

Aggregazione

- Il contenuto delle basi di dati viene spesso aggregato:
 - Il voto massimo nell'esame di basi di dati
 - Il voto medio per ciascun esame
 - Il numero di CFU conseguiti da ciascuno studente

Aggregazione, sintassi e semantica

- Sintassi (usiamo anche GROUP BY al posto di AGG)
AGG *Attributi* *Raggruppamento*; *Funzione (Attributo), ... (Operando)*
 - *Funzione*: count, sum, avg, max, min
- Semantica
 - Il risultato contiene la proiezione sugli attributi di raggruppamento e il valore delle funzioni in corrispondenza al sottoinsieme raggruppato
- simbolo più compatto
 - α (alpha, da “aggregazione”)
 - γ (gamma, da “group by”)

Semantica, in due passi

- Partizione sulla base degli attributi di raggruppamento e calcolo delle funzioni su ciascuna partizione
- Se non ci sono attributi di raggruppamento, le funzioni sono calcolate sull'intera relazione

esami	Matricola	Voto	Codice
	3456	30	04
	3456	26	02
	9283	27	01
	6554	26	01
	6554	26	05

- Il numero totale di esami

AGG NumeroEsami ← count(*) (esami)

NumeroEsami
5

esami	Matricola	Voto	Codice
	3456	30	04
	3456	26	02
	9283	27	01
	6554	26	01
	6554	26	05

- Il voto medio complessivo

AGG MediaGenerale \leftarrow avg(Voto) (esami)

MediaGenerale
26

esami	Matricola	Voto	Codice
	3456	30	04
	3456	26	02
	9283	27	01
	6554	26	01
	6554	26	05

- Il voto medio per ciascuno studente

AGG `Matricola; Media ← avg(Voto) (esami)`

Matricola	Media
3456	28
9283	27
6554	26

esami	Matricola	Voto	Codice
	3456	30	04
	3456	26	02
	9283	27	01
	6554	26	01
	6554	26	05

- Numero esami e voto medio per ciascuno studente

AGG `Matricola; Media ← avg(Voto), NumEsami←count(*)` (esami)

Matricola	Media	NumEsami
3456	28	2
9283	27	1
6554	26	2

Valori nulli

- vengono trattati correttamente (cioè ignorati nelle medie, nelle somme e nei conteggi)

count

- `count(*)` conta le ennuple
- `count(A)` conta le ennuple che hanno il valore di A
 - non i valori distinti di A
 - quindi in assenza di valori il risultato è lo stesso, in presenza di valori nulli per A
 - `count(A)` è il numero di ennuple che non hanno il valore nullo

esami	Matricola	Voto	Codice
	3456	30	04
	3456	NULL	02
	9283	27	01
	6554	26	01
	6554	26	05

- Numero esami e voto medio per ciascuno studente

AGG `Matricola; Media ← avg(Voto), NumEsami ← count(*)` (**esami**)

Matricola	Media	NumEsami
3456	30	2
9283	27	1
6554	26	2

- Numero voti e voto medio per ciascuno studente

AGG `Matricola; Media ← avg(Voto), NumVoti ← count(Voto)` (**esami**)

Matricola	Media	NumVoti
3456	30	1
9283	27	1
6554	26	2