

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)

MySQL Workbench

Local instance MySQL80 x MySQL Model x EER Diagram x

File Edit View Query Database Server Tools Scripting Help

SQL File 2* SQL File 4* orders

Limit to 1000 rows

```

142 -- 21. Calculate the stock remaining after fulfilling all the orders
143 • SELECT B.Book_ID, B.Title, B.Stock, COALESCE(SUM(O.Quantity),0) AS Order_Quantity,
144       B.Stock - COALESCE(SUM(O.Quantity),0) AS Remaining_Quantity
145 FROM Books B
146 LEFT JOIN Orders O ON O.Book_ID=B.Book_ID

```

Result Grid

Book_ID	Title	Author	Genre	Published_Year	Price	Stock
1	Configurable modular throughput	Joseph Crane	Biography	1949	21.34	100
2	Persevering reciprocal knowledge user	Mario Moore	Fantasy	1971	35.80	19
3	Streamlined coherent initiative	Derrick Howard	Non-Fiction	1913	15.75	27
4	Customizable 24-hour product	Christopher Andrews	Fiction	2020	43.52	8
5	Adaptive 5thgeneration encoding	Juan Miller	Fantasy	1956	10.95	16
6	Advanced encompassing implementation	Bryan Morgan	Biography	1985	6.56	2
7	Open-architected exuding structure	Jacqueline Young	Romance	1927	43.63	95
8	Persistent local encoding	Troy Cox	Science Fiction	2019	48.99	84
9	Optimized interactive challenge	Colin Buckley	Fantasy	1987	14.33	70
10	Ergonomic national hub	Samantha Ruiz	Mystery	2015	24.63	25
11	Secured zero tolerance time-frame	Denise Barnes	Fantasy	1998	35.95	10
12	Polarized optimal array	Destiny Scott	Non-Fiction	1989	27.43	63
13	Adaptive 5thgeneration orchestration	Jadyn Miller	Romance	1913	14.04	99
14	Re-engineered demand-driven parallelism	Jeremy Hayes	Science Fiction	1933	6.04	95
15	User-friendly motivating strategy	Keith Smith	Non-Fiction	1997	23.83	58

Books 78 x

Output

Cancelled open file

Activate Windows
Go to Settings to activate Windows.

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

MySQL Workbench

Local instance MySQL80 x MySQL Model x EER Diagram x

File Edit View Query Database Server Tools Scripting Help

SQL File 2* SQL File 4* orders

Limit to 1000 rows

```

142 -- 21. Calculate the stock remaining after fulfilling all the orders
143 • SELECT B.Book_ID, B.Title, B.Stock, COALESCE(SUM(O.Quantity),0) AS Order_Quantity,
144       B.Stock - COALESCE(SUM(O.Quantity),0) AS Remaining_Quantity
145 FROM Books B
146 LEFT JOIN Orders O ON O.Book_ID=B.Book_ID

```

Result Grid

Customer_ID	Name	Email	Phone	City	Country
1	Deborah Griffith	balljoseph@wright-keith.net	1234567891	South Craigfort	Denmark
2	Crystal Clements	kimberlybennett@curtis.com	1234567892	East Derekberg	Nicaragua
3	Susan Fuller	beanmichael@burnett-stewart.net	1234567893	Austinbury	Equatorial Guinea
4	Jamie Ramirez	amandahood@warren.com	1234567894	Dianamouth	Slovenia
5	Marcus Murphy	cornerjohn@yahoo.com	1234567895	Smithbury	Guinea-Bissau
6	Stephen Vasquez	ricemiguel@yahoo.com	1234567896	Hamiltonstad	Rwanda
7	Susan Hicks	jeffrey91@yahoo.com	1234567897	East Rebecca	Montenegro
8	Matthew Johnson	austrikeneth@manning.net	1234567898	Krstenborough	Israel
9	Matthew Williams	jeffrey41@diaz.com	1234567899	Rebeccafurt	Somalia
10	Ronald Osborn	staciekelley@heath.com	1234567900	Lake Benjamin	Cameroon
11	Thomas Garcia	rmiller@gmail.com	1234567901	West Monicabury	Argentina
12	Jennifer Murray	wilsonbrittany@hotmail.com	1234567902	South Ashley...	Zimbabwe
13	Kristine Kim	sarahwilliams@hotmail.com	1234567903	Lake Robert	Nigeria
14	John Wood	johnsonalexander@gmail.com	1234567904	Richardsonville	New Caledonia
15	Vanessa Gaines	tbullock@gmail.com	1234567905	Rodriguezmouth	Wallis and Futuna

Customers 79 x

Output

Query Completed

Activate Windows
Go to Settings to activate Windows.

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

The screenshot shows the MySQL Workbench interface with a SQL query executed. The query is as follows:

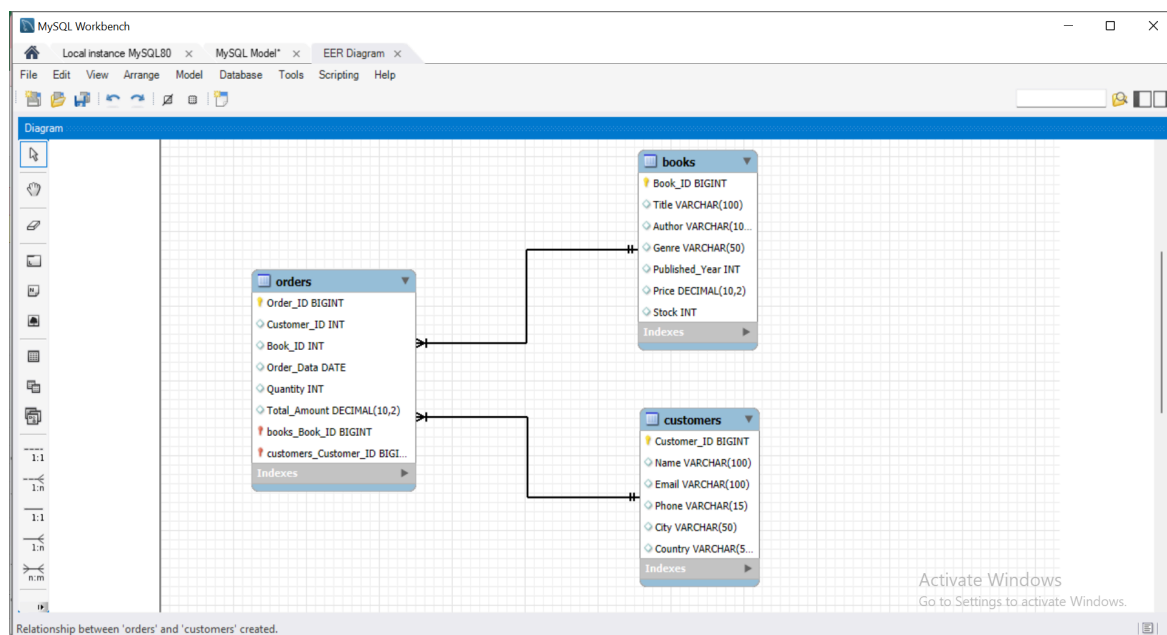
```

145 FROM Books B
146 LEFT JOIN Orders O ON O.Book_ID=B.Book_ID
147 GROUP BY B.Book_ID;
148
149 • Select * from Orders;

```

The results are displayed in the Result Grid, showing 15 rows of data. The columns are Order_ID, Customer_ID, Book_ID, Order_Date, Quantity, and Total_Amount.

Order_ID	Customer_ID	Book_ID	Order_Date	Quantity	Total_Amount
1	84	169	2023-05-26	8	188.56
2	137	301	2023-01-23	10	216.60
3	216	261	2024-05-27	6	85.50
4	433	343	2023-11-25	7	301.21
5	14	431	2023-07-26	7	136.36
6	439	119	2024-10-11	5	249.40
7	195	467	2023-10-23	6	82.92
8	32	159	2024-05-07	4	144.84
9	109	407	2024-01-04	9	379.71
10	94	122	2024-07-09	4	123.00
11	131	206	2023-10-16	1	38.01
12	454	3	2024-06-17	2	31.50
13	420	180	2023-06-08	5	125.45
14	454	319	2023-08-24	2	85.22
15	127	479	2023-01-10	6	229.62



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 x MySQL Model x EER Diagram x

File Edit View Query Database Server Tools Scripting Help

SQL File 4 x

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 published 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
```

Query Completed

Activate Windows
Go to Settings to activate Windows.

Windows taskbar: Search, 9:18 PM 7/16/2025

MySQL Workbench

Local instance MySQL80 x MySQL Model x EER Diagram x

File Edit View Query Database Server Tools Scripting Help

SQL File 4 x

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
```

Query Completed

Activate Windows
Go to Settings to activate Windows.

Windows taskbar: Search, 9:18 PM 7/16/2025

MySQL Workbench

Local instance MySQL80 x MySQL Model x EER Diagram x

File Edit View Query Database Server Tools Scripting Help

SQL File 4 x

Limit to 1000 rows

```
86 -- 11. Find the book with lowest stock
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. Retrieve 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;
```

Query Completed

Activate Windows
Go to Settings to activate Windows.

9:18 PM
7/16/2025

MySQL Workbench

Local instance MySQL80 x MySQL Model x EER Diagram x

File Edit View Query Database Server Tools Scripting Help

SQL File 4 x

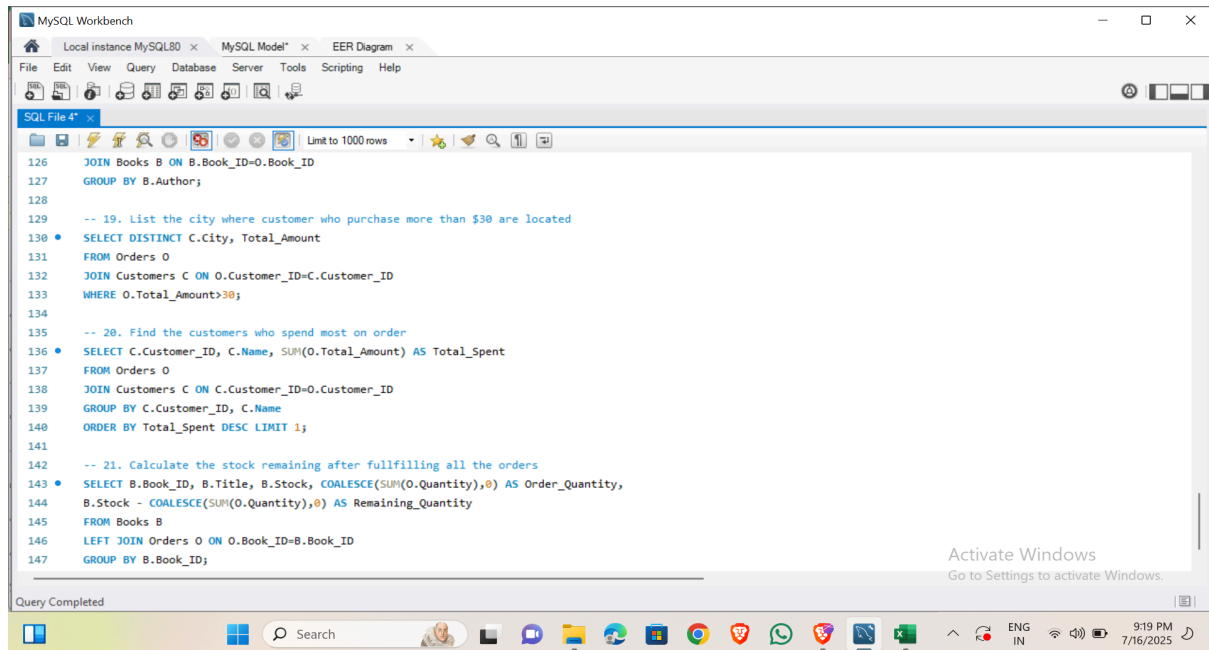
Limit to 1000 rows

```
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
```

Query Completed

Activate Windows
Go to Settings to activate Windows.

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.
