

Hypothesis:
compared to the
upper class.
additionally, we



**ASSIGN
BUSTER**

Hypothesis: Do specific environments influence a person's dietary habits and choices in the United States? Are factors such as socioeconomic status and urbanization correlated with an individual's Body Mass Index (BMI) and obesity? We predict that there will be a higher obesity rate in the lower and middle classes as compared to the upper class. Additionally, we believe that people living in more impoverished areas are more likely to consume cheaper food due to their low income. Moreover, we predict that people who reside in rural areas will have less access to fast food restaurants and processed goods and in turn will likely consume goods that come directly from local farms.

On the contrary, people who live in more urbanized environments are suspected to purchase food that is easily obtainable and more processed which will likely correlate to higher fat concentration and obesity rates. By conducting this experiment, we can better understand the roots of obesity and how certain environments and income contribute to one's BMI.

Introduction: This research would help analyze the effects of dietary intake due to different lifestyles, demographics, and incomes. The understanding of these different factors can lead to a more informed society that reconsiders their eating habits to improve their BMI. Objectively, our findings will help people better understand the factors that influence obesity, which could in turn help decrease obesity rates in America. In a study exploring the calorie density and cost of food it was concluded that there is an "inverse relationship between the energy density (kcal/g) of foods and the energy cost (US\$/1,000 kcal)" (Drewnowski). This data supports the hypothesis that high

calorie and less healthy foods are more accessible to the lower and middle class.

People with a lower income will often choose the high calorie dense diet solely because it is more cost efficient than the recommended diet. Moreover, qualitative reports indicate that rural cultural eating habits tend to have more country and comfort foods which have a higher concentration of fat. Another study done by NHANES found that people in rural areas had higher rates of obesity, despite the other independent factors in the population such as age, education, race/ethnicity, marital status, diet, and physical activity (Befort). "Obesity prevalence significantly differed across rural and urban participants with 39.6% (SE = 1.5) of rural participants being obese compared to 33.4% (SE = 1.1) of urban participants (P = .

006)" (Befort). This refutes our hypothesis of rural areas having lower obesity rates than rural areas. Proposed Methods: To effectively gather and organize our data we will implement the following methodologies: web scraping, data mining, statistical analysis, and data visualization. We would implement web scraping and data mining to obtain a list of the obesity rates per county in every state in the United States.

Information like this is widely available on many public websites; thus, it does not make sense for us to conduct an experiment ourselves to find the obesity rates in each county of the United States. Instead, we would pull and extract the data from the sources listed below using a web scraping algorithm. Figure 1.1 shows the obesity rates of every county in California.

We would implement a crawler to scrape, collect, and input the content in a database to later be analyzed. We are specifically interested in the content shown in the upper right hand corner of Figure 1. 1 which shows the county and its corresponding obesity rate. We would continue to collect the data for every county in the state of California and then repeat this process for the remaining 49 states. Because we are comparing the obesity rates of rural to urban areas, it would be necessary to find the population density of each county in order to categorize each county as either rural or urban. (*Note: this would also be done through web scraping) Figure 1. 2 shows an example of average income per county in the state of California.

We would then use data mining to sort through the counties to determine if a specific county meets the criteria of being either rural or urban based on their population density. Following this we would create a python program that imports the plot. ly library that will allow us to graph the dataset; this would allow for easy data visualization, understanding, and interpretation and would thus allow us to potentially find specific correlations. In order to effectively visualize and interpret the data, we would create 2 histograms (reference Figure 1.

3 and note that Figure 1. 3 does not contain any actual data and is used purely to provide an example of what the histograms distribution could potentially look like). In the first histogram we would graph the distribution of the number of urban counties to the counties obesity rate. The x axis would be the obesity rate (i. e The range values for the bins would be 2% obesity) and the y axis would be the number of counties that fall within a specific obesity rate (i. e.

<https://assignbuster.com/hypothesis-compared-to-the-upper-class-additionally-we/>

the bin). The second histogram would compare the number of rural counties to the counties obesity rate. To maintain reliable data, we would use the same size bins as described in urban histogram.

For the third graph we will create a scatter plot with average income per county on the x axis and average obesity rates per county on the y axis. We would plot the data for both urban and rural areas, differentiating them by color. Based on our findings, we would use statistical analysis to generate a regression for both the urban and rural datasets.

Afterwards, we would determine if the obesity rates between rural and urban areas is statistically significant. Additionally, we would see if any correlation exists between income and obesity rates in urban and rural areas (It is important to note here that if a correlation does exist it will not indicate causation). The first two histograms find the rates of obesity within each county for urban and rural counties respectively. By comparing these two histograms we would be able to determine if there is a difference in obesity rates between rural or urban areas. The scatter plot is used to determine the potential correlation between average income per county and obesity rates per county. Additionally since rural and urban areas are color coded, we can determine if income affects both environments in a similar or different way.

Discussion: After researching different countries obesity rates, average income, and population density, we noticed limitations and obstacles that could arise in both our data and conclusion. Initially it may seem that using BMI as a measure of obesity may be inaccurate, however, after conducting research, it appears that BMI is quite accurate at determining obesity.

There are only a small amount of people who have a BMI indicating obesity but are not actually obese: "When sampling from the general population, over 95% of men and 99% of women identified as obese by BMI were also obese via body fat levels" (Medical News Today). One other limitation in our study is that we categorized counties as either rural or urban rather than categorizing cities or towns. This is a limitation because certain cities or towns of a county may be urban while another city of the same county could be rural. Additionally, counties do not have the same populations and boundaries; many counties are biased due to gerrymandering which can cause our data to contain inaccuracies underlying the research. Perhaps instead we should have selected tighter regions to analyze such as an area's zip code or city/town rather than a large county. After conducting our research, we found that one of our initial predictions made in our hypothesis was incorrect; we discovered that rural areas were in fact more obese than urban areas. This could in part be due to the shift made during the industrialization era.

Prior to the industrial revolution, manual labor accounted for a majority of farming. However, in today's society, machines do a large portion of the labor: "Increased mechanization of rural occupations has reduced these levels of caloric expenditure, which may impact the younger working adults the most" (Befort). Following our research, if we as a society can come to a conclusion of the factors correlated with obesity, we can perhaps alter the public image.

By changing public opinion on the causes of obesity, obesity can be seen as a curable condition rather than a permanent disease. People may be more

<https://assignbuster.com/hypothesis-compared-to-the-upper-class-additionally-we/>

conscious about what they eat, how much they eat, and how often they exercise. Inversely, this study can promote a more understanding society when it comes to the judgement of obese people. By determining the environmental factors that influence obesity, society won't blame the obese individually entirely and will instead look more closely at our society which promotes obese habits and poor dietary choices.