COSC 1436, Dr. Pershwitz

Assignment 5c: 15 points

Write a Program: DNA Sequence Analysis

Every DNA sequence consists of four nucleotides: Adenine, Thymine, Cytosine, and Guanine, referred to by the first letters of their chemical names (A, T, C, and G). I have provided an entire DNA sequence in the file **dnaSequence.txt.**

You goal is to provide a report of the number of each nucleotide within the sequence, and the percent each nucleotide makes up of the total. You should include your output file in the submission in addition to our usual submission format (console output + code).

Your program should have the following:

- You should have 4 comment lines at the top: description of the program, author, section, and date. (1 point)
- Create your variables, use the appropriate type, name them appropriately, and remember to not leave them uninitialized. (1 point)
- Make sure the scope of your variables is as small as possible and does not extend beyond the blocks of code where they are needed (1 point)
- Open the file for input and make sure the operation was successful (2 points)
- Read in nucleotides from the file and calculate counts for each one; if you encounter anything other than A, T, C, or G output an error message, skip over it, and keep going (2 points)
- Calculate the total number of nucleotides in the file (1 point)
- Display a report with total number of nucleotides, the number of each nucleotide, and its percentage to two decimal places in a table format (no tabs) (2 points)
- Open a file for output and make sure the operation was successful (2 points)
- Write your report to the output file in the same format as console output (2 points)
- Close your files (1 point)

Example Output:

```
DNA sequence analysis:
29782 nucleotides in the sequence

Sequence breakdown:
Adenine:  8892      29.86%
Thymine:  9581      32.17%
Cytosine: 5462      18.34%
Guanine:  5847      19.63%
```

Notes:

- Pay attention to where you create and how you initialize your variables. Unsafe code is a 1-point deduction.
- Use named constants as appropriate. Insufficient/improper use of named constants is up to 1-point deduction.
- Comment your code. Uncommented code is a 1-point deduction.
- Remember to include your output file in the submission. An insufficient submission is a 1-point deduction.