Is the U.S. Adult Obesity Rate Affected by Median Income?

Members: Ishita Deshmukh, Tanvi Vijay, Kotra Sree Venkata Naga Sreecharan,
Sarat Chandra Ventrapragada, Claudine Koyama, Pooja Gotur Sathyanarayana

Question and Hypothesis:

Over the years, the obesity rate in the United States has been constantly increasing. There is a growing concern over this increasing rate as obesity tends to have many socio-economic impacts. It also tends to be the root factor behind various other illnesses like diabetes, hypertension and heart diseases. Given obesity's high significance on health and economic costs, it is a clear global health priority, and hence it is vital to understand the many factors that affect obesity to advance coordinated public health policy initiatives to combat this epidemic.

The topic that is addressed in this project is "If the Adult Obesity Rate in the United States is Affected by the Median Income of the Population". Higher income and economic growth are known to be correlated with higher obesity rates in *developing* countries [1]. We seek to find the associations between adult obesity and median income in the United States, being a *developed* country with high economic growth, to see if the same is true, that higher income leads to higher obesity rates or that the reverse is true, and the factors that contribute to that relationship if present in the United States, by state, for the years 2012 to 2021.

There are various other socio-economic factors that influence the rate of obesity. To find the relation, if it exists, between median income and obesity rates we need to control for the other factors which may have an impact on obesity. Some studies have shown a relationship between obesity prevalence and socioeconomic status measured as educational level or income [2,3]. For our research, we have taken the education rate, race (percentage of White or Black) and the minimum wage of a state into account.

Our **Null Hypothesis (Ho)** is that there is no relationship between adult obesity and median income in the United States. Our **Alternate Hypothesis (Ha)** is that there is a relationship between adult obesity and median income in the United States.

Why Data-Driven Approach?

In 2001, obesity rates were higher than ever, with 61% of adults nationwide overweight or obese. In the intervening years, several policies were administered to deal with the problem, yet the prevalence of overweight and obesity rose further, to 68%.[2]

In low-income countries, overweight and obesity is usually more prevalent among higher socioeconomic position groups. For example, in Asian countries like India and Bangladesh, over the years, obesity has shown to have a positive correlation with income levels [3]. On the other hand, in many middle-income countries, like South Africa, obesity tends to decrease with increase in income [4].

Since it is observed that the relation between obesity rates and income vary country to country, it is necessary to follow a data driven approach to find an accurate relationship that relates to the population of the United States. Using a data-driven approach will give us a comprehensive understanding of the different factors that actually influence obesity . It will help us in assessing the amount of influence each factor has and how they vary in different states of the US.

Ideal Experiment:

The ideal experiment would involve examining the relationship between obesity and income where external factors are controlled for. For the ideal experiment, we shall consider a sample of people belonging to the same age group, same state, same income, education qualifications, possessing similar genetic characteristics and eating habits so that the other constraints that influence the obesity rate are controlled. The sample group is then divided randomly into 2 groups where income is increased for one group and retains the same for another group. The impact of the income changes on obesity can then be studied over the years for both groups simultaneously. An ideal model would have no bias, endogeneity, heteroskedasticity, and multicollinearity. But the ideal experiment is not practical and hence a multiple linear regression appears to be the best method to control external variables as well as examine the relationship between obesity and income and whether this relationship is significant.

Methodology:

1) Data Collection:
   a) The data for adult obesity rate is taken from the CDC [link]. It is extracted manually for each year and for each state.
   b) The data for median income is taken from FRED [link] and was web scraped by invoking a selenium safari driver.
   c) The data for minimum wage is extracted manually from the DoL [link] website.
   d) The data for race and education is extracted manually from Census Bureau [link].
   e) Year-on-year data for CPI is downloaded from the census bureau [link].
2) Data Cleaning & Structuring:
   The standard format of data used everywhere is year, state, factor(s).
   a) The data for minimum wage is adjusted to 2021 figures based on CPI data.
   b) The data for race and education is cleaned manually and then structured into the standard format we followed.
   c) Then all the data is joined using state and year as common join variables.
3) Model Building:

Built a multivariate linear regression model using the above data with year, white %, black or african-american %, minimum wage, median income as independent variables and obesity rate as dependent variable.

Result:

All the variables have statistical significance to the relationship with the *adult obesity rate*, with a *p-value of 0.00*.

The negative coefficients from the regression: *percentage of educated (-0.11), median income (-0.10), and minimum wage (-0.73)* demonstrate a negative relationship with adult obesity rates. Higher education levels, higher median incomes, and *higher minimum wage levels* are associated *with lower adult obesity rates*. The positive coefficients from the regression: *race-white (0.08)*, *race-black (0.13)* demonstrate a positive relationship with adult obesity rates. The adult population of race being black has a higher association with obesity compared to white. The R-square and adj R-square values of the model stand at 0.617 and 0.612 respectively.

Using Tableau time-lapse heat maps to visualize the comparison of the *adult obesity rates* and *median income* of the different states in the United States, we observe that overall, the *adult obesity rates increase over time* from 2012 to 2021. We also see that states with *lower median income levels experience higher adult obesity rates in each respective year*.

Conclusion:

Our regression results lead us to *reject the null hypothesis,* and conclude that the alternative hypothesis is true which indicates that there is a significant relationship between adult obesity rate and median income in the United States.

The results show that unlike *developing countries* where *higher income* is associated with *higher obesity*, the United States as a *developed country*, experiences the *reverse effect*, where *lower median income* and *minimum wages levels* are associated with *higher adult obesity rates*.

As a follow up project, it would be beneficial to conduct research on the individual states within the United States to investigate other factors that may contribute to lower income adults within each state experiencing higher obesity rates. This will help policymakers and community leaders be better informed when deciding what steps can be taken to address this issue.

References

1. Sobal J, Stunkard AJ. Socioeconomic status and obesity: A review of the literature. Psychol Bull 105:260–75. 1989.
2. McLaren L. Socioeconomic status and obesity. Epidemiol Rev 29:29–48. 2007.
3. Novak NL, Brownell KD. Role of policy and government in the obesity epidemic. Circulation.2012 Nov 6;126(19):2345-52. Doi: 10.1161/CIRCULATIONAHA.111.037929. PMID: 23129701.
4. Luhar S, Mallinson PAC, Clarke L, Kinra S. Trends in the socioeconomic patterning of overweight/obesity in India: a repeated cross-sectional study using nationally representative data. BMJ Open. 2018 Oct 21;8(10):e023935. doi: 10.1136/bmjopen-2018-023935. PMID: 30344181; PMCID: PMC6196932.
5. Dinsa GD, Goryakin Y, Fumagalli E, Suhrcke M. Obesity and socioeconomic status in developing countries: a systematic review. Obes Rev. 2012 Nov;13(11):1067-79. doi: 10.1111/j.1467-789X.2012.01017.x. Epub 2012 Jul 5. PMID: 22764734; PMCID: PMC3798095.