



BROWN

Broad Trends in Broadband

Raymond Cao, Hyun Choi, Leon Lei, Diane Mutako

CSCI 1951-A (Data Science), Department of Computer Science, Brown University

Questions

How is internet accessibility and technology distributed in the American population?

Do socio-economic factors, specifically income, affect internet accessibility?

Hypotheses

Internet speed in an area is correlated to the yearly household income in the area.

Americans can be split into two groups: rural/suburban residents with few ISP (internet service provider) choices, and urban residents with many choices for service.

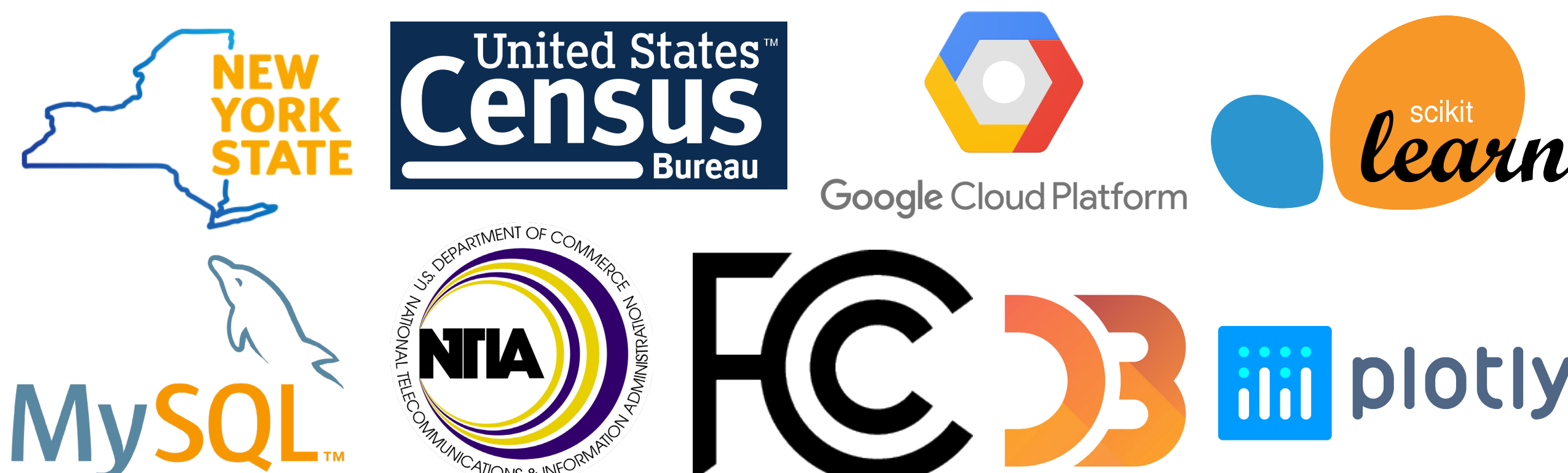
Methodology

- Considered many factors such as technologies (DSL, cable, fiber, etc.), income, and geography to get a clearer understanding of broadband internet availability.
- While the FCC considers satellite and wireless internet to be broadband, some of our analyses exclude these technologies because their high latencies make them infeasible for heavy usage.
- Performed K-means analyses, created choropleth maps, and visualized demographical factors against an area's internet speeds using data from a wide array of government sources.

Challenges

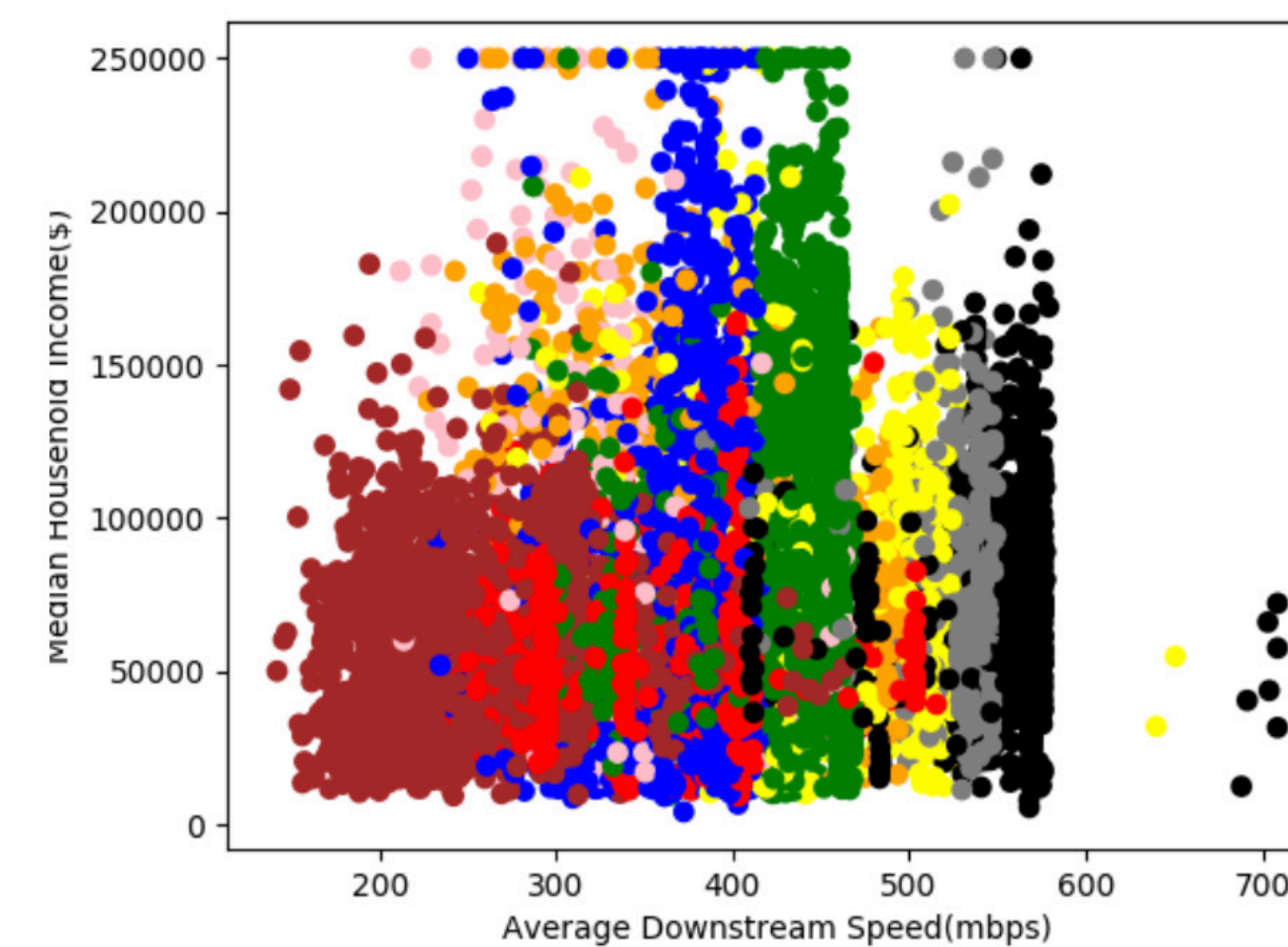
- Initially aimed to analyze nationwide trends at high resolution, as the FCC data and Census data are provided by census block, of which there are 220,333. 70.9 million rows in the FCC dataset.
- Resulted in an impractically large database, and Google Cloud SQL had trouble importing and querying this much data.
- Issues finding data that matched with each other chronologically.
- Focused on county-level data on a few states as a case study.
- MySQL lacks an easily-implemented median function, used avg

Tools, Technologies, and Datasets



NY Income & Speed K-means

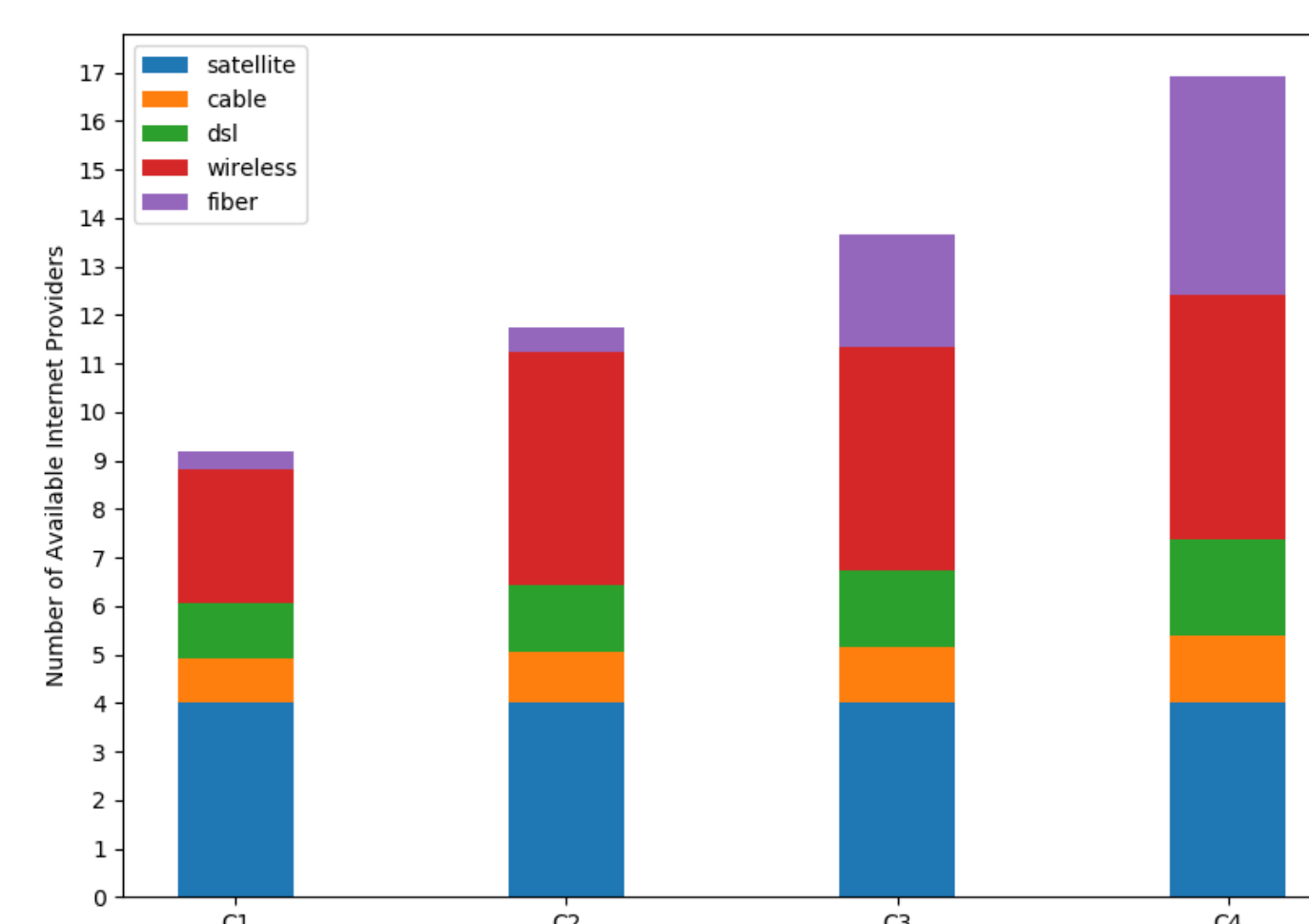
- Due to a high number of service providers and different technologies, average internet speed is not closely correlated to the total income in an area.
- As seen below, the speed in an area does not correlate with the median household income.



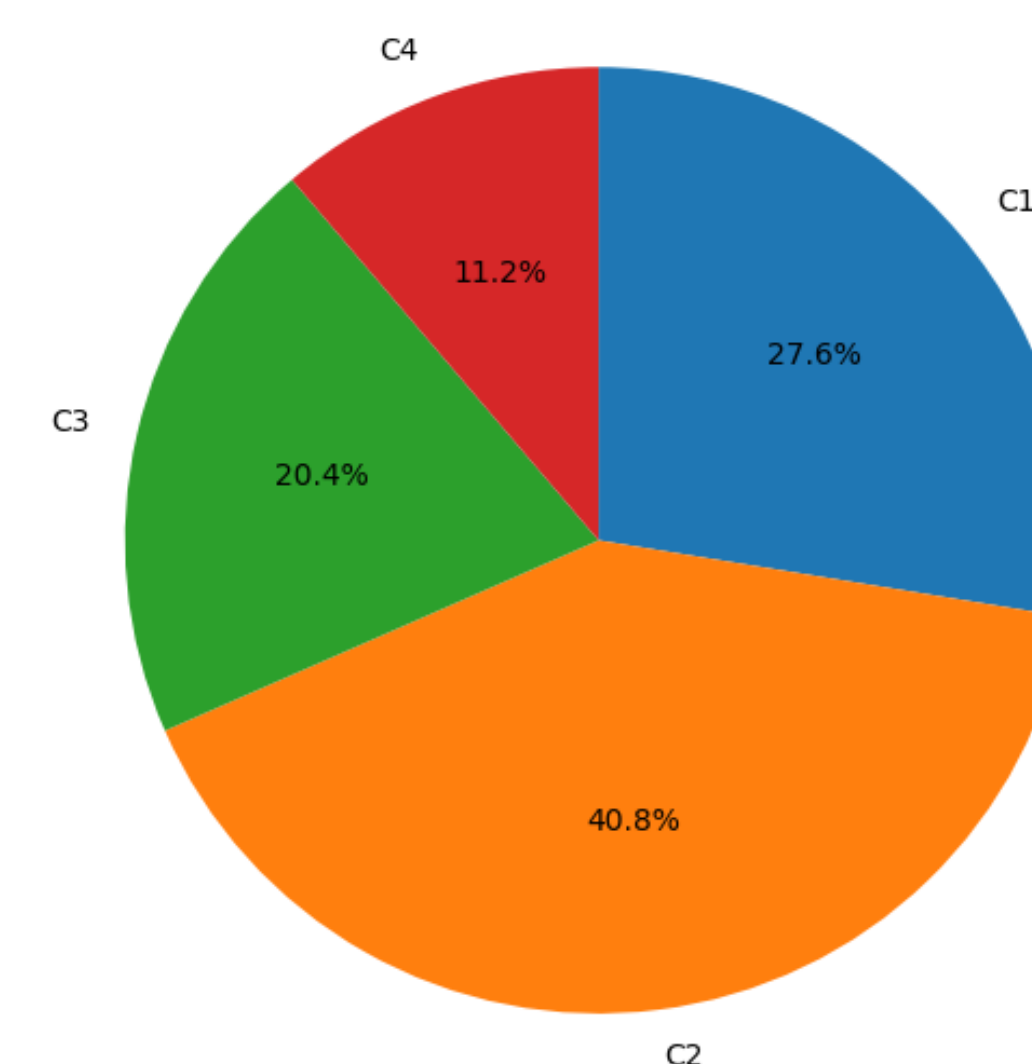
Graph 1. Downstream Speed v. Household Income

NY Internet Availability K-means

- Broke down internet availability in each municipality by technology (DSL, cable, fiber, wireless, satellite) and analyzed the distribution of each access technology across the state.
- Found four distinct clusters: one with few choices overall, one with significant wireless choice, and two with many choices.
- Most people were grouped with the centroid with many wireless internet services, but few other choices.



Graph 2: Internet Availability by Centroid



Graph 3: Population Distribution by Centroid

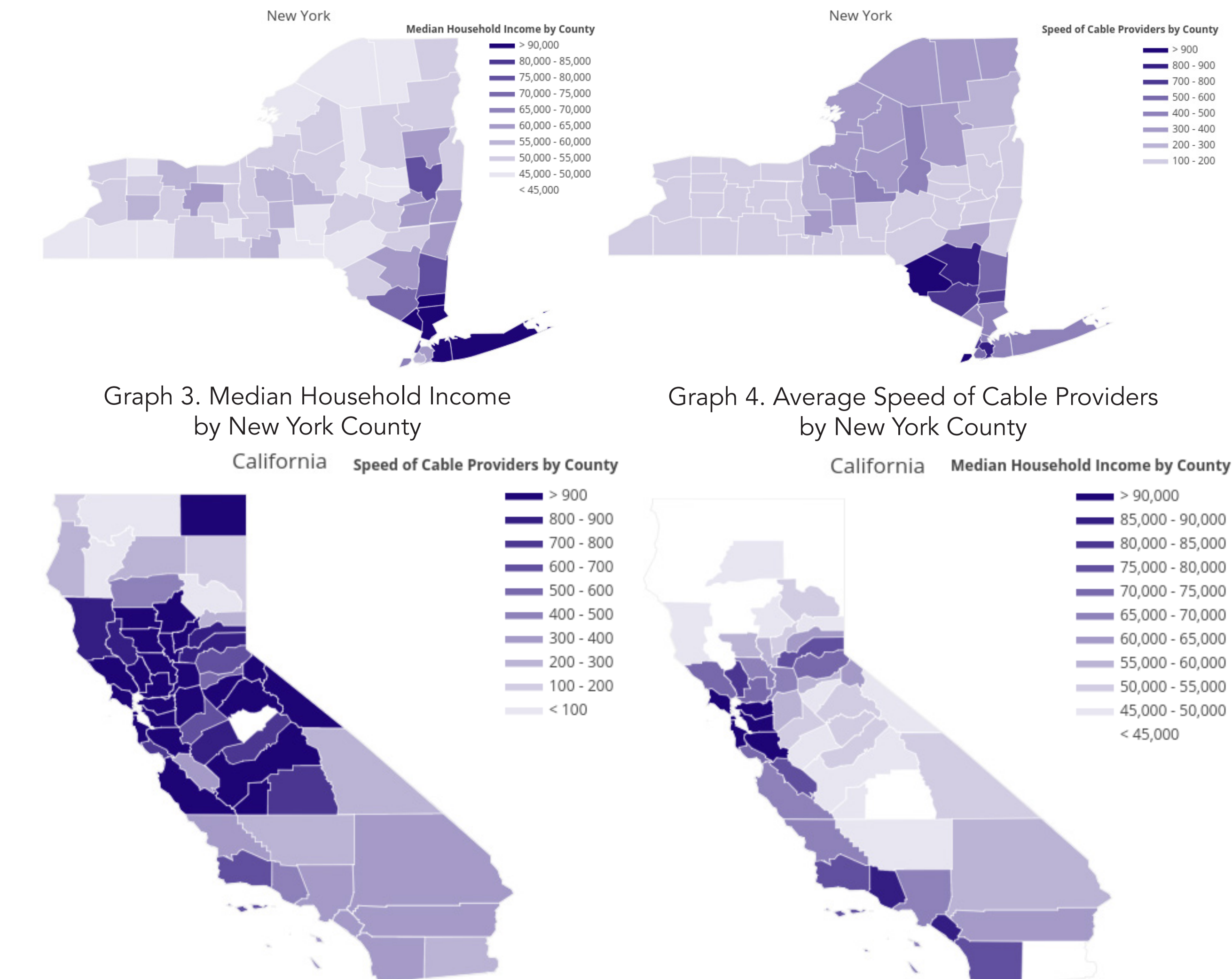
Blog & Interactive Visualizations

- Blog posts and interactive visualizations created over the course of the semester can be viewed here:



tinyurl.com/BrownBroadVis

NY/CA Income & Speed Choropleth



Graph 3. Median Household Income by New York County

Graph 4. Average Speed of Cable Providers by New York County

Graph 3. Median Household Income by California County

Graph 4. Average Speed of Cable Providers by California County

- The choropleth maps further demonstrate the small (if any) correlation between median household income and internet speed.
- Only cable provider speed (excluding DSL, fiber, wireless, satellite) was used for uniformity and because cable is the most common broadband technology.

U.S. Broadband Adoption Rates

- We conducted upper-tailed two-proportion Z-tests for internet usage for people in each income bracket (<\$25k, \$25k-49k, \$50k-74k, \$75k-99k, >\$100k) against all other income brackets to determine if the difference in internet usage was statistically significant ($\alpha = 0.001$).
- Higher income brackets had statistically significantly greater rates of internet usage than lower income brackets. This was true for all years between 2010-2017 (except 2014 and 2016, when no income data was collected).

Data Sources

- Fixed Broadband Deployment Data for New York and California. June 2017. Federal Communications Commission.
- Broadband Availability by Municipality for New York State. 2015. New York Office of IT Services.
- Median Household Income in the Past 12 Months, American Community Survey 2017, Census Bureau
- Broadband Adoption and Computer Use, National Telecommunications and Information Administration