

Core Web Programming
Marty Hall • Larry Brown

Introduction to HyperText Markup Language (HTML)

Agenda

- **Introduction to HTML**
- **Creating and publishing a Web page**
- **Validating a document**
- **Main HTML elements**
- **Block-level HTML elements**
- **Text-level HTML element**
- **Creating hypertext links**
- **Adding images to documents**
- **Building tables**

The World Wide Web

- **The World Wide Web**
 - The set of computers on the Internet that support HTTP
 - Not a separate network.
- **HTML (HyperText Markup Language)**
 - The language used to design web pages
 - Derivative of SGML (standardized markup language) - a metalanguage - that enables the development of other markup languages
 - Authored by Tim Berners-Lee and others at the European Organization for Nuclear Research (CERN) in 1989 to allow a large number of individuals to create, transmit and display electronic documents in a variety of computing platforms in a consistent manner

The World Wide Web

- **HTTP (HyperText Transfer Protocol)**
 - Created by Tim Berners-Lee
 - The language used by a WWW client (e.g. Netscape, Internet Explorer) to request documents from a WWW server (i.e. the program running at Web sites like amazon.com or yahoo.com)
- **Browsers**
 - Programs that load and display the contents of Web documents
 - Early browsers supported only text files
 - In 1992, a group of students and faculty at the Univ of Illinois authored the browser Mosaic which supported HTML 1.0.
 - It provided the ability to display graphical images and operated almost identically on most platforms

HyperText Markup Language

- **Text Mixed with Markup Tags**
 - Tags Enclosed in Angle Brackets
(`<H1>Introduction</H1>`)
- **What Does Markup Describe?**
 - Appearance
 - Layout
 - Content (Can't Enforce an Exact Look)
- **Changes in HTML 3.2 to HTML 4.0**
 - Standardization of frames
 - Deprecation of formatting elements (vs. style sheets)
 - Improved cell alignment and grouping in tables
 - Mouse and keyboard events for nearly all elements
 - Internationalization features

HTML Example

```
<!DOCTYPE HTML PUBLIC "-//W3C//DTD HTML 4.0 Transitional//EN">
<HTML>
<HEAD>
<TITLE>Home Page for Marty Hall</TITLE>
...
<BODY BGCOLOR="#FDF5E6" TEXT="#000000" LINK="#551A8B" ALINK="#FF0000">
<CENTER>
<TABLE BORDER=3 BGCOLOR="#3366FF">
  <TR><TD><STRONG CLASS="title">Home Page
</TABLE>
<P>
<TABLE>
  <TR><TD><STRONG><B>Marty Hall</B><BR>
  President<BR>
  <A HREF="http://www.coreservlets.com">www.coreservlets.com</A><BR>
  coreservlets.com, Inc.</A><BR>
  6 Meadowsweet Ct., Suite B1<BR>
  Reisterstown, MD 21136-6020<BR>
  <I>email:</I>
  <A HREF="mailto:hall@coreservlets.com">hall@coreservlets.com</A><BR>
  <I>Phone:</I> (410) 429-5535<BR>
  <I>Fax:</I> (410) 429-4931</STRONG></TD>
  <TD><IMG SRC="images/Marty-JHU-Head-Small.jpg"
  ALT="Marty" WIDTH="225" HEIGHT="263" HSPACE="5"></TD></TR>
</TABLE>...
```



Creating and Publishing a Web Page

1. Create an HTML document
2. Place the document in a world-accessible directory (often `public_html` or `www`) on a system running an HTTP server

```
Unix> cd
Unix> chmod a+x .    (Note the ".")
Unix> mkdir public_html
Unix> chmod a+x public_html
```

3. Access the web page through ***http://hostname/~username/filename***

- e.g. `http://www.apl.jhu.edu/~hall/test.html`
- If the filename is omitted, a system default filename is assumed (often `index.html`)
 - e.g. `http://www.apl.jhu.edu/~hall/` refers to the file `index.html` in hall's `public_html` directory

Creating and Publishing a Web Page, cont.

4. Validate the Document

- Check the syntax using a formal HTML validator
 - <http://www.htmlhelp.com/tools/validator/>
 - <http://validator.w3.org/>
- The version of HTML against which the document is validated is based on the DOCTYPE
`<!DOCTYPE HTML PUBLIC "-//W3C//DTD HTML 4.0 Transitional//EN">`
- The WWW Consortium recently added advice that Web pages include information on the character set, even though ASCII or Latin-1 is the default. The validator at <http://validator.w3.org/> gives warnings if you omit this. You can ignore such warnings if you wish.

HTML Document Template

```
<!DOCTYPE HTML PUBLIC "-//W3C//DTD HTML 4.0 Transitional//EN">
<HTML>
<HEAD>
  <TITLE>Title</TITLE>
</HEAD>

<BODY>
<H1>Main Heading</H1>
<!-- Rest of page goes here -->

</BODY>
</HTML>
```

← Goes on browser's title bar
May not appear in printouts

← Main heading. Often used as title
Appears in printouts

← HTML comment
Replace with body of
WWW page



Main HTML Elements

1. DOCTYPE

2. HTML

3. HEAD

- TITLE element required
- **<title></title>**
Puts the name of the document in the title bar

Main HTML Elements

4. BODY Element

- Tiled Background: `<BODY BACKGROUND= "URL">`
 - causes the image to tile as the background of the page
- Background Color `<BODY BGCOLOR= "$$$$">`
 - solid background color of the page
- Text Color `<BODY TEXT= "$$$$">`
 - color of the text throughout the page
- Link Color `<BODY LINK= "$$$$">`
 - color of all links throughout the page
- Visited Link `<BODY VLINK= "$$$$">`
 - color of all links that have already been clicked on by visitor
- Active Link `<BODY ALINK= "$$$$">`
 - color of link while being selected

More on Colors

- There are 16,777,216 “true” colors
- \$\$\$\$ is an RGB hex combination; the 216 Web-safe colors are recommended for consistency
- Though not perfectly consistent across all computers, they are close enough
- All modern color monitors support 8-bit color (256 colors)
- Software designers work with color cubes (three dimensions RGB) and $6^3 = 216$ is the largest cube that does not exceed 256
 - Each component is made up of the values 00,33,66,99,CC and FF
<http://www.fortunecity.com/help/development/html/colors.shtml>
 - Valid names are aqua, gray, navy, silver, black, green, olive, teal, blue, lime, purple, white, fuchsia, maroon, red and yellow

META Element

- Placed between the `<HEAD>` and `</HEAD>` tags
- Records document information, forwards and refreshes pages
 - `<META name="author" content="Richard G. Lowe, Jr.">`
 - `<META name="description" content="Everything you wanted to know about stamps, from prices to history.">`
 - `<META name="keywords" content="stamps,stamp collecting, stamp history, prices, stamps for sale">`
 - Informs the search engine of the main topics or points of a document. Due to extreme misuse of this Metatag, its relevance in many search engines has been downgraded. Keywords in the tag should also appear within your content, description, title, alt tags and elsewhere within your document to be considered by the search engines.
 - `<META NAME="refresh" content="1; URL=http://internet-tips.net/About/about.htm">`
 - indicates the document displays for a specified amount of seconds (n) before switching to a new URL. The full absolute URL must be specified.
 - Recommended when you need to move or rename a page on your site. Since search engines may have your old page in their index, it is wise to include a redirect page to not lose those visitors.

Block-Level Elements

- **Headings**
 - H1 ... H6
 - ALIGN
- **Basic Text Sections**
 - P
 - ALIGN
 - PRE
 - WIDTH
 - ADDRESS
 - BLOCKQUOTE

Block-Level Elements, cont.

- **Lists**
 - OL
 - LI
 - UL
 - LI
 - DL
 - DT
 - DD
- **Tables and Forms (Postponed)**
- **Misc.**
 - HR
 - DIV
 - CENTER
 - MULTICOL (Netscape only)

Headings

- **Heading Types**

- `<H1 ...> ... </H1>`

- `<H2 ...> ... </H2>`

- `<H3 ...> ... </H3>`

- `<H4 ...> ... </H4>`

- `<H5 ...> ... </H5>`

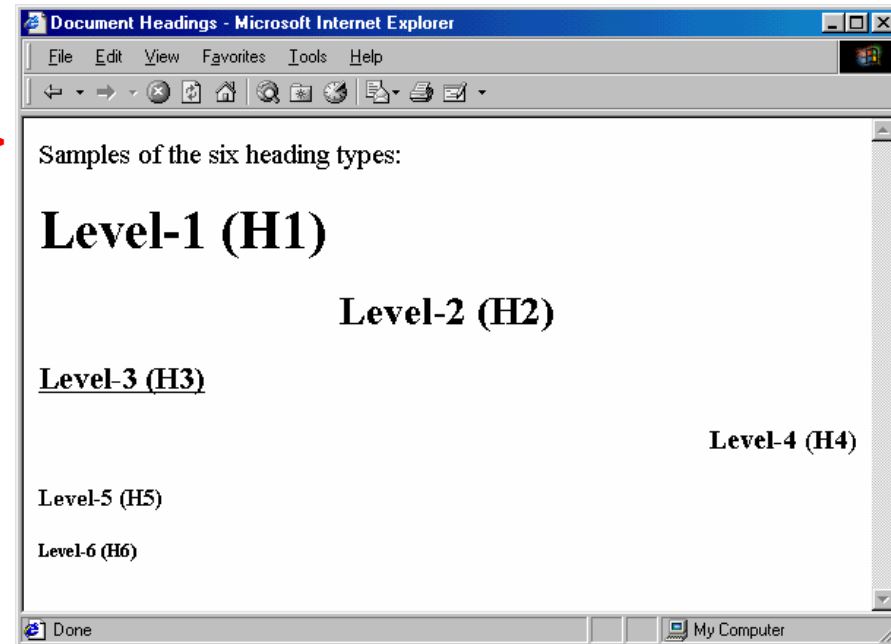
- `<H6 ...> ... </H6>`

- **Attributes: ALIGN**

- Values: LEFT (default), RIGHT, CENTER

Headings, Example

```
<!DOCTYPE HTML PUBLIC "-//W3C//DTD HTML 4.0 Transitional//EN">
<HTML>
<HEAD>
  <TITLE>Document Headings</TITLE>
</HEAD>
<BODY>
Samples of the six heading types:
<H1>Level-1 (H1)</H1>
<H2 ALIGN="CENTER">Level-2 (H2)</H2>
<H3><U>Level-3 (H3)</U></H3>
<H4 ALIGN="RIGHT">Level-4 (H4)</H4>
<H5>Level-5 (H5)</H5>
<H6>Level-6 (H6)</H6>
</BODY>
</HTML>
```



P – The Basic Paragraph

- **Attributes: ALIGN**

- LEFT (default), RIGHT, CENTER. Same as headings.
- Whitespace ignored (use
 for line break)
 - Consecutive <P>'s do not yield multiple blank lines
- End Tag is Optional:

```
<BODY>
<P>
Paragraph 1
</P>
<P>
Paragraph 2
</P>
<P>
Paragraph 3
</P>
</BODY>
```

Fully-Specified

```
<BODY>
Paragraph 1
<P>
Paragraph 2
<P>
Paragraph 3
</BODY>
```

Equivalent with Implied Tags

Preformatted Paragraphs

- **The PRE Element**

- `<PRE> ... </PRE>`

- **Attributes: WIDTH**

- Expected width in characters. Not widely supported.

- **Problem: Special Characters**

`<PRE>`

```
if (a<b) {  
    doThis();  
} else {  
    doThat();  
}
```

`</PRE>`

Desired Character	HTML Required
<	<
>	>
&	&
"	"
Non-breaking space	

OL: Ordered (Numbered) Lists

- **OL Element**

- ``
 `...`
 `...`
 `...`
 ``

- Attributes: TYPE: 1, A, a, I, i START: a positive integer

- **List entries: LI**

- `<LI ...> ... ` (End Tag Optional)
 - Attributes: (When inside OL) VALUE, TYPE

A sample list:

```
<OL>
  <LI>List Item One
  <LI>List Item Two
  <LI>List Item Three
</OL>
```

A sample list:

```
1. List Item One
2. List Item Two
3. List Item Three
```

Nested Ordered Lists

```
<OL TYPE="I">
<LI>Headings
  <LI>Basic Text Sections
  <LI>Lists
    <OL TYPE="A">
      <LI>Ordered
        <OL TYPE="1">
          <LI>The OL tag
            <OL TYPE="a">
              <LI>TYPE
              <LI>START
              <LI>COMPACT
            </OL>
          <LI>The LI tag
        </OL>
      <LI>Unordered
        <OL TYPE="1">
          <LI>The UL tag
          <LI>The LI tag
        </OL>
      <LI>Definition
        <OL TYPE="1">
          <LI>The DL tag
          <LI>The DT tag
          <LI>The DD tag
        </OL>
    </OL>
  <LI>Miscellaneous
</OL>
```

- I. Headings
- II. Basic Text Sections
- III. Lists
 - A. Ordered
 - 1. The OL tag
 - a. TYPE
 - b. START
 - c. COMPACT
 - 2. The LI tag
 - B. Unordered
 - 1. The UL tag
 - 2. The LI tag
 - C. Definition
 - 1. The DL tag
 - 2. The DT tag
 - 3. The DD tag
- IV. Miscellaneous

UL: Unordered (Bulleted) Lists

- **UL Element**

- ``
 `...`
 `...`
 `...`
 ``

- **Attributes: TYPE, COMPACT**

- TYPE is DISC, CIRCLE, or SQUARE

- **List entries: LI (TYPE)**

- TYPE is DISC, CIRCLE, or SQUARE

A sample list:

```
<UL>  
  <LI>List Item One  
  <LI>List Item Two  
  <LI>List Item Three  
</UL>
```

A sample list:

- List Item One
- List Item Two
- List Item Three

UL: Custom Bullets

```
<UL TYPE="DISC">
  <LI>The UL tag
    <UL TYPE="CIRCLE">
      <LI>TYPE
        <UL TYPE="SQUARE">
          <LI>DISC
          <LI>CIRCLE
          <LI>SQUARE
        </UL>
      <LI>COMPACT
    </UL>
  <LI>The LI tag
    <UL TYPE="CIRCLE">
      <LI>TYPE
        <UL TYPE="SQUARE">
          <LI>DISC
          <LI>CIRCLE
          <LI>SQUARE
        </UL>
      <LI>VALUE
    </UL>
  </UL>
```

Unordered Lists

- The UL tag
 - TYPE
 - DISC
 - CIRCLE
 - SQUARE
 - COMPACT
- The LI tag
 - TYPE
 - DISC
 - CIRCLE
 - SQUARE
 - VALUE

Text-Level Elements

- **Physical Character Styles**

- B, I, TT, U, SUB, SUP, SMALL, BIG, STRIKE, S, BLINK
- FONT
 - SIZE
 - COLOR
 - FACE
- BASEFONT
- SIZE

- **Logical Character Styles**

- EM, STRONG, CODE, SAMP, KBD, DFN, VAR, CITE

Text-Level Elements (Continued)

- **Hypertext Links**

- A
 - HREF, NAME, TARGET, ...

- **Images**

- IMG
 - SRC (required), ALT, ALIGN, WIDTH, HEIGHT, HSPACE, VSPACE, BORDER, USEMAP, ISMAP

- **Misc. Text-Level Elements**

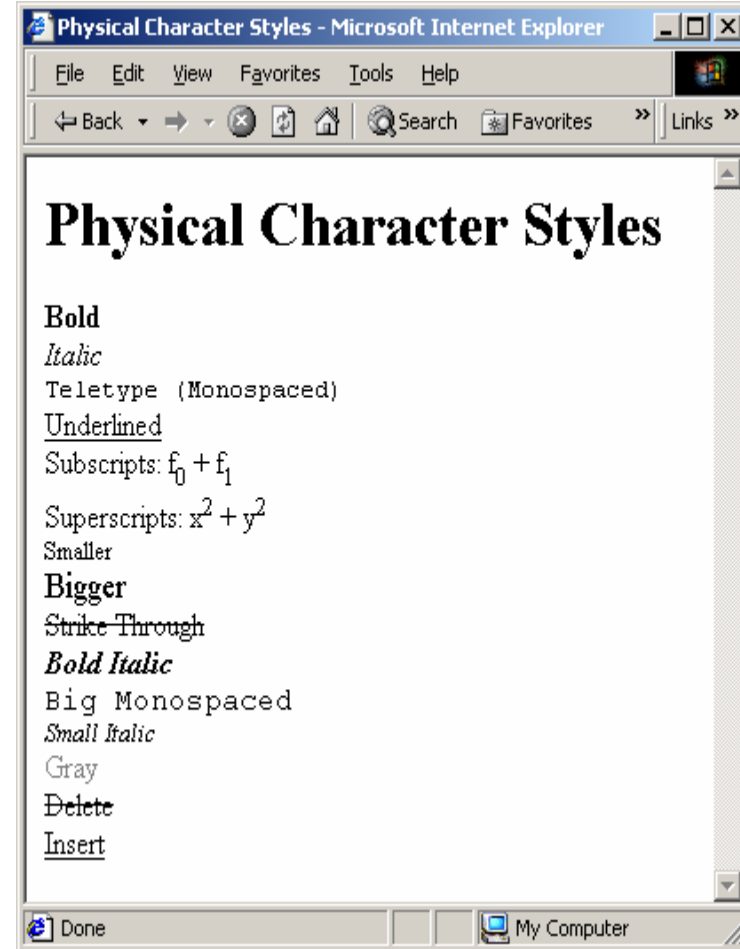
- BR (Explicit line break)
- AREA (Client-side image maps)
- APPLET (Java),
- ...

Physical Character Styles, Example

...

```
<H1>Physical Character Styles</H1>
<B>Bold</B><BR>
<I>Italic</I><BR>
<TT>Teletype (Monospaced)</TT><BR>
<U>Underlined</U><BR>
Subscripts: f<SUB>0</SUB> +
            f<SUB>1</SUB><BR>
Superscripts: x<SUP>2</SUP> +
              y<SUP>2</SUP><BR>
<SMALL>Smaller</SMALL><BR>
<BIG>Bigger</BIG><BR>
<STRIKE>Strike Through</STRIKE><BR>
<B><I>Bold Italic</I></B><BR>
<BIG><TT>Big Monospaced</TT></BIG><BR>
<SMALL><I>Small Italic</I></SMALL><BR>
<FONT COLOR="GRAY">Gray</FONT><BR>
<DEL>Delete</DEL><BR>
<INS>Insert</INS><BR>
```

...

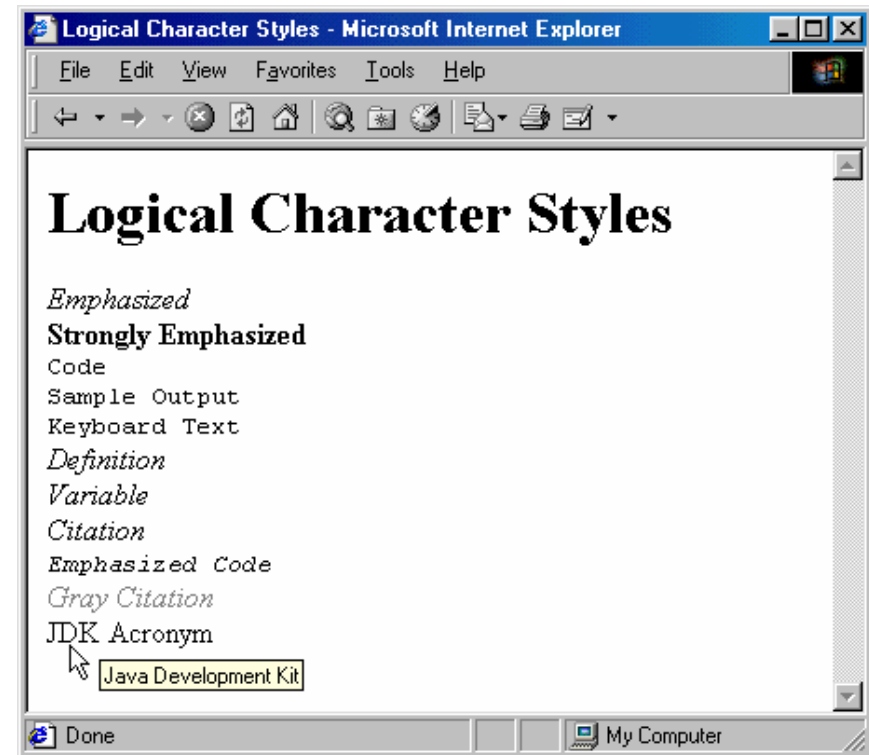


Logical Character Styles, Example

...

```
<H1>Logical Character Styles</H1>  
<EM>Emphasized</EM><BR>  
<STRONG>Strongly  
  Emphasized</STRONG><BR>  
<CODE>Code</CODE><BR>  
<SAMP>Sample Output</SAMP><BR>  
<KBD>Keyboard Text</KBD><BR>  
<DFN>Definition</DFN><BR>  
<VAR>Variable</VAR><BR>  
<CITE>Citation</CITE><BR>  
<EM><CODE>Emphasized  
  Code</CODE></EM><BR>  
<FONT COLOR="GRAY"><CITE>Gray  
  Citation</CITE></FONT><BR>  
<ACRONYM TITLE="Java Development  
  Kit">JDK Acronym</ACRONYM>
```

...



Hypertext Links

- Links can contain images and other text-level elements (i.e., `<A HREF...> ... `)
- Link to Absolute URL
 - Use a complete URL beginning with `http://`
Java is discussed in
``
Chapter 2``.
- Link to Relative URL
 - Use a filename or relative path to filename
 - Interpreted as location of current file
Java is discussed in
``Chapter 2``.

Hypertext Links (Continued)

- **Link to Section**

- Use a section name (see below) preceded by #
Images are discussed in
`Section 2.`

- **Link to Section in URL**

- Use absolute or relative URL, then #, then section name
Images are discussed in
`
Sec. 2 of Chap. 1.`

- **Naming a Section**

- Use `` and do not include the pound sign
`<H2>Images</H2>`

IMG: Embedding Images

- **Example**

```
<IMG SRC="SomeFile.gif" ALT="My Dog"  
      WIDTH=400 HEIGHT=300>
```

- **Attributes:**






- SRC (required)
- ALT (technically required)
- ALIGN (see <BR CLEAR="ALL">)
- WIDTH, HEIGHT
- HSPACE, VSPACE
- BORDER
- USEMAP, ISMAP

Image Alignment, Example

```
<!DOCTYPE HTML PUBLIC "-//W3C//DTD HTML 4.0 Transitional//EN">
<HTML>
<HEAD><TITLE>Image Alignment</TITLE></HEAD>
<BODY>
<H1 ALIGN="CENTER">Image Alignment</H1>
<TABLE BORDER=1>
  <TR><TH>Alignment
    <TH>Result
  <TR><TH><CODE>LEFT</CODE>
    <TD><IMG SRC="rude-pc.gif" ALIGN="LEFT"
      ALT="Rude PC" WIDTH=54 HEIGHT=77>
      This positions the image at the left side,
      with text flowing around it on the right.
  <TR><TH><CODE>RIGHT</CODE>
    <TD><IMG SRC="rude-pc.gif" ALIGN="RIGHT"
      ALT="Rude PC" WIDTH=54 HEIGHT=77>
      This positions the image at the right side,
      with text flowing around it on the left.
    . . .
</TABLE>
</BODY>
</HTML>
```

Image Alignment, Result

The screenshot shows a Netscape browser window with the title "Image Alignment - Netscape". The browser's address bar and menu bar are visible. The main content area displays a table titled "Image Alignment" with five rows, each illustrating a different alignment option for an image within a text block. The table has two columns: "Alignment" and "Result".

Alignment	Result
LEFT	This positions the image at the left side, with text flowing around it on the right. 
RIGHT	This positions the image at the right side, with text flowing around it on the left. 
TOP	Here, the image runs into the paragraph and the line containing the image is aligned with the image top. 
BOTTOM	Here, the image runs into the paragraph and the line containing the image is aligned with the image bottom. 
MIDDLE	Here, the image runs into the paragraph and the line containing the image is aligned with the image center. 

Tables

- **Template**

```
<TABLE BORDER=1>  
  <CAPTION>Table Caption</CAPTION>  
  <TR><TH>Heading1</TH>      <TH>Heading2</TH></TR>  
  <TR><TD>Row1 Col1 Data</TD><TD>Row1 Col2 Data</TD></TR>  
  <TR><TD>Row2 Col1 Data</TD><TD>Row2 Col2 Data</TD></TR>  
  <TR><TD>Row3 Col1 Data</TD><TD>Row3 Col2 Data</TD></TR>  
</TABLE>
```

Table Caption

Heading1	Heading2
Row1 Col1 Data	Row1 Col2 Data
Row2 Col1 Data	Row2 Col2 Data
Row3 Col1 Data	Row3 Col2 Data

TABLE Element Attributes

- **ALIGN**

- The `ALIGN` attribute gives the horizontal alignment of the table as a whole
 - Legal values are `LEFT`, `RIGHT`, and `CENTER`, with `LEFT` being the default

- **BORDER**

- This specifies the width in pixels of the border around the table
- This is in addition to the border around each cell (the `CELLSPACING`).
 - The default is zero, which also results in the visible 3D divider between cells being turned off

- **CELLSPACING**

- This gives the space in pixels between adjacent cells. Drawn as a 3D line if `BORDER` is nonzero, otherwise empty space in the background color is used
 - The default is usually about 3

TABLE Element Attributes (Continued)

- **CELLPADDING**

- CELLPADDING determines the empty space, in pixels, between the cell's border and the table element
- The default is usually about 1

- **WIDTH**

- This specifies the width of the table, either in pixels (<TABLE WIDTH=250>) or as a percentage of the current browser window width (<TABLE WIDTH="75%">)

- **BGCOLOR**

- Specify the background color of the table TABLE (also legal for TR, TD, and TH)

TABLE Element Attributes (Continued)

- **RULES**

- HTML 4.0 attribute that specifies which inner dividing lines are drawn
- All are drawn if this attribute is omitted
- Legal values are NONE, ROWS, COLS, and ALL

- **FRAME**

- Specifies which outer borders are drawn
- All four are drawn if this attribute is omitted
- Legal values are BORDER or BOX (all), VOID (none), ABOVE (top), BELOW (bottom), HSIDES (top and bottom, despite the somewhat confusing name), VSIDES (left and right), LHS (left), and RHS (right)

Table CAPTION

- **Attribute**
 - ALIGN (Values: TOP, BOTTOM)
- **Usage**
 - An enclosing borderless table may give more flexibility than the built-in CAPTION.

Heading1	Heading2
Row1 Col1 Data	Row1 Col2 Data
Row2 Col1 Data	Row2 Col2 Data
Row3 Col1 Data	Row3 Col2 Data

TR: Table Row

- **TR is used to define each row in the table**
- **Each row will then contain TH and/or TD entries**
 - **ALIGN**
 - ALIGN (legal values LEFT, RIGHT, or CENTER) is used to set the default horizontal alignment for table cells
 - **VALIGN**
 - VALIGN (legal values TOP, MIDDLE, or BOTTOM) is used to set the default vertical alignment for table cells
 - **BGCOLOR**
 - Sets the color for the table row, overriding any values set for the table as a whole via the BGCOLOR attribute of TABLE

Table Cells: TH and TD

- **COLSPAN**

- COLSPAN defines a heading or cell data entry that spans multiple columns

```
<TABLE BORDER=1>  
<TR><TH COLSPAN=2>Col 1&2 Heading  
    <TH>Col3 Heading  
<TR><TD>Col1 Data  
    <TD>Col2 Data  
    <TD>Col3 Data  
</TABLE>
```

Col 1&2 Heading		Col3 Heading
Col1 Data	Col2 Data	Col3 Data

Table Cells: TH and TD (Continued)

- **ROWSPAN**
 - ROWSPAN defines a heading or cell data entry that spans multiple rows; similar to COLSPAN
- **ALIGN**
 - LEFT, RIGHT, CENTER, JUSTIFY and CHAR.
 - E.g., the following aligns entries on a decimal point

```
<TD ALIGN="CHAR" CHAR=".">
```
- **VALIGN**
 - TOP, BOTTOM, MIDDLE
- **WIDTH, HEIGHT**
 - Values in pixels only (no percentages officially allowed)
- **BGCOLOR, BACKGROUND**
 - Same as for TABLE and TR