Climate Change and Natural Disasters

- Qunkun Ma, Scott Powell, Jerry Sweitzer

Motivation

Our project aims to examine the relationship between climate and disasters in the United States. Many disasters, including the most destructive disasters such as hurricanes, are climate-driven events. As the temperature of the Earth increases, there should be a change in the frequency and intensity of disasters. Understanding the trends for weather-related disasters can inform hazard mitigation efforts. Additionally, understanding the changes in natural disasters can help with response and recovery planning prior to an incident.

Our group sought to answer the following questions with this project:

- Are natural disasters becoming more frequent due to climate change?
- What types of natural disasters are becoming more frequent?
- Is this change similar across different regions (see map below)?
- Is climate change affecting disaster frequency at certain times of year?

Figure 1 – Map of FEMA Regions – Retrieved from: *https://www.fema.gov/about/organization/regions*

Data Sources

1. Federal Emergency Management Agency (FEMA) Dataset

The primary data source four our project is the Federal Emergency Management Agency Disaster Declarations Summary dataset. This dataset contains a record of all Federal disaster declarations since January 1953. The dataset is updated weekly, and is downloaded as a csv file. As of the beginning of September 2024, there were 66,140 rows in the dataset and 28 features. The features we employed for our project were as follows:

- **Incident start date** - the date the actual incident (eg - flood, fire, etc.) occurred. This is distinct from the declaration date, as the declaration may occur months later.
- **Type of incident** - primary incident type for the declaration.
- **Location** - this is recorded by state, county, or tribal nation. We will employ state-level data for our analysis.
- **Declared incident types** - additional incident types for the disaster declaration. This is distinct from the type of incident, as incidents occur within incidents.

2. NCEI Weather Dataset Continental US

The secondary data sources for our project came from the National Oceanographic and Atmospheric Administration (NOAA) National Centers for Environmental Information (NCEI). These sources included a dataset for statewide time series temperature data and Palmer Drought Severity Index data. The statewide temperature data can be downloaded by state for each month in a selected range. Users can employ a drop-down menu to select the state, time, interval, and variable of interest. This will download a csv file of approximately 14kb. For drought series data, users can either employ the same drop-down menus or download the entire drought series data for all states as a txt file. For this project, our team employed the latter method, and searched the txt file. The files range in size from 72 kb to over 38M depending on the data chosen. More information regarding the download of these files will appear in the data manipulation section.

Links:

- https://www.fema.gov/openfema-data-page/disaster-declarations-summaries-v2
- https://www.ncei.noaa.gov/access/monitoring/climate-at-a-glance/statewide/time-series
- https://www.ncei.noaa.gov/pub/data/cirs/climdiv/

Data Manipulation – FEMA Dataset

Introduction

For the primary dataset, data manipulation begins by downloading the csv file. At the start of our project, the file contained 66,140 rows of data. Examination of this data reveals that each declared disaster can appear more than once. There are multiple reasons for this. In most cases, the declaration appears in the dataset once for each county in the declared state. Additionally, if the declaration proceeded through more than one step – from emergency declaration to disaster declaration – then it would appear for each step in that process. Further complicating matters, after enactment of the Sandy Recovery Improvement Act of 2013, designated Tribal Nations could also receive disaster declarations. Very few Tribal Nations have done so, but the records appear under the state where the tribal reservation is located (e.g. – the Seminole Nation appears as a Florida disaster declaration). This results in large disasters such as hurricanes appearing multiple times in our final dataset. Without conducting archival research and rebuilding the dataset ourselves, there is no way to reconcile this without losing information.

Create Dictionaries for Data Manipulation

The first step in data manipulation involved creating several lists and dictionaries that would come in handy later in the project. For instance, a state dictionary was created that contained state abbreviations and names as key-value pairs. A list of the 48 contiguous states was also created. The weather datasets did not contain data for Alaska, Hawaii, Washington, D.C., or US territories for the 1953-2023 timeframe. Therefore, the analysis was confined to the 48 contiguous states. The state list aided in filtering the data frame later in the process. A list of natural disasters was also created, since they would be employed as columns in the dataframe. The final, and most important dictionary was the disaster dictionary using FEMA codes as keys and disaster types as values. The codes can be found by visiting the disaster declarations summary page and scrolling down to the data description.

```
#dictionary to convert state/territory designators to full word strings
state_dict = {'AL':'Alabama','AK':'Alaska','AZ':'Arizona','AR':'Arkansas','CA':'California','CO':'Colorado','CT':'Connecticut',
              'DE':'Delaware', 'FL':'Florida','GA':'Georgia','HI':'Hawaii','ID':'Idaho','IL':'Illinois','IN':'Indiana','IA':'Iowa',
              'KS':'Kansas','KY':'Kentucky','LA':'Louisiana','ME':'Maine','MD':'Maryland','MA':'Massachusetts','MI':'Michigan',
              'MN':'Minnesota','MS':'Mississippi','MO':'Missouri','MT':'Montana','NE':'Nebraska','NV':'Nevada','NH':'New Hampshire',
              'NM':'New Mexico','NY':'New York','NJ':'New Jersey','NC':'North Carolina','ND':'North Dakota','OH':'Ohio',
              'OK':'Oklahoma','OR':'Oregon','PA':'Pennsylvania','RI':'Rhode Island','SC':'South Carolina','SD':'South Dakota',
               'TN':'Tennessee','TX':'Texas','UT':'Utah','VT':'Vermont','VA':'Virginia','WA':'Washington','WV':'West Virginia',
              'WI':'Wisconsin','WY':'Wyoming','DC':'Washington, DC','GU':'Guam','PR':'Puerto Rico','AS':'American Samoa',
              'MP':'Northern Mariana Islands','FM':'Federated States of Micronesia','MH':'Marshall Islands','PW':'Palau'}
```

Figure 2 – Dictionary with state names and abbreviations for value conversion in the dataframe.

Data Filtering

The next steps in manipulation for the primary dataset include filtering columns of interest, converting date time strings to individual columns, and filtering observations. The columns are then filled with a value of zero as they will be populated later. Next, we filtered for disasters that occurred in the 48 contiguous states from 1953-2023.

```
#dictionary to convert disaster codes to strings representing each type of disaster
disaster_dict = {'0':'Not applicable','1':'Explosion','2':'Straight-Line Winds','3':'Tidal Wave','4':'Tropical Storm',
                 '5':'Winter Storm','A':'Tsunami','B':'Biological','C':'Coastal Storm','D':'Drought','E':'Earthquake',
                  'F':'Flood','G':'Freezing','H':'Hurricane','I':'Terrorist','J':'Typhoon','K':'Dam/Levee Break','L':'Chemical',
                 'M':'Mud/Landslide','N':'Nuclear','O':'Severe Ice Storm','P':'Fishing Losses','Q':'Crop Losses','R':'Fire',
                 'S':'Snowstorm','T':'Tornado','U':'Civil Unrest', 'V':'Volcanic Eruption','W':'Severe Storm','X':'Toxic Substances',
                 'Y':'Human Cause','Z':'Other', '8':'Tropical Depression'}
```

Figure 3 – Dictionary with disaster codes. Retrieved from: https://www.fema.gov/openfema-data-page/disaster-declarations-summaries-v2

```
                #select columns necessary for data analysis, add empty columns for each natural disaster type
column_list = ['femaDeclarationString','state','incidentType','incidentBeginDate','fipsStateCode','region',
               'designatedIncidentTypes','declarationTitle']+natural_disaster
```

Figure 4 – Creating column list for the data frame

Data Cleaning

After creating a dataframe with the required columns and filtering observations, the next step involves populating each disaster column with data. Examination of the FEMA dataset shows that the column of 'designatedIncidentTypes' often has several values. This is because although there is only a single primary incident, many declarations have additional incident types added to them. Not all incidents are a result of the primary incident. For instance, the EF-5 tornado that impacted Joplin, MO was added on to another disaster declaration, so it would be in the 'designatedIncidentTypes' column and not the 'incidentType' column. Some observations do not have data in the 'designatedIncidentTypes', and frequently the 'incidentType' is not reflected either. The 'designatedIncidentTypes' column is a string variable that must be split, and the information converted into an integer and populated in its respective column.

Data Manipulation – FEMA Dataset

Data Cleaning Continued

The below block of code extracts incident types from their respective columns. Of note, ‘try’ and ‘except’ statements are crucial at this step. Since we are examining natural disasters, we do not have columns for incidents such as terrorist attacks. Several natural disasters have non-natural disasters in the ‘designatedIncidentTypes’ field. The ‘try’ and ‘except’ statements allow the program to skip these non-natural disasters.

```
#this block uses the incident codes in disaster_dict to populate the respective incident columns in disaster_df
for i in range(0,len(disaster_df)):
    incident = disaster_df.loc[i,'designatedIncidentTypes']
    incident_list = incident.split(',')

    #if there is only one incident type, add 1 to the corresponding column
    if ((len(incident_list) == 1) & (len(incident_list[0])>1)):
        col = incident_list[0]
        try:
            disaster_df.loc[i, col] += 1
        except:
            continue

    #if there are multiple incident types, add 1 to the corresponding column for each
    else:
        #create a new list, this combines the incident list from designatedIncidentTypes and incidentType
        new_list = []
        for inc in incident_list:
            #add longform name to new_list
            if len(inc) == 1:
                new_list.append(disaster_dict[inc])
        #add incidentType to new_list if not already in new_list
        if disaster_df.loc[i, 'incidentType'] not in new_list:
            new_list.append(disaster_df.loc[i, 'incidentType'])
        #ensures that code skips over non-natural disasters
        for val in new_list:
            try:
                disaster_df.loc[i, val] += 1
            except:
                continue
```

Figure 5 – Populating columns with a tally of incident by type

Final Step

Once this step is complete, we dropped duplicate disaster declaration numbers so that each disaster declaration is counted once per state, and not per county. The final steps involve dropping unnecessary columns and grouping disasters by year, month, and state.

year	month	state	fire	flood	severe	straight_line_winds	winter_storm	hurricane	tornado	tropical_storm
1953	5	Georgia	0	0	0	0	0	0	1	0
		Louisiana	0	1	0	0	0	0	0	0
		Texas	0	0	1	0	0	0	1	0
	6	Iowa	0	1	0	0	0	0	0	0
		Massachusetts	0	0	0	0	0	0	1	0

Figure 6 – Sample output of the FEMA dataframe grouped by year, month and state, listing each type of disaster as a column

```
Index(['Fire', 'Flood', 'Severe Storm', 'Straight-Line Winds', 'Winter Storm',
       'Hurricane', 'Tornado', 'Tropical Storm', 'Mud/Landslide', 'Snowstorm',
       'Coastal Storm', 'Severe Ice Storm', 'Typhoon', 'Freezing', 'Drought',
       'Fishing Losses', 'Other', 'Dam/Levee Break', 'Tropical Depression'],
      dtype='object')
```

Figure 7 – Output columns of the FEMA dataframe, showing the unique types of disasters in the dataframe

	disaster_count
Fire	278
Flood	52
Severe Storm	43
Straight-Line Winds	1
Winter Storm	4

Figure 7 – Sample output of total disasters occurrences CONUS between 1953 - 2023

Having the data frame in the format above will help with joining the FEMA disaster dataset with the weather dataset from NCEI. Additionally, it will allow further transformations to filter by disaster type or displaying disaster occurrences at national level.

Data Manipulation – NCEI Weather Dataset

Introduction

The NCEI Weather Dataset has weather data of states in the US. For this project, the time range of interest is 1953 -2023. The weather data includes features of:

- Average Temperature (avg_temp)
- Minimum Temperature(min_temp)
- Maximum Temperature(max_temp)
- Palmer Hydrological Drought Index -PHDI (phdi) [Measurement of Droughts]
- Palmer Modified Drought Index - PDMI (pmdi) [Measurement of Droughts]
- PDSI (pdsi) Standardized Index Measuring Dryness or Wetness
- Precipitation (pcp) Measurement of rainfall

Data is accessible on the National Centers for Environmental Information (NCEI) website. However, due to Hurricane Helene, its data center has been impacted. Thus, a backup copy has been uploaded to Github.

Github Link: https://github.com/xquinnma/milestone

Data is read in with pd.read_csv with the link to the raw file. Each feature listed above is stored in a separate file.

Example of the raw data file is shown below. The index (first column) is formatted as follows: first three digits indicates the state code, 001 - Alabama, and the last four digits indicates the year of the record, for example, the first row is precipitation data in the state Alabama in year 1985, with the next twelve columns indicating each month of the year.

- climdiv-pcpnst
- climdiv-pdsist
- climdiv-phdist
- climdiv-pmdist
- climdiv-tmaxst
- climdiv-tminst
- climdiv-tmpcst

0010011895	7.52	2.66	7.62	3.58	3.78	5.79	4.51	5.13	1.60	2.22	1.76	4.23
0010011896	4.33	6.16	5.44	3.54	3.24	5.44	4.99	2.82	1.80	2.27	4.42	1.71
0010011897	3.73	6.32	10.66	4.10	1.60	1.97	4.72	6.07	0.75	1.26	1.80	5.38
0010011898	3.92	2.11	3.12	4.49	0.95	3.86	6.05	7.20	3.46	3.84	5.99	3.85
0010011899	6.18	6.40	6.31	2.45	2.03	2.91	6.73	3.86	0.53	2.64	2.97	5.38
0010011900	3.37	8.75	5.96	8.69	2.56	10.38	4.74	3.12	4.06	5.79	3.33	5.10
0010011901	4.97	4.45	6.34	5.19	5.30	3.25	3.49	9.12	4.49	1.22	1.78	7.49
0010011902	3.70	6.93	9.37	2.36	2.52	1.35	2.87	3.77	4.22	3.63	4.11	5.51
0010011903	3.62	10.61	5.62	2.73	6.25	4.75	4.39	4.57	1.58	1.95	2.12	3.23
0010011904	4.24	4.13	3.78	2.32	3.05	3.22	5.01	5.81	1.23	0.26	3.06	4.49

Fig 8. List of files used to build to weather dataframe

Fig 9. Example of the raw file of one of the weather files listed on the left

Creating DataFrame

```python
def format_df(input_url, feature_name, start_year=1953, end_year=2023):
    df = pd.read_csv(input_url, header=None, names=data_column_names, delimiter=r"\s+", dtype= str)
    df = df.set_index('id').stack().reset_index().rename(columns={'level_1':'month', 0:feature_name})

    # create columns for state code and year from id column
    # convert year and month to datetime data
    df['state_code'] = df['id'].astype(str).str[:3].astype(int)
    df['year'] = pd.to_datetime(df['id'].astype(str).str[-4:]).dt.year
    df['month'] = pd.to_datetime(df['month'], format='%m').dt.month

    # limit state code < 48 is continental US
    df = df[df['state_code'] < 49]
    # limit data between year 1953 and 2023
    df = df[(df['year'] >= start_year) & (df['year'] <= end_year)]

    # convert state code to state name
    df['state'] = df['state_code'].astype(str).map(state_code_dict)

    # dropping unnecessary coloumns and shuffle column order and reset index
    df = df.drop(['id', 'state_code'], axis=1)

    # convert column to floats
    df = df.astype({feature_name : 'float'})

    # reorder columns
    df = df[['year', 'month', 'state', feature_name]].reset_index(drop=True)

    #print(df.dtypes)

    return df.copy()
```

Fig 10. Screenshot of the function created to process raw data and convert into dataframes of each weather feature

- After the raw frame is read in as dataframes, a function is created to clean each dataframe. First it will add column names of month 1-12 as the raw file did not contain any header.
- Then, the function extracts state code and year information from the index field creating new columns for state and year.
- The data contains information exceeding the scope of the project. Limiting state code <49 and setting year between 1953 and 2023 restricts the scope of the dataframe.
- Then, replace state code with full state name using a previous dictionary created using the .map() function This creates a new column.
- Lastly, dropping unnecessary columns to keep the dataframe concise.

Data Manipulation – NCEI Weather Dataset

Combine Weather Datasets

	year	month	state	precipitation
0	1953	1	Alabama	5.47
1	1953	2	Alabama	7.48
2	1953	3	Alabama	3.70
3	1953	4	Alabama	7.72
4	1953	5	Alabama	4.49

After running each dataset through the data cleaning function, there would be 7 dataframes. The shape for each dataframe is similar as below. The .stack() function transforms the dataframe from wide format to long format which is easier for joining.

Fig 11. Sample of the individual weather feature data frame

```
# combine all data
df_all_list = [df_precipitation, df_max_temp, df_min_temp, df_avg_temp, df_phdi, df_phmi, df_pdsi]

df_combined = df_all_list[0].copy()

for i in range(len(df_all_list)):
    if i!=0:
        df_combined = pd.merge(df_combined, df_all_list[i], on=['year', 'month', 'state'])


df_combined_grouped = df_combined.groupby(['year', 'month', 'state']).sum()
```

Fig 12. Code that combines 7 individual weather feature data frames into one

The 7 data frames are then combined together into one data frame containing all weather features grouped by year, state and month. This would make it easier to combine weather and FEMA disaster dataframes in the next step.

			precipitation	max_temp	min_temp	avg_temp	pdhi	pdmi	pdsi
year	month	state							
1953	1	Alabama	5.47	60.6	37.3	49.0	-2.12	-1.83	-2.12
		Arizona	0.25	59.5	31.0	45.2	1.51	0.33	-0.66
		Arkansas	3.81	55.3	33.4	44.4	-2.14	-1.58	0.25
		California	4.52	56.4	37.6	47.0	1.96	1.96	1.96
		Colorado	0.89	43.9	18.9	31.4	-2.04	-2.04	-2.04

Fig 13. Sample output of the unified weather data frame with the 7 features as columns

Join Weather and Disaster Datasets Into One

```
#COMBINE INTO ONE DATASET

weather_df = df_combined_grouped.copy()
disaster = disaster_group.copy()

full_df = (disaster.join(weather_df, how = 'outer'))
full_df.fillna(0, inplace = True)
```

After data cleaning and manipulation of both the FEMA disaster dataset and NCEI weather dataset, the two datasets are combined into one unified data frame by year, state, and month.

Fig 14. Code to join FEMA disaster data frame and NCIE weather data frame together.

Joining the FEMA disaster dataframe and the NCEI dataframe yields a dataset with year, month and state indicating location and columns indicating disaster types and climate features. For null data, zero will be filled in place with the .fillna() function because null in this case indicates no such disaster took place.

```
Index(['year', 'month', 'state', 'Fire', 'Flood', 'Severe Storm',
       'Straight-Line Winds', 'Winter Storm', 'Hurricane', 'Tornado',
       'Tropical Storm', 'Mud/Landslide', 'Snowstorm', 'Coastal Storm',
       'Severe Ice Storm', 'Typhoon', 'Freezing', 'Drought', 'Fishing Losses',
       'Other', 'Dam/Levee Break', 'Tropical Depression', 'precipitation',
       'max_temp', 'min_temp', 'avg_temp', 'pdhi', 'pdmi', 'pdsi', 'region'],
      dtype='object')
```

Fig 15. List of columns of the joined data frame

	year	month	state	Fire	Flood	Severe Storm	Straight-Line Winds	Winter Storm	Hurricane	Tornado	...	Dam/Levee Break
0	1953	1	Alabama	0.0	0.0	0.0	0.0	0.0	0.0	0.0	...	0.0
1	1953	1	Arizona	0.0	0.0	0.0	0.0	0.0	0.0	0.0	...	0.0
2	1953	1	Arkansas	0.0	0.0	0.0	0.0	0.0	0.0	0.0	...	0.0
3	1953	1	California	0.0	0.0	0.0	0.0	0.0	0.0	0.0	...	0.0
4	1953	1	Colorado	0.0	0.0	0.0	0.0	0.0	0.0	0.0	...	0.0

Fig 16. Sample output for the joined data frame

Data Analysis and Visualizations

Frequency of Disasters

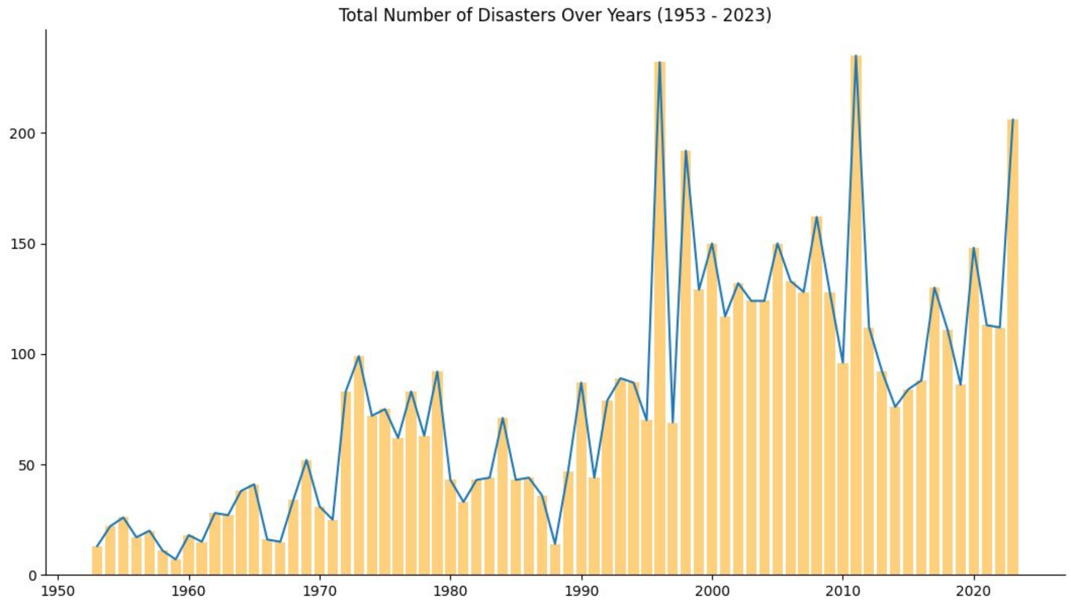


Fig 17. Visualization of disaster occurrences over years between 1953 - 2023

To gain a general understanding of the disasters, visualizations are created at the national level for the 48 contiguous states. Disaster visualizations at state level are available in the notebook.

The first question of the project is - Are natural disasters becoming more frequent due to climate change? The above visualization combines bar and line chart to show the occurrences of disasters over 70 years. It does appear disasters are appearing more frequently between 1990 to 2023 in comparison to 1953 - 1989. However there is not a conclusive and consistent increase over years based on the dataset.

Frequency of Disasters Based on Disaster Type

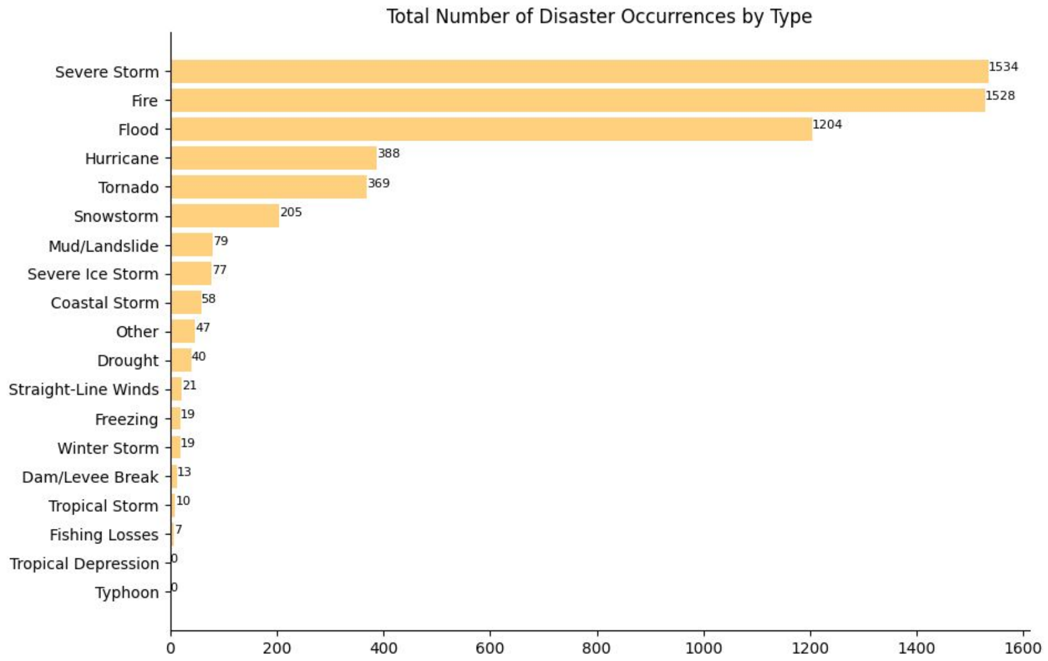


Fig 18. Visualization of disaster occurrences separated by type over years between 1953 - 2023

To answer the second question - “What types of natural disasters are becoming more frequent?” The visualization above provides an overview of the different types of disasters. It appears Severe Storm, Fire, and Flood are the top three most frequent types of disasters in continental US, with all three exceeding over 1,000 occurrences over the span of 70 years (1953-2023).

Severe Storm, Flood and Hurricane could have correlation to weather features such as precipitation. This will be examined further to seek relations. Conversely, Fire disasters may have to do with the lack of precipitation and the general rise of average temperatures due to climate change. This topic will another point of interest to see if there is any relationship between weather and disaster occurrences.

Data Analysis and Visualizations

Types of Disasters by Region

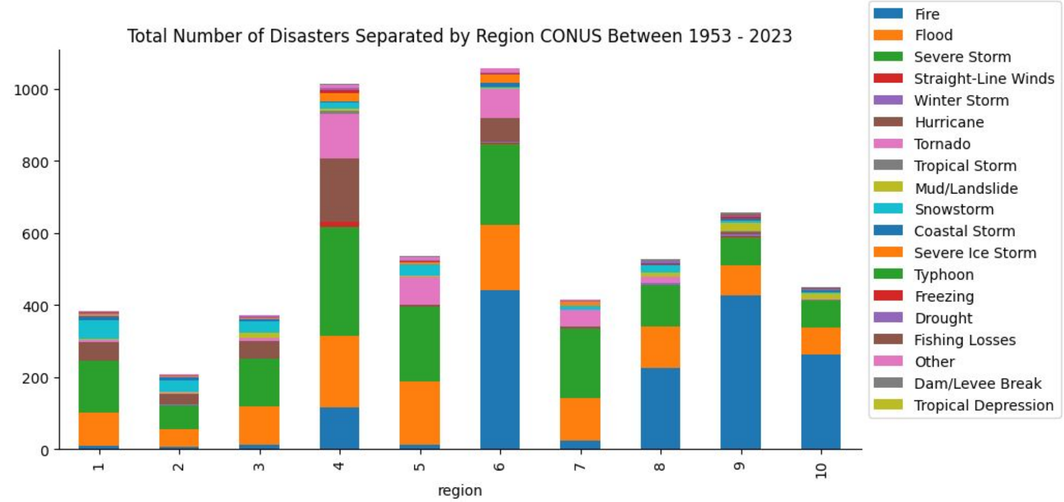
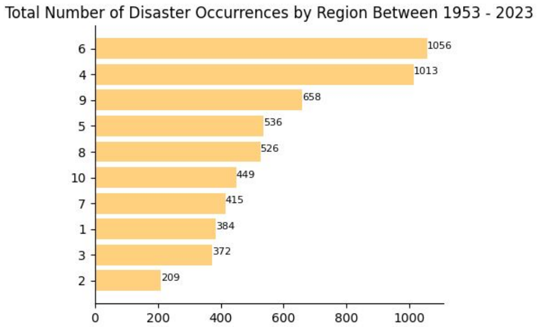


Fig 19. Stacked Bar Charts of Disaster Types by Region



States in Region 4

- Alabama
- Florida
- Georgia
- Kentucky
- Mississippi
- North Carolina
- South Carolina
- Tennessee

States in Region 6

- Louisiana
- New Mexico
- Oklahoma
- Texas
- Arkansas

Fig 20. Bar Charts Ranking Disaster by Region

To answer the third question of this study - “Is this change similar across different regions”? We seek to understand if certain regions are more prone to disasters. The stacked bar chart above helps uncover the total number of disasters and the portion of each type of disaster.

It appears that Region 4 and 6 are most prone to disasters with Region 4 having many severe storms and Region 6 with a great number of fire type disasters.

Frequency of Disasters at State Level

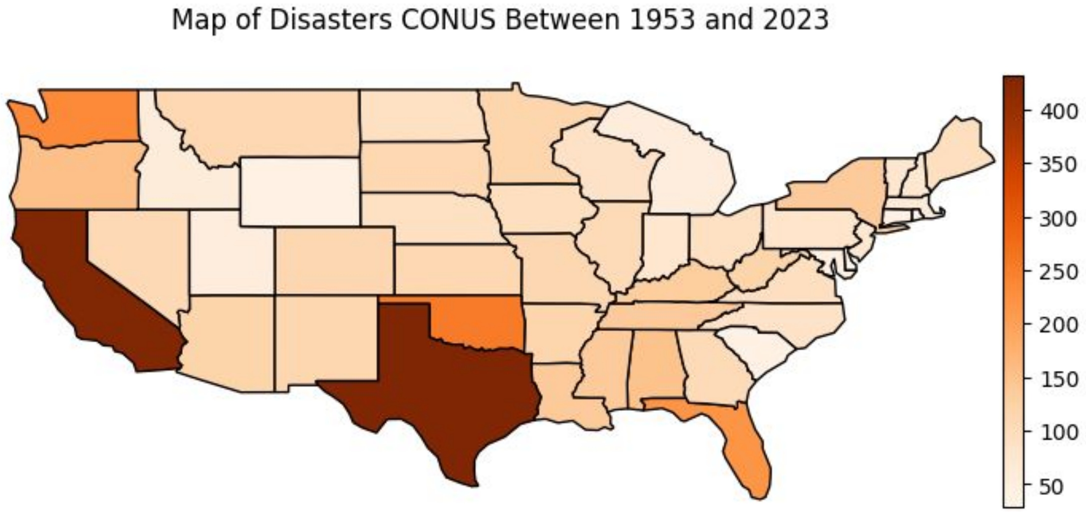


Fig 21. Choropleth Map of Disaster Occurrences at State Level.

After seeing how certain regions in the continental US are more prone to disasters, we can take a closer look at state level. A choropleth map of the US shows the states that have the greatest number of disasters.

In the visualization above, the states of California and Texa seems to have the greatest number of disasters between 1953 and 2023. States such as Washington, Oklahoma and Florida also had a high number of disasters. This map corroborates the visualization on the left with many disaster prone states in Regions 4 and 6.

It is interesting to see that all the top three states are attached to bodies of water. These disaster prone states will be the point of interest to discover relations between weather data and disaster occurrences.

Data Analysis and Visualizations

Types of Disasters by Month

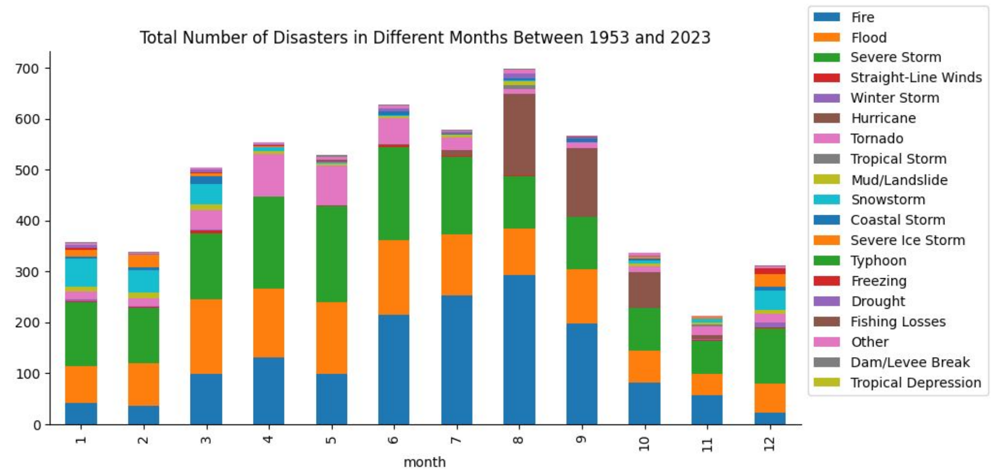


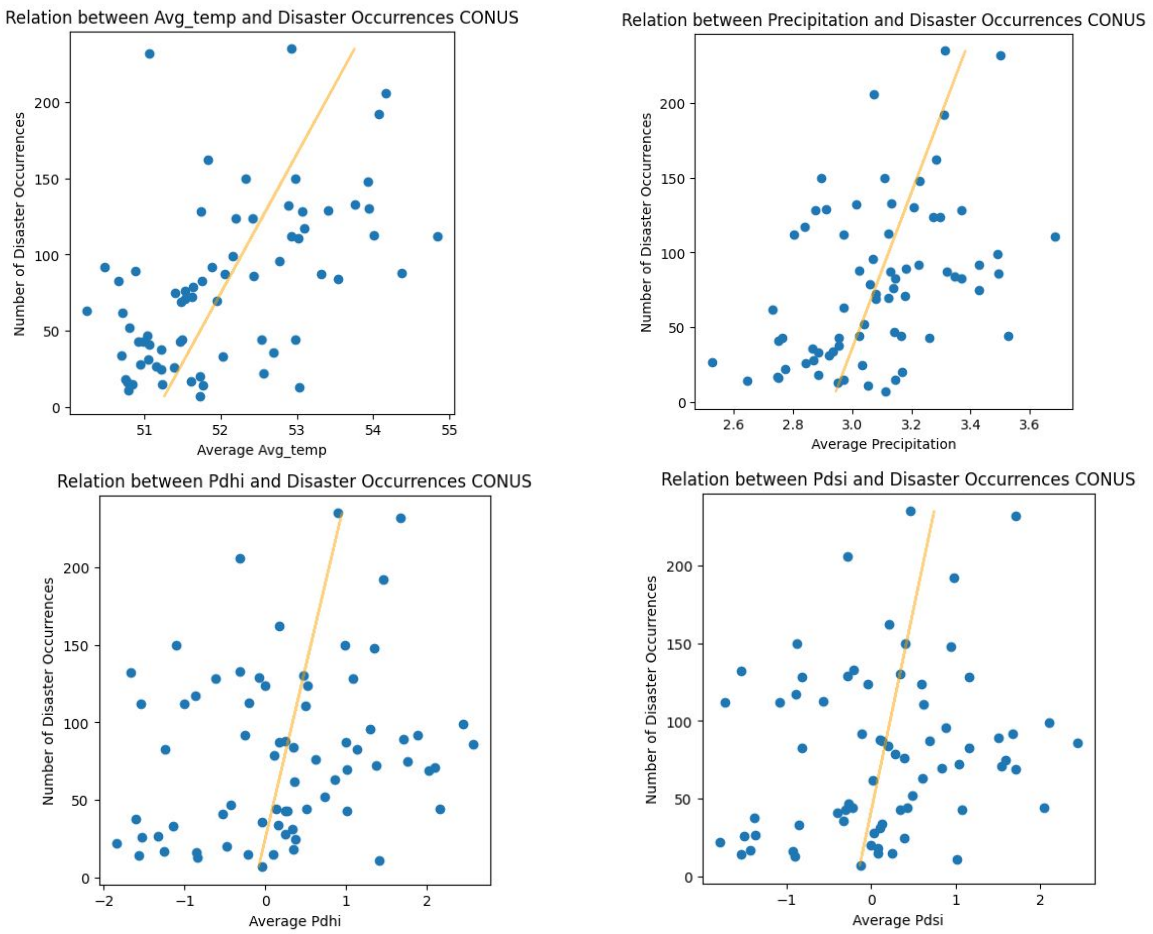
Fig 21. Stacked Bar Charts of Disaster Types by Month

To answer the last question - “Is climate change affecting disaster frequency at certain times of year?” It appears that disasters tend to happen in warmer months of the year around summer time, with the highest number of disasters occuring in August, with fire type disasters taking the highest portion of all disasters. Severe storms and floods hold consistent in Spring months between April and July and slightly drops in colder months.

Relation between Weather and Disaster Data

To discover more insights of how weather could have an effect on disasters within the US, we created a scatter plot of total disaster occurrences versus 4 weather features - average temperature, precipitation and pdsi (which is a good standard metric as it measures temperature, rainfall and drought). A best fit line (yellow) will also be plotted to see if there is any linear relation that could contribute to the more frequent occurrence of disasters.

Fig 22. Scatter plots and best fit lines of weather features against disaster occurrences

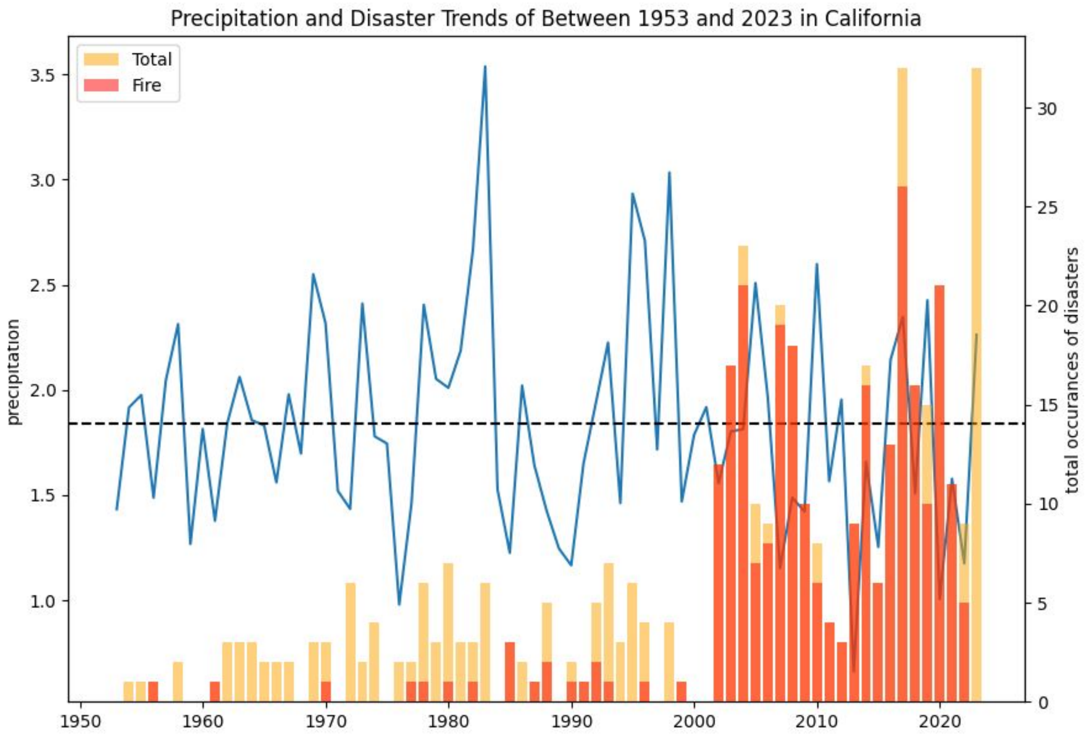


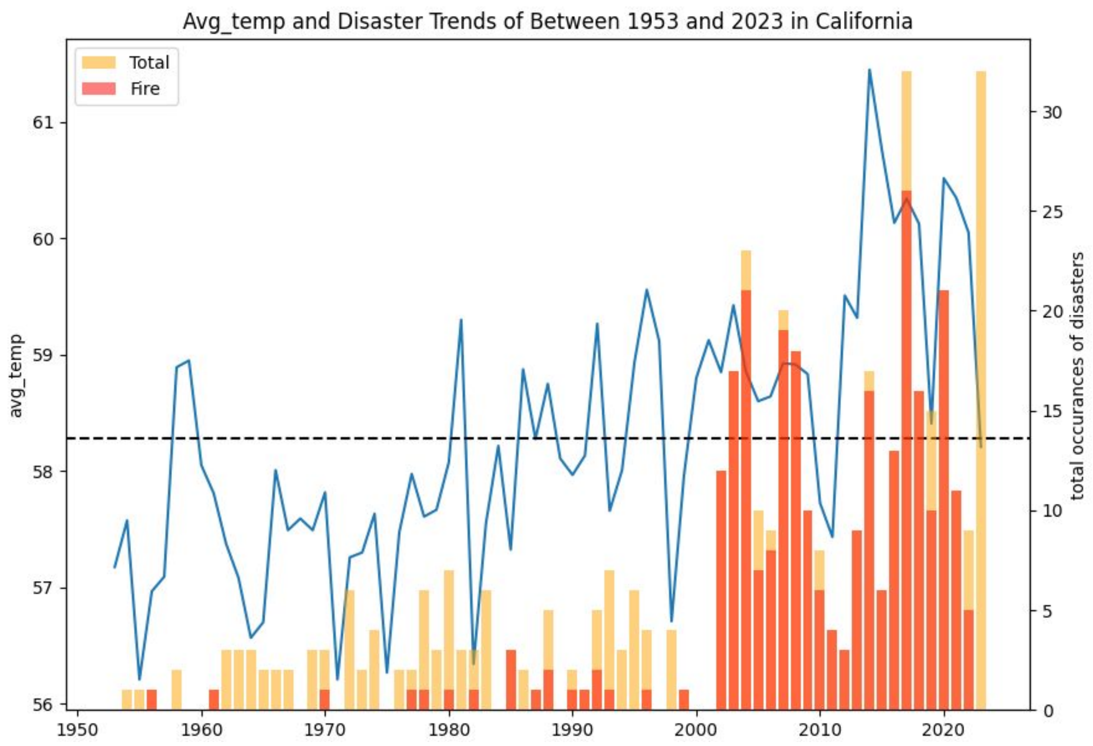
After plotting the data, the scatter plot does not seem to show clear patterns between the weather features versus disaster occurrences. The best fit line only goes through a few data points and does not represent the majority of the data. It is difficult to draw conclusive results from the dataset based on the visualization as it doesn’t show a clear relation.

Data Analysis and Visualizations

Weather and Disaster in Specific States

Fig 23. Combined charts showing weather feature versus disaster occurrences





California was one of the states that had the greatest number of disasters over the span of 1953 and 2023. The combined visualizations above focuses specifically on the state of California to see weather patterns and disaster patterns. The blue line graph is the average precipitation data and average temperature data. The black dotted horizontal line is the state average of the according weather feature over 70 years. Yellow bars represent total number of disasters in California and red bars represent the specific type of disaster which is Fire in this visualization. Above visualizations are a part of an interactive visualization available in the notebook.

As average temperature increases over the years, there seems to be a positive correlation between the number of fire type disasters with the rising average temperature. Also, the amount of rainfall seems to have decreased over the years. The number of fires also increased as precipitation declined and temperatures rose. Although the sample size may be too small to reach firm conclusions, there seems to be some positive correlation between average temperature and rainfall.

Analysis and Conclusion

Data Exploration and Analysis

Exploration of our weather visualizations shows that while average temperatures have risen by approximately two degrees Fahrenheit over the course of the study, that drought severity has not risen. Precipitation has also increased slightly during this timeframe. How does this impact disasters? Nationally, there is no correlation between total disasters and any of the weather variables analyzed. Below is the result of a correlation completed using the Pearson correlation coefficient in Pandas:

```
Correlation between total disasters and avg_temp: -0.005699554656378953
Correlation between total disasters and max_temp: 0.002451359847575512
Correlation between total disasters and min_temp: -0.01441302438668379
Correlation between total disasters and precipitation: -0.0596216422721528
Correlation between total disasters and pdsi: -0.009313633019844334
Correlation between total disasters and phdi: -0.0034220147548141094
Correlation between total disasters and pmdi: -0.00412928147158699
```

Fig 24. Correlation between weather data and disaster occurrences

Although there is no national trend, there certainly is a trend at the region and state level, particularly in the western United States. As shown in our interactive charts, states such as California, Washington and Oregon have seen an increase in disaster declarations in the past 20 years. This is due primarily to an increase in wildfires. There is some evidence to support the hypothesis that wildfires are becoming more frequent since higher temperatures may mean more drought in those states. Large disasters such as hurricanes do not appear to be more frequent, and this is confirmed by reviewing data from NOAA and other researchers (Richie, 2024). One word of caution, due to the peculiarities of the FEMA dataset, the number of hurricanes in our dataset is artificially high. This is because each hurricane is represented once per impacted state. Therefore, readers curious about hurricanes would be better informed by reviewing NOAA data and not by referencing FEMA disaster declarations (see references for list of external sources).

With the exception of the west coast (Regions 9 and 10), most regions are not seeing a noticeable increase in disasters. For individual states, even in hurricane prone areas, the number of declarations in a given year is less than 10, with 5 being a typical number. In recent years, it appears that only California exceeds 10 disasters in most years.

Conclusion

From our analysis, we have shown that there is no demonstrable correlation between total disasters and selected climate factors. One limitation of our study is the use of the FEMA database. This database only includes disasters for which there is a major disaster declaration. Many more disasters occur that never cause enough damage to receive a declaration. For instance, tornadoes less that EF-3 in strength are highly unlikely to receive a declaration, and many EF-3 tornadoes likely do not unless they impact a community directly. Although the Storm Prediction Center maintains detailed databases on all recorded weather incidents, the size of the database was too large for our project. Furthermore, the ability to detect and record natural disasters has dramatically improved during the period of our analysis. As noted by other researchers, the number of disasters (particularly small disasters) has increased over the last century. However, they also note that this is a function of how disasters are reported and is unconnected to climate change (Richie, 2024). As noted by author Michael Shellenberger in his book Apocalypse Never: “There is scant evidence to indicate that hurricanes, floods, tornadoes or drought have become more frequent or intense in the U.S. or globally. In fact, we are in an era of good fortune when it comes to extreme weather.”

Statement of Work and Sources

Statement of Work

Scott Powell

- Downloaded and cleaned the FEMA disaster summary dataset.
- Conducted correlation analysis in pandas.
- Provided write-up for cleaning FEMA data, analysis, and conclusion.
- Organized group work throughout the project.

Jerry Sweitzer

- Assisted with cleaning and manipulation of the NCEI weather dataset.
- Created choropleth of PDSI yearly average.

Qunkun Ma

- Data cleaning and exploration on the NCEI weather dataset.
- Created visualizations of interactive line charts, bar charts, scatter plots, and choropleth maps with the dataset using Matplotlib.
- Refactored code in Jupyter Notebook to ensure efficiency and clarity.
- Added content regarding data cleaning processes of the NCEI weather dataset and visualizations. Revised design of the report slides.

Sources

- Hannah Ritchie and Pablo Rosado (2024) - “Is the number of natural disasters increasing?” Published online at OurWorldinData.org. Retrieved from: 'https://ourworldindata.org/disaster-database-limitations' [Online Resource]
- *Hurricane Landfalls in the United States*. Our World in Data. Retrieved (2024,September 4) from https://ourworldindata.org/grapher/hurricane-landfalls-us.
- *OpenFEMA Dataset: Disaster Declarations Summary*. FEMA.gov. Retrieved (2024, September 16). https://www.fema.gov/openfema-data-page/disaster-declarations-summaries-v2.
- Shellenberger, M. (2020). *Apocalypse never: Why Environmental Alarmism Hurts us all*. Harper.
- *Statewide Time Series*. NOAA.gov. Retrieved (2024, September 16). https://www.ncei.noaa.gov/pub/data/cirs/climdiv/