문항 1. 데이터 분석과 빅데이터 분석의 차이점에 4가지에 대해서 설명하세요.

빅 데이터와 데이터 분석의 차이점은 커다란 데이터는 대량의 복잡한 데이터이며 데이터 분석은 유용한 정보를 인식하고 의사 결정을 지원하기 위해 데이터를 조사, 변환 및 모델링하는 프로세스라는 것이다.

문항 2. 데이터 전처리 기술 3가지가 무엇인지 명칭을 작성하고, 그 개념에 대해 설명하시요. (필요시, 사용되는 세부적인 기술을 설명하시요.)

데이터 전처리 기술은 데이터 필터링, 데이터 변환, 데이터 정제 3가지가 있다.

(1) 데이터 필터링- 오류 발견, 보정, 삭제 및 중복성 검사 등을 수행한다.

(2) 데이터 변환- 다양한 형식으로 수집된 데이터를 분석에 용이하도록 일관성이 있는 형식으로 변환한다. 사용되는 세부적인 기술은 평활화, 집계, 정규화, 일반화 속성 생성이 있다.

(3) 데이터 정제- 수집된 데이터의 불일치성을 교정하기 위한 것이다.