1. **Source**: Source is the raw data that we will use to create our ensembles.
   1. Each row of the source should represent an instance with the columns in each row representing a feature.
   2. By default, the last field is the feature that you want to predict and is called the objective or the response variable.
   3. The first row may be used as a header to provide the feature names.
2. **Data-Sets**: Datasets are a structured version of the source.
   1. Each feature is processed and categorized into numeric, categorical, text or date-time.
   2. Features are further categorized into weather, schedule and building data.
   3. The user can view the data.
3. **Model/Ensembles**- We will build the trees from the data-set. The user could then:
   1. view and plot prediction accuracy stats
   2. view important features
   3. control advanced tree tuning knobs: Pruning, stopping criteria, k fold, #of ensemble trees
   4. Filter trees and gather insights
   5. Explore different methods
4. Baselining
5. Evaluation
6. Synthesis