**West Nile Virus in Chicago Predictions**

Group members:

Zach, Jason, Sam

Project planning description:

To manage our team’s time on this project, we first met together to discuss the project problem statement and explore the data together. After getting a strong grasp of the data and what was necessary to answer the problem statement, the group members discussed what was needed to complete this project and broke the project down into tasks. We then ranked the tasks according to their agreed upon priority. Next, we divided the tasks amongst the group members to work on individually. When we were ready to merge the first round of completed tasks onto our Github repo, we met and filled each other in on how we accomplished the tasks so that each member had an intimate knowledge of each step in the project. We then discussed the next most important steps and divided the next set of tasks amongst the group members to work on individually. We repeated this process while updating our overall task priority list along the way until the project was complete.

We made a list of the tasks and questions to investigate in Microsoft OneNote and then discussed how to divide them up verbally when we met together.

**Here is the list we created:**

Goal:

Given weather, location, testing, and spraying data, this competition asks you to predict when and where different species of mosquitos will test positive for West Nile virus.

Things to do:

* Exploring
  + Weather
    - How does this relate to mosq pop
    - Temp
    - moisture
    - windspeed
    - Correlation between days since last rain?
      * Look at a running average of precipitation over time
      * And temp too
    - How long has it been hot and dry?
      * Does a long dry run lead to a greater increase in mosquitos
  + How is time a factor?
    - What time are mosquitos most active
    - When have they been spraying
  + Is spraying effective at all?
  + Have they been spraying even though we don't have data for it?
  + Look at other datasets
    - Where are the birds at?
      * Species of birds
      * Number of birds around each year
* Get a model running
  + Figure out what model
    - Has to account for autocorrelation and extra predictors
  + Feature engineering
    - Getting onto one timescale
    - Group into week
    - Should we be scaling? Scaling what?
    - Interaction terms
    - Convert location to weights based on past mosquito populations?
  + Combine test stations together
  + Calculate relationship between groups of predictors
  + Calculate prediction intervals

* Today
  + Location
  + Standarizing
  + Optimize clustering
  + Lag weather data
* Deliverables
  + Project planning documentation

Thoughts:

* Make sure we have enough comments
* Pesticide spray w/ weather patterns
* Distance to water a factor?
* Populations of mosquito
  + Predators
  + Rats/ other sources of blood
* Do they prefer certain types of plants?
* CO2 levels?
* Spray method
  + Spray used
  + How is it distributed
* Based on some googling
  + Heat (over 80 degrees) could be important
  + Moisture level
  + Wind speed
* Interesting info
  + They normally don’t travel more than a few hundred yards from where they hatched. -Culex mostiqoes
  + They mostly eat sugary nectar and only use blood for egg development
  + *Culex pipiens*, known as the northern house mosquito. It is the main carrier of **West Nile** virus.
  + [Culex pipiens](https://en.wikipedia.org/wiki/Culex_pipiens)
  + Females live between 40-65 days, males only live for 10
  + Generally, mosquito **activity** will begin when the temperature reaches the 50° F level. Mosquitoes thrive on hot weather. Thus, as the temperature begins to rise, the mosquito volume increases accordingly. The mosquito season reaches its peak during the hot summer months.
  + West nile virus mainly infects birds and that's how