

Übungsaufgabe 2

R Tutorium - Lisa Poggel

Zum 16.5.2022

1. Erstellt ein neues R Skript und speichert es an einem geeigneten Ort. Wiederholt die Schritte aus dem Skript der letzten Stunde mit unserem Märchenkopus.
2. Sucht nach allen Märchen, in denen eine Person Namens “Droßelbart” vorkommt. Hinweis: Märchentext in der Spalte “Text” durchsuchen.
3. Sucht in allen Märchen nach allen Wortkombinationen mit dem Wort “Tochter”. Hinweis: Reguläre Ausdrücke können auch nach Leerzeichen suchen. Das Leerzeichen einfach mit der Tastatur eingeben.
4. Entfernt alle Vorkommnisse aller Formen des unbestimmten Artikels “ein” aus dem Märchen Aschenputtel und speichert den bearbeiteten Märchentext in einer Variable “aschenputtel_bearb”. Hinweis: Groß- und Kleinschreibung beachten!
5. Entfernt nun außerdem das erste Vorkommnis des Namens “Aschenputtel” aus “aschenputtel_bearb”. Speichert den bearbeiteten Text zuletzt in einer neuen Datei “aschenputtel_bearbeitet.txt”
6. Entfernt aus allen Märchen alle Seitenzahlen und speichert alle Märchen in einem neuen Ordner “corpus_cleaned”. Seitenzahlen haben die Form [irgendwelche Zahlen]. Hinweis: In der Übung zur Sitzung 2 haben wir gelernt, wie man einen neuen Ordner erstellt und Dateien schreibt. Teilnehmer*innen des Seminars “Digitale Dramenanalyse” können für diese Aufgabe außerdem auf die Unterlagen zur vorletzten Stunde zurückgreifen.
7. Informiert euch auf dieser Website über das Paket **stringr**. Wiederholt die Übung mit passenden **stringr**-Funktionen. Arbeitet dabei nach Möglichkeit auch mit Zeichenklassen und probiert Lookarounds aus (siehe Rmd. beziehungsweise html-Datei zur letzten Stunde)