빅데이터 텀프로젝트

201622127 전창준

//1.소스코드

install.packages("KoNLP")

library(KoNLP)

install.packages("RColorBrewer")

library(RColorBrewer)

install.packages("wordcloud")

library(wordcloud)

pal2 <- brewer.pal(8, "Dark2")

text <- readLines(file.choose())

text

noun <- sapply(text, extractNoun, USE.NAMES=F)

noun

noun2 <- unlist(noun)

noun2

word\_count <- table(noun2)

word\_count

head(sort(word\_count, decreasing=TRUE), 10)

noun2 <- gsub("and", "", noun2)

noun2 <- gsub("of", "", noun2)

noun2 <- gsub("to", "", noun2)

noun2 <- gsub("their", "", noun2)

noun2 <- gsub("the", "", noun2)on

noun2 <- gsub("for", "", noun2)

noun2 <- gsub("by", "", noun2)

noun2 <- gsub("He", "", noun2)

noun2 <- gsub("has", "", noun2)

noun2 <- gsub("in", "", noun2)

noun2 <- gsub("an", "", noun2)

noun2 <- gsub("or", "", noun2)

noun2 <- gsub("be", "", noun2)

noun2 <- gsub("with", "", noun2)

noun2 <- gsub("on", "", noun2)

noun2 <- gsub("a", "", noun2)

noun2 <- gsub("his", "", noun2)

word\_count <-table(noun2)

wordcloud(names(word\_count), freq=word\_count, scale=c(20,1.2), min.freq=3, random.order=F, rot.per=.1, colors=pal2)

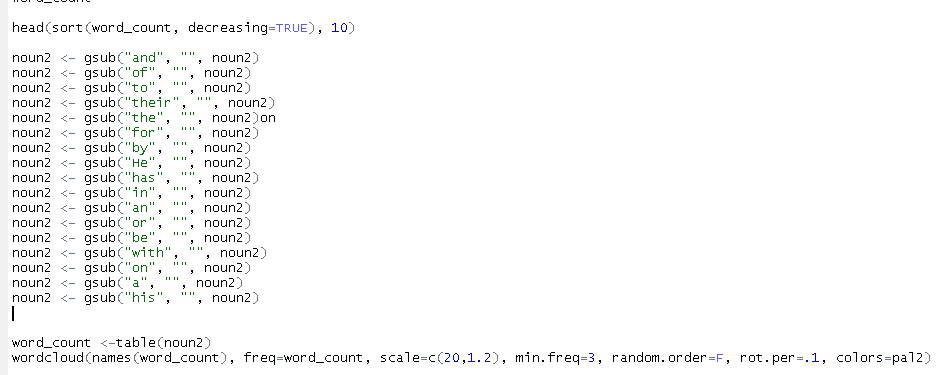
2. 제가 이번에 한 공공데이터 자료는 연설문이며, 공공데이터를 분석하는 것은 1차세계대전 이후의 크고작은 전쟁들을 승리로 이끌었다는 평가를 얻는 명연설에는 무슨 공통점이 있고, 무슨 차이점이 있는지에 대해서 살펴보려고 합니다.

그리고 이것들이 무엇을 시사하는 지에 대해서도 알아보고자 합니다.

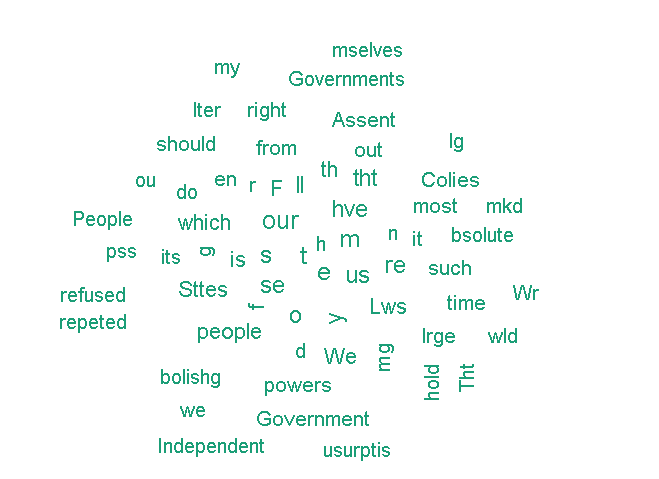
[[1]](#footnote-1)제가 사용한 데이터는 2차세계대전을 승리로 이끌었다고 평가 받는 윈스턴 처칠, [[2]](#footnote-2)포클랜드 전쟁을 승리로 이끈 마가렛 데처, [[3]](#footnote-3)연설문으로써 독일의 통일 발판을 이끌어낸 로널드 레이건, [[4]](#footnote-4)영국으로부터 독립하기 위해 쓴 미국의 독립선언서, [[5]](#footnote-5)걸프 전쟁을 성공적으로 이끌어낸 아버지 부시 그리고 [[6]](#footnote-6)일본의 진주만 기습으로 의회를 설득하는 연설문을 쓴 프랭클린 루즈벨트입니다.

3. 우선 전치사 같은 글자들을 거르기 위해 다음과 같은 방법을 사용하였고,

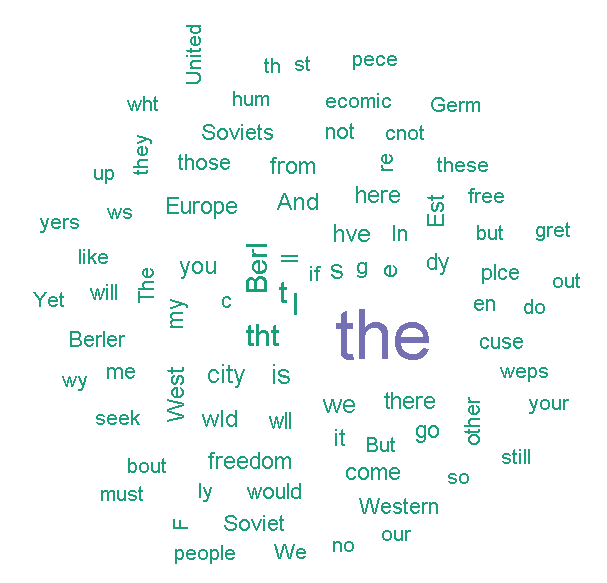
연설문이 짧거나 중복되는 단어가 별로 없을 경우 빈도수를 줄였습니다.



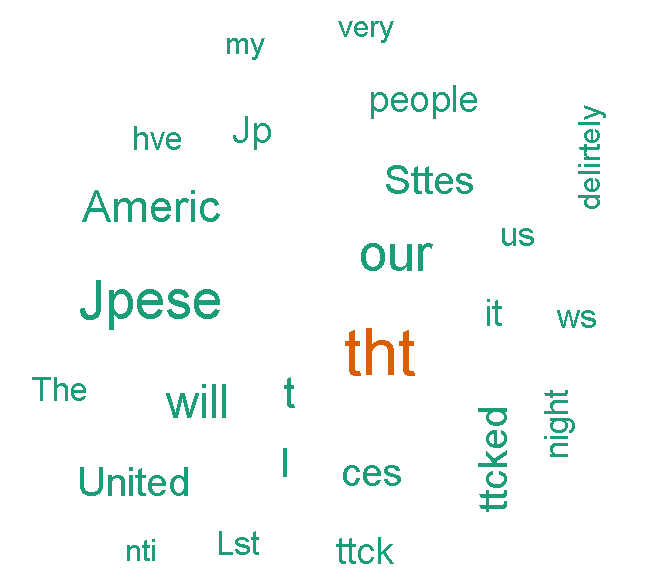
먼저 독립선언문을 보도록 하겠습니다.



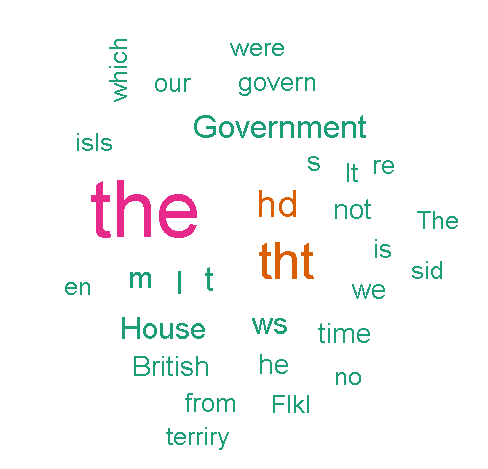
독립선언문에서 볼 수 있는 단어로는 국민, 정부, 독립, 식민지, 승인, 거절, 힘, 완벽한 같은 뜻의 단어가 보이며 그중에서도 특히 정부라는 단어가 많이 보이는 것이 눈에 띕니다.



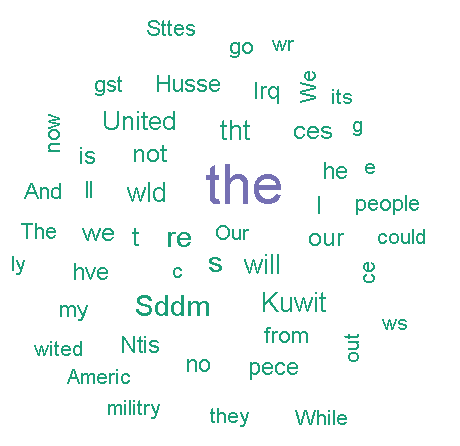
그 다음은 레이건의 연설문인데 자유, 소련, 서방, 동방, 경제, 독일, 통일, 유럽, 베를린, 자유, 무기, 시민과 같은 단어들이 많이 쓰였습니다.



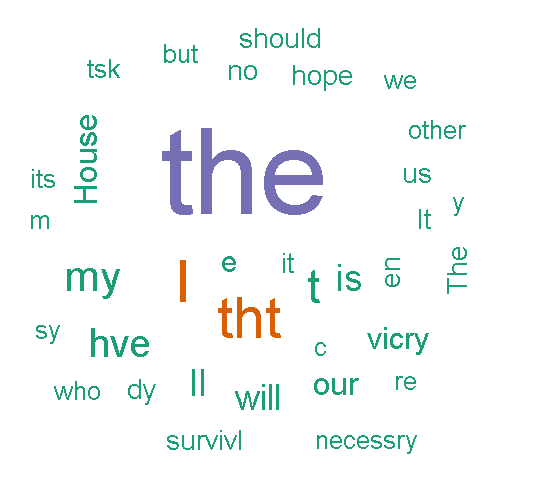
루즈벨트의 의회 연설에서는 일본, 국민, 미국, 침략받다와 같은 단어가 많이 쓰였으며, 특히 공격받았다는 단어가 많이 쓰인 것이 보입니다.



마가렛 대처의 연설에서는 영토, 정부, 영국, 의회(house는 영국 상원 하원을 뜻합니다)과 같은 단어사 사용되었습니다.



아버지 부시의 연설에서는 미국, 이라크, 사담후세인, 쿠웨이트, 군사적인, 평화와 같은 단어들이 많이 쓰였는데 그중에서는 사담 후세인, 쿠웨이트가 독보적입니다.



윈스턴 처칠의 연실에서는 의회, 승리, 필요, 생존, 희망과 같은 단어가 쓰인 것을 확인할 수 있었습니다.

4. 결론

영국 수상들 같은 경우에는 house 즉 의회의 단어 선택 비중이 높았다는 것을 볼 수 있었으며, 자신의 국토가 공격받았을 경우 국토가 공격받았다는 사실에 대한 부분이 많이 이입이되어 있으며, 2차세계대전 이후 슈퍼파워가 된 미국의 경우에는 자신의 영토가 공격받지는 않았음에도 불구하고 처리하는 모습을 볼 수 있었습니다. 이것으로 보아 연설문은 시대의 상황에 따라 바뀌지만 공격받았을 경우 단결을 유도하는 연설문을 작성한다는 것을 알았습니다.

1. "Blood, Toil, Tears and Sweat," *https://winstonchurchill.org*, n.d. 수정, 2018년12월17일 접속, https://winstonchurchill.org/resources/speeches/1940-the-finest-hour/blood-toil-tears-and-sweat-2/. [↑](#footnote-ref-1)
2. "Falkland Islands," *https://www.margaretthatcher.org/*, n.d. 수정, 2018년12월17일 접속, https://www.margaretthatcher.org/document/104910. [↑](#footnote-ref-2)
3. "Tear down this wall," *http://www.historyplace.com/*, n.d. 수정, 2018년12월17일 접속, http://www.historyplace.com/speeches/reagan-tear-down.htm. [↑](#footnote-ref-3)
4. "Declaration of Independence," *https://www.constitution.org*, n.d. 수정, 2018년12월17일 접속, https://www.constitution.org/us\_doi.pdf. [↑](#footnote-ref-4)
5. "Announcing War aganist Iraq," *http://www.historyplace.com/*, n.d. 수정, 2018년12월17일 접속, http://www.historyplace.com/speeches/bush-war.htm. [↑](#footnote-ref-5)
6. "day of infamy speech," *http://www.awon.org/*, n.d. 수정, 2018년12월17일 접속, http://www.awon.org/dc3/. [↑](#footnote-ref-6)