Closing the Digital Divide in the United States

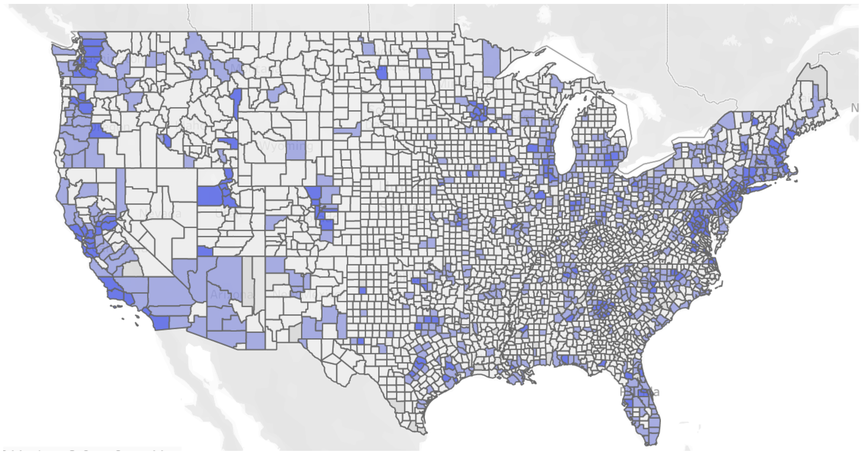
Team 13: Millie O. Symns, Geri Harding, Jayuan (JJ) Ruiz

About 1 in 4 households do not have access to the internet

22%

Households that do not have a desktop or laptop computer

Access to the internet and technology is vital to connect with education and financial resources.



The darker shades highlight the areas doing the best regarding internet and technological devices, which are primarily on the far west and northeast coasts.

01 Introduction

Internet and technology as no longer considered a luxury, but a necessity to navigate recourses to goods for a high quality of life. With the digital gap in the country, there is now Bipartisan Infrastructure Deal has pledged $65 billion to invest in broadband infrastructure deployment, including digital proficiency.

Our goal here is to recommend and highlight the focus areas across the country that should take priority in this funding for infrastructure, internet, and technology.

02 Data + Processing

Data: United States Census Bureau Data (2018), American Community Survey (2018), Iowa & ASU Data Portal (2018)

- Total Population by county
- Median Incomes
- Employment Rates
- Broadband Internet Percentage
- Technology Devices in Households

Processing: Selected 62 columns for all datasets to make the final analytic dataset of 3,144 rows by 83 columns after creating new variables for analysis.

03 Methodology

We focused our analysis on the county level to provide the most detail possible since we expected variability within each state.

We also chose to keep the data as generalizable by only including whole persons' information and not by race, gender, or age because of the limited amount of counties that were surveyed for the ACS data sets.

Additionally, we created categories (population class & income levels) to have deeper insights across the country.

04 Analysis

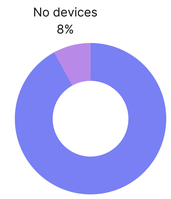
First, we created an accessibility score which is the total score of sub-scores from key metrics related to internet accessibility (the higher the score the lesser the need). Second, we used linear regressions models to predict broadband percent in households and percent households in desktops.

Total Score = broadband pct + income level + desktop pct + smartphone pct + portable pct

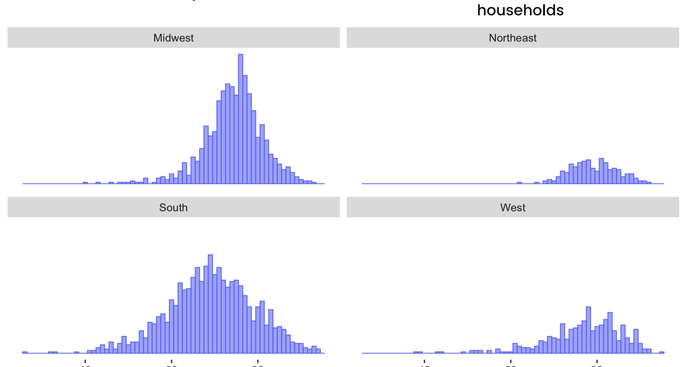
Below shows some results from our exploratory/descriptive analysis.

Region	Avg. Score
Midwest	6.3
Northeast	13.6
South	6.6
West	8.4

The Northeast and West regions are doing best regarding internet accessibity



While there is only about 8% of households without devices, that accounts for 17.8 million households



The South region has the widest range in percent of households with broadband internet and the lowest average compared other regions (69.1%).

05 Results/Findings

In our descriptive analysis, we notice trends by region and population class. Counties in the south were generally behind in their median household income, percent of households with broadband internet and percent of households with devices.

We are able to produce linear regression models to predict the percent of households with broadband (R-squared = 0.821) and percent of households with desktops (R-squared = 0.845). The median income for households was a significant factor in both models.

06 Conclusion + Future Work

Conclusions:

- More research needs to be conducted in rural areas to learn more about the internet and technology needs
- Household income plays a significant factor in having access to the internet and technology devices
- Priority should be given to the South and Midwest regions for infrastructure funding based on our assessment of need

Future Work:

- Incorporate more data on internet and infrastructure pricing and other costs
- Reach out to internet providers for their data on connectivity across the country
- Dig deeper into the variability in specific states and regions by demographic backgrounds

Limitations + Challenges

Data is limited to 2018 to have the most complete information on broadband availability; scoring is incomplete due to missingness of other variables, therefore an incomplete picture for the whole country.