Survey Report

Luyao Liu

2020-10-04

Summary

COVID-19 has had a huge impact on all aspects of our lives. In this survey, we are going to find out the opinions of restarting Ontario for different groups of people in Ontario. The population is all residents in Ontario and we take 1000 samples. We are using stratified sampling method accoring to the percentages of each age group since homogeneous within each age group and heterogeneous between different age groups. For people who are under 60 years old, we will email the questionnaire to each selected person. To encourage respondents to answer the questionnaire, we are willing to mail 10 masks to them once they submitted the questionnaire. In addition, if the person refuse or ignore to finish the questionnaire, we will find out one more respondent from his/her age group untill we receive enough answers. As for those elder people who may not familiar with email, we will print the hard- copy of questionnaire and mail to them with 10 masks. Respondents do not need to fill in any identifiable personal information, and we will also promise to protect the privacy of respondents in the questionnaire.

It is optimistic that we need to contact 1,300 respondents in total. For each respondent, we need to pay for masks and postage, which may cost about 7,000 to 8,000 dollars in total.

1. Introduction

Our company will conduct a sample survey of the opinions of Ontario residents on the restart plan. Statified random sampling will be used as the sampling method and our questionnaire will includes 7 multiple choice questions concerning the gender, age, income, occupation, education, and the specific opinion of restarting. We will make analysis in terms of the relationship among the collected data and give some useful conlusions.

2. Survey Methodology

2.1 Survey Methods Theories

(1) Simple Random Sampling Without Replacement

In simple random sampling without replacement(SRSWOR), each sample unit of the population has only one chance to be selected in the sample. For example, if one draws a simple random sample such that no unit occurs more than one time in the sample, the sample is drawn without replacement.

Suppose we randomly select n distinct units from the population of size N, then

Probability of sample selection $= \dfrac{1}{\left(\begin{array}{c} N \\ n \end{array} \right)}$, where $\left(\begin{array}{c} N \\ n \end{array} \right) = \frac{N!}{(N-n)!}$

Sample Mean $= \bar{y} = \frac{1}{n} \sum_{i \in S} y_i$, which means sum all of the values of sample and then devided by the sample size.

(2) Stratified Random Sampling

Stratified random sampling(SRS) is a method of sampling that involves the division of a population into smaller sub-groups known as strata. In stratified random sampling, or stratification, the strata are formed based on members' shared attributes or characteristics such as income or educational attainment.

Assume there are H (non-overlapping) strata, h will represent a specific stratum, $h = 1, 2, ..., H$

N_h is the population size of the hth stratum.

$$N = \sum_{h=1}^{H} N_h$$

n_h is the sample size of the hth stratum.

$$n = \sum_{h=1}^{H} n_h$$

y_{hi} is value of the study variable y for unit i in stratum h.

$$i = 1, 2, \cdots, N_h; \quad h = 1, 2, \cdots, H$$

Population Mean for stratum h :

$$\mu_{yh} = \frac{1}{N_h} \sum_{i=1}^{N_h} y_{hi}$$

which means sum the sample values in one stratum and then divided by its size.

Overall Population Mean:

$$\mu_y = \frac{1}{N} \sum_{h=1}^{H} \sum_{i=1}^{N_h} y_{hi}$$

which means sum the sample values in every stratum and then divided by the overall population size.

2.2 Survey Sampling Methods

From Stats Canada, we can find out the percentages of each age group in Ontario. The percentages for age groups (0-19, 20-29, 30-39, 40-49, 50-59, 60-69, above 69) are (0.125, 0.135, 0.165, 0.150, 0.140, 0.145, 0.140). We then use Stratified Sampling by these weights to sample the data.

```
## restart
##                     No     Yes and everywhere Yes but only some parts
##                    298                    305                     397
```
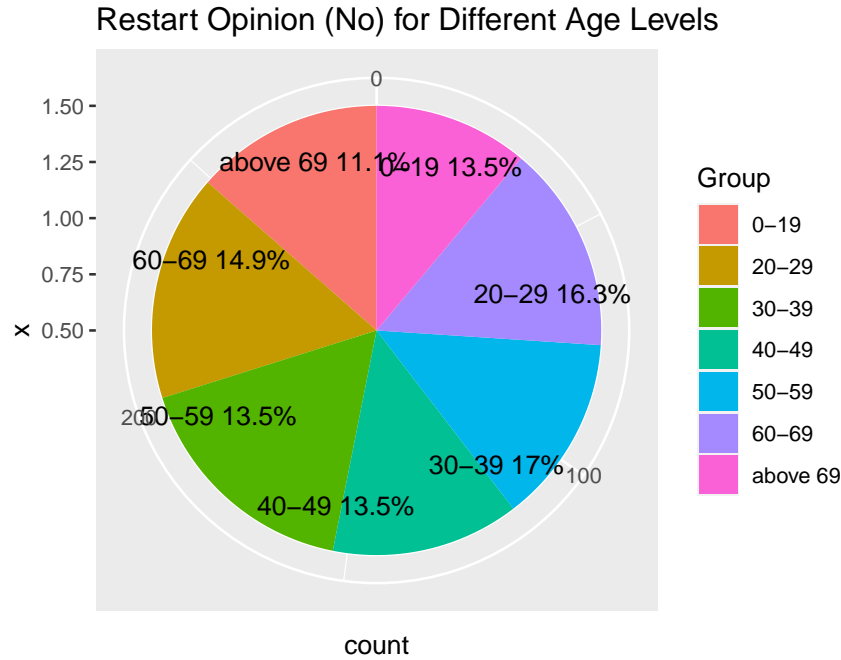
3. Results

This survey collects data from 1000 valid respondents. Since the survey form cannot be submitted unless the respondent selects the answers to all the questions, we avoid missing values.

(1) The frequence tables for Gender with regards to "Restart" are shown as below. We can find most of the respondents agree to restart Ontario but only for some essential parts. Generally speaking, most of the people think Ontario should reopen.

Gender	No	Yes and everywhere	Yes but only some parts
Female	143	152	195
Male	142	155	193
Other	3	1	94
Refuse to answer	0	4	6

(2) Since we split samples' age to 7 groups, pie chart can have a better visualazition patter. The Pie Charts with percentages for Age groups under different opinions to Restart are shown below. From the plots, we can find the respondents who aged between 30 to 39 years old say "No" to restart mostly frenquent. In general, people who are 40 to 59 years old prefer to restart Ontario. This may because they have heavier financial burden so that they need to go back to their work position to make money for family.
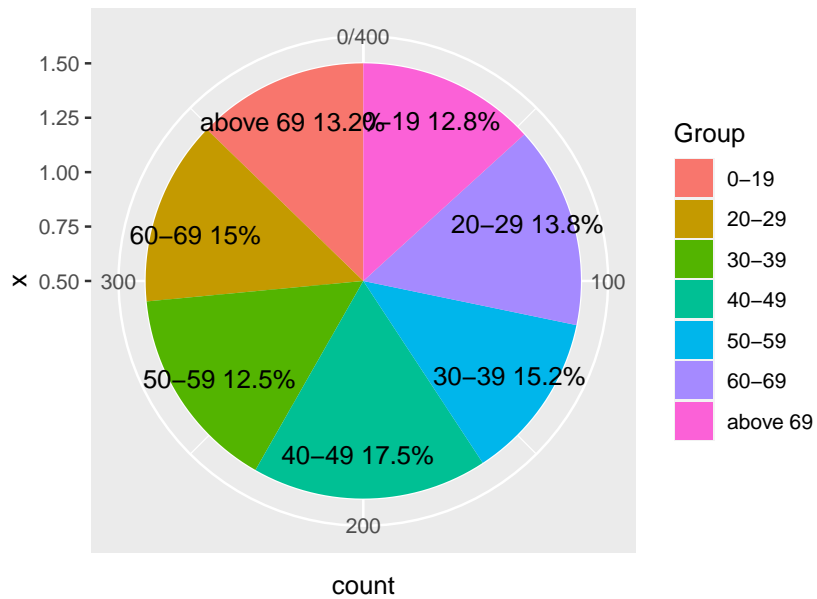
- Restart Opinion (No) for Different Age Levels

Restart Opinion (No) for Different Age Levels

```
## List of 7
##  $ line             : list()
##   ..- attr(*, "class")= chr [1:2] "element_blank" "element"
##  $ rect             : list()
##   ..- attr(*, "class")= chr [1:2] "element_blank" "element"
##  $ text             : list()
##   ..- attr(*, "class")= chr [1:2] "element_blank" "element"
##  $ axis.ticks.length: 'simpleUnit' num 0cm
##   ..- attr(*, "unit")= int 1
##  $ legend.position  : chr "none"
##  $ panel.spacing    : 'simpleUnit' num 0lines
##   ..- attr(*, "unit")= int 3
##  $ plot.margin      : 'simpleUnit' num [1:4] 0lines 0lines 0lines 0lines
##   ..- attr(*, "unit")= int 3
##  - attr(*, "class")= chr [1:2] "theme" "gg"
##  - attr(*, "complete")= logi FALSE
##  - attr(*, "validate")= logi TRUE
```

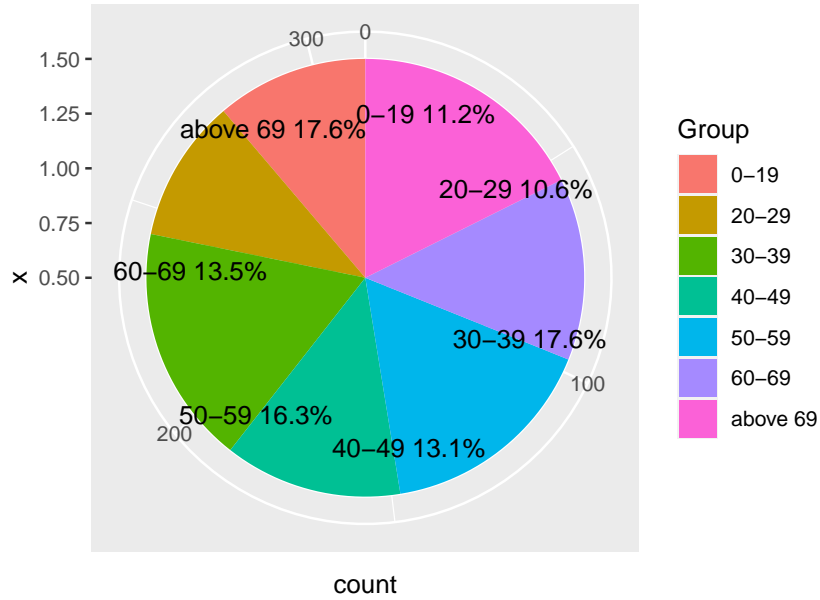- Restart Opinion (Yes But Parts) for Different Age Levels

Restart Opinion (Yes but only some parts) for Different A


```
## List of 7
##  $ line             : list()
##   ..- attr(*, "class")= chr [1:2] "element_blank" "element"
##  $ rect             : list()
##   ..- attr(*, "class")= chr [1:2] "element_blank" "element"
##  $ text             : list()
##   ..- attr(*, "class")= chr [1:2] "element_blank" "element"
##  $ axis.ticks.length: 'simpleUnit' num 0cm
##   ..- attr(*, "unit")= int 1
##  $ legend.position  : chr "none"
##  $ panel.spacing    : 'simpleUnit' num 0lines
##   ..- attr(*, "unit")= int 3
##  $ plot.margin      : 'simpleUnit' num [1:4] 0lines 0lines 0lines 0lines
##   ..- attr(*, "unit")= int 3
##  - attr(*, "class")= chr [1:2] "theme" "gg"
##  - attr(*, "complete")= logi FALSE
##  - attr(*, "validate")= logi TRUE
```

- Restart Opinion (Yes and everywhere) for Different Age Levels

Restart Opinion (Yes and everywhere) for Different Age



```
## List of 7
##  $ line              : list()
##   ..- attr(*, "class")= chr [1:2] "element_blank" "element"
##  $ rect              : list()
##   ..- attr(*, "class")= chr [1:2] "element_blank" "element"
##  $ text              : list()
##   ..- attr(*, "class")= chr [1:2] "element_blank" "element"
##  $ axis.ticks.length: 'simpleUnit' num 0cm
##   ..- attr(*, "unit")= int 1
##  $ legend.position  : chr "none"
##  $ panel.spacing    : 'simpleUnit' num 0lines
##   ..- attr(*, "unit")= int 3
##  $ plot.margin      : 'simpleUnit' num [1:4] 0lines 0lines 0lines 0lines
##   ..- attr(*, "unit")= int 3
##  - attr(*, "class")= chr [1:2] "theme" "gg"
##  - attr(*, "complete")= logi FALSE
##  - attr(*, "validate")= logi TRUE
```

(3) As for Income, Education and Occupation, the graphs with different levels in terms of the opinion of restarting are shown as following. From these graphs, we can find:

- Most of the respondents have income between $50,000 to $74,999, which is also includes the mean and median of the income for Ontario.

- Almost all respondents with high income prefer to restart Ontario since they can make more profits when the market is open. However, the people with low income are not willing to reopen due to the CERB from Canada's govenrment.

- People with lower education level are more willing to restart the Ontario since they have to work physically to make money in most cases while high education peopleu usually have enough ability to make money even if working from home.

- Students prefer to restarting since they need education and social events on campus and some self-employment also agree to restart so that they can make money.
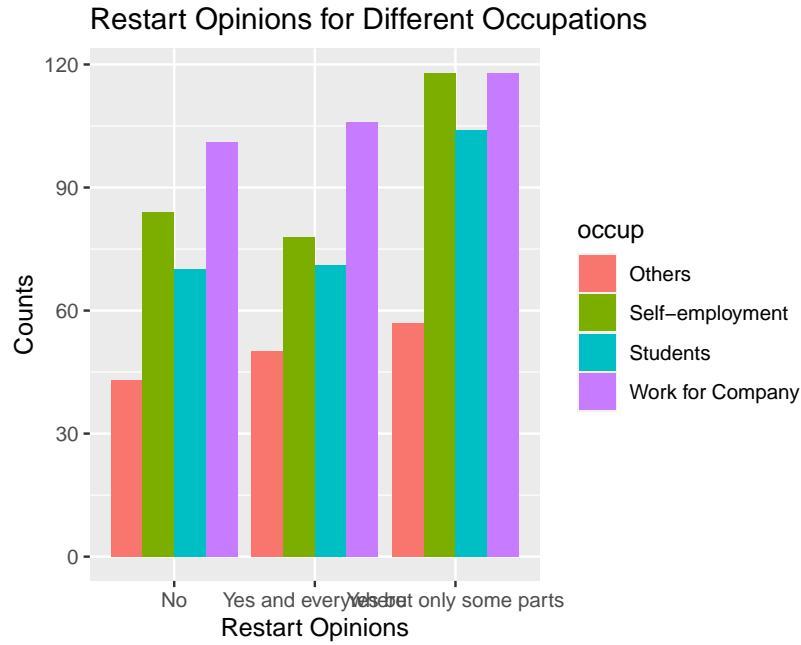
  – Restart Opinions for Different Income Levels
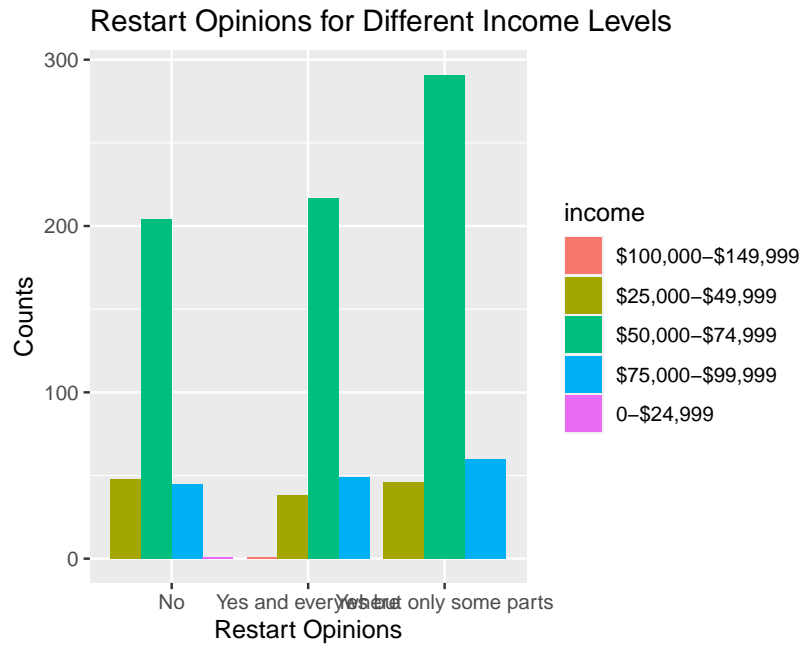
Restart Opinions for Different Income Levels

– Restart Opinions for Different Education Levels

Restart Opinions for Different Education Levels

– Restart Opinions for Different Occupations
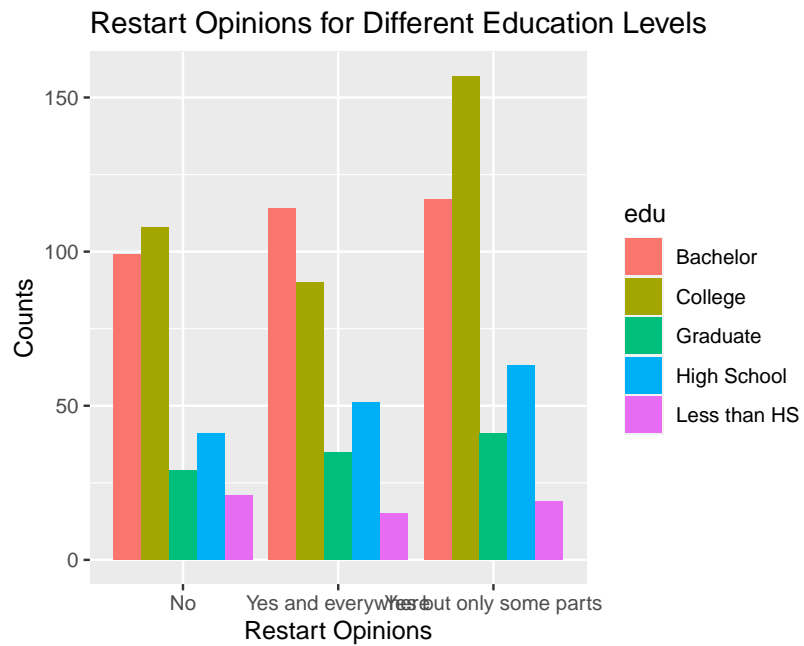
Restart Opinions for Different Occupations

4. Discussion

Weakness and Future Work

1. In order to avoid missing values, the respondents have to finish all questions before summitting. This will reduce the enthusiasm of the respondents and increase the trouble of repeated sampling.

2. We make analysis among the collected data and try to find out the opinions among different groups of people. Since all of our variables are categorical, so it is not proper to do estimation. In the future, we can expend the questionnaire and add some quantitative question to make estimation.

References

- https://www.statista.com/statistics/446025/births-in-canada-by-gender/
- https://www150.statcan.gc.ca/t1/tbl1/en/tv.action?pid=1710000501
- https://www150.statcan.gc.ca/n1/daily-quotidien/200224/dq200224a-eng.html
- https://www12.statcan.gc.ca/census-recensement/2016/dp-pd/hlt-fst/edu-sco/index-eng.cfm
- https://www12.statcan.gc.ca/census-recensement/2016/dp-pd/prof/details/page.cfm?Lang=E&Geo1=PR&Code1=35&Geo2=PR&Code2=35&Data=Count&SearchText=&SearchType=Begins&SearchPR=01&B1=Labour