Visit the following URL for some regular expression practice or "regex play"

```
http://cs.wheatoncollege.edu/regexplay
```

Each regular expression will be matched against every word in the "dictionary", a file called big_english.txt, a somewhat eclectic collection of approximately 235,000 words. (Note: rather than using the .cgi at the website above, students could write a script to use a loop to apply each regex to each word in the file.)

Note: If you ever listen to National Public Radio (NRP) on Sunday mornings (Weekend Edition), you've likely heard Will Shortz' seven-minute feature where he (the "Puzzle Master") challenges readers with word puzzles. Mason Brown, former Wheaton College Reference Librarian, and I have pored over two of Will's books to find puzzles that can be solved by one regexes (or sometimes two regexes in combination).
(0) From Shortz (1996), \#44 "On the Money." The currency name MARK is hidden in the word REMARKABLE, and RAND is hidden in GRANDMA. What common words does DINAR appear in?

Your regex: $\square$

## (Some of) The answer(s):

(1) From Shortz (1996), \#162 "Salty Language." There are common words in the English language that contain the consecutive letters N-A-C-L. Name any three of them.

Your regex: $\square$
(Some of) The answer(s):
(2) From Shortz (1996), \#152 "Two-Tone." Quickly now ... name a common English word that contains the consecutive letter combination T-A-N-T-A-N.

Your regex: $\square$
(Some of) The answer(s):
(3) From Shortz (2003), \#124 "D-Plus."

If you were asked to name a familiar word that contains two S's, followed by another letter, and then two more S's, you might say ASSESSOR. What common word contains two $D$ 's, followed by another letter, and then two more D's?

Your regex: $\square$
The answer:
(4) From Shortz (1996), \#38 Vowel Play." (slightly modified)

There's a well-known old puzzle to name words that contains the vowels A, E, I, $\mathbf{O}$, and $\mathbf{U}$ in order (and only one of each vowel). What are they?

Your regex: $\square$
The answer:
(5) From Shortz (1996), \#76 Telephone Word." (modified)

In an earlier time, you could use the number pad on your phone to enter "text" messages. Most number-keys on a cell phone referred to three or more letters, e.g., (2) ABC. New entering schemes eventually emerged, such as T-9, which stood for Text on 9 keys:

T-9's objective was to make it easier to type text messages. It allowed words to be entered by a single keypress for each letter, as opposed to the multi-tap approach used in the older generation of mobile phones in which several letters are associated with each key, and selecting one letter often required multiple keypresses. It combined the groups of letters on each phone key with a fast-access dictionary of words. It looks up in the dictionary all words corresponding to the sequence of keypresses. (wikipedia.com)

What words can the following phone number 666-7666 spell?

## Your regex:

$\square$
The answer:
(6) (Continuing from \#5 above) What words can your cell number spell? (Note: you may have to shorten your number, e.g., what words can your last four digits spell?)

Your regex: $\square$
The answer:
(7) Assuming a regular keyboard and that you are using your left hand fingers on only the appropriate "left-hand" keys, what is the longest word you can type with only your left hand?

Your regex: $\square$
The answer:
(8) Assuming you are using your right hand fingers on only the appropriate "righthand" keys, what is the longest word you can type with only your right hand?

Your regex: $\square$

## The answer:

(9) Are there any English words that end in the two letters "mt"?

Your regex: $\square$

## The answer:

(10) How many words end in "dous"?

Your regex: $\square$
The answer:
(11) What is the longest word that can be typed with only the letters on one row of a regular QWERTY keyboard?

Your regex:


The answer:
(12) What is the longest word that can be typed with only the letters on the entire home (that is, middle) row of the Dvorak keyboard? (Yeah, you'll have to look that up).

Your regex:


The answer:
(13) Are there any words in English where the letter ' $q$ ' is not immediately followed by a letter that is not a ' $u$ '?

Your regex:


The answer:
(14) Are there any words in which a sequence of three (3) letters is repeated elsewhere in the word?

Your regex: $\square$
The answer:
(15) Are there any words in which a sequence of two letters is directly repeated (with no gaps) at least two more times in the word?

Your regex:


The answer:
(16) Are there any words in which the first three letters of the beginning of a word are the same as the reverse of those letters at the end of the word?

Your regex: $\square$
The answer:
(17) Are there any palindromes of length six? Of length seven? What is the longest palindrome in this dictionary?

Your regex: $\square$
The answer:

## References

LeBlanc, M.D. and Dyer, B.D. (2007). Perl for Exploring DNA. Oxford University Press.

Shortz, Will (1996). The Puzzle Master Presents: 200 Mind-Bending Challenges. Random House.

Shortz, Will (2003). The Puzzlemaster Presents: Will Shortz's Best Puzzles from NPR.
Volume 2. Random House.

