

Predicting severe storm events in the United States

The goal of this capstone project is to predict the number of storm events aggregated at the state level. Building this model can potentially assist states/cities and the general population in preparation for severe storm events (e.g., infrastructure planning, stocking up on supplies, insurance, familiarizing with evacuation routes, etc.). People will likely have experience with these events, but I believe the prevalence will increase with climate change. It is important to start to understand the rising trends because two or more severe storm events within a small window of time can be exceptionally more devastating (e.g., a severe flooding with a hurricane in the following days with no time to recover from the flood).

The National Weather Service provides storm data containing statistics on personal injuries and damage estimates from 1950-present (<https://catalog.data.gov/dataset/ncdc-storm-events-database>). These events include, hurricanes, tornadoes, thunderstorms, hail, floods, drought conditions, lightning, high winds, snow, temperature extremes, etc. There are 51 columns including datetime, location, narrative, damage, etc. I will start by exploring features available in the dataset and determine relevant features to start building the model. This includes visualizing any relationships, as well as the general trends over time per state and overall in the country. I would also like to look at historical meteorological data and analyze any correlations (primarily temperature) and add them to the dataset if they would add to the predictive capabilities. I will provide Jupyter Notebooks with code and some explanations, a PowerPoint presentation, and plan on writing a blog post.