

Restaurant & Consumer Data Analysis using MySQL





Project Statement

- ❖ The project focuses on analyzing restaurant and consumer data using MySQL.
- ❖ Build a structured relational database for restaurant and consumer data
- ❖ Perform data exploration and analysis using SQL queries
- ❖ Derive insights on cuisines, ratings, and consumer behavior
- ❖ Aimed to understand customer preferences and restaurant performance to help make data-driven decisions for better service and business growth.
- ❖ Demonstrate end-to-end SQL workflow — from schema design to advanced analysis



Schema Explanation

ENTITIES :

Consumers: Info like age, budget, occupation, and city.

Restaurants: Info like city, price, alcohol service, parking.

Ratings: Overall, food, and service ratings by each consumer.

Consumer_Preferences: Cuisines liked by each consumer.

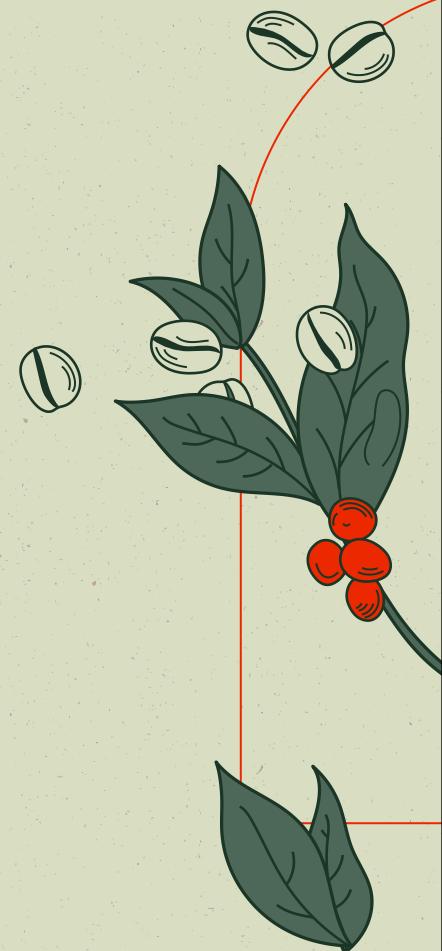
Restaurant_Cuisines: Cuisines served by each restaurant.

Relationships:

One consumer can rate multiple restaurants.

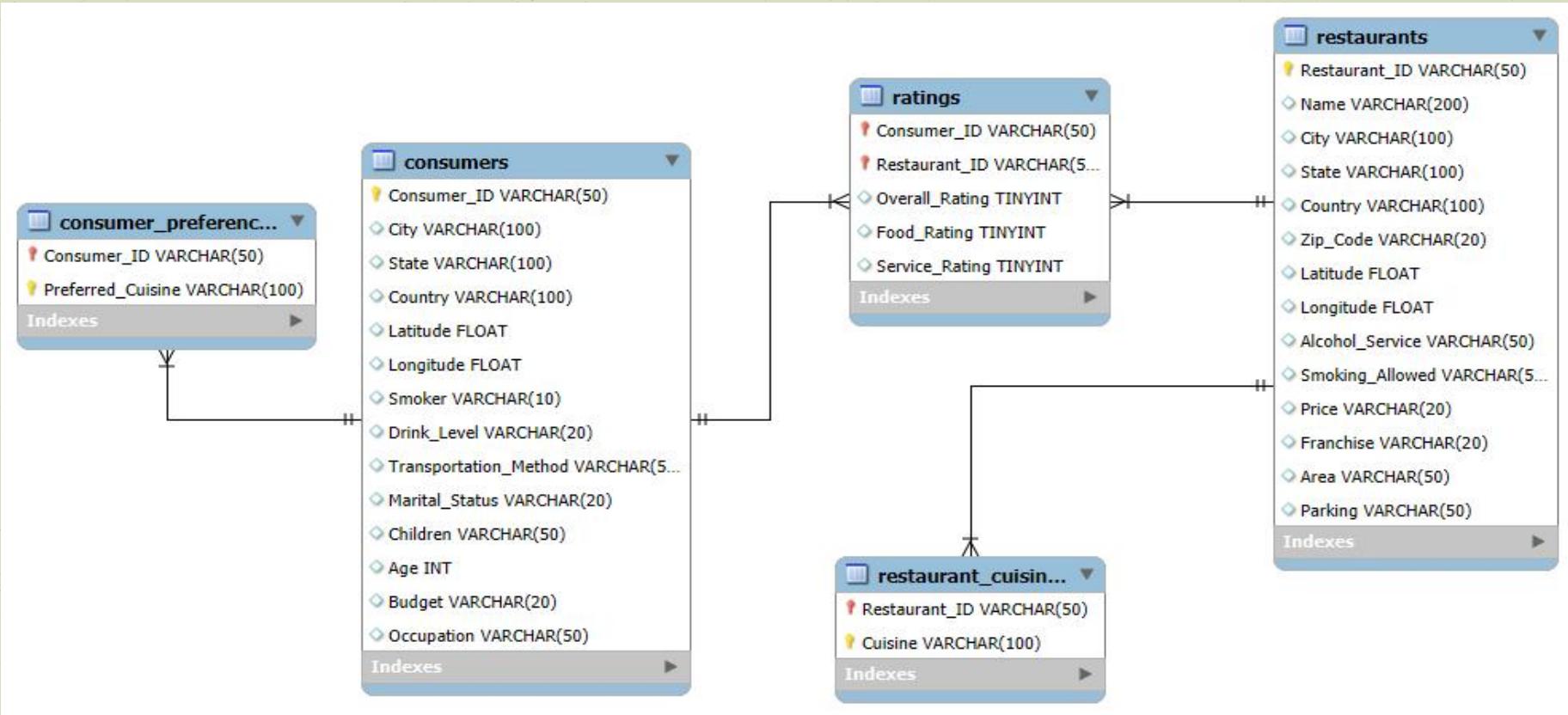
One restaurant can have multiple cuisines.

One consumer can prefer multiple cuisines.





ER Diagram





Key analysis questions (use cases)

City-wise Performance:

Which cities have restaurants with the highest average ratings?

Student Preferences:

What are the top 3 cuisines preferred by low-budget students?

Restaurant Insights:

Which restaurants perform above the overall average rating?

Consumer Behavior:

Which consumers give higher or lower ratings compared to others?

Budget vs Rating:

How does the consumer's budget (Low, Medium, High) affect their restaurant ratings?

Service & Satisfaction:

Does alcohol service ("Wine & Beer") influence overall satisfaction?

Cuisine Popularity:

Which cuisines are most popular and highly rated across all cities?

Performance Comparison:

How do individual consumer ratings compare to each restaurant's average rating?

WHERE Clause :



1. List all details of consumers who live in the city of 'Cuernavaca'

SELECT * FROM CONSUMERS

WHERE CITY = 'CUERNAVACA' ;

	Consumer_ID	City	State	Country	Latitude	Longitude	Smoker	Drink_Level	Transportation_Method	Marital_Status	Children	Age	Budget	Occupation
▶	U1004	Cuernavaca	Morelos	Mexico	18.867	-99.183	No	Abstemious	Public	Single	Independent	72	Medium	Employed
	U1012	Cuernavaca	Morelos	Mexico	18.8133	-99.2437	No	Casual Drinker	Public	Single	Independent	24	Medium	Student
	U1017	Cuernavaca	Morelos	Mexico	18.9526	-99.2016	No	Casual Drinker	Public	Single	Independent	21	Medium	Employed
	U1020	Cuernavaca	Morelos	Mexico	18.8782	-99.223	No	Abstemious	Public	Single	Independent	30	Medium	Employed
	U1030	Cuernavaca	Morelos	Mexico	18.8448	-99.1828	No	Casual Drinker	On Foot	Single	Independent	21	Medium	Student
	U1035	Cuernavaca	Morelos	Mexico	18.8397	-99.2239	Yes	Casual Drinker	Car	Single	Independent	27	Medium	Student
	U1040	Cuernavaca	Morelos	Mexico	18.8952	-99.1804	No	Abstemious	Public	Single	Independent	18	Medium	Student
	U1041	Cuernavaca	Morelos	Mexico	18.9352	-99.2362	No	Abstemious	On Foot	Single	Independent	27	Medium	Employed

2. Find the Consumer_ID, Age, and Occupation of all consumers who are 'Students' AND are 'Smokers'.

```
SELECT Consumer_ID ,Age,Occupation FROM Consumers  
WHERE Occupation = 'Student' AND Smoker = 'Yes';
```

	Consumer_ID	Age	Occupation
▶	U1006	23	Student
	U1015	23	Student
	U1018	23	Student
	U1027	23	Student
	U1029	23	Student
	U1035	27	Student
	U1038	82	Student
	U1048	24	Student
	U1050	22	Student



3. List the Name, City, Alcohol_Service, and Price of all restaurants that serve 'Wine & Beer' and have a 'Medium' price level.

```
SELECT Name ,City,Alcohol_Service,Price FROM Restaurants  
WHERE Alcohol_Service = 'Wine & Beer' AND Price = 'Medium';
```

	Name	City	Alcohol_Service	Price
▶	Mikasa	Cuernavaca	Wine & Beer	Medium
	Restaurant Familiar El Chino	Jiutepec	Wine & Beer	Medium
	Sirlone	San Luis Potosi	Wine & Beer	Medium
	El Pueblito	San Luis Potosi	Wine & Beer	Medium
	Chilis Cuernavaca	Cuernavaca	Wine & Beer	Medium
	Sanborns Casa Piedra	Cuernavaca	Wine & Beer	Medium
	Vips	Cuernavaca	Wine & Beer	Medium
	Restaurant Orizatlán	San Luis Potosi	Wine & Beer	Medium





4. Find the names and cities of all restaurants that are part of a 'Franchise'.

```
SELECT Name,City from Restaurants  
where Franchise = 'Yes';
```

	Name	City
▶	McDonalds Centro	Cuernavaca
	Gorditas Doña Tota	Ciudad Victoria
	Hamburguesas La Perica	Ciudad Victoria
	Pollo Frito Buenos Aires	Ciudad Victoria
	Carnitas Mata	Ciudad Victoria
	La Perica Hamburguesa	Ciudad Victoria
	Little Pizza Emilio Portes Gil	Ciudad Victoria
	Gorditas Dona Tota	Ciudad Victoria

5 . Show the Consumer_ID, Restaurant_ID, and Overall_Rating for all ratings where the Overall_Rating was 'Highly Satisfactory' (which corresponds to a value of 2, according to the data dictionary).

```
SELECT Consumer_ID ,Restaurant_Id ,Overall_Rating  from Ratings  
WHERE Overall_Rating = 2;
```

	Consumer_ID	Restaurant_Id	Overall_Rating
▶	U1001	132825	2
	U1001	135025	2
	U1002	132825	2
	U1002	132862	2
	U1002	132921	2
	U1002	135041	2
	U1003	132723	2
	U1003	132754	2

JOINS AND SUBQUERIES :



1. List the names and cities of all restaurants that have an Overall_Rating of 2 (Highly Satisfactory) from at least one consumer.

```
SELECT R.Name ,R.City, T.Overall_Rating FROM Restaurants R  
JOIN Ratings T on R.Restaurant_ID = T.Restaurant_ID  
WHERE t.Overall_rating = 2;
```

	Name	City	Overall_Rating
▶	Puesto De Tacos	San Luis Potosi	2
	El Rincon De San Francisco	San Luis Potosi	2
	Puesto De Tacos	San Luis Potosi	2
	La Posada Del Virrey	San Luis Potosi	2
	Crudalia	San Luis Potosi	2
	Luna Cafe	San Luis Potosi	2
	Gordas De Morales	San Luis Potosi	2
	Cabana Huasteca	San Luis Potosi	2



2. Find the Consumer_ID and Age of consumers who have rated restaurants located in 'San Luis Potosi'.

```
SELECT C.Consumer_ID ,C.Age FROM Consumers C  
JOIN Ratings R ON C.Consumer_ID =R.Consumer_ID  
JOIN Restaurants T ON R.Restaurant_ID =T.Restaurant_ID  
WHERE T.City='San Luis Potosi';
```



	Consumer_ID	Age
▶	U1006	23
	U1007	23
	U1013	30
	U1033	21
	U1046	22
	U1055	60
	U1061	22
	U1073	23



3. List the names of restaurants that serve 'Mexican' cuisine and have been rated by consumer 'U1001'.

```
SELECT R.NAME FROM Restaurants R  
JOIN Ratings T ON R.Restaurant_ID = T.Restaurant_ID  
JOIN Restaurant_Cuisines C ON R.Restaurant_ID = C.Restaurant_ID  
WHERE C.Cuisine = "Mexican" AND T.Consumer_ID = 'U1001';
```



NAME
Puesto De Tacos
El Rincon De San Francisco



4. Find all details of consumers who prefer 'American' cuisine AND have a 'Medium' budget.

```
SELECT * FROM Consumers C  
JOIN Consumer_Preferences P ON C.Consumer_ID = P.Consumer_ID  
WHERE P.Preferred_Cuisine = 'American' AND C.Budget='Medium';
```

	Consumer_ID	City	State	Country	Latitude	Longitude	Smoker	Drink_Level	Transportation_Method	Marital_Status	Children	Age	Budget	Occupation	Consumer_ID	Preferred_Cuisine
▶	U1001	San Luis Potosi	San Luis Potosi	Mexico	22.14	-100.979	No	Abstemious	On Foot	Single	Independent	23	Medium	Student	U1001	American
	U1005	San Luis Potosi	San Luis Potosi	Mexico	22.1835	-100.96	No	Abstemious	Public	Single	Independent	20	Medium	Student	U1005	American
	U1016	San Luis Potosi	San Luis Potosi	Mexico	22.1562	-100.977	No	Casual Drinker	On Foot	Single	Independent	21	Medium	Student	U1016	American
	U1047	San Luis Potosi	San Luis Potosi	Mexico	22.1424	-100.949	No	Abstemious	Public	Single	Independent	20	Medium	Student	U1047	American
	U1057	San Luis Potosi	San Luis Potosi	Mexico	22.1966	-100.912	Yes	Casual Drinker	Public	Single	Independent	23	Medium	Student	U1057	American
	U1060	Ciudad Victoria	Tamaulipas	Mexico	23.7152	-99.1589	No	Casual Drinker	Public	Single	Independent	21	Medium	Student	U1060	American
	U1073	San Luis Potosi	San Luis Potosi	Mexico	22.15	-100.983	No	Abstemious	Car	Single	Kids	23	Medium	Unemployed	U1073	American
	U1108	San Luis Potosi	San Luis Potosi	Mexico	22.1435	-100.988	No	Abstemious	Public	Single	Independent	29	Medium	Student	U1108	American



5. List restaurants (Name, City) that have received a Food_Rating lower than the average Food_Rating across all rated restaurants.

```
SELECT R.Name, R.City FROM Restaurants R  
JOIN Ratings T ON R.Restaurant_ID =T.Restaurant_ID  
GROUP BY R.Restaurant_ID, R.Name, R.City  
HAVING AVG(T.Food_Rating) < (SELECT AVG(Food_Rating) FROM Ratings);
```

	NAME	City
▶	Puesto de Gorditas	Ciudad Victoria
	Cafe Ambar	Ciudad Victoria
	Cafe Chaires	San Luis Potosi
	McDonalds Centro	Cuernavaca
	Tacos De Barbacoa Enfrente Del Tec	Ciudad Victoria
	Hamburguesas La Perica	Ciudad Victoria
	Pollo Frito Buenos Aires	Ciudad Victoria
	La Perica Hamburguesa	Ciudad Victoria

6. Find consumers (Consumer_ID, Age, Occupation) who have rated at least one restaurant but have NOT rated any restaurant that serves 'Italian' cuisine.

```
SELECT c.Consumer_ID, c.Age, c.Occupation  
FROM Consumers c  
WHERE EXISTS (SELECT 1 FROM Ratings rt WHERE rt.Consumer_ID=c.Consumer_ID)  
AND NOT EXISTS (  
    SELECT 1  
    FROM Ratings rt  
    JOIN Restaurant_Cuisines rc ON rt.Restaurant_ID = rc.Restaurant_ID  
    WHERE rt.Consumer_ID=c.Consumer_ID AND rc.Cuisine='Italian'  
);
```

	Consumer_ID	Age	Occupation
▶	U1001	23	Student
	U1002	22	Student
	U1003	23	Student
	U1004	72	Employed
	U1005	20	Student
	U1006	23	Student
	U1007	23	Student
	U1008	23	Student



7. List restaurants (Name) that have received ratings from consumers older than 30.

```
SELECT R.NAME FROM Restaurants R  
JOIN RatingS T ON R.Restaurant_ID =T.Restaurant_ID  
JOIN Consumers C ON C.Consumer_ID = T.Consumer_ID  
WHERE C.Age >30;
```

NAME
La Parroquia
Emilianos
Tacos Los Volcanes
La Virreina
Cafeteria Y Restaurant El Pacifico
Restaurante Marisco Sam
Restaurante El Cielo Potosino
El RincÃ³n De San Francisco

8. Find the Consumer_ID and Occupation of consumers whose preferred cuisine is 'Mexican' and who have given an Overall_Rating of 0 to at least one restaurant (any restaurant).

```
SELECT C.Consumer_ID ,C.Occupation FROM Consumers C
JOIN Consumer_Preferences P ON C.Consumer_ID = P.Consumer_ID
JOIN Ratings R ON C.Consumer_ID =R.Consumer_ID
WHERE P.Preferred_Cuisine = 'Mexican' AND R.Overall_Rating < 1;
```

	Consumer_ID	Occupation
▶	U1003	Student
	U1003	Student
	U1009	Student
	U1010	Student
	U1010	Student
	U1015	Student
	U1015	Student
	U1015	Student



9. List the names and cities of restaurants that serve 'Pizzeria' cuisine and are located in a city where at least one 'Student' consumer lives.

```
SELECT R.NAME ,R.City FROM Restaurants R  
JOIN Restaurant_Cuisines C ON R.Restaurant_ID = C.Restaurant_ID  
JOIN Ratings T ON R.Restaurant_ID = T.Restaurant_ID  
JOIN Consumers S ON T.Consumer_ID = S.Consumer_ID  
WHERE C.Cuisine = 'Pizzeria' AND R.City=S.City AND S.Occupation ='Student';
```

	NAME	City
▶	Pizza Clasica	San Luis Potosi
	Restaurante Tiberius	San Luis Potosi
	Dominos Pizza	San Luis Potosi
	Restaurante Tiberius	San Luis Potosi
	Restaurante Tiberius	San Luis Potosi
	La Fontana Pizza Restaurante And Cafe	San Luis Potosi
	Dominos Pizza	San Luis Potosi
	Pizza Clasica	San Luis Potosi



10 .Find consumers (Consumer_ID, Age) who are 'Social Drinkers' and have rated a restaurant that has 'No' parking.

```
SELECT C.Consumer_ID , C.Age FROM Consumers C  
JOIN Ratings T ON C.Consumer_ID = T.Consumer_ID  
JOIN Restaurants R ON T.Restaurant_ID = R.Restaurant_ID  
WHERE C.Drink_Level = 'Social Drinkers' AND  
R.Parking = 'none';
```

	Consumer_ID	Age

WHERE Clause and Order of Execution :



1. List Consumer_IDs and the count of restaurants they've rated, but only for consumers who are 'Students'. Show only students who have rated more than 2 restaurants.

```
SELECT C.Consumer_ID ,  
       COUNT(DISTINCT T.Restaurant_ID) AS 'COUNT OF RESTAURANT RATED'  
FROM Consumers C  
JOIN Ratings T ON C.Consumer_ID =T.Consumer_ID  
WHERE C.Occupation = 'Student'  
GROUP BY C.Consumer_ID  
HAVING COUNT(DISTINCT T.Restaurant_ID)>2;
```

	Consumer_ID	COUNT OF RESTAURANT RATED
▶	U1001	9
	U1002	10
	U1003	13
	U1005	9
	U1006	11
	U1007	9
	U1008	9
	U1009	11



2 . We want to categorize consumers by an 'Engagement_Score' which is their Age divided by 10 (integer division).

List the Consumer_ID, Age, and this calculated Engagement_Score, but only for consumers whose Engagement_Score would be exactly 2 and who use 'Public'

```
SELECT Consumer_ID ,Age ,(Age/10) AS Engagement_Score FROM Consumers  
WHERE (Age/10) = 2 AND Transportation_Method='Public';
```

	Consumer_ID	Age	Engagement_Score
▶	U1005	20	2.0000
	U1031	20	2.0000
	U1047	20	2.0000
	U1054	20	2.0000
	U1079	20	2.0000
	U1119	20	2.0000
	U1127	20	2.0000

3. For each restaurant, calculate its average Overall_Rating. Then, list the restaurant Name, City, and its calculated average Overall_Rating, but only for restaurants located in 'Cuernavaca' AND whose calculated average Overall_Rating is greater than 1.0.



```
SELECT R.NAME ,R.City ,
AVG(T.Overall_Rating) AS AVG_RATING FROM Restaurants R
JOIN Ratings T ON R.Restaurant_ID = T.Restaurant_ID
WHERE R.City = 'Cuernavaca'
GROUP BY R.NAME
HAVING AVG(T.Overall_Rating) > 1.0;
```

	NAME	City	AVG_RATING
▶	Mariscos Tia Licha	Cuernavaca	1.6000
	El Cotorreo	Cuernavaca	1.5000
	Rincon Del Bife	Cuernavaca	1.6667
	Log Yin	Cuernavaca	1.7500
	Restaurant Las Mañanitas	Cuernavaca	2.0000
	Sanborns Casa Piedra	Cuernavaca	1.5556
	Kiku Cuernavaca	Cuernavaca	1.6000
	Restaurant Los Pinos	Cuernavaca	1.2500



4. Find consumers (Consumer_ID, Age) who are 'Married' and whose Food_Rating for any restaurant is equal to their Service_Rating for that same restaurant, but only consider ratings where the Overall_Rating was 2.

```
SELECT C.Consumer_ID ,C.Age FROM Consumers C
JOIN Ratings T ON C.Consumer_ID = T.Consumer_ID
WHERE C.Marital_Status = 'Married'
AND T.Food_Rating = T.Service_Rating
AND T.Overall_Rating = 2;
```

	Consumer_ID	Age
▶	U1010	25
	U1044	43
	U1044	43
	U1044	43
	U1055	60



5. List Consumer_ID, Age, and the Name of any restaurant they rated, but only for consumers who are 'Employed' and have given a Food_Rating of 0 to at least one restaurant located in 'Ciudad Victoria'.

```
SELECT C.Consumer_ID , C.Age ,R.NAME FROM Consumers C
JOIN Ratings T ON C.Consumer_ID = T.Consumer_ID
JOIN Restaurants R ON T.Restaurant_ID = R.Restaurant_ID
WHERE C.Occupation = 'Employed'
AND T.Food_Rating = 0 AND R.City='Ciudad Victoria';
```

	Consumer_ID	Age	NAME
▶	U1103	23	Tacos Abi
	U1103	23	Taqueria El Amigo

Advanced SQL Concepts:

1. Using a CTE, find all consumers who live in 'San Luis Potosi'. Then, list their Consumer_ID, Age, and the Name of any Mexican restaurant they have rated with an Overall_Rating of 2.

```
③ WITH MEX_REST AS (SELECT Consumer_ID , Age  
- FROM Consumers WHERE City = 'San Luis Potosi')  
  
SELECT C.Consumer_ID, C.Age ,R.NAME AS REST_NAME ,RC.Cuisine FROM MEX_REST C  
JOIN Ratings T ON C.Consumer_ID = T.Consumer_ID  
JOIN Restaurants R ON T.Restaurant_ID = R.Restaurant_ID  
JOIN Restaurant_Cuisines RC ON R.Restaurant_ID = RC.Restaurant_ID  
WHERE RC.Cuisine='Mexican' AND T.Overall_Rating= 2;
```

Consumer_ID	Age	REST_NAME	Cuisine
U1001	23	Puesto De Tacos	Mexican
U1001	23	El Rincon De San Francisco	Mexican
U1002	22	Puesto De Tacos	Mexican
U1003	23	Gordas De Morales	Mexican
U1003	23	Cabana Huasteca	Mexican
U1003	23	La Estrella De Dimas	Mexican
U1003	23	Puesto De Tacos	Mexican
U1009	21	Gorditas Doa Gloria	Mexican



2. For each Occupation, find the average age of consumers. Only consider consumers who have made at least one rating. (Use a derived table to get consumers who have rated).

```
SELECT c.Occupation, AVG(c.Age) AS Avg_Age  
FROM Consumers c  
JOIN (  
    SELECT DISTINCT Consumer_ID  
    FROM Ratings  
) rated ON c.Consumer_ID = rated.Consumer_ID  
GROUP BY c.Occupation;
```

Occupation	Avg_Age
Student	26.0250
Employed	37.5625
Unemployed	21.5000

3 . Using a CTE to get all ratings for restaurants in 'Cuernavaca', rank these ratings within each restaurant based on Overall_Rating (highest first). Display Restaurant_ID, Consumer_ID, Overall_Rating, and the RatingRank.

```
WITH MY_RATING AS (
    SELECT R.Restaurant_ID ,T.Consumer_ID ,T.Overall_Rating ,R.City
    FROM Ratings T
    JOIN Restaurants R ON T.Restaurant_ID = R.Restaurant_ID
)
SELECT Restaurant_ID,
Consumer_ID,
Overall_Rating,
ROW_NUMBER() OVER (PARTITION BY Restaurant_ID ORDER BY Overall_Rating DESC) AS RatingRank
FROM MY_RATING;
```

Restaurant_ID	Consumer_ID	Overall_Rating	RatingRank
132560	U1087	1	1
132560	U1067	1	2
132560	U1082	0	3
132560	U1050	0	4
132561	U1026	1	1
132561	U1065	1	2
132561	U1129	1	3
132561	U1130	0	4

4. For each rating, show the Consumer_ID, Restaurant_ID, Overall_Rating, and also display the average Overall_Rating given by that specific consumer across all their ratings.

```
SELECT T.Consumer_ID ,  
       T.Restaurant_ID,  
       T.Overall_Rating ,  
       AVG(T.Overall_Rating)  
OVER(PARTITION BY T.Consumer_ID) AS AVG_RATING  
FROM Ratings T;
```

Consumer_ID	Restaurant_ID	Overall_Rating	AVG_RATING
U1001	132825	2	1.1111
U1001	132830	1	1.1111
U1001	135025	2	1.1111
U1001	135033	1	1.1111
U1001	135039	1	1.1111
U1001	135040	1	1.1111
U1001	135045	1	1.1111
U1001	135051	1	1.1111

5. Using a CTE, identify students who have a 'Low' budget. Then, for each of these students, list their top 3 most preferred cuisines based on the order they appear in the Consumer_Preferences table (assuming no explicit preference order, use Consumer_ID, Preferred_Cuisine to define order for ROW_NUMBER).

```
WITH low_budget_students AS (
    SELECT Consumer_ID
    FROM Consumers
    WHERE Occupation = 'Student' AND Budget = 'Low'),
    prefs_ordered AS (
        SELECT cp.Consumer_ID, cp.Preferred_Cuisine,
        ROW_NUMBER() OVER (PARTITION BY cp.Consumer_ID ORDER BY cp.Preferred_Cuisine) AS rn
        FROM Consumer_Preferences cp
        WHERE cp.Consumer_ID IN (SELECT Consumer_ID FROM low_budget_students)
    )
    SELECT Consumer_ID, Preferred_Cuisine
    FROM prefs_ordered
    WHERE rn <= 3
    ORDER BY Consumer_ID, rn;
```

Consumer_ID	Preferred_Cuisine
U1002	Mexican
U1003	Mexican
U1007	Family
U1008	American
U1008	Cafeteria
U1008	Coffee Shop
U1018	Mexican
U1023	Mexican
U1027	Family

6. Consider all ratings made by 'Consumer_ID' = 'U1008'. For each rating, show the Restaurant_ID, Overall_Rating, and the Overall_Rating of the next restaurant they rated (if any), ordered by Restaurant_ID (as a proxy for time if rating time isn't available). Use a derived table to filter for the consumer's ratings first.

```
WITH u1008_ratings AS (
    SELECT Restaurant_ID, Overall_Rating,
    LEAD(Overall_Rating)
    OVER (ORDER BY Restaurant_ID) AS Next_Overall_Rating
    FROM Ratings
    WHERE Consumer_ID = 'U1008'
)
SELECT * FROM u1008_ratings;
```

Restaurant_ID	Overall_Rating	Next_Overall_Rating
132846	1	1
132851	1	2
132854	2	1
132858	1	1
132866	1	1
132869	1	1
135054	1	1
135069	1	1
135108	1	NULL



7. Create a VIEW named HighlyRatedMexicanRestaurants that shows the Restaurant_ID, Name and City of all Mexican restaurants that have an average Overall_Rating greater than 1.5.

```
CREATE VIEW HIGH_RATE_MAX_REST AS
SELECT r.Restaurant_ID, r.Name, r.City
FROM Restaurants r
JOIN Restaurant_Cuisines rc ON r.Restaurant_ID = rc.Restaurant_ID
JOIN Ratings rt ON r.Restaurant_ID = rt.Restaurant_ID
WHERE rc.Cuisine = 'Mexican'
GROUP BY r.Restaurant_ID, r.Name, r.City
HAVING AVG(rt.Overall_Rating) > 1.5;
SELECT * FROM HIGH_RATE_MAX_REST ;
```

Restaurant_ID	Name	City
132755	La Estrella De Dimas	San Luis Potosi
134976	Log Yin	Cuernavaca
135025	El Rincon De San Francisco	San Luis Potosi
135028	La Virreina	San Luis Potosi
135055	La Cochinita Pibil Restaurante Yucateco	San Luis Potosi

8. First, ensure the `HighlyRatedMexicanRestaurants` view from Q7 exists. Then, using a CTE to find consumers who prefer 'Mexican' cuisine, list those consumers (`Consumer_ID`) who have not rated any restaurant listed in the `HighlyRatedMexicanRestaurants` view.

```
WITH mexican_pref_consumers AS (
    SELECT DISTINCT Consumer_ID FROM Consumer_Preferences
    WHERE Preferred_Cuisine = 'Mexican')
    SELECT mpc.Consumer_ID
    FROM mexican_pref_consumers mpc
    WHERE NOT EXISTS ( SELECT 1
        FROM Ratings rt
        JOIN HighlyRatedMexicanRestaurants hrm ON rt.Restaurant_ID = hrm.Restaurant_ID
        WHERE rt.Consumer_ID = mpc.Consumer_ID
    );
    
```

Consumer_ID
U1002
U1006
U1008
U1009
U1010

9. Create a stored procedure GetRestaurantRatingsAboveThreshold that accepts a Restaurant_ID and a minimum Overall_Rating as input. It should return the Consumer_ID, Overall_Rating, Food_Rating, and Service_Rating for that restaurant where the Overall_Rating meets or exceeds the threshold.

```
DROP PROCEDURE IF EXISTS GetRestaurantRatingsAboveThreshold;
DELIMITER $$

CREATE PROCEDURE GetRestaurantRatingsAboveThreshold(
    IN p_restaurant_id VARCHAR(50),
    IN p_min_overall TINYINT
)
BEGIN
    SELECT Consumer_ID, Overall_Rating, Food_Rating, Service_Rating
    FROM Ratings
    WHERE Restaurant_ID = p_restaurant_id
        AND Overall_Rating >= p_min_overall;
END $$

DELIMITER ;

CALL GetRestaurantRatingsAboveThreshold('132560', 1);
-- enter the Restaurant _ id

SELECT DISTINCT Restaurant_ID FROM Ratings LIMIT 10;
```

	Consumer_ID	Overall_Rating	Food_Rating	Service_Rating
▶	U1067	1	0	0
	U1087	1	2	1

Restaurant_ID
132560
132561
132564
132572
132583

10. Identify the top 2 highest-rated (by Overall_Rating) restaurants for each cuisine type. If there are ties in rating, include all tied restaurants. Display Cuisine, Restaurant_Name, City, and Overall_Rating.

```
WITH cuisine_rest_avg AS (
    SELECT rc.Cuisine, r.Restaurant_ID, r.Name, r.City,
    AVG(rt.Overall_Rating) AS avg_overall
    FROM Restaurant_Cuisines rc
    JOIN Restaurants r ON rc.Restaurant_ID = r.Restaurant_ID
    JOIN Ratings rt ON r.Restaurant_ID = rt.Restaurant_ID
    GROUP BY rc.Cuisine, r.Restaurant_ID, r.Name, r.City
),
ranked AS (
    SELECT cr.*,
    RANK() OVER (PARTITION BY Cuisine ORDER BY avg_overall DESC) AS rnk
    FROM cuisine_rest_avg cr
)
SELECT Cuisine,
Name AS Restaurant_Name, City, avg_overall AS Overall_Rating
FROM ranked
WHERE rnk <= 2
ORDER BY Cuisine, Overall_Rating DESC;
```

Cuisine	Restaurant_Name	City	Overall_Rating
American	Tacos Los Volcanes	San Luis Potosi	1.6667
American	KFC	San Luis Potosi	1.4286
Armenian	Little Pizza Emilio Portes Gil	Ciudad Victoria	1.2500
Bakery	Chaires	San Luis Potosi	1.4000
Bar	Restaurant Bar Hacienda Los Martinez	San Luis Potosi	1.6667
Bar	Restaurante Bar El Gallinero	San Luis Potosi	1.5000

11. First, create a VIEW named ConsumerAverageRatings that lists Consumer_ID and their average Overall_Rating. Then, using this view and a CTE, find the top 5 consumers by their average overall rating. For these top 5 consumers, list their Consumer_ID, their average rating, and the number of 'Mexican' restaurants they have rated.

```
CREATE VIEW ConsumerAverageRatings AS
SELECT Consumer_ID, AVG(Overall_Rating) AS avg_overall
FROM Ratings
GROUP BY Consumer_ID;

WITH top5 AS (
    SELECT Consumer_ID, avg_overall
    FROM ConsumerAverageRatings
    ORDER BY avg_overall DESC
    LIMIT 5)
SELECT t.Consumer_ID, t.avg_overall,
    COUNT(DISTINCT CASE WHEN rc.Cuisine = 'Mexican' THEN r.Restaurant_ID END) AS mexican_restaurants Rated
FROM top5 t
LEFT JOIN Ratings rt ON t.Consumer_ID = rt.Consumer_ID
LEFT JOIN Restaurants r ON rt.Restaurant_ID = r.Restaurant_ID
LEFT JOIN Restaurant_Cuisines rc ON r.Restaurant_ID = rc.Restaurant_ID
GROUP BY t.Consumer_ID, t.avg_overall
ORDER BY t.avg_overall DESC;

SELECT * FROM ConsumerAverageRatings;
```

Consumer_ID	avg_overall	mexican_restaurantsRated
U1021	2.0000	3
U1074	2.0000	1
U1100	2.0000	1
U1102	2.0000	0
U1107	2.0000	1



12. Create a stored procedure named **Get ConsumerSegmentAndRestaurantPerformance** that accepts a **Consumer_ID** as input.

The procedure should:

1. Determine the consumer's "Spending Segment" based on their Budget:

'Low' -> 'Budget Conscious'

'Medium' -> 'Moderate Spender'

'High' -> 'Premium Spender'

NULL or other -> 'Unknown Budget'

2. For all restaurants rated by this consumer:

List the **Restaurant_Name**.

The **Overall_Rating** given by this consumer.

The average **Overall_Rating** this restaurant has received from all consumers(not just the input consumer).

A "Performance_Flag" indicating if the input consumer's rating for that restaurant is 'Above Average', 'At Average', or 'Below Average' compared to the restaurant's overall average rating. Rank these restaurants for the input consumer based on the **Overall_Rating** they gave (highest rating = rank 1).

```
DROP PROCEDURE IF EXISTS GetConsumerSegmentAndRestaurantPerformance;
DELIMITER $$

CREATE PROCEDURE GetConsumerSegmentAndRestaurantPerformance(IN p_consumer_id VARCHAR(50))
BEGIN
    SET @budget := (
        SELECT Budget FROM Consumers WHERE Consumer_ID = p_consumer_id LIMIT 1
    );
    SET @segment := CASE
        WHEN @budget = 'Low' THEN 'Budget Conscious'
        WHEN @budget = 'Medium' THEN 'Moderate Spender'
        WHEN @budget = 'High' THEN 'Premium Spender'
        ELSE 'Unknown Budget'
    END;
    WITH rated AS (
        SELECT rt.Restaurant_ID, r.Name AS Restaurant_Name, rt.Overall_Rating AS Consumer_Overall_Rating
        FROM Ratings rt
        JOIN Restaurants r ON rt.Restaurant_ID = r.Restaurant_ID
        WHERE rt.Consumer_ID = p_consumer_id
    ),
    rest_avg AS (
        SELECT Restaurant_ID, AVG(Overall_Rating) AS Restaurant_Avg_Overall
        FROM Ratings
    )
    SELECT @segment, r.Name AS Restaurant_Name, rt.Consumer_Overall_Rating, r.Avg_Rating
    FROM rated rt
    JOIN Restaurants r ON rt.Restaurant_ID = r.Restaurant_ID
    WHERE rt.Consumer_ID = p_consumer_id
    ORDER BY rt.Consumer_Overall_Rating DESC;
```

```
SELECT Restaurant_ID, AVG(Overall_Rating) AS Restaurant_Avg_Overall
FROM Ratings
GROUP BY Restaurant_ID
),
combined AS (
SELECT rd.Restaurant_ID, rd.Restaurant_Name, rd.Consumer_Overall_Rating, ra.Restaurant_Avg_Overall
FROM rated rd
JOIN rest_avg ra ON rd.Restaurant_ID = ra.Restaurant_ID
),
ranked AS (
    SELECT *,ROW_NUMBER() OVER (ORDER BY Consumer_Overall_Rating DESC) AS Consumer_Rank,
    CASE
        WHEN Consumer_Overall_Rating > Restaurant_Avg_Overall THEN 'Above Average'
        WHEN Consumer_Overall_Rating = Restaurant_Avg_Overall THEN 'At Average'
        ELSE 'Below Average'
    END AS Performance_Flag
    FROM combined)
SELECT @segment AS Spending_Segment,Restaurant_Name,
Consumer_Overall_Rating,ROUND(Restaurant_Avg_Overall,2) AS Restaurant_Avg_Overall,
Performance_Flag,Consumer_Rank
FROM ranked
ORDER BY Consumer_Rank;
END $$

DELIMITER ;

SELECT Consumer_ID, Budget FROM Consumers LIMIT 10;
CALL GetConsumerSegmentAndRestaurantPerformance('U1001');

-- enter the Consumer_id
```



Spending_Segment	Restaurant_Name	Consumer_Overall_Rating	Restaurant_Avg_Overall	Performance_Flag	Consumer_Rank
Moderate Spender	Puesto De Tacos	2	1.28	Above Average	1
Moderate Spender	El Rincon De San Francisco	2	1.67	Above Average	2
Moderate Spender	Rincon Huasteco	1	0.92	Above Average	3
Moderate Spender	Restaurant El Muladar De Calzada	1	0.75	Above Average	4
Moderate Spender	Restaurant De Mariscos De Picon	1	1.25	Below Average	5
Moderate Spender	Restaurant Los Compadres	1	0.25	Above Average	6
Moderate Spender	Restaurante La Gran Via	1	1.46	Below Average	7
Moderate Spender	Restaurante Versalles	1	1.43	Below Average	8
Moderate Spender	Tortas La Casa Hacienda	0	1.22	Below Average	9

Consumer_ID	Budget
U1001	Medium
U1002	Low
U1003	Low
U1004	Medium
U1005	Medium
U1006	Medium
U1007	Low
U1008	Low
U1009	Medium
U1010	Medium

CONCLUSION

This project helped us understand how SQL can find useful insights from restaurant and consumer data. We learned to design databases, write queries, and analyze real data. It improved our practical skills in SQL and data analysis for solving real business problems.

Any
QUERIES ?



Thank You...!

