Climate change is arguably the greatest danger facing our society, and its effects could have devastating impacts on every community on the planet. Communities located on oceanic coasts face the likelihood of rising sea levels and hurricanes that will increase in both size and frequency, both of which exacerbate issues of flooding.[[1]](#footnote-1) One such place is New York City, the largest city by population in the United States. New York is an incredibly diverse city, perhaps the most so on Earth, with a majority-minority population and hundreds of spoken languages. However, it is also a city of great inequality, famously home to many of the world’s wealthiest people, but also with a poverty rate higher than the national average. The combination of these facts posed an interesting set of questions for us to explore: when flooding occurs, what kinds of people in New York are most at risk? As climate change worsens, who will be most in danger of these hazards? And lastly, in terms of designing climate change resilience projects, can our findings inform policy decisions?

In order to investigate these questions, we examined both historical and projected data concerning flooding in New York and developed several models to examine the relationship between demographic data and a given area’s risk of flooding. We used census tracts as our unit of analysis and pulled a wide array of demographic data from the American Community Survey, using the 5-year average as of 2020. We then joined this data with the 2050s 500 year floodplain, a map showing what areas of the city are projected to be flooded in the event of a storm that has a 1 in 500 chance of happening in a given year. By using projections based on the sea level rise expected by 2050, we can model what could happen in the event of a storm that is very unlikely to happen in a single year, but likelier over the course of many years or decades. Since a single storm of this magnitude could prove absolutely devastating for the millions of people who live in New York, it is prudent to examine a worst-case scenario, particularly since projections concerning climate change have been repeatedly shown to be overly optimistic. We cleaned both sets of data and performed exploratory data analysis to examine the immediate relationships between flooding and various demographic metrics, finding a statistically significant relationship between the percentage of a census tract living in poverty and its likelihood of lying within the floodplain. We also noticed that areas within the floodplain had lower employment rates. We included visualizations demonstrating the overlap of these two metrics, along with race and income, with the floodplains.

To see if we could devise a model that could predict whether or not a census tract was in the floodplain based on its demographics, we created a decision tree model, a logistic regression model, and a k-nearest neighbors model, finding that the KNN model was the best.

We then looked at historical data, using Superstorm Sandy as it was the most recent weather event to severely flood the city. We used census data from 2010 rather than 2020 to reflect the state of the city at the time when Sandy hit in 2012. We then cleaned a dataset showing the damage and water levels inflicted by Sandy, performed exploratory data analysis, and found that the areas most damaged by the storm were disproportionately white. To see if we could create models that could predict the level of damage based on demographic data, we used a Random Forest model, an elastic net model, and a KNN model, finding that the random forest was the most accurate.

1. <https://www.c2es.org/content/hurricanes-and-climate-change/> [↑](#footnote-ref-1)