Laboratory 7

Topics

1. Reading and writing files
2. Data processing
3. Dictionaries and lists

Discussion

A. What is the difference between a dictionary and a list?
B. Is it possible for two elements in a dictionary to have the same key?

Exercises

Part 1 - Dictionaries

7.1.1 Counting words. Write a program counting all occurrences of a word in a string.
Assume the file contains only alphabetical characters or white spaces.
The program shall output all the words in the file, with the number of occurrences near each word.
[P8.2]

7.1.2 Most frequent words. Extend the program of exercise 11.1.1 so that it shows the 5 most frequent words in the file, not including "the", prepositions and conjunctions. [P8.3]

7.1.3 Cypher. Write a cypher that substitutes each character with a different one. Write a function `encrypt(string)` that can encrypt the input string, and a function `decrypt(string)` that is able to decrypt the string. Use a dictionary to store the values to be substituted, which are the following:

```
A => Q, B => W, C => E, D => R, E => T, F => Y, G => U, H => I, I => O,
J => P, K => A, L => S, M => D, N => F, O => G, P => H, Q => J, R => K,
S => L, T => Z, U => X, V => C, W => V, X => B, Y => N, Z => M
```

It should work both with capital and lowercase characters. Then test it using an input string and show both the outputs on screen.

Part 2 – Files and dictionaries

7.2.1 Counting words part 2. Starting from your program written in **7.1.1**, use a file as input. The input file must be chosen by the user of your program (use the input() function).

7.2.2 Cypher: part 2. Similarly to the previous exercise, read the character transformations from a file. The file structure defines for each row a key-value pair. For instance, if you find the row ‘A Q’, it means that the character ‘A’ shall be substituted with the character ‘Q’.
Then, print the output to a file named “encrypted.txt”.

7.2.2.1 Cypher: part 3. As a follow=up to **7.2.2**, instead of reading the input as a string, the input is included in the input.txt file found in the same folder as **7.2.2**.

7.2.3 Pro capite income. Write a program reading values from the file rawdata_2004.txt inserting them in a dictionary whose keys are the names of Countries, and the values are the annual income pro capite. Please, notice that each field is separated by a tabulation character '\t'. Then, the program shall ask the user to input names of Countries to show each value of pro capite yearly income. Program shall terminate when the string 'quit' is written. Try also working with rawdata_2021.csv performing the same exercise. [P8.17]