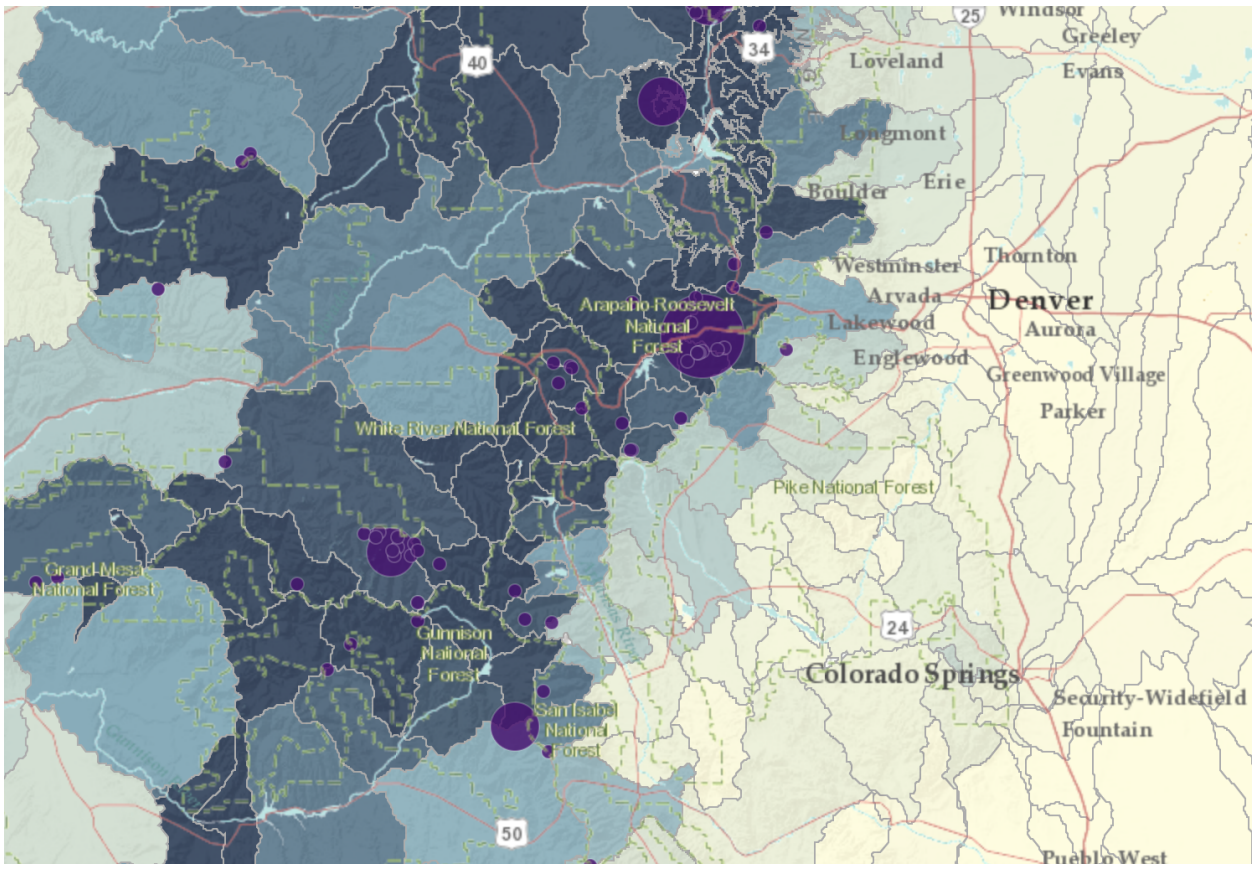

Are we seeing UFOs because we want to? Mapping UFO sightings against possible scientific explanations.

A proposal for CS credited half-credit independent project

9 December 2021

—

Ella Neurohr

Class of May, 2022

Overview

I hope to do an independent half-block using my combined GIS and computer science experiment. I aim to map UFO sightings in the United States against various scientific explanations for the sightings.

Geographic information systems, or GIS, is a visual and spatial means to represent and analyze data. They are interactive, so layers can be manipulated and data points clicked for more information.

I've worked in the GIS lab under Matt Cooney since freshman year, and while he is technically in IT and cannot give credits, he instructs numerous classes including many in the environmental studies department. I first met him in my FYE, where I made this project (I was actually given this project after a student failed to do the celestial mapping the year prior). I want to have something to show for my years working in the GIS lab and that I can apply what I've learned in my computer science studies to interdisciplinary studies.

I'm asking for it to qualify as a 200 level CP credit, since I want to be flexible in deadlines because I will be working on this alongside my block, multiple jobs, and being with my friends senior year.

I imagine the final product to look something like this, but with many more data layers. This is also where I will get my data on sightings, as "the dataset contains over 80,000 records of UFO sightings over the past century (between 1910 and 2014)." Many of these data points have references to the facts of the sighting and are interactive.

Below are some of the natural phenomena that are considered to be likely explanations for UFO sightings. Linked are data sets I've already found.

- Electromagnetic fields
- Location of military installations
- Proximity to metropolitan areas
- Calleda (red) sprites *
- Fireballs
- Meteors
- Atmospheric conditions
    - Thunderstorms account for a number of the potential explanations, including red sprites and ball lighting.
    - Not certain how this will be represented yet. Perhaps zones of atmospheric trends or annual recorded thunderstorms.
- Ball lightning *

        - ➢ An unexplained phenomenon described as luminescent, spherical objects that vary from pea-sized to several meters in diameter. They can last significantly longer than normal lightning.
        - ➢ One of the most common things mistaken for UFOs, and historical accounts of them go back to the 17th century.
        - ➢ Very hard to find spatially consistent data. There are countless anecdotal reports, and I've emailed a professor at New Mexico Tech inquiring about a dataset he may have.
    - ❖ Lenticular clouds
        - ➢ Form when moist, stable air flows over mountains and the temperature at the crest of the wave drops below the local dew point.
        - ➢ One of the most common things mistaken for UFOs.
        - ➢ Will be represented as zones where formation is most likely.
    - ❖ Tectonic faults
        - ➢ The Tectonic Strain Theory developed by Dr. Michael Persinger "predicted that luminous phenomena and associated physical effects were produced by manifestations of tectonic strain that often precede by weeks to months seismic events within the region."

* Data I presume I will need to scrape as I have not yet found a good collection of data to download.

Goals

I. Strengthen my data retrieval and cleaning skills.

II. Become confident in ArcMap Pro (I used ArcMap before the pandemic).

III. Develop and publish an interactive web app.

IV. Learn how to do data analysis in Python.

V. Figure out if UFOs are real once and for all.

Technical Frameworks

These are the technical frameworks, for coding and not, that I plan to use:

- ❖ Python3
    - ➢ Selenium: web driver used for scraping data
    - ➢ Pickle: serialization module
    - ➢ Pandas

    - ➢ Data visualization library, TBD. Most likely some combination of Matplotlib, Seaborn, Numpy, and Plotly.
- ❖ Excel
- ❖ ArcMap Pro
- ❖ ArcGis Web AppBuilder
    - ➢ may end up using an alternative

Milestones

I'd like to begin work on the project over winter break and half block and continue it through block 5. Matt will be teaching the Intro to GIS half block so he will be around the lab then, but I don't need him to be there the entire time I'm working.

I. Winter break/half block

Research, data collection, data cleaning and formatting. This may involve combining acquired data sets to get the intersection of data I require, or scraping websites for geospatial information. I will need to decide on a way to assign a value factor to each data set that when multiplied against distance from a sighting, will approximate the likelihood that the sighting could have been that natural occurrence. This will differ for data points (meteors, military bases) and spatially sparse data (electromagnetic fields, lenticular clouds).

II. Week 1, Block 5

Mapping and visual analysis. This involves physically mapping all of the data, deciding on aesthetics, and making visual analysis in order to get a better idea of how exactly I plan to analyze the data.

III. Week 2-3 Block 5

Data analysis. The most time is allotted for this, because it requires significant research, learning, and experimentation in order to get conclusions that can be visualized and explained. ArcMap attribute tables (the underlying data for the visualization) are CSVs, which I know how to read into a Python program and serialize for more efficient retrieval. One of the goals of the first milestone was to figure out a way to calculate a value for the UFO-alternatives that is based on the spatial relation between the given alternative and nearby sightings. If there are multiple alternatives within a decided range of a sighting, they must be compounded into an aggregated rating of how likely the sighting was to be false.

It will not matter how many explanations are present within a range of the sighting, just the

The scale for likeliness that a sighting was misinterpreted will be the same across the sightings. This means that it will not matter how many explanations are within the given range, just the combined quality of their values, where non-present explanations do not change the score. This is because it has proven difficult to find some datasets consistently across the country (i.e. ball lighting sightings in a county or state) and some layers, like the zones where lenticular clouds form, are approximations.

IV. Week 3-4, Block 5

Compilation. I will combine all of the conclusions and supporting information onto the web app. The maps will be interactive with background information for each of the layers, and there will be a conclusions section with the results of milestone 3.

Final remarks

The support for my conclusions will be correlative, but I hope that in using enough data sets it will reveal a pattern (or not one at all).

There are significantly fewer new tools that I will need to learn in comparison to my team software project, so I am confident in my self-teaching abilities.