Instructions

In this assignment, you will write queries for two databases, TPC-H and CHINOOK and some relational algebra. Both databases are hosted on http://cs338-s13.cs.uwaterloo.ca. You can test your answers there. The application there hosts Microsoft SQL server instead of DB2, but provides a simpler interface to make things easier. Most of your code should work fine. Whenever there is something specific for Microsoft SQL Server, it will be explained within the question. If your query returns too many results, the server will show a message that the query is correct, but there are too many results to show. Please note that in this server message, “correct” means simply “syntactically correct”, not necessarily that your answer to the question is correct.

Part I: The TPC-H database

It is a part-supplier database. Customers make orders that contain multiple items (each item on an order has a record in LINEITEM). PARTSUPP shows which suppliers supply which parts.
Write SQL queries to answer the following:

1- What are the names of the customers who have a balance that is greater than $9000 and are in the “HOUSEHOLD” market segment?

2- Which parts are supplied by at least 1 supplier from the region “MIDDLE EAST”? Include the part name in your answer along with the part key. Eliminate any duplicates and sort your answer based on the part key.

3- How many distinct parts are supplied by European suppliers?

4- Which parts are not supplied by any supplier from EUROPE? Include the part name in your answer.

5- Which customers ordered parts ONLY from suppliers in the same region? Include the customer name, phone and region in your answer, and remove any duplicates.

6- What is the highest extended price for parts that had a discount larger than the tax when ordered?

7- The number of orders that had all of their items received in at most 2 weeks of shipment.

HINT1: avoid counting duplicates.
HINT2: In SQL Server, you can compute the date difference with the DATEDIFF() function. It takes 3 parameters: unit, start date, end date.
For example, SELECT DATEDIFF(day, '2013-09-16', '2013-09-23') returns 7 as the answer. You can put fields in the 2nd and 3rd argument.

8- The number of (distinct) customers who did not order any part that has some American supplier.

9- Which nation does the customer with the highest balance come from? HINT: you can find the maximum value with a simple nested select query.

10- List the names of the countries where some parts took more than 29 days to ship, despite being supplied from a supplier to a customer in this same country.
PART II: The CHINOOK database:

The Chinook Database is about an imaginary video and music store. Each track is stored using one of the digital formats and has a genre. The store has also some playlists, where a single track can be part of several playlists. Just like TPC-H, orders are recorded for customers, but are called invoices. LINEITEM is replaced by INVOICELINE. Every customer is assigned a support employee, and Employees report to other employees.

Write SQL queries to answer the following:

11- Which artists did not make any albums at all? Include their names in your answer.

12- Which artists did not record any tracks of the Latin genre?

13- Which video track has the longest length?
14- Find the names of customers who live in the same city as the top employee (The one not managed by anyone).

15- Find the managers of employees supporting Brazilian customers.

16- How many audio tracks in total were bought by German customers? And what was the total price paid for them?

17- Which playlists have no Latin tracks?

18- What is the space, in bytes, occupied by the playlist “Grunge”, and how much would it cost? (Assume that the cost of a playlist is the sum of the price of its constituent tracks).

19- Which playlists do not contain any tracks for the artists “Black Sabbath” nor “Chico Buarque”?

20- List the names and the countries of those customers who are supported by an employee who was younger than 35 when hired. HINT: use year as the first parameter in DATEDIFF().

PART III: Relational Algebra

Write the relational algebra for the following queries:
1, 2, 4, 5, 10, 11, 14, 15, 20