SQL Questions for Data Engineers

1. **Problem Statement:**
   Assume there are two tables, `orders` and `order_details`. Write an SQL query to retrieve all orders along with their corresponding order details where the total price of the order is greater than 100.

   Create Statement:

```
CREATE TABLE orders (
    order_id INT PRIMARY KEY,
    customer_id INT,
    order_date DATE
);

CREATE TABLE order_details (
    detail_id INT PRIMARY KEY,
    order_id INT,
    product_id INT,
    quantity INT,
    price DECIMAL(10, 2)
);
```

   Insert Data:

```
INSERT INTO orders (order_id, customer_id, order_date) VALUES
    (1, 1, '2021-01-01'),
    (2, 2, '2021-01-02'),
    (3, 1, '2021-01-03'),
    (4, 3, '2021-01-04');

INSERT INTO order_details (detail_id, order_id, product_id, quantity, price) VALUES
    (1, 1, 1, 2, 50),
    (2, 1, 2, 1, 30),
    (3, 2, 3, 3, 20),
    (4, 3, 1, 3, 40),
```

```
    (5, 4, 2, 2, 25),
    (6, 4, 3, 1, 15);
```

2. **Problem Statement:**
   Assume there are two tables, `employees` and `departments`. Write an SQL query to retrieve all employees along with their corresponding department name.

   Create Statement:

```
CREATE TABLE departments (
    dept_id INT PRIMARY KEY,
    dept_name VARCHAR(50)
);

CREATE TABLE employees (
    emp_id INT PRIMARY KEY,
    emp_name VARCHAR(50),
    dept_id INT,
    hire_date DATE,
    salary DECIMAL(10, 2)
);
```

   Insert Data:

```
INSERT INTO departments (dept_id, dept_name) VALUES
    (1, 'IT'),
    (2, 'Finance'),
    (3, 'HR');

INSERT INTO employees (emp_id, emp_name, dept_id, hire_date, salary) VALUES
    (1, 'John', 1, '2020-01-01', 50000),
    (2, 'Jane', 2, '2020-02-01', 60000),
    (3, 'Mark', 1, '2020-03-01', 55000),
    (4, 'Mike', 3, '2020-04-01', 65000);
```

3. **Problem Statement:**
   Assume there is a table `sales` with columns `product_id`, `sale_date`, and `amount`. Write an SQL query to retrieve the total sales amount for each product for the month of January 2021.

   Create Statement:

```
CREATE TABLE sales (
    sale_id INT PRIMARY KEY,
    product_id INT,
    sale_date DATE,
    amount DECIMAL(10, 2)
);
```

Insert Data:

```
INSERT INTO sales (sale_id, product_id, sale_date, amount) VALUES
    (1, 1, '2021-01-01', 100),
    (2, 2, '2021-01-02', 200),
    (3, 1, '2021-01-03', 150),
    (4, 3, '2021-01-04', 300),
    (5, 2, '2021-02-01', 250),
    (6, 3, '2021-02-02', 350);
```

4. **Problem Statement:**
   Assume there is a table `logins` with columns `user_id` and `login_time`. Write an SQL query to retrieve the number of logins for each user for the month of January 2021.

   Create Statement:

```
CREATE TABLE logins (
    login_id INT PRIMARY KEY,
    user_id INT,
    login_time TIMESTAMP
);
```

   Insert Data:

```
INSERT INTO logins (login_id, user_id, login_time) VALUES
    (1, 1, '2021-01-01 12:00:00'),
    (2, 2, '2021-01-01 13:00:00'),
    (3, 1, '2021-01-02 10:00:00'),
    (4, 3, '2021-01-02 11:00:00'),
    (5, 2, '2021-02-01 12:00:00'),
    (6, 3, '2021-02-01 13:00:00');
```

5. **Problem Statement:**
   Assume there are two tables, `customers` and `orders`. Write an SQL query to retrieve

all customers who have placed at least one order.

Create Statement:

```
CREATE TABLE customers (
    customer_id INT PRIMARY KEY,
    customer_name VARCHAR(50),
    address VARCHAR(100)
);

CREATE TABLE orders (
    order_id INT PRIMARY KEY,
    customer_id INT,
    order_date DATE
);
```

Insert Data:

```
INSERT INTO customers (customer_id, customer_name, address) VALUES
    (1, 'John', '123 Main St'),
    (2, 'Jane', '456 Oak Ave'),
    (3, 'Mark', '789 Elm St');

INSERT INTO orders (order_id, customer_id, order_date) VALUES
    (1, 1, '2021-01-01'),
    (2, 2, '2021-01-02'),
    (3, 1, '2021-01-03'),
    (4, 3, '2021-01-04');
```

6. **Problem Statement:**
   Assume there is a table `transactions` with columns `transaction_id`, `user_id`, and `amount`. Write an SQL query to retrieve the average transaction amount for each user.

   Create Statement:

```
CREATE TABLE transactions (
    transaction_id INT PRIMARY KEY,
    user_id INT,
    amount DECIMAL(10, 2)
);
```

   Insert Data:

```
INSERT INTO transactions (transaction_id, user_id, amount) VALUES
    (1, 1, 50),
    (2, 2, 100),
    (3, 1, 75),
    (4, 3, 200),
    (5, 2, 125),
    (6, 3, 150);
```

7. **Problem Statement:**
   Assume there is a table `products` with columns `product_id` and `price`. Write an SQL query to retrieve the top 3 most expensive products.

   Create Statement:

```
CREATE TABLE products (
    product_id INT PRIMARY KEY,
    price DECIMAL(10, 2)
);
```

   Insert Data:

```
INSERT INTO products (product_id, price) VALUES
    (1, 100),
    (2, 200),
    (3, 150),
    (4, 300),
    (5, 250);
```

8. **Problem Statement:**
   Assume there are two tables, `students` and `grades`. Write an SQL query to retrieve the average grade for each student.

   Create Statement:

```
CREATE TABLE students (
    student_id INT PRIMARY KEY,
    student_name VARCHAR(50),
    address VARCHAR(100)
);

CREATE TABLE grades (
    grade_id INT PRIMARY KEY,
```

```
    student_id INT,
    course_name VARCHAR(50),
    grade DECIMAL(10, 2)
);
```

Insert Data:

```
INSERT INTO students (student_id, student_name, address) VALUES
    (1, 'John', '123 Main St'),
    (2, 'Jane', '456 Oak Ave'),
    (3, 'Mark', '789 Elm St');

INSERT INTO grades (grade_id, student_id, course_name, grade) VALUES
    (1, 1, 'Math', 90),
    (2, 2, 'Math', 95),
    (3, 1, 'Science', 80),
    (4, 3, 'Math', 85),
    (5, 2, 'Science', 92),
    (6, 3, 'Science', 88);
```

9. **Problem Statement:**
   Assume there are two tables, `employees` and `salaries`. Write an SQL query to retrieve all employees along with their corresponding salary.

   Create Statement:

```
CREATE TABLE employees (
    emp_id INT PRIMARY KEY,
    emp_name VARCHAR(50),
    hire_date DATE
);

CREATE TABLE salaries (
    salary_id INT PRIMARY KEY,
    emp_id INT,
    salary DECIMAL(10, 2),
    start_date DATE,
    end_date DATE
);
```

   Insert Data:

```
INSERT INTO employees (emp_id, emp_name, hire_date) VALUES
    (1, 'John', '2020-01-01'),
```

```
    (2, 'Jane', '2020-02-01'),
    (3, 'Mark', '2020-03-01'),
    (4, 'Mike', '2020-04-01');

INSERT INTO salaries (salary_id, emp_id, salary, start_date, end_date) VALUES
    (1, 1, 50000, '2020-01-01', '2020-12-31'),
    (2, 2, 60000, '2020-01-01', '2020-12-31'),
    (3, 1, 55000, '2021-01-01', '2021-12-31'),
    (4, 3, 65000, '2021-01-01', '2021-12-31'),
    (5, 2, 70000, '2021-01-01', '2021-12-31'),
    (6, 4, 75000, '2021-01-01', '2021-12-31');
```

10. **Problem Statement:**
    Assume there is a table `orders` with columns `order_id`, `order_date`, and `total_price`. Write an SQL query to retrieve the total sales for each month.

    Create Statement:

```
CREATE TABLE orders (
    order_id INT PRIMARY KEY,
    order_date DATE,
    total_price DECIMAL(10, 2)
);
```

    Insert Data:

```
INSERT INTO orders (order_id, order_date, total_price) VALUES
    (1, '2021-01-01', 100),
    (2, '2021-01-02', 200),
    (3, '2021-02-01', 150),
    (4, '2021-02-02', 300),
    (5, '2021-03-01', 250),
    (6, '2021-03-02', 350);
```