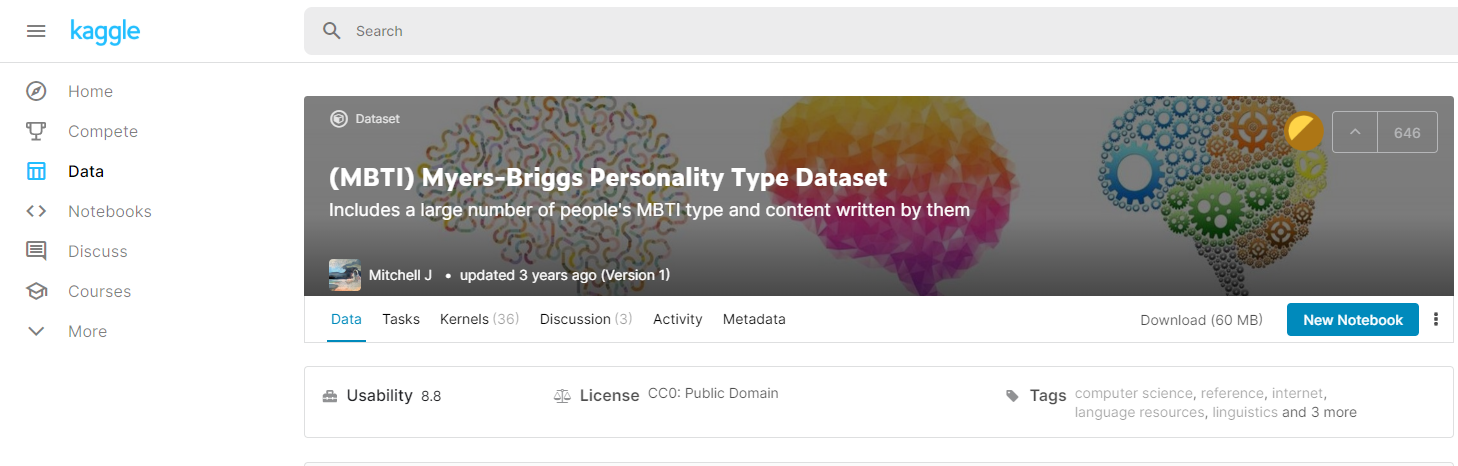
**IOI\_Project1\_20171534\_정진솔**

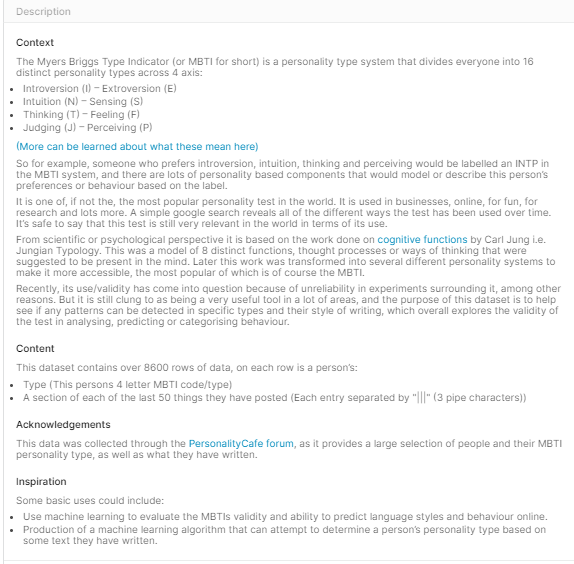
**(MBTI) Myers-Briggs Personality Type Dataset**

<https://www.kaggle.com/laowingkin/mbti-study-personality>

1. 원하는 데이터 셋 선택



1. 데이터 셋 이해



* 마이어스 브릭스 유형 지표(또는 줄여서 MBTI)는 모든 사람을 4축에 걸쳐 16개의 뚜렷한 성격 유형으로 나누는 성격 유형 시스템이다.

내향성(I) – 외향성(E)

직감(N) – 감지(S)

생각(T) – 느낌(F)

판단(J) – 인지(P)

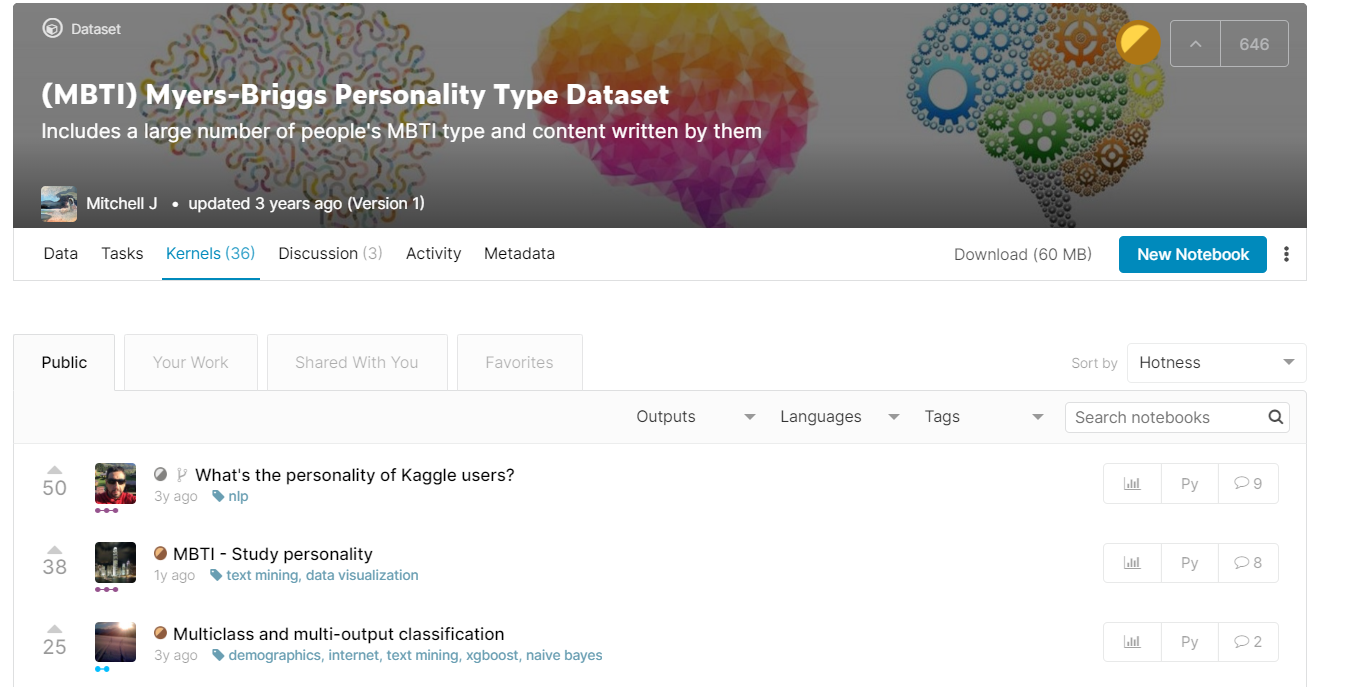
예를 들어, 내성적, 직관적, 사고와 인식을 선호하는 사람은 MB에서 INTP라고 불린다.TI 시스템, 그리고 라벨에 기초하여 이 사람의 선호나 행동을 모델화하거나 설명하는 많은 성격 기반 요소들이 있다.

그것은 세계에서 가장 인기 있는 인성검사 중 하나이다. 그것은 사업, 온라인, 재미, 연구, 그리고 더 많은 것들을 위해 사용된다. 간단한 구글 검색은 그 테스트가 시간이 지남에 따라 사용되어 온 모든 다른 방법을 보여준다. 이 테스트는 사용의 측면에서 여전히 세계와 매우 관련이 있다고 해도 무방하다.

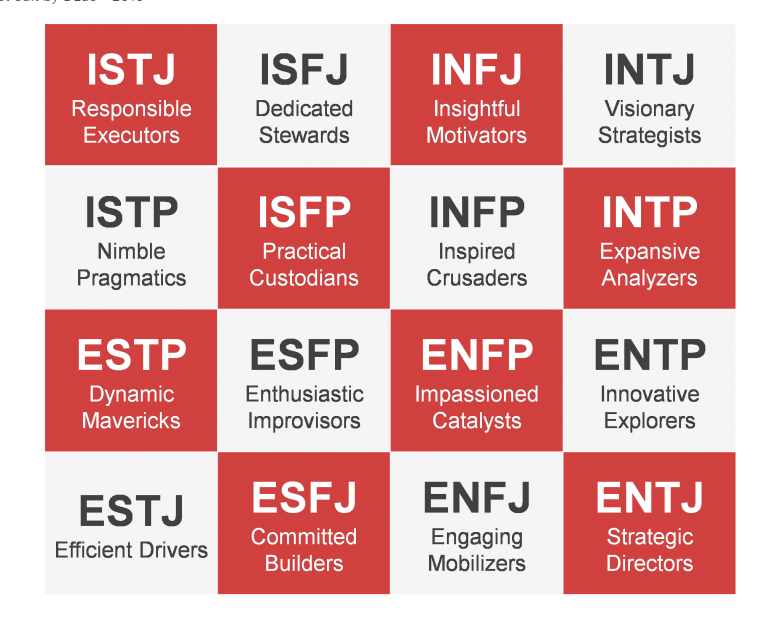
과학적 또는 심리학적 관점에서 칼 융이 인지 기능에 대해 수행한 작업을 바탕으로 한다. Jungian Typology. 이것은 마음속에 존재한다고 제안된 8가지 뚜렷한 기능, 사고 과정 또는 사고 방식의 모델이었다. 후에 이 작품은 접근성을 높이기 위해 몇 가지 다른 성격 체계로 변형되었는데, 그 중 가장 인기 있는 것은 물론 MBTI이다.

최근에, 그것의 사용/유효성은 다른 이유들 중에서도 그것을 둘러싼 실험에서 신뢰할 수 없기 때문에 문제가 되었다. 그러나 이 데이터 집합의 목적은 여전히 많은 영역에서 매우 유용한 도구로서 고수되고 있으며, 이 데이터 집합의 목적은 특정 유형과 특정 유형에서 어떤 패턴이 감지될 수 있는지, 그리고 이들의 쓰기 스타일에서 패턴이 감지될 수 있는지를 확인하는 데 도움이 되며, 이는 전반적으로 행태를 분석, 예측 또는 분류하는 데 있어 테스트의 유효성을 탐구한다.

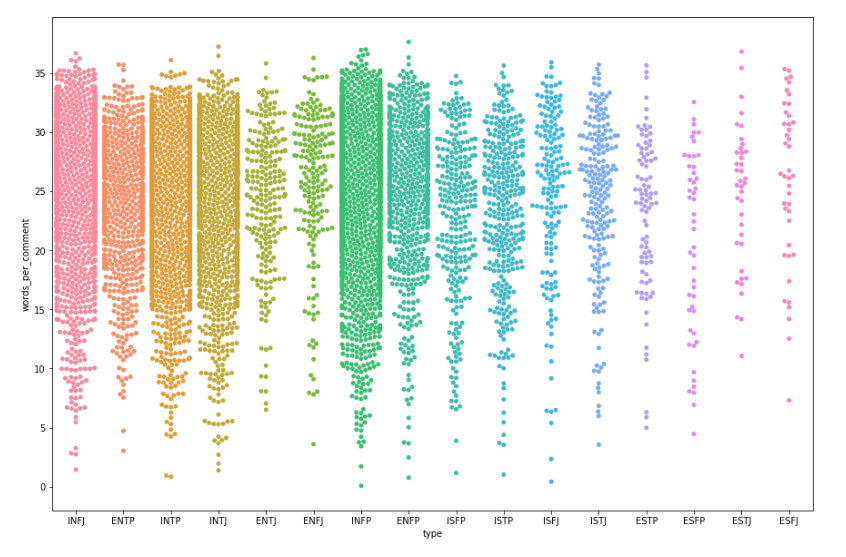
1. 데이터 셋 kernels 중 흥미로운 주제 선택



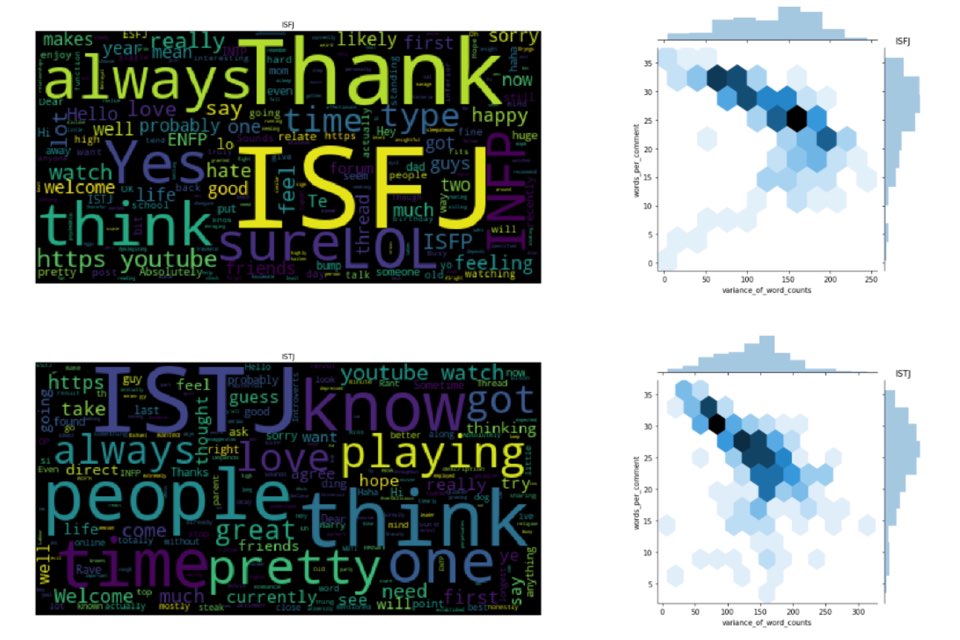
1. 선택한 kernel 분석



* 제 MBTI는 ISFJ(내향형, 감각형, 감정형, 판단형) 입니당 \*^^\* 맞는것 같나용?^^?



* words\_per\_comment를 기반으로 한 MBTI 분포



* words\_per\_comment & variance\_of\_word\_counts 를 기반으로 분석한 ISFJ, ISFP

1. 해당 데이터 활용방안 제시

* 기계 학습을 사용하여 MBTI의 타당성과 온라인에서의 언어 스타일, 행동을 예측하는 능력을 평가함
* 사용자가 작성한 일부 텍스트에 기반하여 개인의 성격 유형을 결정할 수 있는 기계 학습 알고리즘을 제작

= 적절히 조화롭게 섞어 APP출시도 고려해 볼 수 있음