Comparing Ground Beef Prices by Type and Vendor: Optimal Consumer Choices*

Daniel Du Michelle Ji

November 14, 2024

This paper analyzes ground beef prices across major grocery vendors, examining how pricing varies by type (medium, lean, extra lean) and store. We find that medium ground beef is generally the most affordable option, while extra lean is the most expensive, with significant differences observed between vendors. Notably, Voila and Save-On-Foods consistently offer below-average prices, while Metro displays the greatest price variability. These insights provide consumers with a clearer understanding of where to find the best value for their preferred ground beef type, helping them make informed decisions about grocery purchases.

1 Introduction

The grocery market plays an increasingly vital role in daily life, influencing consumer meal planning and budgeting for essential items. As one of the most widely consumed proteins, ground beef is a staple across households, with prices and preferences varying by type and store. Factors such as ground beef type (e.g., lean, medium, extra lean) and retailer play a significant role in determining price, impacting consumer decisions and shopping habits. This paper explores ground beef pricing trends across major vendors, focusing on the price differences between various ground beef types at different grocery stores. Our findings aim to provide consumers with insights into where they can find the best value for their preferred ground beef type, helping them make more informed decisions about where and when to buy.

The paper is broken down into various sections. Section 2 explores the data, looking at ground beef prices of different types and stores over time. Section 3 introduces observations and potential conclusions from the data analysis. Section 4 discusses potential issues to be aware of with the data and our analysis.

*Code and data are available at: https://github.com/danield424/Ground_Beef_Price_Analysis/.

2 Data

We use R Core Team (2023) and Wickham et al. (2019). Data is downloaded from Filipp (2024). We use *SQLite* (2024) to clean the data, filtering for 1lb ground beef products for every store. We look at the price over time of 1 pound of different types of ground beef from various vendors by filtering the data to search for only ground beef products that weigh ~454g (with 5g margin of error). We exclude products like organic, kosher, halal ground beef, and assign ground beef to be of type medium, lean, or extra lean. If there are multiple price updates of the same product on the same date for the same store, we choose the cheapest values, since we assume there was an error with data collection and keep the most accurate.

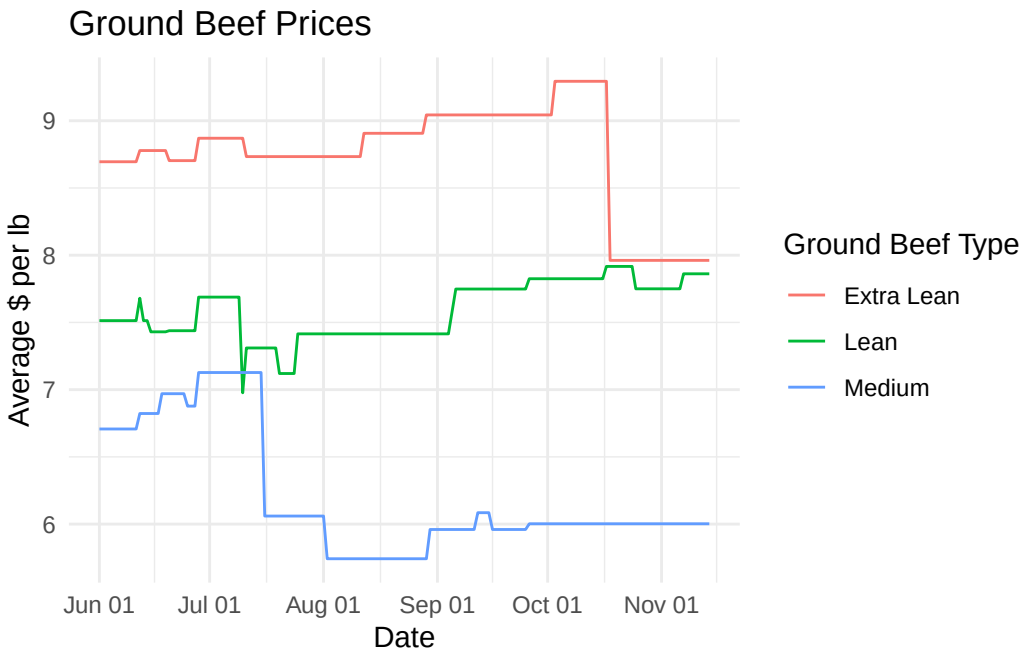

Figure 1: Ground beef prices, all vendors, June 1 - November 14, 2024.

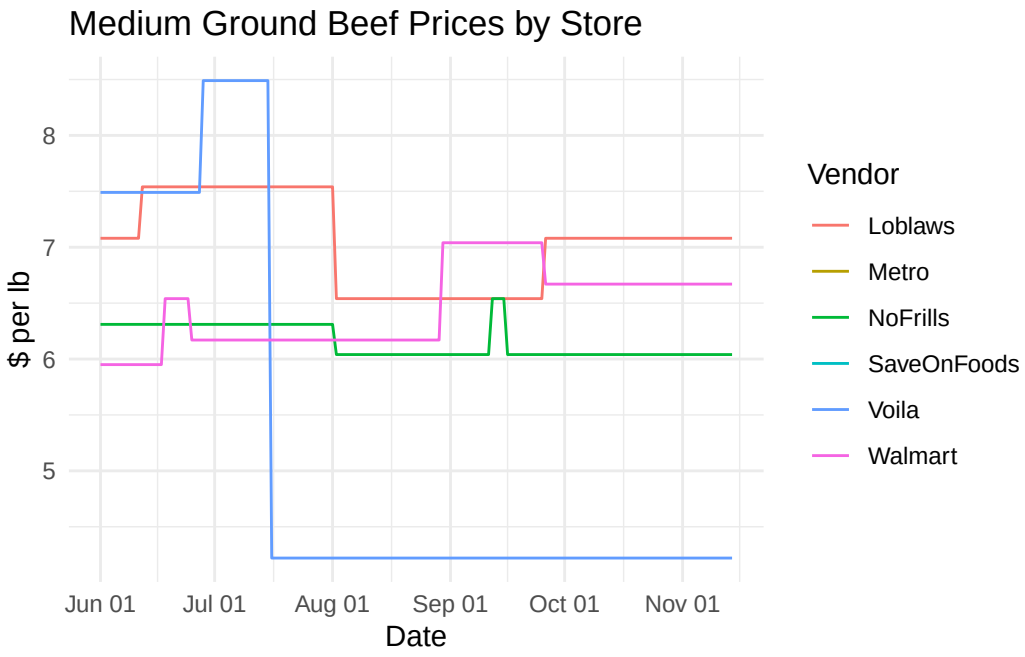

Figure 2: Medium ground beef prices of each vendor from June 1 - November 14, 2024.

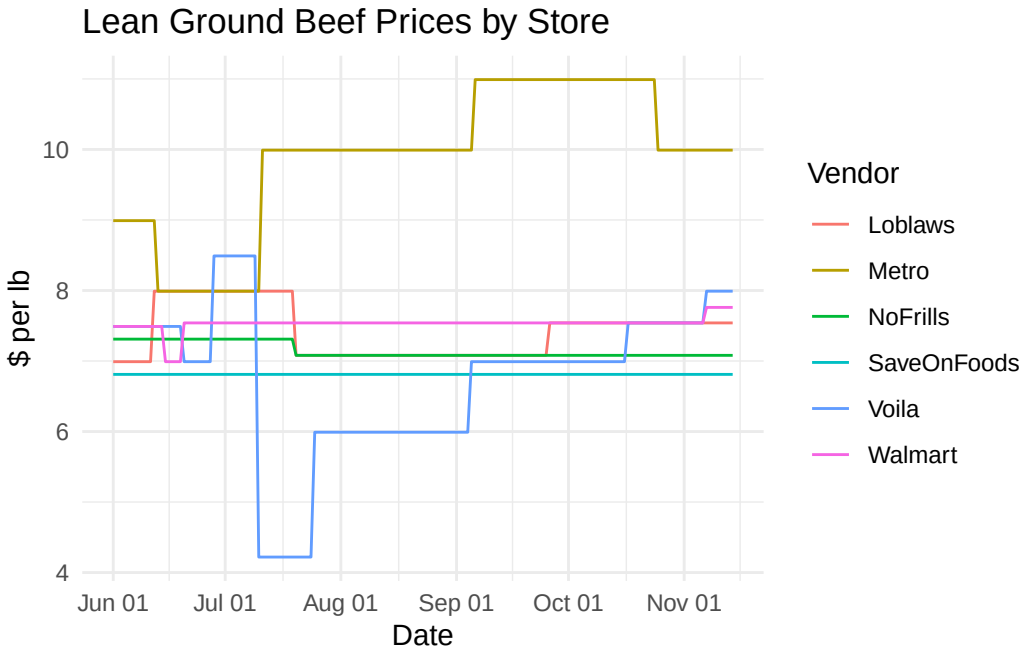

Figure 3: Lean ground beef prices of each vendor from June 1 - November 14, 2024.

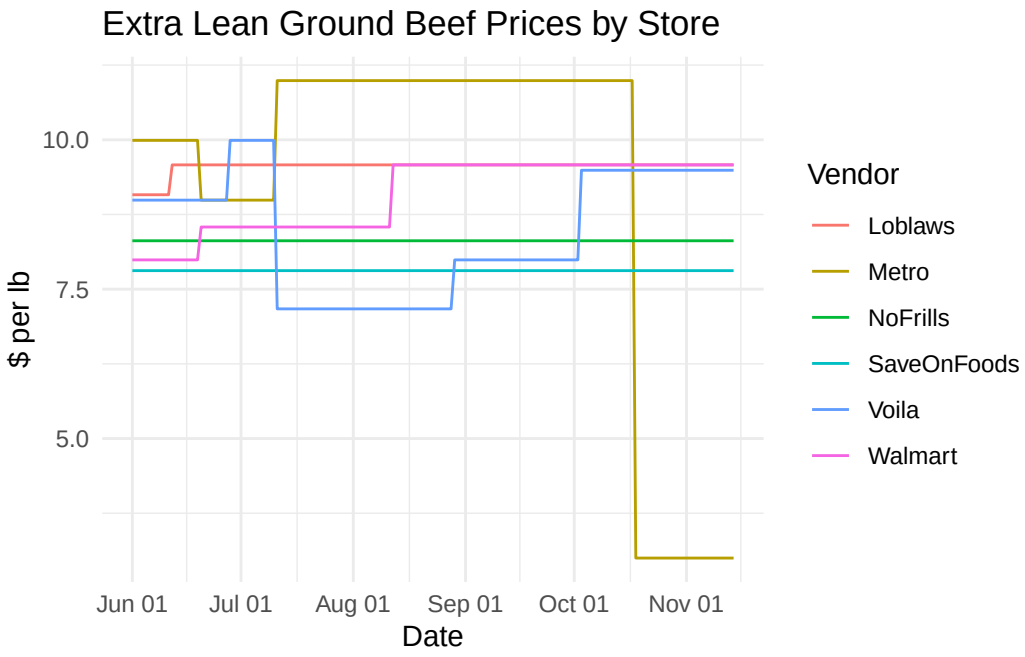

Figure 4: Extra lean ground beef prices of each vendor from June 1 - November 14, 2024.

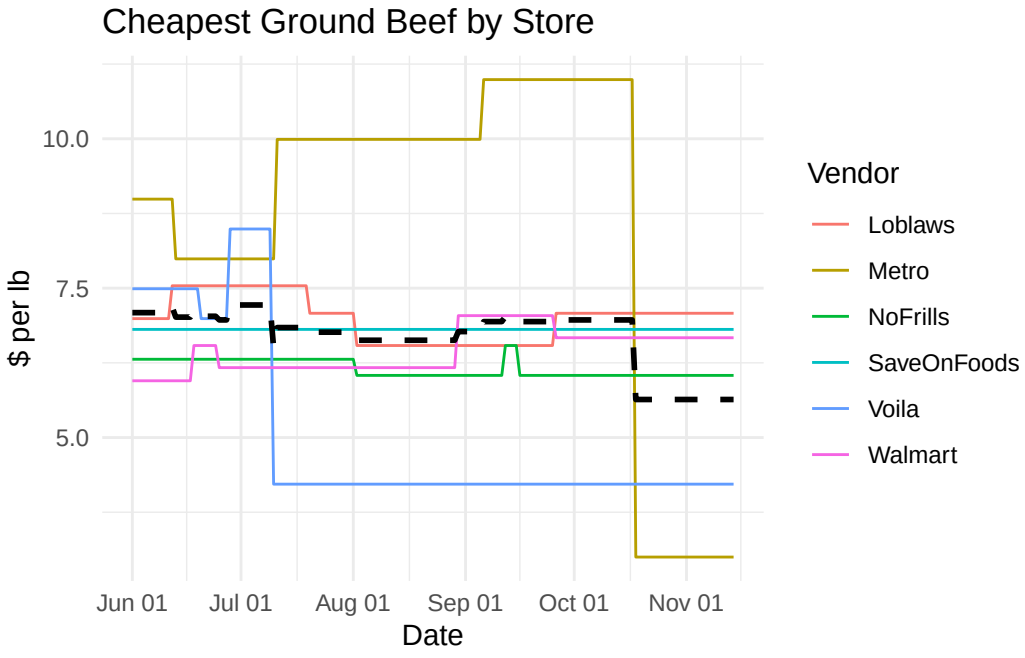

Figure 5: Cheapest ground beef pound at each vendor from June 1 - November 14, 2024.

3 Results

From Figure 1, we find that on average across all vendors, medium ground beef is the cheapest, followed by lean and then extra lean. Figure 2, Figure 3, Figure 4 show the price over time of each type of ground beef for each vendor. Finally, Figure 5 shows the price over time of the cheapest pound of ground beef each vendor has at any given time, as well as the average among vendors of the cheapest pound of ground beef. From this graph, we can see that Voila and Save on Foods usually have below-average prices for ground beef, while Metro has the greatest variance, having both the cheapest and most expensive ground beef on different dates.

4 Discussion

4.1 Correlation vs. Causation

Correlation vs causation should always be taken into account when drawing conclusions from analysis. For example, the analysis reveals a pattern where medium ground beef tends to be the cheapest option across all vendors, followed by lean and extra lean types. However, this observed relationship between ground beef type and price should be interpreted as a correlation, not a causative link. For example, while medium ground beef is consistently priced lower, this does not imply that its type alone directly determines its price. Other factors, such as consumer demand, production costs, and vendor-specific pricing strategies, likely play a role in setting these prices. Therefore, while the data highlights trends in ground beef pricing by type, further investigation would be required to understand the underlying factors driving these price differences.

4.2 Missing Data

One limitation in this analysis is the potential impact of missing data. Price data from each vendor was not collected every day, which could affect our interpretation of price trends over time. While we addressed missing data by interpolating values, this approach has limitations and may not fully capture the true pricing trends. For instance, prices could temporarily drop for promotions or rise due to supply issues, which interpolation would smooth out, potentially obscuring the true variability in prices. Collecting more consistent data would provide a more accurate and nuanced picture of price trends over time.

4.3 Sources of Bias

Several potential sources of bias could influence the findings of this analysis. Selection bias may arise from the choice of vendors included, as different stores may target different consumer demographics or operate in different pricing tiers. Moreover, screen-scraping data from public

websites may capture only advertised prices, which may differ from in-store prices or reflect temporary promotions rather than regular pricing. Finally, regional variations in demand and supply, which influence ground beef pricing, might not be fully represented in a dataset focused on a single region. Recognizing these sources of bias is crucial for interpreting the results with an understanding of the study's limitations.

References

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

R Core Team. 2023. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. https://www.R-project.org/.

SQLite. 2024. https://www.sqlite.org/.

Wickham, Hadley, Mara Averick, Jennifer Bryan, Winston Chang, Lucy D'Agostino McGowan, Romain François, Garrett Grolemund, et al. 2019. "Welcome to the tidyverse." *Journal of Open Source Software* 4 (43): 1686. https://doi.org/10.21105/joss.01686.