HyperText Markup Language/Introduction

HTML is a markup language used in most of the pages of the World Wide Web. HTML files are text files that, unlike completely plain text, contain additional formatting markup—sequences of characters telling web browsers what parts of text should be bold, where the headings are, or where tables, table rows and table cells start and end. HTML may be displayed by a visual web browser, a browser that reads the text of the page to the user, a Braille reader that converts pages to a braille format, email client, or a wireless device like a cellular phone.

Before we start
To author and test HTML pages, you will need an editor and a web browser. HTML can be edited in plain text editors, including those that highlight HTML markup with colors to make it easier to read. There are also WYSIWYG (what you see is what you get) editors of HTML, and complex WYSIWYG editors with website project management and development environments
(It is also a good idea to learn XHTML, Javascript, ASP.NET and/or PHP to make high quality websites, though.) Plain text editors include Notepad (or Notepad++) for Microsoft® Windows, TextEdit for Mac, or Vim, Emacs and others for Linux.

Commercial HTML editors include Adobe Contribute CS5 and Dreamweaver CS5 (both Win/Mac), and Microsoft's Visual Web Developer (Win). There are also free HTML editors out there including Evrsoft First Page (Win), Mozilla KompoZer (Win/Mac/Lin) and Quanta Plus (Lin). It is usually better to gain a basic knowledge of HTML using a code-based HTML editor before delving into the WYSIWYG editors (all of the previous in this paragraph).

To preview your documents, you'll need a web browser. To make your documents look good to the greatest number of readers, test the documents in several browsers. Each browser has slightly different rendering, and most have their quirks, resulting in certain sequences of correctly written HTML rendered incorrectly.

Microsoft Internet Explorer is the most widely used browser, as of November 2010 having a 46% market share. Other common browsers include Google Chrome, Mozilla Firefox, Safari, and Opera. To make sure that your documents are readable in a text only environment, you can use Lynx.

A simple document
Let's start with a simple document. Write this code in your editor (or copy-and-paste it) and save it as "index.html" or "index.htm". The file must be saved with the exact extension, or will not be rendered correctly.

```html
<!DOCTYPE HTML PUBLIC "-//W3C//DTD HTML 4.01//EN" "http://www.w3.org/TR/html4/strict.dtd">
<html>
<head>
    <title> Simple document </title>
</head>
<body>
    <p>The text of the document goes here.</p>
</body>
</html>
```

Now open the document in your browser and see the result. From the above example, we can deduce certain essentials of a HTML document:

- The first line of a valid HTML document must state which version of HTML the document uses. This example uses the strict variant of HTML version 4.01.
- The HTML document begins with a `<html>` tag and ends with its counterpart, the `</html>` tag.
• Within the `<html></html>` tags, there are two main pairs of tags, `<head></head>` and `<body></body>`.
• Within the `<head></head>` tags, there are the `<title></title>` tags which enclose the textual title to be shown in the title bar of the web browser.
• Within the `<body></body>` is a paragraph marked by a `<p></p>` tag pair.
• Most tags must be written in pairs between which the effects of the tag will be applied.
  • `<em>This text is emphasized</em>` – *This text is emphasized*
  • This text includes `<code>computer code</code>` – This text includes `computer code`
  • `<em>This text is emphasized and has <code>computer code</code></em>` – *This text is emphasized and has `computer code`*
• HTML tag pairs must be aligned to encapsulate other tag pairs, for example:
  • `<code><em>This text is both code and emphasized</em></code>` – *This text is both code and emphasized*
  • A mistake: – `<em><code>This mark up is erroneous</code></em>`