R: Spellchecking

Signal Data Science

We'll wrap up the lesson on functional programming with a short project in natural language processing which will draw on all of the concepts in R which you've learned so far.

Spelling correction is one of the most natural and oldest natural language processing tasks. It may seem like a difficult task to you at the moment, but it's surprisingly easy to write a spellchecker that does fairly well. (Of course, companies like Google spend millions of dollars making their spellcheckers better and better, but we'll start with something simpler for now.)

- Read Peter Norvig's How to Write a Spelling Corrector, paying particular attention to the probabilistic reasoning (which is similar to the ideas behind a naive Bayes classifier). Recreate it in R and reproduce his results.

- **After** implementing your own spellchecker, read about this 2-line R implementation of Norvig's spellchecker.