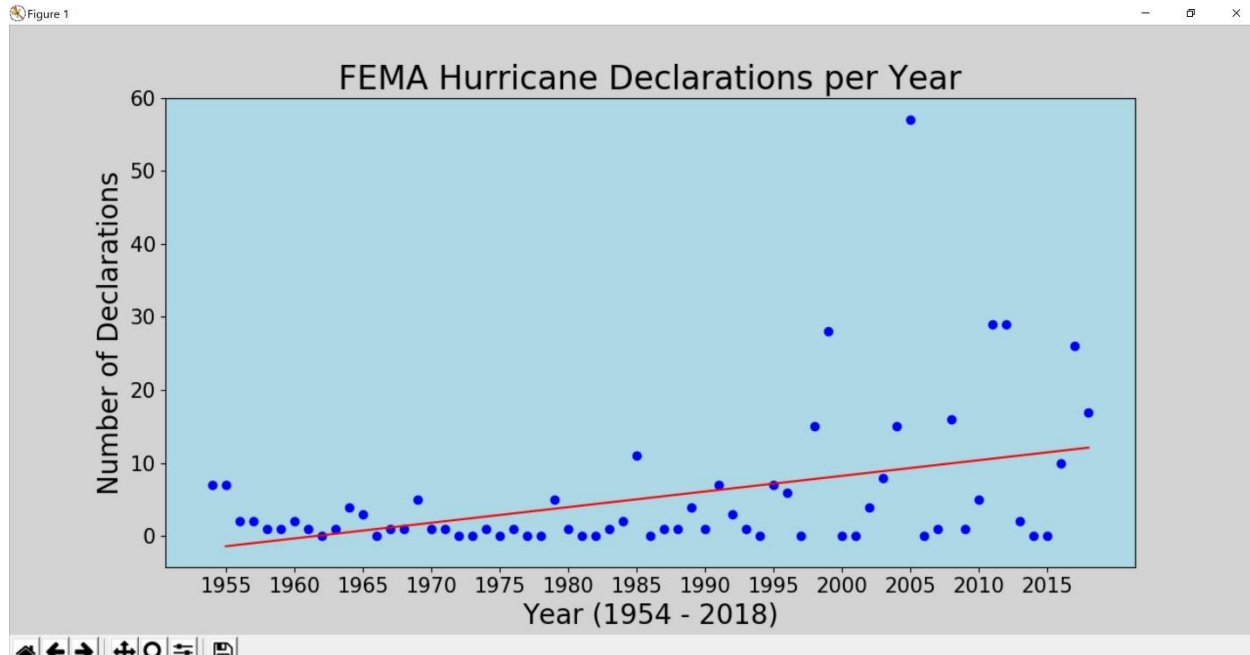# ISTA 131 Final Project: FEMA declarations and Disasters

**Plot 1:** Scatter plot of FEMA Hurricane Declarations per Year

The first plot shows the amount of FEMA declarations from the year 1954 to 2018 (Present time). The plot also includes a line of regression to show how the amount of FEMA declarations have changed throughout time. This shows us that over time there has been a slight increase in the amount of declarations through the years.

To gather the information to create the first plot, the FEMA declaration summary csv file was read and extracted to a data frame. Since the data frame includes multiple recordings with duplicate declarations based off region, the data set was condensed by removing any duplicate declaration numbers, thus creating a series with years as its index and the number of declarations as its value. A regression line was then added by using the statmodels OLS function.

Files used: femaDeclarations.csv and finalprojectdl.py (get_disasters and get_hurricanes functions were used) Link(s) to data set: https://www.fema.gov/openfema-dataset-disaster-declarations-summaries-v1

Plot 2: Average FEMA declaration classifications (Per year)
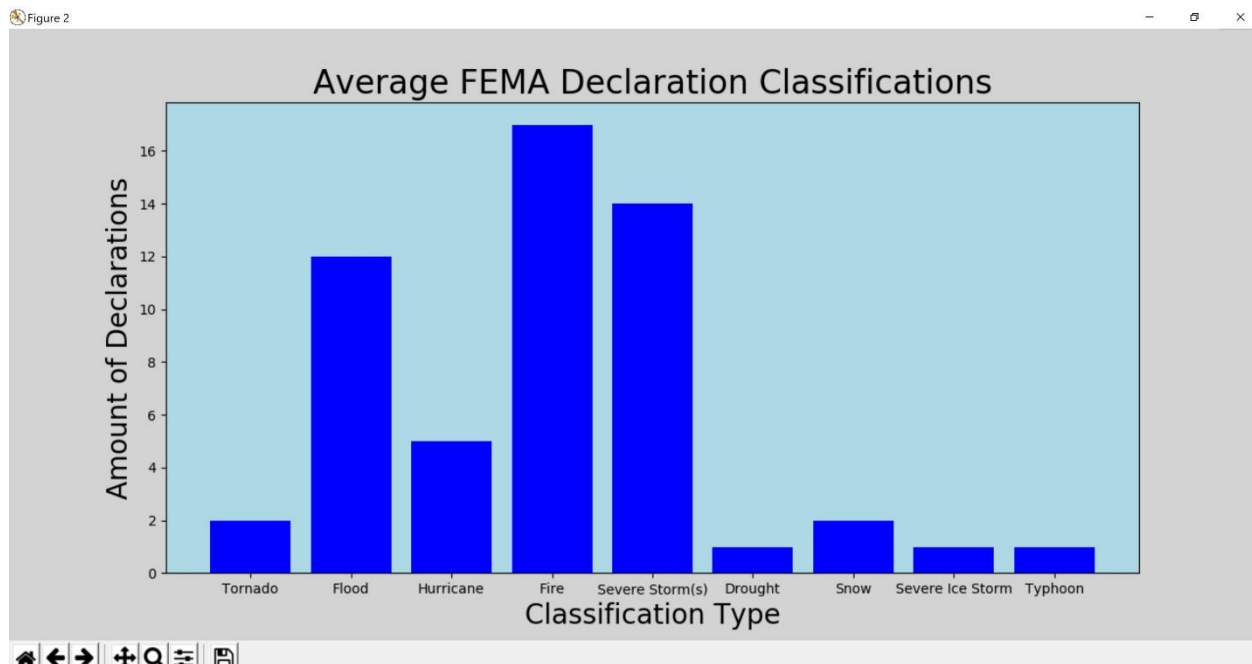
The bar chart displays the average declaration amounts for each classification type found in the disasters data frame created from the get_disasters function in the finalprojectdl.py file. Classifications with an average less than 0.5 are omitted from the bar chart, and all values are rounded up to the nearest whole integer. This gives an estimate of the disasters that are declared throughout an individual year in the U.S.

To generate the data for this plot, each individual classification from the femaDecalarations data set was checked to and added to a data frame indexed by year of occurrence. The data set contains columns based on declaration classifications, i.e. hurricanes, tornadoes, etc. Once all the data was extracted the average from the yearly data was gathered and placed in a separate Series with disaster classification as its index and the yearly average of incidents declared for that classification as its value.

Files used: femaDeclarations.csv and finalprojectdl.py (get_disasters and get_declaration_types function) Links to data set:
https://www.fema.gov/openfema-dataset-disaster-declarations-summaries-v1

Plot 3: Average Yearly Deaths per Disaster
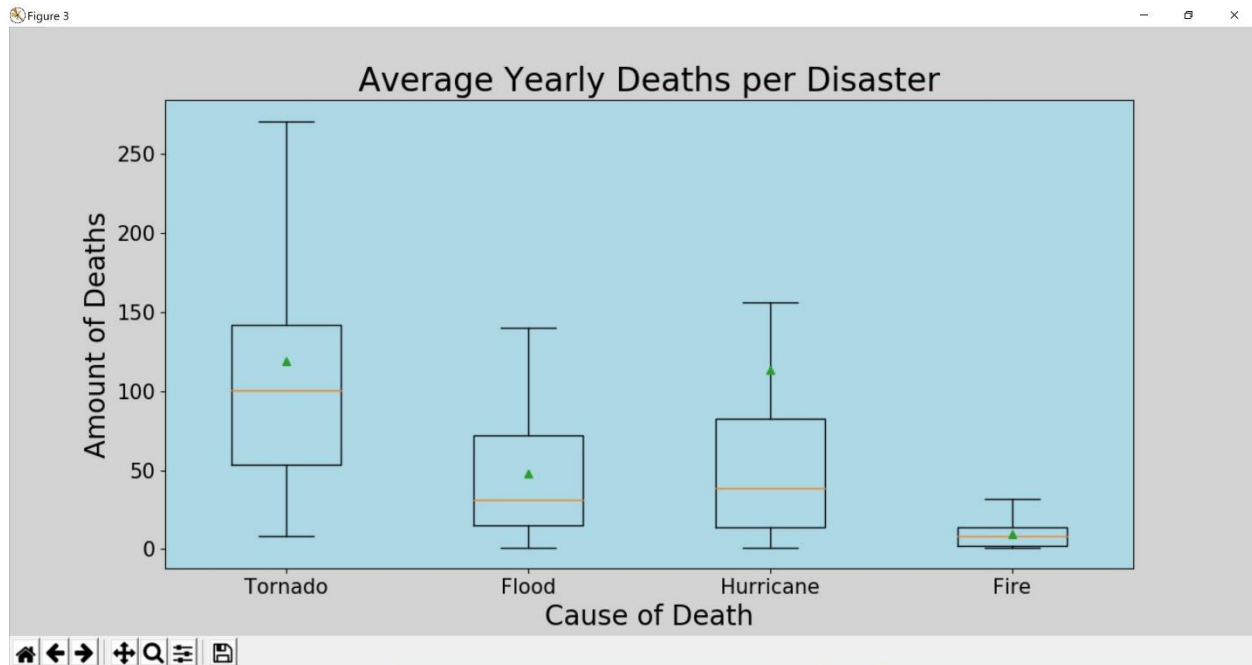
The final plot displays four side by side box plots showing the spread of yearly deaths from data collected since 1953 to 2018. The box plots show the median of the spread along with the average for each disaster. To clean up the plot, all outliers were removed, hence why some of the averages showcased are much higher than the median within the spread. For instance, the hurricane box plot average is thrown off due to 2005's Hurricane Katrina, which resulted in more than 1800 deaths. Yet, disregarding averages, more people tend to die yearly due to tornadoes than any other disaster.

To generate the data used for the third plot, a data frame of yearly information for each disaster type was extracted from its corresponding csv file. Then the data was cleaned up by removing any nan values and extracting only the Deaths columns from each data set.

Files used: us_fire.csv, us_floods.csv, us_hurricanes.csv, us_tornadoes.csv, and finalprojectdl.py (get_data and compare_deaths_by_disaster function) Link(s) to data sets: https://www.emdat.be