Global Temperatures and Natural Disasters

By Jackie Bearnson

Project Overview

Motivation: Climate change is a well documented phenomenon, and one of the major issues facing humanity in current times. Governments around the worlds have been enacting legislation to address this issue.

Background Info: The National Ocean Service states hurricanes form in water over 26.5℃. Intuitively, fires may be more likely to start in hotter and drier conditions. Ocean levels are rising due to melting ice caps.

Hypothesis: Rising temperatures are causing an increase in certain types of natural disasters.

Dataset Description & Sources

Temperature data was from the dataset "Climate Change: Earth Surface Temperature Data" from Kaggle. It contains global average temperature data from 1750 - 2018, put together by Berkeley.

The dataset of natural disasters is titled "Geocoded Disasters (GDIS) Dataset" from SEDAC, a division of NASA. This contains recorded natural disasters around the globe, including country, year, and type of disaster from 1960 - 2015.

Exploratory Data Analysis

Using the Pandas tool in Jupyter Notebooks I ran reports to visualize the layout of the datasets from each source. The datasets were massive containing thousands of entries.

I decided to narrow down to data to a yearly view in order to get the best overall picture, comparing the global average land temperature to the number of disasters of each type.

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

Data Wrangling

The data was cleaned using pandas to get the datasets to display in a similar yearly format.

The disaster data was transformed to be a count of incidents and the temperature was averaged.

The dataset were then combined in order to go to the next step of analysis.

|	year	disastertype_drought	disastertype_earthquake	disastertype_extreme temperature	disastertype_flood	disastertype_landslide	disastertype_mass movement (dry)	disastertype_storm
0	1960	0	7	0	2	1	0	3
1	1961	0	1	0	2	1	0	5
2	1962	0	2	0	1	0	1	3
3	1963	0	2	0	2	1	1	6
4	1964	3	6	0	10	0	0	9

Statistical Analysis - Questions

- How has the average temperature of the Earth changed over time?

- Have the number of natural disasters increased over time?

- How has each type of natural disaster been affected?

- Is there a relationship between the Earth's temperature and the quantity of certain types of natural disasters occuring?

- What will the Earth's average temperature be in the future?

Earth's Temperature Over Time

Linear regression was applied to the data to get a trend of the average temperature over time.

The result had an R-squared value of 0.794 and a p-value under 0.05 indicating we can reject the null hypothesis and accept that the Earth's temperature is increasing.

Global Average Land Temperature vs Year
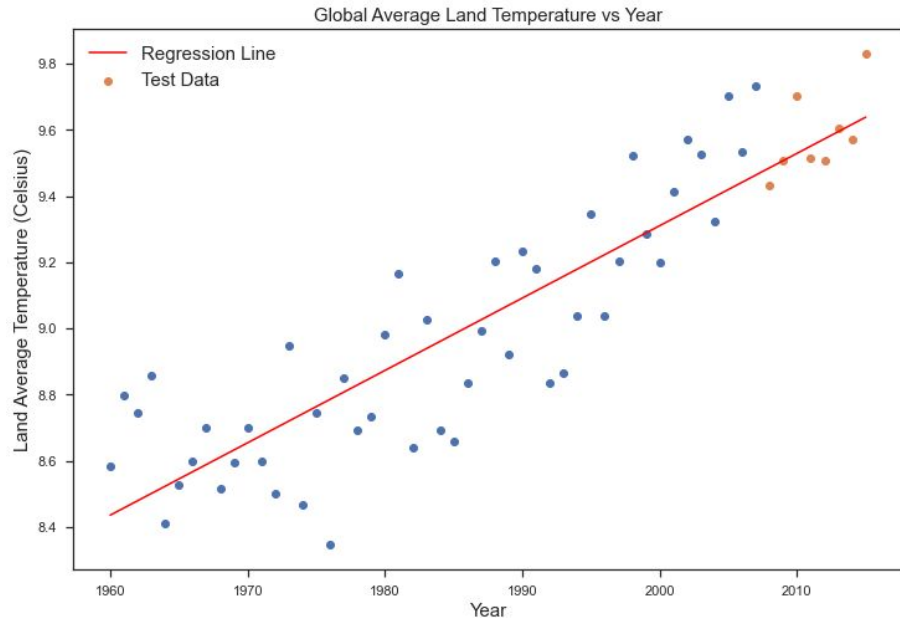
Natural Disasters

The line plot of the number of natural disasters from 1960 - 2015 indicates that floods have increased more than any other type of disaster with storms coming in second.

Extreme temperature events, droughts, and landslides also have visible increases.
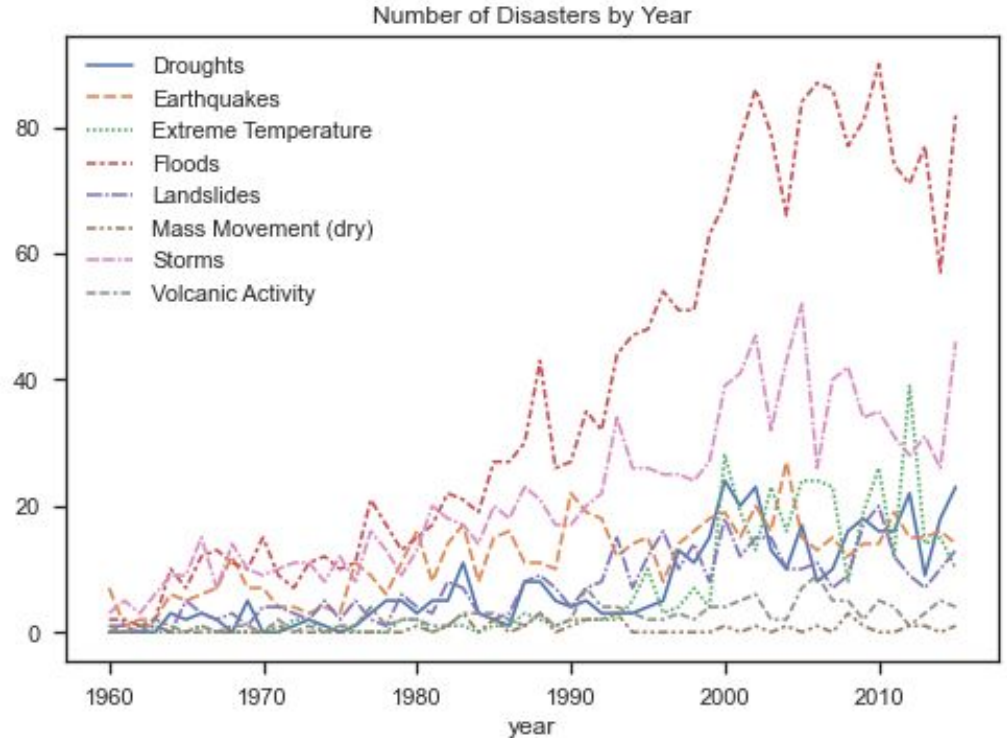
Natural Disasters - Cont

Nearly all disaster types have p values under 0.05 when regression is applied, indicating a statistically significant increases over time.

Since floods could be associated with ocean levels rising, a next step could be to look at locations of reported floods to see if there is a correlation with coastal regions and elevation.
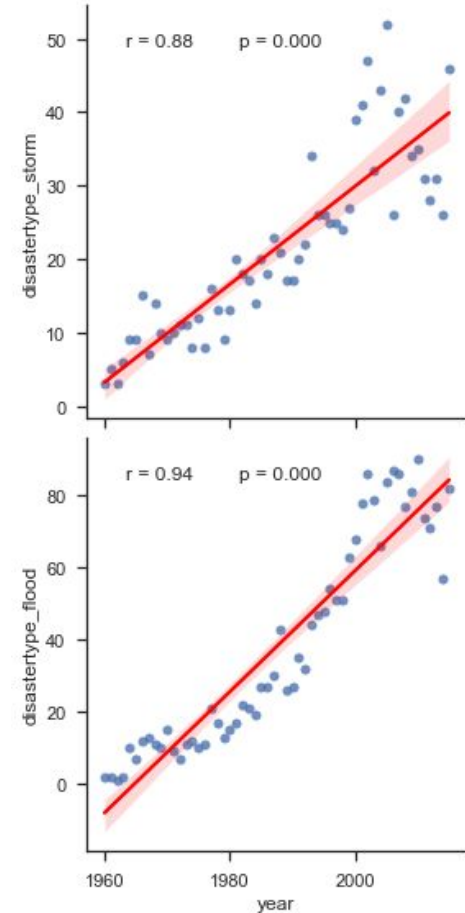
Relationship Between Temperature & Natural Disasters

The p values of each disaster type vs the global average temperature have a similar result with p values less than 0.05.

However, we need to invoke the first difference method in order to really understand the relationship because there may be other variables at play that correlate with time.
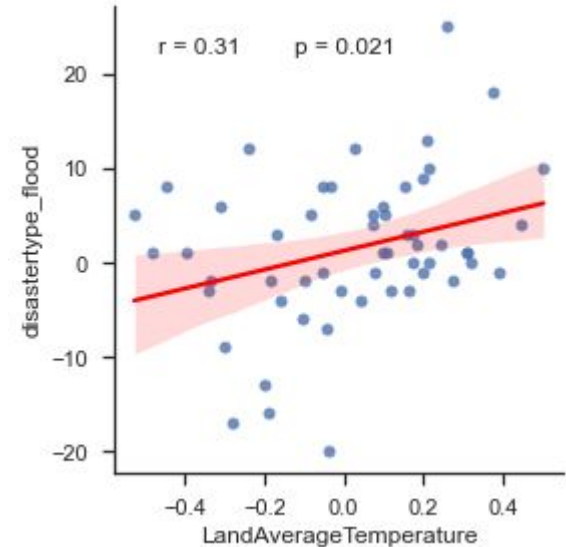
Global Temperature & Flooding

After applying the first difference method, the only relationship with a statistically significant p value is flooding.

The r-squared value of the linear model is 0.096, indicating a high degree of scatter and making it difficult to create a reliable model.

The Durbin-Watson is 2.446 indicating possible multi-collinearity.

The Future

I applied both the linear regression model and the Prophet package in order to get an estimate of the global average temperature in 2040.

Linear Regression: 10.18℃

Prophet: 10.3℃

This equates to an increase of about 0.0393℉ per year. However, some acceleration may be occurring.

Global Average Land Temperature vs Year
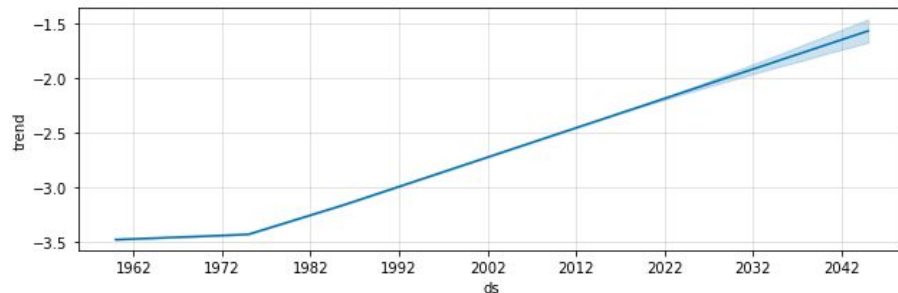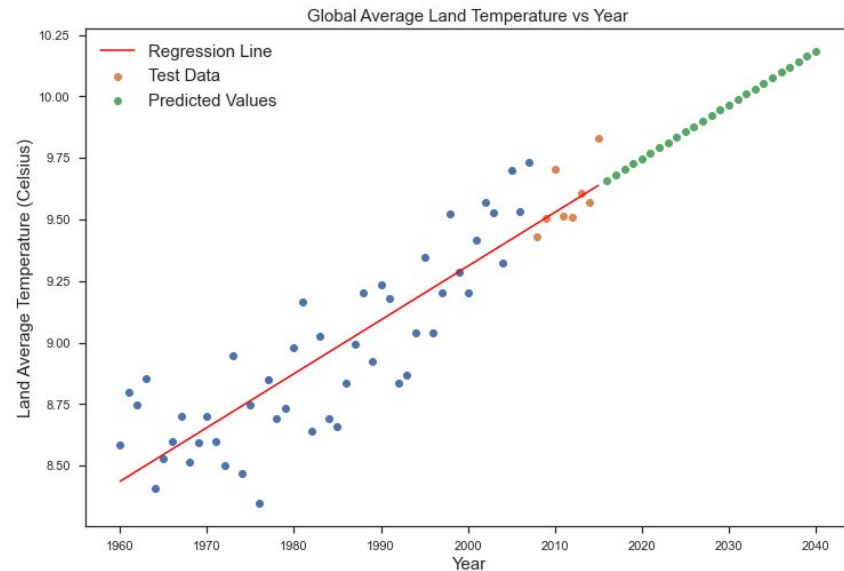
Conclusions

Nearly all types of natural disasters have statistically significant increases over time, with floods, storms, droughts, landslides, and extreme temperature events in the lead.

A possible skewing factor is better reporting with improvements in technology.

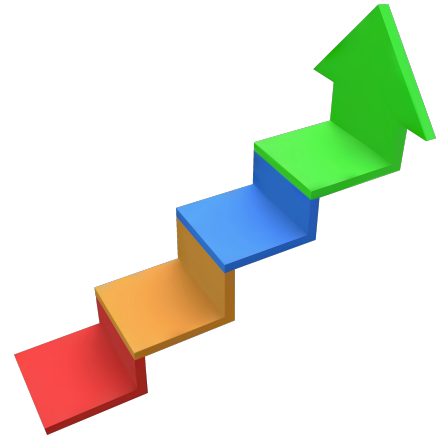Global temperatures are increasing at a rough rate of 0.0393℉ per year.

Floods have the best measurable relationship with changing temperature but more testing is required.

Next Steps

Look at a monthly basis to see if it offers better insights about the relationship between temperature and disasters.

A more indepth look at flooding, including the regions where these events are occurring and a look at their elevations to visualize the effects of ocean level rise and explore other possible causes.

Creating models of disaster occurrences to see where disasters are most likely to occur in the future where applicable.

Links

Github:
https://github.com/JackieAttack/Data-Analytics-Capstone

Temperature Data:
https://www.kaggle.com/datasets/berkeleyearth/climate-change-earth-surface-temperature-data?resource=download

Natural Disaster Data:
https://sedac.ciesin.columbia.edu/data/set/pend-gdis-1960-2018