Lecture 3 & Lab 3

update to tables

1. Insert statements

```
-- should specify all values
insert into {tab1} values ({val1}, {val2}, {val3})

-- furthermore, values match columns one by one
-- other column will be inserted NULL automatically
-- It may occur error for mandatory columns
insert into {tab1}({col1}, {col2}) values({val1}, {val2})

-- if we want to insert a string with a single quote
insert into student(name) values('Bob''t'); -- Bob't here
```

2. delete statements

```
-- remove all the tuples in the tab1
delete from {tab1}
```

Query -- select

1. format

```
select {col1}, {col2}, {col3}... from {tab1}, {tab2}, {tab3}
where {conditions}
```

colomn name don't need `' '`

2. query for all columns

```
select * from {tab1}
```

Remark: This query is frequently used when you don't remember column names. But it would flood the database if the data is large. (In the application programs)

select with restrictions

1. Reason:

   when you are instrested in only a small subset or only want to return some of the rows.

2. Filtering(rows)

    1. overview

        1. perform in the `where` clause

        2. conditions are usually expressed by a column name

3. only rows for which the condition is true will be returned
4. if you want to filter the columns, just filter it behind the "select"

2. example

```
select * from movies where country = 'us'
```

3. what you compare

```
a number
a string constant: must be quoted between single-quotes
another column(same table or another)
result of a function
```

3. select without From or Where
    1. property

       never alter any table

       might be a good choice to figure out an expression is true or false

    2. examples

```
select '437'
-- just return return a cell that contains '437'
-- quite useless
-- may be use in ensure whether an expression is true
select '437' as FOO
-- generate a column name "FOO"
-- do not alter any table
select 'abcd' from {tab1}
-- generate 1 column and n rows with 'abcd' in the cells
-- even if tab1 hasn't such cells
```

Arithmetic Expression

1. **as** clause

   modified the **display** name:

   just show with the column name you want, **never alter the table itself**

```
select {col} as {col_newname} from {tab}
```

2. `and, or`
    1. functions like other programme langugue
    2. precedence: and > or
    3. we can use parentheses to enforce the precedence

```
-- these 2 statement are different
select * from {tab}
where ({cond1} or {cond2}) and ({cond3} or {cond4})

select * from {tab}
where {cond1} or {cond2} and {cond3} or {cond4}
```

3. comparison operators
    1. basic comparison operators

    ```
    <, <=, >, >=, =, !=, <>, =
    ```

    2. `!=, <>`

        Both means not equal to

        Whether they are equivalent depends on the SQL. In postgreSQL, they are equivalent.

    3. `=`

        means **equals**

        never use `==` as **equals** in postgreSQL

    4. **NULL** judgement

    ```
    -- never use {col1} = null as a filter
    where {col1} is null;
    where {col2} is not null
    ```

    5. remark on **bigger** and **smaller**

        means different for different **data types**

    ```
    2 < 10 --true
    '2' < '10' -- false, ASCII + lexicographical order
    '2-JUN-1883' > '1-DEC-2056' -- comapre as strings or dates? It differs in different products
    ```

4. `in()`

    It can be used as the restriction for some columns are discrete range

    a set is in the parentheses

    ```
    where (country = 'us' or country = 'cn')
    where country in('us', 'cn')
    ```

5. intervals
    1. format:

    ```
    where {col1} between {lowerbound} and {upperbound}
    ```

2. remark

   `between {a} and {b}` indicates an interval `[a, b]`

3. example

```
-- query all 5-10 years old kids
where age between 5 and 10
-- query all kids whose name with a first character 'a' or 'b'
where name between 'A' and 'C' and name not like 'c%'
```

6. negation

   1. claim: all comparison can be negated with `not`

   2. examples

```
where country not in ('us', 'gb') or year not between 1940 and 1949;
-- by De Morgan's law, it's equivalent to
where not (country in ('us', 'gb') an year between 1940 and 1949);
```

7. match with **like**

   1. With the usage of **like**, we can restrict a column of string to a specifi patten

   2. wildcard

```
% -- meaning any number of characters, including none

_ -- one and only one character
```

   3. skills

```
-- 1. we don't want any string in the {col} with a character 'A'
where {col} not like '%A%'

-- 2. we don;t want any string in the {col} with 'a' or 'A'
where {col} not like '%A%' and not like '%a%'
where upper({col}) not like '%A%' -- Function called. Slow down the
query
```

8. Date and Datetime

   1. explicit cast is needed to avoid bad surprise when **comparing**

```
where post_date >= '2018-03-12'; -- bad habit
where post_date >= date('2018-03-12'); -- ok
```

   2. More than one patten of string can be converted to same date

```
'YYYY-MM-DD'
'YYYY/MM/DD'
'Mon DD, YYYY'
'Month DD, YYYY'
date('2018-3-12')
date('2018/3/12')
date('Mar 12, 2018')
date('March 12, 2018')
```

3. better way to convert a string to **date** or **timestamp**

    0. Reference:

    > PostgreSQL: Documentation: 12: 9.8. Data Type Formatting Functions

    1. syntax

Function	Return Type	Description	Example
`to_date(text, text)`	`date`	convert string to date	`to_date('05 Dec 2000', 'DD Mon YYYY')`
`to_timestamp(text, text)`	`timestamp with time zone`	convert string to time stamp	`to_timestamp('05 Dec 2000', 'DD Mon YYYY')`

    2. Common format specifiers

    date:

Pattern	Description
`YYYY`	year (4 or more digits)
`Month`	full capitalized month name (blank-padded to 9 chars)
`Mon`	abbreviated capitalized month name (3 chars in English, localized lengths vary)
`MM`	month number (01-12)
`DD`	day of month (01-31)

    timestamp:

Pattern	Description
`HH12`	hour of day (01-12)
`HH24`	hour of day (00-23)
`MI`	minute (00-59)

Pattern	Description
`SS`	second (00-59)
`MS`	millisecond (000-999)

3. More syntax and format supported by postgreSQL

   Reference:

   > PostgreSQL: Documentation: 12: 9.8. Data Type Formatting Functions

4. Comparing **date** and **timestamp**

   A **date** will be automatically converted to **timestamp**, which has higher precision.

   Its **time** will be automatically set to **'00:00:00'**

9. NULL

   1. Boolean in sql can be **true**, **false**, **NULL**.
   2. In SQL, logically, we say NULL means "we don't know whether true or false".
   3. When a NULL is involved in a logicall expression, we are hard to deduce the outcome.
   4. there are some special cases:

      ```
      select true or NULL; -- true
      select false and NULL; -- false
      ```

   5. to filter the **NULLs**, we should use `is NULL` or `is not NULL`;

Functions

1. show DDL of a table: in command line

   ```
   \d movies
   ```

2. Why using functions

   One important feature of SQL is that you don't need to return data **exactly as it was stored**.

   Operators, and many functions allow to return transformed data

   Transforming displayed colomn in Select

   It reduce the performance and sometimes harmful.

   Never store something that can be computed from the original data at any time.

3. concatenating 2 strings

   1. claims

      `concat()` in MySQL, SQL server use `+`

      In most products, including postgreSQL, use `||`

some other types , when becoming the operands of the `||`, will automatically be converted to the string

2. syntax:

```
select [{col1} | {str1}] || [{col2} | {str2}] || [{col3} | {str3}]
from {tab1}
```

3. example:

```
select 'The code of '||c.name||' is '|| c.hex
from color_names;
```

4. displayed colomns

?column?
The code of AliceBlue is #F0F8FF
The code of AntiqueWhite is #FAEBD7
The code of Aqua is #00FFFF
The code of Aquamarine is #7FFFD4

4. **as**

1. claim

   give a **align** to a column and show in the console

2. example

```
select 'The code of ' || c.name || ' is ' || c.hex as 'combine'
from color_names;
```

combine
The code of AliceBlue is #F0F8FF
The code of AntiqueWhite is #FAEBD7
The code of Aqua is #00FFFF
The code of Aquamarine is #7FFFD4

5. type casting

```
select cast({col1} as {type1}) from {tab1}
```

cast the displayed column to {type1}

6. case-end

1. aim: **maps** the **current column cells** to a **enumerate set** and generate a new column **in the console**
2. syntax1: for multi-columns condition

```
case when {cond1} then {result1}
     when {cond2} then {result2}
     ......
     else {resultn}
     end as {new_name}
from [table]
where {conditions}
-- the results will automatically form a column
```

```
-- example
-- two column will be shown in the console
select s.english_name,
case
 when s.latitude is null then 'missing'
 when s.district in ('Luohu','Futian','Nanshan') then 'Inside Area'
 else 'Outside Area' end as area
from stations s;
```

syntax2: for single-column condition

```
case {col}
 when {form1} then {result1}
     when {form2} then {result2}
     ......
     else {resultn}
     end as {new_name}
from [table]
where {conditions}
-- use '=' to match the col and form1, form2...
-- the results will automatically form a column
```

```
-- example
-- only one column will be shown in the console
select '=' ||
case s.district
    when 'Luohu' then 1
    when 'Futian' then 1
    when 'Nanshan' then 1
else 0 end as Inside_or_not
from stations s
```

3. example: we can use it to extend the imformation

```
select Name_
 case gend
     when 'F' then 'female'
     when 'M' then 'male'
     else 'unknown' -- we should deal with the NULL
    end as 'gender'
from people
```

7. String Functions

upper and lower

```
-- convert a column or a string to uppercase or lowercase
upper([{col} | {str}])
lower([{col} | {str}])
```

trim

```
-- delete some sybolic strings in the head or in the tail
select trim([both | leading | trailing] {str_pattern} from [{col},
{str_text}]) from {tab1};
-- 'from {tab1}' can be omitted when no columns appear behind the SELECT

--example
selecet trim(trailing '.' from '10.00.....'); -- 10.00
```

substr

```
-- from the begin_ position intercept string with a length of span_
-- the first character in the string is 1, not 0
substr([{col} | {str_text}], {begin_}, {span_})

--example
select substr('hello world', 5, 3); -- 'o w'
```

replace

```
-- replace all the str_pattern with new_pattern in the text
replace([{col} | {text}], {str_pattern}, {new_pattern})

--example
select replace('sheep', 'ee', 'i');
```

length

```
-- compute the number of bytes of a string
select length('hello'); -- 5
select length('你好');  -- 2 in postgreSQL, depends on the product
```

8. Numerical functions

   round and trunc

```
-- Round to retain specified significant figures
select round(50.123, 2); -- 50.12
select round(50.125, 2); -- 50.13

-- truncates a value to the specified precision
select trunc(123.456, 2); -- 123.45
select trunc(123.456, 0); -- 123
select trunc(123.456, -2); -- 100
```