Grocery Store Sales Performance Analysis

Sydney Holland

October 12, 2025

#### 1. Executive Summary

This report analyzes grocery store sales data to address two key challenges: improving inventory efficiency and assessing the impact of loyalty points on sales.

To address these, over 1,900 transactions records from a publicly available 2025 Kaggle dataset were analyzed using Excel, BigQuery, Tableau, and R Markdown. Data quality issues such as null values and inconsistent formatting were resolved to ensure accurate insights.

#### Key Findings Summary

* **Demand Trends:** Certain products (e.g. apples, cereal) show declining or seasonal demand, while others (e.g. bread, carrots) display unpredictable patterns.
* **Inventory Imbalance:** High-selling products like chicken breast and tomatoes are not consistently fast-moving across stores, indicating uneven performance.
* **Loyalty Points Impact:** Strong correlation found between loyalty incentives and product sales (correlation > 0.69), especially for fresh produce and packaged goods.

#### Recommendations Summary

* **Inventory Optimization:**
  + Increase stock during peak demand months (February - April).
  + Replenish fast-moving products (e.g. cereal, yogurt) frequently across stores.
  + Reduce stocking of slow-moving items (e.g. apples, chicken breast) in underperforming locations. -Expand loyalty point incentives for slower-moving products like ground beef and eggs.

Interactive dashboards support detailed product and store-level insights.

#### 2. Sales Problems

* **Some products frequently go out of stock while others sit on shelves for too long, leading to food waste and lost sales.** 
  + **Objective 1:** Identify demand periods for different products.
  + **Objective 2:** Identify slow-moving vs fast-moving products.
  + **Objective 3:** Create recommendations to optimize product reordering.
* **Are loyalty points affecting sales?**
  + **Objective 1:** Identify which products sell well based on loyalty points.
  + **Objective 2:** Make recommendations to optimize how loyalty points are assigned to maximize sales.

#### 3. Data Overview

* **Dataset Source**: Kaggle Public Data
* **Dataset Name**: Grocery Store Sales Dataset in 2025 - 1900+ Records by Pratyush Puri
* **Data Sources**:
  + **Customer Data**: Customer ID (personal data omitted)
  + **Store Data**: Store Name
  + **Transactional Data**: Transaction Date, Aisle, Product Name, Quantity, Unit Price, Total Amount, Final Amount
  + **Promotional Data**: Discount Amount, Final Amount
* **Dataset Size: 1,980 customers, 10 features**

#### Data Quality Checks:

* **Process**:
  + Performed null value filtering.
  + Standardized quantity formatting.
  + Validated final\_amount and total\_amount calculations using SQL.
* **Issues Found** :
  + Null store\_name values.
  + Mixed formatting in quantity field (“2” vs “2.0”).
* **Action Taken** :
  + Excel used to filter and remove null values from a copy table.
  + Excel used to format quantity column to two decimal places for standardization.
* **Rationale** :
  + Regional analysis is needed for the goal of the project. Null values add no analytical value.
  + Inconsistent data can cause issues in analysis.

#### 4. Analysis:

* **Tools Used**: Microsoft Excel, BigQuery, Tableau, R Markdown
* **Assumptions**
* Perishable products like fresh produce and dairy products should be monitored more closely for overstocking compared to long-shelf-life products like rice and cereal.
* **Loyalty Points vs Demand**:
  + The CORR function was used to calculate the correlation coefficient of the total points issued for a product from each store and the total units sold and total revenue obtain for that product from each store.
    - **My point of reference for a strong correlation is above 0.5.**
    - My analysis obtained a correlation coefficient of 0.766 for units sold and 0.692 for revenue, indicating a strong relationship between the amount of loyalty points earned on each product and the amount of units sold and revenue generated.
      * correlation between units sold and loyalty points: 0.766 > 0.5 (Strong).
      * correlation between revenue and loyalty points: 0.692 > 0.5 (Strong).
  + I used PERCENTILE\_CONT to find the high and low thresholds and median point of the loyalty\_rank column in order to determine my threshold for the loyalty program recommendations.
    - **High threshold** => 254.85
    - **Median Point** => 217.30
    - **Low Threshold** => 179.47

#### 5. Observations

* **Demand Periods By Month**
  + Apples and cereal show significantly lower demand in recent months compared to prior year peaks.
  + Bread and carrots demonstrate highly unpredictable demand patterns. There is no consistent peak month from 2023 - 2025.
  + Potatoes and bananas decline sharply from spring to summer.
  + Rice consistently performs well in April - May.
  + Yogurt has the lowest demand during fall months.
  + Items like pasta, apples, and bread see the most demand early in the year, most notably from February to April.
  + **You can view the interactive Tableau dashboard visualizing Product Units Sold By Month and Total Revenue Per Product By Month** [**here**](https://public.tableau.com/views/GrocerySalesOverTime/Dashboard1?:language=en-US&:sid=&:redirect=auth&:display_count=n&:origin=viz_share_link)**.**
* **Inventory Management:**
  + Chicken breast, tomatoes, potatoes, bread, and onions lead in total units sold across all stores which indicates a strong base demand.
  + Chicken breast and tomatoes have high total sales but have a low fast-moving ratio indicating uneven performance across stores.
  + Products like cereal and yogurt are consistently fast-moving across stores which indicates strong sales.
  + Onions and eggs have good total sales but have low fast-moving ratios.
* **Loyalty Points:**
* Fresh produce and packaged goods showed strong performance in response to point incentives. Among the products that experienced the highest sales and highest loyalty rewards were onions, tomatoes, chicken breast, cereal, and yogurt.
* Ground beef, bread, eggs, carrots, and bananas show low to moderate performance across stores.
* **You can view the interactive Tableau dashboard visualizing Grocery Sales Here** [**here**](https://public.tableau.com/views/GrocerySalesAnalysis_17602981651400/TotalSalesVolumebyProductandStoreName?:language=en-US&:sid=&:redirect=auth&:display_count=n&:origin=viz_share_link)**.**

#### 6. Recommendations

* **Inventory Management:**
  + **Overall Recommendations**
    - Because of the surge in sales early in the year, stores should generally prepare a higher stock from February through April.
    - Products with high sales rank and fast-moving ratios should maintain their safety stock and replenish frequently.
    - Products that are slow-moving should replenish frequently only in top-performing stores.
    - Medium-moving products should be monitored and adjust their replenishment quantity and frequency according to future sales trends. For now, these products should maintain their stocking trends.
  + **Product-Specific Recommendations**
    - Restocking quantities for apples and cereal should be reduced until future analysis indicates a demand rebound.
    - Bread and carrots should keep a slightly higher safety stock to adjust to highly volatile demands.
    - Maintain current reordering for products like potatoes, bread, and onions.
    - Reduce stock for chicken breast and tomatoes in low performing stores and increase in high-performing stores.
    - Replenish fast-moving products like cereal and yogurt frequently across stores.
    - Reduce stocking frequency for items like onions and eggs across stores.
    - Replenish stock of rice more frequently around late spring but lower or maintain safety stock during the off-seasons due to rice’s long shelf life.
* **Store-Specific Recommendations**
* *Increase Stock:*
  + Corner Grocery should restock products like bananas and cheese more frequently.
  + Green Grocer Plaza should restock products like potatoes more frequently.
  + City Fresh Store should restock cereal, yogurt, and milk more frequently and tomatoes less frequently.
  + SuperSave Central should restock tomatoes more frequently.
  + MegaMart Westside should restock apples more frequently.
* *Reduce Stock:*
  + FamilyFood Express should reduce stock in bananas and milk.
  + QuickStop Market should reduce stock in ground beef.
  + FreshMart Downtown should reduce stock in salmon and carrots.
  + City Fresh Store should reduce stock in tomatoes.
* **Loyalty Progran Recommendations**
  + Fresh produce and packaged goods should maintain their current loyalty point incentives.
  + Increase the loyalty point incentive for ground beef, eggs, bread, carrots, and bananas to build engagement.