#!/bin/bash

show\_menu()

{

NORMAL=`echo "\033[m"`

MENU=`echo "\033[36m"` #Blue

NUMBER=`echo "\033[33m"` #yellow

FGRED=`echo "\033[41m"`

RED\_TEXT=`echo "\033[31m"`

ENTER\_LINE=`echo "\033[33m"`

echo -e "${MENU}\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*H1B APPLICATIONS\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*${NORMAL}"

echo -e "${MENU}${NUMBER} 1) ${MENU} Is the number of petitions with Data Engineer job title increasing over time?(MR)${NORMAL}"

echo -e "${MENU}${NUMBER} 2) ${MENU} Find top 5 job titles who are having highest growth in applications.(PIG) ${NORMAL}"

echo -e "${MENU}${NUMBER} 3) ${MENU} Which part of the US has the most Data Engineer jobs for each year?(HIVE) ${NORMAL}"

echo -e "${MENU}${NUMBER} 4) ${MENU} Find top 5 locations in the US who have got certified visa for each year.(HIVE)${NORMAL}"

echo -e "${MENU}${NUMBER} 5) ${MENU} Which industry has the most number of Data Scientist positions?(MR)${NORMAL}"

echo -e "${MENU}${NUMBER} 6) ${MENU} Which top 5 employers file the most petitions each year?(MR) ${NORMAL}"

echo -e "${MENU}${NUMBER} 7) ${MENU} Find the most popular top 10 job positions for H1B visa applications for each year?(HIVE)${NORMAL}"

echo -e "${MENU}${NUMBER} 8) ${MENU} Find the most popular top 10 job positions for Certified H1B visa applications for each year?(HIVE)${NORMAL}"

echo -e "${MENU}${NUMBER} 9) ${MENU} Find the percentage and the count of each case status on total applications for each year. Create a graph depicting the pattern of All the cases over the period of time.(HIVE)${NORMAL}"

echo -e "${MENU}${NUMBER} 10) ${MENU} Create a bar graph to depict the number of applications for each year(MR)${NORMAL}"

echo -e "${MENU}${NUMBER} 11) ${MENU}Find the average Prevailing Wage for each Job for each Year (take part time and full time separate) arrange output in descending order(PIG)${NORMAL}"

echo -e "${MENU}${NUMBER} 12) ${MENU}Which are the employers who have highest success rate in petitions more than 70%and total petions filed more than 1000?(PIG)${NORMAL}"

echo -e "${MENU}${NUMBER} 13) ${MENU}Which are the top 10 job positions which have the success rate more than 70% in petitions and total petitions filed more than 1000? (PIG)${NORMAL}"

echo -e "${MENU}${NUMBER} 14) ${MENU}Export result for option no 12 to MySQL database.${NORMAL}"

echo -e "${MENU}\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*${NORMAL}"

echo -e "${ENTER\_LINE}Please enter a menu option and enter or ${RED\_TEXT}enter to exit. ${NORMAL}"

read opt

}

function option\_picked()

{

COLOR='\033[01;31m' # bold red

RESET='\033[00;00m' # normal white

MESSAGE="$1" #modified to post the correct option selected

echo -e "${COLOR}${MESSAGE}${RESET}"

}

clear

show\_menu

while [ opt != '' ]

do

if [[ $opt = "" ]]; then

exit;

else

case $opt in

1) clear;

option\_picked "1 a) Is the number of petitions with Data Engineer job title increasing over time?";

hadoop fs -rmr /niit/q1;

hadoop jar h1b\_applications.jar Question1a /user/hive/warehouse/h1b\_final /niit/q1;

hadoop fs -cat /niit/q1/p\*

show\_menu;

;;

2) clear;

option\_picked "1 b) Find top 5 job titles who are having highest growth in applications. ";

pig /home/hduser/Documents/H1B\_Analysis/Pig/P1b.pig

show\_menu;

;;

3) clear;

option\_picked "2 a) Which part of the US has the most Data Engineer jobs for each year?";

echo -e "Enter the year (2011,2012,2013,2014,2015,2016)"

read var

hive -e "select worksite, case\_status,count(job\_title) as jobs from h1b\_final where job\_title like '%DATA ENGINEER%' and case\_status='CERTIFIED' and year='$var' group by worksite,year,case\_status order by jobs desc limit 1;"

show\_menu;

;;

4) clear;

option\_picked "2 b) Find top 5 locations in the US who have got certified visa for each year.";

echo -e "Enter the year (2011,2012,2013,2014,2015,2016)"

read var

hive -e " select year, worksite,count(case\_status) as status from h1b\_final where case\_status='CERTIFIED' and year='$var' group by worksite,year order by status desc limit 5;;"

show\_menu;

;;

5) clear;

option\_picked "3) Which industry has the most number of Data Scientist positions?";

hadoop fs -rmr /niit/q3;

hadoop jar h1b\_applications.jar Question3 /user/hive/warehouse/h1b\_final /niit/q3;

hadoop fs -cat /niit/q3/p\*

show\_menu;

;;

6) clear;

option\_picked "4)Which top 5 employers file the most petitions each year?";

hadoop fs -rmr /niit/q4;

hadoop jar h1b\_applications.jar Question4 /user/hive/warehouse/h1b\_final /niit/q4;

hadoop fs -cat /niit/q4/p\*

show\_menu;

;;

7) clear;

option\_picked "5a) Find the most popular top 10 job positions for H1B visa applications for each year?";

echo -e "Enter the year (2011,2012,2013,2014,2015,2016)"

read var

hive -e "select year,job\_title,count(case\_status) as popular from h1b\_final where year ='$var' group by year,job\_title order by popular desc limit 10;"

show\_menu;

;;

8) clear;

option\_picked "5b)Find the most popular top 10 job positions for Certified H1B visa applications for each year ?";

echo -e "Enter the year (2011,2012,2013,2014,2015,2016)"

read var

hive -e "select year,job\_title,case\_status,count(case\_status) as popular from h1b\_final where year ='$var' and case\_status= 'CERTIFIED'group by year,job\_title,case\_status order by popular desc limit 10;"

show\_menu;

;;

9) clear;

option\_picked "6) Find the percentage and the count of each case status on total applications for each year. Create a graph depicting the pattern of All the cases over the period of time.";

read var

hive -e "create table if not exists total(total int,year string) row format delimited fields terminated by ',';"

hive -e "insert overwrite table total select count(\*),year from h1b\_final where h1b\_final.case\_status is not NULL group by year;"

hive -e "select a.case\_status,count(\*) as case\_total,a.year,ROUND((count(\*)/b.total)\*100,2) as perofcase\_status from h1b\_final a left outer join total b on (a.year=b.year) where a.year is not NULL group by a.case\_status,b.total,a.year order by a.year;"

show\_menu;

;;

10) clear;

option\_picked "7) Create a bar graph to depict the number of applications for each year";

hadoop fs -rmr /niit/q7;

hadoop jar h1b\_applications.jar Question7 /user/hive/warehouse/h1b\_final /niit/q7;

hadoop fs -cat /niit/q7/p\*

show\_menu;

;;

11) clear;

option\_picked "8) Find the average Prevailing Wage for each Job for each Year (take part time and full time separate) arrange output in descending order";

echo -e "Enter the case\_status (1-Full time position,2-Part time position)";

read n

case "$n" in

"1")

pig /home/hduser/Documents/H1B\_Analysis/Pig/P8\_FT.pig;

;;

"2")

pig /home/hduser/Documents/H1B\_Analysis/Pig/P8\_PT.pig;

;;

esac

show\_menu;

;;

12) clear;

option\_picked "9) Which are the employers who have highest success rate in petitions more than 70%and total petions filed more than 1000?"

pig /home/hduser/Documents/H1B\_Analysis/Pig/P9.pig;

show\_menu;

;;

13) clear;

option\_picked "10) Which are the top 10 job positions which have the success rate more than 70% in petitions and total petitions filed more than 1000?"

pig /home/hduser/Documents/H1B\_Analysis/Pig/P10.pig;

show\_menu;

;;

14) clear;

option\_picked "11) Export result for option no 12 to MySQL database."

mysql -u root -p'user';

show\_menu;

;;

\n) exit;

;;

\*) clear;

option\_picked "Pick an option from the menu";

show\_menu;

;;

esac

fi

done