**Purpose:** To assign 2008 Public Use Micro-Data (PUMS) / Mortality Disparities in American Communities (MDAC) observations to International Energy Conservation Code (IECC) / U.S. Department of Energy (DOE) climate zone designations used in the Residential Energy Consumption Survey (RECS).

**Background:** PUMS/ MDAC respondents are assigned to Public Use Micro Data Sample Areas (PUMAs), or “special non-overlapping areas that partition a state” (2008 ACS 1-Year PUM Readme, 1). PUMAs are census bureau statistical geographies that contain around 100,000 people. They are recalculated with each decennial census.

IECC climate zones are assigned to 2010 county boundaries using a combination of heating degree days, average temperatures, and precipitation (Guide To Determining Climate Regions By County (2015), 1). Climate zones are simplified to six categories:

1. Hot-Humid
2. Hot-Dry / Mixed Dry
3. Mixed-Humid
4. Marine
5. Cold / Very Cold
6. Subarctic

Each climate zone fits neatly within a county boundary.

Although census tracts can be aggregated to PUMAs and PUMAs nest within states, they cannot be perfectly related to county boundaries. At times, they straddle both county geography and climate zone. Additionally, because PUMAs are recalculated over time, we need to relate 2010 PUMA boundaries to 2000 PUMA boundaries before we can assign PUMS / MDAC respondents to a climate zone.

**Methodology:** PUMS / MDAC assignment to climate region required three data sources:

1. IECC / DOE county climate zone designations[[1]](#footnote-1)
2. 2010 PUMA shape files available from United States Census Bureau / R Tidy Census[[2]](#footnote-2)
3. 2000 – 2010 PUMA relational file from IPUMS USA[[3]](#footnote-3)

To compare counties to PUMAs, we determined the county geography that covered the largest area of a given PUMA using great circle distance and assumed that a PUMS / MDAC respondent was most likely to fall into this county’s corresponding IECC climate zone. We performed all calculations using R’s SF package, intersecting with sf\_intersect, and calculating area with sf\_area.

To relate 2010 PUMAs to 2000 PUMAs, we used the IPUMS USA Consistent Public Use Microdata Areas (ConsPUMAs) approach. ConsPUMAs minimize the difference in populations between PUMA vintages until the population mismatch is smaller than 1 percent. Because ConsPUMAs relate multiple 2000 PUMAs to single 2010 PUMAs, we applied a simple rule and selected the first instance of a 2000 to 2010 relationship.

**Resources:**

1. Guide To Determining Climate Regions By County, (2015). Pacific Northwest National Laboratory: <https://www.energy.gov/sites/prod/files/2015/10/f27/ba_climate_region_guide_7.3.pdf>
2. American Community Survey 2016 – 2020 PUMS File Read Me (2022). American Community Survey Office, US Census Bureau: <https://www2.census.gov/programs-surveys/acs/tech_docs/pums/ACS2016_2020_PUMS_README.pdf>
3. 2000 - 2010 PUMA Crosswalk. IPUMS USA: <https://usa.ipums.org/usa/volii/pumas10.shtml#crosswalk>
4. R SF Documentation <https://cran.r-project.org/web/packages/sf/sf.pdf>

**Appendix**

**Figure 1: Census Geographical Hierarchy**

**A diagram of states and regions

Description automatically generated**

**Figure 2: IECC Climate Zones, 2010 County Boundaries**

A map of the united states

Description automatically generated

**Figure 3: Climate Zones, 2010 PUMA Boundaries**

A map of the united states

Description automatically generated

1. Dataset pulled down from github: <https://gist.github.com/philngo/d3e251040569dba67942>. Added two missing counties – Shannon, North Dakota, Wade Hampton – Alaska. [↑](#footnote-ref-1)
2. Projected CRS boundaries used were Contiguous Albers Equal Area Conic [↑](#footnote-ref-2)
3. See methodological resources here: <https://usa.ipums.org/usa-action/variables/CPUMA0010#description_section> [↑](#footnote-ref-3)