**Project Summary: Ecommserce Sales Analysis using MySQL**

**🔹 Project Overview:**

This project focuses on analyzing ecommerce sales data using **MySQL** for deeper insights.  
The dataset includes information about products, pricing, reviews, and monthly sales over a year.

Key goals:

* Extract meaningful patterns in product sales
* Understand category-wise performance
* Identify top-performing and high-rated products
* Optimize SQL queries for fast and efficient analysis

**🔹 Dataset Details:**

* **File Name:** ecommerce\_sales\_analysis.csv
* **Main Columns:**
  + product\_id, product\_name
  + category, price
  + review\_score, review\_count
  + sales\_month\_1 → sales\_month\_12 (Monthly sales from January to December)

**🔹 Technologies Used:**

* **Database:** MySQL 8.x
* **Language:** SQL
* **Libraries (Optional):** Pandas (for offline testing), Matplotlib (for visualization)
* **Tools:** MySQL Workbench / phpMyAdmin

**🔹 Core Steps Implemented:**

**1. Database and Table Setup**

* Created the ecommerce\_sales table to store the product data.

**2. Data Import**

* Loaded the CSV file into MySQL using LOAD DATA INFILE or manual import.

**3. Data Analysis using SQL**

✅ Simple Queries:

* Retrieve top-priced products.
* Filter products with high review scores.
* Calculate total yearly sales for each product.

✅ Grouping and Aggregation:

* Category-wise sales performance.
* Company-wide total revenue calculation.

✅ Joins:

* Simulated by creating a separate category\_master table for advanced joins.

✅ Subqueries:

* Identified products priced higher than average.
* Found top categories based on total sales.

✅ Views:

* Created a view category\_sales\_view for easy reuse of aggregated data.

✅ Indexes:

* Applied indexing on category, price, and review\_score to speed up performance.

**🔹 Key Insights Found:**

* **Top-Selling Product:** Product\_168 with 10,685 sales.
* **Highest Revenue Category:** Clothing (236,361 units sold).
* **Top Rated Products:** Multiple products with perfect 5.0 review scores.
* **Average Review Score Across All Products:** 3.27 ⭐
* **Total Revenue:** Approximately **$59.45 million** 💰

**🔹 Challenges Faced:**

* Dataset did not have customer and order history, so joins were simulated manually.
* Handling 12 months of sales columns required careful summing operations.
* Optimizing queries for faster analysis on larger datasets.

**🚀 Final Status: Completed Successfully**

This project successfully shows how **real-world ecommerce data** can be cleaned, analyzed, and optimized entirely inside **MySQL** using professional database techniques.