**Finding patterns in food access data**

*Proposed by Louise McMillan, 6/Jul/2023*

This project is about finding patterns in publicly available data from the Food Access Research Atlas in the USA:

<https://www.ers.usda.gov/data-products/food-access-research-atlas/download-the-data/>

<https://www.ers.usda.gov/data-products/food-access-research-atlas/documentation/>

The data list the number of people in each census tract (small regions in the United States of America) have low access to supermarkets, supercentres and large grocery stores, along with other details such as vehicle access that can affect people’s ability to obtain healthy food. The first link lists the latest data, from 2019, and also gives access to the 2015 and 2010 datasets. The second link includes definitions of the fields in the dataset.

The United States Census Bureau also releases data on population sizes in each census tract.

<https://data.census.gov/>

These data can be combined, along with any other relevant datasets with information about different US states, to look at what factors might be linked to the proportion of local population with low access to healthy food.

Students will need to

* Integrate multiple large datasets, some fine-grained and some wide-scale
* Correctly interpret variables with complex definitions
* Carry out an exploratory analysis, choosing a subset of variables and features to report
* Determine which datasets need to be integrated to answer specific questions and/or make specific predictions
* Carry out the analyses required to answer specific questions, and display the chosen model(s) and results
* Consider any ethical aspects when reporting the results