**Online Bookstore Sales Analysis using SQL**

**Project Overview**

This project is a deep dive into the sales data of an online bookstore. The primary objective is to use SQL to query the bookstore's database to uncover key insights into sales performance, customer behavior, and inventory management. By analyzing data across books, customers, and orders, this project aims to answer critical business questions that can help guide marketing strategies, optimize stock levels, and improve overall revenue.

The analysis is broken down into two main parts: **Foundational Analysis** to gather basic metrics and **In-Depth Analysis** to uncover more complex trends and patterns.

**Database Schema**

The project utilizes a relational database designed to simulate a real-world online bookstore. It consists of three primary tables:

* **Books**: Contains details about each book available for sale.
* **Customers**: Stores information about the bookstore's registered customers.
* **Orders**: Records all transaction details, linking customers to the books they purchase.

**Table Structures**

1.

Books Table

| Column | Data Type | Description |
| --- | --- | --- |
| Book\_ID | SERIAL | Unique identifier for each book (Primary Key) |
| Title | VARCHAR(100) | The title of the book |
| Author | VARCHAR(100) | The author of the book |
| Genre | VARCHAR(50) | The genre or category of the book |
| Published\_Year | INT | The year the book was published |
| Price | NUMERIC(10, 2) | The selling price of the book |
| Stock | INT | The current number of units in stock |

2.

Customers Table

| Column | Data Type | Description |
| --- | --- | --- |
| Customer\_ID | SERIAL | Unique identifier for each customer (Primary Key) |
| Name | VARCHAR(100) | The full name of the customer |
| Email | VARCHAR(100) | The customer's email address |
| Phone | VARCHAR(15) | The customer's phone number |
| City | VARCHAR(50) | The city where the customer resides |
| Country | VARCHAR(150) | The country where the customer resides |

3.

Orders Table

| Column | Data Type | Description |
| --- | --- | --- |
| Order\_ID | SERIAL | Unique identifier for each order (Primary Key) |
| Customer\_ID | INT | Foreign Key referencing Customers.Customer\_ID |
| Book\_ID | INT | Foreign Key referencing Books.Book\_ID |
| Order\_Date | DATE | The date the order was placed |
| Quantity | INT | The number of units of a book purchased |
| Total\_Amount | NUMERIC(10, 2) | The total cost of the order |

**Data Analysis & Insights**

The analysis was performed by answering a series of questions, divided into two sections.

**Part 1: Foundational Analysis (Basic Queries)**

This section covers fundamental queries to get a high-level overview of the data.

* **Retrieve books by genre ("Fiction")**: Filters the inventory to understand the offerings within a specific category.
* **Find books published after 1950**: Helps in analyzing the age and relevance of the book inventory.
* **List customers from a specific country (Canada)**: Useful for geographic market segmentation and targeted marketing.
* **Show orders from a specific month (November 2023)**: Allows for tracking sales performance over specific time periods.
* **Calculate total available stock**: A key metric for inventory management and assessing asset value.
* **Find the most expensive book**: Identifies premium products in the catalog.
* **Identify customers ordering more than one item**: Helps in identifying potentially high-value or engaged customers
* **Find orders exceeding a certain value ($20)**: Segments high-value transactions.
* **List all available genres**: Provides an overview of the diversity of the book collection.
* **Find the book with the lowest stock**: Critical for identifying items that need immediate restocking.
* **Calculate total revenue from all orders**: The primary KPI for overall business performance.

**Part 2: In-Depth Analysis (Advanced Queries)**

This section uses more complex queries to derive deeper business intelligence.

* **Retrieve total books sold per genre**: Identifies the most popular genres, guiding purchasing and marketing efforts.
* **Find the average price of "Fantasy" books**: Helps in pricing strategy and understanding the market value of a specific category.
* **List customers with at least 2 orders**: Pinpoints repeat customers, who are crucial for loyalty programs and retention strategies.
* **Find the most frequently ordered book**: Identifies the bookstore's "bestseller" and key product.
* **Show the top 3 most expensive "Fantasy" books**: Narrows down on premium products within a popular genre.
* **Retrieve total books sold by each author**: Measures author popularity and their contribution to sales.
* **List cities of customers who spent over $30**: Identifies key geographic markets with high-spending customers.
* **Find the customer who spent the most**: Identifies the top VIP customer, who could be targeted for special offers.
* **Calculate remaining stock after all orders**: A crucial inventory management query to get a real-time view of available stock vs. sold units.

**Key SQL Concepts Demonstrated**

This project showcases a wide range of SQL skills, including:

* **Data Definition Language (DDL)**: Creating tables and defining the database schema.
* **Data Manipulation Language (DML)**: Retrieving, filtering, and sorting data.
* **Aggregate Functions**: Using SUM(), AVG(), COUNT(), MAX(), and MIN() to perform calculations.
* **Joins**: Combining data from multiple tables (JOIN, LEFT JOIN) to create comprehensive views.
* **Subqueries**: Nesting queries to perform multi-step filtering and analysis, such as finding the most expensive book or the highest-spending customer.
* **Grouping and Filtering**: Using GROUP BY and HAVING to segment data and analyze specific cohorts.
* **Common Table Expressions (CTEs)**: Using the WITH clause to simplify complex queries, such as calculating the remaining stock.
* **Conditional Logic**: Employing CASE WHEN statements to create custom logic for reporting, like showing "Not ordered" for books with no sales.

**Author**

* **Parveen Jalwal**
* **LinkedIn**: <https://www.linkedin.com/in/parveen-jalwal-201a2a302>