Data Science Clinic: Fall 2023
Internet Equity Initiative

In 2022, the Federal Communications Commission (FCC) initiated an expanded Broadband Data Collection project to document existing internet access across the United States. The project allows multiple stakeholders, including state governments, organizations, and individuals, to challenge reported availability based on geo-locations where broadband access should be available ('Broadband Serviceable Locations' (BSLs)). The data collected will help address concerns of data transparency around reported broadband availability, which is why conducting an analysis of the data through a geographic and demographic lens is important.

In order to establish a streamlined dataset to analyze trends, the Internet Equity Initiative team constructed a data pipeline combining FCC challenge resolutions and Census data. The pipeline matches challenge locations and resolution codes to geographic units and census data to explore geographic variation in challenge rejection rates and associations with socioeconomic and demographic characteristics.

The pipeline first downloads 'Fabric' resolution and Census ACS data. By establishing a common coordinate system between FCC and geographic boundary data, and FCC and Census data is then merged at the tract (Census county subdivision), county, and state level. Challenges are then grouped using resolution codes into accepted, rejected, duplicate or withdrawn categories, with aggregated counts for each category. The conclusion of the data pipeline creates a minimal database with all data aggregated for a particular geographic level and conducts preliminary analysis on challenge resolutions and state-by-state comparisons.

In total, **3,048,737** challenges have been filed. **Illinois, Alaska, and West Virginia** have the most total Fabric challenges filed (around 230,000-330,000). **The FCC only accepted around 8%** of all challenges, with the highest number of challenges accepted for West Virginia. Given this analysis, there is substantial potential to further examine challenge resolutions weighted by units (number of households corresponding to one BSL) as well as Census demographic factors that may be correlated with challenge resolutions.