

Regular Expression Tutorial

...

Esha G., Saahil T., Vienna P.

Operands

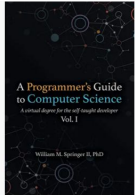
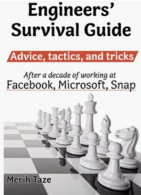
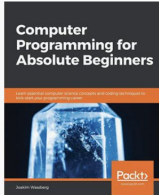

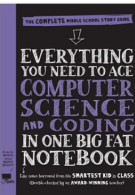
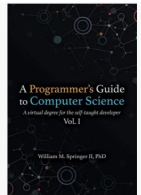
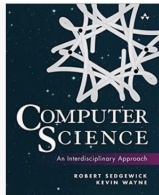

Price - `"$"[0-9,]+\. [0-9]+\. *`

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

			
<p>Sponsored</p> <p>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</p>	<p>Sponsored</p> <p>Engineers Survival Guide: Advice, tactics, and tricks After a decade of working at Facebook, Snapchat, and Microsoft by Merih Taze ★★★★☆ ~ 16 Paperback \$29⁹⁹ \$39.99 ✓prime FREE Delivery Sun, Feb 13</p>	<p>Sponsored</p> <p>Computer Programming for Absolute Beginners: Learn essential computer science concepts and coding techniques to kick-start your... by Joakim Wassberg ★★★★☆ ~ 51 Paperback \$34⁹⁹ ✓prime FREE Delivery Sun, Feb 13</p>	<p>Sponsored</p> <p>Mastering Transformers: Build state-of-the-art models from scratch with advanced natural language processing techniques by Savaş Yıldırım and Meysam Asgari-Chenaghlu ★★★★☆ ~ 9 Paperback \$39⁹⁹ ✓prime FREE Delivery Sun, Feb 13</p>
			
<p>Everything You Need to Ace Computer Science and Coding in One Big Fat Notebook: The Complete Middle School Study Guide (Big Fat...) Part of: Big Fat Notebooks (11 Books) ★★★★☆ ~ 3,685 Paperback \$12⁸⁷ \$16.95 Get 3 for the price of 2 ✓prime FREE One-Day Get it Tomorrow, Feb 12</p>	<p>A Programmer's Guide to Computer Science: A virtual degree for the self-taught developer Book 1 of 2: A Programmer's Guide to Computer Science ★★★★☆ ~ 212 Paperback \$26⁹⁹ \$29.99 ✓prime FREE Delivery Sun, Feb 13 More Buying Choices \$21.64 (15 used & new offers)</p>	<p>Computer Science: An Interdisciplinary Approach by Robert Sedgewick and Kevin Wayne ★★★★☆ ~ 58 Hardcover \$55⁷⁷ \$79.99 ✓prime FREE Delivery Sun, Feb 13 Only 17 left in stock (more on the way). More Buying Choices \$48.63 (20 used & new offers) Other format: eBook</p>	<p>Infinity Chamber 2017 ★★★★☆ ~ 1,112 Prime Video Included with your Prime membership Watch now</p>

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.

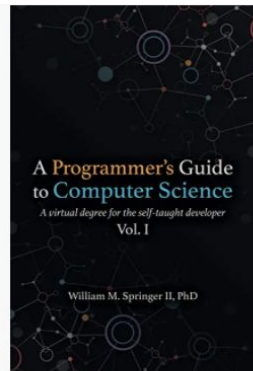
Amazon Shopping: Characters

- Price → \$xxx.xx
- Delivery Date → (Mon-Sun), (Jan-Dec) (1-31)
- Rating → _ out of 5 stars
- FREE Shipping
- Product name



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 */ }
```