데이터분석 기초 SQL 부트캠프

목차

- 1. 데이터베이스와 RDBMS 그리고 SQL
- 2. MYSQL & HEIDISQL 설치
- 3. DataType

문자형, 숫자형, 날짜형

- 4. 데이터베이스
- 5. Table

create, alter, drop, truncate, insert, update

csv file import to table

06. 데이터 불러오기

select, from, alias, limit, offset

1.

데이터베이스와 RDBMS 그리고 SQL

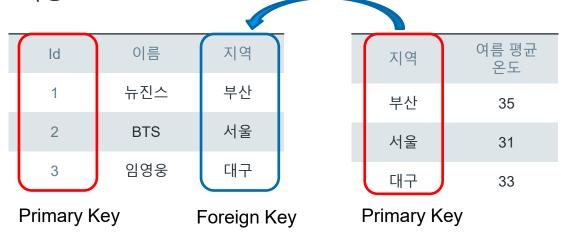
DATABASE(DB)

 Database(DB): 데이터베이스란 컴퓨터에 저장되는 데이터들의 집합, 저장소.

DBMS(DataBase Management System) : DB를 관리하는 시스템
 ex)oracle, mysql, mariadb, ms server

RDBMS(Relational database management system)

 RDB(Relational DataBase): 관계형 데이터베이스는 데이터들 사이의 관계를 가지는 데이터들을 다루며, 테이블 형태로 자료들을 저장.



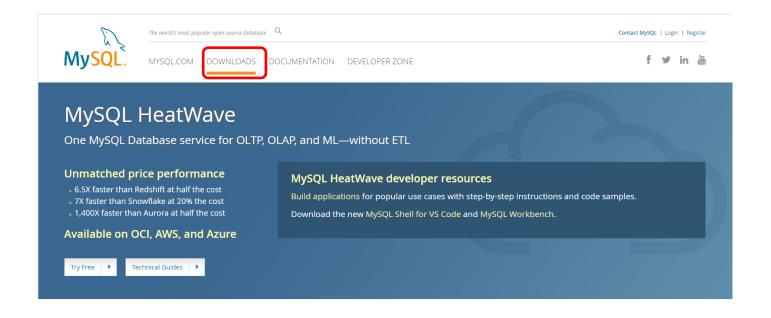
RDBMS(Relational database management system)

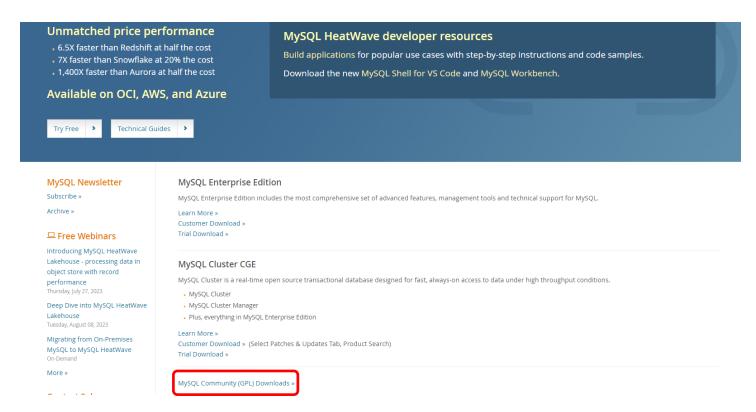
Primary Key		y Foreigr	Foreign Key		Prir	nary Key
	ld	이름	지역		지역	여름 평균 온도
	1	뉴진스	부산		부산	35
	2	BTS	서울		서울	31
	3	임영웅	대구		대구	33

- Primary Key : Table에서 row를 식별하게 해주며, 유일한 값이고 공백일 수 없으며 중복될 수 없음
- Foreign Key: 다른 Table과 연결시켜주는 column. 다른 Table의 primary key를 참조. 공백, 중복 가능.

2. MYSQL 설치

HEIDISQL 설치





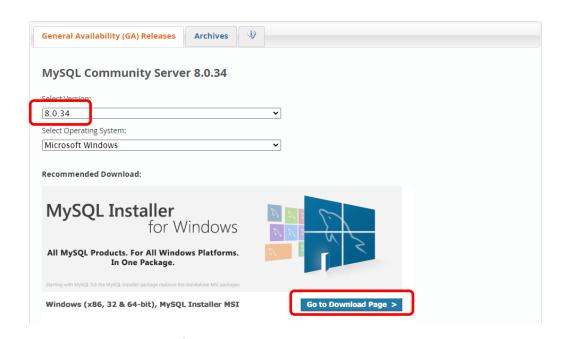
MySQL Community Downloads

- MySQL Yum Repository
- MySQL APT Repository
- MySQL SUSE Repository
- MySQL Community Server
- MySQL Cluster
- MySQL Router
- MySQL Shell
- MySQL Operator
- MySQL NDB Operator
- MySQL Workbench
- MySQL Installer for Windows

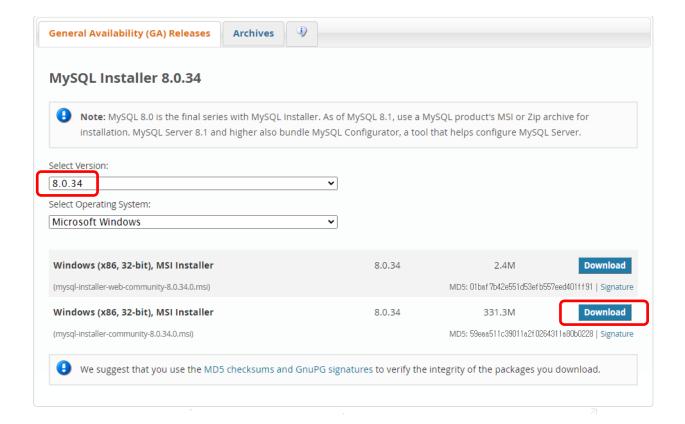
- C API (libmysqlclient)
- · Connector/C++
- Connector/
- Connector/NET
- · Connector/Node.js
- Connector/ODBC
- Connector/Python
- MySQL Native Driver for PHP
- MySQL Benchmark Tool
- Time zone description tables
- Download Archives

MySQL Community Downloads

◀ MySQL Community Server



•



Login Now or Sign Up for a free account.

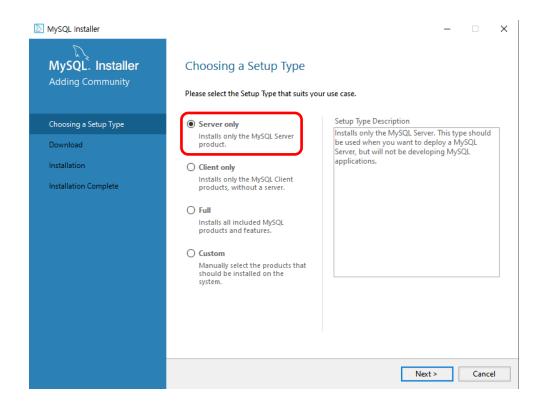
An Oracle Web Account provides you with the following advantages:

- · Fast access to MySQL software downloads
- Download technical White Papers and Presentations
- Post messages in the MySQL Discussion Forums
- · Report and track bugs in the MySQL bug system

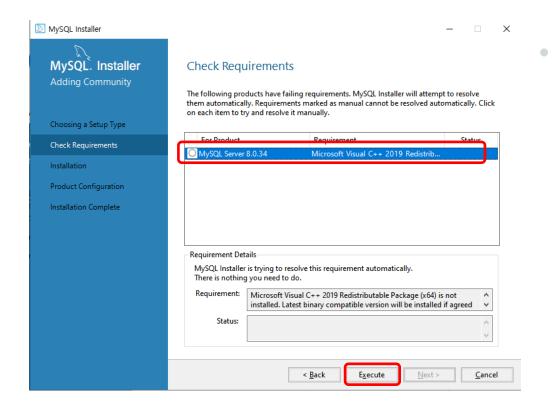


MySQL.com is using Oracle SSO for authentication. If you already have an Oracle Web account, click the Login link. Otherwise, you can signup for a free account by clicking the Sign Up link and following the instructions.

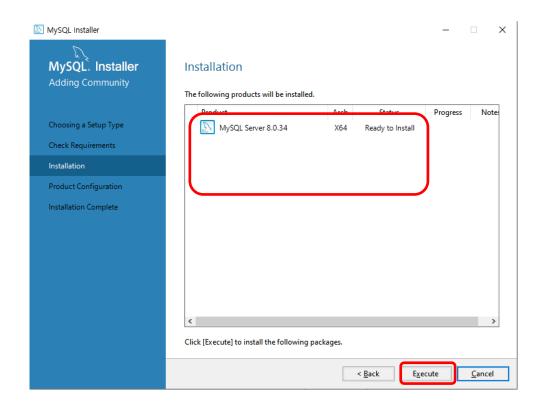
No thanks, just start my download.

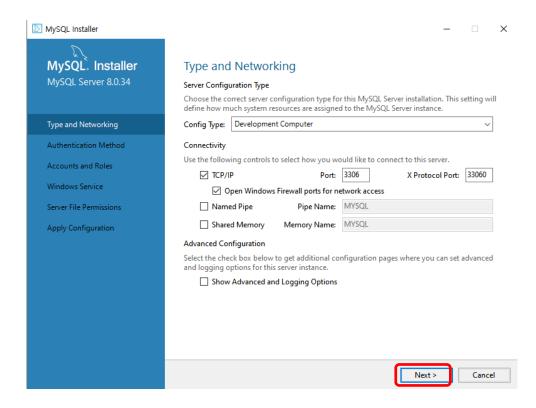


.

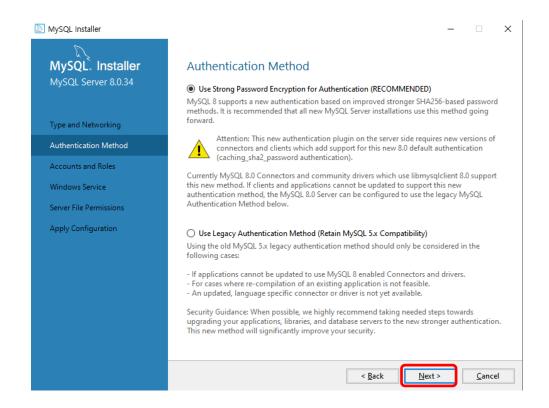


요구되는 프로그램들 모두 Excute. 동의 및 설치.

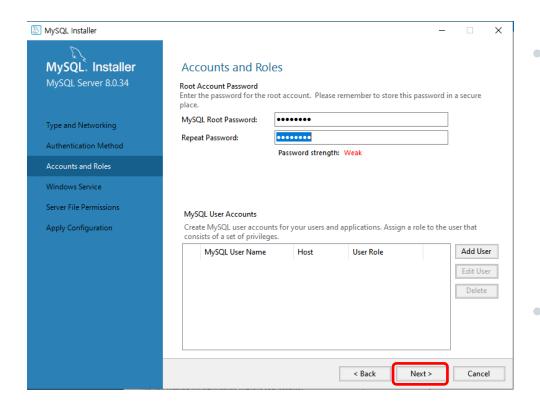




•



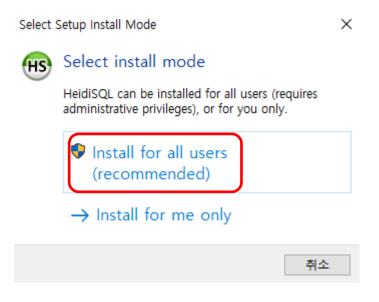
.

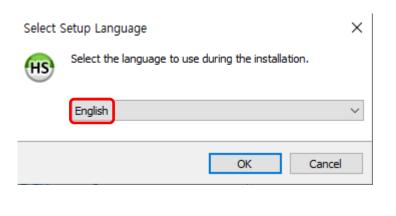


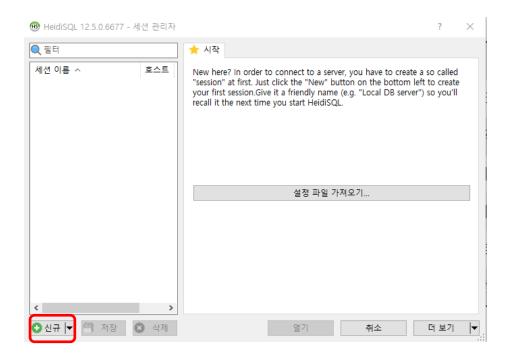
Root Password는 가장 중요한 비밀번호이기 때문에 까먹지 않을 수 있는 비밀번호를 사용.

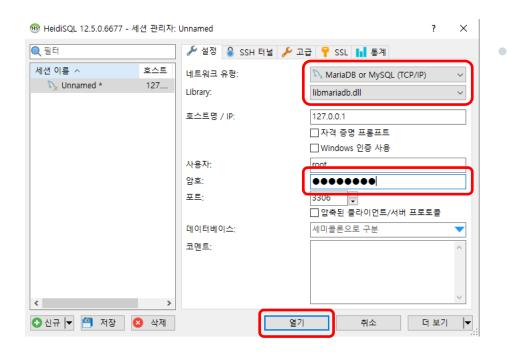
이후 next -> excute->설치완료



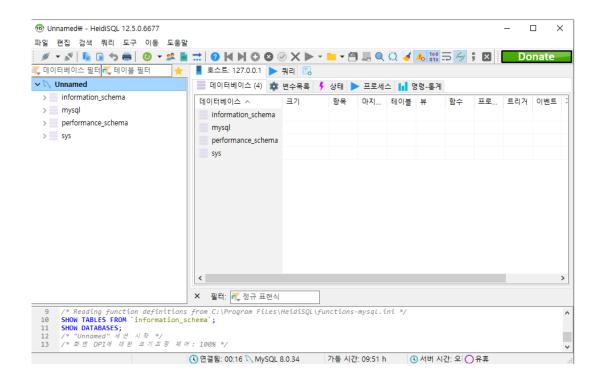








MySQL 설치시 설정했던 root의 암호



3. DataType

문자형, 숫자형, 날짜형

문자형

데이터 타입	Byte	설명
CHAR(n)	1 ~ 255	고정길이 문자형 n : 1~ 255
VARCHAR(n)	1 ~ 65535	가변길이 문자형 n : 1 ~ 65535
TEXT	1 ~ 65535	N 크기의 TEXT 데이터 값

- 강의에서는 encoding type utf8을 사용할 예정.
- Encoding : 문자를 컴퓨터가 이해할 수 있는 신호로 만드는 것
- Utf8에서 영어, 공백은 1byte, 한글은 한글자당 3byte.

숫자형(정수형)

데이터 타입	설명
TINYINT	(1byte) / -128 ~ +127 0 ~ 255수 표현 가능
SMALLINT	(2byte) / -32,768 ~ 32,767 0 ~ 65,536수 표현 가능
MEDIUMINT	(3byte) / -8,388,608 ~ +8,388,607 0 ~ 16,777,215수 표현 가능
INT	(4byte) / -2,147,483,648 ~ +2,147,483,647 0 ~ 4,294,967,295수 표현 가능
BIGINT	(8byte) / 무제한 수 표현 가능

TINYINT SIGNED(생략가능)

-128 ~ 127

TINYINT UNSIGNED

0 ~ 255

숫자형(실수형)

데이터 타입

n: 전체 자릿수(Precision : 정밀도)
d: 소수점 뒷자리수(Scale : 배율)

DECIMAL(n,d)

DECIMAL(4): -9,999 ~ 9,999

DECIMAL(4,1): -999.9 ~ 999.9

DECIMAL(4,2): -99.99 ~ 99.99

• DECIMAL은 최대 65자리까지 지원

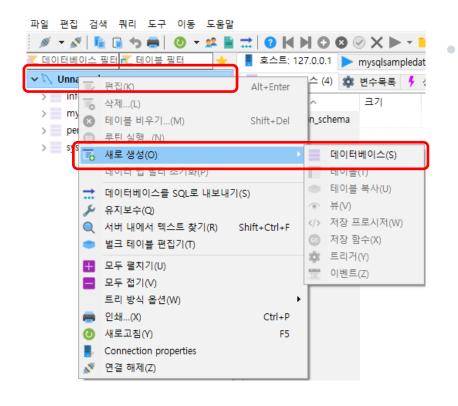
날짜형

데이터 타입	설명
DATE	'YYYY-MM-DD'
DATETIME	'YYYY-MM-DD HH:MM:SS'
TIME	'HH:MM:SS'
YEAR	YYYY

4.

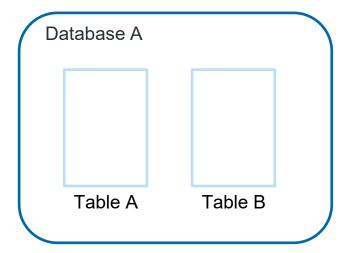
Database

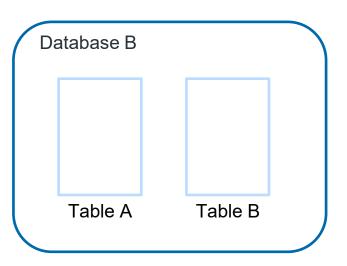
Database 생성



계정명 우클릭 -> 새로생성 -> 데이터베이스

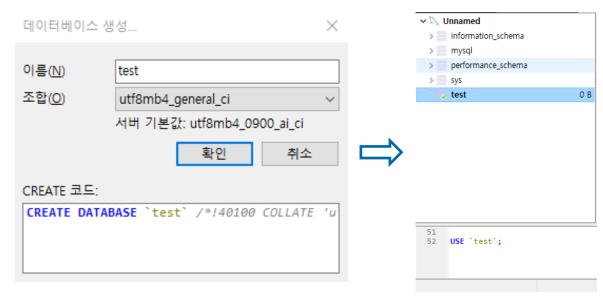
Database





- Database는 데이터를 저장하는 저장소.
- 여러 Database를 만드는 이유는 접근 권한을 쉽게 나누기 위함.

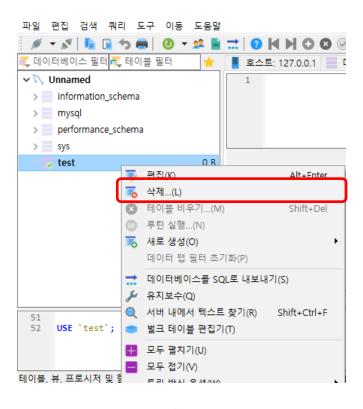
Database 생성



- 조합
- Utf8bm4_general_ci 로 변경

- Test라는 DB가 생성됨.
- 마우스로 클릭시 USE `test`; 쿼리가 자동 실행

Database 삭제



Test라는 DB 마우스 우클릭
 후 삭제.

DROP DATABASE 'TEST';와 같은 효과

5.Table

CREATE, ALTER, DROP, TRUNCATE

CREATE TABLE

```
    CREATE TABLE 테이블이름 (
컬럼명 테이터타입,
컬럼명 테이터타입,
    그로 걸럼명 테이터타입
    );
```

CREATE TABLE

SQL Code

```
CREATE TABLE customers(
customer_number INT,
customer_name VARCHAR(50),
phone VARCHAR(50)
);
```

• 실행결과

#	이름	데이터 유형	길이/설정	부호 없음	NULL 허용	0으로 채움	기본값
- 1	customer_nu	INT	10		~		NULL
2	customer_name	VARCHAR	50		\checkmark		NULL
3	phone	VARCHAR	50		~		NULL

CREATE TABLE

SQL Code

```
CREATE TABLE customers2(
customer_number INT NOT NULL,
customer_name VARCHAR(50) NOT NULL,
phone VARCHAR(50) NOT NULL
);
```

• 실행결과

#	이름	데이터 유형	길이/설정	부호 없음	NULL 허용	0으로 채움	기본값
1	customer_nu	INT	10				기본값 없음
2	customer_name	VARCHAR	50				기본값 없음
3	phone	VARCHAR	50				기본값 없음

INSERT

• INSERT INTO 테이블이름 (컬럼명1, 컬럼명2, 컬럼명3) VALUES (DATA1, DATA2, DATA3) ;

INSERT

SQL Code

```
insert into customers
  (customer_number,customer_name,phone)
values
  (1,'이상훈','010-1234-5678'),
  (2,'김상훈','010-1234-5679'),
  (3,'박상훈','010-1234-5679');
```

• 실행결과

test.customers: 3 행 (총) (대략적)

customer_number	customer_name	phone
1	이상훈	010-1234-5678
2	김상훈	010-1234-5679
3	박상훈	010-1234-5679

DELETE / TRUNCATE

- 둘다 테이블을 지우는 쿼리.
- DELETE : 데이터 삭제
- TRUNCATE : 테이블 초기화
- DELETE FROM `테이블이름` where ~
- DELETE 구문에서 where 문을 생략하면 모든 데이터 삭제

UPDATE

- 이미 존재하는 TABLE 의 내용을 수정하는 구문
- UPDATE 테이블이름 SET 컬럼명1 = DATA1, 컬럼명2 = DATA2 WHERE 조건~~;

ALTER

• UPDATE가 TABLE의 내용을 수정한다면, ALTER는 TABLE의 내용 외적인 부분을 변경.

 TABLE의 이름변경, 컬럼추가,컬럼이름 변경, 데이터타입변경, 컬럼삭제

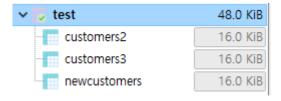
ALTER TABLE 이름 변경

- SQL Code
- ALTER TABLE `테이블명` RENAME `새이름`

```
ALTER TABLE customers rename newcustomers;
```

• 실행결과





ALTER TABLE 컬럼추가

- SQL Code
- ALTER TABLE `테이블명` ADD `새이름` 데이터타입

ALTER TABLE newcustomers ADD 지역 VARCHAR(50);

• 실행결과

#	이름	데이터 유형	길이/설정	부호 없음	NULL 허용	0으로 채움	기본값
- 1	customer_nu	INT	10		~		NULL
2	customer_name	VARCHAR	50		~		NULL
3	phone	VARCHAR	50		~		NULL
4	지역	VARCHAR	50		~		NULL

ALTER TABLE 컬럼 데이터 타입 변경

- SQL Code
- ALTER TABLE `테이블명` MODIFY `컬럼명` 데이터타입

ALTER TABLE newcustomers MODIFY 지역 INT;

• 실행결과

#	이름	데이터 유형	길이/설정	부호 없음	NULL 허용	0으로 채움	기본값
1	customer_nu	INT	10		ightharpoons		NULL
2	customer_name	VARCHAR	50		ightharpoons		NULL
3	phone	VARCHAR	50		~		NULL
4	지역	INT	10		ightharpoons		NULL

ALTER TABLE 컬럼 이름 변경

- SQL Code
- ALTER TABLE `테이블명` CHANGE `기존컬럼명` `새컬럼명` 데이터타입

ALTER TABLE newcustomers CHANGE 지역 REGION VARCHAR(10);

• 실행결과

#	이름	데이터 유형	길이/설정	부호 없음	NULL 허용	0으로 채움	기본값
1	customer_nu	INT	10		~		NULL
2	customer_name	VARCHAR	50		~		NULL
3	phone	VARCHAR	50		~		NULL
4	REGION	VARCHAR	10		~		NULL

ALTER TABLE 컬럼 삭제

- SQL Code
- ALTER TABLE `테이블명` DROP `컬럼명`

ALTER TABLE newcustomers DROP REGION ;

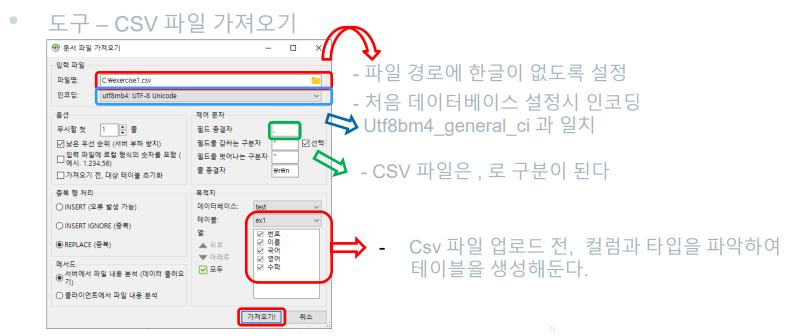
• 실행결과

#	이름	데이터 유형	길이/설정	부호 없음	NULL 허용	0으로 채움	기본값
- 1	customer_nu	INT	10		~		NULL
2	customer_name	VARCHAR	50		~		NULL
3	phone	VARCHAR	50		~		NULL

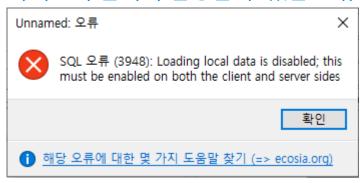
6. 데이터를 테이블에 넣는 방법

Import csv file to table, sql query

 확장자 .CSV 의 데이터를 TABLE에 넣는 기능을 HEIDISQL에서 편리하게 제공.



• 가져오기 클릭시 발생할 수 있는 오류



- LOCAL 에서 데이터를 가져오는 것이 허용되어있지 않기 때문에 허용해줘야 함.

SQL CODE

```
SHOW GLOBAL VARIABLES LIKE 'LOCAL_INFILE';
```

9 결과

```
Variable_name Value local_infile OFF
```

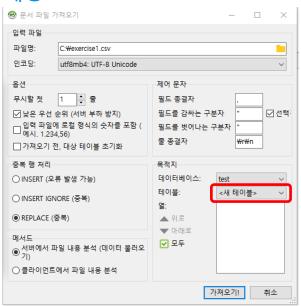
SQL CODE

```
SET GLOBAL LOCAL_INFILE=TRUE
SHOW GLOBAL VARIABLES LIKE 'LOCAL_INFILE';
```

결과

Variable_name Value

 확장자 .CSV 의 데이터를 TABLE에 넣는 기능을 HEIDISQL에서 편리하게 제공.



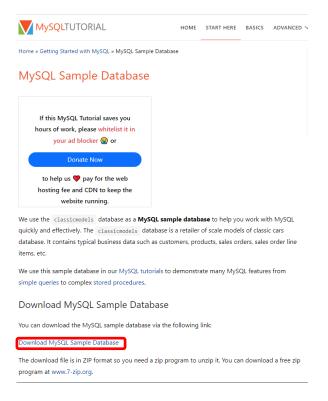
 새 테이블을 선택하면 자동으로 데이터의 타입을 추정해서 제시를 해준다.

```
B CSV 레이아웃 감지 - □ X

파일 찾아보기...

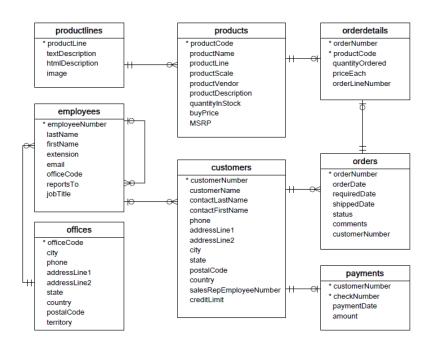
1 CREATE TABLE 'test'.'exercisel' (
2 '.' TINVINT NOT NULL,
3 '.' VARCHAR(10) NOT NULL,
5 '.' TINVINT NOT NULL,
6 '.' TINVINT NOT NULL,
7 ) 8
```

MYSQL SAMPLE DATABASE



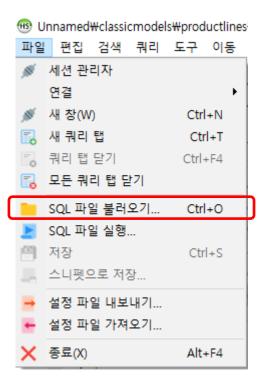
 https://www.mysqltutorial.or g/mysql-sampledatabase.aspx

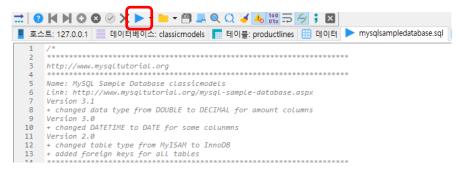
MYSQL SAMPLE DATABASE



MYSQLTUTORIAL 홈페이지에서 제공하는 SAMPLE DATA BASE의 구성.

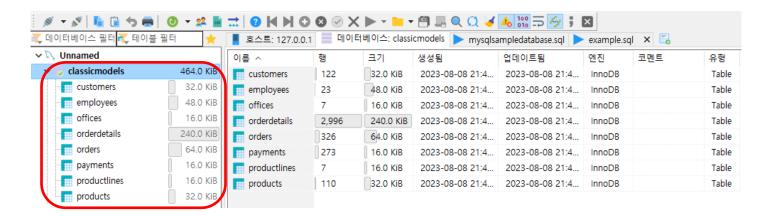
DATABASE 불러오기





• 스크립트의 구성을 보면 CREATE TABLE & INSERT 함수로 테이블 생성과 데이터 입력을 동시에 진행한다.

DATABASE 불러오기



- 스크립트의 실행 결과로 데이터베이스 classicmodels 가 생성.
- 그 데이터베이스 속에 8개의 table이 생성.

DATABASE 불러오기

classicmodels.orderdetails: 2,996 행 (총) (대략적), 제한 수: 1,000								
orderNumber 🥊	productCode 💡	quantityOrdered	priceEach	orderLineNumber				
10,100	518_1749	30	136.0	3				
10,100	518_2248	50	55.09	2				
10,100	518_4409	22	75.46	4				
10,100	524_3969	49	35.29	1				
10,101	518_2325	25	108.06	4				
10,101	518_2795	26	167.06	1				
10,101	524_1937	45	32.53	3				
10,101	524_2022	46	44.35	2				
10,102	518_1342	39	95.55	2				
10,102	518_1367	41	43.13	1				
10,103	510_1949	26	214.3	11				
10,103	510_4962	42	119.67	4				
10,103	512_1666	27	121.64	8				
10,103	518_1097	35	94.5	10				
10,103	518_2432	22	58.34	2				
10,103	518_2949	27	92.19	12				
10,103	518_2957	35	61.84	14				

- Orderdetails Table
- Ordernumber : 주문번호
- productCode : 제품번호
- quantityOrdered : 수량
- priceEach : 개당가격
- orderLineNumber :
 주문라인번호

6. 데이터 불러오기

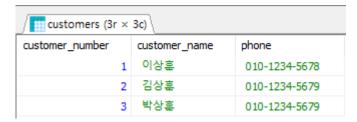
SELECT FROM

SELECT FROM

- SQL Code
- SELECT * FROM 테이블명;
- 여기서 * 는 모든 컬럼

```
SELECT * FROM `customers`
;
```

• 실행결과

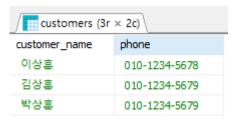


SELECT FROM

- SQL Code
- SELECT 컬럼명1, 컬럼명2 FROM 테이블명;

```
SELECT customer_name, phone FROM customers
:
```

• 실행결과



ALIAS(별칭)

- SQL Code
- SELECT 컬럼명 AS 별칭 FROM 테이블명:

- 사용하는 경우
- SELECT 구문에서 별칭을 사용하는 경우, 컬럼명을 바꾸기 위해서 사용.
- 뒤에서 배우는 서브쿼리에서는 테이블 등에 별칭을 부여할 수 있음.

ALIAS(별칭)

• 원본테이블

수학	영어	이름
75	100	이상훈
99	50	김상훈
40	30	조상훈

• 컬럼의 합



컬럼의 합을 별칭으로 표현

SELECT 이름 FROM ex1 ;	름,영어,	수학, 수	≐학+영어	AS 영수
이름	영어	수학	영수	
이상훈	100	75	175	
김상훈	50	99	149	
조상훈	30	40	70	

LIMIT / OFFSET

- LIMIT과 OFFSET은 주로 결과 집합의 특정 부분을 제한하거나 건너뛸 때 사용
- LIMIT 은 반환되는 레코드의 수를 제한하는 데 사용
- 예를 들어, 결과의 처음 5개 레코드만을 선택하려면 LIMIT 5을 사용

LIMIT / OFFSET

- OFFSET 은 처음 몇 개의 레코드를 건너뛸 것인지를 지정하는 데 사용
- 예를 들어, 처음 5개 레코드를 건너뛰고 그 다음 10개 레코드를 선택하려면 LIMIT 10 OFFSET 5를 사용합니다

LIMIT n

• 원본테이블

번호	이름	국어	영어	수학
1	a	64	2	97
2	b	86	40	33
3	С	92	20	86
4	d	58	27	13
5	e	42	39	2
6	f	75	85	3
7	g	40	71	65
8	h	36	65	57
9	i	6	29	91
10	g	30	28	44
11	k	10	45	88
12	1	51	88	71
13	m	38	50	71
14	j	86	62	5
15	0	16	88	64
16	p	35	80	4
17	q	73	56	34
18	r	72	26	97

SQL Code

```
SELECT * FROM exercise1 LIMIT 5
:
```

- Limit n : 0번째 행부터 5개의 행 출력

• 결과

번호	이름	국어	영어	수학
1	a	64	2	97
2	Ь	86	40	33
3	c	92	20	86
4	d	58	27	13
5	e	42	39	2

LIMIT n, m

• 원본테이블

변호	이름	국어	영어	수학
1	a	64	2	97
2	Ь	86	40	33
3	c	92	20	86
4	d	58	27	13
5	e	42	39	2
6	f	75	85	3
7	g	40	71	65
8	h	36	65	57
9	i	6	29	91
10	g	30	28	44
11	k	10	45	88
12	1	51	88	71
13	m	38	50	71
14	j	86	62	5
15	0	16	88	64
16	р	35	80	4
17	q	73	56	34
18	r	72	26	97

SQL Code

```
SELECT * FROM exercise1 LIMIT 5, 3;
```

- Limit n : 5번째 행부터 3개의 행을 출력

• 결과

번호	이름	국어	영어	수학
6	f	75	85	3
7	g	40	71	65
8	h	36	65	57

LIMIT n OFFSET m

• 원본테이블

변호	이름	국어	영어	수학
1	a	64	2	97
2	Ь	86	40	33
3	С	92	20	86
4	d	58	27	13
5	e	42	39	2
6	f	75	85	3
7	g	40	71	65
8	h	36	65	57
9	i	6	29	91
10	g	30	28	44
11	k	10	45	88
12	1	51	88	71
13	m	38	50	71
14	j	86	62	5
15	0	16	88	64
16	р	35	80	4
17	q	73	56	34
18	r	72	26	97

SQL Code

SELECT * FROM exercise1 LIMIT 3 OFFSET
5;

- OFFSET 으로 5개의 데이터 다음 3개의 데이터 출력.
- LIMIT 5,3 과 같은 결과
- 결과

번호	이름	국어	영어	수학
6	f	75	85	3
7	g	40	71	65
8	h	36	65	57

DISTINCT

• 원본테이블

번호	이름	국어	영어	수학
1	a	64	2	97
2	Ь	86	40	33
3	C	92	20	86
4	d	58	27	13
5	e	42	39	2
6	f	75	85	3
7	g	40	71	65
8	h	36	65	57
9	i	6	29	91
10	g	30	28	44
11	k	10	45	88
12	1	51	88	71
13	m	38	50	71
14	j	86	62	5
15	0	16	88	64
16	р	35	80	4
17	q	73	56	34
18	r	72	26	97

SQL Code

SELECT DISTINCT 수학 FROM exercise1

- 수학 컬럼에서 중복된 항 하나만 남기고 선택
- 단점은 그 행만 보여지는것

	_
수학	
	97
	33
	86
	13
	2
	3
	65
	57
	91
	44
	88
	71
	5
	64
	4
	34

문제 1: customers 테이블에서 customerName을 가져오되,
 중복 없이 가져오세요.

정답

SELECT DISTINCT customerName
FROM customers;

문제 2: products 테이블에서 제품 이름(productName)만
 5개만 가져오세요.

정답

SELECT productName FROM products LIMIT 5;

 문제 3: orders 테이블에서 주문 상태(status)의 종류를 모두 나열하세요. 중복은 제거하세요.

정답

SELECT DISTINCT status
FROM orders;

 문제 4: employees 테이블에서 직원의 성(lastName)을 10개만 가져오되, 5번째부터 시작하세요.

정답

SELECT lastName FROM employees LIMIT 4, 10;

문제 5: products 테이블의 productVendor를 별칭(alias)
 Vendor로 조회하세요.

정답

SELECT productVendor AS Vendor FROM products;

● 문제 6: orders 테이블에서 고객 번호(customerNumber)의 중복 없는 값을 7개만 가져오세요.

정답

SELECT DISTINCT customerNumber FROM orders
LIMIT 7;

 문제 7: employees 테이블에서 employeeNumber를 별칭 Employee_ID로 조회하되, 3번째부터 6개만 가져오세요.

정답

```
SELECT employeeNumber AS Employee_ID
FROM employees
LIMIT 2, 6;
```

 문제 8: offices 테이블에서 국가(country)의 종류를 중복 없이 나열하세요.

정답

SELECT DISTINCT country
FROM offices;

 문제 9: orderdetails 테이블에서 제품 코드(productCode)를 Code라는 별칭으로, 주문 수량(quantityOrdered)를 Quantity라는 별칭으로 조회하세요.

정답

SELECT productCode AS Code, quantityOrdered AS Quantity
FROM orderdetails;

 문제 10: payments 테이블에서 체크 번호(checkNumber)의 중복 없는 값을 10개만 가져오세요.

정답

SELECT DISTINCT checkNumber FROM payments LIMIT 10;