

✓ 100 XP

# Exercise: Complete a challenge to extract, replace, and remove data from an input string

10 minutes

Code challenges reinforce what you've learned and help you gain some confidence before continuing on.

In this challenge, you work with a string that contains a fragment of HTML. You extract data from the HTML fragment, replace some of its content, and remove other parts of its content to achieve the desired output.

If you're unfamiliar with HTML code, review the ["Quick HTML primer"](#) section at the end of this unit.

## Extract, replace, and remove data from an input string

1. Select and delete all code lines in the Visual Studio Code Editor.
2. In Visual Studio Code, add the following "starter" code to get the data for the challenge:

```
C#  
  
const string input = "<div><h2>Widgets &trade;</h2><span>5000</span></div>";  
  
string quantity = "";  
string output = "";  
  
// Your work here  
  
Console.WriteLine(quantity);  
Console.WriteLine(output);
```

If you run the code the output displays blank lines, the starting values for `quantity` and `output` are empty `string` values.

- Take a minute to review the first line of the code containing a `string` of HTML.

C#

```
const string input = "<div><h2>Widgets &trade;</h2><span>5000</span></div>";
```

Notice the tags: `<div>`, `<h2>`, `<span>` and symbol code `&trade;` contained in the `input` variable.

- Examine the **desired output** for the final program output:

Output

Quantity: 5000

Output: `<h2>Widgets &reg;</h2><span>5000</span>`

- Begin adding your solution code to the starter code under the comment `// Your work here`.
- Set the `quantity` variable to the value by extracting the text from between the `<span>` and `</span>` tags.
- Set the `output` variable to the value of `input`, then remove the `<div>` and `</div>` tags.
- Replace the HTML character `™` (`&trade;`) with `®` (`&reg;`) in the `output` variable.
- Run your solution and verify the output matches the expected output.

Output

Quantity: 5000

Output: `<h2>Widgets &reg;</h2><span>5000</span>`

Whether you get stuck and need to peek at the solution or you finish successfully, continue on to view a solution to this challenge.

## Quick HTML primer

In case you're not familiar with HTML, it's the markup language that is used to create all web pages. Skip this section if you have a good understanding of HTML. It's not designed to be a comprehensive tutorial, but provides enough information to complete this challenge.

In HTML, you define the structure of a document using tags. A tag is composed of:

- an opening angle bracket `<`
- a closing angle bracket `>`
- a word describing the type of tag, so for example: `<div>`, `<span>`, `<h2>` etc.

Each tag has a corresponding closing tag that is indicated by a forward slash character `/`. So, if you see `<div>` there should be a corresponding `</div>` tag.

The content between the opening and closing tag is the content of that tag. The content can include text and other tags.

A set of tags can be embedded inside another set of tags, giving an HTML document its hierarchical structure.

---

## Next unit: Review the solution to extract, replace, and remove data from an input string

[Continue >](#)