

UT2_3.- LENGUAJES DE MARCAS EN ENTORNO WEB. HTML5

RESULTADOS DE APRENDIZAJE ASOCIADOS
<ol style="list-style-type: none">1. Reconoce las características de lenguajes de marcas analizando e interpretando fragmentos de código.2. Utiliza lenguajes de marcas para la transmisión de información a través de la Web analizando la estructura de los documentos e identificando sus elementos3. Establece mecanismos de validación para documentos XML utilizando métodos para definir su sintaxis y estructura.
CRITERIOS DE EVALUACIÓN
<ul style="list-style-type: none">- De RA1 – desde CEA hasta CEK- De RA2 – desde CEA hasta CEH- De RA3 – desde CEA hasta CEH

UT2 3.- LENGUAJES DE MARCAS EN ENTORNO WEB. HTML5

Índice de contenido

1.- Etiquetas y atributos HTML – Texto	3
2.- Estructurar el texto	5
2.1.- Párrafos <p>	5
2.2.- Secciones <h1> <h2> <h3> <h4> <h5> <h6>	7
3.- Marcado básico de texto	11
3.1.- Etiquetas y 	11
3.2.- Etiquetas <ins> y 	14
3.3.- Etiqueta <blockquote>	16
4.- Marcado avanzado de texto	18
4.1.- Etiquetas <abbr> y <acronym>	18
4.2.- Etiquetas <dfn>	20
4.3.- Etiquetas <cite>	21
5.- Marcado genérico de texto 	24
6.- Espacios en blanco y nuevas líneas	26
6.1.- Nuevas líneas 	27
6.2.- Espacios en blanco “ ”	29
6.3.- Texto preformateado <pre> y <code>	30
7.- Codificación de caracteres	34

1.- Etiquetas y atributos HTML – Texto

Los atributos son elementos compuestos por dos pares claramente diferenciadas (el nombre del atributo y su valor correspondiente) y se añaden en la etiqueta de un elemento HTML.

La manera de asignar un valor a un atributo es mediante el signo igual (=). El valor puede ponerse con comillas dobles o simples.

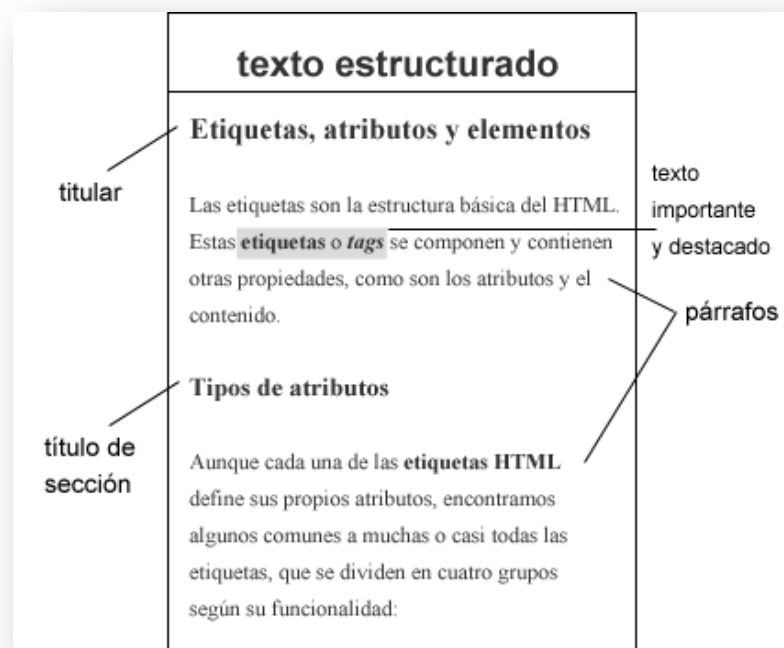
{ Texto

La mayor parte del contenido de las páginas HTML habituales está formado por texto, llegando a ser el más del 90% del código de la página.

HTML define etiquetas para **estructurar** el contenido en secciones y define otras etiquetas para **marcar** elementos importantes dentro del texto

La tarea inicial del editor de contenidos HTML consiste en **estructurar** el texto original definiendo:

- Párrafos
- Titulares
- Títulos de sección



Por tanto, el proceso a seguir es el siguiente:

- 1) Estructurar el **texto simple**. Es decir, indicar las diferentes zonas o secciones de las que se compone el texto.

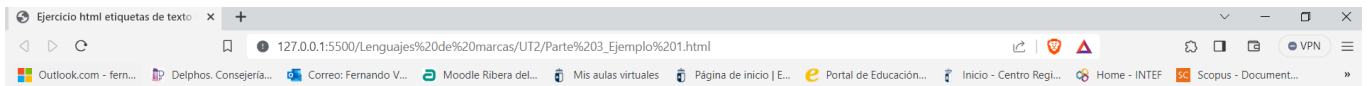


Los textos estructurados utilizan etiquetas para delimitar cada párrafo y títulos de sección para delimitar cada una de las secciones que forman el texto.

- 2) El siguiente paso consiste en marcar los diferentes elementos dentro del propio texto:

- Definiciones
- Abreviaturas
- Textos importantes
- Textos modificados
- Citas a otras referencias
- Etc.

Ejemplo:



Biografía de Tim Berners Lee

Berners-Lee trabajó en el CERN desde junio hasta diciembre de 1980. Durante ese tiempo, propuso un proyecto basado en el *hipertexto* para facilitar la forma de compartir y la puesta al día de la información entre investigadores. En este periodo también construyó un programa llamado *ENQUIRE* que no llegó a ver la luz.

Después de dejar el CERN, en 1980, se fue a trabajar a la empresa de *John Poole Image Computer Systems Ltd.*, pero regresó al CERN otra vez en 1984.

El nacimiento de la web

En 1989, el CERN era el nodo de Internet más grande de Europa y Berners-Lee vio la oportunidad de unir Internet y el hipertexto (**HTTP** y **HTML**), de lo que surgiría la **World Wide Web**.

Desarrolló su primera propuesta de la Web el 12 de marzo de 1989,2 pero no tuvo mucho eco, por lo que en 1990 y con la ayuda de *Robert Cailliau*, hicieron una revisión que fue aceptada por su gerente, *Mike Sendall*.

Usó ideas similares a las que había usado en el sistema *Enquire*, para crear la **World Wide Web**, para esto diseñó y construyó el primer navegador (llamado WorldWideWeb y desarrollado con NEXTSTEP) y el primer servidor Web al que llamó httpd (HyperText Transfer Protocol daemon).

La democratización de la web

El primer servidor Web se encontraba en el CERN y fue puesto en línea el 6 de agosto de 1991. Esto proporcionó una explicación sobre lo que era el World Wide Web, cómo uno podría tener un navegador y cómo establecer un servidor Web. Este fue también el primer directorio Web del mundo, ya que Berners-Lee mantuvo una lista de otros sitios Web aparte del suyo.

Debido a que tanto **el software del servidor como del cliente fue liberado de forma gratuita** desde el CERN, el corazón de Internet Europeo en esa época, su difusión fue muy rápida. El número de servidores Web pasó de veintiséis en 1992 a doscientos en octubre de 1995 lo que refleja cual fue la velocidad de la difusión de internet.

El nacimiento del W3C

En 1994 entró en el *Laboratorio de Ciencias de la Computación e Inteligencia Artificial* del *Massachusetts Institute of Technology*. Se trasladó a EE. UU. y puso en marcha el **W3C**, que dirige actualmente.

El **W3C** es un organismo internacional de estandarización de tecnologías Web dirigido conjuntamente por el *Instituto Tecnológico de Massachusetts*, el *ERCIM francés* y la *Universidad de Keiō* en Japón. Este organismo decidió que todos sus estándares fuesen libres, es decir, que los pudiese utilizar todo el mundo libremente sin coste alguno, lo que sin lugar a dudas fue una de las grandes razones para que la Web haya llegado a tener la importancia que tiene hoy en día.

En su libro *Tejiendo la red*, publicado en 1999, Berners-Lee explica por qué la **tecnología web es libre y gratis**. Se considera al mismo tiempo el **inventor** y el **protector** de la web.

Como se puede observar en el ejemplo anterior, el texto tiene una estructura definida y, además, algunas palabras tienen un formato determinado.

2.- Estructurar el texto

La forma más sencilla de estructurar un texto consiste en separarlo por párrafos.

Además, HTML permite incluir títulos que delimitan cada una de las secciones.

2.1.- Párrafos <p>



Manual:

https://www.w3schools.com/tags/tag_p.asp

Una de las etiquetas más utilizadas en HTML es la etiqueta **<p>**, que permite definir los párrafos que forman el texto de una página.

Ejemplo:



Código de ejemplo:

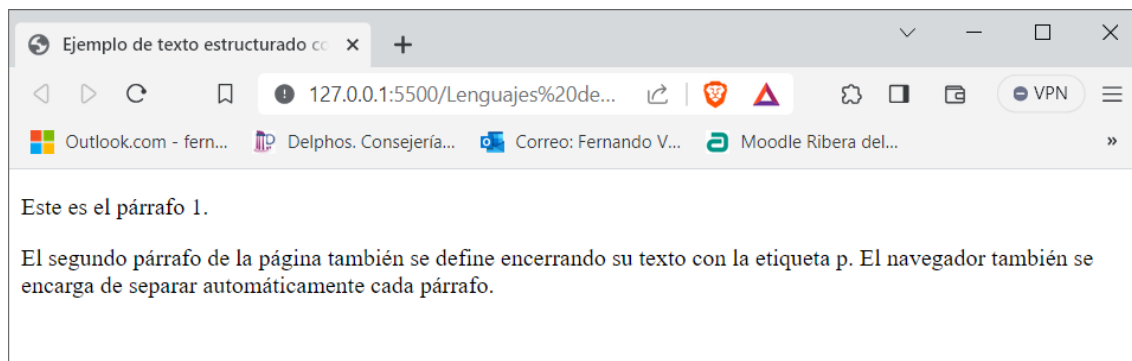
```
<!DOCTYPE html>
<html lang="es">
  <head>
    <title>Ejemplo de texto estructurado con párrafos</title>
  </head>
  <body>
    <p> Este es el párrafo 1. </p>

    <p>El segundo párrafo de la página también se define encerrando
    su texto con la etiqueta p. El navegador también se encarga de
    separar automáticamente cada párrafo. </p>

  </body>
```

```
</html>
```

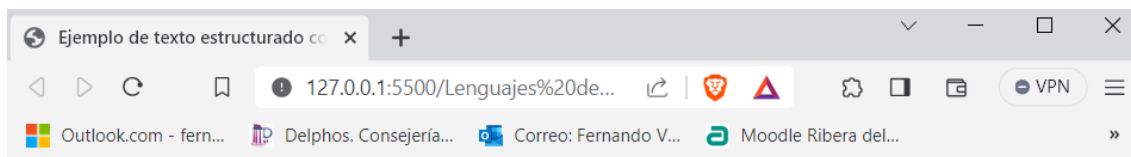
Este ejemplo se visualiza de la siguiente forma:



Otro ejemplo:

```
1 <!DOCTYPE html>
2 <html lang="en">
3 <head>
4   <meta charset="UTF-8">
5   <meta name="viewport" content="width=device-width, initial-scale=1.0">
6   <title>Ejemplo de texto estructurado con párrafos</title>
7 </head>
8 <body>
9   <p>Este es el texto que forma el primer párrafo de la página.
10     Los párrafos pueden ocupar varias líneas y el navegador se encarga
11     de ajustar su longitud al tamaño de la ventana.
12   </p>
13   <p>El segundo párrafo de la página también se define encerrando
14     su texto con la etiqueta p. El navegador también se encarga de
15     separar automáticamente cada párrafo.
16   </p>
17 </body>
18 </html>
```

Lo que veríamos en el navegador, es:



Este es el texto que forma el primer párrafo de la página. Los párrafos pueden ocupar varias líneas y el navegador se encarga de ajustar su longitud al tamaño de la ventana.

El segundo párrafo de la página también se define encerrando su texto con la etiqueta p. El navegador también se encarga de separar automáticamente cada párrafo.

La siguiente tabla recoge la definición de la etiqueta <p>:

Etiqueta	<p>
Atributos comunes	básicos , internacionalización y eventos
Atributos propios	-
Tipo de elemento	Bloque
Descripción	Delimita el contenido de un párrafo de texto

Los párrafos creados con HTML son elementos de bloque, por lo que **siempre ocupan toda la anchura de la ventana del navegador**.

No tienen atributos específicos, pero sí que les pueden asignar los atributos de HTML básicos, de internacionalización y de eventos.

2.2.- Secciones <h1> <h2> <h3> <h4> <h5> <h6>

Las páginas HTML habituales suelen tener una estructura más compleja que la que se puede crear solamente mediante párrafos. De hecho, es habitual que las páginas se dividan en diferentes secciones jerárquicas.

Los títulos de sección se utilizan para delimitar el comienzo de cada sección de la página.



HTML permite crear secciones de hasta seis niveles de importancia.

De esta forma, aunque una página puede definir cualquier número de secciones, sólo puede incluir seis niveles jerárquicos.

Las etiquetas que definen los títulos de sección son **<h1>**, **<h2>**, **<h3>**, **<h4>**, **<h5>** y **<h6>**.

- ☐ La etiqueta **<h1>** es la de mayor importancia y por tanto se utiliza para definir los titulares de la página.
- ☐ La importancia del resto de etiquetas es descendiente, de forma que la etiqueta **<h6>** es la que se utiliza para delimitar las secciones menos importantes de la página.

A continuación, se muestra la definición formal de la etiqueta **<h1>**, siendo idéntica la definición del resto de etiquetas referidas a los títulos de sección:

Etiqueta	<h1>
Atributos comunes	básicos , internacionalización y eventos
Atributos propios	-
Tipo de elemento	Bloque
Descripción	Define los títulos de las secciones de mayor importancia de la página.



Al igual que la etiqueta **<p>, las etiquetas de título de sección son elementos de bloque y no tienen atributos específicos.**

Ejemplo:


```
1 <!DOCTYPE html>
2 <html lang="en">
3 <head>
4   <meta charset="UTF-8">
5   <meta name="viewport" content="width=device-width, initial-scale=1.0">
6   <title>Document</title>
7 </head>
8 <body>
9   <h1>Titular de la página</h1>
10  <p>Párrafo de introducción...</p>
11  <h2>La primera sub-sección</h2>
12  <p>Párrafo de contenido...</p>
13  <h2>Otra subsección</h2>
14  <p>Más párrafos de contenido...</p>
15 </body>
16 </html>
```

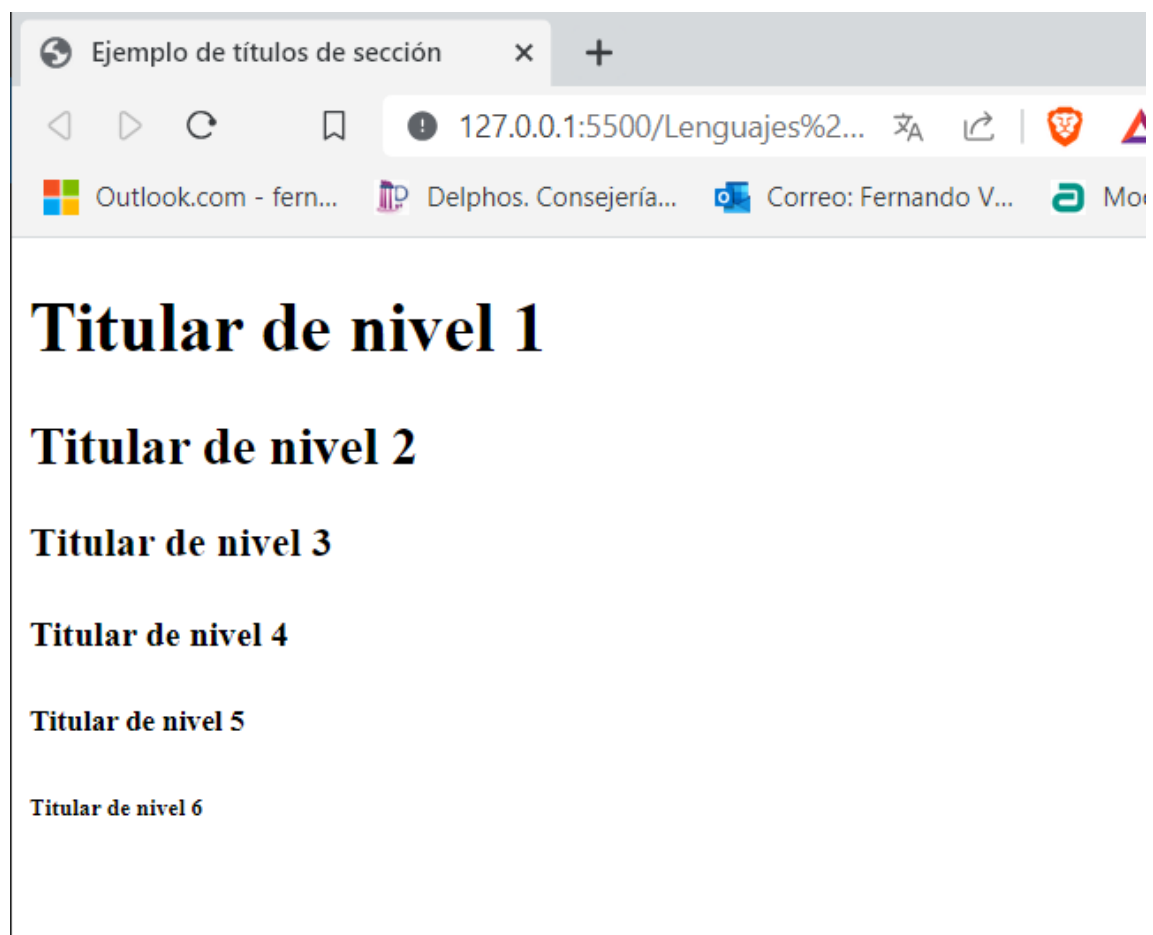
En el navegador, se verá de la siguiente forma:



A continuación, se muestra un ejemplo de uso de las etiquetas h1, h2, h3, h4, h5 y h6:

```
1 <!DOCTYPE html>
2 <html lang="en">
3 <head>
4   <meta charset="UTF-8">
5   <meta name="viewport" content="width=device-width, initial-scale=1.0">
6   <title>Ejemplo de títulos de sección</title>
7 </head>
8 <body>
9   <h1>Titular de nivel 1</h1>
10  <h2>Titular de nivel 2</h2>
11  <h3>Titular de nivel 3</h3>
12  <h4>Titular de nivel 4</h4>
13  <h5>Titular de nivel 5</h5>
14  <h6>Titular de nivel 6</h6>
15 </body>
16 </html>
```

La salida es la siguiente:



3.- Marcado básico de texto

Una vez estructurado el texto en párrafos y secciones, el siguiente paso es el **marcado de los elementos** que componen el texto. Los textos habituales están formados por elementos como:

- ☐ Palabras en negrita o cursiva
- ☐ Anotaciones y correcciones
- ☐ Citas a otros documentos externos

3.1.- Etiquetas `` y ``

Entre las etiquetas más utilizadas para marcar texto se encuentran `` y ``. La definición formal de estas dos etiquetas se muestra a continuación:

Etiqueta	<code></code>
Atributos comunes	básicos , internacionalización y eventos
Atributos propios	-
Tipo de elemento	En línea
Descripción	Realza la importancia del texto que encierra

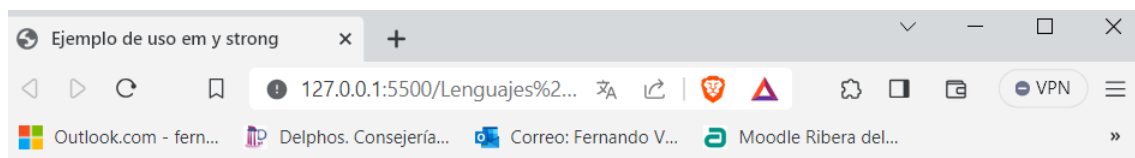
Etiqueta	<code></code>
Atributos comunes	básicos , internacionalización y eventos
Atributos propios	-
Tipo de elemento	En línea
Descripción	Realza con la máxima importancia el texto que encierra

- ☐ La etiqueta `` marca un texto indicando que su importancia es mayor que la del resto del texto.
- ☐ La etiqueta `` indica que un determinado texto es de la mayor importancia dentro de la página.

Ejemplo:

```
1 <!DOCTYPE html>
2 <html lang="en">
3 <head>
4   <meta charset="UTF-8">
5   <meta name="viewport" content="width=device-width, initial-scale=1.0">
6   <title>Ejemplo de uso em y strong</title>
7 </head>
8 <body>
9   <p>
10     El lenguaje HTML permite marcar algunos segmentos de texto
11     como <em>muy importantes</em> y otros segmentos como <strong>
12     más importantes</strong>.
13   </p>
14 </body>
15 </html>
```

La salida es la siguiente:



El lenguaje HTML permite marcar algunos segmentos de texto como *muy importantes* y otros segmentos como **más importantes**.

Por defecto, los navegadores muestran los elementos `` en cursiva para hacer evidente su importancia y muestran los elementos `` en negrita, para indicar que son los más importantes.

ACTIVIDAD1 - UT2_3

Estructurar y marcar el siguiente texto extraído de la Wikipedia para que el navegador lo muestre con el aspecto de la siguiente imagen:

La exploración espacial

La **exploración espacial** designa los esfuerzos del hombre en estudiar el espacio y sus astros desde el punto de vista científico y de su explotación económica. Estos esfuerzos pueden involucrar tanto seres humanos viajando en naves espaciales como satélites con recursos de telemetría o sondas teleguiadas enviadas a otros planetas (orbitando o aterrizando en la superficie de estos cuerpos celestes).

Las personas que pilotan naves espaciales, o son pasajeros en ellas, se llaman astronautas (en Rusia: *cosmonautas*; en China: *taikonautas*). Técnicamente se considera astronauta a todo aquel que emprenda un vuelo sub-orbital (sin entrar en órbita) u orbital a como mínimo 100 km de altitud (considerado el límite externo de la atmósfera).

El cielo siempre ha atraído la atención y los sueños del hombre. Ya en 1634 se publicó la que se considera primera novela de ciencia ficción, *Somnium*, de **Johannes Kepler**, que narra un hipotético viaje a la Luna. Más tarde, en 1865, en una famosa obra de ficción titulada "*De la Terre à la Lune*", **Julio Verne** escribe sobre un grupo de hombres que viajó hasta la Luna usando un gigantesco cañón.

En Francia, **Georges Méliès**, uno de los pioneros del cine, tomaba la novela de Verne para crear "*Le voyage dans la Lune*" (1902), una de las primeras películas de ciencia ficción en la que describía un increíble viaje a la Luna. En obras como "*The War of the Worlds*" (1898) y "*The First Men in The Moon*" (1901), **Herbert George Wells** también se concibieron ideas de exploración del espacio y de contacto con civilizaciones extraterrestres.

Texto plano para usar y formatear:

“La exploración espacial designa los esfuerzos del hombre en estudiar el espacio y sus astros desde el punto de vista científico y de su explotación económica. Estos esfuerzos pueden involucrar tanto seres humanos viajando en naves espaciales como satélites con recursos de telemetría o sondas teleguiadas enviadas a otros planetas (orbitando o aterrizando en la superficie de estos cuerpos celestes). Las personas que pilotan naves espaciales, o son pasajeros en ellas, se llaman astronautas (en Rusia: cosmonautas; en China: taikonautas). Técnicamente se considera astronauta a todo aquel que emprenda un vuelo sub-orbital (sin entrar en órbita) u orbital a como mínimo 100 km de altitud (considerado el límite externo de la atmósfera).

El cielo siempre ha atraído la atención y los sueños del hombre. Ya en 1634 se publicó la que se considera primera novela de ciencia ficción, *Somnium*, de Johannes Kepler, que narra un hipotético viaje a la Luna. Más tarde, en 1865, en una famosa obra de ficción titulada "De la Terre à la Lune", Julio Verne escribe sobre un grupo de hombres que viajó hasta la Luna usando un gigantesco cañón.

En Francia, Georges Méliès, uno de los pioneros del cine, tomaba la novela de Verne para crear "Le voyage dans la Lune" (1902), una de las primeras películas de ciencia ficción en la que describía un increíble viaje a la Luna. En obras como "The War of the Worlds" (1898) y "The First Men in The Moon" (1901), Herbert George Wells también se concibieron ideas de exploración del espacio y de contacto con civilizaciones extraterrestres."

HTML también permite marcar de forma adecuada las modificaciones realizadas en el contenido de una página. En otras palabras, HTML permite indicar de forma clara el texto que ha sido eliminado y el texto que ha sido añadido a un determinado texto original.

3.2.- Etiquetas <ins> y

Las etiquetas utilizadas son **<ins>** y ****, cuya definición formal es la siguiente:

Etiqueta	<ins>
Atributos comunes	básicos , internacionalización y eventos
Atributos propios	<ul style="list-style-type: none"> • <code>cite = "url"</code> - Indica la URL de la página en la que se puede obtener más información sobre el motivo por el que se realizó la modificación. • <code>datetime = "fecha"</code> - Especifica la fecha y hora en la que se realizó el cambio
Tipo de elemento	Bloque y en línea
Descripción	Se emplea para marcar una modificación en los contenidos originales consistente en la inserción de un nuevo contenido

Etiqueta	
Atributos comunes	básicos , internacionalización y eventos
Atributos propios	<ul style="list-style-type: none"> • cite = "url" - Indica la URL de la página en la que se puede obtener más información sobre el motivo por el que se realizó la modificación. • datetime = "fecha" - Especifica la fecha y hora en la que se realizó el cambio
Tipo de elemento	Bloque y en línea
Descripción	Se emplea para marcar una modificación en los contenidos originales consistente en el borrado de cierto contenido



Las dos etiquetas cuentan con los mismos atributos específicos, que opcionalmente se pueden añadir para proporcionar más información sobre los cambios realizados.

- ☐ El atributo **cite** se emplea para indicar la dirección de un documento externo en el que se puede encontrar más información relacionada con la inserción o el borrado de texto.
- ☐ El atributo **datetime** puede utilizarse para indicar la fecha y la hora en la que se realizó cada cambio.

Ejemplo:

```

1 <!DOCTYPE html>
2 <html lang="en">
3 <head>
4   <meta charset="UTF-8">
5   <meta name="viewport" content="width=device-width, initial-scale=1.0">
6   <title>Ejemplo de ins y del</title>
7 </head>
8 <body>
9   <h1>Ejemplo de etiqueta ins y del</h1>
10  <p>El HTML, acrónimo inglés de Hyper Text Markup Language (lenguaje de
11    <del datetime="20091025" cite="http://librosweb.es/xhtml/capitulo_3/marcado_basico_de_texto.html">
12      marcado de hipertexto
13    </del>
14    <ins datetime="20091026" cite="http://librosweb.es/xhtml/capitulo_3.html"> marcas
15      hipertextuales
16    </ins>
17  ) es un lenguaje de marcado diseñado para estructurar textos y presentarlos en forma
18  de hipertexto.
19  </p>
20 </body>
21 </html>

```

En pantalla se ve tal que así:



Por defecto, el texto eliminado (marcado con la etiqueta ****) se muestra tachado de forma que el usuario pueda identificarlo fácilmente como un texto que formaba parte del texto original y que ya no tiene validez.

El texto insertado (marcado con la etiqueta **<ins>**) se muestra subrayado, de forma que el usuario pueda identificarlo como un texto nuevo que no formaba parte del texto original.

3.3.- Etiqueta **<blockquote>**

Por otra parte, en muchos tipos de páginas (artículos, noticias) es habitual citar literalmente un texto externo. HTML define la

etiqueta **<blockquote>** para incluir citas textuales en las páginas web. La definición de la etiqueta HTML con el nombre más largo se muestra a continuación:

Etiqueta	<blockquote>
Atributos comunes	básicos , internacionalización y eventos
Atributos propios	<ul style="list-style-type: none"> cite = "url" - Indica la dirección de la página web original de la que se extrae la cita
Tipo de elemento	Bloque
Descripción	Se emplea para indicar que el texto que encierra es una cita textual de otro texto externo

Permite indicar mediante el atributo **cite** la dirección de un documento del que se ha extraído la cita.

Ejemplo:

```

1  <!DOCTYPE html>
2  <html lang="en">
3  <head>
4      <meta charset="UTF-8">
5      <meta name="viewport" content="width=device-width, initial-scale=1.0">
6      <title>Ejemplo de blockquote</title>
7  </head>
8  <body>
9      <p>Según el W3C, el valor del
10         atributo <em>cite</em> en las etiquetas <strong>blockquote</strong> tiene el
11         siguiente significado:
12     </p>
13
14     <blockquote cite="http://www.w3.org/TR/html401/struct/text.html">El valor de este atributo
15         es una dirección URL que indica el documento original de la cita."
16     </blockquote>
17 </body>
18 </html>

```

La salida en el navegador es la siguiente:



Para indicar de forma clara que el texto es una cita externa, los navegadores muestran por defecto el texto del elemento **<blockquote>** con un gran margen en la parte izquierda.

4.- Marcado avanzado de texto

Las páginas y documentos más avanzados suelen incluir otros elementos importantes que se deben marcar de forma adecuada. Por ello, HTML incluye muchas otras etiquetas que permiten marcar más elementos del texto.

4.1.- Etiquetas <abbr> y <acronym>

La etiqueta **<abbr>** marca las abreviaturas de un texto y la etiqueta **<acronym>** se emplea para marcar las siglas o acrónimos del texto. Su definición es la siguiente:

Etiqueta	<abbr>
Atributos comunes	básicos , internacionalización y eventos
Atributos propios	<ul style="list-style-type: none"> • title = "texto" - Indica el significado completo de la abreviatura
Tipo de elemento	En línea
Descripción	Se emplea para marcar las abreviaturas del texto y proporcionar el significado de esas abreviaturas

Etiqueta	<acronym>
Atributos comunes	básicos , internacionalización y eventos
Atributos propios	<ul style="list-style-type: none"> • title = "texto" - Indica el significado completo del acrónimo o sigla
Tipo de elemento	En línea
Descripción	Se emplea para marcar las siglas o acrónimos del texto y proporcionar el significado de esas siglas

En ambos casos, el atributo **title** se puede utilizar para incluir el significado completo de la abreviatura o sigla.

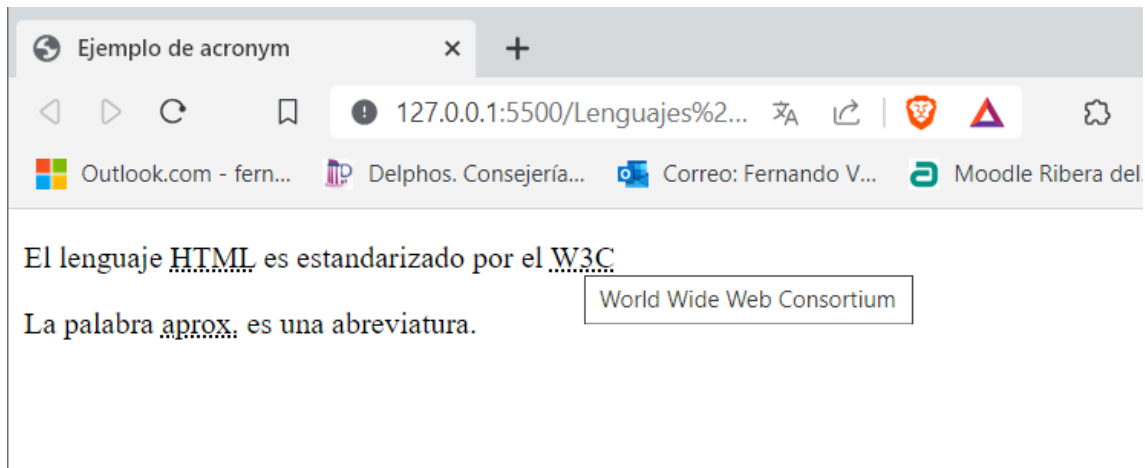
Ejemplo:

```

1  <!DOCTYPE html>
2  <html lang="en">
3  <head>
4      <meta charset="UTF-8">
5      <meta name="viewport" content="width=device-width, initial-scale=1.0">
6      <title>Ejemplo de acronym</title>
7  </head>
8  <body>
9      <p>El lenguaje <acronym title="HyperText Markup Language">HTML</acronym>
10     es estandarizado por el <acronym title="World Wide Web Consortium">W3C</acronym>
11     </p>
12
13     <p>La palabra <abbr title="Aproximación">aprox.</abbr>
14     es una abreviatura.
15     </p>
16 </body>
17 </html>

```

Captura de la salida en el navegador:



La mayoría de los navegadores muestran por defecto un borde inferior punteado para todos los elementos `<abbr>` y `<acronym>`.



Al posicionar el puntero del ratón sobre la palabra subrayada, el navegador muestra un pequeño recuadro (llamado **tooltip** en inglés) con el valor del atributo **title**.

4.2.- Etiquetas `<dfn>`

En ocasiones resulta útil incluir la definición de una palabra extraña o cuyo uso está restringido a un entorno muy determinado.

HTML incluye la etiqueta `<dfn>` para proporcionar al usuario la definición de todas las palabras para las que se considere apropiado. La definición formal de esta etiqueta se muestra a continuación:

Etiqueta	<code><dfn></code>
Atributos comunes	básicos , internacionalización y eventos
Atributos propios	<ul style="list-style-type: none"> <code>title = "texto"</code> - Indica el significado completo del término
Tipo de elemento	En línea
Descripción	Se emplea para marcar las definiciones de ciertos términos y proporcionar el significado de esos términos

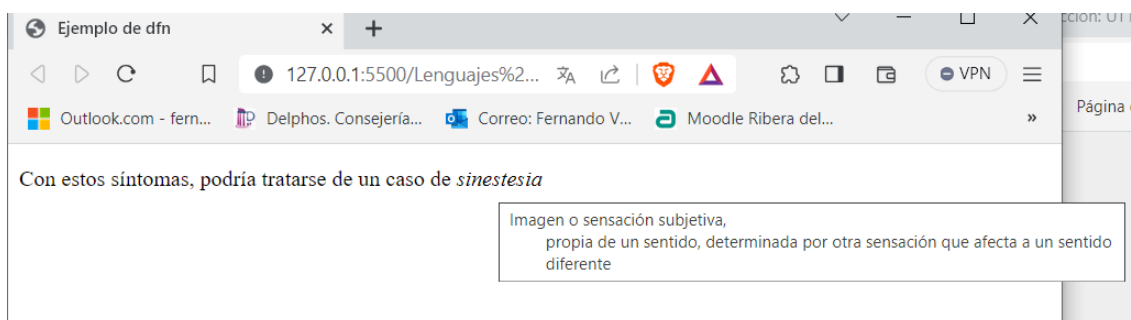
Ejemplo:

```

1 <!DOCTYPE html>
2 <html lang="en">
3 <head>
4   <meta charset="UTF-8">
5   <meta name="viewport" content="width=device-width, initial-scale=1.0">
6   <title>Ejemplo de dfn</title>
7 </head>
8 <body>
9   <p>Con estos síntomas, podría tratarse de un caso de <dfn title="Imagen o sensación subjetiva,
10     propia de un sentido, determinada por otra sensación que afecta a un sentido
11     diferente">sinestesia</dfn>
12   </p>
13 </body>
14 </html>

```

La salida es la siguiente:



4.3.- Etiquetas `<cite>`

Por último, HTML incluye una etiqueta que se puede utilizar para marcar un texto como una cita:

Etiqueta	<code><cite></code>
Atributos comunes	básicos , internacionalización y eventos
Atributos propios	-
Tipo de elemento	En línea
Descripción	Se emplea para marcar una cita o una referencia a otras fuentes

En ocasiones, no está clara la diferencia entre `<cite>` y `<blockquote>`.

- El elemento `<cite>` marca el autor de la cita (persona, documento, etc.) y `<blockquote>` marca el contenido de la propia cita.

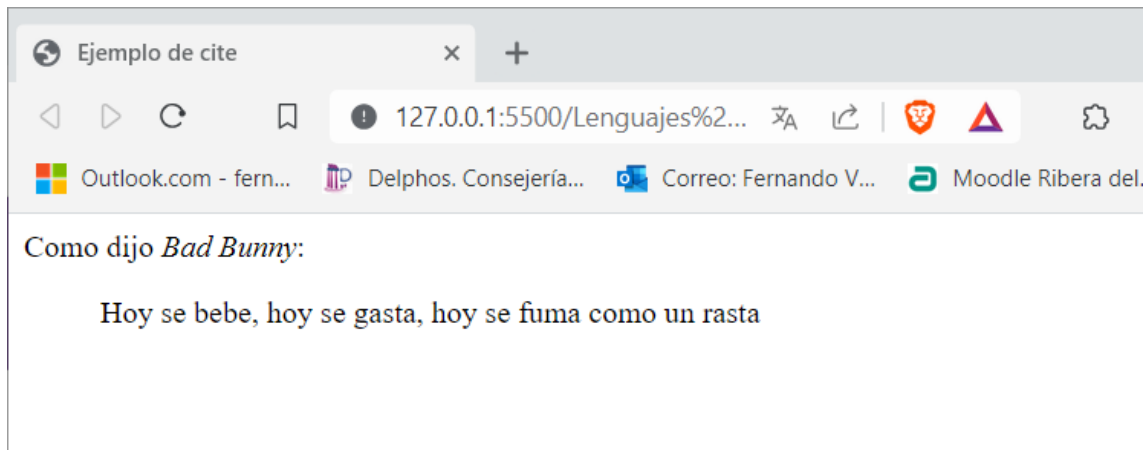
En el siguiente ejemplo, `<blockquote>` encierra el contenido de una frase célebre y `<cite>` encierra el nombre de su autor:

```

1  <!DOCTYPE html>
2  <html lang="en">
3  <head>
4      <meta charset="UTF-8">
5      <meta name="viewport" content="width=device-width, initial-scale=1.0">
6      <title>Ejemplo de cite</title>
7  </head>
8  <body>
9      Como dijo <cite>Bad Bunny</cite>:
10     <blockquote>Hoy se bebe, hoy se gasta, hoy se fuma como un rasta</blockquote>
11 </body>
12 </html>|

```

La salida es la siguiente:



ACTIVIDAD2 - UT2_3

Estructurar y marcar el siguiente texto extraído de la Wikipedia para que el navegador lo muestre con el aspecto de la siguiente imagen:

El Ártico ha perdido el 14% de su hielo marino perenne en un solo año

WASHINGTON.- El hielo perenne del Ártico se redujo en un 14% durante el último año, al perder **720.000 kilómetros cuadrados**, una superficie superior a la Península Ibérica, según datos de la NASA.

Según el JPL, la pérdida del hielo perenne, que debiera mantenerse durante todo el verano, fue todavía mayor y se acercó a un 50% en el momento en que ese hielo se desplazaba desde el Ártico oriental hacia el oeste.

Son Nghiem, investigador del JPL ha declarado que:

"Los cambios registrados en esos años en el hielo ártico son rápidos y espectaculares. De mantenerse la situación, ésta tendrá un impacto profundo en el ambiente, así como en el transporte marino y el comercio."

Texto plano para usar y formatear:

"El Ártico ha perdido el 14% de su hielo marino perenne en un solo año
WASHINGTON. - El hielo perenne del Ártico se redujo en un 14% durante el último año, al perder 720.000 kilómetros cuadrados, una superficie superior a la Península Ibérica, según datos de la NASA.

Según el JPL, la pérdida del hielo perenne, que debiera mantenerse durante todo el verano, fue todavía mayor y se acercó a un 50% en el momento en que ese hielo se desplazaba desde el Ártico oriental hacia el oeste.

Son Nghiem, investigador del JPL ha declarado que:

"Los cambios registrados en esos años en el hielo Ártico son rápidos y espectaculares. De mantenerse la situación, ésta tendrá un impacto profundo en el ambiente, así como en el transporte marino y el comercio"

5.- Marcado genérico de texto

El estándar HTML/XHTML incluye numerosas etiquetas para marcar los contenidos de texto. No obstante, la infinita variedad de posibles contenidos textuales hace que no sean suficientes.

Si se considera el siguiente ejemplo:

Importante: si quiere ponerse en contacto con la empresa ACME, puede hacerlo en el teléfono 900 555 555 o a través de la dirección de correo electrónico contacto@acme.org

El texto del ejemplo anterior contiene elementos de texto:

- Importantes,
- siglas,
- números de teléfono
- direcciones de correo electrónico.

XHTML define la etiqueta **** para marcar los elementos importantes y **<acronym>** para marcar las siglas:

****Importante****: si quiere ponerse en contacto con la empresa **<acronym>**ACME**</acronym>**, puede hacerlo en el teléfono 900 555 555 o a través de la dirección de correo electrónico contacto@acme.org

Desafortunadamente, XHTML **no define** ninguna etiqueta específica para marcar números de teléfono o direcciones de correo electrónico. De la

misma forma, no define etiquetas para otros posibles elementos que se pueden encontrar en los contenidos de texto.

Por este motivo, el estándar HTML/XHTML incluye una etiqueta llamada **** que se emplea para marcar cualquier elemento que no se puede marcar con las otras etiquetas definidas. Siguiendo con el ejemplo anterior, la etiqueta **** se utiliza para marcar el teléfono y la dirección de correo electrónico:

```
<strong>Importante</strong>: si quiere ponerse en contacto con la  
empresa <acronym>ACME</acronym>, puede hacerlo en el teléfono  
<span>900 555 555</span> o a través de la dirección de correo  
electrónico <span>contacto@acme.org</span>
```

➤ ¿Por qué utilizamos esta etiqueta de marcado genérico?

La etiqueta **** se visualiza por defecto con el mismo aspecto que el texto normal. Por tanto, es habitual utilizar esta etiqueta junto con los atributos **id** y **class** para modificar posteriormente su aspecto con CSS:

```
<strong>Importante</strong>: si quiere ponerse en contacto con la  
empresa <acronym>ACME</acronym>, puede hacerlo en el teléfono  
<span class="telefono">900 555 555</span> o a través de la dirección de  
correo electrónico <span class="email">contacto@acme.org</span>
```



La etiqueta **** sólo se puede utilizar para encerrar contenidos y etiquetas en línea.

Cuando se quieren estructurar elementos de bloque, se utiliza la etiqueta **<div>**, tal y como se verá en capítulos posteriores.

6.- Espacios en blanco y nuevas líneas

El aspecto más sorprendente del lenguaje HTML cuando se desarrollan los primeros documentos es el tratamiento especial de los "espacios en blanco" del texto.

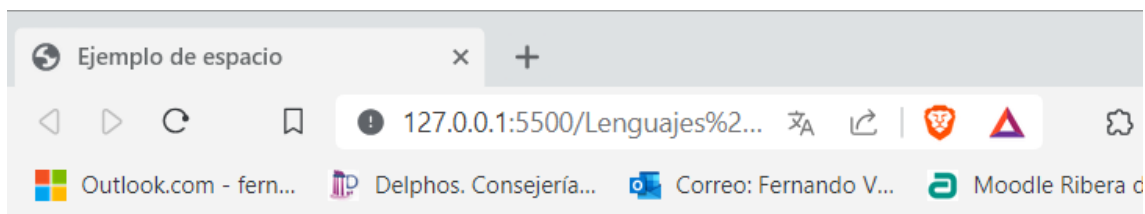


HTML considera espacio en blanco a los espacios en blanco, los tabuladores, los retornos de carro (Return) y el carácter de nueva línea (ENTER o Intro).

Vamos a ver el siguiente ejemplo:

```
1 <!DOCTYPE html>
2 <html lang="en">
3 <head>
4   <meta charset="UTF-8">
5   <meta name="viewport" content="width=device-width, initial-scale=1.0">
6   <title>Ejemplo de espacio</title>
7 </head>
8 <body>
9   <p>Este primer párrafo no contiene saltos de línea ni otro tipo de espaciado.</p>
10
11   <p>Este segundo párrafo sí que contiene saltos
12     de
13     línea
14     y otro tipo de espaciado.
15   </p>
16 </body>
17 </html>
```

Pero fíjate lo que se muestra por pantalla:



Este primer párrafo no contiene saltos de línea ni otro tipo de espaciado.

Este segundo párrafo sí que contiene saltos de línea y otro tipo de espaciado.

Los dos párrafos de la imagen anterior **se ven idénticos**, aunque el segundo párrafo incluye varios espacios en blanco y está escrito en varias líneas diferentes.



La razón de este comportamiento es que HTML ignora todos los espacios en blanco *sobrantes*, es decir, todos los espacios en blanco que no son el espacio en blanco que separa las palabras.

No obstante, HTML proporciona varias alternativas para poder incluir tantos espacios en blanco y tantas nuevas líneas como sean necesarias dentro del contenido textual de las páginas.

6.1.- Nuevas líneas

Para incluir una nueva línea en un punto y forzar a que el texto que sigue se muestre en la línea inferior, se utiliza la etiqueta **
**.

En cierta manera, insertar la etiqueta **
** en un determinado punto del texto equivale a presionar la tecla ENTER (o Intro) en ese mismo punto.

La definición formal de **
** se muestra a continuación:

Etiqueta	

Atributos comunes	<u>básicos</u>
Atributos propios	-
Tipo de elemento	En línea y etiqueta vacía
Descripción	Fuerza al navegador a insertar una nueva línea

La etiqueta **
** es una de las pocas etiquetas especiales de HTML. La particularidad de **
** es que es una etiqueta vacía, es decir, no encierra ningún texto.

De esta forma, la etiqueta debe abrirse y cerrarse de forma consecutiva: **
</br>**

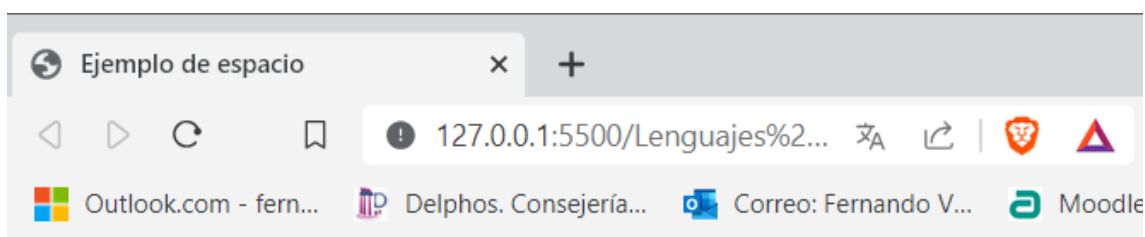


En estos casos, HTML permite utilizar un atajo para indicar que una etiqueta se está abriendo y cerrando de forma consecutiva: **
** (también se puede escribir como **
**).

Utilizando el ejemplo anterior:

```
1 <!DOCTYPE html>
2 <html lang="en">
3 <head>
4   <meta charset="UTF-8">
5   <meta name="viewport" content="width=device-width, initial-scale=1.0">
6   <title>Ejemplo de espacio</title>
7 </head>
8 <body>
9   <p>Este primer párrafo no contiene saltos de línea ni otro tipo de espaciado.</p>
10
11  <p>Este segundo párrafo sí que contiene saltos<br/>
12    de<br/>
13    línea<br/>
14    y otro tipo de espaciado.
15  </p>
16 </body>
17 </html>
```

La salida ahora es la siguiente:



Este primer párrafo no contiene saltos de línea ni otro tipo de espaciado.

Este segundo párrafo sí que contiene saltos
de
línea
y otro tipo de espaciado.

6.2.- Espacios en blanco “ ”

La solución al problema de los espacios en blanco no es tan sencilla como el de las nuevas líneas. Para incluir espacios en blanco adicionales, se debe sustituir cada nuevo espacio en blanco por el texto ** **; (es importante incluir el símbolo **&** al principio y el símbolo **;** al final).

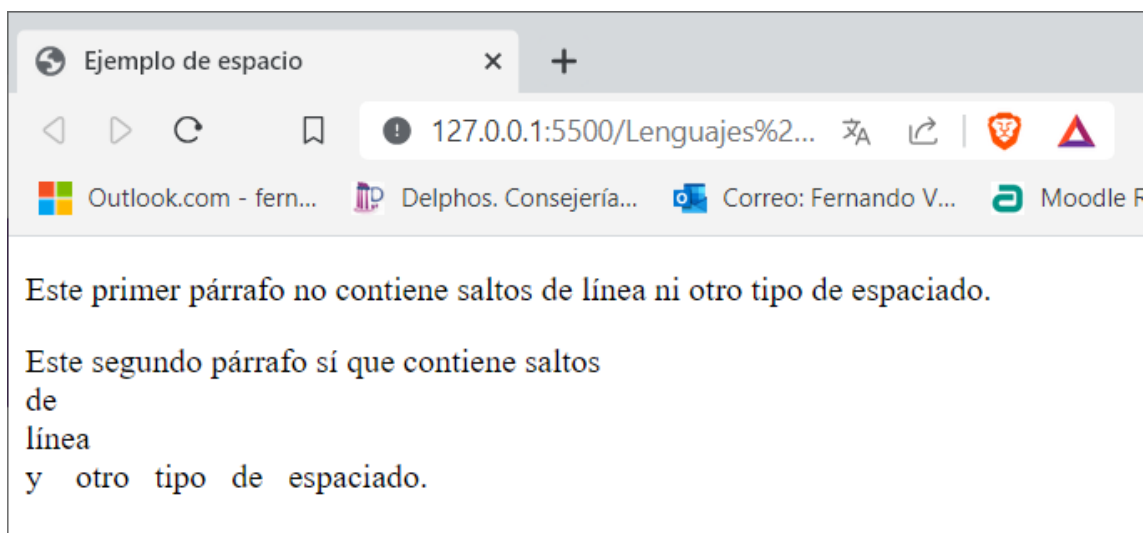
Así, el código HTML del ejemplo anterior se debe rehacer para incluir los espacios en blanco adicionales:

```

1 <!DOCTYPE html>
2 <html lang="en">
3   <head>
4     <meta charset="UTF-8">
5     <meta name="viewport" content="width=device-width, initial-scale=1.0">
6     <title>Ejemplo de espacio</title>
7   </head>
8   <body>
9     <p>Este primer párrafo no contiene saltos de línea ni otro tipo de espaciado.</p>
10
11    <p>Este segundo párrafo sí que contiene saltos<br/>
12      de<br/>
13      línea<br/>
14      y   otro   tipo   de   espaciado.
15    </p>
16  </body>
17 </html>

```

La salida es:



Cada texto ** ** solamente equivale a un espacio en blanco, por lo que se deben escribir tantos ** ** seguidos como espacios en blanco seguidos existan en el texto.

6.3.- Texto preformateado <pre> y <code>

En ocasiones, es necesario mostrar los espacios en blanco de un texto que no se puede modificar. Se trata de un caso habitual cuando una página web debe mostrar directamente el texto generado por alguna aplicación.

En estos casos, se puede utilizar la etiqueta **<pre>**, que muestra el texto tal y como se ha escrito, respetando todos los espacios en blanco y todas las nuevas líneas. La definición formal de la etiqueta se muestra a continuación:

Etiqueta	<pre>
Atributos comunes	básicos , internacionalización y eventos
Atributos propios	-
Tipo de elemento	Bloque
Descripción	Muestra el texto que encierra tal y como está escrito (respetando los espacios en blanco)

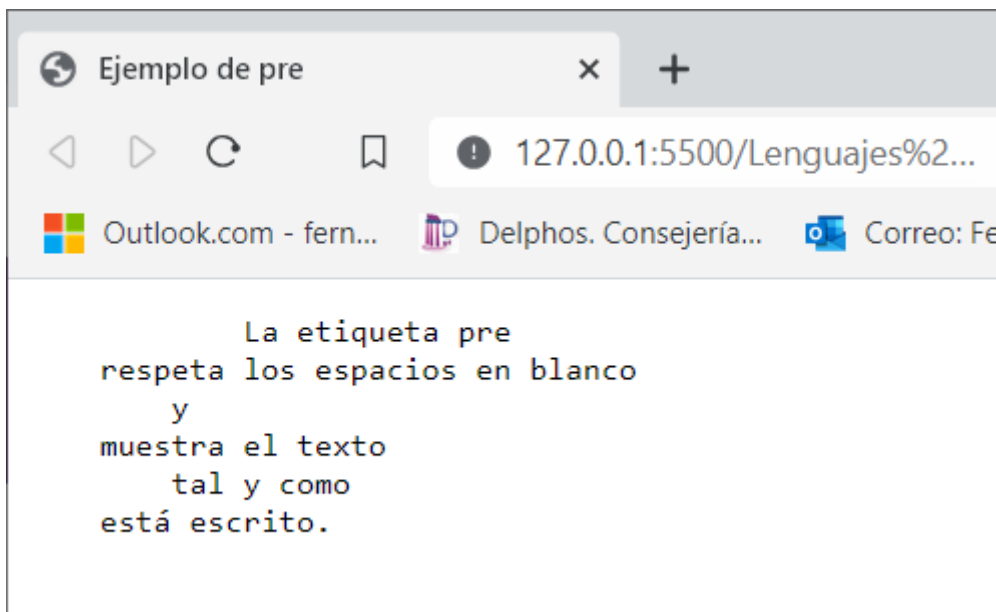
Ejemplo:

```

1  <!DOCTYPE html>
2  <html lang="en">
3  <head>
4      <meta charset="UTF-8">
5      <meta name="viewport" content="width=device-width, initial-scale=1.0">
6      <title>Ejemplo de pre</title>
7  </head>
8  <body>
9      <pre>
10         |   |   La etiqueta pre
11         |   |   respeta los espacios en blanco
12         |   |   y
13         |   |   muestra el texto
14         |   |   tal y como
15         |   |   está escrito.
16         </pre>
17 </body>
18 </html>

```

La salida en el navegador es la que sigue:



Los elementos **<pre>** son especiales, ya que los navegadores les aplican las siguientes reglas:

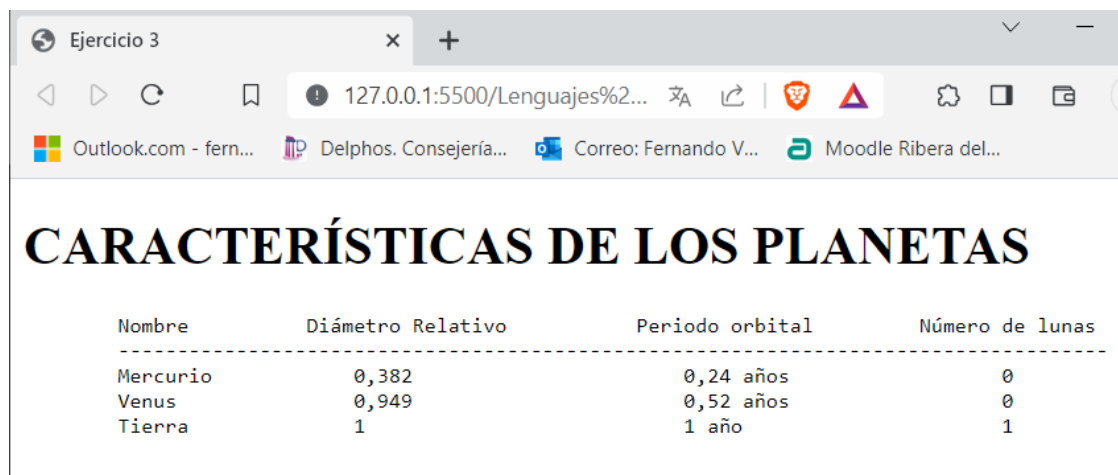
- Mantienen todos los espacios en blanco (tabulaciones, espacios y nuevas líneas).
- Muestra el texto con un tipo de letra especial, denominado "*de ancho fijo*", ya que todas sus letras son de la misma anchura.
- No se ajusta la longitud de las líneas (las líneas largas producen un *scroll* horizontal en la ventana del navegador).



Esta última característica diferencia por completo a los párrafos de los elementos **<pre>**. Como se ha visto, los navegadores ajustan la anchura de los párrafos de texto para que ocupen todo el tamaño de la ventana. Sin embargo, los elementos **<pre>** se muestran tal y como son originalmente, por lo que una línea muy larga dentro de un elemento **<pre>** provoca que la anchura de la página sea superior a la anchura de la ventana del navegador.

ACTIVIDAD3 - UT2_3

Determinar el código HTML que corresponde a la siguiente captura de pantalla:



Otra etiqueta relacionada con `<pre>` es la etiqueta `<code>`, que se utiliza para mostrar código fuente de cualquier lenguaje de programación. La definición formal de `<code>` es la siguiente:

Etiqueta	<code><code></code>
Atributos comunes	básicos , internacionalización y eventos
Atributos propios	-
Tipo de elemento	En línea
Descripción	Delimita el texto considerado un fragmento de código fuente



En la mayoría de las páginas web, no tiene sentido utilizar la etiqueta `<code>`. Sin embargo, en muchas páginas web

técnicas que incluyen listados de programas, trozos de código o etiquetas HTML, lo correcto es emplear la etiqueta `<code>`.

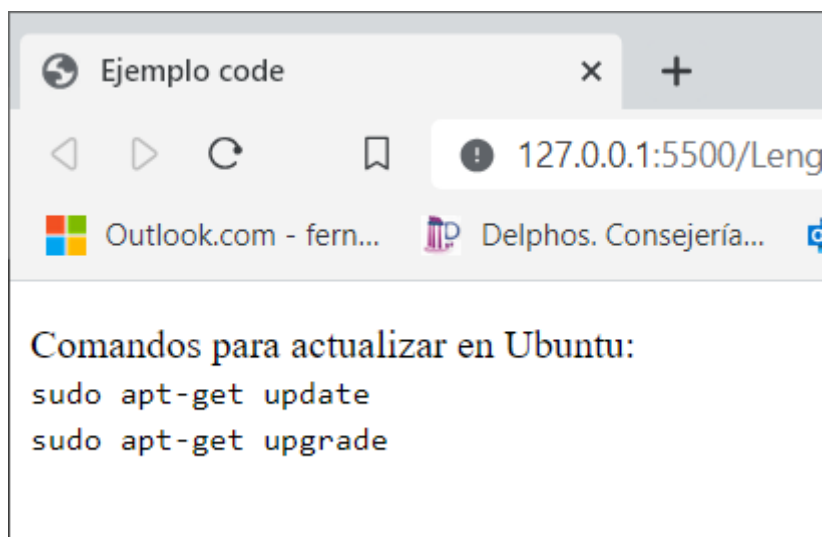
Ejemplo:

```

1  <!DOCTYPE html>
2  <html lang="en">
3  <head>
4      <meta charset="UTF-8">
5      <meta name="viewport" content="width=device-width, initial-scale=1.0">
6      <title>Ejemplo code</title>
7  </head>
8  <body>
9      <p>Comandos para actualizar en Ubuntu:<br/>
10         <code>sudo apt-get update</code><br/>
11         <code>sudo apt-get upgrade</code>
12     </p>
13 </body>
14 </html>

```

La salida es:



Al igual que sucede con los elementos `<pre>`, el texto encerrado por la etiqueta `<code>` se muestra con un tipo de letra especial de ancho fijo.

Por el contrario, el elemento `<code>` no respeta los espacios en blanco ni las líneas, por lo que su comportamiento es similar a la etiqueta `<p>`.

- La última diferencia es que `<code>` es un elemento en línea, mientras que `<pre>` es un elemento de bloque.

7.- Codificación de caracteres

Una consideración importante directamente relacionada con el texto de las páginas HTML es la codificación de los caracteres y la inserción de caracteres *especiales*. Algunos de los caracteres que se utilizan habitualmente en los textos no se pueden incluir directamente en las páginas web:

- ❑ Los caracteres que utiliza HTML para definir sus etiquetas (`<`, `>` y `"`) no se pueden utilizar libremente.
- ❑ Los caracteres propios de los idiomas que no son el inglés (ñ, á, ç, ð, ï, etc.) pueden ser problemáticos dependiendo de la codificación de caracteres utilizada.



La solución a la primera limitación consiste en sustituir los caracteres reservados de HTML por unas expresiones llamadas *entidades HTML* y que representan a cada carácter:

Entidad	Carácter	Descripción	Traducción
<	<	less than	signo de menor que
>	>	more than	signo de mayor que
&	&	ampersand	ampersand
"	"	quotation mark	comillas
 	(espacio en blanco)	non-breaking space	espacio en blanco
'	'	apostrophe	apóstrofo

Los caracteres `<`, `>`, `"` y `&` pueden dar problemas con los textos en HTML

Para mostrar correctamente el texto anterior en una página HTML, se debe sustituir cada carácter especial por su entidad HTML:

`<p>Los caracteres <, >, " y & pueden dar problemas con los textos en HTML</p>`

Por otra parte, los **caracteres propios de los idiomas diferentes al inglés** también **pueden ser problemáticos**. El motivo es que desde que se crea una página web hasta que llega al navegador del usuario, intervienen numerosos procesos:

- El diseñador crea la página web con su editor HTML (por ejemplo, Dreamweaver).
- Si se trata de una aplicación dinámica, el programador recorta la página HTML del diseñador y la mezcla con el resto del código de la aplicación (por ejemplo, PHP).
- El servidor web almacena las páginas HTML estáticas o el código de la aplicación web y sirve las páginas solicitadas por los usuarios.
- El usuario solicita y visualiza las páginas web a través de su navegador.

Si en todos los procesos anteriores se utiliza la misma codificación de caracteres, los caracteres propios de los idiomas se pueden escribir directamente:

`<p>Este párrafo contiene caracteres acentuados y se almacena en formato UTF-8</p>`

La palabra párrafo del ejemplo anterior incluye la letra á. Si el editor HTML del diseñador utiliza la codificación UTF-8, el entorno de desarrollo del

programador también utiliza UTF-8, el servidor web sirve las páginas con esa codificación y el navegador del usuario es capaz de visualizar las páginas con formato UTF-8, el texto anterior se verá correctamente en el navegador del usuario.

Sin embargo, **muchas veces no es posible que todos los procesos involucrados utilicen la misma codificación de caracteres.** Por limitaciones técnicas o por decisiones de los diseñadores y programadores, los textos pueden pasar de codificación UTF-8 a codificación ISO-8859 en cualquier momento. Si se produce este cambio sin realizar una conversión correcta, el navegador del usuario mostrará caracteres extraños en todos los acentos y en todas las letras como la ñ.

➤ ¿Cómo solucionamos esto?

La solución más sencilla para asegurar que todos estos caracteres potencialmente problemáticos se van a visualizar correctamente en el navegador del usuario consiste en sustituir cada carácter problemático por su entidad HTML:

Entidad	Carácter	Descripción oficial
ñ	ñ	latin letter n with tilde
Ñ	Ñ	latin capital n letter with tilde
á	á	a acute
é	é	e acute
í	í	i acute
ó	ó	o acute
ú	ú	u acute
Á	Á	A acute
É	É	E acute
Í	Í	I acute
Ó	Ó	O acute
Ú	Ú	U acute
€	€	euro

Así, el párrafo de texto del ejemplo anterior se podría escribir de la siguiente manera:

```
<p>Este párrafo contiene caracteres acentuados y se almacena  
en formato UTF-8</p>
```

ACTIVIDAD4 - UT2_3

Determinar el código HTML que corresponde a la siguiente captura de pantalla:

