Computer Project #4

Assignment Overview

This assignment focuses on the design, implementation and testing of a Python program to analyze some real data using strings, files, and functions.

It is worth 40 points (4% of course grade) and must be completed no later than 11:59 PM on Monday, October 10.

Assignment Deliverable

The deliverable for this assignment is the following file:

    proj04.py – your source code program

Be sure to use the specified file name and to submit it for grading via the **handin system** before the project deadline.

Assignment Background

We collected national GDP (Gross Domestic Project) data from the Bureau of Economic Analysis (part of the U.S. Department of Commerce) for the years 1969 through 2015. The GDP is one of the primary indicators used to gauge the health of a country's economy. It represents the total dollar value of all goods and services produced over a specific time period; you can think of it as the size of the economy. When the economy is healthy GDP is increasing, but if GDP is decreasing, that can indicate problems. Therefore, the change in GDP is useful to examine and that is the goal of this project.

We put the data in a file named `GDP.txt` that is available in the project directory. The file contains data on the annual change in GDP which may be positive or negative and the GDP value itself for each of the years 1969 through 2015.

If you are interested in where we got our data, here is the link:
http://www.bea.gov/iTable/iTable.cfm?ReqID=9&step=1#reqid=9&step=1&isuri=1

(One note about the data. GDP data gets adjusted in a variety of ways. I have extracted two tables from a few thousand lines of tables in the original file. If you try to derive the annual change in GDP from the data on line 44, it will differ by as many as a few percentage points from the change data we are using from line 9. That is, some adjustments are being done that we are ignoring for this assignment.)

Assignment Specifications

The lines of interest in the file `GDP.txt` are the 9th line which has the annual change in GDP and the 44th line which has the value of GDP for each year. The data starts in column 76 and each of the 46 data items spans 12 columns (there are 46 years inclusively from 1969 through 2015). These numbers

are important because you can use string slicing to extract them. For example, the first data item in a `line` can be extracted using the slice `line[76:76+12]`.

Your task is to find the minimum and maximum change in GDP for the years 1969 through 2015 and to find the GDP value for those two years.   See the sample output below. Your program will prompt for an input file and then display the output.

You must use specific functions as described below and you are not allowed to use collections such as lists, tuples and dictionaries.

Assignment Notes

Divide-and-conquer is an important problem solving technique.   In fact, inside of every challenging problem is a simpler problem trying to get out. Your challenge is to find those simpler problems and *divide* the problem into smaller pieces that you can *conquer.* Functions can assist in this approach. **Hint**: build the following functions one at a time and test each one before moving on to the next.

1. The `open_file()` function takes no arguments and returns a file pointer. It repeatedly prompts for file names until a file successfully opens. Use the `try-except` construct checking for the `FileNotFoundError` exception. You do not need to check that the filename is correct—simply check that the file opened. The simplest version of this function has no error checking so the body of the function is simply
   ```
   fp = open("GDP.txt")
   return fp
   ```
   Hint: Start with that as your function body and add the try-except error checking later.

2. The `find_min_percent(line)` function takes one argument, one line from the GDP.txt file. It returns the minimal value in that line and an index indicating where the minimal value is in the line (you get to decide what value is the index—there are multiple possibilities). You can return two values simply by separating them by commas:
   ```
   return min_value, min_value_index
   ```
   Here is an algorithm to find a minimal value of a series of values that you read. The algorithm is written in pseudo-code so it looks somewhat like Python, but needs details of Python to be complete:
   ```
   min_value = 10000000   # some large value
    for each value
       if value < min_value:   # you have found a smaller value
          min_value = value   # set min_value to that smaller value
   ```
   Hint: Start by having your function simply print all the values in the line. That way you can check that you are examining all the values. Once you can print all the values you can apply the minimum algorithm to find the minimal value.

3. The `find_max_percent(line)` function takes one argument, one line from the GDP.txt file. It returns the maximum value in that line and an index indicating where the maximum value is in the line. This function is nearly identical to the `find_min_percent` function. Instead of starting with a large value you start with a small value and reverse the operator in the Boolean expression.

4. The `find_gdp(line,index)` function takes two arguments, one line from the `GDP.txt` file and an `index` into the line. It returns the GDP value in that line at the specified index. Note that this function can be used to find the GDP value associated with the minimal percent change as well as the maximal percent change—simply call the function with a different index.

1. The `display(min_val, min_year, min_val_gdp, max_val, max_year, max_val_gdp)` function takes six arguments: the minimal value, year and GDP, and the same three for maximum (value, year, GDP). The function does not return anything. It displays the values as shown in the sample below. Note that the GDP values in the file are in billions whereas we expect you to convert billions to trillions for the output.

2. You are not allowed to use collections such as list, tuples and dictionaries. Specifically, do not use the `.split()` method. Use string slicing to extract values from lines in the file.

3. You will be responsible for adhering to items 1-6 of the Coding Standard (http://www.cse.msu.edu/~cse231/General/coding.standard.html)

Suggested Procedure

- *Solve the problem using pencil and paper first.* You cannot write a program until you have figured out how to solve the problem. This first step can be done collaboratively with another student. However, once the discussion turns to Python specifics and the subsequent writing of Python, you must work on your own.

- Use Anaconda to create a new program. Use the required file name (`proj04.py`).

- Write a simple version of the program, e.g. open the file and print all the lines in the file. Run the program and track down any errors. Add one function and retest.

- Use the **handin** system to turn in the first version of your program.

- Cycle through the steps to incrementally develop your program:

    o   Edit your program to add new capabilities.
    o   Run the program and fix any errors.

- Use the **handin** system to submit your final version.

Sample Output

```
In [115]: runfile('/Users/enbody/Documents/cse231/FS16/Projects/Project04/proj04.py',
wdir='/Users/enbody/Documents/cse231/FS16/Projects/Project04')

Enter a file name: xxxx
Error. Please try again

Enter a file name: GDP.txt

Gross Domestic Product
The minimum change in GDP was -2.8 percent in 2009 when the GDP was 14.42 trillion dollars.
The maximum change in GDP was 7.3 percent in 1984 when the GDP was 4.04 trillion dollars.

In [116]:
```

Educational Research

When you have completed the project insert the 5-line comment specified below.

For each of the following statements, please respond with how much they apply to your experience completing the programming project, on the following scale:

1 = Strongly disagree / Not true of me at all
2
3
4 = Neither agree nor disagree / Somewhat true of me
5
6
7 = Strongly agree / Extremely true of me

Please note that your responses to these questions will not affect your project grade, so please answer as honestly as possible.

Q1: Upon completing the project, I felt proud/accomplished

Q2: While working on the project, I often felt frustrated/annoyed

Q3: While working on the project, I felt inadequate/stupid

Q4: Considering the difficulty of this course, the teacher, and my skills, I think I will do well in this course.

Please insert your answers into the <u>bottom</u> of your project program as a <u>comment</u>, formatted exactly as follows (so we can write a program to extract them).

```
# Questions
# Q1: 5
# Q2: 3
# Q3: 4
# Q4: 6
```