RSVP MOVIES' GLOBAL EXPANSION

- SQL CASE STUDY

PRESENTED BY,
SWETHA KIZHAVANA JOSEPH
BUSINESS ANALYST INTERN

INTRODUCTION

❖ RSVP Movies is a renowned Indian film production company known for producing several super-hit movies tailored for the Indian audience, aims to release its next major movie for a global audience in 2024.

❖ The goal of this analysis is to use SQL to derive meaningful insights from past movie data, helping RSVP Movies make data-driven decisions for their new project.

❖ The project is divided into four segments, each focusing on different aspects like release trends, ratings analysis, genre popularity, and industry leaders.

❖ Through this analysis, key insights were obtained on optimal release timing, top-performing genres, successful directors and actors, and strategies for global audience appeal.
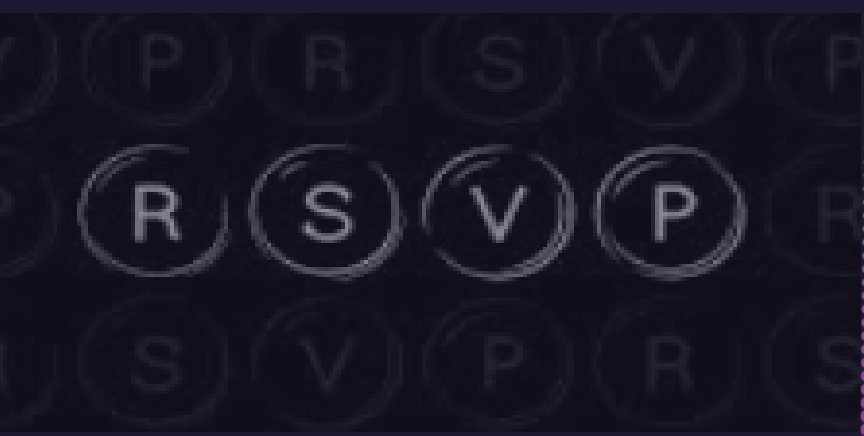
BUSINESS PROBLEM

❖ RSVP Movies has a strong track record of producing super-hit movies for the Indian market but is now looking to expand globally with its next major release in 2024.

❖ The challenge lies in understanding global audience preferences, which differ significantly from the domestic market.

❖ To achieve commercial success globally, RSVP Movies must choose internationally appealing genres, select directors and actors with proven popularity, time the release to align with global trends, and partner with production houses for effective marketing and distribution.

❖ RSVP Movies aims to leverage SQL analysis of historical movie data to identify key trends and guide strategic decisions for their global expansion.

DATASET DESCRIPTION

❖ The analysis uses data from the IMDb database, covering the past three years (2017 to 2019) of movie releases.

❖ Key tables include:

I.   Movie: Contains details like id, title, year, country, duration, worldwide_gross_income.
II. Genre: Lists the genres associated with each movie.
III. Ratings: Includes avg_rating, median_rating, total_votes, and other rating metrics.
IV. Names: Information about actors, directors, and other key personnel.
V.   Director & Role Mapping: Links movies with actors and directors.

TABLE_NAME	TABLE_ROWS
names	25735
genre	14662
role_mapping	14397
movie	7997
ratings	7997
director_mapping	3867

METHODOLOGY

Data Extraction and Cleaning

Queried data from the IMDB SQL Database (Movies, Genres, Ratings tables), addressed null values, and standardized worldwide gross income to USD for consistency.

Data Analysis

Analysed trends in movie releases, genre popularity, and ratings using SQL functions like GROUP BY, JOIN, WINDOW FUNCTIONS (e.g., ROW_NUMBER, DENSE_RANK) and CTEs for efficient data ranking and summarization.

Insight Generation

Derived key insights and recommendations for RSVP Movies' 2024 global release strategy.

ENTITY- RELATIONSHIP DIAGRAM (ERD)

The diagram represents the relationships between key entities in the IMDb database used for analysis. The ERD provides a clear understanding of how movie-related data is structured and connected, allowing for efficient querying and analysis.
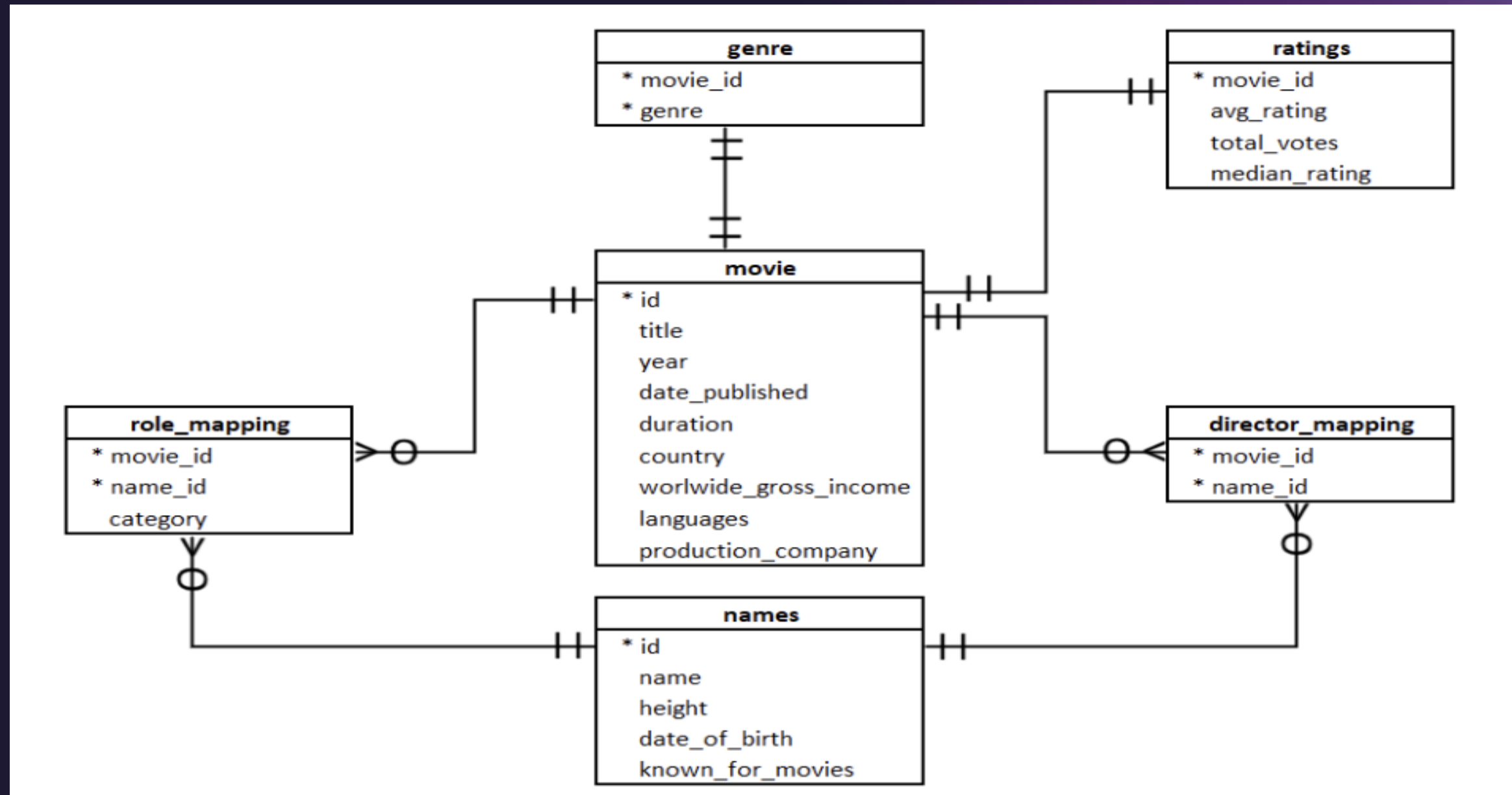
01 SEGMENT – OVERVIEW OF MOVIE RELEASES

Year	Number_of_movies
2017	3052
2018	2944
2019	2001

| Month_Num	Number_of_movies
3	824
9	809
1	804
10	801
4	680
8	678
2	640
11	625
5	625
6	580
7	493
12	438

❖ **Total Number of Movies Released Each Year:**

As displayed in the table and line chart there is a declining trend in the number of movies released each year, with a significant drop from 2018 to 2019.

❖ **Month-wise Trend for Movie Releases:**

The highest number of movies is released in March and the lowest number of movies is released in December, indicating a potential slowdown during the year-end.

Yearly Trend of Movie Releases (2017-2019)
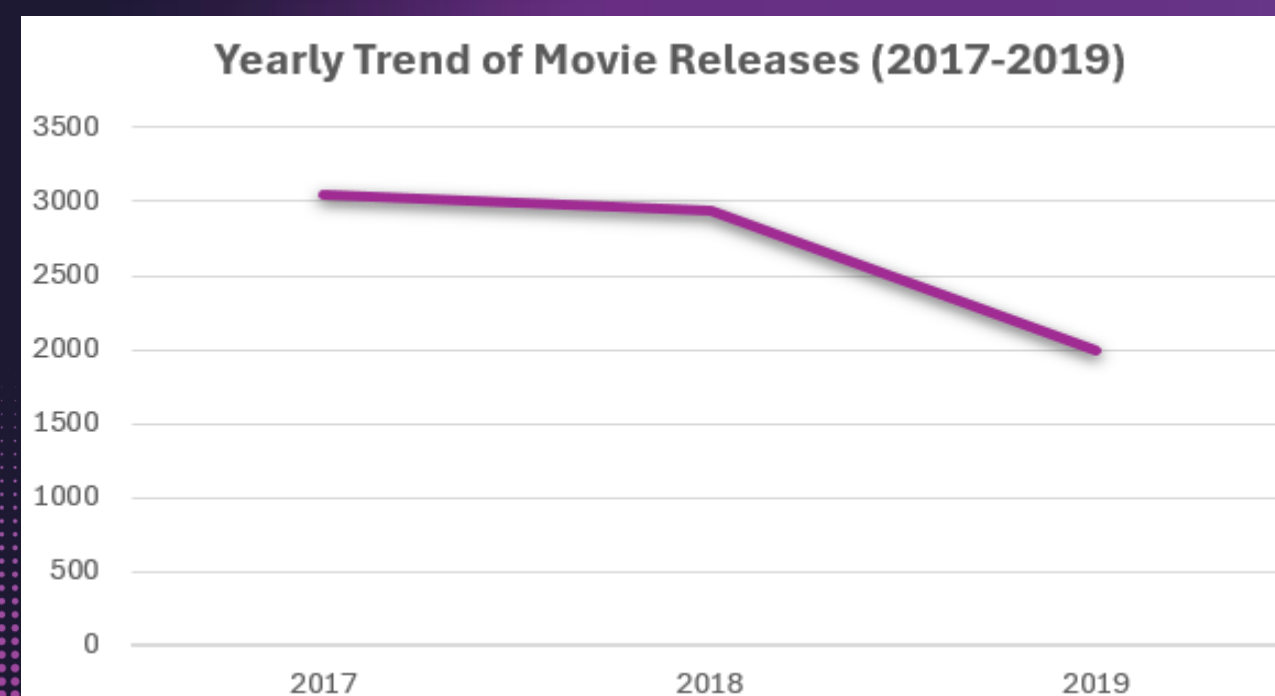
01 SEGMENT – OVERVIEW OF MOVIE RELEASES

Total_Movies
1059

genre
Action
Adventure
Comedy
Crime
Drama
Family
Fantasy
Horror
Mystery
Others
Romance
Sci-Fi
Thriller

genre	Movie_Count
Drama	4285

❖ **Movies Produced in the USA or India in 2019:**

A total of 1,059 movies were produced in 2019 in either the USA or India.

❖ **Unique List of Genres:**

The dataset includes 13 unique genres: Action, Adventure, Comedy, Crime, Drama, Family, Fantasy, Horror, Mystery, Others, Romance, Sci-Fi, and Thriller.

❖ **Top Genre by Movie Count:**

The Drama genre had the highest number of movies produced overall, with a total of 4,285 movies.

01 SEGMENT – OVERVIEW OF MOVIE RELEASES

Movie_With_One_Genre
3289

genre	avg_duration
Action	112.88
Romance	109.53
Crime	107.05
Drama	106.77
Fantasy	105.14
Comedy	102.62
Adventure	101.87
Mystery	101.80
Thriller	101.58
Family	100.97
Others	100.16
Sci-Fi	97.94
Horror	92.72

genre	Movie_Count	Genre_Rank
Thriller	1484	3

❖ **Number of Movies with a Single Genre:**

3,289 movies belong to only one genre.

❖ **Average Duration by Genre:**

Action movies have the longest average duration at 112.88 minutes, while horror is the shortest at 92.72 minutes, with most genres ranging between 100 to 110 minutes.

❖ **Rank of Thriller Genre:**

The Thriller genre is ranked 3rd in terms of the number of movies produced, with 1,484 movies.

02 SEGMENT – RATING ANALYSIS

min_avg_rating	max_avg_rating	min_total_votes	max_total_votes	min_median_rating	max_median_rating
1.0	10.0	100	725138	1	10

title	avg_rating	movie_rank
Kirket	10.0	1
Love in Kilnerry	10.0	1
Gini Helida Kathe	9.8	2
Runam	9.7	3
Fan	9.6	4
Android Kunjappan Version 5.25	9.6	4
Yeh Suhaagraat Impossible	9.5	5
Safe	9.5	5
The Brighton Miracle	9.5	5
Shibu	9.4	6

Distribution of Movies by Median Ratings

Median Rating	Movie_Count
1	94
2	119
3	283
4	479
5	985
6	1975
7	2257
8	1030
9	429
10	346

❖ **Minimum and Maximum Values of the Ratings:**

This data shows that movie ratings range from a minimum of 1.0 to a maximum of 10.0, with total votes varying between 100 and 725,138, and median ratings also ranging from 1 to 10. This implies there are no outliers in the table.

❖ **Top 10 Movies by Average Rating:**

Multiple movies share the same average rating, with "Kirket" and "Love in Kilnerry" tied at rank 1, and several films tied at rank 5 with a 9.5 rating.

❖ **Median Ratings Summary:**

The distribution shows that most movies have a median rating of 7 (2257 movies), with very few rated 1 or 2, suggesting that the majority of movies tend to receive mid-range to high ratings.

SEGMENT – RATING ANALYSIS

02
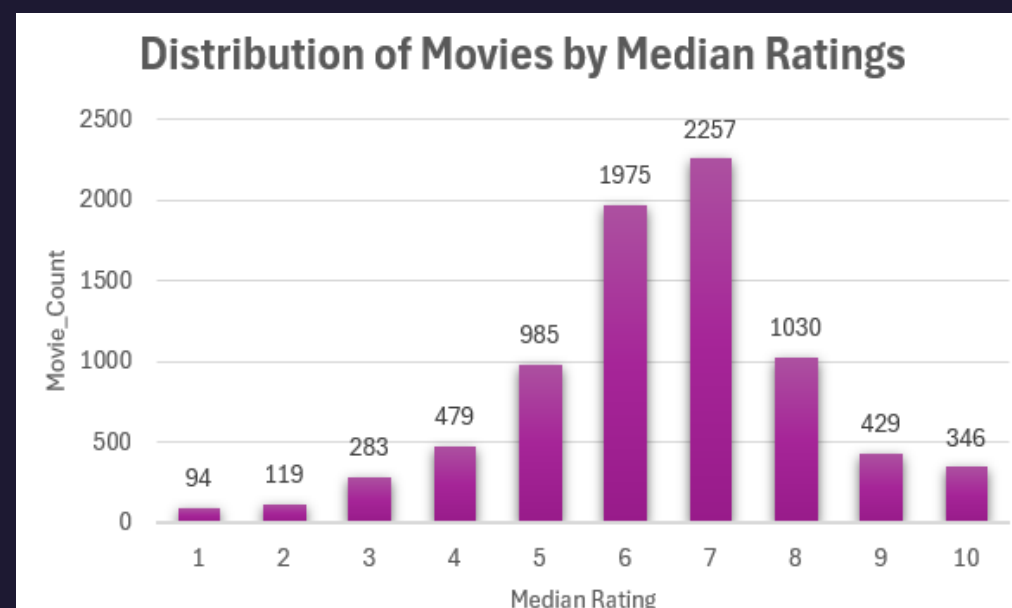
production_company	movie_count	prod_company_rank
Dream Warrior Pictures	3	1
National Theatre Live	3	1

Distribution of Movies by Genre

genre	Movie_Count
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

❖ **Production House with Most Hit Movies:**

Dream Warrior Pictures and National Theatre Live both produced 3 hit movies with an average rating above 8, sharing the top rank.
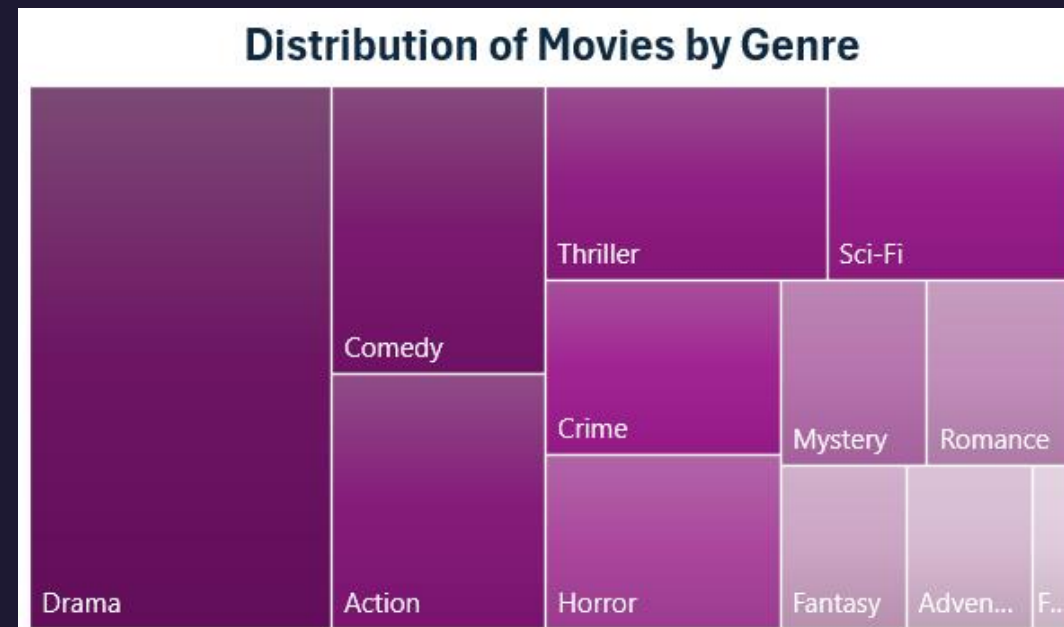
❖ **Movies Released in Each Genre During March 2017 (USA) with >1,000 Votes:**

In March 2017 in the USA, Drama had the most releases with 24 movies having more than 1,000 votes, followed by Comedy with 9 movies, while Family had the fewest with only 1 release. The Tree map visualises the distribution of movie counts across genres, each genre is represented as a rectangle, with its size proportional to the number of movies it contains.

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 Blue Elephant 2	8.8	Horror
The Blue Elephant 2	8.8	Mystery
The Irishman	8.7	Crime
The Irishman	8.7	Drama
The Mystery of Godliness: The Sequel	8.5	Drama
The Gambinos	8.4	Crime
The Gambinos	8.4	Drama
Theeran Adhigaaram Ondru	8.3	Action
Theeran Adhigaaram Ondru	8.3	Crime
Theeran Adhigaaram Ondru	8.3	Thriller
The King and I	8.2	Drama
The King and I	8.2	Romance

median_rating	Movie_released_april2018_april2019
8	361

country	Total_Votes
Germany	106710
Italy	77965

❖ **Movies Starting with 'The' and Having an Average Rating > 8:**
The list features high-rated movies starting with "The" and has an average rating greater than 8 across various genres, with Drama being the most common, including "The Brighton Miracle" (9.5) and "The Colour of Darkness" (9.1).

❖ **Movies with a Median Rating of 8 between 1 April 2018 and 1 April 2019:**
A total of 361 movies with a median rating of 8 were released between April 2018 and April 2019.

❖ **Comparison of Votes - Germany vs. Italy:**
German movies received a total of 106,710 votes, which is higher than the 77,965 votes received by Italian movies.

Director_Name	Movie_Count
James Mangold	4
Joe Russo	3
Anthony Russo	3

❖ **Top 3 Directors in the Top 3 Genres (Avg. Rating > 8):**
The top directors with movies averaging a rating above 8 are James Mangold (4 movies), Joe Russo (3 movies), and Anthony Russo (3 movies).

Actor_Name	Movie_Count
Mammootty	8
Mohanlal	5

❖ **Top 2 Actors (Median Rating ≥ 8):**
The top two actors whose movies have a median rating of 8 or higher are Mammootty with 8 movies and Mohanlal with 5 movies.

Top 2 Actors by Movie Count

5

8

■ Mammootty   ■ Mohanlal
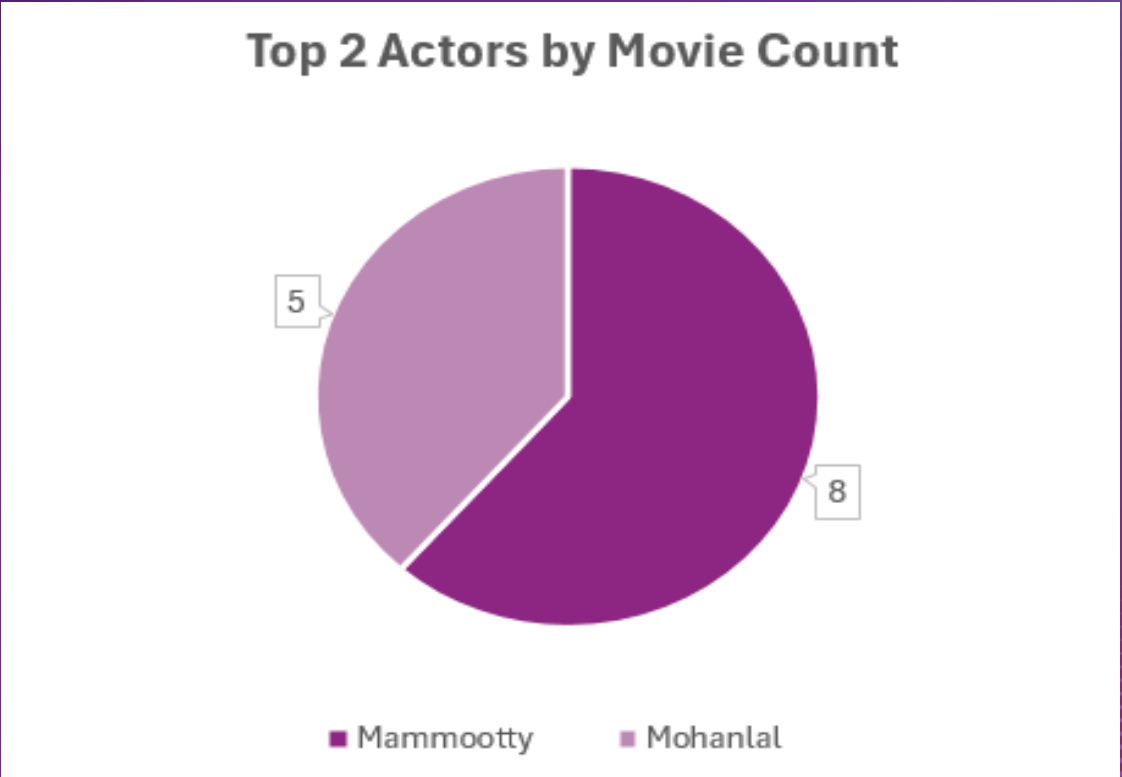
SEGMENT – GENRE & PERFORMANCE ANALYSIS

production_company	Vote_count	prod_comp_rank
Marvel Studios	2656967	1
Twentieth Century Fox	2411163	2
Warner Bros.	2396057	3

❖ **Top 3 Production Houses by Votes:**

The top three production houses based on the number of votes received by their movies are Marvel Studios (2,659,667 votes), Twentieth Century Fox (2,411,163 votes), and Warner Bros. (2,396,057 votes).

❖ **Top Actor in India Based on Average Ratings:**

Vijay Sethupathi ranks 1st among actors with movies released in India, having an average rating of 8.42.

actor_name	total_votes	movie_count	actor_avg_rating	actor_rank
Vijay Sethupathi	23114	5	8.42	1
Fahadh Faasil	13557	5	7.99	2
Yogi Babu	8500	11	7.83	3
Joju George	3926	5	7.58	4
Ammy Virk	2504	6	7.55	5
Dileesh Pothan	6235	5	7.52	6
Kunchacko Boban	5628	6	7.48	7
Pankaj Tripathi	40728	5	7.44	8
Rajkummar Rao	42560	6	7.37	9
Dulquer Salmaan	17666	5	7.30	10
Amit Sadh	13355	5	7.21	11
Tovino Thomas	11596	8	7.15	12
Mammootty	12613	8	7.04	13
Nassar	4016	5	7.03	14
Karamjit Anmol	1970	6	6.91	15

03 SEGMENT – GENRE & PERFORMANCE ANALYSIS

actress_name	total_votes	movie_count	actress_avg_rating	actress_rank
Taapsee Pannu	18061	3	7.74	1
Kriti Sanon	21967	3	7.05	2
Divya Dutta	8579	3	6.88	3
Shraddha Kapoor	26779	3	6.63	4
Kriti Kharbanda	2549	3	4.80	5
Sonakshi Sinha	4025	4	4.18	6

❖ **Top 5 Actresses in Hindi Movies (Based on Average Ratings):**

Taapsee Pannu tops with an average rating of 7.74.

❖ **Classification of Thriller Movies by Rating:**

- Superhit Movies (Rating > 8): The Brighton Miracle (9.5), The Colour of Darkness (9.1).

- Hit Movies (Rating 7-8): Incitement (7.5), Uriyadi 2 (7.3).

- One-time-watch Movies (Rating 5-7): Dukun (6.9), Pet Sematary (5.8).

- Flop Movies (Rating < 5): Fahrenheit 451 (4.9), Staged Killer (3.3).

title	avg_rating	movie_category
Der müde Tod	7.7	Hit movies
Fahrenheit 451	4.9	Flop movies
Pet Sematary	5.8	One-time-watch movies
Dukun	6.9	One-time-watch movies
Back Roads	7.0	Hit movies
Countdown	5.4	One-time-watch movies
Staged Killer	3.3	Flop movies
Vellaipookal	7.3	Hit movies
Uriyadi 2	7.3	Hit movies
Incitement	7.5	Hit movies
Rakshasudu	8.4	Superhit movies
Trois jours et …	6.6	One-time-watch movies
Killer in Law	5.1	One-time-watch movies

SEGMENT – CONTRIBUTORS ANALYSIS

genre	avg_duration	running_total_duration	moving_avg_duration
Action	112.88	112.88	112.88
Adventure	101.87	214.75	107.38
Comedy	102.62	317.37	105.79
Crime	107.05	424.42	106.11
Drama	106.77	531.19	106.24
Family	100.97	632.16	105.36
Fantasy	105.14	737.30	105.33
Horror	92.72	830.02	103.75
Mystery	101.80	931.82	103.54
Others	100.16	1031.98	103.20
Romance	109.53	1141.51	103.77
Sci-Fi	97.94	1239.45	103.29
Thriller	101.58	1341.03	103.16

❖ **Genre-wise Running Total & Moving Average of Movie Duration:**

The running total and moving average of average durations help understand cumulative trends, with genres like Action having the longest average duration (112.88 minutes) and a cumulative running total and moving average that gradually increase as genres are listed.

❖**Top 5 Highest-Grossing Movies Each Year (Top 3 Genres):**

- The 2017 top-grossing movies include The Fate of the Furious ($1,236,005,118) and Despicable Me 3.
- 2018 highlights include Bohemian Rhapsody ($903,635,259) and Venom.
- 2019 saw hits like Avengers: Endgame ($2,797,800,564) and The Lion King.

genre	year	movie_name	worldwide_gross_income ($)	movie_rank
Thriller	2017	The Fate of the Furious	$1236005118	1
Comedy	2017	Despicable Me 3	$1034799409	2
Comedy	2017	Jumanji: Welcome to the Jungle	$962102237	3
Drama	2017	Zhan lang II	$870325439	4
Thriller	2017	Zhan lang II	$870325439	4
Comedy	2017	Guardians of the Galaxy Vol. 2	$863756051	5
Drama	2018	Bohemian Rhapsody	$903655259	1
Thriller	2018	Venom	$856085151	2
Thriller	2018	Mission: Impossible - Fallout	$791115104	3
Comedy	2018	Deadpool 2	$785046920	4
Comedy	2018	Ant-Man and the Wasp	$622674139	5
Drama	2019	Avengers: Endgame	$2797800564	1
Drama	2019	The Lion King	$1655156910	2
Comedy	2019	Toy Story 4	$1073168585	3
Drama	2019	Joker	$995064593	4
Thriller	2019	Joker	$995064593	4
Thriller	2019	Ne Zha zhi mo tong jiang shi	$700547754	5

SEGMENT – CONTRIBUTORS ANALYSIS

04

production_company	movie_count	prod_comp_rank
Star Cinema	7	1
Twentieth Century Fox	4	2

actress_name	total_votes	movie_count	actress_Avg_rating	actress_rank
Parvathy Thiruvothu	4974	2	8.20	1
Susan Brown	656	2	8.95	2
Amanda Lawrence	656	2	8.95	3

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	176.75	5.65	1754	3.7	6.9	613
nm2096009	Andrew Jones	5	190.75	3.04	1989	2.7	3.2	432
nm0001752	Steven Soderbergh	4	254.33	6.77	171684	6.2	7.0	401
nm0425364	Jesse V. Johnson	4	299.00	6.10	14778	4.2	6.5	383
nm0515005	Sam Liu	4	260.33	6.32	28557	5.8	6.7	312
nm0814469	Sion Sono	4	331.00	6.31	2972	5.4	6.4	502
nm0831321	Chris Stokes	4	198.33	4.32	3664	4.0	4.6	352
nm2691863	Justin Price	4	315.00	4.93	5343	3.0	5.8	346
nm6356309	Özgür Bakar	4	112.00	3.96	1092	3.1	4.9	374

❖ **Top 2 Production Houses for Multilingual Hits:**

Star Cinema is the top production company with 7 movies, followed by Twentieth Century Fox with 4 hits.

❖ **Top 3 Actresses with Super Hit Movies in Drama Genre (Avg. Rating > 8):**

Parvathy Thiruvothu ranks first among actresses with an average rating of 8.20 across 2 movies, followed by Susan Brown and Amanda Lawrence, with an 8.95 rating but fewer total votes.

❖ **Top 9 Directors (Based on Number of Movies):**

The top 9 directors, led by A.L. Vijay and Andrew Jones with 5 movies each, vary in average ratings and total votes, with Steven Soderbergh having the highest average rating (6.77) and most votes (171,684).

DATA INSIGHTS

CONCLUSIONS :

❖ Trends in Movie Releases: The number of movie releases has declined from 2017 to 2019, with March being the peak month for releases.

❖ Genre Insights: Drama is the most produced genre, while Action movies tend to have the longest durations.

❖ Ratings & Popularity: Most movies have a median rating of 7, with fewer rated 1 or 2, indicating generally favourable audience reception.

❖ Top Performers: Directors like James Mangold and actors such as Vijay Sethupathi excel in highly-rated films.

❖ Box Office Success: Movies like Avengers: Endgame and production houses like Marvel Studios dominate in gross income.

RECOMMENDATIONS :

❖ Focus on high-quality production in Drama or Action genres to meet international standards and attract a broader audience.

❖ Emphasize strong storytelling and cinematic excellence to ensure positive reception and higher ratings.

❖ Collaborate with top-rated directors and actors to increase the movie's quality and appeal.

❖ Consider releasing new movies in optimal timings for higher visibility and market impact.

❖ Focus on multilingual releases and explore partnerships with leading production houses to leverage their market reach and boost international promotion.

THANK YOU!

✉ swethakjoseph16@gmail.com

🌐 https://www.linkedin.com/in/swetha-kizhavana-joseph-04b68721b/