# Ubuntu Organics – Data Analysis Project

## Project Background

Ubuntu Organics is a South African Direct-to-Consumer (DTC) e-commerce retailer founded in 2025. The business quickly gained traction with its range of health and wellness products but has yet to fully leverage the rich datasets generated by its operations.

This project was designed to extract insights from sales, customer, and campaign data in order to answer five core business questions:
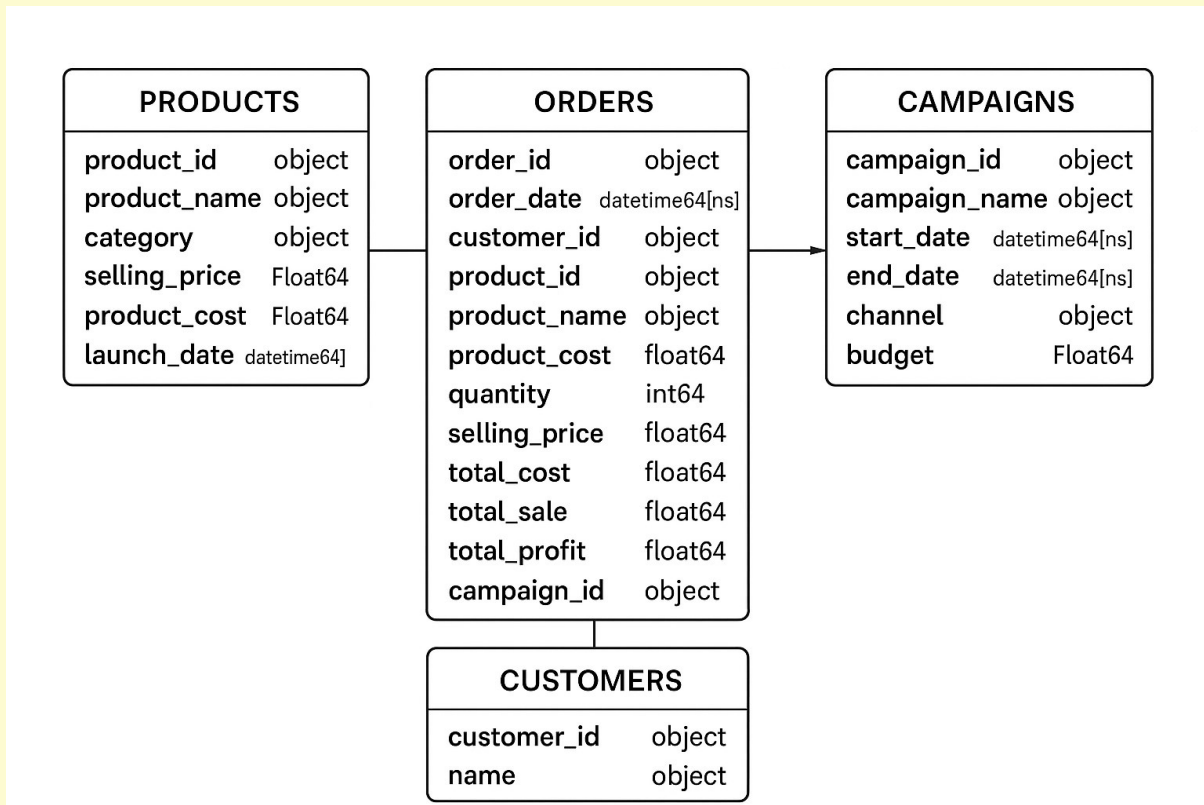
1. What are the monthly sales trends over the past year?
2. How has Average Order Value (AOV) changed over time, particularly during campaigns?
3. Which marketing campaigns delivered the highest ROI?
4. Which acquisition channels are the most cost-effective?
5. Do customers make repeat purchases, and how do they behave compared to first-time buyers?

---

## Data Overview

Analysis was based on four core datasets covering the 2025 calendar year:

- **Products Table** – product IDs, names, categories, selling price, cost, launch date. Used for profitability and lifecycle analysis.
- **Orders Table** – transaction-level sales data with order IDs, dates, customer IDs, product details, quantities, campaign tags, and calculated revenue/profit fields.
- **Campaigns Table** – campaign ID, name, duration, channel (Meta, Google, Referrals, Influencers), and budget. Used for ROI evaluation.
- **Customers Table** – customer ID, acquisition channel, signup date, and region. Used for segmentation and repeat buyer analysis.

**Preprocessing included:** standardizing dates, validating IDs across tables, merging datasets into a unified model, and generating monthly periods for trend analysis.

## PRODUCTS

| | |
|---|---|
| product_id | object |
| product_name | object |
| category | object |
| selling_price | Float64 |
| product_cost | Float64 |
| launch_date | datetime64] |

## ORDERS

| | |
|---|---|
| order_id | object |
| order_date | datetime64[ns] |
| customer_id | object |
| product_id | object |
| product_name | object |
| product_cost | float64 |
| quantity | int64 |
| selling_price | float64 |
| total_cost | float64 |
| total_sale | float64 |
| total_profit | float64 |
| campaign_id | object |

## CAMPAIGNS

| | |
|---|---|
| campaign_id | object |
| campaign_name | object |
| start_date | datetime64[ns] |
| end_date | datetime64[ns] |
| channel | object |
| budget | Float64 |

## CUSTOMERS

| | |
|---|---|
| customer_id | object |
| name | object |

## Executive Summary

Ubuntu Organics' 2025 performance was shaped primarily by **seasonal demand cycles**, with campaigns acting as amplifiers rather than independent growth drivers. Campaign ROI was inconsistent, with only one Meta initiative - **Campaign M08** - standing out.

Acquisition efficiency varied: **Referrals and Google** outperformed, while Meta and Influencers underdelivered. Customer retention emerged as the largest untapped growth lever - **repeat buyers delivered 4× higher value** and responded strongly to campaigns.

**Key Takeaways:**

1. **Seasonality is the main driver of sales.** Campaigns amplify but do not create demand.

2. **Campaign ROI is uneven.** Campaign M08 was a standout success; others performed mediocrely or underperformed.

3. **Repeat buyers are critical.** They are nearly 4x more valuable and much more responsive to marketing.

# Insights Deep Dive

## 1. Sales Trends

- Sales peaked during South African school holidays and year-end shopping periods.
- Campaigns boosted these peaks but were not the underlying cause.
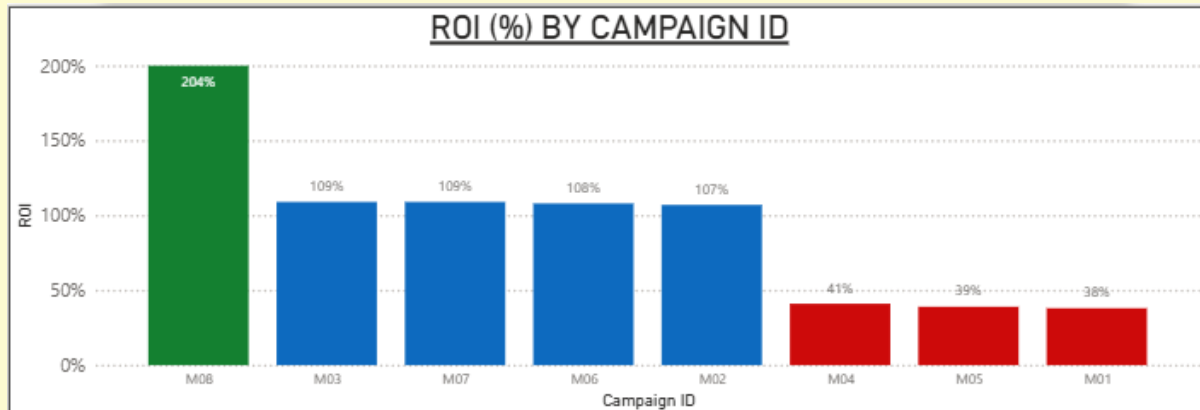- Outside seasonal highs, sales grew only modestly.
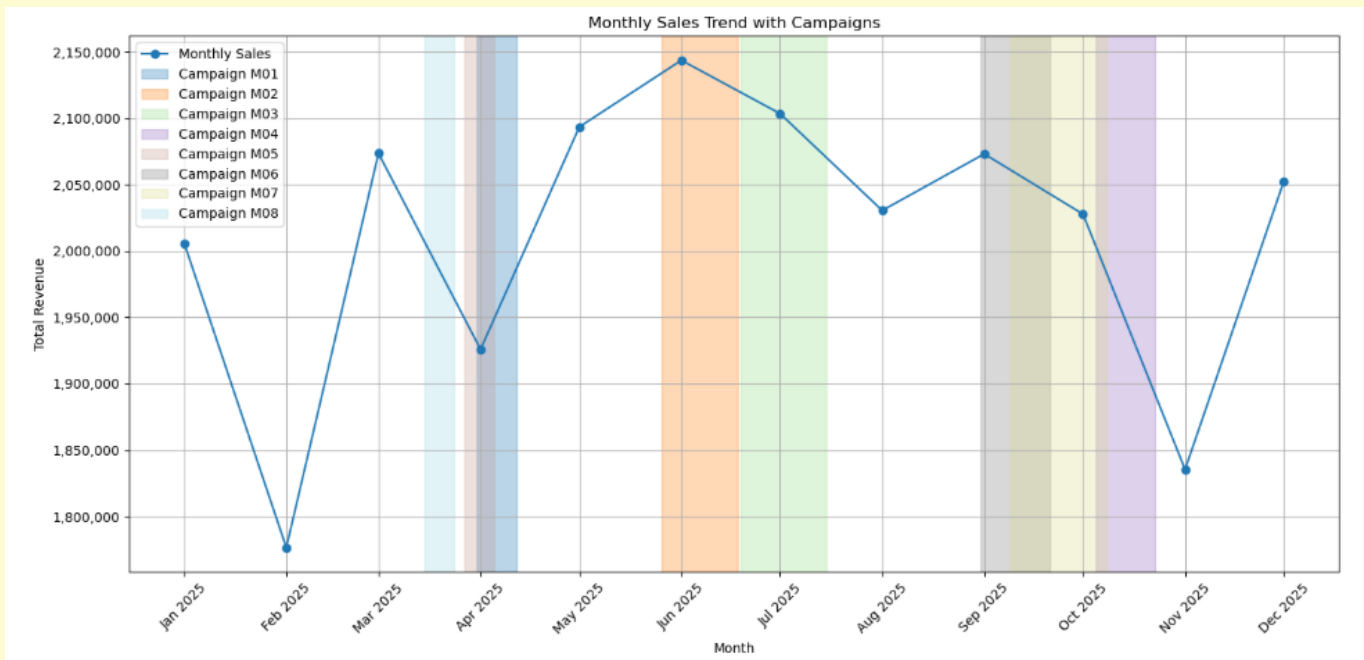


## 2. Average Order Value (AOV)

- AOV followed seasonal cycles rather than campaign timing.
- Campaigns had **no statistically significant impact** on AOV.
- Repeat buyers consistently placed larger orders, raising monthly AOV in high-retention periods.
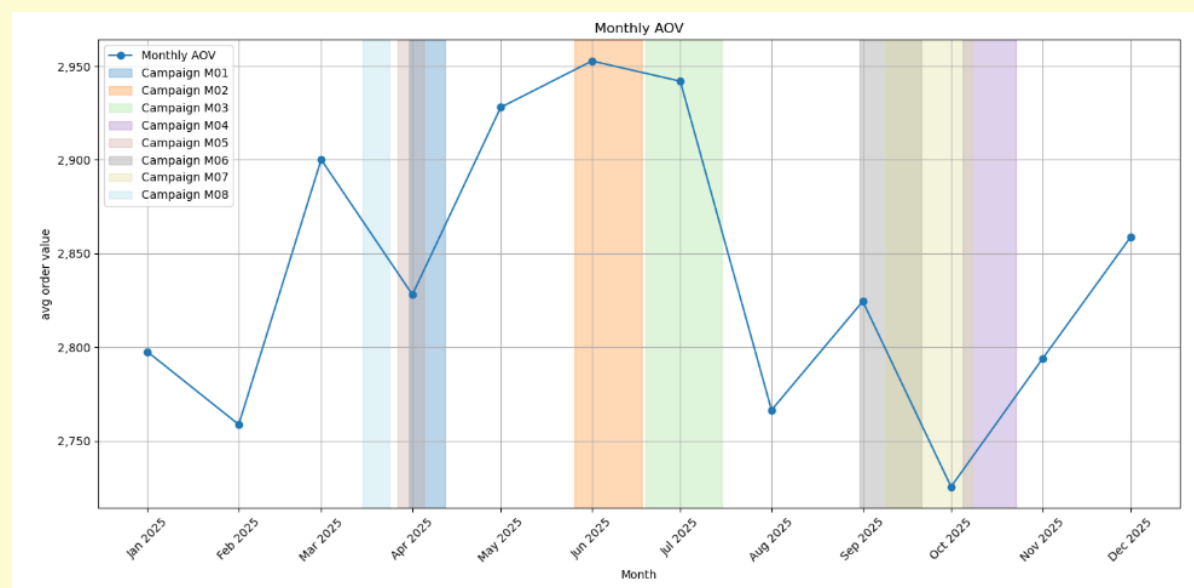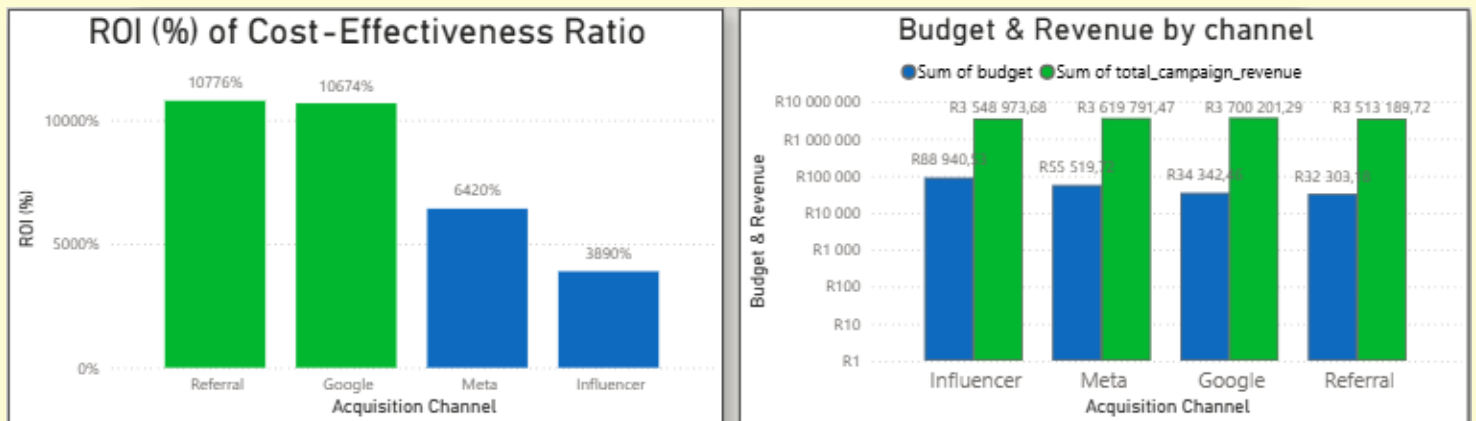
## 3. Campaign ROI

- Most campaigns delivered **average or poor ROI**.
- **Campaign M08** significantly outperformed all others.
- One Meta campaign exceeded expectations; another failed badly, suggesting **execution quality matters more than platform choice**.

### ROI (%) BY CAMPAIGN ID

| Campaign ID | ROI |
|-------------|-----|
| M08 | 204% |
| M03 | 109% |
| M07 | 109% |
| M06 | 108% |
| M02 | 107% |
| M04 | 41% |
| M05 | 39% |
| M01 | 38% |

## 4. Customer Acquisition

- **Referrals and Google** were the most cost-effective channels.
- **Meta and Influencers** generated weaker returns relative to spend.

### ROI (%) of Cost-Effectiveness Ratio

| Acquisition Channel | ROI (%) |
|---------------------|---------|
| Referral | 10776% |
| Google | 10674% |
| Meta | 6420% |
| Influencer | 3890% |

### Budget & Revenue by channel

Sum of budget ● Sum of total_campaign_revenue

| Acquisition Channel | Sum of budget | Sum of total_campaign_revenue |
|---------------------|---------------|-------------------------------|
| Influencer | R88 940,5? | R3 548 973,68 |
| Meta | R55 519,7? | R3 619 791,47 |
| Google | R34 342,4? | R3 700 201,29 |
| Referral | R32 303,1? | R3 513 189,72 |

## 5. Customer Retention

- Repeat buyers generated **~4× higher AOV** than one-time buyers.
- **60% of repeat buyers** had campaign-attributed purchases vs. **20% of one-time buyers**, proving higher responsiveness to marketing.

Average Order Value by Buyer Type



Campaign Acquisition vs Organic Acquisition by Buyer Type
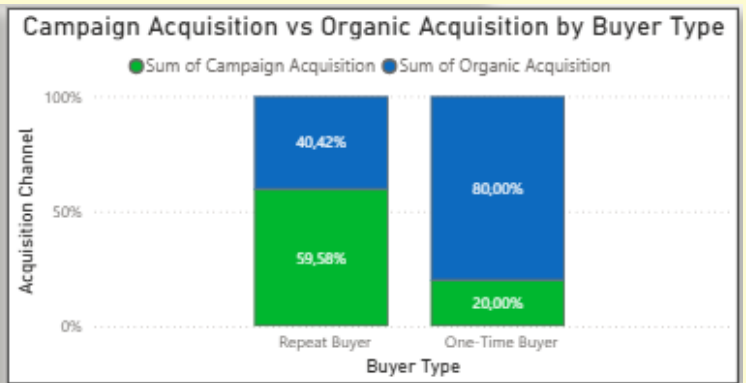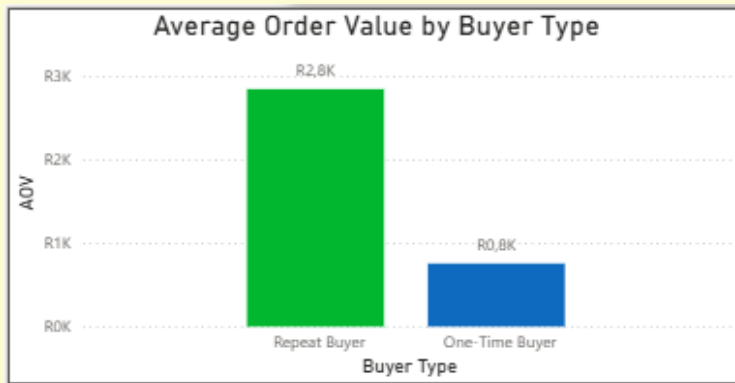
## Recommendations

1. **Align campaigns with seasonality.** Focus on school holidays and year-end peaks to ride existing demand.
2. **Double down on proven tactics.** Replicate strategies from Campaign M08 and the high-performing Meta campaign. If replication fails, shift spend to Referrals and Google.
3. **Prioritize customer retention.** Build loyalty programs, targeted campaigns, and personalized offers to nurture repeat buyers.
4. **Enhance customer profiling.** Capture demographic and behavioral data to enable deeper segmentation and precision targeting.

## Assumptions & Limitations

- Campaign attribution is based on available data; multi-channel interactions may be underreported.
- Repeat buyer definition is limited to 2025 purchase history only.
- Refunds and returns were excluded, potentially overstating sales.
- External market factors (competitors, economic shifts) were not modeled.