Hook Document: Investigating and Predicting Hurricane Activity

GitHub Link: https://github.com/gub4rk/DS4002_CS3

Topic:

This study researches the relationship of climate change variables, such as carbon emissions, on the number of days a hurricane is present in the Atlantic and Pacific oceans. The data for this topic is from 1971 to 2021. After completing this case study, you will be able to answer the question: Can a model predict if the number of days a hurricane is present in the Atlantic Ocean rises over time due to climate change?

Context:

Over the course of July-October 2024, there was a lot of news of hurricane activity and the devastation that resulted. With climate change also being a central issue in the media, it is important to investigate whether climate change variables have an effect on hurricane activity in the Atlantic and Pacific oceans.

Motivation:

Imagine you live in Florida but you are traveling on a work trip. You see the news that a hurricane is hitting and you come home to see your home destroyed. This is a familiar occurrence for citizens of Florida, and Hurricane Milton which occurred during early October in 2024 was no exception. It was the second most intense hurricane on the Atlantic coast in history. With the power of data capturing tools, data scientists are able to use all available resources to aid the improvement of society. Data in this project includes surface temperature, sea temperature, carbon emissions, and more variables that have only been able to be captured recently. Since scientists have gathered this data, it is essential to use it in an impactful way and see how humans can change their climate-related activity in an effort to reduce hurricanes, saving lives, buildings, the economy, and preserving the environment.

Deliverable:

The student will explore the dataset with visualizations and the code for them given by the case creator. The project will hone in on the years 2018-2021 and try to predict the results for these years based on previous years. To do this, the student conducts various tests to study relationships between the independent variables as well as validating these variables. Following this, the student will create a predictive model to gain insight on how the variables influence each other over time. To finalize the deliverables, the student will create visualizations which show the contrast between the variables in actuality compared to how the model forecasts variables. Lastly, the RMSE will be calculated to see how the model is performing. This will be the primary metric for how successful the model is.