

Demographic,Social and Behavioural Factors Negatively Affecting the Mental Health of Canadians during COVID-19*

health, covid, mental, age, social, condition

Jennifer Liu

27 April 2022

Abstract

Using the national survey about Impacts of COVID-19 on Canadians- Mental Health, by Statistics Canada, this paper aims to find the demographic,social and behavioural factors that negatively impact the mental health of Canadians during COVID-19. It is found that women, younger people, those who are concerned about their own health, concerned about how to coop after COVID and concerned about maintainig social ties are correlated with a bad mental health. Understanding this relationship helps government to make changes that helps mental health problems.

*Code and data are available at <https://github.com/jenniferliu888/COVID-Mental-Health>

Introduction

Research from the World Health Organization shows that the current COVID-19 pandemic has caused a 25% increase in anxiety and depression worldwide and that multiple stress factors are contributing to it. Moreover, it has mentioned that young people and women are the worst hit by this mental burden (Organization 2022). Therefore, it becomes crucial to examine the factors that are negatively impacting Canadians' mental health during the pandemic. This will help the government to tackle these factors in order to reduce mental health problem as well as providing useful resources to the most vulnerable ones.

Using the 2020 dataset on the impacts of COVID-19 on Canadians' mental health, collected by Statistics Canada, this paper aims to identify the demographics, social and behavioral factors affecting Canadians' mental health. Raw data used can be access through the ODESI portal. (Canada 2020) Although past research has shown that Canadians are more and more anxious because of the pandemic, specific reasons that are causing it remains uncertain (CMHA 2021), which this paper aims to fulfill this gap.

To analyze the survey, firstly, the original dataset from Statistics Canada was categorized, cleaned and renamed. Non-stated response were also removed from it. Secondly, the methodology and collection process of the data source are evaluated. Thirdly, a linear regression and exploratory data analysis (EDA), are used for further exploration in an attempt to find a relationship. Lastly, the result and any possible limitations are discussed.

It is found that women, younger age people, those who are concerned about their own health, concerned about how to coop after COVID and concerned about maintaining social ties are correlated with a bad mental health. Moreover, concern about own health is the factor that correlates with a highest change in perceived mental health. However, the crowdsourcing nature of the survey data might affect the result and makes it unrepresentative of the Canadian population.

Data

This report is analyzed using R (R Core Team 2020), using tidyverse (Wickham et al. 2019) and dplyr (Wickham et al. 2020) packages. All the tables and graphs are created using ggplot2 (Wickham 2016) and the file is knitted using knitr (Franbois, Henry, and Miller 2021) and patchwork (Pedersen 2020).

Data Source

This report is analyzed using a survey conducted by Statistics Canada, in 2020, called "Crowdsourcing: Impacts of COVID-19 on Canadians-Your Mental Health". It was extracted from the ODESI, which is a data portal for researchers, teachers and students that contains raw data file of multiple Canadians survey of public opinion on social, economic and political issues. (Canada 2020)

From April 3, 2020 Statistics Canada has conducted a series of survey in order to assess Canadians' current economic and social situation. during the COVID-19. The survey is used by organizations such as the Public Health Agency of Canada and Employment and Social Development Canada to find people's need and develop support to fulfill these needs. Some subjects that are covered in these surveys include disability, economic accounts, health, income and mental health. The 2020 survey that is used to analyze this report is emphasized on Canadians' mental health.

Collection

All Canadians above 15 years old from the ten provinces and three territories were able to participate in this survey. Advertisements on social media. news channel and partnership with organization is used to recruit the volunteers. It is a crowdsourcing initiative where self-selected volunteers fill the online survey.

Therefore, no sampling was done and no response rate was found. A total of 45,989 has responded to the survey in between April 24 to May 11, 2020.

Data Characteristics and Cleaning

The dataset has been filtered to only keep the relevant 13 variables. Non-stated response are removed from the dataset. Further, all variables column name has been renamed so that it is easier to read and align with the actual data.

Some of the variables are: mental_before: Mental health compared to before physical distancing feeling: Frequency over the last 2 weeks of feeling nervous, anxious or on edge feeling_afraid: Frequency over the last 2 weeks of feeling afraid as if something awful might happen shopping_grocery: Frequency last week of going shopping at the grocery store or drugstore deliveries: Frequency last week of using delivery service for groceries or drugstore concern_socialities: Concern about impact of COVID-19 - Maintaining social ties concern_loosing: Scale that might lose main job or main self-employment income in the next 4 weeks

Table 1: Summary of Respondents' Mental Health

min	Q1	median	Q3	max	mean	standarddeviation
1	2	3	4	5	2.81398	1.050657

Respondents perceived mental health ranges from Excellent (1) to Poor (5) with a median at Good (3) and a mean of 2.81. Its standard deviation of 1.05 is rather volatile and would be worth examining the factors contributing to this.

Result

This section looks into factors that affect Canadians' mental health by different variables, such as shopping grocery frequency, different type of concern and age group. This section will draw conclusions between variables by first analyzing with graphs and then linear regression model.

Figure 1: Histogram of Respondents' Mental Health

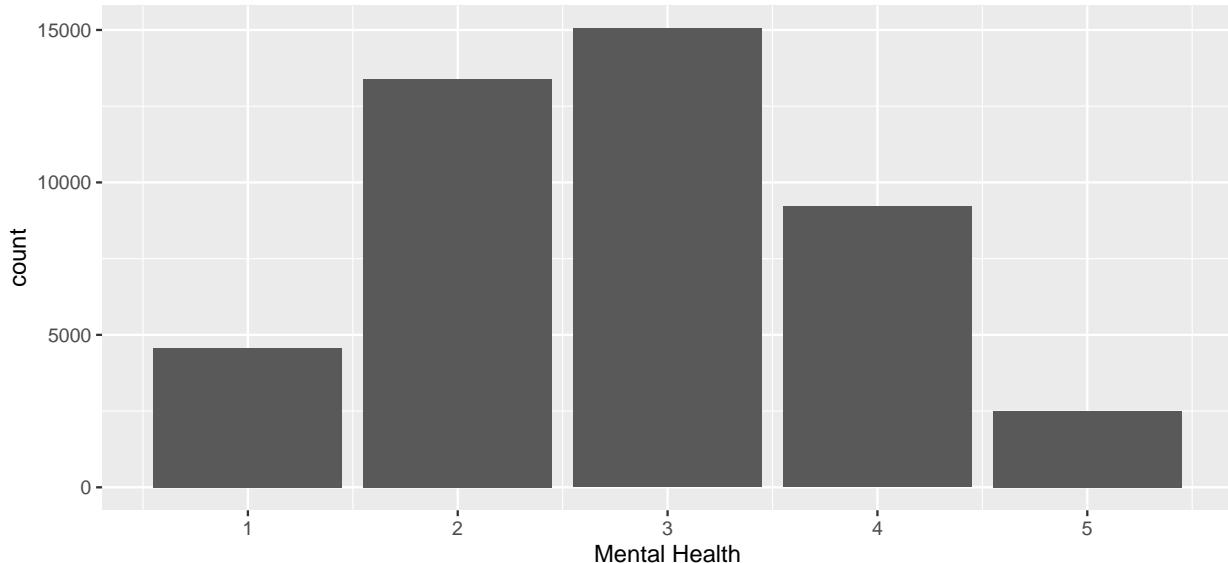


Figure 2: Histogram of Social Ties Concern

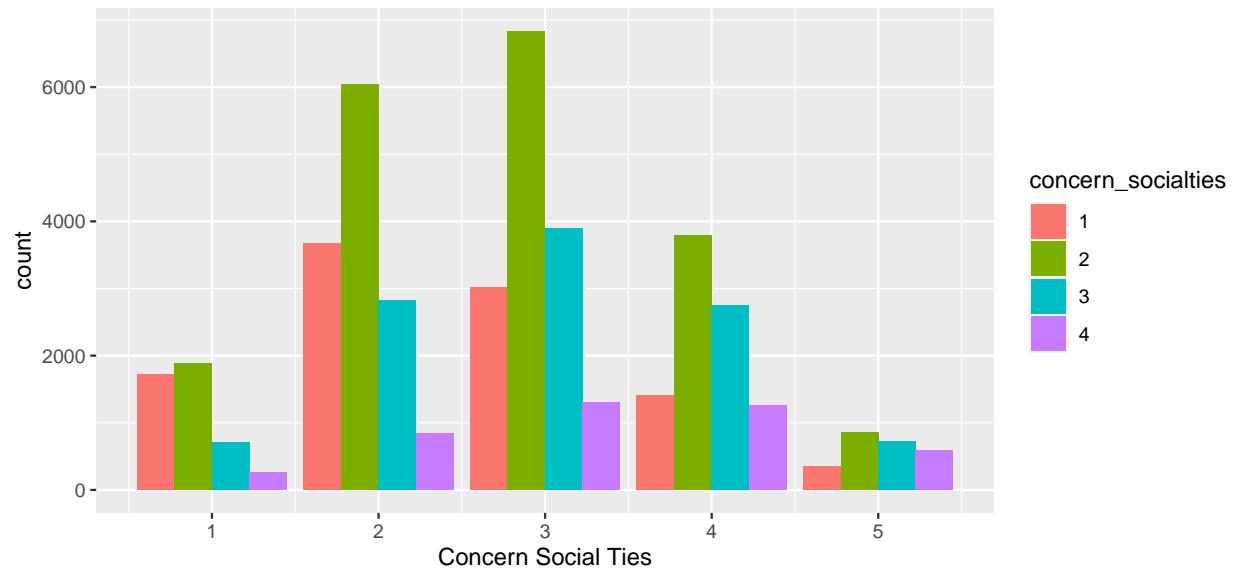


Figure 3: Histogram of Respondents' Gender

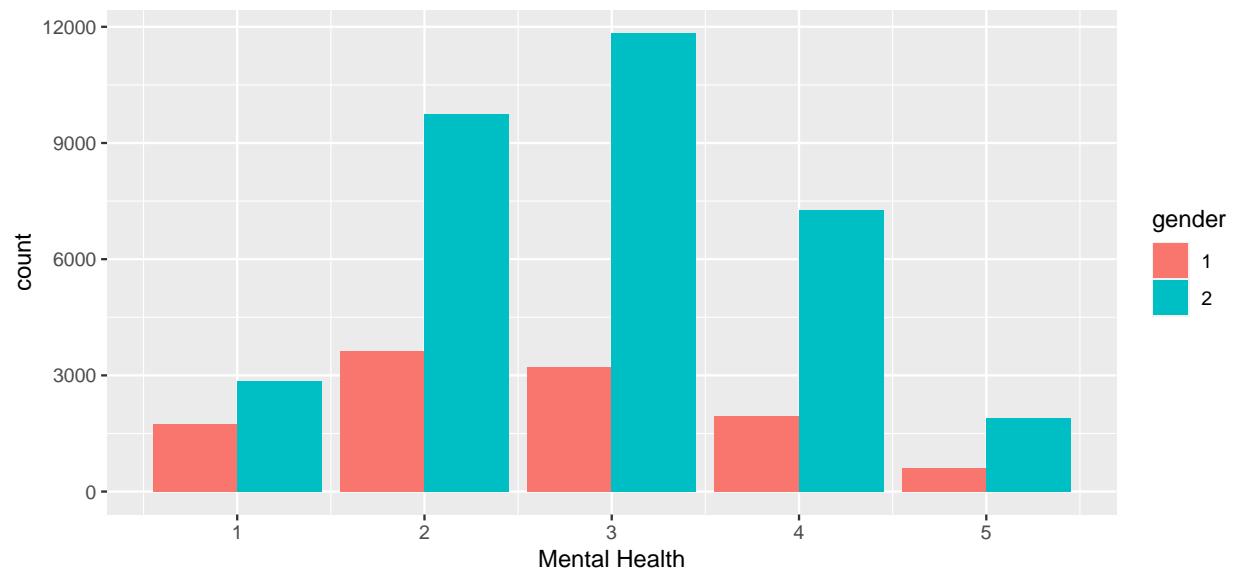


Figure 4: Histogram of Shopping Frequency

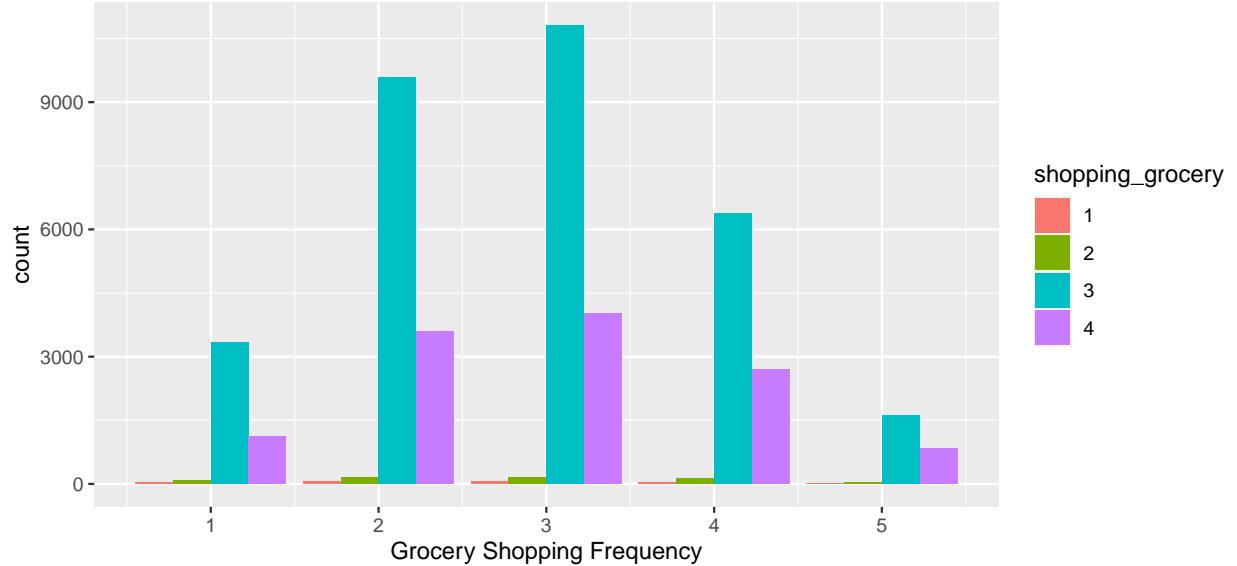


Figure 5: Histogram of Health Concern

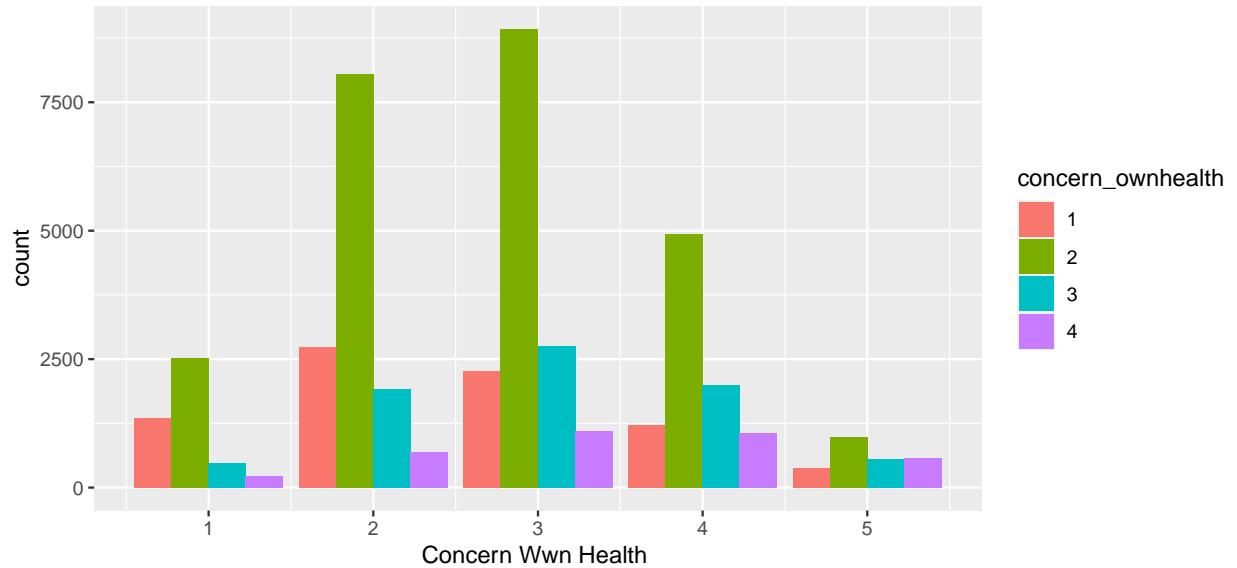


Figure 2 to 5 shows respondents' perceived mental health normal distribution, with 1 being excellent mental health, 2 very good, 3 good, 4 fair and 5 poor, with the higher number of people having good mental health. Each of the histograms is filled with another variable in order to make effective comparison and find possible relationship between both that are going to be used in the linear regression model. Figure 2 is a histogram with respondents' concern about social ties as the variable, 1 being not at all and 4 extremely. It is very interesting to notice that although more people have a good mental health than those with very good, the number of people who are not at all concerned about social ties have decreased. The proportion of people who are very concerned about maintaining social ties is relatively higher with those who have poor mental health.

Figure 3 is a histogram with respondents' gender with 1 being a male and 2 a female. Male respondents with good mental health should be more than those with very good mental health if the proportions remains equal. However, the proportion of male has decreased as the perceived mental health changes from 2 to 3.

Figure 5 shows a histogram of respondents' mental health with their concern about their own health, 1 being not at all and 4 extremely. We can observe that with a worsens in perceived mental health, those who are

not at all concern with their own health has decreased (when perceived mental health goes from 2 to 3). Respondents who have a fair mental health have a higher proportion of those who are extremely concerned about their own heath.

Figure 6: Concern Able to Coop

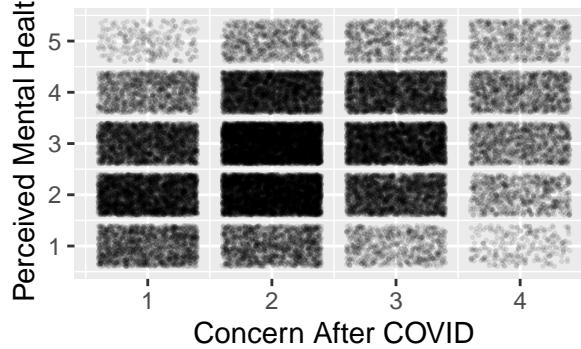


Figure 7: Concern Loosing Job



Figure 8: Concern Family Violence

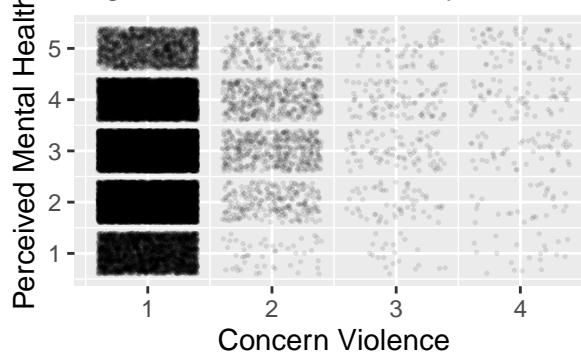


Figure 9: Age Group

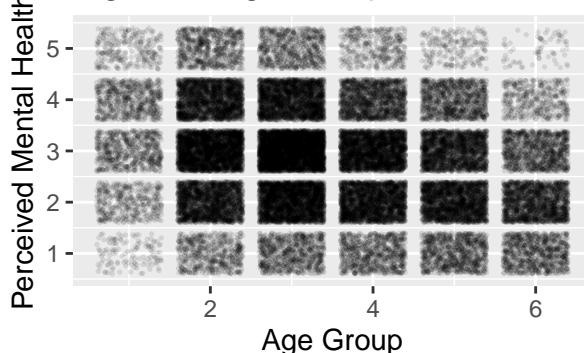
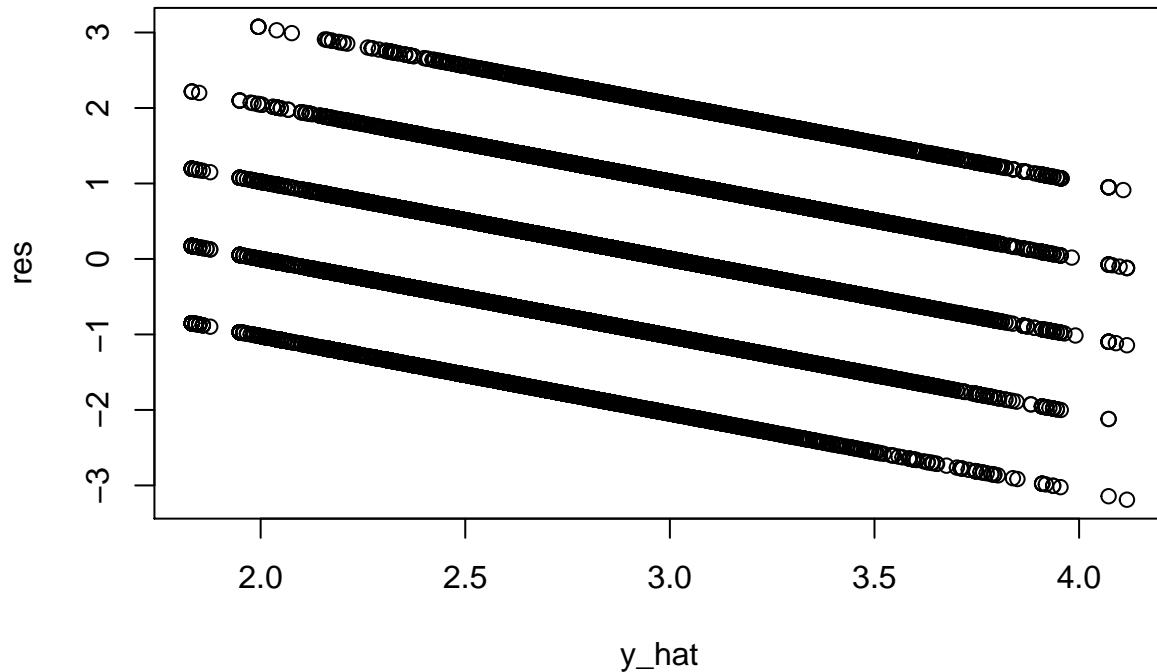


Figure 6-9 are scatterplots of 3 other variables and the perceived mental health where 1 being the least worries, and 4 the most. Because all variaables are discrete, in order to avoid overplotting, a random variation to the location of each point is added. Figure 6 shows a strong positive relationship between concern to be able to live well after COVID and the perceived mental health. Figure 7 shows a weaker positive correlation between concern of loosing job and perceived mental health in figure 7. The more people are concerned with loosing their employment in the next 4 weeks is correlated with weaker mental health. Moreover, those who are extremely concerned with loosing their job has a wide range of perceived mental health spray. In figure 8, no correlation has been found in the scatterplot as no relation is found between concern for family violence and mental health. Figure 9 shows a strong negative correlation between the variables; as age group increases, respondents' perceived mental health gets better

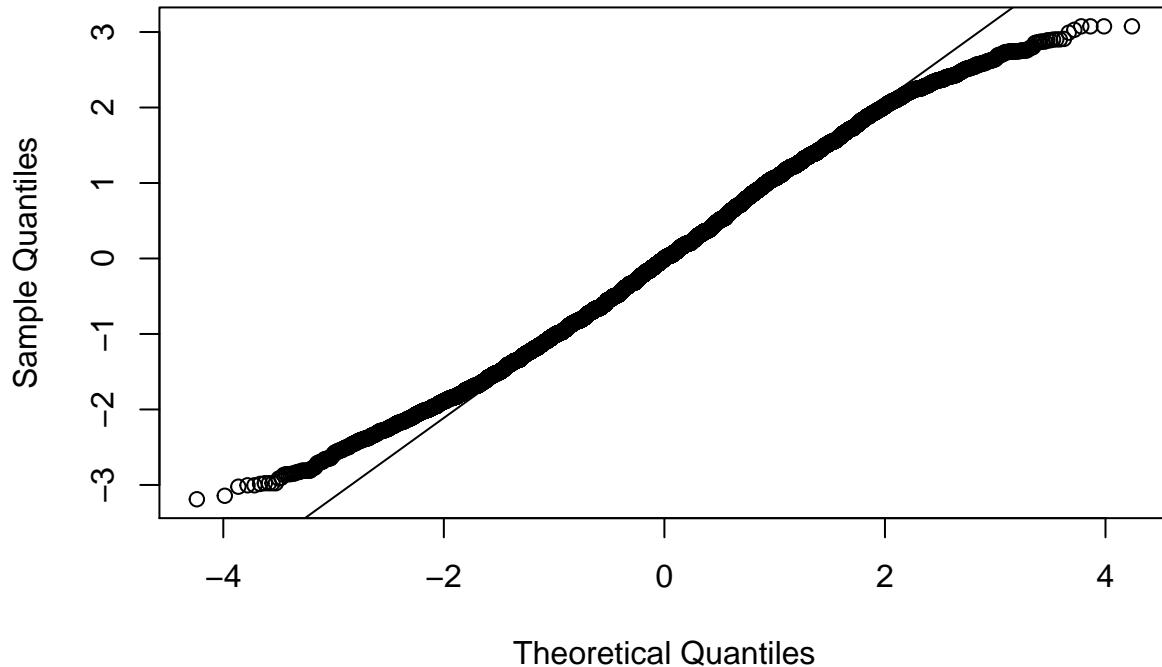
After analyzing graph 1-9, the variables of concern about social ties, concern about their own health, concern to be able to coop after covid, concern of loosing employment and age group are potential factors that are affecting this.

Model



Although this residual plot does not satisfy the assumption of linearity and independence, the variables has been treated as continuous, while they are discrete. The consequence is that the residual plot appears as a series of parallel lines, as shown in the plot. This makes it more difficult since the points clearly will not be randomly scattered. Therefore, we will also perform a residual QQ plot.

Normal Q-Q Plot



There is room of improvement because the dots are slightly deviated from the 45 degree line, but we will assume that normality is not violated in the model because of the large dataset we have.

	Estimate	Std.Error	t-value	Pr(>t)
Intercept	2.383856	0.025847	92.229	< 2e-16
Concern Social Ties	0.116918	0.005847	19.996	< 2e-16
Concern Own Health	0.208381	0.005925	35.168	< 2e-16
Concern Coop After	0.150811	0.005938	25.398	< 2e-16
Concern Loosing Job	-0.008973	0.003289	-2.729	0.00636
Age Group	-0.162380	0.003416	-47.536	< 2e-16

We then constructed a linear regression model for variables selected previously. A mathematical model is useful to explain how each variable affects the mental health. In this model, β_1 represents respondents concern about social ties; β_2 indicates their concern with their own health; β_3 is for their concern to coop after COVID; β_4 indicated concern of loosing employment and β_4 for the age group. The null hypothesis is that the variable has no correlation with mental health. All p-value are higher than the significance level, which means that the sample data does not provide enough evident to reject the null hypothesis. Thus, all variables are affecting respondents' mental health. The residual standard error is 0.9775, which indicated that 97.75% variation in mental health is explained by the model.

$$Y = \beta_0 + 0.116918X_1 + 0.208381X_2 + 0.150811X_3 - 0.008973X_4 - 0.162380X_5$$

- Y : Perceived mental health
- X_1 : Concern about social ties
- X_2 : Concern about own health

- X_3 : Concern how to coop after COVID
- X_4 : Concern loosing job
- X_5 : Age group

From the coefficients, we can find that concern of social ties, concern of their own health and concern to coop after COVID all have positive correlation with perceived mental health. The more they are concerned about these, their mental health becomes weaker. Concern of loosing the job and age group have negative correlation with the interest variable. Those who are younger have a poorer perceived mental health than the older ones and the less concern a person is about loosing its job, the poorer its mental health is. All correlations are aligned with the previous graphs, except for concern of loosing employment, which had a weak positive relationship.

Respondents' concern about their own health has the highest coefficient with concern of loosing job the smallest coefficient.

Discussion

In using data from the Statistics Canada's "Crowdsourcing: Impacts of COVID-19 on Canadians-Your Mental Health", this paper examines the impacts of multiple factors on people's mental health during the pandemic. First, by analyzing graphs, it is found that respondents' concern about their own health and concern about social ties are interesting variables to be analyzed as they have a similar pattern than with an increased in perceived mental health. Moreover, concern to coop after COVID shows a strong positive correlation, concern of loosing job shows a weak positive correlation and age group a strong negative correlation with the perceived mental health.

After conducting a linear regression model, it is found that concern of social ties, concern of their own health and concern to coop after COVID all have positive strong correlation with perceived mental health. Concern of loosing the job have a weak and age group have a strong negative correlation. This finding also makes logical sense. As people becoming more concern with maintaining social ties, more concern with their own health and more concern with how to coop after COVID, their perceived mental health gets worse off. Moreover, as people become younger, their mental health is also getting worse. Concern about their own health is the most impactful variables, when people become more concern for their own health by a rate of 0.21, their mental health increased by 1, moving from good to fair for example.

A discrepancy between the exploratory graphing data analysis and the linear regression model is the variable of rate of concern for loosing a job. It is a weak positive relationship with the scatterplot and a weak negative correlation with the model as it has a negative coefficient. It is reasonable to further analyze people's concern of loosing their job and not including it in the final model because its p-value is the lowest among all variables and its correlation is also the weakest. A change of 1 in mental health, from good to fair, is only correlated with a decreased of -0.008973 of rate of concern for loosing the job. A possible explanation for this discrepancy is that because the graph has been overplotted and that those who are extremely concerned with loosing their job has a wide range of perceived mental health spread, it is difficult to analyze the scatterplot.

Mental health is extremely important for Canadians well-being and government can implement concrete changes based on these results to better ease our mental health. For instance, reducing concern about people's own health by creating meaningful marketing campaign will help. Planning out a concrete plans for a smooth recover also reduces Canadians' concern on how to coop after this pandemic.

From figure 3, we can also see that the proportion of women with more difficult mental health situation is more than men. Further analysis can be done in finding what factors impact Canadians women to have a worse mental health situation during the pandemic than men. Women suffering more than men is an important problem that ought to analyze. A news from CBC on March 11, 2022 also aligned with this finding. It has reported that 60% of women, from 18 to 34 years old affirms that their mental health had worsened throughout the pandemic, and this number is higher than men's (Miller 2022).

Limitation

Multiples issues with this survey should be considered when looking at this paper. To begin, no sampling was done and this survey was only administered on a voluntary basis. The sample of 45,989 respondents might not be representative of the Canadian population. Perhaps, those who exhibit mental health problems has a higher tendency to complete the survey. Moreover, because a major part of the survey promotion has been done through social media, younger people would have a higher response rate than older ones.

Moreover, respondents are asked to rate their perceived mental health on a scale of 1-5. This is highly biased as someone who considers itself to have a good mental health might be different to someone else and it becomes difficult to make comparison. This is similar for questions where respondents are asked to measure their concern about a certain problem.

Appendix

Motivation

1. *For what purpose was the dataset created? Was there a specific task in mind? Was there a specific gap that needed to be filled? Please provide a description.*
 - The dataset was created in order to assess Canadians mental health changes during COVID. No survey has been done before regarding mental health for Canadians.
2. *Who created the dataset (for example, which team, research group) and on behalf of which entity (for example, company, institution, organization)?*
 - Statistics Canada
3. *Who funded the creation of the dataset? If there is an associated grant, please provide the name of the grantor and the grant name and number.*
 - Statistics Canada, government.
4. *Any other comments?*

Composition

1. *What do the instances that comprise the dataset represent (for example, documents, photos, people, countries)? Are there multiple types of instances (for example, movies, users, and ratings; people and interactions between them; nodes and edges)? Please provide a description. The data is in the form of survey with only numbers for the data.*
 - 2. *How many instances are there in total (of each type, if appropriate)? NA*
 - 3. *Does the dataset contain all possible instances or is it a sample (not necessarily random) of instances from a larger set? If the dataset is a sample, then what is the larger set? Is the sample representative of the larger set (for example, geographic coverage)? If so, please describe how this representativeness was validated/verified. If it is not representative of the larger set, please describe why not (for example, to cover a more diverse range of instances, because instances were withheld or unavailable).*
 - NA
4. *What data does each instance consist of? "Raw" data (for example, unprocessed text or images) or features? In either case, please provide a description.*
 - only raw data
5. *Is there a label or target associated with each instance? If so, please provide a description.*
 - na
6. *Is any information missing from individual instances? If so, please provide a description, explaining why this information is missing (for example, because it was unavailable). This does not include intentionally removed information, but might include, for example, redacted text.*
 - na
7. *Are relationships between individual instances made explicit (for example, users' movie ratings, social network links)? If so, please describe how these relationships are made explicit.*
 - na

8. Are there recommended data splits (for example, training, development/validation, testing)? If so, please provide a description of these splits, explaining the rationale behind them.
 - no
9. Are there any errors, sources of noise, or redundancies in the dataset? If so, please provide a description.
 - there are the options of refusal or nonstated in the data
10. Is the dataset self-contained, or does it link to or otherwise rely on external resources (for example, websites, tweets, other datasets)? If it links to or relies on external resources, a) are there guarantees that they will exist, and remain constant, over time; b) are there official archival versions of the complete dataset (that is, including the external resources as they existed at the time the dataset was created); c) are there any restrictions (for example, licenses, fees) associated with any of the external resources that might apply to a dataset consumer? Please provide descriptions of all external resources and any restrictions associated with them, as well as links or other access points, as appropriate.
 - no
11. Does the dataset contain data that might be considered confidential (for example, data that is protected by legal privilege or by doctor-patient confidentiality, data that includes the content of individuals' non-public communications)? If so, please provide a description.
 - all confidential information were not asked.
12. Does the dataset contain data that, if viewed directly, might be offensive, insulting, threatening, or might otherwise cause anxiety? If so, please describe why.
 - no
13. Does the dataset identify any sub-populations (for example, by age, gender)? If so, please describe how these subpopulations are identified and provide a description of their respective distributions within the dataset.
 - the dataset identifies gender and age. Respondents were required to enter this information
14. Is it possible to identify individuals (that is, one or more natural persons), either directly or indirectly (that is, in combination with other data) from the dataset? If so, please describe how.
 - no
15. Does the dataset contain data that might be considered sensitive in any way (for example, data that reveals race or ethnic origins, sexual orientations, religious beliefs, political opinions or union memberships, or locations; financial or health data; biometric or genetic data; forms of government identification, such as social security numbers; criminal history)? If so, please provide a description.
 - yes, it asks for concern about loosing employment.
16. Any other comments?
 - no

Collection process

1. How was the data associated with each instance acquired? Was the data directly observable (for example, raw text, movie ratings), reported by subjects (for example, survey responses), or indirectly inferred/derived from other data (for example, part-of-speech tags, model-based guesses for age or language)? If the data was reported by subjects or indirectly inferred/derived from other data, was the data validated/verified? If so, please describe how.
 - data was collected using online survey

2. What mechanisms or procedures were used to collect the data (for example, hardware apparatuses or sensors, manual human curation, software programs, software APIs)? How were these mechanisms or procedures validated?
 - manual computer collection
3. If the dataset is a sample from a larger set, what was the sampling strategy (for example, deterministic, probabilistic with specific sampling probabilities)?
 - no sampling was done
4. Who was involved in the data collection process (for example, students, crowdworkers, contractors) and how were they compensated (for example, how much were crowdworkers paid)?
 - NA
5. Over what timeframe was the data collected? Does this timeframe match the creation timeframe of the data associated with the instances (for example, recent crawl of old news articles)? If not, please describe the timeframe in which the data associated with the instances was created.
 - 2020-04-24 to 2020-05-11
6. Were any ethical review processes conducted (for example, by an institutional review board)? If so, please provide a description of these review processes, including the outcomes, as well as a link or other access point to any supporting documentation.
 - no
7. Did you collect the data from the individuals in question directly, or obtain it via third parties or other sources (for example, websites)?
 - third parties, through <http://odesi2.scholarsportal.info/webview/>
8. Were the individuals in question notified about the data collection? If so, please describe (or show with screenshots or other information) how notice was provided, and provide a link or other access point to, or otherwise reproduce, the exact language of the notification itself.
 - NA
9. Did the individuals in question consent to the collection and use of their data? If so, please describe (or show with screenshots or other information) how consent was requested and provided, and provide a link or other access point to, or otherwise reproduce, the exact language to which the individuals consented.
 - they voluntary answered the survey
10. If consent was obtained, were the consenting individuals provided with a mechanism to revoke their consent in the future or for certain uses? If so, please provide a description, as well as a link or other access point to the mechanism (if appropriate).
 - NA
11. Has an analysis of the potential impact of the dataset and its use on data subjects (for example, a data protection impact analysis) been conducted? If so, please provide a description of this analysis, including the outcomes, as well as a link or other access point to any supporting documentation.
 - no
12. Any other comments?
 - no

Preprocessing/cleaning/labeling

1. *Was any preprocessing/cleaning/labeling of the data done (for example, discretization or bucketing, tokenization, part-of-speech tagging, SIFT feature extraction, removal of instances, processing of missing values)? If so, please provide a description. If not, you may skip the remaining questions in this section.*
 - yes, only relevant variables are kept.
2. *Was the “raw” data saved in addition to the preprocessed/cleaned/labeled data (for example, to support unanticipated future uses)? If so, please provide a link or other access point to the “raw” data.*
 - yes, <http://odesi2.scholarsportal.info/webview/>
3. *Is the software that was used to preprocess/clean/label the data available? If so, please provide a link or other access point.*
 - using R
4. *Any other comments?*
 - no

Uses

1. *Has the dataset been used for any tasks already? If so, please provide a description.*
 - Because the dataset is publicly open, it might have been used for other research.
2. *Is there a repository that links to any or all papers or systems that use the dataset? If so, please provide a link or other access point.*
 - na
3. *What (other) tasks could the dataset be used for?*
 - paper research, inform Canadians
4. *Is there anything about the composition of the dataset or the way it was collected and preprocessed/cleaned/labeled that might impact future uses? For example, is there anything that a dataset consumer might need to know to avoid uses that could result in unfair treatment of individuals or groups (for example, stereotyping, quality of service issues) or other risks or harms (for example, legal risks, financial harms)? If so, please provide a description. Is there anything a dataset consumer could do to mitigate these risks or harms?*
 - It was collected in a crowdsourcing way, in which the data might not be representative of the Canadians population.
5. *Are there tasks for which the dataset should not be used? If so, please provide a description.*
 - no
6. *Any other comments?*
 - no

Distribution

1. *Will the dataset be distributed to third parties outside of the entity (for example, company, institution, organization) on behalf of which the dataset was created? If so, please provide a description.*
 - no
2. *How will the dataset be distributed (for example, tarball on website, API, GitHub)? Does the dataset have a digital object identifier (DOI)?*

- no
3. *When will the dataset be distributed?*
- no
4. *Will the dataset be distributed under a copyright or other intellectual property (IP) license, and/or under applicable terms of use (ToU)? If so, please describe this license and/ or ToU, and provide a link or other access point to, or otherwise reproduce, any relevant licensing terms or ToU, as well as any fees associated with these restrictions.*
- need to contact ata Liberation Initiative (Statistics Canada) , <http://www.statcan.gc.ca/eng/dli/dli> , ddi- idd@statcan.gc.ca
5. *Have any third parties imposed IP-based or other restrictions on the data associated with the instances? If so, please describe these restrictions, and provide a link or other access point to, or otherwise reproduce, any relevant licensing terms, as well as any fees associated with these restrictions.*
- no
6. *Do any export controls or other regulatory restrictions apply to the dataset or to individual instances? If so, please describe these restrictions, and provide a link or other access point to, or otherwise reproduce, any supporting documentation.*
- no
7. *Any other comments?*
- no

Maintenance

1. *Who will be supporting/hosting/maintaining the dataset?*
- Statistics Canada
2. *How can the owner/curator/manager of the dataset be contacted (for example, email address)?*
- ata Liberation Initiative (Statistics Canada) , <http://www.statcan.gc.ca/eng/dli/dli> , ddi- idd@statcan.gc.ca
3. *Is there an erratum? If so, please provide a link or other access point.*
- no
4. *Will the dataset be updated (for example, to correct labeling errors, add new instances, delete instances)? If so, please describe how often, by whom, and how updates will be communicated to dataset consumers (for example, mailing list, GitHub)?*
- NA
5. *If the dataset relates to people, are there applicable limits on the retention of the data associated with the instances (for example, were the individuals in question told that their data would be retained for a fixed period of time and then deleted)? If so, please describe these limits and explain how they will be enforced.*
- NA
6. *Will older versions of the dataset continue to be supported/hosted/maintained? If so, please describe how. If not, please describe how its obsolescence will be communicated to dataset consumers.*
- no older version available

7. *If others want to extend/augment/build on/contribute to the dataset, is there a mechanism for them to do so? If so, please provide a description. Will these contributions be validated/verified? If so, please describe how. If not, why not? Is there a process for communicating/distributing these contributions to dataset consumers? If so, please provide a description.*

- contact ata Liberation Initiative (Statistics Canada) , <http://www.statcan.gc.ca/eng/dli/dli> , ddi-idd@statcan.gc.ca

8. *other comments?*

- no

Reference

- Canada, Statistics. 2020. *Dataset: Impacts of Covid-19 on Canadians- Mental Health, 2020: Crowdsource File*. <http://odesi2.scholarsportal.info/webview/velocity?v=2&mode=documentation&submode=abstract&study=http%3A%2F%2F142.150.190.128%3A80%2Fobj%2FfStudy%2FICC-13-25-00022020001-E-2020-Version2>.
- CMHA. 2021. *How Are We Feeling? Canadians Are Worried, Bored, Stressed, Lonely and Sad*. <https://cmha.ca/how-are-we-feeling-canadians-are-worried-bored-stressed-lonely-and-sad/>.
- Franbois, Yihui Xiein, Lionel Henry, and Kirill Miller. 2021. *Knitr: A General-Purpose Package for Dynamic Report Generation in R*.
- Miller, Adam. 2022. *2 Years into the Pandemic, Canada's Mental-Health System Is at a Crisis Point*. <https://www.cbc.ca/news/health/canada-mental-health-crisis-covid-19-pandemic-1.6382378#:~:text=Fifty%2Dseven%20per%20cent%20of,health%20implications%20for%20some%20Canadians>.
- Organization, World Health. 2022. *COVID-19 Pandemic Triggers 25*. <https://www.who.int/news/item/02-03-2022-covid-19-pandemic-triggers-25-increase-in-prevalence-of-anxiety-and-depression-worldwide>.
- Pedersen, Thomas Lin. 2020. *Patchwork: The Composer of Plots*. <https://CRAN.R-project.org/package=patchwork>.
- R Core Team. 2020. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. <https://www.R-project.org/>.
- Wickham, Hadley. 2016. *Ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York. <https://ggplot2.tidyverse.org>.
- Wickham, Hadley, Mara Averick, Jennifer Bryan, Winston Chang, Lucy D'Agostino McGowan, Romain Franbois, Garrett Grolemund, et al. 2019. "Welcome to the tidyverse." *Journal of Open Source Software* 4 (43): 1686. <https://doi.org/10.21105/joss.01686>.
- Wickham, Hadley, Romain Franbois, Lionel Henry, and Kirill Miller. 2020. *Dplyr: A Grammar of Data Manipulation*. <https://CRAN.R-project.org/package=dplyr>.