

MILESTONE 2

What is HTML

DOM

HTML and XHTML

Elements, Tags, Attributes and Layouts

Semantic and Non-Semantic Elements

Classes, and Ids

Formatting, Charsets

What is HTML?

- HTML is the Hyper Text Markup Language - template of a webpage - it gives the structure of the website - most widely used language to write Web Pages
- **Hypertext** refers to the way in which Web pages (HTML documents) are linked together. Thus, the link available on a webpage is called Hypertext.
- **Markup Language** - you use HTML to simply "mark-up" a text document with tags that tell a Web browser how to structure it to display.

```
<!DOCTYPE html>
```

```
<html>
```

```
<head>
```

```
<title>This is document title</title>
```

```
</head>
```

```
<body>
```

```
<h1>This is a heading</h1>
```

```
</body>
```

```
</html>
```

Tag	Description
<!DOCTYPE...>	This tag defines the document type and HTML version.
<html>	This tag encloses the complete HTML document and mainly comprises of document header which is represented by <head>...</head> and document body which is represented by <body>...</body> tags.
<head>	This tag represents the document's header which can keep other HTML tags like <title>, <link> etc.
<title>	The <title> tag is used inside the <head> tag to mention the document title.
<body>	This tag represents the document's body which keeps other HTML tags like <h1>, <div>, <p> etc.
<h1>	This tag represents the heading.

- **Void elements** : `<img.../>`, `<hr />` and `
` element (No closing tag)
- An HTML element is defined by a *starting tag*. If the element contains other content, it ends with a *closing tag*.
- An attribute is used to define the characteristics of an HTML element and is placed inside the element's opening tag. All attributes are made up of two parts: a **name** and a **value**

- **Internationalization Attributes**

There are three internationalization attributes, which are available for most XHTML elements. `dir`, `lang`, `xml:lang`

- The **dir** attribute allows you to indicate to the browser about the direction in which the text should flow
- The **lang** attribute allows you to indicate the main language used in a document, but this attribute was kept in HTML only for backwards compatibility with earlier versions of HTML.
- `lang` attribute has been replaced by the **xml:lang** attribute in new XHTML documents
- `<bdo>...</bdo>` element stands for Bi-Directional Override and it is used to override the current text direction

DOM

- The DOM (Document Object Model) is the programming interface for HTML and XML documents, it defines the logical structure of documents and the way a document is accessed and manipulated.
- It is the API used by web browsers and other applications to access the contents of an HTML document.
- The DOM represents the structure as a tree, nodes can contain child-nodes, nodes with no children are said leaf nodes.

HTML and XHTML

- The primary difference between HTML and XHTML lies in how strictly the rules of coding are followed. For instance, HTML is more flexible and omits certain tags without creating a fuss while XHTML demands a well-written code for any function.

Feature	HTML	XHTML
Basis	SGML (Standard Generalized Markup Language)	XML (Extensible Markup Language)
Syntax	More lenient —allows for missing closing tags for certain elements (e.g., <code>
</code>), unquoted attribute values (for certain attributes), attribute minimization (e.g., <code><input type=text></code> instead of <code><input type="text"></code>)	Stricter —All elements to have closing tags, attributes to be quoted, and no attribute minimization
Case sensitivity	Not case-sensitive (e.g., <code>
</code> is equivalent to <code>
</code>)	Case-sensitive (all tags and attributes must be lowercase)
DTD (Document type definition)	Can use multiple DTDs depending on the version of HTML (e.g., HTML 4.01 or HTML5)	Requires a specific DTD (e.g., XHTML 1.0 Strict or XHTML 1.1)

Elements, Tags, Attributes and Layouts

Elements

- HTML elements are the basic building blocks of a webpage.
- An element typically consists of a start tag, content, and an end tag.
- Elements can be nested, meaning one element can contain other elements.
- They define the structure and content of a webpage.
- Examples of elements include paragraphs (`<p>`), headings (`<h1>`), and divs (`<div>`).

Tags

- Tags are used to create HTML elements.
- Most tags come in pairs: an opening tag and a closing tag.
- Some tags are self-closing, like `` and `
`.
- Tags are enclosed within angle brackets, e.g., `<tagname>`.
- Tags define the type of content an element contains, such as text, images, or links.

Attributes

- Attributes provide additional information about an HTML element.
- They are specified in the opening tag of an element.
- Attributes are written as name/value pairs, separated by an equals sign (=).
- Common attributes include `id`, `class`, `style`, `href`, and `src`.

- Attributes can modify the behavior or appearance of an element, such as linking to another page or applying styles.

Layout

- Layouts control the positioning and arrangement of elements on a webpage.
- CSS is commonly used to define layouts, including Flexbox, Grid, and positioning.
- Block elements take up the full width of their container, while inline elements take up only as much space as needed.
- Flexbox allows for the alignment and distribution of space among items in a container along a single axis.
- Grid layout enables complex layouts by dividing the page into rows and columns, making it easier to align content.

Semantic and Non-Semantic Elements

- A semantic element clearly describes its meaning to both the browser and the developer
Eg : `<form>`, `` and `<table>` (Clearly defines its content).
- A non-semantic element doesn't clearly describes its meaning to both the browser and the developer
Eg : `div` and `span` (Tells nothing about its content).

Classes, and Ids

- Classes and IDs make referencing HTML elements from scripts and stylesheets easier.
- The class attribute can be used on one or more tags and is used by CSS for styling - global attribute .
 - `class="class1 class2 class3"`

Uses of class

- Targeting element using CSS: Similar to ids, we can select an element from the document using its class name. We use "." before the class name for targeting a particular class.
- IDs are intended to refer to a single element, meaning the same ID should never be used twice - global attribute
 - `id="uniqueid"`

Uses of ID

- Targeting elements using CSS: Using id, we can easily target the elements using CSS and manipulate their style.
- Manipulating element using Script: We can use *getElementById* and *querySelector(#id)* to manipulate the element with given id.

Class	ID
Multiple element can have same class .	Id is unique in the whole document.
"." is used as selector in CSS	"#" is used as selector in CSS
" getElementsByClassName " is used for targeting element using JS.	" getElementsById " is used for targeting element using JS.

Formatting, Charsets

- HTML Formatting is a process of formatting text for better look and feel.
- HTML provides us ability to format text without using CSS. There are many formatting tags in HTML. These tags are used to make text bold, italicized, or underlined. There are almost 14 options available that how text appears in HTML and XHTML.
- In HTML the formatting tags are divided into two categories:
 - Physical tag: These tags are used to provide the visual appearance to the text.
 - Logical tag: These tags are used to add some logical or semantic value to the text.
- Formatting elements were designed to display special types of text:
 - - Bold text
 - - Important text
 - <i> - Italic text
 - - Emphasized text
 - <mark> - Marked text
 - <small> - Smaller text
 - - Deleted text
 - <ins> - Inserted text

<sub> - Subscript text

<sup> - Superscript text

- HTML Charset is also called HTML Character Sets or HTML Encoding. It is used to display an HTML page properly and correctly because for displaying anything correctly, a web browser must know which character set (character encoding) to use.

HTML CHARACTER ENCODING

- ASCII character set
- ANSI character set
- ISO-8859 character set
- UTF-2 character set