

Problem Set 4: Quanteda Basics und Manifestodaten

Mit den Manifesto Corpus Daten des Videos:

1. Keywords in Context Analyse mit einem „Window“ von 10 Wörtern: Wie wird über „Ausländer“ in den Parteiprogrammen?
 - a. Wie wird der Begriff „Ausländer“ diskutiert? Gibt es Unterschiede zwischen den Parteien?
 - b. Wie werden Begriffe, die mit dem Begriff „Ausländer“ beginnen diskutiert? Gibt es Unterschiede zwischen den Parteien?
 - c. Für die Ergebnisse aus der vorherigen Frage: Sucht euch 5 Sätze aus und extrahiert diese vollständig.

2. Worthäufigkeiten: Welche Worte werden von der FDP genutzt?

Die Analyse sollte ohne Stopwörter und nur mit den Wortstämmen stattfinden. Ihr könnt auch andere Aspekte ausschließen. Reflektiert warum ihr Aspekte ein oder ausschließt.

- a. Findet die 20 meistgenutzten Worte, die von der FDP genutzt werden.
 - b. Stellt diese Worte graphisch dar

3. Co-Occurance Analyse: Welche Worte werden von der AfD in Verbindung genutzt?

Die Analyse sollte ohne Stopwörter und nur mit den Wortstämmen stattfinden. Ihr könnt auch andere Aspekte ausschließen. Reflektiert warum ihr Aspekte ein oder ausschließt

- a. Findet die 20 meistgenutzten Wortpaare, die von der AfD auf Satzebene genutzt werden. (Der Befehl `broom::tidy()` kann euch dabei helfen)

4. Kombinationsaufgabe: Welche Worte werden in Dokumenten (Sätzen), die den Wortstamm „Ausländer“ beinhalten häufig auf?

Die Analyse sollte ohne Stopwörter stattfinden. Ihr könnt auch andere Aspekte ausschließen. Reflektiert warum ihr Aspekte ein oder ausschließt.

- a. Identifiziert welche Dokumente / Sätze den Wortstamm „Ausländer“ beinhalten.
 - b. Wählt die relevanten Dokumente aus dem Corpus aus, und führt eine Text-Frequenz Analyse der 10 meistgenutzten Begriffe durch.