Assignment 2

Due Friday September 16, 2016 by 11:59pm

In this assignment, you will be required to use PostgreSQL. You solutions should include the PostgreSQL statements for solving the problems, as well as the results of running these statements. You can use views but you can not use the `GROUP BY` clause and aggregate functions. You can also not use the `INNER JOIN` (or other joins) operators. Only SQL statements that obey these requirements will receive credit.

1 Miscellaneous Problems

1. (10 points) Let $A(x)$ and $B(x)$ be two unary relation schemas that represent a set A and a set B. (The domain of x in INTEGER.)

   Write a SQL statement that determines whether it is true or not if $A - B$ is empty, $B - A$ is empty, and $A \cap B$ is empty.

   For example, if $A = \{1, 2, 3\}$ and $B = \{1, 3\}$ then your SQL statement should produce the output:

```
 empty_a_minus_b | empty_b_minus_a | empty_a_intersection_b
-----------------+-----------------+-----------------------
 f               | t               | f
(1 row)
```

   because $A - B = \{2\}$, $B - A = \{\}$, and $A \cap B = \{1\}$.

   Your solution should work for arbitrary A and B.

2. (10 points) Let $A(x)$ be the relation schema for a set of positive integers. Write a SQL statement that produces a table that, for each $x \in A$, list the tuple $(x, \sqrt{x}, x^2, 2^x, x!, \ln x)$.

   For example, if $A = \{1, 2, 3, 4, 5\}$ then your SQL statement should produce the following table:

```
  x |  square_root_x   | x_squared | two_to_the_power_x | x_factorial |    logarithm_x
---+------------------+-----------+--------------------+-------------+-------------------
  1 |                1 |         1 |                  2 |           1 |                 0
  2 |  1.4142135623731 |         4 |                  4 |           2 | 0.693147180559945
  3 | 1.73205080756888 |         9 |                  8 |           6 |  1.09861228866811
  4 |                2 |        16 |                 16 |          24 |  1.38629436111989
  5 | 2.23606797749979 |        25 |                 32 |         120 |   1.6094379124341
( 5 rows)
```

3. (10 points) SQL uses 3-valued logic where it concerns the treatment of `NULL` values. (Read your textbook or search the web for relevant information.) Consider relation schemas p(value), q(value), and r(value) where the type of the attribute `value` is `boolean`. Populate each of the relations with the values `true`, `false`, and `NULL` (i.e., the value `unknown`).

   Write a SQL statement that generates the 3-valued truth table for the Propositional Logic formula

$$\neg(p \wedge \neg q) \wedge \neg r$$

   Your statement should return the following answer:

```
 p | q | r | value
---+---+---+-------
 t | t | t | f
 t | t | f | t
 t | t |   |
 t | f | t | f
 t | f | f | f
 t | f |   | f
 t |   | t | f
 t |   | f |
 t |   |   |
 f | t | t | f
 f | t | f | t
 f | t |   |
 f | f | t | f
 f | f | f | t
 f | f |   |
 f |   | t | f
 f |   | f | t
 f |   |   |
   | t | t | f
   | t | f | t
   | t |   |
   | f | t | f
   | f | f |
   | f |   |
   |   | t | f
   |   | f |
   |   |   |
(27 rows)
```

   The blank characters in this table represent the `NULL` (unknown) value.

4. Let $A(x)$, $B(x)$ and C be three unary relation schemas that represent sets A, B and C of integers.

   (a) (10 points) Write a SQL statement that determines whether it is true or not if $A \cap B \neq \{\}$. Provide two answers for this problem. One answer should use the INTERSECT operator, whereas the other answer should not use the INTERSECT operator. For example, if $A = \{1, 2\}$ and $B = \{1, 4, 5\}$ then the result of your statement should be

```
 answer
--------
 t
(1 row)
```

   If, however, $A = \{1, 2\}$ and $B = \{3, 4\}$ then the result of your statement should be

```
 answer
--------
 f
(1 row)
```

   (b) (10 points) Write a SQL statement that determines whether it is true or not if $A \cap B = \{\}$. Provide two answers for this problem. One answer should use the INTERSECT operator, whereas the other answer should not use the INTERSECT operator.

   (c) (10 points) Write a SQL statement that determines whether it is true or not if $A \subseteq B$. Provide two answers for this problem. One answer should use the EXCEPT operator, whereas the other answer should not use the EXCEPT operator.

   (d) (10 points) Write a SQL statement that determines whether it is true or not if $A = B$. Provide two answers for this problem. One answer should use the EXCEPT operator, whereas the other answer should not use the EXCEPT operator.

   (e) (10 points) Write a SQL statement that determines whether it is true or not if $A \neq B$. Provide two answers for this problem. One answer should use the EXCEPT operator, whereas the other answer should not use the EXCEPT operator.

   (f) (10 points) Write a SQL statement that determines whether it is true or not if $|A \cap B| \geq 2$. One answer should use the INTERSECT operator, whereas the other answer should not use the INTERSECT operator.

   (g) (10 points) Write a SQL statement that determines whether it is true or not if $|A \cap B| = 1$.

(h) (10 points) Write a SQL statement that determines whether it is true or not if $(A \cup B) \supseteq C$.

(i) (10 points) Write a SQL statement that determines whether it is true or not if $|(A - B) \cup (B \cap C)| = 2$.

5. (10 points) Consider the relation schema Point(pid, x, y) of a relation of points in the plane. The attribute pid (of type INTEGER) is the identifier of a point, and the attributes x and y, both of type FLOAT, are its x and y coordinates.

   (a) Write a SQL query that returns the pairs of points that are farthest away in distance from each other. Recall that if (x_1, y_1) and (x_2, y_2) are two points in the plane, then the distance between them is given by the formula

   $$\sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2}$$

   For example, if the relation Point is the set of points

```
 pid | x | y
-----+---+---
   1 | 0 | 0
   2 | 0 | 1
   3 | 1 | 0
   4 | 1 | 1
```

   then the result of your query should be

```
 p1 | p2
----+----
  1 |  4
  2 |  3
  3 |  2
  4 |  1
(4 rows)
```

   Notice that the distance between each pair of points in the result is $\sqrt{2} = 1.41421\cdots$.

(b) (10 points) Write a SQL query that returns the pairs of points that are at the next to longest distance away from each other.

For example, if the relation Point is the set of points

```
 pid | x | y
-----+---+---
   1 | 0 | 0
   2 | 0 | 1
   3 | 1 | 0
   4 | 1 | 1
```

then the result of your query should be

```
 p1 | p2
----+----
  1 |  2
  1 |  3
  3 |  4
  2 |  1
  4 |  3
  2 |  4
  4 |  2
  3 |  1
(8 rows)
```

Notice that the distance between each pair of such points is 1.

6. (10 points) Let $W(A, B)$ be a relation schema. The domain of A is INTEGER and the domain of B is VARCHAR(5).

Write a SQL query with returns the A-values of tuples in W if A is a primary key of W. Otherwise, i.e., if A is not a primary key, then your query should return the A-values of tuples in W for which the primary key property is violated. (In this query you should consider creating views for intermediate results.)

For example, consider the following relation instance for W:

W

A	B
1	John
2	Ellen
3	Ann

Then your query should return the following answer since, in this case, A satisfies the primary property for W.

```
a
---
1
2
3
(3 rows)
```

However, if we have the following relation instance for W

W

A	B
1	John
2	Ellen
2	Linda
3	Ann
4	Ann
4	Nick
4	Vince
4	Lisa

then your query should return the following answer because the primary key property of A for W is violated for the A-values 2 and 4.

```
a
---
2
4
(2 rows)
```

2 Writing SQL Queries of Varying Degrees of Difficulty

Use the files student.txt, majors.txt, book.txt, cites.txt, and buys.txt that are provided for this assignment.

Consider the following relation schemas about students and books.

$$\begin{aligned}&\text{Student}(\underline{Sid}, Sname)\\&\text{Major}(\underline{Sid, Major})\\&\text{Book}(\underline{BookNo}, Title, Price)\\&\text{Cites}(\underline{BookNo, CitedBookNo})\\&\text{Buys}(\underline{Sid, BookNo})\end{aligned}$$

The relation Major stores students and their majors. A student can have multiple majors, but we also allow that a student can have no major. A tuple (b, c) in the relation Cites indicates that the book with book number b cites the book with book number c. Note that a book may cite multiple other books. Also, a book does not have to cited.

The primary keys of the relations are the underlined attributes and we assume the following foreign keys:

Attribute in Relation	References Primary Key of Relation
Sid in Major	Sid in Student
BookNo in Cites	BookNo in Book
CitedBookNo in Cites	BookNo in Book
Sid in Buys	Sid in Student
BookNo in Buys	BookNo in Book

Furthermore, assume the following domains for the attributes:

Attribute	Domain
Sid	INTEGER
Sname	VARCHAR(15)
Major	VARCHAR(15)
BookNo	INTEGER
Title	VARCHAR(30)
Price	INTEGER
CitedBookNo	INTEGER

To do this assignment, you will have to create the above relations, including the primary and foreign keys, using the given domain types. You will then also have to copy in the data given in the .txt files.

Write the following queries in SQL.

7. (10 points) Find the titles of books that cost between \$20 and \$40.

8. (10 points) Find the Sid's and Sname's of students who bought a book that cites another book of a lower price.

9. (10 points) Find the Bookno's of books that are cited by a book (or books) that is (are) itself (themselves) cited by another (other) books.

10. (10 points) Find the BookNo's of books that are not cited by another (other) books.

11. (10 points) Find Sid's and Sname's of students who have at least two majors and who only bought books that were cited by other books.

12. (10 points) Find Sid's and majors of students who did not buy any books that cost less than \$30.

13. (10 points) Find each (s, b) pair where s is the Sid of a student and b is the Bookno of a book whose price is the cheapest among the books bought by that student.

14. (10) Without using the `ALL` predicate, list the Bookno's of the cheapest books.

15. (10 points) Find the triples (s1,s2,b) where s1 and s2 are different Sid's of student and b is the BookNo of a book that was bought by student s1 or by student s2, but not by both students.

16. (10 points) Find the tuples (s1,s2) where s1 and s2 are different Sid's of students and such that student s1 and student s2 bought exactly one book in common.

17. (10 points) Find the Bookno's of books that where bought by all students who major in 'Biology'.

18. (10 points) Find the Bookno's of books that where not bought by any students.

19. (10 points) Find the Bookno's of books that where bought buy all but one student.

20. (10 points) Find the pairs (s1,s2) of Sid's of students such that if student s1 bought a book then student s2 also bought that book.