



FMCG Sales Analysis – End-to-End Data Analyst Project

1. Project Objective

The goal of this project was to **understand and demonstrate the complete Data Analyst workflow**—from data understanding and descriptive analysis to diagnostic, predictive, and prescriptive analytics—ultimately delivering **business-driven recommendations** for an FMCG company.

2. Dataset Overview

- **Type:** Structured FMCG business dataset
- **Core Tables Used:**
 - **Sales**
 - **Returns**
 - **Products**
 - **Customers**
 - **Marketing Spend**

Key Columns (Sales & Returns)

Order Number, Order Date, Customer ID & Name, City, Country, Territory, Product Category, Product Code, Product Name, Quantity Ordered, Price Each, Sales, Deal Size, Status

Supporting Lookup Tables (Created)

- Date Table
- Location (Country / Territory)
- Product Category
- Deal Size
- Status

These were used primarily for **slicers, filtering, and time-based analysis**.

3. Initial Data Understanding & EDA

- Before modeling and analysis, **Exploratory Data Analysis (EDA)** was performed across all tables (Sales, Products, Customers, Returns, Marketing Spend).
- **Descriptive statistics** were calculated for all numerical columns (mean, variance, skewness, kurtosis) and **frequency distributions** for categorical variables to understand data spread, concentration, and outliers.
- **Box plots and histograms** were created for each numerical variable to examine distribution behavior, skewness, and potential anomalies.

This step ensured a clear understanding of data characteristics and informed subsequent modeling, aggregation, and interpretation choices.

4. Tools & Techniques

- **Excel**
 - Power Pivot (data modeling, table relationships)
 - Pivot Tables (descriptive analysis)
 - Correlation Analysis
 - Simple Linear Regression
 - Durbin–Watson Test
- **Dashboarding**
 - Interactive slicers (date, location, category, deal size, status)
 - KPI-focused visuals

5. Data Modeling & Measure Creation

- Built a Star Schema using Power Pivot
- Fact tables: Sales, Returns, Marketing Spend
- Dimension tables: Products, Customers, Date, Location, Deal Size, Status

Joins (Logical Assumptions) :

- Sales → Products using Product_Code
- Sales → Customers using Customer_ID
- Sales → Date using Order_Date
- Sales → Location using City

- Returns → Products using Product_Code
- Marketing Spend → Products using Product_Code
- Created calculated measures for analysis:
 - Average Marketing Spend per Quarter
 - Return Rate
 - Average Sales per Quarter

This model enabled consistent slicing, accurate aggregation, and scalable analysis across the dashboard.

6. Analysis Stages Covered

The dashboard below presents a consolidated view of FMCG sales performance, customer concentration, category contribution, geographic distribution, returns, and marketing spend. Interactive slicers were used to explore trends across time, regions, categories, and deal sizes.



A. Descriptive Analytics

Pivot table analysis was performed across **sales, customers, geography, products, returns, and marketing spend** to understand historical performance.

Key Sales & Business Insights

Sales Trends

- Q1 & Q2 show negative growth in most years (except 2003)
- Q3 consistently performs well across all years
- Q4 delivers the highest sales except in 2005

Customer Insights

- Lonsdale Wholesale and Millford Wholesale together contribute **17.54% of sales**
- Top 10 customers (out of 92) contribute **33.65% of total sales**

Geographic Insights

- USA alone contributes **33.84% of sales**
- Top 3 countries contribute **58.98%**
- Except UK and Finland, all other countries contribute **<5%**

Category Insights

- Beverages & Infant Nutrition contribute **64% of sales**
- Top 3 categories account for **82%**
- Remaining categories together contribute only **18%**

Product Concentration

- Top 10 products contribute **74.69%**
- Top 3 products alone contribute **32.27%**
- Remaining 23 products average only **~1.10% each**

Deal Size

- Majority of the business runs on **Small Deal Size**

B. Returns Analysis

- **1.60% of total returns occurred in 2005**
 - Return rate increased sharply with **127% CAGR**
-

Returns by Category

- Beverages: **26.58% of returns**
- Beverages + Dairy Products together: **49.91%**
- Frozen & Instant Foods have **<5% returns**

Returns by Product

- Top 10 products contribute **68.29% of total returns**
- *Milko Gold* alone accounts for **12.15% return share**

C. Marketing Spend Analysis

- Marketing spend **declined by 37% CAGR** till 2005
- Spending dropped consistently year over year

Category-Level Spend

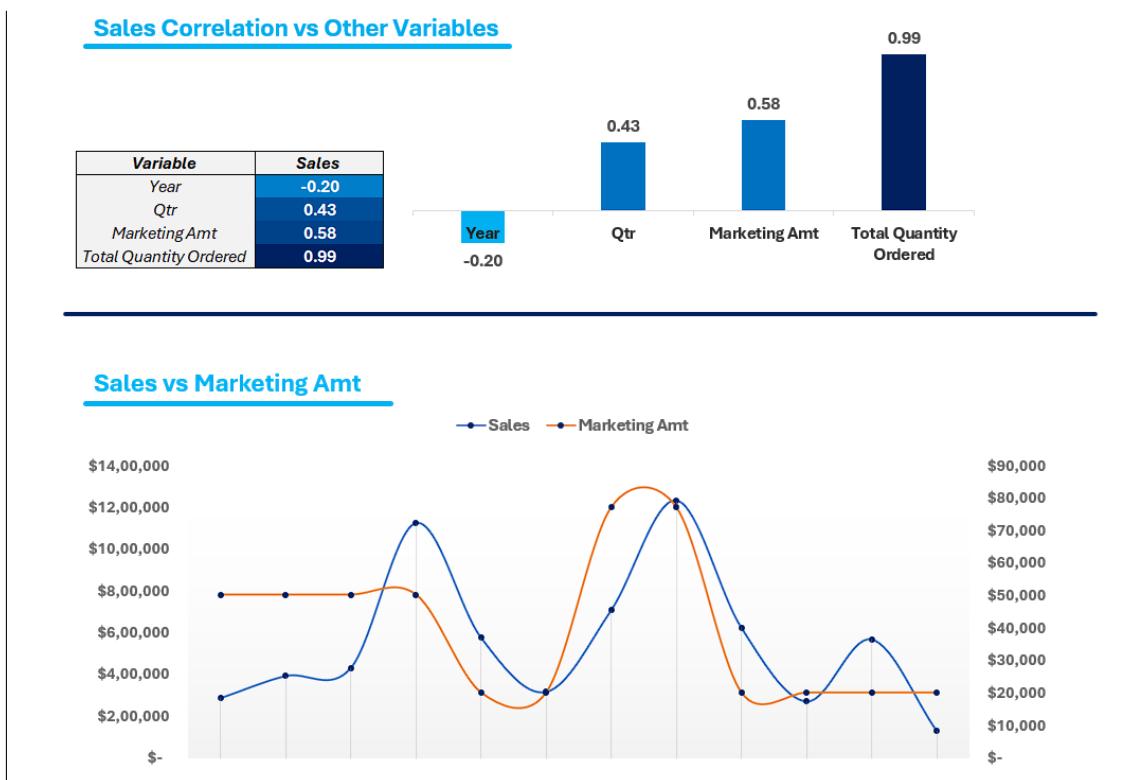
- Beverages receive ~50% of total marketing budget
- Beverages + Infant Nutrition consume **72.72%**
- Frozen Foods, Water & Juices, Snacks receive **<5%**

Product-Level Spend

- Top 10 products consume **80%** of marketing budget
- Top 3 products alone receive **40.17%**

7. Diagnostic Analytics (Why is this happening?)

The following charts visualize the relationships between Sales and key business drivers to identify factors influencing performance and validate drivers observed in descriptive analysis.



Correlation Analysis Insights

Relationship	Correlation
Marketing Spend ↔ Sales	0.58
Marketing Spend ↔ Quantity Ordered	0.57
Year ↔ Marketing Spend	-0.57
Year ↔ Sales	-0.20
Sales ↔ Quantity Ordered	0.99

Interpretation

- Marketing has a **moderate positive impact** on both sales and quantity.
- Over time, **reduced marketing investment** aligns with weaker sales performance.
- Sales growth is primarily driven by **volume growth**.

8. Predictive Analytics

Model Used

- **Simple Linear Regression**
- Dependent Variable: Sales
- Independent Variable: Marketing Spend

Regression Equation

$$\text{Sales} = 8.5888 \times \text{Marketing Spend} + 216,640$$

Key Results

- Every **\$1 increase in marketing spend generates ~\$8.59 in sales**
- Model explains **33% of sales variance ($R^2 = 0.33$)**
- Model is statistically significant at **95% confidence**

Durbin–Watson Test

- $DW = 1.69$
- Indicates **no serious autocorrelation**, supporting model reliability

Interactive Prediction

- Marketing Spend Input → Predicted Sales Output
- Example:
\$10,000 Marketing Spend → ~\$302,528 Predicted Sales

9. Prescriptive Analytics (Business Recommendations)

Business Problem	Root Cause	Recommendations
Low & Irregular Sales	Undiversified, inconsistent marketing	<ul style="list-style-type: none">• Diversify marketing across more products/categories• Maintain stable marketing budgets across years
Rising Return Rates	Quality & supply chain issues	<ul style="list-style-type: none">• Conduct SKU-level quality audits• Improve storage, logistics, and packaging processes• Train supply chain staff

10. Final Outcome

This project demonstrates a **complete Data Analyst workflow**:

- Descriptive → Diagnostic → Predictive → Prescriptive
- Strong focus on **business interpretation**, not just visuals
- Designed to support **data-driven decision-making** for FMCG manufacturers