How Canadian Households' Sense of Belonging Related to Specific Social Identity Factors in 2013

Ziyu Jiang, Yizhen Wang, Zifeng Zhu

2022-03-20

Abstract

Social Identity Survey created and distributed by GSS collected Canadian households' information about social identities. This report will focus on the topic of what variables might Canadian's sense of belonging to Canada depends on. The report has concluded that the level of satisfaction when communicate with their relatives, frequency of following news and current affairs, level of proudness to be Canadian, level of satisfaction of balance between work and life, highest education degree, has been a victim of discrimination in the past five years or not, Income level of the survey respondents are correlated with the their level of sense of belonging to Canada.

Introduction

The sense of belonging is not only about simply getting involved with other people; it is focused on gaining acceptance, attention, and support from members of the group as well as providing the same attention to other members.(Cherry 2021) It was being said that females belong strongly in slightly larger proportions than males, and in Quebec, lower proportions belong strongly to Canada. (Painter 2013) Immigration, Refugees and Citizenship Canada (CIC) has put interests on the topics of sense of belonging in multiple scales for their current and future policies and programs of integration, citizenship and multiculturalism.(Painter 2013) The GSS program, established in 1985, conducts surveys across the 10 provinces. The GSS is recognized for its regular collection of cross-sectional data that allows for trend analysis, and its capacity to test and develop new concepts that address current and emerging issues.(*GSS, Cycle 27: Social Identity* 2013) The GSS created and distributed a social survey about Social Identity and collected the data from the respondents via Statistics Canada.

After analyzing the data in the certain period of time, we can see that the survey respondent's level of satisfaction when communicate with their relatives, frequency of following news and current affairs, level of proudness to be Canadian, level of satisfaction of balance between work and life, highest education degree, has been a victim of discrimination in the past five years or not, Income level are correlated with the survey respondent's level of sense of belonging to Canada. In this report, we will use the social identity survey data from Canadian general social surveys (GSS) to investigate the relationship between the survey respondent's level of sense of belonging to Canada and other independent variables. We will discuss the limitation of the data and the potential bias from the data, apply the logistic regression model on the data. At last, we will discuss the distributions of the independents variables in interests, the resulting influences of the final model, the weakness of our researching processes and the lessons learned from the research and the possible improvement in future.

R statistical programming language(R Core Team 2021) is used for analyzing, and the package `tidyverse`(Wickham et al. 2019) is used for data visualizing and data manipulating in this project. The package `broom`(Robinson, Hayes, and Couch 2022) is also used in this project to convert the statistical objects into tidy tibbles, the package `knitr`(Xie 2014) is used to knit the R markdown file to pdf form and ro create the tables,and the package `ggplot2`(Wickham 2016) is used for creating graphs. Finally, the package `car`(Fox and Weisberg 2019) is used to creating regression models of the variables.

Data

Data Source

The 2013 Canadian General Social Survey on Social Identity has interviewed individuals who were 15 years old and over through Canada's ten provinces, and the survey data are subject to both sampling and non-sampling errors (*GSS, Cycle 27: Social Identity* 2013). This report is based on the data of survey questionnaire results published on Canadian GSS website, and the csv format of the data was loaded into R Studio by R project (R Core Team 2021). The dataset contains several variables that we are interested to study on, such as citizens' sense of social belonging, people's connections with friends and relatives and so on. The collection of data starts from June 2013 to March 2014, and the dataset was the second version of social identity survey. The version 1 was conducted in 2008.

Survey Methodology and Collection

The Social Identity survey we studied was firstly meant to gather data on social trends in order to monitor changes in the living conditions and well-being of Canadians over time, and secondly to provide immediate information on specific social policy issues of current or emerging interest (*GSS, Cycle 27: Social Identity* 2013). In this report, we will use provided survey information to explore how Canadian residents' sense of belonging was related to several specific factors, such as their social networking, mental health and so on.

The target population for the survey included all people who were 15 years and older in Canada's ten provinces in 2013, excluding residents of the Yukon, Northwest Territories, and Nunavut as well as full-time residents of institutions. What's more, in order to carry out sampling, the ten provinces were respectively divided into geographic areas. For instance, many Census Metropolitan Areas were each considered separate areas. Survey interviews were conducted via computer assisted telephone interviewing (CATI) and electronic questionnaire (EQ). And the survey frame was created using two different components, which were lists of telephone numbers in use available to Statistics Canada from various sources and the list of all dwellings within the ten provinces. In each frame, each group of telephone numbers was assigned to a stratum within its province. By conducting indicated survey frame, the target sample size for the survey was 31,973, while the actual number of respondents was 27,695. From the results, coverage of the targeted population by survey frame was estimated to be more than 88 percent complete. The exclusion mostly came from households without eligible telephones, and some GSS respondents were removed for confidentiality reasons. At the same time, the whole sample was divided into three categories. For the regular sample, a household was eligible if it included at least one person 15 years of age or older. For the oversample of immigrants, a household was eligible if it included at least one person 15 years of age or older born outside of Canada. Lastly, for the oversample of youth, a household was eligible if it included at least one person between the age of 15 and 24 years old (*GSS, Cycle 27: Social Identity* 2013).

The data collected for the survey are made up of two components: classification and core content. Classification content (age, sex, education, income) helps to delineate population groups for use in the analysis of core data. Core content is designed to measure changes in society related to living conditions and well-being and to supply data to inform specific policy issues (*GSS, Cycle 27: Social Identity* 2013). Since the survey data collected residents' information all over Canada, the dataset would have a large sample and cover a relatively wider range of participants. Thus, the survey results would be very representative. Also, interviewing through telephone and online questionnaire could make the process quite efficient and convenient. However, because of these fast means of approaching, the amount of non-response would increase. In this case, weights for responding households were adjusted to represent non-responding households. This was done independently within each stratum group. Firstly, adjustment was made for complete non-response. Then the adjustment was done for partial non-response, these households had some auxiliary information which was used to model propensity of response. The combination of these two adjustments were recorded, non-responding households were then dropped (*GSS, Cycle 27: Social Identity* 2013).

Similar datasets could not been found and used, since this Social Identity survey was conducted all over Canada by Statistics Canada and GSS. The data size was massive and only national institutions such as Statistics Canada could obtain the frame (i.e., households telephone numbers) to reach to most households

around the country. So it is hard to find similar datasets.

What's more, some bias may exist in this Social Identity survey. For instance, since the survey was done in 2013, the technology at that time was not as developed as these days. So some questions such as accessing of social networking website may have huge amount of valid skip, since many people did not have social media account at that time. But in general, the survey is relatively comprehensive and objective.

Data Characteristic

There were 790 variables and 27534 observations in the raw dataset with no missing values. The variables this dataset contained were covering excessive topics for this research, it would be simpler and more direct to the interested topic if applying cleaning processes. The first step of the cleaning process was selecting the variables related to the topic in interest; 16 variables were selected by common sense for analysing the interested topic, those are: agegr10, sex, livarr06, icr_30, cwr_40, cwf_60, grp_40, mcr_310,prd_10, wfr_510, dh1ged, sbl_500, discrim, srh_115, relig7, and incm. All of these variables are categorical and the meanings these variables stand for were included in Table 1 which was created with function `kable`(Xie 2014). A dataframe called "df" was created for table1, which includes two new created variables "variables" and "descriptions".

Table 1: Variable names and descriptions.

variables	descriptions
agegr10	Age Group of Survey Respondents
sex	Gender of Survey Respondents
livarr06	Living arrangement of Survey Respondent's household
icr_30	Survey Respondent's frequency of access of social networking sites
cwr_40	Survey Respondent's level of satisfaction when communicate with their relatives
cwf_60	Survey Respondent's level of satisfaction when communicate with their friends
grp_40	Survey Respondent's frequency of participating in group activities and meetings
mcr_310	Survey Respondent's frequency of following news and current affairs
prd_10	Survey Respondent's level of proudness to be Canadian
wfr_510	Survey Respondent's level of satisfaction of balance between work and life
dh1ged	Survey Respondent's highest education degree
sbl_500	Survey Respondent's level of sense of belonging to Canada
discrim	Survey respondent has been a victim of discrimination in the past five years or not
srh_115	Survey respondent's mental health state(self-rated)
relig7	Survey respondent's religion
incm	Survey respondent's Income level

Since the variables selected were all categorical with numeric way of expression in the dataset, the numerical values in the variables were rewritten with the strings of their meanings for the graphing necessity. At the same time, a new dataset was created from the raw dataset which selected the chosen variables and filtered out the observations who did not answer the survey questions clearly.

After that, a new variable called sbl_500_level was created which combined the 4 responses of the variable sbl_500 into two levels: "Very strong" and "Somewhat strong" to the level "Strong"; "Very weak" and "Somewhat weak" to the level "Weak". Then the observations in the level "Strong" were assigned the value 1 for the new variable sbl_500_level, and the observations in the level "Weak" were assigned the value 0 for the new variable sbl_500_level.

Barplots of these 16 variables were created by 'ggplot2' (Wickham 2016). From the plots we can observe that most respondents' ages were from 35 to 64 years old, and female respondents took a bit larger proportion than male, also, over 8000 of them have obtained post-secondary diploma. The respondents mostly lived with their spouse or on their own. And almost every households was quite satisfied with their communication with friends and relatives, only very small proportion of respondents had poor mental health (less than 500). Many

of these residents participated in group activities at least once a week, yet there were still more than 2000 residents did not participate in any group meetings in the past year. Most households had annual income level between $20,000 to $40,000. More than 15000 respondents followed news on daily bases, and almost every participants were quite proud to be Canadian, yet there were still one third of them have experienced any kind of discrimination. In general, most participants had relatively strong sense of belonging in Canada (over 20000), and only a few of them had weak sense of belonging (less than 2500). However, huge amount of valid skip answers appeared in several questions such as social networking and balance between work and life. So our following analysis will focus on data with more concrete answers.

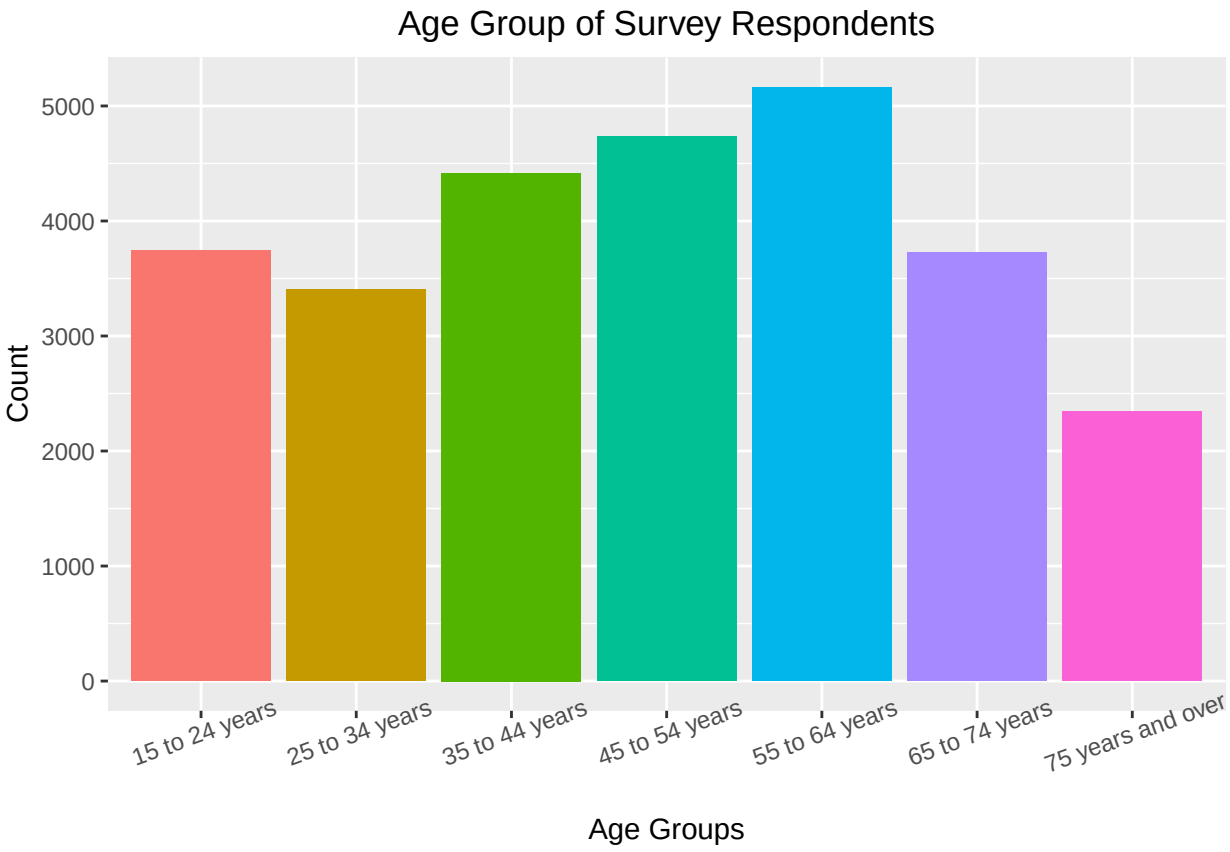


Figure 1: Distribution of the survey respondent's age (in years).

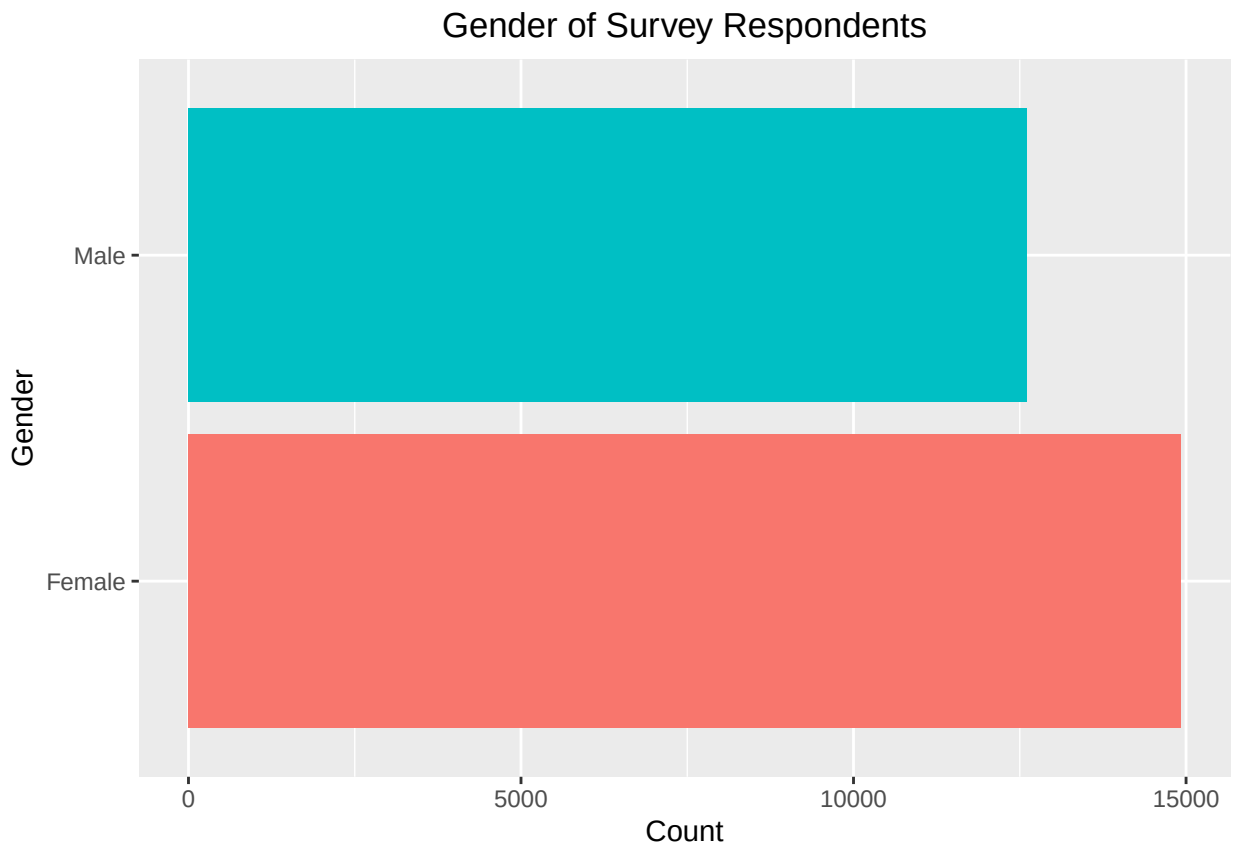


Figure 2: Distribution of Survey Respondent's Gender.

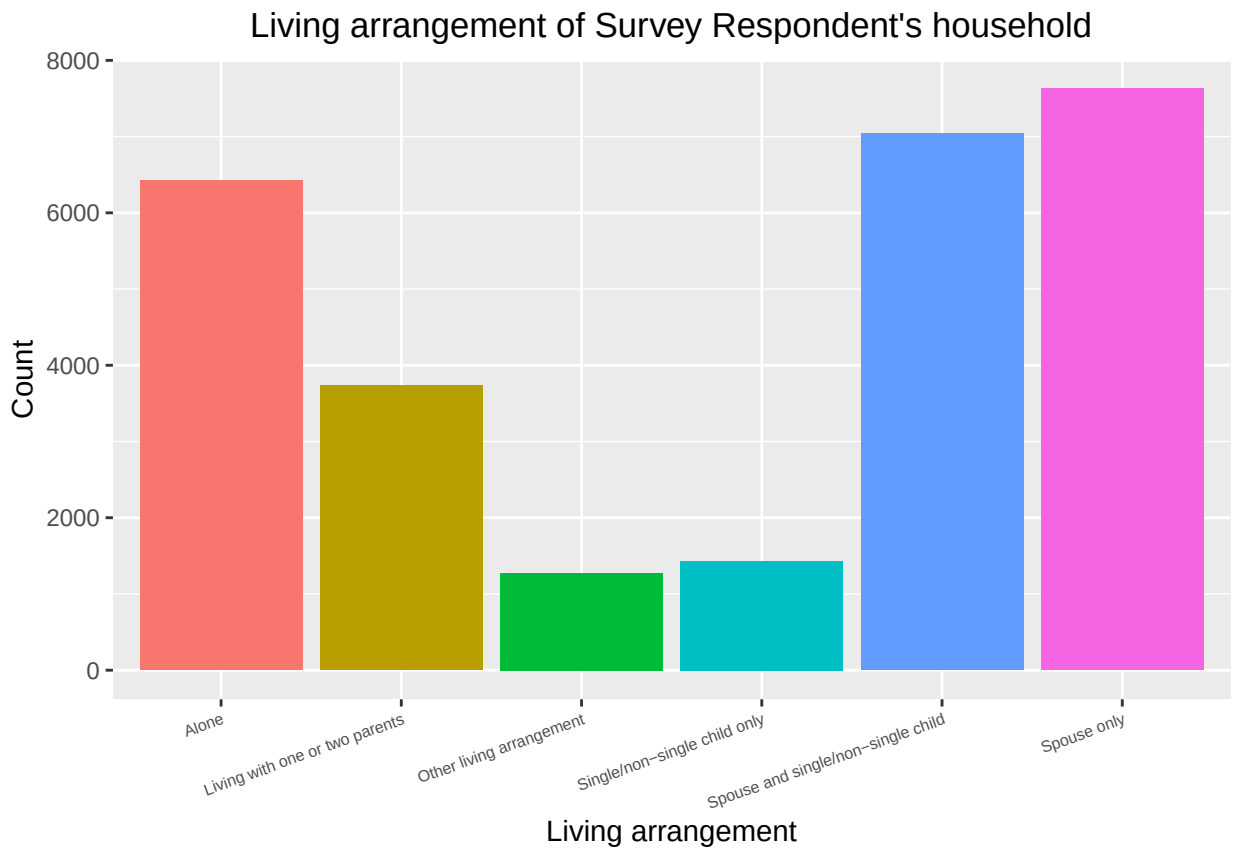


Figure 3: Distribution of living arrangement of Survey Respondent's household.

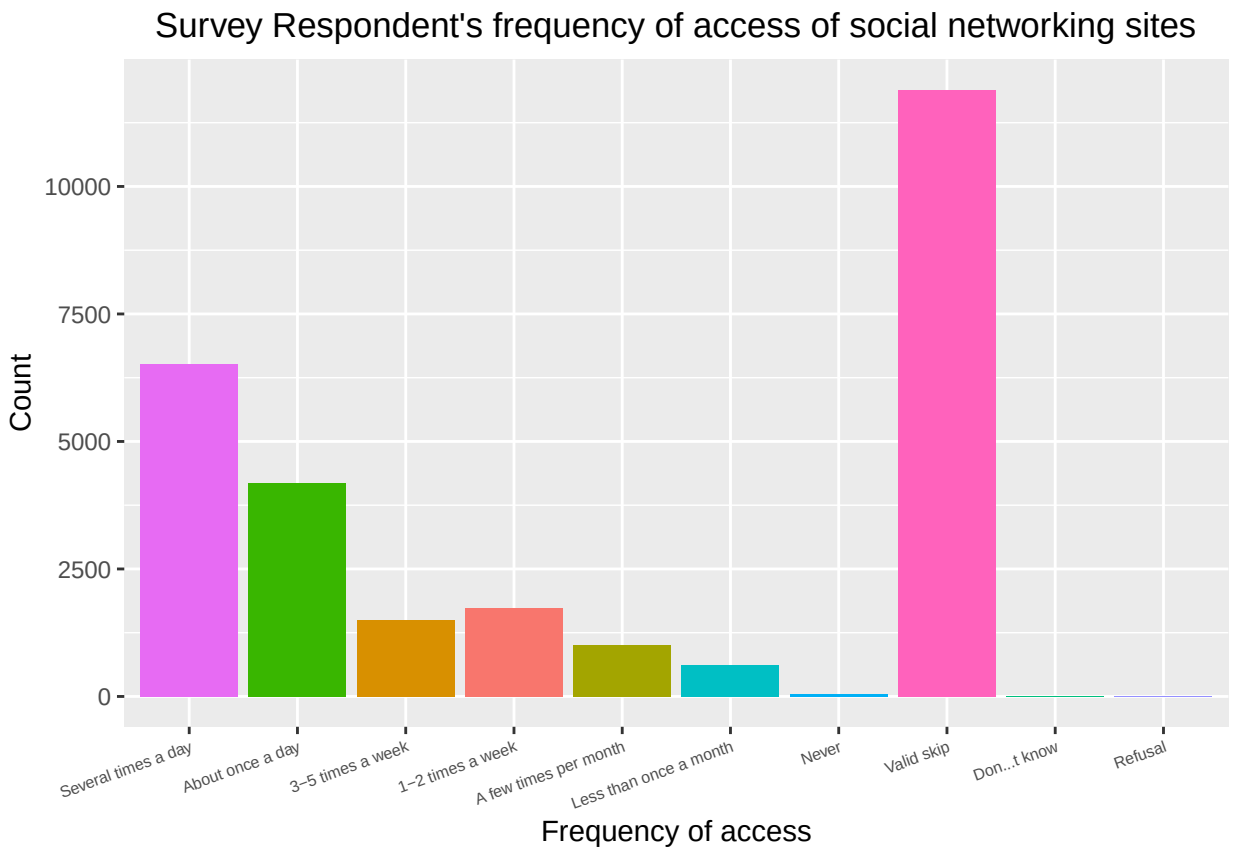


Figure 4: Distribution of survey Respondent's frequency of access of social networking sites.

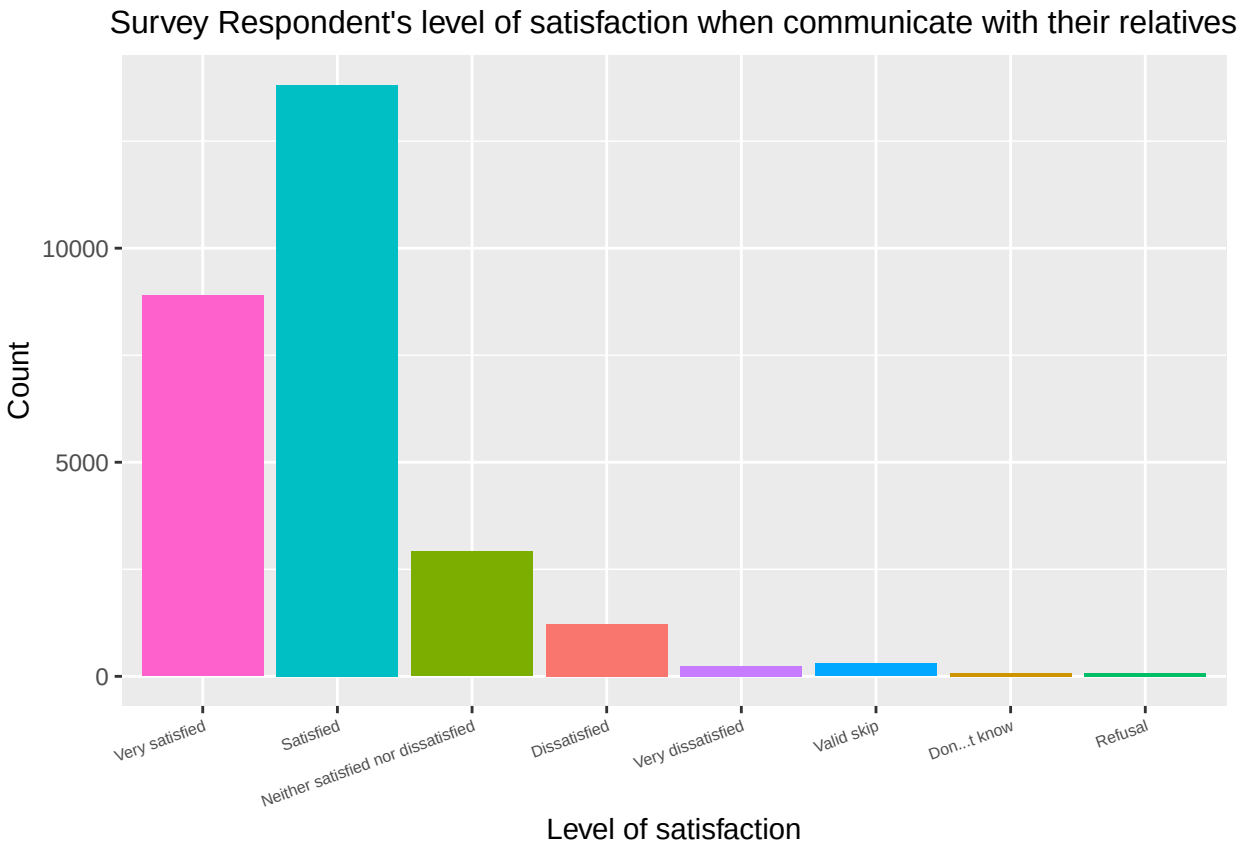


Figure 5: Distribution of survey Respondent's level of satisfaction when communicate with their relatives.

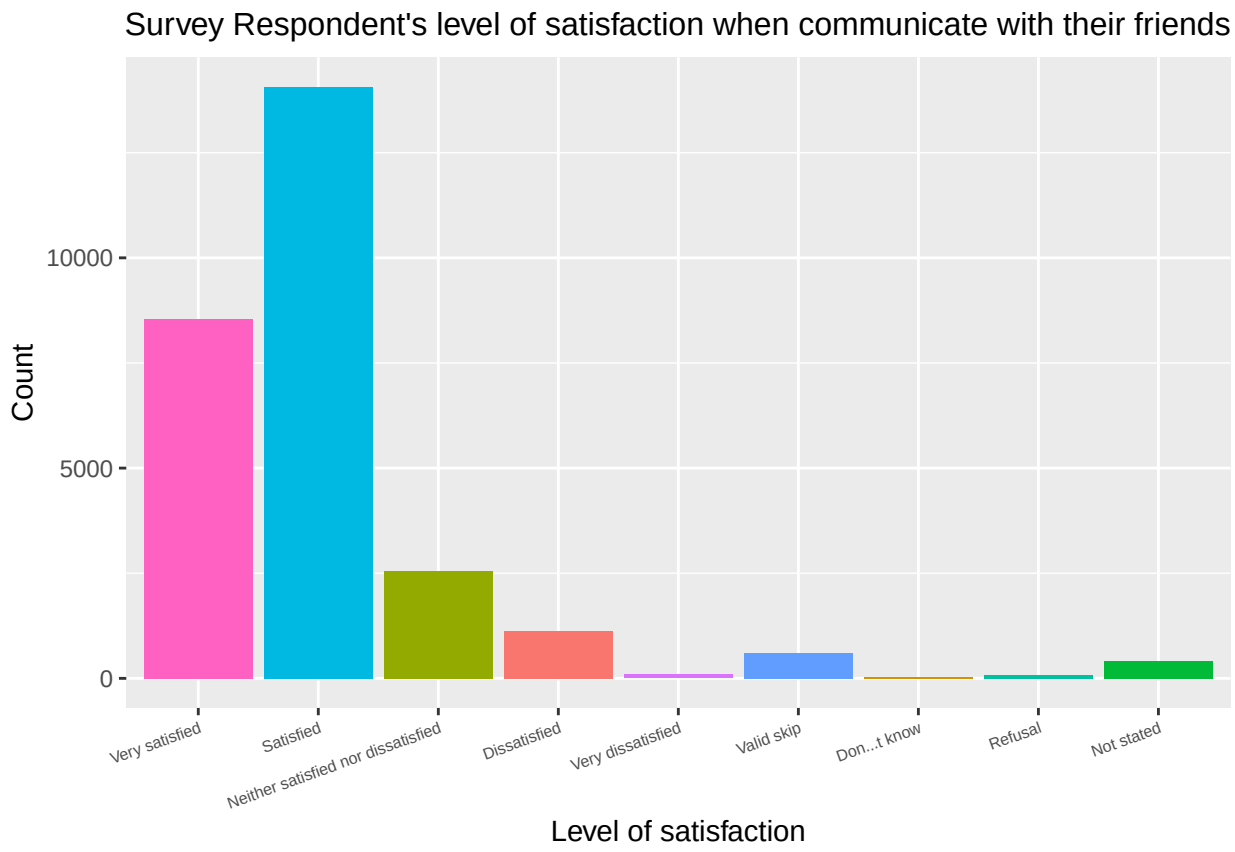


Figure 6: Distribution of survey Respondent's level of satisfaction when communicate with their friends.

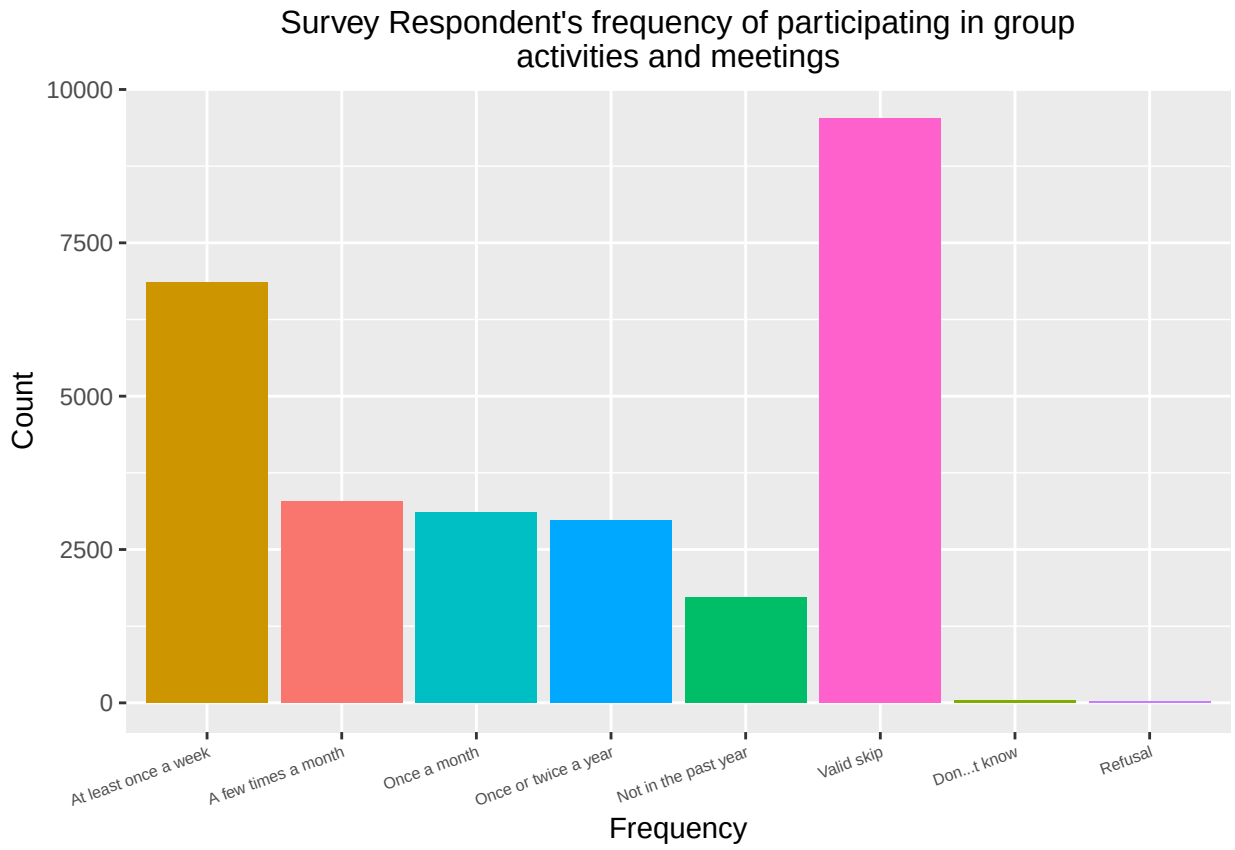


Figure 7: Distribution of survey Respondent's frequency of participating in group activities and meetings.

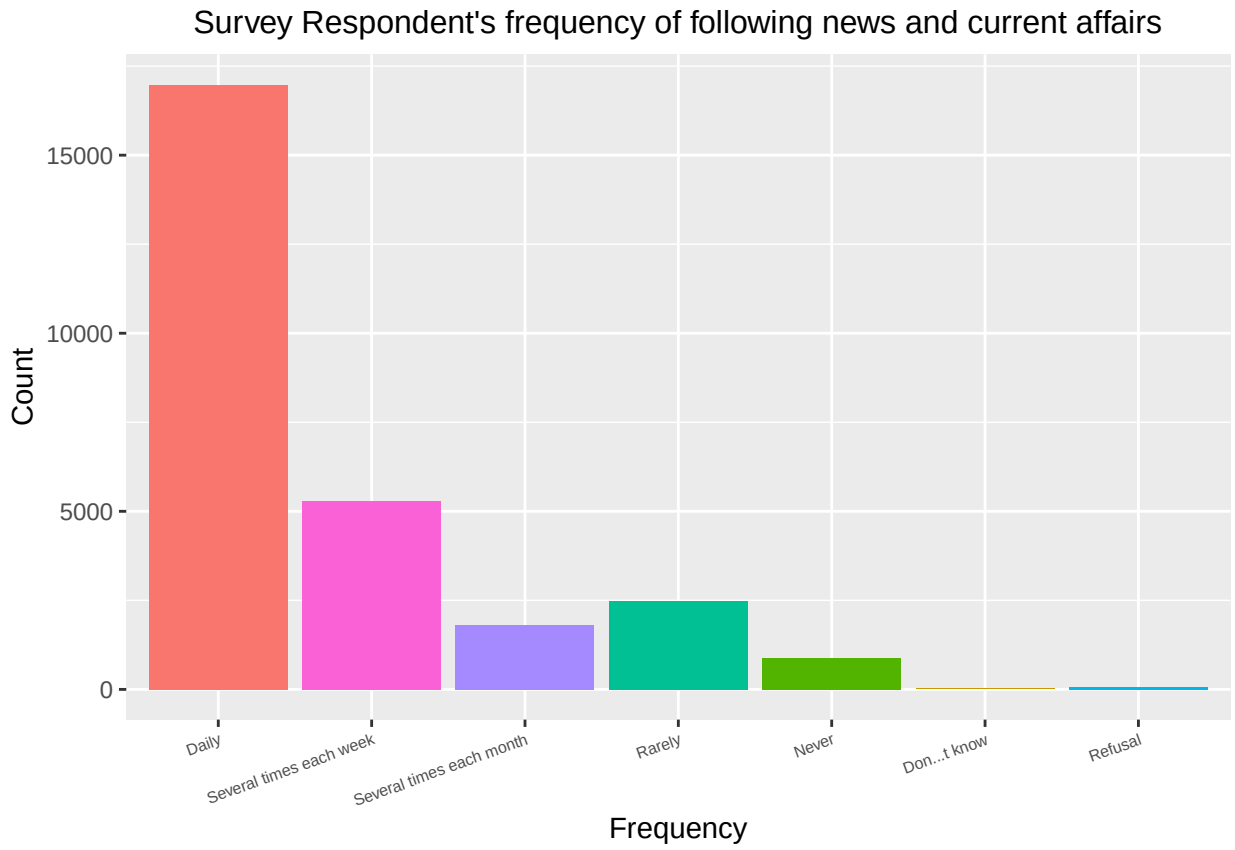


Figure 8: Distribution of survey Respondent's frequency of following news and current affairs.

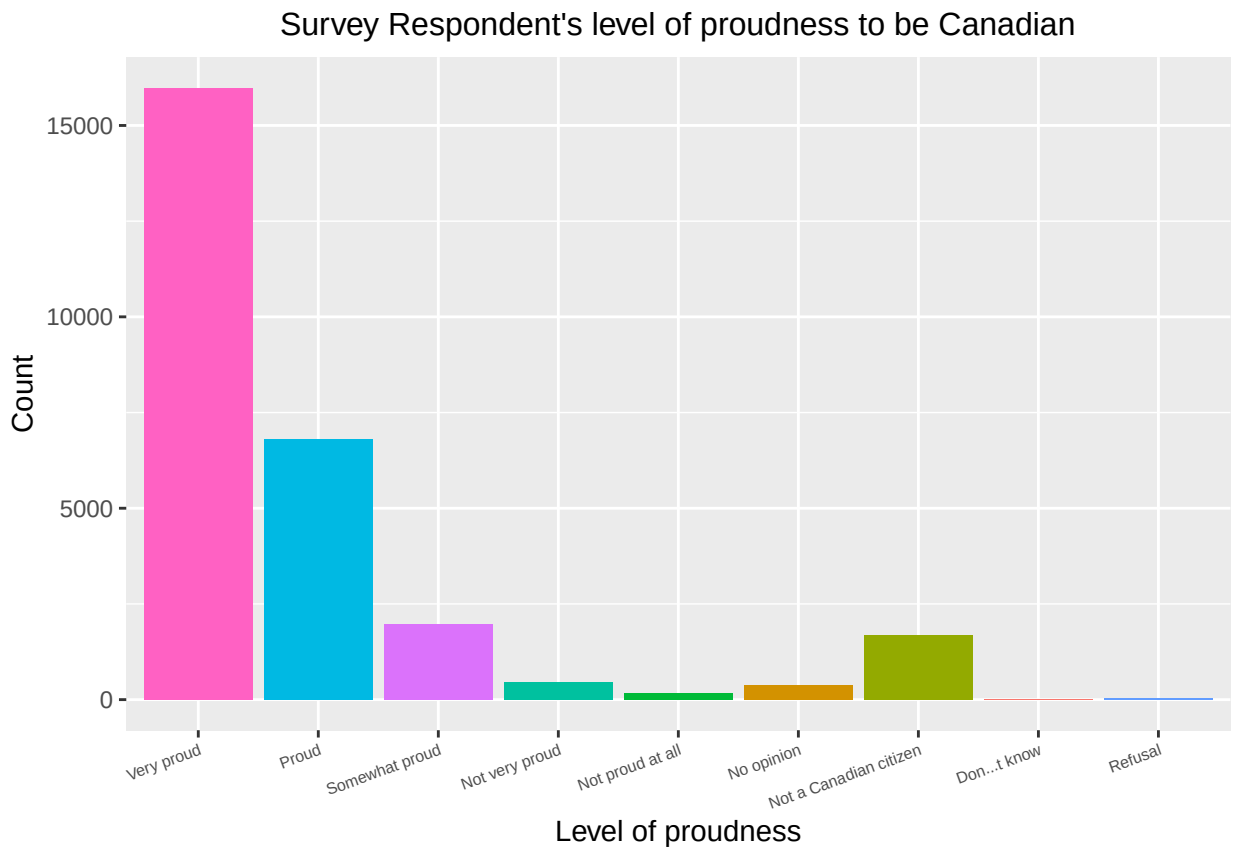


Figure 9: Distribution of survey Respondent's level of proudness to be Canadian.

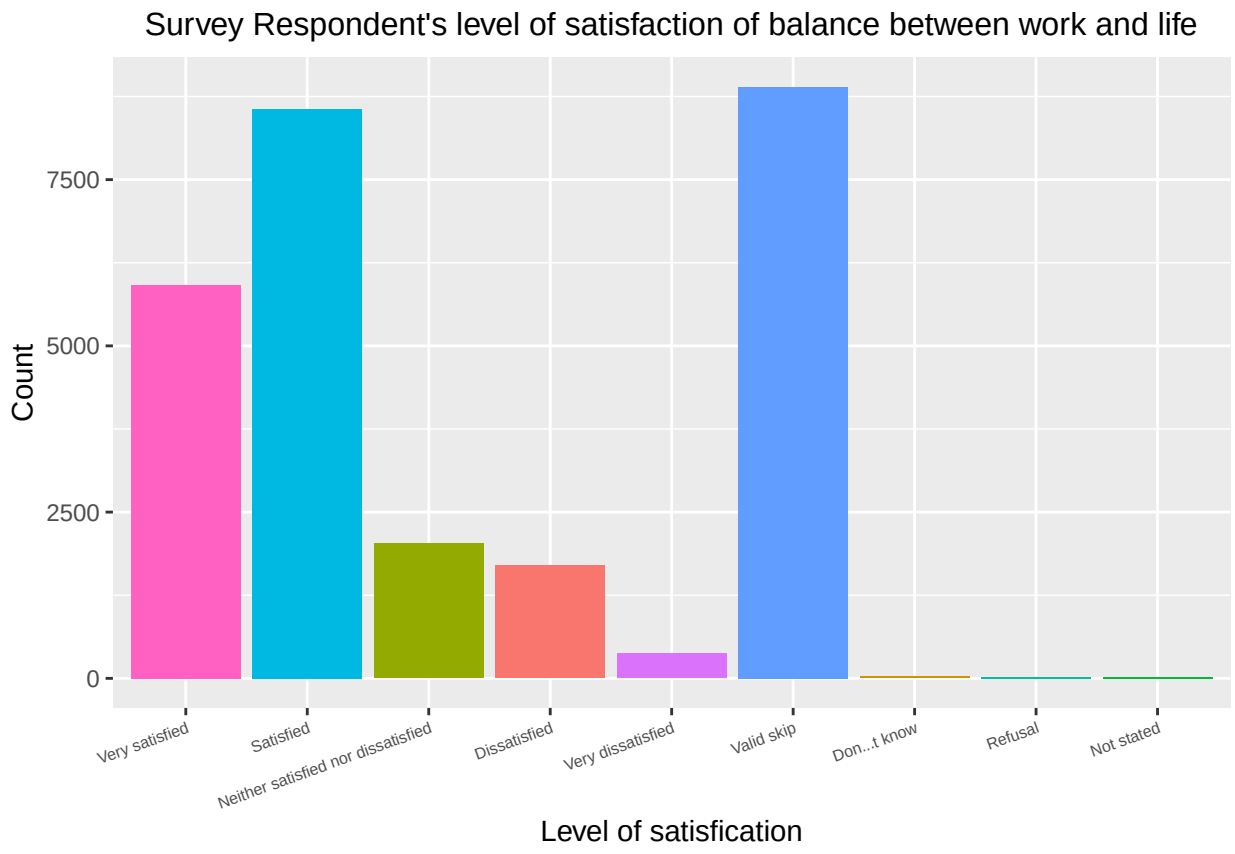


Figure 10: Distribution of survey Respondent's level of satisfaction of balance between work and life.

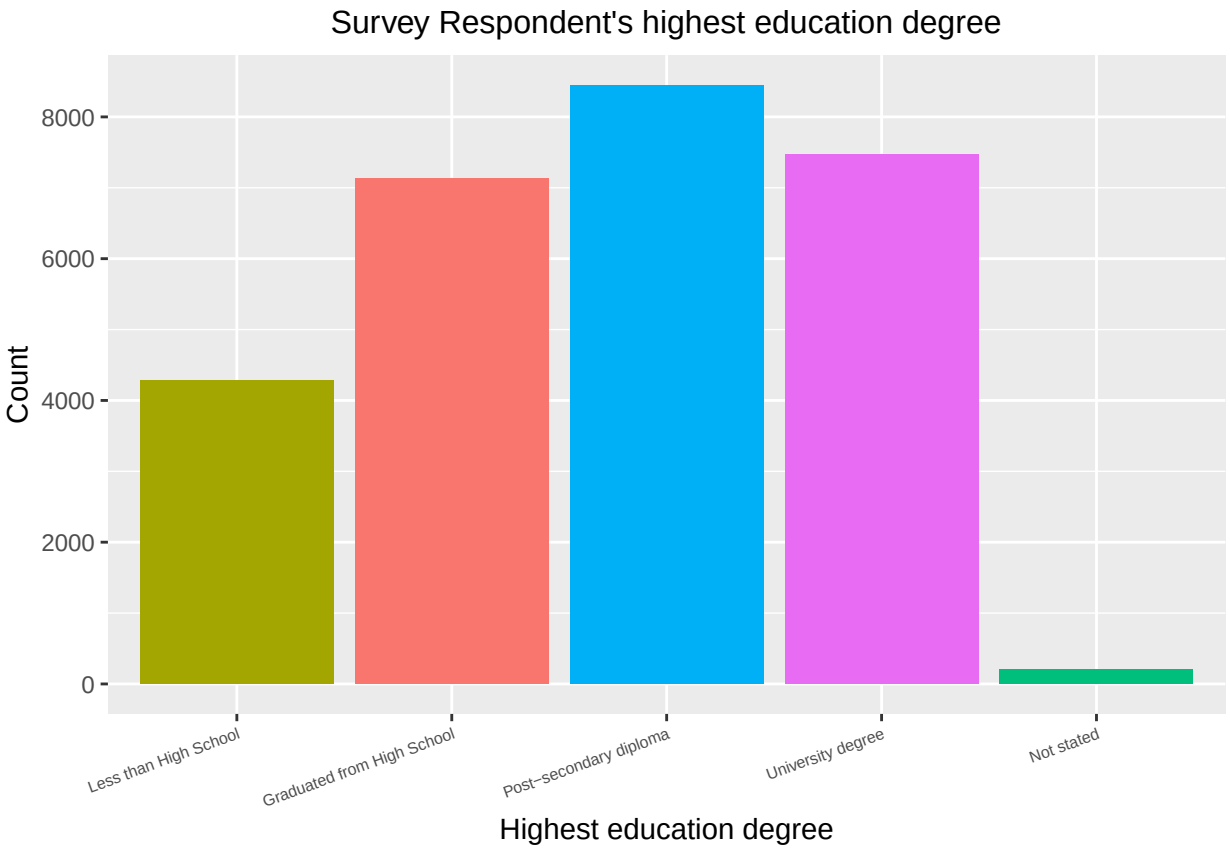


Figure 11: Distribution of survey Respondent's highest education degree.

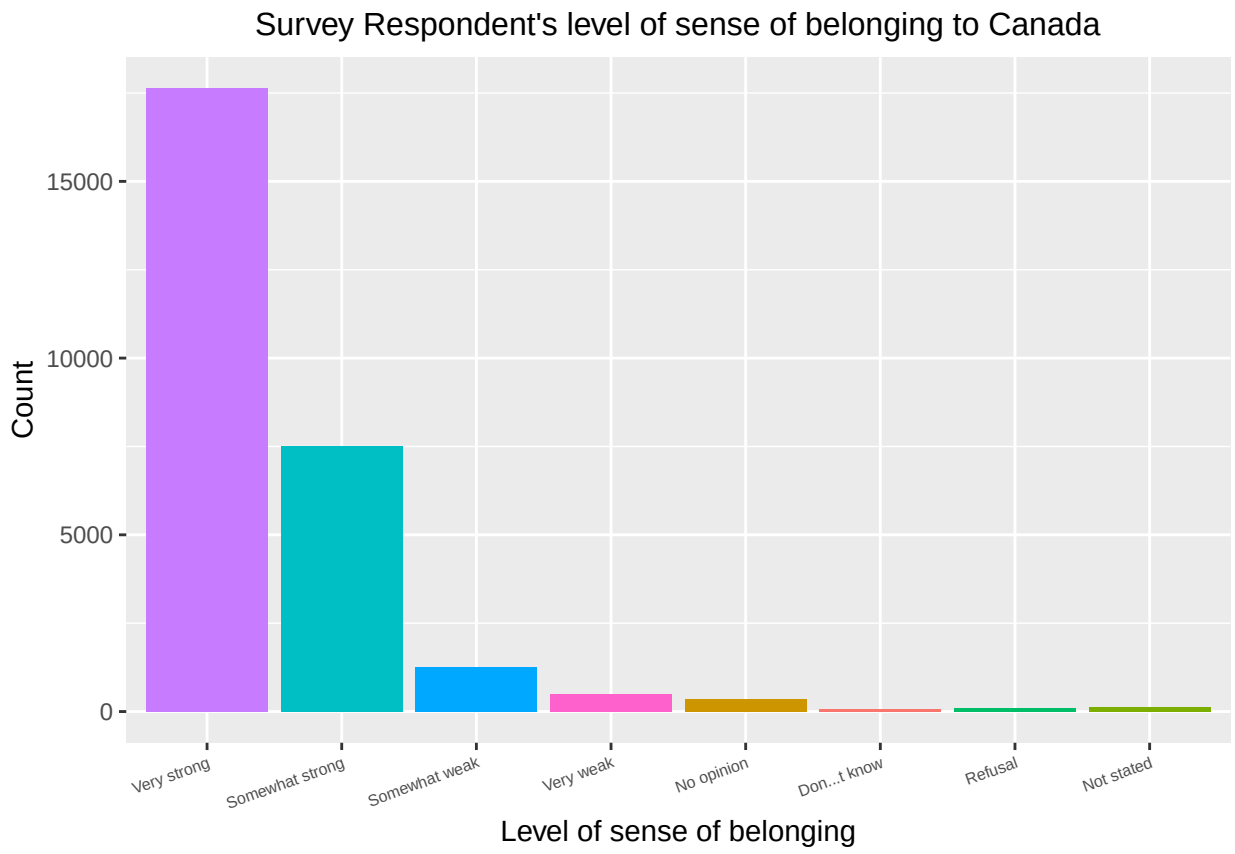


Figure 12: Distribution of survey Respondent's level of sense of belonging to Canada.

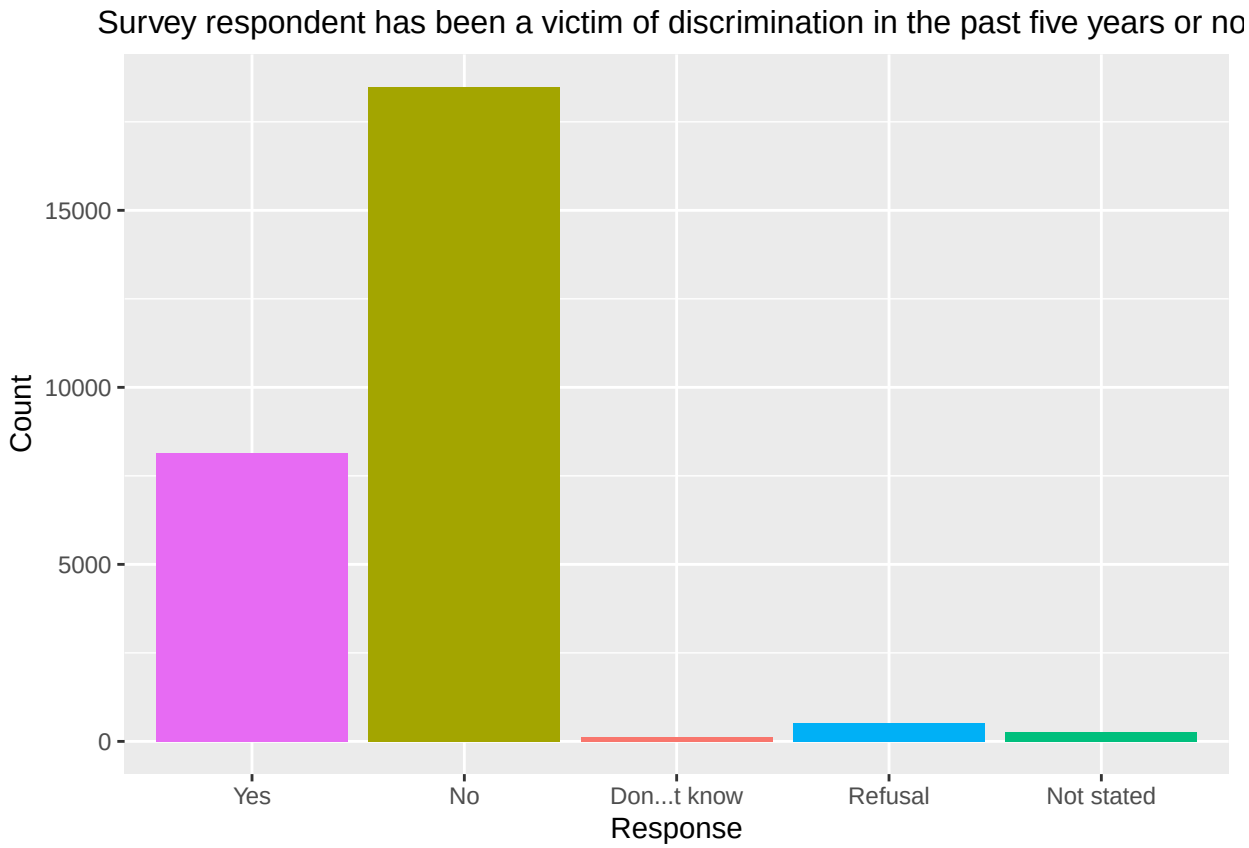


Figure 13: Distribution of survey respondent has been a victim of discrimination in the past five years or not.

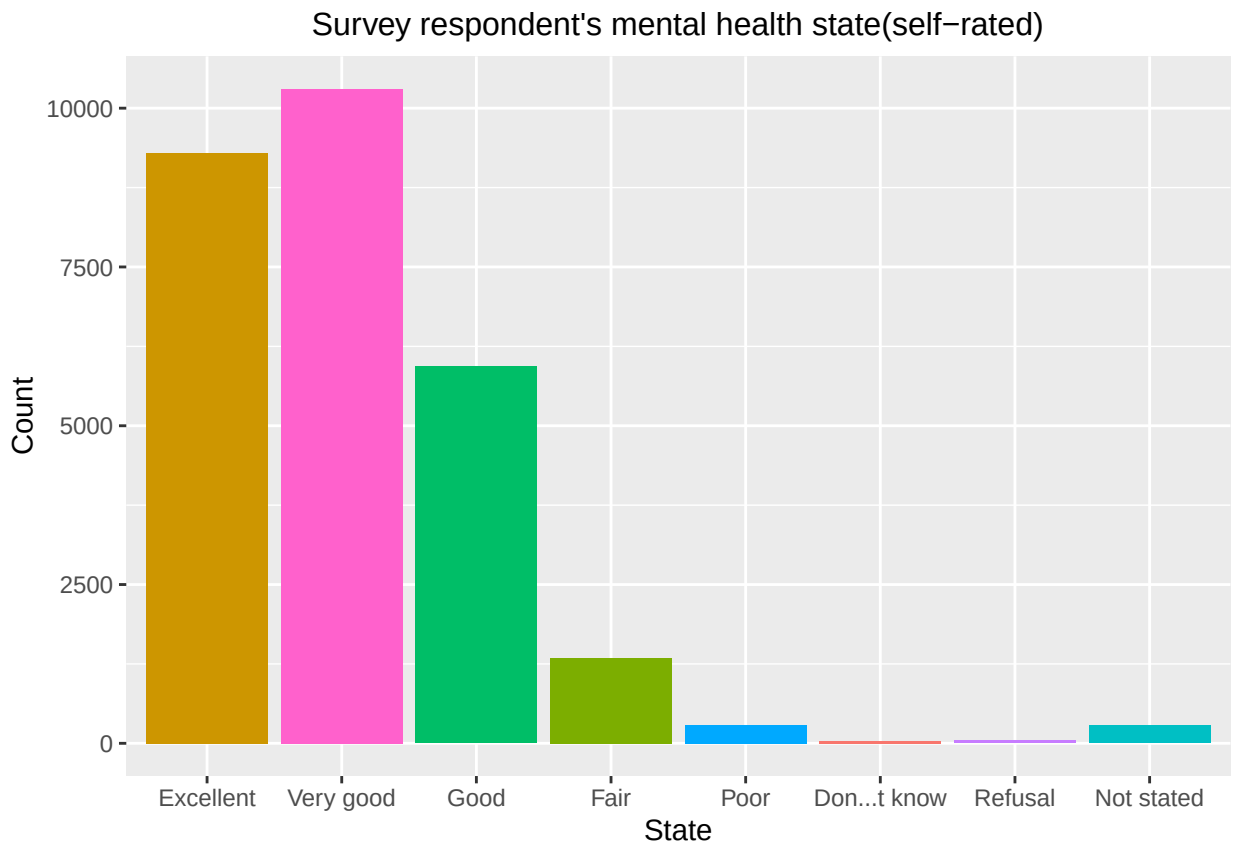


Figure 14: Distribution of survey respondent's mental health state(self-rated).

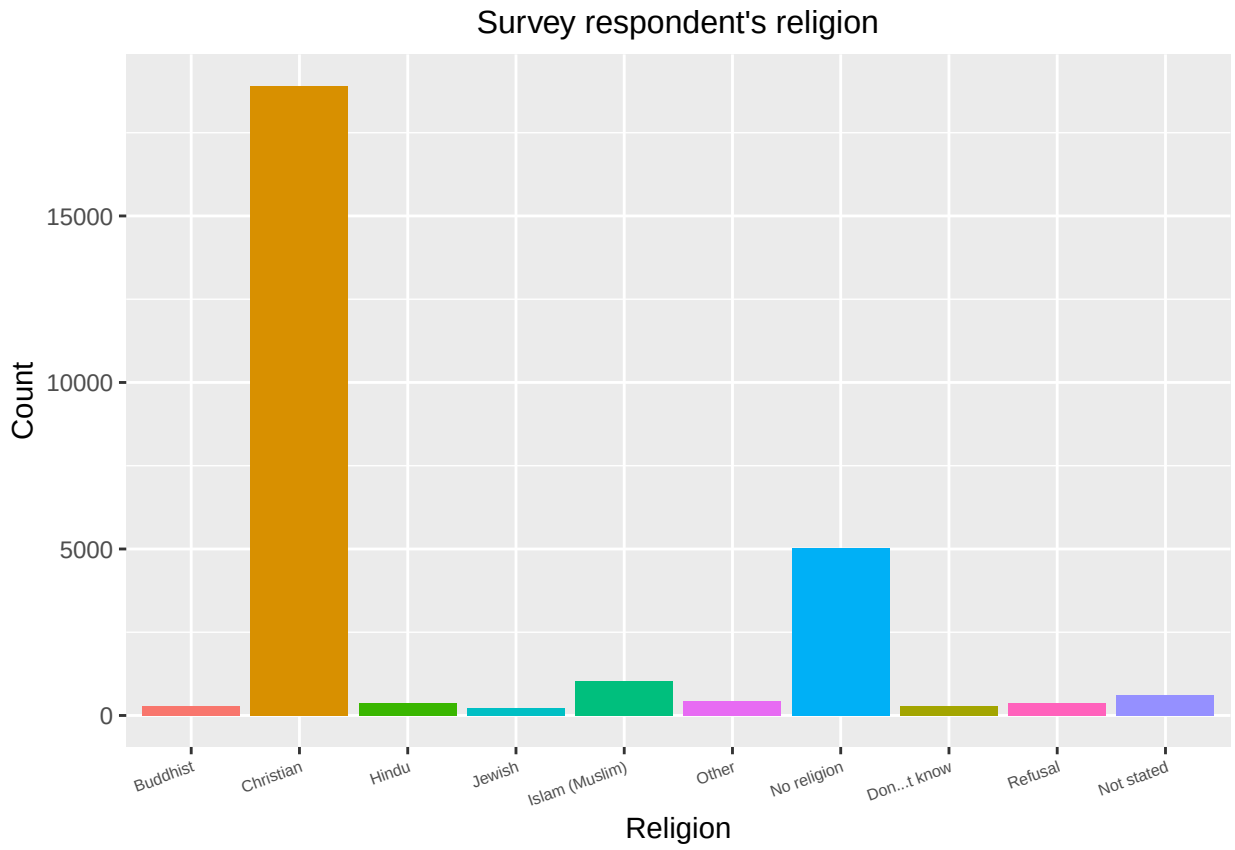


Figure 15: Distribution of survey respondent's religion.

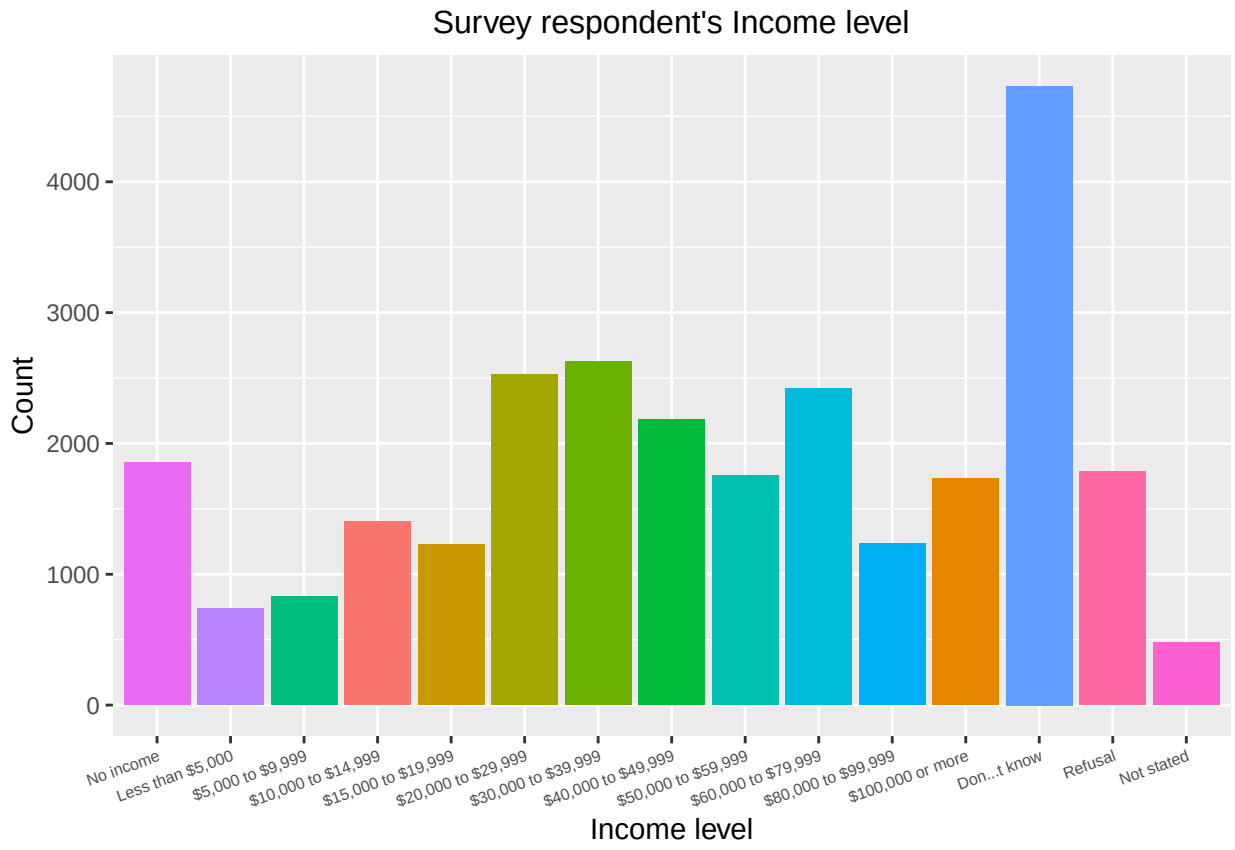


Figure 16: Distribution of survey respondent's Income level.

More specifically, we can take a closer look at the questions regarding the variables that we choose to focus on. For question asked about variable "GRP_40", the question is "[Including participation both on and off the Internet, how/How] often did you participate in group activities and meetings? [Do not include any of your volunteer activities.]". This is a great question because it clearly states that respondent should include activities both online and offline but not volunteer activities. For variable "MCR_10", the question being asked is "How frequently do you follow news and current affairs (e.g. international, national, regional or local)?". This question is not very clear and obvious for participant to answer since sometimes we follow news passively from the TV in the lobby of some buildings or other people, whether we consider "follow" to include unconscious and passive follows is a problem here. The question designed for variable "WFR_510" is "How satisfied [are/were] you with the balance between your job and home life? [Are/Were] you. . . ?". The way this question is asked seems to ignore the fact that some people might work as a housewife or househusband.

Result

Here in the result section, we have conducted logistic regression model to examine the influential variables towards participants' sense of belonging as Canadian. A starting logistic regression model called "model1" with the response variable "sbl_500_level" was created using the function `glm`(Fox and Weisberg 2019) to investigate the relationships between the response variables and the other selected variables. The function `vif`(Fox and Weisberg 2019) was used for checking the multicollinearity between the variables. If a Variance Inflation Factor(VIF) of a independent variable exceeding 5 indicates high multicollinearity between this independent variable and the others.(Bhandari 2020) By checking the VIFs, the variable "agegr10" which represents the age group of the survey respondents was found to have VIF around 7.32 which is larger than 5.

The second logistic regression model with the same response variable was then created by removing the variable "agegr10" from the staring model, the VIFs of the independent variables were all under 5 which means there is no serious multicollinearity problems in the second model. Thus the second model "model2" was chosen to be the final model of this analysis; the model summary was included in Table 2 in the appendix session. Table 2 was also created with function `kable`(Xie 2014), the function `tidy`(Robinson, Hayes, and Couch 2022) was used for converting the statistical objects in the model summary into tidy tibbles.

From the summary of the final model, we can see that there were 7 independent variables contain p-values less than 0.1: "cwr_40", "mcr_310", "prd_10", "wfr_510", "dh1ged", "discrim", and "incm". This would means that when our null hypothesis is "the variable has no correlation with the response variable", with the significance level being 0.1, we have evidence to reject the null hypothesis for these 7 variables.(Frost 2022) Thus we can say that, survey respondent's level of satisfaction when communicate with their relatives, survey respondent's frequency of following news and current affairs, survey respondent's level of proudness to be Canadian, survey respondent's level of satisfaction of balance between work and life, survey respondent's highest education degree, survey respondent has been a victim of discrimination in the past five years or not, survey respondent's income level are correlated with the survey respondent's level of sense of belonging to Canada.

Discussion

Conclusions

The paper has used the 15 chosen variables to examine their relationship with Canadian sense of belonging factor. After we created the bar plots and regression model, we are then able to discuss on the results that they have concluded.

From observations of bar plots, several factors showed lots of valid skip responses. Valid skip indicates that the question is confirmed to be "off path" by respondents because they clearly belong to another group. For instance, some questions such as accessing social networking sites may not be realistic for some participants, since in 2013 they might not have concrete means to visit internet. Also, almost 20000 of households obtained the religion of Christian, this may come from that most residents in Canada are white people, and other races' religion are not as common as Christian in the country.

After we conducted the logistic regression model, seven categorical variables were found to have effect on sense of belonging. They were respectively respondent's level of satisfaction when communicate with their relatives, survey respondent's frequency of following news and current affairs, survey respondent's level of proudness to be Canadian, survey respondent's level of satisfaction of balance between work and life, survey respondent's highest education degree, survey respondent has been a victim of discrimination in the past five years or not and survey respondent's income level. These factors have different level of influences, for instance, respondent's proudness and whether they have experienced discrimination have most obvious correlations. As the number of people not been through discrimination increased, their sense of belonging would increase remarkably. And when households felt proud to be Canadian, their sense of belonging would increase as well. What's more, if respondents were dissatisfied with their communication with relatives, their sense of belonging would decrease slightly. Also, if people were not satisfied with their balance between work and life, their sense of belonging would decrease as well. Notice that following news several times a week had negative

influence on sense of belonging too, this may due to seeing disturbing news and messages. The influence of households annual income and education level were not as obvious as others, yet the results still showed that when people were higher educated and gained higher income, they would feel more belonging.

In this case, in order to improve households' sense of belonging as Canadian, the discrimination in any kinds should be not allowed in the country. People should communicate more with their relatives and friends to create tight bound with cared ones.

Weakness and Areas for Furture

Some limitations are involved as we processing our analysis. For instance, the raw survey data has almost 1000 variables and it could be hard to correctly find the significant factors. We have looked through each variable's data and chosen 16 most representative factors to construct the research. Still, some of these survey questions can be very subjective and may cause potential bias. For example, when asking people about their satisfaction with communication with relatives and friends, some of them may lie about their actual feelings.

As we constructed the logistic regression, we had to eliminate the unclear answers (i.e., valid skip, don't know) to keep the categorical variables more specific and easier to model. Some variables such as social networking may have large amount of unclear answers, and this could cause some level of inaccuracy to our result.

In our future study, we could use our supplementary survey in appendix to obtain more concrete and up-to-date responses. What's more, we could learn more practical and complex regression models to conduct better results for massive and diverse data like this survey.

Appendix

Model summary table for model2

Table 2: Model summary table for model2

term	estimate	std.error	statistic	p.value
(Intercept)	4.3454584	0.8172890	5.3169175	0.0000001
as.factor(sex)2	0.2143026	0.1408641	1.5213430	0.1281738
as.factor(livarr06)2	-0.3211918	0.2242615	-1.4322198	0.1520809
as.factor(livarr06)3	-0.3218559	0.2092333	-1.5382634	0.1239842
as.factor(livarr06)4	-0.4234435	0.2981570	-1.4202031	0.1555486
as.factor(livarr06)5	-0.3299439	0.2720979	-1.2125925	0.2252856
as.factor(icr_30)2	0.1148921	0.1628578	0.7054746	0.4805149
as.factor(icr_30)3	0.2002132	0.2499498	0.8010136	0.4231238
as.factor(icr_30)4	0.2329811	0.2426696	0.9600752	0.3370174
as.factor(icr_30)5	-0.2179005	0.2549935	-0.8545336	0.3928094
as.factor(icr_30)6	0.4763524	0.4197129	1.1349480	0.2563971
as.factor(icr_30)7	-1.3408177	0.9469861	-1.4158789	0.1568109
as.factor(cwr_40)2	0.0086728	0.1800618	0.0481654	0.9615844
as.factor(cwr_40)3	-0.5736540	0.2230230	-2.5721742	0.0101062
as.factor(cwr_40)4	-0.1869284	0.3066692	-0.6095442	0.5421638
as.factor(cwr_40)5	-0.6371610	0.6043274	-1.0543309	0.2917314
as.factor(cwf_60)2	-0.0346909	0.1704181	-0.2035635	0.8386946
as.factor(cwf_60)3	0.3539515	0.2535593	1.3959319	0.1627350
as.factor(cwf_60)4	0.1326340	0.3076061	0.4311813	0.6663365
as.factor(cwf_60)5	0.7769811	1.1089255	0.7006612	0.4835145
as.factor(grp_40)2	0.2079118	0.1916385	1.0849165	0.2779587
as.factor(grp_40)3	0.0723657	0.2033256	0.3559105	0.7219076
as.factor(grp_40)4	0.0559951	0.1908642	0.2933765	0.7692344
as.factor(grp_40)5	-0.3312452	0.2192099	-1.5110869	0.1307663
as.factor(mcr_310)2	-0.3399373	0.1570915	-2.1639453	0.0304686
as.factor(mcr_310)3	-0.1890438	0.2317385	-0.8157634	0.4146355
as.factor(mcr_310)4	-0.1379001	0.2297190	-0.6002990	0.5483070
as.factor(mcr_310)5	0.4041527	0.5605260	0.7210241	0.4708947
as.factor(prd_10)2	-1.1327315	0.1732643	-6.5375947	0.0000000
as.factor(prd_10)3	-2.6207764	0.1795873	-14.5933319	0.0000000
as.factor(prd_10)4	-4.4973176	0.2477737	-18.1509072	0.0000000
as.factor(prd_10)5	-6.2214760	0.5300572	-11.7373677	0.0000000
as.factor(wfr_510)2	-0.0563801	0.1629610	-0.3459726	0.7293633
as.factor(wfr_510)3	0.1695566	0.2410516	0.7034035	0.4818043
as.factor(wfr_510)4	-0.4899613	0.2207526	-2.2195039	0.0264525
as.factor(wfr_510)5	-0.0948527	0.4476673	-0.2118821	0.8321990
as.factor(dh1ged)2	0.0139574	0.2705385	0.0515913	0.9588544
as.factor(dh1ged)3	0.1417910	0.2901755	0.4886386	0.6250976
as.factor(dh1ged)4	0.5668646	0.2988176	1.8970252	0.0578246
as.factor(discrim)2	0.5067717	0.1366684	3.7080393	0.0002089
as.factor(srh_115)2	-0.0156734	0.1627305	-0.0963150	0.9232704
as.factor(srh_115)3	-0.0059363	0.1881930	-0.0315439	0.9748358
as.factor(srh_115)4	-0.0102372	0.2987408	-0.0342679	0.9726635
as.factor(srh_115)5	-0.6737629	0.5781166	-1.1654447	0.2438390
as.factor(relig7)2	-0.9080747	0.6775717	-1.3401898	0.1801837
as.factor(relig7)3	13.4596524	407.1711765	0.0330565	0.9736295
as.factor(relig7)4	0.6005856	1.2399403	0.4843666	0.6281257

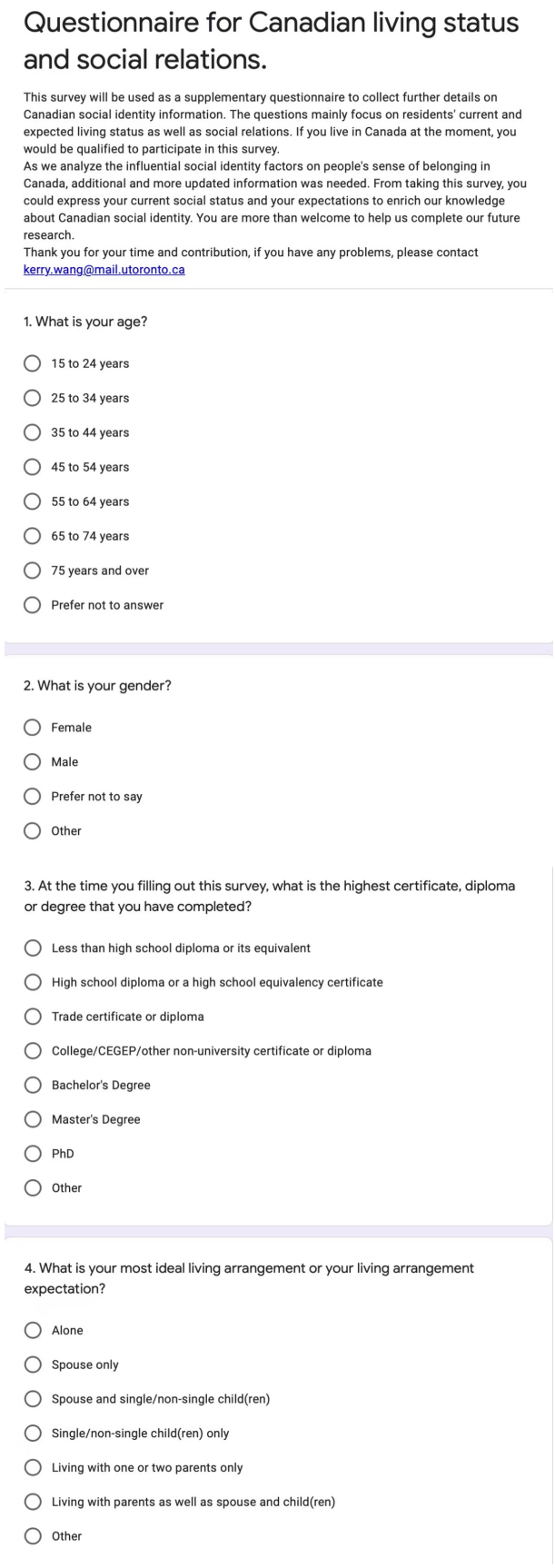
term	estimate	std.error	statistic	p.value
as.factor(relig7)5	-0.2626964	0.8206735	-0.3200986	0.7488936
as.factor(relig7)6	-0.6502254	0.8365460	-0.7772740	0.4369971
as.factor(relig7)7	-0.7544626	0.6830011	-1.1046286	0.2693206
as.factor(incm)2	0.4620133	0.3715647	1.2434263	0.2137108
as.factor(incm)3	0.0131714	0.3686491	0.0357287	0.9714987
as.factor(incm)4	0.2511871	0.3881320	0.6471692	0.5175225
as.factor(incm)5	0.0963203	0.4300395	0.2239801	0.8227728
as.factor(incm)6	0.4986115	0.3832462	1.3010214	0.1932511
as.factor(incm)7	0.0957636	0.3721709	0.2573108	0.7969389
as.factor(incm)8	0.2395010	0.3812658	0.6281733	0.5298904
as.factor(incm)9	0.3797961	0.3888955	0.9766020	0.3287663
as.factor(incm)10	0.6689584	0.3860213	1.7329574	0.0831032
as.factor(incm)11	0.3354238	0.4174932	0.8034235	0.4217300
as.factor(incm)12	0.5175363	0.4152193	1.2464170	0.2126114

Supplementary survey

Follow the link to our survey or scan our QR Code to survey: https://forms.gle/2mu9Gu3j7T98iW7L6

Figure 17: QR code for the survey

The following images show the screenshots of the online form.



Questionnaire for Canadian living status and social relations.

This survey will be used as a supplementary questionnaire to collect further details on Canadian social identity information. The questions mainly focus on residents' current and expected living status as well as social relations. If you live in Canada at the moment, you would be qualified to participate in this survey.
As we analyze the influential social identity factors on people's sense of belonging in Canada, additional and more updated information was needed. From taking this survey, you could express your current social status and your expectations to enrich our knowledge about Canadian social identity. You are more than welcome to help us complete our future research.
Thank you for your time and contribution, if you have any problems, please contact kerry.wang@mail.utoronto.ca

1. What is your age?

- ○ 15 to 24 years
- ○ 25 to 34 years
- ○ 35 to 44 years
- ○ 45 to 54 years
- ○ 55 to 64 years
- ○ 65 to 74 years
- ○ 75 years and over
- ○ Prefer not to answer

2. What is your gender?

- ○ Female
- ○ Male
- ○ Prefer not to say
- ○ Other

3. At the time you filling out this survey, what is the highest certificate, diploma or degree that you have completed?

- ○ Less than high school diploma or its equivalent
- ○ High school diploma or a high school equivalency certificate
- ○ Trade certificate or diploma
- ○ College/CEGEP/other non-university certificate or diploma
- ○ Bachelor's Degree
- ○ Master's Degree
- ○ PhD
- ○ Other

4. What is your most ideal living arrangement or your living arrangement expectation?

- ○ Alone
- ○ Spouse only
- ○ Spouse and single/non-single child(ren)
- ○ Single/non-single child(ren) only
- ○ Living with one or two parents only
- ○ Living with parents as well as spouse and child(ren)
- ○ Other

Figure 18: Screenshot of survey questions. Q1-Q4

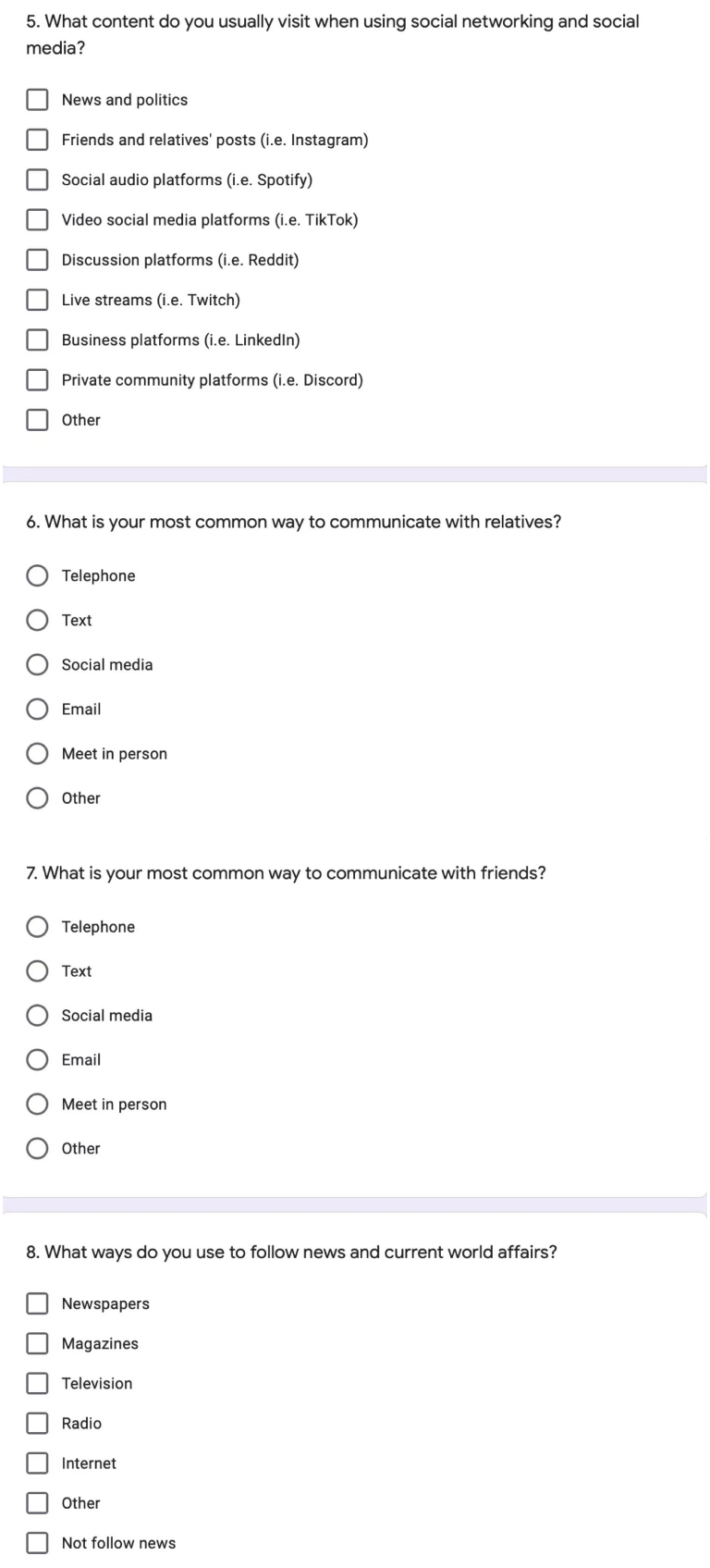


Figure 19: Screenshot of survey questions. Q5-Q8

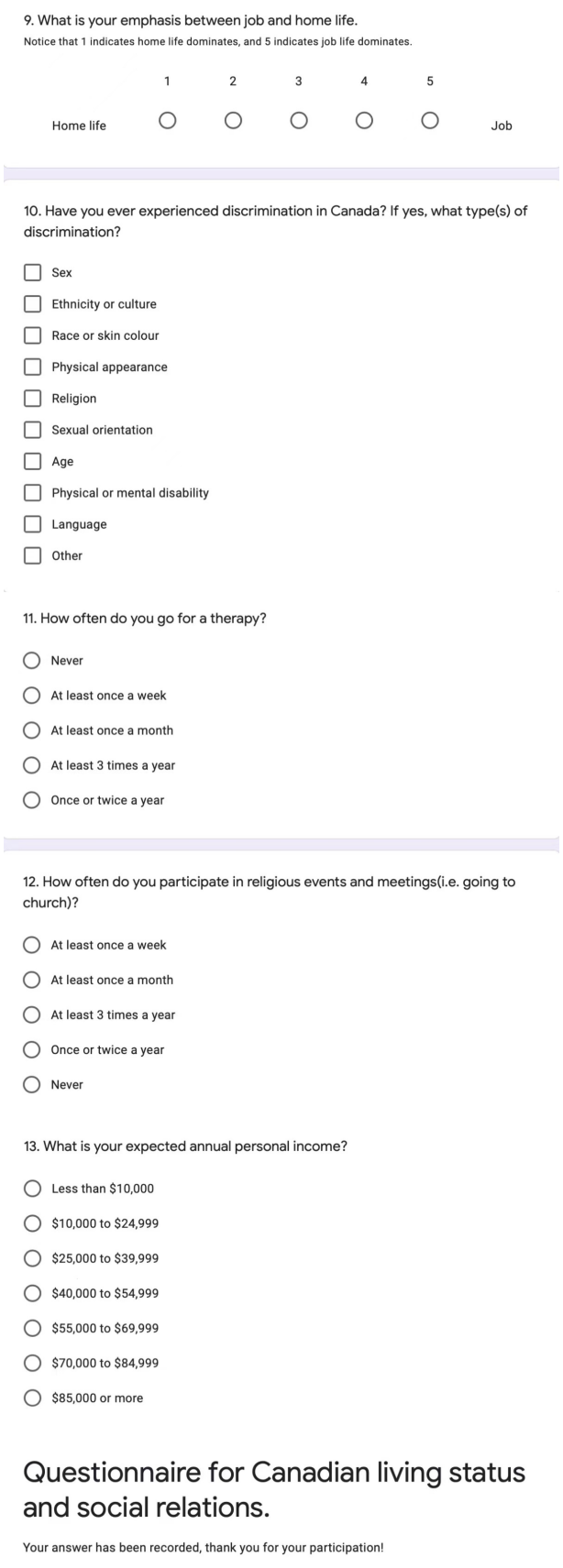


Figure 20: Screenshot of survey questions. Q9-Q13

References

Bhandari, Aniruddha. 2020. *Multicollinearity: Detecting Multicollinearity with VIF*. https://www.analyticsvidhya.com/blog/2020/03/what-is-multicollinearity/.

Cherry, Kendra. 2021. *What Is the Sense of Belonging?* https://www.verywellmind.com/what-is-the-need-to-belong-2795393#:~:text=A%20sense%20of%20belonging%20involves,same%20attention%20to%20other%20members.

Fox, John, and Sanford Weisberg. 2019. *An R Companion to Applied Regression*. Third. Thousand Oaks CA: Sage. https://socialsciences.mcmaster.ca/jfox/Books/Companion/.

Frost, Jim. 2022. *How to Interpret p-Values and Coefficients in Regression Analysis*. https://statisticsbyjim.com/regression/interpret-coefficients-p-values-regression/.

GSS, Cycle 27: Social Identity. 2013. https://sda-artsci-utoronto-ca.myaccess.library.utoronto.ca/sdaweb/dli2/gss/gss27v2/more_doc/en_index.htm.

Painter, Carla Valle. 2013. *Sense of Belonging: Literature Review*. https://www.canada.ca/en/immigration-refugees-citizenship/corporate/reports-statistics/research/sense-belonging-literature-review.html.

R Core Team. 2021. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. https://www.R-project.org/.

Robinson, David, Alex Hayes, and Simon Couch. 2022. *Broom: Convert Statistical Objects into Tidy Tibbles*. https://CRAN.R-project.org/package=broom.

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 François, Garrett Grolemund, et al. 2019. "Welcome to the tidyverse." *Journal of Open Source Software* 4 (43): 1686. https://doi.org/10.21105/joss.01686.

Xie, Yihui. 2014. "Knitr: A Comprehensive Tool for Reproducible Research in R." In *Implementing Reproducible Computational Research*, edited by Victoria Stodden, Friedrich Leisch, and Roger D. Peng. Chapman; Hall/CRC. http://www.crcpress.com/product/isbn/9781466561595.