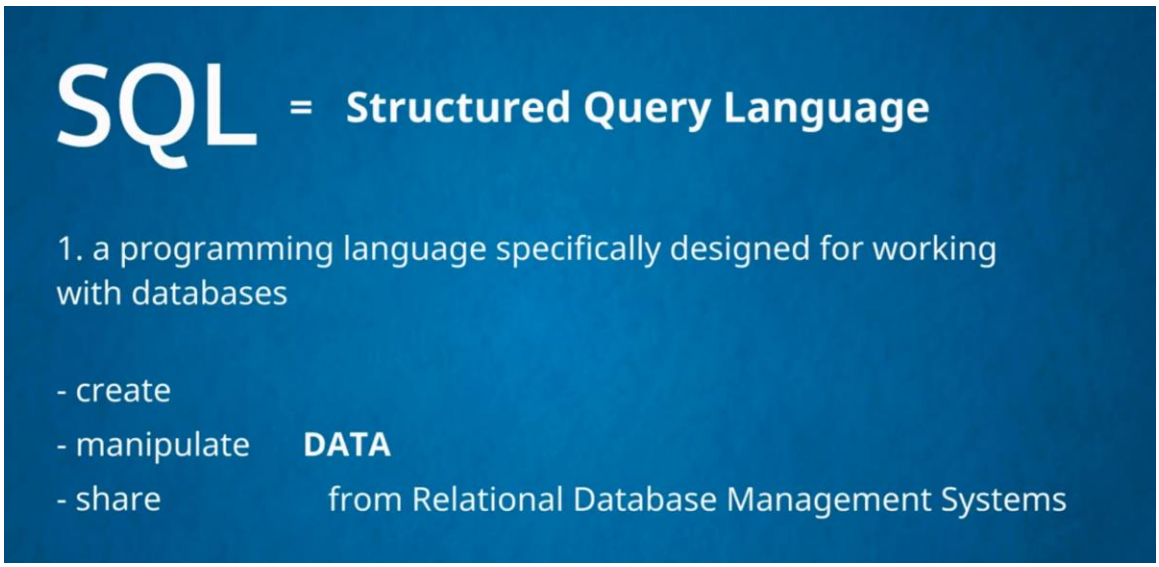
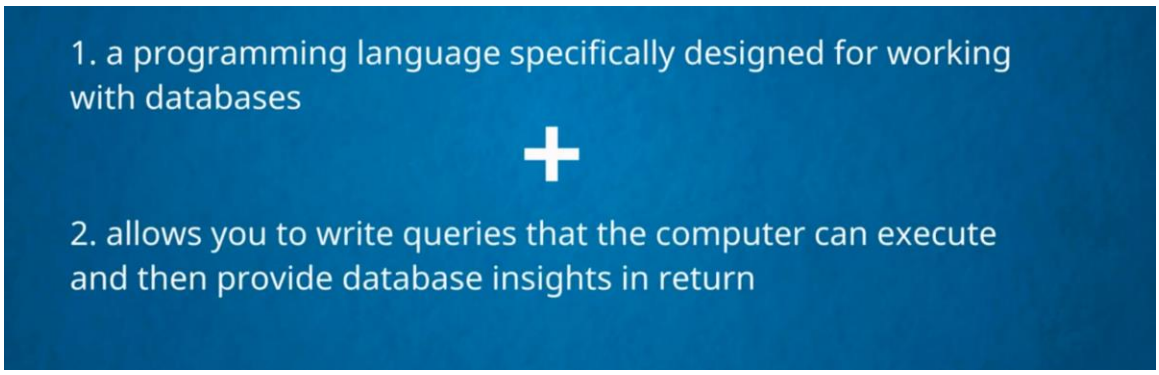
SECTION1: Overview

https://www.udemy.com/sql-mysql-for-data-analytics-and-business-intelligence/learn/lecture/8337984#overview





选择 Database management system！不同 system 之间区别不大。

市面上很多使用 SQL 语言的平台，但是 MySQL 比较好：

1. Open-source/免费
2. 使用排名靠前
3. Reliable, mature


MariaDB
ORACLE
MySQL
Microsoft
SQL Server
SQLite
open-source
(free)
MySQL
reliable
mature
Booking.com
YouTube
LinkedIn

Ranking > Complete Ranking

RSS RSS Feed

DB-Engines Ranking

The DB-Engines Ranking ranks database management systems according to their popularity. The ranking is updated monthly.

Read more about the method of calculating the scores.

trend chart

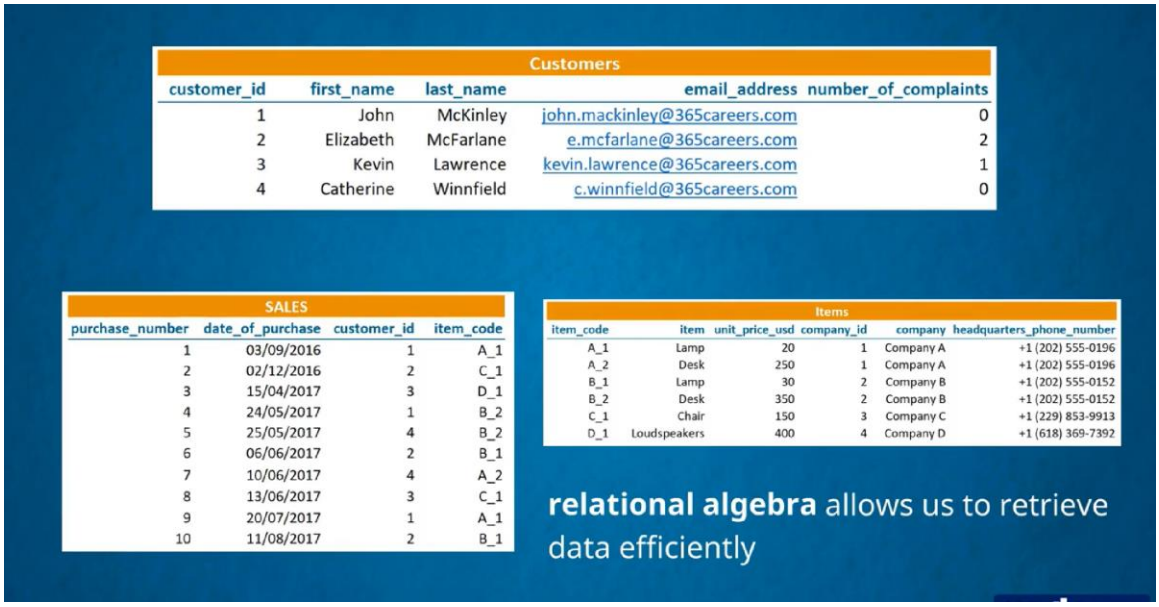
334 systems in ranking, September 2017

Rank Sep 2017	Rank Aug 2017	Rank Sep 2016	DBMS	Database Model	Score Sep 2017	Score Aug 2017	Score Sep 2016
1.	1.	1.	Oracle	Relational DBMS	1359.09	-8.78	-66.47
2.	2.	2.	MySQL	Relational DBMS	1312.61	-27.69	-41.41
3.	3.	3.	Microsoft SQL Server	Relational DBMS	1212.54	-12.93	+0.99
4.	4.	4.	PostgreSQL	Relational DBMS	372.36	+2.60	+56.01
5.	5.	5.	MongoDB	Document store	332.73	+2.24	+16.74
6.	6.	6.	DB2	Relational DBMS	198.34	+0.87	+17.15
7.	7.	↑ 8.	Microsoft Access	Relational DBMS	128.81	+1.78	+5.50
8.	8.	↓ 7.	Cassandra	Wide column store	126.20	-0.52	-4.29
9.	9.	↑ 10.	Redis	Key-value store	120.41	-1.49	+12.61
10.	10.	↑ 11.	Elasticsearch	Search engine	120.00	+2.35	+23.52

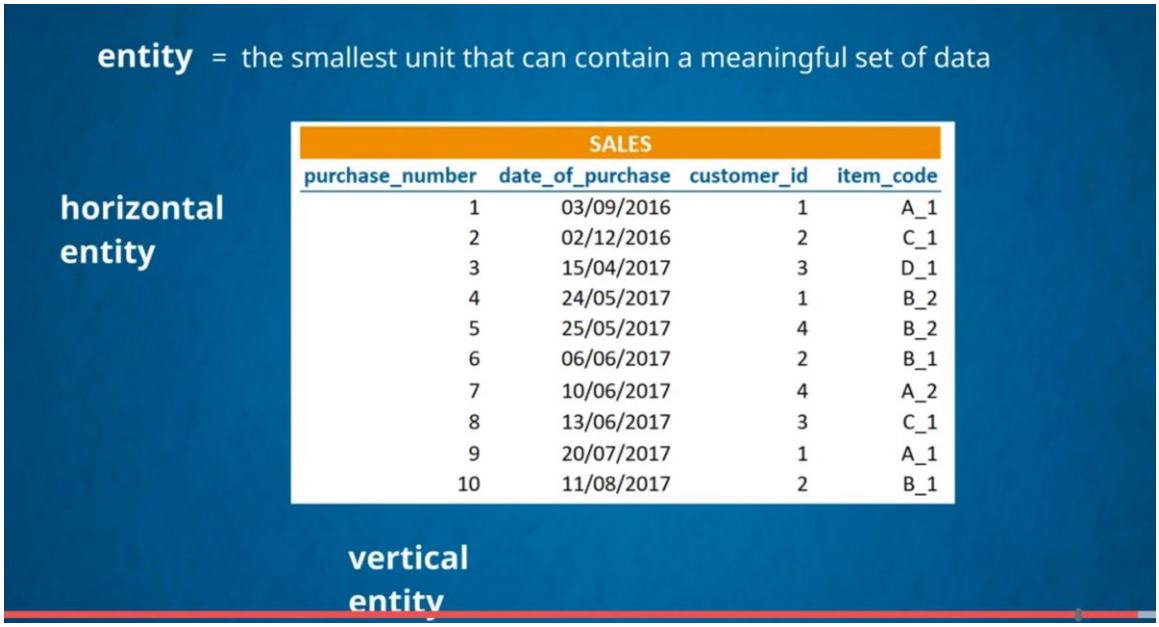
SQL 基于 relational database

Relational database：record*field， a few related tables form relational database

如果把所有数据放在一个 table 里，有太多列，人们不能很好的了解信息，所以按照主题放在不同 table 中

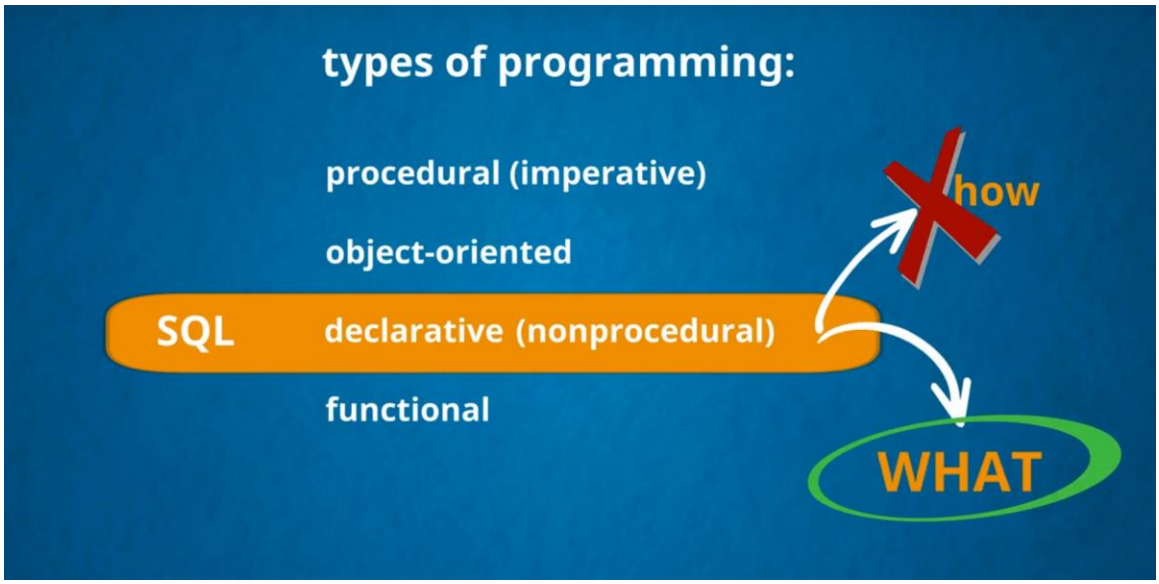


一个最小数据单位/表：entity



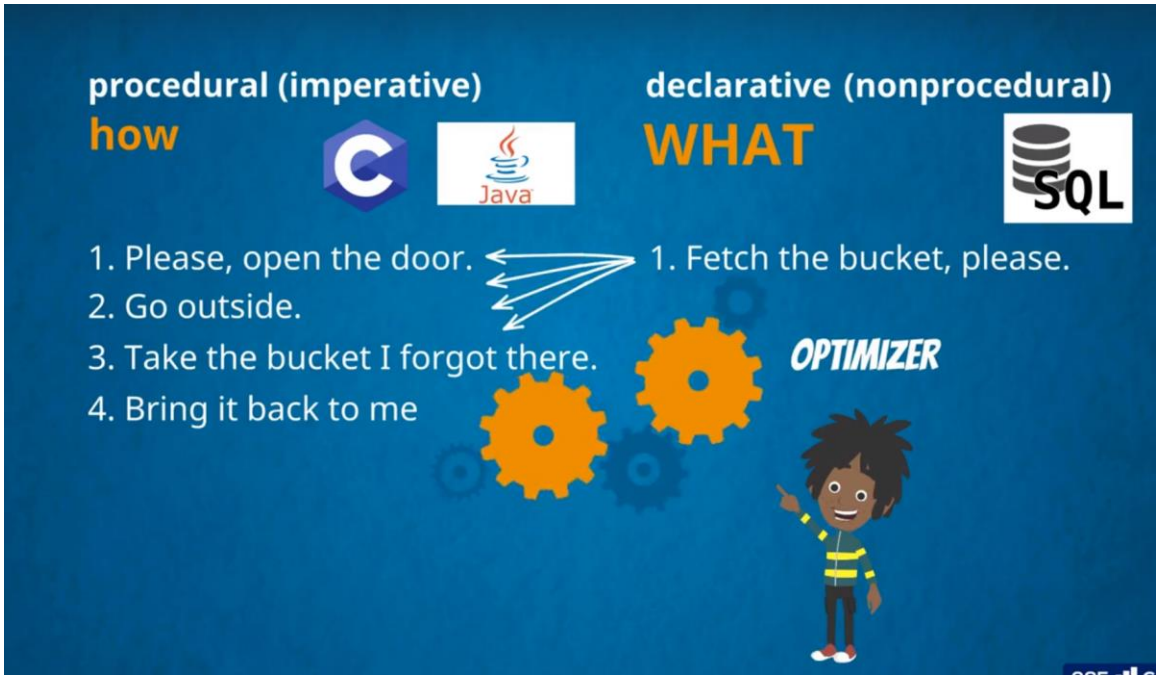
Section 2:

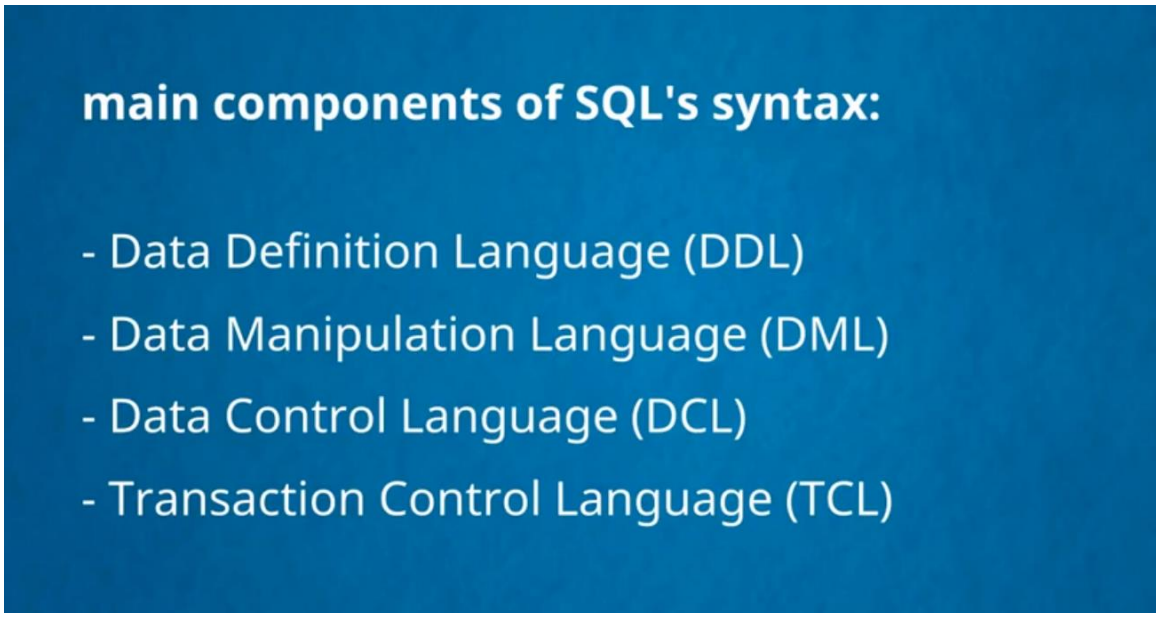
SQL: create & manipulate relational databases



Procedural 就是要好几步，

Declarative 一般一步到位，直接呈现结果 （其实有 built in optimizer 已经帮我们分好步 自行运行了） 所以关注结果/想要什么就行





创建对象：DDL,一般是 create

Data Definition Language

- **Data Definition Language (DDL)**
  - CREATE
  - ALTER
  - DROP
  - RENAME
  - TRUNCATE

Data Definition Language

- **SQL's syntax**

  comprises several types of statements that allow you to perform various commands and operations

- **Data Definition Language (DDL)**
  - a syntax
  - a set of statements that allow the user to define or modify data structures and objects, such as tables

- **the CREATE statement**

  used for creating entire databases and database objects as tables

Data Definition Language

- **the CREATE statement**

  used for creating entire databases and database objects as tables

SQL

```
CREATE object_type object_name;
CREATE TABLE object_name (column_name data_type);
```

Data Definition Language

- the ALTER statement
  used when altering existing objects

- ADD
- REMOVE
- RENAME

Data Definition Language

SQL

```
ALTER TABLE sales
ADD COLUMN date_of_purchase DATE;
```

sales

purchase_number	date_of_purchase

Data Definition Language

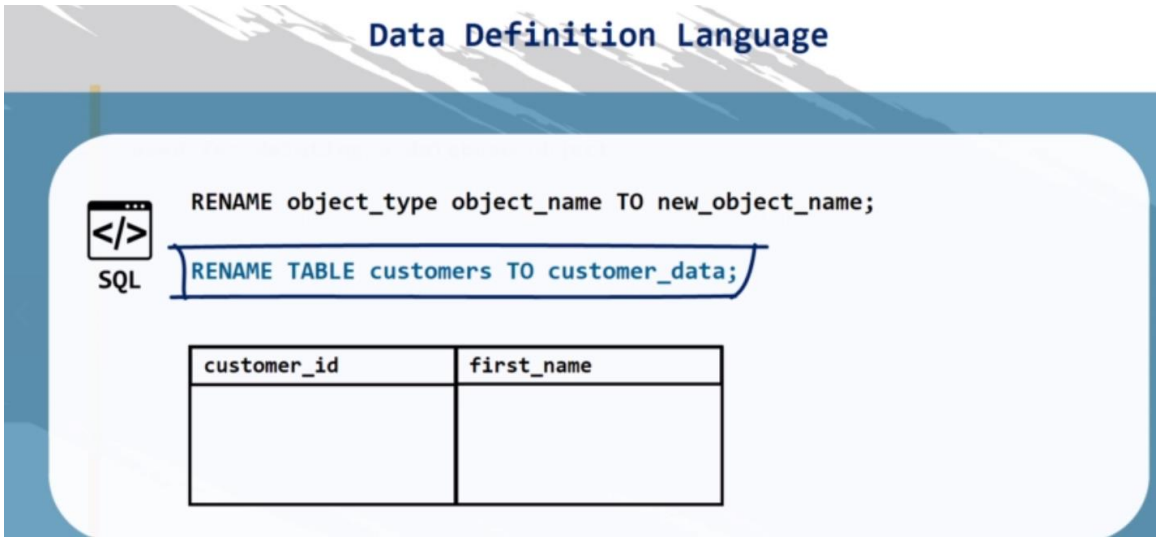
SQL

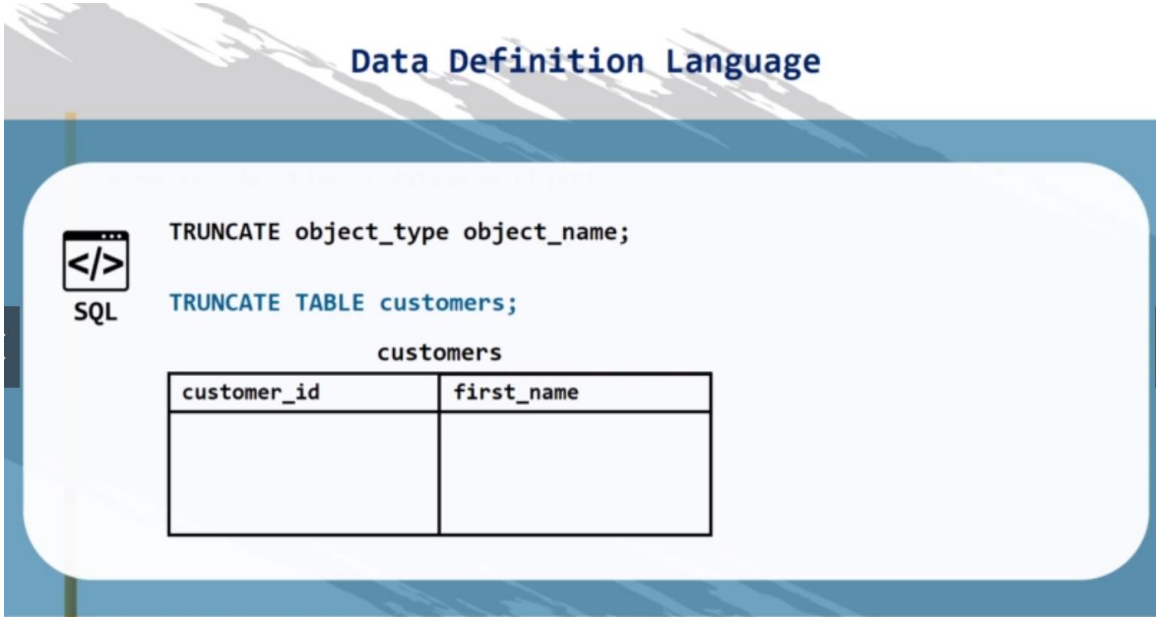
```
DROP object_type object_name;
```

```
DROP TABLE customers;
```

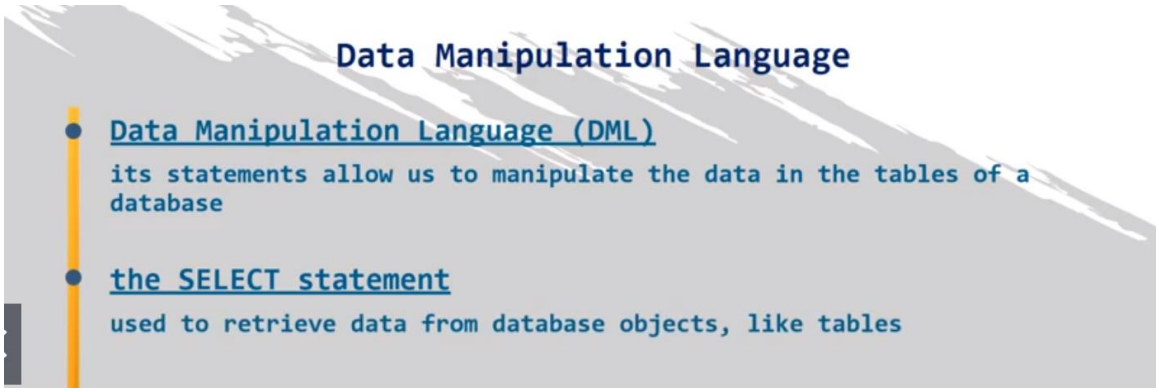
customers

customer_id	first_name





命名 var 时不能用函数



Data Manipulation Language

SQL

```
SELECT… FROM sales…;
```

sales

purchase_number

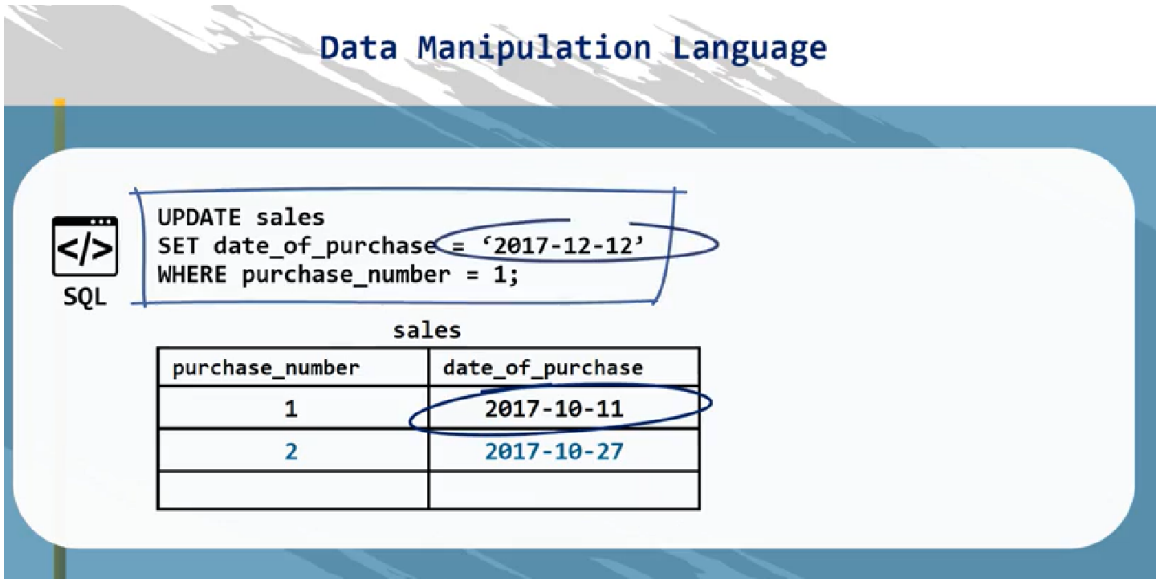
Data Manipulation Language

SQL

```
INSERT INTO sales (purchase_number, date_of_purchase) VALUES
(1, ‘2017-10-11’);
```

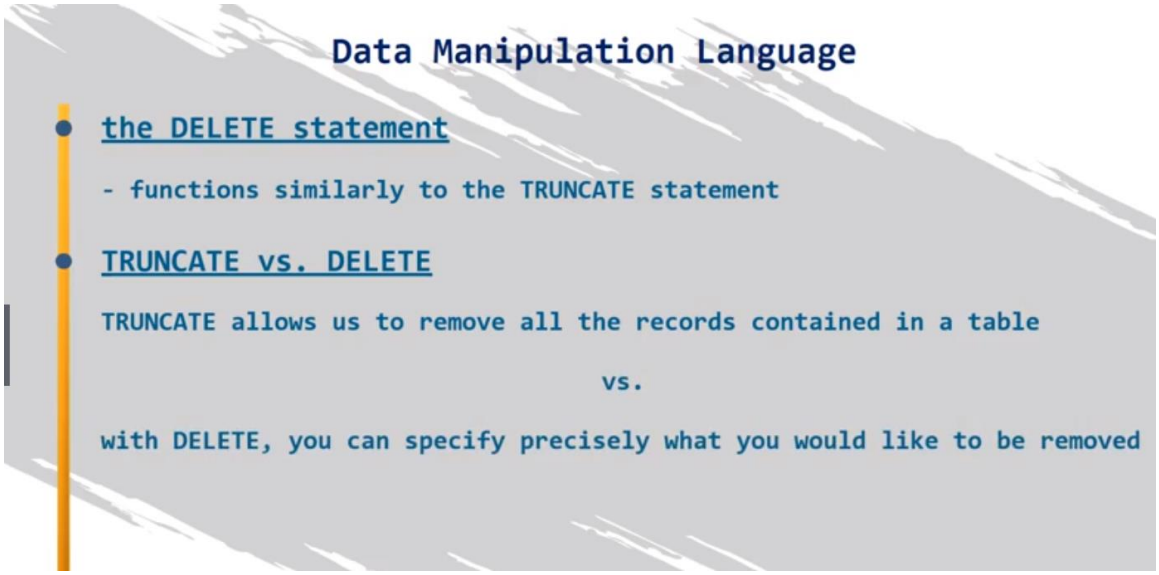
sales

purchase_number	date_of_purchase



Truncate 删所有行

Delete 可以 customize



Data Manipulation Language

SQL

```
DELETE FROM sales
WHERE
    purchase_number = 1;
```

sales

purchase_number	date_of_purchase
2	2017-10-27

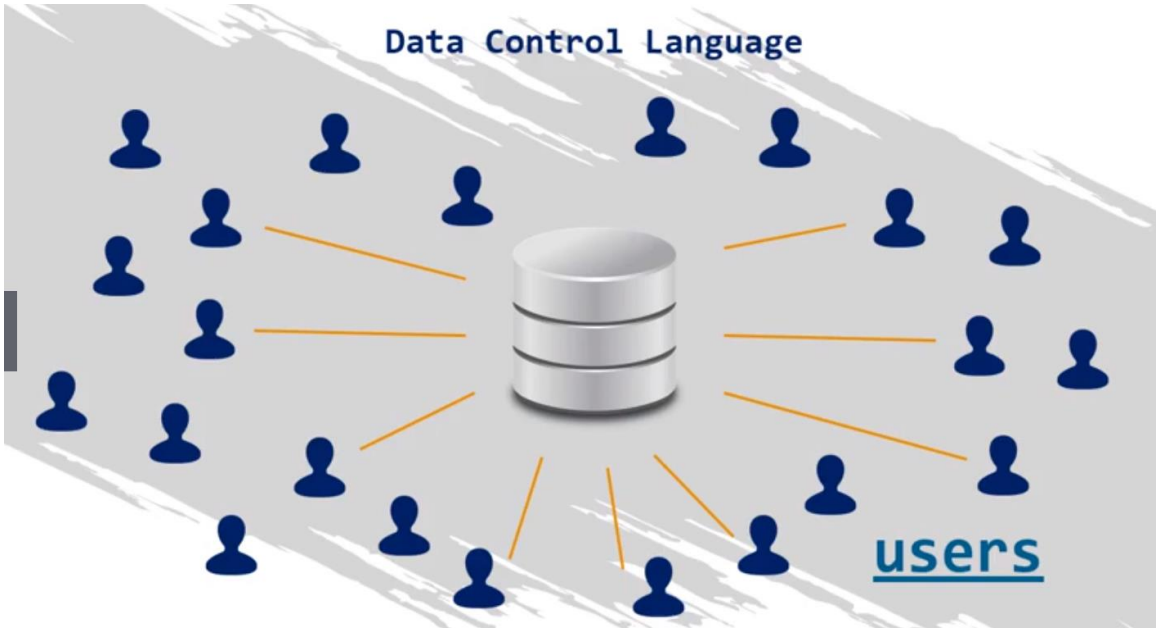
Data Manipulation Language

- Data Manipulation Language (DML)
  - SELECT… FROM…
  - INSERT INTO… VALUES
  - UPDATE… SET… WHERE
  - DELETE… FROM… WHERE

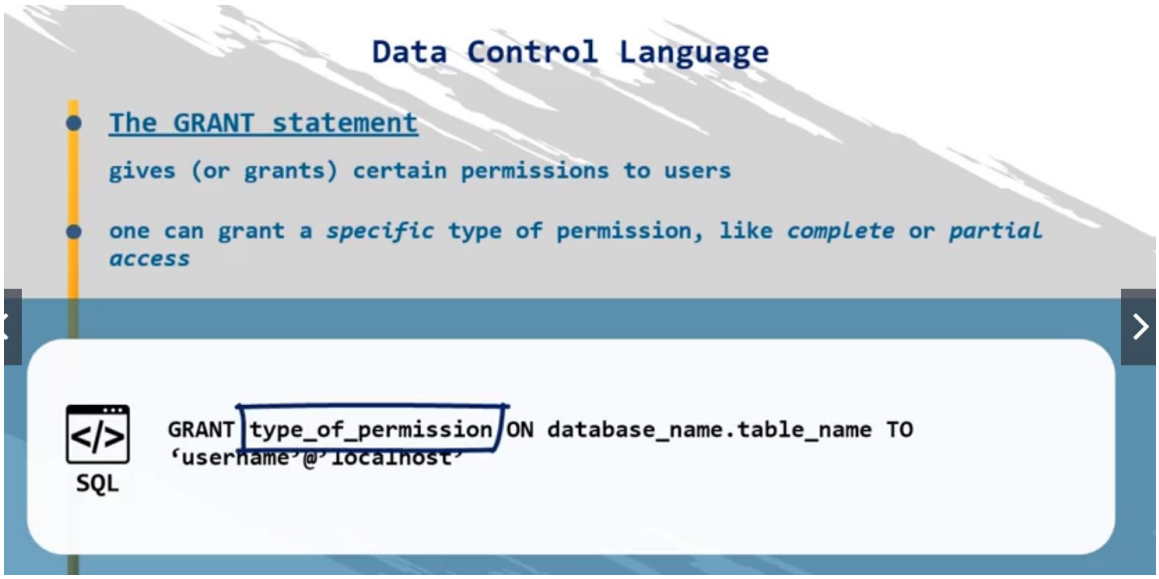
Data Control Language

- Data Control Language (DCL)
- the GRANT and REVOKE statements
  allow us to manage the rights users have in a database

User may have dif right to access



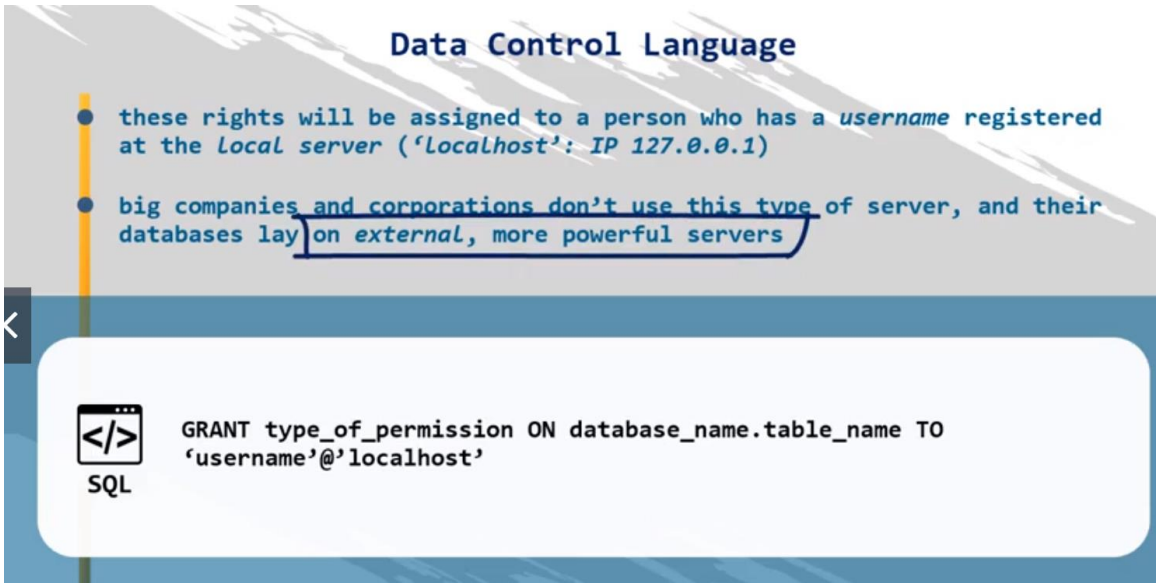
Grant 给某人 弄某表的某级别权限



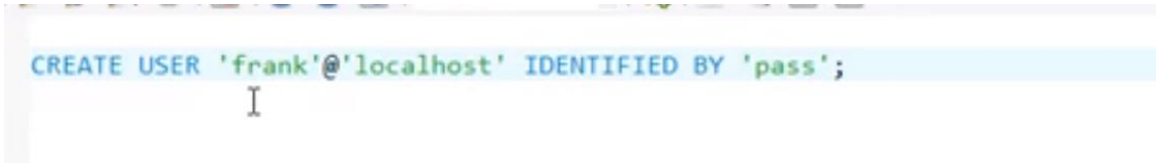
The server is based on our machine and it's commonly knownas "localhost", which is the domain name for the local IP address of our machine 127.0.0.1.

Naturally, big companies and corporations don't use this type of server and their databases lay on external, much more powerful servers specifically designed to store big amounts of data.

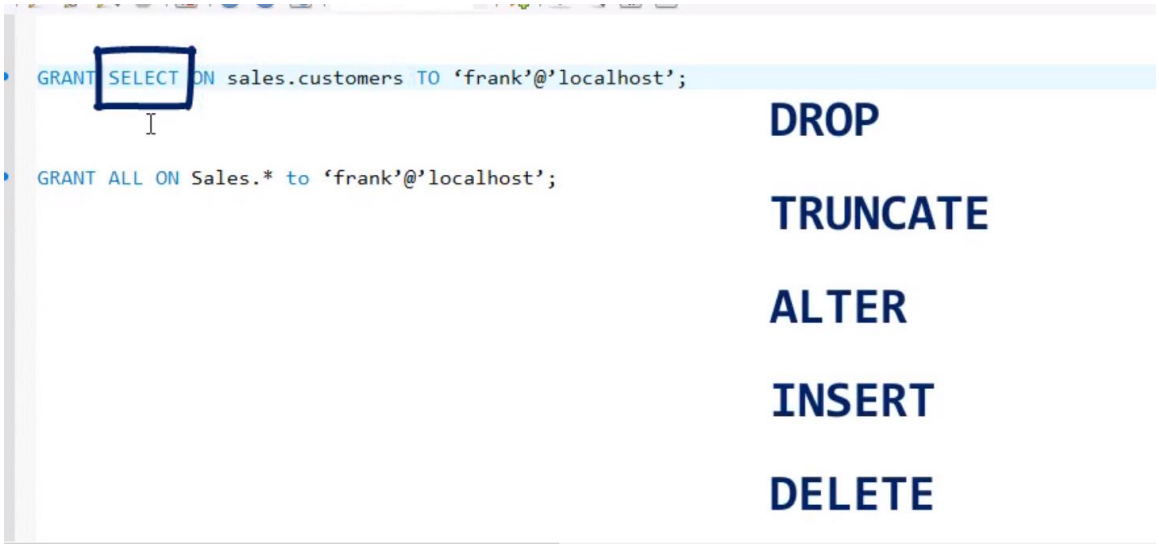
Those servers are not our "localhost" and are accessible on different IP addresses.



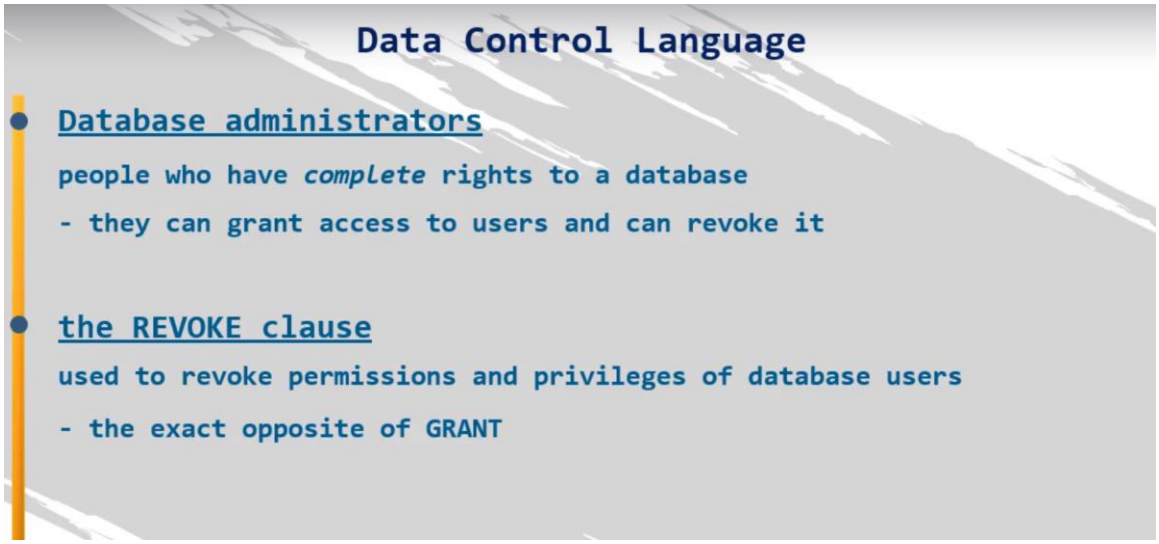
Create user，密码为 pass



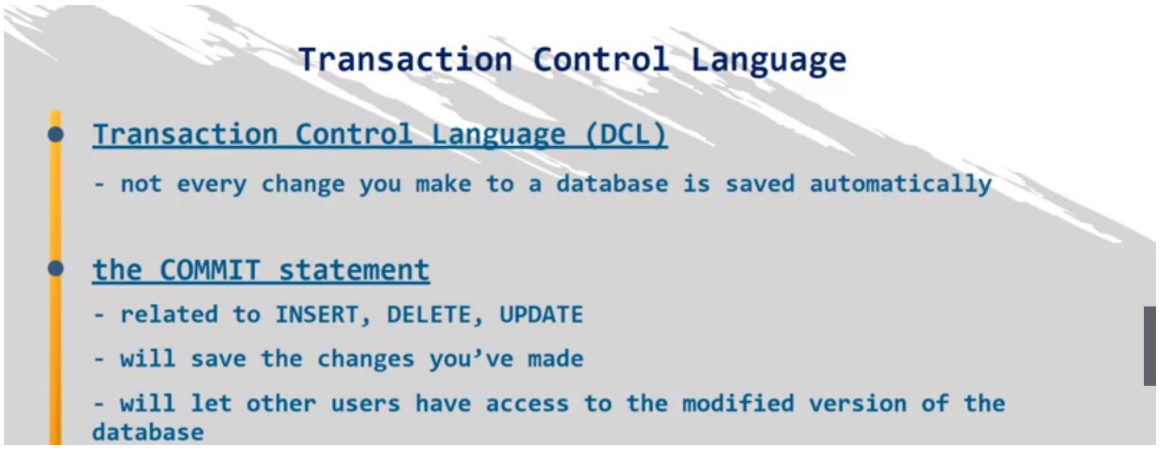
给这个人的权限是 select，及不能用右侧的，只能看到 salse 中的 customer



Administrator 有权授权 grant/撤权 revoke




```
REVOKE SELECT ON sales.customers FROM 'frank'@'localhost';
```



若某 user 更新数据库，如何全面更新，让其他 user 看到更新后的

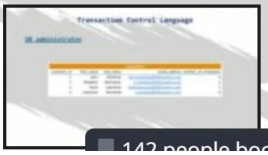
Transaction Control Language

DB administrator

Customers				
customer_id	first_name	last_name	email_address	number_of_complaints
1	John	McKinley	john.mackinley@365careers.com	0
2	Elizabeth	McFarlane	e.mcfarlane@365careers.com	2
3	Kevin	Lawrence	kevin.lawrence@365careers.com	1
4	Catherine	Johnson	c.winnfield@365careers.com	0

Problem:

users



Customers				
customer_id	first_name	last_name	email_address	number_of_complaints
1	John	McKinley	john.mackinley@365careers.com	0
2	Elizabeth	McFarlane	e.mcfarlane@365careers.com	2
3	Kevin	Lawrence	kevin.lawrence@365careers.com	1
4	Catherine	Winnfield	c.winnfield@365careers.com	0

142 people bookmarked this moment

Transaction Control Language

DB administrator

SQL

```
UPDATE customers
SET last_name = 'Johnson'
WHERE customer_id = 4
COMMIT;
```

Customers				
customer_id	first_name	last_name	email_address	number_of_complaints
1	John	McKinley	john.mackinley@365careers.com	0
2	Elizabeth	McFarlane	e.mcfarlane@365careers.com	2
3	Kevin	Lawrence	kevin.lawrence@365careers.com	1
4	Catherine	Johnson	c.winnfield@365careers.com	0

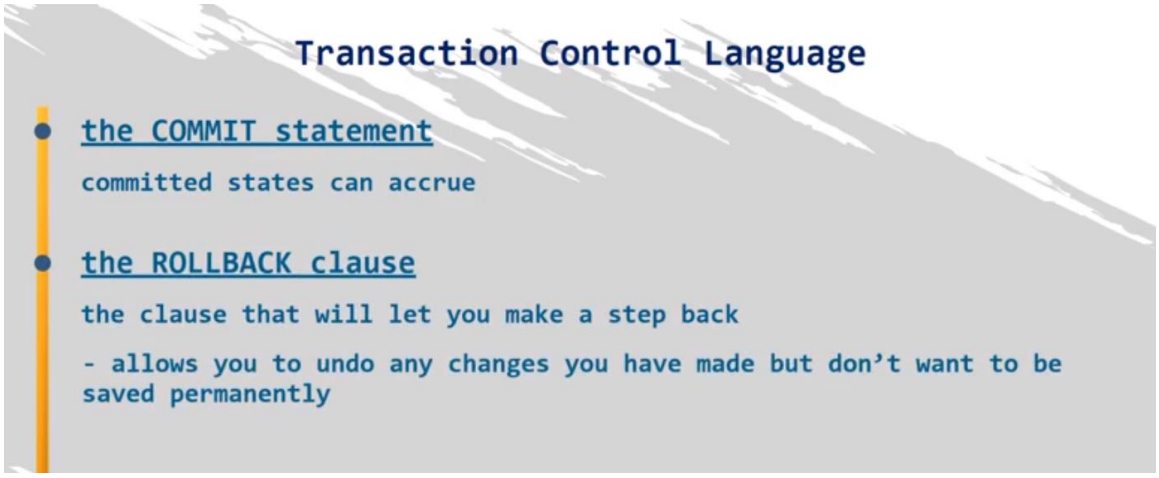
DB administrator

Customers				
customer_id	first_name	last_name	email_address	number_of_complaints
1	John	McKinley	john.mackinley@365careers.com	0
2	Elizabeth	McFarlane	e.mcfarlane@365careers.com	2
3	Kevin	Lawrence	kevin.lawrence@365careers.com	1
4	Catherine	Johnson	c.winnfield@365careers.com	0

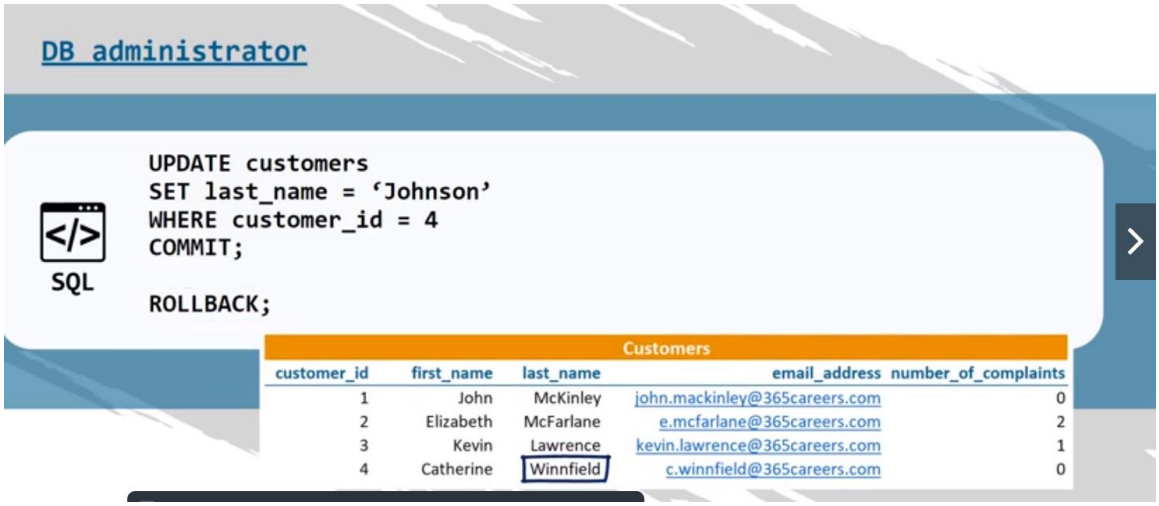
users

Customers				
customer_id	first_name	last_name	email_address	number_of_complaints
1	John	McKinley	john.mackinley@365careers.com	0
2	Elizabeth	McFarlane	e.mcfarlane@365careers.com	2
3	Kevin	Lawrence	kevin.lawrence@365careers.com	1
4	Catherine	Johnson	c.winnfield@365careers.com	0

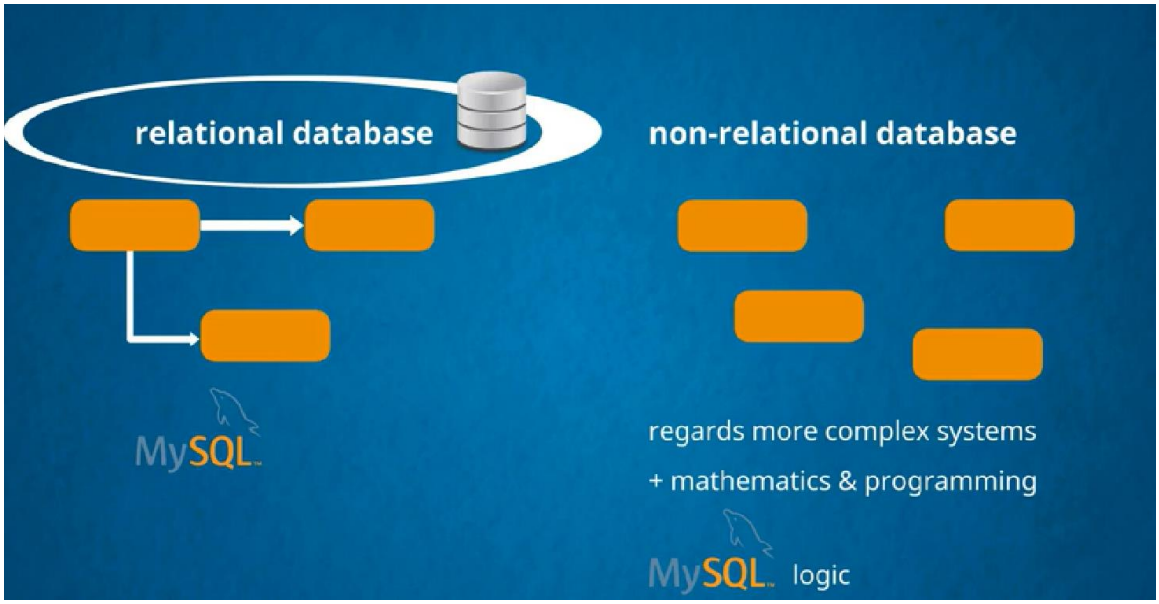
每次都点 commit 太麻烦



Roll back : revert to last non-commit state (不太清楚区别，可以看看 Database 书)

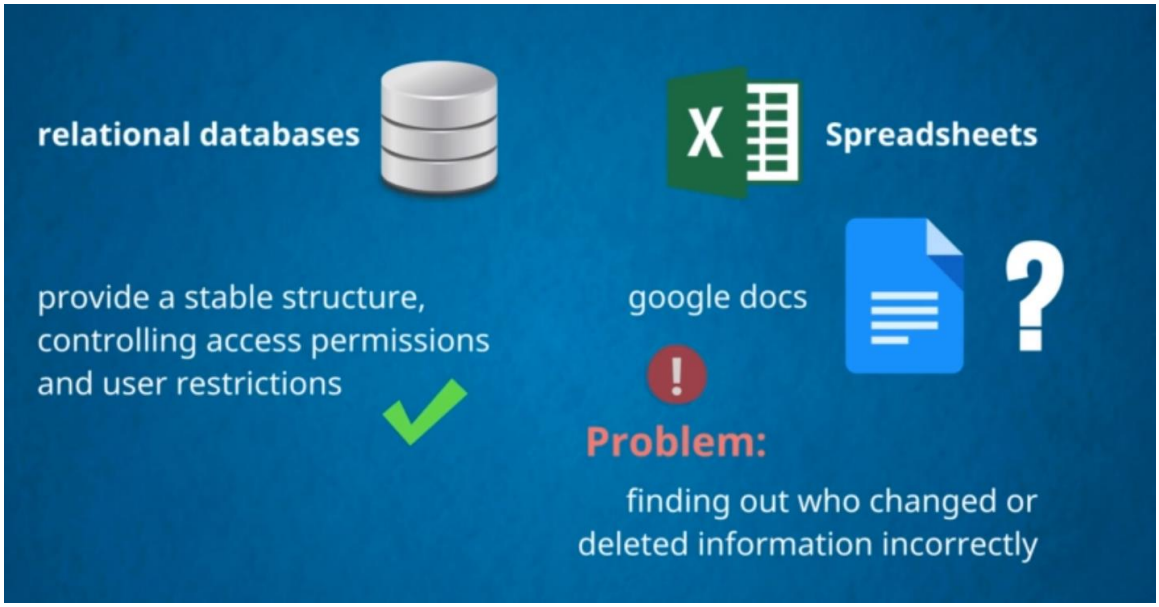


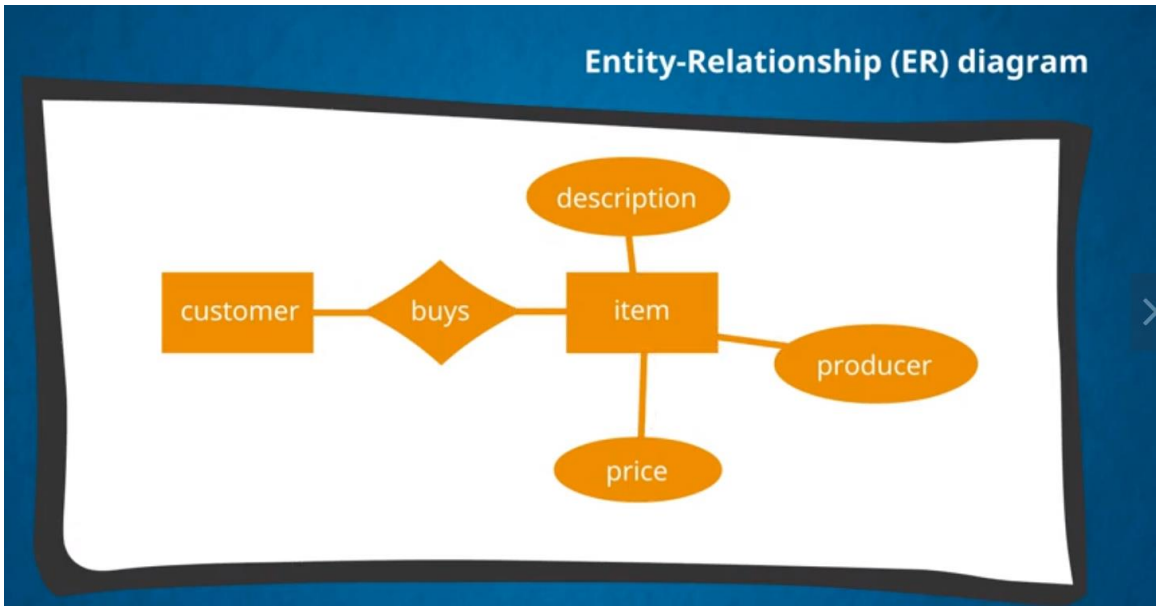
SECTION3: Relational Database

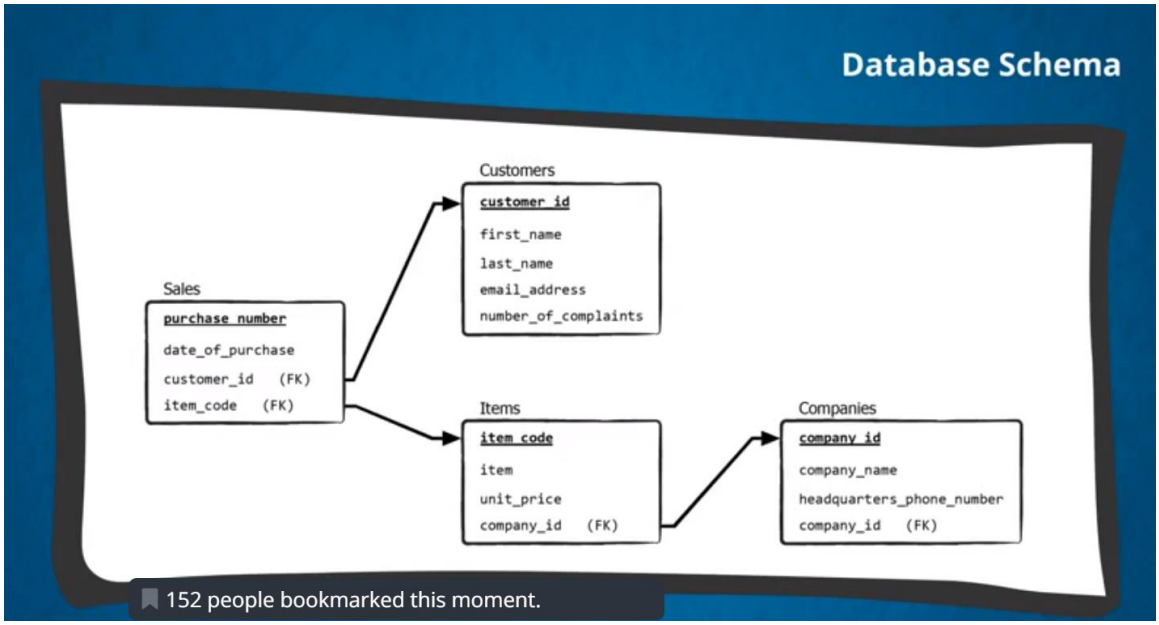


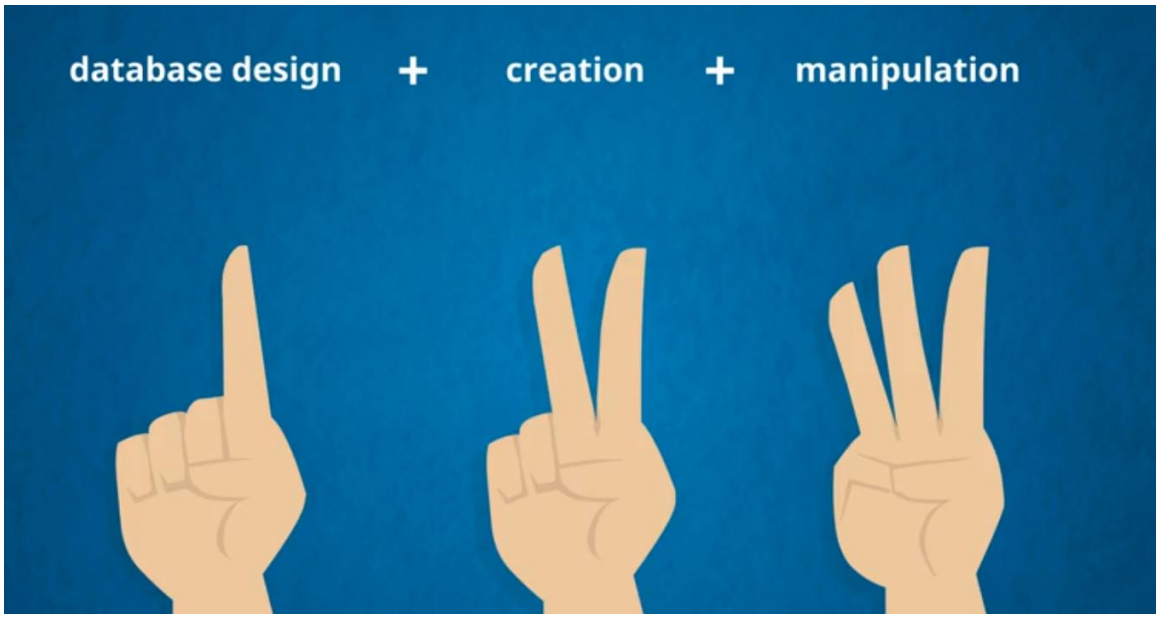
Customers

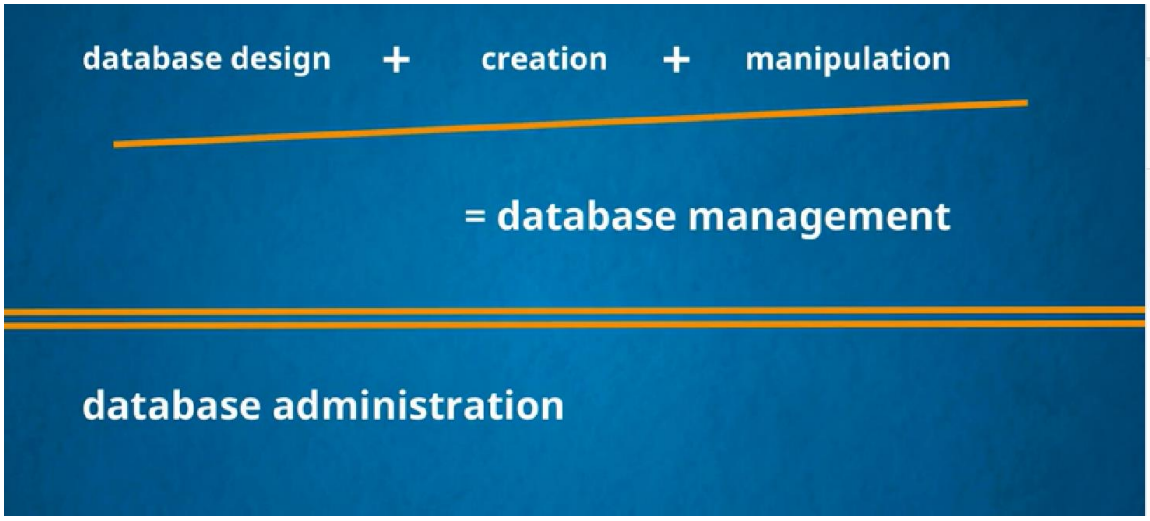
customer_id	first_name	last_name	email_address	number_of_complaints
1	John	McKinley	john.mackinley@365careers.com	0
2	Elizabeth	McFarlane	e.mcfarlane@365careers.com	2
3	Kevin	Lawrence	kevin.lawrence@365careers.com	1
4	Catherine	Winnfield	c.winnfield@365careers.com	0


relational databases
Spreadsheets
provide a stable structure,
controlling access permissions
and user restrictions
google docs
?
Problem:
finding out who changed or
deleted information incorrectly


Entity-Relationship (ER) diagram
description
customer
buys
item
producer
price

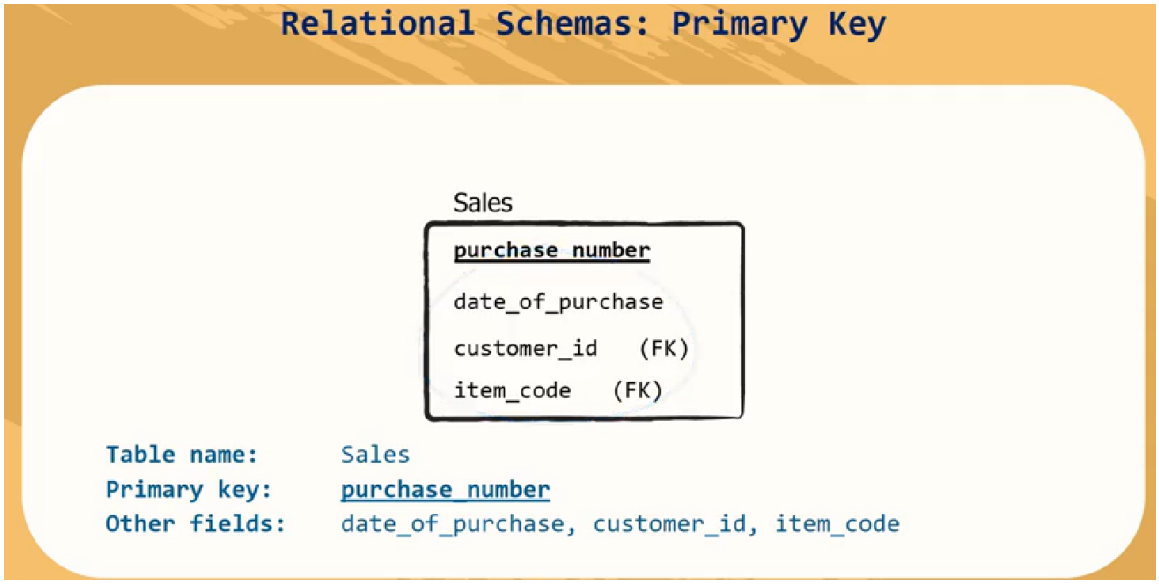

Database Schema
Customers
customer_id
first_name
last_name
email_address
number_of_complaints
Sales
purchase_number
date_of_purchase
customer_id (FK)
item_code (FK)
Items
item_code
item
unit_price
company_id (FK)
Companies
company_id
company_name
headquarters_phone_number
company_id (FK)
152 people bookmarked this moment.


database design + creation + manipulation



daily maintenance for database – administration

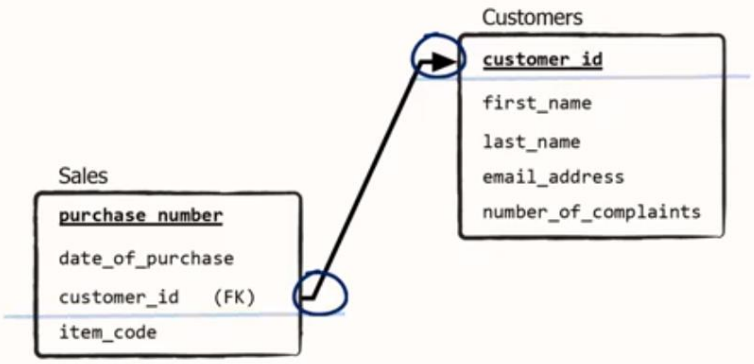
Primary key: 此表的 unique 区分每行的东西，可以是两列的组合 unique



Foreign key 此表连到其他表的 （箭头指向其他表）自己表内的 FK 可以重复（多对一）

Foreign key shows the relationships of databases

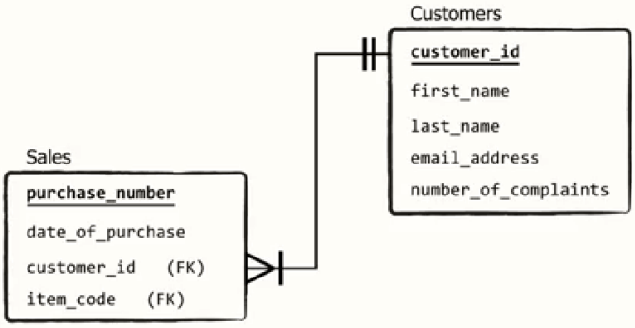
Relational Schemas: Foreign Key

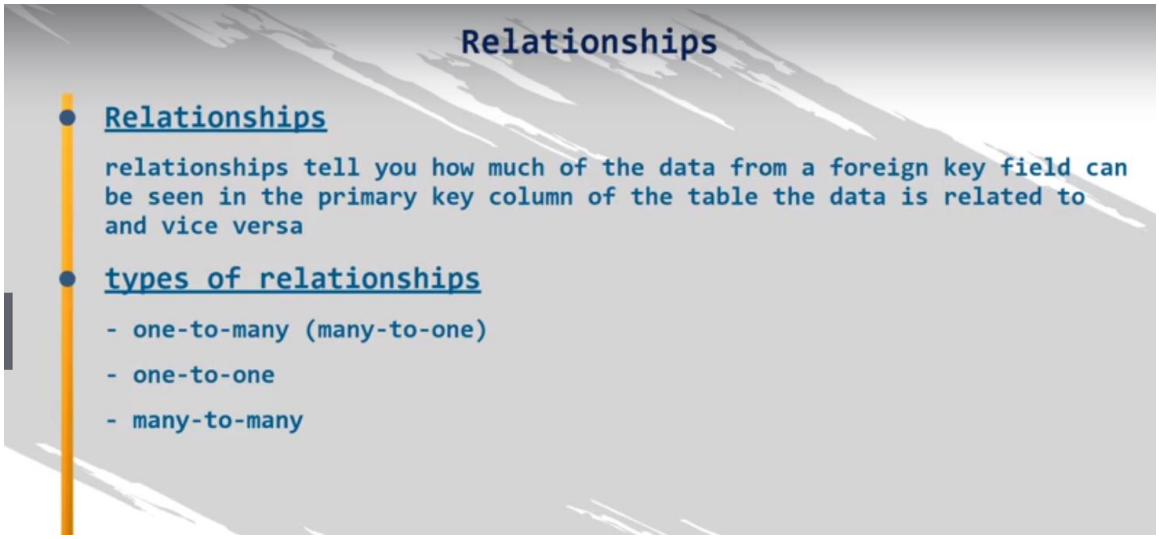


Relationships

Customers to Sales: *one-to-many*

Sales to Customers: *many-to-one*

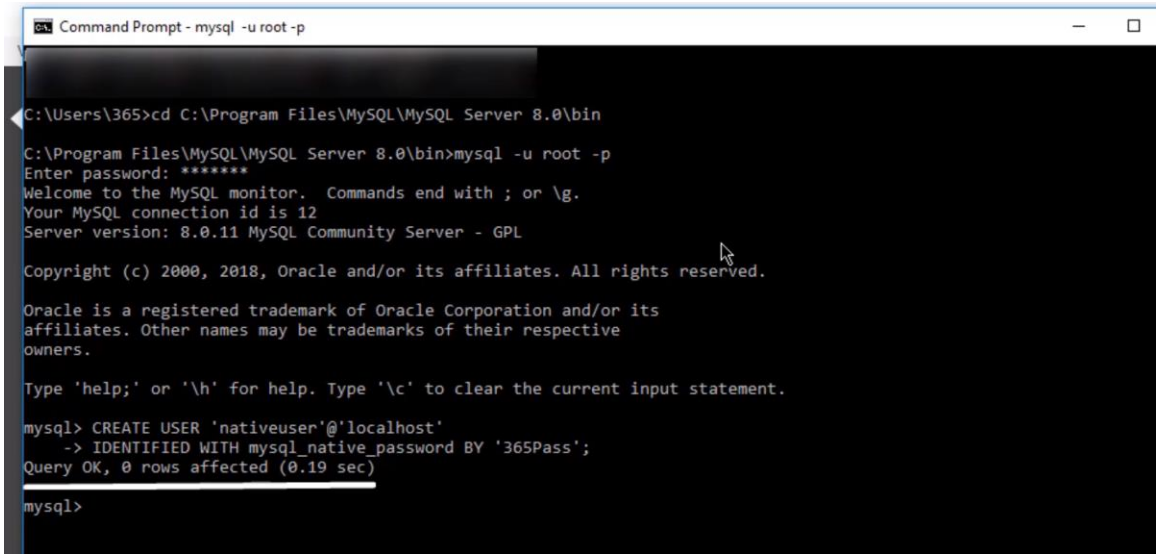


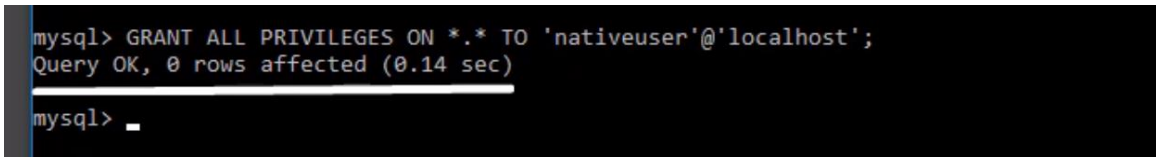


SECTION4: Install MySQL and get started

Mysql connection – 加号—新加 connection

可以 load multiple connections when using workbench





我的电脑里：“C:\Program Files\MySQL\MySQL Server 8.0\bin”

在 prompt 中输入

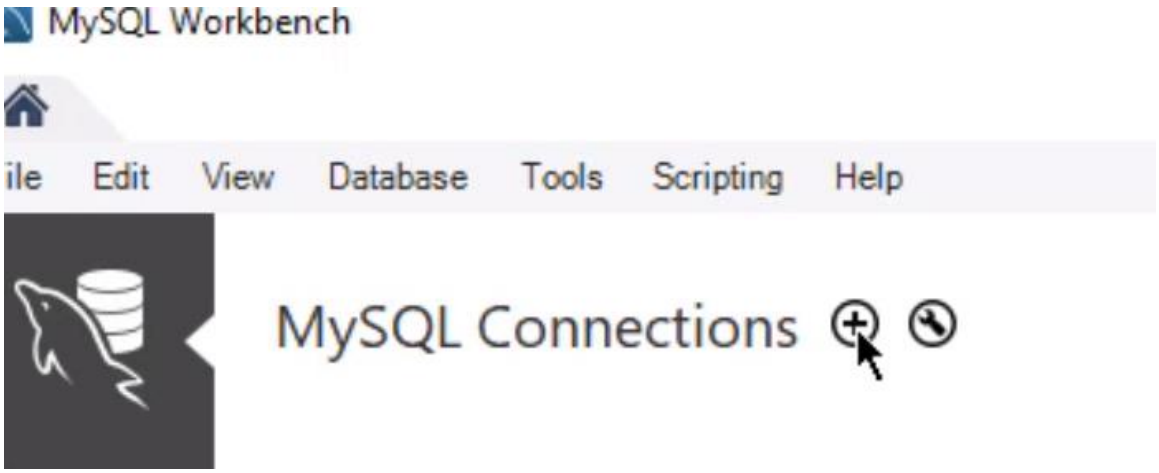
cd C:\Program Files\MySQL\MySQL Server 8.0\bin

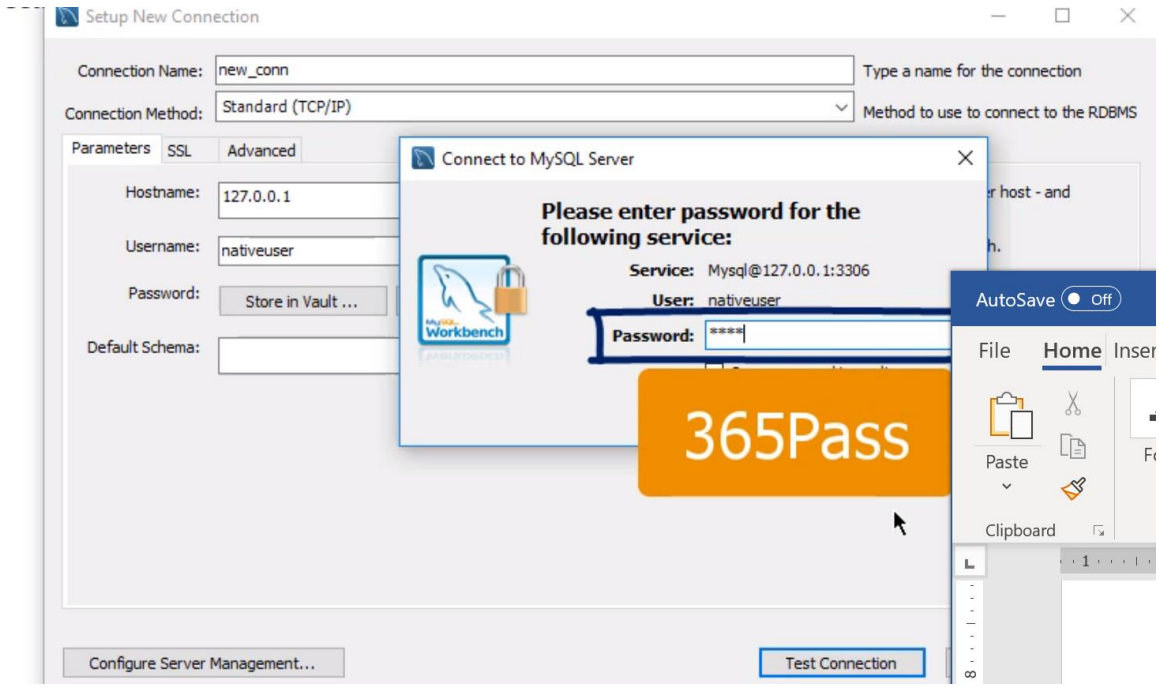
CREATE USER 'user2'@'localhost'

IDENTIFIED WITH mysql_native_password BY 'user2';

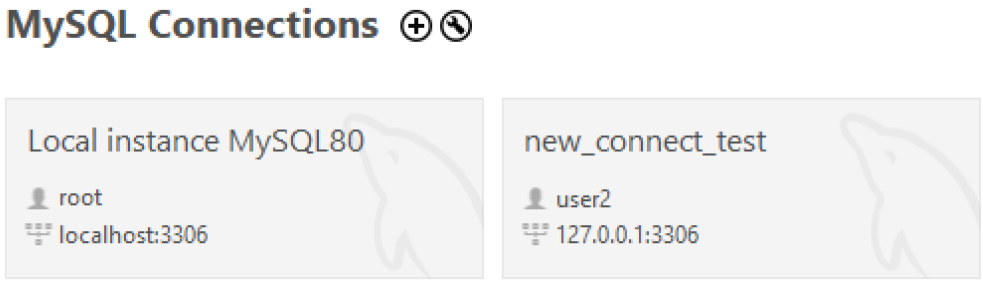
GRANT ALL PRIVILEGES ON *.* TO 'user2'@'localhost';

然后回到 sql 界面点这里建新的连接

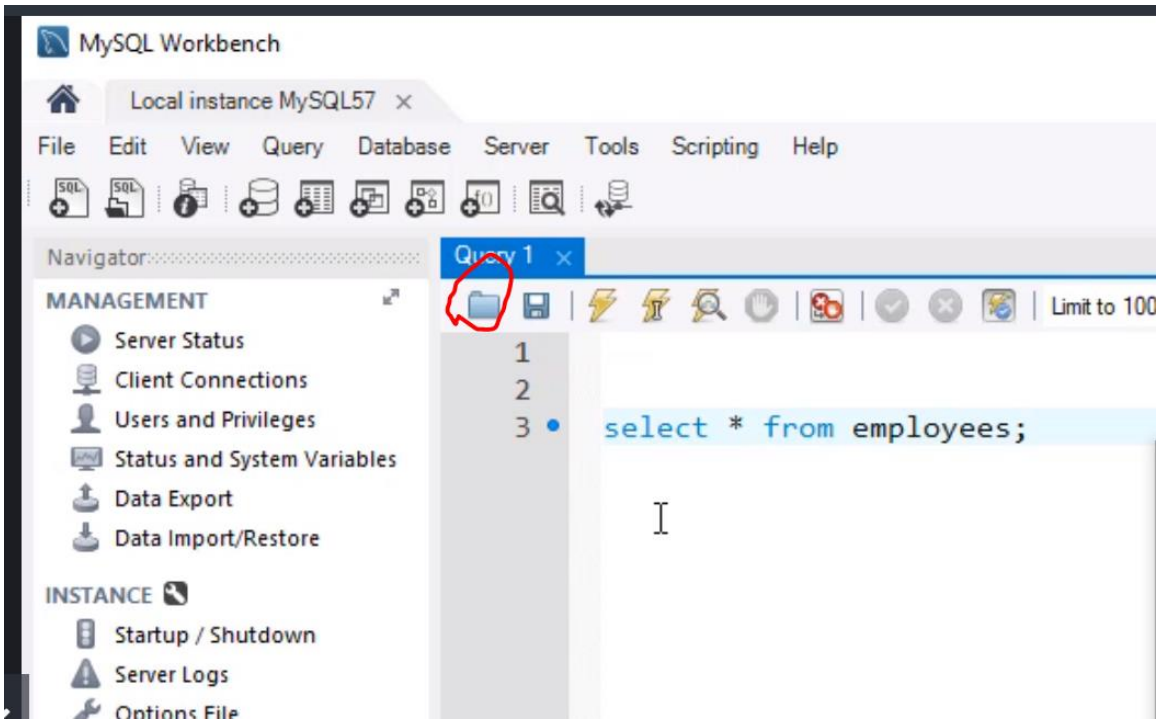




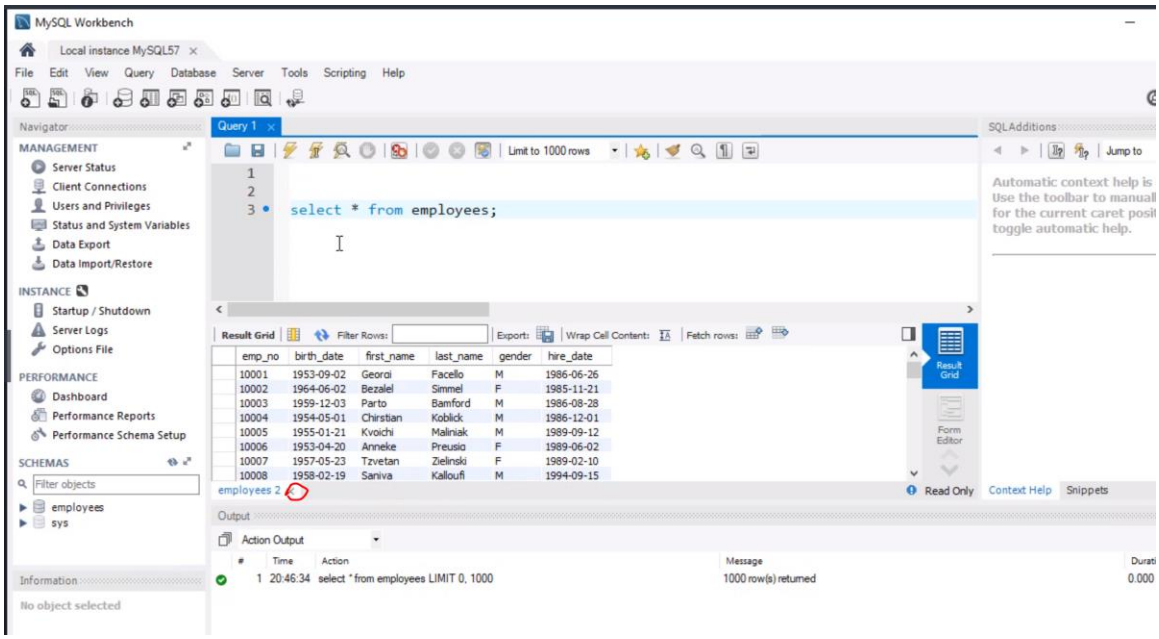
第一行随便填，后面都是按照设置的



熟悉操作界面

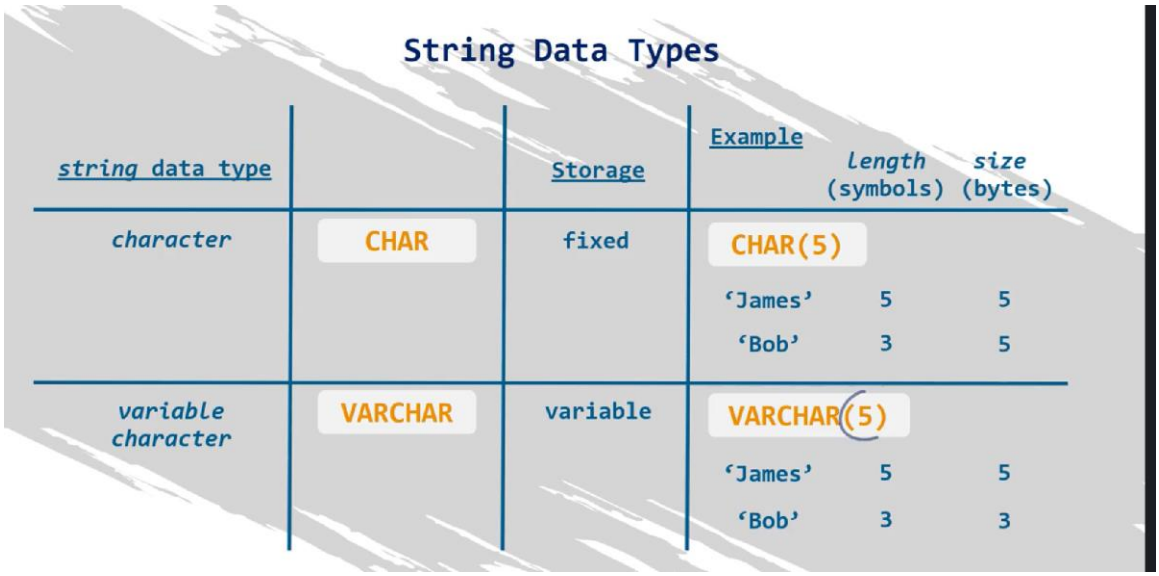


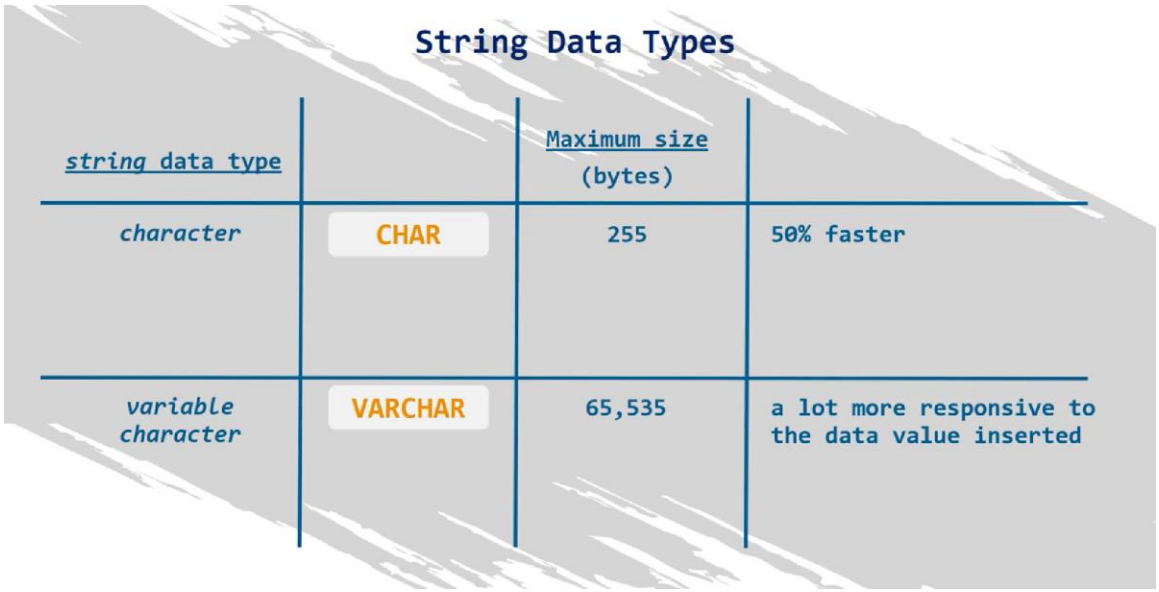
这个是直接打开一个新的 sql 文件



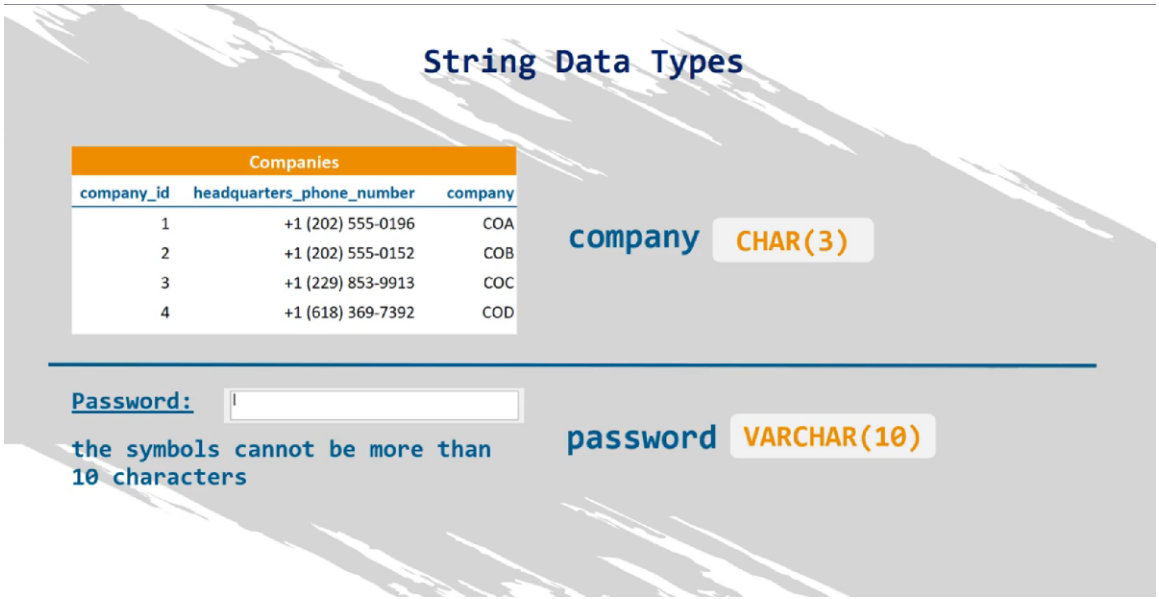
关掉一个 output

Section 5

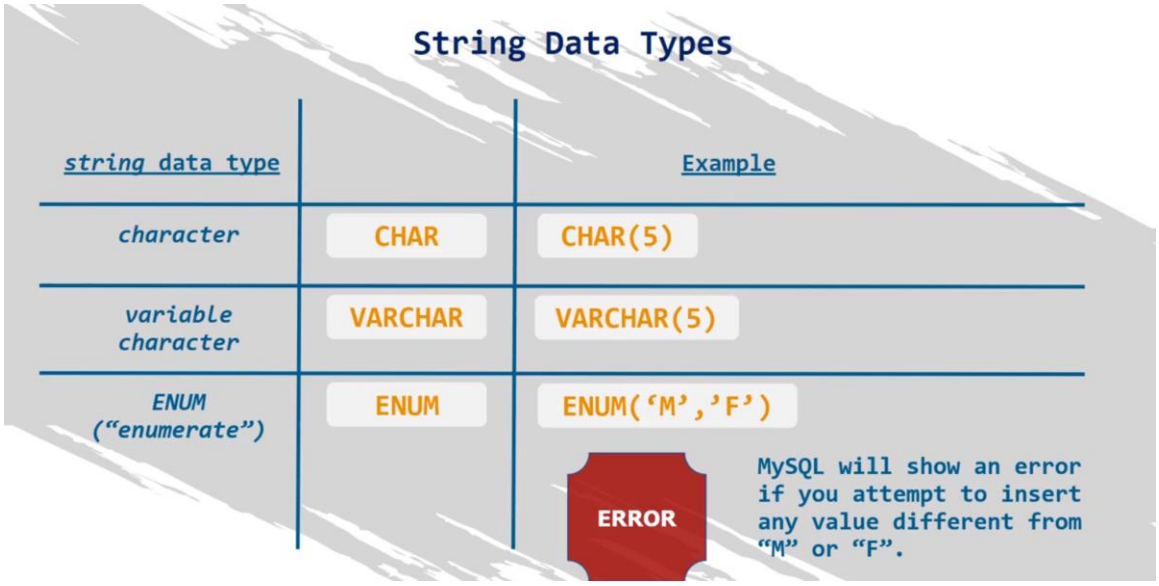




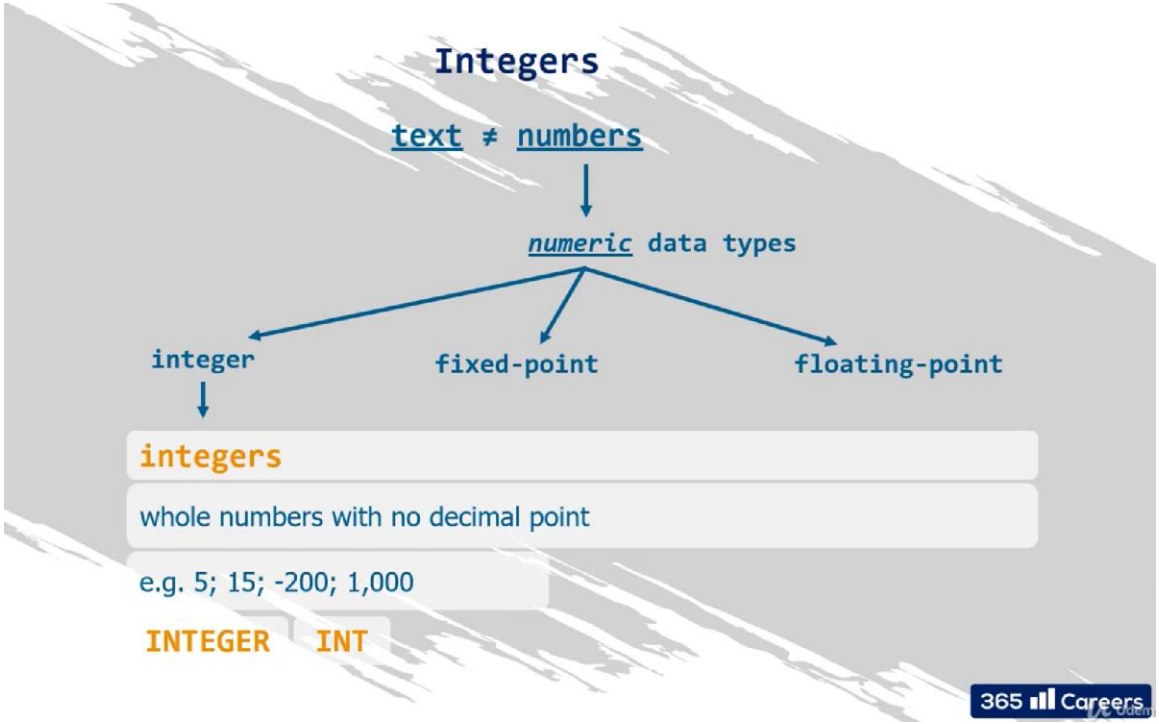
固定的长度：char(3)是刚刚好的三个 digit

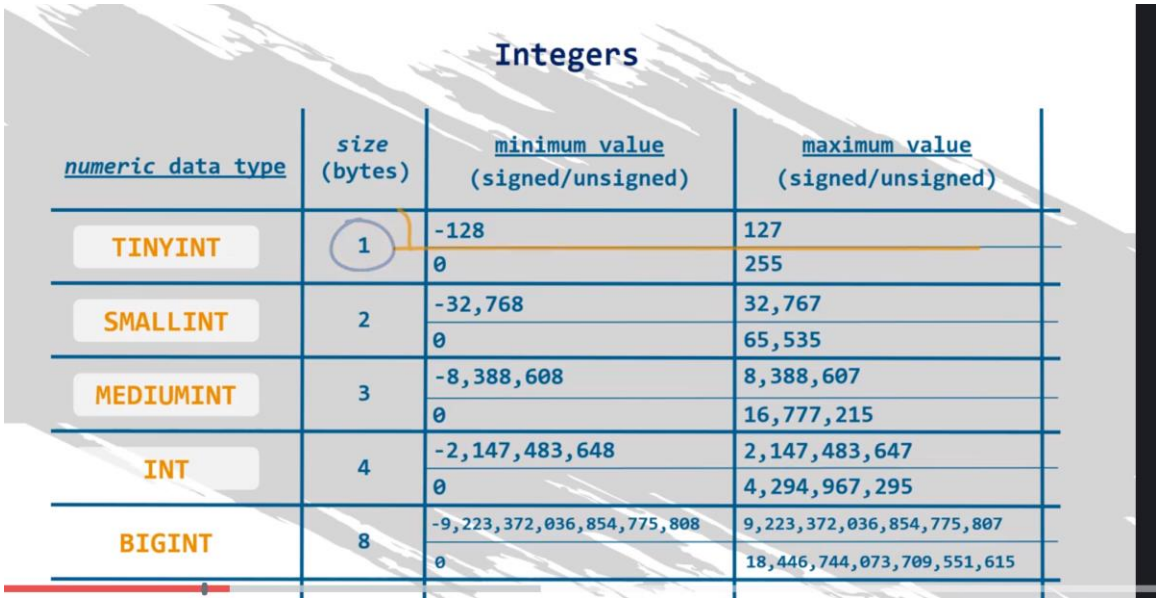


一共有三种 string data type,最后 enum 是枚举，列出能够在这里接受的所有 string


String Data Types
string data type
Example
character
CHAR
CHAR(5)
variable character
VARCHAR
VARCHAR(5)
ENUM ("enumerate")
ENUM
ENUM('M','F')
ERROR
MySQL will show an error if you attempt to insert any value different from "M" or "F".

Integers


Integers
text ≠ numbers
numeric data types
integer
fixed-point
floating-point
integers
whole numbers with no decimal point
e.g. 5; 15; -200; 1,000
INTEGER
INT
365 Careers



Integers

numeric data type	size (bytes)	minimum value (signed/unsigned)	maximum value (signed/unsigned)
TINYINT	1	-128	127
		0	255
SMALLINT	2	-32,768	32,767
		0	65,535
MEDIUMINT	3	-8,388,608	8,388,607
		0	16,777,215
INT	4	-2,147,483,648	2,147,483,647
		0	4,294,967,295
BIGINT	8	-9,223,372,036,854,775,808	9,223,372,036,854,775,807
		0	18,446,744,073,709,551,615

SIGNED: 正负值都可以

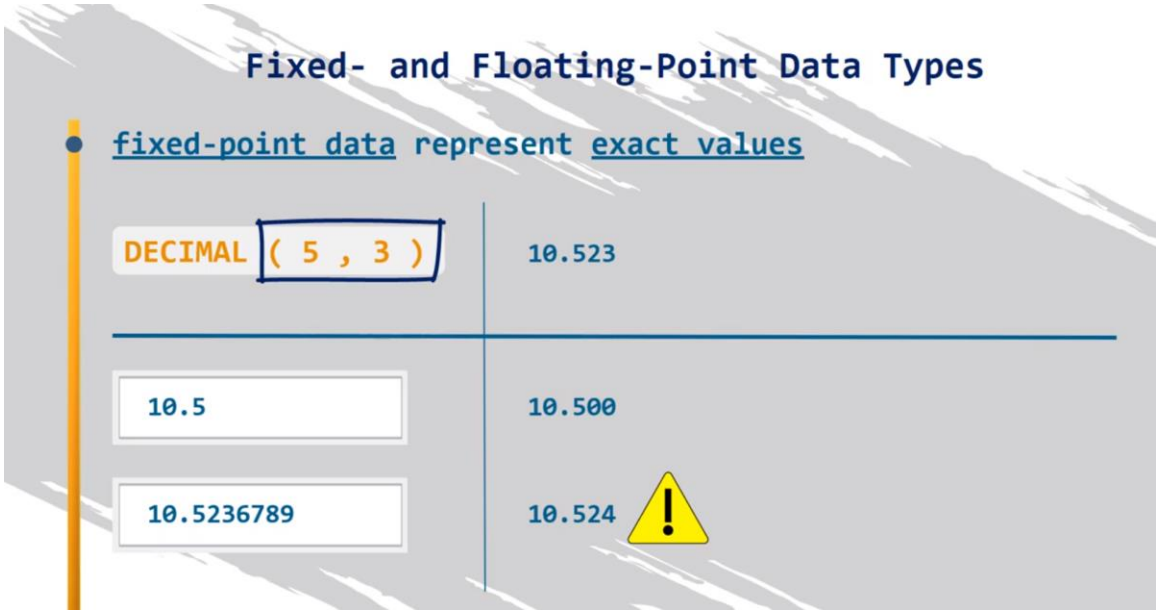
Unsigned: 只能是非负数

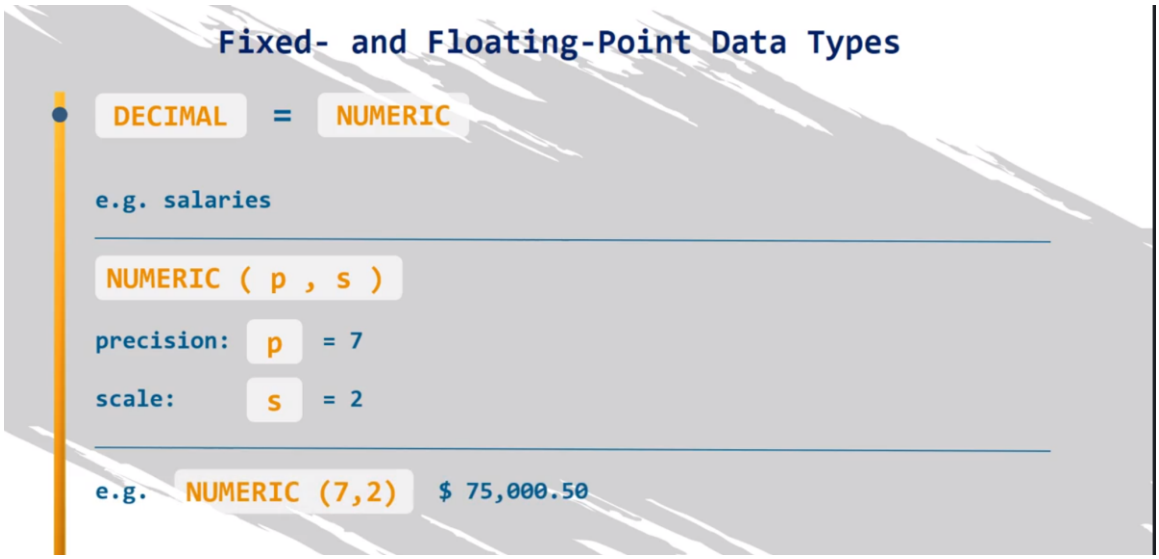
一般是 signed，若想 unsigned 要定义

选择刚好的 string 定义以降低 store space

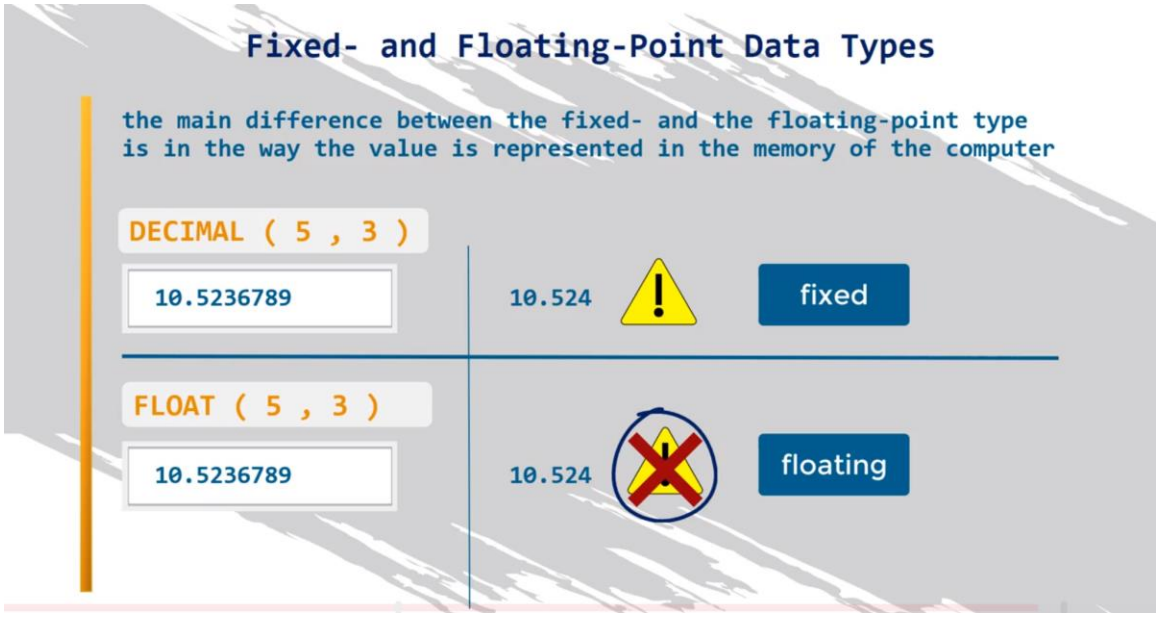
Decimal(总 digit 数，小数后 digit 数)

Decimal(7) 与 decimal(7,0)等价





Float 可以入四舍五入近似，fixed 是确切值



Fixed- and Floating-Point Data Types

Floating-point data type	size (bytes)	precision	maximum number of digits
FLOAT	4	*single*	23
DOUBLE	8	*double*	53

Other Useful Data Types

DATE 

used to represent a date in the format YYYY-MM-DD

1st of January 1000 - 31st of December 9999

---

e.g. 25th of July 2018: ‘2018-07-25’

Other Useful Data Types

DATE +  = DATETIME

next to the date, we could save the time:
YYYY-MM-DD HH:MM:SS[.fraction]

0 – 23:59:59.999999

---

e.g. 25th of July 2018 9:30 a.m.: ‘2018-07-25 9:30:00’

Other Useful Data Types

DATETIME

represents the date shown on the calendar and the time shown on the clock

vs.

TIMESTAMP

used for *a well-defined, exact point in time*

Other Useful Data Types

TIMESTAMP

used for *a well-defined, exact point in time*

1st of January 1970 UTC – 19th of January 2038, 03:14:07 UTC

- records the moment in time as the number of seconds passed after the 1st of January 1970 00:00:00 UTC

e.g. 25th of July 2018: 1,535,155,200

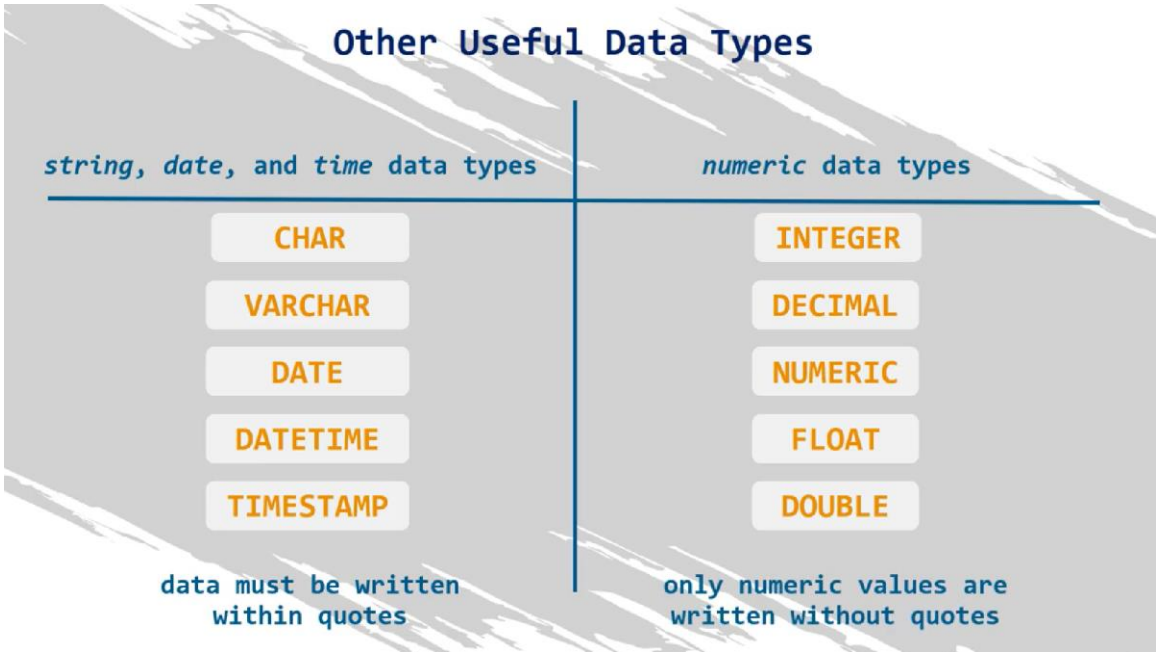
more than 48 years!

Other Useful Data Types

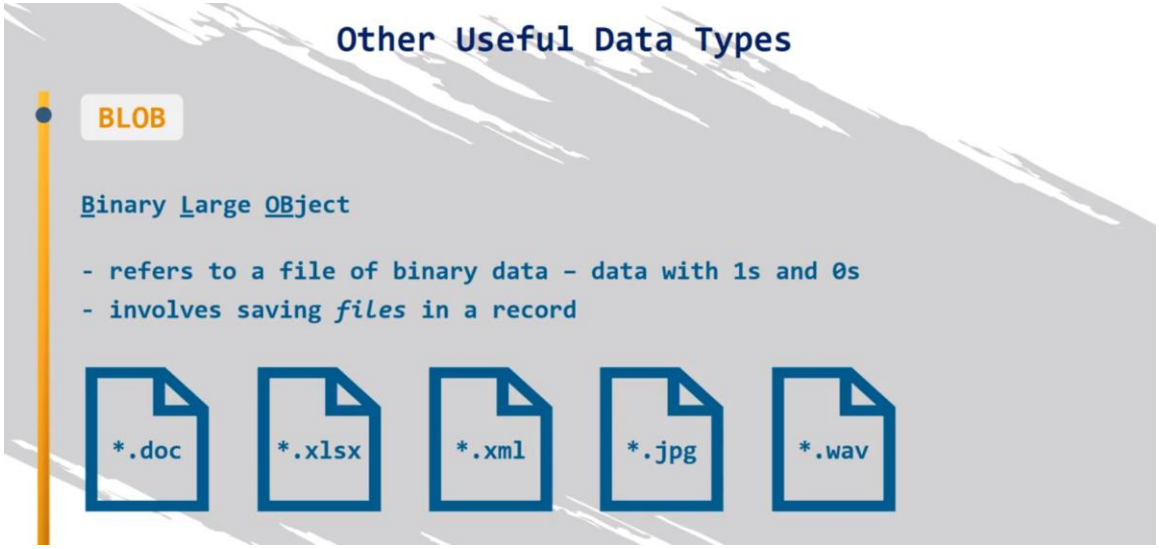
TIMESTAMP

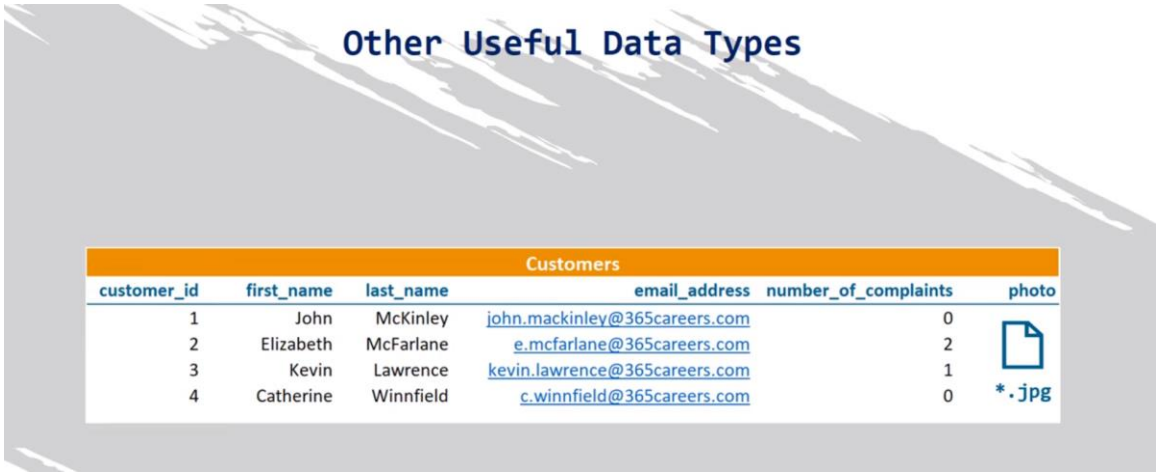
- representing a moment in time as a number allows you to easily obtain the difference between two TIMESTAMP values

e.g. end time:	'2018-07-25 10:30:00' UTC	TIMESTAMP
-		
start time:	'2018-07-25 09:00:00' UTC	TIMESTAMP
	5,400	TIMESTAMP


Other Useful Data Types
string, date, and time data types
numeric data types
CHAR
VARCHAR
DATE
DATETIME
TIMESTAMP
INTEGER
DECIMAL
NUMERIC
FLOAT
DOUBLE
data must be written within quotes
only numeric values are written without quotes

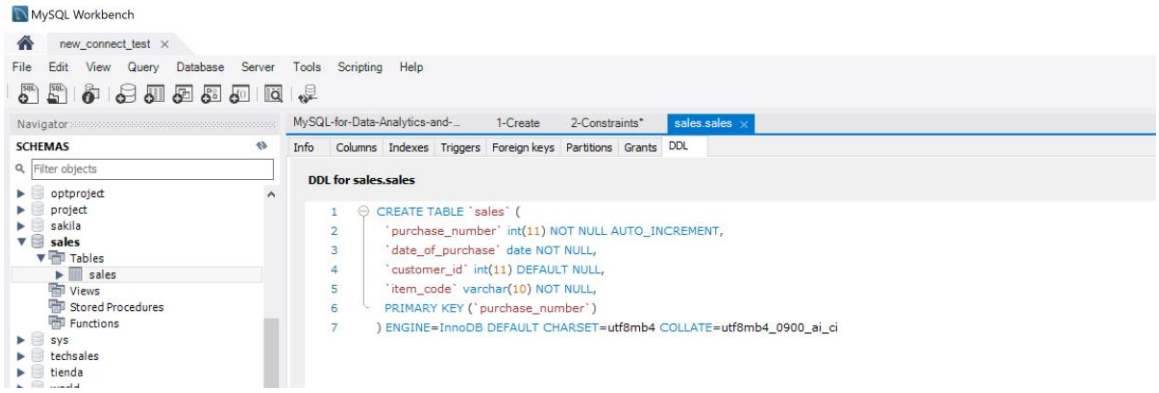
存文件/大文件

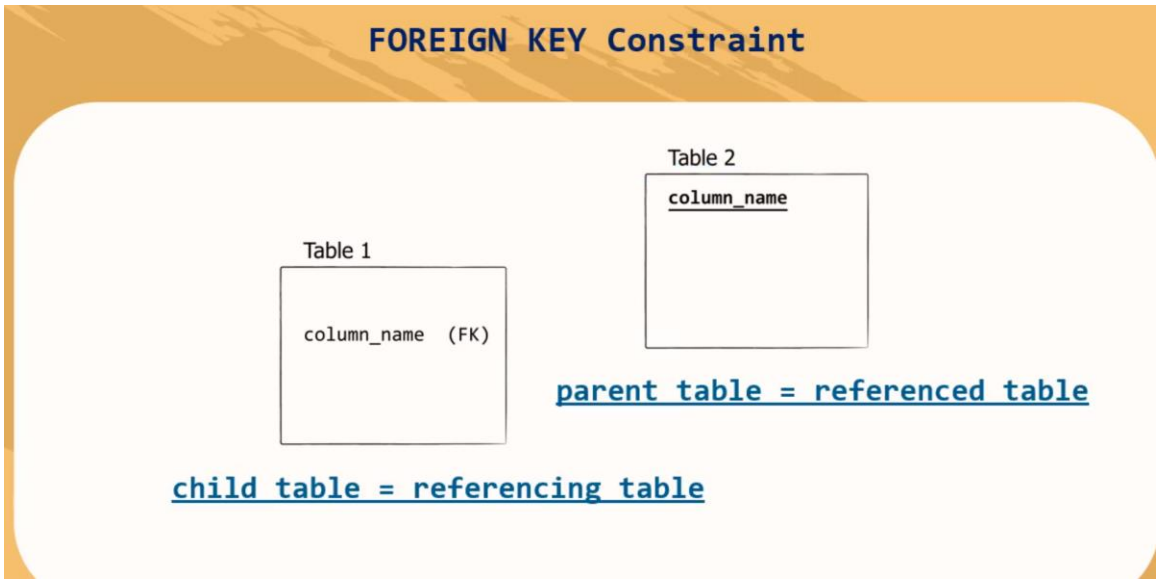

Other Useful Data Types
BLOB
Binary Large OBject
- refers to a file of binary data – data with 1s and 0s
- involves saving files in a record
*.doc
*.xlsx
*.xml
*.jpg
*.wav

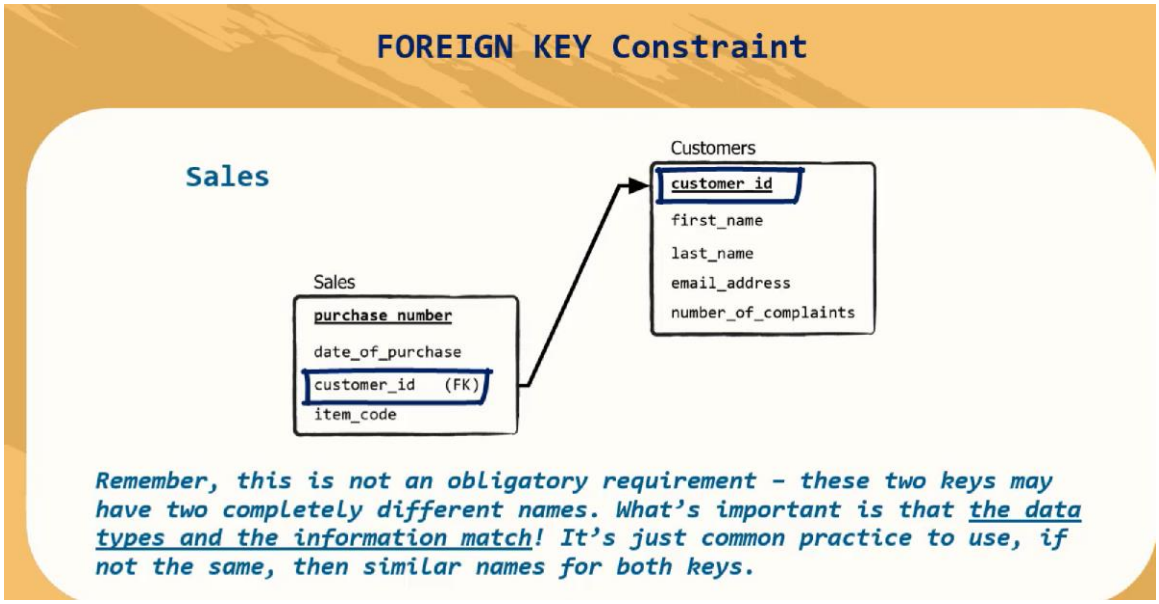


Section 6

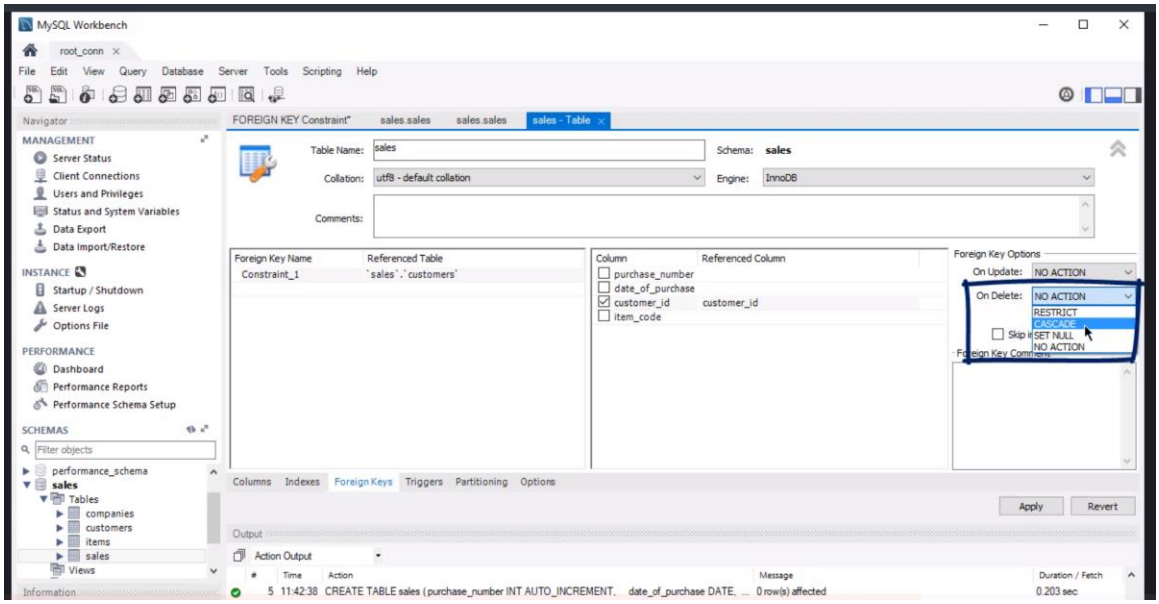
Can see ddl of the table

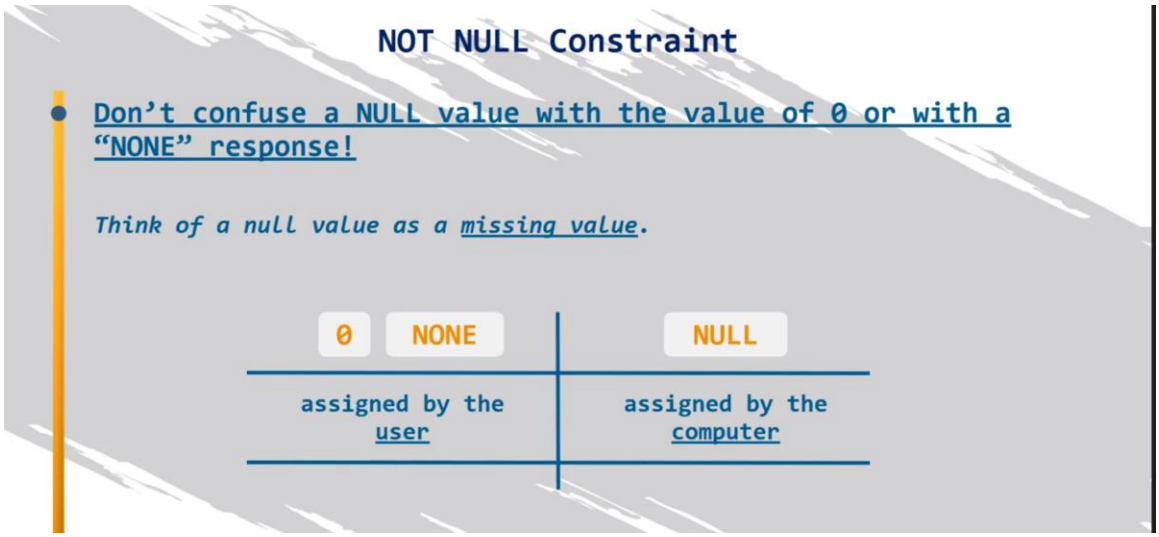




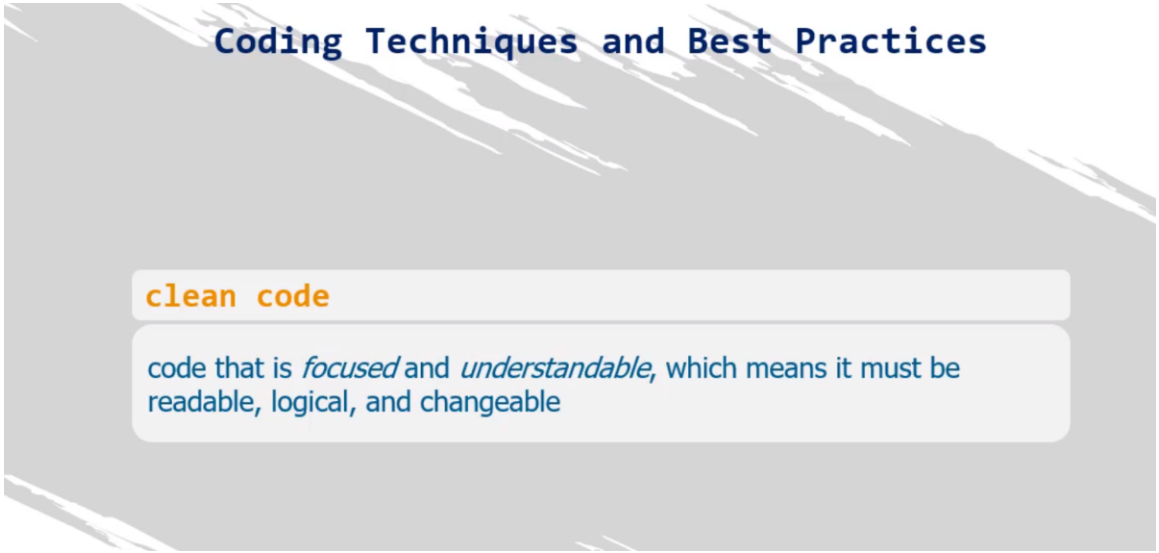


不想写代码可以之间右键 table-alter table 等




NOT NULL Constraint
Don’t confuse a NULL value with the value of 0 or with a “NONE” response!
Think of a null value as a missing value.
0
NONE
NULL
assigned by the user
assigned by the computer

Section 7


Coding Techniques and Best Practices
clean code
code that is focused and understandable, which means it must be readable, logical, and changeable

Coding Techniques and Best Practices

when assigning names to variables or SQL objects,

always chose shorter, meaningful names, conveying *specific information*

↓

pronounceable, where one word per concept has been picked

customer_purchase_unique_number ✗

Sales			
purchase_number	date_of_purchase	customer_id	item_code
1	9/3/2016	1	A_1
2	12/2/2016	2	C_1
3	4/15/2017	3	D_1
4	5/24/2017	1	B_2
5	5/25/2017	4	B_2
6	6/6/2017	2	B_1
7	6/10/2017	4	A_2
8	6/10/2017	3	C_1
9	7/20/2017	1	A_1
10	8/11/2017	2	B_1

Coding Techniques and Best Practices

SQL

```sql
CREATE TABLE sales
(
    purchase_number INT,
    date_of_purchase DATE,
    customer_id VARCHAR(255),
    item_code VARCHAR(255),
PRIMARY KEY (purcase_number)
);
```

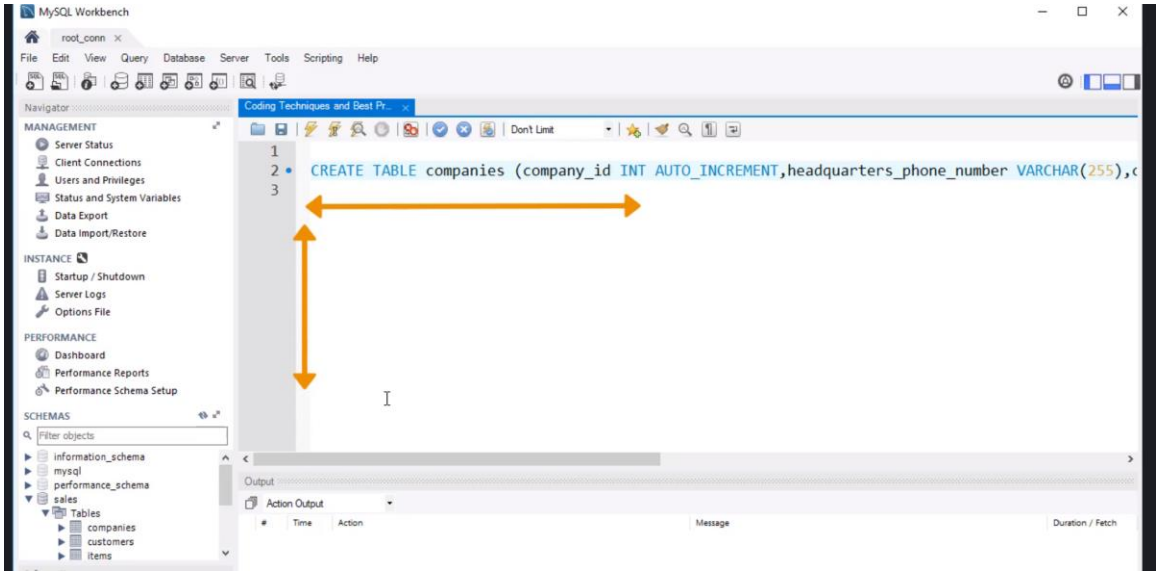
Coding Techniques and Best Practices

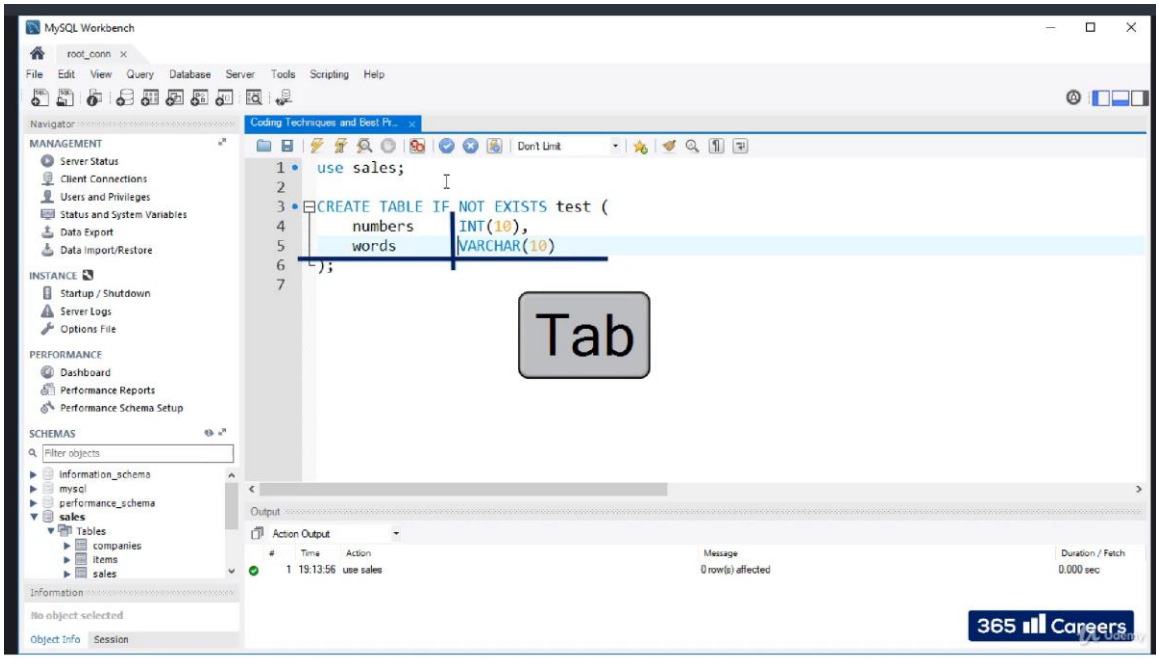
`purchase_number`

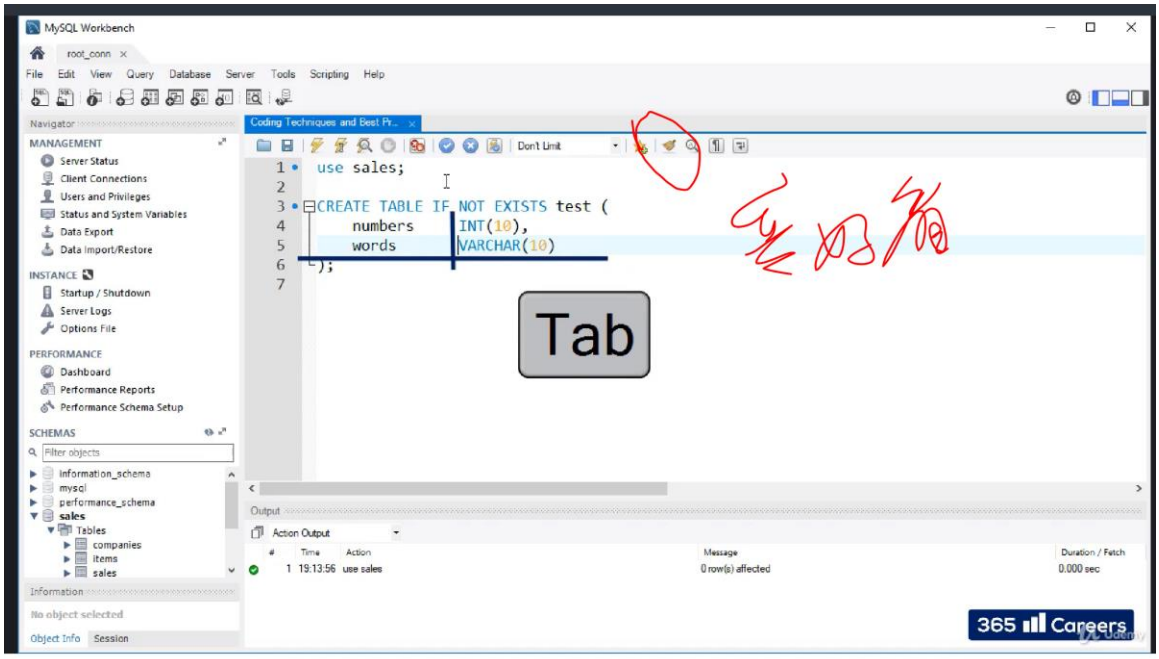
`PurchaseNumber`

`purchase number`

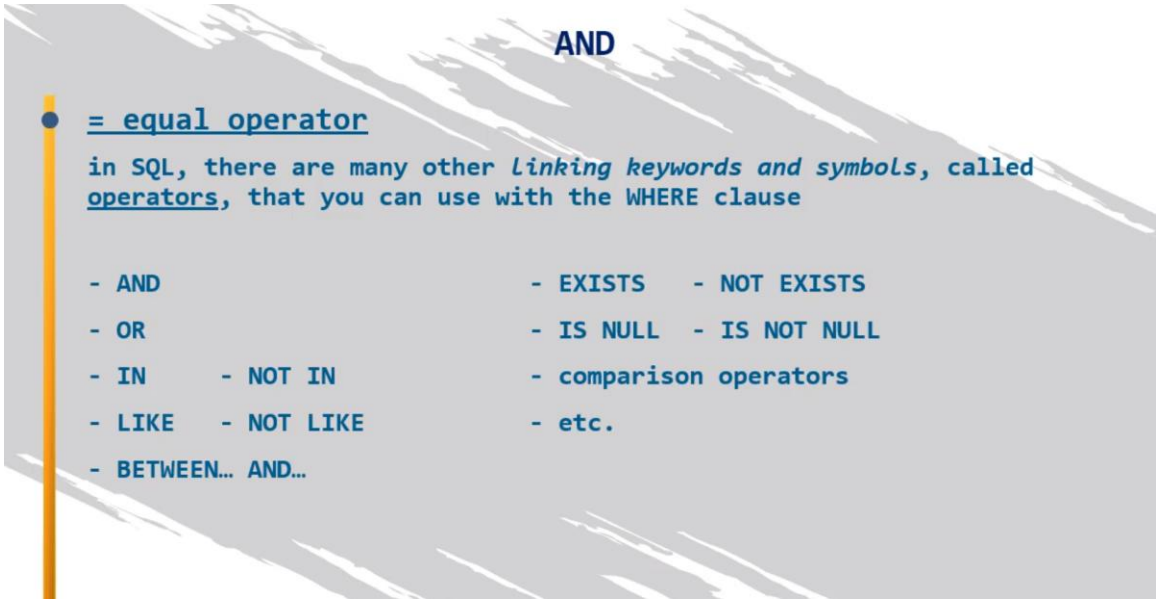
Sales			
purchase_number	date_of_purchase	customer_id	item_code
1	9/3/2016	1	A_1
2	12/2/2016	2	C_1
3	4/15/2017	3	D_1
4	5/24/2017	1	B_2
5	5/25/2017	4	B_2
6	6/6/2017	2	B_1
7	6/10/2017	4	A_2
8	6/10/2017	3	C_1
9	7/20/2017	1	A_1
10	8/11/2017	2	B_1




MySQL Workbench
root_conn
File Edit View Query Database Server Tools Scripting Help
Navigator
MANAGEMENT
Server Status
Client Connections
Users and Privileges
Status and System Variables
Data Export
Data Import/Restore
INSTANCE
Startup / Shutdown
Server Logs
Options File
PERFORMANCE
Dashboard
Performance Reports
Performance Schema Setup
SCHEMAS
Filter objects
information_schema
mysql
performance_schema
sales
Tables
companies
items
sales
Information
No object selected
Object Info
Session
Coding Techniques and Best Pr...
Don't Limit
use sales;
CREATE TABLE IF NOT EXISTS test (
numbers INT(10),
words VARCHAR(10)
);
Tab
Output
Action Output
Time Action Message Duration / Fetch
1 19:13:56 use sales 0 row(s) affected 0.000 sec
365 Careers


MySQL Workbench
root_conn
File Edit View Query Database Server Tools Scripting Help
Navigator
MANAGEMENT
Server Status
Client Connections
Users and Privileges
Status and System Variables
Data Export
Data Import/Restore
INSTANCE
Startup / Shutdown
Server Logs
Options File
PERFORMANCE
Dashboard
Performance Reports
Performance Schema Setup
SCHEMAS
Filter objects
information_schema
mysql
performance_schema
sales
Tables
companies
items
sales
Information
No object selected
Object Info
Session
Coding Techniques and Best Pr...
Don't Limit
use sales;
CREATE TABLE IF NOT EXISTS test (
numbers INT(10),
words VARCHAR(10)
);
Tab
Output
Action Output
Time Action Message Duration / Fetch
1 19:13:56 use sales 0 row(s) affected 0.000 sec
365 Careers

Section 9



Start from 96