July 28

I. Final Project - Update & review
   A. Problem Statement
   B. Data sources
   C. Next up - specify a method >>> Are you able to run your model??? >>> What are your results
II. Responses
   A. Machine Learning Foundations #8, Tokenization for Natural Language Processing
   B. Machine Learning Foundations #9, Using the Sequencing APIs
   C. Word Embeddings
      1. Why is using one-hot encoding an inefficient towards vectorizing a corpus of words? How are word embeddings different? (see this video https://www.youtube.com/watch?v=EEk6OiOOT2c)
      2. Compile and train the model from the tensorflow exercise. Plot the training and validation loss as well as accuracy. Post your plots and describe them.
      3. Stretch Goal: Follow the link to the Embedding Projector provided at the end of the exercise. Produce the visualization of your embeddings. Interpret your visualization. What is it describing? Is there relevance with regard to words that are proximate to each other?
   D. Text Classification with an RNN
      1. Again compile and train the model from the tensorflow exercise. Plot the training and validation loss as well as accuracy. Stack two or more LSTM layers in your model. Post your plots and describe them.
III. For tomorrow
   A. Lecture #10 - Using NLP to build a sarcasm classifier
   B. Text generation with an RNN
   C. Neural machine translation with attention