**Reg no: 21MAI003**

**ASSIGNMENT 2**

1. Write a query to display:  
   a. Employee Name and Manager name (employee’s manager)

create table employees (emp\_id int, first\_name string, last\_name string, email string, phone\_number int, hire\_date timestamp, job\_id string, salary float, commission\_pct float, manager\_id int, department\_id int) row format delimited fields terminated by ',';

**QUERY:**

SELECT e.first\_name, m.first\_name AS manager\_name

FROM employees e

JOIN employees m on e.manager\_id = m.emp\_id;

**SOLUTION:**

Total MapReduce CPU Time Spent: 9 seconds 150 msec

OK

Adam Steven

Karen Steven

Gerald Steven

John Steven

Eleni Steven

Lex Steven

Michael Steven

Matthew Steven

Neena Steven

Kevin Steven

Den Steven

Shanta Steven

Payam Steven

Alberto Steven

Nancy Neena

Jennifer Neena

Shelley Neena

Susan Neena

Hermann Neena

Alexander Lex

Valli Alexander

Diana Alexander

Bruce Alexander

David Alexander

Daniel Nancy

Luis Nancy

John Nancy

Jose Manuel Nancy

Ismael Nancy

Sigal Den

Guy Den

Shelli Den

Karen Den

Alexander Den

Girard Matthew

Irene Matthew

Winston Matthew

James Matthew

Julia Matthew

Steven Matthew

Jean Matthew

Martha Matthew

Mozhe Adam

Laura Adam

Julia Adam

Anthony Adam

Nandita Adam

TJ Adam

James Adam

Alexis Adam

Michael Payam

Kelly Payam

Jennifer Payam

Timothy Payam

Randall Payam

Jason Payam

Hazel Payam

Ki Payam

John Shanta

Renske Shanta

Joshua Shanta

Sarah Shanta

Samuel Shanta

Britney Shanta

Vance Shanta

Stephen Shanta

Randall Kevin

Alana Kevin

Kevin Kevin

Donald Kevin

Douglas Kevin

Trenna Kevin

Peter Kevin

Curtis Kevin

Christopher John

Oliver John

Nanette John

Peter John

David John

Peter John

Sarath Karen

Janette Karen

Patrick Karen

Allan Karen

Lindsey Karen

Louise Karen

Clara Alberto

David Alberto

Danielle Alberto

Amit Alberto

Mattea Alberto

Sundar Alberto

William Gerald

Tayler Gerald

Harrison Gerald

Lisa Gerald

Sundita Gerald

Elizabeth Gerald

Jonathon Eleni

Charles Eleni

Ellen Eleni

Kimberely Eleni

Alyssa Eleni

Jack Eleni

Pat Michael

William Shelley

Time taken: 37.126 seconds

b. Employee Name and Department Name (the employee belongs to)

create table emp (emp\_id int, first\_name string, salary float, department\_id int, department\_name string, location\_id int, city string, country\_id int, country\_name string) row format delimited fields terminated by ',';

**QUERY:**

SELECT e.first\_name, e.department\_name AS department

FROM emp e;

**SOLUTION:**

OK

Steven Executive

Neena Executive

Lex Executive

Alexander IT

Bruce IT

David IT

Valli IT

Diana IT

Nancy Finance

Daniel Finance

John Finance

Ismael Finance

Jose Manuel Finance

Luis Finance

Den Purchasing

Alexander Purchasing

Shelli Purchasing

Sigal Purchasing

Guy Purchasing

Karen Purchasing

Matthew Shipping

Adam Shipping

Payam Shipping

Shanta Shipping

Kevin Shipping

Julia Shipping

Irene Shipping

James Shipping

Steven Shipping

Laura Shipping

Mozhe Shipping

James Shipping

TJ Shipping

Jason Shipping

Michael Shipping

Ki Shipping

Hazel Shipping

Renske Shipping

Stephen Shipping

John Shipping

Joshua Shipping

Trenna Shipping

Curtis Shipping

Randall Shipping

Peter Shipping

John Sales

Karen Sales

Alberto Sales

Gerald Sales

Eleni Sales

Peter Sales

David Sales

Peter Sales

Christopher Sales

Nanette Sales

Oliver Sales

Janette Sales

Patrick Sales

Allan Sales

Lindsey Sales

Louise Sales

Sarath Sales

Clara Sales

Danielle Sales

Mattea Sales

David Sales

Sundar Sales

Amit Sales

Lisa Sales

Harrison Sales

Tayler Sales

William Sales

Elizabeth Sales

Sundita Sales

Ellen Sales

Alyssa Sales

Jonathon Sales

Jack Sales

Charles Sales

Winston Shipping

Jean Shipping

Martha Shipping

Girard Shipping

Nandita Shipping

Alexis Shipping

Julia Shipping

Anthony Shipping

Kelly Shipping

Jennifer Shipping

Timothy Shipping

Randall Shipping

Sarah Shipping

Britney Shipping

Samuel Shipping

Vance Shipping

Alana Shipping

Kevin Shipping

Donald Shipping

Douglas Shipping

Jennifer Administration

Michael Marketing

Pat Marketing

Susan Human Resources

Hermann Public Relations

Shelley Accounting

William Accounting

Time taken: 41.262 seconds

c. Country name and region name (country belongs to)

**SQOOP** :

sqoop import --connect jdbc:mysql://localhost/hr --table departments --target-dir /dept --username root -m 1;

create table emp\_detail(emp\_id int, job\_id string, manager\_id int , department\_id int, location\_id int, first\_name string, last\_name string, salary float, commission\_pct float,department\_name string,job\_title string,city string,state\_province string,country\_name string,region\_name string) row format delimited fields terminated by ',';

**QUERY**:   
SELECT DISTINCT e.country\_name, e.region\_name FROM emp\_detail e;

**SOLUTION**:

Total MapReduce CPU Time Spent: 3 seconds 160 msec

OK

Bavaria Germany

California United States of America

Ontario Canada

Oxford United Kingdom

Texas United States of America

Washington United States of America

null United Kingdom

Time taken: 35.293 seconds

2. How many countries fall under a region, write a query to display the region and total  
countries in that region.

**QUERY:**SELECT region\_name, count(country\_name) FROM emp\_detail group by region\_name;

**SOLUTION:**

OK

Canada 2

Germany 1

United Kingdom 35

United States of America 68

Time taken: 112.149 seconds

3. Write a query to display:  
a. Which department draws maximum salary?

**QUERY**

select department\_name, salary FROM emp group by department\_name, salary SORT BY salary DESC limit 1;

**SOLUTION:**

Total MapReduce CPU Time Spent: 9 seconds 380 msec

OK

Executive 24000.0

Time taken: 69.728 seconds

b. List highest paid employees in each department?

**QUERY:**

select department\_name, MAX(salary) FROM emp group by department\_name;

**SOLUTION:**

Total MapReduce CPU Time Spent: 4 seconds 180 msec

OK

Accounting 12000.0

Administration 4400.0

Executive 24000.0

Finance 12000.0

Human Resources 6500.0

IT 9000.0

Marketing 13000.0

Public Relations 10000.0

Purchasing 11000.0

Sales 14000.0

Shipping 8200.0

Time taken: 23.828 seconds

c. To which country the following departments belong?  
i. Administration  
ii. Executive  
iii. Manufacturing  
iv. Government Sales  
v. Recruiting

The countries with all departments to which employees belong:  
**QUERY:**  
select DISTINCT e.department\_name, e.country\_name from emp e;

**SOLUTION**:

Total MapReduce CPU Time Spent: 2 seconds 910 msec

OK

Accounting United States of America

Administration United States of America

Executive United States of America

Finance United States of America

Human Resources United Kingdom

IT United States of America

Marketing Canada

Public Relations Germany

Purchasing United States of America

Sales United Kingdom

Shipping United States of America

Time taken: 22.884 seconds

Countries that were specific to department in question:

**QUERY**:

In SQL:

create table depts as (select d.department\_name, c.country\_name

from departments d

join locations l

on d.location\_id = l.location\_id

join countries c

on l.country\_id = c.country\_id);

In sqoop:

sqoop import --connect jdbc:mysql://localhost/hr --table depts --target-dir /depts --username root -m 1;

In Hive:

**create table dept1(country\_name string, department\_name string)** row format delimited fields terminated by ',';**;**

**load data inpath '/depts' into table dept1;**

**select country\_name, regexp\_extract(department\_name, from**

select d.country\_name, d.department\_name from dept1 d where d.department\_name IN ('Administration','Executive','Manufacturing','Government Sales','Recruiting');

**SOLUTION:**

Job 0: Map: 1 Cumulative CPU: 0.84 sec HDFS Read: 1183 HDFS Write: 192 SUCCESS

Total MapReduce CPU Time Spent: 840 msec

OK

Administration United States of America

Executive United States of America

Manufacturing United States of America

Government Sales United States of America

Recruiting United States of America

Time taken: 12.986 seconds

4. Pratik has joined the company recently into a new department ‘Learning and  
Development’ as manager. Make necessary updates in the database tables and insert  
the record. Assume necessary data.

**Approach 1:**

**In sql:**

* insert into departments values(280, 'Learning and Development', 207,1500);
* insert into employee values(207, 'pratik',35000,280,'Learning and Development',1700,'Seattle','US','California');

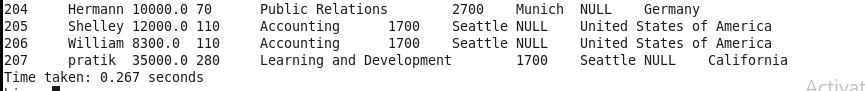
**In sqoop:**

* sqoop import --connect jdbc:mysql://localhost/hr --table departments --hive-import --username root --incremental append --check-column department\_id --last-value 270;
* sqoop import --connect jdbc:mysql://localhost/hr --table employee --hive-import --username root --incremental append --check-column employee\_id --last-value 206;

**In hive:**

load data inpath '/employee' into table emp;

**SOLUTION**:



**Approach 2**:

Loading directly to hive without modifying the table:

created a csv called assignment2.csv

207,’LNDEV’, 100, 80,2500, ’US’,’Pratik’,’mishra’, 39000.00,0.40,’Learning and Development’,’Development Head’, ‘Seattle’,‘ Washington’, ‘United States of America’

**In HIVE:**

load data local inpath '/home/training/Desktop/assignment2.csv' into table emp\_detail;

Value got added

A picture containing text

Description automatically generated