

# Death License Registrations May Indicate a Discrepancy in Reported Data During the COVID-19 Pandemic\*

An Analysis of Death License Registrations in Toronto from 2016-2020

Chloe Thierstein

3 February 2023

## Abstract

Death registry statistics offer us a useful tool to better our healthcare system, from monitoring trends to government policy, it is essential to our understanding of the world around us. The COVID-19 pandemic marked a significant shift in the Canadian mortality rate in the year 2020. This paper utilizes death registry data from three civic centres in the greater Toronto area from 2016 to 2020, alongside secondary sources and research. From this work, we demonstrate that even data from considerably reliable sources may become distorted if enough care is not taken in its recording and the world around the data is not considered.

## 1 Introduction

Canadian death registry statistics offer a useful tool in helping better our healthcare system as by monitoring trends in public health such as infectious diseases, suicide, and unintentional injuries they help the healthcare sector provide better services and resources, like screening and prevention programs (Statistics Canada 2022b). However, maintaining this data may not always be done accurately which could cause confusion over our understanding of Canadian demography and health.

In this paper we look at death registry statistics from three civic centres in the greater Toronto area, Etobicoke, Scarborough and North York between 2016 and 2020. From this data, we find that the death licenses provided by these centres during the first year of the COVID-19 pandemic (2020), are much lower than in previous years (Open Data Toronto 2023). This is significant as Canada experienced an increase of 7.7% in deaths in 2020 (Statistics Canada 2022a), and 5.2% more deaths than would typically be expected when taking into account Canada's aging population (Statistics Canada 2021). We reason that this discrepancy in expected data versus reported data could be the result of strain and complications put onto record-keeping during this time similarly to what was seen in early 2019 with long delays in the distribution of death, birth and marriage certificates due to high online demand (Jeffords 2019). Future work could look specifically at how strain on data reporting centres during the COVID-19 pandemic influenced data and mitigation techniques for the future.

This paper will begin with an overview of its data management, source and cleaning. Next we will briefly overview the drop off in trend from previous years in death registrations to begin our discussion. Next we will consider the relationship between death licenses provided at each civic centre over the period of 2016 to 2020, to better understand how the year 2020 differed from past years. Next, we discuss in more detail, death licenses provided by each civic centre to determine a possible cause for the uncharacteristic decline of death licenses in 2020. We will then look at the mean number of deaths from each civic centre to further gauge how these centres relate to rates of death registrations. Finally, we will consider the limitations of the data, including biases.

---

\*Code and data are available at: [https://github.com/cthierst/death\\_registry\\_analysis.git](https://github.com/cthierst/death_registry_analysis.git).

## 2 Data

### 2.1 Data Management

This paper utilizes the R statistical programming language (R Core Team 2020), along with several packages. These packages are, tidyverse (Wickham et al. 2019), janitor (Firke 2021), here (Müller 2020) and dplyr (Wickham et al. 2022). The data being analyzed comes from Open Data Toronto and it is imported using the opendatatoronto package (Gelfand 2022). All figures have been created using ggplot2 (Wickham 2016) and the tables have been created with knitr (Xie 2023) and kableExtra (Zhu 2021), packages. The color styles in the graphs were created by using the RColorBrewer (Neuwirth 2022) package and any graph combinations were made using the Patchwork (Pedersen 2022) package.

### 2.2 Data Source and Cleaning

The data comes from death registrations which are entered into the Registry Services Tracking System (RSTS) by Registry Services staff who are located at three of the civic centres, Etobicoke, North York and Scarborough (Open Data Toronto 2023). It's creation supports the Vital Statistics Act, a Provincial legislation (Open Data Toronto 2023) which involve the collection of deaths, marriages, stillbirths, and live births (Statistics Canada 2022b). This data set is updated monthly (Open Data Toronto 2023). The variables from this data set represent the civic centres; "Etobicoke", "North York", and "Scarborough", number of death licenses registered in the month, place of death; "Outside City Limits" or "Toronto", and time period by month and year in which the death was registered (Open Data Toronto 2023).

To properly analyze this data for the purposes of this paper certain data was removed including, all data for the years 2011-2015 and 2021-2023. This was done to ensure that the data set being worked with had sufficient and consistent data to draw from without being too broad in scope. Additionally, rows from the civic centre variable labelled "Toronto" were removed as death registrations at this location were especially low and its data input, sporadic. Finally, the variable for place of death was not included in the analysis of this data as it was not relevant to its narrative.

### 2.3 Data Analysis

The province of Ontario represents a large portion of deaths in Canada, taking on an estimated 37% of Canadian mortality from 2016-2020 (Statistics Canada n.d.). It is also estimated that on average, that just over 100 thousand people have died each year in Ontario between 2016 and 2020 (Statistics Canada n.d.). We may suggest then, that Toronto makes up a large portion of this mortality rate, as it makes up approximately 44% of the Ontarian population (Statistics Canada n.d.). Additionally, Canada has an aging population, meaning that it's mortality rate is expected to rise every year although it rose exponentially due to the COVID-19 pandemic (Statistics Canada 2022a). Therefore, we may suspect that the COVID-19 pandemic would increase the mortality rate significantly in Toronto as Canada saw an increase of 7.7% in deaths in 2020 (Statistics Canada 2022a).

When looking at Figure 1, however, we can see that from 2016 to 2018 there was an increase in death licenses registered, as expected, followed by a slight drop in 2019. This drop in 2019 may be indicative of a backlog from Ontario's Vital Statistics in early 2019 which was cited as having delays of over three months, waiting for death, birth and marriage certificates (Jeffords 2019). What is most important in Figure 1 is that when looking at the "2020" bar in Figure 1, we can see that the death licenses registered drops significantly in comparison to any of the previous 4 months. However, the amount of death licenses registered each month in 2020 was inconsistent as can be seen in Figure 2.

Figure 1: Death Licenses Registered at Three Toronto Civic Centres from 2016-2020

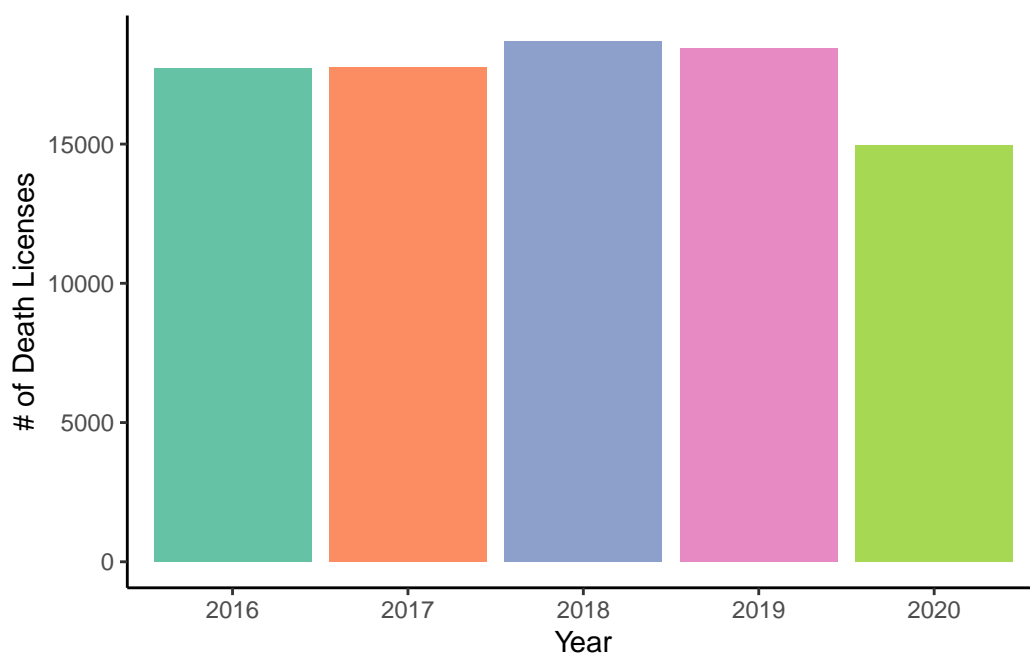
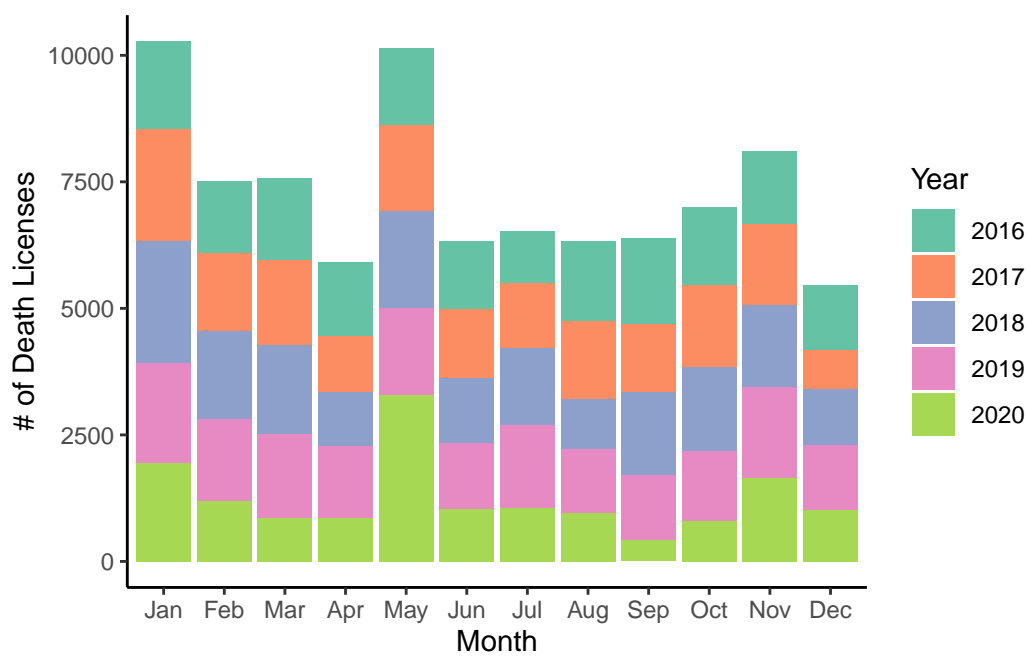


Figure 2: Death Licenses Registered by Month at Three Toronto Civic Centres from 2016-2020



By looking at Figure 2 we can see that, generally, each month is relatively consistent in the amount of death licenses registered from the three civic centres from 2016 to 2019 with a few variations. However, when looking at 2020, we notice significantly lower rates of death license registration each month overall in comparison to past years with the exception of May and November. In the later, we see death license registrations that are nearly balanced with previous years. Whereas in May, there is a large spike in death license registration that far surpasses any month's death license registration in the previous four years.

Looking at May's spike without knowing the context of the years, we may assume that it is a particularly deadly month, however that we can see that the spike in May, shown in Figure 2, is likely indicative of the first wave of COVID-19 in Canada hitting its peak at the beginning of the month before declining (Detsky and Bogoch 2020). This makes sense, however, COVID-19 infections were not the only factor driving the increased death rates in Canada during 2020. There were indirect consequences of the pandemic which lead to the higher death rate in Canada. For example, there were over four-thousand deaths from accidental poisonings which was nearly one-thousand more than in 2019 (Statistics Canada 2022a). By understanding that the increase in the Canadian death rate in 2020 was not only caused by COVID-19 infections, we can see the complexity of the situation. Although the continued decline in death licenses registered from these civic centres during a year that saw an unexpected 5.2% (Statistics Canada 2021) rise in deaths indicates a potential issue in the reporting done at these civic centres.

Figure 3 demonstrates how many death licenses were registered at each civic centre, Etobicoke, North York and Scarborough from 2016 to 2020, both monthly and yearly. From Figure 3 we can see that, generally, Scarborough registers the most death licenses, followed by Etobicoke and finally North York. Interestingly, in 2020, Etobicoke largely overtook Scarborough in death license registrations while simultaneously registered more death licenses than it had in the previous four years. This could indicate that the Etobicoke centre was taking on more work from other civic centres as it was perhaps less overwhelmed during the COVID-19 pandemic. However, more data would be needed to confirm this.

Figure 3: Monthly and Yearly Number of Death Licenses Registered by Each Civic Centre from 2016-2020

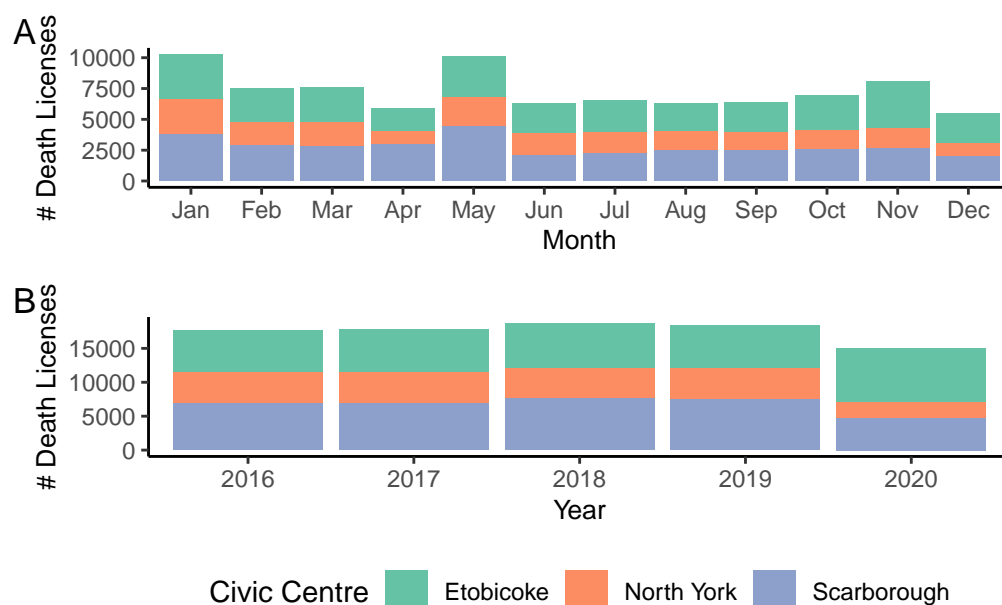


Table 1: Average Number of Deaths Per Month from 2016-2020 at Etobicoke Centre

Month	# of Death License Registrations				
	2016	2017	2018	2019	2020
Jan	342	360	372	346	358
Feb	264	276	275	259	250
Mar	275	281	285	269	237
Apr	236	209	211	237	273
May	315	320	329	320	403
Jun	236	238	235	253	196
Jul	223	250	263	275	235
Aug	241	239	211	230	210
Sep	248	225	251	236	185
Oct	281	275	275	278	226
Nov	348	351	340	356	280
Dec	229	209	233	228	217

Table 2: Average Number of Deaths Per Month from 2016-2020 at Scarborough Centre

Month	# of Death License Registrations				
	2016	2017	2018	2019	2020
Jan	350	366	376	371	359
Feb	266	269	272	289	262
Mar	275	281	273	284	244
Apr	273	262	261	268	299
May	382	396	409	382	419
Jun	214	214	201	206	223
Jul	253	248	263	253	274
Aug	275	279	252	271	288
Sep	252	243	244	236	195
Oct	256	264	258	244	226
Nov	246	256	265	261	301
Dec	214	184	199	203	272

Table 3: Average Number of Deaths Per Month from 2016-2020 at North York Centre

Month	# of Death License Registrations				
	2016	2017	2018	2019	2020
Jan	291	323	330	303	293
Feb	210	214	236	222	193
Mar	224	221	235	225	182
Apr	184	164	155	180	189
May	246	250	259	268	372
Jun	194	193	199	179	180
Jul	176	186	194	211	209
Aug	199	195	165	172	186
Sep	198	181	194	170	135
Oct	183	192	205	176	160
Nov	194	200	204	220	233
Dec	165	139	151	184	189

Many of the trends spoken about previously can be seen explicitly in Table 1, Table 2, and Table 3. These tables present the average number of death licenses registered per month from 2016 to 2020 at each civic centre. These tables allow us to better grasp the variability in death licenses registered each month from 2016-2020. High variations in amount of death licenses registered can be seen mostly between months, although between years these variations become much smaller. This is important as upon first glance one would not assume inherently that the data actually shows much less death license registrations in 2020.

While more data and research needs to be done to fully understand the story of data collection and reporting during the COVID-19 pandemic, this paper looks at a screenshot of death registry data from three civic centres between 2016 and 2020. By doing this, we attempt to better understanding the significance of data collection and how it may be impacted by outside conditions. We suggest that there is data missing from the report due to a variety of factors and reason that pressure and complications but onto data-keeping centres during the COVID-19 pandemic may be to blame. We encourage future inquiries to look specifically at how strain put on data reporting centres during the COVID-19 pandemic influenced data and how to mitigate these strains in the future.

## References

- Detsky, Allan S., and Isaac I. Bogoch. 2020. “COVID-19 in Canada Experience and Response.” JAMA Network. <https://jamanetwork.com/searchresults?author=Isaac+I.+Bogoch&q=Isaac+I.+Bogoch>.
- Firke, Sam. 2021. *Janitor: Simple Tools for Examining and Cleaning Dirty Data*. <https://CRAN.R-project.org/package=janitor>.
- Gelfand, Sharla. 2022. *Opendatatoronto: Access the City of Toronto Open Data Portal*. <https://CRAN.R-project.org/package=opendatatoronto>.
- Jeffords, Shawn. 2019. “Delays in Birth, Death, Marriage Certificates Caused by High Demand, Online Issues.” CTV News Windsor. <https://windsor.ctvnews.ca/delays-in-birth-death-marriage-certificates-caused-by-high-demand-online-issues-1.4305301>.
- Müller, Kirill. 2020. *Here: A Simpler Way to Find Your Files*. <https://CRAN.R-project.org/package=here>.
- Neuwirth, Erich. 2022. *RColorBrewer: ColorBrewer Palettes*. <https://CRAN.R-project.org/package=RColorBrewer>.
- Open Data Toronto. 2023. “Death Registry Statistics.” City of Toronto Open Data Portal. <https://open.toronto.ca/dataset/death-registry-statistics/>.
- Pedersen, Thomas Lin. 2022. *Patchwork: The Composer of Plots*. <https://CRAN.R-project.org/package=patchwork>.
- R Core Team. 2020. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. <https://www.R-project.org/>.
- Statistics Canada. n.d. “Table 17-10-0008-01 Estimates of the Components of Demographic Growth, Annual.” <https://doi.org/10.25318/1710000801-eng>.
- . n.d. “Table 17-10-0135-01 Population Estimates, July 1, by Census Metropolitan Area and Census Agglomeration, 2016 Boundaries.” <https://doi.org/10.25318/1710013501-eng>.
- . 2021. “Provisional Death Counts and Excess Mortality, January 2020 to August 2021.” <https://www150.statcan.gc.ca/n1/daily-quotidien/211108/dq211108a-eng.htm>.
- . 2022a. “Deaths, 2020.” <https://www150.statcan.gc.ca/n1/daily-quotidien/220124/dq220124a-eng.htm>.
- . 2022b. “Frequently Asked Questions on Vital Statistics.” <https://www.statcan.gc.ca/en/about/relevant/vscc/faq>.
- 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, Romain François, Lionel Henry, and Kirill Müller. 2022. *Dplyr: A Grammar of Data Manipulation*. <https://CRAN.R-project.org/package=dplyr>.
- Xie, Yihui. 2023. *Knitr: A General-Purpose Package for Dynamic Report Generation in r*. <https://yihui.org/knitr/>.
- Zhu, Hao. 2021. *kableExtra: Construct Complex Table with ‘Kable’ and Pipe Syntax*. <https://CRAN.R-project.org/package=kableExtra>.