

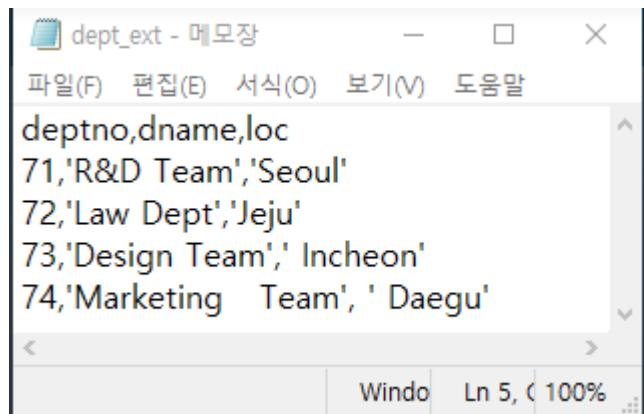
A decorative graphic on the right side of the slide. It features a grid of circles, some of which are filled with a light gray color, while others are empty. To the left of this grid is a solid red vertical bar. The entire graphic is set against a light gray background that also features a large, faint, light gray rectangular shape.

01. 외부 데이터 적재(Loading)

1. Sample Data Loading (SQLDEV)

1 Copy dept_ext.csv to c:\SQLDEV

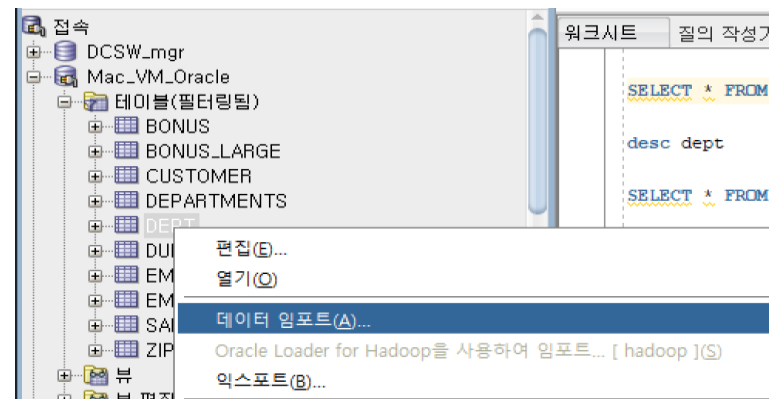
2 Notepad or Excel에서 파일 관찰



	A	B	C
1	deptno	dname	loc
2	71	'R&D Team'	'Seoul'
3	72	'Law Dept'	'Jeju'
4	73	'Design Team'	'Incheon'
5	74	'Marketing Team'	'Daegu'

3 데이터 적재 (적재 대상 테이블 선택)

접속 → devDinkDBMS(개인마다 다른)
→ DEPT → 마우스 왼쪽버튼 클릭
→ 데이터 импорт (Data Import)



1. Sample Data Loading

4

찾아보기 → c:\SQLDEV\dept_ext.csv → 문자 데이터가 ' ~ ' 로 표현

데이터 마라보기

데이터 파일 임포트: C:\SQLDEV\dept_ext.csv

찾아보기(R)...

파일 형식

☒ 헤더 건너 뛴 후

형식(F): csv

인코딩(I): MS949

구분자(D): ,

원쪽 둘러싸기(U): ' ' (highlighted)

오른쪽 둘러싸기(B): " " (highlighted)

행 건너뛰기(K): 0

☒ 마라보기 행 제한(P): 100

행 터미네이터(L): 표준: CR LF, CR 또는 LF

파일 내용

deptno	dname	loc
71	R&D Team	Seoul
72	Law Dept	Jeju
73	Design Te...	Incheon
74	Marketing ...	Daegu

도움말(H) < 뒤로(B) 다음(N) > 완료(F) 취소

5

임포트 방식 선택 (삽입 or 스크립트 삽입)

데이터 임포트 마법사 - 단계 2/4

임포트 방식

데이터 임포트 방식을 선택하십시오. 스테이징 외부 테이블 방식의 경우 대상 테이블을 임포트할 수 있도록 외부 테이블이 스테이징 테이블로 생성됩니다. 기타 임포트 방식의 경우 데이터가 테이블로 바로 임포트됩니다.

임포트 방식: 삽입 (highlighted)

테이블 이름: DEPT

☒ SQL 워크시트로 생성 스크립트 전송

☐ 임포트 행 제한(I): 100

파일 내용

deptno	dname	loc
71	'R&D Team'	'Seoul'
72	'Law Dept'	'Jeju'
73	'Design Team'	'Incheon'
74	'Marketing Team'	'Daegu'
75	'제2 고객 마케팅팀'	'Gunpo'

도움말(H) < 뒤로(B) 다음(N) > 완료(F) 취소

1. Sample Data Loading

6 선택된 열 3개 컬럼 확인후 → 다음

데이터 임포트 마법사 - 단계 3/5

열 선택

데이터 집합에서 임포트할 열을 선택하여 원하는 순서대로 배열하십시오.
사용 가능한 열

선택된 열

deptno
dname
loc

파일 내용

deptno	dname	loc
71	'R&D Team'	'Seoul'
72	'Law Dept'	'Jeju'
73	'Marketing Team'	'Incheon'

다음(N) >

7 문제 해결후 → 재시작한후 다음 단계 진행 * dept_ext.csv에서 데이터 관찰후 보정

데이터 임포트 마법사 - 단계 4/5

열 정의

왼쪽의 소스 데이터 열 목록에 있는 각 열에 대해 오른쪽의 대상 테이블을 선택합니다.

일치 기준 이름

소스 데이터 열

deptno
dname
loc

대상 테이블 열

이름 DNAME
데이터 유형 VARCHAR2
크기/전체 자릿수 14
소수점 이하 자릿수 0
☒ 널 가능? 기본값
설명

데이터

'R&D Team'
'Law Dept'
'Design Team'
⚠ 'Marketing Team'

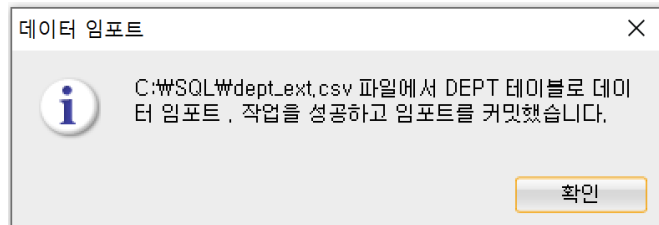
Status
데이터가 열 정의와 호환되지 않거나 널 불가능 열에 사용할 수 없습니다.

다음(N) >

● 1. Sample Data Loading

8

임포트 완료



9

데이터 확인

SELECT * FROM DEPT;

	DEPTNO	DNAME	LOC
1	10	ACCOUNTING	NEW YORK
2	20	RESEARCH	DALLAS
3	30	SALES	CHICAGO
4	40	OPERATIONS	BOSTON
5	71	R&D Team	Seoul
6	72	Law Dept	Jeju
7	73	Design Team	Incheon
8	74	Marketing Team	Daegu

10

데이터 보정

```
UPDATE DEPT SET LOC=LTRIM(LOC);  
SELECT * FROM DEPT;  
COMMIT;
```

● 2. Sample Data Loading (Excel)

① Copy dept_ext.csv to dept_ext2.csv

② Excel 에서 파일 오픈

	A	B	C
1	deptno	dname	loc
2	71	'R&D Team'	'Seoul'
3	72	'Law Dept'	'Jeju'
4	73	'Design Team'	' Incheon'
5	74	'Marketing Team'	'Daegu'

③ = " insert into dept values(" & A2 & "," & B2 & "," & C2 & ");"

	A	B	C	D
1	deptno	dname	loc	SQL
2	71	'R&D Team'	'Seoul'	insert into dept values(71,'R&D Team','Seoul');
3	72	'Law Dept'	'Jeju'	
4	73	'Design Team'	' Incheon'	
5	74	'Marketing Team'	'Daegu'	

* "~" : 문자열 , & : 셀 합성

- **금요일 주말 과제**

주제: 서울시 거주자중

대기오염 물질 배출 시설이 많은

지역(구 단위)고객(Customer) 대상

건강제품 마케팅

기간: 05/07~05/08

● 3. 공공 개방 데이터 Loading

- ① 공공 개방 데이터
서울시 빅데이터 캠퍼스 <https://bigdata.seoul.go.kr>
공공 데이터 포털 <https://www.data.go.kr>
- ② 데이터 수집 : 대기오염물질배출설치 사업장
→ 대기오염물질배출설치사업장.xlsx
- ③ Excel 에서 데이터 관찰후 데이터 축소(Data Reduction)
 - * 고객 데이터(Customer)와 연결 고리를 검토
 - * 축소 대상 ??
- ④ 축소된 데이터를 기준으로 테이블 생성 [각자 공부 해서]
- ⑤ 데이터 적재(Loading)
- ⑥ [Optional]데이터 전처리(축소/변환/정제/통합) , 데이터 정제 (불일치, 결측치, 이상치, 잡음)
- ⑦ 오염물질 배출 설치 사업장이 가장 많은 구 조사(ex 동작구)
- ⑧ 고객(Customer) 데이터와 연계(Join)하여 마케팅 대상자를 파일로 추출하여 제출,
임의의 주제로 분석하여 데이터 시각화 결과 및 SQL 제출 [시각화는 기존 프로젝트 활용]

● etc. SQL

대상	SQL
기존 테이블 복사	<pre>CREATE TABLE DEPT_BK1 AS SELECT * FROM DEPT WHERE 1=0 ; CREATE TABLE DEPT_BK2 AS SELECT DEPTNO,LOC FROM DEPT WHERE DEPTNO IN (10,20); CREATE TABLE DEPT_BK3 AS SELECT DEPTNO,DNAME ' ' LOC AS NAME_LOC FROM DEPT WHERE DEPTNO IN (10,20);</pre>
컬럼 추가	<pre>ALTER TABLE DEPT_BK2 ADD (TELNO VARCHAR2(12)); ALTER TABLE DEPT_BK2 ADD (ADDR VARCHAR2(100) DEFAULT 'SEOUL'); CREATE TABLE DEPT_BK3 AS SELECT DEPTNO, LOC, DEPTNO ',' LOC AS D_LOC FROM DEPT_BK2;</pre>
컬럼 길이 변경	<pre>ALTER TABLE DEPT_BK2 MODIFY (TELNO VARCHAR2(2)); DESC DEPT_BK2; ALTER TABLE DEPT_BK2 MODIFY (TELNO VARCHAR2(22)); ALTER TABLE DEPT_BK2 MODIFY (ADDR VARCHAR2(30)); ALTER TABLE DEPT_BK2 MODIFY (ADDR VARCHAR2(4)); ALTER TABLE DEPT_BK2 MODIFY (ADDR VARCHAR2(22));</pre>
컬럼 삭제	<pre>ALTER TABLE DEPT_BK2 DROP COLUMN TELNO; ALTER TABLE DEPT_BK2 DROP COLUMN ADDR;</pre>

● etc. SQL

대상	SQL
제약사항 추가	<pre>ALTER TABLE DEPT ADD CONSTRAINT DEPT_DEPTNO_PK PRIMARY KEY(DEPTNO); ** 문제 해결후 P.K 추가 !!!! ALTER TABLE DEPT DISABLE PRIMARY KEY; ALTER TABLE DEPT DISABLE CONSTRAINT DEPT_DEPTNO_PK ; ALTER TABLE DEPT ENABLE CONSTRAINT DEPT_DEPTNO_PK ;</pre>

● 4. 내/외부 데이터를 활용한 데이터 분석 및 시각화

개인별 미니 프로젝트

* 금요일 과제를 참고하여 플젝 수행

주 제: 내/외부 데이터를 활용한 데이터 분석 (각자 자율 주제 – 금융/통신)

(내부) 고객(Customer) 데이터, 직원(EMP), 부서(DEPT)

(외부) 공공 개방 데이터

ex) BC 카드 소비 매출 , KT 통신사

ex) 경기 데이터 드림 (<https://data.gg.go.kr>)

공공 데이터 포털 (<https://www.data.go.kr>)

기 간: 2일(05/10~05/11)

산출물: 1) 분석 주제/배경 2) 데이터 출처/개요/주요특징

3) SQL 4) 데이터 시각화 5) 결과에 대한 각자의 통찰력(Insight)