

2 CRISP-DM 분석 방법론

- Cross-Industry Standard Process for Data Mining
- 데이터 마이닝에 대한 산업 표준 프로세스로 1996년 유럽연합의 ESPRIT에서 있었던 프로젝트에서 시작되었다.
- 4개 레벨, 6단계로 구성되어 있다.
- 폭포수 모델처럼 한 방향으로 구성되어 있지 않고 **단계 간 피드백(Feedback)**을 통하여 단계별 완성도를 높이게 구성되어 있다.

[2-1] CRISP-DM 4레벨

단계(Phases)	<ul style="list-style-type: none"> 최상위 레벨, 여러 개의 단계(Phases)로 구성됨 각 Phases는 일반화 태스크(Generic Tasks)를 포함 예) 기획, 데이터 준비, 분석, 평가
일반화 태스크 (Generic Tasks)	<ul style="list-style-type: none"> 데이터 마이닝의 단일 프로세스를 완전하게 수행하는 단위 예) 데이터 정제(Data Cleaning)
세분화 태스크 (Specialized Tasks)	<ul style="list-style-type: none"> 일반화 태스크를 구체적으로 수행하는 단위 예) 일반화 태스크 데이터 정제의 세분화 태스크에는 범주형 데이터 정제, 연속형 데이터 정제 등이 있음
프로세스 실행 (Process Instances)	데이터 마이닝을 위한 구체적인 실행

