COMSW4111_003_2024_03: Midterm

Version 1.0, 2024-10-16 5:00 AM

Instructions

- We will be managing time **very strictly.**
    - Do NOT open this cover page until the exam moderator/proctor states that you can begin.
    - You will be given a 10 minute warning and 5 minute warning before the exam time is complete.
    - You must stop writing immediately when the proctor announces that the exam is over. Failing to stop will result in substantial point deductions on the exam.
- You cannot use, look at, touch, … any electronic devices. All bags, books, coats, etc. should be in one of the collection areas at the front of the classroom.
- Your desk/writing area must be clear of everything other than this exam document, the cheat sheet we provided, and your one page "cheat sheet."
- Write your answers in the spaces provided after questions. If you need additional space, you may use the backside of a page. Please indicate that your answer continues on the backside of the page.

"A man who uses a great many words to express his meaning is like a bad marksman who instead of aiming a single stone at an object takes up a handful and throws at it in hopes he may hit." — Samuel Johnson

We will deduct points if your answers are not concise and to the point.

Written Questions

This section requires short, written answers. No question requires more than 5 sentences or bullet points.

W1 What are the three basic concepts in the ER data model?

W2 Give three benefits of using ER Diagrams for database projects.

W3 SQL has *types,* e.g. *varchar(5).* ER modeling recommends that a columns values come from an *atomic domain*. Briefly explain the difference between a domain and a SQL data type. Briefly explain the concept of atomic domain. Provide a simple example to illustrate your answer.

W4 Explain the concept of *Physical Data Independence.*

W5 A fundamental type of database user is a *Database Administrator.* List three functions or Tasks that a database administrator performs.

W6 Briefly explain the concepts of *natural key* and *surrogate key.*

W7 Consider the SQL statement,

```
delete from instructor
where salary < (select avg (salary)
                from instructor);
```

Briefly explain why executing this statement may produce incorrect results.

W8 Is it always possible to insert, delete or update that base tables that a SQL CREATE VIEW statement references? If not, give two reasons why it is not always possible. You can use the schema that accompanies the recommended textbook to provide an example.

W9 Consider Codd's Rule 10: Integrity Independence:
A database must be independent of the application that uses it. All its integrity constraints can be independently modified without the need of any change in the application. This rule makes a database independent of the front-end application and its interface.

List 3 concepts from the SQL Data Definition Language that implement this rule. What is an example of a problem caused when all integrity rules are only implemented in applications using the database.

W10 Briefly explain the concepts of a *composite primary key*.

W11 Why is examining the data currently in a table not always a valid approach to identifying possible keys?

Relational Algebra

The questions in this section reference the following relations.

Note: The terms like **R.c** or **T.d** simply indicate the difference between a column **c** and an attribute in a row with the value “c.” We will not deduct points based on column names as long as your answer is clear and correct.

R

R.a	R.b	R.c
1	a	d
3	c	c
4	d	f
5	d	b
6	e	f

S

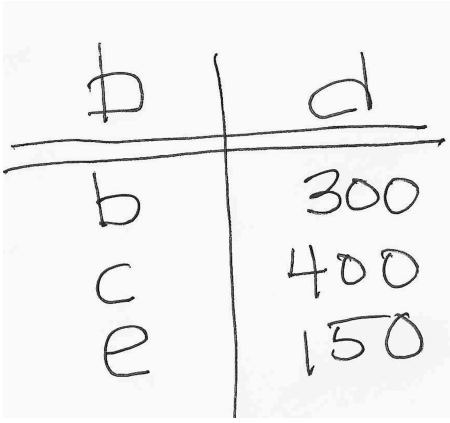
S.b	S.d
a	100
b	300
c	400
d	200
e	150

T

T.b	T.d
a	100
d	200
f	400
g	120

Note: You can hand draw your results. An example might be:

The result of S - T is



Hopefully, your handwriting and drawing skills are better than your professor’s

R1 What is the result of executing $(\pi\ a, b\ (R)) \bowtie (T \cup S)$

R2 What is the result of executing

$$\sigma\ a > 3\ (\ (\pi\ a,b\ (R) \cup \pi\ a,c\ (R))\ ) \rtimes S$$

R3 The relation below is the the result of R anti-join T (R ▷ T)

R.a	R.b	R.c
3	c	c
6	e	f

Write an equivalent relational expression that does not use the anti-join operator: ▷

R4 In relational algebra, the symmetric difference △ of two relations is the set of tuples that are in either relation but not in both. That is S △ T = (S - T) ∪ (T - S). The result of S △ T is

b	d
'b'	300
'c'	400
'e'	150
'f'	400
'g'	120

Write a relational algebra expression that that produces S △ T. You may only use the following relational operators: π, σ, ←, ⋈, ⋊, ⋉, =, ≠. Hint: You may have to use the θ-condition on your joins.

Entity Relationship Modeling

Consider a data model representing concepts from the Harry Potter series of books. The data model has the following entity types:

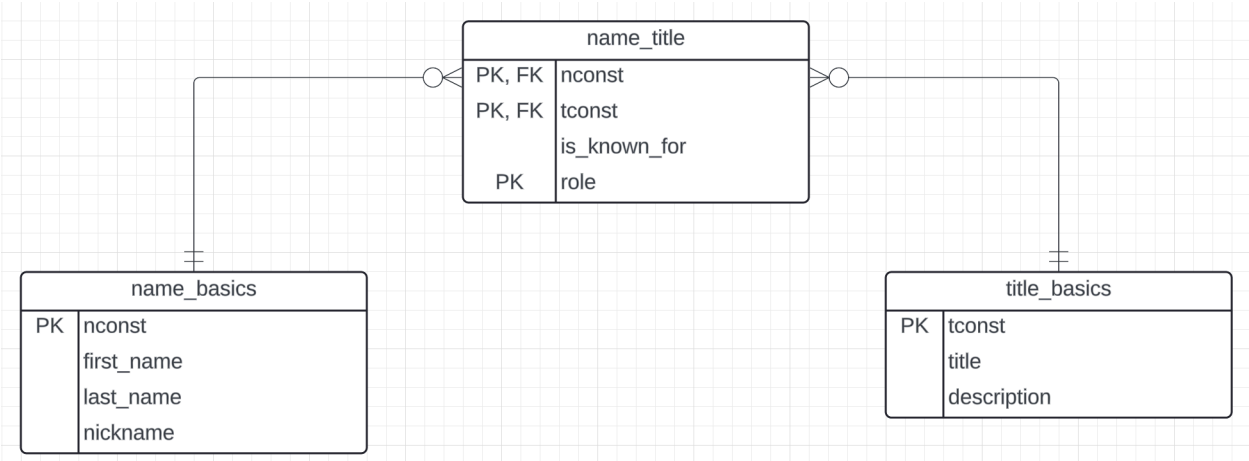
- **Character**:
    - ID: A unique ID number for the character.
    - Name: The name of the character.
- **Book**:
    - Book_No: The volume numbers, i.e. 1, 2, 3, 4, 5, 6, 7
    - Book_Title: The title of the book, e.g. “Harry Potter and the Philosopher's Stone,” “Harry Potter and the Chamber of Secrets.”
- **Chapter**:
    - Chapter_No: The chapter number within the volume.
    - Chapter_Title: The text of the chapter title. For example,
        - “The Boy Who Lived” is the 1st chapter of ““Harry Potter and the Philosopher's Stone.”
        - “At Flourish and Blotts” is the 4th chapter of “Harry Potter and the Chamber of Secrets.”
- **Location**:
    - ID: A unique ID for a location.
    - Name: The name of the location, e.g. “Hogwarts,” “The Forbidden Forest.”
- **Event**:
    - ID: A unique ID for the event.
    - Description: A short description of the event.

An specific **Event** occurs at *exactly one* location, in *exactly one* **Chapter** of *exactly one* **Book.** There *0, 1 or many* **Characters** participating in the event.

Draw a Crow's Foot Logical Model Diagram that realizes the description above. You may have to add attributes to entity types. You must show relationships and foreign keys.

Realizing an ER Model in SQL DDL

Write SQL CREATE TABLE statements that create the schema for the following data model.



Note:

- *is_know_for* is a boolean.
- *role* is one of *Actor, Director, Writer, Producer.*

- All columns except *nickname* are NOT NULL.

You do not need to worry about defining indexes.

SQL

The following are tables from the database associated with the textbook.

Takes

ID	course_i	sec_id	semeste	year	grad
128	CS-101	1	Fall	2017	A
128	CS-347	1	Fall	2017	A-
12345	CS-101	1	Fall	2017	C
12345	CS-190	2	Spring	2017	A
12345	CS-315	1	Spring	2018	A
12345	CS-347	1	Fall	2017	A
19991	HIS-351	1	Spring	2018	B
23121	FIN-201	1	Spring	2018	C+
44553	PHY-101	1	Fall	2017	B-
45678	CS-101	1	Fall	2017	F
45678	CS-101	1	Spring	2018	B+
45678	CS-319	1	Spring	2018	B
54321	CS-101	1	Fall	2017	A-
54321	CS-190	2	Spring	2017	B+
55739	MU-199	1	Spring	2018	A-
76543	CS-101	1	Fall	2017	A
76543	CS-319	2	Spring	2018	A
76653	EE-181	1	Spring	2017	C
98765	CS-101	1	Fall	2017	C-
98765	CS-315	1	Spring	2018	B
98988	BIO-101	1	Summer	2017	A
98988	BIO-301	1	Summer	2018	NULL

Department

dept_nam	building	budget
Biology	Watson	90000
Comp. Sci.	Taylor	100000
Elec. Eng.	Taylor	85000
Finance	Painter	120000
History	Painter	50000
Music	Packard	80000
Physics	Watson	70000

Course

course_ic	title	Dept_Nam	Credits
BIO-101	Intro. to Biology	Biology	4
BIO-301	Genetics	Biology	4
BIO-399	Computational Biology	Biology	3
CS-101	Intro. to Computer Science	Comp. Sci.	4
CS-190	Game Design	Comp. Sci.	4
CS-315	Robotics	Comp. Sci.	3
CS-319	Image Processing	Comp. Sci.	3
CS-347	Database System Concepts	Comp. Sci.	3
EE-181	Intro. to Digital Systems	Elec. Eng.	3
FIN-201	Investment Banking	Finance	3
HIS-351	World History	History	3
MU-199	Music Video Production	Music	3
PHY-101	Physical Principles	Physics	4

Student

ID	name	dept_nam	tot_cred
00128	Zhang	Comp. Sci.	102
12345	Shankar	Comp. Sci.	32
19991	Brandt	History	80
23121	Chavez	Finance	110
44553	Peltier	Physics	56
45678	Levy	Physics	46
54321	Williams	Comp. Sci.	54
55739	Sanchez	Music	38
70557	Snow	Physics	0
76543	Brown	Comp. Sci.	58
76653	Aoi	Elec. Eng.	60
98765	Bourikas	Elec. Eng.	98
98988	Tanaka	Biology	120

SQL1 Produce a query result of the form:

(student_ID, no_of_courses, computed_credits, recorded_credits)

Where:

- student_ID is the student's *ID* from *student.*
- *name* is the student's *name* from *student.*
- *no_of_courses* is the count of the courses the student took from *takes.*
- *compute_credits* is the sum of the credits for courses the student took.
- *recorded_credits* is *tot_cred* from the table *student.*

Please put your SQL below.

For the remaining questions, use the schema for *department* from the sample database associated with the textbook. The definition is:

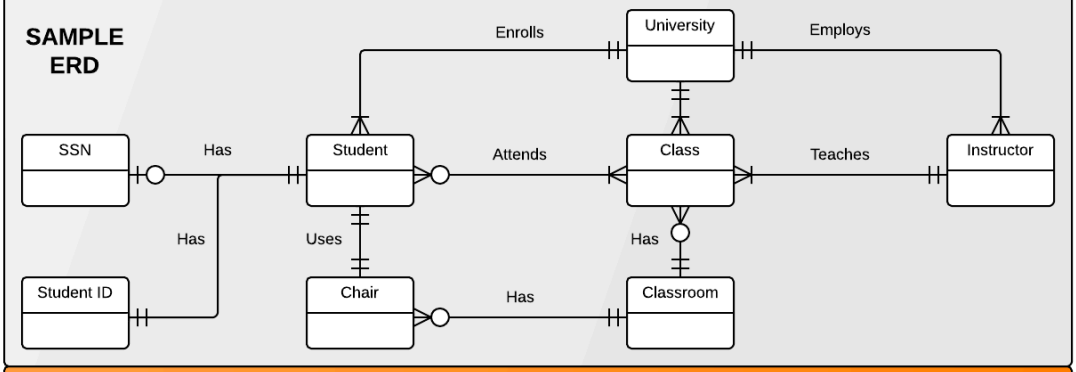
```
create table if not exists db_book.department
(
   dept_name varchar(20)    not null
       primary key,
   building  varchar(15)    null,
   budget    decimal(12, 2) null,
   check (`budget` > 0)
);
```

SQL2 Write a statement to add a row *('Chem. Eng.', 'Taylor', 150,000)* to the table.

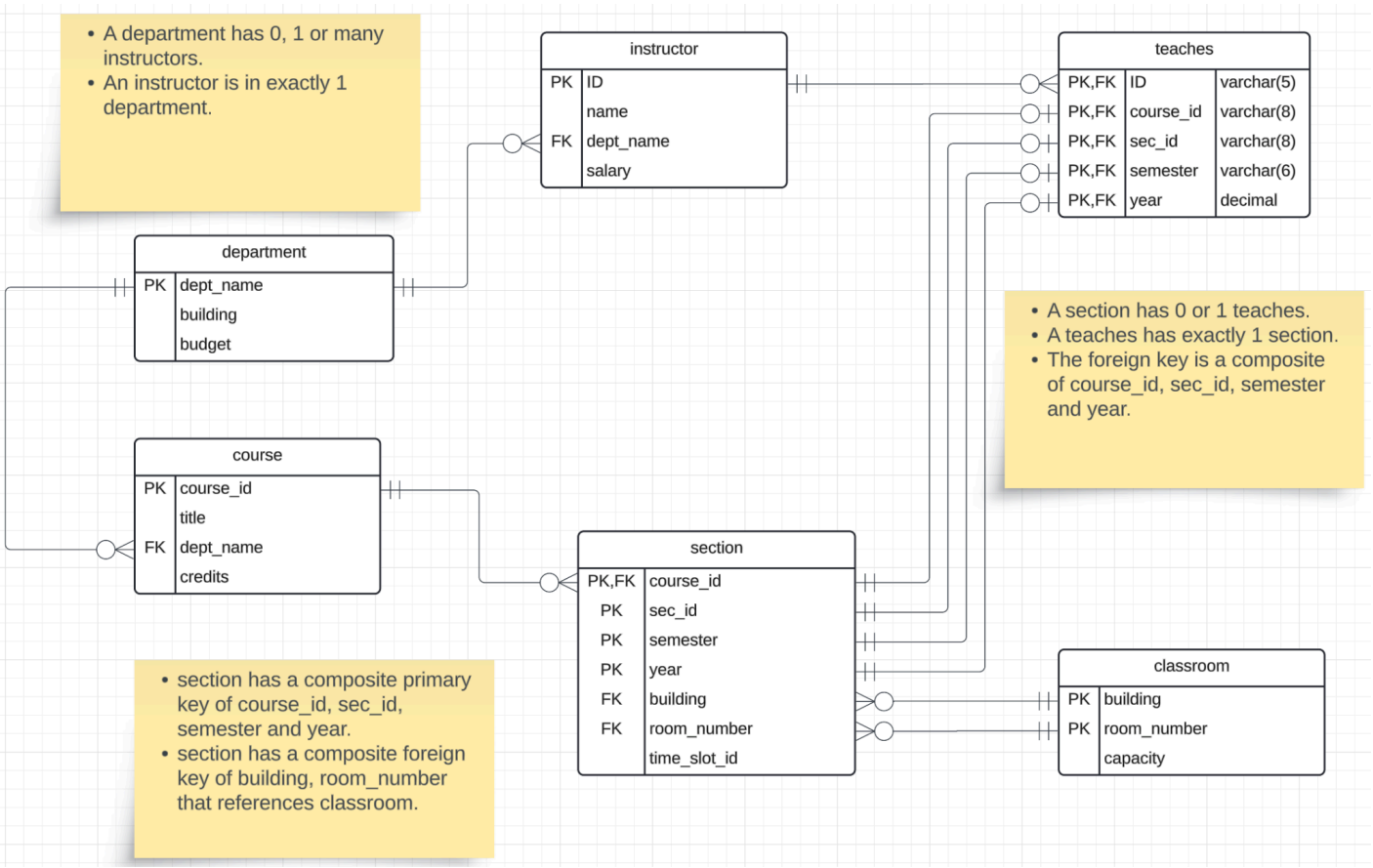
SQL3 Write a statement that changes the building name for a department in *Taylor* to *Northwest* except for the department ‘Comp. Sci.’

ERD "Crow's Foot" Relationship Symbols [Quick Reference]

Created by Vivek M. Chawla | @VivekMChawla | April 7, 2013



Notation	Meaning	Example
	Relationship	Student — Enrolls — University
	One	Student — Has — Student ID Number
	Many	Student — Attends — Class
	One and ONLY One	Student — Uses — Chair
	Zero or One	Student — Has — Social Security Number
	One or Many	Instructor — Teaches — Class
	Zero or Many	Classroom — Has — Chair


• A department has 0, 1 or many instructors.
• An instructor is in exactly 1 department.
instructor
PK ID
name
FK dept_name
salary
teaches
PK,FK ID varchar(5)
PK,FK course_id varchar(8)
PK,FK sec_id varchar(8)
PK,FK semester varchar(6)
PK,FK year decimal
department
PK dept_name
building
budget
• A section has 0 or 1 teaches.
• A teaches has exactly 1 section.
• The foreign key is a composite of course_id, sec_id, semester and year.
course
PK course_id
title
FK dept_name
credits
section
PK,FK course_id
PK sec_id
PK semester
PK year
FK building
FK room_number
time_slot_id
classroom
PK building
PK room_number
capacity
• section has a composite primary key of course_id, sec_id, semester and year.
• section has a composite foreign key of building, room_number that references classroom.

- τ order by
- γ group by
- $\neg$ negation
- $\div$ set division
- $\bowtie$ natural join, theta-join
- $\ltimes$ left outer join
- $\rtimes$ right outer join
- $\ltimes\rtimes$ full outer join
- $\ltimes$ left semi join
- $\rtimes$ right semi join
- $\triangleright$ anti-join

```
σ student_dept='Comp. Sci.' ∨ student_dept='Elec. Eng.'
(
  π student_id, student_name, student_dept,
    advisor_id, advisor_name, advisor_dept
  (
    (
      π student_id←ID, student_name←name,
        student_dept←dept_name, i_id
            (student ⋈ ID=s_id advisor)
    )
    ⋈ i_id=advisor_id
    (
      π advisor_id←ID, advisor_name←name, advisor_dept←dept_name
            (instructor ⋈ ID=i_id advisor)
    )
  )
)
```

Result

student_id	student_name	student_dept	advisor_id	advisor_name	advisor_dept
128	Zhang'	Comp. Sci.'	45565	Katz'	Comp. Sci.'
12345	Shankar'	Comp. Sci.'	10101	Srinivasan'	Comp. Sci.'
76543	Brown'	Comp. Sci.'	45565	Katz'	Comp. Sci.'
76653	Aoi'	Elec. Eng.'	98345	Kim'	Elec. Eng.'
98765	Bourikas'	Elec. Eng.'	98345	Kim'	Elec. Eng.'

SQL CHEAT SHEET http://www.sqltutorial.org

QUERYING DATA FROM A TABLE

```
SELECT c1, c2 FROM t;
```
Query data in columns c1, c2 from a table

```
SELECT * FROM t;
```
Query all rows and columns from a table

```
SELECT c1, c2 FROM t
WHERE condition;
```
Query data and filter rows with a condition

```
SELECT DISTINCT c1 FROM t
WHERE condition;
```
Query distinct rows from a table

```
SELECT c1, c2 FROM t
ORDER BY c1 ASC [DESC];
```
Sort the result set in ascending or descending order

```
SELECT c1, c2 FROM t
ORDER BY c1
LIMIT n OFFSET offset;
```
Skip *offset* of rows and return the next n rows

```
SELECT c1, aggregate(c2)
FROM t
GROUP BY c1;
```
Group rows using an aggregate function

```
SELECT c1, aggregate(c2)
FROM t
GROUP BY c1
HAVING condition;
```
Filter groups using HAVING clause

QUERYING FROM MULTIPLE TABLES

```
SELECT c1, c2
FROM t1
INNER JOIN t2 ON condition;
```
Inner join t1 and t2

```
SELECT c1, c2
FROM t1
LEFT JOIN t2 ON condition;
```
Left join t1 and t1

```
SELECT c1, c2
FROM t1
RIGHT JOIN t2 ON condition;
```
Right join t1 and t2

```
SELECT c1, c2
FROM t1
FULL OUTER JOIN t2 ON condition;
```
Perform full outer join

```
SELECT c1, c2
FROM t1
CROSS JOIN t2;
```
Produce a Cartesian product of rows in tables

```
SELECT c1, c2
FROM t1, t2;
```
Another way to perform cross join

```
SELECT c1, c2
FROM t1 A
INNER JOIN t2 B ON condition;
```
Join t1 to itself using INNER JOIN clause

USING SQL OPERATORS

```
SELECT c1, c2 FROM t1
UNION [ALL]
SELECT c1, c2 FROM t2;
```
Combine rows from two queries

```
SELECT c1, c2 FROM t1
INTERSECT
SELECT c1, c2 FROM t2;
```
Return the intersection of two queries

```
SELECT c1, c2 FROM t1
MINUS
SELECT c1, c2 FROM t2;
```
Subtract a result set from another result set

```
SELECT c1, c2 FROM t1
WHERE c1 [NOT] LIKE pattern;
```
Query rows using pattern matching %, _

```
SELECT c1, c2 FROM t
WHERE c1 [NOT] IN value_list;
```
Query rows in a list

```
SELECT c1, c2 FROM t
WHERE  c1 BETWEEN low AND high;
```
Query rows between two values

```
SELECT c1, c2 FROM t
WHERE  c1 IS [NOT] NULL;
```
Check if values in a table is NULL or not

SQL CHEAT SHEET http://www.sqltutorial.org

MANAGING TABLES

```
CREATE TABLE t (
    id INT PRIMARY KEY,
    name VARCHAR NOT NULL,
    price INT DEFAULT 0
);
```
Create a new table with three columns

```
DROP TABLE t ;
```
Delete the table from the database

```
ALTER TABLE t ADD column;
```
Add a new column to the table

```
ALTER TABLE t DROP COLUMN c ;
```
Drop column c from the table

```
ALTER TABLE t ADD constraint;
```
Add a constraint

```
ALTER TABLE t DROP constraint;
```
Drop a constraint

```
ALTER TABLE t1 RENAME TO t2;
```
Rename a table from t1 to t2

```
ALTER TABLE t1 RENAME c1 TO c2 ;
```
Rename column c1 to c2

```
TRUNCATE TABLE t;
```
Remove all data in a table

USING SQL CONSTRAINTS

```
CREATE TABLE t(
   c1 INT, c2 INT, c3 VARCHAR,
   PRIMARY KEY (c1,c2)
);
```
Set c1 and c2 as a primary key

```
CREATE TABLE t1(
   c1 INT PRIMARY KEY,
   c2 INT,
   FOREIGN KEY (c2) REFERENCES t2(c2)
);
```
Set c2 column as a foreign key

```
CREATE TABLE t(
   c1 INT, c1 INT,
   UNIQUE(c2,c3)
);
```
Make the values in c1 and c2 unique

```
CREATE TABLE t(
  c1 INT, c2 INT,
  CHECK(c1> 0 AND c1 >= c2)
);
```
Ensure c1 > 0 and values in c1 >= c2

```
CREATE TABLE t(
    c1 INT PRIMARY KEY,
    c2 VARCHAR NOT NULL
);
```
Set values in c2 column not NULL

MODIFYING DATA

```
INSERT INTO t(column_list)
VALUES(value_list);
```
Insert one row into a table

```
INSERT INTO t(column_list)
VALUES (value_list),
       (value_list), ....;
```
Insert multiple rows into a table

```
INSERT INTO t1(column_list)
SELECT column_list
FROM t2;
```
Insert rows from t2 into t1

```
UPDATE t
SET c1 = new_value;
```
Update new value in the column c1 for all rows

```
UPDATE t
SET c1 = new_value,
    c2 = new_value
WHERE condition;
```
Update values in the column c1, c2 that match the condition

```
DELETE FROM t;
```
Delete all data in a table

```
DELETE FROM t
WHERE condition;
```
Delete subset of rows in a table

SQL CHEAT SHEET http://www.sqltutorial.org

MANAGING VIEWS

```
CREATE VIEW v(c1,c2)
AS
SELECT c1, c2
FROM t;
```
Create a new view that consists of c1 and c2

```
CREATE VIEW v(c1,c2)
AS
SELECT c1, c2
FROM t;
WITH [CASCADED | LOCAL] CHECK OPTION;
```
Create a new view with check option

```
CREATE RECURSIVE VIEW v
AS
select-statement -- anchor part
UNION [ALL]
select-statement; -- recursive part
```
Create a recursive view

```
CREATE TEMPORARY VIEW v
AS
SELECT c1, c2
FROM t;
```
Create a temporary view

```
DROP VIEW view_name;
```
Delete a view

MANAGING INDEXES

```
CREATE INDEX idx_name
ON t(c1,c2);
```
Create an index on c1 and c2 of the table t

```
CREATE UNIQUE INDEX idx_name
ON t(c3,c4);
```
Create a unique index on c3, c4 of the table t

```
DROP INDEX idx_name;
```
Drop an index

SQL AGGREGATE FUNCTIONS

AVG returns the average of a list

COUNT returns the number of elements of a list

SUM returns the total of a list

MAX returns the maximum value in a list

MIN returns the minimum value in a list

MANAGING TRIGGERS

```
CREATE OR MODIFY TRIGGER trigger_name
WHEN EVENT
ON table_name TRIGGER_TYPE
EXECUTE stored_procedure;
```
Create or modify a trigger

WHEN
- **BEFORE** – invoke before the event occurs
- **AFTER** – invoke after the event occurs

EVENT
- **INSERT** – invoke for INSERT
- **UPDATE** – invoke for UPDATE
- **DELETE** – invoke for DELETE

TRIGGER_TYPE
- **FOR EACH ROW**
- **FOR EACH STATEMENT**

```
CREATE TRIGGER before_insert_person
BEFORE INSERT
ON person FOR EACH ROW
EXECUTE stored_procedure;
```
Create a trigger invoked before a new row is inserted into the person table

```
DROP TRIGGER trigger_name;
```
Delete a specific trigger

> Definitions used throughout this cheat sheet

Primary key:
A primary key is a field in a table that uniquely identifies each record in the table. In relational databases, primary keys can be used as fields to join tables on.

Foreign key:
A foreign key is a field in a table which references the primary key of another table. In a relational database, one way to join two tables is by connecting the foreign key from one table to the primary key of another.

One-to-one relationship:
Database relationships describe the relationships between records in different tables. When a one-to-one relationship exists between two tables, a given record in one table is uniquely related to exactly one record in the other table.

One-to-many relationship:
In a one-to-many relationship, a record in one table can be related to one or more records in a second table. However, a given record in the second table will only be related to one record in the first table.

Many-to-many relationship:
In a many-to-many relationship, records in a given table 'A' can be related to one or more records in another table 'B', and records in table B can also be related to many records in table A.

> Sample Data

Artist Table

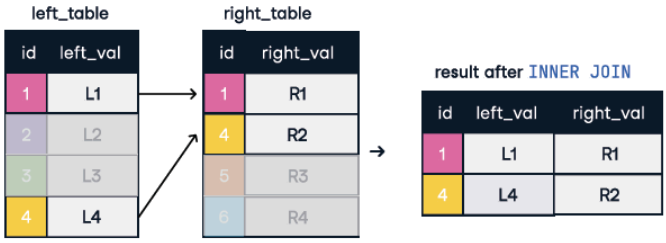
artist_id	name
1	AC/DC
2	Aerosmith
3	Alanis Morissette

Album Table

album_id	title	artist_id
1	For those who rock	1
2	Dream on	2
3	Restless and wild	2
4	Let there be rock	1
5	Rumours	6

INNER JOIN

An inner join between two tables will return only records where a joining field, such as a key, finds a match in both tables.



INNER JOIN join ON one field

```
SELECT *
FROM artist AS art
INNER JOIN album AS alb
ON art.artist_id = alb.artist_id;
```

Result after INNER JOIN:

artist_id	name	title	album_id
1	AC/DC	For those who rock	1
1	AC/DC	Let there be rock	4
2	Aerosmith	Dream on	2
2	Aerosmith	Restless and wild	3

INNER JOIN with USING

```
SELECT *
FROM artist AS art
INNER JOIN album AS alb
USING (artist_id);
```

SELF JOIN

Self-joins are used to compare values in a table to other values of the same table by joining different parts of a table together.

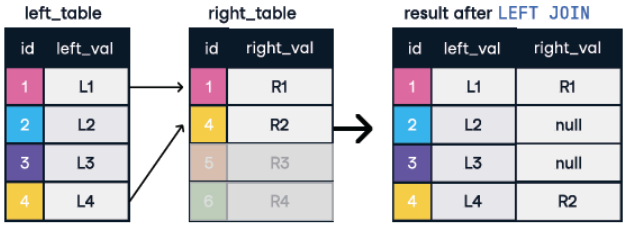
```
SELECT
  alb1.artist_id,
  alb1.title AS alb1_title,
  alb2.title AS alb2_title
FROM album AS alb1
INNER JOIN album AS alb2
ON art1.artist_id = art2.artist_id
WHERE alb1.album_id<>alb2.album_id;
```

Result after Self join:

artist_id	name	album_id	alb2_title
1	AC/DC	1	For those who rock
2	Aerosmith	2	Dream on
2	Aerosmith	3	Restless and wild
1	AC/DC	4	Let there be rock

LEFT JOIN

A left join keeps all of the original records in the left table and returns missing values for any columns from the right table where the joining field did not find a match.



LEFT JOIN on one field

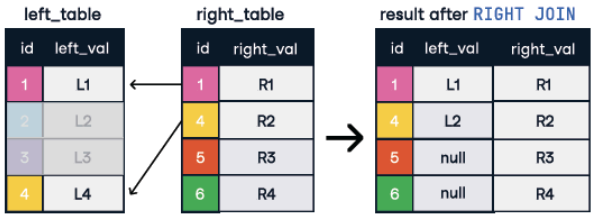
```
SELECT *
FROM artist AS art
LEFT JOIN album AS alb
ON art.artist_id = alb.artist_id;
```

Result after LEFT JOIN:

artist_id	name	album_id	title	name
1	AC/DC	1	For those who rock	1
1	AC/DC	4	Let there be rock	1
2	Aerosmith	2	Dream on	2
2	Aerosmith	3	Restless and wild	2
3	Alanis Morissette	null	null	null

RIGHT JOIN

A right join keeps all of the original records in the right table and returns missing values for any columns from the left table where the joining field did not find a match. Right joins are far less common than left joins, because right joins can always be re-written as left joins.



RIGHT JOIN on one field

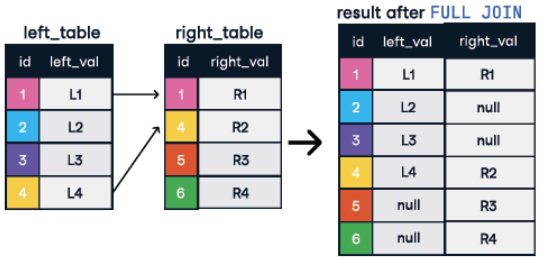
```
SELECT *
FROM artist as art
RIGHT JOIN album AS alb
ON art.artist_id = alb.artist_id;
```

Result after RIGHT JOIN:

artist_id	name	album_id	title	name
1	AC/DC	1	For those who rock	1
1	Aerosmith	2	Dream on	2
2	Aerosmith	3	Restless and wild	2
2	AC/DC	4	Let there be rock	1
3	null	5	Rumours	6

FULL JOIN

A full join combines a left join and right join. A full join will return all records from a table, irrespective of whether there is a match on the joining field in the other table, returning null values accordingly.



FULL JOIN on one field

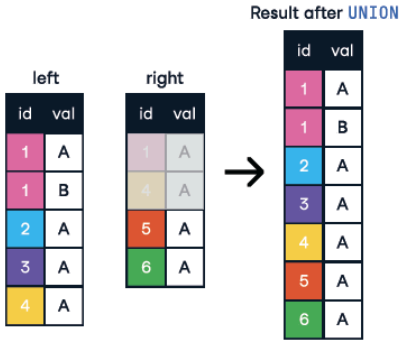
```
SELECT *
FROM artist as art
FULL OUTER JOIN album AS alb
ON art.artist_id = alb.artist_id;
```

Result after FULL JOIN:

artist_id	name	album_id	title	name
1	AC/DC	1	For those who rock	1
1	AC/DC	4	Let there be rock	1
2	Aerosmith	2	Balls to the wall	2
2	Aerosmith	3	Restless and wild	2
3	Alanis Morissette	null	null	null
null	null	5	Rumours	6

UNION

The `UNION` operator is used to vertically combine the results of two `SELECT` statements. For `UNION` to work without errors, all `SELECT` statements must have the same number of columns and corresponding columns must have the same data type. `UNION` does not return duplicates.



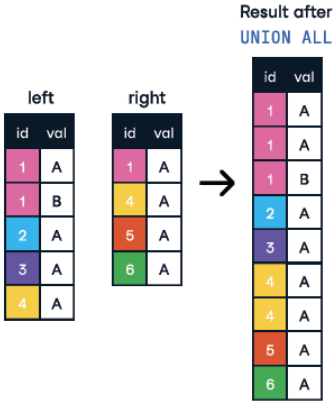

```
SELECT artist_id
FROM artist
UNION
SELECT artist_id
FROM album;
```

Result after UNION:

artist_id
1
2
3
6

UNION ALL

The `UNION ALL` operator works just like `UNION`, but it returns duplicate values. The same restrictions of `UNION` hold true for `UNION ALL`




```
SELECT artist_id
FROM artist
UNION ALL
SELECT artist_id
FROM album;
```

Result after UNION ALL:

artist_id
1
2
3
1
2
2
1
6