

# Project phase 1

Group 18

2025-08-27

Which socio-economic and geographic factors best explain and potentially predict the disparities in food access between urban and rural communities in the United States

## Exploratory Data Analysis

To address the research question of “rural-urban disparities in food access,” the exploratory data analysis can be conducted from several perspectives.

First we conduct comparisons between different years (2010, 2015, and 2019) to examine whether urban-rural disparities are widening or improving.

Table 1: The count and propotion of urban classification in each year

year	Urban	count	proportion
2010	0	17568	0.241
2010	1	55296	0.759
2015	0	17692	0.243
2015	1	55172	0.757
2019	0	17362	0.239
2019	1	55169	0.761

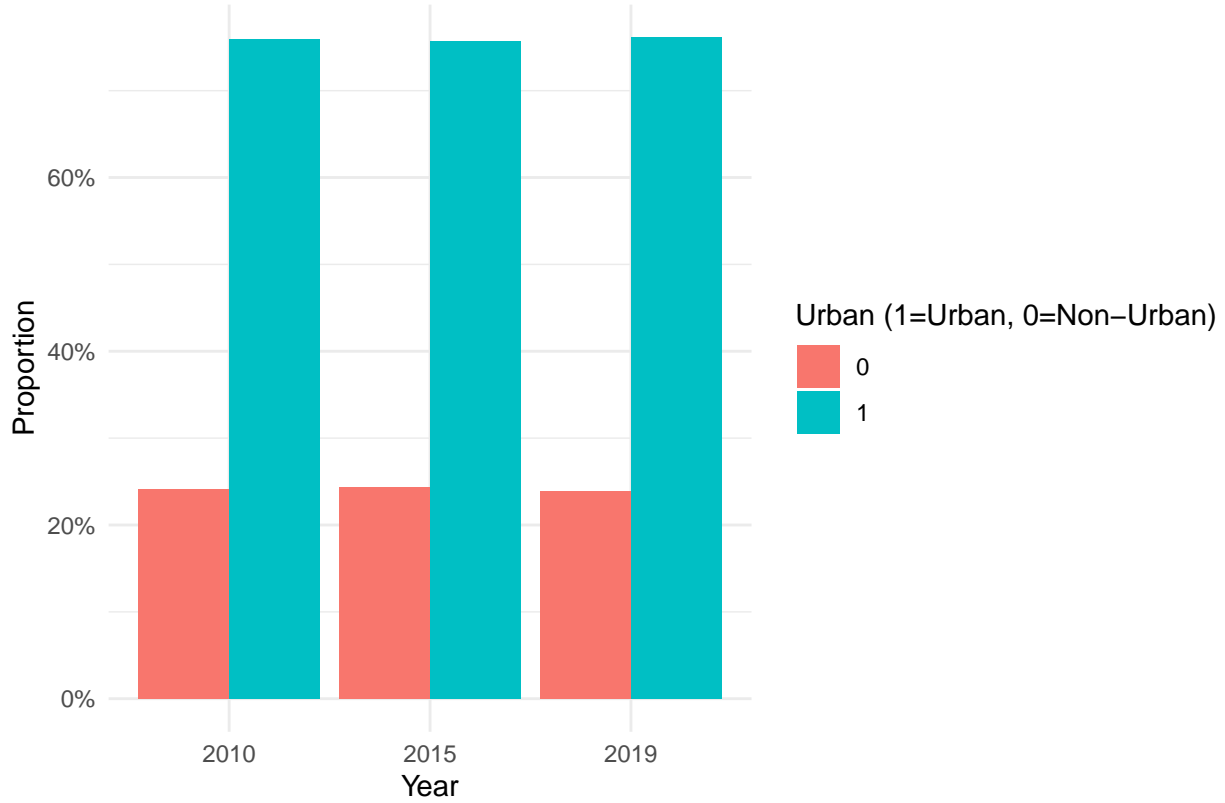


Figure 1: Urban proportion by year

Table 1 and Figure 1 present the distribution of urban status across three years: 2010, 2015, and 2019. The variable Urban is coded as 0 for rural and 1 for urban. Both counts and proportions are reported. In 2010, there were 17,568 rural observations (24.1%) and 55,296 urban observations (75.9%). In 2015, the distribution remained very similar, with 17,692 rural (24.3%) and 55,172 urban (75.7%). In 2019, there were 17,362 rural (23.9%) and 55,169 urban (76.1%). The proportion of urban and rural areas has been relatively stable over the three time points. Urban areas consistently account for about three-quarters of the total, while rural areas make up about one-quarter. There is a small decrease in the number and proportion of rural observations in 2019 compared to 2010 and 2015, suggesting a modest trend toward urbanization. However, the change is minor (less than 1 percentage point). The data indicate that the urban-rural composition has not undergone substantial change over the decade, suggesting that any observed differences in outcomes over time are unlikely to be driven by large shifts in urbanization.

To evaluate the distribution of the proportion of people with low food accessibility across the country, we use the variable `LILATracts_1And10` indicating low income and low access tract measured at 1 mile for urban areas and 10 miles for rural areas.

Table 2: The count and proportion of low income and low food accessibility separated by urban status in each year

year	Urban	LILATracts_1And10	count	proportion
2010	0	0	16153	0.919
2010	0	1	1415	0.081
2010	1	0	47752	0.864
2010	1	1	7544	0.136
2015	0	0	16352	0.924

year	Urban	LILA_Tracts_1And10	count	proportion
2015	0	1	1340	0.076
2015	1	0	47267	0.857
2015	1	1	7905	0.143
2019	0	0	16010	0.922
2019	0	1	1352	0.078
2019	1	0	47228	0.856
2019	1	1	7941	0.144

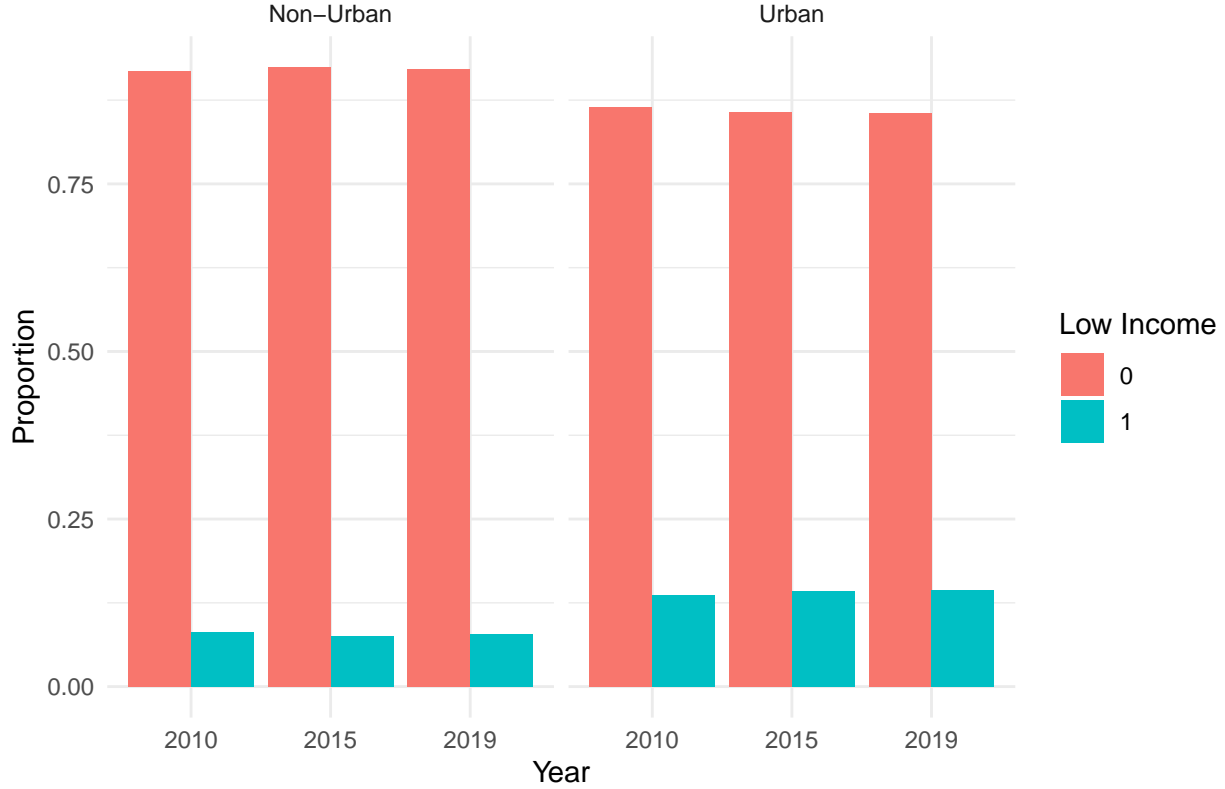


Figure 2: Low Income Proportion and Low Food Accessibility by Year and Urban Status

From Table 2 and Figure 2, it is found in all three years, the proportion of low-income tracts (LILA\_Tracts\_1And10 = 1) is consistently higher in urban areas (13.6%–14.4%) than in non-urban areas (7.6%–8.1%). This suggests that urban areas face slightly higher rates of low income combined with low food accessibility compared to non-urban areas. Across 2010, 2015, and 2019, the proportion of low-income tracts remains relatively stable within both urban and non-urban areas. The values are around 8% for non-urban areas and around 14% for urban areas. This indicates that over the decade, the disparity between urban and non-urban low-income food access has not changed substantially. In absolute numbers, urban areas have more low-income tracts due to higher population and tract density: e.g., in 2019, urban low-income tracts = 7,941 vs. non-urban = 1,352. The majority of tracts in both urban and non-urban areas are not classified as low-income low-access (LILA\_Tracts\_1And10 = 0). Urban areas have roughly 1.7 times higher proportion of low-income low-access tracts than non-urban areas. The stable proportions suggest persistent structural challenges in food accessibility in both urban and non-urban settings.

To evaluate the relationship between median family income and status of low income and low food access, Table 3 and Figure 3 present the analysis results.

Table 3: Summary statistics of median family income by low-income and low food access status (2015 & 2019)

year	LILATracts_1And10	mean_median	sd_median
2015	0	70859.08	34794.70
2015	1	43846.72	13753.58
2019	0	81167.33	38124.42
2019	1	48981.59	14437.49

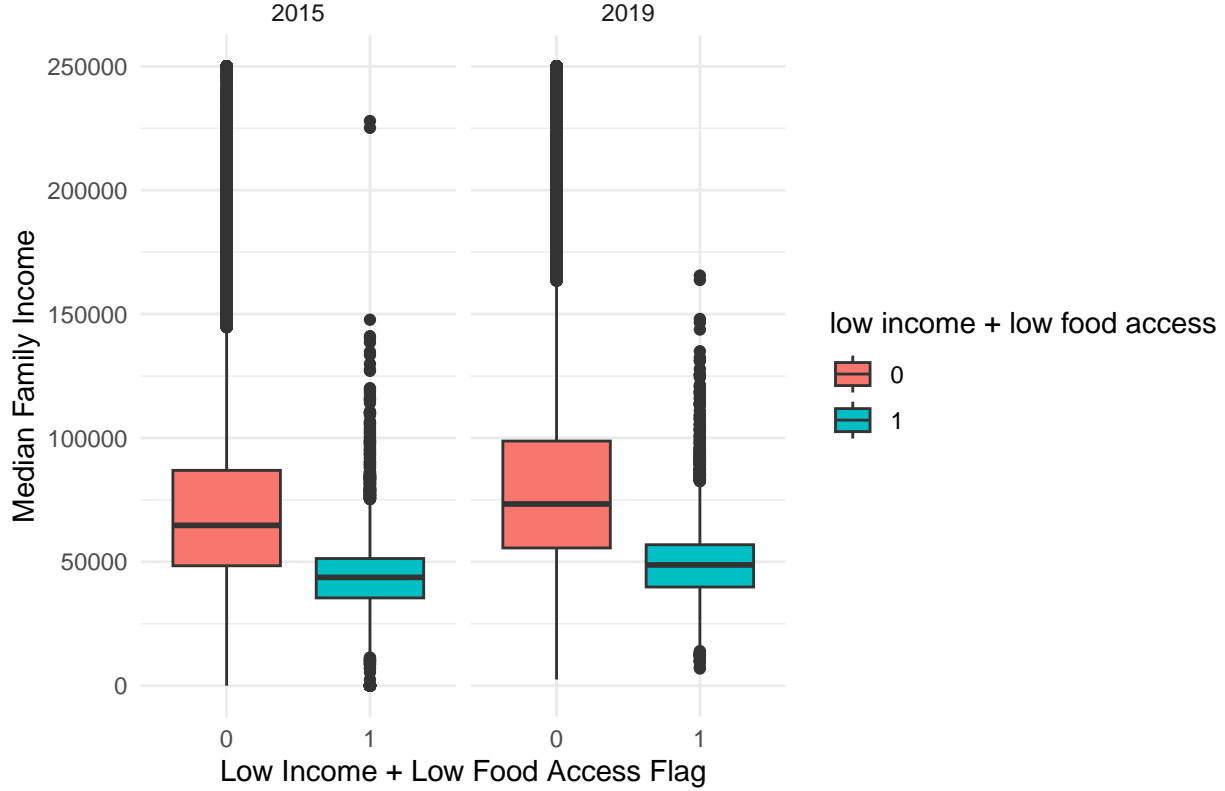


Figure 3: Median Family Income by Low-Income and Low Food Access Status (2015 & 2019)

Across both years, tracts classified as low-income and low-access exhibited substantially lower median family incomes than their counterparts. In 2015, non-flagged tracts had a mean median income of \$70,859 (SD = \$34,795), compared to \$43,847 (SD = \$13,754) in flagged tracts. In 2019, these averages increased to \$81,167 (SD = \$38,124) and \$48,982 (SD = \$14,437), respectively, indicating income growth over time for both groups. However, the absolute income gap widened slightly—from approximately \$27,000 in 2015 to over \$32,000 in 2019—suggesting that economic gains were not evenly distributed. The boxplots illustrate this disparity visually: the distribution of incomes in flagged tracts is consistently shifted downward, with narrower interquartile ranges and fewer high-income outliers compared to non-flagged tracts. Overall, the results demonstrate a persistent and widening socioeconomic divide between areas with and without low food access, underscoring the structural nature of income-related food accessibility challenges.