

Compito di Sistemi Informativi I
(Ing. Informatica + Ing. Elettr. e TLC fino AA 2001/02)
del 19 febbraio 2003

ESERCIZIO 1

Date le relazioni:

ESAMI			
<u>MATRICOLA</u>	<u>ESAME</u>	<u>DATA</u>	<u>VOTO</u>
1234	ANALISI B	2002-03-15	30
2345	FISICA A	2002-11-04	27
4532	ECONOMIA A	2001-06-24	18
1234	ANALISI A	2000-07-02	24
...

STUDENTI

<u>MATRICOLA</u>	<u>NOME</u>	<u>RESIDENZA</u>	<u>DATANASCITA</u>	<u>CDL</u>
1234	CARLO ROSSI	MODENA	1980-10-10	GESTIONALE
2345	ANNA CORLI	FERRARA	1981-12-07	ELETRONICA
4532	MARTA VERDI	BOLOGNA	1979-06-01	GESTIONALE
...

Si scriva una query SQL che fornisca il nome e la media degli studenti che hanno preso **tutti** i voti dal 20 al 30 (compresi). **Facoltativo:** si scriva anche una query SQL che restituisca la città in cui abitano gli studenti che, considerando soltanto esami sostenuti dopo il 1999, hanno dato il maggior numero di esami ed hanno la media più bassa.

ESERCIZIO 2

Si spieghi brevemente da cosa nasce il problema dell'**overflow** nei *file hash statici*, come viene trattato e con quale impatto sulle prestazioni dell'organizzazione.

ESERCIZIO 3

Si illustrino, anche attraverso esempi significativi, i principali **problemi** causati da uno schema relazionale che non è in **2NF**, in **3NF** o in **BCNF**.

ESERCIZIO 4

Date le relazioni dell'**ESERCIZIO 1**, utilizzando *parametri simbolici* (es. NT_{ESAMI}, dim(disk_block), dim(RESIDENZA), NK_{DATANASCITA}, min(DATA) etc.), **dimensionare** i seguenti indici: *unclustered* su **STUDENTI.MATRICOLA**, *clustered* su **DATA** e su **ESAMI.MATRICOLA**, *unclustered (TID ord.)* su **RESIDENZA**. Si scrivano poi le **formule di costo** necessarie all'ottimizzazione dell'espressione:

$$\sigma_{DATA \geq 1/1/98 \wedge DATA \leq 12/31/99 \wedge RESIDENZA \text{ like 'P\%' }} (ESAMI \bowtie STUDENTI)$$

considerando *nested loops*, *merging scans* e *simple hash* come metodi di join.