Digital Equity & Broadband Report

Prepared by students in PLAN 6020 and PLAN 6040 at the University of Virginia

Spring 2023

Table of Contents

Executive Summary

In a world increasingly reliant on digital access, broadband equity is more pressing than ever. Before the 2020s, the lofty goal of "broadband for all" could be dismissed as an idealistic dream. Today, after widespread reliance of schools and workplaces on remote communications technologies became commonplace during and after the COVID-19 pandemic, access to high-functioning broadband has been rightly reframed as a need rather than a luxury. With this reframing has come the realization that a very real and harmful "digital divide" is plaguing many of America's localities.

This project represents an attempt to understand that divide in the hopes of ultimately addressing it. It represents the continued progress of efforts by the Albemarle County Broadband Accessibility and Affordability Office and the Regional Digital Equity Plan Coalition (RDEP), in partnership with the University of Virginia and Penn State University and funded by the federal "Your Home, Your Internet" Affordable Connectivity Outreach Grant Program. Initial analysis was conducted by graduate students in the University of Virginia's Urban and Environmental Planning Program. We analyzed previously collected quantitative data from the RDEP coalition and data from the U.S. Census Bureau to generate conclusions for each separate dataset as well as inter-regional comparisons and composite metrics of broadband accessibility. During the qualitative portion of the project, we conducted original research by conducting focus groups with local community partners. The end goal of this research was to aid and inform the creation of a Regional Digital Equity Plan for Albemarle County, Virginia and surrounding areas.

This report contains our findings. It is subdivided into an introduction to broadband access and the project's goals, a section dedicated to qualitative analysis, and a section dedicated to quantitative analysis.

The qualitative portion of the project first aimed to identify common themes in the focus groups. Systematic content analysis revealed that the themes most frequently mentioned in the focus groups were literacy/digital literacy skills, learning curves, affordability, digital inclusion, and connectivity.

Teams also generated next steps, in the form of suggestions for the County, from the main findings of the focus groups. They include (a) the implementation of digital literacy programming for children and adults; (b) the provision by Albemarle County of easy-to-use desktop computers in central gathering spaces; (c) the establishment of laptop loaner pools; (d) greater advocacy by the County on behalf of residents when collaborating with broadband providers; (e) a basic understanding that not all residents have or use smartphones; and (f) the need to continue researching this topic.

Quantitative reports were generated from four different data sets. These four data sets each address a different set of questions related to digital equity in the region. The four groups focused on (a) comparing Albemarle County to analogous counties nationwide, (b) disaggregating the characteristics of census tracts with different levels of broadband adoption (i.e. income, race/ethnicity, gender, rural/urban, own/rent, job type), (c) examining speeds and preferences from RDEP surveys and associating those data with local demographics, and (d) creating composite metrics of broadband access including schools data, census data, and speed tests. Major findings from the groups included the importance of prioritizing residents' upload and download speeds to improve sentiments about internet in the region. Another product of the quantitative analysis was the creation of a composite metric to determine the quality of broadband access in a given census tract.

Acknowledgements

We want to thank the following individuals and groups for their contributions to this report.

Students:

- Meredith Beavers
- Bo Bian
- Harrison Bluestein
- Chris Chao
- Maddy Duval
- Tatiana Gold
- Heidi Hahn
- Gwendolyn Hellen-Sands
- Mia Hsu
- Dylan Metsch-Ampel
- Nick Meurer
- Yara Mortada
- Kristan Pitts
- Emily Routman
- Maddie Rodgers
- Alan Simpson
- Adam Stevenson
- Trevor Storm
- Lizzie Suffa
- Brian Van

Professors:

- Dr. Barbara Brown Wilson
- Dr. Andrew Mondschein

Community Partners:

- Jason Inofuentes
- Dr. Christopher Ali
- Yancey School Community Center
- Piedmont Housing Alliance
- Jefferson Area Board for Aging
- Starr Hill Pathways
- Bright Stars

PIEDMONT
HOUSING
ALLIANCE

STARR HILL PATHWAYS **BRIGHT STARS**
YANCEY SCHOOL COMMUNITY CENTER

Introduction

As remote work, online learning, and telemedicine have become the new normal, access to high-speed internet has become a necessity for individuals and communities. However, not all individuals and communities have equal access to these resources, which has led to disparities in education, healthcare, and economic opportunities. Therefore, the need for digital equity has become increasingly pressing.

This report presents the findings of a comprehensive study conducted by the students of two University of Virginia classes, PLAN 6020: Methods of Community Research and PLAN 6040: Quantitative Methods of Planning, in partnership with the Regional Digital Equity Plan coalition and the Broadband Accessibility and Affordability Office of Albemarle County. This project aims to inform a Regional Digital Equity Plan (RDEP) for Albemarle County, Virginia, focusing on addressing broadband access and equity in the region. The final product consists of two sections: the qualitative report and the quantitative report.

For the qualitative report, we compiled background information, developed survey questions, conducted focus groups within the community, and transcribed and coded qualitative data from the focus groups.

The community groups were as follows:

- **Jefferson Area Board for Aging (JABA):** "JABA has been supporting seniors, adults with disabilities, and caregivers throughout central Virginia for over 45 years."
- **Piedmont Housing Alliance:** "Piedmont Housing Alliance is dedicated to improving financial outcomes for individuals and families by offering innovative affordable housing solutions."
- **Yancey School Community Center:** " A unique facility that provides recreational programming and services." Yancey provides a community center space that services southern Albemarle County residents.
- **Starr Hill Pathways:** "The Starr Hill Pathways program is designed to offer Charlottesville area youth, career and college exploration and support them, as rising 7th graders through high school."
- **Bright Stars:** "The Bright Stars Program is a comprehensive preschool program provided in collaboration with the Albemarle County Department of Social Services and the Albemarle County Public Schools. In addition to a high-quality preschool experience, the children and their families are provided support and case management by a family coordinator through the Department of Social Services."

For the quantitative report, we conducted multiple analyses on broadband and social data. This report summarizes the findings of these efforts and presents recommendations for developing a Regional Digital Equity Plan (RDEP). The report also highlights the contributions of community members who participated in the focus groups and acknowledges their role in shaping the final product. Ultimately, this project and the cumulative joint report aim to improve broadband access and equity in the region and contribute to the development of a more connected and inclusive community by informing the development of the Regional Digital Equity Plan Coalition (RDEP).

The RDEP was our key partner in this research project. The Coalition is facilitated by the Broadband Affordability and Accessibility Office of Albemarle County but includes members from a diverse array of departments within the County. The Coalition's mission is to advocate

for and facilitate programs that provide affordable and accessible broadband to all. We worked closely with Jason Inofuentes, a Broadband Program Manager within the Broadband Affordability and Accessibility Office, to determine community partners and to organize relevant focus groups. The work drew heavily upon the data, trends and connections already established by the Regional Digital Equity Plan Coalition.

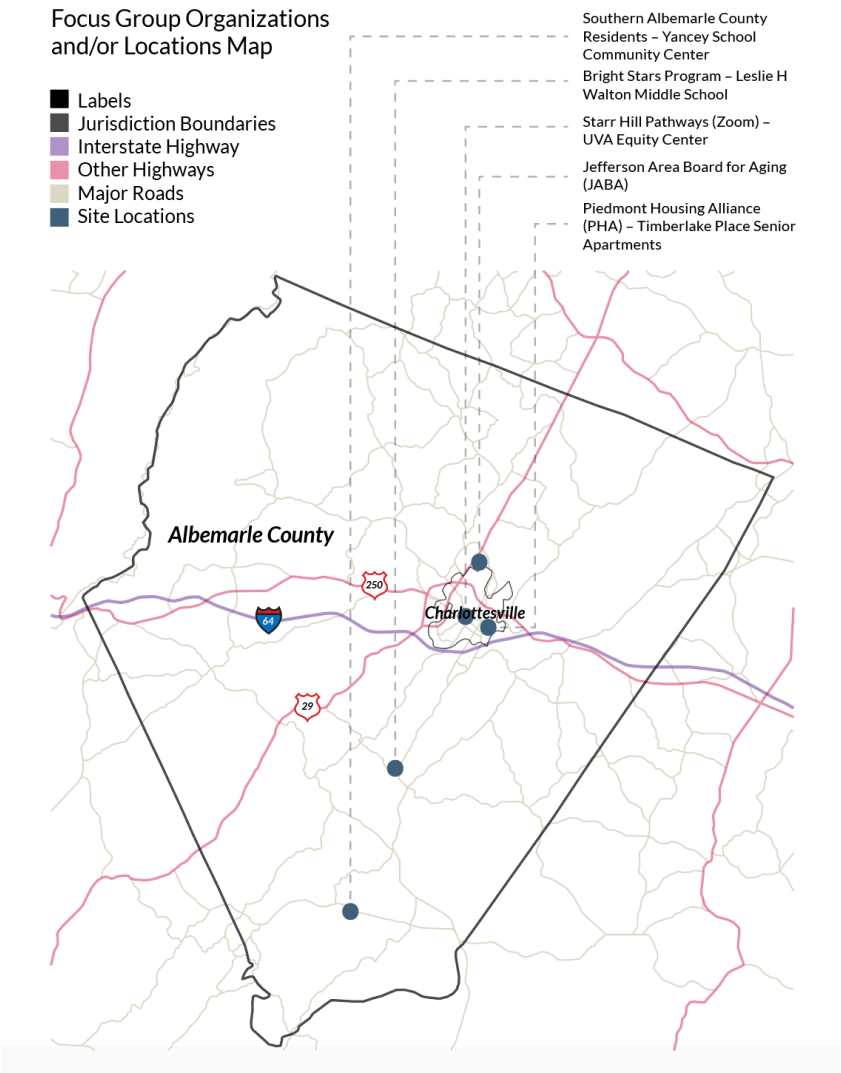

Figure 1. Focus Group Organizations and/or Locations

Qualitative Analysis

MOU OUTLINE

At the project's outset, a Memorandum of Understanding (MOU) agreement was made between the student consultants, the Regional Digital Equity Plan coalition, the Broadband Accessibility and Affordability Office, and Dr. Christopher Ali, Professor of Telecommunications at Penn State University. The MOU defines the project's scope of work and outputs, the responsibilities of the community partners and student consultants, and the terms of intellectual property sharing between all parties.

The MOU includes the following specifics:

- The student consultants will conduct focus groups, transcribe and code qualitative data, prepare survey questions, and generate a final report.
- The community partner will provide bi-weekly status updates, recruit participants, coordinate the timing and location of focus groups, and attend the student consultants' final report presentation.
- The output will be a final report that assesses the results of the focus groups and presents a demographic analysis.
- The community partner will bear the cost of purchasing necessary event materials or outside services and compensate focus group participants for meals and other costs associated with removing barriers.
- Any intellectual property jointly created will be jointly owned and credited appropriately.

METHODS USED BY ALL GROUPS

Before the focus groups met, participants were provided with a consent form and/or a minor assent form. The consent form was intended to ensure that participants understood the research study's goals before participating in the focus group. In addition, we read the consent form aloud at the beginning of each focus group to ensure that everyone understood its contents. The goal of the project was to identify the digital needs and struggles of residents of Albemarle County and the City of Charlottesville. These insights will be used to create the Albemarle County Digital Equity Plan. The focus groups lasted approximately one hour, and there were no known risks associated with participation. Focus group participants were assured that no names or identifying information would be used in the final presentation or report of the study.

The consent form and verbal preamble explicitly stated that participants could withdraw from the study at any time without any impact on their financial incentive. The project team recognized that the time and effort participants allot to a study is invaluable and wanted to offer an incentive for their contributions – a $25 VISA gift card. The consent form and preamble also emphasized that the incentive was still available to the participants regardless of their decision to leave during the focus group. To ensure that participants received their compensation in a timely manner, the project team provided access to the gift card through a virtual form that was filled out during the focus group and subsequently mailed to them.

Once the preamble was completed, the participants were notified that the recording had started. Then, questions began. The questions asked during the focus groups were standardized, with each focus group receiving the same questions. They focused on inquiries about

accessibility to the internet, the participants' thoughts and feelings, and potential struggles with the internet. We aimed to gather in-depth insights into the experiences of participants and the needs of the Charlottesville and Albemarle County area using these questions.

Following an icebreaker question ("How did you spend your weekend?"), the moderator began a series of starting questions centered on how the participants regularly access the internet and why they use it. The questions were as follows:

- How do you regularly access the internet?
- Thinking back over the last two days, if you accessed the internet, why did you? What did you do?

These questions were designed to provide context for the participants' experiences with broadband access and affordability. Depending on the focus group, the participants' experiences may have revealed potential areas of focus regarding usage. Therefore, the next set of questions focused on participants' thoughts and feelings about their current circumstances as they related to affordable subscriptions, hardware, and skills to use the internet effectively. The questions were as follows:

- Do you feel that you have the speeds and connections necessary to do everything you want online?
- Do you feel that your subscriptions are affordable for your household income?
- Do you feel that you have the hardware necessary to do everything you want online?
- Do you feel that you have the skills and training necessary to do everything you want to do online?

These questions aimed to uncover any barriers that the participants may have encountered when accessing the internet. The final set of questions focused on asking participants to share stories about any potential struggles relating to the internet. The questions were as follows:

- What would you love to be able to do with the internet that you are unable to right now?
- Tell us about a time that you struggled doing something online.
- What do you think is standing in the way of you achieving this goal?
- If you could talk directly to the County about broadband and connectivity, what would you tell them?

Ultimately, the questions asked of the participants aimed to provide a deeper understanding of the participants' experiences with broadband access and affordability and to identify any common themes that could inform the development of the Albemarle County Digital Equity Plan. While experiences may vary across communities, the answers provided allowed student consultants to identify patterns in digital inequity across the Albemarle County and Charlottesville region.

Within each focus group, designated transcribers and coders utilized the recorded audio to investigate and identify potential patterns. Transcribers also took care not to name the participants in the written transcripts. Instead, the transcripts would utilize a unique letter to differentiate the participants whilst keeping their involvement as anonymous as possible. Following this, coders would utilize the transcripts and a predetermined code dictionary to identify patterns of behavior within the focus groups.

To code each focus group transcript, the student consultants developed a unified coding dictionary as a group and identified themes and topics in each transcript using HyperResearch, a free qualitative analysis software. The coding dictionary - as described in the quantitative analysis section - contained a list of 9 codes and subcodes. For example, one code that the student consultants identified was 'barriers', and one of its subcodes was 'unreliability'. The student consultants then loaded each transcript into HyperResearch and highlighted words, phrases, and sections of the text, to which we assigned the relevant codes. The resulting coded transcript created a

frequency report that detailed to the student consultants how often a theme/code/topic came up in that individual transcript. By doing so, the student consultants were able to quantify the patterns expressed in the focus groups.

RESULTS

The student research team developed a coding dictionary to analyze the focus group transcripts through a two-fold process. First, we read example digital equity plans from other U.S. cities published through "Partners Bridging the Digital Divide" to establish a running list of important concepts and themes[1]. After conducting the focus groups, we refined and edited the list based on the topics discussed with the participants. Appendix A contains the complete coding dictionary and the coding frequency tables for all groups.

Across all focus groups, the most common themes discussed were digital literacy skills, learning curves, affordability, digital inclusion, and connectivity. Despite similarities in the themes discussed, each group addressed the themes in different ways, highlighting different and valuable aspects of digital equity.

Jefferson Area Board for Aging (JABA)

The topics and themes discussed most frequently at the Jefferson Area Board for Aging focus group were digital literacy skills, learning curves, apathy, and fear. These topics were discussed as barriers to connectivity and digital literacy. Participants expressed the need for technology classes or technological assistance services that are repetitive and slow-paced to facilitate familiarity and learning digital skills. Participants said that grandchildren and other relatives often do not have the necessary time to go over how to do internet-related tasks and technology slowly, and it takes repetition and near-daily use for the skills to stick. Participants also discussed various fears about using technology, especially regarding internet scams or objectionable content. Some of this fear was accompanied by apathy from participants about furthering their digital literacy skills.

Piedmont Housing Alliance (PHA)

The participants in the Piedmont Housing Alliance focus group, who were residents of Timberlake Place Senior Apartments, mostly discussed topics relating to the affordability of services and technology, connectivity, digital inclusion, learning curves, and digital literacy skills. Several participants stated that they rely on community assets such as public libraries to access desktop computers and assistance regarding technology skills. They expressed that although the library enables them to use the internet, they would rather be able to do these things at home with affordable, easy-to-use computers.

Another notable topic of discussion among participants was the tendency for everyday services such as doctors' offices to assume smartphone ownership by requiring mobile check-in or some other mobile internet activity. The participants expressed concern that these societal assumptions are a threat to equity and leave those who cannot afford smartphones or internet connectivity behind. While each participant had home internet access, they were highly concerned with the affordability of their internet service and expressed that they would like a government program to help them access affordable computers to access the internet at home.

Yancey School Community Center

Unreliability, cell coverage, device accessibility, and connectivity were the most notable concerns of those who attended the Yancey School Community Center focus group. Many attendees shared that they frequently experience unreliable internet connection and access. Some shared that they often go hours without a stable connection and have to travel far to access the internet. In addition, they expressed concern over the speed of their internet connection. Some participants shared that their households do not have enough bandwidth to use multiple devices at once, which inhibits remote work capabilities and the ability for children to complete their homework -- most of which is completed online in Albemarle County Public Schools. Though not explicitly related to broadband, poor cell coverage and service throughout the region was another important issue discussed. Due to the area's hilly topography and rural nature, this problem is

1 https://www.pbdd.org/digital-inclusion-planning/

unevenly distributed throughout the area and among the focus group participants. Many residents were acutely aware of the exact areas within Southern Albemarle County where cell phone coverage drops, highlighting the persistent nature of the problem.

Other important themes that emerged from the focus group were affordability and accountability. Residents shared concerns about the expense of having a technician come to their homes, as well as the cost of upgrading their current hardware and/or extending the fiber optic cables to their homes, especially when their residences are located far from the road. Regarding accountability, the attendees felt that they were being left behind by Albemarle County and that resources are not equitably distributed. They added that they feel that Southern Albemarle County specifically doesn't get the same resources that the rest of the County does, highlighting equity and digital redlining concerns. Lastly, many expressed concerns over the lack of competition between broadband internet providers and the consequent lack of choice and affordability. Most focus group participants believe that the County should advocate on their behalf so that broadband providers extend and improve their service to southern Albemarle County.

Starr Hill Pathways

Among the participants from Starr Hill Pathways, digital literacy, connectivity, relevant content and services, and cell coverage were the most discussed themes. The focus group subjects expressed a general attitude of comfort and familiarity with the internet and technology in general. However, this attitude did not translate to perfect digital literacy. Although both participants indicated that they "would feel comfortable" downloading software or filling out an online application, they later admitted that they had difficulty locating the link (forwarded to their inbox by a parent) to log on to the focus group.

The participants also did not distinguish between traditional broadband internet mediums (i.e., wi-fi) and cellular data. Responding to a question about what they most frequently did online, both interviewees first said "text" (not a broadband-dependent activity) and then added, respectively, that they also used the internet to watch videos and use Snapchat. Similarly, questions about the ease and reliability of internet connections elicited responses strictly about hotspots ("I can always use my phone or something, because I have a hotspot.") as well as cellular range ("In the country, they have some spots… where you have no bars."). Overall, this could indicate a more pragmatic approach to connectivity that does not necessarily consider the specifics of where "internet" comes from. More speculatively, however, the interviewees' consistent interpretation of "internet" as cellular data could indicate a lack of consistent broadband access at home, with poor connectivity resulting in a tacit understanding of cellular data as a more reliable alternative to traditional broadband mediums.

Except for the note about non-universal (or uneven) cell coverage in "the country," the focus group participants did not note any major dissatisfactions with their broadband situations. Prompted to consider what they might want to be able to do on the internet that they currently could not, one participant replied that "I already have everything I want." It should be noted that participants may have been socially insulated from broadband issues experienced by their families – neither participant knew, for instance, how affordable their family's internet plan was. Similarly, interview subjects did not express information related to broader issues of equity and accessibility; again, this could have been due to age-related insulation from such complex social issues.

LIMITATIONS

Overall, four common issues across the focus groups limited the efficacy of this study: communication, incentives, sampling, and transcription.

Before beginning the focus group sessions, moderators informed participants that the conversation would be recorded and transcribed. Unfortunately, some group members decided to leave after being told this information, which was unexpected given that we had assumed participants were informed of the logistics beforehand. This issue highlighted the need for effective communication to avoid actions that could result in frustration, loss of trust, and disengagement among participants.

During the project's planning phase, we decided to provide all focus group participants with a $25 VISA gift card as an incentive to participate and as compensation for their time. However, in preparation for our first focus group, JABA, we learned that the incentive would not be available in time for the focus group, due to complications with the Institutional Review Board (IRB). As an alternative, we had to ask participants for their contact information following each focus group in order to mail them an incentive at a later date. For some focus groups this was made clear to participants before the discussion, but for others it was not disclosed until after the discussion. This issue resulted in delayed compensation for our participants, as well as frustration in the need for additional measures to receive their incentive. This issue resulted in delayed compensation for our participants and frustration at the need for additional measures to receive their incentive, which may have deterred some people from participating in our focus group conversations.

Analysis of our focus groups was completed through the transcription and coding of the recording of each focus group. Unfortunately, there were complications hearing the participants on the recordings due to low volumes and background noise, which led to difficulty in transcription. As a result, there were moments during the transcription process in which we had to make educated guesses within context to determine the specific words or phrases said.

Due to the nature of focus groups being an event that takes time and effort to attend, we found it difficult to achieve large numbers of participants. Though the Yancey School Community Focus Group was large, most focus groups were between two and six participants. Due to the small focus group size, there are likely perspectives or demographic groups that are missing from our analysis.

Some focus groups experienced limitations specific to their participants or target demographic populations. These limitations are discussed in greater detail below:

Starr Hill Pathways presented specific limitations due to the nature of our target participants. Starr Hill Pathways is a program that provides support and career exploration opportunities to youth in Charlottesville. The participants in this focus group were all middle school-aged, and the focus group took place over Zoom rather than in person, which led to brief answers from the participants and confusion over broader topics like affordability. Before the focus group, a coordinator explained that these students are not forthcoming with information and that we should tailor the questions to their experiences. However, as the questions needed to be standardized across all the focus groups, we were unable to do so. As a result, there were several questions that the respondents could not answer or required further clarification.

The lack of information voluntarily offered by the Starr Hill Pathways, and the participants' limited engagement may have been contributed by the limited exposure the respondents had to issues regarding broadband affordability. These limitations highlight the need for further research to better understand the specific needs and challenges faced by similar respondents in the community related to digital equity. For instance, while the interviewees expressed a general comfort and familiarity with technology, this did not necessarily translate to experienced digital literacy. In addition, the participants had difficulty locating the link to log on to the focus group. Not only may this have contributed to the lack of respondents in this focus group, but it also reveals the gaps in digital literacy that prevented them from commenting on issues that could have been experienced by similar respondents not included.

The Bright Stars focus group presented one of the biggest limitations in our study. Bright Stars is a preschool program by Albemarle County that provides additional support to families through social services, family coordinators, and case management. We had established Bright Stars' families as a desired focus group and scheduled a meeting time for March 30th, 2023, at 6 pm. We arrived to host the focus group at a local public middle school, but unfortunately no participants showed up. We later realized that this was the last day of school before spring break for these families, and this was likely the reason we did not have any participants. Due to this missed focus group, we are missing data from the Bright Stars families in our qualitative analysis. Because many of these families are considered low-income, live in rural areas, and have young children, they would be an important demographic to consider.

CONCLUSION

At the outset, we partnered with the Regional Digital Equity Plan coalition, Albemarle County's Broadband Accessibility and Affordability Office, and Dr. Christopher Ali to develop a deeper understanding of the state of digital equity and broadband access in Albemarle County. To do so, we conducted focus groups at several community organizations and gathering places: JABA, PHA, Yancey School Community Center, Starr Hill Pathways, and Bright Stars. The participants were demographically diverse, ranging in age, socioeconomic background, ethnicity, racial and gender identity, and participants joined us from homes across the County. Yet, despite the participants' various life experiences and lived realities, there were common themes and topics discussed in each focus group, primarily digital literacy, learning curves, affordability, connectivity, and digital inclusion.

The participants' willingness to share their stories gave important insight into how residents' on the ground experiences are impacted by digital equity disparities across Albemarle County. For example, some residents frequently suffer from unreliable internet access, while others avoid specific areas because they are known to have poor coverage. Participants also shared positive stories and offered helpful suggestions, such as referring other participants to affordable internet plans and offering insightful recommendations for educational programming. As the participants made clear, digital equity and broadband access are about more than just access to the internet. Removing barriers to access and improving digital literacy can have positive impacts on a wide range of issues, including educational outcomes, health care, increased access to vital social networks, and more.

Through analysis of the transcripts and identification of the shared experiences, struggles, and successes, our research informs the recommendations and next steps in this report to help develop the RDEP. We want to explicitly acknowledge that while our research and findings are valuable and inform actions the County should take, the work is ongoing. Issues of digital equity and broadband access should be reevaluated frequently to ensure Albemarle County is aware of residents' realities and working efficiently to improve access and affordability. Moreover, further research may help develop more extensive and specific programming for residents of the County. Increased communication and outreach may contribute towards better relationships between the County and residents.

NEXT STEPS

As a result of the focus groups, transcript coding, and analysis processes, we have identified several recommended next steps for Albemarle County and the Albemarle Broadband Authority. Broadly, these recommendations are focused on affordability, accessibility, educational programming, and appropriate follow-up with focus group participants.

1. **Develop digital literacy programming for children and adults, particularly senior residents.** The programming should be in the form of classes and summary "cheat sheets" that can be distributed to children's parents, guardians, or directly to senior residents. While the cheat sheets would not replace class programming, they could be relied on in the short term for answers to immediate or pressing simple, clarification questions (e.g., the difference between cellular

data and broadband, how to connect to the internet, and a list with contact information of internet providers that service the County). These programs must be flexible enough to adapt to updates and changes to broadband while meeting other goals, such as teaching digital literacy fundamentals, to be successful.

a. **Senior Residents:** There is demand for frequent (ideally daily or weekly), repetitive, patient, and gradual digital literacy education for senior residents in Albemarle County. In addition to focusing on digital literacy basics, the programming should also teach students about internet hygiene and safety with specific emphasis on identifying and avoiding potential scams and inappropriate or adult content. This programming should be offered at multiple locations throughout the County to maximize its impact, effectiveness, and accessibility. Potential locations could include JABA, PHA, or a central location identified and sponsored by the County. Transportation should be provided for adults who do not have access to a car (e.g., new bus routes to the program's location).

b. **Youth:** While the children who participated in our focus groups did not directly express a wish for this, their answers made it evident that there is a need for direct and basic digital literacy programming for children in Albemarle County. The County's public schools could implement this programming, which should focus on internet hygiene, safety, and the development of a fundamental understanding of how to interact with the internet.

2. **Albemarle County should provide easy-to-use desktop computers in central gathering spaces throughout the County, especially in areas with limited broadband access.** These spaces should be staffed with people with digital literacy skills who can assist residents and answer questions as necessary. In addition, transportation should be provided (e.g., a new bus route).

3. **Establish loaner pools of basic laptops (e.g.,** Chromebooks) for adult and youth residents:

a. **Adults:** This gives residents an affordable option to access a computer from home. If this is cost prohibitive, the County could implement an affordable rental fee to offset the costs.

b. **Youth:** These loaner pools should be facilitated through the public school system and reserved solely for students. Youth participants expressed that public schools require online schoolwork for all students. Therefore, there is an onus on the County to ensure that all households with students have access to broadband and hardware sufficient for completing schoolwork without hindering the broadband access or speed of other household users (e.g., siblings or parents who work from home).

4. **Collaborate with the area's internet providers and advocate on behalf of County residents:**

a. Establish initiatives to subsidize the cost or incentivize installing broadband infrastructure in regions without reliable broadband access.

b. Establish affordable pricing initiatives or subsidize the cost of internet access for specific sectors of County residents (e.g., low-income folks and residents in areas with poor broadband access). Moreover, advertise the affordable pricing options that do exist. For example, an older participant shared information about an affordable price plan offered by Xfinity for senior citizens. She mentioned, however, that most folks did not know about it because Xfinity does not advertise it.

c. Hire County technicians who can provide hardware and infrastructure maintenance at affordable costs for residents.

5. **Understand that not all folks have or use smart phones.** The County should promote and advertise alternative and traditional forms (e.g., mailing address, e-mail, landline), for identity confirmation, appointment scheduling, mailing list sign-ups, etc. Residents across the focus groups expressed annoyance that there is a universal assumption in public and private places that residents have and use smart phones. The County can counteract this assumption by not requiring smart phones and providing alternatives for residents within their own programs, as well as by advocating for the adoption of

additional methods in private businesses and places across the County.

6. **Ensure that all focus group participants received their incentives.** If any participant has not received their incentive, prioritize following up and sending it immediately. Since Albemarle residents in multiple focus groups expressed sentiments of feeling like they were being left behind by the County, this is an essential show of good faith that is important for relationship-building, and fundamentally the right thing to do as a thank you for their contributions and time.

7. **Ensure that all participants are offered a copy and given access to this report.** Since it directly concerns them and would not have been possible without their input, participants should have access to the RDEP Report.

8. **Advertise and make this report available to all Albemarle residents interested in learning more about broadband access and digital equity in the County.**

9. **We have also identified a few areas for further research that may aid the County in fully implementing our recommended steps and making progress beyond this report:**

    a. Some participants expressed apathy regarding using and learning how to use the internet. If Albemarle County is interested in increasing such residents' internet usage, further research into why these residents are apathetic is necessary.

    b. For the residents who expressed annoyance at the assumption that all people have and use smart phones, further research is necessary to understand what their preferred alternative method would be. Moreover, while alternative methods like e-mail, landline, and snail mail exist, it would be worthwhile to research whether there are even more options available.

    c. Some participants expressed feeling left behind by Albemarle County. Further research is necessary to identify how widespread this sentiment is (i.e., how many residents feel this way, and where are they?). After obtaining a deeper understanding of this issue, Albemarle County should establish a focused and sustained campaign aimed at these areas and residents to improve their broadband access, digital literacy, and relationship with the County.

Quantitative Analysis

The datasets for this project were obtained from both national and regional sources. The National Telecommunications and Information Administration's (NTIA) ACCESS BROADBAND metrics collected through the U.S. Census Bureau provided information related to select social and economic indicators at the County, tract and block group summary levels. Data from the Regional Digital Equity Plan Coalition (RDEP) through the Albemarle County Broadband Accessibility and Affordability Office (BAAO) and Albemarle Broadband Authority (ABBA) provided other regional statistics and surveys regarding school and housing administrative data as well as broadband indices related to an assortment of speed, provider and sentiment metrics.

DISAGGREGATION

Key Takeaways

- 88.9% of Albemarle County & Charlottesville city households already have access to broadband.
- For our analysis, we supplemented data provided by the Albemarle Regional Digital Equity Plan (RDEP) with data from the American Community Survey 5-year estimates (2017-2021).
- Using a set of independent variables, we ran a correlation analysis and 6 regression analyses to determine their relationship with our dependent variable, percent of households with broadband access. We also created a series of maps to visually and spatially display the data.
- For Albemarle & Charlottesville, as well as for counties surrounding Albemarle, there are statistically significant relationships between the amount of households that have broadband access and 1) individuals who at least have a Bachelor's Degree and 2) increasing median household income
- Due to data limitations, we were unable to examine 1) the relationship between broadband access and populations and 2) compare rural and urban regions. Both of these are good areas for further study.
- Our next step recommendations include digital literacy initiatives to improve broadband access and educational attainment, the implementation of community-owned networks, the development of a broadband infrastructure strategy, and the establishment of private-public partnerships.

Introduction

To better understand socioeconomic factors affecting broadband access and affordability in the region, our group's analysis focused on disaggregated demographic data from Albemarle County, the City of Charlottesville, and surrounding localities. Our analysis aimed to discover how variables such as race, income, and educational attainment impacted access to broadband and other digital technologies. This report will describe the methods we used to collect and investigate data, the results of our correlation and regression analyses, the limitations of those analyses, and suggestions for the next steps to address digital inequity in the region.

Data and Methods

Our main source of data was provided by the Albemarle Regional Digital Equity Plan (RDEP) which included various demographic data points by state, county, and tract for the entirety of the US. These data

tables each included percent of households with broadband as a key variable, along with other demographic variables such as educational attainment level, median home value and self-employment rate. Due to the limited demographic variables included in the RDEP data, we pulled additional demographic data points from the American Community Survey 5-Year Estimates (2017-2021). We narrowed all of our data to focus on Albemarle County and the City of Charlottesville, as well as all counties adjacent to Albemarle County.

We employed the following variables for our analysis:

- Broadband Access per Household (dependent variable)
- Non-Hispanic White Population
- Black Population
- Hispanic or Latino Population
- Population Below Poverty Level
- Median Household Income
- Median Home Value
- Highest Educational Attainment is High School Diploma or GED
- Highest Educational Attainment is Some College or Associate's Degree
- Highest Education Attainment is Bachelor's degree or higher
- Employment Rate
- Labor Participation
- Telework Population
- Self-Employed Population
- Households with Children Under 18
- Population With No Health Insurance Coverage

Once we had determined our list of variables, we began our statistical analysis using R and R Studio software to conduct our correlations and regressions. First, we ran a correlation to determine the relationship between each independent variable and broadband access by household. Then, we ran a series of regressions to further explore the relationship between these variables and broadband access in Albemarle County and Charlottesville City.

The first regression included all our independent variables and broadband access by household as the dependent variable in Albemarle County and Charlottesville. The second regression removed multiple independent variables (highest educational attainment is high school diploma or GED, highest educational attainment is some college or associate's degree, Non-Hispanic white population, Hispanic or Latino population, and labor participation) as they are likely multi-collinear. The coefficients for each variable in the second regression were then scaled and plotted. We ran a third regression with three select independent variables: highest educational attainment is bachelor's degree or higher, Black population, and median household income. Once again, the regression coefficients were scaled and plotted.

Next, we moved on to running regressions with the same independent variables,, but with data from counties and independent cities adjacent to Albemarle County. These include Louisa County, Orange County, Greene County, Fluvanna County, Buckingham County, Nelson County, August County, Rockingham County, Staunton City and Waynesboro City. The fourth regression analyzed all the independent variables, with percent households with broadband as the dependent variable, for all counties and independent cities adjacent to Albemarle County. For the fifth regression, we repeated what we did for the second regression and removed the same variables that are likely multi-collinear. We scaled and charted the coefficients from regression five. For our final and sixth regression, we analyzed the same three selected independent variables

(highest educational attainment is bachelor's degree or higher, Black population, and median household income) within the context of the adjacent counties and independent cities. Again, we scaled and charted these regression coefficients.

To supplement our statistical analysis, we also created a series of maps using ArcGIS Pro. We mapped each independent variable by census tract in both Charlottesville and Albemarle County, as well as our dependent variable, percent of households with broadband.

Analyses

Our quantitative analysis, which we conducted to disaggregate the Albemarle Regional Digital Equity Plan (RDEP) demographic data, consists of three distinct components: mapping, correlation analysis, and regression analysis. Each component provides distinct and important insights into broadband access and affordability in Charlottesville and Albemarle County. However, interpreting the combined analyses creates a much deeper and more thorough understanding regarding digital equity in the region. In this section, we will analyze each component and then discuss the important key takeaways and insights from our analyses.

Maps

In total, we created 11 maps highlighting both Albemarle County and Charlottesville. These maps provide essential context for our correlation and regression analyses, but they also convey important information on their own (see Figures 2-12). They detail the following:

- Percent of households with broadband access

In general, broadband access is concentrated in the center of Albemarle County. However, there are exceptions. One of the outermost census tracts is within the 95-100% access bracket, and there are a handful of census tracts in the core of the County, including Charlottesville, that are within the two lowest access brackets, 76-80% and 81-84%. Moreover, much of Charlottesville itself is only within the 85 – 90% and 91 – 94% access brackets, with just one census tract in the 95 – 100% bracket. While Albemarle County, including Charlottesville, has a relatively high floor of 76% of residents having access, most of the area falls within the 85 – 94% access range. Clearly, this indicates that there is a significant portion of the County and City that do not have access to broadband.

- Distribution of residents' racial identities (white, Black, Hispanic or Latino alone)

These three maps illustrate that Albemarle and Charlottesville are overwhelmingly non-Hispanic white. Indeed, the lowest bracket for percent of non-Hispanic white only population in a census tract is 32 – 58%. The lowest brackets for Black-only and Hispanic/Latino-only residents are 0 – 3% and 0 – 2% respectively. In short, the majority of the outlying and increasingly rural census tracts are majority white, and the County becomes increasingly diverse towards the center and Charlottesville. However, there are exceptions. The southernmost census tract has a relatively high proportion of Black residents, and there are census tracts within Charlottesville that have nearly 100% white residents.

- Educational Attainment (high school or GED, some college, bachelor or master's)

There are increased levels of educational attainment towards the center of the County and Charlottesville. Indeed, the overwhelming majority of residents with a bachelor's or master's degree live in Charlottesville or the immediately surrounding census tracts. While the maps indicate some variance and nuance, folks in Charlottesville and the center of the County are generally more likely to have a higher level of educational attainment. This may be impacted by the presence and location of the University of Virginia.

- Access to Health Insurance

Interestingly, there is no striking pattern regarding percent of population with health insurance. The census tracts with the highest percentages are outside of Charlottesville and scattered across the County, although a few border each other. Moreover, Charlottesville does not have a higher level of health insurance access than the rest of the County.

▶ Median Household Income and Median Home Value

Not surprisingly, these two maps, while not identical, resemble each other and present a pattern. Many of the census tracts within the highest brackets ($110,000 - $170,000 for median household income and $520,000 - $700,000 for median home value) are the same. Likewise, the census tracts within the lowest brackets ($20,000 - $50,000 for median household income and $0 - $240,000 for median home value) are the same. There are no census tracts that split drastically across these two variables (i.e., high median household income, low median home value or vice versa). Again, the center of the County and Charlottesville contains the wealthier census tracts. However, there are also census tracts with lower median household income and lower median home value in Charlottesville and the core of the County.

▶ Percent of Households with Children

In Albemarle County and Charlottesville, the lowest bracket for percent of family households with children younger than 18 years old is 68 – 83%. Therefore, in every census tract, the majority of family households have children younger than 18. However, there is not a particularly striking pattern for the disbursement of the rates.

In general, there is a trend of higher rates of broadband access, educational attainment, wealth, and concentration of white residents in the core of Albemarle County and in Charlottesville. However, as discussed, there are exceptions to this observation. For example, the census tracts in the south-central area of Charlottesville are an isolated pocket with lower rates of median household income, median home values, and low broadband access. Moreover, it is important to compare the maps detailing both percent of broadband access and the maps detailing the distribution of residents' racial identities to the variables displayed in the other maps. To help enable such comparisons, we conducted a correlation analysis of the variables.

Correlation Analysis

The correlation analysis displays the statistical relationship between our dependent variable, broadband access, and a set of 19 independent variables (Table 2). The analysis contains all the variables we used for the maps, as well as additional variables such as percent of residents living in poverty, percent of residents who telework, percent of residents who are employed, percent of residents who are self-employed, percent of labor force participation, percent of residents who are not insured, and total population. There are several variables with strong positive correlations to broadband access, including median household income (0.52), attainment of a bachelor's degree or higher (0.51), percent of households with children (0.37), median home value (0.33), and percent of residents who identify as non-Hispanic white (0.30). There are also several variables with strong negative correlations to broadband access, including attainment of a high school diploma (-0.46), percent of residents living in poverty (-0.45), and percent of residents who identify as black (-0.42). However, there are also several weak correlations, including total population and percent insured which are respectively correlated at -0.08 and 0.08 with broadband access. Here are the key takeaways:

The two independent variables with the strongest positive correlations are median household income and educational attainment of a bachelor's degree or higher. Their correlation scores are 0.52 and 0.51, respectively. On the surface, this makes sense, considering that higher wealth is likely to indicate more of an ability to afford broadband, and increased education and wealth may indicate the potential for reliance on broadband access for work and wealth accumulation. High education may also indicate that residents have experience with the internet and technology and understand their utility and importance.

The three independent variables with the strongest negative correlations are attainment of a high school diploma (-0.46), percent of residents living in poverty (-0.45), and percent of residents who identify as Black alone (-0.42). Considering that higher levels of educational attainment are positively correlated with broadband access, it is reasonable that lower levels of educational attainment

are negatively correlated with broadband access. This may have to do with lack of experience and education around broadband and/or accessibility issues resulting from lower educational attainment. Similarly, it is not surprising that poverty is negatively correlated with broadband access, given the cost of internet access and technological hardware. Moreover, the negative correlation between Black residents and broadband access is important. More research is required to fully determine the underlying issues, but it is not shocking that there is a negative correlation between broadband access and Black residents. Given the legacy and continuing repercussions of state-sanctioned discrimination, Black Americans experience disproportionately high rates of poverty and low rates of generational wealth.

There are several other independent variables with positive correlations to broadband access, including median home value (0.33), percent of residents who identify as Non-Hispanic white (.30), and percent of residents who telework (0.29). There are also a few other independent variables with negative correlations to broadband access, including associate's degree (-0.20) and percent of residents without health insurance (-0.14). Our regression analyses reveal more information about the relationship between broadband access, the dependent variable, and all the independent variables.

Regression Analyses

We completed six regression analyses that can be broadly organized into two categories. The first set, Regressions 1 – 3, analyzes broadband access in Albemarle County and Charlottesville, and the second set, Regressions 4 – 6, analyzes broadband access in the surrounding counties and independent cities for comparison. Each set includes:

- A regression analysis with all 15 independent variables
- A regression analysis with five independent variables removed to account for multicollinearity
- A regression analysis with just three select variables: percent of residents with bachelor's degree or higher, percent of residents who identify as Black, and median household income.

Regression Analysis Set 1 (Albemarle County and Charlottesville): In Regression 1, none of the independent variables have a p-value below 0.05 (Table 3). In fact, the lowest p-value is .10. Therefore, the results are not statistically significant, and we cannot rule out the null hypothesis. Moreover, the regression analysis has a high R^2 value of 0.57, meaning that 57% of the value of the percent of residents with broadband access can be explained by the impact of the independent variables. The R^2 value may be artificially inflated by the inclusion of so many variables, several of which are related and multicollinear (e.g., poverty and median household income, high school diploma and bachelor's or master's degrees).

In Regression 2, one of the independent variables has a p-value lower than 0.05 and is therefore statistically significant: percent of residents with a bachelor's degree or higher (Table 4). This variable's coefficient is 20.98, indicating that with each percentage increase in residents with such educational attainment, the measure of residents with broadband access increases by 20.98%. Like Regression 1, the R^2 value is relatively high at 0.53. Again, this is likely explained by the amount of variables; and, while there are fewer variables, some still overlap (e.g., median home value, median household income, and poverty).

In Regression 3, two of the three independent variables are statistically significant based on their P-Values: median household income and again percent of residents with a bachelor's degree or higher (Table 5). In this regression, the former's coefficient is less than 0.00, while the latter's is 15.85. The coefficient for the percent of residents with a bachelor's degree or higher decreased a bit between Regressions 2 and 3, but it still indicates that with each percentage increase of residents with such educational attainment, the measure of residents with broadband access increases by 15.85%. Despite conducting the regression with only three independent variables, the R^2 value is still relatively high at 0.37. In other words, the impact of the three variables account for 37% of the value of the percent of broadband access in Charlottesville and Albemarle County (Figure 13).

Within the regression analyses focusing on Charlottesville and Albemarle County, only two independent variables had statistically significant impacts. The most impactful independent variable was educational attainment of a bachelor's degree or higher.

Regression Analysis Set 2 (Surrounding Counties and Independent Cities): In Regression 4, the R^2 value is very high at 0.66 (Table 6). In this analysis, there are four statistically significant independent variables: percent of residents with a high school degree or equivalent (coefficient of 54.56), percent of residents with some college (coefficient of 65.70), percent of residents with a bachelor's degree or higher (coefficient of 55.76), and median household income (coefficient less than 0.00). Furthermore, median household income again had a negligible coefficient with a value less than 0.00. So, while each increase in median home value is having an impact on broadband access, its impact is minute. This may be because the units are single dollars whereas median home income is measured in the range of tens to hundreds of thousands of dollars. Therefore, median home income would have to increase by more than just a single unit (i.e., $1) to see a tangible effect.

Like Regression 2, in Regression 5, the two statistically significant values are percent of residents with a bachelor's degree or higher and median household income (Table 7). The former's coefficient is 16.43, and the latter's is less than 0.00, which is also similar to the regressions in set 1. The R^2 continues to be high at 0.54. Lastly, in Regression 6, the R^2 decreases slightly, but remains high at 0.48 (Table 8). Again, percent of residents with a bachelor's degree or higher and median home income are statistically significant. Their respective coefficients are 18.64 and less than 0.00 (Figure 14).

There were more statistically significant independent variables in the second set of regression analyses than in the first, but they are still relatively limited. Moreover, the R^2 values were low across all six regressions. Still, as discussed within the context of the correlation analysis and maps, it makes sense that median household income and educational attainment impact broadband access, even if their impact is relatively slight.

Analysis Summary

Unfortunately, due to the lack of statistically significant independent variables, the insights from the regression analyses are somewhat limited. However, the R^2 values across the six regressions were all relatively high. Moreover, the statistically significant independent variables (e.g., percent of residents with bachelor's degree or higher and median household income) do convey important information and reveal relevant relationships, especially in conjunction with the maps and correlation analysis. In short, the two independent variables with the most significant impacts on broadband access were percent of residents with a bachelor's degree or higher and median household income. Regarding median household income, the fact that its coefficient is less than 0.00 is misleading. While on the surface it indicates a negligible impact, the coefficient is likely skewed by the scaled unit. In other words, the unit for median household income is dollars. So with each increase in $1, the impact is minute. However, if the units were scaled up to be hundreds or thousands of dollars, the coefficient would likely be larger and each increase in median household income would have a larger impact on broadband access. Indeed, this is corroborated by the scaled coefficients, which show that median household income has a larger relative impact that percent of residents with bachelor's degree or more. Our analyses and research reflect the complicated nature of broadband access and the varied social and geographic realities in Albemarle County, Charlottesville, and the surrounding independent cities and counties. Our work highlights the variables, particularly those reflecting racial identity, wealth (i.e., median home value and median household income), and educational attainment that correlate with and have an impact on broadband access here in central Virginia.

Limitations

In the mapping and regression analysis we were limited by the availability of data. For example, we were initially curious if there was any relationship between broadband access and population density, but the American Community Survey data we had access to via tidycensus lacked area data for tracts. Relatedly, we also considered analyzing the relationship between rural and urban areas in broadband access by creating a binary variable to control for the difference between rural and urban areas (by picking a certain population density threshold above which we would code a census tract as urban). Unfortunately, because this analysis depends upon the availability of population density data, we were unable to add this component to our analysis.

The limited number of observations (census tracts) in the counties and independent cities immediately surrounding Albemarle County also made a more granular analysis of the difference in broadband connectivity patterns between Albemarle County and Charlottesville and adjacent counties and municipalities more difficult. For example, we wanted to run a regression analysis on each surrounding county, but because of the limited number of census tracts in any one county, regression analysis could not be performed because of the insufficient number of observations.

Another limitation of our research is the fact that a relatively small percentage of households in Albemarle County and Charlottesville do not have access to broadband internet. For example, 89.3% of Albemarle County residents and 87.8% of Charlottesville residents have broadband access. Because our analysis focused on a relatively small percentage of the population, we have fewer observations to draw from–which in turn affected the statistical significance of our results.

Like every form of analysis that seeks to tease out a relationship between subsets and characteristics of a population and an outcome of interest (in this case, broadband access), there is some amount of directional uncertainty in relationships. For example, our analysis found a statistically significant positive relationship between median household income and broadband access. Although it is probable that higher incomes yield higher broadband access because affordability is less of a barrier for high-income households, there is at least some uncertainty that this is indeed the direction of the relationship. Perhaps there is another variable that we have not controlled for that is related to both increased median income and increased broadband access such that there is not a direct relationship between median income and broadband access. Or perhaps it is the broadband access that, in turn, increases a household's median income.

Next Steps

Our study provides substantial insights into the relationship between broadband connectivity and socioeconomic variables in Albemarle County and Charlottesville; however, there are limitations that should be addressed in future research. As previously mentioned in the limitations section, more granular data would allow for a more detailed examination in future studies to better understand how socioeconomic factors impact broadband availability.

Digital literacy initiatives could improve broadband access and educational attainment. Digital literacy programs can be provided online or in-person. Basic knowledge of computers, introductory software courses, and online job recruitment strategies may be offered. These steps could encourage individuals to enroll in online courses, helping individuals pursue economic mobility.

Community-owned networks are another consideration to bridge the digital divide and boost economic growth.These networks are internet services operated by community organizations, enabling affordable access and enhancing local control and decision-making in underserved communities. Typically, non-profit organizations administer community-owned networks such as Free-Nets and Wired Cities.

Additionally, a broadband infrastructure strategy may improve internet access in areas with limited services.This technique enables communities to identify infrastructure improvements, such as the installation of new fiber optic connections and the expansion of

existing networks in underserved areas. Municipalities might also incentivize broadband providers to service underserved areas while penalizing those that do not.

Establishing partnerships between the public and private sectors may boost internet access in terms of speed quality. Private companies may be able to assist governments in expanding broadband infrastructure and services to underserved areas. In addition, partnerships may assist local governments in expanding internet access when insufficient funds or expertise exist.

Overall, this study identifies limitations to be addressed in future research. Additional data is required to understand how demographics impact broadband access in Albemarle County and Charlottesville. In addition to digital literacy, community-owned networks and public-private partnerships might improve internet access and educational and economic opportunities in underserved regions. If explored and implemented, these options could help bridge the digital divide and boost internet access.

Conclusion

In conclusion, our correlation analysis shows that socioeconomic variables such as higher median household incomes, median home values, and percentages of residents identifying as non-Hispanic white are positively correlated with broadband access. Other positively correlated factors include attaining at least a bachelor's degree and having children in a household. However, our regression analysis found educational attainment and median household income were the only statistically significant variables. The maps included in this report show that, with some exceptions, people who fall under the socioeconomic categories associated with broadband access are generally concentrated in the center of Albemarle County and the City of Charlottesville. They suggest that efforts to improve broadband access should focus on rural Albemarle County and areas of Charlottesville with higher concentrations of people whose socioeconomic profiles are not associated with greater broadband access.

We acknowledge that our analysis has some limitations. We note that because nearly 90% of Albemarle and Charlottesville residents have broadband access, our analysis was limited to a relatively small percentage of the population, which lessened the statistical significance of our results. Additionally, we found some directional uncertainty in the relationship between median household income and broadband access, suggesting that there may be other variables that we could have analyzed to explain this uncertainty.

Finally, we have identified strategies the region should take to improve broadband access rates. These include digital literacy initiatives, community-owned broadband networks, a regional broadband infrastructure improvement strategy, and new public-private partnerships. We hope that our analysis will contribute to a growing body of research on the importance of broadband access and ways to address the digital divide.

Table 1. Variables used by Disaggregation group

Topic	Value	Albemarle County	City of Charlottesville	Region Combined
Population	Total Population	111,438	46,597	158,035
Broadband	Households with Broadband	89.30%	87.80%	88.9%
Race	Non-Hispanic White	76.4%	65.4%	73.2%
	Black	8.8%	18.0%	11.5%
	Hispanic or Latino	5.9%	5.7%	5.8%
Economic	Below Poverty	6.9%	23.1%	11.7%
	Median Household Income	$90,568	$63,470	$81,093
	Median Home Value	$386,400	$355,800	$372,380
Education	High School or GED	14.7%	15.9%	15.0%
	Some College/ Associates	19.7%	18.6%	19.4%
	Bachelor's or More	58.9%	58.0%	58.7%
Employment	Employment Rate	96.7%	95.7%	96.4%
	Telework	15.5%	12.7%	14.6%
	Self-Employed	13.6%	12.6%	13.3%
Children	With Children Under 18	26.9%	22.0%	25.4%
Insurance	No Health Insurance Coverage	5.7%	7.1%	6.1%

Source: American Community Survey 5-Year Estimates (2017-2021)

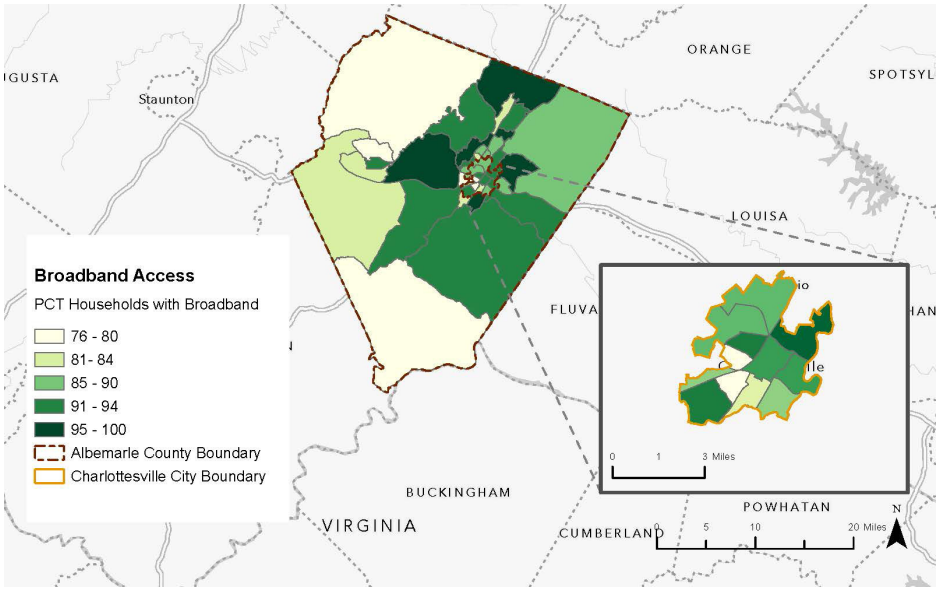

Figure 2. Percent of census tracts with broadband access

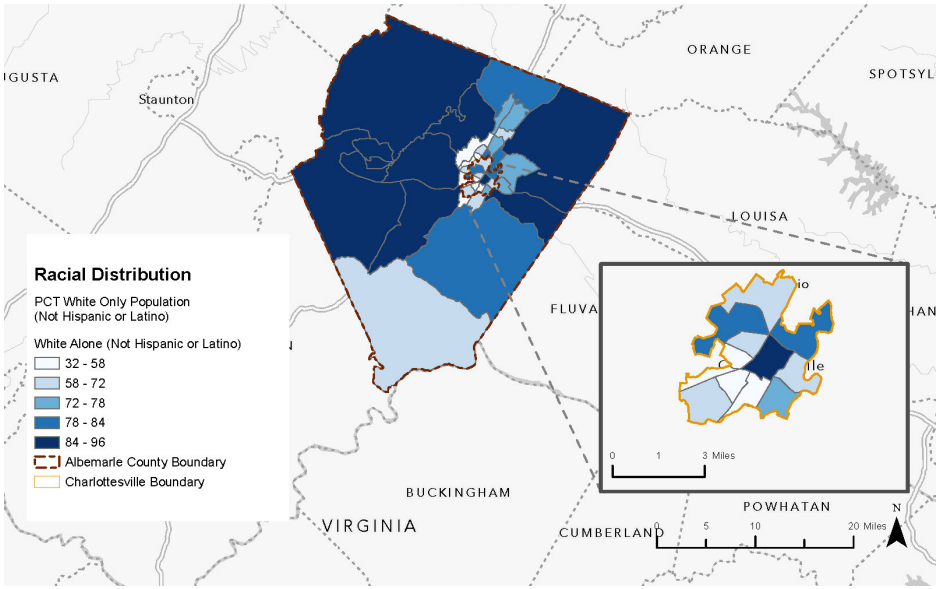

Figure 3. Percent of census tracts that identify as Non-Hispanic White

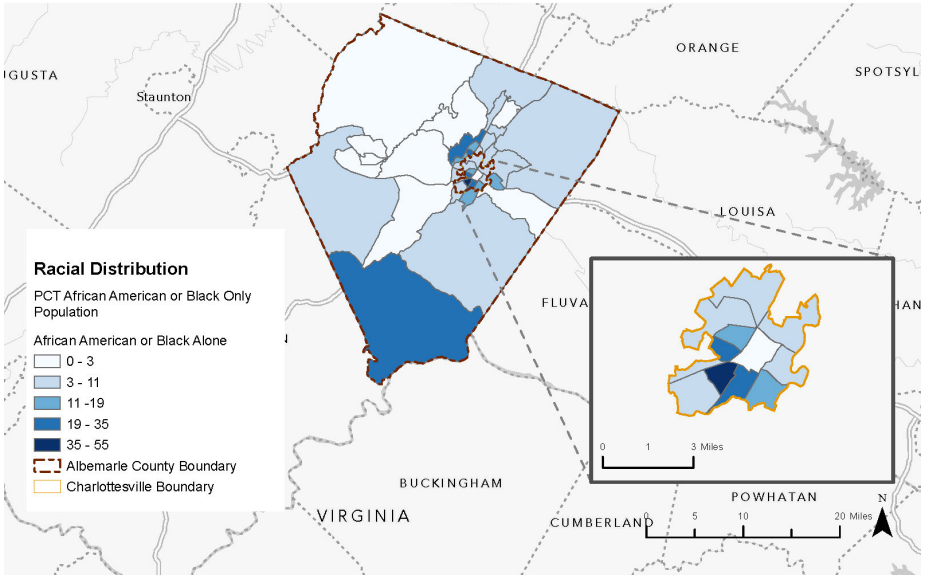

Figure 4. Percent of census tracts that identify as Black alone

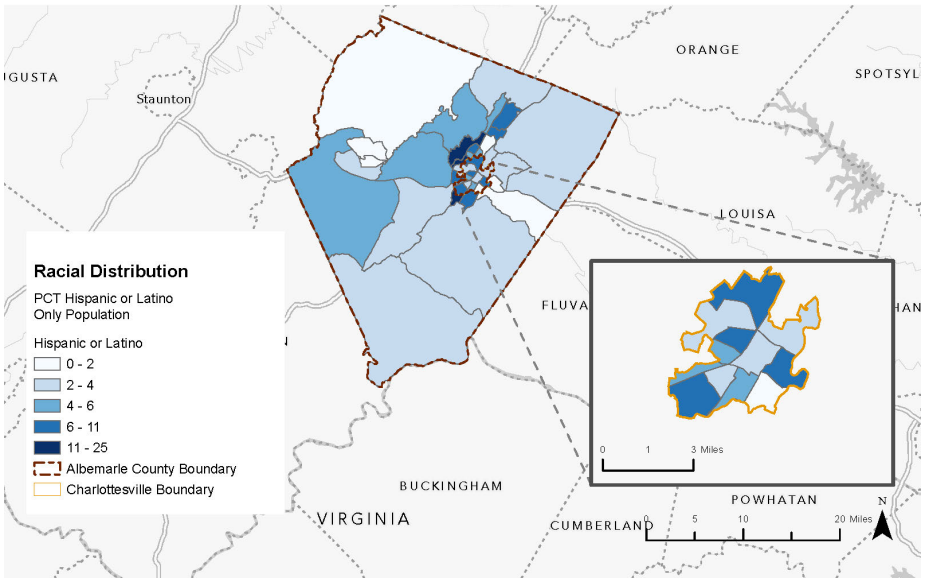

Figure 5. Percent of census tracts that identify as Hispanic

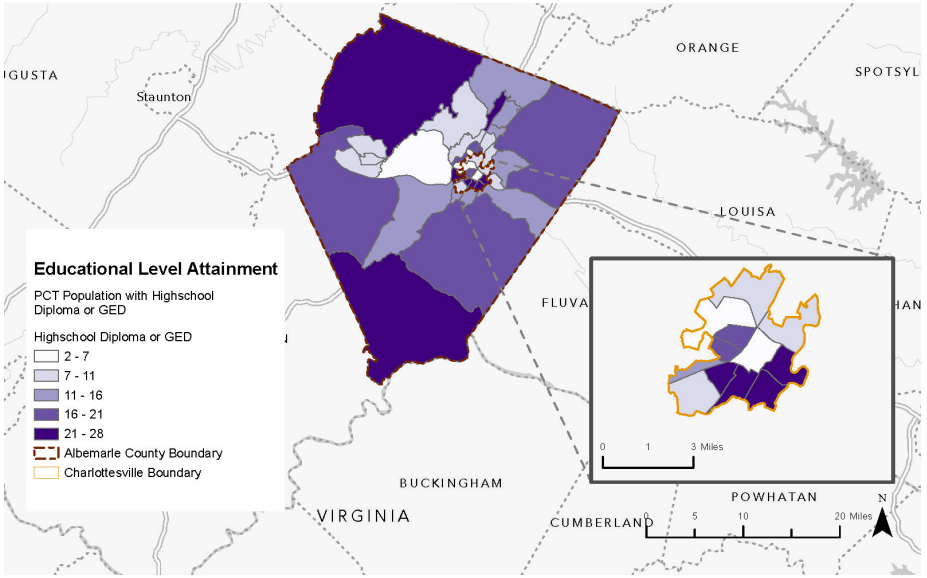

Figure 6. Percent of census tracts whose highest educational attainment is high school or GED

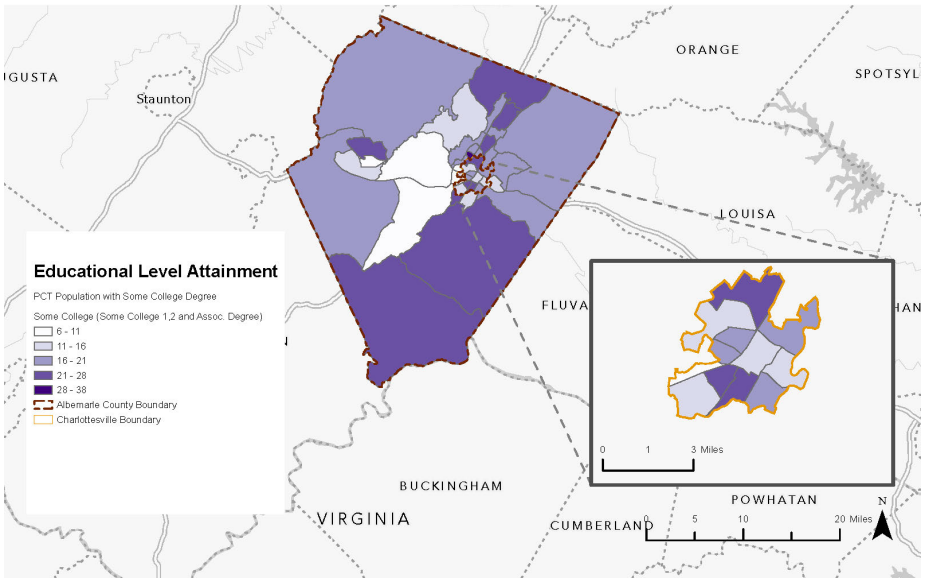

Figure 7. Percent of census tracts whose highest educational attainment is some college or Associate's degree

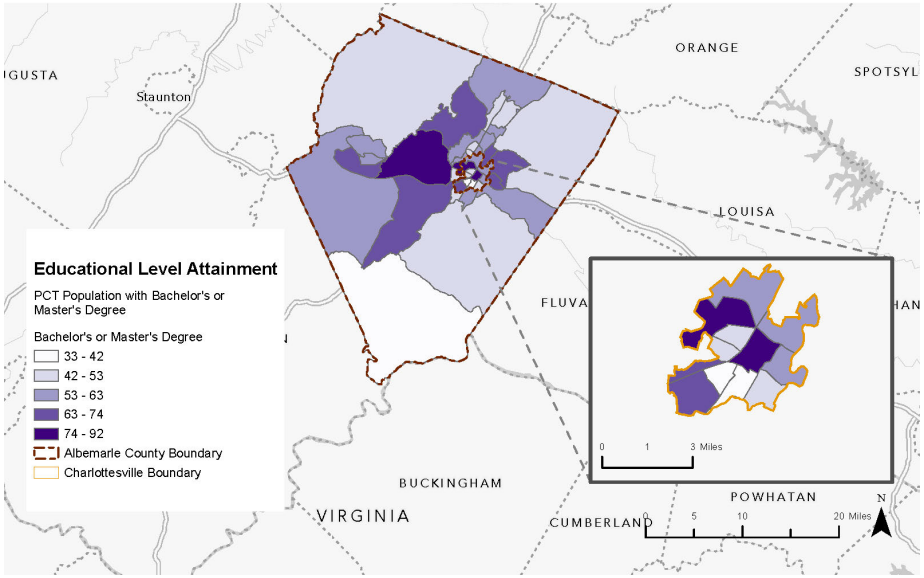

Figure 8. Percent of census tracts whose highest educational attainment is a Bachelor's degree or higher

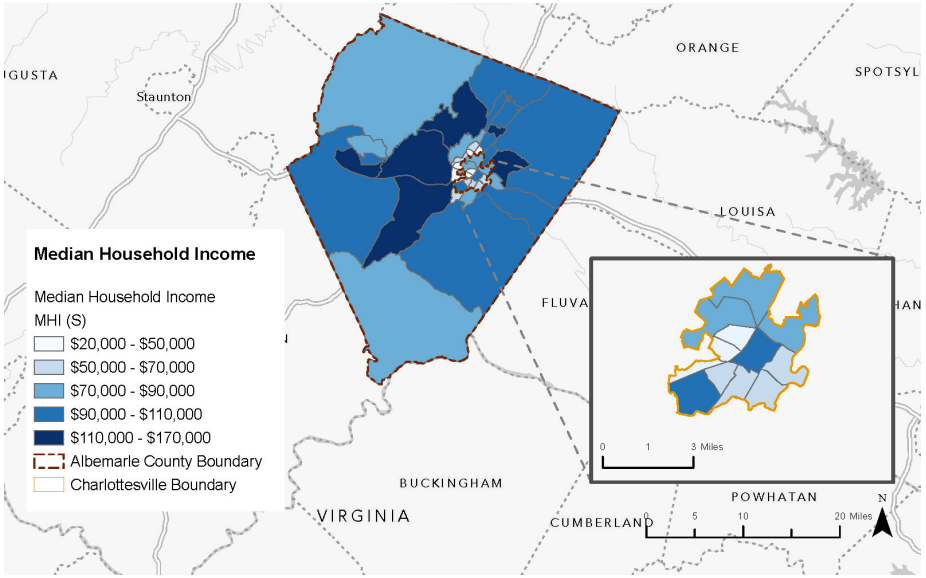

Figure 9. Median household income by tract

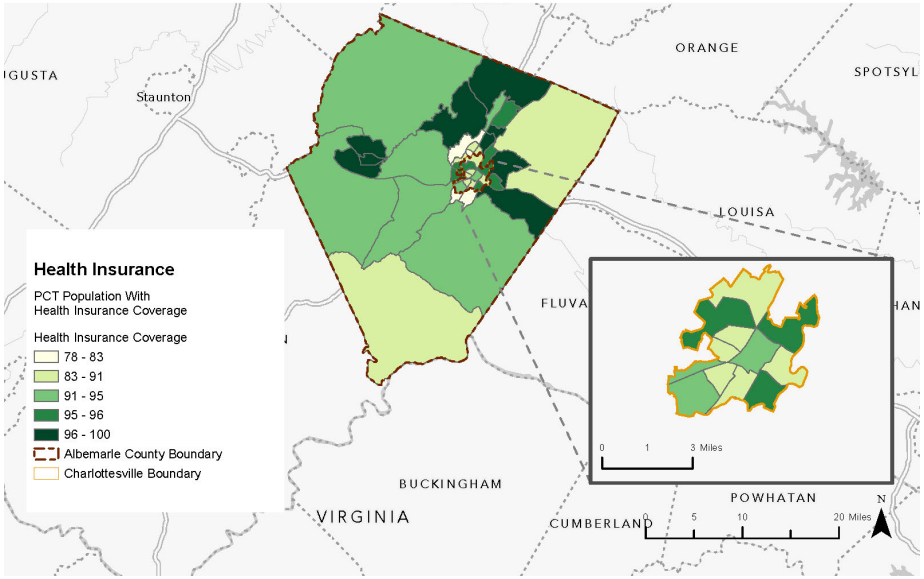

Figure 10. Percent of census tracts who have health insurance coverage

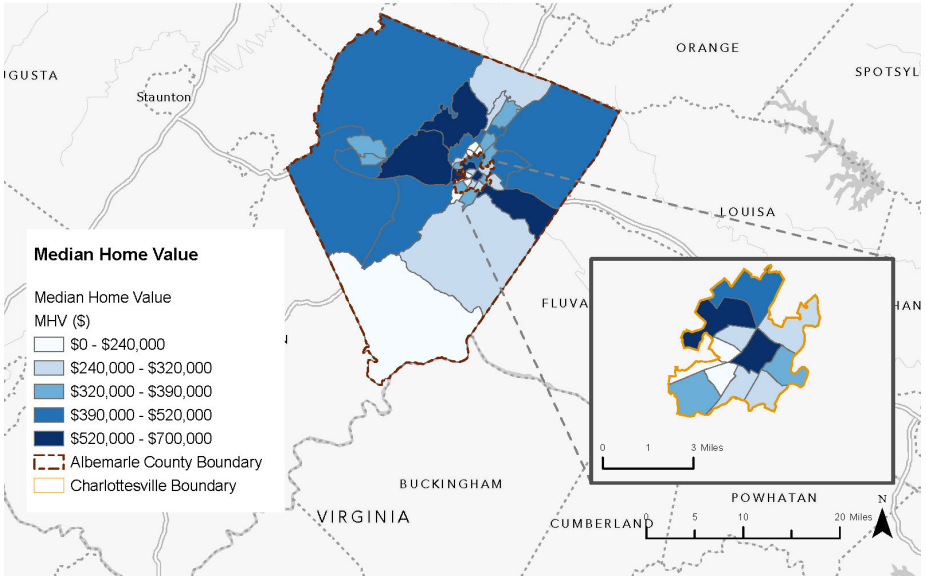

Figure 11. Median home value by tract

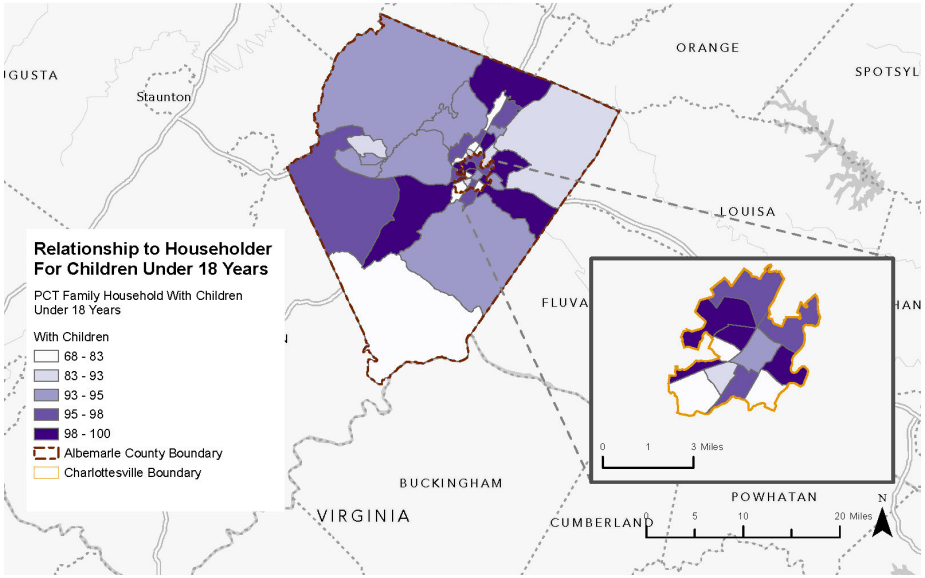

Figure 12. Percent of census tracts with children under 18 years

Table 2. Correlation between % households with broadband and selected ACS variables

Independent Variables	Correlation with % HH with Broadband	Absolute Value of Correlation
Median Household Income	0.52	0.52
Bachelor's Degree or More	0.51	0.51
High School Diploma	-0.46	0.46
% High School or GED	-0.45	0.45
% Poverty	-0.45	0.45
% Black	-0.42	0.42
% with Children	0.37	0.37
Median Home Value	0.33	0.33
% Non-Hispanic White	0.3	0.3
% Telework	0.29	0.29
% Employed	0.28	0.28
% Labor Force Participation	0.27	0.27
Associate's Degree or Some College	-0.25	0.25
% Self-Employed	0.23	0.23
Associate's Degree	-0.2	0.2
% Not Insured	-0.14	0.14
% Hispanic	0.11	0.11
Total Population	-0.08	0.08
% Insured	0.08	0.08

Table 3. Regression 1: Broadband Access in Albemarle County and Charlottesville

Observations: 39

R-squared = 0.573782

Adj. R-squared = 0.295813

Independent Variable (at tract level)	Coefficient.	Standard Error	p-value
Intercept	83.62	48.95	0.1
Percent with High School Degree or GED (as highest ed attainment)	-18.08	34.07	0.6
Percent with Some College	-15.61	30.27	0.61
Percent with Bachelors or more	9.22	24.76	0.71
Percent Black	-5.9	27.87	0.83
Percent Non-Hispanic White	-6.17	21.18	0.77
Percent Hispanic	19.88	41.92	0.64
Percent Insured	-12.4	32.55	0.71
Percent with Children	13.38	14.77	0.37
Percent Employed	-0.17	0.74	0.82
Percent Labor Participation	0.25	0.71	0.73
Percent Self-Employed	0.02	0.3	0.94
Percent Telework	0.17	0.21	0.41
Percent Poverty	-0.12	0.13	0.37
Median Home Value	0	0	0.36
Median Household Income	0	0	0.22

Table 4. Regression 2: Broadband Access in Albemarle County and Charlottesville (removed several variables that could be multi-collinear)

Observations: 39

R-squared = 0.537796

Adj. R-squared = 0.372723

Independent Variable	Coefficient	Standard Error	p-value
Intercept	86.99	23.42	0
Percent with Bachelor's Degree or More	20.98	9.27	0.03**
Percent Black	-8.22	10.77	0.45
Percent Insured	-33.09	20.06	0.11
Percent with Children	12.02	12.25	0.33
Percent Employed	0.12	0.12	0.31
Percent Self-Employed	-0.02	0.28	0.94
Percent Teleworking	0.13	0.19	0.5
Percent in Poverty	-0.06	0.11	0.59
Median Home Value	0	0	0.34
Median Household Income	0	0	0.21

*** significant at the 95% confidence level*

Table 5. Regression 3: Broadband Access in Albemarle County and Charlottesville (with select independent variables)

Observations: 39

R-squared = 0.37

Adj. R-squared = 0.32

Independent Variable	Coefficient	Standard Error	p-value
Intercept	73.05	5.83	0
Percent with Bachelor's Degree or More	15.85	7.88	0.05**
Percent Black	-2.07	9.75	0.83
Median Household Income	0	0	0.03**

*** significant at the 95% confidence level*

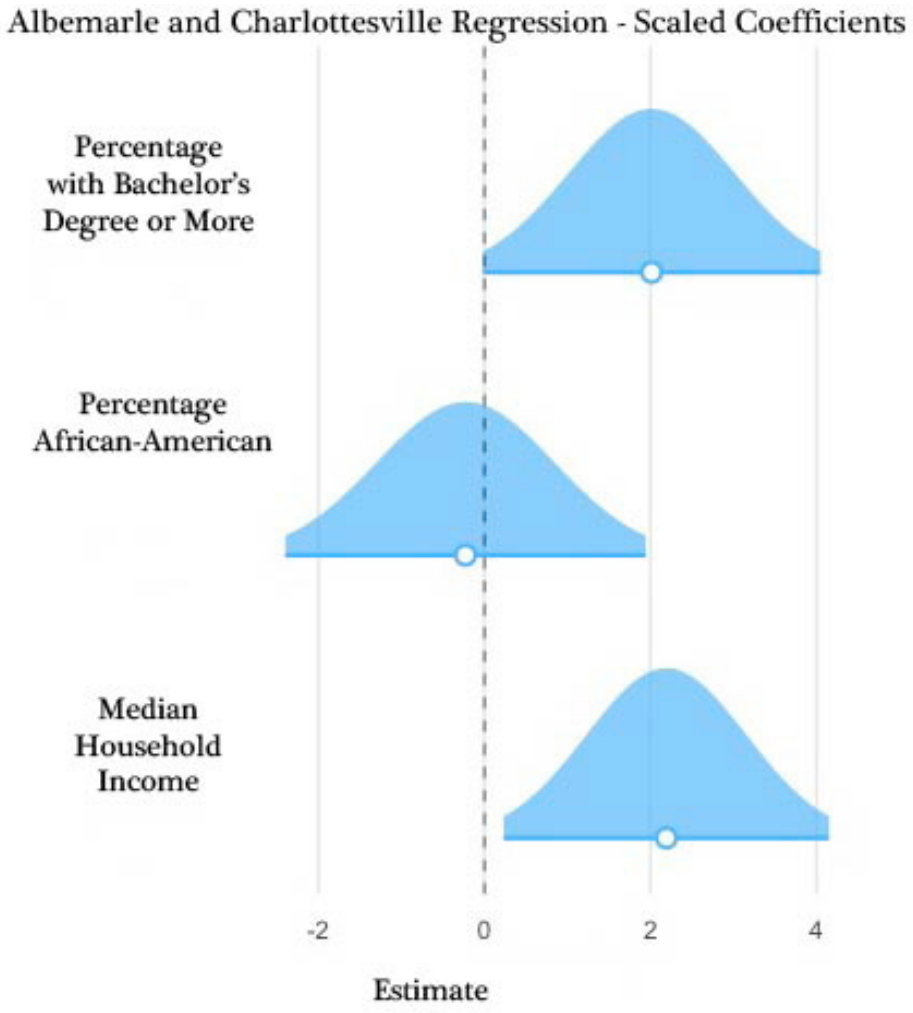

Figure 13. Scaled regression coefficient graph for Albemarle/Charlottesville

Table 6. Regression 4: Broadband Access in Counties and Independent Cities Surrounding Albemarle County

Observations: 87

R-squared = 0.657160

Adj. R-squared = 0.584729

Independent Variable	Coefficient	Standard Error	p-value
Intercept	49.57	27.48	0.08
Percent with High School Degree or GED (as highest ed attainment)	54.56	15.68	0.00***
Percent with Some College	65.7	15.46	0.00***
Percent with Bachelors or more	55.76	13.68	0.00***
Percent Non-Hispanic White	-36.64	23.55	0.12
Percent Black	-43.04	27.01	0.12
Percent Hispanic	-9.92	25.89	0.7
Percent Insured	-7.44	10.96	0.5
Percent with Children	-2.23	6.62	0.74
Percent Employed	0.04	0.36	0.91
Percent Labor Participating	0.09	0.39	0.82
Percent Self-Employed	-0.17	0.21	0.4
Percent Teleworking	-0.12	0.16	0.46
Percent in Poverty	0.07	0.14	0.61
Median Home Value	0	0	0.16
Median Household Income	0	0	0.02**

*** significant at the 95% confidence level*

**** significant at the 99% confidence level*

Table 7. Regression 5: Broadband Access in Counties and Independent Cities Surrounding Albemarle County (removed several variables that could be multi-collinear)

Observations: 87

R-squared = 0.54

Adj. R-squared = 0.37

Independent Variable	Coefficient	Standard Error	p-value
Intercept	47.89	11.46	0
Percent with Bachelors or more	16.43	7.79	0.03**
Percent Black	-5.25	9.1	0.57
Percent Insured	3.41	11	0.76
Percent with Children	1.25	7.04	0.86
Percent Employed	0.13	0.09	0.17
Percent Self-Employed	-0.38	0.22	0.09*
Percent Telework	-0.16	0.17	0.35
Percent in Poverty	0.12	0.89	0.38
Median Home Value	0	0	0.27
Median Household Income	0	0	0.00**

** significant at the 90% confidence level*

*** significant at the 95% confidence level*

**** significant at the 99% confidence level*

Table 8. Regression 6: Broadband Access in Counties and Independent Cities Surrounding Albemarle County (with select dependent variables)

Observations: 87

R-squared: 0.48

Adj. R-squared: 0.46

Independent Variable	Coefficient	Standard Error	p-value
Intercept	61.65	3.24	0
Percent with Bachelor's Degree or More	18.64	5.65	0.00***
Percent Black	-11.34	7.4	0.13
Median Household Income	0	0	0.00***

*** *significant at the 99% confidence level*

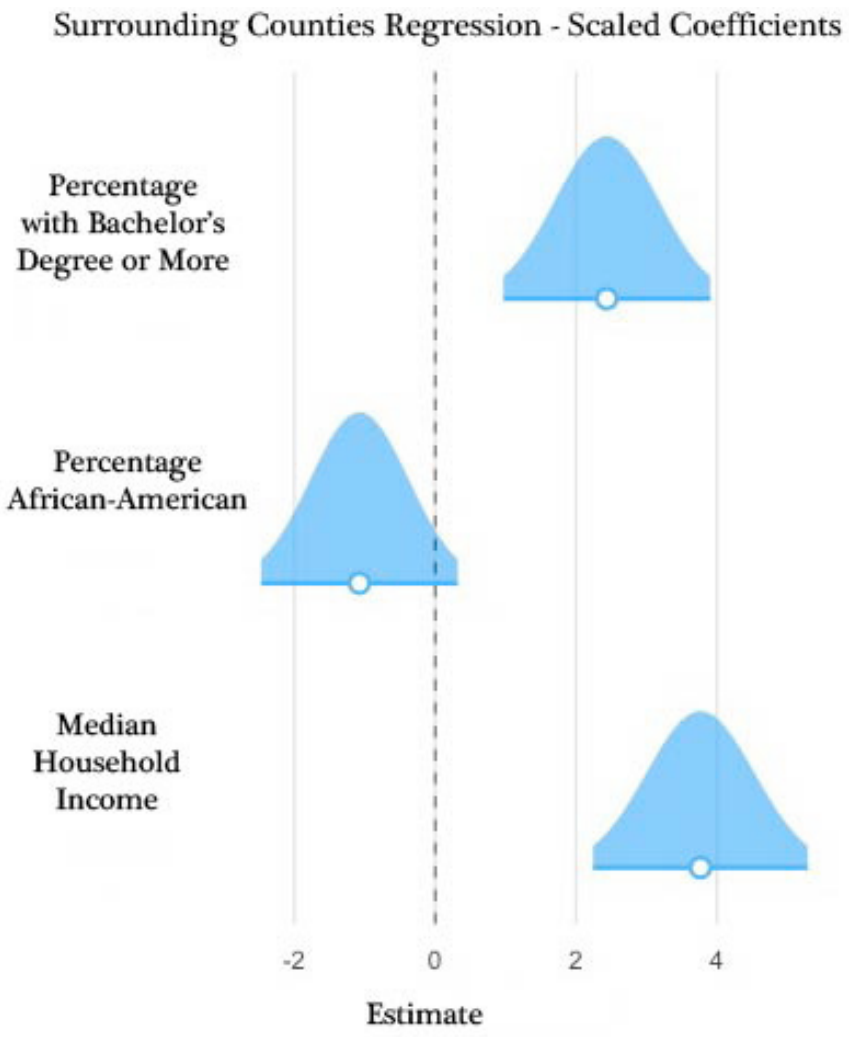

Figure 14. Scaled regression coefficient graph for the surrounding Counties and independent Cities

SPEEDS AND PREFERENCES

Key Takeaways

- Upload and download speeds were statistically significant predictors of 4 sentiment scores: speed, reliability, customer service, and competition.
- Ping was a statistically significant predictor of only one sentiment score: reliability.
- Jitter was not a statistically significant predictor of any sentiment scores.
- Upload speed was the only statistically significant predictor of the affordability sentiment score.
- Faster upload and download speeds improve how people feel about their internet service overall.
- In the development of a Regional Digital Equity Plan (RDEP), Albemarle County and the coalition for broadband equity should prioritize improving upload and download speeds in areas with poor service.

Introduction

The analysis of relevant broadband data is paramount for the RDEP's ability to understand the fundamental and underlying relationships surrounding digital equity throughout Albemarle County and the City of Charlottesville. The main focus of our analysis was to investigate the connections between broadband speed metrics and user sentiments scores. Specifically, the degree to which measurable criteria such as upload and download speeds, ping, and jitter predict responses and attitudes to survey scores regarding speed, reliability, customer service, competition, and affordability.

Data and Metrics

R Studio software was used to run correlation and regression analyses between sentiment survey and RDEP speed test data. R Studio was also used to produce descriptive statistics for both speeds and sentiment survey data (Figure 15). A graph (Figure 16) was also created mapping speeds by provider, relative to an adequate speed of 25 mbps[1]. Five categories of sentiment survey data were chosen as the dependent variables - Speed Score, Reliability Score, Customer Service Score, Competition Score, and Affordability Score. Our independent variables consisted of four RDEP speed test categories – Upload Speed, Download Speed, Jitter, and Ping. Five multivariate regressions were run then between the five dependent variables and the four independent variables to see if there were any relationships, significant or not, between sentiment score data and speed score data.

Results - Regression Analysis

We performed 5 regression analyses with each individual sentiment category (speed, reliability, customer service, competition, and affordability) as the response variables and the speed variables (upload speed, download speed, ping, jitter) as the predictor variables. Figure 17 contains the coefficients for each regression.

For the speed sentiment score regression with an estimate value of 2.34, 14.3% of the variability in speed sentiment score can be explained by download speed, upload speed, ping, and jitter (R^2 = 0.142700). However, ping and jitter were not statistically significant predictors of speed sentiment score. For every one increase in download speed, speed sentiment score increases by 0.0034. For every one increase in upload speed, speed sentiment score increases by 0.0022 (Table 9). Individual effect plots for each predictor variable are shown in Figure 18.

For the reliability sentiment score regression with an estimate value of 2.24, 18.5% of the variability in reliability sentiment score can be explained by download speed, upload speed, ping, and jitter (R^2 = 0.185488). However, jitter was not a statistically significant predictor

1 "Broadband Speed Guide." 2022. July 18, 2022. https://www.fcc.gov/consumers/guides/broadband-speed-guide

of reliability sentiment score. For every one increase in download speed, reliability sentiment score increases by 0.0033. For every one increase in upload speed, reliability sentiment score increases by 0.0041. For every one increase in ping, reliability sentiment score decreases by 0.0014 (Table 10). Individual effect plots for each predictor variable are shown in Figure 19.

For the customer service sentiment score regression with an estimate value of 2.09, 13.7% of the variability in customer service sentiment score can be explained by download speed, upload speed, ping, and jitter (R^2 = 0.137052). However, ping and jitter were not statistically significant predictors of customer service sentiment score. For every one increase in download speed, customer service sentiment score increases by 0.0027. For every one increase in upload speed, customer service sentiment score increases by 0.0035 (Table 11). Individual effect plots for each predictor variable are shown in Figure 20.

For the competition sentiment score regression with an estimate value of 1.62, 4.3% of the variability in competition sentiment score can be explained by download speed, upload speed, ping, and jitter (R^2 = 0.043022). However, ping and jitter were not statistically significant predictors of competition sentiment score. For every one increase in download speed, competition sentiment score increases by 0.0011. For every one increase in upload speed, competition sentiment score increases by 0.0025 (Table 12).Individual effect plots for each predictor variable are shown in Figure 21.

For the affordability sentiment score regression with an estimate value of 3.53, 1.5% of the variability in affordability sentiment score can be explained by download speed, upload speed, ping, and jitter (R^2 = 0.014705). However, download speed, ping, and jitter were not statistically significant predictors of affordability sentiment score. For every one increase in upload speed, affordability sentiment score increases by 0.0028 (Table 13). Individual effect plots for each predictor variable are shown in Figure 22.

Discussion

Our analyses made sense because it's intuitive that people's sentiment scores would increase with higher upload and download speeds. Our analyses showed that this trend existed among all five sentiment score regressions with upload speed, and four of our sentiment score regressions with download speed. Perhaps, the only statistically significant result that didn't make sense was that reliability sentiment score decreased, albeit small, for every one increase in ping. Additionally, the relatively low R^2 values for each of the regressions are expected given that many different variables and factors influence how people feel about their broadband service.

With regards to questions relevant to the Regional Data Equity Plan, an important takeaway is that faster upload and download speeds positively affect how people feel about their speed, reliability, customer service, competition, and affordability. If sentiment is identified as an important metric of future success, then Albemarle County and focus on improving upload and download speeds for its residents. Additionally, if an area of the county has historically and currently, felt left behind and underserved compared to the rest of the county (such as Southern Albemarle County identified by the Yancey School Community Center Focus Group), then improving upload and download speeds in those regions to heal prior feelings of stigma could be an important priority for decision makers to consider.

Limitations

Several limitations inhibit the broad application of these results. For example, there are other factors that influence sentiment scores that were not included in this analysis. The independent variables used in the regression models correspond mainly to speed and reliability metrics (upload/download, ping, jitter) and should not be relied on to predict overall affordability, customer service, or competition sentiment scores. Other predictor variables such as actual internet service prices and customer service response rates should be incorporated to develop a more comprehensive regression analysis of sentiment scores.

Another limitation of the analysis was that we excluded survey data that didn't have responses for both raw speed and sentiment survey data. The original dataset contained 4,495 observations, but we included only 739 observations in the regression analyses. Some of the observations provided only had raw speed data available so we excluded those observations.

Next Steps

The next steps for the Regional Digital Equity Plan should be focusing on how to improve and/or provide internet services with faster upload and download speeds because they were the most reliable predictors of sentiment scores. Another step that might be taken for the RDEP is to collect other relevant data from consumers, such as monthly internet expenses, to be able to better analyze the costs that consumers are paying for different levels of internet quality. It might also be useful to identify the types of internet (i.e. fiber, cable) that each service provider uses to better determine and analyze the impacts of the different types of internet on user satisfaction. There also could be further regressions using different variables to predict sentiment scores, as speed metrics did not result in the most reliable predictors of sentiment scores.

Conclusion

In conclusion, the findings from this regression analysis identify upload and download speeds as areas of opportunity for the RDEP development to increase positive sentiments regarding internet access and service quality among residents in Albemarle County and Charlottesville. To obtain more comprehensive information about what factors impact broadband sentiments, the RDEP team and the coalition should collect additional data and perform more robust

Figure 15. Average sentiment scores from survey

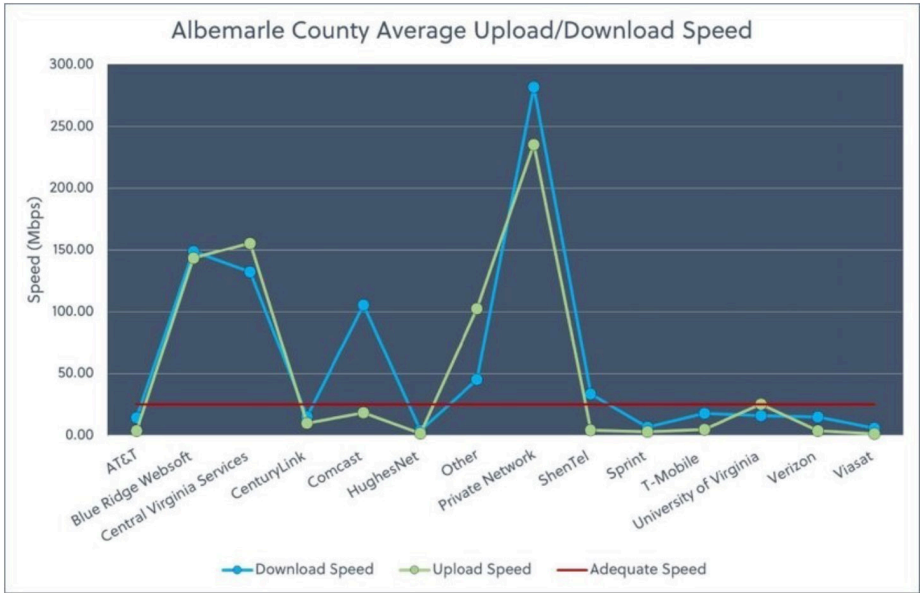

Figure 16. Average upload & download speeds in Albemarle County by provider

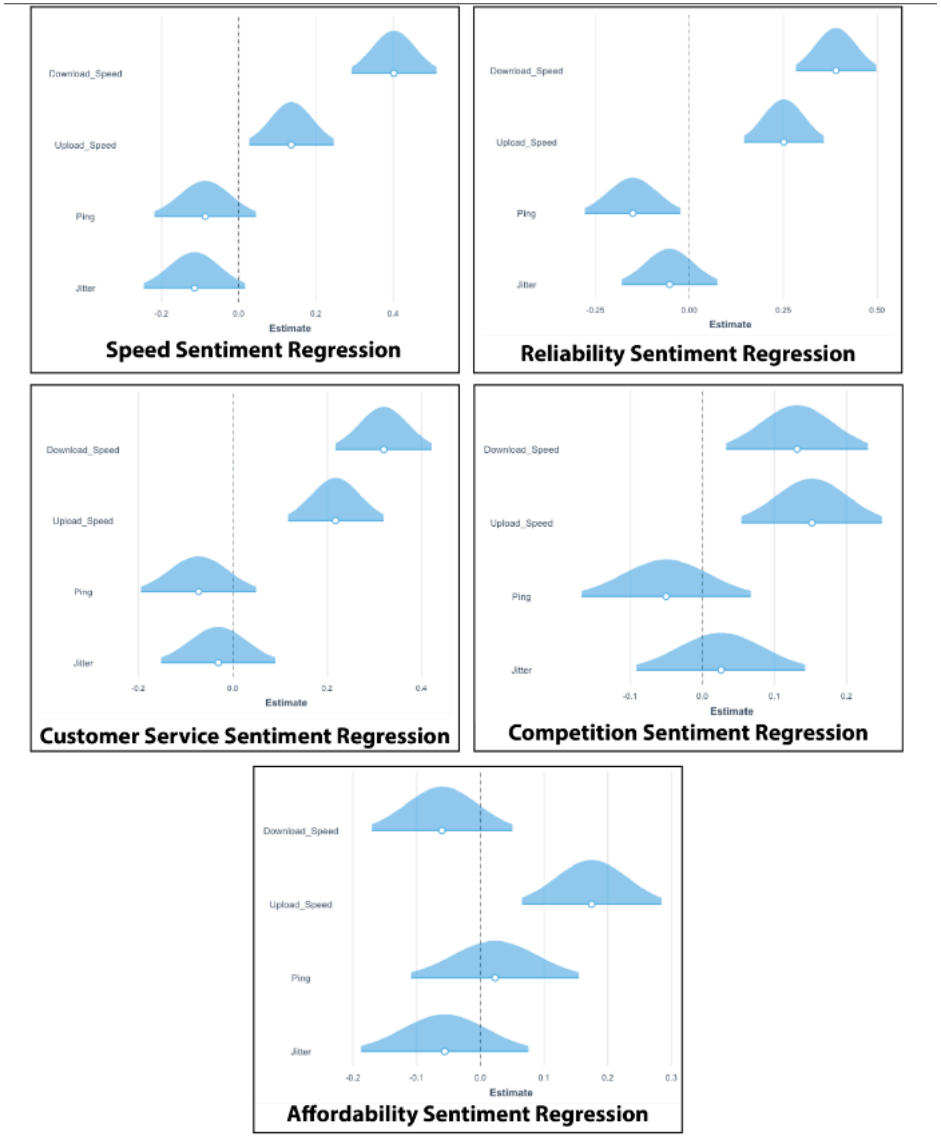

Figure 17. Scaled regression plots

Table 9. Speed score regression

Model Info:
Observations: 739
Dependent Variable: Speed_Score
Type: OLS linear regression

Model Fit:
F (4,734) = 30.544167, p = 0.000000
R^2 = 0.142700
Adj. R^2 = 0.138028

	Est.	S.E.	t val.	p
(Intercept)	2.341110	0.061617	37.994724	0.000000
Download_Speed	0.003423	0.000477	7.170214	0.000000
Upload_Speed	0.002204	0.000900	2.449190	0.014551
Ping	-0.000813	0.000626	-1.300048	0.193993
Jitter	-0.001296	0.000754	-1.719778	0.085894

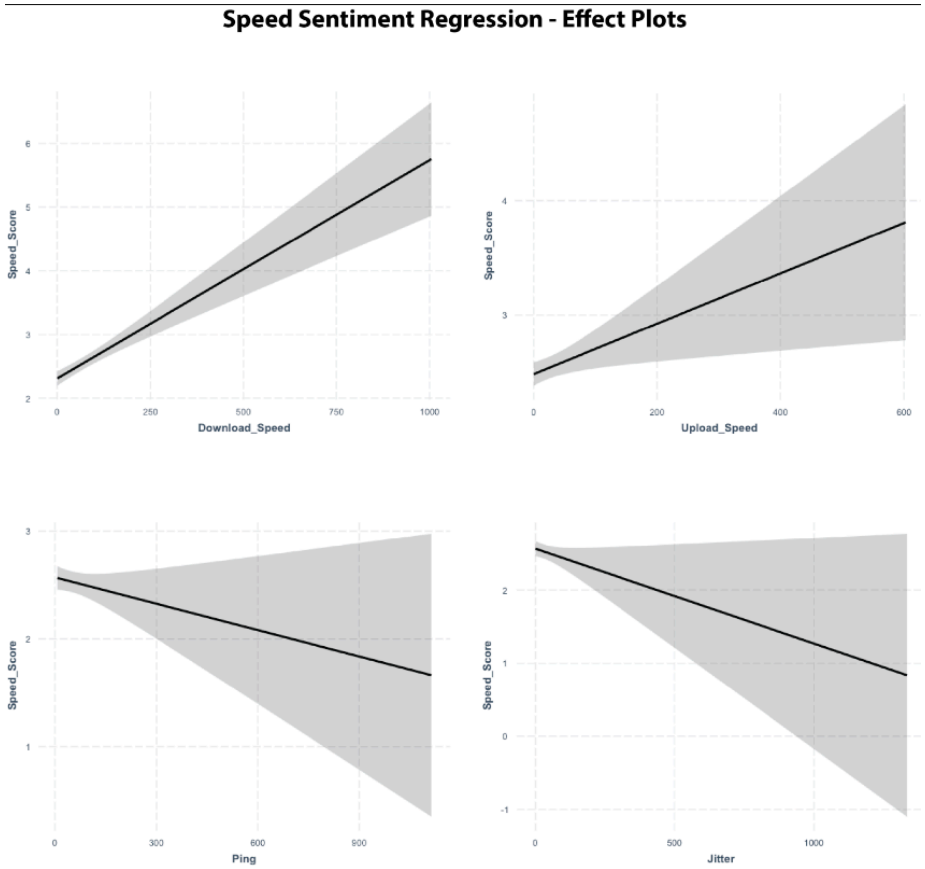

Figure 18. Speed sentiment regression effect plot

Table 10. Reliability score regression

Model Info:
Observations: 739
Dependent Variable: Reliability Score
Type: OLS linear regression

Model Fit:
F (4,734) = 41.788252, p = 0.000000
R^2 = 0.185488
Adj. R^2 = 0.181049

	Est.	S.E.	t val.	p
(Intercept)	2.235021	0.059832	37.355104	0.000000
Download_Speed	0.003322	0.000464	7.167195	0.000000
Upload_Speed	0.004057	0.000874	4.643010	0.000004
Ping	-0.001412	0.000608	-2.323455	0.020427
Jitter	-0.000601	0.000732	-0.820913	0.411963

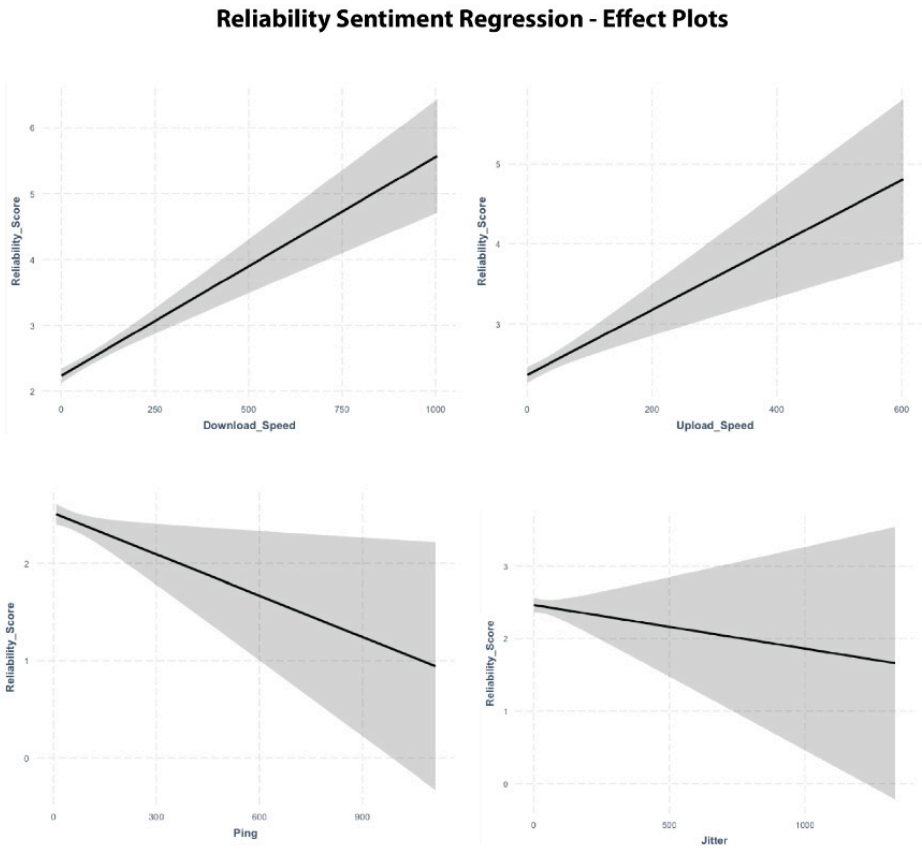

Figure 19. Reliability sentiment regression effect plots

Table 11. Customer service score regression

Model Info:
Observations: 739
Dependent Variable: Customer_Service_Score
Type: OLS linear regression

Model Fit:
F (4,734) = 29.143157, p = 0.000000
R^2 = 0.137052
Adj. R^2 = 0.132349

	Est.	S.E.	t val.	p
(Intercept)	2.093121	0.057060	36.682740	0.000000
Download_Speed	0.002724	0.000442	6.162449	0.000000
Upload_Speed	0.003522	0.000833	4.225696	0.000027
Ping	-0.000678	0.000579	-1.170140	0.242324
Jitter	-0.000350	0.000698	-0.500870	0.616613

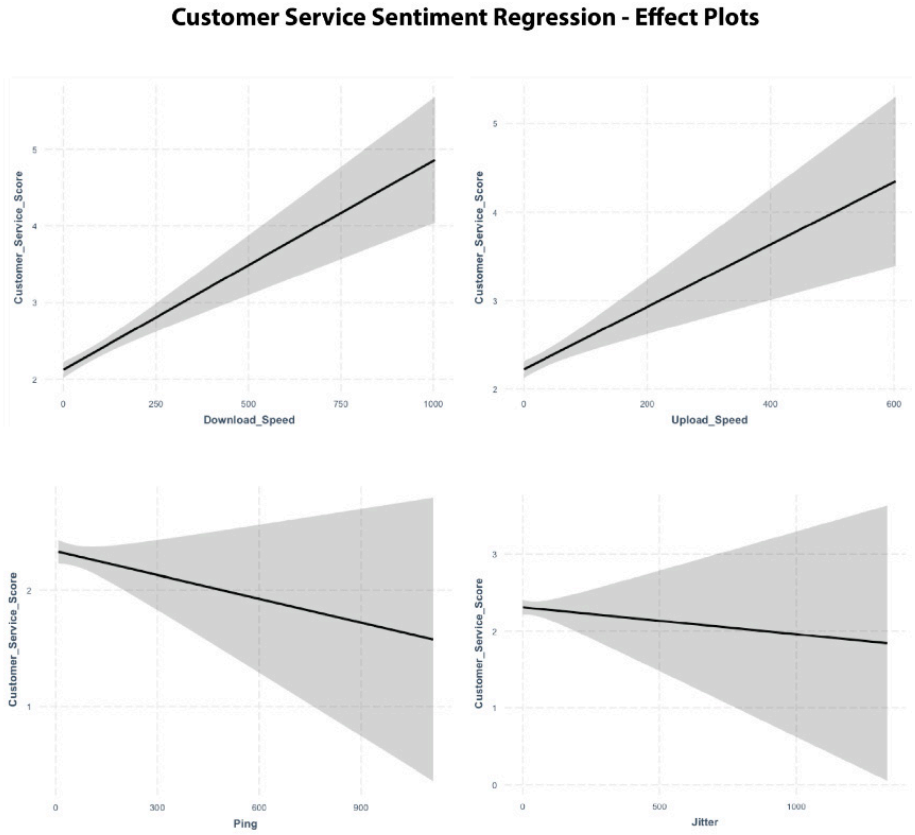

Figure 20. Customer service sentiment regression effect plots

Table 12. Competition score regression

Model Info:
Observations: 739
Dependent Variable: Competition Score
Type: OLS linear regression

Model Fit:
F (4,734) = 8.249396, p = 0.000002
R^2 = 0.043022
Adj. R^2 = 0.037807

	Est.	S.E.	t val.	p
(Intercept)	1.619143	0.054973	29.453396	0.000000
Download_Speed	0.001118	0.000426	2.625451	0.008834
Upload_Speed	0.002451	0.000803	3.052497	0.002351
Ping	-0.000469	0.000558	-0.839346	0.401549
Jitter	0.000291	0.000673	0.432999	0.665143

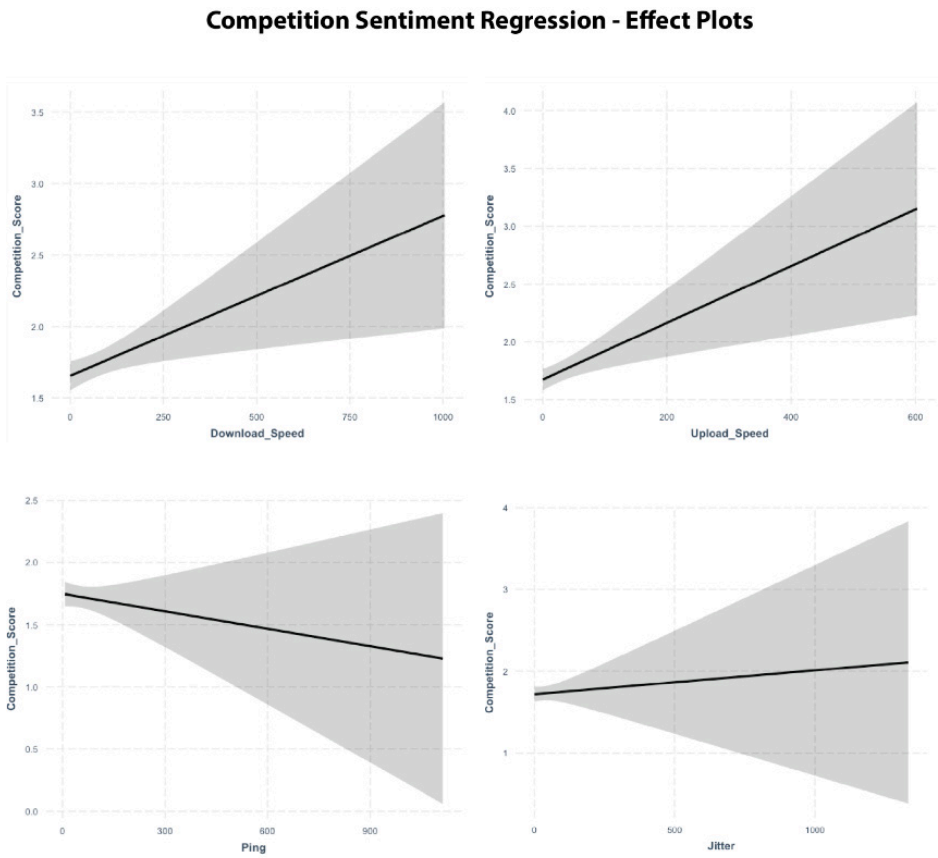

Figure 21. Competition sentiment regression effect plots

Table 13. Affordability score regression table

Model Info:
Observations: 739
Dependent Variable: Affordability_Score
Type: OLS linear regression

Model Fit:
F (4,734) = 2.738666, p = 0.027846
R^2 = 0.014705
Adj. R^2 = 0.009336

	Est.	S.E.	t val.	p
(Intercept)	3.528041	0.061877	57.016540	0.000000
Download_Speed	-0.000513	0.000479	-1.070194	0.284884
Upload_Speed	0.002823	0.000904	3.123674	0.001856
Ping	0.000216	0.000628	0.343358	0.731427
Jitter	-0.000633	0.000757	-0.836353	0.403228

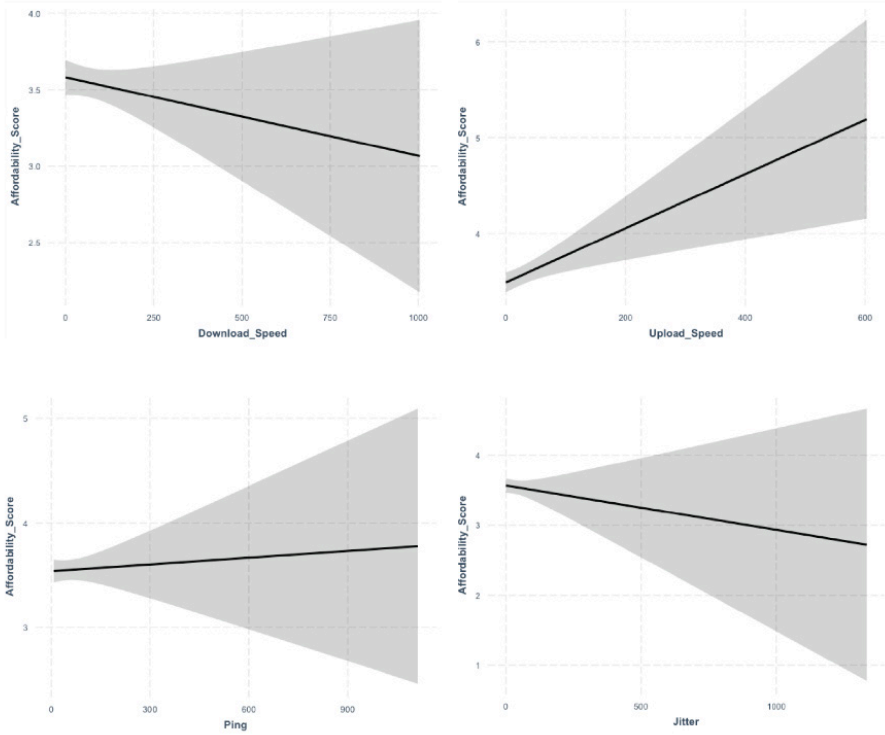

Figure 22. Affordability score regression effect plots

REGIONAL COMPARISONS

Key Takeaways

- Digital equity was assessed through comparative analysis of states and university towns using broadband access as a critical variable.
- Correlation and regression analysis of data from the census and American Community Survey 5-year estimates for 2017-2021 were used.
- Florida, Georgia, South Carolina, and Virginia had moderate to high correlations for most variables related to digital equity.
- Charlottesville/Albemarle areas and Northern Virginia had generally higher percentages of households with broadband subscriptions than other regions in the state.
- The study provides valuable insights into digital equity in various regions and can help support the development of policies to address the digital divide in Charlottesville/Albemarle.

Introduction

Comparative analysis is one way to assess a given community's digital equity. For this comparative analysis, we explored digital equity by looking at factors related to digital equity at a state level and examining similar cities/communities. With Virginia being in the South Atlantic Region, a comparison of the remaining states in that region seemed fitting, which included the following states: West Virginia, North Carolina, South Carolina, Georgia, and Florida. Secondly, we examined the state of digital equity in college towns similar to the University of Virginia in Charlottesville/Albemarle County. The comparative universities and towns included: Montana State University in Bozeman, MT, Michigan State in East Lansing, MI, and the University of Vermont in Burlington, Vermont.

The critical variable we analyzed as an essential factor in digital equity was Broadband access. Across each of our comparative states and communities, we examined the following demographic variables: high school diploma and bachelor's degree attainment, poverty level, employment rate, and median household income. By comparing these variables across the different regions, we gained a more comprehensive understanding of digital equity and how Charlottesville/Albemarle compares to similar university towns and states throughout the region.

Methods and Results

The quantitative methods used for this data analysis include correlation and regression and exploratory data analysis for census data. The data used for analyzing broadband access was the Access Broadband data at the county and state levels. This dataset was joined to census data, specifically the American Community Survey 5-year estimates for 2017-2021, to allow mapping broadband access for each state at the county level. The correlation analysis was used to look at associations between variables in the Access Broadband data; the variables used were the percent of households with a broadband subscription, the employment rate, the poverty rate, median home value, the percent of the population with a high school diploma, and the percent of the population with a bachelor's degree or higher. We then ran a regression to see how each variable predicts the percentage of households with a broadband subscription. This analysis allowed us to see which factors most impacted predicting households with broadband subscriptions. See an overview of the findings below.

South Atlantic States Results Highlights

**Maps showcasing comparative states "Percent of Households with a Broadband Access" for each of the states can be found in Figures 23-28.*

***South Atlantic States data references can be found in Tables 14-16.*

- Virginia – relatively high correlation for all variables except educational attainment (percent high school diploma and percent bachelor's degree or higher) (Table 14)
    - Charlottesville/Albemarle areas, as well as NOVA, seem to have generally higher percentages of households with a broadband subscription compared to other regions in VA

- South and Southwest VA have the lowest percentage of subscriptions – this seems to coincide with urban/rural divides throughout the state (Figure 23)

- From the regression analysis, employment rate and poverty rate were statistically significant in predicting the percentage of households with a broadband subscription (Table 15)

- Florida - relatively high correlation for all variables, yet overall, it appeared to be lower than VA's correlation (Figure 24)
  - From the regression analysis, employment rate and poverty rate were statistically significant in predicting the percentage of households with a broadband subscription (same findings as VA)
  - Poverty was the strongest relationship with a 0.0 p-value (which makes sense - the cost of broadband would logically be a much larger tradeoff for households below the poverty line) (Table 14)
- Georgia – Moderate to High Correlation for all variables (Figure 25)
  - Correlations
    - Two variables showed a strong, positive correlation. These variables were Median Home Value and Percent of people with a Bachelor's degree.
    - The rest of the variables showed a moderate correlation. Two of these variables displayed a moderate, negative correlation (Poverty Rate and Percent of people with a High School Diploma. The other variable, employment rate, displayed a moderate, positive correlation (Table 14)
  - There were three variables statistically significant in Georgia's analysis.
    - Two variables had p-values of 0.00, showing extremely high significance. These variables include employment rate and percent of people with a high school diploma.
    - The other statistically significant variable was the poverty rate, with a p-value of 0.01.
- South Carolina—Moderate to High Correlation for most variables (Figure 26)
  - Two variables demonstrated statistical significance, including high school diploma and poverty rate. (Table 14)
  - There is a strong high correlation between broadband access and employment, median home value, and percentage of bachelor's attainment.
- North Carolina (Figure 27)
  - Strong correlations across the categories as relating to home broadband subscription
    - Median home value has the strongest positive correlation, followed very closely by PCT of the population with a bachelor's degree or higher.
    - The poverty rate has the strongest negative correlation, followed closely by PCT of population with a high school diploma.
  - In running a multivariate regression, poverty and employment rates are the only variables with statistical significance.
- West Virginia (Figure 28)
  - Two variables contained statistically significant p-values—percent employed and percent that held a high school diploma.

- Both percent employed, and median home value has a strong, positive correlation

University Town Results Highlight

**Maps showcasing comparative cities "Percent of Households with a Broadband Access" for each of the regions can be found in Figures 29-32.*

*** Comparative City Data References can be found in Tables 17-19.*

- Charlottesville, VA (tract data is from Charlottesville City and larger Albemarle County) (Figure 29)
  - Correlation
    - Weaker results overall; the strongest correlation with a home broadband subscription is poverty rate (negative) and PCT of the population with a bachelor's degree or higher (positive).
    - Other variables were weak.
  - Regression
    - The poverty rate is the only statistically significant variable.
- Bozeman, MT (tract data is from the larger Gallatin County) (Figure 30)
  - Correlation
    - The strongest correlation between home broadband subscriptions was with poverty rate (negative) followed by PCT employed (positive). Other correlations between the DV (PCT Broadband subscription) were weak
  - Regression
    - The poverty rate is the only statistically significant variable.
- East Lansing, MI (tract data is from Clinton and Ingham Counties because East Lansing is located in both counties) (Figure 31)
  - Correlation
    - The strongest correlation was with the poverty rate (Negative) followed by the Employment rate (positive) and Median home value (positive)
  - Regression
    - The employment rate and poverty rate were the only two significant variables
- Burlington, VT (tract data is from the larger Chittenden County) (Figure 32)
  - Correlation
    - The strongest correlation was with the poverty rate (Negative) followed by the Employment rate (positive)
  - Regression
    - The employment rate and poverty rate were the only two significant variables

Limitations

When considering the limitations of the regional comparisons, first and foremost, it is essential to recognize some of the limitations that arose from our need to merge NTIA data with census data. Census data in and of itself has various limitations associated with it. Some of these limitations include the many possibilities of errors in the census investigation method due to non-response, measurement, and a lack of preciseness of the definition of statistical units, to name a few. One of the other limitations we encountered occurred when comparing Charlottesville, VA, to other college towns. Our basis for these comparisons was the populations of the cities in which the colleges

were located. The issue, though, is that there are not towns that fully encompass the demographics of Charlottesville, so the comparisons themselves are not entirely fluid.

Conclusion

In conclusion, this analysis identified broadband access as a critical variable. The study compared various demographic factors such as education, poverty level, employment rate, and median household income across different states and university towns. Our methods included correlation and regression analysis and exploratory data analysis of data from the census and American Community Survey 5-year estimates for 2017-2021. Despite limitations in the census investigation method of non-response, measurement, and a lack of precision in defining statistical units, some meaningful information was gathered from the analysis.

Across these states, through this comparison, the study created a more comprehensive understanding of the state of digital equity in these areas and how they compare to each other. The findings demonstrated that Florida, Georgia, South Carolina, and Virginia had moderate to high correlations for most variables related to digital equity. In addition, the Charlottesville/Albemarle areas and NOVA in Virginia had generally higher percentages of households with broadband subscriptions than other regions in the state. Although this is positive, there is still work to be done to advance digital equity in Albermarle/Charlottesville.

Overall, the study provides valuable insights into digital equity in various regions and can help support the development of policies to address the digital divide in Charlottesville/Albemarle.

Next steps

The next steps following the regional comparisons would be for the county to look at how these other places throughout the South Atlantic region with similar statistics have been able to generate positive change. This positive change should be centered around developing ways to increase broadband accessibility throughout the county. Looking at the various demographic factors we explored, it is crucial for the county to explore promoting this change in a way that is equitable and does further exacerbate the discrepancies that currently exist.

Table 14. Demographic data comparisions by state

States	Percent Home Broadband Subscription	Educational Attainment: Percent with a Bachelor's Degree or higher	Median Home Value in Dollars	Poverty Rate (%)
Virginia	**87.6**	**40.3**	**295,500**	**9.9**
Florida	87.2	31.5	248,700	13.1
Georgia	86.2	33	206,700	13.9
North Carolina	85.4	33	197,500	13.7
South Carolina	83.2	29.8	181,800	14.5
West Virginia	80.8	21.8	128,800	16.9
***Bolded values indicate highest value for each variable**				

Table 15. Regression coefficients by state

State	Employment Rate	Poverty Rate	Median Home Value	Educational Attainment: Percent with a Bacehlor's Degree or higher	Educational Attainment: Percent with a High School Diploma or higher
Virginia	**0.37**	**-0.26**	0	0.19	**-0.31**
Florida	**0.16**	**-0.77**	0	0.02	-0.34
Georgia	**0.31**	**-0.35**	0	0	**-0.48**
North Carolina	**0.29**	**-0.5**	0	0.01	-0.32
South Carolina	0.15	**-0.96**	0	-0.12	**-0.51**
West Virginia	0.28	-0.03	0	-0.16	-0.32
***bolded values indicate statistical significance**					

Table 16. R^2 values by state

State	Overall R Squared Value
Virginia	0.66
Florida	0.76
Georgia	0.64
North Carolina	0.65
South Carolina	0.80
West Virginia	0.36

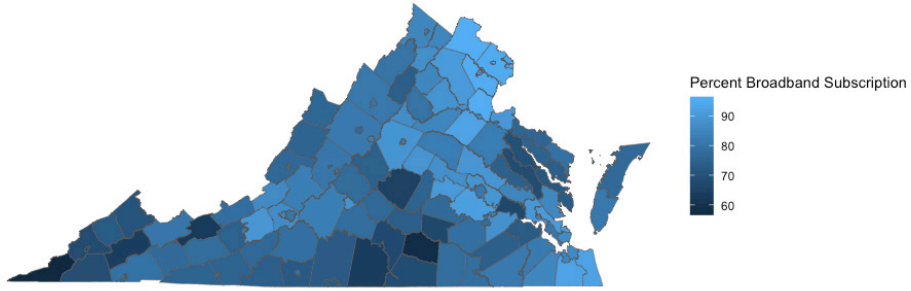

Figure 23. Broadband subscription by county in Virginia

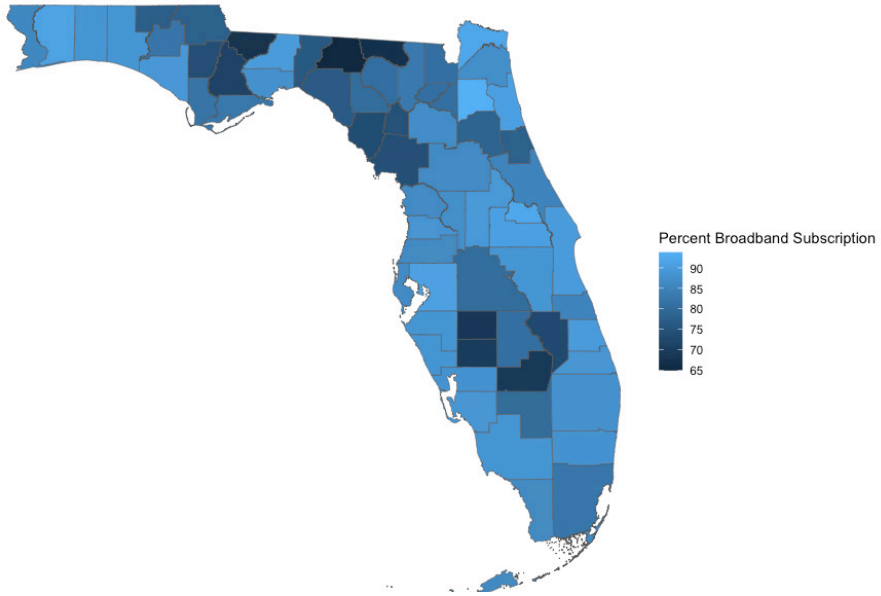

Figure 24. Broadband subscription by county in Florida

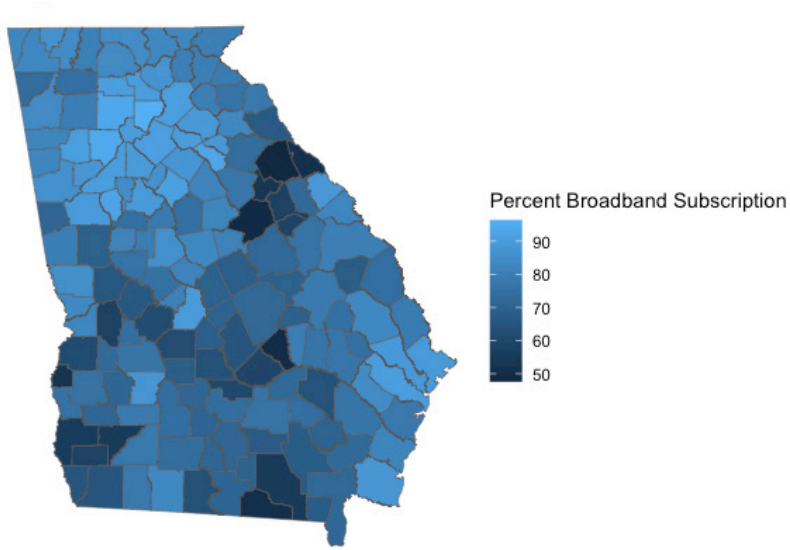

Figure 25. Broadband subscription by county in Georgia

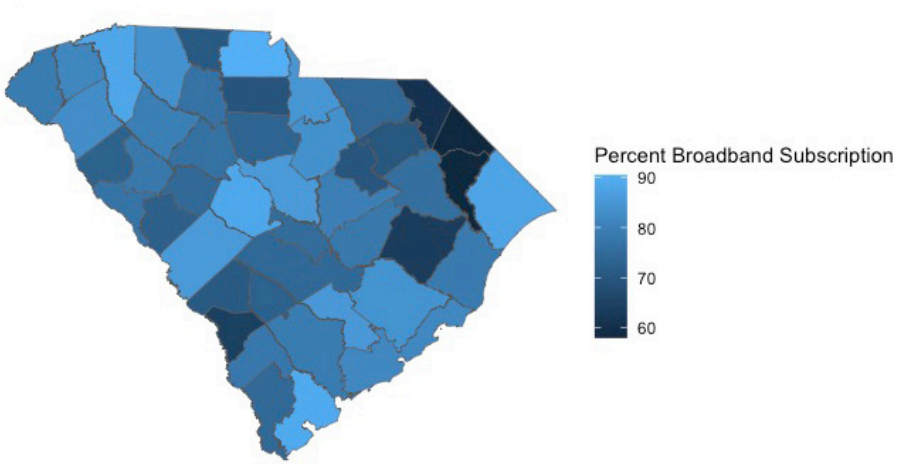

Figure 26. Broadband subscription by county in South Carolina

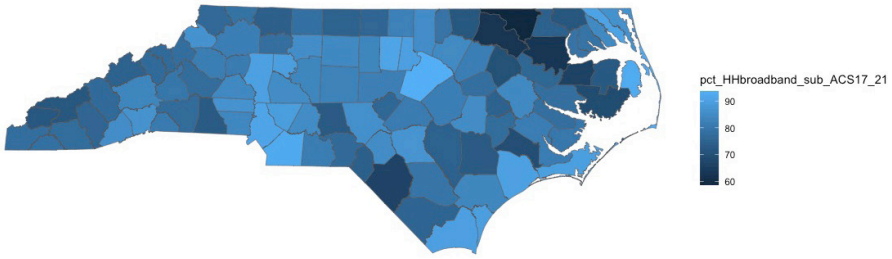

Figure 27. Broadband subscription by county in North Carolina

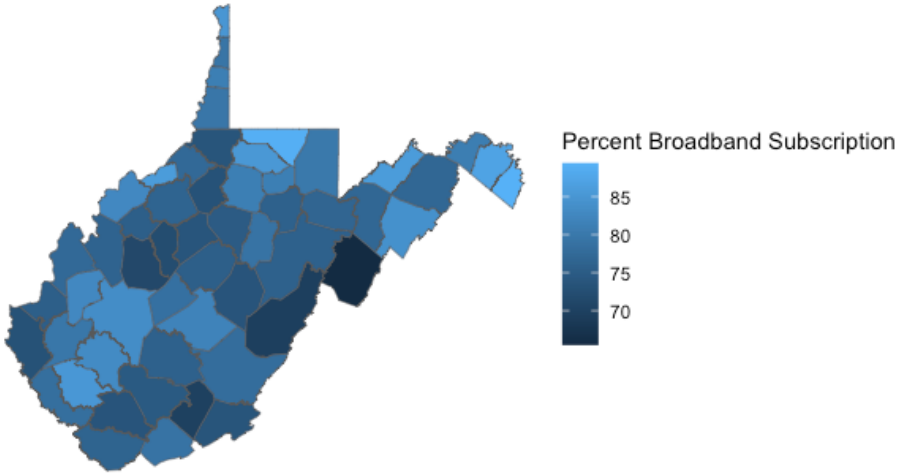

Figure 28. Broadband subscription by county in West Virginia

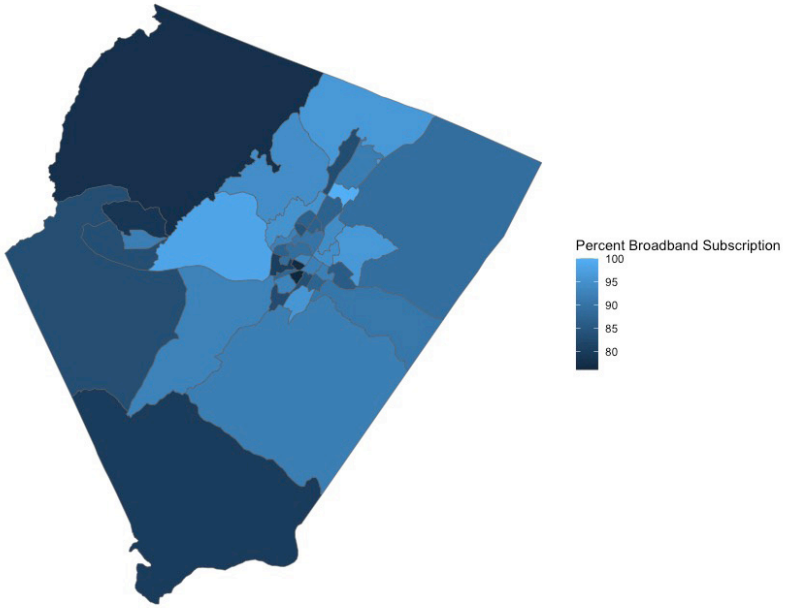

Figure 29. Broadband subscription by county in Albemarle County/ Charlottesville city, VA

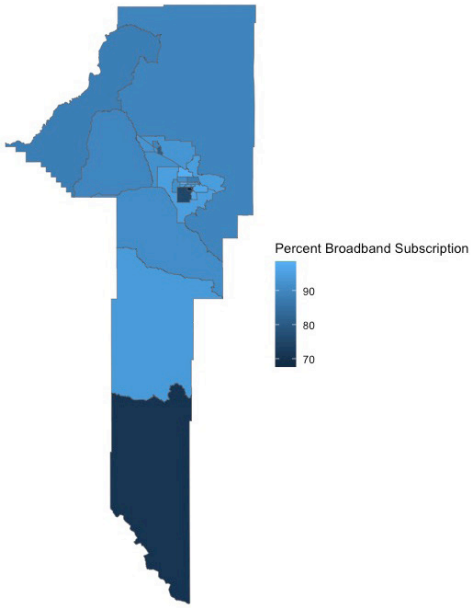

Figure 30. Broadband subscription by county in Gallatin County, MT

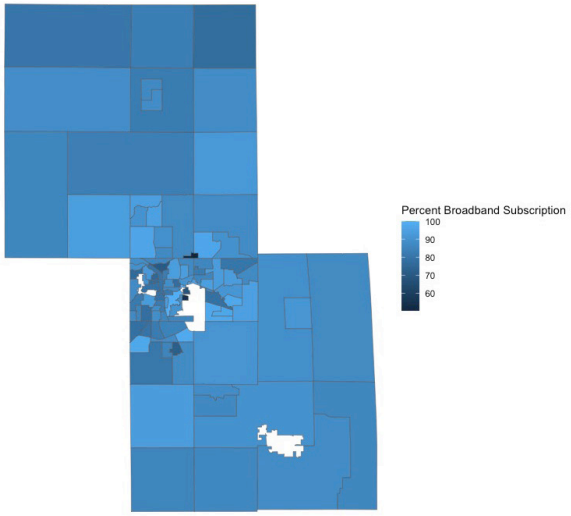

Figure 31. Broadband subscription by county in Clinton/Ingham Counties, MI

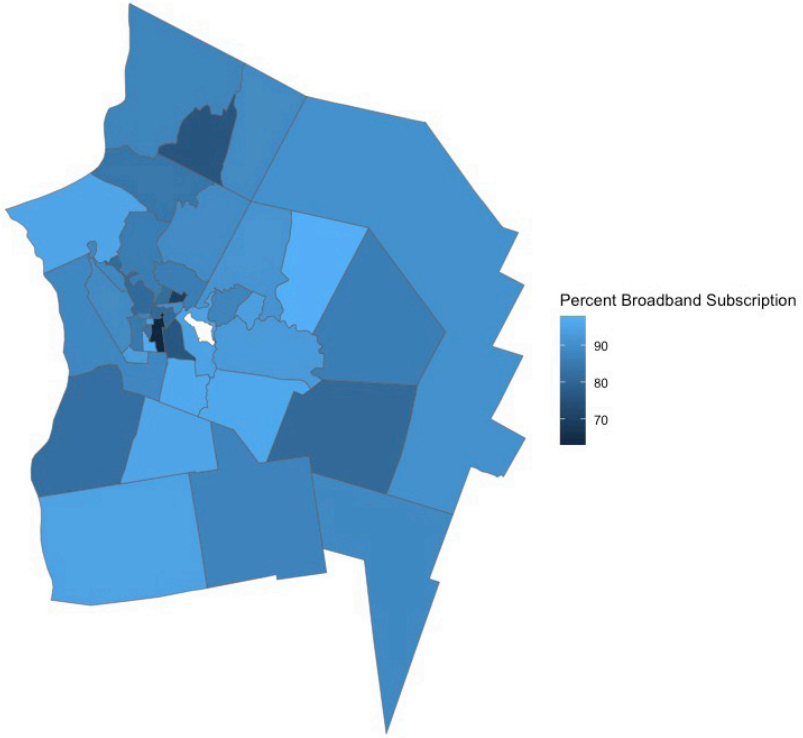

Figure 32. Broadband subscription by county in Chittenden County, VT

Table 17. Demographic data comparisions by city/region

Cities	Percent Home Broadband Subscription	Educational Attainment: Percent with a Bachelor's Degree or higher	Median Home Value in Dollars	Poverty Rate (%)
Bozeman, MT	**91**	52.2	**435,400**	**10.4**
Burlington, VT	88.2	54.1	331,300	11.3
Charlottesville, VA*	87.8	**58.5**	371,100	15
East Lansing, MI*	87.4	36.6	175,300	13
***Bolded values indicate highest for each variable**				
*Charlottesville includes averaged Charlottesville City and Albemarle County data				
*East Lansing data includes averaged Englewood and Clinton Counties				

Table 18. Regression coefficient omparisions by city/region

City	Employment Rate	Poverty Rate	Median Home Value	Educational Attainment: Percent with a Bacehlor's Degree or higher	Educational Attainment: Percent with a High School Diploma or higher
Bozeman, MT	0.31	**-0.53**	0	0.25	-0.03
Burlington, VT	**0.33**	**-0.2**	0	0.19	0.08
Charlottesville, VA*	0.1	**-0.19**	0	0.24	0
East Lansing, MI*	**0.34**	**-0.21**	0	0.06	0.06
***bolded values indicate statistical significance**					
*Charlottesville includes averaged Charlottesville City and Albemarle County data					
*East Lansing data includes averaged Englewood and Clinton Counties					

Table 19. R^2 comparisions by city/region

City	Overall R Squared Value
Bozeman, MT	0.51
Burlington, VT	0.3
Charlottesville, VA*	0.32
East Lansing, MI*	0.38

COMPOSITE METRICS

Key Takeaways

- Strong positive correlations were found between percent household broadband and mean download speeds, mean reliability, and mean affordability.
- Negative correlations were found between percent household broadband and percent high school diploma, sum of reported slow speeds (1-10 Mbps), and mean ping speeds.
- The R^2 value for the regression model is 0.77, indicating that the independent variables explain 77% of the variance in percent of households with broadband
- The overall p-value is 0.27, indicating that the regression is not statistically significant.
- The composite metric analysis used a weighted criteria matrix with four predictor variables and showed that block groups closer to the City of Charlottesville have better results than block groups on the outskirts of the county.

Introduction

This project is intended to analyze a mixture of datasets involving school, broadband, and community data to create a composite metric for the purpose of the Regional Digital Equity Plan. In analyzing this material, a composite metric, or a metric whose values are defined by a mathematical formula involving other metrics, details a broadband access score. Additionally, a weighted criteria matrix evaluates potential options against a list of weighted factors. The composite scores dictate how broadband is affected by its block group.

Data Sources

Data for this project was aggregated from three sources. Since the 2021 ACCESS BROADBAND Act, the National Telecommunications and Information Administration (NTIA) has maintained a dashboard to track information about the economic impacts of broadband investment. Specifically, it tracks 24 American Community Survey (ACS) variables relating to employment levels, education, and commerce. The present project incorporates the ten variables which contain information at the tract level as its first data source.

Data also comes from a survey of internet speeds conducted by the Albemarle County School District. Survey respondents have the option to indicate that their internet is fast, slow, or "none" (i.e. no internet). Number of respondents for each of these three categories, plus the total number of respondents, are aggregated to the block group level.

The final data source included is the RDEP's survey of Albemarle County residents. This survey includes a categorical measure of speed and satisfaction; four numerical measures of upload speed, download speed, ping, and jitter; and a sentiment survey which measures subjective concepts such as reliability and affordability.

The aim of this analysis was first to aggregate and join data from these disparate sources, and then to create a composite measure of broadband access using columns from all three sources.

Aggregating the Data

Because ACS data was given at the tract level, data from the other two datasets (given at the block group level) had to be aggregated before a join could be performed. For numeric values in the RDEP speed test survey, the value given in the table is the mean of all observations collected from block groups within a given census tract. The number of observations per tract (n) is given in a separate column. Categorical data is pivoted wider such that each category (i.e. 11-25 Mbps) becomes a separate column showing the total count of respondents in that category for each census tract. Because school survey data points are counts (per block group) of individual responses, these are summed, rather than averaged, to create the tract-level value.

For ease of processing, null observations have been eliminated from the dataset. This includes fifteen Albemarle census tracts with no data from the school survey or RDEP speed test.

Analyzing the Data

Though the initial combined data table had 62 columns, only seven variables were selected to feed into the composite metric or broadband accessibility. Variables were chosen for the composite score based on a strong correlation with Percent Household Broadband (`pct_HHbroadband_sub_ACS17 _21`). Regression analysis made use of all seven variables; however, three were dropped while computing the composite score due to their inverse correlations.

Variables selected:

- Pct_HS_diploma_ACS17_21 (ACS)*: Percent High School Diploma
- Slow_sum (School Survey)*: Sum of Reported Slow Speeds
- 0-10Mbps_currentInt (RDEP Speed Test – categorical)*: 1-10 Mbps
- AutoSpeedD_mean (RDEP Speed Test – speed metrics): Mean Download Speeds
- AutoPing_mean (RDEP Speed Test – speed metrics)*: Mean Ping Speeds
- Reliabilit_mean (RDEP Speed Test – sentiment survey): Mean Reliability
- Affordabil_mean (RDEP Speed Test – sentiment survey): Mean Affordability

**Used in composite metric*

The composite metric was computed using a weighted criteria matrix. After standardizing all values, each was weighted equally. Values were multiplied by 10 and summed. With four predictor variables, the resulting composite metric indicates overall quality of broadband access by assigning a score from 0 to 40. Higher scores indicate better quality broadband access overall.

Analyses: Correlation

This report explored correlation, regression, and a composite metric analysis between percent household broadband and the remaining variables.

A correlation analysis measures the relationship between two variables describing the strength and direction of that relationship. A strong correlation is closer to +/- 1. A stronger positive correlation is represented as a larger dark blue circle, while a weaker positive correlation is represented as a smaller lighter blue circle. A stronger negative correlation is represented as a larger dark red circle, while a weaker negative correlation is represented as a smaller lighter red circle. A positive correlation between two variables moves in the same direction. A negative correlation is the opposite, as the value of one variable increases, the other decreases. A correlation coefficient of 0 indicates no correlation between the variables. The variables measured are analyzed against percent household broadband. Most of the relationships are very strong, with the negative correlations being weaker than the positive correlations. The variables that have positive correlations are mean download speeds, mean reliability, and mean affordability. The variables that have negative correlations are percent high school diploma, sum of reported slow speeds, 1-10 Mbps, and mean ping speeds (Figure 33).

Analyses: Regression

The R^2 value is the coefficient of determination or the variability of the dependent variable that is explained by an independent variable. The R^2 value for the model is 0.77 which indicates that the independent variables explain 77% of the variance in percent of households with broadband (Table 20). The model explains a big proportion of the variance in the data. The overall p value is 0.27 which means that the regression is not statistically significant. This means that there is no strong evidence to suggest that the independent variables in the model have a significant effect on the percent of households with broadband.

The regression coefficient represents the slope of the linear relationship between the independent variables and the dependent

variable. A negative regression coefficient means that as the percentage of households with broadband increases, the response variable tends to decrease. The variables with negative regression coefficients are percent of population with high school diploma, sum of reported slow speeds, 1-10 Mbps, and mean reliability. A positive regression coefficient means that as the percentage of households with broadband increases, the response variable tends to also increase. The variables with positive regression coefficients are mean auto speed, mean auto ping, and mean affordability. The results of this analysis are listed in the regression table and in the combined effect plot. Effects plots are used to visualize how changes in the independent variable affect the predicted value of the dependent variable. Overall, the R^2 value suggests that the model is a good fit for the data, but the non-significant p-value indicates that the relationship between the independent variables and the dependent variable are not strong enough to be considered statistically significant. The results of the plot sum depict how certain the results of the true regression coefficient are within a 95% confidence interval (Figure 34). The circle on the line visualizes the standardized slopes, while the line shows the margin of error of the slope. For this analysis, many of the standardized slopes are close to zero, meaning there is limited impact from our independent variables on broadband access.

Composite Metric

The composite metric was computed using a weighted criteria matrix. The four predictor variables are percent high school diploma, sum of reported slow speeds, 1-10 Mbps, and mean reliability. A score closer to 40 is better and a score closer to zero is worse. The spatial results of this analysis are that block groups closer to the City of Charlottesville have better results than block groups on the outskirts of the county (Tables 21 and 22).

Maps

The maps pictured below represent the overall data results for each variable (Figures 35-43). They are aggregated to the block group level. Looking at the raw data helps us to spatially orient to the dataset and to also see where data is missing. Maps can help to make complex datasets more accessible and easier to interpret.

Limitations

One significant limitation is the small size of the project's dataset. Although there are 29 census tracts in Albemarle County, only 12 were fully represented on all three surveys (and thus included in the final table). Though many values are aggregated from more robust datasets, the small number of rows severely limits the possibilities for gleaning valid tract-level insights from the composite table.

Generalizability of the composite metric beyond Albemarle County is similarly limited. Many data sources come from programs or initiatives unique to Albemarle County. Although analogous data sources may be found in other localities, methods presented here can only serve as the most basic blueprint for understanding broadband access. Nevertheless, insights from the data into the state of broadband access in rural Albemarle County may well generalize to similar locations across the country.

Lastly, the composite metric needs further adjustments. However, rather than a limitation, we see this as an opportunity for continued research. First, weights can be tuned beyond their current evenly weighted distribution to better capture the relative importance of various factors to overall broadband health. Second, rather than their current standardization, variables can be normalized such that the three negatively correlated variables that were dropped could be added back to the weighted criteria matrix for a more robust composite score.

Conclusion

The most significant outcomes from the composite metrics section of the analysis are the correlation analysis and the framework for the development of a composite score. After an analysis of several variables from across three data sources, seven were picked out as being especially strongly correlated with broadband access (as measured by the 2021 5-year American Community Survey). Of these, four were used to create a prototype of a composite broadband score. As predictors of overall broadband access, careful monitoring of these four variables – percent with a high school diploma, sum of residents reporting a slow internet connection, number of residents reporting speeds from 0-10 Mbps, and mean ping of all county residents – may be useful as a shorthand for understanding the quality of tract-level broadband access.

Next Steps

Based on the analysis conducted, the following recommendations are proposed for the Regional Digital Equity Plan:

- Increase access to reliable and high-speed internet in areas with a lower percentage of households with broadband access.
- Increase both qualitative and quantitative data surveys in the county.
- Prioritize investments in infrastructure and services to improve the reliability and speed of internet connections.
- Focus on affordability of internet services to increase adoption rates, especially in areas with lower median incomes.
- Continue to collect and analyze data to track progress and adjust strategies.

By implementing these recommendations, the Regional Digital Equity Plan can work towards closing the digital divide and ensuring that all residents have equitable access to the internet.

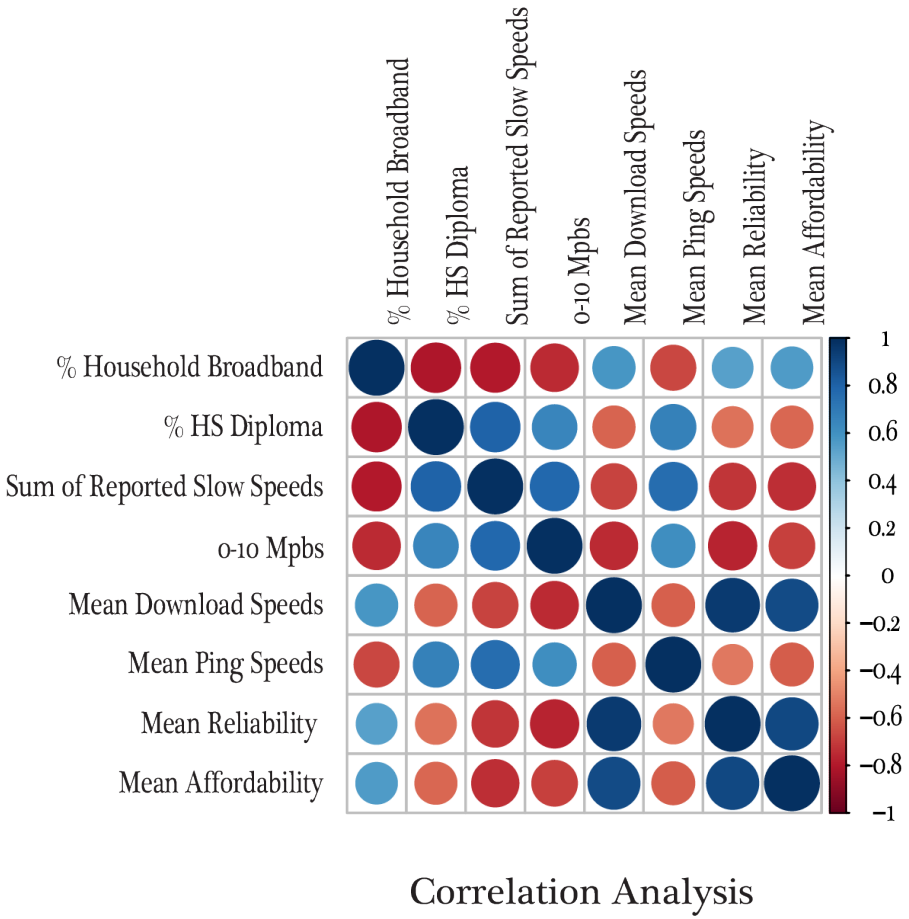

Figure 33. Correlation plot

Table 20. Regression

Dependent Variable: Percent Household Broadband
p=0.27
r^2=0.77

	Est.	S.E.	p
Intercept	112.97	42.81	0.06
Percent High School Diploma	-0.45	0.56	0.47
Slow Sum	-0.08	0.15	0.60
0-10 Mbps	-0.03	0.04	0.43
Mean Speed	0.13	0.26	0.64
Mean Ping	0.01	0.06	0.93
Mean Reliability	-6.21	9.86	0.56
Mean Affordability	0.36	13.37	0.98

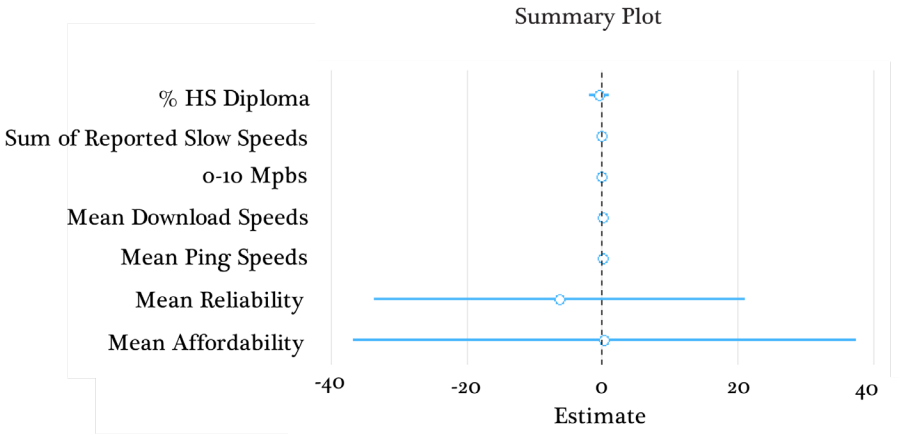

Figure 34. Regression scaled plot

Table 21. Composite scores by GEOID

Data Composite

GEOID	% Household Broadband	Composite Score
51003010100	77.8	17.6656204081633
51003010201	96.4	26.2960273092369
51003010202	94.6	35.5194312236287
51003010401	89.3	21.6347141552511
51003010402	90.7	32.35416
51003010602	94.4	37.2344208333333
51003011000	98.4	31.1559922865014
51003011201	83.4	21.7265717171717
51003011202	92.7	29.6049309200603
51003011301	92	26.8872043010753
51003011303	96.1	36.3131130434783
51003011400	80.1	15.1075086956522

Table 22. Composite scores by GEOID

Data Composite

GEOID	% Household Broadband	% HS Diploma	Sum of Reported Slow Speeds	0-10 Mpbs	Mean Download Speeds	Mean Ping Speeds	Mean Reliability	Mean Affordability	Composite Score
51003010100	77.8	22.5	85	247	15.49017007	78.02615646	1.8	3.489473684	17.6656204081633
51003010201	96.4	13.2	49	130	32.4551004	77.51598394	2.372881356	3.457627119	26.2960273092369
51003010202	94.6	9.9	22	13	54.52202532	28.43088608	2.678571429	3.821428571	35.5194312236287
51003010401	89.3	19	45	186	15.27383562	131.8821461	1.779411765	3.470588235	21.6347141552511
51003010402	90.7	12.8	18	40	35.62485714	81.646	2.882352941	3.794117647	32.35416
51003010602	94.4	10.3	11	0	63.15125	20.4728125	3.2	3.8	37.2344208333333
51003011000	98.4	6.5	34	88	50.65140496	46.0168595	2.487804878	3.885365854	31.1559922865014
51003011201	83.4	17.3	61	195	17.37757576	91.50237374	1.68	3.34	21.7265717171717
51003011202	92.7	14.1	34	94	17.49447964	59.79339367	2.090909091	3.757575758	29.6049309200603
51003011301	92	21.1	65	73	16.83991935	66.65322581	1.696969697	3.337878788	26.8872043010753
51003011303	96.1	14.3	12	7	97.59565217	24.42217391	4.111111111	4.444444444	36.3131130434783
51003011400	80.1	22.6	102	151	23.96646739	202.3122826	2.144444444	3.413888889	15.1075086956522

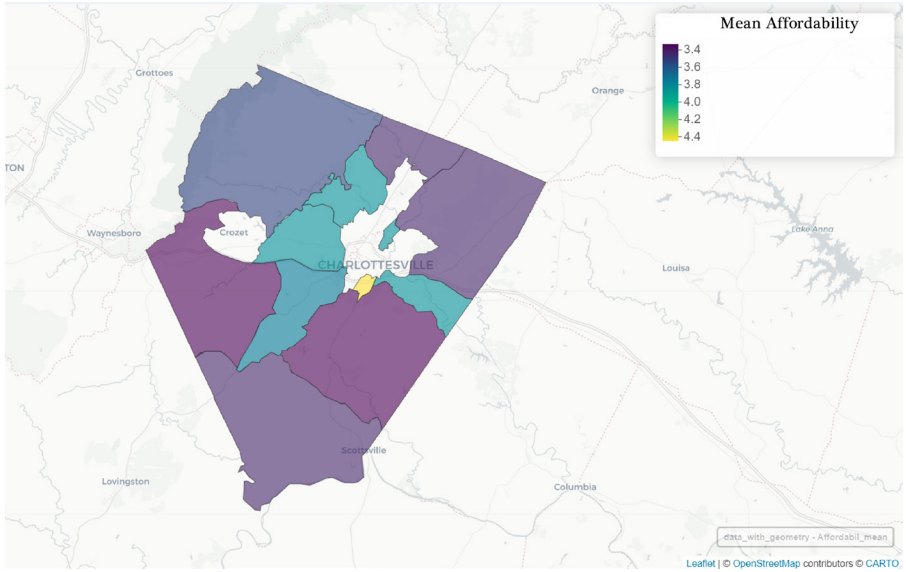

Figure 35. Affordability of broadband by block group

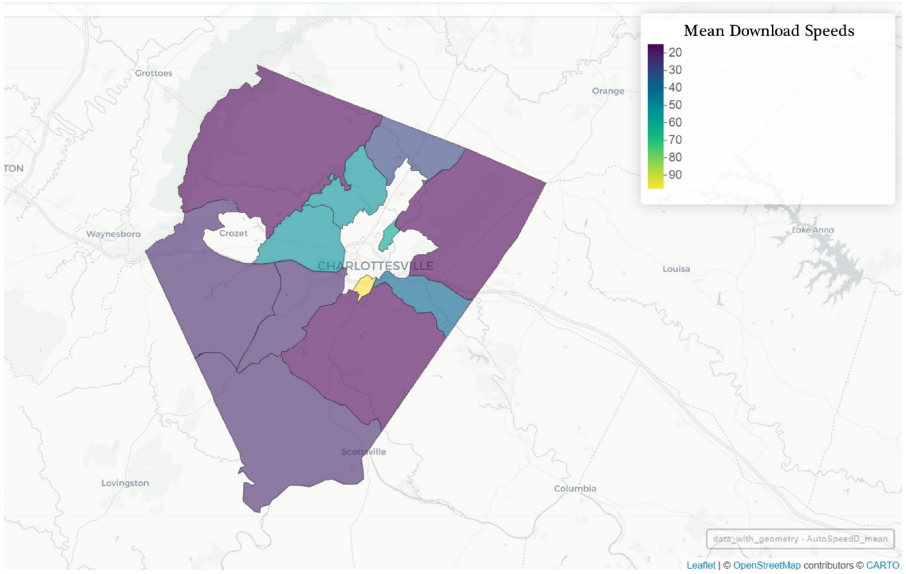

Figure 36. Download speed by block group

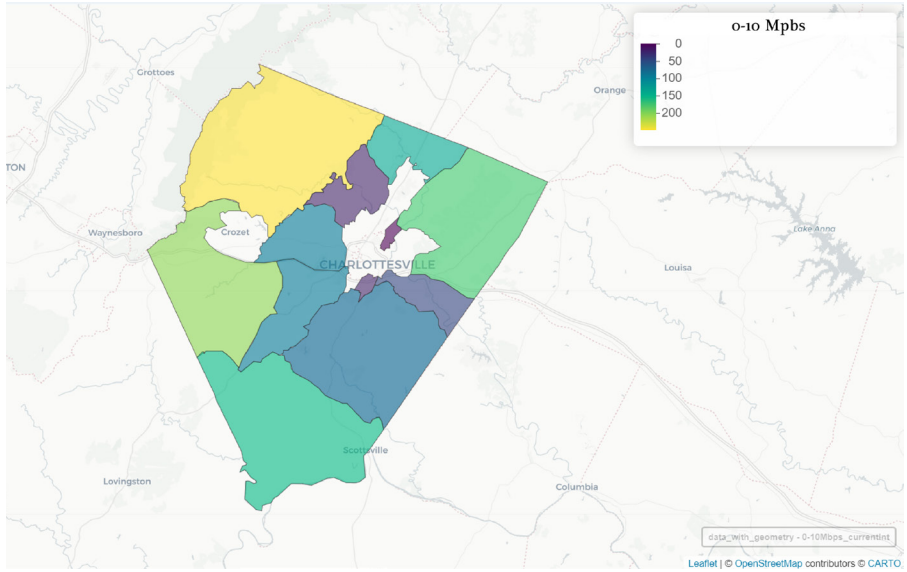

Figure 37. Broadband with less than 10 Mpbs

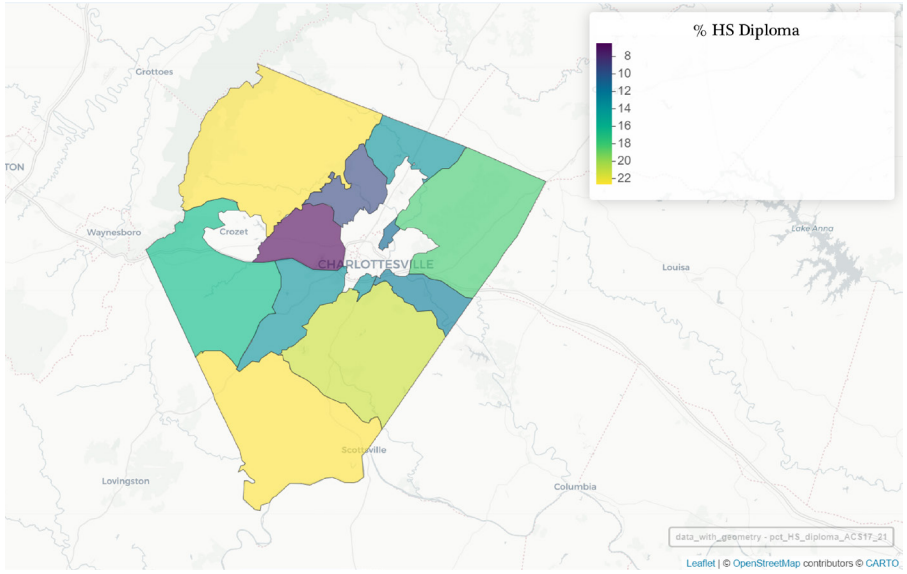

Figure 38. Highest educational attainment - high school diploma only

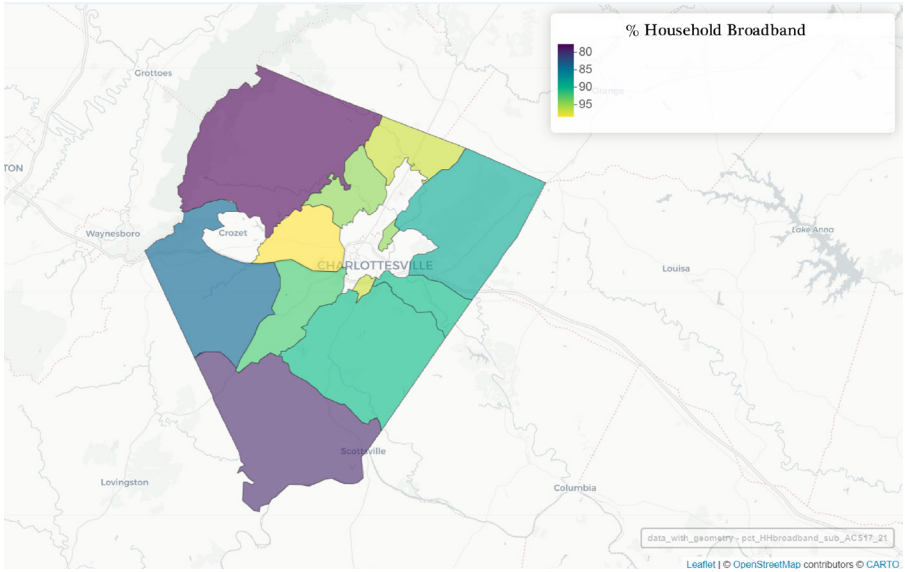

Figure 39. Amount of households with broadband by block group

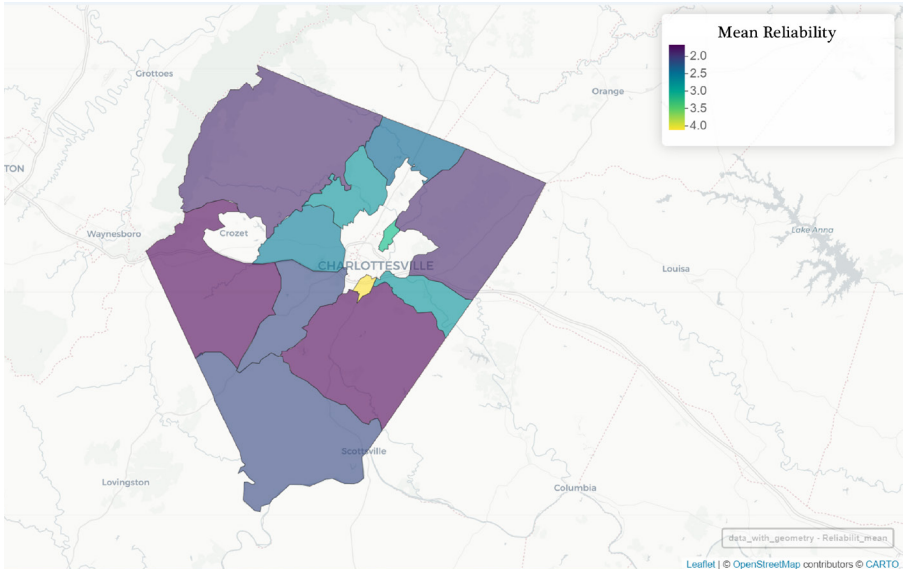

Figure 41. Broadband reliability

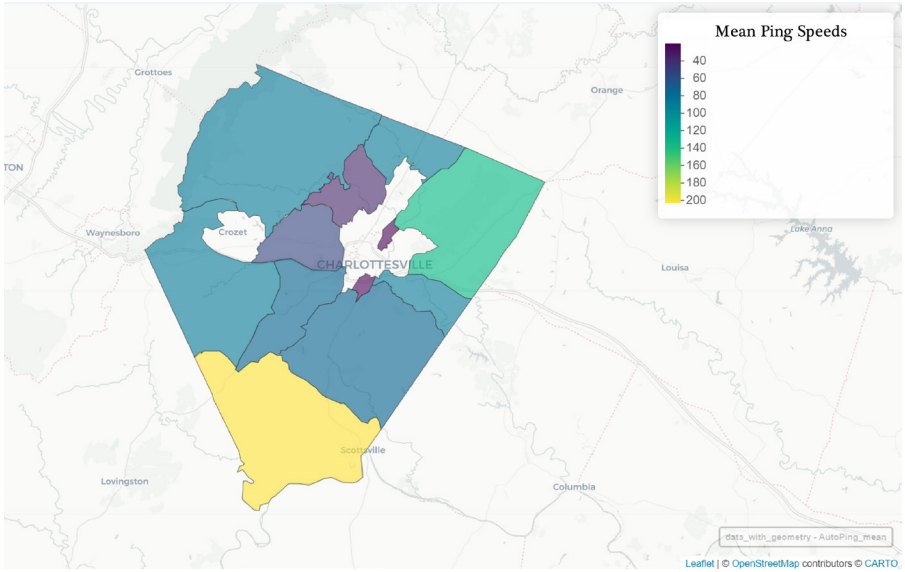

Figure 40. Mean ping speeds

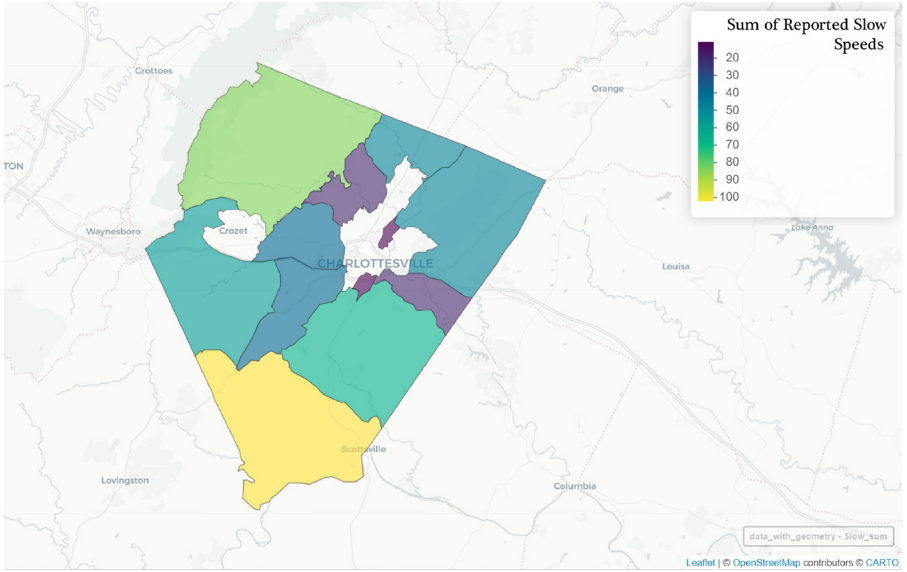

Figure 42. Sum of reported slow speeds

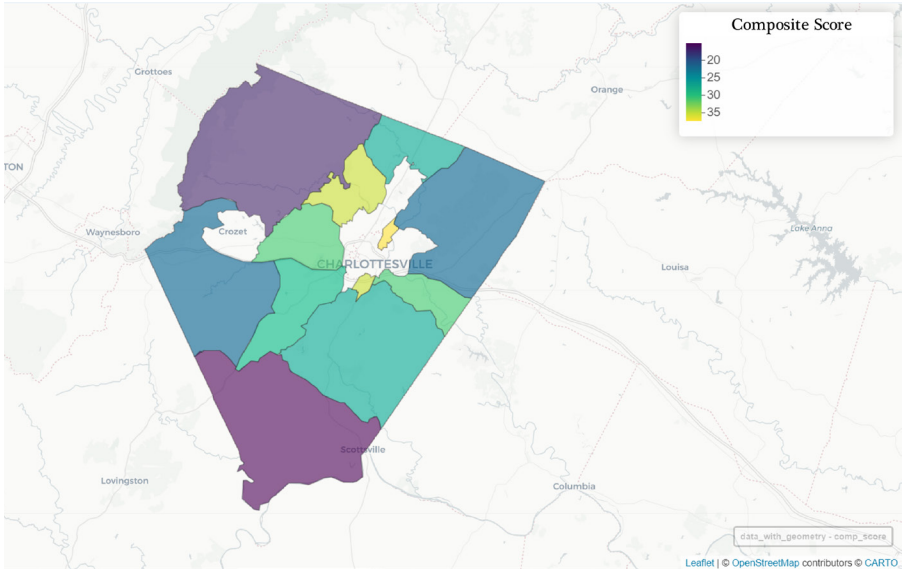

Figure 43. Composite score by block group

Appendix A: Qualitative Coding

Table 23. Coding dictionary used by all focus groups

Code	Subcode (if applicable)	Code or Subcode Definition (s)
Connectivity		Quality and accessibility whilst traveling through the city rather than being limited to their home or mobile device.
Literacy/Digital Literacy Skills		An ability to proficiently use the internet/computers for their own needs/goals/uses
		The ability to interface with digital and online resources in a safe and knowledgeable way.
Leadership/Accountability		Ability of individuals or organizations to take charge and be responsible for ensuring equitable and reliable availability of broadband services, while also being accountable for addressing any gaps or challenges in achieving universal access.
Potential Investment/Suggestions		Theoretical ideas and plans that would improve digital equity (e.g., JABA folks suggested sending "tutors" to JABA to help)
Continuous Communication		Consistent, open communication focused on building trust. (Long Beach, pg. 13).
Relevant content and services		The ability of residents to access safe and relevant online services that meet their needs.
Community Asset		Current municipal infrastructure that holds the potential for broadband usage.
Digital Privacy		The recognition and implementation of governmental protections on internet usage and education on digital literacy.
Miscellaneous Code		Anything that does not fall into one of the above codes, but you believe is relevant/important can be coded miscellaneous.
County Equity		An equitable distribution of resources, specifically broadband access, across all counties in Virginia
Crucial Access		The necessity of having reliable and affordable internet connectivity for roles as as education, job, healthcare, and social relationships.
Device Access		The availability and affordability of broadband internet-accessible devices such as smartphones, tablets, laptops, and desktop PCs.

Future-Proofing		The process of creating and placing broadband infrastructure and services that are scalable and adaptable to future requirements in order to maintain their value and relevance throughout time.
Rural Connectivity		Refers to the accessibility and quality of broadband internet connectivity in rural regions that are frequently underserved or unserved by traditional broadband providers.
Safety Exclusion		The impact of a lack of internet connectivity on public safety, such as the inability to contact 911, obtain weather warnings, or receive important health information
Barriers	Unreliability	Poor or spotty internet quality, or total lack of access/connectedness to the internet
	Learning curve	lack of access to traditional education or exposure to the internet, computer hardware, etc.
	Language	non-native English speakers, poor (digital) literacy, etc.
	Time/Competing priorities	focusing on work, children, family, etc.
	Fear	(e.g., scams, viruses, accessing illegal sites, personal reputation, legal issues, etc.)
	Affordability	making affordable access to essential tech from adequate devices to reliable high-speed internet
	Transit	Relating to the ability to travel to places where they can access the internet or use technology.
	Cell Coverage	Refers to the geographical area or range within which a cellular network provides service and connectivity.t
	Apathy	Not prioritizing increasing their technology usage.
	Competition/Choice	Lack of competition, unclear options
Digital equity		Providing reasonable resources to ensure that everyone has access to the use of information and communications technology (8).
	Universal access	Ensuring that all people who want/need access are able to obtain access in a way that does not cause social, financial, or temporal hardship
	Digital Inclusion	“Proactively ensuring everyone has access and use of computer literacy training, the Internet, technology devices and other digital inclusion resources and services” (Long Beach, pg. 9).
	Digital redlining	Low-income neighborhoods and neighborhoods of color do not have equal access to Broadband and upgrade services. (Long Beach, pg. 12).
	Uneven Geography	Disparities regarding broadband access across geographical regions that can be delineated through geographic or municipal boundaries

Table 24. Coding frequency for each focus group

Topic	JABA	PHA	SHP	Yancey	Total
Affordability	4	6	-	6	16
Apathy	9	3	2	-	14
Cell Coverage	-	-	3	12	15
Community Asset	4	5	-	1	10
Competition/Choice	-	-	-	3	3
Connectivity	6	8	4	9	27
County Equity	-	-	-	2	2
Crucial Access	-	-	-	3	3
Device Access	-	-	-	9	9
Digital Inclusion	2	9	-	2	13
Digital Redlining	-	-	-	3	3
Fear	7	-	-	1	8
Future-Proofing	-	-	-	2	2
Language	1	-	-	-	1
Leadership/Accountability	6	3	-	3	12
Learning Curve	12	15	-	-	27
Literacy/Digital Literacy Skills	19	11	5	-	35
Misc.	-	3	1	6	10
Potential Investment/ Suggestion	-	-	-	1	1
Relevant Content and Services	4	3	3	1	11
Rural Connectivity	-	-	-	4	4
Safety Exclusion	-	-	-	4	4
Time/Competing Priorities	-	1	-	3	4
Transit	-	1	4	1	6
Uneven Geography	-	-	-	5	5
Universal Access	-	5	-	1	6
Unreliability	4	1	-	14	19