Understanding the Intersection of Climate Risk, Food System Resilience, and Racial Equity: A County-Level Data Tool

Technical Appendix

Sonia Torres Rodríguez, Noah Johnson, Kassandra Martinchek, Judah Axelrod, Amy Rogin, Elaine Waxman

Last updated October 2023

This technical appendix describes the data sources and specifications, methodology, and limitations for the underlying data for "Understanding the Intersection of Climate Risk, Food System Resilience, and Racial Equity: A County-Level Data Tool," available at https://urban-institute.shinyapps.io/food-systems-clustering/.

The tool enables communities to assess key aspects of food system resilience and empowers changemakers to invest in places that face outsized challenges to resilience, with a particular focus on communities of color. It is a preliminary tool that brings together diverse datasets on climate hazards, agricultural production, food security and access, and racial equity—and we intend to build on this data tool in future work using researcher and community feedback. The tool includes the following:

- A "national map" tab, which shows areas with elevated climate hazards and food systems challenges. These maps show the intersection of climate hazards that have strong impacts on local food systems, including wildfires, droughts, and floods, food security, and agricultural production. This tab is useful for exploring where climate and food systems risks intersect and which counties face cumulative disadvantages in food system resilience.

- A "county-level data" tab, which shows data on food security and access, demographics, agricultural production, and climate hazards at the county level. Users can toggle between raw values and their county's percentile rank relative to other counties. This tab is useful for exploring county-level data on factors that shape food system resilience to climate hazards and seeing how your county compares with all US counties.

- A "race-disaggregated data" tab, which shows county-level data on factors that shape food system resilience, disaggregated by race. This tab includes data for differences in demographics, social ties, home values, homeownership, employment, poverty, agricultural production, and income by race and ethnicity. It is useful for exploring differences within counties between populations to illustrate the landscape of varied risks by race and ethnicity.

The tool's data are available for download from the Urban data catalog, and the code is available for review in the public GitHub repository, available at https://github.com/UI-Research/food-systems-resilience-and-equity.

More details on the value and applications of the pilot data tool can be found in the following publication:

- Growing Climate Hazards Pose Risks to Food System Resiliency Nationwide. Communities of Color Stand to Bear the Brunt.

The rest of this technical appendix is organized in three sections. First, we describe the data sources and their use in the tool. Second, we describe each variable and the process for their construction. Third, we briefly summarize data and methodological limitations.

Data Sources

This section discusses the data sources used to construct the measures in "Understanding the Intersection of Climate Risk, Food System Resilience, and Racial Equity: A County-Level Data Tool." For each data source in this section, we provide details about the original dataset, including its authoring organization, its purpose, and how we used the data in the tool. The data sources covered include the following:

- Map the Meal Gap

- Food Environment Atlas

- Census of Agriculture

- American Community Survey

- Spatial Hazard Events and Losses Database for the United States (SHELDUS)

- Centers for Disease Control and Prevention (CDC) National Public Health Tracking Network Data

- Urban Institute Mobility Metrics

- Census Bureau American Indian, Alaskan Native, and Native Hawaiian Areas

The data source cleaning can be viewed in the 01_clean-data.R in the Github repository (https://github.com/UI-Research/food-systems-resilience-and-equity).

Map the Meal Gap

About the data source: Map the Meal Gap (MMG) provides estimates of food insecurity and cost per meal at the county level with the goal of better understanding variation in food access and insecurity at the county level. The data source is produced by Feeding America. MMG addresses a data gap in local-level data on food security at lower geographies than the state level, which are not available through the Current Population Survey. According to MMG's technical documentation (available at https://www.feedingamerica.org/research/map-the-meal-gap/how-we-got-the-map-data?_ga=2.170297189.434004663.1697722312-1394693673.1697722312), the dataset is generated using Current Population Survey data to assess the relationship between food insecurity and determinants of food insecurity at the state level. We selected this dataset because of its inclusion of estimates at multiple levels from county to national and includes some racial disaggregation. MMG estimates have been published yearly (since 2017) and the most recent release is from 2021. Upon request, researchers can obtain racially disaggregated estimates of food insecurity rates, although some counties do not have race-disaggregated estimates due to small sample sizes.

Use of these data in this tool: We downloaded county-level data from 2020 (the most recent release when we conducted our analysis). We use these data to measure food insecurity (both for the overall county population and for combined Black and Latinx households) and the average cost per meal at the county level.

Food Environment Atlas

About the data source: The Food Environment Atlas (FEA) contains detailed data on food environment factors like grocery and retail store proximity, food prices, expenditures on fast foods, food taxes, dietary data, and income and poverty maintained by the Economic Research Service in the United States Department of Agriculture (USDA). According to the FEA documentation (available at https://www.ers.usda.gov/data-products/food-environment-atlas/documentation/), the FEA dataset on retail store proximity is created by comparing a directory of supermarkets to population-level data in the Census of Population and Housing and the American Community Survey. The FEA was first released in 2011, and yearly or biyearly data were released until 2017.

Use of these data in this tool: We use 2015 estimates from FEA in this tool, which are the most recently available data from the 2017 data release.

Census of Agriculture

About the data source: The Census of Agriculture (COA) is a census of US farms and ranches and their operations and is conducted by the National Agricultural Statistics Service at the US Department of Agriculture. The census includes any farms with agricultural production valued at $1,000 or more during the defined census year. The COA is conducted once every 5 years and captures a wide range of metrics about farms and ranches, including their land use, land ownership, owner and worker demographics, and production of agricultural goods. The most recently released COA is from 2017 and includes data at the county, state, and national levels.

Use of these data in this tool: We downloaded county-level data from the most recent COA release with data from 2017.

American Community Survey

About the data source: The American Community Survey (ACS) is administered by the US Census Bureau and is the most detailed and complete source for population and housing information in the US. The ACS includes highly detailed and disaggregated representative data at the national level and down to incredibly small geographies (census block groups). Additionally, ACS data include both 1- and 5-year estimates for variables along with measured errors.

Use of these data in this tool: We downloaded data from the ACS with the R package tidycensus (https://walker-data.com/tidycensus/) and used 2021 5-year estimates for all variables. Five-year

estimates are used because of their smaller margins of error and because estimates are available for all counties, even those with small populations. We rely on the ACS for data on a wide range of demographic factors including population by race and ethnicity, poverty, homeownership, and more.

Spatial Hazard Events and Losses Database for the United States

About the data source: The Spatial Hazard Events and Losses Database for the United States (SHELDUS) is a climate hazard database operated by the Arizona State University Center for Emergency Management and Homeland Security. The SHELDUS database covers natural hazards such as thunderstorms, hurricanes, floods, wildfires, and tornados in the United States. These data are helpful in measuring the historical precedence of climate hazards. The database contains information on the date of an event, the affected location (county and state), and the direct losses caused by the event. Direct losses are measured by property and crop losses and are reported in inflation-adjusted per capita US dollars. The database is available at the county level, covers the period from January 1960 to December 2021, and is updated annually.

Use of these data in this tool: We downloaded county-level SHELDUS hazard data for three hazards (flooding, wildfires, and droughts) for the 10-year period between 2012 and 2021. Data on each of these three hazards were selected to better understand historical county-level climate damage and were included in the climate risk section of the tool. Flooding, wildfires, and droughts were of particular interest because of their direct influence on food system resilience. More details about these data, including how to obtain paid user licenses, can be found at the SHELDUS website.

Centers for Disease Control and Prevention National Public Health Tracking Network Data

About the data source: The Centers for Disease Control and Prevention's (CDC's) National Environmental Public Health Tracking Network (NEPHTN) is a system of integrated health, climate event exposure, and climate hazard information and data from a variety of city, state, and national sources. These data are collected and integrated to a core set of health, exposure, and hazards data that inform public health decisionmaking. The modeled extreme precipitation data and heat data for the CDC Data Explorer are obtained by NEPHTN from the North American Land Data Assimilation System. To determine population exposure to extreme precipitation or extreme heat, NEPHTN uses a multistage geo-imputation approach to convert hourly grid-level precipitation data to block, tract, and

county-level estimates. Population-level weights are used to transform data from the grid-level to the block, tract, and county-level estimates.

Use of these data in this tool: We downloaded county-level historical data on the annual counts of days with extreme precipitation and extreme heat days for the year 2019. These data were selected because extreme weather events can have a direct impact on food system resilience, for example, through water resource scarcity or harvest losses, and were included in the county-level section of the tool. The extreme precipitation and heat data in the data explorer are available at the county level nationally, except for counties in Hawaii and Alaska. The NEPHTN data and methodology can be accessed at the CDC Data Explorer.

Urban Institute Mobility Metrics

About the data source: The data bring together carefully selected metrics that are shown by strong evidence to have predictive relationships to upward economic and holistic mobility, such as educational attainment, access to stable housing, and the amount of debt families have. The creation of this dataset was informed by a working group that reviewed a wide range of factors and selected those that most directly influence mobility from poverty over the course of a person's life. The data are constructed using a variety of federal data sources, including Census data (such as American Community Survey and County Business Patterns data), the US Department of Health and Human Services' Area Health Resource File, and the Environmental Protection Agency's Air Quality Index. More detail on the Mobility Metrics can be accessed in the technical report (available at https://upward-mobility.urban.org/boosting-upward-mobility-metrics-inform-local-action-second-edition-technical-report) and in the data documentation (available at https://ui-research.github.io/mobility-from-poverty/).

Use of these data in this tool: We identify and use key variables that serve as economic indicators or could impact individuals and communities' food system resilience, including concentrated poverty, ratio home value, and ratio of average income to living wage. Mobility Metrics data are available at the county and city levels and are both longitudinal and disaggregated by race/ethnicity. We downloaded county-level race-disaggregated data on key metrics using the most recent year available for each variable (ranging from 2018 to 2021).

Census Bureau American Indian, Alaskan Native, and Native Hawaiian Areas

About the data source: The US Census Bureau tabulates and publishes TIGER/Line shapefiles for several geographic entities that include digital representations of key national geographic and cartographic information needed to create maps. One of the available types of TIGER/Line shapefiles covers all American Indian, Alaskan Native, and Native Hawaiian areas in the US. These shapefiles include both legal and statistical entities at the federal and state levels for which the Census Bureau publishes data, such as: American Indian reservations and off-reservation trust lands, tribal jurisdiction statistical areas, Alaska Native Regional Corporations, Alaska Native village statistical areas, tribal designated statistical areas, state designated tribal statistical areas, and Hawaiian home lands. The 2022 TIGER/Line shapefile technical documentation (available at https://www2.census.gov/geo/pdfs/maps-data/data/tiger/tgrshp2022/TGRSHP2022_TechDoc.pdf) can be accessed through the US Census Bureau.

Use of these data in this tool: We downloaded TIGER/Line shapefile data using the R package tigris (https://www.rdocumentation.org/packages/tigris/versions/2.0.4) and the function native_areas (https://www.rdocumentation.org/packages/tigris/versions/1.6.1/topics/native_areas) to estimate the amount of land in a county that is designated as tribal land, inclusive of reservation land and off-reservation trust land.

Variable Definitions

This section discusses the constructed variables used in the tool. It is structured by summarizing the construction of each variable under the following topics: food security and access, agricultural production, climate risk, racial equity, and demographics. The data source cleaning can be viewed in the 01_clean-data.R and 02_build-data.R files in the Github repository (https://github.com/UI-Research/food-systems-resilience-and-equity).

Food Security and Access

The food security and access variables in the tool include household food insecurity rate, cost per meal, share of households receiving Supplemental Nutrition Assistance Program (SNAP) benefits, and share of households with low store access and no vehicles.

HOUSEHOLD FOOD INSECURITY RATE

Definition: The estimated share of households that are food insecure (or may experience reduced quality, variety, desirability of diet, or reduced food intake) over a 12-month period.

Variable names: fi_rate_overall, fi_rate_black_any_eth, fi_rate_hisp_any_race, fi_rate_white_non_hisp, fi_rate_black_and_hispanic

Unit of measurement: Percentage

Data source: 2020 Map the Meal Gap, county-level data

Racial data disaggregation: Overall, Black (any race), Hispanic (any race), white (any race), combined Black/Hispanic

Data details: Household food insecurity rates are estimated by Feeding America using a two-way fixed effects model with data from all 50 states. We use the overall food insecurity rate included in the Map the Meal Gap data directly. Given sample size limitations, a researcher from MMG created a joint food insecurity rate for Black and Hispanic/Latinx groups as a proxy for communities of color, and disaggregated between white communities and communities of color.

Use in tool: Different disaggregations of this indicator can be viewed in the national map, county-level data on food security and access, and race-disaggregated data food insecurity sections of the tool.

COST PER MEAL

Definition: The estimated cost per meal based on 26 categories of food and state and local taxes.

Variable name: cost_per_meal

Unit of measurement: Dollars

Data source: 2020 Map the Meal Gap, county-level data

Racial data disaggregation: Not available

Data details: Cost per meal is calculated by Feeding America using data from NielsenIQ that accounts for differences in the price of food that exist between counties.

Use in tool: This indicator can be found in the county-level food security and access section of the tool.

SHARE OF HOUSEHOLDS RECEIVING SNAP

Definition: The share of households receiving Supplemental Nutrition Assistance Program benefits.

Variable Name: share_hh_receive_snap

Unit of measurement: Percentage

Data source: 2021 5-year ACS, county-level data

Racial data disaggregation: Not available

Data details: The share of households receiving SNAP benefits is determined by dividing the number of households receiving benefits by the number of households for which benefit receipt is determined.

Use in tool: This indicator can be viewed in the county-level food security and access section of the tool.

SHARE OF HOUSEHOLDS WITH LOW STORE ACCESS AND NO VEHICLES

Definition: The share of housing units in a county without a car that are more than one mile from a supermarket or large grocery store.

Variable name: pct_laccess_hhnv15

Unit of measurement: Percentage

Data source: 2015 Food Environment Atlas, county-level data

Racial data disaggregation: Not available

Data details: This variable is taken directly from the 2015 Food Environment Atlas.

Use in Tool: This indicator can be viewed in the county-level food security and access section of the tool.

Agricultural Production

The agricultural production variables in the tool include combined share of edible crop and animal production, food sold directly to consumers, and race and ethnicity of agricultural producers.

EDIBLE CROP AND ANIMAL PRODUCTION

Definition: The share of agricultural production that is dedicated to human-consumable crops and animal products.

Variable name: share_production_edible_crops, share_production_edible_animals, and share_production_edible_overall

Unit of measurement: Percentage

Data source: 2017 Census of Agriculture, county-level data

Racial data disaggregation: Not available

Data details: Share of animal production is calculated by dividing the dollar value of agricultural production of crops and animals suitable for human consumption by the total dollar value of agricultural production (edible and non-edible). Edible agricultural production is defined as agricultural production intended for human consumption, and includes both crops and animals, such as: wheat, rice, grains, vegetables, fruit, chickens, eggs, cattle, milk, and pigs. These variables are an approximation of edible production and do not precisely estimate the shares of each that may be for human consumption, as this is not specified in the data. For instance, corn is excluded from edible agricultural production, because the majority of domestic corn production is used for animal consumption. These variables were constructed from Census of Agriculture data in consultation with researchers from the National Agricultural Statistics Service to determine which agricultural crops and animal products are produced for human consumption. Additionally, a joint share of overall edible food production was created by summing the shares for edible crop and edible animal production.

Use in tool: The share of edible crop and edible animal production indicators can be viewed in the county-level agricultural production section of the tool. The joint share of edible food production can be viewed in the national maps section of the tool.

FOOD SOLD DIRECTLY TO CONSUMERS

Definition: Share of the total dollar value of agricultural production that is sold directly to consumers in US dollars.

Variable name: share_commodity_direct

Unit of measurement: Percentage

Data source: 2017 Census of Agriculture, county-level data

Racial data disaggregation: Not available

Data details: The variable is calculated by dividing the dollar amount of food sold directly to consumers (excluding distributors or nonlocal consumers) by the dollar sum of food sold to all consumers and distributors. This variable was constructed from Census of Agriculture data in consultation with researchers from the National Agricultural Statistics Service.

Use in tool: This indicator can be viewed in the county-level agricultural production section of the tool.

RACE AND ETHNICITY OF AGRICULTURAL PRODUCERS

Definition: Share of agricultural producers by race and ethnicity.

Variable name: share_producers_white, share_producers_black, share_producers_hispanic, share_producers_other_race, share_producers_aapi

Unit of measurement: Percentage

Data source: 2017 Census of Agriculture, county-level data

Racial data disaggregation: White, Black, Hispanic, Asian American and Pacific Islander, other races combined (all categories besides Hispanic are non-Hispanic unless otherwise indicated).

Data details: The share of agricultural producers by race and ethnicity is determined by dividing the number of producers of each race or ethnicity by the total number of producers in a given county. Each farm or site of agricultural production can have more than one producer for whom race and ethnicity is determined. The COA suppressed counties which had fewer than 30 agricultural producers of a given race or ethnicity.

Use in tool: This indicator can be viewed in the county-level agricultural production section of the tool.

Climate Risk

The climate risk variables in the tool include 10-year climate hazard climate damages, extreme precipitation days, and extreme heat days.

CLIMATE HAZARD DAMAGES

Definition: Combined flooding, wildfire, and drought climate hazard crop and property damages in inflation-adjusted US dollars.

Variable name: sheldus_average_p

Unit of measurement: Percentile

Data source: 2012–2021 Spatial Hazard Events and Losses Database for the United States (SHELDUS), county-level data

Racial data disaggregation: Not available

Data Details: This variable was constructed from SHELDUS data. First, county-level hazard data for flooding, wildfire, and drought hazards were downloaded for the 10-year period between 2012 and 2021. Property and crop damage were summed at the county level and over the 10-year-period of interest in inflation-adjusted (to the year 2021) per capita US dollars. The aggregate direct losses are measured by property and crop losses and are reported in inflation-adjusted per capita US dollars. Counties with no crop or property damage in the selected 10-year period are imputed with a zero-damage value; we observe that many counties have zero damages in each of the three categories.

Data suppression: Due to data license restrictions, only the binary quintile variable is available in the tool and publicly-shared data on the Data Catalog. Users who have a SHELDUS data license should be able to fully recreate our analysis by referring to the code repository.

Use in tool: This indicator can be viewed in the national map section of the tool.

EXTREME PRECIPITATION AND EXTREME HEAT DAYS

Definition: Number of extreme precipitation or extreme heat days in a given year by county.

Variable Name: share_ex_precip_days , share_ex_heat_days

Unit of Measurement: Percentage

Data source: Centers for Disease Control and Prevention's (CDC's) National Environmental Public Health Tracking Network (NEPHTN), county-level data

Racial data disaggregation: Not available

Data details: The extreme precipitation indicator was accessed from the "Precipitation & Flooding" Content Area and the "Historical Precipitation" indicator. The extreme heat indicator downloaded from the "Heat & Heat-related Illness (HRI)" Content Area and the "Historical Temperature & Heat Index" indicator; the extreme heat indicators are constructed to only capture values from May through September. These two inputs were used to generate a county-level number of extreme precipitation

and heat days in the year 2019. A day was considered to have "extreme" heat or precipitation if the relative 90th percentile threshold is exceeded—in other words, if in a given county, the weather on a particular day is hotter or rainier than 90 percent of all other days on record for that same county. Percentile thresholds are specific to each geography and are created across data from 1979 to 2020. A share of annual extreme precipitation days was created by dividing the annual number of extreme precipitation days by 365 days. Each county is percentile ranked according to its share of extreme precipitation and heat days, respectively.

Use in tool: This indicator can be viewed in the county-level climate hazards section of the tool.

Racial Equity

The racial equity variables of the tool serve as proxies or helpful contextual points for how structural racism affects communities, especially in the context of food system resilience, and reflect common correlates of food security and access (see https://apps.urban.org/features/disrupting-food-insecurity/). The racial equity variables in the tool include share employed, population below the federal poverty level, concentrated poverty, homeownership, ratio of home value, percentage of neighbors of different race and ethnicity, race and ethnicity share of county population, median household income, and tribal land percentage.

SHARE EMPLOYED

Definition: The share of adults ages 25 to 54 who are employed.

Variable name: share_employed_all, share_employed_black_non_hispanic, share_employed_hispanic, share_employed_white_non_hispanic, share_employed_other_races_and_ethnicities

Unit of measurement: Percentage

Data source: 2021 Urban Institute Mobility Metrics, county-level data

Racial data disaggregation: Overall, Black, white, Hispanic, other races and ethnicities combined (all categories besides Hispanic are non-Hispanic unless otherwise indicated)

Data details: The share of adults ages 25 to 54 who are employed is determined by dividing the number of employed adults by the total number of adults. This measure is different from the unemployment rate because it measures who out of the total adult population (not just those in the labor force) is employed.

Use in tool: This indicator can be viewed in the race-disaggregated data section of the tool.

POPULATION BELOW FEDERAL POVERTY LEVEL

Definition: The percentage of the population below 100 percent of the federal poverty level (FPL).

Variable name: share_employed_all, per_below_fpl_black, per_below_fpl_hisp, per_below_fpl_white_non_hisp, per_below_fpl_native, per_below_fpl_asian, per_below_fpl_pi

Unit of measurement: Percent

Data source: 2021 5-year ACS, county-level data

Racial data disaggregation: Overall, Black, Hispanic, white, Native American, Asian, Asian, Pacific Islander (all categories besides Hispanic are non-Hispanic unless otherwise indicated)

Data details: This variable is defined by dividing the number of people below 100 percent of the FPL by the total number of people for whom poverty status is determined in the ACS for a given county. The Census Bureau determines whether a person is in poverty (see https://www.census.gov/topics/income-poverty/poverty/guidance/poverty-measures.html) by assessing whether their household falls under specific income thresholds that vary by family size and composition.

Use in tool: This indicator can be found in the race-disaggregated data section of the tool.

CONCENTRATED POVERTY

Definition: Concentrated poverty is defined as the share of residents experiencing poverty who live in high-poverty neighborhoods.

Variable name: share_poverty_exposure_all, share_poverty_exposure_black, share_poverty_exposure_hispanic, share_poverty_exposure_white_non_hispanic, share_poverty_exposure_other_races_and_ethnicities

Unit of measurement: Percentage

Data source: 2021 Urban Institute Mobility Metrics, county-level data

Racial data disaggregation: Overall, Black, Hispanic, white, other races and ethnicities combined (all categories besides Hispanic are non-Hispanic unless otherwise indicated)

Data details: This variable is taken directly from the 2021 Urban Institute Mobility Metrics. Additional details can be found in the Mobility Metrics technical report (https://upward-mobility.urban.org/boosting-upward-mobility-metrics-inform-local-action-second-edition-technical-report) and data documentation (https://ui-research.github.io/mobility-from-poverty/). High-poverty

neighborhoods are defined as census tracts in which more than 40 percent of the residents are experiencing poverty.

Use in tool: This indicator can be viewed in the race-disaggregated data section of the tool.

HOMEOWNERSHIP

Definition: The share of owner-occupied homes in a given county out of the total number of occupied homes.

Variable Name: per_h_owner, per_h_owner_black, per_h_owner_hisp, per_h_owner_non_hisp_white, per_h_owner_native, per_h_owner_asian, per_h_owner_pi

Unit of Measurement: Percent

Data source: 2021 5-year ACS, county-level data

Racial data disaggregation: Overall, Black, Hispanic, white, Native American, Asian, Asian, Pacific Islander (all categories besides Hispanic are non-Hispanic unless otherwise indicated)

Data details: Homeownership rates are calculated by dividing the number of owner-occupied homes for each racial/ethnic subgroup by the total number of homes occupied (by renters or owners) for that subgroup in a given county. Additionally, an overall homeownership rate is calculated using totals without racial/ethnic disaggregation.

Use in tool: This indicator can be viewed in the race-disaggregated data section of the tool.

RATIO OF HOME VALUE

Definition: The ratio of the share of a community's housing wealth held by a racial/ethnic group to the share of households of the same group.

Variable name: ratio_black_nh_house_value_households, ratio_hispanic_house_value_households, ratio_white_nh_house_value_households, ratio_other_nh_house_value_households

Unit of measurement: Ratio

Data source: 2021 Urban Institute Mobility Metrics, county-level data

Racial data disaggregation: Black, Hispanic, white, other races and ethnicities combined (all categories besides Hispanic are non-Hispanic unless otherwise indicated)

Data details: This variable is taken directly from the 2021 Urban Institute Mobility Metrics. Additional details can be found in the Mobility Metrics technical report (https://upward-mobility.urban.org/boosting-upward-mobility-metrics-inform-local-action-second-edition-technical-report) and data documentation (https://ui-research.github.io/mobility-from-poverty/).

Use in tool: This indicator can be viewed in the race-disaggregated data section of the tool.

PERCENTAGE OF NEIGHBORS OF DIFFERENT RACE AND ETHNICITY

Definition: The average share of each racial/ethnic group's neighbors who are members of other racial or ethnic groups.

Variable name: share_white_nh_exposure, share_black_nh_exposure, share_hispanic_exposure, share_other_nh_exposure

Unit of measurement: Percentage

Data source: 2021 Urban Institute Mobility Metrics, county-level data

Racial data disaggregation: White, Black, Hispanic, other races and ethnicities combined (all categories besides Hispanic are non-Hispanic unless otherwise indicated)

Data details: This variable was taken directly from the 2021 Urban Institute Mobility Metrics. Additional details can be found in the Mobility Metrics technical report (https://upward-mobility.urban.org/boosting-upward-mobility-metrics-inform-local-action-second-edition-technical-report) and data documentation (https://ui-research.github.io/mobility-from-poverty/). This measure is a segregation proxy and is calculated with a dissimilarity index for the races and ethnicities of a person's neighbors.

Use in tool: This indicator can be found in the race-disaggregated data section of the tool.

RACE AND ETHNICITY SHARE OF COUNTY POPULATION

Definition: The share of a county's population by race and ethnicity.

Variable name: per_non_hisp_black, per_non_hisp_white, per_hisp, per_non_hisp_native, per_non_hisp_asian, per_non_hisp_pi

Unit of measurement: Percentage

Data source: 2021 5-year ACS, county-level data

Racial data disaggregation: White, Black, Hispanic, Native American, Asian, Pacific Islander (all categories besides Hispanic are non-Hispanic unless otherwise indicated)

Data details: For each racial/ethnic group we calculate the percentage of the county population by dividing the subgroup's population by the county's total population.

Use in tool: This indicator can be viewed in the race-disaggregated data section of the tool.

MEDIAN HOUSEHOLD INCOME

Definition: The median, or 50th percentile, household income in a given county.

Variable name: median_hh_income, median_hh_income_black, median_hh_income_hisp, median_hh_income_non_hisp_white, median_hh_income_native, median_hh_income_asian, median_hh_income_pi

Unit of measurement: Dollars

Data source: 2021 5-year ACS, county-level data

Racial data disaggregation: Overall, Black, Hispanic, white, Native American, Asian, Pacific Islander (all categories besides Hispanic are non-Hispanic unless otherwise indicated)

Data details: This variable is taken directly from the 2021 5-year ACS. A household is defined by the US Census Bureau as including all of the people who reside within a housing unit.

Use in tool: This indicator can be viewed in the race-disaggregated data section of the tool.

TRIBAL LAND PERCENTAGE

Definition: The percentage of land within a county that is designated as tribal land by the US Census Bureau.

Variable name: tribal_percent

Unit of measurement: Percentage

Data source: US Census Bureau TIGER/Line shapefiles (tigris R package; see
https://www.rdocumentation.org/packages/tigris/versions/2.0.4)

Racial data disaggregation: Not available

Data details: This indicator was created using US Census Bureau TIGER/Line shapefiles covering all American Indian, Alaskan Native, and Native Hawaiian areas in the US and the county-level shapefiles (accessed through R package tigris). Tribal land area is defined consistently with the US Census Bureau to include both reservation and off-reservation trust land. The area of tribal land in a county is calculated by obtaining the intersection of all tribal land area with a given county's land area. Counties with no tribal land area are imputed with a zero land area value (2,502 counties out of 3,143 US counties). The tribal land percentage indicator is calculated by dividing the percentage of tribal land area by the total land area in a given county.

Use in tool: This indicator can be viewed in the national maps and in the county-level climate hazards sections of the tool.

Demographics

The demographic variables in the tool include ratio of average income to living wage, unemployment rate, and share with a disability.

RATIO OF AVERAGE INCOME TO LIVING WAGE

Definition: The ratio of pay for the average job to the local cost of living.

Variable name: ratio_average_to_living_wage

Unit of measurement: Ratio

Data source: 2021 Urban Institute Mobility Metrics, county-level data

Racial data disaggregation: Not available

Data details: This variable is taken directly from the 2021 Urban Institute Mobility Metrics. Additional details can be found in the Mobility Metrics technical report (https://upward-mobility.urban.org/boosting-upward-mobility-metrics-inform-local-action-second-edition-technical-report) and data documentation (https://ui-research.github.io/mobility-from-poverty/).

Use in tool: This indicator can be viewed in the county-level demographic details section of the tool.

UNEMPLOYMENT RATE

Definition: The share of unemployed people in the labor force out of the total number of people in the labor force.

Variable name: unemp_rate

Unit of measurement: Percentage

Data source: 2021 5-year ACS, county-level data

Racial data disaggregation: Not available

Data details: Unemployment is calculated by dividing the number of unemployed people by the total number of civilians in the labor force (both employed and unemployed people in the labor force, not including those not in the labor force, such as children under 18) in a given county.

Use in tool: This indicator can be viewed in the county-level demographic details section of the tool.

SHARE WITH A DISABILITY

Definition: The share of people with a disability.

Variable name: per_disability

Unit of measurement: Percentage

Data source: 2021 5-year ACS, county-level data

Racial data disaggregation: Not available

Data details: Disability rates are calculated by dividing the number of people with at least one disability by the total population in a given county.

Use in tool: This indicator can be viewed in the county-level demographic details section of the tool.

Methodology

This section discusses the methodological approach used to construct the national maps and the nearest neighbors analysis. It is structured by summarizing the generation of the percentile ranks used in the "county-level" tab, the top quintiles displayed in the national map, and the K-nearest neighbor (KNN) analysis shown by county. The generation of percentiles and top quintiles can be viewed in the 02_build-dataset.R script, while the KNN analysis can be viewed in the 03_generate-knn.R file in the Github repository (https://github.com/UI-Research/food-systems-resilience-and-equity).

Percentile Ranking

Using raw measures generated from the food security and access, agricultural production, climate risk, racial equity, and demographics variables, we assign a percentile rank measure to each county. The percentile rank assigns a given county a score from a scale of 1 to 100 for how it compares to the rest of the raw values of other counties. For example, if a county's raw number of extreme heat days were the greatest of any US county in the dataset, that county would receive a value close to the 100th percentile. We used the R function percent_rank() from the dplyr package, which counts the total number of values less than the values of the county in question, and divides it by the number of observations minus 1.

The quintile map views can be viewed in the county-level county comparisons tab in the "county comparison" drop-down selection of the tool. The percentile variables are denoted by a "_p" variable name suffix in the data catalog.

Quintile County-Level National Maps

To identify communities facing intersecting climate and food systems risks that may have insufficient resources to mitigate against these risks, we construct national-level maps depicting whether counties are in the top quintile of both climate damages and food systems risks, which we define for this analysis as county-level food security rates (overall and among Black and Latinx households specifically), agricultural production, or tribal land presence.

To do this, we first make a binary indicator for whether a county is in the top quintile of counties in terms of its: (1) per capita climate damages, (2) food insecurity rate (overall), (3) food insecurity rate among Black and Latinx households, and (4) share of edible agricultural production. Then, counties with any tribal land within them are designated with a binary indicator. Lastly, we identify counties with both high climate risks and risks to food systems in the following four ways:

1. counties in the top quintile of climate damages and overall food insecurity rates

2. counties in the top quintile of climate damages and food insecurity rates for Black and Latinx households

3. counties in the top quintile of climate damages and share of edible agricultural production

4. counties with tribal land presence in the top quintile of climate damages

The percentile ranking can be viewed in the national maps sections of the tool.

K-Nearest Neighbors (KNN) Analysis

We identify counties that share similar vulnerabilities related to food insecurity, climate hazards, and agricultural production using the K-nearest neighbors (KNN) algorithm. Euclidean distance (see https://rpubs.com/pmtam/knn), the straight-line distance between two points, was used to calculate which observations are closest in proximity to a given data point. The formula for calculating Euclidean distance between two observations *i* and *j* across *p* variables is as follows:

$$D_{ij} = \sqrt{\sum_{m=1}^{p} \left(x_{im} - x_{jm} \right)^2}$$

For the purposes of our analysis, a k-value of 5 is selected, generating the 5 nearest neighbor counties. The KNN algorithm was applied to three primary domains in the dataset: climate hazards, food production, and food security and access. The data for each of these domains were composed of the following county-level indicators:

- **Climate hazard damages:** Combined flooding, wildfire, and drought climate hazard crop and property damages from 2012–2021 (see variable definition above)

- **Agricultural production:** Combined edible crop production and edible animal production (see variable definition above)

- **Food security and access:** Overall food insecurity rate and food insecurity rate for Black and Hispanic/Latinx people (see variable definitions above)

The data were preprocessed for the KNN analysis in the following ways:

- Indicators were normalized before calculating nearest neighbors, so that distances were not driven by differing units of measurement.

- Missing or zero values were also addressed before applying the KNN algorithm.

  » Counties with missing values in climate hazards, food production, and food security and access were dropped, since calculating Euclidean distances requires complete data. For food production, for example, 27 counties were excluded due to having missing values in the original indicator.

  » For the climate hazards data, to prevent the large number of zero values (for counties with no reported crop or property damage in the selected 10-year period) from introducing significant noise into the results, counties with zeroes across all the wildfire, drought, and

flooding indicators were excluded. This left 2,586 counties remaining out of the original 3,143 counties for the climate hazard KNN analysis.

The algorithm was applied using the R function KNN from the dbscan package (see https://cran.r-project.org/web/packages/dbscan/index.html). This nearest neighbor analysis can be viewed in the county-level county comparisons for the climate hazards, agricultural production, and food security and access sections of the tool.

Data and Methodological Limitations

Although this pilot tool seeks to provide an overview of food system resilience concerns at the intersection of climate hazards, food production and access, and structural racism, gaps in data availability and quality limit its scope.

First, there are a few limitations with the data used in the pilot tool which introduce a certain level of error. These include some limitations to the climate and food insecurity data on (1) the CDC National Public Health Tracking Network Data, (2) MMG food insecurity estimates, and (3) SHELDUS data. For example, the CDC National Public Health Tracking Network Data on extreme precipitation and heat excludes Hawaii and Alaska and relies on modeled data (not weather station–based observations). Because Hawaii (https://nca2018.globalchange.gov/chapter/27/) and Alaska (https://nca2018.globalchange.gov/chapter/26/) often disproportionately bear the brunt of climate disasters, we chose to not include these data in anything other than descriptive analysis. The race and ethnicity data for the food insecurity and access measure in the MMG data are also limited in that they do not have separate measures for Asian American and Pacific Islander (AAPI) and American Indian Alaska Native (AIAN) communities due to small sample sizes in the underlying data used to calculate food insecurity rates. A joint Black and Hispanic/Latinx food insecurity rate is used as an imperfect proxy for food insecurity among all communities of color. Lastly, the SHELDUS climate hazard data used in the tool only include historical data, and do not project future risk. Because climate risk is compounding and accelerating in severity (see https://www.oecd.org/env/cc/2482280.pdf), historical climate hazards data are an imperfect proxy for future climate risks.

Second, there are some methodological limitations to the pilot tool. For example, the race- and ethnicity-disaggregated data show how different dimensions of racial inequality and equity (such as concentrated poverty, employment, and homeownership) show up in communities, but they may show an incomplete picture of the full spectrum of racial inequities and how they interact with food systems.

For example, concentrated poverty may show how people can respond individually to a climate or food disaster, but it may not show how structurally racist policies in housing, land use, and planning may affect the community resources available for responding effectively to a crisis. Separately, the analysis in the pilot tool is limited to descriptive, not causal or inferential, analysis.

Despite its limitations, this pilot tool gathers diverse data sources into one data asset, provides methods/analysis to identify counties with shared risks, and enables analysis of counties facing intersecting risks of climate change, food insecurity and access, and agricultural production. For research next steps and extensions of this tool, please refer to our accompanying working paper, "Characterizing Local Food System Resilience to Climate Hazards in Communities of Color," available on the APPAM website.

Acknowledgments

For more information on this project, see Kassandra Martinchek, Judah Axelrod, Sonia Torres Rodríguez, Amy Rogin, Noah Johnson, and Elaine Waxman, "Understanding the Intersection of Climate Risk, Food System Resilience and Racial Equity: A County-Level Data Tool," Urban Institute, October 31, 2023.

URBAN INSTITUTE

500 L'Enfant Plaza SW
Washington, DC 20024
www.urban.org

ABOUT THE URBAN INSTITUTE

The Urban Institute is a nonprofit research organization that provides data and evidence to help advance upward mobility and equity. We are a trusted source for changemakers who seek to strengthen decisionmaking, create inclusive economic growth, and improve the well-being of families and communities. For more than 50 years, Urban has delivered facts that inspire solutions—and this remains our charge today.