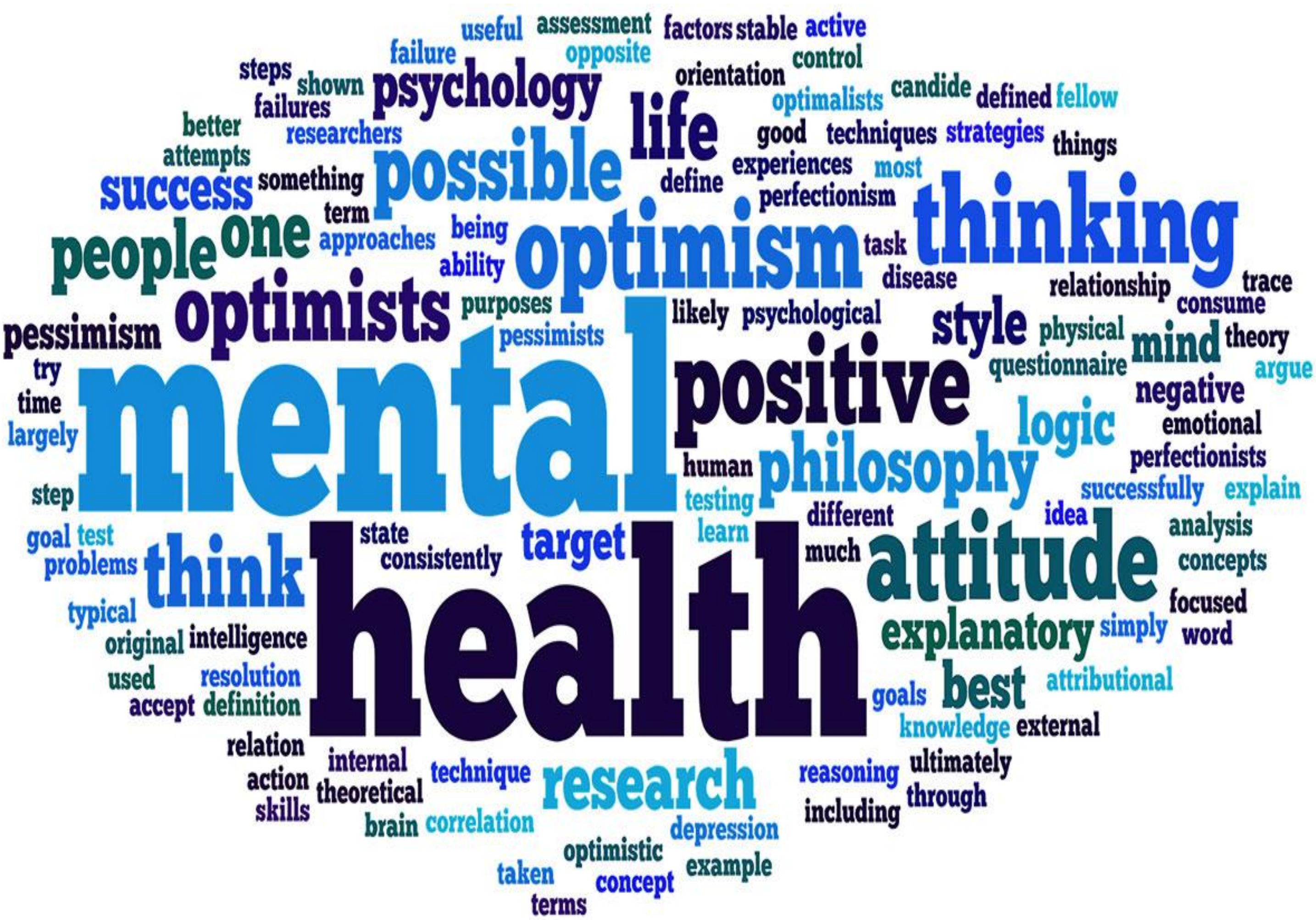
FACTORS AFFECTING MENTAL HEALTH

Group 2 - Harshitha Patlolla, Sai Kiran Batchu, Hema Naga Rupa Sri Maturi, Amit Sonagra

ISM 6137 – Statistical Data Mining



CONTENTS

1. EXECUTIVE SUMMARY

One of the most prevalent medical problems in the US is mental illness. Mental health includes all aspects of physical, psychological, emotional, and social wellbeing. It influences our thoughts, emotions, and behaviors. It also affects the ability to handle stress, interact with people, and make good decisions. Both mental and physical health are crucial aspects of overall health. For instance, depression raises the chance of developing a wide range of physical health issues, especially chronic diseases like diabetes, heart disease, and stroke. Similarly, having chronic illnesses raises the likelihood of developing a mental disease.

This analysis assesses the factors contributing to mental health in the US population. The data was sourced from CDC's Behavioral Risk Factor Surveillance System (BRFSS), the world's largest, ongoing telephonic health survey system that monitors health-related risk behaviors, chronic health conditions, and preventive services across the United States. This data consists of 450k records and 303 features, which are essentially survey questions that gather information on perceived health status (physical and mental health), demographics, economic and social factors, and other health behaviors (smoking, drinking, physical activity). The analysis starts with exploratory data analysis, handling missing values, feature engineering, and reducing the number of features to 39 by performing calculations and combining features to enable complete analysis. Statistical models (Poisson, Quasi Poisson, and negative binomial) were used to estimate the effects of explanatory variables on the number of bad mental health days. The study shows a significant relationship between poor mental health days and general health, difficulty concentrating, age, exercise, and cigarette smoking.

The findings of this study may provide public health officials and populations at risk for mental health with information about the factors contributing to mental illness which will aid them in planning, developing, implementing, and assessing control strategies.

2. PROBLEM STATEMENT AND SIGNIFICANCE

Poor mental health is a major contributor to disability in the world and represents an important public health problem. According to the World Health Organization (WHO), mental illnesses account for more collective disability burden in developed countries than any other group of illnesses, including cancer and heart disease[1]. WHO estimates that about 14 percent of the global burden of all diseases can be attributed to mental, neurological, or substance use disorders.

In the United States, National Alliance on Mental Illness (NAMI) has reported that 1 in 5 adults (21%), have experienced mental illness, and 1 in 20 adults (5.6%) had a severe mental illness in 2020. The U.S. has spent around $280 billion on mental health services in 2020, with the Medicaid program accounting for about 25% of that total.

The objective of this analysis is to find the factors that affect the mental health of an individual. When an individual suffers from mental health issues, our economy as a whole loses on productive work hours and it can cause a cascading effect on their families at the same time.

3. PRIOR LITERATURE

#	Title	Findings	Authors
1	Effect of Inadequate Sleep on Frequent Mental Distress	The study used data from 2018 BRFSS survey. After removing participants who did not fall into specific age bucket and who did not have data on mental distress, they defined mental distress as participants whose mental health was not good for more than 14 days. Using logistic regression authors found that association between self-reported sleep data and mental health.	Blackwelder A, Hoskins M, Huber L. [2]

		Some of the predictors are – Marital Status, Annual HH Income, Binge Drinking in last 30 days, Smoke, Lost health coverage in last year, Age, Gender, Race	
2	A Comparison of Depression and Mental Distress Indicators, Rhode Island Behavioral Risk Factor Surveillance System, 2006	The study tries to find simpler and less time-consuming way to find PHQ8 indicator using BRFSS survey. Using the # of days mental health was not good var, they crated a new var and ran logistic regression. Multiple imputation was done for missing data. They could not establish that BRFSS survey data can replace PHQ8 indicator. Some of the predictors are – Age, Gender, Annual Income, Employment status, Current Smoker, Chronic Drinker, Asthma, Diabetes, Obesity, Disability	Jiang Y, Hesser JE. [3]
3	Using the Behavioral Risk Factor Surveillance System to Assess Mental Health, Travis County, Texas, 2011–2016	Logistic regression models were used to detect relationships between each chronic condition and depression or poor mental health by regulating other demographic factors. Adults who were diagnosed with depression more frequently than poor mental health. They stratified the prevalence of depressive disorder and poor mental health by demographic features, healthcare access, risk behaviors, and chronic disease. When compared to respondents without these problems, respondents with chronic health disorders had a considerably greater prevalence of depression and poor mental health.	Haruna Miyakado-Steger, MS1; Sarah Seidel, DrPH [4]
4	Socioeconomic factors and happiness: evidence from self-reported mental health data	This study analyzes how mental health correlates with economic (income, employment, etc.), demographic and social factors using BRFSS data. It analyzes the impact of these factors on people's happiness by treating the number of bad mental health days as an indirect indicator of the respondent's (un)happiness using linear regression. Some findings include Self-reported mental health changes most with age, employment situation, and marital status. Also, the relationship between relative income and happiness is stronger among women than men. Women's happiness is more related to factors such as height and weight, also found evidence of better mental health in generations born before WWII, after accounting for age and other personal variables.	Jacek Rothert1 · Douglas VanDerwerken2 · Ethan White3 [5]
5	The positive association between employment and self-reported mental health in the USA	This study uses BRFSS data to investigate the association between employment and mental health among US adults by employing logistic regression and marginalized negative binomial regression models. Respondents with arthritis or a stroke had the largest relative risk of mental ill days when compared to those who were not employed. Employed men had 25% lower risk of mental ill days compared to women. Overall, people who have chronic illnesses reported having more mean mental unhealthy days.	Chinaeke Eric, Gwynn Melanie, Hong Yuan, Zhang Jiajia, Olatosi Bankole [6]
6			

4. DATA SOURCE AND PREPARATION

The Behavioral Risk Factor Surveillance System (BRFSS) is a telephone survey administered in all 50 states, the District of Columbia, Puerto Rico, the US Virgin Islands, and Guam with funding and specifications from the Centers for Disease Control and Prevention (CDC). The BRFSS monitors the prevalence of behavioral risks for the leading causes of disease and death among adults in the United States.

Dataset for the year 2021 is used in this study. This data consists of 450k records and 303 features, which are essentially survey questions that gather information on perceived health status (physical and mental health), demographics, economic and social factors, and other health behaviors (smoking, drinking, physical activity).

2021 BRFSS Data (ASCII) - Survey data for the year 2021

2021 BRFSS Overview CDC - Data Dictionary

Data Cleaning:

Performed data cleaning in python

- Considered only the records where mental health was reported
- Removed fields which had more than 70% of the missing data
- For the selected features, rows where response was missing or refused or don't know were dropped as it biases the data if any assumption was made about it
- Dropped variables that were calculated from existing fields
- Dropped fields related to interview conditions like interview month, interview day, interview year
- Derived state names from state fips code

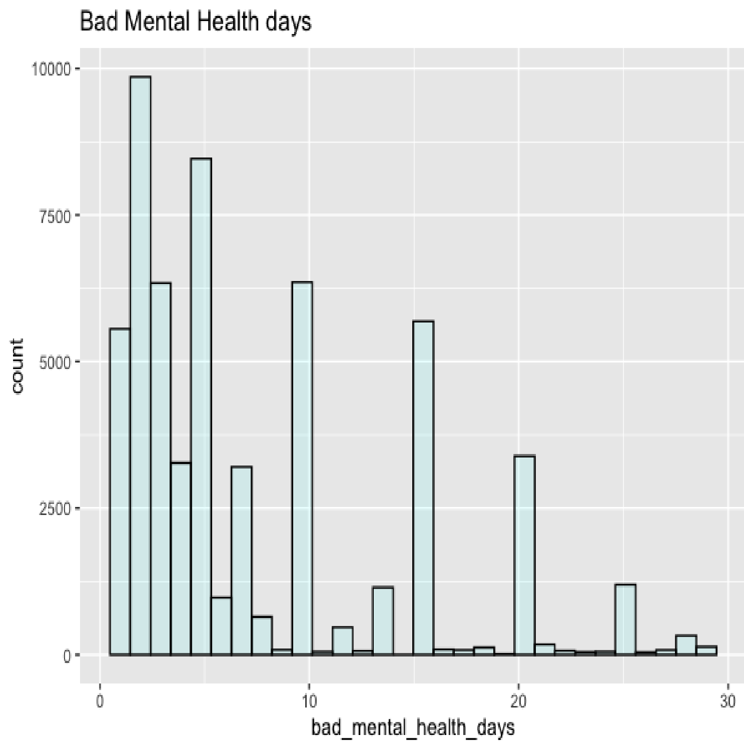
After data cleaning and feature engineering, the final dataset had **39 columns** and **57999 rows**.

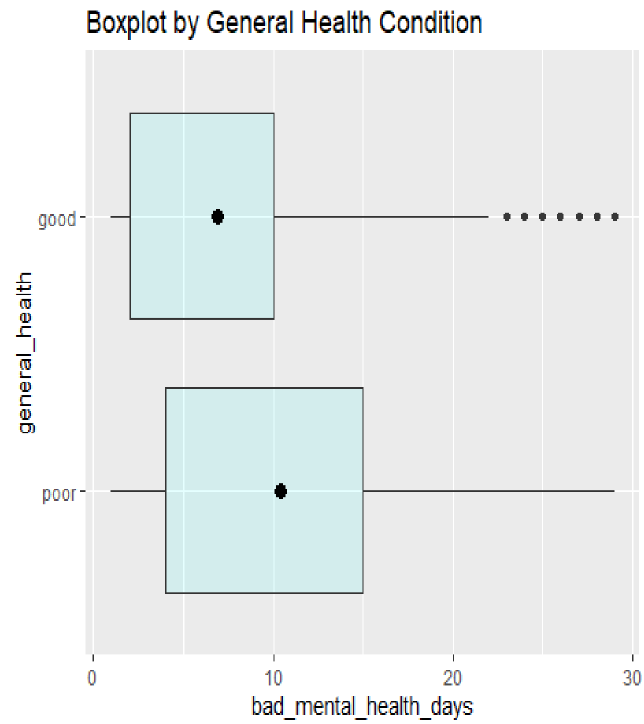
5. VARIABLE CHOICE

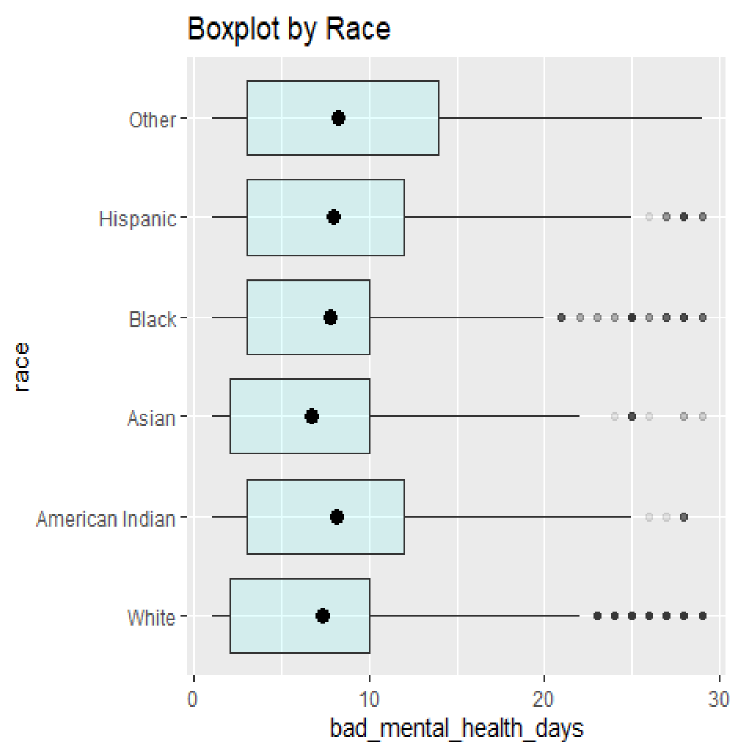
<table>
<tr><th>Predictor</th><th>Sign of effect</th><th>Rationale</th></tr>
<tr><td>State</td><td>+/-</td><td>Mental health issues in people may vary based on state demographics and policies</td></tr>
<tr><td>Sex</td><td>+/-</td><td>Females are more likely to experience more mental health issues mostly due to hormonal changes</td></tr>
<tr><td>General_health</td><td>+</td><td>A person with good general health is less likely to experience mental health disorders</td></tr>
<tr><td>Bad mental health days</td><td></td><td>Dependent Variable</td></tr>
<tr><td>Routine_Checkup</td><td>+</td><td>Regular routine checkups might improve mental wellness.</td></tr>
<tr><td>Exercise</td><td>+</td><td>Regular exercise releases happy hormones that reduce stress</td></tr>
<tr><td>Cholesterol</td><td>-</td><td>High cholesterol increases the risk of heart disease and has a negative impact on mental health</td></tr>
<tr><td>Asthma</td><td rowspan="5">-</td><td rowspan="5">Physical illnesses have a negative impact on mental health</td></tr>
<tr><td>Chronic_Bronchitis</td></tr>
<tr><td>Kidney_Disease</td></tr>
<tr><td>Diabetes</td></tr>
<tr><td>Arthritis</td></tr>
</table>

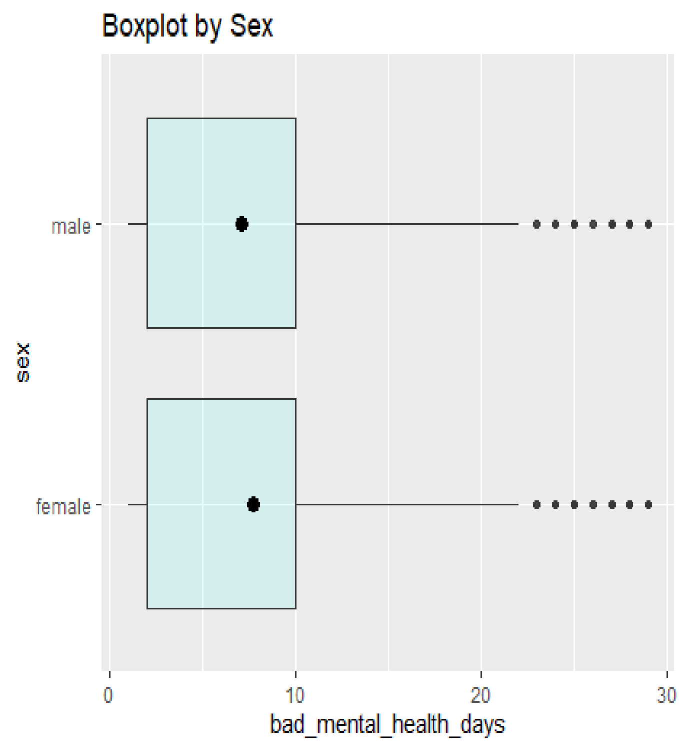
Heart_Diseases		
Cancer		
Difficulty_Doing_Errands		
Difficulty_Hearing		
Difficulty_Seeing		
Difficulty_Concentrating		Difficulty concentrating is correlated with mental health conditions like depression
Marital_status	+/-	Separated/Widowed individuals are likely to be more depressed than married, unmarried
Education_level	+	Higher-educated individuals have better control of their life and experience fewer mental health issues.
House	+/-	People with their own house feel safer and experience fewer mental health problems
Veteran	-	Soldiers generally experience more trauma, which leads to poor mental health.
Employment_status	+/-	Unemployed individuals are more mentally impacted than employed individuals
Children	-	Responsibilities increase as the number of children in a household increases which might lead to mental stress and disorders
Height_Inch	+/-	
Weight_kg		Physical appearance can affect mental health
BMI		
Smoking	-	Smoking increases anxiety and tension and can add to mental illnesses.
Bad physical health days	-	As the number of days with bad physical health increases stress levels also increase that deteriorates mental health
Health_Insurance	+	People without health insurance have higher stress levels than those who are insured.
Personal Health Care Provider	+	Having a personal health care provider gives a secure feeling that might help with mental health conditions
Afford to see doctor	+	Being able to afford medical care has a positive effect on mental health.
Income_Level	+/-	People with lower income have a higher likelihood of developing mental health problems.
Flushot	+	A flu shot can increase one's sense of security and boost their mental health.
County_type	+/-	The prevalence of depression is higher in residents of rural areas compared to urban areas
Race	+/-	Racial and ethnic minority groups have disparities in mental health
Age	+/-	Mental health issues may be higher in younger adults than in older people.
Binge_Drinking	-	Heavy drinking interferes with chemicals in the brain that are vital for good mental health.
Diet	+/-	Healthy diet helps with mental health due to antioxidants and vitamins while fried foods can cause higher risk of developing depression

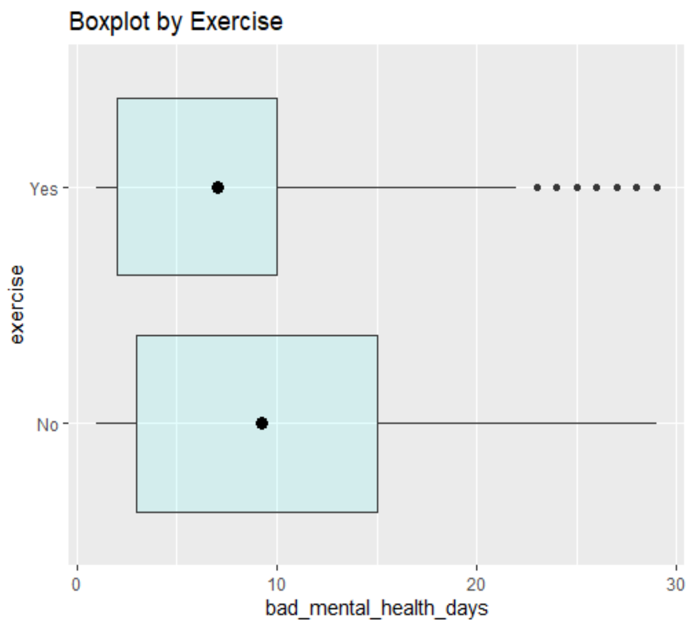
6. EXPLORATORY DATA ANALYSIS & VISUALIZATIONS


Bad Mental Health days
count
bad_mental_health_days
10000
7500
5000
2500
0
0
10
20
30


Boxplot by General Health Condition
general_health
good
poor
bad_mental_health_days
0
10
20
30


Boxplot by Race
race
Other
Hispanic
Black
Asian
American Indian
White
bad_mental_health_days
0
10
20
30


Boxplot by Sex
sex
male
female
bad_mental_health_days
0
10
20
30

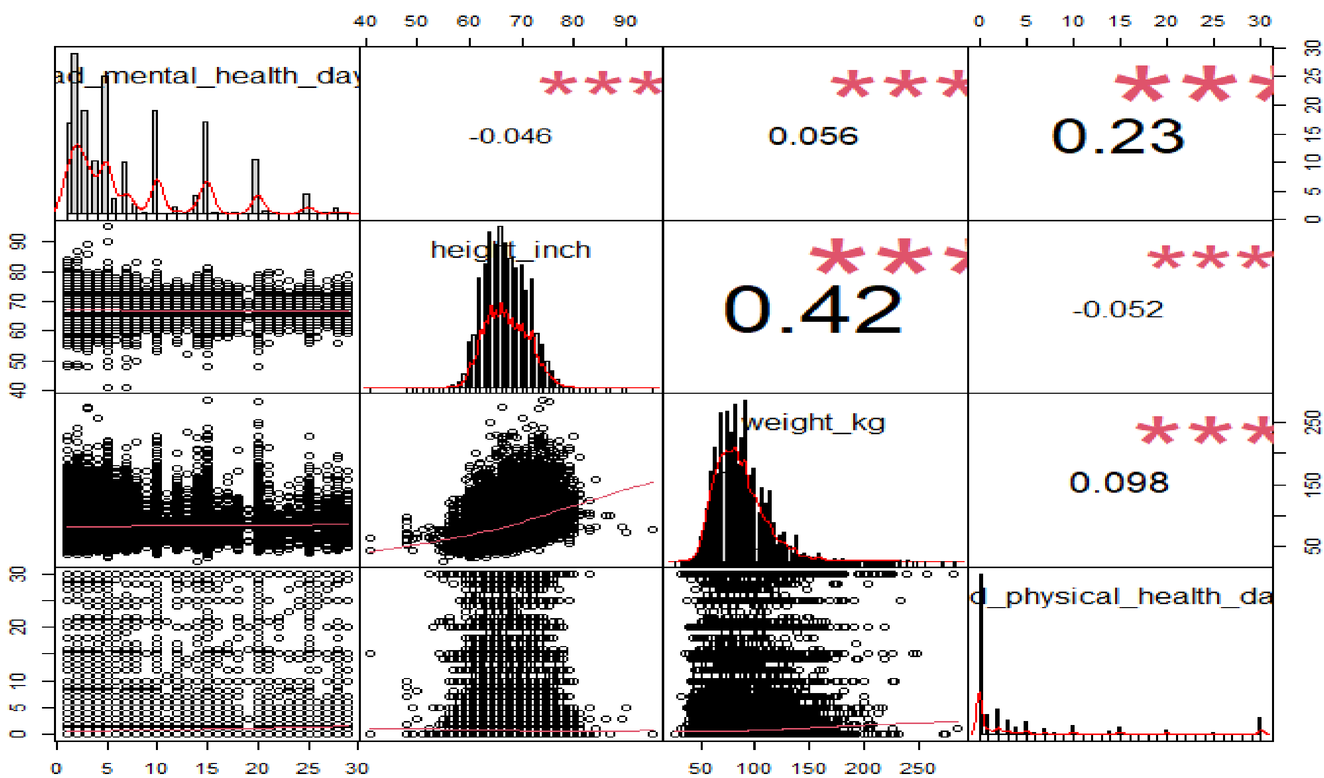


The histogram of bad mental health days is right skewed with some peaks at specific intervals.

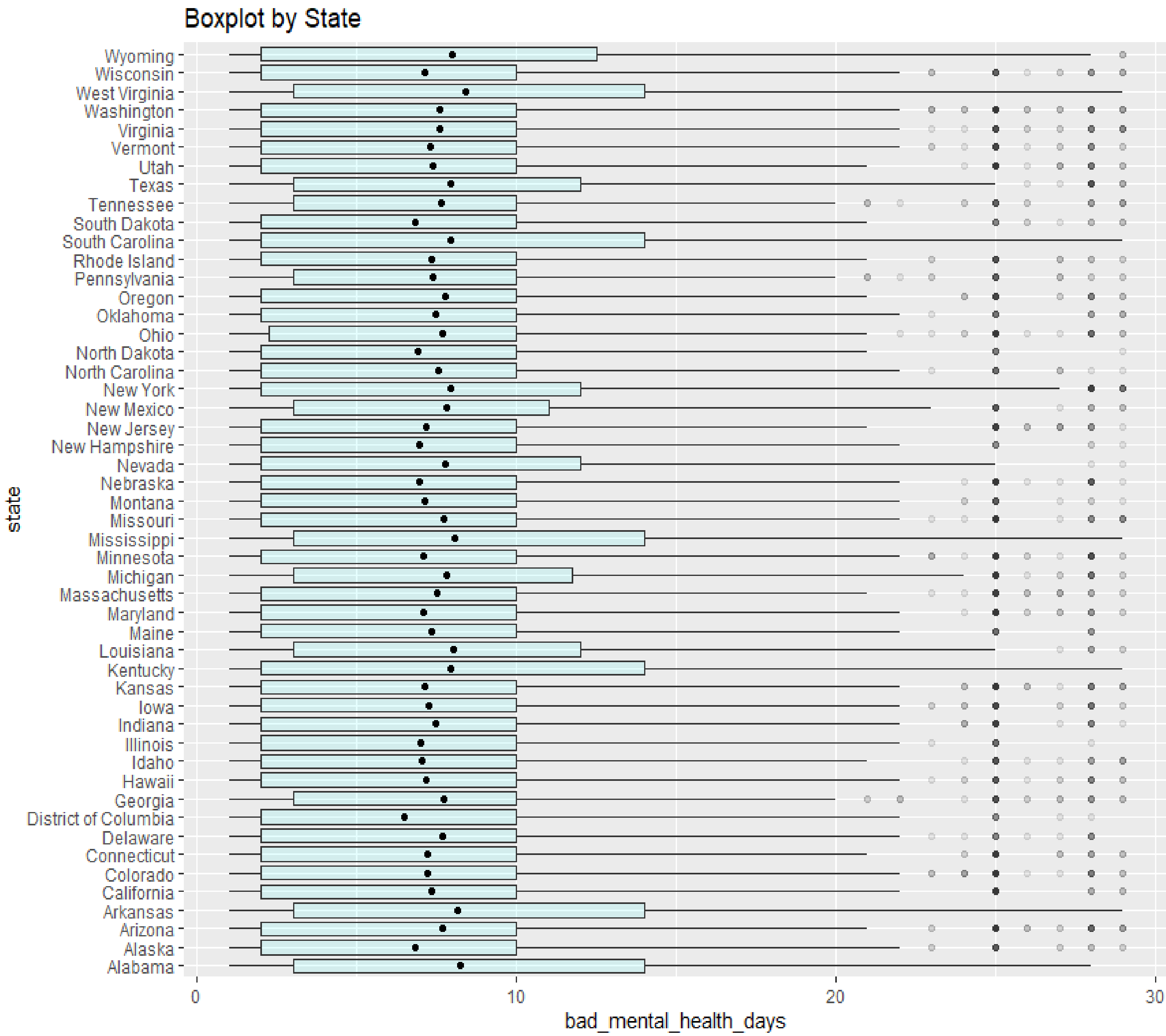
People with poor general health tend to experience more mental illness probably because poor general health condition usually refers to people suffering from any kind of disease or illness. American Indians and other races have more mental illness days when comparing different races.

Females suffer more from mental illness than males as they tend to have more hormonal imbalances. It is apparent that exercise improves physical and mental health hence people without exercise have a significant difference in average bad mental health days.

Correlation plot of Numeric Variables



It is important to check the multicollinearity of the numeric variables as they potentially skew the model outputs. We see that the highest correlation is between height and weight which is inherent. As any of the correlations is not near or beyond 0.7 there shouldn't be any multicollinearity problem and skewed model outputs.



Alabama, West Virginia, and Arkansas are the top 3 states with the highest average bad mental health days when compared to other states.

7. MODELS

The bad_mental_health_days dependent variable is count type hence Poisson models are appropriate to use. Initially Poisson model was used and to overcome the overdispersion problem Quasi Poisson and Negative Binomial models were employed. It was apparent that the negative binomial model performed the best among the three models.

Type	Model
Poisson	poisson = glm(bad_mental_health_days ~ state + general_health + routine_checkup + cholesterol + chronic_bronchitis + diabetes + marital_status + house + employment_status + height_inch + bmi + bad_physical_health_days + personal_health_care_provider + heart_diseases + income_level + difficulty_hearing + difficulty_concentrating + difficulty_doing_errands + difficulty_seeing + metropolitan_county + age + diet + sex + exercise + asthma + kidney_disease + arthritis + education_level + veteran + children + weight_kg + smoking + health_insurance + afford_to_see_doctor + cancer + flushot + race + binge_drinking , family = "poisson" (link=log), data = df)
Quasi Poisson	quasipoisson = glm(bad_mental_health_days ~ state + general_health + routine_checkup + cholesterol + chronic_bronchitis + diabetes + marital_status + house + employment_status + height_inch + bmi + bad_physical_health_days + personal_health_care_provider + heart_diseases + income_level + difficulty_hearing + difficulty_concentrating + difficulty_doing_errands + difficulty_seeing + metropolitan_county + age + diet + sex + exercise + asthma + kidney_disease + arthritis + education_level + veteran + children + weight_kg + smoking + health_insurance + afford_to_see_doctor + cancer + flushot + race + binge_drinking , family = "quasipoisson" (link=log), data = df)
Negative Binomial	negativeBinomial = glm.nb(bad_mental_health_days ~ state + general_health + routine_checkup + cholesterol + chronic_bronchitis + diabetes + marital_status + house + employment_status + height_inch + bmi + bad_physical_health_days + personal_health_care_provider + heart_diseases + income_level + difficulty_hearing + difficulty_concentrating + difficulty_doing_errands + difficulty_seeing + metropolitan_county + age + diet + sex + exercise + asthma + kidney_disease + arthritis + education_level + veteran + children + weight_kg + smoking + health_insurance + afford_to_see_doctor + cancer + flushot + race + binge_drinking, data = df)

8. ASSUMPTIONS TESTING

Dispersion test:

The dispersion test for Poisson model shows the presence of overdispersion as the value 4.817 is more than the optimal value of 1.

```
> dispersiontest(poisson)

        Overdispersion test

data:  poisson
z = 109, p-value < 2.2e-16
alternative hypothesis: true dispersion is greater than 1
sample estimates:
dispersion
  4.817501
```

Durbin Watson Test for multicollinearity:
The DW test specification of 2.0032 for the negative binomial indicates no presence of multicollinearity between the independent variables.

```
> dwtest(negativeBinomial)

        Durbin-Watson test

data:  negativeBinomial
DW = 2.0032, p-value = 0.5713
alternative hypothesis: true autocorrelation is greater than 0
```

VIF Test for Independence Assumption:
As all the values of GVIF^(1/(2*Df)) are less than 5 the independence assumption of the observations is satisfied.

9. INSIGHTS AND RECOMMENDATIONS

From the beta coefficients of negative binomial model, Wyoming, Alabama, and West Virginia are the top 3 states and South Dakota, Illinois and Alaska are the least 3 states with bad mental health days. Adults with good general health tend to be 11.4 % (~ 3.3 days) less mentally ill than adults with bad general health condition.

Adults who are physically active i.e., exercise regularly reported 11.6% (~3.4 days) less bad mental health days. Cigarette smoking and e-cigarette smoking increase bad mental health days by 8.7% (~2.5 days) and 11.5 % (~ 3.3 days) respectively. Females are more mentally ill (9.8%) than men. Higher levels of education add 4.6 days additional mental illness days to an individual.

Recommendations:

- The states of Wyoming, Alabama, and West Virginia should design better healthcare benefits like affordable insurance for their residents as the affordability to see a therapist has a greater impact on the reduction of mental illnesses
- Create awareness of the benefits of exercise and maintaining good general health conditions
- Adults should consider quitting smoking or binge drinking as it can improve their mental health
- Colleges and high schools can establish and promote the campus health and wellness centers or set up counseling services

10. REFERENCES

[1] World Health Organization: Promoting mental health: concepts, emerging evidence, practice (summary report). 2004, Geneva (CH): World Health Organization

[5] Rothert, J., VanDerwerken, D. & White, E. Socioeconomic factors and happiness: evidence from self-reported mental health data. *Empir Econ* **58**, 3101–3123 (2020). https://doi.org/10.1007/s00181-019-01655-y

[6] Chinaeke Eric, Gwynn Melanie, Hong Yuan, Zhang Jiajia, Olatosi Bankole, The positive association between employment and self-reported mental health in the USA: a robust application of marginalized zero-inflated negative binomial regression (MZINB), *Journal of Public Health*, Volume 42, Issue 2, June 2020, Pages 340–352, https://doi.org/10.1093/pubmed/fdaa030

https://www.bu.edu/sph/news/articles/2019/public-health-means-mental-health/#:~:text=Mental%20health%20is%20truly%20a,the%20drivers%20of%20physical%20health.

11. APPENDIX

R – Code:

```
# imports
rm(list = ls())
library(readxl)
library(dplyr)
library(corrplot)
library(lme4)
library(MASS)
library(AER)
library(stargazer)
library(ggplot2)
library(lattice)
library(PerformanceAnalytics)

df = read_xlsx("dataset.xlsx")

colnames(df) = tolower(colnames(df))
colnames(df)

#------------------------------------------------------
# Visualizations
#------------------------------------------------------
# Box plots

# by state
ggplot(df, aes(x=state, y=bad_mental_health_days)) +  coord_flip()+
  geom_boxplot(fatten = NULL, fill="cyan", alpha=0.1) +
  ggtitle("Boxplot by State") +
  stat_summary(fun=mean, geom='point', size = 1.5)

# by health condition
ggplot(df, aes(x=general_health, y=bad_mental_health_days)) +  coord_flip()+
  geom_boxplot(fatten = NULL, fill="cyan", alpha=0.1) +
  ggtitle("Boxplot by General Health Condition") +
  stat_summary(fun=mean, geom='point', size = 3)

#by race
ggplot(df, aes(x=race, y=bad_mental_health_days)) +  coord_flip()+
  geom_boxplot(fatten = NULL, fill="cyan", alpha=0.1) +
  ggtitle("Boxplot by Race") +
  stat_summary(fun=mean, geom='point', size = 3)

# by sex
ggplot(df, aes(x=sex, y=bad_mental_health_days)) +  coord_flip()+
  geom_boxplot(fatten = NULL, fill="cyan", alpha=0.1) +
  ggtitle("Boxplot by Sex") +
  stat_summary(fun=mean, geom='point', size = 3)

# by exercise
```

```
ggplot(df, aes(x=exercise, y=bad_mental_health_days)) +  coord_flip()+
  geom_boxplot(fatten = NULL, fill="cyan", alpha=0.1) +
  ggtitle("Boxplot by Exercise") +
  stat_summary(fun=mean, geom='point', size = 3)


# Histogram
ggplot(df, aes(x=bad_mental_health_days)) +
  geom_histogram(color="darkblue", fill="cyan", alpha =0.25) +
  ggtitle("Histogram of Bad Mental Health days")

# Correlation plot
dtemp = df[c(3 , 14, 15, 18 )]
chart.Correlation(dtemp)

#------------------------------------------------------
# Converting to factor variables and releveling the data
#------------------------------------------------------
df$general_health = factor(df$general_health)
df$general_health = relevel(df$general_health, "poor")

df$routine_checkup = factor(df$routine_checkup)
df$routine_checkup = relevel(df$routine_checkup, "Never")

df$marital_status = factor(df$marital_status)
df$marital_status = relevel(df$marital_status, "Unmarried")

df$education_level = factor(df$education_level)
df$education_level = relevel(df$education_level, "Kindergarten school")

df$house = factor(df$house)
df$house = relevel(df$house, "Rent")

df$employment_status = factor(df$employment_status)
df$employment_status = relevel(df$employment_status, "Unemployed")

df$children = factor(df$children)
df$children = relevel(df$children, "None")

df$bmi = factor(df$bmi)
df$bmi = relevel(df$bmi, "NormalWeight")

df$smoking = factor(df$smoking)
df$smoking = relevel(df$smoking, "no")

df$income_level = factor(df$income_level)
df$income_level = relevel(df$income_level, "less than 25k")

df$race = factor(df$race)
df$race = relevel(df$race, "White")

df$age = factor(df$age)
df$age = relevel(df$age, "18-24")

df$diet = factor(df$diet)
df$diet = relevel(df$diet, "mixed diet")

str(df)
```

```
attach(df)



#------------------------------------------------------
# Models
#------------------------------------------------------
# Poisson model
poisson <- glm(bad_mental_health_days ~ state + general_health + routine_checkup + cholesterol + chronic_bronchitis + diabetes + marital_status + house + employment_status + height_inch + bmi + bad_physical_health_days + personal_health_care_provider + heart_diseases + income_level + difficulty_hearing + difficulty_concentrating + difficulty_doing_errands + difficulty_seeing + county_type + age + diet + sex + exercise + asthma + kidney_disease + arthritis + education_level + veteran + children + weight_kg + smoking + health_insurance + afford_to_see_doctor + cancer + flushot + race + binge_drinking , family = "poisson" (link=log), data = df)

summary(poisson)

# QuasiPoisson model to control over dispersion
quasipoisson <- glm(bad_mental_health_days ~ state + general_health + routine_checkup + cholesterol + chronic_bronchitis + diabetes + marital_status + house + employment_status + height_inch + bmi + bad_physical_health_days + personal_health_care_provider + heart_diseases + income_level + difficulty_hearing + difficulty_concentrating + difficulty_doing_errands + difficulty_seeing + county_type + age + diet + sex + exercise + asthma + kidney_disease + arthritis + education_level + veteran + children + weight_kg + smoking + health_insurance + afford_to_see_doctor + cancer + flushot + race + binge_drinking , family = "quasipoisson" (link=log), data = df)

summary(quasipoisson)

# NegativeBinomial Model to control overdispersion
negativeBinomial = glm.nb(bad_mental_health_days ~ state + general_health + routine_checkup + cholesterol + chronic_bronchitis + diabetes + marital_status + house + employment_status + height_inch + bmi + bad_physical_health_days + personal_health_care_provider + heart_diseases + income_level + difficulty_hearing + difficulty_concentrating + difficulty_doing_errands + difficulty_seeing + county_type + age + diet + sex + exercise + asthma + kidney_disease + arthritis + education_level + veteran + children + weight_kg + smoking + health_insurance + afford_to_see_doctor + cancer + flushot + race + binge_drinking , data = df)

summary(negativeBinomial)

#------------------------------------------------------
#Stargazer Output of the models
#------------------------------------------------------
stargazer(poisson, quasipoisson, negativeBinomial, single.row=TRUE, type="text")

#------------------------------------------------------
#Quality checks
#------------------------------------------------------
# Dispersion Test for poisson model
dispersiontest(poisson)

# for Negative Binomial Models
# Durbin Watson test for autocorrelation
dwtest(negativeBinomial)

# VIF test for independence
vif(negativeBinomial)
```

Stargazer output:

	Dependent variable:		
	bad_mental_health_days		
	Poisson	glm: quasipoisson link = log	negative binomial
	(1)	(2)	(3)
stateAlaska	-0.167*** (0.019)	-0.167*** (0.042)	-0.168*** (0.041)
stateArizona	-0.062*** (0.016)	-0.062* (0.035)	-0.059* (0.035)
stateArkansas	-0.052** (0.021)	-0.052 (0.046)	-0.050 (0.047)
stateCalifornia	-0.094*** (0.017)	-0.094** (0.037)	-0.102*** (0.037)
stateColorado	-0.090*** (0.016)	-0.090*** (0.034)	-0.093*** (0.034)
stateConnecticut	-0.063*** (0.017)	-0.063* (0.037)	-0.063* (0.036)
stateDelaware	-0.065*** (0.021)	-0.065 (0.046)	-0.061 (0.046)
stateDistrict of Columbia	-0.122*** (0.021)	-0.122*** (0.046)	-0.127*** (0.044)
stateGeorgia	-0.093*** (0.018)	-0.093** (0.039)	-0.093** (0.039)
stateHawaii	-0.057*** (0.017)	-0.057 (0.038)	-0.053 (0.038)
stateIdaho	-0.147*** (0.018)	-0.147*** (0.040)	-0.144*** (0.039)
stateIllinois	-0.150*** (0.023)	-0.150*** (0.050)	-0.169*** (0.049)
stateIndiana	-0.093*** (0.016)	-0.093*** (0.036)	-0.098*** (0.036)
stateIowa	-0.078*** (0.017)	-0.078** (0.036)	-0.074** (0.036)
stateKansas	-0.113*** (0.015)	-0.113*** (0.033)	-0.122*** (0.033)
stateKentucky	-0.125*** (0.019)	-0.125*** (0.042)	-0.117*** (0.042)
stateLouisiana	-0.045** (0.018)	-0.045 (0.040)	-0.056 (0.040)
stateMaine	-0.097*** (0.018)	-0.097** (0.040)	-0.107*** (0.040)
stateMaryland	-0.094*** (0.015)	-0.094*** (0.034)	-0.093*** (0.033)
stateMassachusetts	-0.051*** (0.017)	-0.051 (0.037)	-0.056 (0.037)
stateMichigan	-0.048*** (0.016)	-0.048 (0.036)	-0.044 (0.036)
stateMinnesota	-0.104*** (0.015)	-0.104*** (0.033)	-0.105*** (0.033)
stateMississippi	-0.092*** (0.020)	-0.092** (0.044)	-0.068 (0.045)
stateMissouri	-0.106*** (0.016)	-0.106*** (0.035)	-0.106*** (0.035)
stateMontana	-0.094*** (0.019)	-0.094** (0.041)	-0.104*** (0.040)
stateNebraska	-0.122*** (0.016)	-0.122*** (0.034)	-0.131*** (0.034)
stateNevada	-0.063*** (0.022)	-0.063 (0.047)	-0.045 (0.047)
stateNew Hampshire	-0.114*** (0.021)	-0.114** (0.045)	-0.132*** (0.044)
stateNew Jersey	-0.095*** (0.017)	-0.095** (0.038)	-0.090** (0.038)
stateNew Mexico	-0.053*** (0.017)	-0.053 (0.038)	-0.050 (0.038)
stateNew York	-0.032** (0.014)	-0.032 (0.031)	-0.027 (0.032)
stateNorth Carolina	-0.077*** (0.019)	-0.077* (0.041)	-0.068* (0.041)
stateNorth Dakota	-0.142*** (0.021)	-0.142*** (0.046)	-0.153*** (0.044)
stateOhio	-0.081*** (0.015)	-0.081** (0.034)	-0.081** (0.034)
stateOklahoma	-0.121*** (0.020)	-0.121*** (0.043)	-0.124*** (0.043)
stateOregon	-0.059*** (0.018)	-0.059 (0.040)	-0.050 (0.040)
statePennsylvania	-0.076*** (0.017)	-0.076** (0.038)	-0.071* (0.038)
stateRhode Island	-0.079*** (0.018)	-0.079** (0.039)	-0.081** (0.039)
stateSouth Carolina	-0.048*** (0.017)	-0.048 (0.037)	-0.053 (0.038)
stateSouth Dakota	-0.167*** (0.019)	-0.167*** (0.043)	-0.172*** (0.042)
stateTennessee	-0.114*** (0.018)	-0.114*** (0.040)	-0.119*** (0.040)
stateTexas	-0.050*** (0.016)	-0.050 (0.035)	-0.056 (0.035)
stateUtah	-0.042*** (0.016)	-0.042 (0.034)	-0.043 (0.034)
stateVermont	-0.048*** (0.018)	-0.048 (0.039)	-0.055 (0.039)
stateVirginia	-0.054*** (0.017)	-0.054 (0.037)	-0.062* (0.037)
stateWashington	-0.038** (0.015)	-0.038 (0.034)	-0.041 (0.034)
stateWest Virginia	-0.017 (0.018)	-0.017 (0.039)	-0.020 (0.039)
stateWisconsin	-0.079*** (0.018)	-0.079** (0.039)	-0.083** (0.038)

stateWyoming	0.005 (0.023)	0.005 (0.050)	0.010 (0.051)
general_healthgood	-0.116*** (0.005)	-0.116*** (0.010)	-0.114*** (0.011)
routine_checkup5 or more years ago	-0.035 (0.036)	-0.035 (0.078)	-0.045 (0.080)
routine_checkupWithin past 2 years	-0.044 (0.035)	-0.044 (0.077)	-0.052 (0.079)
routine_checkupWithin past 5 years	-0.019 (0.035)	-0.019 (0.078)	-0.034 (0.079)
routine_checkupWithin past year	-0.065* (0.035)	-0.065 (0.077)	-0.075 (0.078)
cholesterolYes	0.021*** (0.005)	0.021** (0.010)	0.020** (0.010)
chronic_bronchitisYes	-0.021*** (0.006)	-0.021* (0.013)	-0.023* (0.014)
diabetesYes	-0.003 (0.005)	-0.003 (0.011)	-0.003 (0.011)
marital_statusMarried	-0.059*** (0.005)	-0.059*** (0.010)	-0.062*** (0.010)
marital_statusSeparated	0.012** (0.005)	0.012 (0.011)	0.017 (0.011)
marital_statusWidowed	0.012 (0.008)	0.012 (0.018)	0.021 (0.018)
houseOther	-0.003 (0.007)	-0.003 (0.016)	-0.0001 (0.016)
houseOwn	-0.027*** (0.004)	-0.027*** (0.009)	-0.028*** (0.009)
employment_statusEmployed	-0.025*** (0.005)	-0.025** (0.011)	-0.027** (0.012)
employment_statusHomemaker	-0.045*** (0.009)	-0.045** (0.020)	-0.054*** (0.020)
employment_statusRetired	-0.077*** (0.007)	-0.077*** (0.016)	-0.084*** (0.016)
employment_statusStudent	0.016* (0.010)	0.016 (0.021)	0.012 (0.022)
height_inch	-0.002*** (0.001)	-0.002 (0.001)	-0.002* (0.001)
bmiObese	0.046*** (0.006)	0.046*** (0.014)	0.046*** (0.014)
bmiOverWeight	0.021*** (0.004)	0.021** (0.010)	0.021** (0.009)
bmiUnderweight	0.063*** (0.013)	0.063** (0.029)	0.072** (0.030)
bad_physical_health_days	0.011*** (0.0002)	0.011*** (0.0005)	0.013*** (0.0005)
personal_health_care_providerYes	-0.001 (0.006)	-0.001 (0.012)	-0.001 (0.012)
heart_diseasesYes	-0.009 (0.006)	-0.009 (0.013)	-0.002 (0.013)
income_level100k-200k	-0.043*** (0.006)	-0.043*** (0.014)	-0.039*** (0.014)
income_level25k-50k	0.009* (0.005)	0.009 (0.011)	0.007 (0.012)
income_level50k-100k	-0.031*** (0.006)	-0.031** (0.012)	-0.035*** (0.013)
income_levelMore than 200k	-0.058*** (0.007)	-0.058*** (0.014)	-0.059*** (0.015)
difficulty_hearingYes	0.029*** (0.006)	0.029** (0.014)	0.034** (0.014)
difficulty_concentratingYes	0.345*** (0.004)	0.345*** (0.009)	0.349*** (0.009)
difficulty_doing_errandsYes	0.039*** (0.005)	0.039*** (0.011)	0.038*** (0.011)
difficulty_seeingYes	0.030*** (0.007)	0.030** (0.015)	0.038** (0.016)
county_typeUrban	0.023*** (0.005)	0.023* (0.012)	0.018 (0.012)
age25-34	-0.094*** (0.007)	-0.094*** (0.015)	-0.092*** (0.016)
age35-44	-0.158*** (0.008)	-0.158*** (0.017)	-0.160*** (0.017)
age45-54	-0.231*** (0.008)	-0.231*** (0.017)	-0.233*** (0.018)
age55-64	-0.280*** (0.008)	-0.280*** (0.019)	-0.290*** (0.019)
age64+	-0.375*** (0.010)	-0.375*** (0.022)	-0.381*** (0.022)
diethealthy diet	0.019*** (0.007)	0.019 (0.014)	0.025* (0.014)
sexmale	-0.092*** (0.005)	-0.092*** (0.010)	-0.098*** (0.010)
exerciseYes	-0.107*** (0.004)	-0.107*** (0.009)	-0.116*** (0.009)
asthmaYes	0.035*** (0.004)	0.035*** (0.009)	0.043*** (0.009)
kidney_diseaseYes	-0.024*** (0.008)	-0.024 (0.018)	-0.023 (0.019)
arthritisYes	0.016*** (0.004)	0.016* (0.009)	0.023*** (0.008)
education_levelCollege	0.131* (0.079)	0.131 (0.173)	0.171 (0.173)
education_levelElementary School	0.121 (0.080)	0.121 (0.176)	0.143 (0.176)
education_levelHigh School	0.131* (0.079)	0.131 (0.173)	0.172 (0.173)
veteranyes	0.056*** (0.006)	0.056*** (0.013)	0.056*** (0.012)
children1	-0.007 (0.005)	-0.007 (0.010)	-0.007 (0.010)
children2	-0.003 (0.005)	-0.003 (0.011)	-0.002 (0.011)
children3	-0.022*** (0.007)	-0.022 (0.016)	-0.018 (0.016)
childrenMore than 3	-0.022** (0.010)	-0.022 (0.022)	-0.021 (0.021)
weight_kg	0.0005*** (0.0001)	0.0005 (0.0003)	0.001** (0.0003)
smokingcigarettes	0.081*** (0.003)	0.081*** (0.007)	0.087*** (0.007)
smokinge-cigarettes	0.103*** (0.011)	0.103*** (0.023)	0.115*** (0.025)
health_insuranceyes	0.015** (0.007)	0.015 (0.016)	0.014 (0.017)
afford_to_see_doctorYes	-0.144*** (0.005)	-0.144*** (0.010)	-0.154*** (0.011)
cancerYes	0.008* (0.005)	0.008 (0.011)	0.007 (0.011)

flushotYes	-0.014*** (0.003)	-0.014** (0.007)	-0.016** (0.007)
raceAmerican Indian	-0.023* (0.012)	-0.023 (0.025)	-0.004 (0.025)
raceAsian	-0.069*** (0.011)	-0.069*** (0.023)	-0.068*** (0.022)
raceBlack	-0.051*** (0.006)	-0.051*** (0.014)	-0.044*** (0.013)
raceHispanic	-0.050*** (0.006)	-0.050*** (0.013)	-0.042*** (0.013)
raceOther	0.009 (0.008)	0.009 (0.017)	0.016 (0.017)
binge_drinkingyes	0.044*** (0.004)	0.044*** (0.009)	0.046*** (0.009)
Constant	2.488*** (0.095)	2.488*** (0.208)	2.498*** (0.209)
Observations	57,999	57,999	57,999
Log Likelihood	-226,655.900		-170,592.000
theta			2.123*** (0.016)
Akaike Inf. Crit.	453,541.900		341,414.000
Note:			*p<0.1; **p<0.05; ***p<0.01