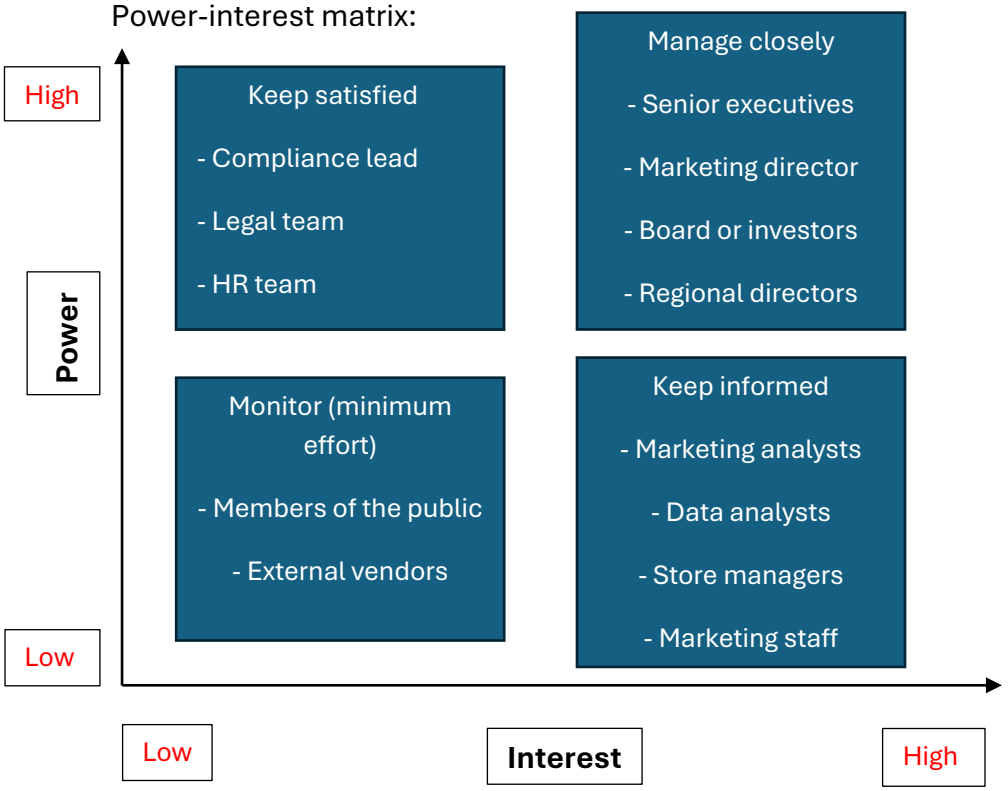
Power-interest matrix:



Background/ context of the business

Business problem

2Market has a critical knowledge gap regarding its customer segments and their purchasing behaviours. This prevents the optimisation of marketing spend and product strategy, creating significant risk to core revenue streams like liquor and slowing the growth of high-potential categories.

Questions that should be asked to better understand the data:

Customer demographics: who are our most valuable customer segments, defined by country, marital status, and education?

- Product performance: which product categories generate the most revenue and attract the most customers within each demographic segment?

- Channel effectiveness: which advertising channels deliver the highest number of lead conversions, and how does this vary by customer segment?

Business Impact

- Risk: Losing market share in our most profitable category (liquor) due to ineffective retention.

- Opportunity: Missing revenue from underperforming categories by failing to identify cross-sell targets.
- Efficiency: Wasting marketing spend on channels that do not convert key customer demographics.

Primary Goal of the Analysis
- Protect and Grow Liquor Revenue: By identifying high-value customer segments for targeted retention campaigns.
- Lift Underperforming Categories: By uncovering data-driven opportunities for cross-selling to increase overall basket value.

Analytical approach

The initial data preparation stage focused on validation and type consistency. Excel was used as a preliminary inspection tool to review the raw dataset structure and confirm that columns were logically aligned with their intended use. Particular attention was paid to identifying numeric fields stored as text, such as income and product spend columns. These fields were converted to numeric formats to ensure they could be reliably aggregated during analysis. Excel was also used to scan for obvious inconsistencies, missing values, and implausible entries before importing into the PostgreSQL. Heavy transformations were deliberately avoided at this stage to preserve the original data structure for reproducibility.

The cleaned dataset was then loaded into a PostgreSQL database, where formal validation and analysis were conducted. SQL was chosen over Excel for the core analysis due to its robustness for aggregation, segmentation, and reproducibility across large datasets. A primary key check was performed in SQL to confirm ID uniqueness; no duplicate customer records were found, validating the dataset's suitability for stakeholder-level analysis. SQL was also used to enforce correct data types and ensure consistency across joins and aggregations.

A key methodological decision involved resolving ambiguity around the concept of "product popularity." Rather than relying on a single metric, popularity was defined using three complementary measures: total revenue per product category, number of unique customers purchasing each product, and average spend per purchasing customer. This multi-metric approach avoided misleading conclusions that could arise from revenue alone, such as high revenue driven by a small number of extreme spenders. When calculating average spend, zero values were explicitly excluded using conditional logic to ensure averages reflected actual purchasing behavior rather than non-buyers.

Further SQL analysis incorporated grouping, filtering, and joins to explore performance across demographics such as country and marital status, as well as advertising effectiveness based on social media conversions.

Finally, Tableau was used to visualise the results and support interpretation. Visualisations were designed to compare revenue composition, identify dominant product categories, and highlight differences across demographic segments.

Dashboard design and development

The dashboard design was developed to support data-driven decision-making across multiple stakeholder groups by presenting complex analytical outputs in a clear, structured, and accessible format. The overall design philosophy prioritises interpretability, comparability, and narrative flow, ensuring that insights can be quickly understood by senior executives while remaining analytically transparent for operational teams.

Three dashboards were planned to reflect distinct analytical objectives. The first dashboard focuses on customer demographics and revenue distribution, combining a map of total spend by country with bar charts segmented by age, marital status and education level. A map visualisation was selected for country-level spend to provide immediate spatial context and highlight geographic concentration and expansion opportunities. Bar charts were chosen for demographic comparisons because they allow accurate comparison of magnitude across categories, avoiding the ambiguity that can arise from pie or radial charts when values differ substantially.

The second dashboard focuses on identifying the most effective advertising channels by directly comparing advertising volume with realised conversions. It contains three charts: a bar chart showing the number of Facebook advertisements by country, a bar chart of social media conversions by marital status colour coded for Instagram and Twitter, and a summary bar chart of total conversions by platform. Facebook is absent from the marital status conversion chart due to negligible conversion volume. This intentional contrast shows that high Facebook advertising activity does not translate into conversions, whereas Instagram clearly outperforms other platforms across all customer segments.

The third dashboard analyses product performance by separating revenue dominance from customer reach. A heat map shows revenue per product category by country, highlighting Spain as the leading liquor market. A revenue bar chart by marital status shows liquor generating the most revenue, while a final bar chart shows meat items attract the most customers overall.

Colour choices were kept consistent across dashboards, with product categories and platforms assigned stable, high-contrast colours to support pattern recognition and accessibility. Layouts follow a top-down narrative, moving from high-level context to detailed breakdowns, aligning with executive reading behaviour.

Pattern, trends, and insights

Several clear and consistent patterns emerged from the analysis across customer demographics, product performance, and marketing effectiveness. Most notably, liquor is the dominant revenue driver across all demographic segments, including country, marital status, and education level. This dominance is visible in both total revenue and average spend metrics, indicating that liquor contributes not only through broad customer participation but also through relatively high per-customer value. This pattern aligns with established retail research, which identifies alcoholic beverages as high-margin anchor products that drive repeat purchasing and basket expansion.

In contrast, meat products consistently attract the largest number of unique customers, particularly among married and cohabiting households. This suggests strong market penetration but comparatively lower average spend per customer, indicating an opportunity to increase value through premium offerings or bundling strategies. Lower-performing categories such as vegetables, chocolates, and commodities show limited revenue contribution across most segments, highlighting gaps in product engagement rather than isolated underperformance.

Geographically, spending is concentrated in a small number of countries, with Spain and South Africa contributing disproportionately to total revenue.