

STATS/CSE 780
Homework Assignment 1

Konrad Swierczek - 001423065

January 31, 2023

Introduction

Throughout 2022, Canadians experienced increased costs for consumer goods. However, many factors influence the price of everyday essentials. This report visualizes trends in the price of consumer products across Canada during 2022 and shows that while prices did increase, both province and product type have an influence on these changes.¹

Methods

The dataset “Monthly average retail prices for selected products” (Statistics Canada, 2020) was retrieved from Statistics Canada. This data is available from 2017 onwards and is collected monthly from all provinces (this report also includes averaged data across provinces). For the purposes of this report, only data from 2022 is considered (Data is at present unavailable for December of 2022). 110 products ranging from produce to hygiene products and the mean of all products are included. Some data are not available in certain provinces; likely due to product availability (e.g., four litre milk bags are not available in Newfoundland & Labrador). All data processing and visualization was performed using R and the tidyverse package (R Core Team, 2020; Wickham et al., 2019).

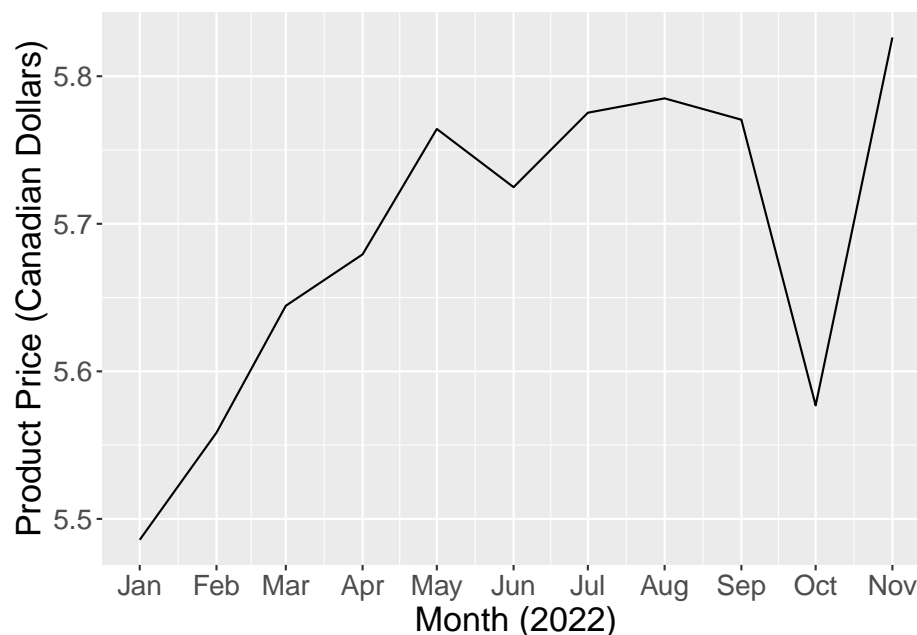


Figure 1: Average price of 110 consumer products in Canada. Prices are in Canadian dollars. Data is collected monthly.

¹All materials and reproducible code used are available at <https://github.com/konradswierczek/STATS780>

Results

Consumer product prices did increase during 2022, with an average price of \$5.61 in January and \$6.02 in November (see Figure 1). However, the data also indicates that a significant decrease in prices occurred in the Fall. The increasing trend, as well as the sudden decrease in prices in the Fall, is generally visible across all provinces (although this decrease is greater in Quebec). Figure 2 shows the average price of all 110 products for each province. In addition to the general increase in prices, in Ontario average prices increased significantly in May before rebounding. Each province also had different baseline prices for most provinces which influences how they change over time. While a normalization could have been performed to offset this, the baseline prices for each province are in of themselves of interest as they are influenced by policy and economic factors.

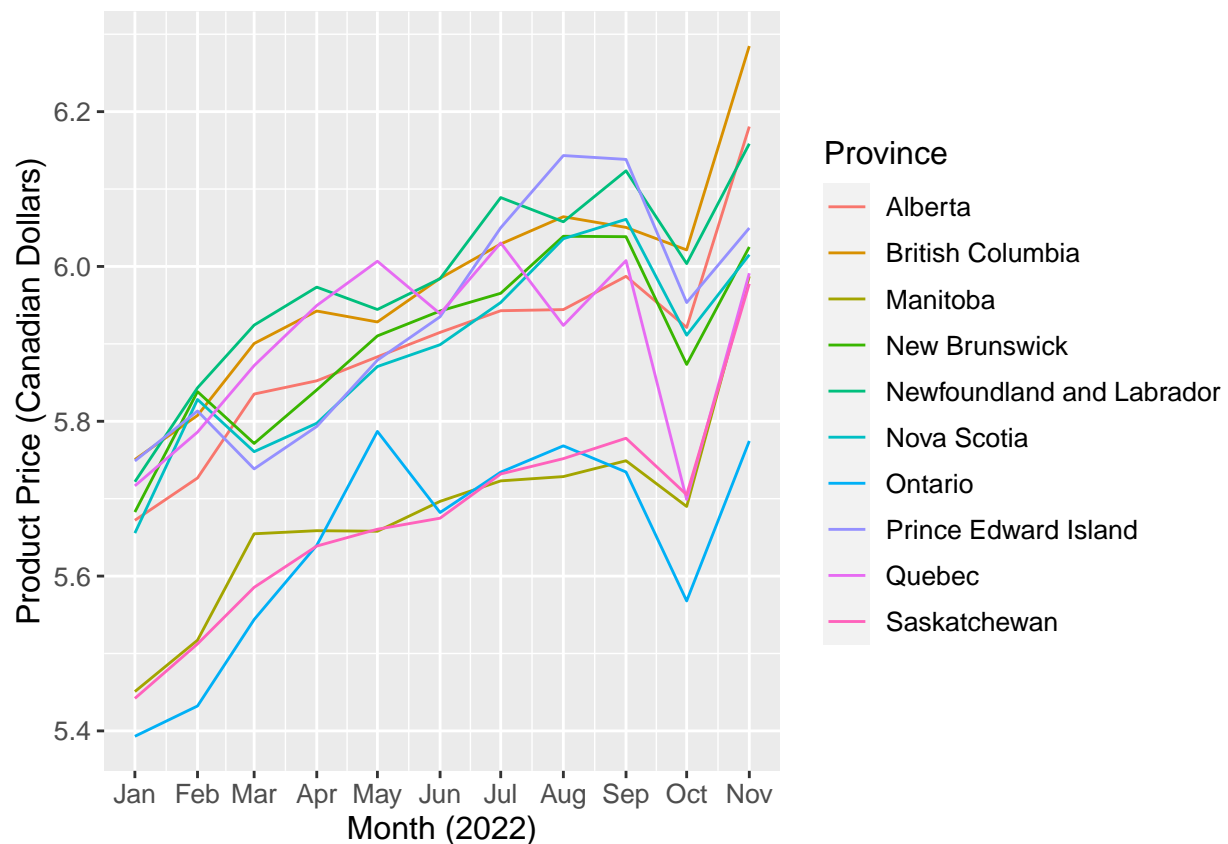


Figure 2: Average price of 110 consumer products across all Canadian provinces. Prices are in Canadian dollars. Data is collected monthly.

Despite the visible increase in prices, the average change over the year is less than a dollar. In order to consider how price increases might impact consumers, the attached [Shiny app](#) displays all 110

products individually. Users can interactively select any product and view the price throughout the year for each province or Canada-wide. When viewed individually, some items nearly doubled in price from January to November (e.g., strawberries) while others actually decreased (e.g., striploin) throughout the year.

Conclusions

This report indicates that Canadians have experienced an increase in cost of consumer goods such as food and hygienic products throughout 2022. This trend is apparent across many of the individual 110 items in the dataset, and is present in all provinces. The cause of this increase should be explored further by analyzing which specific products became more or less expensive, as well as social, economic, and policy changes that may have occurred (for instance, increased interest in the price of groceries by the media and the Competition Bureau Canada (Canada, 2022) and changing global supply line conditions). Finally, data from previous years should be compared to 2022 to determine if these increases are proportional to past increases. The price of consumer goods has an immense impact on the daily lives of Canadians, making this data and the underlying causes incredibly relevant and important to disseminate widely in a format accessible to the average Canadian.

References

- Canada, C. B. (2022). Competition bureau to study competition in canada's grocery sector. In *Canada.ca*. Government of Canada. <https://www.canada.ca/en/competition-bureau/news/2022/10/competition-bureau-to-study-competition-in-canadas-grocery-sector.html>
- R Core Team. (2020). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. <https://www.R-project.org/>
- Statistics Canada. (2020). *Monthly average retail prices for selected products, by province*. Government of Canada. <https://doi.org/10.25318/1810024501-ENG>
- Wickham, H., Averick, M., Bryan, J., Chang, W., McGowan, L. D., François, R., Grolemund, G., Hayes, A., Henry, L., Hester, J., Kuhn, M., Pedersen, T. L., Miller, E., Bache, S. M., Müller, K., Ooms, J., Robinson, D., Seidel, D. P., Spinu, V., ... Yutani, H. (2019). Welcome to the tidyverse. *Journal of Open Source Software*, 4(43), 1686. <https://doi.org/10.21105/joss.01686>

Supplementary Materials

```
# knitr setup
knitr::opts_chunk$set(echo = TRUE)
knitr::opts_chunk$set(fig.pos = "H", out.extra = "")

# imports
library(ggplot2)
library(readr)
library(shiny)
library(dplyr)
library(knitr)
library(lubridate)

# "Monthly average retail prices for selected products"
# https://doi.datacite.org/doi/10.25318/2F1810024501-eng

# Pull Data
temp <- paste(tempfile(), ".zip", sep = "")
download.file("https://www150.statcan.gc.ca/n1/tbl/csv/18100245-eng.zip", temp)
data <- read_csv(unz(temp, "18100245.csv"))
metadata <- read_csv(unz(temp, "18100245_MetaData.csv"))
unlink(temp)

# Tidy Data
tidy_data <- data

# Subset data for only 2022
tidy_data$date <- lubridate::ymd(paste(data$REF_DATE, "-01", sep = ""))
tidy_data <- subset(tidy_data, date > "2021-12-31" & date < "2023-01-01")

# Add entries for means of all products
means <- tidy_data %>%
  group_by(GEO, date) %>%
```

```
  summarise_at(vars(VALUE), list(VALUE = mean))
means$Products <- "All Products"
tidy_data <- bind_rows(tidy_data, means)

# Write .RData for shinyapp
save(tidy_data, file = "shiny/tidy_data.RData")

# Figure 1: Price in Canada have increased
ggplot(data = subset(subset(tidy_data, Products == "All Products"),
                      GEO == "Canada")) +
  geom_line(aes(x = date, y = VALUE)) +
  xlab("Month (2022)") +
  ylab("Product Price (Canadian Dollars)") +
  scale_x_date(date_breaks = "months", date_labels = "%b") +
  theme(text = element_text(size = 17))

# Figure 2: These trends vary by province
ggplot(data = subset(subset(tidy_data, Products == "All Products"),
                      GEO != "Canada")) +
  geom_line(aes(x = date, y = VALUE, colour = GEO)) +
  labs(colour = "Province") +
  xlab("Month (2022)") +
  ylab("Product Price (Canadian Dollars)") +
  scale_x_date(date_breaks = "months", date_labels = "%b") +
  theme(text = element_text(size = 12))

# Shiny App: STATS780-W23 Assignment 1
# Konrad Swierczek
# LINK TO APP: https://swierckj.shinyapps.io/assignment1/
# Imports
library(shiny)
library(ggplot2)
```

```
# Load Data
load("tidy_data.RData")
#####

# App UI
ui <- fluidPage(
  titlePanel("Monthly average retail prices for selected products"),

  mainPanel("First select a product:
    the average of all 110 products is selected by default.
    Next, choose a region.
    Canada displays the average value for all provinces,
    while All Provinces displays colour-coded lines for each province.
    This can help understand how product type and province
    influence the price of a product over time.
    For instance, the price of laundry detergent fluctuates in Quebec,
    but not Ontario."),

  selectInput(inputId = "x_var",
    label = "Select a Product",
    choices = unique(tidy_data$Products),
    selected = "All Products"),
  selectInput(inputId = "which_plot",
    label = "Region",
    choices = c("Canada", "All Provinces"),
    selected = "Canada"),

  plotOutput(outputId = "plot"

)
)
#####
```



```

# App Server
server <- function(input, output) {
  output$plot <- renderPlot({
    title <- "retail prices over time"
    if (input$which_plot == "Canada") {
      ggplot(data = subset(subset(tidy_data, Products == input$x_var),
                            GEO == "Canada")) +
        geom_line(aes(x = date, y = VALUE)) +
        labs(title = paste("Price of ", input$x_var, " over 2022", sep = "")) +
        xlab("Month (2022)") +
        ylab("Product Price (Canadian Dollar)") +
        scale_x_date(date_breaks = "months", date_labels = "%b") +
        theme(text = element_text(size = 20))
    } else {
      ggplot(data = subset(subset(tidy_data, Products == input$x_var),
                            GEO != "Canada")) +
        geom_line(aes(x = date, y = VALUE, colour = GEO)) +
        labs(colour = "Province",
              title = paste("Price of ", input$x_var, " over 2022", sep = "")) +
        xlab("Month (2022)") +
        ylab("Product Price (Canadian Dollar)") +
        scale_x_date(date_breaks = "months", date_labels = "%b") +
        theme(text = element_text(size = 20))
    }
  })
}

#####
# Run App
shinyApp(ui, server)
#####

```