**Basic Plan of Attack for Assignment**

Feature creation/ manipulation (one hot encoding etc)

* Casual/Professional classification (based on time of day, etc)

Handling of categorical data

* One hot encoding, perhaps explore other techniques

Pipeline?

* May be beneficial in testing various options

Determining best mix of train/dev/test splits

* Aspect of request is to only use training data for periods earlier in the year compared to what is being predicted
  + IE Use January and Feb data for March.
  + This is beyond the scope of this course, but an exploration of “on-line” techniques could prove interesting
    - How far back to use information?
* Otherwise, preform a more traditional split
  + Random
  + Based on days of the month (the current method used in the contests test data).

Try out various methods/techniques (complete with parameter optimization)

* Standard regression
* Decision rees
* Bayes etc
* Other Methods