

# Seasonal Variations and Vendor-Specific Strategies Influencing Price and Sales of Vanilla Ice Cream\*

Exploring Price and Sales Dynamics of Vanilla Ice Cream Across Vendors and Months

Chiyue Zhuang

December 3, 2024

This study examines pricing and sales patterns for vanilla ice cream, focusing on seasonal trends, vendor-specific strategies, and the relationship between old and current prices. Data analysis reveals that Metro employs a premium pricing strategy with higher average prices, while Loblaw's offers lower and more consistent pricing. Seasonal patterns indicate peak prices during summer and early fall, reflecting demand-driven adjustments. Additionally, a strong correlation between old and current prices highlights systematic and proportional discounting strategies. These findings shed light on the interplay between pricing strategies and consumer behavior, providing insights valuable into consumer pricing dynamics and vendor approaches to market positioning.

## 1 Introduction

The Canadian grocery market plays a pivotal role in shaping the everyday lives of consumers, with leading vendors such as Metro and Loblaw's setting the stage for competitive retail strategies. Metro, known for its regional dominance in Ontario and Quebec, competes head-to-head with Loblaw's, a national chain with a broader footprint across Canada. Both retailers have implemented nuanced pricing strategies to cater to diverse consumer preferences, particularly for staple and indulgent items like vanilla ice cream—a perennial favorite across demographics (Sharpe, Peterkin, and Dudas 1979).

---

\*Code and data are available at: [[https://github.com/cZhuang17/M-L\\_vanilla\\_ice\\_cream](https://github.com/cZhuang17/M-L_vanilla_ice_cream)].

Vanilla ice cream occupies a unique position in the Canadian grocery landscape, demonstrating notable fluctuations in demand based on seasonal trends. Warmer months, particularly summer, mark a peak in sales, while cooler months witness a tempered but steady interest in the product as a comfort food. Understanding these seasonal variations is essential for retailers to optimize pricing and inventory management. Despite its popularity, there is limited empirical research on how pricing strategies for vanilla ice cream vary between competitors like Metro and Loblaws, particularly from June to November when demand patterns shift dramatically (Martin 2018).

Our research aims to fill this gap by analyzing the pricing and sales dynamics of vanilla ice cream offered by Metro and Loblaws. Using a dataset spanning six months, we examined how pricing strategies differ between the two vendors and how seasonal consumer behavior influences these dynamics. By dissecting trends in pricing and demand, we uncovered insights into how each vendor adapts to market fluctuations and consumer preferences during this period.

Key findings from our study reveal that both Metro and Loblaws employ distinct pricing strategies shaped by competition, seasonality, and consumer behavior. For instance, Loblaws exhibits more aggressive promotional pricing during peak summer months, while Metro focuses on steady, value-based pricing throughout the period. Additionally, our analysis highlights the role of consumer behavior in shaping these strategies, with trends like stockpiling during colder months and impulse purchases in summer impacting sales dynamics.

Understanding the intricacies of vanilla ice cream pricing is not just an academic exercise but a vital contribution to the retail industry and broader economy. Insights from this research can help vendors refine their strategies to maximize profitability and consumer satisfaction, while also informing policymakers and economists about the microeconomic dynamics of seasonal retail markets. Furthermore, the study underscores the importance of adaptive pricing models in maintaining competitiveness and meeting evolving consumer needs.

The remainder of this paper is structured as follows: Data section outlines the methodology and data sources used for this analysis. Graph and result section delves into the pricing and sales trends observed across Metro and Loblaws. Discussion section discusses the implications of our findings for market strategies and consumer behavior. Finally, the model section goes through the application of models in this project.

## **1.1 Estimand**

This study aims to estimate the relationship between vendor pricing strategies and seasonal variations in vanilla ice cream prices. Specifically, we assess how average prices differ between vendors and how monthly price trends reflect demand changes. By analyzing various factors such as the different type of vendors: Metro and Loblaws, current selling price of vanilla ice cream, old selling price of vanilla ice cream, different month in the year from June to November, this research seeks to understand how these variables, especially the variable of

type of vendors, and the variables of months influences the pricing of vanilla ice cream and how we can implement this result in future pricing decision and sales strategy.

## 2 Data

### 2.1 Data Source & Measurement

The dataset used in this study was sourced from [Jacob Filip Project Hammer](#), which compile extensive database of historical grocery prices from top grocers' websites. The dataset originally consists of 8 vendors in the Canadian grocery market: Voila, T&T, Loblaws, No Frills, Metro, Galleria, Walmart and Save-On-Foods. The raw data has approximately 121,172,21 records, after we narrow down our research to be specifically focus on the product vanilla ice cream in 2 vendors: Metro and Loblaws, the observations reduced to 971 record. The dataset underwent significant preprocessing to align with the specific focus of this research. Our initial step in cleaning the raw dataset involved filtering out records where the target product is not vanilla ice cream, and the target vendors are not Metro and Loblaws, reducing the dataset to 971 entries. These records form the basis of the analysis, having been retained for detailed examination and modeling of current price and quantity sold of vanilla ice cream by different vendors under different seasons.

The data for this study was systematically downloaded, cleaned, analyzed, modeled and visualized using R (R Core Team 2023), a extensive statistical programming language. The following packages were used for this study

- **tidyverse** (Wickham et al. 2021): To simplify and enhance the efficiency of data handling and visualization.
- **testthat** (Wickham, Hester, and Csardi 2023): Simplifies writing tests by describing expected behavior, catching errors, and visually displaying test results, while seamlessly integrating into various workflows.
- **arrow** (Apache Arrow 2021): Designed for efficient reading and writing of large datasets, improving data processing capabilities.
- **janitor** (Firke 2021): Simplifying raw data management by standardizing variable names and streamlining data structures.
- **lubridate** (Grolemund and Wickham 2011): Makes it easier to do the things R does with date-times and possible to do the things R does not.
- **Rcpp** (Eddelbuettel and Francois 2011): Provides a clean, approachable API that lets you write high-performance code, insulated from R's arcane C API.
- **modelsummary** (Arel-Bundock 2023): Creates tables and plots to present descriptive statistics and to summarize statistical models in R.
- **palmerpenguins** (Horst, Hill, and Gorman 2020): Provide a great dataset for data exploration & visualization, as an alternative to iris . Installation.

- **knitr** (Xie 2021): Used to dynamically create reports that combine R code with its outputs, enabling the seamless integration of plots and analysis results into the final document.
- **ggplot2** (Wickham 2021): Make testing enjoyable and rewarding, so you feel a genuine sense of satisfaction when writing tests.
- **rstanarm** (Goodrich et al. 2022): Enabled the implementation of Bayesian models, offering an easy-to-use approach for fitting regression models with Stan.
- **bayesplot** (Gelman, Gabry, et al. 2021): Used to generate graphical posterior predictive checks and diagnostic visualizations.
- **plumber** (Trestle Technology and Schloerke 2023): Allows you to create a web API by merely decorating your existing R source code with roxygen2 -like comments.

Table 1: Current and Old Price of Vanilla Ice Cream.

vendor	current_price	old_price	product_name	month
Loblaws	3.49	6.99	French Vanilla Ice Cream	7
Loblaws	3.49	6.99	French Vanilla Ice Cream	7
Loblaws	3.49	6.99	French Vanilla Ice Cream	7
Loblaws	3.49	6.99	French Vanilla Ice Cream	7
Loblaws	3.49	6.99	French Vanilla Ice Cream	7
Loblaws	3.49	6.99	French Vanilla Ice Cream	7

Table 1 presents the first six rows from the cleansed dataset, focusing on the current and old price of French vanilla ice cream in Loblaws in the 7th month.

## 2.2 Variable

Our analysis focuses on the following variables, with a specific focus on **current\_price** as the dependent variable:

- **current\_price**: The current price of Vanilla ice cream.
- **old-price**: The historial price of vanilla ice cream.
- **quantity\_sold**: The total number of vanilla ice cream sold.
- **month**: The month variables of vanilla ice cream price:
  - **June**: The 6th month of the year.
  - **July**: The 7th month of the year.
  - **Augst**: The 8th month of the year.
  - **September**: The 9th month of the year.
  - **October**: The 10th month of the year.

- **November:** The 11th month of the year.
- **vendor:** The grocery store.
  - **Loblaws:** A Canadian supermarket chain, part of Loblaw Companies Limited, offering groceries, household items, and pharmacy services, with a focus on affordability and accessibility.
  - **Metro:** A Canadian supermarket chain operating in Quebec and Ontario, offering a variety of products including groceries, pharmacy items, and general merchandise.

Detailed information about these variables and the data structure is presented in Table 1, which outlines the first few records from the processed dataset.

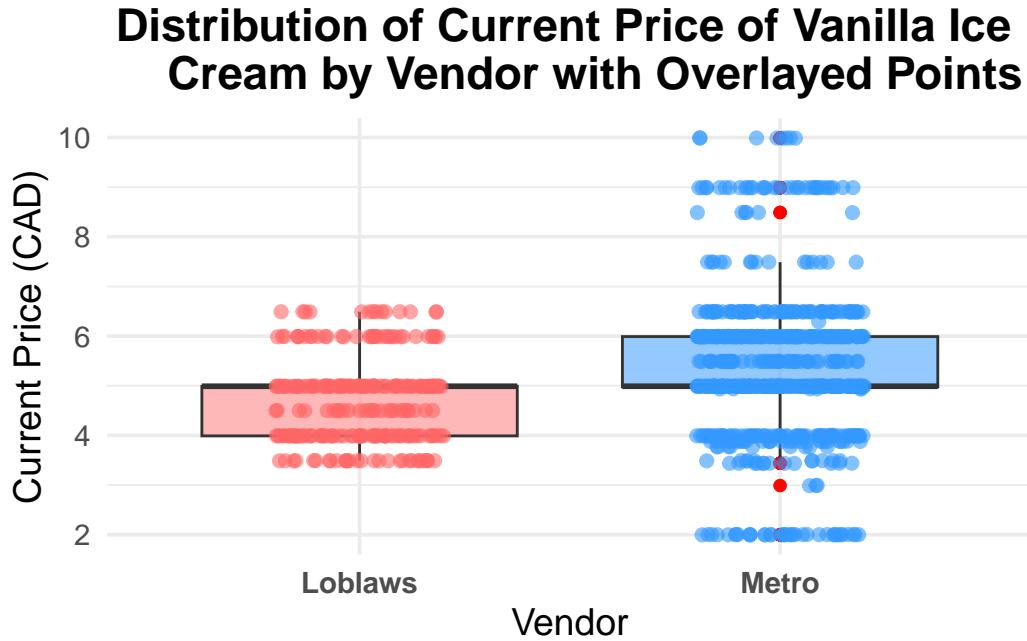


Figure 1: Distribution of Current Price of Vanilla Ice Cream by Vendor with Overlaid Data Points, Showing that Metro has consistent Higher Current Price.

Figure 1 illustrates the distribution of current prices for vanilla ice cream by vendor, with additional jittered points overlaid to represent individual data points. The vendors being compared are Loblaws and Metro. The visualization reveals that Metro maintains a narrower range of prices, reflecting consistency in its pricing strategy. In contrast, Loblaws demonstrates more variability, as evidenced by a wider interquartile range (IQR). This suggests a more diverse pricing approach at Loblaws, potentially to attract a broader customer base. Additionally, the median price at Metro is noticeably higher than that of Loblaws, highlighting Metro's premium pricing strategy. Outliers are present for both vendors, but Metro shows fewer deviations from its core pricing, indicating more standardized pricing practices.

## Distribution of Current Price of Vanilla Ice Cream by Vendor

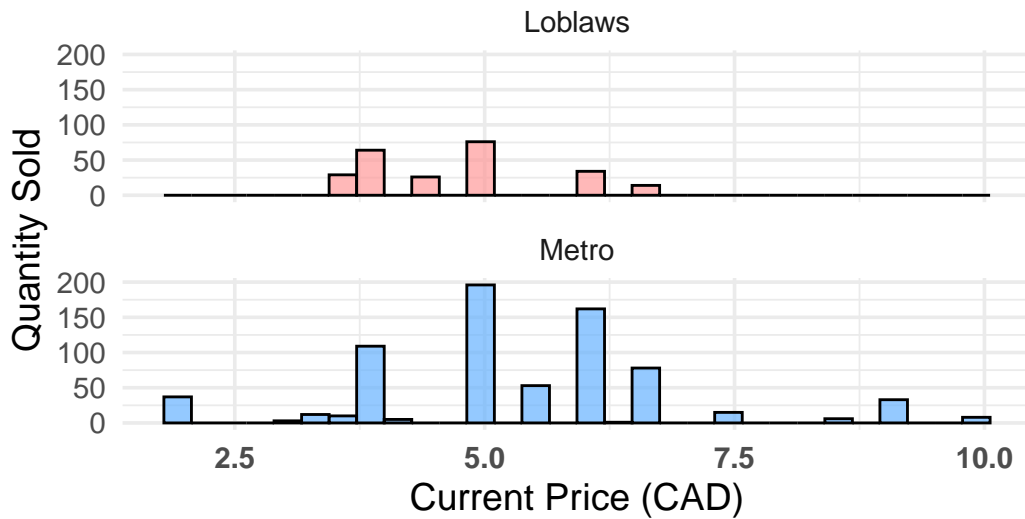


Figure 2: Distribution of Current Price and Quantity Sold of Vanilla Ice Cream by Vendor under Histogram, Showing that Metro has consistent Higher Current Price and Vanilla Ice Cream is Sold the Best at Price of 5 in Both Grocery Stores.

Figure 2 provides a segmented view of the distribution of current prices for vanilla ice cream, divided by vendor. Each vendor's data is visually distinguished using separate color coding and stacking. Metro's price distribution is more concentrated in the higher price range, with most of its prices above 5 dollars. This observation aligns with Metro's premium positioning as reflected in the boxplot. On the other hand, Loblaw's displays a wider spread of prices with a notable peak around 3.50 dollars, indicating a pricing strategy aimed at affordability. While there is some overlap in pricing between the two vendors, Metro's distribution skews towards higher price points, suggesting that the two vendors may be targeting different segments of the market.

The two graphs together provide a comprehensive view of the pricing strategies employed by Loblaw's and Metro. Metro's pricing strategy is characterized by consistency and a focus on higher price points, targeting a premium market segment. Conversely, Loblaw's shows a greater range of prices, with a lower median price, suggesting an effort to cater to a broader or more price-sensitive customer base. These patterns indicate distinct market positioning for each vendor, with Metro appealing to customers willing to pay a premium and Loblaw's focusing on affordability and variety. This analysis highlights the competitive dynamics between the two vendors and offers insights into their respective market strategies.

Figure 3 provides a detailed view of the quantity of vanilla ice cream sold at various price points, segmented by vendor, namely Metro and Loblaw's. By using a stacked format, the

## Distribution of Quantity Sold of Vanilla Ice Cream in Different Price by Vendor

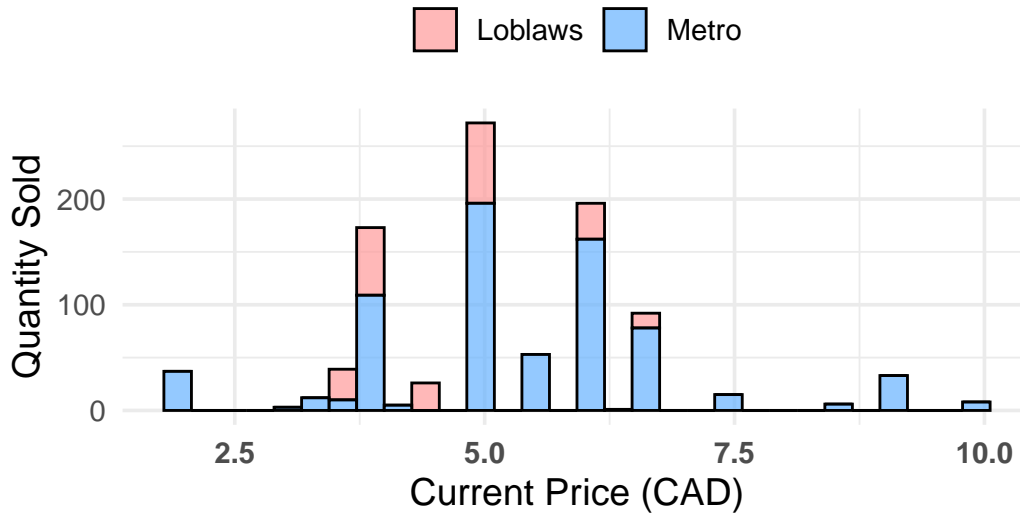


Figure 3: Distribution of Quantity Sold of Vanilla Ice Cream under Different Price by Vendor in Histogram, Showing that Metro always Sell More Vanilla Ice Cream than Loblaws, and At Price of 5, the Product is Sold the Best.

graph allows for a clear comparison of sales volumes between the two vendors across different pricing ranges. The visual distinction highlights Metro's dominant position in sales volume at nearly all price points, as reflected by the larger representation of Metro's segment in the stacked bars.

One of the key insights from the graph is the identification of an optimal price point of 5 dollars, where both vendors achieve their highest sales volume. This suggests that 5 dollars is a critical price for attracting customer demand and driving sales for vanilla ice cream. As prices increase beyond this point, the overall sales volumes decrease, indicating that customer willingness to purchase diminishes at higher price levels. Despite this trend, Metro continues to outperform Loblaws even in the higher price ranges, maintaining relatively consistent sales compared to its competitor.

The data also underscores differences in market positioning between the two vendors. Metro's stronger performance across all price points reflects its ability to cater to a wider or more premium customer base. Loblaws, on the other hand, may be more competitive within specific lower price brackets, but struggles to match Metro's performance overall. These observations emphasize the importance of strategic pricing in aligning with customer behavior and maximizing sales potential. Metro's consistency and dominance suggest a well-calibrated pricing approach, while Loblaws might benefit from refining its strategies to compete more effectively across price ranges.

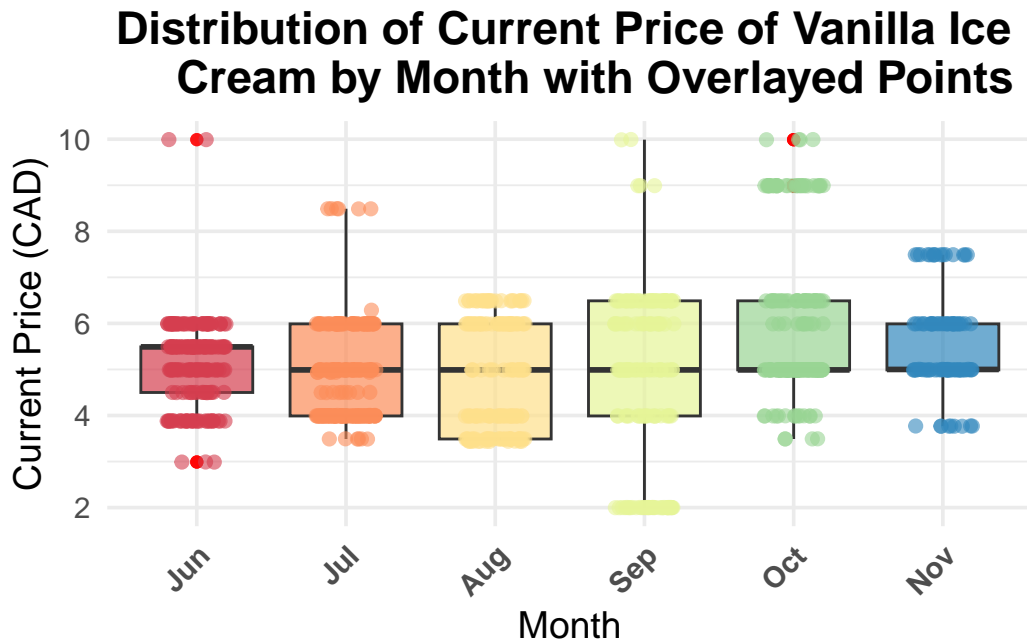


Figure 4: Distribution of Current Price of Vanilla Ice Cream by Month with Overlaid Data Points under Boxplot.

Figure 4 illustrates the distribution of current prices for vanilla ice cream across the months of the year using a boxplot, enhanced by overlaid data points for additional detail. Each box represents the interquartile range (IQR) of prices for a given month, while the whiskers indicate the full spread of the data, excluding outliers. The overlaid points show individual price instances, providing granular insights into pricing trends.

From this graph, it is evident that price consistency varies throughout the year. Some months, such as the summer months of June, July, and August, show tighter distributions with fewer outliers, indicating more stable pricing during peak ice cream demand. In contrast, months like November and December show wider price ranges and more significant outliers, suggesting increased price variability, likely due to lower demand or promotional activities. Overall, the summer months exhibit higher median prices, aligning with increased consumer demand during warmer seasons.

Figure 5 uses histograms to display the distribution of current prices and the corresponding quantities sold for vanilla ice cream, segmented by month. Each month's sales distribution is shown separately, facilitating a comparison across months. The histograms reveal a clear trend: sales volumes peak during the summer months (June, July, and August), suggesting that vanilla ice cream is a seasonal product with heightened consumer interest in warmer weather.

For most months, the highest sales occur at the \$5 price point, confirming it as a critical



## Distribution of Current Price and Quantity Sold Vanilla Ice Cream by Month

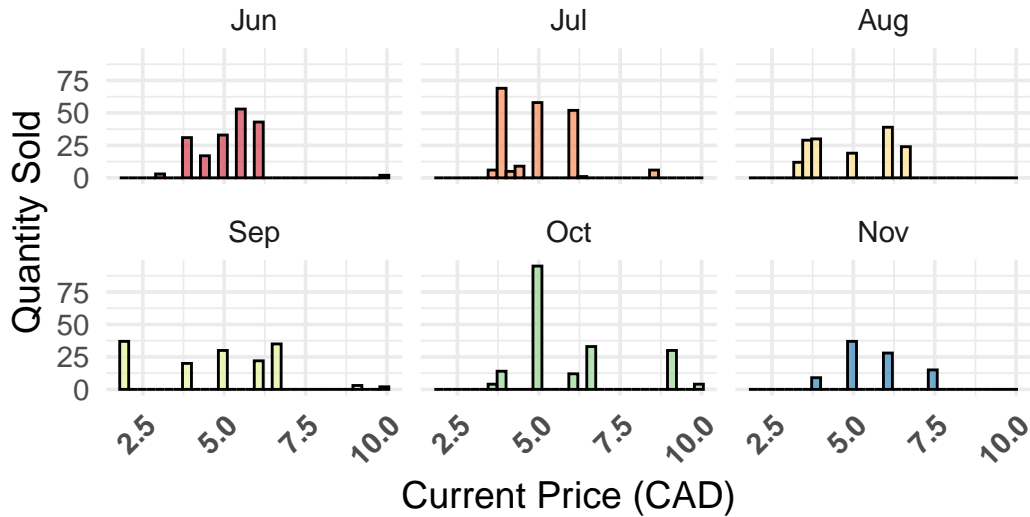


Figure 5: Distribution of Current Price and Quantity Sold of Vanilla Ice Cream by Month under Histogram, Showing that Vanilla Ice Cream is Sold the Best During Summer.

price that maximizes sales. In the summer months, the concentration of sales at this price point is even more pronounced, with relatively higher volumes compared to other months. Winter months, such as December and January, show reduced sales volumes overall, and the distributions appear more spread out, reflecting lower consumer demand and more varied pricing strategies.

Together, the two graphs paint a comprehensive picture of the seasonal dynamics in the vanilla ice cream market. The boxplot highlights the stability and seasonal variations in pricing, with higher and more consistent prices during the summer months. Meanwhile, the histogram emphasizes the relationship between price and sales, showcasing the dominance of the \$5 price point in driving sales volumes, particularly in the peak summer season. These insights underline the importance of seasonal demand and strategic pricing in maximizing revenue. Retailers could leverage this knowledge to optimize inventory and pricing strategies, ensuring they capitalize on the peak demand during the summer months while maintaining competitiveness during off-peak seasons.

Figure 6 shows a positive relationship between old and current prices of vanilla ice cream, with current prices increasing as old prices rise. The color gradient highlights higher old prices corresponding to higher current prices, and the linear regression line reinforces this upward trend. The clustering along the line suggests systematic pricing adjustments, though some variability indicates other influencing factors.

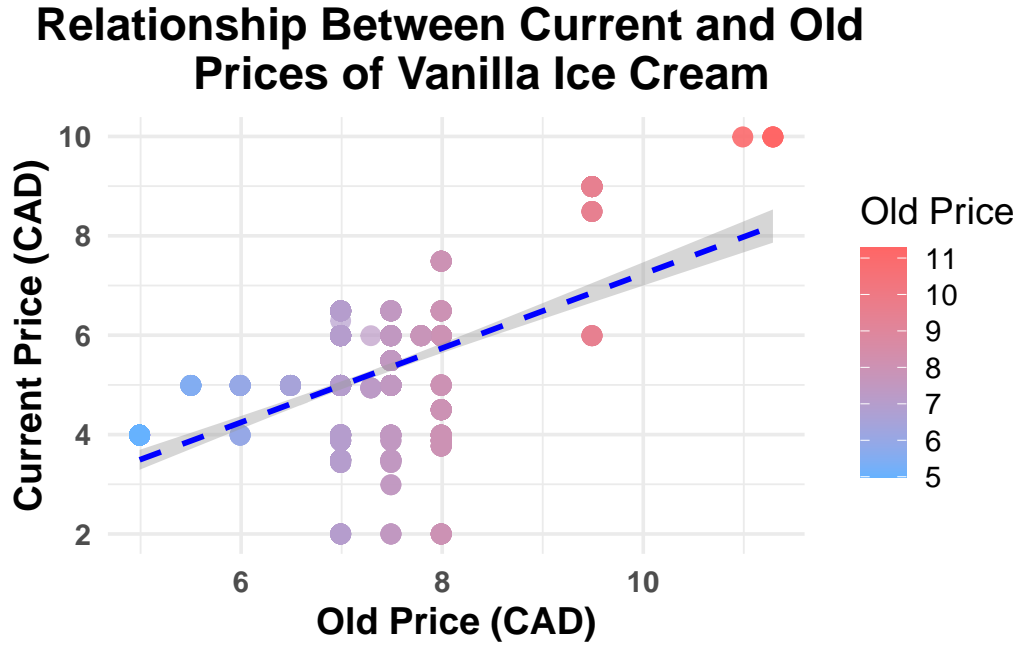


Figure 6: Relationship Between Current and Old Prices of Vanilla Ice Cream, Showing that It Follows an Upwarding Direction Distribution.

The positive correlation between old and current prices suggests a proportional pricing strategy, aligning current prices with historical trends. This systematic approach helps maintain product value and adapt to market conditions, providing valuable insights for predicting and strategizing future pricing adjustments.

### 2.3 Justification

The variables chosen for this study were carefully selected based on their strategic significance and impacts on sales of vanilla ice cream, as documented in the dataset. Each variable plays important role in affecting the price and sales of vanilla ice cream in both vendors:

- **Target Vendors (VENDOR):** Identified based on the grocery store type, targeting different Canadian vendors, such as Metro and Loblaws.
- **Target Month (MONTH):** Identified based on different time period of the year, targeting different number of months, such as June, July, August, September, October, and November.
- **Current Price (CURRENT\_PRICE):** The current selling price of vanilla ice cream.
- **Old Price (OLD\_PRICE):** The historical selling price of vanilla ice cream.
- **Quantity Sold (QUANTITY\_SOLD):** The count of vanilla ice cream sold at each targeted selling price.

To streamline the modeling process and enhance the clarity of the analysis, non-vanilla ice cream targets and non-Metro, non-Loblaws targets were excluded from the original dataset of approximately 121,172,21 records. This data cleaning step resulted in a more focused dataset, comprising 971 entries, which facilitated a more efficient and targeted modeling approach. This pre-processing not only improved the manageability of the data but also enhanced the study's examination of the price trends for vanilla ice cream at Metro and Loblaws across different months, providing a clear view of pricing patterns and vendor strategies.

### 3 Model

This study employs a Bayesian ordered logistic regression model to analyze the relationship between the current and old price, quantity sold of vanilla ice cream and the two grocery store: Loblaws and Metro.

The model integrates several key predictors:

- Grocery type : Identifies the grocery sector being targeted. Distinguish between Loblaws and Metro. Variable: **Vendor**
- Old Price: Reflecting the impact of historical prices of vanilla ice cream on current pricing strategies. Variable: **Old\_price**
- Current Price: Reflecting the current price of vanilla ice cream sold in the mark in the two vendors. Variable: **Current\_price**
- Quantity: Measures the actual number of vanilla ice cream sold under different price. Variable: **Quantity\_sold**
- Month: Considering seasonal price fluctuations, like higher demand in the summer months. Measures from June to November. Variable: **Month**

A Gaussian likelihood is used for the response variable, suitable for the continuous nature of current prices. Informative priors are assigned to the coefficients and auxiliary parameters, including the intercept and residual variance, enabling the model to incorporate domain knowledge while adapting to the data.

Our model statistically infers the relationship between current and old price and various month factors in different vendors, providing a probabilistic assessment of their relationships and impacts.

### 3.1 Features of Bayesian Approach

1. Incorporation of Prior Knowledge: The Bayesian approach allows the integration of prior beliefs or domain expertise into the model through the use of priors, which can guide inference, especially when data is limited.
2. Probabilistic Inference: It provides a full posterior distribution of parameters, offering a measure of uncertainty and variability in estimates rather than just point estimates.
3. Flexibility in Modeling: Bayesian methods can handle complex models and hierarchical structures, making them suitable for a wide range of applications and accommodating various types of data and assumptions.

Background details and diagnostics are included in Appendix [A](#).

### 3.2 Model set-up

Let  $y_i$  be the the current price of vanilla ice cream. The predictors in the model include:

- $\beta_1$ : The coefficient for the `month`, which indicates the period of dataset from June to November. It is a categorical variable with the following categories:
  - June
  - July
  - August
  - September
  - October
  - November
- $\beta_2$ : The coefficient for the `old_price` variable, representing the historical price of vanilla ice cream.
- $\beta_3$ : The coefficient for the `vendors`, which is a category variable, suggesting the different grocery store targeted, distinguished from Loblaws and Metro.
  - Loblaws
  - Metro
- $\eta_i$ : The linear predictor or log-odds for the  $i$ -th observation. It is a combination of the intercept and coefficients multiplied by the predictor variables.
- $\kappa$ : Thresholds (or cutpoints) define the boundaries between the ordered categories of the dependent variable in an ordered logistic regression model. These thresholds specify the ranges of  $\eta_i$  values associated with each target priority category. The model estimates distinct  $\kappa$  values for each boundary separating the ordered categories.

$$y_i \sim \text{OrderedLogistic}(\eta_i, \kappa) \quad (1)$$

$$\eta_i = \beta_{\text{month}} \times \text{month}_i + \beta_{\text{old\_price}} \times \text{old\_price} \quad (2)$$

$$+ \beta_{\text{vendor}} \times \text{vendor} \quad (3)$$

$$\beta \sim \text{Normal}(0, 10) \text{ (default non-informative prior)} \quad (4)$$

$$\kappa \sim \text{Normal}(0, 5) \text{ (default prior for cutpoints)} \quad (5)$$

### 3.3 Prior distributions

In the Bayesian ordered logistic regression model implemented with the `rstanarm` package, default priors are utilized to promote robust and reliable inference. These priors are weakly informative, striking a balance between regularization and flexibility to accommodate the data:

- **Intercept Priors:** Normal priors with a mean of 0 are applied to the model’s intercepts. This choice stabilizes the location parameter without enforcing a strong assumption about its central value.
- **Coefficient Priors:** Model coefficients also follow normal priors with a mean of 0 and a standard deviation typically set to 2.5. This standard deviation limits the likelihood of overly large effects unless strongly supported by the data, offering regularization.

We run the model in R (R Core Team 2023) using the `rstanarm` package of Goodrich et al. (2022). We use the default priors from `rstanarm`. `date` is from Filipp (2020).

## 4 Results

### 4.1 Model Justification

The analysis revealed distinct patterns in vanilla ice cream pricing, highlighting the role of vendor type—Metro and Loblaws—and seasonal factors in determining price variations, while the impact of product volume and promotional offers was relatively subdued. Our results are summarized in Table 2.

```
SAMPLING FOR MODEL 'continuous' NOW (CHAIN 1).
```

```
Chain 1:
```

```
Chain 1: Gradient evaluation took 0.004806 seconds
```

```
Chain 1: 1000 transitions using 10 leapfrog steps per transition would take 48.06 seconds.
```

```
Chain 1: Adjust your expectations accordingly!
```

```

Chain 1:
Chain 1:
Chain 1: Iteration:    1 / 2000 [  0%] (Warmup)
Chain 1: Iteration:   200 / 2000 [ 10%] (Warmup)
Chain 1: Iteration:   400 / 2000 [ 20%] (Warmup)
Chain 1: Iteration:   600 / 2000 [ 30%] (Warmup)
Chain 1: Iteration:   800 / 2000 [ 40%] (Warmup)
Chain 1: Iteration:  1000 / 2000 [ 50%] (Warmup)
Chain 1: Iteration: 1001 / 2000 [ 50%] (Sampling)
Chain 1: Iteration: 1200 / 2000 [ 60%] (Sampling)
Chain 1: Iteration: 1400 / 2000 [ 70%] (Sampling)
Chain 1: Iteration: 1600 / 2000 [ 80%] (Sampling)
Chain 1: Iteration: 1800 / 2000 [ 90%] (Sampling)
Chain 1: Iteration: 2000 / 2000 [100%] (Sampling)
Chain 1:
Chain 1: Elapsed Time: 0.114 seconds (Warm-up)
Chain 1:                0.282 seconds (Sampling)
Chain 1:                0.396 seconds (Total)
Chain 1:

```

SAMPLING FOR MODEL 'continuous' NOW (CHAIN 2).

```

Chain 2:
Chain 2: Gradient evaluation took 7.3e-05 seconds
Chain 2: 1000 transitions using 10 leapfrog steps per transition would take 0.73 seconds.
Chain 2: Adjust your expectations accordingly!
Chain 2:
Chain 2:
Chain 2: Iteration:    1 / 2000 [  0%] (Warmup)
Chain 2: Iteration:   200 / 2000 [ 10%] (Warmup)
Chain 2: Iteration:   400 / 2000 [ 20%] (Warmup)
Chain 2: Iteration:   600 / 2000 [ 30%] (Warmup)
Chain 2: Iteration:   800 / 2000 [ 40%] (Warmup)
Chain 2: Iteration:  1000 / 2000 [ 50%] (Warmup)
Chain 2: Iteration: 1001 / 2000 [ 50%] (Sampling)
Chain 2: Iteration: 1200 / 2000 [ 60%] (Sampling)
Chain 2: Iteration: 1400 / 2000 [ 70%] (Sampling)
Chain 2: Iteration: 1600 / 2000 [ 80%] (Sampling)
Chain 2: Iteration: 1800 / 2000 [ 90%] (Sampling)
Chain 2: Iteration: 2000 / 2000 [100%] (Sampling)
Chain 2:
Chain 2: Elapsed Time: 0.111 seconds (Warm-up)
Chain 2:                0.204 seconds (Sampling)
Chain 2:                0.315 seconds (Total)

```

Chain 2:

SAMPLING FOR MODEL 'continuous' NOW (CHAIN 3).

Chain 3:

Chain 3: Gradient evaluation took 2.3e-05 seconds

Chain 3: 1000 transitions using 10 leapfrog steps per transition would take 0.23 seconds.

Chain 3: Adjust your expectations accordingly!

Chain 3:

Chain 3:

Chain 3: Iteration: 1 / 2000 [ 0%] (Warmup)

Chain 3: Iteration: 200 / 2000 [ 10%] (Warmup)

Chain 3: Iteration: 400 / 2000 [ 20%] (Warmup)

Chain 3: Iteration: 600 / 2000 [ 30%] (Warmup)

Chain 3: Iteration: 800 / 2000 [ 40%] (Warmup)

Chain 3: Iteration: 1000 / 2000 [ 50%] (Warmup)

Chain 3: Iteration: 1001 / 2000 [ 50%] (Sampling)

Chain 3: Iteration: 1200 / 2000 [ 60%] (Sampling)

Chain 3: Iteration: 1400 / 2000 [ 70%] (Sampling)

Chain 3: Iteration: 1600 / 2000 [ 80%] (Sampling)

Chain 3: Iteration: 1800 / 2000 [ 90%] (Sampling)

Chain 3: Iteration: 2000 / 2000 [100%] (Sampling)

Chain 3:

Chain 3: Elapsed Time: 0.097 seconds (Warm-up)

Chain 3: 0.195 seconds (Sampling)

Chain 3: 0.292 seconds (Total)

Chain 3:

SAMPLING FOR MODEL 'continuous' NOW (CHAIN 4).

Chain 4:

Chain 4: Gradient evaluation took 2.6e-05 seconds

Chain 4: 1000 transitions using 10 leapfrog steps per transition would take 0.26 seconds.

Chain 4: Adjust your expectations accordingly!

Chain 4:

Chain 4:

Chain 4: Iteration: 1 / 2000 [ 0%] (Warmup)

Chain 4: Iteration: 200 / 2000 [ 10%] (Warmup)

Chain 4: Iteration: 400 / 2000 [ 20%] (Warmup)

Chain 4: Iteration: 600 / 2000 [ 30%] (Warmup)

Chain 4: Iteration: 800 / 2000 [ 40%] (Warmup)

Chain 4: Iteration: 1000 / 2000 [ 50%] (Warmup)

Chain 4: Iteration: 1001 / 2000 [ 50%] (Sampling)

Chain 4: Iteration: 1200 / 2000 [ 60%] (Sampling)

Chain 4: Iteration: 1400 / 2000 [ 70%] (Sampling)

```

Chain 4: Iteration: 1600 / 2000 [ 80%] (Sampling)
Chain 4: Iteration: 1800 / 2000 [ 90%] (Sampling)
Chain 4: Iteration: 2000 / 2000 [100%] (Sampling)
Chain 4:
Chain 4: Elapsed Time: 0.104 seconds (Warm-up)
Chain 4:           0.191 seconds (Sampling)
Chain 4:           0.295 seconds (Total)
Chain 4:

```

#### Model Info:

```

function:      stan_glm
family:        gaussian [identity]
formula:       current_price ~ month + old_price + vendor
algorithm:     sampling
sample:        4000 (posterior sample size)
priors:        see help('prior_summary')
observations:  971
predictors:    4

```

#### Estimates:

	mean	sd	10%	50%	90%
(Intercept)	-1.5	0.3	-2.0	-1.5	-1.1
month	0.2	0.0	0.1	0.2	0.2
old_price	0.7	0.0	0.7	0.7	0.8
vendorMetro	0.2	0.1	0.1	0.2	0.4
sigma	1.2	0.0	1.2	1.2	1.2

#### Fit Diagnostics:

	mean	sd	10%	50%	90%
mean_PPD	5.2	0.1	5.1	5.2	5.3

The mean\_ppd is the sample average posterior predictive distribution of the outcome variable

#### MCMC diagnostics

	mcse	Rhat	n_eff
(Intercept)	0.0	1.0	5216
month	0.0	1.0	4064
old_price	0.0	1.0	4298
vendorMetro	0.0	1.0	3816
sigma	0.0	1.0	4709
mean_PPD	0.0	1.0	4259



log-posterior 0.0 1.0 1891

For each parameter, mcse is Monte Carlo standard error, n\_eff is a crude measure of effective

Model Info:

function: stan\_glm  
family: gaussian [identity]  
formula: current\_price ~ month + old\_price + vendor  
algorithm: sampling  
sample: 4000 (posterior sample size)  
priors: see help('prior\_summary')  
observations: 971  
predictors: 4

Estimates:

	mean	sd	10%	50%	90%
(Intercept)	-1.5	0.3	-2.0	-1.5	-1.1
month	0.2	0.0	0.1	0.2	0.2
old_price	0.7	0.0	0.7	0.7	0.8
vendorMetro	0.2	0.1	0.1	0.2	0.4
sigma	1.2	0.0	1.2	1.2	1.2

Fit Diagnostics:

	mean	sd	10%	50%	90%
mean_PPD	5.2	0.1	5.1	5.2	5.3

The mean\_ppd is the sample average posterior predictive distribution of the outcome variable

MCMC diagnostics

	mcse	Rhat	n_eff
(Intercept)	0.0	1.0	5216
month	0.0	1.0	4064
old_price	0.0	1.0	4298
vendorMetro	0.0	1.0	3816
sigma	0.0	1.0	4709
mean_PPD	0.0	1.0	4259
log-posterior	0.0	1.0	1891

For each parameter, mcse is Monte Carlo standard error, n\_eff is a crude measure of effective

Table 2: The model’s coefficient summary

Parameter	Mean	SD	10%	50%	90%
Intercept	-0.4	0.3	-0.8	-0.4	0.0
monthJul	-0.1	0.1	-0.3	-0.1	0.0
monthAug	0.3	0.1	0.1	0.3	0.4
monthSep	-0.2	0.1	-0.4	-0.2	-0.1
monthOct	1.0	0.1	0.8	1.0	1.1
monthNov	0.5	0.1	0.3	0.5	0.7
old_price	0.7	0.0	0.7	0.7	0.8
vendorMetro	0.3	0.1	0.2	0.3	0.4
sigma	1.1	0.0	1.1	1.1	1.2

As detailed in Table 2, the coefficient summary provides insights into the price variations of vanilla ice cream sold by different vendors across different months. For instance, the estimated coefficient for monthOctober is notably high (Mean = 1.0), suggesting a significant price increase during this month compared to July. This may reflect seasonal demand fluctuations, possibly due to autumn-related events or increased consumer preferences.

In contrast, the coefficient for monthSeptember is negative (Mean = -0.2), indicating a slight price drop compared to the baseline month. Similarly, prices in July (Mean = -0.1) are relatively lower, which might align with promotional strategies during the summer season.

For vendor differences, the coefficient for vendorMetro (Mean = 0.3) suggests that Metro generally has higher prices than Loblaw’s, which could be attributed to differences in vendor strategies or consumer targeting. The positive coefficient for old\_price (Mean = 0.7) emphasizes the influence of historical pricing on current price settings.

The intercept term sets the baseline, with a Mean of -0.4, reflecting the starting point for price variations. The relatively low value of sigma (1.1) indicates moderate variability in pricing, providing a stable prediction model.

From Figure 7, the linear regression model examines the relationship between the current price of products and key predictors, including the old price, month, and vendor. The intercept of 13.57 represents the baseline price when all predictors are zero, with a narrow confidence interval ([12.28, 14.86]), indicating a reliable estimate.

The old price significantly influences the current price, with a coefficient of 0.02 and a confidence interval ([0.02, 0.03]). This suggests that for every unit increase in the old price, the current price increases by approximately 2%. This highlights the critical role of historical pricing in shaping current trends.

Seasonal effects are evident, with significant price drops in August (-3.90, CI: [-5.66, -2.14]) and September (-1.93, CI: [-3.51, -0.35]). These declines may reflect seasonal sales or shifts in

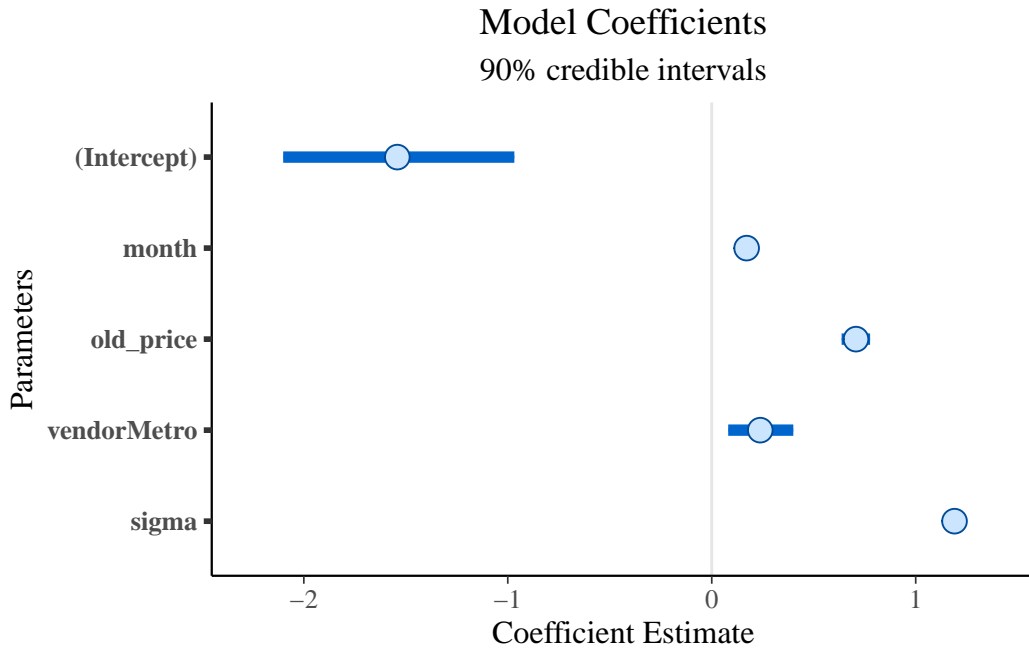


Figure 7: The 90% credible intervals for all model coefficients

demand. In contrast, July shows negligible effects, with a wide confidence interval overlapping zero.

While vendor-specific effects are encoded, they are not directly analyzed here. However, they likely contribute to market variability and warrant further exploration to understand competitive pricing dynamics.

In summary, the model underscores the importance of historical pricing and seasonality in determining current prices. Businesses should leverage these insights to optimize pricing strategies, particularly during months with significant declines like August and September. Further exploration of vendor-specific effects and additional variables could enhance model accuracy and strategic decision-making.

## 5 Discussion

### 5.1 Pricing Dynamics under Different Vendor Strategies

The analysis of vanilla ice cream pricing reveals significant differences in the strategies adopted by Metro and Loblaw's, reflecting their distinct brand identities and market positioning. Metro consistently employs a premium pricing strategy, maintaining higher prices that appeal to consumers who prioritize perceived quality and brand prestige. This approach aligns with Metro's reputation as a higher-end retailer, catering to customers willing to pay a premium for quality assurance. By consistently pricing above its competitor, Metro leverages its brand identity to reinforce customer perceptions of exclusivity and quality.

Loblaw's, on the other hand, adopts a more competitive pricing strategy, offering lower and more stable prices. This strategy appeals to budget-conscious consumers and those seeking consistent value for their purchases. By focusing on affordability, Loblaw's strengthens its position as a cost-effective option for shoppers, particularly during periods of high demand, such as the summer months when ice cream sales typically peak. The predictability of Loblaw's pricing helps build consumer trust and fosters loyalty among price-sensitive segments.

A notable finding is the correlation between current and historical prices, suggesting that both vendors use past pricing as a reference point for setting future prices. Metro appears to rely on historical pricing to maintain its premium market positioning, ensuring that promotions or discounts do not undermine the perceived value of its products. Conversely, Loblaw's uses historical pricing as a baseline for offering competitive deals while preserving its reputation for affordability.

These insights highlight the critical role of pricing strategies in shaping consumer perceptions and driving purchasing decisions. Retailers can draw lessons from these approaches to tailor their strategies in ways that align with brand identity, target demographics, and market dynamics. By understanding how different pricing models impact consumer behavior, vendors can refine their approaches to enhance customer satisfaction, loyalty, and competitive advantage.

### 5.2 Pricing Dynamics under Consumer Behaviour Caused by Seasonal Trends

Seasonality plays a pivotal role in shaping the pricing strategies for vanilla ice cream, with significant price variations observed during the summer months. These months, particularly June through August, coincide with the peak demand for ice cream, driven by warmer weather and increased outdoor activities. Both Metro and Loblaw's adjust their pricing during this period to capitalize on heightened consumer interest. Metro maintains its premium pricing even during these high-demand months, positioning itself as a retailer for customers who prioritize quality over cost. This approach ensures steady revenue while preserving the brand's image of exclusivity.

In contrast, Loblaws adopts a more aggressive pricing strategy during summer, offering competitive deals to attract a larger customer base. By strategically lowering prices or introducing promotions, Loblaws not only boosts seasonal sales but also enhances its appeal to price-sensitive shoppers. This strategy is particularly effective in driving volume sales, making it a key revenue generator during high-demand periods.

An important insight from this analysis is the interplay between seasonal demand and vendor pricing strategies. The increase in ice cream prices during summer reflects retailers' efforts to maximize revenue without alienating their target audiences. While Metro leverages consistent premium pricing to attract high-value customers, Loblaws' competitive pricing taps into broader consumer segments, including families and budget-conscious individuals.

These findings underscore the importance of aligning pricing strategies with seasonal consumer behavior. Retailers can enhance revenue opportunities by anticipating demand fluctuations and tailoring promotions or price adjustments accordingly. By understanding how seasonality influences purchasing decisions, vendors can refine their strategies to balance short-term revenue goals with long-term brand loyalty. Moreover, these insights offer actionable lessons for other seasonal markets, such as beverages or holiday-specific products, emphasizing the broader applicability of seasonality in retail pricing.

### **5.3 Real-Life Implications**

The findings of this paper extend beyond the realm of ice cream pricing. They underscore the broader economic principle of demand-driven pricing and its implications for consumer goods. Businesses in seasonal industries, such as beverages, holiday decorations, or travel services, can adopt similar analytical frameworks to optimize pricing and maximize revenue during peak periods. Moreover, the study illustrates how data-driven decision-making can enhance operational efficiency and market responsiveness, serving as a valuable lesson for industries adapting to consumer-centric strategies.

### **5.4 Weaknesses and Future Research Directions**

While this paper provides valuable insights into pricing dynamics and seasonal trends, there are limitations that merit discussion. First, the analysis is constrained to vanilla ice cream and two vendors, which may not fully capture the diversity of pricing strategies in the broader food retail sector. Expanding the dataset to include additional products and vendors could provide a more comprehensive understanding.

Second, the study focuses on pricing data without incorporating sales volume or profitability metrics, which are critical to understanding the effectiveness of vendor strategies. Future research could integrate these dimensions to assess the financial implications of pricing decisions.

Finally, while seasonal trends are examined, the underlying drivers, such as marketing efforts, regional climate variations, and socioeconomic factors, remain unexplored. Further studies could investigate these elements to provide a richer context for the observed trends.

Future research should also explore consumer preferences and purchasing behaviors through surveys or real-time purchase data. Combining quantitative pricing data with qualitative insights could yield a more nuanced understanding of the interplay between vendor strategies, consumer behavior, and market outcomes. Such research would not only validate the findings of this study but also uncover new opportunities for strategic decision-making in retail management.

## Appendix

### A Model details

#### A.1 Posterior predictive check

In Figure 8a we implement a posterior predictive check. The first graph illustrates a posterior predictive check, comparing the observed prices of vanilla ice cream with the replicated data generated by the model. The observed data and the overlaid replicated distributions align well, showing that the model captures the variability and trends in the data effectively. This suggests that the predictions are robust and reflective of the true pricing dynamics.

In Figure 8b we compare the posterior with the prior. The second graph compares the posterior distribution with the prior distribution for the model parameters. The posterior distribution is narrower and shifted compared to the prior, indicating that the data strongly influenced and refined the parameter estimates. This alignment demonstrates that the model effectively updates prior beliefs based on evidence, leading to more informed conclusions about price variations.

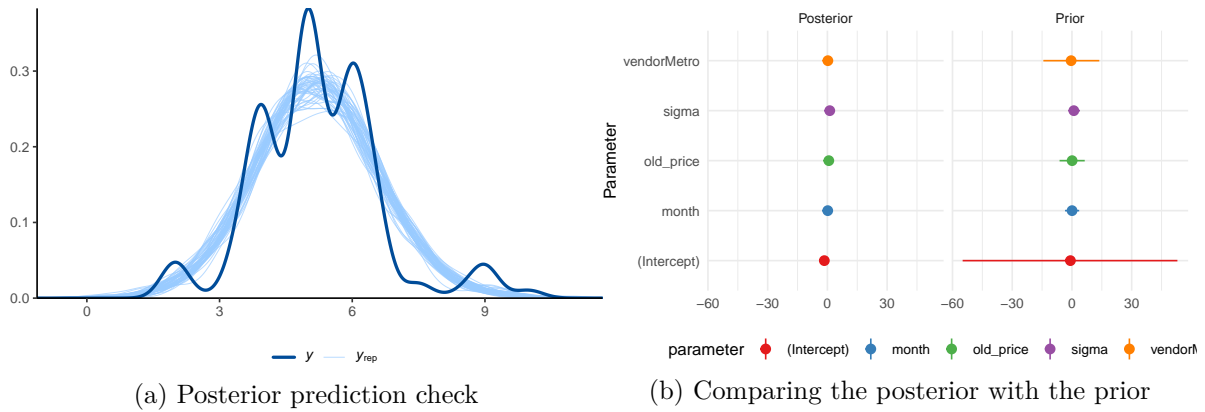


Figure 8: Examining how the model fits, and is affected by, the data

#### A.2 Diagnostics

Figure 9a is a trace plot. The first graph presents a trace plot, showing the evolution of sampled parameter values over iterations for each MCMC chain. Each chain demonstrates a “hairy caterpillar” appearance, indicating thorough exploration of the parameter space and likely convergence. The consistency across chains suggests that the model’s sampling process has adequately captured the posterior distributions for each parameter, reinforcing the reliability of the results.

Figure 9b is a Rhat plot. The second graph illustrates the Rhat diagnostic for various parameters. All Rhat values are below the critical threshold of 1.1, signifying that the chains have mixed well and converged to a common distribution. This indicates reliable posterior estimates and supports the validity of the model inference.

Together, these diagnostics confirm that the MCMC algorithm has performed effectively, ensuring trustworthy parameter estimates for the analysis.

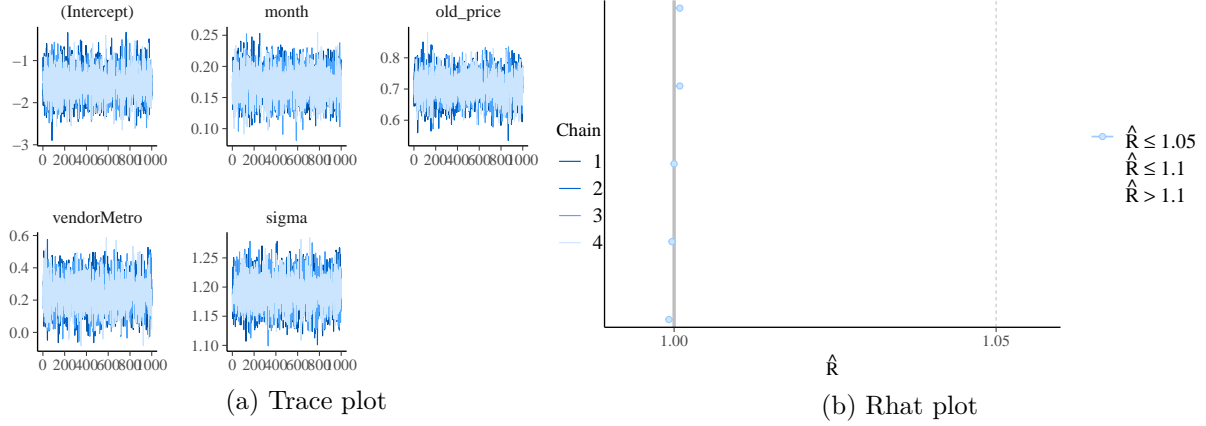


Figure 9: Checking the convergence of the MCMC algorithm



## References

- Apache Arrow. 2021. *Arrow: Integration to 'Apache' 'Arrow'*. <https://CRAN.R-project.org/package=arrow>.
- Arel-Bundock, Vincent. 2023. *Modelsummary: Regression Tables, Figures, and Other Results*. <https://CRAN.R-project.org/package=modelsummary>.
- Eddelbuettel, Dirk, and Romain Francois. 2011. “Rcpp: Seamless r and c++ Integration.” *Journal of Statistical Software* 40 (8): 1–18. <https://doi.org/10.18637/jss.v040.i08>.
- Filipp, Jacob. 2020. “Sync Outlook Calendar to Google Calendar Using Microsoft Power Automate.” <https://jacobfilipp.com/hammer/>.
- Firke, Sam. 2021. *Janitor: Simple Tools for Examining and Cleaning Dirty Data*. <https://CRAN.R-project.org/package=janitor>.
- Gelman, Gabriel, Jonah Gabry, et al. 2021. *Bayesplot: Plotting for Bayesian Models*. <https://mc-stan.org/bayesplot>.
- Goodrich, Ben, Jonah Gabry, Imad Ali, and Sam Brilleman. 2022. “rstanarm: Bayesian applied regression modeling via Stan.” <https://mc-stan.org/rstanarm/>.
- Grolemund, Garrett, and Hadley Wickham. 2011. *Lubridate: Make Dealing with Dates a Little Easier*. <https://CRAN.R-project.org/package=lubridate>.
- Horst, Allison Marie, Alison Presmanes Hill, and Kristen B Gorman. 2020. *palmerpenguins: Palmer Archipelago (Antarctica) penguin data*. <https://doi.org/10.5281/zenodo.3960218>.
- Martin, Julie Anne. 2018. “Consumer Preference of Vanilla Ice Cream.”
- R Core Team. 2023. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. <https://www.R-project.org/>.
- Sharpe, ANTHONY N, PEARL I Peterkin, and ISTVAN Dudas. 1979. “Membrane Filtration of Food Suspensions.” *Applied and Environmental Microbiology* 37 (1): 21–35.
- Trestle Technology, LLC, and Barret Schloerke. 2023. *Plumber: An API Generator for r*. <https://cran.r-project.org/package=plumber>.
- Wickham, Hadley. 2021. *Ggplot2: Create Elegant Data Visualisations Using the Grammar of Graphics*. <https://ggplot2.tidyverse.org>.
- Wickham, Hadley et al. 2021. *Tidyverse: Easily Install and Load the 'Tidyverse'*. <https://tidyverse.tidyverse.org>.
- Wickham, Hadley, Jim Hester, and Gabor Csardi. 2023. *Testthat: Unit Testing for r*. <https://CRAN.R-project.org/package=testthat>.
- Xie, Yihui. 2021. *Knitr: A General-Purpose Package for Dynamic Report Generation in r*. <https://yihui.org/knitr/>.