SQLD 시험 및 수업 요약

1. 데이터 모델링의 이해
   1. 데이터 모델링의 이해
      1. (논리적 – 물리적) 엔티티 – 테이블, 속성 – 열, 인스턴스 – 행
   2. 데이터 모델의 이해
      1. 모델링
         1. 추상적이고 구조화된 형태로 표현
         2. 특징
            1. 단순화 – 불필요한 세부사항 제거
            2. 추상화 – 간략하게 대략적으로 표현
            3. 명확화 – 애매모호함 제거 및 정확하게 현상 기술
         3. 데이터 모델링 3가지 관점
            1. 데이터 관점 – 정적
            2. 프로세스 관점 – 동적
            3. 데이터와 프로세스 관점 – 통합
         4. 데이터 모델링 유의점
            1. 중복
            2. 비유연성 – 사소한 변화에 잦은 모델 변화가 발생하지 않도록
            3. 비일관성
         5. 데이터 모델링 3요소
            1. 대상 - entity
            2. 속성 - attribute
            3. 관게 – relationship
         6. 데이터 모델링 3단계
            1. 개념적 모델링 – 추상화
            2. 논리적 모델링 – 세부 내역
            3. 물리적 모델링 – 논리적 모델링 이후 실제 구현 – 추상화 낮음
         7. 스키마의 3단계 구조
            1. 스키마 – 구조와 제약 조건에 관한 전반적인 명세를 기술한 데이터의 집합
            2. 외부 스키마 – 사용자 관점 데이터베이스 스키마 정의
            3. 개념 스키마 – 외부 스키마를 통합하여 데이터베이스의 전체 논리 구조를 정의( 개체 , 속성, 관계, 데이터 타입 등 정의)
            4. 내부 스키마 – 데이터가 물리적으로 어떻게 저장되는지 정의
            5. \* 독립성

논리적 독립성 – 논리적 데이터 구조가 변경되어도(개념 스키마 변경) – 응용 프로그램 영향 적음

물리적 독립성 – 물리적 구조가 변경되어도 개념/외부 스키마에 영향을 주지 않음

* + - 1. ERD
         1. 대상(entity – table) 들의 관계를 시각적으로 표현
         2. 작성 절차

엔티티를 그린다

엔티티 배치

엔티티 관계 설정

관계명 서술

참여도 기술

필수 여부 확인

1. 엔티티
   1. 엔티티의 개념
      1. 독립적 식별 가능한 객체나 사물을 나타냄
      2. 업무상 분석해야 하는 대상들로 이루어진 집합
      3. 엔티티와 속성, 인스턴스 등의 관계
   2. 엔티티의 특징
      1. 유일한 식별자에 의해 식별 가능
      2. 해당 업무에 필요하고 관리하고자 하는 정보
      3. 인스턴스들의 집합
      4. 엔티티는 반드시 속성을 가짐
      5. 엔티티는 업무 프로세스에 의해 이용
      6. 다른 엔티티와 최소 1개 이상의 관계 성립
   3. 엔티티의 분류
      1. 유형과 무형에 따른 분류
         1. 유형 엔티티 – 물리적 형태가 있음
         2. 개념 엔티티 – 물리적인 형태 없음
         3. 사건 엔티티 – 업무를 수행에 따라 발생하는 엔티티
      2. 발생 시점에 따른 분류
         1. 기본 엔티티 – 그 업무에 원래 존재하는 정보
         2. 중심 엔티티 – 기본 엔티티로부터 발생되고 그 업무에서 중심적인 역할
         3. 행위 엔티티 – 2개 이상의 부모 엔티티로부터 발생
   4. IE & 바커 표기법
2. 속성
   1. 속성의 개념
      1. 업무에서 필요로 하는 고유한 성질, 특징을 의미(관찰 대상) -> 컬럼으로 표현할 수 있는 단위
      2. 더 이상 분리되지 않는 최소의 데이터 단위
   2. 엔티티, 인스턴스, 속성, 속성값의 관계
      1. 한 개의 엔티티는 2개 이상의 인스턴스의 집합이어야 한다.
      2. 한 개의 엔티티는 2개 이상의 속성을 갖는다.
      3. 한 개의 속성은 1개의 속성값을 갖는다.
      4. 속성은 엔티티에 속한 엔티티에 대한 자세하고 구체적인 정보를 나타냄 -> 각 속성은 구체적인 값을 가짐
   3. 속성의 특성
      1. 해당 업무에서 반드시 필요하고 관리하고자 하는 정보여야 한다.
      2. 정해진 주식별자에 함수적 종속성을 가져야 한다.
      3. 하나의 속성은 반드시 하나의 속성값을 가진다. -> 속성의 원자성
   4. 함수적 종속성
      1. 어떤 속성 A의 값에 의해 다른 속성 B도 유일하게 결정된다면, B는 A에 함수적으로 종속됐다 하고, A -> B
      2. 완전 함수적 종속
         1. 구성하는 컬럼이 2개 이상일 경우, PK값 모두에 의한 종속관계를 나타낼 때 완전 함수 종속성 만족
      3. 부분 함수적 종속
         1. 기본키 전체가 아니라, 기본키 일부에 대해 종속될 때
   5. 속성의 분류
      1. 속성의 특성에 따른 분류
         1. 기본 속성 : 업무로부터 추출된 모든 속성
         2. 설계 속성 : 기본 속성 외의 규칙화 혹은 기본 속성을 변경하여 만들어지는 속성
         3. 파생 속성 : 다른 속성에 의해 만들어지는 속성
      2. 엔티티 구성방식에 따른 분류
         1. PK – 인스턴스를 식별할 수 있는 속성
         2. FK – 다른 엔티티 관계에서 포함된 속성
         3. 일반 속성 - 엔티티에 포함되어 있고 PK/FK에 포함되지 않는 속성
      3. 분해 여부에 따른 속성
         1. 단일 속성 – 하나의 의미로 구성
         2. 복합 속성 – 여러 개의 의미로 구성
         3. 다중값 속성 – 속성에 여러 개의 값을 가질 수 있는 경우
   6. 도메인
      1. 각 속성이 가질 수 있는 값의 범위
      2. 데이터 타입과 크기, 제약사항을 지정
3. 관계
   1. 관계의 개념 : 엔티티 간의 연관성
   2. 관계의 종류
      1. 존재적 관계 – 한 엔티티의 존재가 다른 엔티티의 존재에 영향을 미치는 관계
      2. 행위적 관계 – 엔티티 간의 어떤 행위가 있는 것을 의미
      3. ERD에서는 관계를 구분하지 않음
   3. 관계의 구성
      1. 관계명
      2. 차수
      3. 선택성
   4. 관계의 차수 Cardinality
      1. 1:1
         1. 완전 1대 1 관계 – 하나의 엔티티에 관계되는 엔티티가 반드시 하나
         2. 선택 1대 1 관계 – 관계되는 엔티티가 하나 혹은 없음
      2. 1:N
         1. 엔티티의 하나의 행에 다른 엔티티의 값이 여러 개 있음
      3. N:M
         1. 두 엔티티가 다대다의 연결 관계를 가지고 있음
   5. 관계의 페어링 – 엔티티 안의 인스턴스가 개별적으로 관계를 가지는 것
      1. 관계의 차수 : 하나의 엔티티와 다른 엔티티 간의 레코드 연결 방식
      2. 관계의 페어링 : 두 엔티티 간의 특정 연결을 설명
4. 식별자
   1. 개념
      1. 하나의 엔티티에 구성된 여러 속성 중에 엔티티를 대표할 수 있는 속성
      2. 하나의 유일한 식별자가 존재
      3. 식별자(논리 모델링), 키(물리 모델링)
   2. 주식별자의 특징
      1. 유일성
      2. 최소성 – 유일성을 만족하는 최소한의 속성
      3. 불변성 – 한번 지정 후 변경 x – 항상 고유값
      4. 존재성 – NULL 허용 x
   3. 식별자 분류
      1. 대표성 여부에 따른 식별자의 종류
         1. 주 식별자
         2. 보조 식별자 : 대표성이 없음
      2. 생성 여부에 따른 식별자의 종류
         1. 내부 식별자 : 엔티티 내부에서 스스로 생성되는 식별자
         2. 외부 식별자 : 다른 엔티티와 관계로 인하여 만들어지는 식별자
      3. 속성 수에 따른 식별자의 종류
         1. 단일식별자
         2. 복합식별자
      4. 대체 여부에 따른 식별자의 종류
         1. 본질식별자 – 비즈니스 프로세스에서 만들어지는 식별자
         2. 인조식별자 – 인위적으로 만들어지는 식별자
   4. 관계 간 엔티티 구분
      1. 강한 개체 – 독립적으로 존재 가능
      2. 약한 개체
   5. 식별 관계와 비식별 관계
      1. 식별 관계: 하나의 엔티티가 자신의 기본키를 다른 엔티티가 기본키의 하나로 공유하는 관계 – 실선
      2. 비식별 관계: 강한 개체의 기본키를 다른 엔티티에서 기본키가 아닌 일반 속성으로 관계를 가짐 – 점선
   6. 키의 종류
      1. 기본키 – 엔티티를 대표할 수 있는 키
      2. 후보키 – 유일성과 최소성을 만족하는 키- 후보키 중 하나가 기본키, 나머지는 대체키
      3. 슈퍼키 – 유일성은 만족하지만 최소성은 만족하지 않는 키
      4. 대체키 – 여러 후보키 중 기본키가 아닌 키
      5. 외래키 – 다른 테이블의 기본키를 참조하는 키
5. 정규화
   1. 정규화의 개념
      1. 최소한의 데이터만을 하나의 엔터티에 넣는 식으로 데이터를 분해하는 과정
      2. 중복을 제거하고 데이터 모델의 독립성을 확보
      3. 이상현상을 줄이기 위한 데이터베이스 설계 기법
      4. 논리 데이터 모델링 수행 시점에서 고려됨
   2. 이상현상
      1. 정규화를 하지 않아 발생하는 현상( 삽입, 갱신, 삭제 이상)
      2. 삽입이상 – 특정 인스턴스 삽입 시 정의되지 않아도 될 속성까지 반드시 입력되어야 하는 현상
      3. 삭제이상 – 특정 정보 삭제 시 삭제되지 않아도 될 정보까지 삭제됨
   3. 정규화 단계
      1. 제1 정규화(1NF)
         1. 테이블의 컬럼이 원자성을 갖도록 테이블을 분해
         2. 하나의 행과 컬럼의 값이 반드시 한 값만 입력되도록 행을 분리
         3. \* 인스턴스에 두 개 이상의 값이 입력되어 있을 때, 값 하나만 입력되록 만들어줌
      2. 제2 정규화(2NF)
         1. 제 1 정규화를 만족하는 테이블에 대해 완전 함수 종속을 만들도록 테이블을 분해
         2. 완전 함수 종속 : 기본키를 구성하는 모든 컬럼의 값이 다른 컬럼을 결정짓는 상태
         3. PK가 2개 이상일 때 발생하며 PK의 일부와 종속되는 관계가 있다면 분리한다.
      3. 제3 정규화(3NF)
         1. 제 2정규화를 진행한 테이블에 대해 이행적 종속을 없애도록 테이블을 분리
         2. A->B, B->C일 때, A->C가 성립되는 상황에서 (A,B) , (B,C)로 테이블을 분리하는 것
      4. BCNF 정규화
         1. 모든 결정자가 후보키가 되도록 테이블을 분해
      5. 제4 정규화
         1. 여러 컬럼들이 하나의 컬럼을 종속시키는 경우, 분해하여 다중 값 종속성을 제거
      6. 제5 정규화
         1. 조인에 의해서 종속성이 발생되는 경우 분해
6. 관게와 조인의 이해
   1. 관계의 개념
      1. 엔티티의 인스턴스 사이의 논리적인 연관성
   2. 관계의 분류
      1. 존재 관계 : 엔티티 간의 상태 ex)사원 엔티티는 부서 엔티티에 소속
      2. 행위 관계 : 엔티티 간에 어떤 행위가 존재함 ex) 주문은 고객이 주문할 때 발생
   3. 조인의 의미
      1. 테이블은 데이터의 중복을 피하기 위해 정규화됨, 분리되면서 두 테이블을 관계를 가지게 되고, 이 테이블들을 동시에 출력하거나 관계가 있는 테이블을 출력하기 위해 데이터를 연결해야 하는데 이를 조인 이라고 함
   4. 계층형 데이터 모델
      1. 하나의 엔티티 내의 인스턴스끼리 계층 구조를 가짐
      2. 계층 구조를 갖는 인스턴스끼리 연결하는 조인을 셀프 조인이라고 함
   5. 상호 배타적 관계
      1. 두 테이블 중 하나만 가능한 관계 ex) 주문 엔티티의 개인 혹은 법인번호 엔티티 둘 중 하나만 사용될 수 있음
7. 모델이 표현하는 트랜잭션의 이해
   1. 트랜잭션이란?
      1. 하나의 연속적인 업무 단위
         1. A가 B에게 100만원 이체
            1. A의 잔고가 100 이상인지 확인
            2. 100 이상이면 A의 잔고 -100
            3. B의 잔고에 +100
         2. A->C로 진행되는 과정을 자체를 트랜잭션 -부분 commit 불가
   2. 필수적, 선택적 ERD
      1. 두 엔티티의 관계가 필수적 -> 하나의 트랜젝션 형성
         1. IE > 아무것도 안함
         2. 바커 > 실선
      2. 두 엔티티 독립적 -> 선택적 관계
         1. IE > 원을 그림
         2. 바커 > 점선
8. NULL 속성의 이해
   1. NULL이란
      1. NULL 아직 정해지지 않는 값
      2. 모델 설계 시 컬럼별로 허용할지 결정
   2. NULL 특성
      1. NULL을 포함한 연산 결과는 항상 NULL
      2. 그룹함수는 NULL을 제외하고 연산을 수행한다.
         1. SUM, AVG, MIN, MAX 등의 그룹 함수는 항상 NULL을 무시하고 연산
   3. NULL의 ERD 표기법
      1. IE 표기법에서 NULL의 표기 방식 없음
      2. 바커 표기법에서 속성 앞에 동그라미가 NULL 허용 속성을 의미 -> IE 선택적 관계랑 같이 생각하자
9. 본질 식별자와 인조 식별자
   1. 식별자 구분 – 대체 여부에 따른
      1. 본질 식별자
         1. 업무에 의해 만들어지는 식별자
         2. 반드시 필요함
      2. 인조 식별자
         1. 인위적으로 만들어지는 식별자
         2. 반드시 필요하지 않지만 관리의 편의성 등의 이유로 인해 인위적으로 만들어지는 식별자
         3. 행 구분 기본키로 사용 – 자동 증가하는 일련번호 gudx
         4. 인조 식별자의 단점
            1. 중복 데이터의 발생 가능
            2. 불필요한 인덱스 생성