Washington State University

School of Electrical Engineering and Computer Science

CptS 451 – Introduction to Database Systems

Online

Dr. Sakire Arslan Ay

Homework-3
Relational Alrebra

Name: CHARLES NGUYEN

Student Number: 011606177

Question	Max Points	Score
1 through 10	100	
Total	100	

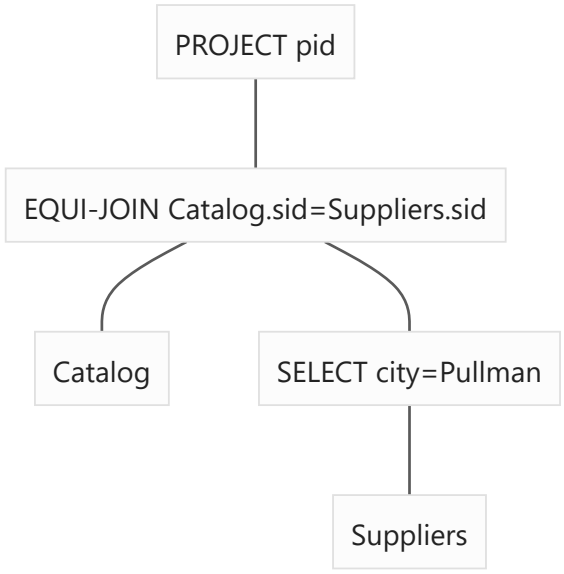
1. (10pts)

- Find all
  - *distinct* **parts**
    - **supplied**
      - by *Pullman* stores
- Return matched pid **parts**.

- Solution:
  - R1: `SELECT * Suppliers WHERE city=Pullman`
  - R2: parts need to be matched with stores, so we choose Catalog
    - `SELECT * Catalog`
  - EQUI-JOIN R1 and R2 on (Catalog.sid = Suppliers.sid)
  - PROJECT on pid

$$\prod_{pid} \sigma_{\text{Catalog.sid=Suppliers.sid}} [Catalog \times \sigma_{city=\text{Pullman}}(Suppliers)]$$

$$\begin{aligned} R_1 &:= \sigma_{city=Tom}(Suppliers) \\ R_2 &:= \sigma_{Catalog.sid=Suppliers.sid}(Catalog \times R_1) \\ RESULT_{pid} &:= \Pi_{pid}(R_2) \end{aligned}$$

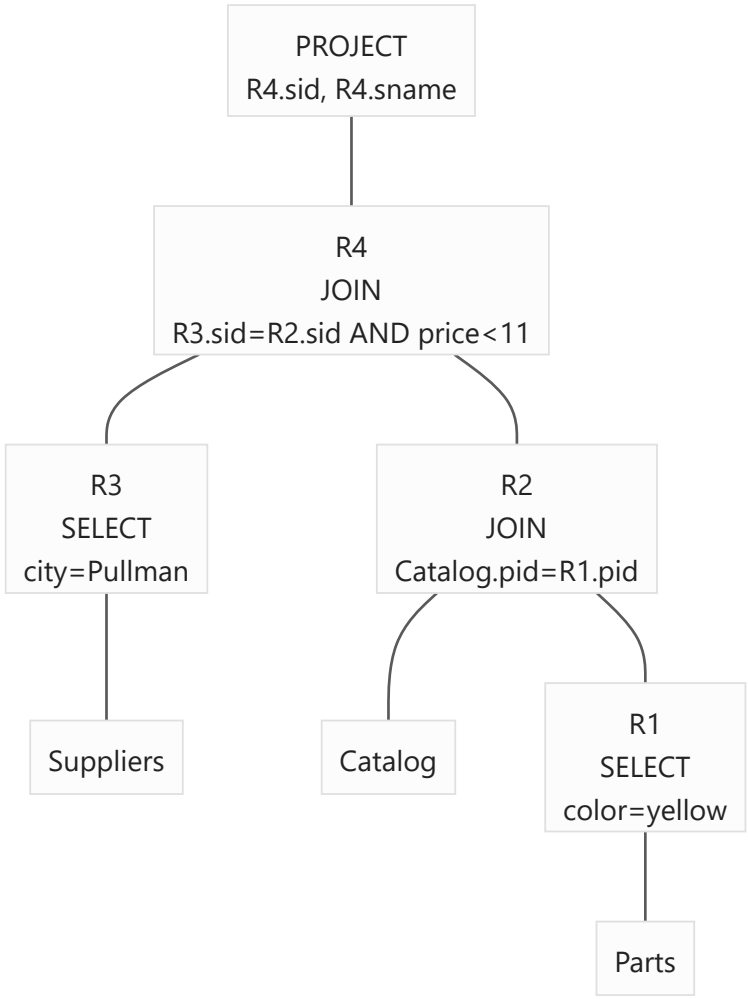


2. (10pts)

- Find
  - the **suppliers**
    - in *Pullman*
  - who supply
    - a *yellow* part
    - for *less than $11*
- Return matched sid and name of **suppliers**.

- Solution:
  - R1: `SELECT Parts WHERE color=yellow`
  - R2: EQUI-JOIN Catalog and R1 on (Catalog.pid=Parts.pid)
  - R3: `SELECT Suppliers WHERE city=Pullman`
  - R4: THETA-JOIN R3 and R2 on (R3.sid=R2.pid AND price < 11)
  - PROJECT on sid, sname *We can also just do Cartesian Product twice and then apply both conditions*:

$$\prod_{\text{(sid, sname)}} [\sigma_{\text{city=Pullman AND color=yellow}}(Suppliers \times Catalog \times Parts)]$$

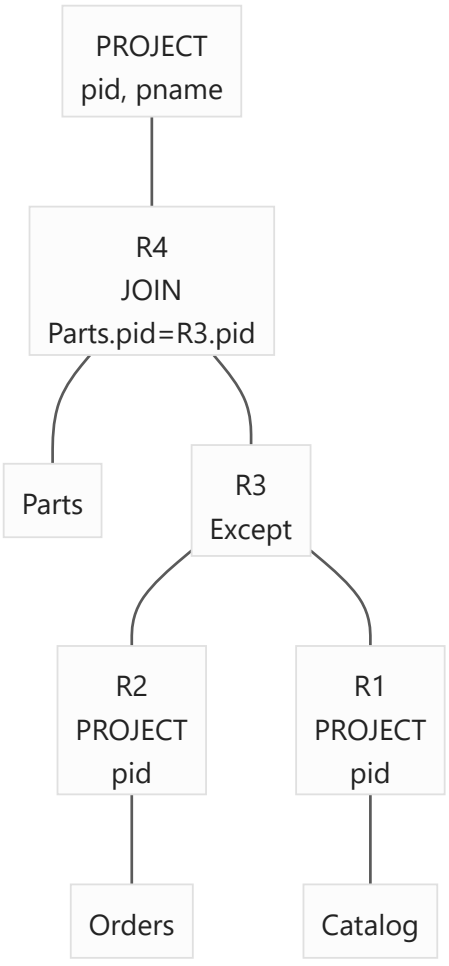

PROJECT
R4.sid, R4.sname
R4
JOIN
R3.sid=R2.sid AND price<11
R3
SELECT
city=Pullman
R2
JOIN
Catalog.pid=R1.pid
Suppliers
Catalog
R1
SELECT
color=yellow
Parts

3. (10pts)

- Find all
  - **parts**,
    - provided by *some suppliers*, i.e. present in **catalog**
      - i.e. *only those found in Catalog*
    - never **ordered** *by a customer*
- Return matched pid and pname of **parts**.

- Solution:
  - R1: `PROJECT pid FROM Catalog`
  - R2: `PROJECT pid FROM Orders`
  - R3: Take set EXCEPT on R1 and R2
  - R4: JOIN Parts and R3
  - PROJECT (pid, pname) from R4

$$\prod_{\text{(pid, pname)}} = Parts \times [\Pi_{pid}(Orders) - \Pi_{pid}(Catalog)]$$

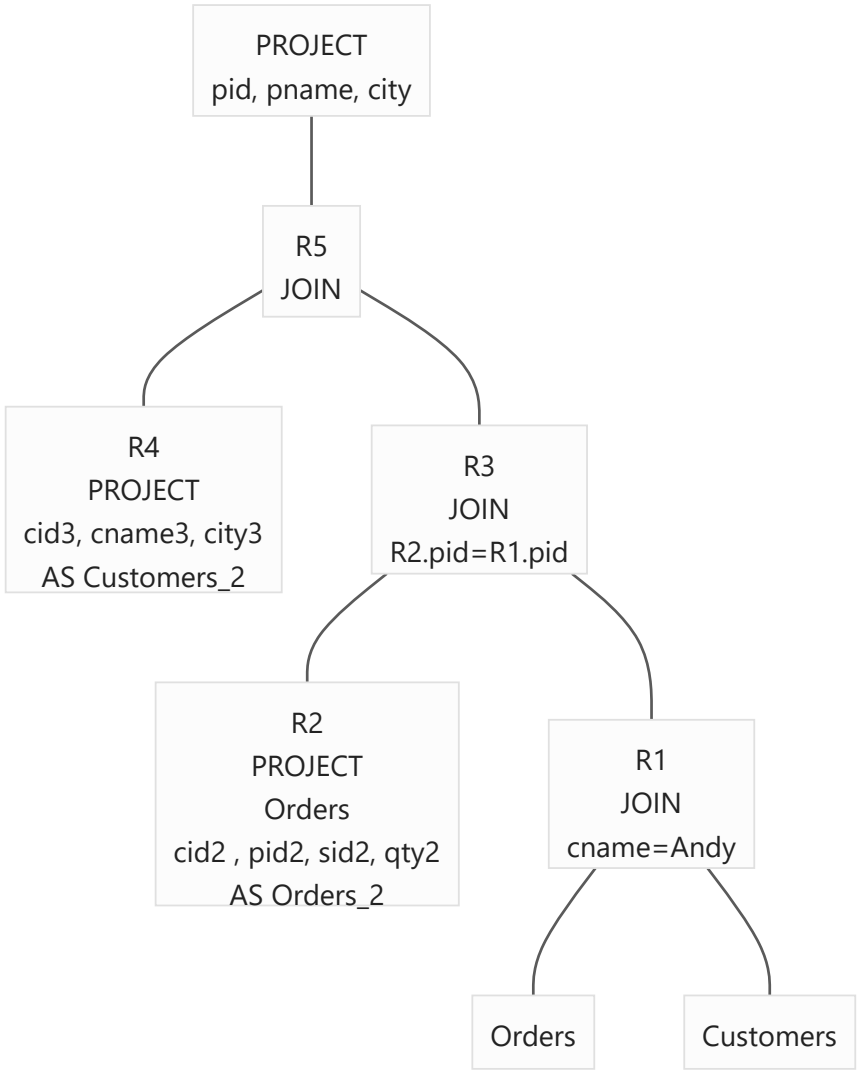


4. (10pts)

- Find all
  - *Customers*, who
    - **ordered** *one of the products that* **Andy ordered**
      - i.e. someone who ordered at least one part as Andy's.
- Return matched (pid, pname, city) of **customers**.

- Solution:
  - R1: JOIN Customers and Order WHERE cname=Andy
  - R2: PROJECT (cid, pid, sid, qty) FROM Orders AS Orders_2
  - R3: JOIN R2 with R1 WHERE R2.pid=R1.pid
  - R4: PROJECT (cid3, cname3, city3) FROM Customers AS Customers_2
  - R5: JOIN R4 with R3
  - PROJECT (pid, pname, city) from R5

$$\prod_{(cname2, city2, pid2)} = (\Pi_{Customers2:=(cid3, cname3, city3)} \times (\Pi_{\text{Order2:=(cid2, pid2, sid2, qty2)}} \times \sigma_{pid2=pid}(Orders \times \sigma_{\text{cname=Andy}}(Customers))))$$

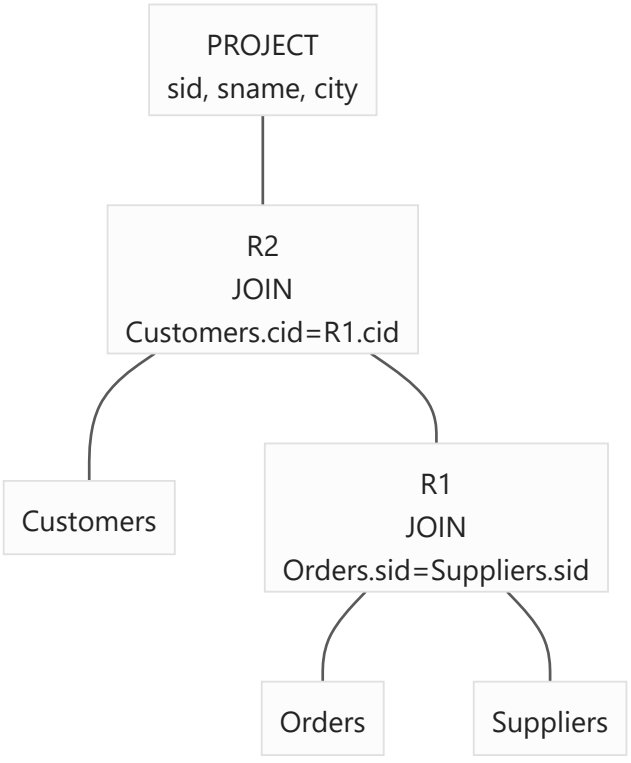

PROJECT
pid, pname, city
R5
JOIN
R4
PROJECT
cid3, cname3, city3
AS Customers_2
R3
JOIN
R2.pid=R1.pid
R2
PROJECT
Orders
cid2 , pid2, sid2, qty2
AS Orders_2
R1
JOIN
cname=Andy
Orders
Customers

5. (10pts)

- Find **suppliers**, who
  - received orders from customers
    - who live in the *same city as the suppliers*
- Return matched (sid, sname, city) of suppliers.

- Solution:
  - R1: JOIN Orders and Suppliers WHERE Orders.sid=Suppliers.sid
  - R2: JOIN Customers and R1 WHERE Customers.cid = R1.cid
  - PROJECT (sid, sname, city) from R2

$$\prod_{(sid,sname,city)} = (\sigma_{Customers.cid=cid}(Customers \times (\sigma_{Orders.sid=Suppliers.sid}(Orders \times Suppliers))))$$

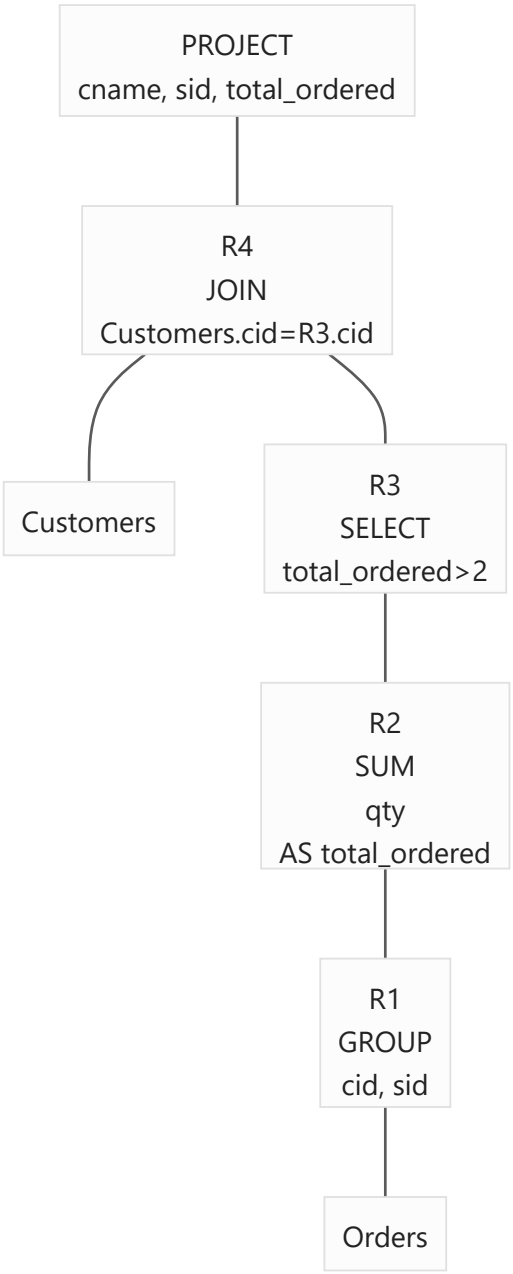


6. (10pts)

- Find **customers**, who
  - ordered qty>2 from a single supplier
- Return matched (cname, sid, total_ordered)

- Solution:
  - R1: GROUP Orders by (cid, sid)
  - R2: Aggregate R1 by SUM the qty attribute in each group in R1 AS total_ordered
  - R3: SELECT from R1 WHERE total > 2
  - R4: JOIN Customers with R3
  - PROJECT (cname, sid, total_ordered)

$$\prod_{(cname, sid, total\_ordered)} = (\sigma_{Customers.cid=cid}(Customers \times (\sigma_{\text{total\_ordered}>2}(\gamma_{(\text{cid, sid, total\_order:=SUM(qty)})}(Orders))))$$

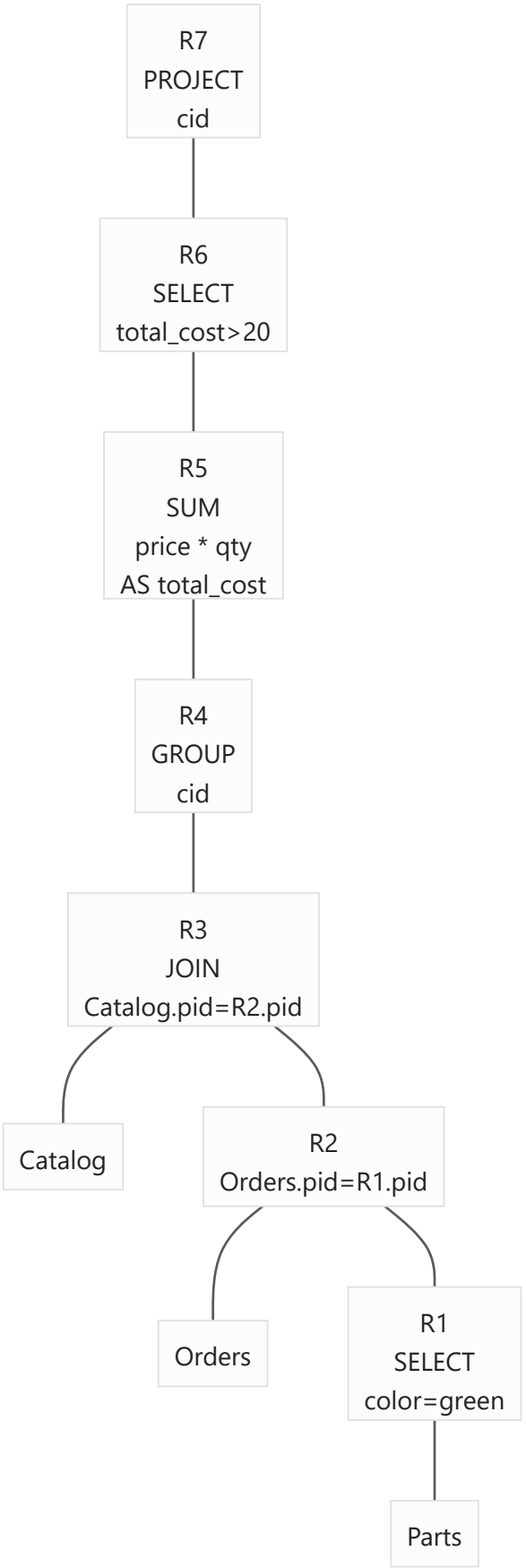

PROJECT
cname, sid, total_ordered
R4
JOIN
Customers.cid=R3.cid
Customers
R3
SELECT
total_ordered>2
R2
SUM
qty
AS total_ordered
R1
GROUP
cid, sid
Orders

7. (10pts)

- Find the customer, who
  - paid more than $20 on 'green' parts
    - total cost of orders on green parts is more than $20
- Return cid from customers.

- Solution:
  - R1: SELECT from Parts WHERE color=green
  - R2: JOIN Orders and R1
  - R3: GROUP R2 by cid
  - R4: Aggregate R3 by SUM by (price * qty) AS total_cost
  - R5: SELECT those with total_cost > 20
  - R6: PROJECT on cid

$$\prod_{cid} = (\sigma_{\text{total\_cost}>20}(\gamma_{(\text{cid, total\_cost:=SUM(price*qty)})}(Catalog \times Orders \times (\sigma_{\text{color=green}}(Parts)))))$$

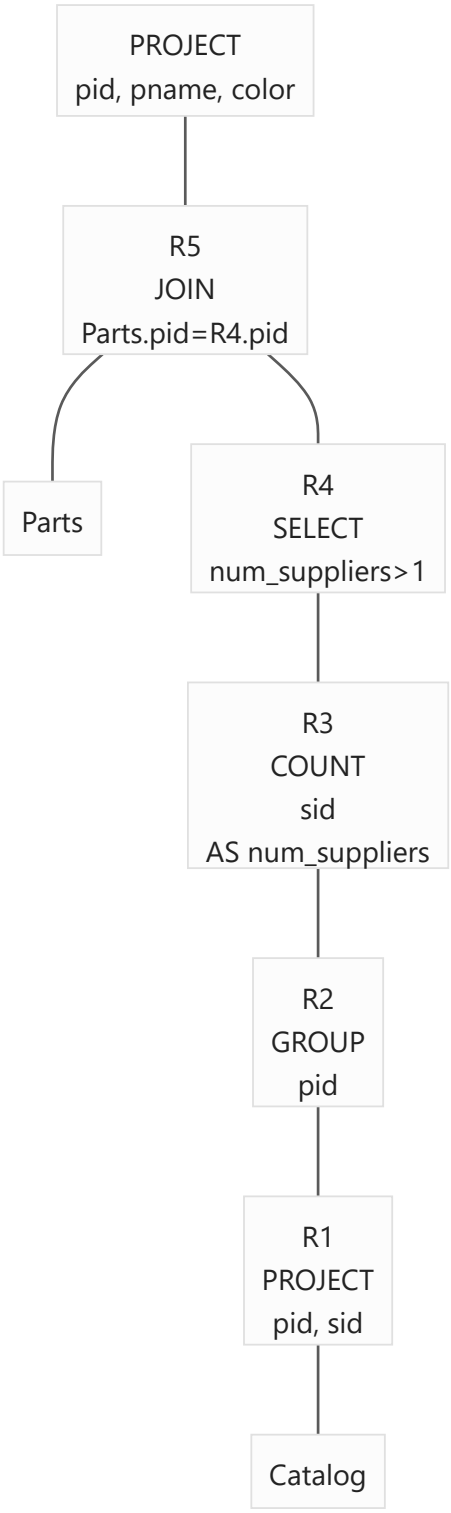

R7
PROJECT
cid
R6
SELECT
total_cost>20
R5
SUM
price * qty
AS total_cost
R4
GROUP
cid
R3
JOIN
Catalog.pid=R2.pid
Catalog
R2
Orders.pid=R1.pid
Orders
R1
SELECT
color=green
Parts

8. (10pts)

- Find the parts, which
  - are supplied by at least 2 different suppliers
    - *market availability, i.e. catalog*
- Return (pid, pname, color) of those parts.

- Solution:
  - R1: PROJECT (pid, sid) from Catalog
  - R2: GROUP R1 by pid
  - R3: Aggregate R2 by COUNT on (sid) for suppliers AS num_suppliers
  - R4: SELECT from R3 WHERE num_suppliers>1
  - R5: JOIN Parts with R4
  - PROJECT (pid, pname, color) from R5

$$\prod_{(pid,pname,color)} = (\sigma_{Parts.pid=pid}(Parts \times \sigma_{num\_suppliers>1}(\gamma_{(\text{pid,num\_suppliers:=COUNT(sid)})}(\Pi_{(pid,sid)}(Catalog)))))$$

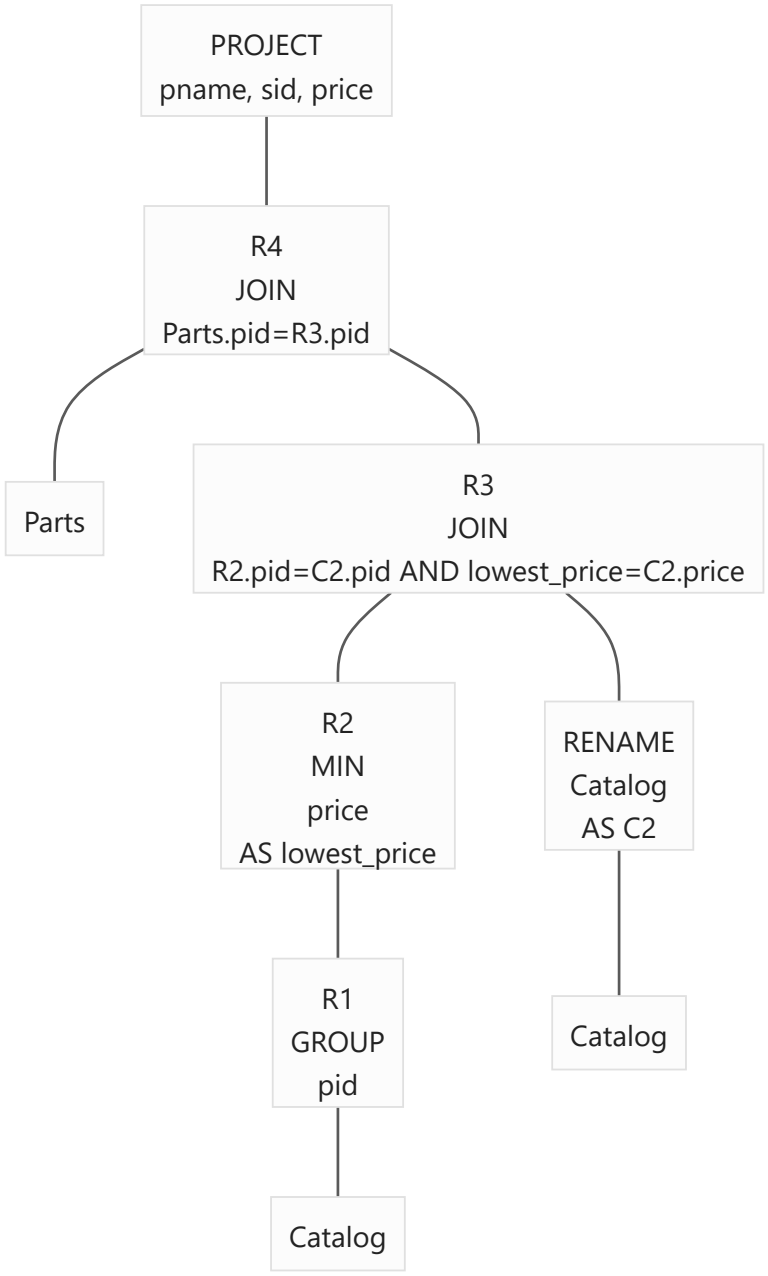

PROJECT
pid, pname, color
R5
JOIN
Parts.pid=R4.pid
Parts
R4
SELECT
num_suppliers>1
R3
COUNT
sid
AS num_suppliers
R2
GROUP
pid
R1
PROJECT
pid, sid
Catalog

9. (10pts)

- Find the supplier of every part in the catalog, who
  - offers the lowest price.
- Return (pname, sid, price) for the parts.

- Solution:
  - R1: GROUP Catalog by (pid)
  - R2: Aggregate R1 by MIN on (price) AS lowest_price
  - RENAME Catalog AS C2
  - R3: JOIN R2 with C2
  - R4: JOIN Parts with R3
  - PROJECT (pname, sid, price) from R4

$$\prod_{(pname,sid,price)} = (\sigma_{Parts.pid=pid}(Parts \times \sigma(\gamma_{(\text{pid, lowest\_price:=MIN(price)})} \times \Pi_{*}(Catalog))))$$


PROJECT
pname, sid, price
R4
JOIN
Parts.pid=R3.pid
Parts
R3
JOIN
R2.pid=C2.pid AND lowest_price=C2.price
R2
MIN
price
AS lowest_price
RENAME
Catalog
AS C2
R1
GROUP
pid
Catalog
Catalog

10. (10pts)

- Find the number of suppliers in each city.

- Solution:
  - R1: GROUP Suppliers by (city)
  - R2: Aggregate R1 by COUNT on (sid) as num_suppliers
  - PROJECT result

$$\prod = \gamma_{\text{city, num\_suppliers:=COUNT(sid)}}(Suppliers)$$