

Diet Compositions and Obesity

Ken Ye

2023-12-16

Introduction

With an alarming 41.9% of the U.S. population classified as obese, as revealed by the 2011 National Health and Nutrition Examination Survey (NHANES) [1], the implications of this health crisis become undeniably significant. However, the impact becomes even more pronounced when considered globally, where obesity contributed to 8% of deaths in 2017—an alarming surge from 4.5% in 1990 [2].

Obesity, beyond its visual manifestations, is intricately linked to severe health conditions, including heart disease, stroke, type 2 diabetes, and various cancers. These conditions stand as leading causes of preventable, premature deaths. The economic burden is equally staggering, with the estimated annual medical cost of obesity in the United States reaching nearly \$173 billion in 2019 dollars. Notably, medical costs for adults grappling with obesity were \$1,861 higher than those for individuals maintaining a healthy weight [1].

As we delve into the complex web of factors contributing to obesity, the role of food emerges as a central player. Continuous overconsumption, a potential precursor to weight gain and obesity, emphasizes the critical importance of understanding food composition. Certain foods, due to their nutrient content, can significantly influence overall health outcomes.

However, the narrative extends beyond food alone. Economic status emerges as a pivotal consideration, acting as a covariate in my analysis. Death rates attributed to obesity exhibit variations among countries, and a fundamental differentiator lies in their economic standing. Developed countries, buoyed by high incomes, may boast advanced medical systems capable of mitigating obesity-related deaths. Paradoxically, these countries may also grapple with higher obesity prevalence. On the contrary, developing nations, constrained by lower incomes, may face a different set of challenges, potentially emphasizing malnutrition over obesity.

In this analysis, the focal point is discerning the effects of a country's food supply and economic status on death rates attributed to obesity. The implications of this study are far-reaching, offering valuable insights for policymakers eager to curb obesity-related deaths and aiding consumers in crafting diets resilient to the detrimental effects of obesity.

A dual-stage methodology will be employed. Initially, simple linear regression sheds light on the potential impact of individual food groups and a country's economic status. This analysis guides us towards identifying statistically significant associations. Subsequently, the methodology advances to Lasso regression, a more sophisticated model aimed at pinpointing the most influential variables while mitigating the risk of overfitting.

Data

Data Sources

Three separate datasets, totaling 26 predictor variables and 1 response variable, are used in this analysis:

- 1) Daily caloric supply (OWID based on UN FAO & historical sources)

Predictor variables (after renaming): `Miscellaneous`, `Alcohol`, `Animal.fat`, `Vegetable.oils`, `Oilcrops`, `Fish.and.seafood`, `Sugar.crops`, `Sugar.sweeteners`, `Starchy.roots`, `Meat.other`, `Meat.sheep.and.goat`, `Meat.pig`, `Meat.poultry`, `Meat.beef`, `Eggs`, `Milk`, `Nuts`, `Fruit`, `Vegetables`, `Pulses`, `Cereals.other`, `Barley`, `Maize`, `Rice`, `Wheat` - each one (25 total) is a food group available for consumption, continuous, unit (kcal per day per capita)

Source: Our World in Data based on UN FAO & historical sources – processed by Our World in Data

Link: <https://www.fao.org/faostat/en/#data/FBS>

- 2) World Bank income classification

Predictor variable (after renaming): `Income.group` - nominal, four categories (Low-income countries, Lower-middle-income countries, Upper-middle-income countries, High-income countries)

Source: World Bank (2022) – processed by Our World in Data

Link: <https://datahelpdesk.worldbank.org/knowledgebase/articles/906519-world-bank-country-and-lending-groups>

- 3) Share of total deaths that are from all causes attributed to high body-mass index, in both sexes aged age-standardized

Response variable (after renaming): `Death.obesity` - continuous, unit (%)

Source: IHME, Global Burden of Disease Study (2019) – processed by Our World in Data

Link: <https://ourworldindata.org/obesity>

Data Cleaning & Merging

Missing Data Imputation

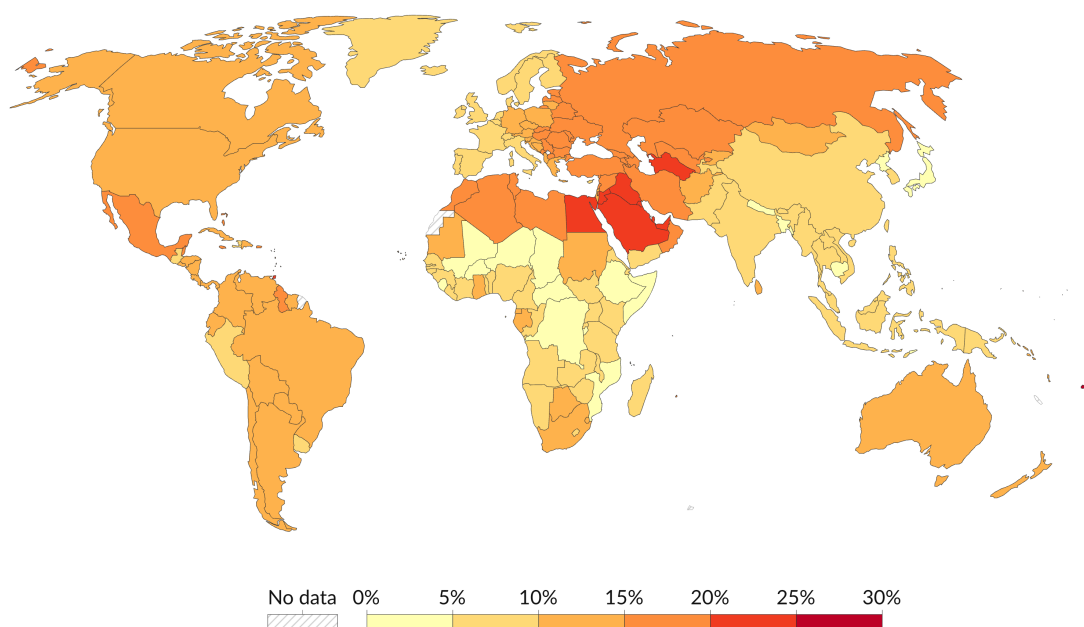
In my analysis, I encountered missing values in certain food groups across countries. Removing these entries would significantly reduce the dataset entries, leaving only 100 out of 174 countries with 45 degrees of freedom for model fitting (if all columns are utilized for the model). This approach would not be ideal for accurately representing global patterns or obtaining robust model estimates.

To address this challenge, I opted for imputation using the “predictive mean matching”(PMM) method from the `mice` package. PMM ensures that imputed values are drawn from observed values with similar predicted values, preserving the distributional characteristics of the data. This method allows us to utilize the full dataset, providing a more comprehensive and representative view of the relationships between food supply patterns and obesity death rates across diverse countries. Imputing missing values with PMM contributes to the validity of our inferences and enhances the generalizability of my findings to a broader global context.

Share of deaths attributed to obesity, 2019



Obesity is defined as having a body-mass index (BMI) equal to or greater than 30. BMI is a person's weight in kilograms divided by their height in meters squared. Shown is the share of total deaths, from any cause, with obesity as an attributed risk factor.

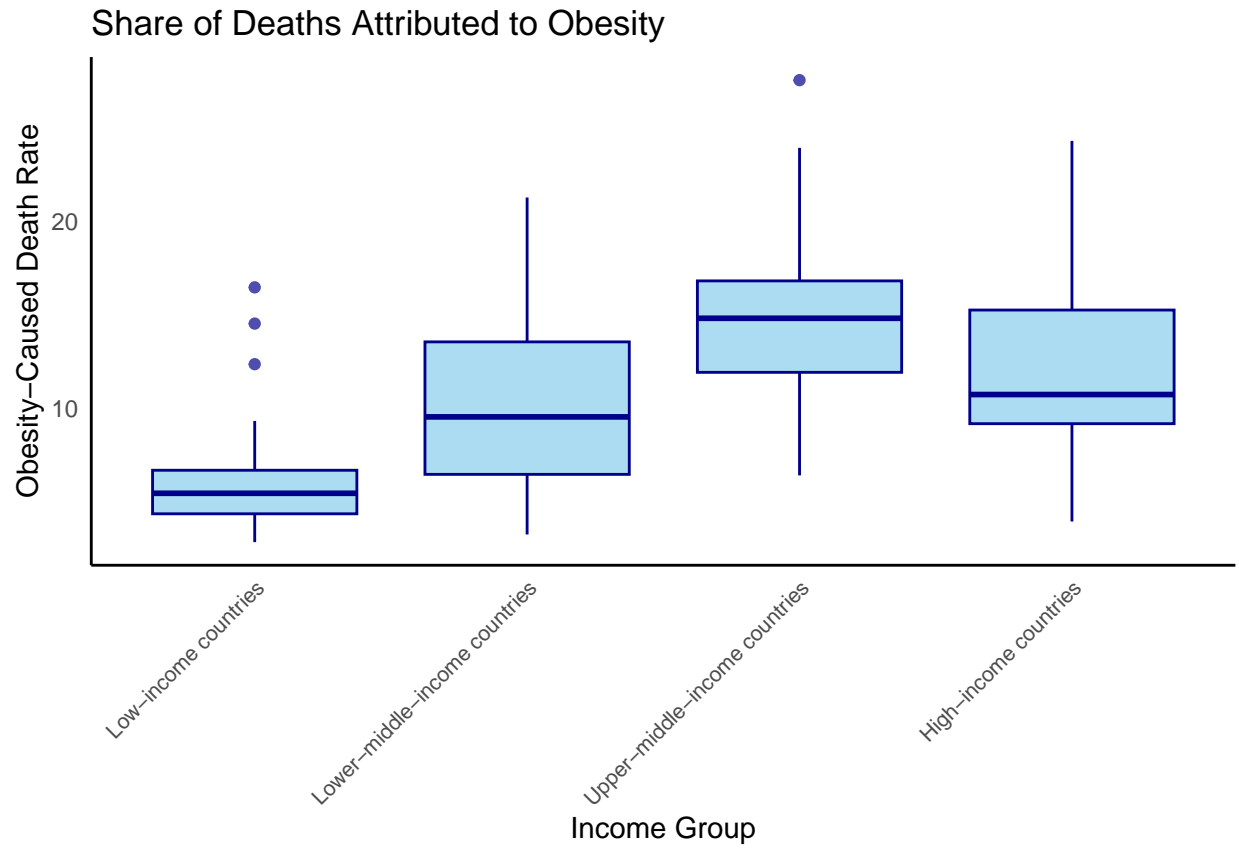


Data source: IHME, Global Burden of Disease (2019)

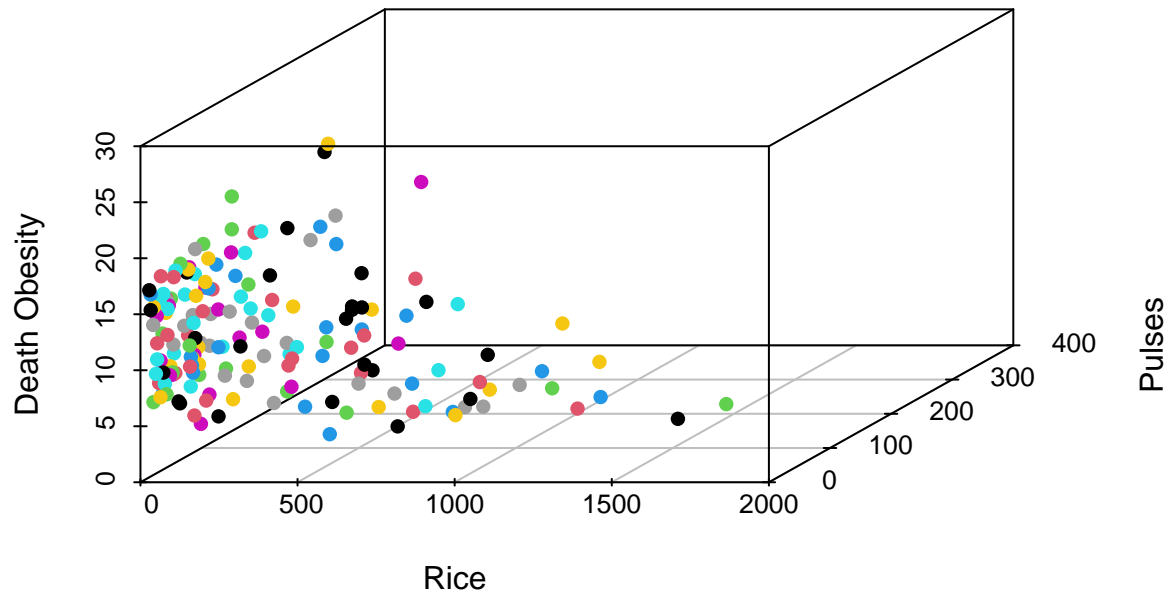
OurWorldInData.org/obesity | CC BY

Figure 1: Share of Deaths Attributed to Obesity Across the Globe

EDA



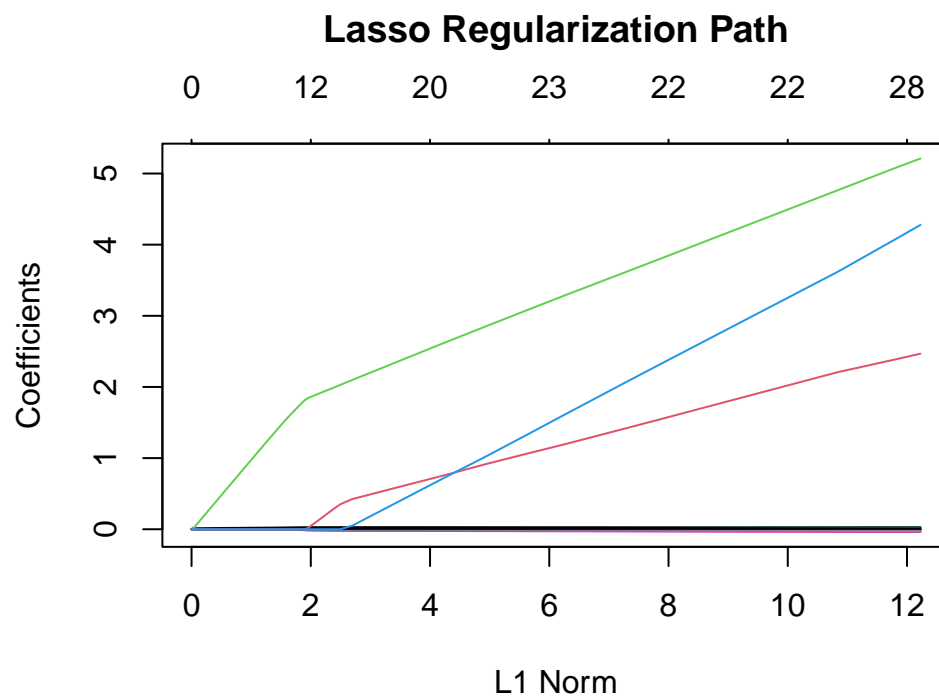
3D Scatter Plot of Rice and Pulses Interaction

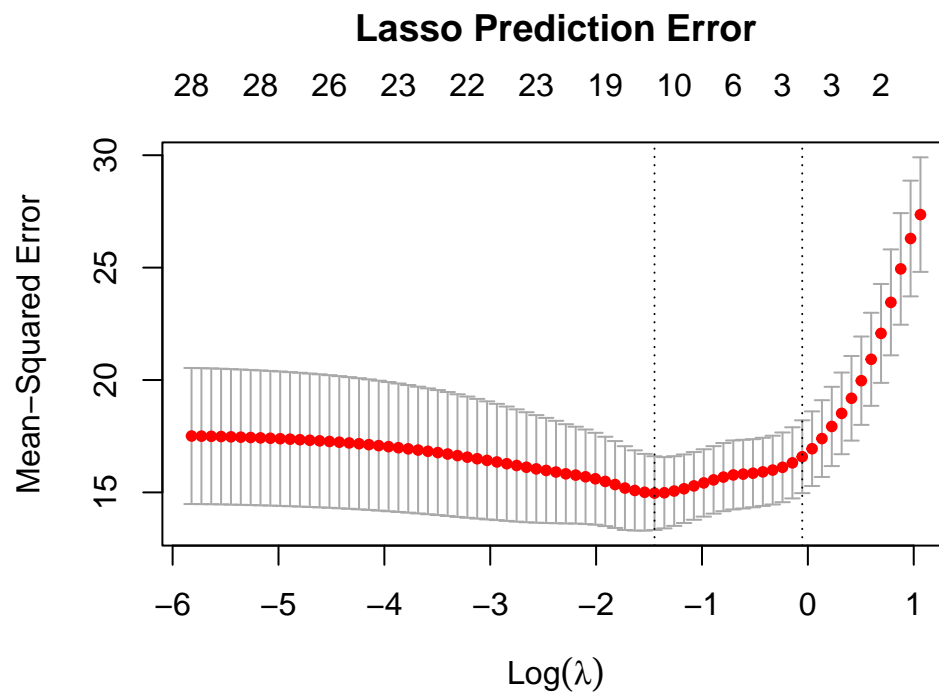


Methodology

Simple Linear Regression Model

Lasso Regression Model





Model Comparison

Final Model

Results

	Coefficient
(Intercept)	5.4800
Oilcrops	0.0080
Fish.and.seafood	-0.0172
Sugar.sweeteners	0.0035
Meat.sheep.and.goat	0.0145
Meat.pig	-0.0032
Meat.poultry	0.0242
Meat.beef	-0.0154
Fruit	0.0004
Pulses	-0.0013
Cereals.other	-0.0021
Barley	0.0002
Wheat	0.0063
Income.groupLower-middle-income countries	0.0279
Income.groupUpper-middle-income countries	1.8858

Conclusion

The USDA Center for Nutrition Policy and Promotion recommends a very simple daily diet intake guideline: 30% grains, 40% vegetables, 10% fruits, and 20% protein. [1] [2]

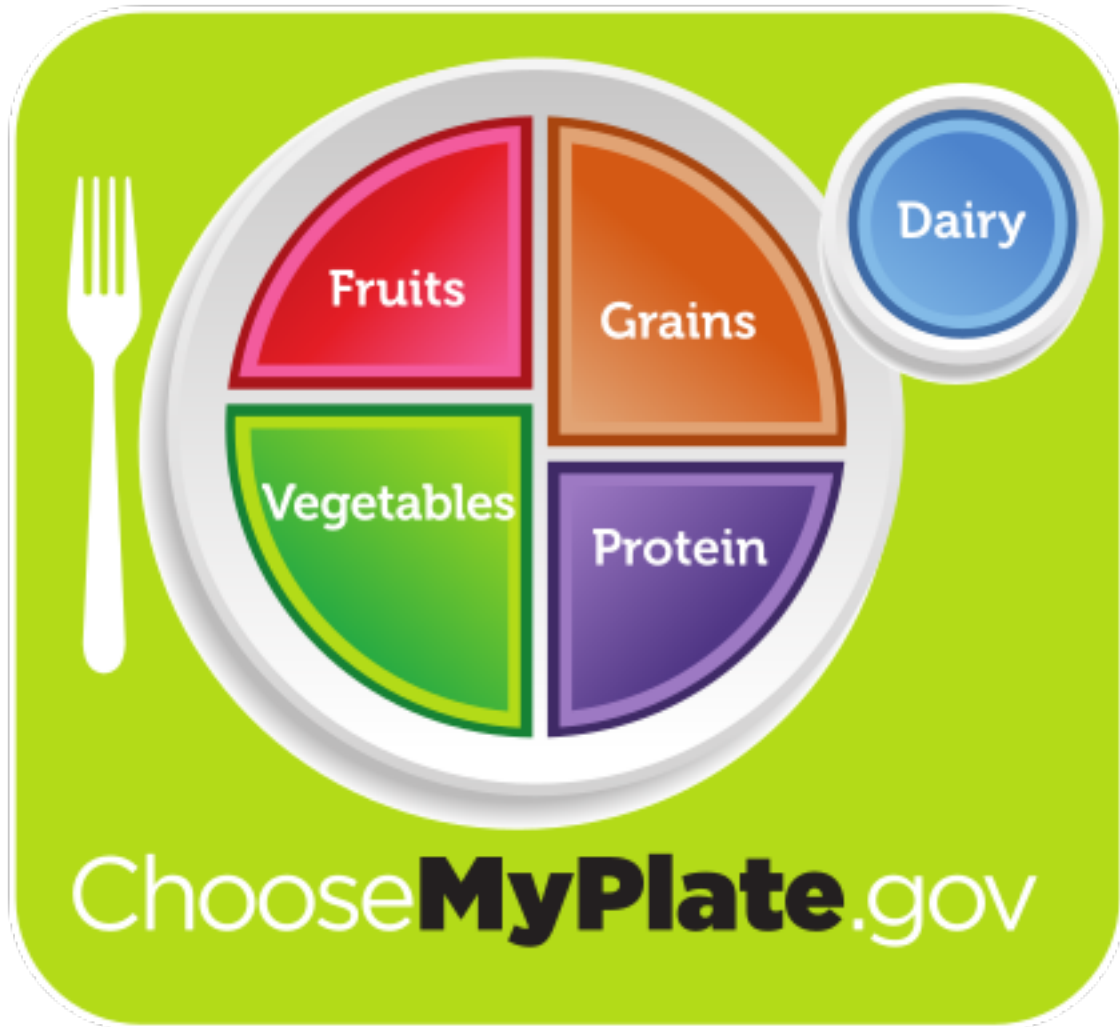
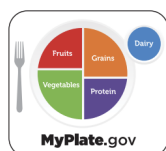


Figure 2: MyPlate Guidelines

Limitations

- 1) supply, not consumption. assuming consumption is proportional to supply.








Start *simple* with **MyPlate Plan**

The benefits of healthy eating add up over time, bite by bite. Small changes matter.
Start Simple with MyPlate.

A healthy eating routine is important at every stage of life and can have positive effects that add up over time. It's important to eat a variety of fruits, vegetables, grains, protein foods, and dairy or fortified soy alternatives. When deciding what to eat or drink, choose options that are full of nutrients. Make every bite count.

Food Group Amounts for 2,400 Calories a Day for Ages 14+ Years

				
2 cups	3 cups	8 ounces	6½ ounces	3 cups
Focus on whole fruits Focus on whole fruits that are fresh, frozen, canned, or dried.	Vary your veggies Choose a variety of colorful fresh, frozen, and canned vegetables—make sure to include dark green, red, and orange choices.	Make half your grains whole grains Find whole-grain foods by reading the Nutrition Facts label and ingredients list.	Vary your protein routine Mix up your protein foods to include seafood; beans, peas, and lentils; unsalted nuts and seeds; soy products; eggs; and lean meats and poultry.	Move to low-fat or fat-free dairy milk or yogurt (or lactose-free dairy or fortified soy versions) Look for ways to include dairy or fortified soy alternatives at meals and snacks throughout the day.



Choose foods and beverages with less added sugars, saturated fat, and sodium.
Limit:

- Added sugars to **less than 60 grams** a day.
- Saturated fat to **less than 27 grams** a day.
- Sodium to **less than 2,300 milligrams** a day.



Be active your way:

Children 6 to 17 years old should move **60 minutes** every day. Adults should be physically active at least **2½ hours** per week.

Figure 3: MyPlate Plan

Citations

- [1] <https://www.cdc.gov/obesity/>
- [2] <https://ourworldindata.org/obesity>
<https://www.dietaryguidelines.gov/>
<https://www.myplate.gov/>