

### Structure keywords

|                         |  |
|-------------------------|--|
| <b>from</b>             | Point to a table for the query. Specify a table name using tags.<br><input checked="" type="checkbox"/> <code>from my.app.web.auth</code>  |
| <b>where</b>            | Where clause to filter results.<br><input type="checkbox"/> <input checked="" type="checkbox"/> <code>where [filter1 expression], [filter2 expression]</code><br><input type="checkbox"/> <input checked="" type="checkbox"/> String values in expressions have to be surrounded by double quotes, single quotes are not allowed.  |
| <b>select</b>           | Add column to result set:<br><input checked="" type="checkbox"/> <code>select source_column or column operation as destination_column</code><br><input checked="" type="checkbox"/> <code>select uriHost(uri) as host</code>   |
| <b>group every / by</b> | Group clause with an optional server and client aggregation period filter:<br><input checked="" type="checkbox"/> <code>group [every server_period] by column</code><br><code>[every client_period]</code><br><input checked="" type="checkbox"/> <code>group by statusCode</code><br><input checked="" type="checkbox"/> <code>group every 10m</code><br><input checked="" type="checkbox"/> <code>group every 10m by statusCode every 1m</code>  |
| <b>every</b>            | Client aggregation filter:<br><input checked="" type="checkbox"/> <code>every period</code><br><input type="checkbox"/> <input checked="" type="checkbox"/> Period syntax: an integer number follow by a symbol indicating the time period: <code>[s:seconds, m:minutes, h:hours, d:days, no suffix :milliseconds - seconds]</code><br><input type="checkbox"/> <input checked="" type="checkbox"/> <code>every 0</code> means no client period.<br><input type="checkbox"/> <input checked="" type="checkbox"/> <code>every 5m by serverIp</code> |
| <b>ifthenelse</b>       | <code>if/then/else</code> equivalent clause to set conditionally column values:<br><input type="checkbox"/> <input checked="" type="checkbox"/> <code>ifthenelse(condition, error_value, success_value)</code><br><input type="checkbox"/> <input checked="" type="checkbox"/> <code>select ifthenelse(statusCode != 200, "Error", "Success") as statusCodeDesc</code>   |

### Structure keywords (cont)

|               |   |
|---------------|---|
| <b>decode</b> | <code>switch/case</code> equivalent clause to set conditionally column values:<br><input type="checkbox"/> <input checked="" type="checkbox"/> <code>decode(column, check_value, value, [check_value2, value2])</code><br><input type="checkbox"/> <input checked="" type="checkbox"/> Each pair of arguments <code>[check_value, value]</code> is equivalent to a case sentence of a switch statement.<br><input type="checkbox"/> <input checked="" type="checkbox"/> <code>decode(statusCode, 200, "Success", 400, "Not Found", 406, "Error", 404, "Error") as statusCodeDesc</code> |
| <b>nvl</b>    | <b>Null-Coalescing</b> operator. Allow to set an alternate value when input value is null.<br><input checked="" type="checkbox"/> <code>nvl(column, alternate_value_when_null)</code>   |

### Aggregation functions and operators

|                        |  |
|------------------------|--|
| <b>avg</b>             | Returns the <b>average</b> of a range of values <code>[avg]</code> or only over <i>not null</i> values <code>[nnavg]</code> of the results on each group:<br><input type="checkbox"/> <input checked="" type="checkbox"/> <code>avg(column); nnavg(column)</code>          |
| <b>count</b>           | Returns the <b>count</b> of results on each group:<br><input type="checkbox"/> <input checked="" type="checkbox"/> <code>count([column])</code><br><input type="checkbox"/> <input checked="" type="checkbox"/> With argument, include only not null entries in the count. |
| <b>first / nnfirst</b> | Returns the <b>first</b> or the <b>not null first</b> entry of the results on each group:<br><input type="checkbox"/> <input checked="" type="checkbox"/> <code>first(column); nnfirst(column)</code>  |



By **Rafa Hernández (elpluto)**  
[cheatography.com/elpluto/](https://cheatography.com/elpluto/)

Published 2nd January, 2023.  
Last updated 2nd January, 2023.  
Page 1 of 8.

Sponsored by **Readable.com**  
Measure your website readability!  
<https://readable.com>

### Aggregation functions and operators (cont)

|   |   |
|---|---|
| <b>last / nlast</b>   | Returns the <b>last</b> or the <b>not null last</b> entry of the results on each group:<br><input type="checkbox"/> <input checked="" type="checkbox"/> <code>last(column); nlast(column)</code>  |
| <b>max / min</b>  | Returns the <b>maximum</b> or the <b>minimum</b> value for the columns provided, on each group:<br><input type="checkbox"/> <input checked="" type="checkbox"/> <code>max/min(col1, [col2], [col3]...)</code>   |
| <b>median</b>   | Returns the statistical <b>median</b> for a column on each group:<br><input type="checkbox"/> <input checked="" type="checkbox"/> <code>median(column)</code><br><input type="checkbox"/> <input checked="" type="radio"/> Restricted to columns of integer type  |
| <b>sum</b>  | Returns the <b>sum</b> of the results on each group:<br><input type="checkbox"/> <input checked="" type="checkbox"/> <code>sum(column)</code>   |
| <b>sum2</b>   | Returns the <b>sum</b> of the squares of the results on each group:<br><input type="checkbox"/> <input checked="" type="checkbox"/> <code>sum2(column)</code>   |
| <b>percentile5</b><br><b>percentile10</b><br><b>percentile25</b><br><b>percentile75</b><br><b>percentile90</b><br><b>percentile95</b> | Returns the specific statistic <b>percentileN</b> , using linear interpolation, of the results on each group:<br><input type="checkbox"/> <input checked="" type="checkbox"/> <code>percentile[N](column)</code>  |
| <b>stddev / nnstddev</b>  | Returns the <b>biased standard deviation [stddev]</b> of the values or not null values <b>[nnstddev]</b> of the results on each group:<br><input type="checkbox"/> <input checked="" type="checkbox"/> <code>stddev / nnstddev(column)</code><br><input type="checkbox"/> <input checked="" type="radio"/> Biased |

### Aggregation functions and operators (cont)

|                            |   |
|----------------------------|---|
| <b>ustddev / nnustddev</b> | Returns the <b>unbiased standard deviation [ustddev]</b> of the values or not null values <b>[nnustddev]</b> of the results on each group:<br><input type="checkbox"/> <input checked="" type="checkbox"/> <code>ustddev / nnustddev(column)</code><br><input type="checkbox"/> <input checked="" type="radio"/> Unbiased         |
| <b>var / nnvar</b>         | Returns the <b>biased variance [var]</b> of the values or not null values <b>[nnvar]</b> of the results on each group:<br><input type="checkbox"/> <input checked="" type="checkbox"/> <code>var/nnvar(column)</code><br><input type="checkbox"/> <input checked="" type="radio"/> Biased   |
| <b>uvar / nnuvar</b>       | Returns the <b>unbiased variance [uvar]</b> of the values or not null values <b>[nnuvar]</b> of the results on each group:<br><input type="checkbox"/> <input checked="" type="checkbox"/> <code>uvar/nnuvar(column)</code><br><input type="checkbox"/> <input checked="" type="radio"/> Unbiased                                 |
| <b>hllpp</b>               | Returns the estimated count of distinct values of the results on each group using the <b>HyperLogLog++</b> algorithm:<br><input type="checkbox"/> <input checked="" type="checkbox"/> <code>hllpp(column)</code><br><input type="checkbox"/> <input checked="" type="radio"/> Applies on <b>DC (distinct count)</b> data types.   |
| <b>hllppcount</b>          | Returns the estimated count of distinct values of the results on each group using the <b>HyperLogLog++</b> algorithm:<br><input type="checkbox"/> <input checked="" type="checkbox"/> <code>hllppcount(column)</code><br><input type="checkbox"/> <input checked="" type="radio"/> Applies on <b>float or integer</b> data types. |



By [Rafa Hernández \(elpluto\)](https://cheatography.com/elpluto/)  
[cheatography.com/elpluto/](https://cheatography.com/elpluto/)

Published 2nd January, 2023.  
Last updated 2nd January, 2023.  
Page 2 of 8.

Sponsored by [Readable.com](https://readable.com)  
Measure your website readability!  
<https://readable.com>

### String operators and functions

**has, [->]** Case sensitive **contains** comparison. Using the operator  
-> only allows check one value:

- `has (column, value1, [value2],...)`
- `column -> value1`

**weakhas** Case insensitive **contains** comparison:  
`weakhas (column, value)`

**in, [-<]** Case sensitive **is contained** comparison. Using the operator '<' allows only one value:

- `in (value1, [value 2], [...], column)`
- `value1 <- column`

**weakin** Case insensitive **is contained** comparison:  
`weakin (value, column)`

**startswith** Returns strings that start with specific value:  
`startswith (column, value)`

**endswith** Returns strings that end with a specific value:  
`endswith (column, value)`

**toktains** Specialized **contains** function for ASCII delimited tokens:  
`toktains (column, value, [bool_ left], [bool_ right])`

**length** Returns the length of a string value:  
`length (column)`

**locate** Returns the position of a substring, **indexOf** function:  
`locate (column, substring_to Locate)`

**lower** Returns the transformation to lower case:  
`lower (column)`

**upper** Returns the transformation to upper case:  
`upper (column)`

### String operators and functions (cont)

**replace** Replaces **only first occurrence** of a string with a substitute string:  
`replace (column, string ToS earch, string ToR eplace)`

**replaceall** Replaces **all occurrences** of a search string with a substitute string:  
`replaceall (column, string ToS earch, string ToR eplace)`

**split** Returns a specific piece of splitting operation by a separator:  
 pieceN umber begin at 0.  
`split (column, separator String, pieceN umber)`

**splitre** Returns a specific piece of splitting operation by a regular expression:  
 pieceN umber begin at 0.  
`splitre (column, re(string) or regexp, pieceN umber)`

**substring** Returns a substring beginning at specific index with the provided length:  
`substring (column, index, length)`

**subs** Returns a string replacing **first substring occurrence** based on a regular expression using a template string as substitution value:  
 FailValue is returned when is provided and no occurrences found  
 `subs (column, regexp, template, [failValue])`  
 `subs (column, re(string), template (string), [failV alue])`



By **Rafa Hernández (elpluto)**  
[cheatography.com/elpluto/](https://cheatography.com/elpluto/)

Published 2nd January, 2023.  
Last updated 2nd January, 2023.  
Page 3 of 8.

Sponsored by **Readable.com**  
Measure your website readability!  
<https://readable.com>

### String operators and functions (cont)

**subsall** Returns a string replacing **all substring occurrences** based on a regular expression using a template string as substitution value:

- 👁 FailValue is returned when is provided and no occurrences found
- ✔ `subs(column, regexp, template, [failValue])`
- ✔ `subs(column, re(string), template(string), [failValue])`

**trim** Returns the result of trimming **both sides**:  
`trim(column)`

**ltrim** Returns the result of trimming **left side**:  
`ltrim(column)`

**rtrim** Returns the result of trimming **right side**:  
`rtrim(column)`

**matches, [-]** Matches function that finds occurrences in a column using a regular expression:

- ✔ `matches(column, re(string) or regexp value)`
- ✔ `column ~ re(string) or regexp value`

**peek** Returns the part of a string based on a regular expression, optionally indicating a specific part occurrence:

- 👁 If no partNumber is provided then returns first part occurrence.

`peek(column, re(string) or regexp, [partNumber])`

**formatnumber** Format a number with a specific mask and locale:  
`format number (numberColumn, mask, locale)`

- 👁 `format number (total Amount, "###.##", "en-GB")`

### String operators and functions (cont)

**damerau** Returns **Damerau** distance:  
`damerau(column, value)`

**hamming** Returns **Hamming** distance:  
`hamming(column, value)`

**levenshtein** Returns **Levenstein** distance:  
`levenshtein(column, value)`

**osa** Returns **osa** distance:  
`osa(column, value)`

**publicsuffix** Returns the main public suffix of a hostname:  
`publicsuffix(hostnameColumn)`

- 👁 `'www.my.site.co.uk' = 'co.uk'`

**rootdomain** Returns the root domain of a hostname part of an url:  
`rootdomain(hostnameUrlColumn)`

- 👁 `'www.my.site.com' = 'site'`

**rootprefix** Returns the root prefix of a hostname part of an url:  
`rootprefix(hostnameUrlColumn)`

- 👁 `'www.my.site.com' = 'www.my.site'`

**rootsuffix** Returns the root suffix of a hostname part of an url:  
`rootsuffix(hostnameUrlColumn)`

- 👁 `'www.my.site.com' = 'my.site.com'`

**subdomain** Returns the subdomain of a hostname part of an url:  
`subdomain(hostnameUrlColumn)`

- 👁 `'www.my.site.com' = 'www'`

**topleveldomain** Returns the top level domain of a hostname part of an url:  
`topleveldomain(hostnameUrlColumn)`

- 👁 `'www.my.site.co.uk' = 'uk'`

### shannonentropy



By **Rafa Hernández (elpluto)**  
[cheatography.com/elpluto/](https://cheatography.com/elpluto/)

Published 2nd January, 2023.  
Last updated 2nd January, 2023.  
Page 4 of 8.

Sponsored by **Readable.com**  
Measure your website readability!  
<https://readable.com>

### Cryptographic functions

md5  
sha1  
sha256  
sha512

### Web functions

urischeme  
urihost  
uriport  
uripath  
urifragment  
uriquery  
uriuser  
urissp  
uriauthority  
absoluteuri  
opaqueuri  
urldecode  
uaurl  
uaname  
uatype  
uaversion  
uaicon  
uarobot  
uainfourl  
uafamily  
uacompany  
uacompanyurl  
uadeviceicon  
uadeviceinfourl  
uadevicetype  
uaosurl  
uaosname  
uaosicon  
uaosfamily  
uaoscompany  
uaoscompanyurl  
uaosversion Possible missing function to filter by OS version.

### Meta functions

pragmavalue  
tablename

### Data Types

**str** String

**int** Integer number: 1, 58, 12598

**float** Floating point number: 24.256

**boolean** Boolean: true, false

**timestamp** Timestamp date in format: yyyy-MM-dd HH:mm:ss.SSS

**boxar(int)** Byte array in hexadecimal string format

**duration** Amount of time: an integer following by a letter [d]ays, [h]ours, [m]inutes, [s]econds, [No suffix ]:m ill ise conds

**geocord** Geographic coordinates set:  
 Latitude/ longitude sexage simal values: 40°24'N 3°41'W  
 Hash representation of coordinates (geohash)

**ip** IPv4 address format: 192.168.5.56

**ip6** IPv6 address format: 2001:0db8:85a3:0000:0000:8a2e:0370:7334

**net4** IPv4 address in format: {x.x.x.x/0}

**net6** IPv6 address in format: x.x.x.x.x.x/s

**regexp** Regular expression: [^\w]

**template** Represents a substitution string mask.

**dc** Represents a estimated count of distinct elements in a data stream.

**image** Image as Base64 encoding image.

**mac** MAC address in format: 00:0a:95:9d:68:16

**namepattern** Represents a part of a table name: my.app, demo, ...

**set(name)** Represents a set of table names: {my.app, p.test, my.app.test2}

**json** String in json format: {"id":345, "name":"John"}



### Data Types (cont)

**jq** Represents a **jq** filter, **jq** is a command line json processor.  
 👁 `.email`

### Common comparison functions and operators

**eq, [=]** **Equals to** function and operator:  
 ✔ `eq(column, value or column)`  
 ✔ `column1 = value or column`

**eqic** Case insensitive **Equals to** function:  
`eqic(column, value or column)`

**ge, [>=]** **Greater or equal** function and operator:  
 ✔ `ge(column, value or column)`  
 ✔ `column >= value or column`

**gt, [>]** **Greater than** function and operator:  
 ✔ `gt(column, value or column)`  
 ✔ `column > value or column`

**le, [<=]** **Less or equal** function and operator:  
 ✔ `le(column, value or column)`  
 ✔ `column <= value or column`

**lt, [<]** **Less than** function and operator:  
 ✔ `lt(column, value or column)`  
 ✔ `column < value or column`

**ne, [!=]** **Not equal** function and operator:  
 ✔ `ne(column, value or column)`  
 ✔ `column /= value or column`

**isnull** Check if **is null** function:  
`isnull (column)`

**isnotnull** Check if **is not null** function:  
`isnotnull (column)`

### Logig Functions

**and**

**or**

**not**

### JSON related functions

**jqeval**

**label**

**jsonparse**

### Math functions and operators

**abs**

**add / [+]**

**sub / [-]**

**mul / [\*]**

**div / [/]**

**rdiv / [/]** Real division function and operator:

**mod / [%%]** Module function:

**rem / [%]** Return the remain of a division operation:

**pow** Power function:

**cbirt** Cube root function:

**sqrt** Square root function:

**ceil**

**floor**

**round**

**signum**

### Statistical and specialised statistical functions

**estimation**

**pack**

**unpackllpp**

### Network functions

**ispublic**

**isprivate**

**ipip4**

**ipprotocol**

**purpose**

**host**

**routing**

**httpstatusdescription**

**httpstatustype**

**reputation**

**score**

**sbl**



### Conversion functions

int

str

bool

float

image

ip4

net4

ip6

net6

mac

to16

from16

to64

from64

toutf8

fromutf8

toz85

fromz85

compatible

mapped

translated

template

timestamp

duration

re

parsedate

formatdate

humansize

mkboxar

### Special comparison functions

**matches** Matches function finds occurrences in a column using a regular expression:

✔ `matches(column, re(string) or regexp value)`

✔ `column ~ re(string) or regexp value`

### Special comparison functions (cont)

**anymatches** Find occurrences in a **set of names** type column against a **namepattern**:

```
anymatches(s etO fNames, nameglob( - string) or namepattern)
```

```
✔ anymatches(s(t ables, nameglob( " - my.a pp.*.* "))
```

**nameglob** Return a formatted string as a **namepattern** to use with **anymatches**:

```
nameglob( string)
```

### Date and time functions

day

dayofweek

dayofyear

month

year

epoch

hour

minute

second

millisecond

today

tomorrow

yesterday

period

### Packet functions

hasio4

hastcp

hasudp

hasether

ip4proto

ip4src

ip4dst

ip4status

ip4ttl

ip4len

ip4payload

ip4flags



### Packet functions (cont)

ip4fragment

ip4cs

ip4hl

ip4ds

ip4ecn

ip4tos

etherdst

ethersrc

etherpayload

etherstatus

ethertag

ethertype

tcpdst

tcpsrc

tcpstatus

tcpflags

tcppack

tcpcs

tcpseq

tcpnl

tcppayload

tcpurg

tcpwin

udpsrc

udpport

udpstatus

udpcs

udplen

udppayload



By **Rafa Hernández** (elpluto)  
[cheatography.com/elpluto/](https://cheatography.com/elpluto/)

Published 2nd January, 2023.  
Last updated 2nd January, 2023.  
Page 8 of 8.

Sponsored by **Readable.com**  
Measure your website readability!  
<https://readable.com>