

Walmart Dataset SQL Query Results

Introduction

This document presents an analytical overview of Walmart's retail transaction data using SQL-based queries. The dataset contains detailed sales records including invoice IDs, product lines, pricing, quantities, taxes, dates, and customer information. Through structured queries, we extract key insights regarding sales trends, revenue performance, and data integrity over specific time periods.

The primary focus is on understanding temporal sales patterns, summarizing transactional metrics, and evaluating data quality, particularly for the first quarter of 2019. The analyses are organized into distinct sections for clarity.

1. Extract Year and Month

Year	Month	Total
2019.0	1.0	548.97
2019.0	3.0	80.22
2019.0	3.0	340.53
2019.0	1.0	489.05
2019.0	2.0	634.38
2019.0	3.0	627.62
2019.0	2.0	433.69
2019.0	2.0	772.38
2019.0	1.0	76.15
2019.0	2.0	172.75

Extract Year and Month

This section extracts the year and month from each transaction date. This preprocessing step enables time-based aggregation and filtering in subsequent analyses.

2. Group by Year & Month (SUM Unit Price)

Year	Month	Total Revenue	Total Orders
2019.0	1.0	19753.89	352.0
2019.0	2.0	17159.52	303.0
2019.0	3.0	18758.72	345.0

Group by Year & Month (SUM Total)

This section performs similar grouping but focuses on the total field (which includes taxes), offering a more comprehensive view of monthly revenue.

3. Group by Year & Month (SUM Total)

Year	Month	Total Revenue	Total Orders
2019.0	1.0	116291.87	352.0
2019.0	2.0	97219.37	303.0
2019.0	3.0	109455.51	345.0

Filter (Q1 2019)

This analysis filters the dataset to include only transactions from January to March 2019. It helps isolate revenue and order counts within this quarter to examine short-term business trends.

4. Filter (Q1 2019)

Year	Month	Total Revenue	Total Orders
2019.0	1.0	116291.87	352.0
2019.0	2.0	97219.37	303.0
2019.0	3.0	109455.51	345.0

Null Checks

To ensure data quality, this section checks for missing values in key fields like invoice_id, date, and total. Identifying and handling null values is critical before performing accurate analysis.

5. Null Checks

Field	Count
Total Rows	1000
Null Invoice ID	0
Null Date	0
Null Total	0