## Executive Summary: E-commerce Data Analysis with Python and SQL

This project showcases a comprehensive analysis of an e-commerce dataset, integrating advanced SQL queries and Python for data transformation, exploration, and visualization. The analysis uncovered actionable insights into customer behavior, sales trends, and business performance.

**1. Data Transformation and Preparation**

Using Python, I automated the process of transforming raw CSV data into a MySQL database, ensuring data integrity and consistency. Key steps included:

- **Imported and cleaned 7 datasets** (e.g., `customers`, `orders`, `payments`, `products`, `geolocation`) containing over **1 million rows** combined.
- Implemented a script to dynamically infer SQL data types based on column characteristics (e.g., `TEXT`, `INT`, `FLOAT`, `DATETIME`).
- Handled missing values by replacing `NaN` with `NULL`, ensuring compatibility with MySQL.

**2. Advanced SQL Analysis:**

- Conducted in-depth analyses using advanced SQL techniques like window functions, CTEs (Common Table Expressions), and aggregation.
- Examples include calculating Year-over-Year (YoY) growth, cumulative sales, moving averages, and customer retention rates over a 36-month window.

**3. Insights and Visualizations:**

- Leveraged Python libraries (e.g., Matplotlib, Seaborn) to create meaningful visualizations, showcasing trends in sales, customer behavior, and seasonal performance.
- Provided strategic insights on customer purchasing patterns, top-performing products, and seller contributions.

**4. Results and Impact:**

- Identified key business drivers, such as high-value customers and growth opportunities.
- Delivered a data-driven approach for enhancing customer retention, optimizing product offerings, and increasing sales revenue.

This project reflects my ability to integrate technical expertise in SQL and Python with analytical skills to deliver impactful solutions for real-world business problems.