Compito Basi di Dati. Tempo concesso: 2 ore			26 Giugno 2019
Nome:	Cognome:	Matricola:	

Si considerino le seguenti specifiche relative alla realizzazione della base dati per la gestione delle riunioni organizzate nell'ambito di un progetto di ricerca.

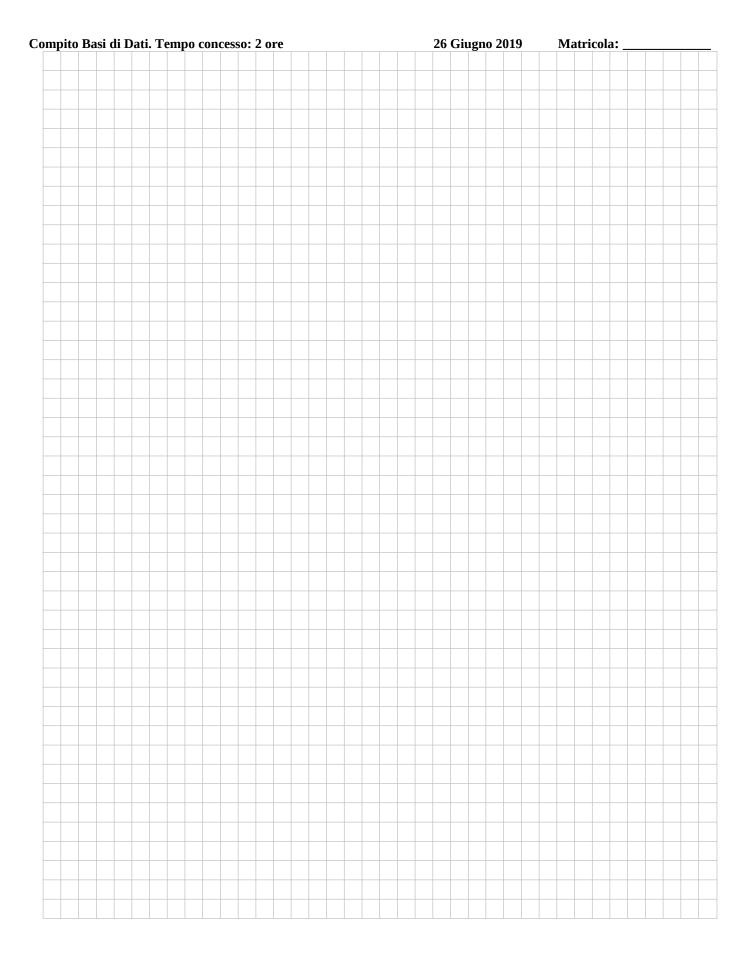
Si vuole realizzare un sistema informativo per la gestione delle riunioni organizzate nell'ambito di un progetto di ricerca. Al progetto partecipano diversi partner (università o aziende) appartenenti all'Unione europea; per ciascun partner si memorizzano una sigla identificativa (es. UNIBO per l'Università di Bologna), il nome completo, la nazione di appartenenza, il legale rappresentante, l'indirizzo e un recapito email.

I partner del progetto si riuniscono periodicamente per discutere dei progressi dell'attività di ricerca svolta. Ciascun partner del progetto può essere rappresentato alle diverse riunioni da una o più persone che devono essere memorizzate nel sistema. Uno stesso partner può essere rappresentato da persone diverse in riunioni diverse.

Per ciascuna riunione si devono memorizzare la data di inizio e di fine, il luogo in cui si svolge (nome, indirizzo, città, nazione) e il programma giornaliero dettagliato. In particolare, il programma di ciascuna giornata consiste in un elenco di eventi, ciascuno dei quali può essere una presentazione o un momento ricreativo (es. coffee break o pranzo). Per ciascun evento si memorizzano l'orario di inizio, l'orario di fine e una descrizione. Inoltre per le presentazioni si memorizzano il partner responsabile e il relatore (una delle persone partecipanti alla riunione in rappresentanza del partner responsabile), mentre per i momenti ricreativi si memorizzano il luogo di svolgimento ed eventuali note.

Si definisca il relativo **schema E/R** (<u>nella metodologia proposta a lezione</u>) e si evidenzino eventuali **vincoli inespressi e dati derivati**.





1. Nell'ambito degli schemi relazionali, descrivere formalmente il concetto di dipendenza funzionale.

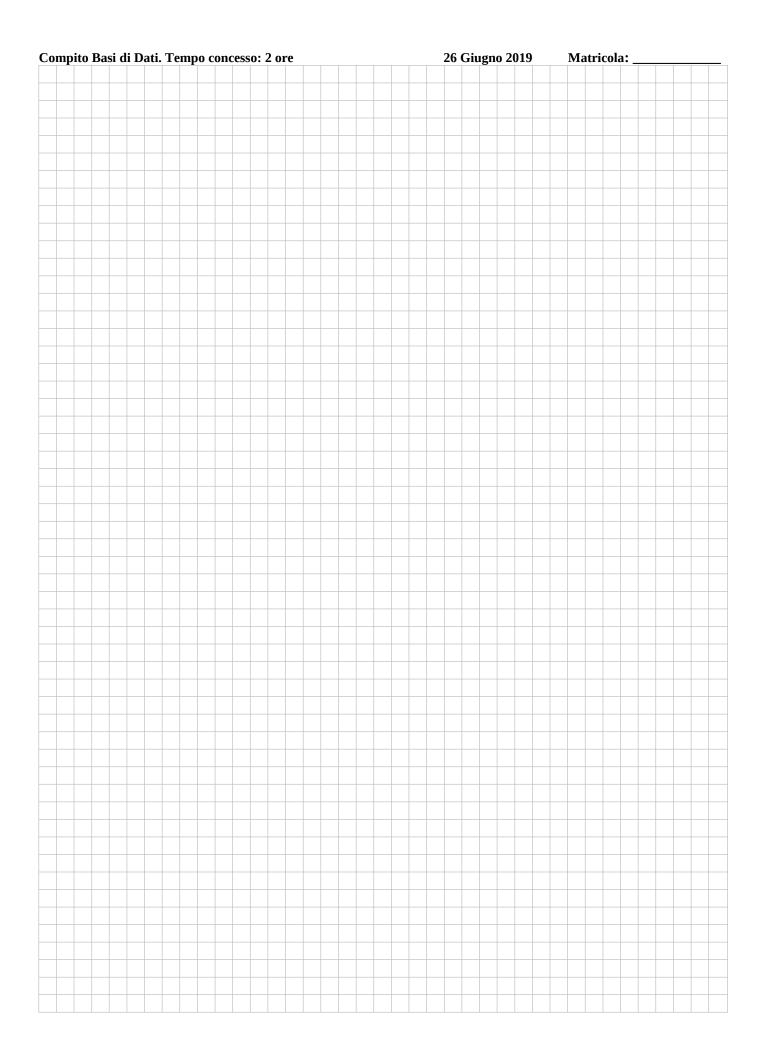
# 2. Dato il seguente schema relazionale:

MOSTREFOTOGRAFICHE (CodFotografia, Titolo, CodFotografo, NominativoFotografo, DataFotografia, DataNascitaFotografo, NazionalitàFotografo, CodMostra, NomeMostra, CittàMostra, IndirizzoMostra, NazioneMostra, DataInizio, DataFine)

## Sapendo che:

- ogni fotografia è identificata da un codice univoco; si memorizzano inoltre il titolo, il fotografo che l'ha scattata e la data;
- per ciascun fotografo, identificato da un codice univoco, si memorizzano la data di nascita e la nazionalità;
- ogni mostra è caratterizzata da codice univoco, nome, indirizzo, città, nazione e periodo di svolgimento (DataInizio, DataFine);
- una fotografia può essere esposta in più mostre e ogni mostra, naturalmente, include più fotografie;

evidenziare la chiave della relazione, elencare tutte le dipendenze funzionali non banali presenti nello schema (specificandone il tipo nel caso siano "problematiche") e decomporre lo schema in terza forma normale.



È data la query:

SELECT punto\_vendita, fattura\_emessa, importo FROM Vendita WHERE fattura\_emessa = 1 AND punto\_vendita = 'P28' ORDER BY importo DESC

#### sulla relazione:

VENDITA(<u>id</u>, punto\_vendita, importo, corrispettivo\_emesso, fattura\_emessa). Il file dati è ordinato su *id*.

Su *punto\_vendita* è costruito un indice unclustered (*B+-tree*), con NL<sub>punto\_vendita</sub> = 653.

### Si determinino:

- la selettività dei predicati (f(p1), f(p2))
- il numero di record attesi dopo l'esecuzione della query (ER<sub>Result</sub>)
- il miglior piano di accesso per la risoluzione della query tra  $C_{\text{Seq}}$  e  $C_{\text{punto\_vendita}}$

tenendo conto dei seguenti dati:

NP = 7677, NR = 450000

 $NK_{punto\_vendita} = 50$ ,  $fattura\_emessa \in \{0,1\}$ ,  $corrispettivo\_emesso \in \{0,1\}$ 

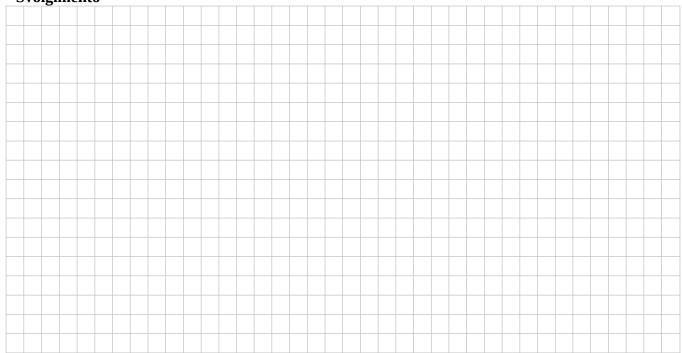
Len(id) = 8 Byte, Len(punto\_vendita) = 40 Byte, Len(importo) = 8 Byte, Len(corrispettivo\_emesso) = 1 Byte, Len(fattura\_emessa) = 1 Byte

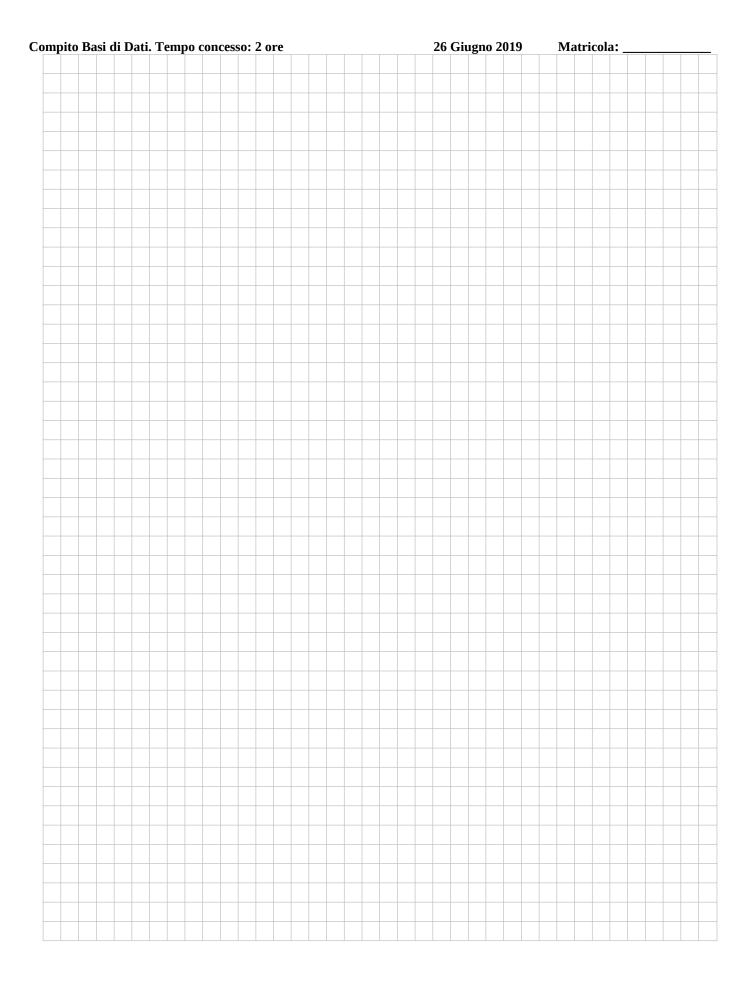
D = 4096 Byte, header della pagina = 96 Byte

Si ipotizzi di utilizzare l'algoritmo Sort-Merge a Z=3 vie per l'eventuale ordinamento del risultato. Nella valutazione del costo di sort si usi la formula approssimata.

Nella valutazione del costo relativo all'indice, si trascuri il costo di discesa dell'albero.







Si consideri il seguente schema relazionale relativo a un sistema di gestione di riunioni di condominio:

RIUNIONI (codRiunione, data, codPresidente: CONDOMINI) N.B. UNIQUE data

Presenze (codRiunione: RIUNIONI, codCondomino: CONDOMINI)

ORDINEDELGIORNO (codRiunione: RIUNIONI, punto, titolo, argomento, codRelatore: CONDOMINI)

CONDOMINI (codCondomino, nome, cognome)

- 1. Scrivere un'espressione di algebra relazionale che visualizzi i condomini che hanno partecipato a tutte le riunioni del 2019 (codCondomino, nome, cognome).
- 2. Scrivere una query SQL che visualizzi i condomini che non erano presenti all'ultima riunione di condominio (nome, cognome).
- 3. Scrivere una query SQL che visualizzi l'elenco dei condomini che hanno presieduto almeno 2 riunioni di condominio nell'anno 2018 (nome, cognome, numeroRiunioni).
- 4. Scrivere una query SQL che visualizzi, per ciascuna riunione, il numero di punti all'ordine del giorno e il numero di relatori <u>distinti</u> (codRiunione, data, numeroPunti, numeroRelatori).

Svolgimento

