

Decreasing Prevalence of COVID-19 Cases in Toronto from 2020 to 2023 Supports the Lifting of Related Regulations.*

Denise Chang

September 27, 2024

In 2020 the coronavirus disease caused worldwide panic due to its fast transmission rate and the health complications for its affected host. This study aims to investigate the trends in COVID-19 cases in the City of Toronto from 2020 to 2023, as well as the severity of each of these cases. Based on the exploration of the data, there is strong evidence to support the hypothesis that the coronavirus disease is not as relevant today as it was during the peak of the pandemic suggesting the safety of lifting related regulations. The results of this study are significant, as they impact the future directions of COVID-19 regulations for businesses, schools and governmental institutions.

1 Introduction

One of Toronto's iconic scenes is the Toronto islands located in Lake Ontario. The Toronto islands is a network of 15 smaller islands connected by pathways and bridges. (the main way is use the ferry (also a fan favourite) since it is both cheapers, convenient and runs all year round.) EXPLAIN FERRY (which terminal is departs from in dt and what are the three different) However, on the logistic side there are a lot of ressources gone into running this service. A the terminal for example

In this paper, I am interested in the trends in ferry usage to the Toronto Islands by seasonality. Using data from Open Data Toronto (SEE COVID HOW FORMUKATE), I explore (specifically diff in tickets sold and redeemed) for insights on which seasons of the year (scheduled season) is the most and least busy.

*Code and data supporting this analysis are available at: https://github.com/DeniseChang9/Covid-19_Cases.git

The remainder of this analysis is organized into 3 distinct sections. Section 2 discussed the data collected and the studied variables. Section 3 presents the results and findings of the exploration of the dataset with the help of visualized data. Section 4 explores further insights from the results and discusses a few weakness of this analysis. This section also briefly mentions potential next steps following this paper.

2 Data

The statistical programming language R (R Core Team 2023) is used to clean and process the raw data. Specifically, statistical libraries such as `tidyverse` (Wickham et al. 2019), `arrow` (Richardson et al. 2023) and `janitor` (Firke 2023) were used in the data cleaning process, while libraries such as `knitr` (Xie 2021) and `here` (Müller 2020) were used for loading and rendering tables and graphs in this analysis.

2.1 Data collection

The dataset used for this paper is titled “Toronto Island Ferry Ticket Counts” and is published by Toronto’s Parks, Forestry & Recreation division. For the analysis, the dataset is retrieved from the City of Toronto Open Data Portal through the R package `opendatatoronto` (Gelfand 2022).

The data collected is the aggregated data of tickets redeemed and sold at the Jack Layton Ferry Terminal, where ferry departs from the City to one of the three islands: This dataset is updated at 15 minutes interval from 6:30AM to 11:30PM, which are the earliest and latest times where the ferry departs from the city to one of the islands.

The data used for this paper was retrieved on September 25 2024 at YYYPM, and was last refreshed on September 25 2024 at XYZPM.

say what date spans this dataset, what subset we used + why (completeness)

2.2 Variables of interest

2.2.1 Tickets Redeemed

One of the variables we are interested in is the average number of tickets redeemed by per month from 2022 to 2023. Tickets to use the Toronto Islands Ferry can be bought online by using the QR code at the ferry terminal or by buying it in advance on the website. Self-serve customers are then given a bar code that can redeem their ticket in person at the ferry terminal. The number of tickets redeemed are measured by the number of bar codes scanned

in-person at the ferry gate. Many stations are opened to scan the bar codes, so the total amount is compiled in the Toronto Ferry’s internal system.

Table 1: First Ten Rows of the Redeemed Tickets Data

Date	Number of Tickets Redeemed
2022-01-01	503
2022-01-02	345
2022-01-03	532
2022-01-04	475
2022-01-05	225
2022-01-06	358
2022-01-07	369
2022-01-08	543
2022-01-09	313
2022-01-10	224

Table 1 is a sample of the ten first rows of the number of redeemed tickets per day from 2022 to 2023. Each row of this dataset represents a different date, ranging from January 1 2022 (2022-01-01) to December 31 2023 (2023-12-31). The “Date” columns adds a sense of succession to the data, allowing readers to identity trends in redeemed tickets throughout time. The “Number of Tickets Redeemed” column is the total number of tickets bought online that were redeemed on the specified date.

2.2.2 Tickets Sold

Another variable we will be looking at is the number of tickets sold in-person. Customers opts to buy tickets in person at the terminal passes by a physical employee and can make their transaction through them. The employee inputs it into their own internal system and collects payment to complete the transaction. The internal system keeps track and of the number of transactions and tickets sold. The number of tickets sold is measured by the internal ferry system and through the transactions and the amount of tickets per transaction. In the raw dataset, this amount is updated at 15 minutes interval, but for the sake of this analysis, we will be looking at the amount of tickets sold per day.

Table 2: First Ten Rows of the Sold Tickets Data

Date	Number of Tickets Sold
2022-01-01	471
2022-01-02	322
2022-01-03	503

Date	Number of Tickets Sold
2022-01-04	412
2022-01-05	201
2022-01-06	324
2022-01-07	339
2022-01-08	477
2022-01-09	289
2022-01-10	198

Table 2 shows the first 10 rows of the number tickets sold per day from 2022 to 2023. Similarly to the sample of number of tickets redeemed, each row of this sample represents a unique date from January 1 2022 and December 31 2023, making trends in sales more identifiable to readers. The “Number of Tickets Sold” column is the total number of tickets sold in person for the Toronto Island Ferry on the specified date.

3 Results

Putting it in graphs, what do I get? Where are there more, where are there less. SEPARATE THEM BY SEASON HERE.

(Maybe discussion?) => Seasonality... Transport not convenient, also weather? not many activities.

References

- Firke, Sam. 2023. *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>.
- Müller, Kirill. 2020. *Here: A Simpler Way to Find Your Files*. <https://CRAN.R-project.org/package=here>.
- R Core Team. 2023. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. <https://www.R-project.org/>.
- Richardson, Neal, Ian Cook, Nic Crane, Dewey Dunnington, Romain François, Jonathan Keane, Dragoş Moldovan-Grünfeld, Jeroen Ooms, Jacob Wujciak-Jens, and Apache Arrow. 2023. *Arrow: Integration to 'Apache' 'Arrow'*. <https://CRAN.R-project.org/package=arrow>.
- 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>.

Xie, Yihui. 2021. *Knitr: A General-Purpose Package for Dynamic Report Generation in r*.
<https://yihui.org/knitr/>.