Week-1 of project 4

* Step one is down loading the data and setting up our work flow management stuff. Trello, Slack, Git repo.
  + Task and benchmark goals were set for the following two weeks
    - We touched base on some of our initial observations in the data sets.
* Data cleaning was our next step.
  + Nulls and Noaa coding in the weather were dealt with.
  + We decided to keep duplicates, a a max observation per trap was 50
  + Converted all date to date time format for concat ease
  + Eliminated useless location data, kept traps and lat lon
  + Made month and year columns
  + Data viz rainfall and temp, sunrise vs temp, wind v departure
  + Tableau viz of sprays and weather, sprays and west nile, and distribution of traps and WN.
* Feature Engineering
  + We first worked on engineering features for weather. We settled on averaging temps and rain fall over a 7 and 14 day period based on life cycle of mosquito.
  + Made dummies for traps and species
    - We also included num mosquitos

Week-2

* Modeling
  + Now that we had our data and concat all our features we began testing some models
  + We ran 3 different random forest models.
    - One with grid search
    - One with gird search and sratKfolds
    - One without grid search and with sratKfold
      * We then examined our feature importance
  + Then passed our top 5 importance from our forest into a logistic regression to get odds ration for there impact on WN
  + We also ran a SVM and XGboost on our data and found our best ROC score with XGboost.