

# How is the number of Children in Canada change?

Qianyuan Huang; Bradley Robert Wong; Jiayi Zheng

03/26/2022

## Abstract

The government has been investing lots of resource on improving the wellbeing of Canadian. GSS is a voluntary survey collected every year with different topics. There are multiple themes life at work, home etc... the survey we used and analyzed has concluded the factors affecting the number of children in Canada. It shows that the number of children is affected by several factors which include religion, marital status and family income. The goal for us is to help policymakers analyze the important factors affecting the number of children to improve the quality of policymaking.

Website for github<sup>1</sup>

## 1 Introduction

Whether we are assessing options for changing the society, attempting to make changes to improve the well-being of children. Policy makers need to make proper assessments on trends in families and children. These data were based on General Social Survey. Our aim is to help the policy makers take out all the useful data and make models to explain certain trends on families and children to ensure they can have accurate information when making policies. Children will be the future of our society and they are very important. We would want this project to help make more comprehensive and improve policies in the current society. Therefore, family and children would both be important in responding to the survey.

A short survey was created and conducted by the government to access the opinions of Ontario people. This survey includes over thousands of respondents from all over the Ontario region. The survey was conducted online in which we would send email and letters to citizens that are willing to participate in the survey. We included the purpose and the link to allow respondents to access the survey. To ensure that each response is unique and to make sure respondents don't do the survey repeatedly. The survey maybe accessed only by each unique email address once. From our responses we have response that shows there are several factors that have an effective change on the number of children in Canada which includes religion, marital status (divorced, married), in which we can see religion has had one the most significant effects on the number of children in a family. Lots of family think that religion has played a huge part in deciding how many children they want and have. This would mean that in order to write policies it would important for policymakers to consider all these different factors. From the survey we have obtained, it was observed that citizens responding towards questions regarding marriage may include response bias. This maybe due to some might not feel comfortable answering certain sensitive quest.

## 2 Data

The data is analysis and cleaned in R (R Core Team 2021) using readxl (Wickham and Bryan 2019), jenitor (Firke 2021), the tidyverse (Wickham et al. 2019), and dplyr packages (Wickham et al. 2021) in the Rstudio. Firstly, the data cleaning is conductd to get the tidy data. The EDA is done using statistical knowledge and the corresponding graphs are generated by ggplot2 (Wickham 2016). Finally, after all the analysis and

---

<sup>1</sup>[https://github.com/Qianyuan-Huang/STA304\\_paper3](https://github.com/Qianyuan-Huang/STA304_paper3)

conclusion, the paper is knitted in Rmarkdown.

## 2.1 Data Source

The report used data source of 20603 observations from 2013 Canadian General Social Survey(GSS), the dataset was obtained from the University of Toronto Computing in the Humanities and Social Sciences (CHASS) Database Centre and can be accessed by everyone with access to UofT library system. GSS General Social Survey was established in 1985 and the overall objectives of the program were gathering data on social trends in order to monitor changes in the living conditions and well-being of Canadians, and to provide information on specific social policy issues. The survey in 2013 was conducted with the Computer Assisted Telephone Interviewing (CATI).

## 2.2 Bias

There are various sources of bias in this survey. One of which is the dataset is from 2013. With events happening throughout the years including COVID-19. The data may not be very accurate. People after covid-19 may have a very large change in their perspective of value which will lead to different factors affecting the number of children (Hardy 2020). Not only the change of perspective but some may not be living together as a household anymore meaning there may be a change of the size of the household or the people living in the household. this can cause inaccurate information from respondents.

Since the survey conducted is completely anonymous, the ethicality of the survey is handled properly. However, it would also be hard for us to do make sure that the surveys are done correctly and who responded to certain questions as respondents have the option to skip the question. This may lead to lack of responds of certain group of people to specific questions. One of the bias may be the different people responding to the surveys. every year, we may have different portion of the sample responding to the survey, this may cause an imbalance sampling for the other years as the people responding to the survey is different. To ensure this doesn't affect the results of our analysis, we would try to keep track of the people responding and account for the trends in the respondents.

## 2.3 Data Overview and Cleaning

The raw data we used contain 461 variables with 20602 observation and we selected 81 variables and 12736 observations that are most representative into the project. Since what we want is a variable name and a variable definition, we first set up a dictionary to rename the variable to be more visualized. Then we delete all the observation with no response to the crucial variable such as total number of children in house, age, or age at the birth. Than we change some variable from strings into numbers. We mutate the question with answer "Yes" or "No" in to "1" or "0" and mutate the variable with "valid skip", "no response", "refusal", or "not stated" into "NA".

As for other variable with complicated answers such as main activity, we group them according to the activity category. We set all the job with proper salary as "Yes" and the activity such as "Looking for paid work" or "Household work" as "No".

## 2.4 Data Summary

Table 1: Table 1: Age Summary Table

min	Q1	median	Q3	max	mean	var
20.3	44.4	58.9	69.8	80	57.29376	231.3837

The summary of the age showed the minimum age of the respondent is 20.3 and the maximum age is 80, which shows the dataset satisfied the object target of the survey with the target respondents from 20-80. The median of the respondent is about 58 years old and with variance of 231, it tells that the age distribution is relatively average from 20 to 80.

Table: Table 2: Age at First Birth Summary Table

Table 2: Table 3: Total Number of Children Summary Table

min	Q1	median	Q3	max	mean	var
18	22.8	26.4	30.3	45	26.86411	29.34967

min	Q1	median	Q3	max	mean	var
1	2	2	3	7	2.36919	1.342134

The first summary shows the age when the respondents gave birth to their first child. With the minimum value 18 and maximum 45, the data from the age at first birth is normally distributed with the peak value around 26. Table 2 shows the summary for the total number of children in Canada. The number of children have the mean value of 2 and variance of 1.34, which represent the distribution of the dataset of children's age is peak at 2.3.

### 3 Result

#### 3.1 Distribution

Figure 1: Histogram of Total Number of Children

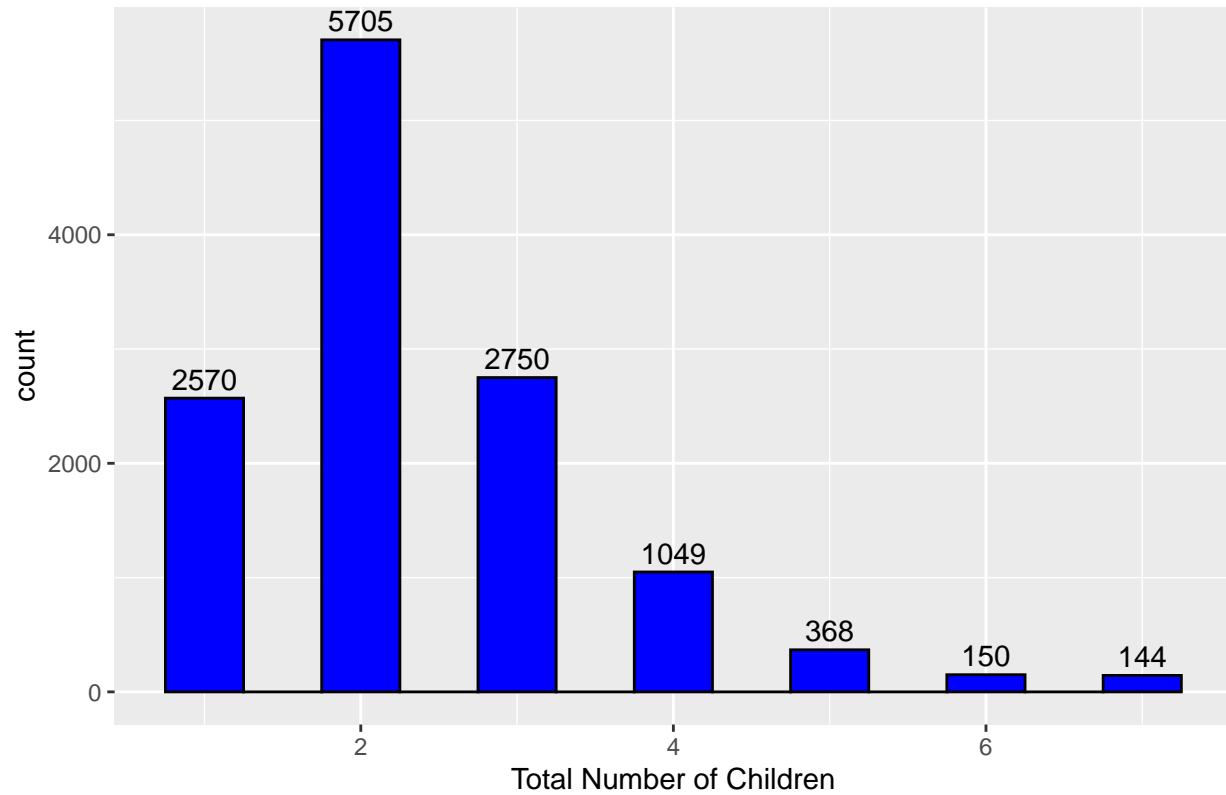


Figure 2: age at first birth distribution



Ordering from the largest frequency to the smallest frequency, the histogram distribution of total number of children can be summarized as:

About 45% of respondents (5705 out of 12736) have a total of 2 children in the family. About 22% of respondents (2750 out of 12736) have a total of 3 children. About 20% of respondents (2570 out of 12736) have a total of 1 child. About 8% of respondents (1049 out of 12736) have a total of 4 children. About 3% of respondents (368 out of 12736) have a total of 5 children. About 1% of respondents (150 out of 12736) have a total of 6 children. About 1% of respondents (144 out of 12736) have a total of 7 children. No respondents have a total of more than 7 children.

From the histogram of the distribution of age at child's first birth, we can see that this distribution is right skewed. About 10% of respondents had their first child at or before the age of 20. About 63% of respondents had their first child at or before the age of 30 and after the age of 20. About 25% of respondents had their first child at or before the age of 40 and after the age of 30. About 2% of respondents had their first child after the age of 40.

Statistics Canada (2014) reported that the average age at first birth in 2011 is 28.5 years. The mean age at child's first birth of our data is 26.86411, which also falls in the range of 25 to 30.

### 3.2 Province

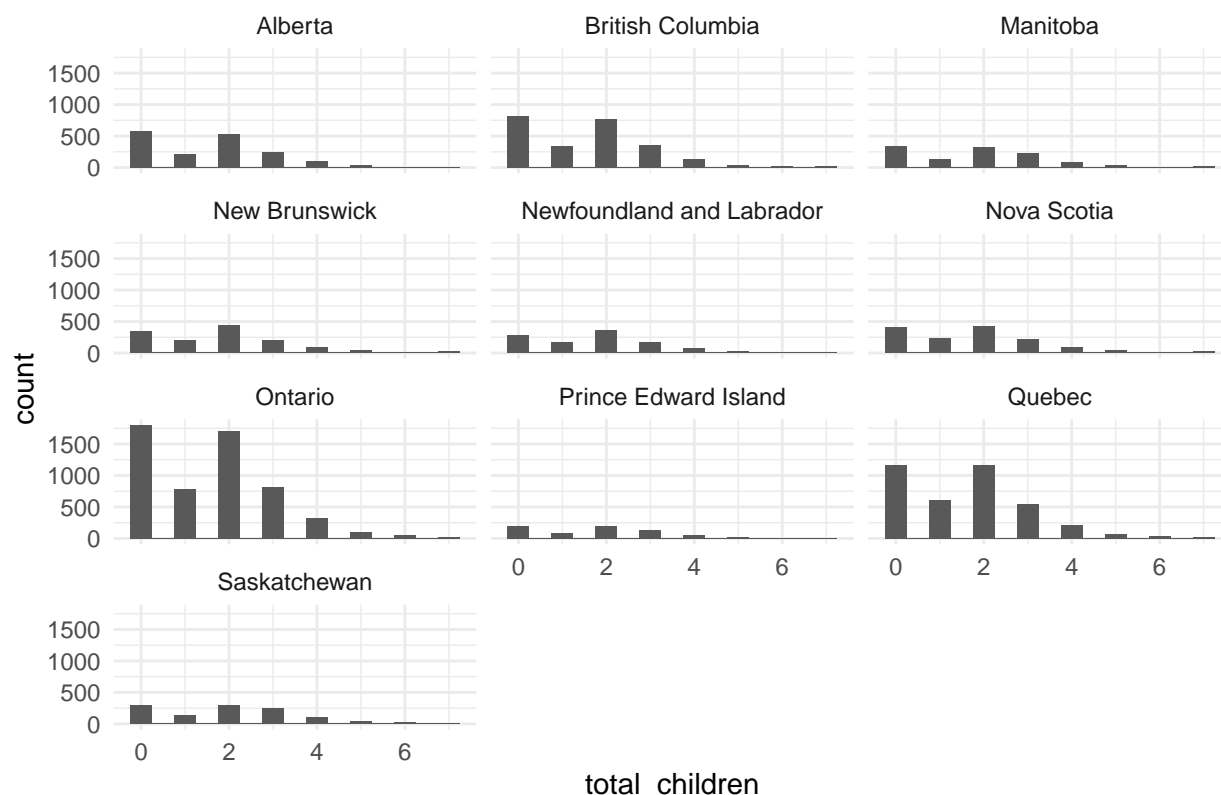


Figure 3: Total Number of Children Distribution in different Province

Grouping by provinces, ranked based on the proportion of respondents having no child, we can see that respondents from Alberta are the most likely to have no child; respondents from Ontario or British Columbia are the second most likely to have no child; respondents from Newfoundland and Labrador are the least likely to have no child.

Based on the modes of the total number of children in each province, we can see that respondents from Alberta, BC, Manitoba, and Ontario are more likely to have 0 child than any other number of children; respondents from New Brunswick, Newfoundland and Labrador, and Nova Scotia are more likely to have 2 children than any other number of children; respondents from Prince Edward Island, Quebec, and Saskatchewan are more likely to have 0 or 2 children than any other number of children.

### 3.3 Marital Status

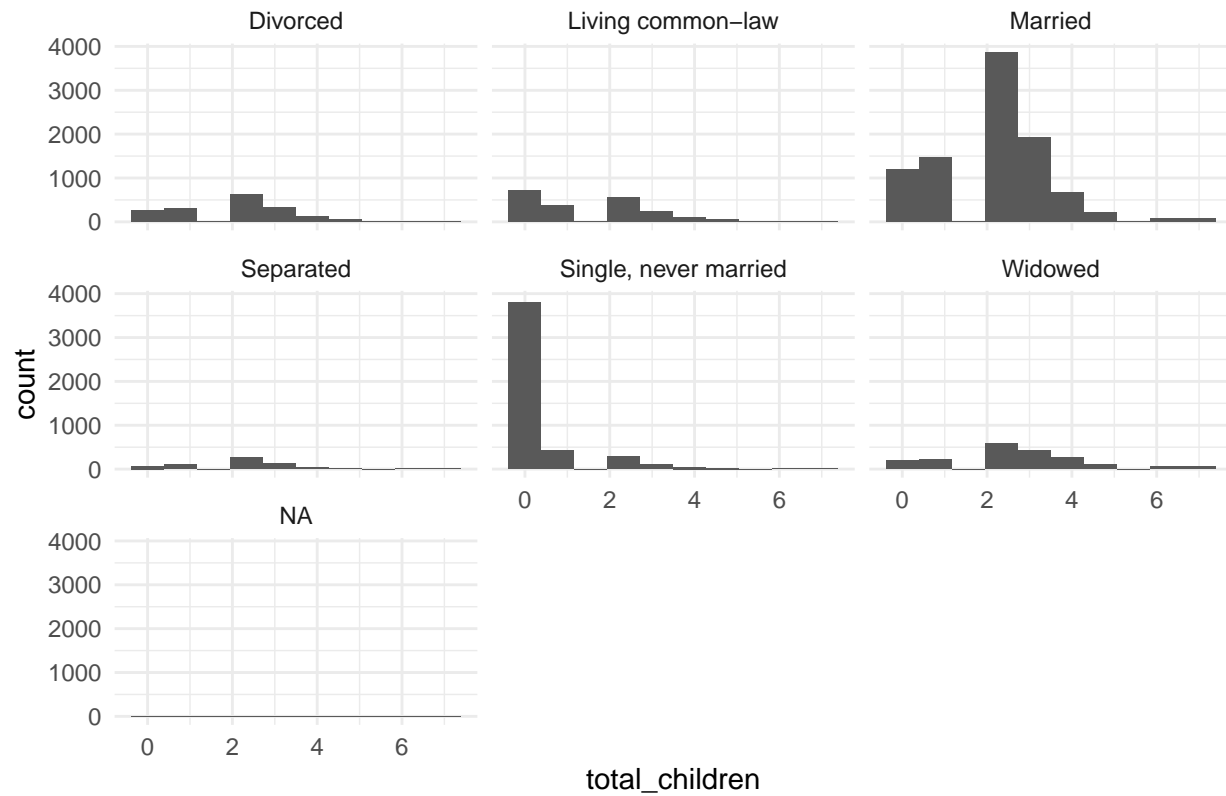


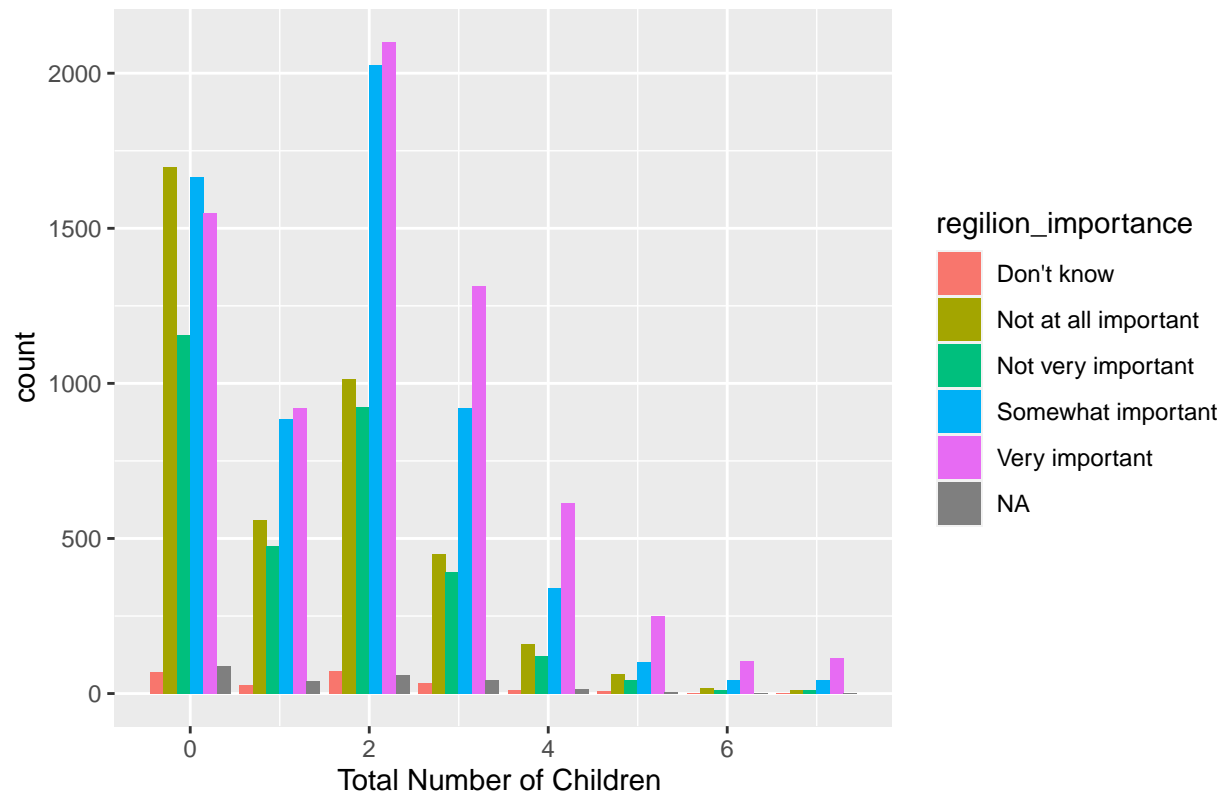
Figure 4: Total Number of Children Distribution for different Marital Status

Grouping by marital status, based on the proportion of respondents having no child, single respondents are the mostly likely to have no child; separated, widowed, and married respondents are among the least likely to have no child.

For divorced respondents, they are most likely to have 2 children and least likely to have 5 or more than 5 children. For living common-law respondents, they are most likely to have 0 child and least likely to have 5 or more than 5 children. For married respondents, they are most likely to have 2 children and least likely to have 5 or more than 5 children. For separated respondents, they are most likely to have 2 children and least likely to have 5 or more than 5 children. For single (never married) respondents, they are most likely to have 0 child and least likely to have 4 or more than 4 children. For widowed respondents, they are most likely to have 2 children and least likely to have 6 or more than 6 children.

### 3.4 Religious

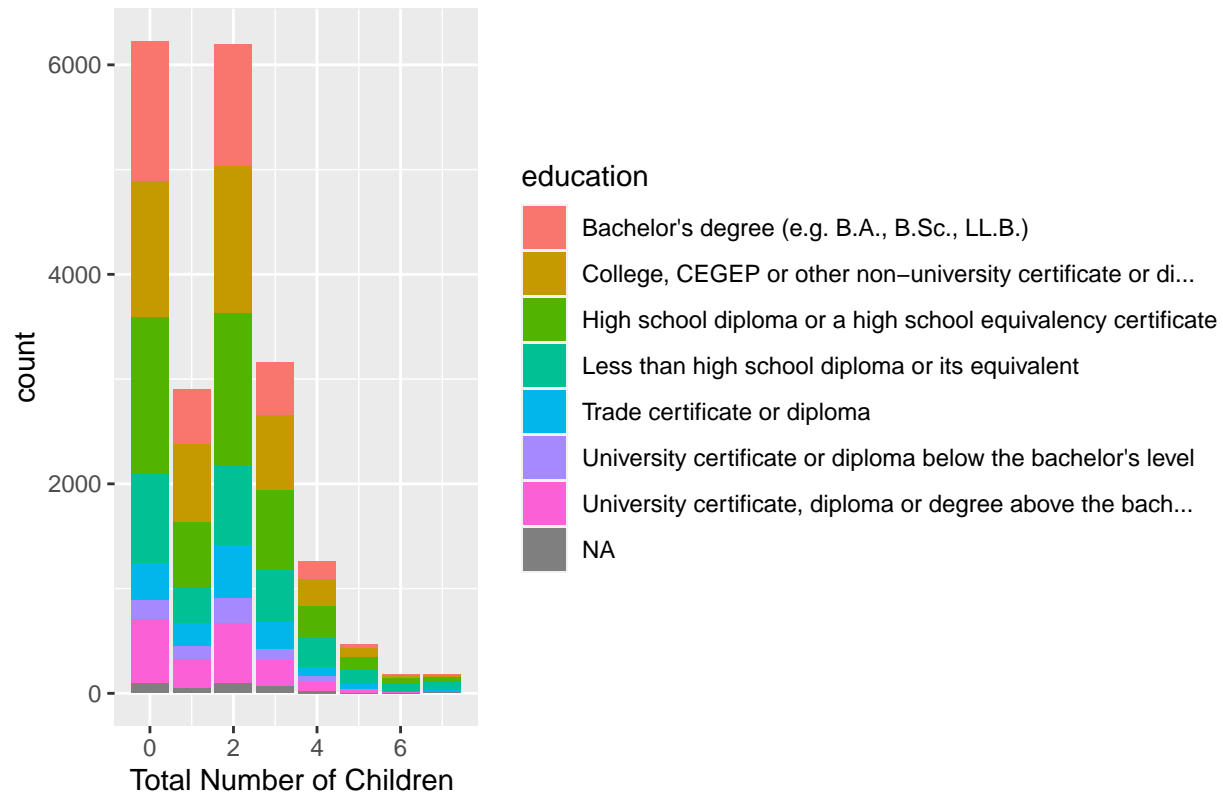
Figure 5: Total Number of Children distribution in different Religion



We can ignore respondents who do not know religion importance or who have provided no response. Among respondents with 0 child, most of them think that religion is not at all important and least of them think that religion is not very important. Among respondents with 1, 2, 3, 4, 5, 6, or 7 child/children alike, most of them think that religion is very important and least of them think that religion is not very important. However, the difference between the proportion of respondents considering religion to be very important and the proportion of respondents considering religion to be not very important is much larger for respondents with 2, 3, 4, 5, 6, or 7 children than that for respondents with 1 child.

### 3.5 Education

Figure 6: Total Number of Children distribution in different education

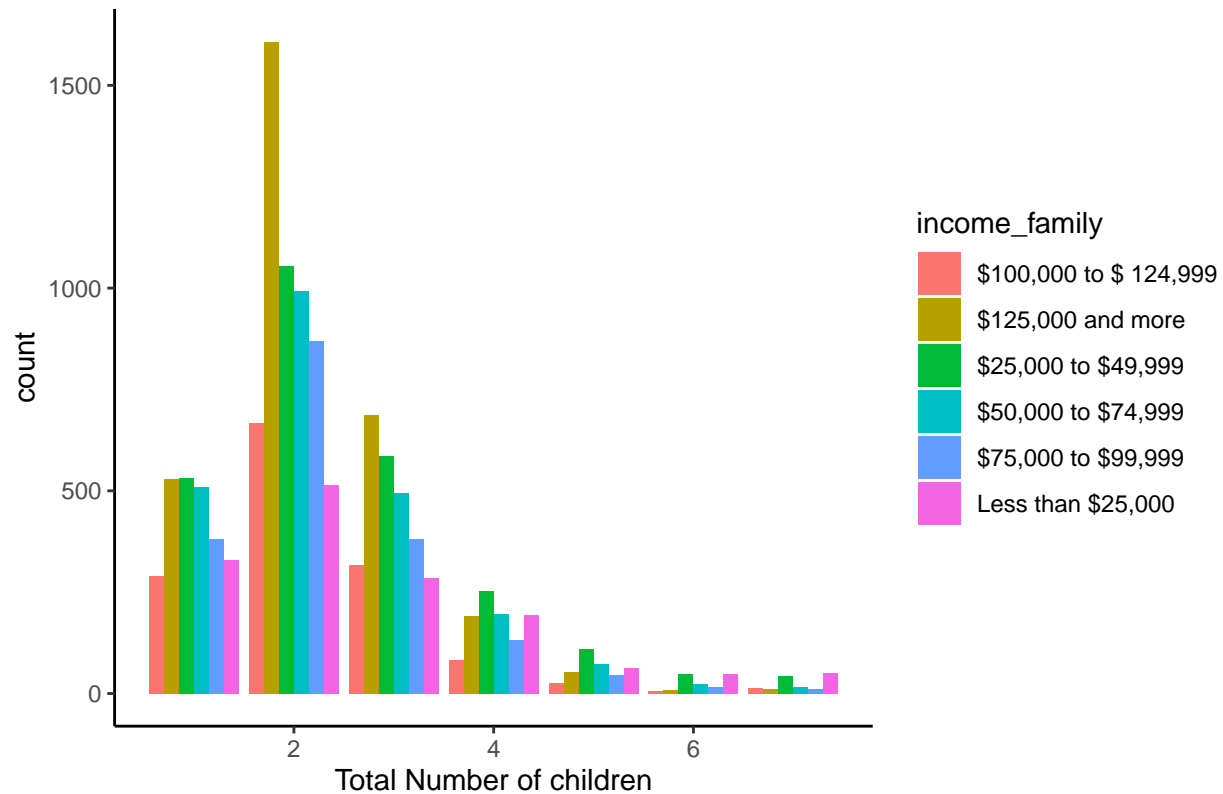


From Figure 6 we could tell that the relationship between the education degree and the total number of children is not significant with family have no more than 2 children. However, as for the family with more than 3 children, more than fifty percent of the respondents have highest education degree of high school diploma or less than high school diploma. Almost all the respondents who have more than 5 children in the house have no more than high school diploma. It is clear that as the education degree grows, the intention of having more than two children would drop significantly. The higher education people take, the less likely they intent to have more than two children.



### 3.6 Family Income

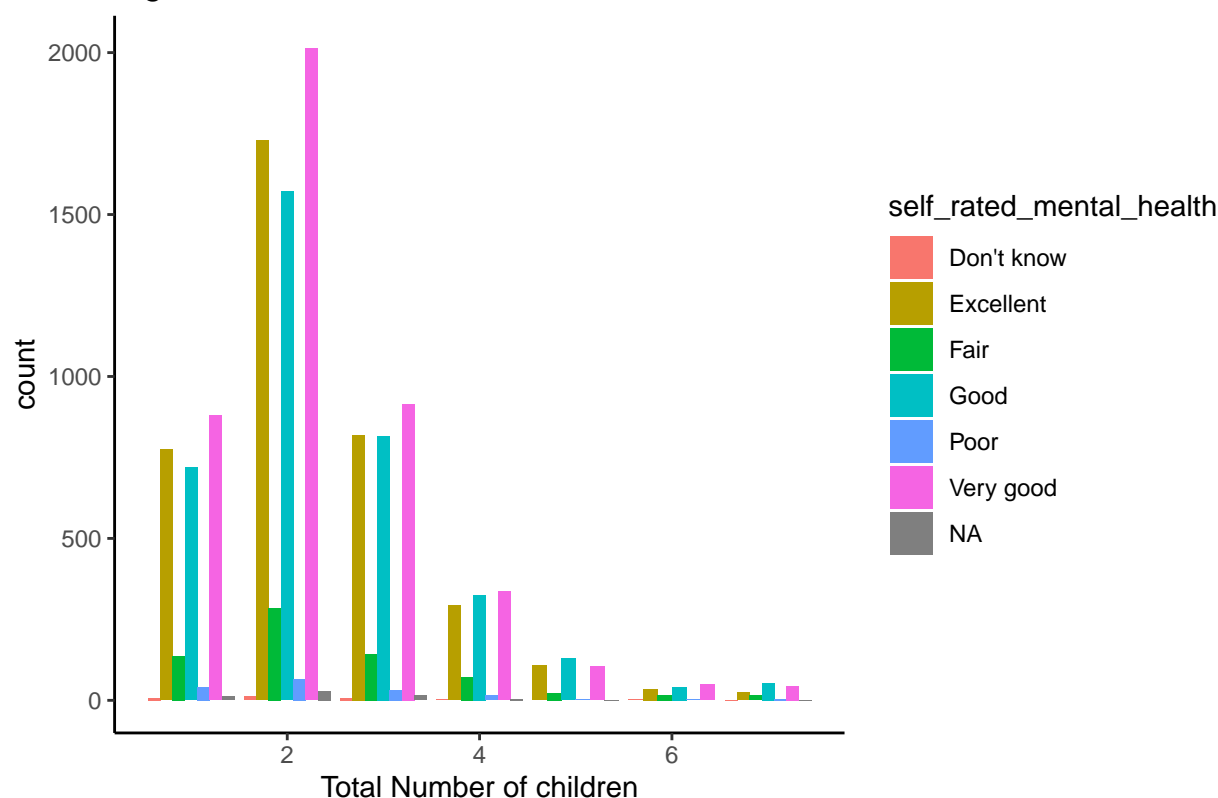
Figure 7: Total number of Children and Family Income



From figure 7 we could tell the distribution for the total number of children with family income follow the basic normal distribution. However the respondents with high income intend to have 2 children and the respondents with relatively low income intend to have more than 4 children. This result correspond with the result we obtained from the relationship between numebr of children and education. The people with relatively weak education background are more likely to have more kids. Due to the lack of education, the family income could be relatively lower than the highly educated repondents. Therefore, Figure 7 shows the similar result with Figure 6.

### 3.7 Self-rated Mental Health

Figure 8: Total number of Children and Mental Health



The result can be separate into two parts. The first part is for the respondents with fewer than 4 childrens. People with two to four children in the house tend to be in a relatively better mental health condition than people with single child. As for the second part, respondents with 5 or more children seems to be in a worse condition than the respondent in the first part. In the part, the people with more children tend to be in a poorer mental health condition.

## 4 discussion

### 4.1 Marital Status

The question gives respondents 6 options to choose from. In the graphs we have created, households that have married couples seem to have a large proportion of all the respondents. They also seem to have the most children on average as over 3500 household that are married have 2 kids. We are also able to clearly the impact of marital status have to the number of children as single house has the most counts of having no children. Contrastly, married households tend to have more children (Canada 2014).

However, the number of respondent having children without marriage are still relatively large. The divorced respondents have about 12.5% of the total number of kids of married respondents. But according to the official number from Statistics Canada, the divorce rate is about 10 percent (Stone 2018). We could find out that people tend to divorce more after giving birth to the first child(Canada 2021).

### 4.2 Religious

The question in the survey gives 6 options to respondents to choose from. As we can see from the graph, except households with 0 kids we are able to see households with 1-6 or more children think that religion is very important and that it does affect how many children they decide to have. We can also see the proportion

of respondents thinking that religion is somewhat important and very important takes up more than 60 percent of the responses meaning that we can conclude religion is a factor that affects the number of children families have.

Another weakness is with regards to how religion affects the number of children. We were not able to ask what religion in particular they trust as the question only asked how they personally think religion affected them. In further investigations, we would add more questions about religion for example how long or what religion they believe in. These questions can make our analysis more comprehensive and trustworthy.

### **4.3 Education**

The choice in investigating the education background is not accurate. Degree from different university could represent different education quality (Christopher 2021). Although degree seems to be the most reasonable way to represent the education background, it would not be able to reveal the real education quality.

### **4.4 Income**

This question in the survey only gives 6 choices. The respondent with extremely high income or low income was excluded from the survey. From the Figure we can find out a strong relationship between the income and number of children. It is generally believed that household with more children are more likely to have more children as they have the ability to raise them. However, the Figure gives an opposite result. The people with relatively low income tend to have more children than the rich, which could be caused by lack of education (Hung-Lin 2005). The respondent with high income rather have two children than one or three children. Therefore, two children might be the most suitable number of children in a family from the perspective of this survey.

### **4.5 Mental Health**

The mental health in this survey was self-rated, which could not be perfectly accurated with no official assessment. Since the choices of “Don’t Know”, “Refusal”, or “Valid Skip” were given in the telephone based survey, respondents with self-rated poor mental health condition might not be willing to answer the question honestly (Fiona 2009). They might choose to skip the question in order not to reveal their own privacy. Therefore, the relationship of mental health condition and number of children might be overestimated in the report.

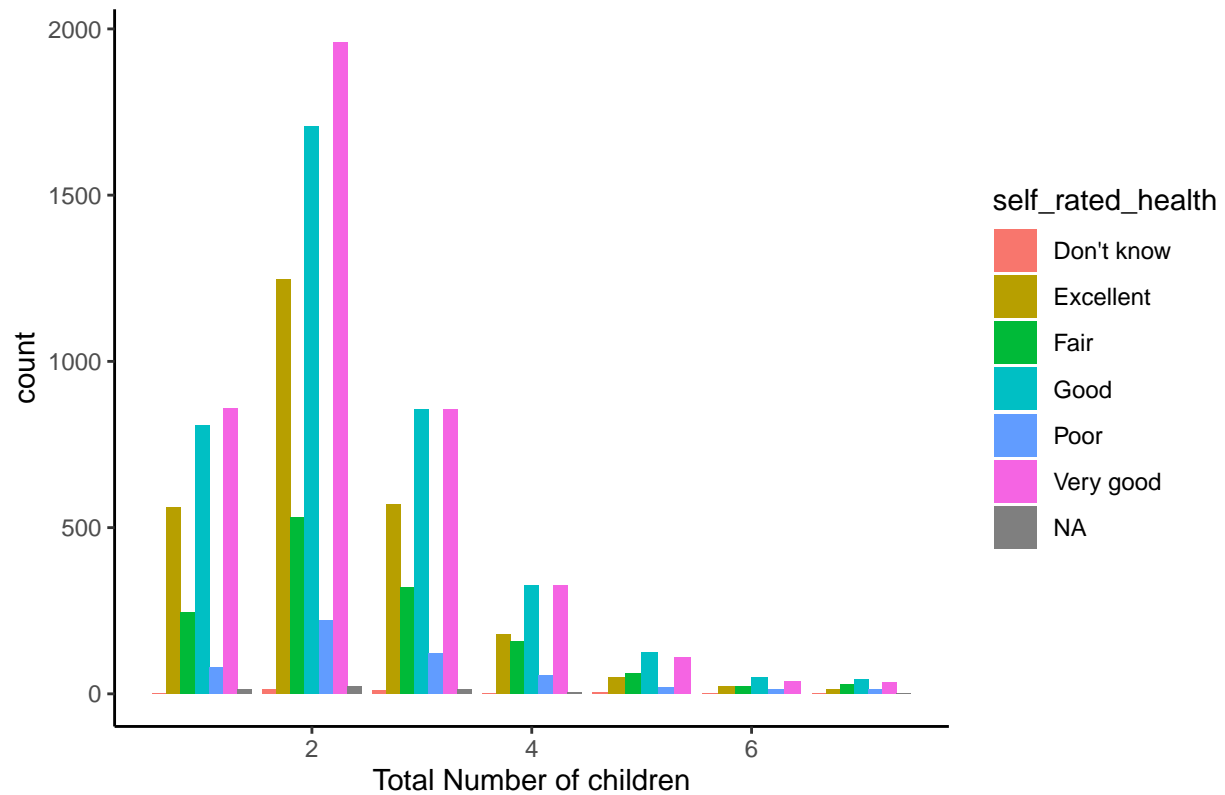
### **4.6 Weakness and Next Step**

As with any surveys, our analysis has some weaknesses as well. Covid-19 has in no doubt shaped a lot of peoples’ life upside down and changed them. Our data in 2013 may not accurately predict what people think now. Furthermore, it is very hard for our interviewers to get in contact with respondents at the moment due to both covid-19 and the household being different as before. For our further survey, we will only be able to do the survey online by sending out emails which will result in a larger proportion of younger people responding, shifting the original weight to a young respondent heavy survey. The General Social Survey we used in this report is from 2013, when the methodology used at the time was Computer Assisted Telephone Interviewing (CATI). Although the collection costs of such method is lower than in-person interview, the reliability of the dataset was not high enough. The households with no telephone coverage or young people with only cellphones were excluded in the survey. Comparing to face-to-face interview, the response rate tend to be lower in the telephone interview. There are even limitations on the amount and type of data which can be collected.

## 5 Appendix

### 5.1 Self-rated Health

Figure 7: Health and total number of children



### 5.2 Supplementary Survey

Our supplementary survey is available here: <https://forms.gle/1vdgHpsiy9b5GbKi8>

Did you plan ahead when deciding to have children?

☐ Yes

☐ No

☐ Somewhat

Did you have a preference on having a boy or girl?

☐ Girl

☐ Boy

☐ No preference

Do you think marital status affects the growth of a child?

☐ Yes

☐ No

☐ Depends

Does the growing environment(neighborhood) affect the growth of a child?

☐ A lot

☐ Slightly

☐ No

Figure 1: A caption

Which aspect in a neighborhood do you think is most important for a child

- ☐ Hygiene
- ☐ Safety
- ☐ Education(Good schools)
- ☐ Crime rate

Do you think going to a private(good) school is important?

- ☐ Yes
- ☐ No
- ☐ Maybe

How many children do you plan to have?

- ☐ 0
- ☐ 1-2
- ☐ 3-4
- ☐ 5 or more

What do you think is the best age to have children?

- ☐ 18-25
- ☐ 26-30
- ☐ 31-35
- ☐ 36 or older

Figure 2: A caption

Would you force your children get good grades?

☐ Yes

☐ No

Would you force your children to develop certain hobbies or learn something?

☐ Yes

☐ No

Do you think growing up in a big city(toronto, Vancouver) is better or in a small city

☐ Big City

☐ Small City

How important do you think education is to a child? (rate from 1 -5 where 1 means less important)

☐ 1

☐ 2

☐ 3

☐ 4

☐ 5

Figure 3: A caption

How much money do you think it will take to raise a child?

您的答案 \_\_\_\_\_

Have you ever thought of giving up your children?

☐ Yes

☐ No

☐ Do not wish to respond

Is there anything you would want to share on children and family?

您的答案 \_\_\_\_\_

Has religion affect on the number of children you plan to have?

☐ Yes

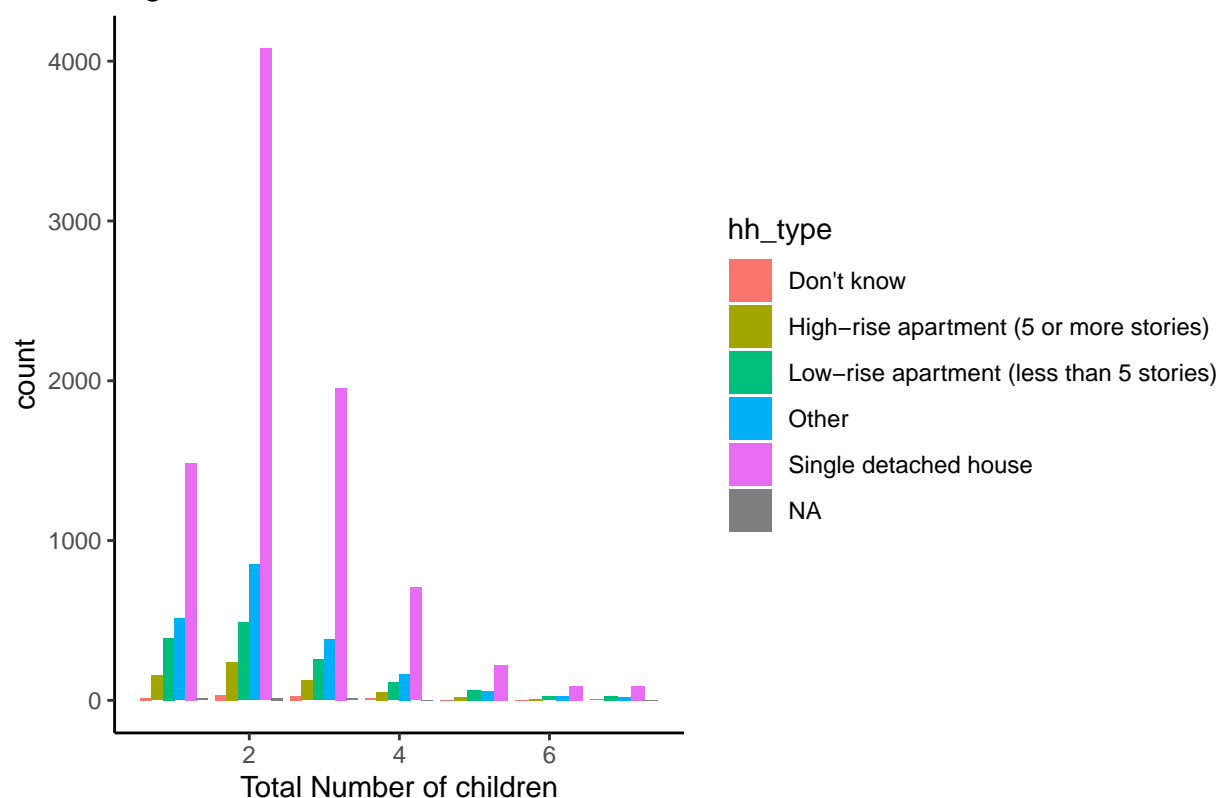
☐ No

Figure 4: A caption



### 5.3 Household

Figure 7: Household and total number of children



## Reference

Canada, Statistics. 2014. “Fertility: Fewer Children, Older Moms.”

———. 2021. “Live Births, by Marital Status of Mother.”

Christopher, J. 2021. “When Salt Ain’t Enough: A Critical Quantitative Analysis of Special Education and Education Degree Production.”

Fiona, Mensah. 2009. “Parents’ Mental Health and Children’s Cognitive and Social Development: Families in England in the Millennium Cohort Study.”

Firke, Sam. 2021. *Janitor: Simple Tools for Examining and Cleaning Dirty Data*. <https://CRAN.R-project.org/package=janitor>.

Hardy, J. 2020. “Connection, Contagion, and Covid-19.”

Hung-Lin, Tao. 2005. “The Effects of Income and Children on Marital Happiness - Evidence from Middle- and Old-Aged Couples.”

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

Stone, Lyman. 2018. “No Ring, No Baby: How Marriage Trends Impact Fertility.”

Wickham, Hadley. 2016. *Ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York. <https://ggplot2.tidyverse.org>.

Wickham, Hadley, and Jennifer Bryan. 2019. *Readxl: Read Excel Files*. <https://CRAN.R-project.org/>

package=readxl.

Wickham, Hadley, Mara Averick, Jennifer Bryan, Winston Chang, Lucy D'Agostino McGowan, Romain François, Garrett Golemund, et al. 2019. “Welcome to the tidyverse.” *Journal of Open Source Software* 4 (43): 1686. doi:10.21105/joss.01686.

Wickham, Hadley, Romain François, Lionel Henry, and Kirill Müller. 2021. *Dplyr: A Grammar of Data Manipulation*. <https://CRAN.R-project.org/package=dplyr>.