

Basi di dati A.A. 2022/23

Esercitazione N. 5

Maurizio Lenzerini, Riccardo Valentini

1. La base di dati già preparata

Si consideri una base di dati in SQL relativa ad una applicazione di social network, con le seguenti tabelle:

```
CREATE TABLE comune (  
  nome VARCHAR(20),  
  provincia VARCHAR(20),  
  regione varchar(20),  
  CONSTRAINT comune_pk PRIMARY KEY(nome)  
);
```

```
CREATE TABLE utente (  
  id INTEGER,  
  eta INTEGER,  
  comune VARCHAR(20),  
  CONSTRAINT utente_pk PRIMARY KEY(id),  
  CONSTRAINT comune_fk foreign key (comune) REFERENCES comune  
);
```

```
CREATE TABLE segue (  
  follower INTEGER,  
  utente INTEGER,  
  CONSTRAINT segue_pk PRIMARY KEY(follower, utente),  
  CONSTRAINT follower_fk foreign key (follower) REFERENCES utente,  
  CONSTRAINT utente_fk foreign key (utente) REFERENCES utente,  
  CONSTRAINT no_auto_followers CHECK (follower <> utente)  
);
```

dove

- la tabella "comune" memorizza nome, provincia e regione dei comuni italiani che sono di interesse nella nostra applicazione;
- la tabella "utente" memorizza codice (id), età e comune di residenza degli utenti che sono di interesse nella nostra applicazione;
- la tabella "segue" memorizza l'informazione su quale utente segue (e quindi è follower di) quale altro utente.

La base di dati è stata già popolata e sia la definizione dello schema sia i comandi di inserimento delle varie tuple nelle varie tabelle si trovano nel file "script.sql" scaricabile da Moodle nella sezione Esercitazione-5.

2. Scrittura e test delle query

Occorre scrivere e provare le query SQL corrispondenti alle seguenti esigenze informative. Evitare in ogni caso i duplicati nei risultati.

1. Per ogni utente T mostrare il suo id e l'id degli utenti che T segue e che risiedono nel Lazio.
2. Per ogni utente T mostrare il suo id e l'id degli utenti che T segue, che sono più giovani di T e che risiedono in comuni diversi da quello di T.
3. Per ogni coppia $\langle T, S \rangle$ di utenti che si seguono a vicenda e che risiedono in province diverse della stessa regione, mostrare l'id di T e di S.
4. Per ogni utente T mostrare il suo id ed il numero di "followers" di T.
5. Per ogni utente T mostrare il suo id ed il numero di "followers" di T che risiedono nello stesso comune di T e sono più giovani di T.
6. Mostrare l'id di ogni utente T che ha solo "followers" residenti nella propria regione.
7. Vogliamo sapere, per ogni regione, la quantità di comuni che mediamente si trovano nelle sue province. Quindi, per ogni regione R, dobbiamo mostrare R ed il numero medio di comuni che si trovano nelle province di R.
8. Un utente si dice "fedele" se è follower esattamente di un utente. Mostrare l'id di ogni utente che ha solo "followers" fedeli.
9. Un utente si dice "influencer in regione" se è seguito da almeno il 75% degli utenti residenti nella sua regione. Per ogni regione R, mostrare R e tutti gli influencer in regione residenti in R.
10. Per ogni utente T, mostrare il suo id, quante sono le province in cui risiedono followers di T e quante sono le regioni in cui risiedono followers di T.

Query più complessa (facoltativa):

11. Per ogni utente T, per ogni provincia P e per ogni regione R, mostrare l'id di T, la provincia P, il numero di followers di T che risiedono in P, la regione R ed il numero di followers di T che non risiedono in R.