Investigating U.S. Food Insecurity Through Data

Conrad Linus Muhirwe Sharon Wanyana Ryann Tompkins
Alex Arevalo

2025-10-22

Table of contents

1. Administration

App Title: Investigating U.S. Food Insecurity Through Data

Team Members: - Conrad Linus Muhirwe - Sharon Wanyana - Ryann Tompkins - Alex Arevalo

Course: DATA-613: Data Science Practicum

2. Project Overview

This project examines the prevalence and distribution of household food insecurity across U.S. states (and, where feasible, counties) and key demographic groups. We link geographic patterns to socioeconomic conditions (income, poverty, employment, education) and local food environment characteristics, and surface insights through an interactive Shiny application designed for policy and program stakeholders.

Objectives

1. Describe levels and trends in food insecurity across places and populations (2023–2025).
2. Explain variation using socioeconomic drivers and food-environment context.
3. Contextualize results with safety-net program reach (SNAP/WIC participation).
4. Communicate findings via an accessible, exportable, and update-ready Shiny tool.

2.1 Problem Context

Food insecurity remains a persistent challenge in the United States, with far-reaching effects on public health, education, and economic mobility. It disproportionately affects low-income households, racial and ethnic minorities, families with children, and communities with limited access to stable employment or affordable food outlets. These disparities reflect broader structural inequities such as income volatility, housing costs, and unequal access to education and social safety nets.

This project addresses food insecurity through three core objectives:

- Quantify its prevalence and geographic distribution across states and demographic groups.
- Analyze relationships between food insecurity and socioeconomic factors such as income, poverty, employment, and education.
- Visualize insights through an interactive Shiny app that maps trends and supports evidence-based decision-making.

The application is designed for policymakers, researchers, and nonprofit practitioners who need accessible, timely data to inform interventions like SNAP outreach, school meal programs, and workforce initiatives. By integrating transparent analysis with interactive visualization, the project enables data-driven strategies to reduce food insecurity and promote social equity.

2.2 Literature Review

National monitoring by the USDA Economic Research Service (ERS) indicates that food insecurity rose to 13.5% of U.S. households in 2023, reversing several years of pre pandemic improvement and reflecting the continued impact of inflation and economic volatility (USDA ERS, 2024). These official estimates provide the baseline for cross-state and subgroup comparisons in this project.

Spatial analyses reveal clear regional patterns: rural and southern counties consistently report the

highest food insecurity rates, driven by economic disinvestment and limited access to affordable food outlets (Feeding America, 2024). Such geographic disparities highlight the relevance of contextual variables rurality, retail food density, and transportation access in understanding food insecurity dynamics.

Extensive empirical work demonstrates that food insecurity remains strongly associated with income, race/ethnicity, and family structure, and is linked to adverse health outcomes throughout the life course (Gundersen & Ziliak, 2015). This evidence supports our use of American Community Survey (ACS) variables poverty, income, unemployment, and education to model exposure risk and analyze subgroup variation (e.g., single-parent households, families with children, and racially minoritized populations).

Finally, program participation plays a critical mitigating role. Rigorous quasi-experimental research by Gundersen, Kreider, and Pepper (2017) finds that participation in the Supplemental Nutrition Assistance Program (SNAP) reduces food insecurity by approximately six percentage points among households with children. This finding motivates inclusion of SNAP participation rates as a key covariate in our state-level models, recognizing that policy access and uptake are central mediators of household vulnerability.

2.3 Proposed Data Source

The project will rely primarily on the USDA Economic Research Service (ERS) Food Access Research Atlas, a comprehensive open dataset that maps food access indicators across U.S. census tracts. This dataset supports the project's objectives of analyzing geographic and socioeconomic disparities in food insecurity and related access conditions.

Dataset	Source	Year	Geographic Unit	No of Variables	No of Observations	Source Link
Food Access Research Atlas	U.S. Department of Agriculture, Economic Research Service (USDA ERS)	2025 (latest release, January 5, 2025)	Census tract (U.S. nationwide)	147 variables	72,000 tracts across 50 states + DC	https://www.ers.usda.gov/data-products/food-access-research-atlas/

2.3.1 Description and Purpose:

The USDA Food Access Research Atlas provides tract level data describing food access and socioeconomic characteristics that influence household food security. It includes measures such as income level, distance to supermarkets, vehicle availability, and urban-rural classification. The dataset also contains demographic attributes (e.g., poverty rate, racial composition, population density), enabling multivariate exploration of social determinants of food insecurity.

Key Variables of Interest:

- **TractFIPS** Census tract identifier for geographic mapping.
- **State** State name for aggregation and filtering.
- **LILATracts_1And10** Low-income and low-access tract indicator.

- **PovertyRate** Percent of population below poverty line.
- **lapophalfshare** Population with low food access within ½ mile.
- **Urban** Urban/rural classification.
- **Nsupermarkets** Count of supermarkets within threshold distance.

3 Use Case and Actor Description

This Shiny application supports public policy analysis and research on food insecurity across the United States by making socioeconomic and food access data interactive and interpretable for decision makers and researchers.

The primary users are public policy analysts and nonprofit researchers in food policy, community development, or social welfare professionals who need to identify geographic disparities, assess high-risk populations, and communicate data-driven insights.

Through the app, users can

- Explore food insecurity data at national, state, and county levels.
- Analyze relationships with socioeconomic factors such as income, education, and employment.
- Create visual summaries and reports for policy and program decisions.

The app combines exploration and analysis tools to support evidence-based decision-making across federal, state, and local contexts.

3.1 Actor Questions of Interest

The Shiny app is designed to help the actor answer key research and policy questions such as:

- How does food insecurity vary across U.S. states and regions over time?
- Are there seasonal or long-term trends that can inform future policy focus?
- What is the relationship between food insecurity and socioeconomic variables (e.g., income, unemployment, education)?
- Which factors appear to be the strongest predictors of food insecurity?
- How do rural vs. urban areas differ in food insecurity rates?
- Which demographic groups such as households with children or racial/ethnic minorities are most affected?
- Can we observe measurable shifts in food insecurity following policy interventions (e.g., SNAP expansions or COVID-era relief programs)?

3.2 Anticipated Workflow for the Actor

The app's design follows the logical steps of a policy analyst's exploratory workflow:

1. Select Scope – Choose a geographic level (national, state, or county) and a time frame (e.g., 2023–2025).
2. Filter Variables – Select socioeconomic indicators (e.g., poverty rate, income, unemployment, education).
3. Visualize Data – Generate interactive charts, heatmaps, and choropleth maps showing temporal and geographic trends.
4. Explore Relationships – View summary statistics, correlations, and regression results dynamically.

5. Generate Insights & Export – Download customized tables, figures, or reports for communication and decision-support.
6. Iterate – Adjust filters or parameters to test new hypotheses or compare across states and years.

4. Ethical Review

Ethical conduct underpins the integrity and societal value of this project. Given the sensitivity of food insecurity data, the team will follow ethical principles throughout all stages of the data life cycle from acquisition to dissemination in alignment with the American Statistical Association's (ASA) Ethical Guidelines for Statistical Practice (2023).

Ethical Considerations Across the Data Life Cycle

1. **Data Acquisition.** The project uses only publicly available, de-identified datasets from trusted sources such as the USDA ERS, U.S. Census Bureau (ACS), and Feeding America. All data comply with open-data licenses, include clear citations and metadata, and exclude any proprietary or personally identifiable information (PII).
2. **Data Storage and Management.** All data will be securely organized in the project's GitHub repository with a clear directory structure and version control to ensure traceability and reproducibility. Only cleaned, approved datasets will be uploaded, while raw data remains unaltered in protected folders. Team members will limit access to necessary files and document all updates through detailed commit messages.
3. **Data Cleaning and Analysis.** All data transformations will be transparently documented to ensure reproducibility and objectivity. The team will follow ASA guidelines for methodological rigor, avoiding biased manipulation of variables or results. Model assumptions and limitations will be clearly stated, and correlations or regressions will be interpreted cautiously without implying causation.
4. **Data Visualization and Interpretation.** The app will follow strict visual ethics to ensure clarity and accuracy. All charts will use correct scaling, clear labels, and accessible color palettes, avoiding any distortion or omission of uncertainty. Language in visuals will remain neutral and respectful, especially when referencing vulnerable groups or regions with high food insecurity.
5. **Dissemination and Communication.** The Shiny app will be used for educational and analytical purposes to inform evidence-based policymaking. Results will include clear notes on limitations and potential errors, encouraging users to interpret findings within their broader socioeconomic context.

5. Application Concept

5.1 General Layout of the App to Support the Actor Workflow

The Shiny application is an interactive dashboard designed for policy analysts, researchers, and nonprofit professionals to explore U.S. food insecurity data through a structured workflow of overview → exploration → analysis.

It features three main tabs:

1. `Overview` – Displays national food insecurity summaries, KPIs, and data definitions.
2. `Exploration` – Offers interactive maps and charts to compare trends across states, counties,

and demographic groups.

3. `Analysis` – Provides statistical tools such as correlations, regressions, and group comparisons to examine relationships with socioeconomic variables.

A consistent layout with sidebar controls and dynamic outputs ensures an intuitive, user-friendly experience for data-driven decision-making.

5.2 Planned Options for User Data Selection and Manipulation

The app will allow users to tailor analyses through interactive inputs that dynamically update visualizations and tables. Key controls include:

- Geographic Filters - Dropdown menus for selecting specific states, regions, or counties.
- Temporal Filters - Sliders or dropdowns to set year ranges (e.g., 2010 – 2024).
- Variable Selection - Checkboxes or multi-select inputs to include/exclude socioeconomic indicators such as median income, unemployment, education, and poverty rate.
- Transformations - Options to apply log transformations or normalization to improve comparability across scales.
- Subgroup Focus - Filters to isolate demographic or regional groups (e.g., rural vs. urban, households with children).

These tools will update outputs in real time, allowing flexible exploration from multiple analytical perspectives without requiring users to write code.

5.3 Planned Options for User Exploratory Numerical / Graphical Analysis

The Exploration Tab will include a range of descriptive and graphical tools to facilitate insight generation:

- Interactive Maps (Choropleths) - Show food-insecurity rates by state or county with color gradients for geographic comparison.
- Trend Analysis - Line and bar plots illustrating temporal changes in food insecurity and related socioeconomic factors.
- Comparative Visuals - Multi-panel or faceted charts to compare multiple states or demographic groups simultaneously.
- Summary Tables - Interactive data tables built with {DT} for sorting, searching, and exporting selected metrics.
- Dynamic Tooltips - Hover-over elements displaying contextual statistics such as exact rates, rankings, or summary values.

Visualizations will be developed using {ggplot2}, {plotly}, and {leaflet}, ensuring interactivity, consistency in color schemes, and accessibility for all users.

5.4 Planned Options for User Statistical Models and Tests

The Analysis Tab will provide lightweight, transparent tools for exploratory statistical modeling and inference:

1. Correlation Analysis
    - Compute Pearson and Spearman correlation coefficients between food insecurity and socioeconomic indicators.

- Display results as interactive correlation matrices or heatmaps.

2. Regression Modeling
   - Allow users to specify dependent and independent variables for simple or multiple linear regression.
   - Present model summaries (coefficients, R^2, p-values) with diagnostic plots and interpretation notes.
3. Group Comparisons
   - Conduct two-sample t-tests comparing mean food-insecurity rates across categories (e.g., rural vs. urban, high- vs. low-income states).
   - Report p-values, confidence intervals, and short text interpretations.
4. Predictive Exploration
   - Offer reactive model fitting to generate predicted food-insecurity rates given socioeconomic inputs.

All analyses will be fully transparent, with clear documentation of assumptions, limitations, and non-causal interpretation. Users will have the option to download model results, tables, and figures for inclusion in reports or policy briefs.

6. Collaboration and Branching Plan

The team will use a Shared Repository Workflow Model to coordinate development and version control through GitHub. The main branch will hold stable, production-ready code, while each member works on dedicated feature branches (e.g., data-cleaning, ui-design, analysis-module).

Team members will commit regularly with clear messages, submit pull requests for peer review, and merge only approved changes. GitHub Issues and project boards will track tasks, and weekly meetings will align progress.

This structured workflow ensures efficient collaboration, accountability, and consistent code quality while supporting independent contributions and smooth integration.

6.1 Team Roles and Responsibilities

Team Member	Course Level	Primary Responsibilities
Conrad Linus Muhirwe	DATA-613	• Repository setup, version control & documentation. • Statistical modeling, validation and interpretation
Sharon Wanyana	DATA-613	Correlation & Regression analysis, hypothesis testing
Ryann Tompkins	DATA-613	• Literature & ethical review • Exploratory data analysis & visualization using `{ggplot2}` and `{plotly}`, summary tables plus interactive maps.
Alex Arevalo	DATA-413	Data acquisition, cleaning, and transformation, UI & UX.

6.2 Schedule and Milestones

Milestone	Target Date	Deliverable / Outcome
Data acquisition and Project Plan	Oct 24 2025	Verified `.rds` datasets stored in `/data` folder; cleaning scripts documented.
App skeleton (UI structure & tabs)	Oct 30 2025	Functional navigation across Overview, Exploration, and Analysis tabs.
Progress Report submission	Nov 11 2025	Updated `.qmd` and HTML with preliminary visuals and workflow.
Statistical modeling integration	Nov 20 2025	Regression and correlation modules functional with outputs.
App refinement & user testing	Dec 3 2025	Fully interactive dashboard tested across user scenarios.
Vignette drafting & demo preparation	Dec 8 2025	Completed vignette, rehearsal of oral demonstration.
Final submission	Dec 9 2025	Final GitHub repo, HTML deliverables, and class demonstration.

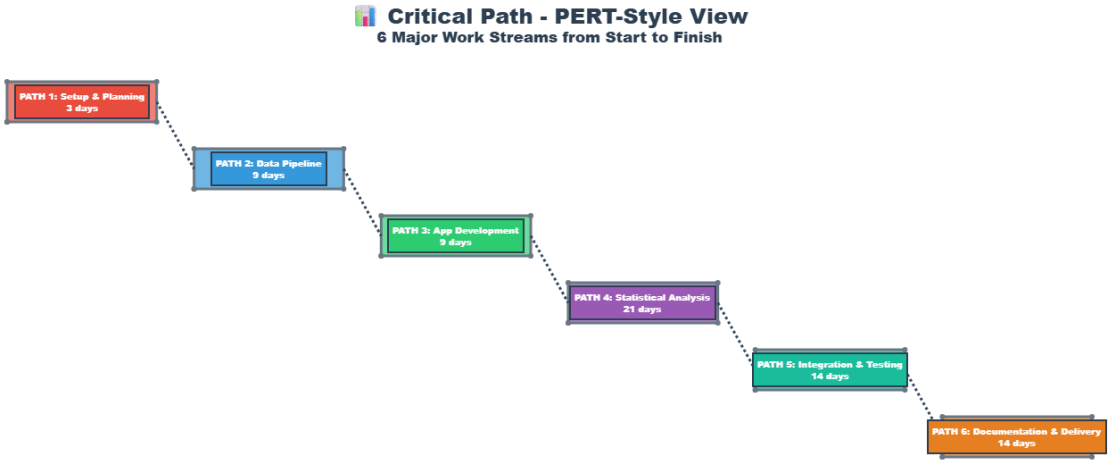

Figure 1: Critical Path

REFERENCES

- Feeding America. (2024). Map the Meal Gap 2024: Executive summary. https://www.feedingamerica.org the-meal-gap/overall-executive-summary
- Gundersen, C., & Ziliak, J. P. (2015). Food insecurity and health outcomes. Health Affairs, 34(11), 1830–1839. https://www.healthaffairs.org/doi/10.1377/hlthaff.2015.0645
- Gundersen, C., Kreider, B., & Pepper, J. (2017). Partial identification methods for evaluating food assistance programs: A case study of the causal impact of SNAP on food insecurity. American Journal of Agricultural Economics, 99(4). https://onlinelibrary.wiley.com/doi/10.1093/ajae/aax026
- U.S. Department of Agriculture, Economic Research Service. (2024). Household food security in the United States in 2023 (ERR-337). https://www.ers.usda.gov/publications/pub-details?pubid=109895