**Capstone #1 Proposal**

**Project Idea #1:**

* **Topic (Title): Earth In Human Hands: The Northwest Passage**
* **High level description of project:** This project is intended to be the first of three that analyze the experiences of and changes to the whole Earth system in the anthropocene. In this lens, datasets regarding arctic sea ice extent (daily and annually to 1979) are combined with monthly temperature data from the northern hemisphere (back to 1850) and *possibly*data from tweets regarding public sentiment on climate change in current times. The goal here is to visualize a rate and direction of change in average monthly temperatures in the northern hemisphere, compare that change against a baseline mean temperature, and view the impact of changes through monthly changes in arctic sea ice as well as the annual minimum extent of arctic sea ice. It would be very interesting to see how rates of change have changed over time, and if there is a relationship or correlation between minimum extents from year to year. Ultimately, I would want to try and predict, based on the trends discovered in this data, if the Northwest Passage is likely to come into existence anytime soon, and if so, a ballpark estimate of when. It would be really interesting to see how public sentiment of climate change in North America holds up against the economic potential of a major shipping route along Canada’s northern coastline.
* **Data Source(s):** 
  + **Websites and/or databases (linked here):**

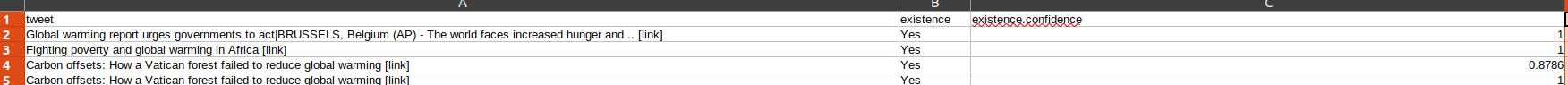
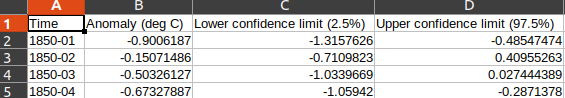
**https://www.kaggle.com/nsidcorg/daily-sea-ice-extent-data**

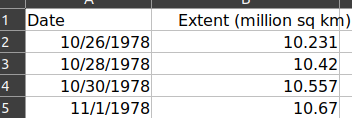
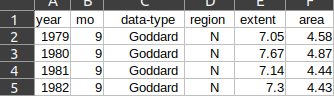
**https://www.metoffice.gov.uk/hadobs/hadcrut5/data/current/download.html**

**https://climate.nasa.gov/vital-signs/arctic-sea-ice/**

**https://data.world/crowdflower/sentiment-of-climate-change/workspace/file?filename=1377884570\_tweet\_global\_warming.csv**

* + **Number of data points (rows, pages scraped): ~21,000**
  + **Screenshot of first 5 rows of the database OR the website with the numbers you need:**



**

* **Potential Future Employer:** Any organization that operates at the intersection of big data and climate change planning/advocacy, such as:

NASA

The Rocky Mountain Institute

USDA/USGS/USFS/US Dept. Of Homeland Security

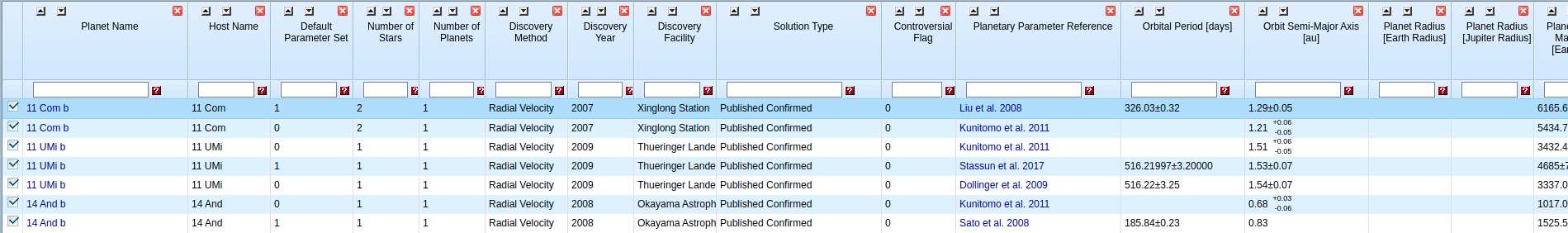
NOAA

Research Universities

Municipalities

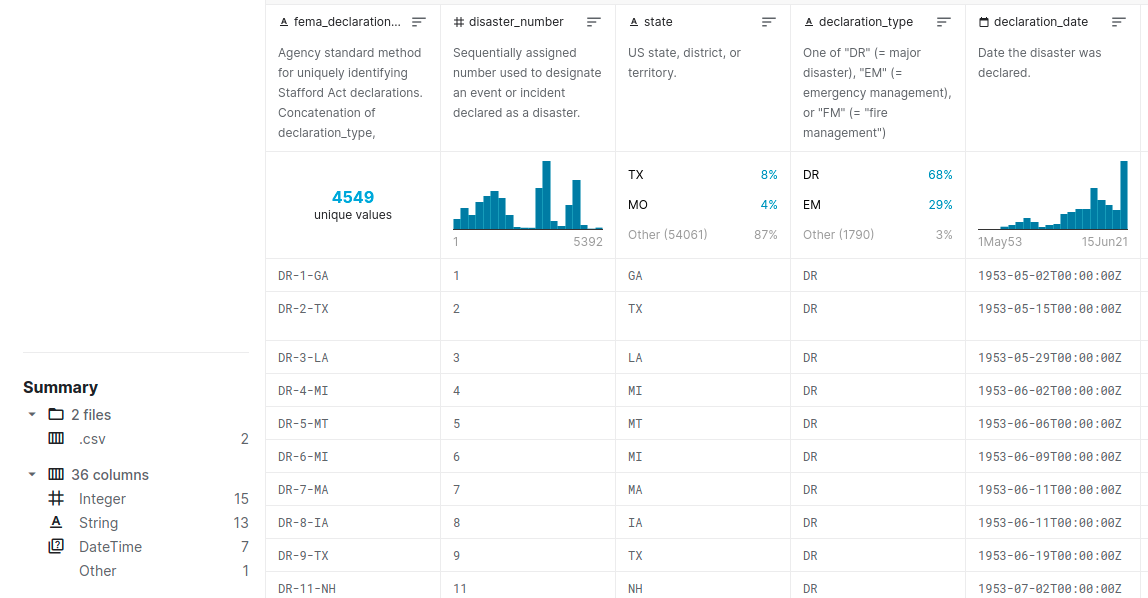
Climate-focused nonprofits

**Project Idea #2:**

* **Topic (Title): The Maximum Likelihood...of SETI Succeeding**
* **High level description of project:** This dataset apparenty contains every exoplanet discovered, current into 2021. With the ability to find exoplanets of similar size and mass to Earth, as well as orbital period, it would be interesting to see how many of the exoplanets discovered “look like” Earth in terms of their characteristics. Such conjecture would allow us to find a sort of “average rate” of the discovery of “Earthy” planets, which we could compare against the Drake Equation and get a (really really) rough estimate of when humans can expect to find a planet that might actually have life on it, or find one that could support our lives (this reasoning based on “Goldilocks zone”/human perceptions of what life is and where it can exist!). Leveraging Gaia’s 1-billion star database to find the distances to interesting candidates isn’t feasible, but certainly it is possible to do some extra work on a few candidate exos to see how(ish) far away they are, if they are worth more detailed observation, and if anyone currently living could actually expect to see anything come of that exploration.
* **Data Source(s):** 
  + **Websites and/or databases (linked here):** 
    - **https://exoplanetarchive.ipac.caltech.edu/cgi-bin/TblView/nph-tblView?app=ExoTbls&config=PS**
  + **Number of data points (rows, pages scraped): 29477**
  + **Screenshot of first 5 rows of the database OR the website with the numbers you need:**
  + 

**Project Idea #3:**

* **Topic (Title): Earth In Human Hands: Naturally Disastrous?**
* **High level description of project:** These two datasets provide coarse and fine resolutions of data regarding the count, type, damage cost, and locality of various natural disasters in the United States from either 1953 – 2021, or 1900 – 2021. It would be really interesting to see how the prevalence of natural disasters has shifted over time, and if there are specific kinds of natural disasters that have shifted in either frequency or spatial distribution in the last 50 or 100 years. Personally, I would love to see if there has been a significant change in the cost of natural disasters and the frequency of natural disasters that cost over a certain threshold. If time and skill allow, I would like to follow the trends of costly or widespread natural disaster types into the regions they occur most frequently.
* **Data Source(s):** 
  + **Websites and/or databases (linked here):** 
    - **https://www.kaggle.com/dataenergy/natural-disaster-data**
    - **https://www.kaggle.com/headsortails/us-natural-disaster-declarations**
  + **Number of data points (rows, pages scraped): ~13,000**
  + **Screenshot of first 5 rows of the database OR the website with the numbers you need:**





* **Potential Future Employer:** FEMA?