Homelessness and Drug Toxicity*

Gavin Crook Samarth Rajani Mohid Sharif

March 15, 2024

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

1 Introduction

You can and should cross-reference sections and sub-sections. We use R Core Team (2023) and Wickham et al. (2019a).

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

2 Data

We obtained our data from the City of Toronto **opendatatoronto** database portal, using the 'opendatatoronto' package (Gelfand 2020) and the statistical programming language **R** (R Core Team 2023). We used the **tidyverse** package for data manipulation (Wickham et al. 2019b) and **kableExtra** for table formatting (Zhu 2021). The header includes two lines of code **"usepackage{float}"** which allow the use of float in R markdown and the line **"floatplacement{figure}{H}"** (user9112767 2018) which keeps the tables and figures locked in the specific place where they are written in R markdown.

This data set records reports of homeless deaths across Toronto and records the details of the deceased. The data records the **year of death, cause of death, age group, gender, and number of deaths** for every report. The data classifies cases by age group rather than specific ages, age groups are grouped by 20 year age gaps (E.g. "20 to 39 Years") except for the first group which is **"<20"** and the last group which is **"60+"**. Gender is identified and reported as **"Male"** or **"Female"**. Year of death is simply as the year the deaths were reported, starting from **2017** up until **2023**. Cause of deaths are classified as follows: **"Accident", "Cancer", "Cardiovascular Disease", "Covid-19", "Drug Toxicity", "Homicide",**

*Code and data are available at: https://github.com/MohidSharif/Drug-Toxicity-and-Homelessness..

“Pneumonia”, “Other”, “Suicide”, and “Unknown/Pending”. And the number of deaths is simply the number of deaths provided in that report.

Since we are only interested in deaths due to drug toxicity, we simplified the data and created two data sets. First the data was cleaned to remove any **“Unknown”** or **empty** values. Then causes of death due to **Drug Toxicity** were isolated and all others were removed from the data set. We then created two data sets, one highlighting the number of deaths per year for each gender and one for the number of deaths per year for each age group. Since each report in the data set can contain multiple deaths, therefore all counts from each report had to be added to their respective grouping to create a new death count variable for the two data sets. Using these new data sets we can now compare and analyze the death trend over the years for each gender and age group.

Table 1 shows the data associated with cases reported as “death due to drug toxicity”. This data sample shows the number of deaths reported for the years 2017 and 2018 for each age group.

Table 1: Age Group Deaths Due to Drug Toxicity

Year	Age Group	Deaths
2017	<20	0
2017	20-39	13
2017	40-59	15
2017	60+	3
2018	<20	0
2018	20-39	16
2018	40-59	15
2018	60+	2

Table 2 shows the data associated with cases reported as “death due to drug toxicity”.This data shows us the number of deaths reported for the years 2017-2020 for each gender.

Table 2: Gender Deaths Due to Drug Toxicity

Year	Gender	Deaths
2017	Male	26
2017	Female	5
2018	Male	22
2018	Female	11
2019	Male	26
2019	Female	13
2020	Male	57

2020	Female	17

(**age-graph?**) visualizes our first data of interest, showing deaths per year for each age group. Looking at this graph teaches us a few things. Firstly, homeless aged 40-59 are the most prone to death due to drug intoxication, followed by age group 20-39. Secondly, deaths by intoxication peaked during the years 2020-2022, these years are when the pandemic was most prevelant. Lastly, there seems to be no improvement in the homeless drug problem over this time period.

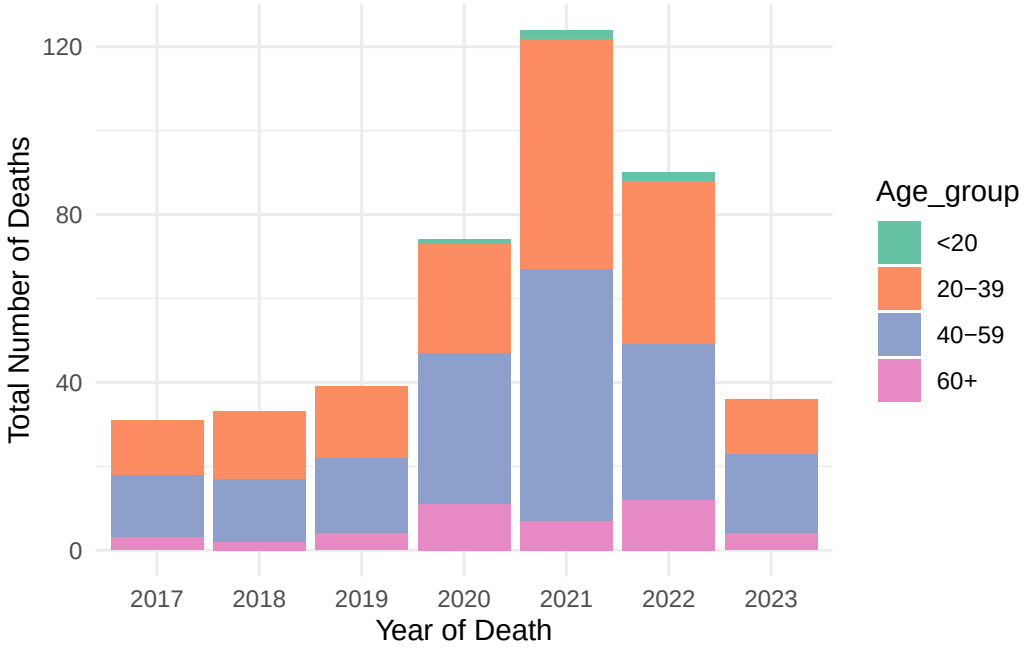


Figure 1: Deaths Per Year for Age Group

(**gender-graph?**) visualizes our second data set of interest, showing number of deaths for males and females from 2017-2023. We can learn a few more things on top of what (**age-graph?**) teaches us. From (**gender-graph?**) we can see that men see more deaths due to intoxication than women. We also see a much drastic increase in deaths by drug intoxication for men during the pandemic years compared to women. We also see a large decrease in deaths of women for the year 2023. Overall, we see no improvement in deaths for men, while we only see an improvement for women in the year 2023.

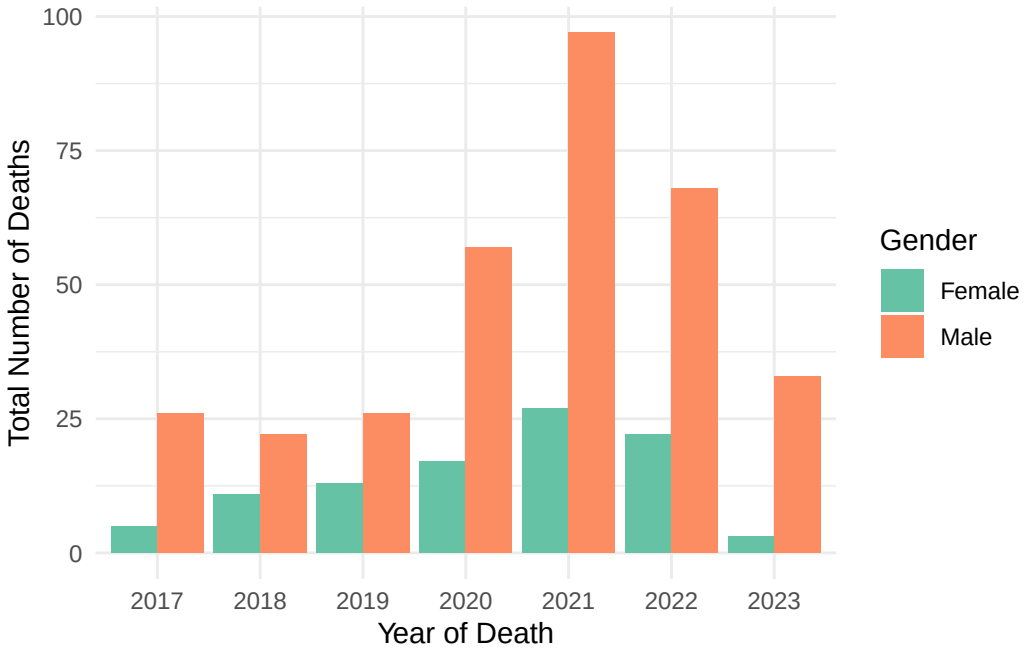


Figure 2: Deaths Per Year per Gender

3 Model

The goal of our modelling strategy is twofold. Firstly, use a poisson distribution and a negative binomial distribution as we have count data.Both distribution can be use on count data, and by using both we are able to compare results. Secondly, use a multilevel or Bayesian model.

I DO NOTKNOW WHAT TO PUT HERE 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 $count_i$ as the count of deaths that resulted from drug toxicity in Toronto. Then $Gender_i$ is the gender for that count of deaths and $Agegroup_i$ is the range of ages for the deaths.

$$\text{Count}_i|\mu_i, \sigma \sim \text{Normal}(\mu_i, \sigma) \tag{1}$$
$$\mu_i = \alpha + \text{Gender}_i + \text{Age\_group}_i \tag{2}$$
$$\alpha \sim \text{Normal}(0, 2.5) \tag{3}$$
$$\text{Gender} \sim \text{Normal}(0, 2.5) \tag{4}$$
$$\text{Age\_group} \sim \text{Normal}(0, 2.5) \tag{5}$$
$$\sigma \sim \text{Exponential}(1) \tag{6}$$

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`.

3.1.1 Model justification

We expect a their to be a larger relationship between males and the number of deaths. This is because there are more males within the homeless population and more males homeless as a result of addiction to drugs. These reason make it clear that that the male relationship is larger than the female.

Secondly, we expect the relationship to be larger between younger individuals and the number of deaths. We expect this result as the older individuals who die when homeless, will die of illness like Covid-19 or cancer. This limits the number of deaths occurring as a result of drugs. This may result in a negative relationship for the older age groups and a positive relationship for the younger age groups.

4 Results

Our results are summarized in Table 3.

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.

Table 3: Explanatory models of flight time based on wing width and wing length

	Gender Only	Age Only	Age and Gender
(Intercept)	1.64	−0.12	−0.89
	(0.10)	(0.44)	(0.47)
GenderMale	0.98		1.10
	(0.11)		(0.11)
Age_group20-39		2.74	2.77
		(0.45)	(0.46)
Age_group40-59		2.77	2.85
		(0.45)	(0.46)
Age_group60+		1.47	1.41
		(0.49)	(0.49)
Num.Obs.	43	43	43
Log.Lik.	−240.304	−202.787	−149.961
ELPD	−249.3	−216.2	−162.5
ELPD s.e.	39.2	39.2	22.5
LOOIC	498.5	432.5	325.0
LOOIC s.e.	78.3	78.4	45.0
WAIC	498.9	432.8	324.3
RMSE	9.93	9.35	7.49

5.2 Second discussion point

5.3 Third discussion point

5.4 Weaknesses and next steps

Weaknesses and next steps should also be included.

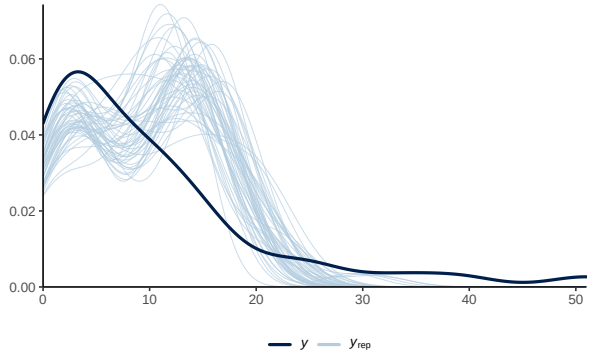
Appendix

A Additional data details

B Model details

B.1 Posterior predictive check

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

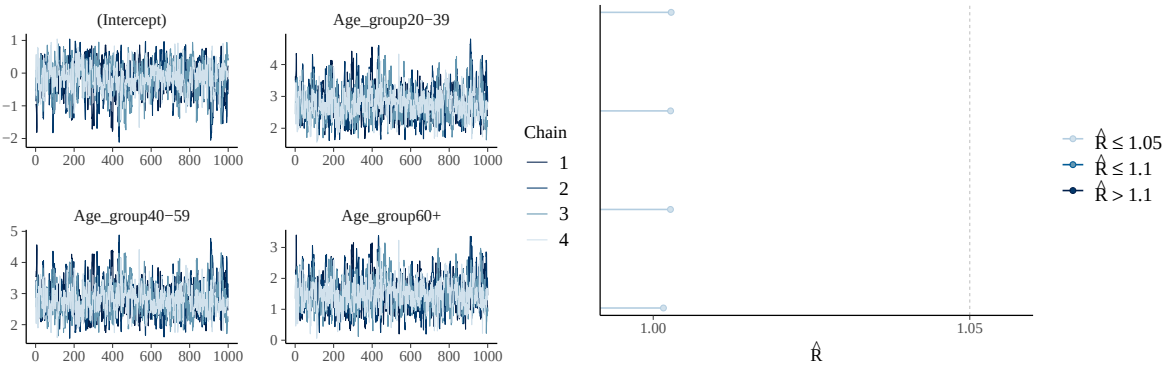


(a) Posterior prediction check

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

B.2 Diagnostics

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



(a) Trace plot (b) Rhat plot

Figure 4: Checking the convergence of the MCMC algorithm

References

Gelfand, Sharla. 2020. *Opendatatoronto: Access the City of Toronto Open Data Portal*.

Goodrich, Ben, Jonah Gabry, Imad Ali, and Sam Brilleman. 2022. "Rstanarm: Bayesian Applied Regression Modeling via Stan." https://mc-stan.org/rstanarm/.

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

user9112767. 2018. "How to Hold Figure Position with Figure Caption in PDF Output of Knitr?" *Stack Overflow*. https://stackoverflow.com/questions/29696172/how-to-hold-figure-position-with-figure-caption-in-pdf-output-of-knitr.

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

———, et al. 2019a. "Welcome to the tidyverse." *Journal of Open Source Software* 4 (43): 1686. https://doi.org/10.21105/joss.01686.

Zhu, Hao. 2021. *kableExtra: Construct Complex Table with 'Kable' and Pipe Syntax*.