Nutrition Study: How Eating Habits Affect One's Personal Characteristics

Group 11 - Andrew Carras, Joe Wuest, Minh Anh Dang, Spencer Lustila

2022-11-13

Contents

Introduction

Background and Motivations

BMI (Body Mass Index) is a measure of weight adjusted for height, calculated as weight in kilograms divided by the square of height in meters (kg/m^2) [1]. The BMI index are sometimes called the Quetelet index, named after the man who originally proposed it in 1832, Adolphe Quetelet [2]. You are of normal weight if your BMI is between 18.5 and 25, overweight if it is between 25 and 30, a BMI of 30 or more are considered obese [3].

Some studies had suggested that BMI is a rather poor indicator of percent of body fat and does not capture information on the mass of fat in different body sites [4]. However, the CDC still consider BMI an appropriate measure for screening for obesity and its health risks [1]. We are intrigued by this, as such, in this project, using the Nutrition Data on 315 patients, we want to examine the effect of habits on ones' BMI and come to our own conclusions, however limited that conclusion might be. Our first research question is: Is BMI a good indicator of healthy habits? We expect calories to be the most correlated factor to the BMI, in that the more calories one consumed a day, the higher their BMI tends to be.

For our next question, we want to know if and how smoking affect other eating habits. Smoking is highly associated with obesity-related behaviors including low leisure time physical activity and low fruit/vegetable intake [5]. To properly answer this question, we will use linear regressions to examine the relationship between nutritions' consumptions and smoking habits.

For our final research question, we want to answer whether there are big differences in the eating habits between genders. A study had suggested that man and woman have different eating styles [6]. However, are there differences in the consumption of any specific nutrition groups? If yes, what are those, and can we make any predictions for the wider population based on our conclusions? We expect men to consume more calories and fat than women daily.

Data

For this project, we are using a Nutrition study on 315 patients undergoing non-cancerous elective surgery. The main subjective of this study was to investigate the relationship between personal characteristics and dietary factors, and plasma concentrations of retinol and carotenoids. This data set consists of 17 variables, of which 7 are documenting various nutrition consumptions of each patient, 2 are about the plasma concentrations of retinol and beta-carotene in the patient's blood, 2 for vitamin usage, 2 for smoking status, and the remaining reported on the personal characteristics of each patient.

Results

Question 1

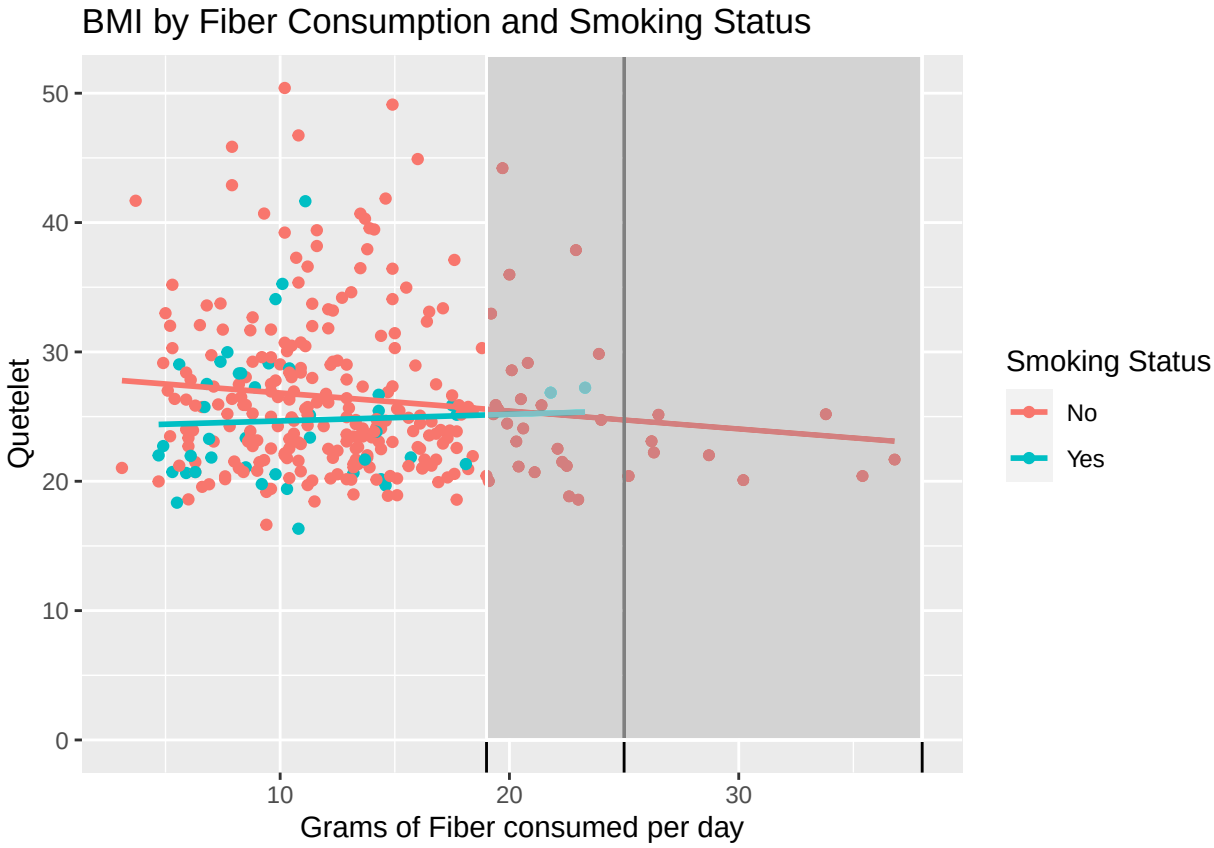
For the first question, we made a full multiple regressions model using all candidate predictors in the data set that concerns variables about habits (Smoking Status, Vitamin Usage, Fat, Fiber, Calories, Alcohol, Cholesterol, Beta-Carotene, and Retinol Consumption). We then use a backward model selection algorithm to find the best fit model.

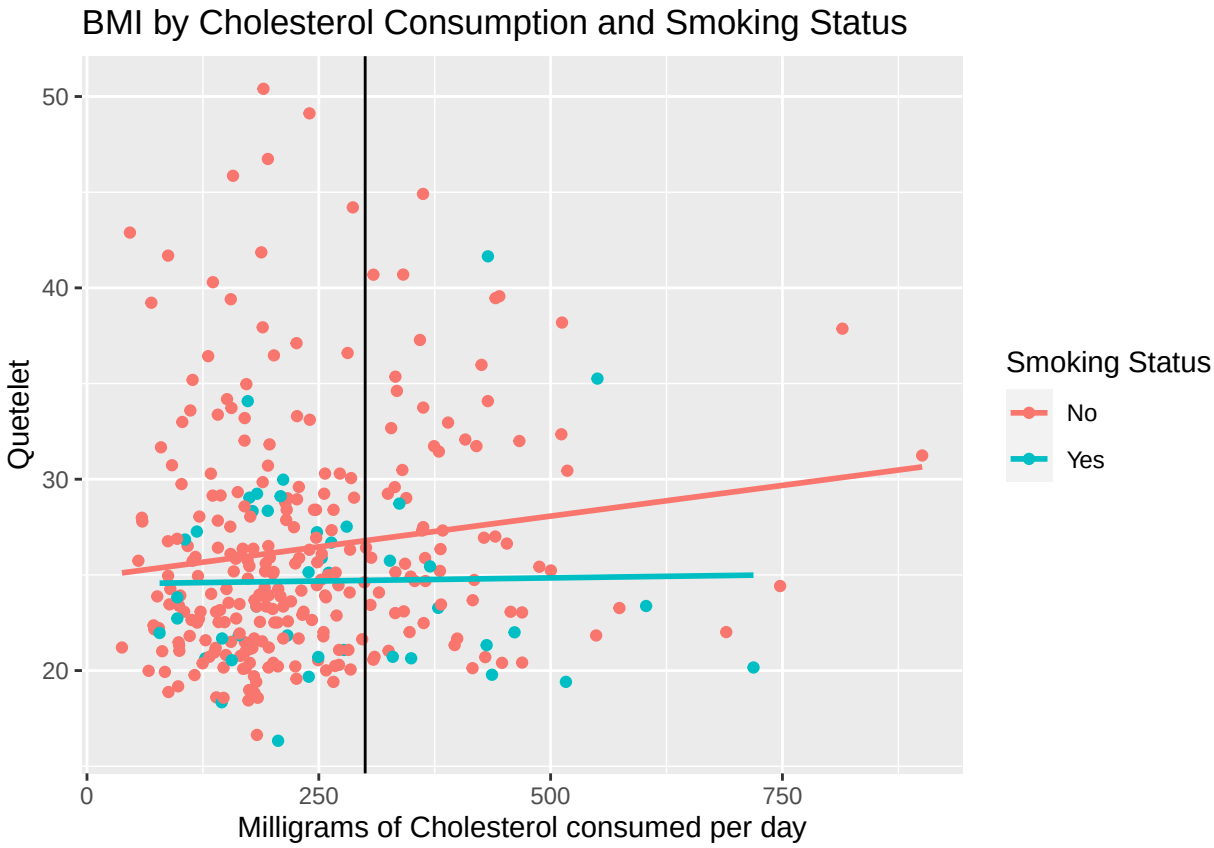
This best fit model has an adjusted R^2 value of nearly 0.036, with the most influential factor being Smoking, averaging a difference of 2.51 between non-smokers and current smokers, 0.88 between non-smokers and ex-smokers, while the least influential being cholesterol consumption, averaging less than 0.01 increase in BMI for each mg of consumption per day.

However, we can see that alcohol consumption has a pretty large p-value of about 0.15, meaning that this result is not replicable and is not statistically significant, as such, we have chosen to remove alcohol consumption from our model. The p-value for ex-smoker is also pretty large (0.23), as such, we will also omit this factor.

We modified our model based on those analyses and the result are as follows: The modified model has an adjusted R^2 value of about 0.031. In this model, on average, for each gram of fiber consumed per day, the BMI decreases by 0.148, and for each mg of cholesterol consumed per day, the BMI increases by 0.006. Using the estimated coefficient, the BMI equation for non-smokers is: BMI = 26.79 - 0.148 * Fiber + 0.006 * Cholesterol

And the equation for current smokers is: BMI = 24.49 - 0.148 * Fiber + 0.006 * Cholesterol


BMI by Fiber Consumption and Smoking Status
Quetelet
50
40
30
20
10
0
10
20
30
Grams of Fiber consumed per day
Smoking Status
No
Yes


BMI by Cholesterol Consumption and Smoking Status
Quetelet
50
40
30
20
0
250
500
750
Milligrams of Cholesterol consumed per day
Smoking Status
No
Yes

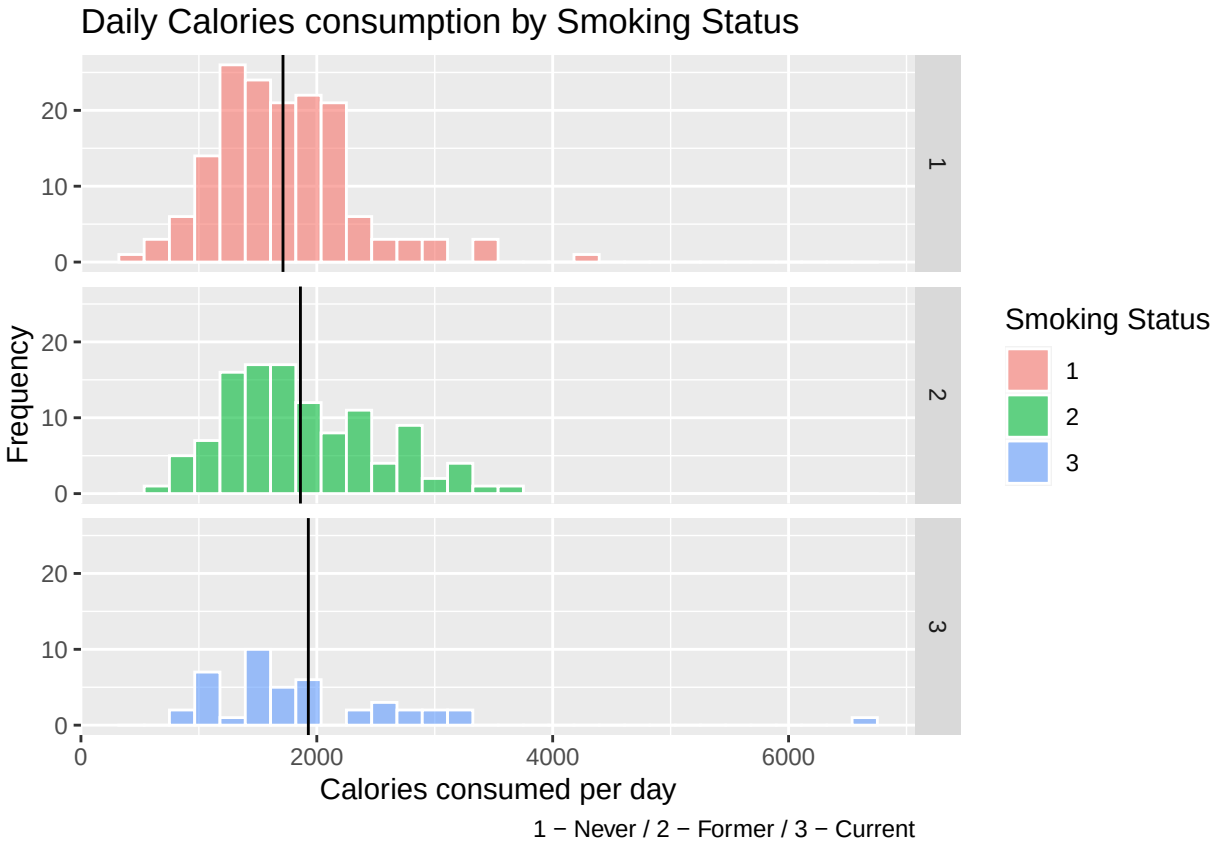
The Institute of Medicine recommended a daily fiber intake of 19 - 38 grams depending on gender and age [7], while the FDA recommended a total dietary fiber intake of about 25 grams a day [8]. Looking at the shaded region in the graph, it is clear that those who consumed enough fiber a day tend to have lower BMI, and are more likely to be a non-smoker. In this data set, no current smoker consumed more than 25 grams of fiber per day.

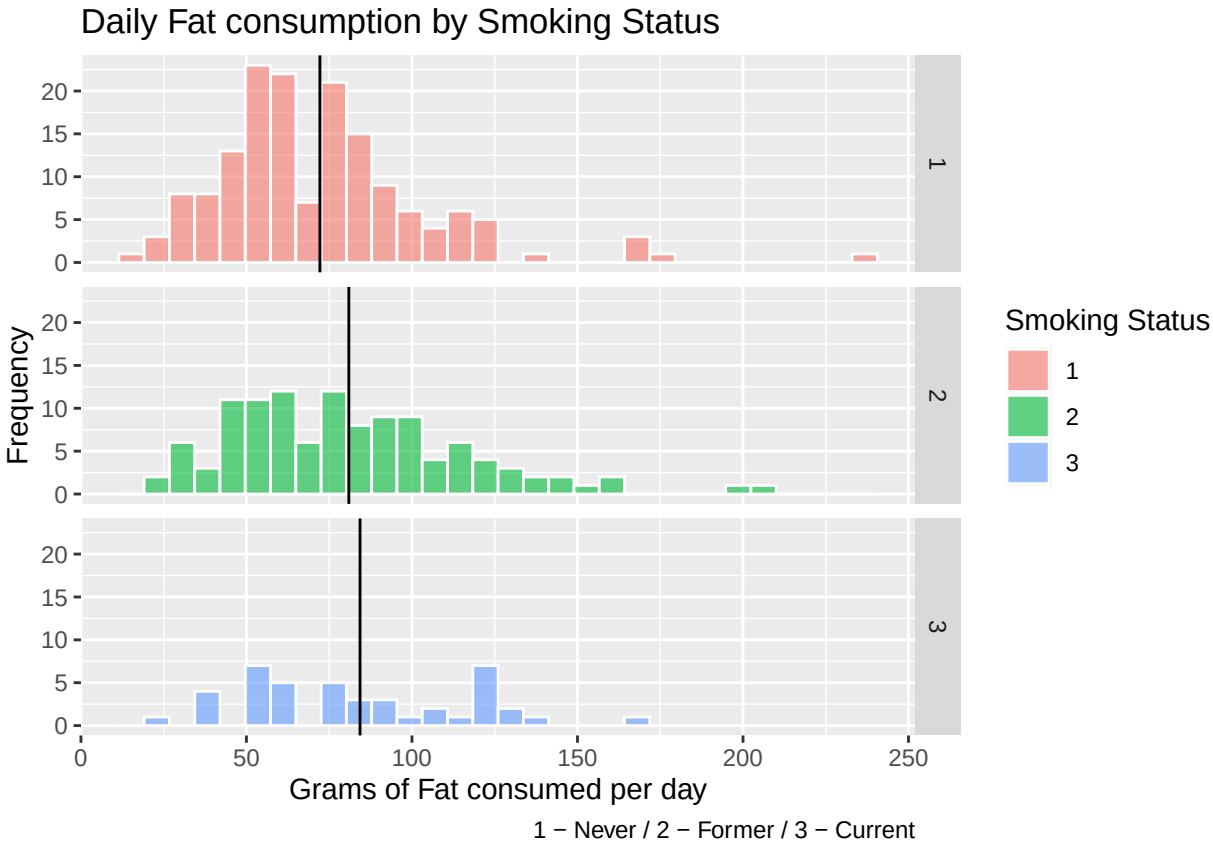
The American Heart Association adopted a recommendation of limiting dietary cholesterol intake to 300 mg/day for healthy individuals in the United States [9]. 32.56% of current smokers consumed more than 300 mg of cholesterol per day, while 25.37% of non-smokers consumed more than the recommended values.

Question 2

In the previous question, it was clear to us that current smokers consumed less fiber than non- and ex- smoker. To investigate this further, we use linear regressions to examine the relationship between fiber consumption, beta-carotene consumption and smoking habits. From these linear regressions, we found that on average, current smokers consumed 2.54 g less fiber and 558.28 mcg less beta-carotene than non-/ex-smokers daily. For reference, non-/ex-smokers consumed on average 13.14 g of fiber and 2261.81 mcg of beta-carotene per day.

Next up, we make linear regressions for all other variables about eating habits (Calories, Fat, Cholesterol, RetinolDiet) and smoking habits. We then get rid of any regression that has a large p-value (> 0.05). The two variables that met our conditions are Calories and Fat, in that there are clear differences in their consumption between current smokers, ex-smokers, and non-smokers. Non-smokers consumed on average 1713.39 calories and 72.19 g of fat per day. Ex-smokers consumed on average 1861.18 calories and 80.93 g of fat per day. And current smokers consumed on average 1928.10 calories and 84.31 g of fat per day.

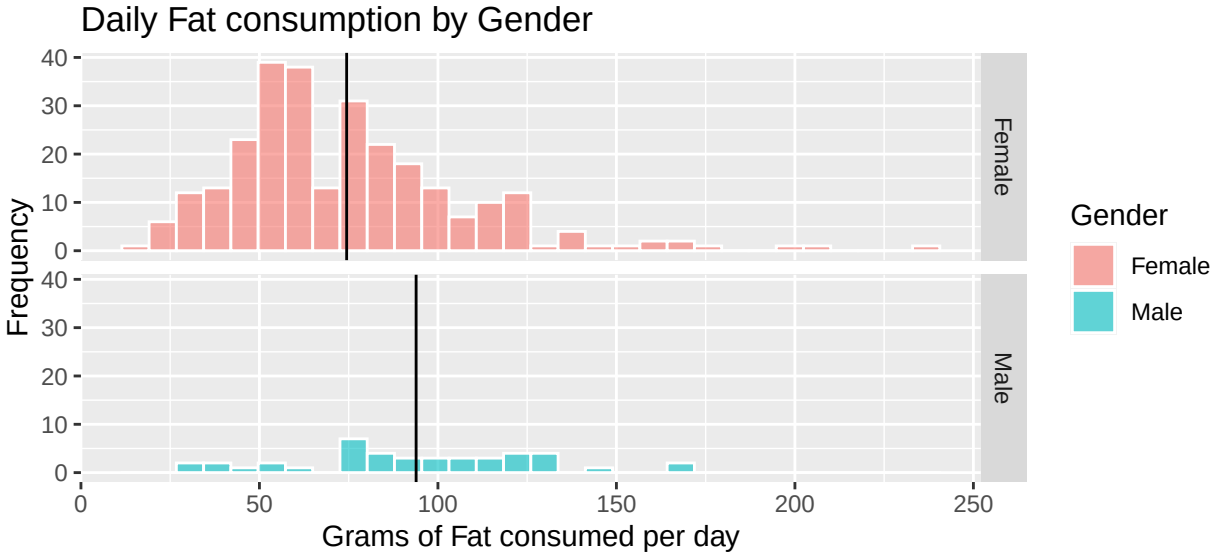


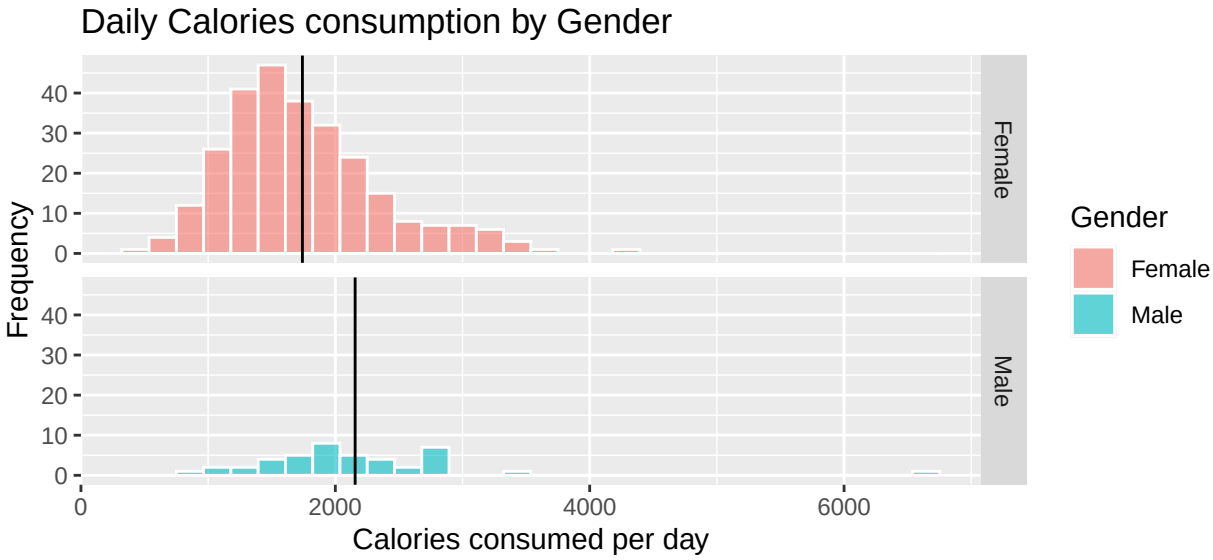


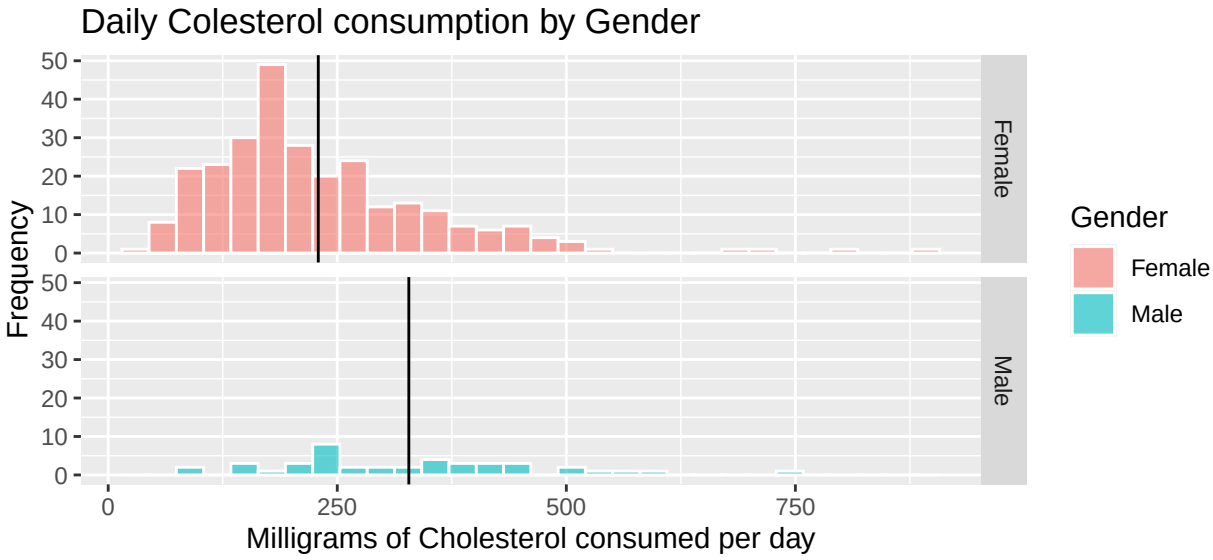
Question 3

For this final question, we use the same method from question 2, in that we make linear regressions again for all habit variables against genders. This time, three variables that met our conditions: Fat, Calories, and Cholesterol.

On average, men consumed 19.45 g more fat, 414.38 more calories, and 98.84 mg more cholesterol per day than women.







Conclusions

Summarization of Results

After reviewing our results, the first observation we made was that calorie consumption doesn't seem to be very influential on BMI. We found that Fiber consumption and smoking status are more significant in determining an individual's BMI, and Cholesterol consumption might have some limited effects on BMI. Based on our results, individuals who consume more fiber and don't smoke tend to have the lowest BMI. However, non-smokers tends to have larger BMI than current smokers. Therefore, it is reasonable to conclude that BMI is not a good indicator of healthy eating habits since the effect of food consumption on it is quite small. We also found that smoking habits do seem to have an effect on eating habits. Individuals who smoke tend to consume more fat and more calories while consuming less fiber and less beta-carotene. These results seem to support the idea that smoking has a negative impact on eating habits. Finally, we observed that

men tend to consume significantly more calories, cholesterol, and fat than women. This is the result that we initially suspected and our evidence clearly supports this.

Critiques, Limitations, and Reliability

For questions 2 and 3, we investigated the effect of categorical predictors on eating habits, we didn't consider whether or not this data set is representative of any particular predictor. For smokers, this data set reported on 43 current smokers out of 315 patients, which accounts for 13.65% of the data set. Compared to the percentage of current smokers in 2001 - 22.8% [10], we can see that this data set might not be representative smoker-wise. Gender-wise, also not. In this data set, there are only 42 men, which is 13.33% of all patients, much less than 50%. We had uses the hypothesis test for both predictors, and both p values are less than 0.05, therefore we had enough evidence to conclude that the data set are not representative smoker-wise and gender-wise. Therefore, the results and trends found for questions 2 and 3 should be used cautiously and should not be applied to a larger population.

Moreover, for all of our models, the calculated correlations were rather small (< 0.1), which means that our models do not display clear linear relationships between the predictors and their responses, but rather trends. For example, those who consumed enough fiber a day tend to have lower BMI, but no clear linear relationship can be established between fiber consumption and BMI.

Suggestions for Future Research

The biggest problem with our results was that the data set is not representative smoker-wise or gender-wise. As such, for future research, we would find a different data set that is more representative. This data set, while big, did not randomly choose its subjects. The group of people being studied are patients undergoing various non-cancerous surgery - quite a descriptive and focused population. Having a more representative data set would certainly give us more ground to apply the results from our data analysis to a larger population.

An interesting step that we can take next is to look at how BMI relates to the concentration of various nutrients in our blood, not just its consumption. Everybody has a different digestion rate, as such, while no clear linear relationship can be established between food consumption and BMI, it might be possible for internal concentrations and BMI. Having such data can also nudge us in the direction of what factor is the most influential over your body's digestion rate of various food.

All in all, food science is a complex field, because human bodies and behaviors are also complex. As such, data analysis in this field should be conducted carefully and its results should be used cautiously.

Resources

1. Body mass index: Considerations for Practitioners. (n.d.). Retrieved November 25, 2022, from https://www.cdc.gov/obesity/downloads/BMIforPactitioners.pdf
2. Eknoyan, G. (2007). Adolphe Quetelet (1796 1874) the average man and indices of obesity. Nephrology Dialysis Transplantation, 23(1), 47–51. https://doi.org/10.1093/ndt/gfm517
3. MediLexicon International. (n.d.). Why BMI is inaccurate and misleading. Medical News Today. Retrieved November 24, 2022, from https://www.medicalnewstoday.com/articles/265215
4. Nuttall, F. Q. (2015). Body mass index. Nutrition Today, 50(3), 117–128. https://doi.org/10.1097/nt.0000000000000092
5. Chiolero, A., Wietlisbach, V., Ruffieux, C., Paccaud, F., & Cornuz, J. (2006). Clustering of risk behaviors with cigarette consumption: A population-based survey. Preventive Medicine, 42(5), 348–353. https://doi.org/10.1016/j.ypmed.2006.01.011
6. Rolls, B. J., Fedoroff, I. C., & Guthrie, J. F. (1991). Gender differences in eating behavior and body weight regulation. Health Psychology, 10(2), 133–142. https://doi.org/10.1037/0278-6133.10.2.133
7. Institute of Medicine, Food and Nutrition Board. Dietary Reference Intakes: Energy, Carbohydrates, Fiber, Fat, Fatty Acids, Cholesterol, Protein and Amino Acids. Washington, DC: National Academies Press; 2005
8. CFR - Code of Federal Regulations Title 21. accessdata.fda.gov. (n.d.). Retrieved November 24, 2022, from https://www.accessdata.fda.gov/scripts/cdrh/cfdocs/cfcfr/cfrsearch.cfm?fr=101.77#
9. Association A.H. Diet and Heart Disease. American Heart Association; Dallas, TX, USA: 1968.
10. Centers for Disease Control and Prevention. (n.d.). Cigarette smoking among adults — United States, 2001. Centers for Disease Control and Prevention. Retrieved November 30, 2022, from https://www.cdc.gov/mmwr/preview/mmwrhtml/mm5240a1.htm
11. Using geom_rect for time series shading in R. Stack Overflow. (1962, May 1). Retrieved December 12, 2022, from https://stackoverflow.com/questions/29648907/using-geom-rect-for-time-series-shading-in-r
12. How to add new line in markdown presentation? Stack Overflow. (1962, December 1). Retrieved December 12, 2022, from https://stackoverflow.com/questions/33191744/how-to-add-new-line-in-markdown-presentation