

KUBIC 자료분석 메뉴얼

로그인

[자료열람](#)[자료분석](#)[커뮤니티](#)[홈페이지소개](#)[Open API](#)[로그인](#)[회원가입](#)[검색조건](#)[기간](#)[기관별](#)[주제별](#)[검색도움말](#)



구글 계정으로 로그인

[검색조건](#)[기간](#)[기관별](#)[주제별](#)[검색도움말](#)

자료검색 및 자료담기



원하는 키워드로
자료검색!

🔍 통일

검색조건

기간

기관별

주제별

검색도움말

홈 > 검색결과

키워드분석 ?

검색필터

결과 내 재검색

키워드 포함

키워드 제외

기간

1일

1주

1개월

3개월

6개월

1년

전체

직접입력

확인

연도-월-일

연도-월-일

"통일"에 대한 검색결과는 4,163건 입니다.

최신순

10건씩 보기

☐ 전체선택

☐ 새 통일 한국의 영.유아 교육 연구

1. 문제제기2. 새 통일 한국 교육의 방향3. 새 통일 한국의 인간관, 영.유아교육관4. 새 통일 한국의 영.유아...

1998 | 임명희

연관문서

☐ 민간 통일 운동의 주요 논의 동향과 통일 정책 수용여부에 관한 연구

1. 서론2. 한국의 민간 통일 운동 배경에 관한 분석틀3. 통일 운동의 성격의 이해4. 통일 운동의 비판적 인...

1997 | 박찬석

📁 담기

검색필터

결과 내 재검색

키워드 포함

키워드 제외



기간

1일

1주

1개월

3개월

6개월

1년

전체

직접입력

확인

연도-월-일



연도-월-일



기관별

통일연구원 1994

통일부 1314

남북하나재단 488

국회 외교통일위원회 163

평화와 통일을 여는 사람들 145

평화를만드는여성회 59

주제별

전체

정치

경제

사회

문화

국제

☒ 전체선택

☒ 새 통일 한국의 영.유아 교육 연구

1. 문제제기2. 새 통일 한국 교육의 방향3. 새 통일 한국의 인간관, 영.유아교육관4. 새 통일 한국의 영.유아...

1998 | 임명희

연관문서

☒ 민간 통일 운동의 주요 논의 동향과 통일 정책 수용여부에 관한 연구

1. 서론2. 한국의 민간 통일 운동 배경에 관한 분석틀3. 통일 운동의 성격의 이해4. 통일 운동의 비판적 인...

1997 | 박찬석

연관문서

☒ 한반도 통일의 미래와 주변 4국의 기대

머리글 요약 Abstract 제1장 서론 I. 연구의 필요성 및 목적II. 연구의 방법III. 연구의 결과

2013 | 김규환

연관문서

☒ 알기쉬운 통일교육 12주제:프리젠테이션-제1부-통일비전

제1부 통일비전제1강 통일 필요성 /2제2강 통일비용과 통일편익 /3제3강 통일비전과 과제 /42

내 보관함에
자료 담기

담기

자료분석 - 메뉴얼



자료분석 페이지의
기본 매뉴얼을
확인할 수 있습니다.

[기관별](#)[주제별](#)[검색도움말](#)[홈](#) > [자료분석](#) > [매뉴얼](#)

매뉴얼

[매뉴얼](#)[전처리](#)[분석](#)

- 자료분석 프로세스

데이터 수집

원하는 검색어로
자료 검색

내 보관함에
자료 담기

원하는 자료
선택

데이터 전처리

전처리 실행

품사 선택

사용자 사전
적용

- 동사, 명사, 형용사

- 불용어, 유의어, 복합어

↓
데이터 분석

[검색조건](#)[기간](#)[기관별](#)[주제별](#)[검색도움말](#)[홈](#) > [자료분석](#) > [매뉴얼](#)

매뉴얼

[매뉴얼](#)[전처리](#)[분석](#)

- 내 보관함: 검색한 키워드에 따라 수집된 문서들이 저장되며, 원하는 키워드의 문서들을 선택하여 분석을 진행할 수 있습니다.
- 사용자 사전: 기본사전과 내 사용자 사전이 있습니다. 기본 사전은 시스템 내에 기본적으로 등록된 사전이며, 내 사용자 사전은 자신이 원하는 단어들을 사전 형식에 맞게 작성하고 csv파일로 업로드하여 전처리를 진행 할 수 있습니다.
 1. 유의어사전: 같은 의미의 다른 표현들을 대표단어로 변경합니다. 제일 첫번째로 작성한 단어가 대표어가 됩니다.
ex) 김정은 위원장, 김정은, 김위원장 → 첫번째 단어인 김정은 위원장으로 단어가 변경됩니다.
 2. 불용어사전: 분석 시 필요 없거나, 의미 없는 단어들을 불용어 사전에 등록해, 그 단어는 전처리 과정에서 제외할 수 있습니다.
대체로 조사, 접속사 등이 불용어가 될 수 있으며, 명사, 형용사 중에서도 제거하고 싶은 경우 불용어 사전에 등록하여 원하는 단어들로만 분석을 진행할 수 있습니다.
ex) 결국, 이것, 첨부파일
 3. 복합어사전: 신조어나 커스터 마이징한 단어들을 분석에 포함시키고 싶을 때, 복합어 사전에 그 단어들을 등록하여 분석을 진행할 수 있습니다.
ex) 김정은 위원장, 김정은 체제, 코로나
- 품사 선택: 명사, 동사, 형용사 중에 원하는 품사를 선택하여 전처리를 진행할 수 있습니다.
- 분석 종류 (마우스 올리면 각각의 분석 종류마다 정의 나오게)

[빈도수 분석](#)[TFIDF](#)[분할군집분석](#)[계층군집분석](#)[N-gram](#)[유의어 분석](#)[의미연결망](#)[토픽 모델링](#)[상세매뉴얼 다운로드](#)

자료분석 - 전처리



검색조건

기간

기관별

주제별

검색도움말

홈 > 자료분석 > 전처리

전처리

자료열람

전처리

분석

내 보관함

	검색 키워드	문서목록	수집날짜	통계	전처리
<input type="radio"/> 북한	- 북한 사람이 들려주는 북한 이야기 - 변하는 북한 변하지 않는 북한 - 북한 핵실험 이후 북한 문제 해법		2021년 07월 08일 20시 46분	총문서수:3개 평균 단어수:230개	
<input type="radio"/> 통일	- 북한 사람이 들려주는 북한 이야기 - 변하는 북한 변하지 않는 북한 - 북한 핵실험 이후 북한 문제 해법		2021년 08월 06일 20시 52분	총문서수:5개 평균 단어수:230개	

전처리

내 보관함

검색
키워드

문서 타이틀

통계
: 총 문서 수,
평균 단어 수

	검색 키워드	문서목록	수집날짜	통계	전처리
<input checked="" type="radio"/>	북한	<ul style="list-style-type: none">- 북한 사람이 들려주는 북한 이야기- 변하는 북한 변하지 않는 북한- 북한 핵실험 이후 북한 문제 해법	2021년 07월 08일 20시 46분	총문서수:3개 평균 단어수:230개	
<input type="radio"/>	통일	<ul style="list-style-type: none">- 북한 사람이 들려주는 북한 이야기- 변하는 북한 변하지 않는 북한- 북한 핵실험 이후 북한 문제 해법	2021년 08월 06일 20시 52분	총문서수:5개 평균 단어수:230개	

검색한 키워드별로
선택하여 분석을
진행할 수 있음

데이터 수집
날짜

전처리실행 버튼을
누르면, 전처리된
결과를 미리보기 &
다운로드 가능

	검색 키워드	문서목록	수집날짜	통계	전처리
<input checked="" type="radio"/>	북한	- 북한 사람이 들려주는 북한 이야기 - 변하는 북한 변하지 않는 북한 - 북한 핵실험 이후 북한 문제 해법	2021년 07월 08일 20시 46분	총문서수:3개 평균 단어수:230개	
<input type="radio"/>	통일	- 북한 사람이 들려주는 북한 이야기 - 변하는 북한 변하지 않는 북한 - 북한 핵실험 이후 북한 문제 해법	2021년 08월 06일 20시 52분	총문서수:5개 평균 단어수:230개	

전처리

- 내 사용자 사전 적용

불용어사전	내 사전 등록	사전삭제	기본사전	사용안함
유의어사전	내 사전 등록	사전삭제	기본사전	사용안함
복합어사전	내 사전 등록	사전삭제	기본사전	사용안함

기본 사전 미리보기

불용어사전 침부, 중략, 하지만, 그러나...

유의어사전 {김정은 위원장: [김정은, 김위원장],
북한: [조선인민민주주의공화국, 북조선]}...

복합어사전 {한동대학교: NNP, 남북 회담: NNP...

- 품사 선택

동사

명사

형용사

전처리 실행

내 사용자 사전

- 내 사전 등록: 원하는 단어들로 구성된 ‘내 사용자 사전’을 등록하여 전처리에 사용할 수 있습니다.
- 사전 삭제: 등록한 내 사용자 사전을 삭제합니다.
- 기본사전: 이미 시스템 내에 등록되어 있는 ‘기본 사전’을 선택합니다. 미리보기를 통해 기본 사전에 저장되어 있는 단어를 확인할 수 있고, 다운로드가 가능합니다.
- 사용 안함: 기본 사전을 포함한 모든 사전을 사용하지 않습니다.

품사 선택

- 동사, 명사, 형용사 중 어떤 품사들을 수집하고 싶은지 선택할 수 있으며, 복수 선택이 가능합니다.

내 사전 등록 방법

	검색 키워드	문서목록	수집날짜	통계	전처리
<input checked="" type="radio"/>	북한	- 북한 사람이 들려주는 북한 이야기 - 변하는 북한 변하지 않는 북한 - 북한 핵실험 이후 북한 문제 해법	2021년 07월 08일 20시 46분	총문서수:3개 평균 단어수:230개	
<input type="radio"/>	통일	- 북한 사람이 들려주는 북한 이야기 - 변하는 북한 변하지 않는 북한 - 북한 핵실험 이후 북한 문제 해법	2021년 08월 06일 20시 52분	총문서수:5개 평균 단어수:230개	

전처리

- 내 사용자 사전 적용

기본 사전 미리보기

불용어사전

stop_file.csv

사전삭제

기본사전 용안함

유의어사전

내 사전 등록

사전삭제

기본사전 용안함

복합어사전

내 사전 등록

사전삭제

기본사전 용안함

유의어사전 {김정은 위원장: [김정은, 김위원장]
북한: [조선인민민주주의공화국, 북조선]...}

복합어사전 {한동대학교: NNP, 남북 회담: NNP...}

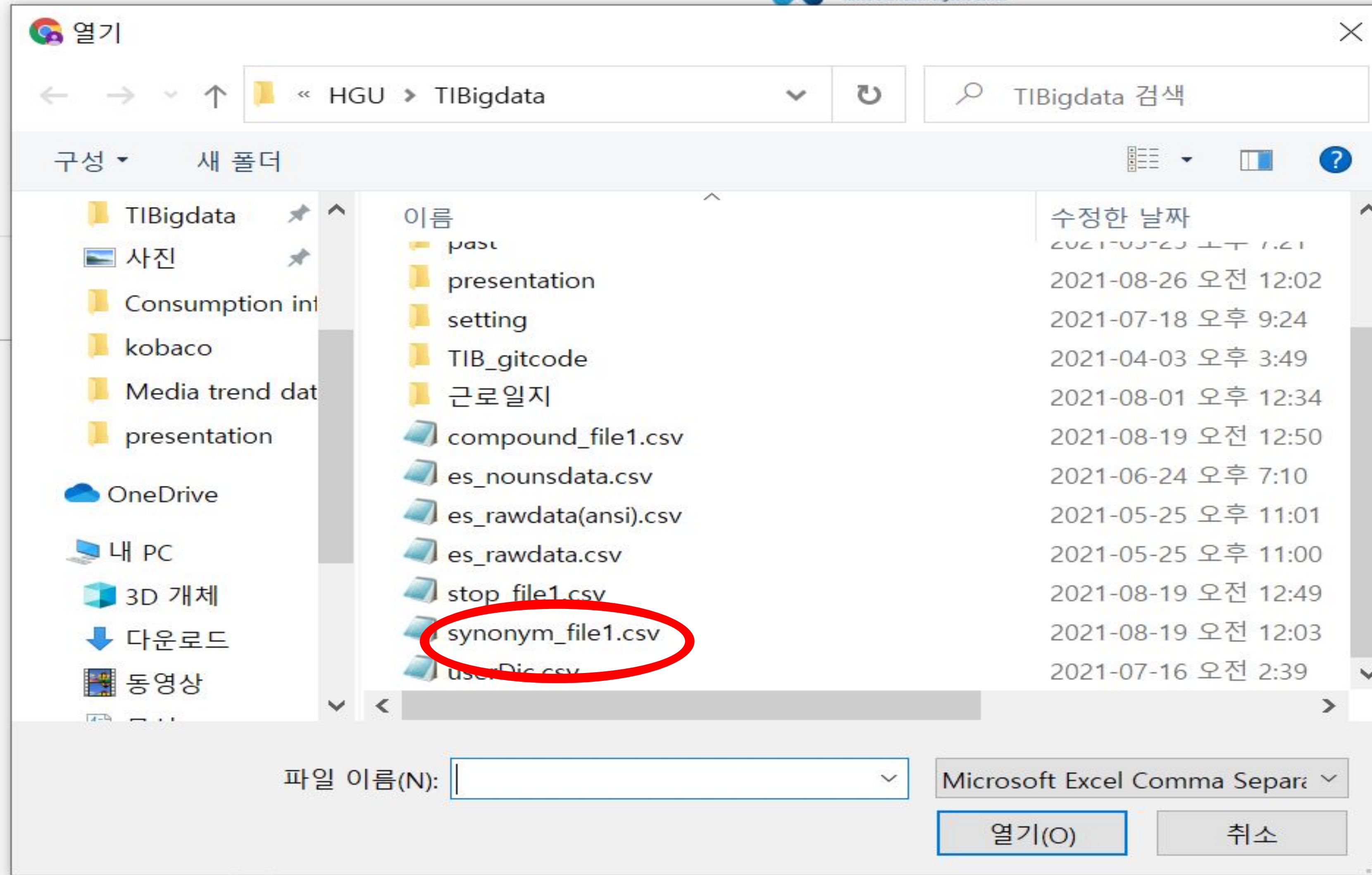
- 품사 선택

동사

명사

형용사

전처리 실행



홈 > 자료분석 > 전처리

통계

전처리

총문서수:3개
평균 단어수:230개총문서수:5개
평균 단어수:230개

	검색 키워드	문서목록	수집날짜	통계	전처리
<input checked="" type="radio"/>	북한	- 북한 사람이 들려주는 북한 이야기 - 변하는 북한 변하지 않는 북한 - 북한 핵실험 이후 북한 문제 해법	2021년 07월 08일 20시 46분	총문서수:3개 평균 단어수:230개	
<input type="radio"/>	통일	- 북한 사람이 들려주는 북한 이야기 - 변하는 북한 변하지 않는 북한 - 북한 핵실험 이후 북한 문제 해법	2021년 08월 06일 20시 52분	총문서수:5개 평균 단어수:230개	

전처리

- 내 사용자 사전 적용

기본 사전 미리보기

불용어사전

stop_file.csv

사전삭제

기본사전

사용안함

유의어사전

synonym_file.csv

사전삭제

기본사전

사용안함

복합어사전

내 사전 등록

사전삭제

기본사전

사용안함

복합어사전 {한동대 학교: NNP, 남북 회담: NNP...}

- 품사 선택

동사 명사 형용사

전처리 실행

유의어 사전

	A	B	C	D	E
1	김정은 위원장	김정은	김 위원장		
2	북한	북조선	조선인민민주주의공화국		
3	남한	대한민국	한국		
4	문재인 대통령	문재인	문대통령		
5					
6					
7					
8					
9					
10					

- 사용자는 첫번째 컬럼(A)에 대표로 등록하고 싶은 단어들을 작성하고, 다음 컬럼들(B,C..)에는 각각의 첫번째 컬럼에 해당하는 단어의 유의어들을 작성한다. 첫번째 컬럼의 단어들이 **대표어**가 되며, 나머지 컬럼의 단어들은 첫번째 컬럼의 단어로 대체된다.

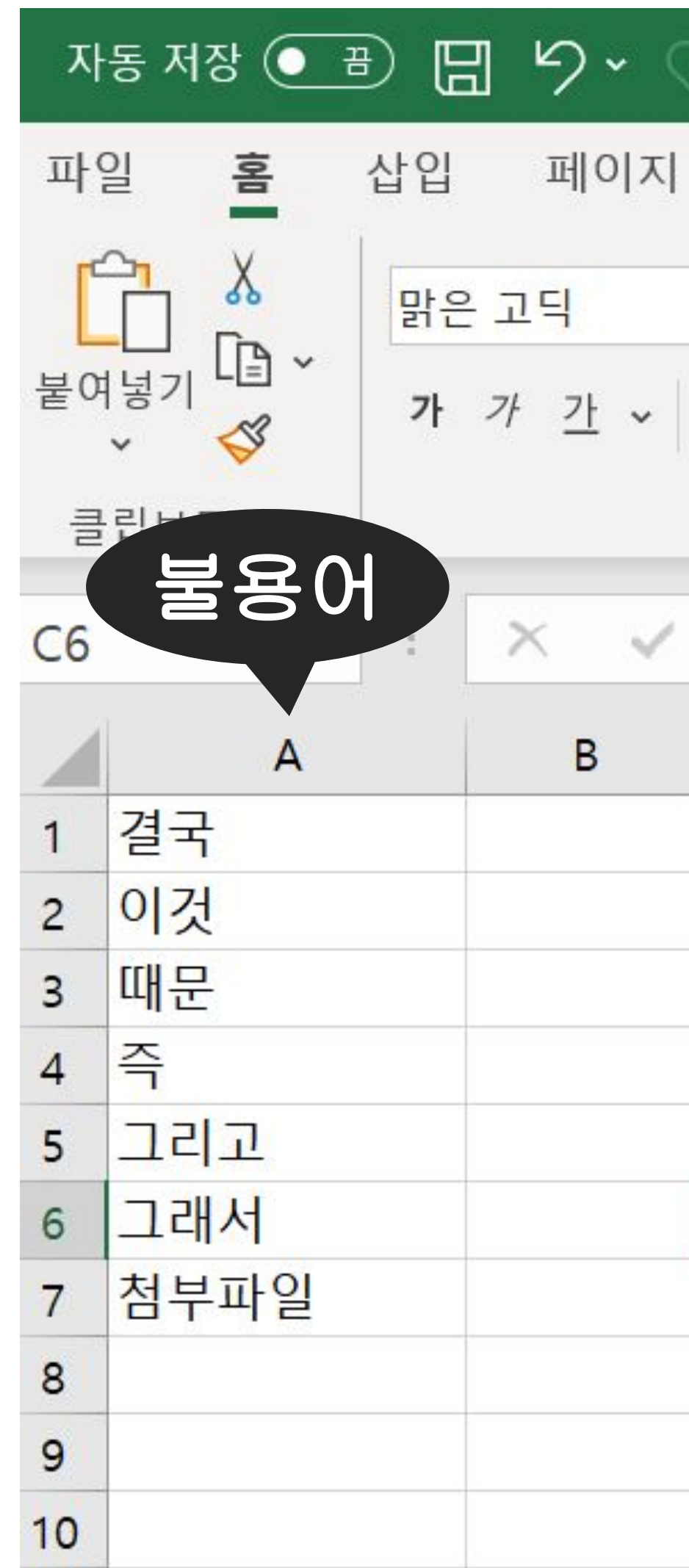
ex) 대표어1: 김정은 위원장

김정은, 김 위원장 □ 김정은 위원장

대표어2: 복한

북조선, 조선인민민주주의공화국 □ 북한

불용어 사전



- 사용자는 분석 시에 제외하고 싶은 단어들을 **불용어**로 등록할 수 있다. 대체로 조사, 접속사 등이 불용어가 될 수 있으나, 제거하고 싶은 명사, 형용사 또한 불용어로 등록하여 전처리시 제거할 수 있다.
- 불용어로 등록하고 싶은 단어들을 **첫번째 컬럼**에 작성하고 파일(csv형식)로 저장한다.

복합어 사전

	복합어	해당 품사
1	코로나	NNG
2	김정은 체제	NNG
3	김정은 국무위원장	NNG
4		
5		
6		
7		
8		
9		
10		

- 사용자는 신조어나 커스터마이징한 단어들을 복합어사전에 등록하여 분석시에 적용할 수 있다.
- **첫번째 컬럼**에 복합어로 등록하고 싶은 **단어**를 쓰고, **다음 컬럼**에는 그 단어의 **품사**를 **메카품사태그**를 참고하여 작성하고 파일(csv형식)로 저장한다.

분석



검색조건

기간

기관별

주제별

검색도움말

홈 > 자료분석 > 자료분석

분석

매뉴얼

전처리

분석

내 보관함

	검색 키워드	문서목록	수집날짜	통계	전처리
<input type="radio"/> 북한	- 북한 사람이 들려주는 북한 이야기 - 변하는 북한 변하지 않는 북한 - 북한 핵실험 이후 북한 문제 해법		2021년 07월 08일 20시 46분	총문서수:3개 평균 단어수:230개	
<input type="radio"/> 통일	- 북한 사람이 들려주는 북한 이야기 - 변하는 북한 변하지 않는 북한 - 북한 핵실험 이후 북한 문제 해법		2021년 08월 06일 20시 52분	총문서수:5개 평균 단어수:230개	

분석

매뉴얼

전처리

분석

내 보관함

	검색 키워드	문서목록	수집날짜	통계	전처리
<input type="radio"/>	북한	- 북한 사람이 들려주는 북한 이야기 - 변하는 북한 변하지 않는 북한 - 북한 핵실험 이후 북한 문제 해법	2021년 07월 08일 20시 46분	총문서수:3개 평균 단어수:230개	
<input type="radio"/>	통일	- 북한 사람이 들려주는 북한 이야기 - 변하는 북한 변하지 않는 북한 - 북한 핵실험 이후 북한 문제 해법	2021년 08월 06일 20시 52분	총문서수:5개 평균 단어수:230개	

분석

빈도수 분석

분석할 단어 수:

분석 실행

TFIDF



🔍 통일

검색조건

기간

기관별

주제별

검색도움말

홈 > 자료분석 > 자료분석

분석

매뉴얼

전처리

분석

내 보관함

	검색 키워드	문서목록	수집날짜	통계	전처리
<input type="radio"/>	북한 <ul style="list-style-type: none">- 북한 사람이 들려주는 북한 이야기- 변하는 북한 변하지 않는 북한- 북한 핵실험 이후 북한 문제 해법		2021년 07월 08일 20시 46분	총문서수:3개 평균 단어수:230개	
<input type="radio"/>	통일 <ul style="list-style-type: none">- 북한 사람이 들려주는 북한 이야기- 변하는 북한 변하지 않는 북한- 북한 핵실험 이후 북한 문제 해법		2021년 08월 06일 20시 52분	총문서수:5개 평균 단어수:230개	

빈도수 분석 ?

TFIDF ?

군집분석 ?

토픽 모델링 ?

N-gram ?

유의어분석 ?

의미연결망 ?

개체명인식 ?

[검색조건](#)[기간](#)[기관별](#)[주제별](#)[검색도움말](#)[홈](#) > [자료분석](#) > [자료분석](#)

분석

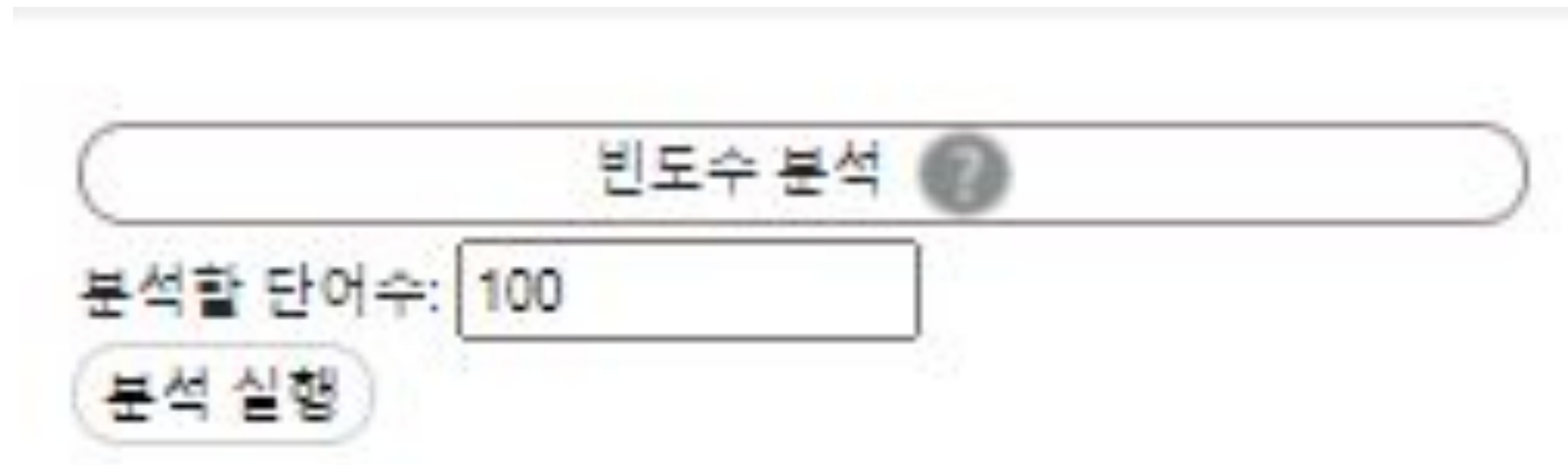
[매뉴얼](#)[전처리](#)[분석](#)

내 보관함

	검색 키워드	문서목록
<input type="radio"/>	북한	<ul style="list-style-type: none">- 북한 사람이 들려주는 북한 이야기- 변하는 북한 변하지 않는 북한- 북한 핵실험 이후 북한 문제 해법
<input type="radio"/>	통일	<ul style="list-style-type: none">- 북한 사람이 들려주는 북한 이야기- 변하는 북한 변하지 않는 북한- 북한 핵실험 이후 북한 문제 해법

[빈도수 분석 ?](#)분석할 단어수: [분석 실행](#)[빈도수 분석 ?](#)분석할 단어수: [분석 실행](#)[TFIDF ?](#)[빈도수 분석 ?](#)[트렌드 분석 ?](#)[N-gram ?](#)[유의어 분석 ?](#)[의미연결망 ?](#)[개체명인식 ?](#)

빈도수 분석



The screenshot shows a web interface for frequency analysis. At the top, there is a horizontal bar with the text "빈도수 분석" (Frequency Analysis) and a question mark icon. Below this bar, there is a label "분석할 단어수:" (Number of words to analyze:) followed by a text input field containing the number "100". Below the input field, there is a button labeled "분석 실행" (Execute Analysis).

- 빈도수 분석은 텍스트에서 많이 등장하는 단어의 등장 빈도를 분석함으로써 **핵심어**를 추출하는 분석방법
- 빈도수 분석 메뉴를 클릭하면, 사용자는 **상위 몇 개**의 단어들로 분석을 진행할 것인지 분석할 단어 수 옵션을 지정할 수 있다.

[검색조건](#)[기간](#)[기관별](#)[주제별](#)[검색도움말](#)[홈 > 자료분석 > 자료분석](#)

분석

[매뉴얼](#)[전처리](#)[분석](#)

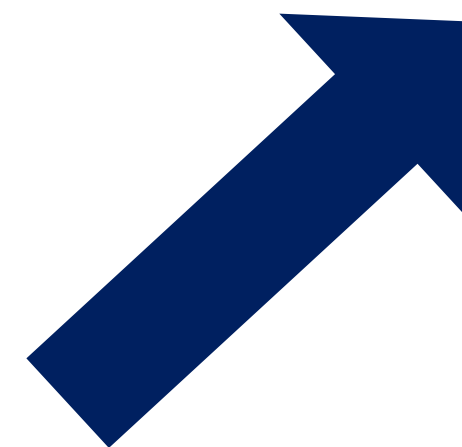
내 보관함

	검색 키워드	문서목록
<input type="radio"/>	북한 <ul style="list-style-type: none">- 북한 사람이 들려주는 북한 이야기- 변하는 북한 변하지 않는 북한- 북한 핵실험 이후 북한 문제 해법	
<input type="radio"/>	통일 <ul style="list-style-type: none">- 북한 사람이 들려주는 북한 이야기- 변하는 북한 변하지 않는 북한- 북한 핵실험 이후 북한 문제 해법	

TFIDF ?

분석할 단어수:

분석 실행



TFIDF ?

분석할 단어수:

분석 실행

빈도수 분석 ?

단어분석 ?

트픽 모델링 ?

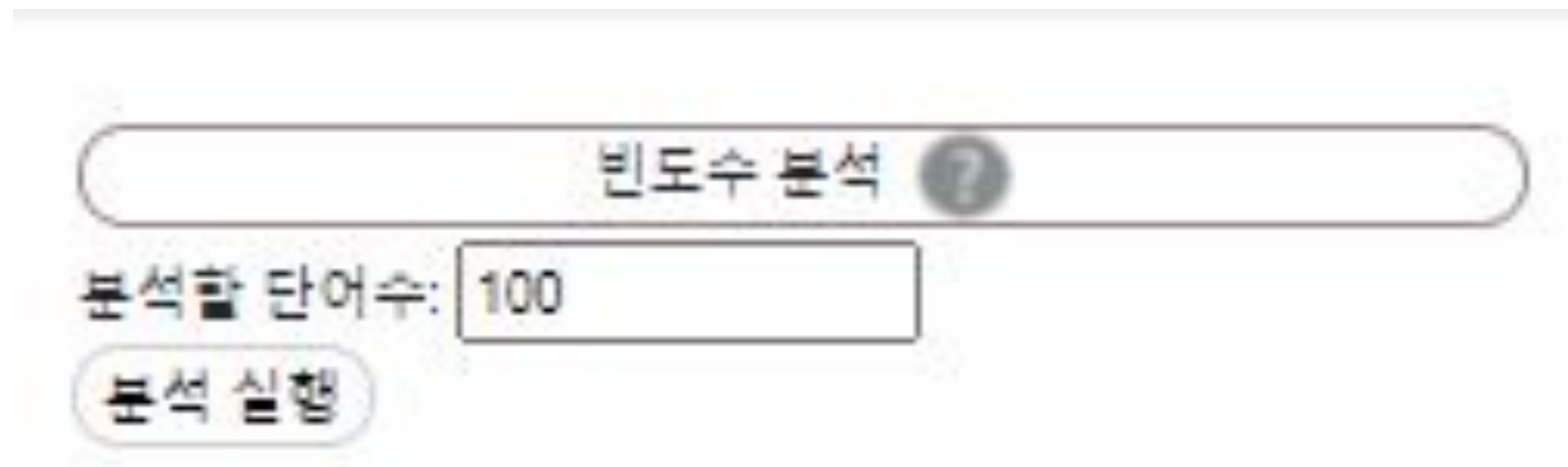
N-gram ?

유사어분석 ?

의미연결망 ?

개체명인식 ?

tfidf 분석



빈도수 분석 ?

분석할 단어수: 100

분석 실행

- Tfidf는 문서 역빈도 분석이라고도 하며, 특정 말뭉치에 등장하는 **어휘의 빈도**를 계산해 단어가 특정 문서에서 얼마나 중요한 **핵심어**인지 추출하는 분석방법이다.
- Tfidf 메뉴를 클릭하면, 사용자는 상위 몇 개의 단어들로 tfidf분석을 진행할 것인지 분석할 단어 수 옵션을 지정할 수 있다.