

[Belge başlığı]

[Belge alt konu başlığı]



[Tarih]

[Şirket adı]

[Şirket adresi]

**TABLE OF CONTENTS**

Chapter 1: Introduction………………………………….………………………..…..2

Chapter 2: Structure of the Project and Algorithm………………………...…………..1

Chapter 3: Functions…………………….…………………………………................1

Chapter 4: Use Case Diagram……………………………………………………...…1

Chapter 5: Project Results……………………………….……………..……..............1

References…………………...……………………….……………..……...................1

**CHAPTER 1**

**INTRODUCTION**

In the project, we planned to do some operations by converting the E-Books downloaded from the internet as .pdf files to .txt format. These processes are to see the 20 most recent words in a book or to see the common and different words between 2 books. The project was completed 90% successfully according to the homework paper and I do not think that the project has deficiencies in the completed part and I could not see it in the tests I did.

The 10% of the project that could not be completed was the E-Book download phase from the internet. Since the methods I tried in this section are not successful and sufficient, I preferred not to put them in my code.

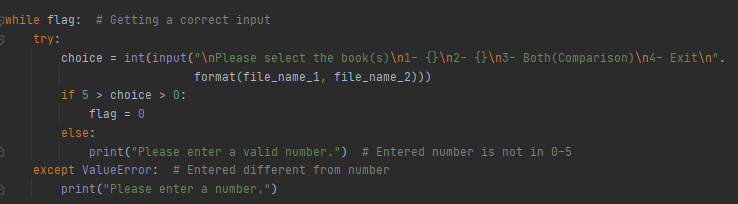
**CHAPTER 2**

**STRUCTURE OF THE PROJECT AND ALGORITHM**

The program consists of 3 main stages: Data entry and verification, selection of the appropriate option for data entry and preparation, main function call. Let's talk about these 3 stages;

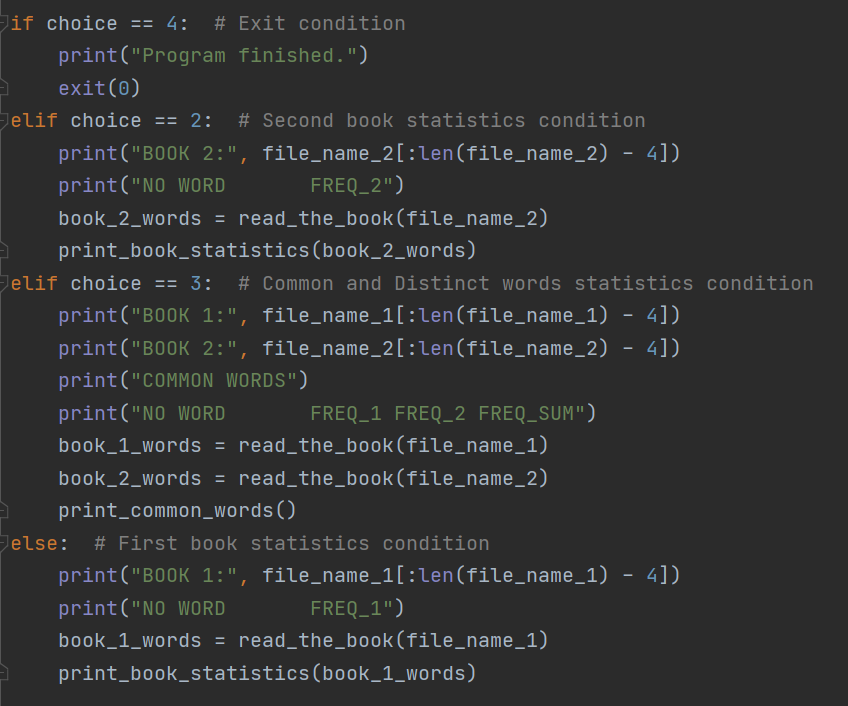
**Data Entry and Verification**

In this part, the user is expected to select a valid option from the menu presented to him/her. There are 4 options offered to the user, I expect her/him to select any of the numbers from 1 to 4. I do not allow the user to move on to the next step until there is a valid data entry. Here is the code ;



**Selection of the Appropriate Option for Data Entry and Preparation**

In this section, we choose the appropriate option according to the correct data entry from above and make a small preliminary preparation for the desired situation. Here is the code ;

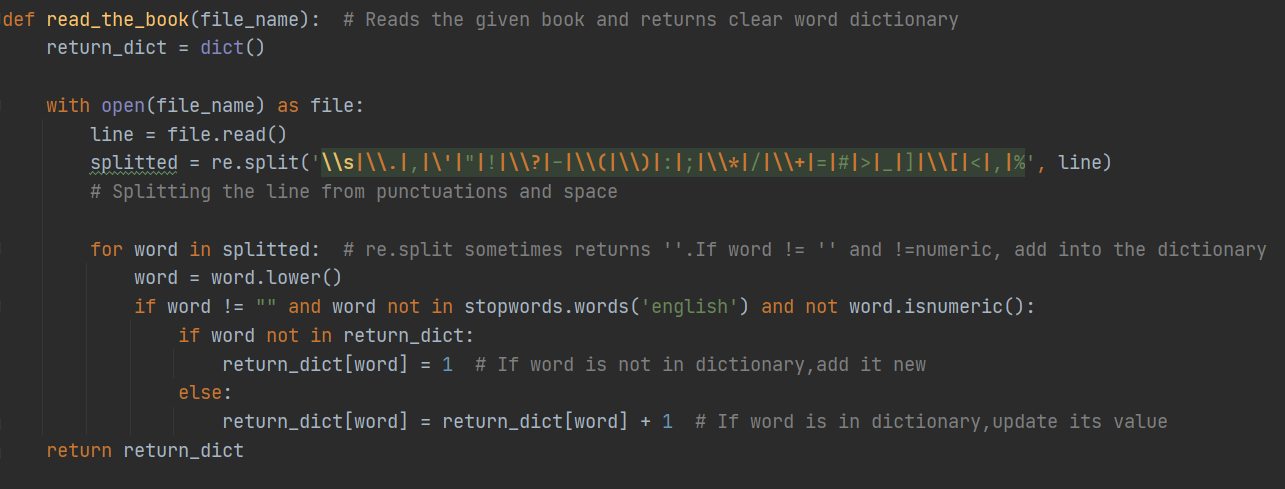


**Main Function Call**

This is the part with the most transactions and the most complex. If the user selects 4, the program closes. If he chooses 1, the statistics of the first book entered into the system are written on the screen. If he selected 2, the statistics of the second book entered into the system are written on the screen. If he chooses 3, the common and different words of the 2 books are printed on the screen. Let's take a closer look at these functions.

**CHAPTER 3**

**FUNCTIONS**



Our first function is to read\_the\_book. It takes the name of the .txt file as a parameter. Reads the file and outputs a dictionary of words and their frequencies. This function is always called except option 4.

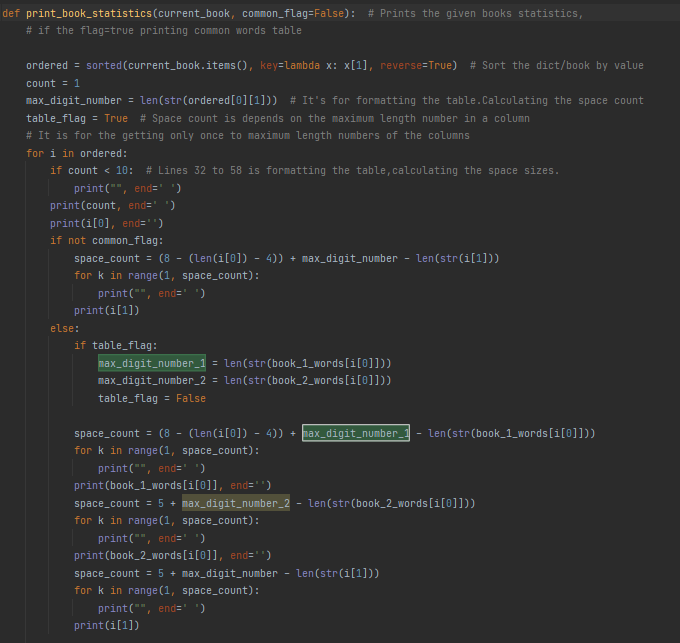
The stages of the function are as follows:

1 - Read one line from the file.

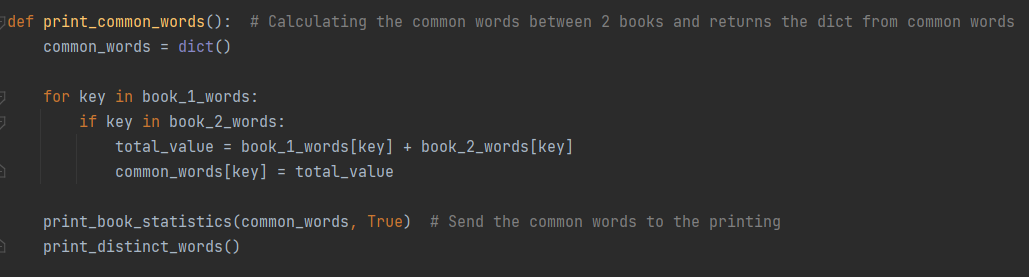
2 - Use the split function of the Re library to remove all punctuation and drop it into the temporary list.

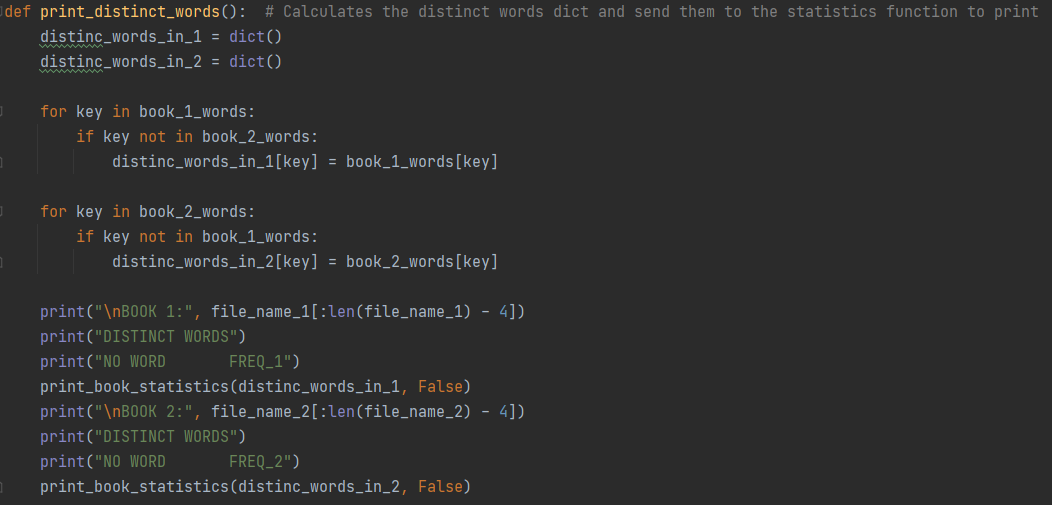
3 - Check each element in the temporary list to see if it is an empty string or contains a numeric character or it is in the stop word list from nltk library. If it meets these requirements, add it to the results dict.

4 – Return the result\_dict.



Function 2 is print book statistics. The functioning of the function is actually very simple. It sorts the incoming dictionary by value and then prints it on the screen. But the operations required to look good on the screen make the function look long. The function has 2 parameters. The first parameter is the book. It takes the dictionary representing the book and sorts it. Then it calculates the required number of spaces and prints it on the screen. The second parameter of the function is the table flag. If I want to print a table of common words with this function, I am sending this flag true. It is default false.



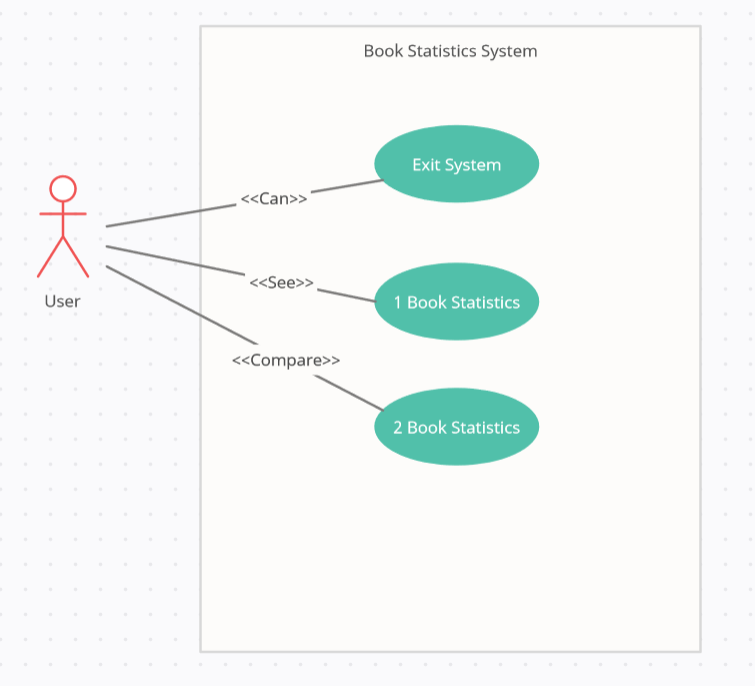


The last 2 functions are our functions that find common and different words. They take two dictionaries created by the file reading function, compare them with each other and produce a dictionary that includes common or different words. Then they complete their operations by sending this dictionary to print\_book\_statistics.

**CHAPTER 4**

**USE CASE DIAGRAM**

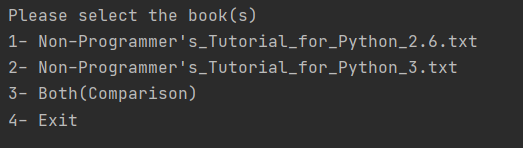
The diagram is small because there are not many functions that the user can do in the project. The user can only perform 3 basic operations.



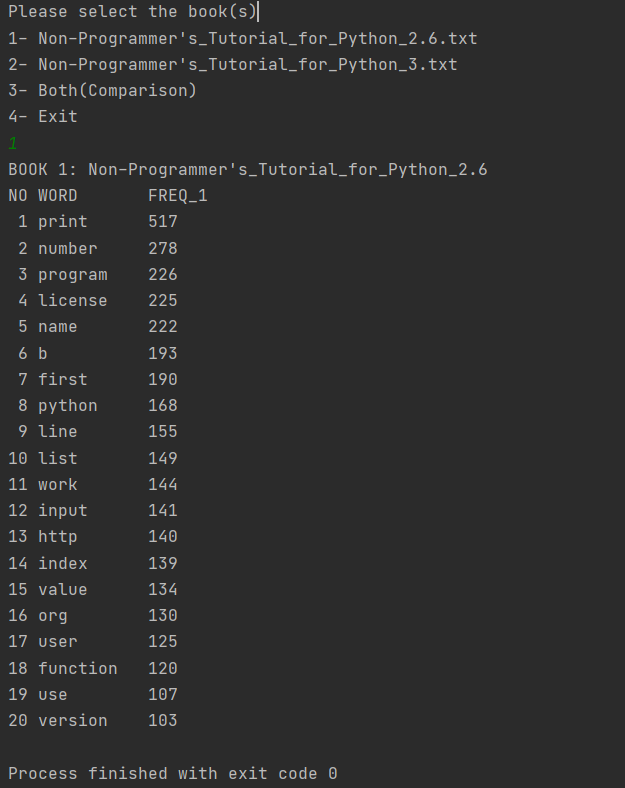
**CHAPTER 5**

**PROJECT RESULTS**

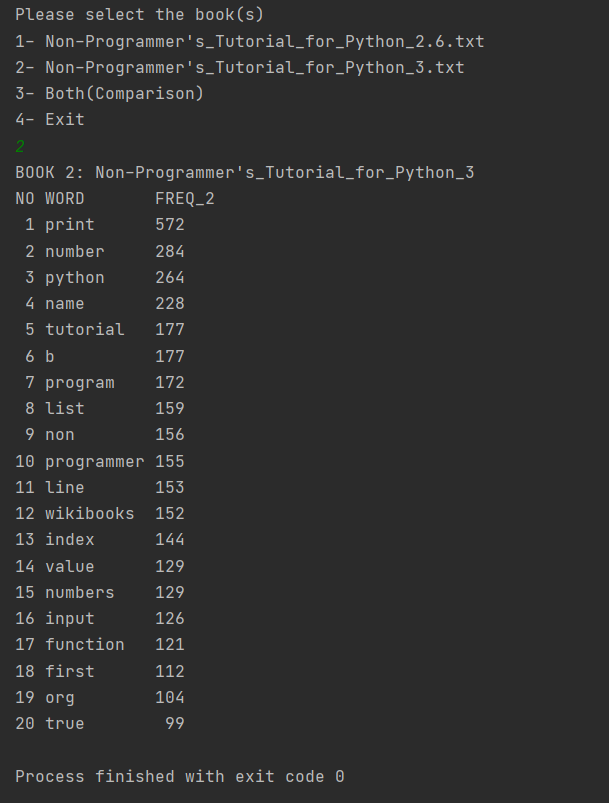
**Menu**

****

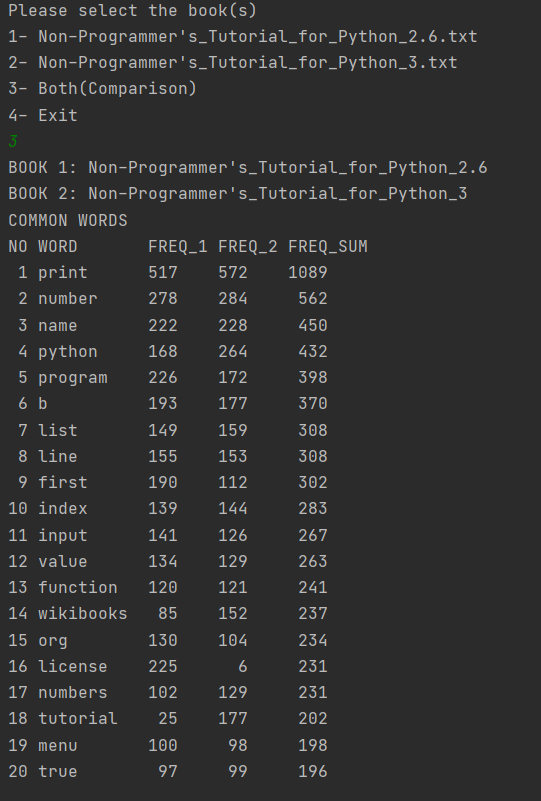
**Option 1**

****

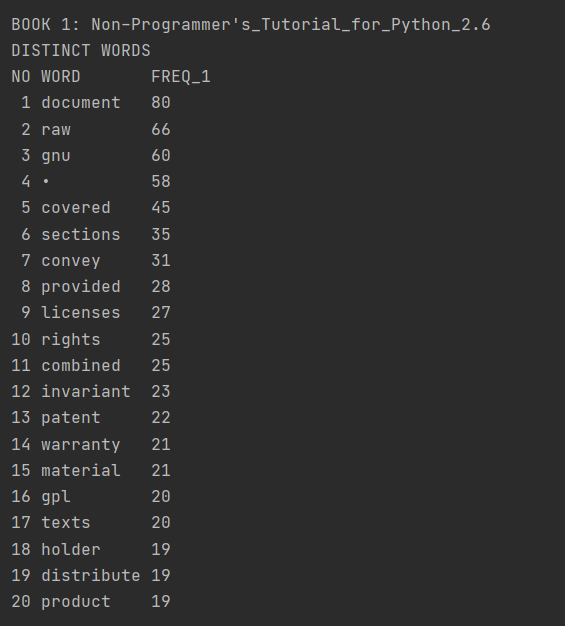
**Option 2**

****

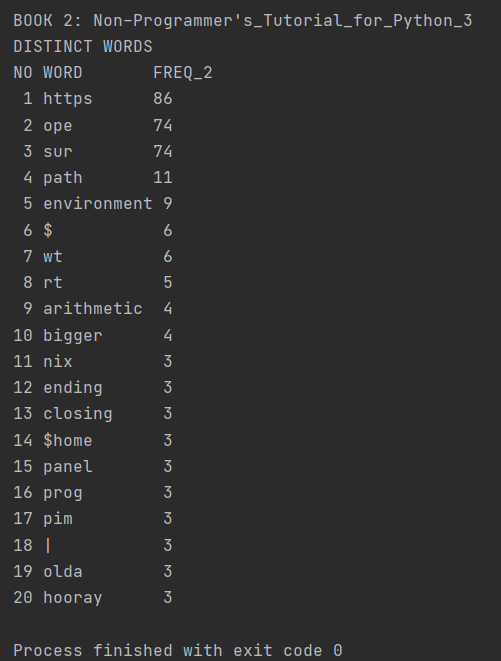
**Option 1/3**

****

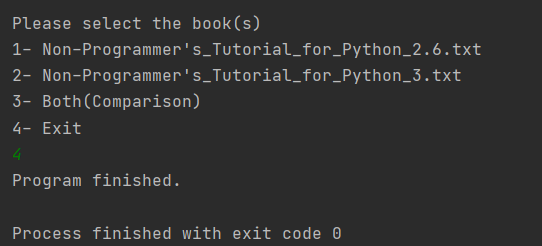
**Option 2/3**

****

**Option 3/3**

****

**Option 4**

****

**REFERENCES**

*Lib/Re*. (n.d.). Retrieved January 4, 2021, from

<https://docs.python.org/3/library/re.html>

*StopWords*. (n.d.). Retrieved January 4, 2021, from

https://www.nltk.org/\_modules/nltk/corpus.html