

Inspection Frequency Reveals Violation Patterns in Toronto Restaurants and Takeout Establishments (2022-2024)*

Jerry Xia

September 27, 2024

This study analyzes Toronto’s DineSafe dataset (2022-2024) to examine the relationship between inspection frequency and food safety infractions in restaurants and takeout establishments. Results show 73.4% of restaurant inspections lead to infractions, compared to 58.6% for takeout locations, with more frequent inspections correlating with higher infraction counts. These findings suggest that regulatory efforts may need to be more targeted, particularly in full-service operations, to address the higher violation rates and improve overall food safety compliance in Toronto.

Table of contents

1	Introduction	2
2	Data	3
2.1	Overview	3
2.2	Results	5
3	Discussion	8
A	Appendix	10
A.1	Dataset and Graph Sketches	10
A.2	Data Cleaning	10
A.3	Attribution Statement	10
	References	11

*Code and data are available at: https://github.com/Jerryx2020/toronto_dinesafe_analysis

1 Introduction

In large urban centers like Toronto, food safety is a critical public health concern, as food-borne illnesses can pose significant risks to the population. The DineSafe program, managed by Toronto Public Health, enforces health and safety regulations by conducting regular inspections of food establishments, including restaurants, food trucks, and takeout locations. These inspections result in outcomes such as a pass, conditional pass, or closure, depending on compliance with food safety standards (Gelfand 2022). According to the Centers for Disease Control and Prevention (CDC), regular and transparent food inspections are essential for reducing foodborne illness outbreaks, as public posting of results encourages compliance (Centers for Disease Control and Prevention 2024). However, disparities in inspection frequency and the severity of infractions across establishment types, particularly between mobile vendors and traditional restaurants, raise concerns about whether current regulatory efforts are sufficient (Analytics 2023).

This paper analyzes the DineSafe dataset from 2022 to 2024 to explore patterns in inspection frequency and compliance outcomes across different types of food establishments. While previous studies have examined broad compliance trends, this research focuses specifically on how the type of establishment—such as restaurants versus takeout locations—affects the likelihood of infractions. Notably, restaurants tend to have more infractions compared to takeout establishments, suggesting that more stringent oversight may be needed for full-service operations (Agency 2023). Additionally, restaurant vendors exhibit high rates of infractions relative to the number of inspections conducted, further underscoring potential regulatory gaps.

To address the variations in inspection frequency and compliance, the DineSafe inspection data for Toronto’s food establishments was cleaned and analyzed, as described in Section 2.1 and Section A.2. The results, detailed in Section 2.2, show that establishments inspected more frequently tend to have higher infraction counts, suggesting that frequent inspections uncover more violations but may not effectively reduce them (Figure 1). Inspection activity fluctuated monthly from 2022 to 2024, with peaks likely driven by public health crises or regulatory changes (Figure 2). Restaurants, in particular, exhibit a higher percentage of inspections resulting in infractions (73.4%) compared to takeout establishments (58.6%), highlighting the greater compliance challenges faced by full-service operations (Figure 3). These findings emphasize the need for targeted regulatory efforts to improve food safety compliance across different sectors.

The remainder of this paper is organized as follows: Section 2.1 provides a detailed overview of the dataset and methodology; Section 2.2 presents the key findings, including inspection frequency and Section 3 concludes with recommendations for improving food safety oversight in Toronto. An appendix with supplementary materials, including cleaning and sketches, ensures full reproducibility of the results.

2 Data

2.1 Overview

This analysis utilizes the DineSafe dataset from Toronto’s Open Data platform, accessed using the `opendatatoronto` package (Gelfand 2022). The dataset includes inspections conducted from 2022 to 2024, providing detailed information on health inspections of food establishments, including restaurants and takeout locations, throughout Toronto.

The DineSafe dataset used in this study contains 20,481 individual inspections from 2022 to 2024, covering 6,231 unique food establishments across Toronto. Among these establishments, 73.5% are categorized as restaurants, while the remaining 26.5% are takeout establishments. The average duration between consecutive inspections for any establishment is 170 days, highlighting a relatively consistent inspection schedule. Of the infractions recorded, 37.2% are categorized as “Significant”, while 17.5% are “Crucial”, indicating a lower proportion of severe food safety violations. Only 45.3% of recorded infractions are labeled “Minor”.

The key variables in the dataset include:

- **Establishment ID:** A unique identifier for each food establishment in Toronto.
- **Establishment Type:** The type of food establishment, such as “Restaurant” or “Food Take Out.” This variable helps categorize the establishments and compare compliance across different types of food services.
- **Severity:** This variable indicates the severity of food safety infractions, categorized as “C - Critical,” “S - Significant,” “M - Minor,” and “No Infraction.” Critical infractions represent immediate risks to public health, while minor infractions are less severe.
- **Inspection Date:** The date on which an inspection occurred. This variable allows for temporal analysis of inspections over time, such as monthly or yearly trends.
- **Establishment Status:** Indicates whether the establishment passed, failed, or received a conditional pass following an inspection.

They serve as the primary indicators for assessing food safety. These regular inspections are crucial for maintaining food hygiene standards, as highlighted by the CDC, which emphasizes the importance of routine inspections and the public posting of results, such as letter grades, to encourage compliance and reduce foodborne illnesses (Centers for Disease Control and Prevention 2024).

Several variables required construction or manipulation to facilitate the analysis. **Counting Infractions:** For the analysis of infraction rates, the Severity variable was used to construct a count of confirmed infractions for each establishment. Specifically, any row where the Severity was not labeled as “No Infraction” was considered to have a confirmed infraction. This constructed variable allowed for a clearer view of the total number of infractions per establishment. **Infraction Percentage:** To compare infraction rates between establishment types (restaurants

vs. takeout), we calculated the Percentage of Inspections with Infractions. This involved dividing the total number of inspections where an infraction occurred by the total number of inspections for each establishment type. This metric allowed us to assess compliance rates across different food service sectors.

DineSafe is subject to limitations in measurement. For example, while most severe infractions are likely to be well-documented, less serious violations may be underreported, or follow-up inspections may not be conducted in a timely manner. This is a limitation noted in broader studies of food safety compliance systems (Analytics 2023). Other relevant datasets in Toronto, such as the Toronto Public Health Inspection Results and COVID-19 Business and Community Closure Data, were considered but not used due to their broader scope or focus on closures rather than detailed food safety inspections. While these datasets offer insights into public health compliance, they lack the detailed inspection frequency and infraction data that the DineSafe dataset provides (Gelfand 2022; Toronto 2024). The DineSafe dataset offers detailed insights into food safety but has limitations. It may underrepresent infractions in lower-risk establishments, which are inspected less frequently, as previous studies indicate that inspections tend to focus on higher-risk venues (Analytics 2023). Furthermore, not all infractions lead to immediate penalties, meaning some compliance issues may persist undetected (Analytics 2023; Agency 2023). The dataset also does not capture asymptomatic outbreaks or unreported violations, which can affect public health (Agency 2023).

For this study, only the inspection data for “Restaurant” and “Food Take Out” establishments that passed their most recent inspection were retained. This analysis exclusively focuses on data for restaurants and takeout establishments, as these are among the most common food service types in Toronto. This focus allows for an in-depth exploration of the correlation between inspection frequency and the occurrence of infractions in establishments that are deemed to be in compliance. The DineSafe dataset is updated regularly by Toronto Public Health, and is considered open data under Toronto’s Open Data Licence (Section A.3), as long as proper attribution is provided (Toronto 2024).

2.2 Results

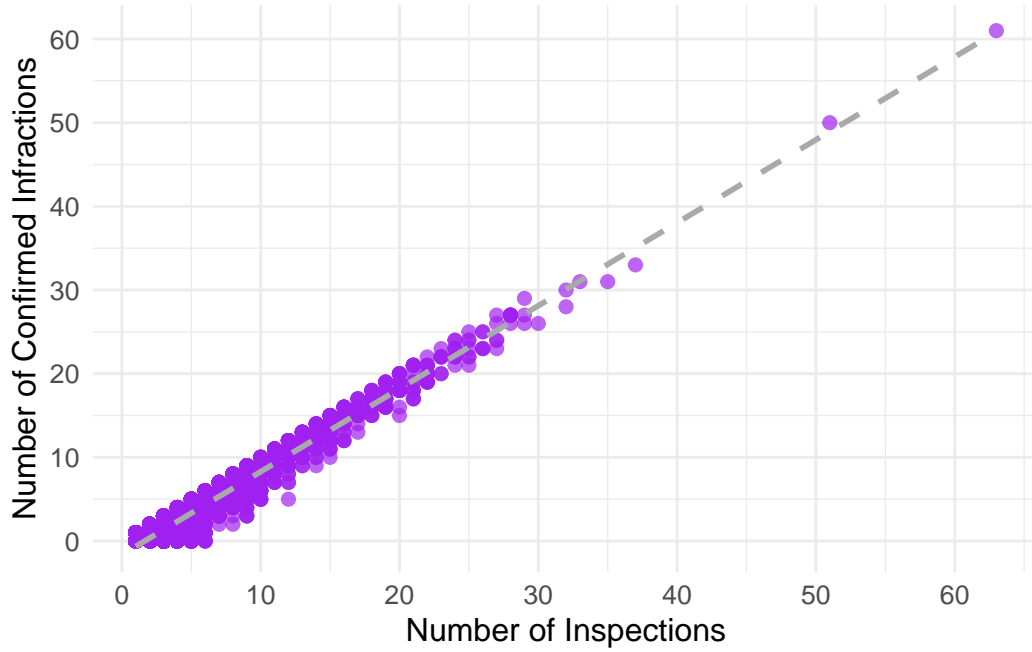


Figure 1: Inspection Count vs Infraction Count per Establishment

As shown in Figure 1, the graph now excludes only inspections where “No Infraction” was recorded. The results show that establishments inspected more frequently tend to have higher counts of infractions. This positive correlation reflects findings from the literature that increased inspection frequency does not necessarily reduce violations but instead highlights pre-existing issues (Public Health 2023; Centers for Disease Control and Prevention 2024).

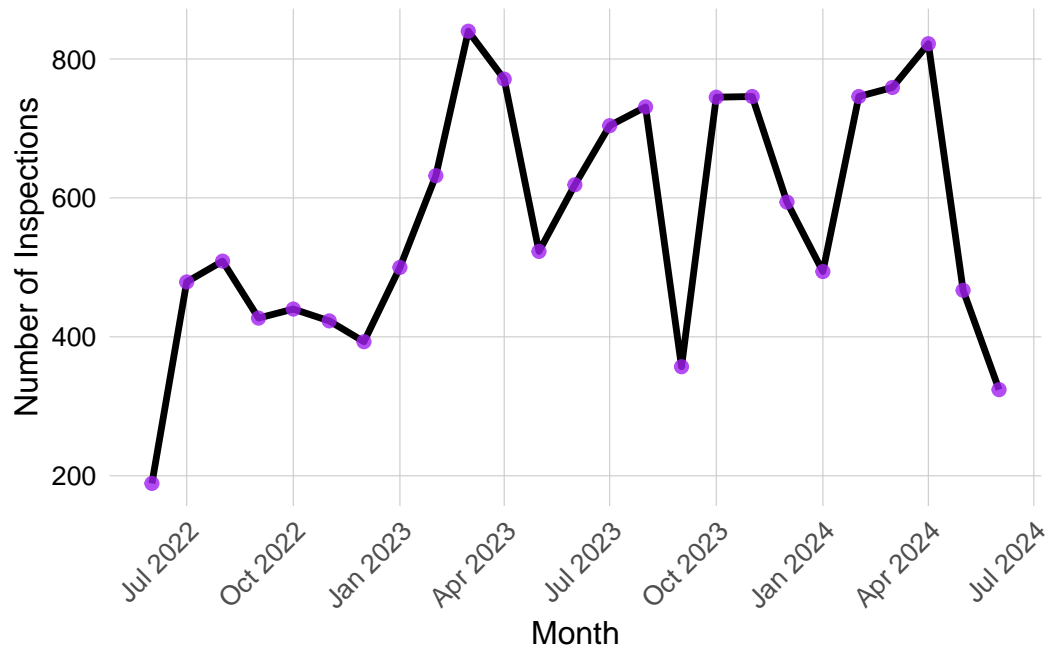


Figure 2: Number of Inspections Conducted Over Time (Monthly)

As shown in Figure 2, the number of inspections fluctuates month by month. Several external factors, such as regulatory changes, seasonal variations, or public health crises like the COVID-19 pandemic, likely influence these inconsistencies in inspection activity (Analytics 2023).

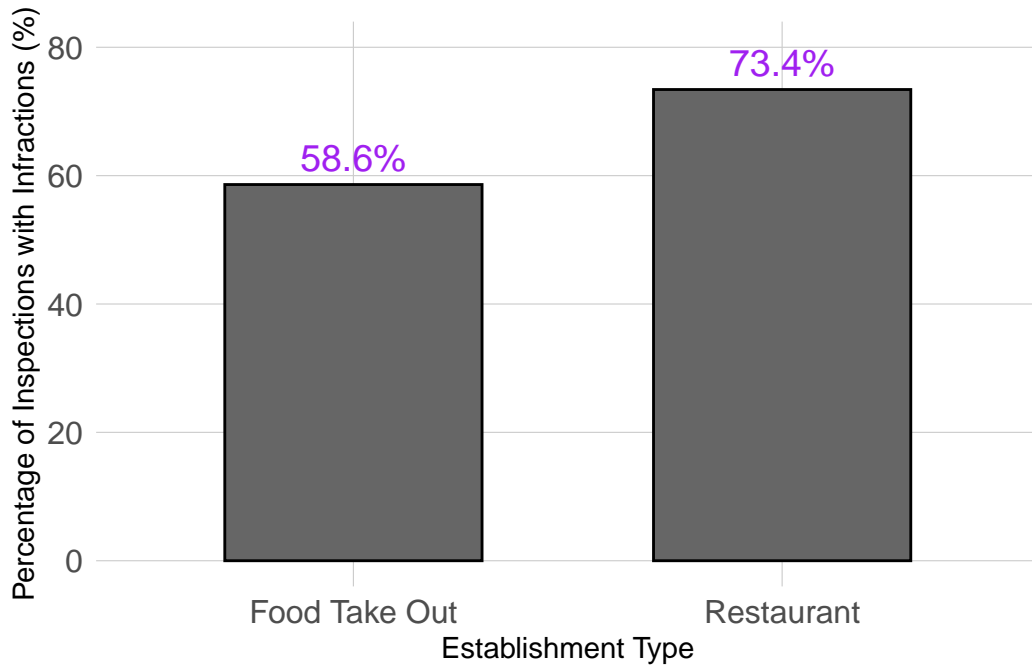


Figure 3: Percentage of Inspections Resulting in Infractions: Restaurants vs. Takeout

As shown in Figure 3, the percentage of inspections resulting in infractions is higher for restaurants (73.4%) compared to takeout establishments (58.6%). This indicates that despite takeout establishments being subject to fewer total inspections due to their smaller numbers, restaurants have a greater percentage of infractions per inspection. This could suggest that full-service restaurants face more complex operational challenges that increase the likelihood of non-compliance with food safety regulations. Research has shown that full-service restaurants, due to the complexity of food handling processes, tend to have higher rates of infractions compared to smaller or more specialized establishments (Analytics 2023; Agency 2023). These findings underscore the need for continued regulatory focus on restaurants, where maintaining compliance appears more difficult, reflecting the operational differences between full-service restaurants and takeout locations (Centers for Disease Control and Prevention 2024). All plots were generated using `ggplot2` (H. Wickham 2016).

3 Discussion

The analysis of the DineSafe data revealed several key insights into food safety practices across Toronto’s restaurants and takeout establishments. One notable observation, as depicted in Figure 1 and discussed in Section 2.2, is the positive correlation between the number of inspections conducted at an establishment and the number of infractions identified. This finding suggests that more frequent inspections tend to uncover more infractions. However, it raises the question of whether these inspections are uncovering ongoing compliance issues rather than fostering improvements. Similar conclusions have been noted in previous studies, which found that increased inspection frequency highlights existing compliance problems without necessarily reducing the overall number of violations (Public Health 2023; Centers for Disease Control and Prevention 2024).

When examining the number of inspections conducted over time, as shown in Figure 2, and detailed in Section 2.2, the data indicates inconsistencies in inspection activity. These variations are likely influenced by external factors such as regulatory changes, public health events (e.g., the COVID-19 pandemic), or operational adjustments (Analytics 2023). While spikes in inspection activity may correspond to heightened public health concerns, ensuring sustained compliance requires consistent follow-up actions. Without continuous enforcement measures, temporary increases in inspections may not lead to long-term reductions in infractions.

Figure 3 in Section 2.2 compares the percentage of inspections resulting in infractions between restaurants and takeout establishments. Restaurants, with a 73.4% infraction rate compared to 58.6% for takeout establishments (Figure 3), face greater compliance challenges, suggesting a need for enhanced regulatory scrutiny. This suggests that full-service restaurants, due to their operational complexity, may face greater compliance challenges with food safety regulations. This observation is consistent with other research that highlights the difficulties full-service restaurants experience in maintaining hygiene and safety standards, as their more complex operations involve multiple stages of food preparation and handling (Analytics 2023; Agency 2023). These findings underscore the importance of focused regulatory oversight on restaurants, where compliance challenges appear more prevalent.

Despite the valuable insights provided by this analysis, certain limitations must be acknowledged. The DineSafe dataset captures only reported infractions during scheduled inspections, and some minor infractions may go unreported due to limited follow-up or enforcement in lower-risk establishments. For instance, Hazel Analytics noted in their 2023 report that establishments classified as low-risk may experience fewer inspections, potentially underrepresenting compliance issues in these sectors (Analytics 2023). This can be easily seen in the data where quick sorting by specific columns results in many thousands of blank entries. Additionally, infractions that do not lead to immediate enforcement actions, such as fines or temporary closures, may allow non-compliance to persist undetected for extended periods (Centers for Disease Control and Prevention 2024). Addressing these limitations will likely require more frequent and unannounced follow-up inspections, as well as extending scrutiny to a wider range of establishments (Agency 2023).

Future research could investigate the application of machine learning techniques to predict establishments most at risk of failing inspections based on historical data, enabling more efficient allocation of resources (Public Health 2023). Further studies examining the impact of enforcement actions, such as fines and mandatory re-inspections, on long-term compliance would provide valuable insights into regulatory effectiveness. Comparative analyses between cities with varying levels of transparency—such as those that publicly disclose inspection scores—could offer additional insights into effective strategies for reducing foodborne illness outbreaks and improving compliance (Agency 2023).

In conclusion, while the DineSafe program has been effective in identifying non-compliant establishments, this study suggests that frequent inspections alone are insufficient to prevent recurring food safety violations. Inconsistent inspection schedules, particularly among high-risk establishments, may further undermine the efficacy of regulatory efforts. To safeguard public health, regulatory bodies should consider implementing more targeted interventions, such as increased oversight of full-service restaurants and sectors with high infraction rates, coupled with stricter penalties for non-compliance (Analytics 2023; Agency 2023).

A Appendix

A.1 Dataset and Graph Sketches

Sketches depicting both the desired dataset and the graphs generated in this analysis are available in the GitHub Repository.

A.2 Data Cleaning

The data cleaning process was essential to prepare the raw DineSafe dataset for accurate analysis. Initially, we filtered the data to focus exclusively on “Restaurant” and “Food Take Out” establishments that had passed their most recent inspection. This ensured that our analysis targeted establishments in good standing, allowing us to assess how inspection frequency correlates with violations in compliant establishments.

Next, we removed irrelevant columns to simplify the dataset and enhance clarity. For example, columns unrelated to inspection outcomes or establishment types were excluded. Additionally, we ensured consistency across the dataset by renaming certain columns for clarity and ease of analysis.

To handle date-related data, we utilized the lubridate package (Grolemund and Wickham 2011), which enabled consistent date formatting and facilitated time-based analysis. The entire data cleaning process was carried out using the tidyverse package (Hadley Wickham et al. 2019), which streamlined the filtering, mutating, and summarizing operations essential for preparing the dataset for further exploration.

A.3 Attribution Statement

“Contains information licensed under the Open Government Licence – Toronto” (Toronto 2024).

References

- Agency, Food Standards. 2023. “The Impact of Displaying Food Hygiene Ratings on Compliance.” *FSA Research and Evidence*. <https://science.food.gov.uk/>.
- Analytics, Hazel. 2023. “Post-Pandemic Food Safety Compliance: Key Insights from Health Inspections.” *Nation’s Restaurant News*. <https://www.nrn.com/food-safety/post-pandemic-food-safety-compliance>.
- Centers for Disease Control and Prevention. 2024. “Posting Inspection Scores Reduces Food-borne Illness Outbreaks in Restaurants.” *Food Safety Insights*. <https://www.cdc.gov/foodsafety/>.
- Gelfand, D. 2022. *Opendatatoronto: Access the Toronto Open Data Portal*. <https://CRAN.R-project.org/package=opendatatoronto>.
- Grolemund, Garrett, and Hadley Wickham. 2011. *Lubridate: Make Dealing with Dates a Little Easier*. <https://CRAN.R-project.org/package=lubridate>.
- Public Health, MDPI Journal of. 2023. “Using Machine Learning to Predict Non-Compliant Food Outlets.” *MDPI Open Access*. <https://www.mdpi.com/journal/publichealth>.
- Toronto, City of. 2024. “Open Government Licence – Toronto.” <https://open.toronto.ca/open-data-license/>.
- Wickham, H. 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>.