CSCI 1100 — Computer Science 1 Homework 6
Files and Sets

Homework Overview

Exterminate!

This homework is worth **90 points** toward your overall homework grade and is due **Thursday November 5, 2015 at 11:59:59 pm**. It consists of one single program. Please review the rules about excess collaboration from HW 3 to make sure you are turning in a program that is your own.

As always, you are expected to match your output to the one given on the submission server. Your grade is going to depend almost completely on code correctness and structure, very little on the actual formating of the output (see below).

You are not expected to use dictionaries to answer this homework (unless you try the bonus part), but you can. It is your call. Overall, you may find it conceptually easier not to use dictionaries and read the input file a few times to construct your answer. You must use sets. It will become quite obvious why as you start working on the program.

As mentioned in previous homeworks, you program must have the following structure:

```
## all imports
import textwrap ##if you use it, see below

## all functions first
def function1(x):
    return x + 1

def function2(y):
    return y -1

## all your main code after this if statement
if __name__ == '__main__':
    z = 10
    print function1(z)
    print function2(z)
```

This homework brings together many little things you have learnt in different parts of this class. But, we will not tell you how to do the whole thing. My solution is a total of 65 lines including white lines for readability. So, it is not a long program if you plan it properly. I recommend constructing it in little pieces.

A final word: This file is from 2013, so it will not have the most recent information. If you are a Dr. Who fan and would like to help me update it, it would be great. We will maintain it for the next generation of CS1 students.

Problem: Villains and stories

You are given a file called `DrWhoVillains.tsv`. It is a TAB separated file containing villains in a science fiction show called Dr. Who. The show is actually known for its extremely low budget and

silly creature effects. Look at the picture of a Dalek (one of the main villains) above to see this for yourself. Each line of this file is a different villain, containing the following information (separated by TABs):

```
1. Villain: name of the villain (e.g. Dalek)
2. Year first: the first year this villain was introduced
3. Year last: the last year this villain was shown
4. Doc. no.: the list of doctor ids with this villain, seperated by commas
5. Doctor actor: the list of actors for each doctor above
6. Episodes: number of episodes featuring this villain
7. Stories, total: number of stories featuring this villain, each story may
    span many episodes
8. Motivation (invasion earth, kill humans, etc): the description of the main
    motivation of this villain
9. Story titles: the titles of the stories that the villain was involved with,
    a list of strings separated by commas.
```

Your program must do the following:

- Read through the file to find the villain names and the (set of) stories for each villain. Be careful, the file is manually created and may have repeated names, extra spaces before after names, etc. You must use sets to find unique stories.
- Find the top 10 most popular villains: those with top 10 highest number of stories listed in the file and display them to the user, numbered 1 to 10, in the decreasing order of popularity. Then, ask the user for a number.
- If the user enters -1, the program exits.
- If the user enters something other than a number between 1 to 10, or -1, the program displays the top 10 list again and asks for another input.
- If the user enters a valid option, numbered between 1 to 10, then you find the corresponding villain and display the following:
  - The number of stories with this villain
  - The names of all other villains that were in a story with this chosen villain, in alphabetical order.
  - The names of all story lines that featured only this villain, and no other villain. If none, you should display there is no such story line.

  Once you are done displaying this, you will print again the top 10 and ask for the next input, until user enters -1.

One thing you want to do is to display the names of matching villains and stories in a reasonable looking paragraph. You do not have to match the output in the submission server exactly. You should however try to break the output into multiple lines of similar (but not necessarily identical) length. We will not take points off for breaking lines mid-word.

There are many ways to do this. For example, what we did in the last homework is one way. My solution involves a module called `textwrap` which wraps text into a list of multiple lines. Try the following to see if you can figure out how to use this module.

```
>>> import textwrap
>>> x = '+'*200
>>> textwrap.wrap(x,40)
```

Deliverables – Final Check

Here is a summary of all the requirements for this homework. Double check this before you submit:

- Your program must use sets, that will also help you quite a lot.
- Submit a single file called `hw6.py` that assumes the existence of a text file called `DrWhoVillains.tsv`.
- The main loop of the program should be able to handle incorrect entry, repeated calls to display the top 10 villains, and exit when -1 entered.
- Print all input that you read immediately.
- The outputs containing multiple items like villains or stories should be a reasonable looking paragraph, wrapping long lines into multiple lines. But the output does not have to be a perfect match for the submission server. Lines do not have to be the same length.
- Your program should print the villain names and story names in alphabetical order.
- Your program should have the required structure.

Expected output

Below you can see the expected functioning of this program with the file we gave you (note: we might change the file in the homework submission server):

```
1. Daleks
2. Cybermen
3. Master (the)
4. Dalek Supreme (Supreme Dalek), inc Progenitor
5. Sontarans
6. Ice Warriors (inc benevolent appearences)
7. Davros
8. Silurians (exc Sea Devils)
9. Kovarian, Madame
10. Cyber-leader
Please enter a number between 1 and 10, or -1 to end
Enter a villain ==> stop
stop

1. Daleks
2. Cybermen
3. Master (the)
4. Dalek Supreme (Supreme Dalek), inc Progenitor
5. Sontarans
6. Ice Warriors (inc benevolent appearences)
7. Davros
8. Silurians (exc Sea Devils)
9. Kovarian, Madame
10. Cyber-leader
Please enter a number between 1 and 10, or -1 to end
Enter a villain ==> 5
5
```

```
Sontarans in 9 stories, with the following other villains:
==================================================
     Alliance - onscreen - Daleks, Cybermen, Sontarans & Autons, (cameos
     Silurians, Sycorax, Judoon, Hoix, Roboforms, microcameos Blowfish,
     Uvodni, Weevils), Atraxi, Borusa, Chessene of the Franzine Grig,
     Cybermen, Dalek Supreme (Supreme Dalek), inc Progenitor, Daleks,
     Dastari, Joinson, Gray, Steve, Harris, Private Carl, Headless Monks,
     Irongron, Kovarian, Madame, Manton, Colonel, Martha Jones- Clone,
     Master (the), Naismith, Joshua, Rassilon, Rattigan, Luke, Red
     Carnivorous Maw, Shockeye of the Quawncing Grig, Silurians (exc Sea
     Devils), Staal, Stike, Vardan, Varl, Varla

The stories that only features Sontarans are:
==================================================
     The Sontaran Experiment


1. Daleks
2. Cybermen
3. Master (the)
4. Dalek Supreme (Supreme Dalek), inc Progenitor
5. Sontarans
6. Ice Warriors (inc benevolent appearences)
7. Davros
8. Silurians (exc Sea Devils)
9. Kovarian, Madame
10. Cyber-leader
Please enter a number between 1 and 10, or -1 to end
Enter a villain ==> 7
7
Davros in 6 stories, with the following other villains:
==================================================
     Dalek Supreme (Supreme Dalek), inc Progenitor, Daleks, Lytton,
     Commander Gustav, Movellans, Nyder, Sharrel

There are no stories with only this villain
==================================================

1. Daleks
2. Cybermen
3. Master (the)
4. Dalek Supreme (Supreme Dalek), inc Progenitor
5. Sontarans
6. Ice Warriors (inc benevolent appearences)
7. Davros
8. Silurians (exc Sea Devils)
9. Kovarian, Madame
10. Cyber-leader
Please enter a number between 1 and 10, or -1 to end
Enter a villain ==> -1
-1
Exiting
```