

BT3051 Assignment 2
Roll number: BS16B019
Collaborators: Nil
Time: 2 hours

1.a. No. 18 is not matched as after it matches 1, it doesn't look in the or condition.
1.b. No. The regex will not check the next line (unless -z is evoked, In this case, it'd match "There are 118 trucks...is Zero" , so still not the target).

2.a. **^(?!F).*el+\$\b**

If the word starts with F, don't match. Then look for words that end with e followed by 1 or more occurrences of l.

2.b. **\b((?!cq)\w)*\b**

If the word contains cq in succession, don't match.

2.c. **\b((?!i).)*[il][nt]i\w*\b**

Match words that have an 'i' or 'l' followed by a 'n' or 't', followed by another 'i' , but doesn't have any 'i' preceding (ie, the [il] should be the first 'i' it encounters). Then match 0 or more characters.

2.d. **\b\d{4,5}\b**

Matches exactly 4 or 5 occurrences of any digit

3.a. Match a '<', then an 'a', then 0 or more occurrences of any alphabet, then an 'e' followed by 0 or more occurrences of any alphabet, then an 'i' followed by 0 or more occurrences of any alphabet and so on and close with '>'.

Eg -

Match - <aeiou>, <abstemious>, <aerious>, <affectious>, <anemious>

Mismatch - <facetious>, <caesious>, affectious, iouea

3.b. Matches any word which has at-least 4 characters (alphanumeric), no spaces allowed. The word must be enclosed in <>

Eg -

Match - <match>, <mismatch>, <hello>, <world>

Mismatch - <ma2ch>, mach, echo, <ca1>

3.c. Starts with 3 digits, followed by any number of digits (0 or more), followed by the exact first 3 digits, and another digit. This set of 4 digits can occur more than once. The string has to end with a '\$'

Eg -

Match - 12312311231\$, 1231231\$, 1233211231\$, 2102107\$

Mismatch - 1231231, 1234561\$, 123123\$, 21021093\$

3.d. Match one or more occurrences of any alphanumeric character, followed by a ' , followed by one or more occurrences of any alphanumeric character again.

Eg -

Match - can't , shouldn't , couldn't , hasn't

Mismatch - wood, could, should' , wanted

3.e Nice one. Starts with 3 digits. All 3 digits are in \1, the first two in \2, and only the first digit in \3

The regex matches 3 digits followed by another digit, then all the three digits, first two digits, first digit and \$

Eg -

Match - 1230123121\$, 7890789787\$, 4561456454\$, 1592159151\$

Mismatch - 1230123, 1230123\$, 1230321\$, 1230123121

4.a. **`^[+-]?[0-9]*\.[0-9]*$`**

4.b. **`^[0-9]{2}([_\.\/])?[0-9]{2}1[0-9]{4}$`**

4.c. To just get the filename at the end **`\w+\.[\w+]`** works. But to check for http(s)

`^http[s]?://(?:[A-Za-z-]+/)*([\w+\.[\w+])` this works, the filename is stored in \1.

4.d **`ATG([ATGC]{3})*(TAA|TAG|TGA)`**

4.e. **`AUG((?!UAG|UAA|UGA)[AUGC]{3})*(UAA|UAG|UGA)`**

4.f.