MARKETING ANALYTICS REPORT

CLIENT: NATA SUPERMARKET

by

Group K

Members:

Manan Agarwal

Poonam Kolur Math

Ishani Sumant Palwanker

Mansi Gupta

Avani Kothari

TABLE OF CONTENTS

PART ONE **4**
MARKET ANALYSIS & CASE DESCRIPTION **4**
- Introduction to The Canadian Grocery Retail Environment 4
- Understanding the Canadian Consumer 4
- Evolving consumer behaviours 5
- The Role of Analytics in Competitive Advantage 5
- Challenges at Nata Supermarkets 5
- Future Direction 6

PART TWO **7**
DATA DESCRIPTION **7**
- First Glimpse Data Observations 7
- Data Preparation 7
  - Program configuration 7
  - Initial inspection and data import 7
  - Managing Missing Values and Recoding Data 8
  - Standardising Numeric Data and Outlier Detection 8
  - Final Notes and Dataset Verification 9
- Descriptive Analysis 9
  - Overall Spending Distribution among Product Categories 9
  - Income as the Motive of Total Consumption 10
  - Spending Practices and Education 11
  - Marriage, Children, and Total Spending 12
  - Age and Product Category 12
  - Total Spending and Engagement Channels 13
  - Relationships among campaigns, products, and channels 14
  - Multiple Variables: Scatterplot Matrix 15
  - Takeaways from Descriptive Analysis 15
- Clustering and Customer Segmentation 15
  - Cluster Interpretation and Customer Profiling 16
  - Important Strategic Notes 17

PART THREE **18**
MAIN ANALYSIS **18**
- Analysing Linear Regression Models 18
  - Drivers of Total Spending 18
  - Category-Level Spending Forecasters 19
  - Strategic Insight Overall 20
- Analysing Logistical Regression Models 21

PART FOUR **24**
DECISIONS & RECOMMENDATIONS **24**

APPENDIX **25**
REFERENCES 28

PART ONE

MARKET ANALYSIS & CASE DESCRIPTION

Introduction to The Canadian Grocery Retail Environment

Businesses today are no longer constrained by generic market trends. With over 400 million terabytes of data created daily (Duarte, 2025), organisations that extract actionable insights are gaining a decisive competitive edge. This is especially true in Canada's grocery market — a CAD 159.9 billion industry in 2023, projected to reach CAD 191.5 billion by 2028.

Nata Supermarkets, founded in 1972 in Brandon, Manitoba, now operates 37 stores nationwide. However, internal reports show declining performance relative to competitors (MarketLine, 2024). The Canadian food retail sector is intensely competitive, with low margins and high price sensitivity (Statista, 2024). Market leaders like Loblaw (29%), Sobeys/Safeway (21%), Costco (11%), Metro/Jean Coutu (10.8%), and Walmart (7%) leverage advanced analytics to optimise stock, pricing, and loyalty programs — advantages that chains like Nata struggle to replicate (MarketLine, 2024).

Understanding the Canadian Consumer

Canada's grocery retail market is shaped by demographic and behavioural trends. Young males aged 14–18 have the highest annual food expenditure at CAD 4,456, with spending remaining strong among 19–30-year-olds before declining in older age groups. Male consumers typically outspend females across most age bands. While granular data on marital status or education is lacking, online shopping trends suggest a tech-savvy, educated base, likely comprising younger singles and smaller households. Income influences behaviour, but price sensitivity remains high across all income levels, with affordability being the top decision-making factor (Statista, 2024).

Evolving consumer behaviours

Despite the rise of e-commerce, 80% of Canadians still prefer supermarkets for groceries (Statista, 2024). Yet, major players like Walmart.ca, Costco.ca, and Amazon.ca are expanding online, targeting younger, convenience-driven consumers via services like Instacart. Online purchases focus on both necessity and indulgence items, with top products including coffee, snacks, fruits, vegetables, rice, and soft drinks.

Price sensitivity shapes not only what consumers buy but also how they shop. About 43.3% seek more promotions, 34.6% use more coupons, and 33.6% rely on loyalty programs. Shopping frequency remains high, with 44% visiting stores weekly and 37% two to three times per week. Moreover, 30.6% switch stores for better deals, and 26.6% shop using special offer coupons.

Nearly half of Canadian shoppers cite price as their top store selection factor, followed by stock availability and product freshness (Statista, 2024).

The Role of Analytics in Competitive Advantage

Case studies show that analytics is redefining success in grocery retail. Walmart, an early adopter, uses machine learning to forecast checkout demand, optimise pharmacy operations, and personalise marketing (Walmart Canada, 2024). Its loyalty program, Walmart+, reduces churn through predictive insights. Loblaw leverages loyalty data to enhance e-commerce engagement, while Sobeys and FreshCo use analytics for regional sourcing and customer responsiveness. Metro is shifting toward wholesale and e-commerce growth, guided by consumer insights (MarketLine, 2024). These examples highlight the clear and profitable value of marketing analytics across Canada's top retailers.

Challenges at Nata Supermarkets

Presently, there are two major challenges with Nata Supermarkets, as identified by the Vice President of Technology Vina Verago. Firstly, poor targeting through promotions and poor stock management of products (MarketLine, 2024). Compared to competitors, who have leveraged data analysis to predict how customers will behave and modify offers accordingly, Nata has historically been unable to follow this trend. It has failed to deeply understand its customers' preferences and forecast their needs.

Future Direction

Nata Supermarkets must leverage its consumer data to remain viable. With the sector projected to grow at a steady 3.7% CAGR through 2028 (MarketLine, 2024), maintaining the status quo is not an option. By analysing customer demographics, preferences, and purchase patterns, Nata can inform smarter promotions, optimise inventory, and elevate the shopping experience.

As competitors adopt data-driven strategies to deliver personalised, omnichannel experiences, Nata must embed analytics into its core operations. This report outlines how Nata can transform its data into actionable insights through segmentation, predictive modelling, and visualisation, positioning the company for sustainable growth in Canada's evolving grocery landscape.

PART TWO

DATA DESCRIPTION

First Glimpse Data Observations

Nata's dataset includes demographic and purchasing data for 2,240 customers, covering birth year, marital status, income, product category spend, and responses to six campaigns. An initial review highlights variables worth exploring for potential correlations and helps identify irrelevant columns.

Income and age may predict purchasing behaviour, but this requires validation through statistical testing. Early identification of high-impact variables, those with strong profit potential or demand, will help focus the analysis. Beyond continuous variables, campaign acceptance data offers valuable insights for future promotional planning.

Categorical variables like education, marital status, and campaign outcomes must be encoded as factors. To begin exploratory and descriptive analysis, the dataset will be loaded and prepared in R.

Data Preparation

The first step of data preparation is understanding its structure, flaws, and then cleaning it thoroughly. Using best practices in data preparation for consumer segmentation and behavioural modelling, we handled missing values, inconsistent formatting, outliers, and irrelevant variables.

Program configuration

For this project, we installed crucial packages from the Comprehensive R Archive Network (CRAN). Packages such as `tidyverse` for data manipulation, `corrplot` for correlation visualisation, `factoextra` for clustering analysis, `gpairs` for scatterplot matrix visualisations, `psych`, `coefplot`, and `summarytools` were used in R Studio to handle data cleansing. Preparing the dataset for exploratory research, modelling, and visualisation required these instruments very definitely.

Initial inspection and data import

The dataset was imported using read_csv() and explored with glimpse() and the first 50 rows. This revealed issues such as incorrect data types (especially for categorical and ordinal variables), NA values, and inconsistencies in column structure.

To improve clarity, several columns were renamed. E.g., ID became CustomerID, Dt_Customer became DateJoined, and campaign columns like AcceptedCmp1 and Response were relabelled for consistency. This streamlined further analysis and improved team readability.

Running summary() exposed key data issues: an unrealistic maximum age of 132, an outlier income value of 666666, 24 NA entries in the income column, and improperly formatted dates.

Managing Missing Values and Recoding Data

Next, we identified NA values in the Income column using colSums(is.na()) and removed those rows, as income is a key predictor and imputation could introduce bias. Categorical columns like Education and Marital_Status were inspected using unique(). Marital status was simplified to ‘Single’ or ‘Together’, encoded as 0 and 1. Invalid entries like “YOLO” and “Absurd” were removed.

Education was converted into an ordered factor reflecting hierarchy (“Basic” to “PhD”), and binary variables like campaign responses and complaints were transformed into labelled factors for clarity and model performance.

Columns Z_CostContact and Z_Revenue were dropped due to zero variance. A new Age column was created by subtracting Year_Birth from 2025. Entries with ages below 18 or above 90 were excluded as likely input errors or outliers.

Standardising Numeric Data and Outlier Detection

The `scale()` function helped to standardise all pertinent numerical variables (apart from household children counts, customer ID, and birth year). Comparability across variables depends on standardising, particularly for methods sensitive to scale, like clustering and principal component analysis.

First, outlier detection was done with boxplots of standardised variables produced with `ggplot2`. Filtered were extreme income values over ±3 standard deviations from the mean. One especially found and eliminated, a very inaccurate income value of 66666. These changes improved the consistency and dependability of further modelling projects.

While outliers existed across product categories, there were no extreme outliers. Thereby, we decided to retain outliers within the spending categories. This decision is reflective of real-world consumer behaviour pattern. Purchasing behaviour seldom follows a normal distribution and extreme variation in some categories is expected.

Final Notes and Dataset Verification

Redundant columns, including `Year`, were discarded after all changes and cleaning activities. Using `dim()` and `summary()`, the last dataset was checked to verify data integrity and shape. Post-outlier removal of the cleaned data was also re-standardised and visualised once again to verify normal distribution and absence of severe abnormalities.

We avoided winsorizing outliers and missing values (which is often opted for) to avoid creating perfect models that may not reflect real-world consumer patterns.

This comprehensive data cleaning technique produces a structured, noise-reducing, analysis-ready dataset.

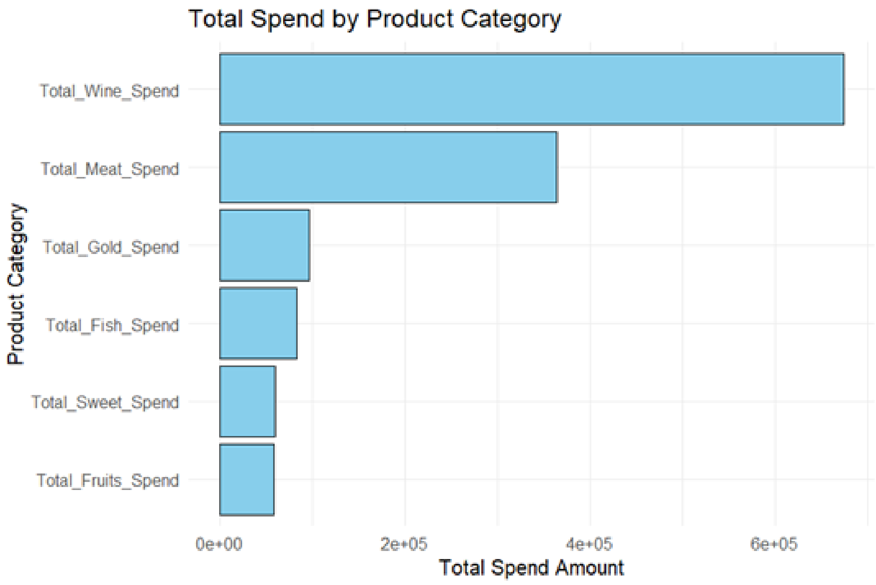
Descriptive Analysis

The initial descriptive analysis of Nata Supermarkets' cleaned dataset reveals a customer base that is predominantly middle-aged to senior, with an average age of 56 and a high level of education—over half hold a university degree or higher. Most customers have no children or teenagers at home, suggesting fewer household obligations and potential for discretionary spending. Income levels vary widely, with a mean of approximately CAD 51,606, reinforcing the importance of income-based segmentation in marketing strategies. In terms of shopping behaviour, store purchases remain the most common channel, followed by web and catalogue purchases, with some customers showing high engagement across all three.

Overall Spending Distribution among Product Categories

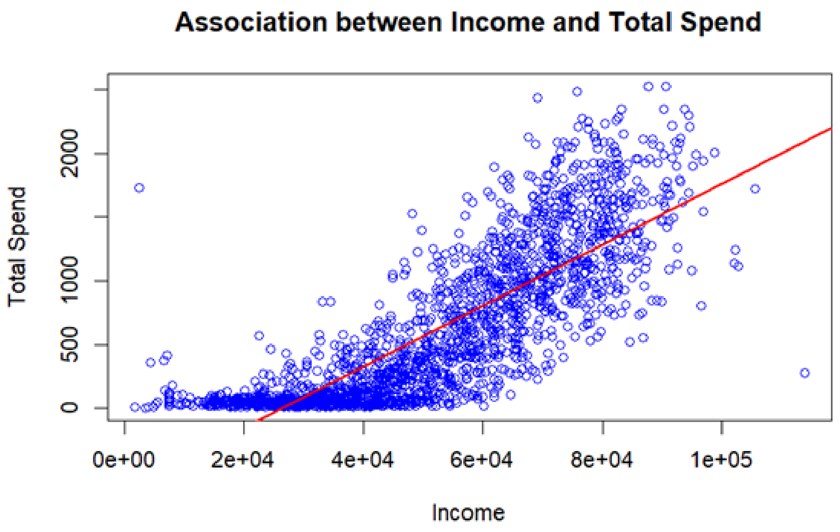
A barplot was used to visualise overall product-wise spending, as it effectively compares total values across product categories. Computed by summing up six expenditure categories — wine, meat, fish, fruits, sweets, and gold products. By total spend, wines ranked highest among all the categories, followed by meat goods and gold objects. On the other hand, seafood, sweets, and fruits were linked to reduced general expenditure.

For Nata, wines and meats generate the most sales, hence marketing initiatives and promotions should give these areas top priority. Furthermore, gold products, despite their niche, have great value and would profit from specific targeting techniques.



Income as the Motive of Total Consumption

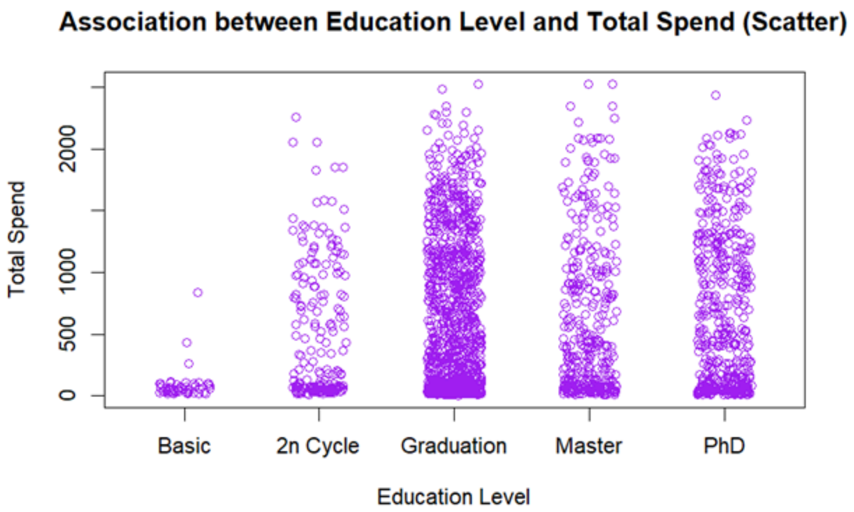
A scatterplot with a regression line was employed for income vs total spend to reveal the strength and direction of their linear relationship. Confirmed by a high correlation value (~0.82), a scatter plot between `Income` and `Total_spends`, a significant positive linear association was noted. With numerous outliers spending over CAD 2,000 yearly on groceries, the regression line fits very well, and high-income consumers were consistently linked with greater total spending.



For this dataset, income is the best indicator of overall expenditure. This result emphasises the need of focussing on high-income consumer groups by means of luxury products and tailored promotions.

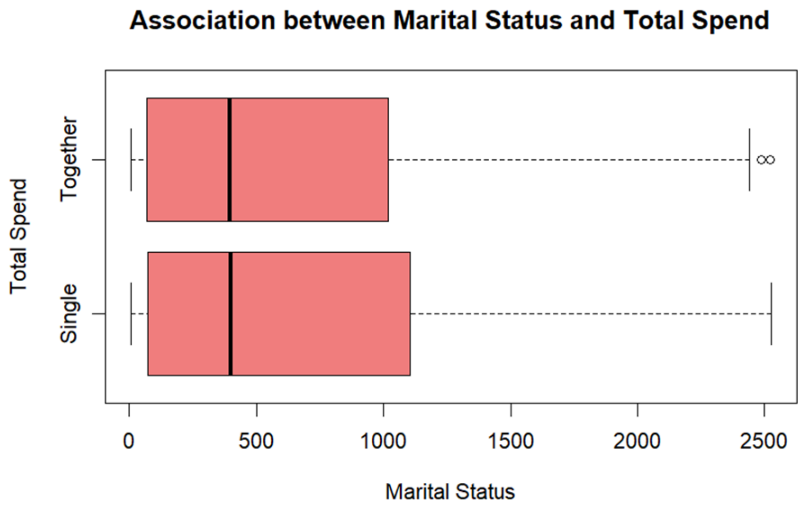
Spending Practices and Education

Jitterplot was chosen for education vs total spend to display individual spending variations across overlapping educational categories. Customers with basic education demonstrated regularly low spending levels according to the jitter plot between `Education` and `Total_Spend`. Those with graduate, master's, or PhD degrees showed both larger and more erratic purchasing behaviours, usually ranging from CAD 500 to CAD 2,000.

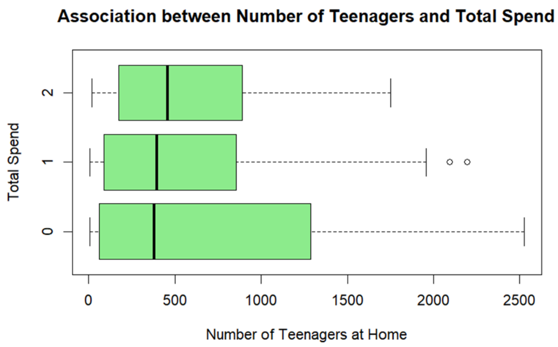


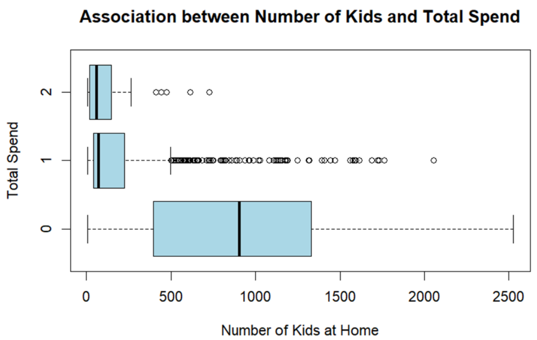
Marriage, Children, and Total Spending

Boxplots for marital status and age vs product spend were used to compare spending distributions across categorical groups while highlighting medians and outliers. Notable patterns were shown by boxplots of `Marital_Status`, `Kidhome`, and `Teenhome` versus `Total_Spend`. Couples and singles spent somewhat similarly statistically. Singles tend to spend slightly more.



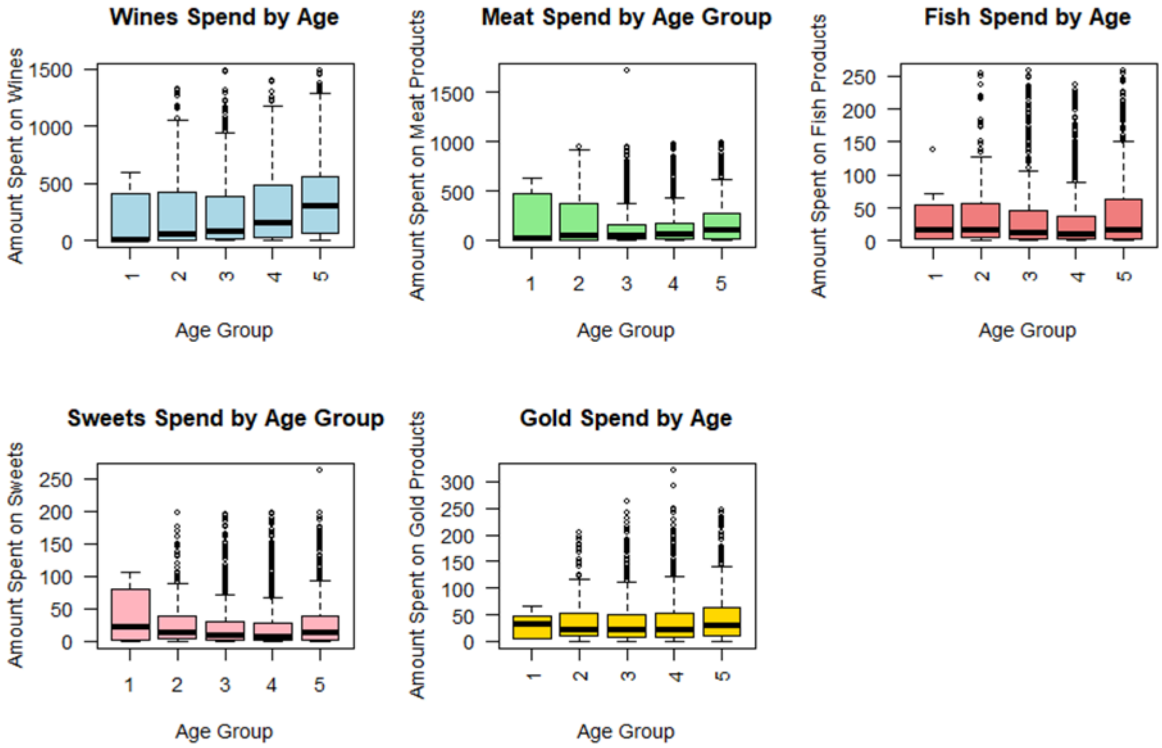
Households without children exhibited the largest spending ranges; those with one or two children showed a significant median and maximum expenditure drop. Adolescents (Teenhome) showed a similar trend, with consumers without adolescents paying more.





Age and Product Category

Five categories were created from age groups: 18–30, 31–40, 41–50, 51–60, and 60+. When product expenditure was broken up by age group:



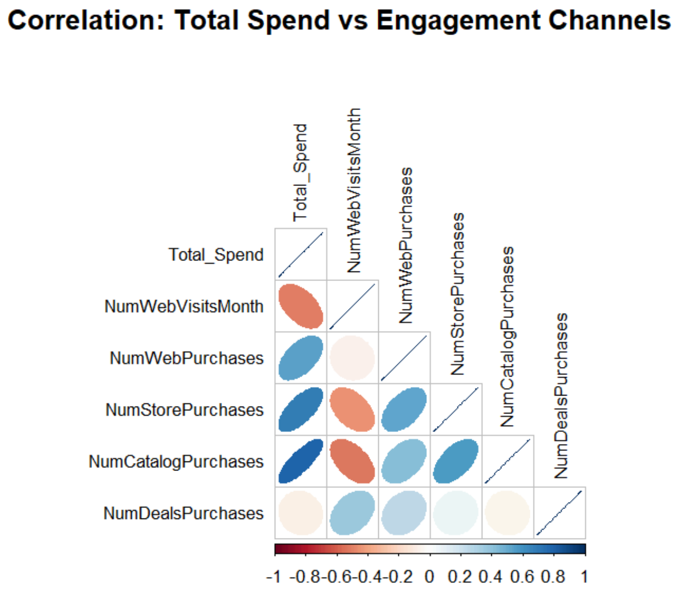
Wines: Spending was constant across all categories; the younger (18–30) and oldest (60+) segments showed modest variation and large outliers.

Meat: Spending fell with age; it peaked in younger groups.

Fish, sweets, and gold: These categories show constant but modest expenditure across all age groups; younger consumers particularly like sweets.

Total Spending and Engagement Channels

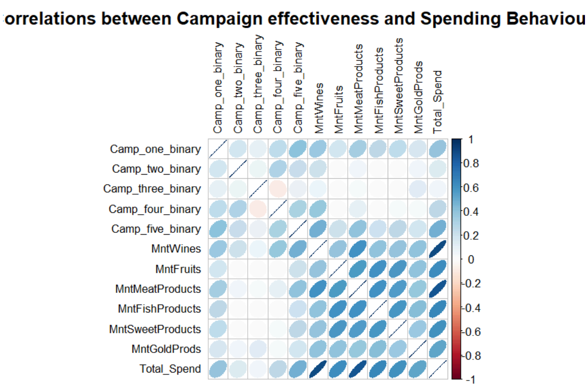
Correlation heatmap was used to analyze total spend vs engagement channels, as it efficiently visualizes multiple pairwise relationships and their strength in a single view. Using correlation analysis, we found that Total_Spend showed the most positive correlations with Catalog purchases, web purchases and store purchases. On the other hand, the number of website visits revealed a poor or even negative association with expenditure and web purchases, implying that visits by themselves do not convert to sales.

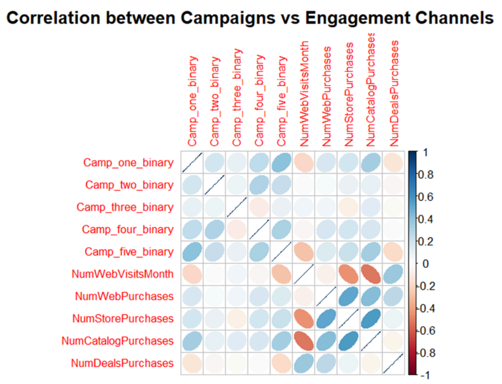


Relationships among campaigns, products, and channels

Two focused heatmaps investigated how campaign acceptability related to both product expenditure and interaction behaviour.

Campaigns 1 was most favourably linked, according to product spending correlations, with wine and meat purchases — the top-spending categories. It also has the most total spend.

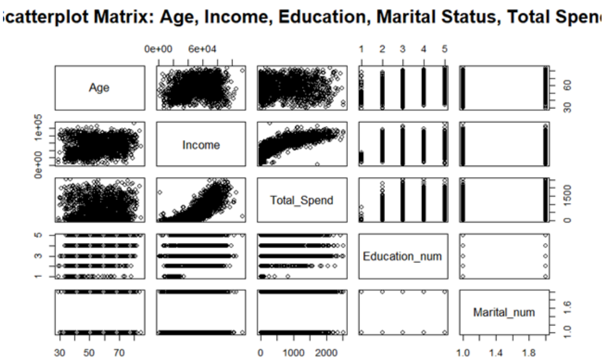


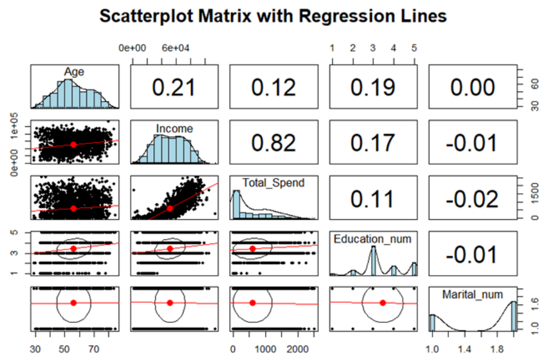


Most Campaigns performed better leading to more Catalog Purchases as compared to other channels. 1 and 5 linked favourably with catalogue and in-store purchases but not with online visitors according engagement channel analysis.

Multiple Variables: Scatterplot Matrix

A hollistic picture was given by a scatterplot matrix including age, income, total expenditure, education (numerically encoded), and marital status. Income was the sole variable displaying a clear, positive linear connection with `Total_Spend`, consistent with past research. In this setting, marital status and education had little impact on prediction.





Takeaways from Descriptive Analysis

Income, product preferences, and channel engagement should be central to future marketing and targeting strategies at Nata Supermarkets, taking precedence over less impactful demographic variables such as marital status. The descriptive analysis confirms that income, education, and transactional engagement channels are the strongest drivers of customer expenditure. Campaign insights highlight a few standout promotions, particularly those aligned with high-value product categories like wines and meats. These products not only dominate overall spend but also correlate positively with successful campaigns and purchase channels like catalogues and in-store interactions. Targeting high-income, well-educated, and child-free customer segments, who consistently display higher spending, should be a key focus for segmentation and personalised marketing.

Clustering and Customer Segmentation

After the data was completely cleansed and standardised, the study concentrated on k-means cluster-based consumer segmentation of Nata Supermarkets. The goal was to separate consumers according to their demographic traits, purchasing behaviour, and interaction style. This division offers a useful basis for tailored marketing plans and resource distribution.

First standardising the dataset helped to guarantee that every numerical variable equally affected the distance computations. After that, the scaled data was subjected to the k-means algorithm using elbow technique to ascertain the ideal number of clusters. Across cluster counts ranging from 1 to 10, the within-cluster sum of squares (WSS) was computed. With an elbow plot showing an inflection point at k = 3, three clusters seemed to be a reasonable choice.

K = 3 was verified as the final cluster count after testing cluster cohesiveness and separation employing the `fviz_cluster()` visualisations. The visualisation confirmed three distinct clusters with the least overlap, tightly packed data points, and high similarity within the clusters. These clusters were then included in the dataset for comparison and interpretation.

Cluster Interpretation and Customer Profiling

Cluster 1: Value, High Engagement Consumers

Particularly in wines, meat, and gold items, this cluster is smaller in size but has the greatest overall spend per consumer. Across all categories, these clients often have the greatest earnings; they are also older, more educated. Their buying patterns show a luxury lifestyle. They are more likely to accept campaigns (especially Campaigns 1, 2, and 5) and participate across several channels, mostly catalogue and retail purchases. Usually devoid of children at home, they have more free will money. Among Nata's segments, this one is both the most profitable and strategically significant. For well-targeted advertising, these clients are quite responsive and involved. They ought to be given priority for VIP service, loyalty programs, special discounts, and premium products. Offering customised product suggestions and continuing to promote premium categories like wine and gold can help to maximise campaign results.

Cluster 2: Underengaged Low-Income, Low-Spending

Customers with below-average earnings and lowest overall expenditure fall into Cluster 3. Their replies to all ads are small and their educational background usually falls on the lower side. They prefer less retail visits and are mainly alienated from Nata's marketing channels — often with little or no internet usage. They are more likely to share bigger homes with teens and children, which might limit spending capability.

This sector presents a difficulty for profitability as it is cost-sensitive and hard to include. Although they are not perfect objectives for high-end marketing, budget-friendly promotions, combined offers, and requirements discounts can help to meet them. Coupons or family-oriented marketing can inspire involvement, therefore improving their buying frequency and general value.

Cluster 3: The core middle spenders

The biggest share of Nata's clientele consists of this group. These people are middle-aged, generally, with modest earnings and average overall expenditure across most product categories. Although they are not very active online, their engagement behaviour consists of a sensible balance of store and catalogue purchases. There is an average response rate for campaigns; acceptance does not show any significant surge. Their household size usually consists of one or two children or teens, implying that family responsibilities may affect their expenditure balance. This group captures Nata's basic clientele: steady, reliable, loyal. Marketing plans should centre on keeping consumers happy, providing value-driven specials, and consistent stock across basic categories such as meat and home goods.

Important Strategic Notes

The cluster study shows that Nata Supermarkets' consumer base is split into three useful categories. While the low-spend section (Cluster 2) requires careful, low-cost involvement tactics, the medium core (Cluster 3) forms the operational backbone, and the premium segment (Cluster 1) gives the best returns should be fostered.

Consistent with past descriptive results, income once more appeared as the most significant factor influencing total consumption across all clusters.

PART THREE

MAIN ANALYSIS

Analysing Linear Regression Models

Following the identification of important consumer segments using clustering, linear regression models were used to investigate the factors influencing spending behaviour, both generally and within particular expenditure categories. These models provide predictive analysis of how factors, including income, product cross-spending, and channels of interaction, affect consumer purchases. Customised marketing plans depending on channel, product category, and consumer profile are built on the results.

Drivers of Total Spending

Starting with income and purchase counts across catalogue, web, and in-store channels as predictors, the first multiple regression model evaluated total expenditure as the outcome variable. With total cost, all predictors were strongly and favourably linked.

With coefficients around **0.40**, income and catalogue purchases were the most powerful factors. This suggests, given other factors constant, a one-unit increase in each predictor almost leads to a 0.40 unit rise in overall expenditure. Though with lesser coefficients (**0.096 and 0.13** respectively), web and shop purchases were still quite important. With a p-value of 2.2e-16 and $R^2 = 0.7899$, this model is statistically significant.

Business Implications: Sales increases cannot be driven without high-income consumers.
The great return on investment of catalogue interaction emphasises the importance of revived catalogue marketing.

Using omnichannel techniques (online, retail, and catalogue) together will help to generate small sales incrementally.

A scenario simulation proved the practical usage of the model: a 500-unit income increase and five more purchases in each channel projected a total spending increase of over 200 units, therefore proving clear value in multichannel engagement and income-based targeting.

Category-Level Spending Forecasters

Wine Products

Income (β = 0.31) and catalogue purchases (β = 0.24) drove wine expenditure most substantially; online (β = 0.19), shop (β = 0.16), and meat product spending (β = 0.05) followed. Indicating high predictive ability, the model explained 63% of the variation in wine spending ($R^2 = 0.633$).

The implication from this regression is clear. Richer consumers find more appeal in wine. For this category, catalogue specials are quite successful. Another insight uncovered was that cross-selling meat items might improve purchasing baskets.

A simulation showed that moderate increases in income by $500 and purchases across categories will increase wine expenditure by over CAD 160, therefore confirming wine as a high-opportunity commodity with wealthy, multi-channel consumers.

Food Products

Spending on fish (β = 0.26), sweet (β = 0.23), and meat goods (β = 0.18) most affected fruit expenditure. Income also had a (β = 0.065) but smaller influence. Store purchases showed a little beneficial effect; catalogue purchases were not noteworthy.

Cross-category marketing helps fruits, especially with regard to sweet products and seafood. The key is in-store specials; catalogue-based fruit campaigns might not be as successful.

When relevant variables were raised, the scenario analysis revealed a total expected rise of 70.5 units in fruit expenditure, therefore emphasising cross-category bundling as a main strategy.

Meat Products

Catalogue purchases (β = 0.31) followed by income (β = 0.24) were the best predictors of meat spending. Additionally, having significant favourable benefits were cross-spending on fish (0.14), fruits (β = 0.12), wines and sweets (β = ~0.07).

Business implication: Catalogue customers are quite meat-inclined. Therefore, this is the most practical avenue for meat promotions.

Fish Products

Fruit ($\beta = 0.23$), sweets ($\beta = 0.22$), meat ($\beta = 0.19$), and gold ($\beta = 0.11$) purchases drove fish expenditure. Though income and retail visits were not significant, catalogue sales were notable.

Fish items appeal more to active consumers in many different food categories than to more affluent consumers, particularly.

Sales may be greatly raised with catalogue marketing and bundling techniques (such as seafood plus sweet snacks or wine). In-store promotions might not have much effect on fish buying.

Gold Products

Fish product purchases ($\beta = 0.21$) and site and catalogue interaction (coefficients ~0.24 each) individually drove gold product purchasing. Income has no meaningful impact.

Purchases of gold seem channel and behavior-driven, instead of income-dependent. Especially for those already involved with fish, digital and catalog-exclusive promotions are ideal.
For this group, behavioural segmentation may prove more helpful than income segmentation.

Sweet Products

Fish ($\beta = 0.25$) and fruits ($\beta = 0.23$) most affected spending on sweets; meat, income, catalogue, and shop interaction had favourable effects as well. About half the variance in sweet consumption could be accounted by the model.

Sweets are great for cross-promotion as they go nicely with other fresh items. Though less so, higher-income consumers still spend more on sweets. For this category, both catalogue and in-store marketing are rather successful.

Strategic Insight Overall

The most constant determinants of consumer value are income, catalogue engagement, and cross-category product spending. Varied categories call for varied contributions from web and store channels; web is great for discretionary goods like gold, and in-store is better for basics like fruits.
Campaign and channel plans should be product-specific; no one-size-fits-all solution exists.
With this analytical basis, Nata Supermarkets can maximise return on marketing spend and customer lifetime value by moving from broad mass marketing to exact, data-based segmentation and targeting.

Analysing Logistical Regression Models

For every marketing campaign (Campaigns 1–6) as well as for customer complaints, logistic regression models were created to evaluate which consumer traits most affect campaign acceptance and complaint behaviour. These models use income, total expenditure, age, family composition, and marital status to forecast the probability of particular consumer actions. The study not only finds statistically significant predictors but also reveals their relative influence, therefore offering evidence-based suggestions for focused consumer involvement and marketing.

Campaign 1

According to this campaign, approval is rather significantly predicted by income as well as total consumer expenditure. More specifically, the likelihood of adopting the campaign rises by a ratio of 4.77 for every unit increase in scaled income; overall expenditure adds an odds increase of 2.05. With an AIC of 710.26, the model fit is robust; all of the predictors have great significance ($p = 0.001$).
Among high-income, high-spending consumers, Campaign One works best. These people probably view the advertising as premium or value-adding, and their overall buying behaviour corresponds with more thorough brand interaction.

Targeting high-value consumers utilising income segmentation and purchase history filters, this campaign should keep focusing especially on channels where these consumers are most engaged, including catalogues and luxury product categories.

Campaign 2

The Campaign 2 model utilised just overall expenditure as a predictor. With every unit increase in total expenditure double the odds of acceptance (odds ratio = 2.63), results revealed a robust, statistically significant link. With an AIC =286.56 the model showed a clear improvement above the null deviance, indicating a strong match.

Regardless of wealth, age, or family status, Campaign 2 does well with consumers who already show great transactional interaction with the brand. Especially recent or high-frequency purchasers, segment and retarget consumers based just on purchasing history, then combine follow-up offers to maximise repeat conversions.

Campaign 3

This yielded more subtle observations. Once more, total expenditure was a substantial positive predictor that raised almost two-fold (odds ratio = 1.95) acceptance probability. Still, wealth and age were negative indicators. With an odds ratio of 0.57 for income and 0.82 for age, younger, lower-income consumers are more likely to welcome this advertising.

Younger, frugal groups find more attraction in this marketing. Its language or service probably speaks to pragmatic necessities or lifestyle-specific preferences (e.g., discounts or digital-first capabilities). Use age and income data to focus campaigns on younger generations, therefore avoiding higher-income consumers who are less receptive.

Campaign 4

This campaign called for age, number of children, and overall expenditure. Once more, total expenditure proved to be the greatest predictor (odds ratio = 2.50), while both age and number of children were just marginally significant positive variables ($p < 0.1$ for children).

Consumers who make more, are older, and have children are somewhat more inclined to embrace this approach. This implies that, appealing to life-stage-specific demands, the ad could provide family-oriented value or convenience. Using store-level promotions and product bundling for household categories like snacks, beverages, and basics, highlight this campaign to older homes with children.

Campaign 5

This model took marital status, income, and number of children. While the number of children had a negative effect (odds ratio = 0.58), income had a large positive impact, greatly raising the likelihood of acceptance (odds ratio = 23.01). Having a relationship (Marital_Status = "Together") somewhat raised probabilities (odds ratio = 1.84).

Among wealthy, paired consumers without children, this approach works best. Larger or more budget-conscious homes might not find appeal in its design. Present this ad as a premium lifestyle choice best suitable for working couples. Unless value-based changes are made, avoid focusing on households having several dependents.

Campaign 6

This model included marital status, income, and expenditure as well. Fascinatingly, overall expenditure was once more highly predictive (odds ratio = 2.52), while income and marriage also exhibited major negative impacts (odds ratios = 0.73 and 0.42, respectively).

High-spending, single, lower-to- middle-income consumers will find most resonance in this ad. It might be more focused on personal preferences or value propositions.
Target independent, technologically engaged consumers in Refocus Campaign Six, then change messaging away from luxury or partnership-based framing.

The Complaint Model

We ran a different logistic regression to grasp complaint behaviour. Fascinatingly, earlier join dates also predicted fewer complaints; higher overall expenditure was linked to a decreased risk of complaints — though marginally significant. This implies that high-spend, longer-tenured consumers are less prone to show discontent.

Spend level and loyalty go hand in hand with satisfaction, or at least with reduced complaint probability. Watch more attentively younger or low-spending consumers for discontent. Early intervention based on loyalty programs and feedback surveys might help to lower the attrition from this at-risk population. Targeting based on total expenditure reliably predicted campaign acceptability across all logistic regression models, so this is the single most accurate behavioural indicator available. Income had two effects: it helped certain campaigns (1 and 5) and harmed others (3 and 6). Age, home composition, and marital status contributed another layer of categorisation. Using customised messaging and intelligent segmentation, these insights enable Nata Supermarkets to transition from generic campaign distribution to precision-targeted marketing, hence generating better ROI.

PART FOUR

DECISIONS & RECOMENDATIONS

Nata Supermarkets should prioritise a dual-segmentation strategy by targeting high-income, high-spend consumers (Cluster 1) with premium offerings (particularly wines, meats, and gold items). Furthermore, we recommend engaging middle-income, value-conscious customers (Cluster 3) with steady catalogue promotions and consistent stock availability. While Cluster 2 is less profitable, targeted family bundles, coupons, and in-store deals can increase their participation cost-effectively.

Campaigns 1, 2, and 5 showed the strongest potential and should be scaled strategically based on segment responsiveness. Campaign 1, in particular, should focus on premium channels and income-based targeting, while Campaigns 3 and 6 require repositioning toward younger, single, value-driven customers.

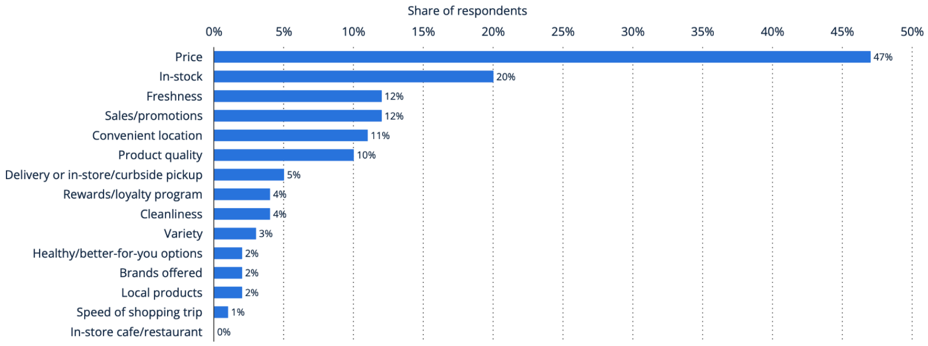
Marketing spend should shift from broad outreach to precision targeting, using income, spend history, and channel preferences as segmentation filters. Catalogue and in-store engagement remain critical drivers of revenue and should be enhanced, while passive channels like site visits need reevaluation.

Finally, campaign success and complaint prevention both correlate with customer spending and loyalty. Introducing a tiered loyalty program and leveraging predictive models to preempt churn will strengthen customer retention and ROI.

APPENDIX

How shoppers rank factors for choosing grocery stores in Canada as of 2023

How Canadian shoppers rank factors for choosing grocery stores 2023

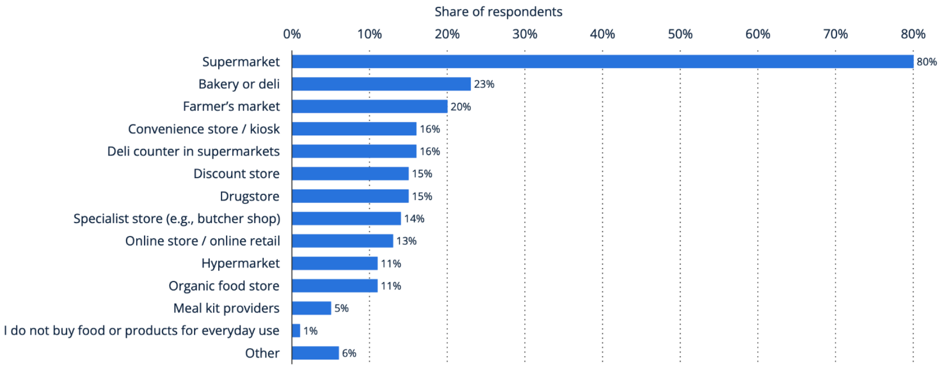


31 **Description:** As of 2023, Canadian consumers are most likely to shop at a grocery retailer based on its price, if products are in-stock, and how fresh the products are. Price is the top factor, however, with almost half of consumers saying this was the most important. Read more
Note(s): Canada; 2023; 1,000 respondents; Primary and shared-decision makers
Source(s): Canadian Grocer; EnsembleIQ

statista

Grocery shopping by store type in Canada as of December 2024

Grocery shopping by store type in Canada 2024

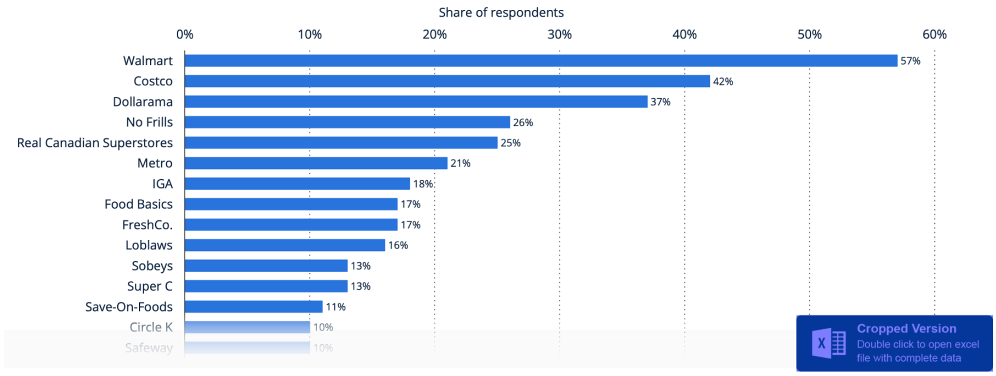


29 **Description:** "Supermarket" and "Bakery or deli" are the top two answers among Canadian consumers in our survey on the subject of "Grocery shopping by store type". Read more
Note(s): Canada; January to December 2024; 2022 respondents; 18-64 years

statista

Grocery shopping by store brand in Canada as of December 2024

Grocery shopping by store brand in Canada 2024

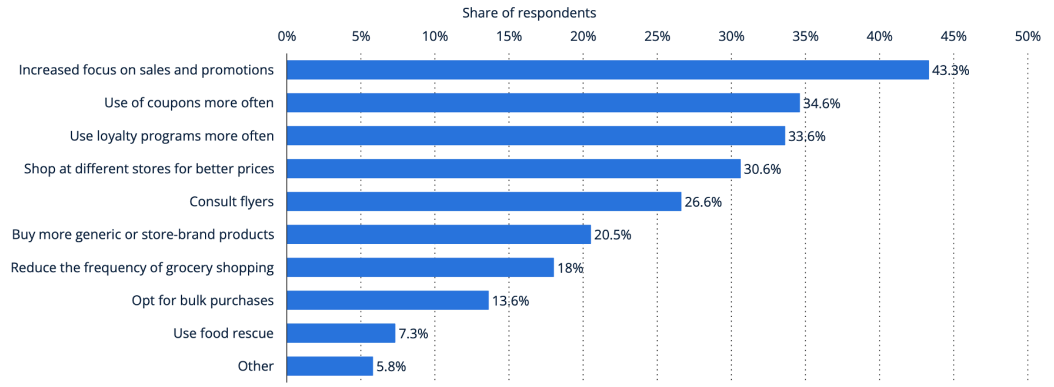


28 **Description:** We asked Canadian consumers about "Grocery shopping by store brand" and found that "Walmart" takes the top spot, while "Longo's" is at the other end of the ranking. Read more
Note(s): Canada; January to December 2024; 2022 respondents; 18-64 years
Source(s): Statista Consumer Insights

statista

How do you expect your grocery shopping habits to change in 2024 due to potential price fluctuations?

Canadians' change in food-buying habits due to future food price fluctuations 2024

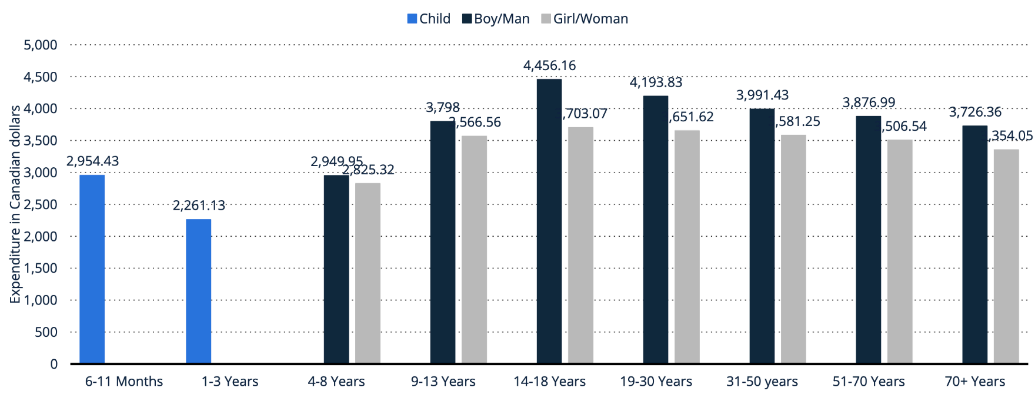


27 **Description:** According to a survey conducted in Canada in 2023, over 43 percent of respondents stated they would increase their focus on sales and promotions in the coming year to compensate for future grocery price fluctuations. Some 34.6 percent will chose to use coupons more often,while 33.6 percent with use loyalty programs more frequently. Read more
Note(s): Canada; 2023; 5,000 respondents
Source(s): Caddle; Dalhousie University

statista

Annual food expenditure in Canada in 2023, by age and gender (in Canadian dollars)

Annual food expenditure in Canada by age and gender 2023

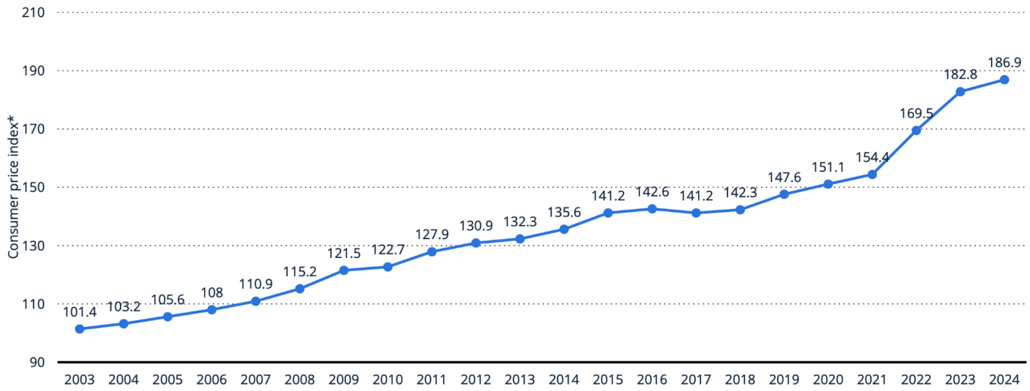


26 **Description:** In 2023, Canadian male consumers aged between 14 and 18 had the highest annual food expenditure among all categories, averaging about 4,456 Canadian dollars. Among the adult population, consumers over the age of 70 appeared to spend less than

statista

Annual consumer price index (CPI) of food purchased from stores in Canada from 2003 to 2024

Consumer price index of food purchased from stores in Canada 2003-2024

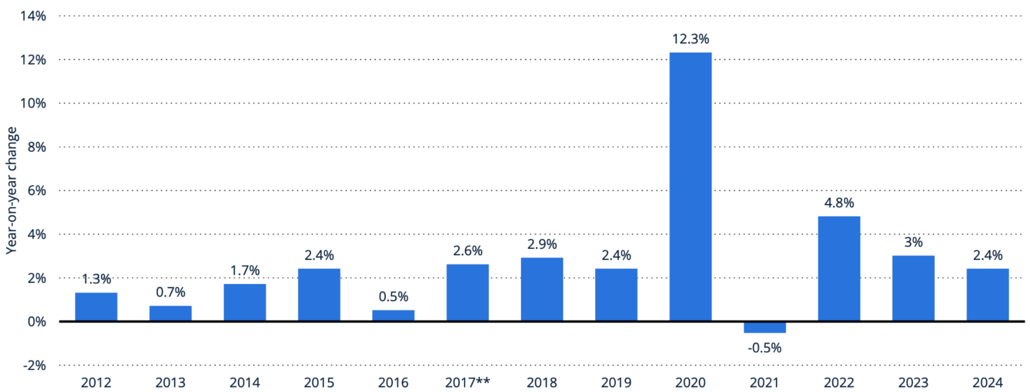


25 **Description:** This statistic shows the annual consumer price index of food purchased from stores in Canada from 2003 to 2024. The annual consumer price index for food purchased from stores was measured at 186.9 in 2024. Read more
Note(s): Canada; 2003 to 2024; *Base: 2002 = 100. The year 2002 is the base year used to calculate the index value.Adapted from Statistics Canada, statcan.gc.ca, 2003 to 2023. This does not constitute an endorsement by Statistics Canada of this [...] Read more
Source(s): StatCan

statista

Change in retail sales of food and beverage retailers in Canada from 2012 to 2024*

Change in retail sales of food and beverage stores in Canada 2012-2024



4 **Description:** Retail sales of food and beverage stores in Canada increased by 2.4 percent in December 2024 compared to December 2023. In 2023, retail sales of food and beverage stores in the country had increased by three percent compared to the same month a year earlier. Read more
Note(s): Canada; 2012 to 2024; *Seasonally adjusted. **As of 2017 the source stopped providing annual change figures, instead providing monthly change figures for December of each year. Data from 2012 to 2023 are taken from previous [...] Read more
Source(s): StatCan

statista

REFERENCES

DLT Labs Inc, 2019. *Walmart Canada and DLT Labs™ launch world's largest full production blockchain solution for industrial application* [Online]. *Newswire.ca*. CNW Group. Available from: https://www.newswire.ca/news-releases/walmart-canada-and-dlt-labs-tm-launch-world-s-largest-full-production-blockchain-solution-for-industrial-application-857165201.html [Accessed 23 April 2025].

Duarte, F., 2025. *Amount of Data Created Daily (2024)* [Online]. *Exploding Topics*. Available from: https://explodingtopics.com/blog/data-generated-per-day [Accessed 25 April 2025].

Khan, J., 2024. *Walmart Luminate arrives in Canada, bringing its ecosystem of insights-driven solutions to Walmart Canada's suppliers and merchants* [Online]. Walmart. Available from: https://www.walmartcanada.ca/news/2024/11/20/walmart-luminate-arrives-in-canada--bringing-its-ecosystem-of-in#:~:text=Walmart%20Luminate%20is%20launching%20in,basket%20pairing%2C%20substitutions%20and%20more./walmart-launches-luminate-data-platform-canada [Accessed 17 April 2025].

MarketLine, 2015. *Fruit & Vegetable in Canada* [Online]. MarketLine. Available from: https://new-marketline-com.ezproxy1.bath.ac.uk/Analysis/ViewasPDF/canada-fruit-vegetables-26827 [Accessed 23 April 2025].

MarketLine, 2024. *Food & Grocery Retail in Canada* [Online]. MarketLine. Available from: https://new-marketline-com.ezproxy1.bath.ac.uk/Analysis/ViewasPDF/canada-food-grocery-retail-ml-221798 [Accessed 20 April 2025].

Statista, 2021. *Convenience Stores in Canada* [Online]. Available from: https://www-statista-com.ezproxy1.bath.ac.uk/study/70658/convenience-stores-in-canada/ [Accessed 20 April 2025].

Statista, 2024a. *Food Retail in Canada* [Online]. Statista. Available from: https://www-statista-com.ezproxy1.bath.ac.uk/study/34639/food-retail-in-canada-statista-dossier/ [Accessed 19 April 2025].

Statista, 2024b. *Supermarkets and Grocery Stores in Canada* [Online]. Available from: https://www-statista-com.ezproxy1.bath.ac.uk/study/34640/supermarkets-and-grocery-stores-in-canada-statista-dossier/ [Accessed 19 April 2025].

T, O., 2024a. *Food shopping behavior in Canada - statistics & facts* [Online]. Statista. Available from: https://www-statista-com.ezproxy1.bath.ac.uk/topics/2871/food-shopping-behavior-in-canada/#topicOverview [Accessed 21 April 2025].

T, O., 2024b. *Loblaws - Statistics & Facts* [Online]. Statista. Available from: https://www-statista-com.ezproxy1.bath.ac.uk/topics/3235/loblaws/#topicOverview [Accessed 20 April 2025].

T, O., 2025. *Number of supermarkets and other grocery retailers in Canada as of December 2024, by employment size* [Online]. Statista. Available from: https://www-statista-com.ezproxy1.bath.ac.uk/statistics/459546/supermarkets-and-other-grocery-stores-by-employment-size-canada/ [Accessed 20 April 2025].