Project Milestone 1

I wrote a script that dynamically loads and categorizes data into pandas data-frames by the feature-set name (e.g. "bag-of-words"). The training data is then split into k portions for future cross-validation. I haven't calculated any statistics across feature values in specific categories. However, that is clearly my next step. I also plan on calculating the entropy of the dataset and attempting to visualize the principal components of certain feature categories to see if the labels are separable in certain feature categories.