Executive Summary: Market-Basket Analysis Using Synthetic Transaction Data

1. How the Synthetic Data Were Generated

To simulate real-world shopping behavior and enable market-basket analysis, we generated synthetic transaction data based on real product listings from two supermarkets. The data generation followed a structured, probability-weighted methodology:

- **Product Source**: Actual supermarket product and pricing data were used.
- **Item Popularity Model**: Product selection probabilities were **inversely proportional to their prices** — a realistic assumption that lower-cost staple items are more commonly bought than premium or niche products.
- **Customer Simulation**:
    - **10,000 synthetic customers** were generated.
    - Each customer's basket contained **1 to 25 products**, with an average of 5.
    - Items in each basket were **randomly drawn**, but **weighted by popularity**.
    - Only the **top 50 most frequent items** were retained to improve analytical focus and reduce memory overload.
- **Data Format**:
    - The final dataset was exported in **wide** format:
        - **Wide Format**: Each row = Customer, each column = Product (1 if purchased)

2. Most Actionable Association Rules

Using the Apriori algorithm, we identified many **statistically significant buying patterns**. The following are the most **actionable and interpretable rules** for business stakeholders:

Antecedent Products	**⇒ Consequent Product**	**Confidence**	**Lift**
Eti Gofret + İçim Kakaolu Süt	⇒ Ülker Badem Kraker	34%	4.50
Ülker Badem Kraker + İçim Süt	⇒ Eti Gofret	29%	4.55
Halley Bisküvi + Kekstra Muffin	⇒ Indomie Noodle	22%	3.16
Eti Gofret + Indomie Gurme	⇒ Ülker Badem Kraker	22%	2.85
Dankek 8 Kek + Yupo Çokojelo	⇒ Dankek Pöti Kek	22%	3.05

These rules reveal **consistent co-purchase behavior**, especially among:

- **On-the-go snacks**
- **Sweet-salty combos**
- **Quick meal pairings**

3. Business Interpretation and Strategic Recommendations

This analysis offers valuable **insights into customer shopping behavior** and suggests several actionable strategies for supermarkets:

Cross-Promotion Opportunities

- **Bundle snack bars, flavored milk, and crackers** into "office snack packs" or "school lunch kits".
- Promote combinations like **noodles + sweet bars** for after-school or late-night meals.

Shelf Placement

- Co-locate high-lift itemsets (e.g., "Eti Gofret", "Ülker Badem Kraker", and "İçim Süt") near checkout lines or convenience aisles to trigger impulse buys.
- Use **"Often Bought Together" signage** in aisles or online stores.

Dynamic Pricing & Combo Offers

- Apply targeted discounts to **frequently paired items**, increasing basket size.
- Implement **loyalty rewards** when customers buy multi-item bundles derived from high-confidence rules.

Flavor/Variety Optimization

- Multiple rules show flavor variation is common (e.g., multiple types of Dankek).
  → Use this insight to promote **multi-flavor bundles** or sampler packs.

Conclusion

Through realistic simulation and rigorous association analysis, this project reveals **tangible opportunities** for product placement, pricing, bundling, and targeted marketing. These insights are not only statistically significant but also directly translatable into **sales-boosting retail strategies**.