Data Analytics

Assignment - 03

Use simple dataset & apply all operations for Casendra and Hive

→ Here lets use simple dataset of employees. with fields 'emp_id' 'name' , 'dept' 'salary'

Casendra

i) create keyspace

Create keyspace company with Replication
emp_id - INT primary KEY,
FIRST - name text,
last - name text,
email .text,
dept text,
Salary INT ; // table created

iii) INSERT Into Data

Insert INTO company employees (emp_id, first name. last-name email , dept, salary)

iv) QUERY DATA

empid	Firstname	last-name	Email	dept	Salary
1.	Alice	John	Toho123@ gmail	HR	110000
2.	John	shree	nithy eshwer @ gmail .sm	Agsi s tant	120000

v) Update Data.

Update company.employees
SET email = 'johnny.doe @gmail.com'
WHERE emp—id = 1;

vi) DELETE DATA

DELETE FROM company.employees
WHERE emp—id = 5;

vii) DROP &TABLE

DROP TABLE company - employees

vii) DROP NAMESPACE

DROP TABLE company

HIVE operations:

i) create Database

CREATE DATABASE company // data, onyx created

(ii) USE DATABASE.
    = USE company;

(iii) CREATE TABLE
CREATE TABLE employees (
    emp_id. #INT,
    firstname STRING,
    last_name STRING,
    salary INT);

(iv) LOAD DATA

Suppose the file / employees / csv is stored
in HDFS @ user / hive / employees.csv.

v) INSERT DATA.

INSERT INTO TABLE employees VALUES

QUERY DATA:

Select * FROM employees;

OUTPUT

emp id	First-name	last-name	email	dept	salary
1.	Alice	.Doe	Doe123@gmail.com	HR	65000
2.	Vaishu	Davi	vaish438@gmail.com	Ansis	75000

vii) UPDATE DATA

Hive doesn't support direct updation.

- you can typically insert new data or create with updates

(viii) DELETE DATA

DELETE FROM employees
WHERE emp-id = 1;

X) QUERY DATA:

Select * FROM employees;

emp id	First name	last name	email	dept	salary
2.	Vaishu	Devi	Vaishy4021 3@gmail.com	Ansis tant	53000.

ix): DROP DATABASE

DROP TABLE E mployees,

x) DROP A TABLE

DROP DATABASE Company DATABASE

ro