



Operatori aggregati e raggruppamenti

- Le funzioni di aggregazione possono essere applicate a **partizioni delle ennuple** delle relazioni
- Per specificare le partizioni delle ennuple, si utilizza la clausola **group by**:

group by *listaAttributi*



Operatori aggregati e raggruppamenti

Il numero di figli di ciascun padre

```
select padre, count(*) as NumFigli  
from paternita  
group by padre
```

paternita

padre	figlio
Sergio	Franco
Luigi	Olga
Luigi	Filippo
Franco	Andrea
Franco	Aldo



padre	NumFigli
Sergio	1
Luigi	2
Franco	2



Semantica di interrogazioni con operatori aggregati e raggruppamenti

1. Si esegue l'interrogazione **ignorando la group by** e gli operatori aggregati:

```
select *  
from paternita
```

2. Si raggruppano le **ennuple che hanno lo stesso valore per gli attributi che compaiono nella group by**, si produce una ennupla del risultato per ogni gruppo, e si applica l'operatore aggregato a ciascun gruppo



Esercizio 7: group by

Massimo dei redditi per ogni gruppo di persone che sono maggiorenni ed hanno la stessa età (indicando anche l'età)

Esprimere la query in SQL

persone

nome	eta	reddito
-------------	------------	----------------



Esercizio 7: soluzione

Massimo dei redditi per ogni gruppo di persone che sono maggiorenni ed hanno la stessa età (indicando anche l'età)

```
select eta, max(reddito)
from   persone
where  eta > 17
group by eta
```



Raggruppamenti e target list

In una interrogazione che fa uso di **group by**, possono comparire nella target list (oltre a funzioni di aggregazione) **solamente** attributi che compaiono nella **group by**.

Esempio:

Scorretta: redditi delle persone, raggruppati per età.

```
select eta, reddito
from persone
group by eta
```

Potrebbero esistere più valori dell'attributo per lo stesso gruppo.

Corretta: media dei redditi delle persone, raggruppati per età.

```
select eta, avg(reddito)
from persone
group by eta
```



Raggruppamenti e target list

La restrizione sintattica sugli attributi nella `select` vale anche per interrogazioni che semanticamente sarebbero corrette (ovvero, per cui esiste un solo valore dell'attributo per ogni gruppo).

Esempio: i padri col loro reddito, e con reddito medio dei figli.

Scorretta:

```
select padre, avg(f.reddito), p.reddito
from   persone f join paternita on figlio = nome
       join persone p on padre = p.nome
group by padre
```

Corretta:

```
select padre, avg(f.reddito), p.reddito
from   persone f join paternita on figlio = nome
       join persone p on padre = p.nome
group by padre, p.reddito
```



Condizioni sui gruppi

Si possono anche imporre le condizioni di **selezione sui gruppi**. La selezione sui gruppi è **ovviamente diversa** dalla condizione che seleziona le tuple che devono formare i gruppi (clausola **where**). Per effettuare la selezione sui gruppi si usa la clausola **having**, che deve apparire dopo la “**group by**”

Esempio: i padri i cui figli hanno un reddito medio maggiore di 25.

```
select padre, avg(f.reddito)
from   persone f join paternita
      on figlio = nome
group by padre
having avg(f.reddito) > 25
```




Esercizio 8: where o having?

I padri i cui figli sotto i 30 anni hanno un reddito medio maggiore di 20



Esercizio 8: soluzione

I padri i cui figli sotto i 30 anni hanno un reddito medio maggiore di 20

```
select padre, avg(f.reddito)
from   persone f join paternita
      on figlio = nome
where  f.eta < 30
group by padre
having avg(f.reddito) > 20
```



Sintassi, riassumiamo

SelectSQL ::=

select *ListaAttributiOEspressioni*
from *ListaTabelle*
[**where** *CondizioniSemplici*]
[**group by** *ListaAttributiDiRaggruppamento*]
[**having** *CondizioniAggregate*]
[**order by** *ListaAttributiDiOrdinamento*]