**Lab Report**

Title: Better Title TBD: Power Grid Vulnerability and Marginalized Communities

Notice: Dr. Bryan Runck

Author: Cecelia Isaac

Date: 2/18/2021

**Project Repository:** https://github.com/CeceliaAi/GIS5572/tree/master/Final%20Project

**Abstract**

For my project, I will be analyzing American electrical grids. The storms in Texas have led to a surge of articles on this topic lately (Douglas, 2021), but the knowledge that our grids are vulnerable to climate change has been studied for many years (Weiss, 2019). On top of this, not all areas of the grid are protected or maintained equally. Marginalized communities are more likely to be impacted by power grid failure (Plumer, 2021). This project will attempt to pinpoint areas of highest risk in the hope that preventative measures will be taken. I will be looking at the ownership and maintenance of our grids against Census data on race, income, and female-headed households, likely at the tract level. I will compare this data to energy grid network data from the Homeland Infrastructure Foundation. Analysis will be done in ArcPro with some work in ArcPy. The result will be a series of maps exploring the grid’s vulnerability and relationship to these demographics.

**Problem Statement**

This lab’s problem is to download the data through a Python script, import it onto a map, and search for connections between the data. The first level of analysis will be straightforward: the electrical grids can be symbolized by owner, or peak capacity, or vulnerability, and overlaid on polygons showing the distribution of certain marginalized communities. After this visual analysis, I will use use ArcPro tools to analyze these relationships to pull out spatial patterns not immediately apparent upon visual examination.

*Table 1. Methodology Outline*

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **#** | **Requirement** | **Defined As** | **Spatial Data** | **Attribute Data** | **Dataset** | **Preparation** |
| 1 | Electrical grid shapefile; symbolized by vulnerability and possible ownership | Vector dataset; features also have associated substations | Polyline network dataset of the entire US | Service status, ownership | https://hifld-geoplatform.opendata.arcgis.com/datasets/electric-power-transmission-lines?geometry=-102.431%2C44.299%2C-86.819%2C46.986 | Possibly clip to study area; possibly run analysis to discover most vulnerable areas |
| 2 | Control Areas, along with electrical line ownership, to help determine responsibility of grids in various regions | Vector dataset | Polygons of retail service areas of energy providers | Websites, peak months, unplanned outages, peak capacity | https://hifld-geoplatform.opendata.arcgis.com/datasets/control-areas | Possibly symbolized separately, or possibly joined with grid data (or both) to give a complete picture |
| 3 | Represent marginalized communities | Census demographic numbers symbolized by tract | Polygon |  | https://www.nhgis.org/ | Possibly calculating columns to get ratios per population |

**Input Data**

I will be using publicly available Census and electrical infrastructure data. I believe Census data at the tract level will be a sufficient level of granularity, but I will be researching further before downloading. Two shapefiles will come from Homeland Infrastructure Foundation-Level Data (HFILD) Open GP – Energy: polyline data on electrical lines, and control areas, which is polygon data that shows which parties are responsible for energy supply of the electrical grids in the area. The electrical line data also shows ownership of the grids. HFLID also has point data for electric substations, but I don’t know yet if that will be relevant, since it contains much the same information as the electrical line shapefiles.

*Table 2. Data*

|  |  |  |  |
| --- | --- | --- | --- |
| **#** | **Title** | **Purpose in Analysis** | **Link to Source** |
| 1 | Control Areas | Energy stakeholders in various regions of the US | https://hifld-geoplatform.opendata.arcgis.com/datasets/control-areas |
| 2 | Electric Power Transmission Lines | Shapefile of the energy grid network; will be analyzed for vulnerability; attribute data includes power company ownership | https://hifld-geoplatform.opendata.arcgis.com/datasets/electric-power-transmission-lines/data?geometry=-102.431%2C44.299%2C-86.819%2C46.986 |
| 3 | Electric Substations | Shapefile of the energy stations; not sure yet how relevant it will be since the information is the same as the power line data | https://hifld-geoplatform.opendata.arcgis.com/datasets/electric-substations?geometry=-110.440%2C43.508%2C-79.920%2C48.828 |
| 4 | Census Data | Not downloaded yet since I haven’t decided which datasets; demographic data at the tract level of marginalized communities; will be used to analyze intersections of these communities with weak infrastructure | https://www.nhgis.org/ |

**Methods**

The project process will start with building an ETL to download data from the respective sources. This data will be loaded into ArcPro where it will be cleaned. Several basic maps will need to be created. The Census data will be symbolized by the various demographic data ratios. The electrical grid data will be symbolized as well, but I have not determined which variables are most relevant. I am interested in displaying vulnerability, ownership, and capacity for the grid. After, the feature layers will be displayed together and overlapping areas of high grid vulnerability and high incidence of marginalized communities will be highlighted.

**Results**

The results will hopefully address the problem statement by showing a spatial pattern. It is my hypothesis that more vulnerable grids will overlap with vulnerable communities, since this pattern is well-established in other studies into infrastructure (for example, the Flint water crisis). I am also interested in seeing how private energy corporations play a role in grid upkeep and if there is a pattern of negligence for any one company.

To this end, I will be producing several maps as my output in order to tell a visual story about the state of our energy grids. At this time, I am planning on writing in a blog or Story Map so the content is accessible to audiences without previous in-depth knowledge.

**Results Verification**

I am not sure how to verify the results of an original study. If my results match up with similar studies on vulnerability of electrical grids and marginalized communities, that will be one angle. However, I have not yet found a study exactly like my project idea, only tangential research. Another possible way to verify my results is to create a reproducible workflow so others can test my conclusions.

**Discussion and Conclusion**

This project was my second idea since I could not find the data I wanted for the first, so I do not yet know exactly which elements will become the focal point of my project. I still need to determine how vulnerability of grids is assessed and whether it is even within the feasible scope of this project to calculate that myself, or if I can find a dataset that already has this information, as a spatial file or to be used as ancillary data. I do not yet know what spatial patterns will appear, but I have several hypotheses since I have been interested in electrical grids for a while and have followed the research on them. I hope to be able to use some of ArcPro’s more advanced analysis tools in my project. Even without all of the details sorted out, I feel confident I can successfully accomplish the main problem statement, which is to pinpoint areas of high risk.

**References**

Erin Douglas, Kate McGee, Jolie McCullough. (2021, February 17). Texas leaders failed to heed warnings that left the state’s power grid vulnerable to winter extremes, experts say. *The Texas Tribune*. <https://www.texastribune.org/2021/02/17/texas-power-grid-failures/>

HIFLD Open GP - Energy. (2020a). *Electric Substations* [Shapefile]. Homeland Infrastructure Foundation-Level Data (HIFLD). <https://hifld-geoplatform.opendata.arcgis.com/datasets/electric-substations?geometry=-110.440%2C43.508%2C-79.920%2C48.828>

HIFLD Open GP - Energy. (2020b). *Control Areas* [Shapefile]. Homeland Infrastructure Foundation-Level Data (HIFLD). <https://hifld-geoplatform.opendata.arcgis.com/datasets/control-areas>

HIFLD Open GP - Energy. (2020c). *Electric Power Transmission Lines* [Shapefile]. Homeland Infrastructure Foundation-Level Data (HIFLD). <https://hifld-geoplatform.opendata.arcgis.com/datasets/electric-power-transmission-lines/data?geometry=-102.431%2C44.299%2C-86.819%2C46.986>

King, L. (2021, January 28). *How America’s Power Grid Is Vulnerable To Undetected Cyberattack*. <https://www.forbes.com/sites/llewellynking/2021/01/28/how-the-supply-chain-in-heavy-bulk-power-equipment-is-vulnerable-to-undetected-cyberattack/?sh=ae9ae027213a>

Matthew Weiss. (2019). An assessment of threats to the American power grid. *Energy, Sustainability and Society*, *9*(18). <https://energsustainsoc.biomedcentral.com/articles/10.1186/s13705-019-0199-y>

Mike M. McMahon. (2017, November). *Mapping the Vulnerability and Strength of the Power Grid*. <https://isen.northwestern.edu/mapping-the-vulnerability-and-strength-of-the-power-grid>

Brad Plumer, Hiroko Tabuchi. (2021, February 17). The Far-Reaching Effects of the Storm, on Power and People. *New York Times*. <https://www.nytimes.com/2021/02/17/climate/storms-power-outage-newsletter.html>

**Self-score**

|  |  |  |  |
| --- | --- | --- | --- |
| **Category** | **Description** | **Points Possible** | **Score** |
| **Structural Elements** | All elements of a lab report are included **(2 points each)**:  Title, Notice: Dr. Bryan Runck, Author, Project Repository, Date, Abstract, Problem Statement, Input Data w/ tables, Methods w/ Data, Flow Diagrams, Results, Results Verification, Discussion and Conclusion, References in common format, Self-score | 28 | **28** |
| **Clarity of Content** | Each element above is executed at a professional level so that someone can understand the goal, data, methods, results, and their validity and implications in a 5 minute reading at a cursory-level, and in a 30 minute meeting at a deep level **(12 points)**. There is a clear connection from data to results to discussion and conclusion **(12 points)**. | 24 | **24** |
| **Reproducibility** | Results are completely reproducible by someone with basic GIS training. There is no ambiguity in data flow or rationale for data operations. Every step is documented and justified. | 28 | **28** |
| **Verification** | Results are correct in that they have been verified in comparison to some standard. The standard is clearly stated **(10 points)**, the method of comparison is clearly stated **(5 points)**, and the result of verification is clearly stated **(5 points)**. | 20 | **20** |
|  |  | 100 | **100** |