National Undergraduate Big Data Challenge 2020

Correlation between COVID-19, Mental Illness and Substance Abuse Prevalence

Robert Joseph[1], Hamza Qureshi[2], and Shonnae Frazer[3]

[1]University of Alberta
[2]Brock University
[3]York University

June 2, 2020

Abstract

The aim of this paper is to determine if the experience of the COVID-19 pandemic correlates with worsened mental health and increased substance use among the affected populations. Data sets from the National Addiction & HIV Data Archive Program (NAHDAP) and Statistics Canada (StatCan) were analyzed, with the aim of determining the trends in mental health and alcohol purchases during the COVID-19 pandemic. Using machine learning model techniques, a strong association between the two was discovered. It was also predicted that similar future situations would yield the same results. Statistical tools within Python, R, and SAS were used to explore the data and analyse it using various techniques., including linear regression. As a result, our findings demonstrate that coronavirus pandemics correlate with increased mental illness and substance use within the affected populations. If its effects are similar to SARS, however, considering that COVID-19 has had a much more widespread impact, significant increases in mental illness and substance abuse can be expected in the months and years to come. The analysis determined that there is a strong correlation between the COVID-19 pandemic, mental illness and alcohol and drug use. Limitations to this conclusion involve the fact that the pandemic is still ongoing and the datasets used were relatively new and small.

Keywords
Isolation, Substance Abuse, Mental Health, Pandemic, COVID-19, SARS

1 Introduction

The COVID-19 pandemic has infected over 6 million people, including almost 400,000 deaths, worldwide[1]. The detrimental impact of COVID-19 on the physical health of those infected and the global economy are clear. However, the less tangible effects of the pandemic are also important. As the world begins a slow and cautious return to some semblance of normality, it is necessary to accurately measure the full scope of the damage COVID-19 has caused, both to address these impacts and to prepare for future pandemics. This led to the question: How has the COVID-19 pandemic affected the mental health of the affected populations? Mental distress is tied to alcohol and drug consumption, so these factors were also analyzed.

In the months since the discovery of COVID-19, preliminary studies have demonstrated that the pandemic has worsened mental health[2] and increased alcohol use[3] among Canadians. Canadians who experienced a decline in their mental well-being were more likely to increase their use of alcohol, tobacco and cannabis [4]. The COVID-19 pandemic has also created new barriers to people suffering from mental illness and substance abuse disorders accessing treatment.

Prior to the COVID-19 pandemic, nearly one in five of U.S. adults (47 million) reported having a mental illness in the past year, and over 11 million had a serious mental illness which frequently resulted in functional impairment and limited life activities [2].

The severe acute respiratory syndrome (SARS) pandemic, which was similar to COVID-19 but on a much smaller scale, was found to result in increased risk of mental illness and drug

use among affected health care workers [5]. We analyzed data from SARS to better evaluate the long-term effects of pandemics on the prevalence of mental illness and drug use.

A broad body of research links several side effects of pandemics - for example, social isolation, stress and unemployment - to adverse mental health outcomes. In addition, epidemics have been shown to induce general stress across a population and may lead to new mental health and substance use issues [6].

2 Materials & Methods

A variety of data sets from were used, from open sources including StatCan, NAHDAP, Kaggle and CDC. The data sets used for our analysis were:

- Alcohol Abuse/Dependence Symptoms Among Hospital Employees Exposed to a SARS Outbreak[2]
- KFF Health Tracking Poll - Early April 2020: The Impact Of Coronavirus On Life In America[6]
- Impact of Alcohol or Drug Use and Incarceration on Child Care in Santa Clara County, California, 2003 (ICPSR 4211)[7]
- Canadian Consumers Adapt to COVID-19: A Look at Canadian Grocery Sales up to April 11[4]
- Epidemiological summary of COVID-19 cases in Canada[1]

The datasets retrieved from the Centers for Disease Control and Prevention (CDC) and Inter-university Consortium for Political and Social Research (ICPSR) provided reported overdose deaths in the United States.

This data was modelled and fit through either a linear or quadratic regression. The data preparation was completed using SAS and the visual plots were completed using Python in Jupyter lab.

While trying to correlate the substance abuse/mental illness with COVID-19, a variety of factors were considered, including previous mental health disorder, age, country, previous substance abuse disorders, and family related issues. All analysis was completed using a 95% confidence interval.

Matplotlib, numpy and seaborn were the libraries used in Python.

Visual heat maps were used to show the strong correlation between the two groups and dynamic metrics were used to justify results.

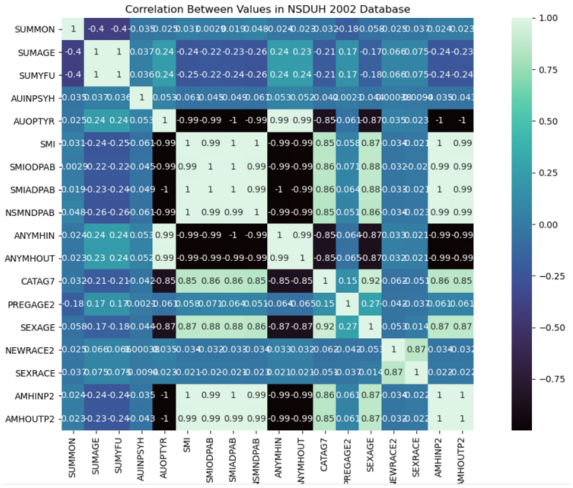

Figure 1: A heat map showing the correlation between selected key columns from the 2002 National Survey on Drug Use and Health[8]

3 Results

The results from the heat maps and the various other plots provide insights into how variables are correlated with each other.

Figure 1, Figure 2 and Figure 3 depict the heat-map that was obtained by correlating the variables with each other. Where $P < 0.5$, there is a weak correlation, while $P > 0.5$ indicates a strong correlation. These heat maps are based off of the NSDUH (National Survery on Drug Use and Health) databases, and were constructed using the seaborn module in python as well as pandas dataframes.

4 Discussion

Statistics Canada tracked data on Canadians' purchases of alcohol weekly from January to mid-April 2020, and its daily tracking of COVID-19 cases is still ongoing. As depicted by figure 2, the independent variable "number of COVID-19 cases" was plotted against the dependent variable "52-week percentage change in alcohol purchase volumes". A linear regression was performed using R to evaluate the correlation between these two variables, and Pearson's product-moment correlation was calculated. This statistical test was chosen because it measures the strength of the linear association between two variables, and the Pearson correlation coefficient, r, gives a clear numerical representation of the strength of the correlation. The ggplot R package was then used to create a linear regression graph. From the statistical analysis, the following values were derived: cor = 0.775668, t = 4.2574, df = 12, p-

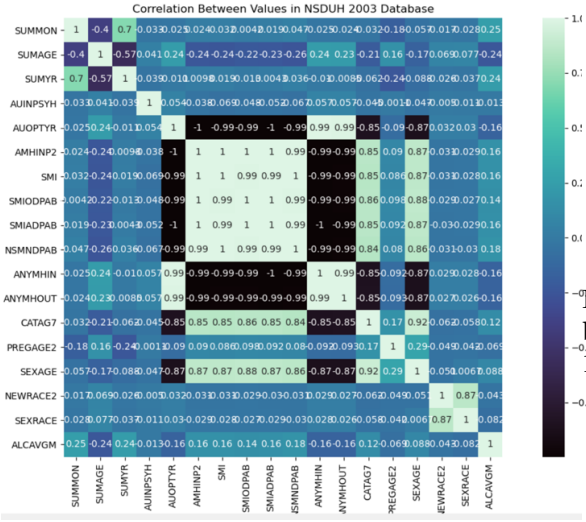

Figure 2: A heat map showing the correlation between selected key columns from the 2003 National Survey on Drug Use and Health[9]

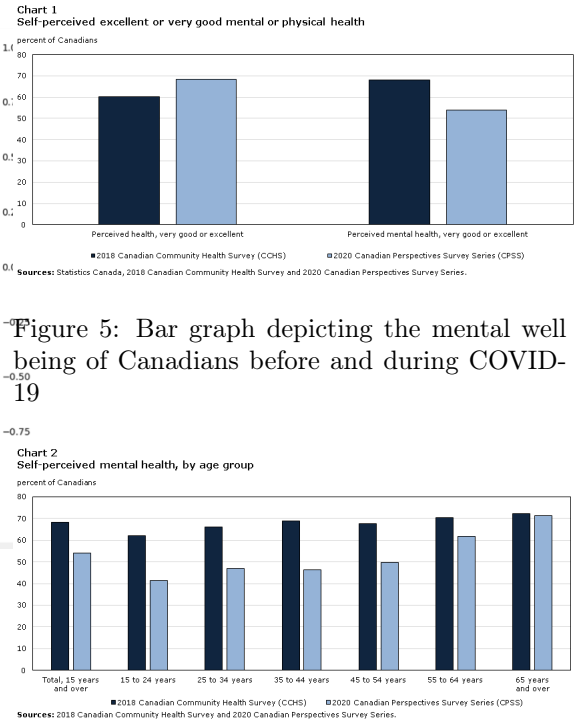

Figure 5: Bar graph depicting the mental well being of Canadians before and during COVID-19

Figure 6: Bar graph depicting the mental well being of Canadians by age group during COVID-19

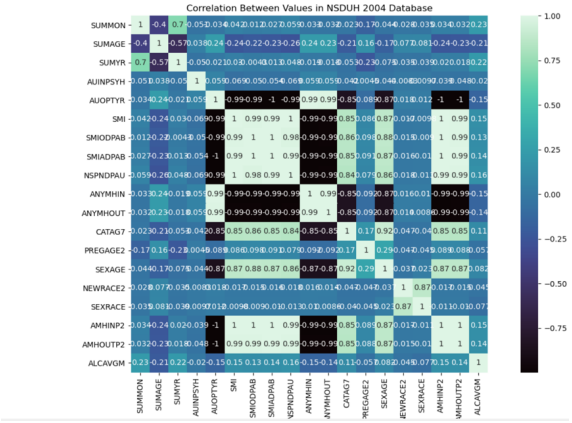

Figure 3: A heat map showing the correlation between selected key columns from the 2004 National Survey on Drug Use and Health[10]

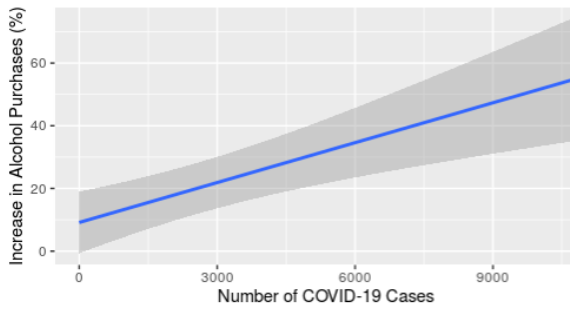

Figure 4: Linear Regression Model Showing Correlation Between Number of COVID-19 Cases and Alcohol Purchases at Grocery Stores in Canada

value = 0.001113, 95 percent confidence interval: (0.4165009, 0.9253971). From these values, it is evident that there is a significant, strong positive correlation between the increase in COVID-19 cases and an increase in Canadians purchasing alcohol.

The reliability of this analysis is limited by the fact that the COVID-19 pandemic is still ongoing and the extent of its effects have not yet been realized or studied. The data sets used are very recent and relatively small, and as stated on the Statistics Canada website, the "data are experimental and are subject to revision" [4]. It is too soon to be certain of the significance of the patterns observed, and certain variables have already begun to trend downward from their peak. For example, alcohol purchases spiked in mid-March 2020, and while they remain significantly higher than the same period in 2019, they have begun to decrease [4]. Furthermore, our analysis only demonstrates correlation, not causation. Using the same example, it is possible that alcohol purchases only increased in-store because citizens could no longer access alcohol from bars, and Canadians are not actually consuming more alcohol than before.

While comparing the various age groups we found out that the youth are most likely to report more mental health disorders. Research shows that substance use among teens often oc-

curs with other risky behaviors and can lead to substance use problems in adulthood. In 2017, more than one in ten high school students reported using illicit drugs (14%) or misusing prescription opioids (14%) [2].

While analyzing the heat maps generated a strong correlation of >0.99 in 2002 and 2004 was found between illicit drug use within populations classified as having a serious mental illness, however, this was not the case with the population that had no serious mental illness present prior to examination. This may indicate that those with a serious mental illness were the most affected during the SARS outbreak.

Conclusions

The findings show that the combination of anxiety and social isolation caused by pandemics can lead to worsened mental health outcomes and an increased prevalence of substance use. These trends were demonstrated during the 2004 SARS pandemic, during the 2020 COVID-19 pandemic, and are likely to persist during any future global virus outbreaks. Public health policy makers need to consider this issue when seeking to address the impact of COVID-19. Future studies should seek effective ways to combat this phenomenon.

Acknowledgements

We would like extend our sincerest gratitude to the Canadian STEM Fellowship for organizing the Big Data Challenge and allowing us to learn more as well as improve our technical skills and our mentor Sanjay Arumugam for guiding us. We also would like to thank the various workshop facilitators for teaching us the required skills needed to present this report in the best way possible.

References

[1] Government of Canada. Epidemiological summary of covid-19 cases in canada. 2020.

[2] Kamal R. Orgera K. Cox C. Garfield R. Hamel L. Muñana C. Panchal, N. and P. Chidambaram. The implications of covid-19 for mental health and substance use. 2020.

[3] D.H. Jernigan. America is drinking its way through the coronavirus crisis – that means more health woes ahead. 2020.

[4] Statistics Canada. Canadian consumers adapt to covid-19: A look at canadian grocery sales up to april 11. 2020.

[5] Smith L. E. Woodland L. Wessely S. Greenberg N. Webster R.K. Rubin G. J. Brooks, S. K. The psychological impact of quarantine and how to reduce it: rapid review of the evidence. 2020.

[6] Kearney A. Hamel L. Kirzinger, A. and M. Brodie. Kff health tracking poll - early april 2020: The impact of coronavirus on life in america. 2020.

[7] James A. Wiley. Impact of alcohol or drug use and incarceration on child care in santa clara county, california, 2003 (icpsr 4211). 2005.

[8] Substance Abuse and Mental Health Services Administration.

[9] Substance Abuse and Mental Health Services Administration.

[10] Substance Abuse and Mental Health Services Administration.