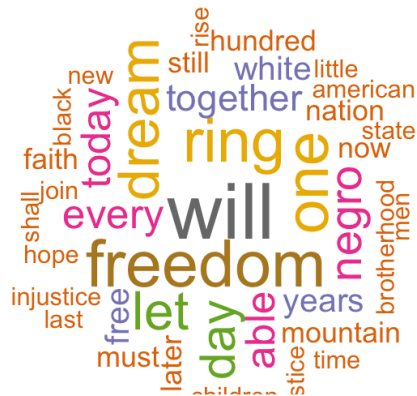


Assignment #6: Web scraping



2. One can organize data using Wordcloud in R. First, one can translate the text data into a data frame for visualization. Then, one can use the function “htmlTreeParse” to read and analyze the text data in paragraphs. After doing that, one can create a corpus for preprocessing, and then remove unnecessary characters from the data frame, like punctuation marks. Last, using Wordcloud, one can organize the data frame by frequency, repetition, etc.

3. Challenge: One can download multiple files using Python and create folders for downloaded content. One can use a download function and call a module titled “requests” to determine which files, if any, had issues downloading.