

DATA NUTRITION LABEL

SAM, MATTHEW, MATTHEW

ABOUT

This dataset contains the obesity percentage of each state, as well as the average walkability index of each block census group in every state, averaged across the state. We plan to establish a correlation between the percentage of obese individuals and the average walkability index.

BACKGROUND

Obesity: 2015 CDC Behavioral Risk Factor Surveillance System Survey
Walkability: Information from each block group in the 2019 Census.

Four variables are weighted in a formula to create the walkability index. The block group's intersection density, proximity to transit stops, employment mix, and the mix of employment and occupied housing. This creates a walkability index ranked 1-20, with larger values indicating a more walkable area.

Data Creation Range

2015-2019

Created By

Obesity: Lake County Illinois
Walkability: U.S. Environmental Protection Agency

Source

<https://catalog.data.gov/dataset/walkability-index1>
<https://catalog.data.gov/dataset/national-obesity-by-state-d765a>

Use Cases

- Public Health Interventions
Target states with high obesity rates, emphasizing physical activity and promoting activity
- Urban Planning
Implement infrastructure improvements in necessary places, including sidewalks, parks, and pedestrian-friendly zones
- Policy Development
Advocate for transportation policies that prioritize pedestrian-friendly infrastructure

Alert Count	6
Completeness	1
Missing States	1
Provenance	0
Collection	3
Socioeconomic Bias	1
Volunteer Bias	1
Response Bias	1
Description	0
Composition	2
Racial Bias	1
Inaccurate Prediction	1

Missing States

The state of Rhode Island is missing from the dataset, which may skew results
Mitigation: N/A

Socioeconomic Bias

Dataset may have bias present in the form of economic divide, with those of lesser socioeconomic status more susceptible to obesity
Mitigation: Acknowledge this is a source of potential bias

Response Bias

Since it was a survey, those ashamed of their weight may not answer truthfully
Mitigation: Acknowledge the limitations of surveys as data sources

Volunteer Bias

Since it was a survey, some may choose not to respond or may not be able to due to literacy, economic, or linguistic barriers. This makes the set not representative of the complete population.
Mitigation: Acknowledge the limitations of surveys as data sources

Racial Bias

Due to existing segregation in the US, racial demographics may be correlated with walkability
Mitigation: Acknowledge the confounding factors, and do not draw direct conclusions

Inaccurate Prediction

Since both datasets are grouped by state, it is hard to predict and draw conclusions over such a vast area.
Mitigation: Use more precise data points, such as zip codes or cities