**- 정형데이터(Structured data/formal data)**
즉시 통계적 분석에 사용될 수 있을 만한 형태로 정리되고 가공된 데이터
고정된 필드에 저장된 데이터(관계형 데이터베이스, 스프레드시트 등)

**- 반정형데이터(semi-structured data)**
파일 형태, 메타데이터(데이터 내부에 정형 데이터의 스키마)
반정형 데이터의 반은 Semi를 의미. 즉, 완전한 정형이 아니라 약한 정형 데이터라는 뜻을 담고 있음. 그렇기 때문에, 고정된 양식은 없으나 어느 정도 구조가 정해져 있는 데이터로, 반정형 데이터의 종류로는 로그 데이터, HTML, XML 등이 있음.

**- 비정형데이터(Unstructured data)**
데이터 세트가 아닌 하나의 데이터가 수집 데이터로 객체화
언어 분석이 가능한 텍스트 데이터, 멀티미디어 데이터 - 동영상, 이미지, 텍스트 등