

**2 Databases (tgg22)**

This question involves the relational movie database used in our SQL practical.

In each part you are given the result of an SQL query together with a possibly incorrect conclusion drawn from this result.

In each case your task is to argue for or against the conclusion. You must clearly justify your reasoning. If the SQL query can be corrected, then do so.

(a) The following query returns 1422.

Conclusion: Our database contains information on 1422 directors.

```
select count(*)
from has_position
where position = 'director';
```

[6 marks]

(b) The following query returns these records:

PERSON_ID	NAME	POSITION	TOTAL
-----	-----	-----	-----
nm0498278	Stan Lee	writer	15

Conclusion: Stan Lee did not produce any of the movies in our database.

```
select person_id, name, position, count(*) as total
from has_position as hp
join people as p on p.person_id = hp.person_id
where position <> 'actor'
and name = 'Stan Lee'
group by person_id, name, position
```

[6 marks]

[continued ...]

- (c) The following query attempts to return records (`role`, `year`, `total`) where Jennifer Lawrence plays the same `role` during the `year` a `total` number of times in different movies. The query returns these records:

ROLE	YEAR	TOTAL
-----	----	-----
Tiffany	2012	1
Mystique	2011	1
Raven	2011	1
Aurora Lane	2016	1
Katniss Everdeen	2012	2
Ree	2010	1
Rosalyn Rosenfeld	2013	1
Katniss Everdeen	2013	2

Conclusion: Jennifer Lawrence played Katniss Everdeen in two movies in 2012.

```
select r1.role, m1.year, count(*) as total
from plays_role as r1
join plays_role as r2 on r2.person_id = r1.person_id
join movies as m1 on m1.movie_id = r1.movie_id
join movies as m2 on m2.movie_id = r2.movie_id
join people as p on p.person_id = r1.person_id
where p.name = 'Jennifer Lawrence'
      and r1.role = r2.role
group by r1.role, m1.year;
```

[8 marks]