Part 1. Web Crawling

* 왜 NYT를 선택했는지 설명 (Target)
* 기사를 crawling 한 키워드, 기간, 개수 설명 (Data collection)
* Data cleaning 방법 설명 – 최종 data 개수

- Number of article

- Number of words(Token)

* Collection한 data를 분류한 방법 (titles, bodies, date)
* Fire, Fell, collapse 등 대표 키워드 빈도 분석

Part 2. Word2Vec

* Word2Vec의 방법 상세히 설명 (size, window, min, workers, sg, iter 등) (parameter)
* Fire, Fell, collapse 등 대표 키워드 관련 유사한 단어 50개 나열 및 분류
* 대표적으로 유사도 높은예, 낮은예 비교로 신뢰도 검증

Part 3. Uniform Manifold Approximation and Projection (UMAP)

* UMAP 분석 방법 상세히 설명 / UMAP이 의미하는 것
* 관련 그래프 제공 및 설명
* 각 대표 키워드 그룹핑 후 상세 설명

<공통>

* 각 방법론 이론적 / 수학적 설명
* 각 part별 literature review