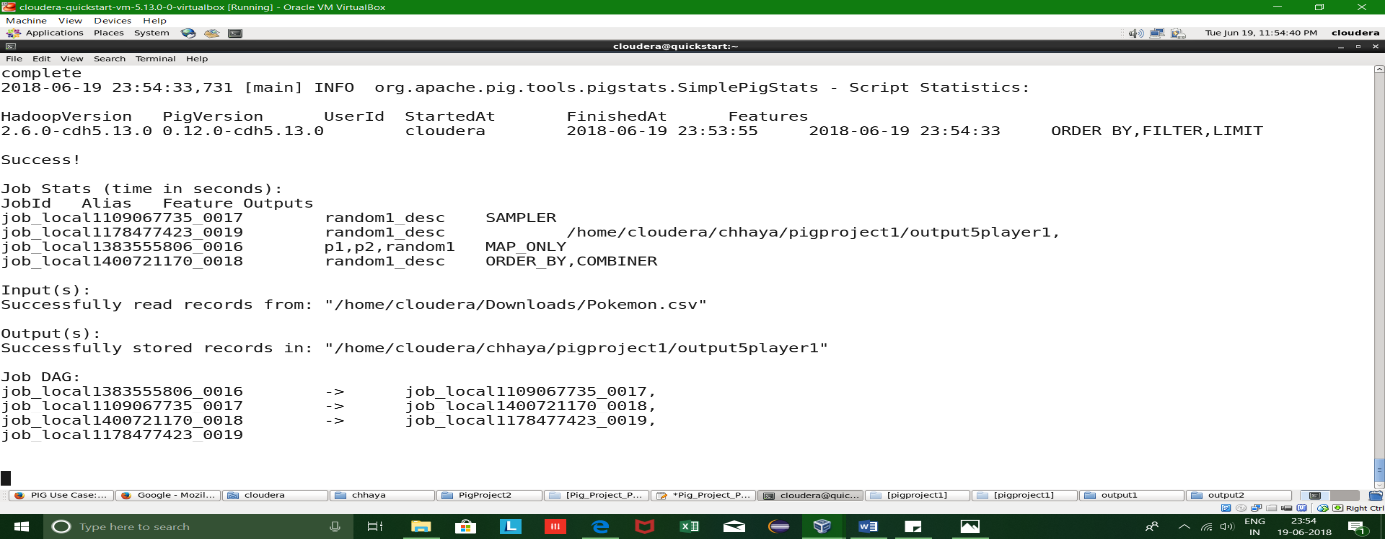
pick\_rand1 = LIMIT random1\_desc 5;

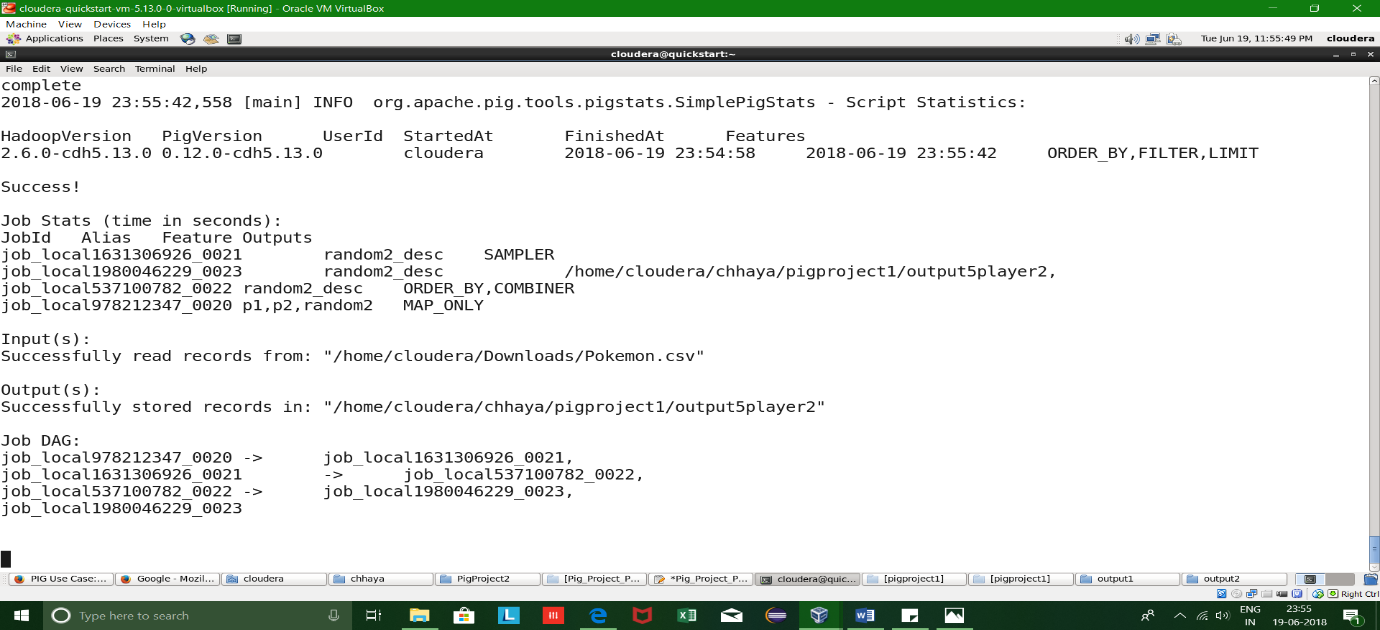
pick\_rand2 = LIMIT random2\_desc 5;

STORE pick\_rand1 INTO '/home/cloudera/chhaya/pigproject1/output5player1' USING PigStorage('\t');

STORE pick\_rand2 INTO '/home/cloudera/chhaya/pigproject1/output5player2' USING PigStorage('\t');

Output Screenshots 5:





================================================================

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).**

Input Pig Commands 6:

-------------------------------

REGISTER '/home/cloudera/Downloads/piggybank-0.17.0.jar';

p1 = LOAD '/home/cloudera/Downloads/Pokemon.csv' USING PigStorage(',')

AS (sr:int , pokemonname:chararray , type\_1:chararray , type\_2:chararray, total:int , hp:int, attack:int , defense:int , specialattack:int , specialdefense:int , speed:int);

p2 = FILTER p1 BY ( defense > 55);

random1 = FOREACH p2 GENERATE RANDOM() as random, pokemonname, type\_1, type\_2 , total, hp, attack, defense, specialattack, specialdefense, speed;

random1\_desc = ORDER random1 by random desc;

random2 = FOREACH p2 GENERATE RANDOM() as random, pokemonname, type\_1, type\_2 , total, hp, attack, defense, specialattack, specialdefense, speed;

random2\_desc = ORDER random2 by random desc;

pick\_rand1 = LIMIT random1\_desc 5;

pick\_rand2 = LIMIT random2\_desc 5;

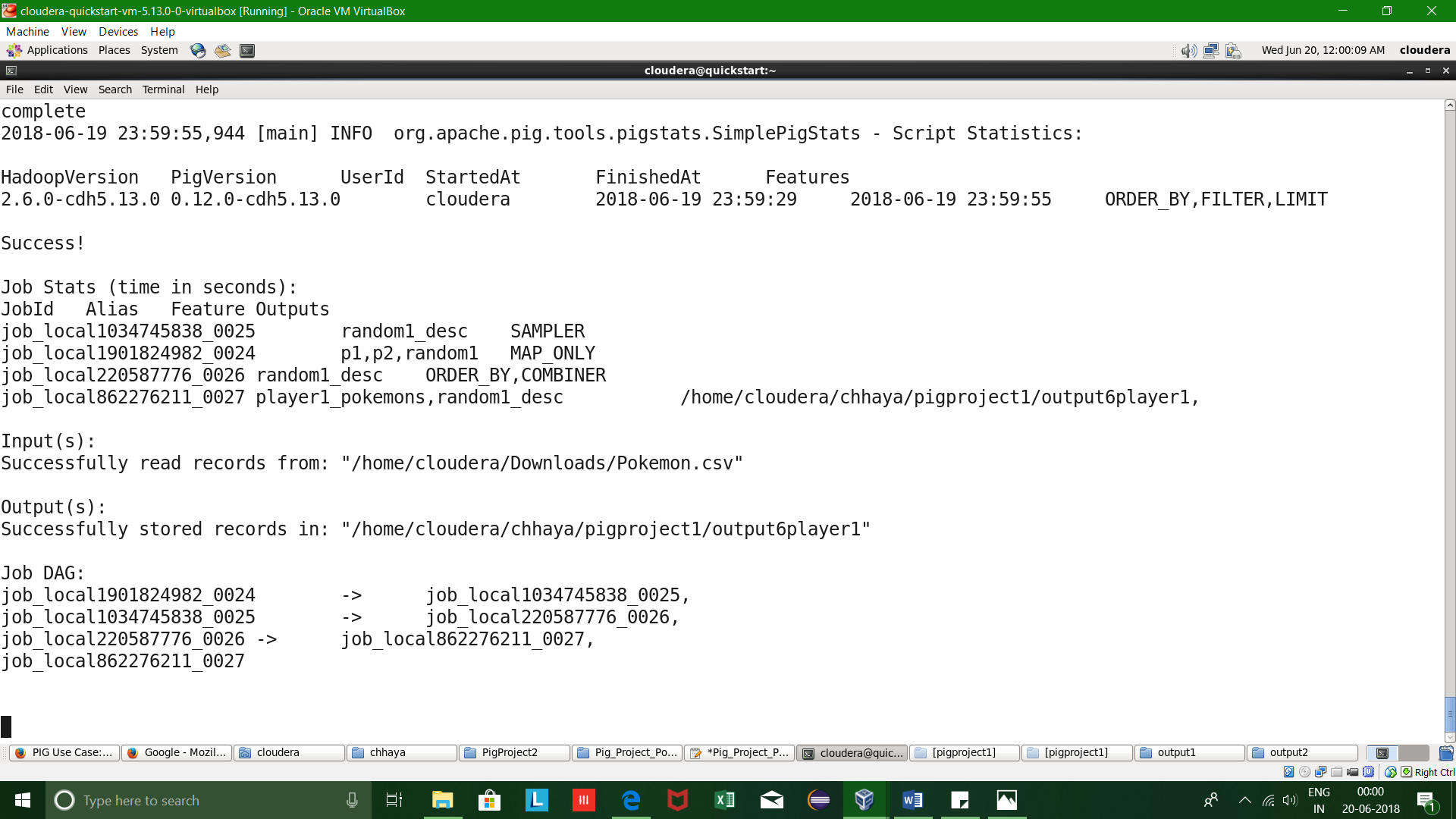
player1\_pokemons = FOREACH pick\_rand1 GENERATE pokemonname, hp;

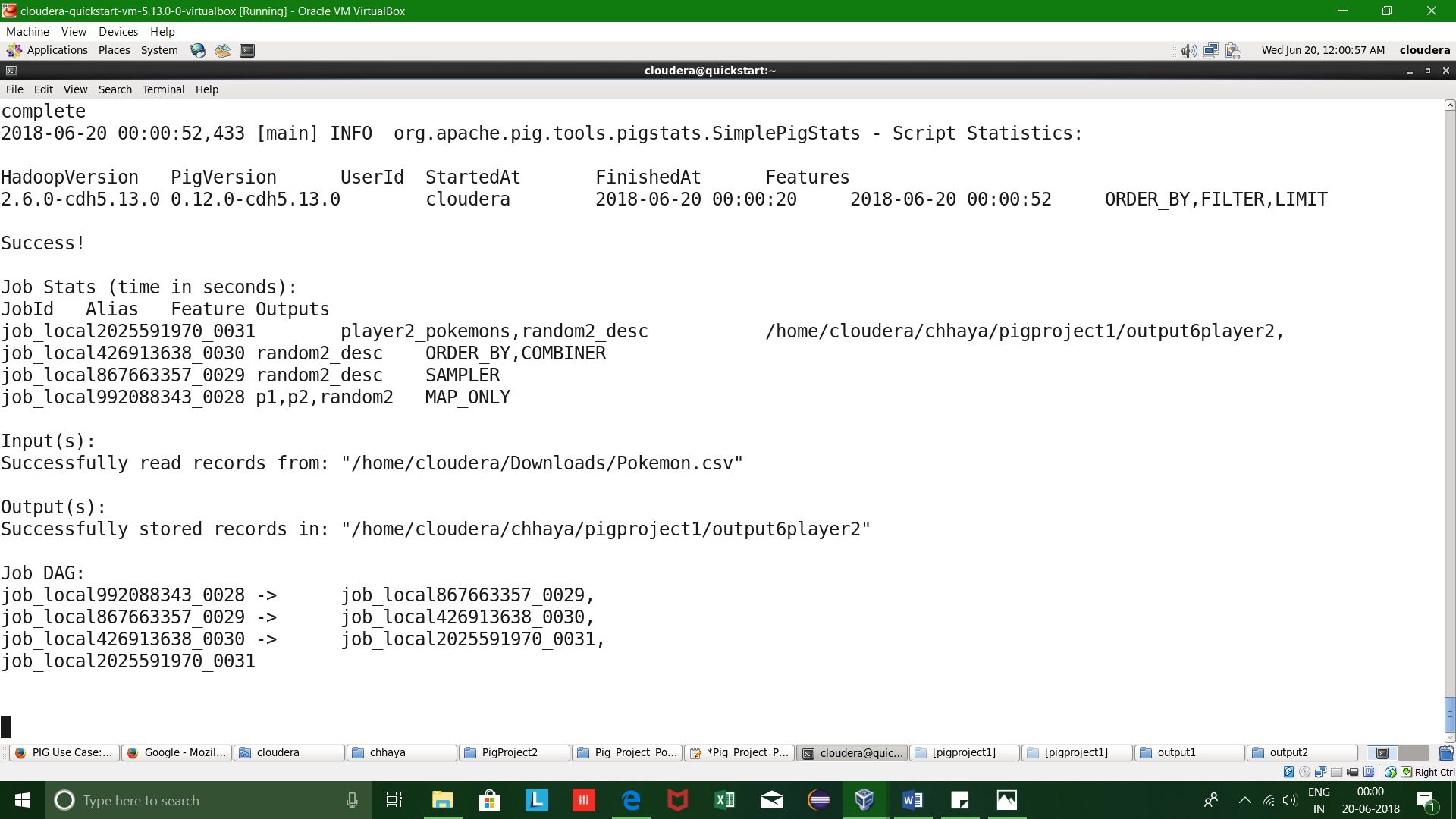
player2\_pokemons = FOREACH pick\_rand2 GENERATE pokemonname, hp;

STORE player1\_pokemons INTO '/home/cloudera/chhaya/pigproject1/output6player1' USING PigStorage('\t');

STORE player2\_pokemons INTO '/home/cloudera/chhaya/pigproject1/output6player2' USING PigStorage('\t');

Output Screenshots 6:





================================================================