



MOVIE DATA SET

A Data Exploration and Insight Report

Prabhakar Singh

CONTENTS



1

DATA Set
Overview

2

Objectives

3

SQL Queries and
result



OBJECTIVES



The objective of this dataset is to explore and analyze movies from multiple film industries using their ratings, and associations with languages and studios. It enables queries across relational tables to draw insights into viewer preferences, studio output, and industry trends.

This analysis can help in understanding trends, identifying low or high-performing movies, and making data-driven decisions in film production and distribution.

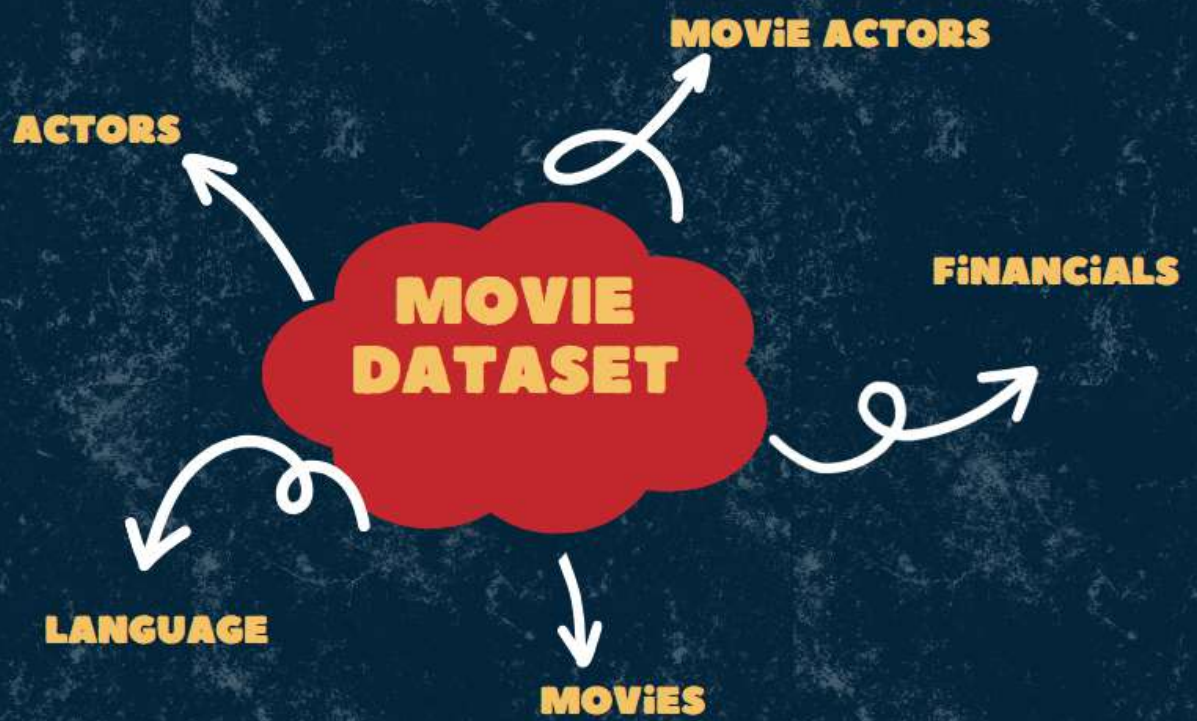


TABLE OVERVIEW

Tables name are as follows : movies, movie_actor, languages, financial, actors

movies

- movie_id
- title
- industry
- release_year
- test
- imdb_rating
- studio
- language_id

movie_actor

- movie_id
- actor_id

languages

- language_id
- name

financial

- movie_id
- budget
- revenue
- unit
- currency

actors

- actor_id
- name
- birth_year



ACTiON

GRAB THE POPCORN!

Select all the rows from the movies table whose imdb_rating is higher than the average rating

```
select * from movies where imdb_rating > (select avg(imdb_rating) from movies);
```

movie_id	title	industry	release_year	imdb_rating	studio	language_id
101	K.G.F: Chapter 2	Bollywood	2022	8.4	Hombale Films	3
106	Sholay	Bollywood	1975	8.1	United Producers	1
107	Dilwale Dulhania Le Jayenge	Bollywood	1995	8.0	Yash Raj Films	1
108	3 Idiots	Bollywood	2009	8.4	Vinod Chopra Films	1
111	The Shawshank Redemption	Hollywood	1994	9.3	Castle Rock Entertainment	5
112	Inception	Hollywood	2010	8.8	Warner Bros. Pictures	5
113	Interstellar	Hollywood	2014	8.6	Warner Bros. Pictures	5
115	The Pursuit of Happiness	Hollywood	2006	8.0	Columbia Pictures	5

MOVIES THAT PRODUCE 500% OR MORE PROFIT AND THEIR RATING WAS LESS THAN AVERAGE RATING OF ALL MOVIES

```
with x as (select *, (revenue-budget)*100 / budget as profit_pct from financials),  
y as (select * from movies where imdb_rating < (select avg(imdb_rating)from movies))  
select x.profit_pct,x.movie_id, y.title,y.imdb_rating  
from x join y  
on x.movie_id = y.movie_id  
where profit_pct >= 500;
```

	title	imdb_rating	profit_pct	movie_id
▶	Titanic	7.9	1001.000000	117
	Avatar	7.8	1101.265823	119



SELECT ALL ACTORS WHOSE AGE IS BETWEEN 70 AND 80

```
with actors_age as (  
  select name, year(curdate()) - birth_year as age from moviesdb.actors  
)  
select name , age from actors_age where age>70 and age<80 ;
```

	name	age
▶	Jaya Bachchan	77
	Liam Neeson	73
	Sam Neill	78
	Mithun Chakraborty	75
	Tommy Lee Jones	79



Select all the rows from the movies table whose imdb_rating is higher than the average rating

```
select * from movies where imdb_rating > (select avg(imdb_rating) from movies);
```

movie_id	title	industry	release_year	imdb_rating	studio	language_id
101	K.G.F: Chapter 2	Bollywood	2022	8.4	Hombale Films	3
106	Sholay	Bollywood	1975	8.1	United Producers	1
107	Dilwale Dulhania Le Jayenge	Bollywood	1995	8.0	Yash Raj Films	1
108	3 Idiots	Bollywood	2009	8.4	Vinod Chopra Films	1
111	The Shawshank Redemption	Hollywood	1994	9.3	Castle Rock Entertainment	5
112	Inception	Hollywood	2010	8.8	Warner Bros. Pictures	5
113	Interstellar	Hollywood	2014	8.6	Warner Bros. Pictures	5
115	The Pursuit of Happiness	Hollywood	2006	8.0	Columbia Pictures	5

SELECT ALL THE MOVIES WITH MINIMUM AND MAXIMUM RELEASE YEAR. NOTE THAT THERE CAN BE MORE THAN ONE MOVIE IN MIN AND A MAX YEAR HENCE OUTPUT ROWS CAN BE MORE THAN 2

```
select * from movies where release_year in (  
(select min(release_year) from movies) ,(select max(release_year) from movies));
```

movie_id	title	industry	release_year	imdb_rating	studio	language_id
101	K.G.F: Chapter 2	Bollywood	2022	8.4	Hombale Films	3
102	Doctor Strange in the Multiverse of Madness	Hollywood	2022	7.0	Marvel Studios	5
105	Thor: Love and Thunder	Hollywood	2022	6.8	Marvel Studios	5
118	It's a Wonderful Life	Hollywood	1946	8.6	Liberty Films	5
133	RRR	It's a Wonderful Life	2022	8.0	DVV Entertainment	2
135	The Karabnis Files	Bollywood	2022	8.3	Zee Studios	1

SELECT ACTOR ID,ACTOR NAME AND TOTAL NUMBER OF MOVIES THEY ACTED IN

```
select actor_id , name,  
(select count(*) from movie_actor where actor_id = actors.actor_id) as actors_count  
from actors order by actors count desc;
```

Result Grid		 Filter Rows:	
	actor_id	name	actors_count
▶	54	Chris Hemsworth	5
	95	Chris Evans	4
	61	Aamir Khan	3
	94	Robert Downey Jr.	2
	51	Sanjay Dutt	2
	55	Natalie Portman	2
	56	Tom Hiddleston	2
	57	Amitabh Bachchan	2
	69	Leonardo DiCaprio	2
	59	Shah Rukh Khan	2
	154	Anushka Sharma	2
	164	Salman Khan	2

SELECT ALL MOVIES WHOSE RATING IS GREATER THAN *ANY* OF THE MARVEL MOVIES RATINGS

```
select * from movies where imdb_rating > any (  
    select imdb_rating from movies where studio = "Marvel Studios"  
);
```

movie_id	title	industry	release_year	imdb_rating	studio	language_id
101	K.G.F: Chapter 2	Bollywood	2022	8.4	Hombale Films	3
102	Doctor Strange in the Multiverse of Madness	Hollywood	2022	7.0	Marvel Studios	5
104	Thor: Ragnarok	Hollywood	2017	7.9	Marvel Studios	5
106	Sholay	Bollywood	1975	8.1	United Producers	1
107	Dilwale Dulhania Le Jayenge	Bollywood	1995	8.0	Yash Raj Films	1
108	3 Idiots	Bollywood	2009	8.4	Vinod Chopra Films	1
109	Kabhi Khushi Kabhie Gham	Bollywood	2001	7.4	Dharma Productions	1
110	Bajirao Mastani	Bollywood	2015	7.2		1
111	The Shawshank Redemption	Hollywood	1994	9.3	Castle Rock Entertainment	5
112	Inception	Hollywood	2010	8.8	Warner Bros. Pictures	5

SELECT ALL MOVIES WHOSE RATING IS GREATER THAN *ANY* OF THE MARVEL MOVIES RATINGS

```
select * from movies where imdb_rating > any (  
  select imdb_rating from movies where studio = "Marvel Studios"  
) ;
```

movie_id	title	industry	release_year	imdb_rating	studio	language_id
101	K.G.F: Chapter 2	Bollywood	2022	8.4	Hombale Films	3
102	Doctor Strange in the Multiverse of Madness	Hollywood	2022	7.0	Marvel Studios	5
104	Thor: Ragnarok	Hollywood	2017	7.9	Marvel Studios	5
106	Sholay	Bollywood	1975	8.1	United Producers	1
107	Dilwale Dulhania Le Jayenge	Bollywood	1995	8.0	Yash Raj Films	1
108	3 Idiots	Bollywood	2009	8.4	Vinod Chopra Films	1
109	Kabhi Khushi Kabhie Gham	Bollywood	2001	7.4	Dharma Productions	1
110	Bajirao Mastani	Bollywood	2015	7.2		1
111	The Shawshank Redemption	Hollywood	1994	9.3	Castle Rock Entertainment	5
112	Inception	Hollywood	2010	8.8	Warner Bros. Pictures	5



SELECT ACTORS WHO WORKED IN MOVIES (101,110,121)

```
select * from actors where actor_id in (  
select actor_id from movie_actor where movie_id in (101,110,121))
```

actor_id	name	birth_year
50	Yash	1986
51	Sanjay Dutt	1959
65	Ranveer Singh	1985
66	Deepika Padukone	1986
86	Christian Bale	1974
87	Heath Ledger	1979
NULL	NULL	NULL

SELECT ACTORS WHOSE AGE >70 AND <85

```
SELECT  
  *  
FROM  
  (SELECT  
    name, YEAR(CURDATE())  
  FROM  
    actors) AS age_actors  
WHERE  
  age > 70 AND age < 85;
```

Result Grid |   Filter Rows:

	name	age
▶	Amitabh Bachchan	83
	Jaya Bachchan	77
	Liam Neeson	73
	Ben Kingsley	82
	Sam Neill	78
	Mithun Chakraborty	75
	Tommy Lee Jones	79



MOVIES WITH LOWEST IMDB_RATING

```
SELECT
    *
FROM
    movies
WHERE
    imdb_rating = (SELECT
        MIN(imdb_rating)
    FROM
        movies);
```

Result Grid		Filter Rows:	Edit:	Export/Import:	Wrap C		
	movie_id	title	industry	release_year	imdb_rating	studio	language_id
▶	139	Race 3	Bollywood	2018	1.9	Salman Khan Films	1
*	NULL	NULL	NULL	NULL	NULL	NULL	NULL

**THANK
YOU**

