

# 1 APPENDIX: Selecting the number of word clusters

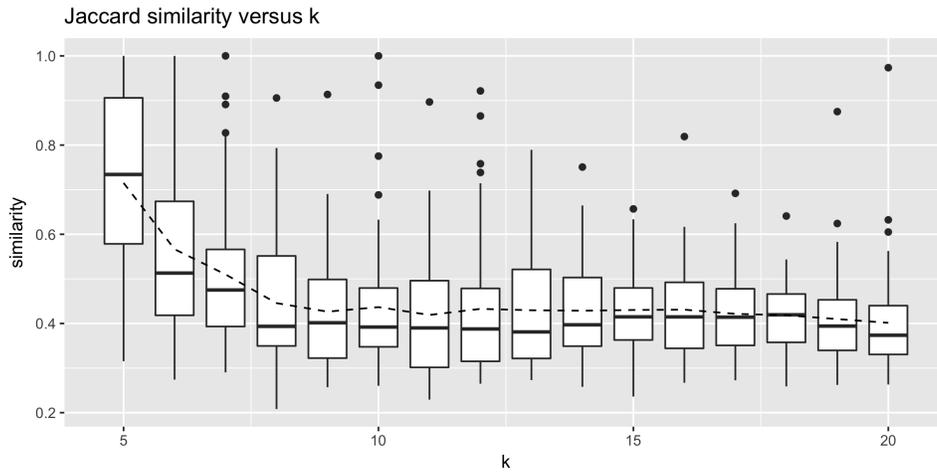


Figure 1: Average pairwise Jaccard Similarity between 100 90% subsamples of the set of word vectors.

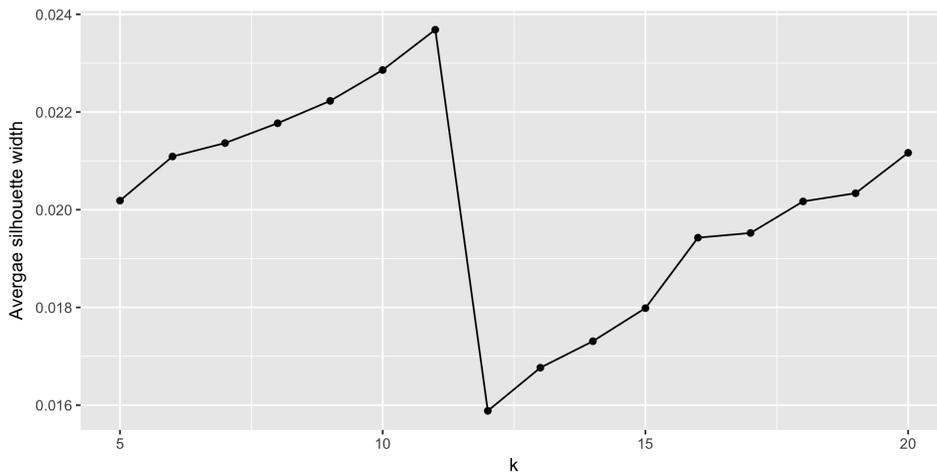


Figure 2: Average Silhouette width based on 100 90% subsamples of the set of word vectors.

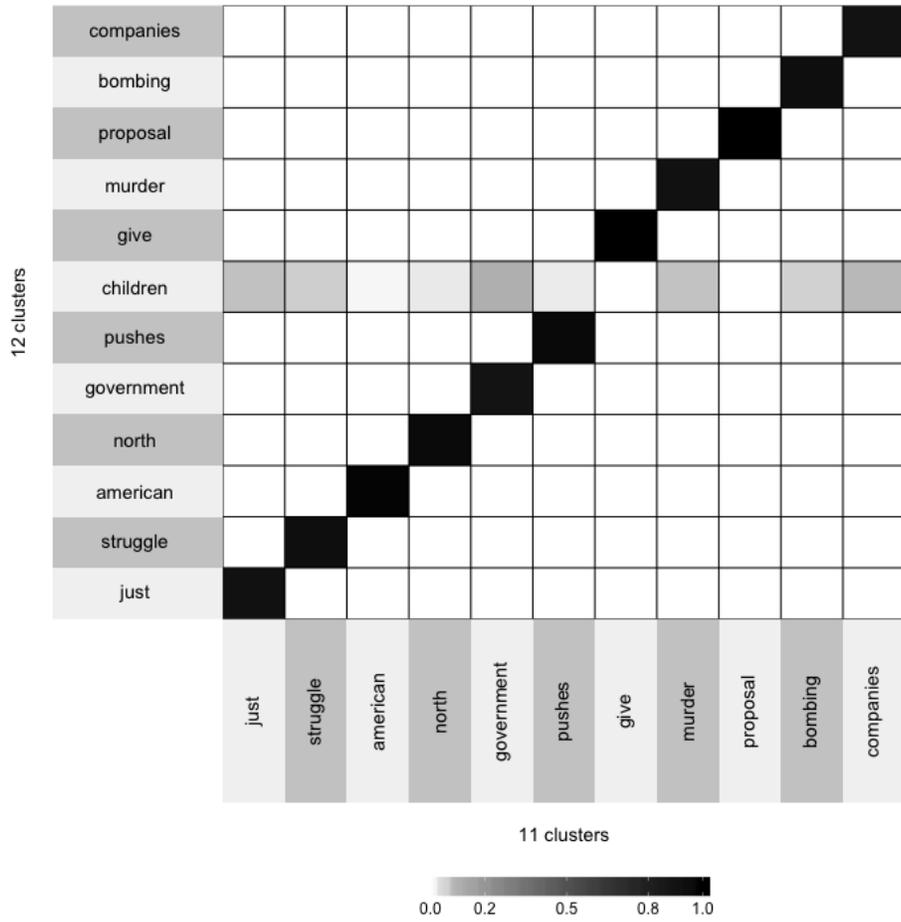


Figure 3: The proportion of words that are common in each cluster when we change  $k$  from 11 to 12.