

Analysis of the National High School Ethic Bowl Spatial Data

Corey Neff, 11/21/20

Introduction

The National High School Ethics Bowl (NHSEB) is a nationwide campaign to engage high school students in respectful and deliberative ethical discourse. The competitions, held yearly, stray from those of typical debate clubs by fostering a more constructive and collaborative atmosphere (The Parr Center for Ethics 2020). Each year, participating schools compete in one of the 39 current regional competitions across the country, and the subsequent winners will participate in the national competition held at the University of North Carolina, Chapel Hill. The NHSEB seeks to expand its program to more high schools, and it seeks data-driven decisions to aid its growth. With that in mind, this work is relevant because it attempts to inform these decisions using geospatial analysis.

The current director of the NHSEB, Alex Richardson, attempts to expand the program every year, but lacks the appropriate data analysis to properly inform these expansions. This work will summarize the current state of the spatial distribution of NHSEB competitions, as well as propose future locations for potentially new schools and regional competitions. I hypothesize that the Northeast U.S. is comparatively over-represented, and that new high schools and regionals should be seeded in the Midwest and West Coast. To do so, I will create normalized kernel density rasters of the U.S. high school population and current participating high schools and subsequently subtract them using the “minus” raster function. I will also summarize the current state of participating high schools and regionals using network and origin-destination analysis.

Methods

First, the data was procured from the director of the NHSEB, which included current regionals and participating high schools as of the 2019-2020 competition cycle. The data was then prepared by formatting in excel by cleaning up column headings and removing unnecessary variables. Next, the data was imported into ArcGIS Pro (ver. 2.6.3), and the tables for both regionals and participating high schools were geocoded to convert their addresses into point feature classes for the analysis (*Figure 1*); each address was checked by hand, which took the bulk of the time. Finally, the data of the U.S. population that will be needed was downloaded from the online ArcGIS Hub and

geocoded as well (ESRI 2020). For this data, the high school-aged population distribution was estimated using the “Ages 15-19” column included in the dataset.

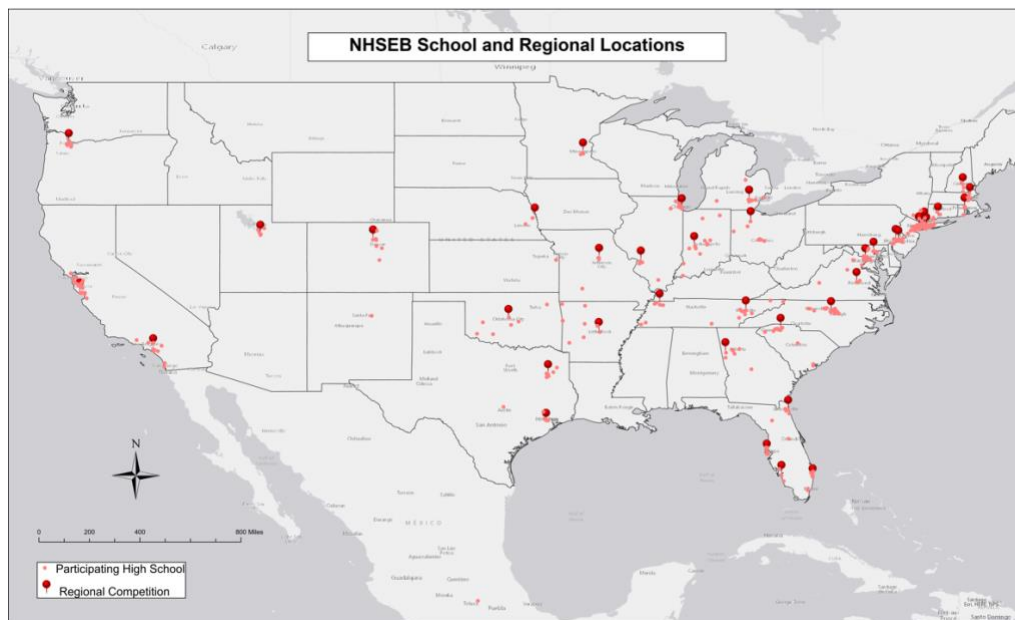


Figure 1. A representation of current participating high schools and regional competitions.

For the main analysis, a density map of the current locations of participating high schools was compared to a density map of the U.S. high school-aged population (ages 15-19). To accomplish this, kernel density maps were created for locations of current participating schools, as well as for the U.S. high school-aged population distribution. In order to allow for comparison, both of the previously mentioned density rasters were then linearly rescaled to normalize them on a scale from 1-10. Finally, using the raster calculator, the density map of participating high schools was “subtracted” from the density map of the high school population. This created a map that emphasizes the differences between the two rasters, which ultimately highlights under-represented and over-represented regions of the U.S in terms of participating high-school locations.

For the secondary analysis, an origin-destination analysis, a buffer analysis, and a network analysis were performed to summarize the current state of the distribution of participating schools. First, an origin-destination analysis was performed to calculate the average distance from each high school to its corresponding regional competition location (*Figure 2*). The origins were the participating high schools, and the destinations were their assigned regional competition. High schools whose distances lie significantly above the total mean distance would, again, be considered for relocation to a possibly closer regional or used as evidence to create a new regional in that particular area. Next, the buffer analysis consisted of creating buffers around each of the 39 regional competitions

equal to twice the combined mean distance (i.e., 77.6 miles) from a high school to its regional competition (*Figure 3*). Again, high schools that lay outside of the buffer area would be considered outliers investigated for relocation to a closer regional or creation of a new regional. Finally, the network analysis was created using each of the 39 regional competition locations as the epicenter. Three drive-time polygons were created that represented the time it takes to drive 30, 60, and 120 minutes from each regional location (*Figure 4*). These times were chosen by the director. Again, participating high schools that lie outside of the polygons would be considered for relocation to a possibly closer regional or used as evidence to create a new regional in that particular area.

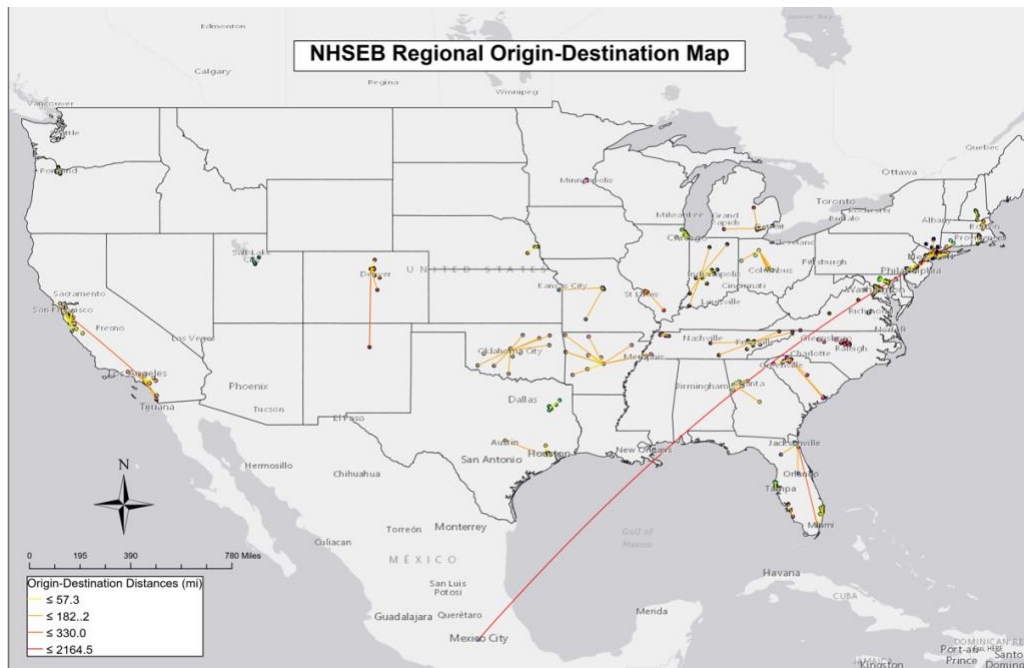


Figure 2. Origin-Destination analysis of participating high schools and their assigned regional competition.

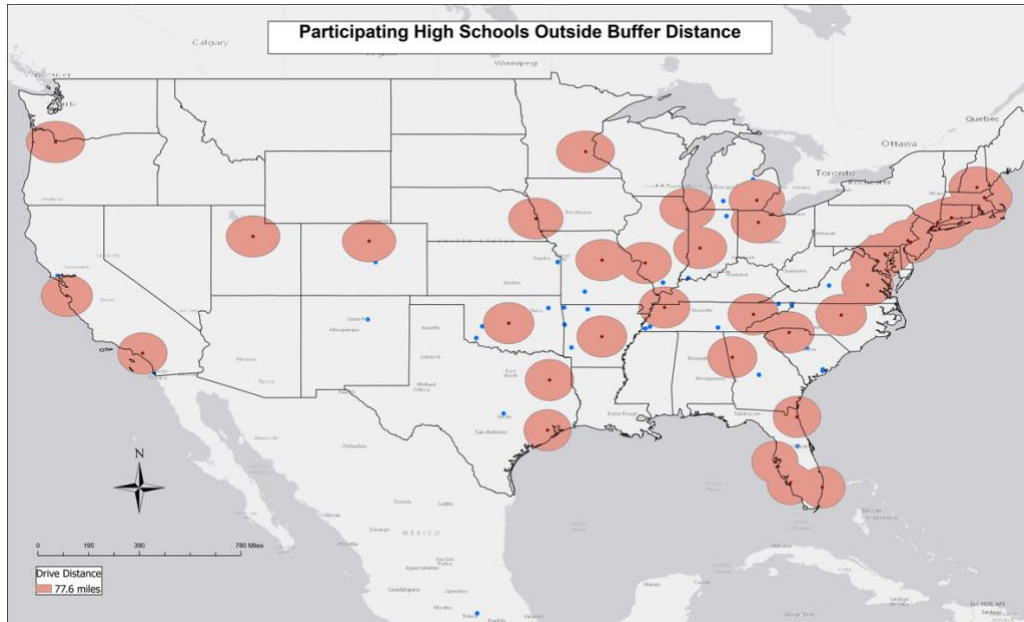


Figure 3. Buffer analysis, creating 77.8 mile radius buffers around each regional competition.

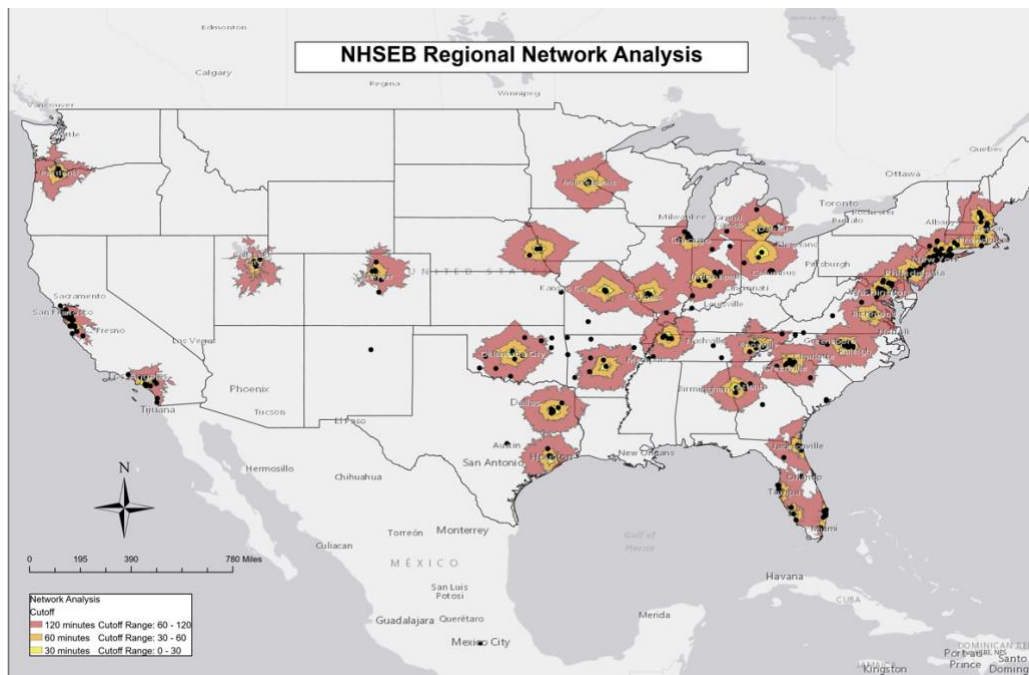


Figure 4. Network Analysis, creating 30-60-120-minute drive-time polygons around each regional competition.

Results

During the 2019 NHSEB season, there were 331 schools competing in 39 regional competitions, according to the data provided. After creating a kernel density map, the majority of regional and participating high school density could be seen along the east coast, followed by the Midwest (Figure 5, top). This is not unlike the density map created for the U.S. high-school aged population; however, the U.S. population density appeared to be

concentrated in smaller, but more numerous metro areas. This results in a more “spotted” appearance upon inspection (Figure 5, bottom). After normalizing both density maps on a scale from 1 to 10 and subtracting the density map of NHESB school density from the map of the U.S. population, areas of clear difference began to emerge (Figure 6). In the resulting map, areas of high density that remained appeared red and indicate a lack of participating high schools. Locations of low density were blue and indicated that creating new regionals and high schools in that area should be of low importance. At its highest, the density differences were as much as 6.60 and -5.55 at the lowest, with the majority of the country remaining near 0, indicating no significant differences in density.

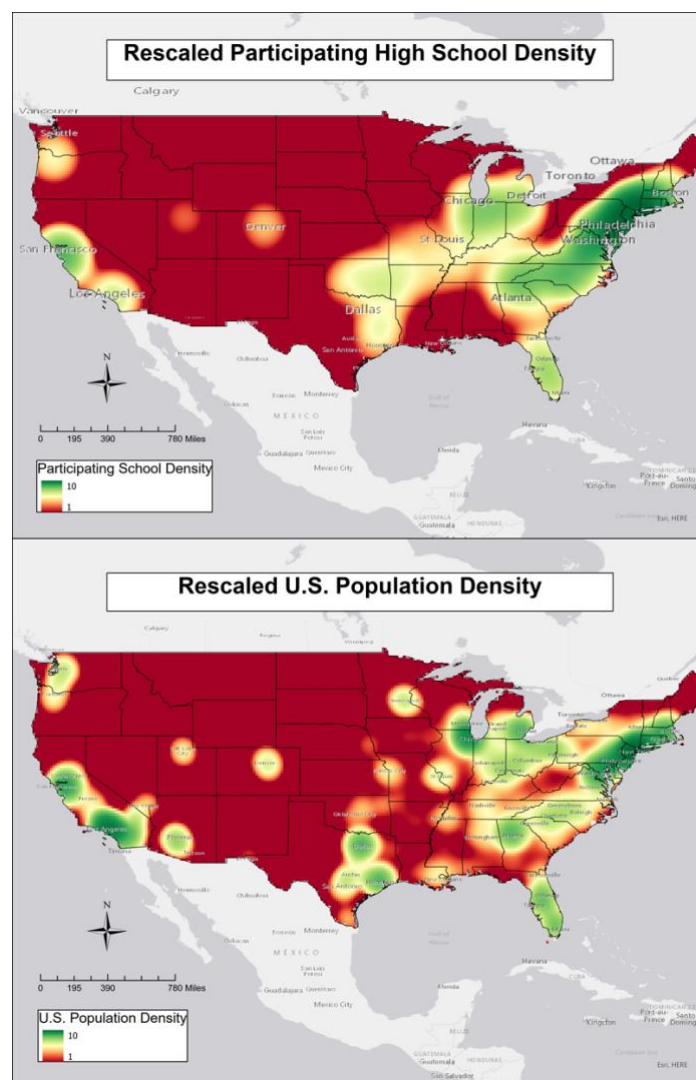


Figure 5. Rescaled density representations of participating high schools (top) and the U.S. population ages 15-19 (bottom).

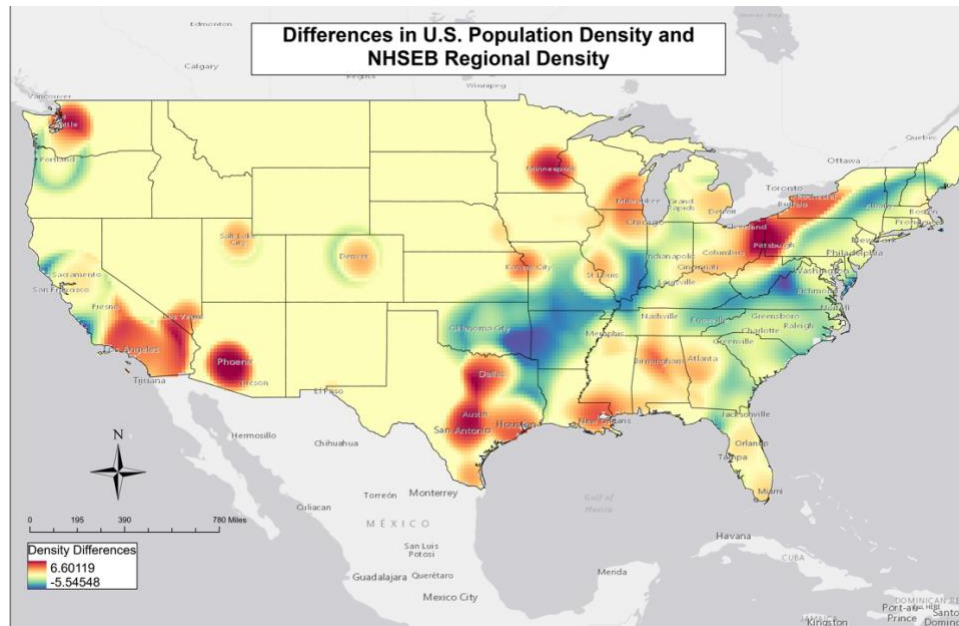


Figure 6. Differences between the two rescaled density maps in Figure 5.

As for the secondary analyses, the mean distance from each school to its corresponding regional location was 38.8, with a comparatively large standard deviation of 128.4. This large standard deviation can be largely attributed to the Connecticut Regional Ethics Bowl, as a single competing school – CIM-ORT in Mexico City, Mexico – had to commute over 2000 miles (Figures 7 and 8). The regionals with the lowest mean distance among their participating schools were almost exclusively in the Northeast Region of the U.S, where there exists a much denser concentration of both regionals and participating high schools, with the Philadelphia Regional Ethics bowl having the lowest mean distance. For the buffer analysis, a total of 33 high schools existed outside of the created buffer, mainly among the Midwestern states (Table 1).

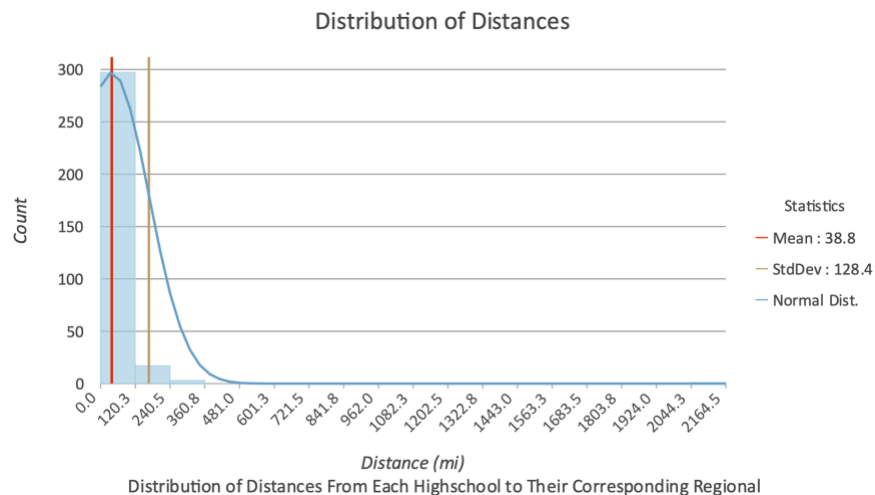


Figure 7. Distribution of distances from each high school to its assigned regional competition.

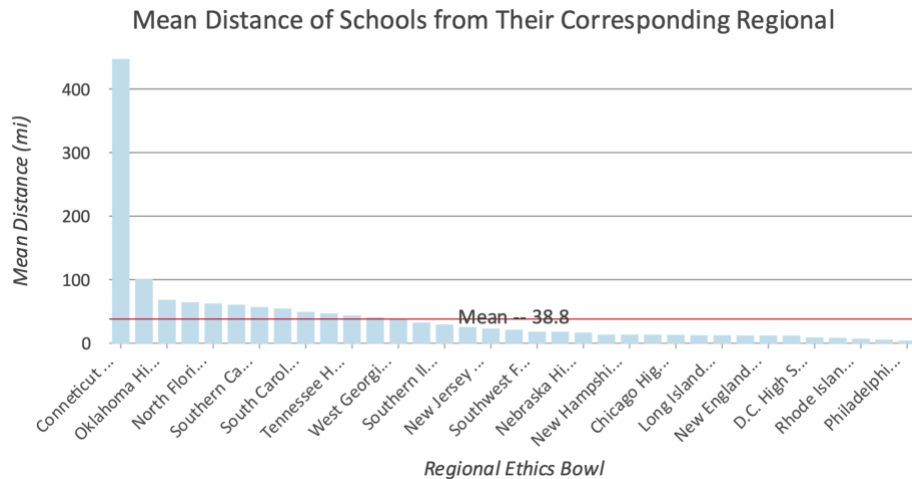


Figure 8. Mean distance from participating schools to each regional.

City	Region	School
Arlington	Tennessee	Arlington High School
Decatur	Arkansas	Decatur High School
Fort Smith	Arkansas	Future School of Fort Smith
Dierks	Arkansas	Dierks High School
Harrison	Arkansas	Harrison High School
Fort Smith	Arkansas	Northside High School
Memphis	Tennessee	Kingsbury High School
Montezuma	New Mexico	United World School
Manitou Springs	Colorado	Manitou Springs High School
Cuajimalpa de Morelos	Ciudad de México	CIM ORT (Mexico)
Evansville	Indiana	Benjamin Bosse High School
Austin	Texas	Westwood High School
Albion	Indiana	Benton Central Jr/Sr High School
Saginaw	Michigan	Saginaw Arts & Sciences Academy
Portage	Michigan	Portage Northern High School
Olathe	Kansas	Olathe South High School
Springfield	Missouri	Springfield Public Schools, Central High School
Apopka	Florida	Wekiva High School
Novato	California	Novato High School
Altus	Oklahoma	Altus High School
Cordell	Oklahoma	Cordell High School
Pryor	Oklahoma	Pryor High School
North Charleston	South Carolina	Palmetto Scholars Academy
Charleston	South Carolina	West Ashley High School
Columbia	South Carolina	J. L. Mann Academy of Math, Science and Technology
San Diego	California	High Tech High
Marion	Illinois	Marion High School
Kingsport	Tennessee	Dobyns-Bennet
Mountain City	Tennessee	Johnson County High School
Sewanee	Tennessee	St. Andrew's-Sewanee School
Buchanan	Virginia	James River High School
Warner Robins	Georgia	Warner Robins High School

Table 1. A list of high schools that existed outside of the 77.8 mile buffers created around each regional competition.

Discussion

The main goal of this project was to make evidence-based decisions concerning future locations for NHSEB expansion. Upon inspection of the map of kernel density differences (Figure 6), areas of the U.S. that should be considered for future regional competitions (as well as new participating high schools) include Seattle, WA; Phoenix, AZ, Seattle, WA; Minneapolis, MN; and Cleveland, OH, among others. Areas of low importance for new regionals and high schools included areas surrounding the Arkansas-Oklahoma border and, to a lesser degree, the East Coast. Because the yearly national competition is located in Chapel Hill, NC, it is understandable that most participating high schools and regional competitions are located near it along the East Coast. The goal of the

secondary analyses was to support the decisions above. The mean distance from a high school to its regional was 38.8. This value was skewed higher due to the very large distance from CIM-ORT to the Connecticut Regional Ethics Bowl, indicating it should be relocated to a closer regional competition – possibly to one of the two regionals in Texas. In addition, most of the high schools that were outside of the 77.6 mile buffer area were in the Midwestern states of Arkansas, Missouri, and Oklahoma, indicating a new regional competition should be created in this region that is central to those outlier high schools. This conclusion is interesting because, as stated earlier, these Midwestern states were of low importance for future regional locations, largely because of their low population density compared to participating high school density. Despite this, a new regional competition should be created in this area so as to not place any unfair burden on these high schools by making them drive much further to their regional than the rest of the country.

In conclusion, areas of high importance for future regionals and high schools include Seattle, WA; Phoenix, AZ, Seattle, WA; Minneapolis, MN; and Cleveland, OH. A new regional should also be created near the Arkansas-Missouri-Oklahoma border, as well, to reduce the distance that high schools in this area are forced to travel. My original hypothesis that the U.S. Northeast was comparatively over-represented turned out to be moderately true. This was exemplified when, after subtracting the two normalized rasters, the northeast region showed mostly low density, indicating that there were too many participating high schools in that region relative to the population density. The analyses of this project were largely limited by the data; had the data provided included more information like the number of team members and demographics for each school, more could be done. In addition, future analyses could examine the growth of the NHSEB over time, as the data provided only included the 2019-2020 competition season. Despite these shortcomings, it was more than possible to make general recommendations for the future directions of the NHSEB program using the data provided.

References

ESRI. 2020. "USA Counties." ArcGIS Hub. 2020.

https://hub.arcgis.com/datasets/48f9af87daa241c4b267c5931ad3b226_0.

The Parr Center for Ethics. 2020. "Mission and History." National High School Ethics Bowl. 2020.

<https://nhseb.unc.edu/mission-history/>.