ASSIGNMENT 1

Basic Python Programming

By Aditya Anil,AM.EN.U4AIE19006

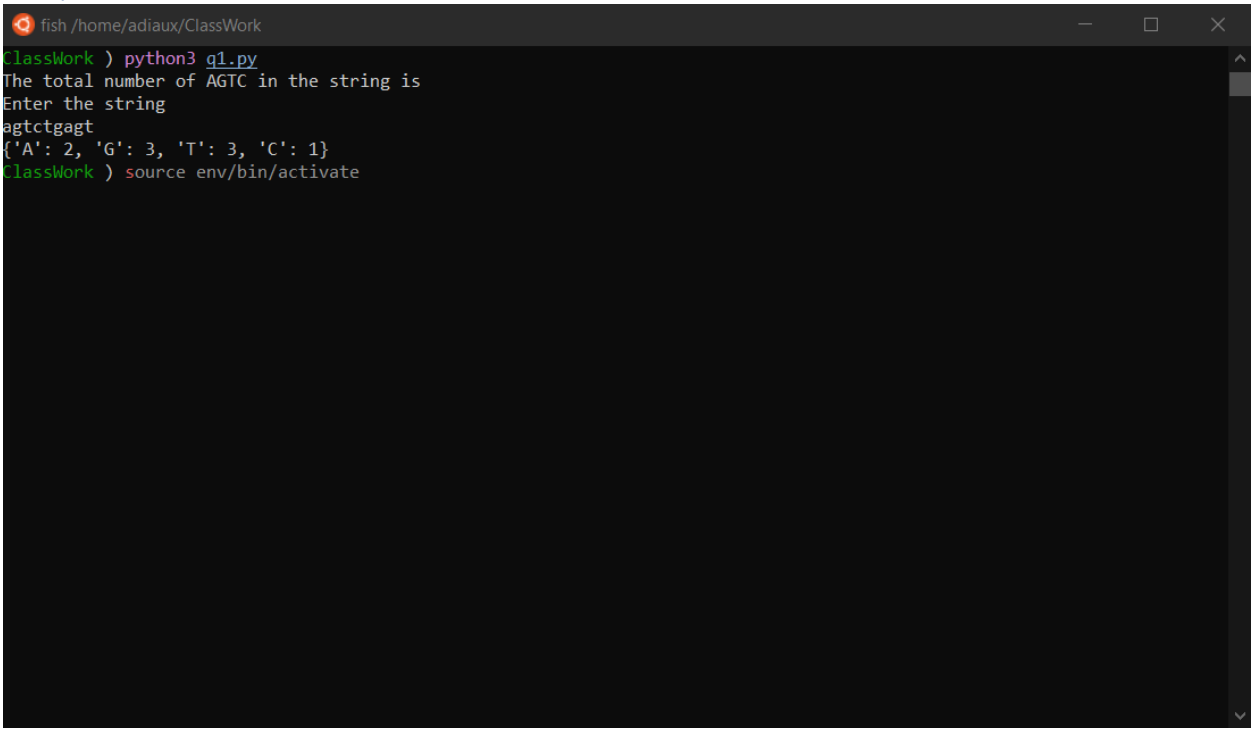
Question 1

Create a Python script using functions to validate a DNA sequence. The sequence should be specified as the parameter to the function and the function should check whether the length of the sequence equals the total number of A, T, G, C alphabets present in the sequence

Code:

```
def calNumber(str):
    baseDict={}
    countA=countG=countT=countC=0
    for x in str:
        if x=='A':
            countA=countA+1
        if x=='G':
            countG=countG+1
        if x=='T':
            countT=countT+1
        if x=='C':
            countC=countC+1
    baseDict["A"]=countA
    baseDict["G"]=countG
    baseDict["T"]=countT
    baseDict["C"]=countC
    return baseDict
print("The total number of AGTC in the string is")
print(calNumber(input("Enter the string \n").upper()))
```

Output:



Question 2

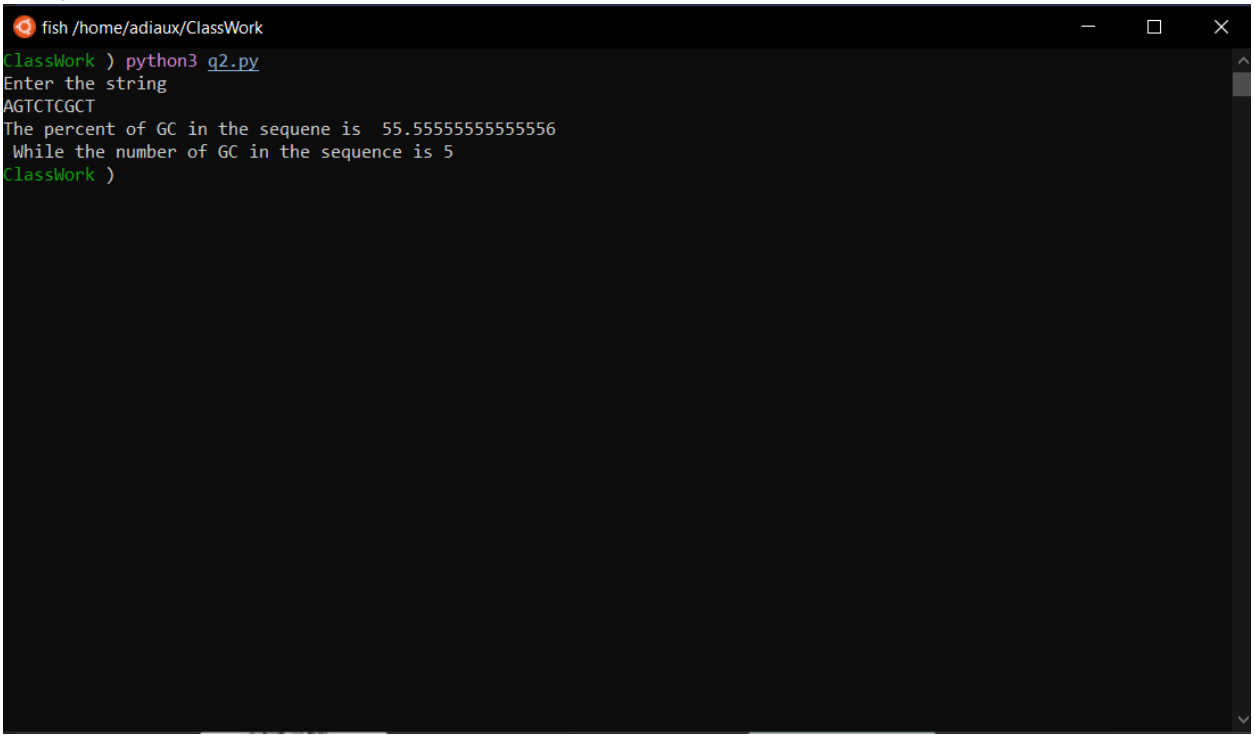
Write a function to calculate and return the GC content of a sequence.

Code:

```
def calGC(str):
    cGC=str.count("G") + str.count("C")
    count = (str.count("G") + str.count("C"))/(str.count("A") + str.count("T")+st
r.count("G") + str.count("C"))
    return [count,cGC]

list=calGC(input("Enter the string \n").upper())
print("The percent of GC in the sequene is ",list[0]*100,"\n While the number of
GC in the sequence is",list[1])
```

Output:



Question 3

Write a function to make single base mutation at random position.

Code:

```
import random
def mutation(str):
    list=['A','G','T','C']
    num=random.randint(0,int(len(str))-1)
    ch=random.choice(list)
    while ch==str[num]:
        ch=random.choice(list)
    str = str[:num] + ch + str[num + 1:]
    print(str)
mutation(input("Enter the DNA sequnce \n").upper() )
```

Output:


fish /home/adiaux/ClassWork
ClassWork) python3 q3.py
Enter the DNA sequnce
ATGCTGTA
ATGCTATA
ClassWork)

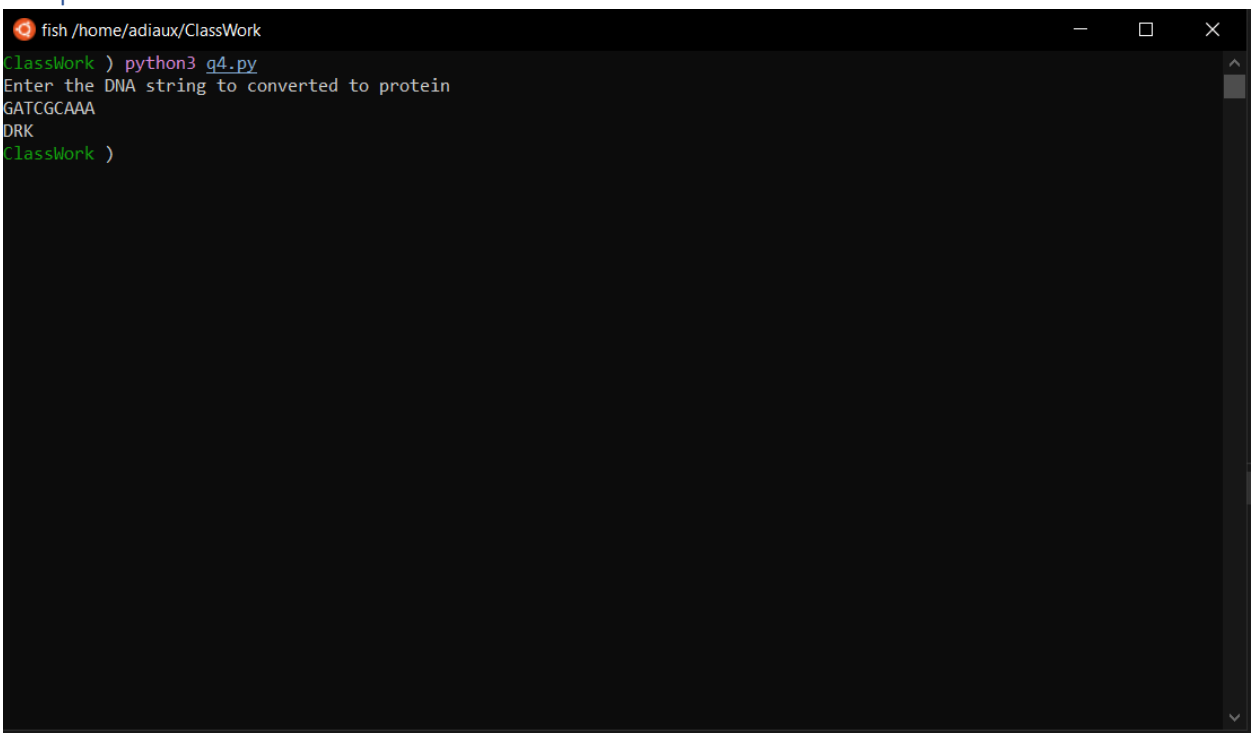
Question 4

Using python dictionaries, create an RNA codon translation table and perform translation of a DNA sequence into a protein.

Code:

```
def convert(str):
    baseDict={
        "UUU":'F',"UUC":'F',"UUA":'L',"UUG":'L',
        "CUU":'L',"CUC":'L',"CUA":'L',"CUG":'L',
        "AUU":'I',"AUC":'I',"AUA":'I',"AUG":'I',
        "GUU":'V',"GUC":'V',"GUA":'V',"GUG":'V',
        "UCU":'S',"UCC":'S',"UCA":'S',"UCG":'S',
        "CCU":'P',"CCC":'P',"CCA":'P',"CCG":'P',
        "ACU":'T',"ACC":'T',"ACA":'T',"ACG":'T',
        "GCU":'A',"GCC":'A',"GCA":'A',"GCG":'A',
        "UAU":'Y',"UAC":'Y',"UAA":'_',"UAG":'_',
        "CAU":'H',"CAC":'H',"CAA":'Q',"CAG":'Q',
        "AAU":'N',"AAC":'N',"AAA":'K',"AAG":'K',
        "GAU":'D',"GAC":'D',"GAA":'E',"GAG":'E',
        "UGU":'C',"UGC":'C',"UGA":'_',"UGG":'W',
        "CGU":'R',"CGC":'R',"CGA":'R',"CGG":'R',
        "AGU":'S',"AGC":'S',"AGA":'R',"AGG":'R',
        "GGU":'G',"GGC":'G',"GGA":'G',"GGG":'G',
    }
    str=str.replace('T','U')
    protein =""
    flag=0
    if len(str)%3 == 0:
        for i in range(0, len(str), 3):
            protein+= baseDict[str[i:i + 3]]
            if baseDict[str[i:i + 3]] =="_":
                break
    else:
        print("\n Error retry \n")
    print(protein)
convert(input("Enter the DNA string to converted to protein \n").upper()
)
```

Output


fish /home/adiaux/ClassWork
ClassWork) python3 q4.py
Enter the DNA string to converted to protein
GATCGCAAA
DRK
ClassWork)