

# Assignment Solution: Task 2 - Business Analysis

[Executive Summary: Wolt Snack Shop Performance](#)

[Customer Retention & Growth Analysis](#)

[Data Source & Methodology](#)

[Chart 1: Customer Retention Heat Map](#)

[Chart 2: Customer Base Evolution by Segment](#)

[Chart 3: New Customer Quality \(Organic vs. Promo\)](#)

[Category & Product Performance Analysis](#)

[Data Source & Methodology](#)

[Chart 1: Category Performance Leaderboard](#)

[Chart 2: Category Health Composition](#)

[Chart 3: Seasonal Consumption Pattern](#)

[Chart 4: Star Products](#)

[Chart 5: Monthly Trend](#)

[Product Affinity & Bundling Analysis](#)

[Chart 1: Top Product Combinations in basket](#)

[Other use case: Finding Product Affinity](#)

# Assignment Solution: Task 2 - Business Analysis

Building on the data models from Task 1, this analysis evaluates the "Wolt Snack Shop" project's viability and growth drivers for the Business Team. The following insights address key questions regarding customer retention, category health, and bundling opportunities.

## Deliverables & Visualization Note:

- **Datasets:** All processed datasets used in this analysis are available in the [Datasets/BI/](#) folder.
- **Dashboarding:** Given more time, I would have deployed these visualizations as an interactive Tableau Dashboard.

## Executive Summary: Wolt Snack Shop Performance

**Overview:** Wolt Snack Shop demonstrates strong product-market fit, driven by habit-forming service rather than discounts. While retention is excellent, revenue is critically dependent on just two categories, creating a significant vulnerability

### Key Findings & Strategic Recommendations

- **High-Quality Growth:** Growth is organic, not promo-driven. New cohorts show strong ~73% Month 1 retention , and >50% of active users have graduated to high-loyalty segments.
- **Category Risks & Seasonality:** Revenue relies heavily on *Crisps* and *Chocolate* (generating 2x other categories). *Chocolate* is highly seasonal, peaking at €72k in Fall before dropping in Spring.
  - *Recommendation:* Aggressively manage inventory flips to balance *Chocolate* (Winter peak) against *Other Confectionary* (Summer peak).
- **Bundling Opportunities:** The *Fuego Mild Dip + Funny-Frisch Chips* pair sells 3.5x more than the next best combination.
  - *Recommendation:* Use star products (like *Funny-Frisch*) in "Mix & Match" promos to anchor and drive sales of lower-volume items.

# Customer Retention & Growth Analysis

## Data Source & Methodology

The following analysis is built using the `mart_customer_monthly_summary` data mart.

This mart provides a complete history of every customer's behavior, allowing us to generate a "`customer_analysis`" dataset. This dataset aggregates millions of raw transaction rows into high-level metrics on **Cohorts** (when they started) and **Segments** (how loyal they are now).

**SQL Query:** The following query was used to generate the dataset feeding the charts below:

```
SQL
SELECT
    -- 1. Time Dimensions
    year,
    month,
    month_start_date,

    -- 2. Cohort Dimensions (For Retention)
    cohort_month_label,
    months_since_first_purchase,

    -- 3. Segment Dimensions (For Growth)
    customer_segment_at_end_of_month,
    is_first_month,

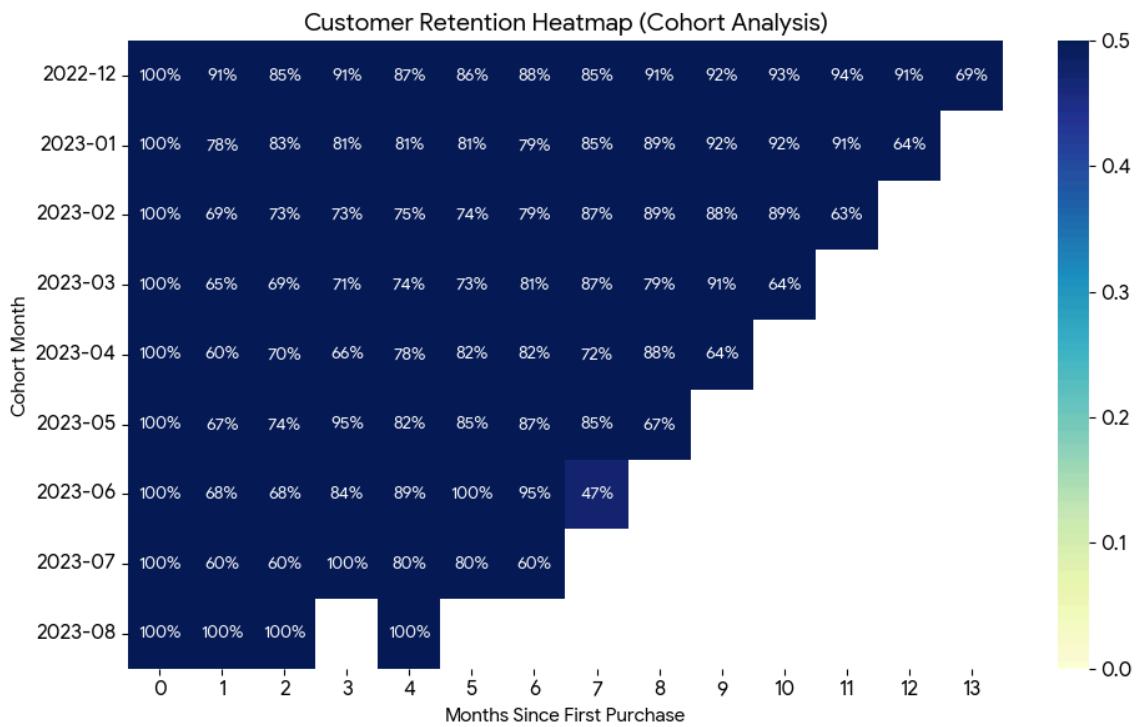
    -- 4. Key Metrics
    COUNT(DISTINCT customer_key) AS active_customers,
    SUM(revenue_this_month) AS total_revenue,

    -- 5. Acquisition Quality Metrics
    SUM(CASE WHEN promotion_only_purchase_first_month = true THEN 1 ELSE 0 END)
AS count_promo_only_new_customers,
    AVG(promotion_usage_percentage_till_date) AS avg_promo_reliance

FROM retail_data_model.marts.mart_customer_monthly_summary
GROUP BY 1, 2, 3, 4, 5, 6, 7
ORDER BY 1, 2
```

The dataset is at: [Datasets/BI/customer\\_analysis.csv location](#)

## Chart 1: Customer Retention Heat Map

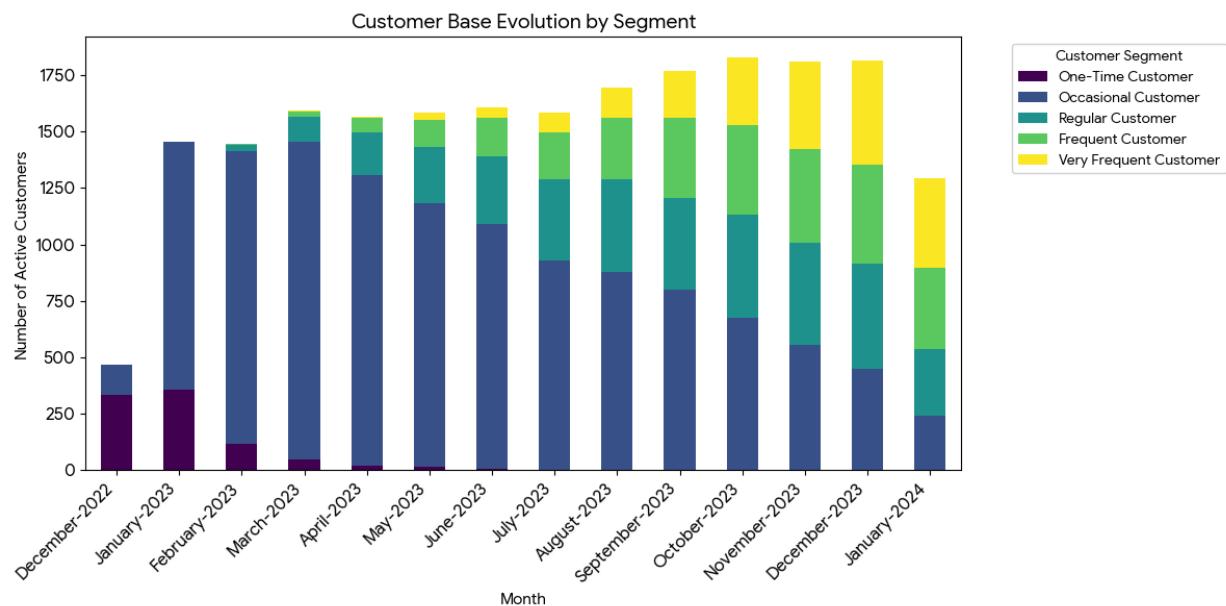


**Purpose:** Answer the question, "Once we acquire a customer, do they come back?"

### Analysis:

- **High "Stickiness":** We see an exceptionally high ~73% Month 1 Retention Rate (average). This means nearly 3 out of 4 new customers return to the store in their second month.
- **Consistent Quality:** The cohorts (rows) show very stable blue patterns. This indicates that despite scaling up, the quality of our service and product has not degraded - new customers today are just as loyal as those who joined months ago.

## Chart 2: Customer Base Evolution by Segment

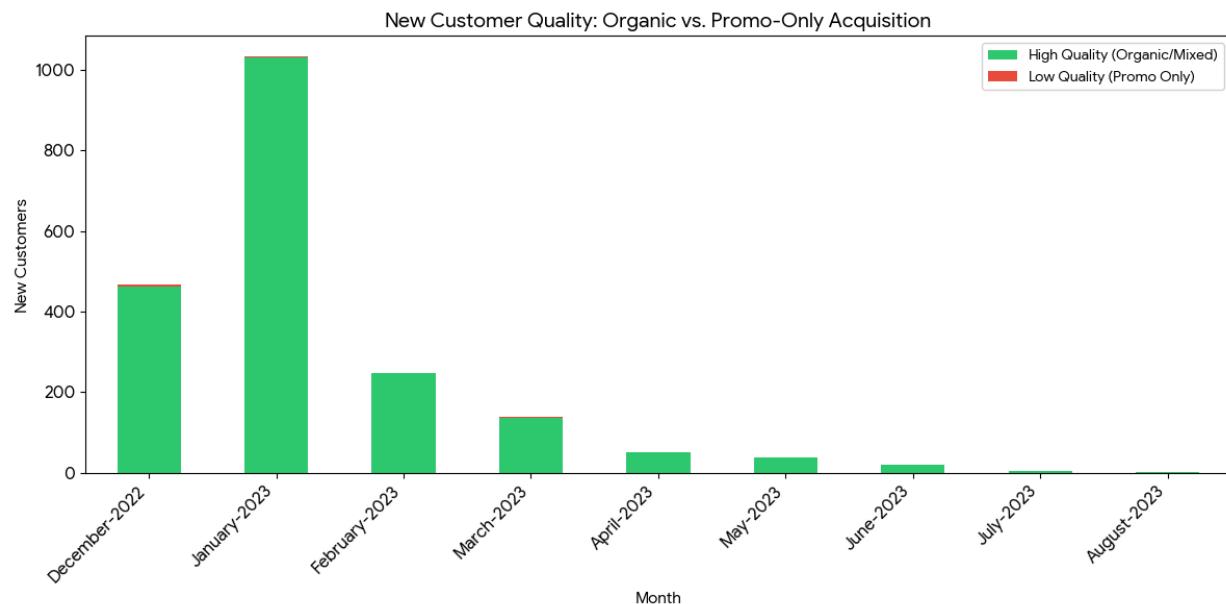


**Purpose:** Answers the question, *"Are our customers growing in value over time?"*

### Analysis:

- **The Graduation Engine:** We are successfully moving customers up the loyalty ladder. The "Very Frequent" (Yellow) and "Frequent" (Green) segments are growing steadily every month.
- **Healthy Maturation:** As time goes on, a larger percentage of our active user base is made up of these high-value, habitual shoppers. In the most recent months, these loyal segments make up over 50% of our active base, proving we are building a permanent, habit-based business.

### Chart 3: New Customer Quality (Organic vs. Promo)



**Purpose:** Answers the question, *"Are we just buying growth with discounts?"*

#### Analysis:

- **Organic First:** The vast majority of new customers (Green bars) are acquiring organically or buying full-priced items. They are **not** "promo hunters."
- **No Discount Dependency:** The "Promo Only" segment (Red bars) is negligible. This confirms that our acquisition strategy is high-quality. We attract users who want the product for its value, not just because they found a coupon code.

## Category & Product Performance Analysis

### Data Source & Methodology

The following analysis is built using the `mart_category_product_performance` data mart.

This mart enables analysis by tracking every product's market share, growth momentum, and profitability. We generated a "`category_health`" dataset to answer questions about assortment efficiency and seasonal demand.

**SQL Query:** The following query was used to generate the dataset feeding the charts below:

```
SQL
SELECT
    -- 1. Dimensions
    year,
    month,
    season,
    item_category,
    performance_category, -- (e.g., 'Star Product', 'Underperformer')

    -- 2. Aggregated Metrics
    COUNT(DISTINCT item_key) AS product_count,
    SUM(revenue) AS total_revenue,
    SUM(quantity_sold) AS total_units_sold,

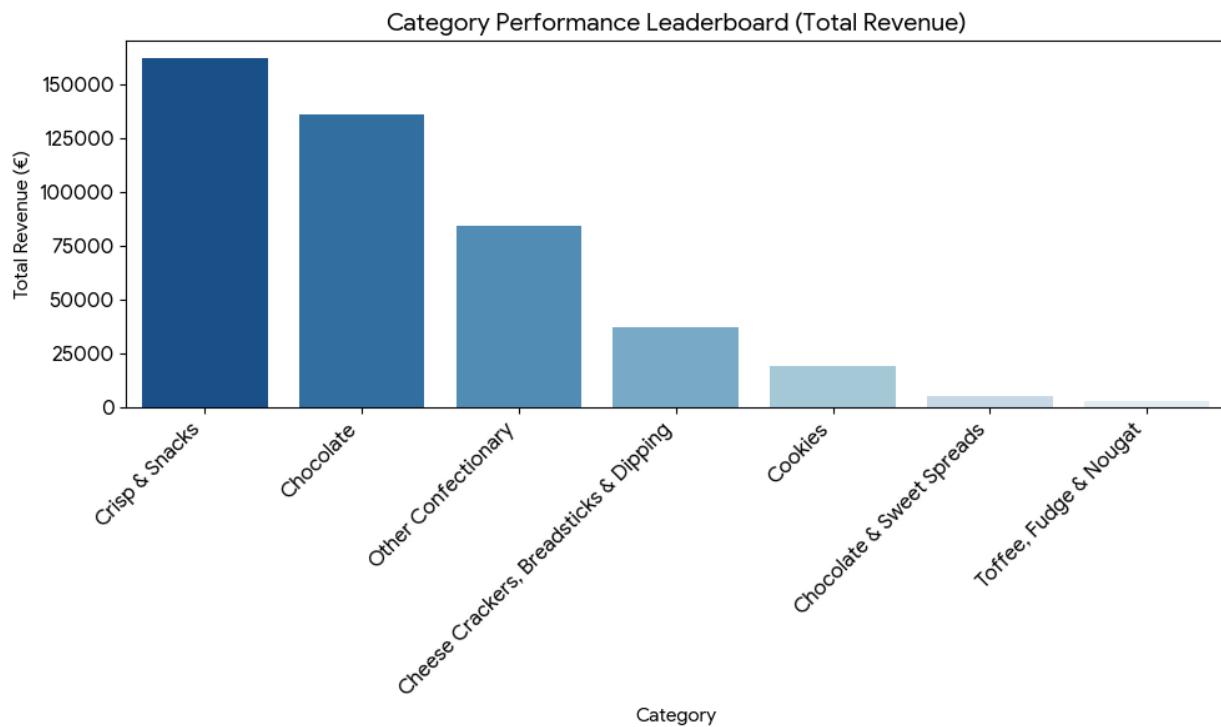
    -- 3. Price Metric (Realized Price)
    -- Calculates what customers actually paid, accounting for discounts
    SUM(revenue) / NULLIF(SUM(quantity_sold), 0) AS avg_realized_price,

    -- 4. Weighted Growth
    SUM(revenue * revenue_growth_mom_pct) / NULLIF(SUM(revenue), 0) AS
    weighted_avg_growth_pct

FROM retail_data_model.marts.mart_category_product_performance
GROUP BY 1, 2, 3, 4, 5
ORDER BY total_revenue DESC
```

The dataset is at: [Datasets/BI/category\\_health.csv location](#)

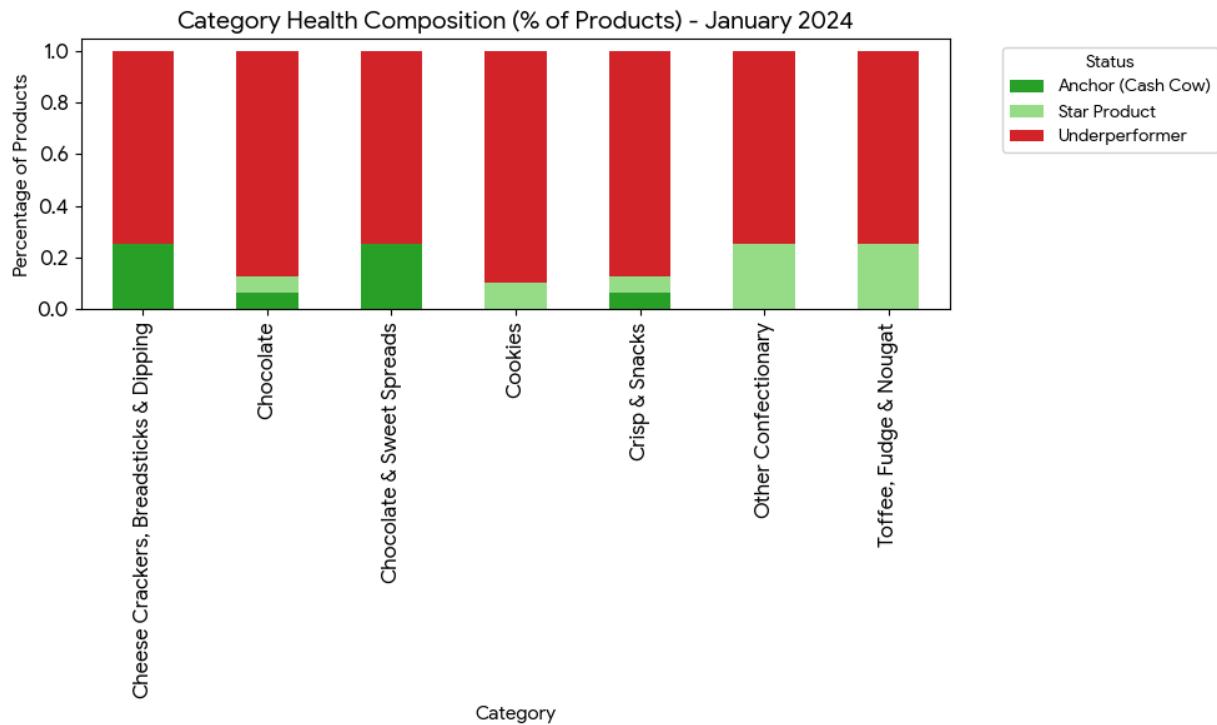
## Chart 1: Category Performance Leaderboard



### Analysis:

- **Crisps & Snacks** and **Chocolate** are the leaders. They generate nearly **2x** the revenue of any other category
- **Risk Concentration:** While this looks good, it means the business is highly vulnerable. A supply chain issue or trend shift in just *one* of these two categories would cripple overall performance. The "Long Tail" categories (Toffee, Spreads) are currently too small to act as a buffer.

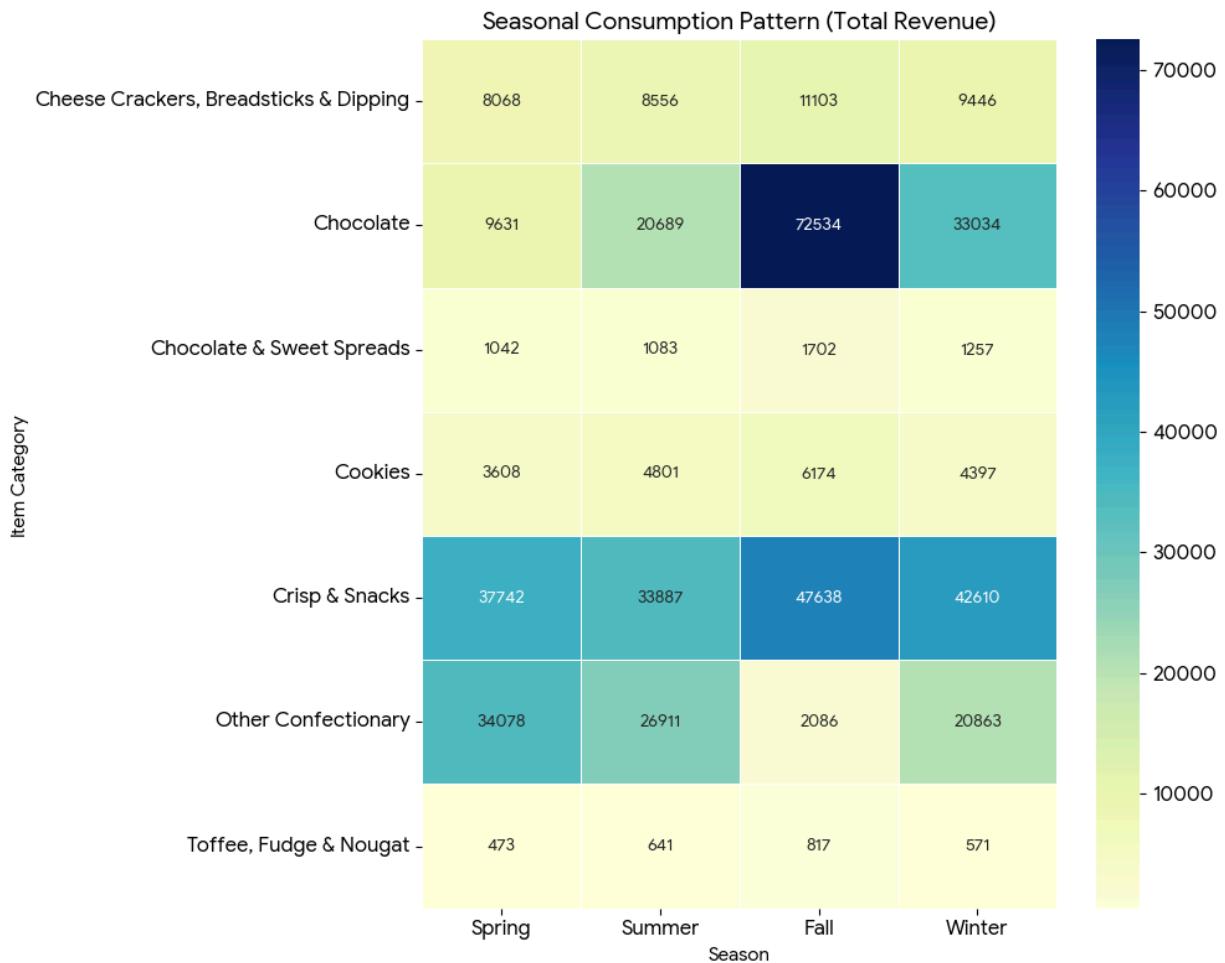
## Chart 2: Category Health Composition



### Analysis and Actions:

- **Biggest categories by revenue, Crisps & Snacks and Chocolate** have a lot of underperforming products which implies high assortment inefficiencies.
- A deep dive on bottom 30% (or X%) of products for inventory and assortment cleanup can be a next step

**Chart 3: Seasonal Consumption Pattern**



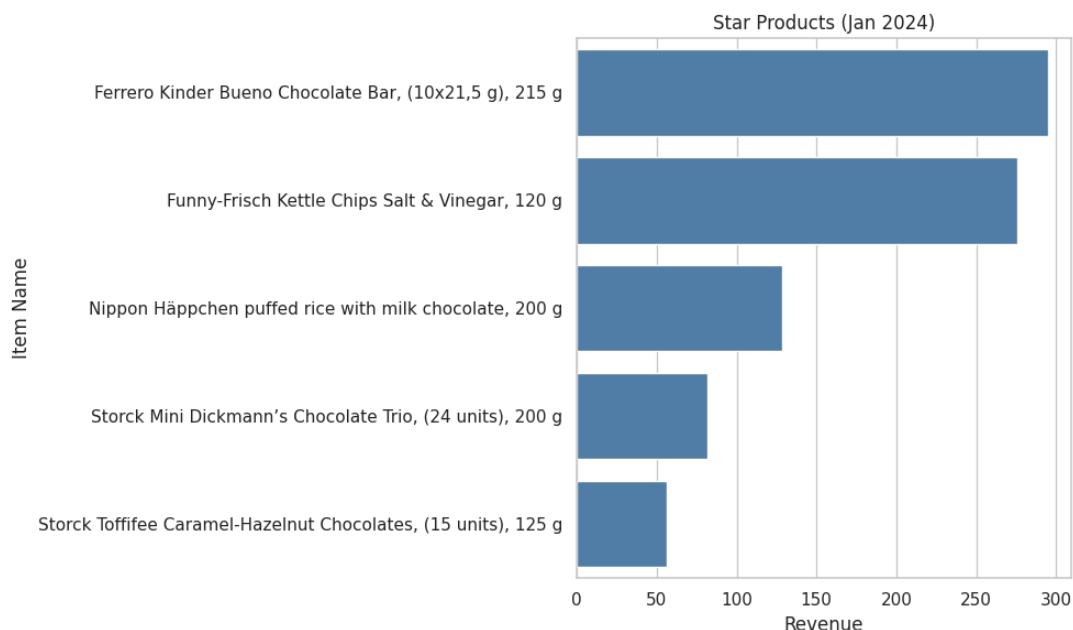
### Analysis:

- **Chocolate:** Highly seasonal. Revenue explodes in Fall (72k) and Winter, but is low in Spring (9.6k)
- **Other Confectionary:** The counter-balance. Peaks in Spring (34k) and Summer, effectively zero in Fall.
- **Crisps & Snacks:** Strong year-round performer, peaking in Fall/Winter (47k).
- **Cookies/Crackers:** Stable baselines with minor Fall uplifts.

## Actions

- **Inventory Flip:** As Summer ends, aggressively clear *Other Confectionary* to make space for *Chocolate* stock.
- **Fall Bundles:** Create "Party Packs" combining *Chocolate* and *Crisps* to leverage their simultaneous Fall peaks.
- **Spring Push:** Launch *Other Confectionary* campaigns early in Spring to offset the post-Winter drop in *Chocolate* sales.
- **Price Tweak:** Test slight price increases on *Cheese Crackers* in Fall when demand is naturally inelastic.

## Chart 4: Star Products



**Analysis:** We only have 5 Star Products

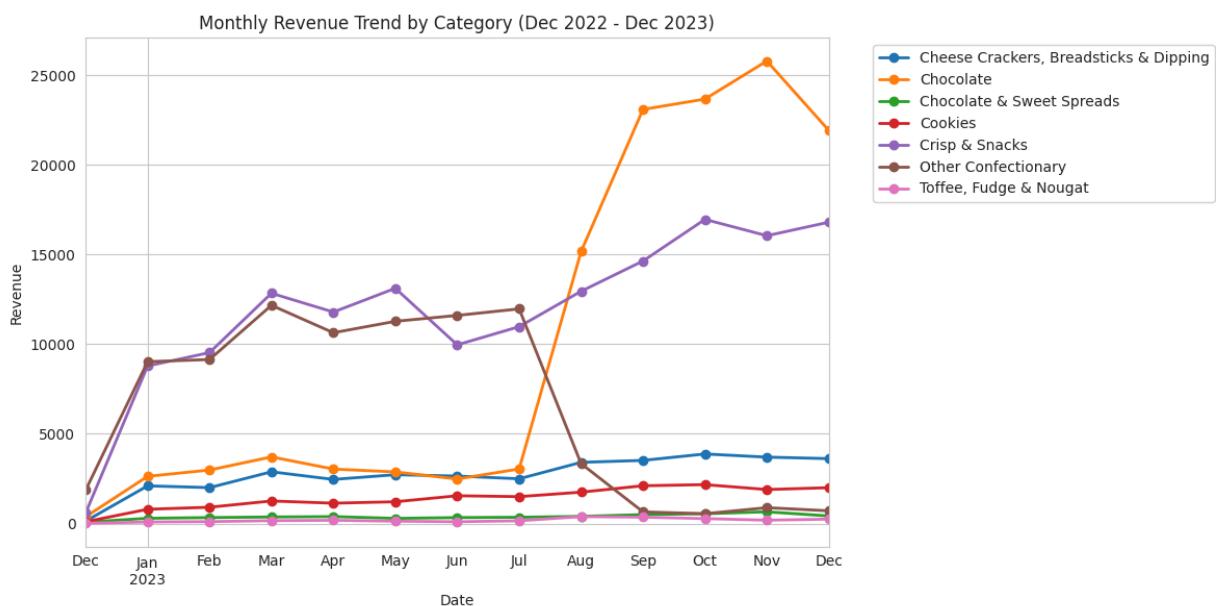
- *Kinder Bueno*
- *Funny-Frisch Kettle Chips*
- *Nippon Puffed Rice*
- *Storck Mini Dickmann's*
- *Storck Toffifee*

**Action:** Protect these 5 items at all costs. Ensure 100% availability. If we run out of stock, we lose your primary growth driver

SQL

```
--SQL code for Chart 4: Star Products
SELECT year,
       month,
       item_name_en,
       SUM(revenue) revenue
FROM retail_data_model.marts.mart_category_product_performance
WHERE performance_category = 'Star Product' --This can come from BI tool
AND year = 2024 AND month = 1 --This can come from BI tool
GROUP BY all
ORDER BY 4 DESC
```

## Chart 5: Monthly Trend



The above line graph gives same insights as mentioned earlier in [here](#), but it gives a monthly snapshot of what's happening at Category Level

### Analysis:

- We can clearly see the "Chocolate" (Orange line) spike beginning from August, validating its seasonal nature.
- "Other Confectionary" (Brown line) drop is visible as a cliff-edge in late summer, confirming the urgent need for investigation.

# Product Affinity & Bundling Analysis

The following analysis is built using the `mart_product_baskets` data mart

This mart provides a transaction-level view of every basket purchased, including the full list of items, total value, and customer details. We utilized this to generate a "**basket\_composition**" dataset. This dataset aggregates individual purchase records to identify exact product combinations that appear most frequently, filtering out single-item purchases to focus specifically on bundling behavior.

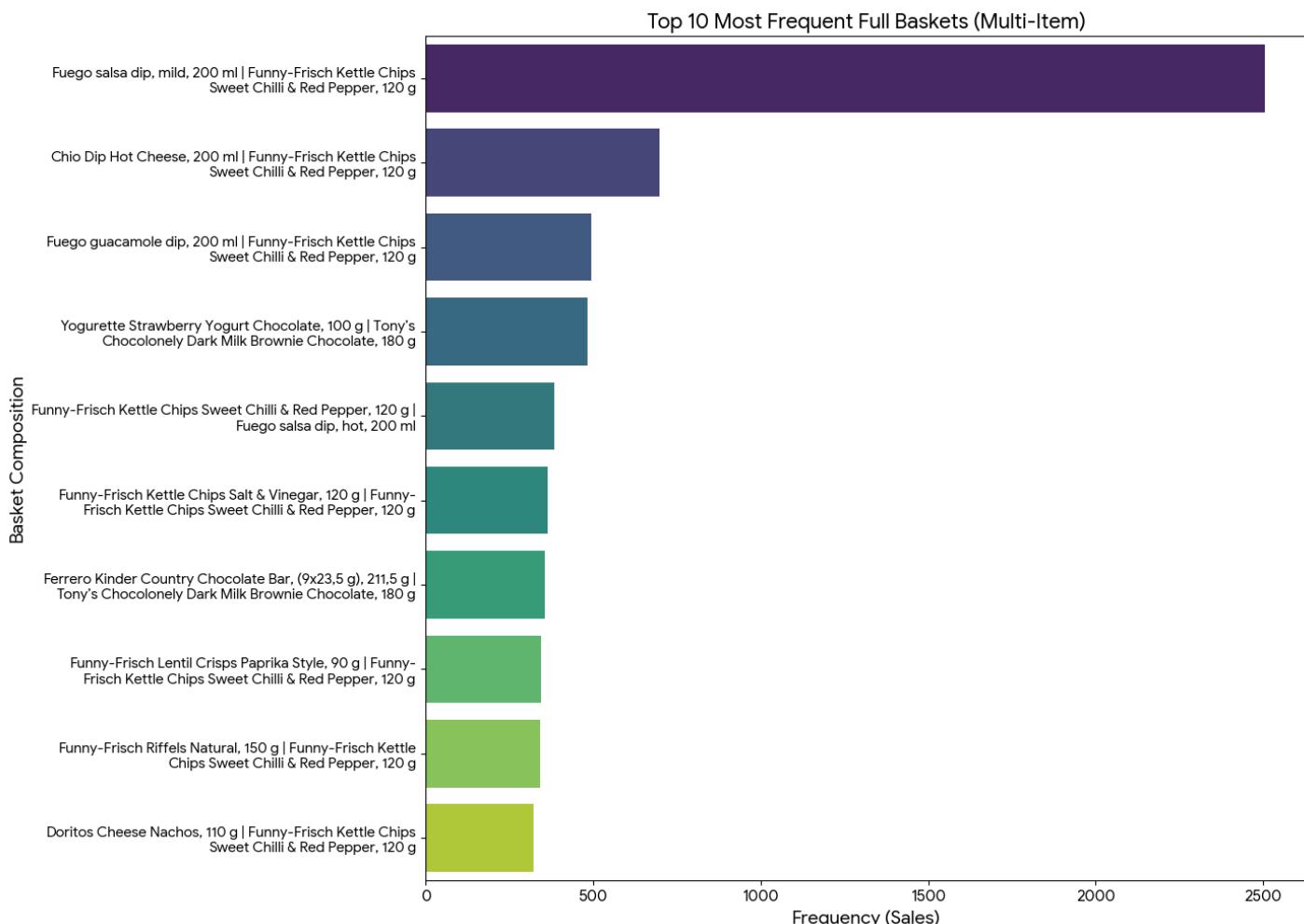
## SQL Query

The following query was used to generate the dataset feeding the charts below:

```
SQL
SELECT
    item_names_list,
    COUNT(*) AS sales
FROM
    retail_data_model.marts.mart_product_baskets
WHERE
    unique_items > 1
GROUP BY
    item_names_list
ORDER BY
    sales DESC
LIMIT 10;
```

The dataset is at: [Datasets/BI/basket\\_composition.csv location](#)

## Chart 1: Top Product Combinations in basket



## Insights & Actions

- **Dominant Combo:** The "Fuego Mild Dip + Funny-Frisch Sweet Chilli Chips" pair leads with 2,504 sales (3.5x higher than #2).
  - **Action:** Create a pre-bundled "Party Pack" SKU with these two items to speed up checkout.
- **Brand Anchoring:** Funny-Frisch chips appear in nearly every top savoury basket.
  - **Action:** Launch a "Mix & Match" promo (2 Chips + 1 Dip) to use this anchor brand to sell lower-volume dips.
- **Chocolate Niche:** A distinct group buys Tony's Dark Milk Brownie + Yogurette Strawberry together (484 sales).
  - **Action:** Trigger a "Perfect Pairing" recommendation for Yogurette whenever Tony's chocolate is added to the cart.

## Other use case: Finding Product Affinity

This data mart empowers business stakeholders to identify high-value product pairings (cross-selling opportunities). By utilizing this dataset, users can select a primary product in the BI tool - such as 'Product A' - and instantly view a ranked list of products most frequently purchased alongside it. This insight is essential for optimizing bundle offers

SQL

```
SELECT trim(split(item_names_list, '|')[1]) AS Product_A,
       trim(split(item_names_list, '|')[2]) AS Product_B,
       COUNT(*) sales
  FROM retail_data_model.marts.mart_product_baskets
 WHERE unique_items > 1
   AND Product_A = 'Fuego salsa dip, mild, 200 ml' --This is an example, this can
                                                 be filtered in BI tool
 GROUP BY all
 ORDER BY 3 DESC
```

The screenshot shows a database interface with a SQL query editor and a results table.

**SQL Query:**

```
1  SELECT trim(split(item_names_list, '|')[1]) AS Product_A,
2      trim(split(item_names_list, '|')[2]) AS Product_B,
3      COUNT(*) sales
4  FROM retail_data_model.marts.mart_product_baskets
5  WHERE unique_items > 1
6  AND Product_A = 'Fuego salsa dip, mild, 200 ml'
7  GROUP BY all
8  ORDER BY 3 DESC
```

**Results:**

12 rows returned in 39ms

T Product_A	T Product_B	123 sales
Fuego salsa dip, mild, 200 ml	Funny-Frisch Kettle Chips Sweet Chilli & Red Pe...	2565
Fuego salsa dip, mild, 200 ml	Lay's Salt & Vinegar Crisps, 150 g	279
Fuego salsa dip, mild, 200 ml	Crunchips Western Style Potato Chips 150 g	268
Fuego salsa dip, mild, 200 ml	Chio Tortillas Original Salted, 110 g	235
Fuego salsa dip, mild, 200 ml	Chio Dip Hot Cheese, 200 ml	171
Fuego salsa dip, mild, 200 ml	Pringles Sour Cream & Onion Chips, 165 g	117
Fuego salsa dip, mild, 200 ml	Fuego guacamole dip, 200 ml	112
Fuego salsa dip, mild, 200 ml	Funny-Frisch Lentil Crisps Paprika Style, 90 g	95
Fuego salsa dip, mild, 200 ml	Crunchips Cheese & Onion Crisps, 150 g	88
Fuego salsa dip, mild, 200 ml	Lay's Crips herb butter flavour, 150 g	75
Fuego salsa dip, mild, 200 ml	Funny-Frisch Kettle Chips Salt & Vinegar, 120 g	55
Fuego salsa dip, mild, 200 ml	Fuego salsa dip, hot, 200 ml	33

Example dataset