The data set that we have used is found at <http://wwbp.org/downloads/public_data/dataset-fb-valence-arousal-anon.csv> and is associated with the paper *Modelling Valence and Arousal in Facebook posts* here: <https://wwbp.org/papers/va16wassa.pdf>. This dataset has 2800+ entries, which we will split into training and testing sets. Each entry in the dataset contains 5 values. The first value in the string that contains the data to be analyzed. The next 2 contain values depicting the sentiment of the data, while the next 2 contain values depicting the arousal of the data (or intensity). Both are on a scale of 1-9, with 1 being negative for valence and neutral for arousal, and 9 for very positive for valence and very high for arousal. Right now, the data reader splits the data into half for training and testing. This will be investigated further as to the best split between training and testing data.