**2021학년도 2학기 중간과제물(온라인제출용)**

**교과목명 : 데이터처리와 활용**

**학 번 : 202135-367895**

**성 명 : 김태정**

**연 락 처 : 010-4172-4516**

o 과제유형 : (공통) 형

o 과 제 명 : 중간과제물 문제

- 이하 과제 작성

**1번문제**

**1.1번문제 – 개체와 개체 타입**

개체(Entity)란 실세계의 식별 가능한 유형적이거나 무형적인 사물을 의미한다. 개체는 속성들의 집합이며 반드시 한 개 이상의 속성을 가져야한다. 속성은 애트리뷰트(Attribute)라고도 한다. 이 때 애트리뷰트를 모은 개체를 개체 타입(Entity Type)이라고한다.

**1.2번문제 – 개체 무결성**

테이블 의 기본 키를 구성하는 애트리뷰트들은 유일한 값을 식별해야하는 의무를 가진다. 그런데 만약 기본 키를 구성하는 애트리뷰트가 null값이라면 유일하게 식별하는 기능을 하지 못하므로 이를 막는다. 이러한 규칙을 개체 무결성(Entity Integrity Rule)이라고 한다.

**1.3번문제 – 데이터베이스 질의어**

관계형 데이터베이스의 데이터에 접근하기 위해서는 질의(Query)를 해야한다. 이러한 질의는 일종의 프로그래밍언어로 되어있으며 이를 SQL(Structured Query Language)라고 한다. SQL을 이용하여 저료의 검색, 관리, 수정, 조작, 제어등이 가능하다.

**1.4번문제 – 스키마**

데이터베이스에서 자료의 구조, 자료의 표현방법, 자료간의 관계를 형식 언어 형태로 정의한 것을 스키마라고한다. 스키마는 3층 구조로 되어 있으며 외부 스키마, 내부 스키마, 개념 스키마로 나뉜다.

**1.5번문제 – 도메인**

개체 타입(Entity Type)은 애트리뷰트(Attribute)의 집합이다. 여기서 애트리뷰트들은 특정한 값을 취하게 되는데 이 때 애트리뷰트가 가질 수 있는 값의 집합을 도메인(domain)이라고 한다.

**1.6번문제 – 튜플**

관계형 데이터베이스는 열과 행으로 이루어져있다. 열의 경우 속성의 집합으로 애트리뷰트라고한다. 행의 경우 특정 개체타입의 인스턴스(Instance)를 가지고 있다. 이 인스턴스를 다른 말로 레코드(Record)혹은 튜플(Tuple)이라고 한다.

**1.7번문제 – 슈퍼 키**

테이블에서 각각의 튜플을 구별하기 위해서는 튜플을 유일하게 식별하기 위한 애트리뷰트 집합이 필요하다. 여기서 어트리뷰트 집합이기에 한개일 수도 있고 복수개일 수도 있다. 이러한 어트리뷰트의 집합을 키(Key) 혹은 슈퍼키(Super Key)라고 한다.

**1.8번문제 – 후보 키**

테이블에서 튜플을 유일하게 구별할 수 있는 애트리뷰트의 집합을 슈퍼키라고 하는데 이 경우 여러개의 슈퍼키가 존재할 수 있다. 가령 (학번)이 슈퍼키일 경우 학번을 포함한 (학번, 날짜)와 (학번, 성별) 역시 슈퍼키가 될 수 있다. 이 때 키 식별에 필요한 최소한의 애트리뷰트의 집합을 후보 키(Candidate Key)라고 한다.

**1.9번문제 – 기본 키**

테이블에서 후보 키가 1개 이상 존재하는데 이 중에서 유일하게 식별하기로 하나의 후보키를 선택할 수 있고 이를 기본 키(Primary Key)라고 한다. 여기서 선택되지 못한 나머지 후보 키를 대체 키(Alternate Key)라고 한다.

**1.10번문제 – 외래 키**

두 테이블을 연결하는데 사용하는 키를 외래 키(Foreign Key)라고한다. 외래키의 대상이 되는 테이블의 키는 반드시 존재해야하며 이를 참조 무결성 규칙(Referential Integrity Rule)이라고 한다.

**2번문제**

**2.1번문제**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 대출 | 대출번호(PK) | 대출액 |  |  |
| 지급 | 지급번호(PK) | 대출번호(FK) | 지급날짜 | 지급금액 |

지급은 약한개체 타입이기 때문에 대출이 없으면 존재할 수 없다. 따라서 대출의 대출번호를 외래키로 대출에 연관이 되어있는 관계형 스키마를 작성할 수 있다.

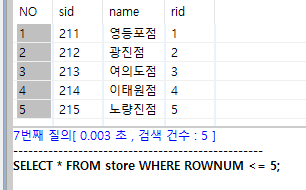
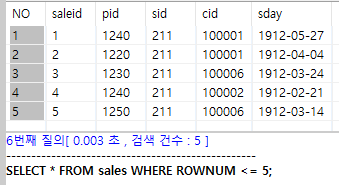
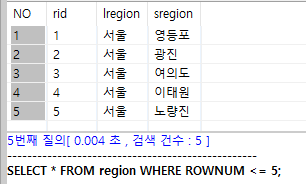
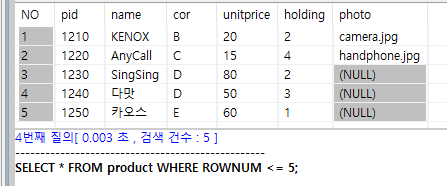
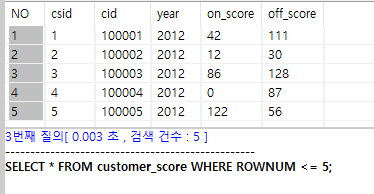
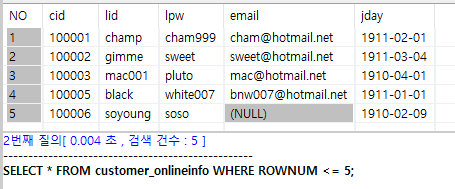
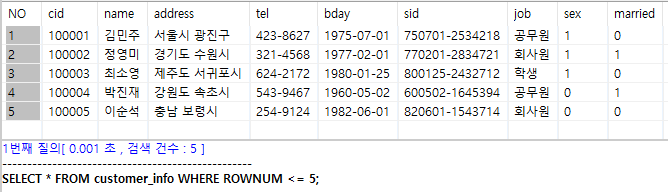
**2.2번 문제**

|  |  |  |  |
| --- | --- | --- | --- |
| 주민 | 주민번호(PK) | 이름 | 주소 |
| 고객 | 주민번호(PK,FK) | 신용도 |  |
| 근로자 | 주민번호(PK,FK) | 월급 |  |
| 임직원 | 주민번호(FK) | 임직원번호(PK) |  |
| 컨설턴트 | 주민번호(PK,FK) | 초과수당 | 주근무일수 |
| 비서 | 주민번호(PK,FK) | 담당임직원 |  |

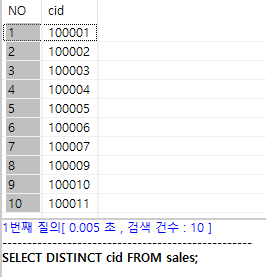
주민번호는 이견의 여지가 없다. ER 다이어그램만 봐서는 여러가지 방향으로 생각해 볼 수 있다. 가령 고객과 근로자를 다른 식별자(고객번호, 근로자번호)를 PK로 두는 방법이 있으나 약한 타입개체라고 가정하고 주민번호를 그대로 사용할 수 있다. 또한 그렇게 생각하는게 이 문제상 자연스럽다. 다만 임직원과 컨설턴트와 비서는 근로자의 하위인데 이를 알게 하려면 근로자 번호같은 식별번호를 두어야한다. 여기서는 이 3종류가 근로자임을 알아야하는지 몰라도 상관없는지가 중요한 쟁점이 될것인데 문제이서는 표현되지 않으므로 그냥 근로자임을 스키마로서 정의할 필요가 없다고 가정한다. 마지막으로 임직원의 경우 주민번호를 그대로 이어받아서 사용해도 되고 임직원 번호를 PK로 사용해도 된다. 일관성을 위해서는 주민번호를 이어받아 식별관계로 해도 되었으나 임직원번호를 PK로 사용하는게 임직원 테이블의 취지에 맞는것 같아서 임직원 번호를 PK로 선택하였다.

**3번문제**

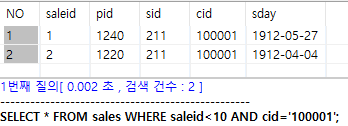
**3.1번문제**



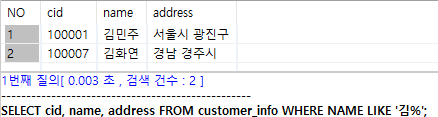
**3.3번문제**



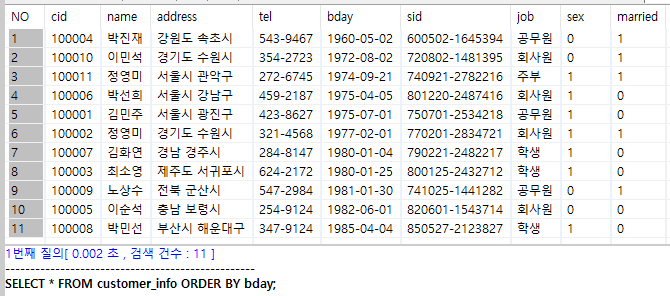
**3.5번문제**



**3.7번문제**



**3.9번문제**



**3.10번문제**

