Basi di dati II, primo modulo — Tecnologia delle basi di dati 24 settembre 2010 — Compito A

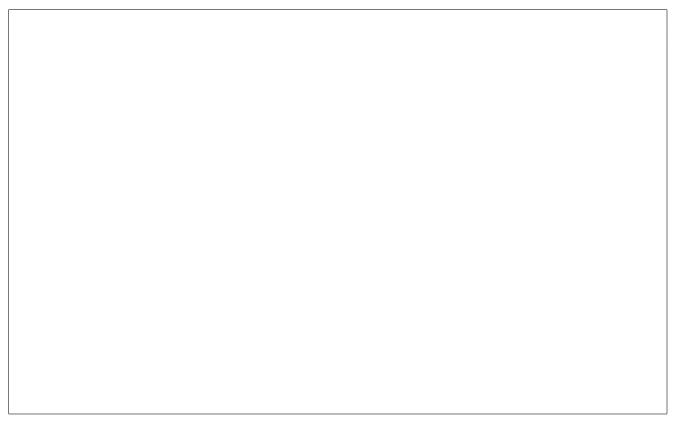
Rispondere alle prime tre domande su questo foglio e alla quarta sul foglio separato. Tempo a disposizione: un'ora e trenta minuti.

Cognome	Nome	Matricola	Ordin
ricerca di tutti i clienti	i consideri una relazione R(CodiceC di una certa regione, un indice second no (mostrare anche esempi che illus	dario sull'attributo Regione potr	ebbe in alcuni casi risultare
a (ad esempio, ma è in	Si consideri una relazione $R(\underline{A} \ B \ C)$ rrilevante, 4Byte), molto più piccolande (N ennuple) e che le operazioni	a della dimensione del blocco pa	ri a P. Si supponga che la
o ₁ SELECT * FROM R	ORDER BY A, con frequenza f_1		
o_2 SELECT A, B, C F	ROM R ORDER BY A, con frequenza j	$f_2 = 10 \times f_1$	
_	i alternative di memorizzazione, cal ll'unità di tempo, in base alle varial	- \ -	ortare la formula che indica
(i) memorizzazione de	lla relazione $R(\underline{A}\ B\ C\ D\ E)$ ordinat	a su A	
costo unitario di	o ₁ :	_ costo unitario di o_2 :	
costo complessivo	D:		
(ii) memorizzazione de	elle proiezioni $R1(\underline{A}\ B\ C)$ e $R2(\underline{A}\ I$	O(E), entrambe ordinate su A	
costo unitario di	o ₁ :	costo unitario di o_2 :	
costo complessivo	o:		

Supporre che il join possa essere eseguito con il metodo merge-join (e che il costo del join stesso sia trascurabile rispetto alle due scansioni).

Domanda 3 (20%) Si consideri una forma di equivalenza fra schedule denominata *final-state*-equivalenza, secondo la quale due schedule S_1 e S_2 sono equivalenti se, per ogni istanza d della base di dati, essi la trasformano nello stesso modo (e quindi $S_1(d) = S_2(d)$, se con S(d) indichiamo l'istanza della base di dati ottenuta applicando a d le operazioni dello schedule S).

- 1. Formulare una definizione per questa proprietà, variante della definizione di view-equivalenza.
- 2. Spiegare, almeno intuitivamente, il rapporto che esiste fra final-state-equivalenza e view-equivalenza (sono equivalenti, incomparabili, oppure una implica l'altra?).



Domanda 4 (40%) Una catena di negozi gestisce le attività utilizzando, in ciascun negozio, una base di dati con le seguenti relazioni (su cui sono definiti gli ovvi vincoli di integrità referenziale):

- Prodotti(<u>CodiceProdotto</u>, Descrizione, Prezzo, Categoria, Marca)
- Marca(Codice, Nome)
- Categorie(Codice, Descrizione, Macro Categoria)
- MacroCategorie(Codice, Descrizione)
- Vendite(NumeroScontrino,Ora)
- DettaglioVendite(<u>NumeroScontrino,CodiceProdotto,Quantità</u>)

Si noti che

- Le informazioni relative alle vendite vengono mantenute solo nel corso della giornata.
- Il prezzo di un prodotto può variare da un giorno all'altro.

Utilizzando tali informazioni, la catena vuole realizzare un data mart relativo alle vendite dei prodotti, avente come misure le quantità vendute e gli incassi, che permetta di effettuare analisi di tipo temporale (incluse, oltre ai giorni, anche le fasce orarie della giornata, ad esempio 9-10, 10-11 e così via, oppure mattina e pomeriggio) e su prodotti (singoli o per categoria e/o per marca) e sui negozi. Allo scopo, specificare un possibile dettaglio del data mart; in particolare

- 1. **specificare esplicitamente la grana scelta**, supponendo che la quantità delle vendite sia tale che si è deciso di non utilizzare il massimo livello di dettaglio, ma solo quello strettamente indispensabile (in altri termini, la grana non deve essere il singolo dettaglio di vendita, ma una opportuna aggregazione);
- 2. mostrare gli schemi delle tabelle (tabella dei fatti e tabelle delle dimensioni) indicare anche le dimensioni i cui dati non provengono dalla base di dati sopra mostrata

(Rispondere sul foglio separato)

Cognome	Nome	Matricola	Ordin.

Basi di dati II, primo modulo — Tecnologia delle basi di dati 24 settembre 2010 — Compito B

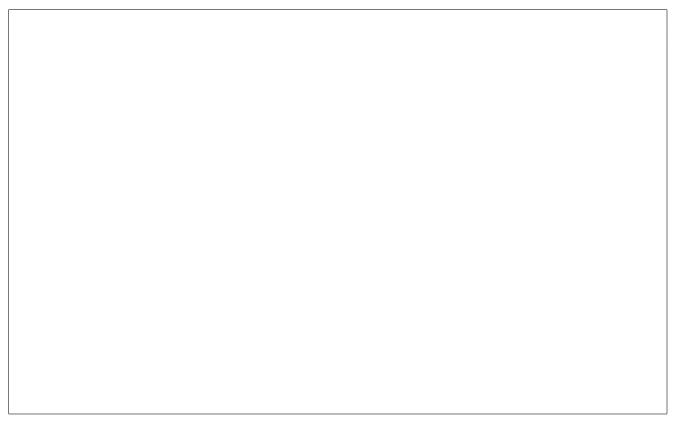
Rispondere alle prime tre domande su questo foglio e alla quarta sul foglio separato. Tempo a disposizione: un'ora e trenta minuti.

Cognome	Nome	Matricola	Ordin
ricerca di tutti i clienti	consideri una relazione R(Codice di una certa città, un indice secon	dario sull'attributo Città potrel	obe in alcuni casi risultare
conveniente e in aitri n	o (mostrare anche esempi che illus	strino le due situazioni da un pu	nto di vista quantitativo).
d (ad esempio, ma è irr	i consideri una relazione $R(\underline{A} \ B \ C)$ rilevante, 4Byte), molto più piccolade (L ennuple) e che le operazioni	a della dimensione del blocco pa	ri a P. Si supponga che la
o_1 SELECT * FROM R O	RDER BY A, con frequenza f_1		
o_2 SELECT A, B, C FR	OM R ORDER BY A, con frequenza	$f_2 = 10 \times f_1$	
~	alternative di memorizzazione, cal l'unità di tempo, in base alle varial	- \ -	rtare la formula che indica
(i) memorizzazione del	la relazione $R(\underline{A}\ B\ C\ D\ E)$ ordinat	ta su A	
costo unitario di α	<i>p</i> ₁ :	$_$ costo unitario di o_2 : $_$	
costo complessivo	:		
(ii) memorizzazione de	lle proiezioni $R1(\underline{A}\ B\ C)$ e $R2(\underline{A}\ I$	O(E), entrambe ordinate su A	
costo unitario di α	p ₁ :	_ costo unitario di o_2 :	
costo complessivo	: - <u></u>		

Supporre che il join possa essere eseguito con il metodo merge-join (e che il costo del join stesso sia trascurabile rispetto alle due scansioni).

Domanda 3 (20%) Si consideri una forma di equivalenza fra schedule denominata *final-state*-equivalenza, secondo la quale due schedule S_1 e S_2 sono equivalenti se, per ogni istanza i della base di dati, essi la trasformano nello stesso modo (e quindi $S_1(i) = S_2(i)$, se con S(i) indichiamo l'istanza della base di dati ottenuta applicando a i le operazioni dello schedule S).

- 1. Formulare una definizione per questa proprietà, variante della definizione di view-equivalenza.
- 2. Spiegare, almeno intuitivamente, il rapporto che esiste fra final-state-equivalenza e view-equivalenza (sono equivalenti, incomparabili, oppure una implica l'altra?).



Domanda 4 (40%) Una catena di negozi gestisce le attività utilizzando, in ciascun negozio, una base di dati con le seguenti relazioni (su cui sono definiti gli ovvi vincoli di integrità referenziale):

- Prodotti(CodiceProdotto,Descrizione,Prezzo,Categoria,Produttore)
- Produttore(Codice, Nome)
- Categorie(Codice, Descrizione, Macro Categoria)
- MacroCategorie(<u>Codice</u>, Descrizione)
- Vendite(NumeroScontrino,Ora)
- DettaglioVendite(<u>NumeroScontrino,CodiceProdotto,Quantità</u>)

Si noti che

- Le informazioni relative alle vendite vengono mantenute solo nel corso della giornata.
- Il prezzo di un prodotto può variare da un giorno all'altro.

Utilizzando tali informazioni, la catena vuole realizzare un data mart relativo alle vendite dei prodotti, avente come misure le quantità vendute e gli incassi, che permetta di effettuare analisi di tipo temporale (incluse, oltre ai giorni, anche le fasce orarie della giornata, ad esempio 9-10, 10-11 e così via, oppure mattina e pomeriggio) e su prodotti (singoli o per categoria e/o per produttore) e sui negozi. Allo scopo, specificare un possibile dettaglio del data mart; in particolare

- 1. **specificare esplicitamente la grana scelta**, supponendo che la quantità delle vendite sia tale che si è deciso di non utilizzare il massimo livello di dettaglio, ma solo quello strettamente indispensabile (in altri termini, la grana non deve essere il singolo dettaglio di vendita, ma una opportuna aggregazione);
- 2. mostrare gli schemi delle tabelle (tabella dei fatti e tabelle delle dimensioni) indicare anche le dimensioni i cui dati non provengono dalla base di dati sopra mostrata

(Rispondere sul foglio separato)

Cognome	Nome	Matricola	Ordin.