Web scrapping script 30/7/2015

The script will scrape a website, parse the information obtained and create cache files should the internet become unavailable. The cache will be refreshed for each search so cached files are upto date.

Requirements

Python

OS of your choice

Installation

Create a directory and place the Homework.py inside that directory. Homework\_comment.py is a commented version of Homework.py for other developers to use.

The script will create cache files inside this folder should web access become available. No results will be available should web access not be available for an initial run of the script.

Command line format

Homework.py <name> <year1> <year2>

Changelog.

0.1 First revision

0.2 Fixed issue where script would not search <year2> but search <year2> -1

0.3 Rounded result of mean to 2 decimal places

0.4 Added caching of web pages

0.5 Minor code changes to parsing

0.6 Added error handling

0.6a Added visible output for errors

0.7 Changed code to hide results obtained while a search was in progress.

0.9 Created commented version of script

0.9a Created user friendly error messages

1.0 First Release