

AC50002 - Programming Languages for Data Engineering

SHEETAL UMESH || 2544848 || PYTHON ASSIGNMENT

Design of the assignment:

- Reading Values of the letters -

Reads values from the file that values_file_path has supplied.

The function read_values(values_file_path), creates a dictionary called sorted_postn_values, which is sorted by value and maps letters to their equivalent values.

Gives back the dictionary after sorting.

- Function to generate abbreviations -

word_least_letter_checker(theword, sorted_postn_values): Takes into account certain circumstances to determine the lowest scoring letter and its score within a word.

sorted_postn_values, name, and least_score_checker_updated: Tracks the lowest scoring letter and its score for every word in a name by using the word_least_letter_checker to each word.

name_abbreviator(name, sorted_postn_values): Utilizing predefined rules, this function generates scores and abbreviations for names.

- Main execution (Abbreviation process) -

The function abbreviator (path, sorted_postn_values) takes names from a file, performs some processing and cleaning on them, uses name_abbreviator to calculate scores and abbreviations, and then writes the results to an output file.

For convenience of processing, the code groups the output data into dictionaries and lists before writing the results to an output file.

The user is prompted to enter the data filename by the script.

To get the letter values, call read_values.

Use the abbreviator function to produce scores and abbreviations, then save the results to an output file.

- Output directory handling -

To keep the generated output files, the code creates an "output" directory if one doesn't already exist.

The path to the result file is printed by the code once processing is complete.

An output file with the original names and their abbreviations is produced by the application.

Evidence of Testing:

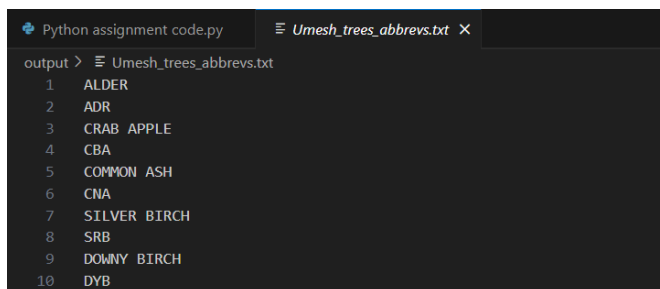
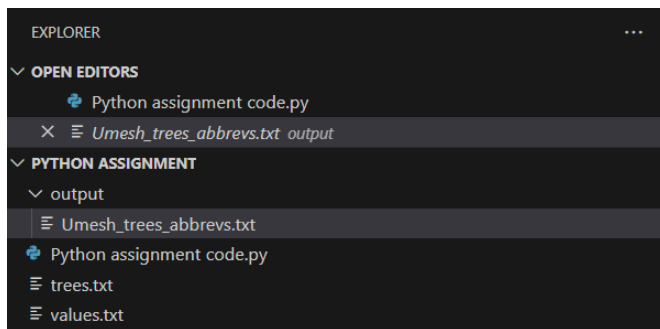
Given values.txt-

{ 'A'-25, 'B'-8, 'C'-8, 'D'- 9, 'E'-35, 'F'-7, 'G'-9, 'H'-7, 'I'-25, 'J'-3, 'K'-6, 'L'-15, 'M'-8, 'N'-15, 'O'-20, 'P'-8, 'Q'-1, 'R'-15, 'S'-15, 'T'-15, 'U'-20, 'V'-7, 'W'-7, 'X'-3, 'Y'-7, 'Z'-1 }

Testing values –

Word	Expected abbreviation	Resulting abbreviation	Scores of each letter of the abbreviation	Total score
ALMOND-LEAVED WILLOW	ALW	ALW	0+0+0	0
MIDLAND HAWTHORN	MDH	MDH	0+5+0	5
OF	" "	" "	Not applicable	Not applicable

SCREENSHOTS –



The Python program to get the word abbreviations has been attached in the Github repository.

<https://github.com/Sheetal-2912/-2544848-Python-assignment.git>