

HOMEWORK BI

DARIO CAROLLA MAT. 807547

MATTEO GAVERINI MAT. 808101

FEDERICO MANENTI MAT. 790032

PREPROCESSING

- ▶ Cambio formato data (YYYY-MM-DD)
- ▶ Aggiunta colonna "Anno_accademico"
- ▶ Gli anni accademici 2014/15 e 2018/19 non sono stati considerati



	CdS	AD	DtAppello	Anno_accademico
1	[G8501R] SCIENZE DELLA FORMAZIONE PRIMARIA	[G8501R033] ATTIVITA DI TIROCINIO DEL III ANNO	2017-07-25	2016/17
2	[E3101Q] INFORMATICA	[E3101Q103] BASI DI DATI	2017-06-16	2016/17
3	[E3101Q] INFORMATICA	[E3101Q103] BASI DI DATI	2017-07-17	2016/17
4	[E3101Q] INFORMATICA	[E3101Q108] LINGUAGGI DI PROGRAMMAZIONE	2017-01-31	2016/17
5	[E3101Q] INFORMATICA	[E3101Q109] ANALISI E PROGETTAZIONE DEL SOFTWARE	2017-06-16	2016/17
6	[E3101Q] INFORMATICA	[E3101Q110] RETI E SISTEMI OPERATIVI	2017-02-22	2016/17
7	[E3101Q] INFORMATICA	[E3101Q110] RETI E SISTEMI OPERATIVI	2017-01-30	2016/17

DOMANDA 1

“Distribuzione del numero degli studenti iscritti nei vari appelli, suddivisa per anni e per corso di laurea”

Considerazioni:

- ▶ Ogni studente è contato per tutte le iscrizioni effettuate
- ▶ Il conteggio delle iscrizioni è suddiviso per anno accademico
- ▶ Rimozione anni accademici 2014/15 e 2018/19 perché non rilevanti

Risultati:

Visualizzazione

DOMANDA 2

“Individuazione della Top-10 degli esami più difficili suddivisi per corso di studi”

Considerazioni:

- ▶ Top-10 per ogni anno accademico
- ▶ Gli assenti ed i ritirati sono considerati come bocciati
- ▶ Si considerano solo gli esami con più di tre iscritti

Risultati:

Visualizzazione

DOMANDA 3

“Individuazione dei corsi di laurea ad elevato tasso di commitment, ovvero appelli di esami diversi ma del medesimo corso di laurea che si sono svolti nello stesso giorno”

Considerazioni:

- ▶ Gli appelli della stessa materia, svolti nel medesimo giorno ma solo con professori diversi sono considerati come unico appello.

Risultati:

Visualizzazione

DOMANDA 4

“Individuazione della Top-3 degli esami con media voti maggiore e minore rispettivamente, calcolati per ogni singolo corso di studi”

Considerazioni:

- ▶ Gli esami sono suddivisi per anno accademico
- ▶ Si considerano solo gli esami con più di tre iscritti

Risultati:

Visualizzazione Top-3

Visualizzazione Bottom-3

DOMANDA 5

“Calcolare la distribuzione degli studenti “fast&furious” per corso di studi, ovvero studenti con il rapporto “votazione media riportata negli esami superati” su “periodo di attività” maggiore”

Considerazioni:

- ▶ Si considerano solo gli studenti che hanno tentato almeno tre appelli e con una carriera di almeno un mese

Risultati:

Visualizzazione

DOMANDA 6

“Individuazione della Top-3 degli esami “trial&error”, ovvero esami che richiedono il maggior numero di tentativi prima del superamento”

Considerazioni:

- ▶ Si considerano solo gli studenti insufficienti che hanno preso parte all'appello
- ▶ Si considerano solo gli esami con più di tre iscritti

Risultati:

Visualizzazione

DOMANDA 7

“Scoprire per ogni esame di ciascun corso di studi il mese con la maggior percentuale di promossi”

Considerazioni:

- ▶ Gli assenti ed i ritirati sono considerati come bocciati
- ▶ Si considerano solo i mesi in cui il totale degli iscritti è maggiore di tre
- ▶ Il tasso di superamento non è calcolato sui singoli anni accademici ma complessivamente

Risultati:

Visualizzazione

CONCLUSIONI

Tempi di esecuzione (MacBook Air 2017):

Query	Normalizzato (ms)	Denormalizzato (ms)
1	1537 ± 46	2016 ± 55
2	17350 ± 265	5136 ± 262
3	436 ± 23	2806 ± 67
4 (bottom e top)	3559 ± 125 e 3591 ± 128	5254 ± 124 e 5084 ± 118
5	4919 ± 105	4381 ± 97
6	22442 ± 215	4989 ± 102
7	18226 ± 245	5246 ± 124

Osservazioni:

- ▶ Nelle query più semplici (es. 1) il DB normalizzato risulta più veloce.
- ▶ Nelle query con più JOIN il DB denormalizzato risulta più veloce.