note

What is XML?

XML is a standard for storing information in a hierarchy. Here is some example XML code:

```
<food>
  <fruit type='apple' />
  <fruit type='banana' />
  <meat type='beef' />
</food>
```

The words contained inside < and > symbols are known as 'tags'.

In this example there is a collection of different food items grouped together inside the <food> tag.

You can see that XML tags can contain other tags and that tags can have extra properties that contain more information, as the 'type' property does in the example.

XML can be used to store all kinds of information. A basic HTML file is just an XML document that tells a web browser what to display on a web page and how to display it.

Lesson 1: Create a very simple web page

HTML is the 'language' used to create web pages and is a form of XML (see sidebar).

Any computer with a web browser and text editor can be used to create and edit HTML documents, but this course will show you how to create HTML documents using only software that comes pre-installed on every Windows computer.

If you are using a different operating system you should still be able to complete this course but you will need to use a text editor on your computer.

You can also use any other text editor, and might prefer to use one that is designed for HTML such as Visual Studio or Notepad++.
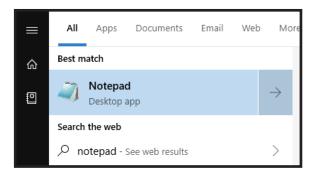
1.   Download the sample files.

The sample files for this course can be downloaded from: **http://learnhtmltoday.com**

You will need the sample files to complete some of the lessons. The sample files are always free to downloads.

2.   Open Notepad.

1. Click the Start Menu or press the Start key on your keyboard.

2. Type **Notepad** on your keyboard.

    Windows searches your computer and displays an icon for the *Notepad* program.
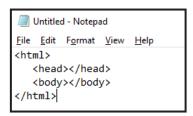

3. Click the *Notepad* icon to open Notepad.

3.   Create the outline of an HTML web page.

All HTML web pages have the same basic framework. Start by typing the following code into Notepad:

```
<html>
    <head></head>
    <body></body>
</html>
```

This is the basic framework of a web page. The *<html>* tag tells the web browser that this is an HTML document and that everything between *<html>* and *</html>* is HTML code.

The *<head>* tag's purpose is to contain information that will not be displayed on the page itself, including the page title and links to other files that may contain styles or scripts.

The *<body>* tag contains the main body of the web page, including the content that will be displayed on the page.

```
Untitled - Notepad
File  Edit  Format  View  Help
<html>
    <head></head>
    <body></body>
</html>
```

4.   Add some text to the page.
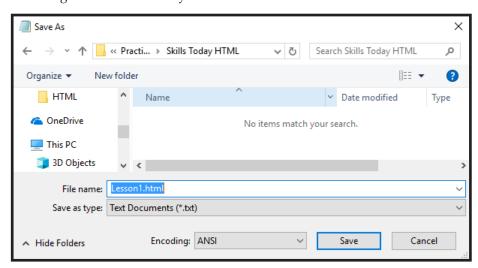
The page is currently completely empty. Add some text between the *<body>* and *</body>* tags as shown:

```
<html>
    <head></head>
    <body>
       Hello world!
    </body>
</html>
```

You now have a web page that will display the words *Hello world!* (or whichever text you chose to add).

5.   Save the web page as: **Lesson1.html**

1.   Click *File > Save As...* from the menu at the top of the Notepad window.

2.   Navigate to the folder of your choice and save the file as: **Lesson1.html**

```
Save As                                                          ×
←  →  ∨  ↑  | « Practi...  ›  Skills Today HTML   ∨  Ö   Search Skills Today HTML    ₽
Organize ▼    New folder                                   ▤▤ ▼    ❓
  | HTML         ^    Name              ^    ∨ Date modified       Type
  ☁ OneDrive
  💻 This PC            No items match your search.
  📁 3D Objects  ∨  <                                           >
       File name: | Lesson1.html                                    ∨
     Save as type: | Text Documents (*.txt)                         ∨
  ∧ Hide Folders     Encoding: | ANSI        ∨  |  Save  |  |  Cancel  |
```

You can save to any folder on your computer, but I have chosen to save into the *C:\Practice\Skills Today HTML* folder.

You will see how to view the page that you've just created in your web browser in the next lesson.

note

.html or .htm?

The letters after the full stop in a file's name are known as the *file extension*. File extensions tell computers the type of the file.

In this lesson you saved a web page with the *.html* file extension, but you could also have used *.htm*. Both file extensions work in exactly the same way, so why are there two options?

The answer is that early Windows computers enforced a 3 character limit on file extensions. This meant that they couldn't use the *.html* extension that was used by Unix web servers.

.html was the original extension and it is the extension used by this book. Some might argue for *.htm* because most other files have 3 character extensions, but there's really no difference as long as you are consistent in the extension you use.