

REGEX (Regular Expression)

Some patterns

`^` -- starts with

`$` -- ends with

`[A-Z]` indicates check all letters from A-Z

`[a-z]` checking lowercase

`[0-9]` checking numeric

Repeaters: (`*`, `+`, `{}`)

- `*` : match characters 0 or more times
- `+` : at least 1 or more times
- `{}` : checking the exact no of length
 - `{3}` check 3 times
 - `{1,3}` means min 1 and max 3

Character Classes

`\w` to match word

`\W` not a word

`\s` match whitespace

`\S` no space

`\d` match digits

`\D` non-digit character

Pattern:: `^[a-zA-Z0-9._+%]+@[a-zA-Z0-9.-]+\.[a-zA-Z]{2,}$`]];

`^` :: starts with

`[a-zA-Z0-9._+%]+` == `[a-zA-Z]` allows letters caps and smalll

`[0-9]` allows digits

`[_+%]` allows only this special symbols in text

meaning of `+` repeat more that 1 time means

sonam.soni (allowed), sonam%soni allowed

`@[a-zA-Z0-9.-]+` ==

`@` is compulsory you have to write it

after `@` again some letters an only `.` - symbol allowed

means `@test-live` (valid) , `@sonam.check` (valid)

`\.[a-zA-Z]{2,}$`

. is compulsory you have to write it (again)

any letter minimum length is 2

.com, .org, .edu (valid)

\$ indicating:: end of line

AWS to validate email address in emails.txt file

```
awk '/^[a-z]+@[a-z]+\.[a-z]+/ { print $1 }' emails.txt
```