Project Wildfire

CSPB 4502 – Group 3

Team

- Natalie Dreher

- Ronald Durham

- Grant Fairbairn

Project Description

- We are utilizing the data set "1.88 Million US Wildfires" for wildfires occurring in the United States between 1992 and 2015.

- The dataset includes attributes relating to discovery date, containment date, fire size, and geographic location.

- This raises interesting questions about whether certain areas are more or less likely to experience wildfires and whether certain regions are more effective in fighting existing fires. We can also explore whether wildfires have increased or decreased in number and severity over time.

Prior Work

- The dataset has been downloaded over 20,000 times.

- Others have attempted to answer the three "inspiration questions" listed on Kaggle.com:

    - Have wildfires become more or less frequent over time?

    - What counties are the most and least fire-prone?

    - Given the size, location and date, can you predict the cause of a wildfire?

- We can review the results of some contributors' analyses on these questions to develop new ones looking for interesting patterns in the data.

Prior Work

- There are also publicly available articles discussing trends in wildfire data over multiple decades.

- This article, for example, notes that the annual number of wildfires decreased between 1991 and 2020, but the area burned in those fires increased sharply, particularly in the western United States:

  - https://www.rff.org/publications/explainers/wildfires-in-the-united-states-101-context-and-consequences/

Datasets:

- List of datasets to use

  - 1.88 Million US Wildfires

  - 24 years of geo-referenced wildfire records

- Where found (URL and who is supplying the data, e.g., NASA)

  - The data was collected using funding from the U.S. Government and can be used without additional permissions or fees and should be cited as such:

  - Short, Karen C. 2017. Spatial wildfire occurrence data for the United States, 1992-2015 [FPAFOD20170508]. 4th Edition. Fort Collins, CO: Forest Service Research Data Archive. https://doi.org/10.2737/RDS-2013-0009.4

  - https://www.kaggle.com/datasets/rtatman/188-million-us-wildfires?resource=download

- Whether it you have it downloaded (on who's machine)

  - Ron has the entirety of the sqlite file downloaded on his local machine

  - Main "fires.csv" file is uploaded and shared to group members

Proposed Work:

- Convert from sqlite to csv file

- We will need to do some data cleaning as some of the rows are incomplete. Will most likely utilize the attributes that will be useful in answering our questions and remove excess or ones not applicable to this project.

- We will need to further understand and preprocess a couple of the attributes specifically, discovery date and cont date as they are supposed to be dates of the fire but are recorded as "2,453,555.500" and similar numbers. However, for the most part the data in each of the other attributes will be usable as is.

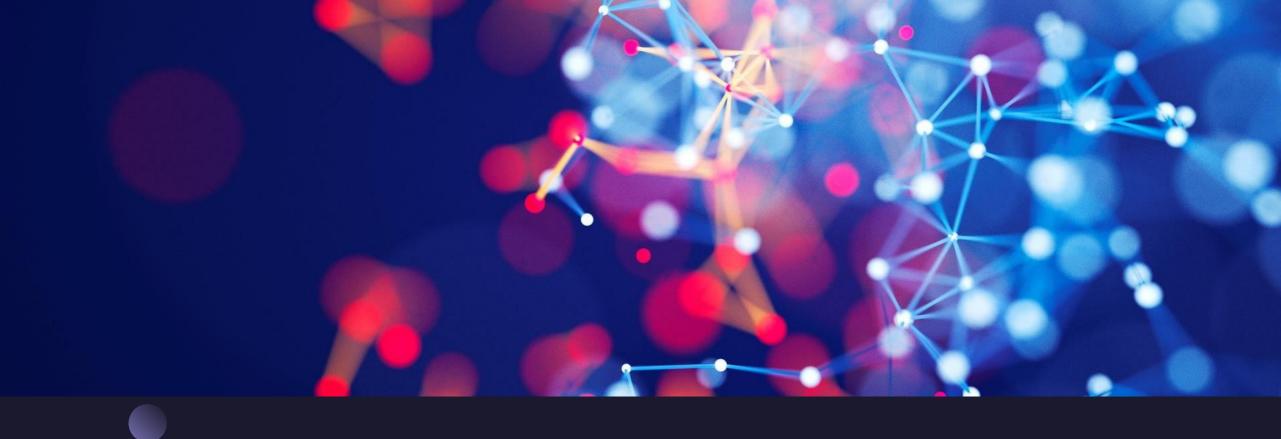# Tools to be used

PYTHON LIBRARIES:

- NumPy

- Pandas

- Matplotlib

- SQLite3

OTHER TOOLS:

- SQLite

Evaluation

- Have we expanded on existing work in a meaningful way?

- Have we answered the questions we set out to answer?

- Can our work be used in a way that's helpful to the wider world?

Summary

[insert text]

Thank You

https://github.com/rodu4835/CSPB4502_ProjectWildfire