

Final Project: Assignment 01

Olivia Gomez

Part 1: Research Question

Have restrictive immigration laws and a rise in anti-immigration rhetoric across the country contributed to poorer health outcomes for immigrant populations?

Part 2: Data

Data Set 1: National Health Interview Survey

[*Link*](#)

This data set is created from a survey conducted each year to capture statistical information on the amount, distribution, and effects of illness and disability in the United States. There is a clear data dictionary, called codebook on the website. The units of measurement are individual adults and children. There are many dimensions of the data, including demographics, socioeconomic factors, and health behaviors and conditions. Missingness in data is unavoidable as this is a survey, but imputation was used to mitigate impacts and missing variables are coded. Limitations of the data include self-reporting bias as well as the fact it cannot be used to measure changes over time.

Data Set 2: American Community Survey

[*Link*](#)

This data set provides demographic details for Americans every year. There is a data dictionary provided through the IPUMS website. The unit of analysis is an individual adult. Dimensions of the data include economic status, migration information, demographics, health insurance status, and many more. Missingness in data is unavoidable as this is a survey, but missing variables are coded to easily remove from evaluation. Limitations of the data include self-reporting bias and non-response/selection bias as this is a voluntary survey.

Data Set 3: Annual National Election Studies

[*Link*](#)

The ANES time series studies conduct interviews during presidential election years to understand voting behavior, public opinion, and attitudes. There is a codebook provided. The unit of analysis is an individual adult. Dimensions of the data include feelings towards different political issues such as immigration, demographic characteristics, and political identifications. Missingness in data is unavoidable as this is a survey, but missing variables are coded to easily remove from evaluation. Limitations of the data include issues with self-reporting bias.