Study on The Effectiveness of Government Regulation of Dine Safety*

Ruikang Wang (1008238872)

September 27, 2024

The 'Dinesafe' data was retrieved from Open Data Toronto to study the health and safety of the restaurant industry in Toronto in recent years, as well as government regulatory efforts. The government's approach to food safety inspections, the inspection criteria and the key inspection targets are all very rigorous and reasonable, but food safety is still an issue that needs to be emphasized in Toronto.

1 Introduction

Humans have been searching for food since they were born. Food has been a vital part of human history, whether for sustenance in poverty or for enjoyment in affluence. Contaminated food often brings disease, and in more serious cases can even kill us outright, and the problem is serious enough to warrant global attention (Anagnostou and Abrams 2023). In 2010 alone, Thirty-one foodborne hazards caused 600 million cases of foodborne illnesses and 420 million deaths (Kuchenmüller et al. 2009). The Canadian government clearly recognizes the seriousness of the food safety issue and regularly inspects the vast majority of the restaurant industry. In this paper, we will based on the data in dataset 'Dinesafe' to find out the safer and less safe areas for food and beverage industry in the Toronto area. Moreover, this paper will try to statistically evaluate how reasonable the Toronto government's food safety efforts are and whether they can be improved.

2 Data

In order to get a better overview of the food safety situation in Toronto restaurants and to know what kind of inspections the Toronto government has done in the restaurant industry, i

*Code and data are available at: https://github.com/Rickenn/Dinesafe.

used the "Dinesafe" data from Open Data Toronto (Gelfand 2022). The raw data of "Dinesafe" contains all the inspection results from 2001 to September 24, 2024, but because the amount of data is too large and too far back in time to analyze the data in today's reality, I chose 32,000 data from 2022 to 2024 years to generated a raw data file. Even so, there's still tons of data in this data that I don't need, such as "_id", "Inspection Date", and so on. In order to make the data more concise, I started cleaning data by the statistic software R (R Core Team 2023), tidyverse (Wickham et al. 2019), and here (Müller 2020). Then this clean data only contains the variables I need like "Severity", "Latitude", "Longitude " and so on.

In order to get a simple idea of the overall hygiene of restaurants in Toronto, I have categorized each restaurant into five types of hygiene problems based on the data: "S - Significant", "C - Crucial", "M - Minor I categorized each shop according to the data into "S - Significant", "C - Crucial", "M - Minor", "NA" and "NA - Not Applicable". "S - Significant" is the most serious problem and 'NA' means this restaurant has no problem, and counts the number of results detected by the Toronto government for each.

Table 1: Number of Occurrences of Each Category in Severity

Severity	count
C - Crucial	849
M - Minor	11837
NA - Not Applicable	1395
S - Significant	6090
NA	11829

According to Table 1 we can find that the restaurants with light health problems (M - Minor) in the Toronto restaurant industry is the most, there are 11,837, accounting for 36.99% of the total number of inspections; the second highest percentage is the restaurants that fully passed the health inspection (NA), there are 11,829, accounting for 36.97% of the total number of inspections, and the rest of the order of the number of stores in order of quantity are the stores with the most serious problems (S - Significant), 6090, 19.03%; stores whose inspection standards do not allow reasonable detection of the store's condition (NA - Not Applicable), 1,395, 4.36%; and stores with more serious problems (C - Crucial), 849, 2.65%.

3 Results

Starting with the conclusion that the Toronto government takes the issue of restaurant sanitation seriously, and that the inspection methods and inspection standards may be reasonable. Because of Table 1, we know that of these 32,000 tests, only 1,395 tests failed to produce a reasonable result, which are NA - Not Applicable. Conversely, more than 95% of these 32,000 tests produced an appropriate result in accordance with existing testing standards, which could

be considered a high success rate. However, does this show that the Toronto government has done a perfect job in food safety regulation? The answer may be no. This can also be clearly understood from Table 1. If the Toronto government is excellent in managing the hygiene of individual restaurants, there should not be so many test results showing that the restaurant has enough hazards to jeopardize human health. Of the 32,000 tests, a total of 6,939 showed that the restaurant would jeopardize the health of the customers, which were S - Significant and C - Crucial, accounting for 21.68% of the total number of tests. By probability, for every five times a customer obtains food at a location other than home, they jeopardize their health once.

4 Discussion

4.1 Distrubution of Offending Restaurant

Since there is nothing wrong with the testing methods and testing standards, what are the reasons for the public's food hygiene not being well protected? My guess might be that it's because the addresses highlighted for inspection are not the same as the ones that really have serious hygiene issues. In order to test this assumption, we first have to know in which region the food hygiene problems are occurring. Use the dataset "Dinesafe" (Gelfand 2022) to flit out the data whose "Severity" is "S - Significant", and use their "Latitude" and "Longitude" as x-axis and y-axis, and then we will generate Figure 1, which is a scatter plot with latitude ranging from 43.55 to 43.90 and longitude ranging from -79.7 to -79.1, where each point is the geographic location of the store with the most serious food hygiene problems, by ggplot2 (Wickham 2016).

Figure 1 shows that a large number of restaurants with serious health problems are clustered in the range of latitude [43.63:43.68] and longitude [-79.45:-79.35], and that there is a clear straight line spread from latitude 43.65 to 43.80 on longitude -79.4. Now we use the same method to make a map of the location of restaurants with more critical health issues (C - Crucial), Figure 2.

Unfortunately, because the amount of data for C - Crucial is so much less compared to S - Sgnificant, based on Figure 2 we can conclude that there are some C - Crucial restaurants clustered around (43.65, -79.4), but that hardly tells us anything. To compensate for the small amount of data, this time we choose to graph the M - Minor restaurant, which has the largest amount of data, and obtain Figure 3.

This time we find that M - Minor Restaurant not only has a very similar coordinate range to S - Sgnificant Restaurant, but also has the straight line trace propagation between [(43.65,-79.5):(43.70,-79.25)], [(43.70,-79.45):(43.75,-79.2)]. Combining these three graphs illustrates that the vast majority of stores detected with hygiene problems are distributed in the range of latitude [43.63:43.68] versus longitude [-79.45:-79.35], and also locate as a stright line from latitude 43.65 to 43.80 on longitude -79.4.

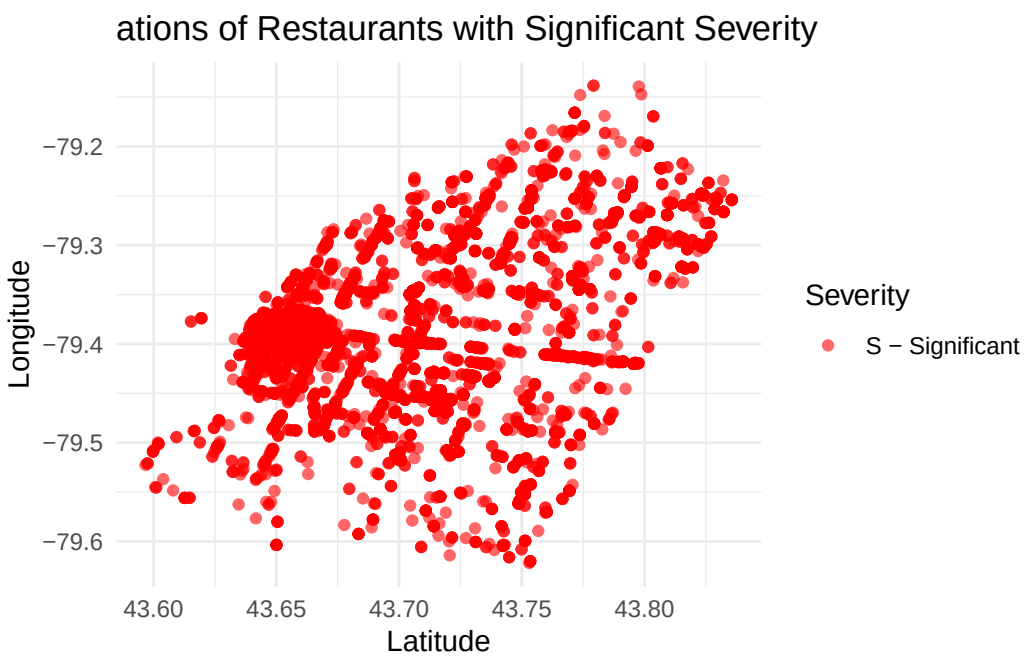

Figure 1

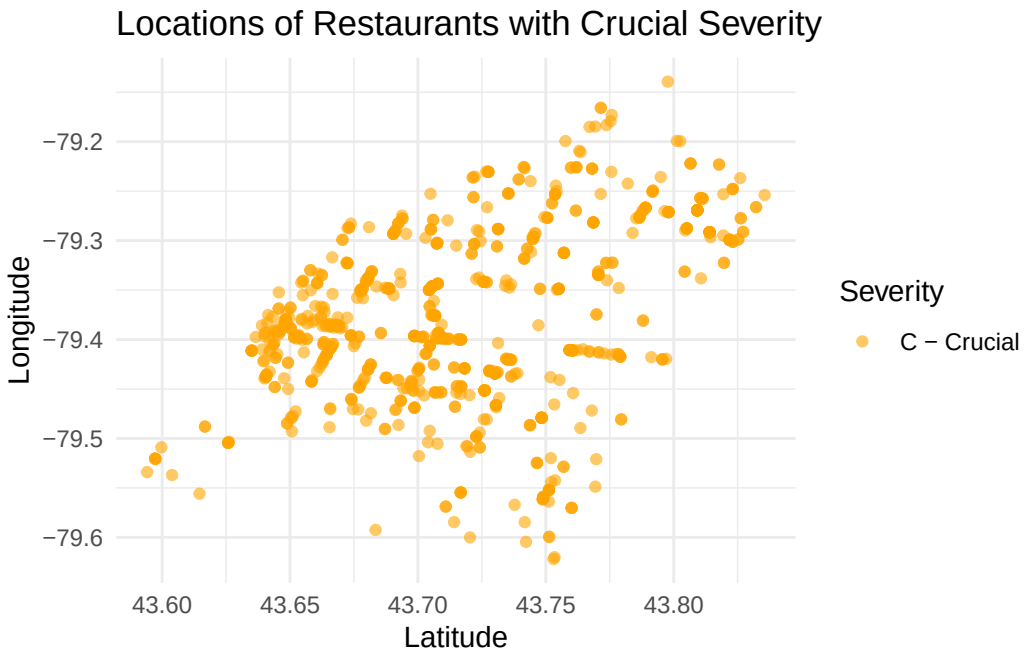

Figure 2

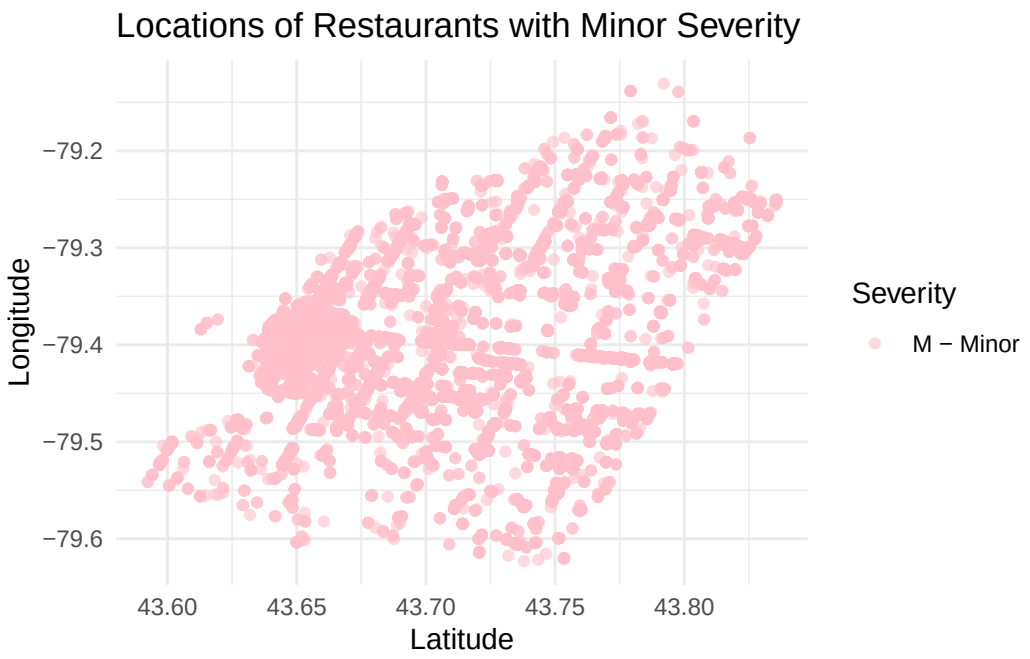

Figure 3

4.2 Focus of Inspections

After figuring out the distribution of restaurants, we need to find those places that the government focuses on monitoring. I took the restaurants in the dataset where "Min. Inspections Per Year" was 3 and 2, divided the 3's into one group and the 2's into another, and plotted them against the latitude and longitude of the corresponding data. I got Figure 4 and Figure 5.

We were surprised to find that whether it was the address of the restaurant that the government focused on three times, or the address that was inspected two times, they were all equally in the range of latitude [43.63:43.68] versus longitude [-79.45:-79.35], and also locate as a stright line from latitude 43.65 to 43.80 on longitude -79.4.

4.3 Why Not Perfect?

By comparing the five figures above, we can see that Figure 4 is not only very similar to Figure 1, but also covers the points of Figure 2. And at the same time, Figure 5 is almost identical to Figure 3. This shows that the government has not only focused on inspecting S - Sgnificant and C - Crucial restaurants, even M - Minor restaurants have not been missed. This shows that the inspection methods of government, the inspection criteria and the targets of inspection are all reasonable. Then there is only one more reason why the food safety problem has not been fully solved, and that is the lack of punishment by the government. For the

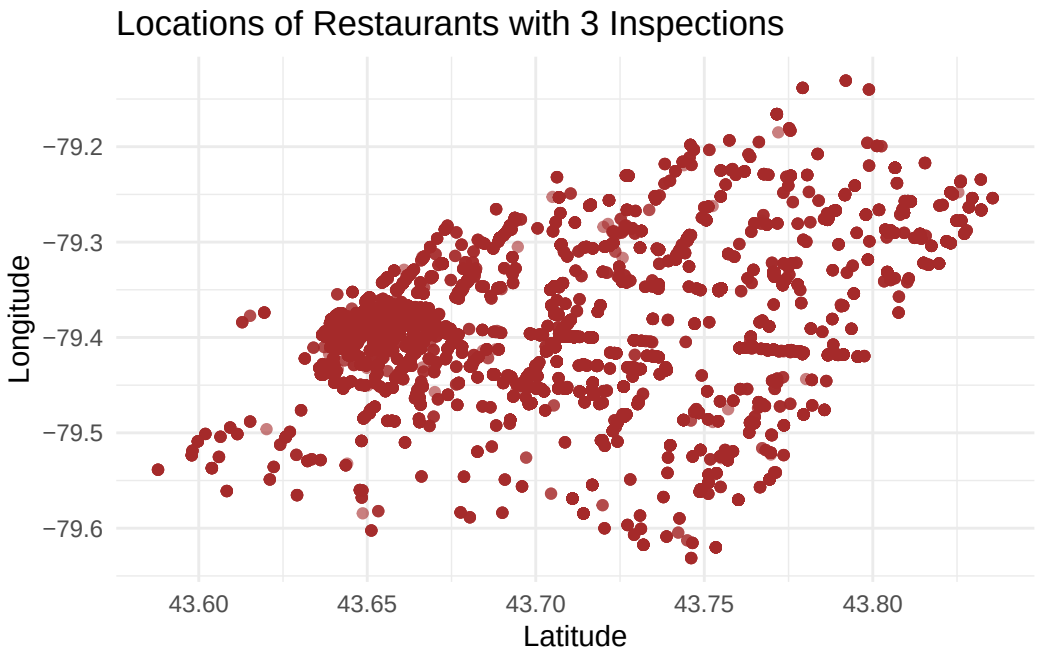

Figure 4

Figure 5

offending restaurants, the cost of rectifying their mistakes is much higher than the cost of government penalties, so they choose not to rectify the situation. If the Government intends to tackle the problem of food safety, it needs to increase the amount of penalties and step up enforcement.

4.4 Weaknesses and next steps

The data and images only show that there is no error in the way the government inspects, the inspection criteria, and the key inspection targets, and the belief that the government is not inspecting enough is only an assumption, and there is not enough data to prove this now. As a next step, we need to look into the data related to government penalties and compare it with the results we are now coming up with.

References

Anagnostou, Aikaterini, and Elissa Abrams. 2023. "Is It Food Poisoning or Allergy?: A Diagnostic Challenge Revisited." *Annals of Allergy, Asthma & Immunology* 130 (4): 403–4. https://doi.org/https://doi.org/10.1016/j.anai.2022.12.019.

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

Kuchenmüller, T, S Hird, C Stein, P Kramarz, A Nanda, and AH Havelaar. 2009. "Estimating the Global Burden of Foodborne Diseases-a Collaborative Effort." *Eurosurveillance* 14 (18): 19195.

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/.

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.