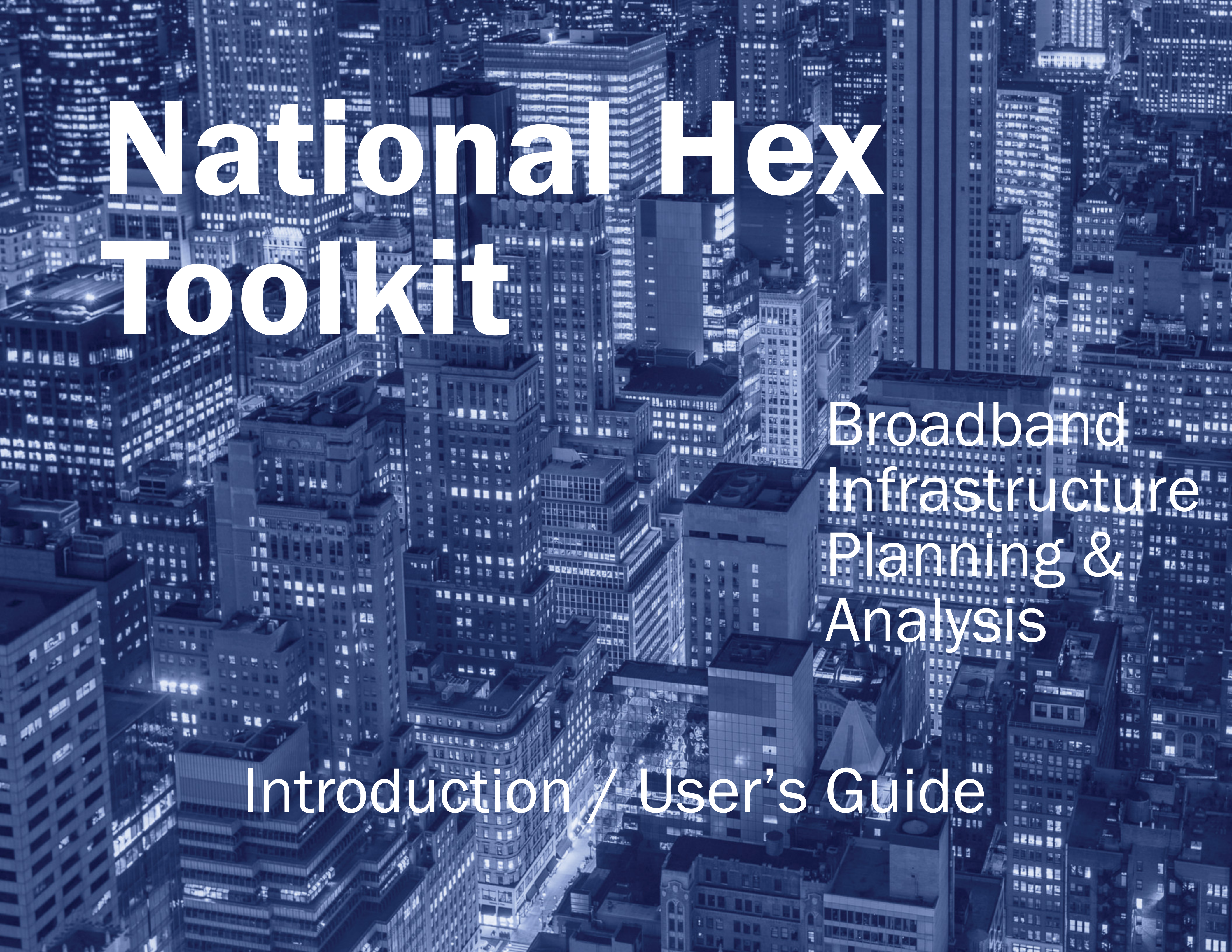
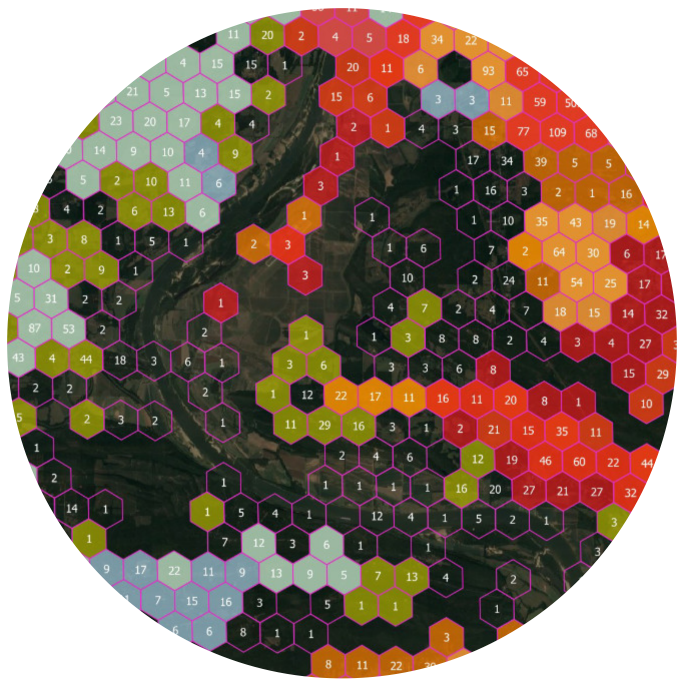


National Hex Toolkit

Broadband Infrastructure Planning & Analysis

Introduction / User's Guide



SIGNALS Analytics, LLC

Table of Contents

List of Figures

Introduction

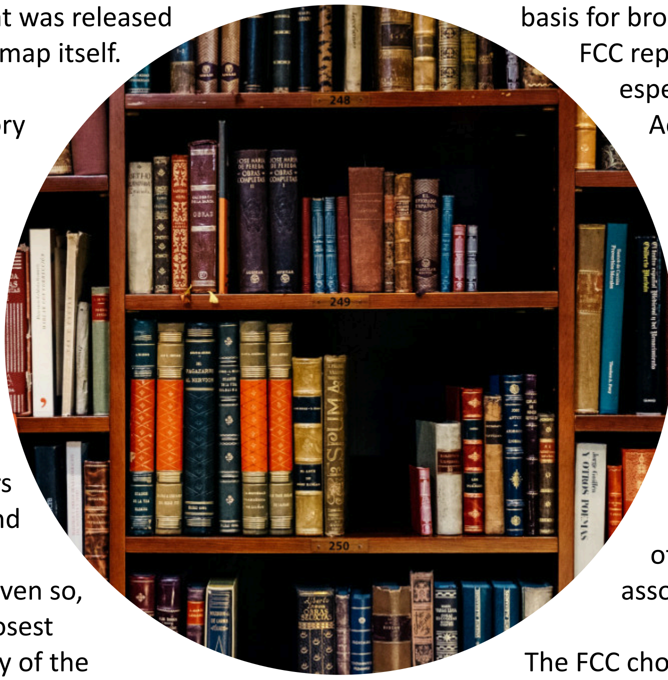
On November 18th 2022 the FCC announced a new map. Broadband providers and telecommunication industry media rushed to see it. Consumers read about it and many visited the site.

Those reporting on the FCC "map" in industry media may have missed the most important piece. The most newsworthy story - that analysts discovered only over time - was the data that was released *in conjunction with* the map. It was not the map itself.

It is hard to write an engaging front page story about a collection of CSV files or a webpage entitled "Data Download". It turns out that the most interesting piece is the data. The online map visualizes only a small portion of the newly collected data. Many of those in the industry who rushed to see the map left underwhelmed. Worse, many missed the riches in the adjoining tab.

The online map is novel. It allows consumers to locate their home and see what broadband providers are reporting coverage. It is configurable with multiple possible views. Even so, it retains an abstract quality at all but the closest zoom levels. Figure 1 contrasts the simplicity of the zoomed out online map to the richness of the data underneath.

This hidden data is described in the coming pages.



Legislative History

The new FCC map, its underlying "Fabric" of locations, and the associated Broadband Data Collection (BDC) process are a function of several legislative initiatives.

The Broadband DATA Act, signed into law on March 23rd, 2020, established the framework of a "Fabric" of locations acting as the basis for broadband reporting. Form 477 - the historic FCC reporting process - was notoriously inaccurate, especially in rural areas. The Broadband DATA Act was passed as health officials began to impose Covid-19 lockdowns and as broadband connectivity became exceptionally important to every segment of society.

The American Rescue Plan (ARP) and the Infrastructure Investment and Jobs Act (IIJA), signed into law on March 11th, 2021 and November 15th 2021, respectively, provided unprecedented levels of broadband funding and tied much of the infrastructure-related funds to metrics associated with the FCC Fabric and the BDC.

The FCC chose to release performance data using two different geometries. Each location is associated with a census block and with an H3 hexagon. Fixed broadband is reported using level 8 hexagons and mobile broadband is reported using level 9. Figure 2 shows the spectacular increase in precision in rural areas that is possible with level 8 hexagons, as compared to census blocks. Census blocks are widely used because they are familiar and convenient.

Figure 1: FCC Map (left) vs FCC Data, visualized by technology layer (right).

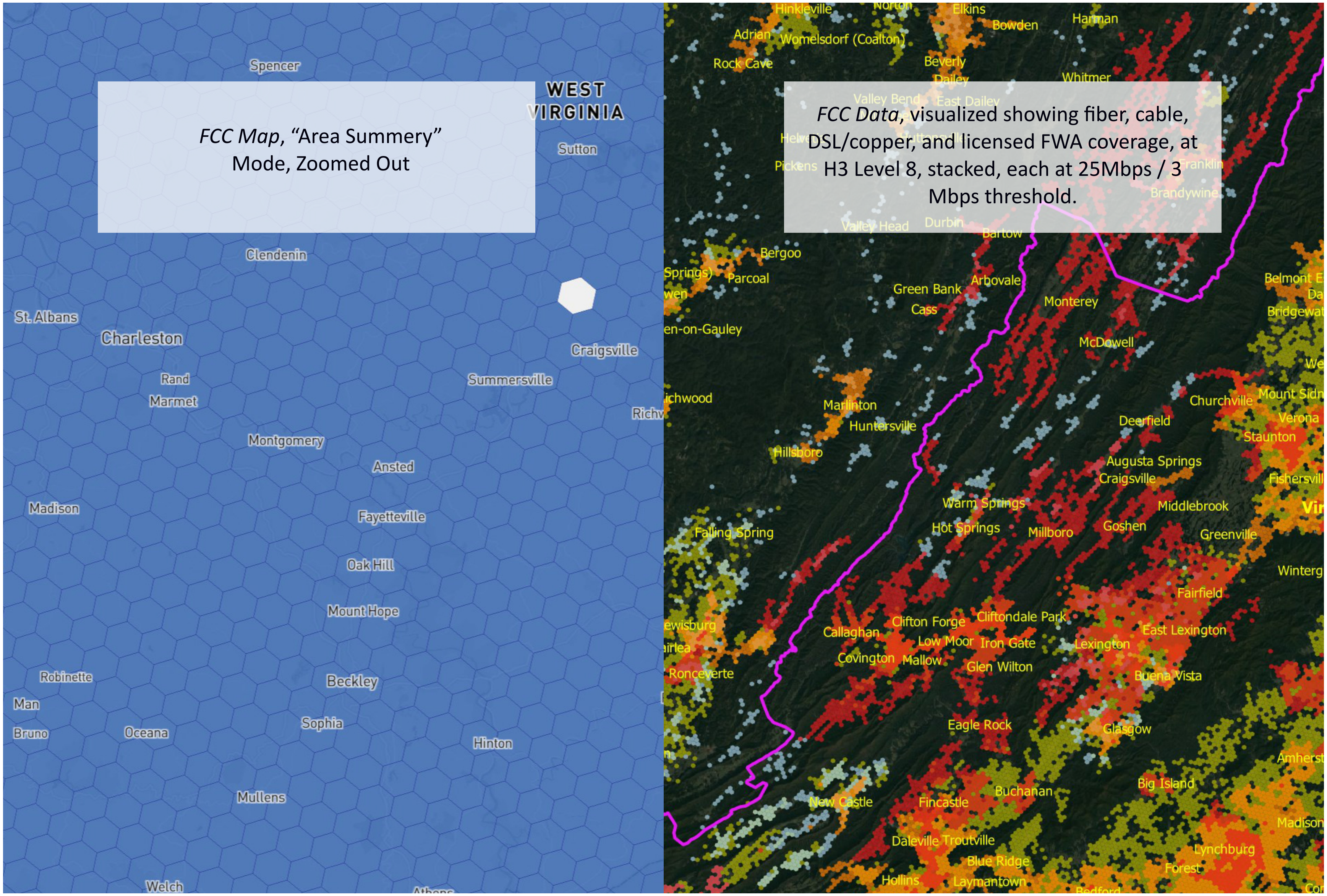
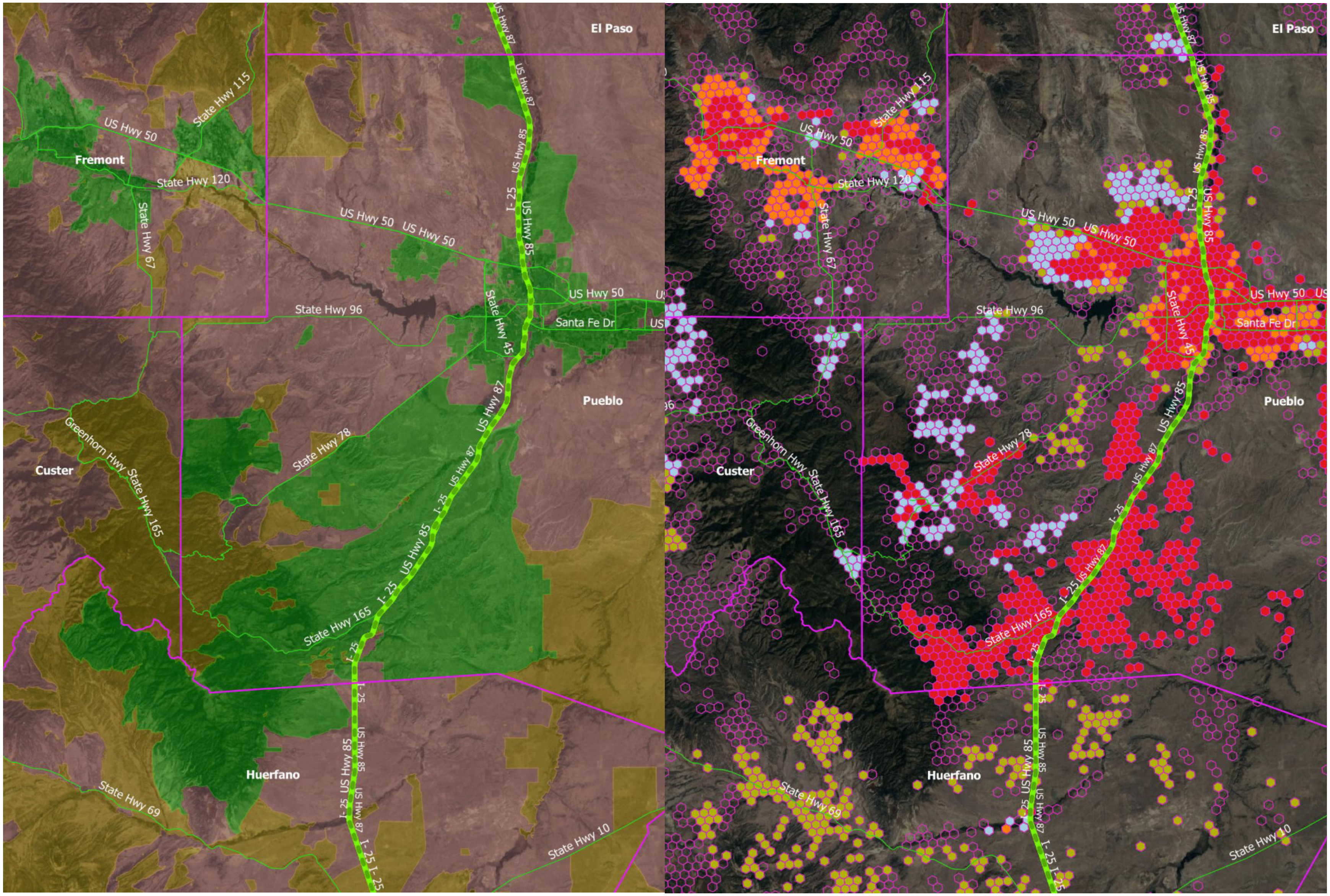


Figure 2: Precision in Rural Areas: Census Block (left) vs. Hex Level 8 (right)



Census blocks combine to form block groups that combine to form tracts that combine to form counties. The hierarchical nature of census blocks and their relationship to political subdivisions (counties and states), is described in Figure 18. Most legacy broadband maps continue to use census blocks. Even so, hexagons offer a huge advantage for precise broadband planning.

The largest piece of infrastructure funding, $42.45 billion, is associated with the Broadband Equity, Access, and Deployment Program (BEAD) that is part of the IIJA and administered by the NTIA.

On May 13th, 2022 the NTIA released a 98-page BEAD Notice of Funding Opportunity (NOFO), laying out the details of the grant program and inviting eligible entities (states and territories) to participate.

The NOFO distinguishes between locations served with fully qualified broadband (100 Mbps down, 20 Mbps up, and latency not more than 100 msec), locations that are underserved (equal or greater than 25 Mbps down, 3 Mbps Up, latency not more than 100 msec, and not fully qualified), and locations that are unserved (any other level of service or no broadband).

It further identifies four "reliable" technologies: fiber, cable, DSL/ copper, and licensed terrestrial wireless (a.k.a. licensed FWA).

The National Hex Toolkit, accordingly, tracks the presence of these four technologies and tracks the availability of service at served/ qualified, underserved, and unserved thresholds based on these four technologies.

Presence means that the technology / speed is available in at least one location in the hexagon (a concept similar to what was used in Form 477, except with much smaller geographies and therefore vastly improved precision). There is the potential for overstating coverage, but only slightly, and primarily in transition areas. In general, if a provider deploys fiber or cable or another technology it will offer it to every home in a neighborhood. The exception is in a "transition" hexagon, where only a portion of the area is actually covered.

Other technologies exist (unlicensed FWA, geosynchronous orbit satellite, and non-geosynchronous orbit satellite) but they are excluded from BEAD funding.

Since BEAD focuses on four technologies and since most of the currently available funding falls under BEAD the Toolkit focuses on the four allowed technologies (which may include many competing providers). An area that doesn't have qualifying coverage using a BEAD technology is at risk of being overbuilt.

Version 1.1 of the Toolkit now includes statistical "depth of coverage" calculations relative to the 100-20 Mbps / low-latency performance threshold. In an area with some qualifying coverage it shows the percentage of locations within each hexagon that meet the performance threshold. It further color-codes the hexagon (purple = low qualification rate to green = high qualification rate) to help the user visualize the ratio while displaying some other metric as text. Finally, it has a binary BEAD flag that turns on if 80% of the locations have qualifying coverage. This enables the user to view a county or a state and quickly see which areas will be fundable under BEAD or not, based on existing coverage. Figures 7, 8, and 9 show examples of this new functionality.

Version 1.1 also has greatly improved performance and a data export capability that makes it easy to select an area then export most of the relevant data to a spreadsheet or to another format.

Figure 3: Areas with 25 Mbps / 3 Mbps / low latency coverage, or better, show as level 8 hexagons (green)

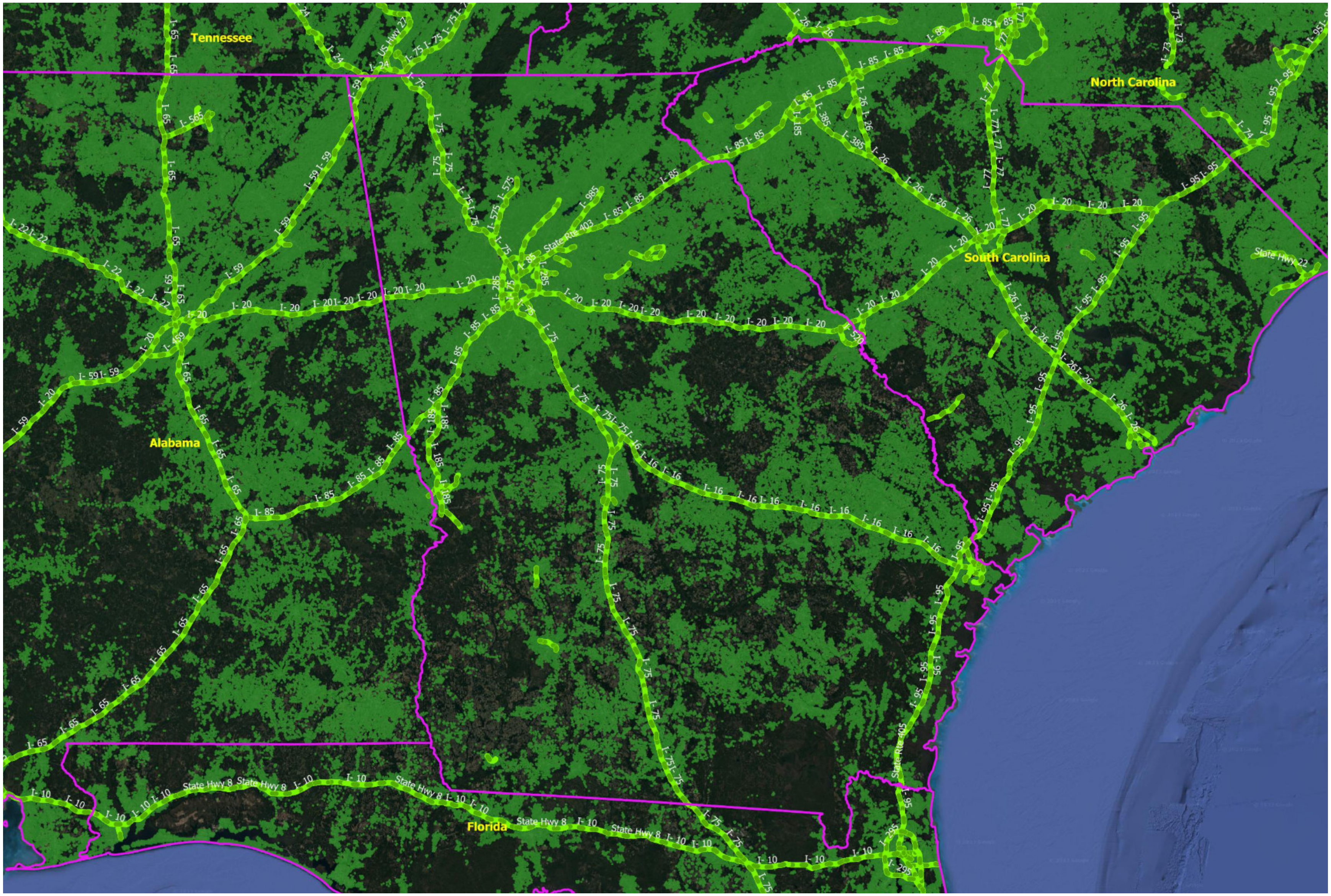


Figure 4: Level 8 hexagons, >=25/3, colored by technology and stacked: fiber (red), cable (orange), DSL/copper (light blue) and licensed FWA (olive)

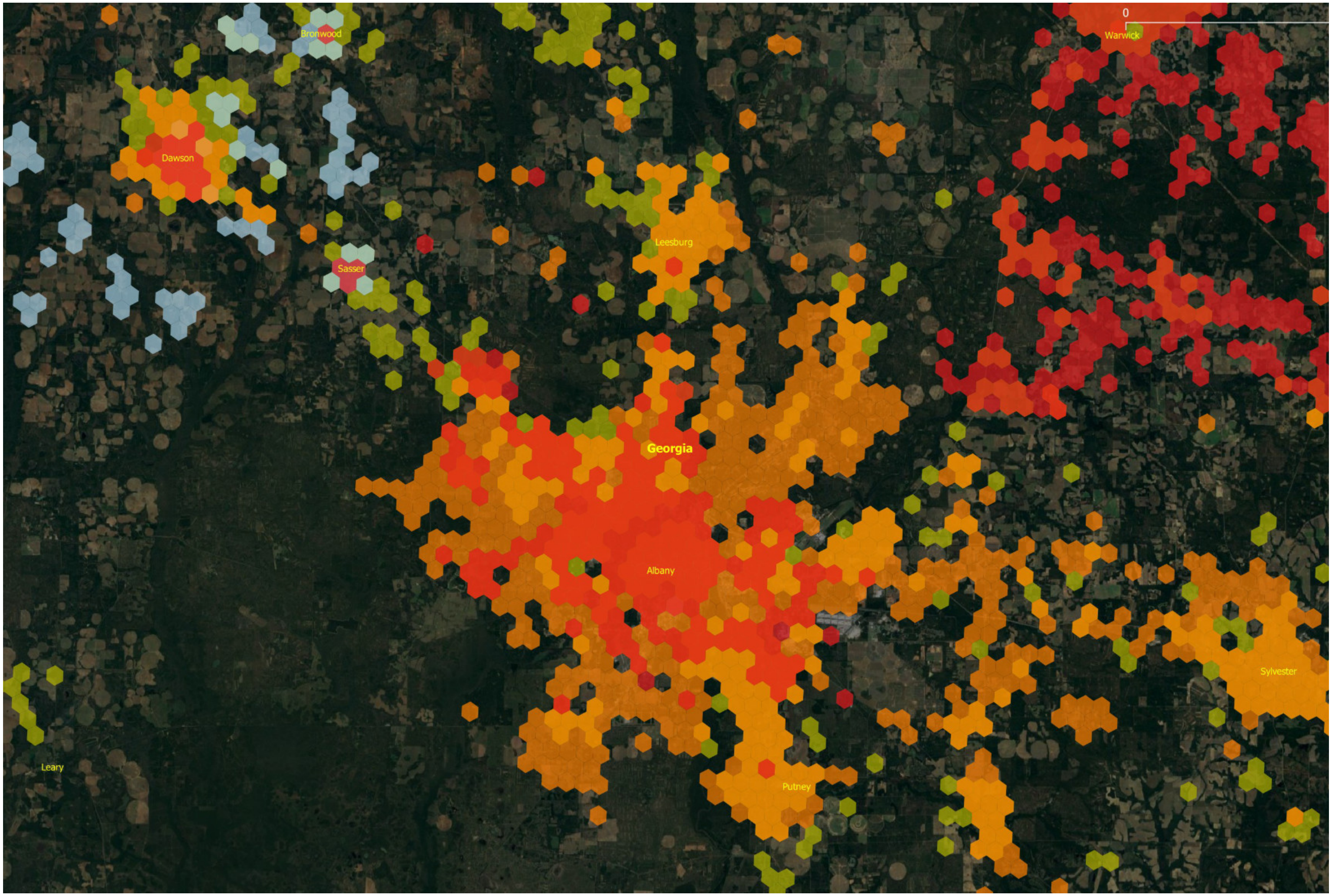
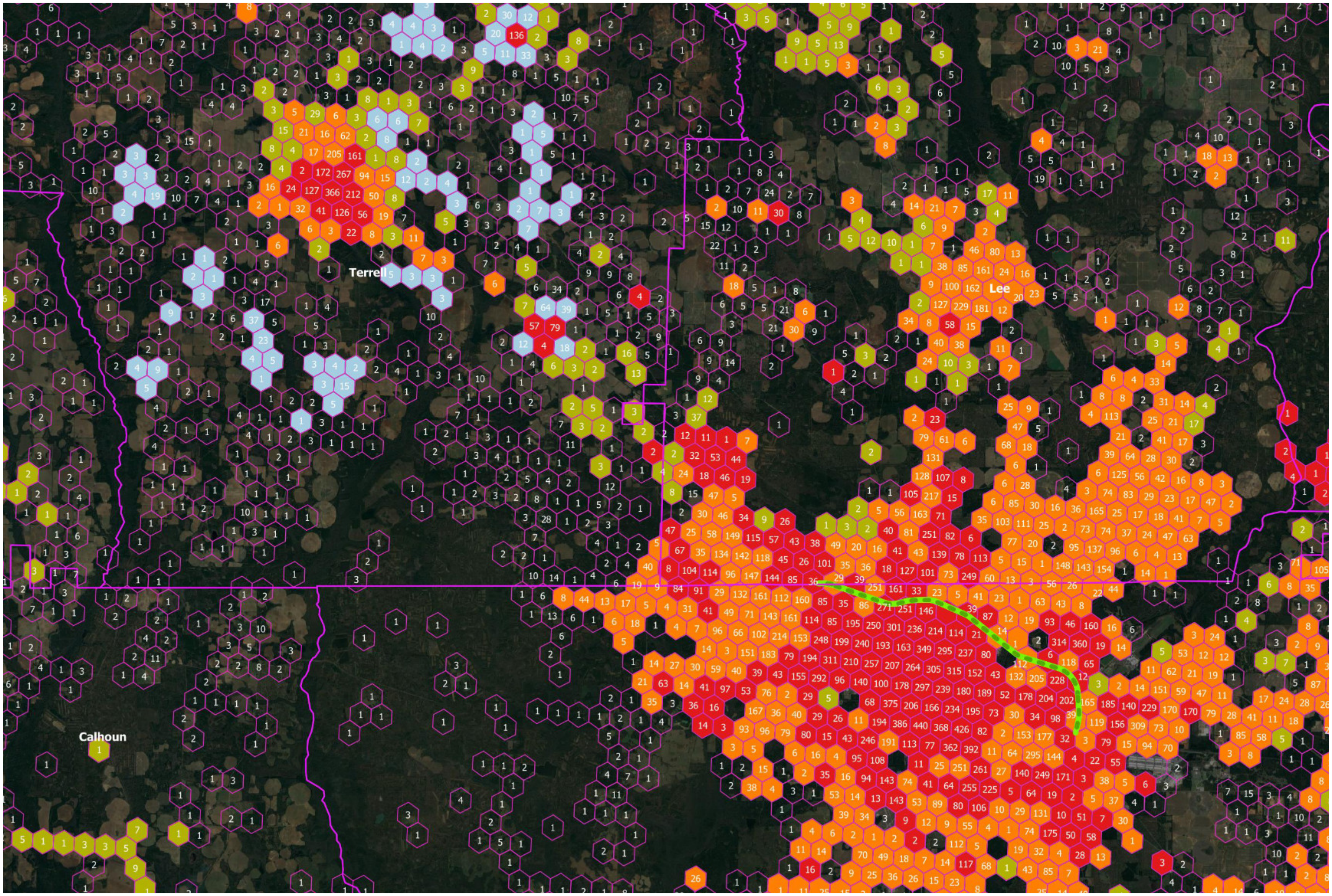


Figure 5: Similar to Figure 4, with hexagon boundaries showing populated areas and numbers showing the total number of locations



Other toolkits include other functionality. Non-BEAD technologies are allowed under other funding programs. As an example, ARP/CPF funds are subject to fewer federal restrictions than IIJA/BEAD funds.

The National Hex Toolkit and the Infrastructure Essentials BEAD Toolkit both includes statistical data on a county and place basis (see Figure 8). These visualizations lack the precision (by a large margin) of a level 8 hexagon.

The BEAD NOFO further restricts funding in areas that have received federal funding that are subject to an "enforceable commitment". Accordingly, the Toolkit tracks several federal programs: the Broadband Infrastructure Program (BIP), RDOF, CAF II, and several related USDA programs, notably ReConnect.

The Opportunity

The magnitude of investment associated with BEAD ($42.45 billion) makes it a once-in-a-generation opportunity. It follows a large number of previous programs (CAF, CAF II, RDOF, ReConnect, etc.) using grants, loans, reverse auctions and other tools to fund infrastructure deployment in challenging high cost areas.

Equally important, BEAD ups the ante by defining "qualifying broadband" as reliable service offering 100 Mbps down / 20 Mbps up, with no more than 100 msecs of latency.

The map of the United States changes with this new definition of broadband. One does not need to look far to find geographies that lack qualified broadband.

Nation at a Glance

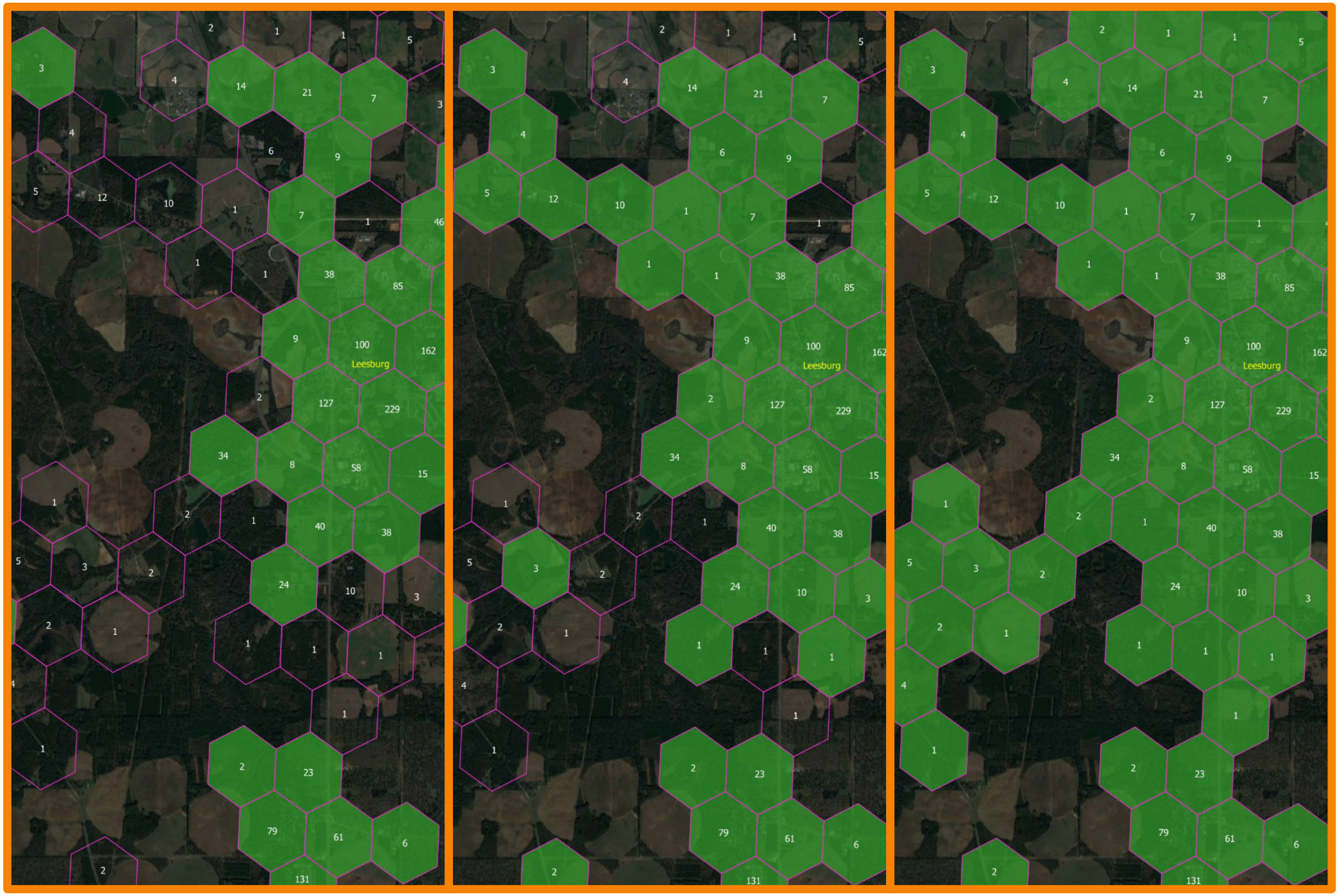
Figures 3, 4, and 5 show show the range of visualizations possible with hexagonal data. Figure 3 shows a multi-state area. Hexagons with coverage at 25-3 and low latency, are shown in green. Everything else is bare (a darkened earth). One can quickly see covered and uncovered areas. One can easily toggle the threshold to 100-20, showing qualified coverage, but over a smaller area. Finally one can toggle it "any performance" to see populated areas with broadband offerings that may or may not represent effective service.

Figure 4 zooms into a small area. It shows four layers of technology (cable, fiber, DSL/copper, and licensed FWA) that can be turned on and off individually. Each is shown at a specified performance threshold (or greater): any performance, 25-3, or 100-20.

Figure 5 is similar to figure 4 except that all populated areas are shown in magenta hexagons and each hexagon includes a label representing the number of locations. The view is powerful because it allows the user to clearly distinguish between populated and unpopulated areas and to identify populated areas that have no coverage at the specified threshold. It is possible for an ISP to quickly see what areas are uncovered at a certain speed or with a certain technology and to count the locations passed in the covered and uncovered areas. Finally it is possible to change locations into units (a close approximation of housing units, except that Fabric "units" also includes non-enterprise businesses). If the view is zoomed out the labels start to overlap.

Figure 6 shows performance (vs. technology) as simple green colors at three different thresholds. One could do something similar using a single technology. Licensed FWA typically covers a large area at low performance thresholds and a smaller area at high performance

Figure 6: Performance: 100 Mbps - 20 Mbps (left) vs. 25 Mbps - 3 Mbps (center) vs. Any Performance (right), with locations as labels.



thresholds. Fiber, in contrast, if its present, shows relatively little variation as a result of the performance threshold.

Figure 7 shows the percentage of locations with qualifying (100-20, low-latency) broadband (a new feature in version 1.1). Each hexagon is also color coded (purple to green) to reflect the ratio.

Figure 8 is similar to Figure 7 except that the percentage value is replaced with another metric, the number of locations per hexagon. The color still reflects the percentage of locations with qualifying coverage.

Figure 9 is similar to Figures 7 and 8 except that the labels have been turn off. It includes a binary (yes/on) indicator that goes green when the percentage of qualifying locations meets or exceeds 80%. At 80% an area ceases to qualify for BEAD, according to the NOFO. This is a very effective tool to be able to see fundable and unfundable areas - at least based on existing qualifying broadband.

After the release of the FCC Fabric, v1, states had an opportunity to identify missing locations (location challenge, type 1). There were 178,009 successful challenges. The missing locations were added to the Fabric v2, which was released to licensees on 1/3/2023. The Fabric v2 is in the hands of ISPs, who are associating their coverage as of 12/31/2022 with each v2 location. The results will be released mid-2023. The points shown in Figure 10 show the missing locations. The data include addresses and associate each point with a census block and with a level 8 hexagon. The image suggests that states expended different levels of effort and engaged different strategies in identifying missing locations.

Figure 11 shows coverage statistics over large areas (relative to the size of a level 8 hexagon) such as counties and places.

Figure 12 shows the result of the most impactful federal subsidy programs: BIP, RDOF, CAF II, and the USDA initiatives (most importantly ReConnect). In this case only qualifying (100-20 Mbps and low latency) RDOF and CAF II areas are visualized in green. This section of the toolkit enables the user to change assumptions and include or exclude the corresponding areas. One can see areas that have been challenged or rescinded, as an example.

Figures 13 and 14 show socio-economic metrics:

- Figure 13 shows median household income per block group. Economic metrics are helpful in demonstrating need.

- Figure 14 shows the percentage of households receiving SNAP (a.k.a. food stamp) benefits. A SNAP recipient also qualifies for a monthly broadband subsidy under the Affordable Connectivity Program (ACP).

Industry data and basic demographics also highlight opportunities:

- Figure 15 shows fully qualifying tracts in green, based on Ookla data published by the NTIA. It also shows the number of housing units at the block group level. The magenta lines, represent block groups, that add up to tracts.

- Figure 16 is similar except that housing units are replaced with the median downlink and uplink speeds measured by Ookla. The numbers and the boundaries are shown at a tract level , consistent with the data.

- Figure 17 shows the lowest priced broadband subscription by Zip Code. It is helpful in that it highlights the relationship between availability and competition and affordability.

Figure 7: Depth of Coverage. Percentage of Locations with Qualifying Coverage, as Indicated by Shading (purple to green) and Labels

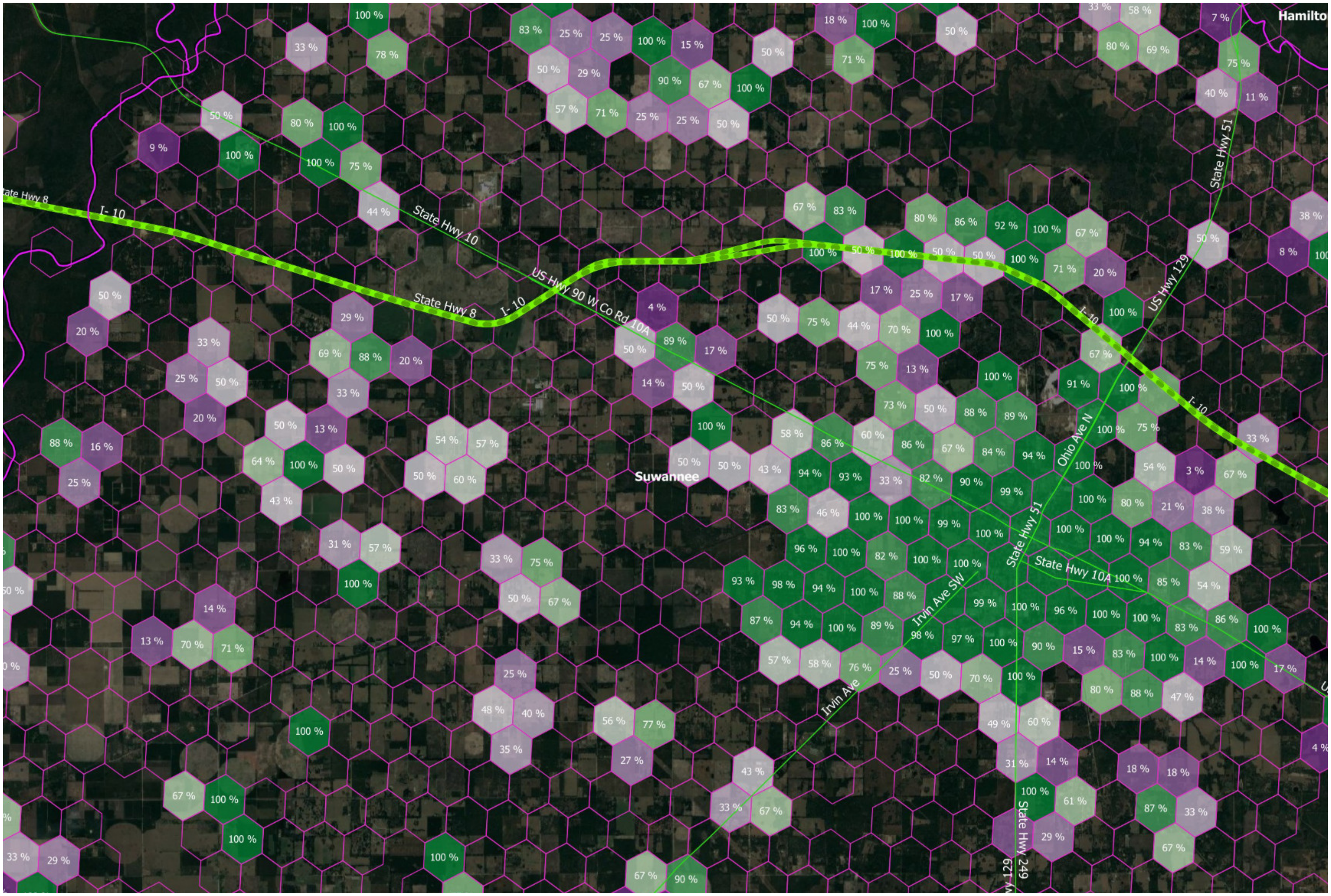


Figure 8: Percentage of Locations with Qualifying Coverage Shown by Shading (purple to green). Number of Locations Shown by Labels.

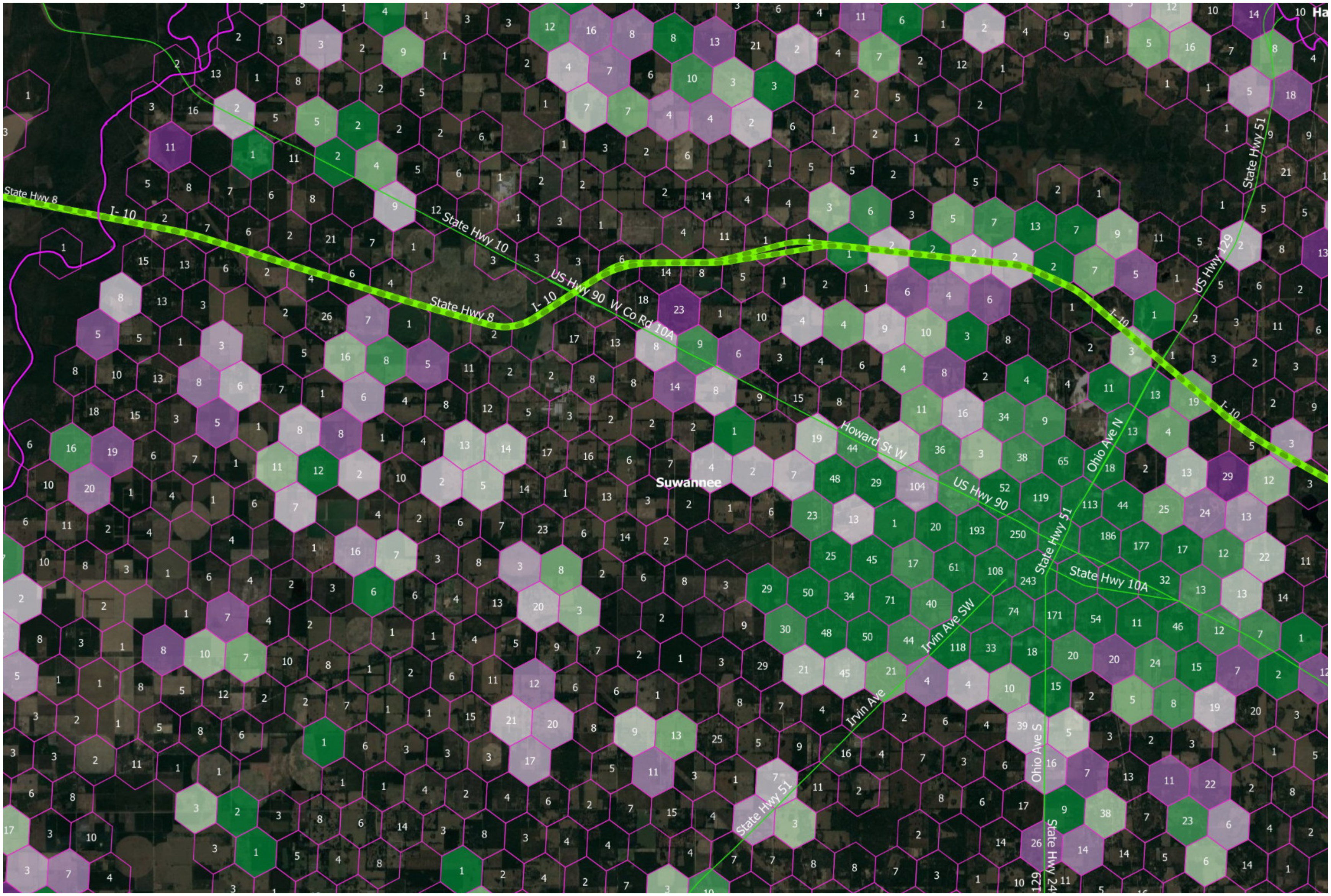


Figure 9: Binary Indicator that 80% of Locations (or more) Have Qualifying Coverage. No Labels.

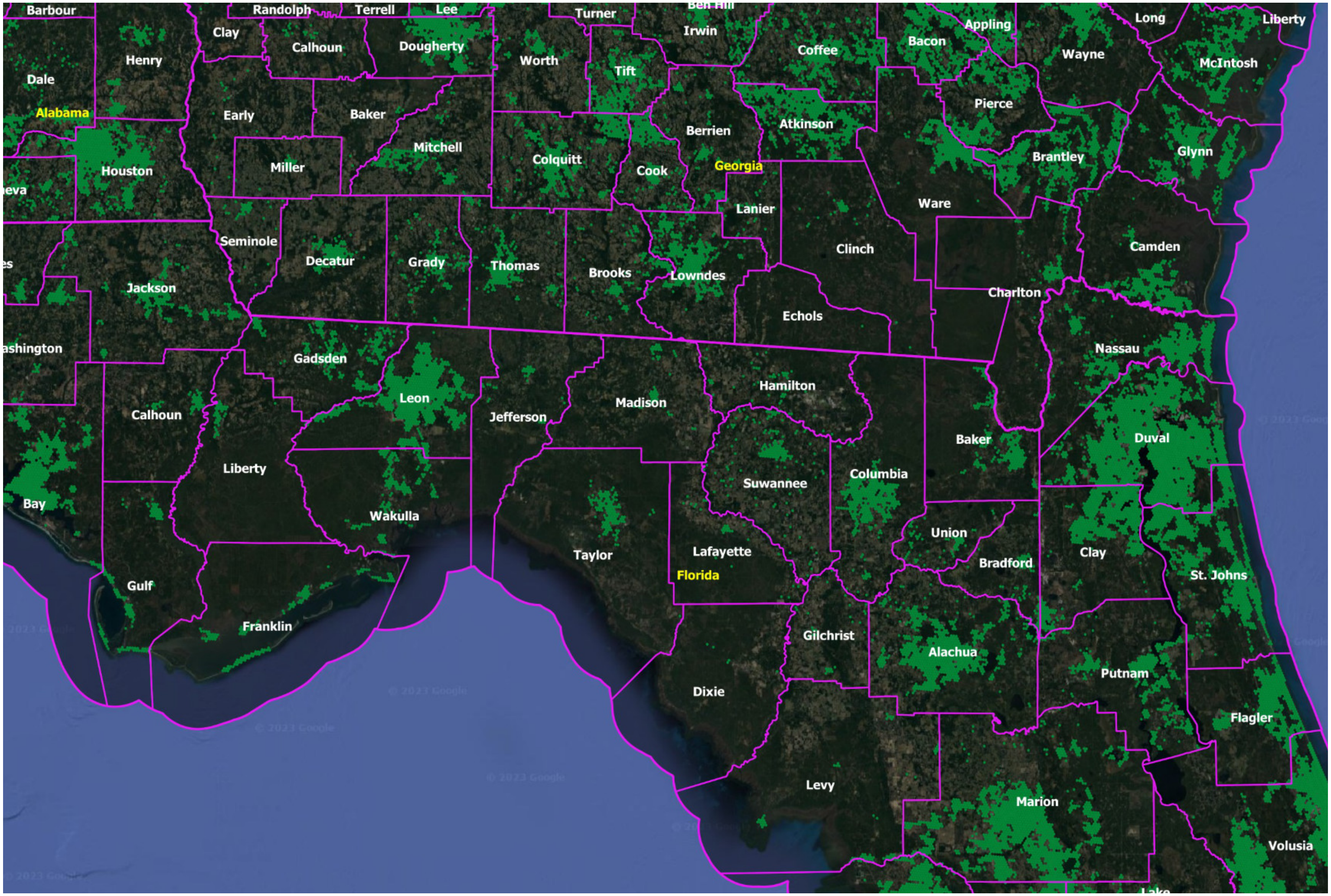


Figure 10: 178,009 Successful Location Challenges (Missing Locations, Type 1) to Fabric v1

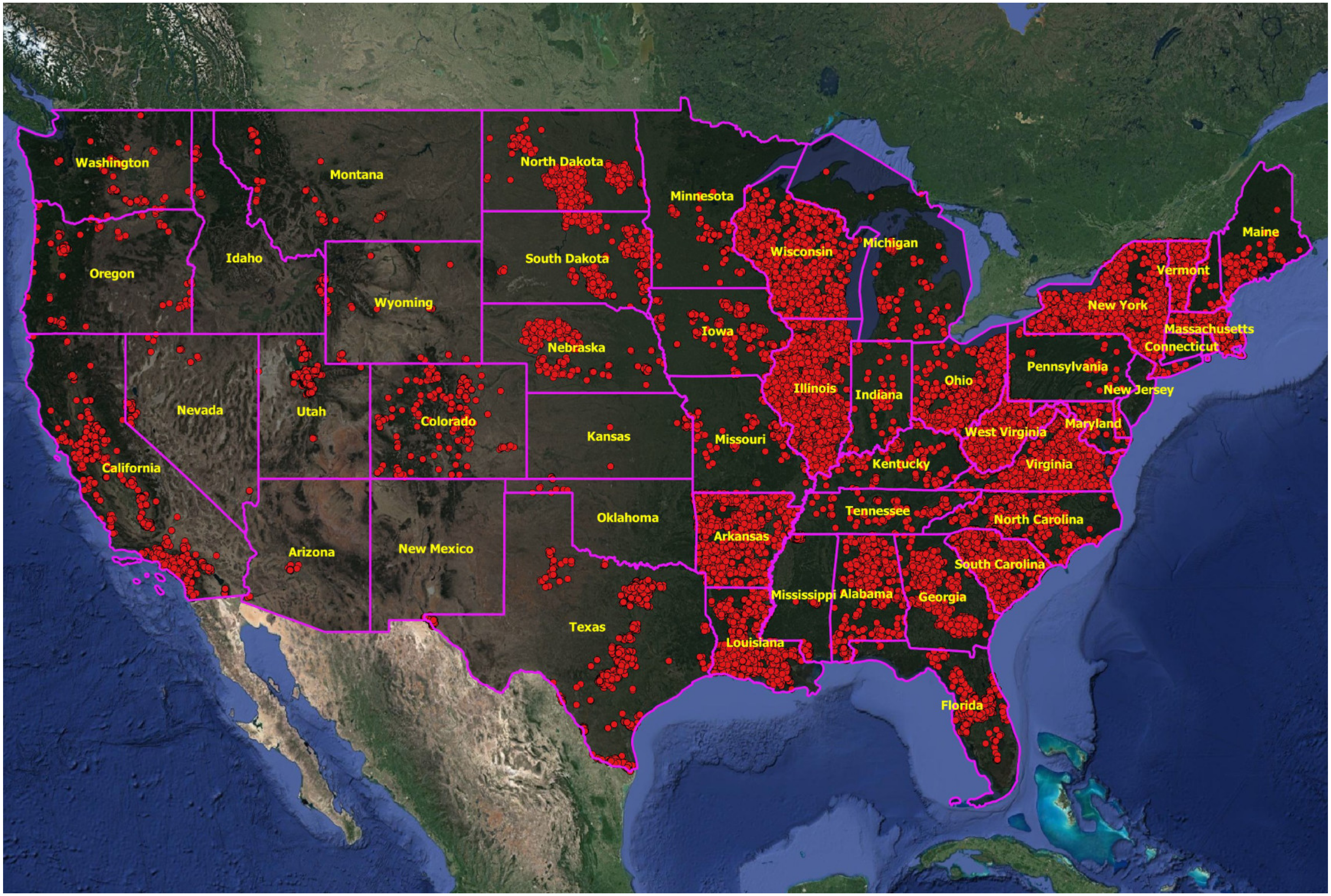


Figure 11: Other Fabric Visualizations. Top: % of locations, 100-20, Bottom Left: % 100-20 Places, Bottom Right: Rooftop Points

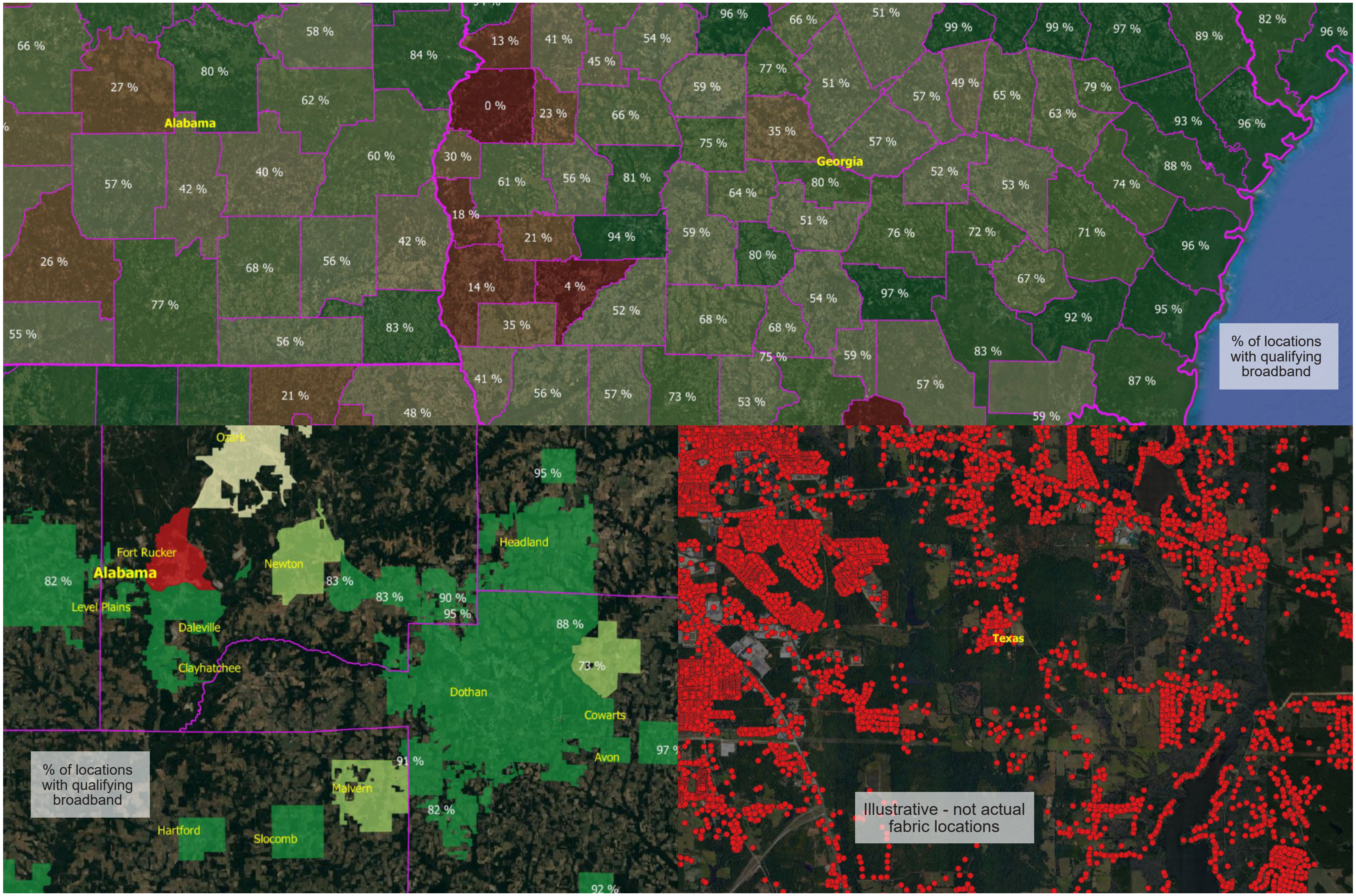


Figure 12: Federal Funding for Broadband via BIP, RDOF, CAF II, and USDA Programs

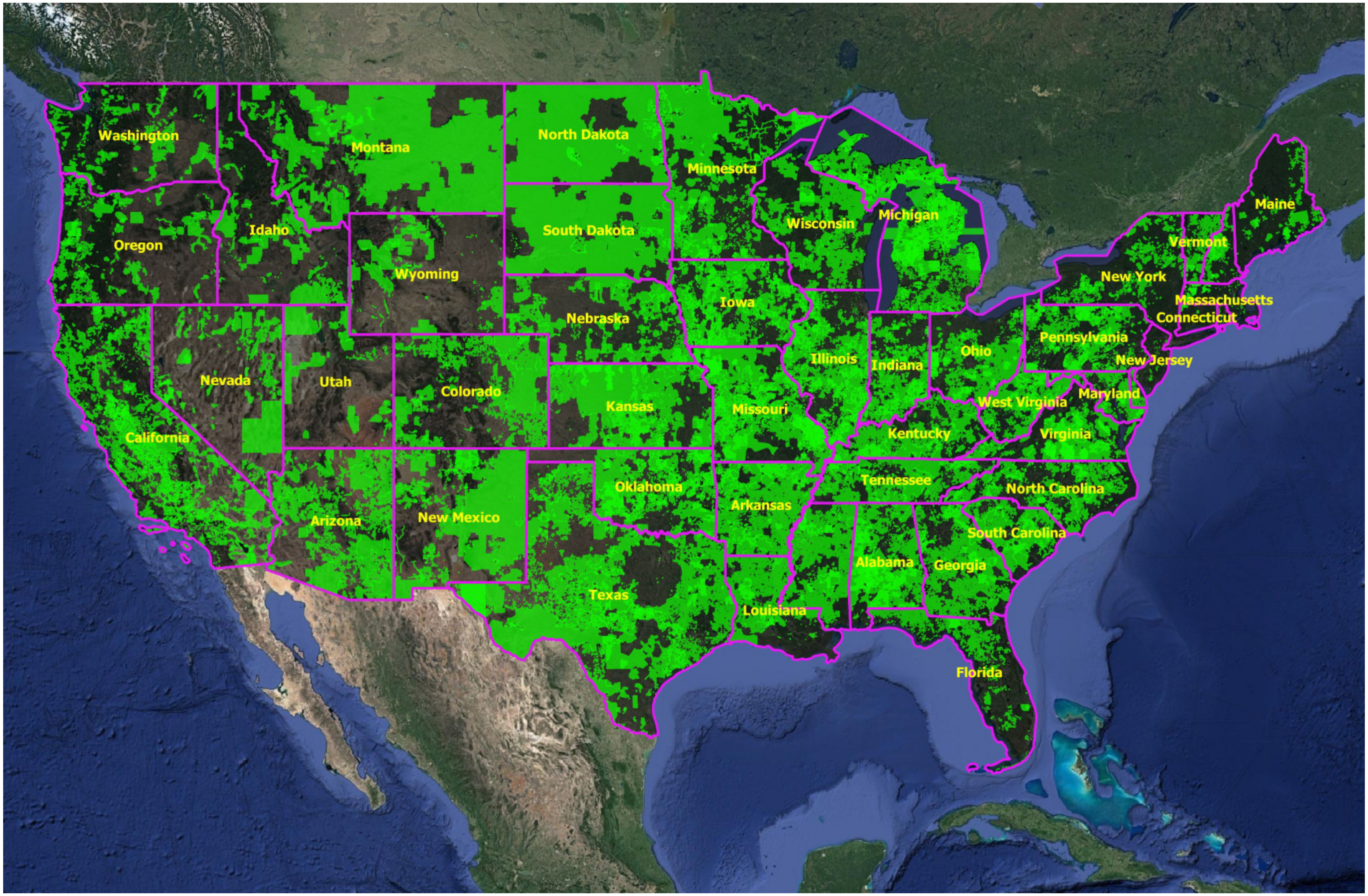


Figure 13: Median Household Income by Block Group

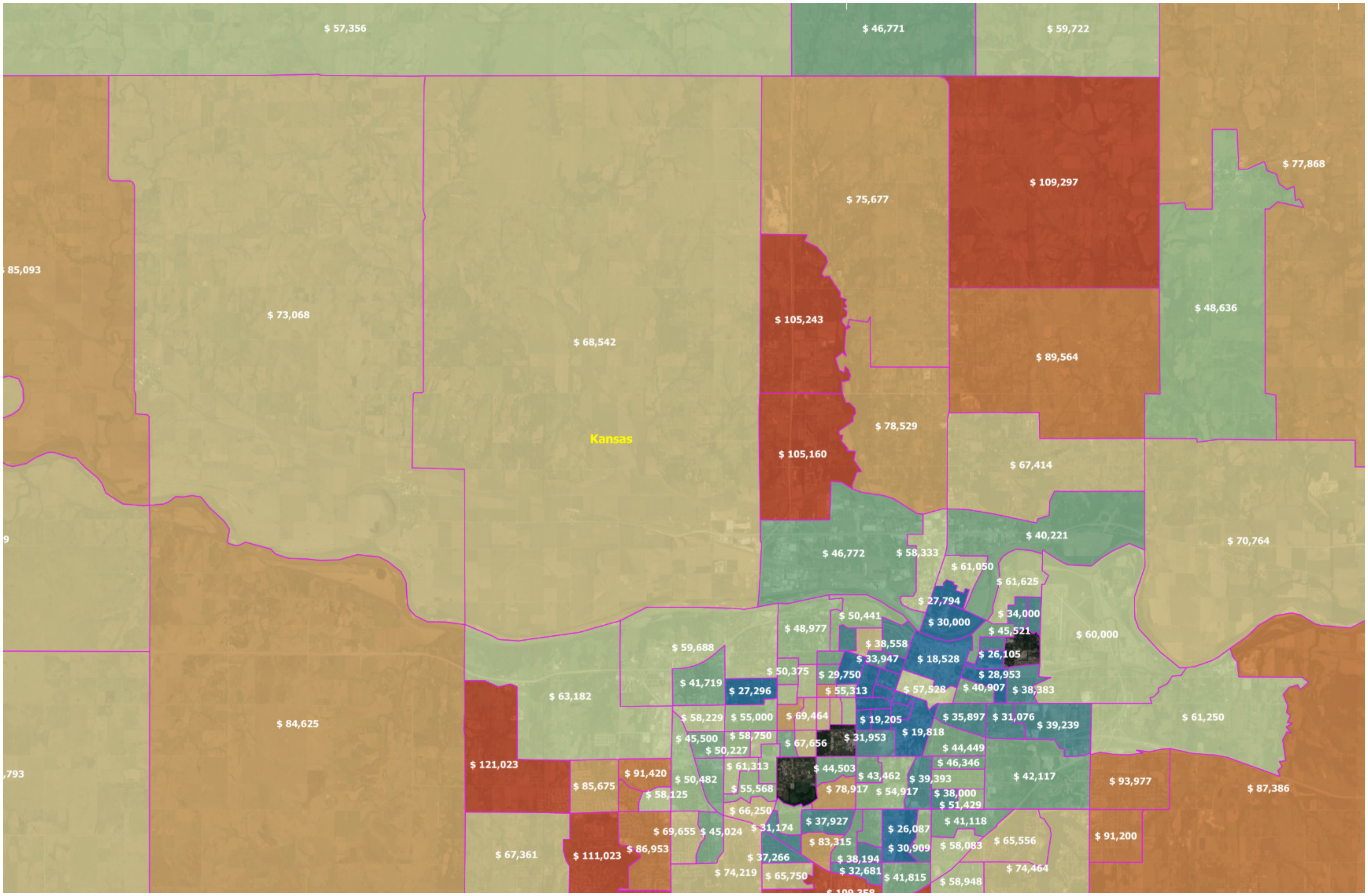


Figure 14: Food Stamp Participate Rate (% of Households) by Block Group

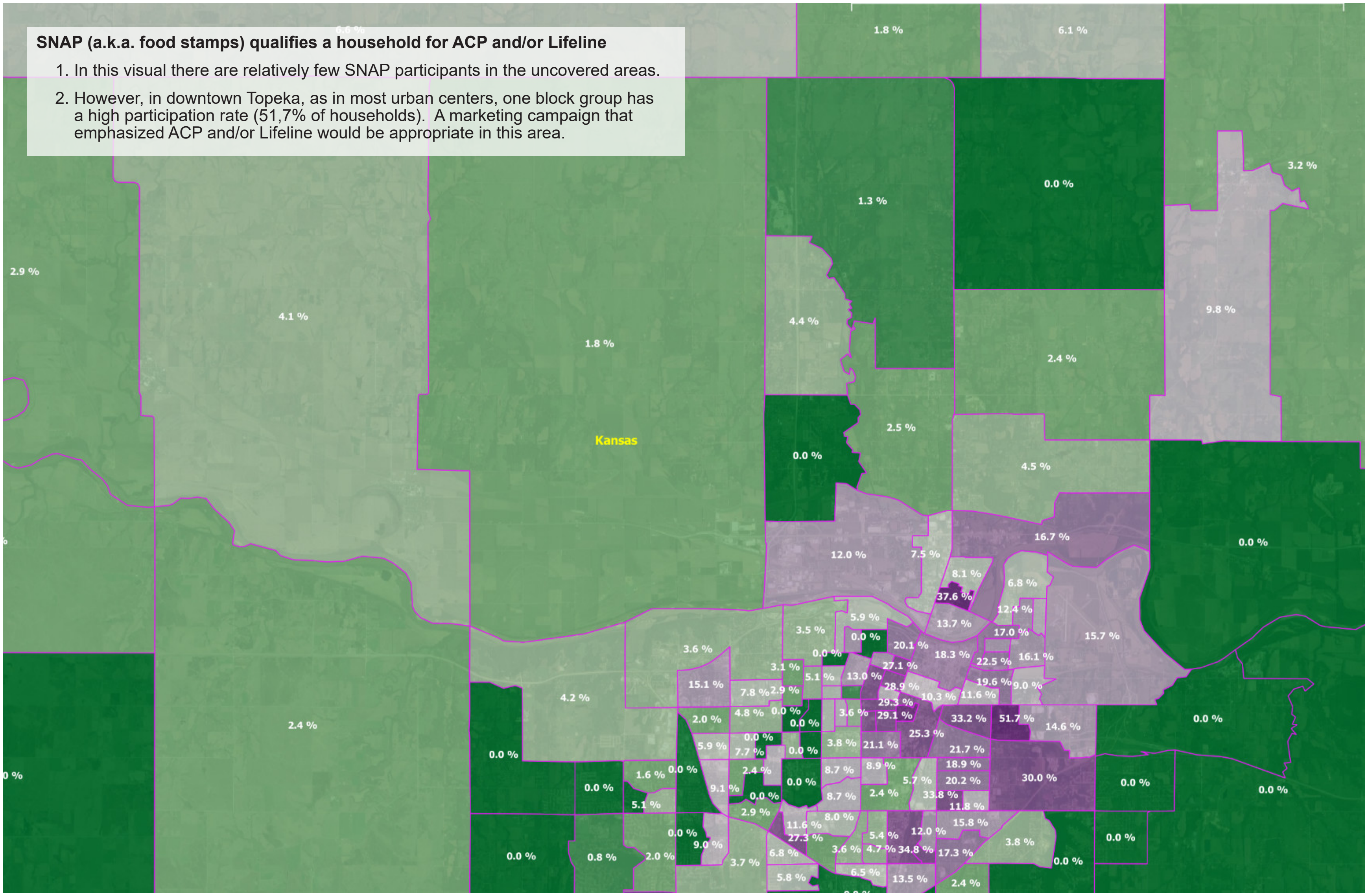


Figure 18 shows Community Anchor Institutions (CAIs). These are red dots in this figure. There are seven different categories. They can be displayed in aggregate or by category, with or without labels.

Measuring Performance

There are a number of factors that greatly influence performance and impact reported speeds. These include:

- *Advertised vs. Measured*. When an ISP reports its performance to the FCC it is asked to describe the "maximum advertised speed". This number of often greater than the average (mean or median) speed delivered. In addition, most ISPs offer a range of plans and many subscribers do not purchase the most expensive plan. Thus, there is often a difference, between the maximum advertised speed and the speed delivered to the average subscriber.

- *A Partially Covered Census Block*. Historically when an ISP reported its performance to the FCC on Form 477 it identified each census block it serves in whole or in part. If it had a single subscriber then that census block was categorized as served. The resulting coverage map – in the eyes of most consumers – overstated coverage. The FCC Fabric does a lot to address this concern.

- *Mean vs. Median*. The "average" speed can be measured many different ways. In most networks performance statistics are skewed to the right. In other words there are likely to be a few subscribers with very high speeds and a large number of subscribers with lower speeds. The median speed (the number above which half of the data points fall and below which half of the data points fall) is generally consider a more meaningful metric than the mean, just as median household income is a better measure of spending power than mean income. We find with performance data some sets of measurements that report mean speeds and others that report median speeds. Means speeds in most cases will be significantly higher than median speeds. Both measurements may be technically correct but the resulting numbers are different. Additionally, one could consider time-of-day differences or use a cumulative distribution function. A service level agreement that promises to deliver a certain speed 99.9% of the time is far more stringent than a promise to deliver the same speed "on average".

- *Geographic Area*. The size of the geographic area considered greatly influences the numerical result and how meaningful it might be. In a large geographic area (especially one that includes both urban or dense suburban geography and rural geography) the urbanized portion is likely to be better served than the rural portion. A single performance number is likely to overstate rural quality of service. Also, since household densities are higher in urban areas the distribution of measurements is likely to disproportionately reflect urban households. The solution is to characterize network performance in as small a geographic area as possible. Blocks are better than block groups which are better than tracts which are better than counties. Level 8 hexagons are the best, especially in rural areas.

- *Modem Speed vs. Wi-Fi Speed*. If one measures performance at the modem the result is likely to be more favorably than the same measurement on a device connected via Wi-Fi. Wi-Fi networks, while convenient, are notoriously unreliable. Like any radio access technology they are subject to signal strength degradation (due to obstructions, distance, and

Figure 15: Housing Units by Block Group and Ookla Measured Speeds (Green is >= 100-20) by Tract

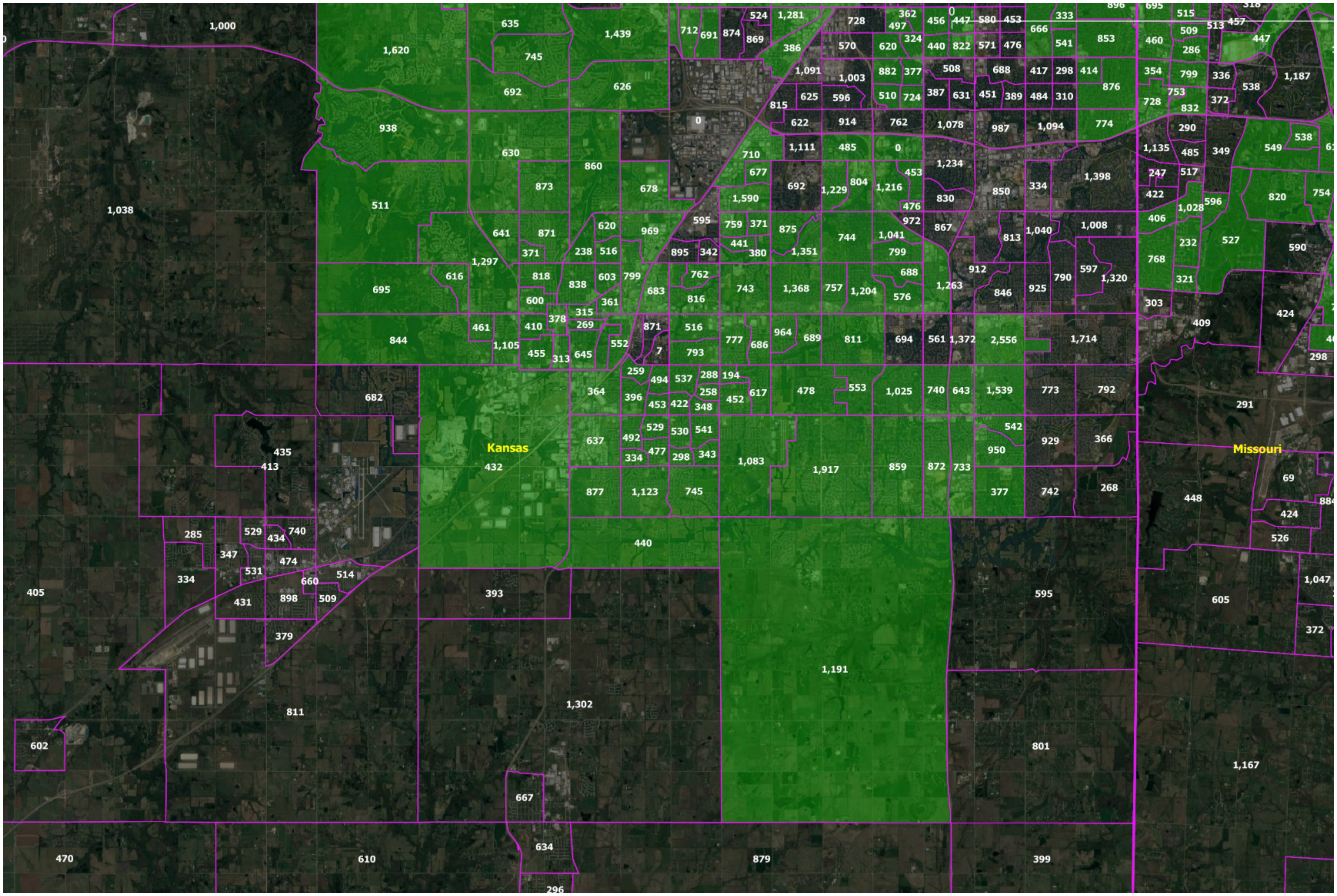


Figure 16: Median Speeds (Down-Up, Ookla) by Census Tract

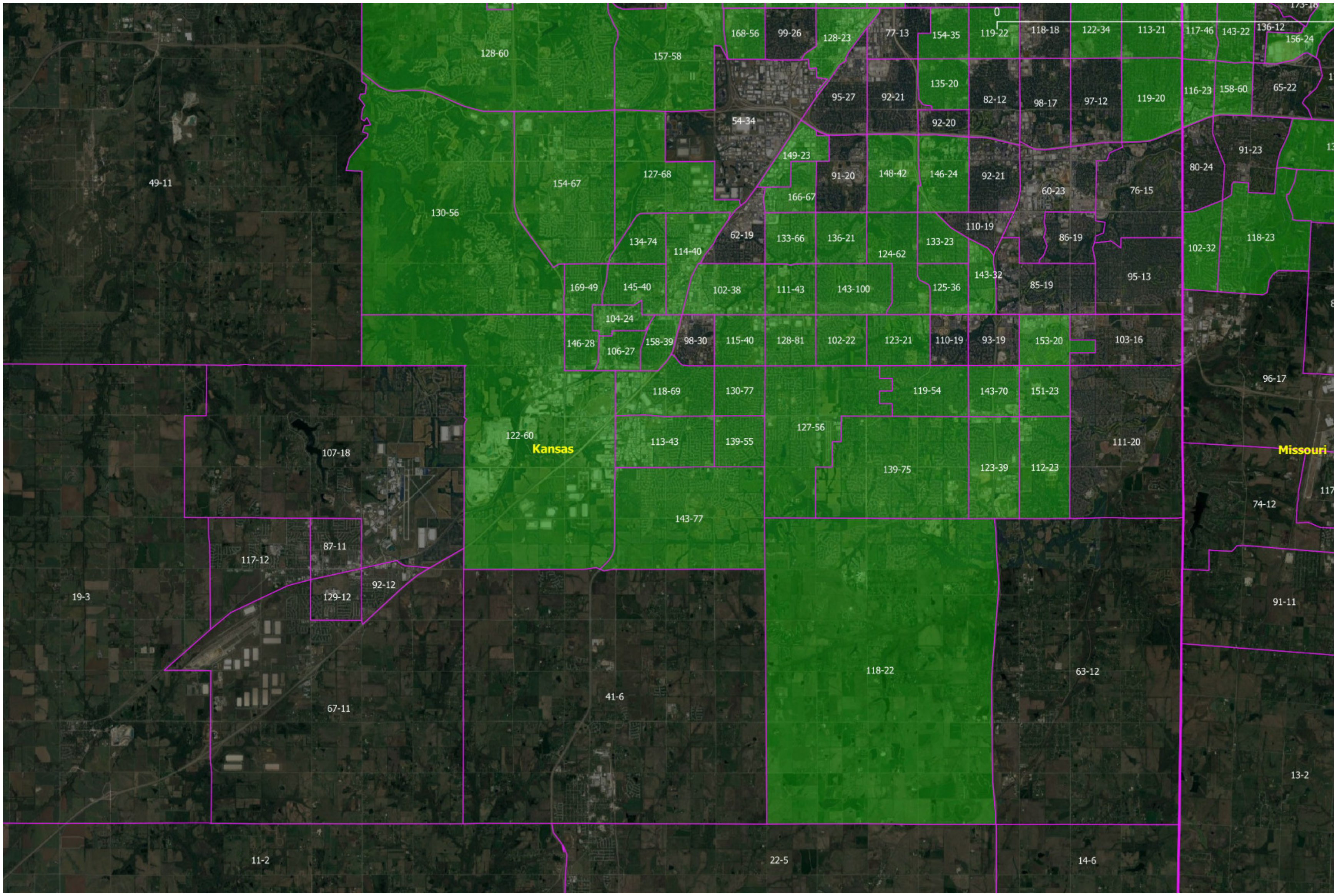


Figure 17: Minimum Monthly Cost of Broadband by Zip Code (BroadbandNow)

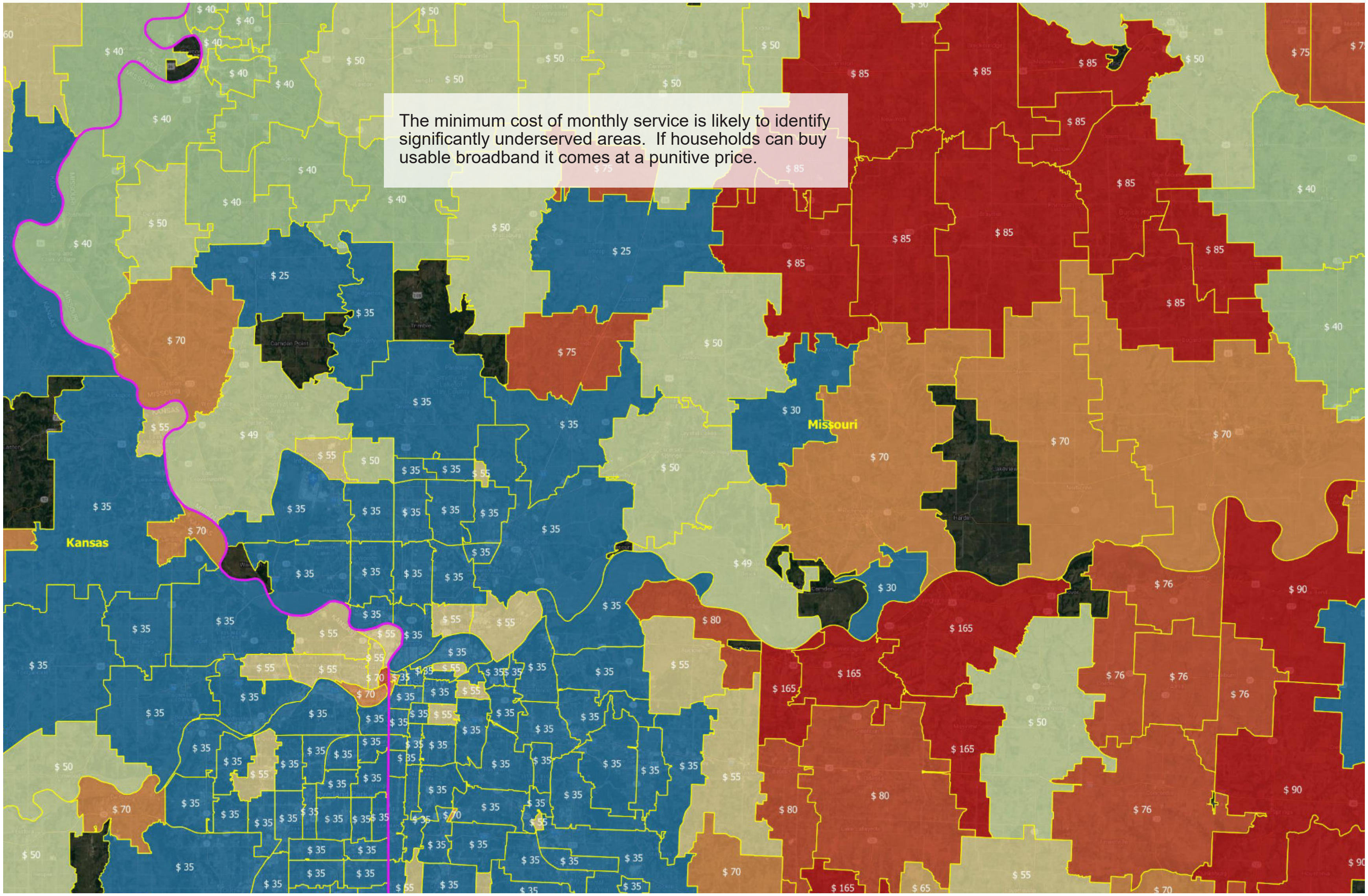
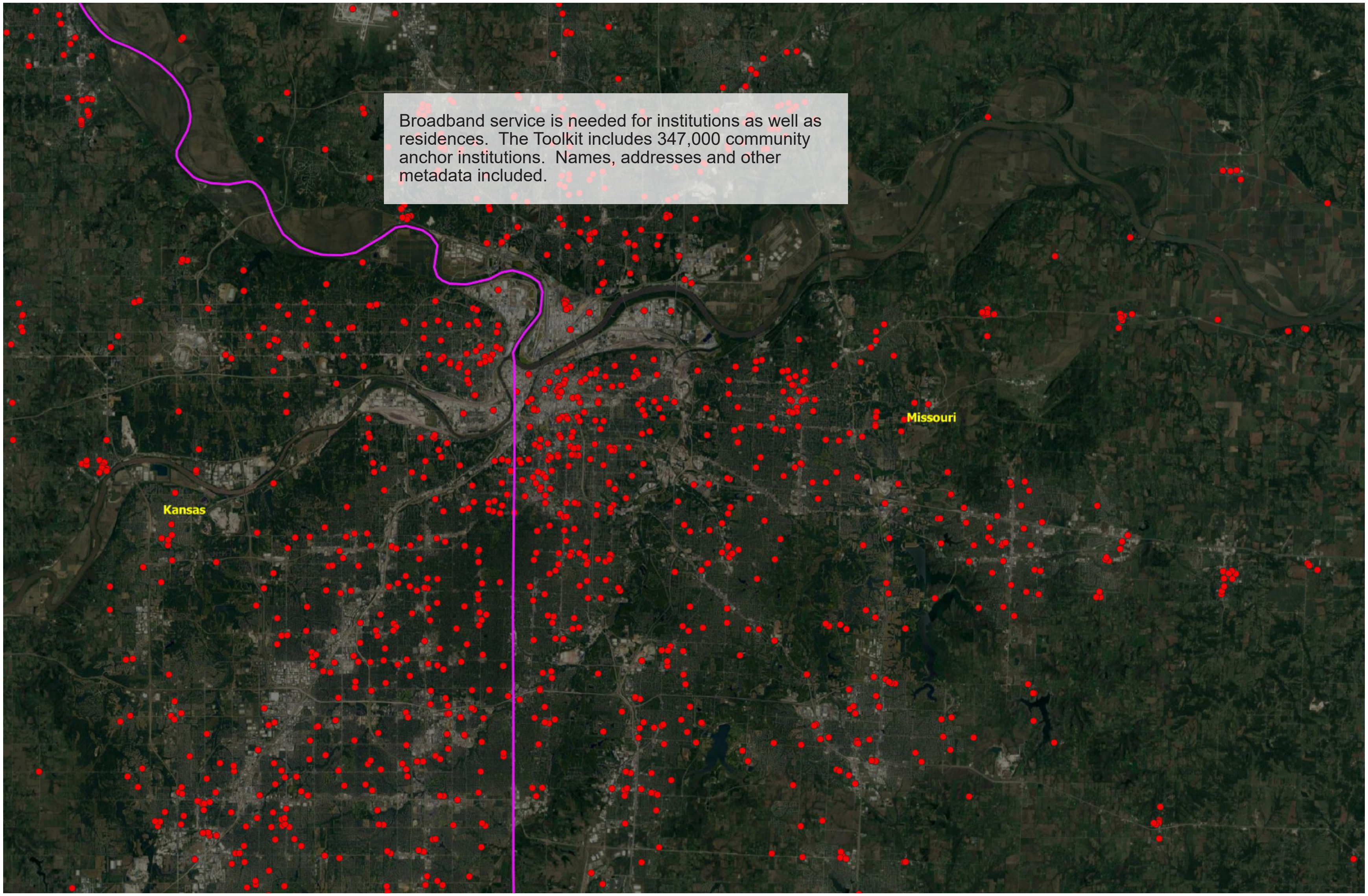


Figure 18: Community Anchor Institutions (7 selectable categories such as schools, libraries, medical facilities, etc.)



reflective fading) and to interference (especially in an urban environment). Equally importantly many consumers don't know how to set up a Wi-Fi network or how to optimize it for performance or how to measure performance. Many view Wi-Fi in binary terms, like electrical power that is either "on" or "off" (a black out). Wi-Fi networks are much more nuanced and can be a major factor limiting one's broadband experience. If a consumer has a poor experience because of a poorly configured Wi-Fi network it is unfair to blame the ISP.

- *Intentionality*. When someone runs a speed test it is likely that they believe they should have reasonable connectivity. In contrast, when machine-to-machine downloads are measured passively the consumer may be unaware that communication is taking place. A laptop user, for instance, might close his laptop and plug it in to charge in a room that has poor Wi-Fi coverage. A measurement of passive download speeds (e.g. operating system updates) may not represent the speeds the user experiences when deliberately connecting to the network.

- *Pricing of Service*. Most consumers do not purchase the most expensive broadband plan available. Consequently, a measurement of consumer speed is not necessarily an indicator of the capabilities of the associated outside plant.

FCC vs. Industry Data

With the above caveats, here are the included data sets:

- *FCC Fabric*. The FCC requires ISPs to report service availability twice a year as part of the Broadband Data Collection (BDC) program. The point-level results are mapped to census blocks and level 8 hexagons. Each point is tagged with the details of each ISP offering coverage. The volume of data, relative to that of Form 477 (the historic reporting mechanism) is large. An analyst using public data does not know the exact location (longitude, latitude, and street address) of each location ID, unless viewing information a few houses at a time on the streaming FCC map, but does know the associated hexagon. Since hexagons offer greatly improved resolution in rural areas relative to census blocks the analysis can be much more precise.

- *Federal Commitments*. The federal government (the FCC, the USDA, and others) as well as state governments have subsidized the deployment of broadband infrastructure. The Toolkit includes data from the four most impactful federal programs. BEAD rules restrict infrastructure grants in these areas to avoid duplication. The Toolkit includes the following calculated layers, with various sub-layers for challenges and rescinded awards:
    - Broadband Infrastructure Program (BIP).
    - RDOF. A map layer that shows all RDOF awards that fit the "low latency" category and deliver 100-20 Mbps.
    - CAF II. A map layer shows all CAF II awards that meet or exceed 100-20 Mbps and are "low latency".
    - USDA. Map layers (protected and pending) showing USDA commitments, most notably ReConnect, including both grant and loan programs.

- *Measured Data*. The NTIA, as part of their National Broadband Mapping Program, collected measured data from multiple commercial and non-profit entities and made that data available to the public. The contributors included Ookla, M-Lab, and Microsoft. Links to each of the source data sets, including additional descriptive information and, in some cases, public use licenses, are included in the "Web Site Links" folder of the Toolkit. The data is provided in multiple resolutions and with a wide variety of metrics. The Toolkit visualizes the most important metrics. Here is a quick summary of each data set:

- *Ookla*. Ookla, the owner of speedtest.net, reports gathering 37 billion speed tests worldwide. The data they provided to the NTIA was collected between January 1st and June 30th of 2020. It is aggregated by census tract and by county. It reports median download and upload speeds.

- *M-Lab*. M-Lab is "a consortium of research, industry and public-interest partners dedicated to: Providing an open, verifiable measurement platform for global network performance. " M-Lab similarly reports median download and upload speeds. The data they provided to the NTIA was similarly collected between January 1st and June 30th of 2020. The results are aggregated on a county basis.

- *Microsoft*. Microsoft reports the performance associated with passive machine-to machine downloads (operating system and other updates). The sample size is spectacular. The results (characterizing the percentage of downloads fulfilled over a 25 Mbps / 3 Mbps or better connection) are more pessimistic than others, possibly because of the passive nature of the communication and the possibility that many devices may be in poor Wi-Fi coverage. The results are aggregated on a county basis.

- BroadbandNow (also BroadbandNow Research) collects a range of metrics on a Zip Code basis. They were not included in the NTIA collection but represent another credible source of measured data. The resolution of a Zip Code is better than that of a county but not as good as that of a tract. The Toolkit shows BroadbandNow download speeds and the "The Lowest Regular Monthly Priced Terrestrial (Wired + Fixed Wireless) Residential Standalone-Internet Broadband (25 Mbps Download / 3 Mbps Upload) Plan available." Areas with poor internet connectivity often have high prices, as shown by this metric.

In its 2021 grant programs the NTIA strongly encouraged applicants to present their own data. This could be data they collected or data acquired in collaboration with another entity.

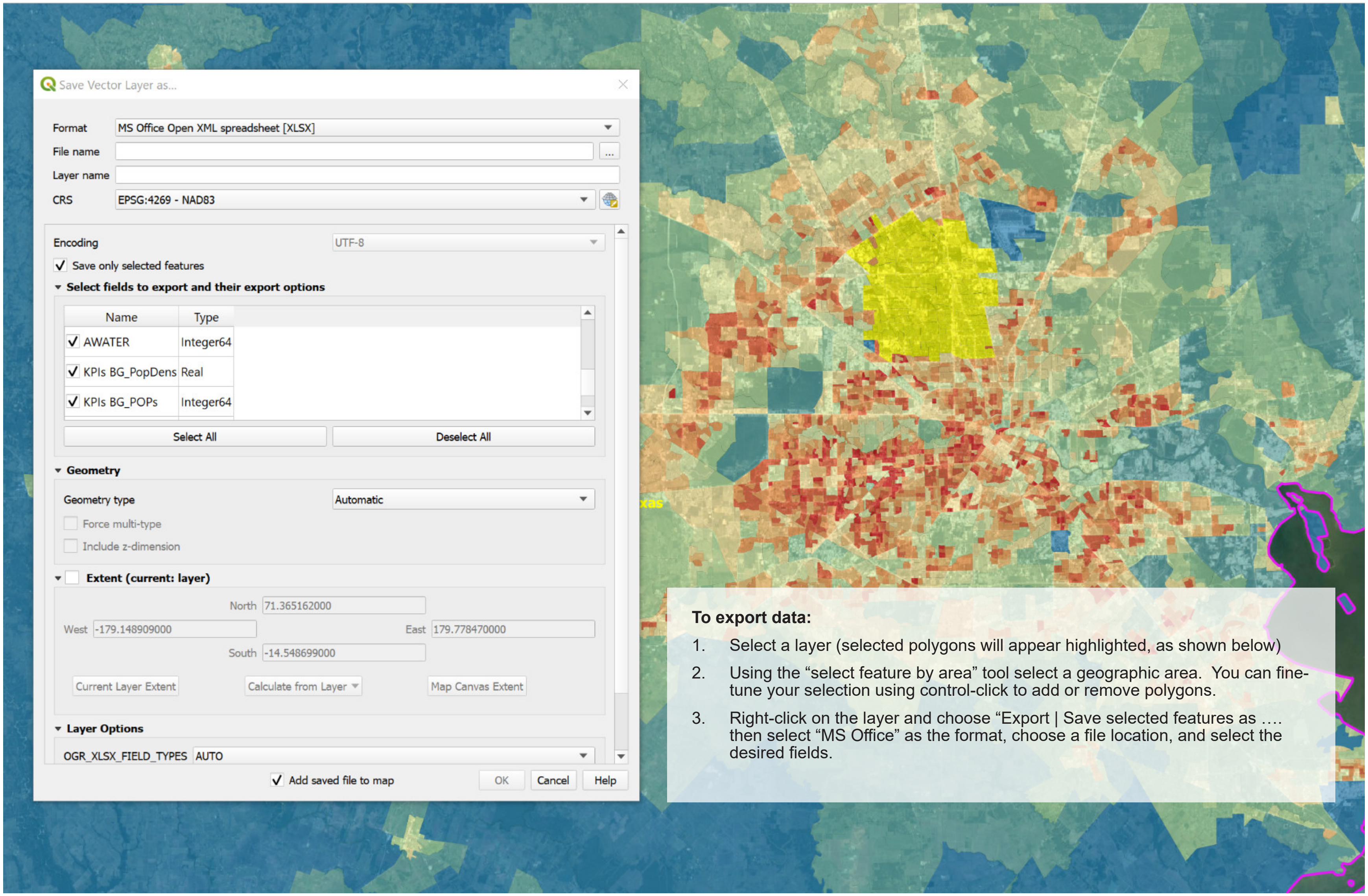
Custom data is helpful in making the case that an area, while theoretically covered based on operator submissions, may not be covered to the required standard.

BroadbandToolkit.com enables interested customers to purchase current crowd-sourced data from Ookla. The data, after being processed as described in the Ookla whitepaper (link) can be integrated as a set of layers into the Toolkit.

Layers of Visualization

- Inputs required to identify eligible geographies (tract and block group):

Figure 19: How to Select and Export Data



- Household size
- Median household income
- Poverty threshold as a function of household size

- Key demographic inputs (block group):
  - Population density
  - Household density
  - Housing unit (physical structures, whether currently occupied or not) density

- Key income inputs (block group):
  - Income per capita
  - Mean income per household
  - Median income per household
  - Average household size
  - Estimated poverty line
  - SNAP (a.k.a. food stamps) household participation rate

- Important contextual data:
  - A wide variety of streamable background maps (physical geography, roads, etc.) from leading map content providers, as well as solid backgrounds (for readability).
  - Geographic boundaries (block groups, tracts, zip code tabulation area, counties, states, and congressional districts) alone, or with associated numerical codes and/or names.
  - FCC license areas (BEAs, BTAs, CMAs, EAGs, MEAs, MTAs, PEAs, REAs, RPCs), relevant to those who may be providing wireless solutions.
  - Roads (primary and/or secondary, with or without labels).

- The data sets take several forms that can be used in any combination:
  - Visual data in a geographical information system tool (where layers can be enabled or disabled).
  - Color-coded layer information (enabling rapid assessment of large geographic areas)
  - Numerical text layers (enabling the user to see exact numbers)
  - Selectable data (enabling the user to select a geographic area and extract that specific data from a large database).
  - Excel spreadsheet data (enabling the user to view and manipulate all of the data that ships with the tool).

Many of these included resources are described in greater detail in the coming pages.

Demographic Data Sets

It is important to understand the significance of the various demographic data sets:

- *Housing Units* reflect the number of physical structures (single family homes, apartments, condominiums, mobile homes, etc.) in which a household could reside. The occupancy rate is the ratio of (rented HUs + owner-occupied HUs) / total HUs.

  In a city with 100% occupancy housing units could – conceptually – equal households, although such a situation rarely, if ever, exists.

- *Households* are the number of groups of people (family and non-family) that live together. A household would generally have a single fixed internet connection to the home. We know the number of households and the population associated with the households for each census block.

- *Group Quarters* are larger groups who do not live in households. Group quarters include university dormitories, nursing homes, and prisons, as examples. The population in group quarters is distinct from the population living in households.

- *Population*. This is the total number of people living in an area, regardless of their housing situation. This total population includes those living in households (the vast majority, whether in family or non-family households) plus those living in group quarters.

Income-Related Data

The Toolkit provides a number of layers of data describing income and income-related programs:

- *Annual Income per Capita*. This is annual aggregate income divided by total population.

- *Mean Annual Household Income*. This is annual aggregate household income divided by total households. The portion of the population that lives in group quarters (college dormitories, nursing homes, and prisons, as examples) is excluded.

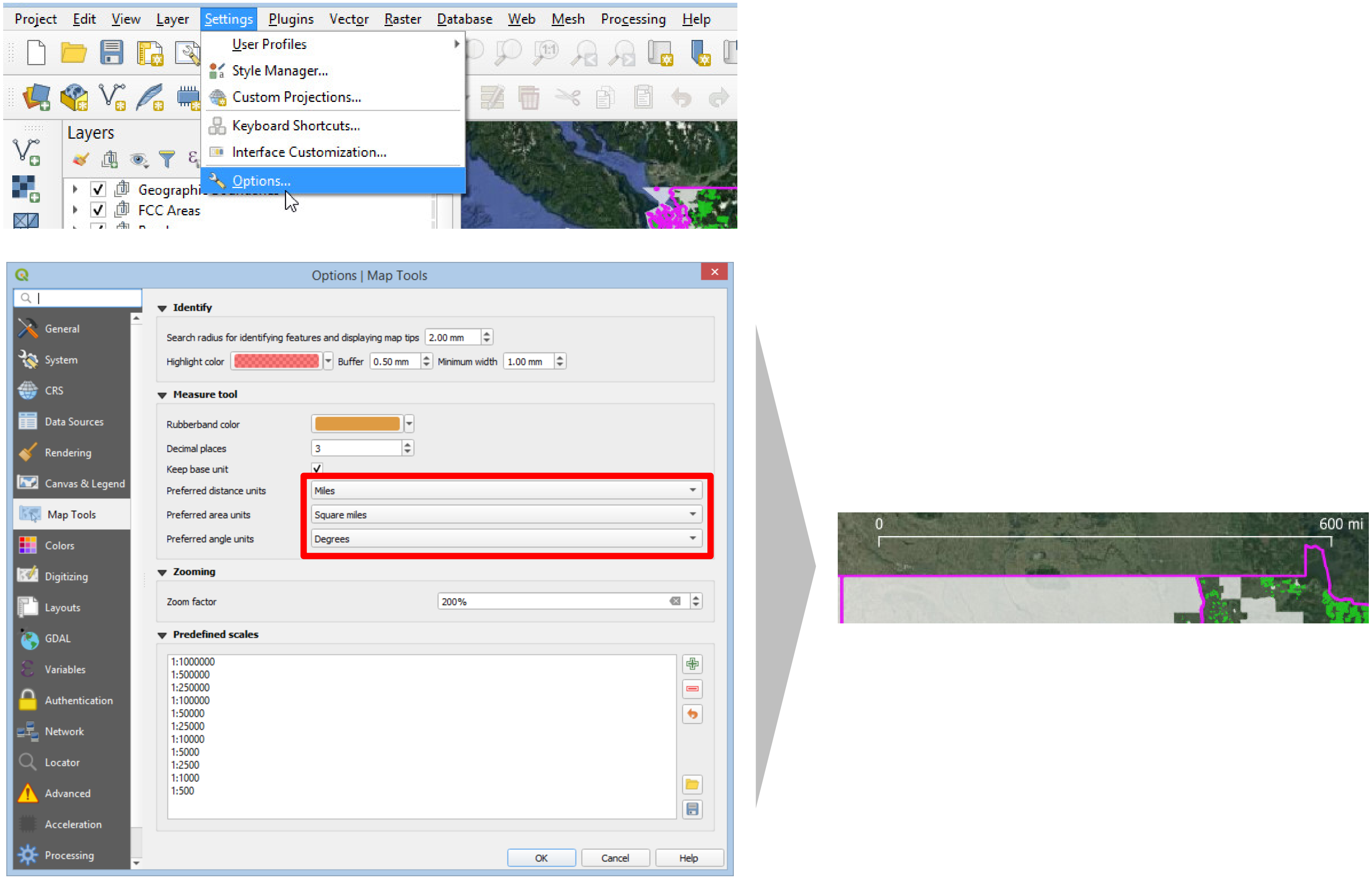
- *Median Annual Household Income*. The median annual household income is a number above which half the households earn more and below which half the households earn less. In most geographic areas median income is lower than mean income because a few larger earners pull the mean upward. Median income is considered the best indicator of household buying power for non-luxury goods.

- *Poverty Line*. Both the US Census Bureau and the US Department of Health and Human Services (HHS) provide poverty metrics. The Census Bureau uses “Poverty Thresholds” for statistical purposes. It is a 48-cell matrix that includes family size, number of children, 1 and 2 person units, and whether or not an individual is elderly. There is no geographic dimension. In contrast, the Department of Health and Human Services uses a relatively simple “Poverty Guidelines”. The latter can be calculated based on family size and geography (Alaska vs. Hawaii vs. the Contiguous 48 States). In its 2021 grant programs the NTIA decided to use a national average of the Census Bureau “Poverty Threshold”. The way the NTIA used it, the calculation was dependent upon household size. The reader was instructed to round the average household size in a geographic area up to the next integer value then performed a lookup.

- *Household Size*. The household size is calculated for each block group using the total number of households and the total population in households (excluding the population in group quarters).

- *SNAP Participation Rate*. The Supplemental Nutrition Assistance Program (SNAP), a.k.a. food stamps, participate rate is shown as a percentage of households within each

Figure 20: Setting the Legend to Miles



block group. A household receiving SNAP benefits is eligible for ACP and Lifeline subsidies..

USDA Data Sets

The US Department of Agriculture (USDA) publishes a number of data sets that represent either evaluation criteria or eligibility criteria. They identify areas of economic need, areas that meet certain definitions of "rurality", and areas that have received or are likely to receive USDA funding:

Measures of rurality:

Frontier and Remote Areas (FAR Level 4) that are:

- 15 minutes or more from an urban area of 2,500-9,999 people
- 30 minutes or more from an urban area of 10,000-24,999 people
- 45 minutes or more from an urban area of 25,000-49,999 people
- 60 minutes or more from an urban area of 50,000 or more people.

FAR is visualized as a set of ZIP Code areas. It is described in detail on the USDA Economic Research Service FAR web page.

Non-Rural Areas are represented as a set of polygons. They tend to coincide – as one might expect – with cities and other urban areas.

The 100-mile butter is, as its name suggests, a buffer around urban areas.

Economic Need:

Small Area Income and Poverty Estimates (SAIPE) are *counties* with a 20% or higher rate of poverty. It is described in detail on the Census Bureau SAIPE web page.

Socially Vulnerable Communities, identified by the Socially Vulnerable Index (SVI), are the most vulnerable census *tracts*, based on 15 measures of economic vulnerability.

Funded Geographies:

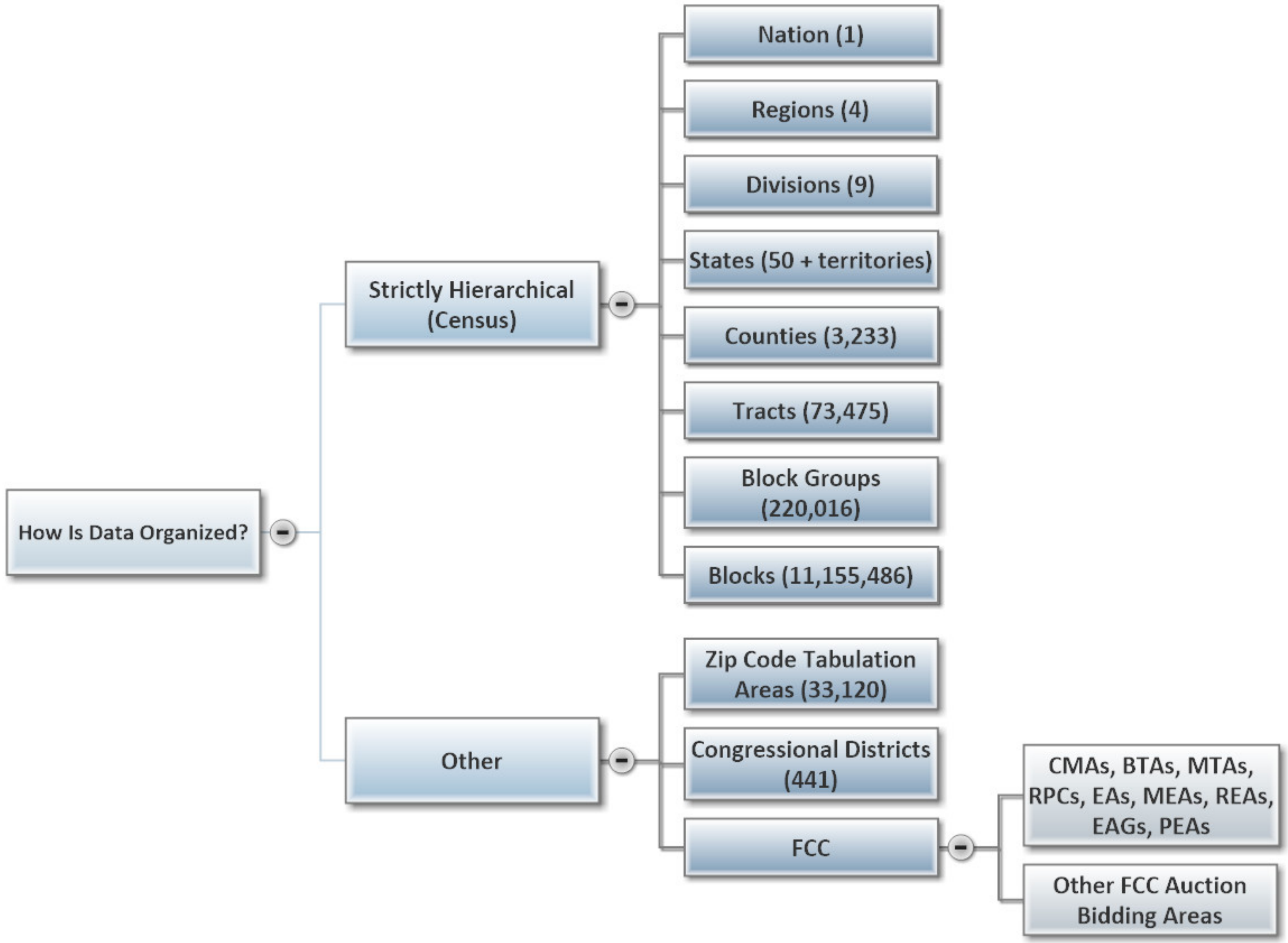
Protected. These are areas that have received USDA funds (ReConnect, Community Connect, Farm Bill Broadband, Telecommunications Infrastructure), either grants or loans, and are "protected" from duplicative funding.

Pending Applications. These are areas with applications for USDA funding currently under consideration.

Technology

The Toolkit explicitly shows which technologies are present in each level 8 hexagon at each of three performance levels. The technologies evaluated are those relevant to BEAD funding: fiber, cable, DSL/copper, and licensed FWA. The performance thresholds are "any performance", 25-3, and 100-20. Populated areas that are not covered by any of the four BEAD "reliable" technologies are blank.

Figure 21: Hierarchy of Geographic Boundaries



Opportunity Zones

Opportunity Zones were created by the 2017 Tax Cuts and Jobs Act to spur economic development and job creation in distressed communities. Opportunity zones have been designated in all 50 states and in every inhabited US territory (American Samoa, Guam, Northern Mariana Island, Puerto Rico, and the US Virgin Islands). Investors benefit from deferral or exclusion of capital gains. The community benefits from financial investment. They are designed to encourage investment in areas with high socioeconomic needs. The statute excludes specific luxury investments (e.g. golf courses) and specific "sin industries", but is otherwise applicable to any business.

The Toolkit visualizes Opportunity Zones either as bright green polygons (if one wishes to find them on a map) or as clear tiles in an ocean of whited out space. Understanding the location of Opportunity Zones relative to other measurable metrics – particularly indications of need – is extremely powerful. It enables a business to deploy capital in locations that are likely to have a favorable social impact while offering investors unique tax incentives. Local governments (urban and rural) may wish to encourage outside investment based on the juxtaposition of Opportunity Zones with important needs identified by other layers of data.

Geocoded Data

The visual portion of the Toolkit includes a number of important features:

- *Scrollable / pannable user interface*. A user can easily zoom in for a closer view or zoom out for a more distant view using the mouse's wheel. Similarly, a user can grab the image and drag it in any direction. One can start in Florida, zoom out then fly to Hawaii or Alaska (with the national tool) then zoom in again for a detailed view. As the user manipulates the screen all of the underlying data scrolls and pans accordingly.

- *Many layers*. The Toolkit includes many layers of data that can be individually toggled on or off. The Toolkit is designed to show one graphical layer at a time. Each layer is translucent, so that the underlying map is visible.

- *Meaningful colors*. Numerical data is displayed using a range of colors. In most cases the scale runs from blue (low values) to red (high values). The color scheme provides a hint as to the type of data. In each case there are ten color thresholds that represent quantiles (an equal number of data points). There is no absolute meaning to "red" or "blue" across data elements, since the color scheme changes with each data element to reflect the underlying range of the data. The colors are designed to provide a visual cue to help the user see patterns and identify outlying values within any particular data set. Each data set has a legend that the user can see by clicking the "expand" triangular icon next to the "Visual" label for the data set. The legend displays the exact range of values associated with each color.

- *Numerical overlays*. Most of the data sets (all of demographic and economic data) display not only colors, but also numbers. With demographic and economic data it is desirable to turn on the numerical overlay to see the exact value of each underlying region, especially when zoomed in to a small geographic area. If the user zooms out it is generally desirable to turn off the numerical display and enjoy a rich mosaic of color. If numerical text continues to be displayed when zoomed out the text associated with adjacent regions starts to overlap and quickly becomes unreadable.

- *Geographic Boundaries*. The Toolkit includes a long list of boundaries that can be turned on or off. These include state lines, county lines, congressional districts, zip code (ZCTA) boundaries, census tracts, census block groups, and census blocks.

- *License Areas*. The FCC has licensed spectrum over the years using different geographic boundaries. Some users of the Toolkit may own spectrum. The toolkit therefore includes boundaries for the most widely used license areas. These include: Cellular Market Areas (CMAs), Basic Trading Areas (BTAs), Major Trading Areas (MTAs), Regional PCS Areas (RPCs), Economic Areas (EAs or BEAs), Major Economic Areas (MEAs), Regional Economic Areas (REAs), Economic Areas Groupings (EAGs), and Partial Economic Areas (PEAs).

- *Boundary names and other data*. In most cases the name of the boundary (e.g. the state or county) can be displayed. Alternatively a code may be displayed. The Census Bureau has a hierarchical numbering scheme called FIPS that begins at the state level (2 digits) then goes to the county level (2+3 = 5 digits) then to the tract level (2 + 3 + 6 = 11 digits) then to the block group level (2 + 3 + 6 + 1 = 12 digits) then to the block level (2+ 3+ 6 +1 +3 = 15 digits). If a user wishes to pull up spreadsheet data that corresponds to a visualize image it is helpful to turn on the numerical display for FIPS, take a screen shot, then look for the corresponding data set of spreadsheet data. Since everything is hierarchical one can select a county (the first 5 digits of the FIPS) by selecting all the block group data with the desired county code in the first 5 digits, as an example.

- *Roads*. The Toolkit includes primary and secondary roads, with or without name labels. Road layers may be turned on or off. Alternatively, one might choose an underlying map that includes road and place labels. Google, Bing, and OpenStreets, in the Maps folder, each include this option.

  Choice of Units for Distance and Area Measurements. In the United States people discuss distances in miles and areas in square miles or perhaps acres. The scientific community tends to use kilometers and square kilometers. The FCC and the Census Bureau have increasingly adopted metric units in their publications.

  1 km = 0.621371 statutory miles. Similarly 1 square kilometer = 0.386102 square miles. A square mile is equal to 640 acres or 258.999 hectors.

  The Toolkit can display the map legend in either kilometers or miles. Figure 20 shows how to change from one set of units to another.

  The internal databases of the Toolkit, including the spreadsheets, represent units in meters, square meters, kilometers, and square kilometers.

Graphical Versus Tabular

To build a business case one needs real data, not just a pretty picture.

The Toolkit includes a comprehensive set of geocoded spreadsheet data that largely matches the demographic and economic data sets in the visual tool. Each data set includes:

- *Numerical Code*. A numerical identifier for the region (typically a FIPS code), either a block or block group or a fragment of a block.
- *Name of Region or Entity*. The name of the region (e.g. a state and county or tribal nation) or an entity (a college, university).
- *Calculated Metrics*. Examples include areas, percentages, prorated metrics, growth factors, median household income, etc.
- *Raw "Counter" Data*. Raw data. The most important elements are housing units (structures), households (people), population (people), and the population included in households. Census polygons are hierarchical. Each block group, for instance, sums all of the blocks within it.
- *Population Estimates / Forecast.* In addition, the Toolkit includes separate population estimates / forecasts for every county in the 50 United States plus for the District of Columbia and Puerto Rico, based on the most recent yearly data from the Census Bureau. Recent county-level growth rates are used to produce near-term forecasts through the date of the auction (July 1, 2020).

Optimizing Performance

Geographical Information Systems, such as the QGIS browser, are data crunching machines and miraculous pieces of software in that they seek to visualize overwhelmingly large amount of data elegantly. The Visual Toolkit, as an application, has been highly optimized for performance, using as few computer resources as possible to achieve its objectives. Even so, it is helpful for every user to be aware of factors that impact performance:

- *Computer Hardware*. Every GIS application demands significant hardware resources. Ideally, one would run the Toolkit on a computer with a 64-bit operating system, lots of RAM, reasonable processing power, and fast disk access (ideally SSD). The current version of QGIS and the current highly optimized version of the Toolkit will both run with modest resources.
- *Application Loading Time*. Expect any GIS application to take a few minutes to load. Think of it as an opportunity to get a fresh cup of coffee. During the load process it connects to all of the linked data sets and prepares to load the associated data on demand. It doesn't work like Microsoft Word, where one clicks on a document and, an instant later, the document appears. Once loaded, though, the application is designed to be responsive, with a few caveats.
- *First Time Loading a Layer*. The first time a user loads a visual layer after launching the application in QGIS the application may pause for a few seconds as it finds the desired data. Afterwards, one can select and deselect that data layer and expect the text and graphics to appear and disappear almost instantaneously, because it has been cached by the

application. One can then sometimes zoom in and out and pan with minimal delay.

- *Streaming Maps*. In theory, streaming maps can cause the user interface to be slow because map data must be retrieved from a remote server. In practice, with the current version of QGIS and with a fast internet connection, the delay is negligible. Do be careful not to enable multiple maps at the same time. Doing so will multiply the volume of data that must be downloaded. Also, the user will see only one map at a time, so most of the effort will be wasted. If the user is in an airplane or has a slow internet connection or no connection at all then the user should disable the map by unchecking the map layer. If one is unsure about the impact of the map it is easy to disable it and to enable, instead, a solid color background (gray, black, white) to see if the application becomes visibly more responsive. Dark solid backgrounds, while not as pretty as a map, are wonderful for reading detailed overlaid data.

- *Text Overlays*. A text overlay can sometime slow the display. This is generally the case when text is enabled with a high resolution data set and the user is zoomed out. Imagine, for instance, looking at the continental United States, viewing data at the block group level, with text enabled. QGIS would attempt to write 200,000 numerical values on the screen, one number for each polygon visualized. The user interface would be slow and the resulting image would be a mess. It is best to turn off text overlays before you zoom out then decide what text is appropriate at the new zoom level. Up close, one might be interested in block group FIPS codes, but zoomed out one might be interested in state boundaries and state names, as an example. High resolution boundaries (e.g. block group boundaries) when zoomed out can also slow the user interface and flood the resulting image with a single color of ink (reflecting the color of the boundaries). The implications of most of these decisions will become obvious the first time one uses the application.

Curated Sources

The carefully curated data in the Toolkit comes from a number of exceptionally high quality government sources as well as other public and private sources already discussed:

- US Census Bureau (many different products)
- Federal Communications Commission (FCC)
- National Telecommunications Information Administration (NTIA)
- U.S. Geological Survey (USGS)
- U.S. Congress (legislative text)

Significant sources include:

- US Census Bureau, American Community Survey. The ACS surveys 3.5 million households + 185,000 persons in group quarters per year on subjects ranging from household income to demographics to physical space and amenities to devices and connectivity to monthly household expenditures. These are converted into single-year and 5-year estimates and other data products. The Toolkit uses the 2019 5-year estimates, which were published on December 10th, 2020.

- US Census Bureau, Population and Housing Estimates (PEP), The International Data Base, County Business Patterns, and a wide rage of geographic boundary products and definitions. Extremely important are well-documented processes, which give the numerical data important context and meaning.

- The Federal Communications Commission. The FCC collects fixed and mobile coverage by technology by operator, as well as other service metrics and publishes detailed auction results. This tool Fabric data (published 11/18/2022), reverse auction (CAF II and RDOF) funding commitments for geographies that meet or exceed the FCC's definition of broadband, as well as USDA (e.g. ReConnect) funding commitments. Visibility into funding commitments is important for infrastructure-oriented grants to prevent a duplication of funding.

Video Tutorials

Video tutorials exist to help new users get started, understand the sophisticated functionality, enable and disable options, and effectively use the various Toolkit products. A two minute video highlights the capabilities of BroadbandToolkit.com.

Next Steps

Be sure to check for the latest whitepapers:

https://broadbandtoolkit.com/pages/whitepapers

The Toolkit team will be happy to answer your questions / discuss your needs by phone or video conference.

You can reach us by phone at *415-346-5393* or by e-mail at *support@broadbandtoolkit.com.*

Updated: April 25th, 2023