David Hall
Project 1: Stepwise Refinement
CS333
10/26/16

Design document for Project Gutenberg Book Indexing Application

- create constants
    - HTML tags and preformatted headers
        - individual tags like <h2> <ul> and <a> with closing tags
        - formatted sections like HTML "wrapper"
        - formatted lines such as for index terms
        - all formatted constants should include variable patterns
- create pages
    - initialize ifstream bookIn to file name given in command args
    - create vector to hold page strings/possibly "page" objects
    -
    - while not end of file
        - create empty string/page
        - set counter linesIn to 0
        - for linesIn < MAX_PER_PAGE AND NOT end of file, read in page
            - read each line with getline()
            - increment counter
        - add HTML "wrapper" to page
            - find title
            - check if page is first or last page
                - Check if vector is empty for first
                - check if readIn is at end of file for last
            - add first part of wrapper with title and links to previous,next,first, and last, excluding as necessary
            - add closing tags to end of string
        - output files to individual HTML files
        - add string/page to vector
- create index page
    - create list of letter sections A-Z
    - create data structure of index terms
        - iterate through book:
        - For each page in book vector:
            - create new entry if:
                - Preceded by '.' or space AND
                - Over 3 characters
            - Convert entry to lower case
            - Record page #

      - Increment a counter of how many times occured
        - increment only once per page

    - iterate through data structure and discard entries with over PAGE_THRESHOLD repeat count
  - sort list of index terms alphabetically (or add to list alphabetically)
    - place index terms in iterable data structure organized by Letter,then term, then page numbers
- output index page to its own HTML file
  - Create header of page with title
  - Create list of letter headers in <h2> format
  - for each letter:
    - loop through index terms and add entries for each term into <ul> </ul> list
    - add pages for each entry to <a></a> list of links to each page
  - add ending code
  - save file to "indexPage.html"