Project Title: Restaurant Data Analysis Management for Informed Decision-Making

Submission Date: 12/15/2023

Submitted By: Group 11

Group Members:

Soumith Roy Akula (02084467)

Uday Reddy Polepally (02115349)

Yashwanth Reddy Kistipati (02086419)

Advisor: Yukui Luo, PhD

Abstract

The Restaurant Data Analysis Management System is a comprehensive software solution designed to empower restaurant owners and managers with data-driven insights for informed decision-making. Recognizing the significance of operational efficiency, enhanced customer experiences, and data-driven decision-making in the competitive restaurant industry, this project focuses on creating a robust database and implementing data analysis techniques.

1. Introduction

1.1 Problem Statement

The Restaurant Data Analysis Management System is a comprehensive solution designed to address the challenges faced by restaurants in today's dynamic and competitive industry. The project aims to leverage the power of data analysis to enhance decision-making processes, streamline operations, and ultimately improve the overall performance and customer satisfaction of restaurants.

1.2 Project Objectives

Operational Efficiency Enhancement: Simplify daily restaurant operations, order handling, and menu management through effective data management and analysis.

Customer Experience Improvement: Understand customer preferences, order history, and ratings to make informed decisions for enhancing the overall customer experience.

Data-Driven Decision-Making: Facilitate data-driven decision-making by acquiring and analysing data related to menus, orders, and customer feedback.

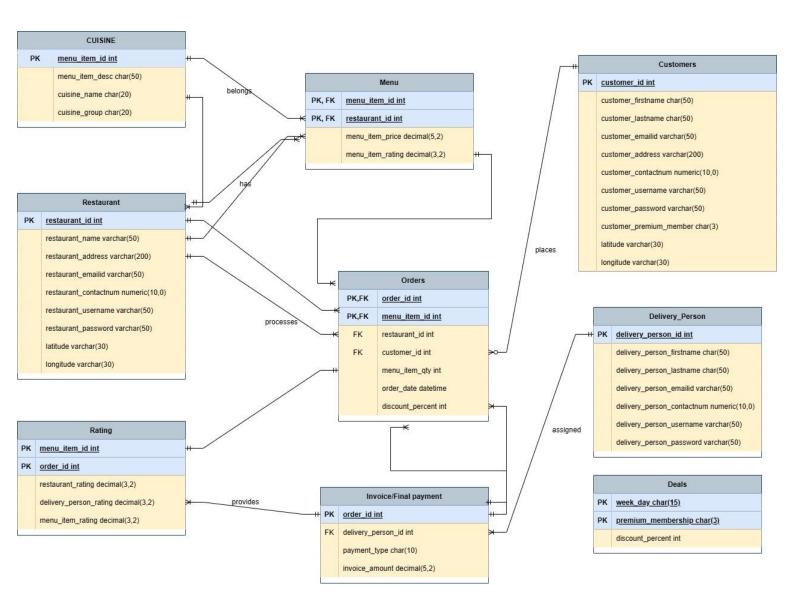
Secure User Access Control: Establish a secure user management system to control access to the database, ensuring data integrity and confidentiality.

Improved Customer Service: Utilize data analysis to identify the most frequently ordered menu items and customer preferences, leading to tailored offerings and improved customer service.

Efficient Order Management: Enable real-time tracking and management of incoming orders for restaurant owners, ensuring efficient order processing.

2. Database Design

2.1 Entity-Relationship Diagram



Relationships:

- **a.** Restaurant-Cuisine (1:M): One restaurant can offer multiple cuisines, represented by the Menu table's foreign key (menu item id) referencing the Cuisine table.
- **b.** Restaurant-Menu (1:M): One restaurant can have multiple items in its menu, represented by the Menu table's foreign key (res_id) referencing the Restaurant table. Foreign Key(s): res_id in the "Menu" table references res_id in the "Restaurant" table. Description: This relationship establishes a connection between restaurants and their menus. It allows each restaurant to have a list of menu items associated with it.
- **c.** Customer-Orders (1:M): One customer can place multiple orders, represented by the Orders table's foreign key (cus_id) referencing the Customer table. Foreign Key(s): cus_id in the "Orders" table references cus_id in the "Customer" table. Description: This relationship links each order to a specific customer, enabling the system to associate orders with individual customers.
- **d.** Restaurant-Orders (1:M): One restaurant can receive multiple orders, represented by the Orders table's foreign key (res_id) referencing the Restaurant table. Foreign Key(s): res_id in the "Orders" table

references res_id in the "Restaurant" table. Description: This relationship associates each order with a specific restaurant. It allows tracking orders back to the restaurant from which they were placed.

- **e.** Delivery_Person-Invoice (1:M): One delivery person can be associated with multiple invoices, represented by the Invoice table's foreign key (dp_id) referencing the Delivery_Person table.
- **f.** Orders-Invoice (1:1): Each order has a corresponding invoice, represented by the Invoice table's foreign key (order_id) referencing the Orders table.
- g. Customer-Rating (1:M): One customer can provide multiple ratings, represented by the Rating table's foreign key (cus_id) referencing the Customer table.
- **h.** Orders-Rating (1:1): Each order can have a corresponding rating, represented by the Rating table's foreign key (order_id) referencing the Orders table. Foreign Key(s): order_id, menu_item_id in the "Rating" table reference order_id, menu_item_id in the "Orders" table. Description: This relationship associates ratings with specific orders and menu items, enabling the system to store and retrieve feedback for both.
- i. Menu-Rating (1:1): Each menu item can have a corresponding rating, represented by the Rating table's foreign key (menu_item_id) referencing the Menu table.

2.2 Table Descriptions

RESTAURANT Table: Stores information about different restaurants.

Attributes: res id: Unique identifier for each restaurant.

res_name: Name of the restaurant.

res_location: Location or address of the restaurant. res_contact: Contact information for the restaurant. res_manager: Name of the restaurant manager.

CUSTOMER Table: Contains details about the customers who place orders.

Attributes: cus id: Unique identifier for each customer.

cus_firstname: First name of the customer. cus lastname: Last name of the customer.

cus_contact: Contact information for the customer.

CUISINE Table: Holds information about different cuisines associated with menu items.

Attributes: cuisine_id: Unique identifier for each cuisine.

cuisine_name: Name of the cuisine.

MENU Table: Represents the menu items offered by each restaurant.

Attributes: menu item id: Unique identifier for each menu item.

menu item desc: Description of the menu item.

res_id: Foreign key referencing the restaurant to which the menu item belongs. cuisine_id: Foreign key referencing the cuisine associated with the menu item.

ORDERS Table: Contains details about the orders placed by customers.

Attributes: order id: Unique identifier for each order.

res_id: Foreign key referencing the restaurant receiving the order.

cus_id: Foreign key referencing the customer placing the order.

menu_item_id: Foreign key referencing the menu item in the order.

order_date: Date and time when the order was placed.

DELIVERY_PERSON Table: Stores information about the delivery personnel.

Attributes: dp_id: Unique identifier for each delivery person.

dp_firstname: First name of the delivery person. dp_lastname: Last name of the delivery person.

RATING Table: Contains ratings provided by customers for various aspects.

Attributes: rating_id: Unique identifier for each rating.

order_id: Foreign key referencing the order for which the rating is given.

res_id: Foreign key referencing the restaurant being rated.

menu_item_id: Foreign key referencing the menu item being rated.

dp_id: Foreign key referencing the delivery person being rated.

restaurant_rating: Rating for the restaurant. menu_item_rating: Rating for the menu item.

delivery_person_rating: Rating for the delivery person.

INVOICE Table: Contains details about invoices generated for orders.

Attributes: invoice_id: Unique identifier for each invoice.

order_id: Foreign key referencing the order for which the invoice is generated. dp id: Foreign key referencing the delivery person associated with the order.

Invoice_amount: Total amount in the invoice.

payment_type: Type of payment used.

DEALS Table: Table stores information about special promotions and discounts offered by restaurants.

Attributes: week_day: Represents the day of the week for which the deal is applicable. premium_membership: Indicates whether the deal is exclusive to premium members. discount_percent: Specifies the percentage of discount offered in the deal.

3. Data Collection

3.1 Sample Data Generation:

For the purpose of meaningful analysis and system testing, realistic sample data has been generated to simulate various scenarios within the Restaurant Data Analysis Management system. This sample data covers diverse aspects of the system, including information about restaurants, customers, menu items, orders, delivery personnel, deals, invoices, and ratings. This ensures that the database is representative of real-world scenarios, allowing for comprehensive testing and analysis.

3.2 Data Initialization Process:

The data initialization process involves populating the tables with the generated sample data. This step is crucial for creating a functional and representative database that mirrors the actual usage of the system. The sample data is inserted into corresponding tables using SQL INSERT statements, establishing relationships between entities through foreign keys. This process ensures that the database is well-structured and contains meaningful information for subsequent analysis and testing.

By executing the provided SQL commands, the system is initialized with sample data, enabling users to interact with the application in a manner that closely resembles real-world usage. This initialization process is fundamental for evaluating the system's performance, functionality, and effectiveness in handling various scenarios.

- i. **CREATE TABLE:** This command is used to create tables and define columns, primary keys, foreign keys, and constraints.
- ii. **ALTER TABLE:** It allows modification of an existing table, including adding or dropping columns, defining primary and foreign keys, etc.
- iii. **JOIN:** SQL uses JOIN operations (e.g., INNER JOIN, LEFT JOIN, RIGHT JOIN) to combine rows from multiple tables based on related columns between them.
- iv. **FOREIGN KEY Constraint:** This constraint ensures the referential integrity of the data in atable by specifying that the values in a column must match values in a related table's primary key.

There are mainly three types of relationships used:

- one-to-one relationships
- one-to-many relationships
- many-to-many relationships

By establishing and maintaining these relationships using keys and SQL commands, databases organize data efficiently and ensure data consistency across multiple tables.

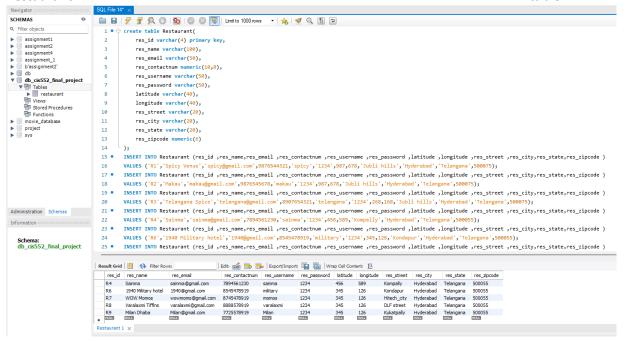
3.3 Data Integrity:

Data integrity in the database is maintained using:

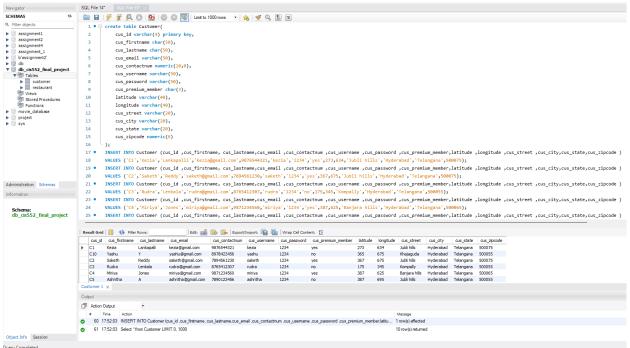
- **Referential Integrity:** This is maintained through foreign keys. It ensures the relationships between tables are preserved. A foreign key in one table points to the primary key in another table.
- **Constraints:** SQL provides various constraints like NOT NULL, UNIQUE, CHECK, and DEFAULT constraints. These enforce rules at the column level.
- Indexes: While not directly related to integrity, indexes improve data retrieval performance. They ensure that data is stored and retrieved efficiently while maintaining the integrity of relationships.
- **Triggers:** Triggers are special stored procedures that are automatically executed or fired when certain events occur (like an INSERT, UPDATE, or DELETE operation). They help maintain data integrity by allowing custom checks, modifications, or actions to be performed when specific conditions are met.

Using above rules for Data collection, Data initialization and Data integrity methods we have created the tables for our database. The following are the screenshots of database tables that we crated and inserted values.

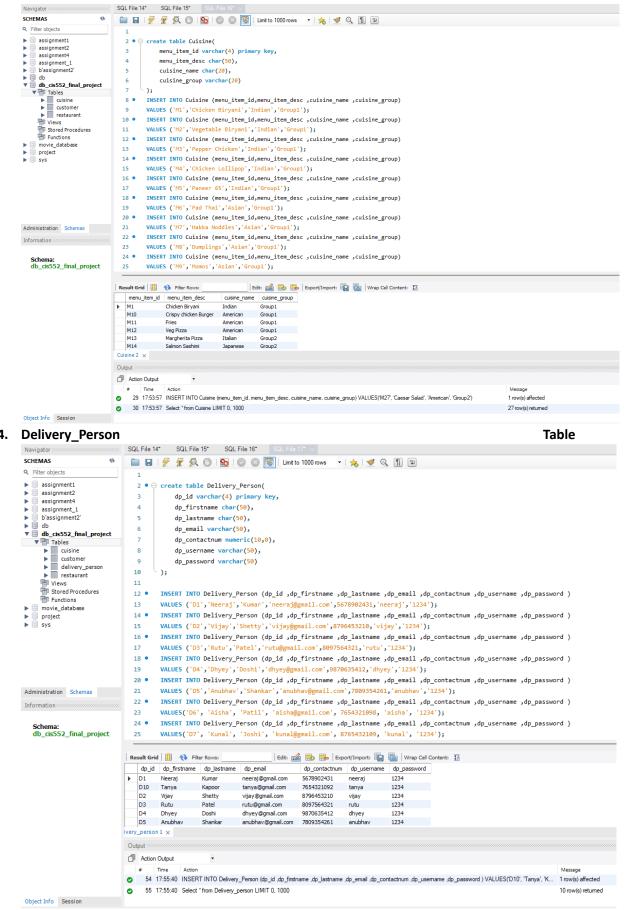
1. Restaurant table



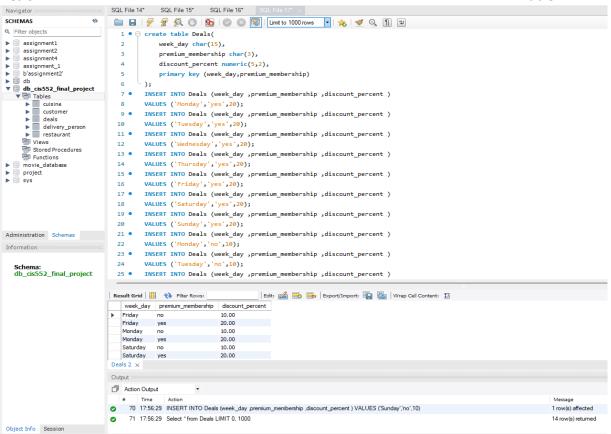
2. Customer Table



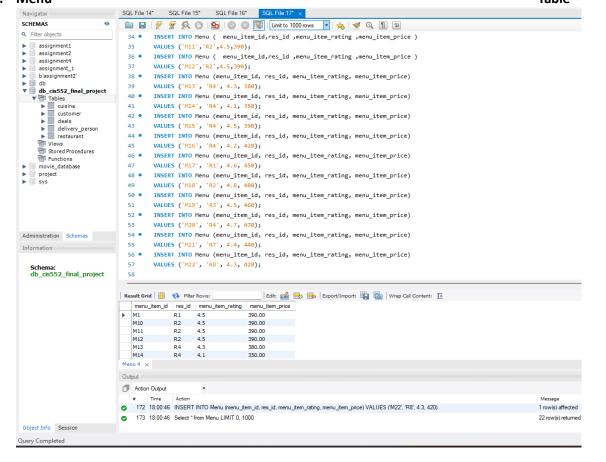
3. Cuisine Table



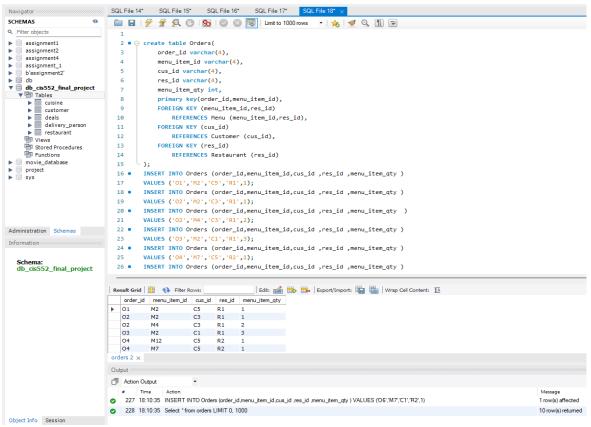
5. Deals Table



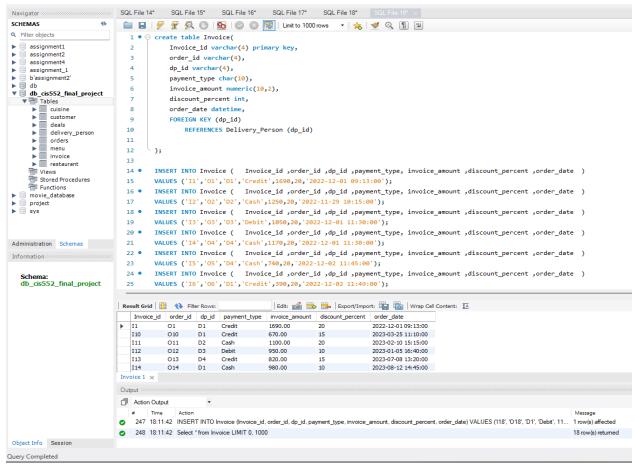
6. Menu Table

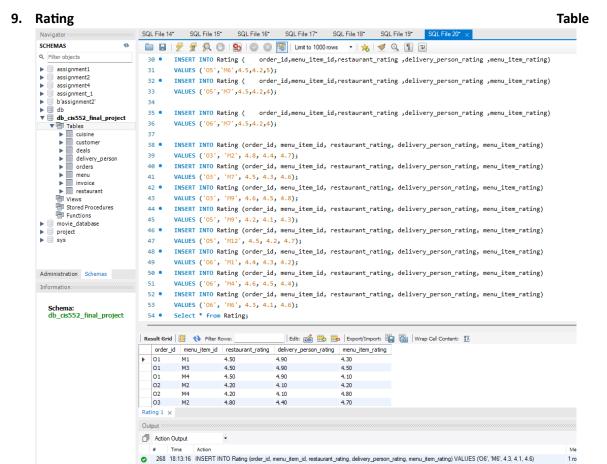


7. Orders Table



8. Invoice Table





19 r

4. SQL Queries

4.1 Query Examples

Object Info Session
Ouery Completed

i. Query to show top 2 best rated restaurants:

The query identifies the top 2 best-rated restaurants. It uses a subquery to calculate the average restaurant rating for each restaurant, assigns a dense rank based on the descending order of average ratings, and then filters the results to include only those with a rank of 1 or 2. Finally, it selects the restaurant name and average rating for display.

269 18:13:17 Select *from Rating LIMIT 0, 1000

```
#Query to show top 2 best rated restaurants

1 #Query to show top 2 best rated restaurants

2 select res_name,Avg_Restaurant_Rating from

3 (select res_name,avg(restaurant_rating) as Avg_Restaurant_Rating,DENSE_RANK() over (order by avg(restaurant_rating) desc) as res_rank from

4 (SELECT distinct res_name,restaurant_rating FROM RESTAURANT

JOIN ORDERS ON restaurant_res_id=orders.res_id

3 JOIN RATING ON orders.order_id=rating.order_id

4 ) tablea

8 group by res_name

9 ORDER BY Avg_Restaurant_Rating DESC ) tableh

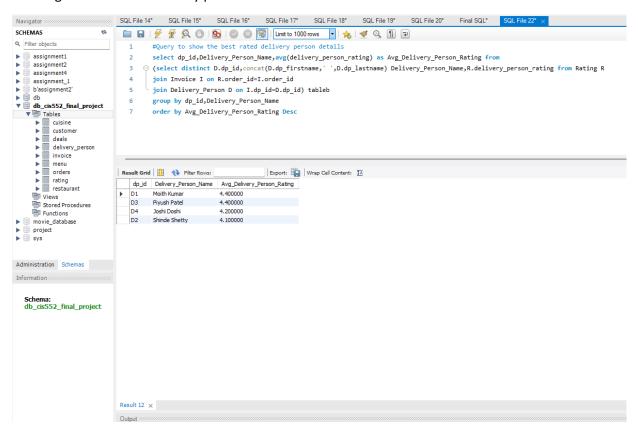
10 where res_rank in (1,2)

Result Grid  Fiber Rows:

| res_name | Avg_Restaurant_Rating | Export: | Exp
```

ii. Query to show the best rated delivery person details:

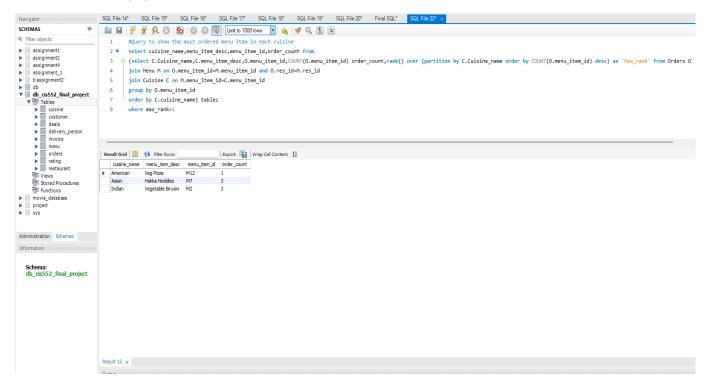
This SQL query retrieves the delivery person ID (dp_id), concatenated full name (Delivery_Person_Name), and the average rating (Avg_Delivery_Person_Rating) for each delivery person. The data is obtained by joining the Rating, Invoice, and Delivery_Person tables. The inner subquery (tableb) ensures distinct combinations of delivery person details. The outer query then calculates the average rating for each delivery person, grouping the results by dp_id and Delivery_Person_Name, and finally, it orders the output by the average rating in descending order, showing the best-rated delivery persons first.



iii. Query to show the most ordered menu item in each cuisine:

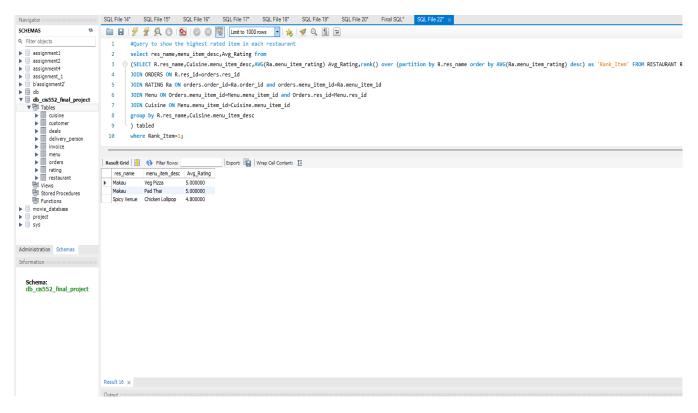
This query retrieves the most ordered menu item in each cuisine by joining the Orders, Menu, and Cuisine tables. It utilizes the RANK() window function to rank menu items based on order count within their respective cuisines. The results are filtered to include only the top-ranked items in each cuisine,

providing information such as cuisine name, menu item description, ID, and the corresponding order count for the most popular menu item in each cuisine.



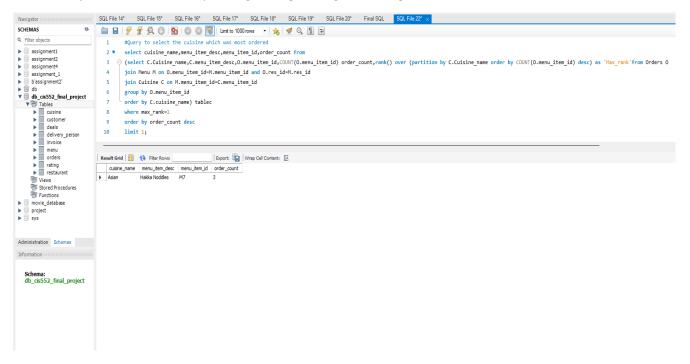
iv. Query to show the highest rated item in each restaurant:

This query retrieves the highest-rated menu item in each restaurant by joining the RESTAURANT, ORDERS, RATING, MENU, and CUISINE tables. It uses the AVG() function to calculate the average rating for each menu item within each restaurant. The RANK() window function is then applied to rank the items based on their average ratings within each restaurant. The final results are filtered to include only the top-ranked items in each restaurant, providing information such as restaurant name, menu item description, and the corresponding average rating for the highest-rated item in each restaurant.



v. Query to select the cuisine which was most ordered:

This query retrieves the highest-rated menu item in each restaurant by joining the RESTAURANT, ORDERS, RATING, MENU, and CUISINE tables. It uses the AVG() function to calculate the average rating for each menu item within each restaurant. The RANK() window function is then applied to rank the items based on their average ratings within each restaurant. The final results are filtered to include only the top-ranked items in each restaurant, providing information such as restaurant name, menu item description, and the corresponding average rating for the highest-rated item in each restaurant.



4.2 Query Optimization Techniques

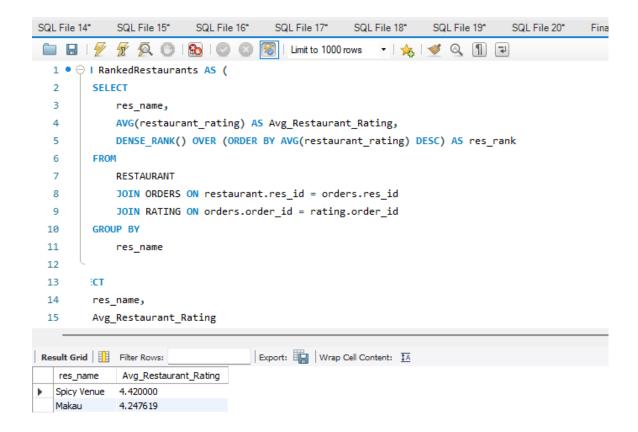
Query optimization is a crucial aspect of database management that aims to enhance the performance of database queries by minimizing execution time and resource usage. Here are some general query optimization techniques:

- Use Indexes: Indexes help speed up data retrieval by providing a quick lookup mechanism.
 Ensure that columns involved in WHERE clauses, JOIN conditions, and ORDER BY clauses are properly indexed.
- ii. *Avoid SELECT: Retrieve only the columns needed for the query rather than using SELECT * to fetch all columns. This reduces the amount of data transferred and processed.
- iii. Optimize JOIN Operations: Use INNER JOINs when possible, as they are generally more efficient than OUTER JOINs. Consider using appropriate JOIN types based on the relationship between tables.
- iv. Use WHERE Clause Effectively: Apply filters in the WHERE clause to limit the number of rows processed. This reduces the amount of data that needs to be read and improves query performance.
- v. Aggregate Functions and GROUP BY: Be mindful of using aggregate functions and GROUP BY clauses. Ensure that they are applied only when necessary to avoid unnecessary calculations.
- vi. Avoid Subqueries if Possible: Subqueries can be resource intensive. Consider using JOINs or other techniques to achieve the same result without the need for subqueries.
- vii. Use EXISTS and IN Sparingly: EXISTS and IN can be inefficient for large datasets. Consider alternative approaches, such as JOINs or EXISTS with correlated subqueries, for better performance.
- viii. Optimize ORDER BY and LIMIT: If possible, avoid sorting large result sets. Use ORDER BY and LIMIT judiciously and ensure that the columns used for sorting are indexed.
- ix. Partitioning: Partition large tables to improve query performance by restricting the amount of data that needs to be scanned.
- x. Use Proper Data Types: Choose appropriate data types for columns to minimize storage space and optimize query performance.
- xi. Database Design: A well-designed database schema with normalized tables and appropriate relationships can contribute to efficient query execution.

Here We tried to Optimize the couple of queries from above Queries using Optimization techniques:

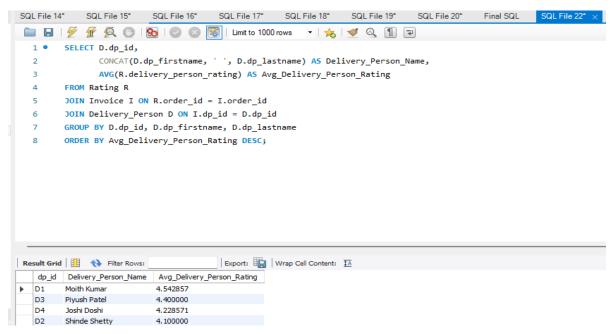
1. Query to show top 2 best rated restaurants:

- Table Aliases: Use meaningful table aliases to improve code readability.
- Avoid Redundant Subquery: You can eliminate the redundant subquery by using the window function directly in the outer query.



2. Query to show the best rated delivery person details:

- Removed Subquery: The subquery was removed, and the join conditions were integrated directly into the main query for better clarity.
- Avoided DISTINCT: The DISTINCT keyword was eliminated, assuming that the join conditions are appropriately set to avoid duplicate records.
- Simplified Concatenation: The CONCAT function was used directly in the SELECT clause for simplicity.

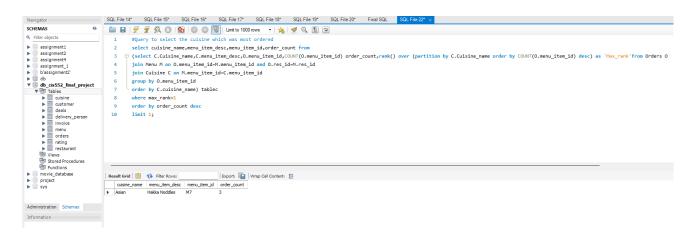


4.3 TRC queries

Tuple Relational Calculus (TRC) is a formal language used for relational databases. It is a non-procedural query language, meaning that it specifies what data to retrieve rather than how to retrieve it. TRC operates on sets of tuples, which represent rows in a relational database table.

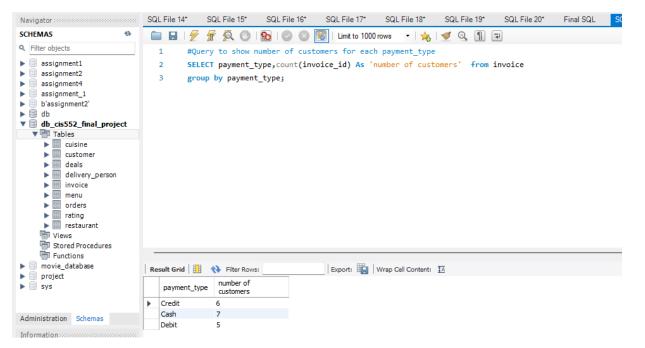
i. TRC to select the cuisine which was most ordered

 $\label{lem:count} \{ (cuisine_name,menu_item_desc,menu_item_id,order_count) | \\ \exists O,M,C(Orders(O)\land Menu(M)\land Cuisine(C)\land O.menu_item_id=M.menu_item_id\land O.res_id=M.res_id\land \\ M.menu_item_id=C.menu_item_id\land C.cuisine_name=cuisine_name\land M.menu_item_desc=menu_item_desc\land O.menu_item_id=menu_item_id \land COUNT(\land COUNT(O.menu_item_id)=order_count \\ RANK(COUNT(O.menu_item_id))=1 \} ORDER BY order_count DESC LIMIT 1 \\ \end{tabular}$



ii. TRC to show number of customers for each payment type

{(payment_type, COUNT(invoice_id))|\(\frac{1}{2}\) invoice(invoice_id,payment_type)}



5. Conclusion

In conclusion, the restaurant management system project has successfully used delivered a robust and well-structured relational database to address the complex operational requirements of a modern restaurant. By designing tables for entities such as restaurants, cuisines, menus, orders, invoices, and ratings, we've created a comprehensive system that ensures data integrity through foreign key relationships. Realistic sample data was generated to facilitate meaningful analysis and testing of the system's functionality. The implementation supports key restaurant management operations, including menu management, order placement, and invoice generation. Furthermore, the system's SQL queries enable valuable insights into restaurant performance, customer preferences, and efficient delivery operations. While the project has achieved its primary objectives, future enhancements could focus on scalability, user interface integration, and heightened security measures to meet the evolving needs of the restaurant industry. Overall, this project provides a strong foundation for a versatile and efficient restaurant management system.

6. Future Recommendations

- Scalability: While the system is designed to handle typical restaurant operations, scalability
 may be a consideration for large-scale enterprises. Future enhancements could involve
 optimizations for performance in such scenarios.
- ii. User Interface: The current implementation focuses on the database and backend functionality. Integration with a user interface for easy interaction with the system would be a logical next step.
- iii. Security Measures: Implementing robust security measures, such as encryption and access controls, would be essential for protecting sensitive customer and business data.
- iv. User Interface (UI) Enhancement: Develop an intuitive and user-friendly interface for both restaurant staff and customers. This can include features such as an interactive menu, easy order placement, and a visually appealing dashboard for restaurant administrators.
- v. Mobile Application Development: Consider creating a mobile application to extend the reach and convenience of the restaurant management system. Mobile apps can offer onthe-go access for customers, delivery personnel, and restaurant staff.
- vi. Integration with Online Platforms: Integrate the system with popular online food delivery platforms to broaden the restaurant's customer base. Seamless integration with platforms like UberEats, DoorDash, or Grubhub can streamline order processing and increase visibility
- vii. Advanced Analytics and Reporting: Enhance the analytical capabilities of the system by incorporating advanced reporting tools and data visualization. This can help restaurant owners and managers make data-driven decisions based on sales trends, customer feedback, and operational performance.

Github:

The complete project details can be found in my github account by following this link,

Restaurant Data Analysis Management for Informed Decision-Making