Executive Summary — IMDB Movies SQL Project

1. INTRODUCTION

The IMDB Movies SQL Project is a comprehensive case study based on the RSVP Movies dataset. It aims to uncover data-driven insights for strategic decision-making in the film industry. The project answers 50 real-world business questions across multiple segments, focusing on movie trends, ratings, genres, actors, directors, and box office performance.

2. PROBLEM STATEMENT

RSVP Movies, following the success of their previous production 'The Sky Is Pink', planned to produce an international film targeting a global audience. To make informed decisions, they needed insights on audience preferences, profitable genres, top-performing talent, and market trends.

3. DATASET OVERVIEW

The database contains 6 core tables with information about movies, ratings, genres, actors, directors, and related mappings. The dataset was provided in both SQL and Excel formats, with an ERD to illustrate table relationships.

4. KEY TABLES

Table Name	Rows	Description
movie	7,997	Movie details including title, year, languages, gross income
ratings	7,997	Ratings and votes
genre	14,662	Genre mapping for movies
names	25,735	Actor/Director names
director_mapping	3,867	Mapping between movies and directors
role_mapping	15,615	Mapping between movies and actors

5. METHODOLOGY & WORKFLOW

The project followed a structured workflow divided into multiple segments:

1. Data Import — Load the SQL dataset into MySQL Workbench.
2. Data Cleaning — Format currency, handle multi-valued fields, create derived columns.
3. Query Execution — Execute 50 queries grouped into 7 analytical segments.
4. Insight Generation — Analyze query outputs to identify trends and patterns.

6. SEGMENT SUMMARY

Segment	Query Range	Focus Area
1	Q1–Q9	Basic filtering and counts
2	Q10–Q17	Actor, director, and genre analysis
3	Q18–Q23	Revenue and ratings
4	Q24–Q29	Complex subqueries and comparisons
5	Q30–Q38	Aggregations and joins
6	Q39–Q45	Box office trends
7	Q46–Q50	Custom KPIs and special metrics

7. KEY INSIGHTS

- Drama and Comedy dominate production volumes, while Action movies yield higher revenues.
- Certain directors consistently achieve above-average ratings.
- US and India lead in movie production.
- Multilingual releases expand audience reach and market penetration.

8. RECOMMENDATIONS

- Focus on Drama/Action genres for global releases.
- Collaborate with top-rated directors for consistent quality.
- Utilize multilingual strategies for broader audience engagement.
- Explore co-productions between US and Indian studios.

9. TOOLS & TECHNOLOGIES USED

- MySQL Workbench — For query execution
- SQL — For data extraction and transformation
- Excel — For dataset exploration and ERD visualization
- AI (ChatGPT) — For query planning and documentation

10. IMPORTANT SETUP INSTRUCTION

Before running any queries:

1. Create a MySQL connection/server in MySQL Workbench.
2. Import the main dataset (IMDB+dataset+import.sql) into your database (schema: imdb).
3. Run the solved project file (IMDB_SQL_Solved_Project_SahilVerma_All_Segments.sql) to execute all queries.

11. ABOUT THE AUTHOR

Sahil Verma — Data Analyst specializing in SQL, Python, and Data Visualization. Experienced in delivering actionable insights through structured data analysis.

12. ENDING NOTE

This project demonstrates the power of SQL in answering real-world business questions. By structuring queries into segments and focusing on key metrics, it provides both technical depth and actionable insights. Aspiring analysts are encouraged to attempt the included practice questions to strengthen their SQL skills.