Bon Appétit: Ontario Restaurants that Opened for Indoor Dining had Significantly More COVID-19 Cases Traced Back to their Business Compared to Restaurants who Remained Take-Out/Delivery Only in March 2021

Laura Cline, Lee Doucet, & Will Trefiak

19/02/2021

Abstract

This paper uses multiple linear regression and multiple logistic regression to explore the relationship between opening Ontario restaurants for in-door dining and the number of COVID-19 cases traced back to that restaurant in March 2021. We used simple random sampling on 38,000 Ontario restaurants in the Canadian Business Registries to pull a sample of ~18,900 restaurants who were split into treatment and control groups. The paper exposes that restaurants who received the treatment - open for indoor dining - had significantly more COVID-19 cases traced back to their location compared to restaurants who did not receive the treatment. Therefore, the paper concludes that the Ontario restaurants should keep restaurants shutdown for indoor dining in order to minimize the spread of the virus and protects Ontarians health.

Contents

1 Introduction

As Public Health Ontario has previously identified restaurants as a primary site of virus transmission and that they should be closed for in-person dining, Public Health Ontario is now recommending to the Ministry of Health that with conditions improving they can begin re-opening restaurants for in-person dining. Further, recommendations state opening at a limited capacity, all staff and customers must wear masks, and contact information must be provided to help monitor community spread. In response, the Ministry of Health has decided to give the order to begin the re-opening of in-person dining in restaurants. Locations that are eligible to participate will be determined by a random lottery draw. Community spread will be monitored and if there is not a corresponding increase in cases of Covid-19, additional restaurants will be allowed to participate.

In this paper, we will use a Randomized Controlled Trial (RCT) to analyze the relationship between opening some Ontario restaurants for indoor dining (treatment group) and the remaining restaurants will remain shuttered to indoor dining. We used simple random sampling to gather a sample to distribute our survey which gathers information about our intervention. The World Health Organization (WHO) defines "contact tracing" as the "process of identifying, assessing, and managing people who have been exposed to a disease to prevent onward transmission." In 2019, Statistics Canada identified over 38,000 legally registered restaurants in Ontario alone and all of these restaurants have been impacted by COVID-19 rules such as restrictions on indoor dining, mandatory mask use, and enhanced sanitation protocols. We will study if opening some restaurants for indoor dining in March 2021 impacts the number of COVID-19 cases traced back to that establishment. We hypothesize that opening restaurants for indoor dining will significantly increase the number of COVID-19 cases traced back to that restaurant because customers must take off their masks to consume their food which increases the risk of spreading the disease in a closed and not adequately ventilated space. We chose opening restaurants for indoor dining as our intervention because the Ontario government is currently moving more Ontario regions into "green zones" where restaurants can be open for indoor dining. Future work could expand upon this study by looking performing a similar study in other provinces across Canada which may have different COVID-19 rules for restaurants which lead to a more nuanced understanding about our intervention and possibly discover other variables contributing to COVID-19 case counts.

The code and data supporting this analysis is available on our GitHub repository for this project: https://github.com/lauracline/Opening-Ontario-Restaurants-for-Indoor-Dining-and-Contact-Traced-COVID-19-Cases.

The remainder of this paper is structured as follows: Section 2 discusses the data, intervention, data gathering methodology, data ethics, descriptive data analysis, and determines if the treatment and control groups are representative samples. Section 3 performs linear regression and logistic regression to test our hypothesis. Finally, Section 4 discusses the study's findings and some weaknesses.

2 Data Description

2.1 The "Name of Our Survey" Dataset

The link to our survey to measure the the affect of our intervention between our treatment and control groups are: https://docs.google.com/forms/d/e/1FAIpQLSdRF-dAem7l3eBsrYyCuOSextEwQusLNHs4r5JR9H-E3Tn4FA/viewform.

2.2 Intervention

2.2.1 Intervention Design

The Ontario Provincial government plans to reopen restaurants for in-person dining as soon as it is feasible to do so while ensuring the data supports the reopening will not put communities at risk of spreading Covid-19. As total daily case counts are declining after the peak of the second wave, Ontario Health believes this is a good opportunity for an intervention to test the impact of reopening in-person dining in select restaurants in Ontario. To implement this intervention, a Random Control Trial (RCT) has been selected to allow certain restaurants to open their in-person dining. Health officials will conduct this intervention for two-weeks and then close all in-person dining again while monitoring the daily case counts for an additional two weeks to determine the impact on community spread of Covid-19. The intervention will take place from March 1st to March 15th and then we will wait until March 29th to distribute the survey. Our participants will received the survey on March 29th and they will have until April 4th to return the survey. This time frame was chosen as it is consistent with the current known medical knowledge of 5 to 14 days for the duration of symptoms of Covid-19 to appear in individuals. Based on the data from the intervention, Ontario Health will evaluate the extent, if any, that the opening of in-person dining lead to an increase of Covid-19 community spread. If there is limited cause for concern, Ontario Health will recommend to the Ontario Provincial government that they can re-open all restaurants or in-person dining.

2.2.2 RCT Selection Process

Through the implementation of an RCT, all legal restaurants in Ontario will be randomly divided into two groups with the same probability of being chosen. Restaurants selected for the treatment group will be allowed to reopen their in-person dining while those in the control group will have their in-person dining remain closed. The control group is the counterfactual, what they represent in this case what happens to community spread of Covid-19 when certain restaurants are not opened for in-person dining. This methodology is both statistically fair while providing a transparent and simple enough process that the general public can understand why that only certain restaurants are being allowed to open for in-person dining. The ease of sorting group members by random chance with the same probability of selection reduces the opportunity for bias, both known and unknown, with group selection by researchers. The usefulness of RCT's have made them the "gold standard of impact evaluation" (Paul Gertler 2016).

2.2.3 Randomized Controlled Trial

An RCT will provide all restaurants the same random chance of being chosen for the intervention which reduces the bias between who becomes selected for the treatment and control group. With the large number of restaurants in Ontario, randomly assigning each one to a treatment or control group will produce a high probability of both groups being statically identical. Which is important for robust results as the RCT's will not only randomly distribute restaurants based on observed characteristics such as size and location, it will also distribute them on harder to measure characteristics such as motivations, and desire for Covid-19 compliance. For more considerations regarding the sampling methodology of our data set and how we ensured randomness in our sample, please proceed to the following section on sampling methods.

2.2.4 Null and Alternative Hypotheses for our Intervention

To better capture our intervention in statistical terms, the following null and alternative hypothesis is included:

H0: There is no relationship between the number of contact-traced COVID-19 cases and if a restaurant has been opened for indoor dining.

H1: There is a relationship between the number of contact-traced COVID-19 cases and if a restaurant has been opened for indoor dining.

2.3 Population, Frame, and Sample

The population in question for this survey is all restaurants in Ontario, and our sampling frame are Ontario restaurants registered under the Canadian Business Registries. From this, our final survey sample will be Ontario restaurant owners who respond to an electronic questionnaire sent via. email. As an incentive, we will inform survey candidates that their responses will help shape policies being put in place by the Ontario Provincial government geared towards relieving the economic impact restaurants face during the COVID-19 pandemic.

2.4 Sampling Strategy

As a means of mitigating bias in our survey, our sampling methodology is somewhat twofold, with the first sample attempting to mitigate the biases introduced in the 'second.' Furthermore, we cannot use cluster or stratified sampling because we are unable to determine whether a restaurant belongs to a particular strata or cluster because this data is not available in Canada's Business Registries. Instead, we are using a *simple random sample without replacement* (SRSWOR) whereby the sample is randomly selected from a subset of the population. Practically, this means each sample observation is individually selected with an equal probability of being drawn. The practical sampling procedure utilized to draw a sample from our population comes from the tidyverse package in the R statistical programming language (Wickham et al. 2019).

In this sampling method, each member of the population has an exactly equal chance of being selected for the survey. This is the optimal sampling method for this survey because it involves a single random selection and requires little knowledge in advance about the demographics of the population. This method also has high internal validity because it reduces the impact of potential confounding variables. With a large enough sample size, a simple random sample also has high external validity because it will represent all the characteristics of the larger population. The sampling design we utilized comes from Wu and Thompson (Wu and Thompson 2020) and will assume population values are fixed, which they will be at the time the sample is drawn from the Canadian Business Registries. It is also assumed that our SRSWOR satisfies the following probability measure:

$$\mathscr{P}(\mathbf{S}) = \begin{cases} 1/\binom{N}{n} & \text{if } |\mathbf{S}| = n\,, \\ 0 & \text{otherwise}\,. \end{cases}$$

Figure 1: Where P(**S**) is the sampling method, N/n is the total size of survey candidates of population N

As previously noted, given a large enough sample size, estimators such as the sample mean (average) and sample variance ('spread-outedness' of the data) can aid us in making inferences about our population as a whole, the sample mean and sample variance can be written as:

$$\bar{y} = \frac{1}{n}\sum_{i\in\mathbf{S}} y_i \quad \text{and} \quad s_y^2 = \frac{1}{n-1}\sum_{i\in\mathbf{S}}\left(y_i - \bar{y}\right)^2 .$$

Figure 2: where i is units sampled and **S** is the sample

Given the above statistical expressions that define the sample mean and probability measure, estimating our population mean through SRSWOR can be achieved as follows when yhat(**S**) and P(**S**) are expanded into their whole expressions and subsequently manipulated in the following proof from Wu and Thompson:

For our purposes, this proof gives us sufficient reason to believe the simple random sampling method we want to use is in fact random and will tell us information that is relevant to the population itself.

While not necessarily a second stage, since we are relying on data gathering techniques that require voluntary response to survey sent via. email, it is still important to note the biases such a selection process can bring with it. In a word, while a simple random sample was introduced to reduce bias, self reporting may reintroduce it to a lesser degree. This could be potentially compounded by the offered incentive, which may provide a larger pool of businesses that have been adversely impacted by the COVID-19 pandemic than those

$$
\begin{aligned}
E\left\{\bar{y}(\mathbf{S})\right\} &= \sum_{\mathbf{S}:\ \mathbf{S}\in\Omega} \bar{y}(\mathbf{S})\mathscr{P}(\mathbf{S}) \\
&= \binom{N}{n}^{-1} \sum_{\mathbf{S}:\ |\mathbf{S}|=n} \frac{1}{n} \sum_{i\in\mathbf{S}} y_i \\
&= \binom{N}{n}^{-1} \frac{1}{n} \sum_{i=1}^{N} \sum_{\mathbf{S}:\ i\in\mathbf{S}} y_i \\
&= \binom{N}{n}^{-1} \frac{1}{n} \sum_{i=1}^{N} \binom{N-1}{n-1} y_i \\
&= \frac{1}{N} \sum_{i=1}^{N} y_i = \mu_y \,.
\end{aligned}
$$

Figure 3: Proof the sample mean is a sufficient estimator of the population mean

who have not. While this is a minor consideration given the emphasis placed on conducting a SRSWOR is a sufficient size in our initial sampling phase, it is still worth noting as a possible bias in the data.

2.5 Data Ethics

In addition to potential biases this data may be subject to, the ethical considerations of our sample data must also be considered carefully as we are dealing with information about a struggling Ontario business community as well as asking questions geared specifically towards racialized groups in Canada. Because of this, our sample data has a unique set of considerations that lie at the intersection of data privacy, equity, and access. First, and in all cases, to protect the individual privacy of respondents, all surveys are anonymous and neither business nor owner names are recorded. In addition, location data will not be collected with enough specificity to identify survey respondents.

In terms of handling the ethical concerns related to our survey data, danah boyd, Solon Barocas, Kate Crawford, and number of other scholars provide guidance in their article "10 Simple Rules for Responsible Big Data Research" (Matthew Zook 2017). The 10 rules can be understood as follows:

1. Acknowledge that data are people and can do harm
2. Recognize that privacy is more than a binary value
3. Guard against the reidentification of your data
4. Practice ethical data sharing
5. Consider the strengths and limitations of your data; big does not automatically mean better
6. Debate the tough, ethical choices
7. Develop a code of conduct for your organization, research community, or industry
8. Design your data and systems for auditability
9. Engage with the broader consequences of data and analysis practices
10. Know when to break these rules

In essence, our sampling strategy, as well as the approach we take in this entire survey design, reflects these 10 rules robustly. First, privacy will be protected through ensuring the survey questions are not designed to collect personally identifiable information, and that no questions regarding location have too much specificity. For example, our survey will gather data on the municipality or region of the restaurant as well as whether the restaurant is rural or urban - but will not be recording addresses or postal codes. Data for this survey will be collected via. Google Forms, a commonly used and secure survey platform that allows researchers to collect survey data electronically with ease.

In terms of ethically sharing our data, the entirety of this survey design will be made publicly available via. Github repository as a means of ensuring our work is reproducible as well as methodologically robust. Utilizing a version control tool such as Github also has the advantage of giving our work (read, the associated systems and data) for increased audibility and transparency, a crucial component of any ethical data work.

In terms of the "tough ethical choices" our team debated while designing this survey, there are two I would like to discuss. For a large majority of this survey design, we were unsure how to proceed with classifying racialized groups in Canada, and wavered back and forth on using the term BIPOC (black, indigenous person of colour) in the survey itself. While one of our initial objectives was to understand the impact COVID restrictions had on BIPOC restaurants, we ran into the issue of the term BIPOC being not readily defined (Wilson 2020). This was also compounded with the fact that the term BIPOC may be unfamiliar some some of our survey respondents, and why we opted to use language of "racialized" groups instead. This certainly will have an impact on the data our survey will be collecting as the inclusion of other racialized groups in our treatment group means the scope is no longer contained to BIPOC businesses. Regardless, our original intent of exploring the impact of COVID restrictions from a perspective of equity and advocacy for racialized groups

Table 1: Total Number of Restaurant Owners in Ontario (March 2021)

Number of Ontario Restaurant Owners
38000

Table 2: Does Your Business Have the Option for Indoor Dining at Your Current Location? (March 2021)

Answer	Number of Restaurants
Yes	37632
No	193
Very Limited (Under 5 Tables)	175

in Canada, who in countless other contexts have been affected by the pandemic the most (Bascaramurty 2021).

The second "tough ethical choice" our team had to make surrounded whether we would try to capture or account for businesses not operating with a business license. While the notorious Adamson BBQ was the first of these types of businesses to come to our minds, upon more discussion we realized there may be a significant chunk of businesses currently operating without licenses in Ontario that may be doing so due issues of systemic inequality and lack of access to business resources. While we discussed that this would most certainly be data that is crucial to collect on an often overlooked population, the fact that our sampling frame is restaurants registered in the Ontario Business Registry means we limited our capacity to study this population from the outset. Having said that we understand our inferences can only be made about a (registered) subset of Ontario's entire restaurant population, which also includes ones that are unregistered.

2.6 Descriptive Data Analysis

2.6.1 38,000 Restaurants in our Sampling Frame

Although we cannot know our entire population of our data frame, we can predict that our sampling frame is ~38,000 restaurants because they is the number of registered restaurants in Ontario. Table 1 counts the number of restaurants in our population (Table 1). Using the information from the Canadian Business Registries and our survey model we can explore the demographics and characteristics of our overall population.

The table was created using R (R Core Team 2020), tidyverse (Wickham et al. 2019), and kableExtra (Zhu 2020).

2.6.2 Over 99% of Ontario Restaurants Have Indoor Dining Space in their Locations

Table 2 shows that ~99% of restaurants can have indoor dining at their current location (Table 2). Consequently, out intervention can be used on the vast majority of our respondents and we can filter out the survey respondents who did not have indoor dining capacity.

The table was created using R (R Core Team 2020), tidyverse (Wickham et al. 2019), and kableExtra (Zhu 2020).

2.6.3 Only ~30% of Ontario Restaurts are Owned by Non-White Visible Minorities

The first graph (Figure 4) demonstrates that the majority of Ontario restaurant owners are not BIPOC visible minorities. The graph was built using R (R Core Team 2020), tidyverse (Wickham et al. 2019), and ggplot2 (Wickham 2016).

The graph demonstrates the diversity within Canada's restaurant industry, but does not reveal the differences in access to capital, government programs, or differences in how these two groups were impacted by the lockdown. For instance, Vidya Rao's article (2020) on the impact of California's lockdown on restaurant

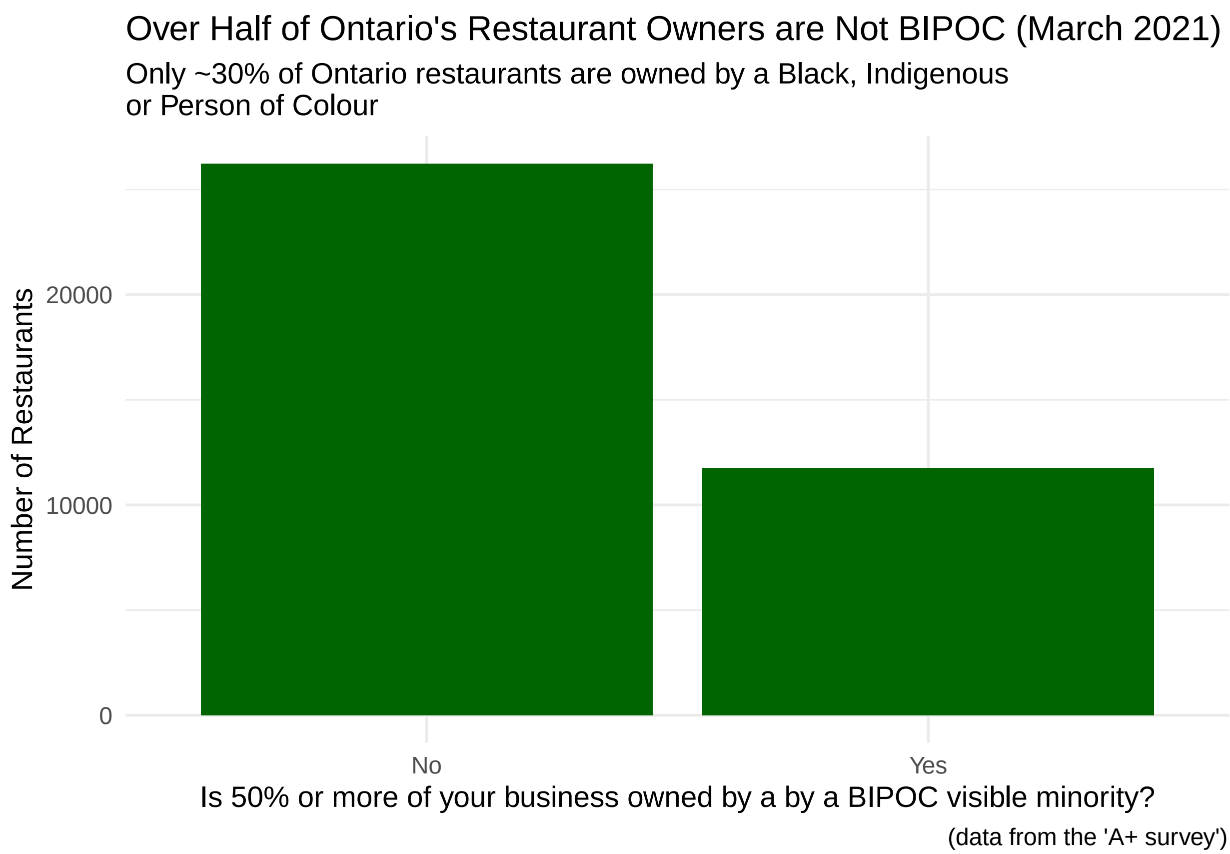


Figure 4: Is 50% or more of your business owned by a BIPOC visible minority? (March 2021)

owners in San Francisco revealed that 41% BIPOC-owned restaurants permanently closed in 2020 because these groups have less access to bank credit and loans due to banks' long history of discrimination such as red-lining, predatory subprime loans, and other exclusionary practices. Additionally, BIPOC-owned restaurants are typically located in ethnic community centers and their customers typically prefer indoor-dining in order to interact with other community members. Since indoor dining is eliminated due to lockdowns, BIPOC restaurants have had less customers and sales (Rao 2020).

Similarly, the graph (Figure 5) below illustrates that the majority (~31%) of Ontario visible restaurant owners are South Asian, followed by Chinese, Black and Multiple Visible Minority. The data reveals that the distribution of restaurant owners in Ontario is similiar to the overall ethnic distribution for the province. This means that there are a proportional number of restaurant owners by ethnicity in Ontario and there is an equitable distribution. The graph was built using R (R Core Team 2020), tidyverse (Wickham et al. 2019), and ggplot2 (Wickham 2016).

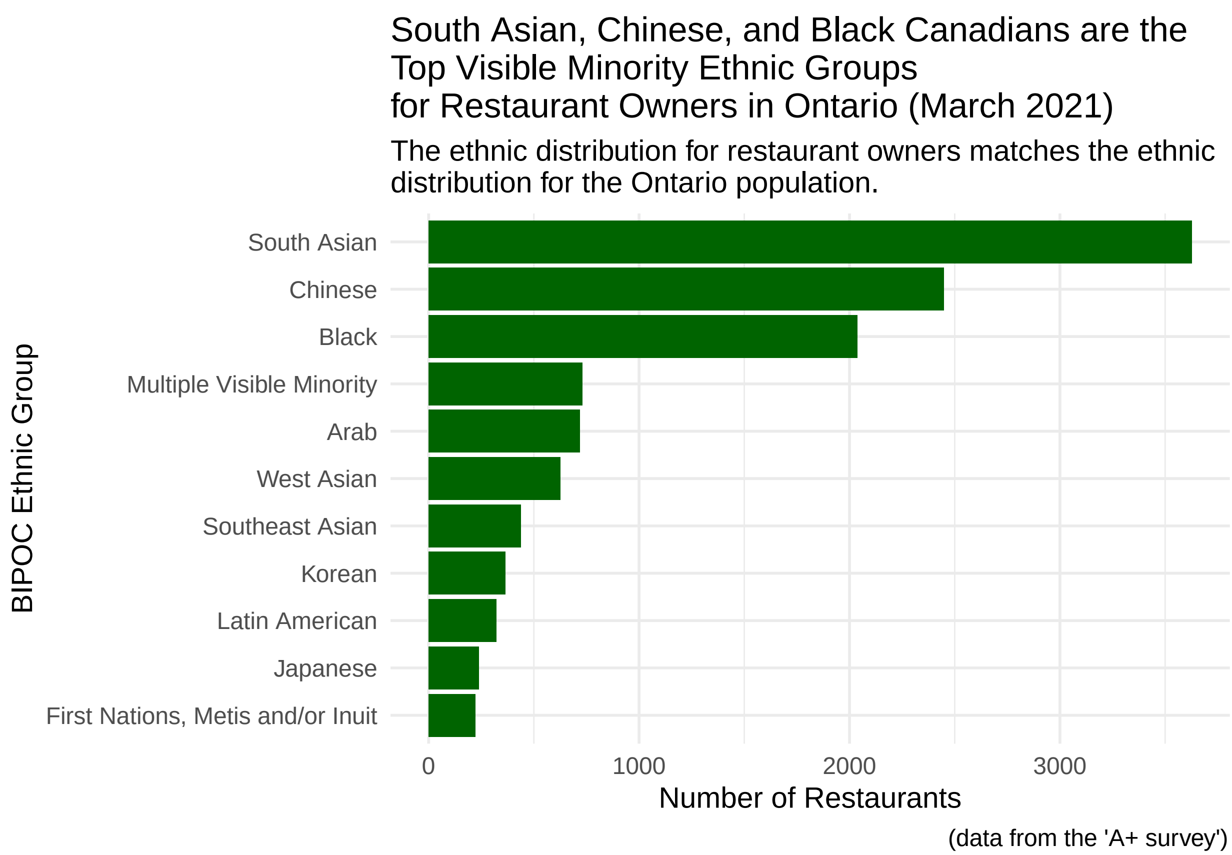


Figure 5: Ethnic Distributions for BIPOC Restaurant Owners in Ontario (March 2021)

2.6.4 Ontario Restaurants Sales, Gross Revenue, Gross Expenses, Net Revenue and Number of Employees have each been Severely Impacted by COVID-19 Lockdowns

Furthermore, the four graphs (Figure 6) below reveal how Ontario restaurants were impacted by the pandemic and lockdowns between March 2020 to March 2021. Despite most restaurants experiencing a the same number of sales through the year, a majority of restaurants experienced a decrease in gross revenues combined with an increase in gross expenses. This led to '66% of restaurants having a decrease in net revenues. Although restaurants were able to maintain their number of sales through methods like take-out and delivery, the graphs demonstrates that the increase in expenses combined with the decrease in revenues caused by customers ordering cheaper meals, led to less restaurants profiting. The graph was built using R (R Core Team 2020), tidyverse (Wickham et al. 2019), ggplot2 (Wickham 2016), gridExtra (Auguie 2017) and RColorBrewer

(Neuwirth 2014).

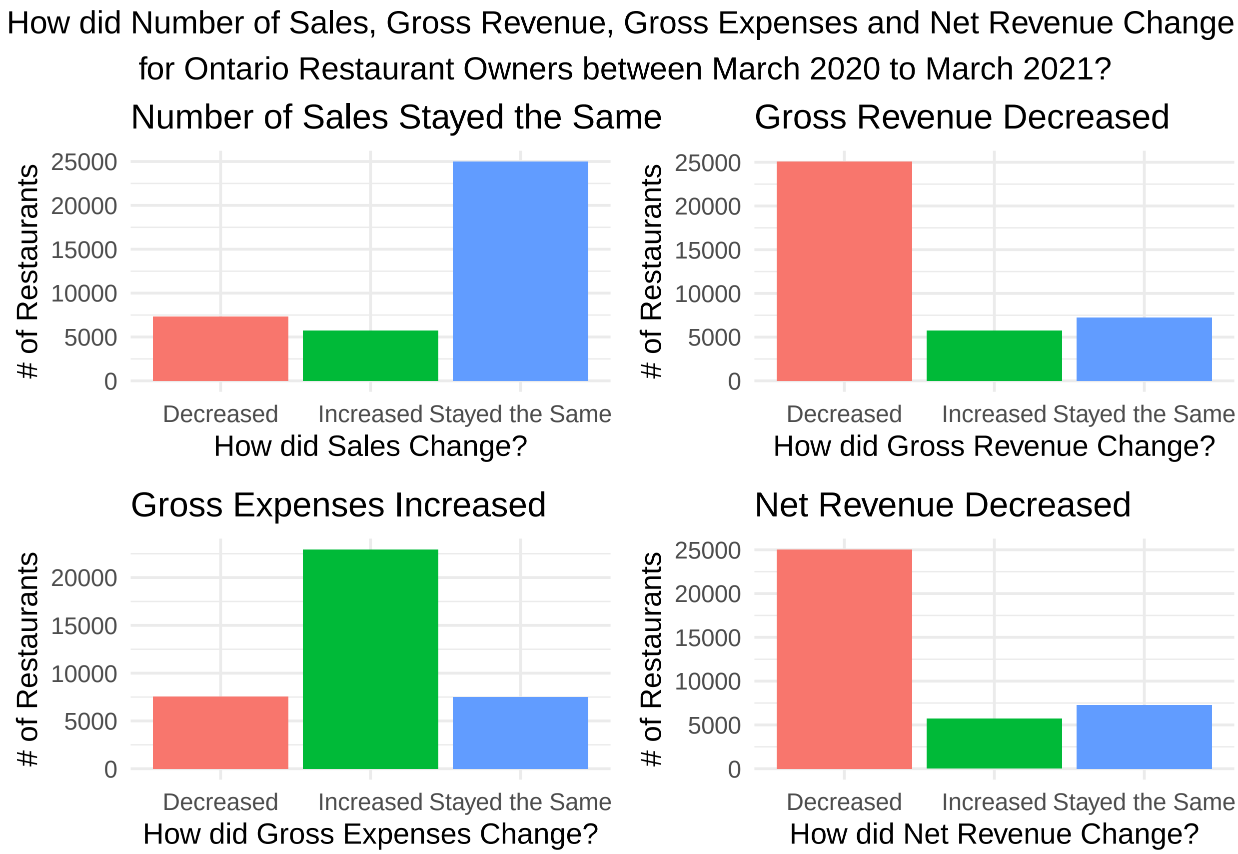


Figure 6: The COVID-19 Pandemic's impact on Number of Sales, Gross Revenue, Gross Expenses, and Net Revenue between March 2020 to March 2021

Moreover, the graph (Figure 7) below demonstrates that a majority of Ontario restaurants maintained the same number of employees during the pandemic. However, these restaurants may have been able to maintain their current staffing levels by using government programs like the 10% wage subsidy, rent subsidy, and other regional programs. The graph was built using R (R Core Team 2020), tidyverse (Wickham et al. 2019), and ggplot2 (Wickham 2016).

2.6.5 Majority of Ontario Restaurants have Access to a Food Delivery/Take-Out Platform

We also analyzed if restaurants had access to food delivery apps for take-out orders (Figure 8). Although over 80% of restaurants are using food delivery platforms like UberEats, Skip the Dishes, and Door Dash, the data above demonstrates that these platforms are not enough to maintain these business's net revenue. The graph was built using R (R Core Team 2020), tidyverse (Wickham et al. 2019), and ggplot2 (Wickham 2016).

2.6.6 Most Ontario Restaurants will not be able to Survive Another Year Given their Current Revenue and Expenses

Lastly, the graph (Figure 9) below reveals that 85% of Ontario restaurants are considering permanent bankruptcy or closure in the next year given their current revenues and expenses. 25% of Ontario restaurants are considering permanent closure or bankruptcy in less than a month. The data demonstrates that if current lockdown restrictions are not eased or the government does not provide additional funding, Ontario's restaurant industry may disappear, making millions of additional Ontarians unemployed and reducing the number of potential employers during Ontario economic recovery to reduce the labour surplus. The graph was built using R (R Core Team 2020), tidyverse (Wickham et al. 2019), and ggplot2 (Wickham 2016).

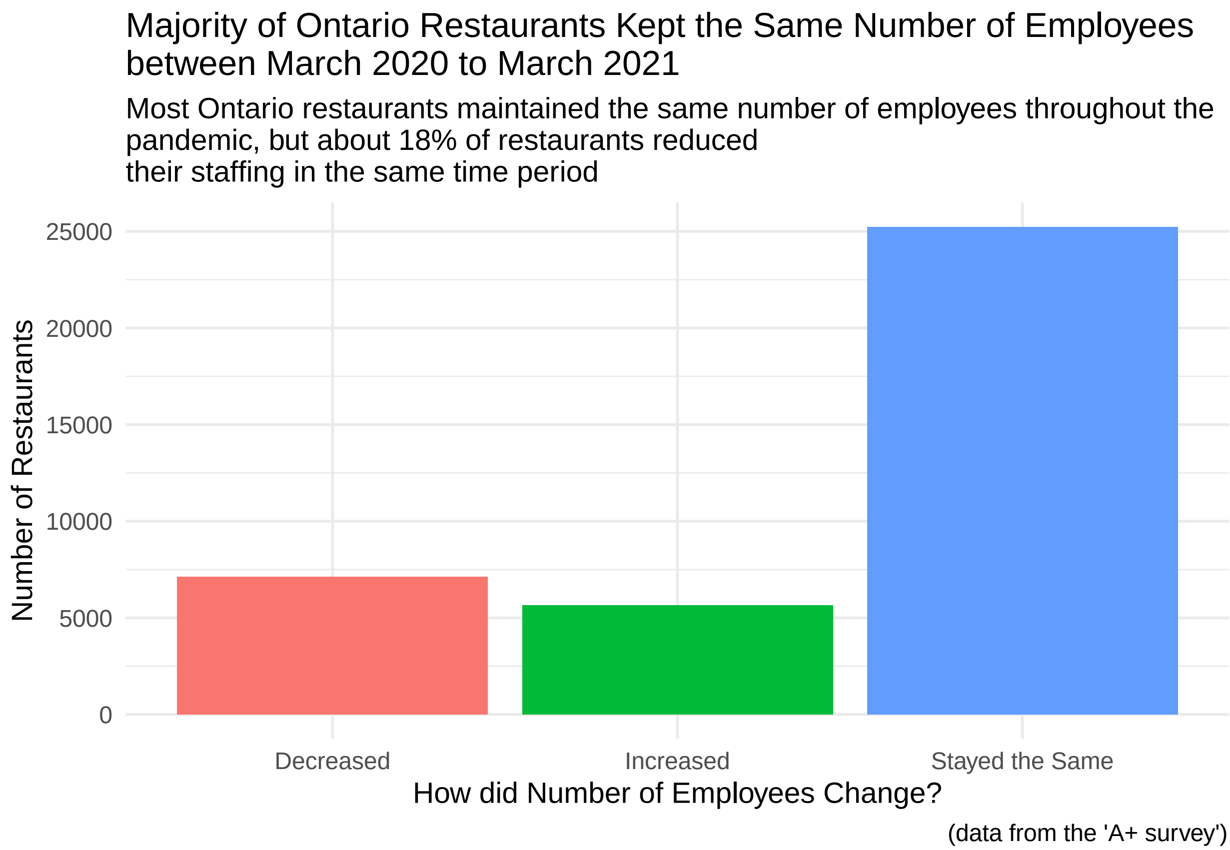


Figure 7: Over the past year (March 2020 to present), how has each of the number of employees changed for this business?

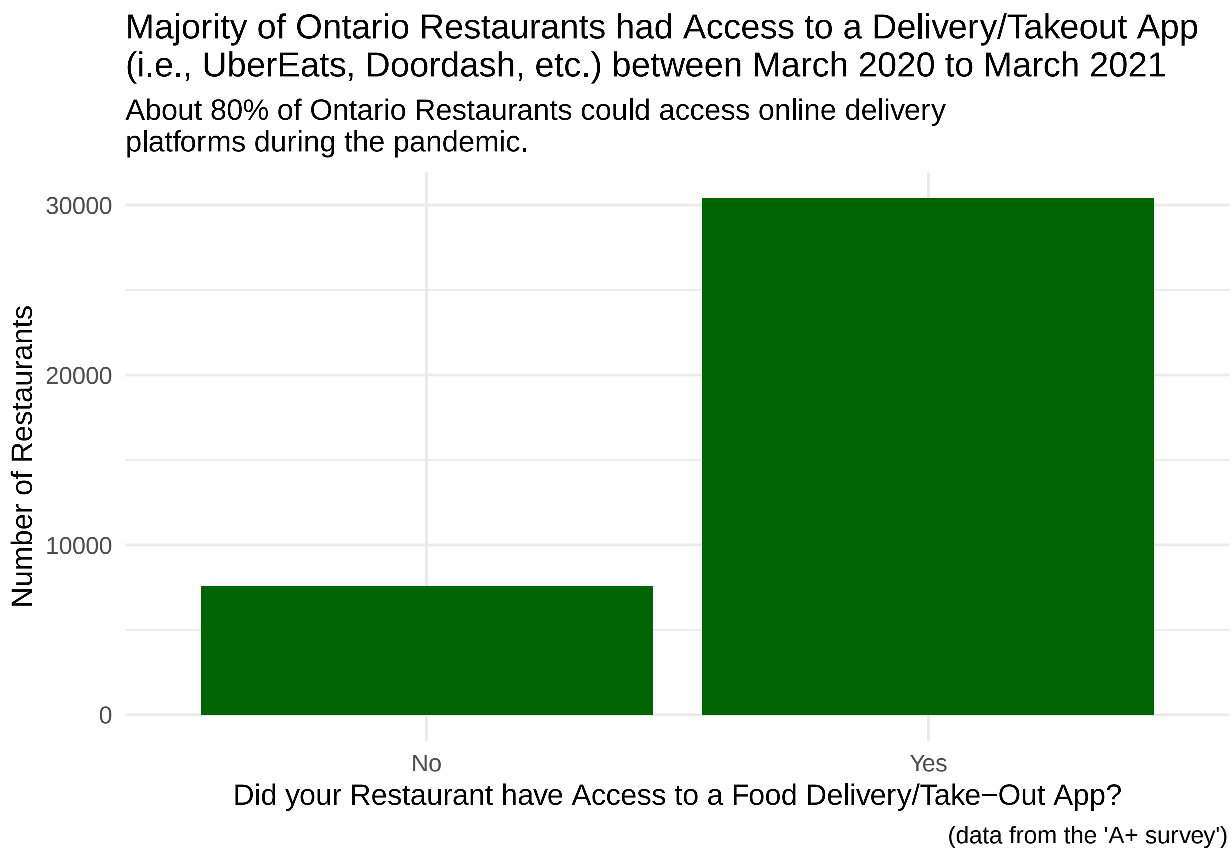


Figure 8: Does your business have access to platforms like UberEats, Doordash, or SkipTheDishes to complete delivery and takeout orders? (March 2021)

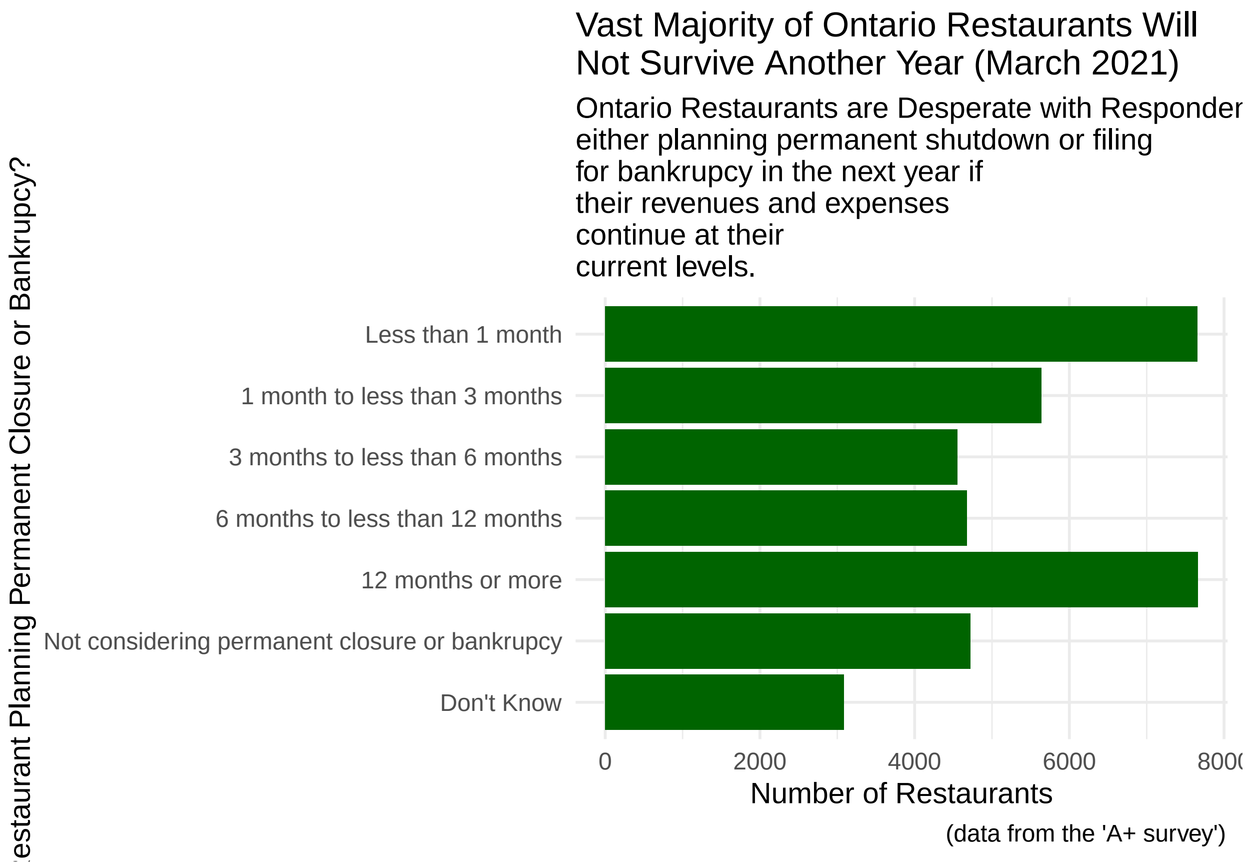


Figure 9: How long can this business continue to operate at its current level of revenue and expenditure before having to consider permanent closure or bankruptcy? (March 2021)?

Table 3: Number of Restaurants who have Access to a Delivery or Takeout Platform per Group

Group	Access to Delivery Service and Takeout Apps	Number of Restaurants
1	No	1995
1	Yes	7635
2	No	1858
2	Yes	7657

2.7 The Treatment and Control Groups are Representative Samples

Our treatment and control groups have high external validity because the Canadian Business Registries list all ~38,000 legal restaurants in Ontario. Thus, our sampling frame encompasses all legal restaurants in the province. Additionally, 18,977 restaurant owners responded to our survey. Thus, we have captured the responses of 50% of our population. Although there may have been factors that may have caused some restaurant owners not to respond to our survey that could impact our external validity, our random sampling method ensures that our evaluation sample accurately reflects the population of Ontario restaurant owners so the impacts of reopening restaurants for in-door dining on the number of contact traced covid-19 cases can be extrapolated from our population.

If our treatment and control groups have the same characteristics in order to control for confounding variables that may lead to differences in the treatment and control group - outside of the treatment. If we have high internal validity, then our treatment is truly independent and we can estimate the 'average treatment effect' of re-opening restaurants for indoor dining on contact-traced covid-19 numbers:

$ATE = E[y|d = 1] - E[y|d = 2]$

That is, the difference between the treated group (d = 1) and the control group (d = 2), when measured by the expected value of our outcome group (number of contact-traced COVID-19 cases). So the mean causal effect is simply the difference between the two expectations.

We will look at the mean of the treatment and control groups from the sampling frame, grouped by restaurants that have access to a delivery/take-out app like UberEats, Skip the Dishes, or Doordash. We can tell from the table (Table 3)below that the two groups have very similar values. The table was created using R (R Core Team 2020), tidyverse (Wickham et al. 2019), and kableExtra (Zhu 2020).

We can use a t-test to test if the treatment and control group do not have confounding variables that will impact our response. The t-test was created using R (R Core Team 2020) and tidyverse (Wickham et al. 2019). Our null hypothesis is that there is no difference between the treatment and control group. Our alternative hypothesis is that there is a difference between the treatment and control group.

We will use a Welch's t-test because it tests if the variances of two samples are equal. The Welch's t-test defines the statistic t by the following formula:

$$t = \frac{\overline{X}_1 - \overline{X}_2}{\sqrt{\frac{s_1^2}{N_1} + \frac{s_2^2}{N_2}}}$$

Figure 10: Welch's T-Test Formula

Where $\bar{X}_j$, s_j, and N_j are the same mean mean, sample standard deviation, and sample size respectively for the two samples. In this case, the treatment and control group (Alexander 2021).

Another reason we are using the t-test is because it gives us the p-value which is a “statistical summary of the compatibility between the observed data and what we would predict or expect to see if we knew the entire statistical model (all the assumptions used to compute the p-value) were correct” (Sander Greenland and Altman 2016).

If we set the alpha level for a statistically significant p-value at 0.01, the p-value (0.04) in the t-test (Table **??**) suggests that the data is not unusual if all the assumptions used to compute the p-value (including the test hypothesis) were correct. A p-value of 0.04 means the discrepancy between the hypothesis prediction would be larger or as large as the observed more than 4% of the time if only chance was creating the discrepancy. However, even if the test hypothesis is wrong, the p-value may be large because it was inflated by a large random error or due to other erroneous assumptions. All in all, the t-test implies that the treatment and control group have high internal validity (Sander Greenland and Altman 2016).

```
## # A tibble: 1 x 10
##   estimate estimate1 estimate2 statistic p.value parameter conf.low conf.high
##      <dbl>     <dbl>     <dbl>     <dbl>   <dbl>     <dbl>    <dbl>     <dbl>
## 1  -0.0119     0.793     0.805     -2.05  0.0401    19141.  -0.0233 -0.000537
## # ... with 2 more variables: method <chr>, alternative <chr>
```

3 Results

3.1 Linear Regression for Contact-Traced COVID-19 Cases

We will use multiple linear regression to determine if there is a relationship between the number of contact-traced COVID-19 cases to a restaurant and opening a restaurant for in-door dining (Table **??**). We are using a multiple linear regression because the response variable is quanitative (number of contact-traced COVID-19 cases. Also, we commonly have more than one predictor variable in real life. Thus, multiple linear regression gives each predictor a separate slope coefficient in a single model to determine of the relationship between the response and predictor variables, decide on important variables, create a better model fit, and make predictions. The multiple linear regression was created using R (R Core Team 2020). The formula for multiple linear regression is:

$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + ... + \beta_p X_p$

Where β_0 is the y-intercept, X_i represents the ith predictor and β_i quantifies the association between the variable and the response (Gareth James and Tibshirani 2017).

Our null and alternative hypothesis are:

H0: There is no relationship between the number of contact-traced COVID-19 cases and if a restaurant has been opened for indoor dining.

H1: There is a relationship between the number of contact-traced COVID-19 cases and if a restaurant has been opened for indoor dining.

We also have a racist uncle, Eddie who is very xenophobic. Eddie argues that we should “re-open” the economy by opening all restaurants for in-door dining and closing all restaurants owned by a non-white visible minority. He has been telling people on chatrooms not to order from any restaurants owned by a BIPOC. Thus, we will also test the relationship between the number of contact-traced COVID-19 cases and if more than 50% of the restaurant is owned by a Black, Indigenous or Person of Colour, including their ethnicity.

We will set our alpha level to 0.05.

```
##
## Call:
## lm(formula = contact_traced_covid_cases ~ group + Q12 + Q13,
##     data = simulated_dataset)
##
## Residuals:
```

```
##     Min      1Q  Median      3Q     Max
## -6.237 -2.036 -0.077  1.825  9.934
##
## Coefficients: (1 not defined because of singularities)
##                                         Estimate Std. Error t value Pr(>|t|)
## (Intercept)                            6.0660195  0.0274187 221.236   <2e-16 ***
## group2                                -2.9890107  0.0351638 -85.003   <2e-16 ***
## Q12Yes                                 0.0981947  0.1309916   0.750   0.4535
## Q13Black                              -0.1390501  0.1501508  -0.926   0.3544
## Q13Chinese                            -0.0175664  0.1462993  -0.120   0.9044
## Q13First Nations, Metis and/or Inuit -0.2485359  0.2743771  -0.906   0.3650
## Q13Japanese                           -0.2563663  0.2515794  -1.019   0.3082
## Q13Korean                             -0.0004967  0.2239101  -0.002   0.9982
## Q13Latin American                     -0.4096575  0.2283392  -1.794   0.0728 .
## Q13Multiple Visible Minority           0.0510359  0.1801245   0.283   0.7769
## Q13NA                                         NA         NA      NA       NA
## Q13South Asian                         0.0731000  0.1412876   0.517   0.6049
## Q13Southeast Asian                    -0.3179582  0.2103032  -1.512   0.1306
## Q13West Asian                          0.0042603  0.1890312   0.023   0.9820
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 2.432 on 19132 degrees of freedom
## Multiple R-squared:  0.2746, Adjusted R-squared:  0.2741
## F-statistic: 603.5 on 12 and 19132 DF,  p-value: < 2.2e-16
```

The intercept is the number (6.08) of contact_traced covid-19 cases we would expect if a restaurant opened for in-door dining. If we closed a restaurant to in-door dining, we would expect the number of contact-traced COVID-19 cases to decrease by -2.98 cases. Our p-value for group 2 (<2e-16)is well below our alpha-level threshold of 0.05. Therefore, the p-value flags the data as being unusual if all the assumptions used to compute it (including the test hypothesis) were correct. A p-value less than 0.05 suggests that a discrepancy from the hypothesis prediction would be as large or larger than the observed no more than less than 5% of the time if only chance was creating the discrepancy. However, the low p-value may also be caused by a large random error or some assumptions other than the test hypothesis were violated. Thus, we can reject the null hypothesis that there is no difference between the number of contact traced COVID-19 cases and if a restaurant is open for in-door dining (Sander Greenland and Altman 2016).

The p-values for Q12Yes and for each ethnicity are all above the alpha threshold. The large p-values suggest that the data is not unusual if all the assumptions used to compute the p-value (including the test hypothesis) were correct. Thus, we can conclude that Uncle Eddie is wrong because we do not have enough evidence to reject the null hypothesis that there is no difference between the number of contact_traced COVID-19 cases and if the restaurant is owned by a non-white visible minority.

The R-Squared value illustrates that our model explains ~27% of the response data around its mean.

Racist Uncle Eddie did not like our results and asked us to run another test.

3.2 Logistic Regression to Predict if a Restaurant is Open for Indoor Dining

Similarly, we can also use a logistic regression to determine if there is a relationship between a restaurant being open for indoor-dining and the number of contact-traced COVID-19 cases (Tab **??**). We will also test racist Uncle Eddie's hypothesis again too. The logistic regression was created using R (R Core Team 2020). We will use logistic regression because our response variable is binary (open for indoor dining: yes or no). Additionally, there is no natural way to convert a qualitative response variable with two levels into a quantitative response that is ready for linear regression. Consequently, logistic regression is preferable because it is suited for qualitative response variables (Gareth James and Tibshirani 2017).

The formula for logistic regression is:

$$p(X) = \frac{e^{\beta_0+\beta_1 X}}{1+e^{\beta_0+\beta_1 X}}.$$

Figure 11: Logistic Regression Formula

Where P is the probability of X (the mean of the Dependent Variable), e is the base of the natural logarithm (about 2.718), and β_0 and β_1 are the parameters of the model (y-intercept and slope respectivly) (Gareth James and Tibshirani 2017).

We will set our alpha level to 0.05.

```
##
## Call:
## glm(formula = group ~ contact_traced_covid_cases + Q12 + Q13,
##     family = "binomial", data = simulated_dataset)
##
## Deviance Residuals:
##     Min       1Q   Median       3Q      Max
## -2.2214  -0.8512  -0.1469   0.8929   2.2329
##
## Coefficients: (1 not defined because of singularities)
##                                       Estimate Std. Error z value Pr(>|z|)
## (Intercept)                           2.256300   0.040300  55.988   <2e-16 ***
## contact_traced_covid_cases           -0.514166   0.008160 -63.013   <2e-16 ***
## Q12Yes                               -0.007644   0.127758  -0.060    0.952
## Q13Black                             -0.027842   0.145936  -0.191    0.849
## Q13Chinese                            0.110485   0.142846   0.773    0.439
## Q13First Nations, Metis and/or Inuit -0.174791   0.267052  -0.655    0.513
## Q13Japanese                          -0.044075   0.241274  -0.183    0.855
## Q13Korean                             0.127813   0.222852   0.574    0.566
## Q13Latin American                    -0.013868   0.221557  -0.063    0.950
## Q13Multiple Visible Minority          0.072031   0.177514   0.406    0.685
## Q13NA                                       NA         NA      NA       NA
## Q13South Asian                        0.130079   0.137787   0.944    0.345
## Q13Southeast Asian                   -0.259220   0.202346  -1.281    0.200
## Q13West Asian                        -0.088036   0.185783  -0.474    0.636
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##     Null deviance: 26540  on 19144  degrees of freedom
## Residual deviance: 20364  on 19132  degrees of freedom
## AIC: 20390
##
## Number of Fisher Scoring iterations: 4
```

The logistic regression results demonstrate that for everyone one unit increase in the number of contact-traced COVID-19 cases, the log odds that the restaurant was not open for indoor dining decrease by 0.50. Our p-value for contract-traced COVID-19 cases (<2e-16)is well below our alpha-level threshold of 0.05. Therefore, the p-value flags the data as being unusual if all the assumptions used to compute it (including the test hypothesis) were correct. Thus, we can reject the null hypothesis that there is no difference between a restaurant being open for indoor dining and the number of contact-traced COVID-19 cases (Sander Greenland and Altman 2016).

The p-values for Q12Yes and for each ethnicity are all above the alpha threshold. The large p-values suggest that the data is not unusual if all the assumptions used to compute the p-value (including the test hypothesis) were correct. Thus, we can conclude that Uncle Eddie is wrong because we do not have enough evidence to reject the null hypothesis that there is no difference between the number of contact_traced COVID-19 cases and if the restaurant is owned by a non-white visible minority. Statistics have shows racist Uncle Eddie again that his test hypothesis is very likely wrong!

3.3 Visualizing the Relationship between Opening Restaurants for Indoor Dining and Contact-Traced COVID-19 Cases

The histogram below (Figure 12) illustrates that restaurants who have opened for in-door dining (treatment group = 1) have a larger mean in contact-traced COVID-19 cases than restaurants that have not opened for in-door dining (control group = 2). The treatment group distribution is also more spread out indicating that is has a larger standard deviation. Thus, the histogram explains why both the linear regression and logistic regression returned a low p-value for the relationship between opening restaurants for indoor dining and the number of contact-traced COVID-19 cases. The histogram was created using R (R Core Team 2020), tidyverse (Wickham et al. 2019), ggplot2 (Wickham 2016), and RColorBewer (Neuwirth 2014).

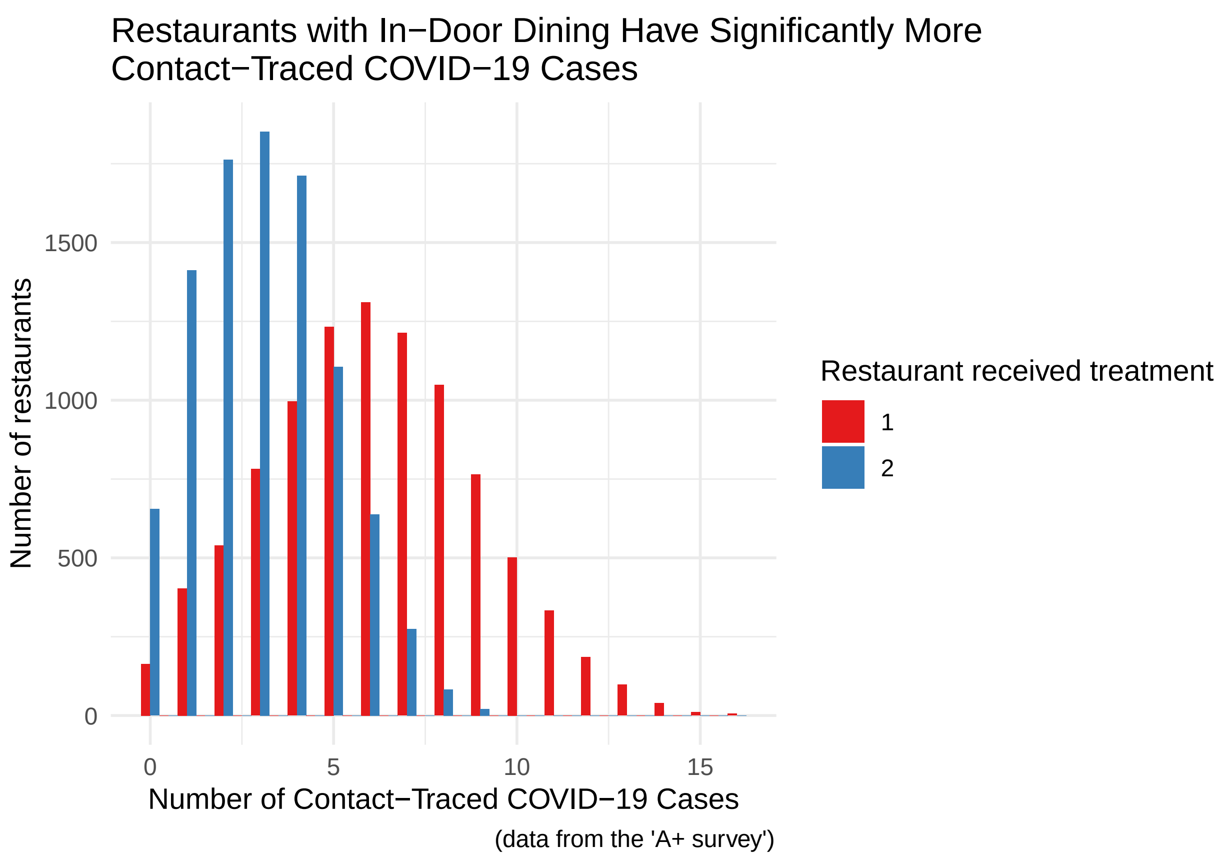


Figure 12: Relationship between Opening Treatment Restaurants for Indoor Dining and the Number of COVID-19 Cases Traced Back to that Restaurant (March 2021)

The table (Table 4) below further demonstrates the large difference in contact-traced COVID-19 cases

Table 4: Total and Average Number of COVID-19 Cases Traced back to the Treatment Group (1) versus the Control Group (2)

group	Number of COVID-19 Cases	Number of Restaurants	Average Number of COVID-19 Cases per Restaurant
1	58602	9630	6.085358
2	29469	9515	3.097110

between our treatment and control group. The average number of cases for restaurants that opened for indoor dining is double the number restaurants that did not receive the intervention. Therefore, the disease is less transmissible when restaurants remain shuttered to indoor dining.

The table was created using R (R Core Team 2020), tidyverse (Wickham et al. 2019), and kableExtra (Zhu 2020).

4 Discussion and Future Research

Our intervention on Ontario restaurants examined allowing them to re-open for in-person dining over a duration of two-weeks. Asking the question that would lead to increased occurrences of Covid-19 community spread that could be traced back to the restaurants that participated in our intervention. Using simulated data to explore this scenario, the results forecasted that community spread would occur from restaurants if the province allowed them to open in-person dining. Our recommendation to not have in-person dining open yet is unfortunate for the restaurant industry as they are struggling with financial stability on account of their increased debt and inability to generate sufficient revenue as they are among the most impacted industries during the pandemic (Morden 2021)). Restaurant Canada, a non-profit organization working with restaurants across the country, is petitioning all levels of government for help. Referencing a survey of 1,000 Canadian restaurants they conducted showing the level of desperation that the industry faces as one out of every two independent restaurants could close permanently depending on what happens over the course of the next few months (Mace 2020). Restaurants are going to be struggling for revenue if not already as they adapt to post pandemic life. As hope for a quick re-opening diminishes, the burden of responsibility must shift to the government to provide adequate assistance until life resumes to what could be understood as normal.

In the broader research context, our intervention corroborates what has been observed by other researchers. Amesh Adalja of John Hopkins Center for Health Security has stated that the current trends in how Covid-19 is spreading is consistent with what health officials predicated from the beginning, that certain activities leading to more spread than others (Cooney 2020) Restaurants, especially those that are full-service, fall into this category as one of the largest predicted impacts of infection, "due to the large number of restaurantsas well as their visit densities and long dwell times" (Cooney 2020). The nature of how Covid-19 is spreading means that opening full-service restaurants would be considered high-risk risk for increased spread of Covid-19 and that for it to be considered safe that there would have to be a reduction in the maximum occupancy with other health precautions put in place (Chang and Pierson 2021). As we are not even a year into the Pandemic, further research would be useful trying to calculate a safe occupancy that allows restaurants to survive economically. A recommendation would be an A/B test for the maximum safe allowable patrons in a restaurant with rigorous safety precautions put in place to limit the chance of spread.

While research is being conducted to determine safe occupancy in restaurants, not enough is being done financially to help them weather the pandemic. In Portugal, the Portuguese Restaurants and Hotel Owners Association has stated that 70% of the industry will not be able to pay salaries without timely government support (Madeira 2020). The National Association of Restaurants in Brazil conducted a survey where over 50% of restaurants are experiencing revenue drops between 50% to 90% and that they are in financial trouble (Freitas and Stedefeldt 2020). Globally with data used from OpenTable, an online restaurant reservation company, shows that after the World Health Organization declared the Covid-19 pandemic on March 11,2020 that business dropped between 10% and 20% immediately for them and then by March 18th, 2020 that business had declined by 90% and has been very slow to recover (Kaitano Dube and Chikodzi 2020). Where in-person dining is not allowed, to generate revenue, restaurants have tried expanding to delivery options.

Although this model may work for some restaurants, others cannot cover their costs or market conditions may be against them as the popularity of fine-dining $35-dollar pasta plates may not be attractive to consumers. For those that are allowed to open, often at reduced capacity, they must take on addition expenses with contactless payment, menu boards and sanitization while offering discounts as an incentive to attend (Kaitano Dube and Chikodzi 2020). As restaurants do whatever they can to survive, governments must consider their role in ensuring their survival and value both directly and indirectly to the economy as they may determine their fate through funding and grants.

In this paper, we explored the relationship between opening a restaurant for indoor dining and the number of COVID-19 cases traced back to that business within the 18,977 restaurant owners that participated in our Randomized Controlled Experiment and responded to our survey in March 2021. First, we discovered that restaurants that open for indoor dining are statistically more likely to have more contact-traced COVID-19 cases compared to restaurants that remained closed. Although these results could be explained by the population density of where our restaurants were located, our simple random sampling method controlled for this confounding variable. Instead, the increase in COVID-19 cases is likely due to customers having to take off their masks when they eat their food which releases the virus into the air. Also, the customers are in a closed environment and the restaurant may have poor ventilation which both contribute to disease transmission. Secondly, the results illustrated that 50% or more of a restaurant being owned by a BIPOC visible minority does not contribute to an increase in COVID-19 cases being traced back to a restaurant. This result was likely due to Uncle Eddie's test hypothesis being a racist assumption and there is no evidence from external literature that ethnicity causes COVID-19 transmission. Lastly, we uncovered that the vast majority Ontario restaurant owners are considering permanent closure or bankruptcy in the next year if their current revenues and expenses continue. Consequently, the Ontario government must find alternative solutions other than opening restaurants for indoor dining to financially support these businesses. The results show that the economic gain of opening restaurants is ethnically not worth of the public health risk.

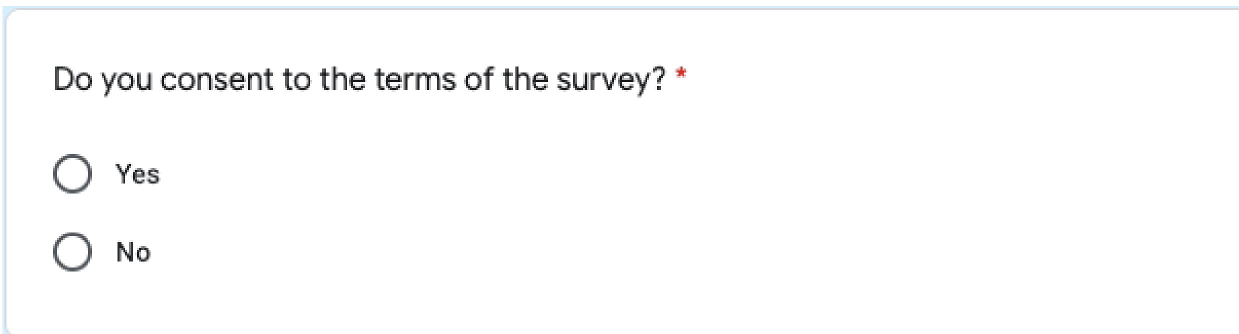
There are a few limitations to our survey and proceeding experiment worth discussing, all of which circulate around a long standing tension between sample anonymity and sample accuracy. This tension is probably best laid out in work done but Lisa Austin and David Lie (2019), which speaks more directly to the balancing act that comes with making sure your data is both accurate and private. Speaking in terms of anonymization and de-anonymization, Austin and Lie engage with methods such as k-anonymity and differential privacy, measuring their respective strengths and limitations. Very briefly, k-anonymity is a process where specific data features are omitted from a sample (or for Austin and Lie, a dataset) to ensure no re-identification of respondents is possible (Austin and Lie 2016, 592). Differential privacy is slightly more sophisticated as it requires the development of a model that measures significant "privacy loss" (Austin and Lie 2016, 583) in a sample and accounts for this metric in the proceeding statistical analysis.

In a broad sense, we approached our sampling strategy and survey design from a perspective of 'proto k-anonymity,' where the survey itself was developed with protection against data re-identification as a top priority. As a result, the survey data collected not only omits certain features, but excludes them entirely. This mitigates virtually any possibility of this at-risk data falling into the wrong hands and avoids risks of re-identification from the outset. While keeping at-risk data out of a sample is the safest way to ensure re-identification, it also has the limitation of reducing data quality compared to the k-anonymity and differential privacy methods. With this being said, however, our team firmly stands by the sampling strategy and privacy-first survey design employed, as the liabilities associated with data re-identification are far more costly than producing sample data of a marginally lower quality.

Looking forward, new research into this particular intervention could be conducted with a revised sampling strategy that incorporates k-anonymity, differential privacy, or both, into its into it's initial design. Keep in mind, developing such a strategy will require more time and resources to account for both modelling optimal privacy metrics as well as liability concerns that come with more sensitive information being collected. Given the reproducible nature of our work, it is our hope that other researchers also attempt to utilize —and perhaps expand upon— our initial sample design with an integrated differential privacy score or selected list of k-anonymization features. Given the time sensitive nature of our own intervention analysis, however, we must leave this further work to like-minded researchers.

When considering the implications of our intervention, however, there are a number of directions for further research that could all contribute unique insights into the ongoing COVID-19 pandemic and its relationship with local restaurants. For example, more analysis could be conducted on the treatment group, and specifically which sample features could contribute to virus transmission outside of the ones observed. In addition, further research and interventions geared towards economic relief such as provincial rent subsidies, wage subsidies, and tax benefits could be measured on the same sample group with the enhanced benefit the privacy methods discussed above. As we continue to navigate the uncertainty of this pandemic, it is imperative that these forms of experimentation are further pursued.

5 Appendix: Survey Details

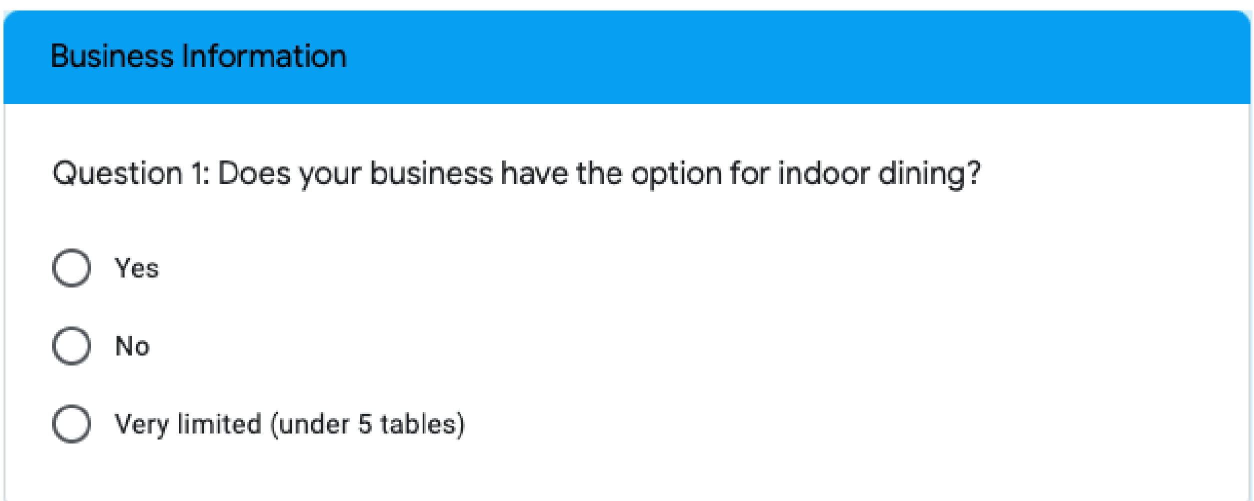


Do you consent to the terms of the survey? *

Yes

No

Figure 13: Consent



Business Information

Question 1: Does your business have the option for indoor dining?

Yes

No

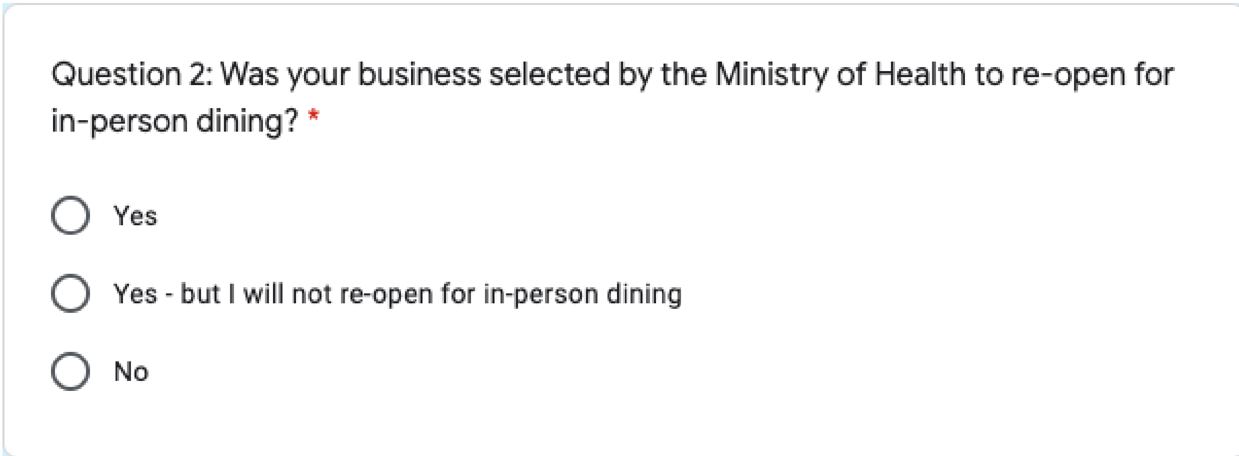
Very limited (under 5 tables)

Figure 14: Question 1

Bibliography

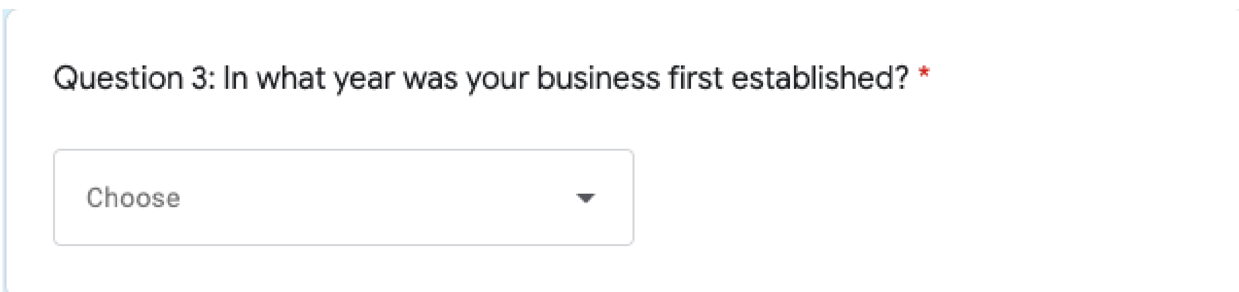
Alexander, Rohan. 2021. *Telling Stories with Data.* https://www.tellingstorieswithdata.com.

Auguie, Baptiste. 2017. *gridExtra: Miscellaneous Functions for "Grid" Graphics.* https://CRAN.R-project.org/package=gridExtra.


Question 2: Was your business selected by the Ministry of Health to re-open for in-person dining? *

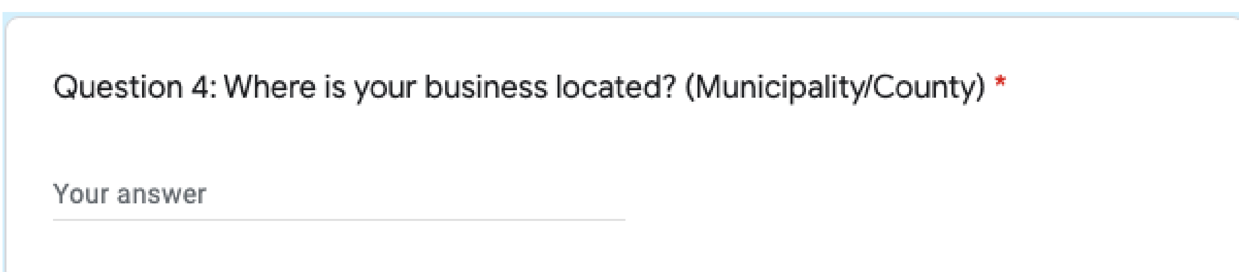
- Yes
- Yes - but I will not re-open for in-person dining
- No

Figure 15: Question 2


Question 3: In what year was your business first established? *

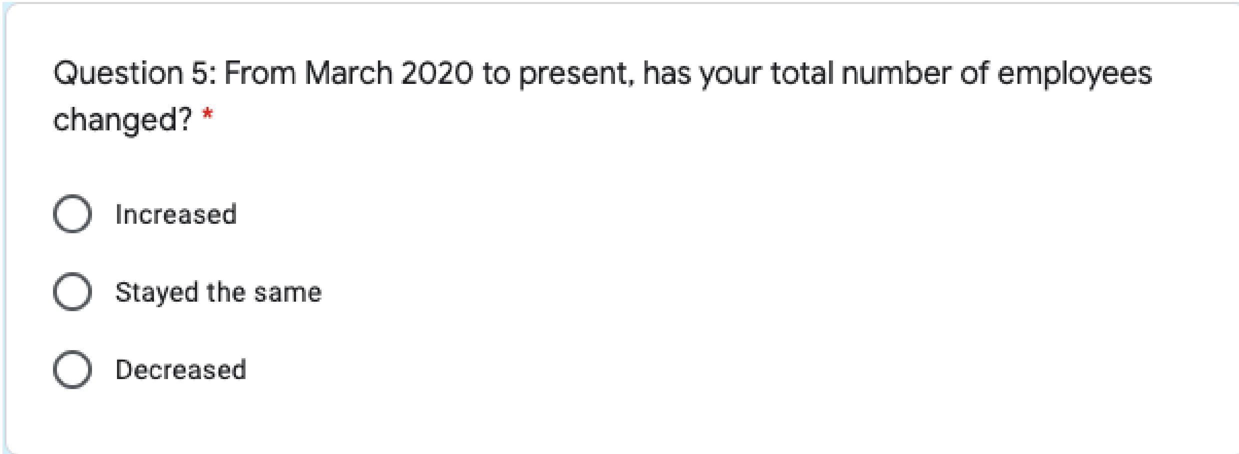
Choose

Figure 16: Question 3


Question 4: Where is your business located? (Municipality/County) *

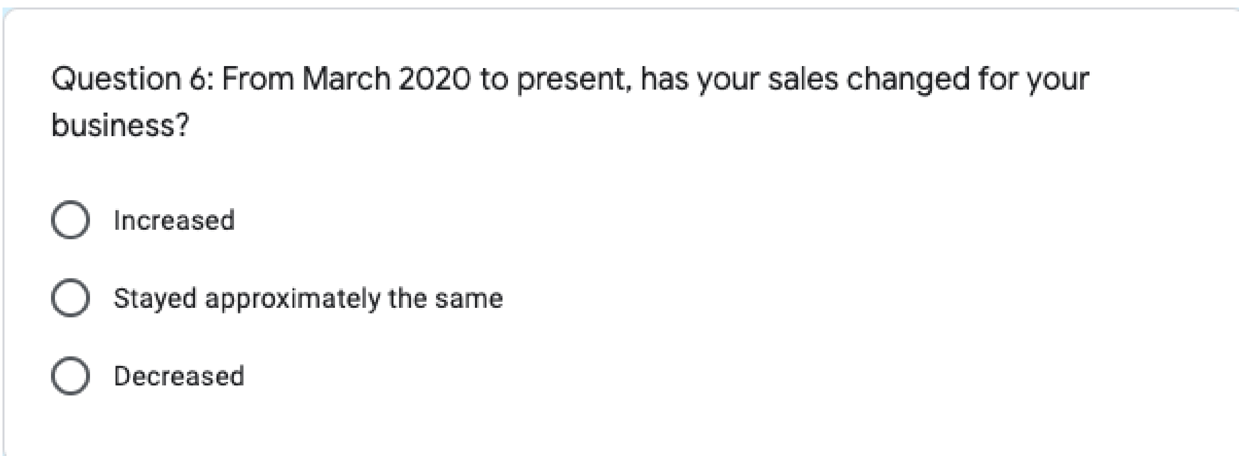
Your answer

Figure 17: Question 4


Question 5: From March 2020 to present, has your total number of employees changed? *

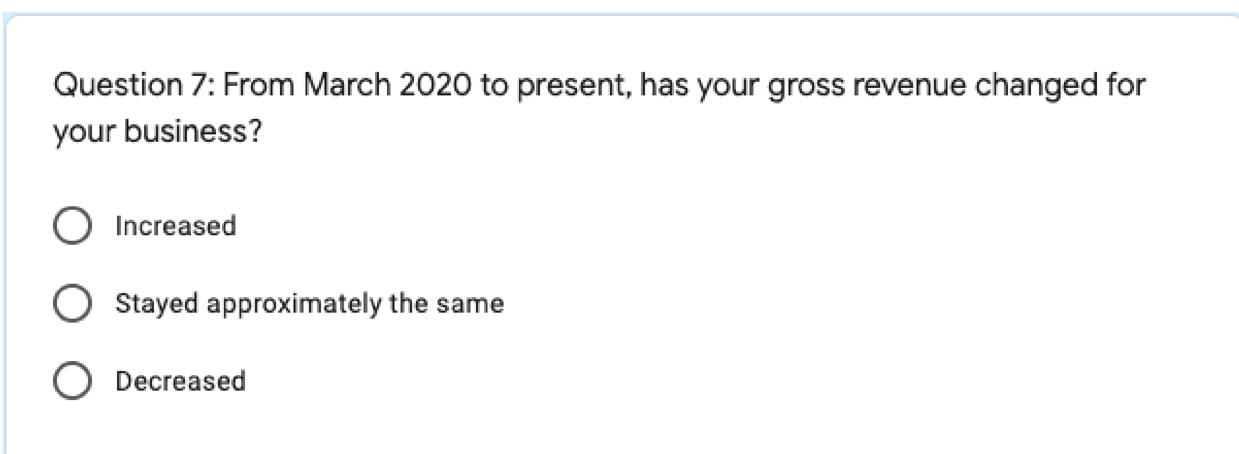
- Increased
- Stayed the same
- Decreased

Figure 18: Question 5


Question 6: From March 2020 to present, has your sales changed for your business?

- Increased
- Stayed approximately the same
- Decreased

Figure 19: Question 6


Question 7: From March 2020 to present, has your gross revenue changed for your business?

- Increased
- Stayed approximately the same
- Decreased

Figure 20: Question 7

Question 8: From March 2020 to present, has your gross expenses changed for your business?

- Increased
- Stayed approximately the same
- Decreased

Figure 21: Question 8

Question 9: From March 2020 to present, has your net revenue changed for your business?

- Increased
- Stayed approximately the same
- Decreased

Figure 22: Question 9

Question 10: Which of the following have been obstacles to your business since March 2020? (Select all that apply) *

- [ ] Shortage of labour force
- [ ] Recruiting and training skilled employees
- [ ] Supply chain challenges
- [ ] Insufficient and/or fluctuating customer demand
- [ ] Cost of insurance
- [ ] Increased competition
- [ ] Maintaining cash flow
- [ ] Debt
- [ ] Rent expenses
- [ ] Managing online presence
- [ ] Cost of delivery and take-out service options
- [ ] None of the above
- [ ] Other: ______

Figure 23: Question 10

Government Funding or Credit

Question 11: Due to Covid-19, was funding or credit for your business approved or received through any of the following sources? (Select all that apply) *

- [ ] Canada Emergency Business Account (CEBA)
- [ ] Temporary 10% wage subsidy
- [ ] Canada Emergency Wage Subsidy (CEWS)
- [ ] Canada Emergency Rent Subsidy (CERS)
- [ ] Canada Emergency Commercial Rent Assistance (CECRA)
- [ ] Region relief and recovery fund
- [ ] Provincial, territorial, or municipal government programs
- [ ] Grant or loan funding from donors or mutual-aid sources
- [ ] Financial institution (e.g., term loan or line-of-credit)
- [ ] Loan from family or friends
- [ ] None of the above
- [ ] Other: ____________________

Figure 24: Question 11

Question 12: If you answered "None of the above" for Question 11, which of the following reasons has the business not accessed any funding or credit due to Covid-19?

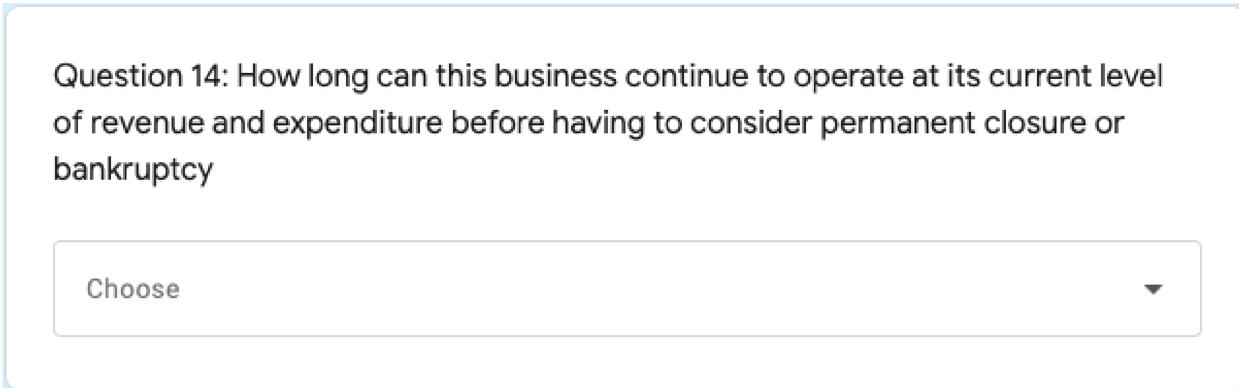
- ○ Funding or credit not needed
- ○ Waiting for approval or submitting application
- ○ Did not meet eligibility requirements
- ○ Unable to fill out application
- ○ Lack of awareness
- ○ Other: ______

Figure 25: Question 12

Question 13: Does your business have the ability to take on more debt?

- ○ Yes
- ○ No
- ○ Don't know

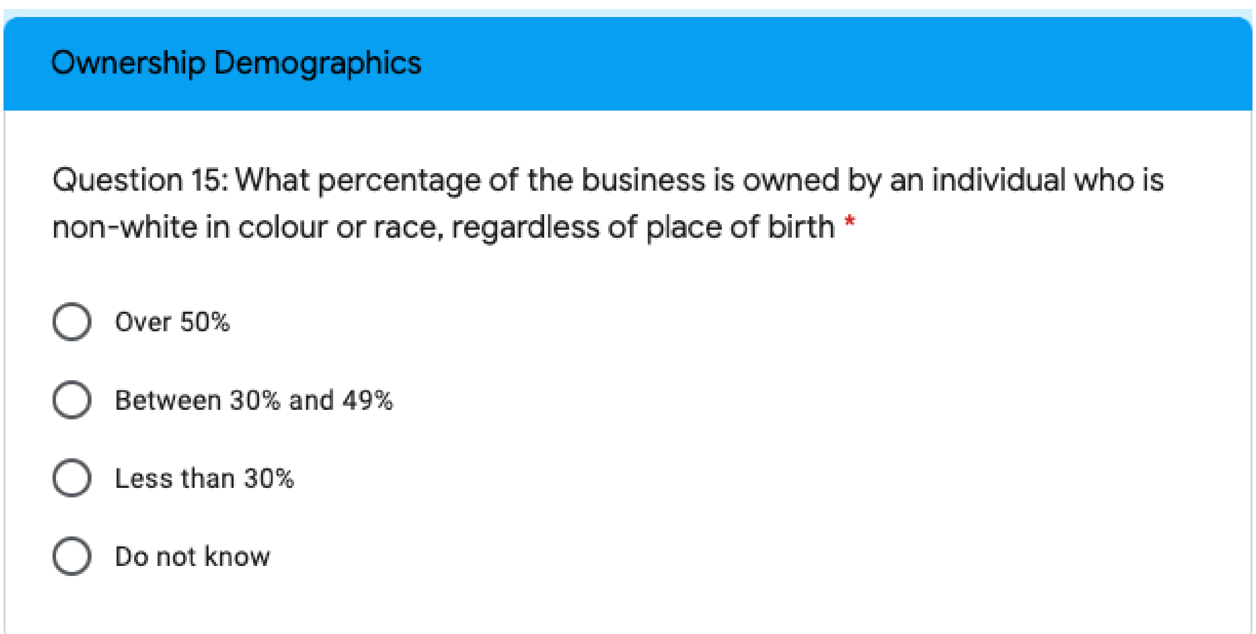
Figure 26: Question 13



Question 14: How long can this business continue to operate at its current level of revenue and expenditure before having to consider permanent closure or bankruptcy

Choose

Figure 27: Question 14



Ownership Demographics

Question 15: What percentage of the business is owned by an individual who is non-white in colour or race, regardless of place of birth *

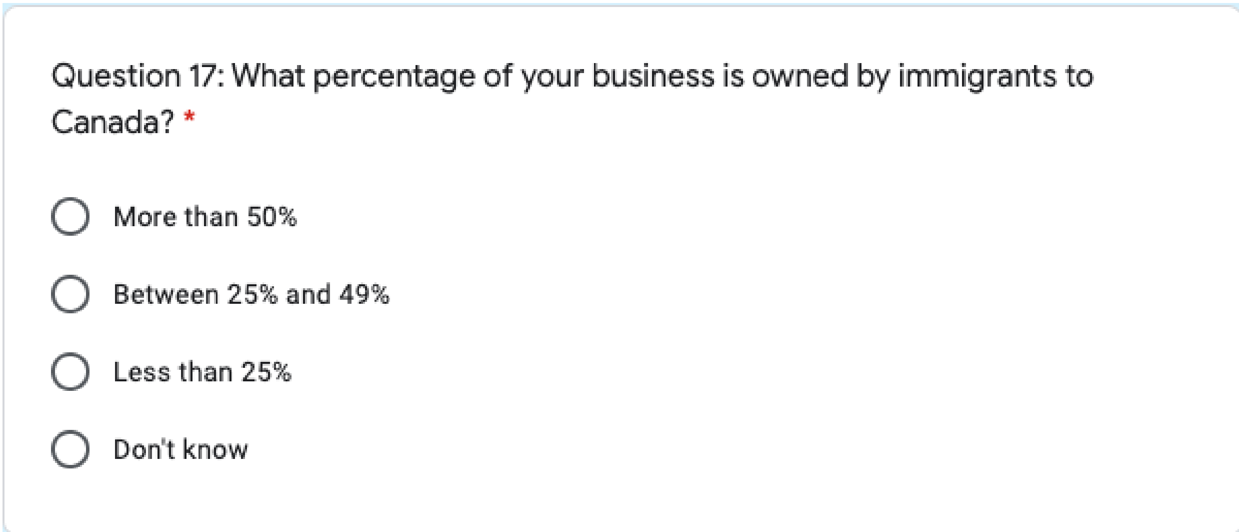
- Over 50%
- Between 30% and 49%
- Less than 30%
- Do not know

Figure 28: Question 15

Question 16: What ethnic groups are the business owners apart of? (Select all that apply) *

- [ ] Black, African, and/or Afro-Canadian
- [ ] Chinese
- [ ] Filipino
- [ ] First Nations, Metis, and/ore Inuit
- [ ] Japanese
- [ ] Korean
- [ ] Latin American
- [ ] South Asian (i.e., East Indian, Pakistani, Bangladeshi)
- [ ] Southeast Asian (i.e., Vietanamese, Thai, Cambodian)
- [ ] West Asian (i.e, Iranian, Afgan, Turkish)
- [ ] Prefer not to say
- [ ] Not applicable
- [ ] Other: ____________

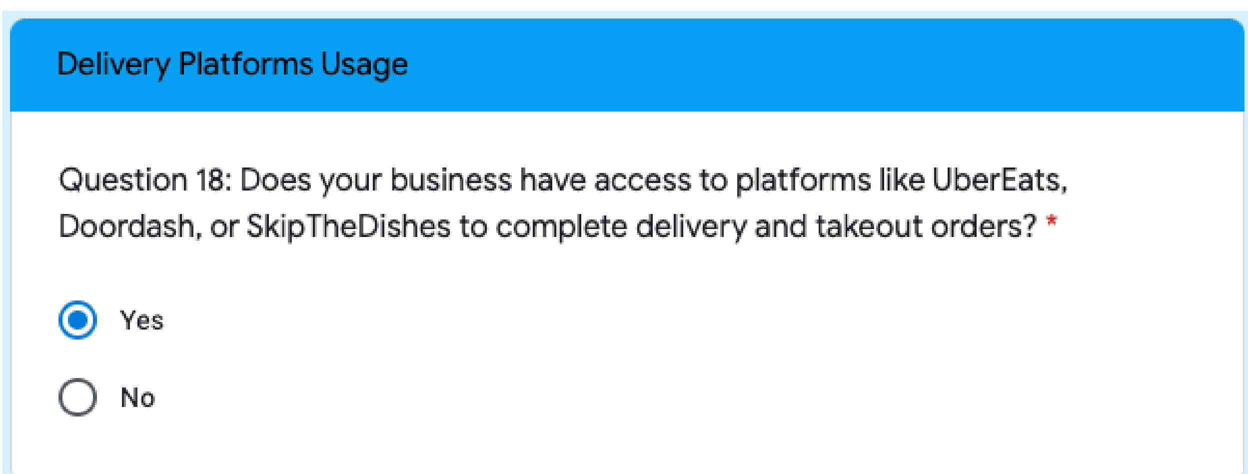
Figure 29: Question 16



Question 17: What percentage of your business is owned by immigrants to Canada? *

- More than 50%
- Between 25% and 49%
- Less than 25%
- Don't know

Figure 30: Question 17

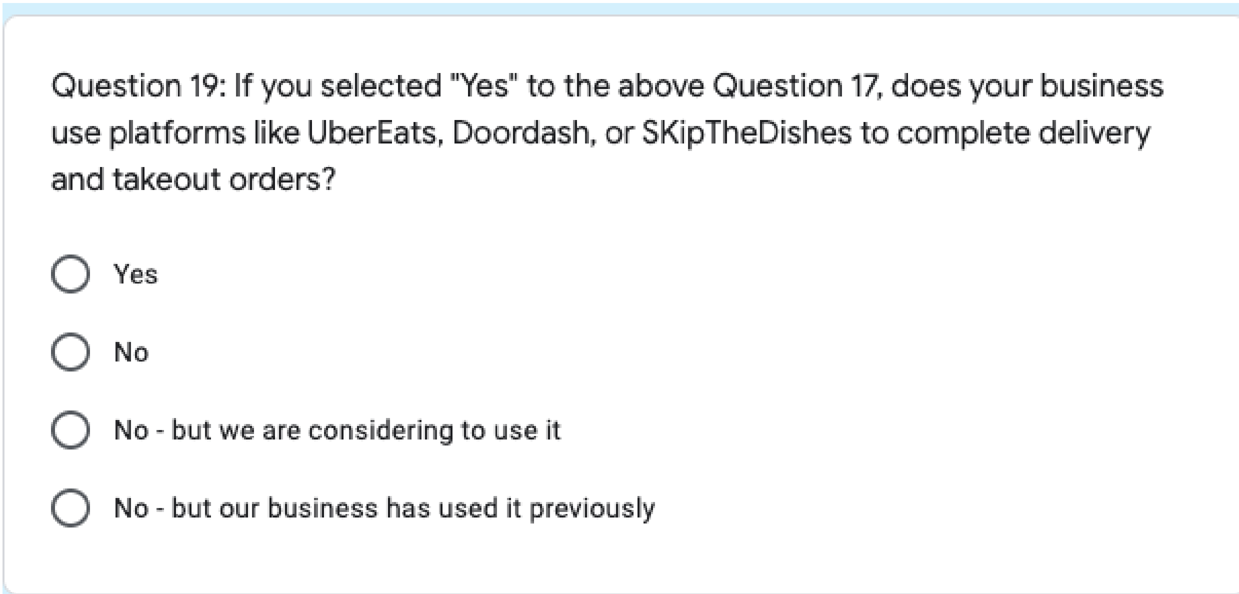


Delivery Platforms Usage

Question 18: Does your business have access to platforms like UberEats, Doordash, or SkipTheDishes to complete delivery and takeout orders? *

- Yes
- No

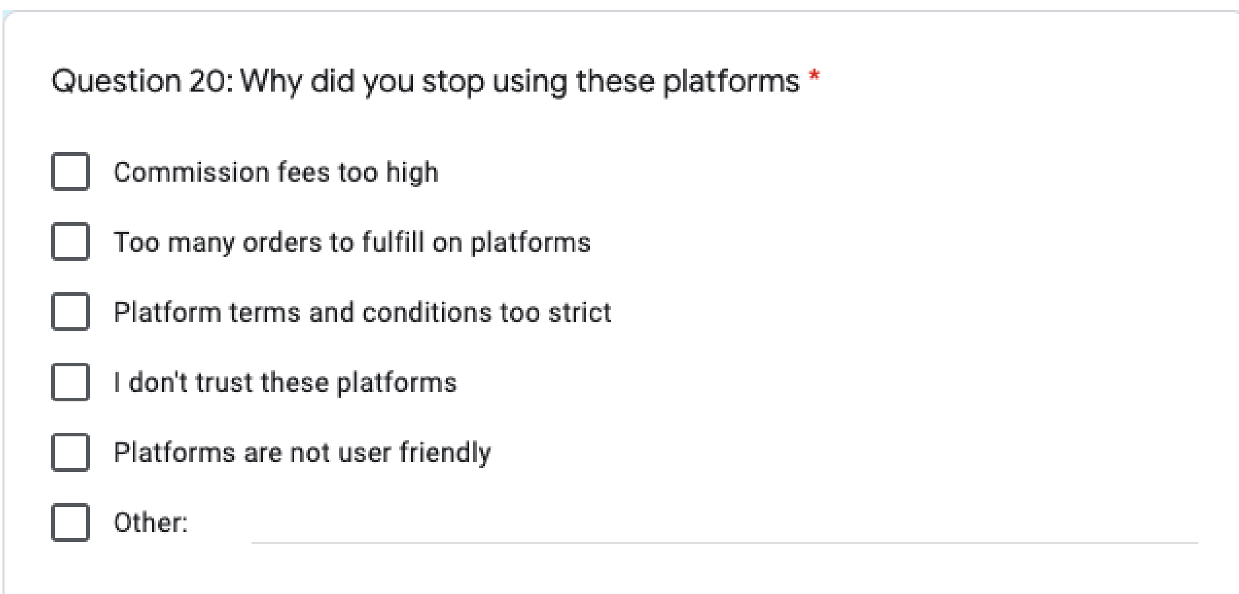
Figure 31: Question 18



Question 19: If you selected "Yes" to the above Question 17, does your business use platforms like UberEats, Doordash, or SKipTheDishes to complete delivery and takeout orders?

- ○ Yes
- ○ No
- ○ No - but we are considering to use it
- ○ No - but our business has used it previously

Figure 32: Question 19



Question 20: Why did you stop using these platforms *

- ☐ Commission fees too high
- ☐ Too many orders to fulfill on platforms
- ☐ Platform terms and conditions too strict
- ☐ I don't trust these platforms
- ☐ Platforms are not user friendly
- ☐ Other:

Figure 33: Question 20

Contact Tracing

Question 21: How many contact tracing cases of Covid-19 can be traced back to your restaurant from the dates of March 1st until March 15th? *

Your answer

Figure 34: Question 21

Statement of Consent and Privacy

You agree to participate in this survey which is intended to collect information related to the experience of restaurants in Ontario in relation to Covid-19.

Your participation in this survey is voluntary. You may refuse to take part in the research or exit the survey at any time. You are free to decline to answer any particular question you do not wish to answer for any reason.

There are no direct benefits for completing the survey. The information you provide may be used for policy action on your behalf.

There is some risk where you may find the questions involving financial information too sensitive to disclose. Those are all optional if you do not wish to provide that information.

Your survey answers will be sent to a link on Google Forms where data will be stored in a password protected electronic format. Petit Poll does not collect identifying information such as your name, email address, or IP address. Therefore, your responses will remain anonymous. No one will be able to identify you or your answers, and no one will know whether or not you participated in the study.

No personally identifiable information (PII) such as addresses, names, or postal codes shall be collected in this survey. For more information on PII and how we keep your information safe, visit: https://www.ontario.ca/document/freedom-information-and-privacy-manual/privacy-protection

Figure 35: Terms

Figure 36: The Moral of the Story

Austin, L. M., and D. Lie. 2016. *Safe Sharing Sites*. https://www.nyulawreview.org/issues/volume-94-number-4/safe-sharing-sites/.

Bascaramurty, Dakshana. 2021. *Racialized Canadians Have Some of the Highest Rates of COVID-19 Infections in the Country. Who Can Allay Their Doubts about Taking the Vaccine?* https://www.theglobeandmail.com/canada/article-racialized-canadians-need-the-covid-19-vaccine-more-urgently-than-most/.

Chang, Serina, and Emma Pierson. 2021. *Mobility Network Models of COVID-19 Explain Inequities and Inform Reopening*.

Cooney, Elizabeth. 2020. *Restaurants and Gyms Were Spring 'Superspreader' Sites. Occupancy Limits Could Control COVID, New Study Predicts*. https://www.statnews.com/2020/11/10/restaurants-gyms-were-spring-superspreader-sites-occupancy-limits-could-control-spread/.

Freitas, R. S. G. de, and Elke Stedefeldt. 2020. *COVID-19 Pandemic Underlines the Need to Build Resilience in Commercial Restaurants' Food Safety*.

Gareth James, Trevor Hastie, Daniela Witten, and Robert Tibshirani. 2017. *An Introduction to Statistical Learning with Applications in r*. https://static1.squarespace.com/static/5ff2adbe3fe4fe33db902812/t/6009dd9fa7bc363aa822d2c7/1611259312432/ISLR+Seventh+Printing.pdf.

Kaitano Dube, Godwell Nhamo, and David Chikodzi. 2020. *COVID-19 Cripples Global Restaurant and Hospitality Industry*.

Mace, Jon. 2020. *Half of Independent Restaurants Could Close Permanently, Industry Group Warns*. https://www.bnnbloomberg.ca/half-of-independent-restaurants-could-close-permanently-due-to-covid-19-industry-group-warns-1.1425910.

Madeira, A. 2020. *The Impact of Pandemic Crisis on the Restaurant Business*.

Matthew Zook, danah boyd, Solon Barocas. 2017. *Ten Simple Rules for Responsible Big Data Research*. https://journals.plos.org/ploscompbiol/article/comments?id=10.1371/journal.pcbi.1005399.

Morden, James. 2021. *The Impact of COVID-19 on the Restaurant Industry Outlook*. https://search.proquest.com/openview/6e5401711aff71c0ea5f6b2d8fdeb9bd/1?pq-origsite=gscholar&cbl=33486.

Neuwirth, Erich. 2014. *RColorBrewer: ColorBrewer Palettes*. https://CRAN.R-project.org/package=RColorBrewer.

Paul Gertler, Patrick Premand, Sebastian Martinez. 2016. *Impact Evaluation in Pratice: Second Edition*.

R Core Team. 2020. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. https://www.R-project.org/.

Rao, Vidya. 2020. *Immigrant and BIPOC Restaurateurs Face an Even Tougher COVID-19 Reality*. https://pos.toasttab.com/blog/on-the-line/immigrant-bipoc-restaurateurs-even-tougher-covid-reality.

Sander Greenland, Kenneth J. Rothman, Stephen J. Senn, and Douglas G. Altman. 2016. *Statistical Tests, p Values, Confidence Intervals, and Power: A Guide to Misinterpretations*. https://www.tellingstorieswithdata.com.

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.

Wilson, Kerissa. 2020. *Nearly 80*. https://toronto.ctvnews.ca/nearly-80-of-covid-19-cases-in-toronto-were-among-racialized-groups-in-november-1.5276918.

Wu, Changbao, and Mary Thompson. 2020. *Sampling Theory and Practice*. https://www.springer.com/gp/book/9783030442446.

Zhu, Hao. 2020. *kableExtra: Construct Complex Table with 'Kable' and Pipe Syntax*. https://CRAN.R-project.org/package=kableExtra.