Regular Expression Tutorial

• • •

Esha G., Saahil T., Vienna P.

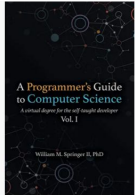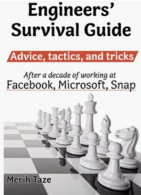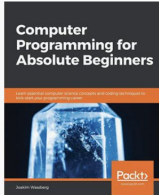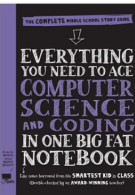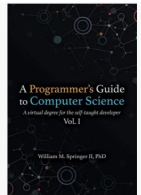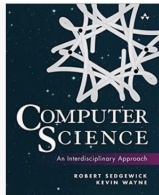# Operands

Price - `"$"[0-9,]+\.[0-9]+.*`

Shipping date - `.*(Mon | Tue | Wed | Thu | Fri | Sat | Sun)", "(Jan | Feb | Mar | Apr | Jun | Jul | Aug | Sep | Oct | Nov | Dec)" "[0-9]+`

Rating - `[0-9]\.[0-9]" out of 5 stars".*`

Free shipping - `"FREE Shipping".*`

Application: Amazon Shopping

In our demo, we will scan through a series of products from Amazon.

For instance, the image to the left is an example of how items are displayed after the customer specifies interest in "computer science"-related products.

This format is typical of many online retailers other than Amazon.

Amazon Shopping: Characters

- Price → $xxx.xx

- Delivery Date → (Mon-Sun), (Jan-Dec) (1-31)

- Rating → _ out of 5 stars

- FREE Shipping

- Product name

Sponsored ⓘ

A Programmer's Guide to Computer Science: A virtual degree for the self-taught developer
by Dr. William M Springer II, Nicholas R Allgood, et al.

★★★★⯪ ⌄ 212
Paperback
$26⁹⁹ $29.99
prime FREE Delivery **Sun, Feb 13**

Input: Example

- Price → $26.99

- Delivery Date → Sun, Feb 13

- Rating → 4 out of 5 stars

- FREE Delivery

- A Programmer's Guide to Computer Science: A virtual degree for the self-taught developer
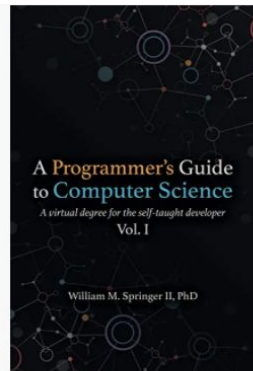
Sponsored ⓘ

A Programmer's Guide to Computer Science: A virtual degree for the self-taught developer
by Dr. William M Springer II, Nicholas R Allgood, et al.

★★★★☆ ⌄ 212

Paperback
$26⁹⁹ $29.99
prime FREE Delivery **Sun, Feb 13**

Challenges

- Product name size

- No product name uniformity

- Prices range varies in length (e.g. x.xx, xx.xx, xxx.xx)

- Shipping times

JFlex: Lexical Rules

```
%%
/**
 * lexical rules
 */
"$"[0-9,]+\.[0-9]+.*           {return "Price: " + yytext();}
.*(Mon | Tue | Wed | Thu | Fri | Sat | Sun)", "(Jan | Feb | Mar | Apr | Jun | Jul | Aug | Sep | Oct | Nov | Dec)" "[0-9]+ {return "Shipping:" + yytext();}
[0-9]\.[0-9]" out of 5 stars".* {return "Rating: "+yytext();}
"FREE Shipping".* {return "Free shipping";}
.+      {return "Product Name:" + yytext();}
{WhiteSpace}              {}
.                        { /* do nothing */ }
```