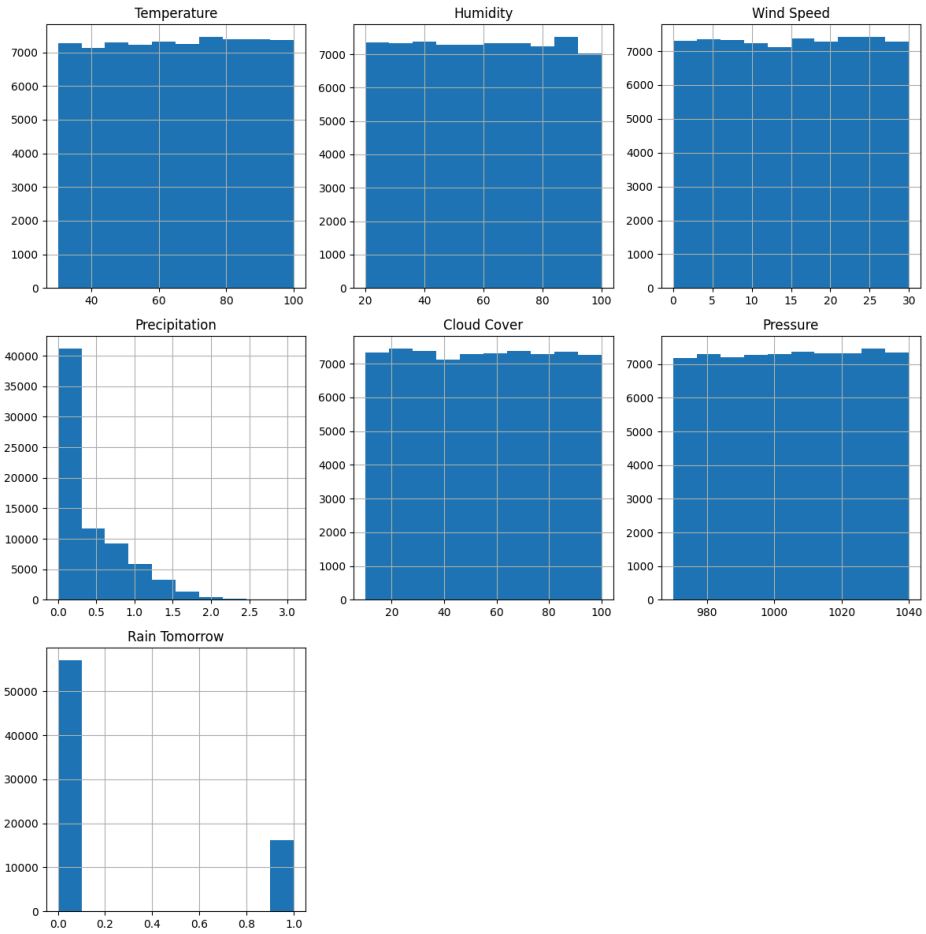
Project Intermediate Update

So, machines can, in fact, learn how to visualize data and use it to then be trained to make predictions from what it is programmed to do. For this project intermediate update, I will go over the preprocessed dataset, issues with implementation and training dataset, and challenges that we face to our knowledge.

Preprocessed dataset and exploratory analysis



We have managed to download the dataset from Kaggle, visualize it, and make a PNG of the image. From the visualization above, we were able to collect the data from twenty major USA cities' rainfall prediction datasets on Kaggle. With this dataset, we will build and train a predictive model to analyze weather trends. As far as we are concerned, we are still figuring out

how we could expand this idea to create a near-future forecast that predicts flash floods and warns people for disaster preparedness.

The person in charge of the dataset has tasked himself to implement MongoDB as well as the dataset from the National Oceanic and Atmospheric Administration to have a global hourly dataset to then get an accurate severe weather prediction across the globe.

Initial model implementation and training results

We have yet to implement a model and train it due to the dataset within the program not being imported, yet, which was a bit of miscommunication within the project team members. We plan on being more communicative as time progresses. After this weekend, we should have a model and training going, and you will be able to look into our GitHub repo and find the changes necessary.

Challenges faced and next steps

One of the main challenges so far was preparing the dataset, as it was giving a lot of lip. For more accurate prediction data, we are on the lookout for another dataset that'll produce better results. For this, we seek to use MongoDB; if there are any recommendations, we are also open to listening. At the moment, we have yet to build and train the model, but as mentioned will soon bring those changes.