

Universita' di Milano Bicocca
Corso di Basi di dati 1 in eLearning
C. Batini
7. SQL DML
7.4 Select - 4

Gestione duplicati in SQL: la clausola DISTINCT

- In SQL, al contrario dell'algebra relazionale, la tabella finale mantiene eventuali tuple distinte duplicate.
- Quindi, per eliminarle, occorre utilizzare la clausola DISTINCT

Consideriamo lo schema di relazione

Impiegati (Matricola, Cognome, Filiale, Stipendio)

Vogliamo trovare cognome e filiale di tutti gli impiegati

In algebra relazionale abbiamo

$\text{PROJ}_{\text{Cognome, Filiale}}(\text{Impiegati})$

Cognome e filiale di tutti gli impiegati

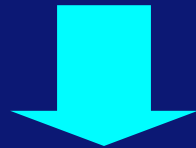
Esecuzione in algebra relazionale

Cognome	Filiale
Neri	Napoli
Neri	Milano
Rossi	Roma

PROJ Cognome, Filiale (Impiegati)

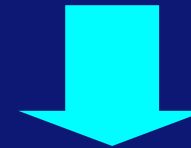
Semantica in SQL della clausola DISTINCT

Select
cognome, filiale
from impiegati



Cognome	Filiale
Neri	Napoli
Neri	Milano
Rossi	Roma
Rossi	Roma

Select distinct
cognome, filiale
from impiegati



Cognome	Filiale
Neri	Napoli
Neri	Milano
Rossi	Roma

Selezione, proiezione e join insieme

Torniamo all'esempio principale

Maternità

Madre	<u>Figlio</u>
Luisa	Maria
Luisa	Luigi
Anna	Olga
Anna	Filippo
Maria	Andrea
Maria	Aldo

Paternità

Padre	<u>Figlio</u>
Sergio	Franco
Luigi	Olga
Luigi	Filippo
Franco	Andrea
Franco	Aldo

Persone

<u>Nome</u>	Età	Reddito
Andrea	27	21
Aldo	25	15
Maria	55	42
Anna	50	35
Filippo	26	30
Luigi	50	40
Franco	60	20
Olga	30	41
Sergio	85	35
Luisa	75	87

Selezione, proiezione e join nell'algebra relazionale

I padri di persone che
guadagnano più di
venti mila €

Quali relazioni
Vanno utilizzate?

PROJ_{Padre}(Paternità)
JOIN_{Figlio = Nome}
SEL_{Reddito > 20}(Persone)

N.B. Qui abbiamo
fatto prima la SEL
E poi il Join

Maternità	Madre	Figlio
	Luisa	Maria
	Luisa	Luigi
	Anna	Olga
	Anna	Filippo
	Maria	Andrea
	Maria	Aldo

Paternità	Padre	Figlio
	Sergio	Franco
	Luigi	Olga
	Luigi	Filippo
	Franco	Andrea
	Franco	Aldo

Persone	Nome	Età	Reddito
	Andrea	27	21
	Aldo	25	15
	Maria	55	42
	Anna	50	35
	Filippo	26	30
	Luigi	50	40
	Franco	60	20
	Olga	30	41
	Sergio	85	35
	Luisa	75	87

Selezione, proiezione e join in SQL

I padri di persone che guadagnano più di venti
mila €

```
Select distinct padre  
from persone, paternita  
where Figlio = Nome and Reddito > 20
```

Risultato

Luigi
Franco

N.B. qui non
abbiamo necessita'
di ridenominazioni

Maternità	Madre	Figlio
	Luisa	Maria
	Luisa	Luigi
	Anna	Olga
	Anna	Filippo
	Maria	Andrea
	Maria	Aldo

Paternità	Padre	Figlio
	Sergio	Franco
	Luigi	Olga
	Luigi	Filippo
	Franco	Andrea
	Franco	Aldo

Persone		
Nome	Età	Reddito
Andrea	27	21
Aldo	25	15
Maria	55	42
Anna	50	35
Filippo	26	30
Luigi	50	40
Franco	60	20
Olga	30	41
Sergio	85	35
Luisa	75	87

Join naturale

- Padre e madre di ogni persona

Algebra: Paternità JOIN Maternità

SQL

```
select Paternità.figlio, padre, madre  
from Maternità, Paternità
```

```
where Paternità.figlio = Maternità.figlio
```

N.B. Trova
Prima le relazioni

Maternità	Madre	Figlio	Persone		
			Nome	Età	Reddito
	Luisa	Maria	Andrea	27	21
	Luisa	Luigi	Aldo	25	15
	Anna	Olga	Maria	55	42
	Anna	Filippo	Anna	50	35
	Maria	Andrea	Filippo	26	30
	Maria	Aldo	Luigi	50	40
			Franco	60	20
			Olga	30	41
			Sergio	85	35
			Luisa	75	87

Paternità	Padre	Figlio
	Sergio	Franco
	Luigi	Olga
	Luigi	Filippo
	Franco	Andrea
	Franco	Aldo

Le persone che guadagnano più dei rispettivi padri; mostrare nome, reddito e reddito del padre

```

PROJNome, Reddito, RP (SELReddito>RP
  (RENNP,EP,RP ← Nome,Età,Reddito(Persone)
    JOINNP=Padre
    (paternita JOINFiglio =Nome persone)))
  
```

Maternità

Madre	Figlio
Luisa	Maria
Luisa	Luigi
Anna	Olga
Anna	Filippo
Maria	Andrea
Maria	Aldo

Paternità

Padre	Figlio
Sergio	Franco
Luigi	Olga
Luigi	Filippo
Franco	Andrea
Franco	Aldo

Persone

Nome	Età	Reddito
Andrea	27	21
Aldo	25	15
Maria	55	42
Anna	50	35
Filippo	26	30
Luigi	50	40
Franco	60	20
Olga	30	41
Sergio	85	35
Luisa	75	87

Le persone che guadagnano più dei rispettivi padri; mostrare nome, reddito e reddito del padre

```
select f.nome, f.reddito, p.reddito  
from persone p, paternita, persone f  
where p.nome = padre  
and figlio = f.nome  
and f.reddito > p.reddito
```

Maternità

Madre	Figlio
Luisa	Maria
Luisa	Luigi
Anna	Olga
Anna	Filippo
Maria	Andrea
Maria	Aldo

Paternità

Padre	Figlio
Sergio	Franco
Luigi	Olga
Luigi	Filippo
Franco	Andrea
Franco	Aldo

Persone

Nome	Età	Reddito
Andrea	27	21
Aldo	25	15
Maria	55	42
Anna	50	35
Filippo	26	30
Luigi	50	40
Franco	60	20
Olga	30	41
Sergio	85	35
Luisa	75	87

Domanda 7.4.1

Trovare le persone che guadagnano meno delle rispettive madri;
mostrare nome, reddito e reddito della madre.

La precedente interrogazione era:

```
select f.nome, f.reddito, p.reddito
from persone p, paternita, persone f
where p.nome = padre
and figlio = f.nome
and f.reddito > p.reddito
```

Risposta: la nuova e'

```
select f.nome, f.reddito, p.reddito
from persone m, maternita', persone f
where m.nome = madre
and figlio = f.nome
and f.reddito < m.reddito
```

Esercizio 7.4.1

Trovare le persone che guadagnano meno sia dei rispettivi padri che delle rispettive madri. Mostrare nome, reddito, reddito del padre, reddito della madre.

Esaminiamo con maggiore precisione
l'operatore di ridenominazione

Utilita' delle ridenominazioni

Schema R1(A1,A2), R2 (A3,A4)

```
REN B1,B2 ← A1,A4 (
  PROJ A1,A4 (SEL A2 = A3 AND A4 = C1(
    R1 JOIN R2 JOIN REN C1,C2 ← A1,A2 (R1))))
```

Si puo' tradurre in

```
select X.A1 AS B1, Y.A4 AS B2
from R1 X, R2 Y, R1 Z
where X.A2 = Y.A3 AND Y.A4 = Z.A1
```


Ridenominazioni e variabili

Una ridenominazione svolge in SQL il ruolo che le variabili svolgono nei linguaggi programmatici.

Es.

.... From Impiegato I1, Impiegato I2

(puo' essere utile nei Join, vedi piu' avanti)

E' come se fossero create due versioni diverse della stessa tabella Impiegato

Per questa ragione useremo nel seguito per le ridenominazioni anche il termine variabili

SELECT, con ridenominazione del risultato

```
select figlio, f.reddito as reddito,  
p.reddito as redditoPadre  
from persone p, paternita, persone f  
where p.nome = padre  
and figlio = f.nome  
and f.reddito > p.reddito
```

Concetti introdotti

- Clausola DISTINCT
- Join naturale in SQL
- Utilita' delle ridenominazioni