REGULAR EXPRESSION

What is Regular Expression

- Regular Expressions, commonly known as "regex" or "RegExp", are a specially formatted text strings used to find patterns in text.

- Regular expressions are one of the most powerful tools available today for effective and efficient text processing and manipulations.

- For example, it can be used to verify whether the format of data i.e. name, email, phone number, etc. entered by the user was correct or not, find or replace matching string within text content, and so on.

PHP's built-in pattern-matching functions

Function	What it Does
preg_match()	Perform a regular expression match.
preg_match_all()	Perform a global regular expression match.
preg_replace()	Perform a regular expression search and replace.
preg_grep()	Returns the elements of the input array that matched the pattern.
preg_split()	Splits up a string into substrings using a regular expression.

Character Classes

- Square brackets surrounding a pattern of characters are called a character class e.g. [abc].

RegExp	What it Does
[abc]	Matches any one of the characters a, b, or c.
[^abc]	Matches any one character other than a, b, or c.
[a-z]	Matches any one character from lowercase a to lowercase z.
[A-Z]	Matches any one character from uppercase a to uppercase z.
[a-Z]	Matches any one character from lowercase a to uppercase Z.
[0-9]	Matches a single digit between 0 and 9.
[a-z0-9]	Matches a single character between a and z or between 0 and 9.

Predefined Character Classes

- Some character classes such as digits, letters, and whitespaces are used so frequently that there are shortcut names for them.

Shortcut	What it Does
.	Matches any single character except newline \n.
\d	matches any digit character. Same as [0-9]
\D	Matches any non-digit character. Same as [^0-9]
\s	Matches any whitespace character (space, tab, newline. Same as [\t\n]
\S	Matches any non-whitespace character. Same as [^ \t\n\r]
\w	Matches any word character (definned as a to z, A to Z,0 to 9, and the underscore). Same as [a-zA-Z_0-9]
\W	Matches any non-word character. Same as [^a-zA-Z_0-9]

Position Anchors

- There are certain situations where you want to match at the beginning or end of a line, word, or string. To do this you can use anchors.

- Two common anchors are caret (^) which represent the start of the string, and the dollar ($) sign which represent the end of the string.

RegExp	What it Does
^p	Matches the letter p at the beginning of a line.
p$	Matches the letter p at the end of a line.

Pattern Modifiers

- A pattern modifier allows you to control the way a pattern match is handled.

- Pattern modifiers are placed directly after the regular expression, for example, if you want to search for a pattern in a case-insensitive manner, you can use the i modifier, like this: /pattern/i. The following table lists some of the most commonly used pattern modifiers.

Pattern Modifiers(contd.)

Modifier	What it Does
i	Makes the match case-insensitive manner.
m	Changes the behavior of ^ and $ to match against a newline boundary (i.e. start or end of each line within a multiline string), instead of a string boundary.
g	Perform a global match i.e. finds all occurrences.
o	Evaluates the expression only once.
s	Changes the behavior of . (dot) to match all characters, including newlines.
x	Allows you to use whitespace and comments within a regular expression for clarity.