

SQL Data Analysis Project

By

Yash Shahane

Objective:-

THE SITUATION

You and your business partner were recently approached by another local business owner who is interested in purchasing Maven Movies. He primarily owns restaurants and bars, so he has lots of questions for you about your business and the rental business in general. His offer seems very generous, so you are going to entertain his questions.

THE OBJECTIVE

Use MySQL to:

Leverage your SQL skills to extract and analyze data from various tables in the Maven Movies database to answer your potential Acquirer's questions. Each question will require you to write a multi-table SQL query, joining at least two tables.

**1. 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

```
first_name, last_name, address, district, city, country
from
    store
left join staff on store.manager_staff_id=staff.staff_id
left join address on staff.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;
```

Result Grid Filter Rows: Search Export:

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

2. 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
    store_id, inventory_id, title, rating, rental_rate, replacement_cost
from
    inventory
left join film on inventory.film_id=film.film_id;
```

store_id	inventory_id	title	rating	rental_rate	replacement_c...
1	1	ACADEMY DINOSAUR	PG	0.99	20.99
1	2	ACADEMY DINOSAUR	PG	0.99	20.99
1	3	ACADEMY DINOSAUR	PG	0.99	20.99
1	4	ACADEMY DINOSAUR	PG	0.99	20.99
1	16	AFFAIR PREJUDICE	G	2.99	26.99
1	17	AFFAIR PREJUDICE	G	2.99	26.99
1	18	AFFAIR PREJUDICE	G	2.99	26.99
1	19	AFFAIR PREJUDICE	G	2.99	26.99
1	26	AGENT TRUMAN	PG	2.99	17.99
1	27	AGENT TRUMAN	PG	2.99	17.99
1	28	AGENT TRUMAN	PG	2.99	17.99
1	32	AIRPLANE SIERRA	PG-13	4.99	28.99
1	33	AIRPLANE SIERRA	PG-13	4.99	28.99
1	41	ALABAMA DEVIL	PG-13	2.99	21.99
1	42	ALABAMA DEVIL	PG-13	2.99	21.99
1	43	ALABAMA DEVIL	PG-13	2.99	21.99
1	46	ALADDIN CALENDAR	NC-17	4.99	24.99
1	47	ALADDIN CALENDAR	NC-17	4.99	24.99
1	48	ALADDIN CALENDAR	NC-17	4.99	24.99
1	49	ALADDIN CALENDAR	NC-17	4.99	24.99
1	53	ALAMO VIDEOTAPE	G	0.99	16.99
1	54	ALAMO VIDEOTAPE	G	0.99	16.99
1	55	ALAMO VIDEOTAPE	G	0.99	16.99
1	56	ALAMO VIDEOTAPE	G	0.99	16.99

3. 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.

```
Select
    store_id, rating,
    Count(inventory_id) as No_of_inventory_items
from
    inventory
inner join film on inventory.film_id=film.film_id
group by
    store_id, rating
order by
    store_id;
```

Result Grid			Filter Rows:	Search	Export:
store_id	rating	No_of_inventory_ite...			
1	G	394			
1	PG	444			
1	PG-13	525			
1	R	442			
1	NC-17	465			
2	G	397			
2	PG	480			
2	PG-13	493			
2	R	462			
2	NC-17	479			

4. 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.

```
Select
    store_id,category.name as film_category,
    count(film.film_id) as no_of_films,
    avg(replacement_cost) as Avg_replacement_cost,
    sum(replacement_cost) as sum_replacement_cost
From
    Film
    left join inventory on film.film_id = inventory.film_id
    left join film_category on inventory.film_id = film_category.film_id
    left join category on film_category.category_id=category.category_id
group by
    store_id, category.name
order by
    sum_replacement_cost desc;
```

Result Grid



Filter Rows:



Search

Export:



	store_id	film_catego...	no_of_films	Avg_replacement_c...	sum_replacement_c...	
	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	
	1	Games	128	21.130625	2704.72	
	1	Classics	131	20.615954	2700.69	
	1	Documentary	130	20.728462	2694.70	
	2	Horror	136	19.563529	2660.64	

5. 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.

```
Select
```

```
    first_name, last_name, store_id, active, address, city , 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;
```

Result Grid Filter Rows: Search Export:

first_name	last_name	store_id	active	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	ANDERS...	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 King...
DONNA	THOMPS...	1	1	270 Toulon Boulevard	Elista	Russian Fe...
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
LAURA	RODRIG...	1	1	28 Charlotte Amalie Street	Sal	Morocco
SARAH	LEWIS	2	1	1780 Hino Boulevard	Liepaja	Latvia
KIMBERLY	LEE	2	1	96 Tafuna Way	Crdoba	Argentina

6. 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.

```
Select
    first_name, last_name,
    count(rental.rental_id) as Total_lifetime_rentals,
    sum(amount) as Total_payment_collected
From customer
left join rental on customer.customer_id = rental.customer_id
left join payment on rental.rental_id = payment.rental_id
group by
    first_name, last_name
order by
    Total_payment_collected desc;
```

Result Grid Filter Rows: Search Export:

	first_name	last_name	Total_lifetime_rent...	Total_payment_collec...
	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
	GORDON	ALLARD	32	160.68
	BRITTANY	RILEY	28	159.72
	GUY	BROWN...	32	159.68
	WARREN	SHERROD	33	159.67
	MARGIE	WADE	36	159.64

7. 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 as company_name
from
    advisor;
```

Result Grid				
Filter Rows: <input type="text"/> Search Export:				
	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

8. 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?

```
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;
```

Result Grid Filter Rows: Search Export:

number_of_awards	pct_w_one_film
3 awards	0.5714
2 awards	0.9242
1 award	0.8333