Canadian GSS 2017: Factors that Impact Happiness*

Anna Li and Mohammad Sardar Sheikh

20 March 2022

Abstract

Happiness is a measure of the quality of life, the desired state that all aim to achieve. Yet, it is unclear what are the true factors impact happiness, both positively and negatively. Particularly, factors that determine happiness may vary country by country due to different cultures and policies. This paper looks at 2017 Canadians, specifically how income, education, health, family, and marital status impact their happiness. Income, health and marriage provide a positive impact on how happy a person is, therefore being the areas that should be focused on to improve life quality on both the individual and government level.

1 Introduction

Happiness can be considered a measurement of life quality. The main goal of life can be summarized as achieving happiness, whether it is through higher income, education, great social interaction, or more. However, there are many factors that people deem to be crucial for a successful, happy life that is not necessarily related to actually achieving the end goal. Parents with children are found to be less happy than single individuals "Parenthood and Happiness: Effects of Work-Family Reconciliation Policies in 22 OECD Countries" (2016), despite the fact that society's impression of an ideal life is constructing a family with a certain amount of children. Finding out what are the actual factors that are related and can increase happiness is crucial for Canada to increase the life quality of their citizens. It is also important for individuals, as they set goals for their future and visualize what it means to be successful.

The paper focuses upon five factors and their relationship with happiness: education, income, health, marital status, and family size. This was chosen in consideration of literature and also what seems to be an ideal life: Healthy, rich, highly educated, married and living happily with children. Literature has revealed a relationship of each factor with education, with varying conclusions of whether it is a positive or negative relationship. An increase in individual income increases happiness, although the effect is rather temporal Latif et al. (2015). Marriage is shown to be correlated with higher happiness as well Stack and Eshleman (1998). However, having children seems to lead to a decrease in marital happiness WHITE, BOOTH, and EDWARDS (1986). Those who are happy also are shown to live a longer life and have better health in general Siahpush, Spittal, and Singh (2008). Education can lead to higher happiness, but only after people reach 30; before then, education seems to play a negative role Nikolaev and Rusakov (2016).

Studies conducted have shown the relationship between particular variables and happiness individually, as well as conducting large studies for happiness and all potentially affecting variables. This report fills the gap in the literature in two ways. First, it focuses on Canada in 2017, pre-pandemic. This would lead to the assumption that it could be that family and social interaction are not as highly valued as they would be currently in 2022, after years of pandemic and social distancing. This provides insight into how Canada specifically views the various factors and how they can impact happiness. Moreover, it provides a niche understanding of how family income and the desired family size can impact happiness.

This report utilizes a combination of graphs, summary statistics, and regression to find the relationship between happiness and the various variables. The paper finds that education is negatively related to happiness

---

*Code and data are available at https://github.com/annadlli/canada-gss-happiness.git.

when income is held constant. Income, health, total children desired, and marriage is positively related to happiness. Therefore to improve quality of life, income, mental and physical health, and a good romantic relationship should be prioritized. Education can be considered a means to achieve income, which leads to higher levels of happiness. Total children desired reflects the desire of family, rather than the actual burden that having a child might bring that lowers happiness. Therefore, while these two factors do show a relationship with happiness, they should not be the priority when trying to find the major factors that impact happiness. To truly improve quality of life, a correct emphasis on the related factors to happiness is essential for effective policy and desirable results.

The remainder of the paper is split into five sections. Section 2 explains the data source and collection methodology, potential bias and ethics issues, as well as our selected data's characteristics. Section 3 discusses the methods used to produce the wanted results to investigate the relationship between happiness and the various factors. Section 4 presents the findings from our methodology. Section 5 discusses what the findings in results mean, the weaknesses and limitations in interpreting the findings of the paper, as well as next steps to enhance and expand on the findings of the report. Section 5.5.2 includes testing assumptions for regression, alternative graphs of the same findings, and a supplementary survey that could improve and add to the study conducted in this report.

2   Data

This report was created using the R statistical programming language R Core Team (2020). Analysis and figures were created with the packages ggplot2 Wickham (2016), knitr Xie (2021), kableExtra, Zhu (2021), jtools Long (2020), gtsummary, Sjoberg et al. (2021), patchwork Pedersen (2020), and tidyverse Wickham et al. (2019). Package car Fox and Weisberg (2019) was utilized in the appendix to check the regression model. Data was obtained from CHASS, and cleaned using code provided by professor Rohan Alexander Alexander (2022).

2.1   Data Source and Collection

This report uses 6476 observations from the 2017 Canadian General Social Survey Microdata. The data was obtained from the University of Toronto Computing in the Humanities and Social Sciences (CHASS) Database Centre, available to licensed University of Toronto users Uoftlibraries (n.d.). CHASS utilizes Berkeley's software Survey Documentation and Analysis (SDA) to provide access to the microdata files.

2.1.1   Canadian GSS

The Canadian General Social Survey (GSS) is a voluntary survey collected every year, with the theme that the data is being collected for changing between topics. In total, there are seven different themes, and these include life at work and home; families; caregiving and care receiving; giving, volunteering and participating; victimization; social identity and time use Canada (2019). The data that we are using is regarding families. Each theme is explored every five to seven years. The survey collects data from people aged 15 and older from all the provinces of Canada. The provinces are divided into strata. A simple random sample was then performed in each stratum. The target sample size is 20,000 respondents, as this number was deemed sufficient by Statistics Canada to capture the various demographic aspects of the population. The sampling frame defines households that have a land-line or a cell phone registered at the same address, and these are chosen randomly. The data collection primarily takes place via computer-assisted telephone interviews. The interviewers were instructed to extract all the required information and to contact the respondent again if the interview could not take place or if it was not able to be completed at the time. Data for income is gathered through the respondents' tax records (after asking for consent). Data has not been readily collected for indigenous populations, so it is often aggregated total of the First Nations people, Metis, and Inuit to inflate the total indigenous population counts. Ever since the GSS has switched to more modern methods to collect data in 2013, the response rate has dropped, and it has been harder to collect reliable data from respondents.

2.1.2   Strengths of Canadian GSS

For each variable, there was a great deal of variation in the choices given to the respondents. For example, many questions offer five different choices as well as the additional option of "Don't Know" to account for uncertainty. This helps us get more accurate information from the respondents regarding their situation. Furthermore, if the respondents did not give accurate information regarding a question using the online questionnaire, they would be contacted using a telephone interview to get the relevant information from them. Another strength the survey collection method had was that it had a proportionate sample from all the different provinces of Canada, with special care taken to include the aboriginal population in the survey as well. As such, the results can be used to generalize to the entire population.

2.1.3   Potential Ethical and Bias Issues with Survey Methodology

The Canadian General Social Survey has a few issues that should be considered. First is the existence of non-sampling errors. Non-sampling errors occurred due to participants choosing not to respond to the survey (as it is a voluntary survey) or did not give accurate responses to the questions. This can happen when the respondents do not understand what a question implies and hence gives a "wrong," undesired answer in terms of the purpose of the survey.

There is also the potential for sampling errors to occur as well. Sampling errors happen when the sample does not accurately represent the population. For this survey, 68 out of 100 times the difference between the sample estimate and the population parameter was less than one standard error Canada (2020). As the survey data was used to generalize findings to the entire population of Canada, non-responses for questions and replies such as 'don't know' poses potential problems, as this assumes that the proportion of non-responses would be similar in the entire population as well.

Ethically, there are two factors that must be approached cautiously in the survey - race and gender. For gender, the respondents were only given two options, male or female. They were also given the option to use "don't know" or "refuse to answer," but this likely groups together a lot of the other gender groups that exist. Using gender as a variable could lead to interpretations not accurately capturing the current demographic where people can identify as genders not restricted to the binary. Moreover, this question might demotivate respondents to give further information as they may believe that they will not be accurately represented in the population. Another issue to be cautious about is the lack of information from the indigenous population. The responses received are likely aggregated and estimations to represent the population better instead of being actual responses from people. This limits the applicability of results to the indigenous population, particularly to those who are not a part of the three indigenous groups the GSS sampled.

2.1.4   Questionnaire Details

A few questions have very subjective choices. When a candidate reports their mental/physical health, they are given the choices: Don't know, poor, fair, good, very good and excellent. The responses depend on the respondents' own interpretation of the question as well as each choice. One person's interpretation of "fair" may differ greatly from another's understanding. Therefore, it is likely that people's subjective understanding of the questions and options can cause our results for such factors to be biased. These questions should instead be structured in a more objective way that limits subjective interpretability, perhaps giving the candidate a numeric scale instead of a categorical choice.

Another issue with the questionnaire is with the question asking respondents about their feelings for life, which is interpreted by our report as happiness. The specific wording of the question is "what are your feelings for life as a whole." The wording of the question can be a bit unclear and also asks for respondents' feelings over their entire life. This extreme long time horizon can cause respondents to have difficulty coming up with their actual feelings on the spot. Perhaps it would be better to ask about the respondents' feelings on a number of different time horizons, such as the past month, past year, and entire life. This could provide more insight into how certain world events could be affecting their experiences, as well as the relationship of the investigated variable's impact on happiness.

2.1.5 Variable Selection

12 variables were selected to conduct analysis for our report. Some were chosen to investigate the selected data set's characteristics, seeing whether it generally matches what is expected for a representative data set about Canada, such as visible minority. Others were chosen to investigate their relationship and effect on happiness. For our study, we wanted to focus on 5 different factors that supposedly impact happiness: education, income, health, marital status, and family.

Happiness was measured using the survey question asking respondents how satisfied they were in life. We assumed that a respondent that was satisfied with their life was equivalent to being happy. The income of both the family of the respondent and the respondent themselves was measured. Income asked for the before tax CAD income, with a total of 6 categories. Self-rated health and self-rated mental health were used to gain an understanding of the health situation of respondents. This was measured by a simple scale of "poor, fair, good, very good, and excellent." We assumed the total number of children desired could provide a glimpse of how large of family the respondents desire and would like to obtain. The total children desired asked how many children the respondent wanted in total, including any children that they already have. Therefore, we deemed it representative of our investigation of family.

There were no clear outliers within the observations that were eliminated.

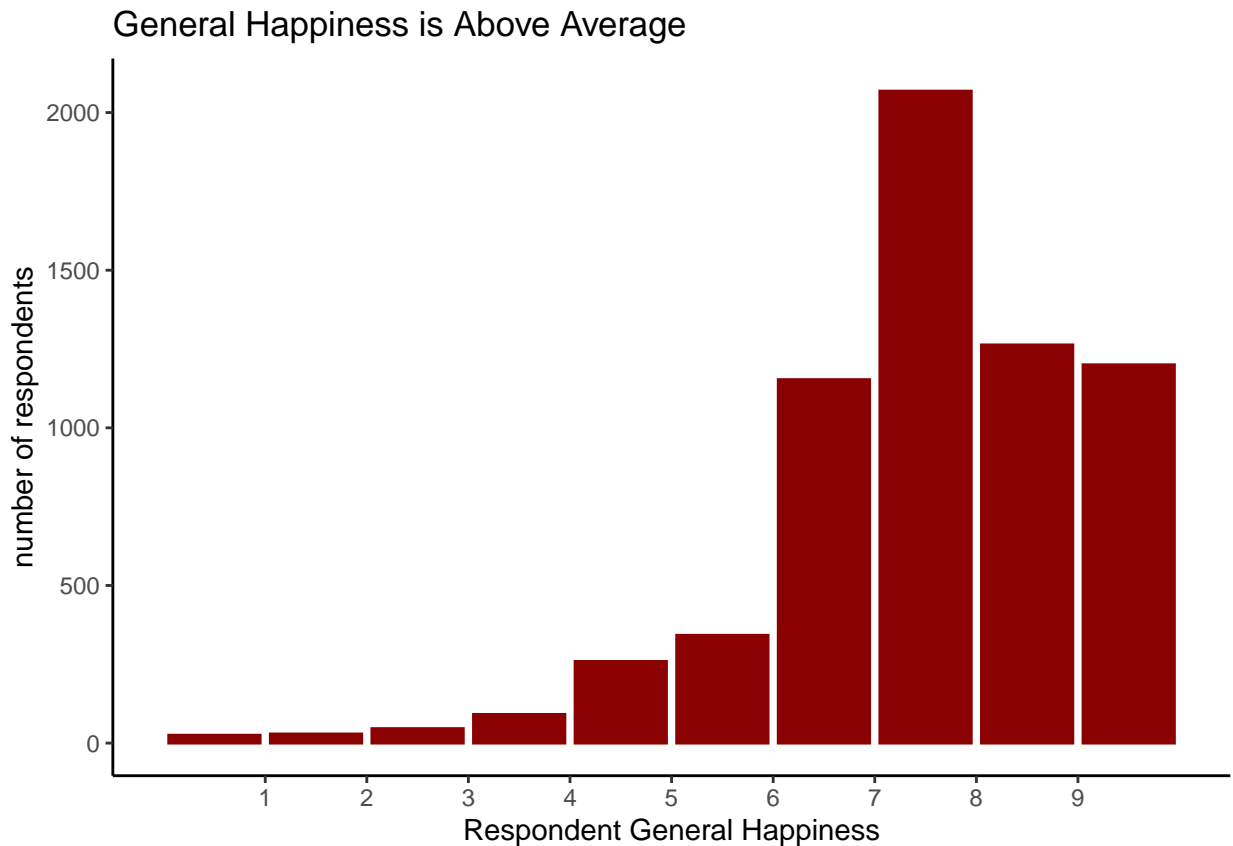## 2.2 Happiness: People are Abover Average Satisfied with Life

Figure 1: Feelings of life as a whole is above average satisfied.

Figure 1 is created through ggplot(Wickham (2016)). There are only a few responses overall in proportion to the total sample that feel unhappy in life, as demonstrated by the left-skewed distribution and few responses to the left of 5, the midpoint. The majority of the response is concentrated around 7-10, which can be

interpreted as feeling above average about how they are doing in life. It is interesting that the most common answer is 7. While this means the respondents are feeling above average, this also can indicate they are not fully satisfied and have space to improve their happiness.

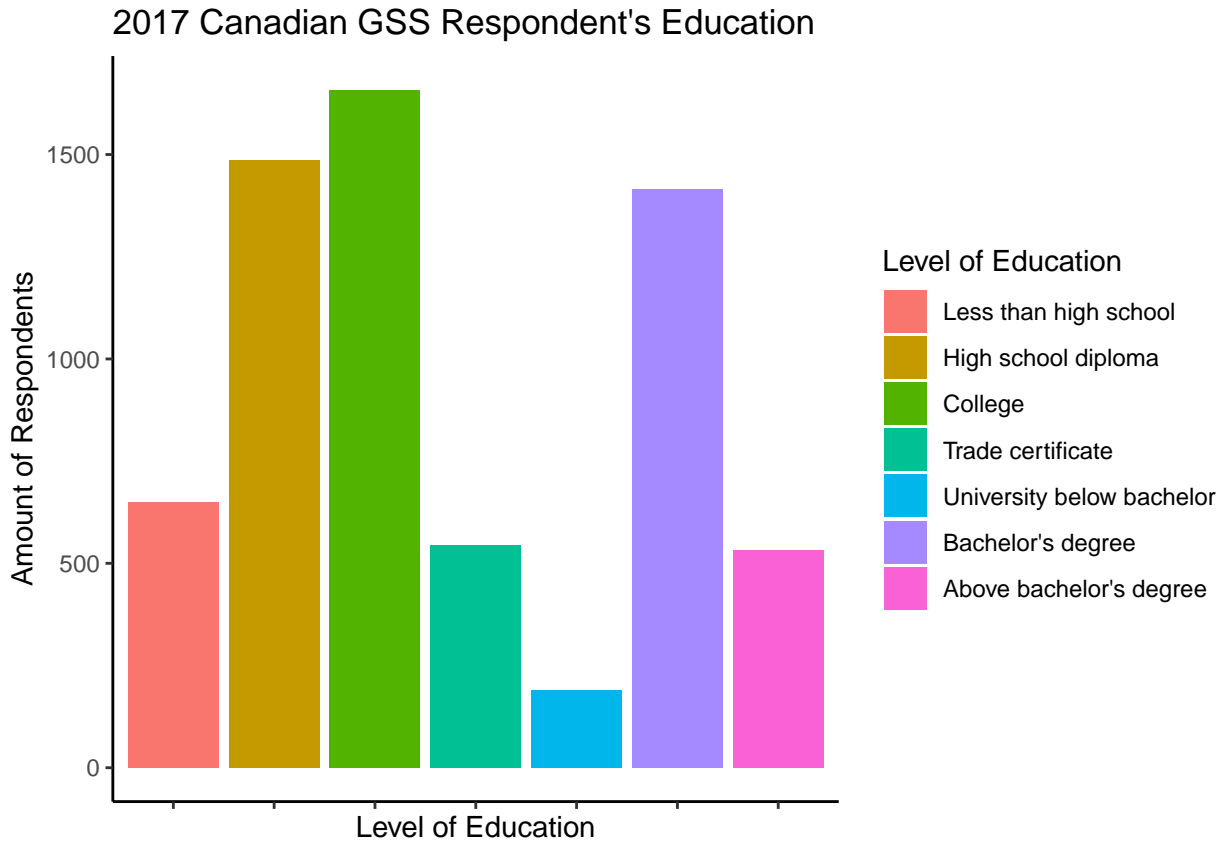## 2.3 Education: High school diploma, College, and Bachelor's Degree are Most Common

Figure 2: Majority of Respondents have either a high school diploma, college degree, or bachelor's degree

Figure 2 is created through ggplot(Wickham (2016)). A large portion of the respondents has either a high school diploma, college education, or bachelor's degree. The number of respondents who have an education less than high school and above bachelor's degree is quite similar. The least amount of respondents has a university below bachelor's degree, with around 250 responses. All other bins have at least 500 respondents. There seems to be a good amount of respondents in each education level to investigate education as a possible factor in a relationship with happiness.

2.4   Health: Majority Believe They Are Above Average For Health

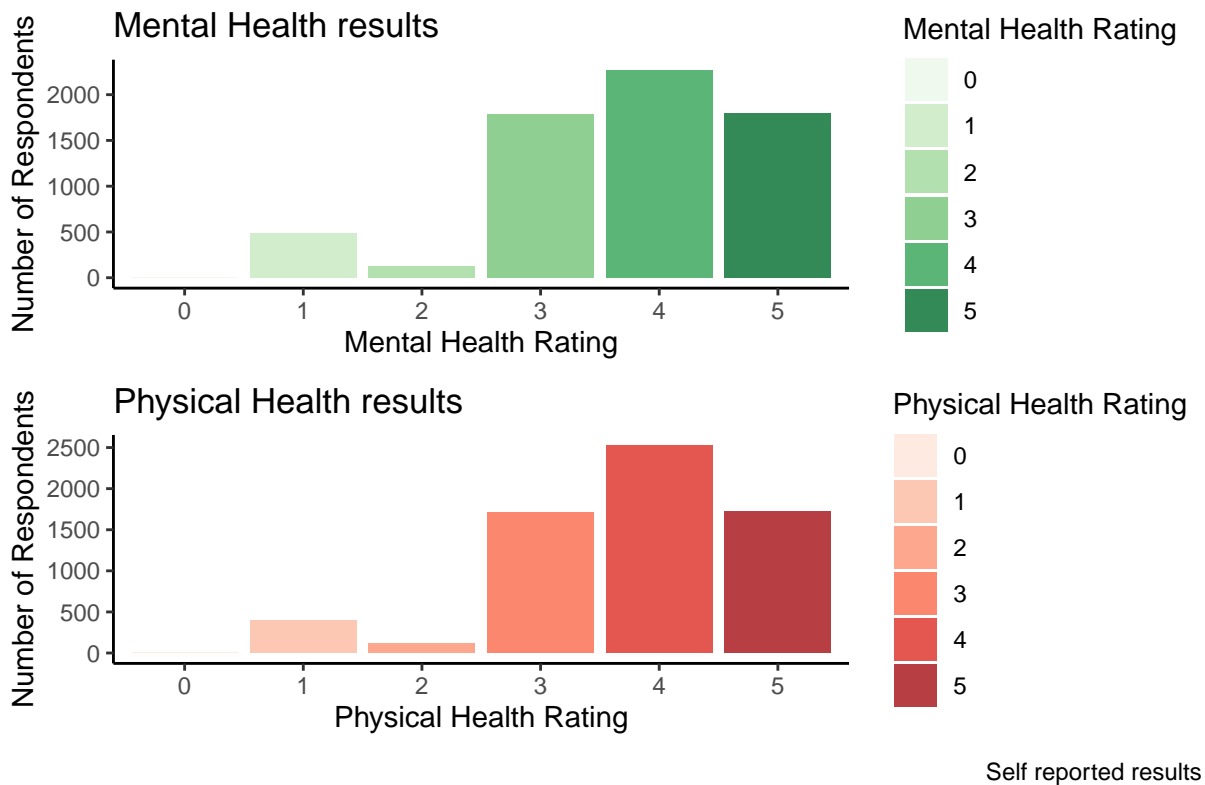Respondent's Belief about Own Health (Physical & Mental)

Figure 3: Similar Trend for Mental and Physical Health

Figure 3 is created through ggplot(Wickham (2016)) and patchwork (Pedersen (2020)). The 2 bar plots describe respondents' self-reported results in terms of their physical health and mental health. The trend for physical and mental health are very similar, with a majority rating themselves 4 out of 5, having very good health—few rate themselves as having fair health, a 2 out of 5. An interesting result is that there are more respondents who responded poor or good (1 or 3 respectively on the scale) than fair. This could be due to the way that respondents interpret the word fair.

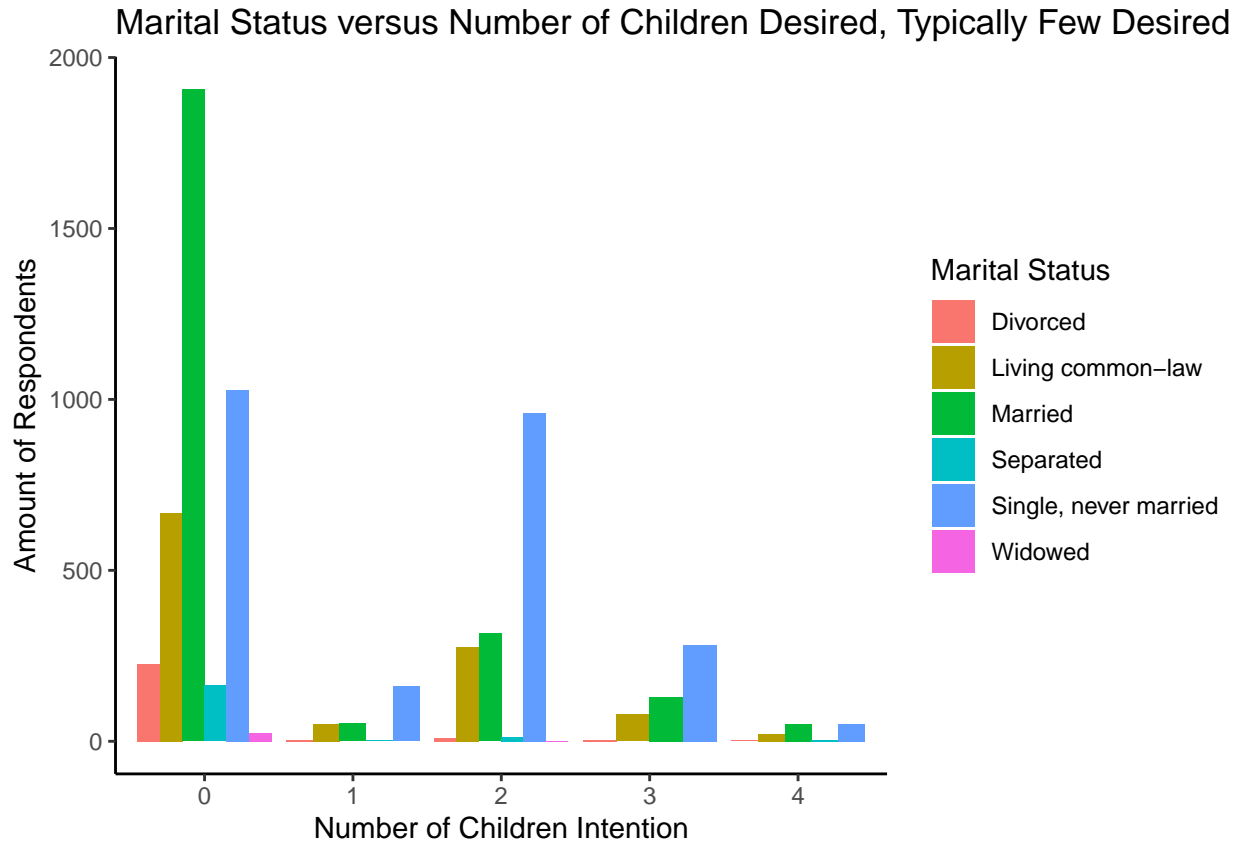## 2.5 Marital status: Fewer Children Are Desired

Figure 4: Majority do not want children, regardless of Marital Status

Figure 4 is created through ggplot(Wickham (2016)). The barplot looks at how marital status can be related to how many children the respondent intends to have. The majority of the respondents, regardless of their marital status, intend to have 0 children. Out of the respondents who do want children, the most common marital status are single, never married, married, and living common-law. Very few respondents who are divorced want children, with the only visible bin being at 0 children. Respondents who are widowed show a similar trend, but with fewer respondents overall. This could be due to the overall minimal number of respondents who are widowed in the sample.

2.6 Income: Different Trend in Respondent and Family Income

Comparison of Family and Respondent's Income

Figure 5: Different trends for family and respondent income, distribution shows opposite skews.

Figure 5 is created through ggplot(Wickham (2016)) and patchwork (Pedersen (2020)). The 2 bar plots compare the difference between the family's income group and the respondent's income group before tax. The highest bin demonstrates that the respondent is most likely to have a before-tax income of less than 25,000, with 25,000 to 49,999 being the next prominent. In contrast, for the family's income group, the most common group is 125,000 or more. The trends overall seem to suggest that the respondent is not the main person who is bringing in the majority of the family's income. This can be a suggestion that the respondent's burden to provide for their family is not as major as someone who is the sole moneymaker within their family.

2.7 Summary Table: Equal Male and Female Respondents, Few of Aboriginal or Visible Minority Groups

Table 1: Data set: Number of Respondents by Sex and Race

	Yes	No	Don't Know
Male	3077	3399	0
Visible Minority	392	6057	27
Aboriginal	375	6073	28

Table 1 is created through the kable function in the knitr package Xie (2021) and kableExtra Zhu (2021) as well. The number of respondents who are male and female is quite similar, about half split. Around 30

respondents do not know if they are visible minority or aboriginal. There were around 400 respondents who answered yes to being of a visible minority or aboriginal status. As in comparison to our 6476 observations, we believed that this was a small enough proportion that it would be hard to find significant results. It is worth noting that all respondents were classified as male or female; nobody was classified as "don't know." This could mean that those of different gender statuses went by their birth sex, or some way similar to deal with the other possible gender classifications.

2.8 Summary Table 2: Proportion of Visible Minority Respondents Similar to Aboriginal

Table 2: Data set: Proportion of Respondents by Sex and Race

	Yes	No	Don't Know
Male	0.4751	0.5249	0.0000
Visible Minority	0.0605	0.9353	0.0042
Aboriginal	0.0579	0.9378	0.0043

Table 2 was created using kable in knitr Xie (2021) and kableExtra package Zhu (2021). It reveals the proportions of the results in Table 1. Gender is quite evenly split, with around 50% for each gender. Around 6 percent of the respondents answered yes to visible minority and aboriginal. This corresponds with how the GSS survey collects information about the general Canadian population, where there is a small percentage of people in Canada who are of visible minority or aboriginal.

2.9 Summary Table 3: Age of each happiness group is relatively the same

Table 3: Mean and standard deviation of age for all happiness groups are very similar.

Happiness rating	Age (mean)	Age (sd)
1	34.7	10.64
2	38.07	7.45
3	32.8	9.9
4	33.79	9.26
5	37.61	10.28
6	34.57	9.5
7	35.09	9.57
8	35.53	10.28
9	35.35	10.28
10	35.47	10.3

Table 3 was created using kable in knitr Xie (2021) and kableExtra package Zhu (2021). It summarizes the age distribution for each happiness group, the desired response variable. All groups had relatively the same standard deviation and mean, with the mean age being around 34 to 37 and the standard deviation being around 9-10. The group that is the most different from the overall trend is the happiness group 2 (rated themselves being 2 out of 10 in terms of happiness), where the mean age is around 38, and the standard deviation is 7.45. The group that answered 3 out of 10 for happiness also has a relatively lower mean of 32. Nevertheless, the deviations are quite small, leading to the overall relatively similar age characteristics for all happiness groups.

3 Methodology

Based on the literature review, we narrowed down the variables to those that had a significant impact on feelings of life (happiness). We used visual aids such as faceted histograms and bar plots, alongside summary statistics, as well as individual relationships between the feelings of life and other relevant variables to further narrow down our list. We then checked for individual and pairwise relationships between our variables and for whether a linear regression model would be the right fit for the data that we have.

We used power transformation Fox and Weisberg (2019) to check if any of our variables needed to be transformed for a more accurate fit to a linear regression model. We fit a baseline model with all the variables that we chose, which we then reduced using likelihood ratio tests (since our models were nested). We finally chose a model that seemed to explain the variation in feelings of life, and we checked to see if the VIF (variance inflation factor) for our model was not greater than 5 for any one variable.

See appendix Figure 15 for the specific checking of assumptions.

4 Results

4.1 Education, Income, and Happiness

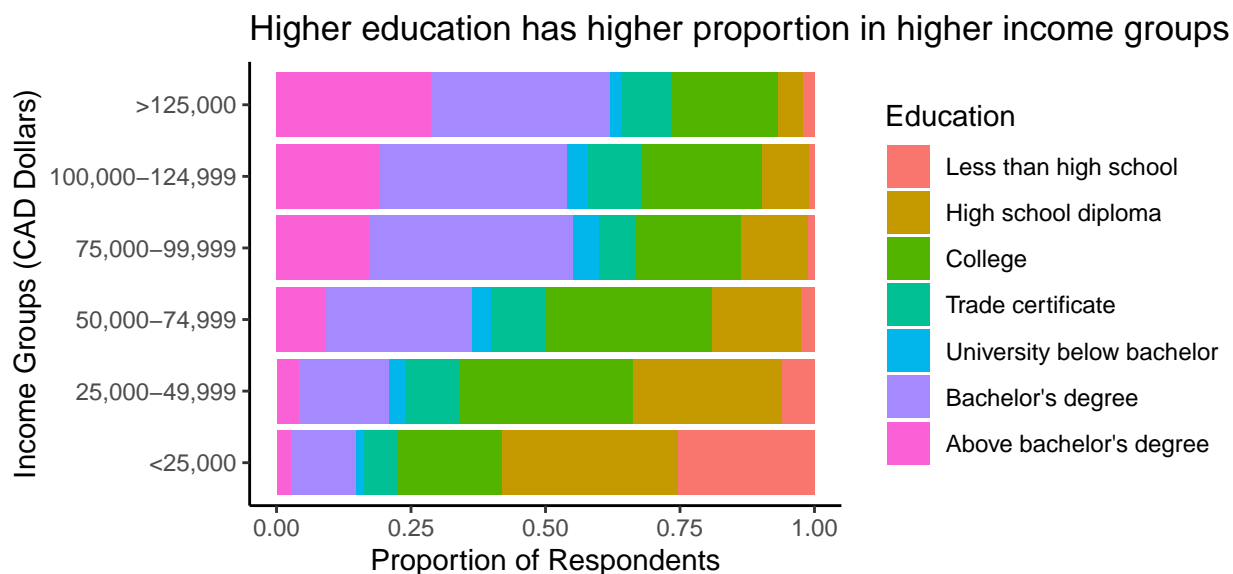### 4.1.1 Positive Relationship between Income and Education

Figure 6: Education and income has a positive, direct relationship.

Figure 6 compares the proportion of different education levels for each income group. There is an increasing trend for the bachelor's degree group and above bachelor's degree group (pink and purple), as the proportion of respondents from these two education levels increases as the income group increases. The opposite can be said for less than high school and high school diploma groups (red and yellow). As the income group increases, the proportion of respondents who have an education equal to or less than high school decreases. College, trade certificate, and university below bachelor does not show as prominent of a trend as the education levels on both ends of the scale.

4.1.2 Income: Higher Income, High Concentrated of High Happiness Levels
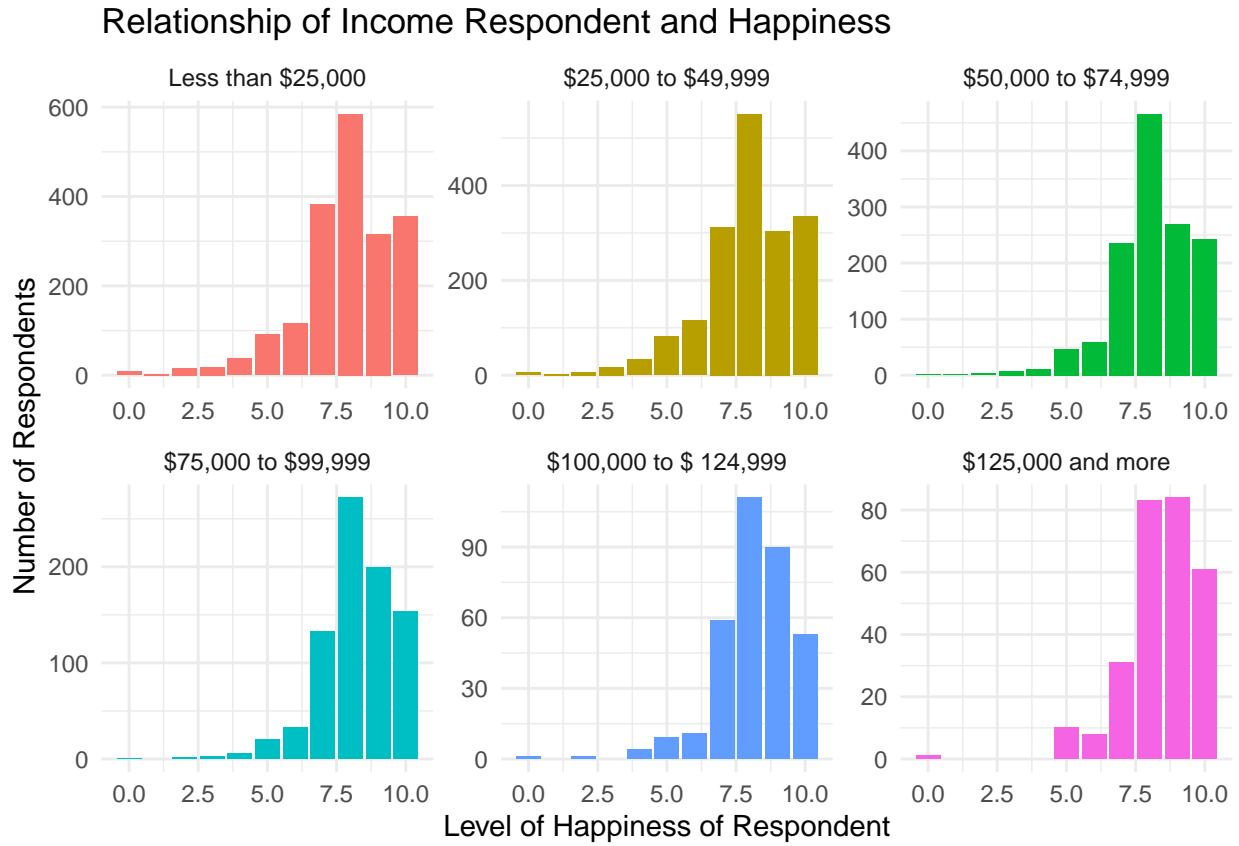

Figure 7: As income increases, the left-skewed tail decreases.

Figure 7 reveals similar distributions of happiness groups for each income level. All income groups show a left-skewed graph, with the majority of the respondents answering 8 as the most common level of happiness. However, as the income group increases, the tail decreases and a larger proportion of the respondents are concentrated at a higher level of happiness. The group of 125,000 CAD or more has a slightly different shape in comparison with the rest of the income groups. It has 2 modes at 8 and 9 for happiness groups. There seems to be a slightly positive relationship between income and happiness, although it is not very obvious.

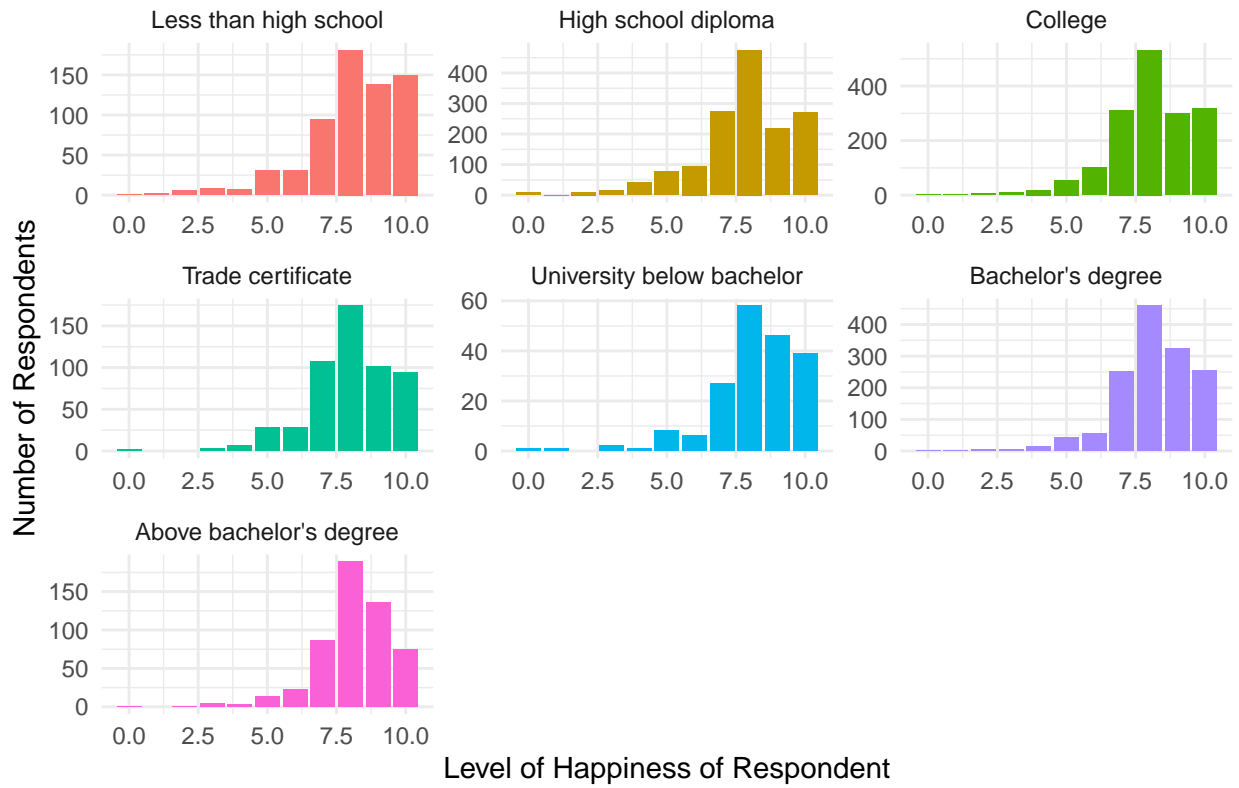#### 4.1.3 Education: No Significant Effect on Happiness

Figure 8: Generally same distribution of life satisfaction regardles of eduation level.

Figure 8 demonstrates relatively similar distributions for all different education levels. All education levels show a peak at 7-8, along with a left-skewed distribution. It is quite interesting that for those who have an education less than high school, the respondents have one of the highest proportions for the maximum happiness level of 10 within the different education groups. Overall, there are very few who are rated below 5 for happiness regardless of the education level. There does not seem to be a significant relationship between education and happiness level.

4.2 Health: Higher Ratings of Health Seems to Result in Higher Happiness
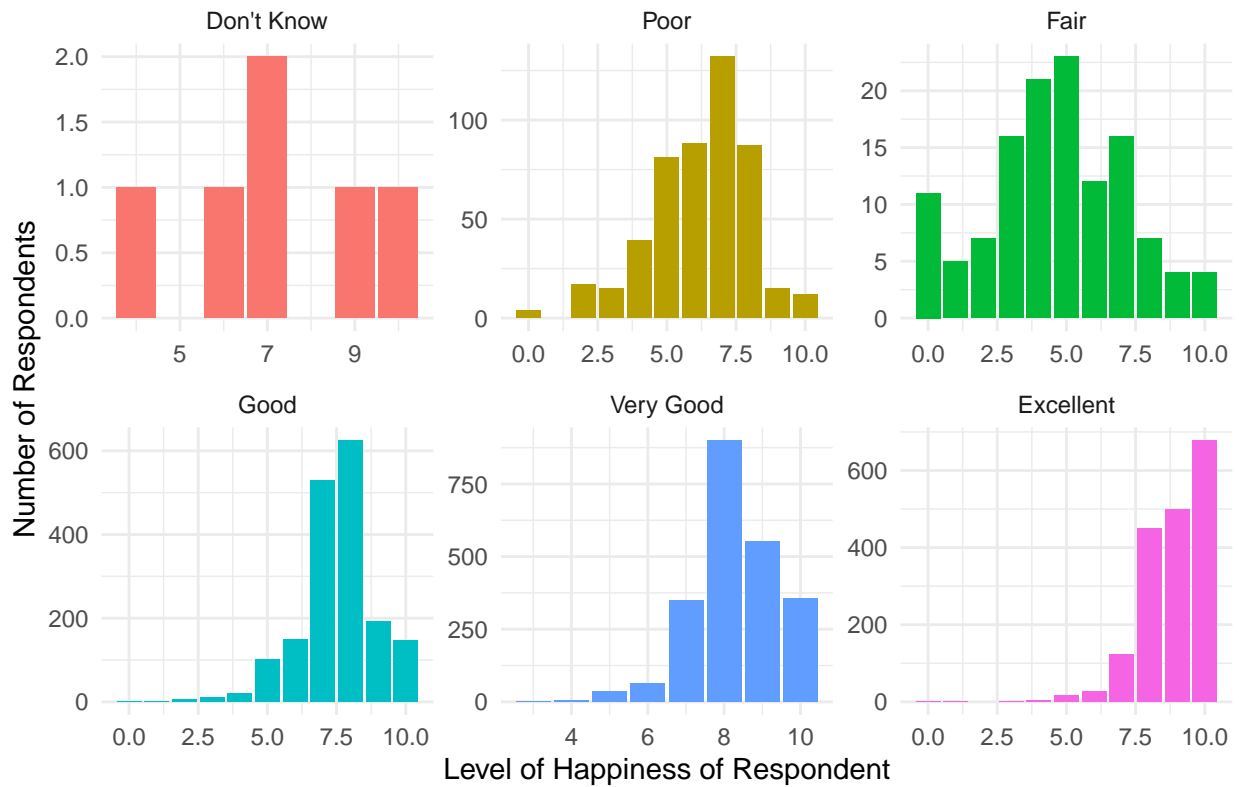
4.2.1   Mental Health Results

Figure 9: At higher levels of mental health, generally larger proportion are more satisfied with life.

Figure 9 investigates the relationship between the level of happiness and the different mental health self-rated groups. As the self-rated results for mental health increase, the distribution of the level of happiness becomes more left-skewed. In comparison, those who answered fair were relatively evenly spread out, with a rating of 5/10 being most common. Those who answered "excellent" were concentrated around 8-10 on the happiness scale, with few respondents answering less than this range. It is worth noting that the respondents who answered poor were actually more left-skewed than those who answered fair for their feelings of life, with a mode around 7 while fair had a mode around 5. It seems that the general trend cannot be generalized for each specific level yet.
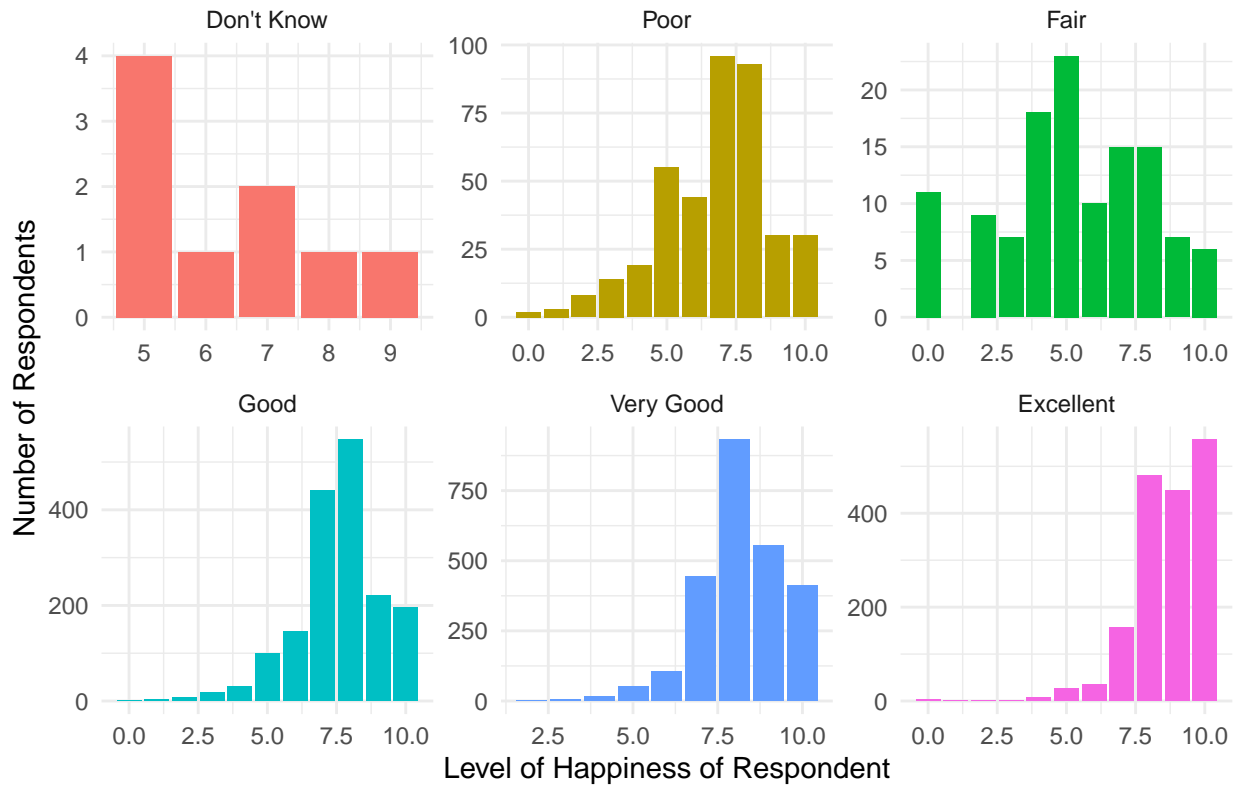
4.2.2 Physical Health Results

Figure 10: Higher health has high proportion of people with higher life satisfaction.

Figure 10 investigates the relationship between the level of happiness and different self-rated health groups. Like Figure 9, there seems to be a similar trend that as people self-rate themselves in higher health groups, the respondents are more highly concentrated at the top end of the scale of happiness. Similarly, it seems that when comparing poor and fair health groups, the fair group overall is less happy than poor, as demonstrated by the lower mode at 5 and the larger bins in the range 0 to 5. Moreover, comparing good and very good groups, the respondents who answered good for health have a higher proportion greater than 7. Those who answered very good have a higher concentration answering 10 out 10, being totally satisfied and happy with their life. There still seem to be differences between specific groups that deviate from the overall trend.

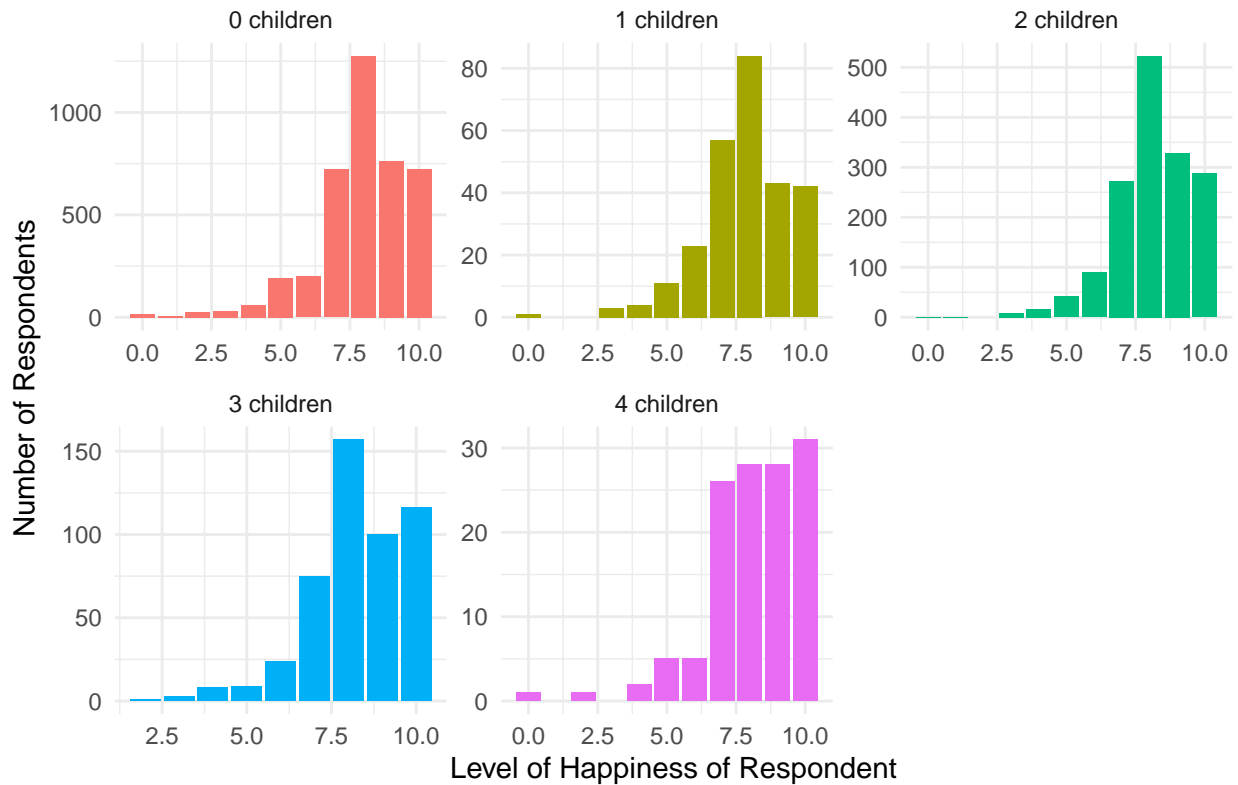## 4.3 Children Desired and Happiness: No Significant Relationship, Excluding Those Who Desire 4 Children

Figure 11: Those who want 4 children show a different trend, where most respondents are very satisfied with life.

Figure 11 shows that there is a significantly different happiness level distribution for the respondents that want 4 children in comparison to the groups that want 0 to 3 children. Respondents who want 0 to 3 children show a left-skewed distribution, with a prominent mode at 8. In contrast, respondents who want 4 children show multiple modes around 7 to 10. While also being left-skewed, the bins to the right are quite uniformly distributed. It is likely that findings will be similar for those who desire 0 to 3 children, with respondents who desire 4 children displaying a different trend. Looking at the distributions of the different groups, it seems that desiring 4 children leads to a higher proportion of respondents to have higher happiness levels, specifically scoring 7 or above. Nevertheless, it seems that there does not seem to be one overall trend for all groups between happiness and number of children desired that can be observed from Figure 11.

4.4 Marriage and Happiness: Similar Distribution, Excluding Widowed

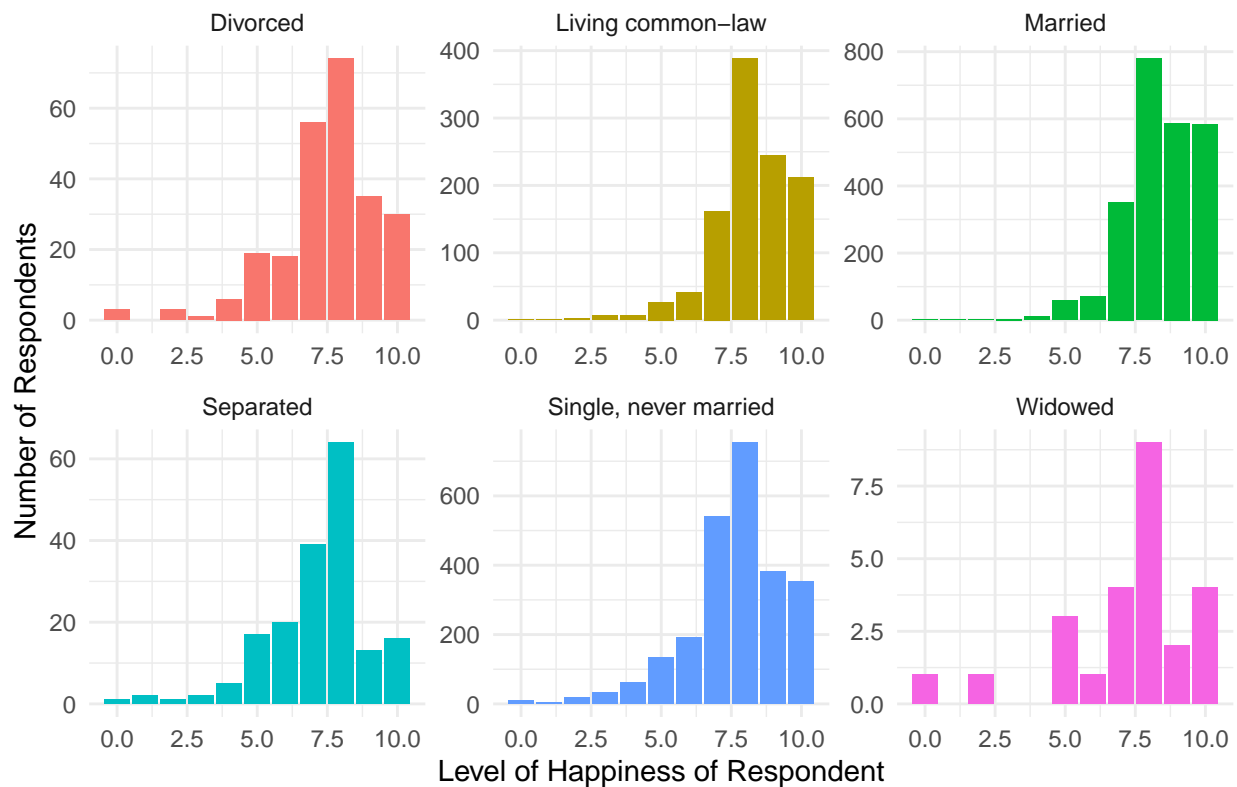### Relationship between Marital Status and Happiness

Figure 12: Generally same distribution for all marital status, except for widowed.

Figure 12, similar to figure 11, generally shows similar distributions for all marital status groups except for widowed. All marital status groups demonstrate a left-skewed distribution, with a mode around 8 for happiness level. Although it is relatively small, respondents who are a part of widowed shows the highest proportion of responses below 5 for happiness. This can be noted by the visible 2 bins at 0 and 2, while the other groups do not have any real visible bins below 5. Another unique feature regarding the widowed group is how there are few respondents who answered 6 or 9. All other groups show a more smooth distribution, with 5 and 9 being the second or third most likely response. This could be due to the smaller sample group of those who are widowed. While the overall trend is similar, there do seem to be unique features for the widowed group that may explain some variation in our results with happiness. Altogether, however, it does not seem that changing marital status has any obvious relationship with happiness level.

4.5 Regression Model

Table 4 was created using jtools Long (2020) and gtsummary Sjoberg et al. (2021). We can see that education is negatively correlated with happiness, with the baseline being the education level less than high school. Each subsequent level of education reduces happiness, holding all else equal. These results are very significant. Age is not a significant variable in the regression.

We can see that the respondent's income levels are negatively correlated with happiness, with the baseline being earning less than 25,000 CAD before tax. Moving up the income chain for a respondent slightly reduces happiness, holding all the other variables constant. These results, however, are not very significant. Family income levels, on the other hand, are positively correlated with happiness, with the baseline being earning less than 25,000 CAD before tax. As the respondent's family earns a higher combined income, the average

Table 4: Linear Regression Results

Characteristic	**Beta**	**95% CI**	**p-value**
education			
Less than high school			
High school diploma	-0.33	-0.45, -0.21	<0.001
College	-0.31	-0.44, -0.19	<0.001
Trade certificate	-0.44	-0.59, -0.29	<0.001
University below bachelor	-0.31	-0.52, -0.10	0.003
Bachelor's degree	-0.39	-0.52, -0.26	<0.001
Above bachelor's degree	-0.44	-0.60, -0.29	<0.001
income_respondent			
Less than $25,000			
$25,000 to $49,999	-0.08	-0.17, 0.02	0.13
$50,000 to $74,999	-0.13	-0.23, -0.02	0.020
$75,000 to $99,999	-0.07	-0.20, 0.05	0.2
$100,000 to $ 124,999	-0.22	-0.38, -0.05	0.010
$125,000 and more	-0.17	-0.35, 0.01	0.064
income_family			
Less than $25,000			
$25,000 to $49,999	0.08	-0.07, 0.23	0.3
$50,000 to $74,999	0.17	0.03, 0.32	0.018
$75,000 to $99,999	0.17	0.03, 0.32	0.018
$100,000 to $ 124,999	0.23	0.08, 0.38	0.002
$125,000 and more	0.30	0.16, 0.44	<0.001
self_rated_mental_health			
Don't Know			
Poor	-1.0	-2.0, -0.01	0.047
Fair	-2.3	-3.3, -1.3	<0.001
Good	0.04	-1.0, 1.1	>0.9
Very Good	0.64	-0.37, 1.6	0.2
Excellent	1.1	0.12, 2.1	0.029
self_rated_health			
Don't Know			
Poor	0.85	0.02, 1.7	0.046
Fair	0.04	-0.82, 0.89	>0.9
Good	1.2	0.37, 2.0	0.005
Very Good	1.4	0.54, 2.2	0.001
Excellent	1.5	0.72, 2.4	<0.001
marital_status			
Divorced			
Living common-law	0.34	0.16, 0.51	<0.001
Married	0.47	0.30, 0.64	<0.001
Separated	-0.19	-0.43, 0.05	0.12
Single, never married	-0.08	-0.25, 0.09	0.3
Widowed	-0.18	-0.69, 0.34	0.5
number_total_children_intention			
0 children			
1 children	0.01	-0.15, 0.16	>0.9
2 children	0.12	0.04, 0.20	0.003
3 children	0.27	0.14, 0.39	<0.001
4 children	0.08	-0.15, 0.30	0.5

respondent's happiness, holding all the other variables constant, increases. These results are very significant as well.

For mental health, if we take the baseline to be a respondent that has stated his mental health status as 'good,' for an increase in mental health, there is an increase in average happiness, holding all else equal. Similarly, for a decrease in mental health from the baseline response of 'good,' there is a decrease in average happiness, holding all other variables constant. For physical health, we can see that every increase in reported health rating increases happiness significantly, holding all other variables constant.

5   Discussion

5.1   Education and Happiness

Based on the results that we get from the regression model, we find that education is actually negatively correlated with feelings of happiness. The majority of the respondents to the survey were high school graduates or had a bachelor's degree from a university or college. The results of our regression that the happiest respondents were those that had an education lower than high school. To further investigate this trend, the age of the respondents was looked at in the appendix, Figure 14. The figure reveals that those who have an education less than high school were around age 20 or less. At this youth age range, people generally tend to be happier due to having to deal with less stress that comes with a professional life, whether it be work or exams and difficult educational work from post-secondary institutions. Moreover, there is less financial burden and responsibilities for youth, as they do not have to provide for the family and have limited responsibilities as a minor. Rather than education being the cause, it is likely that the overall less stressful situation for the typical age of those who have an education level less than high school is a reason behind the happiest results. Logically, if we were to follow this assumption, as respondents increase in age, facing more responsibilities and difficult work, this would cause more stress and less happiness.

Another explanation for the negative relationship can come from considering education's relationship with income. To get the relationship between education and happiness, all other variables, such as income, are held constant. Therefore a more accurate interpretation is that for people in the same income group, those that are more educated are generally less happy. This can be considered by utilizing a comparison mindset. If income was used as a level to measure happiness and success Latif et al. (2015), education could be considered an expense that helps achieve a certain income level. This is due to education requiring both financial resources and the valuable resources of time and effort of respondents. People would be happy if they had to spend less, have less education, but also get a high income, high happiness level. Therefore, within the same income group, as respondents with higher education spend more but obtain the same income as someone who put in less effort, these respondents would be unhappy at the unequal results of their effort.

It is worth noting that the trend found within the model could be limited in its application to real-life due to the interaction between education and income. Our results graph Figure 6 demonstrate a direct, positive relationship between education and income. Therefore, it is not plausible to assume that income would stay constant as education increases, which is the assumption that our regression model utilizes. Therefore, the results cannot be generalized to the overall public or the effect of education on happiness itself.

5.2   Income and Happiness

Our regression results demonstrate conflicting results for family and respondent income. The model shows that a higher level of respondents' income is negatively correlated to happiness. However, these results are insignificant. In contrast, the regression model shows significant results that for each level increase in family income, the respondent's happiness generally increases. The differing trends show how where income comes from can be an important factor in happiness. For example, it is possible that a reason for the different trends is that respondents that have a higher personal income may have to take a greater share of the household expenses. As they are the main source of income for their family, this leads to higher stress and pressure, and consequently more negative emotions with less happiness.

Focusing on literature about income, there are two main findings that interact with our report. First, the

conclusion that having a higher income is related to higher happiness in Latif et al. (2015) does not hold for our report. As our results are insignificant, it cannot be said that this report is against the results of Latif et al. (2015). However, as Latif et al. (2015) does state that the increase in happiness is temporal, this can be supported by our report, which demonstrates that the positive impact on happiness by personal income does not exist. Second, for respondents that belong to the same income group, those at the higher end of the group will be in general happier than those at the lower end of the group. The results for this report look at income groups in general, rather than what is between income group Easterlin (2001). This does provide a suggestion that there could be differing trends within the overall trend between groups that we have identified.

Our finding of household income can provide some insight regarding how family can be important to happiness. Marriage is known to increase happiness while having children decreases happiness Stack and Eshleman (1998) & WHITE, BOOTH, and EDWARDS (1986). Looking from an income perspective, these findings do seem to connect with each other. Marriage increases the number of income earners within the household, increasing household income. This would then lead to an increase in happiness, according to the results from our model. Having children, on the other hand, will probably lead to a decrease in family income. This can be due to the various expenses associated with having children, such as education, childcare costs, and more. Therefore, as having children decreases household income, this would lead to a decrease in happiness. The findings in this report regarding household income support the relationships in literature, providing insight into how the different factors interact and impact happiness.

An interesting consideration would be whether the income of children would be considered household income. Household income is defined as the sum of all income of people living in the same house who are 15 years or older. Considering the life of a typical Canadian, it is likely that once a child is able to earn a stable income, such as taking on a full-time job, they will be living independently, in a different place from their parents. Hence, despite their potential ability to increase household income, when they are able to do so, they will not be counted. The household income would only reflect the expenses the family incurred before the child became a stable income earner.

Finally, focusing on our Figure 7, we can see that income and education are positively correlated. Education is negatively correlated with happiness, yet income can be positively related to happiness. A potential explanation is that people that are more educated on average earn a higher income. Subsequently, this higher personal income leads to a higher household income, which significantly results in more happiness.

5.3 Happiness and Health (Physical and Mental Health)

Healthier individuals are on average happier than those that do not consider themselves healthy. This result is significant for both categories that we have, mental health and physical health, according to the regression model utilized. Respondents that stated their mental health as being 'good' or above were generally happier than those that responded with lower self-rated observations. This observation can be due to the fact that happiness can be considered an emotional concept, which depends on emotional and mental well-being. As the results are self-rated, respondents' responses reflect their current state of mind and how well they are feeling. These are very close feelings to what it means to be happy. Therefore, those who feel that they are struggling with mental health are probably likely to feel less happy as well. This echoes the results in the literature, where it was reported that life satisfaction was strongly associated with self-reported mental health Lombardo et al. (2018).

For physical health, respondents that reported a higher level were generally happier, holding all else equal. This relationship supports what was found in Siahpush, Spittal, and Singh (2008), where people who are happier have better health and longevity. It reflects the general idea that having better health grants people more freedom to do what they want. Someone who is constantly bed-ridden will most likely be less happy than someone who is healthy and active.

Considering the current environment where many Canadian citizens are stressed on a constant basis, this is an important finding that should be shared with people, especially the youth. While it may be true that having a higher income being "successful" can lead to happiness, happiness also depends on physical health and mental health. Moreover, being happy can also impact health and longevity, meaning that it is not only something desirable but is something crucial to survival as well. Therefore, to improve quality of life, there

should be a focus on finding ways to support mental and physical health. Ways could include more accessible mental and health resources, changing curriculum and adjusting company culture to show the importance of having a good balance between work, education, and life, maintaining both physical and mental health.

Altogether, it is essential for people to be healthy both mentally and physically in order to achieve happiness. Both aspects must be viewed as important, and activities should adjust to accommodate these as priorities.

5.4   Happiness, Marital Status, Total Children Desired

The respondent's marital status also has an effect on their happiness. Marital status has been categorized into different groups. This can be seen in figure 4. From our regression model, we see that the happiest group are those that are married, followed closely by those that are living common-law. We can also see that those that were separated were, on average, less happy compared to those who had decided to divorce their partner. These results seem to be similar to those that were found in our literature article, where being married had a positive impact on the happiness level of a person Stack and Eshleman (1998).

Further, we can see that the total number of children that are intended also seems to affect happiness, as happiness levels increase with each additional child that the respondent intends to have. This result could provide a different perspective on the effect children have on happiness. In WHITE, BOOTH, and EDWARDS (1986), the authors found that having children leads to less happiness in comparison to those without children. As the variable of interest within this paper is the number of children intended, it could be that there are different reasons and factors at play that actually lead to an increase in happiness. This could reflect the idealized version of life that respondents perceive before they face the actual burden and responsibility of having as many children as they actually desire.

We also inspect a correlation between marital status and the number of children that are intended. While the majority of the respondents did not intend to have any children, the groups that did want to have more were primarily comprised of married respondents. Coincidentally, these respondents reported being the happiest, holding all other factors constant. This suggests that there is a possibility that the effect of being married could outweigh the noticed negative impact of having children on happiness observed in the literature, as some of the married respondents do have children already. The respondents that reported that they were married also reported that they had a higher household income, but this was accounted for in the regression so that it does not confound the results that we get.

5.5   Weaknesses and Next Steps

Weaknesses and limitations that the report faces will be discussed in this section, including the next steps to address these limits. The next steps will also include ways to expand the results of our study to be applicable for a larger range of people, as well as providing a deeper investigation of factors impacting happiness.

5.5.1   Potential Implication Limits

There are a few limitations that must be considered when considering the results of the paper. While the model we opted for was a linear regression model, not all assumptions were fully met. Trying to transform the response variables was unhelpful in solving the situation, as the independent variables investigated were factors. Therefore, for the sake of interpretation and meaningful results, we decided to go forth with the linear regression.

Another limitation is the accuracy of the data collection process. As mentioned in the ethics and bias, many of the results were self-reported, with no objective, clear definition of each level. It could be possible that due to the subjective understanding of the various choices, such as what "fair" and "good" mean, respondents answered in a way that does not best represent their actual status.

Moreover, omitting the empty responses from our data set could have led to a disproportionate sample of respondents from different provinces in Canada. This limits the applicability of our results from being generalized to Canada overall.

5.5.2   Next Steps

In response to the limitations we faced, there are multiple steps that can be taken to clear up the concerns and constraints. First, a follow-up study that ensures the cleaned data set has both no missing values and a proportionate amount of respondents from each geographical region would help amplify the usability of the study. Second, a follow-up survey with clearer definitions of what each of the choices meant would be good. An example is included in the appendix of what such a survey could entail. Third, to address the issue of linear responses, collecting information about the variable happiness in a continuous way would help meet all the linear regression assumptions.

Expanding on our report, it would be useful to do an investigation of how gender and race can play a role in happiness. Despite issues to combat inequality and discrimination, there still is systemic discrimination in Canada Yum (n.d.). It is highly plausible that race and gender could be crucial factors that impact the respondents' happiness levels. Moreover, it would be great to see the interaction of gender and race with the investigated variables in this study. Depending on a person's gender and race, it could be that their ability to obtain certain results is impacted, hence impacting their happiness.

Moreover, an investigation of how work impacts happiness would be a good addition to gaining a better understanding of factors that affect happiness. After education, a large portion of a citizens' life would be spent at work. Therefore, looking at factors such as working hours, work-life balance, pressure could provide a better insight into other important factors for happiness.

In addition, an investigation of how the COVID pandemic has impacted the factors that affect happiness would help provide insight into how the pandemic has changed peoples' lives. It is possible that the dynamic between social life and happiness, as well as other factors, has changed due to the social restrictions the COVID pandemic has brought upon society. This would help provide insight on how people's values have changed due to COVID and whether there should be a pivot in policy to accommodate the new norms people are used to now.

Finally, a comparison with other countries around the world regarding happiness and related factors that impact the satisfaction level for life could be valuable. The comparison could help reveal how much culture and policy play a role in the meaningfulness of certain factors to impact happiness. A comparison with the United States, specifically, could be interesting, as there are many similar values in a culture that Canada and the United States hold in common. A larger study in comparison to other countries would also help identify whether such a thing is inherent to Canada's situation or applicable to the world in general.
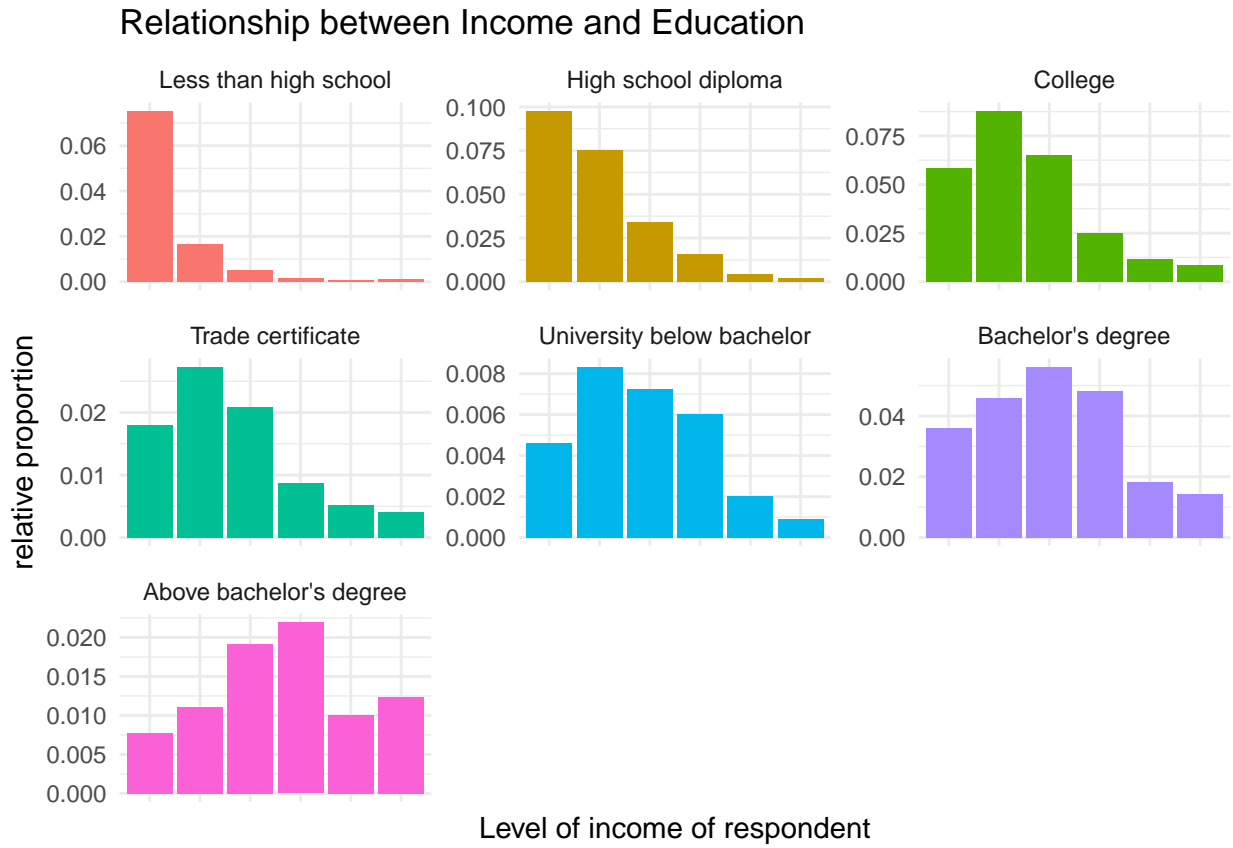
Appendix

.1 Figures

Figure 13: Higher education has higher proportion of respondents with higher income.
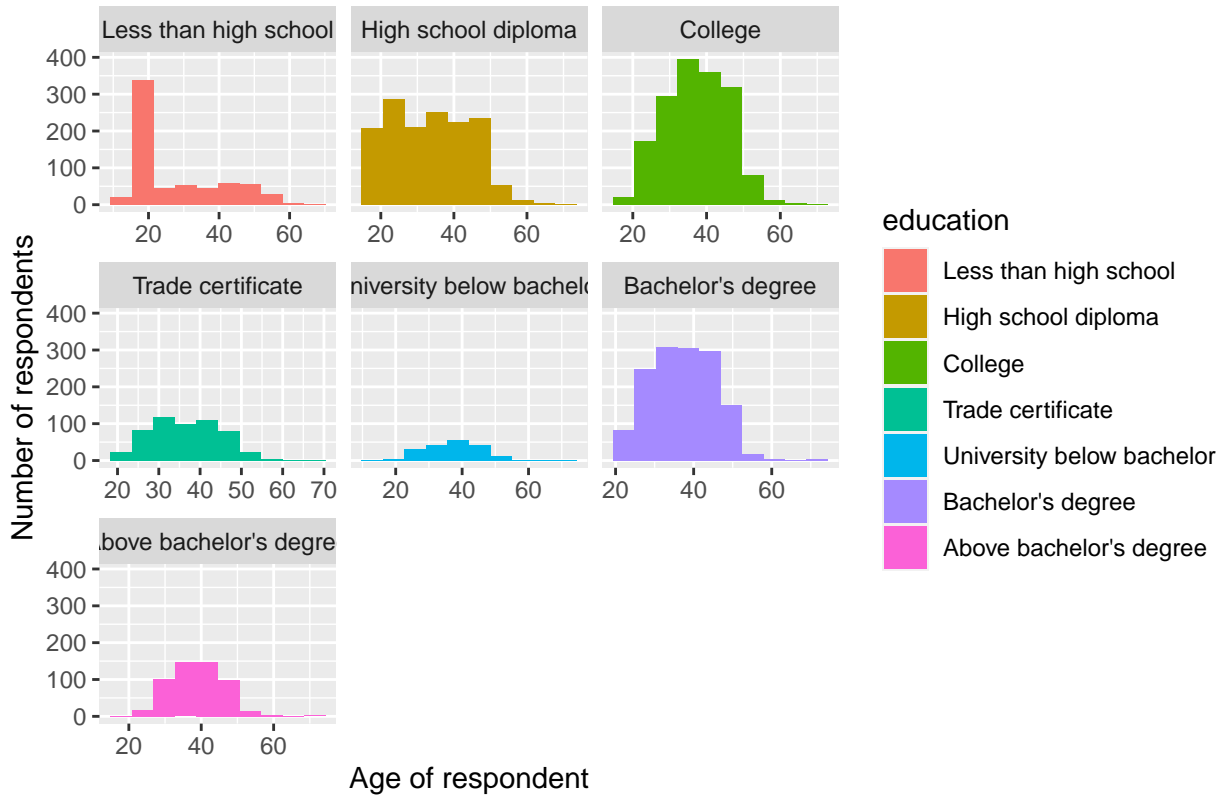
Figure 14: Majority of respondents who have less than high school education are less than 20 year old

Y vs fitted
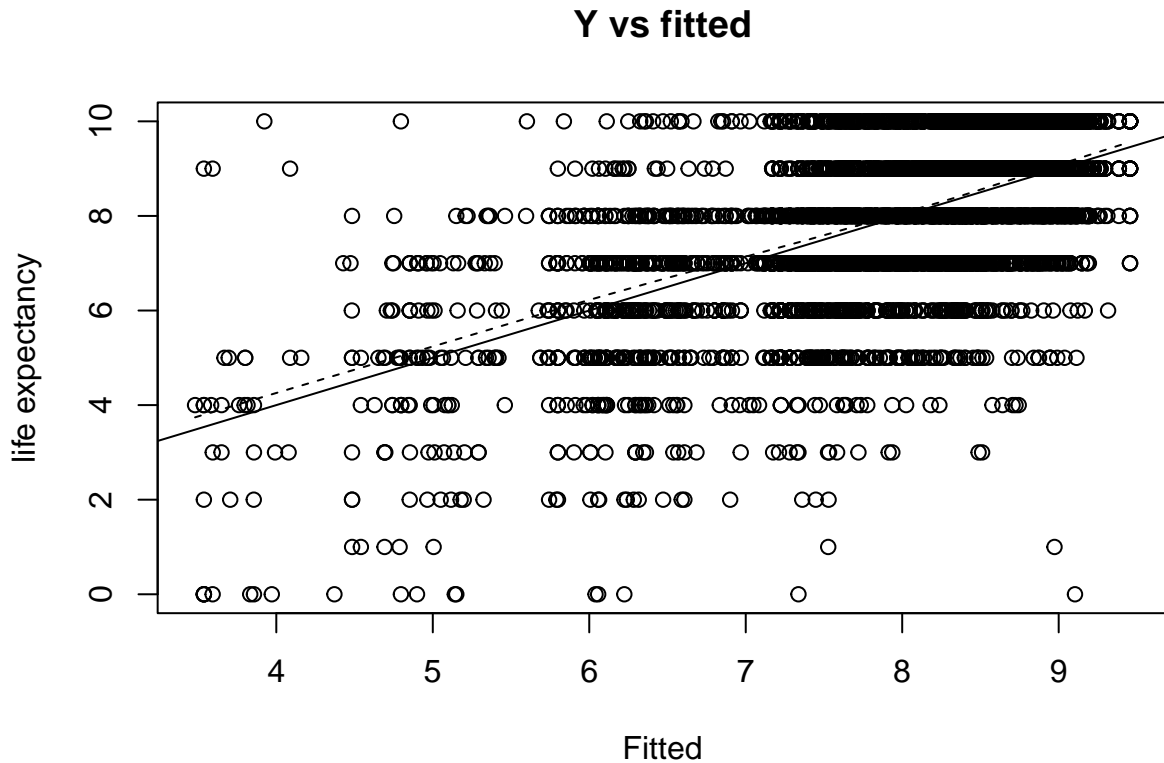

Figure 15: Checking linear assumptions: linearity not quite met.

```
##                             GVIF Df GVIF^(1/(2*Df))
## education               1.391428  6        1.027910
## income_respondent       2.860818  5        1.110834
## income_family           2.402046  5        1.091586
## self_rated_mental_health 1.860319  5       1.064042
## self_rated_health       1.895862  5        1.066058
```

.2 Supplementary Survey

Our supplementary survey is available here: https://forms.gle/u4VEqsUUZH6r7A6u5.

.2.1 Preamble

The purpose of this survey is to investigate further factors related to happiness. We are interested in more detailed gender and race data than what is within the 2017 Canadian GSS and the impact it has on happiness. We are also interested in a more detailed understanding of how the pandemic might have impacted factors related to happiness. Questions regarding health, happiness in different social spaces, and feelings during different stages of the pandemic are asked to gain a better understanding of COVID's impact upon Canada.

By proceeding with the survey, you understand that we will use your responses to understand better the effect of race, gender, and the COVID pandemic upon happiness. The survey is voluntary, and if you decide to participate, you can skip questions and withdraw at any time.

.2.2 Questions

1.What is your gender?

- Male

- Female
- Transgender
- Nonbinary
- Genderqueer
- Other (please specify)

2. If you are of visible minority, what is your specific group?

- Black
- Chinese
- Filipino
- Japanese
- Korean
- South Asian-East Indian
- Southeast Asian
- North African
- Arab
- Not a part of visible minority groups
- Other (please specify)

3. On a scale of 1-5, how happy do you feel at your workplace/school? 1 is the least happiest, and 5 the happiest.

- 0
- 1
- 2
- 3
- 4
- 5

4. On a scale of 1-5, how happy do you feel at home? 1 is the least happiest, and 5 the happiest.

- 0
- 1
- 2
- 3
- 4
- 5

5. On a scale of 1-5, how stressed do you feel? 1 is the least stressed, 5 the most.

- 0
- 1
- 2

- 3
- 4
- 5

6. Do you have any chronic illnesses?

- Yes
- No

7. In the last 3 months, how many times have you gotten sick?

- 0
- 1
- 2
- 3
- 4
- greater than 5
- don't know

8. Have you been diagnosed with any mental illnesses?

- yes
- no

9. At the present time, how would you describe your emotional health?

- poor
- fair
- good
- very good
- excellent

10. On a scale of 1-5, how would you describe your current emotional health state? 1 is the worst, 5 the best.

- 1
- 2
- 3
- 4
- 5

11. Over the COVID pandemic 2020-2021, how would you describe your emotional health?

- poor
- fair
- good
- very good
- excellent

12. Over the COVID pandemic 2020-2021, how would you describe your emotional health on a scale of 1-5? 1 being the worst, 5 the best

- 1
- 2
- 3
- 4
- 5

13. After all COVID restrictions were removed in March 2022, how would you describe your emotional health?

- poor
- fair
- good
- very good
- excellent

14. After all COVID restrictions were removed in March 2022, how would you describe your emotional health on a scale of 1-5? 1 being the worst and 5 the best

- 1
- 2
- 3
- 4
- 5

References

Alexander, Rohan. 2022. "Cleaning for CA GSS 2017." https://piazza.com/class/kyby191sd475p4?cid=167.

Canada, Statistics. 2019. "This Document Provides a Comprehensive Reference to the Information Available from the General Social Survey (GSS). It Provides a Description of the Content of Each of the 18 GSS Cycles (e.g. Time Use, Social Support, Education, the Family), as Well as Background Information, Target Population and Collection Methodology. A List of the Products and Services Available from Each Cycle Is Also Included." *Statistics Canada: Canada's National Statistical Agency / Statistique Canada : Organisme Statistique National Du Canada.* Government of Canada, Statistics Canada. https://www150.statcan.gc.ca/n1/pub/89f0115x/89f0115x2019001-eng.htm.

———. 2020. "General Social Survey Cycle 31 : Families Public Use Microdata File Documentation and User's Guide." Statistics Canada.

Easterlin, Richard A. 2001. "Income and Happiness: Towards a Unified Theory." *The Economic Journal (London)* 111 (473): 465–84.

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

Latif, Ehsan et al. 2015. "Happiness Adaptation to Income: Evidence from Canada." *Economics Bulletin* 35 (3): 1477–87.

Lombardo, Patrick, Wayne Jones, Liangliang Wang, Xin Shen, and Elliot Goldner. 2018. "The Fundamental Association Between Mental Health and Life Satisfaction: Results from Successive Waves of a Canadian National Survey." *BMC Public Health* 18 (March). https://doi.org/10.1186/s12889-018-5235-x.

Long, Jacob A. 2020. *Jtools: Analysis and Presentation of Social Scientific Data.* https://cran.r-project.org/package=jtools.

Nikolaev, Boris, and Pavel Rusakov. 2016. "Education and Happiness: An Alternative Hypothesis." *Applied Economics Letters* 23 (January): 827–30. https://doi.org/10.1080/13504851.2015.1111982.

"Parenthood and Happiness: Effects of Work-Family Reconciliation Policies in 22 OECD Countries." 2016. *American Journal of Sociology* 122 (3): 886–929. https://doi.org/10.1086/688892.

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/.

Siahpush, Mohammad, Matt Spittal, and Gopal K. Singh. 2008. "Happiness and Life Satisfaction Prospectively Predict Self-Rated Health, Physical Health, and the Presence of Limiting, Long-Term Health Conditions." *American Journal of Health Promotion* 23 (1): 18–26. https://doi.org/10.4278/ajhp.061023137.

Sjoberg, Daniel D., Karissa Whiting, Michael Curry, Jessica A. Lavery, and Joseph Larmarange. 2021. "Reproducible Summary Tables with the Gtsummary Package." *The R Journal* 13: 570–80. https://doi.org/10.32614/RJ-2021-053.

Stack, Steven, and J. Ross Eshleman. 1998. "Marital Status and Happiness: A 17-Nation Study." *Journal of Marriage and Family* 60 (2): 527–36. http://www.jstor.org/stable/353867.

Uoftlibraries. n.d. "University of Toronto Libraries." *CHASS Data Centre.* University of Toronto. http://datacentre.chass.utoronto.ca.myaccess.library.utoronto.ca/.

WHITE, LYNN K., ALAN BOOTH, and JOHN N. EDWARDS. 1986. "Children and Marital Happiness: Why the Negative Correlation?" *Journal of Family Issues* 7 (2): 131–47. https://doi.org/10.1177/019251386007002002.

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. 2021. *Knitr: A General-Purpose Package for Dynamic Report Generation in r.* https://yihui.org/knitr/.

Yum, Hyejung Jessie. n.d. "The Myth of Happiness: A Commentary on Racism and Mental Health in Canada." *The Canadian Journal of Theology Mental Health and Disability*, Spring 2021, 1 (1). https://jps.library.utoronto.ca/index.php/cjtmhd/article/view/36231/27492.

Zhu, Hao. 2021. *kableExtra: Construct Complex Table with 'Kable' and Pipe Syntax.* https://CRAN.R-

project.org/package=kableExtra.