# SQL Concepts

A **database** is a container that holds tables and other SQL structures related to those tables.

A **column** is a piece of data stored by our table.

A **row** is a single set of columns that describe attributes of a single thing.

Columns and rows together make up a table

The word **field** is often used instead of column

The word **record** is often used instead of row

## CREATE DATABASE

**CREATE DATABASE table\_name**;

**USE DATABASE table\_name**;

Capitalization and underscores help our program in SQL (even though SQL doesn’t need them)

## CREATE TABLE

**CREATE TABLE table\_name (column\_name DATATYPE, … )**

## DATATYPES

CHAR(CHARACTER)

It’s a rigid and prefers its data to be a set length

INT(INTEGER)

The number should be whole and it can be negative too

BLOB

It takes large amount of text data

DATE

It keeps track of our dates and time

VARCHAR

It holds text data of up to 255 characters in length

DATETIME or TIMESTAMP

Keeps track of the dates

DEC(DECIMAL)

It gives numbers with decimal places

### DEC(3,2)

The first argument is the total number of digits of precision, and the second argument is the number of digits after the decimal point.

## DESCRIBE

**DESC table\_name**

Describes the structure of the table

## DROP TABLE

**DROP TABLE table\_name**

It deletes the table and any data in it

## INSERT DATA

**INSERT INTO table\_name(column name, …)**

**VALUES(‘value1’, …)**

**INSERT INTO table\_name**

**VALUES(‘value1’, …)**

### INSERTING DATA

Insert data with single quotes on it.

**Handle quotes with a backlash**

INSERT INTO table\_name VALUES(‘Grover**\’s** MILL’, ‘NJ’)

**Handle quotes with an extra single quote**

INSERT INTO table\_name VALUES(‘Grover**’’s** MILL’, ‘NJ’)

## INNER NULL

CREATE TABLE table\_name (column\_name DATATYPE **NOT NULL**, …)

## DEFAULT

CREATE TABLE table\_name (column\_name **DEFAULT 1.00**)

Using a **default** value fills the empty columns with a specified value

## SELECT

**SELECT \* from table\_name**

Select all the data from all the columns from table\_name

### WHERE

SELECT \* FROM table\_name **WHERE column\_name = ‘column\_data’**

Select all the data from the selected column where that column consists of selected data

The **VARCHAR, CHAR, BLOB, DATA** and **TIME** data types need single quotes. The numeric types **DEC** and **INT**, do not

### SPECIFIC COLUMN

**SELECT column\_name, column\_name, … FROM table\_name**

### BOOLEAN

#### AND

SELECT column\_name FROM table\_name WHERE column\_name = ‘column\_data’

**AND** column\_name = ‘column\_data’

#### OR

SELECT column\_name FROM table\_name WHERE column\_name = ‘column\_data’

**OR** column\_name = “column\_data’

#### NOT

SELECT column\_name FROM table\_name **WHERE column\_name NOT IN (‘column\_data\_1’, ‘column\_data\_2’)**

SELECT column\_name FROM table\_name **WHERE NOT column\_name BETWEEN number\_1 and number\_2**

SELECT column\_name FROM table\_name **WHERE NOT column\_name LIKE ‘A%’**

### COMPARISON OPERATOR

#### Greater than >

SELECT column\_name FROM table\_name column\_name = ‘column\_data’ AND

column\_name **>** column\_value

#### Less than <

SELECT column\_name FROM table\_name column\_name = ‘column\_data’ AND

column\_name **<** column\_value

#### Less than and Equals <=

SELECT column\_name FROM table\_name column\_name = ‘column\_data’ AND

column\_name **<=** column\_value

#### Greater than and Equals >=

SELECT column\_name FROM table\_name column\_name = ‘column\_data’ AND

column\_name **>=** column\_value

### NULL

SELECT column\_name FROM table\_name WHERE column\_name **IS NULL**

### LIKE

It is a wildcard operator

SELECT \* FROM table\_name WHERE column\_name **LIKE ‘%CA’**

The percent is a stand-in for any number of unknown characters

SELECT \* FROM table\_name WHERE column\_name **LIKE ‘\_CA’**

The underscore is a stand-in for just one unknown character

### BETWEEN

SELECT column\_table FROM table\_name WHERE column\_name **BETWEEN number\_value\_1 and number\_value\_2**

### IN

SELECT column\_table FROM table\_name WHERE **column\_name IN (‘data\_1’, ‘data\_2’)**

## DELETE

We can use WHERE clauses with DELETE statements the same way we use them with INSERT statements

**DELETE FROM table\_name WHERE table\_column = ‘table\_data’**

## UPDATE

**UPDATE table\_name SET column\_name = ‘column\_data’ WHERE column\_name = ‘column\_data’**

Update statements can replace DELETE/INSERT combinations. Update statements can be used on multiple records in our table. We can use them with basic math operators to manipulate our numeric values

UPDATE table\_name **SET column\_name = column value + 1**

WHERE column\_name = ‘column\_data’

OR column\_name = ‘column\_data’

UPDATE table\_name **SET column\_name = RIGHT(another\_column\_name, 2)**

## ATOMIC DATA

A little piece of information that can’t or shouldn’t be divided. Data should be broken into the smallest pieces of data that can’t or shouldn’t be divided.

**Rule 1:** A column with atomic data can’t have several values of the same type of data in that column

**Rule 2:** A table with atomic data can’t have multiple columns with the same type of data.

## NORMAL

Making our atomic is the first step in creating a NORMAL table.

**Benefit 1:** Normal tables won’t have duplicate data, which will reduce the size of our database

**Benefit 2:** With less data to search through our queries will be faster

Each row of data must contain atomic values

Each row of data must have a unique identifies, known as a PRIMARY KEY

## PRIMARY KEY

A primary key is a column in our table that makes each record unique.

**Rule 1:** A primary key can’t be NULL

**Rule 2:** A primary key must be given a value when the record is inserted

**Rule 3:** The primary must be compact

**Rule 4:** The primary key values can’t be changed

CREATE TABLE table\_name(

**column\_name\_1 DATATYPE NOT NULL AUTO\_INCREMENT,**

column\_name\_2 DATATYPE,

**PRIMARY KEY (column\_name\_1)**

)

Auto Increments just automatically fill the column with value that starts on row 1 with a value of 1 and goes up in increments of 1.

## SHOW

**SHOW CREATE TABLE table\_name**

## BACKTICK

It allows us to use whatever field name we wish when designing your table. Sometimes it makes a lot of sense to name a field `key`, `order`, or `values`... all of which require backticks when referring to them.

## ALTER

CHANGE: Changes both the name and data type of an existing column

MODIFY: Changes the data type or position of an existing column

ADD: Add a column to our table and it needs data type

DROP: Drop the column from the table.

### RENAME

ALTER TABLE **old\_table\_name RENAME TO new\_table\_name**

### CHANGE

ALTER TABLE table\_name **CHANGE COLUMN old\_column\_name new\_column\_name** INT NOT NULL AUTO\_INCREMENT, **ADD PRIMARY KEY** **(new\_column\_name)**

ALTER TABLE table\_name **CHANGE** **COLUMN old\_column\_name new\_column\_name** **NEW\_DATA\_TYPE**

### MODIFY

ALTER TABLE table\_name **MODIFY COLUMN column\_name DATATYPE**

### ADD

ALTER TABLE table\_name **ADD COLUMN column\_name DATATYPE AFTER specific\_column\_name**

**ALTER TABLE table\_name ADD COLUMN column\_name DATATYPE FIRST**

### DROP

ALTER TABLE table\_name **DROP COLUMN column\_name**

## STRING FUNCTION

Text values and values stored in CHAR or VARCHAR columns are known as strings. String functions allow us to select part of a text column.

We can use string functions in combination with SELECT, UPDATE, and DELETE

### RIGHT

SELECT **RIGHT(column\_name, count\_number\_from\_right)** FROM table\_name

### SUBSTRING

SELECT **SUBSTRING(COLUMN\_NAME, start\_position, length)** FROMtable\_name

### SUBSTR

UPDATE table\_name SET column\_name = **SUBSTR(column\_name\_1, LENGTH(column\_name\_2) + 2)**

### SUBSTRING INDEX

SELECT **SUBSTRING\_INDEX(column\_name, ‘delimiter’, which\_one)** FROM table\_name

**which\_one,** if it is “1”, it means it is looking for the first comma. If it were “2” it would keep going until finds a second comma and grab everything in front of that.

### UPPER

SELECT **UPPER(column\_name)** FROM table\_name

### LOWER

SELECT **LOWER(column\_name)** FROM table\_name

### LTRIM

SELECT **LTRIM(column\_name)** FROM table\_name

Removes any string with extra spaces removed from the left

### RTRIM

SELECT **RTRIM(column\_name)** FROM table\_name

Removes any string with extra spaces removed from the right

### REVERSE

SELECT **REVERSE(column\_name)** FROM table\_name

Reverses the order of letters in the string

### LENGTH

SELECT **LENGTH(column\_name)** FROM table\_name

Returns a count of hominy characters are in the string

## CONDITIONS

UPDATE table\_name

SET column\_name =

**CASE**

**WHEN column\_name\_1 = some\_value\_1**

**THEN new\_value\_1**

**WHEN column\_name\_2 = some\_value\_2**

**THEN new\_value\_2**

**ELSE new\_value\_3**

**END**

## ORDER BY

SELECT column\_name FROM table\_name WHERE column\_name = ‘column\_data’ **ORDER BY column\_name**

It allows us to alphabetically order any columns

### ORDER TWO COLUMNS

**SELECT** **column\_name\_1, column\_name\_2 FROM table\_name** **ORDER BY order\_column\_name\_1, order\_column\_name\_2**

column\_name\_1 will get ordered by order\_column\_name\_1

column\_name\_2 will get ordered by order\_column\_name\_2

### ORDER WITH MULTIPLE COLUMNS

**SELECT \* FROM table\_name ORDER BY order\_column\_name\_1, order\_column\_name\_2, order\_column\_name\_3**

First the columns are ordered by order\_column\_1 then order\_column\_2 and at the end order\_column\_3

We can sort by as many columns as we need

### SQL RULES OF ORDER

Non-alphabet characters show up before and after numbers.

Numbers show up before text characters.

NULL values show up before numbers and alphabet characters

Uppercase characters show up before and lowercase letters

A 1 will show up before A1

Order: **!= & ( \* + - ? @ ~**

### DESC

Keyword DESC after the column name in ORDER BY clause reverses the order of the result

SELECT \* FROM table\_name ORDER BY order\_column\_name\_1 **DESC**

### ASC

SELECT \* FROM table\_name ORDER BY order\_column\_name\_1 **ASC**

We can put ASC but it is not the necessary since the default order is ASC

## ARITHMETIC FUNCTION

### SUM

SELECT **SUM(column\_name)** FROM table\_name WHERE another\_column\_name = “column\_data”

### GROUP BY

SELECT **column\_name\_1,** **SUM(column\_name\_2)** FROM table\_name **GROUP BY column\_name\_1 ORDER BY SUM(column\_name\_2) DESC**

### AVERAGE

SELECT **column\_name\_1,** **AVG(column\_name\_2)** FROM table\_name **GROUP BY column\_name\_1**

### MIN AND MAX

SELECT **column\_name\_1,** **MAX(column\_name\_2)** FROM table\_name **GROUP BY column\_name\_1**

SELECT **column\_name\_1,** **MIN(column\_name\_2)** FROM table\_name **GROUP BY column\_name\_1**

### COUNT

SELECT **COUNT(column\_name)** FROM table\_name

It will return the number of rows in a column

### DISTINCT

SELECT **DISTINCT(column\_name)** FROM table\_name ORDER BY **another\_column\_name**

SELECT **COUNT(DISTINCT column\_name)** FROM table\_name

### LIMIT

SELECT **column\_name FROM table\_name LIMIT 2**

Lets only the first two data to show up

SELECT **column\_name FROM table\_name LIMIT 0,5**

Let us see the data from row 0 to row 5

## COMMENT

**/\*Select all the columns**

**of all the records**

**in the Customers table:\*/**

**--SELECT \* FROM Customers;**

## SCHEMA

A descript of the data (the columns and tables) in our database, along with any other related objects and the way they all connect is known as **schema**

Creating a diagram of our table lets us keep the design of the table separate from the data that inside of it.

## FOREIGN KEY

The **foreign key** is a column in a table that references the **primary key** of another table

We will be able to insert values into our foreign key that exist in the table the key came from, the parent table. This is called **referential integrity.**

Creating a foreign key as a constraint in our table give us definite advantages. We’ll get errors if we violate the rules, which will stop us accidentally doing anything to break the table.

Example:

**CREATE TABLE interest (**

--Adding the primary key command to the line where we set it up is quicker way to designate

**int\_id INT NOT NULL AUTO\_INCREMENT PRIMARY KEY,**

**interest VARCHAR(50) NOT NULL,**

**contact\_id INT NOT NULL,**

--The key came from (my\_contacts) 🡪 contact\_id and that it’s a foreign key

**CONSTRAINT my\_contacts\_contact\_id\_fk**

**FOREIGN KEY (contact\_id)**

**REFERENCES my\_contact (contact\_id)**

## RELATIONSHIP BETWEEN TABLES

|  |  |
| --- | --- |
|  | One to One |
|  | One to Many |
|  | Many to Many |

### ONE-TO-ONE

**TABLE B**

**TABLE A**

**ONLY ONE** of these rows -- Matches TO -- **ONLY ONE** of these rows

One-to-One: Exactly one row of a parent table is related to one row of child table

### ONE-TO-MANY

**TABLE B**

**TABLE A**

**ONLY ONE** of these rows -- Matches TO -- **MANY** of these rows

One-to-Many: A record in TABLE A can have many matching records in TABLE B, but a record in TABLE B can only match one record in TABLE A

### MANY-TO-MANY

**TABLE B**

**TABLE A**

**MANY** of these rows -- Matches TO -- **MANY** of these rows

### JUNCTION TABLE

Many-to-Many 🡪 A junction table holds a key from each table

## NORMAL FORM REVISITED

### ATOMIC DATA

A little piece of information that can’t or shouldn’t be divided. Data should be broken into the smallest pieces of data that can’t or shouldn’t be divided.

**Rule 1:** A column with atomic data can’t have several values of the same type of data in that column

**Rule 2:** A table with atomic data can’t have multiple columns with the same type of data.

### NORMAL

Making our atomic is the first step in creating a NORMAL table.

**Benefit 1:** Normal tables won’t have duplicate data, which will reduce the size of our database

**Benefit 2:** With less data to search through our queries will be faster

Each row of data must contain atomic values

Each row of data must have a unique identifies, known as a PRIMARY KEY

### FIRST NORMAL FORM

**1NF**

**Rule 1:** Columns contain only atomic values

**Rule 2:** No repeating groups of data

A key made of two or more columns is known as a **composite key**. A composite key is a primary key composed of multiple columns, creating a unique key. When a column’s data must change when another column’s data is modified

When a column’s data must change when another column’s data is modified, the first column is **functionally dependent** on the second.

**{Name + Power} 🡺 Composite Key**

|  |  |  |
| --- | --- | --- |
| **Name** | **Power** | **Weakness** |
| Superman | Fly | Dumb |
| Batman | Rope | Gadgets Dependent |
| Spiderman | Web | Monster |

**Shorthand Notations**

**T.x ->; T.y**

Technical term for this a shorthand notation

Relational Table is called T

Column y is functionally dependent on column x

**Partially functional dependency:** It means that a non-key column is dependent on some, but not all, of the columns in a composite primary key.

**{Name + Power} 🡺 Composite Key {Name is dependent 🡪 Initials}**

|  |  |  |  |
| --- | --- | --- | --- |
| **Name** | **Power** | **Weakness** | **Initials** |
| Superman | Fly | Dumb | SM |
| Batman | Rope | Gadgets Dependent | BM |
| Spiderman | Web | Monster | SM |

**Transitive function dependency:** If changing any of the non-key columns might cause any of the other columns to change, we have a transitive dependency.

When any non-key column is related to any of the other non-key columns.

arch\_enemey\_id 🡺 arch\_enemy\_city

|  |  |  |
| --- | --- | --- |
| **Name** | **arch\_enemy\_id** | **arch\_enemy\_city** |
| Superman | 4 | La La City |
| Batman | 5 | Arkham City |
| Spiderman | 8 | Times City |

### SECOND NORMAL FORM

**2NF**

**Rule 1:** Be in 1NF

**Rule 2:** Have no partial functional dependencies

Your 1NF table is also 2NF, if all the columns in the table are part of the primary key or It has a single column primary key.

### THIRD NORMAL FORM

**3NF**

**Rule 1:** Be in 2NF

**Rule 2:** Have no transitive dependencies

If your table has an artificial primary key and no composite primary key, it’s in 2NF

### SUMMARY NORMAL FORM

1. A column with atomic data can’t have several values of the same type of data in that column
2. A table with atomic data can’t have multiple columns with the same type of data.
3. No repeating groups of data
4. Have no partial functional dependencies
5. Have no transitive dependencies

## MULTIPLE QUERIES

### CREATE TABLE, THEN INSERT WITH SELECT

CREATE TABLE profession (id INT(11) NOT NULL AUTO\_INCREMENT PRIMARY KEY, profession varchar(20))

* Fill up the **profession column** of the **profession table** with the values from our **SELECT**

**INSERT INTO profession (profession)**

new\_**table new\_column**

**SELECT profession FROM my\_contacts**

existing\_column existing\_table

**GROUP BY profession**

**ORDER BY profession**

### CREATE TABLE WITH SELECT, THEN ALTER TO ADD PRIMARY KEY

* Create the **profession table** with **one column**, full of the values from the **SELECT**

**CREATE TABLE profession AS**

**SELECT profession FROM my\_contacts**

**GROUP BY profession**

**ORDER BY profession**

ALTER TABLE profession

ADD COLUMN id INT NOT NULL AUTO\_INCREMENT FIRST,

ADD PRIMARY KEY(id)

### CREATE, SELECT AND SELECT

* Create the profession table with both a primary key and a profession column, and fill the profession column with the values from the **SELECT**

**CREATE TABLE profession(**

**id INT(11) NOT NULL AUTO\_INCREMENT PRIMARY KEY,**

**profession varchar(20)**

**) AS**

**SELECT profession FROM my\_contacts**

**GROUP BY profession,**

**ORDER BY profession**

## ALIASES

Table aliases are also called **correlation names**

SELECT profession **AS mcprof**

FROM my\_contacts **AS mc**

**GROUP BY mc\_prof**

**ORDER BY mc\_prof**

Removing AS makes no difference but it means the same thing

SELECT profession **mcprof**

FROM my\_contacts **mc**

**GROUP BY msc\_prof**

**ORDER BY mc\_prof**

## CARTESIAN JOIN

The **cross-join** returns from one table crossed with every row from the second

**boy**

Davey

Bobby

**toy**

hula hoop

balsa glider

toy soldiers

harmonica

baseball cards

2\*5 = 10 combinations

|  |  |
| --- | --- |
| **toy** | **boy** |
| hula hoop | Davey |
| balsa glider | Bobby |
| toy soldiers | Davey |
| harmonica | Bobby |
| baseball cards | Davey |
| hula hoop | Bobby |
| balsa glider | Davey |
| toy soldiers | Bobby |
| harmonica | Davey |
| baseball cards | Bobby |

**SELECT t.toy, b.boy**

**FROM toy AS t**

**CROSS JOIN**

**boys as b**

### INNER JOIN

An **inner join** is a cross join with some result rows removed by a condition in the query.

An inner join combines the records from two tables using comparison operators in a condition

**SELECT column\_name FROM table\_1**

**INNER JOIN table\_2 ON some\_condition**

**Table Boys**

|  |  |  |
| --- | --- | --- |
| **boy\_id** | **boy** | **toy\_id** |
| 1 | Davey | 2 |
| 2 | Bobby | 1 |
| 3 | Beaver | 2 |

**Table Toys**

|  |  |
| --- | --- |
| **toy\_id** | **toy** |
| 1 | hula hoop |
| 2 | balsa glider |
| 3 | toy soldiers |

EQUIJOIN INNER JOIN

Test for equality

**SELECT boys.boy, toys.toy FROM boys**

**INNER JOIN toys**

**ON boys.toy\_id = toys.toy\_id**

**Result**

|  |  |
| --- | --- |
| **boy** | **Toy** |
| Beaver | balsa Glider |
| Bobby | hula hoop |
| Davey | balsa glider |

### NON EQUIJOIN INNER JOIN

test for inquality

**SELECT boys.boy, toys.toy FROM boys**

**INNER JOIN toys**

**ON boys.toy\_id <> toys.toy\_id**

**Result**

|  |  |
| --- | --- |
| **boy** | **Toy** |
| Beaver | hula hoop |
| Beaver | toy soldiers |
| Bobby | balsa glider |
| Bobby | toy soldiers |
| Davey | hula hoop |
| Davey | toy soldiers |

### NATURAL JOIN INNER JOIN

Identify matching column names, Column we’re joining by has the same name in both tables

**Result**

|  |  |
| --- | --- |
| **boy** | **Toy** |
| Beaver | balsa Glider |
| Bobby | hula hoop |
| Davey | balsa glider |

## SUBQUERY

A **subquery** is a query that is wrapped within another query. It’s also called INNER query.

OUTER QUERY + INNER QUERY = Query with a subquery

**OUTER QUERY**

**SELECT mc.firstname, mc.last\_name, cj.title**

**FROM current\_job AS cj**

**NATURAL JOIN my\_contacts as mc**

**WHERE cj.title IN (*SELECT title from job\_l istings*);**

***INNER QUERY***

SELECT mc.first\_name, mc.last\_ame, cj.salary

FROM my\_contact as mc

NATURAL JOIN current\_job as cj

WHERE **cj.salary = (SELECT MAX(cj.salary) FROM current\_job cj)**

If a subquery is used as column expression in a SELECT statement, it can only return one value from one column.