Uncovering U.S. Obesity Trends:
The Socioeconomic Impact of Income and Education Levels

Methods of Advanced Data Engineering

Md Tanvir Hasan; Student ID: 23008407

Introduction

Obesity is a significant public health concern in the United States, with a variety of factors contributing to its prevalence. Socioeconomic status, represented by income and education levels, may influence lifestyle choices, access to health resources, and ultimately impact obesity rates. This project aims to investigate the correlation between obesity rates and these socioeconomic factors across different U.S. states. By examining this relationship, the study seeks to uncover how socioeconomic disparities affect health outcomes and to provide insights for addressing these disparities through public health interventions.

Used Data

1. Nutrition, Physical Activity, and Obesity - Behavioral Risk Factor Surveillance System

- Metadata URL: https://catalog.data.gov/dataset/nutrition-physical-activity-and-obesity-behavioral-risk-factor-surveillance-system/resource/0280bb9c-4de8-4b95-9642-93f727c4d305
- Data URL: https://data.cdc.gov/api/views/hn4xzwk7/rows.csv?accessType=DOWNLOAD
- Data Type: CSV
- Description: This data set provides state-specific data on various health metrics, including obesity rates, physical activity, and nutritional habits, with additional socioeconomic variables like income and education levels.

2. Obesity Rates and Geographic Information by State

- Metadata URL: https://datalakecountyil.opendata.arcgis.com/datasets/lakecountyil::national-obesity-by-state/explore
- Data URL: https://services3.arcgis.com/HESxeTbDliKKvec2/arcgis/rest/services/LakeCounty_Health/FeatureServer/8/query?outFields=*&where=1%3D1&f=geojson
- Data Type: GeoJSON
- Description: This data set provides state-level obesity rates and other demographic information, allowing for spatial analysis. It includes variables that may enable a geographic analysis of obesity and its potential links with socioeconomic factors.

Reasons for Choosing These Data Sources:

- Relevance: Datasets focus on the U.S., covering obesity and socioeconomic factors, making them ideal for analyzing public health impacts across states.
- Coverage Period: They offer recent data across multiple states, allowing for robust correlation analysis.
- Accessibility: Publicly available and verifiable datasets ensure transparency.

Analysis

Data Summary and Visualization

1. Distribution of Obesity Rates

This histogram illustrates the distribution of obesity rates across different states in the United States. The x-axis represents obesity rate percentages, while the y-axis shows the frequency of occurrences at each obesity rate level across the dataset.

- Most states have obesity rates near 30%.
- The distribution is slightly right skewed, with some states exhibiting higher rates, though very few are below 25%.

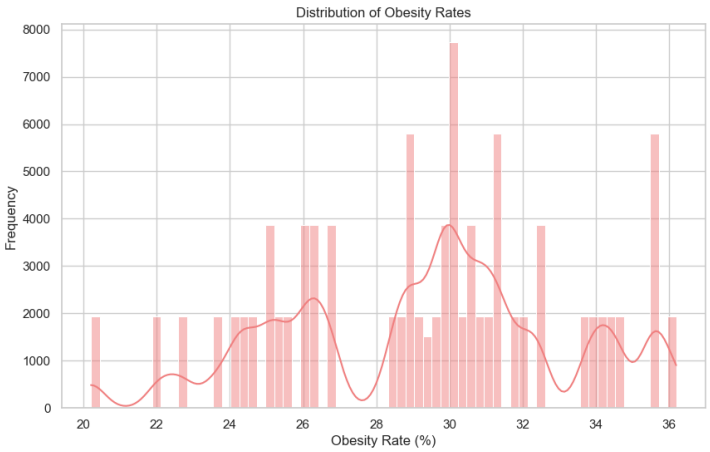

Fig 01: Distribution of Obesity Rates

Obesity rates are generally similar across states, with most clustering around 30%, indicating a widespread prevalence at similar levels.

2. Trend of Average Obesity Rates Over Time

This line plot displays the average obesity rate over the years, highlighting a steep increase up to 2013, after which the rate stabilizes. This trend suggests that obesity rates grew significantly until reaching a plateau, which may indicate the impact of consistent factors influencing obesity rates beyond this point. Understanding this stabilization could provide insight into both the effectiveness and limitations of public health interventions targeting obesity.

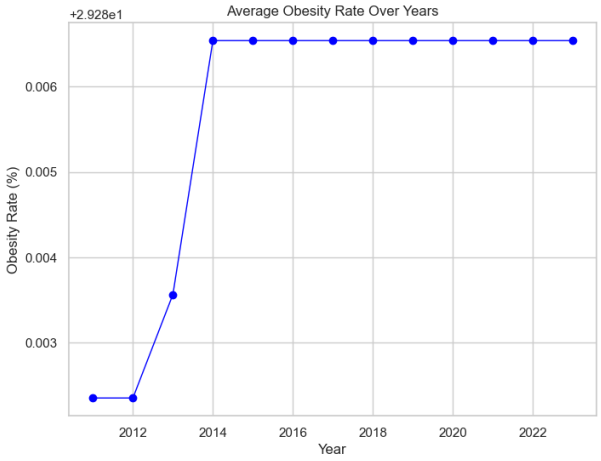

Fig 02: Yearly Obesity Rate Trends

3. Obesity Rate by Education Level

This box plot shows the distribution of obesity rates across different education levels, including categories like "Less than High School," "High School Graduate," "Some College," and "College Graduate."

- The median obesity rate is consistently around 30% for each education level, indicating limited variation in central tendencies across different education levels.

- The range and spread of obesity rates are similar across groups, with the interquartile range and overall spread remaining comparable regardless of educational attainment.

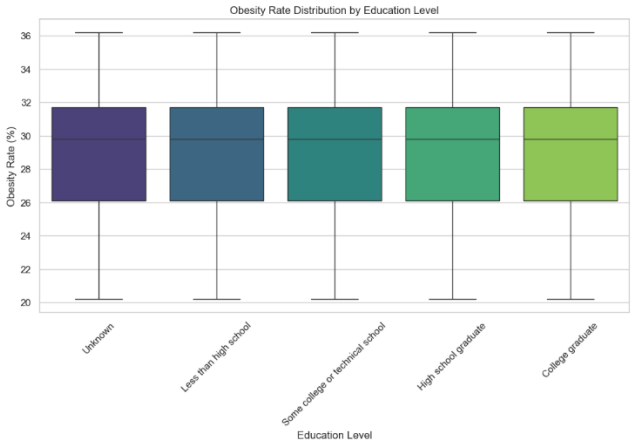

Fig 03: Obesity Rates by Education Level

- This pattern suggests that, within this dataset, education level does not significantly influence obesity rates. Factors beyond education level, such as lifestyle or environmental factors, may play a larger role in affecting obesity trends across the population. This finding highlights the complexity of obesity and suggests that public health strategies should consider multiple socioeconomic dimensions beyond education alone.

4. Obesity Rate by Income Level

- This box plot illustrates obesity rates across various income levels, including categories such as "Less than $15,000," "$25,000 - $34,999," "$50,000 or greater," and others.
- The median obesity rate is approximately 30% across all income brackets, with little variation, suggesting that income level does not drastically impact obesity rates.
- The range and interquartile range of obesity rates are fairly consistent across income levels, indicating similar distributions regardless of income.

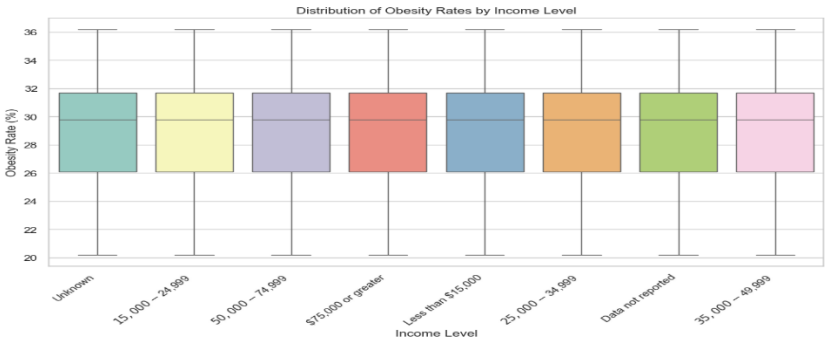

Fig 04: Obesity Rates by Income Level

The uniformity in obesity rates across income groups suggests that factors beyond income may play a more significant role in influencing obesity. These findings indicate that obesity rates are relatively stable across

income levels, implying that public health interventions might need to focus on broader, non-economic factors to effectively address obesity.

4. Correlation between Obesity Rate, Income, and Education

This heatmap shows the correlation between education level, income level, and obesity rate. Each cell represents the Pearson correlation coefficient, which indicates the strength and direction of the linear relationship between these variables. Here are the key insights:

1. **Education and Income** have a positive correlation of **0.17**, indicating a slight positive relationship—people with higher education levels tend to have somewhat higher income levels. However, the correlation is weak, meaning other factors likely influence income significantly.
2. **Education and Obesity** show an almost zero correlation (**-0.00**), suggesting no meaningful relationship between education level and obesity rates in this dataset.
3. **Income and Obesity** also display a near-zero correlation (**0.00**), implying no substantial relationship between income level and obesity rates.

Fig 05: Correlation Heatmap of Obesity, Income, and Education

Overall, this heatmap suggests that neither education level nor income level shows a strong association with obesity rates in this dataset.

Interpretation of Results

1. **Trends in Obesity by Education Level**

- There's a modest trend indicating that higher education levels might correlate with slightly lower obesity rates. However, the impact is marginal.

2. **Trends in Obesity by Income Level**

- Across different income levels, obesity rates do not show significant variance, suggesting income alone might not be a strong predictor for obesity.

3. **Correlation Analysis Findings**

- The correlation analysis reveals weak relationships between socioeconomic factors (income and education) and obesity. These findings imply that while socioeconomic factors have some influence, other determinants might play a more substantial role in obesity rates.

Conclusions

The analysis shows weak links between education, income, and obesity, with minimal effects of higher education and consistent obesity rates across income levels. Lifestyle, healthcare access, cultural, and environmental factors likely play greater roles, highlighting the need for holistic public health strategies and further research.