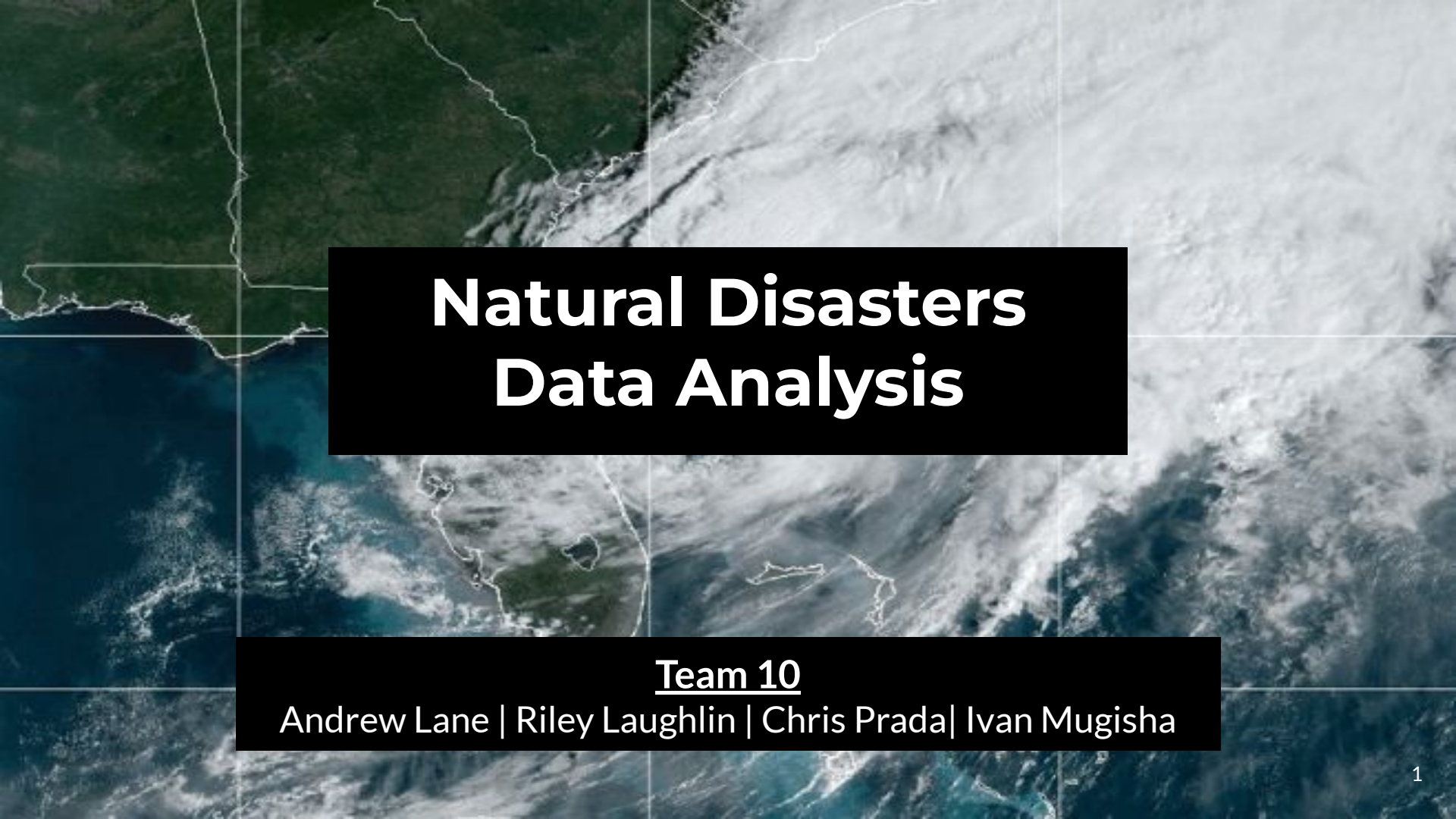


Natural Disasters Data Analysis

Team 10
Andrew Lane | Riley Laughlin | Chris Prada| Ivan Mugisha

Dataset Descriptions

1. The dataset encompasses 39,953 locations for 9,924 disasters that occurred worldwide in the years 1960 to 2018. Disaster types include floods, storms (typhoons, monsoons etc.), earthquakes, landslides, droughts, volcanic activity and extreme temperatures.

2. The dataset contains U.S. disaster cost assessments of the total, direct losses ($ in millions) inflicted by: tropical cyclones, inland floods, drought & heat waves, severe local storms (i.e., tornado, hail, straight-line wind damage), wildfires, crop freeze events and winter storms.

Key Questions

- Do natural disasters increase over time?
- Is there a relationship between disaster duration and economic impact?
- Which areas are most affected by natural disasters?
- What are the financial and human impacts of major disasters?

Evaluating Natural Disaster Frequency Trends

Hypothesis:

Natural disasters occur more frequently every year over time.

1. The Geocoded Disasters (GDIS) Dataset is a geocoded extension of a selection of natural disasters from the Centre for the Research on the Epidemiology of Disasters' (CRED) Emergency Events Database(EM-DAT).

2. Part of NASA's Earth Observing System Data and Information System (EOSDIS)

3. Data is collected at Columbia University' Climate School Center for Integrated Earth System Information (CIESIN)

1. NCEI maintains one of the most significant archives on Earth, with comprehensive oceanic, atmospheric, and geophysical data. They archive over 229 terabytes of data each month from over 130 observing platforms.

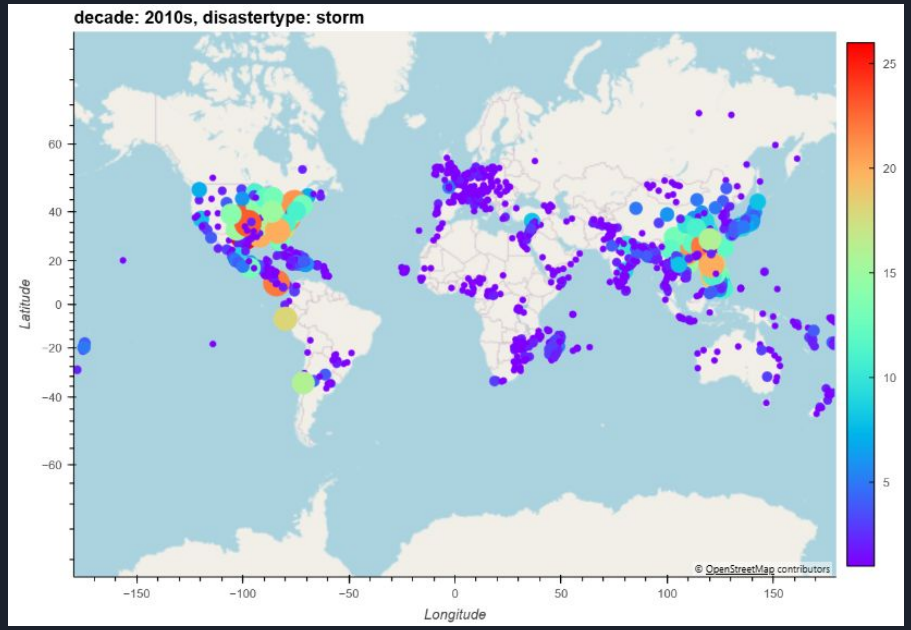
2. NCEI provides access to an extensive archive of environmental data through several platforms. They deliver the climate, coastal, oceanographic, and geophysical data you need in a variety of formats.

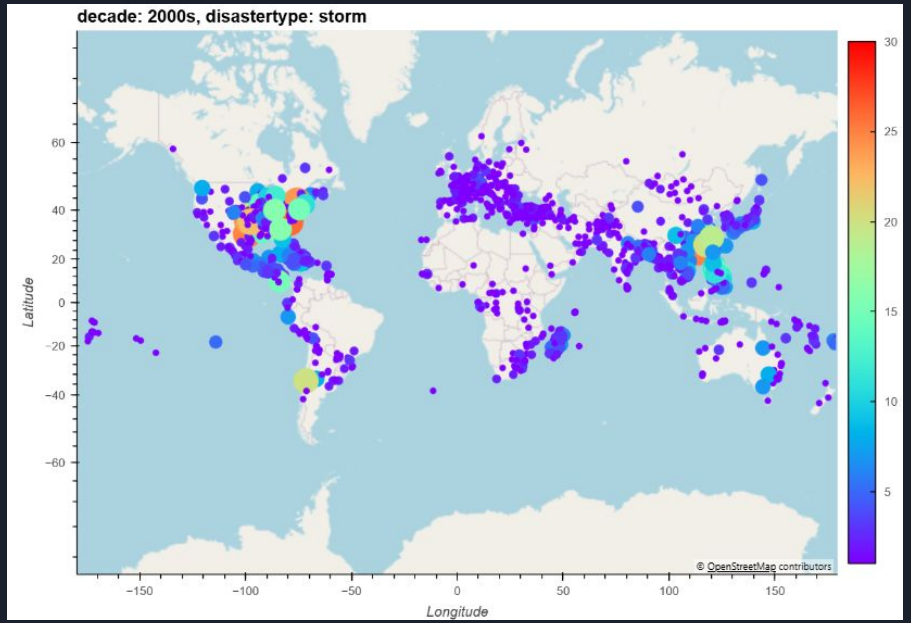
Dataset 1 - Data Cleaning and Processing

	id	country	iso3	gwno	year	geo_id	geolocation	level	adm1	adm2	adm3	location	historical	hist_country	disastertype	disasterno	latitude	long
0	109	Albania	ALB	339.0	2009	346	Ana E Malit	3	Shkoder	Shkodres	Ana E Malit	Ana E Malit	0	NaN	flood	2009-0631	42.020948	19.4
1	109	Albania	ALB	339.0	2009	351	Bushat	3	Shkoder	Shkodres	Bushat	Bushat	0	NaN	flood	2009-0631	41.959294	19.5
2	175	Angola	AGO	540.0	2001	760	Onjiva	3	Cunene	Cuanhama	Onjiva	Onjiva	0	NaN	flood	2001-0146	-17.093484	15.6
3	187	Angola	AGO	540.0	2009	710	Evale	3	Cunene	Cuanhama	Evale	Evale	0	NaN	flood	2009-0092	-16.531533	15.7
4	187	Angola	AGO	540.0	2009	749	Mupa	3	Cunene	Cuvelai	Mupa	Mupa	0	NaN	flood	2009-0092	-16.200065	15.8

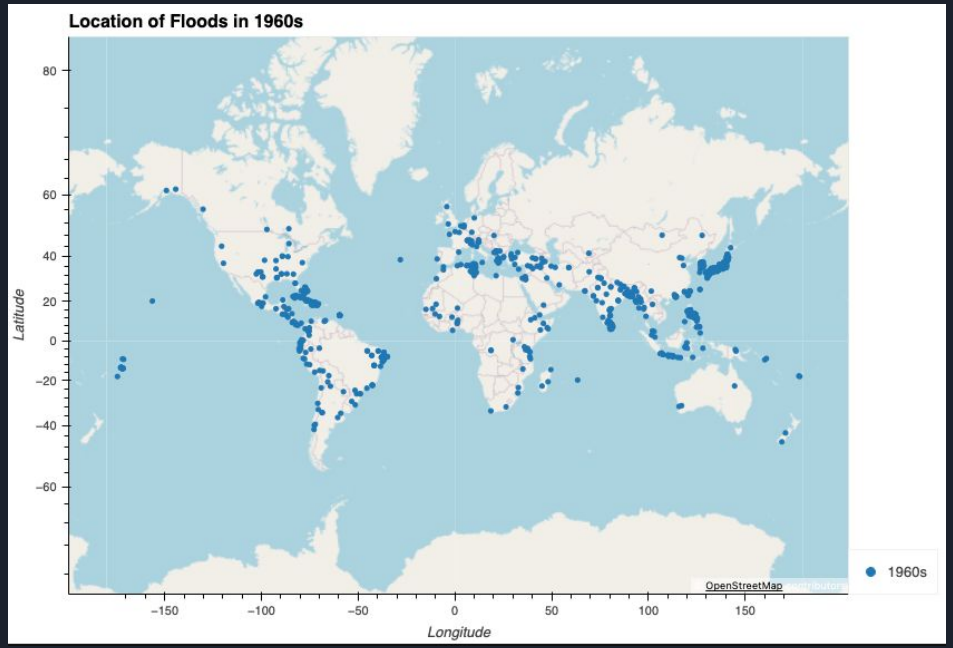
	id	country	year	level	location	disastertype	disasterno	latitude	longitude
0	109	Albania	2009	3	Ana E Malit	flood	2009-0631	42.020948	19.418317
1	109	Albania	2009	3	Bushat	flood	2009-0631	41.959294	19.514309
2	175	Angola	2001	3	Onjiva	flood	2001-0146	-17.093484	15.665758
3	187	Angola	2009	3	Evale	flood	2009-0092	-16.531533	15.773987
4	187	Angola	2009	3	Mupa	flood	2009-0092	-16.200065	15.844189

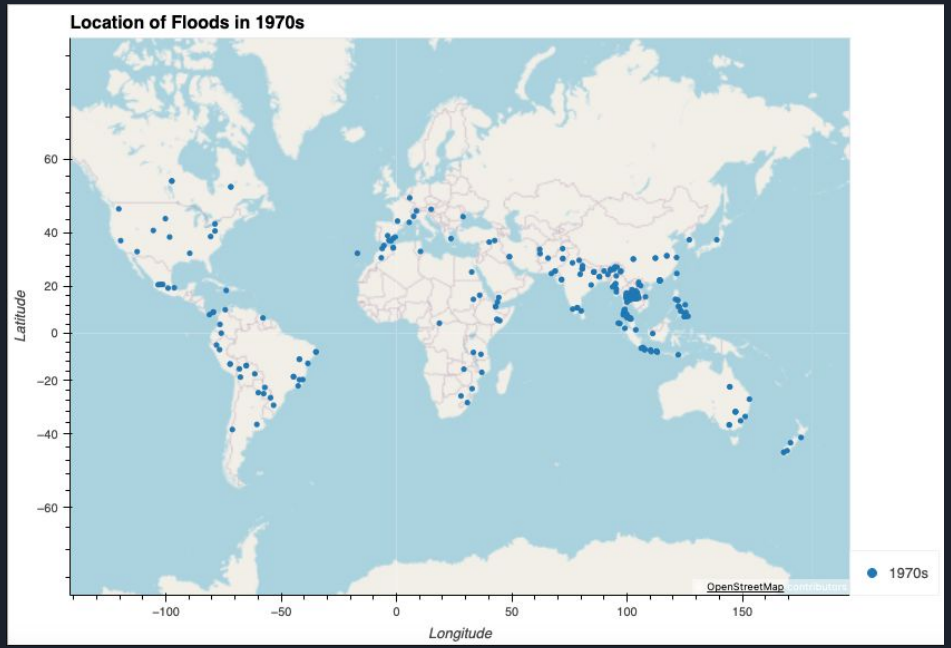
Visualization 1: Regions Affected by Storms (2010s and 2000s)



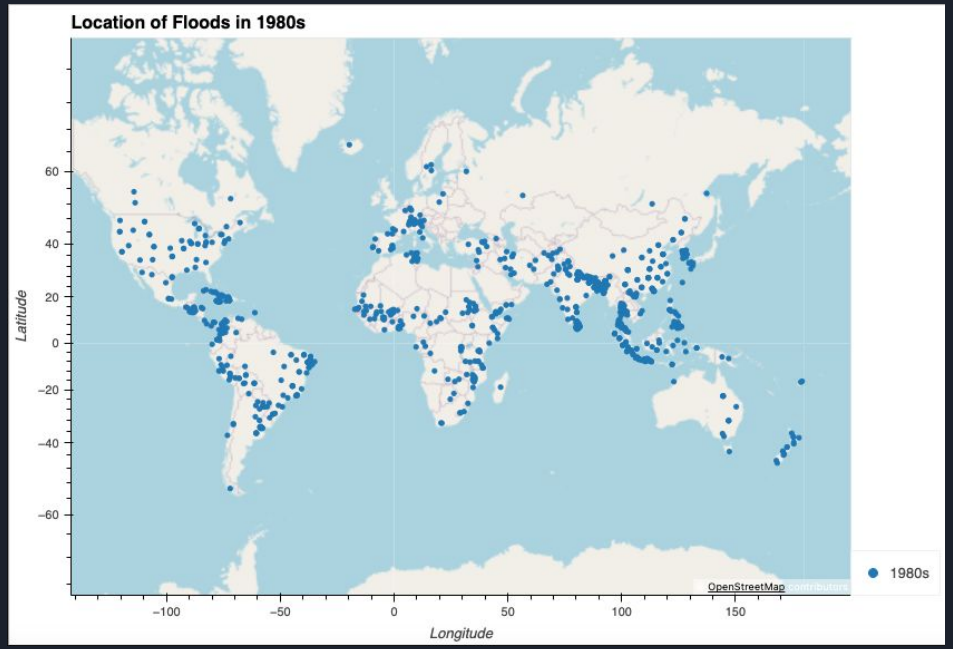


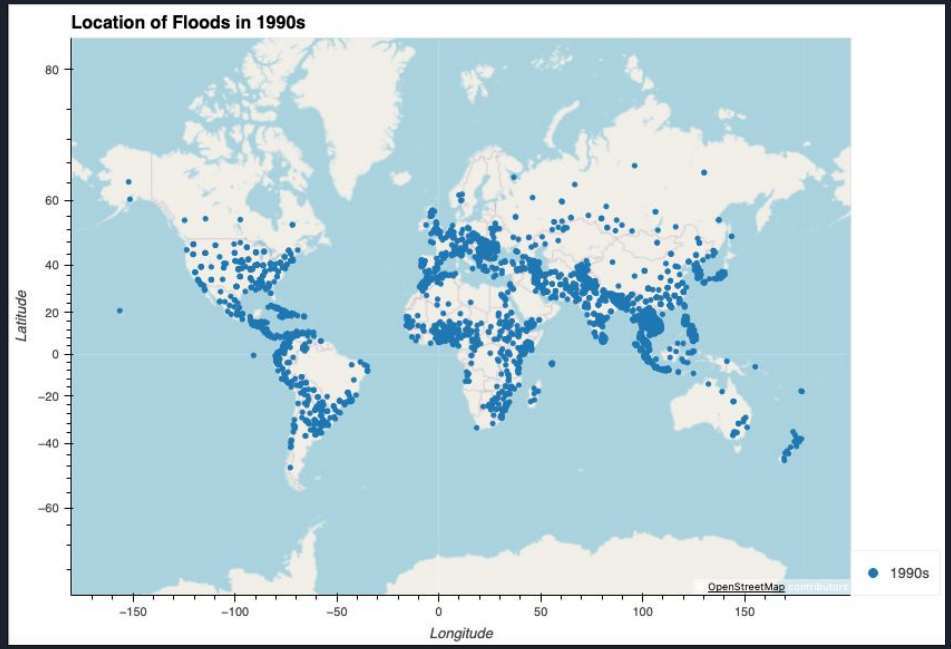
Visualization 2: Floods by Decades (1960s and 70s)



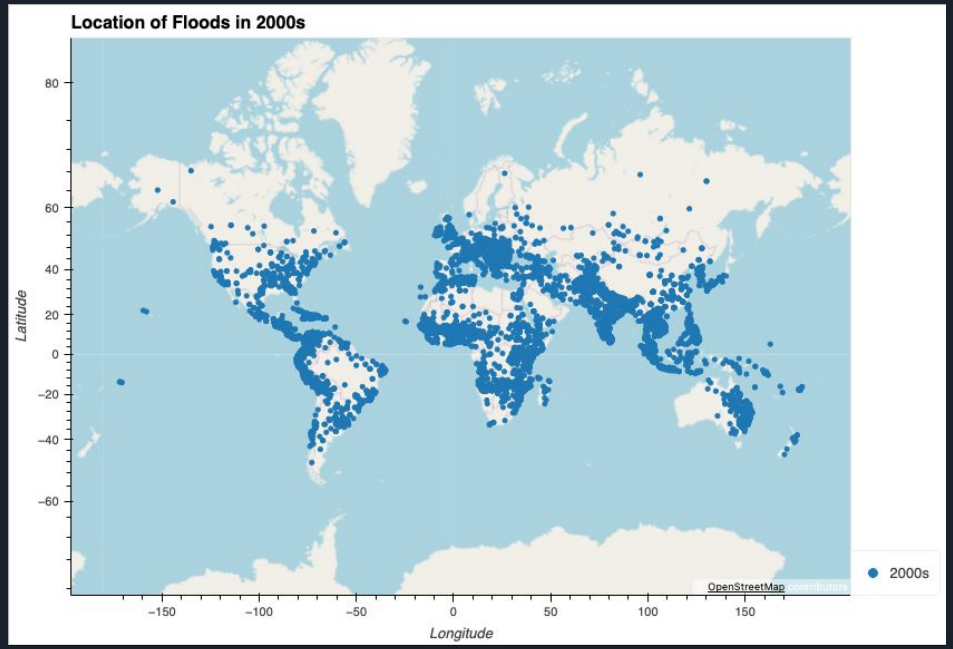


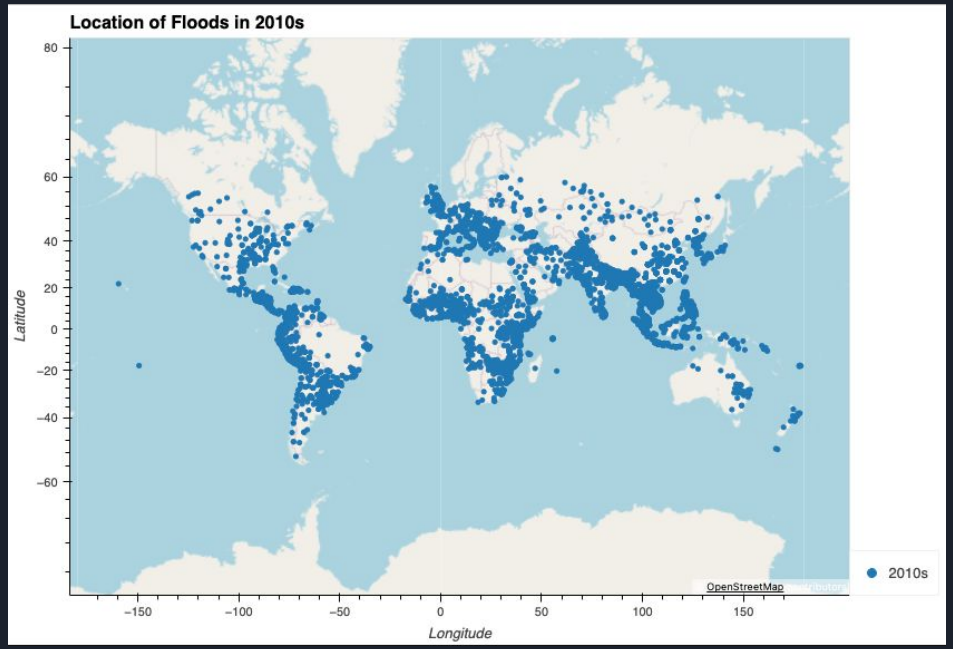
Visualization 2: Floods by Decades (1980s and 90s)



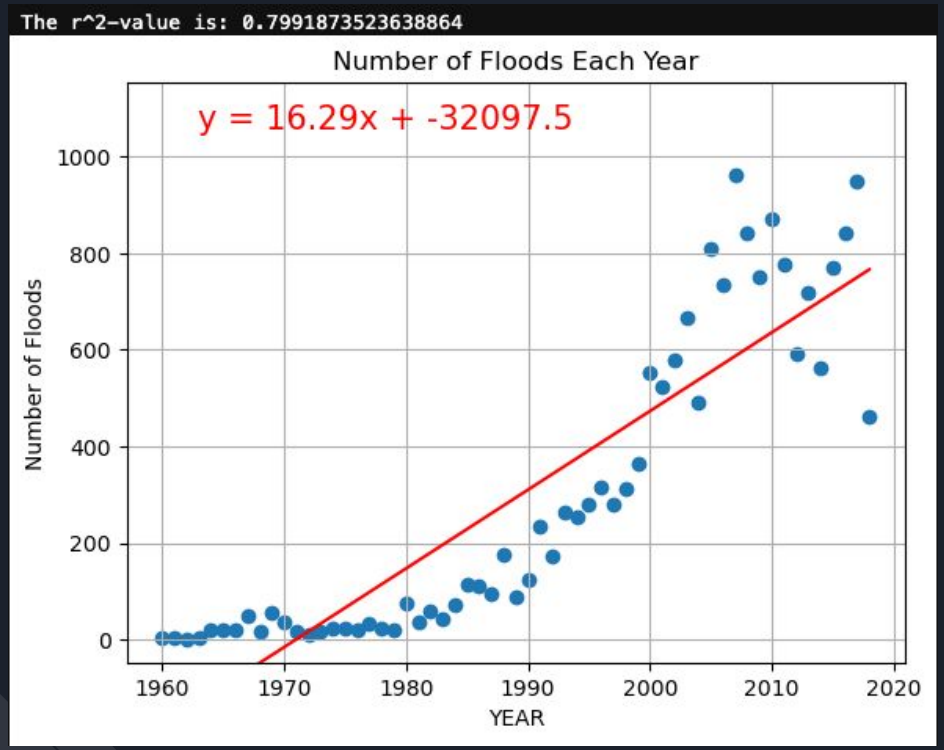


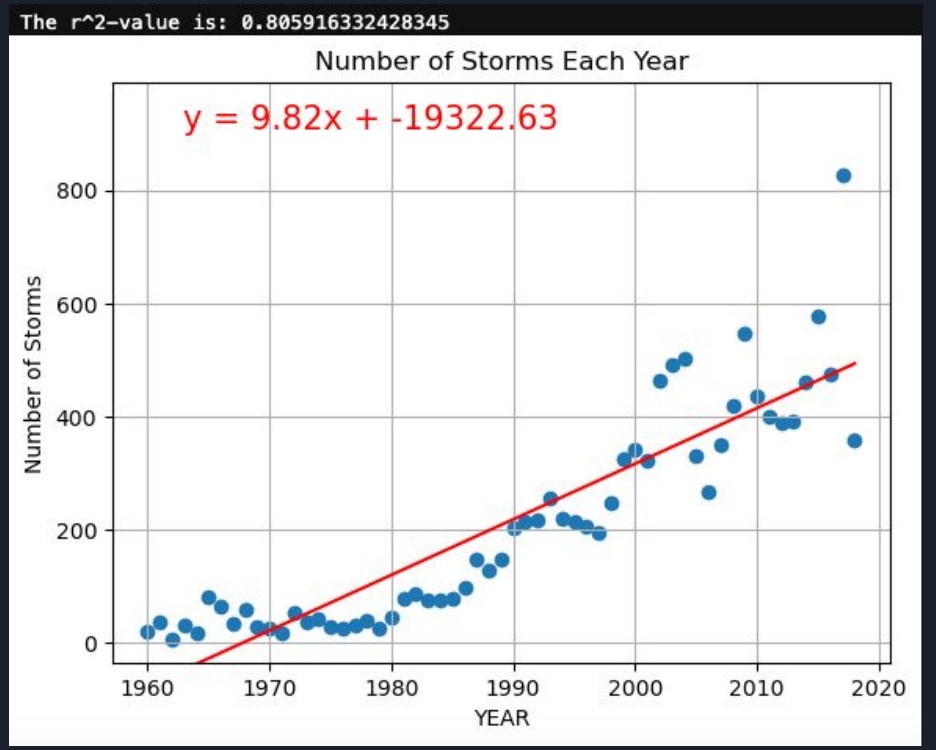
Visualization 2: Floods by Decades (2000s and 2010s)





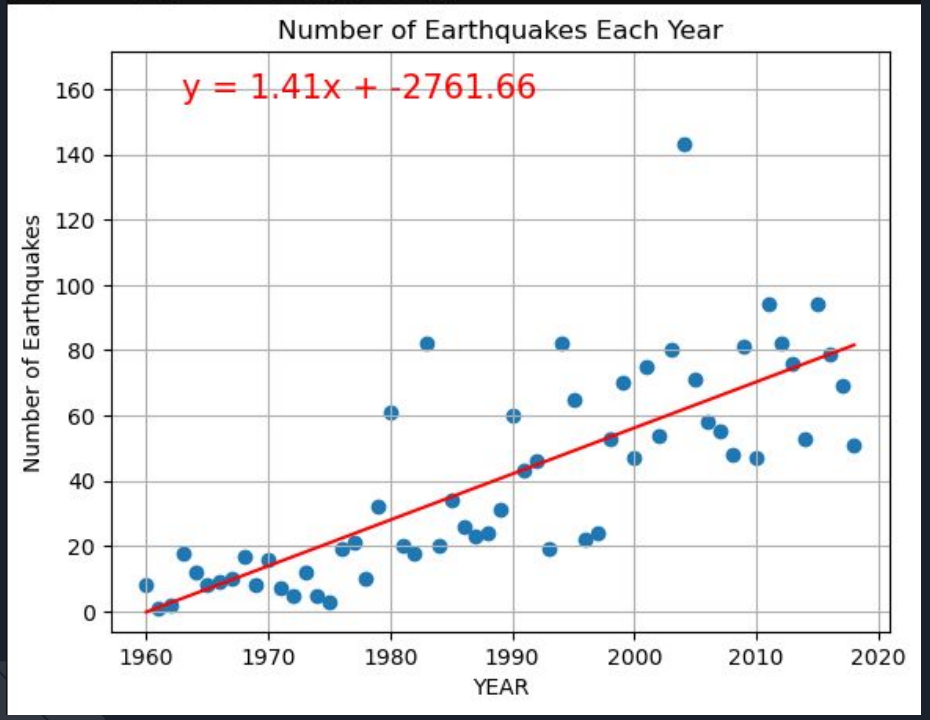
Visualization 3: Floods & Storms Regression Analysis



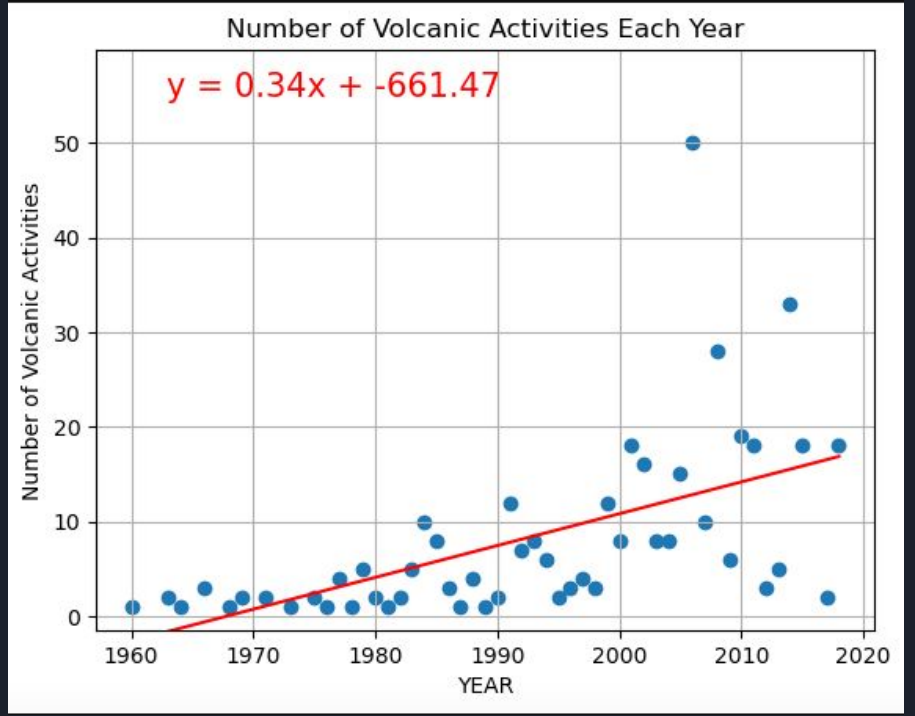


Visualization 4: Earthquakes & Volcanic Activity Regression Analysis

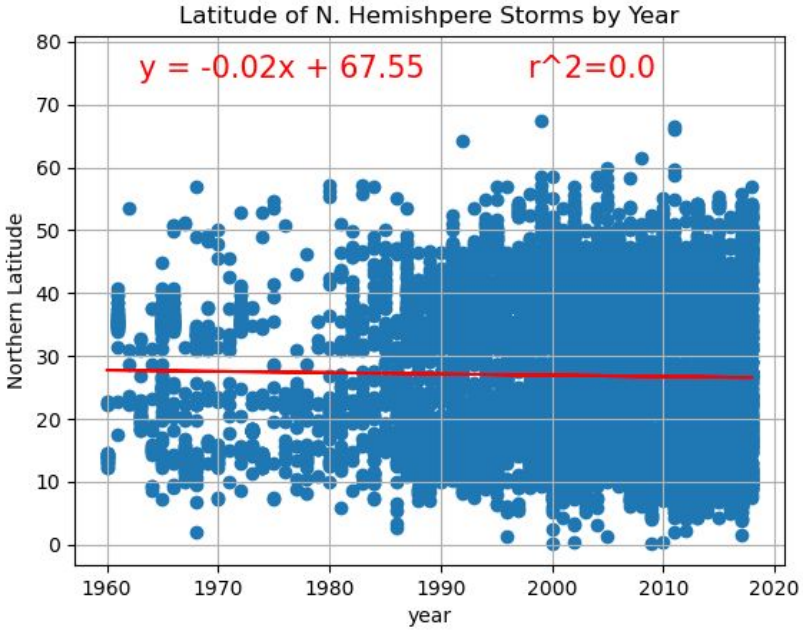
The r^2-value is: 0.6120364642004856

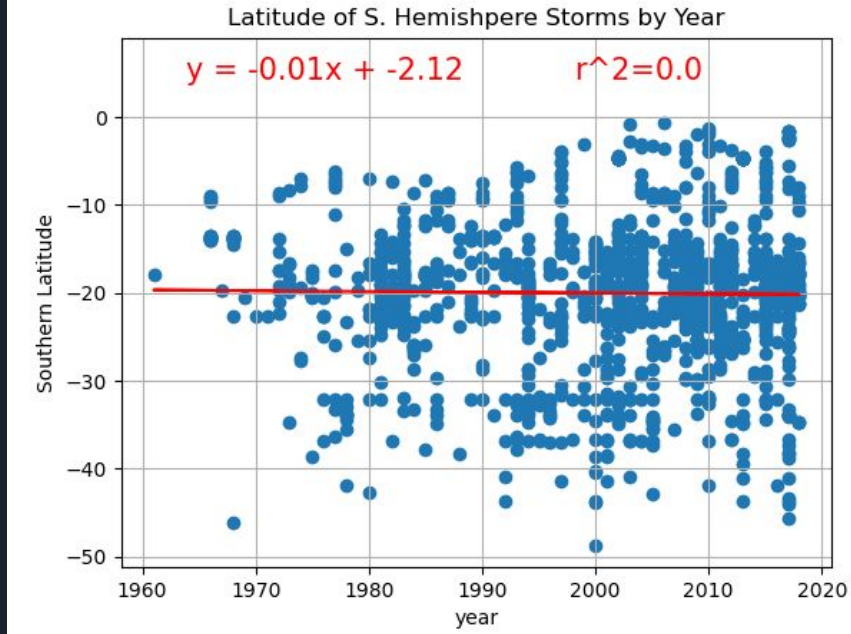


The r^2-value is: 0.32156106562606984

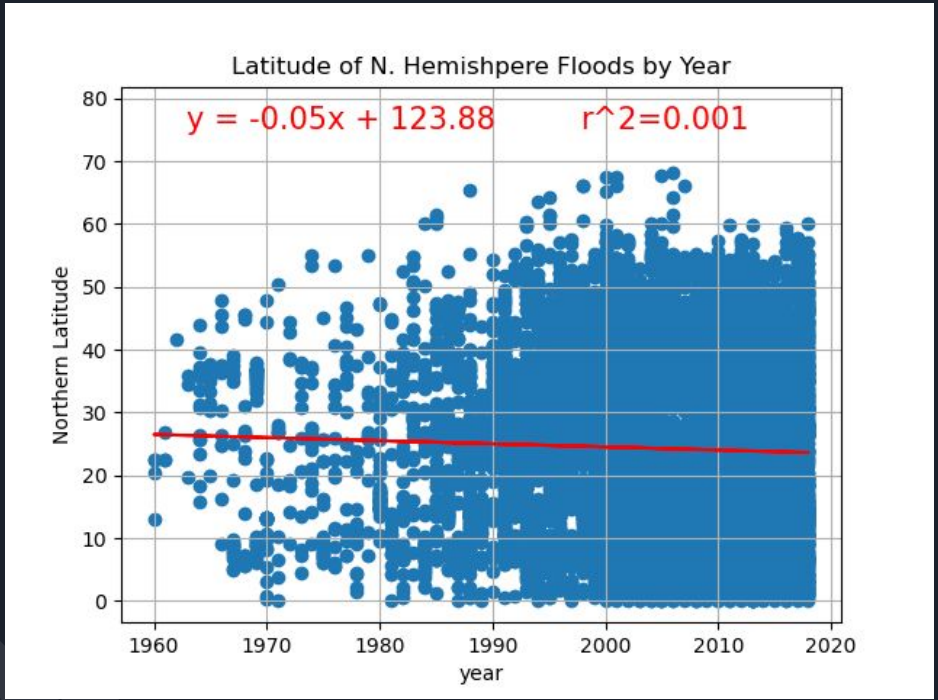


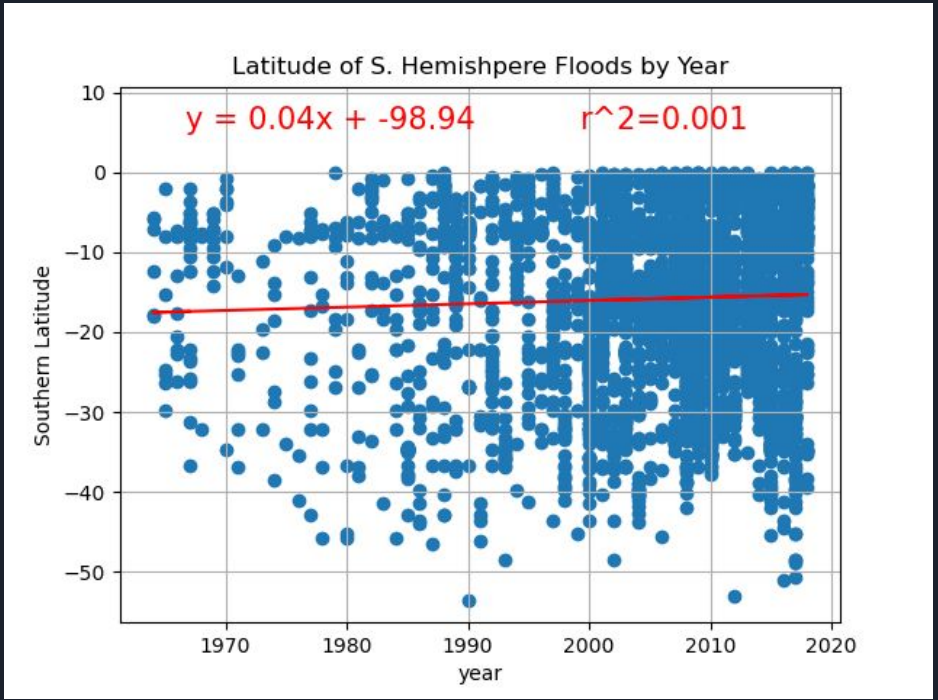
Visualization 5: Spread of Storms by Latitude





Visualization 6: Spread of Floods by Latitude



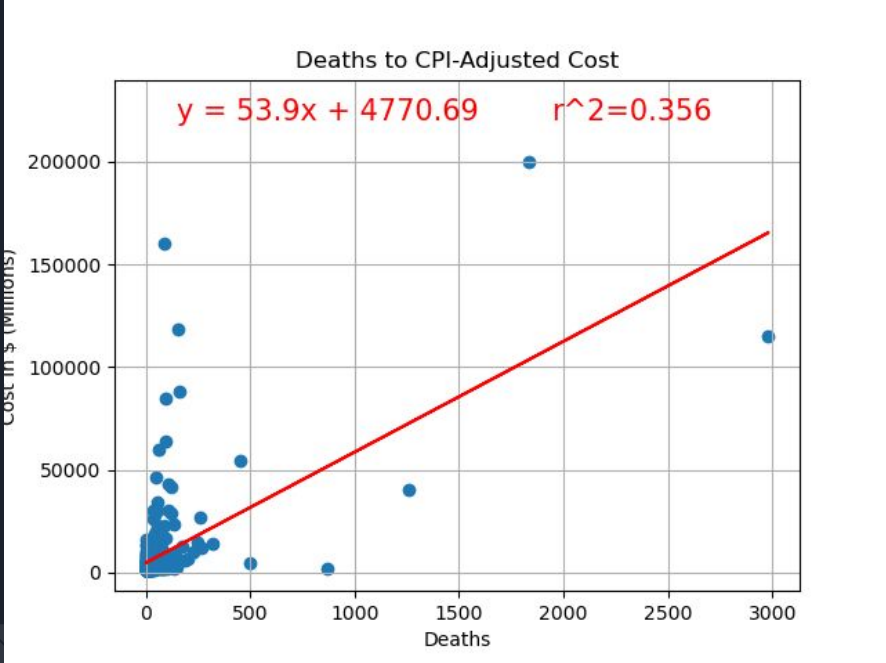


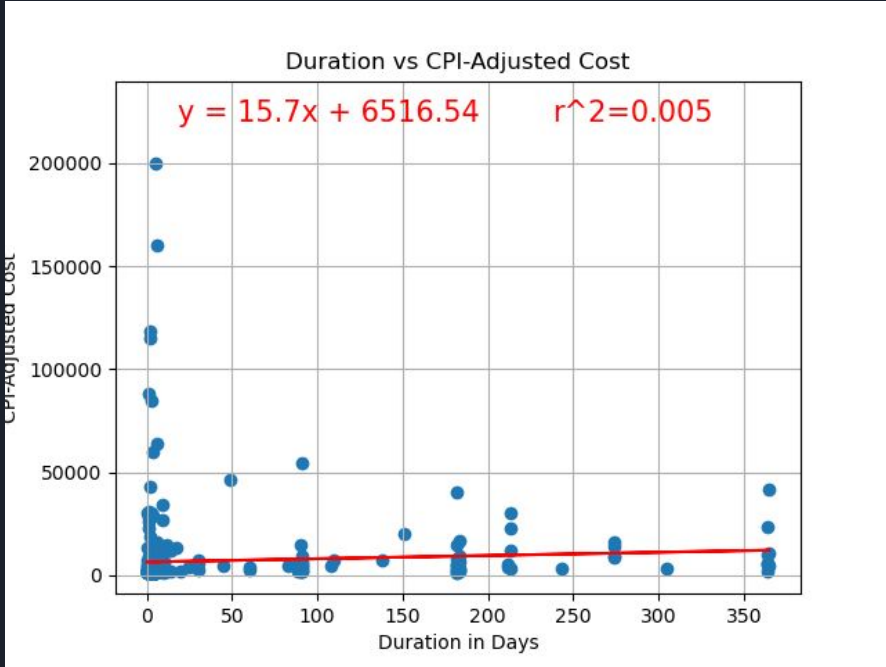
Dataset 2 - Data Cleaning and Processing

1	Name	Disaster	Begin Date	End Date	CPI-Adjusted Cost	Unadjusted Cost	Deaths
0	Southern Severe Storms and Flooding (April 1980)	Flooding	19800410	19800417	2742.3	706.8	7.0
1	Hurricane Allen (August 1980)	Tropical Cyclone	19800807	19800811	2230.2	590.0	13.0
2	Central/Eastern Drought/Heat Wave (Summer-Fall...	Drought	19800601	19801130	40480.8	10020.0	1260.0
3	Florida Freeze (January 1981)	Freeze	19810112	19810114	2070.6	572.0	0.0
4	Severe Storms, Flash Floods, Hail, Tornadoes (...	Severe Storm	19810505	19810510	1405.2	401.4	20.0

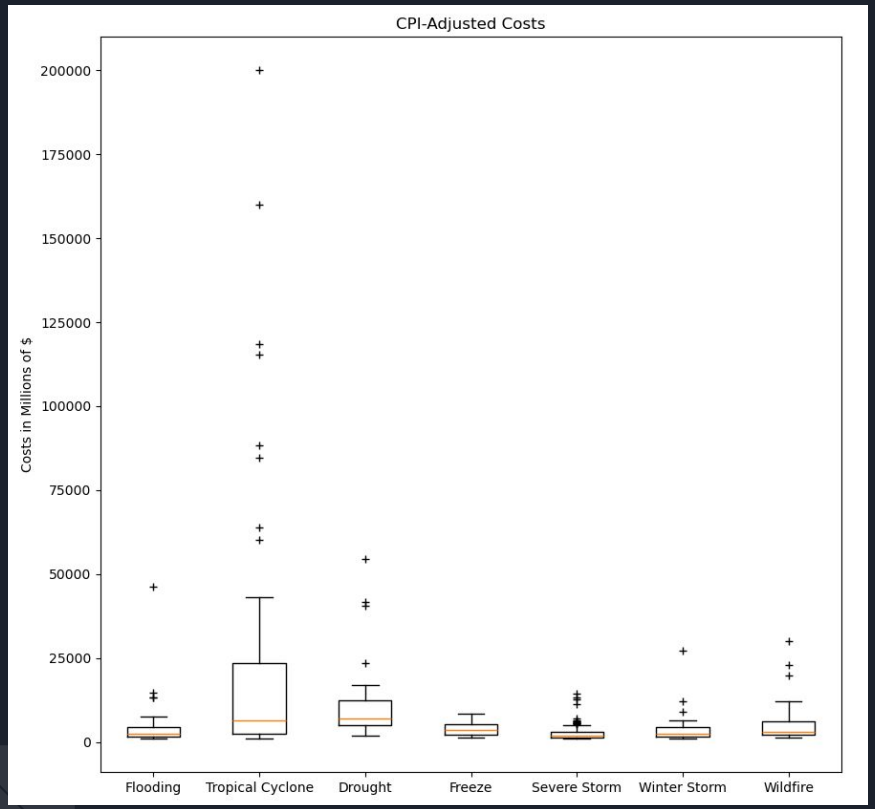
1	Name	Disaster	Begin Date	End Date	CPI-Adjusted Cost	Unadjusted Cost	Deaths	Duration
255	Hurricane Maria (September 2017)	Tropical Cyclone	2017-09-19	2017-09-21	115200.0	90000.0	2981.0	2
119	Hurricane Katrina (August 2005)	Tropical Cyclone	2005-08-25	2005-08-30	200047.3	125029.5	1833.0	5
2	Central/Eastern Drought/Heat Wave (Summer-Fall...	Drought	1980-06-01	1980-11-30	40480.8	10020.0	1260.0	182
64	Central, Southern and Northeast Drought/Heat W...	Drought	1995-07-01	1995-09-30	2017.6	984.0	872.0	91
88	Eastern Drought/Heat Wave (Summer 1999)	Drought	1999-06-01	1999-08-31	4772.8	2498.9	502.0	91

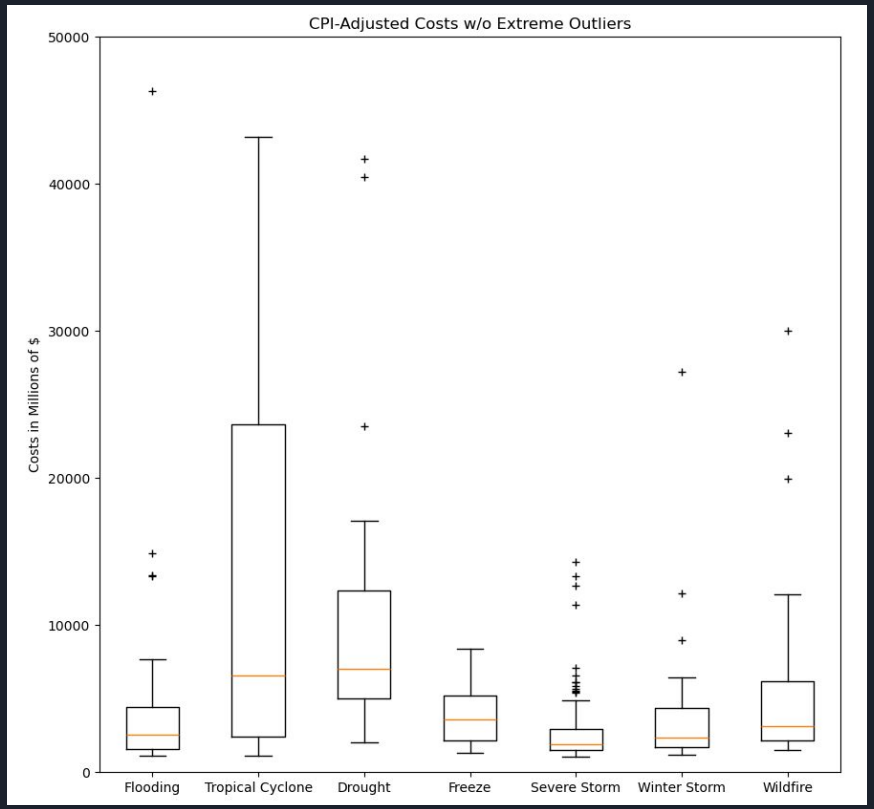
Visualization 7: Deaths and Duration to CPI-Adjusted Cost



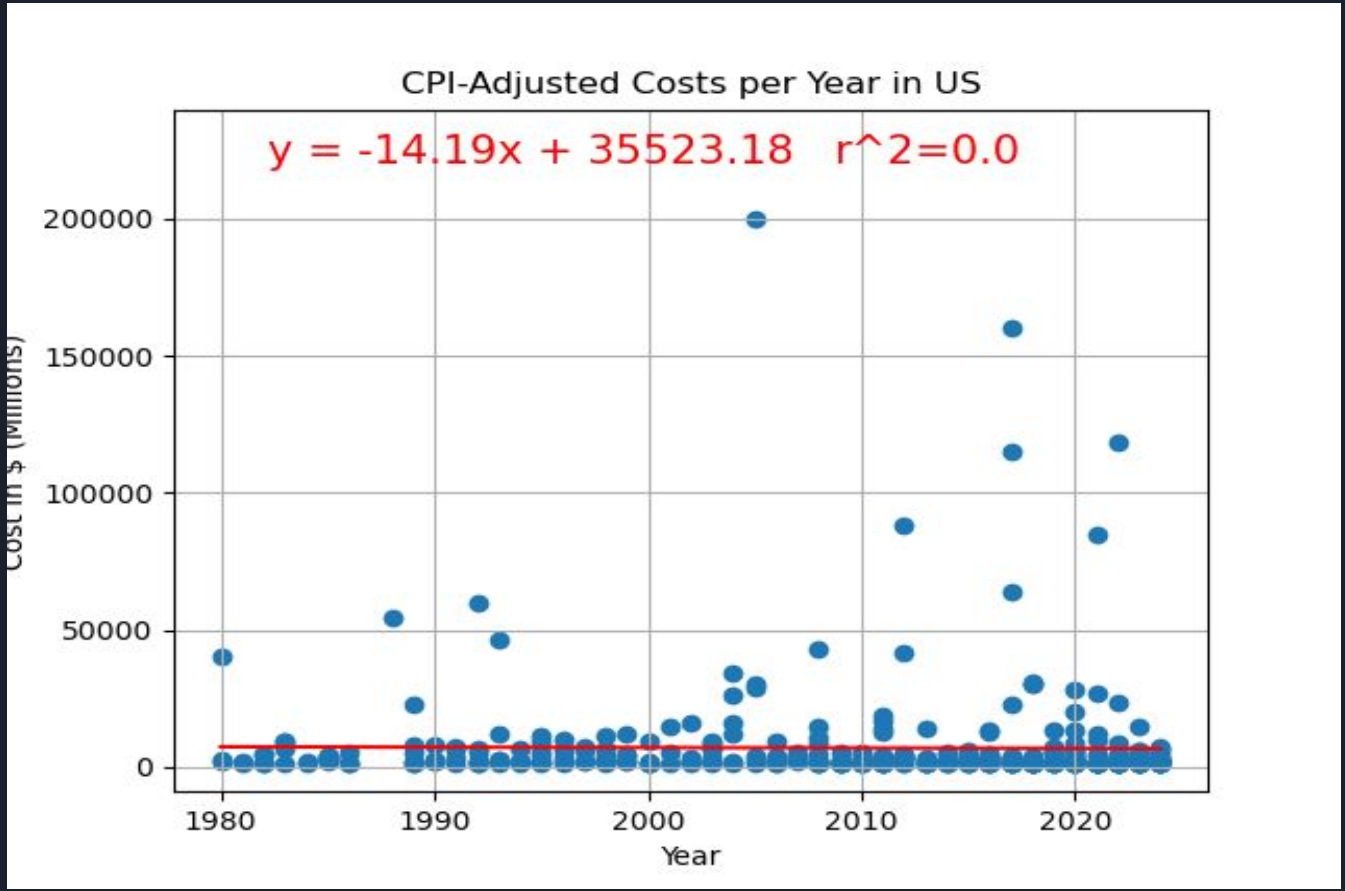


Visualization 8: CPI-Adjusted Costs by Disaster Type

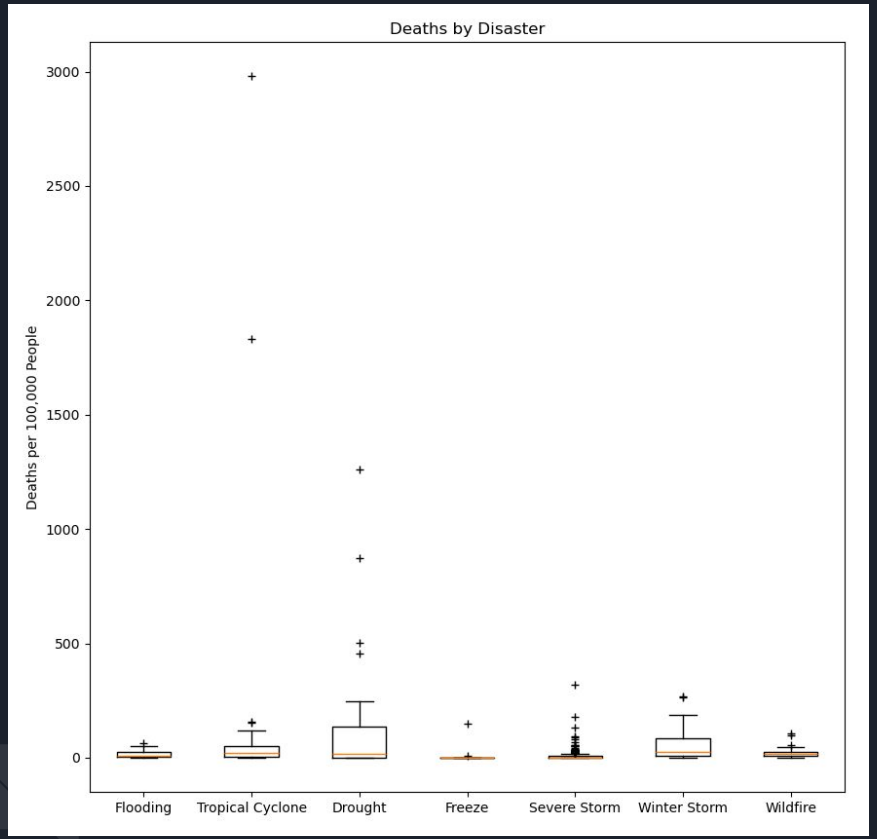


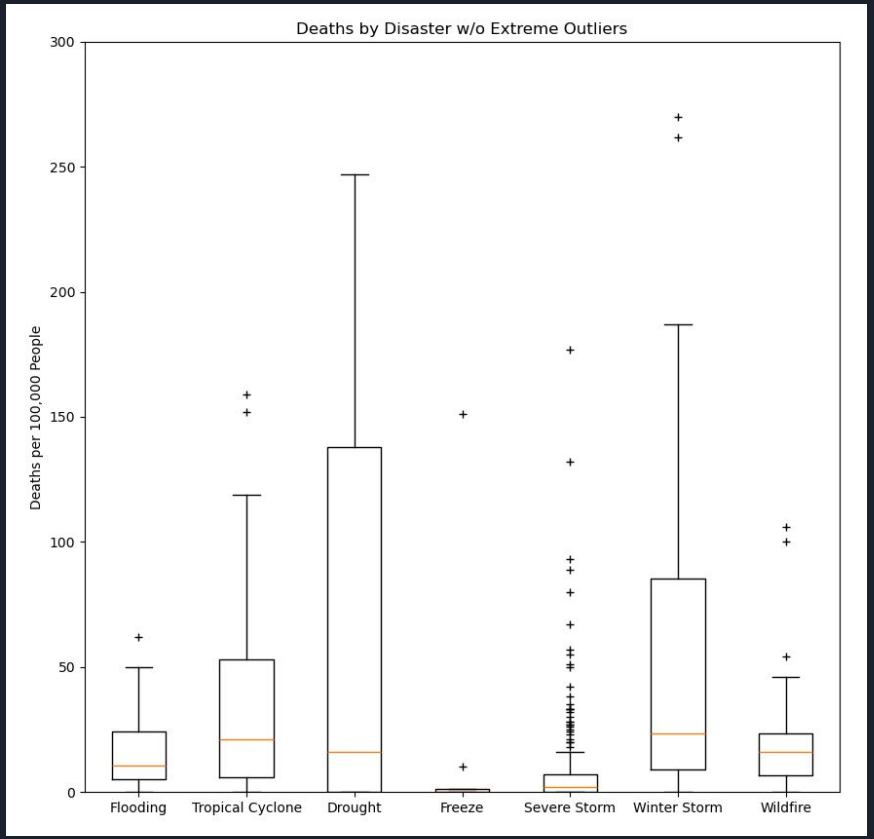


Visualization 9: CPI-Adjusted Costs by Disaster Type

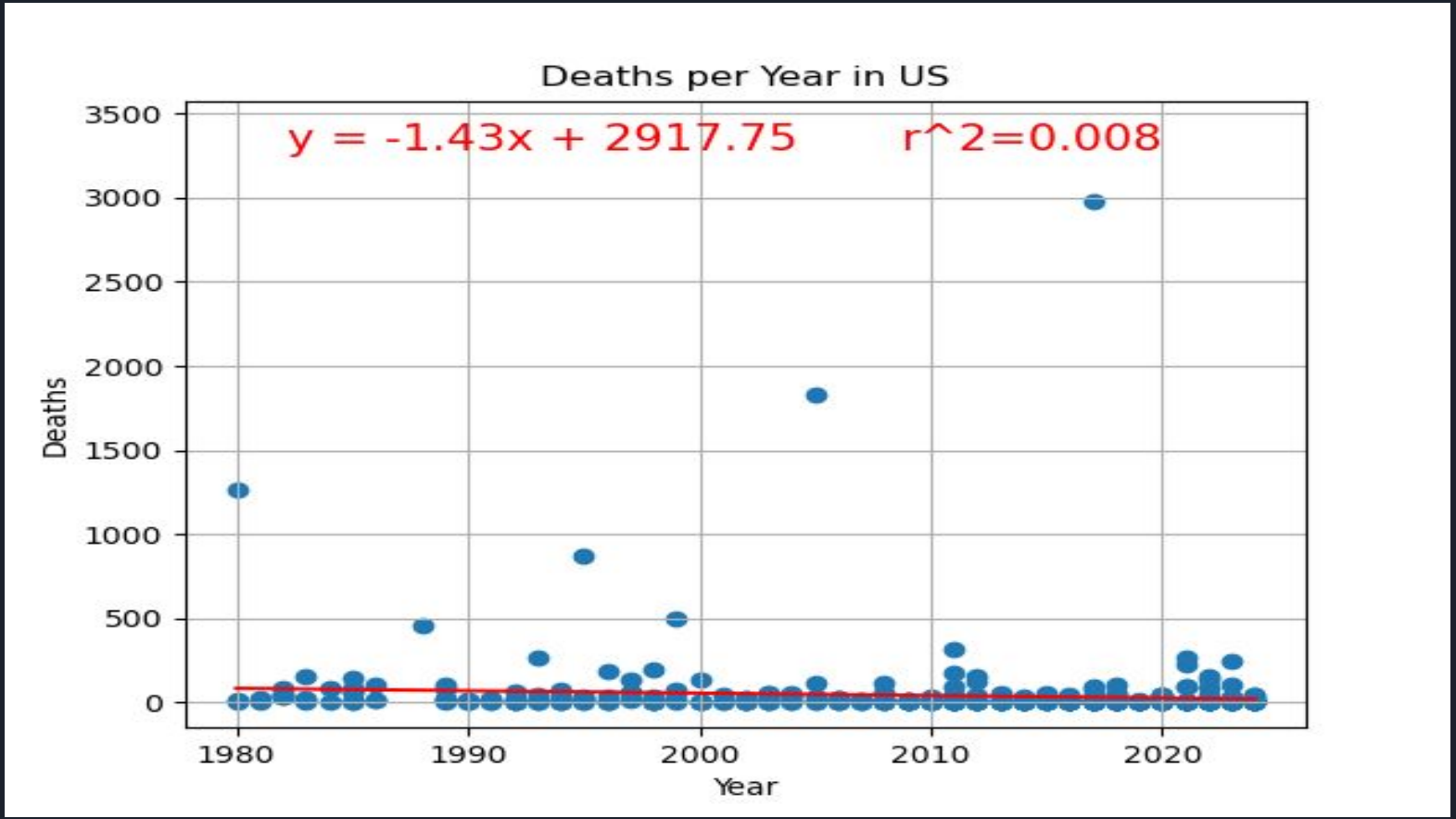


Visualization 10: Deaths by Disaster Type





Visualization 11: Deaths by Disaster Type



Findings / Conclusion

Rising Disaster Counts: The frequency of natural disasters, especially floods and storms, has increased over time, with both climate change and improved tracking likely contributing to the trend.

High-Impact Regions and Events: Brady, Texas, faced the most disasters from 1960–2018. Hurricane Katrina incurred the highest economic cost ($200 billion), while Hurricane Maria had the highest death toll (2,981).

Weak Cost-Death Correlation: Disaster costs show only a low correlation with death tolls, and no link was found between disaster duration and cost.

Data Limitations: Without storm intensity data, we cannot confirm assumptions about links between storm strength, cost, and fatalities.


END