

SQL Repository - 2020

Q1. [TYPE: Data Manipulation, CONTRIBUTOR: Malay Bhattacharyya] Consider the following schema representing the Indian Power League winners and Player of Series of all time.

IPL = <year : integer, *venue* : string, *winning_team* : string, *player_of_series* : string>

POS = <po_series : string, *team* : string>

The primary keys are underlined in the corresponding tables. Note that a Player of Series is not necessarily from the winning team. Assume that none of the Indian Power League players can change their teams. Write the following queries in SQL.

- (i) Find the names of all the Players of Series who do not belong to the winning team.
- (ii) Find the name of the most winning team in Indian Power League. Note that, there can be multiple most winning teams.
- (iii) Find the name of the team that won the previous to the last edition of Indian Power League. Note that Indian Power League might not have occurred each year.

Ans. The database schema in SQL:

```
create table POS(  
    po_series char(40),  
    team char(40),  
    primary key (po_series)  
);  
create table IPL(  
    year int(8) not null,  
    venue char(40),  
    winning_team char(40),  
    player_of_series char(40),  
    primary key (year),  
    foreign key (player_of_series) references POS(po_series)  
);
```

The queries in SQL:

(i) **OPTION 1 –**

```
select t1.player_of_series from POS t1  
where exists  
(  
select * from IPL t2 where t1.player_of_series = t2.player_of_series and t1.team<>t2.winning_team  
);
```

OPTION 2 –

```
select distinct a.player_of_series  
from IPL a  
natural join POS b  
where b.team <> a.winning_team;
```

OPTION 3 – ...;

OPTION 4 – ...;

(ii) **OPTION 1 –** select t2.winning_team from(

```
select t1.winning_team,max(t1.win_count) from
(
select winning_team,count(winning_team) as win_count from ipl group by winning_team
)t1)t2;
```

```
OPTION 1 – select WinTeam
            from (select WinTeam, count(*) as times
                  from IPL
                  group by WinTeam) freq
            where times = (select max(times)
                          from (select WinTeam, count(*) as times
                                from IPL
                                group by WinTeam) freq);
```

```
;
(iii) OPTION 1 –
select WINNER from IPL order by YEAR DESC LIMIT 1,1;
```

```
OPTION 2 –
select df1.winning_team
from
    (select winning_team, year
     from ipl
     order by year DESC limit 2) df1
order by df1.year ASC limit 1;
```

Q2. [TYPE: Data Definition, CONTRIBUTOR: Malay Bhattacharyya] Consider the following schema representing the names of Nobel Prize winners of all time.

NOBEL = <year : integer, category : string, name : string, country : string>

Create this table following all possible basic integrity constraints. Assign a primary key thoughtfully even though it is not mentioned in the schema. Note the following things:

- A year cannot be negative
- There are cases where no Nobel Prize was awarded for a particular category in a year
- The same person can win multiple Nobel prizes and that is even in different categories

Ans. The database schema in SQL:

...

```
OPTION 1 -
Create Table NOBEL(
Year int(4) not null,
Category varchar(100),
Winner ID int(4),
Name varchar(100),
Country varchar(100),
Primary key(Year, Category, Winner ID),
```

```
check( Category in ('---','---',.....), Country in ('---','---',.....))
);
```

Q3. [TYPE: Data Manipulation, CONTRIBUTOR:20BM6JP54]

Consider the table worker having the following attributes and Integrity constraints:

```
CREATE TABLE Worker (
    WORKER_ID INT NOT NULL PRIMARY KEY,
    FIRST_NAME CHAR(25),
    LAST_NAME CHAR(25),
    SALARY INT(15),
    JOINING_DATE DATETIME,
    DEPARTMENT CHAR(25)
);
```

- i)Write an SQL query to fetch nth max salaries from the Worker table.
- ii)Write an SQL query to print the FIRST_NAME from the Worker table after removing white spaces from the right side.
- iii)Write an SQL query to show only odd rows from the Worker table.
- iv)Write an SQL query to fetch the list of employees with the same salary from the Worker table.

Ans. ...

Q4. [TYPE: Data Manipulation , CONTRIBUTOR: 20BM6JP37]

Database scheme Hackers (hacker_id: int, name: string)

Submission (submission_date:date,submission_id: int,hacker_id: int,score:int)

Write a query to get the total number of unique hackers who made at least a submission each day (starting on the first day of the contest), and find the hacker_id and name of the hacker who made maximum number of submissions each day. If more than one such hacker has a maximum number of submissions, print the lowest hacker_id. The query should print this information for each day of the contest, sorted by the date.

Ans. ...

Q5. [TYPE: Data Manipulation, CONTRIBUTOR: 20BM6JP05] Consider the following schema CITIES representing the cities and their details.

CITIES=<ID: integer, CITY: string, STATE: string, LAT_N: integer, LONG_W: integer>

Query the two cities with shortest and longest city names as well as their respective lengths i.e. number of characters in city name. If there is more than one smallest or largest city, choose the one that comes first when ordered alphabetically.

Ans. ...

Q6. [TYPE: Data Manipulation, CONTRIBUTOR: 20BM6JP20] ...

Consider the following schema, Credit from a bank(XYZ), this schema includes customer ids, name (full), Age (in years), Bank Balance in INR (no negative balances), Profession- (Managers,Teacher,Doctor etc.), Existing_loan (is a yes/no column), Credit_approval is a Boolean value (0 for no and 1 for yes), Credit limit is the maximum amount that can be dispatched, credit limit is NULL for those with no credit approvals.

Credit = <Customer_id : string, Name : string, Age : integer, Bank_Balance : integer, DOB : Date, Profession : string, Existing_loan : string, Credit_approval : number, Credit_Limit : integer>

```
CREATE TABLE Credit(  
    Customer_id INT NOT NULL PRIMARY KEY,  
    Name varchar(25),  
    Age INT(4),  
    Bank balance INT NOT NULL(10),  
    DOB DATETIME,  
    Profession varchar(25)  
    Existing_Loan varchar(4)  
    Credit_approval tinyint(1)  
    Credit_limit INT(10)  
);
```

Q1) Find the person with the highest balance amount, check whether the same person has the highest approved credit limit.

Q2) Find the number of senior citizens(age >= 60) who have taken a loan (i.e having a yes in the Existing_loan column).

Q3) Find the proportion of people with credit approval granted (Credit_approval is an int expression with values either 1 or 0, 1 for a yes and 0 for a no).

Ans.

Q7. [Type: Data Manipulation , Contributor: 20BM6JP46]

Given a Relation R(A,B,C) where each column/attribute specifies the length of a triangle. Write a query identifying the type of each record in the TRIANGLES table using its three side lengths. Output one of the following statements for each record in the table:

- Equilateral: It's a triangle with sides of equal length.
- Isosceles: It's a triangle with sides of equal length.
- Scalene: It's a triangle with sides of differing lengths.
- Not A Triangle: The given values of A, B, and C don't form a triangle.

Q8. [TYPE: Data Manipulation, CONTRIBUTOR: 20BM6JP19] Chef TJ conducted a competitive coding competition consisting of 3 challenges (Breakfast Challenge, Lunch Challenge and Dinner Challenge) for 61 engineers. Out of the 3 challenges, only the Dinner challenge was a team event.

BREAKFAST = <roll: integer, name: string, language: string, score: int>

LUNCH = <roll: integer, name: string, language: string, score: int >

DINNER = <roll: integer, name: string, language: string, team: string, rank: int, score: int >

Write the SQL queries for following operations.

(i) Find the total number of unique languages used in the overall competition.

- (ii) Find the number of engineers with scores greater than 90 in both Breakfast and Lunch Challenge.
- (iii) Find the difference between highest and lowest total score.
- (iv) Find the overall winner - highest scoring member among the top 3 teams of Dinner Challenge.

Ans.

Q9. [TYPE: Data Manipulation, CONTRIBUTOR: 20BM6JP33]

The following relations keep track of airline flight information:

Flights = <flno: integer, from: string, to: string, distance: integer, departs: time, arrives: time, price: real>

Aircraft = <aid: integer, aname: string, cruisingrange: integer>

Certified = <eid: integer, aid: integer>

Employees = <eid: integer, ename: string, salary: integer>

Note that the Employees relation describes pilots and other kinds of employees as well; every pilot is certified for some aircraft, and only pilots are certified to fly. Write each of the following queries in SQL

1. Find the names of aircraft such that all pilots certified to operate them have salaries more than \$80,000.
2. Find the names of pilots whose salary is less than the price of the cheapest route from Los Angeles to Honolulu.
3. Identify the routes that can be piloted by every pilot who makes more than \$100,000
4. Compute the difference between the average salary of a pilot and the average salary of all employees (including pilots)

Ans. ...

Q10. [TYPE: Data Manipulation, CONTRIBUTOR: 20BM6JP38]

Consider the following three schemas representing student details, subject details and marks of students respectively.

STUDENT = <stid: integer, stname: string>

SUBJECT = <subid: integer, subname:string>

MARKS = <stid: integer, subid: integer, mark: integer>

Find all the student names who have scored the maximum marks in each subject.

Ans. ...

Q11. [TYPE: *Data Manipulation*, CONTRIBUTOR: *20BM6JP47*]

The following relations keep track of Award Winners from 1950 to 2020:

OSCAR <year: integer, category: string, name: string, movie: string, nominations: integer, genre: char>

ACTOR <name: string, year: integer, nominations: integer, genre: char>

DIRECTOR <name: string, year: integer, nominations: integer, genre: char>

1. How many Oscars did the Genre “*thriller*” receive from the year 1960 to 1990
2. Who were the directors and actors who got the maximum number of Oscars between the years 1971 to 2000
3. Which Genre movies has seen the most drastic drop in number of Oscars received from the 1980's to 2000's

Ans. ...

Q12. [Type: DDL, Triggers, CONTRIBUTOR: 20BM6JP04]

After having studied FunDS from Prof. Malay Bhattacharyya, Gabbar Singh and his gang were inspired, and feeling the pain of slowdown, they decided to get organized to fatten the thinning profit margins. They felt the need to create a database of all the monthly collection money that they forcibly procured from the poor people of Ramgarh and other nearby villages. His first lieutenant – Saambha – in consultation with the local goon - Kaalia, suggested the following schema –

GANGMEMBER = <g_id: integer, name: character>

VILLAGE = <v_id: integer, v_name: character>

COLLECTION = <house_id, v_id, g_id , collection_tgt, collection_achieved, month: integer, year: integer>

REVENUE = <v_id, year, month, v_revenue, v_tgt, surplus_or_deficit>

Gabbar Singh made it clear that none of the attributes in any table should take a Null value. All primary/composite keys have been underlined above.

1. Continuously tormented by the goons, Thakur and his family packed their bags and moved out from Ramgarh to Las Vegas to seek new beginnings in the hope of a bright future. Write a query to remove all his entries in the COLLECTION table.
2. Irritated by the clumsiness of Kaalia, Gabbar singh decided it was high time they added some integrity control. Create a mechanism such that an entry isn't allowed to be made to the COLLECTION and REVENUE tables if the month value is either above 12 or below 1. Also, do the same if year is less than 2020.
3. Write a trigger to automatically update the REVENUE table corresponding to the given village, whenever a collection_tgt or collection_achieved entry is made to the COLLECTION table. NOTE that the v_revenue

will be the sum of all collection_achieved tuples corresponding to the v_id in the VILLAGE table. Similarly, the v_tgt will be the sum of all collection_tgt tuples corresponding to the given v_id in the VILLAGE table.

4. Define the “surplus_or_deficit” attribute in REVENUE table such that the difference between the revenue and target gets displayed there automatically. Default value for Revenue is Zero.

Q.13 [TYPE: Data Manipulation, CONTRIBUTOR: 20BM6JP17]

Consider the following schema representing the tables of an Employee Database:

- **Company_Employees** = <emp_id : integer, emp_name : string, job_name : string, manager_id : integer, hire_date : Date, salary : Decimal(10,2), Commission : Decimal(7,2), dep_id : integer>
- **Company_Department** = <dep_id : integer, dep_name : string, dep_location : string>
- **Pay_Grade** = <grade : integer , min_salary : integer , max_salary : integer>

1. Write a query in sql to list the employees whose salary is more than 3000 after giving 30% increment.
2. Write a query in SQL to list the ID, name, salary, and job_name of the employees for whom:
 - a. Annual salary is below 30000 but receiving some commission which should not be more than the salary.
 - b. And designation is SALESMAN and working for department 3001.
3. Write a query in SQL to list the employee ID, name, salary, department name of all the 'MANAGERS' and 'CONSULTANTS' working in DELHI, MUMBAI with an exp more than 4 years without receiving the commission and display the list in ascending order of location.
4. Write a query in SQL to list the employees of grade 3 and 4 working in the department of SALES or FINANCE and whose salary is more than the salary of NIKHIL and experience is more than SACHIN. List the result in the ascending order of experience.

Q14. [Type: Data Manipulation, CONTRIBUTOR: 20BM6JP60]

A competition is organised for the students of 5 IITs in India (B, D, K, M, KGP) to award prize money based on their performance. Each candidate is given a score with upto 2 decimal points out of 100 (Higher the score, better the performance). No two candidates can have the exact same score. Finally, prizes are distributed to the following students:

1. Top 5 highest marks scorers from each IIT where participant count was > 100 and Top 2 highest marks scorers from each IIT where participant count was between 50 and 100 (both included). Prize money = 50000 each student
2. Top 5 highest marks scorers from across the country. Prize money = 100000 each student
3. Next 5 highest marks scorers from across the country. Prize money = 50000 each student

A candidate who is eligible for prize money in both the categories (inter-college and intra-college) will receive the total amount. Display in tabular format the names of all the students who are eligible for prize along with the name of their IIT and the total amount they will receive. Students from the same IIT should be displayed together. Within each IIT, students should be displayed in the decreasing order of their total prize amount won.

Use the following schema to write appropriate queries.

SCORE: <score : double, s_name : string, IIT : string, state : string>

Q15. [Type: Data Manipulation, CONTRIBUTOR: 20BM6JP57]

Consider the following relational schema of employees.

EMP=<eid: integer, ename: string, age: integer, salary: real>

WORKS=<eid: integer, did: integer, pct_time: integer>

DEPT=<did: integer, budget: real, managerid: integer>

1. Define a table constraint on Emp that will ensure that every employee makes at least \$12,000.
2. Define a table constraint on Dept that will ensure that all managers have age ≥ 30 .
3. Define an assertion on Dept that will ensure that all managers have age ≥ 30 . Compare this assertion with the equivalent table constraint. Explain which is better.
4. Write SQL statements to delete all information about employees whose salaries exceed that of the manager of one or more departments that they work in. Be sure to ensure that all the relevant integrity constraints are satisfied after your updates.

Ans. ...

Q16. [Type: Data Manipulation, CONTRIBUTOR: 20BM6JP06]

You are the database administrator for an analytics course jointly offered by three institutes with one semester at each institute.

SEMESTERS = <Sem: integer, Institute : string Start Date : date, End Date : date, Duration : integer >

During the Sem1 semester, the students of the course found the schedule to be too demanding and requested for an extension of 1 month in each semester.

Update the Semesters table to increase the duration of each semester by one month. This is the table currently:

Sem	Institute	Start Date	End Date	Duration
0	Institute3	15-10-2020	15-11-2020	1
1	Institute1	15-11-2020	15-02-2021	3

2	Institute2	15-02-2021	15-05-2020	3
3	Institute3	15-05-2021	15-08-2021	3
4		15-08-2021	15-02-2022	6

Ans. ...

Q17. [Type: Data Manipulation, CONTRIBUTOR: 20BM6JP35]

Let's suppose after a few years down the line after PGDBA, you started a firm that manages and regulates the service providers(SALESMAN) for several other registered companies(CUSTOMER). You, as a boss after a break of 10 days want to know the whereabouts of the progress and implement some ideas that came into your mind. Your firm maintains the database having following schemas:

SALESMAN: <salesman_id : integer, name : string, city : string, commission : decimal(3,2)>

CUSTOMER: <customer_id : integer, cust_name : string, city : string, salesman_id : integer>

ORDERS: <ord_no : integer, purch_amt : float(7,2), ord_date : date, customer_id : integer, salesman_id : integer>

1. You want to conduct a Fun Meet of all stakeholders of your company after you realise the need for a break and make employees indulge in informal talks to know each other better. Hence, write a query to find those customers with their name and those salesmen with their name and city who live in the same city so that you can mail them the venue to meet.

Coming back to work:

2. You want to find the names of all customers along with the salesmen who work for them so that you can suggest them on how to improve their efficiency depending on the customer they cater.

3. You want to hire a few salesmans to speed up the service. So to get an idea about vacancy, write a query that displays all those orders by the customers not located in the same cities where their salesmen live.

4. To keep track of operations,

- Write the query that finds out each order number followed by the name of the customers who made the order in order.
- Write a query that produces all customers with their name, city, salesman and commission, who is served by a salesman and the salesman works at a rate of the commission within 12% to 14%.

Q18. [Type: Data Definition, Data Manipulation, CONTRIBUTOR: 20BM6JP29]

match: match_id, team_id, opponent_id, season_id, venue_id, country, toss_winner_id, match_winner_id, Win_Type, outcome, man_of_the_match_id, win_margin

match_id is the primary key and team_id, Venue_id are foreign keys for the table match

Team: team_id, team_name

Venue: Venue_id, Venue_name

Win_Type values as by_runs indicate that the team bowling first has won.

Create tables match, Team and Venue with necessary constraints and Find the venue(s) where the team bowling first has won the match across seasons. If there are more than 1 venues, list all of them in order of the number of wins (by the bowling team).

Q19. [Type: Data Manipulation, CONTRIBUTOR: 20BM6JP14]

Take a look at the following tables consisting of marks of students for mid semester and end semester.

Mid Sem Marks: student_id, subject_id, mid_semester_marks

End Sem Marks: student_id, subject_id, end_semester_marks

Student: student_name, student_id

Subject: subject_id, subject_name

Find out the correlation coefficient between the mid sem marks and final semester marks of students from the students table. Do the same subject wise. Also create a new relation rank and calculate the rank of the students based on mid sem and end semester marks.

Q20. [Type: Data Manipulation, CONTRIBUTOR: 20BM6JP40]

The emails of students of a course are stored in a column in the format

abcdef2022@email.iimcal.ac.in

In the schema

Student = <email: string, name:string>

The first text string can be of any length, followed by 4 characters of the passing year of the student, followed by the domain address '@email.iimcal.ac.in' which is same for all the students

Write an SQL query to add a column with only the year of passing of each student which is to be extracted from the email address.

Q21. [Type: Data Manipulation, CONTRIBUTOR: 20BM6JP34]

The stats of players are given in the following table.

Football_Players = <Player_id : string, Name : string, Age : integer, Contract_Amount : integer, Career_rating : integer, Nationality : string, Current_club : string, Position : string, Weight_kg : integer, Height_cm : integer >

```
create table Football_Players (  
    Player_id CHAR(20) NOT NULL PRIMARY KEY,  
    Name CHAR(25),  
    Age INT(4),  
    Contract_Amount INT NOT NULL(10),  
    Career_rating INT NOT NULL(4),  
    Nationality CHAR(25),  
    Current_club CHAR(64),  
    Position CHAR(20),  
    Weight_kg INT(4),  
    Height_cms INT(4)  
);
```

Q1) Find the highest paid player country wise.

Q2) Find the maximum BMI (Body mass index) equivalent.(use the formula Weight/Height for simplicity)

Q3) Find the highest paid player and the lowest paid player, also find the difference between their ratings.

Q22. [Type: Data Manipulation, CONTRIBUTOR: 20BM6JP03]

Consider below the schemas in the database of a million dollar start-up of the PGDBA alumni of batch-2. The HR head of the company wants to know a few important things about the situation in the company and he calls you for doing some analysis as he knows that you have done the FDS course in PGDBA. Help him analyze the situation.

EMPLOYEE PERSONAL DETAILS : < EMPLOYEE_ID : integer ,FIRST_NAME : string , LAST_NAME : string , EMAIL : string >

EMPLOYEE PROFESSIONAL DETAILS : < EMPLOYEE_ID : integer , JOB_TITLE : string , SALARY : integer , MANAGER_ID : integer, DEPARTMENT_ID : integer>

The questions that he needs the answers to , at priority basis, are mentioned below:

- 1) Display the first and last name, salary, and department ID for those employees who earn less than the average salary, and also work at the department where the employee Malay (first name) Bhattacharaya (last name) is working.
- 2) Write a query to display the employee ID and name (first name and last name) of all employees who works in such a department where the employee(s) are working whose first name contains the expression 'ALAY'.
- 3) Write a query to display the employee number, name(first name and last name) and job title for all employees whose salary is smaller than any salary of those employees whose job title is Senior_Consultant. Exclude Job title Senior_Consultant.
- 4) Write a query to display the employee id, name (first name and last name), salary and the SalaryStatus column with a title HIGH and LOW respectively for those employees whose salary is more than and less than the average salary of all employees.
- 5) Write a query which is looking for the names of all employees whose salary is greater than 50% of their department's total salary bill and who report to Amitava Banarjee.

Q23. [Type: Data Manipulation, CONTRIBUTOR: 20BM6JP50]

The PGDBA committee at ISI decided to record the students' entry and exit time for each online class in a table. Using this data professors decided to measure the attendance and engagement level of each student in this online program.

The **Student_Log** table has 5 columns Student_Name, Roll_Number, Subject, Action(enter/Exit), Created at (time at which student enters or leaves a class)

Student_Log: <Roll_Number : integer, Student_Name : string, Subject: string, Action : string, Created_at : DEFAULT TIMESTAMP>

The primary key is underlined in the student_log table. Note that the first entry in the action column for each student is 'enter' and every 'enter' is succeeded by an 'exit'.

Create table Student_Log

```
(
    Roll_Number int(4),
    Student_Name char(40),
    Subject char(40),
    Action char(40),
    Created_at datetime
```

```
);
```

Sample Data in the table

Roll_Number	Student_Name	Subject	Action	Created_at
20BM6JP50	Kritica Meena	Funds	enter	10-01-2021 12:00:05
20BM6JP50	Kritica Meena	Funds	exit	20-04-2021 12:40:11
20BM6JP50	Kritica Meena	SSD	enter	10-01-2021 16:10:34

Q1: Subject and week wise ,Avg time spent in the class at student level.

Q2: Find the no of students inside Funds class at '2020-02-09 18:30:00' (Assuming class started at '2020-02-09 18:00:00')

Q3: Prof wants to reward students with highest attendance (in terms of hrs spent) ,Measure total amount of time spent by each student in online classes since the day they started (first class)

Q4: Measure amount of hrs spent by each student on last day (9th Feb)

Q24. [Type: **TRIGGERS** , CONTRIBUTOR: **20BM6JP48**]

Consider the following Relational Schema:

Employee (*emp_id*: integer, *emp_name*: string, *age*: integer, *salary*: real)

Works (*emp_id*: integer, *dept_id*: integer, *time*: integer)

Department (*dept_id*: integer, *budget*: real, *manager_id*: integer)

Here, time = percentage of time a given employee works.

Write SQL triggers to ensure each of the following requirements.

- 1) Employees must make a minimum salary of Rs 5,00,000.
- 2) A manager must always have a higher salary than any employee that he or she manages.
- 3) Whenever an employee is given a raise, the manager's salary must be increased to 1.2 times of the raise.
- 4) If (3) holds then, whenever an employee is given a raise, the department's budget must be increased to be greater than the sum of salaries of all employees in the department.

Q25. [Type: **Data Manipulation**, CONTRIBUTOR: **20BM6JP23**]

Pivot the Occupation column in OCCUPATIONS so that each Name is sorted alphabetically and displayed underneath its corresponding Occupation. The output column headers should be Doctor, Professor, Singer, and Actor, respectively.

Note: Print NULL when there are no more names corresponding to an occupation.

Input Format

The OCCUPATIONS table is described as follow:

Column	Type
Name	String
Occupation	String

Occupation will only contain one of the following values: Doctor, Professor, Singer or Actor.

Note:

The first column is an alphabetically ordered list of Doctor names.

The second column is an alphabetically ordered list of Professor names.

The third column is an alphabetically ordered list of Singer names.

The fourth column is an alphabetically ordered list of Actor names.

The empty cell data for columns with less than the maximum number of names per occupation (in this case, the Professor and Actor columns) are filled with NULL values.

Q26. [Type: **Data Manipulation**, CONTRIBUTOR: **20BM6JP42**]

Suppose Mr. Das has a business and he wants to analyse the sales data to target the valuable customers and to provide them the best possible after sales services to retain them in future.

ORDER = <order_id: integer, purchase_amount: integer, cust_id: string>

CUSTOMER= <cust_id: string, cust_name: string, city: string>

Q1. Find customer names and their city who have purchased a total above INR 5,00,000 worth of product?

Q2. Find the order ids and customer names who are from Kolkata and Mumbai?

Q3. Find the customer ids who have placed more than 500 orders?

Q4. Find the name and city of top 10 customers in terms of order value?

Q27.[TYPE: **Data Manipulation**, CONTRIBUTOR: **20BM6JP15**]

Consider the following schema of a bank for storing customer information.

BANK = <customer_id:string,name:string,customer_relationship_value:int,country:string>

The field customer_id is the primary key.

Write a single sql query which classifies all the customers of the bank into categories of "High Net Worth Client", "Medium Net Worth Client" and "Low Net Worth Client" based on the criteria of customer_relationship_value.

Criteria for classification:

customer_relationship_value greater than 100,00,00 are "High Net Worth Client"

customer_relationship_value between 100,000 and 100,00,00 are "Medium Net Worth Client"

customer_relationship_value less than 100,000 are "Low Net Worth Client"

The query should return name, customer_relationship_value and customer_classification for all the customers situated in India.

Example: For the following table and data, the required output should be like the one mentioned below:

```
create table BANK(  
customer_id int primary key,  
name varchar(50),  
customer_relationship_value bigint,  
country varchar(50)  
);
```

```

insert into BANK values(1,"Ambani",88000000000,"India");
insert into BANK values(2,"Adani",25000000000,"India");
insert into BANK values(3,"Nilay",120000,"India");
insert into BANK values(4,"Shubham",120,"India");
insert into BANK values(5,"Bezos",193000000000,"US");
insert into BANK values(6,"Musk",150000000000,"US");

```

The result of the required query should be like:

	name	customer_relationship_value	customer_classification
▶	Ambani	88000000000	High Net Worth Client
	Adani	25000000000	High Net Worth Client
	Nilay	120000	Medium Net Worth Client
	Shubham	120	Low Net Worth Client

[Type: Data Manipulation, CONTRIBUTOR: 20BM6JP02]

Q28. Consider the following schema pertaining to sales of Lenovo laptops across India.

Sales = <purchase_id:int, model_id:int,dop:date>

purchase_id is primary key and model_id is foreign key referencing Laptop table

Schema of Laptop table

Laptop=<model_id=int, category:string,cost:int>

Category is one of these three – Economic, Premium and Gaming

Write an SQL Query to find out:

1. Which category generated maximum revenue in 2020 and how much is the revenue?
2. List top 5 models according to revenue generated till now?
3. How many of laptops were sold from each of the category of market in the year 2019 and 2020

[Type: Data Manipulation, CONTRIBUTOR: 20BM6JP28]

Q29. The Constituency wise Lock Sabha Election 2019 Result shown in Below Table.

Election 2019 Result <Serial No: integer, Constituency: string, State: string, Voting turnout: float, Elected MP: string, Party of Elected MP: string, Runner-up : string, Runner-up Party :string, Win margin by votes: integer, Win margin by percentage: float>

```
create table Election_2019_Result (
```

```
    Serial_No INT(20) NOT NULL PRIMARY KEY,
```

```
    Constituency CHAR(25) NOT NULL,
```



```
State CHAR(25),  
Voting_turnout FLOAT(8),  
Elected_MP CHAR(25),  
Party_of_Elected_MP CHAR(25),  
Runner-up CHAR(25),  
Party_of_Runner-up CHAR(25),  
Win margin INT(16),  
Win_margin_by_percentage FLOAT(8),  
);
```

Q1) Find the State with Maximum Number of Constituency.

Q2) Find the name of Elected MP who wins with maximum margin percentage wise.

Q3) Find the constituency with maximum voting turnout.

[TYPE: Data Manipulation, CONTRIBUTOR: 20BM6JP59]

Q30. Consider the following schema.

TRANSACTIONS = <transaction_id, user_id, transaction_date, product_id, quantity>

1) Write a SQL query to find the number of users who purchased the same products.

2) Write a SQL query to find the number of users who purchased products on multiple days.

(Note that a given user can purchase multiple products on a single day).

Q31. [Type: **Data Manipulation**, CONTRIBUTOR: **20BM6JP43**]

Club selection vacancies for 3 year are posted on a website. Applicants can view the ad and fill form to apply for the position they feel they best suit for. The candidate data is stored in the given format. The table Job_details table contains information about the job details like the club contact person and club name and club requirement descriptions

Table **Applicant_Details** (Application_Id: Job_Id: String, Qualification: String, Candidate_Name: String, Aadhaar_Id: String, Enrollment_Year: integer, Age: integer, Gender: String)

Table **Job_Details** (Job_Id: String, Job_Description - String, Interviewer - String, Club_Name: String)

create table Job_Details(Job_Id: varchar(10) not null, Job_Description – varchar(100) not null, Interviewer - String, Club_Owner: varchar(100) not null, primary key (Job_Id));

insert into Job_Details values("0000000002", "Cultural Fest Publicity Head", "Rajni", "Cultural Fest");

insert into Job_Details values("0000000001", "Cultural Fest Event manager", "Ramesh", "Cultural Fest");

insert into Job_Details values("0000000003", "Technology Fest Event manager", "Ritesh", "Technology Fest");

insert into Job_Details values("0000000004", "Technology Fest Publicity Head", "Ramesh", "Cultural Fest");

1. Ensure the data while inserting has the following limitations from the Job Owner, Enrollment_Year >2016 and Enrollment_Year <2020, show error message for "Invalid Applicant".
2. Find the number of applicants for each job
3. Allocate the candidate data to the Interviewer.
4. Find the candidate who applied for more than two jobs.

Q32. [TYPE: Data Manipulation, CONTRIBUTOR: 20BM6JP36]

End semester examinations were conducted in ABC institute of technology. We have tables consisting of information regarding the student, information related to the room (classrooms where exams were held) and data corresponding to the center assignment. Every room has a unique No., the Department and Floor to which it belongs. The student table consists of the student ID, name, Department to which he/she belongs to. The center assignment details contain the information - S.No, student ID, Room No., Department of the room and Exam name.

STUDENT = <student_id: integer, student_name: string, department: string>

ROOM = <room_no: integer, department: string, floor: integer>

CENTER = <s_no: integer, student_id: integer, room_no: integer, department: string, exam_name: string>

1. Write an SQL query to find the names of students who belong to the 'Mechanical Engineering' department.
2. Write an SQL query to find out the names of the exams that were conducted in the 'Electrical Engineering' Department.
3. Write an SQL query to produce the names of all the students who wrote 'Database systems' examination in the 'Computer Science' Department.

Ans. The database schema in SQL:

```
create table student (  
    student_id int(10) not null,  
    student_name char(30),  
    department char(30),  
    primary key (student_id)  
);  
  
create table room (  
    room_no int(10) not null,  
    department char(30),  
    floor int(3),  
    primary key (room_no)  
);  
  
create table center (  
    s_no int(10) not null,  
    student_id int(10),  
    room_no int(10),  
    department char(30),  
    exam_name char(30),  
    primary key (s_no),  
    foreign key (student_id) references student(student_id),  
    foreign key (room_no) references room(room_no)  
);
```

Q33. [TYPE: Data Manipulation, CONTRIBUTOR:20BM6JP07]

You are the head of the accounts division of the company XYZ. It has been found that recently there has been an increase in travel reimbursement and you need to prepare a report summarizing key points. Following tables are available with you

Employee = <Emp_ID: integer, Emp_Name: string, Department: string>

Here Emp_ID is the primary key

Approvals = <Request_ID: integer, Emp_ID: int, Amount: int, Date: date>

Here Request_ID is the primary key

Emp_ID is foreign key referencing Employee table

Date contains date at which the amount was reimbursed for the Request_ID

None of the attributes should be null and following key points should be extracted

1. What is the total amount reimbursed for the year 2018, 2019 and 2020?
2. In the current year which department has availed the maximum amount of reimbursement and what is the amount?
3. List out Emp_ID, Name and Department of 5 employees who got most reimbursement requests approved in the last month.

Q34. [TYPE: Data Manipulation, CONTRIBUTOR:20BM6JP18]

Trevor is a food vlog maker who travels all over the world and tastes various cuisines. When he first came to India he fell in love with Indian food. Since then he has eaten food at thousands of outlets over the past 15 years and he has rated them on a scale of 1 to 10. He has also recorded the level of spiciness on a scale of 1-10 along with other things like Veg/Non-Veg etc. Ramesh who watches Trevor's videos decides to try out the best cuisines in Calcutta. Unfortunately, he cannot tolerate spicy food, so he decided to avoid food whose spice levels are above 6. He does not want to waste money on bland food so he decided to have food whose ratings are above 7. Since Ramesh was a vegetarian he would not touch Non-Veg food.

FOOD=<ID: string, name: string, type: string, cuisine: string, outlet: string, address: string, city: string, food_rating: integer, spice_level: integer, price: float>

Write queries for the following

- a) What constraints did Trevor have to implement while defining the FOOD table?
- b) Ramesh wants food_rating above 8, spice_level below 7 and type of food should be Veg. He wants to search only in Calcutta. Write a query that returns the name of the food, name of outlet, its address and the food rating.
- c) Write a query so that he can find food of South Indian cuisine whose price is below Rs. 1000 in addition to the above conditions.

d) Suppose Ramesh wants to compare biryani ratings of Calcutta to those of other cities. Find out the average ratings of all 'biryani' from Calcutta, Delhi, Lucknow and Hyderabad. Present the city names in the descending order of their average ratings.

e) Find out the average ratings of Veg and Non-Veg food in the city of Calcutta and compare.

f) Ramesh wants to have an idea whether there is any association between food rating and spice level. Compute the average food ratings of food whose spice level is 10. Similarly do it for spice levels of 9,8,...,1 respectively. Arrange spice levels according to the ascending order of their average food ratings. Comment.

g) Ramesh wants to see the relation between price and food rating. So he divides the price into bins: (0-250), (250-500), (500-1000), (1000-3000), (>3000). Compute the average food ratings for these bins and comment.

h) Which cuisine has the highest average food ratings, South Indian, Punjabi, Chinese or Italian?

Q35. [TYPE: Data Manipulation, CONTRIBUTOR:20BM6JP49]

Consider an online grocery shopping company whose database contains:

grocery_data = <order_id , user_id, city_id , status , transaction_date, amount >

The following metrics need to be calculated :

1. Days with the highest failure rates in transactions (Failure Rate = No. of Failed Transactions / Total Transactions)
2. Top 3 cities in terms of number of successful orders in 2020
3. Top 5 cities with the largest userbase in 2020
4. No. of Days in the month of March 2020 where no transaction happened
5. Identify the city with the least value of total orders in 2020.

Q36. [TYPE: Data Manipulation, CONTRIBUTOR:20BM6JP08]

A Professor named Virus conducts two exams for the students in his class every year and selects top three students in the class (based on total marks) and rewards them with prizes.

Exam-marks data = <Roll-no, Exam1_marks, Exam2_marks >

Roll-no is unique for each student(Primary key). Help Professor Virus in selecting the top three students in the class for rewarding them.

Q37. [TYPE: Data Manipulation, CONTRIBUTOR:20BM6JP27]

Student (*snum*: integer, *sname*: string, *major*: string, *level*: string, *age*: integer)

Class (*name*: string, *meets at*: time, *room*: string, *fid*: integer)

Enrolled (*snum*: integer, *cname*: string)

Faculty (*fid*: **integer**, *fname*: **string**, *deptid*: **integer**)

1. Express each of the following integrity constraints in SQL unless it is implied by the primary and foreign key constraint; if so, explain how it is implied. If the constraint cannot be expressed in SQL, say so. For each constraint, state what operations (inserts, deletes, and updates on specific relations) must be monitored to enforce the constraint.

- (a) Every class has a minimum enrollment of 5 students and a maximum enrollment of 30 students.
- (b) At least one class meets in each room.
- (c) Every faculty member must teach at least two courses.
- (d) Only faculty in the department with deptid=33 teach more than three courses.
- (e) Every student must be enrolled in the course called Math101.
- (f) The room in which the earliest scheduled class (i.e., the class with the smallest meets at value) meets should not be the same as the room in which the latest scheduled class meets.
- (g) Two classes cannot meet in the same room at the same time.
- (h) The department with the most faculty members must have fewer than twice the number of faculty members in the department with the fewest faculty members.
- (i) No department can have more than 10 faculty members.
- (j) A student cannot add more than two courses at a time (i.e., in a single update).
- (k) The number of CS majors must be more than the number of Math majors.
- (l) The number of distinct courses in which CS majors are enrolled is greater than the number of distinct courses in which Math majors are enrolled.
- (m) The total enrolment in courses taught by faculty in the department with deptid=33 is greater than the number of Math majors.
- (n) There must be at least one CS major if there are any students whatsoever.
- (o) Faculty members from different departments cannot teach in the same room.

Q38. [TYPE: Data Manipulation, CONTRIBUTOR:20BM6JP12]

Suppose you are working as a Risk Analyst for an Investment Bank. Your company makes an investment in a different portfolio. Create a database of the portfolio management for the investment your firm has made so that you can calculate the daily risk no from changes in daily market price.

Portfolio: <ISIN: string, Company Name: String, Sector: String, Invested Amount: Integer>

Price:<ISIN: String, Current Price: Integer, Previous Price: Integer>

- 1) From the Portfolio table name the company where the invested amount maximum.
- 2) Group the sector and find the name of the sector with maximum investment.
- 3) For a particular day find the name of the company that has a maximum % increase in the current price.

Q39. [TYPE: Data Manipulation, CONTRIBUTOR: **20BM6JP61**]

Consider the insurance database, where the primary keys are underlined.

PERSON (pid: integer, name: string, address: string)

CAR (license: integer, year: integer, model: string)

ACCIDENT (date: date, driver: string, damage-amount: integer)

OWNS (pid: integer, license: integer)

LOG (license: integer, date: date, driver: string)

Construct the following SQL queries for this relational database:

- (i) Find the total number of people whose cars were involved in accidents in 2020.
- (ii) Find the number of accidents in which the cars belonging to "John Smith" were involved.
- (iii) Add a new accident record for the Toyota belonging to "Jones"

40. [TYPE: Data Manipulation, Contributor: **20BM6JP11**]

Consider the following database schema:

CITY (ID: Number, Name: Varchar2(20), CountryCode: Varchar2(3), District: Varchar2(20), Population: Number)
COUNTRY (CODE: Varchar2(3), Name: Varchar2(50), Continent: Varchar2(25), SurfaceArea: Number, Population: Number, GNP: Number, GovernmentForm: Varchar2(50), Capital: Varchar2(20))

Construct the following SQL queries for this relational database:

- (i) Given the CITY and COUNTRY tables, query the sum of the populations of all cities where the CONTINENT is 'Asia'.
- (ii) Given the CITY and COUNTRY tables, query the names of all cities where the CONTINENT is 'Africa'.
- (iii) Given the CITY and COUNTRY tables, query the names of all the continents (COUNTRY.Continent) and their respective average city populations (CITY.Population) rounded down to the nearest integer.

Q 41. [Type: Data Manipulation, CONTRIBUTOR: **20BM6JP45**]

Ravi Shastri, head Indian coach after having four shots of Tequila in a party decided to quit his career in cricket coaching and thought to become a football manager in ISL and approached Sachin Tendulkar, owner of Kerala Blasters for a chance. Recognizing Ravi's selection skill, Sachin offered him job on two conditions. Number one is to stop drinking and Second Condition is to start his manager career by scouting youngsters all over the world for the upcoming ISL. Shastri decided to check for Under 19 superstars who excelled in FIFA U-19 world cup. He asked stats from the team analyst exclusively about the under 19 players who signed for ISL auctions.

Analyst summarised the whole data in 3 tables.

1. Availabiltychart_2020 = <Country: string , BusyMonth: integer >
2. U19_PlayerStats_2019=<Playerid:integer,PlayerName:string,Country:string, Goalsavesuccess:integer, Assist:integer, Goalscored: Integer, Tacklesuccess: Integer, Passsuccess:integer, Yellowcard: integer, Redcard:integer, Fouls: integer>
3. ISL_Auction_List =<Auctionid:integer, Playerid:integer, Pricetagincrore:integer>

Ravi who is expert in SQL has to find out players with his matching criteria by querying the table given by the analyst.

1. Player who is available for this season and must be an aggressive (Fouls and cards can be ignored) first rated (Atleast 12 goals)striker.
2. Player who is available/not available for this season and midfielder (Passsuccess should be greater than 75)without any red card.
3. Goalkeeper with maximum Goalsavesuccess and should be available for this season.

Ans:

The database schema in SQL:

```
create table Availabiltychart_2020
( Country VARCHAR(30),
  BusyMonth TINYINT NOT NULL,
  PRIMARY KEY (Country) );
```

```
create table U19_PlayerStats_2019
(Playerid TINYINT,
  PlayerName VARCHAR(50) NOT NULL,
  Country VARCHAR(30),
  Goalsavesuccess TINYINT NOT NULL,
  Assist TINYINT NOT NULL,
  Goalscored TINYINT NOT NULL,
  Tacklesuccess TINYINT NOT NULL,
  Passsuccess TINYINT NOT NULL,
  Yellowcard TINYINT NOT NULL,
  Redcard TINYINT NOT NULL,
  Fouls TINYINT NOT NULL,
  PRIMARY KEY (Playerid),
  FOREIGN KEY(Country) );
```

```
create table ISL_Auction_List
(Auctionid TINYINT ,
  Playerid TINYINT NOT NULL,
  Pricetagincrore TINYINT NOT NULL
  PRIMARY KEY (Auctionid)
```


FOREIGN KEY (Playerid));

Q 42. [Type: Data Manipulation, CONTRIBUTOR: 20BM6JP13]

Consider the two tables below:

Employee table : (Employee_id:string,Employee_name:string,RecruitedBy:string)

Recruiter table: (Recruiter_id:string,Recruiter_name:string)

- a) Write a query that retrieves all employees recruited by a recruiter named X.
- b) Write a query to retrieve all employees that were not recruited by any recruiter

Q 43. [Type: **Data Manipulation**, CONTRIBUTOR: **20BM6JP44**]

Consider the two tables below:

COMPANY table : (company_id:int, company_name:string, Company_location:string)

ITEMS_PRODUCED table: (item_id:int, item_name:string, item_unit:string, company_id:int)

The table COMPANY has company_id as primary key and ITEMS_PRODUCED has item_id as primary key

- a) Write a query that retrieves to retrieve all the unique company names, their location and unit name from which it is produced.
- b) Write a query to retrieve the number of companies which produces the item 'salty chips'.
- c) Write a query to retrieve the number of distinct companies located in each location.

Q 44. [Type: Data Manipulation, CONTRIBUTOR: **20BM6JP56**]

Consider the following relational schema. An employee can work in more than one department; the pct time field of the Works relation shows the percentage of time that a given employee works in a given department. Emp(eid: integer, ename: string, age: integer, salary: real) Works (eid: integer, did: integer, pct time: integer) Dept(did: integer, budget: real, managerid: integer) Write SQL-92 integrity constraints (domain, key, foreign key, or CHECK constraints; or assertions) or SQL:1999 triggers to ensure each of the following requirements, considered independently.

1. Employees must make a minimum salary of \$1000.
2. Every manager must also be an employee.
3. The total percentage of all appointments for an employee must be under 100%.
4. A manager must always have a higher salary than any employee that he or she manages.

5. Whenever an employee is given a raise, the manager's salary must be increased to be at least as much.
6. Whenever an employee is given a raise, the manager's salary must be increased to be at least as much. Further, whenever an employee is given a raise, the department's budget must be increased to be greater than the sum of salaries of all employees in the department.

Q 45. [Type: Data Manipulation, CONTRIBUTOR: **20BM6JP32**]

Konoha academy is the institute present in 'Hidden Leaf Village' in the country 'Land of Fire' which grooms the upcoming ninja/shinobi to fight evil. Each shinobi is given a unique shinobi id upon his/her entry to the academy. There are 4 departments (Genjutsu, Ninjutsu, Taijutsu, Senjutsu) which teach the shinobis to control these 4 major powers. Among each department there are a fixed number of courses. Each shinobi can take as many courses from any department.

Consider the following database schema:

SHINOBI : (shinobi_id int(4), shinobi_name varchar (40))

COURSE: (course_id int (4), course_name varchar(20), department_name varchar(40), course_strength int(4))

TAKES: (course_id int(4), shinobi_id int(4), grade char(1))

- Find the set of all the courses offered in the "Genjutsu" department.
- Find the set of all the courses that a shinobi with the name "Itachi Uchiha" has taken.
- Find all shinobis who have taken all courses offered in the "Ninjutsu" department.

Q 46. [TYPE: **Data Manipulation**, CONTRIBUTOR: **20BM6JP09**]

Schema:

member <member_id : int(4), name : char(10), age : int(4)>

book <isbn : char(10), title : char(10), authors : char(10), publisher : char(10)>

borrowed <member_id : int(4), isbn : char(10), date : date>

Questions:

- Print the names of members who have borrowed any book published by "JK Rowling".
- Print the names of members who have borrowed all the books published by "JK Rowling".
- For each publisher, print the names of members who have borrowed more than 3 books of that publisher.

4. Print average number of books borrowed per member.

(Take into account that if a member does not borrow any books, then that member does not appear in the *borrowed* relation at all)

Q47: [TYPE: **Data Manipulation**, CONTRIBUTOR: **20BM6JP30**]

An ardent music lover developed two Relations(Movies and Songs)

Schema:

Movies = <Year: int, Movie: char, Actor: char, Actress: char> Movie is the Primary Key

Songs = <Song_id: int, Singer: char, Music_Director: char, Song: char, Movie: char, Language: char> Song_id is the Primary Key and Movie is the foreign key to Movies Relation.

Questions:

1. For a given Actor(ex: Mahesh Babu), Actress(ex: Kajal) and Music_Director(ex: Anirudh) combination, write a query to output the song name which starts with a specific string of characters (ex: "pyaar").
2. For given period between two different years(ex: between 2016 and 2020), write a query to get the name of Music_Director who composed second most number of songs.
3. Write a query to get the name of music director who composed songs in most number of languages.

Q48. [TYPE: **Data Manipulation**, CONTRIBUTOR: **20BM6JP10**]

Consider the following relational schema. An employee can work in more than one department; the pct time field of the Works relation shows the percentage of time that a given employee works in a given department –

Emp(eid: integer, ename: string, age: integer, salary: real)

Works(eid: integer, did: integer)

Dept(did: integer, dname: string, budget: real, managerid: integer)

Write the following queries in SQL:

1. Print the names and ages of each employee who works in both the Hardware department and the Software department.
2. Print the name of each employee whose salary exceeds the budget of all of the departments that he or she works in.
3. Find the managerids of managers who control the largest amounts.

Q49. [TYPE: **Data Manipulation**, CONTRIBUTOR:20BM6JP24]

Consider the following tables. Person (Aadhar_no:integer, Name:String, Address:String, age:int, No. of Ailments:int)

Insurance (Aadhar_no:int, Company:String, Plan:String, Sum Insured:int, Years:int)

- (i) Find the maximum number of Ailments grouped by Age
- (ii) Delete tuples from the table Person where no. of Ailments is 0.
- (iii) Add your details in the Person and Insurance Database
- (iv) Find the average sum insured per year for all the people together.

Q 50. [TYPE: **Data Manipulation**, CONTRIBUTOR: 20BM6JP55] Consider the following schema representing the Delhi Metro System.

METRO_STOPS	= < <u>S_ID</u> : Integer, Name : Char(30)>
METRO_LINES	= < <u>L_ID</u> : Integer, Colour: Char(30)>
METRO_ROUTE	= < <u>L_ID</u> : Integer, <u>Stop_Num</u> : Integer, Stop: Integer>

The primary keys are underlined in the corresponding tables.

S_ID is the Primary Key for METRO_STOPS

L_ID is the Primary Key for METRO_LINES

L_ID, Stop_Num together form the Primary Key for METRO_ROUTE

Stop is a Foreign Key in METRO_ROUTE referencing METRO_STOPS via S_ID

- (i) Write an SQL Query to List all the Stops (Names) that you can visit from “Hauz Khas” without changing the Metro Line.
- (ii) Find the travel routes from “Hauz Khas” to “Chandni Chowk” involving two Metro Lines. Show the Metro Line Colour for the first Metro, the Name of the Metro Stop for the Transfer, and the Metro Line Colour for the second Metro.

Q51. [TYPE: **Data Manipulation**, CONTRIBUTOR:20BM6JP21]

Consider the following relational schema regarding the Prime Ministers of India.

PM(Serial No.: int, Name: String,Tenure: String, Party: String, Gender: String)

Here tenure is a string which is of the form DD/MM/YYYY - DD/MM/YYYY.

- (i) Find the PM who served the longest tenure. (If the same PM served multiple times, consider the sum of all tenures as his/her tenure).
- (ii) Find the total number of female PMs who served India.
- (iii) Find the party with maximum number of PMs and find the party with maximum tenure of PMs.
- (iv) Find the average tenure of the PM until last completed tenure. (2014 - 2019 can be considered as the last tenure).

Q52. [TYPE: **Data Manipulation**, CONTRIBUTOR:20BM6JP58]

A cricket Academy trains aspiring cricketers in different departments (ex: Batting ,Bowling ,Allrounder, wicket keeper) and assigns individual Trainer for each player. Each Player can be trained under one coach. A Trainer can take many players.

There are 3 tables

Player: PlayerID (Int), Name(string),DOB(date),Department(String), TrainerID(int). Primary key(Player ID)

Trainer: TrainerID (int), Name(String),DOB(date),Department(String) Primary Key(Trainer ID)

PlayerStats:PlayerID(int),MatchesPlayed(int),RunsScored(int),Hundreds(int),Fifties(int),Wickets taken(int),Catches Taken(int) . PrimaryKey(Player ID)

Questions :

- 1.Who is the Trainer of the player who scored most runs?
- 2.Who are the players who scored Runs above 2000 and took at least 50 wickets.
3. Who is the trainer who trains the most Number of students?
4. who is the player taken most wickets trained under Trainer #####?
- 5.The youngest player in the Academy is trained under?

Q53. [TYPE: **Data Manipulation**, Contributor: 20BM6JP51]

Consider the following schema on customer account information stored in the database of any Bank.

Customer Data: (Name: String, DOB: Date, Type: String, Address: String, CustId: Int(Primary Key))

KYC Data: (PAN: String, Aadhaar: Int, Voter ID: String, Other Documents: String)

Questions:

1. Write a SQL query to find the name of the customer from KYC data and vice versa.
2. Write a SQL query to find the custids whose KYC data are not seeded properly. Further, within the same query write a command to update if the KYC data are available but not mapped to that particular custid.
3. Write a SQL query to check for any duplicate entries in both customer data and KYC data and also write a command to delete the redundant data if duplicates are found.

Q54. [TYPE: **Data Definition and Data Manipulation**, Contributor: **20BM6JP01**]

A HR researcher wants to calculate correlation between pay of an employee and number of companies in which he/she has worked. All the employees in the study have same number of years of work experience.

The researcher has a relation R with 2000 tuples where

$R = \langle \text{uniqueID: integer, name: string, companies: string, salary: integer} \rangle$

Details of attributes:

uniqueID: Unique ID given to every employee

name: Name of Employee

salary: Current pay of the employee in rupees

companies: This is a Multivalued attribute and it contains a list of companies that an employee worked in. Each company is separated by a Comma. For example, Suresh worked in Infosys, TCS, Accenture. Ramesh worked in HCL. They are listed in the table as follows

uniqueID	name	companies	salary
12901	Suresh	Infosys,TCS, Accenture	100000
23012	Ramesh	HCL	80000

Note: There is no attribute in the above relation containing the number of companies an employee worked in.

For his study, Analyst wants to categorize employees into 4 categories – **A, B, C, D**

A consists of employees who worked only in a single company.

B consists of employees who worked in 2 or 3 companies

C consists of employees who worked in 4-6 companies

D consists of employees who worked in more than 6 companies

Question: Write a SQL code which outputs a relation containing 3 attributes – Category, Number of employees in the category, Average salary

Q55. [TYPE: Data Manipulation, CONTRIBUTOR: 19BM6JP41]

Suppose you have a database consisting of a single table with student name, roll number, section name, and marks with the following schema. CLASS= The primary key is underlined in the above schema. Write the following queries in SQL.

(i) Fetch the section name and number of students in that section.

(ii) Fetch the section name and total marks obtained by the students altogether in that section.

Q.56. [TYPE: **Data Definition and Manipulation**, CONTRIBUTOR: **20BM6JP26**]

In the wake of the Digital India campaign, your city police station wanted to maintain all the records of crime digitally. As an independent consultant, you have been invited to set up a database for the local police officials. The details of the schema are as follows:

OFFICERS = <id : integer, *name*: string, *age*: integer, *mobile_number*: integer, *department_id*: string, *zone_id*: integer>

DEPARTMENT = <id: string, *name*: string>

CRIMINALS = <id : string, *team* : string, *name* : string, *age*: integer, *crime_details*: string, *date*: date, *zone_id* : integer, *officer_id*: integer>

ZONE = <id: integer, *name*: string, *city*: string, *pin*: string, *population*: integer>

PAROLE APPLICATIONS = <id: integer, *application_date*: date, *criminal_id*: string, *status*: string, *expiry_date*: date>

Create all the above tables by defining suitable primary keys and foreign keys. After creating the tables, let's assume that it has some data filled as expected. Now, answer the following questions:

1. The police station wants to give rewards to most dedicated police officials. Find out the name of the officer who has caught the maximum number of criminals in the year 2021.

2. Expiry_date columns in PAROLE APPLICATIONS indicate the time by which the criminal needs to return to the police station. Find out all the criminals that need to report to the police station in the current month so that police officers can easily keep track of the defaulters.
3. Government needs to redistribute the number of police officers in each zone based on the rate of increment of crime. To help the government, find out the zone name where the number of criminals (per 100 people) in 2020 as compared to 2019, have increased maximum.

Q.57. [TYPE: Data Definition and Manipulation, CONTRIBUTOR: 20BM6JP16]

Consider the following schema

Emp(eid: integer, ename: string, age: integer, salary: real)

Works(eid: integer, did: integer, pct_time: integer)

Dept(did: integer, dname: string, budget: real, managerid: integer)

An employee can work in more than one department; the pct_time field of the Works relation shows the percentage of time that a given employee works in a given department.

1. Find the enames of managers who manage only departments with budgets larger than Rs10 Lakh, but at least one department with a budget less than Rs50 Lakh.
2. Find the enames of managers who manage the departments with the largest budgets.
3. Find the managerid(s) of managers who manage only departments with budgets greater than Rs10 Lakh.
4. Print the names and ages of each employee who works in both the Clerical department and Warehouse department.

Q. 58. [TYPE: Data Manipulation, CONTRIBUTOR: 20BM6JP25]

10 college students, all from multiple backgrounds, want to be self independent. So they joined a part time job in 4 different companies. two schemas are given below containing students' job details and their academic details.

JOB = <roll: integer, student name: string, company: string, salary: int, age: int >

STUDENT = <roll: integer, student name: string, department: string, state: string >

Write the SQL queries for following operations.

- i) Find the name of the student whose age is above 19 and whose salary is less than Rs. 5000
- ii) Find the average salary offered by each unique company
- iii) Find the age of the student who is from Physics department and working in 'ABC Corporation'
- iv) Find all those students who are older than 18 years and are from West Bengal and whose salary is more than Rs. 3500

Q.59 [TYPE: Data Manipulation, CONTRIBUTOR: 20BM6JP41]

Consider the following schema representing details about the online customers who place orders in different outlets of Bawarchi Restaurant.

CUSTOMER = <name: string, age: integer, gender: string, address: string>

ORDERS = <name: string, biryani: string>

DELIVERS = <outlet: string, biryani: string, cost: real>

The keys corresponding to each relation are underlined. Write the following queries in SQL.

- (i) Find the names of all males who live at AU,Vizag and orders both “Mutton keema” and “Chicken Tandoori” biryani.
- (ii) Find the names of all customers who have never placed an order.
- (iii) Find all the outlets that delivers at least one biryani that can be ordered by “Varma” in less than 300.

Q.60 [TYPE: Data Definition/Manipulation, CONTRIBUTOR: 20BM6JP52]

Consider the following schema representing some of the data storage and retrieval about a soccer tournament based on FIFA WORLD CUP 2018.

a) **soccer_country** = < country_id: integer, country_abbr: string, country_name: string>

b) **match_details** = < match_no: integer, play_stage: char, team_id: integer, win_lose: char, decided_by: char, goal_score: integer, penalty_score: integer, ass_ref: integer, player_gk: integer>

Answer the following:

1. Write a query in SQL to find the teams played the first match of FIFA world cup 2018.
2. Write a query in SQL to find the winner of FIFA world cup 2018.
3. Write a query in SQL to find the match no in which Germany played against Russia.
4. Write a query in SQL to find the team which was defeated by France in FIFA world cup 2018 final.

Question:61

[TYPE: **Data Manipulation**, CONTRIBUTOR: **20BM6JP39**]

A MNC wants to identify its Region-Department Level (APAC / US&Canada / Europe & Middle East) earnings at a quarterly level. As an analyst, you have been asked to extract the necessary data for the annual report. The details of the schema are as follows:

City_mapping = <city_id : integer, name: string, country: string, region: integer, continent: string, zone_id: integer>

DEPARTMENT = <id: string, name: string>

EARNINGS = <department_id : string, city_id : integer, report_date : date, gross_earnings: float, revenue: int, expenses: float>

- (i) Identify the city with highest profit in the last 3 years.
- (ii) Identify the department burning cash with least returns as investment in the previous year.
- (iii) Rank the countries in the descending order of gross earnings.