Instacart Grocery Basket Analysis Executive Summary

Author: Spencer Davis
Tools: Python (pandas, NumPy), Jupyter, matplotlib/seaborn
Dataset: Instacart grocery orders (2017). *(Some customer attributes may be course-provided/synthetic for learning.)*

Objective

Identify when demand peaks and how customer segments differ to support marketing timing, promotions, and operational planning.

What I Did

- Cleaned and merged order, product, and customer tables; checked data consistency.
- Engineered features (e.g., price bands, customer profiles, region) for segmentation.
- Built aggregations/crosstabs and visualized insights for stakeholder decisions.

Key Findings

- **Demand by day:** Weekends lead. Saturday ~6.21M and Sunday ~5.67M order lines; Tuesday/Wednesday ~3.8M are lowest.
- **Demand by hour:** Strong late-morning to afternoon peak; busiest hour ~10:00 (~2.76M) with a peak window around 10:00–15:00.
- **Pricing mix:** Purchases are mostly mid-range. $5–$15 ≈ 67.4%, then ≤$5 ≈ 31.2%, and >$15 ≈ 1.3%.
- **Segmentation:** Customer profiles vary by region, supporting region-specific targeting (e.g., parents vs. singles vs. seniors).

Recommendations

- **Marketing timing:** Run broad promos around weekend + 10:00–15:00 peaks; use Tue/Wed for retention nudges and experimentation.
- **Promo strategy:** Optimize around mid-range staples; treat premium/high-price items as targeted offers (small share).
- **Segmentation:** Combine profile + region for tailored messaging and bundles.