# Регулярки

## Задание для: Поляков Игорь Андреевич

**|** Версия: 2

Надо написать программу, которая извлекает из текстового файла данные и сохраняет их в табличной форме.

* Открываете файл по ссылке <https://github.com/rust-lang/regex/blob/1.4.5/src/testdata/basic.dat#L1-L221>, сохраняете выделенный желтым цветом текст в файл.
* Смотрите на входной файл и на таблицу снизу, таблица - это то, что должно в итоге получиться.
* Часть 1. Пишете регулярку, которая находит во входном файле интересующие куски, один кусок соответствует одной строке в выходной таблице, группами помечаете ячейки.
  + С помощью замены показываете, что получили правильные значения
  + Делать можно в студии или на сайте типа <https://regexr.com/> (еще есть <https://www.debuggex.com/> и <https://regex101.com/>), но в студии должно работать в любом случае.
  + **Вы должны уметь воспроизвести регулярку по памяти за пару минут и объяснить каждый символ в ней.**
* Часть 2. Пишете консольную программу, которая
  + читает текст из скачанного файла;
  + читает регулярку из файла с регуляркой (так удобнее дебажить);
  + с помощью Regex находит нужные куски;
    - из свойства Match.Groups достает нужные компоненты, каждой компоненте делаете s = s.Replace("\t", "\\t") (чтобы совпало с таблицей из задания);
    - добавляете слева "строка {номер куска}";
    - печатает кусок в консоль для отладки (должно быть видно структуру);
  + записывает куски в \*.csv файл.
    - Для каждой ячейки: static string EscapeCsvValue(string s) => "\"" + s.Replace("\"", "\"\"") + "\"";
    - Разделитель ячеек: ;, кодировка: Encoding.Default. Если открыть полученный файл в Excel, должна получиться выходная таблица.
    - Если в Excel не будет открываться или нету Excel, то сделайте разделитель табуляцией, потом Ctrl-C из выходного текстового файла и Ctrl-V в пустую таблицу любого аналога Excel, должно работать.
  + Читать файлы можно только функцией File.ReadAllText, записывать - File.WriteAllText.

Выходная таблица (показаны только первые и последние строки):

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Номер | Группа 1 | Группа 2 | Группа 3 | Группа 4 |
| строка 1 | BE | abracadabra$ | abracadabracadabra | (7,18) |
| строка 2 | BE | a...b | abababbb | (2,7) |
| ... | ... | ... | ... | ... |
| строка 216 | E | .\*(\\000).\* | \000 | (0,4)(0,4) |
| строка 217 | E | \\000 | \000 | (0,4) |