User’s Documentation

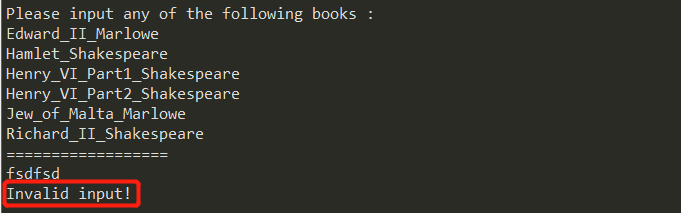
*Kewen Deng, 29330440, FIT9133 Assignment 3*

1 Instruction

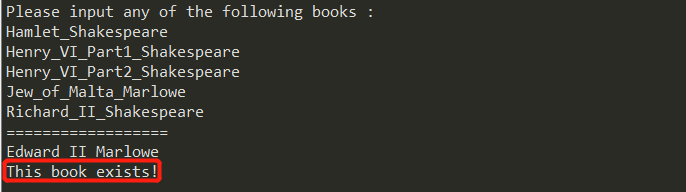
This program contains 5 files named as preprocessor\_29330440.py, character\_29330440.py, word\_29330440.py, visualiser\_29330440.py and main\_29330440.py. Each python file gives answer to one single task. Users can have a clear view of the 5 files by following steps:

1. Open main\_29330440.py.
2. Click on ‘Run’.
3. Input name of book according to the hints
4. Hints change and input another book
5. After entered all the books, wait for processing.
6. Check answers. (Both .csv files and pictures)

It is necessary to mention that if users entered a wrong name, the program would say the input is invalid.



Also, if users entered a same name for twice, the program would say the name already exists.

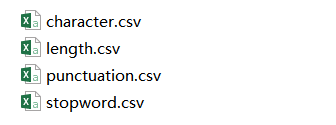


2 Limitation

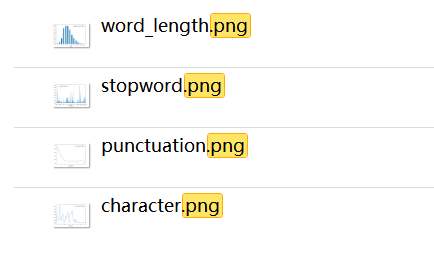
1. Before all the 6 books has been entered, the user can not quit the program naturally.
2. The punctuation in this program has been defined as: ',', '.', '?', '!', ';', ':', '\'', '-'.
3. Words like ‘Who’s’, ‘We’ve’ has been considered as a whole word like who’s and we’ve.
4. In the analyze of characters, this program would only consider the characters of English letters(A-Z) and numbers(0-9).
5. In the analyze of words, this program could just count the occurrences of different spellings of words, regardless of their meanings. For example, the different tenses and plural numbers of a same word might be considered as different word.
6. In the analyze of visualization, this program uses line plot to show characters, punctuation and stop-words, and uses bar plot to show word length.
7. It might take users long time waiting for the output, please be patient.

3 Assumption

In order to make sure that all the outputs are user-friendly, the 4 frequencies of different tasks have been saved as different .csv file in the same directory.



The 4 different pictures plotted have been saved as well.



4 Demonstration

In order to further illustrate where is the program processing, the program would kindly show you what has been done like:

