1. Trudy Garrett

Data 601

Fall 2025

**Data Science Goals**

I have been a contract data scientist for over 10 years. I have analyzed data for organizations that have received grants from the National Science Foundation, the National Institute of Health, the Substance Abuse and Mental Health Services Administration, the Department of Education, and the Department of Defense and other organizations. In most cases, the data could fit in excel.  More frequently, employers want information gleaned from big data. I am now a National Oceanic and Atmospheric Administration (NOAA) CESSRST Fellow.  I am pursuing a Graduate Certificate in Data Science to expand my skillset, specifically to analyze NOAA’s big data.

  Big data refers to large structured and unstructured data that would overwhelm traditional data processing systems. Examples of big data include internet clickstream data, social media posts, sensor data, financial transaction logs, healthcare records, satellite imagery, and multimedia files like videos and audio.  Organizations use big data analytics to uncover trends and make better decisions.

NOAA collects environmental big data.  The data is derived from satellites, weather models, ships, and ground-based sensors. NOAA data includes satellite imagery, radar data, weather models, atmospheric measurements, ocean data, remotely sensed imagery of land and coastlines, and lightning strike data. In Data 601, I want to understand how I can use machine learning to explore the relationship between lightning and wildfires using NOAA’s big data. I am also interested in using AI to build models to predict extreme weather.  I am particularly interested in generating models to predict weather impacts on the island of Antigua.

Additionally, I am interested in analyzing social media posts of minority responses to and understanding of weather alerts. Currently, NOAA uses sample survey data to measure population response to weather warnings.  The survey population includes mostly white affluent males. Understanding weather preparation behavior for minority populations using big data is in line with NOAA’s mission to protect life.

There are some ethical concerns with the use of big data including the inability to get consent for the use of subject data and knowledge extraction. Using big data to analyze social media data does not give subjects the opportunity to consent to their posts being used in a study. This is unethical because it violates a person’s right to consent and privacy. Some social media posts contain community knowledge. Extracting that information without consent or compensation could be perceived as community theft of ideas. Because of these ethical concerns, analysis using big data should handle the data with care, ensure privacy, and consent with community members about the use and publication of any culturally unique information.

In Data 601, I hope to learn about the ways in which I can analyze NOAA’s big data and social media data that is relevant to NOAA’s mission. My goal is to predict extreme weather events and to explore population differences in weather alert interpretations and responses. Additionally, I hope to learn strategies for respecting subject privacy and rights to community knowledge.