Kieran Connolly

Parker Lazear

J Lei

Project Proposal

Our main motivating question is whether we can discern patterns that could indicate the impact of internet access on educational attainment. We plan to work through this through multiple streams using geospatial analysis and machine learning. We are planning to use a combination of data sources from [IPUMS](https://usa.ipums.org/usa/), [Pew Research Center](https://www.pewresearch.org/internet/dataset/2021-core-trends-survey/), and the [Urban Institute](https://educationdata.urban.org/data-explorer/explorer?educationLevel=K-12&dataLevel=School%20Districts&fipsNumbers=12&endYear=2021&timeframeType=most-recent&togglesList=80-135%2C80-134%2C80-131%2C80-130%2C101-1139%2C101-1142%2C29-235%2C30-333%2C30-334%2C30-335%2C30-336%2C30-337). These data sets were pulled from a variety of sources including census data, CommonCore of Data, EdFacts, and the Pew Survey.

To produce the least confounding results, we have chosen to specify West Virginia as our area of interest. We chose this state for its lack of broadband access paired with its high percentage of rural areas and lower geographic mobility. While other states also suffer from a lack of broadband, it is often a result of economic status, which also impacts educational attainment and success.

Our visual and geospatial analysis will map patterns of educational attainment by the school district against the level of internet access available to people in those school districts. This will attempt to demonstrate a relationship between internet availability and school performance, as we believe that internet access is key to student achievement, such as the ability to complete assignments at home, being able to use educational internet resources and tutoring, and developing technological literacy skills.

Our machine-learning approach will utilize West Virginia census data to predict attainment for individuals based on connectivity. From this prediction, we will be able to create a geospatial visualization of predicted educational attainment throughout the state. Once this is created, we will create visualizations of actual educational attainment throughout the state based on Urban Institute data and connectivity data from the Pew Research Survey. We will then compare the predicted visualization with the actual visualization to determine the accuracy of our machine-learning model. In terms of completing this, we are most concerned with both the relative missingness in data, as well as being able to compare technical shapefiles for school districts throughout the state. We are also concerned about how much we will be able to realistically complete with machine learning.

While there is a large push for research in this area currently, so much of the impacts are new and the data is relatively incomplete. In the 2023 recommendations from the FCC on digital learning, nearly all of the data provided is from qualitative interviews and leaves room for large errors and differences in experiences (Digital Empowerment and Inclusion Working Group of the Communications Equity and Diversity Council. Federal Communications Commission, 2023). Most of the quantitative research that has been done on this was done prior to the pandemic, and not only has internet connectivity and internet learning changed vastly since then, but we have also learned, importantly, that asking “Do you have internet in your home?” does not capture differences in broadband connectivity versus unreliable dial-up. This causes large errors in earlier data, like a report done by the Milken Institute that tries to find the exact models we are trying to show. Beyond that, the Milken report uses incomplete data with high degrees of missingness (Lee, 2017). Our report should be able to fill a gap in the education policy space being able to come closer to an actual answer to the question of how internet connectivity impacts education.

Sources Cited

Digital Empowerment and Inclusion Working Group of the Communications Equity and Diversity Council. Federal Communications Commission. (2023). *Connecting Opportunity Communities to Broadband During the COVID-19 Pandemic: Lessons Learned and Recommendations.* <https://www.fcc.gov/sites/default/files/cedc-digital-empowerment-inclusion-wg-broadband-access-report-06152023.pdf>

Lee, Joe. (2017). *Internet Usage Effect on Educational Attainment: Evidence of Benefits*. Milken Institute.<https://milkeninstitute.org/sites/default/files/reports-pdf/Internet-Usage-and-Educational-Attainment-FINAL.pdf>