LDA LSA는 시장동향 및 석사논문에서 가장 많이 쓰이는 분석 도구이다. (트랜드 분석 논문을 찾아보면 대부분 토픽모델링으로 LDA 및 LSA를 사용한다.)

LDA = 토픽 별로 어떤 단어 분포를 가지고 있는가

어떻게?

* 단어 분포를 가지고 주제를 분류(문장/문단 안에 함께 나오는 단어들이 하나의 토픽으로 묶임)

사용자가 설정하는 값: 알파, 베타, K(토픽의 갯수)

함수 식 = 문서 안에서 단어가 몇개 등장했는가?(우리가 알 수 있는 값), 사용자가 설정하는 값(우리가 알고 있는 값)