My title*

My subtitle if needed

Amy Jin Maggie Zhang

October 4, 2024

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

1 Introduction

Overview paragraph

Estimand paragraph

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 2023)…. Our data (Toronto Shelter & Support Services 2024)…. Following Alexander (2023), we consider…

Overview text

A sense of the dataset should be communicated to the reader. The broader context of the dataset should be discussed. Explain if there were similar datasets that could have been used and why they were not.

*Code and data are available at: https://github.com/aj3616/Forcasting-US-Elections.

2.2 Variables

All variables should be thoroughly examined and explained.
If variables were constructed then this should be mentioned, and high-level cleaning aspects of note should be mentioned, but this section should focus on the destination, not the journey. It is important to understand what the variables look like by including graphs, and possibly tables, of all observations, along with discussion of those graphs and the other features of these data.

2.3 Summary statistics & Relationships

Summary statistics should also be included, and well as any relationships between the variables. If this becomes too detailed, then appendices could be used. Basically, for every variable in your dataset that is of interest to your paper there needs to be graphs and explanation and maybe tables

2.4 Measurement

Some paragraphs about how we go from a phenomena in the world to an entry in the dataset. A thorough discussion of measurement, relating to the dataset, is provided in the data section. Please ensure that you explain how we went from some phenomena in the world that happened to an entry in the dataset that you are interested in.

2.5 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 (Figure 1), from Horst, Hill, and Gorman (2020).

Talk more about it.

And also planes (Figure 2). (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.

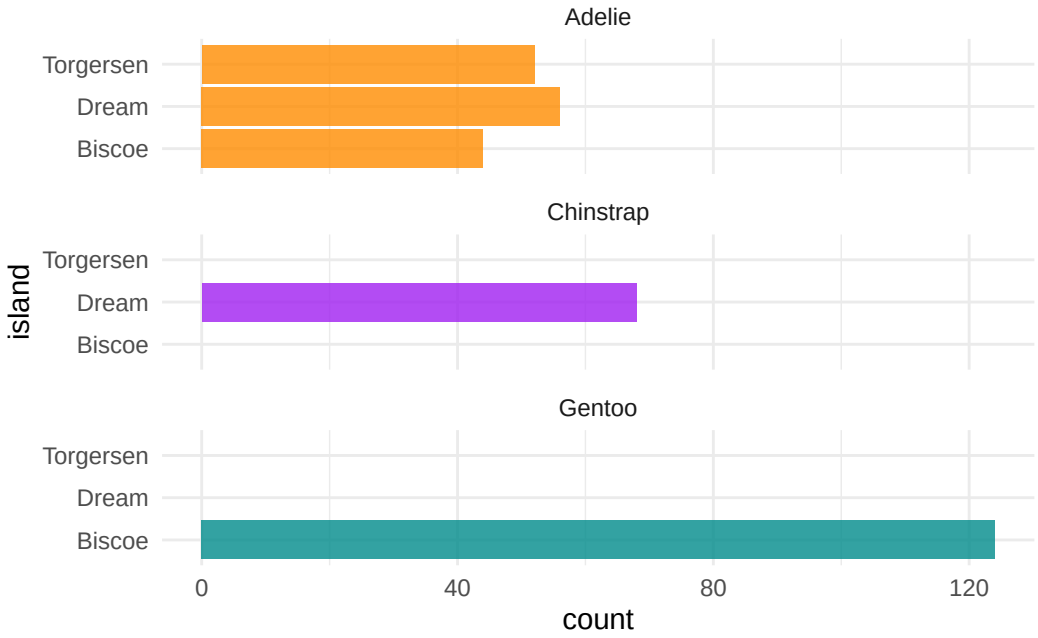


Figure 1: Bills of penguins

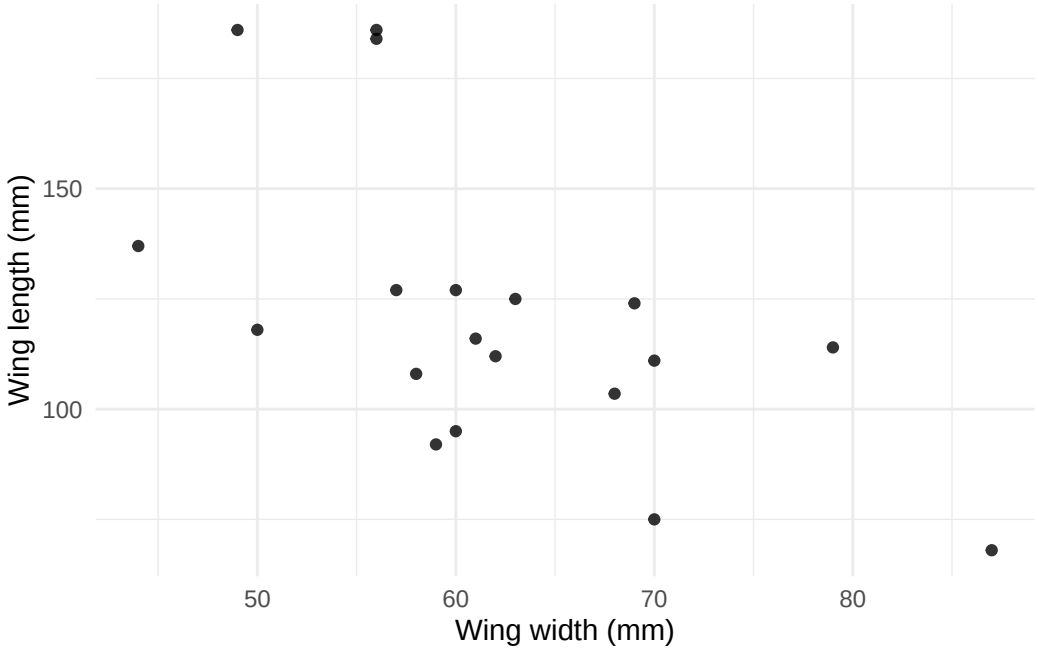


Figure 2: Relationship between wing length and width

2.6 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.

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.

The model should be nicely written out, well-explained, justified, and appropriate.

Detail the statistical model used, defining and explaining each aspect and its importance, and ensure that variables are well-defined and correspond with those discussed in the data section.

The model should have an appropriate balance of complexity—neither overly simplistic nor unnecessarily complicated—and be justified as suitable for the situation.

3.1 Modeling decisions

Explain how decisions made in modeling reflect the aspects discussed in the data section, including why specific features are included (e.g., why use age rather than age-groups, treating province effects as levels, categorizing gender, etc?).

3.2 Mathematical notations

Present the model using appropriate mathematical notation supplemented with plain English explanations, defining every component. If applicable, define sensible priors for Bayesian models.

3.3 Assumptions, limitations, circumstances

Clearly discuss the underlying assumptions, potential limitations, and circumstances where the model may not be appropriate.

3.4 Software & model validation

Mention the software used to implement the model, and provide evidence of model validation and checking—such as out-of-sample testing, RMSE calculations, test/training splits, or sensitivity analyses—while addressing model convergence, diagnostics, and any alternative models or variants considered,

3.5 Strenth and weeknesses and final choice

including their strengths and weaknesses and the rationale for the final model choice.

3.6 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.

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

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

3.6.1 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.

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

4 Results

Our results are summarized in Table 1.

Results will likely require summary statistics, tables, graphs, images, and possibly statistical analysis or maps. There should also be text associated with all these aspects.

Show the reader the results by plotting them where possible. Talk about them. Explain them. That said, this section should strictly relay results.

Regression tables must not contain stars.

5 Discussion

5.1 What is done in this paper?

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 What is something that we learn about the world?

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

5.3 What is another thing that we learn about the world?

5.4 Weaknesses and next steps

What are some weaknesses of what was done? What is left to learn or how should we proceed in the future?

Weaknesses and next steps should also be included.

Appendix

The deep dive provides a thorough understanding of how something goes from being a person's opinion to part of a result for this pollster. Please provide a thorough overview and evaluation of the pollster's methodology, highlighting both its strengths and limitations.

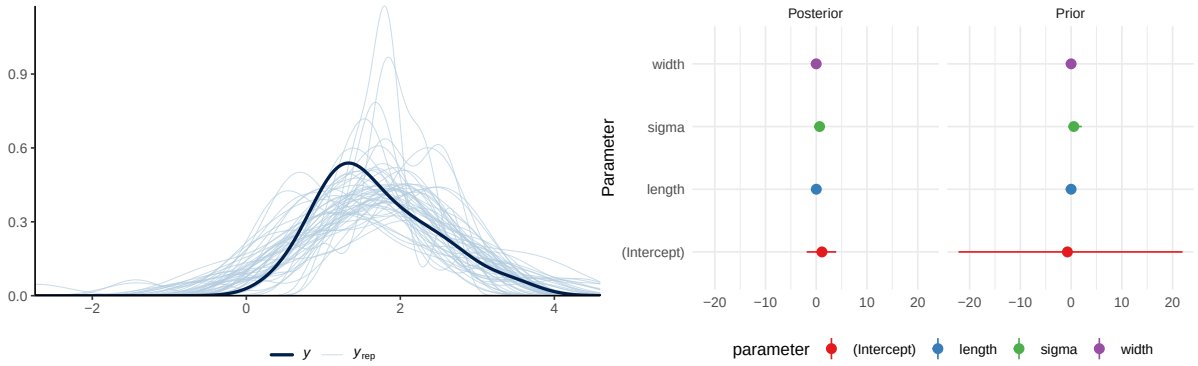
A Additional data details

B Model details

B.1 Posterior predictive check

In Figure 3a we implement a posterior predictive check. This shows...

In Figure 3b we compare the posterior with the prior. This shows...



(a) Posterior prediction check

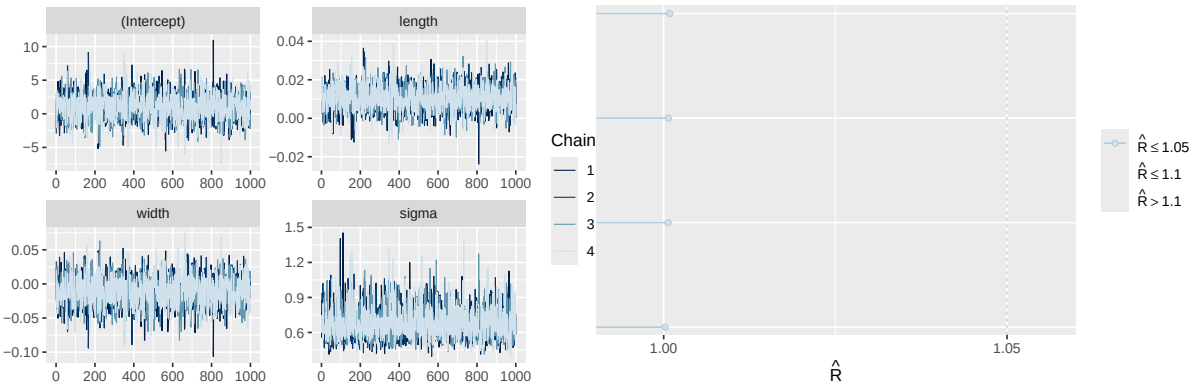
(b) Comparing the posterior with the prior

Figure 3: Examining how the model fits, and is affected by, the data

B.2 Diagnostics

Figure 4a is a trace plot. It shows... This suggests...

Figure 4b is a Rhat plot. It shows... This suggests...



(a) Trace plot (b) Rhat plot

Figure 4: Checking the convergence of the MCMC algorithm

References

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

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

Horst, Allison Marie, Alison Presmanes Hill, and Kristen B Gorman. 2020. *palmerpenguins: Palmer Archipelago (Antarctica) penguin data*. https://doi.org/10.5281/zenodo.3960218.

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

Toronto Shelter & Support Services. 2024. *Deaths of Shelter Residents*. https://open.toronto.ca/dataset/deaths-of-shelter-residents/.