1) Which of the following is not a valid Python expression:
   a. 5+2*3.0
   b. "a"+"bcd"
   c. str(3)+"tree"
   d. 3.0+"aaa"

2) What is the result of running the following Python code:

   ```
   a = 10
   while a > 2 :
       a = a - 1
    print(a)
   ```

   a. 5
   b. 3
   c. 2
   d. 9

3) Write a small Python code that calculates the sum of numbers 1 to *n*

4) Consider the list L = [1, 2, 3, 4, 5, 6, 7, 8, 9, 10]. Suppose I want to create a new list, containing only the even numbers in L. Which of the following expressions should I use?

   a. L[1:-1]
   b. L[1:]
   c. L[1::2]
   d. L[0::2]

5) Consider the following code to write two lines of text to a file:

   ```
   f = open("myfile.txt", "r")
   f.write("line 1 – asasdfsdfsadfa")
   f.write("line 2 – gksgdkfgdfkgjdfg")
   ```

   Spot 3 errors in the above code

6) Calculate the logic value (True or False) of the following Python expression:

   ```
   ((3>2) and not (6 <4)) or (10==7)
   ```

7) Will the following code ever stop? Why?

   ```
   i = 10
   while i > 5:
       i = i + 1
   ```

8) Consider the following code to read a FASTA file from disk:

```
f = open("file.fasta", "r")
lines = f.read().splitlines()
Header = lines[1:]
Sequence = ''
for line in lines[1:]:
    Sequence = Sequence + line
```

Using the following table for molecular weights of the 20 aminoacids:

A	89.1
C	121.2
D	133.1
E	147.1
F	165.2
G	75.1
H	155.2
I	131.2
K	146.2
L	131.2
M	149.2
N	132.1
P	115.1
Q	146.2
R	174.2
S	105.1
T	119.1
V	117.1
W	204.2
Y	181.2

change the above code to calculate the molecular weight of a protein based on its sequence (consider the fact that for each peptide bond that is formed between two a aminoacids, a water molecule (molecular weight 18) is eliminated, and a polypeptide sequence N residues long has N-1 peptide bonds. (**HINT:** create a dictionary with (amino, weight) pairs).

9) Write a Python function that will accept a string variable as argument and check that this string variable contains a valid protein sequence. (**HINT:** valid protein sequences must contain only characters corresponding to the one-letter codes of the 20 aminoacids.)

10) Suppose the variable "Sequence" contains a nucleic acid sequence. Write a conditional expression that checks if the sequence is DNA or RNA (or if it cannot be known for sure).

Answers to exercises:

1) "d" is not a valid expression, because you can't add a string and a float point number together

2) The result is 2 (c.), because the body of the "while" loop will only be executed if a > 2, so the last time it can be run is when a is 3, and then the statement "a=a-1" will turn a into 2, and the body will not be executed, a and "a" will still be 2 when the "print(a)" statement is executed.

3)
```
sum = 0
for i in range(1, n+1):
    sum = sum + i
print(sum)
```

4) L[1::2], because I should start on the first even number on the list (2), and the print every second number.

5) Error 1: the "open" statement should contain a "w" argument, and not "r", because we are writing to the file, not reading it. Error 2: we should add a "\n" newline character at the end of each string argument in the "f.write" commands, otherwise there will be no line breaks and all the characters will be on a single long line. Error 3: after the "f.write" we should issue a "f.close()" command to ensure the new file contents are written to disk.

6) True

7) It won't ever stop. As was written, the while statement will execute until i is lesser or equal to to 5, but since "I" started with value 10 and it is increasing by 1 for each pass of the while loop, it will keep increasing and never reach 5.

8)
```
mw_dict = { 'A' : 89.1,
            'C' : 121.2,
            'D' : 133.1,
            'E' : 147.1,
            'F' : 165.2,
            'G' : 75.1,
            'H' : 155.2,
            'I' : 131.2,
            'K' : 146.2,
            'L' : 131.2,
            'M' : 149.2,
            'N' : 132.1,
            'P' : 115.1,
            'Q' : 146.2,
            'R' : 174.2,
            'S' : 105.1,
```

```python
                        'T' : 119.1,
                        'V' : 117.1,
                        'W' : 204.2,
                        'Y' : 181.2 }

    f = open("plmn_human.fasta", "r")
    lines = f.read().splitlines()
    Header = lines[1:]
    Sequence = ''
    for line in lines[1:]:
        Sequence = Sequence + line

    # sum the molecular weight of each amino
    MW = 0
    for amino in Sequence:
        MW = MW + mw_dict[amino]

    seqlen = len(Sequence)
    # Remove the water molecule weight
    # from each peptide bond formed
    MW = MW - 18*(seqlen-1)

    print("Molecular Weight: ", MW)
```

9)

```python
def isValidProtSeq(sequence):
    valid_aminos = 'ACDEFGHIKLMNPQRSTVWY'
    for amino in sequence:
        if amino not in valid_aminos:
            return False
    return True
```

10)

```python
Sequence = 'ATGCCCCCATTTGGGGCCAAAGGGCAGAAGCCATGACATTTAGGA'

if 'U' in Sequence:
    if 'T' in Sequence:
        print("Sequence contains both DNA and RNA bases?")
    else:
        print("It's an RNA sequence.")
elif 'T' in Sequence:
    print("It's a DNA sequence.")
else:
    print("Could be a DNA or RNA sequence.")
```