A cross-sectional observational study: Impact of social distance restrictions and relevant policies on COVID-19 pandemic daily confirmed cases

Submitted to

Instructor: Shion Guha

By Group 9
Audrey Aw
Jing Yan
Jia Lei Qian

Table of Contents

1. Introduction

COVID-19 has been wreaking havoc in North America for almost a year. Both the US and Canadian governments have introduced many policies in response. Among these policies, lockdown restrictions have been one of the most controversial ones. Therefore, this study aims to use the real historical data to examine the impact and effectiveness of the stay-at-home order in New York and Ontario, two regions that COVID-19 has heavily impacted. There are two reasons for choosing those two places for comparison. First, New York is in close proximity to Ontario, and both of them are regions of rapid financial growth. Second, while the stringency index in containment and closure policies are similar in both places, the propagation of COVID-19 in New York is far greater than that of Ontario.

In this analysis, the data will be retrieved from COVID-19 Data Hub (Guidotti and Ardia, 2020). It is a worldwide, unified dataset that aggregates more than a dozen data sources such as Johns Hopkins Center for Systems Science and Engineering, Oxford COVID-19 Government Response Tracker, and Public Health Agency. One can instantly download the latest, structured, and daily historical data across several official sources, and conduct analyses to aid in better understanding COVID-19. To be specific, the original dataset contains 12 policy measures and 9 COVID-19 variables of 199 countries. It records the data from 2020-01-01 and is updated on an hourly basis. In this analysis, the dataset's location narrows down to two leading states in the US and Canada: New York and Ontario. Detailed data descriptions will be discussed in the third section of the report.

2. Literature Review & Research Questions

COVID-19 is an infectious disease that is most commonly spread through close contact with an infected individual. There are several non pharmaceutical intervention strategies that have been adopted to reduce the transmission of the virus, including physical distancing, the use of face masks and stay-at-home orders. Existing studies have shown that transmission of the virus is lowered when a physical distance of 1m or more is observed (Chu et al., 2020). Stay-at-home orders are associated with a reduction in population movement, and a decrease in close interpersonal contact outside the household, thus limiting one's potential exposure to the virus (Moreland, 2020). Before pharmaceutical interventions such as the COVID-19 vaccine are widely available, non-pharmaceutical interventions are essential in maintaining control over the virus and reducing the burden on the healthcare system.

Castillo et al. (2020) suggests that state-level stay-at-home orders in the U.S. are associated with a consistent reduction in infection rates across the 42 states that have implemented a stay-at-home order. The logged infection rate was 0.113/day pre-stay-at-home-order, as compared to 0.047/day post-stay-at-home-order. A separate study conducted by Sen et al. (2020) on the effect of stay-at-home orders and COVID-19 associated hospitalizations in 4 U.S. states noted that actual cumulative hospitalizations deviated from projected cumulative hospitalizations after a stay-at-home order was implemented, with slower exponential hospitalizations growth in all 4 states studied. Additionally, the timing of implementation of stay-at-home orders following the first reported COVID-19 case has an effect on the peak number of COVID-19 cases and deaths. An observational analysis on various countries and U.S. states with known stay-at-home orders demonstrated that a delayed implementation of stay-at-home orders is generally associated with a longer delay in reaching peak number of cases and deaths, as well as a larger overall regional burden of infection (Medline et al., 2020). A Canadian study led by Yuan et al. (2020) on the efficacy of the City of Toronto's stay-at-home order found that the implementation of stay-at-home orders allow for outbreaks to be quickly controlled, but that its effects generally manifest fully by the 65-day mark and extension of the stay-home-order beyond this point provides no additional benefit.

Most existing literature has primarily focused on the effect of stay-at-home-orders in U.S. states, with limited literature on that of Canadian provinces. As the COVID-19 response in Canada is similar to that in the U.S., where individual provinces / states work with regional health units to determine containment and mitigation strategies (Detsky and Bogoch, 2020), this study will provide a more in-depth, comparative analysis on the timing of implementation of stay-at-home orders on COVID-19 cases in New York and Ontario. Thus, the following research questions are proposed:

1. Does there exist a difference in **daily confirmed cases** under stay-at-home restriction between New York and Ontario?

2. Do the stay-at-home restrictions have an effect on the **daily confirmed cases** in New York and Ontario respectively?

3. Among all policy measures, which features will affect the **daily confirmed cases** in New York and Ontario respectively?
    - Does there exist an interaction effect between stay-at-home restrictions and information campaigns that will affect the daily confirmed cases for both places?

3. Data Description

We use the Covid-19 dataset collected from 2020-03-03 to 2021-01-31 with a sample size of 672 in total. Since we mainly focus on studying the changes in the number of daily confirmed cases in New York and Ontario, we divided the data into two data frames about these two places with a sample size of 336 in each. Our target dependent variable is daily confirmed cases and also has several independent variables.

Here is the detailed description of all variables in New York and Ontario datasets:

Covid-19 Variables (Continuous Variables):

- Deaths: Cumulative number of death under epidemic
- Confirmed: Cumulative number of covid confirmed cases
- Tests: Cumulative number of tests
- Vaccine: Cumulative number of people vaccinated
- Recovered: Cumulative number of people who restored health
- Hosp: Number of patients in the hospital on date
- Icu: Number of patients in ICUs on date
- Vent: Number of patients requiring invasive ventilation on date
- Population: Total population of each province/city.

Policy Measures(Categorical Variables):

- ***School closing***: Four levels, from taking no measures to close all schools.
- ***Workplace closing:*** Four levels, from taking no measures to close all-but-essential workplaces.
- ***Cancel events:*** Three levels, from taking no measure to have to cancel all events.
- ***Gathering restrictions:*** Five levels, from no restrictions to cannot gather more than 10 people.
- ***Transport closing:*** Three levels, from no restrictions to require closing all public transport.
- ***Stay home restrictions:*** Four levels, 0 is taking no measures to restrict, 1 is recommended to stay at home, 2 is required not leaving the house except essential shopping and exercise, 3 is required not leaving the house with minimal exceptions.
- ***Internal movement restrictions:*** Three levels, from no restrictions to require closing
- ***International movement restrictions:*** Five levels, from no restrictions to require closing all borders.
- ***Information campaigns:*** Three levels, from no public covid information campaign to coordinated public information campaigns.
- ***Testing policy:*** Four levels, from no testing policy to open public testing.
- ***Contact tracing:*** Three levels, from no contact tracing to comprehensive contact tracing.

Policy Measures(Continuous Variables):

- ***Stringency index:*** Stringency of government responses.

	Count	**Mean**	**Standard deviation**	**Minimum**	**Maximum**
Confirmed	336	483618.8	3.036410e+05	0	1410656
Recovered	311	73987.4	23025.019310	2045	127798
Deaths	323	23358.5	7646.78	3	35178

Population	336	23628065	0	23628065	23628065
Tests	336	483618.8	3.036410e+05	0	1410656
Vaccine	0	N/A	N/A	N/A	N/A
Hosp	321	4323.56	4795.05	325	18825
ICU	312	1089.07	1348.81	109	5225
Vent	270	409.86	481.04	47	2425

Table 1: Descriptive Summary table of New York Covid Variables

	Count	Mean	Standard deviation	Minimum	Maximum
Confirmed	334	66768.61976	66798.103637	15	268211
Recovered	299	63551.75251	57995.735644	8	242807
Deaths	334	2677.679641	1446.910027	0	6188
Population	334	14711827.0	0	14711827.0	14711827.0
Tests	327	3238261	2752671	3394	9295815
Vaccine	0	N/A	N/A	N/A	N/A
Hosp	0	N/A	N/A	N/A	N/A
ICU	0	N/A	N/A	N/A	N/A
Vent	0	N/A	N/A	N/A	N/A

Table 2: Descriptive Summary table of Ontario Covid Variables

4. Data Visualization

An exploratory data analysis is further performed in this part, and several exciting phenomena are found in the COVID-19 Data Hub dataset. Notice that this study narrows the investigation scope into New York and Ontario and selects the data ranging from 2020-03-03 to 2021-01-31. The visualizations are all based on data from these two distinct regions.

First, we compare the cumulative confirmed cases in New York and Ontario in their respective countries. Figure 1 and Figure 2 indicate the cumulative confirmed cases of COVID in terms of states in the US and Canada, respectively. According to those two figures, New York had the fourth-highest number of diagnoses in the United States, while

Ontario had the highest number of diagnoses in Canada. However, the total number of confirmed cases in New York (n=1,410,656) was five times in Ontario (n=272,917). As states with many people diagnosed, it is intriguing and meaningful to analyze and compare the impact of the restriction orders on both places.

Figure 1: Top 7 states with the highest cumulative confirmed cases in the US before 2021-01-31

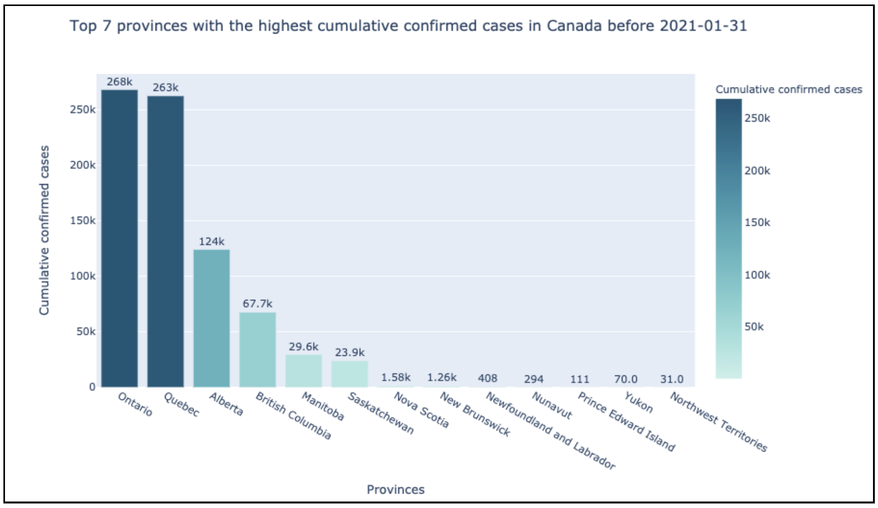

Figure 2: Top 7 provinces with the highest cumulative confirmed cases in Canada before 2021-01-31

Also, Figure 3 shows a similar stringency index pattern over the same period in both places. Stringency index captures variations in containment and closure policies. A similar pattern provides another reason for us to compare and analyze those two together.

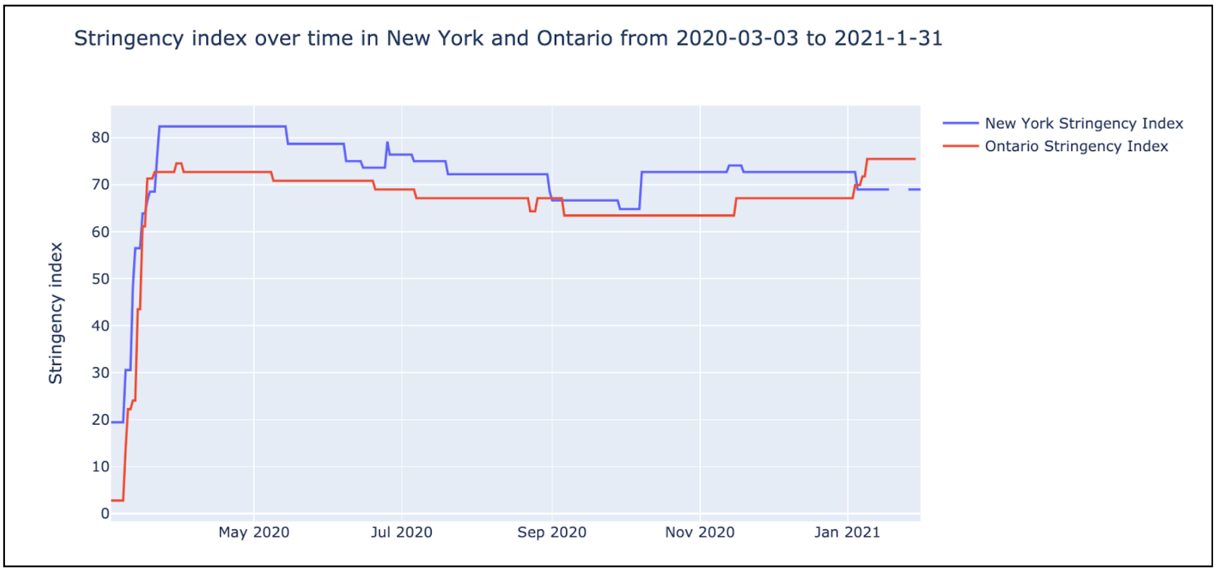

Figure 3: Stringency index over the same period of time in New York and Ontario

Besides, this study illustrates the effect of stay-home-order restrictions on daily confirmed cases in both regions. According to Figure 4, both New York and Ontario have a majority of level 1 measures, which recommends not leaving the house. The difference is that New York has stricter restrictions, while Ontario mainly has two levels of restriction.

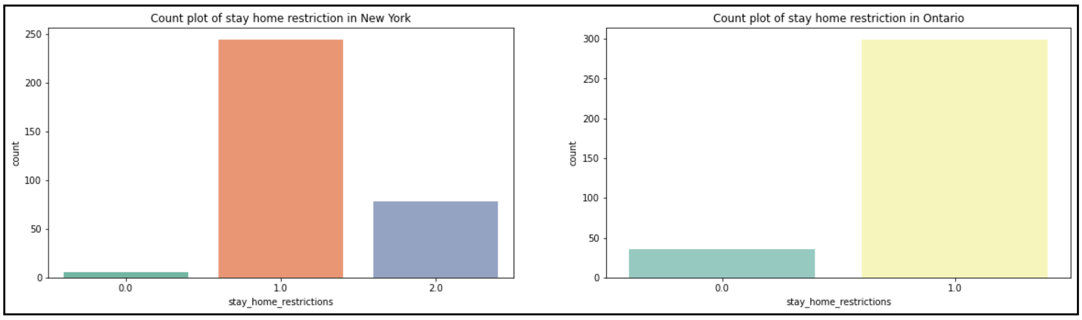

Figure 4: Count plot of stay home restrictions in New York and Ontario

Take a further look, the dynamic plot in New York in Figure 5 indicates a crucial period, March to June, as the restriction level increases to level 3. However, the plot in Ontario has no change in the level of restrictions, and Guidotti and Ardia (2020) did not indicate the reason for it.

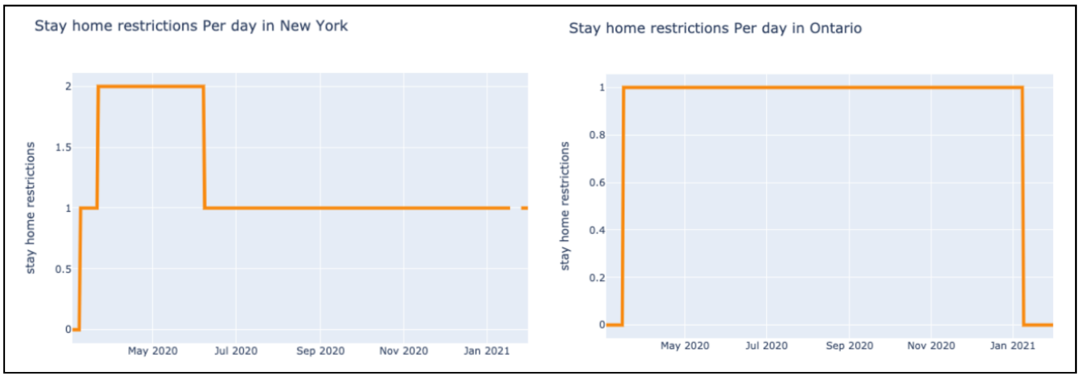

Figure 5: Stay home restrictions per day in New York and Ontario

To further investigate the effect, we draw a time series plot to show the COVID trend in both states with highlights of stay-at-home order/ shutdown restriction. According to Figure 6 shown below, there is an overall increasing trend of COVID-19 in New York from March to middle December. Both places have an increasing trend from November to January next year. Notice that although the overall trend of COVID-19 in Ontario shows an increase, the growth rate is much lower than the one in New York. Besides, in this plot, several periods are highlighted. The red one indicates that the first movement-control-order from the NY government does not perform well; the number of confirmed cases increases rapidly (Kerr, A., 2020). Therefore, the NY government extended the stay-home-order. The second yellow highlight is between the extended stay-home order and phase 1 reopening. The third orange and fourth purple highlights are the period of the shutdown with extension (COVID-19 pandemic in Ontario., 2020). Notice that as the time of movement restriction increases, the growth rate of confirmed cases becomes slower. It implies the effectiveness of quarantine as it effectively controls the social distancing among people.

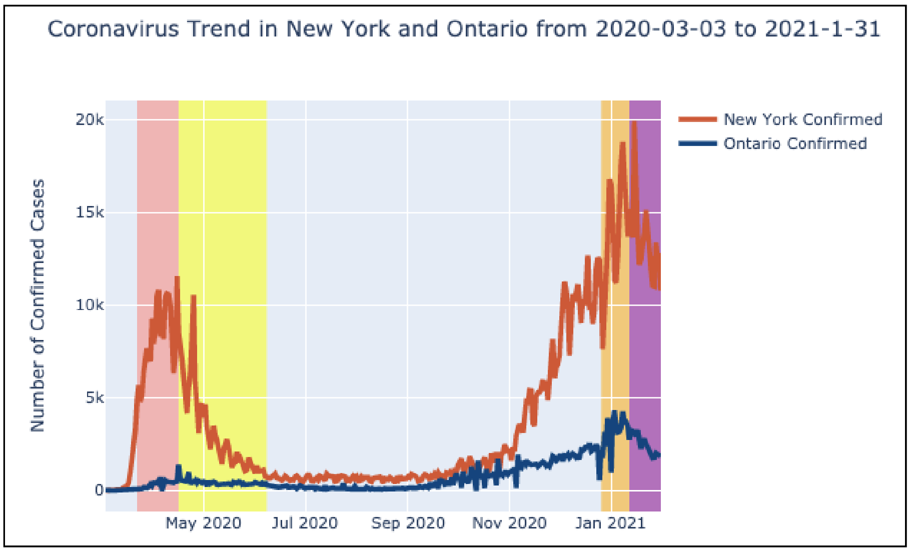

Figure 6: COVID-19 Trend in New York and Ontario from 2020-03-03 to 2021-01-31

Lastly, before using statistical models to answer the second research question, this study checks distribution of both response and explanatory variables. Figure 4.7 indicates the distributions of policy measures in New York and Ontario besides. In contrast, Figure 8 only shows the relevant continuous variables in New York as the records of those variables in the Ontario dataset are missing. Based on Figure 7, it is worth noticing that all policy measures are categorical. New York measures are usually more than one level, while the levels of measures in Ontario are all binary. Meanwhile, the variables from Figure 8 are all right-skewed. Notice that skewed data would limit model choices and affect the result's performance since the tail regions act as outliers of models; the data transformation might be required to process further analysis.

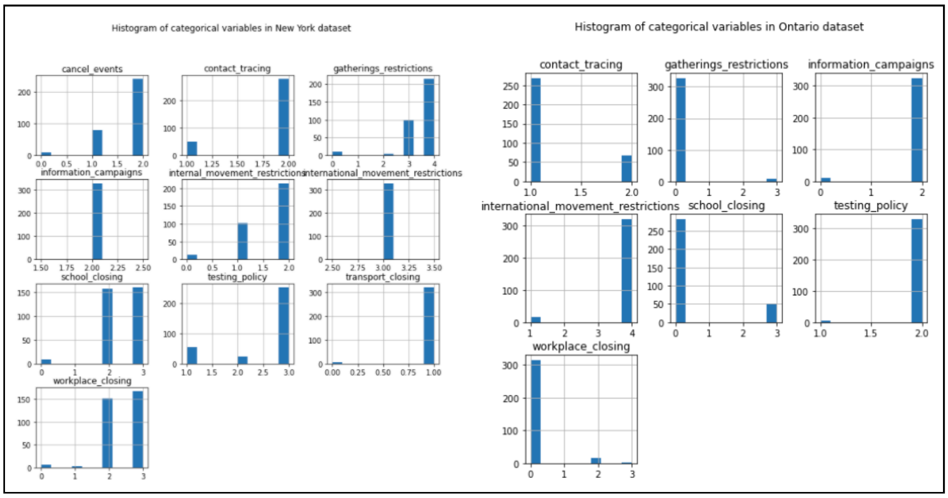

Figure 7: Distribution of categorical variables in New York and Ontario datasets

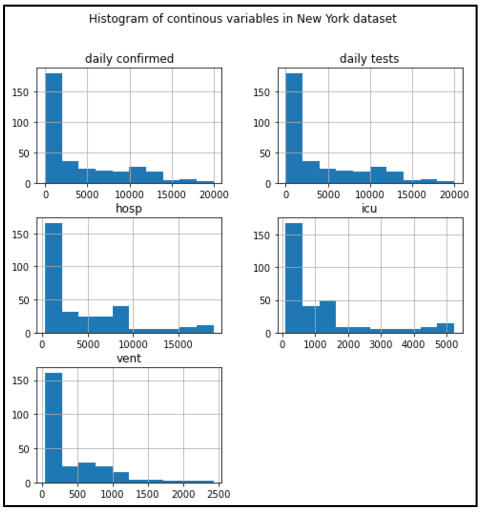

Figure 8: Distribution of continuous variables in New York dataset

5. Methodology

In the situation that specific vaccines and antivirals are not popularized, the purpose of this analysis is to add growing evidence to the topic of the relationship between staying home restrictions and the spread of COVID-19 and provides insightful suggestions based on that. This analysis will demonstrate concepts covered in the lectures and use real noisy historical data to investigate the effect of risk factors related to COVID-19. This project will explore how data science techniques help us design, execute, and analyze scientific experiments.

The detailed methodology is as follows. In this analysis, we performed an observational study based on the COVID-19 dataset from March 2020 to January 2021 and used several statistical methods to address three research questions mentioned previously. The path will be followed by data preprocessing, assumption verification, model construction, and conclusion for each research question.

In the first part, we performed an independent sample t-test to check whether there exists a difference in daily confirmed cases under stay home restriction between New York and Ontario. For the second part, this study will use two separate one-way ANOVA with corresponding post hoc tests to detect whether the stay home restrictions impact the daily confirmed cases in New York and Ontario, respectively. Lastly, this study performed a factorial analysis of variance (ANOVA) and multiple linear regressions to answer the third research question. Specifically, we investigated which level of features (policy measures) affect the daily confirmed cases in both places. We also tested the main and interaction effects between stay home restrictions and the presence of information campaigns.

6. Analysis and Discussion

6.1 Research Question 1

In the first portion of our analysis, we focused on our first research question: does there exist a difference in **daily confirmed cases under stay home restriction** between New York and Ontario? Guidotti and Ardia (2020) provided the following definitions for the levels of stay home restrictions. Level 0: no measures; Level 1: recommend not leaving house; Level 2: require not leaving house with exceptions for daily exercise, grocery shopping, and “essential” trips and Level 3: require not leaving house with minimal exceptions (e.g. leaving the house only once every few days, or only one person can leave at a time, etc.).

As observed in Figure 5 previously, New York has stricter restrictions ranging from level 0 to level 2, whereas Ontario has more lax restrictions ranging from level 0 to 1. However, as both areas were under level 1 restriction for most of the dates in our dataset, (n=244 and n=299 days for New York and Ontario respectively), we compared them on the basis of level 1 restriction using an independent sample t-test.

Approach

1. **Data Preprocessing**

Prior to performing the t-test, we reviewed the dataset for missing values under the 'daily confirmed' attribute, which indicates the number of confirmed cases per day in each administrative area. No missing values were detected. Additionally, apart from selecting a subset of the original dataframe where the stay home restriction was equal to level 1, no further preprocessing of data was required at this stage.

2. **Assumption Verification**

There are several assumptions that should be checked for an independent sample t-test analysis: (1) independence of errors, (2) statistical power, (3) homoscedasticity, (4) normality assumption.

(1) Independence of errors

Data collection for the dataset was random and performed separately, thus we may assume the independence of errors in both samples.

(2) Statistical power

A power analysis was performed to assess the adequacy of sample sizes for this analysis. The Cohen's d effect size obtained was **0.961** and the calculated required sample size for the analysis was **18.014**. As the sample size for New York and Ontario are n = 244 and n = 299 respectively, the obtained sample size is sufficient for the study. Additionally, power curves for varying effect sizes in a two sample t-test were generated as seen in Figure 9 below. At an effect size of 0.9 and above, beyond an estimated sample size of 80, the experiment reaches a plateau in terms of the power it stands to gain from an increase in sample size. Therefore, the obtained sample sizes in the dataset are sufficient to ensure good power and minimisation of the occurrence of a type II error.

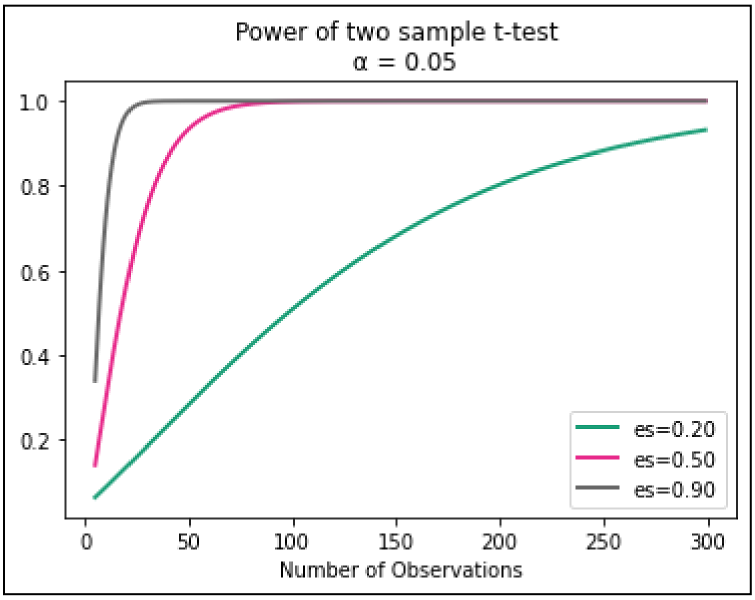

Figure 9: Power curves for varying effect sizes in a two sample t-test

(3) Homoscedasticity

Levene's test was used to assess homoscedasticity in both samples.The hypotheses for the Levene’s Test are:

H_0: The variances are equal across all samples / groups
H_a: The variances are not equal across all samples / groups

The obtained W statistic of **99.700** and p-value of **1.160e-21** (****) indicate that we may reject the null hypothesis, and that the variances are unequal for both samples.

(4) Normality assumption.
A vital assumption for t-tests is the normality assumption - that samples should be normally distributed. To verify that the data follows a Gaussian distribution, we first visualize the distribution of data with a kernel density plot.

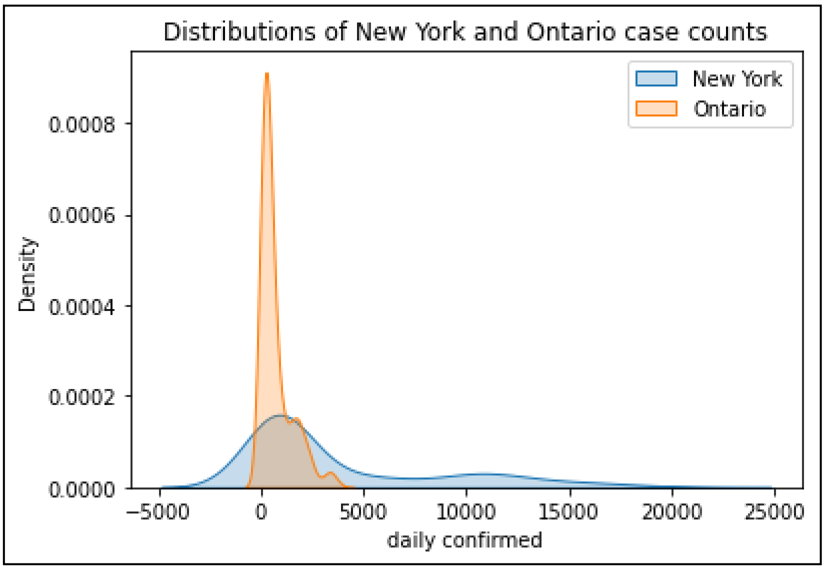

Figure 10: Kernel density distribution of New York and Ontario daily confirmed case counts

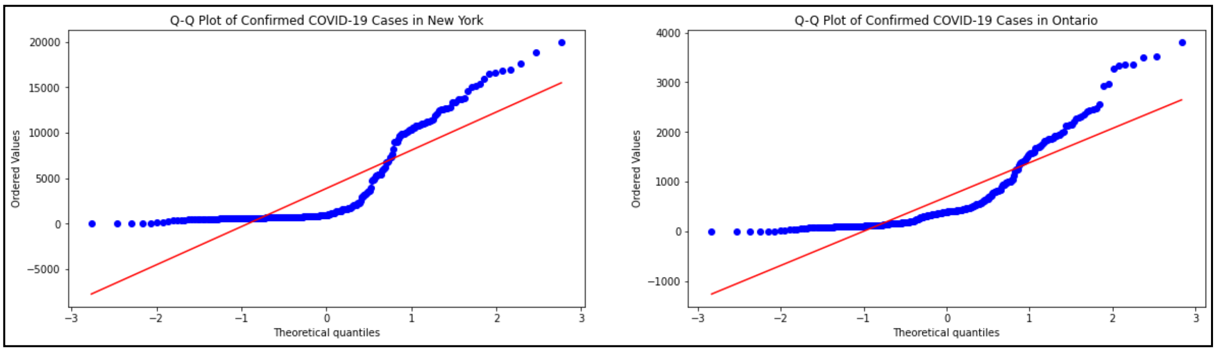

Figure 11: Q-Q Plots of New York and Ontario daily confirmed case counts

In Figure 10 above, we see that the data is skewed and the distribution appears to be non-Gaussian. The Q-Q plots in Figure 11 further demonstrate that both datasets for New York and Ontario are highly right-skewed. Additionally, the Shapiro-Wilk test was performed to check the normality of the distribution. The results obtained in the test are summarised in Table 3 as follows:

Variable: daily confirmed cases	W-statistic	P-value
New York	0.794	**≤ 0.001 (***)**
Ontario	0.787	**≤ 0.001 (***)**

(Bold numbers indicate the significant p-values; ns: p> 0.05, *: p≤ 0.05, **: p≤ 0.01, ***: p≤ 0.001, ****: p≤ 0.0001)

Table 3: Shapiro-Wilk test result for normality of distribution

When comparing two samples with non-normal distributions, the Mann–Whitney U test (Wilcoxon rank sum test) is often cited as a non-parametric test option. However, Skovlund and Fenstad (2000) propose that in the event of samples with unequal variances and skewed distributions, the Mann-Whitney U test is not appropriate. Instead, transformations are recommended to transform data to become closer to normality. Therefore, to address the non-normality issue, a Box-Cox transformation was performed.

3. Model Construction and Hypotheses

As sample sizes and variances were unequal, a Welch's two-sample t-test was performed. In the descriptive statistics section, we noted that the average number of daily confirmed cases in New York was much greater than that of Ontario. Therefore, we performed a right-tailed test to check if similarly, there is a greater number of daily confirmed cases in New York even under a stay home restriction. The hypotheses may be expressed as follows:

H_0: There is no difference in daily confirmed cases under stay home restriction between New York and Ontario
($\mu_{\text{New York}} = \mu_{\text{Ontario}}$)
H_1: The number of daily confirmed cases under stay home restriction in New York is greater than that in Ontario
($\mu_{\text{New York}} > \mu_{\text{Ontario}}$)

4. Results and Discussion

Variable: daily confirmed cases	T-statistic	P-value
New York, Ontario	4.902	**≤ 0.0001 (****)**

(Bold numbers indicate the significant p-values; ns: p> 0.05, *: p≤ 0.05, **: p≤ 0.01, ***: p≤0.001, ****: p≤ 0.0001)

Table 4: Independent two-sample t-test result on daily confirmed cases in New York and Ontario

The obtained values of t statistic = 4.902 and p-value 6.288e-07 ($p < 0.0001$) in Table 4 above suggest that the null hypothesis ($\mu_{New\ York} = \mu_{Ontario}$) may be rejected. Hence, we may conclude that the number of daily confirmed cases in New York is greater than that of Ontario under the same level of stay-home restriction (level 1). This suggests that New York residents were possibly less likely to heed the stay-home recommendation. As the COVID-19 virus is transmitted through airborne particles and contact with infected individuals, this could have led to increased opportunities for virus transmission when outside the safety of one's household.

6.2 Research Question 2

For research question 2, we committed to find whether the **stay home restrictions** have an effect on the **daily confirmed cases** in New York and Ontario respectively. Therefore, we performed two separate one-way factorial analysis of variance (ANOVAs) for both New York and Ontario to estimate the relationship between the categorical independent variable (stay home restrictions) and the dependent variable (daily confirmed cases). However, as ANOVA is an omnibus test, it tests for a difference overall between all groups, but cannot tell which group is different. Therefore, following the one-way ANOVA tests, we conducted post-hoc comparisons to identify the groups that differ.

<u>**Approach**</u>

1. **Data preprocessing**
   As mentioned in the data description and research question 1 sections, there are four levels of stay home restrictions, where New York has restrictions ranging from level 0 to level 2, and Ontario only has level 1 restrictions. We noted that there are some missing values in the Ontario dataset, likely indicating that there is no restriction measure at that time. Therefore, we set the missing value of stay_home_restrictions to level 0. After that, we re-coded the values of the stay_home_restrictions column from numeric to string.

2. **Assumption Verification**

(1) Independence
The independence assumption requires the absence of a relationship between the participants in any of the groups and that all groups are mutually exclusive. In our data, this condition was met.

(2) Normality
In order to test this assumption, the Shapiro-Wilk test was used. The null hypothesis is that the data is normally distributed. The results are listed in the following Table 5. For the New York dataset, we obtained a W test statistic of 0.860, and p-value of 1.2583e-16. Due to the p value being less than the chosen alpha level 0.05, the null hypothesis is rejected and there is evidence that the data tested is not normally

distributed. For the Ontario dataset, the obtained W test statistic is 0.843, and the p-value is less than 5% significant level, similarly indicating a non-normal distribution. To fulfill the normality assumption, a Box-Cox transformation was performed in order to transform the data to the normality form.

Variable	W-statistic	P-value
Stay home restrictions (New York)	0.860	**≤ 0.00001 (****)**
Stay home restrictions (Ontario)	0.843	**≤ 0.00001 (****)**

(Bold numbers indicate the significant p-values; ns: p> 0.05, *: p≤ 0.05, **: p≤ 0.01, ***: p≤ 0.001, ****: p≤ 0.0001)

Table 5: The results of Shapiro-Wilk test of normality for stay home restrictions in New York and Ontario

(3) Homoscedasticity
For testing equality of variances, Levene's Test was used. The null hypothesis is that all groups have equal variances. The results are listed in the following table:

Variable	P-VALUE
Stay home restriction(New York)	**≤ 0.05 (*)**
Stay home restriction(Ontario)	**≤ 0.00001(****)**

(Bold numbers indicate the significant p-values; ns: p> 0.05, *: p≤ 0.05, **: p≤ 0.01, ***: p≤ 0.001, ****: p≤ 0.0001)

Table 6: The results of Levene's Test of Homogeneity of variance for stay home restrictions

As the p-values of both New York and Ontario were less than the significant level of 0.05, the null hypothesis was rejected and indicated that the variance of the groups differ significantly. However, although the variances are unequal, it was not feasible to switch to using non-parametric ANOVA as non-parametric tests assume that the shape of the data distribution is the same in each group. Therefore, non-parametric ANOVAs are not appropriate for use in the analysis of groups with different standard deviations.

3. **Model Constructions and Hypotheses**

The following are the ANOVA hypotheses that apply to both New York and Ontario:
H_0: All group means are equal. (ie: $\mu_1 = \mu_2 = \ldots\ldots = \mu_n$)
H_a: At least one of the groups' means are different from the others.

Since the purpose of this study was to test for a difference between the stay home restriction levels of New York and Ontario, we utilized two separate one-way Anova as our model to answer the research question. The test statistic that was used was the F-statistic, which compares the mean square between samples (MS_B) to the mean square within samples(MS_w).

4. **Results and Discussion**

First of all, let's take a high level look at the variables.

Stay home restrictions	**N**	**Mean**	**Standard deviation**	**Standard error**	**95% Confidence Interval**
Level 0	6	21.0000	24.025	9.808	[-4.2127, 46.2127]
Level 1	244	546820.984	276330.537	17690.250	[511975.1824, 581666.7849]
Level 2	78	256100.769	115001.999	13021.412	[230171.8249, 282029.7136]

Table 7: Summary of stay home restrictions levels for New York dataset

Stay home restrictions	**N**	**Mean**	**Standard deviation**	**Standard error**	**95% confidence interval**
Level 0	34	164598.706	116368.303	19956.999	[123995.885, 205201.526]
Level 1	299	55867.385	47542.859	2749.475	[50456.538, 61278.231]

Table 8: Summary of stay home restrictions levels for Ontario dataset

From Table 7 and Table 8 above, it can be found that for both New York and Ontario, out of 336 sample sizes for each, more than 200 days are in the Level 1 (recommended to stay at home). For New York, the mean confirmed cases in level 1 is around 546821, and 256100 confirmed cases in level 2. For Ontario, the confirmed cases in level 0 is 164598, and confirmed cases in level 1 is 55867. It is apparent that the number of confirmed cases are lower with a stricter level of stay home restriction.

	Sum of squares	**Degree of freedom**	**F statistics**	**P-value**
Stay home restrictions	6.332×10^{12}	2.0	52.570	**≤ 0.0001 (****)**
Residual	1.957×10^{13}	325.0	N/A	N/A

(Bold numbers indicate the significant p-values; ns: p> 0.05, *: p≤ 0.05, **: p≤ 0.01, ***: p≤ 0.001, ****: p≤ 0.0001)

Table 9: Values obtained after conducting One-way ANOVAs for New York dataset

	Sum of squares	**Degree of freedom**	**F statistics**	**P-value**
Stay home restrictions	3.609×10^{11}	1.0	106.623	**≤ 0.0001 (****)**
Residual	1.120×10^{12}	331.0	N/A	N/A

(Bold numbers indicate the significant p-values; ns: p> 0.05, *: p≤ 0.05, **: p≤ 0.01, ***: p≤ 0.001, ****: p≤ 0.0001)

Table 10: Values obtained after conducting One-way ANOVAs for Ontario dataset

In order to determine the relationship between variables of stay home restrictions and daily confirmed cases, one-way Anova of New York and Ontario datasets were conducted. According to the above Table 9 and Table 10, the obtained values of F-statistics is 52.570 for New York and 106.623 for Ontario. In addition, the p-values generated by these two datasets are both quite small, which are close to 0. These indicated that the null hypothesis should be rejected, meaning that there is a statistically significant difference between the groups (levels of stay home restrictions) and their effects on the daily confirmed cases, and at least one of the groups' mean is different from the others.

To discover which groups differed significantly from each other, conducted post-hoc tests with the approach of Tukey Honestly Significant Difference. Tukey's HSD is able to test all pairwise group comparisons while controlling for the multiple comparison. The results of Tukey test are as follows, where "group1" and "group2" columns are the groups being compared, and “reject” is the decision rule based on the corrected p-value :

Group 1	**Group 2**	**Mean difference**	**Corrected P-value**	**95% confidence interval**	**Reject**
Level 0	Level 1	546799.983	**≤ 0.001(***)**	[223713.036, 869886.930]	True
Level 0	Level 2	256079.769	> 0.05 (ns)	[-75155.568, 587315.106]	False
Level 1	Level 2	-290720.214	**≤ 0.001(***)**	[-392416.635, -189023.793]	True

(Bold numbers indicate the significant p-values; ns: p> 0.05, *: p≤ 0.05, **: p≤ 0.01, ***: p≤ 0.001, ****: p≤ 0.0001)

Table 11: Multiple Comparison of Means - Tukey HSD of New York dataset

Group 1	**Group 2**	**Mean difference**	**Corrected P-value**	**95% confidence interval**	**Reject**
Level 0	Level 1	204502.615	**≤ 0.01 (**)**	[20518.952, 388486.277]	True

(Bold numbers indicate the significant p-values; ns: p> 0.05, *: p≤ 0.05, **: p≤ 0.01, ***: p≤ 0.001, ****: p≤ 0.0001)

Table 12: Multiple Comparison of Means - Tukey HSD of Ontario dataset

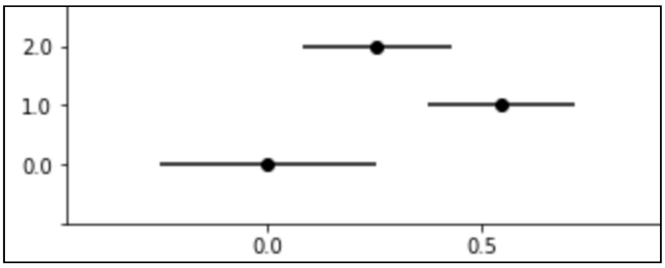

Figure 12: Multiple Comparisons Between All Pairs of New York dataset

The Tukey HSD indicates that, in the New York dataset, there is a statistically significant difference in confirmed cases between stay home restriction of level 0(no measures to restrict) and level 1(recommended to stay at home), level 1(recommended to stay at home) and level 2(required not leaving the house except essential shopping and exercise), because the decision rules of these two groups are true. In the Ontario dataset, since there are only two restriction levels, the comparison of means between these two levels is also significantly different, no other groups differ significantly. In conclusion, the mean of confirmed cases in each restriction level are different.

6.3 Research Question 3

To further explore the relationship between daily confirmed cases and different policy measures and answer the third research question, this study mainly performed two statistical analyses on both New York and Ontario datasets: **factorial analysis of variance (ANOVA) linear regression**. To be specific, multi-factors ANOVAs were conducted to test the main and interaction effects among selected categorical policy measures features for both places. Instead of using post-hoc tests, multiple linear regressions were performed followed by the ANOVA to investigate which level of features (policy measures) will contribute to the daily confirmed cases in both places. Also, this study utilized the Helmert encoding method for each predictor. This coding method allowed us to compare each level of a variable with the mean of the variable's previous levels. The entire statistical analysis path was followed by data preprocessing, assumption verification, model construction, and discussion.

<u>**Approach**</u>

1. **Data Preprocessing**
   This study set the data preprocessing as the first step of the approach. We illustrated bar plots of missing values for both datasets. The x-axis represents policy measures, the left y-axis shows the percentage of missing values, and the left y-axis means the sample size. According to Figure 13 shown below, the New York dataset did not contain missing values, while seven variables in Ontario datasets had missing data of more than 70%. Therefore, there was no need to clean the New York dataset, and all policy measure features were selected for further analysis. Meanwhile, the variables with a massive amount of missing values were removed, and this analysis only chose six explanatory variables for analyzing the Ontario case. The existing missing values for each variable were replaced by 0, indicating no information or no detailed policy measure. In this case, the selected features were: stay home restrictions, international movement restrictions, testing policy, contact tracing, information campaigns, and stringency index.

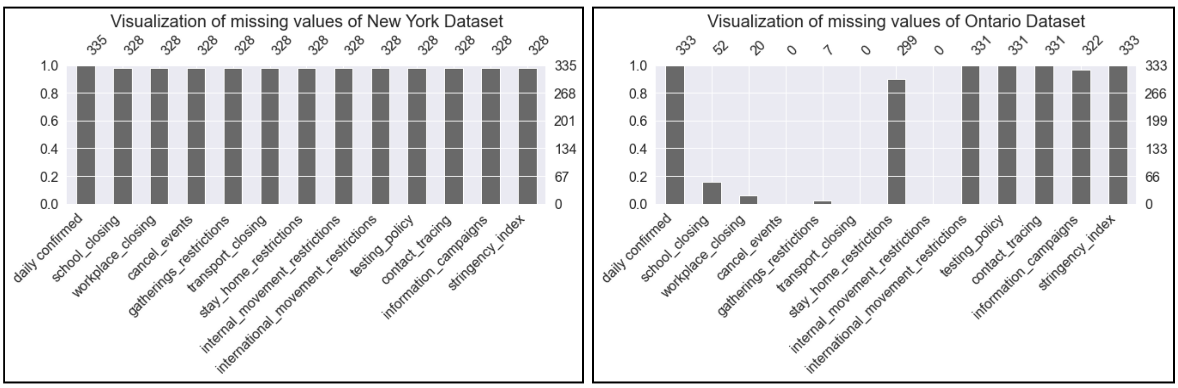

Figure 13: Visualization of missing values of New York and Ontario datasets

2. Assumption Verification

Note that the ANOVA model is a special case of a regression model in which all predictors are categorical variables. Indeed, a (multiple) linear regression model and an ANOVA model are mathematically the same, except for their encoding method and ways of interpretation. Before constructing linear regression and factorial ANOVA models, it is critical to check and fulfill four key assumptions: (1) Observations are independent of each other (2) No multicollinearity, (3) Normality of dependent variable, and (4) Homoscedasticity.

(1) Independence of response variables

In this study, the response variables are the daily confirmed cases in New York and Ontario. It is worth noticing that the original dataset is time-series data with a cumulative number of cases. Converting the cumulative cases to daily cases can avoid repeated measures. Also, all treatment groups in our analysis are mutually exclusive. Therefore, based on the designed study, the first assumption is satisfied.

(2) No multicollinearity

Correlation matrices for both New York and Ontario were built to examine the existence of multicollinearity. There is the fact that an absolute correlation coefficient over 0.7 among two or more variables indicates a strong correlation and the presence of multicollinearity. Hence, the analysis selected the explanatory variables with a correlation lower than 0.7 of other explanatory variables and a higher correlation with the dependent variable. The heatmaps of correlation matrices for both New York and Ontario scenarios after removing highly correlated predictors are demonstrated in Figure 14 and Figure 15. Both figures used a color scale to illustrate the correlation among all dependent and independent variables. The lighter the color, the less relationship between two variables, and vice versa. According to those two heatmaps, it is interesting to notice that all predictors have a small association with the response variables in both scenarios. Also, the variable daily confirmed cases in New York have a slight positive correlation with gathering restrictions and testing policy. In contrast, Ontario's one has a small positive relationship with both stringency index and information campaigns.

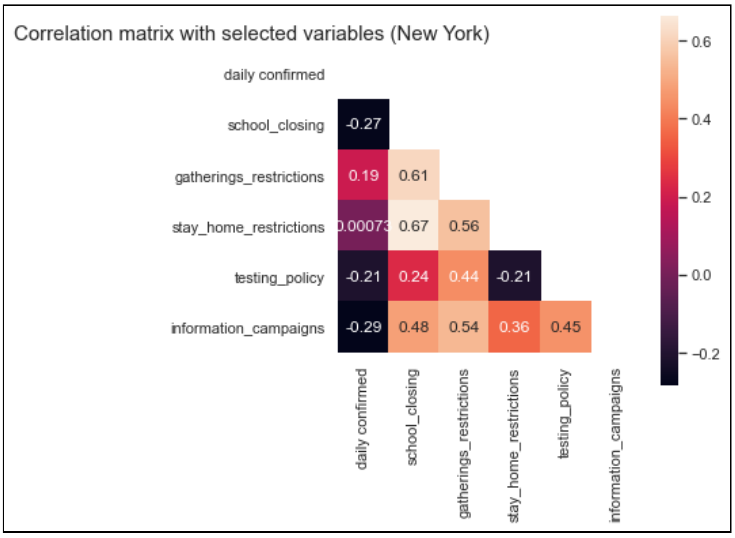

Figure 14: Correlation matrix visualization of daily confirmed cases in New York

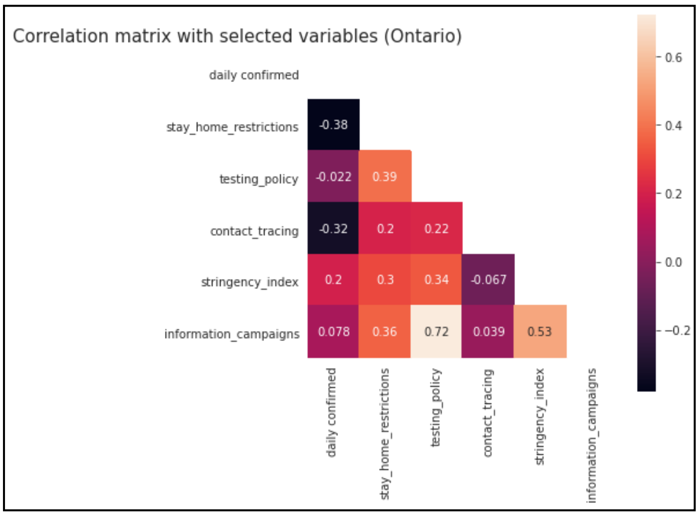

Figure 15: Correlation matrix visualization of daily confirmed cases in Ontario

(3) Normality

Previous sections have discussed this assumption. The distribution plot in the data description section (Figure 8) shows that both response variables are right-skewed. In response to the two previous research questions, Shapiro-Wilk tests were conducted to prove this conclusion further. Performing Box-Cox transformation is one of the ways to address the non-normality issue. Specifically, it is a data-transform method that can perform a range of power transformations on non-normal dependent variables. It evaluates a suitable transformation automatically and selects the best fit to transfer.

(4) Homoscedasticity

Moreover, this analysis carried out a series of Levene's Tests for each explanatory variable to verify the assumptions of homoscedasticity. Instead of Barlett's Tests, Levene's Test is preferred for a dataset that does not follow normal distribution.

The hypotheses for the Levene's Test are:
H_0: The variances are equal across all sample/ groups
H_a: The variances are not equal across all samples/ groups

Table 13 lists the results of Levene's Test of Homogeneity of variance for all explanatory variables with corresponding p-values. According to Levene's Tests, only the testing policy and information campaigns in both New York and Ontario datasets are insignificant, indicating that those four variables satisfy the equal variance assumption. However, this analysis decided not to remove any explanatory variables for further research because several studies oppose testing homogeneity and choose whether to assume it. For instance, Coombs et al. (1996, p. 148) have suggested that a parametric test is generally acceptable in controlling the type I error rate with a large sample size, even if the assumption of equal variance is violated.

Explanatory variables (New York)		**P-value**	**Explanatory variables (Ontario)**		**P-value**
Equal variance	Testing policy	> 0.05 (ns)	**Equal variance**	Testing policy	> 0.05 (ns)
	Information campaigns	> 0.05 (ns)		Information campaigns	> 0.05 (ns)
Unequal variance	Stay home restrictions	≤ 0.05 (*)	**Unequal variance**	Stay home restrictions	**≤ 0.00001 (****)**
	School closing	**≤ 0.0001 (****)**		Contacting tracing	**≤ 0.00001 (****)**
	Gathering restrictions	**≤ 0.0001 (****)**		Stringency index	**≤ 0.00001 (****)**

(Bold numbers indicate the significant p-values; ns: p> 0.05, *: p≤ 0.05, **: p≤ 0.01, ***: p≤ 0.001, ****: p≤ 0.0001)

Table 13: The results of Levene's Test of Homogeneity of variance for all explanatory variables

Furthermore, it is essential to perform a power analysis before running the models because the experimental results with small power will lead to wrong conclusions, which will affect the decision-making process. This section used a software called G.Power (Faul et al.) to perform a power analysis for factorial ANOVA and linear regression. Figure 16 shows the power analysis result: The total sample size of 75 is enough to get the effect size of 0.4 with 0.05 error probability and 80% power, which means the current datasets are feasible and able to help us to obtain strong power in this analysis.

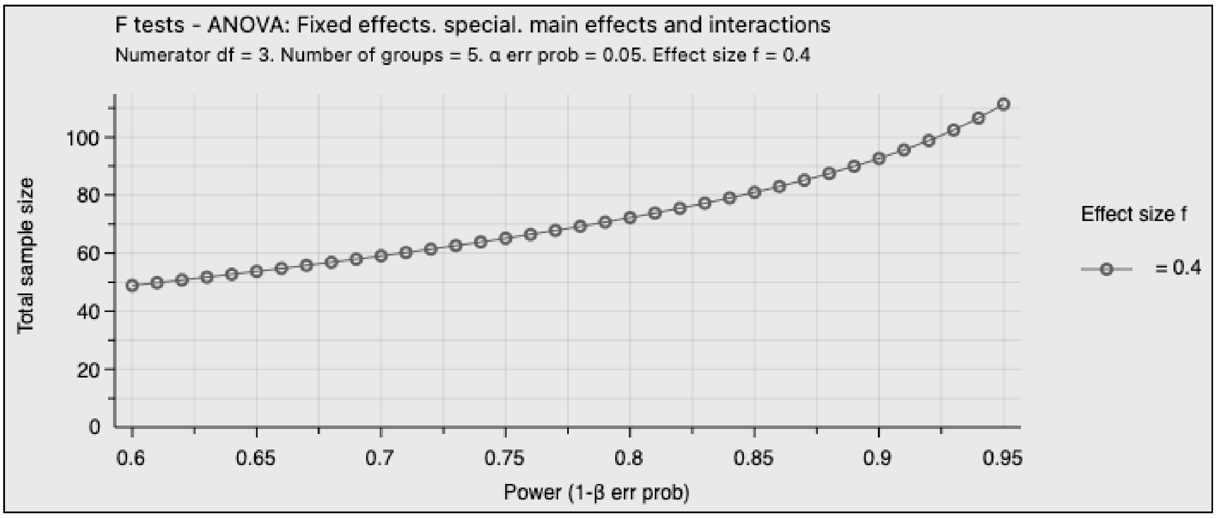

Figure 16: The results of power analysis obtained by G.Power

3. **Model Constructions and Hypotheses:**
The models and hypotheses for factorial ANOVAs and multiple linear regressions are discussed in the following two parts:

(1) Multi-factor ANOVAs and Multiple linear regression without interaction term

We first conducted factorial ANOVAs with five independent variables for both New York and Ontario cases. An experimental design with factorial ANOVAs allows researchers to explore multiple independent variables effect change while considering individual main effects for each other. For this research question, the linear models for detecting the main effects of different policies on daily confirmed cases in New York and Ontario can be expressed as Equation (1) and Equation (2), where βs are the regression parameters and ε is the random error.

Model (Main Effect)
Daily confirmed cases in New York = $\beta_0+\beta_1\cdot$(testing policy)+$\beta_2\cdot$(information campaigns)+ $\beta_3\cdot$(stay home restrictions)+$\beta_4\cdot$(school closing) + $\beta_5\cdot$(gathering restrictions) +ε ---(1)

Daily confirmed cases in Ontario= $\beta_0+\beta_1\cdot$(testing policy)+$\beta_2\cdot$(information campaigns)+ $\beta_3\cdot$(stay home restrictions)+$\beta_4\cdot$(contacting tracing) + $\beta_5\cdot$(stringency index) +ε --(2)

Hypotheses:
H_0: The means are equal across all sample/ groups (There is no main effect for the explanatory variable)
H_a: The means are not equal among all samples/ groups (there exists a main effect for the explanatory variable)

(2) Two-factor ANOVAs and multiple linear regression with interaction term

This analysis also detects the interaction effect between stay home restriction and public information campaign using two-way ANOVAs and multiple linear regression models for both places. The interaction effects model can be referred to as Equation (3), where β3 represents the unique effect of information campaigns on stay-at-home restrictions.

Model (Interaction Effect)

Daily confirmed cases in New York/Ontario = $\beta_0+\beta_1\cdot$(stay home restrictions)$+\beta_2\cdot$(information campaigns)$+\beta_3\cdot$(stay home restrictions) $\cdot$(information campaigns)$+\varepsilon$--(3)

Hypotheses:

H_0: There is no interaction effect of stay home restrictions and information campaigns ($\beta_3=0$)

H_a: There exists a main effect stay home restrictions and information campaigns ($\beta_3\neq0$)

4. Results and Discussion:

(1) The main effect of policy measures on daily confirmed cases

The use of multi-way factorial ANOVA without interaction term on both New York and Ontario datasets with each policy measure variable as grouping variables and the daily confirmed cases as dependent variables returned results shown in Table xxx. According to Table 14, the p-values for all policy measures under the New York dataset, except stay-at-home restrictions, are less than the 5% significant level. It implies the existence of main effects for testing policy, public information campaigns, school closing, and gathering restrictions on the daily confirmed cases in New York. Interestingly, for Ontario, the results for both testing policy and public information campaigns are insignificant, presenting that only stay-at-home restrictions, contacting tracing, and stringency index have main effects on the daily confirmed cases in Ontario, respectively.

New York			Ontario		
Variables	**F statistics**	**P-value**	**Variables**	**F statistics**	**P-value**
Testing policy	163.446	**≤ 0.0001 (****)**	Testing policy	1.082	> 0.05 (ns)

Information campaigns	122.093	**≤ 0.0001 (****)**	Information campaigns	2.219	> 0.05 (ns)
Stay home restrictions	1.166	> 0.05 (ns)	Stay home restrictions	26.000	**≤ 0.0001 (****)**
School closing	58.776	**≤ 0.0001 (****)**	Contacting tracing	15.295	**≤ 0.0001 (****)**
Gathering restrictions	94.743	**≤ 0.0001 (****)**	Stringency index	33.444	**≤ 0.0001 (****)**

(Bold numbers indicate the significant p-values; ns: $p > 0.05$, *: $p \leq 0.05$, **: $p \leq 0.01$, ***: $p \leq 0.001$, ****: $p \leq 0.0001$)

Table 14: Values obtained after conducting Multi-Factor ANOVAs for both New York and Ontario datasets

Following the factorial ANOVAs, we performed multiple linear regression analyses for both the New York and Ontario datasets to detect the effect of a detailed level of features that contribute to the daily confirmed cases. The results with coefficients and corresponding p-values are shown in Table 15. The intercepts estimate the average treatment effect on the daily confirmed cases, and the levels of features with significant results have been bolded. The variables with bolded results mean it has a main effect on the daily confirmed cases. Since the response variables have been transformed and the explanatory variables have been encoded using the Helmert method, it is hard to interpret coefficients' meaning. Besides, Table 15 shows the coefficient of multiple determination (R^2) in both cases. R^2 represents the percentage of daily confirmed cases where the selected predictors can explain variation. In this case, the coefficient of multiple determination for the New York model (0.72) is much higher than the one for the Ontario model, indicating that the New York model fits the data better.

New York (R^2 = 0.720)			**Ontario (R^2 = 0.329)**		
Variables	**Coefficient**	**P-values**	**Variables**	**Coefficient**	**P-values**
Intercept	15.886	**≤ 0.0001 (****)**	**Intercept**	2.284	> 0.05 (ns)
Testing policy (level 1)	-8.255	**≤ 0.0001 (****)**	**Testing policy (level 1)**	-1.100	> 0.05 (ns)

Testing policy (level 2)	-3.735	**≤ 0.0001 (****)**	**Testing policy (level 2)**	0.163	> 0.05 (ns)
Testing policy (level 3)	-2.447	**≤ 0.0001 (****)**			
Information campaigns (level 2)	-1.965	**≤ 0.0001 (****)**	**Information campaigns (level 2)**	1.056	> 0.05 (ns)
Stay home restrictions (level 1)	1.414	> 0.05 (ns)	**Stay home restrictions (level 1)**	-1.436	**≤ 0.0001 (****)**
Stay home restrictions (level 2)	0.428	> 0.05 (ns)			
School closing (level 2)	0.249	> 0.05 (ns)	**Contact tracing (level 1)**	-0.731	> 0.05 (ns)
School closing (level 3)	-1.256	**≤ 0.05 (*)**	**Contacting tracing (level 2)**	-0.944	**≤ 0.0001 (****)**
Gathering restrictions (level 2)	1.363	> 0.05 (ns)	**Stringency index**	0.120	**≤ 0.0001 (****)**
Gathering restrictions (level 3)	4.098	**≤ 0.0001 (****)**			
Gathering restrictions (level 4)	3.179	**≤ 0.0001 (****)**			

(Bold numbers indicate the significant p-values; ns: p> 0.05, *: p≤ 0.05, **: p≤ 0.01, ***: p≤ 0.001, ****: p≤ 0.0001)

Table 15: Values obtained after conducting Multiple Linear Regressions for both New York and Ontario datasets

(2) The interaction effect between stay home restrictions and public information campaigns on daily COVID confirmed cases

A two-way ANOVA with interaction was performed for both New York and Ontario datasets respectively to detect the interaction effect between stay-at-home restrictions and public information campaigns on daily COVID confirmed cases. The result is presented in Table 16. It is noteworthy that the results of the New York scenario are all insignificant ($p > 0.05$), while the Ontario results are all significant ($p < 0.05$). The interaction plot shown in Figure 17 also proves no interaction effect for New York cases since there is only one downward line and a dot. Simultaneously, Figure 18 shows the interaction effect is significant between those two policies on Ontario's daily confirmed cases as the lines are not parallel. One can conclude that interaction of both information campaigns and stay home restrictions significantly affect the daily confirmed cases in Ontario. The relationship between stay home restrictions and Ontario daily confirmed cases depends on the levels of public information campaigns.

	New York		Ontario	
Variable	**F statistics**	**P-value**	**F statistics**	**P-value**
Stay home restrictions	4.095×10^{-10}	> 0.05 (ns)	22.451	**≤ 0.0001 (****)**
Information campaigns	-2.880×10^{-10}	> 0.05 (ns)	36.977	**≤ 0.0001 (****)**
Stay home restrictions × Information campaigns	0.703	> 0.05 (ns)	4.411	**≤ 0.05 (*)**

(Bold numbers indicate the significant p-values; ns: p> 0.05, *: p≤ 0.05, **: p≤ 0.01, ***: p≤ 0.001, ****: p≤ 0.0001)

Table 16: Results of two-way ANOVA with interaction for both New York and Ontario datasets

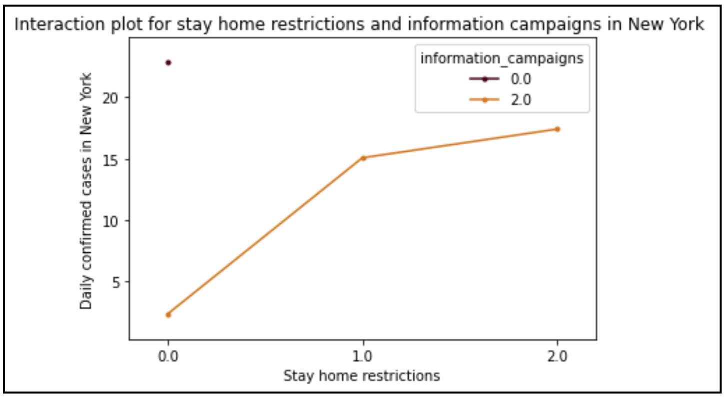

Figure 17: Interaction plot for interaction effect between stay home restrictions and information campaigns in New York

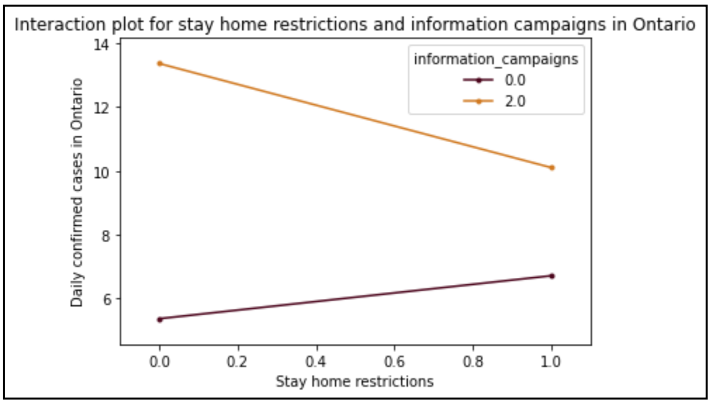

Figure 18: Interaction plot for interaction effect between stay home restrictions and information campaigns in Ontario

Furthermore, this study used a multiple linear regression model to detect the detailed interaction effects between stay home restrictions and public information campaigns, and the results are presented in Table 17. From Table 17, we can conclude that:

- The intercept means that if there is no public information campaigns and no testing policy, we can expect an average of 5.374 increases of daily confirmed cases in Ontario.
- The coefficient on level 1 stay home restrictions (i.e., Recommend not leaving home) is positive and insignificant, which means we cannot reject any hypothesis that said that the policy of recommending not leaving home would lead to more daily confirmed cases in Ontario in this model.
- The coefficient on the interaction term is negative and statistically significant, implying that the interaction of the policy of recommending not leaving home and the presence of coordinated public information campaigns have an negative impact on the Ontario daily confirmed cases.

Ontario (R^2=0.132)		
	Coefficient	**P-value**
Intercept	5.374	**≤ 0.0001 (****)**
Stay home restrictions (level 1)	1.345	> 0.05 (ns)
Information campaigns (level 2)	7.988	**≤ 0.0001 (****)**
Stay home restrictions (level 1) × information campaigns (level 2)	-4.605	**≤ 0.05 (*)**

(Bold numbers indicate the significant p-values; ns: p> 0.05, *: p≤ 0.05, **: p≤ 0.01, ***: p≤ 0.001, ****: p≤ 0.0001)

Table 17: Results of multiple linear regression with interaction for Ontario dataset

7. Conclusion

In this study, we investigated the effectiveness of stay-at-home restriction in New York and Ontario, using statistical analysis techniques such as t-tests, one-way ANOVAs, multi-factors ANOVAs, and multiple linear regressions to answer our research questions.

In our first research question, we noted that the number of daily confirmed cases in New York is greater than that of Ontario under the same level of stay-home restriction (level 1). As discussed in section 6.1 above, the significantly more rampant spread of COVID-19 in New York over Ontario is an indication of the efficacy of state / province-level response to the pandemic, as well as the response of individuals to the pandemic. This could pose further interesting social science research questions, such as the effects of individualist versus collectivist cultural norms in the event of an epidemic.

For the second research question, first of all, according to the summary tables, we found the number of confirmed cases decreased with a stricter level of stay home restriction in both New York and Ontario. From the results of ANOVA tables, it shows that stay home restrictions do affect confirmed cases, and at least one of the restriction levels is different from others. In order to discover which specific groups differed significantly, we did further research with post-hoc tests and came to the conclusion that stay home restriction of level 0 and level 1, level 1 and level 2 are differ from each other in New York dataset, and level 0 and level 1 are also significantly different in Ontario dataset.

According to the third research question results, it is interesting to see that the results were quite different in New York and Ontario. Specifically, the stay home restrictions in a multi-factor model do not significantly influence the daily confirmed cases in New York.

Instead, all other selected policies (testing policy, information campaigns, school closing, and gathering restrictions) significantly impact the daily confirmed cases in New York. Also, there is no interaction effect between stay-at-home regulation and public information campaigns in the New York case. On the other hand, stay home restriction significantly impacts Ontario's daily confirmed cases in a multi-factor model. Also, there is a negative interaction effect of the policy of recommending not leaving home and coordinating public information campaigns towards the Ontario daily confirmed cases.

Shortcomings and limitations exist in this study as well. First, the original dataset does not satisfy the normality and homoscedasticity assumptions for traditional ANOVAs and linear models. With unbalanced cell sizes, the categorical variables fail to fulfill the assumption of equal variance, which leads to biased results for the linear regression, and dramatically affecting statistical power and Type I error rates. We could use resampling methods such as bootstrapping to address this concern. Second, in research question 3, we selected the predictor based on the correlation coefficient to minimize the intercorrelation among independent variables. However, as noted, the coefficients of determination (R^2s) for the multiple linear regression models for Ontario cases are small, indicating that Ontario's policy measure features are not explaining much in the variation of the daily confirmed cases. Therefore, some better approaches exist and are also worth trying. For example, we may try the Principal Component Analysis (PCA) to construct the linearly uncorrelated variables for good model performances in the future.

Through this case study, we were able to explore and practice experimental design concepts covered in the course. For example, Helmert encoding which was used to code categorical variables when conducting Multiple Regression Analysis in section 6.3. This project provided us with the opportunity to demonstrate our understanding of the following concepts: pre-experimental tests to ensure robustness of statistical models, post-hoc tests to support the results of ANOVAs, other vital concepts such as interaction effects between variables, as well as skills not covered in the course (i.e. dealing with samples which have a non-Gaussian distribution). Additionally, we honed our data visualisation skills, which is a necessary skill in every data scientist's toolkit, as compelling data visualisations are extremely helpful when communicating statistical findings to the general population.

While the COVID-19 pandemic has significantly affected the lives of individuals globally in the past year, it has provided us with the opportunity to glean important lessons in relation to success and failures in virus containment strategies from a statistical perspective. This pandemic follows a string of other coronaviruses, from the SARS outbreak in 2003 to the emergence of MERS in 2012, and has resulted in more deaths than both SARS and MERS combined. It is our hope that this study will provide insight into the efficacy of stay-at-home orders in two regions of geographical proximity, and aid in decision-making processes relating to virus containment, if necessary, in the future.

References

Castillo, R. C., Staguhn, E. D., & Weston-Farber, E. (2020). The effect of state-level stay-at-home orders on COVID-19 infection rates. *American Journal of Infection Control*, *48*(8), 958–960. https://doi.org/10.1016/j.ajic.2020.05.017

Chu, D. K., Akl, E. A., Duda, S., Solo, K., Yaacoub, S., Schünemann, H. J., Chu, D. K., Akl, E. A., El-harakeh, A., Bognanni, A., Lotfi, T., Loeb, M., Hajizadeh, A., Bak, A., Izcovich, A., Cuello-Garcia, C. A., Chen, C., Harris, D. J., Borowiack, E., … Schünemann, H. J. (2020). Physical distancing, face masks, and eye protection to prevent person-to-person transmission of SARS-CoV-2 and COVID-19: A systematic review and meta-analysis. *The Lancet*, *395*(10242), 1973–1987. https://doi.org/10.1016/S0140-6736(20)31142-9

COVID-19 pandemic in Ontario. (2020). Retrieved from https://en.wikipedia.org/wiki/COVID-19_pandemic_in_Ontario

Detsky, A. S., & Bogoch, I. I. (2020). COVID-19 in Canada: Experience and Response. *JAMA*, *324*(8), 743–744. https://doi.org/10.1001/jama.2020.14033

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods, 39*, 175-191.

Guidotti, E. & Ardia, D., (2020), "COVID-19 Data Hub", *Journal of Open Source Software 5 51*:(2376), https://doi.org/10.21105/joss.02376

Kerr, A. (2020). A Historical Timeline of COVID-19 in New York City. Retrieved from https://www.investopedia.com/historical-timeline-of-covid-19-in-new-york-city-5071986

Medline, A., Hayes, L., Valdez, K., Hayashi, A., Vahedi, F., Capell, W., Sonnenberg, J., Glick, Z., & Klausner, J. D. (2020). Evaluating the impact of stay-at-home orders on the time to reach the peak burden of Covid-19 cases and deaths: Does timing matter? *BMC Public Health*, *20*(1), 1750. https://doi.org/10.1186/s12889-020-09817-9

Moreland, A. (2020). Timing of State and Territorial COVID-19 Stay-at-Home Orders and Changes in Population Movement—United States, March 1–May 31, 2020. *MMWR. Morbidity and Mortality Weekly Report*, *69*. https://doi.org/10.15585/mmwr.mm6935a2

Sen, S., Karaca-Mandic, P., & Georgiou, A. (2020). Association of Stay-at-Home Orders With COVID-19 Hospitalizations in 4 States. *JAMA*, *323*(24), 2522–2524. https://doi.org/10.1001/jama.2020.9176

Skovlund, E., & Fenstad, G. U. (2001). Should we always choose a nonparametric test when comparing two apparently nonnormal distributions? *Journal of Clinical Epidemiology, 54*(1), 86–92. https://doi.org/10.1016/S0895-4356(00)00264-X

Yuan, P., Li, J., Aruffo, E., Li, Q., Zheng, T., Ogden, N., Sander, B., Heffernan, J., Gatov, E., Gournis, E., Collier, S., Tan, Y., Li, J., Arino, J., Bélair, J., Watmough, J., Kong, J. D., Moyles, I., & Zhu, H. (2020). Efficacy of “stay-at-home” policy and transmission of COVID-19 in Toronto, Canada: A mathematical modeling study. *MedRxiv*, 2020.10.19.20181057. https://doi.org/10.1101/2020.10.19.20181057

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis for the social, behavioral, and biomedical sciences. Behavior Research Methods, 39, 175-191.

Coombs WT, Algina J, Oltman D. (1996). Univariate and multivariate omnibus hypothesis tests selected to control type I error rates when population variances are not necessarily equal. Rev Educ Res 66:137–79.