

HTML DESCRIBES THE STRUCTURE OF PAGES

In the browser window you can see a web page that features exactly the same content as the Word document you met on the page 18. To describe the structure of a web page, we add code to the words we want to appear on the page.

You can see the HTML code for this page below. Don't worry about what the code means yet. We start to look at it in more detail on the next page. Note that the HTML code is in blue, and the text you see on screen is in black.

```
<html>
 <body>
   <h1>This is the Main Heading</h1>
   This text might be an introduction to the rest of
      the page. And if the page is a long one it might
      be split up into several sub-headings.
   <h2>This is a Sub-Heading</h2>
   Many long articles have sub-headings so to help
      you follow the structure of what is being written.
      There may even be sub-sub-headings (or lower-level
      headings).
   <h2>Another Sub-Heading</h2>
   Here you can see another sub-heading.
 </body>
</html>
```

The HTML code (in blue) is made up of characters that live inside angled brackets — these are called HTML elements. Elements are usually made up of two tags: an opening tag and a closing tag. (The closing tag has an extra forward slash in it.) Each HTML element tells the browser something about the information that sits between its opening and closing tags.

HTML USES ELEMENTS TO DESCRIBE THE STRUCTURE OF PAGES

Let's look closer at the code from the last page. There are several different elements. Each element has an opening tag and a closing tag.

CODE

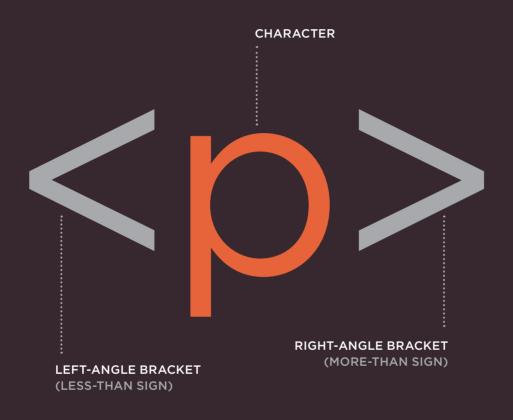


Tags act like containers. They tell you something about the information that lies between their opening and closing tags.

DESCRIPTION

| The opening https://www.ntml">https://www.ntml tag is HTML code |
|---|
| The <body> tag indicates that anything between it and the closing ······</body> |
| tag should be shown inside the main browser window. |
| Words between <h1> and </h1> are a main heading. |
| A paragraph of text appears between these and tags. |
| |
| Words between <h2> and </h2> form a sub-heading. |
| Here is another paragraph between opening and closing tags. |
| ······································ |
| Another sub-heading inside <h2> and </h2> tags. |
| Another paragraph inside and tags. |
| The closing tag indicates the end of what should appear in the main browser window. |
| The closing tag indicates that it is the end of the HTML code. |

A CLOSER LOOK AT TAGS

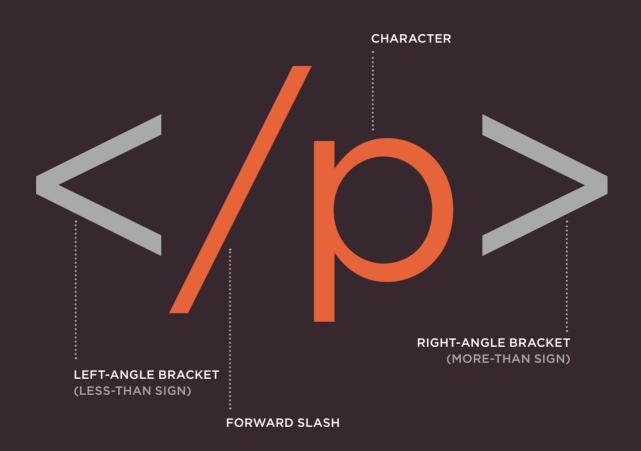


OPENING TAG -

The characters in the brackets indicate the tag's purpose.

For example, in the tags above the p stands for paragraph.

The closing tag has a forward slash after the the < symbol.



CLOSING TAG

The terms "tag" and "element" are often used interchangeably.

Strictly speaking, however, an element comprises the opening

tag and the closing tag and any content that lies between them.

ATTRIBUTES TELL US MORE ABOUT ELEMENTS

Attributes provide additional information about the contents of an element. They appear on the opening tag of the element and are made up of two parts: a name and a value, separated by an equals sign.



The attribute name indicates what kind of extra information you are supplying about the element's content. It should be written in lowercase.

The **value** is the information or setting for the attribute. It should be placed in double quotes. Different attributes can have different values.

Here an attribute called lang is used to indicate the language used in this element. The value of this attribute on this page specifies it is in US English.

HTML5 allows you to use uppercase attribute names and omit the quotemarks, but this is not recommended.



The majority of attributes can only be used on certain elements, although a few attributes (such as lang) can appear on any element. Most attribute values are either pre-defined or follow a stipulated format. We will look at the permitted values as we introduce each new attribute.

The value of the lang attribute is an abbreviated way of specifying which language is used inside the element that all browsers understand.

BODY, HEAD & TITLE

<body>

You met the <body> element in the first example we created. Everything inside this element is shown inside the main browser window.

<head>

Before the <body> element you will often see a <head> element. This contains information about the page (rather than information that is shown within the main part of the browser window that is highlighted in blue on the opposite page). You will usually find a <title> element inside the <head> element.

<title>

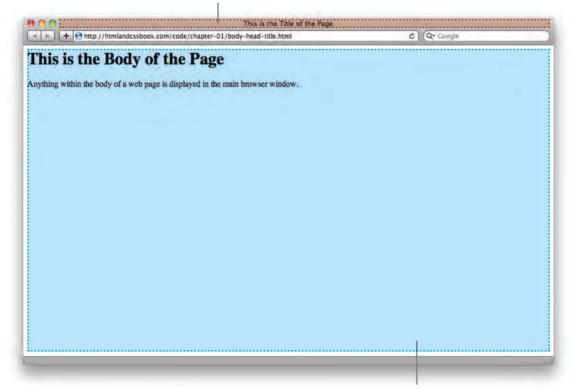
The contents of the <title> element are either shown in the top of the browser, above where you usually type in the URL of the page you want to visit, or on the tab for that page (if your browser uses tabs to allow you to view multiple pages at the same time).

RESULT

This is the Body of the Page

Anything within the body of a web page is displayed in the main browser window,

Anything written between the <title> tags will appear in the title bar (or tabs) at the top of the browser window, highlighted in orange here.



Anything written between the <body> tags will appear in the main browser window, highlighted in blue here.

You may know that HTML stands for HyperText Markup Language. The HyperText part refers to the fact that HTML allows you to create links that allow visitors to move from one

page to another quickly and easily. A markup language allows you to annotate text, and these annotations provide additional meaning to the contents of a document. If you think of a web

page, we add code around the original text we want to display and the browser then uses the code to display the page correctly. So the tags we add are the markup.

CREATING A WEB PAGE ON A PC

To create your first web page on a PC, start up Notepad. You can find this by going to:

Start
All Programs (or Programs)
Accessories
Notepad

You might also like to download a free editor called Notepad++ from notepad-plus-plus.org.



Type the code shown on the right.

9





Go to the File menu and select **Save as...** You will need to save the file somewhere you can remember. If you like, you could create a folder for any examples that you try out from this book.

Save this file as first-test. html. Make sure that the *Save* as *type* drop down has *All Files* selected.



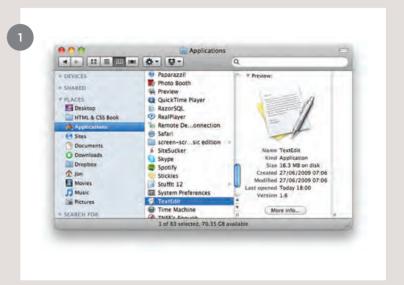
Start your web browser. Go to the *File* menu and select *Open*. Browse to the file that you just created, select it and click on the *Open* button. The result should look something like the screen shot to the left.

If it doesn't look like this, find the file you just created on your computer and make sure that it has the file extension .html (if it is .txt then you need to go back to Notepad and save the file again, but this time put quote marks around the name "first-test.html").

CREATING A WEB PAGE ON A MAC

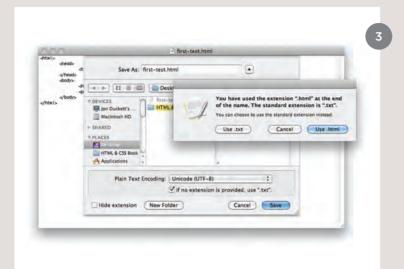
To create your first web page on a Mac, start up TextEdit. This should be in your **Applications** folder.

You might also like to download a free text editor for creating web pages called TextWrangler which is available from barebones.com.



Type the code shown on the right.





Now go to the *File* menu and select *Save as...* You will need to save the file somewhere you can remember.

If you like, you could create a folder for any examples that you try out from this book. Save this file as first-test.html. You will probably see a window like the screen shot to the left.

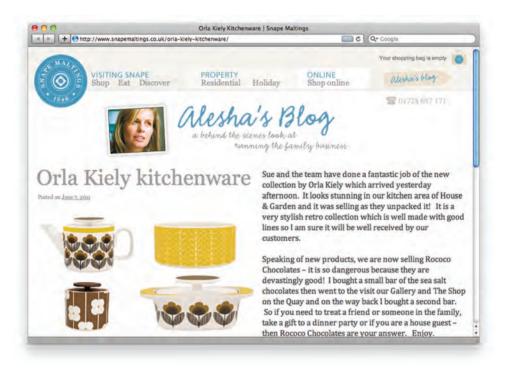
You want to select the *Use.html* button.



Next, start your web browser, go to the *File* menu, and select *Open*. You should browse to the file that you just created, select it and click on the *Open* button. The result should look like the screen shot to the left.

If it doesn't look like this, you might need to change one of the settings in TextEdit. Go to the TextEdit menu and select **Preferences**. Then on the preferences for **Open and Save**, tick the box that says **Ignore rich text commands in HTML files**. Now try to save the file again.

CODE IN A CONTENT MANAGEMENT SYSTEM



If you are working with a content management system, blogging platform, or e-commerce application, you will probably log into a special administration section of the website to control it. The tools provided in the administration sections of these sites usually allow you to edit parts of the page rather than the entire page, which means you will rarely see the <ntml>, <nead>, or <body> elements.

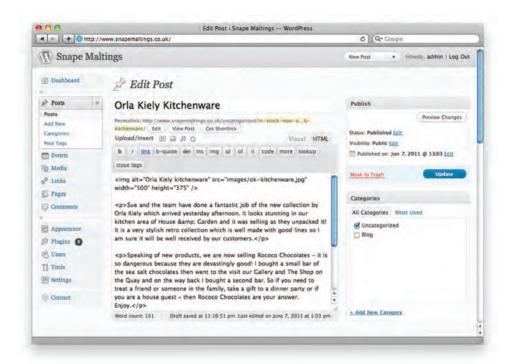
Looking at the content management system on the opposite page, you have a box that allows you to enter a title for the page, another box for the main article, a way to enter a publication date, and something to indicate which section of the site this page belongs in.

For an e-commerce store, you might have boxes that allow you to enter a title for the product, a description of the product, its price, and the quantity available.

That is because they use a single 'template' to control all of the pages for a section of the site. (For example, an e-commerce

system might use the same template to show all of their products.) The information you supply is placed into the templates.

The advantage of this approach is that people who do not know how to write web pages can add information to a website and it is also possible to change the presentation of something in the template, and it will automatically update every page that uses that template. If you imagine an e-commerce store with 1,000 items for sale, just



altering one template is a lot easier than changing the page for each individual product. In systems like this, when you have a large block of text that you can edit, such as a news article, blog entry or the description of a product in an e-commerce store, you will often see a text editor displayed.

Text editors usually have controls a little like those on your word processor, giving you different options to style text, add links or insert images. Behind the scenes these editors are adding HTML code to your text, just like the code you have seen earlier in this chapter. Many of these editors will have an option that allows you to see (and edit) the code that they produce.

Once you know how to read and edit this code, vou can take more control over these sections of your website.

In the example above, you can see that the text editor has a tab for Visual / HTML views of what the user enters. Other systems

might have a button (which often shows angle brackets) to indicate how to access the code.

Some content management systems offer tools that also allow you to edit the template files. If you do try to edit template files you need to check the documentation for your CMS as they all differ from each other. You need to be careful when editing template files because if you delete the wrong piece of code or add something in the wrong place the site may stop working entirely.

LOOKING AT HOW OTHER SITES ARE BUILT

When the web was first taking off, one of the most common ways to learn about HTML and discover new tips and techniques was to look at the source code that made up web pages.

These days there are many more books and online tutorials that teach HTML, but you can still look at the code that a web server sends to you. To try this out for yourself, simply go to the sample code for this chapter, at www.htmlandcssbook.com/code/ and click on the link called "View Source."

Once you have opened this page, you can look for the *View* menu in your browser, and select the option that says *Source* or *View source*. (The title changes depending on what browser you are using.)

You should see a new window appear, and it will contain the source code that was used to create this page.

You can see this result in the photograph on the right. The page you see is the window at the top; the code is below.

At first this code might look complicated but don't be discouraged. By the time you have finished the next chapter of this book, you will be able to understand it.

All of the examples for this book are on the website, and you can use this simple technique on any of the example pages to see how they work.

You can also download all of the code for this book from the same website by clicking on the "Download" link.



SUMMARY STRUCTURE

- ▶ HTML pages are text documents.
- HTML uses tags (characters that sit inside angled brackets) to give the information they surround special meaning.
- ▶ Tags are often referred to as elements.
- Tags usually come in pairs. The opening tag denotes the start of a piece of content; the closing tag denotes the end.
- Opening tags can carry attributes, which tell us more about the content of that element.
- Attributes require a name and a value.
- To learn HTML you need to know what tags are available for you to use, what they do, and where they can go.

TEXT

- ► Headings and paragraphs
- ▶ Bold, italic, emphasis
- ▶ Structural and semantic markup

When creating a web page, you add tags (known as markup) to the contents of the page. These tags provide extra meaning and allow browsers to show users the appropriate structure for the page.

In this chapter we focus on how to add markup to the text that appears on your pages. You will learn about:

- Structural markup: the elements that you can use to describe both headings and paragraphs
- Semantic markup: which provides extra information; such as where emphasis is placed in a sentence, that something you have written is a quotation (and who said it), the meaning of acronyms, and so on



HEADINGS

<h1><h2><h3><h4><

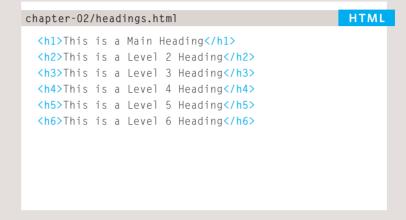
HTML has six "levels" of headings:

<h1> is used for main headings

<h2> is used for subheadings

If there are further sections under the subheadings then the <h3> element is used, and so on...

Browsers display the contents of headings at different sizes. The contents of an <h1> element is the largest, and the contents of an <h6> element is the smallest. The exact size at which each browser shows the headings can vary slightly. Users can also adjust the size of text in their browser. You will see how to control the size of text, its color, and the fonts used when we come to look at CSS.



This is a Main Heading

RESULT

This is a Level 2 Heading

This is a Level 3 Heading

This is a Level 4 Heading

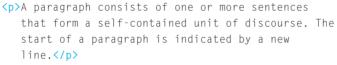
This is a Level 5 Heading

This is a Level 6 Heading

PARAGRAPHS

HTML

chapter-02/paragraphs.html



Text is easier to understand when it is split up into units of text. For example, a book may have chapters. Chapters can have subheadings. Under each heading there will be one or more paragraphs.



To create a paragraph, surround the words that make up the paragraph with an opening tag and closing tag.

By default, a browser will show each paragraph on a new line with some space between it and any subsequent paragraphs.

RESULT

A paragraph consists of one or more sentences that form a self-contained unit of discourse. The start of a paragraph is indicated by a new line.

Text is easier to understand when it is split up into units of text. For example, a book may have chapters. Chapters can have subheadings. Under each heading there will be one or more paragraphs.

BOLD & ITALIC



By enclosing words in the tags and we can make characters appear bold.

The element also represents a section of text that would be presented in a visually different way (for example key words in a paragraph) although the use of the element does not imply any additional meaning.

chapter-02/bold.html

HTML

This is how we make a word appear bold.

This is how we make a word appear bold.

RESULT

Inside a product description you might see some **key features** in bold.

<i>>

By enclosing words in the tags <i> and </i> we can make characters appear italic.

The <i> element also represents a section of text that would be said in a different way from surrounding content — such as technical terms, names of ships, foreign words, thoughts, or other terms that would usually be italicized.

chapter-02/italic.html

HTML

This is how we make a word appear italic.

RESULT

It's a potato Solanum teberosum.

Captain Cook sailed to Australia on the Endeavour.

SUPERSCRIPT & **SUBSCRIPT**

HTML

chapter-02/superscript-and-subscript.html

On the 4th of September you will learn about E=MC².

The amount of CO(sub>2</sub> in the atmosphere grew by 2ppm in $2009 \le \text{sub} \le 1 \le \text{sub} \le . \le \text{p}$

RESULT

On the 4th of September you will learn about E=MC2.

The amount of CO₂ in the atmosphere grew by 2ppm in 2009₁.

<sup>

The <sup> element is used to contain characters that should be superscript such as the suffixes of dates or mathematical concepts like raising a number to a power such as 2^2 .

<sub>

The <sub> element is used to contain characters that should be subscript. It is commonly used with foot notes or chemical formulas such as H₂0.

WHITE SPACE

In order to make code easier to read, web page authors often add extra spaces or start some elements on new lines.

When the browser comes across two or more spaces next to each other, it only displays one space. Similarly if it comes across a line break, it treats that as a single space too. This is known as white space collapsing.

You will often see that web page authors take advantage of white space collapsing to indent their code in order to make it easier to follow.

The moon is drifting away from Earth.

RESULT

The moon is drifting away from Earth.

The moon is drifting away from Earth.

LINE BREAKS & HORIZONTAL RUI FS

HTML

chapter-02/line-breaks.html

The Earth
gets one hundred tons heavier every day
due to falling space dust.

The Earth gets one hundred tons heavier every day due to falling space dust.

HTML

chapter-02/horizontal-rules.html

Venus is the only planet that rotates clockwise. <hr /> Jupiter is bigger than all the other planets combined.

RESULT

Venus is the only planet that rotates clockwise.

Jupiter is bigger than all the other planets combined.

As you have already seen, the browser will automatically show each new paragraph or heading on a new line. But if you wanted to add a line break inside the middle of a paragraph you can use the line break tag
.

<hr />

To create a break between themes — such as a change of topic in a book or a new scene in a play — you can add a horizontal rule between sections using the <hr /> tag.

There are a few elements that do not have any words between an opening and closing tag. They are known as **empty elements** and they are written differently.

An empty element usually has only one tag. Before the closing angled bracket of an empty element there will often be a space and a forward slash character. Some web page authors miss this out but it is a good habit to get into.

VISUAL EDITORS & THEIR CODE VIEWS

Content management systems and HTML editors such as Dreamweaver usually have two views of the page you are creating: a visual editor and a code view.

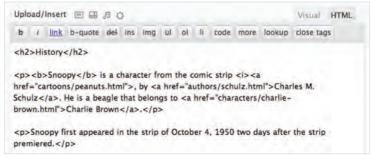
Visual editors often resemble word processors. Although each editor will differ slightly, there are some features that are common to most editors that allow you to control the presentation of text.

- Headings are created by highlighting text then using a drop-down box to select a heading.
- Bold and italic text are created by highlighting some text and pressing a b or i button.
- New paragraphs are created using the return or the enter key.
- Line breaks are created by pressing the shift key and the return key at the same time.
- Horizontal rules are created using a button with a straight line on it.

If you copy and paste text from a program that allows you to format text (such as Word) into a visual editor, it may add extra markup. To prevent this copy the text into a plain text editor first (such as Notepad on a PC or TextEdit on a Mac) and then copy it from that program and paste it into the visual editor.

Code views show you the code created by the visual editor so you can manually edit it, or so you can just enter new code yourself. It is often activated using a button with an icon that says HTML or has angled brackets. White space may be added to the code by the editor to make the code easier to read.





SEMANTIC MARKUP

There are some text elements that are not intended to affect the structure of your web pages, but they do add extra information to the pages — they are known as semantic markup.

In the rest of the chapter you will meet some more elements that will help you when you are adding text to web pages. For example, you are going to meet the element that allows you to indicate where emphasis should be placed on selected words and the <blockquote> element which indicates that a block of text is a quotation.

Browsers often display the contents of these elements in a different way. For example, the content of the element is shown in italics, and a <blockquote> is usually indented. But you should not use them to change the way that your text looks; their purpose is to describe the content of your web pages more accurately.

The reason for using these elements is that other programs, such as screen readers or search engines, can use this extra information. For example, the voice of a screen reader may add emphasis to the words inside the element, or a search engine might register that your page features a quote if you use the <blockquote> element.

STRONG & EMPHASIS

The use of the element indicates that its content has strong importance. For example, the words contained in this element might be said with strong emphasis.

By default, browsers will show the contents of a element in bold.

chapter-02/strong.html

HTML

Beware: Pickpockets operate in this area. This toy has many small pieces and is not suitable for children under five years old.

Beware: Pickpockets operate in this area.

RESULT

This toy has many small pieces and is not suitable for children under five years old.

The element indicates emphasis that subtly changes the meaning of a sentence.

By default browsers will show the contents of an element in italic

chapter-02/emphasis.html

HTML

I think Ivy was the first. I think Ivy was the first. I think Ivy was the first.

I think Ivy was the first.

RESULT

I think Ivy was the first.

I think Ivy was the first.

QUOTATIONS

HTML

chapter-02/quotations.html

<blockquote cite="http://en.wikipedia.org/wiki/</pre> Winnie-the-Pooh">

Did you ever stop to think, and forget to start again?

</blockguote>

As A.A. Milne said, <q>Some people talk to animals. Not many listen though. That's the problem.</q>

RESULT

Did you ever stop to think, and forget to start again?

As A.A. Milne said, "Some people talk to animals. Not many listen though. That's the problem."

There are two elements commonly used for marking up quotations:

<blookly </pre>

The <blockguote> element is used for longer quotes that take up an entire paragraph. Note how the element is still used inside the <blockquote> element.

Browsers tend to indent the contents of the <blockquote> element, however you should not use this element just to indent a piece of text — rather you should achieve this effect using CSS.



The <g> element is used for shorter quotes that sit within a paragraph. Browsers are supposed to put quotes around the <q> element, however Internet Explorer does not therefore many people avoid using the <q> element.

Both elements may use the cite attribute to indicate where the quote is from. Its value should be a URL that will have more information about the source of the quotation.

ABBREVIATIONS & ACRONYMS

<abbr>>

If you use an abbreviation or an acronym, then the <abbr> element can be used. A title attribute on the opening tag is used to specify the full term.

In HTML 4 there was a separate <acronym> element for acronyms. To spell out the full form of the acronym, the title attribute was used (as with the <abbr> element above). HTML5 just uses the <abbr> element for both abbreviations and acronyms.

chapter-02/abbreviations.html

HTML

<abbr title="Professor">Prof</abbr> Stephen
 Hawking is a theoretical physicist and
 cosmologist.

<acronym title="National Aeronautics and Space Administration">NASA</acronym> do some crazy space stuff.

Prof Stephen Hawking is a theoretical physicist and cosmologist.

RESULT

NASA do some crazy space stuff.

National Aeronautics and Space Administration

CITATIONS & **DEFINITIONS**

HTML

chapter-02/citations.html

<cite>A Brief History of Time</cite> by Stephen Hawking has sold over ten million copies worldwide.

A Brief History of Time by Stephen Hawking has sold over ten million copies worldwide.

HTML

chapter-02/definitions.html

A <dfn>black hole</dfn> is a region of space from which nothing, not even light, can escape.

RESULT

A black hole is a region of space from which nothing, not even light, can escape.

<cite>

When you are referencing a piece of work such as a book, film or research paper, the <cite> element can be used to indicate where the citation is from

In HTML5, <cite> should not really be used for a person's name — but it was allowed in HTML 4, so most people are likely to continue to use it.

Browsers will render the content of a <cit.e> element in italics.

<dfn>

The first time you explain some new terminology (perhaps an academic concept or some jargon) in a document, it is known as the defining instance of it.

The <dfn> element is used to indicate the defining instance of a new term.

Some browsers show the content of the <dfn> element in italics. Safari and Chrome do not change its appearance.

AUTHOR DETAILS

<address>

The <address> element has quite a specific use: to contain contact details for the author of the page.

It can contain a physical address, but it does not have to. For example, it may also contain a phone number or email address.

Browsers often display the content of the <address> element in italics.

You may also be interested in something called the hCard microformat for adding physical address information to your markup.

ONLINE EXTRA:

You can find out more about hCards on the website accompanying this book.

homer@example.org

RESULT

742 Evergreen Terrace, Springfield.

CHANGES TO CONTENT

HTML

chapter-02/insert-and-delete.html

It was the worst <ins>best</ins> idea she had ever had.

<ins>

The <ins> element can be used to show content that has been inserted into a document, while the element can show text that has been deleted from it.

The content of a <ins> element is usually underlined, while the content of a <de1> element usually has a line through it.

It was the worst best idea she had ever had.

HTML

chapter-02/strikethrough.html

Laptop computer: $\langle p \rangle \langle s \rangle Was $995 \langle /s \rangle \langle /p \rangle$ $\langle p \rangle$ Now only $375 \langle p \rangle$



The <s> element indicates something that is no longer accurate or relevant (but that should not be deleted).

Visually the content of an <s> element will usually be displayed with a line through the center.

Older versions of HTML had a <u> element for content that was underlined, but this is being phased out.

RESULT

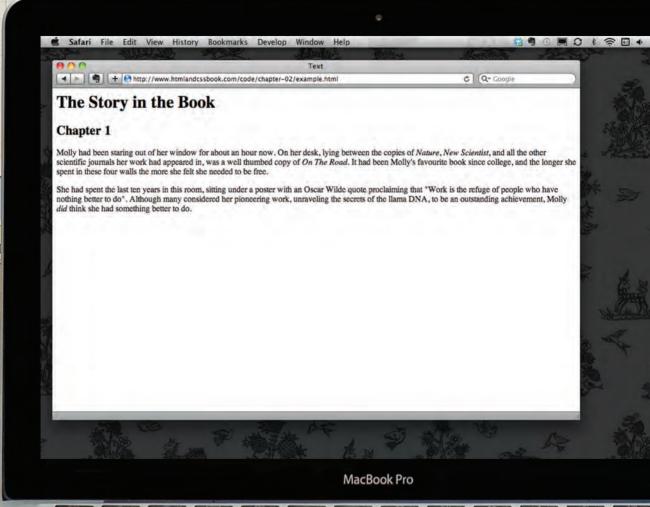
Laptop computer:

Was \$995

Now only \$375









EXAMPLE

This is a very simple HTML page that demonstrates text markup.

Structural markup includes elements such as <h1>, <h2>, and .
Semantic information is carried in elements such as <cite> and .

```
<ht.ml>
 <head>
    <title>Text</title>
  </head>
  <body>
    <h1>The Story in the Book</h1>
    <h2>Chapter 1</h2>
    Molly had been staring out of her window for about
       an hour now. On her desk, lying between the copies
       of <i>Nature</i>. <i>New Scientist</i>. and all
       the other scientific journals her work had
       appeared in, was a well thumbed copy of <cite>On
       The Road</cite>. It had been Molly's favorite book
       since college, and the longer she spent in these
       four walls the more she felt she needed to be
       free.
    She had spent the last ten years in this room,
       sitting under a poster with an Oscar Wilde guote
       proclaiming that <q>Work is the refuge of
       people who have nothing better to do\langle /q \rangle. Although
       many considered her pioneering work, unraveling
       the secrets of the llama <abbr
       title="Deoxyribonucleic acid">DNA</abbr>, to be an
       outstanding achievement, Molly <em>did</em> think
       she had something better to do.
  </body>
</html>
```

SUMMARY TEXT

- > HTML elements are used to describe the structure of the page (e.g. headings, subheadings, paragraphs).
- They also provide semantic information (e.g. where emphasis should be placed, the definition of any acronyms used, when given text is a quotation).