

RegexOne

These notes have been taken from <https://regexone.com>. a

Lesson 1: An Introduction, and the ABC's

Regular expressions are extremely useful in extracting information from text. The first thing to recognize is that **everything is a character**.

regex allows you to match characters anywhere in the string.

```
regex: abc abcdef abcde abcd abc
regex: 123 abc123xyz define "123" var g = 123;
```

Lesson 2: The Dot

The dot operator `.` is a wildcard and can match any *single* character. In order to match an actual period character, we need to escape the operator by using a slash, ie. `\.` The following shows a way to match the first three strings but now the last.

```
regex: .... match: cat. match: ?=+. match: 896. skip: abc1
```

Lesson 3: Matching Specific Characters

Sometimes `.` is too powerful for our use. For example, if we wanted to validate a phone number but we used `.`, it would allow *(abc) def-ghif* as being a valid phone number!

In order to match specific characters, we define them inside square brackets. For example, the pattern `[abc]` will only match a **single** 'a', 'b', or 'c' letter and nothing else.

Below are a few lines, match only the first three strings using regex, but not the last three.

```
regex: [cmf]an match: can match: man match: fan skip: dan skip:
ran skip: pan
```

Lesson 4: Excluding Specific Characters

In other cases, we might want to exclude specific characters instead of including them. To represent this, we use the hat operator `^`. For example, the pattern `[^abc]` will match any **single** character *except for* the letters 'a', 'b', or 'c'.

Match the lines that match with hog and dog, but not bog

```
regex: [^b]og match: hog match: dog skip: bog
```

Lesson 5: Character Ranges

Use a dash to indicate a character range of sequential characters to match. For example, the pattern `[0-6]` will match a single digit from 0 to 6, and the pattern `[^n-p]` will match a single character **except** for the letters from n through p.

Multiple character ranges can be used in the same set of brackets, along with individual characters. An example of this is the alphanumeric `\w` metacharacter which is equivalent to the character range `[A-Za-z0-9_]` and is often used to match characters in English text.

Lesson 6: Repititions of Characters to Match

Instead of