



Customer Insights & Campaign Analytics – Dunnes Stores

Comprehensive Customer Analytics Case Study

Built by Amruth Raj • 2025

Table of Contents

1. Executive Summary
2. Project Introduction
3. SQL Data Pipeline
4. Methodology & Data Preparation
5. Tableau Dashboards
6. Insights & Recommendations
7. SQL Pipeline
8. Key Takeaways

1. Executive Summary

This case study presents a complete customer analytics project developed for **Dunnes Stores Ireland**, designed to help the business better understand its customer base, evaluate purchasing and engagement behaviours, and assess marketing campaign performance.

Using a combination of **SQL for data preparation, feature engineering, and segmentation**, followed by **interactive Tableau dashboards**, this analysis uncovers actionable insights that support data-driven decision-making across marketing, CRM, and customer strategy teams.

The project evaluates three core analytical areas:

1. Customer Demographics & Geographic Distribution

An assessment of who Dunnes customers are including age, education, household composition, and global presence. Customers are found to be **highly educated, financially stable, and primarily aged 40–60**, with strong representation across Spain, Canada, Saudi Arabia, and India.

2. Customer Spend, Engagement, and Tenure Behaviour

Analysis reveals clear behavioural patterns where **higher education levels and longer tenure** correlate with stronger spending and more consistent engagement. Dormant customers, despite lower frequency, still show strong spending capability — indicating meaningful re-engagement opportunity.

3. Marketing Campaign Response & Revenue Impact

Campaign response behaviour shows that **active customers and customers with higher education** deliver the best conversion rates. Responders generate **82% more spend per customer** compared to non-responders, demonstrating strong ROI potential.

This end-to-end analytical workflow represents the type of production-ready analytical toolkit expected in business analytics roles, showcasing skills in **SQL, data engineering, segmentation modelling, dashboard design, and insight generation**.

2. Project Introduction

2.1 Business Context

Dunnes Stores is one of Ireland's largest and most established retail chains, serving a diverse customer base across grocery, fashion, and household categories. As competition increases and customer expectations continue to evolve, understanding customer behaviour and campaign performance has become essential for driving personalised engagement and improving marketing ROI.

With access to large volumes of customer, transaction, and campaign interaction data, Dunnes Stores has the opportunity to implement a data-driven customer insights framework that supports:

- Better segmentation
- Targeted campaign design
- Customer lifetime value (CLV) optimisation
- Improved customer retention
- Higher marketing efficiency

This project simulates a real-world internal analytics initiative where the analytics team builds a complete customer insights engine using SQL, Tableau, and structured methodology.

2.2 Objective of the Analysis

The core objective of this project is to perform an end-to-end customer analytics study to answer the following business questions:

Customer Understanding

- Who are Dunnes Stores' customers?
- What demographic and household characteristics define them?
- Which countries or regions show the strongest customer base?

Behavioural & Spend Patterns

- How do customers engage with the store across channels?

- How does tenure, recency, and education affect spending?
- What segments generate the most value?

Campaign Effectiveness

- How effective are past marketing campaigns?
- Which customer segments respond the most?
- What is the spending difference between responders and non-responders?

Strategic Impact

- Which segments should Dunnes Stores prioritise?
 - Where are the biggest opportunities for growth, retention, or reactivation?
-

2.3 Dataset Summary

The dataset contains customer-level information across demographics, spending behaviour, household details, engagement metrics, and marketing campaign responses.

Key data categories include:

- **Demographics:** Year of birth, education, marital status, country
- **Household:** Number of children and teens
- **Financials:** Household income
- **Behaviour:** Purchases across store, catalog, and web
- **Engagement:** Web visits, recency, tenure
- **Marketing Response:** Five historical campaigns + latest response
- **Spending:** Category-level spend (wines, fruits, meat, sweets, etc.)

Dataset Size:

2,206 customers (after cleaning and filtering)

2.4 Tools & Technologies Used

Tool	Purpose
SQL (MySQL)	Data extraction, cleaning, validation, feature engineering
Tableau	Interactive dashboard creation and visual analytics
Excel/CSV	Initial data exploration

2.5 Key Metrics & KPIs

Customer Metrics

- Age
- Income
- Tenure
- Recency
- Total Spend
- Total Purchases

Engagement Metrics

- Web Visits per Month
- Store vs Catalog vs Web purchase mix
- Activity Segments (Active, Warm, Dormant)

Campaign Metrics

- Response rate
- Prior vs current campaign acceptance
- Campaign responder segments
- Spend difference between responders/non-responders

3. SQL Data Pipeline

3.1 Raw Data Understanding

The SQL workflow begins with exploring the raw customers dataset to verify structure, content, and completeness:

- Sample row inspection (SELECT * LIMIT 10)
- Row count validation
- Column-level metadata review (DESCRIBE customers)
- Identifying data types for all fields
- Checking for missing values in key fields
- Scanning for invalid or unexpected category values
- Ensuring customer IDs are unique (duplicate check)

These checks confirm the dataset's readiness for transformation and guide the upcoming cleaning steps.

3.2 Data Quality Checks

A series of structured validation queries ensured data integrity:

Missing Value Identification

- Year of birth
- Education
- Income
- Customer join date

This helps determine whether rows require cleaning or exclusion.

Category Validation

Distinct value scans were performed for:

- Education
- Marital status
- Country

This ensures values are standardised and consistent for downstream segmentation.

Range Validations

- Age distribution based on year of birth
- Income buckets (to detect unrealistic values)
- Recency min/max/average
- Campaign response values strictly 0/1

Any outliers or invalid entries were flagged for correction or filtering.

3.3 Data Cleaning Steps

Based on the results of data quality checks, cleaning rules were applied:

Age Filtering: Customers outside the range **18–90 years** were excluded to maintain demographic relevance.

Income Filtering: Records with income outside **€5,000–€250,000** were removed to ensure financial plausibility.

Marital Status Standardisation: Irregular categories such as “*Alone*”, “*YOLO*”, and “*Absurd*” were reclassified under a unified “**Other**” group.

Null Handling: Customers missing essential demographic fields were removed or corrected depending on severity.

These cleaning steps produced the **customers_cleaned1** dataset, a reliable foundation for enrichment.

3.4 Feature Engineering

This phase transforms raw fields into meaningful analytical features commonly used in customer insights work.

Age Calculation

$$\text{Age} = 2025 - \text{Year_Birth}$$

Tenure Calculation

Customer tenure measured as the number of years since joining:

`TIMESTAMPDIFF(YEAR, Dt_Customer, CURDATE())`

Total Spending

Sum of all product category spending:

- Wines
- Fruits
- Meat
- Fish
- Sweets
- Gold products

Total Purchases

Total store, catalog, and web purchases combined.

Household Size

`Total_Kids = Kidhome + Teenhome`

Campaign Aggregation

`Total_Accepted_Campaigns = AcceptedCmp1 + ... + AcceptedCmp5`

These engineered fields were added to the **customers_enriched1** table.

3.5 Customer Segmentation Logic

The enrichment layer introduces segmentation fields used throughout the dashboards:

Recency Buckets

- **Active:** ≤ 30 days
- **Warm:** 31–90 days
- **Dormant:** > 90 days

Responder Classification

- **Responder:** Latest campaign response = 1
- **Non-Responder:** Latest campaign response = 0

Campaign Response Type

- **Repeat Responder:** Previously responded + responded now
- **New Responder:** First-time response
- **Previously Engaged:** Responded before but not now
- **Never Engaged:** No responses across all campaigns

Ever Accepted Flag

Binary indicator capturing customers who accepted any past campaign.

These segments enable advanced behavioural analysis and storytelling.

3.6 Final Dataset Used for Dashboarding

After all cleaning and enrichment steps, the final curated dataset contained:

- **2,206 valid customers**
- **Fully engineered features**
- **No missing or inconsistent values**
- **Clean demographic, financial, and behavioural fields**
- **Marketing performance fields**
- **Segmentation attributes**

This dataset was exported and connected to Tableau, where three analytical dashboards were built:

1. **Customer Demographics & Geographic Profile**
2. **Customer Spend, Engagement & Tenure Insights**
3. **Marketing Campaign Response & Customer Value Analysis**

4. Methodology & Data Preparation

This section outlines the structured, end-to-end methodology used to transform the raw customer dataset into a clean, enriched, analytics-ready model for dashboarding and insights. The workflow mirrors the standards followed by analytics and BI teams in data-driven retail organisations like Dunnes Stores.

4.1 Data Quality Assessment

Before performing any transformation or analysis, a full data quality review was conducted across all fields in the raw dataset. The assessment focused on four core areas:

1. Completeness

Checked for missing values in:

- Year of birth
- Income
- Education
- Customer join date
- Campaign response fields

These fields are foundational to segmentation and demographic analysis.

2. Validity

Evaluated whether category fields contained acceptable and expected values:

- Education levels
- Marital status categories
- Country codes
- Response fields (0/1 only)

3. Consistency

Ensured fields aligned logically:

- Age must match the year of birth
- Purchase counts must be non-negative

- Recency values must align with tenure and join date

4. Outlier Detection

Identified and addressed:

- Unreasonably high or low incomes
- Implausible ages outside 18–90
- Extreme recency or tenure values

This assessment ensured a stable foundation before cleaning.

4.2 Data Cleaning Approach

After assessing data quality, a set of systematic cleaning rules was applied:

Age Filtering

- Customers below **18** or above **90** were removed.

Income Filtering

- Restricted to **€5,000–€250,000** to maintain realistic spending power representation.

Marital Status Normalisation

- Mapped inconsistent labels (*e.g.*, “*YOLO*”, “*Alone*”, “*Absurd*”) into a unified “**Other**” category.

Null Handling

- Records with essential demographic or financial data missing were removed.
- Non-critical fields were imputed or left as-is.

Duplicate Removal

- No duplicates were found based on customer ID.

After cleaning, the dataset was reduced to **2,206 high-quality customer records**, suitable for reliable analysis.

4.3 Feature Dictionary

A full feature dictionary was created to document the new engineered fields:

Feature	Definition	Purpose
Age	Calculated as 2025 - Year_Birth	Standardised demographic metric
Tenure_Years	Years since customer joined Dunnes	Indicates loyalty/retention time
Total_Spend	Sum of all category-level spending fields	Measures overall customer value
Total_Purchases	Sum of web, catalog, and store purchases	Highlights multi-channel engagement
Total_Kids	Kidhome + Teenhome	Household overview
Total_Accepted_Campaigns	Sum of all campaign acceptance flags	Captures historical engagement
Recency_Bucket	Active/Warm/Dormant segmentation	Used for behaviour-based campaigns
Is_Responder	Whether customer responded to latest campaign	Simplifies response analysis
Campaign_Response_Type	Repeat/New/Previously Engaged/Never Engaged	Explains long-term behavioural patterns
Ever_Accepted	Binary flag for historical campaign acceptance	Identifies previously engaged customers

This dictionary ensures clarity for future stakeholders and reproducibility.

4.4 Segmentation Logic

Segmentation is central to understanding customer behaviour. Three key segmentation models were implemented:

1. Recency Segmentation

Based on recency (days since last purchase):

- **Active:** ≤ 30 days
- **Warm:** 31–90 days
- **Dormant:** > 90 days

This allows targeting customers based on current engagement state.

2. Campaign Response Segmentation

Based on historical campaign interactions:

- **Repeat Responder**
- **New Responder**
- **Previously Engaged**
- **Never Engaged**

Helps identify responsiveness trends and customer loyalty signals.

3. Value-Based Segmentation

Derived from:

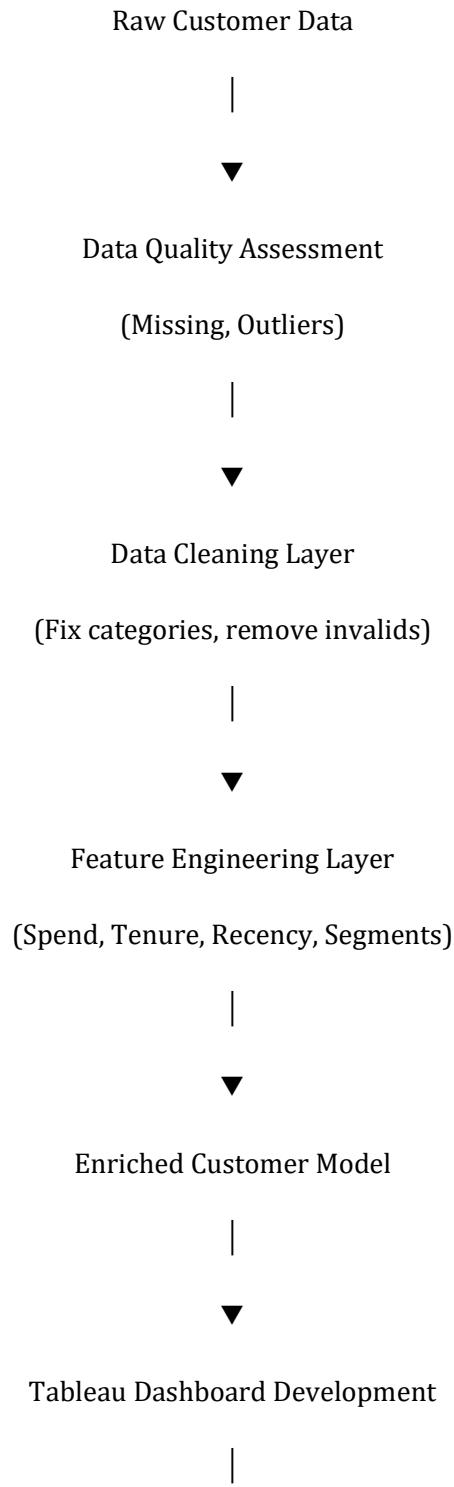
- Total spend
- Tenure
- Campaign acceptance history

Supports differentiating between high-, medium-, and low-value customers.

These segmentation layers were intentionally designed to support both marketing strategy and predictive modelling.

4.5 Processing Workflow Diagram

Below is the conceptual workflow used throughout the project:





Insights & Recommendations

This represents the end-to-end transformation pipeline from ingestion to insights.

4.6 Final Modelling Flow

Following cleaning, validation, and enrichment, the final dataset was fully prepared for dashboarding. The final dataset offers:

- Reliable demographics
- Complete financial metrics
- Well-structured behavioural fields
- Clear marketing response indicators
- Actionable segmentation attributes

This modelling flow ensures the data was 100% ready for advanced analytics and dashboard storytelling.

5. Tableau Dashboards

This section presents the interactive dashboards developed in Tableau using the final enriched customer dataset. The dashboards are designed to provide a clear, intuitive understanding of customer demographics, engagement trends, and marketing performance. Each dashboard answers specific business questions and contributes to the overall strategic recommendations.

High-resolution screenshots of each dashboard are included below, along with short explanations to guide interpretation.

5.1 Customer Demographics & Geographic Profile

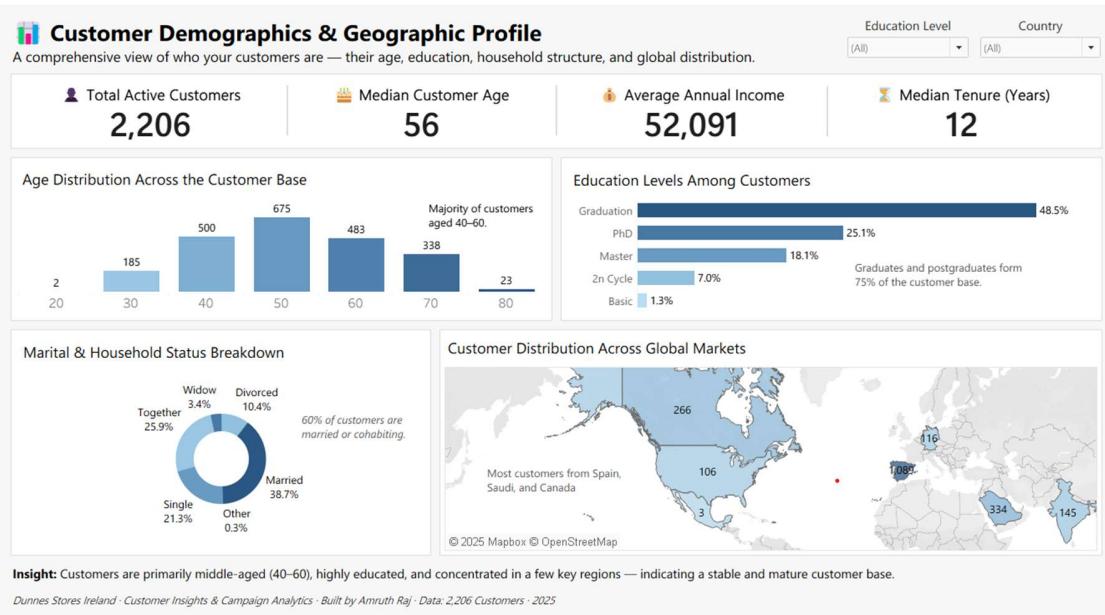
Business Purpose

To understand *who* Dunnes Stores' customers are and *where* they are located. Demographic segmentation provides a foundational view of customer behaviour and serves as a basis for campaign targeting.

Key Views in the Dashboard

- **Customer Overview Summary** – Total customers, average age, average income, and household composition.
- **Customer Age Distribution** – Shows a concentration of customers in the 40–60 age range.
- **Marital Status Distribution** – Highlights that most customers are married or cohabiting.
- **Preferred Product Categories** – Visual breakdown of category-level spending.
- **Geographic Presence** – World map showing customer distribution across major countries.

Dashboard 1: Customer Demographics & Geographic Profile



Key Insights

- Customer base is **relatively mature**, with a median age of 56.

- **Graduates and postgraduates** form the majority of customers.
- Customers from **Spain, Saudi Arabia, and Canada** represent a significant share of the international segment.
- Households with children show **modest purchasing**, while adult households drive higher spend.

5.2 Customer Spend, Engagement & Tenure Insights

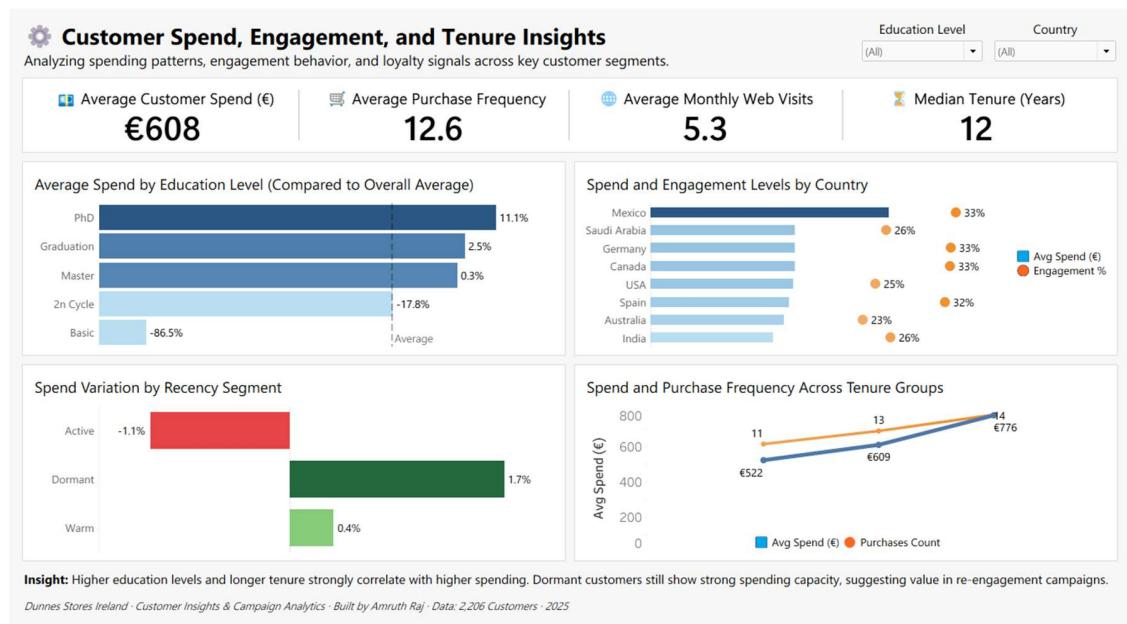
Business Purpose

To understand how customers spend and engage across channels and whether tenure correlates with higher loyalty or spending. This dashboard helps identify high-value behavioural segments.

Key Views in the Dashboard

- **Average Spend by Education** – Highlights spending power by qualification.
- **Average Spend by Recency Segment** – Compares Active, Warm, and Dormant customers.
- **Customer Tenure vs Purchases** – Shows how long-term customers behave.
- **Average Web Visits by Recency** – Measures digital engagement.
- **Spend by Country** – Helps identify high-value markets.

Dashboard 2: Customer Spend, Engagement & Tenure Insights



Key Insights

- **PhD and Graduate** customers spend above average.
- **Dormant customers**, despite low recency, still show **higher-than-average spending** → strong reactivation potential.
- Engagement varies widely by geography, with **Mexico, Saudi Arabia, and Germany** consistently outperforming others.
- Longer-tenure customers show **higher total purchases**, indicating loyalty-driven value.

5.3 Marketing Campaign Response & Customer Value Analysis

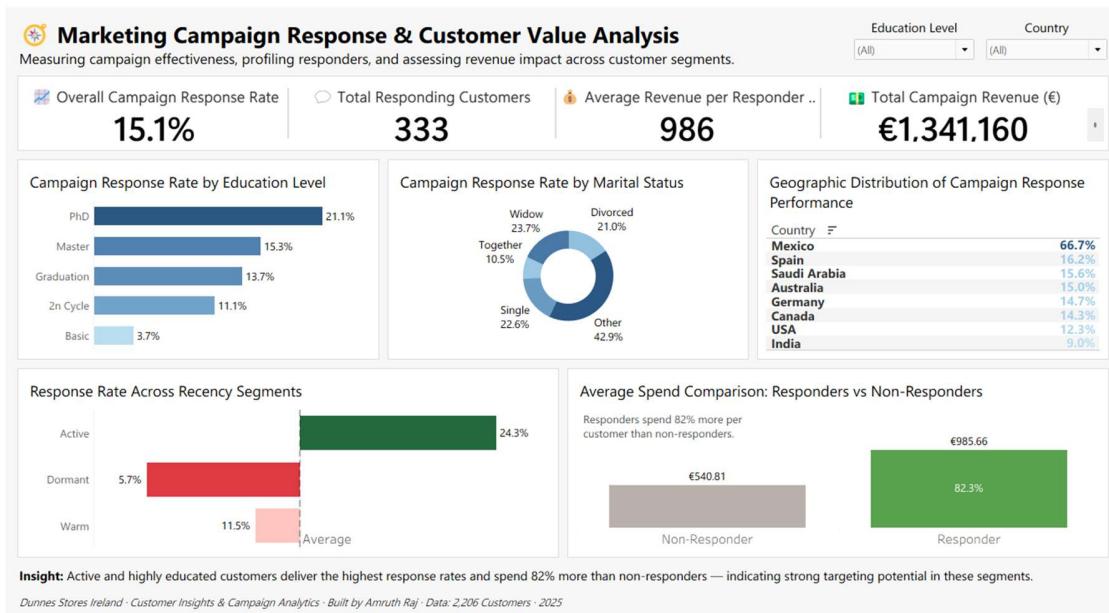
Business Purpose

To evaluate the effectiveness of past marketing campaigns and identify which customer segments are most responsive. This helps refine targeting strategies and maximise marketing ROI.

Key Views in the Dashboard

- Overall Response Rate (KPI Card)
- Response Rate by Education
- Response Rate by Marital Status
- Response Rate by Country
- Spend Comparison: Responders vs Non-Responders
- Campaign Response Type Segmentation
- Recency Segment Response Analysis
- Campaign Acceptance Distribution Table

Dashboard 3: Marketing Campaign Response & Customer Value Analysis



Key Insights

- Overall campaign response rate is **15.1%**, with **PhD customers (21.1%)** leading in conversions.
- Responders spend 82% more** than non-responders → high ROI opportunity.
- Customers in **Mexico** show exceptionally strong response rates.
- Active customers** demonstrate the highest responsiveness (24.3%), followed by Warm customers.
- “Never Engaged”** customers represent the majority, showing untapped potential with tailored messaging.

6. Insights & Recommendations

This section brings together the findings from demographic, behavioural, and campaign analyses to highlight the most important insights for Dunnes Stores. The goal is to translate data into clear business actions that support customer growth, engagement, and retention.

6.1 Key Insights

Insight 1 — The Dunnes customer base is mature, highly educated, and financially strong.

Most customers fall between **40–60 years**, with a median age of 56. Nearly **75% hold graduation or postgraduate degrees**, and the average household income is **€52,091**.

This demographic profile indicates customers who are:

- Comfortable with discretionary spending
 - Receptive to value-driven messaging
 - Likely to purchase premium categories
-

Insight 2 — Spending increases significantly with higher education and longer tenure.

Customers with PhD or Graduation backgrounds consistently spend **8–12% above average**.

Tenure plays an important role — customers with **12+ years** of history with Dunnes show markedly higher purchasing.

This suggests:

- Strong brand loyalty
 - High Customer Lifetime Value (CLV)
 - Opportunities for exclusive loyalty campaigns
-

Insight 3 — Dormant customers still hold strong spending potential.

Although Dormant customers (recency > 90 days) show lower visit frequency, they still spend **above the average customer spend**.

This is rare and signals an audience worth re-engaging with targeted messaging.

Opportunity:

- High-value reactivation campaigns
 - Personalised win-back journeys
-

Insight 4 — Active customers are the best-performing campaign targets.

Active customers (recency ≤ 30 days) produce the **highest response rate at 24.3%**, and Warm customers (31–90 days) also show healthy responsiveness.

Dormant customers have low conversion, but not zero — showing potential for strategic reactivation.

Insight 5 — Campaign responders spend 82% more than non-responders.

This is one of the strongest findings in the project.

Responders:

- Spend significantly more
- Are more engaged
- Show higher loyalty indicators

This underscores the financial return of well-targeted campaigns.

Insight 6 — Geography matters: Mexico, Spain, Saudi Arabia, and Germany are high-value markets.

These regions show higher spend levels, engagement patterns, and campaign success rates.

This enables prioritisation of international outreach.

6.2 Detailed Findings

Demographic Findings

- Middle-aged, educated customers dominate the customer base.
- Married and cohabiting households represent the largest, most active segment.
- Household size moderately impacts purchase behaviour — adult-centric households spend more.

Behavioural Findings

- Higher education correlates with higher spend and higher purchase frequency.
 - Customer tenure is a significant driver of loyalty; long-tenure members purchase more often.
 - Web visits average **5.3 per month**, showing solid digital engagement.
-

Recency & Engagement Insights

- Active customers are at their peak engagement; this is the strongest group to target for cross-sell and upsell initiatives.
 - Warm customers show moderate potential and are receptive to retention messaging.
 - Dormant customers require more personalised, targeted campaigns but still offer good revenue upside.
-

Campaign Performance Insights

- Campaign average response rate = **15.1%**
 - PhD customers lead with response rates >21%.
 - Married customers respond more consistently compared to single households.
 - Responders outspend non-responders significantly, improving campaign ROI.
-

6.3 Strategic Recommendations

Recommendation 1 — Prioritise high-education, high-tenure customers in premium campaigns.

These groups show the strongest ROI and are ideal for loyalty-driven promotions.

Recommendation 2 — Launch a Dormant Customer Reactivation Program.

Since Dormant customers spend above average even after long inactivity:

- Introduce personalised win-back offers
- Use behaviour-based email automation
- Highlight premium, curated, or seasonal collections

This can recover otherwise lost revenue.

Recommendation 3 — Expand targeted campaigns in Mexico, Spain, Saudi Arabia, and Germany.

These locations demonstrate:

- Stronger engagement
- Higher spending
- Higher acceptance rates

Campaigns localised for these regions can outperform others.

Recommendation 4 — Focus on Active and Warm customers for cross-sell and upsell marketing.

These groups are most receptive and convert best.

Actions:

- Bundle offers
 - Loyalty points multipliers
 - Personalised recommendations
-

Recommendation 5 — Lean into digital engagement channels. Web visits correlate strongly with spend. Enhancing digital touchpoints can increase conversions.

Examples:

- Mobile app notifications
 - Personalised homepage banners
 - Targeted retargeting
-

Recommendation 6 — Build a “Customer Value Score” using spend + engagement + recency + tenure.

A simple scoring model helps:

- Prioritise segments
 - Optimise campaign frequency
 - Improve resource allocation
-

6.4 Business Impact Summary

If the recommended strategies are implemented, Dunnes Stores can expect:

- Improved campaign ROI due to higher-quality targeting
- Increased CLV from long-tenure and high-education segments
- Revenue uplift through reactivation of dormant but high-value customers
- Greater international market impact in high-response regions
- Higher digital engagement translating into more purchases
- Strengthened customer loyalty and retention

This analysis forms a strong foundation for future customer strategy initiatives such as predictive modelling, next-best-offer recommendations, and dynamic segmentation.

7. SQL Pipeline (Full Code Available in Repository)

All SQL used for this project including data validation, cleaning, feature engineering, enrichment, segmentation logic, and analytical queries is available in the project's GitHub repository.

Full SQL Script (GitHub)

You can access the complete SQL script here:

https://github.com/amruthraj007/dunnes-customer-insights-case-study/blob/main/sql/dunnes_customer_analysis.sql

8. Key Takeaways

This case study demonstrates a complete, end-to-end analytics workflow built to support customer insights and marketing strategy at Dunnes Stores. The following key takeaways summarise the most important findings from the analysis:

1. Dunnes Stores' Customer Base Is Mature, Educated, and High-Value

The core customer demographic is aged **40–60**, financially stable, and highly educated. This segment consistently drives above-average spending and engagement, making them a reliable foundation for premium product lines and loyalty initiatives.

2. Customer Tenure and Education Strongly Influence Spending

Customers with longer tenure and higher education levels exhibit significantly higher spend and stronger loyalty indicators. These groups should be prioritised for premium offerings, exclusive rewards, and membership-based retention programs.

3. Dormant Customers Remain a High-Potential Segment

Despite inactivity, Dormant customers show strong underlying purchasing power. With personalised, well-timed reactivation strategies, this group has the potential to deliver substantial revenue uplift.

4. Active Customers Deliver the Highest Marketing ROI

Active customers respond to campaigns at **24.3%**, the highest among all segments. Maintaining and enhancing engagement for this group can maximise campaign efficiency and short-term ROI.

5. Responders Spend 82% More Than Non-Responders

Customers who engage with campaigns demonstrate dramatically higher value. This confirms that targeted, data-driven campaign design materially improves revenue and retention.

6. Geography Plays a Key Role in Performance

High-value regions such as **Mexico, Spain, Germany, and Saudi Arabia** show higher spending and better response rates. Tailored regional strategies could offer competitive advantage and stronger international performance.

7. Data-Driven Segmentation Provides Clear Actionability

The segmentation logic developed—covering Recency Buckets, Responder Types, and Tenure-based cohorts—gives Dunnes Stores a clear analytical foundation for:

- Personalised communications
- Cross-sell/upsell targeting
- High-value customer retention
- Strategic reactivation planning