

## Exercises

2. Let's use the text of "Oliver Twist: The Parish Boy's Progress" Charles Dickens. This is book number 730 in the *gutenbergr* package.

(a) To investigate how often word  $X$  is followed by word  $Y$ , separate the text into two columns: "word1" and "word2". In this process, remove all cases where either word1 or word2 is a stop word.

(b) Visualize all bi-gram relationships that occur at least 4 times ( $n > 4$ ) simultaneously (following the guidelines in slide #16 of today's lecture).

(c) Visualize the networks ( $n > 4$ ) of the co-occurring words as in slide #17.

(d) Pick four common words ("no", "not", "never", "neither") that negate the subsequent term, and use the joining and counting approach to examine all of them at once.