1BM19CS073

KIRAN M K

Big Data Analytics Laboratory

6B2

Program 1. Perform the following DB operations using Cassandra.

1. Create a key space by name Employee

```
cqlsh> CREATE KEYSPACE employee WITH replication = {'class':'SimpleStrategy', 'replication_factor':1};
cqlsh> describe keyspace
No keyspace specified and no current keyspace
cqlsh> describe Employee;

CREATE KEYSPACE employee WITH replication = {'class': 'SimpleStrategy', 'replication_factor': '1'}  AND durable_writes = true;
cqlsh> create table Employee.employee_info(emp_id int Primary Key, emp_name text, designation text, date_of_joining timestamp, salary double,
dept_name text);
cqlsh> select * from Employee.employee_info;

 emp_id | date_of_joining | dept_name | designation | emp_name | salary
--------+-----------------+-----------+-------------+----------+--------
```

2. Create a column family by name Employee-Info with attributes Emp_Id Primary Key, Emp_Name,
Designation, Date_of_Joining, Salary, Dept_Name

```
cqlsh> begin batch insert into Employee.employee_info(emp_id, date_of_joining, dept_name, designation, emp_name, salary) values (1, '2021-06-0
3', 'Deploymemt', 'Manager', 'KIRAN', 1500000.50); apply batch;
cqlsh> select * from Employee.employee_info;

 emp_id | date_of_joining                 | dept_name  | designation | emp_name | salary
--------+---------------------------------+------------+-------------+----------+---------
      1 | 2021-06-02 18:30:00.000000+0000 | Deploymemt |     Manager |    KIRAN | 1.5e+06

(1 rows)
cqlsh> begin batch
   ... insert into Employee.employee_info(emp_id, date_of_joining, dept_name, designation, emp_name, salary) values (2, '2020-09-03', 'Develop
ment', 'Web Developer', 'VAISHAK', 1700000.50);
   ... insert into Employee.employee_info(emp_id, date_of_joining, dept_name, designation, emp_name, salary) values (121, '2019-05-03', 'Resea
rch and Development', 'Intern', 'VISHVESH', 2000000.50);
   ... apply batch;
cqlsh> select * from Employee.employee_info;

 emp_id | date_of_joining                 | dept_name                | designation   | emp_name | salary
--------+---------------------------------+--------------------------+---------------+----------+---------
      1 | 2021-06-02 18:30:00.000000+0000 |              Deploymemt |       Manager |    KIRAN | 1.5e+06
      2 | 2020-09-02 18:30:00.000000+0000 |             Development | Web Developer |  VAISHAK | 1.7e+06
    121 | 2019-05-02 18:30:00.000000+0000 | Research and Development |        Intern | VISHVESH |   2e+06

(3 rows)
```

3. Insert the values into the table in batch

```
cqlsh> update Employee.employee_info SET emp_name = 'PRAMOD', dept_name = 'Testing' WHERE emp_id = 121;
cqlsh> select * from Employee.employee_info;

 emp_id | date_of_joining                 | dept_name   | designation   | emp_name | salary
--------+---------------------------------+-------------+---------------+----------+---------
      1 | 2021-06-02 18:30:00.000000+0000 |  Deployment |       Manager |    KIRAN | 1.5e+06
      2 | 2020-09-02 18:30:00.000000+0000 | Development | Web Developer |  VAISHAK | 1.7e+06
    121 | 2019-05-02 18:30:00.000000+0000 |     Testing |        Intern |   PRAMOD |   2e+06
```

4. Update Employee name and Department of Emp-Id 121

```
cqlsh> create table Employee.emp(emp_id int, emp_name text, designation text, date_of_joining timestamp, salary double, dept_name text, prim
y key(emp_id, salary) );
cqlsh> begin batch
   ... insert into Employee.emp(emp_id, salary, date_of_joining, dept_name, designation, emp_name) values (1, 1500000.50, '2021-06-03', 'Dep
yment', 'Manager', 'KIRAN');
   ... insert into Employee.emp(emp_id, salary,date_of_joining, dept_name, designation, emp_name) values (2, 1100000.50, '2022-05-03', 'Deve
pment', 'Web Developer', 'VAISHAK');
   ... insert into Employee.emp(emp_id, salary, date_of_joining, dept_name, designation, emp_name) values (121, 1900000.50, '2022-05-05', 'R
earch and Development', 'Intern', 'VISHVESH');
   ... apply batch;
cqlsh> select * from Employee.emp;

 emp_id | salary  | date_of_joining                 | dept_name                | designation   | emp_name
--------+---------+---------------------------------+--------------------------+---------------+----------
      1 | 1.5e+06 | 2021-06-02 18:30:00.000000+0000 |               Deployment |       Manager |    KIRAN
      2 | 1.1e+06 | 2022-05-02 18:30:00.000000+0000 |              Development | Web Developer |  VAISHAK
    121 | 1.9e+06 | 2022-05-04 18:30:00.000000+0000 | Research and Development |        Intern | VISHVESH

(3 rows)
cqlsh> alter table Employee.employee_info add projects set<text>;
cqlsh> select * from Employee.emp;

 emp_id | salary  | date_of_joining                 | dept_name                | designation   | emp_name
--------+---------+---------------------------------+--------------------------+---------------+----------
      1 | 1.5e+06 | 2021-06-02 18:30:00.000000+0000 |               Deployment |       Manager |    KIRAN
      2 | 1.1e+06 | 2022-05-02 18:30:00.000000+0000 |              Development | Web Developer |  VAISHAK
    121 | 1.9e+06 | 2022-05-04 18:30:00.000000+0000 | Research and Development |        Intern | VISHVESH
```

5. Sort the details of Employee records based on salary

```
cqlsh> update Employee.employee_info set projects = projects + {'Hospital Management', 'Payment Interface'} where emp_id = 1;
cqlsh> select * from Employee.employee_info;

 emp_id | date_of_joining                 | dept_name   | designation   | emp_name | projects                                          | salary
--------+---------------------------------+-------------+---------------+----------+---------------------------------------------------+---------
      1 | 2021-06-02 18:30:00.000000+0000 |  Deployment |       Manager |    KIRAN | {'Hospital Management', 'Payment Interface'} | 1.5e+06
      2 | 2020-09-02 18:30:00.000000+0000 | Development | Web Developer |  VAISHAK |                                              null | 1.7e+06
    121 | 2019-05-02 18:30:00.000000+0000 |     Testing |        Intern |   PRAMOD |                                              null |   2e+06

(3 rows)
cqlsh> update Employee.employee_info set projects = projects + {'Lab', 'Helper'} where emp_id = 121;
cqlsh> select * from Employee.employee_info;

 emp_id | date_of_joining                 | dept_name   | designation   | emp_name | projects                                          | salary
--------+---------------------------------+-------------+---------------+----------+---------------------------------------------------+---------
      1 | 2021-06-02 18:30:00.000000+0000 |  Deployment |       Manager |    KIRAN | {'Hospital Management', 'Payment Interface'} | 1.5e+06
      2 | 2020-09-02 18:30:00.000000+0000 | Development | Web Developer |  VAISHAK |                                              null | 1.7e+06
    121 | 2019-05-02 18:30:00.000000+0000 |     Testing |        Intern |   PRAMOD |                                {'Helper', 'Lab'} |   2e+06
```

6. Alter the schema of the table Employee_Info to add a column Projects which stores a set of Projects done by the corresponding Employee.

```
cqlsh> insert into Employee.employee_info(emp_id, date_of_joining, dept_name, designation, emp_name, salary) values (11, '2019-05-05', 'Resear
ch and Development', 'Intern', 'SHREYAS', 1000000.50) using TTL 15;
cqlsh> select * from Employee.employee_info;

 emp_id | date_of_joining                 | dept_name                | designation   | emp_name | projects
 | salary
--------+---------------------------------+--------------------------+---------------+----------+-----------------------------------------
-+---------
     11 | 2019-05-04 18:30:00.000000+0000 | Research and Development |        Intern |  SHREYAS |                                    null
 |   1e+06
      1 | 2021-06-02 18:30:00.000000+0000 |               Deployment |       Manager |    KIRAN | {'Hospital Management', 'Payment Interface'}
 | 1.5e+06
      2 | 2020-09-02 18:30:00.000000+0000 |              Development | Web Developer |  VAISHAK |                                    null
 | 1.7e+06
    121 | 2019-05-02 18:30:00.000000+0000 |                  Testing |        Intern |   PRAMOD |                          {'Helper', 'Lab'}
 |   2e+06

(4 rows)
cqlsh> select * from Employee.employee_info;

 emp_id | date_of_joining                 | dept_name   | designation   | emp_name | projects                                          | salary
--------+---------------------------------+-------------+---------------+----------+---------------------------------------------------+---------
      1 | 2021-06-02 18:30:00.000000+0000 |  Deployment |       Manager |    KIRAN | {'Hospital Management', 'Payment Interface'} | 1.5e+06
      2 | 2020-09-02 18:30:00.000000+0000 | Development | Web Developer |  VAISHAK |                                              null | 1.7e+06
    121 | 2019-05-02 18:30:00.000000+0000 |     Testing |        Intern |   PRAMOD |                                {'Helper', 'Lab'} |   2e+06
```

7. Update the altered table to add project names.

```
cqlsh> update Employee.employee_info set projects = projects + {'API', 'Mood detection'} where emp_id = 2;
cqlsh> paging off;
Disabled Query paging.
cqlsh> select * from Employee.emp where emp_id in (1,2,121) order by salary;

 emp_id | salary  | date_of_joining                 | dept_name                | designation    | emp_name
--------+---------+---------------------------------+--------------------------+----------------+----------
      2 | 1.1e+06 | 2022-05-02 18:30:00.000000+0000 |              Development | Web Developer |   VAISHAK
      1 | 1.5e+06 | 2021-06-02 18:30:00.000000+0000 |               Deployment |        Manager |     KIRAN
    121 | 1.9e+06 | 2022-05-04 18:30:00.000000+0000 | Research and Development |         Intern | VISHVESH
```