Question 1. [6 MARKS]

Part (a) [4 MARKS]

Suppose that

- $R1$ is a relation with $t1$ tuples and $a1$ attributes.
- $R2$ is a relation with $t2$ tuples and $a2$ attributes.
- L is a list of n attributes.
- c is a boolean expression involving the attributes of $R1$.

Assume that the expressions below are legal expressions of relational algebra. Fill in the table to indicate the size of the relation that is the result of each expression.

Solution:

	Number of tuples		Number of attributes	
Expression	minimum	maximum	minimum	maximum
$\Pi_L R1$	1	$t1$	n	n
σ_c R1	0	$t1$	$a1$	$a1$
$R1XR2$	$t1 \times t2$	$t1 \times t2$	$a1 + a2$	$a1 + a2$
$R1 \bowtie R2$	0	$t1 \times t2$	$\max(a1, a2)$	$a1 + a2$

Part (b) [2 MARKS]

Suppose R and S are relations. Which of the following statements are true? Circle one answer for each. **Do not guess.** There is 1 point for each correct answer, -1 for each incorrect answer, and 0 points if you leave the answer blank.

1. If R and S have no attributes in common, $R \times S = R \bowtie S$.

   True False

2. If R and S have at least one attribute in common, it cannot be true that $R \times S = R \bowtie S$.

   True False

Solution:

Part 1 is true.

Part 2 is also true, but only because the schemas of the two relations are necessarily different: $R \times S$ includes each common attribute twice, while $R \bowtie S$ does not. But it *is* possible to have an R and an S where the tuples that are included are the same. This occurs, for example, if on each common attribute, both relations have a single value and it's the same value.

Question 2. [8 marks]

Consider the following database:

P

A	B	C
1	5	1
4	6	5
2	8	1
3	4	1
1	2	3
3	3	2

Q

C	D
1	5
2	6

Assuming set semantics, give the result (schema and data) returned by the following queries. Use the same tabular format as above; do **not** describe the result in English.

Part (a) [2 marks]

$(\Pi_C P - \Pi_C Q) \cap (\Pi_C P - \Pi_C(P \bowtie \sigma_{D=5} Q))$

Solution:

C
5
3

Part (b) [2 marks]

$T := \sigma_{P1.A<P2.A \wedge P1.C=P2.C}(\rho_{P1}(P) \times \rho_{P2}(P))$

$Answer := \Pi_C P - \Pi_{P1.C} T$

Solution:

Strictly speaking, this query is ill-formed. The left operand of the set difference has attribute C, whereas the right operand has attribute $P1.C$. A rename would have fixed this.

C
5
3
2

Part (c) [2 MARKS]

$T := (\Pi_A P \times \Pi_C Q) - (\Pi_{A,C}(P \bowtie Q))$

$Answer := \Pi_A P - \Pi_A T$

Solution:

A
3

Part (d) [2 MARKS]

$P1(A, B, C) := P$

$P2(A, B, C) := P$

$T := \sigma_{P1.C=P2.C \wedge P1.B>P2.B}(P1 \times P2)$

$Answer := P - \Pi_{P1.A,P1.B,P1.C} T$

Solution:

Again, strictly speaking, this query is ill-formed. The left operand of the set difference has attribute A, B, C, whereas the right operand has attribute $P1.A, P1.B, P1.C$. A rename would have fixed this.

A	B	C
4	5	6
3	4	1
1	2	3
5	3	2

Question 3. [10 MARKS]

Consider the following schema for a hair salon. Keys are underlined.

- Clients(<u>CID</u>, name, phone).
  CID is the ID of a client, *name* and *phone* are their name and phone number.
- Staff(<u>SID</u>, name).
  SID is the ID of a staff member and *name* is their name.
- Appointments(<u>CID, date, time</u>, service, SID)
  CID is the ID of the client whose appointment it is, *date* and *time* indicate when the appointment happens, *service* is the name of the service they have at this appointment, and *SID* is the ID of the staff member providing the service for this apppointment. *CID* is a foreign key on Clients and *SID* is a foreign key on Staff. That is, the following inclusion dependencies hold:
  Appointments[CID] $\subseteq$ Clients[CID], and
  Appointments[SID] $\subseteq$ Staff[SID].

Which of the following queries correctly find the name of every client who has not had a haircut in 2010? Circle one answer for each. **Do not guess.** There are 2 points for each correct answer, -1 for each incorrect answer, and 0 points if you leave the answer blank.

1. $A := (\Pi_{CID} Clients) - (\Pi_{CID}(\sigma_{date.year=2010 \wedge service=\text{``haircut''}} Appointments))$
   $Answer := \Pi_{name}(A \bowtie Clients)$

   $\boxed{\text{Correct}}$ Incorrect

2. $A := (\Pi_{CID,name} Clients) - (\Pi_{CID,name}(\sigma_{date.year=2010 \wedge service=\text{``haircut''}}(Clients \bowtie Appointments)))$
   $Answer := \Pi_{name} A$

   $\boxed{\text{Correct}}$ Incorrect

3. $A := (Clients \bowtie Appointments) - (\sigma_{date.year=2010 \wedge service=\text{``haircut''}}(Clients \bowtie Appointments))$
   $Answer := \Pi_{name} A$

   Correct $\boxed{\text{Incorrect}}$

   Note: This query would inappropriately omit clients who have never had an appointment.

4. $A := (\Pi_{CID} Clients) - (\Pi_{CID} \sigma_{date.year=2010 \wedge service \neq \text{``haircut''}} Appointments)$
   $Answer := \Pi_{name}(A \bowtie Clients)$

   Correct $\boxed{\text{Incorrect}}$

5. $A := (\Pi_{CID} \sigma_{date.year=2010 \wedge service \neq \text{``haircut''}} Appointments)$
   $Answer := \Pi_{name}(A \bowtie Clients)$

   Correct $\boxed{\text{Incorrect}}$

Question 4. [8 marks]

This question assumes the same schema as for question 3.

Write the following queries using only the basic Relational Algebra operators $\Pi, \sigma, \bowtie, \times, \cap, \cup, -, \rho$. Assume the set semantics (not bag semantics) for Relational Algebra.

1. CID of all clients who have never had an appointment for both a haircut and another, different, service on the same date.

   Solution:

   $Pairs := \rho_{A1}(Appointments) \times \rho_{A2}(Appointments)$
   $Have(CID) := \Pi_{A1.CID}(\sigma_{A1.CID=A2.CID \wedge A1.service=\text{"}haircut\text{"} \wedge A2.service \neq \text{"}haircut \wedge A1.date=A2.date}(Pairs))$
   $Answer := (\Pi_{CID} Clients) - Have$

2. Name and phone number of the client who had staff member Guilano's first appointment.

   Solution:

   $Guiliano := \Pi_{CID,date,time} \sigma_{name=\text{"}Guiliano\text{"}}(Appointments \bowtie Staff)$
   $Pairs := \rho_{G1}(Giuliano) \times \rho_{G2}(Giuliano)$
   $Beaten(CID, date, time) := \Pi_{G1.CID,G1.date,G1.time}(\sigma_{G1.date>G2.date \vee (G1.date=G2.date \wedge G1.time>G2.time)}(Pairs))$
   $Answer := \Pi_{name}((Giuliano - Beaten) \bowtie Clients)$