

230924 7기 SQL반 수업

BDA 7기 SQL반 수업 자료

데이터베이스 사용자

03 데이터베이스 사용자

❖ 데이터베이스 사용자

- 데이터베이스를 이용하기 위해 접근하는 모든 사람
- 이용 목적에 따라 데이터베이스 관리자, 최종 사용자, 응용 프로그래머로 구분

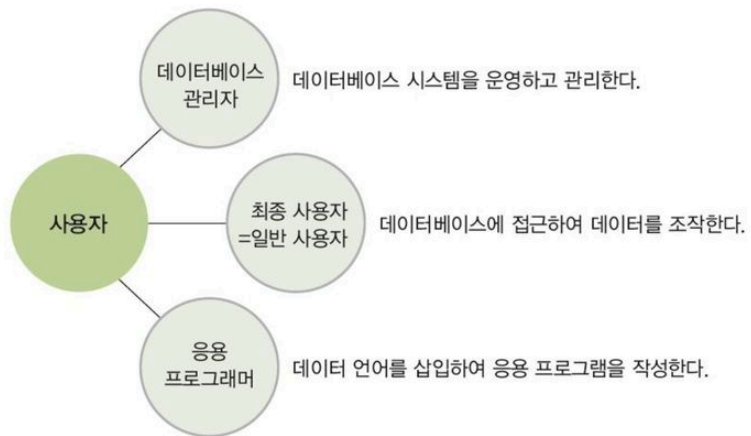


그림 3-7 데이터베이스 사용자

데이터 베이스 관리자

데이터베이스 관리자 DBA (DataBase Administrator)

데이터베이스 구성 요소 선정

우리 조직에 필요한 데이터는?

단순하게 DB를 select * from abc where 단순히 쿼리만 날리는 게 아니라 데이터베이스를 설계 시스템을 운영 관리 설계랑 구축

데이터베이스 스키마 정의

컬럼의 구성

DBA 관리자의 경우는 관리라는 key -> 운영, 유지보수 설계 구축에도 같이 참여한다.

물리적 저장 구조와 접근 방법 결정

데이터 베이스 설계 할 때 배울 예정이지만 논리적, 물리적 데이터 베이스 설계하고 어떤 시스템으로 접근할지를 다양하게 선택 할 수 있다. 빅쿼리, AWS, snowflake, 하둡 등 이런 것들 결정 조직 차원에서는 다방면으로 고민, 비용도 고민해야 함

무결성 유지를 위한 제약조건 정의

데이터가 중복, 잘못된 정보 절대 안 된다. 제약 조건 정확성, 유효성 유지

보안 및 접근 권한 정책 결정

백업 및 회복 기법 정의

시스템 데이터베이스 관리

시스템 성능 감시 및 성능 분석

기존에 KPI 지표가 고객 유입률 -> 유입률에 대한 정보만 확인하고 있는데 -> KPI 이탈율로 바뀜 -> 이탈율을 관리 이탈율에 대한 데이터를 만들지 않았다면, 이탈율을 만들어야 한다.

데이터베이스 재구성

아주 귀찮은 작업

최종 사용자, 응용 프로그래머

엔드 포인트에서 데이터 결과를 확인하거나 자기가 원하는 것만 보면 된다.

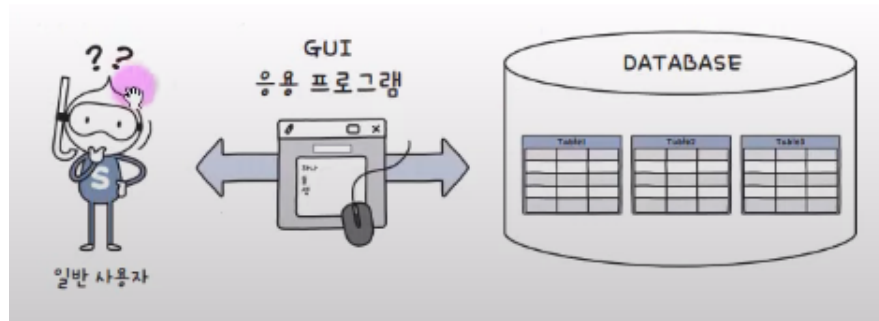
- 최종 사용자는 데이터 분석가 또는 다른 부서 실무진들
- 응용프로그래머 -> SQL등을 이용해서 응용프로그램을 만들고 더욱 필요한 데이터를 쉽게 찾을 수 있도록 만듦

SQL 연동할 수 있는 부분이 많다.

Python 연동, snowflake 머신러닝 이 부분 연동, 태블로 연동, AWS 환경 쿼리 연동

e.g) 마케터가 태블로 시각화로 데이터를 보고싶다 -> 고객 AARRR 지표를 시각화해 주세요!
쿼리를 통해 데이터를 정리하고 -> 태블로 넣고 대시보드화 -> 공유하는 과정

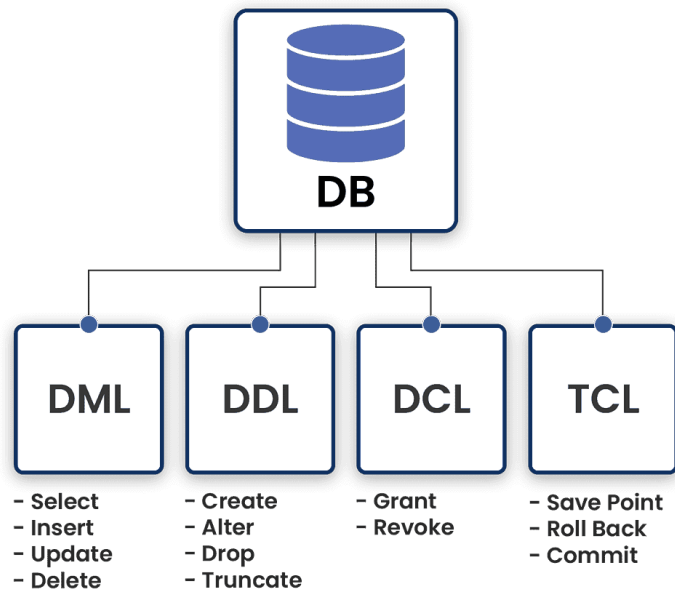
Results		Messages				
	id	employee_id	customer_id	start_time	end_time	call_outcome_id
1	1	1	4	2020-01-11 09:00:15.000	2020-01-11 09:12:22.000	2
2	2	1	2	2020-01-11 09:14:50.000	2020-01-11 09:20:01.000	2
3	4	1	1	2020-01-11 09:24:15.000	2020-01-11 09:25:05.000	3
4	5	1	3	2020-01-11 09:26:23.000	2020-01-11 09:33:45.000	2
5	6	1	2	2020-01-11 09:40:31.000	2020-01-11 09:42:32.000	2
6	8	1	1	2020-01-11 09:42:32.000	2020-01-11 09:46:53.000	3
7	3	2	3	2020-01-11 09:02:20.000	2020-01-11 09:18:05.000	3
8	7	2	4	2020-01-11 09:41:17.000	2020-01-11 09:45:21.000	2
9	9	2	1	2020-01-11 09:46:00.000	2020-01-11 09:48:02.000	2
10	10	2	2	2020-01-11 09:50:12.000	2020-01-11 09:55:35.000	2



데이터 언어

스키마에 대해서 쿼리로 수정하거나 삭제하거나 등등 가능

- DDL 데이터 정의어 : 스키마 정의, 수정 삭제
- DML 데이터 조작어 : 데이터 삽입, 수정, 삭제, 검색 등 처리 요구 데이터 추출 등 다양한 기본적인 쿼리 대표적인 select 문
- DCL 데이터 제어어 : 내부적 규칙, 기법 정의
내부에 필요한 규칙이나, 권한 등등
- TCL : 트랜잭션을 제어하는 명령어
쿼리를 작성하고 실제 프로그램이 진행되면서 일의 가장 작은 프로세스



데이터 언어

데이터 정의어

스키마 구조, 제약조건 변경, 삭제 등을 삭제
내가 원하는 table을 조합해서 만들 수도 있고, 여러가지로 데이터베이스를 구축하기 위한 언어

데이터 조작어

내가 원하는 데이터를 추출할 경우에 많이 사용
인스턴스 (값) 값들을 출력하는 경우
전체 고객 데이터 -> 8월의 A제품을 구매한 고객의 평균 방문빈도 ?
Python 충분히 가능하다.

절차적 데이터 조작어 : 구체적으로 알려준다.
비절차적 데이터 조작어 : 큰 지시만 내리는 상황
e.g) 8월의 A제품을 구매한 고객의 평균 방문빈도 ? 쿼리를 작성할 때 정해진 건 없다. 최소한의 쿼리문이나, 작동시간, 원리를 이해하고 최대한 효율적으로 작성할 순 있다. 정답이 있는 건 아니다.

1. 8월 데이터만 추출하고 데이터베이스로 만들고 -> 고객에대해서 접근하고 -> A제품 보고
2. 바로 현재 테이블에서 쿼리문으로 조건들을 정리해서 데이터를 뽑을 수도 있다.

데이터 제어어

무결성, 일관성 등등 중요한 부분
규칙과 기법 -> 다양한 사람들이 사용할 예정 약속의 개념으로 정의 -> 스키마 정의, 데이터 베이스 접근 권한 등
시스템 사용에 대해서 정리를 하는 것
-데이터베이스 관리자나 권한이 많은 경우에만 부여하거나 이런 식으로 진행
대부분 사람들이 제어어를 접근할 순 없다.

데이터베이스 관리 시스템의 구성

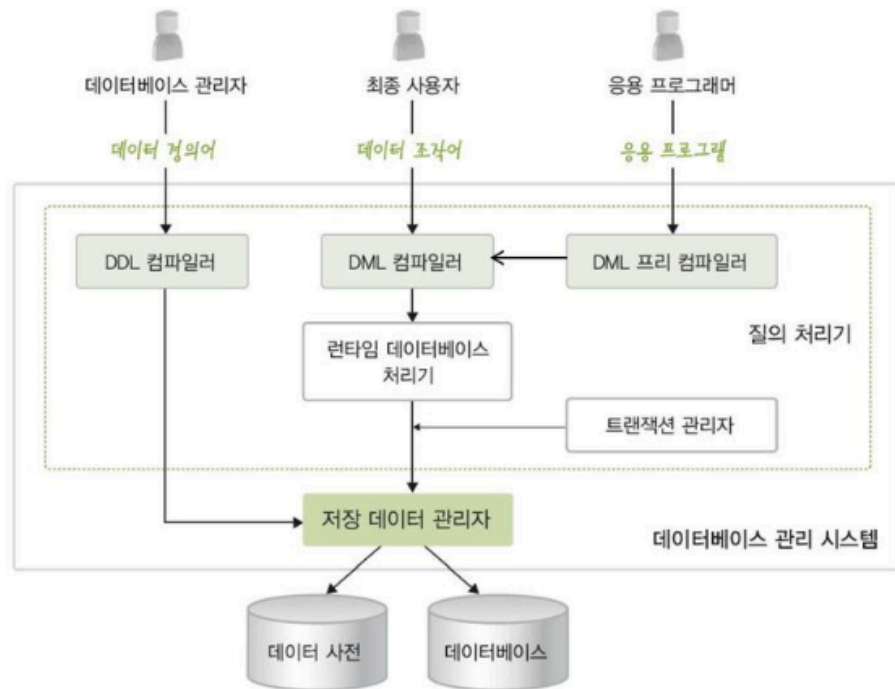


그림 3-11 데이터베이스 관리 시스템의 구성

컴파일러? (Compiler) 번역기의 개념

특정 프로그래밍 언어를 -> 다른 프로그래밍 언어로 바꿔주는 것

데이터 모델링

서점

서점에는 발생할 수 있는 비즈니스가 정말 다양하다.

교보문고

온라인

오프라인

데이터 모델

문고 큰 개념으로 잡으면

도서 고객 주문하고 그 도서를 받는 과정

현실의 문제를 -> 추상화하면서 컴퓨터가 이해할 수 있게 만드는 과정

이 프로세스를 데이터베이스화 하려면

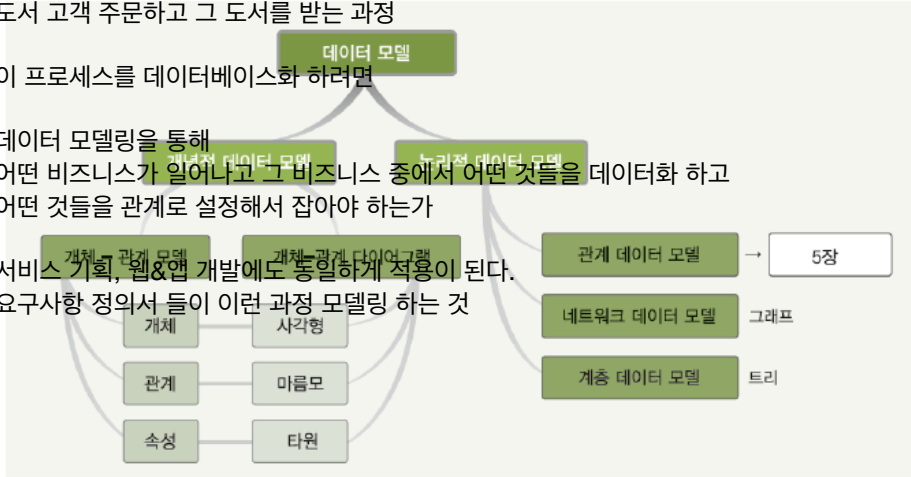
데이터 모델링을 통해

어떤 비즈니스가 일어나고 그 비즈니스 중에서 어떤 것들을 데이터화 하고

어떤 것들을 관계로 설정해서 잡아야 하는가

서비스 기획, 웹&앱 개발에도 동일하게 적용이 된다.

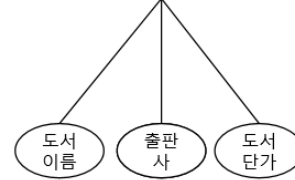
요구사항 정의서 등이 이런 과정 모델링 하는 것



- ▶ 데이터 모델링과 데이터 모델의 개념을 이해한다.
- ▶ 개념적 데이터 모델인 개체 - 관계 모델을 이용해 모델링을 하는 방법을 익힌다.
- ▶ 개체 - 관계 모델을 개체 - 관계 다이어그램으로 작성하는 방법을 익힌다.
- ▶ 논리적 데이터 모델의 종류와 특징을 이해한다.

도서 테이블

도서이름 / 출판사 / 도서단가
abc / 한빛 / 15,000



고객 테이블

고객이름 주소 전화번호
홍길동 서울 01001111

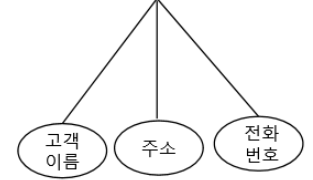


그림 6-5 개념적 모델링의 예

고객은 -> 주문을 통해 도서를 받는
다.

주문 테이블

주문일자 / 주문고객 / 주문상품

환불

취소

케이스가 발생하면

이 케이스를 어떤 식으로 데이터베이스로 만드는가?

데이터 모델링

데이터 모델링이란 무엇인가요?

데이터 모델링은 조직의 정보 수집과 관리 시스템을 정의하는 시각적 표현 또는 청사진을 생성하는 프로세스입니다. 이 청사진 또는 데이터 모델은 데이터 분석자, 과학자, 엔지니어와 같은 다양한 이해관계자들이 조직의 데이터에 대한 통일된 개념을 생성할 수 있게 돕습니다. 이 모델은 해당 비즈니스가 수집하는 데이터, 서로 다른 데이터 세트 사이의 관계, 데이터를 저장하고 분석하는 데 사용되는 방식을 설명합니다.

데이터 모델링이 중요한 이유는 무엇인가요?

오늘날 조직은 다양한 소스에서 많은 양의 데이터를 수집합니다. 그러나 원시 데이터로는 충분하지 않습니다. 수익성 있는 비즈니스 결정을 내리는 데 도움이 되는 실행 가능한 인사이트를 얻기 위해 데이터를 분석해야 합니다. 정확한 데이터 분석을 위해서는 효율적인 데이터 수집, 저장 및 처리가 필요합니다. 여러 데이터베이스 기술과 데이터 처리 도구가 있으며 데이터 세트마다 효율적인 분석을 위한 도구가 다릅니다.

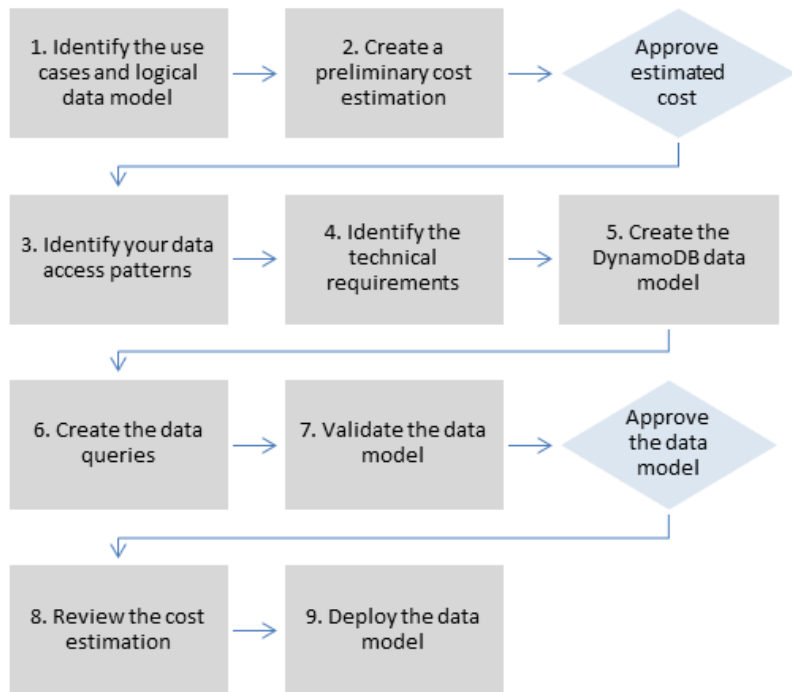
데이터 모델링은 데이터를 이해하고 이 데이터를 저장 및 관리하기 위한 올바른 기술 선택을 할 수 있는 기회를 제공합니다. 건축가가 집을 짓기 전에 청사진을 설계하는 것과 같은 방식으로, 비즈니스 이해관계자는 조직을 위한 데이터베이스 솔루션을 엔지니어링하기 전에 데이터 모델을 설계합니다.

데이터 모델링은 다음과 같은 이점을 제공합니다.

- 데이터베이스 소프트웨어 개발 오류 감소
- 데이터베이스 설계 및 생성 속도와 효율성 촉진
- 조직 전체에서 데이터 문서화 및 시스템 설계의 일관성 조성
- 데이터 엔지니어와 비즈니스 인텔리전스 팀 간의 커뮤니케이션 촉진

데이터를 수집해야 하는데 어떤 데이터를 수집해야 하나?
비즈니스 관계를 잘 이해하고

AWS 모델링 예시



데이터 모델

- 추상화 작업
- 현실에 있는 모든 상황 또는 존재하는 것을 모델링 하는 것은 쉽지 않음

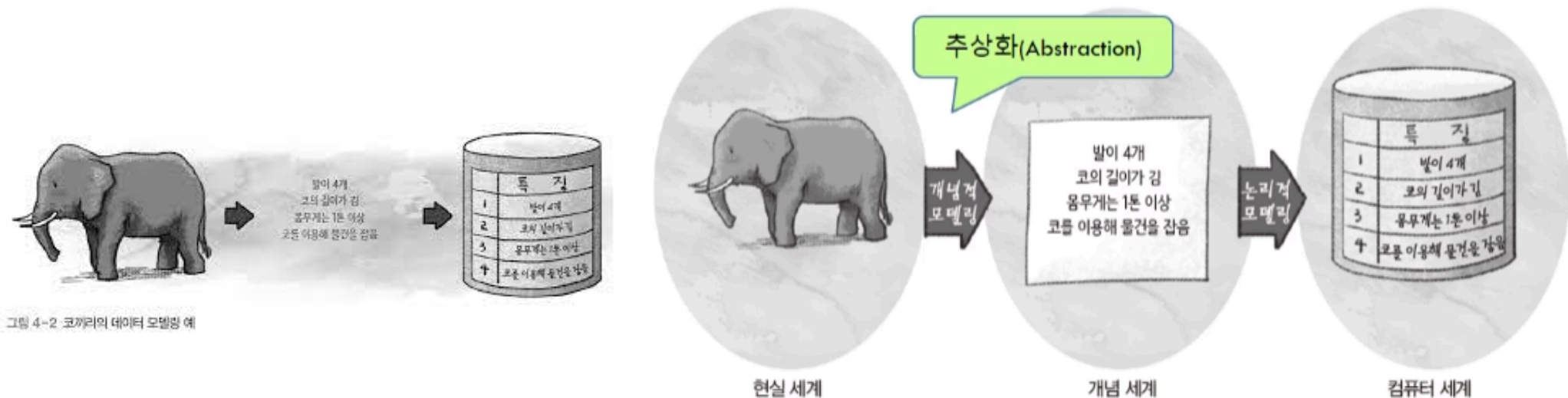


그림 4-2 코끼리의 데이터 모델링 예

그림 4-3 코끼리의 2단계 데이터 모델링 예

자동차 대리점 사례로 예시를 들면?

자동차 대리점에서 비즈니스를 하게 되는데 우리가 실제 데이터베이스를 만들어야 한다.

모델링을 진행해야 하고

우선적으로 개념적 모델링을 진행해야 하는데 어떤식으로 접근할까?

고객은 자동차대리점에서 주문하고 자동차를 직접 받는다.

대리점에는 판매직원에 대한 기준

고객과 판매한 대리점사원은 연결되어 있어야 한다.

자동차에는 여러 브랜드와 차량 가격

모든 고객의 정보 중 이메일은 필수적 있어야 한다.

• 개념적 모델링은?

자동차번호 자동차 재료 -> 크게 보기에는 자동차라는 큰 개념으로 접근하면 세분화 할 수 있는 부분
•
조금 더 넓은 의미에서 고민을 해야 한다.

자동차 Cars 테이블 -> 엔터티

고객 Customers 테이블 -> 엔터티

대리점 정보 테이블 -> 엔터티

대리점 직원 정보 테이블 -> 엔터티

판매 정보 테이블 -엔터티

• 논리적 모델링은?

스키마에 대한 정보까지도 고민하면서 테이블을 설계해야 한다.
•
티들의 관계들을 통해서 생성될 수 있는 테이블 고민해야 한다.

이메일로 접근을 한다면

고객에 대한 이메일은 소문자와숫자만 사용가능, @모든 소문자.com,kr 최대 길이는 50자 또는 100자

주문 테이블에 주문금액은 정수형 데이터로 만든다.

주문 테이블에 들어간 주문날짜들은 ->날짜 형식으로 만들어진다.

대리점에 존재하는 직원들의 수는 정수형이 10명미만으로 표현

Car

차량/ 수량/ 제조연도 / 출시일 (스키마 구조에 맞춰서)

데이터 베이스가 쌓아지게 되는 것

고객

고객ID/ 고객이메일/ 고객 휴대전화번호 (- -온리 숫자 ,길이) 스키마구조

주문

주문일 / 주문고객/ 주문차량 (스키마구조)

데이터 베이스가 될 예정

물리적 모델링 DBMS 데이터베이스 내에 해당 테이블이 만들어지는 것

데이터베이스가 만들어지는 과정

개체, 관계에 대한 모델을 다음시간 부터 배울 예정

개체, 속성 관계 이런 것들을 실제 비즈니스나 사례에서 발생하는 것을 모델링해서

데이터 베이스로 만드는 과정

E-R 다이어그램

관계에 대한 차수

데이터가 1:N 다대다 관계인지 1:1 관계인지 이런 것들이 중요해 진다.

PK FK 유니크한 값을 식별할 수 있는 것이 무엇인지?

어떤 컬럼이 PK 될 것인가

필수과제 1.

그동안 배웠던 개념들을 본인만의 노트(필기 또는 수기로 필기 등)로 정리해서 업로드

• PDF파일로 만들어 주시고 그동안 배운 개념 정리 잘 부탁드립니다.

• 지금까지 배운 전체 내용

• 기한은 과제출 기한은 동일

• 복습과제는 필수과제로 보시면 됩니다.

• 시험기간은 정리할 예정

•