

Overview and Objectives

Our goal was to identify strong, actionable associations between products—pairs of items that frequently sell together—in order to inform cross-selling strategies, promotional bundles, and store layout decisions. To do this, we analyzed the complete set of association rules extracted from transaction data (available in `association_rules.csv`) and applied conservative thresholds to filter down to only the most reliable relationships. Below is a detailed narrative of our methodology, findings, and recommendations that can be shared with stakeholders.

1. Methodology & Filtering Criteria

1. Dataset Description

The full CSV contains hundreds of rules of the form “Antecedent \rightarrow Consequent,” where:

- **Antecedent** is the item or item-set on the left (e.g., `{Bread}`).
- **Consequent** is the item or item-set on the right that appears in the same basket (e.g., `{Eggs}`).
- Each rule is accompanied by metrics: **support**, **confidence**, **lift**, and various secondary measures (leverage, conviction, etc.).

2. Threshold Selection

To ensure that we only focus on robust, frequently occurring associations, we applied the following cutoffs:

- **Support > 0.10**: The pair must appear in at least 10% of all transactions, so it isn't a rare combination.
- **Confidence > 0.50**: Among all transactions containing the antecedent, more than half also contain the consequent, indicating a strong directional relationship.
- **Lift > 1.20**: The observed co-occurrence is at least 20% more likely than if the products were purchased independently. A lift significantly above 1 suggests a genuine inter-dependence rather than random chance.

After filtering, only five rules (out of the original hundreds) satisfied all three criteria. We prioritized these because they are both common and strong—making them ideal candidates for cross-sell initiatives.

2. Key Metrics Distributions

Before diving into specific rules, it's helpful to understand how support, confidence, and lift are distributed across the entire rule set:

1. Support Distribution

- The vast majority of rules have very low support (below 5%), showing that most item combinations are rare.
- A small tail of rules exhibits support between 10% and 35%. We selected our threshold (10%) at a point where we capture these more common patterns while excluding the long tail of rare combinations.

2. Confidence Distribution

- Many rules cluster below 0.30–0.40 confidence, indicating that even if two items appear together at all, it is usually in a minority of baskets that contain the antecedent.
- Fewer rules achieve confidence above 0.50. By choosing 0.50 as our cutoff, we ensure that more than half of transactions containing “Antecedent” also contain “Consequent.”

3. Lift Distribution

- Most rules have lift values hovering around 1.0–1.10, indicating that they co-occur roughly as often as would be expected by pure chance.
- A much smaller set of rules exhibits lift above 1.20. These are the rules where co-occurrence is meaningfully higher than random, suggesting a true association worth capitalizing on.

4. Support vs. Confidence Scatter (colored by Lift)

- When plotted on a two-axis chart (support on the X-axis, confidence on the Y-axis), most rules cluster in the bottom left corner (low support, low confidence).
- Only a handful of points lie in the quadrant where support > 0.10 and confidence > 0.50. The color gradient (highest lift values appearing in a darker shade) highlights exactly which rules not only meet thresholds on support and confidence but also have elevated lift. Those points become our focus.

3. Detailed Findings: Top Filtered Rules

After applying our thresholds, we identified **five** rules that truly stand out. They are listed below in descending order of lift (i.e., strength of association relative to independent purchase probability). For each, we describe the metric values, interpret them in business terms, and suggest potential actions.

3.1 {Cheese} → {Bread}

- **Support:** 0.127 (12.7% of all transactions contain both cheese and bread)
- **Confidence:** 0.737 (73.7% of all transactions that include cheese also include bread)
- **Lift:** 2.191
 - This lift signifies that shoppers who buy cheese are over **2.19 times** more likely to buy bread compared to a shopper chosen at random.
- **Interpretation:** Cheese purchasers almost always pick up bread as well—beyond what we'd expect purely by overall purchase rates.
- **Actionable Insights:**
 - i. Display cheese and bread side-by-side (end-cap shelf, combo signage).
 - ii. Offer a “Cheese & Bread” bundle discount (e.g., 10% off when both items are bought together).
 - iii. Promote “perfect pairing” recipes or sampling events (e.g., highlight grilled cheese sandwiches).

3.2 {Bread} → {Eggs}

- **Support:** 0.282 (28.2% of all transactions contain both bread and eggs)
- **Confidence:** 0.838 (83.8% of transactions containing bread also contain eggs)
- **Lift:** 1.486
 - Customers who buy bread are about 1.49 times more likely to buy eggs than if the two items were uncorrelated.
- **Interpretation:** Bread is a strong predictor that eggs will be in the same basket—an unsurprising but quantifiably reliable breakfast combo.
- **Actionable Insights:**
 - i. Create a “Breakfast Starter Pack” with bread and eggs (and potentially butter or milk) featured together in promotional materials.
 - ii. Place egg cartons next to bread shelves to cue shoppers to add both.
 - iii. Tie in a loyalty-point incentive (“Buy bread & eggs today to earn a bonus point”).

3.3 {Bread} → {Milk}

- **Support:** 0.273 (27.3% of transactions have both bread and milk)
- **Confidence:** 0.811 (81.1% of baskets with bread also have milk)

- **Lift:** 1.416
 - Shoppers who buy bread are 1.42 times more likely than average to also purchase milk.
- **Interpretation:** Bread often goes hand-in-hand with milk—another staple pairing for households stocking up on basics.
- **Actionable Insights:**
 - i. Bundle signage: “Bread + Milk for Easy Breakfasts.”
 - ii. Run a limited-time “Buy bread, get 10% off milk” coupon.
 - iii. Feature bread and milk in circular ads or emails targeting families.

3.4 {Cheese} → {Urban}

- **Support:** 0.118 (11.8% of transactions have both cheese and the product “Urban”)
- **Confidence:** 0.684 (68.4% of cheese-buying baskets include “Urban”)
- **Lift:** 1.298
 - Cheese customers are 1.30 times more likely than a random shopper to also buy “Urban.”
- **Interpretation:** While “Urban” may represent a specialty or premium item, it is frequently sold alongside cheese. This could indicate that “Urban” complements cheese—perhaps as a particular kind of spread, cracker, or complementary gourmet product.
- **Actionable Insights:**
 - i. Consider co-placement of “Urban” next to cheese displays, especially if “Urban” is a high-margin item.
 - ii. Introduce sampling stations where customers can taste cheese with “Urban” to spur incremental purchases.
 - iii. Include “Urban” in meal-prep recipe cards featuring cheese (e.g., cheese board recommendations).

3.5 {Juice} → {Urban}

- **Support:** 0.209 (20.9% of all transactions have both juice and “Urban”)
- **Confidence:** 0.639 (63.9% of juice-buying baskets include “Urban”)
- **Lift:** 1.212
 - Purchasers of juice are about 21% more likely than average to also buy “Urban.”
- **Interpretation:** There may be an implied lifestyle or thematic connection—perhaps shoppers picking up “Urban” (a gourmet item, snack, or specialty brand) also like premium or fresh juices.
- **Actionable Insights:**

- i. Cross-merchandise “Urban” with juice in a “Healthy & Gourmet” end-cap.
- ii. Create an in-store demo: pair juice samples with “Urban” tastings to highlight how they complement each other.
- iii. Bundle “Urban” with assorted juices in a gift pack for special occasions or corporate events.

4. Business Implications & Recommendations

1. High-Impact Cross-Sell Opportunities

- The three “Bread → X” rules particularly stand out because bread is in over 33% of all baskets. By targeting bread as a “gateway” item (given its high prevalence), we can influence three separate but strong cross-sell relationships:
 - **Bread & Eggs** (Lift 1.49, Confidence 0.84)
 - **Bread & Milk** (Lift 1.42, Confidence 0.81)
 - **Bread & Cheese** (via Cheese → Bread, Lift 2.19, Confidence 0.74)
- Essentially, each loaf of bread presents a 70–80% chance to also move eggs, milk, or cheese. Promotional strategies around bread will therefore likely yield high incremental revenue.

2. Targeted Promotional Bundles

- **Breakfast Bundle:** Combine bread, eggs, and milk into a package deal or offer a “buy any two, get 10% off the third” discount.
- **Cheese & Bread Premium Pair:** Since the lift of 2.19 indicates a very strong affinity, consider a “Cheese Lover’s Pair” with two complementary cheese types and a discount on artisan bread.

3. Store Layout and Merchandising

- **Adjacency Placement:** Shelving and end-caps can be arranged so that:
 - Bread is near the eggs and dairy sections (milk and cheese).
 - Specialty “Urban” items are placed adjacent to cheese and juice displays.
- This subtle product adjacency encourages “grab what you see” incremental purchases rather than sending customers on a quest to locate a second item elsewhere in the store.

4. Marketing & Communication

- **Digital Promotions:** Send targeted email or app notifications to loyalty members who recently purchased cheese—suggesting they add bread for a 10% off combo.
- **In-Store Signage:** Use signboards that explicitly highlight “Did you know 74% of cheese buyers also buy bread? Grab them together today!”; similarly for bread & eggs (84%) or bread & milk (81%).

- **Recipe Cards & Sampling:** Offer recipe cards in produce or dairy sections that illustrate how to use pairs like cheese + “Urban,” cheese + bread, or juice + “Urban,” with tear-off coupons attached.

5. Limitations & Next Steps

1. Transactional Context

- These rules reflect only co-occurrence and do not capture *why* customers buy items together. For example, “Urban” may be a seasonal item or part of a rotating promotion. Any further promotional strategy should treat these insights as one input among others (e.g., margin data, vendor agreements, seasonal trends).

2. Time Sensitivity

- The underlying data may have been collected over a specific time window (e.g., last quarter). If product assortments or customer behavior shifts—say, due to a seasonal event or a special promotion—the rules’ strength could change.
- **Recommendation:** Re-run the association rule mining process every 3–6 months (or after a major promotion/season) to confirm that these associations remain stable.

3. Subset Analysis

- We used relatively conservative thresholds (support > 10%, confidence > 50%, lift > 1.20). If stakeholders want to explore more niche cross-sells (e.g., product affinities within a particular category or during a particular promotional period), thresholds could be temporarily lowered (e.g., support > 5%) to surface less frequent but potentially high-margin opportunities.

4. Additional Metrics

- We focused on support, confidence, and lift—three of the most common metrics. However, secondary measures like **leverage** (absolute difference between observed and expected co-occurrence) and **conviction** (how much more often the rule holds than it would if antecedent and consequent were independent) could also identify “unexpected” pairings worth exploring.
- **Recommendation:** In a follow-up exercise, we can sort rules by highest leverage or conviction to see if any high-margin or promotional items surface outside our current shortlist.

6. Summary & Stakeholder Takeaways

- **Key Findings**

- i. Five high-confidence, high-lift rules were identified, with the standout being {Cheese} → {Bread} (lift = 2.19).
- ii. Three “Bread → X” rules (Eggs, Milk, and via Cheese → Bread) illustrate bread’s power as a hub item, present in roughly one-third of all transactions.
- iii. Two “Urban” rules tie a specialty product to both juice and cheese, suggesting an emerging “gourmet” or lifestyle cluster to explore.

- **Business Impact**

- By capitalizing on these strong associations—especially bread-centric pairings—we can drive incremental revenue through cross-promotions, bundling, and more persuasive store layouts.
- Targeted marketing campaigns (both digital and in-store) around these specific item pairs are likely to see high ROI, given the already elevated probability of joint purchase.

- **Next Steps**

- i. Implement cross-sell displays and bundled promotions for bread + eggs, bread + milk, and cheese + bread starting in the next quarter.
- ii. Monitor uplift in unit sales of eggs/milk/cheese when advertised alongside bread to quantify promotional effectiveness.
- iii. Recompute association rules periodically (quarterly) and especially after major campaigns, to ensure pairings remain strong.

Overall, this data-driven approach gives us confidence that the proposed cross-selling tactics will resonate with our core customers—because they align directly with historic basket behavior. By focusing on these top-tier rules, we minimize the risk of pushing irrelevant pairings and maximize the chance of incremental sales growth.