

Домашнее задание 1

Вы можете выполнять задание, воспользовавшись сайтом regex100.com или же взять какой-нибудь текстовый редактор, например, Notepad++ (на Windows) или Sublime (на Mac OS). Наиболее уверенные в себе, могут выполнить его на Python (понадобится модуль `re` и функция `re.findall()` из него). Выбор инструмента не влияет на оценку.

Для выполнения части 2 вам понадобятся файлы AppleRus.txt и AppleEng.txt из архива.

Часть 1

1. Найдёт ли регулярное выражение `sl[e]py?`, последовательность `sleepy`? Ответ объясните.
2. Дано регулярное выражение `н*его`. Какие последовательности символов оно найдёт? Опишите словами, приведите примеры.
3. Какие последовательности найдёт регулярное выражение `шо?ко?лад?`? Перечислите пять таких последовательностей.
4. Дано регулярное выражение `var[iou]+s`. Какие из перечисленных вариантов оно позволит найти?
a) various b) variis c) varius d) vars e) varaaaaas f) varooois
5. Что найдёт регулярное выражение `[^Nn][A-Za-z]+`? Какие последовательности оно не найдёт?
6. Какие последовательности найдёт регулярное выражение `\b[тТ]епло[a-я]+\b`?
7. Дано регулярное выражение `#[a-я]+_[0-9]?[a-я]+` (знаки `#` и `_` обозначают сами себя, т.е. хэштег и нижнее подчеркивание)
a) Какие последовательности символов оно найдёт? Опишите словами, приведите примеры.
b) Измените регулярное выражение так, чтобы оно могло находить те же последовательности, что и `#[a-я]+_[0-9]?[a-я]+`, а также последовательности вроде `#журналистика_`, т.е. заканчивающиеся на нижнее подчеркивание.

Часть 2

1. Напишите регулярное выражение, которое найдёт два слова с заглавной буквы подряд
2. Напишите регулярное выражение, которое найдёт даты в формате *оп – месяц – число – , – год*, например, *On March 5, 1953*. Не забудьте учесть регистры и границы. Используйте текст AppleEng.txt

3. Напишите регулярное выражение, которое будет вытаскивать из текста "AppleRus.txt" последовательности вроде “заместитель председателя”, “руководитель Алтайского”, “руководитель Осетинского” и т.п., то есть должность на -тель плюс первое слово после него.

4. а) Напишите регулярное выражение, которое найдет дробные числа с любым количеством знаков после точки. Регулярное выражение должно находить числа полностью, а не части других чисел.

б) Напишите регулярное выражение, которое найдет числа с одной десятой, например, 199.5 или 1.8. Это должны быть отдельные числа, а не части другого числа, например, 199.5 из 199.52.

с) Напишите регулярное выражение, которое найдет Все указания различных значений в миллионах и миллиардах, например, 199.5 million, 100 million, 80 billion, 1.89 billion (для текста из файла AppleEng.txt).