My title*

My subtitle if needed

First author Another author

December 1, 2024

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

1 Introduction

#Data

The Canadian Grocery Price Data(Filipp 2024) is a comprehensive dataset designed to analyze pricing trends across eight major grocery vendors: Voila, T&T, Loblaws, No Frills, Metro, Galleria, Walmart Canada, and Save-On-Foods. It includes historical price data collected through web scraping, starting from February 28, 2024, with updates up to the latest extract. The data is structured in two key formats: CSV files for product metadata and time-series prices, and an SQLite database that combines these for efficient querying. This dataset offers valuable insights into pricing dynamics, sale patterns, and vendor strategies, making it a critical tool for economic analysis, legal research, and policymaking aimed at fostering competition in Canada's grocery sector.

The dataset was simulated, cleaned, analyzed, and tested using the R programming language (R Core Team 2023a), tidyverse (Wickham et al. 2019), knitr (Xie 2014), ggplot2 (Wickham 2016) for plots, gt(Iannone et al. 2024) for tables, tidyr(Wickham and Henry 2023), arrow(Richardson and Labs 2023) for parquet, here(Müller 2023), rstanarm(Goodrich et al. 2023), broom(Robinson et al. 2023), loo(Vehtari et al. 2023), lubridate (Grolemund and Wickham 2023), while tibble (Wickham and Müller 2023) helped simplify data frame management. The testthat package (Wickham et al. 2023) was essential for unit testing and ensuring code reliability, and we employed styler (Walthert and Meyer 2023) for reformatting and maintaining a consistent code style.

Estimand paragraph

*Code and data are available at: https://github.com/aj3616/Canadian_Groceries_SQL/tree/main

Results paragraph

Why it matters paragraph

Telegraphing paragraph: The remainder of this paper is structured as follows. Section 2….

2 Data

2.1 Overview

We use the statistical programming language R (R Core Team 2023b)…. Our data (Filipp 2024)…. Following Alexander (2023), we consider…

Overview text

2.2 Measurement

Some paragraphs about how we go from a phenomena in the world to an entry in the dataset.

2.3 Outcome variables

Add graphs, tables and text. Use sub-sub-headings for each outcome variable or update the subheading to be singular.

Some of our data is of penguins (**?@fig-bills**), from (**palmerpenguins?**).

Talk more about it.

And also planes (Figure 1). (You can change the height and width, but don't worry about doing that until you have finished every other aspect of the paper - Quarto will try to make it look nice and the defaults usually work well once you have enough text.)

Talk way more about it.

2.4 Predictor variables

Add graphs, tables and text.

Use sub-sub-headings for each outcome variable and feel free to combine a few into one if they go together naturally.

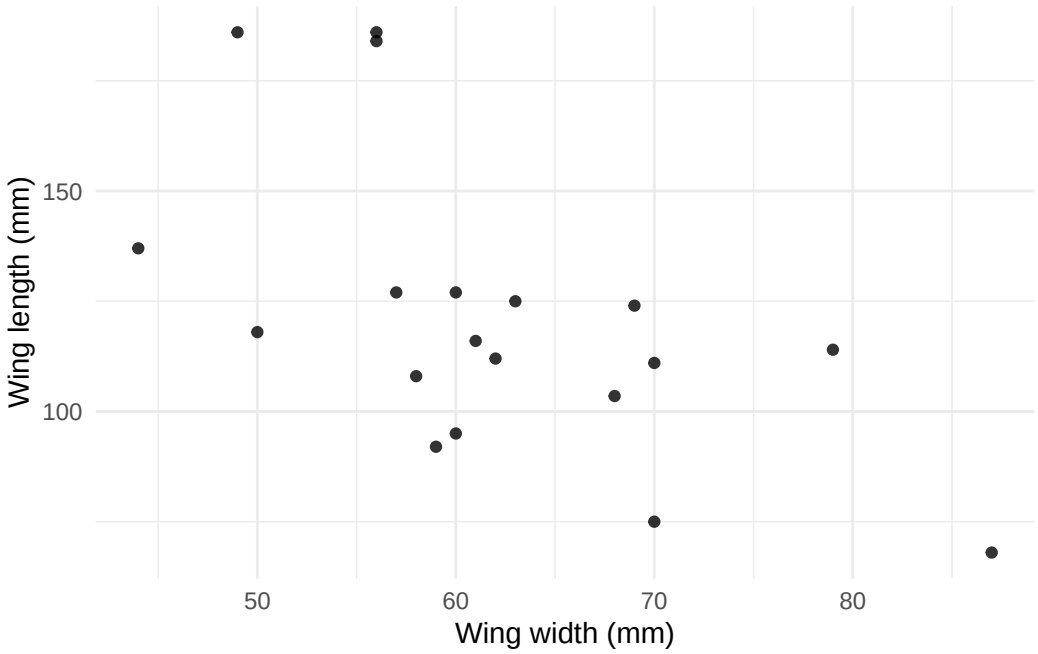


Figure 1: Relationship between wing length and width

3 Model

The goal of our modelling strategy is twofold. Firstly,…

Here we briefly describe the Bayesian analysis model used to investigate… Background details and diagnostics are included in Appendix B.

3.1 Model set-up

Define y_i as the number of seconds that the plane remained aloft. Then β_i is the wing width and γ_i is the wing length, both measured in millimeters.

$$y_i|\mu_i, \sigma \sim \text{Normal}(\mu_i, \sigma) \tag{1}$$
$$\mu_i = \alpha + \beta_i + \gamma_i \tag{2}$$
$$\alpha \sim \text{Normal}(0, 2.5) \tag{3}$$
$$\beta \sim \text{Normal}(0, 2.5) \tag{4}$$
$$\gamma \sim \text{Normal}(0, 2.5) \tag{5}$$
$$\sigma \sim \text{Exponential}(1) \tag{6}$$

Table 1: Explanatory models of flight time based on wing width and wing length

	First model
(Intercept)	1.12
	(1.70)
length	0.01
	(0.01)
width	−0.01
	(0.02)
Num.Obs.	19
R2	0.320
R2 Adj.	0.019
Log.Lik.	−18.128
ELPD	−21.6
ELPD s.e.	2.1
LOOIC	43.2
LOOIC s.e.	4.3
WAIC	42.7
RMSE	0.60

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

3.1.1 Model justification

We expect a positive relationship between the size of the wings and time spent aloft. In particular...

We can use maths by including latex between dollar signs, for instance θ.

4 Results

Our results are summarized in Table 1.

5 Discussion

5.1 First discussion point

If my paper were 10 pages, then should be be at least 2.5 pages. The discussion is a chance to show off what you know and what you learnt from all this.

5.2 Second discussion point

Please don't use these as sub-heading labels - change them to be what your point actually is.

5.3 Third discussion point

5.4 Weaknesses and next steps

Weaknesses and next steps should also be included.

Appendix

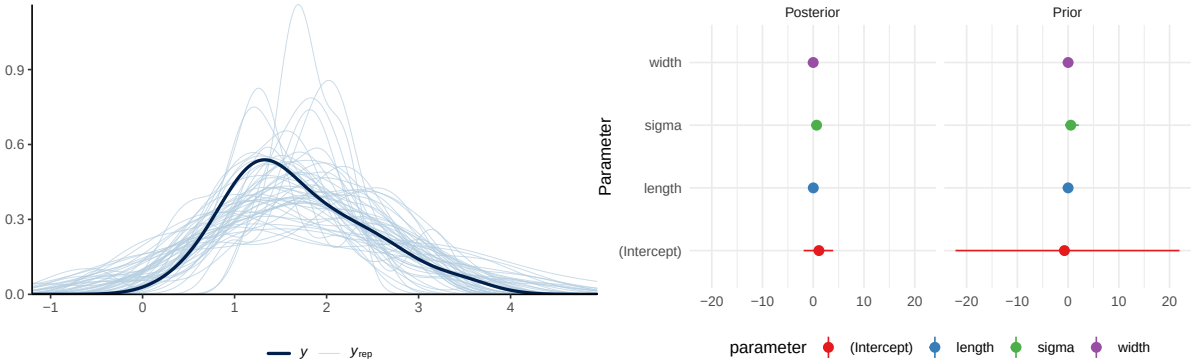
A Additional data details

B Model details

B.1 Posterior predictive check

In Figure 2a we implement a posterior predictive check. This shows...

In Figure 2b we compare the posterior with the prior. This shows...



(a) Posterior prediction check

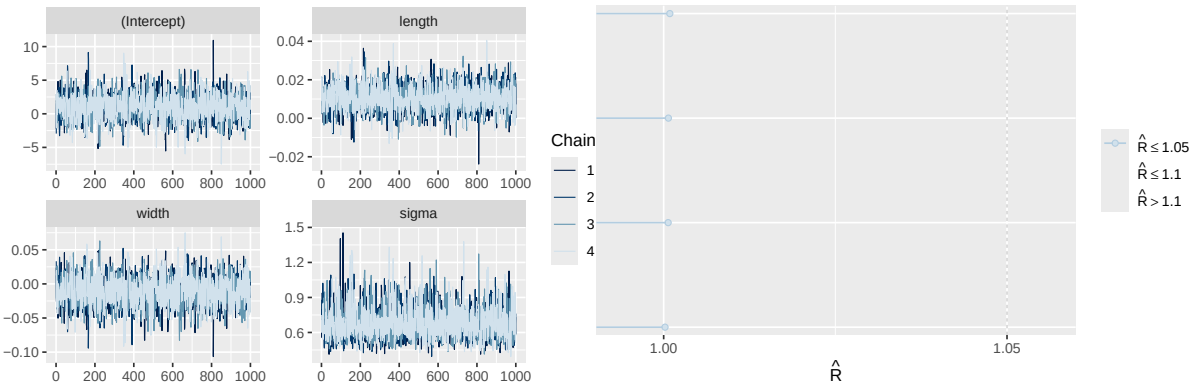
(b) Comparing the posterior with the prior

Figure 2: Examining how the model fits, and is affected by, the data

B.2 Diagnostics

Figure 3a is a trace plot. It shows... This suggests...

Figure 3b is a Rhat plot. It shows... This suggests...



(a) Trace plot (b) Rhat plot

Figure 3: Checking the convergence of the MCMC algorithm

References

Alexander, Rohan. 2023. *Telling Stories with Data*. Chapman; Hall/CRC. https://tellingstorieswithdata.com/.

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

Goodrich, Ben, Jonah Gabry, Imad Ali, and Sam Brilleman. 2022. "rstanarm: Bayesian applied regression modeling via Stan." https://mc-stan.org/rstanarm/.

Goodrich, Ben, Jonah Gabry, Imad Ali, Sam Brilleman, and Rok Češnovar. 2023. *Rstanarm: Bayesian Applied Regression Modeling via Stan*. https://mc-stan.org/rstanarm/.

Grolemund, Garrett, and Hadley Wickham. 2023. *Lubridate: Make Dealing with Dates a Little Easier*. https://CRAN.R-project.org/package=lubridate.

Iannone, Richard, Joe Cheng, Barret Schloerke, Ellis Hughes, Alexandra Lauer, JooYoung Seo, Ken Brevoort, and Olivier Roy. 2024. *Gt: Easily Create Presentation-Ready Display Tables*. https://CRAN.R-project.org/package=gt.

Müller, Kirill. 2023. *Here: A Simpler Way to Find Your Files*. https://CRAN.R-project.org/package=here.

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

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

Richardson, Neal, and Ursa Labs. 2023. *Arrow: Integration to 'Apache' 'Arrow'*. https://CRAN.R-project.org/package=arrow.

Robinson, David, Alex Hayes, Simon Couch, Nicholas Tierney, and Max Kuhn. 2023. *Broom: Convert Statistical Objects into Tidy Tibbles*. https://CRAN.R-project.org/package=broom.

Vehtari, Aki, Jonah Gabry, Yuling Yao, Andrew Gelman, Mans Magnusson, and Paul-Christian Bürkner. 2023. *Loo: Efficient Leave-One-Out Cross-Validation and WAIC for*

Bayesian Models. https://mc-stan.org/loo/.

Walthert, Lorenz, and Nicolas Meyer. 2023. *Styler: Non-Invasive Pretty Printing of r Code*. https://CRAN.R-project.org/package=styler.

Wickham, Hadley. 2016. *Ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York. https://ggplot2.tidyverse.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.

Wickham, Hadley, and Lionel Henry. 2023. *Tidyr: Tidy Messy Data*. https://CRAN.R-project.org/package=tidyr.

Wickham, Hadley, Jim Hester, Lionel Henry, and Winston Chang. 2023. *Testthat: Unit Testing for r*. https://CRAN.R-project.org/package=testthat.

Wickham, Hadley, and Kirill Müller. 2023. *Tibble: Simple Data Frames*. https://CRAN.R-project.org/package=tibble.

Xie, Yihui. 2014. "Knitr: A Comprehensive Tool for Reproducible Research in R." In *Implementing Reproducible Computational Research*, edited by Victoria Stodden, Friedrich Leisch, and Roger D. Peng. Chapman; Hall/CRC.