

# Natural Disasters over Time: A Study of the Frequency and Economic Impact of Climate Related Disasters

by Haley Egan, Maggie Houck, Will  
Johnson, Diana Morris



# Motivation

- ▶ Climate change is thought to be associated with an increase in certain natural disasters, among them drought, wildfires, flooding, and extreme temperatures.
- ▶ Our project tracks the frequency and severity of such disasters (measured in numbers and total economic impact) over time.

# Research Questions

Are natural disasters increasing over time?

Are natural disasters increasing in severity over time?

Are natural disasters and their severity linked to Climate Change?

# The Data

- ▶ A subset from Kaggle of a more comprehensive disaster dataset maintained by EM-DAT, the International Disaster Database.
- ▶ Tracks the total number of disasters of each type along with the total economic impact.

# Explanation of Data

## Criteria for a Disaster's Inclusion

At least 10 deaths

At least 100 people affected (including injured or homeless)

The declaration of a state of emergency or an appeal for international assistance.

## Economic Damage

Includes total estimated losses, reconstruction costs, insured losses. All amounts in thousands of U.S. dollars (2019 value).

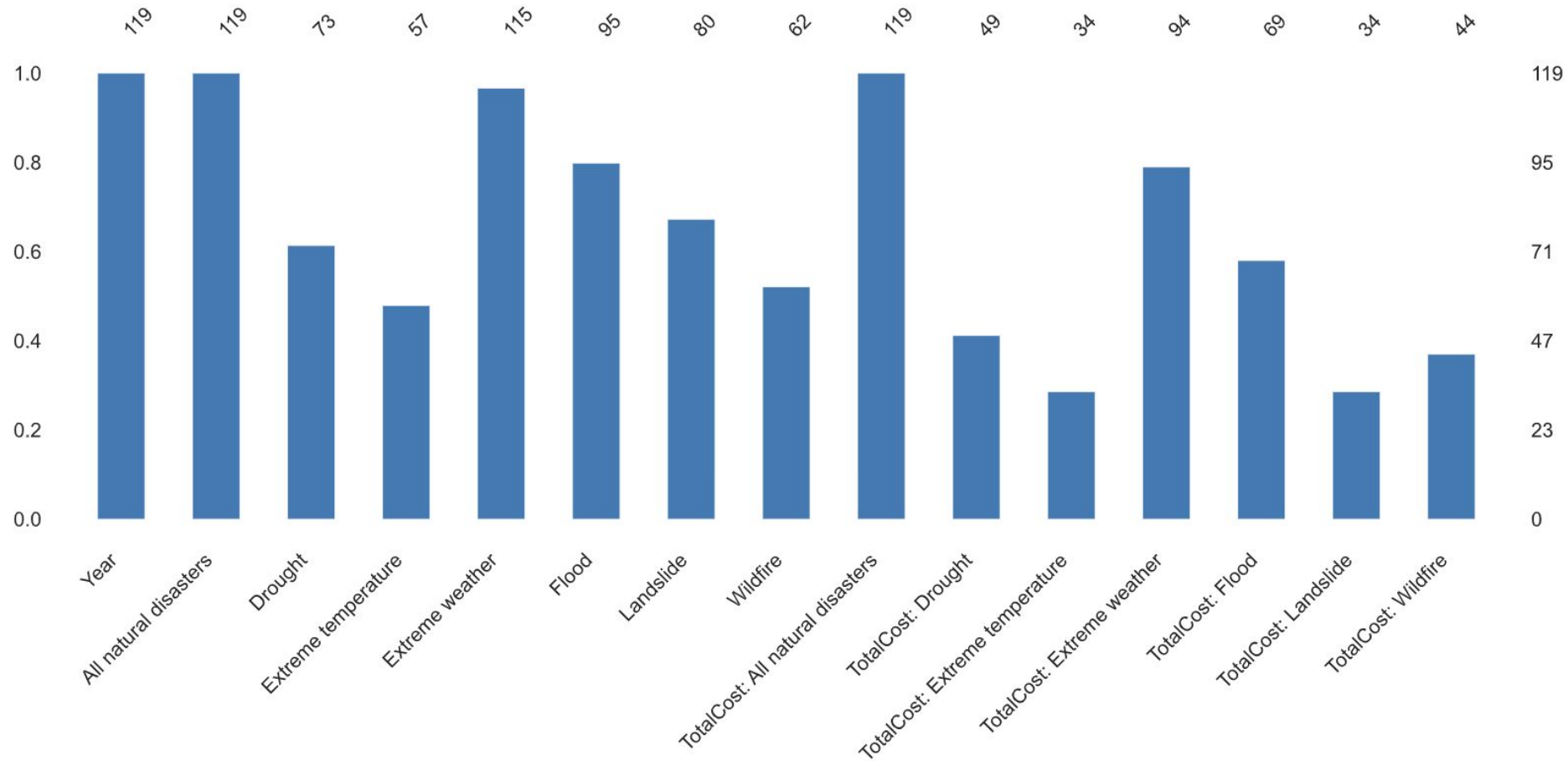
# Data Cleaning

Not all disasters are  
climate-change related

- ▶ Volcanos,
- ▶ earthquakes,
- ▶ tsunamis

Missing Data present

# Data Cleaning



A simple visualization of nullity by column.

# Data Cleaning

- ▶ Disasters unrelated to climate change were removed.
- ▶ Economic damages and frequency of disasters were combined in a single dataframe.
- ▶ Years prior to 1970 were removed due to many missing data points.
- ▶ Years after 2015 were removed due to reporting lags in the source data set.



```
import unittest

class CorrectColumnsTest(unittest.TestCase):

    def test_deleted(self):
        check1 = DisasterData()
        check1.cleanDf()
        self.assertNotIn('Earthquake', check1.df.columns)

    def test_retained(self):
        check2 = DisasterData()
        check2.cleanDf()
        self.assertIn('Extreme weather', check2.df.columns)

class ColumnsTypeTest(unittest.TestCase):

    def test_type(self):
        check3 = DisasterData()
        check3.cleanDf()
        self.assertEqual(check3.df['Extreme temperature'].dtypes, float)

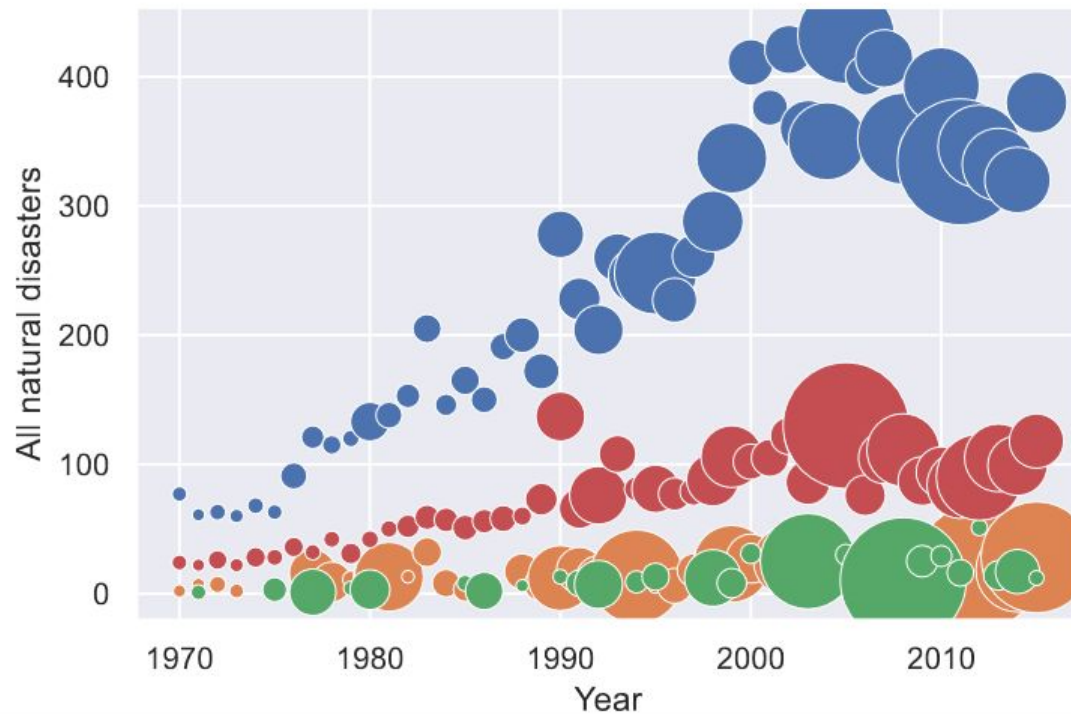
    def test_type2(self):
        check4 = DisasterData()
        check4.cleanDf()
        self.assertEqual(check4.df['Extreme weather'].dtypes, float)

if __name__ == '__main__':
    unittest.main()
```

# Unit Tests

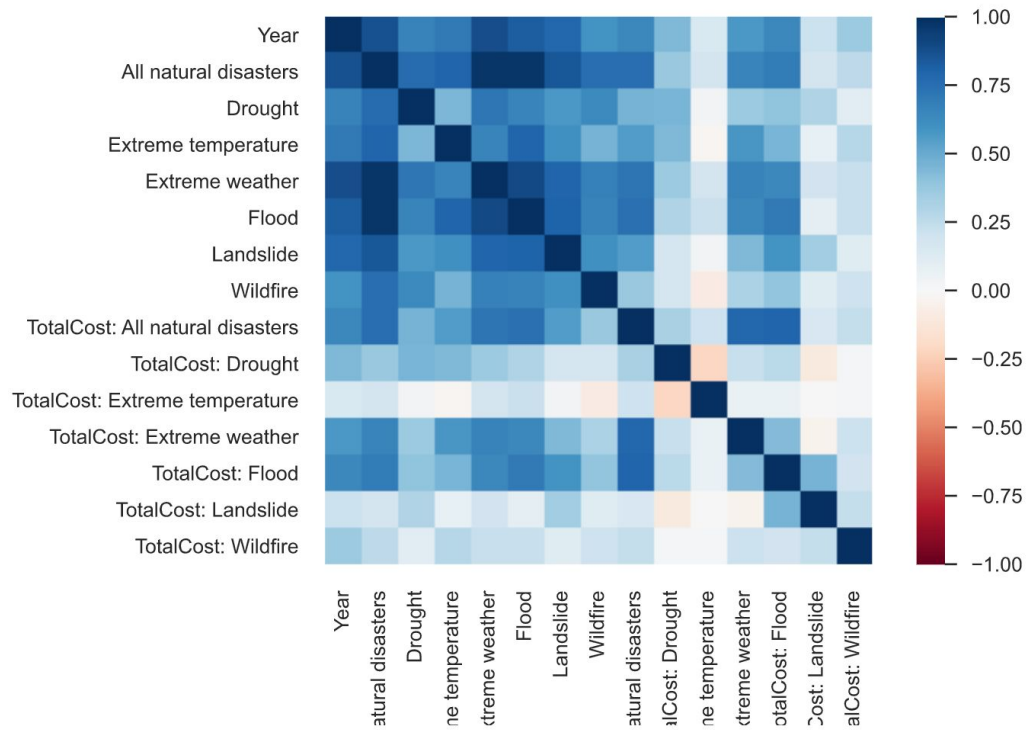
- ▶ We confirmed that our cleaning had in fact retained the relevant columns, excluding disasters not linked to climate change.
- ▶ We further tested that the resulting columns yielded the correct data type.

# All Natural Disasters



- All Natural Disasters
- Extreme Weather
- Drought
- Extreme Temperature

This graph depicts the trend of disasters over time. The size of bubbles represents the total cost, with different colors tracking different types of disaster.



# Correlation Matrix

# Conclusions

- ▶ A clear upward trend in the number of natural disaster over time, for all natural disaster types, as well as an increase in economic impact.
- ▶ The increase in economic impact may indicate the increase in severity of the natural disasters over time.
- ▶ With our current dataset, we are unable to conclude that these increases are directly tied to Climate Change, something that will take further analysis with a larger scope.

# Where to go from here

- ▶ This can be used to make budgetary decisions on future disaster relief.
- ▶ Bringing in carbon emissions data on top of our current dataset as a covariate in our regression analysis.
- ▶ Controlling for economic/population growth over this time period. Maybe there has been more for disasters to destroy.
- ▶ Geospatial data to control for where these disasters are occurring.