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

(A) 1. Which is true?
   (A) A database is a collection of related data. (B) Data after processing is knowledge. (C) Database schema changes every time the database is updated. (D) None of the above

(D) 2. All of the following are properties of metadata except:
   (A) data definitions. (B) data structures. (C) rules or constraints. (D) processing logic.

(B) 3. 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.

(C) 4. Which is the application of data warehouses?
   (A) Shipping of information. (B) Order processing. (C) Decision support. (D) File updating.

(D) 5. With the database approach, data descriptions are stored in a central location known as a(n):
   (A) Index. (B) Data warehouse. (C) Data server. (D) Repository.

(B) 6. Which is in DBMS-independent design process?
   (A) Transaction implementation. (B) Requirement analysis. (C) Application program design. (D) None of the above.

(B) 7. Which is a SQL DCL command?
   (A) Delete. (B) Grant. (C) Update. (D) None of the above.

(C) 8. Which SQL command is used to empty a table?
   (A) Delete table. (B) Drop table. (C) Truncate table. (D) None of the above.

(A) 9. Which is used to order the result set in SQL?
   (A) Order by. (B) Group by. (C) Sort by. (D) Having.

(D) 10. In MySQL which command can show the schema of a table?
   (A) Show. (B) Display. (C) Present. (D) Describe.

(A) 11. Which returns the number of rows in SQL?
   (A) COUNT(). (B) NUMBER(). (C) NUM(). (D) None of the above.

(D) 12. What does the following SQL statement do? `select name from student where major = 'Computer Science';`?
   (A) Retrieves all majors whose name is Computer Science from the student table. (B) Retrieves all students who likes Computer Science from the student table. (C) Retrieves the name of all students majoring in Computer Science. (D) None of the above.

(A) 13. Which SQL command can be used to change a table definition?
   (A) ALTER TABLE. (B) CHANGE TABLE. (C) MODIFY TABLE. (D) None of the above.

(D) 14. In the LIKE operator of SQL, which can represent any string?
   (A) * (B) ; (C) % (D) _

(D) 15. Which is not a aggregation function in SQL?
   (A) FIRST(). (B) SUM(). (C) AVG(). (D) TOTAL().

(A) 16. When you post a message on your Facebook account, which SQL command will be used?
   (A) Insert. (B) Grant. (C) Select. (D) Delete.

(A) 17. What is a virtual table in SQL?
   (A) View. (B) Vision. (C) Screen. (D) None of the above.

(B) 18. In MySQL which is used to execute a SQL script?
   (A) Use. (B) Source. (C) EXECUTE (D) None of the above.

(C) 19. Which is the rule activated by updates to the table?
   (A) Constraint. (B) Business rule. (C) Trigger. (D) None of the above.

(D) 20. Which is not a meaning for null values?
   (A) Attribute does not apply to this tuple. (B) Value exists but is not available. (C) Attribute value is unknown. (D) Value is beyond the domain range.

(C) 21. Which is true?
   (A) A super key is a key. (B) A superkey should be minimal. (C) Candidate keys can be designated as unique keys. (D) None of the above.

(B) 22. Which is an entity in a gallery?
   (A) Open hours. (B) Exhibition items. (C) Ticket price. (D) None of the above.

(D) 23. Which is the process of storing the join of higher normal form relations as a base relation?
   (A) Specialization. (B) Generalization. (C) Realization. (D) Denormalization.

(A) 24. Which is usually to represent a relationship in an ER diagram?
   (A) Verb. (B) Noun. (C) Adjective. (D) Proposition.

(D) 25. An entity whose existence depends on another entity is called:
   (A) Codependent entity. (B) Variant entity. (C) Strong entity. (D) Weak entity.

(A) 26. The _____ of a relationship type is the number of participating entity instances.
   (A) Cardinality. (B) Degree. (C) Identification. (D) Participation.

(B) 27. Which type of relationships between a course and a student?
   (A) One-to-many. (B) Many-to-many. (C) Many-to-one. (D) One-to-one.

(D) 28. Which is a type of semi-structured data?
   (A) Images. (B) Text. (C) Video. (D) XML.

(B) 29. Which rule states that an entity instance can simultaneously be a member of two (or more) subtypes?
   (A) Partial specialization. (B) Overlap. (C) Disjoint. (D) Total specialization.
(C) 30. Which of the following is false?
   (A) The EER is a type of conceptual data models. (B) E. F. Codd first proposed the relation model.
   (C) An entity can exist in the database merely by being a member of a subclass. (D) none of the above

(C) 31. Which is persistent data? (A) SQL statements (B) work queues (C) HTML documents (D) none of above

(A) 32. Which constraint may the delete operation violate?
   (A) Referential constraint (B) Entity constraint (C) Key integrity (D) None of the above

(A) 33. In the following EER diagram, which is true?

(A) A person can be a camper and a runner. (B) A person must be a camper, a biker, and a runner.
   (C) A person must be a camper, a biker, or a runner. (D) None of the above

(B) 34. Which of the following finds those groups meeting stated conditions?
   (A) group by (B) having (C) using by (D) none of the above

(B) 35. Which is an approach to map a n-ary relationship type in the ER diagram to the relational schema?
   (A) Merged relation (B) A relationship relation and n foreign keys
   (C) A relation set of simple component attributes (D) None of the above

(D) 36. Which integrity constraints can trigger a sequence of operations?
   (A) restrict (B) set default (C) set null (D) cascade

Part 2: Questions and Answers (106 points)

1. (16 points) Briefly explain these terminologies. If they are acronyms, also write what they stand for.
   (a) (3 points) entity integrity (b) (3 points) functional dependency (c) (3 points) ontology
   (d) (4 points) XML (e) (3 points) data mining

   (a) The entity integrity indicates the values of primary key attributes in a relation cannot be null.
   (b) Functional dependency specifies that the value of an attribute in a table determine the value of other attribute in the same table.
   (c) Ontology means using conceptual modeling and other tools to develop a specification of a conceptualization.
   (d) EXtensible Markup Language (XML) is a language used to specify the data content.
   (e) The data mining can be defined in either one as shown in below:
      • The discovery of new information in terms of patterns or rules from vast amounts of data.
      • The process of finding interesting structure in data.
      • The process of employing one or more computer learning techniques to automatically analyze and extract knowledge from data.

2. (a) (4 points) What is data model?
   (b) (6 points) Describe the three-schema architecture of databases.

   (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) In the three-schema architecture, schemas can be defined at three levels: internal schema, conceptual schema, and external schemas.

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 music store database:

<table>
<thead>
<tr>
<th>employee table</th>
<th>works_on table</th>
</tr>
</thead>
<tbody>
<tr>
<td>employee_no</td>
<td>employee_no</td>
</tr>
<tr>
<td>no</td>
<td>project_no</td>
</tr>
<tr>
<td>name</td>
<td>hours</td>
</tr>
<tr>
<td>position</td>
<td></td>
</tr>
<tr>
<td>department_no</td>
<td></td>
</tr>
<tr>
<td>E99022</td>
<td>E99022</td>
</tr>
<tr>
<td>Lady Gaga</td>
<td>10</td>
</tr>
<tr>
<td>manager</td>
<td>60000</td>
</tr>
<tr>
<td>MK100</td>
<td></td>
</tr>
<tr>
<td>E99145</td>
<td>E99145</td>
</tr>
<tr>
<td>Taylor Swift</td>
<td>8</td>
</tr>
<tr>
<td>clerk</td>
<td>30000</td>
</tr>
<tr>
<td>SL200</td>
<td></td>
</tr>
<tr>
<td>E99262</td>
<td>E99262</td>
</tr>
<tr>
<td>Amy Winehouse</td>
<td>12</td>
</tr>
<tr>
<td>secretary</td>
<td>40000</td>
</tr>
<tr>
<td>HR300</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>department table</th>
<th>project table</th>
</tr>
</thead>
<tbody>
<tr>
<td>department_no</td>
<td>project_no</td>
</tr>
<tr>
<td>department_name</td>
<td>project_title</td>
</tr>
<tr>
<td>manager_no</td>
<td></td>
</tr>
<tr>
<td>MK100</td>
<td>PJ101</td>
</tr>
<tr>
<td>Marketing</td>
<td>World Peace</td>
</tr>
<tr>
<td>E95022</td>
<td></td>
</tr>
<tr>
<td>SL200</td>
<td>PJ202</td>
</tr>
<tr>
<td>Sales</td>
<td>Save the Earth</td>
</tr>
<tr>
<td>E90123</td>
<td></td>
</tr>
<tr>
<td>HR300</td>
<td>PJ311</td>
</tr>
<tr>
<td>Human Resources</td>
<td>Save the Children</td>
</tr>
<tr>
<td>E88234</td>
<td></td>
</tr>
</tbody>
</table>

where primary keys are underlined. department_no in the employee table is a foreign key referencing to the department table. employee_no and project_no in the works_on table are foreign keys referencing to the employee and project 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 ('E95262', 'Ketty Perry', 'Marketing', MK100) into the employee table.
ii. (2 points) Change the project_no of 'PJ202' in the works_on table from 'PJ202' to 'PJ212'.
iii. (2 points) Remove the row ('SL200', 'Sales', E95022) from the department table.
iv. (2 points) Change the project name in the project table from 'World Peace' to 'Stop Violence'.

(b) Use SQL to answer the following questions.

i. (3 points) Create the employee table with the required constraints.
ii. (3 points) Add a constraint of 0 ≤ hour ≤ 40 in the works_on table.
iii. (2 points) Add an attribute gender into the employee table.
iv. (2 points) Insert ('RH400', 'Research', 'E93202') into the department table.
v. (3 points) For each project, list the project name and the total hours per week (by all employees) spent on that project.
vi. (2 points) Change the position of Taylor Swift from 'clerk' to 'staff'.
vii. (2 points) List the name of the employee who has highest salary and her or his salary.
viii. (3 points) Remove all projects Amy Winehouse works on.
ix. (3 points) List the department name and the number of employees in the order of the number of employees.
x. (3 points) List the department name and the average salary which is higher than 50000.

(a) i. It violates the key constraint because the employee_no 'E95262' already existed.
ii. It violates the referential integrity because the foreign key, works_on 'PJ212' in the class table will have no primary key to reference to in the project table.
iii. It violates the referential integrity because the foreign key, department_no 'SL2638' in the employee table will have no primary key to reference to in the department table.
iv. It violates no constraint.

(b) i. create table employee (
   employee_no char(6) primary key not null,
   name varchar(30),
   position varchar(15),
   department_no char(6));
ii. alter table works_on add constraint hour_constraint check (0 <= hour and hour <= 40)
iii. alter table employee add gender char(1);
iv. insert into department values ('RH400', 'Research', 'E93202');
v. select pname, sum(hours) from project, works_on where pnumber=pno group by pnumber, pname
vi. update employee set position = 'staff' where name = 'Taylor Swift';
vii. select name, max(salary) as max_salary from employee where salary = max_salary;
viii. delete from project where employee_no = 
(select employee_no from employee, works_on
 where project.project_no = works_on.project_no and name = 'Amy Winehouse');
ix. select department_name, count(*) as num_of_employee from employee, department
 where employee.department_no = department.department_no
 group by department_name order by num_of_employee;
x. select department_name, avg(salary) as avg_salary from employee, department
 where employee.department_no = department.department_no
 group by department_name having avg_salary > 50000;

6. (8 points) Consider a database in a university with the following functional dependencies:

StdSSN \rightarrow \text{StdCity, StdClass}
OfferNo \rightarrow \text{Semester, Year, CourseNo, CourseName}
CourseNo \rightarrow \text{CourseName}
StdSSN, OfferNo \rightarrow \text{Score}

Decide which type of normalized form it is. If it is not a 3NF, transform it to a 3NF.

(a) (3 points) course ((OfferNo, Semester, Year, CourseNo), course(CourseNo, courseName))
(b) (5 points) grade_report (stdSSN, StdClass, OfferNo, Year, CourseNo, CourseName, score)
(a) 2NF, offer((OfferNo, Semester, Year, CourseNo), course(CourseNo, courseName))
(b) 1NF, student(StdSSN, StdClass), offer((OfferNo, Year, CourseNo), score(StdSSN, CourseNo, score),
course(CourseNo, courseName))

7. (16 points) Consider the following schema for the enrollment information. Normalize it to 3NF relations.
enrollment(student_id, student_name, course_no, title, credit, section_no, classroom, time, employee_id, instructor_name, office)

where primary keys are underlined.
(a) (8 points) Draw the ER diagram.
(b) (8 points) Show the relation schema and indicate the functional dependency.

(a)