



Introduction to Python I (Exercises 05)

Regex

- 1) Write a program that prompts for a string. Make sure the string is 10 or more characters in length. The program will then check the entered string for the occurrence of the substring '**code-**', at the beginning of the string. If you find the '**code-**' followed by 2 digits, then print '**code-??**', where ?? are the characters at position 8th and 9th of the string. Otherwise (if the regex pattern is not found), print the last two characters of the string.

- 2) Write a script that calls a function called **checkPhone**(passing a phone string). This function will use the regex module (re) to validate a pattern. Find a suitable regex expression on the net (google it).

The script will prompt for a phone number and validate it with the function. Return the phone number string if the format is correct. Otherwise return None.

Sample phone string: (514) 990-0909, 1 (438) 887-9090, 57 (300)794-5529

if the phone is valid, print it. Otherwise, print "Invalid phone" (Do not print inside the function).

- 3) Write a program that prompts for a file name (codefile.txt). The file (codefile.txt) is provided in the documents under the exercises section. Opens the file, scans all the lines that start with the characters 'AA' followed by an integer between 1-9, followed by one of the characters a, b, c, d and followed by another integer between 1 and 9. If the pattern is matched at the beginning of the line, add up (sum) the values of the 10th and 11th characters of each line. Print the total number of lines read, and the total of the sum.

- 4) Write a function that checks for Canadian Postal Codes.

The postal code is a six-character code defined and maintained by Canada Post Corporation (CPC) for the purpose of sorting and delivering mail. The characters are arranged in the form 'ANA NAN', where 'A' represents an alphabetic character and 'N' represents a numeric character (e.g., K1A 0T6). The postal code uses 18 alphabetic characters and 10 numeric characters. Postal codes do not include the letters D, F, I, O, Q or U, and the first position also does not make use of the letters W or Z.

Hint: [ABCEGHJ-NPRSTVXY] checks for any of these characters. For this example no lower case characters.