

# Cleaning and Organizing a Customer Sales Dataset with SQL

## Scenario

You are a data analyst for a retail company. The company has provided you with a dataset containing customer details, sales transactions, and product information. The data is messy and inconsistent, and your task is to clean and organize it for meaningful analysis and reporting.

---

## Data Cleaning

### 1. Customer Name Standardization:

- Remove leading and trailing spaces using TRIM.
- Convert all names to uppercase using UPPER.

### 2. Product Name Standardization:

- Correct inconsistent names using REPLACE.

### 3. Handling Missing Values:

- Replace NULL values in the email column with a placeholder using COALESCE.
- Treat 0 in the price column as missing data using NULLIF.

### 4. Phone Number Standardization:

- Remove special characters or spaces from phone numbers using REPLACE.
- 

## Data Organization

### 1. Group Sales Data:

- Use GROUP BY to calculate total sales by region.

### 2. Filter High-Performing Regions:

- Use the HAVING clause to focus on regions with total sales exceeding \$50,000.

### 3. Rank Products by Popularity:

- Use RANK or ROW\_NUMBER to rank products based on sales volume.

### 4. Segment Customers:

- Categorize customers as "High-value," "Medium-value," or "Low-value" based on total purchase amounts using CASE.