성공한 기업인의 졸업식 연사 텍스트 분석

20120863 박정현

주제 소개 및 선정 이유: 세계적으로 성공한 기업인들의 대학교 졸업식 연사 텍스트를 분석해 공통점, 차이점, 특이점 등을 발견하고자 함. 이를 기반으로 성공한 기업인들의 인사이트를 배울 수 있기를 희망.

기업인 및 텍스트 소개:

주요 분석 인물 및 텍스트

1. 빌 게이츠 (William Henry Gates)

출생: 1955년 10월 28일 (미국 워싱턴 주 시애틀)

업적: 마이크로소프트 설립자이자 기업인

텍스트 소개: 2007년 6월 7일, 하버드 대학 졸업식 연설문

1. 마크 주커버그 (Mark Elliot Zuckerberg)

출생: 1984년 5월 14일 (미국 뉴욕 주 화이트플레인스)

업적: 페이스북 설립자 및 현 CEO

텍스트 소개: 2017년 5월 25일, 하버드 대학 졸업식 연설문

1. 스티브 잡스 (Steven Paul Jobs)

출생: 1955년 2월 24일 (미국 캘리포니아 주 샌프란시스코)

업적: 애플 창시자, 픽사 애니메이션 창시자

텍스트 소개: 2005년 6월 12일, 스탠포드 대학 졸업식 연설문

1. 마윈 (馬雲)

출생: 1964년 9월 10일 (중국 저장 성 항저우 시)

업적: 알리바바 그룹 창시자 및 회장

텍스트 소개: 2013년 8월 11일, 홍콩 과기대학 졸업식 연설문

부가적 분석 인물 및 텍스트 (비교 대상)

1. 셰릴 샌드버그 (Sheryl Sandberg)

출생: 1969년 8월 28일 (미국 워싱턴 D.C.)

업적: 현 페이스북 최고 운영 책임자 (COO), 페이스북 최초 이사회 여성 임원

텍스트 소개: 2016년 5월 14일, 캘리포니아 버클리 대학 졸업식 연설문

1. 버락 오바마 (Barack Hussein Obama)

출생: 1961년 8월 4일 (미국 하와이 주 호놀룰루)

업적: 미국 제44대 대통령

텍스트 소개: 2016년 5월 7일, 미국 하워드 대학 졸업식 연설문

1. 윌리엄 맥레이븐 (William Harry McRaven)

출생: 1955년 11월 6일 (미국 노스캐롤라이나 주 파인허스트)

업적: 전 미국 특수작전사령관

텍스트 소개: 2014년 5월 17일, 미국 텍사스 대학 졸업식 연설문

분석 방법:

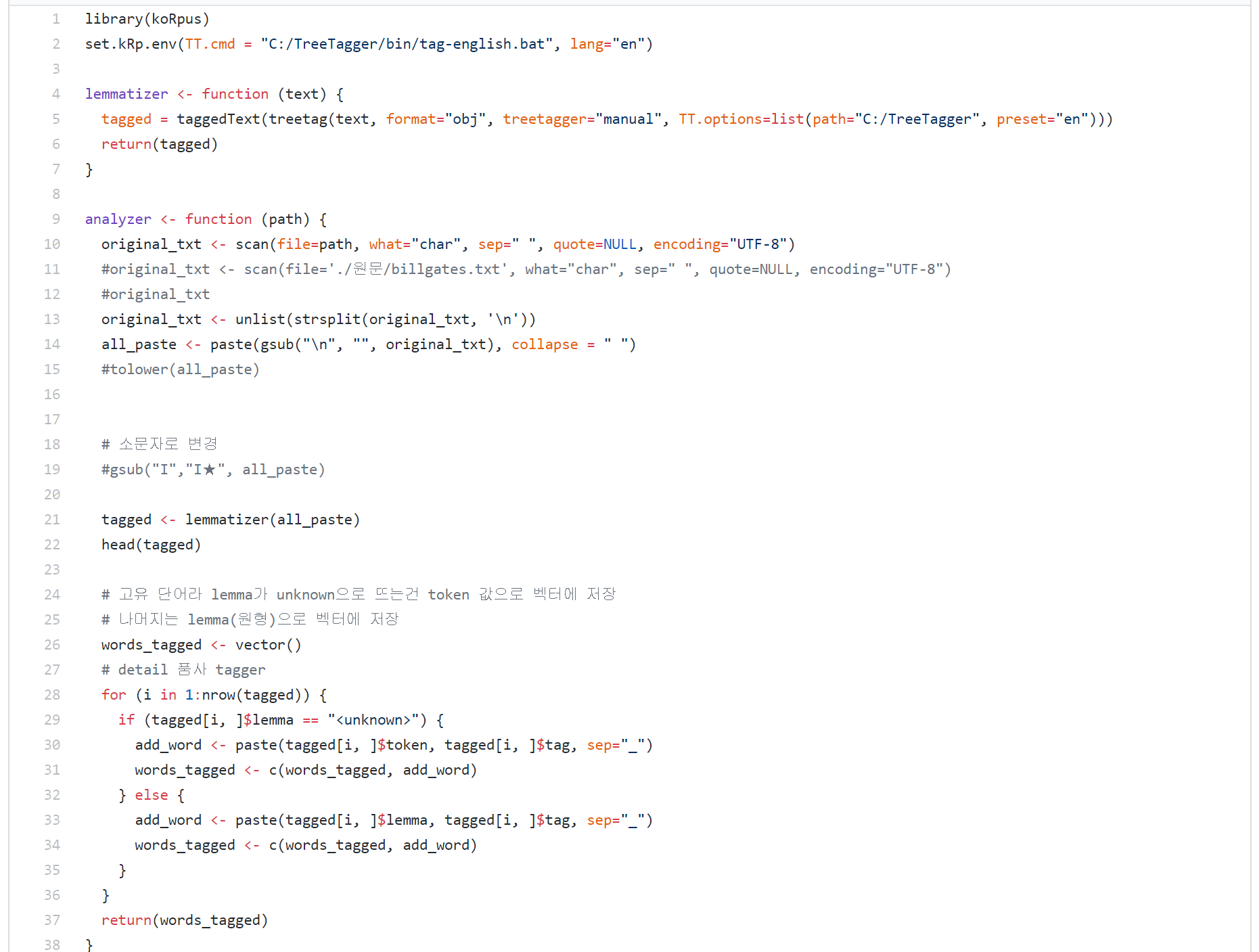
빈도분석, n-gram, 키워드부분석 ~~

1. 빈도분석

분석 방법:

1 – lemmatizer을 이용해 텍스트의 모든 단어에 형태소 태그를 단다.

\*analyzer2.R 파일 소스코드



\*출력 예시

> head(analyzer(‘billgates.txt’))

[1] “President\_NP” “Bok\_NP” “,\_,” “former\_JJ” “President\_NP” “Rudenstine\_NP”

2 – 형태소 태그가 달려있는 문자열 벡터에서 필요한 품사만 grep으로 뽑아내 데이터 프레임으로 변경한다. 이후 워드클라우드로 출력한다.

\* 빌게이츠 텍스트 분석하는 프로그램 소스코드 (이하 모든 프로그램에서 텍스트 원문 불러오는 부분 제외하고 모두 동일)

- 내용어를 구성하는 명사(\_N\*\*), 동사(\_V\*\*, have와 be 동사 제외), 형용사(JJ), 부사(RB)를 추출



\*출력 예시

> head(Freq.tag, 5)

Freq

people\_NNS 26

world\_NN 22

Harvard\_NP 20

n’t\_RB 18

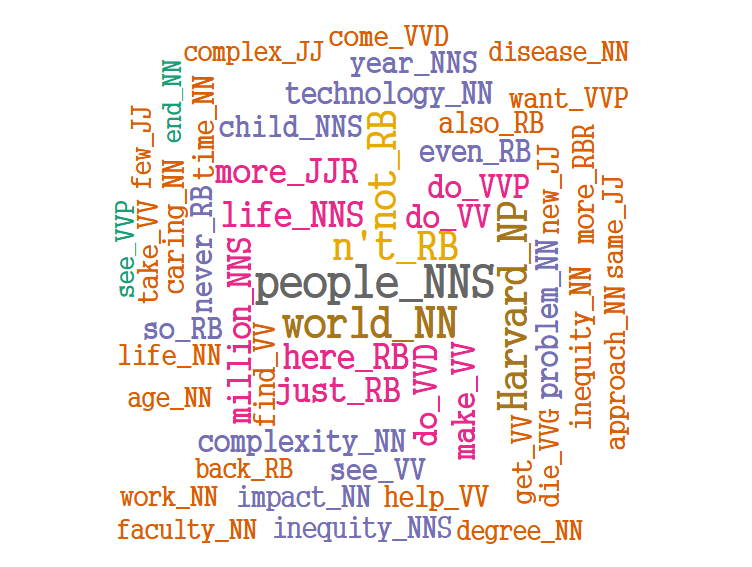
not\_RB 17

> makeWordcloud(Freq.tag)

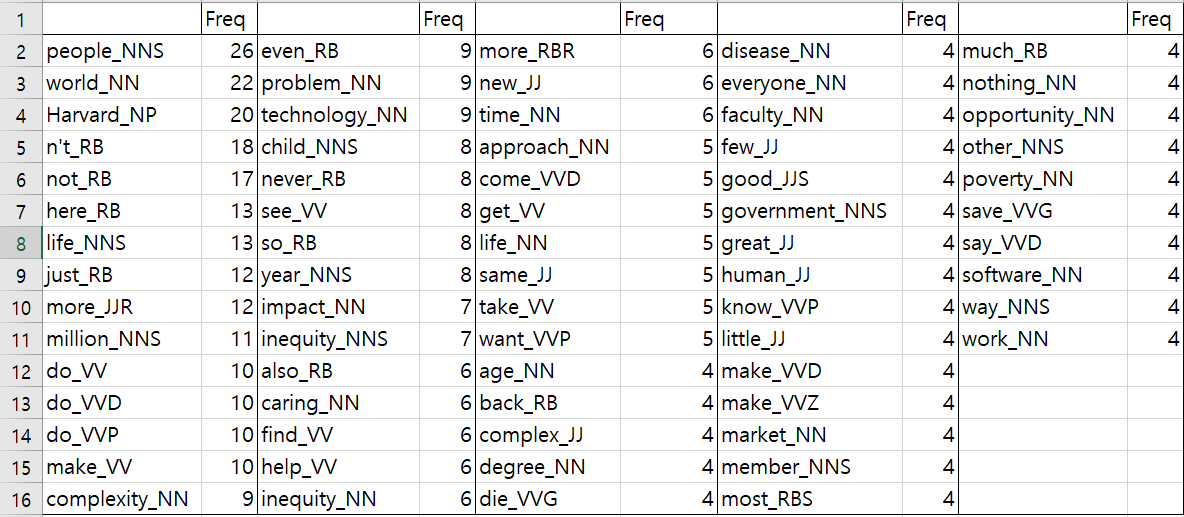


1) 빌게이츠 텍스트 빈도분석

> 빌게이츠 연설문의 워드클라우드



> 빈도 수 4회 이상의 단어들



> 해석

1, 2번 째로 나온 people와 world는 영어에서 자주 쓰이는 명사 Top5에 포함되고, (<http://www.talkenglish.com/vocabulary/top-1500-nouns.aspx>) 4, 5번 째로 나온 n’t와 not 또한 가장 자주 쓰이는 부사 Top1이기 때문에 이를 고려해야 한다. 즉, 이러한 단어들은 단순 절대 빈도분석 보다는 다른 텍스트들과의 상대 빈도 분석을 할 때 의미를 찾을 수 있을 것이라 생각한다. (이후 분석에서 비교)

또한 here, just, more, million, make 등의 단어들은 문맥을 알지 못하면 단어 자체만으로 의미를 많이 찾을 수 없기에 이 단어들은 연어 또는 n-gram 등 문맥을 함께 분석하는 분석에서 상대적으로 더 유의미할 것이다.

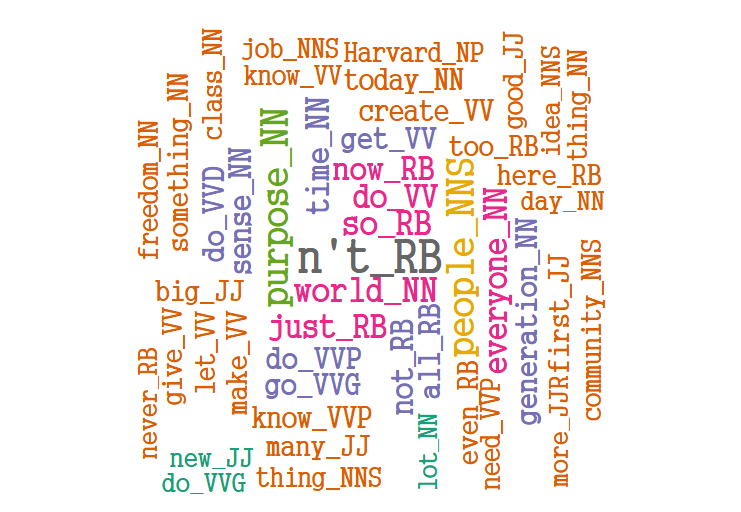
이를 제외하면 Harvard, life, inequity (NN + NNS를 합치면 빈도수 총 13회로 6순위), complexity, problem, technology, child 을 분석 대상으로 뽑을 수 있다.

하버드 대학의 졸업식 연설이고, 빌게이츠도 자퇴를 하긴 했지만 하버드 대학 출신이었기 때문에 대학교와 관련된 일화 등을 많이 이야기 한 것으로 볼 수 있다. 그리고 life라는 단어를 자주 언급함으로써 연설문 전체가 삶의 방향에 대한 이야기임을 알 수 있다. 그리고 그 방향에 inequity, complexity, problem 등의 단어를 써서 불공평한 사회 문제와 연관 짓고 있다. 그리고 이를 technology를 통해서 해결할 수 있다고 말한다. 이후 n-gram 분석에서 확인하겠지만 불공평한 사회 문제를 더욱 부각하기 위해 어린 아이들(child)이 죽어 나가는 (die, disease 빈도수 4회) 예시를 많이 사용했다. 이외에도 빈도수는 4회로 상대적으로 낮지만 caring, help, poverty 등의 단어들을 사용한 것을 보면 이러한 해석이 설득력 있다고 볼 수 있다.

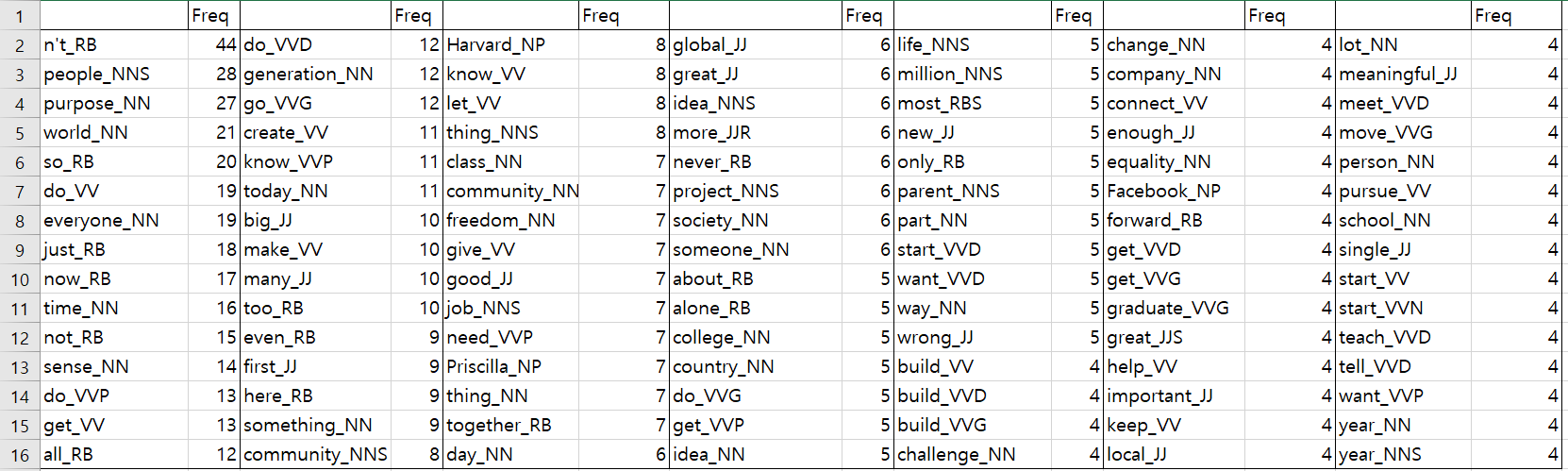
즉, 빌게이츠는 동문인 하버드 대학 졸업생들에게 세상의 불공평함을 인지하고 이를 해결하기 위한 삶을 살아 달라는 연설을 한 것으로 해석할 수 있다.

2) 마크 주커버그 텍스트 빈도분석

> 마크 주커버그 연설문의 워드클라우드



> 빈도 수 4회 이상의 단어들



> 해석

빌게이츠 연설문과 마찬가지로 주커버그의 연설문에서도 1,2,4 번 째로 많이 나온 n’t, people, world는 원래 영어에서 빈도수가 높기 때문에 이후 상대 빈도 분석을 통해 파악해야 할 것이다.

단순 절대 빈도에서는 purpose, everyone, sense, all, generation, idea (NN + NNS 합쳐 11회), community, Harvard, freedom, Priscilla, together, global, society 등을 의미 있게 해석해 볼 수 있겠다. 주커버그는 일반적인 단어는 아닌 purpose라는 단어를 압도적으로 많이 사용함으로써 ‘목적 있는 삶’을 강조했다. 이때의 목적은 단순히 자신의 목적이 아닌, ‘모든 사람들이 (everyone, all, together) 목적을 가진 삶’을 의미한다. 즉, 모든 사람들이 목적을 가진 삶을 살 수 있도록 더 나은 세상을 만들자는 이야기를 하고 있다.

이를 달성하기 위해 주커버그는 ‘공동체’를 강조한다. (community, together, global, society 등) 재미있는 것은 사실 주커버그는 자신의 회사인 ‘Facebook’을 소개할 때도 회사의 미션이 전 세계 공동체를 하나로 묶는 것이라고 자주 이야기 한다는 사실이다. 즉, 대놓고 드러내지는 않았지만 위에서 말한 세상을 위해선 페이스북이 반드시 필요하다는 얘기를 하고 있다고도 볼 수 있다.

그 외에도 고유명사이자 자신의 아내인 Priscilla를 7회나 언급하면서 아내에 대한 사랑을 드러내기도 했고 주커버그도 빌게이츠와 마찬가지로 Harvard 대학 자퇴생으로서 Harvard 대학 졸업 연설을 했다.

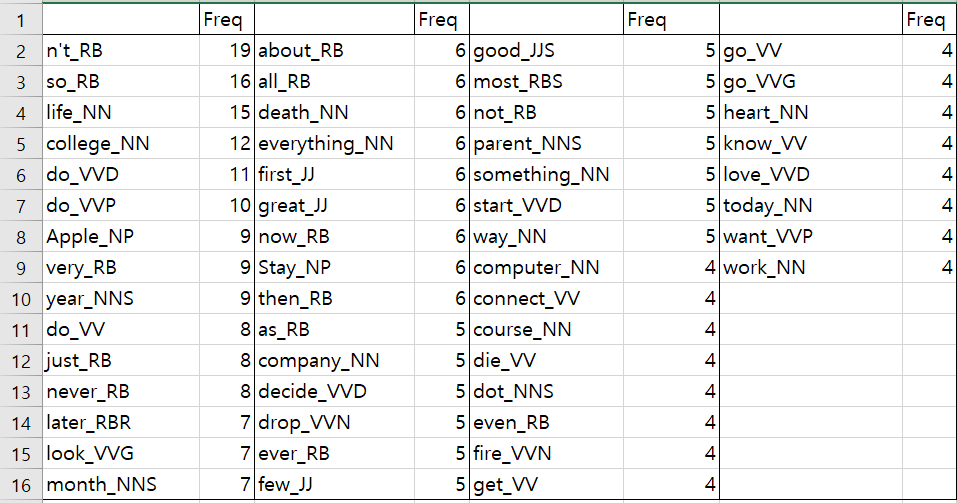
추가적으로 주커버그가 최근 미국 대통령을 준비하고 있다는 소문이 돌고 있는데, 이를 알고 다시 한번 연설문을 보면 주커버그의 연설문은 기업가가 아닌 대통령이 할 법한 연설을 하고 있다고도 볼 수 있다. 재미있는 것은 우연이겠지만 6명의 기업가들 중에서 주커버그의 연설문만 압도적으로 긴데, (4개 연설문 평균 단어 수 2,822개 / 주커버그 연설문 단어 수 4,150개) 전 미국 대통령인 오바마의 연설문도 총 5,998 단어로 대학교 졸업 연설문 중에 긴 편에 속했다.

3) 스티브 잡스 텍스트 빈도분석

> 스티브 잡스 연설문의 워드클라우드



> 빈도수 4회 이상의 단어들



> 해석

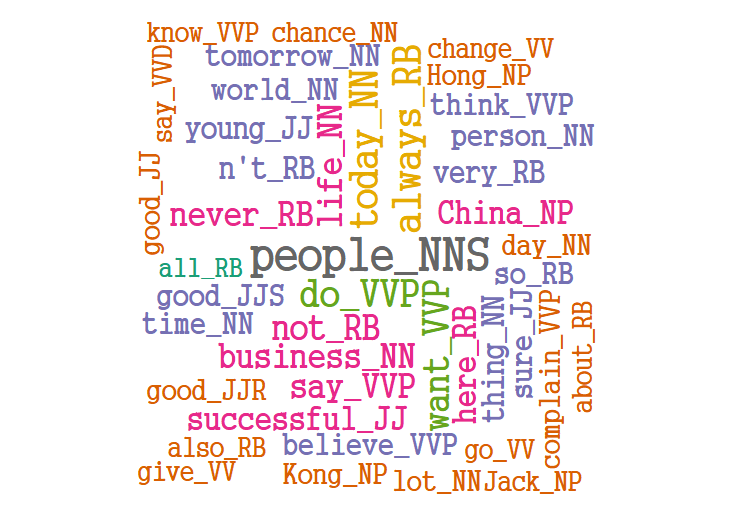
1,2 번 째에 나온 n’t와 so는 원래 영어에서 높은 빈도수를 보이므로 절대 빈도 분석에서는 해석하지 않겠다. 하지만 앞서 분석한 빌게이츠, 주커버그의 연설문에서 주로 쓰인 people와 world는 스티브 잡스의 연설문에서는 거의 보이지 않았다. (people 2회, world 3회) 원래 영어에서 높은 빈도수를 보이는 단어 임에도 거의 쓰지 않은 것은 해석해 볼 여지가 있다. (이후 분석)

여기서는 life, college, death, Stay 등을 뽑아볼 수 있겠다. 스티브 잡스도 빌게이츠처럼 삶(life)에 대한 이야기를 주로 했는데, 삶의 방향성을 제시한 빌게이츠와는 다르게 잡스는 자신의 삶에 대한 이야기만 했다. 삶에 대한 3가지 이야기를 했는데 이 중 첫 번째 이야기는 주로 자신의 대학시절(college) 이야기였다. 여기서 스티브 잡스도 빌게이츠/주커버그와 마찬가지로 대학 중퇴자이지만 둘과는 다르게 대학에서 배운 것들에 대해 많이 이야기를 했다. 또 그는 다른 기업가들이 거의 언급하지 않은 death를 많이 언급했는데, 이는 그의 세 번째 이야기가 죽음과 관련한 이야기이기 때문이다. 그리고 대문자로 시작하는 명사 ‘Stay’가 6회나 언급된 것은 주목할 만한데, 이 연설에서 바로 스티브 잡스의 모토로 유명한 ‘Stay Hungry, Stay Foolish’가 유명해졌기 때문이다.

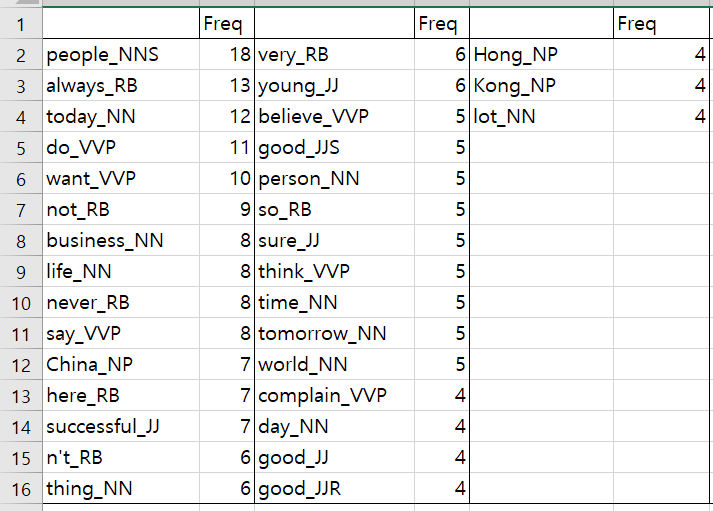
스티브 잡스의 연설문은 절대 분석만으로는 특이점을 찾기 어려웠다. 따라서 차후에 다각적인 분석을 통해 더 깊게 살펴보고자 한다.

4) 마윈 텍스트 빈도분석

> 마윈 연설문의 워드클라우드



> 빈도 수 4회 이상의 단어들

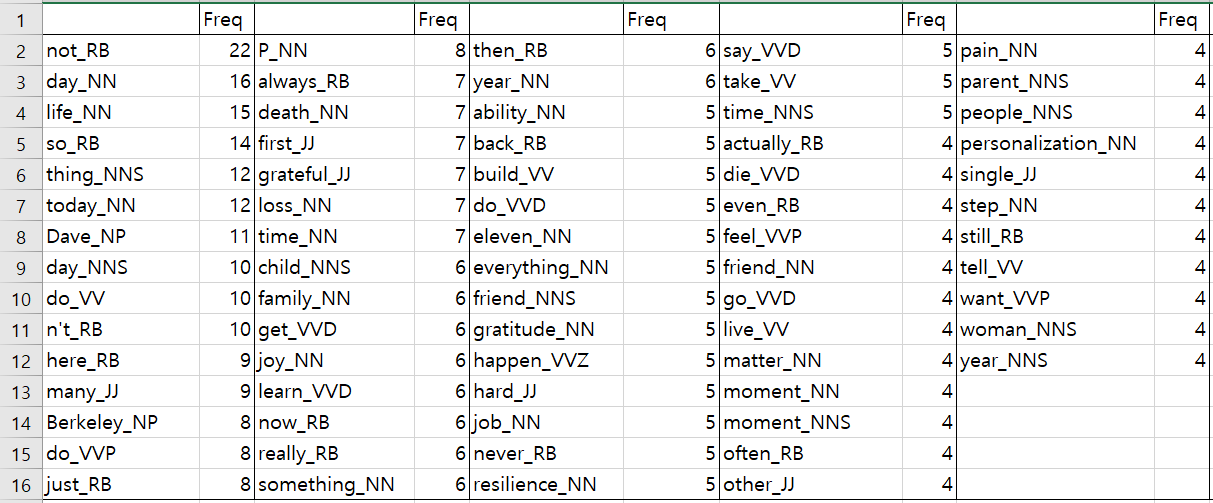


> 해석

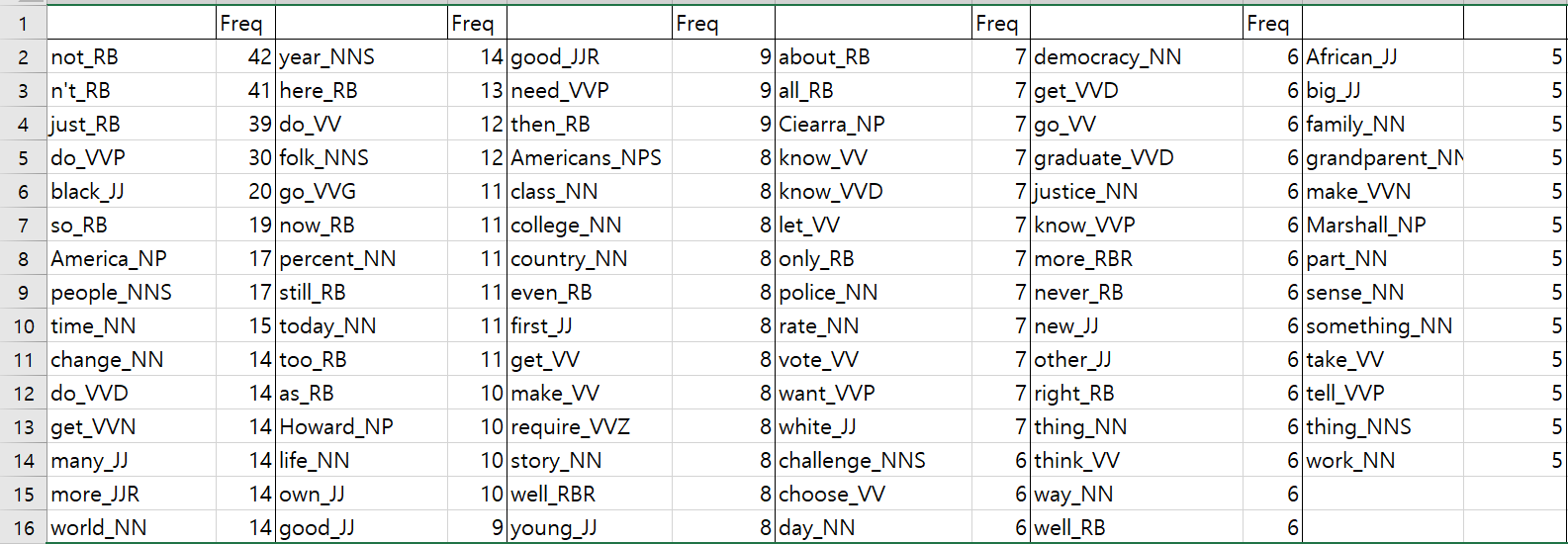
우선 마윈의 연설문은 4개의 연설문 중 가장 짧은 연설문(1,201 단어)이다. 또한 마윈은 앞서 언급한 기업인들과 다르게 중국 출신으로 영어가 모국어가 아니다. 이러한 사항들을 고려해서 해석해야 한다.

7) 셰릴 샌드버그, 버락 오바마, 윌리엄 맥레이븐 텍스트 빈도분석

\* 셰릴 샌드버그 (4회 이상)



\* 버락 오바마 (5회 이상)



\* 윌리엄 맥레이븐 (4회 이상)

