Retail Store Inventory Forecasting Dataset

About Dataset

This dataset provides synthetic yet realistic data for analyzing and forecasting retail store inventory demand. It contains over 73000 rows of daily data across multiple stores and products, including attributes like sales, inventory levels, pricing, weather, promotions, and holidays.

The dataset is ideal for practicing machine learning tasks such as demand forecasting, dynamic pricing, and inventory optimization. It allows data scientists to explore time series forecasting techniques, study the impact of external factors like weather and holidays on sales, and optimize supply chain performance.

Key Data Features:

Date: Daily records from [start\_date] to [end\_date].

Store ID & Product ID: Unique identifiers for stores and products.

Category: Product categories like Electronics, Clothing, Groceries, etc.

Region: Geographic region of the store.

Inventory Level: Stock available at the beginning of the day.

Units Sold: Units sold during the day.

Demand Forecast: Predicted demand based on past trends.

Weather Condition: Daily weather impacting sales.

Holiday/Promotion: Indicators for holidays or promotions.

Based on your dataset and business context, here are problem statements for each of the key areas—demand forecasting, dynamic pricing, and inventory optimization—with an emphasis on what SQL queries could help analyze:

**1.Combine Product and Region Performance:**

**Problem Statement:**

Combine sales performance by product and region to analyze which region sells which product category best.

**2.Daily Forecast Accuracy**

**Problem Statement:**

Calculate the daily forecast error for each product and return only days with major deviations.

**3.Running Total of Units Sold per Product**

**Problem Statement:**

Track the running total of units sold for each product over time.

4.**Identify Restocking Needs Based on Sales Trend:**

**Problem Statement:**

Identify products whose average 7-day sales exceed inventory level, signaling restocking needs.

**5. Demand Forecasting:**

**Problem Statement:** What is the average forecast error per product and region? Which weather conditions lead to the highest variance in demand?

**6.Dynamic Pricing**

**Problem Statement:** During which holidays or promotions did demand spike significantly? Can we find high-demand products under specific weather or regional conditions?

**7.Inventory Optimization**

**Problem Statement:** Which products frequently had stockouts (units sold equal to or greater than inventory)? How often does actual demand exceed forecast, leading to understock?

**8.Impact of Discounts on Demand and Revenue**

**Problem Statement:**

How do different discount levels influence product demand and revenue across categories?

**9.Identify Optimal Price Ranges Based on Revenue**

**Problem Statement:**

What price ranges generate the highest revenue per product category?

**10.Pricing Efficiency vs Competitor Pricing**

**Problem Statement:**

Are we pricing products competitively? Are we losing sales when our prices exceed those of competitors? Which categories are most affected by higher pricing?

**11.Revenue Lost Due to Underpricing (High Demand, Low Price)**

**Problem Statement:**

Are we underpricing high-demand products, resulting in lost revenue opportunities?

**12.Forecasting Revenue Based on Price Sensitivity**

**Problem Statement:**

How does revenue fluctuate when prices are adjusted up or down relative to demand?