# **Retail Analytics Project — Business Report**

**Overview.** This project applies customer segmentation and market-basket analysis to a retail transaction dataset and makes actionable insights from the data.

**Customer segmentation findings and business impact** - 5 Customer Groups

* **Loyal Frequent Shoppers** — recent activity, shop often, broad variety. They should receive VIP perks and early access rather than discounts.
* **Bulk High-Spend Buyers** — larger baskets and higher spend. They respond to bundles, subscribe-and-save, and replenishment nudges.
* **Steady Long-Term Buyers** — return regularly over months. They fit milestone coupons, free-shipping nudges, and “buy again” widgets.
* **Variety Seekers (Lapsed)** — used to browse widely but haven’t purchased lately. They need win-back offers, samplers, and “new arrivals.”
* **One-Off Low-Spend Buyers** — low frequency and low spend. Keep outreach cheap and stop after two tries if there’s no response.

These groups matter because they point to different promotion levels, channels, and guardrails (for example, keep margins high for loyal shoppers).

## **Association rules & recommendation strategy -** Apriori mined rules on a basket-by-product matrix with tunable support, confidence, and lift. Suggest add-ons with **lift ≥ 1.5** show 1 or 2 suggestions. Email or SMS when timing is important. Do not include low-margin consequences.

**Methodology** - I cleaned the data, filled gaps with medians, standardized features, and dropped IQR outliers. Customers were clustered with K-Means and sanity-checked with 2D/3D PCA, while Apriori mined basket rules with add-on recommendations and quick ROI estimates.

## **Limitations & future work** - Modeling returned items, rare products that don’t sell as much. Another feature to add would be to personalize it based on customer history rather than just segmenting customers into groups.

## **ROI potential & implementation**

ROI approximates incremental attach as *(lift − 1) × base attach*; revenue ≈ *eligible baskets × incremental attach × average price × margin*. With realistic assumptions several rules yield positive ROI after a modest campaign cost; sensitivity remains favorable unless margin drops sharply or costs rise.