

# Final Programming Assignment

Start Assignment

**Due** Dec 4 by 11:59pm    **Points** 20    **Submitting** a file upload    **File Types** py, doc, and docx  
**Available** Nov 20 at 1:30pm - Dec 6 at 12pm 16 days

For the final programming assignment, you will be working individually to scrape information from the [classicshorts.com](http://classicshorts.com) (<http://classicshorts.com/>) website.

## Part 1

Your program should start by asking the user to name a short story by path on the [classicshorts.com](http://classicshorts.com) (<https://www.classicshorts.com/>) website. For example, if the user enters 'aos' then the program will access the website <http://classicshorts.com/stories/aos.html> (<http://classicshorts.com/stories/aos.html>). If the user enters an invalid short story (path, such as 'aps'--note the typo), the program will say "Short story not found" and quit. (Hint: try-except)

If a story is found, state that the story was found and how many vocabulary terms there are.

## Part 2

The program will store the vocabulary words in such a way that the user can optionally update the word's definition. **Eliminate duplicates.** By default, the definition should be an empty string (""). It will ask the user if he/she would like to update any definitions. If the user enters 'Y' (for yes), then program will ask for the term and definition. Afterwards, the program will ask the user if he/she would like to update additional definitions. If the user answers 'Y' again, then the program will repeat this step until the user enters 'N' (for no).

Note, for words that already have definitions, the program **MUST** warn the user and state what the current definition is. For words not in the vocabulary list, the program will simply state that the word was not found and ask if the user would like to update additional definitions until 'N' is encountered.

## Part 3

As a final step, the program will ask the user for a file name to save the vocabulary words. The text must be saved using format() such that there are enough spaces for the longest vocabulary word, followed by a space, a dash, and a space before the definition. The vocabulary words should be left aligned. See below for an example.

Here is an example of what the program should look like to the user, with red as the user-input:

```
What's the short story path? araby
Short story found. There are 20 unique vocabulary words.
Would you like to update a definition (Y/N)? Y
Term: sombre
Definition: gloomily dark
Would you like to update a definition (Y/N)? Y
Term: gauntlet
Definition: glove
Would you like to update a definition (Y/N)? Y
Term: gauntlet
WARNING! gauntlet is currently defined as 'glove'
Definition: a glove
Would you like to update a definition (Y/N)? Y
Term: proffer
ERROR! Term not found.
Would you like to update a definition (Y/N)? N
What would you like to save the file as? myvocab1.txt
File saved!
```

Here is an example of the corresponding output file: [myvocab1.txt](https://my.uclaextension.edu/courses/36536/files/5638413/download?download_frd=1) ↓ ([https://my.uclaextension.edu/courses/36536/files/5638413/download?download\\_frd=1](https://my.uclaextension.edu/courses/36536/files/5638413/download?download_frd=1))

Note that even though there are many instances of 'bazaar' in Araby, it is only counted once in the unique vocabulary words list.

Here is another example of the program:

```
What's the short story path? araby
Short story found. There are 20 unique vocabulary words.
Would you like to update a definition (Y/N)? N
What would you like to save the file as? myvocab2.txt
File saved!
```

Here is an example of a short story with no vocabulary words:

```
What's the short story path? souls
Short story found. There are no vocabulary words.
```

Here is an example of a short story that could not be located:

```
What's the short story path? python
Short story not found.
```

Hints:

- All vocabulary words are associated with a dictionary.com hyperlink. Use the dictionary.com hyperlink to extract the vocabulary work when you first go through the original [classicshorts.com](https://www.classicshorts.com/) (<https://www.classicshorts.com/>) page.
- Not all key terms refer to dictionary.com. For example, "Hongkong" in [Beyond the Wall](https://www.classicshorts.com/stories/btw.html) (<https://www.classicshorts.com/stories/btw.html>) links to a Wikipedia page. It does not need to be added to our vocabulary list.
- DO NOT TRY TO ACCESS ALL OF THE SHORT STORIES AS YOU COULD GET LOCKED OUT OF THE WEBPAGE
- KEEP IT SIMPLE!

Code Hints:

The following code may serve as a helpful example for this assignment, but is not intended to be included in the assignment. In other words, you'll need to modify it.

```
from bs4 import BeautifulSoup
from urllib.request import urlopen

response = urlopen('https://www.classicshorts.com/stories/aos.html')
soup = BeautifulSoup(response, 'html.parser')

atag = soup.findAll('a')
for tag in atag:
    if "dictionary" in tag.get('href'):
        print(tag.get('href'))
        print(tag.get('href').replace("http://dictionary.reference.com/browse/", ""))
```

### EXTENSION (5 points):

This is your own customization of the original assignment. Think of 3+ that you can make the current program more flexible or user-friendly. These changes cannot be coding efficiency improvements (as your original program should already be at its best). The changes are open-ended and MUST be visible to someone using your program. For some improvements, it might help to look in the upcoming chapters. For the "extension" portion of this assignment, you will need to download and complete this "Programming Extension Form": [extension\\_template.docx](https://my.uclaextension.edu/courses/36536/files/5638400/download?download_frd=1) ↓ ([https://my.uclaextension.edu/courses/36536/files/5638400/download?download\\_frd=1](https://my.uclaextension.edu/courses/36536/files/5638400/download?download_frd=1))

In all, you should submit three files:

1. Original assignment exactly as prompted above (.py file)
2. [Programming Extension Form \(.docx file\)](https://my.uclaextension.edu/courses/36536/files/5638400/download?download_frd=1) ↓ ([https://my.uclaextension.edu/courses/36536/files/5638400/download?download\\_frd=1](https://my.uclaextension.edu/courses/36536/files/5638400/download?download_frd=1))
3. Modified .py file with changes as listed in the Extension Form

Keep in mind the deadline for this assignment. Because this is a "Final" Assignment, late work will not be accepted.

Final Programming

Criteria	Ratings				Pts
Part 1 Task	<b>5 pts</b> <b>Full Marks</b>  Fully addresses all required elements of the task, without error	<b>3 pts</b> <b>Half Marks</b>  Mostly addresses required elements of the task. The program runs, but the program is unable to identify valid short story links with vocabulary, valid short story links without vocabulary, and/or invalid short story links.			<b>0 pts</b> <b>No Marks</b>  5 pts
Part 2 Task	<b>4 pts</b> <b>Full Marks</b>  Fully addresses all required elements of the task, without error		<b>2 pts</b> <b>Half Marks</b>  Mostly addresses required elements of the task. The program runs, but vocabulary definition update does not run exactly as required in the prompt.		<b>0 pts</b> <b>No Marks</b>  4 pts
Part 3 Task	<b>3 pts</b> <b>Full Marks</b>  Fully addresses all required elements of the task, without error		<b>2 pts</b> <b>Half Marks</b>  Mostly addresses required elements of the task. The program runs, but the saved file has some error.		<b>0 pts</b> <b>No Marks</b>  3 pts
Readability	<b>3 pts</b> <b>Full Marks</b>  Code is well-commented and easy to follow. Variables are appropriately named. Code is free of redundant (unused) variables and lines. Loops and functions are used to avoid retyping multiple lines.				<b>0 pts</b> <b>No Marks</b>  3 pts
Extension	<b>5 pts</b> <b>Full Marks</b>  The programmer submits both a original .py file and an extension .py file. It is clear in the header file which file is the original and which file is the extension. The programmer also submits an accompanying "Program Extension Form" clearly detailing changes to the programs, and why the changes make the extension program even better than the original. The changes are exceptional and well-tested for a variety of user-inputs.	<b>3 pts</b> <b>Above Average</b>  The programmer submits both a original .py file and an extension .py file. It is clear in the header file which file is the original and which file is the extension. The programmer also submits an accompanying "Program Extension Form" detailing changes to the programs, but it is not clear why the changes make the extension program even better than the original. There about 2-3 changes to the original program, and the program can handle a variety of user-inputs.		<b>1 pts</b> <b>Meets Extension Requirement</b>  The programmer submits both a original .py file and an extension .py file. It is clear in the header file which file is the original and which file is the extension. The programmer also submits an accompanying "Program Extension Form" detailing changes to the programs, but it is not clear why the changes make the extension program even better than the original. There is only one minor change to the original program.	<b>0 pts</b> <b>No Marks</b>  5 pts
Total Points: 20					