Online Store Analytics – SQLite Database Report

Project Description

This project involves the design, creation, and population of a relational SQLite database for an online store. The database captures key aspects of the store, including customers, products, and orders, providing a platform to analyze customer behavior, sales trends, and revenue performance.

The primary objective is to demonstrate practical skills in SQL, database design, and data management using Python. The simplified structure ensures focus on core analytics while maintaining relational integrity through primary and foreign keys. All data is synthetically generated ensuring privacy compliance.

Data Generation Process

Data for this project was programmatically generated using Python, the Faker library for realistic names and dates and random for numeric and categorical attributes. No external datasets were used.

1. Customers Table: Generates names, ages, genders, and loyalty levels. Loyalty level is an ordinal variable ranging from Bronze to Platinum.
2. Products Table: Stores product information including category, price and stock. Categories are nominal while price and stock are ratio level data.
3. Orders Table: Links customers to products via purchase transactions. Each order contains a list of purchased products, total amount, payment method and status. Status is ordinal while total amount is ratio level data. Order date is treated as interval data.

This approach ensures over 1000 rows of realistic, synthetic data per table with foreign key relationships maintaining referential integrity. Python scripts automate data insertion into the SQLite database *online_store.db*.

Database Schema and Tables

The database is structured with three main tables, designed with constraints to maintain data integrity and ensure accurate relationships.

- Customers Table

Attribute	Type	Constraint
customer_id	Integer	PRIMARY KEY
name	Text	NOT NULL

age	Integer	CHECK(age >= 0)
gender	Text	CHECK Constraint
loyalty_level	Text	CHECK Constraint

- Products Table

Attribute	Type	Constraint
product_id	Integer	PRIMARY KEY
name	Text	NOT NULL
category	Text	CHECK Constraint
price	Real	CHECK(price >= 0)
stock_quantity	Integer	CHECK Constraint

- Orders Table

Attribute	Type	Constraint
order_id	Integer	PRIMARY KEY
customer_id	Integer	FOREIGN KEY REFERENCES Customers(customer_id)
order_date	Text	NOT NULL
products_ordered	Text	JSON or comma-separated list of product IDs and quantities
total_amount	Real	CHECK Constraint
payment_method	Text	CHECK Constraint
status	Text	CHECK Constraint

Conclusion

This project demonstrates the creation of a simplified online store database using SQLite, with three interrelated tables: Customers, Products, and Orders. The database supports analytics on customer behavior, product sales and revenue trends while maintaining referential integrity and realistic synthetic data. The Python script automates the generation and insertion of data, ensuring reproducibility, scalability and privacy compliance. This project provides a practical example of relational database design, synthetic data generation and foundational data analytics for e-commerce scenarios.

Appendices:

1.