Predicting Fruit and Vegetable Consumption from Demographic Factors:

Do social environmental factors drive healthy decisions?

Abigail Strick

Dr. Ryan Sperry, Faculty Advisor

Queens College, CUNY

MA in Data Analytics and Applied Social Research

December 1, 2020

Abstract

Using the 2019 Health Information National Trends Survey (HINTS) dataset, this study attempted to examine previous studies' findings on relationships between one's social environment and daily consumption of fruits and vegetables, in addition to finding new relationships. The literature review found many significant relationships between fruit and vegetable consumption, and socioeconomic status, region and urbanicity, demographic, and various lifestyle factors which were tested in the study as well. Linear modeling results show that the most positive relationship exists between higher fruit and vegetable consumption is high self-efficacy, including being a non-smoker. The most negative relationship found is being male. The findings are important for nutrition education and for more efficient healthcare spending.

Introduction

A 2019 study found that a poor diet is costing Americans $50 billion a year, and are attributable to chronic disease (Jardim et. al). The present study aims to focus on the various factors that may affect the amount of fruits and vegetables Americans consume daily. According to The Centers for Disease Control (CDC), this has an impact on overall health. However, the Mediterranean Diet has been found to be very effective for weight loss and disease prevention by being restricted to a mostly plant-based, low processed diet (D'Innocenzo, et. al, 2019). It is also one of the most researched diet, which has been proven over the years to keep its promises, which serves as a benchmark for this study. The Centers for Disease Control (CDC) recommends consuming 1.5-2 cups of fruits and 2-3 cups of vegetables per day for healthy weight management and to maintain a healthy lifestyle and to ignore the fad diets. However, according to a Centers for Disease Control study in 2017, only 1 in 10 American adults meet this requirement (*CDC Press Releases*, 2016).

This study attempts to provide evidence between social environmental factors in relation to daily consumption of fruits and vegetables. Prior research has shown that there are many factors involved when making a food choice, most notably socioeconomic status, region and urbanicity, demographics, and lifestyle. This study will be studying those factors as well to determine the relationships. The data analysis used for this study is from the Health Information

National Trends Survey (HINTS) dataset from 2019. This study will examine different relationships and their impact on the amount of fruits and vegetables individuals consume via t-testing and ANOVA for bivariate analysis, and a heatmap and linear modeling for multivariate analysis. Results show that of the factories in this study, the one that was found to be most closely related to higher daily fruit and vegetable consumption is high efficacy,

This study seeks to confirm prior findings in addition to finding new associations and relationships when there are many factors controlled for. This study hopes to reveal the relationship of how one's lifestyle and environment determines making healthy choices (or lack thereof). It also hopes to direct researchers for future studies, and for health education and health related monetary spending.

Purpose of study

The present study aims to identify various relationships between healthy eating behavior and several environmental influences, to build on the work of prior studies on this topic, and perhaps find new relationships.

The research questions this study will examine are:

1. Are healthy habits, such as consuming fruits and vegetables, predictable based on lifestyle factors?
2. Do people with higher socioeconomic status consume more fruits and vegetables than people with lower socioeconomic status?
3. Does location matter when it comes to consuming fruits and vegetables?
4. Do other healthy habits that do not involve eating have an effect on fruit and vegetable consumption?

Literature Review

Accepted Healthy Diet Standard

A review of the nutrition literature found that the diets that most promote health are the Mediterranean and the DASH (Dietary Approaches to Stop Hypertension) diets. These diets are

extremely similar, in that they predominantly require people to eat plant-based options, exclude mostly animal based-options, and include lifestyle factors such as exercise (Cena et. al, 2020).

Race/Ethnicity

Race and ethnicity play a role in determining fruit and vegetable consumption. A recent study which followed people from 1981-1991, 2011-2012, and 2013-2014, found that vegetable intake has decreased for young adults during the last 23 years. There were increases in all sources of food in terms of quality, except in restaurants (Patetta et al. 2019). Non-Hispanic Whites had the most increase in food quality, while Mexican Americans did not increase the quality of their food. When poverty level is held constant, African American neighborhoods were found to have the least number of supermarkets, while white neighborhoods had the most (Bower et. al, 2013). It is interesting to note that Hispanic neighborhoods in general had more grocery and convenience stores regardless of poverty level. Race and poverty level were not found to be factors in rural areas. Non-white rural people were found to consume more fruits and vegetables than rural white people (Lutfiyya et. al, 2012). When using the Minorities' Diminished Return theory as a guide, a study found that when African Americans attain higher education, they still do not consume more fruits and vegetables than their less educated counterparts due to racial inequality in educational gains (Assari and Lankarani, 2018). When socioeconomic status was controlled for, a study found that Black-White disparities in Healthy Eating Index scores became smaller, while White-Hispanic disparities became greater (Wang, Y., & Chen X., 2011). Poor quality food availability in disadvantaged areas with majority ethnic minorities with low socioeconomic status increases the risk of obesity. While psychosocial factors do not explain ethnic differences in diet, exercise or weight, socioeconomic factors explain them better. The study also found that women and older Americans overall have higher Healthy Eating Index scores. Areas with more African Americans have mainly access to small grocery and convenience stores which have higher prices and less variety than supermarkets (Dunn et. al, 2011). However, this is not true for Hispanics. The authors are led to believe that this may explain the differences between rates of obesity for white people and African Americans. Racial differences have been shown to create health inequality.

Gender

Research has shown that there are significant differences between male and female levels of physical activity and eating healthy. Differences have been found in females exercising more often than males, and females exercise more for the intent of being healthier (Craft et. al, 2014). Rural females outperform rural males when it comes to healthy eating and consuming more than 5 servings of fruits and vegetables per day (Lutfiyya et. al, 2012). Women and older Americans overall have higher Healthy Eating Index scores (Wang, Y., & Chen X., 2011). Eating food away from home is more common among males than females (Seguin et. al, 2016). Higher BMI has been found to be linked with people who regularly consume food away from home and eat less fruits and vegetables. A relationship was found between smartphone health app usage and fruit and vegetable intake (Carroll et. al, 2017). One study found that young adults who are healthy have higher education and higher income and are more likely to use health apps. App users were also more likely to increase their intake in fruits and vegetables, and exercise more. However, males were also found to be less likely to have a health app on their smartphones. "Males, those with lower education, and older US adults were less likely to engage in a number of eHealth activities." When compared to males, females have been shown to be overall more health-conscious.

Socioeconomic status

It has been observed that obesity is more frequent among low income, low educated people. A study tested the Economic Utility theory, where consumers are constrained by price and caloric energy found in foods, and people try to maximize those two for marginal benefit, until marginal cost displays diminishing returns (i.e. the consumer no longer derives benefit from another unit of food because of price) (Drewnowski et. al, 2005). Research has shown that people with constrained incomes and fixed budgets for food will derive the most benefit from high energy density, low cost, more convenient, albeit less satiating foods such as grains, and added sugars and fats, not from fruits and vegetables, which have lower energy density. A study measuring various metrics' effects relating to socio-economic status, and both perceived barriers of food prices (factors respondents consider important when buying a food, obtained via a Likert scale) and perceived benefit of diet quality (a diet and health-related knowledge scale) on diet quality indicators and indices found that socio-economic barriers do exist when it comes to

healthy eating (Beydoun et. al, 2008). While perceived barriers of food price (with socio-economic status held constant) is associated with an increase in sodium and a decrease in fiber intake, perceived benefit of diet quality pointed towards better nutritional habits overall, namely lower energy intake from fat and saturated fat, and an increased intake of fruits and vegetables. Perceived barriers of food price was found to be negatively related to socio-economic status, and behaved as a mediator between socio-economic status and diet quality indicators and indices. Fruit and vegetable intake are generally low across all socioeconomic groups in the United States and by state. A study found that "overall, the prevalence of meeting the fruit intake recommendation was highest among women (15.1%), adults aged 31–50 years (13.8%), and Hispanics (15.7%); the prevalence of meeting the vegetable intake recommendation was highest among women (10.9%), adults aged ≥51 years (10.9%), and persons in the highest income group (11.4%) (Lee-Kwan et. al, 2017)." Another study found a positive relationship between income and food quality (Patetta et al. 2019). An annual income of $35,000 and over for a household and being highly educated were found to be positive factors in consuming fruits and vegetables (Lutfiyya et. al, 2012). A separate study found that there was an association found that high educational attainment was associated with high health literacy (Kuczmarski et. al, 2016). An interesting finding was that Americans have difficulty reading nutrition labels due to poor math skills. A positive relationship was found between neighborhood income and fresh produce price, with race held constant (Dunn et. al, 2011). One research team based their hypotheses on the economic theories of supply, demand, and income elasticity to see trends in rural consumption of fresh produce. If demand goes up for fresh produce, then the price for fresh produce will also increase (theory of supply and demand). Prices may reflect differences in quality of the fresh produce. One study related to health literacy found that it is possible to predict if a person has a healthy diet (DASH-style diet versus Standard Western Diet) based on nutrition literacy (Taylor et. al, 2019). Adding calorie information to menus does not affect a consumer's food choice. However, adding other relevant supplemental nutritional information did (Sinclair el. Al, 2014). Socioeconomic status has been demonstrated to influence healthy lifestyle choices.

Region and Urbanicity

Rural adults were found to be less likely to consume at least five servings of fruits and vegetables than urban adults. If rural adults consume at least 5 servings of fruits and vegetables a

day, they were most likely to be non-white, and they were engaging in some moderate physical activity, and over age 65 (Lutfiyya et. al in 2012). They had a BMI of less than 30, good access to healthcare, and felt they were in good to excellent health. In 37 states, rural residents were less likely to eat 5 servings of fruits and vegetables per day than their non-rural counterparts. In 11 states, they found that rural residents ate at least five servings of fruits and vegetables every day while non-rural adults did not. Rural residents are more likely to live in a food desert and have low access to stores that sell fresh produce. Healthy people with a healthy lifestyle at a lower BMI were more likely to exercise more and eat at least 5 servings of fruits and vegetables daily (2012). Research has shown that West Virginia is the state with the lowest intake of fruits and vegetables on average. "Overall, 12.2% of adults met fruit recommendations ranging from 7.3% in West Virginia to 15.5% in DC, and 9.3% met vegetable recommendations, ranging from 5.8% in West Virginia to 12.0% in Alaska (Lee-Kwan et. al, 2017)." Rural communities were found to have less price competition than more densely populated areas, meaning that high competition will keep the prices relatively lower (Dunn et. al, 2011). One's residential location in the United States may influence their intake of fruits and vegetables.

Other healthy lifestyle influences

One's lifestyle may influence health decisions. It has been demonstrated that people who have high self-efficacy are more likely to eat more fruits and vegetables (Satia, 2009). Non-smokers are more likely to meet the recommended 5 a day consumption of fruits and vegetables than smokers do, due to smokers having lower self-efficacy and motivation than non-smokers (McClure et. al, 2009). It is possible to predict one's level of physical activity based on environmental and demographic factors, namely, young urban black people with low levels of education, who were below the poverty line, were more likely to be less active, and were also more likely to be smokers compared to white people (Hawes et. al, 2019). During a nine-week study, participants who performed high intensity interval workouts in increasing amounts each time, developed a spontaneous unconscious switch in diet to more healthy choices such as fruits and vegetables (Zeppa et. al, 2020). Lifestyle's influence on health has been shown to be a related factor in determining healthy choices.

Data and Methods:

The data used for this study comes from the Health Information National Trends Survey (HINTS). It was run for the National Cancer Institute (NCI) in early 2019. The target population was the normal adult civilian population of the United States greater than 18 years old. The goal was to obtain 3,500 completed questionnaires on paper and an additional 2,046 online. The survey explores health related issues, including diet.

The mailing procedure involved an initial mailing, a reminder postcard, and if necessary, a second and possibly a third mailing. Of the 14,730 paper forms, 3,439 (23.3%) responded. Of the 8,700 web users, 1,222 responded (14%) on paper and 813 (9.3%) responded digitally. The sampling frame was a database of addresses used by Marketing Systems Group (MSG) to provide a random sample of addresses. All non-vacant residential addresses were subject to sampling. In rural areas, without street addresses, other methods were used to get questionnaires to the inhabitants. The data was stratified into high (>=34%) or low (<34%) minority population concentrations, so minorities could be over samples to increase data precision. In the final sample of 23,430, 16,740 (71.4%) were from the high minority areas, and 6,690 (28.6%) were from the low minority areas. All participants selected received a paper copy of the survey, and they were given an option to complete the survey online.

The incentive given for those who received the survey by mail was a $2 bill, and those who used the web received a $10 e-gift card. One adult was chosen in each household to represent the entire household in the survey. The survey was conducted in English or Spanish by paper, but only English online.

Dependent variables:

Fruit and Vegetables are used as the dependent variables for this study. The variables Fruit and Vegetables represent the trends of consumption (0=none, 6=4 cups or more per day). The two variables will then be added together to create a new variable called meetdailyreq. The maximum possible amount of Fruit and Vegetable to create meetdailyreq is 12, which translates into at least 8 cups of fruit plus vegetables consumed per day. Although ordinal, they will be used as continuous variables to see trends.

Independent variables:

Socioeconomic Status:

Educational level was reported as: (1) less than high school; (2) high school graduates; (3) Vocation; (4) some college; (5) College Graduate; it is used as a continuous variable to find trends. Originally IncomeRanges, Household Income was reported as 1 through 9, recording income from $0 (1) through $200,000 or more (9); it is used as a continuous variable for this study to find trends.

Demographics:

Originally SelfGender, Gender was recoded as 1=Male and 2=Female; it is treated as a nominal variable. The minimum age of the participants surveyed were at least 18 years old; it is a continuous interval variable. Race was recoded as: 1 = White, 2 = African American, 3 = Hispanic,4 = Asian, 5 = Other; it is a nominal variable.

Region and Urbanicity:

Originally, RUC2013 RuralUrban represents how urban or rural the area of the participant is located in. This variable was recoded as: 1 = Urban Metro, 2 = Urban Metro, 3 = Urban Metro, 4 = Urban NonMetro, 5 = Urban NonMetro, 6 = Urban NonMetro, 7 = Urban NonMetro, 8 = Rural, 9 = Rural. This recoding turned it into a nominal variable. All of the similar urban statuses were grouped together to form 3 groups used for this study. Region was originally CENSREG. It represents the census regions. and was recoded: 1 = Northeast, 2 = Midwest, 3 = South, 4 = West. It is a nominal variable.

Health Habits:

BMI is a continuous Interval variable for measuring obesity. OwnAbilityTakeCareHealth represents self-efficacy to take care of one's own health; it was recoded as: 1 = Completely Confident, 2 = Very Confident, 3 = Somewhat Confident, 4 = A Little Confident, 5 = Not Confident At All; it is an ordinal variable. SeekHealthInfo represents if the survey participant seeks health information. It was recoded as 1 = Yes, 2 = No; it is a dichotomous variable. In the

survey, NoticeCalorieInfoOnMenu represents eating away from home and whether or not the participant notices calorie information next to the food on the menu or menu board. It was recoded as 1 = Yes, 2 = No; it is a dichotomous variable. Tablet_AchieveGoal represents answers of participants if they use a smartphone app to track a health-related goal. It was recoded as 1 = Yes, 2 = No; it is a dichotomous variable. smokeStatus represents one's current smoking status: 1 = Current, 2 = Former, 3 = Never. HowLongModerateExerciseMinutes is a continuous interval variable of how long survey participants exercise on the days they exercise, in minutes. *TimesModerateExercise* represents how many times per week participants do moderate exercise; it is a continuous interval variable.

Results

The heatmap below (Figure 1) reveals the relationships of the continuous variables to fruits and vegetables, and to each other. The variable Vegetables is shown to have the strongest positive correlations (albeit modest) between the number of times exercised per week, education, and income (shown in dark blue), while education shows a weaker correlation (pale blue). Another interesting correlation shown is a negative relationship between income, education, and BMI (shown in pale red).

Figure 1. Heatmap of all continuous variables and their relationships with fruit and vegetable consumption
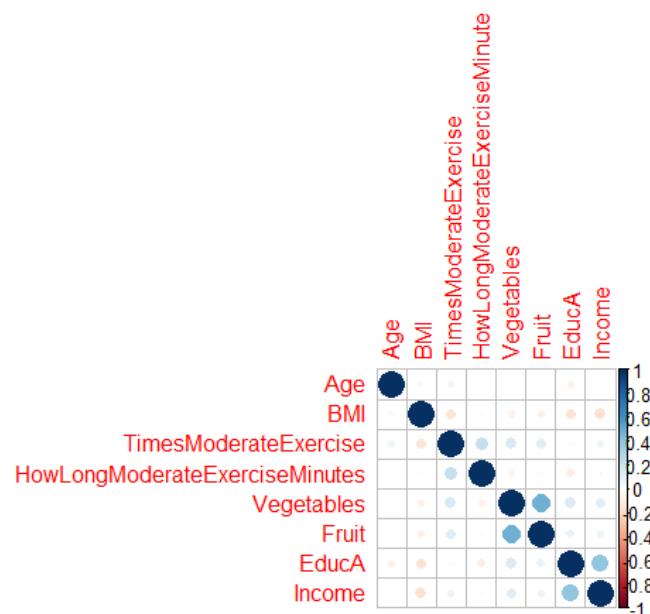
Looking at tables 11 through 14s' t-tests, they reveal several significant relationships between certain behaviors and consuming more fruits and vegetables. The mean for females (5.61) is greater than the mean for males (5.04). The means for using a health app (5.60), noticing calorie information on menus (5.54), and seeking health information (5.43) are greater than for those who do not use a health app (5.17), those who do not notice calorie information on menus (5.15), and those who do not seek health information (4.84). These means reveal that on average, being female and being active in taking care of one's health are associated with the consumption of more fruits and vegetables.

From the ANOVA mean values on tables 15 through 19, it is clear that the strongest associations to consuming more fruits and vegetables are living in the Northeast (5.71), in an urban metropolitan area (5.38), being white (5.45), completely confident in your ability to take care of your own health (5.79), and never having been a smoker (5.5), on average. However, being from an urban or rural area was not found to be statistically significant (p value of .28) while the rest of the ANOVA analyses were statistically significant.

Linear Modeling

The following models were split up by fruit, vegetables, and fruit + vegetable consumption to find relationships separately and together. The models were built up by topic to achieve the optimal r squared. Each subsequent model adds on more control variables. Model 1 represents socioeconomic status, Model 2 adds on region and urbanicity, Model 3 adds on demographics, and Model 4 adds on personal healthy habits. The fruit + vegetable model and the vegetable model both have an r squared of .11. The fruit model has an r squared of only .07.

Table 20. Fruit model

	Model 1	Model 2	Model 3	Model 4
(Intercept)	1.80 ***	1.63 ***	1.76 ***	1.16 **
	(0.12)	(0.29)	(0.31)	(0.39)
EducA	0.14 ***	0.14 ***	0.14 ***	0.10 **
	(0.04)	(0.04)	(0.04)	(0.04)
Income	0.03	0.02	0.03 *	0.00
	(0.01)	(0.01)	(0.01)	(0.01)
RegionNortheast		0.03	0.04	0.05
		(0.10)	(0.09)	(0.09)
RegionSouth		-0.17 *	-0.18 *	-0.16 *
		(0.07)	(0.08)	(0.07)
RegionWest		-0.02	-0.01	0.00
		(0.08)	(0.08)	(0.08)
RuralUrbanUrbanMetro		0.28	0.25	0.17
		(0.27)	(0.27)	(0.26)
RuralUrbanUrbanNonMetro		0.14	0.11	0.04
		(0.28)	(0.28)	(0.28)
GenderMale			-0.25 ***	-0.22 ***
			(0.05)	(0.05)
Age			0.00	0.00
			(0.00)	(0.00)
RaceEthn5Asian			-0.18	-0.18
			(0.14)	(0.14)
RaceEthn5Hispanic			-0.16	-0.19
			(0.10)	(0.10)
RaceEthn5Other			-0.05	-0.02
			(0.16)	(0.16)
RaceEthn5White			-0.13	-0.16
			(0.08)	(0.08)
OwnAbilityTakeCareHealthCompletelyConfident				0.48 **
				(0.18)
OwnAbilityTakeCareHealthNotConfidentAtAll				0.04
				(0.41)
OwnAbilityTakeCareHealthSomewhatConfident				0.20
				(0.18)
OwnAbilityTakeCareHealthVeryConfident				0.37 *
				(0.17)
BMI				-0.01
				(0.00)
TimesModerateExercise				0.10 ***
				(0.02)
HowLongModerateExerciseMinutes				-0.00
				(0.00)
smokeStatusFormer				0.40 ***
				(0.10)
smokeStatusNever				0.45 ***
				(0.09)

	Model 1	Model 2	Model 3	Model 4
Tablet_AchieveGoalYes				0.11 *
				(0.06)
NoticeCalorieInfoOnMenuYes				-0.00
				(0.05)
SeekHealthInfoYes				0.08
				(0.08)
R^2	0.01	0.02	0.03	0.07
Adj. R^2	0.01	0.01	0.02	0.06
Num. obs.	2673	2673	2673	2673

Standard errors in parentheses

*** $p < 0.001$; ** $p < 0.01$; * $p < 0.05$

No significant effect on Income, Age, how long the participant partook in moderate exercise, urbanicity, noticing calories on the menu, and seeking health information. Moderate significance was found at the .1 level for BMI, and being Hispanic or White.

On average, as education level goes up by 1, fruit consumption goes up by .1. Compared to the Midwest, Southerners consume .16 fewer fruit. Compared to females, males consume .22 fewer fruit. Compared to those who are not at all confident in their ability to take care of their health, those who are completely confident and very confident in taking care of their own health consume .48 and .37 more fruit, respectively. As BMI goes up by 1, fruit consumption goes down by .008. As The number of times one does moderate exercise per week goes up by 1, fruit consumption goes up by .1. Compared to those who currently smoke, consumption of fruit increases for former smokers and those who never smoked by .40 and .45 more respectively. Hispanics consume .19 fewer fruits than African Americans. Whites consumer .16 fewer fruits than African Americans. Compared to those who do not use an app to assist with their health goals, those that do use an app consume .11 more fruits.

Table 21. Vegetable model

	Model 1	Model 2	Model 3	Model 4
(Intercept)	1.87 ***	1.71 ***	1.78 ***	1.30 ***
	(0.12)	(0.29)	(0.31)	(0.38)
EducA	0.22 ***	0.22 ***	0.20 ***	0.16 ***
	(0.03)	(0.03)	(0.04)	(0.04)
Income	0.05 ***	0.05 ***	0.06 ***	0.03 *
	(0.01)	(0.01)	(0.01)	(0.01)
RegionNortheast		0.22 *	0.24 *	0.25 **
		(0.09)	(0.09)	(0.09)

RegionSouth		-0.02	-0.01	0.01
		(0.07)	(0.07)	(0.07)
RegionWest		-0.01	0.01	0.04
		(0.08)	(0.08)	(0.08)
RuralUrbanUrbanMetro		0.15	0.14	0.09
		(0.27)	(0.26)	(0.26)
RuralUrbanUrbanNonMetro		0.25	0.22	0.19
		(0.28)	(0.28)	(0.27)
GenderMale			-0.38 ***	-0.35 ***
			(0.05)	(0.05)
Age			0.00	0.00
			(0.00)	(0.00)
RaceEthn5Asian			-0.19	-0.15
			(0.14)	(0.13)
RaceEthn5Hispanic			-0.16	-0.19
			(0.10)	(0.10)
RaceEthn5Other			0.11	0.13
			(0.16)	(0.16)
RaceEthn5White			0.02	-0.05
			(0.08)	(0.08)
OwnAbilityTakeCareHealthCompletelyConfident				0.35 *
				(0.17)
OwnAbilityTakeCareHealthNotConfidentAtAll				-0.01
				(0.40)
OwnAbilityTakeCareHealthSomewhatConfident				-0.03
				(0.17)
OwnAbilityTakeCareHealthVeryConfident				0.18
				(0.17)
BMI				-0.00
				(0.00)
TimesModerateExercise				0.13 ***
				(0.01)
HowLongModerateExerciseMinutes				-0.00 ***
				(0.00)
smokeStatusFormer				0.38 ***
				(0.10)
smokeStatusNever				0.21 *
				(0.09)
Tablet_AchieveGoalYes				0.04
				(0.05)
NoticeCalorieInfoOnMenuYes				0.12 *
				(0.05)
SeekHealthInfoYes				0.10
				(0.08)

R^2	0.03	0.03	0.06	0.11
Adj. R^2	0.03	0.03	0.05	0.10
Num. obs.	2673	2673	2673	2673

Standard errors in parentheses

===

*** p < 0.001; ** p < 0.01; * p < 0.05

No significant effects were found on BMI, age, urbanicity, using an app to track health, seeking health information. There was a moderate effect found for Hispanics at the .1 level. As education goes up by 1, consumption of vegetables goes up by .16. On average, as Income goes up by 1, consumption of vegetables goes up by .03. Compared to the Midwest, consumption of vegetables goes up by .25 in the Northeast. Compared to females, male consumption is lower by .35. Compared to those not being confident at all about taking care of their own health, people who are completely confident in being able to take care of their own health eat .35 more vegetables. As the number of times someone does moderate exercise per week goes up by 1, vegetables consumption goes up by .13. As the number of minutes of moderate exercise goes up by 1, people are less likely to eat vegetables by .002. Compared to current smokers, consumption of vegetables increases for former smokers and those who never smoked by .38 and .21 more respectively. Hispanics are .19 less likely to eat vegetables than African Americans. Those who notice calorie information on menus eat .12 more vegetables than those who don't.

Table 22. Combined Fruits and Vegetables Model

	Model 1	Model 2	Model 3	Model 4
(Intercept)	3.67***	3.34***	3.54***	2.46***
	(0.21)	(0.50)	(0.53)	(0.65)
EducA	0.36***	0.36***	0.34***	0.26***
	(0.06)	(0.06)	(0.06)	(0.06)
Income	0.07**	0.07**	0.09***	0.03
	(0.02)	(0.02)	(0.02)	(0.02)
RegionNortheast			0.25	0.28	0.29
		0.25	0.28	0.29
		(0.16)	(0.16)	(0.16)
RegionSouth		-0.19	-0.19	-0.16
		(0.13)	(0.13)	(0.12)
RegionWest		-0.03	-0.00	0.04
		(0.14)	(0.14)	(0.14)
RuralUrbanUrbanMetro		0.43	0.39	0.26
		(0.46)	(0.46)	(0.44)

RuralUrbanUrbanNonMetro		0.39	0.33	0.22
		(0.48)	(0.48)	(0.46)
GenderMale			-0.63***	-0.58***
			(0.09)	(0.09)
Age			0.00	0.00
			(0.00)	(0.00)
RaceEthn5Asian			-0.37	-0.33
			(0.23)	(0.23)
RaceEthn5Hispanic			-0.32	-0.38*
			(0.17)	(0.17)
RaceEthn5Other			0.06	0.10
			(0.28)	(0.27)
RaceEthn5White			-0.11	-0.22
			(0.14)	(0.14)
OwnAbilityTakeCareHealthCompletelyConfident				0.82**
				(0.29)
OwnAbilityTakeCareHealthNotConfidentAtAll				0.04
				(0.69)
OwnAbilityTakeCareHealthSomewhatConfident				0.17
				(0.30)
OwnAbilityTakeCareHealthVeryConfident				0.55
				(0.29)
BMI				-0.01
				(0.01)
TimesModerateExercise				0.23***
				(0.03)
HowLongModerateExerciseMinutes				-0.00**
				0.00
smokeStatusFormer				0.77***
				(0.17)
smokeStatusNever				0.66***
				(0.16)
Tablet_AchieveGoalYes				0.16
				(0.09)
NoticeCalorieInfoOnMenuYes				0.12
				(0.09)
SeekHealthInfoYes				0.19
				(0.14)
R^2	0.03	0.03	0.05	0.11
Adj. R^2	0.03	0.03	0.05	0.10

| Num. obs. | | 2673 | 2673 | 2673 | 2673 |

Standard errors in parentheses

==

* p < .05   **p < .01   ***p<.001

As education level increases by 1, consumption of fruits and vegetables increases by .26. Compared to the Midwest, the Northeast region consumes .29 more fruits and vegetables. Compared to females, males consume .58 fewer fruits and vegetables. Compared to those who are a little confident in their own ability to take care of their health, those who are completely confident consume .82 more fruits and vegetables. As the number of times one does moderate exercise per week goes up by 1, fruit and vegetable consumption goes up by .233. As moderate exercise activity goes up by 1, fruit and vegetable consumption go down by .002. Compared to those who currently smoke, former smokers and those who never smoked consume .77 and .66 more fruits and vegetables, respectively. Compared to African Americans, Hispanics consume .38 fewer fruits and vegetables. Compared to those who don't use an app to track their health, those who do consume .16 more fruits and vegetables.

Discussion and Conclusions

Given the variables tested, it was interesting to find that they have more of an effect on the model for vegetable consumption than on the model for fruit consumption. An interesting conclusion one may draw from this study is that given income restrictions, people will be more likely to consume fruit than vegetables. Education was found to be a highly statistically significant factor and is positively correlated with eating more fruits and vegetables as education levels increase. Since prior research has shown a significant relationship between higher socioeconomic status and increased fruit and vegetable consumption, it was surprising to find Model 4 in the combined model (table 22) was not statistically significant (Beydoun et. al, 2008; Drewnowski et. al, 2005; Dunn et. al, 2011; Wang, Y., & Chen X., 2011).

Urban versus rural living was not found to have a significant effect on consumption of fruits and vegetables on the linear modeling or t-testing. This was not such of a surprise but makes sense when viewed in the statistical sense, given both areas have constraints per the literature review (Lutfiyya et. al, 2012). However, a more accurate testing would have included an interaction between income and urbanicity. The research indicated that being rural and poor would be significant, however this was not something that was tested in this study. The variables BMI,

Seeking Health Information, and Age also did not come up as statistically significant in any of the linear models, which was a disappointment. The literature review pointed out that Seeking Health Information and Age could potentially be statistically significant (Lutfiyya et. al, 2012; Zeppa, et. al, 2020). However, Seeking Health Information was found to be statistically significant according to the t-test.

This study found that using a health app was found to be statistically significant and positively correlated for the fruit linear model but not for the vegetable model, which is only semi consistent with Carroll et. al's findings (2017). Noticing calorie information on the menu was found to be statistically significant and positively correlated for the vegetable model but not the fruit model, which can add on to some of the research by Seguin et. al (2016).

For the variable region, the South was found to be statistically significant and negatively correlated with fruit consumption and the Northeast was found to be statistically significant and is positively correlated for vegetable consumption. Even though rural versus urban living was not found to be significant, living in the Northeast region of the United States was statistically significant, as expected per Lee-Kwan et. al's study (2017). The Northeast is the most urbanized and wealthiest region in the nation per Bloomberg.com (Badger, 2013).

Unsurprisingly, there is a significant impact that exercise can have on consumption of fruits and vegetables. However, finding a negative correlation between how many minutes people exercise during a week and consumption of fruits and vegetables was unexpected, since prior research found only positive correlations between fruit and vegetable consumption with exercise (Lutfiyya et. al, 2012; Zeppa et. al, 2020).

In the combined model, being completely confident in one's own ability to care of one's own health proved to have the strongest positive correlation (Satia, 2009). Giving up smoking and never having smoked had slightly lower correlations (McClure et. al, 2009).

It was completely expected that females consume more fruits and vegetables than males do. Being male gave the most significant negative correlation in the linear modeling (Carroll et. al, 2017; Lutfiyya et. al, 2012; Seguin et. al, 2016).

Bivariate analysis using t-testing revealed that the additional dichotomous variables are all statistically significant when tested independently from the other variables in the fruit and

vegetables model. Using a health app, noticing calorie information on the menu, and seeking health information proved to all be statistically significant per their respective t-tests. ANOVA testing found the variables with many factors to be statistically significant as well (except for urban versus rural). For both ANOVA and t-testing, it is important to point out that all of the mean values listed here are below the median of 6. The ideal mean would be to score a 7, the equivalent of 1.5-2 cups for fruits (scored as a 3 per the HINTS codebook) and 2-3 cups for vegetables (coded as a 4 per the HINTS codebook). Even though these groups may be statistically significant, the CDC calls for even greater consumption of fruits and vegetables than what people are already eating.

It is interesting to also point out that African Americans consumed more fruits and vegetables than all races/ethnicities except for those coded as "Other" in the combined model (table 22), which confirms Lutfiyya et. al's study (2012). However, ANOVA testing (table 17) found that Whites consume the most fruits and vegetables out of any other race/ethnicity, which is consistent with other findings (Assari and Lankarani, 2018; Bower et. al, 2018; Dunn et. al, 2011; Wang, Y., & Chen X., 2011; Zeppa et. al, 2020). This contradiction may exist since race/ethnicity was tested alone with the dependent variable for the ANOVA test, while other variables were being controlled for in the linear model. This would need to be investigated more in future studies.

Overall, the findings of this study indicate that the majority of people do not meet the CDC's daily recommended amounts for eating fruits and vegetables, which is in accordance with Lee-Kwan et. al's study (2017). However, this may be due to several lifestyle factors, as demonstrated in this study. One of the most interesting findings was confirming that using a health-related app regularly is beneficial to meeting health goals. In addition, various demographic factors may give certain groups of people an advantage to reach the daily amount of fruits and vegetables. Emphasizing and encouraging people of certain demographics to meet the daily recommended amount can be beneficial to the American health care system. Providing greater access to fresh produce to those in need would significantly help increase intake as well. This would not only help on the individual level but also the collective.

Limitations

This study has various limitations. This study does not take into account any intersectionality. The regional level data collected could have been more precise, meaning that more than four regions could have been used. There were several variables to choose from for

education and income in this dataset. The ones chosen for this study had the most factors which yielded the best results in terms of statistical significance and r squared. The education variable, although ordinal, was used to discern a trend. limitation was that income had to be used as an ordinal variable since the dataset did not have it coded as a continuous variable. A limitation with the fruit and vegetable variables for this study was that they are ordinal variables and not continuous variables, since in the dataset they are in a range rather than them being divided up by one cup at a time. Because of this, the analyses are less precise than they potentially can be. There are some unanswered questions about differences among ethnicities such as delving deeper into cultural and traditional eating habits of different groups of people. Further research can help elaborate more on the discrepancies in diet among different ethnic groups. This current study may be used as a means to assist in future research to help study how the consumption of fruit and vegetables may help mitigate the effects of chronic diseases in ethnic minorities and men among all groups.

References

Assari, S., & Lankarani, M. (2018). Educational Attainment Promotes Fruit and Vegetable Intake for Whites but Not Blacks. *J*, *1*(1), 29–41. MDPI AG. Retrieved from http://dx.doi.org/10.3390/j1010005

Badger, E. (2013, December 13). *America's Wealth Is Staggeringly Concentrated in the Northeast Corridor*. Bloomberg.Com. https://www.bloomberg.com/tosv2.html?vid=&uuid=580d0350-127a-11eb-b8ca-6109ada4a67f&url=L25ld3MvYXJ0aWNsZXMvMjAxMy0xMi0xMy9hbWVyaWNhLXMtd2VhbHRoLWlzLXN0YWdnZXJpbmdseS1jb25jZW50cmF0ZWQtaW4tdGhlLW5vcnRoZWFzdC1jb3JyaWRvcg==

Beydoun MA, Wang Y. How do socio-economic status, perceived economic barriers and nutritional benefits affect quality of dietary intake among US adults? European Journal of Clinical Nutrition. 2008 Mar;62(3):303-313. DOI: 10.1038/sj.ejcn.1602700.

Bower KM, Thorpe RJ Jr, Rohde C, Gaskin DJ. The intersection of neighborhood racial segregation, poverty, and urbanicity and its impact on food store availability in the United States. Preventive Medicine. 2014 Jan;58:33-39. DOI: 10.1016/j.ypmed.2013.10.010.

Carroll, J. K., Moorhead, A., Bond, R., Leblanc, W. G., Petrella, R. J., & Fiscella, K. (2017). Who Uses Mobile Phone Health Apps and Does Use Matter? A Secondary Data Analytics Approach. *Journal of Medical Internet Research, 19*(4). doi:10.2196/jmir.5604

CDC Press Releases. (2016, January 1). CDC. https://www.cdc.gov/media/releases/2017/p1116-fruit-vegetable-consumption.html

Cena, H., & Calder, P. C. (2020). Defining a Healthy Diet: Evidence for the Role of Contemporary Dietary Patterns in Health and Disease. *Nutrients*, *12*(2), 334. MDPI AG. Retrieved from http://dx.doi.org/10.3390/nu12020334

Craft, B. B., Carroll, H. A., & Lustyk, M. K. (2014). Gender Differences in Exercise Habits and Quality of Life Reports: Assessing the Moderating Effects of Reasons for Exercise. International journal of liberal arts and social science, 2(5), 65–76.

Adam Drewnowski, Nicole Darmon, Food Choices and Diet Costs: an Economic Analysis, *The Journal of Nutrition*, Volume 135, Issue 4, April 2005, Pages 900–904, https://doi.org/10.1093/jn/135.4.900

D'Innocenzo, S., Biagi, C., & Lanari, M. (2019). Obesity and the Mediterranean Diet: A Review of Evidence of the Role and Sustainability of the Mediterranean Diet. *Nutrients*, *11*(6), 1306. https://doi.org/10.3390/nu11061306

Dunn, R.A., Sharkey, J.R., Lotade-Manje, J. *et al.* Socio-economic status, racial composition and the affordability of fresh fruits and vegetables in neighborhoods of a large rural region in Texas. *Nutr J* **10,** 6 (2011). https://doi.org/10.1186/1475-2891-10-6

Hawes, A. M., Smith, G. S., McGinty, E., Bell, C., Bower, K., LaVeist, T. A., Gaskin, D. J., & Thorpe, R. J., Jr (2019). Disentangling Race, Poverty, and Place in Disparities in Physical Activity. *International journal of environmental research and public health*, *16*(7), 1193. https://doi.org/10.3390/ijerph16071193

Jardim TV, Mozaffarian D, Abrahams-Gessel S, Sy S, Lee Y, Liu J, et al. (2019) Cardiometabolic disease costs associated with suboptimal diet in the United States: A cost analysis based on a microsimulation model. PLoS Med 16(12): e1002981. https://doi.org/10.1371/journal.pmed.1002981

Kuczmarski, M. F., Adams, E. L., Cotugna, N., Pohlig, R. T., Beydoun, M. A., Zonderman, A. B., & Evans, M. K. (2016). Health Literacy and Education Predict Nutrient Quality of Diet of Socioeconomically Diverse, Urban Adults. *Journal of epidemiology and preventive medicine*, *2*(1), 13000115. https://doi.org/10.19104/jepm.2016.115

Lee-Kwan SH, Moore LV, Blanck HM, Harris DM, Galuska D. Disparities in State-Specific Adult Fruit and Vegetable Consumption — United States, 2015. MMWR Morb Mortal Wkly Rep 2017;66:1241–1247. DOI: http://dx.doi.org/10.15585/mmwr.mm6645a1external icon.

Lutfiyya, M.N., Chang, L.F. & Lipsky, M.S. A cross-sectional study of US rural adults' consumption of fruits and vegetables: do they consume at least five servings daily?. *BMC Public Health* **12,** 280 (2012). https://doi.org/10.1186/1471-2458-12-280

McClure, J. B., Divine, G., Alexander, G., Tolsma, D., Rolnick, S. J., Stopponi, M., Richards, J., & Johnson, C. C. (2009). A comparison of smokers' and nonsmokers' fruit and vegetable intake and relevant psychosocial factors. *Behavioral medicine (Washington, D.C.)*, *35*(1), 14–22. https://doi.org/10.3200/BMED.35.1.14-22

Patetta, M.A., Pedraza, L.S. & Popkin, B.M. Improvements in the nutritional quality of US young adults based on food sources and socioeconomic status between 1989–1991 and 2011–2014. *Nutr J* **18,** 32 (2019). https://doi.org/10.1186/s12937-019-0460-4

Persoskie A, Hennessy E, Nelson WL. US Consumers' Understanding of Nutrition Labels in 2013: The Importance of Health Literacy. Prev Chronic Dis 2017;14:170066. DOI: http://dx.doi.org/10.5888/pcd14.170066external icon.

Satia J. A. (2009). Diet-related disparities: understanding the problem and accelerating solutions. *Journal of the American Dietetic Association*, *109*(4), 610–615. https://doi.org/10.1016/j.jada.2008.12.019

Rebecca A. Seguin, Anju Aggarwal, Francoise Vermeylen, Adam Drewnowski, "Consumption Frequency of Foods Away from Home Linked with Higher Body Mass Index and Lower Fruit and Vegetable Intake among Adults: A Cross-Sectional Study", *Journal of Environmental and Public Health*, vol. 2016, Article ID 3074241, 12 pages, 2016. https://doi.org/10.1155/2016/3074241

Sinclair, S. E., Cooper, M., & Mansfield, E. D. (2014). The Influence of Menu Labeling on Calories Selected or Consumed: A Systematic Review and Meta-Analysis. *Journal of the Academy of Nutrition and Dietetics, 114*(9). doi:10.1016/j.jand.2014.05.014

Taylor, M. K., Sullivan, D. K., Ellerbeck, E. F., Gajewski, B. J., & Gibbs, H. D. (2019). Nutrition literacy predicts adherence to healthy/unhealthy diet patterns in adults with a nutrition-related chronic condition. *Public health nutrition*, *22*(12), 2157–2169. https://doi.org/10.1017/S1368980019001289

Wang, Y., & Chen, X. (2011). How much of racial/ethnic disparities in dietary intakes, exercise, and weight status can be explained by nutrition- and health-related psychosocial factors and socioeconomic status among US adults?. *Journal of the American Dietetic Association*, *111*(12), 1904–1911. https://doi.org/10.1016/j.jada.2011.09.036

Donati Zeppa, S., Sisti, D., Amatori, S., Gervasi, M., Agostini, D., Piccoli, G., Bertuccioli, A., et al. (2020). High-intensity Interval Training Promotes the Shift to a Health-Supporting Dietary Pattern in Young Adults. *Nutrients*, *12*(3), 843. MDPI AG. Retrieved from http://dx.doi.org/10.3390/nu12030843

Tables

Descriptive Variables

Table 1. Descriptive Statistics for Variables Used in Analysis of Relationships with Fruit and Vegetable Consumption (2020)

	N	Mean	Std Deviation	Range
Continuous Variables				
Vegetables	2,673	2.91	1.34	0-6
Fruit	2,673	2.44	1.35	0-6
Age	2,673	51.99	15.96	18-97
BMI	2,673	27.81	5.82	14.9-62.9
TimesModerateExercise	2,673	3.73	1.77	1-7
HowLongModerateExerciseMinutes	2,673	55.08	56.86	0-600
EducA	2,673	3.41	0.8	1-4
Income	2,673	6.18	2.05	1-9

Table 2

Gender	N	%
Female	1496	56%
Male	1177	44%
Total	**2673**	**100%**

Table 3

RuralUrban	N	%
Rural	25	1%
UrbanMetro	2419	90%
UrbanNonMetro	229	9%
Total	**2673**	**100%**

Table 4

Race	N	%
African American	314	12%
Asian	145	5%
Hispanic	398	15%
Other	87	3%
White	1729	65%
Total	**2673**	**100%**

Table 5

Region	N	%
Midwest	443	17%
Northeast	363	14%

South	1145	43%
West	722	27%
Total	**2673**	**100%**

Table 6

OwnAbilityTakeCareHealth	N	%
A Little Confident	62	2%
Completely Confident	779	29%
Not Confident At All	12	0%
Somewhat Confident	524	20%
Very Confident	1296	48%
Total	**2673**	**100%**

Table 7

smokeStatus	N	%
Current	232	9%
Former	664	25%
Never	1777	66%
Total	**2673**	**100%**

Table 8

Tablet_AchieveGoal	N	%
No	1364	51%
Yes	1309	49%
Total	**2673**	**100%**

Table 9

NoticeCalorieInfoOnMenu	N	%
No	1248	47%
Yes	1425	53%
Total	**2673**	**100%**

Table 10

SeekHealthInfo	N	%
No	323	12%
Yes	2350	88%
Total	**2673**	**100%**

T-Testing

Table 11. Mean Differences in Fruit and Vegetable Consumption by Gender

Gender	Fruits+Vegetables	
	Mean	N
Female	5.61***	1,496
Male	5.04***	1,177

Table 12. Mean Differences in Fruit and Vegetable Consumption by Health App Usage

Health App Usage	Fruits+Vegetables	
	Mean	N
Yes	5.60***	1,309
No	5.17***	1,364

Table 13. Mean Differences in Fruit and Vegetable Consumption by Noticing Calorie Information on the Menu

Noticing Calorie Information on the Menu	Fruits+Vegetables	
	Mean	N
Yes	5.54***	1,425
No	5.15***	1,248

Table 14. Mean Differences in Fruit and Vegetable Consumption by Seeking Health Information

Seeking Health Information	Fruits+Vegetables	
	Mean	N
Yes	5.43***	2,350
No	4.84***	323

ANOVA Testing:

Table 15. Average Fruit and Vegetable Consumption by Region

Region	Fruit+Vegetable Consumption (Mean values)
Midwest	5.38
Northeast	5.71
South	5.21
West	5.4
P value for F test	.004

Table 16. Average Fruit and Vegetable Consumption by urbanicity

Urbanicity	Fruit+Vegetable Consumption (Mean values)
Rural	4.76
Urban Metro	5.38
Urban Non-Metro	5.23
P value for F test	.28

Table 17. Average Fruit and Vegetable Consumption by Race/Ethnicity

Race/Ethnicity	Fruit+Vegetable Consumption (Mean values)
African American	5.38
Asian	5.26
Hispanic	4.99
Other	5.39
White	5.45
P value for F test	0.01

Table 18. Average Fruit and Vegetable Consumption by One's Own Ability to Take Care of Health

Own Ability to Take Care of Health	Fruit+Vegetable Consumption (Mean values)
Not Confident at All	4.5
A Little Confident	4.34
Somewhat Confident	4.77
Very Confident	5.4
Completely Confident	5.79
P value for F test	0.00000000000000165

Table 19. Average Fruit and Vegetable Consumption by Smoker Status

Smoker Status	Fruit+Vegetable Consumption (Mean values)
Current	4.46
Former	5.42
Never	5.5
P value for F test	0.00000000359