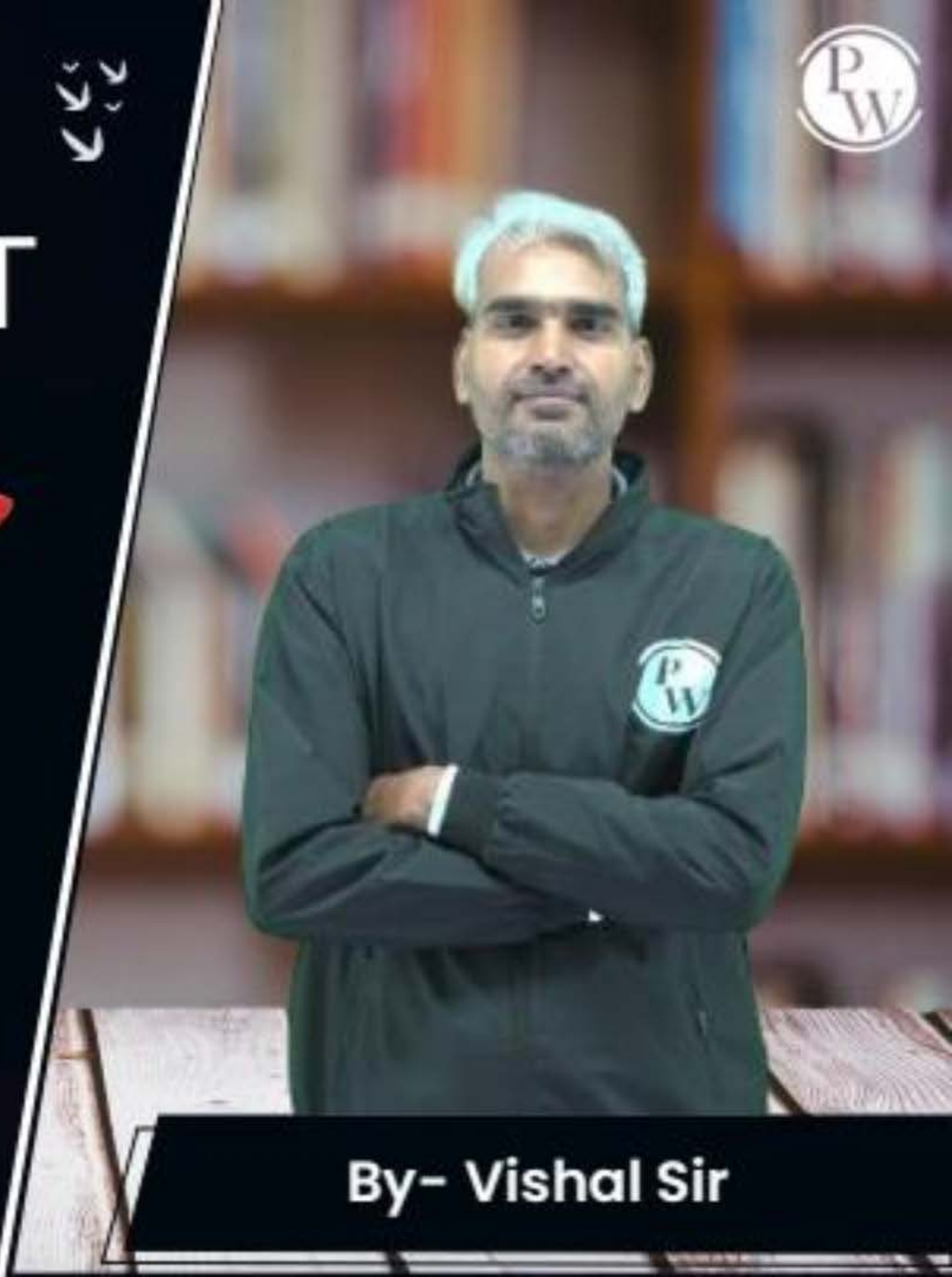
Computer Science & IT

Database Management System

Query Languages

Lecture No. 04



By- Vishal Sir

Recap of Previous Lecture

Topic — Division operation

Topic — Practice questions

Slide

Topics to be Covered

Topic

Practice questions

Slide

Topic : Division (÷)

It is a derived relational algebra opn.

Topic : Division (÷)

Division operation is used whenever the query is with respect to every or all.

Division :-

Student (S)

Sid	Sname
S_1	A
S_2	A
S_3	B
S_4	C

Enroll (E)

Sid	Cid	fee
S_1	C_1	
S_2	C_1	
S_1	C_2	
S_2	C_2	
S_1	C_3	
S_3	C_3	

Course (C)

Cid	Cname
C_1	OS
C_2	DM
C_3	DBMS

Division :-

Query :- Retrieve Sids of all the students

π_{sid} (Student) =

Sid
S_1
S_2
S_3
S_4

Student (S)

Sid	Sname
S_1	A
S_2	A
S_3	B
S_4	C

Enroll (E)

Sid	Cid	fee
S_1	C_1	
S_2	C_1	
S_1	C_2	
S_2	C_2	
S_1	C_3	
S_3	C_3	

Division :-

Enroll (E)

Sid	Cid	fee
S_1	C_1	
S_2	C_1	
S_1	C_2	
S_2	C_2	
S_1	C_3	
S_3	C_3	

Course (C)

Cid	Cname
C_1	OS
C_2	DM
C_3	DBMS

Query :- Retrieve Sids of students who enrolled for some Courses.

At least one (one or more)

π_{sid} (Enroll) =

O/P

Sid
S_1
S_2
S_3

Division :-

Enroll (E)

Sid	Cid	fee
S_1	C_1	
S_2	C_1	
S_1	C_2	
S_2	C_2	
S_1	C_3	
S_3	C_3	

Course (C)

Cid	Cname
C_1	OS
C_2	DM
C_3	DBMS

Query :- Retrieve Course ids of all the Courses

π_{cid} (Course) =

O/P

Cid
C_1
C_2
C_3

Query :-

π_{cid} (Enroll) =

O/P

Cid
C_1
C_2
C_3

It will o/p the Cids of Courses for which at least one student has Enrolled

Division :-

Enroll (E)

Sid	Cid	fee
S_1	C_1	
S_2	C_1	
S_1	C_2	
S_2	C_2	
S_1	C_3	
S_3	C_3	

Course (C)

Cid	Cname
C_1	OS
C_2	DM
C_3	DBMS

Query:- Retrieve Sids of students who have enrolled for all Courses.

Here we are looking for Sids in Enroll table which are associated with all Cids present in Course table.

By observing the given relational table we can say that o/p will be S_1, but we need generalize query.

w.r.t. all we can use division operation

Correct query will be

$$\pi_{Sid,Cid}(Enroll) \div \pi_{Cid}(Course)$$

If we execute this query on the given data in the relational tables then o/p will be S_1

Division :-

Enroll (E)

Sid	Cid	fee
S_1	C_1	
S_2	C_1	
S_1	C_2	
S_2	C_2	
S_1	C_3	
S_3	C_3	

Course (C)

Cid	Cname
C_1	OS
C_2	DM
C_3	DBMS

Cross product

$\pi_{Sid}(Enroll) \times \pi_{Cid}(Course) =$

Sid
S_1
S_2
S_3

Sids of students who enrolled for some Courses.

$\times$

Cid
C_1
C_2
C_3

Cids of all the Courses

$=$

Sid	Cid
S_1	C_1
S_1	C_2
S_1	C_3
S_2	C_1
S_2	C_2
S_2	C_3
S_3	C_1
S_3	C_2
S_3	C_3

It is the universal set

In this relation the students who enroll for at least one Coure have been associated with all Courses

i.e. the situation in which Every student of Enrolled table enrolled for every Course

Division :-

Enroll (E)

Sid	Cid	fee
S_1	C_1	
S_2	C_1	
S_1	C_2	
S_2	C_2	
S_1	C_3	
S_3	C_3	

Course (C)

Cid	Cname
C_1	OS
C_2	DM
C_3	DBMS

π_{sid} [$\pi_{sid}(E) \times \pi_{cid}(C)$ − $\pi_{sid, cid}(E)$]

Sid
S_2
S_3

Sids of Students who did not enroll for all Courses.

Sid	Cid
S_1	C_1
S_1	C_2
S_1	C_3
S_2	C_1
S_2	C_2
S_2	C_3
S_3	C_1
S_3	C_2
S_3	C_3

When Every Student Enrolled for Every Course

Sid	Cid
S_2	C_3
S_3	C_1
S_3	C_2

In this table we will get the information that Which student did not Enroll for what Courses

Sid	Cid
S_1	C_1
S_2	C_1
S_1	C_2
S_2	C_2
S_1	C_3
S_3	C_3

Actual Enrollment information

In the result of this subtraction, we will have the Sids of only those students who did not Enroll for all Courses.

↓

i.e. we will get Sids of students who enrolled for proper subset of Courses, along with the Course ids for which they did not Enroll

Division.

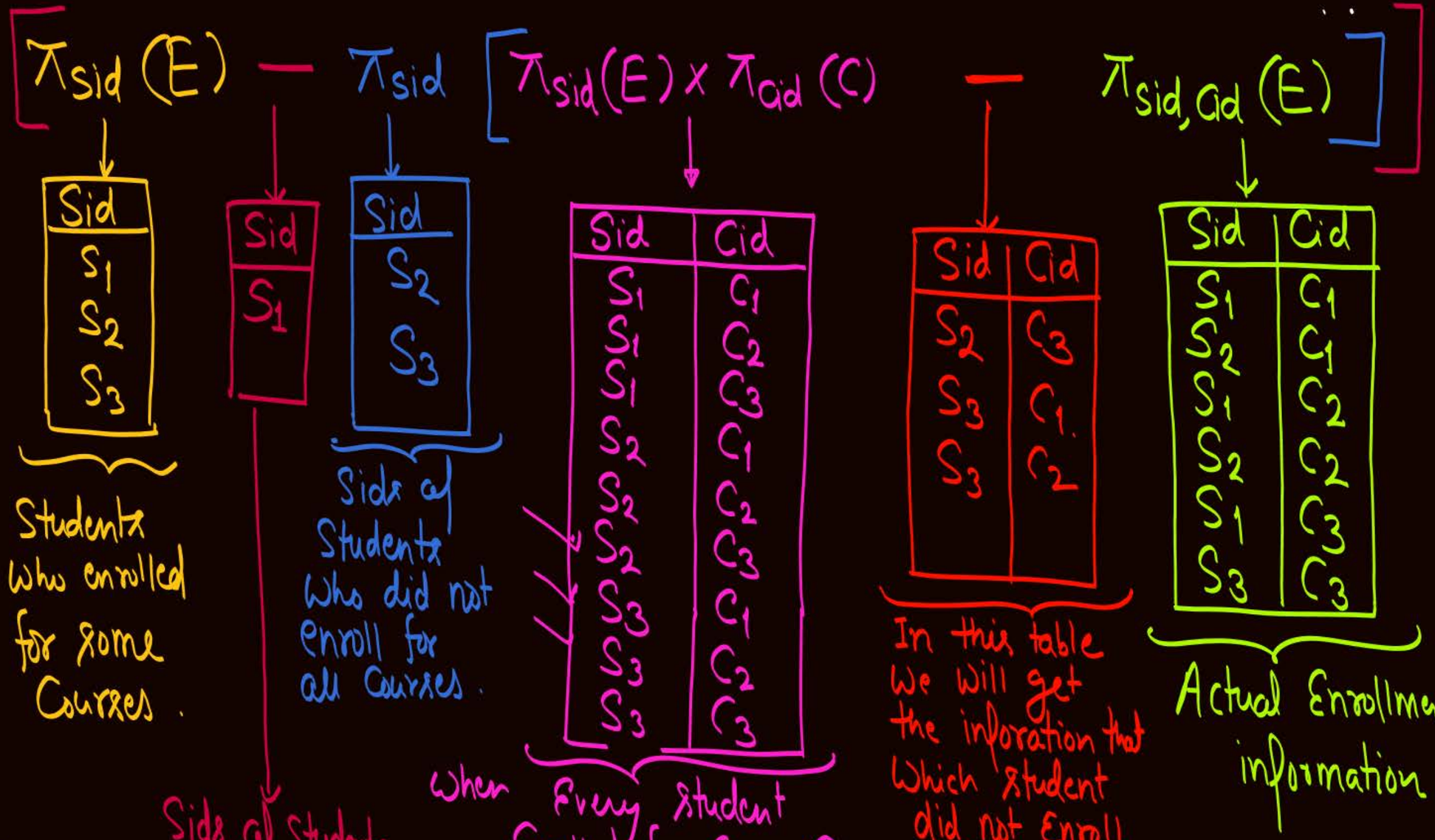
Enroll (E)

Sid	Cid	fee
S_1	C_1	
S_2	C_1	
S_1	C_2	
S_2	C_2	
S_1	C_3	
S_3	C_3	

Course (C)

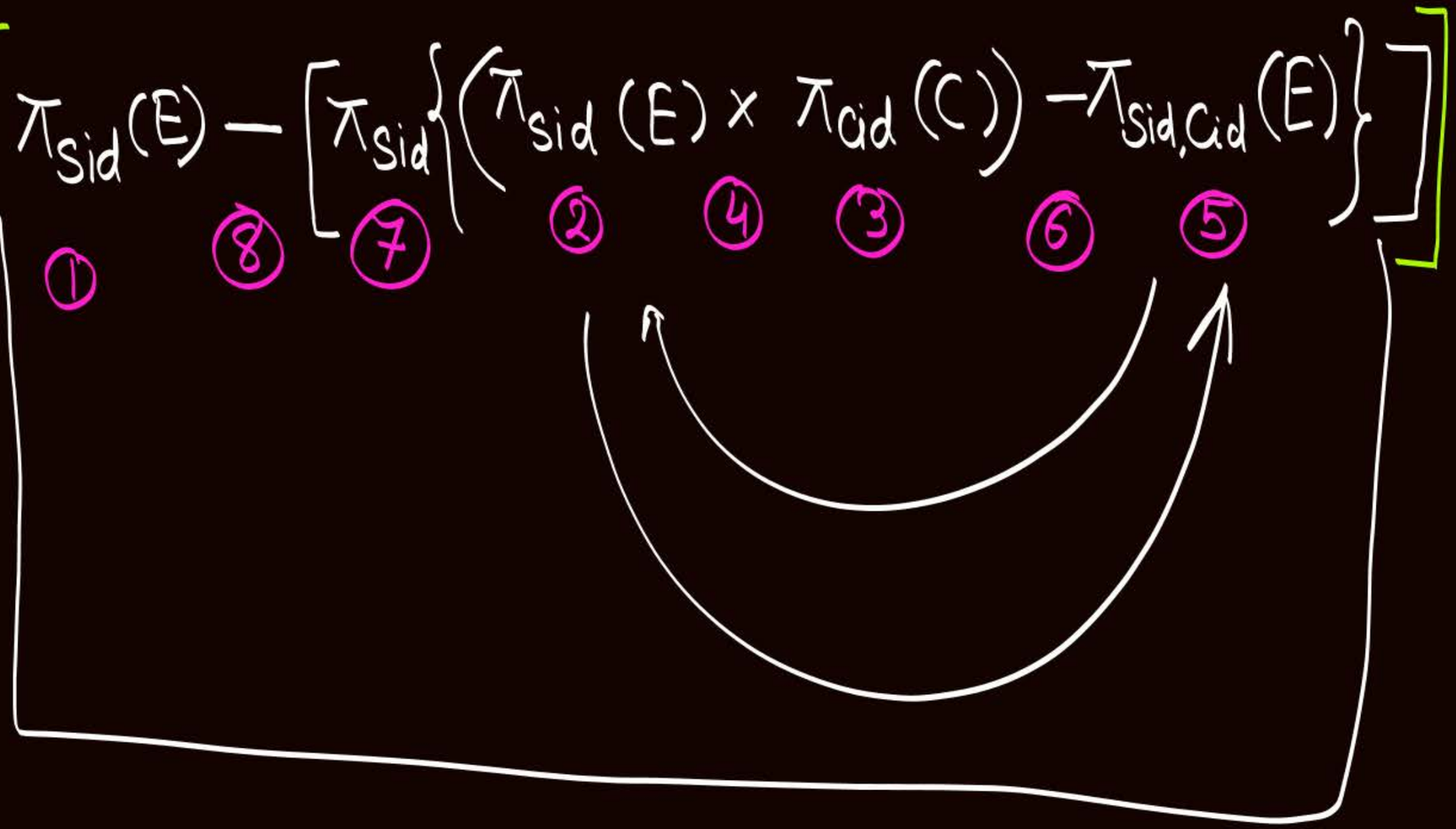
Cid	Cname
C_1	OS
C_2	DM
C_3	DBMS

$$\left[\pi_{Sid}(E) - \pi_{Sid}\left[\pi_{Sid}(E) \times \pi_{Cid}(C) - \pi_{Sid,Cid}(E)\right]\right]$$



$$\{\pi_{sid,cid}(E) \div \pi_{cid}(C)\} = [\pi_{sid}(E) - [\pi_{sid}\{(\pi_{sid}(E) \times \pi_{cid}(C)) - \pi_{sid,cid}(E)\}]]$$

① ⑧ ⑦ ② ④ ③ ⑥ ⑤



Consider the following relational tables:

Supplier (Sid, Sname, Rating)

Contains the information about all suppliers

i.e. Primary Key

Parts (Pid, Pname, Color)

info about all parts

Primary Key

Catalog (Sid, Pid, Cost)

Supplied

it Contains the information that which Supplier has supplied what parts along with Cost.

Together they define Primary Key

Supplier

Sid	Sname	Rating
S_1	A	3
S_2	A	5
S_3	B	7
S_4	C	0

Parts

Pid	Pname	Color
P_1	ABC	Red
P_2	XYZ	Gracen
P_3	KBC	Red

Catalog

Sid	Pid	Cost
S_1	P_1	20
S_1	P_2	30
S_2	P_2	30
S_3	P_2	20
S_3	P_3	10

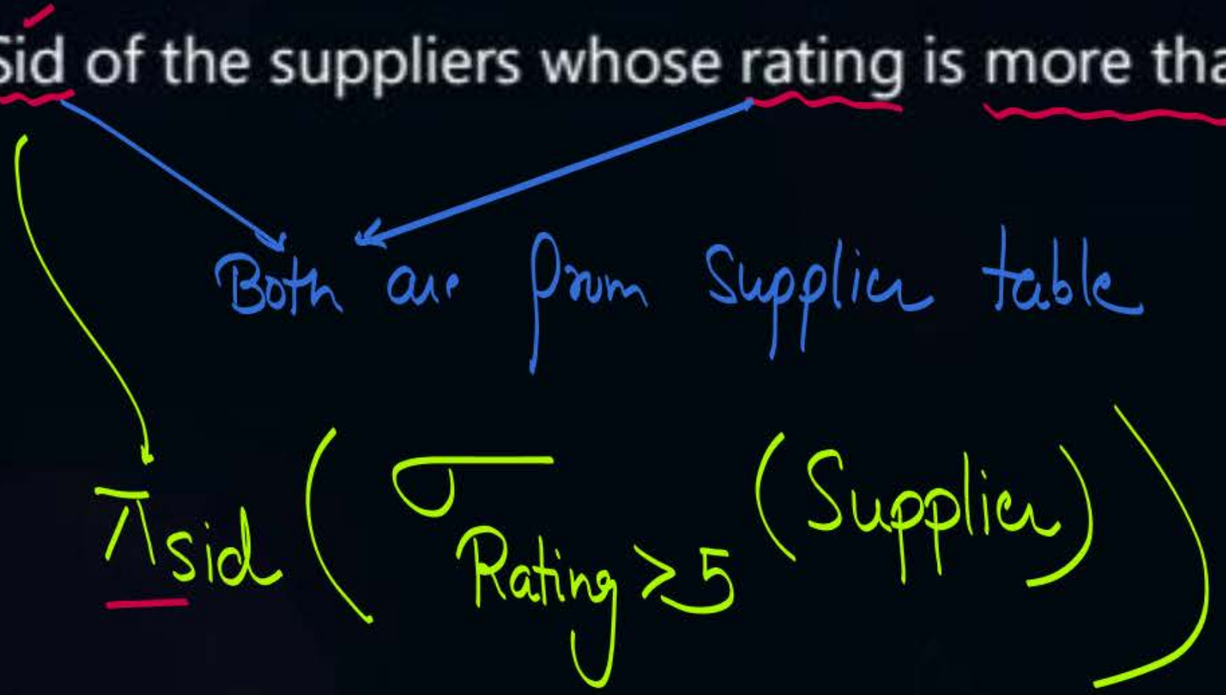
Q: Retrieve Sids of all Supplies

✓ π_{sid} (Supplies)

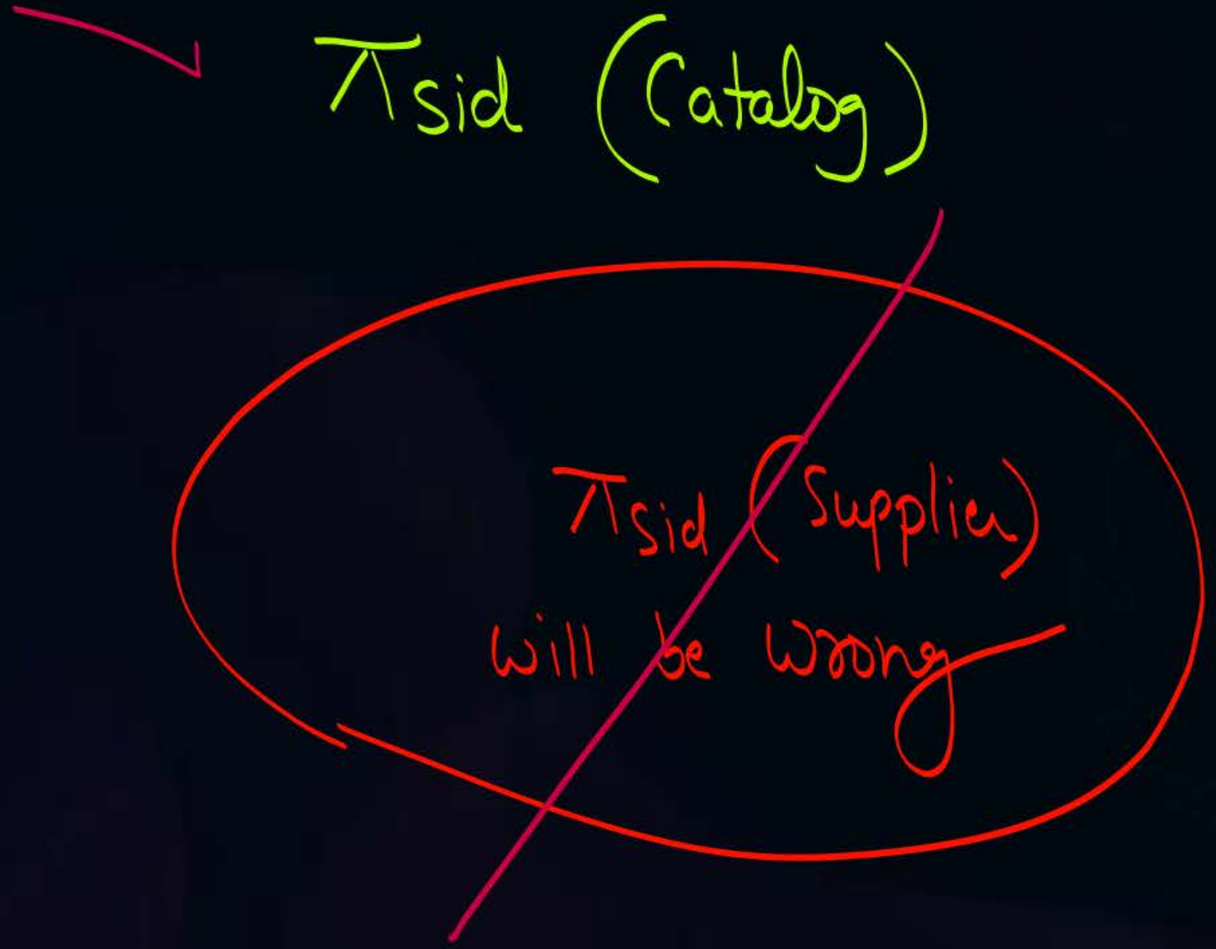
Q: Retrieve Sids of Supplies who supplied some Parts

✓ π_{sid} (Catalog)

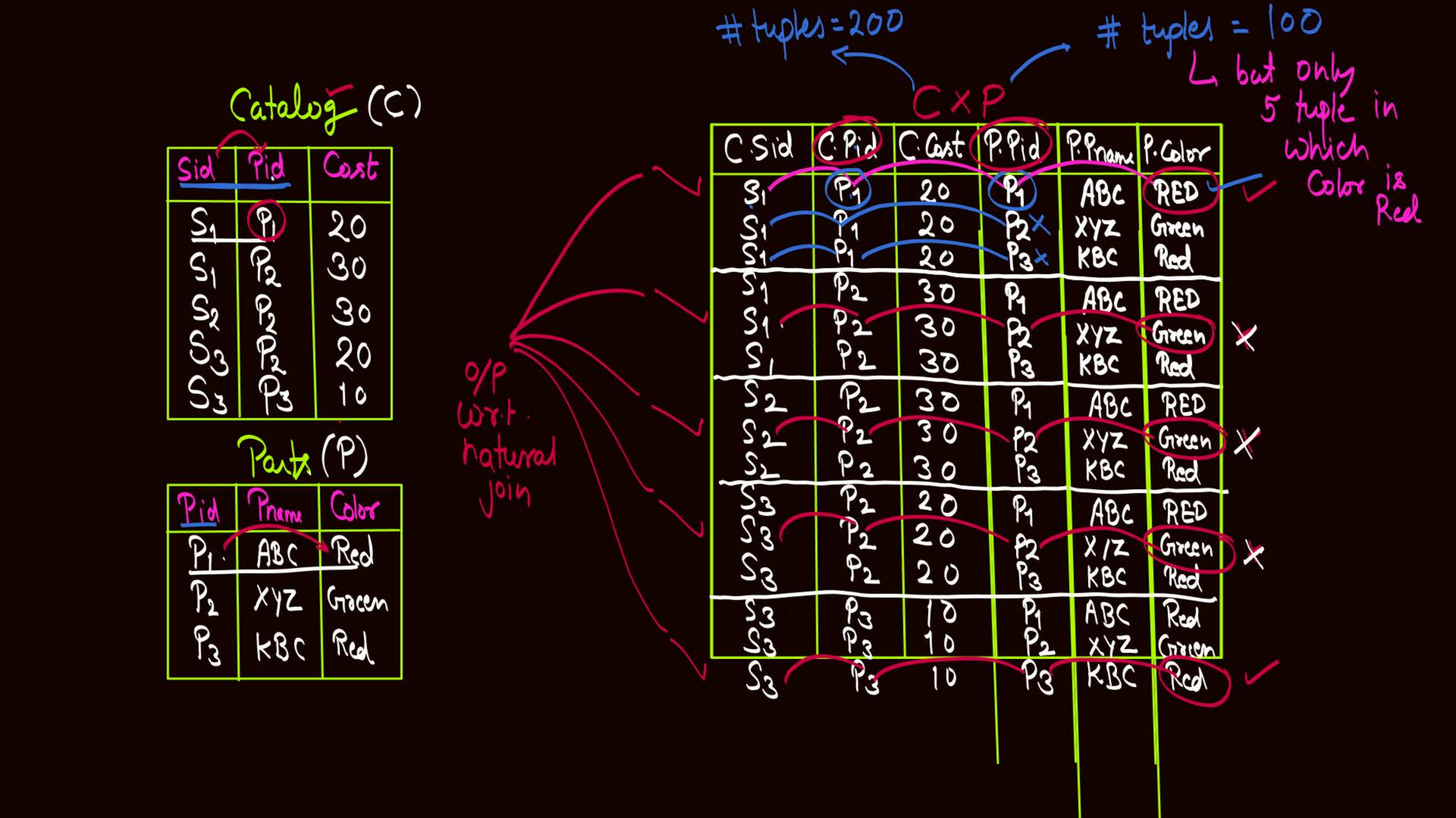
#Q. Retrieve Sid of the suppliers whose rating is more than 5.



#Q. Retrieve Sid of the suppliers who have supplied some parts.



#Q. Retrieve Sid of the suppliers who have supplied some Red color parts.


Catalog (C)
Sid | Pid | Cost
S1 | P1 | 20
S1 | P2 | 30
S2 | P2 | 30
S3 | P2 | 20
S3 | P3 | 10
Parts (P)
Pid | Pname | Color
P1 | ABC | Red
P2 | XYZ | Green
P3 | KBC | Red
o/p w.r.t. natural join
tuples = 200
tuples = 100
but only 5 tuple in which Color is Red
C × P
C.Sid | C.Pid | C.Cost | P.Pid | P.Pname | P.Color
S1 | P1 | 20 | P1 | ABC | RED
S1 | P1 | 20 | P2 | XYZ | Green
S1 | P1 | 20 | P3 | KBC | Red
S1 | P2 | 30 | P1 | ABC | RED
S1 | P2 | 30 | P2 | XYZ | Green
S1 | P2 | 30 | P3 | KBC | Red
S2 | P2 | 30 | P1 | ABC | RED
S2 | P2 | 30 | P2 | XYZ | Green
S2 | P2 | 30 | P3 | KBC | Red
S3 | P2 | 20 | P1 | ABC | RED
S3 | P2 | 20 | P2 | XYZ | Green
S3 | P2 | 20 | P3 | KBC | Red
S3 | P3 | 10 | P1 | ABC | Red
S3 | P3 | 10 | P2 | XYZ | Green
S3 | P3 | 10 | P3 | KBC | Red

(i) $\pi_{C.Sid}\left(\sigma_{C.Pid = P.Pid \wedge P.Color = 'Red'}\ (C \times P)\right)$

20,000 tuples

(ii) $\pi_{Sid}\left[\sigma_{Color = 'Red'}\ (C \bowtie P)\right]$

on 20,000 tuples Equality Condⁿ on Common attribute

(iii) $\pi_{Sid}\left\{C \bowtie \left(\sigma_{Color = 'Red'}\ (P)\right)\right\}$

200 tuples 100 tuples

5 tuples

1000 tuples

Most efficient query among three!

#Q. Retrieve Sid of the suppliers who have supplied some Red or some Green color parts.

Catalog (C)

Sid	Pid	Cost
S_1	P_1	20
S_1	P_2	30
S_2	P_2	30
S_3	P_2	20
S_3	P_3	10

Parts (P)

Pid	Pname	Color
P_1	ABC	Red
P_2	XYZ	Green
P_3	KBC	Red

C × P

C.Sid	C.Pid	C.Cost	P.Pid	P.Pname	P.Color
S_1	P_1	20	P_1	ABC	RED
S_1	P_1	20	P_2	XYZ	Green
S_1	P_1	20	P_3	KBC	Red
S_1	P_2	30	P_1	ABC	RED
S_1	P_2	30	P_2	XYZ	Green
S_1	P_2	30	P_3	KBC	Red
S_2	P_2	30	P_1	ABC	RED
S_2	P_2	30	P_2	XYZ	Green
S_2	P_2	30	P_3	KBC	Red
S_3	P_2	20	P_1	ABC	RED
S_3	P_2	20	P_2	XYZ	Green
S_3	P_2	20	P_3	KBC	Red
S_3	P_3	10	P_1	ABC	Red
S_3	P_3	10	P_2	XYZ	Green
S_3	P_3	10	P_3	KBC	Red

(i) $\pi_{C.Sid}\{\sigma_{C.Pid = P.Pid \wedge (P.Color = Red \vee P.Color = Green)}(C \times P)\}$ ✗

(ii) $\pi_{Sid}\left(\sigma_{Color = Red \vee Color = Green}(C \bowtie P)\right)$

(iii) $\pi_{Sid}\{C \bowtie (\sigma_{Color = Red \vee Color = Green}(P))\}$

Q. Retrieve Sids of the suppliers who have supplied some Red or Green Color Parts.

(i) $\pi_{C_1.Sid}\left(\sigma_{C_1.Pid = P_1.Pid \wedge P_1.Color = 'Red'}\left(C_1 \times P_1\right)\right) \cup \pi_{C_2.Sid}\left(\sigma_{C_2.Pid = P_2.Pid \wedge P_2.Color = Green}\left(C_2 \times P_2\right)\right)$

(ii) $\pi_{Sid}\left[\sigma_{Color='Red'}\left(C \bowtie P\right)\right] \cup \pi_{Sid}\left[\sigma_{Color=Green}\left(C \bowtie P\right)\right]$

(iii) $\pi_{Sid}\left\{\underbrace{C}_{200\ tuples} \bowtie \left(\sigma_{Color='Red'}(P)\right)\right\} \cup \pi_{Sid}\left\{\underbrace{C}_{200\ tuples} \bowtie \left(\sigma_{Color=Green}(P)\right)\right\}$

o/p = S_1, S_3

o/p = S_1, S_2, S_3

o/p = S_1, S_2, S_3

#Q. Retrieve Sid of the suppliers who have supplied some Red and some Green color parts.

Catalog (C)

Sid	Pid	Cost
S_1	P_1	20
S_1	P_2	30
S_2	P_2	30
S_3	P_2	20
S_3	P_3	10

Parts (P)

Pid	Pname	Color
P_1	ABC	Red
P_2	XYZ	Green
P_3	KBC	Red

C × P

C.Sid	C.Pid	C.Cost	P.Pid	P.Pname	P.Color
S_1	P_1	20	P_1	ABC	RED
S_1	P_1	20	P_2	XYZ	Green
S_1	P_1	20	P_3	KBC	Red
S_1	P_2	30	P_1	ABC	RED
S_1	P_2	30	P_2	XYZ	Green
S_1	P_2	30	P_3	KBC	Red
S_2	P_2	30	P_1	ABC	RED
S_2	P_2	30	P_2	XYZ	Green
S_2	P_2	30	P_3	KBC	Red
S_3	P_2	20	P_1	ABC	RED
S_3	P_2	20	P_2	XYZ	Green
S_3	P_2	20	P_3	KBC	Red
S_3	P_3	10	P_1	ABC	Red
S_3	P_3	10	P_2	XYZ	Green
S_3	P_3	10	P_3	KBC	Red

(i) $\pi_{C.Sid}\{\sigma_{C.Pid = P.Pid \wedge (P.Color = Red \wedge P.Color = Green)}(C \times P)\}$

(ii) $\pi_{Sid}\left(\sigma_{Color = Red \wedge Color = Green}(C \bowtie P)\right)$

(iii) $\pi_{Sid}\{C \bowtie \left(\sigma_{Color = Red \wedge Color = Green}(P)\right)\}$

↳ Color can not be Red as well as Green in the same tuple.

∴ O/p will always be empty w.r.t. above queries.

(i) $\pi_{C_1.Sid}\left(\sigma_{C_1.Pid = P_1.Pid \wedge P_1.Color = 'Red'}\ (C_1 \times P_1)\right) \cap \pi_{C_2.Sid}\left(\sigma_{C_2.Pid = P_2.Pid \wedge P_2.Color = Green}\ (C_2 \times P_2)\right)$

(ii) $\pi_{Sid}\left[\sigma_{Color='Red'}\ (C \bowtie P)\right] \cap \pi_{Sid}\left[\sigma_{Color=Green}\ (C \bowtie P)\right]$

(iii) $\pi_{Sid}\left\{\underbrace{C}_{200\ tuples} \bowtie \left(\sigma_{Color='Red'}\ (P)\right)\right\} \cap \pi_{Sid}\left\{\underbrace{C}_{200\ tuples} \bowtie \left(\sigma_{Color=Green}\ (P)\right)\right\}$

O/P :

S_1
S_3

$\cap$

O/P :

S_1
S_3

O/P :

S_1
S_2
S_3

Catalog (C)

Sid	Pid	Cost
S_1	P_1	20
S_1	P_2	30
S_2	P_2	30
S_3	P_2	20
S_3	P_3	10

Parts (P)

Pid	Pname	Color
P_1	ABC	Red
P_2	XYZ	Green
P_3	KBC	Red

$C_1 \times P_1$

C.Sid	C.Pid	C.Cost	P.Pid	P.Pname	P.Color
S_1	P_1	20	P_1	ABC	RED
S_1	P_1	20	P_2	XYZ	Green
S_1	P_1	20	P_3	KBC	Red
S_1	P_2	30	P_1	ABC	RED
S_1	P_2	30	P_2	XYZ	Green
S_1	P_2	30	P_3	KBC	Red
S_2	P_2	30	P_1	ABC	RED
S_2	P_2	30	P_2	XYZ	Green
S_2	P_2	30	P_3	KBC	Red
S_3	P_2	20	P_1	ABC	RED
S_3	P_2	20	P_2	XYZ	Green
S_3	P_2	20	P_3	KBC	Red
S_3	P_3	10	P_1	ABC	Red
S_3	P_3	10	P_2	XYZ	Green
S_3	P_3	10	P_3	KBC	Red

$C_2 \times P_2$

C.Sid	C.Pid	C.Cost	P.Pid	P.Pname	P.Color
S_1	P_1	20	P_1	ABC	RED
S_1	P_1	20	P_2	XYZ	Green
S_1	P_1	20	P_3	KBC	Red
S_1	P_2	30	P_1	ABC	RED
S_1	P_2	30	P_2	XYZ	Green
S_1	P_2	30	P_3	KBC	Red
S_2	P_2	30	P_1	ABC	RED
S_2	P_2	30	P_2	XYZ	Green
S_2	P_2	30	P_3	KBC	Red
S_3	P_2	20	P_1	ABC	RED
S_3	P_2	20	P_2	XYZ	Green
S_3	P_2	20	P_3	KBC	Red
S_3	P_3	10	P_1	ABC	Red
S_3	P_3	10	P_2	XYZ	Green
S_3	P_3	10	P_3	KBC	Red

$$\pi_{C_1.Sid}\left[\sigma_{(C_1.Sid = C_2.Sid)\ \wedge\ \left((C_1.Pid = P_1.Pid \wedge P_1.Color = Red) \wedge (C_2.Pid = P_2.Pid \wedge P_2.Color = Green)\right)}\left\{(C_1 \times P_1) \times (C_2 \times P_2)\right\}\right]$$

We can also use $C_2.Sid$

w.r.t OR

Q: Retrieve Sids of the suppliers who supplied some Red **OR** Some green Color Parts

$$\pi_{C_1.Sid}\left[\sigma_{(C_1.Sid = C_2.Sid)\ \wedge\ \left((C_1.Pid = P_1.Pid \wedge P_1.Color = Red) \vee (C_2.Pid = P_2.Pid \wedge P_2.Color = Green)\right)} \left\{(C_1 \times P_1) \times (C_2 \times P_2)\right\}\right]$$

We can also use $C_2.Sid$

#Q. Retrieve Sid of the suppliers who have supplied all parts.

∴ division

Catalog (Sid, Pid, Cost)

Parts (Pid, Pname, Color)

We are looking for Sids in the Catalog table which are associated with all Pids in Parts table

$$\therefore \pi_{sid,Pid}(C) \div \pi_{Pid}(P)$$

$$= \pi_{sid}(C) - \pi_{sid}\left[\left(\pi_{sid}(C) \times \pi_{Pid}(P)\right) - \pi_{sid,Pid}(C)\right]$$

#Q. Retrieve Sid of the suppliers who have supplied all Red color parts.

$$\pi_{Sid, Pid}(Catalog) \div \pi_{Pid}(\sigma_{Color = Red}(P))$$

$$= \pi_{Sid}(C) - \pi_{Sid}\left[\left(\pi_{Sid}(C) \times \pi_{Pid}(\sigma_{Color = Red}(P))\right) - \pi_{Sid, Pid}(C)\right]$$

#Q. Retrieve Sid of the suppliers who have supplied at least two parts.

✓ Catalog (C1) X ✓ Catalog (C2)

Sid	Pid	Cost
S_1	P_1	20
S_1	P_2	30
S_2	P_2	30
S_3	P_2	20
S_3	P_3	10

Sid	Pid	Cost
S_1	P_1	20
S_1	P_2	30
S_2	P_2	30
S_3	P_2	20
S_3	P_3	10

We will have to compare Sid in each tuple of Catalog table with Sids with other tuples of Catalog table

If there is any Sid which is present in more than one tuple then that Supplier has Supplied at least two parts

$\rho_{C_1}(\text{catalog}) \times \rho_{C_2}(\text{catalog})$

$$\pi_{C_1.Sid}\{\sigma_{(C_1.Sid = C_2.Sid \wedge C_1.Pid <> C_2.Pid)}(C_1 \times C_2)\}$$

<> : Not Equal (!=)

If we do not use this condⁿ then all Sids of the Catalog table will be present in output, (Each tuple will Combine with itself)

#Q. Retrieve Sid of the suppliers who have supplied exactly one part.

Suppliers who have supplied Exactly one Part = Suppliers who have supplied At least one Part − Suppliers who have supplied At least two Parts

{ i.e. 1, 2, 3, or more Parts } − { i.e. 2, 3, 4, ... or more Parts } = Exactly 1

$= \pi_{sid}(Catalog) - \pi_{sid}\left(\sigma_{(C_1.Sid = C_2.Sid \wedge C_1.Pid <> C_2.Pid)}(C_1 \times C_2)\right)$

#Q. Retrieve Sid of the suppliers who have supplied at most one part.

{i.e. 0 or 1 part}

= Suppliers who have supplied 0 or more parts — Suppliers who have supplied two or more parts

{0, 1, 2, ...∞ more} − {2, 3, ...∞ more} = {0 or 1}

$$= \pi_{Sid}(Supplier) - \pi_{Sid}\left(\sigma_{C_1.Sid = C_2.Sid}(C_1 \times C_2)\right)$$

$C_1.Pid <> C_2.Pid$

If supplier did not supply any part then, Sid of that supplier is not present in catalog

#Q. Retrieve Sid of the suppliers who have supplied at least three parts.

$$\pi_{C_1.Sid}\Big[\sigma\big((C_1.Sid = C_2.Sid) \wedge (C_2.Sid = C_3.Sid) \wedge (C_1.Sid = C_3.Sid)\big)(C_1 \times C_2 \times C_3) \wedge \big((C_1.Pid \neq C_2.Pid) \wedge (C_2.Pid \neq C_3.Pid) \wedge (C_1.Pid \neq C_3.Pid)\big)\Big]$$

it is optional

it is mandatory

#Q. Retrieve Sid of the suppliers who have supplied most expensive parts.

$$\pi_{sid}\left[\left(\begin{array}{c}\text{All tuples}\\ \text{of}\\ \text{Catalog table}\end{array}\right) - \left(\begin{array}{c}\text{Tuples of the catalog table}\\ \text{in which Cost is not maximum}\end{array}\right)\right]$$

Sid of the suppliers who have supplied most expensive parts.

It will give all the tuples from the catalog table, in which Cost is maximum

2 mins Summary

Topic Practice questions

Slide

THANK - YOU