XML

XML stands for **E**xtensible **M**arkup **L**anguage. It is a text-based markup language derived from Standard Generalized Markup Language (SGML).

XML tags identify the data and are used to store and organize the data, rather than specifying how to display it like HTML tags, which are used to display the data. XML is not going to replace HTML in the near future, but it introduces new possibilities by adopting many successful features of HTML.

There are three important characteristics of XML that make it useful in a variety of systems and solutions −

* **XML is extensible** − XML allows you to create your own self-descriptive tags, or language, that suits your application.
* **XML carries the data, does not present it** − XML allows you to store the data irrespective of how it will be presented.
* **XML is a public standard** − XML was developed by an organization called the World Wide Web Consortium (W3C) and is available as an open standard.

## XML Usage

A short list of XML usage says it all −

* XML can work behind the scene to simplify the creation of HTML documents for large web sites.
* XML can be used to exchange the information between organizations and systems.
* XML can be used for offloading and reloading of databases.
* XML can be used to store and arrange the data, which can customize your data handling needs.
* XML can easily be merged with style sheets to create almost any desired output.
* Virtually, any type of data can be expressed as an XML document.

# XML - Syntax

syntax rules to write an XML document. Following is a complete XML document −

<?xml version = "1.0"?>

<contact-info>

<name>Tanmay Patil</name>

<company>TutorialsPoint</company>

<phone>(011) 123-4567</phone>

</contact-info>

You can notice there are two kinds of information in the above example −

* Markup, like <contact-info>
* The text, or the character data

## XML Declaration

The XML document can optionally have an XML declaration. It is written as follows −

<?xml version = "1.0" encoding = "UTF-8"?>

Where *version* is the XML version and *encoding* specifies the character encoding used in the document.

### Syntax Rules for XML Declaration

* The XML declaration is case sensitive and must begin with "**<?xml>**" where "**xml**" is written in lower-case.
* If document contains XML declaration, then it strictly needs to be the first statement of the XML document.
* The XML declaration strictly needs be the first statement in the XML document.

## Tags and Elements

An XML file is structured by several XML-elements, also called XML-nodes or XML-tags. The names of XML-elements are enclosed in triangular brackets < > as shown below −

<element>

### Syntax Rules for Tags and Elements

**Element Syntax** − Each XML-element needs to be closed either with start or with end elements as shown below −

<element>....</element>

or in simple-cases, just this way −

<element/>

**Nesting of Elements** − An XML-element can contain multiple XML-elements as its children, but the children elements must not overlap. i.e., an end tag of an element must have the same name as that of the most recent unmatched start tag.

The Following example shows incorrect nested tags −

<?xml version = "1.0"?>

<contact-info>

<company>GRHU

</contact-info>

</company>

**EXAMPLE OF AN XML DOCUMENT**

Suppose we are making xml for a college and inside college we have classes we will represent in hierarchical structure like following:

<? Xml version =1.0 character encoding. “ISO-8895-1”> //DECLARATION OF XML

<College> //ROOT ELEMENT

<class1>

<name>Deepika</name> **CHILD ELEMENTS**

<roll no>23</roll no>

</class1>

<class2>

<name>D</name>

<roll no>24</roll no>

</class2>

</College>