Executive Summary

Problem Statement

RSVP Movies is an Indian film production company which has produced many super-hit movies. They have usually released movies for the Indian audience but for their next project, they are planning to release a movie for the global audience in 2022.

The production company wants to plan their every move analytically based on data and have approached you for help with this new project. You have been provided with the data of the movies that have been released in the past three years. You have to analyse the data set and draw meaningful insights that can help them start their new project.

You are a data analyst and an SQL expert. You have to use SQL to analyse the given data and give recommendations to RSVP Movies based on the insights. For your convenience, the entire analytics process has been divided into four segments, where each segment leads to significant insights from different combinations of tables. The questions in each segment with business objectives are written in the script given below. You have to write the solution code below every question and submit the same SQL script file with the solution in the 'Submission' segment.

This summary describes our analysis and mentions suggestions to RSVP based on the questions given in the assignment

Questions	Results
Q1	Total no. of rows 1. Director_mapping: 3867 2. Genre: 14662 3. Movie: 7997 4. Names: 25735 5. Ratings: 7997 6. Role_mapping: 15615
Q2	Movie table: columns having null values are Country (20), Worlwide_gross_income(3724), Languages (194), and Production_Company (528)
Q3	1. Total number of movies released in the year 2017, 2018 and 2019 respectively are 3052, 2944 and 2001. 2. Month-wise, number of movies released every year in every month: January 804 February 640 March 824 April 680 May 625 June 580 July 493 August 678 September 809 October 801 November 625 December 438
Q4	Number of movies produced in the USA or India in the year 2019 is 1818.
Q5	The unique genres are Adventure, Action, Sci-Fi, Crime, Mystery, Drama, Fantasy, Thriller, Comedy, Horror, Family, Romance and Others
Q6	The Drama genre has the highest movies produced (4285) in all time.

Q7	Movies that were made in only one genre are 3289 in number.
Q8	Avg duration of movies based on genres is: Drama 106.7746 Fantasy 105.1404 Thriller 101.5761 Comedy 102.6227 Horror 92.7243 Family 100.9669 Romance 109.5342 Adventure 101.8714 Action 112.8829 Sci-Fi 97.9413 Crime 107.0517 Mystery 101.8000 Others 100.1600
Q9	Movie count for the genre Thriller is 1484 making it the 3rd ranking in all genres
Q10	Min avg rating: 1.0 Max avg rating: 10.0 Min total votes: 100 Max total votes: 725138 Min median rating: 1 Min median rating: 10 Which means there are no outliers

Q11 Top 10 movies:

Movie	Avg rating
Kirket	10.0
Love in Kilnerry	10.0
Gini Helida Kathe	9.8
Runam	9.7
Fan	9.6
Android Kunjappan Version 5.25	9.6
Yeh Suhaagraat Impossible	9.5
Safe	9.5
The Brighton Miracle	9.5
Shibu	9.4

Q12

Median rating	Movie count
7	2257
6	1975
8	1030
5	985
4	479
9	429
10	346
3	283

2	119
1	94

Q13	Production companies at Rank 1 are Dream Warrior Pictures and National Theatre Live
Q14	No. of movies that were released in the USA which had more than 1000 votes belong to the following genres:

Genre	No. of movies
Drama	24
Comedy	9
Action	8
Thriller	8
Sci-Fi	7
Crime	6
Horror	6
Mystery	4
Romance	4
Fantasy	3
Adventure	3
Family	1

Drama genre had most no. of releases which had 1000 votes in the USA

Q15			
	Movie Title	**Avg rating**	**Genre**
	The Brighton Miracle	9.5	Drama
	The Colour of Darkness	9.1	Drama
	The Blue Elephant 2	8.8	Drama
	The Irishman	8.7	Crime
	The Mystery of Godliness: The Sequel	8.5	Drama
	The Gambinos	8.4	Crime
	Theeran Adhigaaram Ondru	8.3	Action
	The King and I	8.2	Drama

Q16	361 movies were given a median rating of 8 between the period of 1 April 2018 and 1 April 2019.
Q17	Yes, votes for German language (4421525) are more than the votes received for Italian language movies (2559540).
Q18	Columns height, date_of_birth and known_for_movies in the table names have 17335, 13431 and 15226 null values respectively.
Q19	The top three directors in the top three genres whose movies have an average rating > 8 are James Mangold, Anthony Russo, Soubin Shahir
Q20	The top two actors whose movies have a median rating >= 8 are Manmmootty and Mohanlal
Q21	The top three production houses based on the number of votes are 1. Marvel Studios 2. Twentieth Century Fox

	3. Warner Bros.
Q22	The top ranking actor based on their avg ratings is Vijay Sethupathi
Q23	The top ranking actress based on their avg ratings is Taapsee Pannu
Q24	Labelling thriller genre movies based on their avg ratings
Q25	Output shows the genre-wise running total and moving average of the average movie duration
Q26	The top five movies in the top three genres in 2017, 2018 and 2019 are:

Q26 table:

Genre	Year	Title	Gross_Income	Rank
Action	2017	Star Wars: Episode VIII - The Last Jedi	1332539889	1
Action	2017	The Fate of the Furious	1236005118	2
Comedy	2017	Despicable Me 3	1034799409	3
Action	2017	Jumanji: Welcome to the Jungle	962102237	4
Action	2017	Spider-Man: Homecoming	880166924	5
Action	2018	Avengers: Infinity War	2048359754	1
Action	2018	Black Panther	1346913161	2
Action	2018	Jurassic World: Fallen Kingdom	1308467944	3
Action	2018	Incredibles 2	1242805359	4
Action	2018	Aquaman	1148161807	5
Action	2019	Avengers: Endgame	2797800564	1
Drama	2019	The Lion King	1655156910	2
Action	2019	Spider-Man: Far from Home	1131845802	3
Action	2019	Captain Marvel	1128274794	4
Comedy	2019	Toy Story 4	1073168585	5

Q27	The top two production houses that have produced the highest number of hits (median rating >= 8) among multilingual movies are Star Cinema and Twentieth Century Fox
Q28	The top 3 actresses based on number of Super Hit movies (average rating >8) in drama genre are Parvathy Thiruvothu, Susan Brown, Amanda Lawrence
Q29	The top 9 directors based on number of movies are: 1. Andrew Jones 2. A.L. Vijay 3. Sion Sono 4. Chris Stokes 5. Sam Liu 6. Steven Sodenbergh 7. Jesse V. Johnson 8. Justin Price 9. Özgür Bakar

Recommendations for a successful venture:

1. **Production House:** The RSVP Movies should consider collaborating with Dream Warrior Pictures and National Theatre Live as they are ranking no. 1 based on no. of movies produces and Marvel Studios based on the number of global votes. Also, for a multilingual approach, to produce a high rating (>8) movie, they should collaborate with Star Cinema and Twentieth Century Fox

2. **Genre:** Based on past trends, we may assume that if more no. of movies were produced in the Drama genre (globally) it might be an audience favourite and producing a movie in the Drama genre may be profitable. Also, according to the voting trends in the USA, Drama has received more than 1000 votes for 24 movies. If the production houses want to look at genres apart from Drama, they should look at the other two genres as well in the global top three: Drama, Action and Comedy.

3. **Time of Release:** Similarly, since most movies were released in the month of March globally, we might assume that this decision was taken in the past as film makers must have observed that March might be the highest grossing month and people were willing to come out to see movies more often in March. Hence, RSVP Movies should look at a date in March to release their upcoming movie for more collection.

4. **Name of the Movie:** The name of the movie should start with "The" for a much better audience appeal and rating.

5. **Duration:** In any genre, the duration of the movie made should ideally be at least 102.4 and at most be 112.9 minutes based on the moving average grouped by each genre.

6. **Language preference:** To suit the global audience better, RSVP movies should consider making the movie in German language along with the other languages chosen due to a higher number of audience speaking the German language.

7. **Directors:** The directors that RSVP Movies should be working with are James Mangold, Anthony Russo and Soubin Shahir. Furthermore, if they want to look at a director who is the most experienced, they should also consider working with Andrew Jones.

8. **Actors:** Based on median ratings the actors that should be hired to work in the upcoming movie should be Manmmootty and Mohanlal. However, based on average ratings, the best choice of actor and actress who should be hired are Vijay Sethupathi and Taapsee Pannu. Along with this, the actresses who have worked and given superhit movies and Parvathy, Thiruvothu, Susan Brown and Amanda Lawrence.