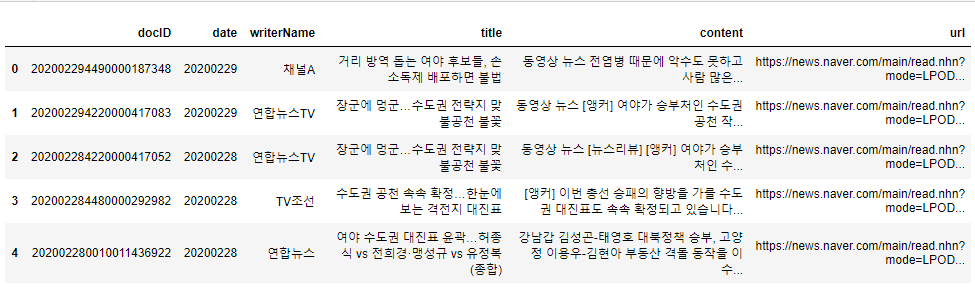
의원별 대표뉴스 추출

1. 크롤링한 데이터



1. 뉴스 본문 토큰화 -> Doc2Vec을 이용하여 벡터화 -> k-means를 이용하여 클러스터링
2. 각 뉴스의 벡터값과 k-means의 중앙값의 코사인 유사도를 구함



1. 날짜를 내림차순으로 정렬한 후 그룹별 코사인 유사도가 가장 높으면서 뉴스 제목에 ‘나경원’이 들어간 대표 뉴스 5개 추출

