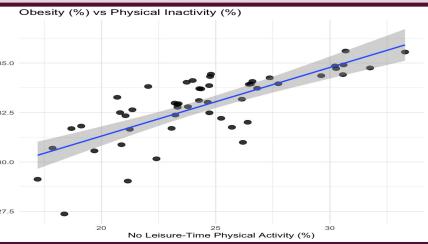


U.S. Department of Health & Human Services

- Original dataset contained:
 - 106,260 obs and 33 col
- Filtered to relevant parameters & dropped NA values
- Missing values have been imputed
- Chosen Parameters: Year, State, Obesity Value, Income

Multiple Linear Regression

What is the association between median income and the mean BMI value across states after controlling for sex and race/ethnicity?



Ridge Regression

How well do the variables age, sex, race/ethnicity, income, and education level predict the percentage of adults who achieve at least 150 minutes a week of moderate-intensity aerobic physical activity across US states?

Beyond the Plate

A Statistical Modeling Approach to Analyzing Sociodemographic Predictors of Obesity in the U.S.

By: Akhila Vemuri, Daniel Tapp, and Veena Seshadri

What factors are correlated with the rise of obesity in America?

Ethics

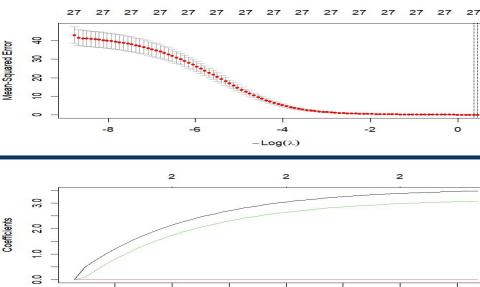
The data being provided contains responses to surveys, meaning that all conclusions are subject to human error. The conclusions drawn from this study use the results from the data as well as prior research

Findings

- Physical inactivity is a key behavioral factor strongly associated with higher obesity prevalence at the state level.
- Lasso regression selected Sex Female, Sex Male, and Education High School Graduate as the strongest predictors
- There is a statistically significant negative relationship between age and fruit/vegetable consumption, meaning that adults are eating less healthy as they get older.
- States with higher physical inactivity tend to be classified as "high obesity."
- The LDA model properly predicts the obesity class based on the consumption of fruits and vegetables
- Higher incomes and low inactivity rates see the lowest probability of obesity while high inactivity and low income have a higher probability. These are only predictive classifications, not determinate.

Logistic/Random Forest Regression

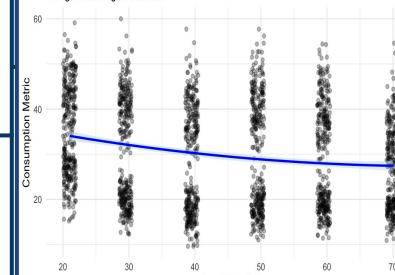
Can we classify whether a respondent is obese ($BMI \geq 30$) or not obese, based on physical activity levels, education, income, and race/ethnicity?



Polynomial Regression Model

How do the healthy eating habits (vegetable and fruit consumption) change with age?

Age vs. Consumption (Polynomial Fit)
Using Robust Age Extraction



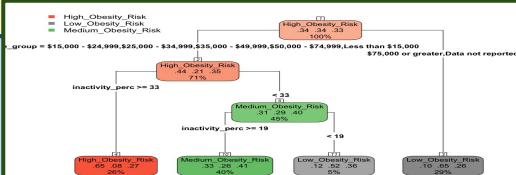
Study Choice: The study data provides long-term analysis of financial factors that contribute to obesity. The parameters chosen describe the non-food related factors that contribute to obesity rates.

Centers for Disease Control and Prevention, Nutrition, Physical Activity, and Obesity – Behavioral Risk Factor Surveillance System, last modified September 16, 2025, <https://catalog.data.gov/dataset/nutrition-physical-activity-and-obesity-behavioral-risk-factor-surveillance-system>

Recommendations
This study would recommend further research regarding an individual's work and life schedule, seeing the impact working house has on the median BMI. It would also be recommended that further research into the pre-existing health conditions of individuals, seeing how this interacts with their median BMI.

Classification Tree

Is there a relationship between an individual's inactivity rate & their income that impacts their BMI?



LDA/QDA

Can the % of adults in a state who engage in healthy behaviors be used to classify whether a state's obesity rate is near the national median?



| | High | Low |
|------|------|-----|
| High | 94 | 13 |
| Low | 15 | 36 |