Part 1 - Select, Project, Join

1. Statement A: It can be useful to compose two selection operators.
   For example:

   $$\sigma_{state='CA'}(\sigma_{enr>8000}\text{College})$$

   Statement B: It can be useful to compose two projection operators.
   For example:

   $$\pi_{GPA}(\pi_{sID,GPA,HS}\text{Student})$$

   ANSWER: Both statements are false.

   EXPLANATION: Two selection operators in a row can always be replaced by a single selection operator whose condition is the "and" of the two selection conditions. If there are two projection operators in a row, the attribute list of the second (outer) projection must be a subset of the attribute list of the first (inner) projection. Thus, the first projection can be removed without changing the result of the expression.

   ---

2. Which of the following expressons does NOT return the names and GPAs of students with HS>1000 who applied to CS and were rejected?

   A) $\pi_{sName,GPA}(\sigma_{Student.sID=Apply.sID}(\sigma_{HS>1000}(Student) \times \sigma_{major='CS'\wedge dec='R'}(Apply)))$

   B) $\pi_{sName,GPA}(\sigma_{Student.sID=Apply.sID\wedge HS>1000\wedge major='CS'\wedge dec='R'}(Student \times \pi_{sID,major,dec}Apply))$

   C) $\sigma_{Student.sID=Apply.sID}(\pi_{sName,GPA}(\sigma_{HS>1000}Student \times \sigma_{major='CS'\wedge dec='R'}Apply))$

   ANSWER: C

   EXPLANATION: The first two expressions are equivalent to the expression given in the video for this query. The third expression is invalid because the sID attributes are not included in the result of the expression to which condition Student.sID=Apply.sID is applied.

   ---

3. Which of the following English sentences describes the result of this expression:

   $$\pi_{sName,cName}(\sigma_{HS>enr}(\sigma_{state='CA'}College \bowtie Student \bowtie \sigma_{major='CS'}Apply))$$

   A) All student-college name pairs, where the student is applying to major in CS at the college, the college is in California, and the college is smaller than some high school

   B) Students paired with all California colleges to which the student applied to major in CS, where at least one of those colleges is smaller than the student's high school

   C) Students paired with all colleges smaller than the student's high school to which the student applied to major in CS, where at least one of those colleges is in California

D) Students paired with all California colleges smaller than the student's high school to which the student applied to major in CS

ANSWER: D

EXPLANATION: The inner natural join connects students with the colleges to which they've applied, allowing only California colleges and CS-major applicaitons. The outer selection condition filters out all applications except those where the high school is bigger than the college, and the final projection keeps the student and college names.

---

4. Which of the following expressions finds the IDs of all students such that some college bears the student's name?

A) $\pi_{sID}(College \bowtie Student)$

B) $\pi_{sID}(\sigma_{cName=sName}(College \times Student))$

C) $\pi_{sID}(\pi_{cName}College \bowtie \pi_{cName}(\sigma_{sName=cName}Student))$

D) $\pi_{sID}(\sigma_{cName=sName}(\pi_{sID}Student \times College \times Student)$

ANSWER: B

EXPLANATION: The first choice returns the IDs of all students in the database. The third choice is invalid because cName is not an attribute of Student. The fourth choice yields all sIDs in Student.

Part 2 - Set Operators, Renaming, Notation

1. Three of the following four expressions finds the names of all students who did not apply to major in CS or EE. Which one finds something different? (Hint: You should not assume student names are unique.)

   A) $\pi_{sName}(Student \bowtie (\pi_{sID}Student - (\pi_{sID}(\sigma_{major='CS'}Apply) \cup \pi_{sID}(\sigma_{major='EE'}Apply))))$

   B) $\pi_{sName}(Student \bowtie (\pi_{sID}Student - \pi_{sID}(\sigma_{major='CS'\lor major='EE'}Apply)))$

   C) $\pi_{sName}(\pi_{sID,sName}Student - \pi_{sID,sName}(Student \bowtie \pi_{sID}(\sigma_{major='CS'\lor major='EE'}Apply)))$

   D) $\pi_{sName}Student - \pi_{sName}(Student \bowtie \pi_{sID}(\sigma_{major='CS'\lor major='EE'}Apply))$

   ANSWER: D

   EXPLANATION: If there are two students named Susan, one who did not apply to CS or EE, and one who did, then 'Susan' will correctly be included in the result of the first three expressions, but will incorrectly be omitted from the result of the last one.

   ---

2. Which of the following English sentences describes the result of this expression:

   $$\pi_{cName}College - \pi_{cName}(Apply \bowtie (\pi_{sID}(\sigma_{GPA>3.5}Student) \cap \pi_{sID}(\sigma_{major='CS'}Apply)))$$

   A) All colleges with no GPA>3.5 applicants who applied for a CS major at that college

   B) All colleges with no GPA>3.5 applicants who applied for a CS major at any college

   C) All colleges where all applicants either have GPA>3.5 or applied for a CS major at that college

   D) All colleges where no applicants have GPA>3.5 or no applicants applied for a CS major at that college

   ---

   ANSWER: B

   EXPLANATION: The intersection finds the IDs of all students who have GPA>3.5 and applied for a CS major at any college. The natural-join with Apply finds all colleges those students applied for. The minus finds all colleges except the ones those students applied for.

   ---

3. Suppose relation Student has 20 tuples. What is the minimum and maximum number of tuples in the result of this expression:

   $$\rho_{s1(i1,n1,g,h)}Student \bowtie \rho_{s2(i2,n2,g,h)}Student$$

   A) minimum = 0, maximum = 400

   B) minimum = 20, maximum = 20

   C) minimum = 20, maximum = 400

   D) minimum = 40, maximum = 40

ANSWER: C

EXPLANATION: If every student has a unique GPA-HS combination, then students join only with themselves, and there are 20 tuples in the result (minimum). If every student has the same GPA and HS as every other student, then all pairs join and the result has $20 * 20 = 400$ tuples.

---

4. Suppose relations College, Student, and Apply have 5, 20, and 50 tuples in them respectively. Remember that cName is a key for College. Do not assume sName is a key for Student. Do assume that college names in Apply also appear in College. What is the minimum and maximum number of tuples in the result of this expression:

$$\pi_{cName}College \cup \rho_{cName}(\pi_{sName}Student) \cup \pi_{cName}Apply$$

A) minimum = 5, maximum = 25

B) minimum = 5, maximum = 75

C) minimum = 25, maximum = 45

D) minimum = 75, maximum = 75

ANSWER: A

EXPLANATION: Recall that duplicates are eliminated automatically in relational algebra. If all students have names that are also college names, then there are only 5 names altogether (minimum). If every student has a unique name and none of them are college names, then there are 5+20=25 names altogether (maximum). Since all college names in Apply are also in College, the third term of the expression cannot add any new names.