

제1절 분석 기획 방향성 도출

○ 분석 기획

- 분석 수행 전, 과제의 정의 및 의도한 결과를 도출 할 수 있도록 관리방안을 사전계획
- 어떠한 목표(What)를 달성하기 위하여(Why) 어떠한 데이터를 가지고 어떤 방식으로(How) 수행할지에 대한 일련의 계획 수립 과정

1. 분석 기획 특징

○ 분석기획에 필요한 역량 3가지

- 수학/통계학적 지식
- 정보기술(IT, 해킹, 통신기술)
- 해당 비즈니스에 대한 이해와 전문성

○ 분석 대상과 방법에 따른 분석 유형 4가지 * 중요 *

	분석대상 known	분석대상 unknown
분석방법 known	Optimization(최적화) 둘다알때	Insight 분석대상 찾기
분석방법 unknown	Solution 방법을 모르니 찾기	Discovery 둘다 찾아내는 발견

○ 목표 시점별 분석 유형 2가지

- **과제 중심적인 접근방식** : 당면 과제 빠르게 해결 짧고 빠른 해결
 - (1차목표) Speed, Test / (과제유형) Quick, Win / (접근방식) Problem Solving
 - **장기적인 마스터 플랜 방식** : 지속적인 분석 내재화, (+)
 - (1차목표) Accuracy, Deploy / (과제유형) Long Term View / 오랜기간 적용
 - (접근방식) Problem Definition
- 융합하여 적용하는 것이 중요

○ 의미 있는 분석

- 분석기술 + it 및 프로그래밍 + 도메인 전문성 + 의사소통 중요
- 분석대상 및 방식에 따른 다양한 분석 주제를 과제 단위 혹은 마스터플랜 단위로 도출

2. 분석 기획시 고려사항

○ 분석 기획시 고려사항 3가지

- 데이터 available data

- 확보가 우선적, 유형 분석(유형별 적용 가능 솔루션 및 방법 다양)
- 정형 데이터(DB데이터), 비정형 데이터(이메일, 보고서, 소셜미디어 데이터), 반정형 데이터(센서 중심으로 스트리밍되는 머신데이터)

처리프로세스 필요

→ 메타데이터

- 활용방안과 활용 가능한 유즈케이스 탐색 proper business use case

- 기존에 잘 구현되어 활용되고 있는 유사 분석 시나리오 및 솔루션 최대한 활용

- 장애요소들의 대한 사전계획 low barrier of execution

- 조직역량의 내재화를 위하여 충분하고 지속적인 교육 및 활용방안 등의 변화관리 고려

진입장벽 제거

제2절 분석 방법론

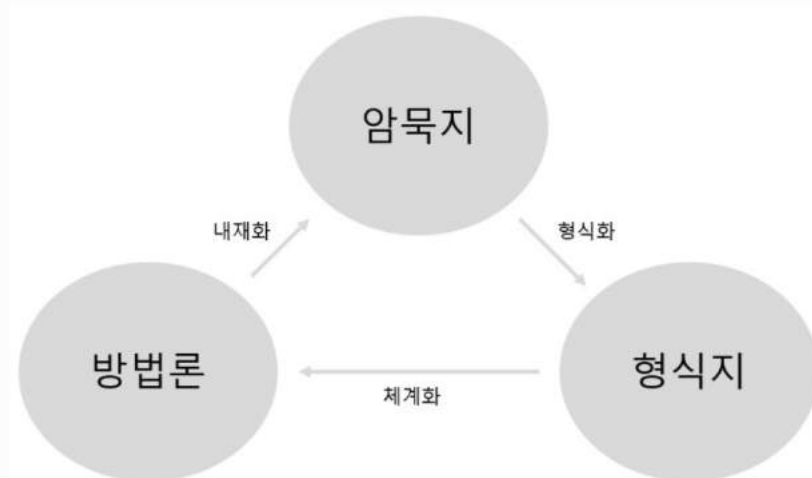
→ KDD, CRISP-DM, 빅데이터분석

1. 분석 방법론 개요

- 데이터를 분석을 체계화된 절차와 방법으로 정리
(데이터 분석이 효과적으로 정착하기 위해 필요)
- 산출물, 프로젝트 성공 가능성을 확보, 제시 가능해야함
- 구성 4가지 : 절차, 방법, 도구와 기법, 템플릿과 산출물

○ 방법론 생성과정

- 암묵지 >(형식화)> 형식지 >(체계화)> 방법론 >(내재화)> 암묵지



○ 방법론 적용 업무의 특성에 따른 모델 3가지

- 폭포수 모델 : 순차적 진행, 이전단계 완료 후 다음단계 진행, 문제발생시 피드백 수행
- 나선형 모델 : 반복을 통해 점증적 개발, 처음 시도하는 프로젝트에 용이, 복잡도 상승 가능성
- 프로토타입 모델 : 폭포수 모델 단점 보완, 일부분만 우선 개발 → 시험사용 → 개선

○ 방법론 구성 : 단계 → 테스크 → 스텝

- 단계 : 단계별 산출물
- 테스크 : 단계 구성하는 단위활동, 물리적·논리적 단위로 품질검토의 항목
- 스텝 : 각 단계는 기준선으로 설정되어 관리, 버전관리를 통해 통제

○ KDD 분석 방법론, CRISP-DM 분석 방법론, 빅데이터 분석 방법론

2. KDD 분석 방법론 (Knowledge Discovery in Databases)

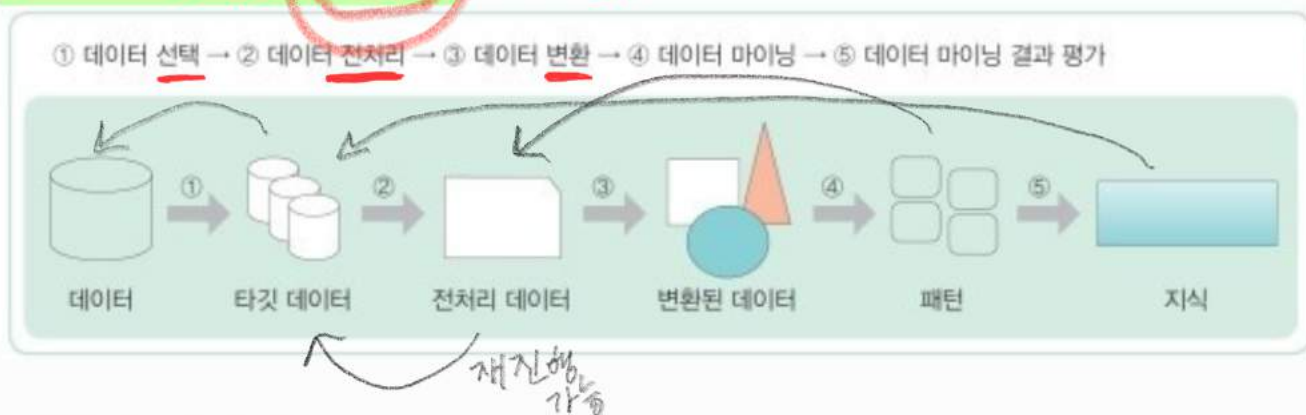
○ 1996년 Fayyad가 프로파일링 기술 기반으로 데이터로부터 통계적

패턴이나 지식을 찾기 위해 활용할 수 있도록 체계적으로 정리한

데이터 마이닝 프로세스

○ 데이터마이닝, 기계학습, 인공지능, 패턴인식, 데이터 시각화 등에서 응용 될 수 있는 구조

○ KDD 분석 절차 5단계 중요!!! 선.전.변



- 1) 데이터셋 선택

데이터마이닝 - 이하 데.마

- 사전에 분석 대상의 비즈니스 도메인 이해와 프로세스 목표 설정 필수
- 데.마에 필요한 목표데이터 구성

- 2) 데이터 전처리 preprocessing

- 정제 : 잡음, 이상치, 결측치 식별·제거·재처리
- 추가 데이터셋이 필요한 경우 데이터 선택 프로세스 재진행

-3) 데이터 변환 transformation

→ 변수 선택!!

- 분석목적에 맞게 변수 생성·선택, 데이터 차원 축소
- 학습용 데이터와 검증용 데이터로 데이터 분리

- 4) 데이터마이닝

→ 분석용 데이터, 테스트용 데이터

- 학습용 데이터 이용, 데.마 기법 선택과 알고리즘 적용하여 데.마 진행
- 필요에 따라 데이터 전처리와 데이터 변환프로세스를 추가로 실행

- 5) 데이터마이닝 결과 평가

- 데.마 결과 해석과 평가, 분석 목적과의 일치성 확인
- 데.마로 얻은 지식을 업무 활용하기 위한 방안 마련
- 필요에 따라 데이터 선택 프로세스~ 데.마 프로세스 반복 수행

3. CRISP-DM 분석 방법론

(Cross Industry Standard Process for Data Mining)

○ 1996년 유럽연합의 ESPRIT에 있었던 프로젝트에서 시작

(5개 업체 주도 : Daimler-Chrysler, SPSS, NCR, Teradata, OHRA)

○ CRISP-DM 프로세스의 4개 레벨

- 단계 Phases : 최상위 레벨은 여러개의 단계로 구성. 각 단계는 일반화 테스크 포함
- 일반화 테스크 Generic Tasks : 데.마의 단일 프로세스를 완전하게 수행하는 단위
- 세분화 테스크 Specialized Tasks : 구체적으로 수행하는 레벨
- 프로세스 실행 Process Instances : 데.마 실행을 위한 구체적인 실행 포함

○ CRISP-DM 프로세스 6단계

vs 선전반
이.준.모



- 업무이해 : 프로젝트 목적과 요구사항 이해단계

- 업무목적 파악, 상황 파악, 데.마 목표설정, 프로젝트 계획 수립

- 데이터 이해 : 데이터 수집, 속성 이해단계

- 초기 데이터 수집, 데이터 기술 분석, 데이터 탐색, 데이터 품질 확인

도메인 이해
+
목표설정

+ 데이터 선택, 전처리

- **데이터 준비** : 분석기법에 적합한 데이터에 편성단계, 시간 다소비 = **데이터 변환**
 · 분석용 데이터 셋 선택, 데이터 정제, **분석용 데이터 셋** 편성, 데이터 통합, 데이터 포매팅
- **모델링** : 다양한 모델링 기법과 알고리즘 선택하고 파라미터 최적화하는 단계,
테스트용 데이터 셋으로 과적합 문제 확인 = **데이터마이닝** 분석용, 테스트용 데이터셋
 · 모델링 기법 선택, 모델 테스트 계획 설계, 모델 작업, **모델 평가**
- **평가** : 모델링 결과가 프로젝트 목적에 부합하는지 평가하는 단계
 · 분석결과 평가, 모델링 과정 평가, **모델 적용성 평가**
- **전개** : 모델을 **실업무에 적용하기 위한 계획을 수립**하고 모니터링과 모델의 **유지보수 계획** 마련 단계
 · 전개 계획 수립, 모니터링과 유지보수 계획 수립, 프로젝트 종료보고서 작성, 프로젝트 리뷰

○ KDD와 CRISP-DM 비교

KDD	CRISP-DM
분석 대상 비즈니스 이해	업무 이해(Business Understanding)
데이터셋 선택(Selection) 데이터 전처리(Preprocessing) 데이터 변환(Transformation)	데이터 이해(Data Understanding) 데이터 준비(Data Preparation)
데이터 마이닝(Data Mining)	모델링(Modeling)
데이터 마이닝 결과 평가 (Interpretation / Evaluation)	평가(Evaluation)
데이터 마이닝 활용	전개(Deployment)

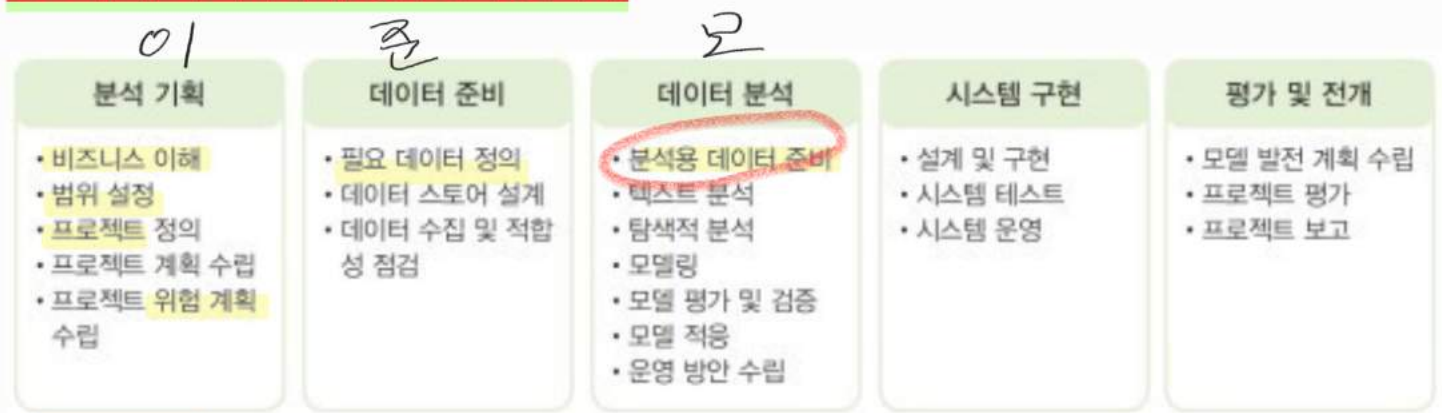
*** 아주 중요, 자주 출제 ***

4. 빅데이터 분석 방법론 (계층적 프로세스 모델)

○ 빅데이터 분석을 위한 3개 계층 단계 → 테스크 → 스텝

- 단계 : 프로세스 그룹을 통하여 완성된 **단계별 산출물** 생성
- 테스크 : 단계를 구성하는 단위 활동, 물리적 또는 논리적 단위로 **품질 검토의 항목** 가능
- 스텝 : WBS의 워크 패키지 해당, 단위 프로세스(입력자료, 처리 및 도구, 출력 자료로 구성)

○ 빅데이터 분석 방법론 5단계



- **분석 기획** : 비즈니스 이해 및 범위 설정 / 프로젝트 정의 및 계획 수립 / 프로젝트 위험계획 수립
- **데이터 준비** : 필요 데이터 정의 / 데이터 스토어 설계 / 데이터 수집 및 정합성 검증
- **데이터 분석** : 분석용 데이터 준비 / 텍스트 분석 / 탐색적 분석 / 모델링 / 모델 평가 및 검증 / 모델 적용 및 운영방안 수립
- **시스템 구현** : 설계 및 구현 / 시스템 테스트 및 운영
- **평가 및 전개** : 모델 발전계획 수립 / 프로젝트 평가 및 보고

1) 분석 기획 * 출력자료를 중점으로

○ 비즈니스 이해 및 범위 설정

- 비즈니스 이해

- 입력 자료 : 업무 매뉴얼, 전문가 지식, 빅데이터 분석 대상 도메인 관련 자료
- 프로세스 및 도구 : 자료 수집 및 비즈니스 이해
- 출력 자료 : 비즈니스 이해 및 도메인 문제점

- 프로젝트 범위 설정

- 입력 자료 : 중장기 계획서, 빅데이터 분석 프로젝트 지시서, 비즈니스 이해 및 도메인 문제점
- 프로세스 및 도구 : 자료 수집 및 비즈니스 이해, 프로젝트 범위 정의서 작성 절차
- 출력자료 : 프로젝트 범위 정의서(SOW)

○ 프로젝트 정의 및 계획 수립

- 데이터 분석 프로젝트 정의

- 입력자료 : 프로젝트 범위 정의서, 빅데이터 분석 프로젝트 지시서
- 프로세스 및 도구 : 프로젝트 목표 구체화, 모델 운영 이미지 설계
- 출력자료 : 프로젝트 정의서, 모델 운영 이미지 설계서, 모델 평가 기준 → WBS 작성

○ 프로젝트 위험계획 수립

- 데이터 분석 위험 식별

· 입력자료 : 프로젝트 범위 정의서, 프로젝트 수행 계획서, 선행 프로젝트 산출물 및 정리자료

· 프로세스 및 도구 : 위험 식별 절차, 위험영향도·발생가능성 분석, 위험 우선순위 판단

· 출력자료 : 식별된 위험 목록

- 위험 대응 계획 수립

· 입력자료 : 식별된 위험 목록, 프로젝트 범위 정의서, 프로젝트 수행 계획서

· 프로세스 및 도구 : 위험 정량적 분석, 위험 정성적 분석

· 출력자료 : 위험관리 계획서(회피, 전이, 완화, 수용으로 구분)

2) 데이터 준비 = 데이터 이해, 준비

○ 필요 데이터 정의

- 데이터 수집 및 저장

· 입력자료 : 정의프로젝트 수행 계획서, 시스템 설계서, ERD, 메타정형데이터 정의서, 문서자료

· 프로세스 및 도구 : 내·외부, 정형·비정형·반정형 데이터 정의

· 출력자료 : 데이터 정의서

- 데이터 획득방안 수립

· 입력자료 : 데이터 정의서, 시스템 설계서, ERD, 메타데이터 정의서, 문서자료, 데이터 구입

· 프로세스 및 도구 : 데이터 획득 방안 수립

· 출력자료 : 데이터 획득 계획서

○ 데이터 스토어 설계

- 정형 데이터 스토어 설계 : 일반적으로 관계형 데이터베이스인 RDBMS 사용

· 입력자료 : 데이터 정의서, 데이터 획득 계획서

· 프로세스 및 도구 : DB 논리설계/물리설계(구분이유:효율적 저장과 활용), 데이터 매핑

· 출력자료 : 정형 데이터 스토어 설계서, 데이터 매핑 정의서

- 비정형 데이터 스토어 설계 (하둡, NoSQL 이용)

· 입력자료 : 데이터 정의서, 데이터 획득 계획서

· 프로세스 및 도구 : 비정형·반정형 데이터베이스 논리설계/물리설계

· 출력자료 : 비정형 데이터 스토어 설계서, 데이터 매핑 정의서

○ 데이터 수집 및 정합성 점검

- 데이터 수집 및 저장

- 입력자료 : 데이터 정의서, 데이터 획득 계획서, 데이터 수집 스크립트
- 프로세스 및 도구 : 데이터 크롤링 도구, ETL 도구, API, 데이터 수집 스크립트(데이터 수집)

· 출력자료 : 수집된 분석용 데이터, 데이터 스토어에 저장

- 데이터 정합성 점검

- 입력자료 : 수집된 분석용 데이터
- 프로세스 및 도구 : 데이터 품질확인, 데이터 정합성 점검 리스트
- 출력자료 : 데이터 정합성 점검 보고서(주기적 확인)

3) 데이터 분석 = 모델링 = 데이터마이닝

○ 분석용 데이터 준비

- 비즈니스 룰 확인

- 입력자료 : 프로젝트 정의서, 프로젝트 수행계획서, 데이터 정의서, 데이터 스토어
- 프로세스 및 도구 : 프로젝트 목표 확인, 비즈니스 룰 확인
- 출력자료 : 비즈니스 룰, 분석에 필요한 데이터 범위

- 분석용 데이터셋 준비

- 입력자료 : 데이터 정의서, 데이터 스토어
- 프로세스 및 도구 : 데이터 선정/변환, ETL 도구
- 출력자료 : 분석용 데이터셋(작업공간(play ground, sand box), 전사 차원 데이터 스토어로 분리가능)

○ 텍스트 분석

- 텍스트 데이터 확인 및 추출

- 입력자료 : 비정형 데이터 스토어
- 프로세스 및 도구 : 분석용 텍스트 데이터 확인, 텍스트 데이터 추출
- 출력자료 : 분석용 텍스트 데이터

- 텍스트 데이터 분석 : 분석하고 모델 구축

- 입력자료 : 분석용 텍스트 데이터, 용어사전(용어 유의어 사전, 불용어 사전 등)
- 프로세스 및 도구 : 분류체계 설계, 형태소 분석, 키워드 도출, 토픽·감정·의견·네트워크 분석

· 출력자료 : 텍스트 분석 보고서(시각화 사용)

○ 탐색적 분석

- 탐색적 데이터 분석 = EDA

- 입력자료 : 분석용 데이터셋
- 프로세스 및 도구 : EDA도구, 통계분석, 변수간 연관성 분석, 데이터 분포 확인
- 출력자료 : 데이터 탐색 보고서(모델링을 위한 기초 자료)

→ 데이터의 통계적 특성

- 데이터 시각화

- 입력자료 : 분석용 데이터셋
- 프로세스 및 도구 : 시각화 도구 및 패키지, 인포그래픽, 시각화 방법론
- 출력자료 : 데이터 시각화 보고서(시스템 구현을 위한 사용자 인터페이스, 프로토타입으로 활용 가능)

○ 모델링

- 데이터 분할

- 입력자료 : 분석용 데이터셋
- 프로세스 및 도구 : 데이터 분할 패키지
- 출력자료 : 훈련용 데이터, 테스트용 데이터 (모델의 과적합과 일반화를 위하여)

홀드아웃
K-fold

- 데이터 모델링

- 입력자료 : 분석용 데이터셋
- 프로세스 및 도구 : 통계 모델링 기법, 기계학습, 모델 테스트 (훈련용 데이터를 이용해 모델을 만든 후 운영 시스템에 적용)

→ 모델링 용

- 출력자료 : 모델링 결과 보고서

- 모델 적용 및 운영방안

- 입력자료 : 모델링 결과 보고서
- 프로세스 및 도구 : 모니터링 방안 수립 알고리즘 설명서 작성
- 출력자료 : 알고리즘 설명서 (운영시스템에 적용하기 위해), 모니터링 방안(안정적 운영을 위해)

시스템 구현 단계에서 중요한 입력 자료
주관식@ 의사코드 수준의 자세한 작성 필요

○ 모델 평가 및 검증

- 모델평가

→ 모델 평가용 (과적합 등...)

- 입력자료 : 모델링 결과 보고서, 평가용 데이터 (테스트용 데이터나 별도사용)
- 프로세스 및 도구 : 모델 평가, 모델 품질관리, 모델 개선사업
- 출력자료 : 모델 평가 보고서

- 모델 검증

- 입력자료 : 모델링 결과 보고서, 모델 평가 보고서, 검증용 데이터 (실 운용 데이터)
- 프로세스 및 도구 : 모델 검증
- 출력자료 : 모델 검증 보고서

5 : 3 : 2

→ 실제로 활용 능력 검증

4) 시스템 구현 → 해도 되고 안해도 된다

○ 설계 및 구현

- 시스템 분석 및 설계

- 입력자료 : 알고리즘 설명서, 운영 중인 시스템 설계서
- 프로세스 및 도구 : 정보시스템 개발방법론(커스터마이징)
- 출력자료 : 시스템 분석 및 설계서

- 시스템 구현

- 입력자료 : 시스템 분석 및 설계서, 알고리즘 설명서
- 프로세스 및 도구 : 시스템 통합 개발도구(IDE), 프로그램 언어, 패키지
- 출력자료 : 구현 시스템

⇒ 구현 및 운영(메뉴얼, 유지보수)

○ 시스템 테스트 및 운영

- 시스템 테스트

- 입력자료 : 구현 시스템, 시스템 테스트 계획서
- 프로세스 및 도구 : 품질관리 활동(객관성, 완전성 확보)
- 출력자료 : 시스템 테스트 결과보고서

주관식@

시스템 테스트는 품질관리 차원에서 진행하여
시스템의 객관성, 안정성 확보

- 시스템 운영 계획

- 입력자료 : 시스템 분석 및 설계서, 구현 시스템
- 프로세스 및 도구 : 운영계획 수립, 운영자 및 사용자 교육
- 출력자료 : 운영자 매뉴얼, 사용자 매뉴얼, 시스템 운영 계획서(지속적 활용을 위해)

5) 평가 및 전개

○ 모델 발전 계획 수립

- 모델 발전 계획

- 입력자료 : 구현 시스템, 프로젝트 산출물
- 프로세스 및 도구 : 모델 발전 계획 수립(모델의 지속성, 기능향상)
- 출력자료 : 모델 발전 계획서(모델의 계획성 확보)

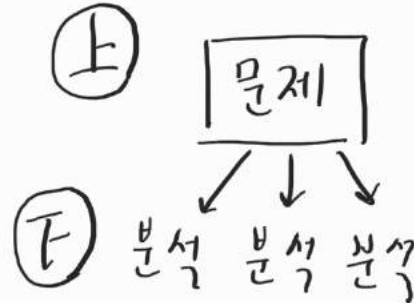
○ 프로젝트 평가 및 보고

- 프로젝트 성과 평가

- 입력자료 : 프로젝트 산출물, 품질관리 산출물, 프로젝트 정의서, 프로젝트 수행 계획서
- 프로세스 및 도구 : 프로젝트 평가 기준, 프로젝트 정량적 평가, 프로젝트 정성적 평가
- 출력자료 : 프로젝트 성과 보고서

- 프로젝트 종료

- 입력자료 : 프로젝트 산출물, 품질관리 산출물, 프로젝트 정의서, 프로젝트 수행 계획서, 프로젝트 성과 보고서
- 프로세스 및 도구 : 프로젝트 지식자산화 작업, 프로젝트 종료
- 출력자료 : 프로젝트 최종 보고서



제3절 분석 과제 발굴

- 과제 발굴 2가지 방식 : 하향식 접근 방식, 상향식 접근 방식
- **하향식 접근 방식** : 문제가 주어지고 해결하기 위해 각 과정이 체계적으로 단계화 되어 수행하는 방식 문제 존재 → 데이터 분석
- **상향식 접근 방식** : 문제의 정의가 어려운 경우, 데이터 기반으로 문제의 재정/해결방안 탐색하고 지속적으로 개선하는 방식 데이터 분석
- 실제분석 과정에서 혼용되어 활용경우가 많다 (**디자인 사고 프로세스**) → 문제발굴
Design Thinking process

1. 하향식 접근법(Top Down Approach)

○ 하향식 접근법 4단계

- 문제 탐색 - 데이터 문제로 정의 - 해결방안 탐색 - 분석의 타당성 평가

1) 문제 탐색 단계

○ 비즈니스 모델기반 문제 탐색

- **비즈니스 모델 캔버스를 활용한 과제 발굴을 위한 5가지 영역**

- 비즈니스 모델 캔버스의 9가지 블록 단순화 + 관리영역

(업무, 제품, 고객 영역) + (규제와 감사, 지원 인프라 영역)

- **업무** : 생산하기 위해 운영하는 내부 프로세스 및 주요 자원 관련 주제 도출
- **제품** : 제품·서비스 개선하기 위한 관련 주제 도출
(제품 주요 기능 개선, 서비스 모니터링 지표도출 등)
- **고객** : 고객, 채널의 관점에서 주제 도출(Call 대기 시간 최소화, 영업점 위치 최적화)
- **규제와 감사** : 생산 및 전달과정 프로세스 중 발생하는 규제 및 보완 관점에서 주제 도출(품질 이상 징후 관리, 새로운 규제 시 예상되는 제품)
- **지원 인프라** : 분석하는 시스템, 인력 관점 주제 도출(EDW 최적화, 적정 운영 인력 도출 등)

- 새로운 관점의 접근을 통해 새로운 유형 분석 기회 발굴

· **거시적 관점(STEEP)**에서 비즈니스 분석하는 5가지 영역

- * **사회(Social) 영역** : 현재 고객을 확장하여 전체 시장을 대상으로 사회적, 문화적, 구조적 트렌드 변화에 기반한 분석 기회 도출(노령화, 저출산에 따른 해당 사업 모델의 변화)
- * **기술(Technological)영역** : 최신 기술 등장 및 변화에 따른 역량 내재화와 제품·서비스 개발에 대한 분석 기회도출(나노 기술, 제품의 Smart화, IT 융합 기술 등)
- * **경제(Economic)영역** : 산업과 금융 전반 변동, 경제 구조 변화에 따른 시장 흐름을 통해 분석 기회 도출(원자재 가격/환율/금리 변동에 따른 구매 전략의 변화 등)
- * **환경(Environmental)영역** : 탄소 배출 규제 및 거래 시장 등장에 따른 원가 절감 및 정보 가시화 등
- * **정치(Political)영역** : 대북관계 동향에 따른 원자재 구매 거래 선의 다변화 등

· **경쟁자 확대 관점에서 분석 기회 발굴 3가지 영역**

- * **대체제 영역** : 자사 제품·서비스를 온라인으로 제공하는 것에 대한 탐색 및 잠재적 위험 파악
- * **경쟁자 영역** : 식별된 주요 경쟁사의 제품·서비스 카탈로그 및 전략 분석하고 이에 대한 잠재적 위험 파악
→ 아예 써먹을!
- * **신규 진입자 영역** : 파괴적인 역할을 수행할 수 있는 신규진입자 동향 파악

· **시장의 니즈 탐색 관점에서 분석 기회 발굴 3가지 영역**

- * **고객 영역** : 고객 구매 동향 및 컨텍스트 이해하여 제품·서비스 개선에 필요한 분석 기회 도출(철강기업은 조선 산업과 자동차 산업탐색)
- * **채널 영역** : 자체 운영 채널 + 경로에 존재하는 채널별로 분석 기회 탐색(은행은 인터넷 전문은행 등 온라인 채널 탐색)
- * **영향자들 영역** : 주주·투자자·협회 등
(M&A 시장 확대에 따른 유사 업종의 신규 기업 인수 기회 탐색 등)

· **역량의 재해석 관점에서 분석 기회 발굴 2가지 영역**

- * **내부 역량 영역** : 지적 재산권, 기술력, 지식·스킬·노하우 재해석, 분석 기회 탐색
(자사 소유 부동산을 활용한 부가가치 창출 기회 발굴 등)
자사가
→ 가진 것
- * **파트너와 네트워크 영역** : 관계사, 공급사 역량 활용을 통해 분석 기회 도출 → 주변
(수출입·통관·노하우를 활용한 추가 사업기회 탐색 등) 회사 연결

○ 외부 참조 모델 기반 문제 탐색

- 유사 동종 사례 벤치마킹을 통한 분석기회 발굴 (산업별, 서비스별 분석 테마 후보 그룹(=pool)에서 Quick and Easy 방식으로 얻고 기업에 적용할 분석 테마 후보 목록을 브레인스토밍을 통해 빠르게 도출)
- 가치 발굴 사례를 풀로 만들어 놓으면 과제 발굴 및 탐색 시 빠르고 의미 있는 분석 기회 도출 가능

비즈니스 모델 및 외부참조모델 -> 분석유즈케이스 -> 구체적인 과제로 정의

○ 분석 유즈 케이스 정의

- 도출한 분석 기회들을 구체적인 과제로 만들기 전 분석 유즈 케이스 정의
- 풀어야할 문제에 대한 상세한 설명 및 해당 문제를 해결시 발생하는 효과 명시
- 향후 데이터 분석 문제로의 전환 및 적합성 평가에 활용

- 분석 유즈 케이스 예시

- 재무 업무 : 자금 시재 예측(문제설명), 구매 최적화(효과)
- 고객 업무 : 서비스 수준 유지, 고객만족 달성
- 판매 업무 : 파이프라인 최적화, 영업성과 분석

2) 문제 정의 단계

- 식별된 비즈니스 문제를 데이터의 문제로 변환하여 정의하는 단계
- 필요한 데이터 및 기법(HOW)을 정의하기 위한 데이터 분석의 문제로의 변환
- 고객 이탈 증대 -> 이탈 영향을 미치는 요인 식별, 이탈 가능성 예측으로 변환
- 재고관리 및 적정 가격 어려움 -> 수요 예측 가능한 인자 추출 및 모델링을 통한 수요 예측으로 변환

3) 해결방안 탐색 단계

- 기존 정보시스템의 단순한 보완으로 분석이 가능한지 고려
- 엑셀 등 간단한 도구로 분석이 가능한지 고려
- 하둡 등 분산병렬처리를 활용한 빅데이터 분석 도구를 통해 보다 체계적이고 심도 있는 방안 고려

		분석 역량 (Who)	
		확보	미확보
분석 기법 및 시스템 (HoW)	기존 시스템	기존 시스템 개선 활용	교육 및 채용을 통한 역량 확보
	신규 도입	시스템 고도화	전문 업체 Sourcing

시스템과 역량 여부에 따른 4가지 해결방안