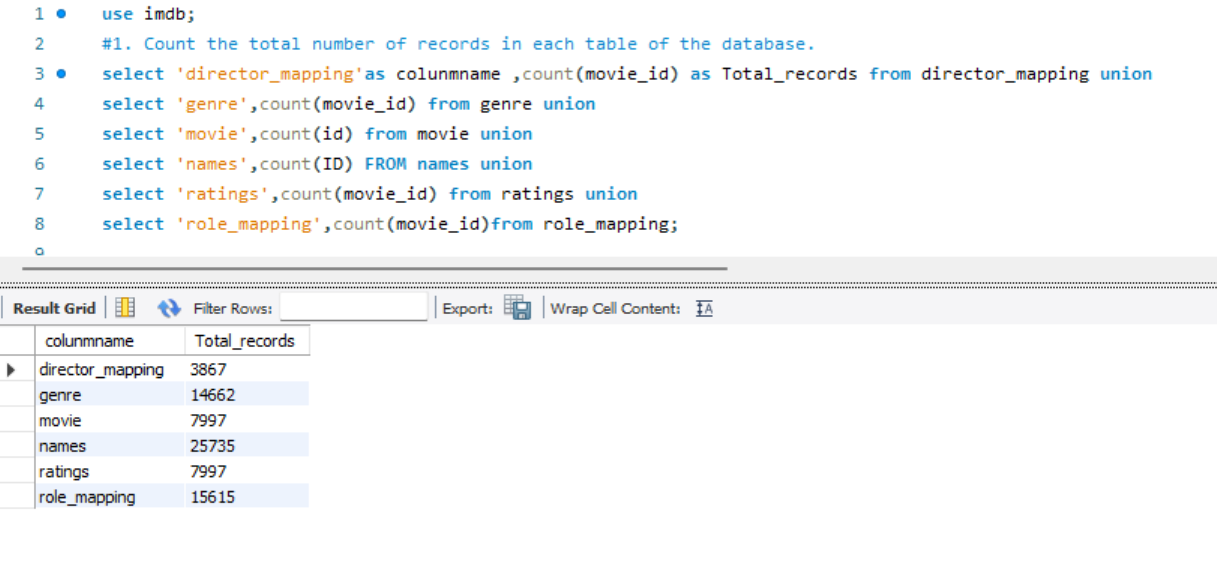
**1. Count the total number of records in each table**

**Purpose**: Determine the size of each table in the IMDB database.

* UNION to combine counts from different tables.
* For tables like director mapping, genre, etc., count the movie id or primary key.

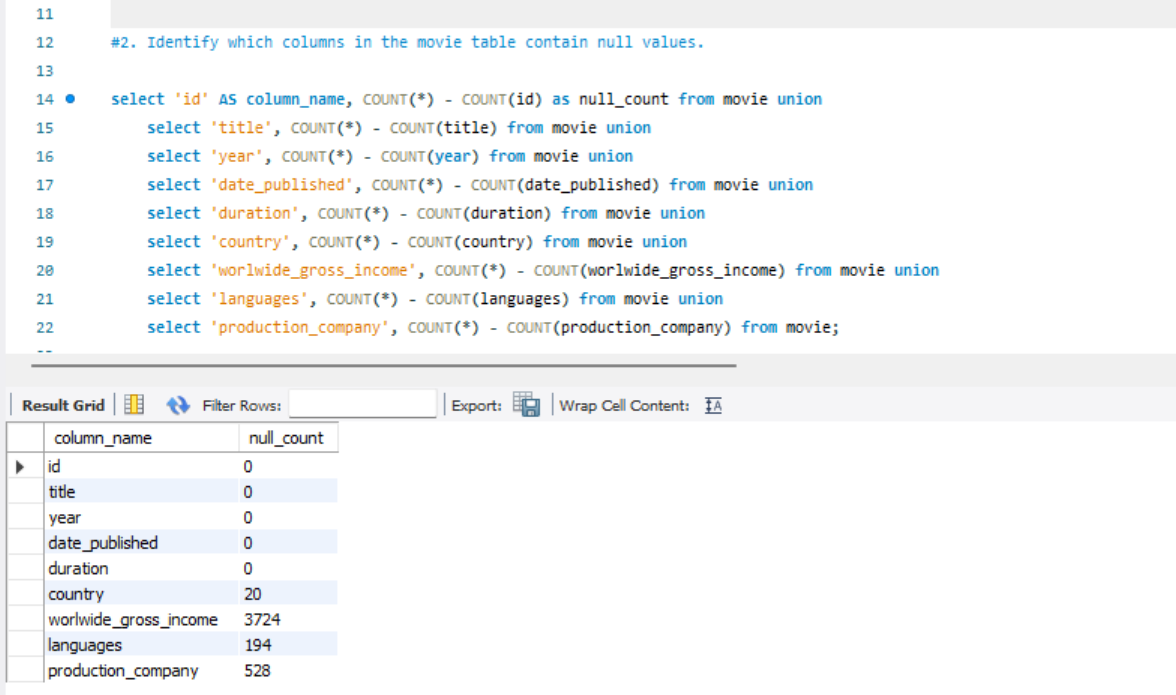


**2. Identify null values in the movie table**

**Purpose**: Check which columns have missing data.

checking for null values in tables. The method here is subtracting counts.

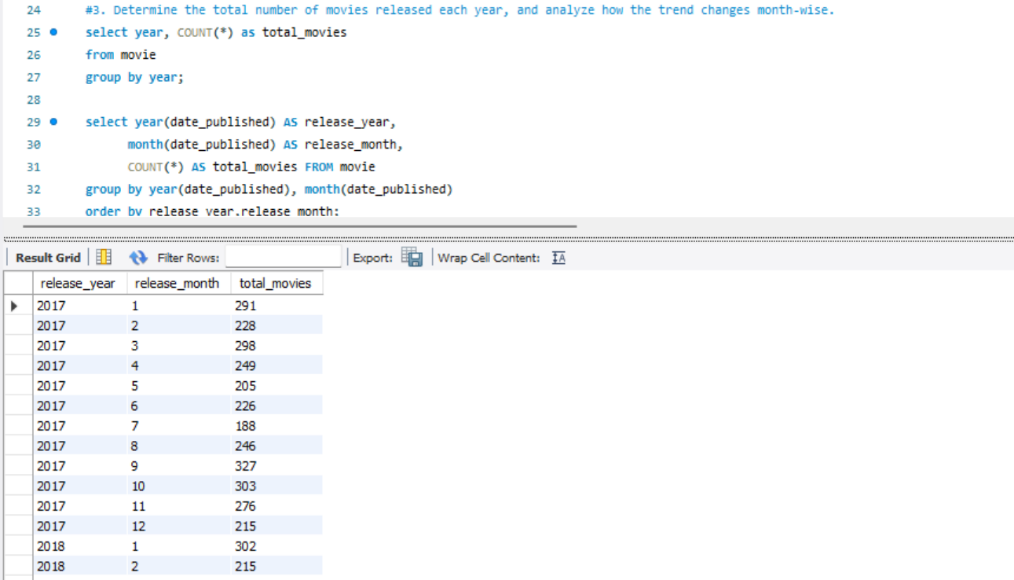
Columns like ‘worldwide gross income’ , ‘Production Company’ and ‘languages’ have higher null counts.



**3. Movies released yearly and monthly trends**

**Purpose**: Analyse movie release patterns over time.

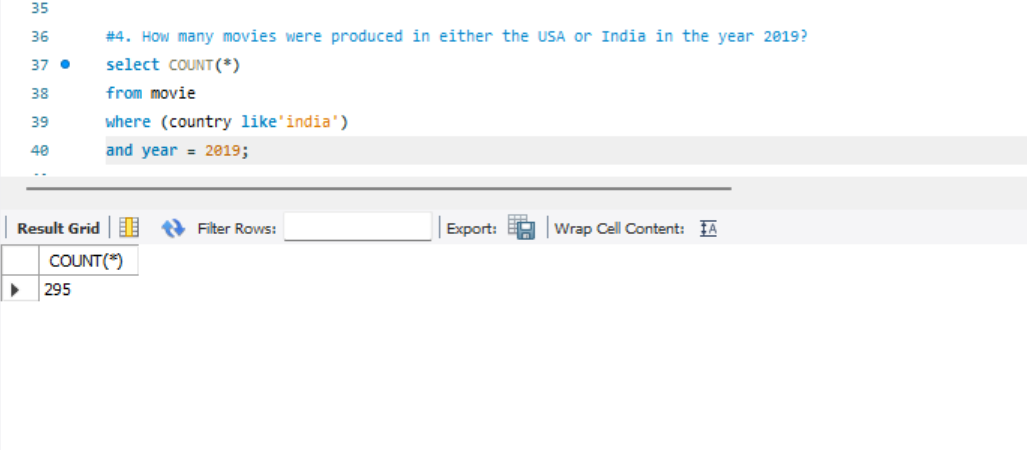
* Group by year for annual counts.
* Use MONTH (date published) for monthly trends.
* Usage of the analyse is how many movies released per month.
* Highest movies release in 2017



**4. Movies produced in the USA or India (2019)**

**Purpose**: Filter movies from specific countries in a given year.

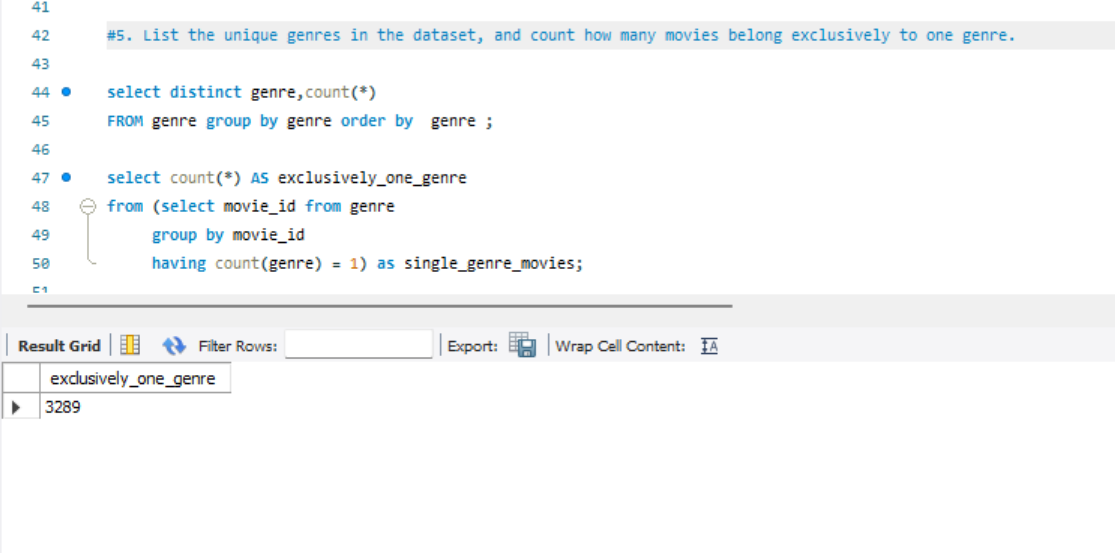
* The syntax using WHERE country LIKE ‘INDIA’ AND year = 2019.
* From this query I get a how many movies produced in India on 2019



**5. Unique genres and movies with one genre**

**Purpose**: List genres and count movies with only one genre.

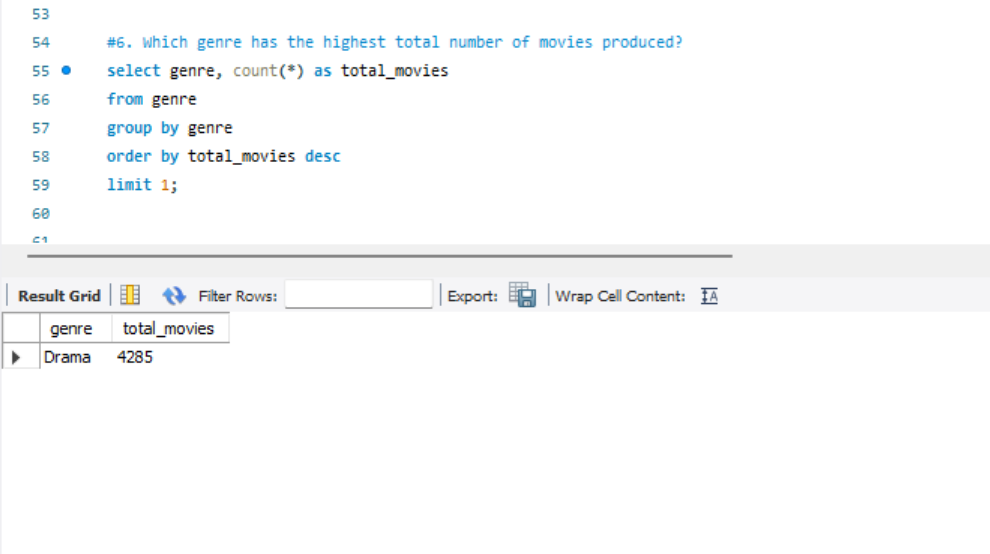
* Using DISTINCT to get a unique values from genre .
* Subquery to count movies with HAVING COUNT(genre) = 1



**6. Genre with the highest movies produced**

**Purpose**: Find the most popular genre.

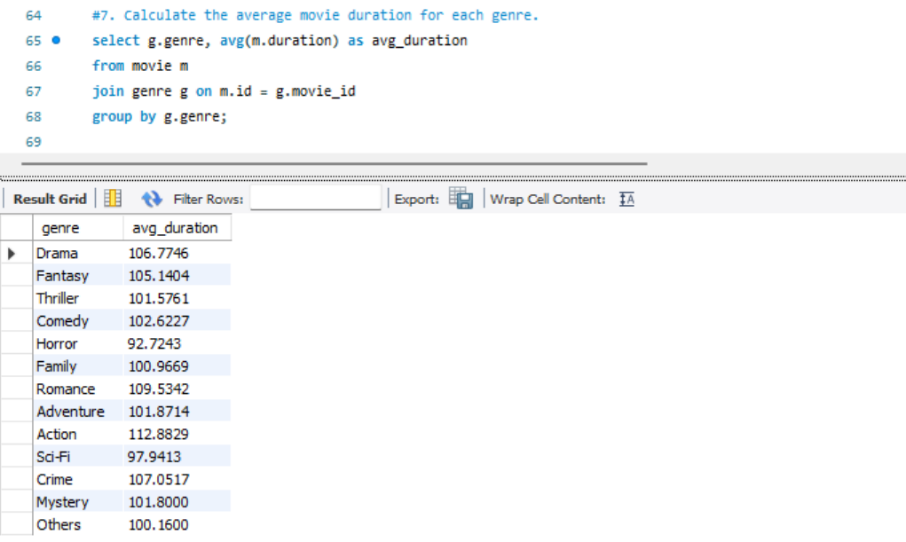
* Group by genre, order by count in descending order, and use LIMIT 1.
* Drama genre has the highest total number of movies produced.



**7. Average duration per genre**

**Purpose**: Compare movie lengths across genres.

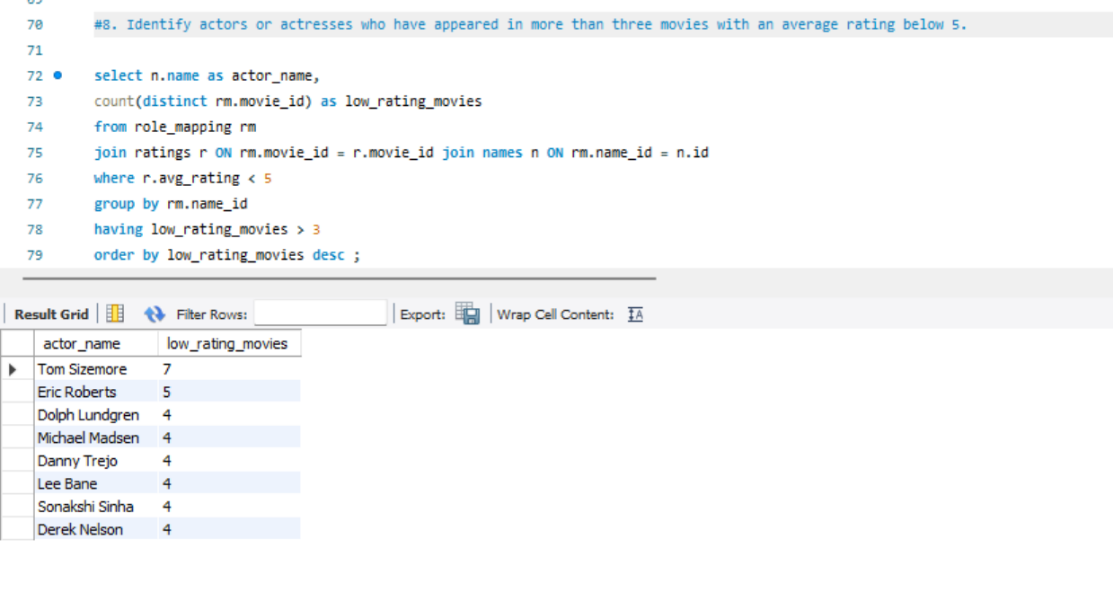
* Join movie and genre tables, then group by genre.  
  **Usage Notes**:
* Genre ‘Drama’ having longer duration.



**8. Actors in >3 movies with avg rating <5**

**Purpose**: Identify actors in poorly rated movies.

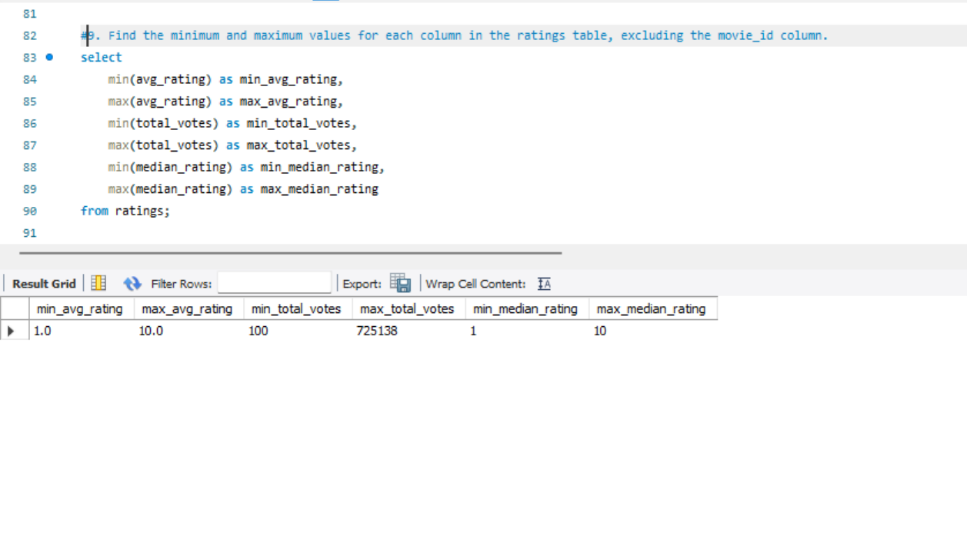
* Join role mapping, ratings, and names tables.
* Filter with HAVING COUNT(movie id) > 3 AND average rating < 5.
* From this analyse getting underperforming actors.



**9. Min/Max values in ratings table**

**Purpose**: Explore the range of ratings and votes.

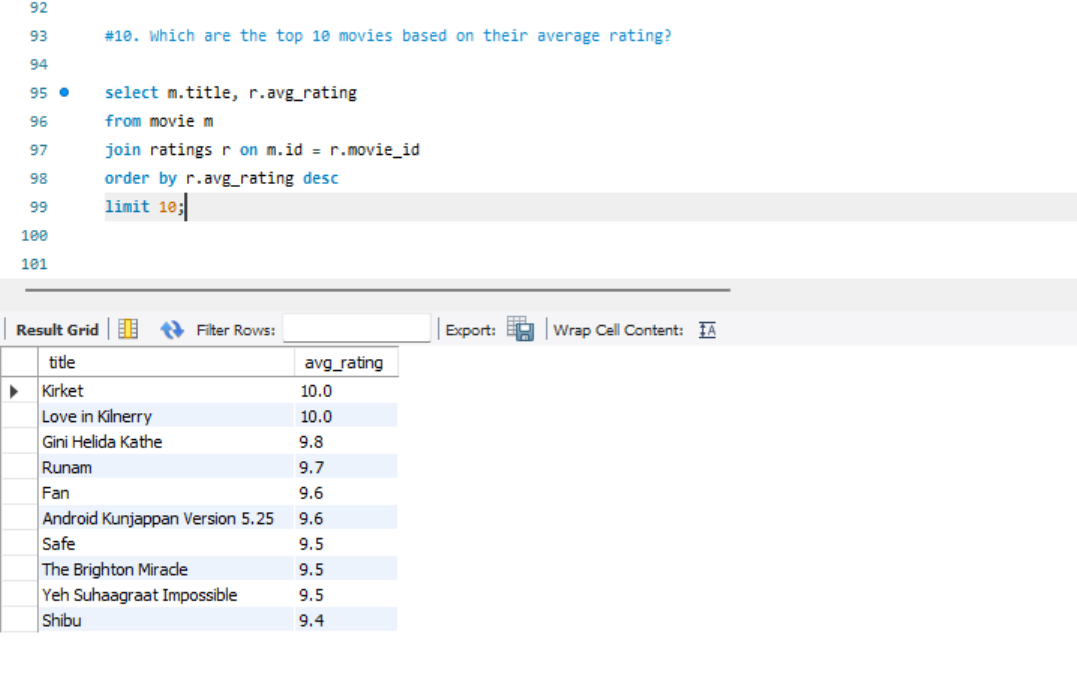
* Calculate min/max for average rating, total votes, and median rating



**10. Top 10 movies by average rating**

**Purpose**: Highlight highest-rated movies.

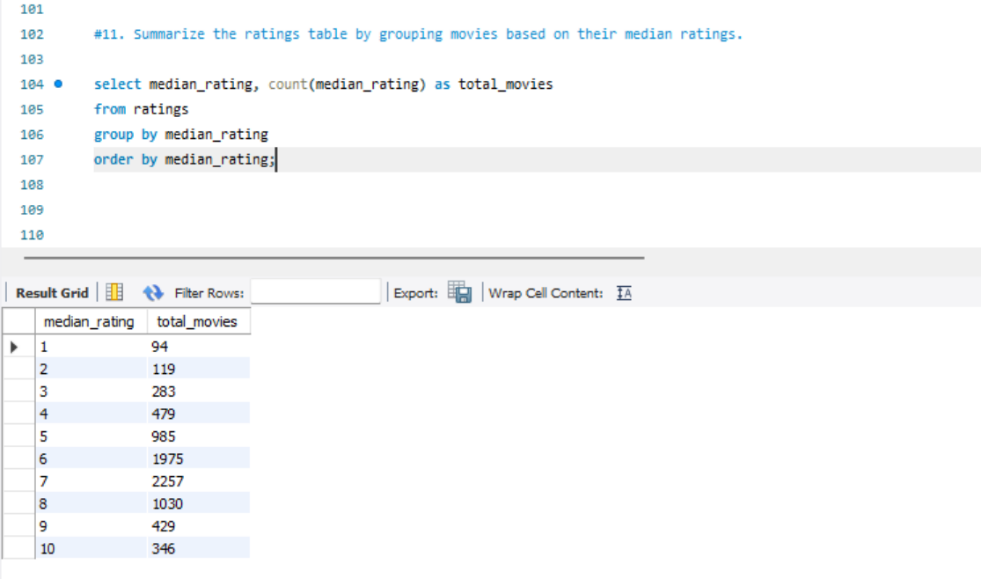
* Join movie and ratings, then order by average rating DESC with LIMIT 10.
* From this query I got a ‘Kirket’ and ‘Love in Kilnerry’ is highest average rating movies.



**11. Movies grouped by median rating**

**Purpose**: Summarize rating distribution.

* Group by median rating and count movies.
* Useful for rating analysis

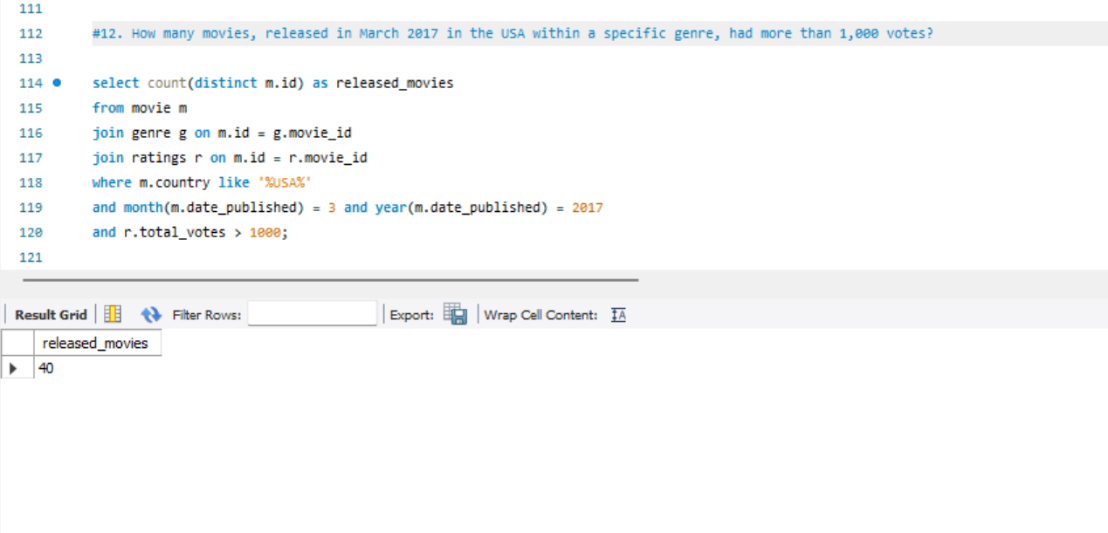


**12. March 2017 USA movies with >1,000 votes**

**Purpose**: Filter high-engagement movies.

* Join movie, genre, and ratings tables.
* Filter by total votes > 1000 and date constraints & for exact country match am using

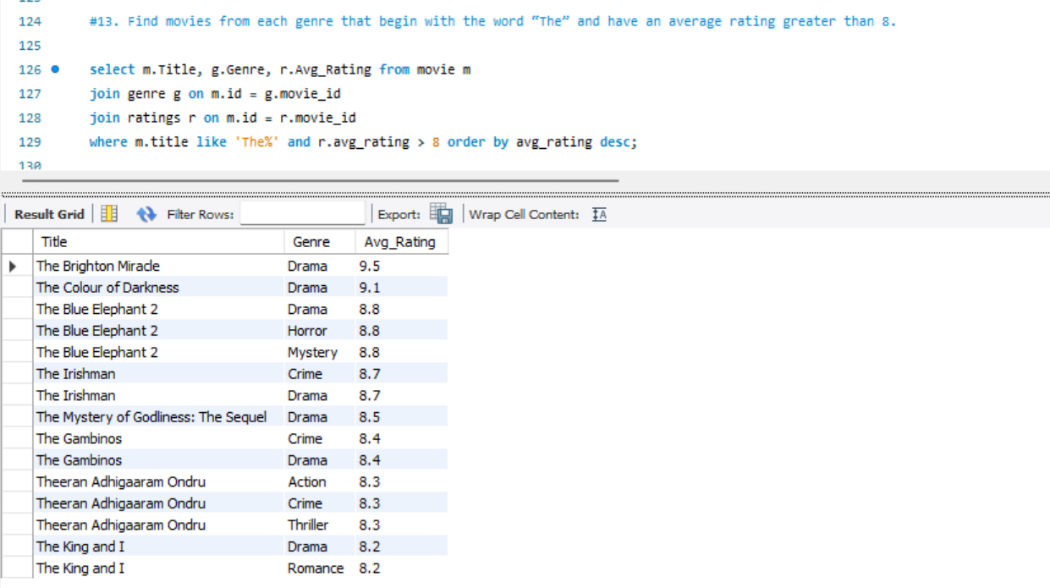
LIKE and using %USA%.



**13. High-rated movies starting with “The”**

**Purpose**: Find top-rated movies with a title pattern.

* Use Where title Like ‘The%' And average rating > 8.
* Useful for marketing analysis.



**14. Movies with median rating 8 (April 2018–2019)**

**Purpose**: Count movies meeting specific criteria.

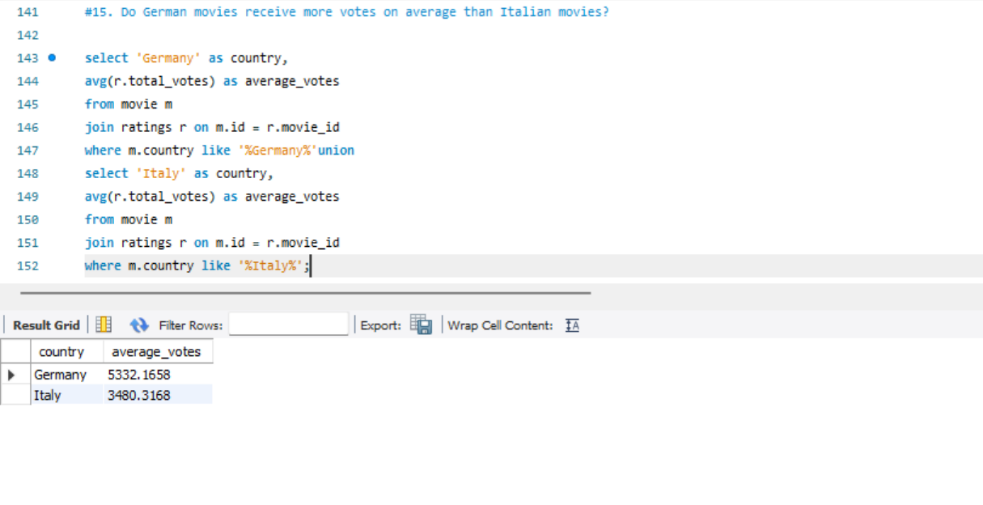
* Join movie and ratings, filter by date range and median rating = 8.
* Useful for to find between that month and year how many movies getting median rating.



15.**German vs. Italian movie votes**

**Purpose**: Compare audience engagement.

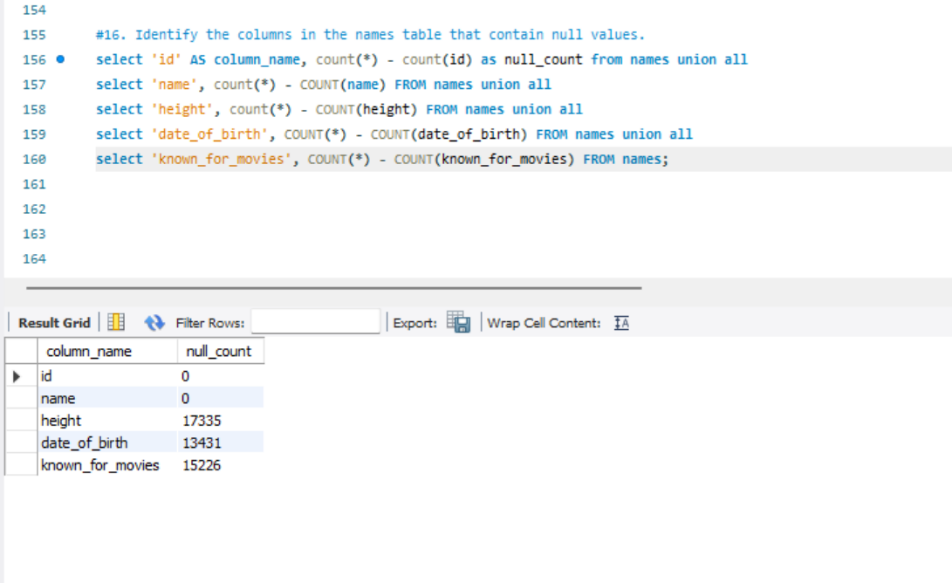
* Use CASE statements to separate German and Italian movies.  
  **Usage Notes**:
* Results may vary based on sample size.



**16. Null values in names table**

**Purpose**: Identify missing data in actor/director profiles.

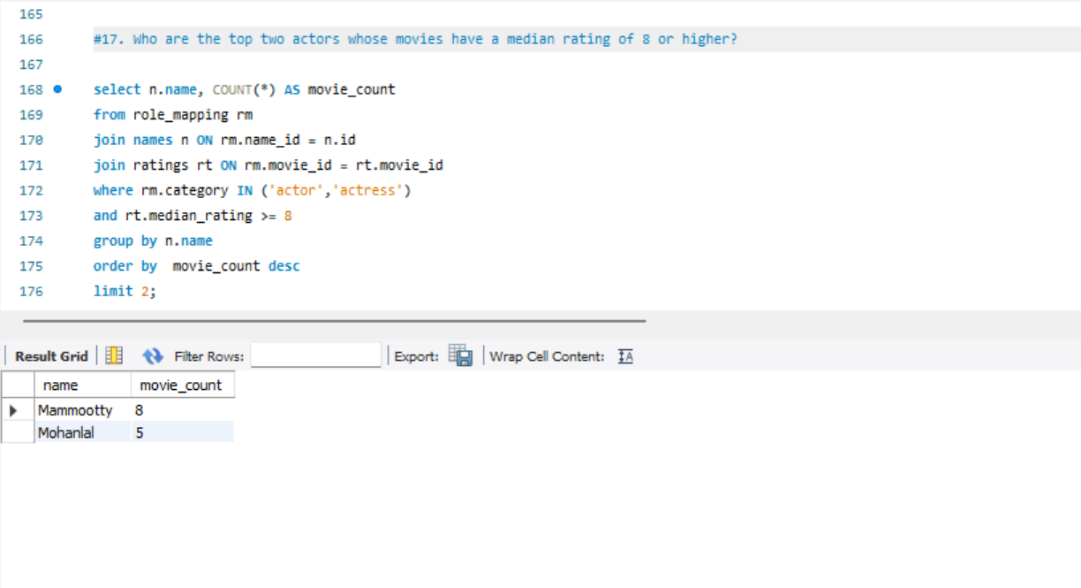
* Columns height, date of birth , Known for movies having null values.



17.**17. Top 2 actors with median rating ≥8**

**Purpose**: Recognize high-performing actors.

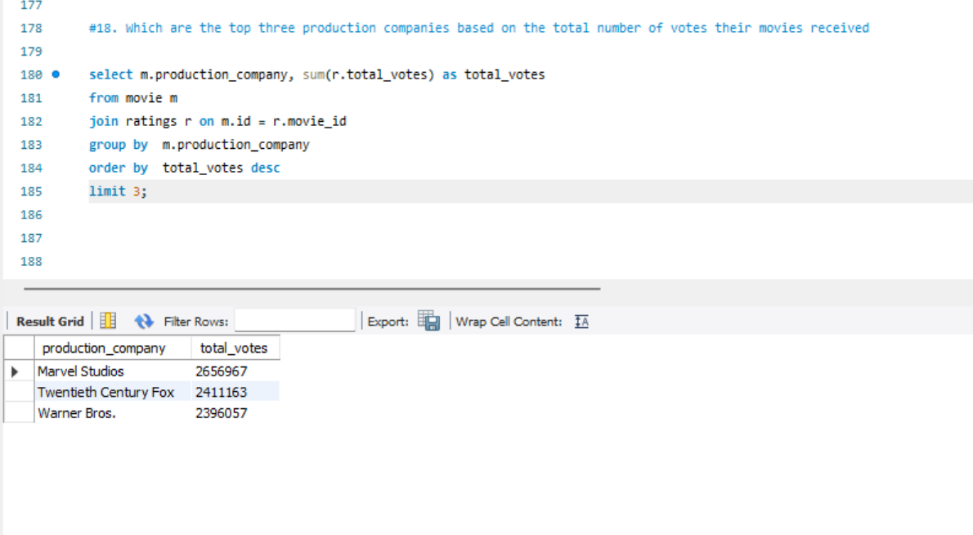
* Join role mapping, ratings, and names tables.
* Filter by category = 'actor' and median rating >= 8.
* Limits to two actors; expandable with LIMIT.



18.**18. Top 3 production companies by votes**

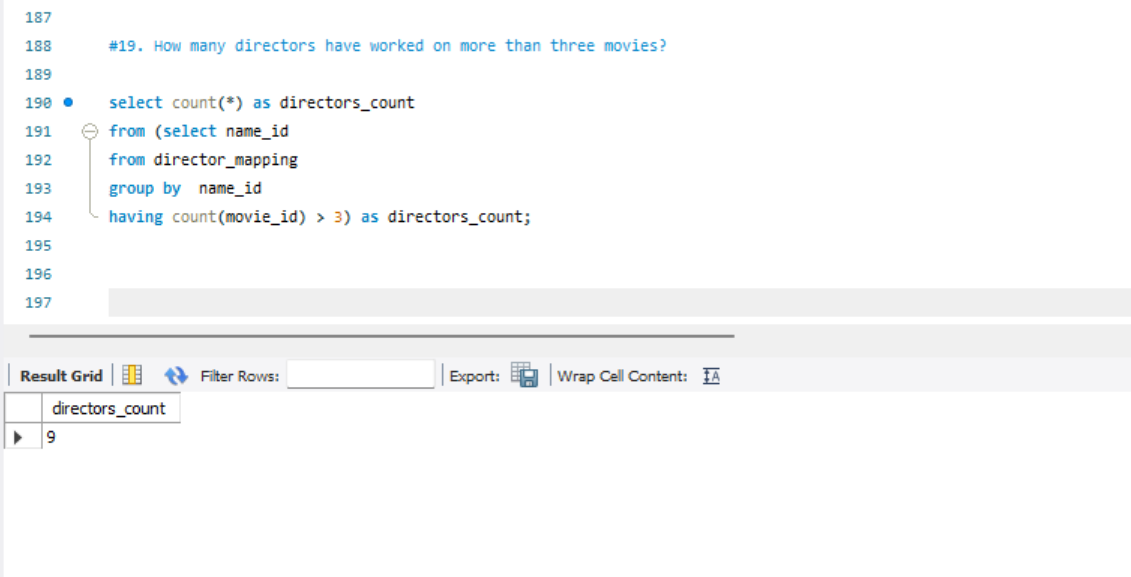
**Purpose**: Rank companies by audience engagement.

* Sum total votes per company and order descending.
* Reflects popularity, not necessarily quality.



19. **Purpose**: Identify prolific directors.

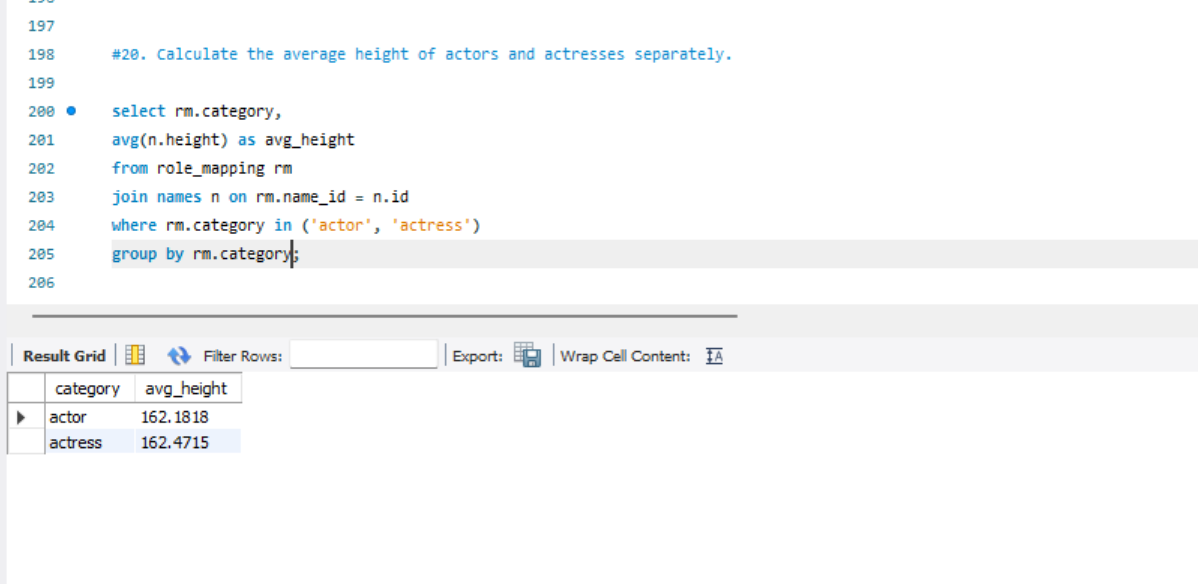
* Subquery to count movies per director.
* Using COUNT(movie id) for accuracy.



**20. Average height of actors/actresses**

**Purpose**: Compare physical attributes by gender.

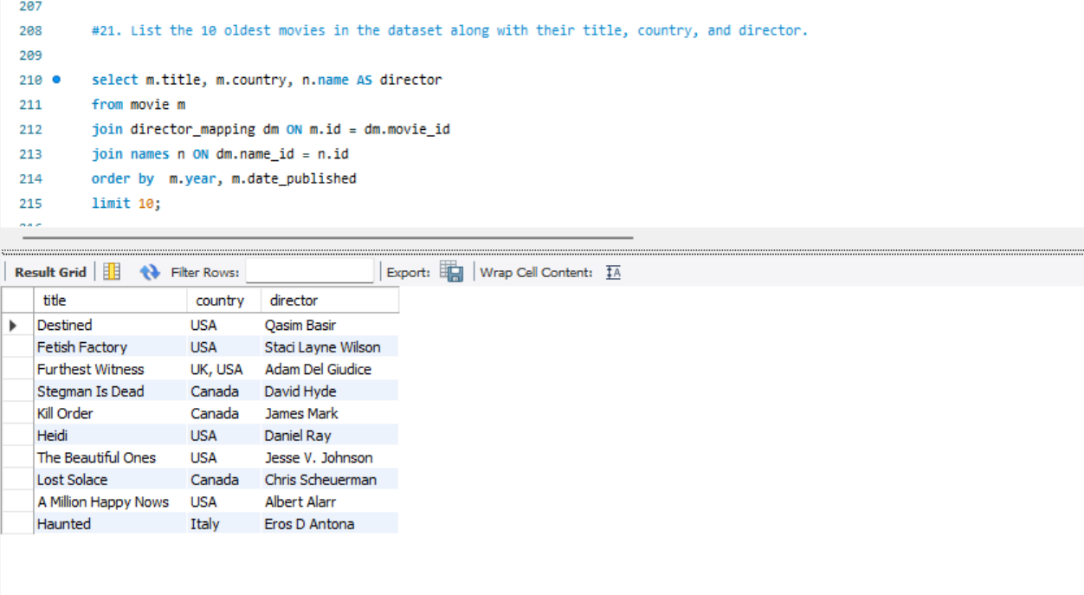
* Join role mapping and names, filter by category.



**21. 10 oldest movies with directors**

**Purpose**: List historical movies.

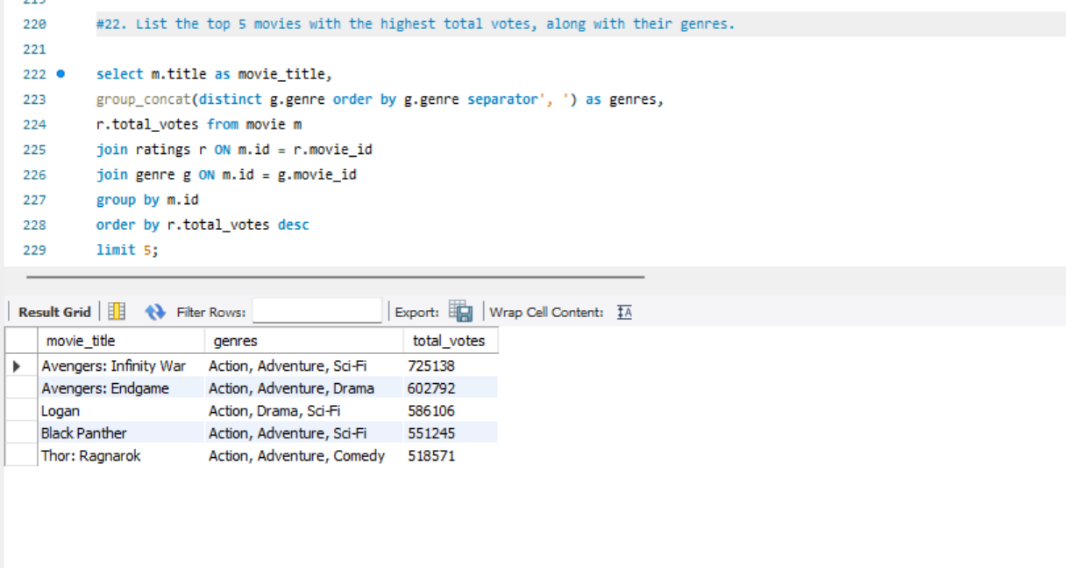
* Join movie and director\_mapping, order by year and date\_published.
* "Oldest" is based on year, not publication date.



**22. Top 5 movies by votes (with genres)**

**Purpose**: Highlight popular movies and their genres.

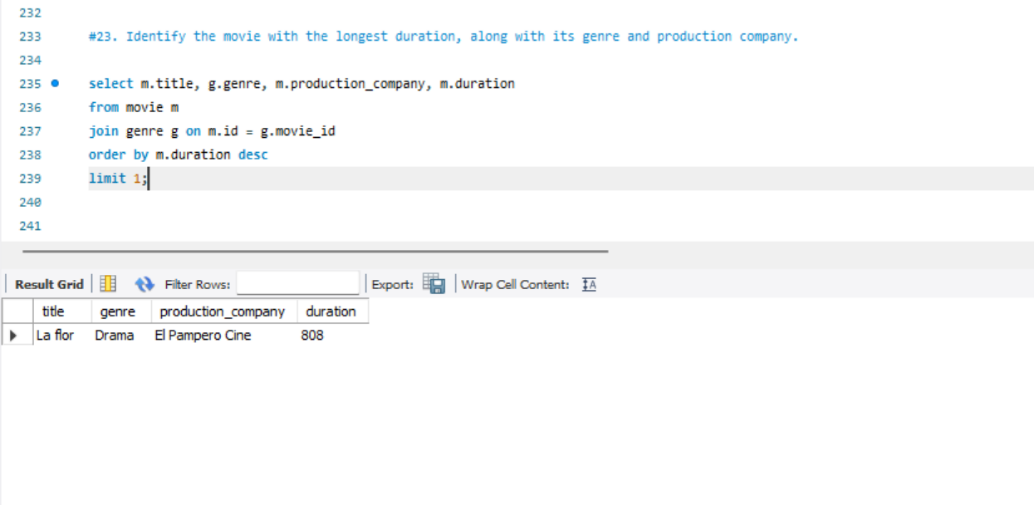
* Use GROUP\_CONCAT to combine genres per movie.
* Genres are concatenated .
* From this Avengers: Infinity War was getting highest vote by people



**23. Longest movie with genre and company**

**Purpose**: Identify the longest movie.

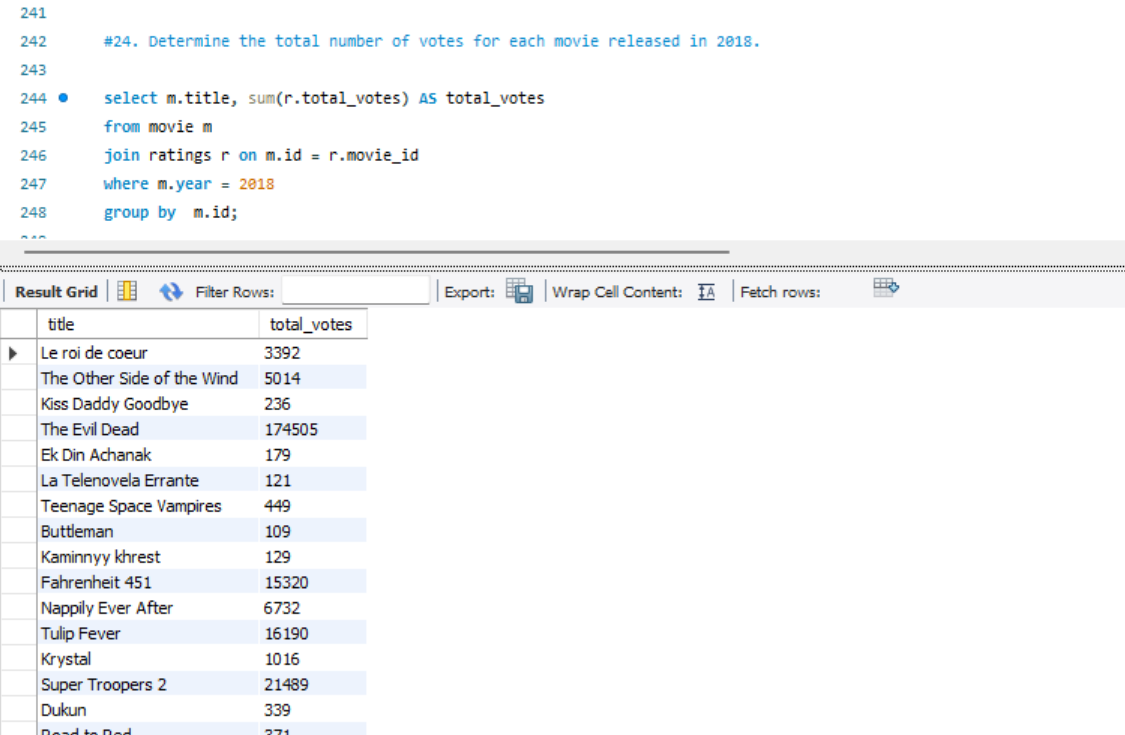
* Order by duration DESC and LIMIT 1.
* Longest Movie Name is La Flor – 808 in Drama Genre



**24. Total votes for 2018 movies**

**Purpose**: Measure audience engagement per movie.

* Join movie and ratings, group by movie.
* Use SUM(total votes) for accuracy.



25.**25. Most common movie language**

**Purpose**: Determine the dominant language.

* Group by languages and count movies.
* The Most Common Language is English

