



Online Bookstore Project

on SQL

Presented by: Himanshi Chauhan



Objective

The aim of this project is to analyze data from an online bookstore using SQL. The focus is on understanding customer behavior, book sales, and order patterns to gain business insights.



About the Dataset



Books

Details of books like
book ID, title, author,
genre, published year,
price, stock



Customers

Customer information
such as customer ID,
name, email, phone,
city, country, location



Orders

Order- related data like
order ID, customer ID,
book ID, quantity,
order date and total
amount



Books

Create Tables

Customers

```
1 Create Table Books(Book_ID Serial Primary Key,  
2   Title Varchar(100),  
3   Author Varchar(100),  
4   Genre Varchar(50),  
5   Published_Year Int,  
6   Price Numeric(10, 2),  
7   Stock Int);
```

```
9 Create Table Customers(  
10   Customer_ID Serial Primary Key,  
11   Name Varchar(100),  
12   Email Varchar(100),  
13   Phone Varchar(15),  
14   City Varchar(50),  
15   Country Varchar(150));
```

```
17 Create Table Orders(  
18   Order_ID Serial Primary Key,  
19   Customer_ID Int References Customers(Customer_ID),  
20   Book_ID Int References Books(Book_ID),  
21   Order_Date Date,  
22   Quantity Int,  
23   Total_Amount Numeric(10, 2));
```

Orders



Import Data

```
25 -- Import Data into Books Table
26 Copy Books from 'C:\Program Files\PostgreSQL\16\data\data resource\Books.csv' csv header;
27
28 --Import Data into Customers Table
29 Copy Customers from 'C:\Program Files\PostgreSQL\16\data\data resource\Customers.csv' csv header;
30
31 -- Import Data into Orders Table
32 Copy Orders from 'C:\Program Files\PostgreSQL\16\data\data resource\Orders.csv' csv header;
```



1) Retrieve all books in the "Fiction" genre:

```
38 select * from Books
39 where Genre = 'Fiction';
```

	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	4	Customizable 24hour product	Christopher Andrews	Fiction	2020	43.52	8
2	22	Multi-layered optimizing migration	Wesley Escobar	Fiction	1908	39.23	78
3	28	Expanded analyzing portal	Lisa Coffey	Fiction	1941	37.51	79
4	29	Quality-focused multi-tasking challenge	Katrina Underwood	Fiction	1905	31.12	100
5	31	Implemented encompassing conglomerati...	Melissa Taylor	Fiction	2010	21.23	44
6	39	Optimized national process improvement	Megan Goodwin	Fiction	1978	10.99	42
7	40	Adaptive didactic interface	Natalie Gonzalez	Fiction	1923	25.97	94
8	47	Reverse-engineered directional conglomer...	John Christian	Fiction	2006	20.37	90
9	62	Re-contextualized real-time strategy	Nicole Lynch	Fiction	1953	26.34	23
10	63	Polarized heuristic database	Franklin Mack	Fiction	1989	22.38	56



Basic Questions

2) Find books published after the year 1950:

```
42 Select * from Books
43 where Published_Year >1950;
```

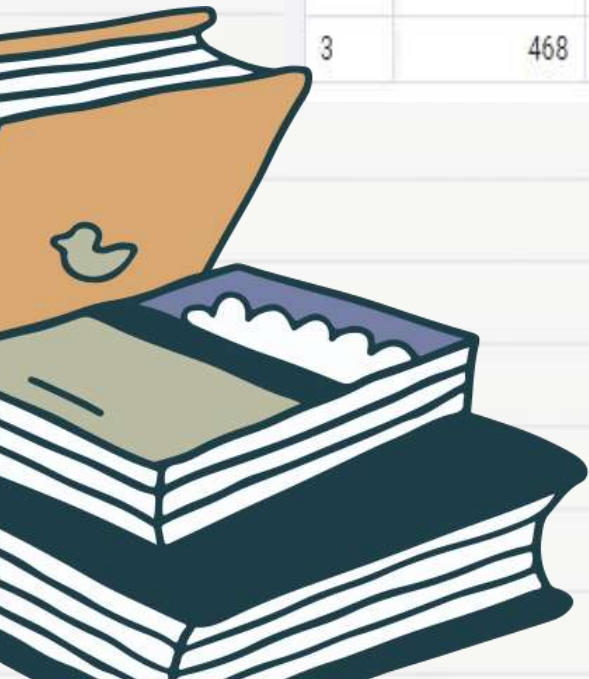
	book_id [PK] integer	title character varying (100)	author character varying (100)	genre character varying (50)	published_year integer	price numeric (10,2)	stock integer
1	2	Persevering reciprocal knowledge user	Mario Moore	Fantasy	1971	35.80	1
2	4	Customizable 24hour product	Christopher Andrews	Fiction	2020	43.52	
3	5	Adaptive 5thgeneration encoding	Juan Miller	Fantasy	1956	10.95	1
4	6	Advanced encompassing implementation	Bryan Morgan	Biography	1985	6.56	
5	8	Persistent local encoding	Troy Cox	Science Fiction	2019	48.99	8
6	9	Optimized interactive challenge	Colin Buckley	Fantasy	1987	14.33	7
7	10	Ergonomic national hub	Samantha Ruiz	Mystery	2015	24.63	2
8	11	Secured zero tolerance time-frame	Denise Barnes	Fantasy	1998	35.95	1
9	12	Polarized optimal array	Destiny Scott	Non-Fiction	1989	27.43	6



3) List all customers from the Canada:

```
46  Select * from Customers
47  where Country = 'Canada';
```

	customer_id [PK] integer	name character varying (100)	email character varying (100)	phone character varying (15)	city character varying (50)	country character varying (150)
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



4) Show orders placed in November 2023:

```
50 select * from Orders
51 where Order_Date between '2023-11-01' and '2023-11-30';
52
```

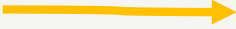
	order_id [PK] integer	customer_id integer	book_id integer	order_date date	quantity integer	total_amount numeric (10,2)
1	4	433	343	2023-11-25	7	301.21
2	19	496	60	2023-11-17	9	316.26
3	75	291	375	2023-11-30	5	170.75
4	132	469	333	2023-11-22	7	194.32
5	137	474	471	2023-11-25	8	363.04
6	163	207	384	2023-11-23	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



5) Retrieve the total stock of books available:

```
54 Select sum(Stock) as Total_Stock from Books;
```

```
55
```

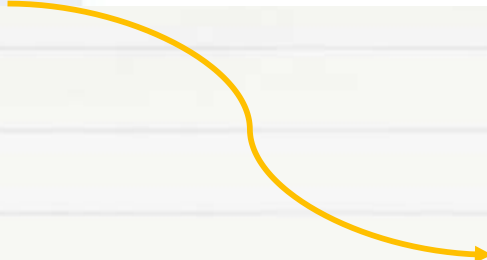


	total_stock
1	25056

6) List all genres available in the Books table:

```
70 Select distinct Genre from Books;
```

```
71
```



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



7) Find the details of the most expensive book:

```
57  Select * from Books
58  order by Price desc
59  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	340	Proactive system-worthy orchestration	Robert Scott	Mystery	1907	49.98	88



8) Show all customers who ordered more than 1 quantity of a book:

```
62 Select * from Orders
63 where Quantity >1;
```

	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



9) Retrieve all orders where the total amount exceeds \$20:

```
66 select * from Orders
67 where Total_amount>20;
68
```

	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



10) Find the book with the lowest stock:

```
73 Select * from Books
74 order by Stock 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

11) Calculate the total revenue generated from all orders:

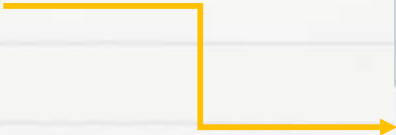
```
77 Select sum(Total_amount) as Revenue
78 from Orders;
```

	revenue numeric
1	75628.66



1) Retrieve the total number of books sold for each genre:

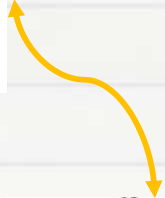
```
83 Select b.Genre, sum(o.Quantity) as Total_Books_sold
84 from Orders o
85 join Books b on o.book_id = b.book_id
86 group by b.Genre;
```



	genre character varying (50)	total_books_sold bigint
1	Romance	439
2	Biography	285
3	Mystery	504
4	Fantasy	446
5	Fiction	225
6	Non-Fiction	351
7	Science Fiction	447

2) Find the average price of books in the "Fantasy" genre:

```
89 Select avg(Price) as Avg_Price_Books
90 from Books
91 Where Genre='Fantasy';
92
```



	avg_price_books numeric
1	25.9816901408450704



3) List customers who have placed at least 2 orders:

```

94 Select o.Customer_ID, c.name, count(o.Order_ID) as order_count
95 from Orders o
96 join Customers c on c.Customer_ID=o.Customer_ID
97 group by o.customer_id, c.name
98 having count(Order_id) >=2;

```

	customer_id integer	name character varying (100)	order_count bigint
1	225	Christopher Mccullough	2
2	418	Kiara Blankenship MD	3
3	322	William Cameron	3
4	325	Emily Vargas	4
5	376	Justin Donaldson	2
6	486	Melanie Kelly	2
7	461	Crystal Pierce	3
8	2	Crystal Clements	2
9	149	Jason Robinson	3
10	173	Victoria Dixon	2

4) Find the most frequently ordered book:

```

101 Select o.Book_ID, b.Title, count(o.Order_ID) as Order_count
102 from Orders o
103 join Books b on b.Book_ID=o.Book_ID
104 group by o.Book_ID, b.Title
105 order by Order_count desc limit 1;
106

```

	book_id integer	title character varying (100)	order_count bigint
1	88	Robust tangible hardware	4



5) Show the top 3 most expensive books of 'Fantasy' Genre :

```
108 Select * from Books
109 where Genre='Fantasy'
110 order by Price desc limit 3;
111
```

	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

6) Retrieve the total quantity of books sold by each author:

```
113 Select b.Author,sum(o.Quantity) as Total_books_sold
114 from Orders o
115 join Books b on b.Book_ID=o.Book_ID
116 group by b.Author
117 order by Total_books_sold desc;
118
```

	author character varying (100)	total_books_sold bigint
1	Patrick Contreras	28
2	Melissa Taylor	27
3	Emily James	24
4	Thomas Trujillo	24
5	Ellen Doyle	23
6	Sheena Harris	23
7	Erica Parker	23
8	Valerie Moore	23
9	Rachel Gibbs	22
10	Amanda Wilson	22

7) List the cities where customers who spent over \$30 are located:

```
120 Select Distinct c.City, o.Total_amount
121 from Orders o
122 join Customers c on o.Customer_ID=c.Customer_ID
123 where o.Total_amount>30;
124
```

	city character varying (50)	total_amount numeric (10,2)
1	Taylorfurt	189.45
2	Leeport	141.39
3	Port Jasonview	149.12
4	Port Aaronstad	145.44
5	Matthewfurt	328.50
6	Angelaside	42.19
7	Lindaburgh	325.92
8	Stephanieberg	156.60
9	Freemanland	198.75
10	Natashaville	399.04

8) Find the customer who spent the most on orders.

```
126 Select c.Customer_ID,c.Name,sum(Total_amount) as Total_spent
127 from orders o
128 join Customers c on c.Customer_ID=o.Customer_ID
129 group by c.Customer_ID,Name
130 order by Total_spent desc limit 1;
131
```

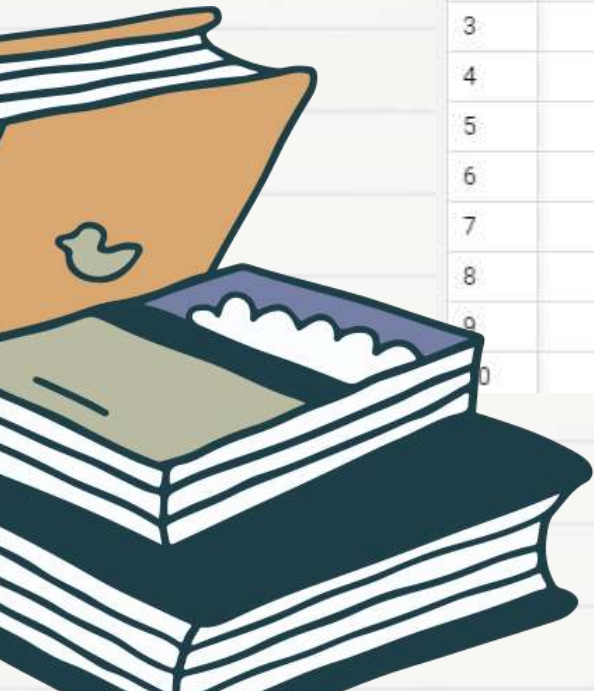
	customer_id [PK] integer	name character varying (100)	total_spent numeric
1	457	Kim Turner	1398.90



9) Calculate the stock remaining after fulfilling all orders:

```
133 Select b.Book_ID, b.Title, b.Stock, Coalesce(Sum(o.Quantity),0) as Order_quantity,  
134 b.Stock- Coalesce(sum(o.Quantity),0) as Remaining_quantity  
135 from Books b  
136 left join Orders o on b.Book_ID=o.Book_ID  
137 group by b.Book_ID  
138 order by b.Book_ID;
```

	book_id [PK] integer	title character varying (100)	stock integer	order_quantity bigint	remaining_quantity 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



Thanks For
Watching

