Assignment 2: SQL SELECT Queries

Install MySQL on your computer ("sudo apt-get install mysql-server" on Ubuntu Linux). Put the following lines in ~/.my.cnf so that you do not need to enter your username/password each time.

```
[client]
user=root
password=your_password
```

Download the attached zip file, which contains all data you'd need for this assignment. Run "mysql < createdb.sql" to create database schema. Run "./loaddata.sh" to load data into the databases.

If you do not use linux or have trouble in loading the data, there is a data dump file datadump.sql prepared for your convenience. Run "mysql < datadump.sql" to create tables and load all data in a single step.

In this assignment, you are asked to write SQL SELECT queries to answer questions listed under the following databases. Test your SQL statement in MySQL to verify its correctness.

To keep query results short, use "SELECT DISTINCT" in the outmost query block if necessary to eliminate duplicates.

To submit your homework, put your answers in a **single plain text file (.txt)** and submit to e-learning. For each question, you should include in your submission:

1. the question number and the question itself in full text;
2. your SQL statement;
3. query result from MySQL.

Database 1: Computer

A computer database 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)
```

1. Find the model number, speed, and hard-disk size for all PC 's whose price is under $1000.
2. List the manufacturers of all printers.
3. Find the model number, memory size, and screen size for laptops costing more than $1500.
4. List all color printers.
5. Find the model number and hard-disk size for those PC 's that have a speed of 3.2 and a price less than $2000.
6. Give the manufacturer and speed of laptops with a hard disk of at least 30.
7. Find the model number and price of all products (of any type) made by manufacturer B.
8. Find those manufacturers that sell Laptops, but not PC 's.

9. Find those hard-disk sizes that occur in two or more PC 's.
10. Find those pairs of PC models that have both the same speed and RAM. A pair should be listed only once; e.g., list (i , j) but not (j,i).
11. Find those manufacturers of at least two different computers (PC's or laptops) with speeds of at least 3.0.
12. Find the makers of PC 's with a speed of at least 3.0.
13. Find the printers with the highest price.
14. Find the laptops whose speed is slower than that of any PC.
15. Find the model number of the item (PC, laptop, or printer) with the highest price.
16. Find the maker of the least expensive color printer.
17. Find the maker(s) of the PC(s) with the fastest processor among all those PC's that have the smallest amount of RAM.
18. Find the average speed of PC 's.
19. Find the average speed of laptops costing over $1000.
20. Find the average price of PC's made by manufacturer "A."
21. Find the average price of PC's and laptops made by manufacturer "D."
22. Find, for each different speed, the average price of a PC.
23. Find for each manufacturer, the average screen size of its laptops.
24. Find the manufacturers that make at least three different models of PC.
25. Find for each manufacturer who sells PC's the maximum price of a PC.
26. Find, for each speed of PC above 2.0, the average price.
27. Find the average hard disk size of a PC for all those manufacturers that make printers.

Database 2: Battleship

A database concerning World War II capital ships. It involves the following relations:

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

28. Find the class name and country for all classes with at least 10 guns.
29. Find the names of all ships launched prior to 1918, but call the resulting column shipName.
30. Find the names of ships sunk in battle and the name of the battle in which they were sunk.
31. Find all ships that have the same name as their class.
32. Find the names of all ships that begin with the letter "R."
33. Find the names of all ships from the Outcomes table whose name consists of three or more words (e.g., King George V).
34. Find the ships heavier than 35,000 tons.
35. List the name, displacement, and number of guns of the ships engaged in the battle of Guadalcanal.
36. Find those countries that have both battleships and battlecruisers.
37. Find those ships that were damaged in one battle, but later fought in.
38. Find those battles with at least three ships of the same country.
39. Find the countries whose ships had the largest number of guns.
40. Find the classes of ships, at least one of which was sunk in a battle.
41. Find the names of the ships with a 16-inch bore.

42. Find the battles in which ships of the Kongo class participated.
43. Find the names of the ships whose number of guns was the largest for those ships of the same bore.
44. Find the number of battleship (type='bb') classes.
45. Find the average number of guns of battleship (type='bb') classes.
46. Find the average number of guns of battleships (type='bb'). Note the difference to the previous question. In this question, you need to weight a class by the number of ships of that class.
47. Find for each class the year in which the first ship of that class was launched.
48. Find for each class the number of ships of that class sunk in battle.
49. Find for each class with at least three ships the number of ships of that class sunk in battle.
50. The weight (in pounds) of the shell fired from a naval gun is approximately one half the cube of the bore (in inches). Find the average weight of the shell for each country's ships.

Database 3: Product Catalog

This 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.
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.

Again, to make grading easier, put your answers in a **single plain text file (.txt)** and submit to e-learning. For each question, include in your submission: 1. the question itself; 2. your SQL statement; and 3. query result from MySQL. Failure to follow this format may cause your homework not graded.