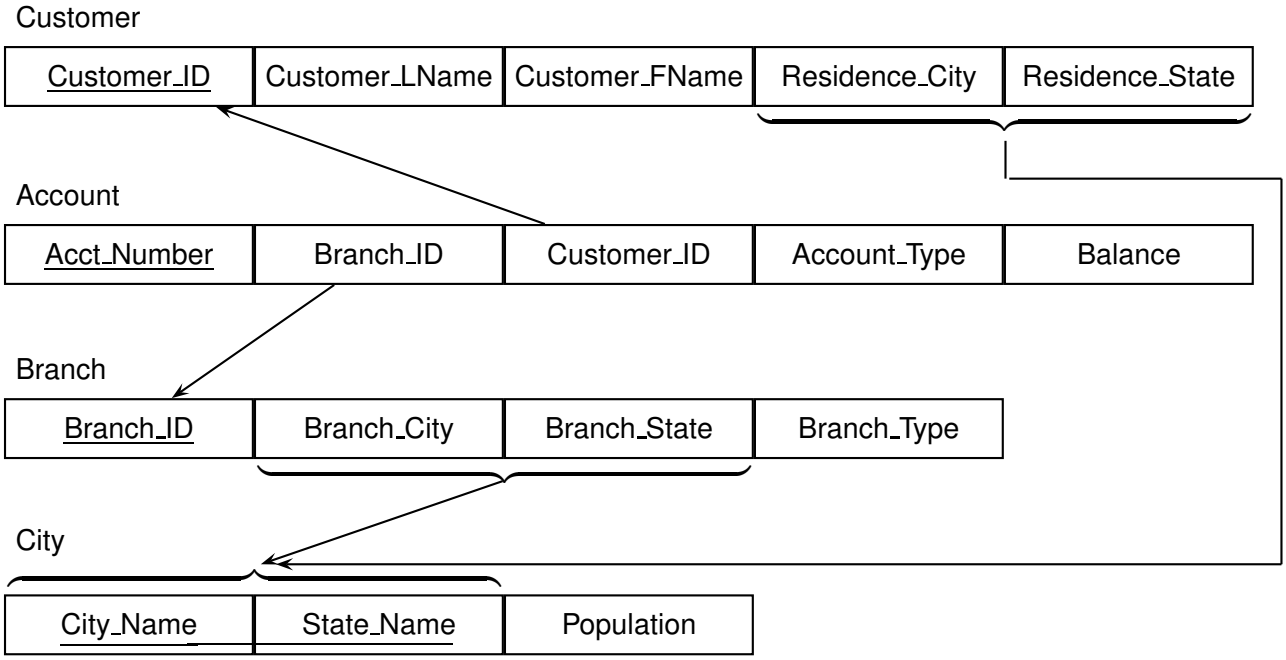
5DV021 Principles of Database Systems
Further Class Exercises on SQL

Solutions to an SQL query from the Examination of December 21, 2006 which requires some special techniques.



Problem 10:

(b) For each (City_Name,State_Name) pair in the City relation, list the total number of customers who live in the city defined by that (City_Name,State_Name) pair, the total number of accounts which are held by such customers, and the average balance over all accounts held by such customers. If there are no accounts associated with a given (City_Name,State_Name) pair, list the average as zero.

Corrected: September 14, 2009

5DV021, Further SQL exercises, page 2

First, here is a try at a solution.

```
Select Residence_City, Residence_State,
       count(distinct Customer.Customer_ID) as N_Cust,
       count(Acct_Number) as N_Acct, avg(Balance) as Avg_Bal
From   Customer, Account
Where  (Customer.Customer_ID=Account.Customer_ID)
Group by Residence_City, Residence_State
Union
Select City_Name, State_Name, 0, 0, 0
From   City
Where  Not Exists
        (Select *
         From Customer
         Where (City_Name=Residence_City) and
               (State_Name=Residence_State));
```

- It works as long as each customer has at least one account, but it misses customers who have no accounts.

5DV021, Further SQL exercises, page 3

- To solve this query correctly, an approach is to solve two separate queries and join the results together.

- The first query finds the number of customers in each city.

```
Select Residence_City, Residence_State,
       count(distinct Customer.Customer_ID) as N_Cust
From   Customer
Group by Residence_City, Residence_State
Union
Select City_Name, State_Name, 0
From City
Where  Not Exists
        (Select *
         From Customer
         Where (City_Name=Residence_City) and
               (State_Name=Residence_State));
```

- The second query finds the number of accounts and average balance for each city.

```
Select Residence_City, Residence_State,
       count(Acct_Number) as N_Acct, avg(Balance) as Avg_Bal
From   Customer, Account
Where  (Customer.Customer_ID=Account.Customer_ID)
Group by Residence_City, Residence_State
Union
Select City_Name, State_Name, 0, 0
From   City
Where  Not Exists
        (Select *
         From Customer, Account
         Where (Customer.Customer_ID=Account.Customer_ID) and
               (Customer.Residence_City=City_Name) and
               (Customer.Residence_State=State_Name));
```

How can these be glued together?

- A simple way is to generate two temporary results and then glue them together.

```
Select Residence_City, Residence_State,
       count(distinct Customer.Customer_ID) as N_Cust
Into Temporary T1
From   Customer
Group by Residence_City, Residence_State
Union
Select City_Name, State_Name, 0
From City
Where  Not Exists
        (Select *
         From Customer
         Where (City_Name=Residence_City) and
               (State_Name=Residence_State));

Select Residence_City, Residence_State,
       count(Acct_Number) as N_Acct, avg(Balance) as Avg_Bal
Into Temporary T2
From   Customer, Account
Where  (Customer.Customer_ID=Account.Customer_ID)
Group by Residence_City, Residence_State
Union
Select City_Name, State_Name, 0, 0
From   City
Where  Not Exists
        (Select *
         From Customer, Account
         Where (Customer.Customer_ID=Account.Customer_ID) and
               (Customer.Residence_City=City_Name) and
               (Customer.Residence_State=State_Name));

Select TA.Residence_City, TA.Residence_State, N_Cust, N_Acct, Avg_BAl
From   T1 as TA, T2 as TB
Where  (TA.Residence_City=TB.Residence_City) and
       (TA.Residence_State=TB.Residence_State);
```

Shortcomings:

- Not standard SQL.
- Side effect of temporary tables.
- Solutions involving temporary tables will not be allowed on the examination.

Corrected: September 14, 2009

These can be combined into one query by using subqueries in the `From` clause.

```
Select RC1 as City, RS1 as State, N_Cust, N_Acct, Avg_Bal
From
     (Select Residence_City as RC1, Residence_State as RS1,
            count(distinct Customer.Customer_ID) as N_Cust
      From   Customer
      Group by Residence_City, Residence_State
      Union
      Select City_Name, State_Name, 0
      From City
      Where  Not Exists
              (Select *
               From Customer
               Where (City_Name=Residence_City) and
                           (State_Name=Residence_State))) as Pointless_1,
     (Select Residence_City as RC2, Residence_State as RS2,
            count(Acct_Number) as N_Acct, avg(Balance) as Avg_Bal
      From   Customer, Account
      Where  Customer.Customer_ID=Account.Customer_ID
      Group by Residence_City, Residence_State
      Union
      Select City_Name, State_Name, 0, 0
      From   City
      Where  Not Exists
             (Select *
              From Customer, Account
              Where (Customer.Customer_ID=Account.Customer_ID) and
                    (Customer.Residence_City=City_Name) and
                    (Customer.Residence_State=State_Name)))
                                                        as Pointless_2
Where (RC1=RC2) and (RS1=RS2);
```

Warning:

- Solutions involving subqueries in the `From` clause may not be accepted on the examination.
- These solutions are shown for illustration of advanced techniques in SQL only.