

# ONLINE BOOK STORE ANALYSIS

PRESENTED BY : KOMAL SHARMA

# PROJECT OVERVIEW

**Project Type:** Semi-guided SQL project

**Platform:** pgAdmin (PostgreSQL)

**Objective:**

Analyze book sales, customer behavior, and inventory performance using SQL queries.

**Key Skills Demonstrated:**

- Writing SQL queries (basic to advanced)
- Joins, aggregations, grouping
- Business-oriented data exploration
- Drawing actionable insights from raw data



# DATASET OVERVIEW

## 1) Books Table (.csv)

- **Columns :** Book\_ID, Title, Author, Genre, Published\_Year, Price, Stock
- **Purpose :** Stores book information and inventory details

## 2) Customers Table (.csv)

- **Columns :** Customer\_ID, Name, Email, City, Country
- **Purpose :** Contains customer contact and location data

## 3) Orders Table (.csv)

- **Columns :** Order\_ID, Customer\_ID, Book\_ID, Order\_Date, Quantity, Total\_Amount
- **Purpose :** Tracks purchases, customer-book relationships, and order values



**Dataset Source :** Kaggle

**Focus Areas :** Sales, Customer Behavior, Inventory Management

# BUSINESS PROBLEM

The online bookstore wants to improve decision-making through SQL-driven insights into customer behavior, sales trends, and inventory levels.

## **Customer Behavior :**

- Identify top customers
- Find high-revenue cities

## **Sales Performance :**

- Track best-selling books and genres
- Analyze total revenue

## **Inventory Status :**

- Spot low-stock books
- Monitor stock remaining after sales

## **Goal :**

Help the business uncover patterns in buying behavior, optimize inventory, and focus on the most profitable areas.

# CUSTOMER INSIGHTS

**Query 1 :** List customers who have placed at least 2 orders.

```
select c.customer_id, c.name, count(o.order_id)
from customers c
join orders o
on c.customer_id = o.customer_id
group by c.customer_id
having count(o.order_id) >= 2;
```



	customer_id [PK] integer	name character varying (100)	count bigint
1	184	Richard McLaughlin	2
2	272	Carl Smith	3
3	22	Stacey Adams	3
4	173	Victoria Dixon	2
5	189	Jason Bell	2
6	117	Molly Murphy	2
7	125	Rebecca Perez	2
8	120	Rita Wallace	2
9	57	Nicolas Joseph	2
10	160	Robert Brooks	2

Total rows: 139 of 139    Query complete 00:00:00.325

# CUSTOMER INSIGHTS

**Query 2 :** List the cities where customers who spent over \$30 are located.

```
select c.city, o.total_amount as Total_spent  
from orders o  
join customers c  
on o.customer_id = c.customer_id  
where o.total_amount > 30;
```



	city character varying (100)	total_spent numeric (10,2)
1	Lake Paul	188.56
2	North Keith	216.60
3	Kelseyfort	85.50
4	East David	301.21
5	Richardsonville	136.36
6	Ramosstad	249.40
7	Rogersborough	82.92
8	New Carlosbury	144.84
9	Ravenberg	379.71
10	West Anthony	123.00

Total rows: 443 of 443

Query complete 00:00:0

# CUSTOMER INSIGHTS

Query 3 : Find the customer who spent the most on orders.

```
select c.customer_id, c.name,  
sum(o.total_amount) as Total_spent  
from orders o  
join customers c  
on o.customer_id = c.customer_id  
group by c.customer_id, c.name  
order by Total_spent desc  
limit 1;
```



	customer_id [PK] integer	name character varying (100)	total_spent numeric
1	457	Kim Turner	1398.90

# SALES INSIGHTS

Query 4 : Find the most frequently ordered book.

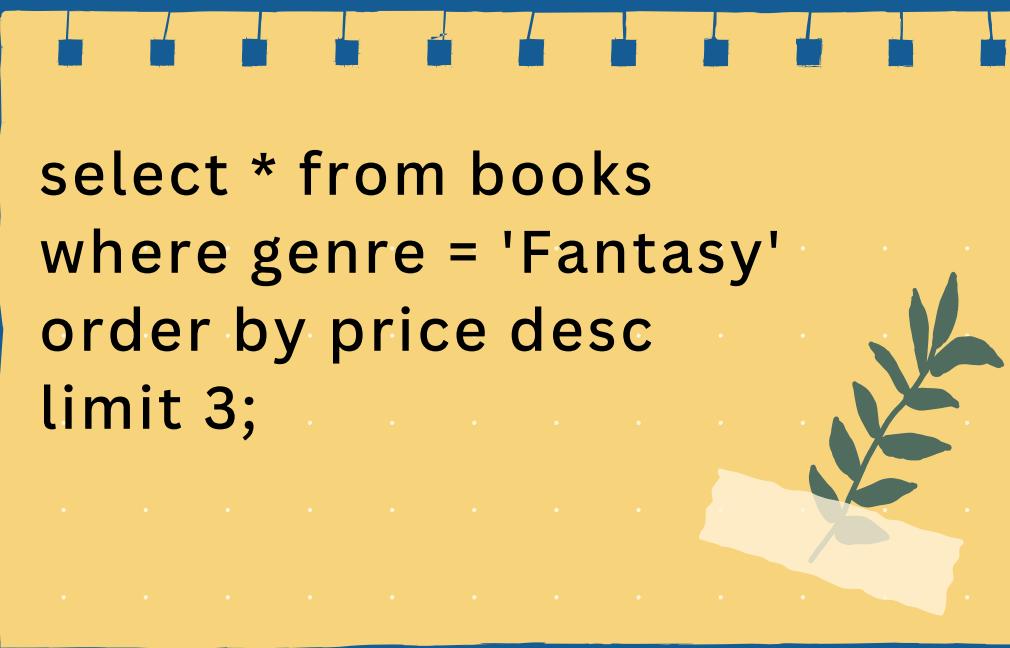
```
select o.book_id, b.title, count(o.order_id) as  
order_count  
from orders o  
join books b  
on o.book_id = b.book_id  
group by o.book_id, b.title  
order by count(o.order_id) desc  
limit 1;
```



	book_id	title	order_count
1	88	Robust tangible hardware	4

# SALES INSIGHTS

Query 5 : Show the Top 3 most expensive books of "Fantasy" genre.



	book_id [PK] integer	title character varying (100)	author character varying (100)	genre character varying (50)	published_year integer	price numeric (10,2)	stock integer
1	240	Stand-alone content-based hub	Lisa Ellis	Fantasy	1957	49.90	41
2	462	Innovative 3rdgeneration database	Allison Contreras	Fantasy	1988	49.23	62
3	238	Optimized even-keeled analyzer	Sherri Griffith	Fantasy	1975	48.97	72

# SALES INSIGHTS

Query 6 : Calculate the total revenue generated from all orders.

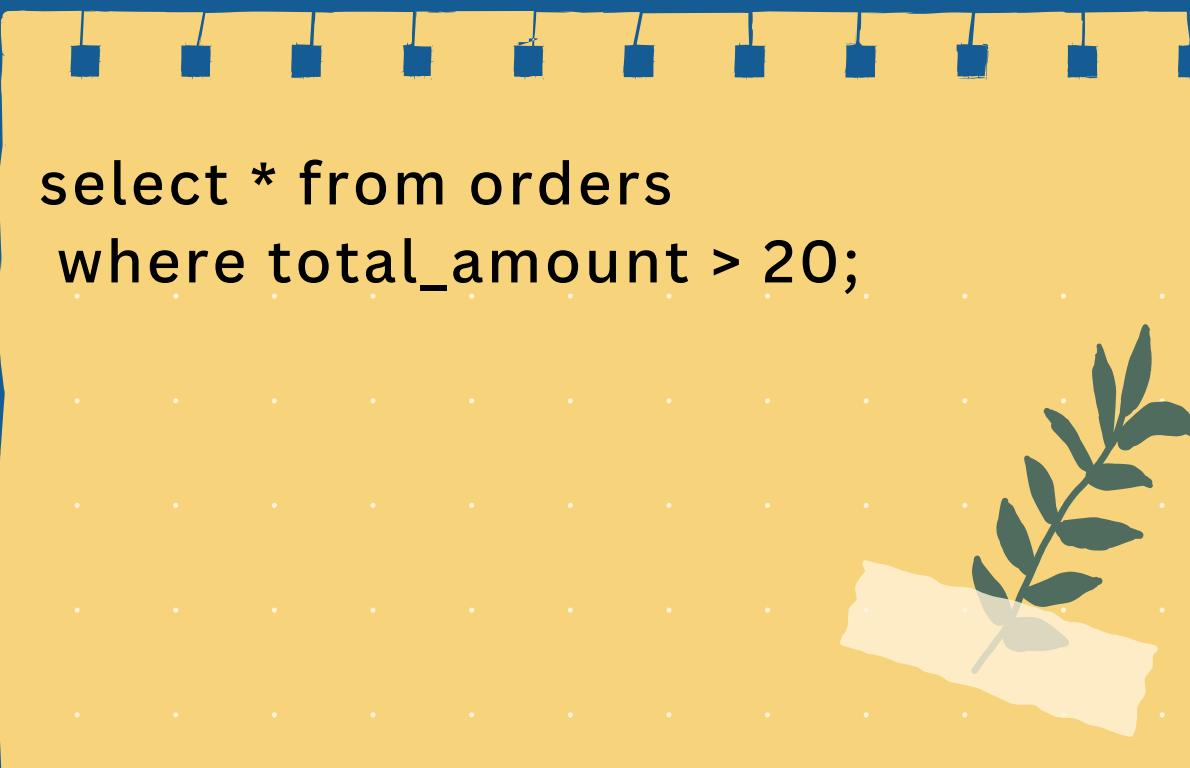
```
select sum(total_amount) as Total_Revenue  
from orders;
```



	total_revenue	numeric
1	75628.66	

# SALES INSIGHTS

Query 7 : Retrieve all orders where the total amount exceeds \$20.



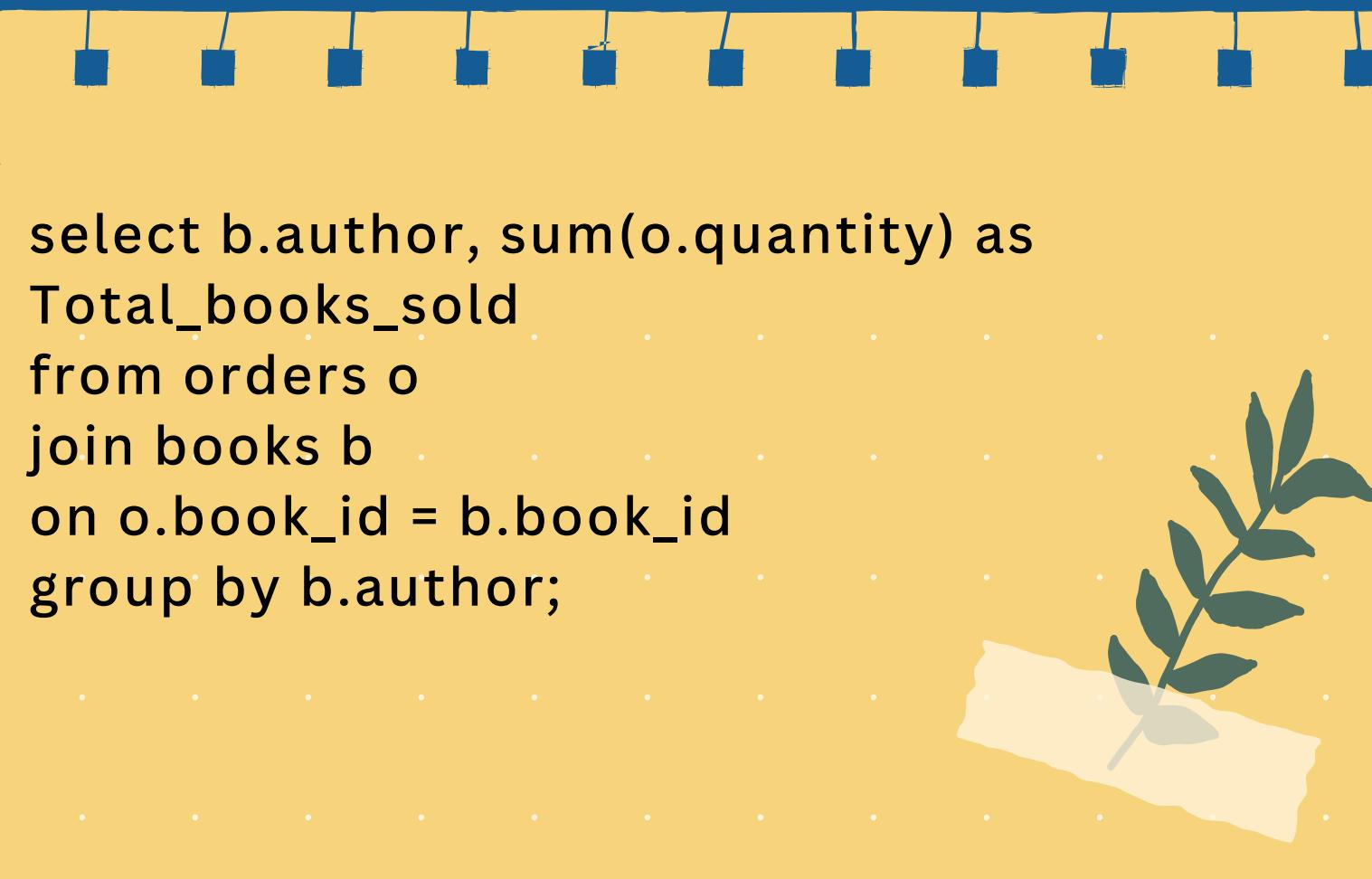
	order_id [PK] integer	customer_id integer	book_id integer	order_date date	quantity integer	total_amount numeric (10,2)
1	1	84	169	2023-05-26	8	188.56
2	2	137	301	2023-01-23	10	216.60
3	3	216	261	2024-05-27	6	85.50
4	4	433	343	2023-11-25	7	301.21
5	5	14	431	2023-07-26	7	136.36
6	6	439	119	2024-10-11	5	249.40
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	123.00

Total rows: 473 of 473    Query complete 00:00:00.082

# INVENTORY INSIGHTS

Query 8 : Retrieve the total quantity of books sold by each author.

```
select b.author, sum(o.quantity) as  
Total_books_sold  
from orders o  
join books b  
on o.book_id = b.book_id  
group by b.author;
```



	author character varying (100)	total_books_sold bigint
1	Jared Cortez	10
2	Tracy Parker	11
3	Taylor Wang	9
4	Cathy Knight	6
5	Bianca Matthews	3
6	Douglas Malone	6
7	James Alvarado	9
8	Betty Cross	6
9	Michael Hill	20
10	Steven McDonald	15

Total rows: 314 of 314

Query complete 00:00:00.06

# INVENTORY INSIGHTS

Query 9 : Find the book with the lowest stock.

```
select * from books  
order by stock asc  
limit 1;
```



	book_id [PK] integer	title character varying (100)	author character varying (100)	genre character varying (50)	published_year integer	price numeric (10,2)	stock integer
1	44	Networked systemic implementation	Ryan Frank	Science Fiction	1965	13.55	0

# INVENTORY INSIGHTS

Query 10 : Calculate the stock remaining after fulfilling all orders.

```
select b.book_id, b.title, b.stock as  
Initial_Stock,  
coalesce(sum(o.quantity),0) as Ordered_Stock,  
b.stock - coalesce(sum(o.quantity),0) as  
Remaining_Stock  
from books b  
left join orders o  
on b.book_id = o.book_id  
group by b.book_id  
order by b.book_id;
```



	book_id [PK] integer	title character varying (100)	initial_stock integer	ordered_stock bigint	remaining_stock bigint
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
9	9	Optimized interactive challenge	70	0	70
10	10	Ergonomic national hub	25	1	24

Total rows: 500 of 500    Query complete 00:00:00.277

# KEY INSIGHTS & RECOMMENDATIONS

## INSIGHTS

- 139 repeat customers → Strong potential for customer retention.
- 443 high-spending users → Most located in urban cities.
- Top customer spent \$1,398.90 → Significant outlier in revenue.
- 'Robust Tangible Hardware' is the most frequently ordered book.
- Total revenue : \$75,628.66.

## RECOMMENDATIONS

- Launch a loyalty program to retain repeat buyers.
- Focus targeted ads on high-value cities.
- Create VIP offers for top spenders.
- Promote bestsellers through bundles or homepage highlights.
- Use monthly revenue tracking to spot growth trends.

**THANK  
YOU VERY  
MUCH!**