Relationship between education and life satisfaction

1002564664

October 19, 2020

Relationship between education and life satisfaction

1002564664

October 19, 2020

Abstract

I explored the dataset of "General Social Survey", cycle 31, which is focused on family, to analyze relationship between education levels and life satisfaction among Canadian citizens. Statistical model such as simple linear regression is used to show how education combined with income and marital status affecting respondents life satisfaction. The results lead to the conclusion that higher education levels influence people life satisfaction indirectly due to higher earnings. In addition, based on the key findings, governments, both federal and local, should encourage people to continue their education.

Introduction

The study of life satisfaction among citizens in developed nations is becoming increasingly important for policy-makers. With increased economic prosperity, people are more focused on other aspects of their lives, such as mental health and life satisfaction. One of the most important and guaranteed way for prosperity is the access to higher education. So it is important to check relationship between education levels and socio-economic status. The expansion of education has often been proposed as a priority in developed nations. The quality of education provided is as important as the access for it, especially among the poorest parts of the population.

Canada has quite a unique place in OECD (Organisation for Economic Co-operation and Development) reports due combination of high quality of life and high education attainments among OECD members. But it has its own drawbacks. More people now have access to higher education, thanks to the loans and grants. However, with increased education attainment, there are also more supply then demand for University-educated people. It's not uncommon for someone with Master degree to work on low-level jobs, due to oversaturation of the chosen field. Such big difference between reality and expectations harms people mental health and overall life satisfaction. So, I want to find the relationship between life satisfaction and education. Given that the life satisfaction may be income related, I will use GSS recorded data about Canadians, including their age and residence.

Throughout this report, I am going to figure out some key factors on how education is affecting someone's life satisfaction. Shortly, in the following sections, I will use statistical methods for data analysis to build a regression model of life satisfaction, and interpret which regression outputs effect the most. To further support my analysis, I will include graphical visualizations such as plots. More detailed information about the data analysis will be explained in Data and Model sections.

Data

I will use the General Social Survey(GSS), cycle 31. The data set contains responses of the General Social Survey from 2017. The contents of the survey are focused on family characteristics in Canada. For example, their health conditions, life satisfaction, education levels, economic status and many others.

Data downloaded from the CHASS website, which is available for all University of Toronto students and employees. Data will help to analyze some main factors affecting Canadians life satisfactions, and how education level is significant part for them. Linear model is used to determine the significant factors with plots are drawn to show the relationship between variables.

Taken directly from GSS 2017 summary, the target population includes all non-institutionalized persons 15 years of age and older, living in the 10 provinces of Canada. The frame population is everyone who is registered by cellular and landline phones in Canada. The sampled population is everyone reached by phone. The 10 provinces of the target population are divided into 27 strata. Simple random sampling without replacement was performed in each stratum.

From the summary of the GSS report, the primary goal of the General Social Survey is to monitor the well-being of Canadians over time. In addition, GSS 2017 is the most recent available survey and it has new section "Health and subjective well being", which is important for our data analysis.

The main characteristic of data set is that the majority of the variables are categorical. Such data is limiting what statistical models we can use. This is mainly because quantitative analysis cannot be performed on such data, without transforming it. Quick summary also show that some observations are NA(not available) or missing. Excluding or including them into analysis might affect our linear model, since some of them might be influential points or outliers.

The original data set was cleaned and modified due to its size and number of variables.

Model

I will use simple linear regression model instead of multiple linear regression (I just don't know how to interpret results from MLR). In addition, simple linear regression provides numeric output, which is easier to interpret rather than logistic regression output.

In this study, The responding variable that I will analyze is "feelings_life", which is a numeric variable that measures the life satisfaction of each individual. I will test it against "education", "income_respondent", and "marital_status" and separate fourth model, "income_respondent" versus "education" to see extra relationship.

```
##       (Intercept)      educationSchool educationUniversity
##        8.14798851          -0.12831637          0.02291961

##                          (Intercept)  as.factor(income_respondent)25000
##                            7.8933248                          0.1806960
##  as.factor(income_respondent)50000  as.factor(income_respondent)75000
##                            0.3284857                          0.4302850
## as.factor(income_respondent)110000 as.factor(income_respondent)125000
##                            0.4623334                          0.6033639

##       (Intercept)    marital_statusL    marital_statusM    marital_statusS
##       7.702072539        0.518384264        0.731260794       -0.004246452
## marital_statusSep    marital_statusW
##      -0.432176345        0.254803918

##       (Intercept)      educationSchool educationUniversity
##          44355.364          -17102.905           5860.816
```

Results

Through the analysis, We find positive linear relationships between a dependent variable, life_satisfaction rate and University level education. Based on the Simple Linear Regression, for University degree, there is an increase 0.02 increase in life satisfaction. In addition, there is a significant increase in life satisfaction for income higher than 125000. For such income bracket, life satisfaction is increased by 0.6 points. Worth to mention, that there is a positive relationship between income and University education. Difference is quite big, for University education, income increased by 5860.816, while for School and lower is decreased by 17102.905. There is also high satisfaction rate for Married couples, where it's 0.73.
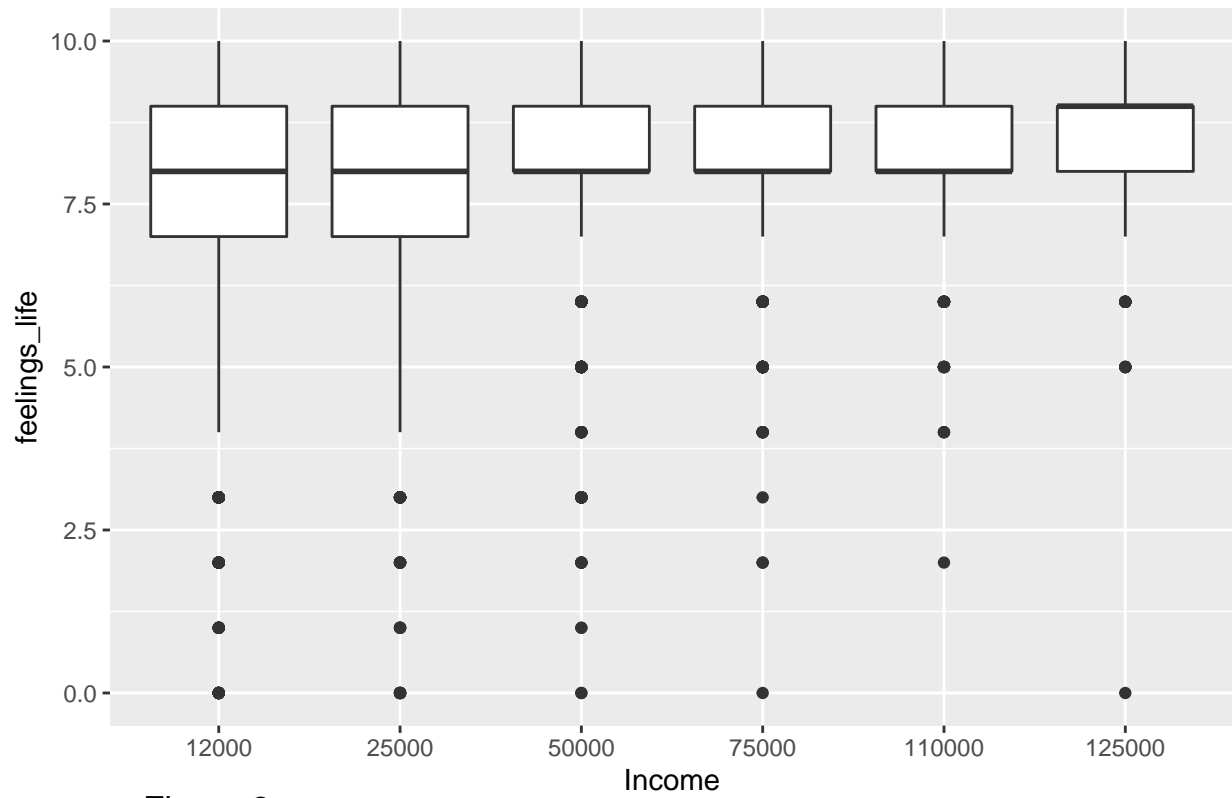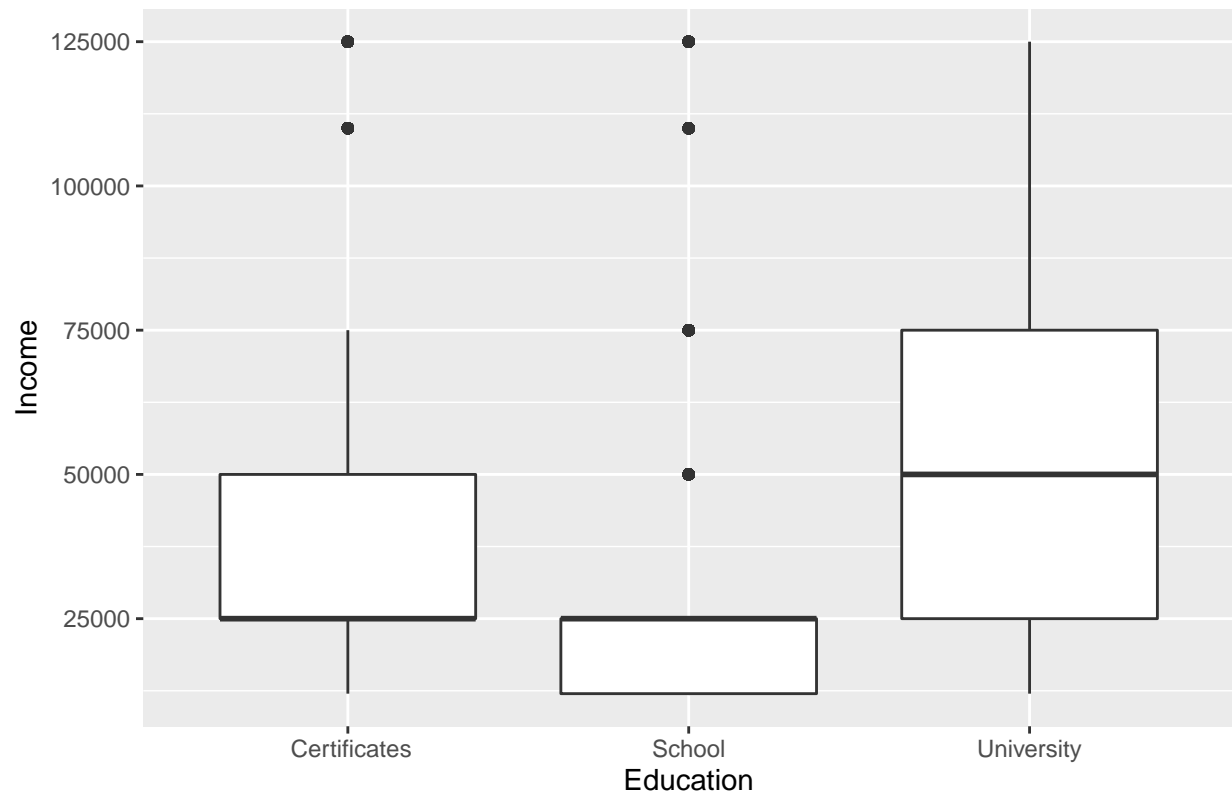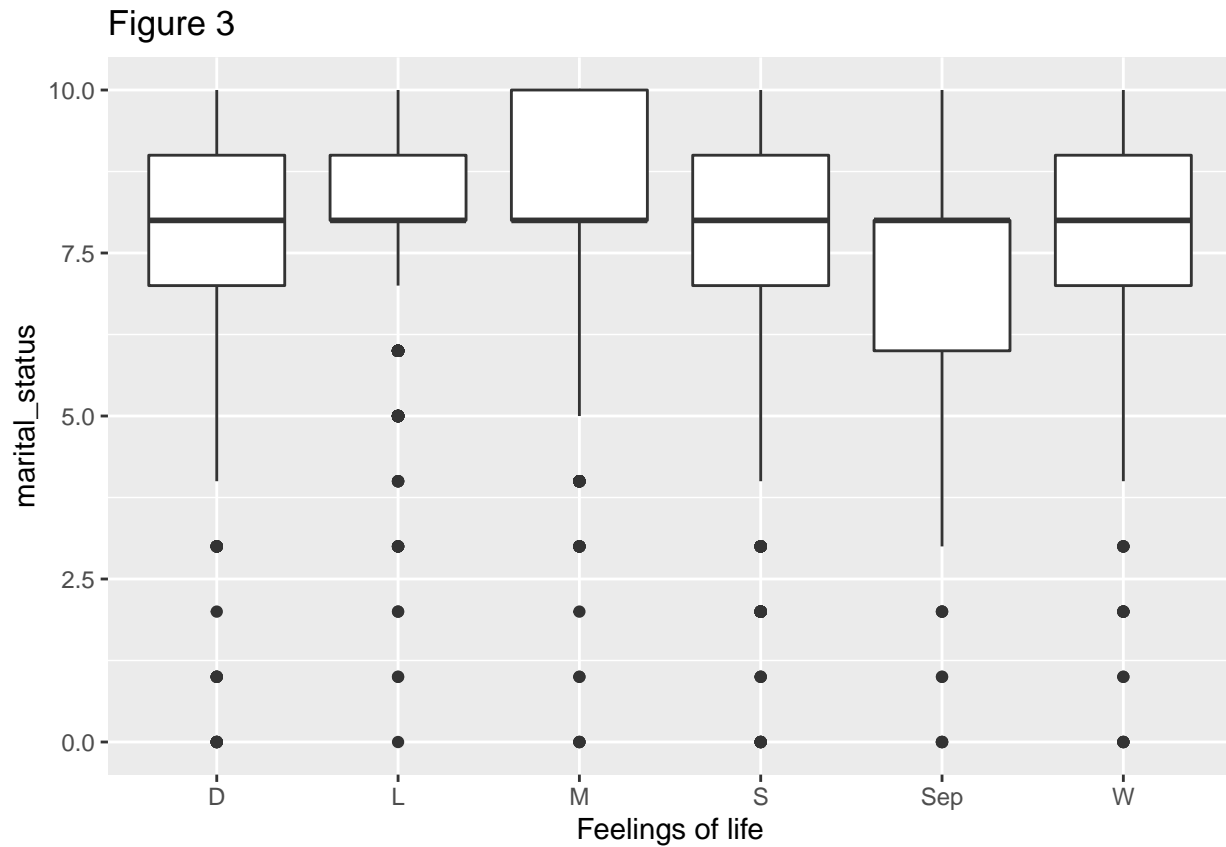
Figure 1

Figure 2

Figure 3

In the Figure 1, we see that life satisfaction start to increase after $50,000$, while in Figure 2 there is a big difference in earnings between University educated versus Certificate holders and High School graduates. Lastly, there is a significant life improvement for Married people verses such as Separated.

Discussion

I analyzed data from the 2017 General Social Survey (GSS) on the Family. We were interested in finding out whether there is a relationship between education and life satisfaction. Despite small improvement between satisfaction and higher education levels, there is a significant difference in incomes for University educated people and high school or less. Based on the data above, income plays an important role in general life satisfaction and higher education can help to meet it. We see a similar case with every income group. As we see in the Figure 1, making more money could lead to a higher happiness rating. Overall, the indicators of Income and Education are good predictors for life satisfaction.

Weaknesses

Like it was mentioned in Data section, data set has lots of categorical variables, which makes it difficult to perform quantitative analysis. Since I omitted lots of NA data points, I am not sure if they have correlation with dependent variable. Some of this data points might have an importance to our model. Also, like in the case of education level and life satisfaction, there is a big difference in scores between "Less than high school diploma or its equivalent" and "Bachelor's degree (e.g. B.A., B.Sc., LL.B.)". This is because our the target population includes all non-institutionalized persons 15 years of age and older. Another weakness is that Simple Linear regression is sensitive to an outliers.

Next Steps

Since the data is for 2017, for the next steps, it is better to compare the current data analysis with post COVID-19 survey. We could explore how people's vision about higher education changed due to COVID-19 pandemic. With Online classes, increased tuition and fewer opportunities in big cities, the value of higher education is expected to decrease. In addition, we could explore the effects of work from home on mental health and general life satisfaction among people in suburbs and cities.

References

1. Wu, Changbao, and Mary E. Thompson. "Basic Concepts in Survey Sampling." Sampling Theory and Practice. Springer, Cham, 2020. 3-15.

2. 2017 GSS Data: General Social Survey, Cycle 31: 2017: Family. (n.d.). Downloaded October 17, 2020, from https://sda-artsci-utoronto-ca.myaccess.library.utoronto.ca/cgi-bin/sda/hsda?harcsda4+gss31

3. Summary about GSS Data: General Social Survey, Cycle 31: 2017: Family. (n.d.). Used October 17, 2020, from https://www23.statcan.gc.ca/imdb/p2SV.pl?Function=getSurvey&Id=335816

4. Data Cleaning help: Alexander, Rohan, and Caetano, Samantha .(2019, Sept 16). "gss_cleaning.R". Used October 16, 2020, from https://www.tellingstorieswithdata.com/01-03-r_essentials.html

5. Higher education among OECD countries: Population with tertiary education. (n.d.). Used October 16, 2020, from https://data.oecd.org/eduatt/population-with-tertiary-education.htm#indicator-chart

6. Education at glance, OECD report: Education at a Glance: OECD Indicators, 2019.(n.d.) Used October 17, 2020, from https://www.oecd.org/education/education-at-a-glance/EAG2019_CN_CAN.pdf

7.RStudio Team (2020). RStudio: Integrated Development for R. RStudio, PBC, Boston, MA http://www.rstudio.com/

8. Knitr package: Xie, Y. (n.d.). Knitr v1.30. Downloaded October 18, 2020, from https://www.rdocumentation.org/packages/knitr/versions/1.30