Project Impala

1. Question 1
    a. **Creating a database**
        i. create database proj_ayush location '/user/f201605673498/myproj';
    b. **Creating table pokemon and loading data**

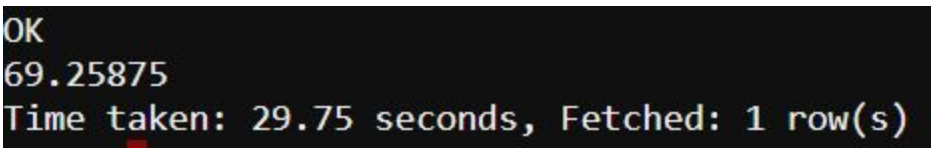
    create external table pokemon(pokemonid int,name string,type1 string,type2 string,total int,HP int,attack int,defence int,sp_atk int,sp_def int,speed int)
    row format delimited
    fields terminated by ','
    stored as textfile
    location '/user/f201605673498/project_impala';
    c. **Verify the table**
        i. select * from pokemon;

```
OK
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
9       Blastoise       Water           530     79      83      100     85      105     78
9       BlastoiseMega Blastoise Water           630     79      103     120     135     115     78
10      Caterpie        Bug             195     45      30      35      20      20      45
```

2. Question 2

   a. **Average HP of the pokemon**

      i. Select avg(HP) from pokemon;



3. Question 3

a. create table pokemon1 with an extra column power_rate

create external table pokemon2(pokemonid int,name string,type1 string,type2 string,total int,HP int,attack int,defence int,sp_atk int,sp_def int,speed int,power_rate string)
row format delimited
fields terminated by ','
stored as textfile;

b. Populate the table pokemon1

insert overwrite table pokemon1
select *, case when HP > 69.25 then "powerful" else "moderate" end power_rate from pokemon;

NOTE: Here to create and populate the power_rate column, I am comparing the HP of the pokemon. Thus, the pokemon having their HP greater than the average HP (i.e 69.25 taken from 2nd query) is populated as "powerful" and less than the average HP is populated as "moderate" in the power_rate column.

c. Verify the loading of data

i. select * from pokemon1;

```
OK
1       Bulbasaur       Grass   Poison  318     45      49      49      65      65      45      moderate
2       Ivysaur Grass   Poison  405     60      62      63      80      80      60      moderate
3       Venusaur        Grass   Poison  525     80      82      83      100     100     80      powerful
3       VenusaurMega Venusaur   Grass   Poison  625     80      100     123     122     120     80      powerful
4       Charmander      Fire            309     39      52      43      60      50      65      moderate
5       Charmeleon      Fire            405     58      64      58      80      65      80      moderate
6       Charizard       Fire    Flying  534     78      84      78      109     85      100     powerful
6       CharizardMega Charizard X       Fire    Dragon  634     78      130     111     130     85      100     powerful
6       CharizardMega Charizard Y       Fire    Flying  634     78      104     78      159     115     100     powerful
7       Squirtle        Water           314     44      48      65      50      64      43      moderate
8       Wartortle       Water           405     59      63      80      65      80      58      moderate
9       Blastoise       Water           530     79      83      100     85      105     78      powerful
9       BlastoiseMega Blastoise Water           630     79      103     120     135     115     78      powerful
10      Caterpie        Bug             195     45      30      35      20      20      45      moderate
```

4. Question 4

   a. **Count the number of powerful and moderate HP pokemon**

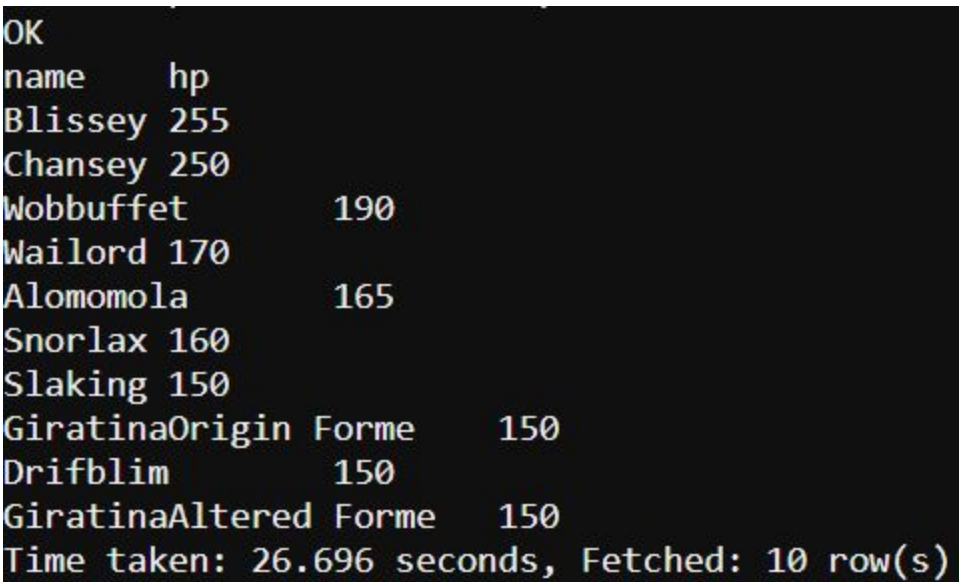
      i. select power_rate,count(power_rate) from pokemon1 group by power_rate;

```
OK
power_rate      cnt
moderate        422
powerful        378
Time taken: 25.305 seconds, Fetched: 2 row(s)
```

5. Question 5

a. Top 10 pokemon by their HP;

i. select name,HP from pokemon order by HP desc limit 10;



```
OK
name    hp
Blissey 255
Chansey 250
Wobbuffet       190
Wailord 170
Alomomola       165
Snorlax 160
Slaking 150
GiratinaOrigin Forme    150
Drifblim        150
GiratinaAltered Forme   150
Time taken: 26.696 seconds, Fetched: 10 row(s)
```