Background

In the 2016 U.S. Presidential election, there was much debate on the effect of the media and social media on voting behavior of Americans. Commentators discussed the development of “echo-chambers,” i.e. the phenomenon whereby people express and discuss their views in forums of similarly minded people. While this discussion revolved on the influence of social media on shaping voter preferences, there can also be geographic factors at play in influencing voter preferences. This project thus seeks to look at the relationship between the isolation of American voters and their voting preferences in the 2016 Presidential Election.

Methodology

A measure of connectivity was created by aggregating information of internet access, and proximity to U.S. road and rail systems. Using county-level data internet access was indicated by the percentage of households owning both a home computer and home broadband. Proximity to U.S. interstate roads was calculated through Euclidean distance, while proximity to U.S. rail stations included all counties within a 120 mile buffer around each station. Internet access was reclassed between 1 and 3, as was road proximity. The rail station buffer, however, was counted as either within (1) or outside (0) the buffer. Together these were set on a scale of 1 to 7, 7 being most connected, and 1 most isolated. A 0 score or whitened-out county indicates a lack of data for internet access.

As for election data, polarization was measured by the ratio between Democratic vote percentage over Republican vote percentage. The voting gap, meanwhile, was the difference between the same values respectively. These were both measures of partisanship.

Results and Limitations

After conducting several regressions, the case that connectivity correlated with partisan voting patterns was not significant. This too was not proven by looking at the individual variables for connectivity, i.e. internet access, interstate road proximity or rail station proximity. This project thus suggests that more variables should be taken into account, including perhaps access to other forms of media—e.g. radio, television, print media (though data for these are difficult to obtain or find), population density, or past voting history of a county. While County-level data was the most readily available data, perhaps the better spatial unit to use would have been congressional districts. Despite their shift over the years, this would have made sense to study voting behavior.