

Example Answers for Assignment 1

1. Explain why databases need to consider null values, and how out join handles this.

In databases, some values for a particular attribute do not exist or are unknown. We only can put null values there. The outer join captures the entities that would otherwise be lost by filling in null values for the attributes from the other relation. Specifically a left outer join will retain the tuples that would be lost from the left relation, a right outer join retains the tuples that would be lost from the right, and a full outer join captures both.

2. Using the bank example in the class (also in the textbook), write relational-algebra queries to find the smallest balance for the accounts in Downtown (branch) using aggregate functions and without using aggregate functions.

Using Aggregate functions: $G_{\min(\text{balance})}(\sigma_{\text{branch_name} = \text{Downtown}}(\text{account}))$

Without Using Aggregate functions:

$$\begin{aligned} t1 &\leftarrow \sigma_{\text{branch_name} = \text{Downtown}}(\text{account}) \\ t2 &\leftarrow \Pi_{t1.\text{balance}}(\sigma_{t1.\text{balance} > d.\text{balance}}(t1 \times \rho_d(t1))) \\ \text{result} &\leftarrow \Pi_{\text{balance}}(t1) - t2 \end{aligned}$$

3. Give expressions in relational algebra for the following expressions:

3.1 The name and city of all of the people who are managed by Bob Smith or John Douglas.

$$\Pi_{\text{person_name}, \text{employee.city}}(\sigma_{\text{manages.manager_name} = \text{Bob Smith} \vee \text{manages.manager_name} = \text{John Douglas}}(\text{employee} \bowtie \text{manages}))$$

3.2 Find the name, street, and city of people who make > \$45,000 and work in a company located in Eugene.

$$\Pi_{\text{person_name}, \text{employee.street}, \text{employee.city}}(\sigma_{\text{company.city} = \text{Eugene} \wedge \text{salary} > 45000}(\text{employee} \bowtie \text{works} \bowtie \text{company}))$$

3.3 A company may locate in different cities. Find cities that have every company that Eugene has.

$$company \div \Pi_{company_name} (\sigma_{city = Eugene} (company))$$

4. The management wants you to make some changes to the data related to First Bank Corporation by writing relational queries.

4.1 Give all employees managed by John Douglas a 2% raise in salary.

$$t1 \leftarrow \Pi_{person_name, company_name, salary} (\sigma_{manager_name = John\ Douglas} (works \bowtie manages))$$

$$works \leftarrow (works - t1) \cup (\Pi_{person_name, company_name, salary * 1.02} (t1))$$

4.2 Give John Douglas a 10% raise in salary.

$$t1 = \sigma_{person_name = John\ Douglas} (works)$$

$$works \leftarrow (works - t1) \cup (\Pi_{person_name, company_name, salary * 1.1} (t1))$$

4.3 Give all other managers (except John Douglas) a 5% deduction. If, as a result of the deduction, the manager will make less than \$100,000 give the manager a 2% deduction, instead.

$$t1 \leftarrow works \bowtie \Pi_{manager_name \text{ as } person_name} (\sigma_{manager_name \neq John\ Douglas} (manages))$$

$$t2 \leftarrow \Pi_{person_name, company_name, salary * 0.95} (\sigma_{salary * 0.95 \geq 100000} (t1))$$

$$t3 \leftarrow \Pi_{person_name, company_name, salary * 0.98} (t1 - (t1 \bowtie \Pi_{person_name} (t2)))$$

$$works \leftarrow (works - t1) \cup t2 \cup t3$$

4.4 Bob Smith is retired and his group is laid off. To management decides to delete the information of Bob Smith and the employees he managed from the company database.

$$t \leftarrow \Pi_{person_name} (\sigma_{manager_name = Bob\ Smith} (manages) \cup \sigma_{person_name = Bob\ Smith} (manages))$$

$$employee \leftarrow employee - (t \bowtie employee)$$

$$works \leftarrow works - (t \bowtie works)$$

$$manages \leftarrow manages - (t \bowtie manages)$$

5. Complete the following relational-algebra expressions using aggregate functions:

5.1 Find the company located in the biggest number of cities.

$t1 \leftarrow \Pi_{company_name} \text{Gcount-distinct (city) as number_cities (company)}$

$maximum \leftarrow \text{Gmax (number_cities) as max-city-count (t1)}$

$result \leftarrow \Pi_{company_name} (\sigma_{number_cities = max-city-count} (t1 \times maximum))$

5.2 Find the company with the smallest payroll in Eugene.

$t1 \leftarrow \Pi_{company_name} \text{Gsum (salary) as sum_salary} (\Pi_{company_name} (\sigma_{city = Eugene} (company)) \bowtie works)$

$minimum \leftarrow \text{Gmin (sum_salary) as min-sum-salary (t1)}$

$result \leftarrow \Pi_{company_name} (\sigma_{sum_salary = min-sum-salary} (t1 \times minimum))$

5.3 For those companies in Eugene or Springfield, find the company whose employees earn a higher average salary than the average salary of Symantec in Springfield.

$Average_symantec \leftarrow \text{G}_{avg (salary) as avg_symantec} (\sigma_{company_name = Symantec} (works))$

$t1 \leftarrow \Pi_{company_name} (\sigma_{city = Eugene \vee city = Springfield} (company)) \bowtie works)$

$t2 \leftarrow \Pi_{company_name} \text{G}_{avg (salary) as avg_salary} (t1)$

$result \leftarrow \Pi_{company_name} (\sigma_{avg_salary > avg_symantec} (t2 \times Average_symantec))$

