

My title*

My subtitle if needed

First author

Another author

November 14, 2024

First sentence. Second sentence. Third sentence. Fourth sentence.

#Data

The preparation and manipulation of the original data is conducted using SQL [CITE SQL]. The steps to conduct the present analysis on the prepared data are carried out using the statistical programming language, R [CITE R].

The data for this analysis was obtained provided by Project Hammer [CITE DATA], and it contains data about product details (e.g. product names, brands, current prices, and old prices, and units for products) for 8 different grocery vendors (Voila, T&T, Loblaws, No Frills, Metro, Galleria, Walmart Canada, and Save-On-Foods). The data is collected through the process of screen-scraping website UIs from vendors within a neighbourhood in Toronto. Data from Project Hammer is available from February 28, 2024, and as this data is collected by extracting information from grocers' websites, it is possible to have missing data, which is more thoroughly addressed in [CROSS REFERENCE DISCUSSION SECTION]. Through the method of screen scraping, the original data aims to measure and reflect grocery vendor decisions and trends in product pricing.

For this analysis, three vendors - Voila, Metro, and No Frills - are considered along with their respective pricing for products that are 100g. The "Price" variable for each vendor in this analysis reflects the current price of their products at the time in which the data was downloaded. The variable, "Units", represents the size of products measured using weight metrics (i.e. grams or kilograms) and counts (i.e. the number of items within a given product). In this case, a weight of 100g is selected to ensure consistency in units across products and vendors.

*Code and data are available at: https://github.com/RohanAlexander/starter_folder.



