

```

1  #include <iostream>
2  #include <fstream>
3  #include <string>
4  #include <cstring>
5  #include <alloca.h>
6  #include <map>
7  #include <set>
8
9  using namespace std;
10
11 //removes all the nonalpha characters and replace them with spaces.
12 char* cleanUp(char* line, int lineLength){
13     int i ;
14     for(i=0; i<lineLength; i++){
15         if(!isupper(line[i]) && !islower(line[i])){
16             line[i] = ' ' ;
17
18         }
19     }
20     return line ;
21 }
22
23
24
25 int main(int argc, char* argv[])
26 {
27
28     map <string, set<int> > invertedIndex; //created a map with set Integers
29     set <string> sortedKey ; //creates a string set which we are going to use to sort
    alphabetically
30     char* pointer ;
31     char* pointer1 ;
32     char* str ;
33     char* str1 ;
34     char* newLine ;
35     int lineSize ;
36     string line ;
37     string fileline ;
38     int index = 0 ;
39     set<int> currentWord ;
40
41     ifstream inputText(argv[1]) ; //takes the second arg as filename
42
43
44     while(inputText.is_open()){ //checks if the input.txt file is_open
45
46         while(getline(inputText, line)){
47             str = strdup(line.c_str()) ;
48             pointer = strtok(str, " ") ; //space as delimiter
49             while(pointer != NULL){ //while there are still tokens, next filedocument
50                 if(pointer != " "){ //if there is a file.
51                     ifstream file(pointer) ; //opens up file
52                     //process and put into map
53                     while(getline(file, fileline)){ //gets the line inside textfile.
54                         lineSize = fileline.size() ; //gets length of line
55                         str1 = strdup(fileline.c_str()) ;
56                         newLine = cleanUp(str1, lineSize) ; //cleaned up line
57
58                         //tokenizes the cleanedup line.
59                         pointer1 = strtok(newLine, " ") ;
60                         while(pointer1 != NULL){ //while there are more words.
61
62                             sortedKey.insert(pointer1) ; //inserts the string into
                                sortedKey string set for later traversal output use
63
64                             currentWord = invertedIndex[pointer1] ;
65                             currentWord.insert(index) ;//inserts the document number
                                into currentword.
66                             invertedIndex[pointer1] = currentWord ; //inserts into map

```

```

67
68
69             //test
70             //cout << "Inserting word " << pointer1 << " from document
             " << index << "!" <<endl ;
71             pointer1 = strtok(NULL, " ") ; //points to next token
72         }
73     } //end while
74
75     file.close() ;
76     index++ ; //keeps track of filecount/index
77 }
78 pointer = strtok(NULL, " ") ; //reassigns pointer so it goes to the
    next textfile.
79 }
80 }
81 inputText.close(); //closes file
82 }
83
84 set<string> :: iterator it; //creates an iterator of set<string> called it
85
86 for (it = sortedKey.begin(); it != sortedKey.end(); it++) {
87     string key = *it;
88     cout << key << ":" ; //prints out the key already sorted in set string
89
90     set<int> :: iterator iter; // creates an iterator for set<int>
91
92     for (iter = (invertedIndex.find(key)->second).begin(); iter != (invertedIndex.
        find(key)->second).end(); iter++) {
93         cout << " " << *iter ; //prints out the value using the set <string> above
            to search for.
94     }
95     cout << endl ;
96 }
97
98 return 0;
99 }
100
101
102
103
104
105
106
107
108

```