Affect.csv: Affect on per paper-day basis
WordCounts.csv: counts of various key words on a paper-day basis
Content.csv: Measures of economic articles on a paper-day basis

Notes:

Content.csv has several measures, each of which is a percentage.

ArticleCount expresses the number of Economic articles divided by the total number of articles in the daily edition

WordCount expresses the number of words in Economic articles divided by the total number of words in the daily edition

As you can imagine, these measures are very similar

Thresholded: Any article with an economic score of more than 40% was counted as being "Economic"

Adjusted: Counts are produced by weighting with the actual economic score. Thus the word count for a particular article would be multiplied by its score.

Again, these measures are a bit different, but the data is pretty similar. I found that the 40% threshold was the best performer and squares pretty well with what an average reader would consider to be an article about economics. In practice, economic scoring ranges from 0 to about 75% with the occasional article going higher. Below 30%, things start getting noisy, but above 30% content, my test set scores just about perfectly. Thus, I have a high degree of confidence in both these measures. Of the two measures, I think thresholded is a bit more accurate, but adjusted works well.

Affect.csv Looks exactly like Content.csv except I have not included any thresholded counts as the adjusted counts are accurate as my affect scoring system is consistent across the range of possible affect values. Higher numbers are more negative. As you can note, the media is unsurprisingly negative on average, though the affect does change with respect to events.

WordCounts.csv: A count of various words on a paper-day basis. Each variable is a paper+a word.

Assorted Changes:

My measures are a bit better for these data as improved methods and more computational horsepower has allowed me to use larger training sets. I've also changed my scoring system a bit in an effort to eliminate as much random noise as possible. I've also dropped the weekend editions from the analysis as these are substantially different than the daily editions (a lot more leisure content in the supplements) and are frequently not retrievable from Lexis Nexis in any case.