Assignment 3: SQL Database Modification and Advanced SELECT Query

This assignment uses the same database schema as used in Assignment 2. Answer the following questions by writing SQL statements.

PART 1: SQL Database Modification (12 points)

Database 1: computer. A computer database schema consists of four relations, whose schemas are:

```
Product (maker, model, type)
PC (model, speed, ram, hdisk, price)
Laptop (model, speed, ram, hdisk, screen, price)
Printer (model, color, type, price)
```

Write the following database modifications in SQL, based on the above database schema.

1. Using two INSERT statements, store in the database the fact that PC model 1100 is made by manufacturer C, has speed 3.2, RAM 1024, hard disk 180, and sells for $2499.
2. Insert the facts that for every PC there is a laptop with the same manufacturer, speed, RAM, and hard disk, a 17-inch screen, a model number 1100 greater, and a price $500 more.
3. Delete all PC 's with less than 100 gigabytes of hard disk.
4. Delete all laptops made by a manufacturer that doesn't make printers.
5. Manufacturer A buys manufacturer B. Change all products made by B so that they are now made by A.
6. For each PC, double the amount of RAM and add 60 gigabytes to the amount of hard disk.
7. For each laptop made by manufacturer B, add one inch to the screen size and subtract $100 from the price.

Database 2: battleship. A database of World War II warships has the following relations:

```
Classes (class, type, country, guns, bore, displacement)
Ships (name, class, launched)
Battles (name, bdate)
Outcomes (ship, battle, result)
```

Write the following database modifications in SQL, based on the above database schema.

8. The two British battleships of the Nelson class — Nelson and Rodney — were both launched in 1927, had nine 16-inch guns, and a displacement of 34,000 tons. Insert these facts into the database.
9. Two of the three battleships of the Italian Vittorio Veneto class — Vittorio Veneto and Italia — were launched in 1940; the third ship of that class, Roma, was launched in 1942. Each had nine 15-inch guns and a displacement of 41,000 tons. Insert these facts into the database.
10. Delete from Ships all ships sunk in battle.
11. Modify the Classes relation so that gun bores are measured in centimeters (one inch = 2.54 centimeters) and displacements are measured in metric tons (one metric ton = 1.1 tons). (Note: you may have to change the data type of bore and displacement accordingly.)
12. Delete all classes with fewer than three ships. (Note: This deletion may be rejected due to violation of foreign key constraints.)

PART 2: SQL SELECT Query (Advanced, 50 points)

This part continues Assignment 2 by asking you to answer more questions by writing SQL SELECT queries. Answer questions in this part in the same way as you did in Assignment 2. Questions are numbered continuously after the last question of Assignment 2.

Database 3: catalog

This product catalog database contains the following relations:

```
Suppliers(sid:integer, sname:string, address:string)
Parts(pid:integer, pname:string, color:string)
Catalog(sid:integer, pid:integer, cost:real)
```

The Catalog relation lists the prices charged for parts by suppliers. Write the following queries in SQL:

51. Find the pnames of parts for which there is some supplier.
52. Find the snames of suppliers who supply every part.
53. Find the snames of suppliers who supply every red part.
54. Find the pnames of parts supplied by "Acme Widget Suppliers" and no one else.
55. Find the sids of suppliers who charge more for some part than the average cost of that part (averaged over all the suppliers who supply that part).
56. For each part, find the sname of the supplier who charges the most for that part.
57. Find the sids of suppliers who supply only red parts.
58. Find the sids of suppliers who supply a red part or a green part.
59. Find the sids of suppliers who supply a red part and a green part.
60. Find the sids of suppliers who supply at least two red parts and at least two green parts.
61. For every supplier that only supplies green parts, print the name of the supplier and the total number of parts that she supplies.
62. For every supplier that supplies a green part and a red part, print the name and price of the most expensive part that she supplies.

Database 4: company

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, dname:string, budget:real, managerid:integer)
```

63. Print the names and ages of each employee who works in both the Hardware department and the Software department.
64. For each department with more than 20 full-time-equivalent employees (i.e., where the part-time and full-time employees add up to at least that many full-time employees), print the did together with the number of employees that work in that department.

65. Print the name of each employee whose salary exceeds the budget of all of the departments that he or she works in.
66. Find the managerids of managers who manage only departments with budgets greater than $1 million.
67. Find the enames of managers who manage the departments with the largest budgets.
68. If a manager manages more than one department, he or she controls the sum of all the budgets for those departments. Find the managerids of managers who control more than $5 million.
69. Find the managerids of managers who control the largest amounts of budgets (A manager may manage more than one department).
70. Find the enames of managers who manage only departments with budgets larger than $1 million, but at least one department with budget less than $5 million.

Database 5: university

The university database has the following relations:

```
Student(snum:integer, sname:string, major:string, level:string,
age:integer)
Class(cname:string, meets_at:string, room:string, fid:integer)
Enrolled(snum:integer, cname:string)
Faculty(fid:integer, fname:string, deptid:integer)
```

The meaning of these relations is straightforward; for example, Enrolled has one record per student-class pair such that the student is enrolled in the class. Write the following queries in SQL. No duplicates should be printed in any of the answers.

71. Find the names of all junior students (level = JR) who are enrolled in a class taught by Ivana Teach.
72. Find the names of students not enrolled in any class.
73. Find the students who are enrolled in two classes taught by the same faculty. List student names together with the two class names.
74. Find the names of faculty members who teach in every room in which some class is taught.
75. Print the level and the average age of students for that level, for each level.
76. Print the level and the average age of students for that level, for all levels except JR.
77. For each faculty member that has taught classes only in room R128, print the faculty member's name and the total number of classes she or he has taught.
78. Find the names of students enrolled in the maximum number of classes.
79. Find the names of faculty members for whom the combined enrollment of the courses that they teach is less than five.
80. Find the age of the oldest student who is either a History major or enrolled in a course taught by Ivana Teach.
81. Find the names of all classes that either meet in room R128 or have five or more students enrolled.
82. For each age value that appears in the Student table, find the level value that appears most often. For example, if there are more FR level students aged 18 than SR, JR, or SO students aged 18, you should print the pair (18, FR).

Database 6: flights

The flights database keeps track of airline flight information:

```
Flights(flno:integer, origin:string, destination:string,
  distance:integer, departs:datetime, arrives:datetime, price:integer)
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:

83. Find the names of aircraft such that all pilots certified to operate them earn more than $80,000.
84. For each pilot who is certified for more than three aircraft, find the eid and the maximum cruising range of the aircraft for which she or he is certified.
85. Find the names of pilots whose salary is less than the price of the cheapest route from Los Angeles to Honolulu.
86. For all aircraft with cruising range over 1000 miles, find the name of the aircraft and the average salary of all pilots certified for this aircraft.
87. Find the aids of all aircraft that can be used on routes from Los Angeles to Chicago.
88. Identify the routes that can be piloted by every pilot who makes more than $100,000.
89. Print the names of pilots who can operate aircrafts with cruising range greater than 3000 miles but are not certified on any Boeing aircraft.
90. Compute the difference between the average salary of pilots and the average salary of all employees (including pilots).
91. Print the name and salary of every non-pilot whose salary is more than the average salary for pilots.
92. Print the names of employees who are certified only on aircrafts with cruising range longer than 1000 miles.
93. Print the names of employees who are certified only on aircrafts with cruising range longer than 1000 miles, but on at least two such aircrafts.
94. Print the names of employees who are certified only on aircrafts with cruising range longer than 1000 miles and who are certified on some Boeing aircraft.

Database 7: Lahman's baseball database

Lahman's Baseball Database is a well-known database containing complete batting and pitching statistics since 1871. The complete database is available at http://www.seanlahman.com/baseball-archive. Schema description is given at http://www.seanlahman.com/files/database/readme2017.txt.

The file lahmans.sql in this homework assignment contains a few tables focusing on batters. The following questions are not technically challenging. But they represent typical questions in data analysis.

95. For all players in the people table who were born in 1990 or later (borthYear >= 1990), group them by the birthYear and report the birthYear, average height, and number of players in each birth year. Order the results by birthYear in ascending order. Hint: use ORDER BY.
96. Using the batting and people tables, list the top 10 number of homeruns (HR) scored by a player in a season. In each record, include player (playerID, nameFirst, nameLast), yearID, and number of homeruns. Hint: use ORDER BY and LIMIT.

97. Add a new column to the result of the previous question to report the salary of the player in that year in each record. (Pay attention that some records may have NULL values).
98. Using the batting and people tables, calculate each player's average number of hits (H) over the years they played, and display the players (playerID, nameFirst, nameLast) with the 10 highest hit averages (in descending order of hit average), along with their hit averages.
99. Using the batting, people and salaries tables, calculate each player's average number of hits (H) and average salary over the years they played, and display the players (playerID, nameFirst, nameLast), hit averages, and average salaries with the 10 highest salary averages.
100. Using the halloffame, people, collegeplaying, schools tables, find the people who were successfully inducted into the Hall of Fame and played in college at a school located in Michigan (state='MI'). For each person, return the playerID, nameFirst, nameLast, schoolID, name_full, and the first year (smallest yearID) in the school in descending order of yearID.

Again, to make grading easier, put your answers in a **single plain text file (.txt)** and submit to e-learning. For each question, include the following in your submission:

1. the question itself;
2. your SQL statement;
3. query result from MySQL.

Failure to follow this format may cause your homework not graded.