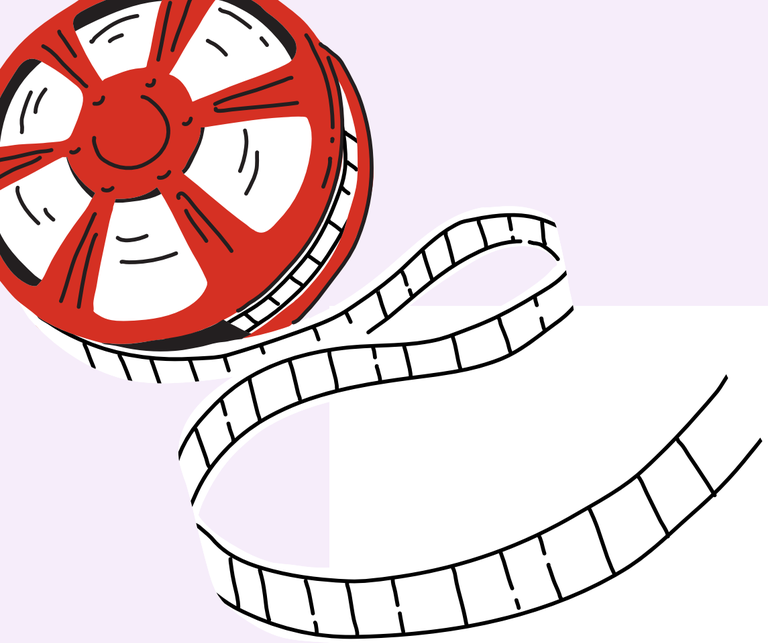
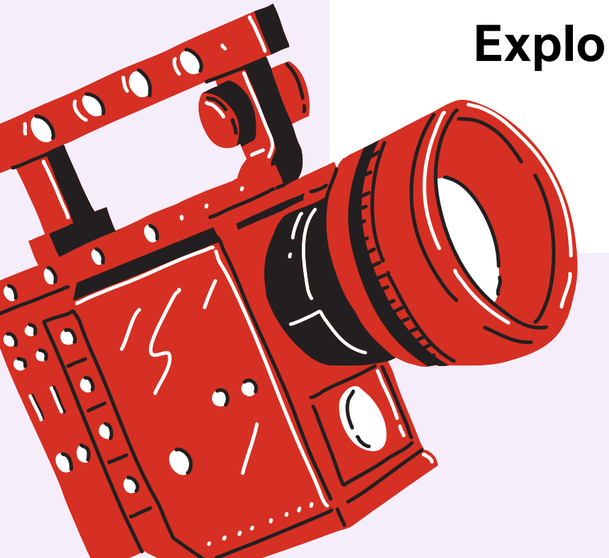


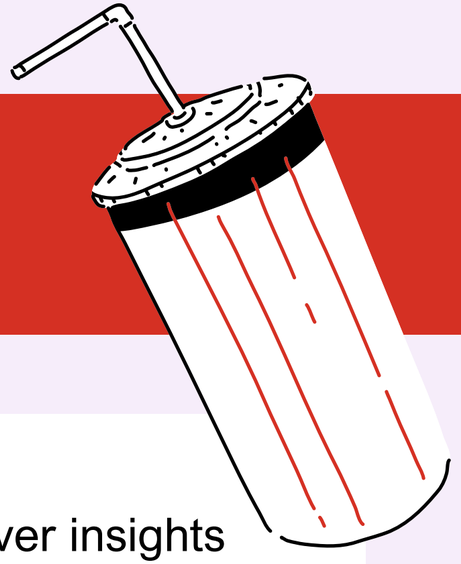
Analysing Maven movies

Exploring Business Insights Using Structured Query Language (SQL)

BY HITESH GUPTA



Introduction



Purpose of the Project:
This project was conducted to explore and analyze the Maven Movies database to uncover insights that can help drive strategic decisions related to customer behavior, rental patterns, film inventory, and business performance.

Objective:
To solve specific business problems using SQL queries and interpret the findings in a meaningful way for business stakeholders.

Dataset Overview:

- **Source:** Maven Analytics
- **Description:** The Maven Movies dataset simulates a DVD rental business. It includes tables such as film, rental, customer, payment, staff, store, and more.

Tools Used: SQL, MySQL Workbench

Project Purpose & Strategic Context

This analysis supports a potential acquisition team aiming to understand the operational health, financial risk, and customer engagement of the Maven Movies DVD rental business. Using SQL, we answered a series of questions covering core business areas.

The questions fall into these main categories:

- Staff & Store Management
- Inventory & Film Diversity
- Payments & Revenue
- Customer Behavior & Value
- Security & Compliance Risk
- Content Coverage (Awarded Actors)
- Operational Insights for Decision-Making

What the Analysis Delivers

By solving the provided business questions using SQL, the analysis delivers:

- A clear view of store-level inventory and customer distribution.
- Identification of high-value customers and revenue-driving staff.
- Insights into content strengths and gaps (e.g. award-winning actors).
- Evaluation of replacement cost liabilities.
- Data to guide fraud detection systems and email privacy risk assessment.

Q1. We will need a list of all staff members, including their first and last names, email addresses, and the store identification number where they work.

```
SELECT first_name, last_name, email, store_id
FROM staff;
```

first_name	last_name	email	store_id
Mike	Hillyer	Mike.Hillyer@sakilastaff.com	1
Jon	Stephens	Jon.Stephens@sakilastaff.com	2

Q2. We will need separate counts of inventory items held at each of your two stores.

```
SELECT COUNT(inventory_id), store_id
FROM inventory
GROUP BY store_id;
```

	count_of_inventory_id	store_id
▶	2270	1
	2311	2

Q3. We will need a count of active customers for each of your stores. Separately, please.

```sql
SELECT store_id, COUNT(customer_id) AS active_customers
FROM customer
WHERE active = 1
GROUP BY store_id;
```

store_id	active_customers
1	318
2	266

Q4. In order to assess the liability of a data breach, we will need you to provide a count of all customer email addresses stored in the database.

```
SELECT COUNT(email) AS count_of_customer_email
FROM customer;
```

count_of_customer_email
599

Q5. We are interested in how diverse your film offering is as a means of understanding how likely you are to keep customers engaged in the future. Please provide a count of unique film titles you have in inventory at each store and then provide a count of the unique categories of films you provide.

```sql
SELECT store_id, COUNT(DISTINCT film_id) AS unique_films
FROM inventory
GROUP BY store_id;
```

store_id	unique_films
1	759
2	762

```sql
SELECT COUNT(DISTINCT name) AS unique_categories
FROM category;
```

unique_categories
16

Q6. We would like to understand the replacement cost of your films. Please provide the replacement cost for the film that is least expensive to replace, the most expensive to replace, and the average of all films you carry.

```sql
SELECT MIN(replacement_cost) AS least_expensive,
    MAX(replacement_cost) AS most_expensive,
    AVG(replacement_cost) AS avg_replacement_cost
FROM film;
```

least_expensive	most_expensive	avg_replacement_cost
9.99	29.99	19.984000

Q7. We are interested in having you put payment monitoring systems and maximum payment processing restrictions in place in order to minimize the future risk of fraud by your staff. Please provide the average payment you process, as well as the maximum payment you have processed.

```
SELECT AVG(amount) AS avg_payment,
       MAX(amount) AS max_payment
FROM payment;
```

avg_payment	max_payment
4.200667	11.99

Q8. We would like to better understand what your customer base looks like. Please provide a list of all customer identification values, with a count of rentals they have made all-time, with your highest volume customers at the top of the list.

```sql
SELECT customer_id, COUNT(rental_id) AS no_of_rentals
FROM rental
GROUP BY customer_id
ORDER BY COUNT(rental_id) DESC;
```

customer_id	no_of_rentals
148	46
526	45
144	42
236	42
75	41
197	40
469	40
137	39
178	39
468	39
5	38
295	38
410	38
459	38
176	37
198	37

Q9. My partner and I want to come by each of the stores in person and meet the managers. Please send over the managers' names at each store, with the full address of each property (street address, district, city, and country please).

```
SELECT
    staff.first_name, staff.last_name,
    address.address, address.district,
    city.city, country.country
FROM
    staff
        INNER JOIN
    address ON staff.address_id = address.address_id
        INNER JOIN
    city ON city.city_id = address.city_id
        INNER JOIN
    country ON country.country_id = city.country_id;
```

first_name	last_name	address	district	city	country
Mike	Hillyer	23 Workhaven Lane	Alberta	Lethbridge	Canada
Jon	Stephens	1411 Lillydale Drive	QLD	Woodridge	Australia

Q10. I would like to get a better understanding of all of the inventory that would come along with the business. Please pull together a list of each inventory item you have stocked, including the store_id number, the inventory_id, the name of the film, the film's rating, its rental rate and replacement cost.

```
SELECT
    inventory.inventory_id, inventory.store_id,
    film.title, film.rating, film.rental_rate,
    film.replacement_cost
FROM
    inventory
        LEFT JOIN
    film ON inventory.film_id = film.film_id;
```

inventory_id	store_id	title	rating	rental_rate	replacement_cost
1	1	ACADEMY DINOSAUR	PG	0.99	20.99
2	1	ACADEMY DINOSAUR	PG	0.99	20.99
3	1	ACADEMY DINOSAUR	PG	0.99	20.99
4	1	ACADEMY DINOSAUR	PG	0.99	20.99
16	1	AFFAIR PREJUDICE	G	2.99	26.99
17	1	AFFAIR PREJUDICE	G	2.99	26.99
18	1	AFFAIR PREJUDICE	G	2.99	26.99
19	1	AFFAIR PREJUDICE	G	2.99	26.99
26	1	AGENT TRUMAN	PG	2.99	17.99
27	1	AGENT TRUMAN	PG	2.99	17.99
28	1	AGENT TRUMAN	PG	2.99	17.99
32	1	AIRPLANE SIERRA	PG-13	4.99	28.99
33	1	AIRPLANE SIERRA	PG-13	4.99	28.99
41	1	ALABAMA DEVIL	PG-13	2.99	21.99
42	1	ALABAMA DEVIL	PG-13	2.99	21.99
43	1	ALABAMA DEVIL	PG-13	2.99	21.99
46	1	ALADDIN CALENDAR	NC-17	4.99	24.99
47	1	ALADDIN CALENDAR	NC-17	4.99	24.99

Q11. From the same list of films you just pulled, please roll that data up and provide a summary level overview of your inventory. We would like to know how many inventory items you have with each rating at each store.

```sql
SELECT
    inventory.store_id,
    film.rating,
    COUNT(inventory.inventory_id) AS Count_of_inventory
FROM
    inventory
        LEFT JOIN
    film ON inventory.film_id = film.film_id
GROUP BY inventory.store_id , film.rating;
```

store_id	rating	Count_of_inventory
1	PG	444
1	G	394
1	PG-13	525
1	NC-17	465
1	R	442
2	PG	480
2	G	397
2	NC-17	479
2	PG-13	493
2	R	462

Q12. Similarly, we want to understand how diversified the inventory is in terms of replacement cost. We want to see how big of a hit it would be if a certain category of film became unpopular at a certain store. We would like to see the number of films, as well as the average replacement cost, and total replacement cost, sliced by store and film category.

```sql
SELECT
    store_id,
    category.name AS Category,
    COUNT(inventory.inventory_id) AS films,
    AVG(film.replacement_cost) AS avg_replacement_cost,
    SUM(film.replacement_cost) AS total_replacement_cost
FROM
    inventory
        LEFT JOIN
    film ON film.film_id = inventory.film_id
        LEFT JOIN
    film_category ON film_category.film_id = film.film_id
        LEFT JOIN
    category ON category.category_id = film_category.category_id
GROUP BY store_id , category.name
ORDER BY SUM(film.replacement_cost) DESC;
```

store_id	Category	films	avg_replacement_cost	total_replacement_cost
2	Sports	181	20.697182	3746.19
1	Action	169	21.191183	3581.31
1	Drama	162	21.934444	3553.38
2	Animation	174	19.995747	3479.26
2	Documentary	164	20.544878	3369.36
1	Sports	163	20.578957	3354.37
2	Sci-Fi	163	20.493067	3340.37
1	Animation	161	20.387516	3282.39
1	Sci-Fi	149	21.795369	3247.51
1	Family	157	20.537771	3224.43
2	Action	143	21.500490	3074.57
2	Games	148	20.773784	3074.52
2	Family	153	19.512876	2985.47
2	Drama	138	21.461014	2961.62
2	Classics	139	21.292158	2959.61
1	New	148	19.267027	2851.52
1	Foreign	153	18.558627	2839.47
1	Comedy	142	19.440704	2760.58
2	Foreign	147	18.636259	2739.53
2	Children	140	19.504286	2730.60

Q13. We want to make sure you folks have a good handle on who your customers are. Please provide a list of all customer names, which store they go to, whether or not they are currently active, and their full addresses – street address, city, and country.

```sql
SELECT
    CONCAT(first_name, ' ', last_name) AS full_name,
    store_id,
    active,
    address.address AS street_address,
    city.city,
    country.country
FROM
    customer
        LEFT JOIN
    address ON customer.address_id = address.address_id
        LEFT JOIN
    city ON address.city_id = city.city_id
        LEFT JOIN
    country ON city.country_id = country.country_id;
```

full_name	store_id	active	street_address	city	country
MARY SMITH	1	1	1913 Hanoi Way	Sasebo	Japan
PATRICIA JOHNSON	1	1	1121 Loja Avenue	San Bernardino	United States
LINDA WILLIAMS	1	1	692 Joliet Street	Athenai	Greece
BARBARA JONES	2	1	1566 Inegl Manor	Myingyan	Myanmar
ELIZABETH BROWN	1	1	53 Idfu Parkway	Nantou	Taiwan
JENNIFER DAVIS	2	1	1795 Santiago de Compostela Way	Laredo	United States
MARIA MILLER	1	1	900 Santiago de Compostela Parkway	Kragujevac	Yugoslavia
SUSAN WILSON	2	1	478 Joliet Way	Hamilton	New Zealand
MARGARET MOORE	2	1	613 Korolev Drive	Masqat	Oman
DOROTHY TAYLOR	1	1	1531 Sal Drive	Esfahan	Iran
LISA ANDERSON	2	1	1542 Tarlac Parkway	Sagamihara	Japan
NANCY THOMAS	1	1	808 Bhopal Manor	Yamuna Nagar	India
KAREN JACKSON	2	1	270 Amroha Parkway	Osmaniye	Turkey
BETTY WHITE	2	1	770 Bydgoszcz Avenue	Citrus Heights	United States
HELEN HARRIS	1	1	419 Iligan Lane	Bhopal	India
SANDRA MARTIN	2	0	360 Toulouse Parkway	Southend-on-...	United Kingdom
DONNA THOMPSON	1	1	270 Toulon Boulevard	Elista	Russian Fed...
CAROL GARCIA	2	1	320 Brest Avenue	Kaduna	Nigeria
RUTH MARTINEZ	1	1	1417 Lancaster Avenue	Kimberley	South Africa
SHARON ROBINSON	2	1	1688 Okara Way	Mardan	Pakistan
MICHELLE CLARK	1	1	262 A Corua (La Corua) Parkway	Tangail	Bangladesh

Q14. We would like to understand how much your customers are spending with you, and also to know who your most valuable customers are. Please pull together a list of customer names, their total lifetime rentals, and the sum of all payments you have collected from them. It would be great to see this ordered on total lifetime value, with the most valuable customers at the top of the list.

```sql
SELECT
    CONCAT(first_name, ' ', last_name) AS full_name,
    COUNT(rental.rental_id) AS total_rentals,
    SUM(payment.amount) AS total_amount
FROM
    customer
        LEFT JOIN
    rental ON customer.customer_id = rental.customer_id
        LEFT JOIN
    payment ON rental.rental_id = payment.rental_id
GROUP BY full_name
ORDER BY SUM(payment.amount) DESC;
```

full_name	total_rentals	total_amount
KARL SEAL	45	221.55
ELEANOR HUNT	46	216.54
CLARA SHAW	42	195.58
RHONDA KENNEDY	39	194.61
MARION SNYDER	39	194.61
TOMMY COLLAZO	38	186.62
WESLEY BULL	40	177.60
TIM CARY	39	175.61
MARCIA DEAN	42	175.58
ANA BRADLEY	34	174.66
JUNE CARROLL	37	173.63
DIANE COLLINS	35	169.65
LENA JENSEN	32	168.68
ARNOLD HAVENS	33	167.67
CURTIS IRBY	38	167.62
MIKE WAY	35	166.65
DAISY BATES	38	162.62
TONYA CHAPMAN	32	161.68
LOUIS LEONE	35	161.65

Q15. My partner and I would like to get to know your board of advisors and any current investors. Could you please provide a list of advisor and investor names in one table? Could you please note whether they are an investor or an advisor, and for the investors, it would be good to include which company they work with.

```
SELECT
    'investor' AS TYPE, first_name, last_name, company_name
FROM
    investor
UNION SELECT
    'advisor' AS TYPE, first_name, last_name, NULL
FROM
    advisor;
```

TYPE	first_name	last_name	company_name
investor	Montgomery	Burns	Springfield Syndicators
investor	Anthony	Stark	Iron Investors
investor	William	Wonka	Chocolate Ventures
advisor	Barry	Beenthere	NULL
advisor	Cindy	Smartypants	NULL
advisor	Mary	Moneybags	NULL
advisor	Walter	White	NULL

Q16. We're interested in how well you have covered the most-awarded actors. Of all the actors with three types of awards, for what % of them do we carry a film? And how about for actors with two types of awards? Same questions. Finally, how about actors with just one award?

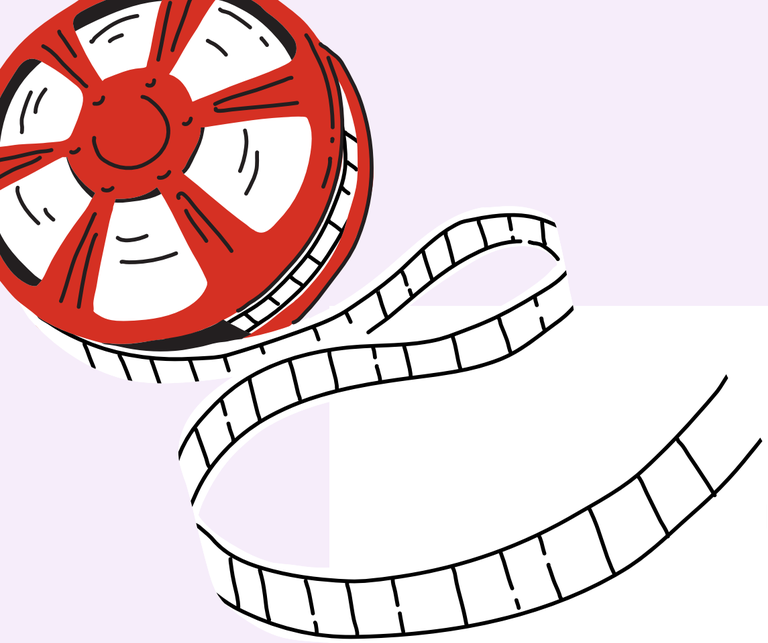
```sql
SELECT
    CASE
        WHEN actor_award.awards = 'Emmy, Oscar, Tony ' THEN '3 awards'
        WHEN actor_award.awards IN ('Emmy, Oscar','Emmy, Tony', 'Oscar, Tony') THEN '2 awards'
        ELSE '1 award'
    END AS number_of_awards,
    AVG(CASE WHEN actor_award.actor_id IS NULL THEN 0 ELSE 1 END) AS pct_w_one_film

FROM actor_award

GROUP BY
    CASE
        WHEN actor_award.awards = 'Emmy, Oscar, Tony ' THEN '3 awards'
        WHEN actor_award.awards IN ('Emmy, Oscar','Emmy, Tony', 'Oscar, Tony') THEN '2 awards'
        ELSE '1 award'
    END
```

number_of_awards	pct_w_one_film
3 awards	0.5714
2 awards	0.9242
1 award	0.8333

Conclusion

This analysis demonstrates how SQL can be effectively used to explore real business datasets and derive data-driven insights. By answering targeted questions, we identified key customer trends, performance metrics, and business opportunities that can guide decision-making in a movie rental business environment.

The project also reinforces the importance of data storytelling—presenting technical results in a format that stakeholders can act on.



Thank You!

Appreciate your time and attention.

BY HITESH GUPTA