

# **Revealing Price Patterns: A Analysis of Price Dynamics and Vendor Strategies in the Canadian Grocery Sector in 2024\***

Jin Zhang      Yi Tang      Siyuan Lu

November 14, 2024

In this paper, we collect historical grocery data from top Canadian grocery sector, clean it, and analyze the pricing strategies. Many items have remained the same price or become cheaper, showing that the grocery vendors tend to use competitive pricing strategies. The research can help drive more competition and reduce collusion in the Canadian grocery sector.

## **Table of contents**

<b>1</b>	<b>Introduction</b>	<b>2</b>
<b>2</b>	<b>Data</b>	<b>2</b>
2.1	Overview . . . . .	2
2.2	Measurement . . . . .	3
2.3	Graphs . . . . .	4
<b>3</b>	<b>Results</b>	<b>6</b>
<b>4</b>	<b>Discussion</b>	<b>7</b>
	<b>References</b>	<b>8</b>

---

\*Code and data are available at: <https://github.com/YiTang2/Canadian-grocery-price.git>

# 1 Introduction

The Canadian grocery market is highly competitive, with a variety of pricing strategies emerging, demonstrating the efforts of merchants to compete and avoid collusion. This study aims to shed light on these pricing strategies by analyzing historical data from major Canadian grocery chains, focusing on how prices change over time and how different merchants position themselves in the market. Through this study, we hope to provide insights into the competitive dynamics of the retail market, thereby promoting more intense competition and reducing collusive behavior within the industry.

This paper explores the current and historical price relationships of different merchants and finds that prices for many products remain the same or even decrease. These findings suggest that Canadian grocers often adopt competitive pricing strategies, either in response to market demand dynamics or as a means to gain a larger market share. These pricing behaviors are critical to understanding the competitive environment and providing policymakers with directions for fair competition.

The significance of this study lies in its potential impact on economic research and practical policy making. By analyzing pricing strategies and their effects in detail, this study provides valuable insights to enhance competitive practices and reduce anti-competitive behavior.

The remainder of this paper is structured as follows. Section 2 discusses the overview and variables of the data, including graphs and analysis. Section 3 explains the findings of the paper. Section 4 consists references of all data, software, and material. Finally, the Discussion section explores the implications of our findings, the study's limitations, and potential avenues for future research.

## 2 Data

### 2.1 Overview

We got the Canadian Grocery Price Data from Jacob Filipp(Filipp 2024). We used R programming language (R Core Team 2023) to analyze the data and plot the graphs for this paper. We also used SQL (9075:2016(E) 2016) to clean the raw data and generated the analysis dataset. The folder structure for this paper follows the startup folder created by (Wickham et al. 2019a). It is very helpful in keeping everything organized. I also referenced many of the examples and techniques provided by Telling Stories with Data (Alexander 2023), which helped me understand how to visualize the data and communicate the findings effectively. In addition, the `tidyverse` (Wickham et al. 2019b) packages is essential for simplifying data cleaning and analysis.`styler` (Müller and Walthert 2024) used in arranging code style. These tools help to organize and present data efficiently. I also plotted the graphs with `ggplot2`

(Wickham 2016) and read the csv file with `here` (Müller 2020). We also use the `comma` function from the `scales`(Wickham, Pedersen, and Seidel 2023) library to format the data, thereby avoiding scientific notation in our charts. We used the `knitr`(Xie 2014) to show the table. And we used `styler` (Müller and Walthert 2024) to do format arrangement in script.

All datas aim to foster competition and reduce market collusion in the Canadian grocery sector by assembling a database of historical grocery prices from prominent vendors. This dataset serves academic and legal purposes, enhancing transparency in grocery pricing strategies. The data encompasses prices from eight major grocery vendors in Canada, including Voila, TandT, Loblaws, and others. Below is the description for the variable we are taking to research:

- vendor: One of the 7 grocery vendors.
- current\_price: Price at time of extract.
- old\_price: An “old” struck-out price. This indicates that there was a Sale on. This is how you can differentiate a price-drop advertised as a Sale vs. a “quiet” decrease without a Sale.
- units: Units (grams, kg, number of items in package). May be blank for some vendors/products.

The dataset allows us to analyze price trends over time for items like Becel Margarine and Sourdough Bread through visualizations. Each chart helps compare how different vendors price these products at various times. The initial data exploration reveals pricing patterns, with summary stats showing average prices, standard deviations, and price ranges. Comparing vendors or time periods can highlight specific trends or unusual pricing behaviors.

## 2.2 Measurement

The measurement process turns grocery pricing data into a structured dataset. This starts by collecting price data from various Canadian grocery sites using screen-scraping. Each price entry captures the “current price” at the time and the “old price,” showing the previous cost before any discounts. Prices are gathered for “in-store pickup” in Toronto through vendor website screen-scraping, focusing on customer-facing prices but possibly missing details in internal APIs. This approach allows us to track price changes over time and study market behaviors.

To keep the data accurate and relevant, we focus on items for in-store pickup in a specific Toronto area, ensuring a consistent location context. The “units” for each product, like grams or kilograms, are recorded to standardize price comparisons. This careful data process helps us deeply analyze pricing strategies across vendors, uncovering trends and any potential collusion in the market.

## 2.3 Graphs

Some of our data is of current price (Figure 1), from JacobFilipp Filipp (2024).

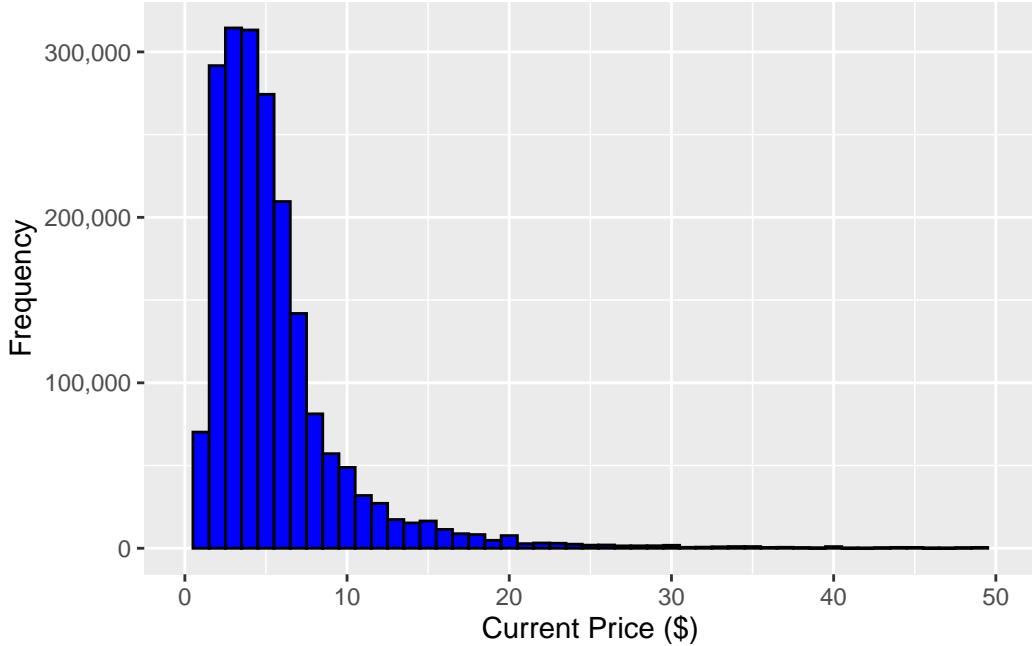


Figure 1: Distribution of Current Prices in Canadian Grocery Sector

Figure 1 shows the distribution of current prices in the Canadian grocery sector. The x-axis of this histogram represents the range of current prices in Canadian dollars, while the y-axis indicates the frequency of each price point. This visualization allows us to see that most grocery items are priced under \$10, showcasing a significant decline in frequency as prices increase. This distribution shows a concentration of more affordable items, with fewer high-priced items, indicating a price-sensitive market where cheaper products dominate the sales volume.

Some of our data is of price (Figure 2), from JacobFilipp Filipp (2024).

Figure 2 shows the logarithmic distribution of prices by different vendors in the Canadian grocery sector, using a box plot for each vendor. The x-axis categorizes the vendors, and the y-axis shows the logarithm of current prices, which shows the wide range of prices and highlighting the differences in price distributions across vendors. Each box plot provides the median, interquartile range (IQR), and potential outliers, revealing variability in pricing strategies among vendors. For instance, some vendors show a wider range of prices, indicating a diverse product mix, while others have more compact distributions, suggesting more consistent pricing across their products.

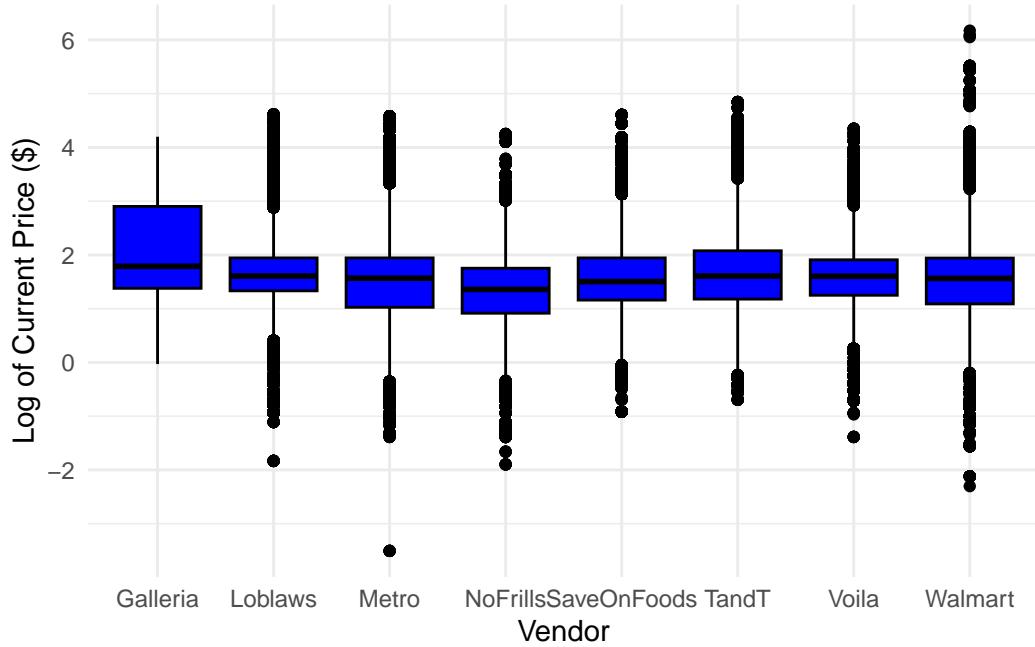


Figure 2: Distribution of Prices by Vendor in Canadian Grocery Sector

Some of our data is of comparison of price (Figure 3), from JacobFilipp Filipp (2024).

Figure 3 visualizes the relationship between old and current prices of grocery items in a scatter plot. The x-axis represents the old prices, and the y-axis the current prices, both in Canadian dollars. The diagonal red dashed line indicates where the old and current prices are equal. Points below this line represent price drops, while points above indicate price increases. This graph reveals a significant cluster of points along and below the line, suggesting that many items have remained the same price or become cheaper. This could indicate competitive pricing tactics or responses to market demand dynamics.



Figure 3: Comparison of Old and Current Prices in Canadian Grocery Sector

### 3 Results

Table 1: Summary Statistics for Current Prices in the Canadian Grocery Sector

	Mean	Median	SD	Min	Max
	6.755484	4.79	9.957137	0	479.99

Table 1 for current prices in the Canadian grocery sector show a varied and competitive market. The mean and median prices give an idea of typical costs, suggesting a focus on affordability. The standard deviation shows a wide range of prices, from low to high, meaning there are options for different budgets. The minimum and maximum prices capture the full range, from budget-friendly to premium choices, highlighting the market's ability to cater to diverse needs. Together, these stats show an active market where vendors use different pricing tactics to appeal to a wide consumer base.

Table 2: Summary Statistics for vendor Prices in the Canadian Grocery Sector

vendor	Mean	Median	SD	Min	Max
Galleria	2.238724	1.943049	0.8883970	0.6780335	4.214495

Table 2: Summary Statistics for vendor Prices in the Canadian Grocery Sector

vendor	Mean	Median	SD	Min	Max
Loblaws	1.842668	1.790091	0.5191117	0.1484200	4.626442
Metro	1.820652	1.756132	0.7173577	0.0295588	4.589955
NoFrills	1.608753	1.587192	0.4952125	0.1397619	4.263102
SaveOnFoods	1.767844	1.704748	0.5158906	0.3364722	4.615022
TandT	1.928242	1.790091	0.7188706	0.4054651	4.851952
Voila	1.770481	1.786747	0.4636231	0.0000000	4.360675
Walmart	1.779771	1.752672	0.5741976	0.0953102	6.175846

Table 2 for log-transformed current prices by vendor highlight the different pricing approaches in the Canadian grocery market. The mean and median on the logarithmic scale show where each vendor positions itself in terms of pricing. The standard deviation reveals the variety in each vendor’s pricing, from budget to premium products. The minimum and maximum values reflect the lowest and highest prices offered, showcasing the full range of options available. These stats emphasize the competitive nature of the market, showing how vendors use pricing to stand out and attract various consumer groups.

## 4 Discussion

This paper examines pricing behaviors in the Canadian grocery market by comparing current and past prices across different vendors. Using visual and statistical tools, it identifies patterns and strategies in how retailers set their prices. The findings show a highly competitive market, with vendors aggressively pricing products to attract consumers. The focus on lower-priced items suggests a strategic response to consumer demand for affordability in everyday purchases.

Differences in pricing strategies among vendors highlight varying approaches to market segmentation. Some vendors offer a wide range of prices to appeal to diverse consumer groups, while others focus on narrower price ranges, targeting specific market segments.

A key limitation of this study is the reliance on screen scraping for data collection, which might miss some pricing details or short-term promotions, potentially affecting the accuracy of the analysis.

Future studies could improve the findings by using direct API access for real-time pricing data and collecting data from a wider geographic area. Further research could also examine how external economic factors influence pricing strategies, offering deeper insights into market dynamics and consumer behavior.

## References

- 9075:2016(E), ISO/IEC. 2016. “Structured Query Language.” International Standard.
- Alexander, Rohan. 2023. *Telling Stories with Data: With Applications in r*. Chapman; Hall/CRC.
- Filipp, Jacob. 2024. “Hammer Dataset.” <https://jacobfilipp.com/hammer/>.
- Müller, Kirill. 2020. *Here: A Simpler Way to Find Your Files*. <https://CRAN.R-project.org/package=here>.
- Müller, Kirill, and Lorenz Walthert. 2024. *Styler: Non-Invasive Pretty Printing of r Code*. <https://CRAN.R-project.org/package=styler>.
- R Core Team. 2023. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. <https://www.R-project.org/>.
- Wickham, Hadley. 2016. *Ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York. <https://ggplot2.tidyverse.org>.
- Wickham, Hadley, Mara Averick, Jennifer Bryan, Winston Chang, Lucy D’Agostino McGowan, Romain François, Garrett Grolemund, et al. 2019a. “Welcome to the tidyverse.” *Journal of Open Source Software* 4 (43): 1686. <https://doi.org/10.21105/joss.01686>.
- , et al. 2019b. “Welcome to the tidyverse.” *Journal of Open Source Software* 4 (43): 1686. <https://doi.org/10.21105/joss.01686>.
- Wickham, Hadley, Thomas Lin Pedersen, and Dana Seidel. 2023. *Scales: Scale Functions for Visualization*. <https://CRAN.R-project.org/package=scales>.
- Xie, Yihui. 2014. “Knitr: A Comprehensive Tool for Reproducible Research in R.” In *Implementing Reproducible Computational Research*, edited by Victoria Stodden, Friedrich Leisch, and Roger D. Peng. Chapman; Hall/CRC.