The Futuer of Canadian Population: What Encourages Population Growth

Hyunseok Rha

2020-10-19

The Futuer of Canadian Population: What Encourages Population Growth

Hyunseok Rha

2020-10-19

Abstract

With Statistics Canada's General Social Survey 2017 data, we have found that an individual's number of children is positively affected by these 3 factors; age, as an individual's age goes up he/she tends to have more children, if an individual was born in Canada then he/she tends to have more children than who doesn't and as individual's family income goes up, he/she tends to have more children than who does not have higher family income.

Introduction

As of 2020, the estimated world population is about 7.8 Billion @wor. It may seems overwhelming, but the annual population growth rate is decreasing. And for the most of the developed contury, the number of population is starting decreasing instead of increasing.

For example, Japan has around 5.1 trillion US dollars of Gross Domestic Product(GDP) for 2019 `[1]`(https://tradingeconomics.com/japan/gdp) and Japan has 3rd largest GDP in the worl. However, Japanese population was at peak at 2010 with around 128 million and as of 2020 the population is around 126 million which is decreased about 2 million over the last 5 years. https://www.worldometers.info/world-population/japan-population/

On top of that, for Japan's median age of population was increased from 44.7 for 2010 to 48.4 for 2020, which indicates that the over all population is getting old and the rate of aging is higher than the rate of new born.

> Japan's natural population decline is among the fastest in the world although some European countries with high emigration ??? such as Romania and Bulgaria ??? are shrinking even faster.

https://www.ft.com/content/29d594fa-5cf2-11e9-9dde-7aedca0a081a

Then what about Canada?

What makes Canadian to have more children and leads to population growth?

In this paper,I will use Statistics Canada's General Social Survey data for 2017 to figure out which factor affects individual's number of children.

Data

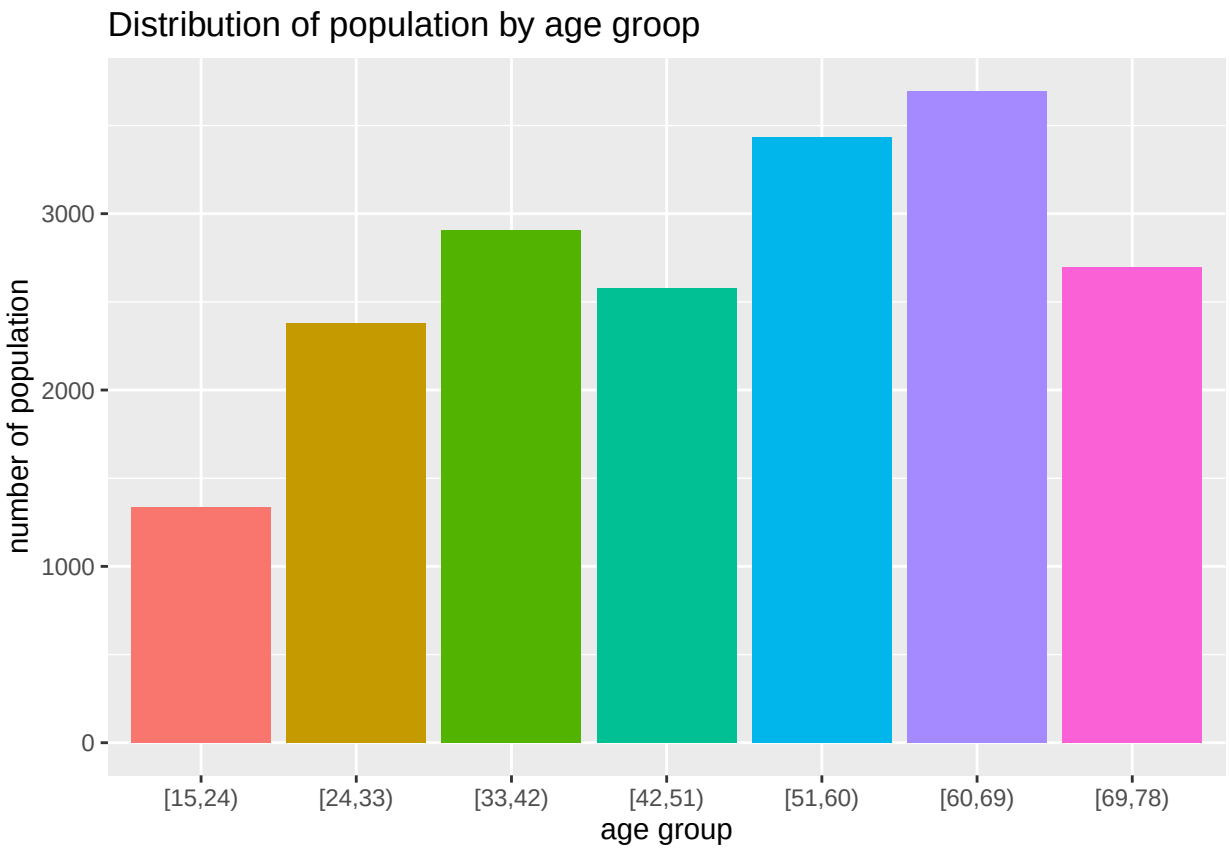
The data is being used for this paper is `General Social Survey`(and I will call it as `GSS` from now on) data for 2017 provided by Statistics Canada that has all non-institutionlaized persons 15 years of age and older, and living in the 10 provinces of Canada as its' target population. `GSS` has two primary objectives that. 1) To gather data on social trend in order to monitor changes in the living conditions and well-beings of Canadians over time; 2) To provice informations on specific social policy issues of current or emerging inters. And since

`GSS` is well collected data that covers large Canadian population and many different areas of interes, it is a perfect data to work with to answer my question `What encourages population growth?`

The data I have used is a cleaned version of original `GSS` 2017 data that the name of columns are changed into more intuitive names and the numer labeled values are converted into more human friendly names such like for the age column, from values of `1, 2` to `15 to 19 years, 20 to 24 years`.

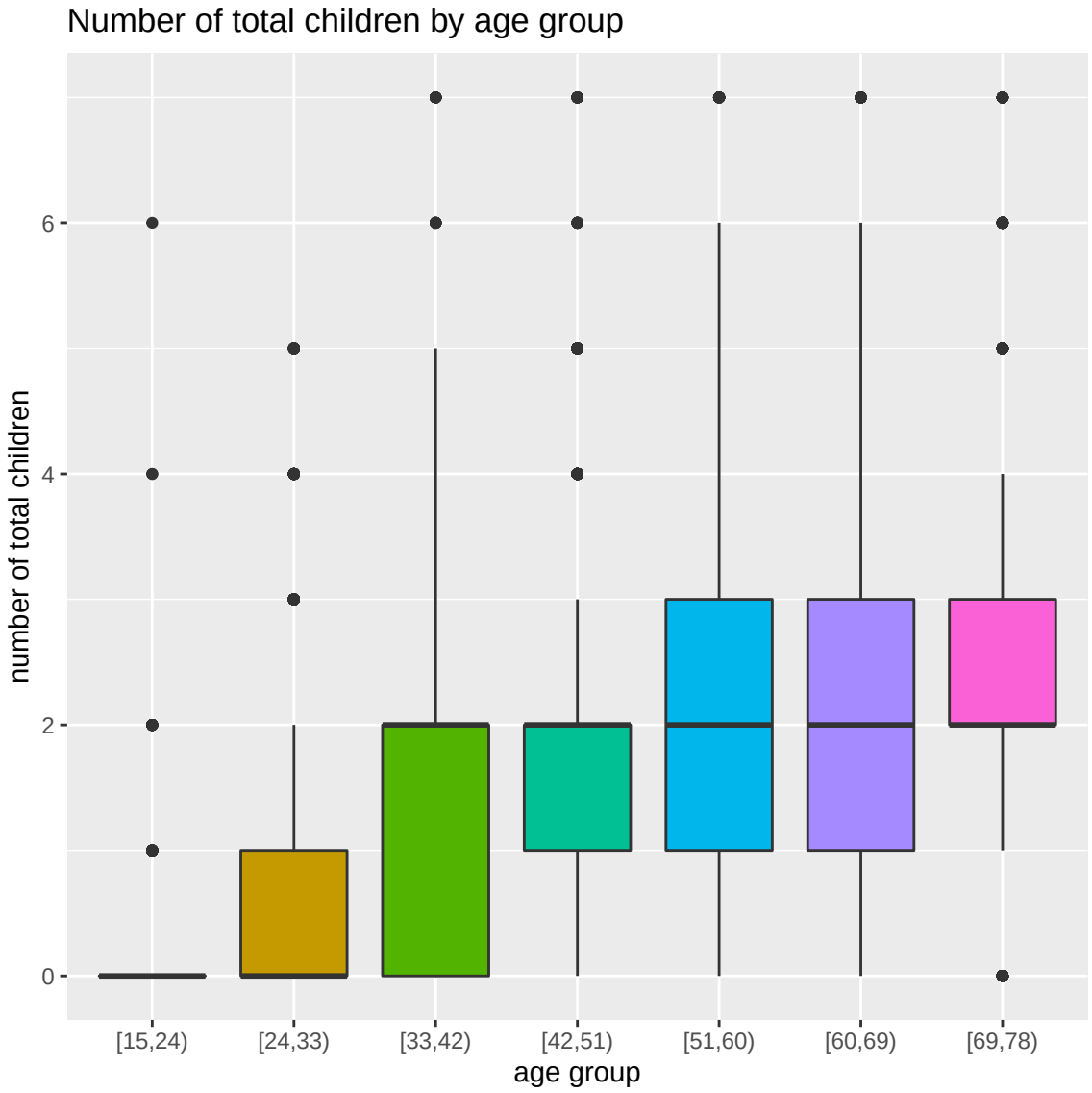
`GSS` 2017 data contains total 20602 entries with 81 columns.

For example, we can see the distribution of population based on their age group.

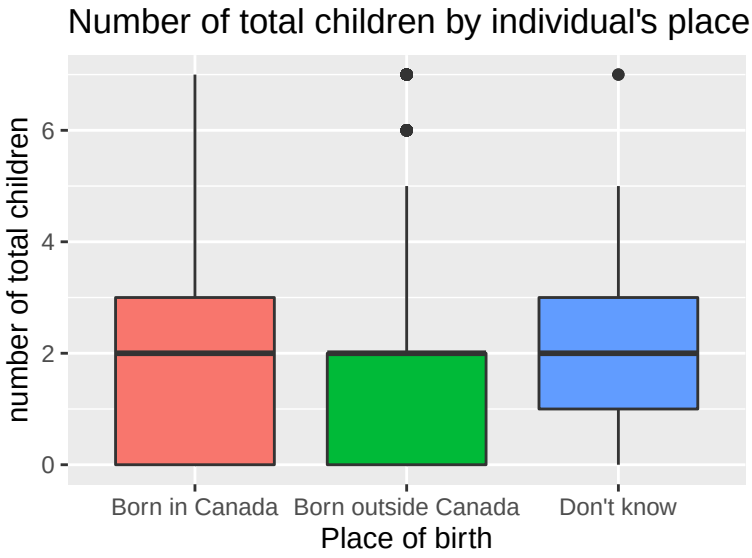


We can have some insights about what affects the number of total children based on different factor. And here's some graph to show how each factor affects the total number of children.

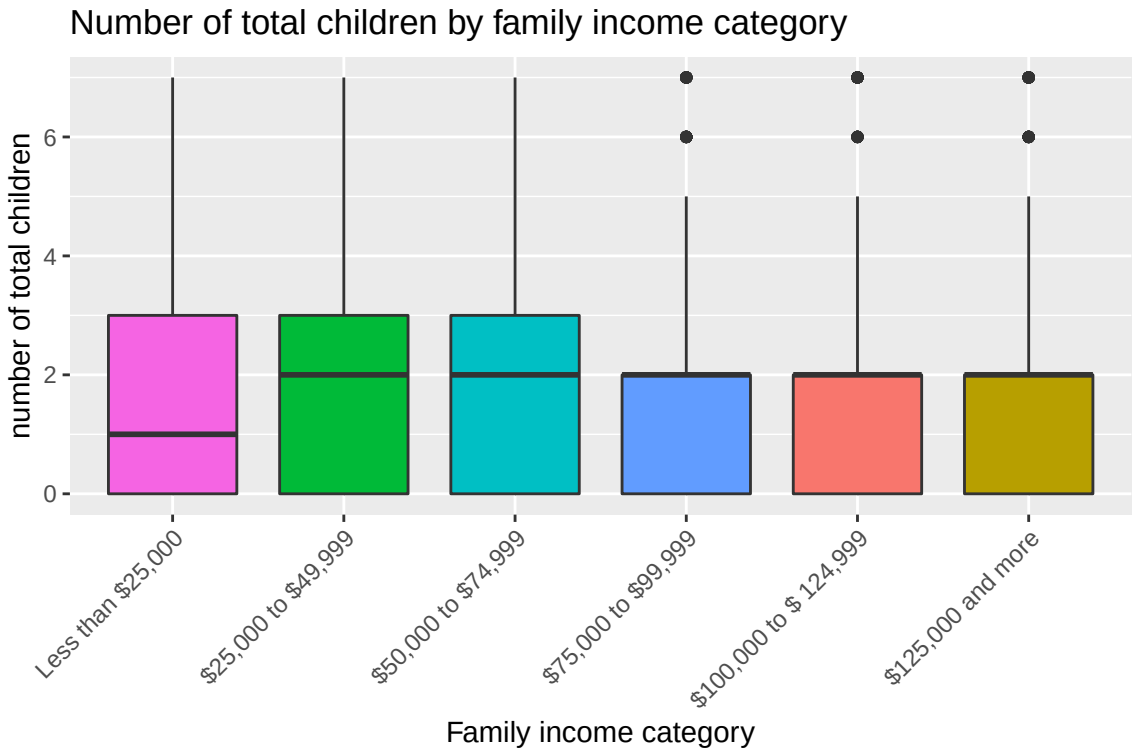
For the age group,



For the individual's place of birth,



For the family income category,



Model

The model will be used is survey weighted regression.

Since the `GSS` is a survey data, we can think the entire Canadians as population and `GSS` data as sample.

The general form of the model looks like $y_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \cdots + \beta_k x_{ki} + \epsilon_i$ where y_i is the value of the study, in our case it is the total number of children. And each β_k represents the factor that affects the total number of children.

For our model, we will use 5 parameters such as age, place_birth_canada, income_family, self_rated_health and self_rated_mental_health to estimate the number of total children and see what is the relationship between each parameter and the total number of children.

Note: the number of population is from https://www150.statcan.gc.ca/n1/pub/12-581-x/2018000/pop-eng.htm

```
## MODEL INFO:
## Observations: 20448
## Dependent Variable: num_of_children
## Type: Survey-weighted linear regression
##
## MODEL FIT:
## R² = 0.21
## Adj. R² = 0.21
##
## Standard errors: Robust
## ------------------------------------------------------------------------
##                                          Est.   S.E.   t val.      p
## ---------------------------------------- ------- ------ -------- ------
## (Intercept)                               0.43   0.29     1.50   0.13
## age                                       0.04   0.00    76.14   0.00
## as.factor(place_birth_canada)Born        -0.05   0.02    -2.25   0.02
## outside Canada
## as.factor(place_birth_canada)Don't        0.87   0.23     3.79   0.00
## know
## as.factor(income_family)$125,000          0.06   0.03     2.00   0.05
## and more
## as.factor(income_family)$25,000          -0.28   0.04    -8.05   0.00
## to $49,999
## as.factor(income_family)$50,000          -0.24   0.04    -6.73   0.00
## to $74,999
## as.factor(income_family)$75,000          -0.08   0.04    -2.32   0.02
## to $99,999
## as.factor(income_family)Less             -0.35   0.04    -8.76   0.00
## than $25,000
## self_rated_healthExcellent               -0.52   0.22    -2.31   0.02
## self_rated_healthFair                    -0.53   0.23    -2.33   0.02
## self_rated_healthGood                    -0.53   0.22    -2.39   0.02
## self_rated_healthPoor                    -0.44   0.23    -1.91   0.06
## self_rated_healthVery good               -0.58   0.22    -2.58   0.01
## self_rated_mental_healthExcellent        -0.14   0.21    -0.70   0.48
## self_rated_mental_healthFair             -0.03   0.21    -0.16   0.87
## self_rated_mental_healthGood             -0.03   0.21    -0.16   0.87
## self_rated_mental_healthPoor             -0.19   0.22    -0.85   0.40
## self_rated_mental_healthVery             -0.11   0.21    -0.52   0.60
## good
## ------------------------------------------------------------------------
##
## Estimated dispersion parameter = 1.75
```

Results

Now the model is avaliable and we have an insights for each parameter that how they affect the total number of children.

Let's have a look at each of them to interpret the result.

age

We can see from the model output that as an individual's age goes up by 1, the number of total children goes up by 0.04. And it makes sence in general that the change of having a child is depending on individual's age.

place_birth_canada

It is quite interesting that, for the `place_birth_canada`, response `Don't Know`(which is equivalent to `Born in Canada`) has quite significant impact that the number of total children increases by 0.87 if it is the case.

income_family

We can see as individual's family income increase, the estimated number of total children increase. The difference between the lowest income category versus highes income category is around 0.41

self_rated_health

We can see when an individual's self rated health is poor it has more positive impact on the total number of children than when the health excellent.

self_rated_mental_health

We can see when an individual's self rated metal health is poor it has more negative impact on the total number of children which is quite different behaviour compare to the self_rated_health metric.

The factors that have positive impact on the total number of children are higher age, born in Canada and high income.

Discussion

???

Weaknesses

The weakneeses of my analysis are that the number of factor is limited and treating total number of children with all the age group can add noise to the result.

As we can see, age acts as a significant factor that affects total number of children and we can have more percise insight when we break down the reponse into age group and do the analysis for each of them.

Next Steps

The next step will be, change the way of analysis that we analyze the response for each age group.On top of that, we can do the similar analysis with different year's `GSS` data. This will give us insight that how's each factor's impact has been changed overtime such as modern people's number of children tend to be affected by family income.

Also with different year's data, we can also predict how's the number of population would change in future based on data.

References

[1] https://tradingeconomics.com/japan/gdp

https://www.worldometers.info/world-population/