**Lab Assignment 1**

**Continental Data Computation**

|  |  |
| --- | --- |
| **Due Date**  **(a two-week LA)** | |
| **Sections(540,543,544,545)** | **1/28/22 @ 11:59pm** |

**Concepts**

* Review of CS1110 Concepts such as basic data structures, control flow structures, and functions in Python
* Use flow chart to design a function

**Problem Specification**

* Your program is to calculate the continent information based on an input file.
* Each input file contains the meta-information in the following order.  
  1) Name\_of\_country

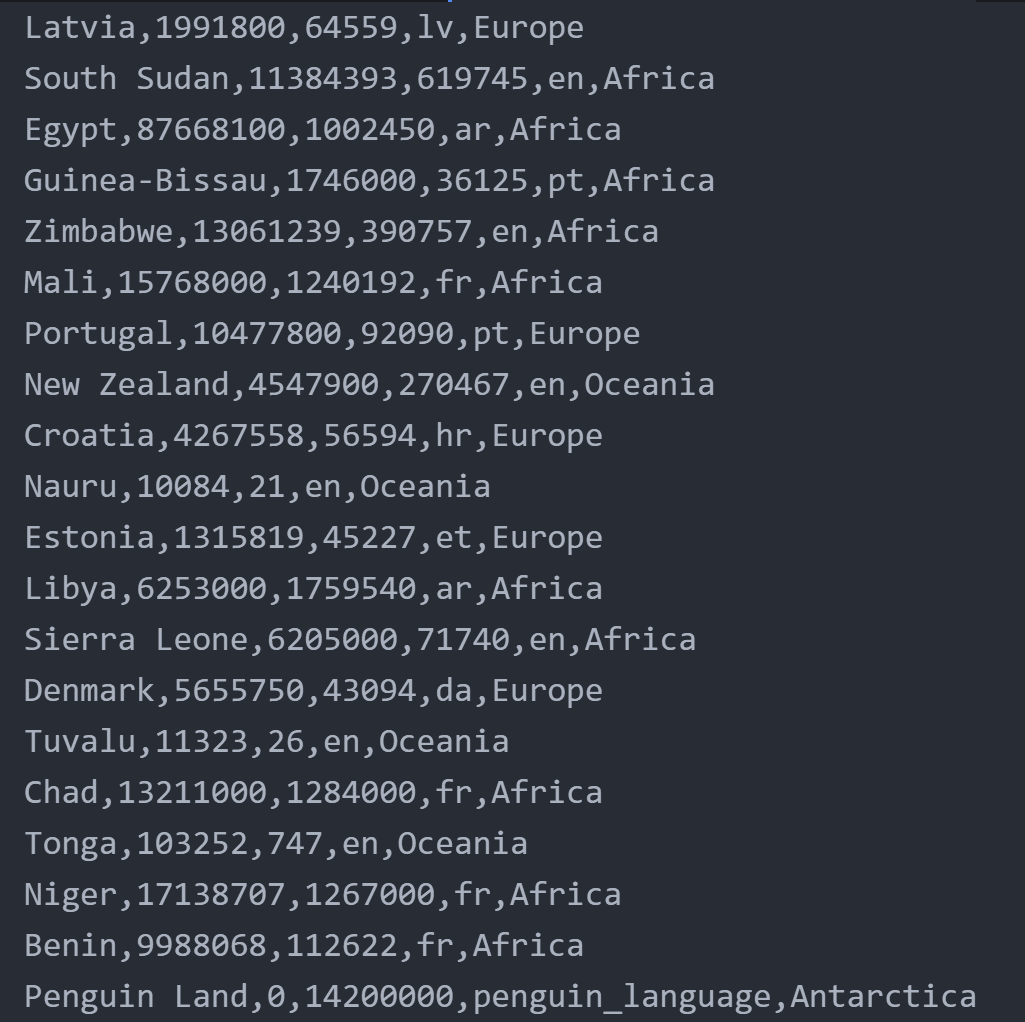
2) Population\_of\_country

3) Area\_of\_country

4) Main\_language

5) Continent

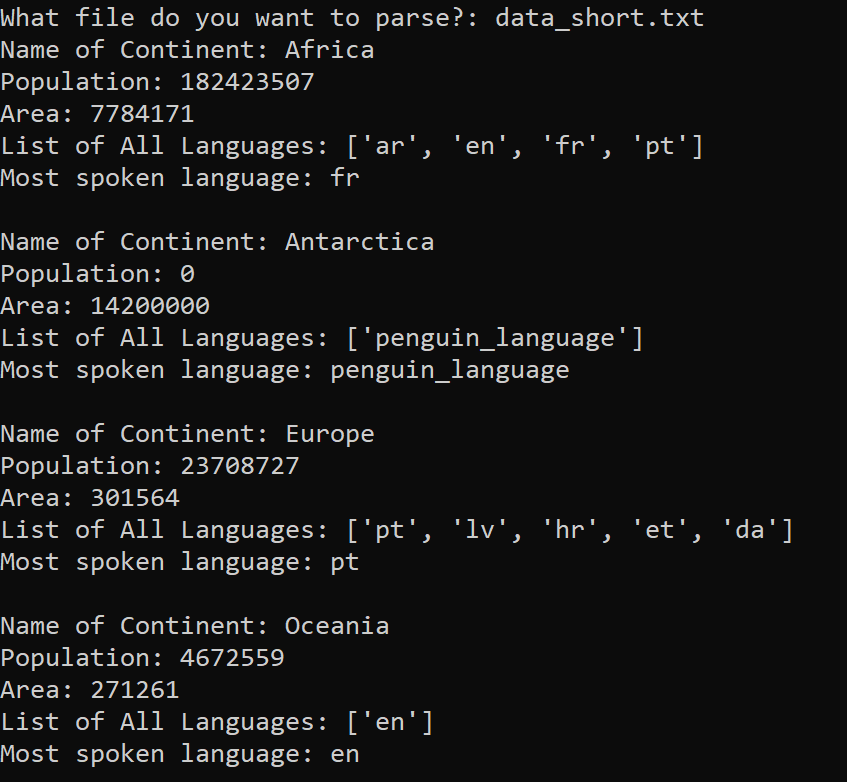
For example, the data\_short.txt as input looks like:



* You take in the whole input file in .txt format and print the desired output in the console.
* You need to parse each line for the given txt file and find out the total population, total area, languages spoken for each continent that must be sorted. The output should contain the metainformation in the following order:

1. Name of Continent: String
2. Population: int
3. Area: int
4. List of All Languages: list of String
5. Most Spoken Language: String

* The output for data\_short.txt is as follows:



* Name of Continent is the name for each continent. In the console, the name shall be sorted.
* Population is the total population of each continent.
* Area is the total area of each continent.
* List of all languages is the collection of all languages spoken in that continent.
* Most spoken language is the language spoken by most countries in the continent. For instance, let’s take an example of the continent Africa in data\_short.txt and manually find out its most spoken language.

A picture containing text, plaque

Description automatically generated

* If you build your own split function without using python’s split, that also counts towards your extra credit, 10 points.

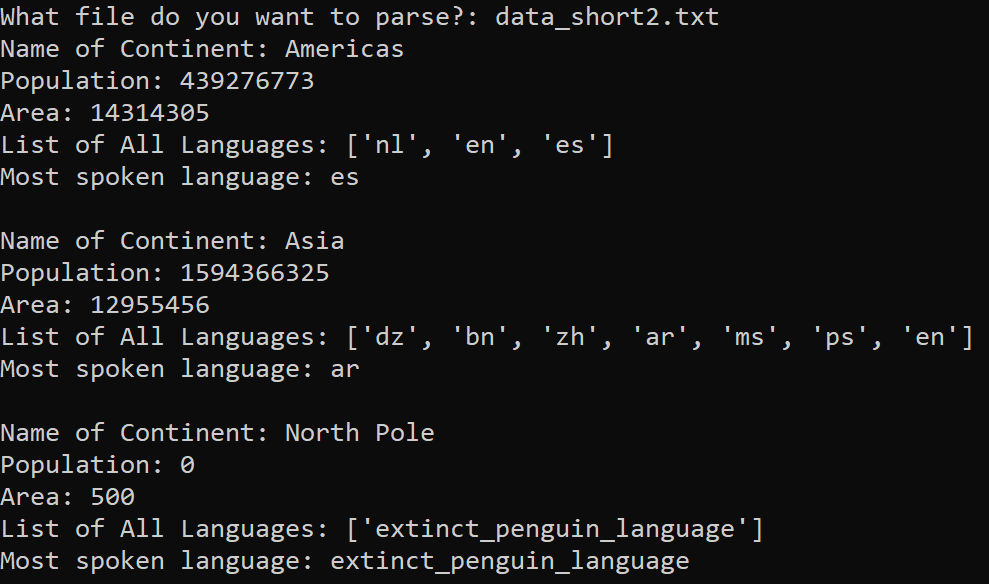
**How to Test**

* We have provided two test files, i.e., data\_short.txt and data\_short2.txt.
* The output for data\_short.txt should look like the following in your console:

Text

Description automatically generated

* The output for data\_short2.txt should look like the following in your console:



* Once the find the correct output for data\_short.txt and data\_short2.txt, finish up your work by doing the same for data.txt. The data.txt is a larger file with five continents.

**Design Requirements**

**Basic Structure**

Your code should be designed in a similar fashion to the following template. A more detailed template has been given to you in the assignment folder.

Text

Description automatically generated

**Flowchart**

You must design a flowchart for each function to illustrate how function is designed. Also, you need to design a flowchart to show how the drive code is designed in terms of how inputs are received, and the functions are called.

**Implementation Phase**

Using the pseudocode developed, write the Python code for your assignment. This is a two-week assignment.

**Testing Phase**

* Another data\_short2.txt is given to you with its output. The output should match exactly.
* Build your program incrementally, carefully testing each function as you go.

**Grading**

* You can earn about 110 points in this assignment. The 10 points counts towards your extra credit.
* A correct program with correct output, with proper use of all the function is worth 100 points.
* If you chose to not use python’s .split and build your own split function you will earn extra 10 points.

**Assignment Submission**

* Generate a .zip file that contains all your files including:
  + Program Files
  + Any input or output files
  + Flowchart.
* Submit the .zip file to the appropriate folder on E-Learning.