STEVENS INSTITUTE OF TECHNOLOGY

Stevens Institute of Technology
Castle Point on Hudson
Hoboken, NJ 07030

YOUR NAME: Yu Zhang

2

1. (30 pts.) Express the following ad-hoc OLAP query in SQL (similar to the ones from the programming assignment).

Please indicate how many scans of the table are required to process the query, and justify your answers as to why you think the numbers of scans are required for the given query.

Use the schema, **sales (prod, cust, yr, mo, day, state, quant)**

Query: For 2009, show for each customer and product, the *total* sales quantities for the 4 quarters (in four columns)

Example

CUSTOMER	PRODUCT	Q1_TOTAL	Q2_TOTAL	Q3_TOTAL	Q4_TOTAL
Bloom	Pepsi	328923	243241	231873	42325
Emily	Milk	14239	9872	12142	2435
Helen	Apple	4235	13242	3128	872

```
create view V1 as
select cust, prod, sum(quant) as Q1
from sales
where yr=2009 and mo between 1 and 3
group by cust, prod

create view V2 as
select cust, prod, sum(quant) as Q2
from sales
where yr=2009 and mo between 4 and 6
group by cust, prod

create view V3 as
select cust, prod, sum(quant) as Q3
from sales
where yr=2009 and mo between 7 and 9
group by cust, prod

create view V4 as
select cust, prod, sum(quant) as Q4
from sales
where yr=2009 and mo between 10 and 12
group by cust, prod

select *
from V1 natural full outer join V2
     natural full outer join V3
     natural full outer join V4
using cust, prod
```

I think 4 scans of the table are required to process the query. For the given query, one output of a quarter needs one scan of the table. Thus, for Q1, Q2, Q3, Q4, 4 scans of the table are needed to get the output.

STEVENS INSTITUTE OF TECHNOLOGY

Stevens Institute of Technology
Castle Point on Hudson
Hoboken, NJ 07030

YOUR NAME: Xu Zhang

18

2. (20 points) Provide an expression in the relational algebra to express each of the following queries. Use the following relational database (5 points each):

employee (person-name, street, city)
works (person-name, company-name, salary)
company (company-name, city)
manages (person-name, manager-name)

5. a) Find the names, street addresses, and cities of residence of all employees who work for First Bank Corporation and earn more than $10,000.

$\Pi_{person\text{-}name,\ street,\ city}(\sigma_{(company\text{-}name = \text{"First Bank Corporation"} \wedge salary > 10000)}(works \bowtie employee))$

3. b) Find the names of all employees who live in the same city and on the same street as do their managers.

$t1 \leftarrow \Pi_{person\text{-}name,\ street,\ city}(\sigma_{employee.person\text{-}name = manages.manager\text{-}name}(employee \times manages))$

$\Pi_{person\text{-}name}(\rho_{manager\_address(person\text{-}name,\ street,\ city)}(t1) \bowtie employee)$

5. c) Find the names of all employees who earn more than every employee of Small Bank Corporation.

$t1 \leftarrow max_{salary}(\sigma_{company\text{-}name = \text{"Small Bank Corporation"}}(works))$

$\Pi_{person\text{-}name}(\sigma_{works.salary > t1.salary}(works \times t1))$

5 d) Find the company with the most employees

$t1 \leftarrow {}_{company\text{-}name}\mathcal{G}_{count\text{-}distinct\ person\text{-}name}(employee)$

$t2 \leftarrow max_{num\text{-}employee}(\rho_{t2(company\text{-}name,\ num\_employee)}(t1))$

$\Pi_{company\text{-}name}(\rho_{t3(company\text{-}name,\ num\text{-}employee)}(t1) \bowtie \rho_{t4(num\_employee)}(t2))$

STEVENS INSTITUTE of TECHNOLOGY

Stevens Institute of Technology
Castle Point on Hudson
Hoboken, NJ 07030

YOUR NAME: Xu Zhang

14

3. (15 points) Answer the following questions

4 a) (4 points) Define the relational operator, 'intersection' using another relational operator.

$r \cap s = r - (r - s)$

2 b) (3 points) Describe the following expression in English.

$\Pi_{\text{customer\_name}}$ ($\sigma_{\text{branch\_name = "Brooklyn" and customer\_name like "\%son"}}$ (
$\sigma_{\text{borrower.loan\_number = loan.loan\_number}}$ (borrower x loan)))

Find ~~the~~ all customers ~~whose names include the string "son"~~ who ~~and~~ have a loan at the Brooklyn branch and their names include the string "son".

8 end in c) (8 points) The expression above requires a full Cartesian product between "borrower" and "loan". Provide a relational expression equivalent to the one above (in (b)) but more "efficient".

$\Pi_{\text{customer\_name}}$ ($\sigma_{\text{borrower.loan\_number = loan.loan\_number}}$
(($\sigma_{\text{branch\_name = "Brooklyn"}}$ (loan)) X ($\sigma_{\text{customer\_name like "\%son"}}$ (borrower))))

STEVENS INSTITUTE OF TECHNOLOGY

Stevens Institute of Technology
Castle Point on Hudson
Hoboken, NJ 07030

YOUR NAME: Xu Zhang

10 **4.** (10 points) Express the following SQL queries in English:

5 a) (5 points)

```
select distinct S.customer_name
        from borrower as S
        where exists (
                (select branch_name
                from branch
                where branch_city = 'Perry Ridge')
                except
                (select R.branch_name
                from borrower as T, loan as R
                where T.loan_number = R.loan_number and
                        S.customer_name = T.customer_name ))
```

Find all customers who donot have a loan at the Perry Ridge branch.

5 b) (5 points)

```
select T.customer-name
from depositor as T
where not exists (
        select R.customer-name
        from account, depositor as R
        where T.customer-name = R.customer-name and
                R.account-number = account.account-number and
                account.branch-name in ('Perryridge', 'Brooklyn'))
```

Find all customers who donot have account in the Perryridge and Brooklyn branch.

STEVENS INSTITUTE OF TECHNOLOGY

Stevens Institute of Technology
Castle Point on Hudson
Hoboken, NJ 07030

YOUR NAME: Xu Zhang

5. (10 points) Answer the questions below regarding the following SQL query.

```
select branch_name, avg (balance)
  from account
group by branch_name
having avg(balance) > 1200
```

a) (3 points) Express the query in English.

Find all branches with their average balance where the average account balance is more than 1200.

b) (7 points) Express the above query without using a 'having' clause

```
Select branch_name, avg_balance
from
    (select branach_name, avg(balance)
     from account
     group by branch_name
    ) as branch_avg ( branch_name, avg_balance)

where avg_balance > 1200
```

STEVENS INSTITUTE OF TECHNOLOGY

Stevens Institute of Technology
Castle Point on Hudson
Hoboken, NJ 07030

YOUR NAME: Xu Zhang

15

6. (15 points) Given the following 2 relations, 'loan' and 'borrower', provide the results to the following join queries – draw the resulting relations (3 points each).

loan:

branch-name	*loan-number*	*amount*
Downtown	L-170	3000
Redwood	L-230	4000
Perryridge	L-260	1700

borrower:

customer-name	*loan-number*
Jones	L-170
Smith	L-230
Hayes	L-155

a) loan **inner join** borrower **on** loan.loan-number = borrower.loan-number

branch-name	loan-number	amount	customer-name	loan-number
Downtown	L-170	3000	Jones	L-170
Redwood	L-230	4000	Smith	L-230

b) loan **left outer join** borrower **on** loan.loan-number=borrower.loan-number

branch-name	loan-number	amount	customer-name	loan-number
Downtown	L-170	3000	Jones	L-170
Redwood	L-230	4000	Smith	L-230
Perryridge	L-260	1700	null	null

c) loan **natural inner join** borrower

branch-name	loan-number	amount	customer-name
Downtown	L-170	3000	Jones
Redwood	L-230	4000	Smith

d) loan **natural right outer join** borrower

branch-name	loan-number	amount	customer-name
Downtown	L-170	3000	Jones
Redwood	L-230	4000	Smith
null	L-155	null	Hayes

~~loan-number~~

e) loan **full outer join** borrower **using** (loan-number)

branch-name	loan-number	amount	customer-name
Downtown	L-170	3000	Jones
Redwood	L-230	4000	Smith
Perryridge	L-260	1700	Hayes
null	L-155	null	null