

Basic HTML Rules

Basic Syntax:

- HTML document are composed of *elements*, optionally followed by *attributes*, which may have values.
- These units are enclosed within angle brackets and are called *tags*.
- Most of the tags have a starting tag (eg.), some text or other tags, followed by an ending tag (ending tags starts with a / character eg.).
- Tag elements and attributes are *not* case-sensitive.
- In general, multiple white-spaces, line breaks, indentation etc. are ignored.
- Even in case of invalid or incorrect document content, the browser tries to render the document with best-effort approach (error messages, warnings etc. are not displayed).

Document Structure:

- The document should start with a DOCTYPE declaration (check the page source of the first link).
- There are *five* types of elements:
 1. *Void Elements*: They contain only the start tag, no end tag. Eg. br, hr, img, embed, input, link, meta, param etc.
 2. *Raw-text Elements*: These are used to include external text files, eg. script, style.
 3. *Escapable raw-text Elements*: These can't use ambiguous ampersand character, along with other restrictions, eg. textarea, title.
 4. *Foreign Elements*: Elements from MathML and SVG belongs to this category.
 5. *Normal Elements*: All other elements belong to this type.
- Content is placed between the start tag and the end tag.
- Void elements can't have contents (since there is no end tag).
- Attributes:
 - They are placed inside the element's start tag.
 - Attributes are Name-Value pairs.
 - An equal sign is placed between the Name and the Value.
 - Attribute names are *not* case-sensitive.
 - Multiple attributes can be placed in a tag, separated by one or more blank spaces.
 - Attributes can be specified in *four* ways:
 1. *Just the attribute name*: <hr noshade>
 2. *Unquoted*: <hr size=2>
 3. *Single Quoted*: <hr size='2'>
 4. *Double Quoted*: <hr size="2">
- Newline: Both DOS style (CR+LF), and Unix style (CR) newline characters are allowed, apart from a third style which consists of LF character.
- Character References (escape characters) can be specified in *three* ways:
 1. *Named Character Reference*: & (&), © (©), ™ (™) etc.
 2. *Decimal Numeric Character Reference*: & (&), ₹ (₹) etc.
 3. *Hexadecimal Numeric Character Reference*: & (&), & (&), ₹ (₹) etc.
- Comments are multi-line by default. Comments start with the <!-- tag, and ends with the --> tag.

For more details, check the w3.org site: [The HTML Syntax](https://www.w3.org/html/)