

## 第2课 学习HTML

### 1、HTML 是什么

HTML ( Hyper Text Markup Language ) 是用来描述网页的一种语言，也叫超文本标记语言。

#### 1-1、查看网页的 HTML 代码

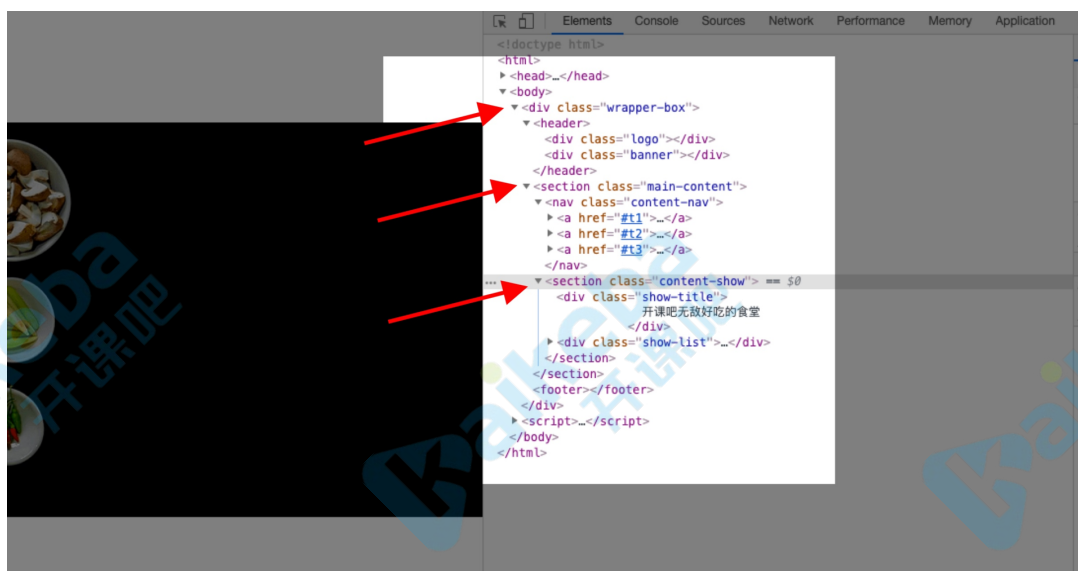
##### ①、显示网页源代码

在网页任意地方点击鼠标右键，然后点击“显示网页源代码”。（Windows系统的电脑还可以使用快捷键ctrl+u来查看网页源代码）

##### ②、检查

- windows：在网页的空白处点击右键，然后选择“检查”（快捷方式是ctrl+shift+i）；
- mac：在网页的空白处点击右键，然后选择“检查”（快捷键 command + option + l(大写 I)）

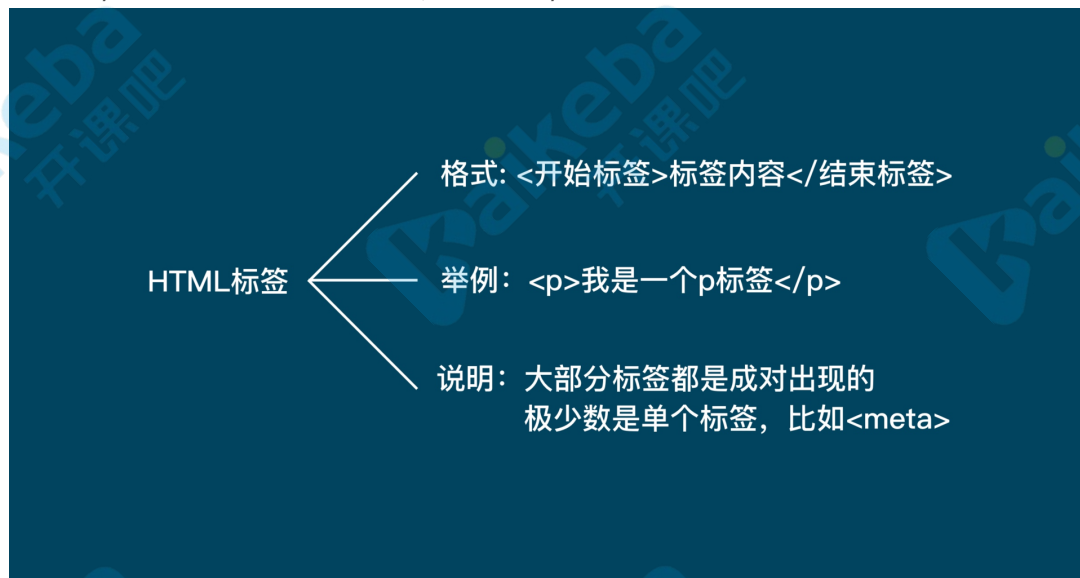
#### 1-2、HTML的层级



## 2、HTML 的组成

### 2-1、标签和元素

(1) 标签：夹在尖括号<>中间的字母，标签通常是成对出现的：前面的是【开始标签】，比如<body>；后面的是【结束标签】，如</body>；



(2) 元素：开始标签+结束标签+中间的所有内容组成。

经常使用的HTML元素		
开始标签	标签意义	结束标签
<h1>	一级标题	</h1>
<h2>	二级标题	</h2>
<a>	链接	</a>
<div>	块, 用来包含其他标签	</div>
<p>	段落	</p>

注意：HTML标签是可以嵌套标签的，而且可以多层嵌套。

### 2-2、<head>和<body>

## HTML基本结构

```
<html>

  <head>
    网页头的内容，这里的内容不会显示在网页里
  </head>

  <body>
    网页体的具体内容
  </body>

</html>
```

HTML文档的最外层标签一定是<html>，里面嵌套着<head>元素与<body>元素。  
<head>元素代表了【网页头】，<body>元素代表了【网页体】，这是最基本的网页结构。

- (1) 【网页头】的内容不会被直接呈现在浏览器里的网页正文中；
- (2) 【网页体】的内容是会直接显示在网页正文中的。

```
<head>
  <meta charset="utf-8">
  <title>title标签里面包含的就是网页的名字</title>
</head>
```

注：

<meta>标签定义了这个网页的编码格式  
<title>元素就是用来设置这个网页的标题，会显示在浏览器的选项卡上

## 2-3、属性

注意：HTML的属性和Python中的属性不是一个东西



### HTML属性

属性的格式为：属性名 = “属性值”  
举例：style=“color:#ff0000”

style属性可以用来给文本类的标签定义样式，包括但不限于字体，字体颜色，字体大小，字体间距，文字对齐方式

### ①、h1 标签及其对应的属性 style

```
1 <h1 style="color:#20b2aa;">这个书苑不太冷</h1>
```

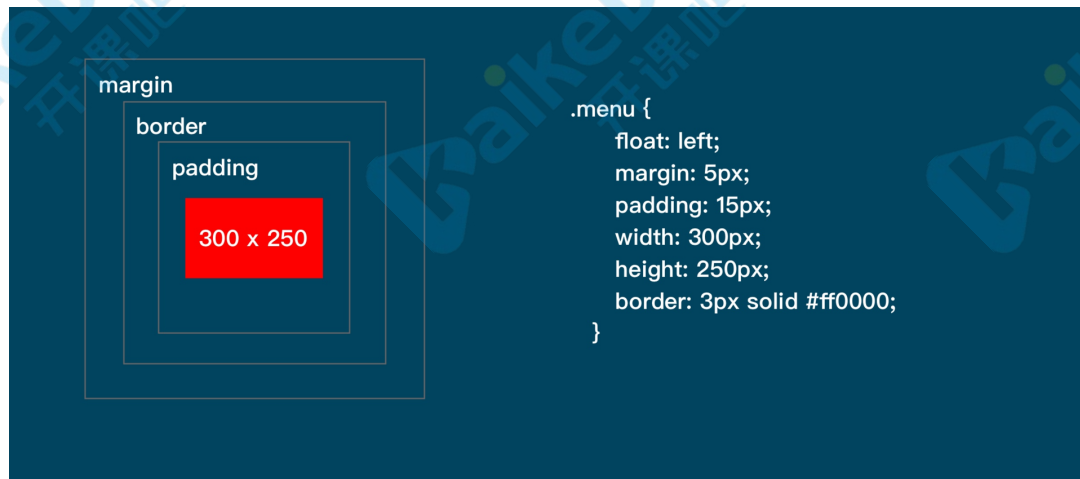
### ②、a 标签及其对应的属性 href

```
1 <a href="https://wordpress-edu-3autumn.localprod.forc.work/">我是一个链接，  
点我试试</a>
```

## 2-4、class和id

目的：为了查找、定位元素，或者为元素设置样式

### 2-4-1、class



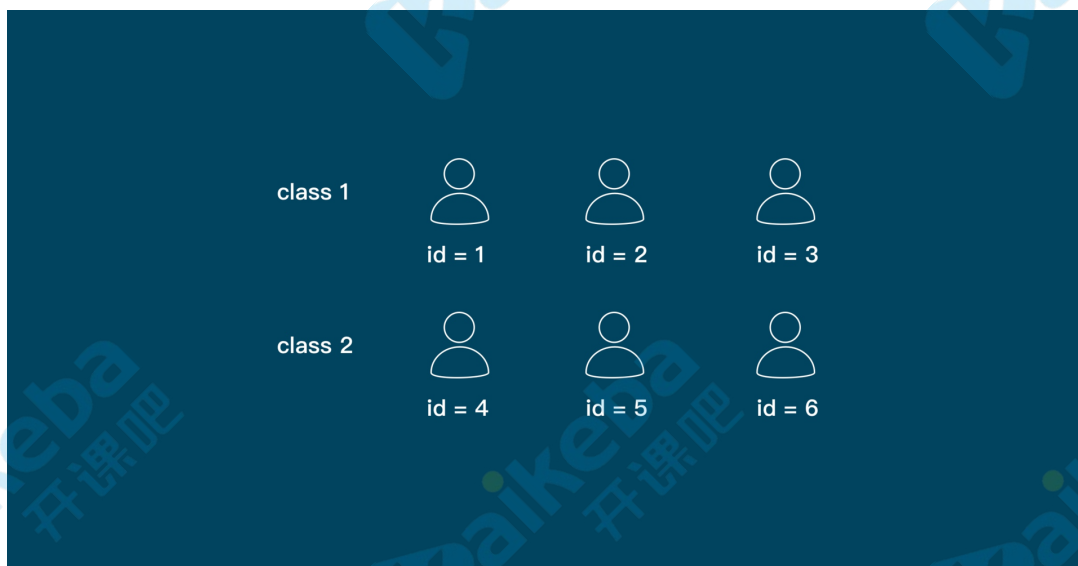
```
1 <style>  
2 .book {  
3 /*以下是.book的具体样式规定*/  
4 float: left; /*控制元素浮动*/  
5 margin: 5px; /*外边距为5像素*/  
6 padding: 15px; /*内边距为15像素*/  
7 width: 300px; /*宽度为350像素*/  
8 height: 250px; /*高度为240像素*/  
9 border: 3px solid #ff0000; /*边框为3像素，颜色为ff0000*/  
10 }  
11 </style>
```

在HTML中，“class”属性可以被多个不同的标签使用，我们在<head>的<style>中定义的“.menu”样式，只要一个标签的class = “menu”，那么它都会继承这个样式

### 2-4-2、id

id属性用于标识「唯一」的元素，而class用于标识「一系列」的元素

“id”就像是学生的学生证号码，每个人都是唯一的，而学生们可以属于同一个班级，班级就像class



总结：

常见属性	
属性	知识点
class	为html元素设置类名
id	为html元素设置唯一的id
href	链接标签的url
style	定义元素的样式

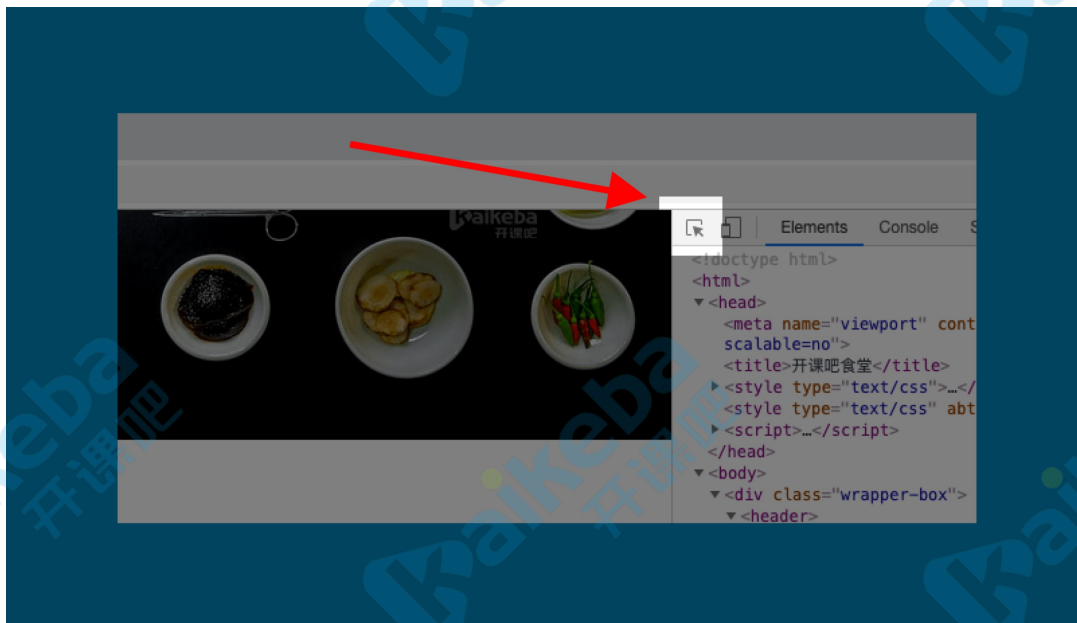
## 2-5、阅读HTML

- 1、点击URL：<https://xiaoke.kaikeba.com/example/canteen/index.html>
- 2、在网页上点击右键，选择「检查」（或者是直接按键盘的F12，适用于Chrome浏览器），打开开发者工具
- 3、我们可以看到在<head>，中有<meta>标签，还有<style>标签。在<style>中定义了很多属性

## 2-6、修改HTML

首先，我们来通过Chrome浏览器的开发者工具来修改HTML代码：





当我们用这个按钮找到网页对应的源代码后，就可以在开发者工具中随时修改，这样就会改变这个网页。

这些改动，都只是在当前的浏览器里改动。无法影响网页所在的服务器的源代码，所以当你重新打开这个网页的时候会恢复原状

## 2-7、用爬虫从HTML提取数据

可以用response.text来获得文本数据再打印出来

```
1 import requests
2 res =
  requests.get('https://xiaoke.kaikeba.com/example/canteen/index.html')
5 print(res.text)
```