# Examining the Relationship between Body Mass Index (BMI) and Life Satisfaction*

### Individuals with Higher Life Satisfaction Tend to Have Lower BMI

Chenhang Huang

April 19, 2024

Life satisfaction not only affects an individual's mood but also has an impact on their physical health. This study conducted explanatory and empirical analysis using the 2017-2018 Canadian Community Health Survey (CCHS) data through a Bayesian regression model. The results indicate that higher levels of life satisfaction are associated with lower average BMI values. Therefore, it is advisable to strive for satisfaction in daily life and maintain a positive attitude towards life, which can help reduce the likelihood of weight gain and obesity.

## 1 Introduction

The high cost of living in Canada has led to increased stress for many individuals, particularly among the middle class. Canadians' life satisfaction is declining, which is a significant indicator of a country's overall happiness index. Life satisfaction not only impacts individuals' mental health but also their physical well-being. Obesity is a prevalent issue affecting people's lives, as it can lead to various health problems such as high cholesterol, hypertension, and others. Moreover, obesity may subject individuals to ridicule from peers, adversely affecting their daily lives.

Body Mass Index (BMI) is a measure used to assess body weight relative to height. It is calculated by dividing a person's weight in kilograms by the square of their height in meters (BMI = weight / height^2). BMI is commonly used as a screening tool to classify individuals into categories such as underweight, normal weight, overweight, and obese. These categories are typically defined as follows:

Underweight: BMI less than 18.5

*Code and data are available at: https://github.com/chenhanghuang001/BMI.git

Normal weight: BMI between 18.5 and 24.9

Overweight: BMI between 25 and 29.9

Obese: BMI 30 or greater

BMI provides a simple and quick method for assessing body weight and is widely used in clinical settings, public health research, and population studies to identify individuals or populations at risk of weight-related health issues such as cardiovascular disease, diabetes, and certain cancers.

The estimand in this paper is the effect of the life satisfaction on the individual's body mass index. However, there are many factors contributing to obesity, such as diet, age, gender, marital status, and others. Similarly, life satisfaction is associated with these factors, including diet, age, gender, and marital status. When studying the impact of life satisfaction on BMI, it's essential to control for these influencing factors to minimize bias in estimating the effect of life satisfaction on BMI. Therefore, during the research process, it's necessary to collect information on diet, age, gender, marital status, and other relevant factors simultaneously.

The remainder of this paper is structured as follows. Section 2 will do the data analysis, including the data source, the number of variables, and their definition. Section 3 will introduce the regression model and the corresponding assumptions. Section 4 focus on the analysis of the results of the regression. Discussion on these results are proposed in Section 5.

# 2 Data

Data analysis is performed in R (R Core Team 2023), and I also use some other package to help me plot the figrues and build the tables. such as readr (Wickham, Hester, and Bryan 2023), ggplot2 (Wickham 2016), tidyverse (Wickham et al. 2019),knitr [knitrR], modelsummary (Arel-Bundock 2022). Data comes from the Canadian Community Health Survey (CCHS) (Statistics Canada 2018).

## 2.1 Data source and clean

The CCHS selects a sample of 130,000 respondents every two years: 120,000 respondents covering the population aged 18 and older, and 10,000 respondents covering the population aged 12 to 17. The CCHS manual provides detailed definitions and measurement methods for each variable. In this study, we selected BMI and life satisfaction as the response variable and predictor, respectively. BMI is calculated based on weight and height. Due to the large number of observations in the original data, we retained only approximately 2000 observations during the data cleaning process. After disregarding observations with missing values, the final dataset consists of 1709 observations and 6 variables.

## 2.2 Measurements

The CCHS employs a multi-stage sample allocation strategy to distribute the sample relatively evenly across the provinces and territories. For each age group (18 and older, 12 to 17), initially, 0.75 of the allocation weight is used to distribute the sample among the provinces based on their respective population sizes. Then, using 0.35 of the allocation weight, the sample for each province is further allocated within the province based on its population size.

The CCHS sample is selected using two different frames: the Area frame and the Canada Child Tax Benefit (CCTB) frame. The "Area" frame is used to select residential samples targeting the population aged 18 and older. During data collection, all members of the household are listed, and individuals aged 18 and older are automatically selected using various selection probabilities based on age and household composition. The CCTB frame is utilized for sampling the population aged 12 to 17. When selecting CCHS residential samples, the sampling scheme of the LFS must be taken into consideration. The LFS plan is a complex two-stage stratified design. Initially, the LFS employs a probability proportional to size (PPS) sampling method to select clusters, and then a systematic sampling of households within the clusters is used to select the final sample. During the selection process, the sample size is expanded to account for non-response and units outside the coverage scope (such as vacant dwellings, institutions, children ineligible due to age or death, etc.) (Statistics Canada 2018).

## 2.3 Explantory data analysis

The variable "life satisfaction" consists of 3 levels: "Satisfied", "Neither satisfied nor dissatisfied", and "Dissatisfied". Among these levels, "Satisfied" has the largest proportion, accounting for approximately 91% of individuals who are satisfied with their current lives, while around 5% of individuals are dissatisfied with their lives, as shown in Table 1. Figure 1 illustrates the relationship between life satisfaction and BMI. It shows that individuals satisfied with their lives have a slight right skewness in BMI distribution, with only a small proportion having excessively high BMI values. However, among those dissatisfied with their lives, some individuals have a BMI exceeding 40, significantly higher than other dissatisfied individuals.

Approximately 28% of individuals rent their dwelling, as indicated in Table 1, and no significant differences in BMI were observed between those who rent and those who do not, as shown in Figure 2. In terms of marital status, married individuals have a smaller proportion of individuals with excessive BMI compared to other groups, as depicted in Figure 3. Comparing individual income and BMI, it appears that individuals with lower incomes are more likely to have a higher BMI, as illustrated in Figure 4. Regarding gender differences, females have a higher probability of having both smaller and larger BMIs compared to males, as shown in Figure 5. The sampled data consists of 861 females and 848 males, with a gender ratio close to 1:1, indicating that the data collected through random sampling has the representatives of the Canadian population.

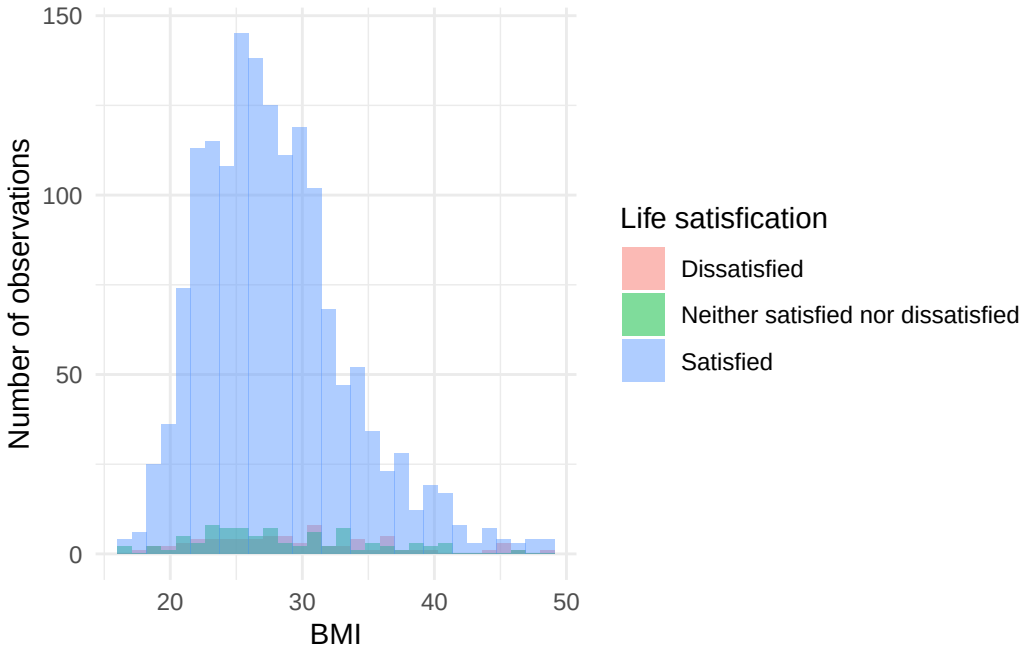


Figure 1: The distribution of BMI grouped by life satisficaiton

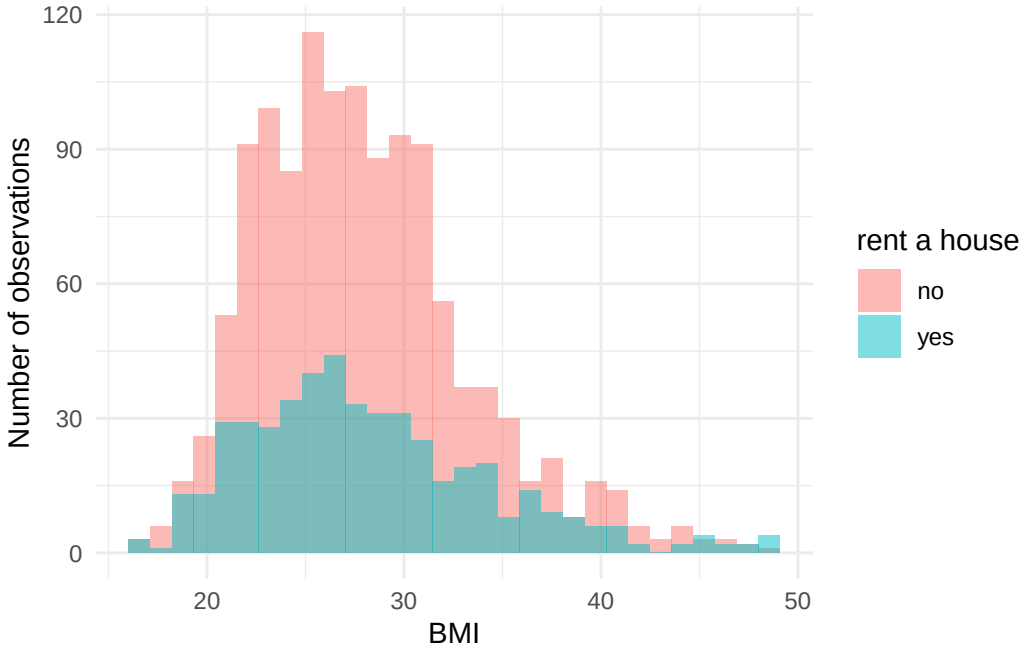


Figure 2: The distribution of BMI grouped by rent a house

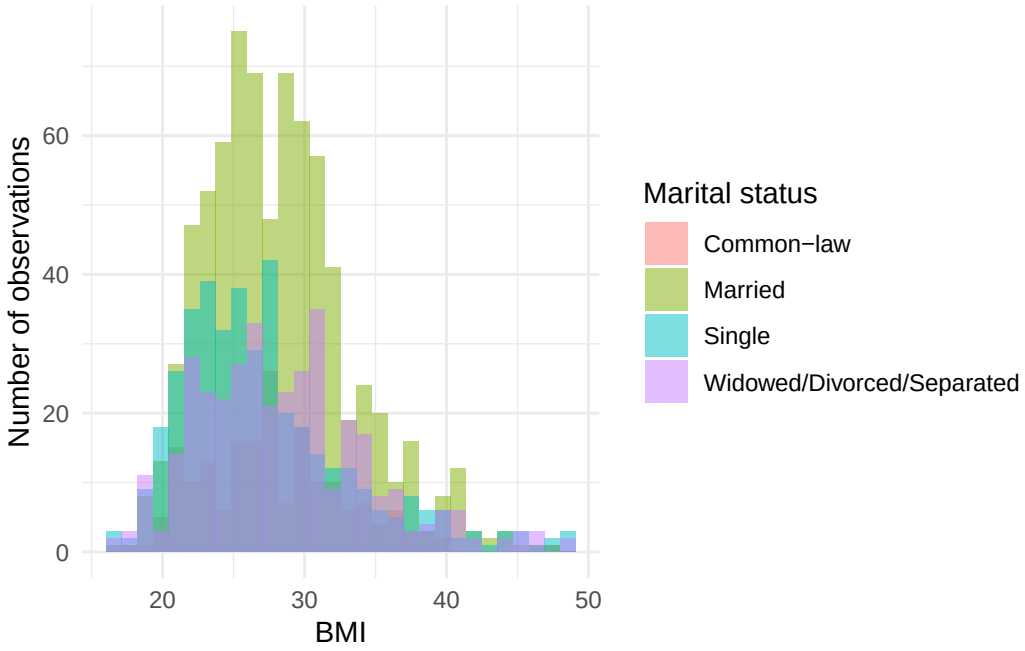


Figure 3: The distribution of BMI grouped by marrital status

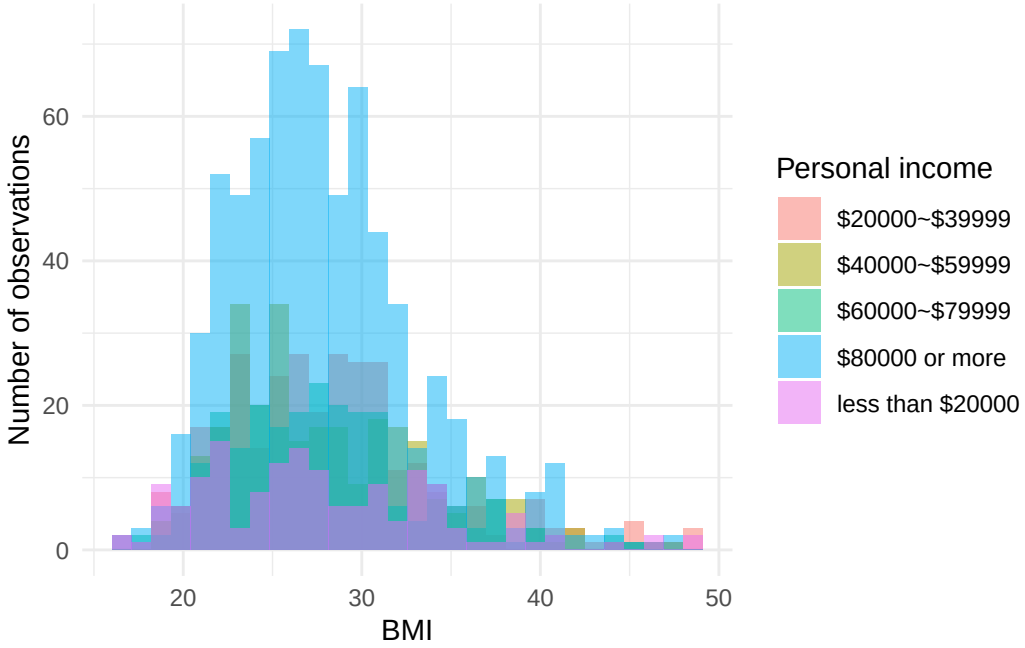


Figure 4: The distribution of BMI grouped by income levels

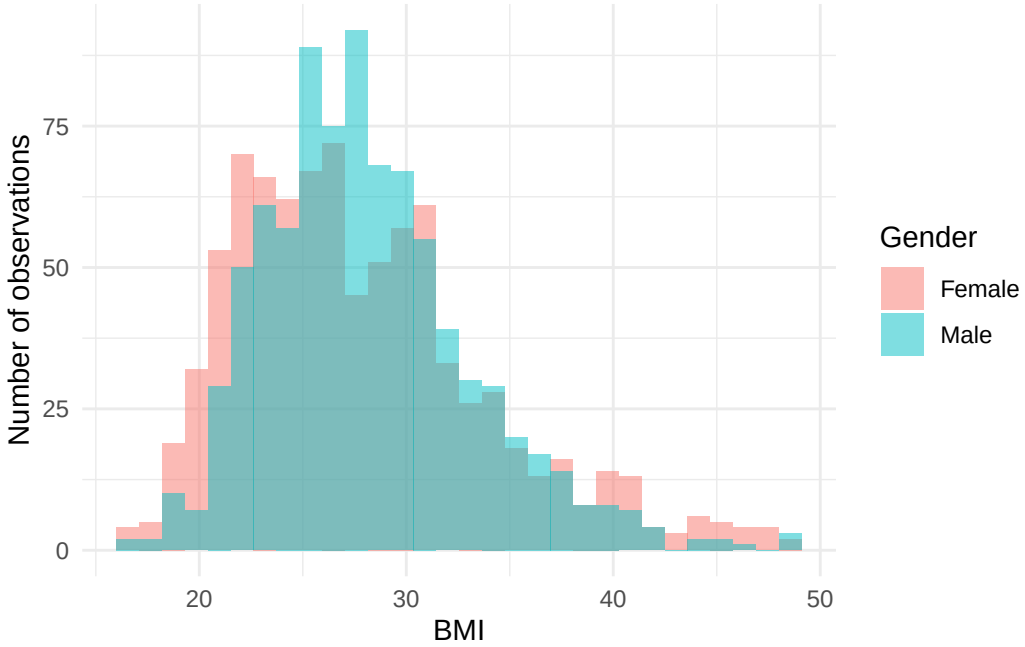


Figure 5: The distribution of BMI grouped by gender

Table 1: Summary Statistic

| | | N | % |
|---|---|---|---|
| lifesatisfication | Dissatisfied | 72 | 4.2 |
| | Neither satisfied nor dissatisfied | 83 | 4.9 |
| | Satisfied | 1554 | 90.9 |
| income | $20000~$39999 | 309 | 18.1 |
| | $40000~$59999 | 283 | 16.6 |
| | $60000~$79999 | 236 | 13.8 |
| | $80000 or more | 727 | 42.5 |
| | less than $20000 | 154 | 9.0 |
| gender | Female | 861 | 50.4 |
| | Male | 848 | 49.6 |
| marry | Common-law | 187 | 10.9 |
| | Married | 751 | 43.9 |
| | Single | 407 | 23.8 |
| | Widowed/Divorced/Separated | 364 | 21.3 |
| rent | no | 1233 | 72.1 |
| | yes | 476 | 27.9 |

# 3 Methods and Models

Here I will introduce the Bayesian anlaysi model used in this paper

## 3.1 Model set-up

From the histogram in Figure 1, it appears that the group of people who are dissatisfied with life are more likely to experience overweight. But we cannot discern the effect of life satisfaction on BMI merely from the graph. Here, I construct a simple linear regression model to further investigate the relationship between life satisfaction and BMI. Taking into account other factors such as income, marital status, living environment, etc., that may affect BMI, I also develop a multiple linear regression model with all explanatory variables included.

### 3.1.1 Simple Linear Regression (SLR)

Define $y_i$ as the BMI for individual i, then the SLR model is:

$$
\begin{aligned}
y_i|\mu_i, \sigma &\sim \text{Normal}(\mu_i, \sigma) && (1)\\
\mu_i &= \beta_0 + \beta_1 lifestaisfication_i + \gamma_i && (2)\\
\beta_0 &\sim \text{Normal}(0, 25) && (3)\\
\beta_1 &\sim \text{Normal}(0, 10) && (4)\\
\sigma &\sim \text{Exponential}(1) && (5)
\end{aligned}
$$

We run the model in R (R Core Team 2023) using the `rstanarm` package of Goodrich et al. (2022). We use the default priors from `rstanarm`.

### 3.1.2 Multiple Linear Regression (MLR)

Include other predictors in the model, the MLR is in the following:

$$y_i|\mu_i, \sigma \sim \text{Normal}(\mu_i, \sigma) \tag{6}$$
$$\mu_i = \beta_0 + \beta_1 lifestaisfication_i + \beta_2 income_i + \beta_3 gender_i + \beta_4 marry_i + \beta_5 rent_i \tag{7}$$
$$\beta_0 \sim \text{Normal}(0, 25) \tag{8}$$
$$\beta_1 \sim \text{Normal}(0, 2.5) \tag{9}$$
$$\beta_2 \sim \text{Normal}(0, 2.5) \tag{10}$$
$$\beta_3 \sim \text{Normal}(0, 2.5) \tag{11}$$
$$\beta_4 \sim \text{Normal}(0, 2.5) \tag{12}$$
$$\beta_4 \sim \text{Normal}(0, 2.5) \tag{13}$$
$$\sigma \sim \text{Exponential}(1) \tag{14}$$

### 3.1.3 Model Justification

We anticipated that individuals dissatisfied with life would have higher BMI. Thus, in the simple linear regression (SLR) model, we only considered the effect of life satisfaction on BMI to compare the differences in BMI among different groups of life satisfaction. Additionally, we wanted to investigate the differences in BMI among different groups of life satisfaction when including control variables. Therefore, multiple linear regression (MLR) can eliminate potential omitted variable bias in the SLR model.

## 4 Results

Table 2 displays the differences in BMI among different life satisfaction groups. The reference group is the dissatisfied level group. The coefficient for the "Neither satisfied nor dissatisfied" group is -1.173, indicating that individuals in this group have a BMI value 1.173 units smaller than those in the dissatisfied level group. Conversely, the "Satisfied" group has a BMI value 1.822 units smaller than the dissatisfied level group. These results demonstrate a relationship between life satisfaction and BMI: as life satisfaction increases, BMI tends to decrease.

Table 3 presents the results of the multiple linear regression (MLR) after considering the influence of other control variables. In the MLR model, holding other predictors constant, individuals in the "Neither satisfied nor dissatisfied" group have a BMI value 1.154 units smaller than those in the dissatisfied level group, while individuals in the "Satisfied" group have a BMI value 1.739 units smaller than the dissatisfied level group. The similarity between the simple linear regression (SLR) and MLR model results indicates that these models are robust in estimating the relationship between life satisfaction and BMI.

Table 3 can be used to analyze the influence of other factors on BMI. The coefficients for income are all negative, indicating that individuals in the income range of 20000~40000 tend

Table 2: The basic Regression results without control

| | BMI |
|---|---|
| (Intercept) | 29.706 |
| | (0.664) |
| lifesatisficationNeither satisfied nor dissatisfied | −1.173 |
| | (0.900) |
| lifesatisficationSatisfied | −1.822 |
| | (0.672) |
| Num.Obs. | 1709 |
| R2 | 0.005 |
| R2 Adj. | 0.001 |
| Log.Lik. | −5366.094 |
| ELPD | −5370.3 |
| ELPD s.e. | 34.4 |
| LOOIC | 10 740.6 |
| LOOIC s.e. | 68.8 |
| WAIC | 10 740.6 |
| RMSE | 5.59 |

to have higher BMI relative to other income levels. Regarding gender, males have a higher average BMI compared to females. In terms of marital status, single individuals have lower BMI than those in other marital statuses, while individuals in less fortunate marital statuses, such as Widowed/Divorced/Separated, have higher average BMI than other groups. Lastly, we observe that individuals who rent tend to have higher average BMI than those who own houses.

Finally, in Figure 6, the credible intervals show that the satisfied life satisfaction group's interval is to the left of 0 and does not include 0. This result indicates that individuals with satisfied life satisfaction have lower BMI values, which is advantageous for reducing the occurrence of obesity. Additionally, for individuals with a marital status of single, their credible interval is also to the left of 0, indicating that the average BMI of single individuals is lower than average.

# 5 Discussion

## 5.1 Key Findings

From the exploratory data analysis (EDA) and the results tables, we observe a negative correlation between BMI and life satisfaction. As life satisfaction increases, the average BMI tends

Table 3: The Regression results with control

| | BMI |
|---|---|
| (Intercept) | 29.634 |
| | (0.898) |
| lifesatisficationNeither satisfied nor dissatisfied | −1.154 |
| | (0.901) |
| lifesatisficationSatisfied | −1.739 |
| | (0.686) |
| income$40000~$59999 | −0.047 |
| | (0.457) |
| income$60000~$79999 | −0.531 |
| | (0.504) |
| income$80000 or more | −0.416 |
| | (0.414) |
| incomeless than $20000 | −0.702 |
| | (0.538) |
| genderMale | 0.324 |
| | (0.270) |
| marryMarried | 0.280 |
| | (0.461) |
| marrySingle | −0.846 |
| | (0.515) |
| marryWidowed/Divorced/Separated | 0.446 |
| | (0.524) |
| rentyes | 0.486 |
| | (0.345) |
| Num.Obs. | 1709 |
| R2 | 0.020 |
| R2 Adj. | 0.000 |
| Log.Lik. | −5358.591 |
| ELPD | −5371.0 |
| ELPD s.e. | 35.1 |
| LOOIC | 10 742.0 |
| LOOIC s.e. | 70.2 |
| WAIC | 10 741.9 |
| RMSE | 5.56 |

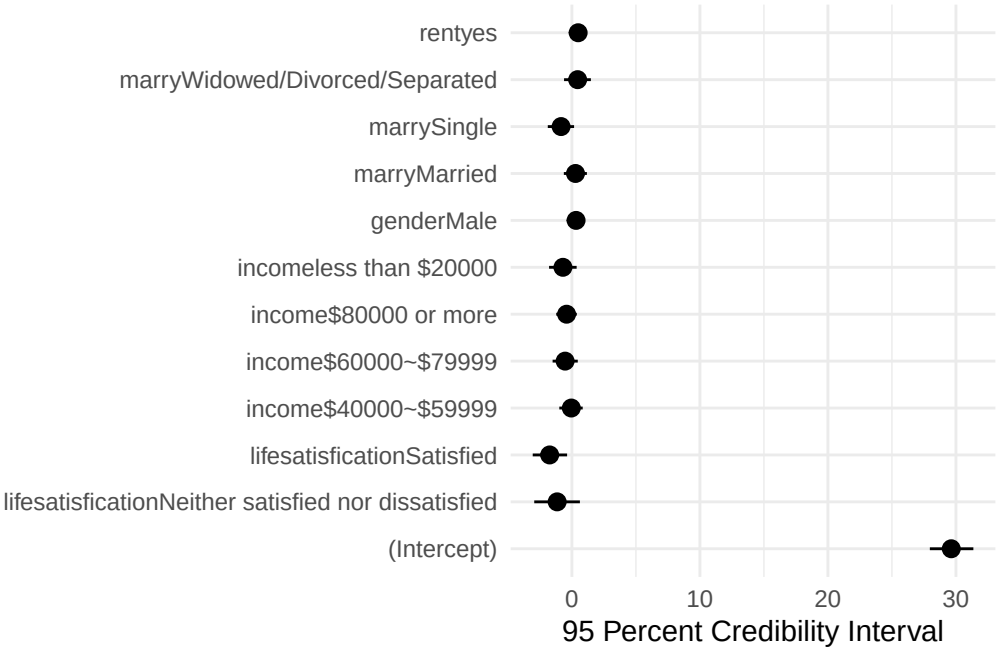


Figure 6: Credible intervals for the regression results with control

to decrease, and this result demonstrates robustness as it remains relatively stable even with the introduction of other variables. There are several possible reasons why individuals with higher life satisfaction tend to have smaller BMIs. Firstly, individuals with higher life satisfaction may prioritize their physical health, thereby avoiding overeating and other behaviors that lead to weight gain and associated health issues. Secondly, people may experience greater satisfaction with life due to their body image, which could contribute to maintaining a lower BMI. Therefore, the relationship between life satisfaction and BMI may be bidirectional, and from this study alone, we cannot establish a causal relationship between life satisfaction and BMI.

### 5.1.1 Marital status will influence the relationship between BMI and life satisfaction

Single individuals often tend to be happier, thus having lower BMIs. Marriage often involves issues such as managing relationships and caring for children, which can sometimes bring worries. On the other hand, single individuals do not have these concerns and can lead simpler and happier lives, resulting in higher life satisfaction. The worst-off marital status group is the Widowed/Divorced/Separated group. These individuals have lost normal marital relationships, which is often a difficult experience regardless of the reasons behind it. Therefore, these individuals may experience a decrease in life satisfaction, lack self-control, and engage in overeating, leading to weight gain.

### 5.1.2 Income between $20000-$40000 will have higher BMI

Individuals with incomes in the $20,000 to $40,000 range tend to have higher BMIs compared to other income levels. This group of individuals may only earn enough to meet basic needs, and in order to save money, they may choose inexpensive but high-calorie foods such as fries and burgers. Consuming these types of foods may increase the risk of weight gain.

## 5.2 Weaknesses and Limitations

This study investigates the impact of life satisfaction on BMI, with the results indicating that dissatisfied individuals are more likely to experience obesity. However, several weaknesses must be acknowledged. Firstly, the measurement of life satisfaction is challenging, as the study only provides three levels and cannot adequately capture the differences in individual life satisfaction. Secondly, in selecting prior distributions for the simple linear regression (SLR) and multiple linear regression (MLR) models, there is a lack of sufficient material to narrow down the range of parameter estimates.

## 5.3 Future Steps

In future research, life satisfaction could be further subdivided into finer categories, such as using a scale of 1-10 to measure life satisfaction, and appropriate questionnaires could be designed for this purpose. Regarding instances where model assumptions are not met, I will attempt simple transformations of the output variable to ensure that the assumptions of the regression model hold. The data used in this study are from the 2017-2018 CCHS dataset, and considering the changes in lifestyle habits post-COVID-19 in 2020, such as the widespread adoption of work from home arrangements, exploring the relationship between life satisfaction and BMI in the latest data will be one of the future research steps.

## References


Arel-Bundock, Vincent. 2022. "modelsummary: Data and Model Summaries in R." *Journal of Statistical Software* 103 (1): 1–23. https://doi.org/10.18637/jss.v103.i01.

Goodrich, Ben, Jonah Gabry, Imad Ali, and Sam Brilleman. 2022. "Rstanarm: Bayesian Applied Regression Modeling via Stan." https://mc-stan.org/rstanarm/.

R Core Team. 2023. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. https://www.R-project.org/.

Statistics Canada. 2018. "Canadian Community Health Survey (CCHS) - Annual Component." %5BURL%5D.

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.

Wickham, Hadley, Jim Hester, and Jennifer Bryan. 2023. *Readr: Read Rectangular Text Data*. https://CRAN.R-project.org/package=readr.