**BDA Practical Question Bank**

**For practical exam you will be getting one problem from category I and other from category II**

**Category I**

Solve using Sqoop/Hive/HBase -- 10 marks

Q.1. Employee( ename ,street ,city)

Works(ename ,cname ,salary)

Company(cname ,city)

Manages(ename ,mname)

Consider above database draw E-R Diagram

Apply following constraints

Workers salary should be within range 10000 to 35000

Customer id should begin with letter c.

Primary Key and foreign key.

And write SQL query for following Statements.

ü Find all employees from database who earn more than their manager.

ü List the names of company which has maximum number of employees.

ü Find cname and no of employees working only for those company in which at least 5 employees are working.

ü List all the employees who live in the same cities as the company for which they works in.

**ANSWERS:**

1. HIVE
   1. **Create tables**

create table Employee( ename varchar(20) CHECK(ename like 'c%'), street varchar(20), city varchar(20));

create table Company (

    -> cname varchar(20) PRIMARY KEY,

    -> city varchar(15));

create table works( ename varchar(20), cname varchar (20) REFERENCES Company(cname), salary int(7) CHECK (salary BETWEEN 10000 AND 35000));

**Query 1 : Find all employees from database who earn more than their manager.**

mysql> select \* from works w1,manages,works w2 where w1.ename=manages.ename and manages.mname=w2.ename and w1.salary>w2.salary;

ename | cname | salary | ename | mname  | ename | cname | salary |

| Charu | Jio   | 31000 | Charu | Chaman | chaman | Jio   | 30000 |

| Char  | JPMC |  29000 | Char  | Chirag | Chirag | JPMC  | 28000 |

**Query 2 : List the names of company which has maximum number of employees**

mysql> select count(ename),cname from works group by cname order by count(ename) desc LIMIT 1;

           4 | JPMC  |

**Query 3 : Find cname and no of employees working only for those company in which at least 5 employees are working**

mysql> select count(ename),cname from works group by cname having count(ename)>=4;

count(ename) | cname |

|            4 | JPMC

**Query 4:  List all the employees who live in the same cities as the company for which they works in.**

mysql> select \* from works w,Employee e, Company c where w.cname=c.cname AND w.ename=e.ename AND e.city=c.city;

-------------------------------------------------------------------------------------------------------------------

Q.2. Customer(cid,cname,city)

Deposits(cid,accno)

Account(branchname,accno,balance)

Borrows(lno,cid)

Loan(lno,amount)

Consider above database draw E-R Diagram

ü Apply following constraints

Account balance should be within range 10000 to 25000

Customer id should begin with letter c.

Primary Key and foreign key.

write SQL query for following Statements

**Query 1 : Find customer names those are having account balance more than loan**

hive> select \* from customer c,deposits d,account a,borrows b,loan l where c.cid=d.cid and d.cid=b.cid and d.accno=a.accno and b.lno=l.lno and a.balance>l.amount

**Query 2 : Find branch name with minimum assets.**

hive> select bname,sum(balance) as assets from account group by bname order by assets asc limit 1;

**Query 3 : Find customer id, customer name of customers those are having at least 2 accounts and at least one loan.**

hive> select d.cid,c.cname from customer c, deposits d where c.cid=d.cid group by d.cid,c.cname having count(d.cid)>=2;

**Query 4 : Find those accounts whose balance is more than all accounts at dadar branch.**

**Query 5 : Delete all accounts which belongs to dadar branch and balance is more than account balance of accounts of john**

Q.3. A library has the following relations

Library(code ,name, no\_of\_books )

Person(id,name,age)

Imember(code,id,Dateofjoining)

Books(Accessionno,title,author,price)

Borrowedby(Accessionno,id,Date\_of\_borrow)

ü Consider above database draw E-R Diagram.

ü Apply following constraints

Personid should begin with letter P.

Primary Key and foreign key.

write SQL query for following Statements

**Query 1 : Give details of person who has borrowed at least two books.**

hive> select p.name,count(bb.id) from person p,borrowedby bb where bb.id=p.id group by p.name having count(bb.id)>=2;

**Query : Give details of person who has borrowed at least one book along with database concepts.**

hive> select p.name,count(bb.id) from person p,books b,borrowedby bb where bb.id=p.id and b.title='db' and bb.accession=b.accession group by p.name having count(bb.id)>=2;

ü Find name of book which has been borrowed minimum number of times

ü Delete all entries from borrowedby of database book borrowed by pid P101.

ü Find person details of persons who has borrowed database books with author korth and navathe.

Q.4 Employee(ssn ,ename ,salary ,superssn ,dno,pno)

Dept(dno,dname)

Project(pno,pname,dno)

Dependent(ssn,dependentname,relationship)

Apply following constraints

SSN should be exactly length 3.

Primary Key and foreign key.

Consider the above database draw E-R Diagram and write SQL statements for the following queries.

ü Retrieve employee name and supervisor name of employees those are earning salary more than their respective supervisors.

ü Retrieve employee details of employees those are earning salary more than

Average salary of department for which employee is working.

ü Give 15% raise in salary if salary is greater than 20000, 10% raise if salary

is within range 10000 to 20000 else 5% raise.

ü Retrieve employee details of employee those belongs to IT department and working on at least one project controlled by IT department.

ü Retrieve employee details those are working on Inventory project but does not Belongs to computer department.

ü Give one example of multiple table based view.

ü Consider schema

Branch (bname0, assets, city) and Accounts (accno,balance,bname)

Q.5. Customer(cid,cname,city,accno)

Account(bname,accno,balance)

Borrows(lno,cid)

Loan(lno,amount)

Consider above database draw E-R Diagram

ü Apply following constraints

Account balance should be within range 10000 to 25000

Customer id should begin with letter c.

Primary Key and foreign key.

Write SQL query for following Statements

ü Find customer details of customers those are having account balance more than

20000 at dadar branch and at least one loan. .

ü Find bname of branches those are having at least 2 accounts.

ü Find customer id, customer name those are having at least 1 accounts and at least 2 loan.

ü Give one example of left outer Join.

ü Delete all accounts which belongs to dadar branch and balance is more than account balance of accounts of john

üü Give example of multi table based view and show it’s updation

Q6. In sqoop

Create table in mysql, import tables in sqoop , export tables from sqoop

Q7. Hue for data analysis

**Category II**

Solve **pyspark/mapreduce** program 15 marks

1. Program to count 4-lettered words

2. Program to count words starting with ‘I’

3. Program to count 3-lettered words

4. Program to count 2-lettered words

5. Program to give matrix-vector multiplication

6. Program to implement join of tables:

7. Program to sort the given dataset

8. Program to find given word string in the dataset

9. Program to find average temperature/user rating

10. Program to implement k-means

11. Program to implement pagerank