

## Algoritmos y Estructuras de Datos

### Trabajo práctico

Un sitio de venta de artículos online pronto a ser puesto en marcha necesita confeccionar su sistema de ventas y reportes. Si bien esta versión inicial se utilizará para hacer pruebas conceptuales del negocio, por lo tanto no contendrá interfaces web, será útil para entender la visión del cliente. Se implementarán funcionalidades básicas para la actualización de montos de usuarios y administración de sus compras. Este prototipo es una versión “simplificada” del sistema a final, no se tendrán en cuenta datos de los medios de pago, datos del artículo comprado (salvo su Nro. de Artículo), rigurosidad en performance e interacción con otros sistemas de la organización. Los tipos de datos principales que se deberán manejar son la “Cuenta” y la “Operación” que presentan la siguiente estructura:

Usuario	
UsuarioID	99999
FechaCreacion	AAAAMMDD
Activo	True/false
TotalImporteCompras	999.999999
eMail	usrmail@server.com

Registro “Usuario”

Compra	
CompraID	9999
FechaHora	AAAAMMDDHH:MM
Monto	999.99999
UsuarioID	99999
NroArtículo	99999

Registro “Compra”

Al comenzar la jornada se “levantan” todos los usuarios desde el archivo “Clientes.bin”. Cada vez que el usuario requiera se procesará un “lote de compras”. Se desconoce cuántas compras puede haber en cada lote, pero puede ser tan grande que haga difícil su manejo en un vector. El procesamiento de un lote de compras implica la actualización de los montos totales de aquellos usuarios que hayan realizado alguna compra. Además, cada vez que se procesa un lote, deben “trasladarse” los registros procesados a un archivo llamado “procesados.bin”, el cual contiene todas las compras de todos los lotes procesados en el día.

Al finalizar la jornada se reescribe el archivo “Clientes.bin” con los saldos actualizados de las cuentas activas únicamente.

Ud. y su equipo deben confeccionar un programa que permita:

1. “Levantar” los clientes del archivo “Clientes.bin”.
2. Cargar un nuevo cliente.
3. Desactivar un usuario existente.
4. Buscar un cliente por ID o por mail.
5. Listar todos los clientes activos ordenados por total del importe.
6. Procesar un lote de compras.
7. Mostrar por pantalla todas las compras realizadas de un cliente dado (desde el archivo procesados.bin).
8. Mostrar todas las compras realizadas entre dos fechas en un reporte escrito en formato html. También mostrar el total de las compras.
9. Mostrar el mismo reporte que el punto 8 en formato CSV.
10. Finalizar jornada.

### Entrega:

Se establecen dos fechas de entrega. Quienes presenten el trabajo en la primera fecha y no aprueben, tienen posibilidad de presentarse en la segunda. Las fechas son 23/10 y 6/11. La entrega debe ser con el programa compilado y funcionando. La misma debe incluir:

- Informe, con carátula que indique los legajos, nombres, apellidos y correos de los integrantes del grupo. En el informe debe indicarse una descripción de la solución con hipótesis tenidas en cuenta para la misma, división de tareas en el equipo, diagrama de bloques de subprogramas. La presentación debe ser prolija, con hojas abrochadas y en folio o carpeta.
- Código fuente en un repositorio en github del grupo, casos de validación y versión digital del informe.

### Composición de los grupos:

Los grupos serán de hasta 4 integrantes sin excepción, no pudiendo haber grupos de 1 persona. Para la defensa , el día de entrega, deben estar presentes todos los integrantes del grupo

## 1.1.Anexo: El lenguaje HTML

Para poder aplicar web scraping, obviamente es fundamental entender la materia prima a tratar, esto es: archivos en HTML.

El lenguaje HTML es el utilizado para la confección de páginas web. Sus siglas representan las iniciales de *HyperText Markup Language* (lenguaje de marcas de hipertexto). HTML es una instancia, una especificación de un lenguaje, utilizando XML.

Si bien, hoy en día, el HTML se utiliza junto a otros lenguajes (como javascript, JQUERY desde el cliente y algún otro lenguaje que pueda ejecutarse desde el servidor, como PHP, ASP.NET, JSP, etc), la estructura de las páginas está especificada puramente en HTML.

Como todo lenguaje formal cumple un estándar normado por [World Wide Web Consortium \(W3C\)](https://www.w3.org/) o Consorcio WWW (<https://www.w3.org/>), organización dedicada a la estandarización de casi todas las tecnologías ligadas a la web, sobre todo en lo referente a su escritura e interpretación. Esta estandarización, permite que los navegadores web deban cumplir el estándar, permitiendo la existencia distintos productos para interpretar HTML.

En definitiva, un HTML es un archivo de texto (se puede crear con el notepad), cuya característica principal es el uso de etiquetas para crear los elementos que componen la página. Dichos elementos no pueden estar en cualquier posición, por ejemplo, el elemento raíz debe ser una etiqueta <html> y todo el contenido de la página debe cerrarse cerrando dicha etiqueta con la etiqueta </html> (todos los elementos deben ser cerrados, aunque algunos pueden ser cerrados dentro de la misma etiqueta que se abren).

## 1.2.Etiquetas html básicas

- <html>: define el inicio del documento HTML, le indica al [navegador](#) que lo que viene a continuación debe ser interpretado como código HTML. Esto es así *de facto*, ya que en teoría lo que define el tipo de documento es el [DOCTYPE](#), que significa la palabra justo tras DOCTYPE el tag de raíz.
- <script>: incrusta un [script](#) en una web, o llama a uno mediante `src="url del script"`. Se recomienda incluir el [tipo MIME](#) en el atributo `type`, en el caso de [JavaScript](#) `text/javascript`.
- <head>: define la [cabecera](#) del documento HTML; esta cabecera suele contener información sobre el documento que no se muestra directamente al [usuario](#) como, por

ejemplo, el título de la ventana del navegador. Dentro de la cabecera `<head>` es posible encontrar:

- `<title>`: define el título de la página. Por lo general, el título aparece en la barra de título encima de la ventana.
- `<body>`: define el contenido principal o cuerpo del documento. Esta es la parte del documento html que se muestra en el navegador; dentro de esta etiqueta pueden definirse propiedades comunes a toda la página, como `color` de fondo y márgenes. Dentro del cuerpo `<body>` es posible encontrar numerosas etiquetas. A continuación se indican algunas a modo de ejemplo:
- `<h1>` a `<h6>`: encabezados o títulos del documento con diferente relevancia.
- `<table>`: define una tabla.
- `<tr>`: fila de una tabla.
- `<td>`: celda de una tabla (debe estar dentro de una fila).

Estas etiquetas se utilizan de forma anidada. Por ejemplo, una tabla de una sola fila que contiene una sola celda sería:

```
<table><tr><td>Contenido de una celda</td></tr></table>
```

- `<a>`: [hipervínculo](#) o enlace, dentro o fuera del [sitio web](#). Debe definirse el parámetro de pasada por medio del atributo `href`. Por ejemplo: `<a href="http://www.example.com">Ejemplo</a>` se representa como ejemplo.
- `<img>`: imagen. Requiere del atributo `src`, que indica la ruta en la que se encuentra la imagen. Por ejemplo: ``. Es conveniente, por accesibilidad, poner un atributo `alt="texto alternativo"`.
- `<li><ol><ul>`: etiquetas para listas.
- `<b>`: texto en negrita (*Deprecado. Se reemplazó por la etiqueta `<strong>`*).
- `<i>`: texto en cursiva (*Deprecado. Se reemplazó por la etiqueta `<em>`*).
- `<s>`: texto tachado (*Deprecado. Se reemplazó por la etiqueta `<del>`*).
- `<u>`: Subrayado.

Nuevamente, la mayoría de etiquetas deben cerrarse como se abren, pero con una barra (/). Note, por ejemplo, la apertura y cierre del elemento `<br/>` del ejemplo de la figura 1.

### 1.3.Ejemplo HTML básico

El siguiente código muestra un ejemplo de un HTML básico que incluye en su cuerpo algunos elementos básicos listados en el punto anterior: un encabezado de tamaño 1, un párrafo, una tabla con dos filas y dos columnas y, finalmente, un vínculo a la página de la facultad. Por simplicidad, y para obviar por ahora los entities html, no se utilizan tildes dentro del código. Ud. puede crear este código, o uno similar, utilizando notepad u otro editor de texto (no procesador de texto como word o write, sino editor). Guarde el archivo con extensión .HTML para que el sistema operativo lo asocie directamente para ser abierto con el navegador predeterminado.

```
1 <html>
2   <head>
3     <title>Sintaxis y semantica de los lenguajes: Trabajo practico</title>
4   </head>
5   <body>
6     <h1>Trabajo practico #1</h1>
7     <p>Vamos a aplicar el conocimiento adquirido en nuestra materia a un lenguaje utilizado actualmente.</p>
8     <table>
9       <tr>
10        <td>
11          Columna 1
12        </td>
13        <td>
14          Columna 2
15        </td>
16      </tr>
17      <tr>
18        <td>
19          1
20        </td>
21        <td>
22          2
23        </td>
24      </tr>
25    </table>
26    <br/> <!-- El br es un ejemplo de etiqueta que se abre y se cierra inmediatamente -->
27    <a href="https://www.frba.utn.edu.ar/en/">Ir a pagina de la facu.</a>
28  </body>
29 </html>
```

Figura 1 - Código HTML de la página de ejemplo

Al abrir el archivo descrito en la figura anterior se visualiza en el navegador web lo siguiente:

## Trabajo practico #1

Vamos a aplicar el conocimiento adquirido en nuestra materia a un lenguaje utilizado actualmente.

Columna 1 Columna 2

1 2

[Ir a pagina de la facu.](https://www.frba.utn.edu.ar/en/)

Figura 2 - Visualización del código de la figura 1