

Major Changes: There have been no major changes

Accomplishments since last meeting:

- Discovered that there is easily generalizable way to gather the data. As this was an anticipated roadblock, as per my milestones, I have decided to focus on using data from 3 states. The three states that I initially selected last week were PA, MD, and NC, as these are three heavily gerrymandered states. After contacting the respective authorities at PA, MD and NC, I discovered that data from MD will be impossible for us to get easily, as none of the research group are MD residents. In substitution of MD, I have downloaded data from Wisconsin, which is freely and publically available. In addition to these steps, I sent the required snail mail requests (which is the only way to get it) to the PA Department of State's Bureau of Elections and NC's State Board of Elections.
- Tested cartogram algorithms on a dataset from the PA Department of Elections from 2010, provided to me by Wesley on Thursday. Based off outdated 2010 presidential data, the topology preserving cartogram algorithms took 22 hours to run and had numerous errors that make the cartogram unusable. I suspect an issue with the tolerance parameter in the Carto3F algorithm, as other algorithms work (albeit slowly) though do not preserve topology. I have contacted Sun to figure out how I can tune my parameters to meet the needs of topology preservation with such a large number of states.
- Reformatted parts of site so that more line spacing appears

Meeting Your Milestone: I met my milestone for this week, though I did not meet the milestone on Friday.

Surprises:

- Sun's Cartogram algorithm implementation was not robust and required extensive filtering of the shape file for preprocessing. It also ran significantly slower than expected, does not work across platforms as advertised, and will likely have other issues that make generating Cartograms from it difficult.
- Although I expected no generalized format for data across states, I expected that the data for individual states would be available for download. This was rather naïve, as states seem to make it purposely difficult to get this data.

Resources Needed: A really good Windows would be nice, as many of the programs that I have to use for the Cartogram part are built primarily for Windows machines. Running a VM on my Mac is inconvenient and makes the algorithms run slower. This is not a huge need, but would make my life easier.