**Python and Bio-Python**

Assignment 9

Practical Solving:

Perform the following questions using regular expressions

1. [Check whether a string starts and ends with the same character or not](https://www.geeksforgeeks.org/python-check-whether-a-string-starts-and-ends-with-the-same-character-or-not-using-regular-expression/)

Algorithm:-

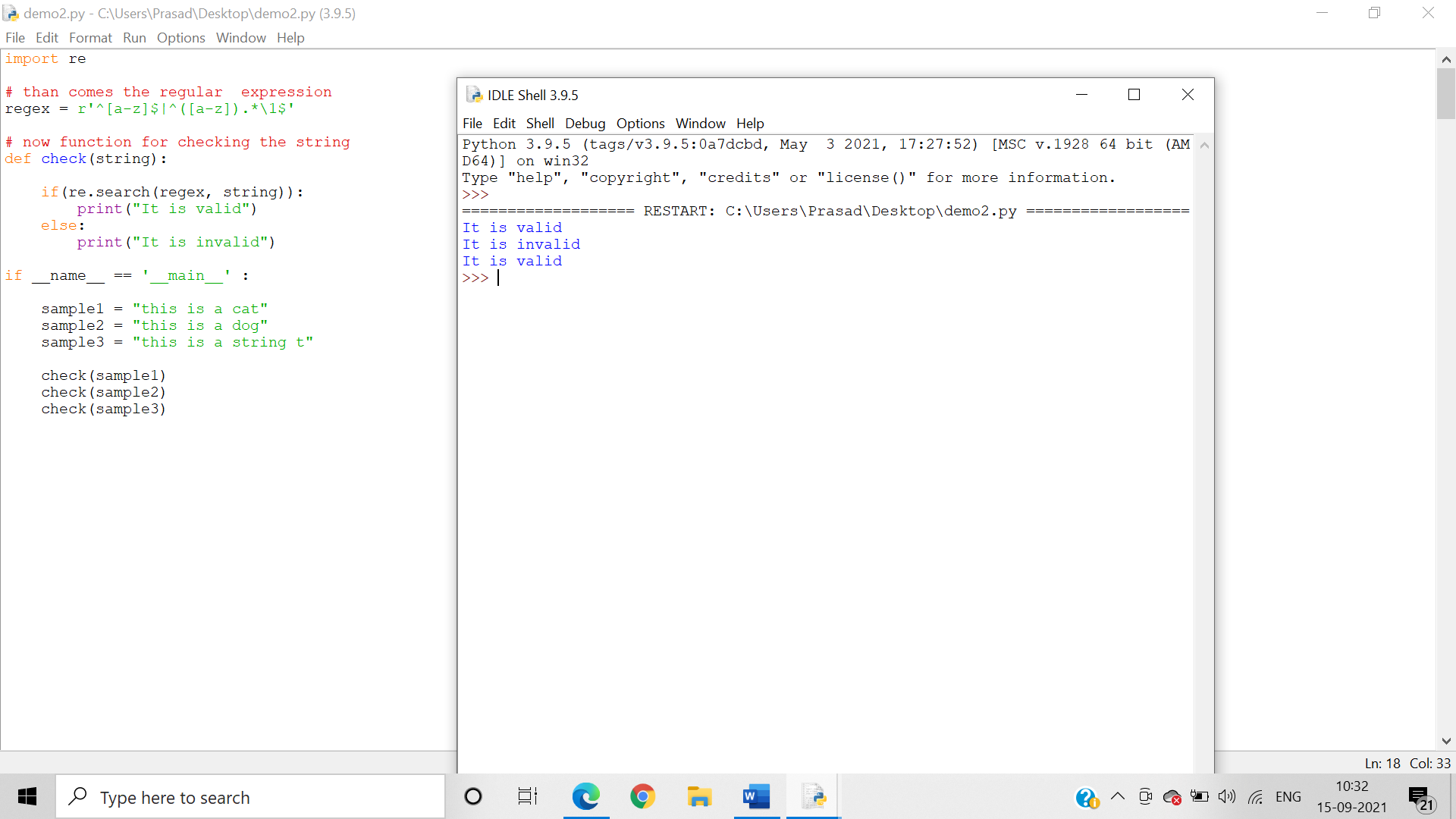
Step 1: Import the re module

Step 2: Define the regex pattern

Step 3: Check the pre-defined inputs for the pattern

Step 4: IF the pattern exists : print (“Valid”)

Step 5: Else: print (“Invalid”)



1. Write a program to Extract IP address from file using Python

Algorithm:-

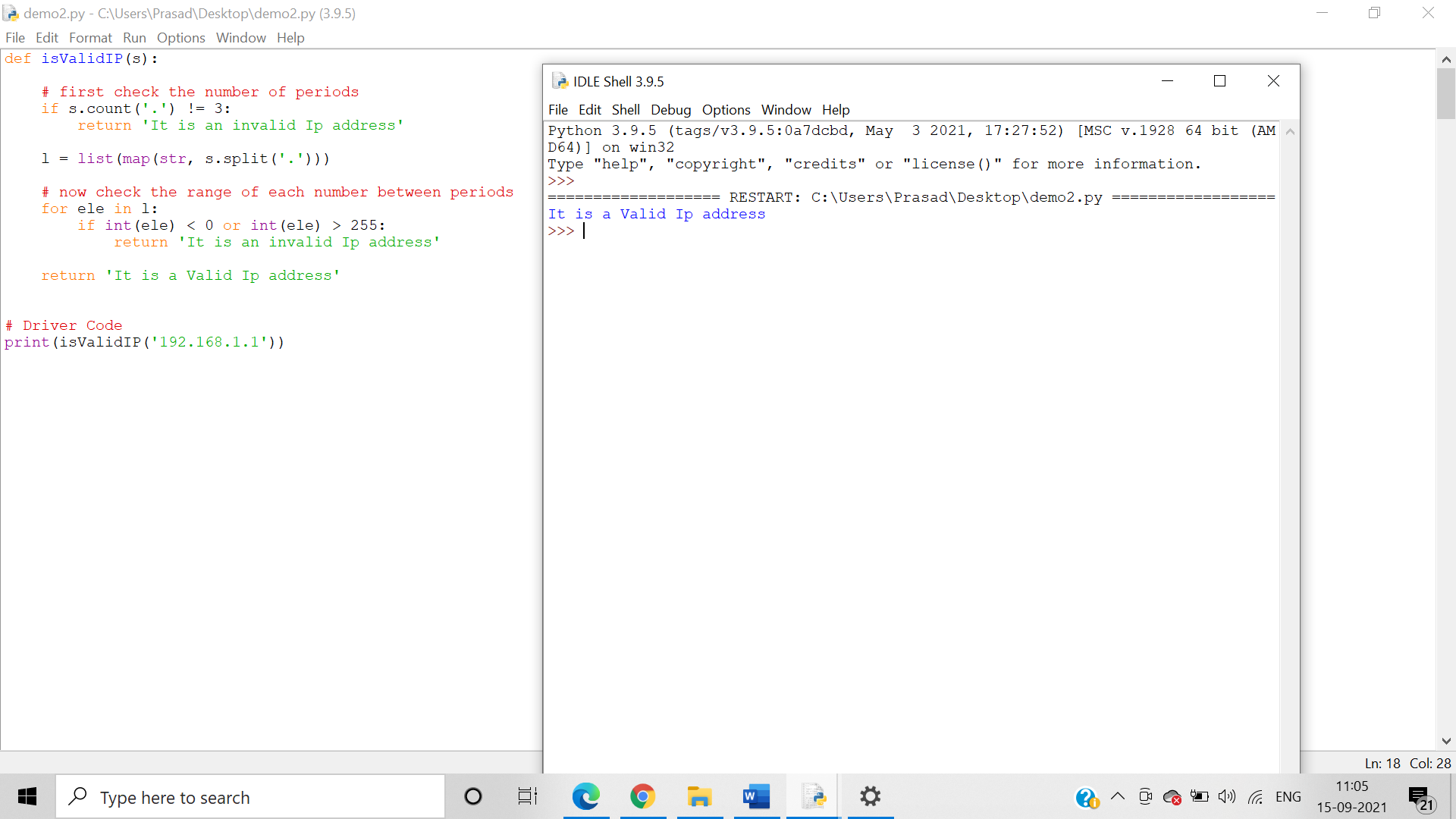
Step 1: import re

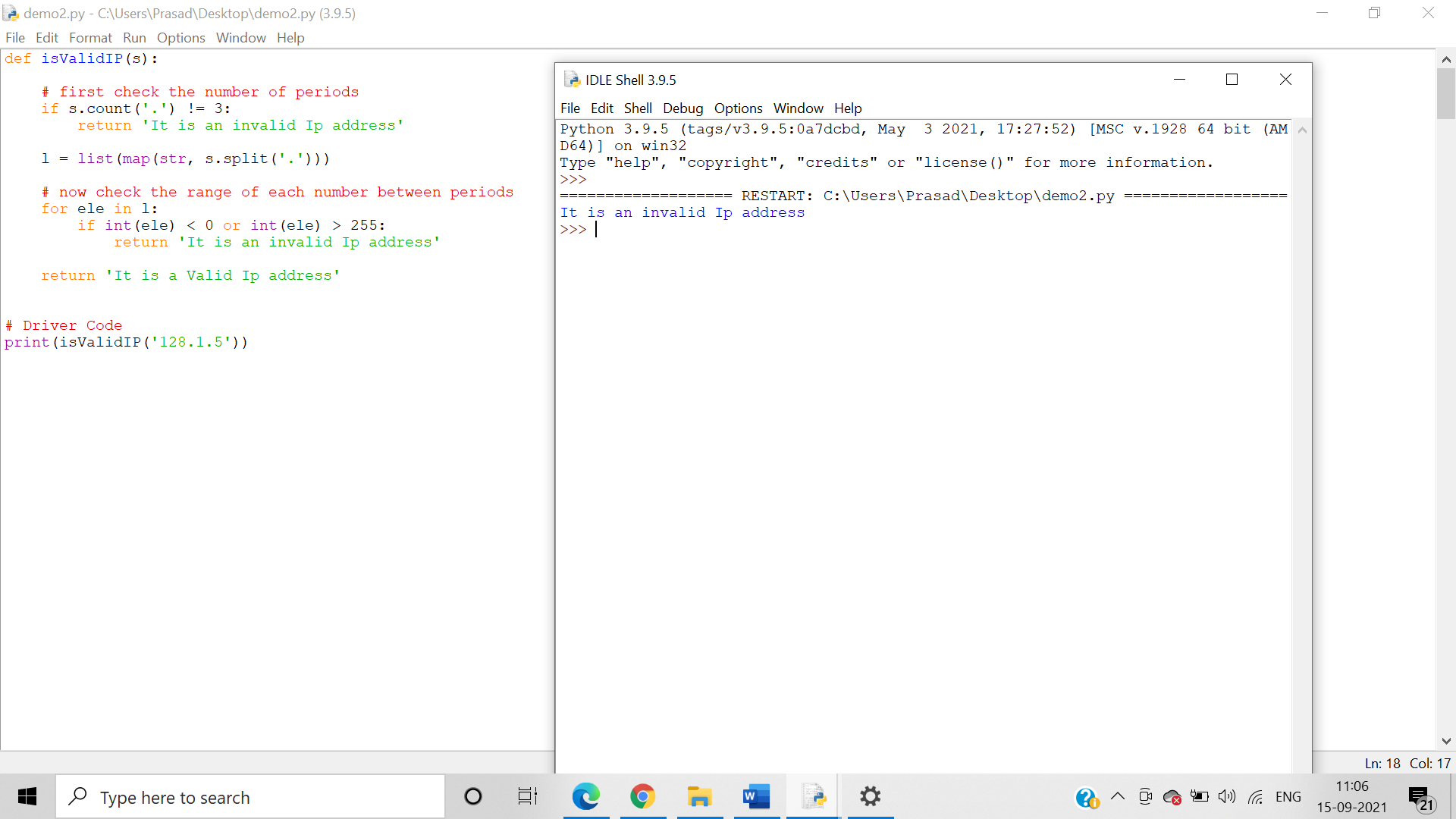
Step 2: open the .txt file

Step 3: Define the regex pattern for an ip address

Step 4: Search for pattern within the file

Step 5: if found : print(empty)





1. Write a program to [Categorize Password as Strong or Weak using Regex in Python](https://www.geeksforgeeks.org/categorize-password-as-strong-or-weak-using-regex-in-python/)

Algorithm:-

Step 1: Import re

Step 2: Check the length of password

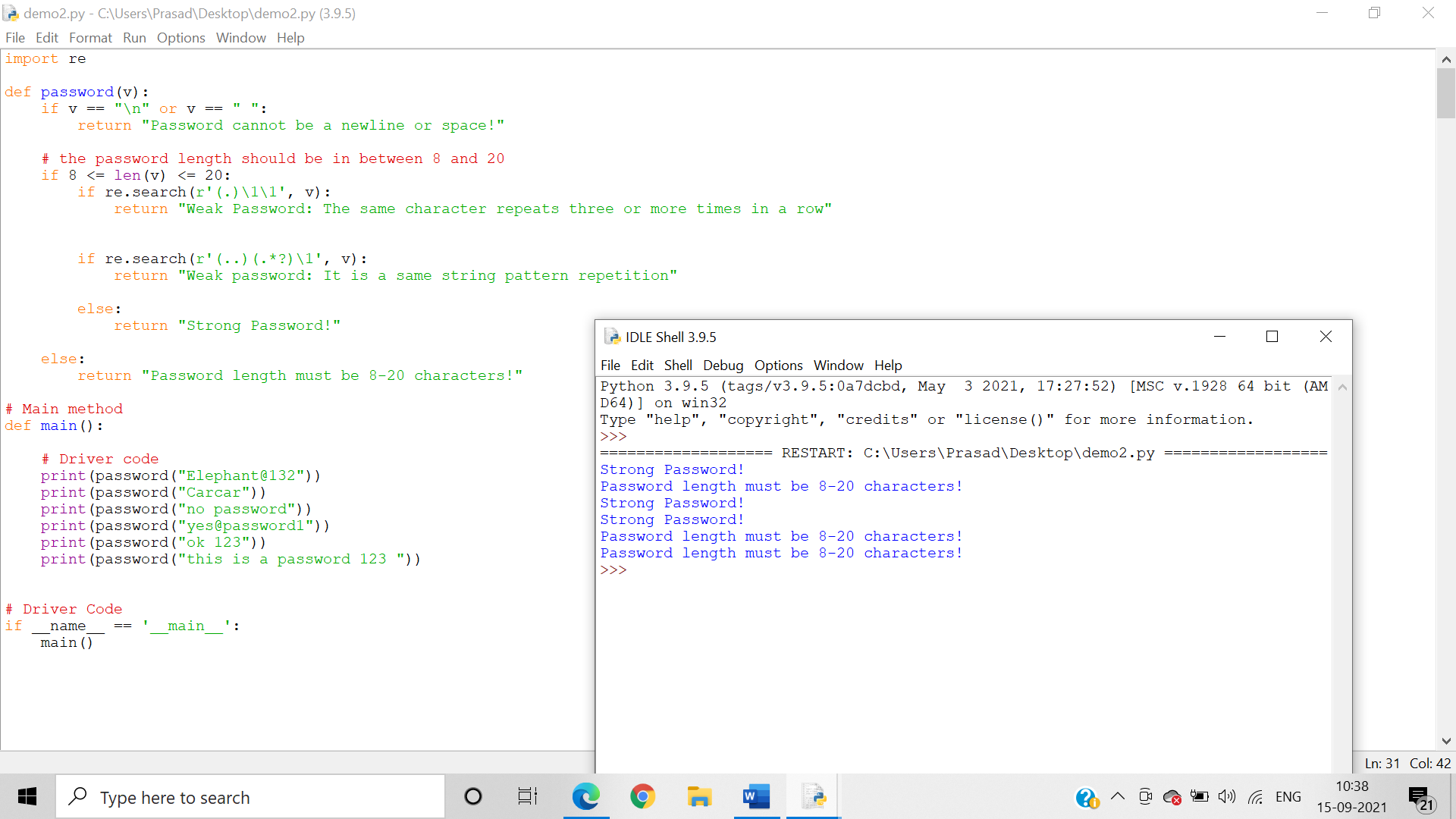
Step 3: if length >=8:

Step 4: Check for special characters, capital letters,small letters ,numbers within the input

Step 5: if all of the above are present: print(“your password is strong”)

Step 6: else : print(“your password is weak”)

Step 7: else : print(“Invalid password”)



1. Write a program to validate an Email address entered by the user

Algorithm:-

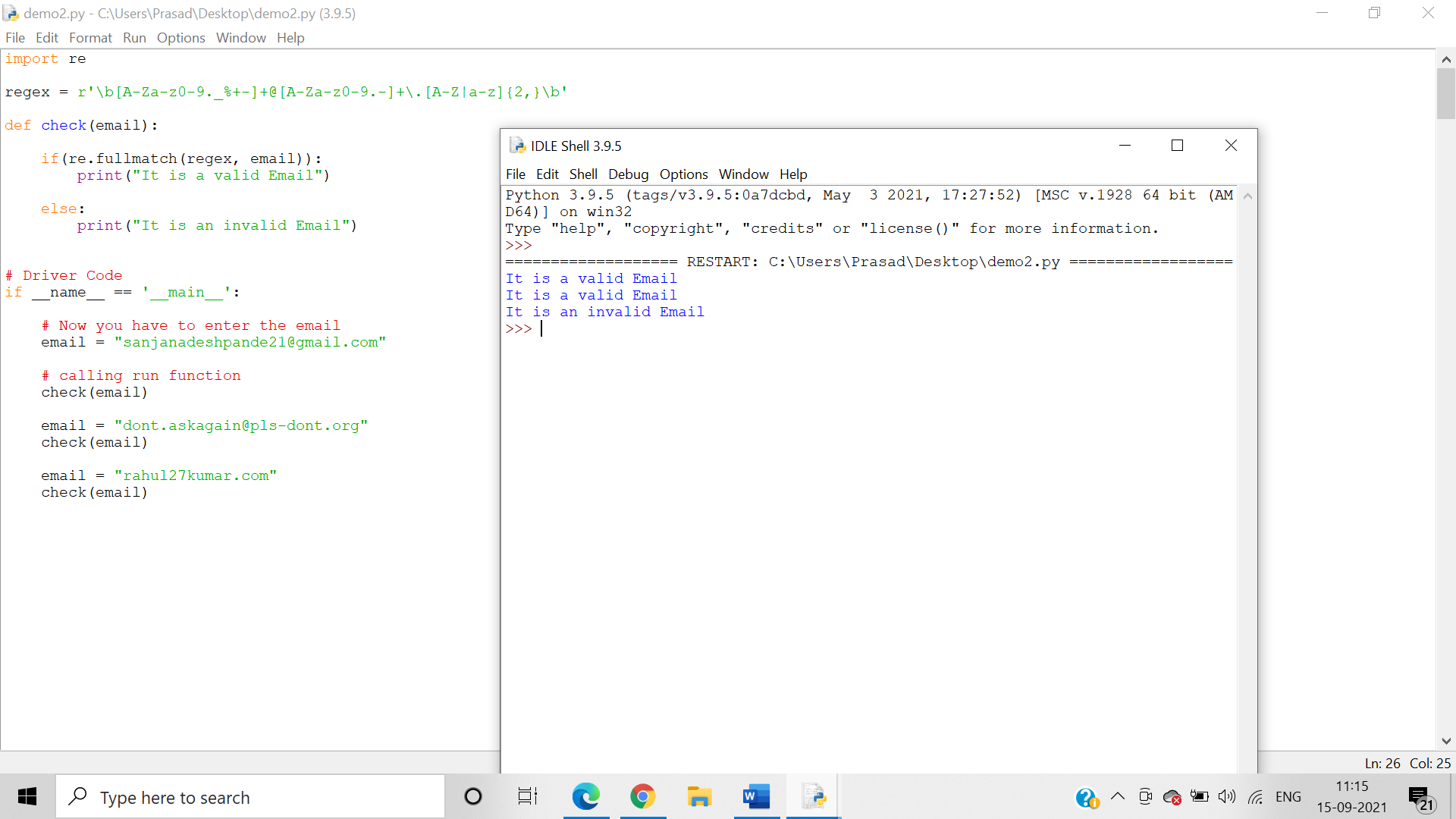
Step 1: Import re

Step 2: Take user defined input

Step 3: Define the regex pattern

Step 4: if (input==regex pattern): print(“valid email id”)

Step 5: else: print (“Invalid”)



1. Write a program to identify no of motifs and their positions in a protein sequence entered by the user. Print the motifs present along with count and positions.(5 motifs)

Algorithm:

Step 1: import re

Step 2: define the regex pattern

Step 3: initialize x = True

Step 4: result1 = re.findall(pattern,sequence)

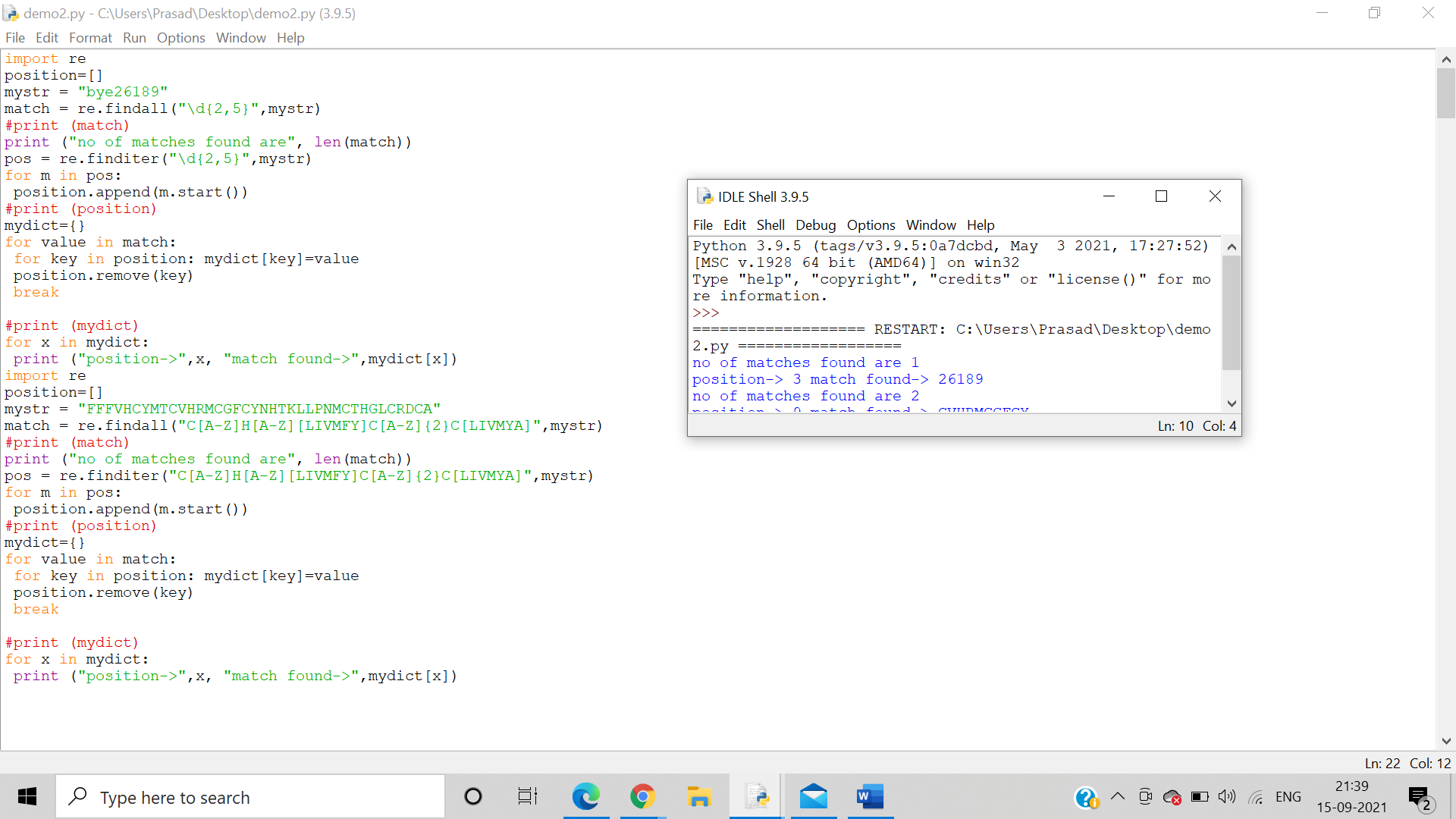
Step 5: result2 = refinditer(pattern,sequence)

Step 6: initialize b = 0

Step 7: for x in result2: print(“Motif position = and motif sequence = “)

Step 8: b = b+1

Step 9: calculate the total no of results using len()



6. Write a program that parses a fasta sequence file (use file handling) and print the raw sequence onto the console. Also parse the accession Id and print it.

7.     Write a program that parses a genbank sequence file (use file handling) and print the annotation in one file and raw sequence in another file.

8.     Write a program that parses helix or sheet regions from the PDB file and print onto the console