Pricing Approaches for Green Tea Products*

A Study of Canadian Grocery Retailers

David Qi Haobo Ren Xinrui Xie Ruikang Wang
Sakura Hu

November 21, 2024

This study uses canadian grocery Price data to analyzes the pricing strategies of green tea-related products across grocery stores in Canada. We finds T&T offers a mix of premium and affordable teas, targeting its Asian customer base, while Walmart maintains low prices through economies of scale. Loblaws uses dynamic pricing and promotions to attract budget-conscious consumers. These strategies reflect different approaches to catering to diverse consumer segments.

Table of contents

*Code and data are available at: https://github.com/UTDQi/grocery_analysis

1 Introduction

Many people think that only Asians are obsessed with tea, but after the UK passed tea around the world, North America are beginning to embrace the culture of tea. This paper will examine the price of green tea and green tea related products among various suppliers, explore the reasons for the price changes, and study the pattern of price changes among various supermarkets through the vehicle of green tea.

By examining the distribution of prices for green tea related products in Canada, we were able to identify differences in the pricing and stocking strategies implemented by various grocery stores. Our findings reveal that T&T offers a wider range of higher-end teas, along with a selection of lower priced options. This aligns with our initial assumption that T&T targets a significant portion of Asian consumers, who tend to favor high-quality teas. We also observed Walmart's consistent low pricing, which reflects its strategy of offering affordable options across the board. Additionally, Loblaws employs pricing tactics such as frequent price adjustments and promotional strategies, which can be seen as attempts to attract cost-conscious consumers through price tricks.

The graphing scripts in this paper are written in R (R Core Team 2023), with the tidyverse package (Wickham et al. 2019) The graphs are generated with ggplot2 (Wickham 2016). Data cleaning was done in SQL (International Organization for Standardization 2023).

The rest of the paper is structured as follows: in Data section we will discuss the various aspects of the data we use, how they were gathered and the variable we shall use. In the Results section, we present the findings we observe from the data. In the Discussion section we will discuss the interpretation to the findings, their possible causation and limitation to this study.

2 Data

2.1 Mesurement

The data used in this paper were obtained from the websites of various food retail companies, including Voila, T&T, Loblaws, No Frills, Metro, Galleria, Walmart Canada, Save-On-Foods. The data was gathered by using screen-scrape of website UI, starting from Feb.28 to Nov.14. This data captures the product name, units, and prices of each vendors (Filipp 2024). This

data set is designed to be sited for suitable for academic analysis on market behaviors and for legal actions.

This data has some limitations. There are missing data on certain days for certain vendors,some missing points occur on days which those specific extracts failed. While data collection for some products starts form July. This data didn't include all vendors in the markets, which may cause bias to our analysis.And as the author of the data set states, "Prices are for the"in store pickup" option, as set for a neighbourhood in Toronto.", which limits the effective range of our data to one city, or even one neighbourhood.

2.2 Key Observation Variables

2.2.1 Current Time

The current time(`current_price`) is the time when this data is retrieved.

2.2.2 Current Price

The current price(`current_price`) is the price of the product retrieved at the time accessed.

2.2.3 Old Price

The current price(`old_price`) is the "old price" of the product retrieved at the time accessed that is stuck out by the vendor. It appears whenever there is a sales on and the vendor wants to show the change in price. This is the key variable to differentiate between sale and silent price drops.

2.3 Overview

Figure 1 is a graph showing the overall distribution for price of green tea related products for each vendor. The term average price refers to the mean price of a product observed over the time where data is available. Most vendors exhibit a right-skewed distribution in green tea related product prices, with the majority priced on the lower end of the spectrum. This pattern suggests that, while there are some high-priced teas, they are relatively uncommon compared to more affordable options.

Figure 2 shows the change in price of Lipton Ginseng Lemon Green Tea 72 Counts in three different retailers. It serves as an example for us to understand the change in price over the data set. Some of the dates are missing for reasons stated above, the data collection of Walmart in this case started on July 25^{th}. The figure suggests that for even for the same product, there may exist huge difference in pricing schemes. Walmart (blue) consistently offers the product

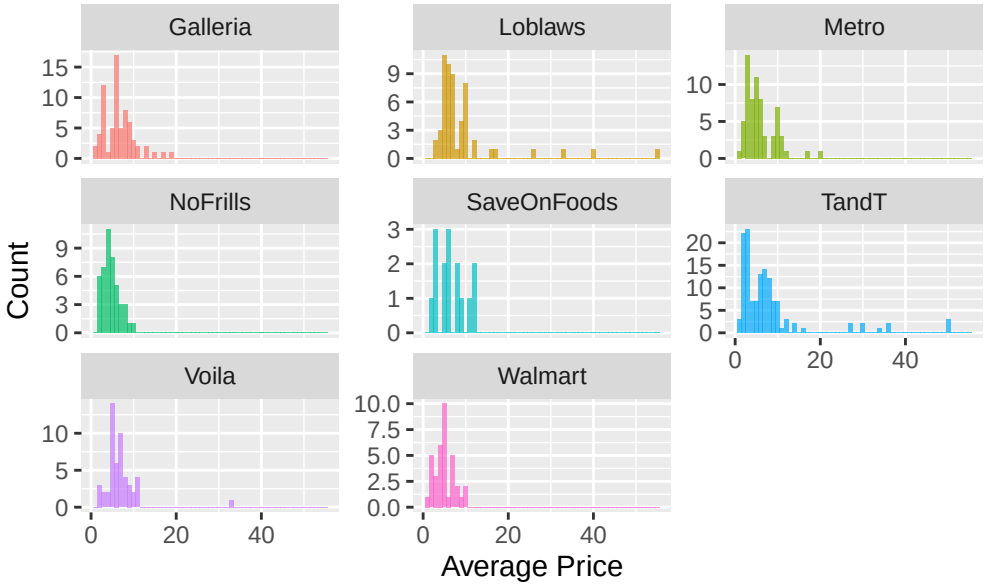

Figure 1: Distribution of Average Prices by Vendor

at a lower price point (around $7). Metro (green) has price fluctuations primarily between $8 and $9, showing some variation but generally maintaining a mid-tier price. Loblaws (red) displays the highest price range, between $9 and $11, with more noticeable rises over time. It is noticeable that we observe a case where Loblaws rises the price before beginning a sale on July 25^{th}, this action that aims to makes the sale seem more impressive is often considered unethical.

3 Results

3.1 Stock Distributions

With Figure 1, we may examine the distribution of different priced products each vantor provides and draw out some trends. Galeria and Voila have a concentration of lower-priced teas, mostly under $10, with very few teas above this price point. Loblaws and Metro also follow a similar pattern, though Metro shows a slightly wider spread, with a few teas priced above $20. TandT and Loblaws have a larger variety of prices, with TandT showing a wider distribution of green tea related product prices up to and beyond $40, indicating a broader range of products that likely include premium or specialty teas. NoFrills and Walmart has the narrowest range in green tea related product prices, concentrated under $10, indicating a limited selection

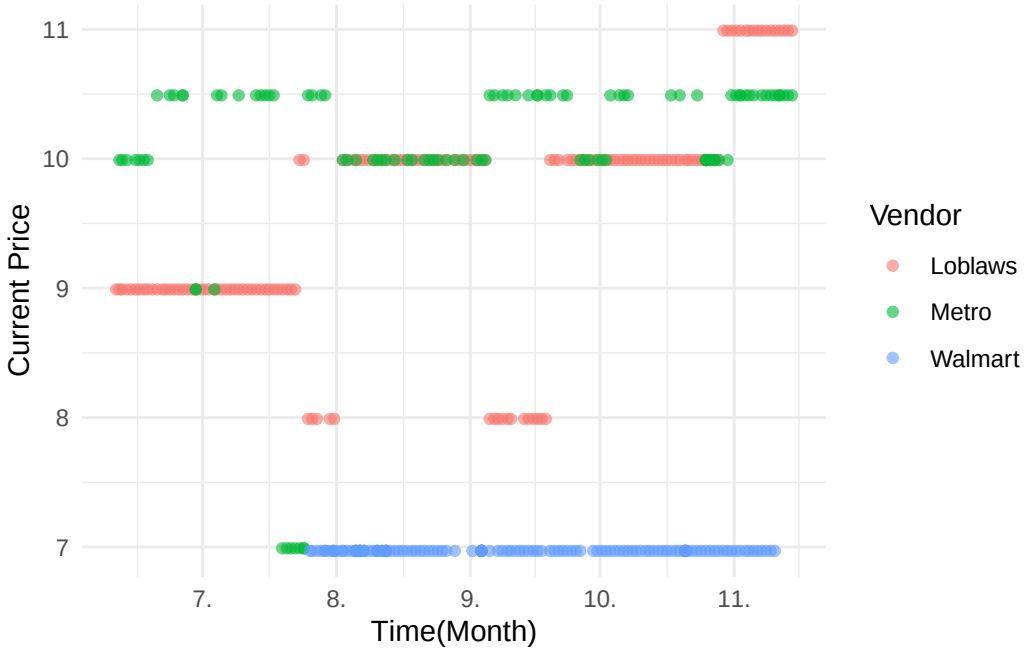

Figure 2: Price Change of Lipton Ginseng Lemon Green Tea 72 Counts

focused on more affordable options. TandT has the most significant spread of green tea related product prices, suggesting it offers both standard and high-end teas, appealing to a diverse range of consumers. Walmart and Galeria have relatively low prices, which may attract budget-conscious consumers looking for affordable tea options. Loblaws and Metro appear to offer a slightly higher variety than some of the other vendors, with a few products in a higher price range, possibly catering to both budget and quality-focused customers. Implications for Consumers: For consumers, these distributions indicate that vendors like TandT provide a wider price range, suitable for those seeking premium options, while stores like SaveOnFoods and Walmart cater more to budget-friendly shoppers with a limited range of low-cost teas.

3.2 Differentiated Pricing

Figure 2 shows the change in price for the same product(Lipton Ginseng Lemon Green Tea 72 Counts) under different grocers. This proved us another possible interpretation for the difference in stock, that possibly as the largest retailer in the world, Walmart is able to use its negotiating power and economy of scale to sell the same items at a lower price. While some grocers, like Loblaws are into constantly changing price and using price tricks to attract costumer. These finding all made the previous analysis on price less reliable.

4 Discussion

One of the most notable trends emerging from the results is the varying pricing strategies employed by different vendors. Walmart, with its vast scale and negotiating power, is able to consistently offer lower prices, reflecting its competitive advantage in the market. Walmart's strategy of maintaining lower price points makes it a preferred choice for price-sensitive consumers, particularly those seeking affordable options in the grocery segment.

On the other hand, Loblaws appears to adopt a more dynamic pricing strategy, using frequent price changes and promotional tactics to attract customers. Loblaws' approach may be particularly effective in attractive in customers who value perceived savings from discounts and deals, though it can create some uncertainty about actual costs and ethical doubts. These pricing shifts can make it difficult for consumers to predict the total cost of their grocery bill, potentially leading to frustration or a sense of distrust in the pricing system.

The presence of T&T with a broad price range for premium teas suggests that it is positioning itself as a specialty retailer, catering to consumers who are not only price-conscious but also looking for high-end or niche green tea products. This could indicate a focus on niche markets, especially on the small population that have higher requirement on teas. T&T's strategy have successfully targeted its main costumers, which are Asian population with a high preference for tea. Which aligns with our initial hypothesis that Asian population favors high quility tea.

The relatively narrow price range found in NoFrills suggests a different strategy focused on affordability and simplicity. With most of its tea offerings priced under $10, NoFrills likely targets budget-conscious consumers who prioritize basic grocery needs. This positions NoFrills as a more accessible option for consumers with limited budgets, but it may also limit its appeal to those seeking variety or premium products.

4.1 Limitations and Improvements

There may be several important limitations in our study and data which restrict the effectiveness of our result, we will discuss three of them below and also provide possible direction for future studies to address these limitations.

4.1.1 Correlation Versus Causation

In the data we have seen that T&T is willing to offer more product in the relative medium price range and high price green tea products, comparing to Walmart, who focuses on offing low price product. Despite the clarity of the trend, we cannot definitively determine its actual causation. It is is possible that Walmart in in Toronto happens to locate in neighborhoods that favors this stock, and is willing to provide similar high-priced products if it is in the location

where T&T is now. Excessive supplementary research will be required to identify possible influence factors, analysis the influence of each of them, and finally reach a conclusion of what caused this difference.

And as we stated, there may not be a great different in types of goods, the difference may be in the pricing. Resolving the problem of missing identification for identical items may show us if there is a real difference in stocks.

4.1.2 Missing Data

There where many data points missing due to problems during data collection and relatively late start of data collection for many products. This may cause challenges in drawing accurate conclusions or identifying trends, as the incomplete dataset might not represent the true patterns or variability within the data. The gaps could introduce biases, skewing results and potentially affecting the reliability of analyses. Furthermore, the late start of data collection for certain products might result in insufficient time for capturing seasonal or time-sensitive changes, further limiting the accuracy of the findings. It is known (Charlebois 2024) that price freeze during November to February is a common practice, however this data is unable to capture this and similar seasonal patterns as the interval is out of scope.

Also the difference in naming across different vendors has also caused difficulties in identifying same product sold by different vendors. This can be viewed as the lack of a product id column that is consistent across vendors.The lack of such data may cause error in analyzing, such as error in identifying identical products sold at different stores.Such error may be reduced with better data collection and cleaning process.

4.1.3 Sources of Bias.

The data set may have implicitly introduced source of bias in some ways. Firstly the organizer of this dataset have explicitly picked 8 vendors to collect data from. Which does not cover all of Canada's grocery market, and since the vendors examined are all major players in the Canadian retail and grocery industry, the data set may become highly unrepresentative for small private businesses. Which may skew the data, overlooking the practices of independent or regional grocery stores that often serve different customer bases, operate on smaller scales, and follow different business models.

A more comprehensive study that includes these smaller retailers is required to understand small firms that does receive less benefit from economy of scale and may be focused to serve local residents.

as it is likely that the dataset is skewed toward the most commonly sold items at major chains, which are typically mass-produced and mainstream products. And are easier to collect data and remember to include. Smaller stores often feature more niche items, locally sourced

goods, or specialty products that cater to specific communities or tastes, which may not be represented in this dataset. These unique products could differ in terms of pricing, availability, and consumer demand, and excluding them would result in an incomplete understanding of the diversity in consumer preferences and retail offerings. A more comprehensive study can also focus on including more types of good for a full analysis.

References

Charlebois, Sylvain. 2024. "Are Grocers Still Colluding?" *Cochrane Times Post*. https://www.cochranetimespost.ca/opinion/columnists/charlebois-are-grocers-still-colluding.

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

International Organization for Standardization. 2023. *ISO/IEC 9075-1:2023 Information Technology — Database Languages SQL*. International Organization for Standardization.

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. 2019. "Welcome to the tidyverse." *Journal of Open Source Software* 4 (43): 1686. https://doi.org/10.21105/joss.01686.