

QUERIES

1. Can you tell me which platform has the most movies represented?

- Included in the email:

```
SELECT COUNT(netflix) AS netflix_count FROM film_info WHERE
netflix = 1;
SELECT COUNT(prime) AS prime_count FROM film_info WHERE prime=
1;
SELECT COUNT(hulu) AS hulu_count FROM film_info WHERE hulu= 1;
SELECT COUNT(disney_plus) AS disney_count FROM film_info WHERE
disney_plus= 1;
```

- Not included in the email:

```
SELECT COUNT(netflix) AS netflix_and_prime FROM film_info WHERE
netflix = 1 AND prime=1;
SELECT COUNT(netflix) AS netflix_and_hulu FROM film_info WHERE
netflix = 1 AND hulu=1;
SELECT COUNT(netflix) AS netflix_and_disney FROM film_info WHERE
netflix = 1 AND disney_plus=1;
SELECT COUNT(prime) AS prime_and_hulu FROM film_info WHERE prime
= 1 AND hulu=1;
SELECT COUNT(prime) AS prime_and_disney FROM film_info WHERE
prime = 1 AND disney_plus=1;
SELECT COUNT(hulu) AS hulu_and_disney FROM film_info WHERE hulu=
1 AND disney_plus=1;
SELECT COUNT(netflix) AS netflix_prime_hulu FROM film_info WHERE
netflix= 1 AND prime=1 AND hulu=1;
SELECT COUNT(netflix) AS netflix_prime_disney FROM film_info
WHERE netflix= 1 AND prime=1 AND disney_plus=1;
SELECT COUNT(netflix) AS netflix_hulu_disney FROM film_info
WHERE netflix= 1 AND hulu=1 AND disney_plus=1;
SELECT COUNT(prime) AS prime_hulu_disney FROM film_info WHERE
prime= 1 AND hulu=1 AND disney_plus=1;
SELECT COUNT(netflix) AS all_platforms FROM film_info WHERE
netflix=1 AND prime= 1 AND hulu=1 AND disney_plus=1;
```

(Sum of above returned counts combined adds up to 40, meaning that 40 films overlap)

2. Can you tell us which platform we should partner with to provide the most enjoyable content to our customers?

- Rotten Tomato

```
SELECT ROUND(AVG(rotten_tomato), 1) AS netflix_rt_rating FROM
film_info WHERE netflix = 1;
SELECT ROUND(AVG(rotten_tomato), 1) AS prime_rt_rating FROM
film_info WHERE prime = 1;
SELECT ROUND(AVG(rotten_tomato), 1) AS hulu_rt_rating FROM
film_info WHERE hulu = 1;
```

```
SELECT ROUND(AVG(rotten_tomato), 1) AS disney_rt_rating FROM
film_info WHERE disney_plus = 1;
```

- IMDb

```
SELECT ROUND(AVG(imdb), 2) AS netflix_imdb_rating FROM film_info
WHERE netflix = 1;
SELECT ROUND(AVG(imdb), 2) AS prime_imdb_rating FROM film_info
WHERE prime = 1;
SELECT ROUND(AVG(imdb), 2) AS hulu_imdb_rating FROM film_info
WHERE hulu = 1;
SELECT ROUND(AVG(imdb), 2) AS disney_imdb_rating FROM film_info
WHERE disney_plus = 1;
```

- Can you find movies from countries that are not well represented in the database, let's say if the country only appears once, and the platform those movies are on?

```
SELECT c.name, fi.netflix, fi.prime, fi.hulu, fi.disney_plus
FROM film_info fi JOIN country c on c.id = fi.country_id WHERE
fi.country_id IN (SELECT country_id FROM film_info fi GROUP BY
fi.country_id HAVING COUNT(fi.country_id)=1);
```

- Which country has the most representation in the database for movies for adult audiences and which has the most for young children?

- Young Audiences:

```
SELECT count(1) AS count_young_audiences, c.name FROM country c
JOIN film_info fi on c.id = fi.country_id JOIN age a on a.id =
fi.age_id WHERE a.name IN ('all', '7+') GROUP BY c.name ORDER BY
count_young_audiences DESC;
```

- Adult Audiences:

```
SELECT count(1) AS count_adult_audiences, c.name FROM country c
JOIN film_info fi on c.id = fi.country_id JOIN age a on a.id =
fi.age_id WHERE a.name IN ('16+', '18+') GROUP BY c.name ORDER
BY count_adult_audiences DESC;
```

- ADDITIONAL: thirteen and older audiences (PG13 in US)

```
SELECT count(1) AS count_pg13_audiences, c.name FROM country c
JOIN film_info fi on c.id = fi.country_id JOIN age a on a.id =
fi.age_id WHERE a.name IN ('13+') GROUP BY c.name ORDER BY
count_pg13_audiences DESC;
```

- Can you tell me the most and least popular genres in the data set? Please provide two tables with the movie titles, movie languages, and genre. Choose whichever rating system. By the way, order it by year but for do not provide the

years. My team wants to guess the date of release of the movies for our team bonding activity. We may ask you for the years later.

- Most popular genre:

```
SELECT fi.name AS title, l.name AS language, g.name AS genre
FROM film_info fi
      JOIN genre g on g.id = fi.genre_id JOIN link_lang_film llf
ON fi.id = llf.film_id JOIN lang l ON llf.lang_id = l.id
      INNER JOIN (SELECT AVG(fi.imdb), g.id FROM genre g JOIN
film_info fi on g.id = fi.genre_id
GROUP BY fi.genre_id ORDER BY AVG(fi.imdb) DESC LIMIT 1) AS g2
ON g2.id = g.id ORDER BY fi.year ASC;
```

- Least popular genre:

```
SELECT fi.name AS title, l.name AS language, g.name AS genre
FROM film_info fi
      JOIN genre g on g.id = fi.genre_id JOIN link_lang_film llf
ON fi.id = llf.film_id JOIN lang l ON llf.lang_id = l.id
      INNER JOIN (SELECT AVG(fi.imdb), g.id FROM genre g JOIN
film_info fi on g.id = fi.genre_id
GROUP BY fi.genre_id ORDER BY AVG(fi.imdb) ASC LIMIT 1) AS g2 ON
g2.id = g.id ORDER BY fi.year ASC;
```

6. We would like to publicly release our partnership through a marketing campaign on a holiday using features from the movie or a movie from Steven Spielberg, since his work is very famous. We want to be sure the movie is holiday related, so if the title contains Halloween, Christmas, or Hanukkah then we would like to have the title, platform availability, and the director's name.

```
SELECT fi.name AS title, d.fname AS director_firstname, d.lname
AS director_lastname, fi.netflix, fi.prime, fi.hulu,
fi.disney_plus FROM director d JOIN film_info fi on d.id =
fi.director_id WHERE (d.fname = 'Steven' AND d.lname =
'Spielberg')
OR LOWER(fi.name) LIKE '%halloween%' OR LOWER(fi.name) LIKE
'%christmas%' OR LOWER(fi.name) LIKE '%hanukka%' OR
LOWER(fi.name) LIKE '%hannuka%' OR LOWER(fi.name) LIKE
'%hanuka%';
```

7. Which subscription provider has the most family movies?

(since I did not provide a chart in the response, I will provide screenshots here.)

```
SELECT COUNT(1) AS disney_count FROM film_info fi JOIN genre g
on g.id = fi.genre_id WHERE g.name = 'Family' GROUP BY
fi.disney_plus HAVING fi.disney_plus=1;
```

disney_count ▾	
1	7

```
SELECT COUNT(1) AS hulu_count FROM film_info fi JOIN genre g on
g.id = fi.genre_id WHERE g.name = 'Family' GROUP BY fi.hulu
HAVING fi.hulu=1;
```

hulu_count ▾	
1	7

```
SELECT COUNT(1) AS prime_count FROM film_info fi JOIN genre g on
g.id = fi.genre_id WHERE g.name = 'Family' GROUP BY fi.prime
HAVING fi.prime=1;
```

prime_count ▾	
1	7

```
SELECT COUNT(1) AS netflix_count FROM film_info fi JOIN genre g
on g.id = fi.genre_id WHERE g.name = 'Family' GROUP BY
fi.netflix HAVING fi.netflix=1;
```

netflix_count ▾	
1	7

8. Are there any unusual movies where a given rating system have much higher scores than the other?

```
SELECT fi.name, ROUND(ABS(fi.rotten_tomato - fi.imdb*10),1) AS
rating_difference, fi.rotten_tomato, fi.imdb, c.name AS country,
g.name AS genre FROM film_info fi JOIN country c ON c.id =
fi.country_id JOIN genre g ON fi.genre_id = g.id WHERE
ROUND(ABS(fi.rotten_tomato - fi.imdb*10),1) >=40 ORDER BY
rating_difference;
```

9. My son loves thriller movies and is learning Japanese in school. So, I want you to find us a good thriller (>75% approval). I would prefer the movie is in Japanese, but if it is in English I do not mind. I cannot sit still for too long so the movie must be less than 2 hours

```
SELECT fi.name AS title, g.name AS genre, fi.rotten_tomato,
l.name AS language, fi.runtime, a.name AS age from genre g JOIN
film_info fi on g.id = fi.genre_id
JOIN link_lang_film llf ON fi.id = llf.film_id JOIN lang l ON
llf.lang_id = l.id JOIN age a on a.id = fi.age_id
WHERE fi.rotten_tomato > 75 AND g.name = 'Thriller' AND
fi.runtime < 120 AND (l.name = 'Japanese' OR l.name =
'English');
```

10. I also want you to find a terrible, foreign, horror film. Many of my colleagues enjoy watching bad horror films for the humor (e.g. Attack of the 50 Foot Cheerleader) and I want to find them one they have not seen before – from a different country.

```
SELECT fi.name AS title, fi.rotten_tomato, c.name AS country
FROM film_info fi JOIN country c on fi.country_id = c.id JOIN
genre g on fi.genre_id = g.id WHERE c.name != 'United States'
AND g.name = 'Horror' ORDER BY fi.rotten_tomato, fi.imdb LIMIT
3;
```

Code for Exporting the Data:

```
mysqldump --no-tablespaces -h xx.xxx.xx.xxx --port 13315 -u madeline-rw -p movie >
movie.sql;
```

<ENTER PASSWORD>