/* So, RSVP Movies plans to make a movie of one of these genres.
Now, wouldn't you want to know which genre had the highest number of movies
produced in the last year?
Combining both the movie and genres table can give more interesting insights.
*/

-- Q6.Which genre had the highest number of movies produced overall?
-- Type your code below:

/* So, based on the insight that you just drew, RSVP Movies should focus on the
'Drama' genre.
But wait, it is too early to decide. A movie can belong to two or more genres.
So, let's find out the count of movies that belong to only one genre.*/

-- Q7. How many movies belong to only one genre?
-- Type your code below:

/* There are more than three thousand movies which has only one genre
associated with them.
So, this figure appears significant.
Now, let's find out the possible duration of RSVP Movies' next project.*/

-- Q8.What is the average duration of movies in each genre?
-- (Note: The same movie can belong to multiple genres.)

/* Output format:

+--------------+-----------------+

```
| genre              |      avg_duration  |
+------------------+---------------
|    thriller    |         105            |
|     .              |          .            |
|     .              |          .            |
+--------------+------------------+ */
-- Type your code below:
```

/* Now you know, movies of genre 'Drama' (produced highest in number in 2019) has the average duration of 106.77 mins.
Lets find where the movies of genre 'thriller' on the basis of number of movies.*/

-- Q9.What is the rank of the 'thriller' genre of movies among all the genres in terms of number of movies produced?
-- (Hint: Use the Rank function)

```
/* Output format:
+--------------+------------------+--------------------+
| genre        |      movie_count  |      genre_rank   |
+--------------+------------------+--------------------+
|drama         |    2312          |         2          |
+--------------+------------------+--------------------+*/
-- Type your code below:
```
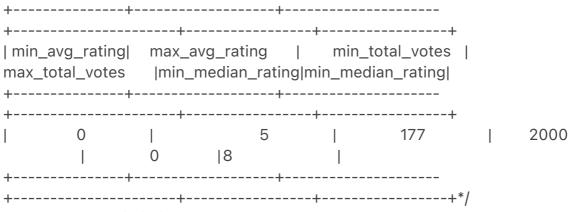
/*Thriller movies is in top 3 among all genres in terms of number of movies
 In the previous segment, you analysed the movies and genres tables.
 In this segment, you will analyse the ratings table as well.
 To start with lets get the min and max values of different columns in the table*/

-- Segment 2:

-- Q10. Find the minimum and maximum values in each column of the ratings table except the movie_id column?
/* Output format:
```
+--------------+-----------------+-------------------
+--------------------+---------------+----------------+
| min_avg_rating|   max_avg_rating    |    min_total_votes  |
max_total_votes    |min_median_rating|min_median_rating|
+--------------+-----------------+-------------------
+--------------------+---------------+----------------+
|          0        |                5        |          177          |      2000
            |        0      |8             |
+--------------+-----------------+-------------------
+--------------------+---------------+----------------+*/
```
-- Type your code below:

/* So, the minimum and maximum values in each column of the ratings table are in the expected range.
This implies there are no outliers in the table.
Now, let's find out the top 10 movies based on average rating.*/

-- Q11. Which are the top 10 movies based on average rating?
/* Output format:
```
+--------------+-----------------+-------------------+
| title          |            avg_rating    |         movie_rank   |
+--------------+-----------------+-------------------+
| Fan          |          9.6        |              5         |
|    .         |         .        |         .           |
|    .         |         .        |         .           |
|    .         |         .        |         .           |
+--------------+-----------------+-------------------+*/
```
-- Type your code below:

-- It's ok if RANK() or DENSE_RANK() is used too

/* Do you find you favourite movie FAN in the top 10 movies with an average rating of 9.6? If not, please check your code again!!
So, now that you know the top 10 movies, do you think character actors and filler actors can be from these movies?
Summarising the ratings table based on the movie counts by median rating can give an excellent insight.*/

-- Q12. Summarise the ratings table based on the movie counts by median ratings.
/* Output format:

```
+---------------+------------------+
| median_rating |    movie_count    |
+------------------+----------------
|     1         |         105       |
|     .         |          .        |
|     .         |          .        |
+--------------+------------------+ */
```
-- Type your code below:
-- Order by is good to have

/* Movies with a median rating of 7 is highest in number.
Now, let's find out the production house with which RSVP Movies can partner for its next project.*/

-- Q13. Which production house has produced the most number of hit movies (average rating > 8)??
/* Output format:
```
+------------------+-------------------+--------------------+
```

```
|production_company|movie_count        |prod_company_rank|
+----------------+-----------------+------------------+
| The Archers    |        1        |        1         |
+----------------+-----------------+------------------+*/
```
-- Type your code below:

-- It's ok if RANK() or DENSE_RANK() is used too
-- Answer can be Dream Warrior Pictures or National Theatre Live or both

-- Q14. How many movies released in each genre during March 2017 in the USA had more than 1,000 votes?
/* Output format:

```
+---------------+------------------+
| genre         |    movie_count   |
+------------------+----------------
|    thriller    |        105       |
|        .       |         .        |
|        .       |         .        |
+---------------+------------------+ */
```
-- Type your code below:

-- Lets try to analyse with a unique problem statement.
-- Q15. Find movies of each genre that start with the word 'The' and which have an average rating > 8?
/* Output format:
```
+---------------+------------------+--------------------+
| title         |    avg_rating    |      genre         |
+---------------+------------------+--------------------+
| Theeran       |        8.3       |      Thriller      |
|      .        |        .         |         .          |
```

```
|       .          |       .          |          .         |
|       .          |       .          |          .         |
+--------------+----------------+------------------+*/
```
-- Type your code below:

-- You should also try your hand at median rating and check whether the 'median rating' column gives any significant insights.
-- Q16. Of the movies released between 1 April 2018 and 1 April 2019, how many were given a median rating of 8?
-- Type your code below:

-- Once again, try to solve the problem given below.
-- Q17. Do German movies get more votes than Italian movies?
-- Hint: Here you have to find the total number of votes for both German and Italian movies.
-- Type your code below:

-- Answer is Yes

/* Now that you have analysed the movies, genres and ratings tables, let us now analyse another table, the names table.
Let's begin by searching for null values in the tables.*/

-- Segment 3:

-- Q18. Which columns in the names table have null values??
/*Hint: You can find null values for individual columns or follow below output format
+--------------+-----------------+-------------------
+---------------------+
| name_nulls | height_nulls |date_of_birth_nulls |
known_for_movies_nulls|
+--------------+-----------------+-------------------
+---------------------+
| 0 | 123 | 1234 | 12345
        |
+--------------+-----------------+-------------------
+---------------------+*/
-- Type your code below:

/* There are no Null value in the column 'name'.
The director is the most important person in a movie crew.
Let's find out the top three directors in the top three genres who can be hired
by RSVP Movies.*/

-- Q19. Who are the top three directors in the top three genres whose movies
have an average rating > 8?
-- (Hint: The top three genres would have the most number of movies with an
average rating > 8.)
/* Output format:

+--------------+-----------------+
| director_name | movie_count |
+--------------+-----------------|
James Mangold	4
.	.
.	.
+--------------+-----------------+ */

-- Type your code below:

/* James Mangold can be hired as the director for RSVP's next project. Do you remeber his movies, 'Logan' and 'The Wolverine'.
Now, let's find out the top two actors.*/

-- Q20. Who are the top two actors whose movies have a median rating >= 8?
/* Output format:

```
+--------------+------------------+
| actor_name   |    movie_count   |
+------------------+----------------
|Christain Bale |          10             |
|     .               |         .             |
+--------------+------------------+ */
```
-- Type your code below:

/* Have you find your favourite actor 'Mohanlal' in the list. If no, please check your code again.
RSVP Movies plans to partner with other global production houses.
Let's find out the top three production houses in the world.*/

-- Q21. Which are the top three production houses based on the number of votes received by their movies?
/* Output format:

```
+-----------------+------------------+-------------------+
|production_company|vote_count        |        prod_comp_rank|
+-----------------+------------------+-------------------+
| The Archers      |        830        |        1             |
|     .            |         .         |              .       |
|     .            |         .         |              .       |
+-----------------+------------------+-------------------+*/
```

-- Type your code below:

-- Q22. Rank actors with movies released in India based on their average
ratings. Which actor is at the top of the list?
-- Note: The actor should have acted in at least five Indian movies.
-- (Hint: You should use the weighted average based on votes. If the ratings
clash, then the total number of votes should act as the tie breaker.)

/* Output format:

actor_name	total_votes	movie_count	actor_avg_rating	actor_rank
Yogi Babu	3455	11	8.42	1
.	.	.	.	
.				
.	.	.	.	
.				
.	.	.	.	
.				

*/
-- Type your code below:

-- Top actor is Vijay Sethupathi

-- Q23.Find out the top five actresses in Hindi movies released in India based on their average ratings?
-- Note: The actresses should have acted in at least three Indian movies.
-- (Hint: You should use the weighted average based on votes. If the ratings clash, then the total number of votes should act as the tie breaker.)
/* Output format:

actress_name	total_votes	movie_count	actress_avg_rating	actress_rank
Tabu	3455	11	8.42	1
.	.	.	.	
.				
.	.	.	.	
.				
.	.	.	.	
.				

*/
-- Type your code below:

/* Taapsee Pannu tops with average rating 7.74.
Now let us divide all the thriller movies in the following categories and find out their numbers.*/

/* Q24. Select thriller movies as per avg rating and classify them in the following category:

    Rating > 8: Superhit movies
    Rating between 7 and 8: Hit movies
    Rating between 5 and 7: One-time-watch movies
    Rating < 5: Flop movies
---*/
-- Type your code below:

/* Until now, you have analysed various tables of the data set.
Now, you will perform some tasks that will give you a broader understanding of the data in this segment.*/

-- Segment 4:

-- Q25. What is the genre-wise running total and moving average of the average movie duration?
-- (Note: You need to show the output table in the question.)
/* Output format:
+---------------+------------------+-------------------+---------------------+
| genre | avg_duration |running_total_duration| moving_avg_duration |
+---------------+------------------+-------------------+---------------------+
comdy	145	106.2	128.42
.	.	.	.
.	.	.	.
.	.	.	.
+---------------+------------------+-------------------+---------------------+*/
-- Type your code below:

-- Round is good to have and not a must have; Same thing applies to sorting

-- Let us find top 5 movies of each year with top 3 genres.

-- Q26. Which are the five highest-grossing movies of each year that belong to the top three genres?
-- (Note: The top 3 genres would have the most number of movies.)

/* Output format:
+--------------+-----------------+-------------------+--------------------+----------------+
| genre | year | movie_name | worldwide_gross_income|movie_rank |
+--------------+-----------------+-------------------+--------------------+----------------+
comedy	2017	indian	$103244842	1
.	.	.	.	.
.	.			
.	.	.	.	.
.	.			
.	.	.	.	.
.	.			
+--------------+-----------------+-------------------+--------------------+----------------+*/
-- Type your code below:

-- Top 3 Genres based on most number of movies

-- Finally, let's find out the names of the top two production houses that have produced the highest number of hits among multilingual movies.
-- Q27. Which are the top two production houses that have produced the highest number of hits (median rating >= 8) among multilingual movies?
/* Output format:

```
+------------------+------------------+--------------------+
|production_company |movie_count       |        prod_comp_rank|
+------------------+------------------+--------------------+
| The Archers      |       830        |        1           |
|      .           |        .         |           .        |   |
|      .           |        .         |           .        |   |
+------------------+------------------+--------------------+*/
```

-- Type your code below:

-- Multilingual is the important piece in the above question. It was created using POSITION(',' IN languages)>0 logic
-- If there is a comma, that means the movie is of more than one language

-- Q28. Who are the top 3 actresses based on number of Super Hit movies (average rating >8) in drama genre?
/* Output format:

```
+--------------+------------------+-------------------- +---------------------+----------------+
| actress_name    |     total_votes      |      movie_count         | actress_avg_rating  |actress_rank    |
+--------------+------------------+-------------------- +---------------------+----------------+
|     Laura Dern   |              1016|        1          |      9.60        | 1        |
|        .         |        .         |     .    |    .
|        .         |
|        .         |        .         |     .    |    .
|        .         |
+--------------+------------------+-------------------- +---------------------+----------------+*/
```

-- Type your code below:

/* Q29. Get the following details for top 9 directors (based on number of movies)
Director id
Name
Number of movies
Average inter movie duration in days
Average movie ratings
Total votes
Min rating
Max rating
total movie durations

Format:

director_id	director_name	number_of_movies	avg_inter_movie_days	avg_rating	total_votes	min_rating	max_rating	total_duration
nm1777967	A.L. Vijay	5	5.65	1754	3.7	6.9	613	
.		.		.		.		.
	.	.	.	.	.	.	.	
.		.		.		.		.
	.	.	.	.	.	.	.	
.		.		.		.		.
	.	.	.	.	.	.	.	
.		.		.		.		.
	.	.	.	.	.	.	.	
.		.		.		.		.
	.	.	.	.	.	.	.	

```
           |
|      .              |          .              |               .            |          .
      |        .    |  .          |  .        |  .          |           .
           |
|      .              |          .              |               .            |          .
      |        .    |  .          |  .        |  .          |           .
           |
|      .              |          .              |               .            |          .
      |        .    |  .          |  .        |  .          |           .
           |
+--------------+-----------------+-------------------
+--------------------+-------------+-------------+------------
+-----------+---------------+


-----------------------------------------------------------------------------
-------------*/
-- Type you code below:
```