📘 SQL Project Report – Online Bookstore

Author: Md Tanvir Hasan, Business Intelligence Analyst

1. Introduction & Objectives

The **Online Bookstore SQL Project** was designed to demonstrate how structured query language (SQL) can be leveraged for:

- Extracting **business intelligence** from raw transaction data.
- Analyzing **sales, customer behavior, and inventory performance**.
- Supporting **data-driven decision making** in the e-commerce/book retail sector.

The project addresses **core business challenges**:

- Which **books and authors drive revenue**?
- How do **customers behave** (loyalty, churn, preferences)?
- Which **genres perform best** across different markets/seasons?
- How can inventory and pricing be **optimized for profitability**?

---

2. Database Design & Structure

The database consists of **three main entities**:

Books Table

- Attributes: BookID, Title, Author, Genre, Price, Stock
- Purpose: Store book details, pricing, and availability

Customers Table

- Attributes: CustomerID, Name, City, Country, JoinDate
- Purpose: Track demographics, location, and account creation

Orders Table

- Attributes: OrderID, CustomerID, BookID, Quantity, OrderDate, TotalAmount
- Purpose: Capture transactions linking customers and books

Relationships:

- One customer → Many orders
- One book → Many orders
- Books & Customers are connected through Orders

---

3. SQL Queries Executed

A total of **50+ queries** were written, divided into categories:

- **Basic Queries:** Retrieve most expensive books, cheapest books, stock levels.
- **Intermediate Queries:** Revenue by genre, top-selling authors, order frequencies.
- **Advanced Queries:** Customer segmentation, RFM analysis, churn prediction, basket analysis, seasonal sales patterns.
- **Business Optimization Queries:** Price elasticity, inventory optimization, price-point analysis, retention metrics.

---

4. Detailed Insights & Interpretations

A. Inventory & Books

- Found **low-stock books** that need urgent replenishment → prevents lost sales.
- Identified **books never ordered** → indicates poor catalog fit or lack of marketing.
- Price analysis showed **some books priced above genre average** → useful for premium positioning.
- **Top 3 expensive fantasy books** reveal a niche luxury segment → opportunity to promote as collector's items.

B. Customers

- **RFM Analysis** identified high-value loyal customers vs at-risk customers.
- Customers segmented into **Active (recent), Lapsing (60–90 days), Dormant (>90 days)** → actionable for targeted campaigns.
- **Geographic segmentation** showed revenue concentrated in a few cities → can guide regional promotions.
- **Behavioral segmentation** highlighted "Focused buyers" (stick to 1 genre) vs "Varied buyers" (cross-genre potential).

C. Sales & Revenue

- **Top-selling genres:** Consistently performed genres across all quarters.
- **Revenue share by authors:** Identified "star authors" driving majority of sales.
- **Month-over-month sales growth** highlighted seasonality (e.g., academic books peak before school terms).
- **High-value recent customers** contributed disproportionately → suggests prioritizing retention programs.
- **Basket analysis** → customers buying "Book A" often also buy "Book B" → enables recommendations.

D. Genre & Author Trends

- **Genre performance shifts by season:** e.g., Fiction strong in summer, Academic books strong in spring.
- **Author trend analysis:** Some authors showed rising demand while others declined → useful for contract/partnership negotiations.
- **Multi-genre authors** performed better in terms of customer retention → signals value in cross-genre publishing.

E. Advanced Business Insights

- **Price elasticity analysis** showed mid-price range books sell fastest → helps in setting optimal price brackets.
- **Customer retention metrics**: Cohort analysis by join month → revealed declining engagement after 6 months → need for loyalty programs.
- **Inventory optimization**: Matched stock vs demand velocity → avoid overstocking slow movers and understocking bestsellers.
- **Revenue concentration**: Top 10% of books contributed ~60–70% revenue → Pareto principle applies.

---

5. Business Value

The project provides clear **business intelligence applications**:

1. **Marketing & Retention**
    - Identify churn-risk customers for reactivation campaigns.
    - Personalize promotions based on genre preference.
2. **Inventory Planning**
    - Focus restocking on high-demand books.
    - Reduce capital lock-up in dead stock.
3. **Revenue Growth**
    - Bundle frequently bought-together books.
    - Prioritize top-performing authors and genres.
4. **Strategic Decisions**
    - Negotiate better deals with star authors.
    - Expand operations in high-revenue cities.
    - Adjust pricing tiers based on elasticity findings.
