CS106 W21 - Assignment 05

Due: Friday, March 5, 11:59 PM

Assignment 5 is graded out of 27 marks.

In CS106 this semester we have been using a file"words.txt". It has about 375,000 words. You are to create a sketch to do the following. For starter code you may use any sketches from the course.

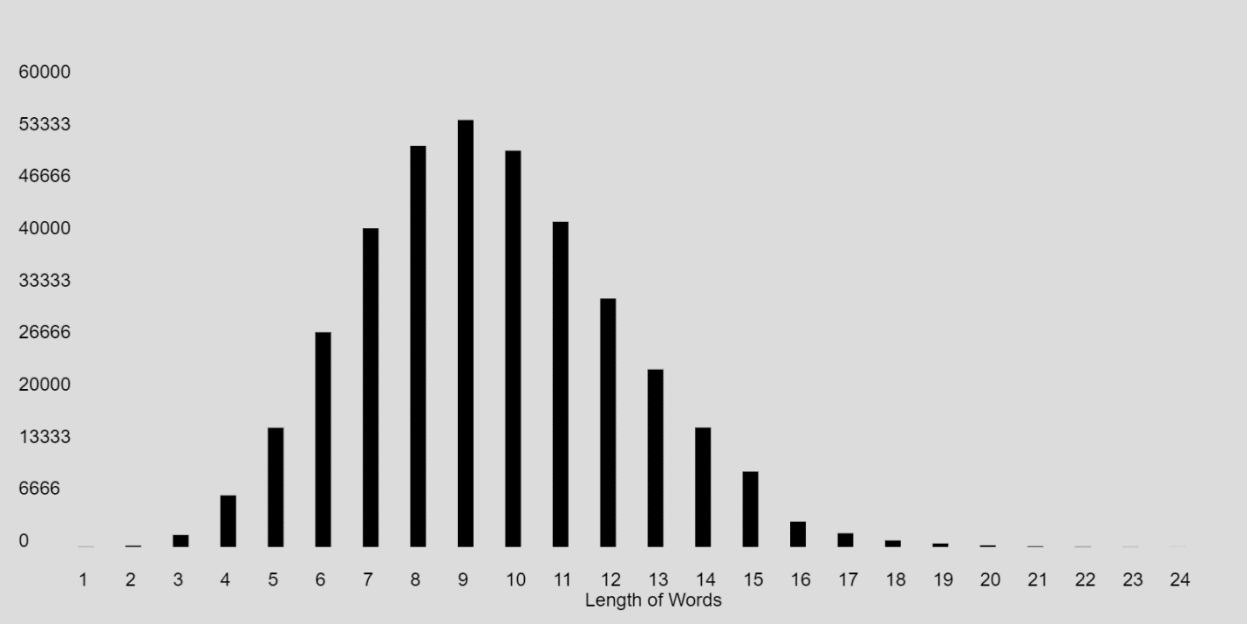
a) [1 mark] Set the background to gray and canvas size to 800x400.

b) [10 marks] In "words.txt" determine how many words there are of each length. If a word is longer than 24 letters then just count it as being of length 24. Store your results in an array. Your array should have values similar to the values on the following page. The first column shows the length of the word and the second column shows how many words of that length. For example, in "words.txt" there are 26 words of length 1, and 149 words of length 2.

c) [1 mark] Use print() to print out your array. It will display in the console and it will look similar to the image at the top of the next page. You do not need to include the total (but you can if you like).

```
1 26
2 149
3 1518
4 6579
5 15241
6 27467
7 40779
8 51317
9 54630
10 50708
11 41613
12 31778
13 22713
14 15265
15 9640
16 3223
17 1738
18 815
19 417
20 194
21 81
22 40
23 16
24 5
Total: 375952
```

d) [10 marks] Graph the data in your array. In creating the graph, you can assume that there are no values larger than 60,000 (that is, in the image above we can assume all values in column two are between 0 and 60,000). You graph should look similar to the image below. It shows, for example, that words of length 9 are the most common, and there are almost no words of length 24.



[5 marks] Coding Style and Efficiency

Follow the course coding style for whitespace and comments. Consult the **"Code Style Guide"** on LEARN. For example:

1) [0.5] Include your name on the first line of code and your student ID number on the second line of code.
2) [0.5] Leave the third line blank.
3) [0.5] Comment your code appropriately. Avoid superfluous comments.
4) [0.5] Correctly and consistently indent your code blocks.
5) [0.5] Use correct inline spacing for variable declaration and assignment.
6) [0.5] Use good line spacing to chunk sections of your code.
7) [0.5] There are no variables that are declared or assigned, but not used.
8) [0.5] There are no unnecessary variables that are duplicates of other variables.
9) [0.5] There is no unnecessarily repeating the same code in multiple places.
10) [0.5] Semicolons were used appropriately (i.e. at the end of most lines).

Restrictions

- You may not use any functions or statements not covered in lecture or labs.
  This includes, but is not limited to:
  - No translate(), rotate(), or scale() functions.

Submitting

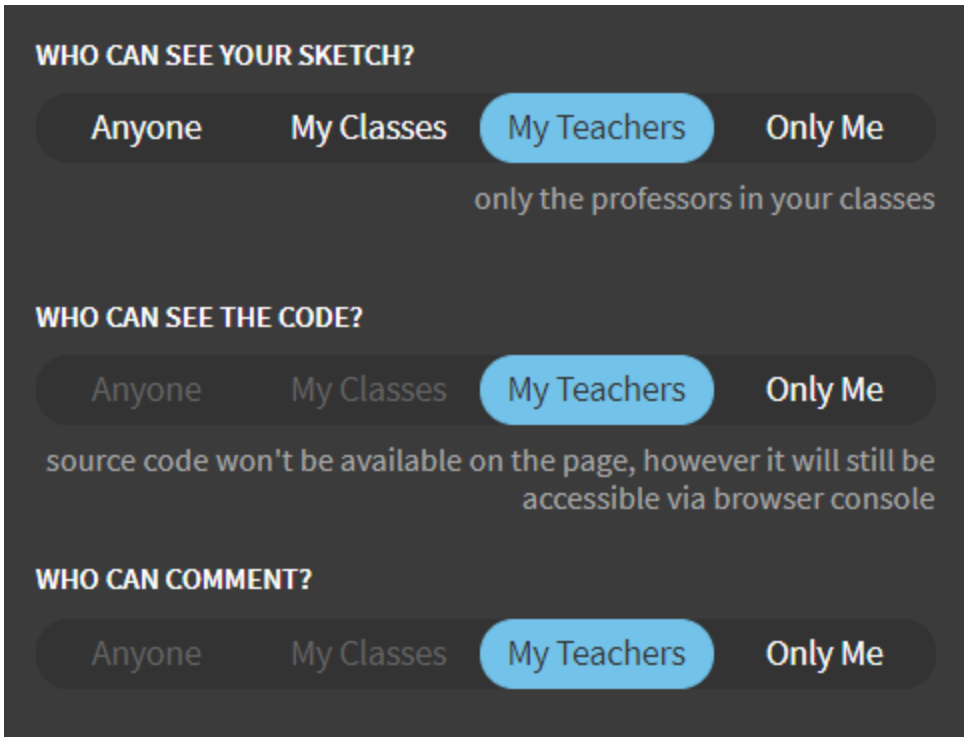
Use the template file in Word "CS106 Assignment Template" in LEARN to create your Assignment 05 submission.

Then convert your Word file to pdf. Please ensure that your URLs are hot links. The TAs need to be able to click on each link in your pdf and go directly to your sketch.

So for example, don't have a link like this: https://openprocessing.org/sketch/1050954

but rather have that link as a hot link as follows: https://openprocessing.org/sketch/1050954

Ensure that each URL you submit has its settings so that the access is as follows:



Submit that pdf file to the Assignment 05 dropbox on LEARN.

An example of how to submit a Lab is shown in the following video:
https://vault.cs.uwaterloo.ca/s/9Xx7AGsewaea773

It is your responsibility to submit to the correct dropbox with the correct file before the deadline. Otherwise you may receive a mark of 0.

Academic Integrity

All assignments in CS106 are done individually. Group work and sharing of code is not allowed.

Detecting Plagiarism:

- We monitor Reddit, File Trading Sites, past year CS106 assignments, etc.
- Measure Of Software Similarity (MOSS)
  - automatic system for determining the similarity of code

Discipline

- Discipline (Policy 71)
  - https://uwaterloo.ca/secretariat-general-counsel/policies-procedures-guidelines/policy-71