Outbreaks in Toronto: Examining Disease Outbreaks Trends Between 2020 to 2024*

An exploration of the location and frequency of disease outbreaks in Toronto between the years 2020 and 2024

Sindhu Priya Mallavarapu

April 18, 2024

This paper analyzes the disease outbreak patterns across the city of Toronto in the years between 2020 - 2024 based on location and models future outbreak trends. It is found that the highest number of disease outbreaks within these five years happened Long-Term Care Facilities. The results of this paper can help government officials see where more healthcare funding and stricter safety practices are needed.

1 Introduction

The pandemic has put a clear spotlight on all the cracks in the medical system. As Toronto is the most populous city in Canada, there is particular drain on this city's medical system. Thus, identifying where there are higher outbreaks can show where the funding needs to be directed to in order to improve the medical system. This paper uses disease outbreak data form Open Data Toronto (Gelfand 2022) to show where the highest number of outbreaks took place in Toronto and model the future outbreak trends for that setting. The estimand in this paper is the total disease outbreaks in each setting.

It is found that no matter what year between 2020 to 2024 were examine, the most outbreaks occurred in Long-Term Care Facilities and the least amount of outbreaks happened in Transitional Care. Moreover, between the examined years, 2022 was year with the highest amount of total disease outbreaks. These findings are important as they can help government officials see where there is a need for more funding and health care professionals.

The paper is structured in the following way. The Data section, Section 2, goes into detail about the data that was used in this paper, including details about its measurement. The

*Code and data are available at: https://github.com/MSindhuPriya/toronto_outbreaks.

Model section, Section 3, talks about the model the model used in the paper. The Results section, Section 4, shows what was found from analyzing the data. Further discussion of the results, limitations, and future research pathways are detailed in the Discussion section, Section 5. The Conclusion and Appendix provide a brief summary, as well as, additional details that were not included in the main body of the paper.

2 Data

We use R (R Core Team 2022) and it's packages [list packages here] to analyze the data.

Some of our data is of penguins (**?@fig-bills**), from Horst, Hill, and Gorman (2020).

Talk more about it.

And also planes (**?@fig-planes**). (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.

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)	5.08
	(0.00)
outbreak_settingHospital-Chronic Care	−0.26
	(0.01)
outbreak_settingHospital-Psychiatric	−2.41
	(0.03)
outbreak_settingLTCH	0.99
	(0.00)
outbreak_settingRetirement Home	0.00
	(0.00)
outbreak_settingTransitional Care	−2.65
	(0.06)
Num.Obs.	3418
Log.Lik.	−97 016.372
ELPD	−97 095.1
ELPD s.e.	1621.8
LOOIC	194 190.2
LOOIC s.e.	3243.6
WAIC	194 190.0
RMSE	123.76

We run the model in R (R Core Team 2022) 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

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 **?@fig-ppcheckandposteriorvsprior-1** we implement a posterior predictive check. This shows…

In **?@fig-ppcheckandposteriorvsprior-2** we compare the posterior with the prior. This shows…

Examining how the model fits, and is affected by, the data

Figure 1: **?(caption)**

B.2 Diagnostics

?@fig-stanareyouokay-1 is a trace plot. It shows… This suggests…

?@fig-stanareyouokay-2 is a Rhat plot. It shows… This suggests…

Checking the convergence of the MCMC algorithm

Figure 2: **?(caption)**

References

Gelfand, Sharla. 2022. *Opendatatoronto: Access the City of Toronto Open Data Portal*. https://CRAN.R-project.org/package=opendatatoronto.

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. 2022. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. https://www.R-project.org/.