Analyzing Grocery Pricing Trends of Bacel Products in Toronto*

Insights from Vendors: Voila, Metro, and No Frills

Julia Lee   Vandan Patel   Haowei Fan   Tianning He

Max LI   Shuangyuan Yang

November 14, 2024

This paper analyzes grocery product pricing data from three Toronto vendors (Voila, Metro, and No Frills) using data collected through screen-scraping from Project Hammer. Prices were standardized to 100g units to ensure comparability across vendors. The findings show that there is no correlation between the vendors and the Becel products that are sold. However, a slight pattern is seen that most customers buy products that are in between the range $1.1 and $1.3 per. 100g.

1 Introduction

In this analysis, we explore grocery product pricing data from three major Toronto retailers: Voila, Metro, and NoFrills. Using a dataset collected through screen-scraping from Project Hammer, we aim to understand differences in product pricing and distribution patterns across these vendors. The data includes pricing information for various products, which we standardized to 100g units to ensure comparability between stores. By performing data cleaning and visualization, this analysis seeks to compare price distributions across vendors and identify any notable trends.

This present analysis first segments the data by vendor, creating subsets for Voila, Metro, and No Frills. For each subset, we visualize the distribution of product prices at each store. These visualizations provide insights into the pricing strategies of each vendor, highlighting the frequency of different price ranges and enabling a side-by-side comparison of pricing trends across these retailers.

Additionally, the analysis examines whether any relationships exist between the price distributions of the three vendors. This enables a direct comparison of prices for similar products

---

*Code and data are available at: https://github.com/JuliaJLee/grocery_vendor_analysis

across vendors, facilitating a clearer understanding of how pricing might vary between Voila, Metro, and No Frills. Although no correlation was found specifically for Becel products across vendors, a slight pattern emerged, showing that most customers tend to purchase products priced between \$1.10 and \$1.30 per 100g.

The remainder of this report details the data used in the analysis (Section 2), the results of the visualizations (Section 3), and the conclusions drawn from these comparisons (Section 4), highlighting pricing patterns and potential pricing disparities among these vendors.

2 Data

The preparation and manipulation of the original data is conducted using SQL (Chamberlin and Boyce 1970). The steps to conduct the present analysis on the prepared data are carried out using the statistical programming language, R (R Core Team 2023).

The data for this analysis was obtained provided by Project Hammer (Filipp 2024), and it contains data about product details (e.g. product names, brands, current prices, and old prices, and units for products) for 8 different grocery vendors (Voila, T&T, Loblaws, No Frills, Metro, Galleria, Walmart Canada, and Save-On-Foods). The data is collected through the process of screen-scraping website UIs from vendors within a neighbourhood in Toronto. Data from Project Hammer is available from February 28, 2024, and as this data is collected by extracting information from grocers' websites, it is possible to have missing data, which is more thoroughly addressed in Section 4. Through the method of screen scraping, the original data aims to measure and reflect grocery vendor decisions and trends in product pricing.

For this analysis, three vendors - Voila, Metro, and No Frills - are considered along with their respective pricing for products that are 100g. The "Price" variable for each vendor in this analysis reflects the current price of their products at the time in which the data was downloaded. The variable, "Units", represents the size of products measured using weight metrics (i.e. grams or kilograms) and counts (i.e. the number of items within a given product). In this case, a weight of 100g is selected to ensure consistency in units across products and vendors. It is also important to highlight that this analysis only considers products from the brand, Becel, across the three vendors.
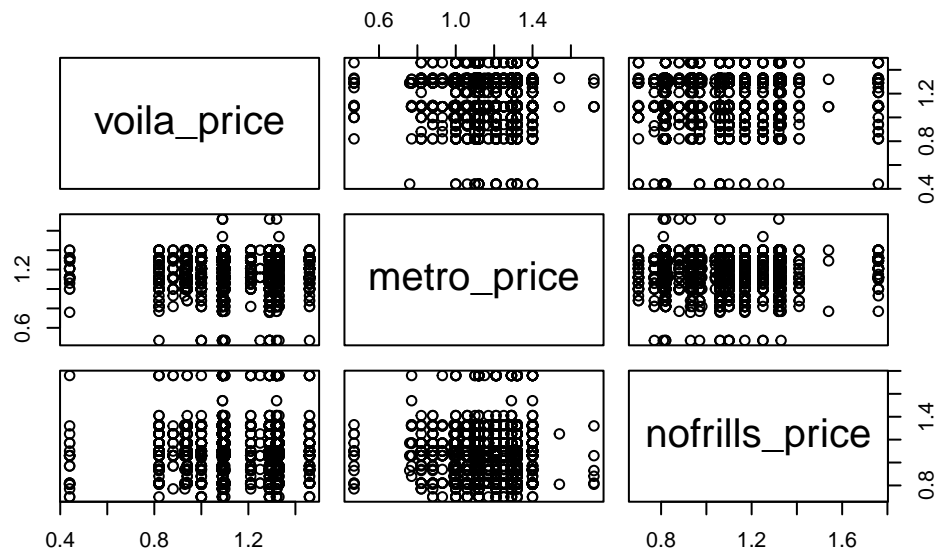
3 Results

Figure 1: The Relationships between Becel Product Prices from fVoila, Metro, and No Frills

Figure 1 shows a pairwise scatter plot matrix of the price distribution of Voila, Metro, and No Frills. It can be seen from the figure that the prices of Voila and Metro are mainly concentrated between 1.0 and 1.3, while the price of No Frills is bimodal, with peaks at 0.9 and 1.2 respectively. Overall, there is a certain overlap in the pricing strategies of the three brands (in the range of 1.0 to 1.3). The prices of Voila and Metro are more concentrated, which may indicate that their pricing is more uniform and their target markets may be similar. However, No Frills has some differences from them. The bimodal distribution of No Frills' prices reflects that it has formulated diversified pricing strategies for different consumer groups, which may attract low-price consumers and cover part of the mid-to-high-price market. This difference indicates that the three may have different competitive strategies in terms of market positioning and target consumers.
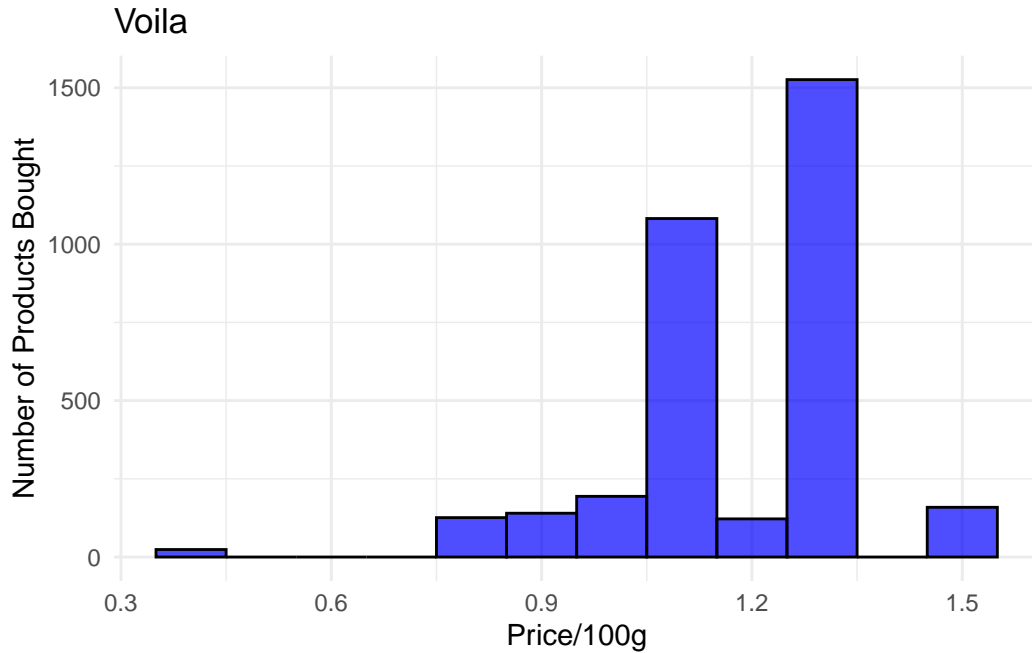
Figure 2: Distribution of the Number of Becel Products Sold Across Different Prices from Voila

Figure 2 illustrates the distribution of Voila's price per 100g. The horizontal axis represents the price per 100g, while the vertical axis denotes the frequency of occurrence. The distribution is right-skewed, with most values concentrated between approximately 1.0 and 1.2 price units per 100g. The highest frequency, exceeding 1,500 occurrences, is observed around the 1.2 price mark. There are fewer products priced below 1.0, and very few exceed 1.5. This indicates that Voila's prices are predominantly clustered around 1.2 per 100g, with some outliers on the lower and higher ends.
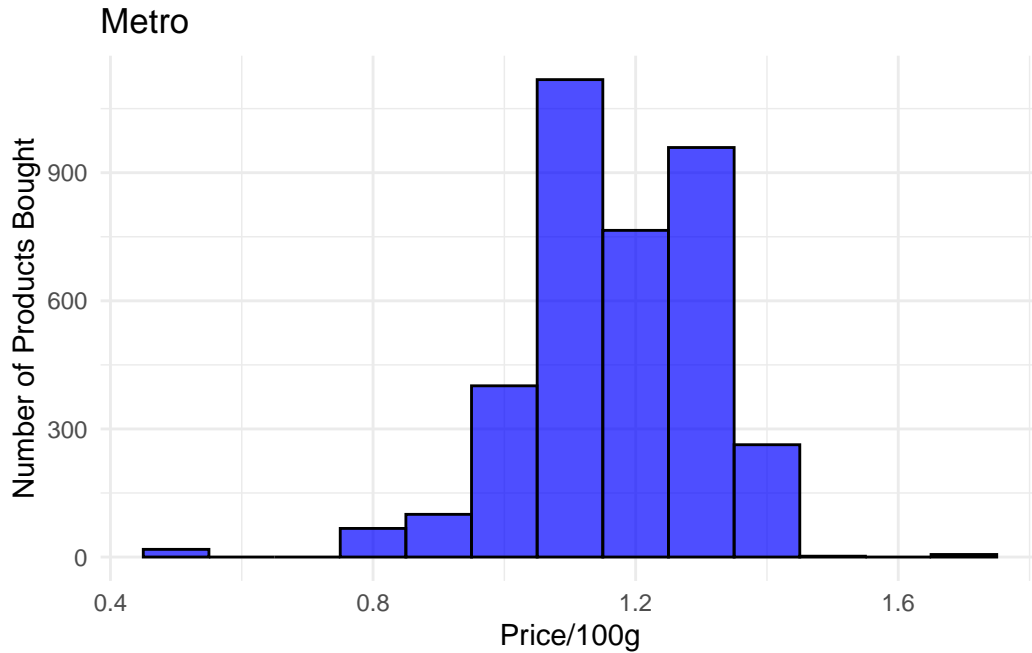
Figure 3: Distribution of the Number of Becel Products Sold Across Different Prices from Metro

Figure 3 represents the distribution of Metro's price per 100g. The horizontal axis shows the price per 100g, while the vertical axis indicates the frequency of occurrences. The distribution is slightly right-skewed, with the majority of prices clustered between approximately 1.0 and 1.3 price units per 100g. The highest frequency, around 900 occurrences, is observed near the 1.2 price mark. A smaller number of items are priced below 1.0, and there are few outliers above 1.5. This suggests that Metro's prices predominantly fall within the 1.0–1.3 range, with a peak at 1.2.
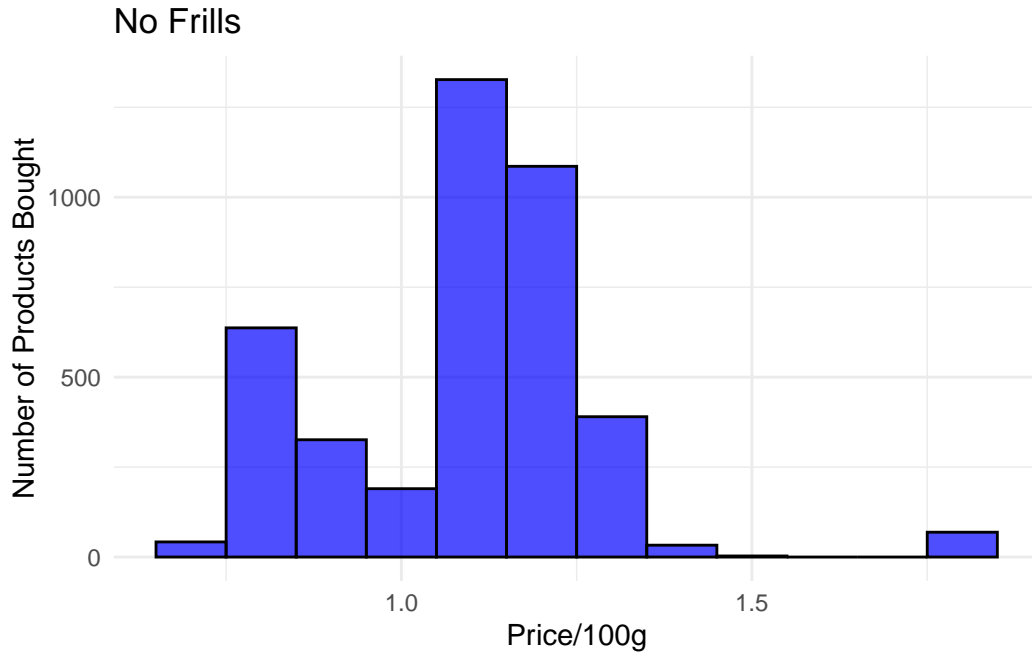
Figure 4: Distribution of the Number of Becel Products Sold Across Different Prices from No Frills

Figure 4 illustrates the distribution of No Frills' price per 100g. The horizontal axis represents the price per 100g, and the vertical axis indicates the frequency of occurrences. The distribution is bimodal, with significant peaks around 0.9 and 1.2 price units per 100g. The highest frequency, exceeding 1,000 occurrences, is observed at the 1.2 price mark. A smaller but notable peak is visible around 0.9, with fewer items priced below this range or above 1.5. This suggests that No Frills' prices are primarily concentrated at two levels, with the higher peak being more dominant.

4 Discussion

References

Chamberlin, Donald, and Raymond Boyce. 1970. *Structured Query Language (SQL)*. IBM.

Filipp, Jacob. 2024. "Project Hammer." https://jacobfilipp.com/hammer/.

R Core Team. 2023. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. https://www.R-project.org/.