* What would a 2nd book look like?
* Explain Tokenisation – removing odd characters, collapsing to stem words?
* Need to click “Update data” at bottom of wrangling paragraph before the graphics work
* ***Issue: choice of stop words lists*** or import your own
* ***Change*** control of frequency plot from min freq count to max no. of words displayed (consistency then also with word clouds)
* ***It would be nice if***
  + some freq displayed (maybe table) let you click on words and have them added to stop words and even nicer if everything immediately updated
  + You could get a list of the stop words you added
  + Stop-word list names and added stop words appear under graphs
  + Would be nice to be able to click on a frequent word/type in a word and see all the lines with that word in it – better still if that word could be highlighted in some way
  + Ability to save text after wrangling
* Control on size of word clouds
* Refreshing word clouds (is there a random element in layout???)
* Think about control of sentiment words displayed on initial view
* Add control to Bing sentiments display
* Think about some way of finding natural divisions for “sentiment scores by groups of lines”, eg. Begin line Chapter sp
* Sentiment negation not working
* Choosing sentiment lists and import your own
* Ability to add words to stop words list from sentiment para

Relationships

* Network Graph not working
* Control of Co-occurrence Counts displayed
* First Word, Second Word almost always gets characters’ manes back to front. Is there a reason for the ordering?
* Nice to add bigram freq displays & maybe even trigrams
* Correlation Tables: you can type in your own word. Would be nice to also be able to choose for a frequency table of common words
  + Would be nice to be able to click on a bigram and see all the lines with that bigram in it – better still if that could be highlighted in some way – better still if the target words lined up vertically so you can concentrate on what comes before and after

ADDING MULTIPLE FILES

What if we had several books

* Frequencies by “book”
  + Scatterplots for 2 books
* Sentiments by book
* Sentiments by book by “chunk”
* Correlations by book
* Topic modelling by book
* Text classification????
* Working with whole sentences

Other things

* Word frequencies
* Sentence lengths
* Line graphs for word frequencies over time\
* Natural clusters of words used together in documents
* *Get Cassidy to look at Chapt 10 of … (esp for the examples)  
  2015 Taylor Arnold & Lauren Tilton Humanities Data in R:Exploring Networks 15.Humanities Data in R.pd*f