빅테이터는 정형 데이터, 비정형 데이터, 반정형 데이터가 있습니다.

정형 데이터는 고정된 필드에 저장된 데이터를 말하며 관계형 데이터 베이스와 스프레드시트 등을 예로 들 수 있습니다. 정형 데이터의 경우는 데이터베이스를 설계한 기술자에 의해 수집되는 정보의 형태가 정해지게 됩니다. 한정된 정보들 속에서 고객의 정보와 상품 분석, 인기 품목에 대한 정보를 분석할 수 있습니다.

비정형 데이터는 고정된 필드에 저장되지 않은 데이터를 의미하며 페이스북과 트위터, 유튜브 영상, 이미지 파일, 음원파일, 워드 문서, PDF 문서 등을 예로 들 수 있습니다. 비정형 데이터의 경우는 페이스북, 트위터, 네이버, 다움 등에서 생성되는 실시간 정보들을 통해서 더 많은 정보들을 수집하고 분석할 수 있습니다. 예를 들면 특정 지역의 날씨 정보, 유동 인구의 수, 이들의 판매 정보등을 수집할 수 있습니다. 형태가 정해지지 않는 정보속에서 분석 방향에 따라 다양한 정보를 수집할 수 있는 것입니다.

반정형 데이터는 고정된 필드에 저장된 데이터는 아니지만 XML, HTML 텍스트 등 메타데이터 및 스키마를 포함하는 데이터입니다.

빅데이터의 85% 가량이 형태가 정해지지 않은 비정형 데이터입니다. 소셜 네트워크 이용자 수의 증가로 비정형 데이터는 급속도로 확산되고 있는 추세지만, 정형 데이터 분석을 위해서 이용되고 있는 많은 기술들이 비정형 데이터에서는 활용할 수 없다는 한계를 가지고 있습니다.