

Online Bookstore SQL Project

A portfolio project analyzing sales, customer behavior, and inventory using PostgreSQL.

[**Javed Hussain, Data Analyst**](#)

Project Objectives

-  **Build Database:** Design a relational schema, create tables, and import raw CSV data into PostgreSQL.
-  **Analyze Data:** Write complex SQL queries to join tables and aggregate data for analysis.
-  **Extract Insights:** Identify actionable business insights related to sales trends, customer behavior, and inventory management.
-  **Visualize Results:** Translate raw statistical data from SQL queries into clear, professional charts.

Technology & Database Schema



Books Table (PK: Book_ID)

Title Author Genre Published_Year
Price Stock



Customers Table (PK: Customer_ID)

Name Email Phone City Country



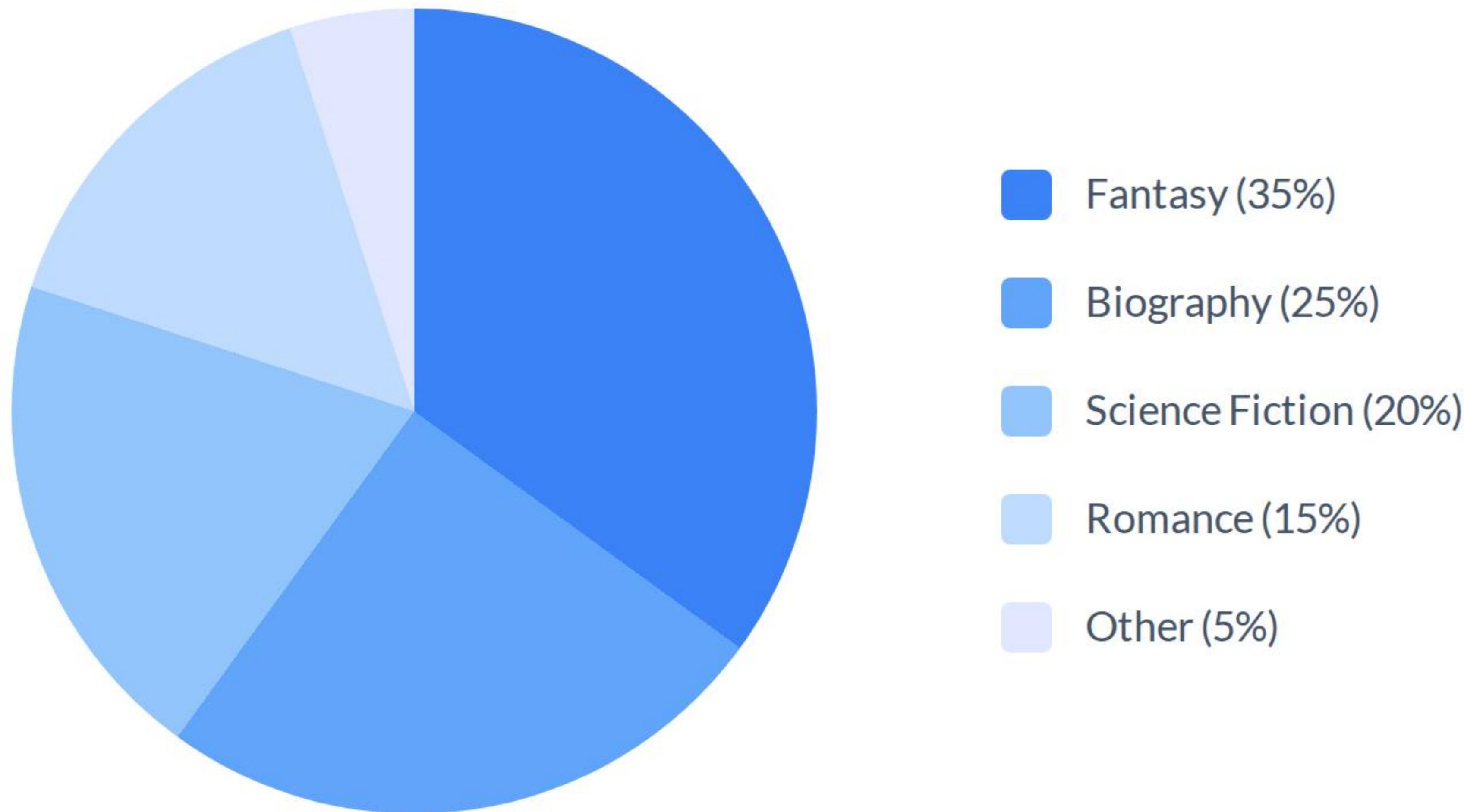
Orders Table (PK: Order_ID)

Customer_ID (FK) Book_ID (FK)
Order_Date Quantity Total_Amount

Business Insights & Analysis

Using SQL to translate raw data into strategic insights and statistics.

Analysis 1: Top Selling Genres by Quantity



The 'Fantasy' genre is the dominant seller, accounting for over a third of all books sold, suggesting a strong area for marketing focus.

SQL for Genre Analysis

Business Question

What is the total number of books sold for each genre, sorted by popularity?

SQL Query (JOIN & GROUP BY)

```
SELECT b.genre, SUM(o.quantity) AS total_sold FROM Books AS b JOIN Orders AS o ON b.book_id = o.book_id GROUP BY b.genre ORDER BY total_sold DESC;
```

Analysis 2: Top 5 Bestselling Books by Order Count



Identifying specific high-frequency sellers allows for targeted stock management and promotion of similar titles.

SQL for Bestseller Analysis

Business Question

What are the top 5 most frequently ordered books across all transactions?

SQL Query (JOIN, GROUP BY, LIMIT)

```
SELECT o.book_id, b.title, COUNT(o.order_id) AS  
order_count FROM Books AS b JOIN Orders AS o ON  
b.book_id = o.book_id GROUP BY o.book_id, b.title  
ORDER BY order_count DESC LIMIT 5;
```

Analysis 3: Top 5 Authors by Total Units Sold



Author popularity is a key driver of sales. This data can inform which authors to feature in marketing campaigns.

SQL for Author Analysis

Business Question

Which authors have sold the most total books (quantity)?

SQL Query (JOIN & GROUP BY)

```
SELECT b.author, SUM(o.quantity) AS  
total_units_sold FROM Books AS b JOIN Orders AS o  
ON b.book_id = o.book_id GROUP BY b.author ORDER BY  
total_units_sold DESC LIMIT 5;
```

Analysis 4: Top 5 High-Value Customers by Spend



Identifying top spenders is vital for loyalty programs and personalized marketing to high-value customers.

SQL for Customer Value Analysis

Business Question

Find the top 5 customers who have spent the most money in the store.

SQL Query (JOIN & GROUP BY)

```
SELECT c.customer_id, c.name, SUM(o.total_amount)
AS total_spent FROM Customers AS c JOIN Orders AS o
ON c.customer_id = o.customer_id GROUP BY
c.customer_id, c.name ORDER BY total_spent DESC
LIMIT 5;
```

Analysis 5: Top 5 Cities by Total Customer Spend



Geographic sales data reveals high-performing regions, perfect for targeted local advertising or pop-up events.

SQL for Geographic Analysis

Business Question

Which cities generate the most revenue? List all cities that have collectively spent over \$460.

SQL Query (GROUP BY & HAVING)

```
SELECT c.city, SUM(o.total_amount) AS total_spent  
FROM Customers AS c JOIN Orders AS o ON  
c.customer_id = o.customer_id GROUP BY c.city  
HAVING SUM(o.total_amount) > 460 ORDER BY  
total_spent DESC;
```

Analysis 6: Key Store-wide Metrics

\$26.45

Avg. Price of 'Fantasy' Books

24,800

Total Units in Stock

SQL for Key Metrics

Avg. Price (Fantasy)

Calculates the average price for all books in the 'Fantasy' genre.

```
SELECT AVG(price) AS avg_fantasy_price FROM Books  
WHERE genre = 'Fantasy';
```

Total Stock

Calculates the total number of all books available in inventory.

```
SELECT SUM(stock) AS total_stock FROM Books;
```

Technical SQL Showcase

Demonstrating advanced queries for inventory management and complex filtering.

Technical 1: Inventory Management (Remaining Stock)

Book ID	Title	Original Stock	Total Sold	Remaining Stock
169	Networked tertiary approach	45	8	37
301	Polarized high-level installation	77	13	64
261	Intuitive content-based toolset	52	6	46
343	De-engineered grid-enabled...	86	16	70
119	Switchable modular moratorium	76	13	63

SQL for Inventory (LEFT JOIN)

Business Question

Calculate the remaining stock for every book after all orders are fulfilled.
This query is crucial for identifying which books to re-order.

SQL Query (LEFT JOIN & COALESCE)

```
SELECT b.book_id, b.title, b.stock,  
COALESCE(SUM(o.quantity), 0) AS sold_quantity,  
(b.stock - COALESCE(SUM(o.quantity), 0)) AS  
remaining_stock FROM Books AS b LEFT JOIN Orders AS  
o ON b.book_id = o.book_id GROUP BY b.book_id,  
b.title, b.stock;
```

Technical 2: Finding Repeat Customers (HAVING)

Business Question

How can we find our loyal customers? List all customers who have placed two or more separate orders.

SQL Query (GROUP BY & HAVING)

```
SELECT c.customer_id, c.name, COUNT(o.order_id) AS  
order_count FROM Customers AS c JOIN Orders AS o ON  
c.customer_id = o.customer_id GROUP BY  
c.customer_id HAVING COUNT(o.order_id) >= 2;
```

Technical 3: Basic Filtering (WHERE)

Question

List all customers who are located in 'Canada'.

SQL Query

```
SELECT customer_id, name, city, country FROM  
Customers WHERE country = 'Canada';
```

Technical 4: Date Filtering (BETWEEN)

Question

Show all orders placed within the month of November 2023.

SQL Query

```
SELECT * FROM Orders WHERE order_date BETWEEN  
'2023-11-01' AND '2023-11-30';
```

Conclusion & Future Scope

This SQL project successfully demonstrates the end-to-end process of database management and analysis. The insights gathered from statistical analysis can directly inform inventory, marketing, and customer relationship strategies for the online bookstore.

Future Scope:

1. **Time-Series Analysis:** Analyze sales patterns by month or quarter.
2. **Customer Segmentation:** Perform RFM (Recency, Frequency, Monetary) analysis.
3. **Recommendation Engine:** Find which books are frequently purchased together.

About the Analyst



Javed Hussain

Data Analyst passionate about transforming raw data into clear, actionable insights and visual stories.

Skilled in SQL, data cleaning, and visualization, I enjoy tackling complex problems and building data-driven solutions.

Thank You

Let's Connect!

Javed Hussain | Data Analyst

 linkedin.com/in/ijavedhussain

 github.com/Jawebdata