

RegEx in SQL

What is a regular
expression?

regex

\bcat\b

The fat cat ran down the street

**How do they work and
what is a regex engine?**

RegEx operators

Anchor operators

`^`

`$`

`\b`

`\B`

Quantifier operators

*

+

?

*

{n}

{n, }

{n,m}

Groups and Ranges

operators

.

()

(x|y)

[]

[^]

[n-m]^{*}

Assertion operators

$x(?=y)$

$x(?!y)$

$(?<=y)x$

$(?<!y)x$

Character Classes

\d

\w

\s

\D

\W

\S

RegEx functions in SQL

REGEXP_LIKE(column_name, 'pattern')

column ~ or ~* 'pattern'

**REGEXP_REPLACE (string, ‘pattern’,
‘replacement’)**

**REGEXP_SUBSTR (string, 'pattern',
start_position, occurrence,
match_parameter)**

substring(column FROM 'pattern')

Differences in RegEX in the "dialects" of SQL

Practical examples

id	email
1	name.surname@gmail.com
2	ivanov123@yahoo.com
3	fake_email@@test

Select all real emails

```
SELECT email  
FROM users  
WHERE REGEXP_LIKE(email, '^[A-Za-z0-  
9._%+-]+@[A-Za-z0-9.-]+\.\.[A-Za-z]  
{2,}\$');
```

Anonymize users

```
SELECT REGEXP_REPLACE(email, '@.*',
  '@example.com') AS masked_email
FROM users;
```

Extract only email domains

```
SELECT REGEXP_SUBSTR(email, '@[A-Za-  
z0-9.-]+') AS domain  
FROM users;
```

Quiz time

Which of the strings
given match the pattern?

regex

ha+

1. ha

2. haa

3. haaa

4. h

regex

ha+

1. ha

2. haa

3. haaa

4. h

regex

```
\d{3}-\d{3}-\d{4}
```

1. 123-4567-8901
2. 12-3456-7890
3. 123-456-7890
4. 1234-567-8901

regex

```
\d{3}-\d{3}-\d{4}
```

1. 123-4567-8901
2. 12-3456-7890
3. 123-456-7890
4. 1234-567-8901

regex

```
^Hello.*world!$
```

1. Hello, world.
2. Oh, Hello big world!
3. Hello. world!
4. Hello amazing world!

regex

```
^Hello.*world!$
```

1. Hello, world.
2. Oh, Hello big world!
3. Hello. world!
4. Hello amazing world!

regex

[A-Z][a-z]*

1. Bob

2. alice

3. J

4. John

regex

[A-Z][a-z]*

1. Bob

2. alice

3. J

4. John

What would the given regex do?

regex

```
^.*(cat|dog).*$
```

What would the given regex do?

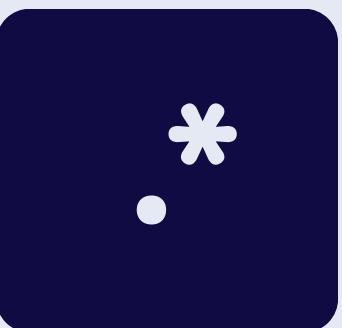
regex

```
^.*(cat|dog).*$
```

Finds any line containing cat or dog.

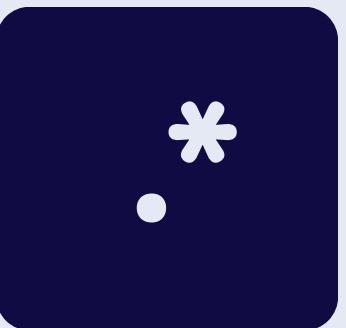
What does this pattern match?

regex



What does this pattern match?

regex



Everything