

R을 통해 PDF 속에 저장된 데이터도 읽을 수 있습니다.

데이터 분석의 과정은 다음과 같습니다.

우선 데이터를 입력 받습니다.

데이터 입력은 txt파일, csv파일, 엑셀 파일, pdf파일 등 다양한 형태로 입력 받을 수 있습니다.

데이터를 입력 받은 후 분석에 적합한 데이터로 가공하는 작업을 거쳐야 하며, 이를 데이터 전처리라고 부릅니다.

데이터 전처리는 데이터 분석의 80%이상을 차지하는 중요한 과정입니다.