

ESERCIZIO CINEMA P.138 LIBRO

Il database è disponibile sulla piattaforma per testare le query (file: cinema.sql)

41

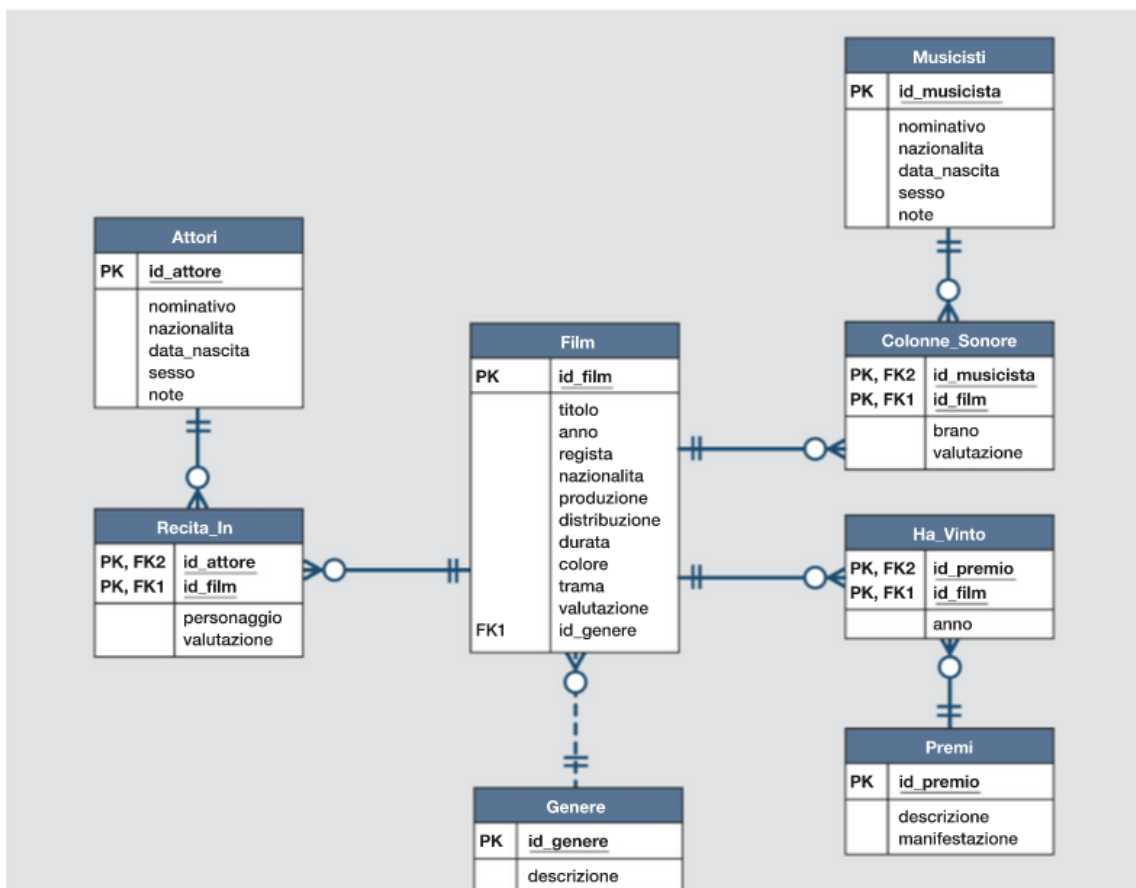


A SCUOLA DI LAVORO Dato lo schema relazionale di FIGURA 4, relativo a una banca dati cinematografica, scrivere i comandi DDL necessari per realizzarne il DB-schema e quindi formulare le seguenti query SQL:



- genere/i con il massimo numero di film specificandone il numero;
- id_attore* e nominativo degli attori che hanno recitato almeno due volte in film del genere «Horror»;
- titolo e numero di premi vinti da ogni film premiato nell'anno 2010;
- tra i vari film di genere DRAMMATICO trovare quello/i di minor durata;
- elenco in ordine di anno dei titoli dei film in cui ha recitato Samuel L Jackson
- elenco dei titoli dei film premiati la cui colonna sonora sia stata scritta da musicisti di nazionalità italiana;
- per ogni attore, il numero di film in cui ha recitato;
- per ogni nazionalità, il numero di premi Oscar vinti dal 2000 al 2011;

- titolo dei film italiani premiati nel 2009 con indicazione di premio (descrizione) e manifestazione;
- anno e titolo dei film in cui abbiano recitato Samuel L Jackson e Salma Hayek.
- il/i film/s di maggior durata;
- per ogni attrice si vuole conoscere il numero di film in cui ha recitato quando era ancora minorenne;
- id_musicista* e nominativo dei musicisti che hanno scritto brani per colonne sonore di film premiati del genere «Commedia»;
- titolo, nazionalità e numero di attori di film del 2011 il cui cast sia composto da meno di 10 attori;
- dati di attori che non hanno mai recitato in film del genere «Horror»;
- per gli anni tra 2000 e il 2010 la valutazione media annua ricevuta dagli attori che hanno recitato in film USA (elenco del tipo *anno, nominativo, valutazione_media*);
- la/le nazionalità dei film con il maggior numero di premi vinti dal 2000 al 2009 alla manifestazione degli Oscar con indicazione del numero di premi.



SOLUZIONE

QUERY

- a) il genere con il massimo numero di film specificandone il numero

Soluzione corretta, processo mentale:

- join fra film e genere per avere a disposizione la descrizione del genere di ciascun film
- raggruppo il risultato per descrizione del genere, chiamo il risultato di questa query T1
- conto il numero di righe per ciascun gruppo (chiamo questa colonna *numero_film*)
- calcolo il massimo fra i valori contati.

A questo punto il risultato di tutto quello che ho fatto è un numero, vale a dire il numero massimo di film per il genere con più film.

Con una selezione su T1 (che devo ricalcolare chiamandola T2) ottengo le righe che hanno come valore di *numero_film* il numero massimo precedentemente ottenuto.

```
select descrizione, numero_film FROM(  
    SELECT genere.descrizione, count(*) as numero_film  
    from film inner join genere using (id_genere)  
    group by genere.descrizione)  
as t1
```

```
WHERE numero_film=(select max(numero_film)  
    FROM (
```

```
    SELECT genere.descrizione, count(*) as numero_film  
    from film inner join genere using (id_genere)  
    group by genere.descrizione)  
as t2
```

```
);
```

Per verificare i risultati ho modificato il DB cambiando il genere di alcuni film in Horror.

Si poteva fare anche in altri modi, ad esempio con Having, forse più complessi:

```
select descrizione, numero_film FROM  
(  
    SELECT genere.descrizione, count(*) as numero_film  
    from film inner join genere using (id_genere)  
    group by genere.descrizione having numero_film=  
  
(
```

```

select max(numero_film) FROM
  (SELECT genere.descrizione, count(*) as numero_film
   from film inner join genere using (id_genere)
   group by genere.descrizione ) as t1
)
) as t2;

```

b) id_attore e nominativo degli attori che hanno recitato almeno 2 volte in film del genere horror.

- join fra attori, recita_in, film e anche genere
- così per ogni attore ho tutte le interpretazioni con tutti i dati
- seleziono solo le righe con genere.descrizione='horror'
- raggruppo per id attore

la query agirà sul risultato della tabella precedente selezionando id_attore e nominativo solo per i gruppi che hanno (Having) un numero di righe ≥ 2 :

```

select id_attore, nominativo
from ((attori inner join recita_in using (id_attore)) INNER join film using
(id_film))inner join genere using (id_genere)
WHERE descrizione='horror'
GROUP BY id_attore
HAVING COUNT(titolo) $\geq$ 2;

```

c) Titolo e numero di premi per ogni film premiato nell'anno 2010 (per il test aggiungo qualche premio del 2010 ai film)

```

select titolo, count(id_film) as numero_premi_2010
from film INNER JOIN ha_vinto using (id_film)
where ha_vinto.anno=2010
GROUP by id_film

```

d) Tra i vari film di genere "drammatico" trovare quello/quelli di minor durata

- seleziono tutti i film di genere cortometraggio
- determino il minimo valore di minuti
- Seleziono dalla tabella (film join genere) quelli che sono drammatici ed hanno durata pari al risultato della query precedente

```

SELECT titolo
FROM film inner join genere using (id_genere)
WHERE descrizione='drammatico' and durata =

(SELECT min(durata)
FROM film inner join genere using (id_genere)
where descrizione='drammatico');

```

OPPURE

```
SELECT titolo
FROM film INNER JOIN genere USING (id_genere)
WHERE
descrizione="drammatico" and durata<=all(SELECT DISTINCT durata FROM
film INNER JOIN genere USING (id_genere)
WHERE descrizione="drammatico")
```

- e) elenco, in ordine di anno dei titoli dei film in cui ha recitato Samuel L. Jackson

```
select titolo, anno
from (attori join recita_in USING (id_attore)) join film USING
(id_film)
WHERE nominativo='Samuel L. Jackson'
ORDER by anno;
```

- f) elenco dei titoli dei film premiati la cui colonna sonora sia stata scritta da musicisti di nazionalità italiana

- join fra film, ha_vinto, premi, colonna_sonora, musicisti
- selezione dei film premiati in cui musicisti sono italiani

```
select DISTINCT titolo
from ((film INNER join ha_vinto using (id_film)) inner join colonne_sonore
using (id_film)) inner join musicisti using (id_musicista)
where musicisti.nazionalita='italia';
```

OPPURE

```
SELECT id_film, titolo
FROM film INNER JOIN ha_vinto USING (id_film)
WHERE id_film IN (SELECT id_film
FROM ((film INNER JOIN colonne_sonore USING (id_film)) INNER JOIN
musicisti USING (id_musicista))
WHERE musicisti.nazionalita="italia")
```

- g) per ogni attore il numero di film in cui ha recitato

- join film, recita_in
- raggruppamento (GROUP BY) per attore con conteggio per id_attore

```
select id_attore, nominativo, count(id_attore)
from attori inner join recita_in using (id_attore)
group by id_attore, nominativo
```

h) per ogni nazionalità il numero dei premi oscar vinti dal 2000 al 2011

- join fra film, ha vinto, premi
- selezione (where) premi.manifestazione=oscar e ha_vinto.anno BETWEEN 2000 e 2011
- raggruppamento per film.nazionalita

```
select film.nazionalita, count(id_film)
from (film INNER join ha_vinto using (id_film)) INNER join premi using
(id_premio)
WHERE premi.manifestazione='oscar' and ha_vinto.anno BETWEEN 2000
and 2011
GROUP by film.nazionalita;
```

i) titolo dei film italiani premiati nel 2009 con indicazione del premio e della manifestazione

```
select titolo, descrizione, manifestazione
from (film INNER join ha_vinto using (id_film)) inner join premi using
(id_premio)
WHERE film.nazionalita='italia' and ha_vinto.anno=2009
```

j) Anno e titolo dei film in cui abbiano recitato Samuel L Jackson e Salma Hayek.

Soluzione 1:

- Seleziono tutti i TITOLI dei film in cui recita Samuel L. Jackson dalla tabella join fra *film – recita_in – attori* (query interna)
- Query esterna: Seleziono tutti i film in cui recita Salma Hayek dalla tabella join fra *film – recita_in – attori* mettendo come clausola WHERE che il titolo sia compreso fra i risultati della query interna

```
select titolo, anno
from (recita_in INNER join film using (id_film)) INNER join attori
using (id_attore)
where nominativo='Salma Hayek' and titolo IN
(select titolo
from (recita_in INNER join film using (id_film)) INNER join attori
using (id_attore)
where nominativo='Samuel L. Jackson');
```

Soluzione 2 (più compatta ma meno intuitiva)

- poiché la condizione da imporre nel WHERE riguarda PIU' RIGHE della stessa tabella *recita_in*, si può operare un self join su questa tabella impostando come condizione $x.id_film=y.id_film$ (x e y sono i due alias di *recita_in*). In questo modo ad ogni riga di *recita_in* sono aggiunti i campi di *recita_in* relativi allo stesso film.

- A questo punto faccio un join di questo risultato con la tabella film (sul campo id_film) per avere le informazioni sui titoli di ciascun film della tabella precedente.
- A questo punto mi mancano le informazioni sui nomi degli attori. Per associare all'id_attore della tabella x il nome dell'attore faccio un join fra il risultato precedente e la tabella attori (che chiamo z) uguagliando x.id_attore=z.id_attore. Per conoscere i nomi degli attori della tabella y ripeto quest'ultima azione facendo un join fra il risultato precedente e un nuovo alias della tabella attori (che chiamo w) uguagliando w.id_attore e y.id_attore.
- Ora faccio la selezione mettendo come clausola where che l'attore di z sia Samuel L Jackson e l'attore di w sia Salma Hayek.

```
SELECT film.titolo, film.anno
FROM (select * from attori INNER JOIN recita_in USING (id_attore))as x,
      (select * from attori INNER JOIN recita_in USING (id_attore))as y,
      film
WHERE x.id_film=y.id_film
AND y.id_film=film.id_film
AND x.nominativo="Samuel L. Jackson"
AND y.nominativo="Salma Hayek"
```

k) il/i film/s di maggior durata

- Seleziono il valore massimo di durata dalla tabella film
- seleziono dalla tabella film il film che ha la durata pari al risultato della query precedente.

```
select id_film, titolo, durata
from film
where durata=
(select max(durata) from film)
```

l) per ogni attrice si vuole conoscere il numero di film in cui ha recitato quando era ancora minorenne

```
select nominativo, count(id_film)
from (attori inner join recita_in using (id_attore)) INNER join film using
(id_film)
where attori.sesso='f' and (film.anno - year(attori.data_nascita))<=18
group by id_attore;
```

- m) id_musicista e nominativo dei musicisti che hanno scritto brani per colonne sonore di film premiati del genere commedia
- faccio un join fra tutte le tabelle che contengono le informazioni e poi seleziono le commedie

```
select id_musicista, nominativo
from (((musicisti INNER join colonne_sonore using (id_musicista)) INNER
JOIN film USING (id_film)) INNER join ha_vinto USING (id_film))
INNER join genere using (id_genere)
where descrizione='commedia';
```

- n) Titolo, nazionalità, numero di attori di film del 2011 il cui cast sia composto da meno di 10 attori

- raggruppo la tabella (film join attori) per film
- conto il numero di attori,
- having numero attori <10

```
SELECT titolo, film.nazionalita, count(id_attore) as numero_attori
FROM (film INNER join recita_in using (id_film)) INNER JOIN attori using
(id_attore)
WHERE FILM.ANNO=2011
group BY (id_film)
having numero_attori<10;
```

- o) Dati di attori di film che non hanno mai recitato in film del genere Horror

La soluzione del libro è sbagliata, restituisce elenco vuoto:

Seleziona solo le interpretazioni di film Horror

```
SELECT Attori.Id_Attoe, Attori.Nominativo
FROM Genere, Film, Attori, Recita_In
WHERE Genere.Id_Genere = Film.Id_Genere
AND Attori.Id_Attoe = Recita_In.Id_Attoe
AND Film.Id_Film = Recita_In.Id_Film
AND Genere.Descrizione='Horror'
AND Film.Id_Film NOT IN (SELECT Id_Film
FROM Genere, Film
WHERE Genere.Id_Genere = Film.Id_Genere
AND Genere.Descrizione='Horror');
```

Esclude dalla tabella precedente, le intrpretazioni dei fil horror, cioè tutte. Il risultato è una tabella vuota

Soluzione:

- query1: con una query interna seleziono gli id dei film del genere horror (dalla tabella film join genere)
- query2: Seleziono gli id_attore degli attori che recitano nei film horror con un select distinct (dalla tabella attori join recita_in join film) WHERE id_film è presente nel risultato della query precedente.

- query 3: seleziono gli attori dalla tabella attori che hanno id_attore non presente nella query2

```
SELECT * from attori
where id_attore not IN
(
  SELECT distinct attori.id_attore
  from (attori INNER JOIN recita_in using (id_attore)) inner join film using
  (id_film)
  where id_film IN
  (select id_film
  FROM film inner join genere using (id_genere)
  where descrizione='horror')
)
;
```

- p) per gli anni tra 2000 e 2016, la valutazione media annua ricevuta dagli attori che hanno recitato in film USA (anno, nominativo, valutazione_media)

- join fra attori, recita_in, film
- selezione dei soli film USA fra 2000 e 2016
- raggruppamento per id_attore e anno con calcolo della media valutazione (AVG)

```
SELECT attori.nominativo, film.anno, avg(recita_in.valutazione) as
valutazione_media
FROM attori, recita_in, film
WHERE attori.id_attore=recita_in.id_attore
AND recita_in.id_film=film.id_film
AND film.nazionalita="USA"
AND film.anno BETWEEN 2000 and 2016
GROUP BY attori.nominativo, film.anno
```

- q) La/Le nazionalità dei film con maggior numero di premi vinti dal 2000 al 2009 alla manifestazione degli Oscar con indicazione del numero di premi

Query iniziale

1. join fra film, ha_vinto_premi
2. seleziono solo i premi Oscar dal 2000 al 2009
3. raggruppo per nazionalità del film e conto i premi vinti
4. cerco il massimo fra i risultati precedenti.

tutto questo mi dà il numero X che è il massimo numero di premi vinti. Ora effettuo la ricerca sulla stessa query del punto 3 con un HAVING che verifica che il numero di premi vinti sia X

```
SELECT film.nazionalita, COUNT(premi.id_premio) as oscar_vinti
FROM film, ha_vinto,premi
WHERE film.id_film=ha_vinto.id_film
AND ha_vinto.id_premio=premi.id_premio
AND film.anno BETWEEN 2000 and 2009
AND premi.manifestazione="oscar"
GROUP BY film.nazionalita
HAVING
oscar_vinti=(SELECT max(oscar_vinti) FROM
(SELECT film.nazionalita, COUNT(premi.id_premio) as oscar_vinti
FROM film, ha_vinto,premi
WHERE film.id_film=ha_vinto.id_film
AND ha_vinto.id_premio=premi.id_premio
AND film.anno BETWEEN 2000 and 2009
AND premi.manifestazione="oscar"
GROUP BY film.nazionalita) As T)
```