

Movies Data Analysis SQL Project

This document presents a series of SQL queries designed to analyze the movies database. The queries cover a range of operations including basic retrievals, string operations, numerical functions, and joins.

Basic Queries

How would you retrieve all records from the `movies` table?

```
4 • SELECT * FROM moviesdb.movies;
```

5

Result Grid | Filter Rows: | Edit: | Export/Import: | Wrap Cell Content: [IA](#)

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
103	Thor: The Dark World	Hollywood	2013	6.8	Marvel Studios	5
104	Thor: Ragnarok	Hollywood	2017	7.9	Marvel Studios	5
105	Thor: Love and Thunder	Hollywood	2022	6.8	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 Entert...	5
112	Inception	Hollywood	2010	8.8	Warner Bros. Pictures	5

movies 24 x

How would you select all records from the `financials` table?

```
5 • SELECT * FROM moviesdb.financials;
```

Result Grid | Filter Rows: | Edit: |

movie_id	budget	revenue	unit	currency
101	1.00	12.50	Billions	INR
102	200.00	954.80	Millions	USD
103	165.00	644.80	Millions	USD
104	180.00	854.00	Millions	USD
105	250.00	670.00	Millions	USD
107	400.00	2000.00	Millions	INR
108	550.00	4000.00	Millions	INR
109	390.00	1360.00	Millions	INR
110	1.40	3.50	Billions	INR
111	25.00	73.30	Millions	USD
113	165.00	701.80	Millions	USD
114	205.00	365.30	Millions	USD

financials 26 x

String Operations

How would you retrieve a list of distinct studios from the `movies` table?

```
7 • SELECT distinct studio FROM moviesdb.movies;
8
9
```

<

Result Grid | | Filter Rows: | Export: | Wrap Cell Content:

	studio
▶	Hombale Films
	Marvel Studios
	United Producers
	Yash Raj Films
	Vinod Chopra Films
	Dharma Productions
	Castle Rock Entertainment
	Warner Bros. Pictures
	Columbia Pictures
	Universal Pictures
	Paramount Pictures

movies 27 x

How would you count the number of distinct studios in the `movies` table?

```
8 • SELECT COUNT(distinct studio) FROM moviesdb.movies;
9
```

<

Result Grid | | Filter Rows: | Export: | Wrap Cell Content:

	COUNT(distinct studio)
▶	23

How would you find all movies that belong to the "hollywood" industry?

9 • `SELECT * FROM moviesdb.movies where industry="hollywood";`

movie_id	title	industry	release_year	imdb_rating	studio	language_id
102	Doctor Strange in the Multiverse of Madness	Hollywood	2022	7.0	Marvel Studios	5
103	Thor: The Dark World	Hollywood	2013	6.8	Marvel Studios	5
104	Thor: Ragnarok	Hollywood	2017	7.9	Marvel Studios	5
105	Thor: Love and Thunder	Hollywood	2022	6.8	Marvel Studios	5
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 Happyness	Hollywood	2006	8.0	Columbia Pictures	5
116	Gladiator	Hollywood	2000	8.5	Universal Pictures	5
117	Titanic	Hollywood	1997	7.9	Paramount Pictures	5
118	It's a Wonderful Life	Hollywood	1946	8.6	Liberty Films	5
119	Avatar	Hollywood	2009	7.8	20th Century Fox	5

movies 29 x

How would you count the number of movies in the "hollywood" industry?

10 • `SELECT count(*) FROM moviesdb.movies where industry="hollywood";`

11

12

count(*)
21

How would you find all movies that do not belong to the "hollywood" industry?

12 • `SELECT * FROM moviesdb.movies where industry!="hollywood";`

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
109	Kabhi Khushi Kabhie Gham	Bollywood	2001	7.4	Dharma Productions	1
110	Bajirao Mastani	Bollywood	2015	7.2		1
127	Pather Panchali	Bollywood	1955	8.3	Government of West Bengal	7
128	Taare Zameen Par	Bollywood	2007	8.3		1
129	Munna Bhai M.B.B.S.	Bollywood	2003	8.1	Vinod Chopra Productions	1
130	PK	Bollywood	2014	8.1	Vinod Chopra Films	1
131	Sanju	Bollywood	2018	NULL	Vinod Chopra Films	1
132	Pushpa: The Rise - Part 1	Bollywood	2021	7.6	Mythri Movie Makers	2

movies 32 v

How would you find all movies with titles that include the word "thor"?

13 • `SELECT * FROM moviesdb.movies where title like "%thor%";` # Use wildcard search

Result Grid

	movie_id	title	industry	release_year	imdb_rating	studio	language_id
▶	103	Thor: The Dark World	Hollywood	2013	6.8	Marvel Studios	5
	104	Thor: Ragnarok	Hollywood	2017	7.9	Marvel Studios	5
	105	Thor: Love and Thunder	Hollywood	2022	6.8	Marvel Studios	5
*	NULL	NULL	NULL	NULL	NULL	NULL	NULL

How would you find all movies produced by either "Hombale Films" or "Marvel Studios"?

13 • `SELECT * FROM moviesdb.movies where studio in("Hombale Films","Marvel Studios");`

Result Grid

	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
	103	Thor: The Dark World	Hollywood	2013	6.8	Marvel Studios	5
	104	Thor: Ragnarok	Hollywood	2017	7.9	Marvel Studios	5
	105	Thor: Love and Thunder	Hollywood	2022	6.8	Marvel Studios	5
	125	Avengers: Endgame	Hollywood	2019	8.4	Marvel Studios	5
	126	Avengers: Infinity War	Hollywood	2018	8.4	Marvel Studios	5
	137	Captain America: The First Avenger	Hollywood	2011	6.9	Marvel Studios	5
	138	Captain America: The Winter Soldier	Hollywood	2014	7.8	Marvel Studios	5

movies 2 x

Numerical Functions

How would you find all movies with an IMDB rating between 7 and 9?

18 `###Numerical Functions`

19 • `SELECT * FROM moviesdb.movies where imdb_rating between 7 and 9;`

Result Grid

	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
	112	Inception	Hollywood	2010	8.8	Warner Bros. Pictures	5

movies 3 x

How would you find all movies released in the years 2017, 2010, or 2022?

```
19 • SELECT * FROM moviesdb.movies where release_year in (2017,2010,2022);
```

Result Grid | Filter Rows: | Edit: | Export/Import: | Wrap Cell Content: |

	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
	105	Thor: Love and Thunder	Hollywood	2022	6.8	Marvel Studios	5
	112	Inception	Hollywood	2010	8.8	Warner Bros. Pictures	5
	133	RRR	Bollywood	2022	8.0	DVV Entertainment	2
	135	The Kashmir Files	Bollywood	2022	8.3	Zee Studios	1
*	NULL	NULL	NULL	NULL	NULL	NULL	NULL

How would you find all movies where the studio field is empty?

```
18 • SELECT * FROM moviesdb.movies where release_year in (2017,2010,2022);
```

```
19
```

Result Grid | Filter Rows: | Edit: | Export/Import: | Wrap Cell Content: |

	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
	105	Thor: Love and Thunder	Hollywood	2022	6.8	Marvel Studios	5
	112	Inception	Hollywood	2010	8.8	Warner Bros. Pictures	5
	133	RRR	Bollywood	2022	8.0	DVV Entertainment	2
	135	The Kashmir Files	Bollywood	2022	8.3	Zee Studios	1
*	NULL	NULL	NULL	NULL	NULL	NULL	NULL

Operations

How would you find the maximum IMDB rating for "hollywood" movies?

```
24 #####Operations
```

```
25 • SELECT max(imdb_rating) FROM moviesdb.movies where industry="hollywood";
```

Result Grid | Filter Rows: | Export: | Wrap Cell Content: |

	max(imdb_rating)
▶	9.3

How would you find the minimum IMDB rating for "hollywood" movies?

26 • `SELECT min(imdb_rating) FROM moviesdb.movies where industry="hollywood";`

27

<	
Result Grid	Filter Rows: <input type="text"/> Export: Wrap Cell Content:
min(imdb_rating)	
▶ 6.8	

How would you find the average IMDB rating for "hollywood" movies?

27 • `SELECT avg(imdb_rating) FROM moviesdb.movies where industry="hollywood";`

<	
Result Grid	Filter Rows: <input type="text"/> Export: Wrap Cell Content:
avg(imdb_rating)	
▶ 8.16190	

How would you count the number of movies in each industry?

29 • `SELECT industry,count(*) FROM moviesdb.movies group by industry;`

<	
Result Grid	Filter Rows: <input type="text"/> Export: Wrap Cell Content:
industry	count(*)
▶ Bollywood	18
Hollywood	21

How would you count the number of movies in each industry and label the count as industry_count?

29 • `SELECT industry,count(*) as industry_count FROM moviesdb.movies group by industry;`

<	
Result Grid	Filter Rows: <input type="text"/> Export: Wrap Cell Content:
industry	industry_count
▶ Bollywood	18
Hollywood	21

How would you count the number of movies in each industry where the count is 20 or more?

33 • `SELECT industry,count(*) as industry_count FROM moviesdb.movies group by industry having industry_count>=20;`

industry	industry_count
Hollywood	21

Financial Analysis

How would you select all records from the `financials` table?

33 • `SELECT * FROM moviesdb.financials;`

movie_id	budget	revenue	unit	currency
101	1.00	12.50	Billions	INR
102	200.00	954.80	Millions	USD
103	165.00	644.80	Millions	USD
104	180.00	854.00	Millions	USD
105	250.00	670.00	Millions	USD
107	400.00	2000.00	Millions	INR
108	550.00	4000.00	Millions	INR
109	390.00	1360.00	Millions	INR
110	1.40	3.50	Billions	INR
111	25.00	70.00	Millions	USD

financials 14 x

How would you calculate the profit as `revenue - budget` for each movie in the `financials` table?

33 • `SELECT *, revenue-budget AS profit FROM moviesdb.financials;`

movie_id	budget	revenue	unit	currency	profit
101	1.00	12.50	Billions	INR	11.50
102	200.00	954.80	Millions	USD	754.80
103	165.00	644.80	Millions	USD	479.80
104	180.00	854.00	Millions	USD	674.00
105	250.00	670.00	Millions	USD	420.00
107	400.00	2000.00	Millions	INR	1600.00
108	550.00	4000.00	Millions	INR	3450.00
109	390.00	1360.00	Millions	INR	970.00
110	1.40	3.50	Billions	INR	2.10
111	25.00	70.00	Millions	USD	45.00

Result 15 x

How would you convert revenue to INR if the currency is USD?

```
33 • SELECT *,if (currency="USD",revenue*77,revenue) as revenue_inr FROM moviesdb.financials;
```

Result Grid

	movie_id	budget	revenue	unit	currency	revenue_inr
▶	101	1.00	12.50	Billions	INR	12.50
	102	200.00	954.80	Millions	USD	73519.60
	103	165.00	644.80	Millions	USD	49649.60
	104	180.00	854.00	Millions	USD	65758.00
	105	250.00	670.00	Millions	USD	51590.00
	107	400.00	2000.00	Millions	INR	2000.00
	108	550.00	4000.00	Millions	INR	4000.00
	109	390.00	1360.00	Millions	INR	1360.00
	110	1.40	3.50	Billions	INR	3.50
	111	25.00	75.00	Millions	USD	5644.40

Result 16 x

How would you normalize the revenue based on the unit (Billions, Millions, Thousands)?

```
35 • SELECT *,
36 case
37 when unit='Billions' then revenue*1000
38 when unit='Thousands' then revenue/1000
39 when unit='Millions' then revenue
40 end as rev_all from moviesdb.financials;
```

Result Grid

	movie_id	budget	revenue	unit	currency	rev_all
▶	101	1.00	12.50	Billions	INR	12500.00
	102	200.00	954.80	Millions	USD	954.80
	103	165.00	644.80	Millions	USD	644.80
	104	180.00	854.00	Millions	USD	854.00
	105	250.00	670.00	Millions	USD	670.00
	107	400.00	2000.00	Millions	INR	2000.00
	108	550.00	4000.00	Millions	INR	4000.00
	109	390.00	1360.00	Millions	INR	1360.00
	110	1.40	3.50	Billions	INR	3500.00
	111	25.00	75.00	Millions	USD	75.00

Result 17 x

