



SQL: RAGGRUPPAMENTI, QUANTIFICAZIONE, METODO

Patrizio Dazzi
a.a. 2017 - 2018

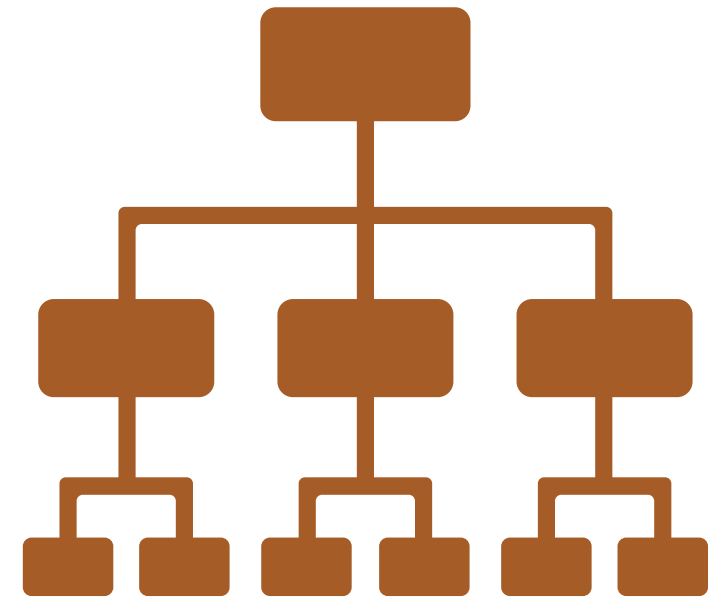
COMUNICAZIONI

- Prove in itinere
 - Primo compitino il 6 novembre
 - Secondo compitino il 19 dicembre
- Homeworks
 - Poche consegne del II... per adesso...
 - Primo homework restituito
 - voti generosi
 - valido per chi ha preso più di 5
 - chi e' stato contattato, venga a parlare con me dopo la lezione



PICCOLO RIASSUNTO DELLA PUNTATA PRECEDENTE

- SQL
 - introduzione
 - concetti
 - SELECT FROM WHERE
 - mapping dell'algebra relazionale
 - Operatori insiemistici
 - unione, intersezione, differenza



**GROUP BY
E
HAVING**



INTERROGAZIONI CON RAGGRUPPAMENTI

- **Nucleo della SELECT**

- SELECT, FROM, [WHERE]

- **Clausola aggiuntiva**

- [ORDER BY]

- **Ulteriori clausole aggiuntive**

- [GROUP BY]

- [HAVING]

CLAUSOLE GROUP BY E HAVING

➤ **GROUP BY**

- operatore di “raggruppamento”

➤ **Sintassi**

- GROUP BY <attributi di raggruppamento>

➤ **Semantica**

- raggruppamento della tabella

- divisione in gruppi delle ennuple

- raggruppamento sulla base dei valori comuni per gli attributi di raggruppamento

CLAUSOLE GROUP BY E HAVING

Esempio: raggruppamento della tabella studenti per ciclo (GROUP BY ciclo)

Studenti

matr	cognome	nome	ciclo	anno	relatore
111	Rossi	Mario	laurea tr.	1	null
222	Neri	Paolo	laurea tr.	2	null
333	Rossi	Maria	laurea tr.	1	null
444	Pinco	Palla	laurea tr.	3	FT
77777	Bruno	Pasquale	laurea sp.	1	FT
88888	Pinco	Pietro	laurea sp.	1	CV

gruppo A
ciclo='laurea tr.'

gruppo B
ciclo='laurea sp.'

CLAUSOLE GROUP BY E HAVING

Esempio: raggruppamento della tabella studenti per ciclo e anno
(GROUP BY ciclo, anno)

Studenti

matr	cognome	nome	ciclo	anno	relatore
111	Rossi	Mario	laurea tr.	1	null
333	Rossi	Maria	laurea tr.	1	null
222	Neri	Paolo	laurea tr.	2	null
444	Pinco	Palla	laurea tr.	3	FT
77777	Bruno	Pasquale	laurea sp.	1	FT
88888	Pinco	Pietro	laurea sp.	1	CV

gruppo A
laurea tr., 1

gruppo B
laurea tr., 2

gruppo C
laurea tr., 3

gruppo D
laurea sp., 1

CLAUSOLE GROUP BY E HAVING

Esempio: raggruppamento della tabella studenti per matricola
(GROUP BY matr)

Studenti

matr	cognome	nome	ciclo	anno	relatore
111	Rossi	Mario	laurea tr.	1	null
333	Rossi	Maria	laurea tr.	1	null
222	Neri	Paolo	laurea tr.	2	null
444	Pinco	Palla	laurea tr.	3	FT
77777	Bruno	Pasquale	laurea sp.	1	FT
88888	Pinco	Pietro	laurea sp.	1	CV

una ennupla
per ogni
gruppo

CLAUSOLE GROUP BY E HAVING

- **Caratteristiche dei gruppi**
 - collezioni di ennuple
 - valori comuni per gli attributi di raggruppam.
- **Operazioni interessanti sui gruppi**
 - funzioni aggregative
 - analisi della distribuzione di valori tra i gruppi
se: numero di studenti per ciclo o per anno
 - OLAP (“On Line Analytical Processing”)

CLAUSOLE GROUP BY E HAVING

- **Interrogazioni con raggruppamento**
 - attributi di raggruppamento (nella GROUP BY)
 - proiezioni su attributi di raggruppamento e funzioni aggregative applicate al gruppo (nella SELECT)
 - condizioni sui gruppi (che coinvolgono funzioni aggregative) (nella HAVING)

CLAUSOLE GROUP BY E HAVING

➤ Esempio: numero di studenti per ciclo

- `SELECT ciclo, COUNT(*)`
- `FROM Studenti`
- `GROUP BY ciclo;`

ciclo	COUNT(*)
laurea tr.	4
laurea sp.	2

➤ Semantica

- viene valutata la clausola FROM
- viene effettuato il raggruppamento secondo la GROUP BY
- viene valutata la clausola SELECT per ciascun gruppo
(ogni gruppo contribuisce ad **UNA** sola ennupla del risultato)

CLAUSOLE GROUP BY E HAVING

➤ **Esempio:** numero di studenti per ciclo

- `SELECT COUNT(*)`
- `FROM Studenti`
- `GROUP BY ciclo;`

COUNT(*)
4
2

CLAUSOLE GROUP BY E HAVING

➤ Vincoli sintattici sulla SELECT

- se c'è una GROUP BY, solo gli attributi di raggruppamento possono comparire nella SELECT

```
SELECT ciclo, count(*)  
FROM Studenti  
GROUP BY ciclo;
```

```
SELECT count(*)  
FROM Studenti  
GROUP BY ciclo;
```

```
SELECT anno, count(*)  
FROM Studenti  
GROUP BY ciclo;
```

CLAUSOLE GROUP BY E HAVING

- Esempio: distribuzione per anno degli studenti della laurea triennale

```
SELECT anno, count(*) as numstud  
FROM Studenti  
WHERE ciclo='laurea tr.'  
GROUP BY anno;
```

CLAUSOLE GROUP BY E HAVING

I passo: WHERE ciclo='laurea tr.'

matr	cognome	nome	ciclo	anno	relatore
111	Rossi	Mario	laurea tr.	1	null
333	Rossi	Maria	laurea tr.	1	null
222	Neri	Paolo	laurea tr.	2	null
444	Pinco	Palla	laurea tr.	3	FT

Risultato finale:
SELECT anno,
count(*) as numstud

anno	numstud
1	2
2	1
3	1

II passo: GROUP BY anno

matr	cognome	nome	ciclo	anno	relatore
111	Rossi	Mario	laurea tr.	1	null
333	Rossi	Maria	laurea tr.	1	null
222	Neri	Paolo	laurea tr.	2	null
444	Pinco	Palla	laurea tr.	3	FT

CLAUSOLE GROUP BY E HAVING

- **Esempio:** distribuzioni delle medie, solo per i corsi con più di 2 esami
- non è possibile usare la WHERE
- **HAVING:** condizioni aggregate su gruppi

```
SELECT corso, avg(voto) as votomedio
FROM Esami
GROUP BY corso
HAVING count(voto)>2;
```

CLAUSOLE GROUP BY E HAVING

Esami

studente	corso	voto	lode
111	PR1	27	false
88888	PR1	30	false
77777	PR1	21	false
111	INFT	24	false
88888	INFT	30	true
222	ASD	30	true
77777	ASD	20	false
88888	ASD	28	false

I passo:
raggruppamento
secondo la
GROUP BY

II passo:
selezione dei
gruppi secondo
la HAVING

III passo:
proiezione e rid.
secondo la
SELECT

Risultato finale

corso	votomedio
PR1	26
ASD	26

FORMA GENERALE DELLA SELECT

- Forma generale della **SELECT**
 - **SELECT** [DISTINCT] <risultato>
 - **FROM** <join o prodotti cartesiani>
 - [WHERE <condizioni>]
 - [GROUP BY <attributi di raggruppamento>]
 - [HAVING <condizioni sui gruppi>]
 - [ORDER BY <attributi di ordinamento>]

FORMA GENERALE DELLA SELECT

- **Se la GROUP BY manca**
 - tutte le ennuple ottenute dopo la WHERE vengono considerate un unico gruppo
 - in questo caso le funzioni aggregative producono un unico valore e non sono ammessi attributi ordinari nella SELECT

- **Nota sulla semantica**
 - tutti i valori NULL normalmente vengono raggruppati assieme

FORMA GENERALE DELLA SELECT

- Una semantica operativa
 - viene valutata la clausola FROM
 - join o prodotti cartesiani >> unica tabella
 - viene valutata la clausola WHERE
 - selezione delle ennuple della tabella
 - viene valutata l'eventuale GROUP BY
 - raggruppamento delle ennuple della tabella
 - viene valutata l'eventuale HAVING
 - selezione dei gruppi della tabella

FORMA GENERALE DELLA SELECT

- Una semantica operativa (continua)
 - viene valutata la clausola SELECT
 - proiezioni, espressioni e funzioni aggregative
 - ridenominazioni
 - eventuale eliminazione di duplicati
 - viene valutata la clausola ORDER BY
 - ordinamenti finali

FORMA GENERALE DELLA SELECT

- **Esempio:** medie in ordine decrescente degli studenti della laurea specialistica che hanno sostenuto almeno due esami

```
SELECT matr, cognome, nome, avg(voto)
FROM Studenti JOIN Esami ON matr=studente
WHERE ciclo='laurea sp.'
GROUP BY matr, cognome, nome
HAVING count(*)>=2
ORDER BY avg(voto) DESC;
```

FORMA GENERALE DELLA SELECT

Studenti

matr	cognome	nome	ciclo	relat
111	Rossi	Mario	laurea tr.	null
333	Rossi	Maria	laurea tr.	null
222	Neri	Paolo	laurea tr.	null
444	Pinco	Palla	laurea tr.	FT
77777	Bruno	Pasquale	laurea sp.	FT
88888	Pinco	Pietro	laurea sp.	CV

Esami

studente	corso	voto	lode
111	PR1	27	false
222	ASD	30	true
111	INFT	24	false
77777	PR1	21	false
77777	ASD	20	false
88888	ASD	28	false
88888	PR1	30	false
88888	INFT	30	true

FORMA GENERALE DELLA SELECT

Passo 1: FROM Studenti JOIN Esami ON matr=studente

matr	cognome	nome	ciclo	relat	studente	corso	voto	lode
111	Rossi	Mario	laurea tr.	null	111	PR1	27	false
111	Rossi	Mario	laurea tr.	null	111	INFT	24	false
222	Neri	Paolo	laurea tr.	null	222	ASD	30	true
77777	Bruno	Pasquale	laurea sp.	FT	77777	PR1	21	false
77777	Bruno	Pasquale	laurea sp.	FT	77777	ASD	20	false
88888	Pinco	Pietro	laurea sp.	VC	88888	ASD	28	false
88888	Pinco	Pietro	laurea sp.	VC	88888	PR1	30	false
88888	Pinco	Pietro	laurea sp.	VC	88888	INFT	30	true

FORMA GENERALE DELLA SELECT

Passo II: WHERE ciclo='laurea sp.'

matr	cognome	nome	ciclo	relat	studente	corso	voto	lode
77777	Bruno	Pasquale	laurea sp.	FT	77777	PR1	21	false
77777	Bruno	Pasquale	laurea sp.	FT	77777	ASD	20	false
88888	Pinco	Pietro	laurea sp.	VC	88888	ASD	28	false
88888	Pinco	Pietro	laurea sp.	VC	88888	PR1	30	false
88888	Pinco	Pietro	laurea sp.	VC	88888	INFT	30	true

Passo III: GROUP BY matr, cognome, nome

matr	cognome	nome	ciclo	relat	studente	corso	voto	lode
77777	Bruno	Pasquale	laurea sp.	FT	77777	PR1	21	false
77777	Bruno	Pasquale	laurea sp.	FT	77777	ASD	20	false
88888	Pinco	Pietro	laurea sp.	VC	88888	ASD	28	false
88888	Pinco	Pietro	laurea sp.	VC	88888	PR1	30	false
88888	Pinco	Pietro	laurea sp.	VC	88888	INFT	30	true

FORMA GENERALE DELLA SELECT

Passo IV: HAVING count(*) >= 2

matr	cognome	nome	ciclo	relat	studente	corso	voto	lode
77777	Bruno	Pasquale	laurea sp.	FT	77777	PR1	21	false
77777	Bruno	Pasquale	laurea sp.	FT	77777	ASD	20	false
88888	Pinco	Pietro	laurea sp.	VC	88888	ASD	28	false
88888	Pinco	Pietro	laurea sp.	VC	88888	PR1	30	false
88888	Pinco	Pietro	laurea sp.	VC	88888	INFT	30	true

Passo V: SELECT matr, cognome, nome, avg(voto)

matr	cognome	nome	avg(voto)
77777	Bruno	Pasquale	20,5
88888	Pinco	Pietro	29,66

Passo VI: ORDER BY avg(voto) DESC

matr	cognome	nome	avg(voto)
88888	Pinco	Pietro	29,66
77777	Bruno	Pasquale	20,5

FINE DELLA LEZIONE