Project 1

Description

For this project you will read in words from a given file into your program. A word consists of any non-whitespace character. An example would be this!is@A&Word!. Each word should be cleaned of punctuation and converted to lowercase so that only lowercase letters remain. The result of cleaning the above example is thisisaword. If you read in a word that consists of only punctuation you should ignore it.

Once you have cleaned all of the words, you should display the total number of words read. Your next task is to determine how many unique words appear in the file and display that amount. Finally, you will then prompt the user to enter a word and report how many occurences occur in the file. You will continue to prompt the user until they enter the EOF character.

Requirements

Please carefully read the following requirements:

- You must supply a makefile that supports the commands "make" and "make clean"
- You must use C++ streams for all I/O (You cannot use Scanner from cs100)
- You must format your output as shown in the example below.
- You must do your own work, you must not share code.
- You must submit the project by September 1st before the beginning of class.
- Your must submit your project in a zipfile as specified under submission

Example

Here is a given document sample.txt

```
Cryptography is both the practice and study of the techniques used
to communicate and/or store information or data privately and securely,
without being intercepted by third parties. This can include processes
such as encryption, hashing, and steganography. Until the modern era,
cryptography almost exclusively referred to encryption, but now
cryptography is a broad field with applications in many critical areas
of our lives.
```

You must be able to run the program as shown below and get the identical output:

```
$ make
g++ -Wall -std=c++11 main.c -o project1
$ ./project1 sample.txt
The number of words found in the file was 63
The number of unique words found in the file was 53

Please enter a word: of
The word of appears 2 times in the document

Please enter a word: is
The word is appears 2 times in the document

Please enter a word: andor
The word andor appears 1 times in the document

Please enter a word: or
The word or appears 1 times in the document

Please enter a word: ^C
$
```

We highly recommend that you create your own tests to make sure you have covered all possibilities. Those students that do not test their code have a higher chance of having errors during grading.

Submission

Remember that you must submit the project before the beginning of class on the day it is due. Any exceptions made will include a 10% penalty per day late. Please read the syllabus for more information.

For this project you must compress the čontents of your project0 directory into a zip file. This means if you unzip this file you would get only the contents of the directory and not the directory itself. The format of the zip file should be LAST-NAME_FIRSTNAME.zip. You can test this on the cs-intro server with the command (change to your name)

```
$ unzip Robinson_Jeffrey.zip
Archive:  Robinson_Jeffrey.zip
  inflating: main.cpp
  inflating: makefile
  inflating: sample.tex
$ ls
main.cpp  makefile  Robinson_Jeffrey.zip  sample.tex
$
```

Once you have created your zip file please submit it to blackbaord for gradinng.

Failure to follow the submission instructions will result in a reduction of your grade.