

regex

#Загрузите данные о землетрясениях

```
anss <- readLines("https://raw.githubusercontent.com/SergeyMirvoda/MD-DA-2017/master/data/earthquakes_2011.html", warn=FALSE)
```

#Выберите строки, которые содержат данные с помощью регулярных выражений и функции grep

2002/01/01 10:39:06.82,-55.2140,-129.0000,10.00,6.00,Mw,78,,,1.07,NEI,2002010140

#внутри <pre>""</pre>

#сплит через запятые

#[...] - character list

#[^...] - inverted character list

#\\d чтобы отсечь заголовок

```
data<-strsplit(anss[gre(pattern="^[^<]\\d" , anss)], ",")
typeof(data)
```

```
## [1] "list"
```

```
attributes(data)
```

```
## NULL
```

#первая строка

```
data[1]
```

```
## [[1]]
```

```
## [1] "2002/01/01 10:39:06.82" "-55.2140"
```

```
## [3] "-129.0000" "10.00"
```

```
## [5] "6.00" "Mw"
```

```
## [7] "78" ""
```

```
## [9] "" "1.07"
```

```
## [11] "NEI" "2002010140"
```

#конкретный элемент

```
data[[1]][[2]] #всё очень плохо
```

```
## [1] "-55.2140"
```

#Проверьте что все строки (all.equal) в результирующем векторе подходят под шаблон.

#?какой шаблон? старый? новый?

```
wtf <- all.equal(data,gre(pattern="^[^<]\\d",anss))
```