In-class Exercise Relational Algebra
Consider the following schema:

Suppliers(*sid:* integer, *sname:* string, *address:* string)
Parts(*pid:* integer, *pname:* string, *color:* string)
Catalog(*sid:* integer, *pid:* integer, *cost:* real)

The key fields are underlined, and the domain of each field is listed after the field name. Therefore *sid* is the key for Suppliers, *pid* is the key for Parts, and *sid* and *pid* together form the key for Catalog. The Catalog relation lists the prices charged for parts by Suppliers. Write the following queries in relational algebra.

1. Find the *name*s of suppliers who supply some red part.

2. Find the *sid*s of suppliers who supply some red or green part.

3. Find the *sid*s of suppliers who supply some red part or are at 221 Packer Street.

4. Find the *sid*s of suppliers who supply some red part and some green part.

5. Find the *sid*s of suppliers who supply every part.

6. Find the *sid*s of suppliers who supply every red part.

7. Find the *sid*s of suppliers who supply every red or green part.

8. Find the *sid*s of suppliers who supply every red part or supply every green part.

9. Find pairs of *sid*s such that the supplier with the first *sid* charges more for some part than the supplier with the second *sid*.

10. Find the *pid*s of parts supplied by at least two different suppliers.

11. Find the *pid*s of the most expensive parts supplied by suppliers named Yosemite Sham.

Consider the following schema:

Suppliers(*sid:* integer, *sname:* string, *address:* string)
Parts(*pid:* integer, *pname:* string, *color:* string)
Catalog(*sid:* integer, *pid:* integer, *cost:* real)

The key fields are underlined, and the domain of each field is listed after the field name. Therefore *sid* is the key for suppliers, *pid* is the key for Parts, and *sid* and *pid* together form the key for Catalog. The Catalog relation lists the prices charged for parts by Suppliers. Write the following queries in relational algebra.

1. Find the *name*s of suppliers who supply some red part.

$\pi_{sname}(\pi_{sid}((\pi_{pid}\sigma_{color=\_'red'\_}Parts) \bowtie Catalog) \bowtie Suppliers)$

Catalog

SID	PID	Cost
1	1	$10.00
1	2	$20.00
1	3	$30.00
1	4	$40.00
1	5	$50.00
2	1	$9.00
2	3	$34.00
2	5	$48.00

Parts

PID	Pname	Color
1	Red1	Red
2	Red2	Red
3	Green1	Green
4	Blue1	Blue
5	Red3	Red

SID	Sname	Address
1	Yosemite Sham	Devil's canyon, AZ
2	Wiley E. Coyote	RR Asylum, NV
3	Elmer Fudd	Carrot Patch, MN

Ok, let's break this down.

$\sigma_{color='red'}Parts$ *gives us*

PID	Pname	Color
1	Red1	Red
2	Red2	Red
5	Red3	Red

$\pi_{pid}\sigma_{color='red'}Parts$ gives us

PID
1
2
5

$((\pi_{pid}\sigma_{color='red'}Parts) \bowtie Catalog)$ gives us:

SID	PID	Cost
1	1	$10.00
1	2	$20.00
1	5	$50.00
2	1	$9.00
2	5	$48.00

$\pi_{sid}((\pi_{pid}\sigma_{color='red'}Parts) \bowtie Catalog)$ gives us:

SID
1
2

$\pi_{sid}((\pi_{pid}\sigma_{color='red'}Parts) \bowtie Catalog) \bowtie Suppliers$) gives us:

SID	Sname	Address
1	Yosemite Sham	Devil's canyon, AZ
2	Wiley E. Coyote	RR Asylum, NV

And finally $\pi_{sname}(\pi_{sid}((\pi_{pid}\sigma_{color='red'}Parts) \bowtie Catalog) \bowtie Suppliers)$

Gives us:

Sname
Yosemite Sham
Wiley E. Coyote

2. Find the *sid*s of suppliers who supply some red or green part.

$\pi_{sid}(\pi_{pid}(\sigma_{color='red' \vee color='green'} Parts) \bowtie catalog)$

3. Find the *sid*s of suppliers who supply some red part or are at 221 Packer Street.

$\rho(R1, \pi_{sid}((\pi_{pid}\sigma_{color='red'} Parts) \bowtie Catalog))$
$\rho(R2, \pi_{sid}\sigma_{address='221PackerStreet'} Suppliers)$
$R1 \cup R2$

4. Find the *sid*s of suppliers who supply some red part and some green part.

$\rho(R1, \pi_{sid}((\pi_{pid}\ \sigma_{color='red'} Parts) \bowtie\!| Catalog))$
$\rho(R2, \pi_{sid}((\pi_{pid}\ \sigma_{color='green'} Parts) \bowtie Catalog))$
$R1 \cap R2$

5. Find the *sid*s of suppliers who supply every part.

$(\pi_{sid,pid}\ Catalog)/(\pi_{pid}\ Parts)$

Given:

Parts

PID	Pname	Color
1	Red1	Red
2	Red2	Red
3	Green1	Green
4	Blue1	Blue
5	Red3	Red

Catalog

SID	PID	Cost
1	1	$10.00
1	2	$20.00
1	3	$30.00
1	4	$40.00
1	5	$50.00
2	1	$9.00
2	3	$34.00
2	5	$48.00
3	1	$11.00

π_{pid} *Parts gives us:*

PID
1
2
3
4
5

$\pi_{sid,pid}$ *Catalog gives us:*

SID	PID
1	1
1	2
1	3
1	4
1	5
2	1
2	3
2	5
3	1

($\pi_{sid,pid}$ *Catalog*)/(π_{pid} *Parts*)

Asks the question – what sids in catalog contain all the part numbers in the divisor. There is only one sid that has all the part numbers, 1.

6. Find the *sid*s of suppliers who supply every red part.

($\pi_{sid,pid}$ *Catalog*)/(π_{pid} $\sigma_{color='red'}$ *Parts*)

7. Find the *sid*s of suppliers who supply every red or green part.

$(\pi_{sid,pid}\ Catalog)/(\pi_{pid}\ \sigma_{color='red' \vee color='green'}\ Parts)$

Let's look at this one:

Catalog

SID	PID	Cost
1	1	$10.00
1	2	$20.00
1	3	$30.00
1	4	$40.00
1	5	$50.00
2	1	$9.00
2	3	$34.00
2	5	$48.00
3	1	$11.00

Parts

PID	Pname	Color
1	Red1	Red
2	Red2	Red
3	Green1	Green
4	Blue1	Blue
5	Red3	Red

$\sigma_{color='red' \vee color='green'}\ Parts$ *gives us 1,2,3,5*

$(\pi_{sid,pid}\ Catalog)/(\pi_{pid}\ \sigma_{color='red' \vee color='green'}\ Parts)$ then says :

What supplier supplies all 4 of those parts?
The answer is only supplier 1.

8. Find the *sid*s of suppliers who supply every red part or supply every green part.

$\rho(R1,\ ((\pi_{sid,pid}\ Catalog)/(\pi_{pid}\ \sigma_{color='red'}\ Parts)))$
$\rho(R2,\ ((\pi_{sid,pid}\ Catalog)/(\pi_{pid}\ \sigma_{color='green'}\ Parts)))$
$R1 \cup R2$

OK – lets look at this one then:

Catalog

SID	PID	Cost
1	1	$10.00
1	2	$20.00
1	3	$30.00
1	4	$40.00
1	5	$50.00
2	1	$9.00
2	3	$34.00
2	5	$48.00
3	1	$11.00

Parts

PID	Pname	Color
1	Red1	Red
2	Red2	Red
3	Green1	Green
4	Blue1	Blue
5	Red3	Red

$\pi_{pid}\ \sigma_{color='red'}\ Parts$ *gives us 1,2,5*

$\pi_{sid,pid}\ Catalog)/(\pi_{pid}\ \sigma_{color='red'}\ Parts)$ *gives us 1*

$\pi_{pid}\ \sigma_{color='green'}\ Parts$ *gives us 3*

$\pi_{sid,pid}\ Catalog)/(\pi_{pid}\ \sigma_{color='green'}\ Parts)$ gives us 1,2

The union of the 2 give us 1,2

9. Find pairs of *sid*s such that the supplier with the first *sid* charges more for some part than the supplier with the second *sid*.

$\rho(R1, Catalog)$
$\rho(R2, Catalog)$
$\pi_{R1.sid, R2.sid}(\sigma_{R1.pid=R2.pid \wedge R1.sid != R2.sid \wedge R1.cost > R2.cost}(R1 \times R2))$

Let's look at an example of this one:

Catalog:

SID	PID	Cost
1	1	$10.00
2	1	$9.00
2	3	$34.00
3	1	$11.00

R1 x R2 gives us:

SID	PID	Cost	SID	PID	Cost
1	1	$10.00	1	1	$10.00
1	1	$10.00	2	1	$9.00
1	1	$10.00	2	3	$34.00
1	1	$10.00	3	1	$11.00
2	1	$9.00	1	1	$10.00
2	1	$9.00	2	1	$9.00
2	1	$9.00	2	3	$34.00
2	1	$9.00	3	1	$11.00
2	3	$34.00	1	1	$10.00
2	3	$34.00	2	1	$9.00
2	3	$34.00	2	3	$34.00
2	3	$34.00	3	1	$11.00
3	1	$11.00	1	1	$10.00
3	1	$11.00	2	1	$9.00
3	1	$11.00	2	3	$34.00
3	1	$11.00	3	1	$11.00

At this point, we are selecting for the 3 and clauses. The first ($\sigma_{R1.pid=R2.pid}$) gives us:

SID	PID	Cost	SID	PID	Cost
1	1	$10.00	1	1	$10.00
1	1	$10.00	2	1	$9.00
1	1	$10.00	3	1	$11.00
2	1	$9.00	1	1	$10.00
2	1	$9.00	2	1	$9.00
2	1	$9.00	3	1	$11.00
2	3	$34.00	2	3	$34.00
3	1	$11.00	1	1	$10.00
3	1	$11.00	2	1	$9.00
3	1	$11.00	3	1	$11.00

The second and clause ($\sigma_{R1.pid=R2.pid \wedge R1.sid != R2.sid}$) *gives us:*

SID	PID	Cost	SID	PID	Cost
1	1	$10.00	3	1	$11.00
1	1	$10.00	2	1	$9.00
2	1	$9.00	1	1	$10.00
2	1	$9.00	3	1	$11.00
3	1	$11.00	1	1	$10.00
3	1	$11.00	2	1	$9.00

Adding in the third clause ($\sigma_{R1.pid=R2.pid \wedge R1.sid != R2.sid \wedge R1.cost > R2.cost)}$ gives us:

SID	PID	Cost	SID	PID	Cost
1	1	$10.00	2	1	$9.00
3	1	$11.00	1	1	$10.00
3	1	$11.00	2	1	$9.00

And finally, projecting the pairs gives us:

SID	SID
1	2
3	1
3	2

10. Find the *pid*s of parts supplied by at least two different suppliers.

$\rho(R1, Catalog)$
$\rho(R2, Catalog)$
$\pi_{R1.pid}\, \sigma_{R1.pid=R2.pid \wedge R1.sid != R2.sid}\, (R1 \times R2)$

Using the following:

SID	PID	Cost
1	1	$10.00
2	1	$9.00
2	3	$34.00
3	1	$11.00

R1 x R2 gives us:

SID	PID	Cost	SID	PID	Cost
1	1	$10.00	1	1	$10.00
1	1	$10.00	2	1	$9.00
1	1	$10.00	2	3	$34.00
1	1	$10.00	3	1	$11.00
2	1	$9.00	1	1	$10.00
2	1	$9.00	2	1	$9.00
2	1	$9.00	2	3	$34.00
2	1	$9.00	3	1	$11.00
2	3	$34.00	1	1	$10.00
2	3	$34.00	2	1	$9.00
2	3	$34.00	2	3	$34.00
2	3	$34.00	3	1	$11.00
3	1	$11.00	1	1	$10.00
3	1	$11.00	2	1	$9.00
3	1	$11.00	2	3	$34.00
3	1	$11.00	3	1	$11.00

$\sigma_{R1.pid=R2.pid}$ *gives us:*

SID	PID	Cost	SID	PID	Cost
1	1	$10.00	1	1	$10.00
1	1	$10.00	2	1	$9.00
1	1	$10.00	3	1	$11.00
2	1	$9.00	1	1	$10.00
2	1	$9.00	2	1	$9.00
2	1	$9.00	3	1	$11.00
2	3	$34.00	2	3	$34.00
3	1	$11.00	1	1	$10.00
3	1	$11.00	2	1	$9.00
3	1	$11.00	3	1	$11.00

$\sigma_{R1.pid=R2.pid \wedge R1.sid != R2.sid}$ *gives us:*

SID	PID	Cost	SID	PID	Cost
1	1	$10.00	2	1	$9.00
1	1	$10.00	3	1	$11.00
2	1	$9.00	1	1	$10.00
2	1	$9.00	3	1	$11.00
3	1	$11.00	1	1	$10.00
3	1	$11.00	2	1	$9.00

Projecting on PID gives us a single part number – 1 (eliminating the duplicates)

11. Find the *pid*s of the most expensive parts supplied by suppliers named Yosemite Sham.

$\rho(R1, \pi_{sid}\, \sigma_{sname='YosemiteSham'}\, Suppliers)$
$\rho(R2, R1 \bowtie Catalog)$
$\rho(R3, R2)$
$\rho(R4(1 \rightarrow sid, 2 \rightarrow pid, 3 \rightarrow cost), \sigma_{R3.cost < R2.cost}(R3 \times R2))$
$\pi_{pid}(R2 - \pi_{sid,pid,cost}\, R4)$

Given:

Suppliers

SID	Sname	Address
1	Wiley E. Coyote	Acme Testing Ground, NV
2	Yosemite Sham	Devil's Canyon, AZ
3	Elmer Fudd	Carrot Patch, MN

and Catalog:

SID	PID	Cost
1	1	$10.00
2	1	$9.00
2	2	$21.00
2	3	$34.00
3	1	$11.00

$\rho(R1, \pi_{sid}\, \sigma_{sname='YosemiteSham'}\, Suppliers)$

Gives us the value 2.

$\rho(R2, R1 \bowtie Catalog)$

Gives us:

SID	PID	Cost
2	1	$9.00
2	2	$21.00
2	3	$34.00

Let's look at:
$\rho(R4(1 \rightarrow sid,\ 2 \rightarrow pid,\ 3 \rightarrow cost),\ \sigma_{R3.cost < R2.cost}(R3 \times R2))$

R3 x R2

SID	PID	Cost	SID	PID	Cost
2	1	$9.00	2	1	$9.00
2	1	$9.00	2	2	$21.00
2	1	$9.00	2	3	$34.00
2	2	$21.00	2	1	$9.00
2	2	$21.00	2	2	$21.00
2	2	$21.00	2	3	$34.00
2	3	$34.00	2	1	$9.00
2	3	$34.00	2	2	$21.00
2	3	$34.00	2	3	$34.00

$\sigma_{R3.cost < R2.cost}(R3 \times R2))$ gives us:

SID	PID	Cost	SID	PID	Cost
2	1	$9.00	2	2	$21.00
2	1	$9.00	2	3	$34.00
2	2	$21.00	2	3	$34.00

$\rho(R4(1 \rightarrow sid,\ 2 \rightarrow pid,\ 3 \rightarrow cost),\ \sigma_{R3.cost < R2.cost}(R3 \times R2))$

Gives us:

SID	PID	Cost
2	1	$9.00
2	2	$21.00

$R2 - \pi_{sid,pid,cost} R4$ gives us:

SID	PID	Cost
2	3	$34.00

And projecting the PID gives us 3 as Yosemite Sham's most expensive part.