Regular Expression:

Regular expression (regex) is a string that contains special characters and symbols to retrieve data.

\w represents every alphanumeric characters (a-z, A-Z, 0-9)

\W represents any non-alphanumeric characters.

\d represents any numeric digits (0-9)

\D represents any non-numeric digits.

. represents any characters except<Enter>

+ means one or more occurrences

‘-‘ zero or more occurrences.

? 0 or 1 occurrences

[abc]= a or b or c

[] this is used to retrieve multiple characters in or like

Ex:[an] means a or n.

{0,4}=0 to4

^ search in the beginning

$ search in the ending.

Methods:

Search() : return first occurrence of the matching str=obj=group()# it will search for whole string but will return first occurrence only.

findall(): returns all matching strings=list=for loop

match():it will only match the pattern with first string only. Searches only in the beginning of the main string=obj=group().

Sub()=replace the old string with new

Split()=to split the string.

Web scrapping: Retrieving data from a html page.

Retriving required data from any source.

1.regex

2.JSON