In this assignment, we use regular expressions to hunt for words, expressions and sentences in a a text document. Your solutions will all receive part of this text on stdin: Alice in Wonderland. For your testing convenience, use `wget` or `curl` to save it to a local file.

1 Most commonly used words (20%)

Write a bash script `words.sh` to extract the 10 most common words in the input, together with their frequency. You may use any of the tools we've discussed in class so far, but you may find `sort`, `sort -n`, `uniq -c` handy.

Below is the correct output for the full text. Note that we are ignoring capitalization.

```
 411 you
 462 said
 514 of
 545 i
 553 she
 595 it
 632 a
 729 to
 872 and
1644 the
```

2 Most commonly used double-vowel words (40%)

Write a bash script `doubles.sh` using `sed` and other tools to find the 10 most commonly used double-vowel words in the text, together with their frequency.

A double-vowel word is a word that contains the a vowel repeated twice, directly adjacent, such as `see` and `soon`. Use `sed -E` with a parenthesis expression to match one vowel, and its twin directly after.

Below is the correct output for the full text.

```
 26 good
 27 seemed
 28 look
 28 three
 30 door
 32 looking
 38 been
 45 looked
 66 see
 74 Queen
```

3 All the curious things (40%)

Alice likes to say things like “What a curious hat!”. Write a script `curious.sh` which extracts from `stdin` all of the curious things Alice observes. To keep things simple, let's say this means every word that comes right after "curious" without any punctuation between.

The trick will be to first get the text into a convenient shape - say with one "curious" per line, then to pull out the curious words. Below are all of those words for the full text. Watch out - some curious words come right after a line break.

```
appearance
as
child
creatures
croquet-ground
dream
feeling
plan
sensation
song
thing
to
today
```