Regular Expressions

Regular Expressions are string patterns that describe text
The basic patterns include character classes, quantifiers and meta-characters
Java applies RE in following classes

- String
- Pattern
- Matcher

RE may be used for

- Searching
- Validating
- Extracting
- Modifying

Character Classes

. - (Dot) Any single character
[ ] - Set of individual characters [abc] or range [a-c]
[^] - Negation of set or range [^abc] or [^a-c]
\\d - any decimal digit [0-9]
\\D - any symbol except digit [^0-9]
\\s - any whitespace symbol
\\S - any non-whitespace symbol
\\w - any alphanumeric symbol [a-zA-Z0-9_]
\\W - any non-alphanumeric symbol [^a-zA-Z0-9_]

Quantifiers

How many times the last specified character can be repeated

+ - one or more
* - 0 or more
? - 0 or 1
{m} - exactly m times
{n,m} - no less than n but no more than m
{n,} - no less than n

Meta-Characters

| - X|Y either X or I
() -grouping matching subsequences are remembered and numbered from left to right beginning from 1. Example, “(KU)\1” matches KUKU
^ - beginning of text
$ - end of text

Class String and Regex

matches (regex) - returns true if a string matches a given regex
split(regex) - returns array of tokens separated by delimiters matching regex
replaceAll(regex, substring) replace all the string's
parts matching regex with the given substring
replaceFirst(regex, substring) replace first encountered the string's part matching regex with the given substring