Part 1: Multiple Choice (108 points - 3 points per question)

1. Which is true?
   (A) Database system is a collection of related data. (B) Metadata is the state of data.
   (C) Database instance changes every time the database is updated. (D) none of the above

2. Which is not a function of a DBMS?
   (A) database manipulation (B) database construction (C) database definition (D) data abstraction

3. Which is a DBMS? (A) Excel (B) Access (C) Apache (D) none of the above

4. Which is a job of a DBA?
   (A) Defining the database constraints (B) Authorizing access to the database
   (C) Defining the database transactions (D) none of the above

5. In the three-schema architecture, which schema can have multiple view?
   (A) External schema (B) Conceptual schema (C) Internal schema (D) none of the above

6. Which is in DBMS-independent design process?
   (A) requirement analysis (B) application program design (C) transaction Implementation (D) none of the above

7. Which stores the presentation layer of the application?
   (A) Web interface (B) database server (C) application server (D) none of the above

8. Which is true?
   (A) A relation has only one key. (B) A superkey must be minimal. (C) Any key is a superkey. (D) none of the above

9. Which constraint may delete violate?
   (A) referential integrity (B) key constraint (C) domain constraint (D) none of the above

10. Which integrity constraints can trigger a sequence of operations?
    (A) restrict (B) set default (C) cascade (D) none of the above

11. Which is a SQL DCL command? (A) delete (B) grant (C) update (D) none of the above

12. Which SQL command is used to empty a table?
    (A) delete table (B) drop table (C) truncate table (D) none of the above

13. Which is used to to sort the result set in SQL? (A) order by (B) group by (C) sort by (D) having

14. Which returns the number of rows in SQL? (A) count() (B) number() (C) num() (D) none of the above

15. Which of the following finds those groups meeting stated conditions?
    (A) group by (B) using by (C) having (D) none of the above

16. What is a virtual table in SQL? (A) view (B) vision (C) screen (D) none of the above

17. To eliminate duplicate rows in a query, which can be used? (A) unique (B) index (C) distinct (D) none of the above

18. Which is a join condition in the following SQL commands?
    select name from employee, department
    where department.name = 'Research' and employee.department_no = department.department_no;
    (A) select name from employee, department (B) department.name = 'Research'
    (C) employee.department_no = department.department_no (D) none of the above

19. Which of the following represents all attributes of a table in a SQL statement? (A) % (B) <> (C) & (D) *

20. Which SQL command is used to search for a specified pattern in a column?
    (A) as (B) like (C) match (D) none of the above

21. In MySQL which command can show the schema of a table? (A) show (B) display (C) present (D) describe

22. In MySQL which is used to execute a SQL script? (A) use (B) \\e (C) \ (D) none of the above

23. When you log into your Faceplayer, which SQL command will be used? (A) insert (B) update (C) delete (D) select

24. A rule that database users should obey is called a: (A) regulation. (B) principle. (C) privilege. (D) constraint.

25. Which is a benefit of denormalization?
    (A) performance improvement (B) higher security (C) less storage (D) none of the above

26. Which is an entity in a conference? (A) location (B) session (C) date (D) none of the above

27. Which is usually to represent a relationship in an ER diagram? (A) adjective (B) noun (C) proposition (D) verb

28. Which is the process of grouping a set of subclasses into a superclass?
    (A) Specialization (B) Identification (C) Aggregation (D) Generalization

29. In the following EER diagram, which is true?
    (A) A student must be a graduate or an undergraduate student.
    (B) An undergraduate student cannot be a special student.
    (C) A student must be a graduate student, an undergraduate, or a special student.
    (D) none of the above
(D) 30. In ER digram what does a double ellipse represent?
   (A) simple attribute (B) associative attribute (C) derived attribute (D) none of the above
(C) 31. Which specifies the number of instances of one entity that can be associated with each instance of another entity?
   (A) degree (B) counter constraint (C) cardinality constraint (D) limit
(B) 32. Which type of relationships between instructors and courses?
   (A) one-to-many (B) many-to-many (C) many-to-one (D) one-to-one
(B) 33. Which refers to the interpretation of attribute values in a tuple? (A) instance (B) semantics (C) syntax (D) ontology
(D) 34. Which is a type of semi-structured data? (A) images (B) text (C) video (D) XML
(B) 35. Which provides a standard software API for using DBMS? (A) CASE (B) ODBC (C) OLAP (D) none of the above
(A) 36. Which is persistent data? (A) HTML documents (B) SQL statements (C) work queues (D) none of above

Part 2: Questions and Answers (115 points)

1. (28 points) Briefly explain these terminologies. If they are acronyms, also write what they stand for.
   (a) data model (b) trigger (c) functional dependency (d) identifying relationship (e) ontology (f) semantic Web (g) XML
      (a) A set of concepts to describe the structure of a database, the operations for manipulating these structures, and certain
         constraints that the database should obey.
      (b) A trigger is a statement that is automatically executed in response to certain events on a particular table or view in a
         database.
      (c) Functional dependency specifies that the value of one attribute (the determinant) determines the value of another
         attribute in the same table.
      (d) An identifying relationship is a relationship that links a strong entity to a weak entity.
      (e) Ontology means using conceptual modeling and other tools to develop “a specification of a conceptualization”.
      (f) Semantic Web is a framework that includes semantic content in Web pages allowing meaningful information exchange
         and search among machines.
      (g) EXtensible Markup Language (XML) is a language used to specify the data content.

2. (8 points) Explain the key constraint, domain constraint, entity integrity constraint, and referential constraint.
   • The key constraint means there is no duplicate key in any relation.
   • The domain constraint indicates every value in a tuple must be from the domain of its attribute.
   • The entity integrity indicates the values of primary key attributes in a relation cannot be null.
   • The referential integrity constraints indicate any attribute of a foreign key in a table can contain only either values
     from the corresponding parent table’s primary key or the null value.

3. (a) (4 points) What does SQL stand for? Explain it.
   (b) (6 points) Based on the functions how can SQL be classified into three categories?
      (a) Structured Query Language (SQL) is a standard language used to retrieve, update and delete data from relational
          database management systems (DBMS).
      (b) Data Definition Language (DDL) is used to define databases.
          Data Manipulation Language (DML) is used to manipulate databases.
          Data Control Language (DCL) is used to control databases.

4. (a) (3 points) Explain normalization for a relational database.
   (b) (9 points) Explain the 1NF, 2NF, and 3NF.
      (a) The process of structuring relations by decomposing their attributes into smaller relations.
      (b) i. 1NF is the relation that has no composite attributes, multivalued attributes, and nested relations.
          ii. In 2NF, every non-prime attribute is fully functionally dependent on the primary key in the relation.
          Other possible answers:
          There are no two keys in the relation.
          There is no partial functional dependency in the relation.
          iii. Third normal form (3NF) is a normal form in which there is no transitive functional dependency in the relation.

5. (33 points) Consider the following swimming competition database:

<table>
<thead>
<tr>
<th>player table</th>
<th>competition</th>
<th>event table</th>
</tr>
</thead>
<tbody>
<tr>
<td>player_no</td>
<td>competition</td>
<td>event_no</td>
</tr>
<tr>
<td>event_no</td>
<td>player_no</td>
<td>date</td>
</tr>
<tr>
<td>101022</td>
<td>Lady Gaga</td>
<td>USA</td>
</tr>
<tr>
<td>101145</td>
<td>Taylor Swift</td>
<td>USA</td>
</tr>
<tr>
<td>101262</td>
<td>Lily Allen</td>
<td>UK</td>
</tr>
<tr>
<td>E101</td>
<td>101022</td>
<td>2014-06-17</td>
</tr>
<tr>
<td>E202</td>
<td>101145</td>
<td>2014-06-17</td>
</tr>
<tr>
<td>E303</td>
<td>101262</td>
<td>2014-06-17</td>
</tr>
<tr>
<td>E101</td>
<td>50m Freestyle</td>
<td></td>
</tr>
<tr>
<td>E202</td>
<td>100m Breaststroke</td>
<td></td>
</tr>
<tr>
<td>E303</td>
<td>200m Butterfly</td>
<td></td>
</tr>
</tbody>
</table>

where primary keys are underlined. player_no and event_no in the competition table are foreign keys referencing to the
player and event table respectively.
(a) If the following operations are taken, check if domain constraints, key constraints, entity integrity, or referential integrity is violated. If there is any violation, explain it.

i. (2 points) Insert ('101262', 'Ketty Perry', 'USA', 3) into the player table.
ii. (2 points) Change the event_no of 'E202' in the event table from 'E202' to 'E212'.
iii. (2 points) Remove the row (E202, '100m Breaststroke') from the event table.
iv. (2 points) Change the event name in the event table from '200m Butterfly' to '100m Backstroke'.

(b) Use SQL to answer the following questions.

i. (3 points) Create the player table with the required constraints.
ii. (2 points) Insert an 'USA' player, 'Justin Timberlake' with ID '101168' and 5 medals into the player table.
iii. (2 points) Change '50m Freestyle' to '100m Butterfly' for the event number E101 in the event table.
iv. (2 points) Add a attribute club into the player table.
v. (3 points) Specify a constraint that restricts time between 9:00 and 21:00 in the competition table.
vi. (3 points) Specify event_no, player_no as the primary key in the competition table.
vii. (3 points) Remove all competitions attended by 'Lady Gaga' from the competition table.
viii. (3 points) List all events attended by 'Katy Perry'.
ix. (4 points) List the name and the number of events which each player attends.

(a) i. It violates the key constraint because the player_no '100262' already existed.
ii. It violates the referential integrity because the foreign key, event_no 'E202' in the perform table will have no primary key to reference to in the event table.
iii. It violates the referential integrity because the foreign key, event_no 'E202' in the perform table will have no primary key to reference to in the event table.
iv. It violates no constraint.

(b) i. create table player ( 
    player_no char(6) primary key not null,
    name varchar(30),
    country varchar(15),
    medal integer);
ii. insert into player values ('101168', 'Justin Timberlake', 'USA', 5)
iii. update event set event = '50m Freestyle' where event_no = 'E101'
iv. alter table player add club varchar(30);
v. alter table competition add constraint time
    constraint check (time >='9:00' and time <= '21:00')
vi. alter table competition add primary key (event_no, player_no)
vii. delete from competition where player_no = (select player_no from player where name = 'Lady Gaga')
viii. select event from player, competition, event
     where player.player_no = competition.player_no and competition.event_no = event.event_no and
     name = 'Katy Perry'
ix. select player_no, name, count(*) from player, competition
    where player.player_no = competition.player_no group by player_no

6. (14 points) Consider the world cup football game. There are player, teams, and matches. Each player has player number, name, and birthday. Each player will join a country team with confederation. Each match has date, time, and location. There are two types of players: professional and amateur. Each professional player has the professional year information. Each amateur players has the club name information.

(a) (7 points) Add the necessary relationships to draw a EER diagram to represent the relationships among teams, players, and matches.

(b) (7 points) Transform the above EER diagrams to 3NF relations and show the 3NF relation schema.
7. (10 points) Consider the following schema for the information about students borrowing books. Normalize it to 3NF relations.

**student_borrow_book** (student_id, lastname, firstname, gender, depart_id, department, isbn, booktitle, author, publisher, purchase_date, checkout, checkout_date, return_date)

department: (depart_id, department)

student: (student_id, lastname, firstname, gender, depart_id)

checkout: (student_id, isbn, checkout, checkout_date, return_date)

book: (isbn, booktitle, author, publisher, purchase_date)