

```
!pip install konlpy
```

```
Requirement already satisfied: konlpy in /usr/local/lib/python3.7/dist-packages (0.5.2)
Requirement already satisfied: beautifulsoup4==4.6.0 in /usr/local/lib/python3.7/dist-packages (from konlpy)
Requirement already satisfied: colorama in /usr/local/lib/python3.7/dist-packages (from konlpy)
Requirement already satisfied: JPype1>=0.7.0 in /usr/local/lib/python3.7/dist-packages (from konlpy)
Requirement already satisfied: lxml>=4.1.0 in /usr/local/lib/python3.7/dist-packages (from konlpy)
Requirement already satisfied: tweepy>=3.7.0 in /usr/local/lib/python3.7/dist-packages (from konlpy)
Requirement already satisfied: numpy>=1.6 in /usr/local/lib/python3.7/dist-packages (from konlpy)
Requirement already satisfied: typing-extensions in /usr/local/lib/python3.7/dist-packages (from konlpy)
Requirement already satisfied: requests-oauthlib>=0.7.0 in /usr/local/lib/python3.7/dist-packages (from konlpy)
Requirement already satisfied: six>=1.10.0 in /usr/local/lib/python3.7/dist-packages (from konlpy)
Requirement already satisfied: requests[socks]>=2.11.1 in /usr/local/lib/python3.7/dist-packages (from konlpy)
Requirement already satisfied: oauthlib>=3.0.0 in /usr/local/lib/python3.7/dist-packages (from konlpy)
Requirement already satisfied: urllib3!=1.25.0,!1.25.1,<1.26,>=1.21.1 in /usr/local/lib/python3.7/dist-packages (from konlpy)
Requirement already satisfied: idna<3,>=2.5 in /usr/local/lib/python3.7/dist-packages (from konlpy)
Requirement already satisfied: certifi>=2017.4.17 in /usr/local/lib/python3.7/dist-packages (from konlpy)
Requirement already satisfied: chardet<4,>=3.0.2 in /usr/local/lib/python3.7/dist-packages (from konlpy)
Requirement already satisfied: PySocks!=1.5.7,>=1.5.6 in /usr/local/lib/python3.7/dist-packages (from konlpy)
```

```
# 메밀꽃.txt에서 나온 횟수가 5번이상인 단어 출력
```

```
from konlpy.tag import Kkma
```

```
from collections import Counter
```

```
infile = open("메밀꽃.txt", encoding='utf-8')
```

```
data=infile.read()
```

```
kkma = Kkma()
```

```
words=kkma.morphs(data)
```

```
vocab=Counter(words)
```

```
print(vocab)
```

```
vo=[print(f'-vlaue {k} appears multiple times ({v} times)') for k,v in  
vocab.items() if v > 4]
```

```
del vo[:1]
```

```
-vlaue 그 appears multiple times (18 times)
```

```
-vlaue 으면 appears multiple times (5 times)
```

```
-vlaue 같이 appears multiple times (10 times)
```

```
-vlaue 계집 appears multiple times (12 times)
```

```
-vlaue 로 appears multiple times (15 times)
```

```
-vlaue ..... appears multiple times (10 times)
```

```
-vlaue 집 appears multiple times (7 times)
```

```
-vlaue 말 appears multiple times (12 times)
```

```
-vlaue 야 appears multiple times (16 times)
```

```
-vlaue 듯이 appears multiple times (5 times)
```

```
-vlaue 구 appears multiple times (7 times)
```

```
-vlaue 그럴 appears multiple times (9 times)
```

```
-vlaue 동이 appears multiple times (19 times)
```

```
-vlaue 세 appears multiple times (7 times)
```

```
-vlaue 충춇집 appears multiple times (5 times)
```

```
-vlaue 녀석 appears multiple times (9 times)
```

```
-vlaue 알 appears multiple times (5 times)
```

```
-vlaue 길 appears multiple times (6 times)
```

```
-vlaue 만 appears multiple times (14 times)
-vlaue 수 appears multiple times (20 times)
-vlaue 마음 appears multiple times (5 times)
-vlaue 버리 appears multiple times (6 times)
-vlaue 때 appears multiple times (16 times)
-vlaue 치 appears multiple times (6 times)
-vlaue 부터 appears multiple times (7 times)
-vlaue 앞 appears multiple times (5 times)
-vlaue 면 appears multiple times (12 times)
-vlaue 하나 appears multiple times (6 times)
-vlaue 겠 appears multiple times (11 times)
-vlaue 란 appears multiple times (5 times)
-vlaue 니 appears multiple times (10 times)
-vlaue 같 appears multiple times (8 times)
-vlaue 면서 appears multiple times (5 times)
-vlaue 젊 appears multiple times (6 times)
-vlaue 그것 appears multiple times (5 times)
-vlaue 러 appears multiple times (6 times)
-vlaue 오 appears multiple times (5 times)
-vlaue 짐승 appears multiple times (10 times)
-vlaue 눈 appears multiple times (6 times)
-vlaue 나귀 appears multiple times (15 times)
-vlaue 젓 appears multiple times (7 times)
-vlaue 까 appears multiple times (5 times)
-vlaue 울 appears multiple times (6 times)
-vlaue 속 appears multiple times (7 times)
-vlaue 아이 appears multiple times (7 times)
-vlaue 아니 appears multiple times (5 times)
-vlaue 저 appears multiple times (5 times)
-vlaue 재천 appears multiple times (7 times)
-vlaue 라고 appears multiple times (6 times)
-vlaue 물 appears multiple times (13 times)
-vlaue 다고 appears multiple times (6 times)
-vlaue 평 appears multiple times (6 times)
-vlaue 지금 appears multiple times (7 times)
-vlaue 이야기 appears multiple times (7 times)
-vlaue 처녀 appears multiple times (6 times)
-vlaue 어머니 appears multiple times (5 times)
-vlaue 쯤 appears multiple times (6 times)
```

```
# 소나기.txt에서 나온 횟수가 5번이상인 단어 출력
```

```
from konlpy.tag import Kkma
```

```
from collections import Counter
```

```
infile = open("소나기.txt", encoding='utf-8')
```

```
data=infile.read()
```

```
kkma = Kkma()
```

```
words=kkma.morphs(data)
```

```
vocab=Counter(words)
```

```
print(vocab)
```

```
vo=[print(f'-vlaue {k} appears multiple times ({v} times)') for k,v in
    vocab.items() if v > 4]
```

```
del vo[:1]
```

```

-vlaue 알 appears multiple times (22 times)
-vlaue 쪽 appears multiple times (5 times)
-vlaue 안 appears multiple times (8 times)
-vlaue 의 appears multiple times (35 times)
-vlaue 같 appears multiple times (16 times)
-vlaue 뵈 appears multiple times (7 times)
-vlaue 덜 appears multiple times (5 times)
-vlaue 집 appears multiple times (7 times)
-vlaue 주머니 appears multiple times (5 times)
-vlaue 일 appears multiple times (6 times)
-vlaue □ appears multiple times (5 times)
-vlaue 잡 appears multiple times (6 times)
-vlaue 전 appears multiple times (9 times)
-vlaue 더 appears multiple times (17 times)
-vlaue 두 appears multiple times (6 times)
-vlaue 으로 appears multiple times (27 times)
-vlaue 번 appears multiple times (5 times)
-vlaue 말 appears multiple times (12 times)
-vlaue 느냐 appears multiple times (5 times)
-vlaue ' appears multiple times (12 times)
-vlaue 어서 appears multiple times (9 times)
-vlaue 없 appears multiple times (8 times)
-vlaue 없이 appears multiple times (5 times)
-vlaue 소리 appears multiple times (6 times)
-vlaue 오 appears multiple times (12 times)
-vlaue 니 appears multiple times (9 times)
-vlaue 보이 appears multiple times (5 times)
-vlaue 못 appears multiple times (6 times)
-vlaue 뜰 appears multiple times (10 times)
-vlaue 이것 appears multiple times (8 times)
-vlaue ? appears multiple times (16 times)
-vlaue 자기 appears multiple times (6 times)
-vlaue 눈 appears multiple times (6 times)
-vlaue 참 appears multiple times (8 times)
-vlaue 산 appears multiple times (6 times)

-vlaue 우리 appears multiple times (5 times)
-vlaue 멀 appears multiple times (5 times)
-vlaue 면 appears multiple times (5 times)
-vlaue 허수아비 appears multiple times (6 times)
-vlaue 들 appears multiple times (7 times)
-vlaue 대 appears multiple times (5 times)
-vlaue 달려가 appears multiple times (5 times)
-vlaue 그 appears multiple times (14 times)
-vlaue ㄴ다는 appears multiple times (5 times)
-vlaue 곳 appears multiple times (6 times)
-vlaue 밑 appears multiple times (6 times)
-vlaue 수숫단 appears multiple times (5 times)
-vlaue 꽃 appears multiple times (14 times)
-vlaue …… appears multiple times (13 times)
-vlaue 그럴 appears multiple times (5 times)
-vlaue 크 appears multiple times (5 times)
-vlaue 가지 appears multiple times (5 times)
-vlaue 송아지 appears multiple times (6 times)
-vlaue 비 appears multiple times (7 times)
-vlaue 이사 appears multiple times (6 times)
-vlaue 호두 appears multiple times (6 times)
-vlaue 아버지 appears multiple times (5 times)

```

```

# 메밀꽃.txt에서 줄풍어를 제거하고 나온 윗수가 5번이상인 단어 출력
from konlpy.tag import Kkma
from collections import Counter
infile = open("메밀꽃.txt", encoding='utf-8')
data=infile.read()
stopfile = open("불용어.txt", encoding='utf-8')
stopwords = stopfile.read()

kkma = Kkma()
words=kkma.morphs(data)
stopwords=kkma.morphs(stopwords)

result=[]
for w in words:
    if w not in stopwords:
        result.append(w)

vocab=Counter(result)
print(vocab)
vo=[print(f'-vlaue {k} apppears multiple times ({v} times)') for k,v in
    vocab.items() if v > 4]
del vo[:1]

```

```

Counter({'으나': 25, '생원': 25, '동이': 19, '나귀': 15, '장': 13, '꼴': 13, '이나': 12, '계
-vlaue 장 apppears multiple times (13 times)
-vlaue 이란 apppears multiple times (5 times)
-vlaue 장판 apppears multiple times (6 times)
-vlaue 벌써 apppears multiple times (7 times)
-vlaue 뒤 apppears multiple times (8 times)
-vlaue 으나 apppears multiple times (25 times)
-vlaue 이나 apppears multiple times (12 times)
-vlaue 축 apppears multiple times (5 times)
-vlaue ㄴ지 apppears multiple times (6 times)
-vlaue 스럽 apppears multiple times (8 times)
-vlaue 군 apppears multiple times (5 times)
-vlaue 생원 apppears multiple times (25 times)
-vlaue 조선 apppears multiple times (9 times)
-vlaue 달 apppears multiple times (10 times)
-vlaue 봉 apppears multiple times (9 times)
-vlaue 을까 apppears multiple times (5 times)
-vlaue 대화 apppears multiple times (5 times)
-vlaue 밤 apppears multiple times (10 times)
-vlaue 걸 apppears multiple times (8 times)
-vlaue 달이 apppears multiple times (5 times)
-vlaue 소리 apppears multiple times (9 times)
-vlaue 시작 apppears multiple times (10 times)
-vlaue 꼴 apppears multiple times (13 times)
-vlaue 계집 apppears multiple times (12 times)
-vlaue 듯이 apppears multiple times (5 times)
-vlaue 동이 apppears multiple times (19 times)
-vlaue 세 apppears multiple times (7 times)
-vlaue 충춇집 apppears multiple times (5 times)
-vlaue 녀석 apppears multiple times (9 times)
-vlaue 길 apppears multiple times (6 times)
-vlaue 버리 apppears multiple times (6 times)
-vlaue 치 apppears multiple times (6 times)

```

```
-vlaue 란 appears multiple times (5 times)
-vlaue 젊 appears multiple times (6 times)
-vlaue 짐승 appears multiple times (10 times)
-vlaue 눈 appears multiple times (6 times)
-vlaue 나귀 appears multiple times (15 times)
-vlaue 젓 appears multiple times (7 times)
-vlaue 울 appears multiple times (6 times)
-vlaue 재천 appears multiple times (7 times)
-vlaue 라고 appears multiple times (6 times)
-vlaue 다고 appears multiple times (6 times)
-vlaue 평 appears multiple times (6 times)
-vlaue 이야기 appears multiple times (7 times)
-vlaue 처녀 appears multiple times (6 times)
-vlaue 어머니 appears multiple times (5 times)
-vlaue 쪼 appears multiple times (6 times)
```

```
# 소나기.txt에서 불용어를 제거하고 나온 횟수가 5번 이상인 단어 출력
from konlpy.tag import Kkma
from collections import Counter
infile = open("소나기.txt", encoding='utf-8')
data=infile.read()
stopfile = open("불용어.txt", encoding='utf-8')
stopwords = stopfile.read()

kkma = Kkma()
words=kkma.morphs(data)
stopwords=kkma.morphs(stopwords)

result=[]
for w in words:
    if w not in stopwords:
        result.append(w)

vocab=Counter(result)
print(vocab)
vo=[print(f'-vlaue {k} appears multiple times ({v} times)') for k,v in
    vocab.items() if v > 4]
del vo[:1]
```

```
Counter({'소녀': 72, '소년': 49, '날': 14, '꽃': 14, '손': 11, '앗': 10, '들': 10, '개울가':
-vlaue 소년 appears multiple times (49 times)
-vlaue 개울가 appears multiple times (8 times)
-vlaue 소녀 appears multiple times (72 times)
-vlaue 운 appears multiple times (5 times)
-vlaue 초시 appears multiple times (5 times)
-vlaue 개울 appears multiple times (6 times)
-vlaue 손 appears multiple times (11 times)
-vlaue 듯이 appears multiple times (6 times)
-vlaue 돌아오 appears multiple times (7 times)
-vlaue 길 appears multiple times (8 times)
-vlaue 더니 appears multiple times (6 times)
-vlaue 오늘 appears multiple times (5 times)
-vlaue 징검다리 appears multiple times (6 times)
-vlaue 앓 appears multiple times (10 times)
```

-vlaue 버리 appears multiple times (5 times)
-vlaue 날 appears multiple times (14 times)
-vlaue 얼굴 appears multiple times (8 times)
-vlaue 움키 appears multiple times (5 times)
-vlaue 는지 appears multiple times (6 times)
-vlaue 건너 appears multiple times (5 times)
-vlaue 다가 appears multiple times (6 times)
-vlaue 조약돌 appears multiple times (5 times)
-vlaue 고는 appears multiple times (8 times)
-vlaue 바보 appears multiple times (7 times)
-vlaue 머리 appears multiple times (6 times)
-vlaue 갈발 appears multiple times (5 times)
-vlaue 가을 appears multiple times (5 times)
-vlaue 갈꽃 appears multiple times (7 times)
-vlaue 다고 appears multiple times (5 times)
-vlaue 뵈 appears multiple times (7 times)
-vlaue 덜 appears multiple times (5 times)
-vlaue 주머니 appears multiple times (5 times)
-vlaue 잡 appears multiple times (6 times)
-vlaue 번 appears multiple times (5 times)
-vlaue 느냐 appears multiple times (5 times)
-vlaue 없이 appears multiple times (5 times)
-vlaue 소리 appears multiple times (6 times)
-vlaue 못 appears multiple times (6 times)
-vlaue 뜰 appears multiple times (10 times)
-vlaue 눈 appears multiple times (6 times)
-vlaue 산 appears multiple times (6 times)
-vlaue 멀 appears multiple times (5 times)
-vlaue 허수아비 appears multiple times (6 times)
-vlaue 달려가 appears multiple times (5 times)
-vlaue ㄴ다는 appears multiple times (5 times)
-vlaue 밑 appears multiple times (6 times)
-vlaue 수숫단 appears multiple times (5 times)
-vlaue 꽃 appears multiple times (14 times)
-vlaue 송아지 appears multiple times (6 times)
-vlaue 이사 appears multiple times (6 times)
-vlaue 호두 appears multiple times (6 times)
-vlaue 아버지 appears multiple times (5 times)

✓ 4초 오전 4:48에 완료됨

