**A REPORT ON  A RESTAURANT DATABASE**

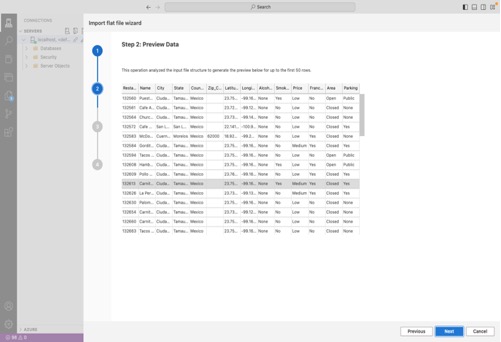
**INTRODUCTION**

This report shows how data was imported and also the T-SQL analysis for four related tables, which I applied FK\_constraints to link together. These tables are Restaurant, Consumer, Ratings and Restaurant\_Cuisine. This import involved downloading four separate files as named above and importing them into SQL server tables, and subsequently into the Restaurant Database I had created, followed by the T-SQL analysis to solve the questions that were listed and also gain insight from the datasets.

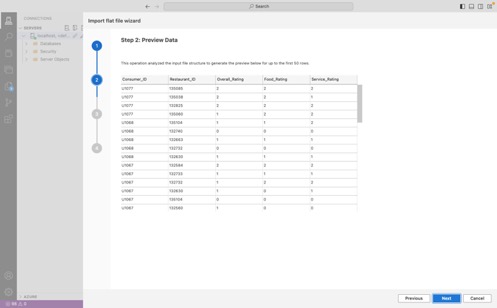
Data Import Steps

This import consists of multiple steps

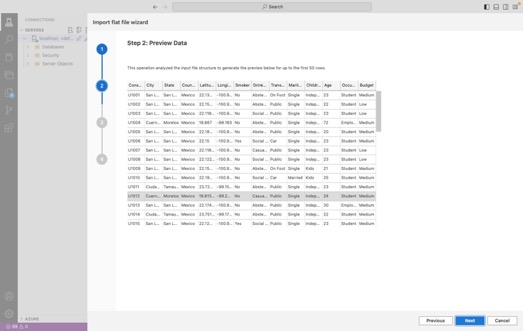
Steps A:



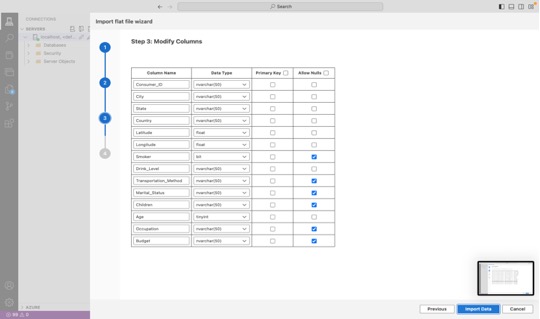
Restaurant table



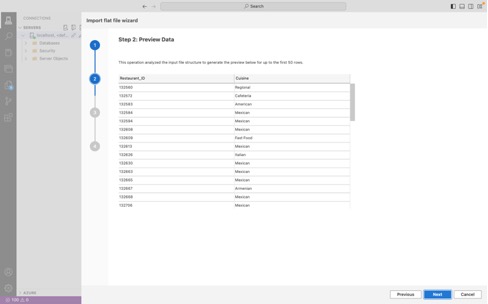
Ratings table



Consumer table



Consumer table



Restaurant\_Cuisine Table

After clicking on import, I reviewed the structure and contents of the csv file  to  understand the data schema and point out primary and foreign key also choose data type and applied constraint were needed, also identify potential issues.

Using the Import File Wizard in Azure, I imported the data from each CSV file into its respective table. I mapped the columns from the CSV files to the corresponding columns in the database tables and specified appropriate data types.

**Validation**

After importing the data, I performed validation checks to ensure data integrity and consistency. This included verifying primary key constraints, checking for NULL values, and examining data distributions.

**T-SQL Analysis**

After importing the data, I conducted T-SQL analysis to gain insights. Below are some example T-SQL statements along with their explanations.

The query counts the total number of restaurants per country from the Restaurants table and orders the results in descending order based on the total number of restaurants.

SELECT

    Country,

    COUNT(\*) AS TotalRestaurant

FROM

    Restaurants

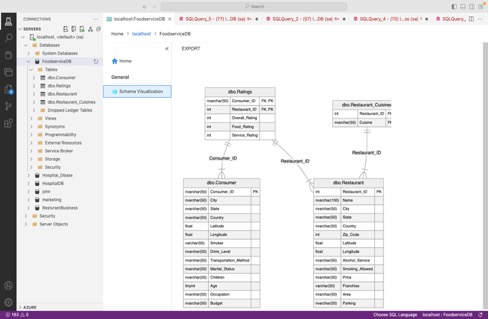
GROUP BY

    Country

ORDER BY

    TotalRestaurants DESC;

Below is a schema visualization for the relationship shared between the tables



**These tables store information about restaurants, consumers, ratings, and cuisines, respectively. Understanding their relationships is crucial for performing meaningful data analysis and ensuring data integrity.**

The dbo.Restaurant table holds detailed information about each restaurant. The Restaurant\_ID column is the primary key, ensuring that each restaurant entry is unique

The dbo.Ratings table contains ratings given by consumers to restaurants. The Consumer\_ID and Restaurant\_ID columns form a composite primary key, indicating that each rating entry is unique per consumer-restaurant pair. These columns are also foreign keys, linking to the dbo.Consumer and dbo.Restaurant tables, respectively.

The dbo.Consumer table contains information about consumers. The Consumer\_ID column is the primary key, ensuring that each consumer entry is unique.

The dbo.Restaurant\_Cuisines table stores the types of cuisines each restaurant offers. The Restaurant\_ID and Cuisine columns form a composite primary key. The Restaurant\_ID column is also a foreign key, linking to the dbo.Restaurant table.

**Relationships Between Tables**

1.  **dbo.Restaurant and dbo.Ratings**:

o   **Relationship**: One-to-Many

o   **Explanation**: Each restaurant can have multiple ratings, but each rating entry corresponds to a unique consumer-restaurant pair. This relationship is enforced by the foreign key Restaurant\_ID in the dbo.Ratings table referencing the Restaurant\_ID in the dbo.Restaurant table.

o

2.  **dbo.Consumer and dbo.Ratings**:

o   **Relationship**: One-to-Many

o   **Explanation**: Each consumer can rate multiple restaurants, but each rating entry corresponds to a unique consumer-restaurant pair. This relationship is enforced by the foreign key Consumer\_ID in the dbo.Ratings table referencing the Consumer\_ID in the dbo.Consumer table.

o

3.  **dbo.Restaurant and dbo.Restaurant\_Cuisines**:

o   **Relationship**: One-to-Many

o   **Explanation**: Each restaurant can offer multiple types of cuisines, but each cuisine entry corresponds to a unique restaurant-cuisine pair. This relationship is enforced by the foreign key Restaurant\_ID in the dbo.Restaurant\_Cuisines table referencing the Restaurant\_ID in the dbo.Restaurant table.

Normalization has been achieved, step by step process to reduce the degree of data redundancy. Normalization will improve the performance of the delta operation (aka. DML operation); UPDATE, INSERT, SELECT.

**Overview of DataSet**

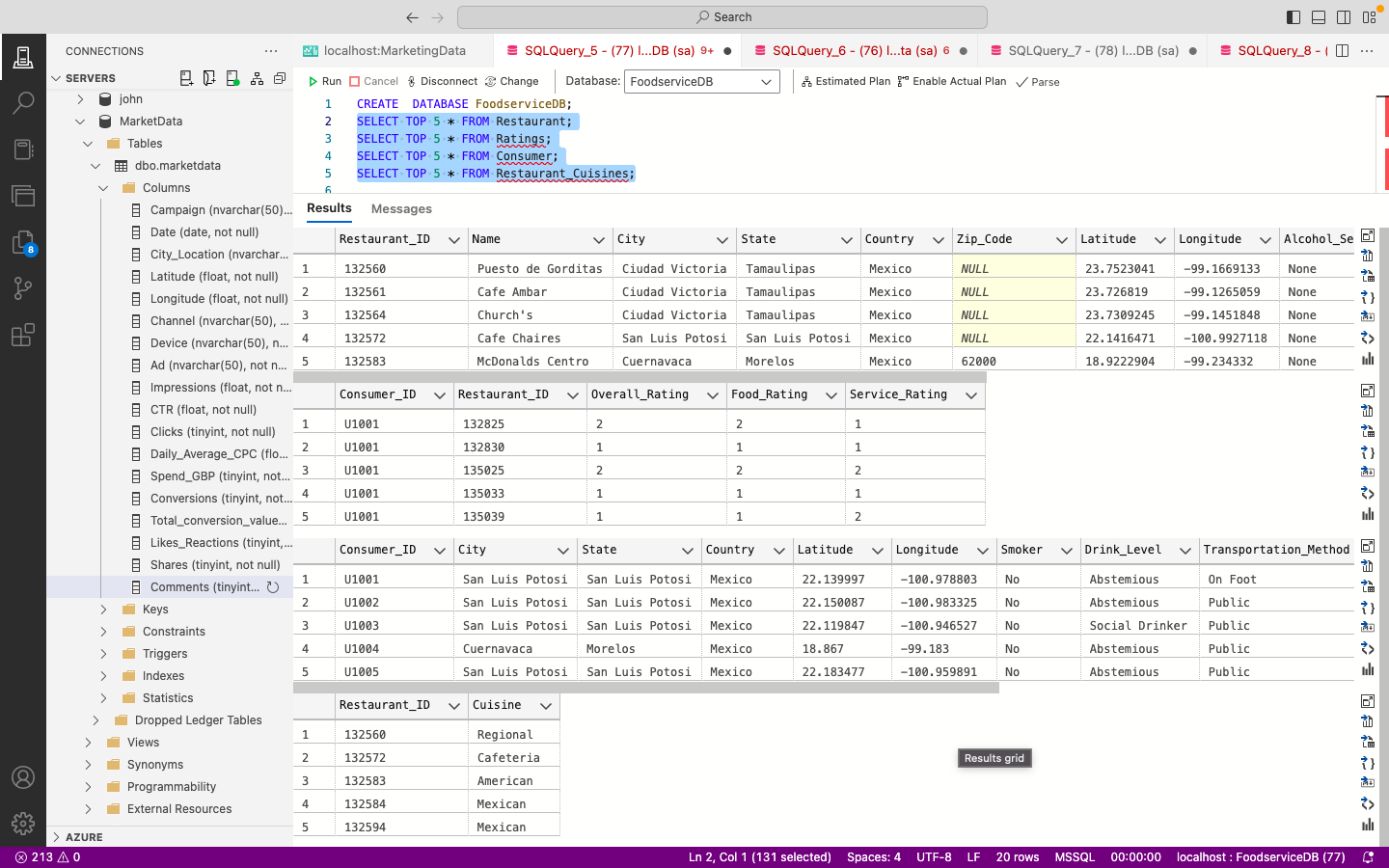
--SELECTING TOP5 --

SELECT TOP 5 \* FROM Restaurant;

SELECT TOP 5 \* FROM Ratings;

SELECT TOP 5 \* FROM Consumer;

SELECT TOP 5 \* FROM Restaurant\_Cuisines;

****

Overview of dataset

**Adding Constraint**

-- Adding primary and foreign key constraints --

ALTER TABLE Ratings

ADD CONSTRAINT FK\_Ratings\_Consumer

FOREIGN KEY (Consumer\_id) REFERENCES Consumer(Consumer\_id);

ALTER TABLE Ratings

ADD CONSTRAINT FK\_Ratings\_Restaurant

FOREIGN KEY (Restaurant\_id) REFERENCES Restaurant(Restaurant\_id);

ALTER TABLE Restaurant\_Cuisines

ADD CONSTRAINT FK\_Cuisines\_Restaurant

FOREIGN KEY (Restaurant\_id) REFERENCES Restaurant(Restaurant\_id);

**Results**

**1) Write a query that lists all restaurants with a medium range price with open area, serving Mexican food?**

SELECT R.Restaurant\_ID,RC.Cuisine,R.Price,R.Area

FROM Restaurant R

JOIN Restaurant\_Cuisines RC

ON R.Restaurant\_ID = RC.Restaurant\_ID

WHERE Price = 'Medium' AND Area = 'Open' AND Cuisine = 'Mexican' ;

|  |  |  |  |
| --- | --- | --- | --- |
| **Restaurant\_ID** | **Cuisine** | **Price** | **Area** |
| 135018 | Mexican | Medium | Open |
| 135106 | Mexican | Medium | Open |

**2) Write a query that returns the total number of restaurants who have the overall rating as 1 and are serving Mexican food. Compare the results with the total number of restaurants who have the overall rating as 1 serving Italian food please give explanations on their comparison.**

-- for Mexican food --

SELECT COUNT (\*) AS Total, Cuisine

FROM Restaurant\_Cuisines RC

JOIN Ratings RS

ON RC.Restaurant\_ID = RS.Restaurant\_ID

WHERE Overall\_Rating = '1' AND Cuisine = 'Mexican'

GROUP BY Cuisine

UNION ALL

-- For Italian food --

SELECT COUNT (\*) AS Total, Cuisine

FROM Restaurant\_Cuisines RC

JOIN Ratings RS

ON RC.Restaurant\_ID = RS.Restaurant\_ID

WHERE Overall\_Rating = '1' AND Cuisine = 'Italian'

GROUP BY Cuisine;

|  |  |
| --- | --- |
| **Total** | **Cuisine** |
| 87 | Mexican |
| 11 | Italian |

**Analysis and Comparison**

* **Quantity Comparison:** There are significantly more Mexican restaurants (87) with a low rating compared to Italian restaurants (11).
* **Possible Explanations:**
  + **Number of Restaurants:** There might be a higher total number of Mexican restaurants in the dataset, increasing the likelihood of more low-rated establishments.
  + **Rating Distribution:** Mexican restaurants might have a broader rating distribution, leading to more low-rated establishments.
  + **Customer Expectations and Preferences:** Cultural differences in customer expectations and preferences might result in varied ratings for Mexican versus Italian food.
  + **Quality and Consistency:** Potentially, the quality and consistency of Mexican restaurants in the dataset could be more variable, leading to more ratings of 1.

**Additional Insights**

* **Location and Demographics:** Further analysis could involve looking at the location and demographics of these restaurants to understand if geographic or demographic factors play a role in the ratings.
* **Price and Service:** Examining the price levels and service ratings could provide additional context for the overall ratings.
* **Consumer Feedback:** Analysing consumer feedback and reviews could give qualitative insights into why certain restaurants receive low ratings.

**3)Calculate the average age of consumers who have given a 0 rating to the 'Service\_rating'column.**

SELECT AVG(Age) AS AvgAge

FROM Consumer C

JOIN Ratings RS

ON C.Consumer\_ID = RS.Consumer\_ID

WHERE Service\_Rating = '0';

|  |
| --- |
| **AvgAge** |
| 26 |

**4)Write a query that returns the restaurants ranked by the youngest consumer. You**

**should include the restaurant name and food rating that is given by that customer to the restaurant in your result. Sort the results based on food rating from high to low.**

SELECT R.Name AS RestaurantName, RS.Food\_Rating, RS.Food\_Rating AS ConsumerFoodRating

FROM Restaurant R

JOIN Ratings RS ON R.Restaurant\_ID = RS.Restaurant\_ID

JOIN Consumer C ON RS.Consumer\_ID = C.Consumer\_ID

WHERE C.Age = (

SELECT MIN(Age)

FROM Consumer

)

ORDER BY RS.Food\_Rating DESC;

|  |  |  |
| --- | --- | --- |
| **RestaurantName** | **Food\_Rating** | **ConsumerAge** |
| Giovannis | 2 | 18 |
| Restaurant Bar Coty Y Pablo | 2 | 18 |
| El Cotorreo | 1 | 18 |
| Kiku Cuernavaca | 1 | 18 |

**5)Write a stored procedure for the query given as:**

**Update the Service\_rating of all restaurants to '2' if they have parking available, either**

**as 'yes' or 'public'**

CREATE PROCEDURE UpdateServiceRatingSForParking

AS

BEGIN

-- Update the Service\_rating to '2' for restaurants with parking available as 'yes' or 'public'

UPDATE Ratings

SET Service\_Rating = 2

WHERE Restaurant\_ID IN (

SELECT Restaurant\_ID

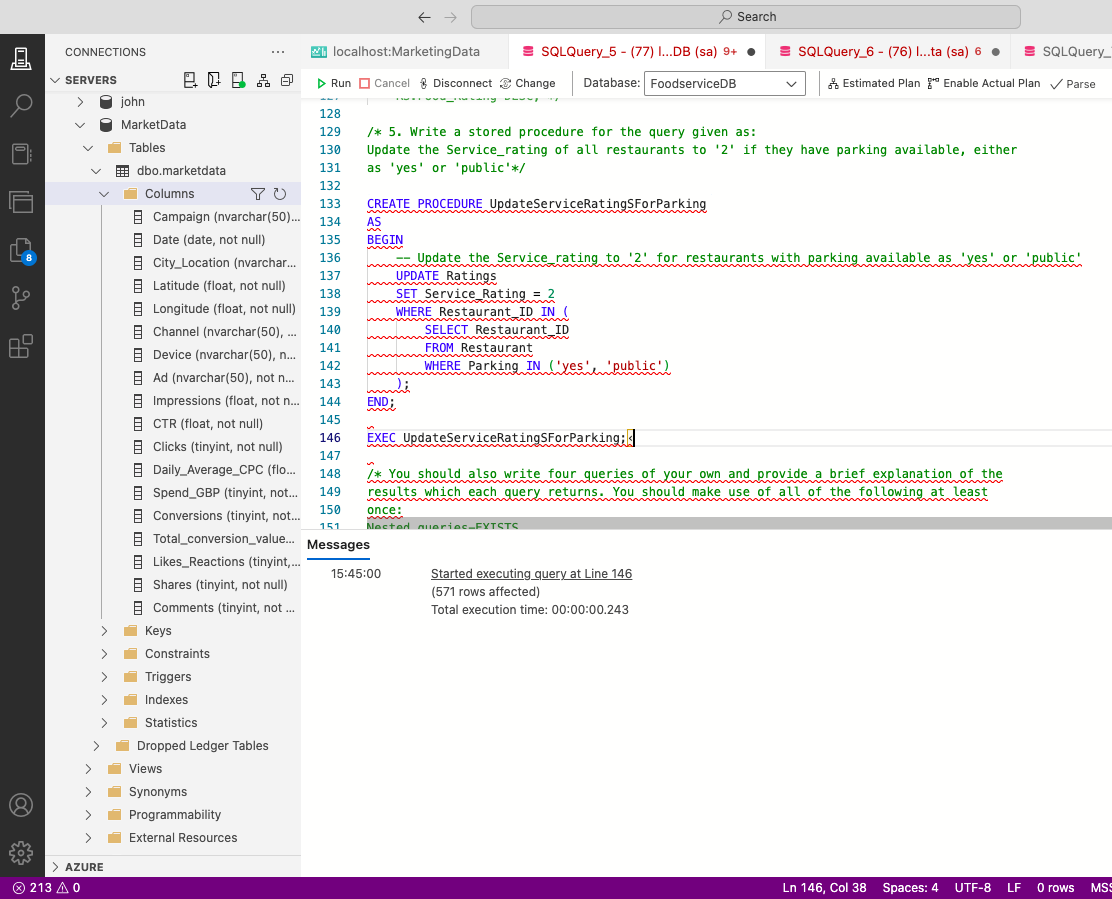
FROM Restaurant

WHERE Parking IN ('yes', 'public')

);

END;

EXEC UpdateServiceRatingSForParking;



**6.) You should also write four queries of your own and provide a brief explanation of the**

**results which each query returns. You should make use of all of the following at least**

**once:**

**Nested queries-EXISTS**

**Nested queries-IN**

**System functions**

**Use of GROUP BY, HAVING and ORDER BY clauses.**

A) Nested queries-EXITS; The query returns the consumer ID, city, and marital status of consumers who use a car for transportation and have rated a restaurant that serves Mexican cuisine.

The EXISTS clause ensures that only those consumers who meet these conditions are included in the result.

SELECT C.Consumer\_ID AS ConsumerID, C.City, C.Marital\_Status

FROM Consumer C

WHERE EXISTS (

SELECT 1

FROM Ratings R

JOIN Restaurant\_Cuisines RC ON R.Restaurant\_ID = RC.Restaurant\_ID

WHERE R.Consumer\_ID = C.Consumer\_ID

AND C.Transportation\_Method = 'Car'

AND RC.Cuisine = 'Mexican'

);

|  |  |  |
| --- | --- | --- |
| **ConsumerID** | **City** | **Marital\_Status** |
| U1006 | San Luis Potosi | Single |
| U1021 | Ciudad Victoria | Single |
| U1022 | San Luis Potosi | Single |
| U1023 | Ciudad Victoria | Single |
| U1035 | Cuernavaca | Single |
| U1045 | San Luis Potosi | Single |
| U1050 | Ciudad Victoria | Single |
| U1055 | San Luis Potosi | Married |
| U1061 | San Luis Potosi | Single |
| U1062 | San Luis Potosi | Single |
| U1071 | San Luis Potosi | Single |
| U1073 | San Luis Potosi | Single |
| U1074 | Cuernavaca | Single |
| U1075 | San Luis Potosi | Single |
| U1078 | San Luis Potosi | Single |
| U1080 | Ciudad Victoria | Single |
| U1085 | San Luis Potosi | Single |
| U1090 | San Luis Potosi | Single |
| U1091 | San Luis Potosi | Single |
| U1093 | Cuernavaca | Single |
| U1101 | San Luis Potosi | Single |
| U1104 | San Luis Potosi | Single |
| U1106 | Cuernavaca | Single |
| U1107 | Ciudad Victoria | Single |
| U1111 | San Luis Potosi | Single |
| U1112 | San Luis Potosi | Single |
| U1136 | San Luis Potosi | Single |

B) Nested queries-IN; The query returns the consumer ID, city, and marital status of consumers who use a car for transportation and have rated a restaurant with a price categorized as 'yes' or 'Low'.

The IN clause ensures that only consumers meeting these conditions are included in the result.

SELECT C.Consumer\_ID AS ConsumerID, C.City, C.Marital\_Status

FROM Consumer C

WHERE C.Consumer\_ID IN (

SELECT DISTINCT RS.Consumer\_ID

FROM Ratings RS

JOIN Restaurant R ON RS.Restaurant\_ID = R.Restaurant\_ID

WHERE R.Price IN ('yes', 'Low')

AND C.Transportation\_Method = 'Car'

|  |  |  |
| --- | --- | --- |
| **ConsumerID** | **City** | **Marital\_Status** |
| U1006 | San Luis Potosi | Single |
| U1021 | Ciudad Victoria | Single |
| U1022 | San Luis Potosi | Single |
| U1023 | Ciudad Victoria | Single |
| U1035 | Cuernavaca | Single |
| U1042 | Cuernavaca | Single |
| U1044 | Cuernavaca | Married |
| U1045 | San Luis Potosi | Single |
| U1046 | San Luis Potosi | Single |
| U1050 | Ciudad Victoria | Single |
| U1055 | San Luis Potosi | Married |
| U1061 | San Luis Potosi | Single |
| U1062 | San Luis Potosi | Single |
| U1071 | San Luis Potosi | Single |
| U1073 | San Luis Potosi | Single |
| U1074 | Cuernavaca | Single |
| U1075 | San Luis Potosi | Single |
| U1078 | San Luis Potosi | Single |
| U1080 | Ciudad Victoria | Single |
| U1085 | San Luis Potosi | Single |
| U1086 | San Luis Potosi | Single |
| U1090 | San Luis Potosi | Single |
| U1091 | San Luis Potosi | Single |
| U1093 | Cuernavaca | Single |
| U1101 | San Luis Potosi | Single |
| U1104 | San Luis Potosi | Single |
| U1106 | Cuernavaca | Single |
| U1107 | Ciudad Victoria | Single |
| U1110 | Cuernavaca | Single |
| U1111 | San Luis Potosi | Single |
| U1112 | San Luis Potosi | Single |
| U1131 | San Luis Potosi | Married |
| U1133 | Jiutepec | Single |
| U1136 | San Luis Potosi | Single |

c)System Function \* Use count to find the total number of people with marital status married

SELECT COUNT(\*) AS MarriedConsumersCount

FROM Consumer

WHERE Marital\_Status = 'Married';

|  |
| --- |
| **MarriedConsumersCount** |
| 10 |

D)Use of GROUP BY, HAVING and ORDER BY clauses; these are cities where the average age of consumers who have a

'Drink\_Level' of 'Abstemious' is less than 30, along with the corresponding average age, ordered by the average age in descending order.

 This can provide insights into which cities have relatively younger consumers who are abstemious.

 SELECT

    C.City,

    AVG(C.Age) AS AverageAge

FROM

    Consumer C

WHERE

    C.Drink\_Level = 'Abstemious'

GROUP BY

    C.City

HAVING

    AVG(C.Age) < 30

ORDER BY

    AverageAge DESC;

|  |  |
| --- | --- |
| **City** | **AverageAge** |
| Jiutepec | 24 |
| Ciudad Victoria | 22 |

**CONCLUSION**

**The analysis of restaurant ratings, particularly focusing on Mexican and Italian cuisines, reveals several insights with broader implications for restaurant owners, customers, and the food industry as a whole. The disparity in the number of low-rated restaurants between these cuisines highlights potential areas for improvement and strategic decision-making.**

**FINDINGS**

* **Higher Number of Low-Rated Mexican Restaurants:** There are 87 Mexican restaurants with an overall rating of 1 compared to only 11 Italian restaurants.
* **Customer Expectations:** Differences in customer expectations and cultural preferences may influence how ratings are assigned.
* **Quality and Consistency:** Variability in quality and consistency of Mexican restaurants might be higher.

**For Restaurant Owners and Managers:**

o   **Quality Improvement:** The high number of low-rated Mexican restaurants suggests a need for improvement in quality and consistency. Owners should focus on standardizing operations, training staff, and enhancing customer service.

o   **Customer Feedback Analysis:** Understanding the reasons behind low ratings through customer feedback can provide actionable insights. Restaurants should implement regular feedback mechanisms and act on customer suggestions.

o   **Market Positioning:** Italian restaurants with fewer low ratings might have better market positioning or meet customer expectations more consistently. Mexican restaurant owners might consider re-evaluating their market strategies to better align with customer preferences.

o   **Operational Focus:** Attention to operational details such as cleanliness, prompt service, and food quality could help in improving ratings.

**For Customers:**

* **Informed Choices:** Customers can use this information to make more informed dining choices based on ratings and reviews. They might prefer restaurants with consistently higher ratings for a better dining experience.
* **Setting Expectations:** Understanding that certain cuisines might have more variability in ratings can help set realistic expectations when choosing where to dine.

· For **the Food Industry:**

* **Benchmarking:** The analysis can serve as a benchmark for different types of cuisine. Industry players can use these insights to compare performance and identify best practices.
* **Culinary Training and Standards:** There might be a need for enhanced culinary training programs focusing on cuisines with more low-rated establishments. This can help improve overall standards and customer satisfaction.
* **Innovation and Trends:** Tracking trends in customer preferences and ratings can help the industry innovate and adapt to changing tastes. For example, introducing fusion dishes or healthier options might attract higher ratings.

· For **Researchers and Analysts:**

* **Further Studies:** The initial findings open avenues for more in-depth research into factors influencing restaurant ratings. Future studies could explore the impact of geographic location, price levels, service quality, and demographic factors on ratings.
* **Data-Driven Decisions:** Using data analytics to study consumer behavior and preferences can lead to more data-driven decisions in restaurant management and marketing strategies.

· Economic **Implications:**

* **Revenue Impact:** Low ratings can directly impact a restaurant's revenue. Restaurants with consistently low ratings might see a decrease in customer footfall, affecting their profitability.
* **Employment and Training:** Restaurants might need to invest more in staff training and development to improve service quality, which can have economic implications for the business.

· Cultural **and Social Implications:**

* **Cultural Representation:** The analysis might reflect broader cultural trends and preferences. For example, the popularity and perceived quality of different cuisines can be tied to cultural representation and acceptance.
* **Community Engagement:** Restaurants often play a key role in community engagement. Ensuring high-quality dining experiences can foster better community relationships and customer loyalty.

* **Targeting Young Consumers:** The data suggests a mixed level of satisfaction among the youngest consumers (age 18). Restaurants need to understand the preferences and expectations of this age group better. This could include menu options, pricing, ambiance, and service quality that appeal to younger patrons.

**Economic Implications:**

o   **Revenue Impact:** Restaurants with lower ratings from young consumers might see a decline in this demographic's patronage, affecting revenue. Conversely, restaurants that effectively cater to young customers can boost their business.

o   **Investment in Quality:** Improving food quality and overall dining experience may require additional investment in training, ingredients, and service, but can lead to higher ratings and increased customer loyalty.

**RECOMMENDATIONS**

* **Engage with Young Customers:** Conduct surveys or focus groups with young consumers to gather more detailed feedback on their dining preferences and experiences.
* **Quality Improvement Programs:** Implement programs focused on improving food quality and consistency. This could include staff training, better ingredient sourcing, and recipe optimization.
* **Marketing Strategies:** Develop marketing campaigns targeted at younger demographics, highlighting aspects of the restaurant that appeal to this age group, such as trendy menu items, social media presence, and youth-oriented events.