**Milestones**

### Milestone 1: 10/30:

* Construct Google Folder (and give TA’s, graders, and Sowers access and URL)
* Download some data
  + Debugging dataset: small enough to test code with; reasonable code should run in 2 minutes
  + Working dataset: large enough to do the problem on (training should run no more than 40 minutes)
  + Convert these datasets to pandas
    - I suggest that you convert datetime to pandas timestamps (allows for time deltas and time manipulation)
    - Pickle (<https://pandas.pydata.org/docs/reference/api/pandas.DataFrame.to_pickle.html>) the data. That converts it to a binary file which can be loaded directly (must faster) into the correct datatypes
* Make a README.md file
  + Listing the team members
  + Explaining the problem (as well as you understand at this point)
  + Stating a license

### Milestone 2: 11/20:

* Colab notebook giving some visualization of the data and some descriptive statistics. Explain what you are doing in text cells.
* For reference, carry out some sort of linear or logistic regression (to be used as a benchmark). Details left to you, but explain what you are doing in text cells in the notebook.

### Milestone 3: 11/29

* + Build a deep learning model for the dataset
  + Investigate effects of mini-batch learning
  + Investigate effects of different optimizers
  + Tune hyperparameters (training testing and validation)

### Milestone 4: 12/9

* + Documentation and cleanup of files
  + Conversion to repo
  + Video summary of project.
    - Should be between 5 and 7 minutes long. Note: we won’t watch the video beyond 7 minutes.
    - **Each slide should be labelled with list of group members who contributed to that slide.**