

# The Ultimate MySQL Bootcamp: Go from Expert to Pro

Link: [Online Sql Tool](#)

slide-1: [presentation](#)

slide-1 contains: create db, insert, drop db, primary key, auto-increment, not null, default, not null default.

slide-2: [presentation](#) also known as Refining Selections

slide-2 contains: string functions, distinct, order by , limit, like,

Correlated subqueries: [Correlated subqueries](#)

## Note about the SQL editor

from the next lecture will give you errors if you're using firefox or any other browser other than Chrome, Safari, or Opera. This is because the SQL Try-It Editor is using WebSQL which isn't compatible with all browsers.

We recommend using chrome for those exercises and for the remainder of the course, if possible.

## First five minutes of SQL

Presentation link: [Presentation](#)

### Where We're Going

```
SELECT
    username,
    photos.id,
    photos.image_url,
    COUNT(*) AS total
FROM photos
INNER JOIN likes
    ON likes.photo_id = photos.id
INNER JOIN users
    ON photos.user_id = users.id
GROUP BY photos.id
ORDER BY total DESC
LIMIT 1;
```

```
SELECT first_name,
       last_name,
       Count(rating)                      AS COUNT,
       Ifnull(Min(rating), 0)              AS MIN,
       Ifnull(Max(rating), 0)              AS MAX,
       Round(Ifnull(Avg(rating), 0), 2)    AS AVG,
       CASE
           WHEN Count(rating) >= 10 THEN 'POWER USER'
           WHEN Count(rating) > 0 THEN 'ACTIVE'
           ELSE 'INACTIVE'
       end                                AS STATUS
FROM   reviewers
       LEFT JOIN reviews
              ON reviewers.id = reviewsreviewer_id
GROUP  BY reviewers.id;
```

## Getting started overview and installation

### What is a db?



Figure 1: Phonebook

Another e.g of db is a phonebook.

Something archaic like this is Rolodex.



Figure 2:  
rolodex

### A Quick Example

Andrews, Archie	- (949)345-2222
Cooper, Betty	- (212)246-9846
Flanders, Ned	- (415)987-3451
Jones, Jughead	- (415)888-3777
Lodge, Veronica	- (714)332-0981
Snow, Jon	- (949)621-1908
Stark, Ned	- (310)119-6501

Find Ned Flanders' Phone Number

Find People With First Name "Ned"

Find All Phone Numbers With Area Code 415

Find All People Who Have a 3-letter First Name

Think of phonebook as a dictionary.

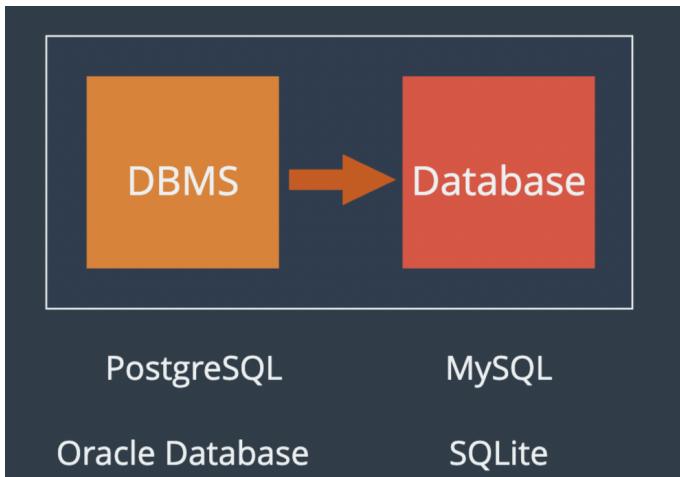
Considering it is ordered based on Last name.

It's a lot difficult to find all people with the First Name "Ned".

### DIAGRAM TIME!



"Database"



are database management system.

## What Is A Database?

A structured set of computerized data with an accessible interface

## SQL VS MYSQL

Sql is the language we use to talk to db.

Crud basically.

PostgreSQL	MySQL	Uses (SQL)structured query language.
Oracle Database	SQLite	

Point is sql is not unique to mysql, that's why they all have SQL in the name.

They shared the **same** language.

## MYSQL VS PostgreSQL

Are different dbms, but they both use Sql.

You'll be writing sql when you interact with this database.

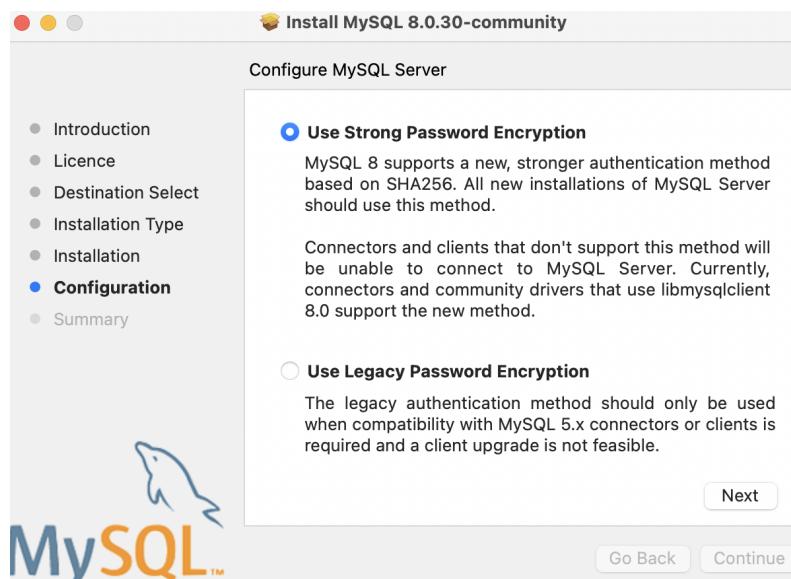
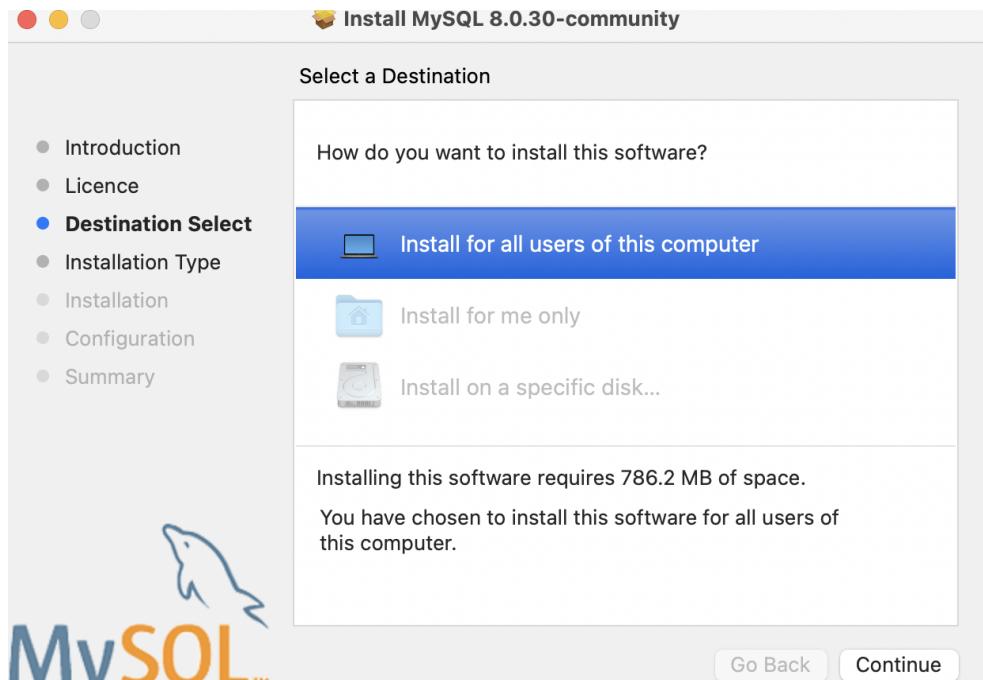
There is **slight** difference in Syntax, they are not identical.

There is a sql standard and all this dbms are tasked with implementing the standard.

### Two takeaways

1. It's pretty easy to switch.
2. what makes dbms unique, Is the feature they offer. Not the language.

Factors like secure, fast, download, permission works



```

kazimsyed@kazims-MacBook-Air ~
% ls /usr/local/mysql/bin
ibd2sdi          mysql_upgrade
innoschecksum    mysqladmin
libprotobuf-lite.3.19.4.dylib mysqlbinlog
libprotobuf.3.19.4.dylib   mysqlcheck
lz4_decompress   mysqld
my_print_defaults mysqld-debug
myisam_ftdump    mysqld_multi
myisamchk        mysqld_safe
myisamlog        mysqldump
myisampack       mysqldumpslow
mysql            mysqldump
mysql_config     mysqldumpimport
mysql_config_editor mysqlshow
mysql_migrate_keyring  mysqlslap
mysql_secure_installation perror
mysql_ssl_rsa_setup  zlib_decompress
mysql_tzinfo_to_sql

kazimsyed@kazims-MacBook-Air ~
% path[8]=!:1
path[8]="/usr/local/mysql/bin"

```

```

kazimsyed@kazims-MacBook-Air ~
% mysql bhavans9911
ERROR 1045 (28000): Access denied for user 'kazimsyed'@'localhost' (using password: NO)
If u
kazimsyed@kazims-MacBook-Air ~
% mysql -u root -p
Enter password:
Welcome to the MySQL monitor.  Commands end with ; or \g.
Your MySQL connection id is 10
Server version: 8.0.30 MySQL Community Server - GPL
forgot ur password, u can change it.

```

```

mysql> show databases;
+-----+
| Database |
+-----+
| information_schema |
| mysql |
| performance_schema |
| sys |
+-----+
4 rows in set (0.00 sec)

```

**mysql> use mysql**

```

mysql> describe user;
+-----+-----+-----+-----+-----+-----+
| Field      | Type       | Null | Key | Default | Extra |
+-----+-----+-----+-----+-----+-----+
| Host       | char(255)  | NO  | PRI |          |       |
| User       | char(32)   | NO  | PRI |          |       |
| Select_priv | enum('N','Y') | NO  |      |          |       |
| Insert_priv | enum('N','Y') | NO  |      |          |       |
| Update_priv | enum('N','Y') | NO  |      |          |       |
| Delete_priv | enum('N','Y') | NO  |      |          |       |
| Create_priv | enum('N','Y') | NO  |      |          |       |
| Drop_priv  | enum('N','Y') | NO  |      |          |       |
| Reload_priv | enum('N','Y') | NO  |      |          |       |
+-----+-----+-----+-----+-----+-----+

```

I need to checkout the mac installatio to learn

ALTER USER '[root@localhost](#)' identified by 'password'

## Creating databases and tables

### Creating databases

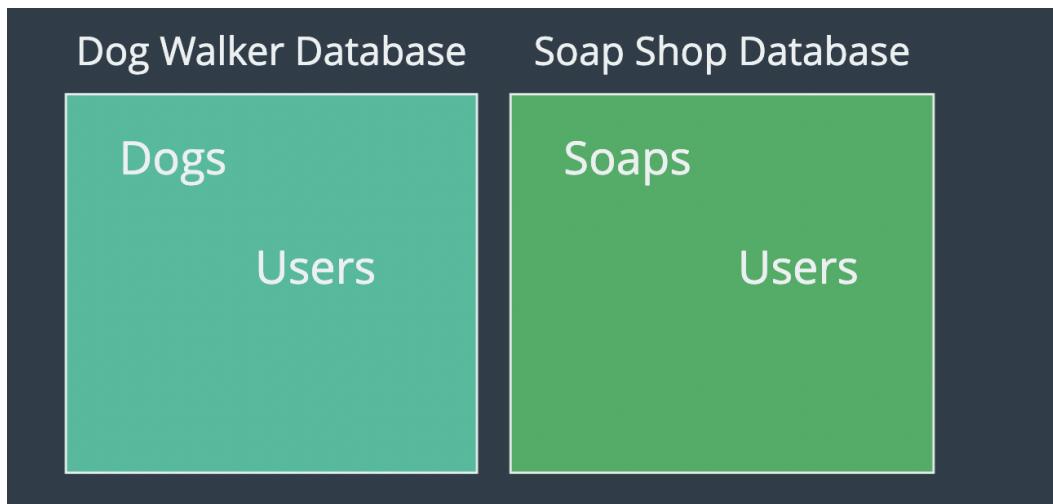
**How to work with databases and tables:**

<b>Create database [if not exists] &lt;db_name&gt;;</b>	Create database
<b>Drop database [if exists] &lt;db_name&gt;;</b>	Drop database
<b>Drop table [if exists] &lt;db_name&gt;.&lt;table_name&gt;</b>	Drop table while being outside the database.

Best practices:

1. Multiple db maybe running on a single Mysql server, you usually want to select the database before you begin working with it.
2. For large database project, DDL statements would be handled by a database specialist or DBA.
3. For small project, SQL programmer have to serve as DBA.
4. And for small projects, SQL programmer have to use DDL statements to create smaller tables that are needed for testing.





*Figure 3: show databases;*

Our dog walking app has users. But so does our Soap shop.

Our dog walking app has payments. But so does our Soap shop app.

They have to be **seperated**.

If we have one db and everything was using on this server, there would be a lot of crossovers and issues.

```
CREATE DATABASE <name>;
```

```
CREATE DATABASE soap_store;
```

```
CREATE DATABASE DogApp;
```

```
CREATE DATABASE My App;
```

*Figure 4: space isn't allowed, use underscore instead or Pascal casing.*

```
DROP DATABASE <name>;
```

```
USE <database name>;
```

```
SELECT database();
```

A database is just a  
bunch of **tables**

*In a relational database, at least*

Giant ER diagram: [ER diagram](#)

### Datatypes

#### Numeric Types

- INT
- SMALLINT
- TINYINT
- MEDIUMINT
- BIGINT
- DECIMAL
- NUMERIC
- FLOAT
- DOUBLE
- BIT

#### String Types

- CHAR
- VARCHAR
- BINARY
- VARBINARY
- BLOB
- TINYBLOB
- MEDIUMBLOB
- LONGBLOB
- TEXT
- TINYTEXT
- MEDIUMTEXT
- LONGTEXT
- ENUM

#### Date Types

- DATE
- DATETIME
- TIMESTAMP
- TIME
- YEAR

**INT**  
A Whole Number  
with a max value of 4294967295

**varchar**  
A Variable-Length String  
Between 1 and 255 characters

# Draw a Tweets Table

At a minimum the columns must include:

- A username (max 15 chars)
- The tweet content (max 140 chars)
- Number of favorites

Make sure to specify correct MySQL datatypes!

Username (max 15 chars)	Content (max 140 chars)	Favorites
		
varchar(15)	varchar(140)	int
username	content	favorites
'coolguy'	'my first tweet!'	1
guitar_queen	'I love music :) '	10
'lonely_heart'	'still looking 4 love'	0

```
CREATE TABLE tablename
(
    column_name data_type,
    column_name data_type
);
```

```
CREATE TABLE cats
(
    name VARCHAR(100),
    age   INT
);
```

Commands	Description
Show tables;	See tables inside the database.
Select columns from table_nm; or desc table_nm	Learn about table
Drop table;	Delete table
Insert into table_nm (col_1, col_2) values (<>,<>);	Insert values into a table
Insert into table_nm (col_1, col_2) values (<>,<>),(<>,<>), ...	Multiple inserts
Show warnings;	

Syntax of the Create table statement:

```
CREATE TABLE [db_name].table_name
(
    column_1 datatype [column attributes]
    [, column_2 datatype [column attributes] ]
    [, table_level_constraints]
)
```

common column attributes

Attribute	Description
Not Null	Indicates that col doesn't accept null values. If ommited the column can accept null values.
Unique	Specifies that each value in the col must be unique.
Default default_value	Specifies a default value for the column.
auto_increment	Identifies a column whose value is automatically incremented by Mysql when row is added. Auto-increment col can be (float or int)

To indicate that each row in a column must contain a unique value, you can code the unique attribute.

Since two null values aren't considered the same, a unique column can contain null values.

Best practice: to use **not null** and **unique** attribute.

--

Autoincrement: mysql starts numbering with 1, you can start with a value other than 1 by coding the auto\_increment attribute like:

auto\_increment=3

Auto-increment value can be specified for only one column in the table and that column can be defined as either the primary key or unique key.

## SHOW TABLES;

```
SHOW COLUMNS FROM <tablename>;
```

Or...

```
DESC <tablename>;
```

## Deleting Tables

```
DROP TABLE <tablename>;
```

## INSERT

```
INSERT INTO cats(name, age)
VALUES ('Jetson', 7);
```

## MULTIPLE INSERT

```
INSERT INTO cats(name, age)
VALUES ('Charlie', 10),
       ('Sadie', 3),
       ('Lazy Bear', 1);
```

## Warnings:

1 Warning??! 😱

Query OK, 1 row affected, 1 warning (0.01 sec)

Let's Take A Look

SHOW WARNINGS;

Not null, default, not null

# What's Up With This?

Field	Type	Null	Key	Default	Extra
name	varchar(5)	YES		NULL	
age	int(11)	YES		NULL	

-The value isn't known.

-Null does not mean zero.

Right now, we could do this...

```
INSERT INTO cats(name)
VALUES ('Alabama');
```

Or This! *gasp*

```
INSERT INTO cats()
VALUES ();
```

```
CREATE TABLE cats2
(
    name VARCHAR(100) NOT NULL,
    age   INT NOT NULL
);
```

Notice The Difference!

Field	Type	Null	Key	Default	Extra
name	varchar(100)	NO		NULL	
age	int(11)	NO		NULL	

Figure 5: Use of not null keyword

# To Set Default Values

```
CREATE TABLE cats3
(
    name VARCHAR(100) DEFAULT 'unnamed',
    age   INT DEFAULT 99
);
```

## What's Up With This?

Field	Type	Null	Key	Default	Extra
name	varchar(5)	YES		NULL	
age	int(11)	YES		NULL	

Figure 6: Use of default keyword. Here we can see that we can manually pass NULL values.

-- so u may feel:

## Isn't This Redundant?

```
CREATE TABLE cats4
(
    name VARCHAR(100) NOT NULL DEFAULT 'unnamed',
    age   INT NOT NULL DEFAULT 99
);
```

## No!

We can still manually set things to NULL if we don't specify NOT NULL

```
INSERT INTO cats3(name, age)
VALUES(NULL, 3);
```

Primary key:

# What's Up With This?

Field	Type	Null	Key	Default	Extra
name	varchar(5)	YES		NULL	
age	int(11)	YES		NULL	

**Primary Key**  
A Unique Identifier

Right now, this could happen!

Name	Breed	Age
Monty	Tabby	10

How Do We Make Each Unique?

Name	Breed	Age	CatID
Monty	Tabby	10	1
Monty	Tabby	10	2
Monty	Tabby	10	3
Monty	Tabby	10	4

Code:

```
CREATE TABLE unique_cats2 (cat_id INT NOT NULL AUTO_INCREMENT  
                      ,name VARCHAR(100)  
                      ,age INT  
                      ,PRIMARY KEY (cat_id));
```

Question:

Table 1: sol-1) when you define a constraint in a column definition is known as column level constraint. You can also define a constraint at the table level using the constraint keyword. when you code a table level constraint, you can provide a name for the constraint. Reasons for using a name for a table-level constraint: reasons to name a constraint

Define an Employees table,  
with the following fields:

- `id` - number(automatically increments), mandatory, primary key
- `last_name` - text, mandatory
- `first_name` - text, mandatory
- `middle_name` - text, not mandatory
- `age` - number mandatory
- `current_status` - text, mandatory, defaults to 'employed'

## THE SOLUTION

```
CREATE TABLE employees (
    id INT NOT NULL AUTO_INCREMENT,
    last_name VARCHAR(255) NOT NULL,
    first_name VARCHAR(255) NOT NULL,
    middle_name VARCHAR(255),
    age INTEGER NOT NULL,
    current_status VARCHAR(100) NOT NULL DEFAULT 'employed',
    PRIMARY KEY (id)
);
```

COPY

## THE SOLUTION (with a slight difference)

```
CREATE TABLE employees (
    id INT NOT NULL AUTO_INCREMENT PRIMARY KEY,
    last_name VARCHAR(255) NOT NULL,
    first_name VARCHAR(255) NOT NULL,
    middle_name VARCHAR(255),
    age INTEGER NOT NULL,
    current_status VARCHAR(100) NOT NULL DEFAULT 'employed'
);
```

The syntax of a column-level primary key constraint:

**column\_name data\_type primary key column\_attributes**

The syntax of a table-level primary key constraint

[ constraint [constraint\_name]]  
**primary key** (**col\_1[,col\_2]...**)

A table with column level constraints

```
Create table vendors
(
    vendor_id      INT PRIMARY KEY AUTO_INCREMENT,
    Vendor_name    VARCHAR(50) NOT NULL UNIQUE
)
```

A table with table-level constraints

```
Create table vendors
(
    vendor_id  int auto-increment
    vendor_name varchar(50) not null

    constraint vendors_pk primary key ( vendor_id),
    constraint vendor_name_uq UNIQUE (vendor_name)
)
```

A table with two column primary key constraints

```
create table invoice_line_items
(
    invoice_id int not null,
    invoice_sequence int not null,
    line_item_description varchar(100) not null,

    constraint line_items_pk primary key( invoice_id, invoice_sequence)
)
```

Here's some pretty basic reasons.

- (1) If a query (insert, update, delete) violates a constraint, SQL will generate an error message that will contain the constraint name. If the constraint name is clear and descriptive, the error message will be easier to understand; if the constraint name is a random guid-based name, it's a lot less clear. Particulary for end-users, who will (ok, might) phone you up and ask what "FK\_\_B\_\_B\_COL1\_\_75435199" means.
- (2) If a constraint needs to be modified in the future (yes, it happens), it's very hard to do if you don't know what it's named. (ALTER TABLE MyTable drop CONSTRAINT um...) And if you create more than one instance of the database "from scratch" and use system-generated default names, no two names will ever match.
- (3) If the person who gets to support your code (aka a DBA) has to waste a lot of pointless time dealing with case (1) or case (2) at 3am on Sunday, they're quite probably in a position to identify where the code came from and be able to react accordingly.

*Figure 7: reasons to name a constraint*

## How to code a foreign key constraint:

The syntax of a column-level foreign key constraint

[column attributes] references table-name(column\_name)

[ON DELETE (CASCADE | SET NULL) ]

The syntax of a table-level foreign key constraint

[constraint constraint\_name]

FOREIGN KEY ( col\_name [,col\_name] ... )

Referecnes table\_name (col\_name [,col\_name] ...)

[ ON DELETE (cascade | set null ) ]

It also let you to set foreign key for multiple columns.

e.g

A table with column-level foreign key constraint

Create table invoices

(

invoice\_id int primary key  
vendor\_id int references vendors ( vendor\_id ),  
invoice number varchar(50) not null unique

)

A table with table-level foreign key constraint

create table invoices

(

invoice\_id int primary key,  
vendor\_id int not null,  
invoice number varchar not null unique,

constraint invoices\_fk\_vendors

foreign\_key (vendor\_id) references vendors (vendor\_id)

)

A constraint that uses the ON DELETE clause

constraint invoices\_fk\_vendors

foreign\_key (vendor\_id) references vendors (vendor\_id)

ON DELETE CASCADE

Foreign key enforces referential integrity.

In some cases though, you may want to automatically delete the related rows in the invoices table when a row in the vendors table is deleted.

To do that, you can code the on delete clause on the foreign key as in the last example.

So, when you delete a row from the primary key table , the delete is cascaded to the related rows, in the foreign key table.

For e.g you delete a row from the vendors table, all related rows in the invoices table will also be deleted.

**CAUTION:**

Cascading delete makes it easier to delete data that you didn't intend to delete, use it with caution.

You can also use not SET NULL On the ON DELETE clause.

So when you delete a row from a primary key table, the values in the foreign key column of foreign key table are set to null.

Since this creates rows in the foreign key table that aren't related to the primary key table, you rarely want to use this option.

## How to alter the columns of a table

After you create tables, you may need to change the columns of table.

Not table cuz for table you can use create or replace table while being cautious of the data in the table and relationships.

You may need **AMD**:

add, modify, delete.

Syntax for modifying the columns of a table.

*Table 2: MySQL won't allow you to change a column if that change would cause data to be lost.*

```
ALTER TABLE [db_name.]table_name
(
  ADD  column_name data_type [column_attributes] |
  DROP COLUMN column_name |
  MODIFY column_name datatype [column_attribute]
)
```

Warning: You should never alter or other database object in a production database without first consulting the DBA.

## How to alter constraints of a table.

After creating a table, you may need to add or drop a constraint.

The syntax for modifying the constraints of a table.

```
Alter table [db_name.]table_name
(
add primary key constraint_definition |
add [constraint constraint_name] foreign key constraint definition |
drop primary key
drop foreign key constraint name.
)
```

Mysql may not allow you to drop the primary key for a table. That's true if the primary key is an auto-increment column or it's referred to by a foreign key.

A table can have multiple foreign key, you must know the name of the key you want to drop.

If you don't know its name, you can use MySQL Workbench to look It up.

## How to rename, truncate, and drop a table.

It is useful if you want to change the name of a table without modifying its column definitions or the data that's stored in the table.

Rename table vendors to vendor	Renaming a table
Truncate table vendor OR Delete * from vendor	Truncate to delete all of the data from a table without deleting the column definitions of the table.

Note:

1-You can't truncate or drop a table if a foreign key constraints in another table refers to that table.

2-When you drop a table, all of its data, constraints, and indexes are deleted.

Warnings:

You shouldn't use these statements on a production database without first consulting the DBA.

## **Strategies for backing up and restoring a database**

PENDING

# MySQL DATA TYPES

DATE TYPE	SPEC	DATA TYPE	SPEC
CHAR	String (0 - 255)	INT	Integer (-2147483648 to 2147483647)
VARCHAR	String (0 - 255)	BIGINT	Integer (-9223372036854775808 to 9223372036854775807)
TINYTEXT	String (0 - 255)	FLOAT	Decimal (precise to 23 digits)
TEXT	String (0 - 65535)	DOUBLE	Decimal (24 to 53 digits)
BLOB	String (0 - 65535)	DECIMAL	"DOUBLE" stored as string
MEDIUMTEXT	String (0 - 16777215)	DATE	YYYY-MM-DD
MEDIUMBLOB	String (0 - 16777215)	DATETIME	YYYY-MM-DD HH:MM:SS
LONGTEXT	String (0 - 4294967295)	TIMESTAMP	YYYYMMDDHHMMSS
LONGBLOB	String (0 - 4294967295)	TIME	HH:MM:SS
TINYINT	Integer (-128 to 127)	ENUM	One of preset options
SMALLINT	Integer (-32768 to 32767)	SET	Selection of preset options
MEDIUMINT	Integer (-8388608 to 8388607)	BOOLEAN	TINYINT(1)

LOB ( LARGE OBJECT) data types that are useful for storing images, sound, video, and large amounts of character data.

We can divide numbers into integers and real numbers. Integers are numbers that don't have a decimal point, and real numbers are numbers that have a decimal point.

*Table 3: this tables shows how unsigned and zerofill attributes work with negative and positive number.*

Data type	Original value	Value stored	Value displayed
INT	99	99	99
INT	-99	-99	-99
INT UNSIGNED	99	99	99
INT UNSIGNED	-99	error	error
INT ZEROFILL	99	99	0000000099
INT(4) ZEROFILL	99	99	0099

The BOOL AND BOOLEAN types are synonyms for TINYINT(1).

## The fixed-point and floating point types

### The fixed point type:

DECIMAL(M,D) : M specifies the maximum of total digit, D specifies the number of digits to the right of the decimal. M>D. By default D=0.

### The floating point type:

Float(4 bytes), double(8 bytes)

--

### Description:

real number unlike integers can store negative values in the column.

If the zero fill attribute for a real number is set, the number is displayed with zeroes padded from the left, and the unsigned attribute is automatically set.

DEC, NUMERIC, FIXED are all synonyms for the DECIMAL TYPE

REAL is a synonym for Double type

## The date and time types

DATE (YYYY-MM-DD), TIME( HH:MM:SS) , DATETIME(YYYY-MM-DD HH:MM:SS) , TIMESTAMP, YEAR[ (2I4) ]

The interesting one that need explanation is

YEAR(2I4) i.e Year in 2-digit or 4-digit format.

How MySQL interprets literal date/time values

Literal value	Value stored in Date column
'2000-05-27'	2000-05-27
'2000-5-27'	2000-05-27
'00-5-27'	2000-05-27
20000527	2000-05-27
2000.05.27	2000-05-27
00/5/27	Error
'2000-27-5'	Wrong format YYYY-MM-DD

Literal value	Value stored in time column
'7:32'	07:32:00
'19:32:11'	19:32:11
193211	19:32:11

Literal value	Value stored in datetime in timestamp column
'2000-05-27 12:00:00'	2000-05-27 12:00:00
'2000-05-27'	'2000-05-27 00:00:00'

## The ENUM and SET TYPES

The main difference between the ENUM and SET types is that an ENUM column can store exactly one value, but a SET column can store, zero, one or upto 64 different values.

You can use ENUM type to store values that are mutually exclusive, such as Yes, No or Maybe.

IOW, you can use the ENUM type to represent a choice of one value, but not two.

For e.g, delivery or pickup, cash, credit, or debit, small, medium, or large; paper or plastic; soup or salad.

Although I assume you might want both soup and salad. For that, you could use a set column.

You can use a SET column when you want to choose more than one value. For e.g, the toppings on a pizza, the software on a computer, or the features of a car could be SET values.

To store a value for an ENUM column, you code a single string.

When you add a row to a table that contains an ENUM column, MySQL assigns a default value to that column, If you don't explicitly specify a value. If the column allows null values, MySQL assigns a null value to the column.

If the column doesn't allow null values, MySQL assigns the first value as the default value, then, you'll want to code that value as the first valu in the set.

To store values in a SET column, you code a single string with one or more values seperated by commas. Then, MySQL stores each acceptable value and ignores any other values. Since commas are used to seperate values, you cant use commas within a value when you define the SET column.

When storing multiple values in a SET column, the order of the values doesn't matter. That's because MySQL stores the values in the same order as in column definition.

It doesn't matter if you repeat a value because MySQL doesn't store duplicate values.

Table 4: How values are stored in SET columns

Value	Stored in column SET('Pepperoni','Mushrooms','Olives')
'Pepperoni'	'Pepperoni'
'Mushrooms'	'Pepperoni'
'Pepperoni,'Corn'	'Pepperoni'
'Olives,Pepperoni'	'Pepperonoi,Olives'

### The large object types

BLOB ( Binary Large object) types store strings of binary data. This data type is often used to store any type of binary data.

The TEXT types work similarly to the BLOB types, they store strings of character. As a result, in other database systems, they are sometimes referred to as CLOB.

TEXT upto 65kb in length.

## The basic syntax of the SELECT statement

Clause	Description
SELECT	Describes the columns in the result set.
From	Name the base table from which the query retrieves the data.
Where	Specifies the condition that must be met for a row to be included in the result set
ORDER BY	Specifies how to sort the rows in the result set
LIMIT	Specifies the number of rows in return

Select invoice\_number, invoice\_date, invoice\_total from invoices order by invoice\_total DESC

Select invoice\_id, invoice\_total, credit\_total + payment\_total as total\_credits from invoices where invoice\_id=17

Select invoice\_number, invoice\_date, invoice\_total from invoices where invoice\_date Between '2011-6-01' and '2001-06-30'

Select invoice\_number, invoice\_date, invoice\_total from invoices where invoice\_total > 5000

Source	Option	Syntax
Base table value	All columns	*
	Column name	column_name
Calculated value	Result of a calculation	Arithmetic expression e.g Select invoice_id, invoice_total, <b>credit_total + payment_total as total_credits</b> from invoices where invoice_id=17
	Result of a function	Functions e.g select concat(fname,' ,lname) as full_name;

## How to insert default values and null values

If a column allows null values, you can use the INSERT statement to insert a null value into that column.

To insert a null value into a column, you can use the NULL keyword.

If a column is defined with a default value, you can use the insert statement to insert that value.

Finally, if a column is defined as an auto increment column, you can have MySQL generate a value for that column.

To insert a default value or to have MySQL generate a value for an auto\_increment column, you can use the default keyword.

Insert into color_sample values (DEFAULT, DEFAULT, NULL)
--

## How to update existing rows

The syntax of the update statement

```
UPDATE table_nm  
set column_name=expression [, column_name=expression]...  
[ where search_condition ]
```

By default, MySQL Workbench runs in safe update mode. This prevents you from updating rows if the Where clause is omitted or doesn't refer to a primary key or foreign key column.

To get around the restrictions of safe update mode, you can turn this mode off.

To do that, select the Edit → Preferences command, select the SQL editor tab, change the “safe update” option, and restart the MySQL Workbench.

Warning:

If you turn off safe update mode and omit the where clause, all rows in the table will be updated.

## How to delete existing rows

To delete one or more rows from a table, you can use the DELETE statement. If necessary, you can use subqueries in a delete statement to help identify the rows to be deleted.

### How to delete rows

By default, MySQL Workbench runs in safe update mode. This prevents you from updating rows if the Where clause is omitted or doesn't refer to a primary key or foreign key column.

To get around the restrictions of safe update mode, you can turn this mode off.

To do that, select the Edit → Preferences command, select the SQL editor tab, change the “safe update” option, and restart the MySQL Workbench.

### Warning:

If you turn off safe update mode and omit the where clause, all rows in the table will be updated.

If you want to make sure that you've selected the correct rows before you delete them, you can code a select statement that retrieves those rows.

Then, once you're sure the select statement is retrieving the correct rows, you can convert the select statement to a delete statement.

## How to use a subquery in a delete statement

```
Delete from invoice_line_items  
where invoice_id in  
(  
select invoice_id  
from invoices  
where vendor_id=115 )
```

## An introduction to subqueries

A subquery is a select statement that's ended within another sql statement.

### Where to code subqueries?

a subquery can be coded, introduced in the WHERE, HAVING, FROM, SELECT clause of a select statement.

-When a subquery returns a single value, you can use it anywhere you would normally use a single value. However, a subquery can also return a list of values( a result set that has one column) using the IN operator.

-A subquery can return a table of values( a result set that has multiple columns). In that case, you can use the subquery in the **FROM** clause.

Four ways to introduce a subquery in a select statement.

-In a WHERE clause as a search condition

-In a HAVING clause as a search condition

-In a FROM clause as a table specification

- In the select clause as a column specification.

-subqueries can be nested within other subqueries.

**-A subquery can't include an ORDER BY clause.**

You need to treat the subquery in the **FROM** clause, as a set of rows in some unspecified and undefined order, and put the **ORDER BY** on the top-level **SELECT**.

Workaround: I have discussed this in Onenote book and its using limit. But execution time increase.

## When to use subqueries

Subqueries provides a way of processing that can't be done any other way.

However most subqueries can be restated as joins and most joins can be restated as subqueries.

A query that uses an inner join

```
select invoice_number, invoice_date, invoice_total  
from invoices join vendors  
on invoices.vendor_id= vendors.vendor_id  
order by invoice_date
```

Same query reinstated with a subquery

```
Select invoice_number, invoice_date, invoice_total  
from invoices where vendor_id in (select vendor_id from vendors)  
order by invoice_date
```

So if you have a choice, which technique should you use?

In general, we recommend that you use the technique that results in the most readable code.

For e.g join tends to be more intuitive than a subquery when it uses an existing relationship between two tables.

On the other hand, subquery tends to be more intuitive when it uses an ad hoc relationship.

Another good point

you should realize that when you use a subquery in a where clause, its results cant be included in the final result set.

For instance, the second e.g above cant include the vendors name from the vendors table. That's because the vendors table isn't named in the FROM clause of the main query(saying select col\_nm **from players**). So if you need to include information from both tables in the result set, you need to use a join.

Advantages of joins:

- Select clause of a join can include columns from both tables.
- A join tends to be more intuitive when it uses an existing relationship between the two tables, such as primary key to foreign key relationship.

Advantages of subqueries:

- You can use a subquery to pass an aggregate value to the main query.
- A subquery tends to be more intuitive when it uses an adhoc relationship between the two tables.
- long, complex queries can be sometimes be easier to code using subqueries.

## How to code subqueries in the WHERE clause

### How to use the IN OPERATOR

In operator is used to test whether an expression is contained in a list of values.

One way to provide the list of values is to use a subquery.

*Table 5: The syntax of a WHERE clause that uses an IN phrase*

where text_expression [NOT] in (subquery)
---

e.g

*Table 6 Subquery returns a list of ID's for each vendor that's in the invoices table. Then the main query returns some data about the vendors whose ID's aren't in the list*

Get vendors without invoices
------------------------------

```
Select vendor_id, vendor_name, vendor_state  
from vendors  
where vendor_id not in  
(  
select distinct vendor_id  
from invoices)  
order by vendor_id
```

Description:

- when you use the **in** operator, the subquery must return a single columns of values.
- A query that uses the NOT IN operator with a subquery can typically be restated as an outer join.

## How to use the comparison operators

Table 7: Syntax of WHERE clause that use a comparision operator.

WHERE expression comparison_operator [some any all] (subquery)
--

Description:

If you code a search condition without the ANY, SOME, and ALL Keywords, the subquery must return a single value. --idk

If you include the ANY, SOME, or ALL keyword, the subquery can return a list of values.

\*\* I think both the statements above are looking to something basic from the opposite directions.

For e.g if you have a single value as result of a subquery, you won't use any, some, all.

And if you have multiple values as a result of a subquery, you may use all, some, any

### How to use the ALL keyword:

To use the ALL keyword to modify the comparison operator so the condition must be true for all the values returned by a subquery.

Condition	Equivalent expression	Desc
X > all(1,2)	X > 2	Evaluates to true if x is greater than the maximum value returned by the subquery
X < ALL(1,2)	X < 1	Evaluates to true if x is less than the minimum value returned by the subquery.
X = ALL(1,2)	X=1 and X=2	
X <> ALL(1,2)	X not in (1,2)	

Q.Get invoice larger than the largest invoice for vendor 34?

Also, get other info as well such as the name of the vendors.

So we will use a join as well.

To get the largest invoice of vendor 34.

select Max(invoice\_total) from invoices where vendor\_id=34.

But since for example sake we need a list of values.

select invoice\_total from invoices where vendor\_id=34.

```
Select vendor_name, invoice_number, invoice_total  
from invoices I JOIN vendors on I.vendor_id = v.vendor_id  
where invoice_total > ALL(select invoice_total from invoices where  
vendor_id=34)  
order by vendor_name;
```

>ALL(select invoice\_total from invoices where vendor\_id=34)

is basically equivalent to

> Max(invoice\_total)

Descriptions:

if no rows are returned by the subquery, a comparision that uses the ALL keyword is always true.

If all of the rows returned by the subquery contains a null value, a comparision that uses the ALL keyword is always false.

## How to use the ANY and SOME keyword

How to use the ANY or SOME keywords to test whether a comparison is true for any of values returned by a subquery.

Get invoices smaller than the largest invoice for vendor 15.

Same way to find largest invoice, just like before.

This is the better way cuz easier read and recommended whenever possible

But for e.g sake.

select invoice\_total from invoices where vendor\_id=34.

```
Select vendor_name, invoice_number, invoice_total  
from invoices I JOIN vendors on I.vendor_id = v.vendor_id  
where invoice_total < ANY(select invoice_total from invoices where  
vendor_id=34)  
order by vendor_name;
```

How to use ANY keyword works:

Condition	Equivalenet expressions	Description
X>ANY(1,2)	X>1	
X<ANY(1,2)	X<2	
X=ANY(1,2)	X IN (1,2)	
X <> ANY(1,2)	(x <>1) or (x<> 2)	Evaluates to true if x is not equal to at least one of the values returned by the subquery.

Description:

If the subquery doesn't return any values, or if it only return null values, a comparison that uses the any keyword evaluates to false.

## Correlated subqueries

[Correlated subqueries](#) To learn extensively click here.

However, think twice before using a correlated subquery in SQL. They can be slow,

Let's begin with an example of a *correlated subquery in SQL*. Suppose we want to find all employees with a salary higher than their average **departmental salary**. We would use the following query. Once again, I've bolded the subquery:

Code

```
SELECT
    lastname,
    firstname,
    salary
FROM employee e1
WHERE e1.salary > (SELECT avg(salary)
                    FROM employee e2
                    WHERE e2.dept_id = e1.dept_id)
```

### Example 2: A correlated subquery in SQL

The main difference between a SQL correlated subquery and a simple subquery is that correlated subqueries reference columns from the outer table. In the above example, `e1.dept_id` is a reference to the outer subquery table.

To identify a correlated query, just look for these kinds of references. If you find at least one, you have a SQL correlated subquery!

Let's look at another example. Suppose we want to obtain the names of departments that have more than 10 employees. We can use the following SQL correlated subquery:

Code

```
SELECT deptname
FROM department d1
WHERE 10 < (SELECT count(*)
              FROM employee e
              WHERE e.dept_id = d1.dept_id)
```

### Example 3: Another correlated subquery in SQL

Here's the code:

Code

```
SELECT
    lastname,
    firstname,
    salary,
    (SELECT avg(salary)
     FROM employee e2
     WHERE e2.dep_id = e1.dep_id) AS avg_dept_salary
FROM employee e1
```

#### Example 4: A SQL correlated subquery in the SELECT list

Suppose we have a table called “ `assigned_to_project` ” that stores the names of employees assigned to projects. We want to find all employees who are *not* assigned to any projects. The solution is the following query:

Code

```
SELECT
    lastname,
    firstname,
    salary
FROM employee e1
WHERE NOT EXISTS (SELECT project_id
                   FROM assigned_to_project
                   WHERE employee_id = e1.employee_id)
```

Uncorrelate subquery is executed once for the entire query.

However you can also code a correlated query that is executed once for each row that's processed by the main query.

Since correlated subqueries can be difficult to code, you may want to test a subquery separately before using it within another select statement.

To do that, however, you'll need to substitute a constant value for the variable that refers to a column in the outer query.

Once, you are sure that the subquery works on its own, you can replace the constant with a reference to the outer query so you can use it within a select statement.

Description:

A correlated query is a subquery that is executed once for each row in the main query.

In contrast, an uncorrelated subquery is executed only once.

A correlated subquery refers to a value that's provided by a column in the main query. For each different value that's returned by the main query for that column, the subquery returns a different value.

To refer to a column in the main query, you can qualify the column with a table name or alias. If a correlated subquery uses the same table as the main query, you can use table alias to remove ambiguity.

## How to use the EXISTS operator

The query tests whether the subquery returns a result set.

IOW, it tests whether the result set exists.

When you use this operator, the subquery doesn't actually return a result set to the outer query.

Instead it returns an indication of whether any rows satisfy the search condition of the subquery.

Because of that, queries that use this operator execute quickly.

--

You typically use the EXISTS operator with the correlated query.

For e.g

The query retrieves all the vendors in the Vendors table that don't have invoices in the invoices table.

An alternative is:

Subquery returns a list of ID's for each vendor that's in the invoices table. Then the main query returns some data about the vendors whose ID's aren't in the list

It executes more quickly than either of those queries.

In this e.g, the correlated subquery selects all invoices that have the same vendor\_id as the current vendor in the outer query.

Since the subquery doesn't actually return a result set, it doesn't matter what columns are included in the select clause. As a result, it's customary to just code an asterisk.

After the subquery is executed, the search condition in the WHERE clause of the main query uses the NOT EXISTS operator to test whether any invoices were found for the current vendor. IF not vendor row is included in the result set.

*Table 8: Syntax of the subquery that uses the EXISTS operator.*

Where [NOT] exists (subquery)
-------------------------------

*Table 9: ALL vendors that dont have invoices*

SELECT Vendor_id, vendor_name, vendor_state from vendors where NOT EXISTS ( Select * from invoices where vendor_id = vendor.vendor_id)
---

Description:

- You can use the EXISTS operator to test that one or more rows are returned by the subquery.
- You can use the NOT EXISTS operator to test that no rows are returned by the subquery.
- When you use these operators with a subquery, it doesn't matter what columns are specified in the SELECT clause. As a result, you just code an asterisk(\*)..

## How to code subqueries in the select clause

To do that, you code the subquery in place of column specification. As a result, the subquery must return a single value for that column.

In most cases, you code correlated subqueries in the Select clause.

e.g. The subquery calc the maximum invoice date for each vendor in the vendors table.

To do that, subquery refers to the vendor\_id column from the vendors table in the main query.

-Difficult to read.

-You can replace with a join.

Get the most recent invoice date for each vendor

```
Select vendor_name, ( select max(invoice_date) from invoices where  
vendor_id = vendors.vendor_id) as latest_inv  
from vendors  
order by latest_inv desc
```

The same query restated using a join

```
Select vendor_name, max(invoice_date) as latest_inv  
from vendors v  
LEFT join invoices l on v.vendor_id = l.vendor_id  
Group by vendor_name  
order by latest_inv DESC
```

Description:

-When you code a subquery in the SELECT clause, the subquery must return a single value.

-When you code a subquery in the select clause, you typically uses a correlated subquery.

-A query that includes a subquery in its SELECT clause can typically be restated using a join instead of the subquery. Because a join is usually faster and easier to read, subqueries are seldom coded in the select clause.

## How to code subqueries in the FROM clause

To do that you code a subquery in place of a table specification.

When you code a subquery in the from clause, it can return a result set that contains any numbers of rows and columns. This result set is sometimes referred to as an inline view.

Subqueries are typically used in the FROM clause to create inline views that provides summarized data to another summary query.

When you code a subquery in the FROM clause, you must assign a table alias to the subquery. This is required even if you don't use a table alias of t (for temporary table) to the subquery.

Get the largest invoice total for the top vendors in each state

```
select vendor_state, max(sum_of_invoices) as max_sum_of_invoices
from
(
  select vendor_state, vendor_name,
  sum(invoice_total) as sum of invoices
  from vendors v join invoices I
  on v.vendor_id = I.vendor_id
  Group by vendor_state, vendor_name
) t
group by vendor_state
```

Description:

When you code a subquery in the FROM clause, you must assign an alias to it. Then, you can use that alias just as you would any other table name or alias.

When you code a subquery in the FROM clause, you should use an alias for any columns in the subquery that performs calc. Then, the inline view can use these aliases as the column names of the table.

e.g.

sum(invoices_total) as sum of invoices	max(sum of invoices)
---	----------------------

## How to code the ORDER BY clause

The order by clause specifies the sort order for the rows in a result set.

### -How to sort by a column name:

The expanded syntax of the ORDER BY clause

```
Order by expression [ASC | DESC] [, expression [ ASC | DESC]] ...
```

```
Select vendor_name,  
concat(vendor_city, ' ', 'vendor_state', ' ', vendor_zip_code) as address  
from vendors  
order by vendor_name DESC
```

An order by clause that sorts by three columns.

```
Select vendor_name,  
concat(vendor_city, ' ', 'vendor_state', ' ', vendor_zip_code) as address  
from vendors  
order by vendor_state, vendor_city, vendor_name
```

Description:

By default, in an ascending order, special characters appear first in the sort sequence, followed by number, then letters. This sort order is determined by the character set used by the server, which you can change when you start the server.

Null values appear first in the sort sequence, even if you're using DESC.

You can sort by any columns in the base table regardless of whether its included in the SELECT Clause.

## How to sort by an alias, expression, or column numbers

Table 10: An order by clause that uses an alias

```
Select vendor_name,  
concat(vendor_city, ' ', 'vendor_state', ' ', vendor_zip_code) as address  
from vendors  
order by address, vendor_name
```

Illustration 1: Note that within the address column, the result set is also sorted by the vendor\_name column

Table 11: An Order by clause that uses an expression

```
Select vendor_name,  
concat(vendor_city, ' ', 'vendor_state', ' ', vendor_zip_code) as address  
from vendors  
order by concat( vendor_contact_last_name, vendor_contact_first_name)
```

Illustration 2: You can also use an arithmetic or string expression in the order by clause. Also note that neither of these columns is included in the SELECT CLAUSE

Table 12: An order by clause that uses column positions

```
Select vendor_name,  
concat(vendor_city, ' ', 'vendor_state', ' ', vendor_zip_code) as address  
from vendors  
order by 2,1
```

Illustration 3: To use this technique, you code the numbers that corresponds to the column of the result set, where 1 is the first column and so on.

Description:

The order by clause can include a column alias that's specified in the select clause if the column alias does not include spaces.

## How to code the limit clause

The limit clause specifies the maximum number of rows that are returned in the result set.

For most queries, you want to see the entire result set you wont use this clause. However, there maybe times when you want to retrieve just a subset of a larger result set.

*Table 13: The expanded syntax of the LIMIT clause*

LIMIT [offset,] row_count
---------------------------

How to limit the number of rows.

mysql>(query) limit 5

How to return a range of rows

mysql> (query) limit 100,100

Description:

You can use the Limit clause to limit the number of rows returned by the select statement. This clause takes one or more two integers arguments.

If you code a single argument, it specifies the maximum row count, beginning within the first row. If you code both arguments, the offset specifies the first row to return where the offset of the first row is 0.

if you want to retirve all of the rows from a certain offset to the end of the result set, code -1 for the row count.

Typically, you'll use an ORDER by clause whenever you use the limit clause.



## How to work with aggregate function

In chapter 3, you learned how to use scalar func, which operates on a single value and return a single value.

How to use aggregate functions, which operates on a series of values and return a summary of a value.

Because aggregate func typically operate on the values in columns, they are sometimes referred to as column functions.

A query that contains one or more aggregate functions is typically referred to as a summary query.

### How to code aggregate functions:

*Table 14: A summary query that counts unpaid invoices and calculates the total\_due*

```
Select count(*) as number of invoices, sum(invoice_total – payment_total – credit_total) as total_due from invoices  
where invoice_total- payment_total – credit_total > 0
```

when you use these functions, you can also code the ALL or Distinct keyword.

The ALL keyword is the Default, which means that all values are included in the calculation. The exceptions are null value, which are always excluded from these functions.

If you don't want duplicate values included, you can code the distinct keyword. In most case, you'll use Distinct only with the count function where a column is typically used as an argument.

But you can't use the ALL or Distinct keyword or an expression with Count(\*) .

The value of this functions are numbers of rows in the base table that satisfy the search condition of the query, including rows with null values.

You wont use distinct on Min or Max because it has no effect on those functions.

**Everything above is in regard with Aggregate functions.**

And it doesn't make sense to use it with the AVG and sum functions.

Description:

The expression you specify for the AVG and SUM functions must result in a numeric value. The expression for the MIN, MAX and Count functions can result in a numeric, date or string value.

By default all values are included in the calc regardless of whether they're duplicated. If you want to omit duplicate values, code the distinct keyword. The keyword typically used with the count function.

All of the aggregate func except the count(\*) ignore null values.

If you code an aggregate function in the select clause, that clause can't include non-aggregate columns from the base table unless the columns are named on a **group by** clause. -- understood

e.g.

```
select aisle, sum(quantity) from groceries group by aisle;
```

In some db's, It selects the first value in the group and allows you to include non-aggregate columns.(personal experience)

## Queries that uses aggregate functions

Interesting usecase:

*Illustration 4: A summary query that uses the count(\*), avg and sum functions*

```
Select 'After 1/1/2011' as selection_date, count(*) as number_of_invoices,  
round(avg(invoice_total),2) as avg_invoice_amt ,  
sum(invoice_total) as total_invoice_amt  
from invoices  
where invoice_date > '2011-01-01'
```

Here, the string ' after 1/1/2011' is used to tell a proper summary.

Another usecase

*Illustration 5: A summary query that uses the distinct keyword*

```
Select 'After 1/1/2011' as selection_date,  
count(distinct vendor_id) as number_of_vendors,  
count(*) as number_of_invoices,  
round(avg(invoice_total),2) as avg_invoice_amt ,  
sum(invoice_total) as total_invoice_amt  
from invoices  
where invoice_date > '2011-01-01'
```

Description:

To count all of the selected rows, you typically use the count(\*) function. Alternatively(meaning to get the same result as before), you can also use the count function with the name of any column that can't contain null values. -- understood

To count only the rows with unique values in a specified column, you can code the count function with the distinct keyword followed by the name of the column.

## How to group and summarize data

The groupby clause determines how selected rows are grouped.

Selected using the where condition.

Having clause determines which groups are included in the final result.

So,

These clauses are coded after the WHERE clause but before the ORDER BY clause.

That makes sense because the where clause is applied before the rows are grouped, and the ORDER by clause is applied after the rows are grouped.

## Syntax

The syntax for the GROUP BY clause in SQL is:

```
SELECT expression1, expression2, ... expression_n,  
      aggregate_function (aggregate_expression)  
FROM tables  
[WHERE conditions]  
GROUP BY expression1, expression2, ... expression_n  
[ORDER BY expression [ ASC | DESC ]];
```

Link: [click here](#)

Description:

-The groupby clause groups the rows of a result based on one, or more columns or expressions. To include two or more columns or expressions, seperate them by commas.

-If you include aggregate func in the select clause, the aggregate is calculated for each group specified by the group by clause.

-The having clause specifies a search condition for a group or an aggregate. Mysql applies this condition after it gorups the rows that satisfy the search condition in the where clause.

- when a select statement includes a group by clause, the select clause can include the columns used for grouping, aggregate functions, and expressions that result in a constant value.

*Illustration 6: summary query that calc the avg invoice amount by vendor*

```
Select vendor_id, round(avg(invoice_total),2) as average_invoice_amt  
from invoices  
group by vendor_id  
having avg(invoice_total)>2000  
order by average_invoice_amount desc
```

## Queries that uses the group by and having clause

*Illustration 7: A summary query that limits the group to those with two or more invoices.*

```
Select vendor_state, vendor_city, count(*) as invoice_qty,  
round(avg(invoice_total),2) as invoice_avg  
from invoices join vendors  
on invoices.vendor_id = vendors.vendor_id  
groupby vendor_state, vendor_city  
having count(*) > 2
```

By default, the groupby clause sorts the columns in ascending order. Usually, that's what you want. However, you can change this sort order by coding the DESC keyword after the column name in the groupby clause.

Or, if u don't want to sort the result set at all or you wanna get your results faster. You can code an ORDER BY NULL clause to avoids sorting overhead, it can improve the performance of your query.

## How the having clause compares to the Where clause

- A where clause can refer to any column in the base table.
- A having clause can only refer to a column included in the select clause.
- A where clause can't contain aggregate functions
- A having clause can contain aggregate functions.

*Illustration 8: A summary query with a search condition in the WHERE clause*

```
Select vendor_name,  
count(*) as invoice_qty,  
round(avg(invoice_total),2) as invoice_avg  
from vendors join invoices  
on vendors.vendor_id = invoices.vendor_id  
where invoice_total >500  
group by vendor_name  
order by invoice_qty DESC
```

## How to code compound search conditions

*Illustration 9: A summary query with a compound condition in the having clause*

```
Select  
invoice_date,  
count(*) as invoice_qty,  
sum(invoice_total) as invoice_sum  
from invoices  
group by invoice_date  
having invoice_date between '2011-05-01' and '2011-05-31'  
and count(*) > 1  
and sum(invoice_total)>100  
order by invoice_date DESC
```

The same query,

*Illustration 10: a summary query with a where condition*

```
Select  
invoice_date,  
count(*) as invoice_qty,  
sum(invoice_total) as invoice_sum  
from invoices  
where invoice_date between '2011-05-01' and '2011-05-31'  
group by invoice_date  
having count(*) > 1  
and sum(invoice_total)>100  
order by invoice_date DESC
```

Description:

You can use the AND or OR operator to code compound search conditions in a having clause just as you can in a where clause.

If a search condition includes an aggregate function, it must be coded in the having clause. Otherwise, it can be coded in either the HAVING or the WHERE clause.

## How to use the WITH ROLLUP operator

MySQL provides an extension to standard sql that's useful for summarizing data; **WITH ROLLUP** operator.

You can use this to **add one or more summary rows** to the end of the result set that uses grouping and aggregates.

This row summarizes all the aggregate columns in the result set.

It summarizes the invoice\_count and invoice\_total\_columns.

Since the vendor\_id column cant be summarized, it assigned a null value.

*Illustration 11: A summary query that includes a final summary row*

```
Select vendor_id, count(*) as invoice_count, sum(invoice_total) as  
invoice_total  
from invoices  
group by vendor_id with ROLLUP
```

*Illustration 12: A summary query that includes a summary row for each grouping level*

```
Select vendor_state, vendor_city, count(*) as qty_vendors  
from vendors  
where vendor_state in ('IA,'NJ')  
group by vendor_state ASC, vendor_city ASC with ROLLUP
```

Description:

When you use the WITH ROLLUP operator, you can't use the order by clause. To start individual rows, you can code the ASC OR DESC keyword after the column in the GROUP BY clause.

When you use the WITH ROLLUP operator, you can't use the DISTINCT KEYWORD in any of the aggregate functions.

## **How to work with inner joins**

Inner join: only those rows that satisfy the join condition are included in the result set.

A join condition names the column in each of the two tables involved in the join and indicates how the two columns should be compared.

In most cases, you use the equal operator to retrieve rows with matching columns. However, you can also use any of the other comparision operators in a join condition. --saw it once in swati-maurya lec.

Tables are typically joined on the relationship between the primary key in one table and a foreign key in the other table.

However, you can also join tables based on relationships not defined in the database. These are called ad-hoc relationships.

If the two columns in a join condition have the same name, you must qualify them with the table name so Mysql can distinguish between them.

Note: inner keyword is optional and is seldom used.

## **How to use table aliases**

If you assign an alias to a table, you must use that alias to refer to the table throughout your query. You can't use the original table name.

## How to join to a table in another database

All tables pertaining to accounts payable such as the Vendors and Invoices tables are stored in the database, or schema, name AP.

All tables pertaining to order management are stored in a database named op.

All tables that are used by the smaller examples presented in this book are stored in a database named EX.

You may occasionally need to join to a table that's in another database.

Description:

MySQL server can store tables in multiple databases. These db's are sometimes referred to as schemas.

When you run a SELECT statement against one database, you can join a table in another database if you have appropriate permissions.

To do that, you must prefix the table name in the other database with the name of that database.

## How to use compound join conditions

Although a join condition typically consist of single comparision, you can include two or more comparisions in a join condition using the AND or OR operators.

For e.g

uses the AND operator to return the first and last names of all customers in the customers table whose first and last names also exist in the Employees table.

*Illustration 13: An inner join with two conditions*

```
Select customer_fname, customer_lname from customer c  
join employee e  
on c.customer_fname=e.fname  
and c.customer_lname=e.lname;
```

## How to use a self join

*Illustration 14: A self-join that returns vendors from cities in common with other vendors*

```
Select distinct v1.vendor_name, v1.vendor_city,  
v1.vendor_state  
from vendors v1 join vendors v2  
on  
v1.vendor_city=v2.vendor_city  
and  
v1.vendor_state=v2.vendor_state  
and  
v1.vendor_name <> v2.vendor_name  
order by v1.vendor_state, v1.vendor_city
```

output:

vendor_name	vendor_state	vendor_city
Kazim	AZ	pheonix
sahil	AZ	pheonix
faraaz	NY	New york
...	...	...

## How to use the implicit inner join syntax

Instead of coding a join condition in the from clause, you can code it in the where clause along with any search conditions. In that case, you list the tables in the FROM clause separated by commas.

*Illustration 15: join the vendors and invoices table*

```
Select invoice_number, vendor_name  
from vendors v, invoices i  
where v.vendor_id = i.vendor_id  
order by invoice_number;
```

## How to work with String functions

Function	Description
concat(str1[,str2]...)	
concat_ws(sep,str1[,str2] ...)	
LTRIM	Return the string with any leading spaces removed.
RTRIM	Return the string with any trailing spaces removed.
TRIM ( [BOTH   LEADING   TRAILING ] [remove] from str)	Return the string without leading or trailing occurrence of the specified removed string. If remove string is omitted spaces are removed.
LENGTH(str)	
locate(find,search[,start])	Returns the position of the first occurrence of the find string in the search string, starting at the specified start position. If the start position is omitted, the search starts at the beginning of the string. If the string isn't found, the function returns zero.
LEFT(str,length)	Returns the specified number of characters from the beginning of the string.
RIGHT(str,length)	Returns the specified number of characters from the end of the string.
substring_index(str,delimiter,count)	Returns the substring before the specified number of occurrences of the specified delimiter string. If count is positive, it returns from the beginning of the string. If count is negative, it returns from the end of the string.
substring(str,start[,length])	Returns the specified number of characters from the string starting at the specified start position. If length is omitted, it returns from the start position to the end of the string.
REPLACE(search, find, replace )	Replace the search string with all the occurrence of the find string with the replace string.
INSERT(str, start, length, insert)	Returns the string with the specified insert string into it starting at the specified start position to the end of the string.
Reverse	
Lower	
Upper	
LPAD(str, length, pad)	Returns the string padded on the left with the specified pad string until its the specified length. If the string is longer than the length. It's truncated
RPAD	
SPACE(count)	Returns the space character repeated count times
Repeat(str,count)	Returns the specified string repeated count times.

Function	Result
concat('last' , 'First')	'lastFirst'
concat_ws('_', 'last', 'first')	'last_first'
LTRIM(' MySql ')	' MySql '
RTRIM	
TRIM( BOTH '*' from '***** MySql ***'	' MySql '
left(' MySql ',3)	'MyS'
right(' MySql ',3)	'Sql'
Substring('555) 555-1212', 7, 8)	'555-1212'
substring_index('https://www.youtube.com/', 'Youtube.com' ', -2)	
length('Mysql')	5
Length(' MySql ')	8
locate('sql', ' mysql')	5
Locate('-', '(559) 555-1212')	10
Replace(right('559) 555-1212', 13), (', '-')	559-555-1212
insert(' MySql ',1,0,'Murach's)	'Murach's MySql'



## How to sort by a string column that contains number

When you attempt to sort string data that's stored in a numeric column.

Table 15: Sorted by the emp\_id column explicitly cast as an integer

```
Select *  
from string_example  
order by cast(emp_id as signed)
```

Table 16: Sorted by the emp\_id column implicitly cast as an integer

```
Select *  
from string_example  
order by emp_id + 0
```

Table 17: Sorted by the emp\_id column after it has been padded with leading zeroes

```
Select LPAD(emp_id,2,'0') as emp_id, emp_name  
from string_example  
order by emp_id
```

## How to parse a string

Another problem you may

substring\_index(col\_name,'delimiter',1)

```
Select emp_name,  
substring_index(emp_name,' ',1) as first_name,  
substring_index(emp_name,' ', -1) as last name  
from string_sample;
```

```
Select emp_name,  
locate(' ',emp_name,1) as first_name, substring_index(emp_name,' ', -1) as  
last name from string_sample;
```

locate(' ',col\_name,1) as first\_name

locate(' ',col\_name,locate(' ',col\_name,1)+1) as last name

```
select emp_name,  
substring(col_name,index,length)  
substring(col_name, locate)
```

```
Select emp_name,  
substring(emp_name,1,locate(' ',emp_name) ) as first_name  
substring(emp_name,locate(' ',emp_name) + 1) as last_name  
from string_sample
```

Desc:

if a string consists of two or more components, you can parse it into its individual components. To do that, you can use the SUBSTRING\_INDEX, SUBSTRING and locate functions.

## How to work with numeric data

Functions	Description
Round(number [, length] )	Returns the number rounded to the precision specified by len. If len is zero, the decimal digits are omitted. If len < 0 then the digits left to the decimal points are rounded.
Truncate(number [,length] )	Same as above. Just len < 0 isn't defined in the book definition.
Ceiling(number)	Return the smallest integer that is <b>greater</b> than or equal to that number
floor(number)	Return the largest integer that is less than or equal to that number.
ABS(number)	Return the absolute value of the number.
SIGN(number)	Return the sign for the number as -1 for a negative number. 1 for positive number ,and zero if the number is 0
SQRT(number)	
Power(number , power)	
RAND([integer])	It returns a floating point number between 0-1. Integer provides a seed value for the random number generator.

Function	Result
Round(12.49,0)	12
Round(12.50,0)	13
Round(12.49,1)	12.5
Truncate(12.51,0)	12
Truncate(12.49,1)	12.4
Ceiling(12.5)	13
ceiling(-12.5)	12
Floor(-12.5)	-12
floor(12.5)	12
ABS(-1.25)	1.25
ABS(1.25)	1.25
SIGN(-1.25)	-1
SIGN(1.25)	1

## How to search for floating-point numbers

floating point types such as double or float types stores approximate values, not exact value.

The details of why are beyond the scope of the book.

From a practical pov, you don't want to search for exact values when you're working with floating point numbers. If you do, you'll miss values that are approximately equal to the value you're looking for.

e.g

float_id	float_value
1	0.9999999999999999
2	1
3	0.10000000000000001
4	12345678

Table 18: A search for an exact value that doesn't include two approximate values

```
>  
select *  
from float_sample  
where float_value=1
```

float_id	float_value
2	1

```
1 record  
query successful (0.23 s)
```

Instead search for approximate values,

Select * from float_sample where float_value between 0.99 and 1.01
Select * from float_sample where round(float_value,2)=1.0

float_id	float_value
1	0.9999999999999999
2	1
3	0.10000000000000001





## Section 5: CRUD COMMANDS

Crud is an acronym for create, read, update, delete.

Update and delete: how do we change things in a db if we make mistakes.

Lets say User forgot their password. We gonna send a reset link.

We need to change it in the db.

Or lets say user leave our website for some really bad reason, so we can delete that user.

### Read

How do we retrieve and search data?

```
UPDATE cats SET breed='Shorthair'  
WHERE breed='Tabby';
```

Best practices:

- Try selecting before you update.

--

### How to change datatype of a column.

```
alter table shirts modify column shirt_size varchar(2);
```

### How to update multiple columns together

```
update shirts set shirt_size = "XS" , color = "off white" where color='white';
```

## The world of string functions.

things I could do with the data.

For numbers, I could add them together.

Or I might to reverse a string.

Or replace all spaces with certain character, maybe \_.

Or conjoin first name or last name.

We can concatinate.

We can substring.

We can uppercase, lowercase.

## Running SQL files

```
| source file_name.sql |
```

## Text functions

```
% cat data1
+-----+
| author_fname | author_lname |
+-----+
| Jhumpa      | Lahiri      |
| Neil        | Gaiman      |
| Neil        | Gaiman      |
| Jhumpa      | Lahiri      |
| Dave        | Eggers      |
| Dave        | Eggers      |
| Michael     | Chabon      |
| Patti       | Smith       |
| Dave        | Eggers      |
| Neil        | Gaiman      |
| Raymond     | Carver      |
| Raymond     | Carver      |
| Don          | DeLillo     |
| John         | Steinbeck   |
| David        | Foster Wallace |
| David        | Foster Wallace |
| Adam         | Smith       |
+-----+
```

```
kazimsyed@kazims-MacBook-Air ~/Desktop/mysql_learn
% sed -n 's/|/"/1
s/|/", "/1
s/|"/)/1
s/ //g
quote> 4,$p' data1
("Jhumpa","Lahiri")
("Neil","Gaiman")
("Neil","Gaiman")
("Jhumpa","Lahiri")
("Dave","Eggers")
("Dave","Eggers")
("Michael","Chabon")
("Patti","Smith")
("Dave","Eggers")
("Neil","Gaiman")
("Raymond","Carver")
("Raymond","Carver")
("Don","DeLillo")
("John","Steinbeck")
("David","FosterWallace")
("David","FosterWallace")
("Adam","Smith")
```

code:

```
sed -n 's/|/(/1
s/|/", "/1
s/|/)/1
s/ //g
4,$p' data1
```

```
kazimsyed@kazims-MacBook-Air ~/Desktop/mysql_learn/filter
% gawk 'BEGIN{ORS=","}{print $0}' data1
("Jhumpa", "Lahiri"), ("Neil", "Gaiman"), ("Neil", "Gaiman"), ("Jhumpa", "Lahiri"), ("Dave", ")
"), ("Patti", "Smith"), ("Dave", "Eggers"), ("Neil", "Gaiman"), ("Raymond", "Carver"), ("Raymo
nbeck"), ("David", "FosterWallace"), ("David", "FosterWallace"), ("Adam", "Smith"), +-----
```

**gawk 'BEGIN{ORS=", "}{print \$0}' data1**



The screenshot shows a Gedit text editor window with the title bar "string\_fun.sql (~/Desktop/mysql\_learn) - gedit". The code in the editor is as follows:

```
1
2 create database learn_str;
3 use learn_str;
4 create table names(id int auto_increment primary key, fname varchar(30), lname varchar(30));
5
6 insert into names (fname,lname) values ("Jhumpa","Lahiri"),("Neil","Gaiman"),("Neil","Gaiman"),("Jhumpa","Lahiri"),
    ("Dave","Eggers"),("Dave","Eggers"),("Michael","Chabon"),("Patti","Smith"),("Dave","Eggers"),("Neil","Gaiman"),
    ("Raymond","Carver"),("Raymond","Carver"),("Don","DeLillo"),("John","Steinbeck"),("David","FosterWallace"),
    ("David","FosterWallace"),("Adam","Smith");
```

data:

```
mysql> select * from names;
+----+-----+-----+
| id | fname | lname |
+----+-----+-----+
| 1  | Jhumpa | Lahiri |
| 2  | Neil   | Gaiman |
| 3  | Neil   | Gaiman |
| 4  | Jhumpa | Lahiri |
| 5  | Dave   | Eggers |
| 6  | Dave   | Eggers |
| 7  | Michael | Chabon |
| 8  | Patti  | Smith  |
| 9  | Dave   | Eggers |
| 10 | Neil   | Gaiman |
| 11 | Raymond | Carver |
| 12 | Raymond | Carver |
| 13 | Don    | DeLillo |
| 14 | John   | Steinbeck |
| 15 | David  | FosterWallace |
| 16 | David  | FosterWallace |
| 17 | Adam   | Smith  |
+----+-----+-----+
17 rows in set (0.00 sec)
```

## text functions:

```

mysql> select concat_ws(' ',fname, lname, 'alaska') from names;
+-----+
| concat_ws(' ',fname, lname, 'alaska') |
+-----+
| Jhumpa_Lahiri_alaska
| Neil_Gaiman_alaska
| Neil_Gaiman_alaska
| Jhumpa_Lahiri_alaska
| Dave_Eggers_alaska
| Dave_Eggers_alaska
| Michael_Chabon_alaska
| Patti_Smith_alaska
| Dave_Eggers_alaska
| Neil_Gaiman_alaska
| Raymond_Carver_alaska
| Raymond_Carver_alaska
| Don_DeLillo_alaska
| John_Steinbeck_alaska
| David_FosterWallace_alaska
| David_FosterWallace_alaska
| Adam_Smith_alaska
+-----+
17 rows in set (0.00 sec)

```

Figure 8: concat with separator

```

mysql> select concat_ws(' ', fname, substr(lname,1,3), '...') from names;
+-----+
| concat_ws(' ', fname, substr(lname,1,3), '...') |
+-----+
| Jhumpa Lah ...
| Neil Gai ...
| Neil Gai ...
| Jhumpa Lah ...
| Dave Egg ...
| Dave Egg ...
+-----+

```

Figure 9: substr

-count starts from 1.

- if the last argument isn't given, then it is consider till end.

e.g1-

```
SELECT SUBSTRING( 'Hello World' , 7 );
```

World

e.g2-

```
SELECT SUBSTRING( 'Hello World' , -3 );
```

rld

```
SELECT  
    REPLACE('cheese bread coffee milk', ' ', ' and ');
```

cheese and bread and coffee and milk

Figure 10: Replace

## Like keyword

There's a book I'm looking for...

But I can't remember the title!

I know the author's first name is David...

Or wait, maybe it's Dan...or Dave

# LIKE

```
WHERE author_fname LIKE '%da%'
```

↑ ↑  
WILDCARDS

# LIKE

```
WHERE author_fname LIKE 'da%'
```

"I want names that start with da "

# LIKE

```
WHERE stock_quantity LIKE '_____'
```

Yes, that is 4 underscores

# But What If...

I'm searching for a book with a '%' in it

I'm searching for a book with an '\_' in it

```
WHERE title LIKE '%\_%'
```

Figure 11: we need to escape.

## Aggregate functions:

```
[mysql]> select distinct count(writer_name) from books_view;
+-----+
| count(writer_name) |
+-----+
|          16 |
+-----+
1 row in set (0.00 sec)

[mysql]> select count(distinct writer_name) from books_view;
+-----+
| count(distinct writer_name) |
+-----+
|          9 |
+-----+
1 row in set (0.00 sec)

[mysql]> select count(all writer_name) from books_view;
+-----+
| count(all writer_name) |
+-----+
|          16 |
+-----+
1 row in set (0.00 sec)
```

All or distinct keywords specify whether duplicate rows are returned.

