Project Progress Report

Hannah Weber, Jacob Paul, Marissa Kelley

12.7.18

We began our project by pulling the dataset from Kaggle. We found that it was quite large, so we began by doing a Sequel query to get the data out and also reduce the mass amount of data. We started by doing a groupby to look at which state had the most amount of wildfires, and turned it into an interactive bar charts. We’ve decided to predict looking at the causes of the wildfires, and look at it in the context of the state. We haven’t run into any problems or challenges yet. Our next steps are to begin doing a random forest, perceptron and logistic regression on the data.

We are planning on doing an oral presentation, but if need be, we can just do a paper (if there are too many groups presenting).