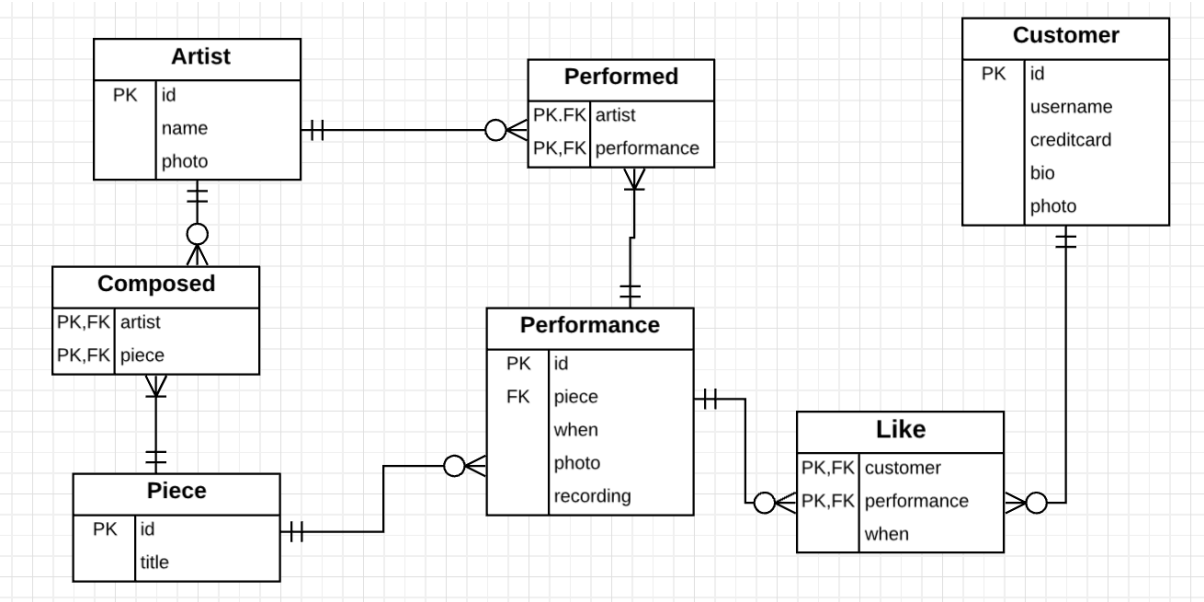
Q1



Q2

1. SELECT lastname FROM Customer

   WHERE instr(lastname,'e') > 0

   ORDER BY length(lastname) DESC

    LIMIT 1;

2. SELECT name, price

   FROM product

   WHERE price > 3 AND category in

        (SELECT id FROM productcategory

        WHERE name = "food")

   ORDER BY price DESC;

3. SELECT ROUND(AVG(numItems),2) AS AvgItems,

    FROM (

        SELECT purchase, count(*) AS numItems

        FROM PurchaseItem

        GROUP BY purchase

) AS PurchaseCounts;

4. SELECT firstname, lastname, COUNT(DISTINCT(paymethod)) AS numMethods

   FROM Customer INNER JOIN Purchase ON Customer.id = Purchase.customer

   GROUP BY Customer.id

   HAVING numMethods =

      (SELECT COUNT(*) FROM PaymentMethod);

5. SELECT COUNT(*) as TotalPurchases

   FROM Purchase

   WHERE checkout =

      (SELECT MAX(id) FROM Checkout

      WHERE Checkout.store = Purchase.store);

Q3

Delivery(deliveryId, restaurantId, deliverymanId, foodId, paymentMethod)

Restaurant(restaurantId, restaurantName, restaurantPhone, restaurantAddress,)

Deliveryman(deliverymanId, deliverymanName)

Food(foodId, foodName, foodPrice)

Q4

1. INT
2. DATE
3. BLOB
4. BLOB
5. VARCHAR(50)
6. TIMESTAMP
7. CHAR(20)
8. CHAR(19)
9. VARCHAR(500)

Q5

A) Inconsistent retrieval occurs when one transaction calculates some aggregate functions over a set of data, while other transactions are updating the data.
For example, some data may be read after they are changed and some data may be read before they are changed, yielding inconsistent results.

B) 1. Database-level lock
– Good for batch processing but unsuitable for multi-user DBMSs
– SQLite, Access
2. Table-level lock
– Can cause bottlenecks, even if transactions want to access different parts of the table and would not interfere with each other
– Not suitable for highly multi-user DBMSs
3. Page-level lock
– An entire disk page is locked (a table can span several pages and each page can contain several rows of one or more tables)
– Not commonly used now
4. Row-level lock
– Improves data availability but with high overhead (each row has a lock that must be read and written to)
– Currently the most popular approach (MySQL, Oracle)
5. Field-level lock
– Allow concurrent transactions to access the same row, as long as they access different attributes within that row
– Most flexible lock but requires an extremely high level of overhead
– Not commonly used

Q6

A) • Presentation logic - User interface
– input (keyboard, touchscreen, voice)
– output (large screen, printer, phone, ATM)
• Business logic - Application Software
– input and command handling
– enforcement of business rules
• Storage logic - Data Storage
– persistent storage of data
– enforcement of data integrity

B) • Advantages
– Scalability
– Technological flexibility (can change business logic easily)
– Can swap out any single component fairly easily
– Long-term cost deduction
– Improved security - customer machine dose presentation only
• Disadvantages

– High short-term cost

– Tools and training

– Complex to design

– Variable standards

Q7

A) Possible threats include: loss of integrity, loss of availability and loss of confidentiality.
Solution:
• Access control
– grant different types of discretionary privileges to users.
– views are an important discretionary authorization mechanism.
• Encryption
– Data is encoded using an algorithm
– Particular tables or columns may be encrypted to protect sensitive data (e.g. password)

B) A cold backup is done when there is no user activity going on with the system. Also called as offline backup, is taken when the database is not running and no users are logged in. All files of the database are copied and no changes during the copy are made.
While a hot backup is taken when the database needs to run all the time. It is an online backup. All files of the database are copied and there may be changes to the database during the copy.

C) • Primary defences
– prepared statements (parameterized queries)
– stored procedures (both mean SQL is no longer dynamic)
• Additional defences
– Principle of least privilege (do not give application accounts DBA privileges)
– White list input validation (check input is from a list of acceptable values)

Q8

A) • Distributed database

– A single logical database physically spreads across multiple computers in multiple locations connected by a data communications link

– Appears to users as though it is one database

• Decentralized database

– a collection of independent databases which are not networked together as one logical database

– appears to users as though many databases

B) Advantages

1. good fit for geographically distributed organizations and users
2. data located near site with greatest demand
3. faster data access (to local data)
4. faster data processing

Disadvantages

1. Complexity of management and control
2. Data integrity
3. Security
4. Lack of standards

C) 1. Need more storage space

2. Integrity
   • can retrieve incorrect data if updates have not arrived
3. Takes time for update operations
   • high tolerance for out-of-date data may be required
   • updates may cause performance problems for busy nodes

Q9

A) NoSQL has several advantages over relational database:
   – designed to run on distributed servers
   – most are open-source
   – built for modern web
   – schema-less (though there may be implicit schema).
   e.g. MongoDB

B) Consistency - Availability - Partition tolerance

- In the presence of a partition, one is then left with two options: consistency or availability.
  - When choosing consistency over availability, the system will return an error or a time-out if particular information cannot be guaranteed to be up-to-date due to network partitioning.
  - When choosing availability over consistency, the system will always process the query and try to return the most recent available version of the information, even if it cannot guarantee it is up-todate due to network partitioning.