Bookstore Analysis Project Report

Project Title: Online Bookstore Analysis Using SQL
Prepared by: Aelina Taqvi

---

1. Project Objective The primary goal of this project was to build and analyze a structured database for an online bookstore. By leveraging SQL, this project focuses on deriving meaningful insights into book inventory, customer demographics, order trends, sales performance, and revenue generation.

---

2. Tools & Technologies Used

- **Database:** MySQL
- **Skills Applied:**
  - SQL (DDL & DML)
  - Data Analysis & Aggregation
  - Joins, Grouping, Filtering
  - Business Intelligence Thinking

---

3. Dataset Overview Three CSV files were considered for the project:

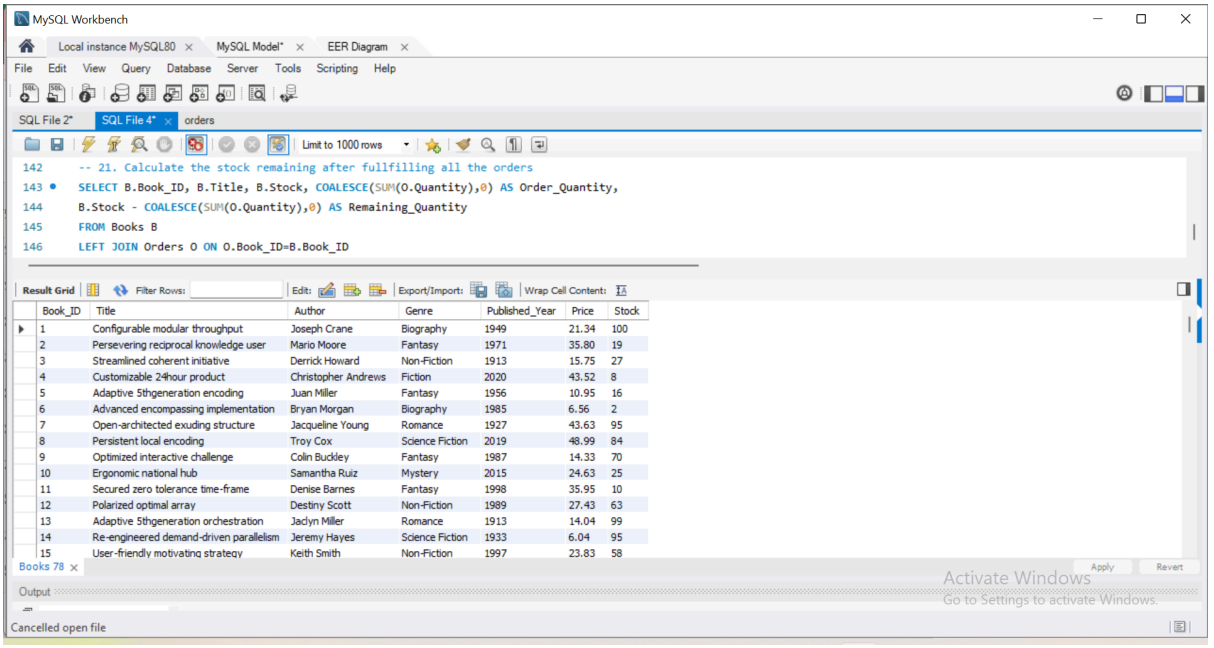
- **Books.csv:** Contains details like Book ID, Title, Author, Genre, Published Year, Price, and Stock.
- **Customers.csv:** Includes Customer ID, Name, Email, Phone, City, and Country.
- **Orders.csv:** Records Order ID, Customer ID, Book ID, Order Date, Quantity, and Total Amount.

These datasets were imported and converted into relational tables in a MySQL database named `BookStore`.

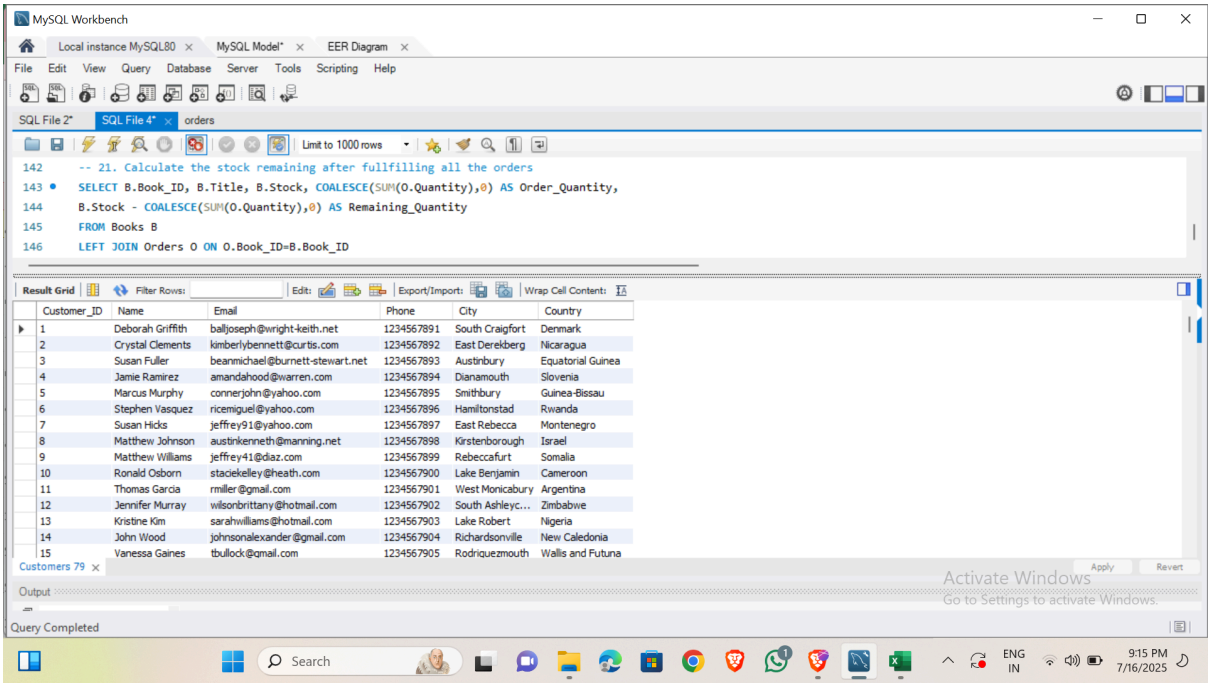
---

4. Database Schema Design Tables created:

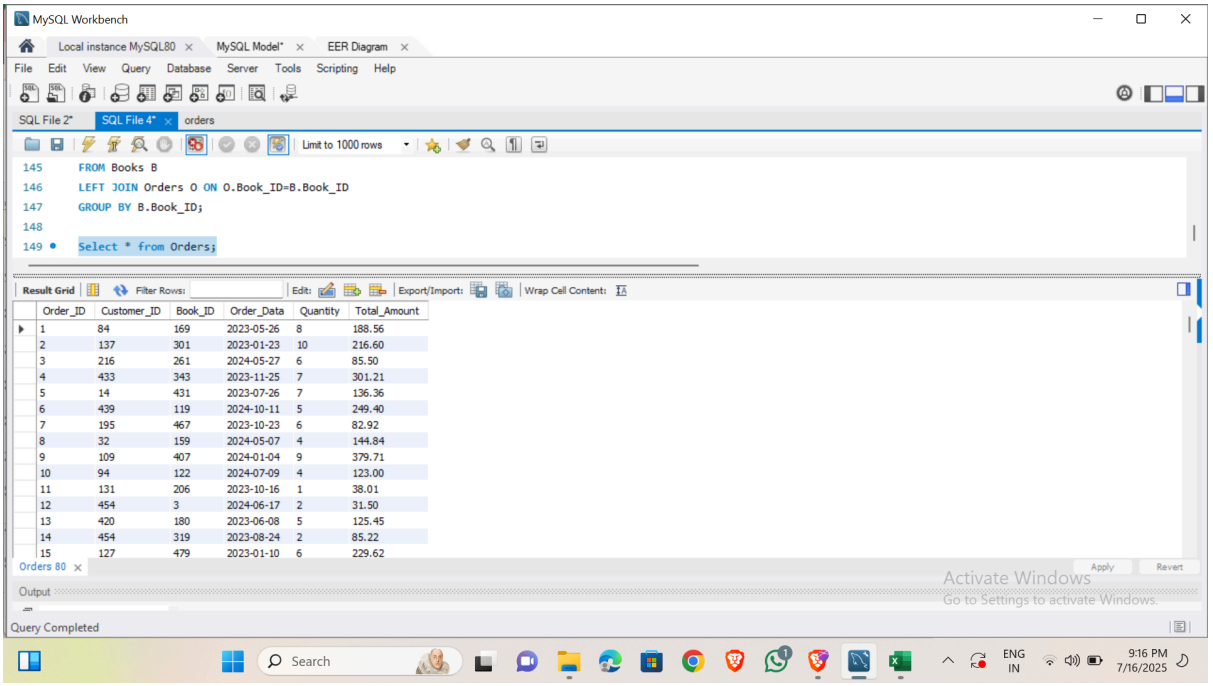
- `Books (Book_ID, Title, Author, Genre, Published_Year, Price, Stock)`

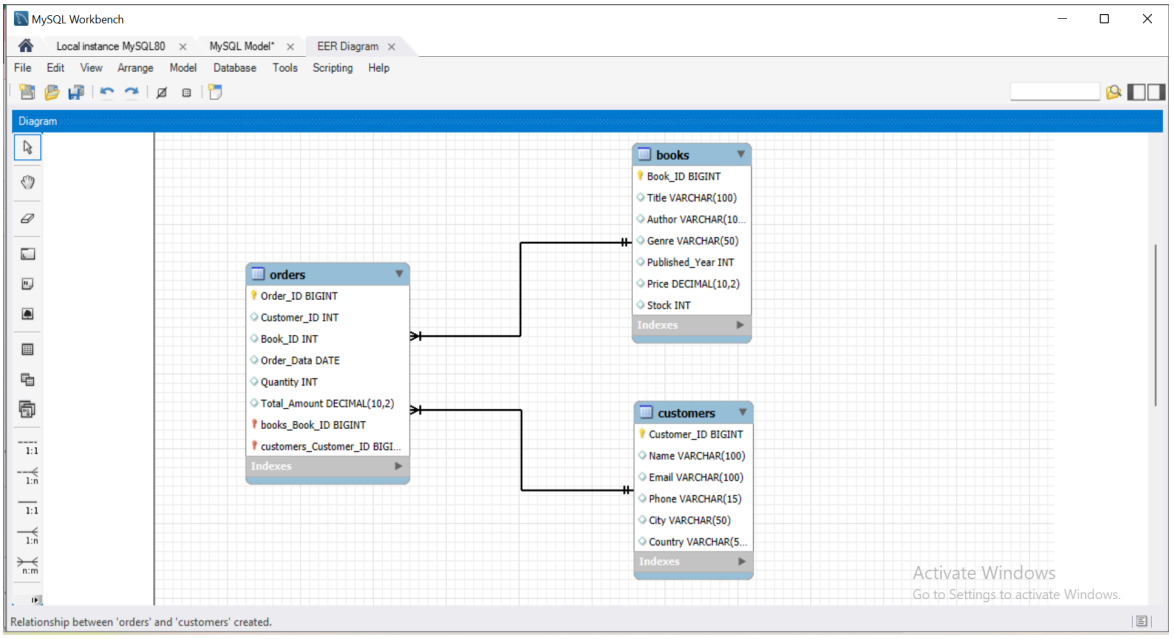


- Customers (Customer_ID, Name, Email, Phone, City, Country)



- Orders (Order_ID, Customer_ID, Book_ID, Order_Date, Quantity, Total_Amount)





Relationships:

- `Orders` table references `Books` and `Customers` through foreign keys.

---

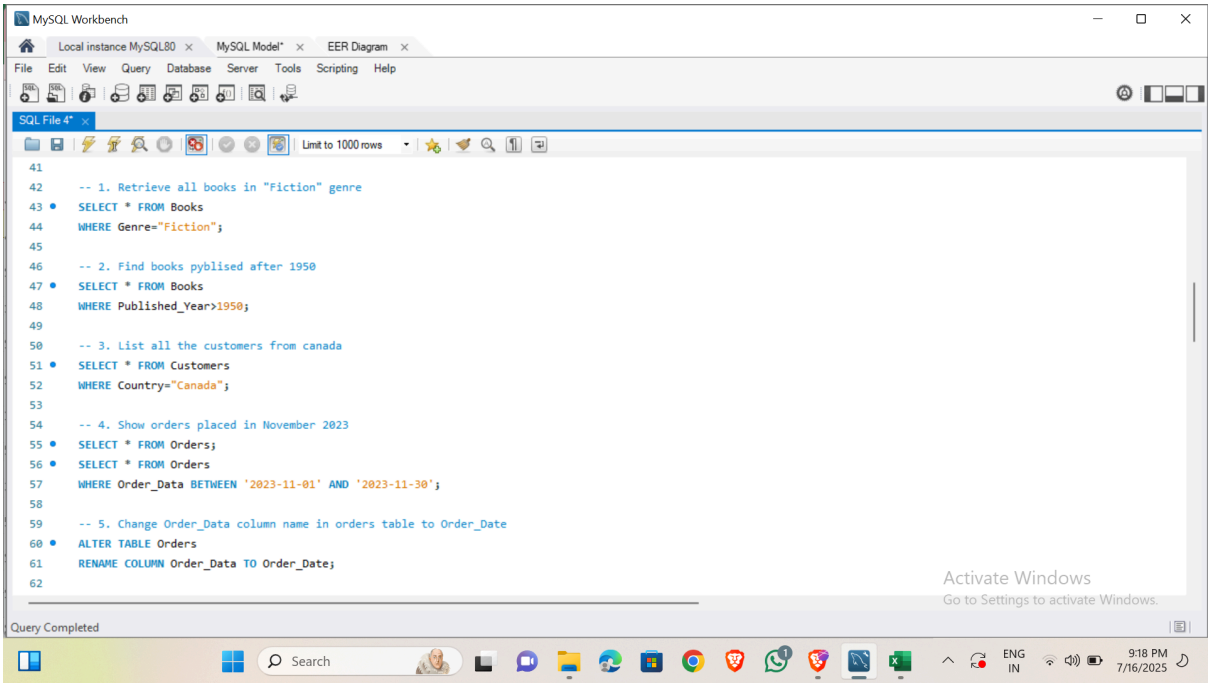
5. Key SQL Queries and Insights

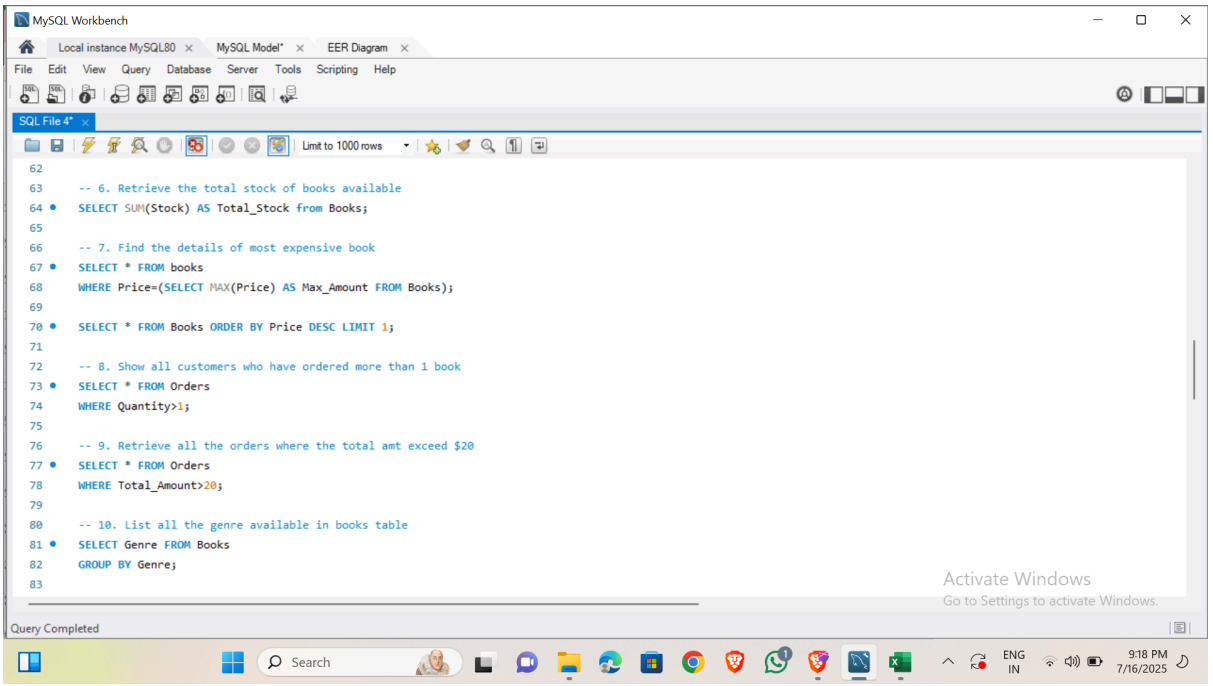
Basic Analysis:

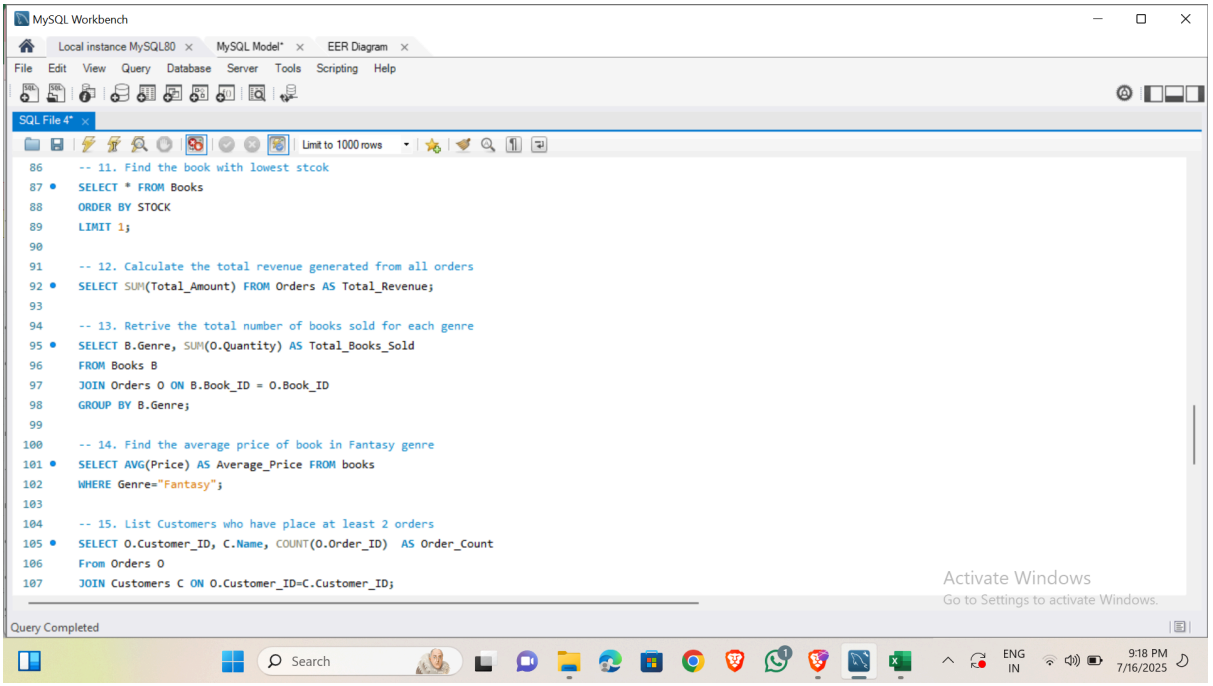
1. **Fiction Books:** Retrieved all books in the Fiction genre
2. **Modern Publications:** Found books published after 1950.
3. **Canadian Customers:** Listed all customers from Canada.
4. **Order Trend:** Identified orders placed in November 2023.
5. **Change Column Name:** Used `ALTER TABLE to rename column.`
6. **Stock Availability:** Calculated total stock using `SUM(Stock)`.
7. **Most Expensive Book:** Extracted book with the maximum price.
8. **Bulk Buyers:** Showed customers ordering more than 1 quantity.
9. **High-Value Orders:** Retrieved orders exceeding $20.
10. **Genres Available:** Used `DISTINCT` and `GROUP BY` to list unique genres.
11. **Low Stock Alert:** Identified books with the least stock.
12. **Total Revenue:** Aggregated all order amounts to compute revenue.

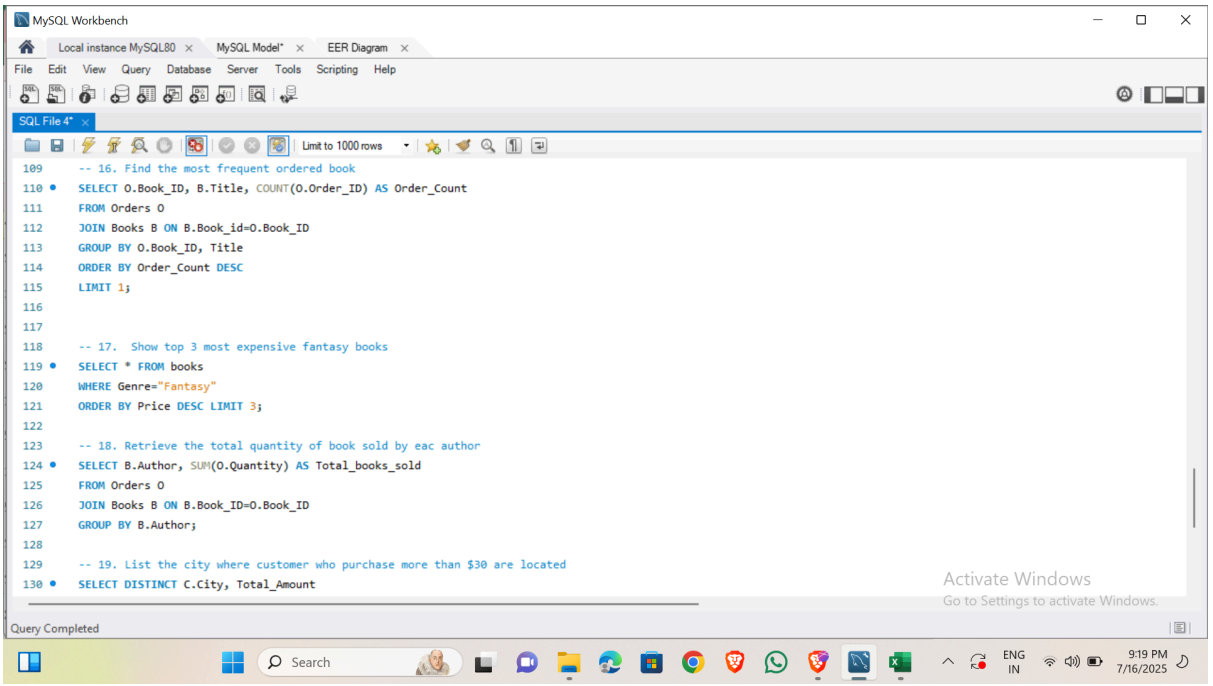
Advanced Analysis:

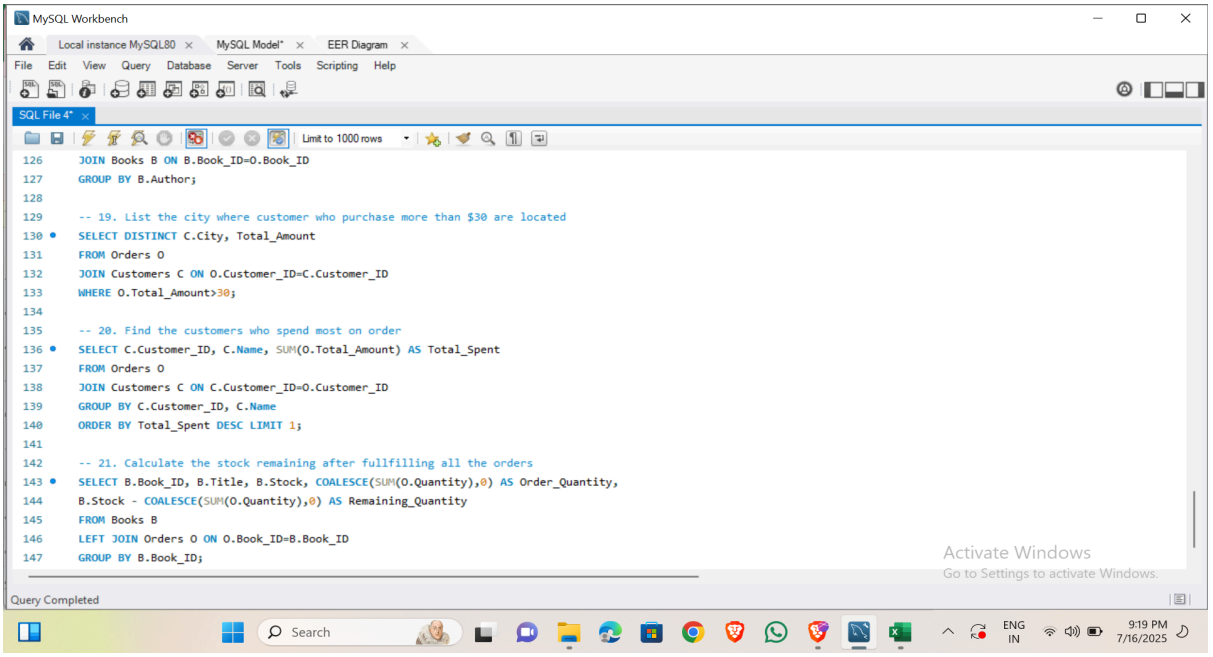
13. **Books Sold by Genre:** Joined `Books` and `Orders` to get quantity sold.
14. **Average Fantasy Price:** Calculated average price of Fantasy books.
15. **Repeat Customers:** Identified customers with 2+ orders.
16. **Popular Book:** Found most frequently ordered book.
17. **Top Fantasy Picks:** Listed 3 most expensive Fantasy books.
18. **Author Performance:** Computed books sold by each author.
19. **Top Cities:** Retrieved cities with customers spending over $30.
20. **Top Spender:** Found customer who spent the most.
21. **Remaining Stock:** Used `LEFT JOIN` and `COALESCE` to calculate stock after all orders.


MySQL Workbench
Local instance MySQL80
MySQL Model*
EER Diagram
File Edit View Query Database Server Tools Scripting Help
SQL File 4*
Limit to 1000 rows
41
42 -- 1. Retrieve all books in "Fiction" genre
43 SELECT * FROM Books
44 WHERE Genre="Fiction";
45
46 -- 2. Find books pyblised after 1950
47 SELECT * FROM Books
48 WHERE Published_Year>1950;
49
50 -- 3. List all the customers from canada
51 SELECT * FROM Customers
52 WHERE Country="Canada";
53
54 -- 4. Show orders placed in November 2023
55 SELECT * FROM Orders;
56 SELECT * FROM Orders
57 WHERE Order_Data BETWEEN '2023-11-01' AND '2023-11-30';
58
59 -- 5. Change Order_Data column name in orders table to Order_Date
60 ALTER TABLE Orders
61 RENAME COLUMN Order_Data TO Order_Date;
62
Activate Windows
Go to Settings to activate Windows.
Query Completed
Search
ENG
IN
9:18 PM
7/16/2025


MySQL Workbench
Local instance MySQL80
MySQL Model*
EER Diagram
File Edit View Query Database Server Tools Scripting Help
SQL File 4*
Limit to 1000 rows
62
63 -- 6. Retrieve the total stock of books available
64 SELECT SUM(Stock) AS Total_Stock from Books;
65
66 -- 7. Find the details of most expensive book
67 SELECT * FROM books
68 WHERE Price=(SELECT MAX(Price) AS Max_Amount FROM Books);
69
70 SELECT * FROM Books ORDER BY Price DESC LIMIT 1;
71
72 -- 8. Show all customers who have ordered more than 1 book
73 SELECT * FROM Orders
74 WHERE Quantity>1;
75
76 -- 9. Retrieve all the orders where the total amt exceed $20
77 SELECT * FROM Orders
78 WHERE Total_Amount>20;
79
80 -- 10. List all the genre available in books table
81 SELECT Genre FROM Books
82 GROUP BY Genre;
83
Activate Windows
Go to Settings to activate Windows.
Query Completed
Search
ENG
IN
9:18 PM
7/16/2025


MySQL Workbench
Local instance MySQL80
MySQL Model*
EER Diagram
File Edit View Query Database Server Tools Scripting Help
SQL File 4*
Limit to 1000 rows
86 -- 11. Find the book with lowest stcok
87 SELECT * FROM Books
88 ORDER BY STOCK
89 LIMIT 1;
90
91 -- 12. Calculate the total revenue generated from all orders
92 SELECT SUM(Total_Amount) FROM Orders AS Total_Revenue;
93
94 -- 13. Retrive the total number of books sold for each genre
95 SELECT B.Genre, SUM(O.Quantity) AS Total_Books_Sold
96 FROM Books B
97 JOIN Orders O ON B.Book_ID = O.Book_ID
98 GROUP BY B.Genre;
99
100 -- 14. Find the average price of book in Fantasy genre
101 SELECT AVG(Price) AS Average_Price FROM books
102 WHERE Genre="Fantasy";
103
104 -- 15. List Customers who have place at least 2 orders
105 SELECT O.Customer_ID, C.Name, COUNT(O.Order_ID) AS Order_Count
106 From Orders O
107 JOIN Customers C ON O.Customer_ID=C.Customer_ID;
Activate Windows
Go to Settings to activate Windows.
Query Completed
Search
ENG IN
9:18 PM 7/16/2025


MySQL Workbench
Local instance MySQL80
MySQL Model*
EER Diagram
File Edit View Query Database Server Tools Scripting Help
SQL File 4*
Limit to 1000 rows
109 -- 16. Find the most frequent ordered book
110 SELECT O.Book_ID, B.Title, COUNT(O.Order_ID) AS Order_Count
111 FROM Orders O
112 JOIN Books B ON B.Book_id=O.Book_ID
113 GROUP BY O.Book_ID, Title
114 ORDER BY Order_Count DESC
115 LIMIT 1;
116
117
118 -- 17. Show top 3 most expensive fantasy books
119 SELECT * FROM books
120 WHERE Genre="Fantasy"
121 ORDER BY Price DESC LIMIT 3;
122
123 -- 18. Retrieve the total quantity of book sold by eac author
124 SELECT B.Author, SUM(O.Quantity) AS Total_books_sold
125 FROM Orders O
126 JOIN Books B ON B.Book_ID=O.Book_ID
127 GROUP BY B.Author;
128
129 -- 19. List the city where customer who purchase more than $30 are located
130 SELECT DISTINCT C.City, Total_Amount
Activate Windows
Go to Settings to activate Windows.
Query Completed
Search
ENG IN
9:19 PM 7/16/2025



---

6. Challenges Faced

- Ensuring referential integrity between tables.
- Handling NULL values using COALESCE.
- Structuring grouped aggregations for meaningful insights.

---

7. Key Learnings

- Gained proficiency in writing efficient SQL queries.
- Understood data relationships in real-world scenarios.
- Learned how to derive actionable business insights through structured querying.

---

8. Future Enhancements

- Integrate Power BI or Tableau for visualization.
- Use Python for advanced analytics or predictions.
- Add feedback analysis to assess customer satisfaction.

---

9. Conclusion This project demonstrates how structured SQL queries can turn raw bookstore data into actionable insights. It showcases the strength of relational databases in supporting business intelligence and decision-making processes.
