Project Title: Adidas US Sales Performance Analysis

Project Summary: Conducted a multi-stage analysis of a 9,600+ record dataset to evaluate Adidas' market performance across the United States. This project integrates Python for data cleaning and EDA, MySQL for advanced querying and database management, and Power BI for interactive visualization to uncover regional trends, product profitability, and channel efficiency.

---

1. Technical Stack Used

- **Data Analysis:** Python (Pandas, NumPy, Matplotlib, Seaborn)
- **Database Management:** SQL (MySQL)
- **Data Visualization:** Power BI
- **Database Connectivity:** SQLAlchemy (Python-to-SQL bridge)

---

2. Workflow & Key Contributions

Phase 1: Exploratory Data Analysis (Python)

- **Data Wrangling:** Handled a dataset of **9,651 transactions**, performing data type conversions (dates, currency) and cleaning missing values.
- **Statistical Analysis:** Used statistics and numpy to calculate mean, median, and variance across sales metrics to identify outliers and distributions.
- **Visual Discovery:** Created heatmaps and distribution plots using **Seaborn** and **Matplotlib** to identify correlations between price per unit, units sold, and operating profit.
- **Pipeline Integration:** Automated the data export from Python directly into a **MySQL database** using SQLAlchemy for streamlined data warehousing.

Phase 2: Database Management & SQL Querying (MySQL)

- **Schema Design:** Designed and implemented the Adidas_sales database schema.
- **Advanced Querying:** Developed complex SQL scripts to perform:
    - **Trend Analysis:** Monthly and yearly sales growth using DATE_FORMAT.
    - **Segmentation:** Used CASE statements to categorize sales into performance tiers (Low, Medium, High).
    - **Ranking:** Identified top 5 retailers, states, and products by total revenue and profit margin using GROUP BY and ORDER BY.

Phase 3: Interactive Dashboarding (Power BI)

- Developed a dynamic multi-page dashboard to track **Key Performance Indicators (KPIs)**: Total Sales, Operating Profit, and Sales Volume.

- **Geospatial Analysis:** Visualized sales density across US regions (Northeast, West, etc.) and specific cities.
- **Channel Performance:** Created comparative charts to analyze the effectiveness of sales methods (In-store, Online, Outlet).

---

3. Key Insights Discovered

- **Product Performance:** Identified which footwear and apparel categories drove the highest volume vs. the highest profit margins.
- **Regional Trends:** Pinpointed high-performing states (e.g., New York, California) and optimized regional sales strategies.
- **Channel Optimization:** Determined that specific sales methods (like Online) offered higher operating margins despite lower individual unit prices.

---

4. Resume Bullet Points (Copy & Paste Ready)

- **Engineered an end-to-end sales analysis pipeline** for 9,600+ Adidas US transaction records using Python, MySQL, and Power BI.
- **Cleaned and pre-processed data in Python**, utilizing Pandas and Seaborn to conduct exploratory data analysis (EDA) and identify key sales drivers.
- **Developed and optimized a MySQL database**, writing complex queries and CASE statements to segment sales performance and calculate regional profitability.
- **Designed an interactive Power BI dashboard** that enabled stakeholders to visualize sales trends, regional performance, and channel-specific profit margins.
- **Automated data migration** from Python to SQL using SQLAlchemy, reducing manual data entry and ensuring data integrity across the analytical stack.