

## Expressões regulares (regex)

Regex é uma ferramenta usada para **buscar, identificar, validar, substituir ou extrair padrões de texto** em strings. Ela funciona como uma "linguagem" especial para definir padrões que podem ser simples (como uma palavra específica) ou complexos (como e-mails, números de telefone, CEPs, etc.).

### - Caracteres especiais básicos

**.** (**ponto**): Corresponde a qualquer caractere, exceto uma nova linha.

Exemplo: **a.b** corresponde a "acb", "axb", "a1b", etc.

**^** (**caret**): Marca o início de uma linha.

Exemplo: **^abc** corresponde a qualquer linha que comece com "abc".

**\$** (**dólar**): Marca o final de uma linha.

Exemplo: **abc\$** corresponde a qualquer linha que termine com "abc".

**\*** (**asterisco**): Corresponde a zero ou mais ocorrências do caractere anterior.

Exemplo: **a\*b** corresponde a "b", "ab", "aab", etc.

**+** (**mais**): Corresponde a uma ou mais ocorrências do caractere anterior.

Exemplo: **a+b** corresponde a "ab", "aab", "aaab", etc.

**[ ]** (**colchetes**): Define uma classe de caracteres, ou seja, uma lista de caracteres que pode ser correspondida.

Exemplo: **[aeiou]** corresponde a qualquer vogal.

**[^]** (**colchetes com acento circunflexo**): Corresponde a qualquer caractere **exceto** os definidos dentro dos colchetes.

Exemplo: **[^0-9]** corresponde a qualquer caractere que não seja um número.

**|** (**pipe**): Alternância (ou). Corresponde a um padrão ou outro.

Exemplo: **abc|def** corresponde a "abc" ou "def".

### - Quantificadores

Definem quantas vezes um caractere ou grupo deve aparecer para ser considerado uma correspondência.

**{n}**: Corresponde exatamente a **n** ocorrências do caractere anterior.

Exemplo: **a{3}** corresponde a "aaa".

**{n,}**: Corresponde a **n** ou mais ocorrências do caractere anterior.

Exemplo: **a{2,}** corresponde a "aa", "aaa", "aaaa", etc.

**{n,m}**: Corresponde a entre **n** e **m** ocorrências do caractere anterior.

Exemplo: **a{2,4}** corresponde a "aa", "aaa" ou "aaaa".

## - Grupos e Captura

Agrupam partes de uma expressão regular. Eles são delimitados por parênteses e podem ser usados para capturar partes do texto correspondido.

**() (parênteses)**: Cria um grupo para capturar.

Exemplo: **(abc)** captura "abc".

**\**: Usado para referenciar grupos capturados anteriormente (ex: **\1**, **\2**).

Exemplo: **(\d)\1** corresponde a dois dígitos idênticos, como "11", "22", etc.

## - Metacaracteres adicionais

**\d**: Corresponde a qualquer dígito (equivalente a **[0-9]**).

Exemplo: **\d{2}** corresponde a dois dígitos.

**\D**: Corresponde a qualquer caractere que não seja um dígito (equivalente a **[^0-9]**).

Exemplo: **\D+** corresponde a uma sequência de caracteres não numéricos.

**\w**: Corresponde a qualquer caractere alfanumérico (letras e números) e o sublinhado (equivalente a **[a-zA-Z0-9\_]**).

Exemplo: **\w+** corresponde a uma palavra.

**\W**: Corresponde a qualquer caractere que não seja alfanumérico ou sublinhado (equivalente a **[^a-zA-Z0-9\_]**).

Exemplo: **\W+** corresponde a caracteres não alfanuméricos.

**\s**: Corresponde a qualquer caractere de espaço em branco (espaços, tabulação, nova linha).

Exemplo: **\s+** corresponde a um ou mais espaços em branco.

**\S**: Corresponde a qualquer caractere que não seja espaço em branco.

Exemplo: **\S+** corresponde a uma sequência de caracteres não espaciais.

**\b**: Corresponde a uma borda de palavra (início ou fim de uma palavra).

Exemplo: **\babc\b** corresponde à palavra "abc" isolada, mas não a "abc123".