

CSCE100 Introduction to Informatics
Fall 2021

Programming Assignment 4: Covid-19 Data Analyzer with Functions

Points: 100 points. Assignment Date: October 5, 2021 Due Date: October 12, 2021

Objectives

1. To familiarize with writing and running Python programs and the Python environment
2. To familiarize with designing programmer-defined functions
3. To familiarize with problem decomposition
4. To familiarize with parameter passing with functions
5. To familiarize with the use of online documentations on Python

Problem

You have been given a program that reads in an input file "Covid19_01_01_2021.csv" (downloaded from the GitHub site of Johns Hopkins University on Covid-19), processes the data, the generates and prints out some statistics. The program is "dataAnalyzer.py". The program runs and works properly.

However, the program has been poorly written in that it does not utilize functions properly, leading to repeated use of same lines of code, making it difficult to maintain.

You are required to improve the program by moving lines of code into functions, thereby making the solution more modular, more maintainable. Also, your revised program should also be significantly shorter. Define your functions accordingly. Here are some additional requirements:

- Your new program is required to perform *exactly* the same way as the program given to you to revise, in terms of what users see when running the program. (5 points)
- Your program is required to have *at least* two programmer-defined functions. (10 points)
- You are required to describe, in a Table, each function that you have in your solution. (5 points). For Example, if we provide the description for the readFile() function that is already in the solution: might be:

Name	Input Parameters	Returns	Purpose
readFile	1. The filename in string 2. The variable list to store the data	The number of rows as an integer	To open a file and read row-by-row the content of the file, storing it in the variable list.

- You are required to write an analysis of your functions, discussing the usefulness of each function: Does it improve Modularity? Readability? Maintainability? (10 points)

- You are also required to turn in a report on:
 1. Why do you choose those specific lines of code to convert into functions? (10 points)
 2. How general are the functions? How applicable are your functions to other problems (such as computing other statistics)? (10 points)
- You must document your program (see <https://devguide.python.org/documenting/>).
 - Name, Date, Affiliation, a description of the program, what inputs does it need, what outputs does it generate (5 points)
 - Inline comments in the program (5 points)

Example Input/Output: None

Bonus (20 points)

Come up with at least three relevant statistics and implement them as functions. The idea here is to use these statistics to help us understand the data better, to inform. As part of this Bonus aspect, please also include a short report on what the statistics reveal about the data and their implications.

Handin

1. The submission deadline for all handins is October 12, 2021, 11:00 AM. **Late handins will not be accepted or graded.**
2. You are required to handin a screen capture of your “testing session” using your program. (10 points)
3. You are required to handin all program files. (5 points)
4. You are required to handin all input and output files. (5 points)
5. You are required to handin your description file that consists of two parts: the table of functions and the analysis. (5 points)
6. You are required to handin online the above using <http://cse.unl.edu/handin/>

Think About

Now, think about the usefulness of functions. Defining a function improves modularity, maintainability, and reusability of our program. Furthermore, imagine a function that gets called numerous times at the different occasions in a program. If that function has 25 lines of code, how would the overall program code look like? Does it mean that if the function is called 4 times, then that means we would have to see $25 + 25 + 25 + 25 = 100$ lines of repeated code in the program. Is that bad? Why? (*Hint: Reusability of a function.*)

Think about when we approach a problem to find a solution. One useful approach is to break the problem down to smaller subproblems. Why? And think about how the solution to a subproblem and a function are related. Indeed, programmers often think about how to break a problem down by identifying the functions that need to be designed first, before they start programming. This is especially so when the solution could involve many lines.