

Goals: importing, more programming practice.

For this homework, put all your functions in one “.py” file (rather than “.ipynb” file). Test code may be in this same file or in a separate “.py” but make sure any test code is in the body of an “if” statement: `if __name__ == "__main__":` Note that some points may be deducted if your code could obviously be written in a “more Pythonic” way.

For this assignment you will write code to decode the “Beale Cypher.” (Google it.) Your code will consist of the following functions. (The last is for the “A” grade).

1) A function with prototype `def readCodeFile(filename = 'BealeCodeFile.txt'):` which will read the code file containing the numbers, remove any non-digits and return a list of “int”. Note that multiple numbers are on each line of the file.

2) A function with prototype `def readBaseDoc(filename = 'D0Ioriginal2a.txt'):` which will read the base document and return a list of words to be used for decoding the message. Specifics: change any hyphens to spaces, split into words, remove any characters that are not upper or lower case letters, change to lower case and return as a list of words in the order given in the file (with no empty strings).

3) A function with prototype `def decodeMsg(thewords, thenums):` which will read the numbers and words from the previous two routines and will return a string of the decoded message, (taking the first letter of the indicated word). Note that `thewords` is one long string without spaces. Also note that the `thenums` are word numbers with the first word corresponding to number 1.

4) A function with prototype `def findwords(textstring, wordfile = 'mostcommon3000.txt', minlen=1):` which will read a file of “common” words (name given by `wordfile`), containing one word per line, and will return a tuple. The first element of the returned tuple is a dictionary whose keys are the words that were found in the text string and the values are an `int` indicating how many times that word appeared in the string. The second tuple element is the sum of all the values (total words found. Note that some words may be contained inside or overlap other words. `minlen` is the minimum length word that is considered.

5) A function with prototype `def dictsort(d):` which will return a list of key-value pairs of the dictionary as tuples (key, value), reverse sorted by value (highest first) and where multiple keys with the same value appear alphabetically (lowest first).

6) A function with prototype `def encodeMsg(thestring, thewords):` which will encode a message (character string) in the style of the “Beale Cypher”, returning a list of the numbers. Each character in `thestring` that corresponds to a letter (‘a’ to ‘z’) results a number in the list corresponding to the index (plus 1) of a word in `thewords` starting with that letter. If more than one word is possible, a random one is selected. Note that spaces and any other non-letters in `thestring` are ignored, as is the case of the letter (upper or lower).