1. SQL - (빅데이터분석)최적화 된 도구(R/파이썬)를 활용한 애널리스트 양성과정

**2021/03/11 목요일 수업**

IPv4 공유주소: 192.168.0.102

윈도우+R키 누르고 [\\192.168.0.102](file:///\\192.168.0.102) 엔터

PART2: SQL 기본과 활용

DB vs DBMS

테이블Table의 구조

행(Row): Record라고도 함 = “튜플”이 정식 명칭

열(Column) = 필드(Field) = “속성”이 정식 명칭

행 + 열 = 테이블

테이블에 따라 제약조건 별로 기본키, 외부키 등으로 나눔

SQL: Structured Query Language

오라클

Oracle 11g r2 download (Oracle Database Software Downloads)

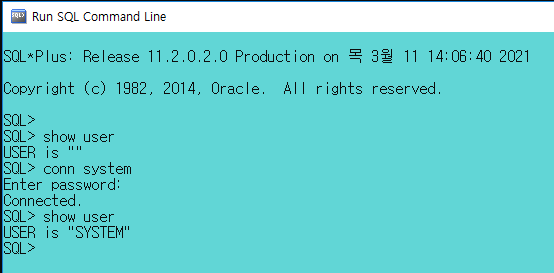
Oracle Database 11gR2 Express Edition for Windows x64

PW: 111111

Window+E -> 내컴퓨터 마우스 오른쪽 버튼 클릭 -> 속성클릭 -> 시스템종류 확인 32비트 or 64비트

윈도우 키 누르고 검색 -> Run SQL Command Line 입력

Run SQL Command Line



USER is “SYSTEM”으로 시스템권한을 가져야 쓸 수 있다.

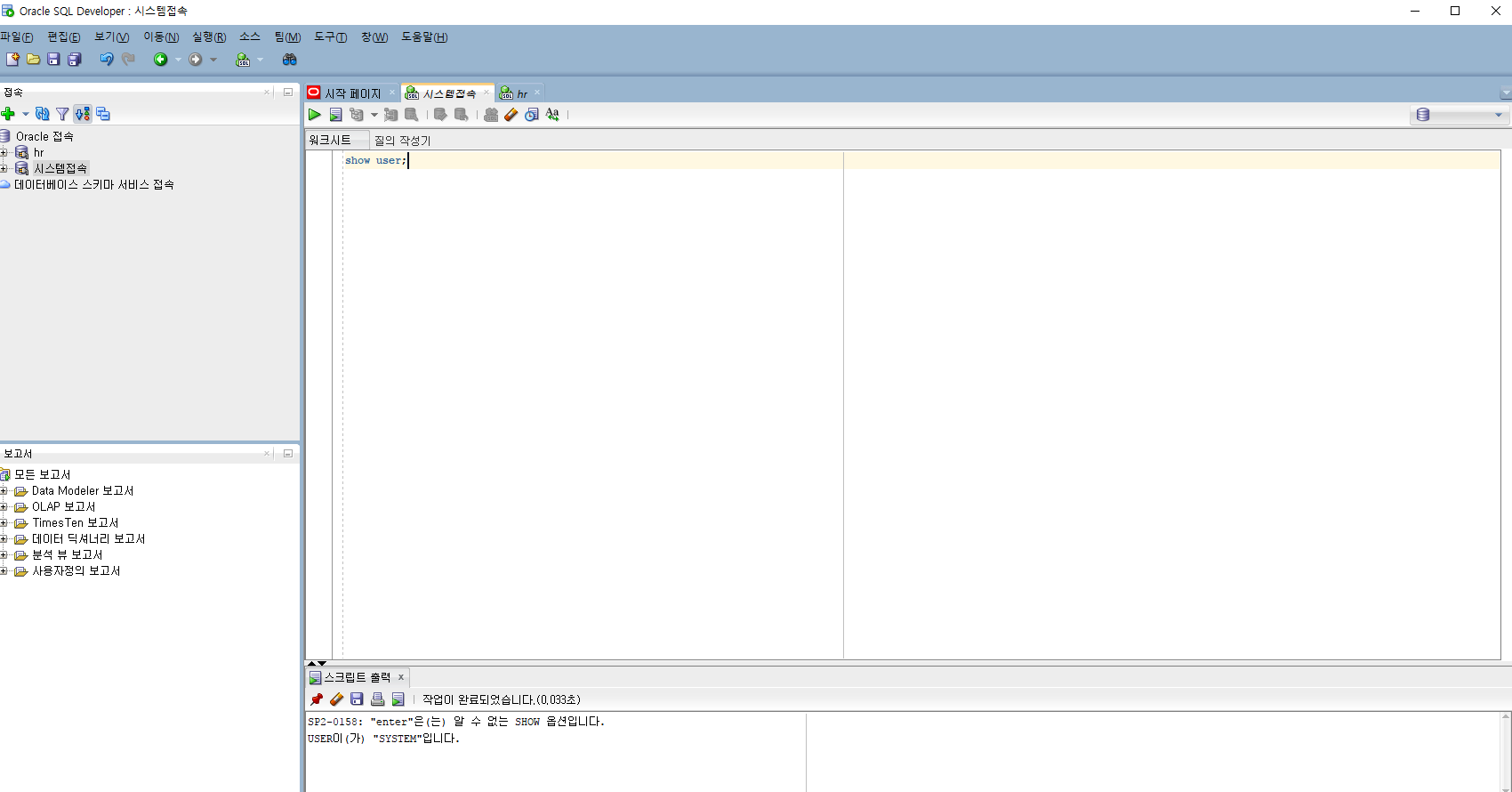
Select\*from tab; 엔터 누르면 Run SQL Command Line에서 그래프를 볼 수 있음

비밀번호를 바꾸고 싶을 때: Alter USER hr account unlock identified by

Alter USER hr account unlock identified by 123456;

“hr이라는 유저의 잠금을 풀고 그 계정의 비밀번호는 123456으로 변경(alter)하겠다.”라는 의미

Show user 누르고 conn system 입력 후 마무리



인터넷 sqldeveloper download 검색

Oracle SQL Developer Downloads: Windows 64-bit with JDK 8 included

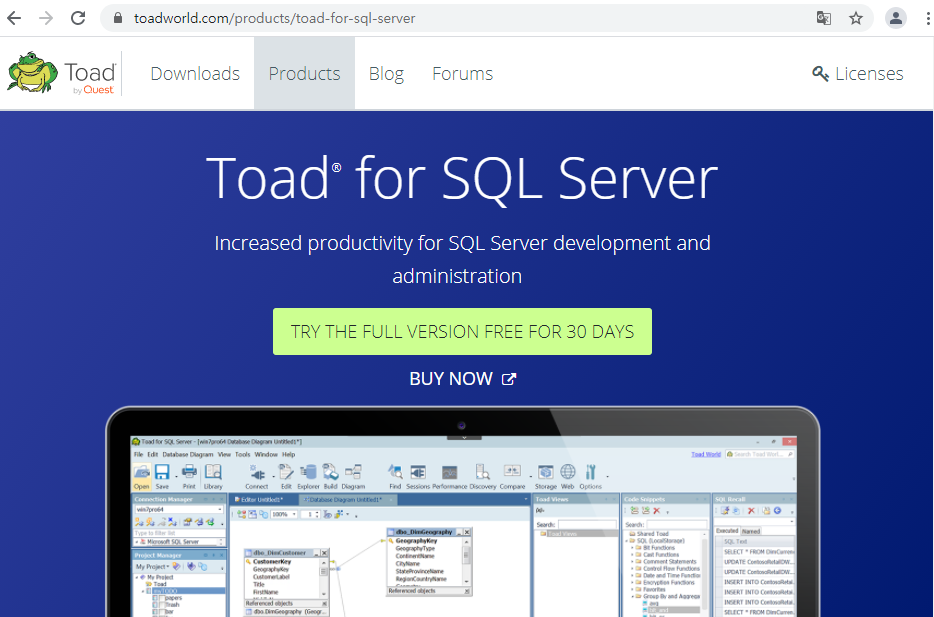
with JDK 8 included -> 컴퓨터에 자바가 설치되어 있지 않아도 괜찮게 기능들을 포함해놓았다는 뜻

다운받은 후 편하게 쓰기 위해서는 저장용량 D드라이브에 settingfiles파일을 만든 뒤, 압축을 푸는 것이 좋음

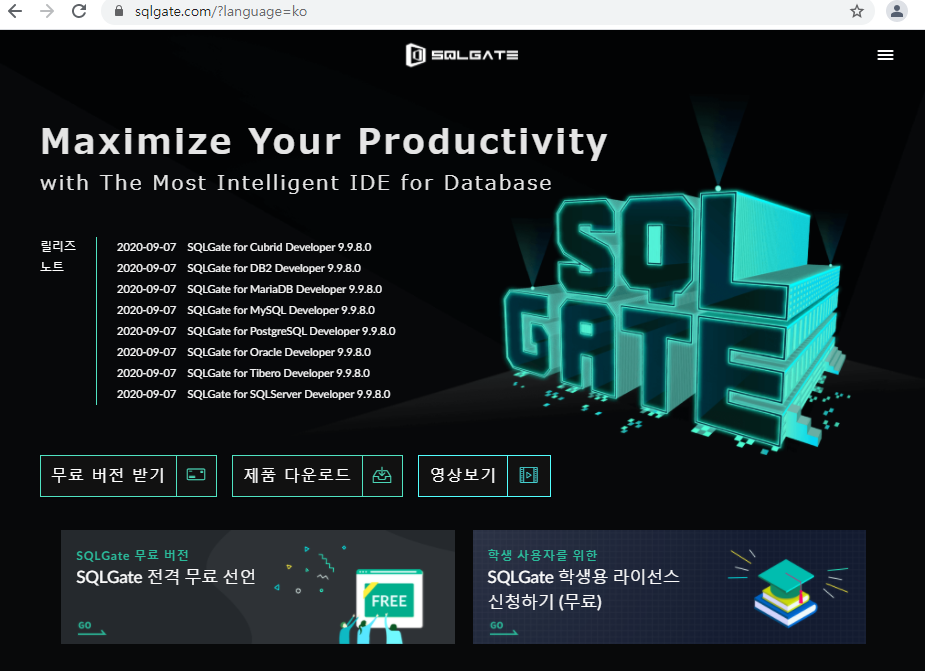
데이터베이스 관리

Toad, sqlgate

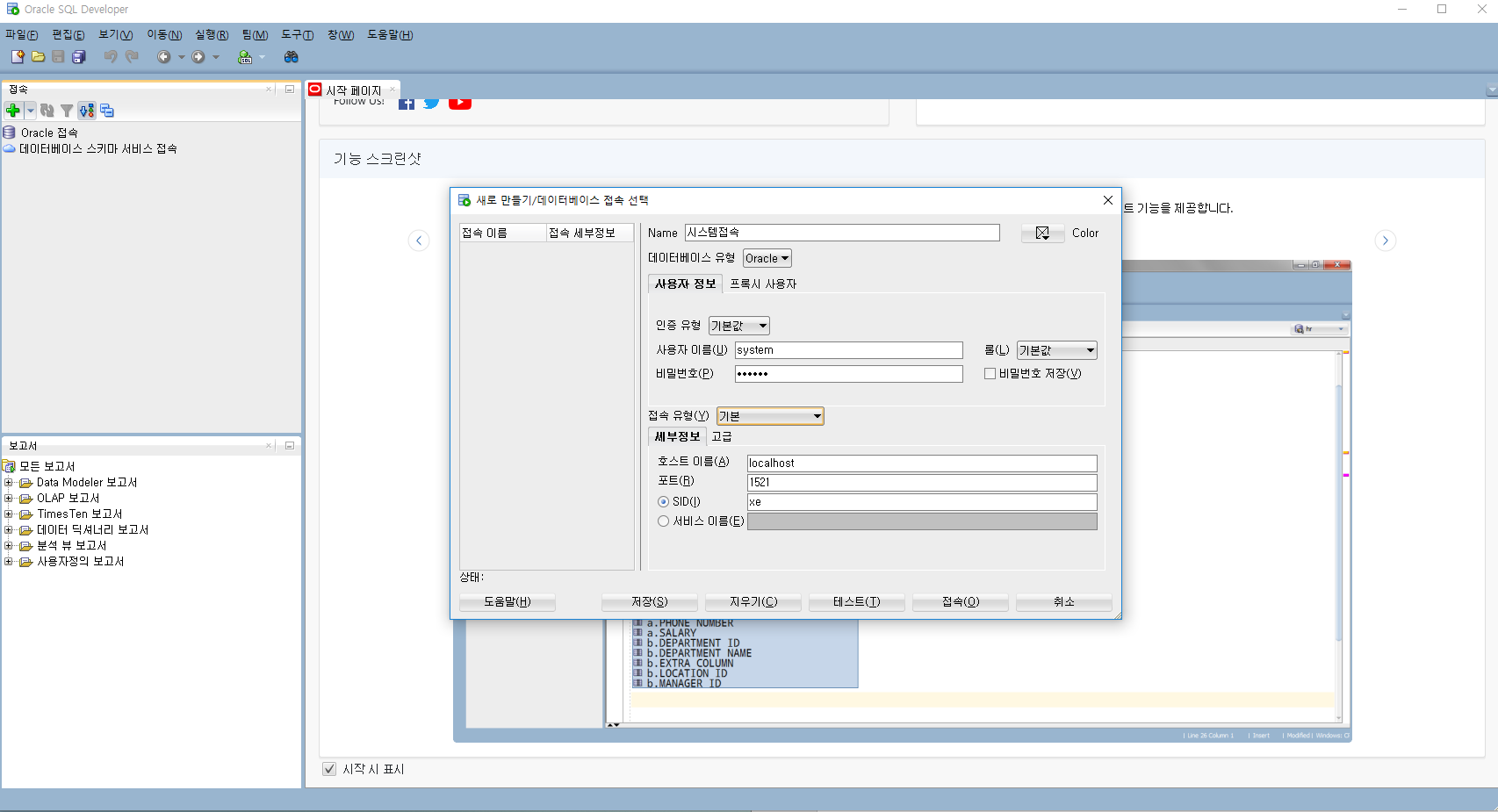
Toad



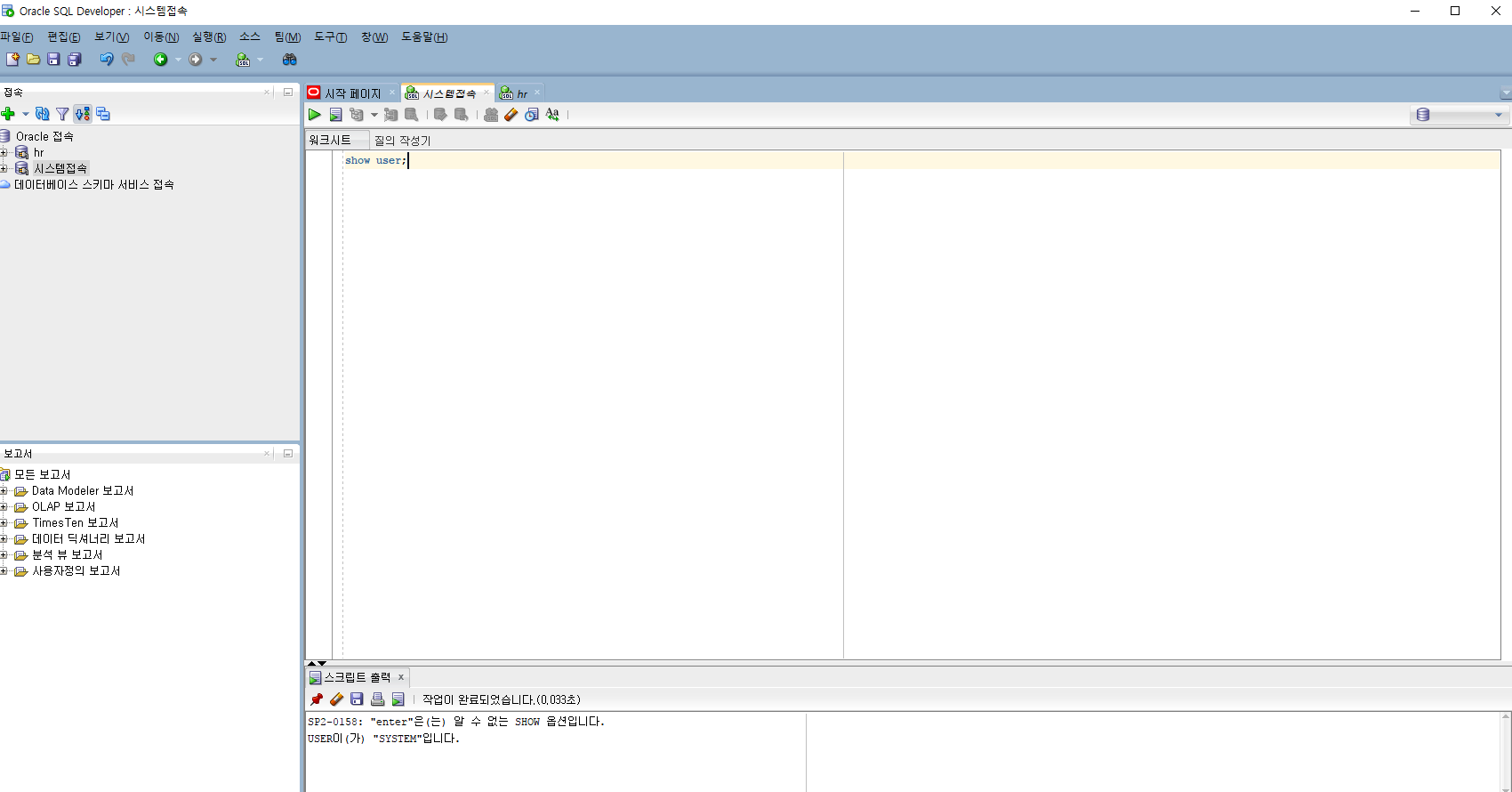
Sqlgate



Oracle SQL Developer 설치후



접속



Show user; 누르고 ctrl+enter

Select\*from tab;

Select\*from countries;

문장의 끝은 항상 세미클로(;)

오라클 데이터 타입:

<https://coding-factory.tistory.com/416>

문자데이터 타입 중에

CHAR[n]: 전화번호나 주민번호 같이 정해져 있는 숫자에 적용

VARCHAR2[n]: 확장성

varchar2(20): 문자열의 범위가 20열이라는 뜻

숫자형데이터는

Number(n)로 표현

머리 속으로 다이어 그램(그림)을 생각하고 코드를 짜는 것이 중요!! 그림이 매우 중요!!

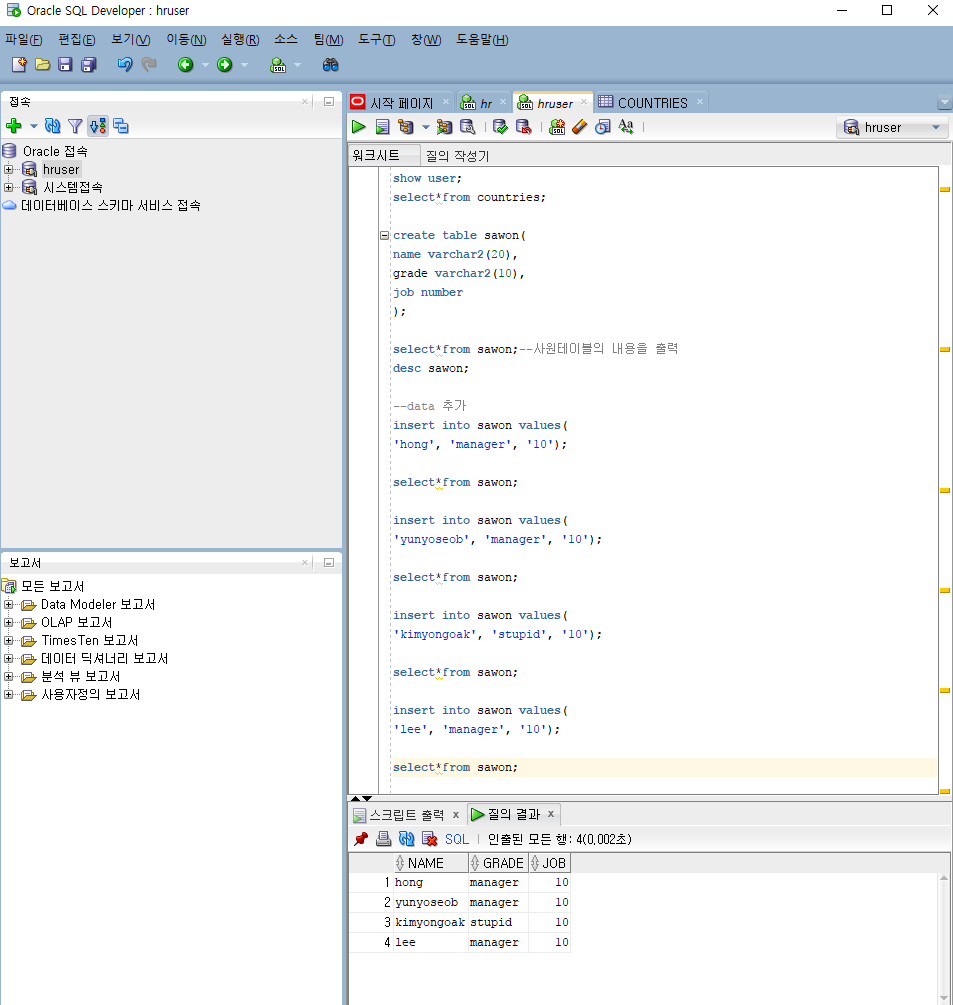
Oracle SQL Developer에서 VALUE는 무조건 전부 대문자로 표기

N테이블을 보고 싶을 땐 select\*from N;

--: 주석 (R에서 ##과 같은 의미)

Desc N; ->N테이블 내림차순

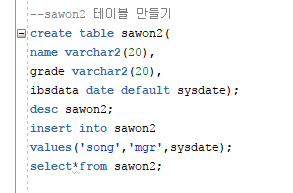
Insert into N values: 행 삽입



칼럼 수와 ‘n’의 개수가 일치하지 않으면 오류 남.

그러나 a,b,c 칼럼이 있어도 insert into N(a,b) values (‘c’,’d’); 이런 식으로 칼럼 수와 ‘n’의 개수를 맞추면 오류는 나지 않지만 c 칼럼(열)에 (null)값으로 표기 됨.

예시) insert into sawon(name, grade) values ('hoon', 'marketing');



날짜를 삽입하고 싶으면, ibsdata date default sysdate를 추가한 뒤, values(‘song’, ‘mgr’, sysdate); 혹은 values(‘song’, ‘mgr’, ‘yy/mm/dd’); 로 입력

날짜는 date(n)으로 표기하는 것이 아니라 date으로만 표기

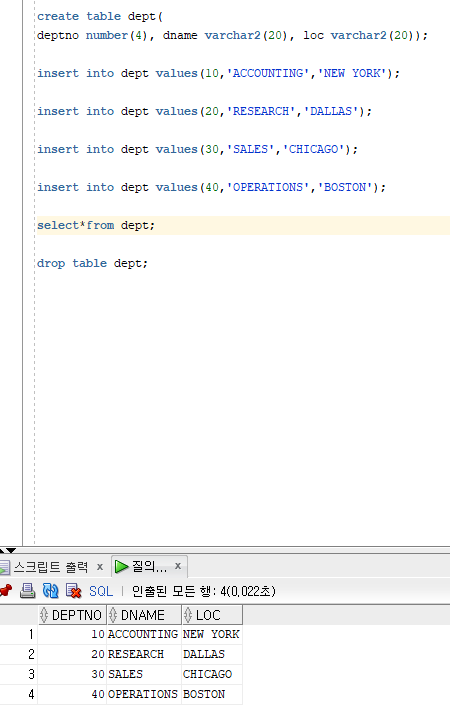
a,b,c(현재 날짜) 칼럼이 있을 때, insert into N(a,b) values (‘c’,’d’); 이런 식으로 칼럼 수와 ‘n’의 개수를 맞추면 디폴트 값으로 지정된 현재 날짜(c)가 날짜 칼럼에 적힘

commit; -> 커밋 완료.

교재추천: 오라클 SQL과 PL/SQL

**2021/03/12 금요일 수업**

기본키, 외래키(외부키)



테이블이 잘못 생성되면, drop table N(테이블명); 으로 삭제 후에

Create table N(테이블명); 으로 다시 만든 후, insert into N(테이블명) values()으로 다시 삽입하면 된다.

Select\*from tab; 테이블을 만들고 지운 흔적 보기

Show recyclebin; 쓰레기통 보기

Purge recyclebin; 쓰레기통 비우기

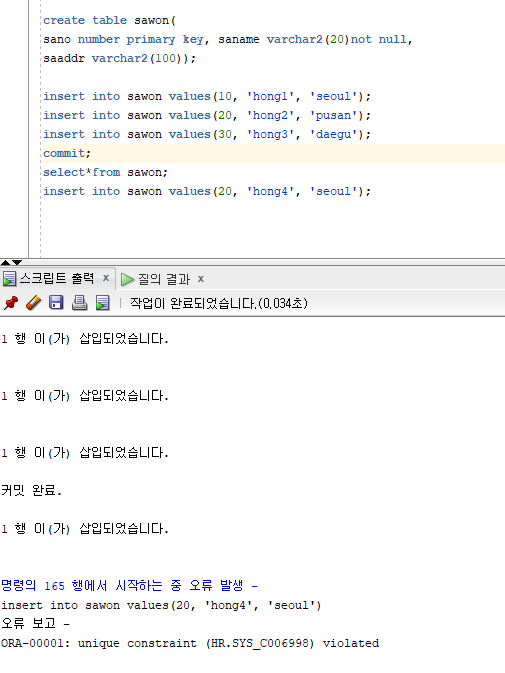
drop table N(테이블명) purge; 테이블과 테이블기록 지우기

CONSTRAINT 제약조건 사용하기

칼럼명 앞에 ‘constraint’를 입력하여 제약조건을 설정한다.

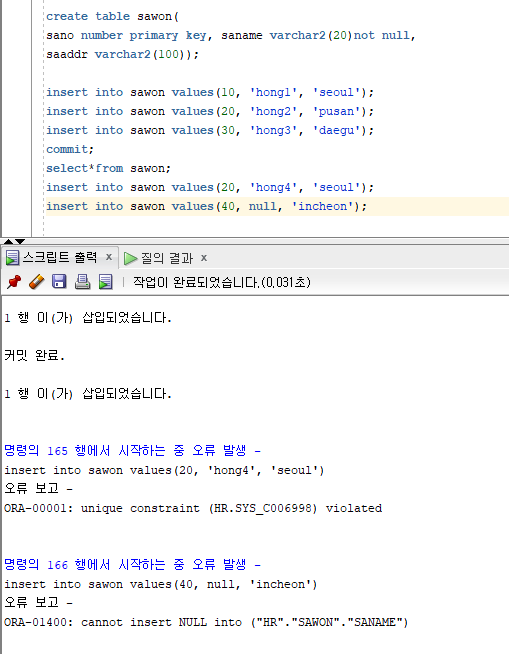
UNIQUE의 뜻은 그 값이 고유하여 중복되지 않아야 함을 의미함. (EX: 아이디 중복검사 등등)

1. Primary key 제약조건이 지켜지지 않은 경우

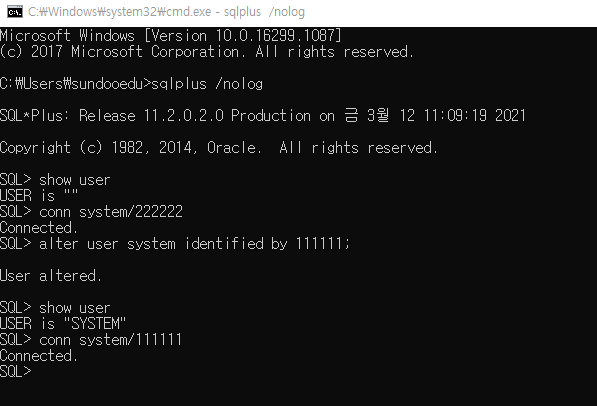


Unique constraint 오류: 유니크 값 오류 발생 (why? Sano가 겹쳐서)

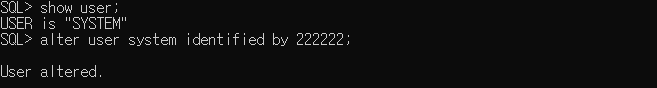
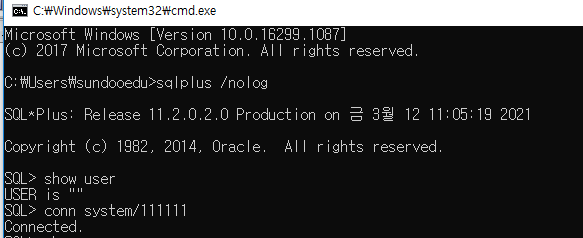
1. Not null 제약조건이 지켜지지 않은 경우



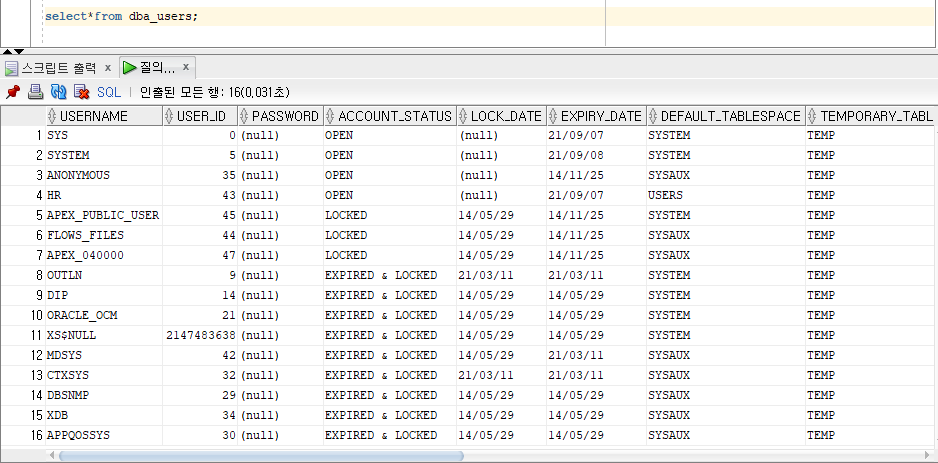
Cannot insert NULL into 오류 발생

SQL 프로그램 비밀번호변경

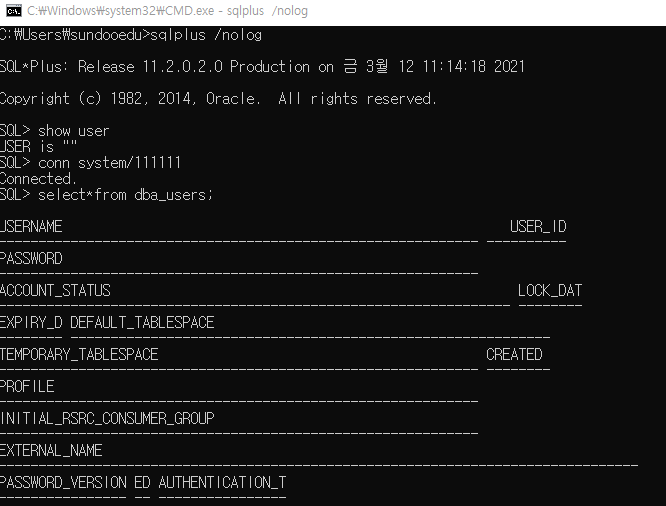
원상 복귀

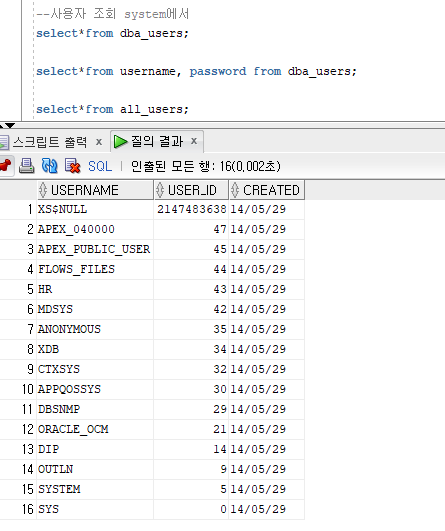
윈도우+R -> CMD -> 

Select\*from dba\_users; -> 사용자 조회



명령 프롬포트(cmd 에서도 확인가능)





@c:\empdept.sql; 커밋완료

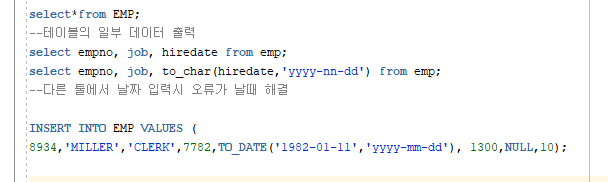
Select empno, job, hiredate from emp;

Select empno, job, to\_char(hiredate,’yyyy-mm-dd’) from emp;

INSERT INTO EMP VALUES (7934,'MILLER','CLERK',7782,'82-01-11',1300,NULL,10);

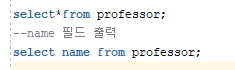
날짜 정보 형식이 툴에 안 맞을 수 있음

이럴 때는

INSERT INTO EMP VALUES (8934,'MILLER','CLERK',7782,to\_date'1982-01-11','yyyy-mm-dd'), 1300,NULL,10);

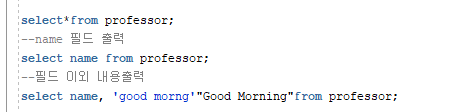
이렇게 to\_date'1982-01-11','yyyy-mm-dd'), 로 날짜 형식을 지정해줘야 함

특정 칼럼(열)만 출력하려면



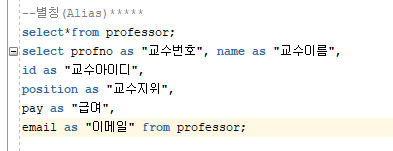
이런 식으로 select 출력하고 싶은 내용 from 테이블명; 으로 출력하면 된다.

필드 이외 내용을 출력하고 싶을 때는

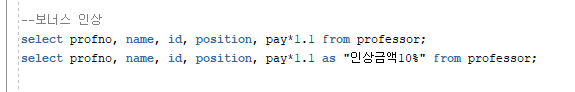


이렇게 하면 된다.

별칭을 주고 싶으면, select 이전별칭 as “바꿀별칭” from 테이블명; 으로 바꿀 수 있다.



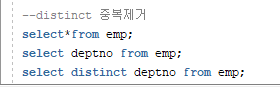
테이블에 숫자속성은 연산을 할 수 있다. 가령 보너스를 1.1배로 올린 후 별칭을 인상금액 10%로 바꾸고 싶으면 pay\*1.1로 올릴 수 있다.



별칭의 경우 as, “”가 빠져도 오류가 나지 않는다. 단, 생략시에는 한단어만 가능하다. 띄어쓰기가 존재하는 단어의 경우 “”가 반드시 필요하다.

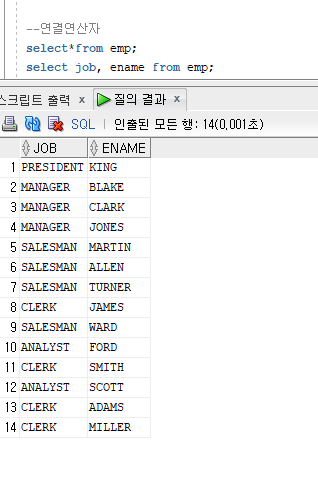
Distinct 중복제거

Select\*frame

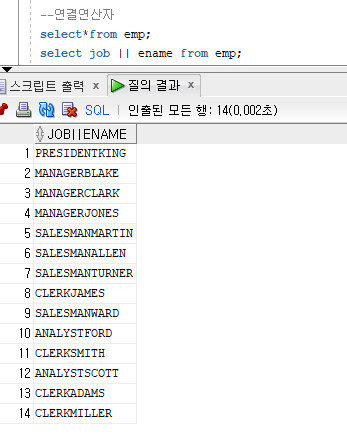


테이블을 출력한 후, 칼럼을 출력 한 뒤, select distinct 칼럼명 from 테이블; 을 통해 중복값을 제거한다.

연결 연산자 사용 이전



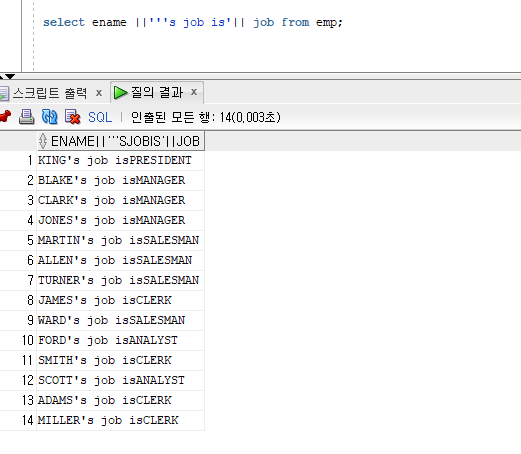
연결 연산자 사용 이후



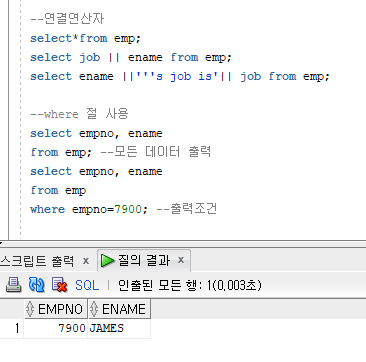
select ename ||'''s job is'|| job from emp;

작은 따옴표 3개는 역할을 뒤에 문자들으로 이전한다는 의미이다.

||’a’|| 이런 식으로 연결연산자를 사용한다.



Where절 사용 (출력 조건절) -> select~from~where~;

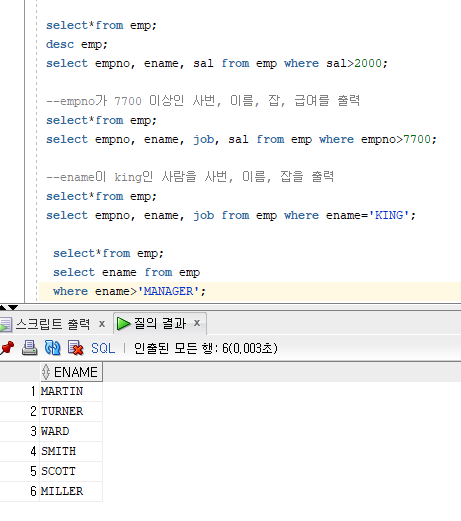


Emp 테이블에서 empno, ename 중에 empno가 7900인 값을 출력하라

Where절에서 문자열에 조건을 다는 경우 value값의 대문자, 소문자는 매우 중요하다.

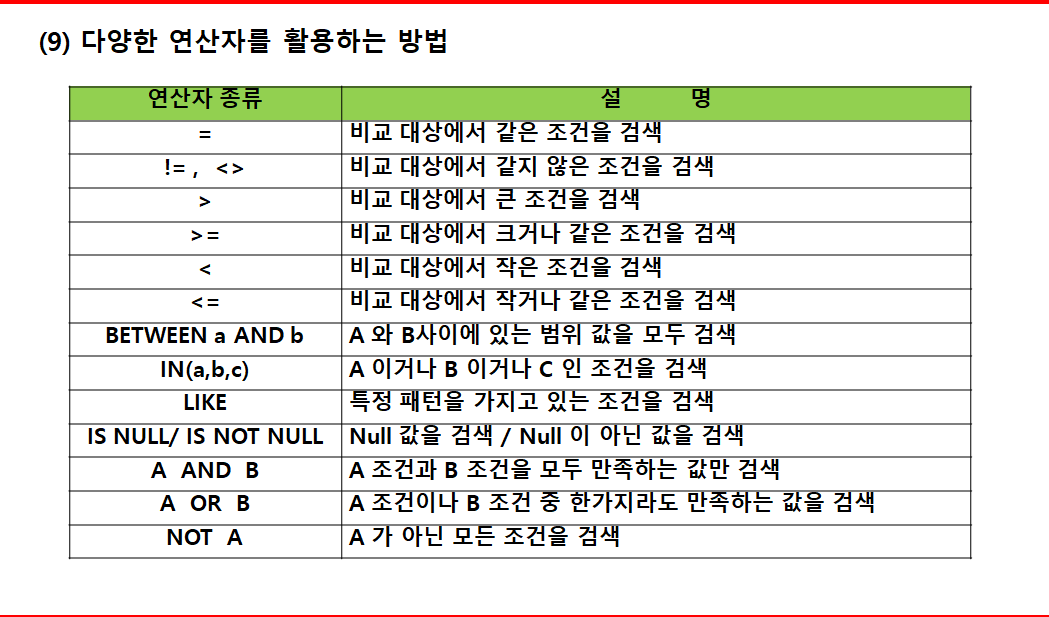
예시로 이름이 KING일 경우 where ename= ‘king’;으로 하면 안 나오나, where ename=’KING’;으로 하면 나온다.

문자도 부등호의 사용이 가능함. Where ename > ‘MANAGER’; 검색시 결과 값 나옴.

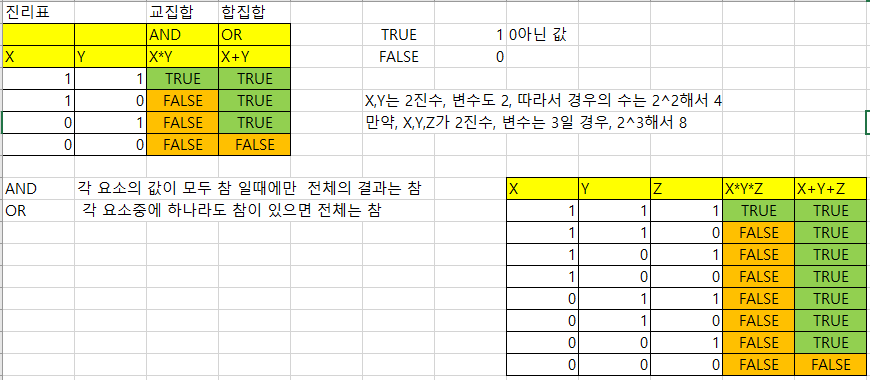


날짜의 경우 where hiredate >= ‘81/12/10’;이면 81년 12월 10일 기준으로 같거나 그 이후를 의미한다.

연산자 활용 방법



진리표



연산자 SQL 활용 예시 1 (and, or)



Select\*from where은 해당 열의 각각의 행까지 모두 보여줌 (전체 열 확인)

select~from~where~은 where에 해당하는 열만 보여줌 (사용하는 테이블 조회)

연산자에서도 사칙연산 우선순위가 똑같이 적용되기 때문에 괄호는 매우 중요하다!!

다만, AND연산자가 OR연산자보다 우선시 되기 때문에 AND우선으로 할거면 괄호가 없어도 상관없다.

연산자 SQL 활용 예시 2 (in, like)



Like 연산자의 경우 키워드 검색시 사용 된다.

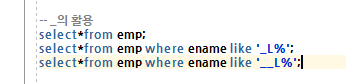
Like 연산자에서 날짜의 경우 월은 숫자로 표현하지 않고, JULY 등으로 표기한다.



D2 폰트 다운로드 (D2 coding 글꼴)

<https://m.blog.naver.com/PostView.nhn?blogId=12fac&logNo=221211884527&proxyReferer=https:%2F%2Fwww.google.com%2F>

‘\_’의 활용

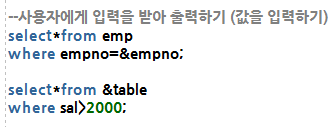


\_가 한 개면 L이 한 글자 들어간 데이터 추출

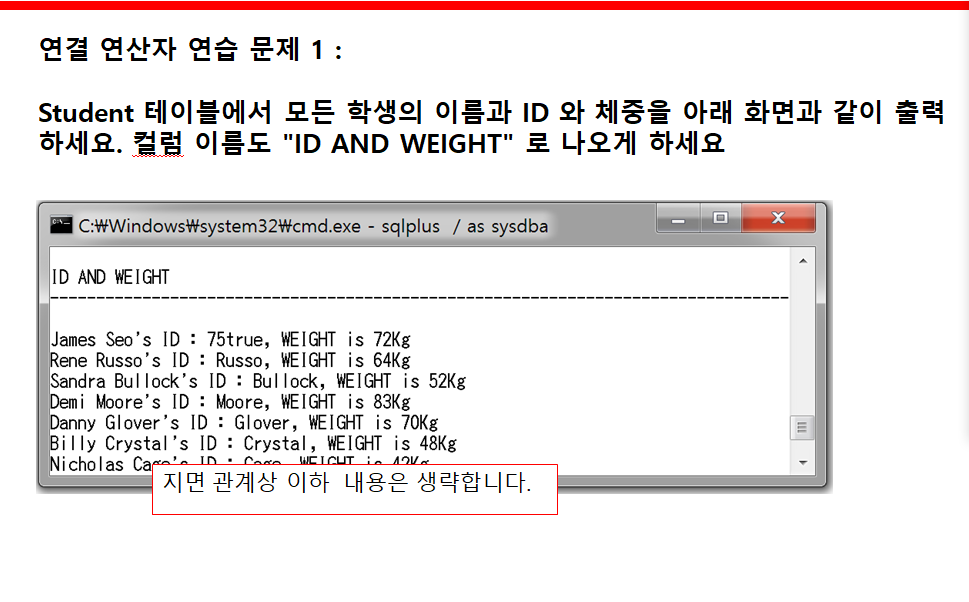
\_가 두 개면 L이 두 글자 들어간 데이터 추출

Null의 활용



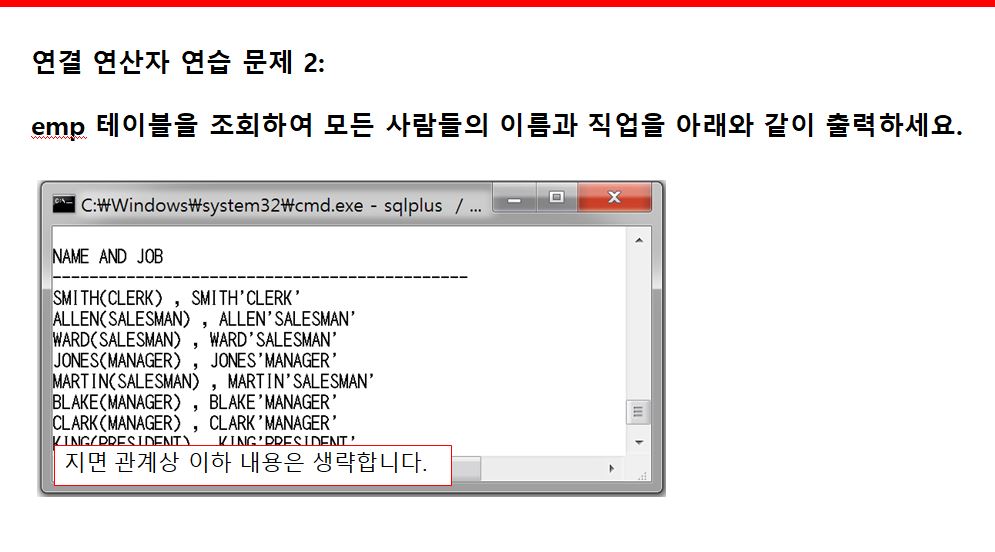


20210312 과제



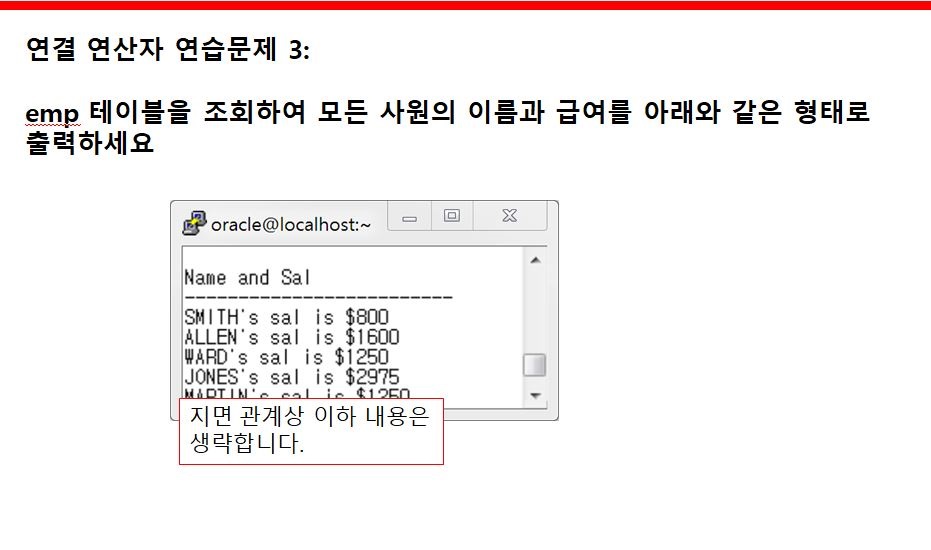
1번 문제 해답

select name||'''s ID:'|| id||', WEIGHT is'|| weight||'kg' from student;



2번 문제 해답

select ename ||'('||job||'), '||ename|| ''''||job||''''AS "NAME AND JOB"from emp;



3번 문제 해답

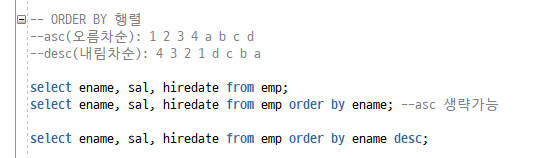
select ename ||'''s sal is'||'$'|| sal as "Name and Sal" from emp;

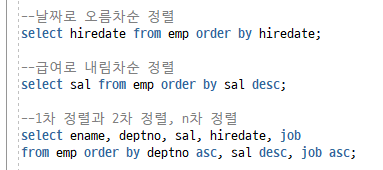
ORDER BY 행렬 (순서대로 정렬할 때 사용)

Asc(오름차순): 1 2 3 4 a b c d default

Desc(내림차순): 4 3 2 1 d c b a

Asc의 경우 생략 가능하다.





**2021/03/15 월요일 수업**

상업용 무료 에디터 구글 검색

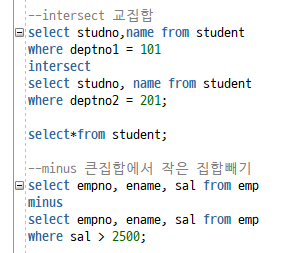
Notepad-plus-plus -> Download -> Notepad++7.8.3 release -> Download 64-bit x64

<https://notepad-plus-plus.org/>

union 중복된 값을 제거하고 출력

union all 중복된 값을 제거없이 모두 출력

select a,b, from table where c=111, select d,e from table where f=111 이런 문구를 블록잡고 F5누르면 아래 출력됨. 교집합이 있는지 intersect 문구삽입을 통해 확인 가능하며 minus를 통해 큰 집합에서 작은 집합을 뺄 수 있음

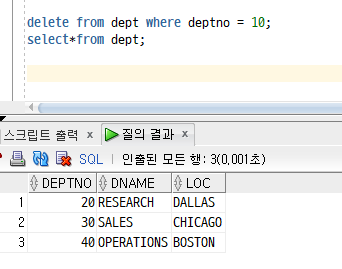


Drop table: 테이블 자체를 지우기

Delete table: 테이블 틀은 남겨두고 테이블 안에 있는 행(value)들만 삭제

Rollback: delete table 다시 살리기(복구) but, drop table은 다시 살릴 수 없으며, commit;을 해버리면 복구 불가

1행만을 제거 하고 싶을 때는 delete명령문에 조건을 달면 된다.



2개 이상의 행을 제거 하고 싶을 때는

where조건절 뒤에 Or을 붙이거나 IN(a,b…), like 등의 연산자를 활용하면 된다.

Ex) Delete from dept where deptno=20 or deptno=30;

Delete from dept where deptno in(10,40);

Delete from dept where deptno like ‘10’;



질의 결과 창에서 선택 후 삭제도 가능 (ctrl+좌클릭 하면 복수 선택 가능)

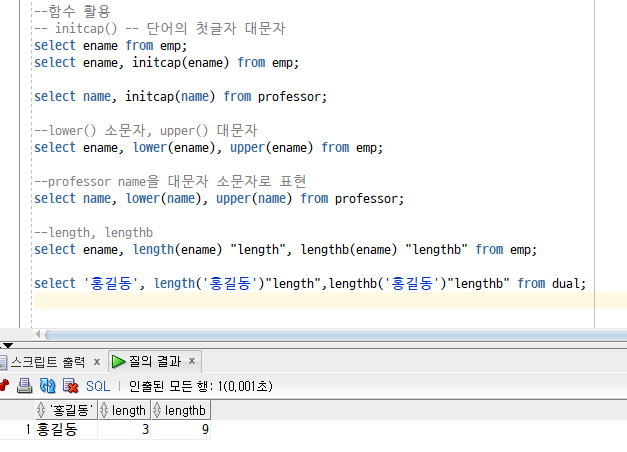
SQL함수

Initcap() – 단어의 첫 글자 대문자

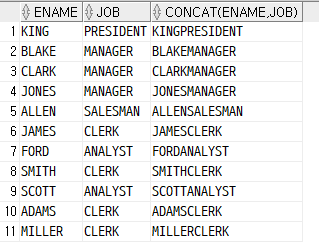
Lower() – 소문자, upper() -- 대문자

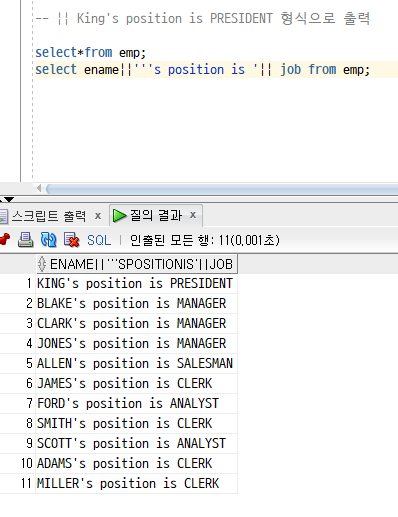
Length – 글자 길이를 의미 (자음, 모음 합쳐서 각각 글자들의 합)

lengthb – 글자 개수를 의미 (자음, 모음 각각 따로 전부 합친 것. 영문자일 때는 length와 차이가 없으나, 한글에서 차이가 남)



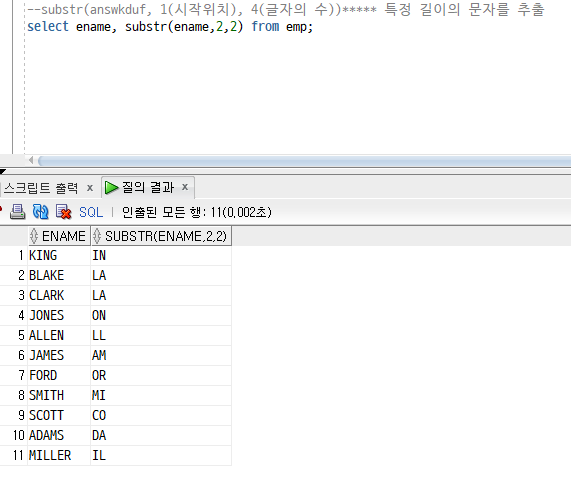
Concat() – 괄호안에 해당하는 밸류들을 합쳐줌





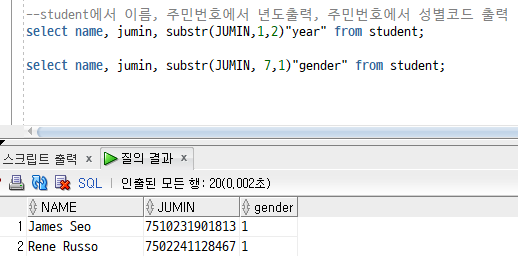
Substr(문자열, 1(시작위치), 4(글자의 수))\*\*\*\*\* 특정 길이의 문자를 추출

: substr 괄호안에 들어가는 문자를 시작위치로부터 하여금 글자의 수만큼 추출하는 함수

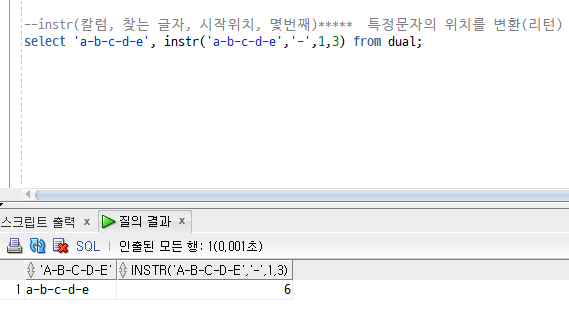


Substr(‘abcdef’, -3, 2) “-3,2” from dual; 은 뒤에서부터 3번째로 시작해서 2글자를 출력한다.

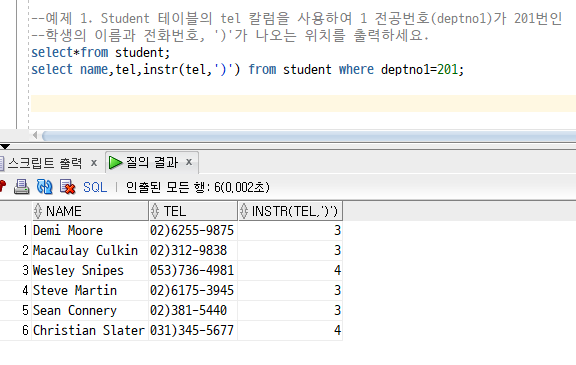
질의 결과 이름을 바꾸고 싶으면 from 앞에 “바꿀이름”을 입력하여 수정



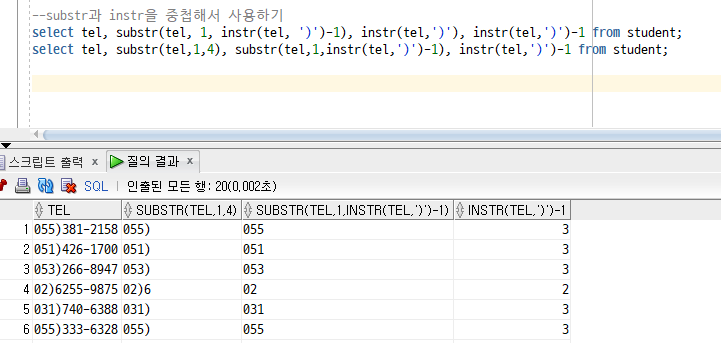
Instr(칼럼, 찾는 글자, 시작위치, 몇번째)\*\*\*\*\* 특정문자의 위치를 변환(리턴)



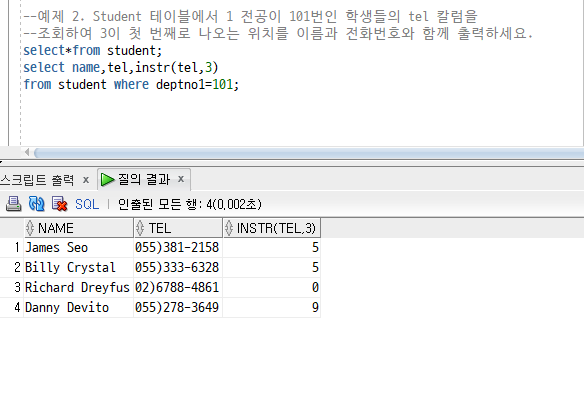
a-b-c-d-e 에서 첫 번째 글자를 시점으로 3번째 ‘-‘가 나오는 지점은 바로 6번째(a-b-c다음) 지점이다. 참고: <https://joke00.tistory.com/97>



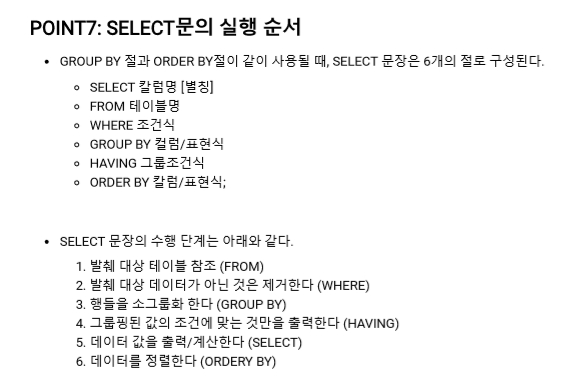
Substr & instr의 활용( 국번 뒤에 ‘)’ 없애고 표현하기 )



예제2)



db입문0.pdf p.15 반드시 암기할것!!



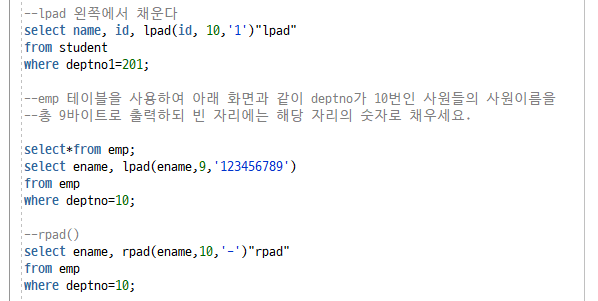
문장을 구성하는 순서(사진 첫 번재 단락)와 문장을 파싱(해석)하는 순서(사진 두 번째 단락)는 다름

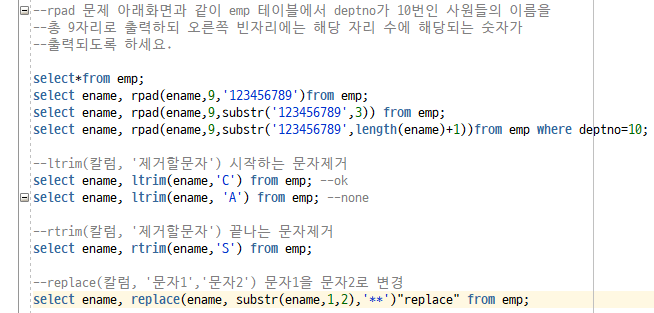
--lpad(해당칼럼, 채울글자수, 채울내용) 글자 수를 채우는데 왼쪽에서 채운다

--rpad(해당칼럼, 채울글자수, 채울내용) 글자 수를 채우는데 오른쪽에서 채운다

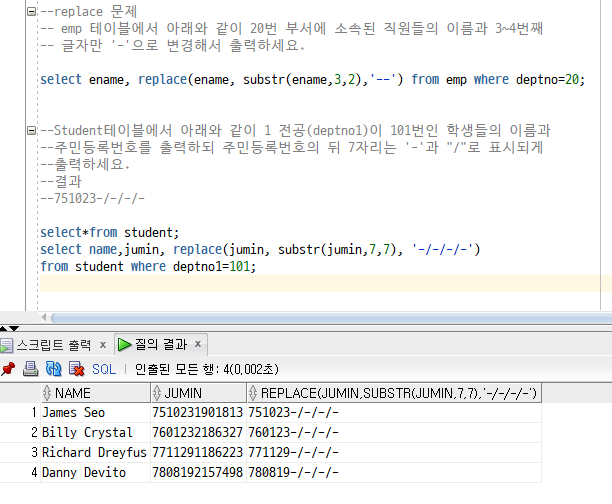
--Ltrim(칼럼, ‘제거할문자’) 시작하는 문자제거

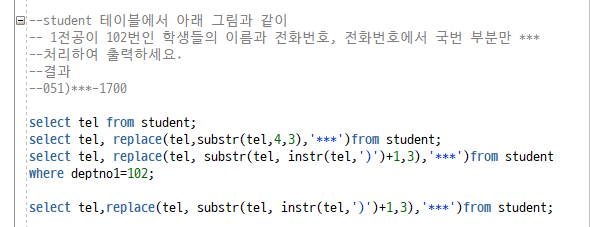
--rtrim(칼럼, ‘제거할문자’) 끝나는 문자제거



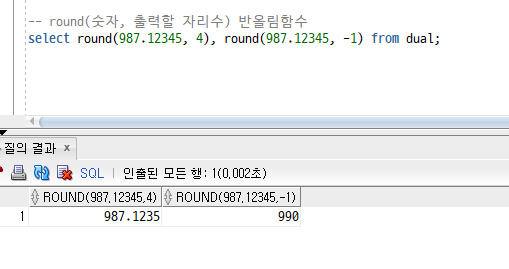


--replace(칼럼, ‘문자1’, ‘문자2’) 문자1을 문자2로 변경



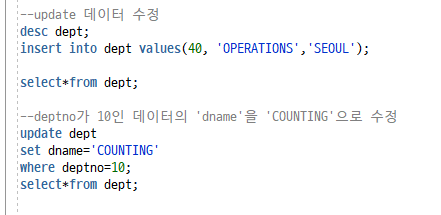


Round(숫자, 출력할 자리수) 반올림함수



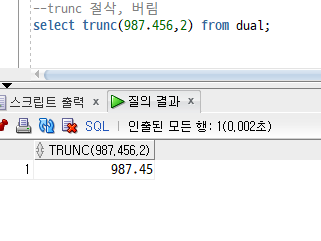
분산이나 표준편차에서 소수점이 나올 때 유용하게 사용

Update 데이터 수정 (update~테이블명~ set~ 내용~ where~위치~;)



**2021/03/16 화요일 수업**

Trunce: 절삭, 버림

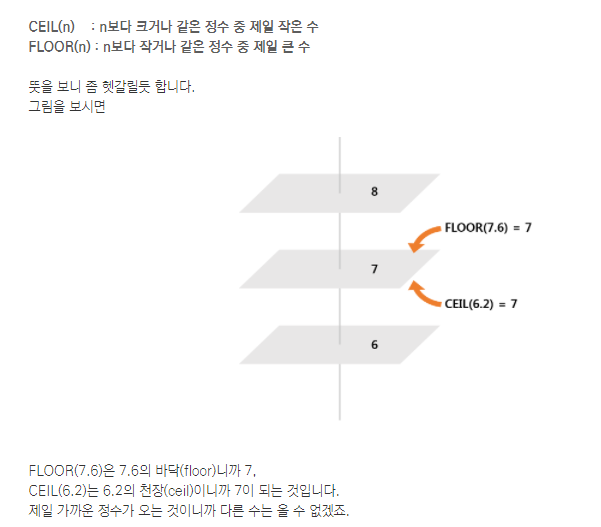


Mod\*\*\*: 나머지 함수

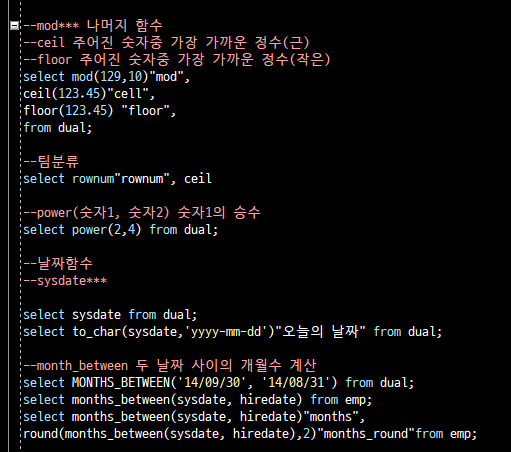
Ceil: 주어진 숫자중 가장 가까운 정수(큰)

Floor: 주어진 숫자중 가장 가까운 정수(작은)

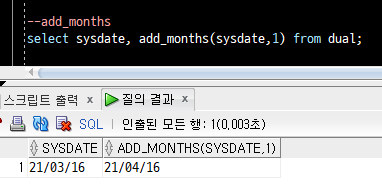
Power(숫자1, 숫자2): 숫자1의 승수



Months\_between(큰 수, 작은 수) 두 날짜 사이의 개월 수 계산



Add\_month, next\_day



Select sysdate, add\_months(sysdate, 추가할 개월 수) from dual;

DAY 같은 경우

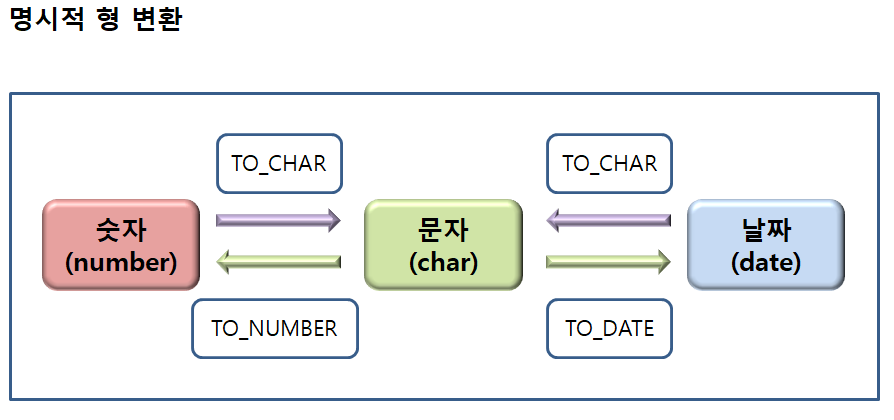
Ex) select next\_day(sysdate, ‘월’) from dual; 돌아오는 월요일의 날짜

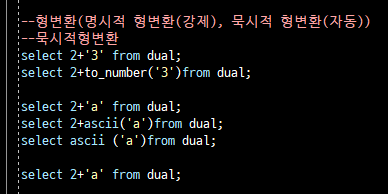
Ex) select sysdate, last\_day(sysdate) from dual; 이번 달 마지막 날짜

형변환(명시적 형변환(강제), 묵시적 형변환(자동))

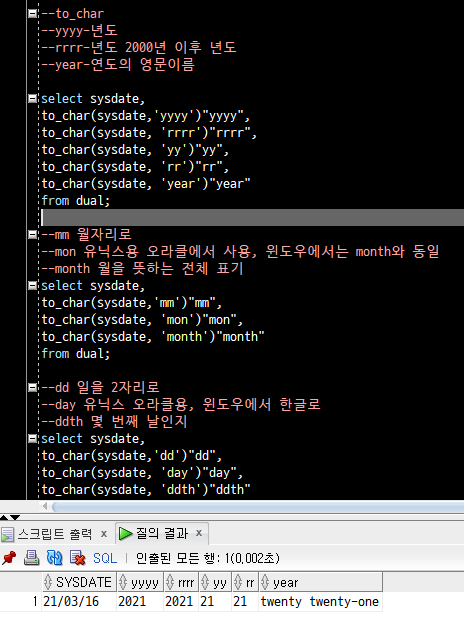
묵시적형변환

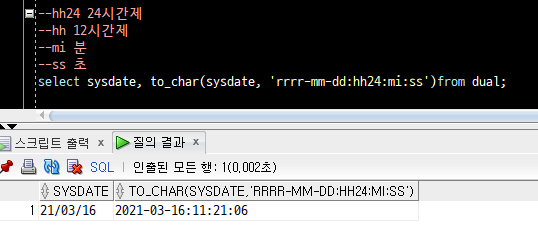
Select 2+’3’ from dual; -> 5



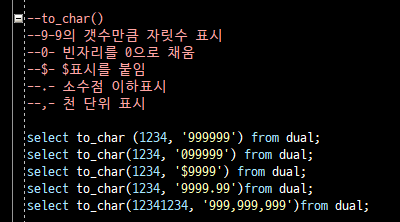


To\_char(sysdate)

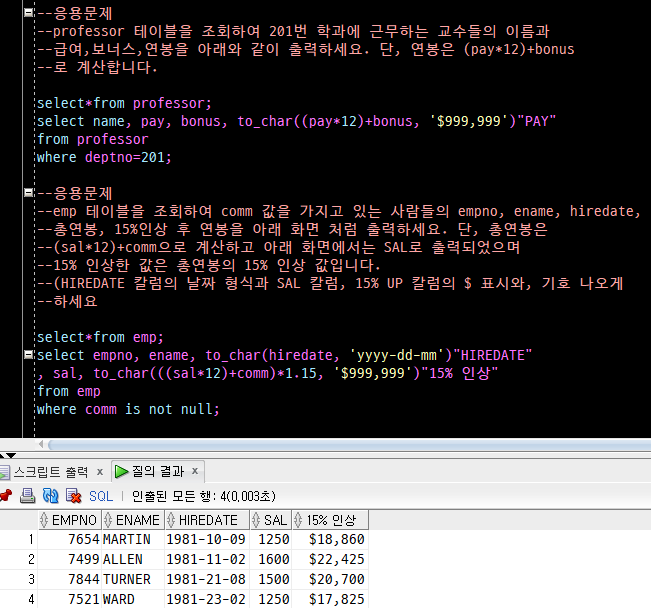




To\_char()

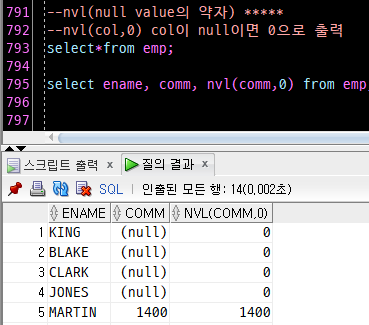


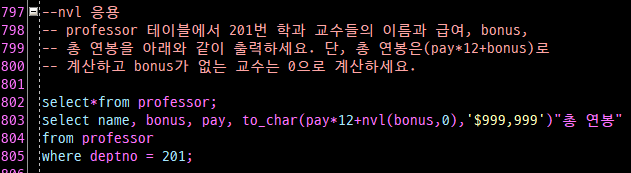
To\_char 응용문제



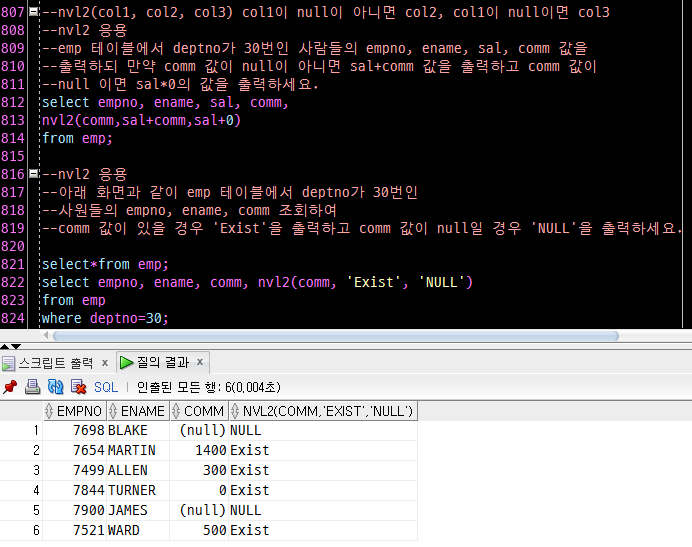
Nvl(null value의 약자)\*\*\*\*\*

Null 값은 처리가 어려우므로 다른 값으로 대체하는 것이 좋음

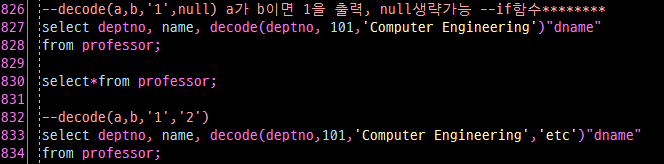


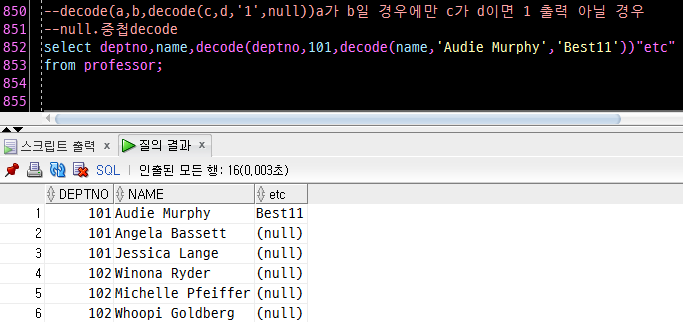


Nvl2(col1, col2, col3): col1이 null이 아니면 col2, col1이 null이면 col3

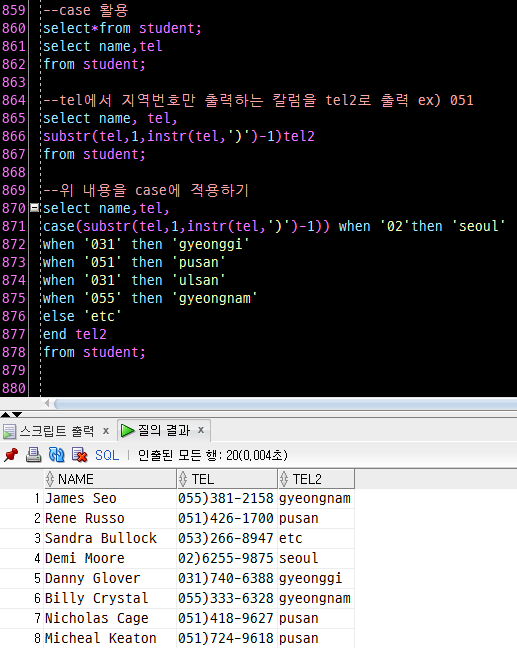


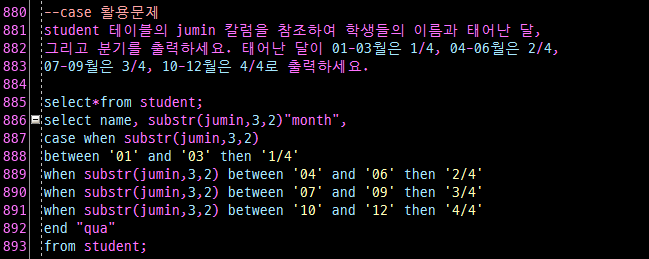
--decode(a,b,'1',null) a가 b이면 1을 출력, null생략가능 --if함수\*\*\*\*\*\*\*\*





Case 활용





ERD Cloud: <https://www.erdcloud.com/>

테이블 모양 좌클릭 후 테이블 생성

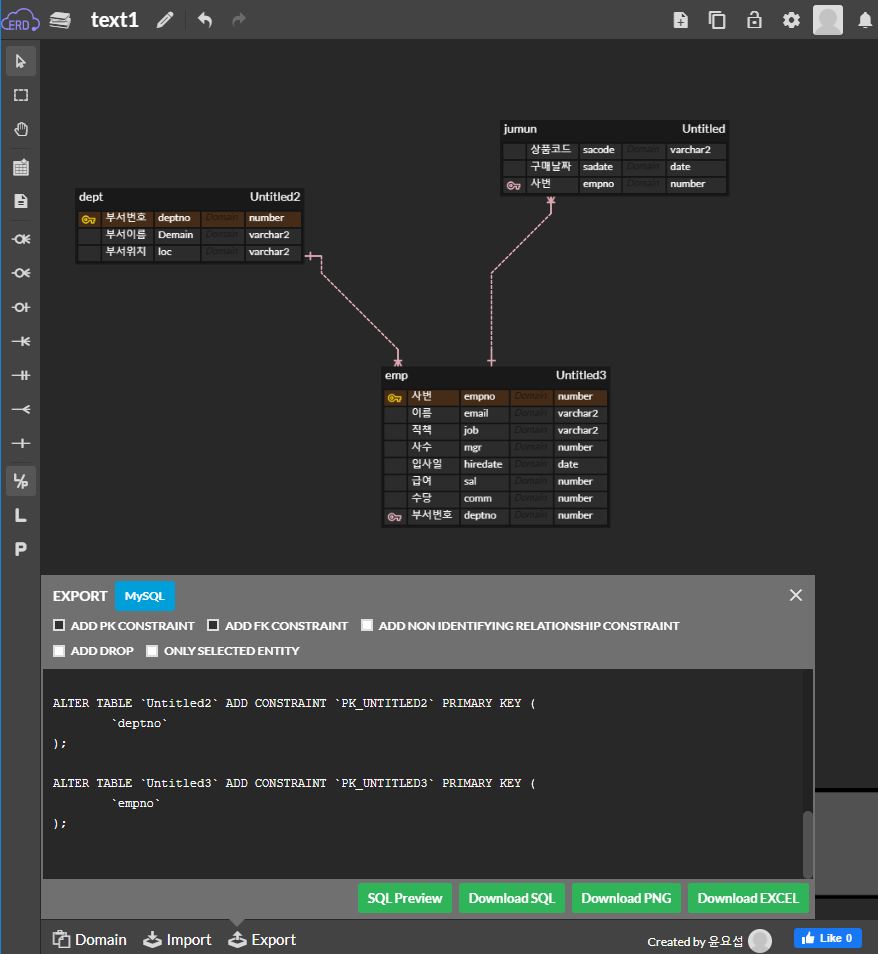
테이블 안에 내용 채우기

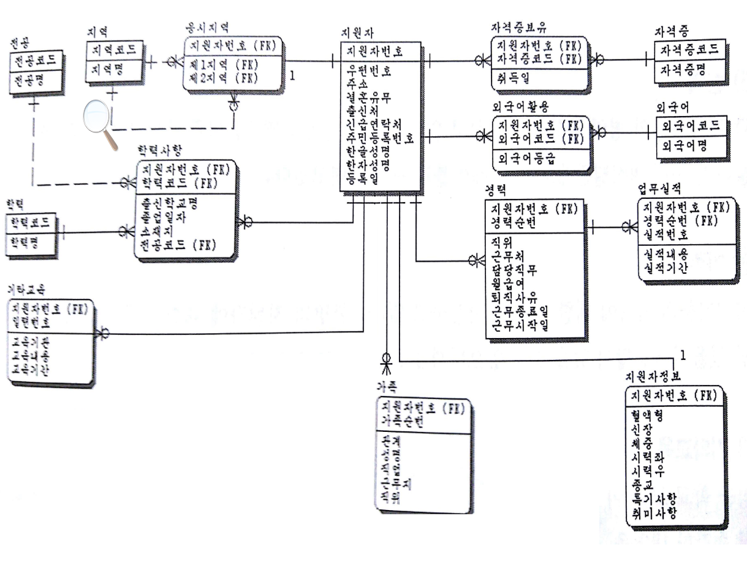
One or many(4) 를 통해 테이블간 연간관계 설정 (relationship or non-relationship)

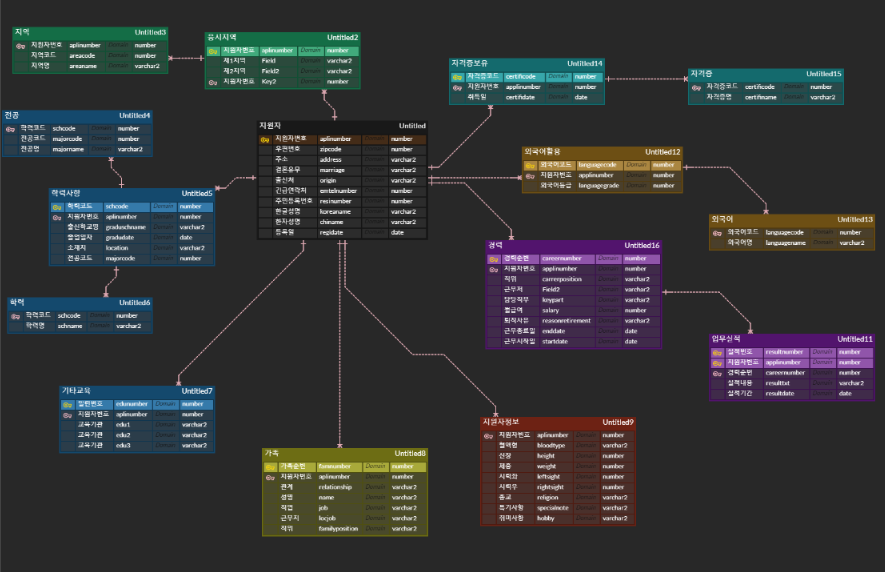
테이블끼리 연관지었을 때 1대 다 관계(즉, 값이 중복되면 안 됨)

과제: 관계설정은 non-relationship으로 다 연결 시킬 것

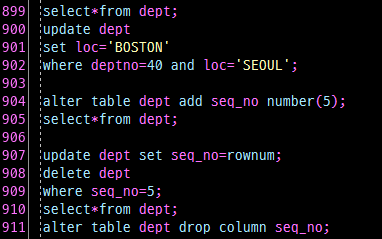
ERD-CLOUD 테이블을 만드는 이유 왼쪽 하단에 export누른 후 SQL Privew로 전환







Sql 활용



Rownum: 결과집합에 대한 가상의 순번