**Ai and Ml Python Week 2**

* Regression is a statistical process for  
  estimating the relationships among  
  variables
* Logistic regression is where the label or target variable is binary
* When to use
  + Should have fairly clean data, not too many outliers, missing values, complex relationships
  + Benchmark binary algorithm
  + Don’t use it when you have a continuous variable
  + Don’t use when you have a lot of data
  + Not the best performing algorithm
  + Summary
    - Logistic regression is a good algorithm for a fast, transparent baseline model forbinary classification problems
    - It doesn’t do too well when you have a lot of data or “messy” data
* Hyperparameters
  + Every algorithm has some levers, things  
    that you can tune or “tweak” about the  
    algorithm to see if the outcome changes
  + C hyperparameter for logistic regression is a regularization parameter which controls how closely the model fits the training data
* Regularization
  + Overfitting: when we fit our data into a model we have and get good results, but then we get poor results using live dataset
  + Trying to mitigate overfitting is called regularization