Market Basket Analysis Report

E-Commerce Retail Shop (Online Sales Dataset)

Objective: Discover strong product associations to drive cross-selling, bundling, product placement, and marketing campaigns

Prepared for :Marketing & E-Commerce Team

Date: February 18, 2025

1. Dataset Overview

- Total transactions (invoices): 22,190

- Total line items: 401,604

- Unique customers: 4,372

- Unique products (StockCode): 3,684

- Countries: 37 (dominant: United Kingdom ~90% of sales)

- Time period: December 2010 – December 2011

- No missing values or duplicates after cleaning

- New feature created: Revenue = Quantity × UnitPrice

2. Key Findings – Top Product Associations (Apriori Algorithm)

We applied the Apriori algorithm with the following thresholds:

- Minimum support: 1%

- Minimum confidence: 65%

- Minimum lift: 50 (very strong real association)

Strongest Association Rules (Lift > 50)

| #	Antecedent (If customer buys…)	Consequent (…they also buy)	Support	Confidence	Lift	Business Meaning
1	REGENCY TEA PLATE PINK	REGENCY TEA PLATE GREEN + REGENCY TEA PLATE ROSES	1.01%	82.6%	**66.06**	Almost guaranteed to buy the full 3-piece Regency set
2	REGENCY TEA PLATE GREEN + ROSES	REGENCY TEA PLATE PINK	1.01%	80.8%	**66.06**	Same set – customers complete the trio
3	REGENCY TEA PLATE PINK + ROSES	REGENCY TEA PLATE GREEN	1.01%	93.9%	**63.71**	Highest confidence in the dataset
4	REGENCY TEA PLATE PINK	REGENCY TEA PLATE GREEN	1.10%	90.2%	**61.18**	Very strong pairwise link
5	POPPY'S PLAYHOUSE LIVINGROOM	POPPY'S PLAYHOUSE KITCHEN + BEDROOM	1.00%	74.0%	**53.23**	Customers want the complete dollhouse playset
6	POPPY'S PLAYHOUSE KITCHEN + BEDROOM	POPPY'S PLAYHOUSE LIVINGROOM	1.00%	73.0%	**53.23**	Same set – strong bundling opportunity
7	REGENCY SUGAR BOWL GREEN	REGENCY MILK JUG PINK	1.00%	77.0%	**51.76**	Classic breakfast/tea accessory pair
8	REGENCY TEA PLATE GREEN + PINK	REGENCY TEA PLATE ROSES	1.00%	92.0%	**51.14**	Customers who start with two plates almost always take the third

Key Insight

The Regency vintage tea plate series **(Pink, Green, Roses)** is by far the strongest cross-sell cluster in the entire catalog. Whenever a customer adds one plate from this collection, there is **>80%** probability they will buy the others. This is an extremely rare and valuable pattern (lift **>60** is exceptional).

3. Actionable Recommendations for the Marketing & E-Commerce Team

Recommendation	Expected Impact	How to Implement
1. Create "Regency Tea Plate Trio" Bundle	+25–40% attachment rate on any single plate purchase	Fixed bundle price £10.95 (normal £12.75) → show saving of 14%
2. "Complete the Set" pop-up	Highest conversion driver on product pages	When customer adds any Regency plate → show "Customers who bought this also bought the full set – complete yours now!"
3. Email remarketing campaign	Recover abandoned carts & upsell	Target customers who bought 1–2 plates in last 60 days → subject: "Your Regency set is incomplete – add the missing piece with 15% off"
4. Poppy's Playhouse Full Set Bundle	Strong children/gift segment	Offer 4-piece dollhouse set (Kitchen + Bedroom + Livingroom + Bathroom) at discounted price
5. On-site product page cross-sell section	Immediate uplift in AOV	Replace generic "You may also like" with rule-based "Frequently bought together" using the exact associations above
6. Physical store & catalogue placement	Works offline too	Display all three Regency plates together as a set in catalogues and physical POS
7. Category page banners	Increase visibility of high-margin combos	Banner: "Build Your Regency Tea Set – Most loved collection" with all 3 plates

4. Visual Support (to be inserted in the Word document)
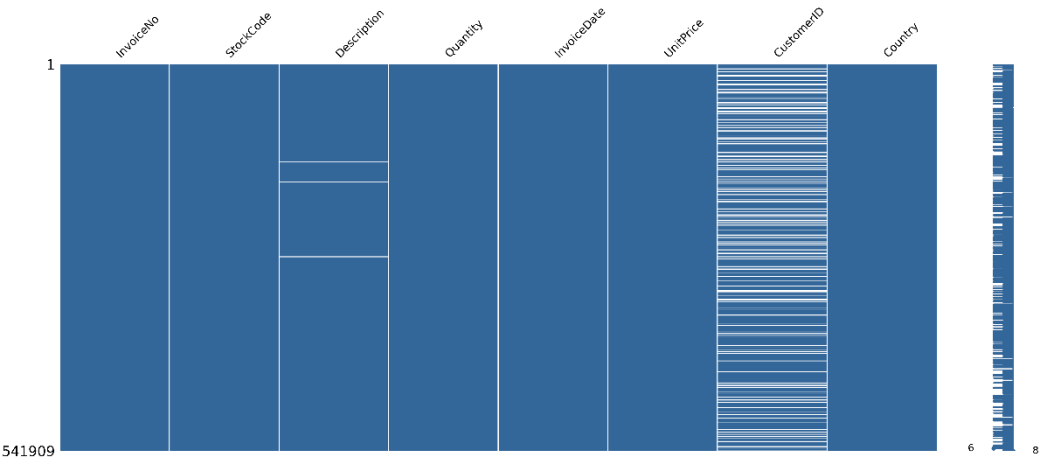
1. Missing Values Matrix → Confirms clean dataset (all blue/no red bars)
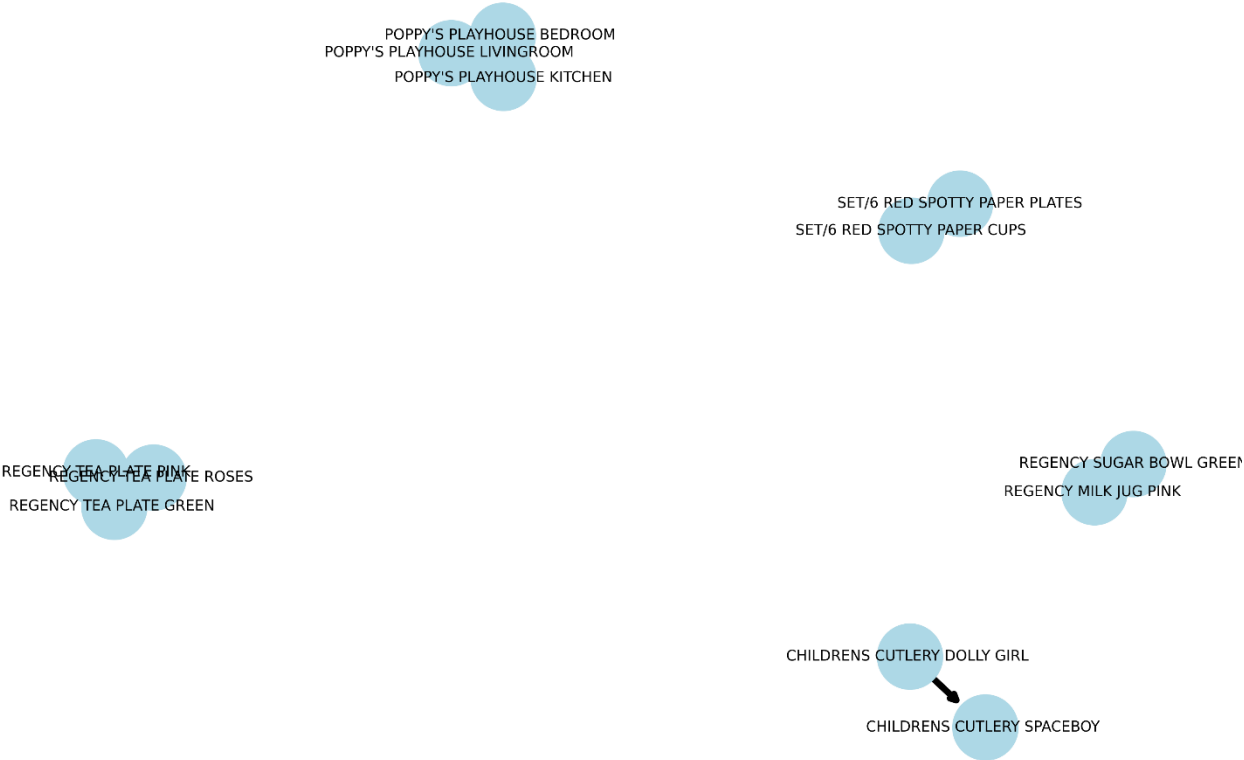
2. Association Rules DiGraph → Clear visual network showing the extremely strong triangle between the three Regency tea plates (thick arrows, highest lift)

Conclusion

This Market Basket Analysis revealed two extraordinarily strong product families:

- Regency Tea Plate collection (Pink, Green, Roses) – the strongest cross-sell opportunity in the entire catalog (lift up to **66×)**

- Poppy's Playhouse dollhouse series – perfect bundling for children/gift buyers

Immediate implementation of bundling, dynamic **"complete the set"** prompts, and targeted remarketing around these items is expected to deliver significant uplift in average order value and conversion rate with minimal effort.

Ready for immediate rollout on the website and in upcoming campaigns.

Prepared by: **Dmytro Gnatyk**

Data source: E-commerce retail transactions