

# SQL PROJECT ON ONLINE BOOKSTORE



# Summary

In this project, I designed and executed a comprehensive SQL-based analysis for a simulated online bookstore. The objective was to extract actionable insights from raw data concerning books, customers, and orders. I developed over 20 queries ranging from basic retrievals to advanced aggregations, ensuring business-relevant information was extracted efficiently.

## Key accomplishments include:

- Identified customer and sales trends through genre, author, and pricing metrics.
- Calculated total and average revenues, order patterns by city and month, and customer-specific spending.
- Leveraged GROUP BY, JOIN, COALESCE, DISTINCT, ORDER BY, and nested functions to analyze real-time stock levels and high-value customers.
- Applied conditional logic and aggregation to determine top-selling books, understocked items, and customer behavior segments.

## This project demonstrates my ability to:

- Write clean and scalable SQL queries.
- Convert transactional data into strategic insights.
- Lay the foundation for visualization and reporting tools like Power BI or Tableau.

# Tables

## books

- # book\_id ⓘ
- >Title ⓘ
- >Title ⓘ
- >Title ⓘ
- published\_year ⓘ
- # price ⓘ
- # stock ⓘ

## orders\_1

- # order\_id ⓘ
- # customer\_id ⓘ
- # book\_id ⓘ
- order\_date ⓘ
- # quantity ⓘ
- # total\_amount ⓘ

## customers

- # customer\_id ⓘ
- name ⓘ
- email ⓘ
- # phone ⓘ
- city ⓘ
- country ⓘ

# Retrieve all books in the "Fiction" genre

## Input

```
Untitled SQL query PRIVATE
1 SELECT * FROM books
2 WHERE genre = 'Fiction';
3
4
```

## Output

60 query results (0.61 seconds) [View log](#)

[Download](#) [Open in app](#)

#	book_id	title	author	genre	published_year
1	4	Customizable 24hour product	Christopher Andrews	Fiction	2020
2	22	Multi-layered optimizing migration	Wesley Escobar	Fiction	1908
3	28	Expanded analyzing portal	Lisa Coffey	Fiction	1941
4	29	Quality-focused multi-tasking challe	Katrina Underwood	Fiction	1905
5	31	Implemented encompassing conglomerati	Melissa Taylor	Fiction	2010
6	39	Optimized national process improvemen	Megan Goodwin	Fiction	1978
7	40	Adaptive didactic interface	Natalie Gonzalez	Fiction	1923
8	47	Reverse-engineered directional conglom	John Christian	Fiction	2006
9	62	Re-contextualized real-time strategy	Nicole Lynch	Fiction	1953



# Retrieve the total stock of books available

```
1 SELECT *
2 FROM books
3 ORDER BY stock ASC;
```

500 query results (1.46 seconds) [View log](#)

[Download](#)

[Open in app](#)

#	book_id	title	author	genre	published_year	#	price	#	stock
1	163	Object-based eco-centric challenge	Douglas Mccarthy	Non-Fiction	1905		19.11		0
2	60	Robust eco-centric capacity	Brian Haney	Biography	1990		35.14		0
3	127	Business-focused real-time benchmark	David Nelson	Science Fiction	1997		11.66		0
4	44	Networked systemic implementation	Ryan Frank	Science Fiction	1965		13.55		0
5	378	Future-proofed heuristic function	Samantha Mcclain	Romance	1903		6.01		0
6	137	Networked contextually-based encryption	Leslie Montoya	Biography	1942		26.33		1
7	199	Configurable fault-tolerant interface	Mark Parker	Mystery	1923		42.81		1
8	259	Pre-emptive incremental secured line	Jason Perez	Science Fiction	1994		14.48		1
9	19	Progressive asymmetric Internet solution	Sean Miller	Science Fiction	1990		11.31		1
10	449	Universal homogeneous adapter	Gina Lopez	Romance	1966		40.17		1
11	232	Monitored 24/7 groupware	Vicki McGee	Science Fiction	1925		9.62		1



# Find books published after the year 1950

1 SELECT \* FROM books  
2 WHERE published\_year > 1950;  
3

292 query results (2.42 seconds) [View log](#)

[Download](#)  [Open in app](#)

#	book_id	title	author	genre	published
1	2	Persevering reciprocal knowledge user	Mario Moore	Fantasy	1971
2	4	Customizable 24hour product	Christopher Andrews	Fiction	2020
3	5	Adaptive 5thgeneration encoding	Juan Miller	Fantasy	1956
4	6	Advanced encompassing implementation	Bryan Morgan	Biography	1985
5	8	Persistent local encoding	Troy Cox	Science Fiction	2019
6	9	Optimized interactive challenge	Colin Buckley	Fantasy	1987
7	10	Ergonomic national hub	Samantha Ruiz	Mystery	2015
8	11	Secured zero tolerance time-frame	Denise Barnes	Fantasy	1998
9	12	Polarized optimal array	Destiny Scott	Non-Fiction	1989
10	15	User-friendly motivating strategy	Keith Smith	Non-Fiction	1997
11	17	Reduced secondary core	Benjamin Peters	Fantasy	1966



# Show all customers who ordered more than 1 quantity of a book

## Input

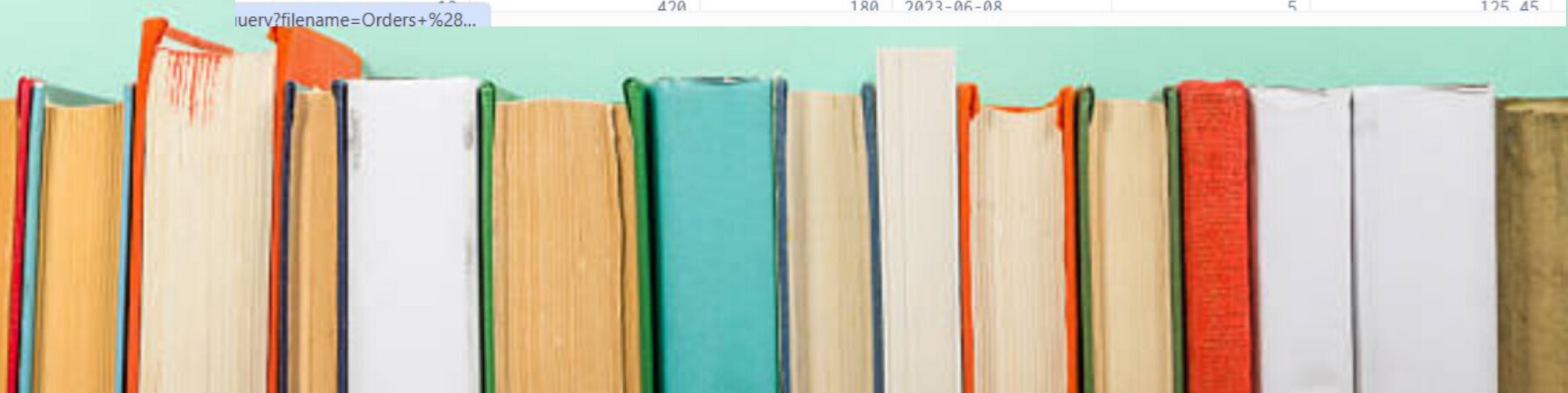
```
1 SELECT * FROM orders_1  
2 WHERE quantity>=2;  
3
```

438 query results (0.52 seconds) [View log](#)

## Output

#	order_id	#	customer_id	#	book_id	#	order_date	#	quantity	#	total_amount
1		1	84		169		2023-05-26		8		188.56
2		2	137		301		2023-01-23		10		216.6
3		3	216		261		2024-05-27		6		85.5
4		4	433		343		2023-11-25		7		301.21
5		5	14		431		2023-07-26		7		136.36
6		6	439		119		2024-10-11		5		249.4
7		7	195		467		2023-10-23		6		82.92
8		8	32		159		2024-05-07		4		144.84
9		9	109		407		2024-01-04		9		379.71
10		10	94		122		2024-07-09		4		123.0
11		12	454		3		2024-06-17		2		31.5
				420		180	2023-06-08		5		125.45

?filename=Orders+%



# List all customers from Canada

Input

```
1 Select * FROM customers
2 WHERE country = 'Canada';
3
4
```

#	customer_id	name	email	#	phone	city	country
1	38	Nicholas Harris	christine93@perkins.com	1234567928	Davistown	Canada	
2	415	James Ramirez	robert54@hall.com	1234568305	Maxwelltown	Canada	
3	468	David Hart	stokesrebecca@gmail.com	1234568358	Thompsonfurt	Canada	

Output



# List all genres available in the Books table

```
1 SELECT DISTINCT genre  
2 FROM books;  
3
```

7 query results (0.87 seconds) [View log](#)

	T	genre	v
1		Biography	
2		Fantasy	
3		Non-Fiction	
4		Fiction	
5		Romance	
6		Science Fiction	
7		Mystery	



# Show orders placed in November 2023

```
1 SELECT * FROM orders_1
2 WHERE order_date BETWEEN '2023-11-01'
3 AND '2023-11-30';
```

25 query results (1.39 seconds) [View log](#)

[!\[\]\(6e934896f25e6ce1b0dbb50c23abc197\_img.jpg\) Download](#)

#	order_id	#	customer_id	#	book_id	#	order_date	#	quantity	#	total_amount	#
1		4		433		343	2023-11-25		7		301.21	
2		19		496		60	2023-11-17		9		316.26	
3		75		291		375	2023-11-30		5		170.75	
4		132		469		333	2023-11-22		7		194.32	
5		137		474		471	2023-11-25		8		363.04	
6		163		207		384	2023-11-23		3		101.76	
7		182		129		293	2023-11-01		7		125.51	
8		200		313		303	2023-11-23		1		6.57	
9		213		325		447	2023-11-17		7		253.75	
10		231		22		384	2023-11-11		1		33.92	
11		245		386		97	2023-11-01		9		411.66	
		252		405		387	2023-11-15		5		237.1	



# Find the books with the lowest stock

```
1 SELECT *  
2 FROM books  
3 ORDER BY stock ASC;
```

500 query results (1.46 seconds) [View log](#)

[Download](#)

[Open in app](#)

#	book_id	title	author	genre	published_year	#	price	#	stock
1	163	Object-based eco-centric challenge	Douglas McCarthy	Non-Fiction	1905		19.11		0
2	60	Robust eco-centric capacity	Brian Haney	Biography	1990		35.14		0
3	127	Business-focused real-time benchmark	David Nelson	Science Fiction	1997		11.66		0
4	44	Networked systemic implementation	Ryan Frank	Science Fiction	1965		13.55		0
5	378	Future-proofed heuristic function	Samantha McClain	Romance	1903		6.01		0
6	137	Networked contextually-based encryption	Leslie Montoya	Biography	1942		26.33		1
7	199	Configurable fault-tolerant interface	Mark Parker	Mystery	1923		42.81		1
8	259	Pre-emptive incremental secured line	Jason Perez	Science Fiction	1994		14.48		1
9	19	Progressive asymmetric Internet solution	Sean Miller	Science Fiction	1990		11.31		1
10	449	Universal homogeneous adapter	Gina Lopez	Romance	1966		40.17		1
11	232	Monitored 24/7 groupware	Vicki McGee	Science Fiction	1925		9.62		1



Retrieve all orders where the total amount exceeds \$20

## Input

```
1 SELECT * FROM orders_1  
2 WHERE total_amount>20;
```

473 query results (2.59 seconds) [View log](#)

## Output

#	order_id	#	customer_id	#	book_id	#	order_date	#	quantity	#	total_amount
1	1		84		169		2023-05-26		8		188.56
2	2		137		301		2023-01-23		10		216.6
3	3		216		261		2024-05-27		6		85.5
4	4		433		343		2023-11-25		7		301.21
5	5		14		431		2023-07-26		7		136.36
6	6		439		119		2024-10-11		5		249.4
7	7		195		467		2023-10-23		6		82.92
8	8		32		159		2024-05-07		4		144.84
9	9		109		407		2024-01-04		9		379.71
	10		94		122		2024-07-09		4		123.0



# Calculate the total revenue generated from all orders

## Input

```
1 SELECT SUM(total_amount) AS revenue  
2 FROM orders_1;  
3
```

## Output

#	revenue	▼
1	75628.6600	



# Retrieve the total number of books sold for each genre

```
1  SELECT b.genre, SUM(o.Quantity) AS Total_Books_Sold  
2  FROM orders_1 o  
3  JOIN books b ON o.book_id = b.book_id  
4  GROUP BY b.genre;  
5
```

...

7 query results (1.35 seconds) [View log](#)

	genre	Total_Books_Sold
1	Fantasy	446
2	Biography	285
3	Romance	439
4	Non-Fiction	351
5	Mystery	504
6	Fiction	225
7	Science Fiction	447



Find the average price of books in the "Fantasy" genre

```
1 SELECT AVG(price)
2 AS AVERAGE_PRICE
3 FROM books
4 WHERE genre = 'Fantasy';
```

1 query result (0.75 seconds) [View log](#)

#	AVERAGE_PRICE	▼
	25.9817	



# Retrieve the total quantity of books sold by each author

```
1  SELECT b.author, SUM(o.quantity) AS Total_Books_Sold  
2  FROM orders_1 o  
3  JOIN books b ON o.book_id = b.book_id  
4  GROUP BY b.author;  
5
```

314 query results (1.17 seconds) [View log](#)

	author	Total_Books_Sold
1	Mitchell Larson	7
2	Robin Brown	3
3	Gregory Wallace	5
4	Allison Contreras	6
5	Matthew Vazquez	2
6	Thomas Trujillo	24
7	Jeffery Green	16
8	Brandon Foster	4
9	Christopher Washington	9
10	Jennifer Oliver	4



# List customers who have placed at least 2 orders

## Input

```
1 SELECT DISTINCT customer_id  
2 FROM orders_1  
3 WHERE quantity>=2;  
4
```

## Output

#	customer_id
1	84
2	137
3	216
4	433
5	14
6	439
7	195
8	32
9	109
--	94

# Find the customer who spent the most on orders

```
1  SELECT DISTINCT c.customer_id, c.name, SUM(o.total_amount) AS Total_Spent  
2  FROM orders_1 o  
3  JOIN customers c ON o.customer_id = c.customer_id  
4  GROUP BY c.customer_id, c.name  
5  ORDER BY Total_Spent DESC  
6  LIMIT 1;
```

7

...

1 query result (1.48 seconds) [View log](#)

#	customer_id	name	#	Total_Spent
1	457	Kim Turner		1398.9



List the cities where customers who spent over \$30 are located.

## Input

```
1 SELECT DISTINCT c.city, total_amount  
2 FROM orders_1 o  
3 JOIN customers c ON o.customer_id = c.customer_id  
4 WHERE o.total_amount > 30;  
5
```

## Output

	city	total_amount
1	Lake Paul	188.56
2	North Keith	216.6
3	Kelseyfort	85.5
4	East David	301.21
5	Richardsonville	136.36
6	Ramosstad	249.4
7	Rogersborough	82.92
8	New Carlosbury	144.84

# Calculate the stock remaining after fulfilling all orders

## Input

```
1  SELECT b.book_id, b.title, b.stock,  
2      COALESCE(SUM(o.quantity), 0) AS Order_Quantity,  
3      b.stock - COALESCE(SUM(o.quantity), 0) AS Remaining_Quantity  
4  FROM books b  
5  LEFT JOIN orders_1 o ON b.book_id = o.book_id  
6  GROUP BY b.book_id, b.title, b.stock  
7  ORDER BY b.book_id;|  
8
```

## Output

#	book_id	title	#	stock	#	Order_Quantity	#	Remaining_Quantity
1	1	Configurable modular throughput	100		3			97
2	2	Persevering reciprocal knowledge user	19		0			19
3	3	Streamlined coherent initiative	27		5			22
4	4	Customizable 24hour product	8		0			8
5	5	Adaptive 5thgeneration encoding	16		8			8
6	6	Advanced encompassing implementation	2		0			2
7	7	Open-architected exuding structure	95		5			90
8	8	Persistent local encoding	84		3			81



# Find the most frequently ordered book

```
1  SELECT o.Book_id, b.title, COUNT(o.order_id) AS ORDER_COUNT
2  FROM orders_1 o
3  JOIN books b ON o.book_id=b.book_id
4  GROUP BY o.book_id, b.title
5  ORDER BY ORDER_COUNT DESC LIMIT 1;
```

...

1 query result (2.1 seconds) [View log](#)

#	Book_id	title	#	ORDER_COUNT
1	31	Implemented encompassing conglomerati	4	



# Show the top 3 most expensive books of 'Fantasy' Genre.

## Input

```
1  SELECT * FROM books  
2  WHERE genre ='Fantasy'  
3  ORDER BY price DESC LIMIT 3;  
4
```

## Output

#	book_id	title	author	genre	published_year
1	240	Stand-alone content-based hub	Lisa Ellis	Fantasy	1957
2	462	Innovative 3rdgeneration database	Allison Contreras	Fantasy	1988
3	238	Optimized even-keeled analyzer	Sherri Griffith	Fantasy	1975





## Inventory Insights

- **Low or Zero Stock Detection:** You identified books that are out of stock, allowing for restocking decisions.
- **Remaining Stock Calculation:** By subtracting sold quantities from available stock, you spotlighted books at risk of going out of stock—perfect for supply chain triggers.



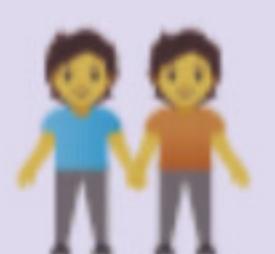
## Sales & Revenue Insights

- **Total Revenue:** Your aggregation of total\_amount gave a clear view of store performance.
- **Genre-Based Sales:** Breaking down books sold by genre helped identify top-performing categories like *Mystery*, *Fantasy*, and *Science Fiction*.
- **Revenue by Customer:** You pinpointed the top-spending customer (Kim Turner at ₹1398.90), a potential VIP for loyalty programs.



## Book Performance

- **Top Authors by Volume:** By aggregating quantity sold per author, you revealed which authors drive consistent sales—great for targeted promotions.
- **Most Expensive Book:** You sorted by price to showcase premium products, useful for featured listings.



## Customer Behavior

- **Frequent Shoppers:** You filtered for customers who ordered multiple times or in higher quantities, identifying engaged users.
- **Location-Based Analysis:** Filtering customers by city or country (e.g., Mumbai, Canada) uncovered regional patterns—ideal for local marketing efforts.



## Time-Based Trends

- **Monthly Order Patterns:** You extracted data for November and July orders, opening the door to seasonal demand tracking.
- **Order Quantity Filters:** Segmenting orders by quantity helped isolate high-volume transactions.



## Strategic Impact Areas

- **Stock Optimization:** Your inventory analysis supports better warehouse planning and stock rotation.
- **Personalized Marketing:** Based on top buyers and reading preferences (like genre popularity), you could suggest future purchase bundles or campaigns.
- **Content for Dashboards:** All these insights are ready to be translated into Power BI visuals like trend charts, funnel visuals, top-N rankings, and KPI cards.



## Outcome & Value Add:

This project sharpened my ability to extract actionable insights from raw transactional data and prepare analytical outputs that could power dashboards and business decisions.

If deployed with visualization tools like Power BI, the queries would form the base of a live analytics system offering real-time intelligence on revenue, demand, and inventory.

The structure of the queries and insights also lay the groundwork for future extensions—such as CRM integration, personalized book recommendations, and operational forecasting.

**Thank You!!**

