Detailed Report – Team HunaTek-Kalman

1. How does your visual inform a decision or action that furthers one or more of the key competition SDGs (zero hunger, clean water and sanitation, climate action)?

The increase in both frequency and intensity of natural disasters is one of the most conspicuous and immediately devasting consequences of anthropogenic global warming. As our climate's ever-increasing volatility will cause disasters to continue to propagate, we think it will become increasingly important to evaluate their impacts in areas outside of "typical" territories. We chose to focus on wildfires, a threat that is both dramatically striking and relatively suited for Earth-observation data analysis, and their presence in the eastern half of our continent. Our visualization involves not only the spread of the fires themselves, but also tracks the corresponding effects on air quality that can impact populations far removed from the fire itself. By raising awareness of under-studied Eastern US wildfires and providing a tool to track this trend, we demonstrate the growing dual risk of catastrophic localized fire damage and persistent broad-reaching air quality detriment. We underscore the need for the development proactive and flexible fire management techniques such as those discussed by Prichard, et al. (2021). Adaptive fuel treatments, prescribed burns, forest thinning, post-fire management, and more all need to be carefully considered and appropriately applied to at-risk regions. A sensitive fire management regime in the Eastern US will help mitigate future dangers and guard against potential losses of biodiversity and forest cover that would in turn exacerbate our climate crisis.

2. How did you create your submission? Include the tools you used (e.g., Python, Excel, specific python packages), how you processed the data, and (if applicable) how you managed your codebase. If you have a public repository with code, you can share a link here.

Our visualization leverages data from the Environmental Protection Agency (EPA)'s Air Quality System (AQS) data repository, which aggregates air pollution data collected by the EPA and other agencies from monitors across the United States. Additionally, we pulled burned area data from the MODIS Terra and Aqua combined MCD64A1 Climate Modeling Grid Burned Area data product. The burned area data was downloaded as Hierarchical Data Format (HDF) files from the University of Maryland sftp server. Also from MODIS, we used the archive of MODIS C6.1 Active Fire data directly from the Fire Information for Resource Management System (FIRMS) website going back to 2000, collecting data for the entire United States and Canada. We processed the air quality and the burned area data using the Anaconda distribution of Python 3 through Jupyter Notebook, using the osgeo/gdal, numpy, os, and pandas packages to ingest and transform the data. In order to focus our analysis, we narrowed down each dataset so that they only contain datapoints spatially located within the geographic bounding box [-87.75, 25.0, -52.75, 62.5], encompassing the Eastern United Dates and Canada.

Our air quality data analysis involved gathering, cleaning, and combining several datasets into one. The first piece of the puzzle consisted of 23 separate Comma-Separated Variable (CSV) files of daily air quality indices from the EPA's AQS repository. We processed the data using the pandas package in Python, looping through each of the files to create one consolidated data frame with a narrowed focus on particulate matter of 2.5 micrometers or less (PM2.5) and carbon monoxide (CO). The second piece of the puzzle involved reading the dataset of Core Based Statistical Areas (CBSAs) and associated global coordinates from the United States Census Bureau into a data frame. We used pandas to standardize the CBSA naming convention of both datasets and merge the two data frames into one. Though there were some missing coordinates and some wayward CBSA names, we were able to troubleshoot all of the problems until we ended up with a complete, concise, AQI by global coordinates database.

To process the MODIS Burned Area data the HDF files were pulled into Python, using osego and pandas. The osego/gdal library was utilized in navigating through the MODIS Burned Area HDF files. The MODIS Burned Area data is formatted in 720 x 1440 array, where each element in the array represents a grid of 0.25 degrees. We subset this array based on the bounding box mentioned above and reshaped it into a data frame where each row represents the burned area of a single grid for a given month, with the latitude and longitude representing the center of the grid.

The visualization was constructed in Power BI, using the three aforementioned datasets. In PowerQuery, we binned the data to reduce the temporal granularity down to one day and the spatial granularity down to 0.01°, which markedly improved performance of the visualizations and would allow for faster refreshes in the future with more years of data. All of the visuals are included with the standard free version of Power BI Desktop, with custom measures to explore the variables; i.e. Average AQI by Year, Cumulative Area Burned, and Cumulative Fire Radiative Power. The most geographically pertinent data came from the MODIS Burned Area set, which was displayed using the standard ArcGIS map, with the built-in heat map clustering and symbology on the Streets basemap. Our visualization, associated datasets, and code base are stored on this Github repository.

3. What motivated you to choose this topic?

Our team was initially drawn to the topic of wildfires by our recollection of Canada's extreme 2023 fire season. Since we are all based in Northern Virginia, hurricane warnings might feel relatively par for the course, but receiving air quality alerts last June was quite the foreign experience – one that drove home the climate change-driven proliferation of natural disasters in recent years. Perhaps unsurprisingly, most research and attention on US wildfires have focused on the western part of the country (including all but one of our sources cited below). With that in mind, we decided to create a wildfire data product centered around the inclusion of the Eastern US. We intend for this visualization both to expand upon the existing, Western-US-based discussions of fire management and to promote awareness of those discussions in regions less accustomed to handling fire disasters.

Figure 1: An unusual – and alarming – image, as wildfire haze shrouds the Capital monuments on June 7, 2023. Photo by Win McNamee via Getty Images

4. How did you learn about the broader context of your chosen issue (e.g., historical, social, political)? This could include drawing on the lived experiences of team members, reading articles and literature, conducting interviews with community members, etc. Did what you learned change your approach?

As mentioned above, we arrived at this project through reflecting on our own and our communities' experiences with the 2023 wildfire season. To expand our perspective, we reviewed several scholarly articles on fires and their impacts, especially as connected to climate change. Pausas and Keeley (2021) and Williams, et al. (2019) both discuss the correlation of anthropogenic global warming to the increased extent of wildfires, addressing factors such as droughts, dry fuel availability, fire weather, and non-climate related human impacts. Initially, we considered focusing our visualization on wildfires and forests specifically, and Davis, et al. (2019) and Stevens-Rumann, et al. (2018) both develop a particularly vicious cycle: as fires are augmented by climate change, it becomes harder for forests to recover post-blaze, while this loss of forested land in turn exerts negative effects on the climate. However, we eventually settled on air quality as our parameter of interest, looking at sources such as Liu, et al. (2016), who offer some excellent maps – that are abruptly cut off at a mid-Dakotan longitude. Relatively late in our process, we came across a recent letter from Donovan, et al. (2023) that felt quite validating. The authors discuss the upwards trend in large wildfire activity in the Eastern US, particularly the southeast, and the proportionately high risk to human populations in these densely populated areas. Finally, Prichard, et al. (2021) close their solutions-oriented paper on a striking contextual note that bears repeating here: the "close alignment" of modern wildfire resiliency strategies with Indigenous knowledge (21). The problems highlighted in our visualization (problems that have been propelled by European-colonial industrialization) will require solutions that fuse modern scientific knowledge with Indigenous fire-management skills.

5. What are the ethics and/or equity issues you considered? What are some possible strategies or approaches for addressing them?

In crafting our visualization, we strove to present the data as intuitively and accessibly as possible. While some degree of tech literacy is required to access our project, the information does not require any significant degree of technical knowledge to understand, and we have provided our conceptual and contextual explanations in straightforward, digestible language. Our code for gathering/processing the data and our PowerBI visualization setup are readily available at the GitHub site linked in section 2 and are intended to be easily reproducible and modifiable.

In terms of the content itself, fires would initially seem to be agnostic in their destructive nature. Yet, a number of biases persist in terms of disproportionate impact. While rural areas are more likely to be directly affected by the flames, urban areas already experience worsened air quality that can become drastically unhealthy with the addition of wildfire haze. Socioeconomic discrepancy arises as wealthier individuals are more capable of removing themselves from dangerous regions. Disabled persons and those with pre-existing conditions may also be at higher risk of suffering from wildfire effects. Our data, operating at quite a high level, does not account specifically for all of these variables, and future work should improve on addressing these and other inequities.

6. Would your team like to share the URL of an interactive visualization?

Please reference our Power BI file to interact with the visualization.

References

Davis, Kimberley T., Solomon Z. Dobrowski, Philip E. Higuera, Zachary A. Holden, Thomas T. Veblen, Monica T. Rother, Sean A. Parks, Anna Sala, and Marco P. Maneta. "Wildfires and climate change push low-elevation forests across a critical climate threshold for tree regeneration." *Proceedings of the National Academy of Sciences* 116, no. 13 (2019): 6193-6198.

Donovan, Victoria M., Raelene Crandall, Jennifer Fill, and Carissa L. Wonkka. "Increasing large wildfire in the eastern United States." *Geophysical Research Letters*, 50 (2023): e2023GL107051.

Liu, Jia Coco, Loretta J. Mickley, Melissa P. Sulprizio, Francesca Dominici, Xu Yue, Keita Ebisu, Georgiana Brooke Anderson, Rafi FA Khan, Mercedes A. Bravo, and Michelle L. Bell. "Particulate air pollution from wildfires in the Western US under climate change." *Climatic change* 138 (2016): 655-666.

Pausas, Juli G., and Jon E. Keeley. "Wildfires and global change." *Frontiers in Ecology and the Environment* 19, no. 7 (2021): 387-395.

Prichard, Susan J., Paul F. Hessburg, R. Keala Hagmann, Nicholas A. Povak, Solomon Z. Dobrowski, Matthew D. Hurteau, Van R. Kane et al. "Adapting western North American forests to climate change and wildfires: 10 common questions." *Ecological applications* 31, no. 8 (2021): e02433.

Stevens-Rumann, Camille S., Kerry B. Kemp, Philip E. Higuera, Brian J. Harvey, Monica T. Rother, Daniel C. Donato, Penelope Morgan, and Thomas T. Veblen. "Evidence for declining forest resilience to wildfires under climate change." *Ecology letters* 21, no. 2 (2018): 243-252.

Williams, A. Park, John T. Abatzoglou, Alexander Gershunov, Janin Guzman-Morales, Daniel A. Bishop, Jennifer K. Balch, and Dennis P. Lettenmaier. "Observed impacts of anthropogenic climate change on wildfire in California." *Earth's Future* 7, no. 8 (2019): 892-910.