Female Students, Disabled Students in Canada in Worse Mental Health State Comparing to Their Peers

Haili Su, Xingyu Yu, Jiaxiang Miao, Minghui Chen

19/10/2020

Abstract

In order to study the interaction of social identites and its impact on Canadian students' mental health, we used the data from 2016 Canadian General Social Survey and performed multiple logistic regression analysis and fit a prediction model for students' self-rated mental health state based on their age, sex/gender, immigration status, visible minority status and disability status. Overall, we find the sex/gender and disability status, especially whether a student has mental/psychological or learning disability, have stronger relationship with the students mental health states among the factors studied with female students and students with disability have more negative mental health state. Meanwhile the results on other factors remain inconclusive.

Introduction

In recent years, mental health has been seen less as a tabooed subject as the public awareness on the issue raises in Canadian society (Davis, 2014). In the Western society, historically the mentally ill were (and for a large part still are) stigmatized, dehumanized, and subjected to confinement in an effort to erase their existence and insulate the public from the threat of their unreason (Foucault, 1965).Today more people are aware that this stigmatization and the cruelty it unleashed is unjustified and harmful, and that mental health issues are much more widespread in our communities. We, as a society, urgently need recognition and open discussions on this issue.

One's mental wellness is significantly impacted by their place in the community. And thanks to decades of intersectionalist research, it is better understood that every individual's mental health state is the result of the cross-effects of different aspects in their identity rather than of only one of the dimensions (Seng et al., 2017). Meanwhile, there are more and more attention on students', especially college students', mental health, as they are generally in a less advantageous economic situation and generally younger (Cooke, 2014. Castillo, 2013). So how is a student's mental wellness affected by the different factors and how do they interact with each other? Can we predict a student's perception of their own mental health state based on their social identity?

To study this question in a Canadian context, in this research we use mainly the data from the results of General Social Survey Cycle 30: Canadians at Work and Home, conducted in 2016 by Statistics Canada (2017). And we use the results to investigate the association between a respondent's self-rated mental health and their age, sex, immigration status, visible minority status, main activity in the 12 months prior to the survey, disability status and their highest education level. Overall the sample consists of 19,609 responses, and among them, 1,201 respondents younger than 44 indicated that their main activity in the past year is "going to school". In this study we call them 'student respondents' or students, and analyse the data of them.

In the 2016 GSS, a respondent can indicate their mental health as "Excellent", "Very Good", "Good", "Fair", "Poor". They can also refuse to answer or say they don't know (87 respondents chose these two responses, which is 0.45% of the sample). and We chose to categorize the "Excellent", "Very Good" and

“Good” responses as “in Good mental health state” while the “Fair” and “Poor” ones as “not in Good mental health state”. A more detailed explanation on the data we used can be found in the Data section below.

With the binary category we performed multiple logistic regression analysis with R language (R Core Team, 2020) to develop a Bayesian hierarchical model with the help of R package brms (Bürkner, 2018).

With our model, we found that female students, especially those with mental/psychological or learning disability, or other disability generally have worse mental health than their peers, while being younger or non-immigrant or non-visible minority might be contributing factors, though further analysis of our model shows that the results on those factors are less conclusive.

This report is written and produced on an R Markdown (Allaire et al., 2020) file with the help of the following R packages: tidyverse (Wickham et al., 2019) and knitr (Xie, 2020), and other packages mentioned below. This report, the complete codes and relevant files including data that support our analysis can be found in here: https://github.com/hlisu/student_mh_canada

Data

In the process of data importing and cleansing, we used the janitor R package (Firke, 2020). Our data importing codes are modified based on the work of Alexander & Caetano (2020).

The data used in this paper is derived from General Social Survey (GSS) on Canadians at Work and Home (cycle 30), which is conducted by Statistics Canada in 2016. It collects information from individuals at the age of 15 and over living in the ten provinces of Canada, excluding full-time residents of institutions. Statistics Canada’s common telephone frame, which combines landline and cellular telephone numbers from the Address Register, the Census of Population and various administrative sources and which has been integrated with Statistic Canada’s common dwelling frame, is used for the GSS data. The frame links a group of telephone numbers of the household members with an address, and contains the cell phone only households.

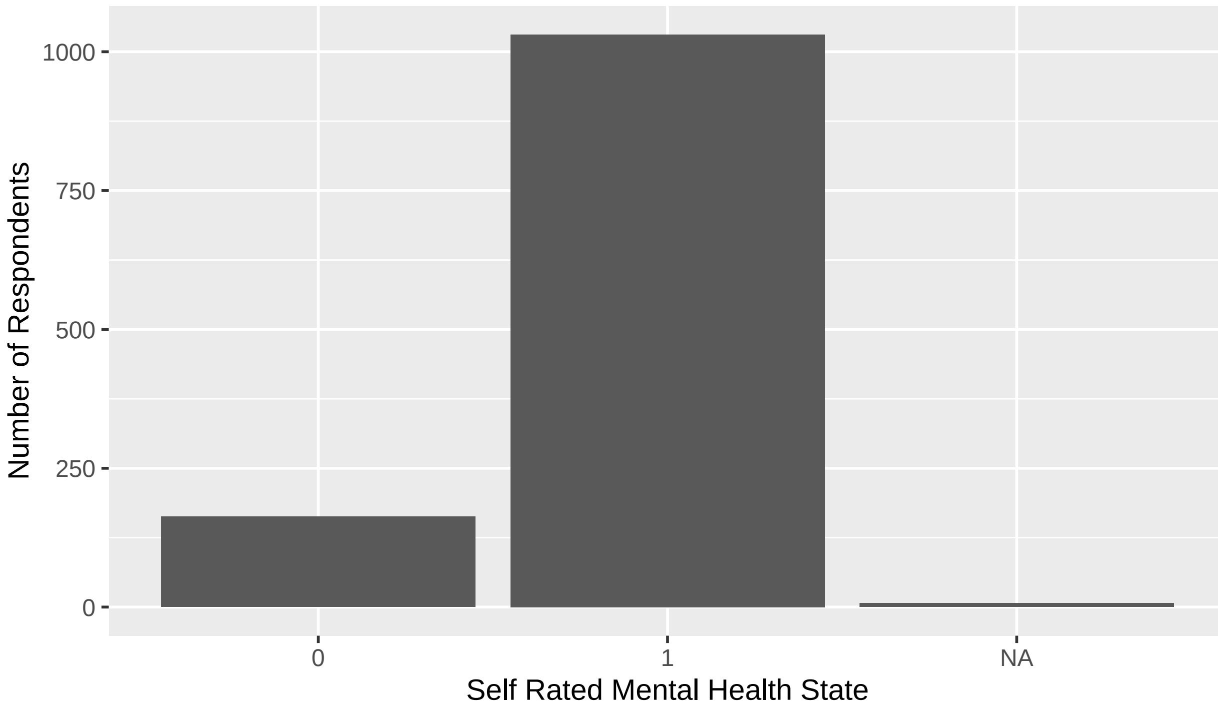
The GSS data have been collected using a combination of self-completed online questionnaires and telephone interviews since 2013. In order to maximize the data accuracy and reduce respondent burden, rather than direct survey questions, the permissions to link the respondents’ survey information to personal tax records and administrative files are asked by Statistics Canada from respondents, in order to collect household information and respondent’s personal information. In this cycle, the GSS program targets a sample size of 19,609 respondents with 2,194 variables, and has a response rate of 50.8% this year. The respondents are invited by using traditional household rostering method, meaning that the GSS program would send an invitation letter for one member in randomized households, and let them log in to the online survey for completion. If the invited member does not response, then a member from another household would be asked for their responses.

Indeed, by using this survey method, the response bias are minimized due to the access of tax records and administrative data sources. Meanwhile, as the frame has been updated and both online survey and telephone interviews are provided as options, the under-coverage issue should be improved because of the inclusion of households that only using telephones and/or no internet connection (e.g., elderly, households with low incomes), and individuals only having cell phones (e.g., younger groups). However, on the other hand, a “hand-off” issue should occur due to the usage of traditional household rostering method, when the selected household respondent is not the one who logs in to the survey, leading to a skip of the selected household so that selection bias and under-coverage issue would happen.

We define student as people who answered “Going to school” as their main activity for the last 12 months (prior to the survey) and younger than 44 – not because older people are not valid students, but because the sample size of students over 44 is too small while we wish to examine the impact of age. Overall, our sub-dataset has 1,201 observations. The tables for the specific numerical value of the proportion of each demographic groups can be found in appendix B.

Except self-rated mental health status (which is our response variable), there are 6 variables are extracted for this paper in order to do an analysis on the response variable. A detailed description on the categories and distribution of selected variables is provided in Appendix B.In the raw data of GSS, the self-rated mental health status is a categorical variables with 5 possible values: "excellent", "very good", "good", "fair", and "poor", respectively. But in order to fit our model, we have changed the response variable into a binary categorical variable, by considering "excellent", "very good" and "good" as a group of "good mental health state" and assigning a number of "1", while putting "fair" and "poor" into a category of "not a good mental health state", which is assigned a integer of "0". Figure 1 is the bar chart for distribution of GSS data after the variation, showing that 85.85% of respondents are in the group of "good mental health state", with an assigned number of 1.

Fig. 1. Distribution of Self Rated Mental Health in the Student Data



Notes: '1' indicates 'in good mental health state',
and '0' indicates 'not in good mental health state'.
Source: General Social Survey 2016, cycle 30, Statistics Canada (2017)

Meanwhile, age group (Figure 2), gender (Figure 3), and disability status (Figure 6) are the other four variables to be considered as corresponding variables to the self-rated mental health. We do not mutate the possible values for the sex, main activity over the past year, and age. On the other hand, the data for disability status are combinations of 6 variables specifying the type of disability and 1 variable asking whether the respondent has a disability or not in original GSS data. The six types of disability are "hearing", "seeing", "mental/psychological", "learning", "physical", and "unknown", respectively. We then combine the "mental/psychological" and "learning" disability as one category named "mental/psychological/learning disability", and define a second possible value "other disability" to be the consolidation of "seeing", "hearing", "physical", "unknown" disability. Furthermore, as the survey has not provided the reason for a "valid skip", we conclude the "valid skip", along with "refusal", "don't know" and "not stated", into a third category of "other situations".

Similar mutations as the adaptation for the self-rated mental health state have been conducted to the variables of visible minority and immigrant status (Figure 4 & 5). Due to the fact that the two variables are binary categorical variables, we assign a number of "1" for a "yes" answer, and a "0" for "no"s. That is, each individual who is an immigrant or a visible minority will be categorized as "1" in the corresponding variable, and each individual who is not an immigrant or not a visible minority will have "0" in the result. We want to highlight that there are two similar survey variables in the GSS data for immigration status: a derived variable immigration status indicating whether respondent is an immigrant or not, and a direct survey question variable with a question “Are you now, or have you ever been a landed immigrant in Canada?”. We have decided to choose the derived immigration status instead of the one from direct survey question, primarily due to the fact that there are no “valid skip” in the immigration status. A more detailed explanation on our decision is provided in Appendix A.

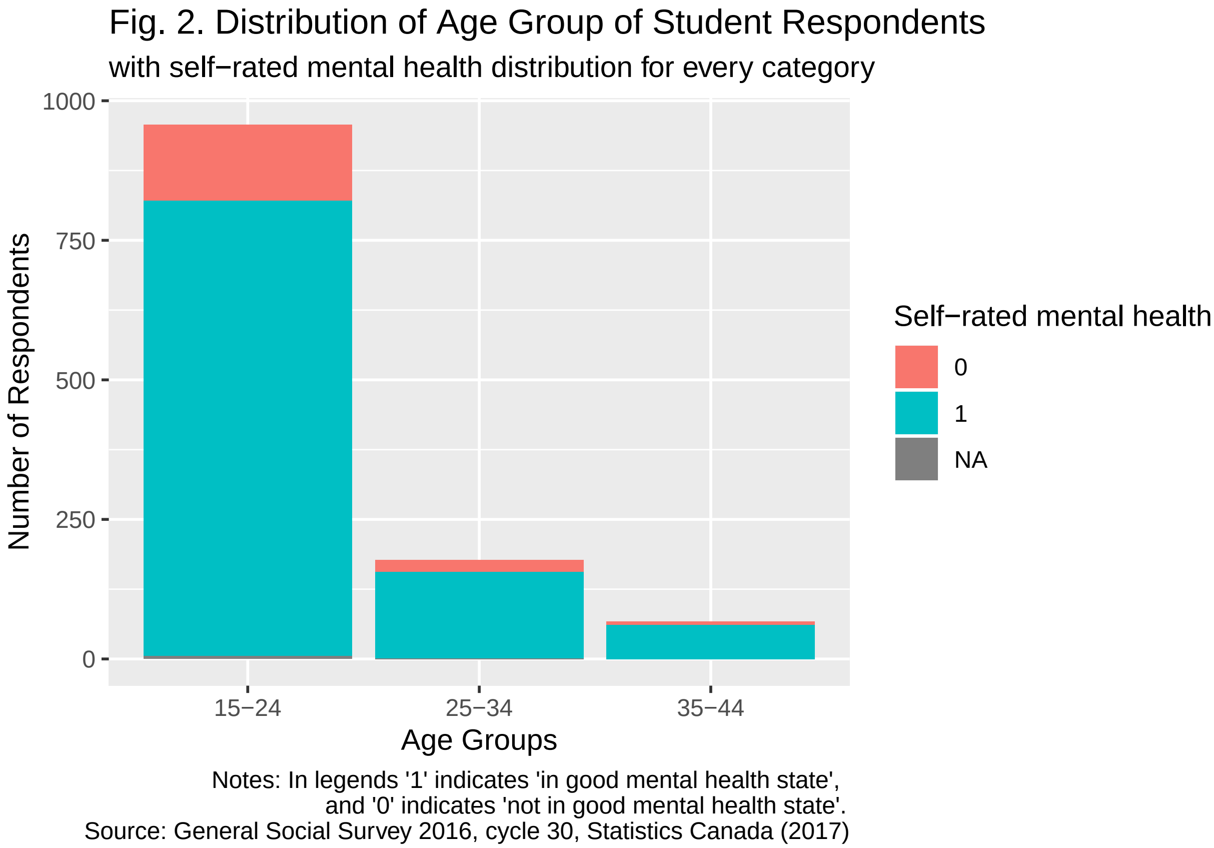
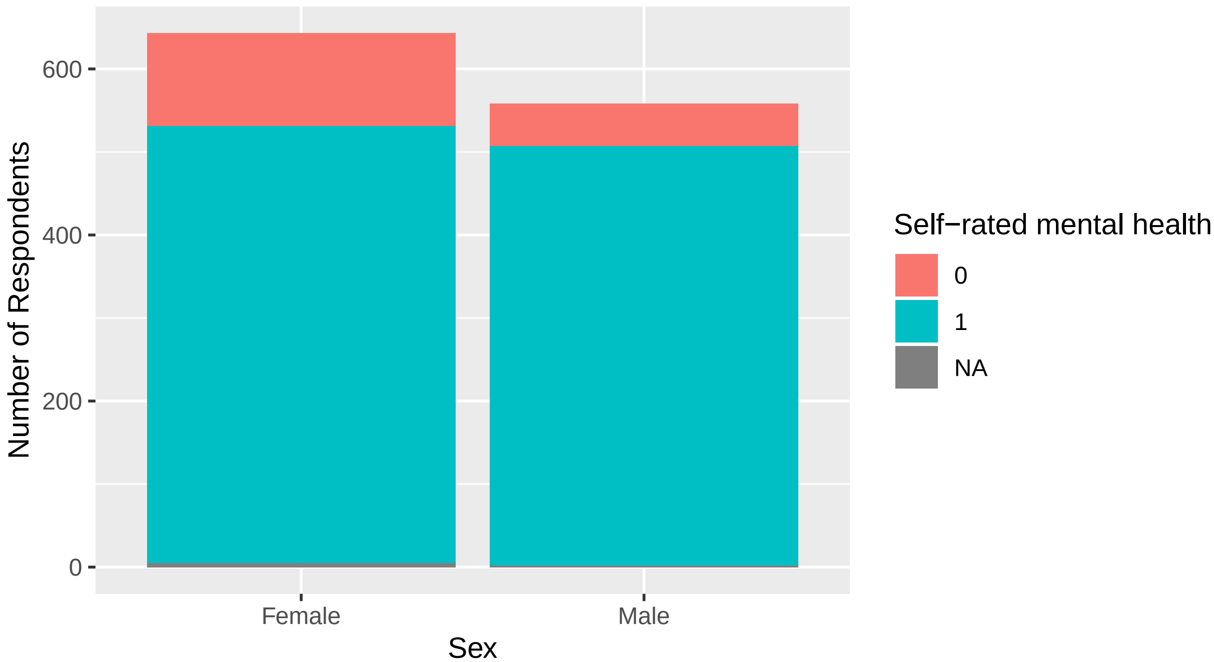


Fig. 3. Distribution of Sex of Student Respondents

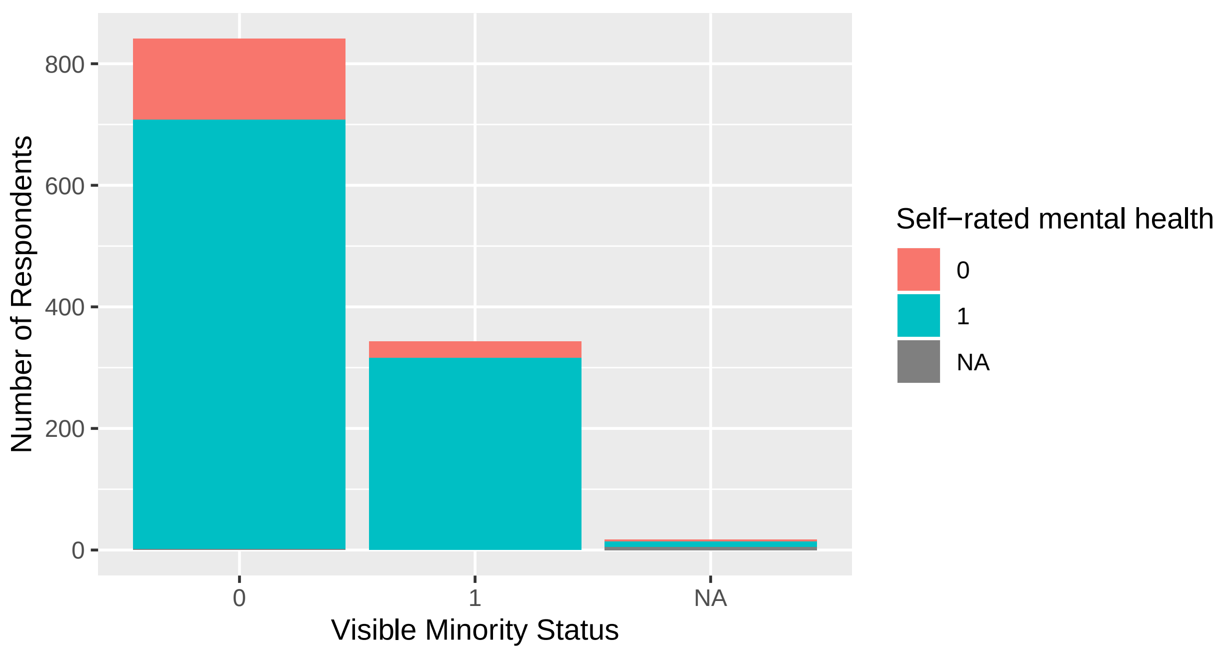
with self−rated mental health distribution for every category



Notes: In legends, '1' indicates 'in good mental health state',
and '0' indicates 'not in good mental health state'.
Source: General Social Survey 2016, cycle 30, Statistics Canada (2017)

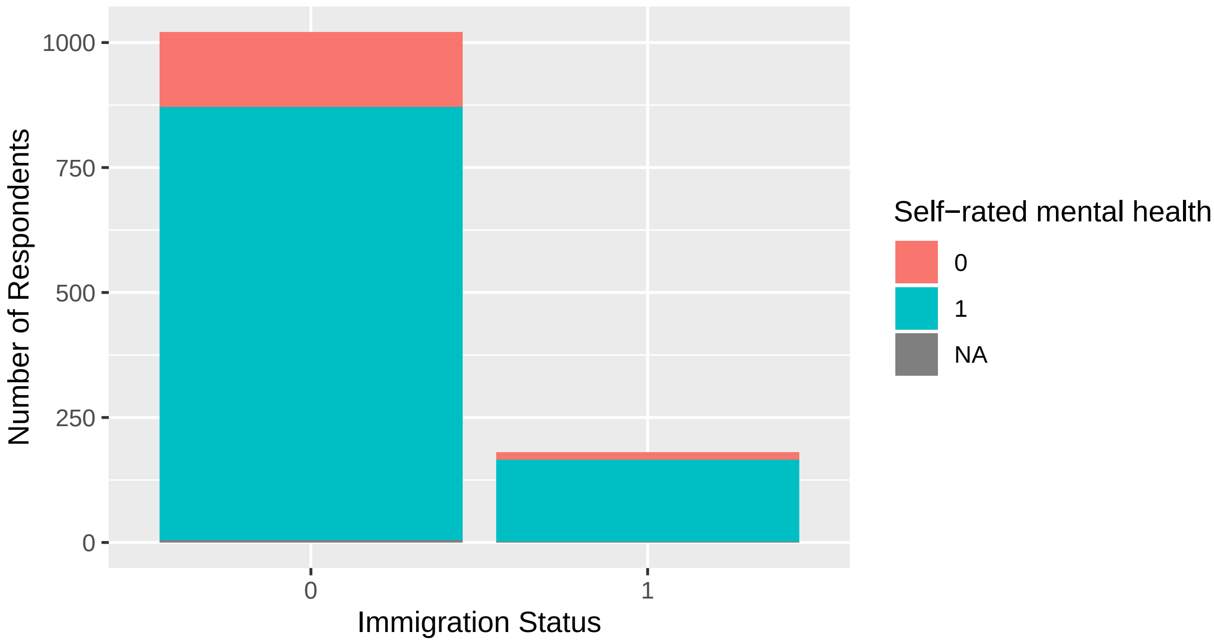
Fig. 4. Distribution of Visible Minority Status of Student Respondents

with self−rated mental health distribution for every category



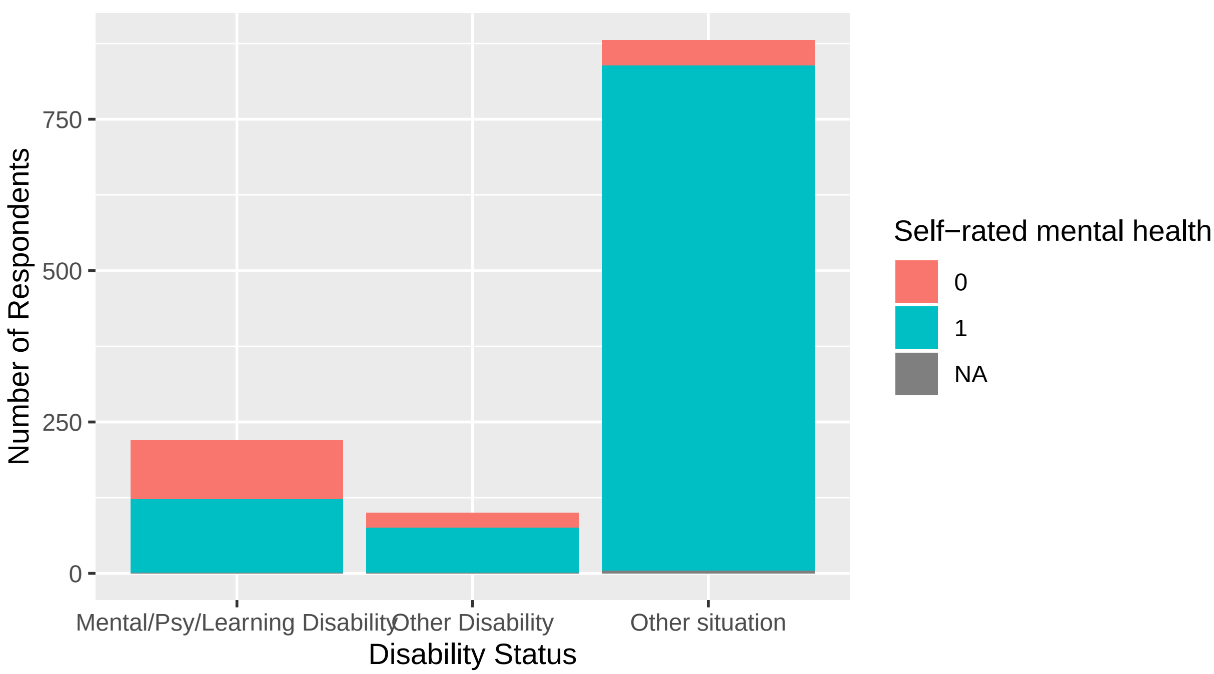
Notes: On x axis, '1' indicates 'yes' and '0' indicates 'no' for visible minority status.
In legends, '1' indicates 'in good mental health state',
and '0' indicates 'not in good mental health state'.
Source: General Social Survey 2016, cycle 30, Statistics Canada (2017)

Fig. 5. Distribution of Immigration Status of Student Respondents
with self−rated mental health distribution for every category



Notes: On x axis, '1' indicates 'yes' and '0' indicates 'no' for immigration status.
In legends, '1' indicates 'in good mental health state',
and '0' indicates 'not in good mental health state'.
Source: General Social Survey 2016, cycle 30, Statistics Canada (2017)

Fig. 6. Distribution of Disability Status of Student Respondents
with self−rated mental health distribution for every category



Notes: In legends, '1' indicates 'in good mental health state',
and '0' indicates 'not in good mental health state'.
Source: General Social Survey 2016, cycle 30, Statistics Canada (2017)

Model

Survey estimation for Standard Logistic Regression

We built a model on people who say their main activity last year is going to school, excluded the people who are older than 54 (too few of them), and have 1201 observations;

In order to fit a model, we performed multiple logistic regression analysis with R language (R Core Team, 2020) to develop a Bayesian hierarchical model with the help of R package brms (Bürkner, 2018).

Firstly, we want to discover that how a person's self-rated mental health are, based on their age, sex, visible minority status, immigration status and disability status.

$$Pr(y_i = 1) = (logit)^{-1}\left(\alpha^{age}_{a[i]} + \alpha^{sex}_{b[i]} + \alpha^{min}_{c[i]} + \alpha^{imm}_{d[i]} + \alpha^{dis}_{e[i]}\right)$$

where the α are their age, sex, main activity over the past year, visible minority, personal income,immigration status, disability status,and highest education level.

Respectively, the notation $a[i]$ refers to the age group a to which individual i belongs. These are modeled as normal distribution:

$$\alpha^{age}_{a} \sim N(0, \sigma_{age}) \text{ for } a = 1, 2, \ldots, A$$

where A is the total number of age-groups.

Respectively, the notation $b[i]$ refers to the sex group b to which individual i belongs. These are modeled as binomial distribution:

$$\alpha^{sex}_{b} \sim B(n, p)$$

Respectively, the notation $c[i]$ refers to the group of visible minority status c to which individual i belongs. These are modeled as:

$$\alpha^{min}_{c} \sim N(0, \sigma_{min}) \text{ for } c = 1, 2, \ldots, C$$

where G is the total number of visible minority.

Respectively, the notation $d[i]$ refers to the group of immigration status d to which individual i belongs. These are modeled as binomial distribution:

$$\alpha^{imm}_{d} \sim N(0, \sigma_{imm}) \text{ for } d = 1, 2, \ldots, D$$

where D is the total number of the groups of immigration status.

Respectively, the notation $e[i]$ refers to the group of disabilities status e to which individual i belongs. These are modeled as:

$$\alpha^{dis}_{e} \sim N(0, \sigma_{dis}) \text{ for } h = 1, 2, \ldots, H$$

where H is the total number of groups of disabilities status.

A preliminary summary of our model fitted with brms is shown below, the detailed results and interpretation of them will be shown in the following sections:

```
##  Family: binomial
##   Links: mu = logit
## Formula: self_rated_mental_health ~ age + sex + immigration_status + visible_minority + disability_s
##    Data: my_data_2 (Number of observations: 1182)
## Samples: 4 chains, each with iter = 2000; warmup = 1000; thin = 1;
##          total post-warmup samples = 4000
```

```
##
## Population-Level Effects:
##                                   Estimate Est.Error l-95% CI u-95% CI Rhat
## Intercept                             0.00      0.16    -0.33     0.32 1.00
## age25M34                              0.16      0.29    -0.39     0.73 1.00
## age35M44                              0.58      0.50    -0.34     1.66 1.00
## sexMale                               0.41      0.21     0.01     0.82 1.00
## immigration_status                    0.29      0.37    -0.42     1.01 1.00
## visible_minority                      0.24      0.27    -0.28     0.78 1.00
## disability_statusOtherDisability      0.91      0.29     0.35     1.48 1.00
## disability_statusOthersituation       2.68      0.22     2.27     3.10 1.00
##                                   Bulk_ESS Tail_ESS
## Intercept                             5930     3359
## age25M34                              6440     2917
## age35M44                              5226     2610
## sexMale                               6859     2741
## immigration_status                    4567     2840
## visible_minority                      4836     2902
## disability_statusOtherDisability      6791     2965
## disability_statusOthersituation       3596     3146
##
## Samples were drawn using sampling(NUTS). For each parameter, Bulk_ESS
## and Tail_ESS are effective sample size measures, and Rhat is the potential
## scale reduction factor on split chains (at convergence, Rhat = 1).
```

Results

Identify the regression equation

$$\log \frac{p}{1-p} = \beta_0+\beta_1 X_{age25-34}+\beta_2 X_{age35-44}+\beta_3 X_{Male}+\beta_4 X_{immigrant}+\beta_5 X_{minority}+\beta_6 X_{OtherDisability}+\beta_7 X_{OtherSituation}$$

Fill out the coefficients:

$$\log \frac{p}{1-p} = 0.006+0.145 X_{age25-34}+0.518 X_{age35-44}+0.409 X_{Male}+0.27 X_{immigrant}+0.233 X_{minority}+ 0.887 X_{OtherDisability}+2$$

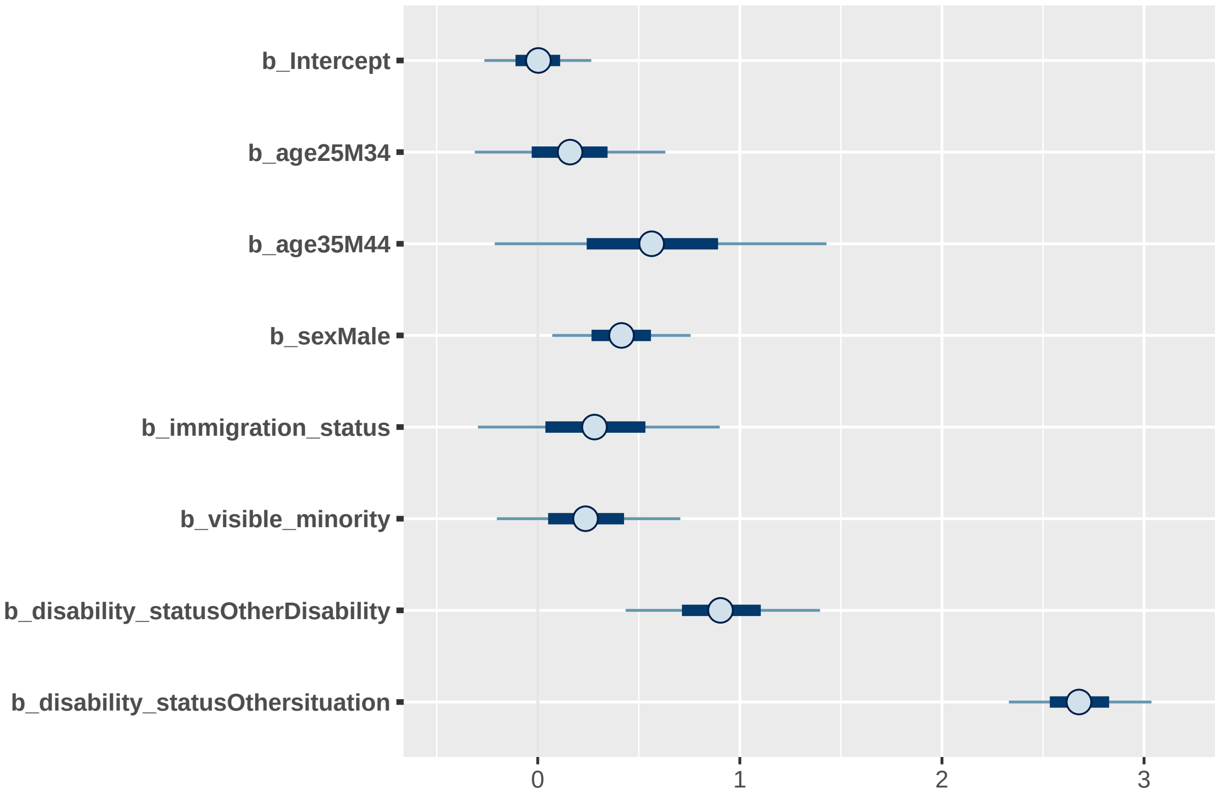
A summary of our parameter estimations and 95% confidence intervals:

Table 1: Parameter Estimates in our final multiple logistic regression model

Factor	Parameter Estimate	Error Estimate	95% CI: lower	95% CI: upper
Intercept	0.0019377	0.1630194	-0.3250230	0.3241694
Age: 25-34	0.1606620	0.2866085	-0.3920922	0.7328615
Age: 35-44	0.5837051	0.5031378	-0.3422545	1.6560639
Male	0.4121768	0.2115589	0.0077490	0.8220227
Immigrant	0.2880618	0.3653142	-0.4174605	1.0086435
Visible Minority	0.2404603	0.2740580	-0.2765570	0.7807198
Disability: Other	0.9113572	0.2925253	0.3531617	1.4806066
Disability: Other Situation	2.6807777	0.2155719	2.2688201	3.1028113

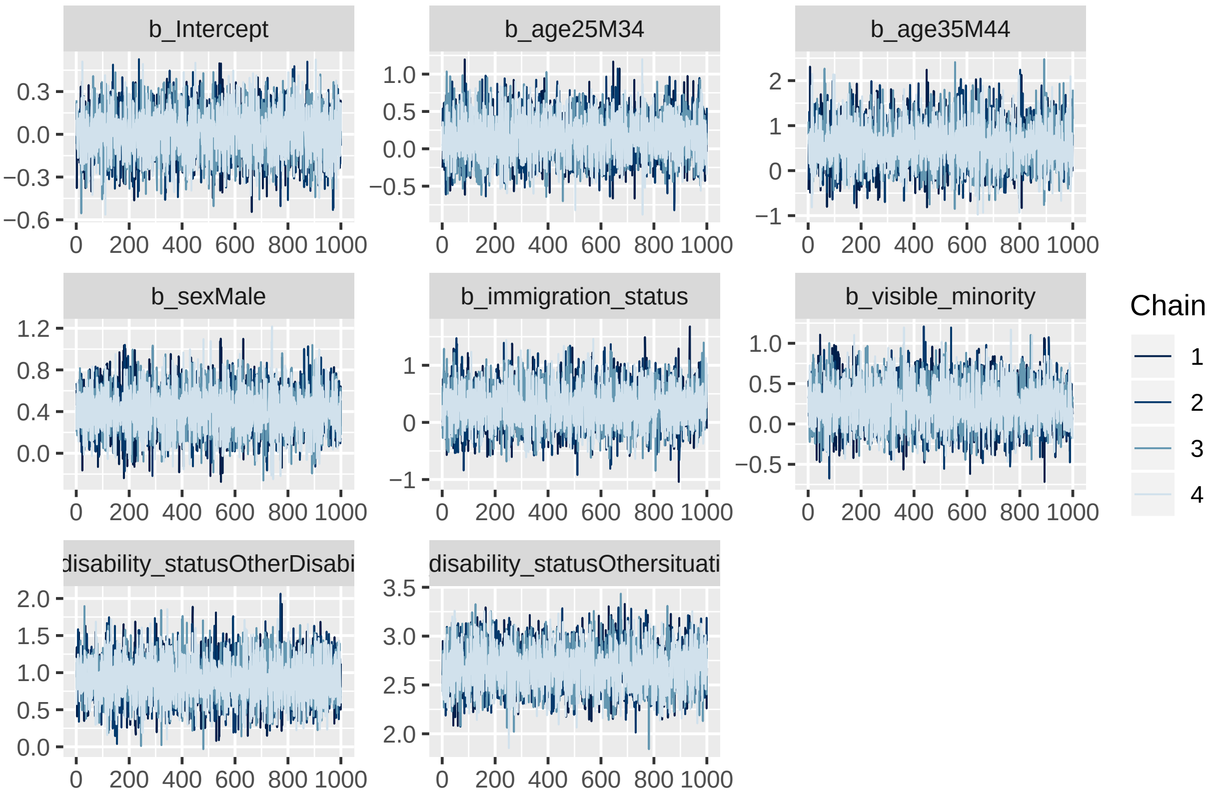
MCMC Plot: Visualizing estimates and 95% CIs:

Fig. 7: A MCMC Plot for Model Parameter Estimations



MCMC Plot: Visualizing the distribution of our prediction:

Fig. 8: MCMC trace plots for our model



Discussion

As the table for parameter estimation in the result section shows, the interception is about 0.0019, which is the log odds of being in a good mental health state for a student who is 15-24 years old, female, non-immigrant, non-visible minority and has mental/psychological or learning disability. The real estimated odds for her to be in a good mental state, $\hat{p}$ for this person would be, by our formula in the model section, $\frac{exp(\hat{\beta_0})}{1-exp(\hat{\beta_0})}$ is near zero–a result with grim indication.

The parameter of each factor that's in our model shows how much the demographic information could impact a student's mental state. For age group factor, the older you are, the higher the odds. As the estimated β_1 when someone is 25-34 years old is 0.16, while estimated β_2 for 35-44 is 0.58. Both are positive showing that younger students are generally in a worse mental state. However as noticed in fig.2, the sample size of students older than 24 are significantly smaller. Fig.7, which shows the 95% confidence interval of our estimates, also shows that the parameter estimations for these two age groups have a wider spread.

When it comes to student's sex/gender identity, our model shows that female students have generally worse mental states. In a patriarchal society where sex-and-gender-based discrimination and violence are still widespread, this result is a reflection of the sad societal reality, affirming the results of numbers of previous studies (Koss 1990, Lennon & Rosenfield, 1992) on this issue. Mental health care on campus needs to address the issues that female students are facing.

And when analyzing the ethnic identity and immigration experience, however, the model shows surprising results, as being a visible minority and/or immigrant have positive effects to the odds of having a good mental health state for students. These results contradict the mainstream perception and prior studies (Pumariega, Rothe & Pumariega 2005, Gary 2005). Considering the relatively small sample size for immigrant-background and visible minority students, as shown in the data section, and the relatively wide confidence intervals for the estimations as shown in fig.7, we should critically interpret these results.

As for the disability status of students, it is unsurprising that students with mental/psychological or learning disability are worse off than other peers. While being non-disable ("other situation") has the most positive effect on a student's mental health among all the factors, the effect is far more positive than students who have non-mental or learning disabilities. Considering the confidence interval for this parameter estimation is one of the narrowest, this shows that students with disability experience significantly negative impacts on their mental health.

Overall, from fig.8, our model seems to be making relatively consistent and satifactory prediction of a student's mental health state. Overall, we identify that female students, especially those with mental/psychological or learning disability, or other disability generally have worse mental health than their peers, while being younger or non-immigrant or non-visible minority might be contributing factors.

Weakness

The binary categorization for mental wellness might be an oversimplification. In fact, the original answers include five levels. A more detailed multilevel modeling could further our understanding on this issue. People may not tell us their real thoughts by some personal issues, they are not willing to do it. In terms of the GSS data, It was collected using a combination of self-completed online questionnaires and telephone interviews since 2013 which was really inconvenient and may lose the reliability and accuracy.

Also, the nature of "other situations" in the disability, as the data section introduced, might not be as simple as "not having disability". A survey designed to target the student population of Canada could provide us with a clearer picture. Or, we might need stratification based on the real world data: although the survey itself adopted a stratified sampling method, when it comes to students in the sample, the demographic make-up doesn't fully reflect the reality given its small sample size. Without a specifically designed survey on students' mental health, this might be the next best thing we can do.

Appendices

A.Why do we use Landed Immigration Status instead of Citizenship Status when studying the impact of immigration experience?

In the questionnaire, BPR_Q16 is a question about if the respondent has ever been a landed immigrant ("Are you now, or have you ever been a landed immigrant in Canada?"), which means if the respondent has ever been granted permanent residency by the Government of Canada. The data also includes a "citizenship status" of which the original question is not found on the survey questionnaire website. The answers for this question include: - (non-immigrants) - Canadian citizen by birth and other citizenship(s) - Canadian citizen by birth only - (immigrants) - Canadian citizen by naturalization and other citizenship(s) - Canadian citizen by naturalization only - Other citizenship(s) non-Canadian only -(undetermined) - Undetermined.

Ideally the answers for citizenship questions are better for investigating the impact of immigration experience on one's mental health, because not only the answer design allows us to include the landed immigrants (who might or might not later become a Canadian citizen), but also the temporary residents who are on study/work/dependent permits. These temporary residents also experience most of the barriers that landed immigrants face, like language barriers, social isolation, discrimination and so on.

However the "undetermined" category adds uncertainty to the data. Here are the proportions of each response to citizenship status question in the sample:31.47% of respondents chose "undetermined", and we don't really know the nature of these responses.

Meanwhile, other immigrants categories make up 12.41%. In the data session we saw 17.25% of the respondents are or were landed immigrants. According to Statistics Canada: "According to the 2016 Census, 7,540,830 people, that is, 21.9% of the Canadian population, were foreign-born (immigrants), 26,412,610 (76.6%) were

Canadian-born (non-immigrants) and 506,625 (1.5%) were non-permanent residents." (Statistic Canada, 2017)

Therefore 17.25% of respondents being landed immigrants is closer to the 21.9% in census results, and should be the more ideal indicator for investigating the mental health impact of immigration here.

B. Summary of the numeric value of proportion of each demographic group in the sample:

Table 2: Self-rated mental health state results in the student data

mental health state	n	proportion %
0	163	13.57
1	1031	85.85
NA	7	0.58

Table 3: Age group distribution in the student data

age	n	proportion %
15-24	957	77.43
25-34	177	14.32
35-44	67	5.42
45-54	31	2.51
55-64	4	0.32

Table 4: Sex distribution in the student data

sex	n	proportion %
Female	643	53.54
Male	558	46.46

Table 5: Immigration status distribution in the student data

immigration state	n	proportion %
0	1021	85.01
1	180	14.99

Table 6: Visible minority distribution in the student data

visible minority	n	proportion %
0	841	70.02
1	343	28.56
NA	17	1.42

Table 7: Disability status distribution in the student data

disability	n	proportion %
Mental/Psy/Learning Disability	220	18.32
Other Disability	100	8.33
Other situation	881	73.36

Reference

- Alexander, R. & Caetano, S (2020). GSS cleansing code. URL, https://q.utoronto.ca/courses/184062/files/9416214/download?wrap=1.
- Allaire J. et al., (2020). rmarkdown: Dynamic Documents for R. R package version 2.4. URL https://rmarkdown.rstudio.com.
- Auguie, B. (2017). gridExtra: Miscellaneous Functions for "Grid" Graphics. R package version 2.3. https://CRAN.R-project.org/package=gridExtra
- Bürkner, P. (2018). Advanced Bayesian Multilevel Modeling with the R Package brms. The R Journal, 10(1), 395-411. doi:10.32614/RJ-2018-017
- Castillo, L. G., & Schwartz, S. J. (2013). Introduction to the Special Issue on College Student Mental Health. Journal of Clinical Psychology, 69(4), 291–297. https://doi.org/10.1002/jclp.21972
- Cooke, R. et al., (2004) Student debt and its relation to student mental health, Journal of Further and Higher Education, 28:1, 53-66, DOI: 10.1080/0309877032000161814
- Davis, S. (2014). Community mental health in Canada: Theory, policy, and practice.
- Firke, S. (2020). janitor: Simple Tools for Examining and Cleaning Dirty Data. R package version 2.0.1. https://github.com/sfirke/janitor
- Foucault, M. (1965). Madness and civilization: A history of insanity in the age of reason.
- Gary, F. A. (2005). STIGMA: BARRIER TO MENTAL HEALTH CARE AMONG ETHNIC MINORITIES. Issues in Mental Health Nursing, 26(10), 979–999. https://doi.org/10.1080/01612840500280638
- Koss, M. P. (1990). The women's mental health research agenda: Violence against women. American Psychologist, 45(3), 374–380. https://doi.org/10.1037/0003-066X.45.3.374
- Lennon, M., & Rosenfield, S. (1992). Women and Mental Health: The Interaction of Job and Family Conditions. Journal of Health and Social Behavior, 33(4), 316-327. Retrieved October 20, 2020, from http://www.jstor.org/stable/2137311
- Pumariega, A.J., Rothe, E. & Pumariega, J.B. Mental Health of Immigrants and Refugees. Community Ment Health J 41, 581–597 (2005). https://doi.org/10.1007/s10597-005-6363-1
- R Core Team (2020). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.
- Seng, J. S., Lopez, W. D., Sperlich, M., Hamama, L., & Reed Meldrum, C. D. (2012). Marginalized identities, discrimination burden, and mental health: Empirical exploration of an interpersonal-level approach to modeling intersectionality. Social Science & Medicine, 75(12), 2437–2445. https://doi.org/10.1016/j.socscimed.2012.09.023
- Statistics Canada. (2017). General social survey (GSS), 2016: Cycle 30, Canadians at Work and Home. [Public use microdata file and codebook]. Ottawa, ON: Statistics Canada. Retrieved from https://www23.statcan.gc.ca/imdb/p2SV.pl?Function=getSurvey&SDDS=5221
- Statistics Canada (2017). 2016 Census of Population. URL, https://www12.statcan.gc.ca/census-recensement/2016/dp-pd/index-eng.cfm.
- Wickham, H. et al., (2019). Welcome to the tidyverse. Journal of Open Source Software, 4(43), 1686, https://doi.org/10.21105/joss.01686
- Xie, Y. (2020). knitr: A General-Purpose Package for Dynamic Report Generation in R. R package version 1.27.