

# Analyzing Grocery Pricing Trends of Bancel 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 Bancel 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.

## Table of contents

<b>1</b>	<b>Introduction</b>	<b>1</b>
<b>2</b>	<b>Results</b>	<b>3</b>
	<b>Discussion</b>	<b>5</b>

## 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

---

\*Code and data are available at: [https://github.com/JuliaJLee/grocery\\_vendor\\_analysis](https://github.com/JuliaJLee/grocery_vendor_analysis)

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 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 ([?@sec-data](#)), the results of the visualizations (Section 2), and the conclusions drawn from these comparisons (Section 2), highlighting pricing patterns and potential pricing disparities among these vendors.

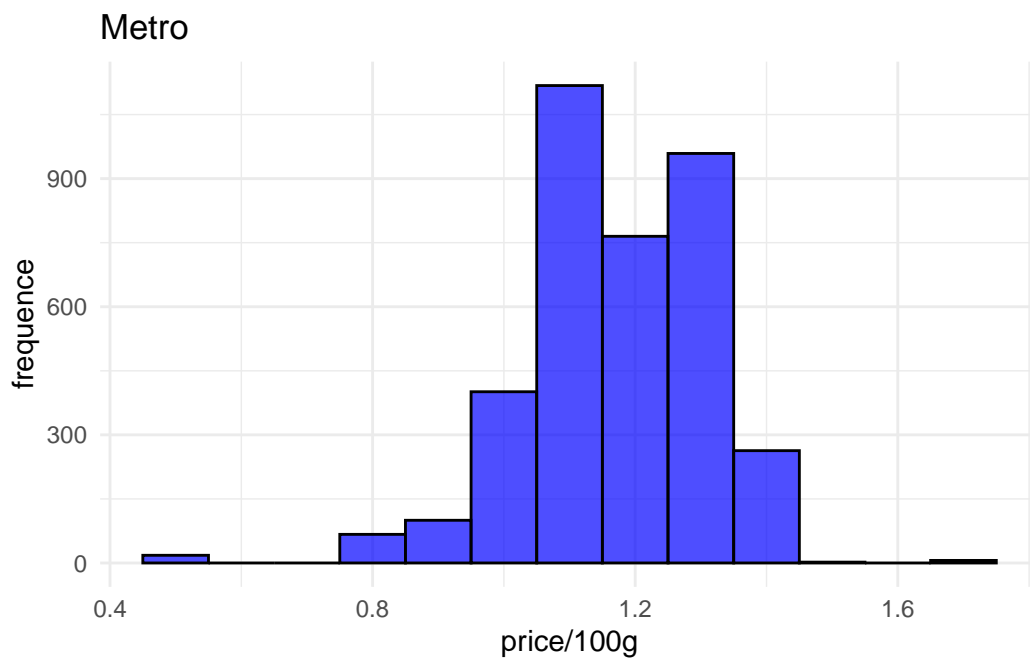
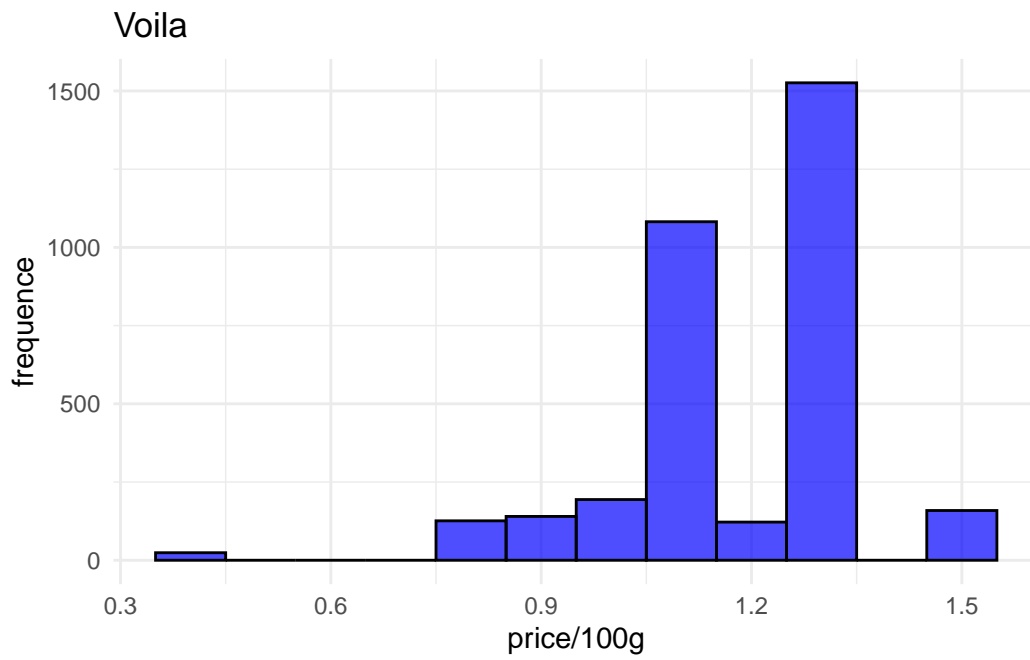
#Data {#sec-data}

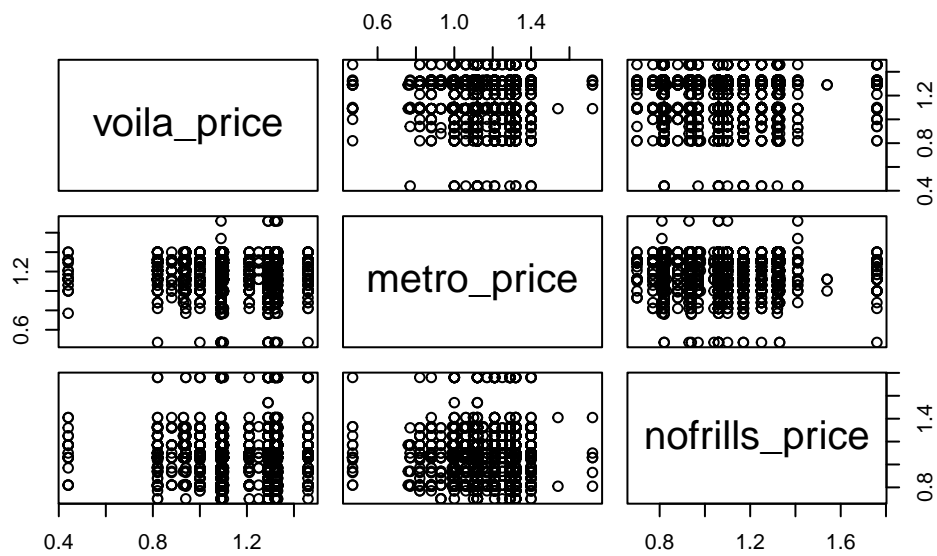
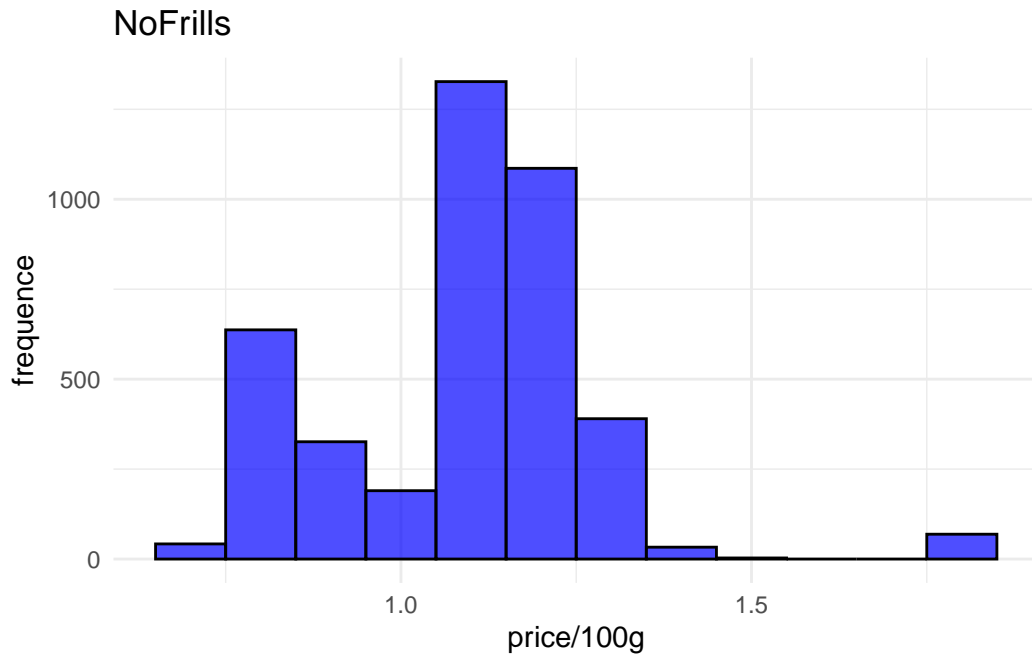
The preparation and manipulation of the original data is conducted using SQL (Chamberlin and Boyce 1970). The steps to conduct the present analysis on the prepared data are carried out using the statistical programming language, R (R Core Team 2023).

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 2. 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.

## 2 Results





## Discussion

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/>.